MATH 6461
Functional Analysis I

Paul Skoufranis

November 30, 2023

Preface:

These are the first edition of these lecture notes for MATH 6461 (Functional Analysis I). Consequently, there may be several typographical errors, missing exposition on necessary background, and more advance topics for which there will not be time in class to cover. Future iterations of these notes will hopefully be fairly self-contained provided one has the necessary background. If you come across any typos, errors, omissions, or unclear explanations, please feel free to contact me so that I may continually improve these notes.

Contents

Chapter 1

Normed Linear Spaces

The most challenging aspects of a functional analysis course is where to begin and how deep to go into each topic. Functional analysis, being the study of infinite dimensional topological vector spaces over the real and complex numbers, requires substantial knowledge from linear algebra, analysis, measure theory, and topology. At the graduate level, it is expected that students are familiar with linear algebra (specifically, vector spaces, subspaces, spans, linear independent, bases, dimension, linear maps, and spectral theory for normal matrices) and have some exposure to analysis. The questions remain, "how much analysis, measure theory, and topology will students have seen, and how much detail should be provided on each topic?"

In this section, we desire to examine the basics of normed linear spaces, whereas we will leave essential study of Banach spaces to Chapter 2. For some students, this will be mostly a review. For others, there may be many topics that are quite new and challenging. While proceeding through this chapter, we will assume that students have seen the basics of metric spaces (definitions, convergent sequences, the metric topology, completeness, continuity) and the only topological notions required for this section are those for metric spaces. A few results will be mentioned using general topology and measure theory, but students unfamiliar with measure theory will be able to proceed without issue.

Thus we begin this chapter by reviewing the concept of a normed linear space including many standard examples and results minimal proof. Then we will transition to constructing other normed linear spaces, such as the bounded linear maps between normed linear spaces and examining further constructions on these spaces. Often we will include results that are more advanced than one might see at the undergraduate level and the results and examples that are more essential to the study of functional analysis.

1.1 Normed Linear Spaces

Again, we begin by reviewing the concept of a normed linear space. As often we will want to discuss results for both the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$, we will use $\mathbb{K}$ to denote the real or complex numbers (i.e. any result where $\mathbb{K}$ is used works for both the real and complex numbers). We chose $\mathbb{K}$ over $\mathbb{F}$ as $\mathbb{F}$ is usually used to denote an arbitrary field in mathematics whereas results in functional analysis require the real or complex numbers.

We begin with the concept of a norm, which is a simple generalization of the absolute value on $\mathbb{K}$ to vector spaces over $\mathbb{K}$.

Definition 1.1.1. Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. A *norm* on $\mathcal{V}$ is a function $\|\cdot\| : \mathcal{V} \to [0, \infty)$ such that

1. for $\vec{v} \in \mathcal{V}$, $\|\vec{v}\| = 0$ if and only if $\vec{v} = \vec{0}$,
2. $\|\alpha\vec{v}\| = |\alpha| \, \|\vec{v}\|$ for all $\alpha \in \mathbb{K}$ and $\vec{v} \in \mathcal{V}$, and
3. (triangle inequality) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ for all $\vec{v}, \vec{w} \in \mathcal{V}$.

Of course, we desire to study vector spaces with a fixed pre-described norm, so we make the following definition.

Definition 1.1.2. A *normed linear space* is a pair $(\mathcal{V}, \|\cdot\|)$ where $\mathcal{V}$ is a vector space over $\mathbb{K}$ and $\|\cdot\|$ is a norm on $\mathcal{V}$.

Note we will often abuse notation by saying that $\mathcal{V}$ is a normed linear space without specifying the norm.

Remark 1.1.3. Any normed linear space $(\mathcal{V}, \|\cdot\|)$ automatically becomes a metric space with respect to the metric $d : \mathcal{V} \times \mathcal{V} \to [0, \infty)$ defined by

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| .$$

As such, the basic topological properties of metric spaces immediately apply to normed linear spaces. In particular:

- a sequence $(\vec{x}_n)_{n\geq 1}$ of elements of $\mathcal{V}$ *converges* to $\vec{x} \in \mathcal{V}$ if
$$\lim_{n\to\infty} \|\vec{x}_n - \vec{x}\| = 0;$$
that is, for every $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that $\|\vec{x}_n - \vec{x}\| < \epsilon$ for all $n \geq N$.
- the *open ball* centred at $\vec{x} \in \mathcal{V}$ of radius $r > 0$ is the set
$$B(\vec{x}, r) = \{\vec{v} \in \mathcal{V} \mid \|\vec{v} - \vec{x}\| < r\}.$$
- a subset $A \subseteq \mathcal{V}$ is *open* in $\mathcal{V}$ if for all $\vec{x} \in A$ there exists an $r > 0$ such that $B(\vec{x}, r) \subseteq A$.

- a subset $F \subseteq \mathcal{V}$ is *closed* in $\mathcal{V}$ if $F^c = \mathcal{V} \setminus F$ is open in $\mathcal{V}$.
- a subset $F \subseteq \mathcal{V}$ is closed if and only if whenever $(\vec{x}_n)_{n\geq 1}$ is a sequence of elements of F that converges to some element $\vec{x} \in \mathcal{V}$, then $\vec{x} \in F$.
- a sequence $(\vec{x}_n)_{n\geq 1}$ of elements of $\mathcal{V}$ is *Cauchy* if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\|\vec{x}_n - \vec{x}_m\| < \epsilon$ for all $n, m \geq N$.
- $\mathcal{V}$ is said to be *complete* if every Cauchy sequence in $\mathcal{V}$ converges in $\mathcal{V}$.
- if $\mathcal{W}$ is a normed linear space, a function $f : \mathcal{V} \to \mathcal{W}$ is said to be *continuous at a point* $\vec{v} \in \mathcal{V}$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $\vec{x} \in \mathcal{V}$ and $\|\vec{x} - \vec{v}\|_{\mathcal{V}} < \delta$, then $\|f(\vec{x}) - f(\vec{v})\|_{\mathcal{W}} < \epsilon$. Consequently, it can be show that f is *continuous* (i.e. continuous at every point in $\mathcal{V}$) if and only if $f^{-1}(U)$ is open in $\mathcal{V}$ for every open subset $U \subseteq \mathcal{W}$.

The immediate added benefit of looking at normed linear spaces over metric spaces is that the norm properties intertwine with the vector space operations with respect to continuity and convergence.

Proposition 1.1.4. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space over $\mathbb{K}$. If $(\vec{x}_n)_{n\geq 1}$ and $(\vec{y}_n)_{n\geq 1}$ are sequences of elements of $\mathcal{V}$ that converge to $\vec{x}$ and $\vec{y}$ respectively, and let $(\alpha_n)_{n\geq 1}$ be a sequence of elements of $\mathbb{K}$ that converge to α, then*

- *$(\vec{x}_n + \vec{y}_n)_{n\geq 1}$ converges to $\vec{x} + \vec{y}$, and*
- *$(\alpha_n \vec{x}_n)_{n\geq 1}$ converges to $\alpha \vec{x}$.*

Proof. Let $\epsilon > 0$. Since

$$\begin{aligned} \|(\vec{x}_n + \vec{y}_n) - (\vec{x} - \vec{y})\| &\leq \|\vec{x}_n - \vec{x}\| + \|\vec{y}_n - \vec{y}\| \text{ and} \\ \|\alpha_n \vec{x}_n - \alpha \vec{x}\| &\leq |\alpha_n| \, \|\vec{x}_n - \vec{x}\| + |\alpha_n - \alpha| \, \|\vec{x}\| \end{aligned}$$

for all n, and since convergence sequences are bounded in their norms by a simple triangle inequality argument, we may chose N sufficiently large so that the right-hand sides of both inequalities is less than ϵ and thus the result follows. ∎

Proposition 1.1.5. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space over $\mathbb{K}$. For all $\vec{x}, \vec{y} \in \mathcal{V}$, we have*

$$|\|\vec{x}\| - \|\vec{y}\|| \leq \|\vec{x} - \vec{y}\|.$$

This inequality is often called the reverse triangle inequality*. Consequently, the function $f : \mathcal{V} \to \mathbb{R}$ defined by $f(\vec{x}) = \|\vec{x}\|$ for all $\vec{x} \in \mathcal{V}$ is continuous.*

Proof. Given $\vec{x}, \vec{y} \in V$, notice by the triangle equality that

$$\begin{aligned} \|\vec{x}\| &= \|\vec{y} + (\vec{x} - \vec{y})\| \leq \|\vec{y}\| + \|\vec{x} - \vec{y}\| \quad \text{and} \\ \|\vec{y}\| &= \|\vec{x} + (\vec{y} - \vec{x})\| \leq \|\vec{x}\| + \|\vec{y} - \vec{x}\| = \|\vec{x}\| + \|\vec{x} - \vec{y}\| \end{aligned}$$

as $\|-\vec{z}\| = |-1| \, \|\vec{z}\| = \|\vec{z}\|$ for all $\vec{z} \in \mathcal{V}$. By rearranging this equation, we obtain the reverse triangle inequality. The continuity of the norm immediately follows from the reverse triangle inequality and the definition of continuity. ∎

1.2 Examples of Normed Linear Spaces

Of course, having a mathematical object is only good if there is a plethora of examples. In this section, we will look at some of the most important examples of normed linear spaces in functional analysis. Note it is necessary for each example to not only verify that the defined norm is indeed a norm, but the set is actually a vector space. We note that the vector space operations on each normed linear space is the canonical one.

Example 1.2.1. The *absolute value function* $|\cdot| : \mathbb{K} \to [0, \infty)$ is a norm on $\mathbb{K}$ (where $\mathbb{K}$ is viewed as a vector space over $\mathbb{K}$ with the canonical operations). In fact, every norm on $\mathbb{K}$ is clearly seen to be a positive scalar multiple of the absolute value function. As such, when we refer to $\mathbb{K}$ as a normed linear space, we always do so with the absolute value function as the norm.

Although there is only one norm on $\mathbb{K}$ upto multiples, if we go up in vector space dimension, several norms that are not multiples are in existence.

Example 1.2.2. For $n \in \mathbb{N}$, recall $\mathbb{K}^n$ is a vector space over $\mathbb{K}$ with respect to coordinate-wise addition and scalar multiplication. For a $p \in [1, \infty)$, define $\|\cdot\|_p : \mathbb{K}^n \to [0, \infty)$ by

$$\|(z_1, \ldots, z_n)\|_p = \left(\sum_{k=1}^n |z_k|^p \right)^{\frac{1}{p}}$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$. Then $\|\cdot\|_p$ is a norm on $\mathbb{K}^n$ called the *p-norm*. In the case $p = 2$, the above norm is called the *Euclidean norm*.

Example 1.2.3. For $n \in \mathbb{N}$, recall $\mathbb{K}^n$ is a vector space over $\mathbb{K}$ with respect to coordinate-wise addition and scalar multiplication. Define $\|\cdot\|_\infty : \mathbb{K}^n \to [0, \infty)$ by

$$\|(z_1, \ldots, z_n)\|_\infty = \sup_{1 \leq k \leq n} |z_k|$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$. We call $\|\cdot\|_\infty$ the sup-*norm* or the ∞-*norm*.

Of course, it is necessary to check that the p-norms are indeed norms. Other than the triangle inequality, all other properties of a norm are easy to verify. To see that the triangle inequality holds, we refer the reader to Appendix D.

Furthermore, it is elementary to see that the various p-norms on $\mathbb{K}^n$ are not multiples of each other. Indeed notice that

$$\|(1,0,\ldots,0)\|_p = 1 \quad \text{whereas} \quad \|(1,1,\ldots,1)\|_p = \begin{cases} n^{\frac{1}{p}} & \text{if } p \neq \infty \\ 1 & \text{if } p = \infty \end{cases}$$

and thus no p-norm on $\mathbb{K}^n$ is a multiple of the other. However, the various p-norms are related to each other in another topological way.

Definition 1.2.4. Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on $\mathcal{V}$ are said to be *equivalent* if there exists $k_1, k_2 \in (0,\infty)$ such that

$$k_1 \|\vec{v}\|_1 \leq \|\vec{v}\|_2 \leq k_2 \|\vec{v}\|_1$$

for all $\vec{v} \in V$.

It is elementary to verify that the notion of equivalent norms from Definition 1.2.4 is an equivalence relation on the set of norms on $\mathcal{V}$. Furthermore, it is elementary to see that if two norms on $\mathcal{V}$ are equivalent, then they define the same topology on $\mathcal{V}$ and they have the same set of Cauchy sequences. Thus, for almost all intents and purposes in functional analysis, two equivalent norms on a vector space "produce the same" normed linear space.

Example 1.2.5. For a fixed $n \in \mathbb{N}$ and $p \in [1,\infty)$, notice for all $(z_1, z_2, \ldots, z_n) \in \mathbb{K}^n$ that

$$\|(z_1, z_2, \ldots, z_n)\|_\infty \leq \|(z_1, z_2, \ldots, z_n)\|_p \leq n^{\frac{1}{p}} \|(z_1, z_2, \ldots, z_n)\|_\infty$$

simply by using the fact that $\|(z_1, z_2, \ldots, z_n)\|_p = (\sum_{k=1}^n |z_k|^p)^{\frac{1}{p}}$. Hence all of the p-norms on $\mathbb{K}^n$ are equivalent.

However, this raises another question that we will answer later: "Are there other norms on $\mathbb{K}^n$ that are not equivalent to the Euclidean norm?"

For now, we turn our attention to infinite dimensions where the p-norms produce different vector spaces and thus different normed linear spaces.

Example 1.2.6. Given a measure space $(X, \mathcal{A}, \mu)$ and a $p \in [1,\infty)$, the *L_p-space of* $(X, \mathcal{A}, \mu)$, denote $L_p(X,\mu)$, is the normed linear space over $\mathbb{K}$ of all measurable functions $f : X \to \mathbb{K}$ such that

$$\int_X |f|^p \, d\mu < \infty$$

modulo the set of all functions equal to zero μ almost everywhere equipped with pointwise addition, pointwise scalar multiplication, and the *p-norm* $\|\cdot\|_p : L_p(X,\mu) \to [0,\infty)$ defined by

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}}$$

for all $f \in L_p(X,\mu)$.

Example 1.2.7. Given a measure space $(X,\mathcal{A},\mu)$, the *L_∞-space of* $(X,\mathcal{A},\mu)$, denote $L_\infty(X,\mu)$, is the normed linear space over $\mathbb{K}$ of all measurable functions $f : X \to \mathbb{K}$ for which there exists an $M \geq 0$ such that

$$\mu(\{x \in X \mid |f(x)| > M\}) = 0$$

modulo the set of all functions equal to zero μ almost everywhere equipped with pointwise addition, pointwise scalar multiplication, and the *∞-norm* $\|\cdot\|_\infty : L_\infty(X,\mu) \to [0,\infty)$ defined by

$$\|f\|_\infty = \inf\{M \in [0,\infty) \mid \mu(\{x \in X \mid |f(x)| > M\}) = 0\}$$

for all $f \in L_\infty(X,\mu)$.

Of course, it is necessary to verify that L_p-spaces are indeed vector spaces and the norms as defined are indeed norms. We refer an interested reader to Appendix D.

Remark 1.2.8. For most measure spaces $(X,\mathcal{A},\mu)$, $L_p(X,\mu)$ are different vector spaces for each value of p. Indeed in the case that $(X,\mathcal{A},\mu)$ is the Lebesgue measure on $\mathbb{R}$, if $q \in [1,\infty)$ and $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f_q(x) = \begin{cases} 0 & \text{if } x < 1 \\ x^{-\frac{1}{q}} & \text{if } x \geq 1 \end{cases},$$

then $f_q \in L_p(X,\mu)$ if and only if $p > q$.

Example 1.2.9. Let $p \in [1,\infty]$ and let $n \in \mathbb{N}$. If X is an n-point space, $\mathcal{A} = \mathcal{P}(X)$, and μ is the counting measure on X, then $L_p(X,\mu) = \mathbb{K}^n$ equipped with the p-norm. Hence Examples 1.2.2 and 1.2.3 are subsumed by L_p-spaces.

Example 1.2.10. Let $p \in [1,\infty)$, let $X = \mathbb{N}$, let $\mathcal{A} = \mathcal{P}(X)$, and let μ be the counting measure on X. Then $L_p(X,\mu)$ can be identified with the set of all sequences $(z_n)_{n\geq 1}$ of elements of $\mathbb{K}$ such that

$$\sum_{k=1}^\infty |z_k|^p < \infty.$$

We use $\ell_p(\mathbb{N})$ (or $\ell_p(\mathbb{N},\mathbb{K})$ to specify the field $\mathbb{K}$) to denote the space of all such sequence. Thus $\ell_p(\mathbb{N})$ is a normed linear space with respect to entry-wise addition, entry-wise scalar multiplication, and norm $\|\cdot\|_p : \ell_p(\mathbb{N}) \to [0,\infty)$ defined by

$$\|(z_n)_{n\geq 1}\|_p = \left(\sum_{k=1}^{\infty} |z_k|^p\right)^{\frac{1}{p}}.$$

Example 1.2.11. Let $X = \mathbb{N}$, let $\mathcal{A} = \mathcal{P}(X)$, and let μ be the counting measure on X. Then $L_\infty(X,\mu)$ can be identified with the set of all bounded sequences $(z_n)_{n\geq 1}$ of elements of $\mathbb{K}$. We use $\ell_\infty(\mathbb{N})$ (or $\ell_\infty(\mathbb{N},\mathbb{K})$ to specify the field $\mathbb{K}$) to denote the space of all such sequence. Thus $\ell_\infty(\mathbb{N})$ is a normed linear space with respect to entry-wise addition, entry-wise scalar multiplication, and norm $\|\cdot\|_\infty : \ell_\infty(\mathbb{N}) \to [0,\infty)$ defined by

$$\|(z_n)_{n\geq 1}\|_\infty = \sup_{n\in\mathbb{N}} |z_n|.$$

Remark 1.2.12. It is not difficult to see that $\ell_1(\mathbb{N}) \subseteq \ell_p(\mathbb{N}) \subseteq \ell_q(\mathbb{N}) \subseteq \ell_\infty(\mathbb{N})$ for all $p,q \in (1,\infty)$ with $p < q$.

Often in this course we will stick with these "little" ℓ_p-spaces oppose to the measure theoretic L_p-spaces due to convenience for those that have not taken measure theory. Surprisingly, most results that work for ℓ_p-spaces can be extended to L_p-spaces provided the measures involved are sufficiently nice.

While we are on the topic of sequence spaces, we note we can produce new normed linear spaces in a very simple way.

Proposition 1.2.13. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and let $\mathcal{W}$ be a vector subspace of $\mathcal{V}$. The restriction of $\|\cdot\|$ to $\mathcal{W}$ is a norm on $\mathcal{W}$. Hence $(\mathcal{W}, \|\cdot\||_{\mathcal{W}})$ is a normed linear space.*

Example 1.2.14. Let c (or $c(\mathbb{K})$ if we desire to specify the field) be the set

$$c = \left\{ (z_n)_{n\geq 1} \,\middle|\, z_n \in \mathbb{K}, \lim_{n\to\infty} z_n \text{ exists} \right\}.$$

As c is a vector subspace of $\ell_\infty(\mathbb{N},\mathbb{K})$, c is a normed linear space with respect to the ∞-norm. We call c the *convergent sequence space*.

Example 1.2.15. Let c_0 (or c if we desire to specify the field) be the set

$$c_0 = \left\{ (z_n)_{n\geq 1} \,\middle|\, z_n \in \mathbb{K}, \lim_{n\to\infty} z_n = 0 \right\}.$$

As c_0 is a vector subspace of $\ell_\infty(\mathbb{N},\mathbb{K})$, c_0 is a normed linear space with respect to the ∞-norm. We call c_0 the *convergent to 0 sequence space*.

Example 1.2.16. Let c_{00} (or c_{00} if we desire to specify the field) be the set

$$c_{00} = \{(z_n)_{n\geq 1} \in c \mid \text{ there exists an } N \in \mathbb{N} \text{ such that } z_n = 0 \text{ for all } n \geq N\}.$$

As c_{00} is a vector subspace of $\ell_\infty(\mathbb{N}, \mathbb{K})$, c_{00} is a normed linear space with respect to the ∞-norm. We call c_{00} the *eventually 0 sequence space.*

Example 1.2.17. Let $(X, \mathcal{A})$ be a measurable space and let $\mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ be the set of all complex measures on $(X, \mathcal{A})$. Then $\mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ is a vector space over $\mathbb{C}$ with respect to pointwise addition and scalar multiplication. Moreover, if we define $\|\cdot\| : \mathcal{M}_{\mathbb{C}}(X, \mathcal{A}) \to [0, \infty)$ by

$$\|\mu\| = \sup \left\{ \sum_{n=1}^{\infty} |\mu(A_n)| \;\middle|\; \{A_n\}_{n=1}^{\infty} \in \mathcal{A} \text{ pairwise disjoint with union } X \right\}$$

for all $\mu \in \mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$, then $\|\cdot\|$ is a norm on $\mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ called the *total variation norm.* Hence $\mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ is a normed linear space.

In fact, it can be shown that for any $\mu \in \mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ that there exists a unique measure $|\mu| : \mathcal{A} \to [0, \infty)$ and a measurable function $\varphi : X \to \mathbb{C}$ such that $|\varphi(x)| = 1$ for all $x \in X$ that is unique upto μ-measure zero sets such that

$$\mu(A) = \int_A \varphi \, d|\mu|$$

for all $A \in \mathcal{A}$. It can be verified that $\|\mu\| = |\mu|(X)$.

Example 1.2.18. Let $(X, \mathcal{A})$ be a measurable space and let $\mathcal{M}_{\mathbb{R}}(X, \mathcal{A})$ be the set of finite signed measures. Then $\mathcal{M}_{\mathbb{R}}(X, \mathcal{A})$ is a real vector subspace of $\mathcal{M}_{\mathbb{C}}(X, \mathcal{A})$ with respect to pointwise addition and scalar multiplication and thus a normed linear space with the restriction of the total variation norm. In fact, if $\mu \in \mathcal{M}_{\mathbb{R}}(X, \mathcal{A})$, it is known there exists a unique pair of singular finite measures $\mu_\pm : \mathcal{A} \to [0, \infty)$ such that $\mu(A) = \mu_+(A) - \mu_-(A)$ for all $A \in \mathcal{A}$. In this case, it can be verified that $\|\mu\| = \mu_+(X) + \mu_-(X)$.

To end our initial set of examples of normed linear spaces, we include the most obvious: continuous functions.

Example 1.2.19. Let $(X, \mathcal{T})$ be a topological space and let $(\mathcal{Y}, \|\cdot\|_Y)$ be a normed linear space over $\mathbb{K}$. A continuous function $f : X \to \mathcal{Y}$ is said to be *bounded* if there exists an $M \in \mathbb{R}$ such that $\|f(x)\|_Y \leq M$ for all $x \in X$.

The set of all bounded continuous functions from X to $\mathcal{Y}$, denoted $C_b(X, \mathcal{Y})$, is a normed linear space over $\mathbb{K}$ with respect to pointwise addition, pointwise scalar multiplication, and the norm $\|\cdot\|_\infty : C_b(X, \mathcal{Y}) \to [0, \infty)$ defined by

$$\|f\|_\infty = \sup \{\|f(x)\|_Y \mid x \in X\}$$

for all $f \in C_b(X, \mathcal{Y})$. The norm $\|\cdot\|_\infty$ often called the sup-*norm* or the ∞-*norm* as it agrees with the one from Example 1.2.7.

Example 1.2.20. If $X = \mathbb{N}$ is equipped with the discrete topology, then $C_b(X, \mathbb{K}) = \ell_\infty(\mathbb{N}, \mathbb{K})$ and the two infinity norms agree.

Example 1.2.21. Let $(X, \mathcal{T})$ be a compact topological space and let $(\mathcal{Y}, \|\cdot\|_Y)$ be a normed linear space over $\mathbb{K}$. As every continuous function $f : X \to \mathcal{Y}$ is automatically bounded, $C_b(X, \mathcal{Y})$ is the set of all continuous functions from X to $\mathcal{Y}$ and is denoted by $C(X, \mathcal{Y})$. Thus $C(X, \mathcal{Y})$ is a normed linear space with the infinity norm from Example 1.2.19.

In the case that X is the closed interval $[a, b]$ and $\mathcal{Y} = \mathbb{R}$, we will use $C[a, b]$ to denote $C([a, b], \mathbb{R})$.

Example 1.2.22. Let $(X, \mathcal{T})$ be a locally compact Hausdorff topological space and let $(\mathcal{Y}, \|\cdot\|_Y)$ be a normed linear space over $\mathbb{K}$. A continuous function $f : X \to \mathcal{Y}$ is said to be *vanish at infinity* if for all $\epsilon > 0$ there exists a compact subset $K \subseteq X$ such that $\|f(x)\|_Y < \epsilon$ for all $x \in X \setminus K$. The set of all continuous functions from X to $\mathcal{Y}$ that vanish at infinity is denoted $C_0(X, \mathcal{Y})$.

As it is not difficult to verify that $C_0(X, \mathcal{Y})$ is a vector subspace of $C_b(X, \mathcal{Y})$, we obtain that $C_0(X, \mathcal{Y})$ is a normed linear space with respect to the ∞-norm.

Example 1.2.23. If $X = \mathbb{N}$ is equipped with the discrete topology, then $C_0(X, \mathbb{K}) = c_0$ and the two infinity norms agree.

1.3 Constructing Normed Linear Spaces

As seen above, there are many normed linear spaces. Furthermore, we have already examined one way of creating normed linear spaces from others; take a vector subspace. In this section, we will examine how to take direct sums of normed linear spaces and quotients of normed linear spaces to obtain new normed linear spaces.

Proposition 1.3.1. *For each $n \in \mathbb{N}$, let $(\mathcal{X}_n, \|\cdot\|_n)$ be a normed linear space over $\mathbb{K}$. Let*

$$\mathcal{X} = \{(\vec{x}_n)_{n\geq 1} \mid \vec{x}_n \in \mathcal{X}_n \text{ for all } n \in \mathbb{N}\}.$$

1. *For $p \in [1, \infty)$, let*

$$\bigoplus_{n\in\mathbb{N}}^{p} \mathcal{X}_n = \left\{ (\vec{x}_n)_{n\geq 1} \in \mathcal{X} \,\middle|\, \sum_{n=1}^{\infty} \|\vec{x}_n\|_n^p < \infty \right\}.$$

Then $\bigoplus_{n\in\mathbb{N}}^{p} \mathcal{X}_n$ is a normed linear space over $\mathbb{K}$ together with the norm $\|\cdot\|_p : \bigoplus_{n\in\mathbb{N}}^{p} \mathcal{X}_n \to [0, \infty)$ defined by

$$\|(\vec{x}_n)_{n\geq 1}\|_p = \left(\sum_{n=1}^{\infty} \|\vec{x}_n\|_n^p \right)^{\frac{1}{p}}$$

for all $(\vec{x}_n)_{n\geq 1} \in \bigoplus_{n\in\mathbb{N}}^p \mathcal{X}_n$. *The space* $\bigoplus_{n\in\mathbb{N}}^p \mathcal{X}_n$ *is called the* ℓ_p-direct sum of $\{(\mathcal{X}_n, \|\cdot\|_n)\}_{n=1}^\infty$.

2. *Let*

$$\bigoplus_{n\in\mathbb{N}}^{\infty} \mathcal{X}_n = \left\{ (\vec{x}_n)_{n\geq 1} \in \mathcal{X} \,\middle|\, \sup_{n=1} \|\vec{x}_n\|_n < \infty \right\}.$$

Then $\bigoplus_{n\in\mathbb{N}}^\infty \mathcal{X}_n$ *is a normed linear space over* $\mathbb{K}$ *together with the norm* $\|\cdot\|_\infty : \bigoplus_{n\in\mathbb{N}}^\infty \mathcal{X}_n \to [0,\infty)$ *defined by*

$$\|(\vec{x}_n)_{n\geq 1}\|_\infty = \sup_{n\in\mathbb{N}} \|\vec{x}_n\|_n$$

for all $(\vec{x}_n)_{n\geq 1} \in \bigoplus_{n\in\mathbb{N}}^\infty \mathcal{X}_n$. *The space* $\bigoplus_{n\in\mathbb{N}}^\infty \mathcal{X}_n$ *is called the* ℓ_∞-direct sum of $\{(\mathcal{X}_n, \|\cdot\|_n)\}_{n=1}^\infty$.

3. *Let*

$$c_0(\mathcal{X}) = \left\{ (\vec{x}_n)_{n\geq 1} \in \mathcal{X} \,\middle|\, \lim_{n\to\infty} \|\vec{x}_n\|_n = 0 \right\}.$$

Then $c_0(\mathcal{X})$ *is a vector subspace of* $\bigoplus_{n\in\mathbb{N}}^\infty \mathcal{X}_n$ *and thus a normed linear space with the* ∞*-norm. The space* $c_0(\mathcal{X})$ *is called the* c_0-direct sum of $\{(\mathcal{X}_n, \|\cdot\|_n)\}_{n=1}^\infty$.

Proof. The proof that all of these spaces are vector spaces and that the described norms are indeed norms is straightforward (very similar to the proof in Appendix D that the p-norms are indeed norms). ∎

Remark 1.3.2. It is not difficult to see that the ℓ_p-direct sum and c_0-direct sum of $\mathbb{N}$ copies of $\mathbb{K}$ are $\ell_p(\mathbb{N}, \mathbb{K})$ and $c_0(\mathbb{K})$ respectively.

Opposed to building bigger spaces, we can mod-out normed linear spaces by closed subspaces.

Theorem 1.3.3. *Let* $(\mathcal{V}, \|\cdot\|)$ *be a normed linear space over* $\mathbb{K}$ *and let* $\mathcal{W}$ *be a closed vector subspace of* V. *Consider the quotient space* $\mathcal{V}/\mathcal{W}$*; that is*

$$\mathcal{V}/\mathcal{W} = \{\vec{v} + \mathcal{W} \mid \vec{v} \in \mathcal{V}\},$$

where

$$\vec{v} + \mathcal{W} = \{\vec{v} + \vec{w} \mid \vec{w} \in \mathcal{W}\}.$$

Then $\mathcal{V}/\mathcal{W}$ *is a vector space over* $\mathbb{K}$ *together with the operations*

$$(\vec{v}_1 + \mathcal{W}) + (\vec{v}_2 + \mathcal{W}) = (\vec{v}_1 + \vec{v}_2) + \mathcal{W} \quad \text{and} \quad \alpha \cdot (\vec{v} + \mathcal{W}) = (\alpha\vec{v}) + \mathcal{W}.$$

Define $p : \mathcal{V}/\mathcal{W} \to [0,\infty)$ *by*

$$p(\vec{v} + \mathcal{W}) = \inf\{\|\vec{v} + \vec{w}\| \mid \vec{w} \in \mathcal{W}\}$$

for all $\vec{v} \in \mathcal{V}$. *Then* p *is a well-defined norm on* $\mathcal{V}/\mathcal{W}$ *called* the quotient norm.

Proof. It is necessary to demonstrate that p is well-defined. Indeed if $\vec{v}_1, \vec{v}_2 \in \mathcal{V}$ are such that $\vec{v}_1 + \mathcal{W} = \vec{v}_2 + \mathcal{W}$, then $\vec{v}_1 = \vec{v}_2 + \vec{w}_0$ for some $\vec{w}_0 \in \mathcal{W}$. Hence

$$\begin{aligned}\inf\{\|\vec{v}_1 + \vec{w}\| \mid \vec{w} \in \mathcal{W}\} &= \inf\{\|\vec{v}_2 + \vec{w}_0 + \vec{w}\| \mid \vec{w} \in \mathcal{W}\} \\ &= \inf\{\|\vec{v}_2 + \vec{w}'\| \mid \vec{w}' \in \mathcal{W}\}\end{aligned}$$

as $\mathcal{W}$ is a vector subspace of $\mathcal{V}$. Hence p is well-defined.

To see that p is a norm, notice if $\vec{v} \in \mathcal{V}$ then clearly

$$0 \leq \inf\{\|\vec{v} + \vec{w}\| \mid \vec{w} \in \mathcal{W}\} \leq \left\|\vec{v} + \vec{0}\right\| < \infty$$

as $\vec{0} \in \mathcal{W}$. Hence $p : \mathcal{V}/\mathcal{W} \to [0, \infty)$. Next clearly $p(\vec{0} + \mathcal{W}) = 0$. Conversely, suppose $\vec{v} \in \mathcal{V}$ is such that $p(\vec{v}+\mathcal{W}) = 0$. Then for each $n \in \mathbb{N}$ there exists an $\vec{w}_n \in \mathcal{W}$ such that $\|\vec{v} + \vec{w}_n\| < \frac{1}{n}$. Hence $\vec{v} = \lim_{n\to\infty} -\vec{w}_n$. Since $-\vec{w}_n \in \mathcal{W}$ for all $n \in \mathbb{N}$ as $\mathcal{W}$ is a vector subspace, and since $\mathcal{W}$ is closed, $\vec{v} \in \mathcal{W}$. Hence $\vec{v} + \mathcal{W} = \vec{0} + \mathcal{W}$ as desired.

To see the second property of a norm, let $\vec{v} \in V$ and $\alpha \in \mathbb{K}$ be arbitrary. If $\alpha = 0$, clearly

$$p(\alpha \cdot (\vec{v} + \mathcal{W})) = p((\alpha\vec{v}) + \mathcal{W}) = p(\vec{0} + \mathcal{W}) = 0 = |\alpha| p(\vec{v} + \mathcal{W}).$$

Otherwise, if $\alpha \neq 0$ notice that

$$\left\{ \frac{1}{\alpha}\vec{w} \,\middle|\, \vec{w} \in \mathcal{W} \right\} = \mathcal{W}$$

as $\mathcal{W}$ is a vector subspace. Thus

$$\begin{aligned}p(\alpha \cdot (\vec{v} + \mathcal{W})) &= p((\alpha\vec{v}) + \mathcal{W}) \\ &= \inf\{\|\alpha\vec{v} + \vec{w}\| \mid \vec{w} \in \mathcal{W}\} \\ &= \inf\left\{ |\alpha| \left\|\vec{v} + \frac{1}{\alpha}\vec{w}\right\| \,\middle|\, \vec{w} \in \mathcal{W}\right\} \\ &= |\alpha| \inf\left\{ \left\|\vec{v} + \frac{1}{\alpha}\vec{w}\right\| \,\middle|\, \vec{w} \in \mathcal{W}\right\} \\ &= |\alpha| \inf\{\, \|\vec{v} + \vec{w}'\| \mid \vec{w}' \in \mathcal{W}\} \\ &= |\alpha| p(\vec{v} + \mathcal{W}).\end{aligned}$$

Hence p satisfies the second property of being a norm.

Finally, to see that p satisfies the triangle inequality, notice for all $\vec{v}_1, \vec{v}_2 \in V$ that

$$\begin{aligned}p((\vec{v}_1 + \mathcal{W}) + (\vec{v}_2 + \mathcal{W})) &= p((\vec{v}_1 + \vec{v}_2) + \mathcal{W}) \\ &= \inf\{\|\vec{v}_1 + \vec{v}_2 + \vec{w}\| \mid \vec{w} \in \mathcal{W}\} \\ &= \inf\{\|\vec{v}_1 + \vec{v}_2 + \vec{w}_1 + \vec{w}_2\| \mid \vec{w}_1, \vec{w}_2 \in \mathcal{W}\} \\ &\leq \inf\{\|\vec{v}_1 + \vec{w}_1\| + \|\vec{v}_2 + \vec{w}_2\| \mid \vec{w}_1, \vec{w}_2 \in \mathcal{W}\} \\ &= \inf\{\|\vec{v}_1 + \vec{w}_1\| \mid \vec{w}_1 \in \mathcal{W}\} + \inf\{\|\vec{v}_2 + \vec{w}_2\| \mid \vec{w}_2 \in \mathcal{W}\} \\ &= p(\vec{v}_1 + \mathcal{W}) + p(\vec{v}_2 + \mathcal{W})\end{aligned}$$

(where the third equality followed as $\mathcal{W}$ is a vector subspace). Hence p is a norm on $\mathcal{V}/\mathcal{W}$. ∎

Example 1.3.4. It is not difficult to show that the quotient space $c(\mathbb{K})/c_0(\mathbb{K})$ is simply $\mathbb{K}$ in disguise via the map that sends $(z_n)_{n\geq 1} + c_0(\mathbb{K}) \mapsto \lim_{n\to\infty} z_n$.

1.4 Bounded Linear Operators

Of course there are many other ways to construct normed linear spaces. One of the most important ways for functional analysis is to look at the morphisms between normed linear spaces. In particular, a morphism between two normed linear spaces should preserve the vector space structure and thus be a linear map. Moreover, to preserve the norm, we would expect some form of continuity. Continuous linear maps are a staple of functional analysis and thus will be introduced in this section for use throughout the course. To begin, we first desire to put a norm structure on certain linear maps between two normed linear spaces.

Definition 1.4.1. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces over $\mathbb{K}$. A linear map $T : \mathcal{X} \to \mathcal{Y}$ is said to be *bounded* if

$$\sup\left\{\|T(\vec{x})\|_{\mathcal{Y}} \mid \vec{x} \in \mathcal{X}, \|\vec{x}\|_{\mathcal{X}} \leq 1\right\} < \infty.$$

If T is bounded, we write

$$\|T\| = \sup\{\|T(\vec{x})\|_{\mathcal{Y}} \mid \vec{x} \in \mathcal{X}, \|\vec{x}\|_{\mathcal{X}} \leq 1\}.$$

The quantity $\|T\|$ is called the *operator norm* of T. Furthermore, the set of bounded linear maps from $\mathcal{X}$ to $\mathcal{Y}$ is denoted $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Remark 1.4.2. Note we can only discuss bounded linear maps between normed linear spaces over the same field. Thus throughout these notes, this will be a standing assumption when discussing bounded linear maps.

In addition, note that $\|T\|$ is a measure of how large the unit ball (the ball of radius 1 centred at $\vec{0}$) in $\mathcal{X}$ is scaled by applying T.

Unsurprisingly, the operator is in fact a norm thereby yielding the following.

Theorem 1.4.3. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces over $\mathbb{K}$. Then $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a normed linear space over $\mathbb{K}$ with the operator norm as defined in Definition 1.4.1.*

To see that the operator norm is indeed a norm, we note that the only non-trivial property of Definition 1.1.1 to verify is that if $\|T\| = 0$, then T is the zero linear map. Note the following lemma yields the result.

Lemma 1.4.4. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces over $\mathbb{K}$ and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then*

$$\|T(\vec{x})\|_{\mathcal{Y}} \leq \|T\| \, \|\vec{x}\|_{\mathcal{X}}$$

for all $\vec{x} \in \mathcal{X}$.

Proof. Since $\left\|T(\vec{0})\right\|_{\mathcal{Y}} = \left\|\vec{0}\right\|_{\mathcal{X}} = 0$, the result holds when $\vec{x} = \vec{0}$. If $\vec{x} \neq \vec{0}$, then $\|\vec{x}\|_{\mathcal{X}} \neq 0$. Consequently, as

$$\left\| \frac{1}{\|\vec{x}\|_{\mathcal{X}}} \vec{x} \right\|_{\mathcal{X}} = \frac{1}{\|\vec{x}\|_{\mathcal{X}}} \|\vec{x}\|_{\mathcal{X}} = 1,$$

we obtain from the definition of the operator norm that

$$\frac{1}{\|\vec{x}\|_{\mathcal{X}}} \|T(\vec{x})\|_{\mathcal{Y}} = \left\| \frac{1}{\|\vec{x}\|_{\mathcal{X}}} T(\vec{x}) \right\|_{\mathcal{Y}} = \left\| T\left(\frac{1}{\|\vec{x}\|_{\mathcal{X}}} \vec{x} \right) \right\|_{\mathcal{Y}} \leq \|T\| .$$

Therefore $\|T(\vec{x})\|_{\mathcal{Y}} \leq \|T\| \, \|\vec{x}\|_{\mathcal{X}}$ as desired. ∎

This easily yields that the composition of bounded linear maps is bounded.

Corollary 1.4.5. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, and $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ be normed linear spaces over $\mathbb{K}$. If $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, then $ST \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$ with $\|ST\| \leq \|S\| \, \|T\|$.*

Proof. Notice for all $\vec{x} \in \mathcal{X}$ that

$$\|(ST)(\vec{x})\|_{\mathcal{Z}} = \|S(T(\vec{x}))\|_{Z} \leq \|S\| \, \|T(\vec{x})\|_{\mathcal{Y}} \leq \|S\| \, \|T\| \, \|\vec{x}\|_{\mathcal{X}} .$$

Thus taking the supremum of this over all $\vec{x} \in \mathcal{X}$ with $\|\vec{x}\|_{\mathcal{X}} \leq 1$ yields the result. ∎

The reason we have been analyzing bounded linear maps in reference to continuous linear maps is the following.

Theorem 1.4.6. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces over $\mathbb{K}$ and let $T : \mathcal{X} \to \mathcal{Y}$ be linear. The following are equivalent:*

(1) T is continuous.

(2) T is continuous at 0.

(3) T is bounded.

Proof. Clearly (1) implies (2). To see that (2) implies (3), let $\epsilon = 1$. Since T is continuous at 0, there exists a $\delta > 0$ such that if $\|\vec{x}\|_{\mathcal{X}} \leq \delta$ then $\|T(\vec{x})\|_{\mathcal{Y}} \leq 1$. Therefore, if $\vec{x} \in \mathcal{X}$ is such that $\|\vec{x}\|_{\mathcal{X}} \leq 1$, then $\|\delta\vec{x}\|_{\mathcal{X}} \leq \delta$ so

$$\delta \|T(\vec{x})\|_{\mathcal{Y}} = \|\delta T(\vec{x})\|_{\mathcal{Y}} = \|T(\delta\vec{x})\|_{\mathcal{Y}} \leq 1.$$

Hence $\|\vec{x}\|_{\mathcal{X}} \leq 1$ implies $\|T(\vec{x})\|_{\mathcal{Y}} \leq \delta^{-1}$ so T is bounded with $\|T\| \leq \delta^{-1}$ by definition.

To see that (3) implies (1), let $\vec{x}_0 \in \mathcal{X}$ be arbitrary. To see that T is continuous at x, let $\epsilon > 0$ and let $\delta = \frac{\epsilon}{\|T\|+1} > 0$. If $\vec{x} \in \mathcal{X}$ is such that $\|\vec{x} - \vec{x}_0\|_{\mathcal{X}} < \delta$, then Lemma 1.4.4 implies that

$$\|T(\vec{x}) - T(\vec{x}_0)\|_{\mathcal{Y}} = \|T(\vec{x} - \vec{x}_0)\|_{\mathcal{Y}} \leq \|T\| \, \|\vec{x} - \vec{x}_0\|_{\mathcal{X}} < \|T\| \frac{\epsilon}{\|T\| + 1} < \epsilon.$$

Therefore T is continuous at $\vec{x}_0$ as $\epsilon > 0$ was arbitrary. Therefore, as $\vec{x}_0 \in \mathcal{X}$ was arbitrary, T is continuous on $\mathcal{X}$. ∎

Perhaps it is surprising at this point in the course, but $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is one of the most important normed linear spaces! Thus it is useful to include some examples. Note it is often quite difficult to actually compute the operator norm of a bounded linear map. However, in the finite dimensional world, things are not too bad (in theory; in practice, computing the necessary quantities can be challenging in dimensions exceeding 4).

Theorem 1.4.7. *Let $A \in \mathcal{M}_n(\mathbb{C})$ and define $L_A : \mathbb{C}^n \to \mathbb{C}^n$ by $L_A(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{C}^n$ (where we write $\vec{x}$ as a column vector and use matrix multiplication). Then*

$$\|L_A\| = \max\left\{\sqrt{\lambda} \;\middle|\; \lambda \text{ an eigenvalue for } A^*A\right\}.$$

Proof. Exercise. ∎

Example 1.4.8. For $p \in [1, \infty]$, define $F, B : \ell_p(\mathbb{N}) \to \ell_p(\mathbb{N})$ by

$$F((x_1, x_2, x_3, \ldots)) = (0, x_1, x_2, \ldots)$$
$$B((x_1, x_2, x_3, \ldots)) = (x_2, x_3, x_4, \ldots).$$

Then F and B are bounded linear operators with $\|F\| = \|B\| = 1$. The operator F is called the *unilateral forward shift* and the operator B is called the *unilateral backward shift*.

Example 1.4.9. For $p \in [1, \infty]$, define $F, B : \ell_p(\mathbb{Z}) \to \ell_p(\mathbb{Z})$ by

$$F((x_n)_{n\in\mathbb{Z}}) = (x_{n-1})_{n\in\mathbb{Z}}$$
$$B((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}.$$

Then F and B are bounded linear operators with $\|F\| = \|B\| = 1$. The operator F is called the *bilateral forward shift* and the operator B is called the *bilateral backward shift*.

Example 1.4.10. Given a compact Hausdorff space $(X, \mathcal{T})$ and a fixed $f \in C(X)$, define $M_f : C(X) \to C(X)$ by

$$(M_f(g))(x) = f(x)g(x)$$

for all $x \in X$ and $g \in C(X)$. Clearly M_f is a linear map such that $\|M_f(g)\|_\infty \leq \|f\|_\infty \|g\|_\infty$ for all $g \in C(X)$. Moreover, by setting $g(x) = 1$ for all $x \in X$, we see that $\|g\|_\infty = 1$ and $\|M_f(g)\|_\infty = \|f\|_\infty$. Hence M_f is a bounded linear map with $\|M_f\| = \|f\|_\infty$ when $C(X)$ is equipped with the infinity norm.

Example 1.4.11. Let $(X, \mathcal{A}, \mu)$ be a measure space, let $p \in [1, \infty]$, and fix $f \in L_\infty(X, \mu)$. Define $M_f : L_p(X, \mu) \to L_p(X, \mu)$ by

$$(M_f(g))(x) = f(x)g(x)$$

for all $x \in X$ and $L_p(X, \mu)$. We note that M_f will be well-defined as the product of measurable functions is measurable and as

$$\begin{aligned}\|M_f(g)\|_p &= \left(\int_X |f(x)g(x)|^p \, d\mu\right)^{\frac{1}{p}} \\ &\leq \left(\int_X \|f\|_\infty^p \, |g(x)|^p \, d\mu\right)^{\frac{1}{p}} \\ &= (\|f\|_\infty^p)^{\frac{1}{p}} \left(\int_X |g(x)|^p \, d\mu\right)^{\frac{1}{p}} = \|f\|_\infty \|g\|_p .\end{aligned}$$

As M_f is clearly linear, we see that M_f is a bounded linear map with $\|M_f\| \leq \|f\|_\infty$.

Using techniques from measure theory, it is not difficult to verify that $\|M_f\| = \|f\|_\infty$ in the case that μ is inner regular.

Example 1.4.12. Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Recall $\mathcal{V}/\mathcal{W}$ is a normed linear space with respect to the quotient norm from Theorem 1.3.3. Define $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ by $q(\vec{v}) = \vec{v} + \mathcal{W}$. Clearly q is a linear map. Moreover, as $\|q(\vec{v})\| \leq \|\vec{v}\|$ by definition, q is a bounded linear map with $\|q\| \leq 1$.

Of course, there are a plethora and inexhaustible list of bounded linear maps that cannot possibly be written down. To complete this section, there are a couple additional examples that should be discussed. In particular, not every nice linear map is bounded.

Example 1.4.13. Let $\mathcal{P}(\mathbb{R})$ denote the set of polynomials with real coefficients. As $\mathcal{P}(\mathbb{R}) \subseteq C[0, 1]$, we see that $\mathcal{P}(\mathbb{R})$ is a normed linear space when equipped with the infinity norm. Define $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ by

$$D(p) = p'$$

for all $p \in \mathcal{P}(\mathbb{R})$ (i.e. take the derivative). Clearly D is a linear map. However, D is not bounded since $\|x^n\|_\infty = 1$ for all $n \in \mathbb{N}$ yet $D(x^n) = nx^{n-1}$ so $\|D(x^n)\|_\infty = n$ for all $n \in \mathbb{N}$.

Example 1.4.14. Define $T : c_{00} \to \mathbb{K}$ by

$$T((z_n)_{n\geq 1}) = \sum_{n=1}^{\infty} z_n$$

for all $(z_n)_{n\geq 1} \in c_{00}$. Clearly T is a well-defined linear map as for all $(z_n)_{n\geq 1} \in c_{00}$ there exists an $N \in \mathbb{N}$ such that $z_n = 0$ for all $n \geq N$. However, T is not bounded. To see this, note if $\vec{x}_m = (z_{m,n})_{n\geq 1} \in c_{00}$ are defined such that

$$z_{m,n} = \begin{cases} 1 & \text{if } n \leq m \\ 0 & \text{if } n > m \end{cases},$$

then $\|\vec{x}_m\|_\infty = 1$ whereas $|T(\vec{x}_m)| = m$. Hence $\|T\| \geq m$ for every $m \in \mathbb{N}$ and thus T is not bounded.

However, the question of whether or not a linear map is bounded depends completely on the norms under consideration.

Example 1.4.15. Fix $x_0 \in [0, 1]$ and define $T_{x_0} : C[0, 1] \to \mathbb{R}$ by

$$T_{x_0}(f) = f(x_0)$$

for all $f \in C[0, 1]$. Clearly T_{x_0} is a linear map. Whether T_{x_0} is bounded or not depends on the norm we place on $C[0, 1]$ (we will always use the absolute value on $\mathbb{R}$).

If $C[0, 1]$ is equipped with the infinity norm, then T_{x_0} is bounded with $\|T_{x_0}\| = 1$. To see this, notice for all $f \in C[0, 1]$ that

$$|T_{x_0}(f)| = |f(x_0)| \leq \|f\|_\infty$$

by definition of the infinity norm. Hence T_{x_0} is bounded with $\|T_{x_0}\| \leq 1$. To see that $\|T_{x_0}\| = 1$, notice the function $g(x) = 1$ for all $x \in [0, 1]$ is an element of $C[0, 1]$ with $\|g\|_\infty = 1$. Since

$$|T_{x_0}(g)| = |g(x_0)| = 1,$$

However, if $C[0, 1]$ is equipped with the 1-norm, then T_{x_0} is not bounded. To see this, for each $n \in \mathbb{N}$, define $f_n \in C[0, 1]$ by

$$f_n(x) = \begin{cases} 2n^2\left(x - \left(x_0 - \frac{1}{2n}\right)\right) & \text{if } x \in \left[x_0 - \frac{1}{2n}, x_0\right] \\ -2n^2\left(x - \left(x_0 + \frac{1}{2n}\right)\right) & \text{if } x \in \left[x_0, x_0 + \frac{1}{2n}\right] \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in [0,1]$. It is not difficult to see that

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

regardless of the value of x_0 (in fact, if $x_0 \in \{0,1\}$ then $\|f_n\|_1 = 1$ for all $n \in \mathbb{N}$, and if $x_0 \notin \{0,1\}$ then $\|f_n\|_1 = 1$ for sufficiently large n). Furthermore, as

$$|T_{x_0}(f_n)| = |f_n(x_0)| = n$$

we obtain that

$$\sup\{|T_{x_0}(f)| \mid f \in C[0,1], \|f\|_1 \leq 1\} = \infty$$

so T_{x_0} is unbounded.

1.5 Dual Spaces

Not only does Example 1.4.15 show the norm under consideration affects whether or not a linear map is bounded, but it leads us to an important class of bounded linear maps. Given any normed linear space, the bounded linear maps into the scalars play a vital role in functional analysis as will be seen in this course. Thus they deserve some special treatment.

Definition 1.5.1. Given a normed linear space $(\mathcal{X}, \|\cdot\|)$ over $\mathbb{K}$, the *dual space* of $(\mathcal{X}, \|\cdot\|)$, denoted $\mathcal{X}^*$, is $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{K})$. The elements of $\mathcal{X}^*$ are called *continuous linear functionals*.

Of course there are many examples of continuous linear functionals. We begin in the finite dimensional world.

Example 1.5.2. Let $\mathcal{X} = \mathbb{C}^n$ equipped with the Euclidean norm. Recall if $f : \mathcal{X} \to \mathbb{C}$ is a non-zero linear function, then there exists a unique $\vec{a} = (a_1, \ldots, a_n) \in \mathbb{C}^n$ such that

$$f((z_1, \ldots, z_n)) = \sum_{k=1}^{n} a_k z_k$$

for all $(z_1, \ldots, z_n) \in \mathbb{C}^n$. By the Cauchy-Schwarz inequality, we know for all $\vec{z} \in \mathbb{C}^n$ that

$$|f(\vec{z})| \leq \|\vec{a}\|_2 \|\vec{z}\|_2$$

with equality if and only if $\vec{z}$ is a multiple of $\vec{a}$. Hence we easily see that $\|f\| \leq \|\vec{a}\|_2$ so every linear functional is continuous. Furthermore, we see by taking $\vec{z} = \frac{1}{\|\vec{a}\|_2}\vec{a}$ (so $\|\vec{z}\|_2 = 1$) that $|f(\vec{z})| = \|\vec{a}\|_2$ and hence $\|f\| =$

$\|\vec{a}\|_2$. Thus $\mathcal{X}^*$ is isometrically isomorphic to $\mathcal{X}$ via the map that sends $(a_1, \ldots, a_n) \in \mathbb{C}^n$ to the linear functional $f : \mathcal{X} \to \mathbb{C}$ defined by

$$f((z_1, \ldots, z_n)) = \sum_{k=1}^{n} a_k z_k$$

for all $(z_1, \ldots, z_n) \in \mathbb{C}^n$

For another example, Example 1.4.15 can be extended.

Example 1.5.3. Let $(X, \mathcal{T})$ be a compact Hausdorff space and let $x_0 \in X$ be fixed. Define $\delta_{x_0} : C(X) \to \mathbb{R}$ by

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

for all $f \in C(X)$. Then δ_{x_0} is a continuous linear functional with $\|\delta_{x_0}\| = 1$ when $C(X)$ is equipped with the infinity norm. We call δ_{x_0} the *point-mass linear functional at* x_0.

Of course, there are more interesting continuous linear functionals to consider. In fact, for post ℓ_p-spaces, we can completely describe their dual spaces with other ℓ_p-spaces.

Theorem 1.5.4 (Riesz Representation Theorem, $\ell_p(\mathbb{N})$). *Let $p \in [1, \infty)$ and $q \in (1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For each vector $\vec{y} = (y_n)_{n\geq 1} \in \ell_q(\mathbb{N})$, define $\varphi_{\vec{y}} : \ell_p(\mathbb{N}) \to \mathbb{K}$ by*

$$\varphi_{\vec{y}}((x_n)_{n\geq 1}) = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Then $\varphi_{\vec{y}} \in (\ell_p(\mathbb{N}))^*$. Moreover, the map $\Phi : \ell_q(\mathbb{N}) \to (\ell_p(\mathbb{N}))^*$ defined by

$$\Phi(\vec{y}) = \varphi_{\vec{y}}$$

for all $\vec{y} \in \ell_q(\mathbb{N})$ is a bijective linear map such that $\|\Phi(\vec{y})\| = \|\vec{y}\|_q$ for all $\vec{y} \in \ell_q(\mathbb{N})$.

Proof. To see that $\varphi_{\vec{y}}$ is a well-defined bounded continuous linear functional, note since $\vec{y} = (y_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ we have for all $(x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ that $(x_n y_n)_{n\geq 1} \in \ell_1(\mathbb{N})$ by Hölders' inequality. Hence

$$\sum_{n=1}^{\infty} |x_n y_n| < \infty$$

and thus $\varphi_{\vec{y}}((x_n)_{n\geq 1})$ is a well-defined element of $\mathbb{K}$. Hence $\varphi_{\vec{y}}$ is well-defined. Furthermore, the fact that $\varphi_{\vec{y}}$ is linear follows from basic properties of convergent series.

To see that $\varphi_{\vec{y}}$ is bounded, notice for all $(x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ that

$$|\varphi_{\vec{y}}((x_n)_{n\geq 1})| = \left|\sum_{n=1}^{\infty} x_n y_n\right| \leq \sum_{n=1}^{\infty} |x_n y_n| \leq \left\|(x_n)_{n\geq 1}\right\|_p \left\|\vec{y}\right\|_q$$

by Hölder's inequality. Hence we easily see that $\varphi_{\vec{y}}$ is bounded and $\left\|\varphi_{\vec{y}}\right\| \leq \left\|\vec{y}\right\|_q$.

The above implies that Φ is well-defined. Moreover, it is elementary to verify that Φ is a linear map. Hence it remains only to show that Φ is bijective and that $\|\Phi(\vec{y})\| = \left\|\vec{y}\right\|_q$ for all $\vec{y} \in \ell_q(\mathbb{N})$. We will show that Φ is surjective and in the process show that $\|\Phi(\vec{y})\| \geq \left\|\vec{y}\right\|_q$ thereby completing the equality. The result will then follow as the norm equality shows the kernel of Φ is simply the zero vector and thus Φ will be injective.

Let $\varphi \in (\ell_p(\mathbb{N}))^*$ be arbitrary. For each $n \in \mathbb{N}$, let $\vec{e}_n$ be the sequence with a 1 in the n^{th} entry and zeros everywhere else. As $\vec{e}_n \in \ell_p(\mathbb{N})$ for all $n \in \mathbb{N}$, the element

$$y_n = \varphi(\vec{e}_n) \in \mathbb{K}$$

is well-defined.

Let $\vec{y} = (y_n)_{n\geq 1}$. We claim that $\vec{y} \in \ell_q(\mathbb{N})$ and that $\varphi = \varphi_{\vec{y}}$. To see that $\vec{y} \in \ell_q(\mathbb{N})$, first consider the case where $p = 1$ and $q = \infty$. As $\|\vec{e}_n\|_1 = 1$ for all $n \in \mathbb{N}$, we obtain that

$$|y_n| = |\varphi(\vec{e}_n)| \leq \|\varphi\|$$

for all $n \in \mathbb{N}$. Hence $\vec{y} \in \ell_\infty(\mathbb{N})$ with $\left\|\vec{y}\right\|_\infty \leq \|\varphi\|$. Thus, once $\varphi = \varphi_{\vec{y}}$ is established, we will have that

$$\left\|\vec{y}\right\|_\infty \leq \|\varphi\| = \left\|\varphi_{\vec{y}}\right\| \leq \left\|\vec{y}\right\|_\infty$$

thereby completing the norm equality.

For $p \neq 1$, for each $n \in \mathbb{N}$ let

$$x_n = \begin{cases} 0 & \text{if } y_n = 0 \\ \overline{y_n}(|y_n|)^{\frac{p}{p-1}-2} & \text{if } y_n \neq 0 \end{cases}$$

and for $N \in \mathbb{N}$ consider the sequence

$$\vec{z}_N = (x_1, x_2, \ldots, x_n, 0, 0, \ldots).$$

Clearly $\vec{Z}_N \in \ell_p(\mathbb{N})$ with

$$\begin{aligned}\left\|\vec{z}_N\right\|_p &= \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}} \\ &= \left(\sum_{n=1}^{N} |y_n|^{p+\frac{p^2}{p-1}-2p}\right)^{\frac{1}{p}} \\ &= \left(\sum_{n=1}^{N} |y_n|^{\frac{p}{p-1}}\right)^{\frac{1}{p}} \\ &= \left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{p}}\end{aligned}$$

as $\frac{1}{p} + \frac{1}{q} = 1$ so $q = \frac{p}{p-1}$. Therefore, as

$$\varphi(\vec{z}_N) = \sum_{n=1}^{N} x_n y_n = \sum_{n=1}^{N} |y_n|^{\frac{p}{p-1}}$$

due to linearity, we obtain that

$$\sum_{n=1}^{N} |y_n|^q = \sum_{n=1}^{N} |y_n|^{\frac{p}{p-1}} = \|\varphi(\vec{z}_N)\| \leq \|\varphi\| \left\|\vec{z}_N\right\|_p = \|\varphi\| \left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{p}}.$$

Hence

$$\left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{q}} = \left(\sum_{n=1}^{N} |y_n|^q\right)^{1-\frac{1}{p}} \leq \|\varphi\|$$

for every $N \in \mathbb{N}$. Therefore, by taking the supremum over N, we obtain $\vec{y} \in \ell_q(\mathbb{N})$ and $\|\vec{y}\|_q \leq \|\varphi\|$. Thus, once $\varphi = \varphi_{\vec{y}}$ is established, we will have that

$$\|\vec{y}\|_q \leq \|\varphi\| = \|\varphi_{\vec{y}}\| \leq \|\vec{y}\|_q$$

thereby completing the norm equality.

To see that $\varphi = \varphi_{\vec{y}}$, let $\vec{x} = (x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ be arbitrary. For each $N \in \mathbb{N}$, let

$$\vec{x}_N = (x_1, x_2, \ldots, x_N, 0, 0, \ldots).$$

Clearly $\vec{x}_N \in \ell_p(\mathbb{N})$, $\lim_{N\to\infty} \|\vec{x} - \vec{x}_N\|_p = 0$ as $p \neq \infty$, and $\varphi(\vec{x}_N) = \varphi_{\vec{y}}(\vec{x}_N)$ for all $N \in \mathbb{N}$ by the definition of $\vec{y}$ and by linearity. Therefore, as φ and $\varphi_{\vec{y}}$ are bounded linear functionals and thus continuous, we obtain that

$$\varphi(\vec{x}) = \lim_{N\to\infty} \varphi(\vec{x}_N) = \lim_{N\to\infty} \varphi_{\vec{y}}(\vec{x}_N) = \varphi_{\vec{y}}(\vec{x}).$$

Hence, as $\vec{x} \in \ell_p(\mathbb{N})$ was arbitrary, $\varphi = \varphi_{\vec{y}}$ thereby completing the proof. (i.e. $c_{00}(\mathbb{N})$ is dense in $\ell_p(\mathbb{N})$ for $p \neq \infty$ and thus as φ and $\varphi_{\vec{y}}$ agree on $c_{00}(\mathbb{N})$ by linearity, continuity yields the result.) ∎

The map Φ in Theorem 1.5.4 is an important example of maps that preserve the vector space and norm structures between normed linear spaces and thus deserves a name.

Definition 1.5.5. Let $(\mathcal{X}, \|\cdot\|_X)$ and $(\mathcal{Y}, \|\cdot\|_Y)$ be normed linear spaces. A map $\Phi : \mathcal{X} \to \mathcal{Y}$ is said to be an *isomorphism* if Φ is a bijective bounded linear map with bounded inverse. In this case $(\mathcal{X}, \|\cdot\|_X)$ and $(\mathcal{Y}, \|\cdot\|_Y)$ are said to be *isomorphic*.

In the case that $\|\Phi(\vec{x})\|_{\mathcal{Y}} = \|\vec{x}\|_{\mathcal{X}}$ for all $\vec{x} \in \mathcal{X}$, it is said that Φ is an *isometric isomorphism* and that $(\mathcal{X}, \|\cdot\|_X)$ and $(\mathcal{Y}, \|\cdot\|_Y)$ are said to be *isometrically isomorphic*.

Remark 1.5.6. It is clear that if $(\mathcal{X}, \|\cdot\|_X)$ and $(\mathcal{Y}, \|\cdot\|_Y)$ are normed linear spaces and $\Phi : \mathcal{X} \to \mathcal{Y}$ is an isometric isomorphism, then $(\mathcal{X}, \|\cdot\|_X)$ and $(\mathcal{Y}, \|\cdot\|_Y)$ are truly the same normed linear space as Φ being a bijective linear map means $\mathcal{X}$ and $\mathcal{Y}$ are the same vector space and Φ being isometric implies the norms are identical under this identification.

When Φ is only an isomorphism, notice for all $\vec{x} \in \mathcal{X}$ that

$$\|\Phi(\vec{x})\|_{\mathcal{Y}} \leq \|\Phi\| \, \|\vec{x}\|_{\mathcal{X}}$$

and

$$\|\vec{x}\|_{\mathcal{X}} = \left\|\Phi^{-1}(\Phi(\vec{x}))\right\|_{\mathcal{X}} \leq \left\|\Phi^{-1}\right\| \|\Phi(\vec{x})\|_{\mathcal{Y}}.$$

Hence

$$\frac{1}{\|\Phi^{-1}\|} \|\vec{x}\|_{\mathcal{X}} \leq \|\Phi(\vec{x})\|_{\mathcal{Y}} \leq \|\Phi\| \, \|\vec{x}\|_{\mathcal{X}}$$

for all $\vec{x} \in \mathcal{X}$ (note $\|\Phi^{-1}\| \neq 0$ for otherwise Φ^{-1} would be the zero linear map). This shows, upto identifying the vector spaces $\mathcal{X}$ and $\mathcal{Y}$, the norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$ are equivalent.

Of course, there are other dual spaces of sequence spaces we can identify.

Theorem 1.5.7 (Riesz Representation Theorem, c_0). *For each vector $\vec{y} = (y_n)_{n\geq 1} \in \ell_1(\mathbb{N})$, define $\varphi_{\vec{y}} : c_0 \to \mathbb{K}$ by*

$$\varphi_{\vec{y}}((x_n)_{n\geq 1}) = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in c_0$. Then $\varphi_{\vec{y}} \in (c_0)^*$. Moreover, the map $\Phi : \ell_1(\mathbb{N}) \to (c_0)^*$ defined by

$$\Phi(\vec{y}) = \varphi_{\vec{y}}$$

for all $\vec{y} \in \ell_1(\mathbb{N})$ is a bijective linear map such that $\|\Phi(\vec{y})\| = \|\vec{y}\|_1$ for all $\vec{y} \in \ell_1(\mathbb{N})$.

Proof. To see that $\varphi_{\vec{y}}$ is a well-defined bounded linear functional, note since $\vec{y} = (y_n)_{n\geq 1} \in \ell_1(\mathbb{N})$ we have for all $(x_n)_{n\geq 1} \in c_0 \subseteq \ell_\infty(\mathbb{N})$ that $(x_n y_n)_{n\geq 1} \in \ell_1(\mathbb{N})$ by Hölders' inequality. Hence

$$\sum_{n=1}^{\infty} |x_n y_n| < \infty$$

and thus $\varphi_{\vec{y}}((x_n)_{n\geq 1})$ is a well-defined element of $\mathbb{K}$. Hence $\varphi_{\vec{y}}$ is well-defined. Furthermore, the fact that $\varphi_{\vec{y}}$ is linear follows from basic properties of convergent series.

To see that $\varphi_{\vec{y}}$ is bounded, notice for all $(x_n)_{n\geq 1} \in c_0$ that

$$|\varphi_{\vec{y}}((x_n)_{n\geq 1})| = \left|\sum_{n=1}^{\infty} x_n y_n\right| \leq \sum_{n=1}^{\infty} |x_n y_n| \leq \|(x_n)_{n\geq 1}\|_\infty \|\vec{y}\|_1$$

by Hölder's inequality. Hence we easily see that $\varphi_{\vec{y}}$ is bounded and $\|\varphi_{\vec{y}}\| \leq \|\vec{y}\|_1$.

The above implies that Φ is well-defined. Moreover, it is elementary to verify that Φ is a linear map. Hence it remains only to show that Φ is bijective and that $\|\Phi(\vec{y})\| = \|\vec{y}\|_1$ for all $\vec{y} \in \ell_1(\mathbb{N})$. We will show that Φ is surjective and in the process show that $\|\Phi(\vec{y})\| \geq \|\vec{y}\|_1$ thereby completing the equality. The result will then follow as the norm equality shows the kernel of Φ is simply the zero vector and thus Φ will be injective.

Let $\varphi \in (c_0)^*$ be arbitrary. For each $n \in \mathbb{N}$, let $\vec{e}_n$ be the sequence with a 1 in the n^{th} entry and zeros everywhere else. As $\vec{e}_n \in c_0$ for all $n \in \mathbb{N}$, the element

$$y_n = \varphi(\vec{e}_n) \in \mathbb{K}$$

is well-defined.

Let $\vec{y} = (y_n)_{n\geq 1}$. We claim that $\vec{y} \in \ell_1(\mathbb{N})$ and that $\varphi = \varphi_{\vec{y}}$. To see that $\vec{y} \in \ell_1(\mathbb{N})$, for each $n \in \mathbb{N}$ let

$$x_n = \begin{cases} 0 & \text{if } y_n = 0 \\ \frac{|y_n|}{y_n} & \text{if } y_n \neq 0 \end{cases}$$

and for $N \in \mathbb{N}$ consider the sequence

$$\vec{z}_N = (x_1, x_2, \ldots, x_n, 0, 0, \ldots).$$

Clearly $\vec{x}_N \in c_0$ with $\|\vec{z}_N\|_\infty = 1$. Therefore, as

$$\varphi(\vec{z}_N) = \sum_{n=1}^{N} x_n y_n = \sum_{n=1}^{N} |y_n|$$

due to linearity, we obtain that

$$\sum_{n=1}^{N} |y_n| = |\varphi(\vec{z}_N)| \leq \|\varphi\| \|\vec{z}_N\|_\infty = \|\varphi\|$$

for every $N \in \mathbb{N}$. Therefore, by taking the supremum over N, we obtain $\vec{y} \in \ell_1(\mathbb{N})$ and $\|\vec{y}\|_1 \leq \|\varphi\|$. Thus, once $\varphi = \varphi_{\vec{y}}$ is established, we will have that

$$\|\vec{y}\|_1 \leq \|\varphi\| = \|\varphi_{\vec{y}}\| \leq \|\vec{y}\|_1$$

thereby completing the norm equality.

To see that $\varphi = \varphi_{\vec{y}}$, let $\vec{x} = (x_n)_{n\geq 1} \in c_0$ be arbitrary. For each $N \in \mathbb{N}$, let

$$\vec{x}_N = (x_1, x_2, \ldots, x_N, 0, 0, \ldots).$$

Clearly $\vec{x}_N \in c_0$, $\lim_{N\to\infty} \|\vec{x} - \vec{x}_N\|_\infty = 0$ as $\vec{x} \in c_0$, and $\varphi(\vec{x}_N) = \varphi_{\vec{y}}(\vec{x}_N)$ for all $N \in \mathbb{N}$ by the definition of $\vec{y}$ and by linearity. Therefore, as φ and $\varphi_{\vec{y}}$ are bounded linear functionals and thus continuous, we obtain that

$$\varphi(\vec{x}) = \lim_{N\to\infty} \varphi(\vec{x}_N) = \lim_{N\to\infty} \varphi_{\vec{y}}(\vec{x}_N) = \varphi_{\vec{y}}(\vec{x}).$$

Hence, as $\vec{x} \in c_0$ was arbitrary, $\varphi = \varphi_{\vec{y}}$ thereby completing the proof. (i.e. c_{00} is dense in c_0 and thus as φ and $\varphi_{\vec{y}}$ agree on c_0 by linearity, continuity yields the result.) ∎

However, there is another normed linear space that has $\ell_1(\mathbb{N})$ as its dual space.

Theorem 1.5.8 (Riesz Representation Theorem, c). *Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For each vector $\vec{y} = (y_n)_{n\geq 0} \in \ell_1(\mathbb{N}_0)$, define $\varphi_{\vec{y}} : c \to \mathbb{K}$ by*

$$\varphi_{\vec{y}}((x_n)_{n\geq 1}) = y_0 \left(\lim_{n\to\infty} x_n\right) + \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in c$. Then $\varphi_{\vec{y}} \in c^*$. Moreover, the map $\Phi : \ell_1(\mathbb{N}) \to c^*$ defined by

$$\Phi(\vec{y}) = \varphi_{\vec{y}}$$

for all $\vec{y} \in \ell_1(\mathbb{N})$ is a bijective linear map such that $\|\Phi(\vec{y})\| = \|\vec{y}\|_1$ for all $\vec{y} \in \ell_1(\mathbb{N})$.

Proof. Exercise. ∎

Remark 1.5.9. The above Riesz Representation Theorems raise the question on what exactly is $(\ell_\infty(\mathbb{N}))^*$? Of course, it is not too difficult to see that if $\vec{y} = (y_n)_{n\geq 1} \in \ell_1(\mathbb{N})$, then $\varphi_{\vec{y}} : \ell_\infty(\mathbb{N}) \to \mathbb{K}$ defined by

$$\varphi_{\vec{y}}((x_n)_{n\geq 1}) = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in \ell_\infty(\mathbb{N})$ is a bounded linear map with $\|\varphi_{\vec{y}}\| = \|\vec{y}\|_1$ by the same arguments as used in Theorem 1.5.7 and Theorem 1.5.8. The issue comes down demonstrating every continuous linear functional is of this form as, in fact, it is not true.

To construct the dual space of $\ell_\infty(\mathbb{N})$, one actually requires some measure theory which we leave to Appendix D.4. In fact, there are generalizations of the above Riesz Representation Theorems to $L_p(X,\mu)$ using measure theory that can also be found in Appendix D.4.

Alternatively, in topology it can be shown that $\ell_\infty(\mathbb{N})$ is isometrically isomorphic to the continuous functions on the Stone-Čech compactification of the natural numbers. Thus it is possible to describe the dual space of $\ell_\infty(\mathbb{N})$ using the dual space of the continuous functions on a compact Hausdorff space. Unfortunately, describing such dual spaces also requires measure theory (see Theorem D.4.9) and we note are related to $\mathcal{M}_{\mathbb{K}}(X,\mathcal{A})$ where $\mathcal{A}$ is the Borel σ-algebra.

1.6 Canonical Embedding and Adjoints

We can take things a step farther with dual spaces. Indeed, the dual space of a normed linear space is a normed linear space and thus has a dual space. Such dual spaces are surprisingly useful and worthy of a name.

Definition 1.6.1. The *double dual space* (or *second dual space*) of a normed linear space $(\mathcal{X},\|\cdot\|)$ is the normed linear space $(\mathcal{X}^*)^*$ and is denoted $\mathcal{X}^{**}$.

Example 1.6.2. By Theorem 1.5.4, we see for all $p\in(1,\infty)$ that

$$(\ell_p(\mathbb{N}))^{**} = (\ell_q(\mathbb{N}))^* = \ell_p(\mathbb{N})$$

where $q\in(1,\infty)$ is such that $\frac{1}{p}+\frac{1}{q}=1$. Similarly, by Theorem 1.5.4, Theorem 1.5.7, and Theorem 1.5.8, we see that $(c_0)^{**}=\ell_\infty(\mathbb{N})$ and $c^{**}=\ell_\infty(\mathbb{N})$.

In Remark 1.5.9 we saw that every element of $\ell_1(\mathbb{N})$ defined an element of $(\ell_\infty(\mathbb{N}))^*$. However, we know that $(\ell_1(\mathbb{N}))^*=\ell_\infty(\mathbb{N})$ by Theorem 1.5.4. Combining these two facts, every element of $\ell_1(\mathbb{N})$ defines an element of $(\ell_\infty(\mathbb{N}))^*=(\ell_1(\mathbb{N}))^{**}$. Similarly, Example 1.6.2 shows that $(c_0)^{**}=\ell_\infty(\mathbb{N})$ and $c^{**}=\ell_\infty(\mathbb{N})$. Therefore, since $c_0\subseteq c\subseteq\ell_\infty(\mathbb{N})$, ever element of c_0 and c define elements of their second dual spaces. This is not a coincidence as the following shows (and note we will now often drop the vector notation as it will become cumbersome).

Theorem 1.6.3 (Canonical Embedding into Double Dual). *Let $(\mathcal{X},\|\cdot\|)$ be a normed linear space. For each $x\in\mathcal{X}$, define $\hat{x}:\mathcal{X}^*\to\mathbb{K}$ by*

$$\hat{x}(f)=f(x)$$

for all $f\in\mathcal{X}^*$. Then $\hat{x}\in\mathcal{X}^{**}$ and $\|\hat{x}\|\leq\|x\|$ for all $x\in\mathcal{X}$. Moreover, if $\mathcal{J}:\mathcal{X}\to\mathcal{X}^{**}$ is defined by

$$\mathcal{J}(x)=\hat{x}$$

for all $x \in \mathcal{X}$, then $\mathcal{J}$ is a contractive linear map. We call $\mathcal{J}$ the canonical embedding *of $\mathcal{X}$ into $\mathcal{X}^{**}$.*

Proof. It is elementary to see that if $x \in \mathcal{X}$, then $\hat{x}$ is a well-defined linear map on $\mathcal{X}^*$. Moreover, for all $f \in \mathcal{X}^*$, we see that

$$|\hat{x}(f)| = |f(x)| \leq \|f\| \, \|x\|$$

so $\hat{x}$ is bounded (and thus in $\mathcal{X}^{**}$) with $\|\hat{x}\| \leq \|x\|$. The fact that $\mathcal{J}$ is linear follows as for all $x, y \in \mathcal{X}$ and $\alpha \in \mathbb{K}$,

$$\widehat{\alpha x + y}(f) = f(\alpha x + y) = \alpha f(x) + f(y) = \alpha \hat{x}(f) + \hat{y}(f)$$

for all $f \in \mathcal{X}^*$ so $\widehat{\alpha x + y} = \alpha \hat{x} + \hat{y}$. ∎

Remark 1.6.4. Recall by Theorem 1.5.4 that if $p, q \in (1, \infty)$ are such that $\frac{1}{p} + \frac{1}{q} = 1$, then $(\ell_p(\mathbb{N}))^* = \ell_q(\mathbb{N})$ via the map $\Phi : \ell_q(\mathbb{N}) \to (\ell_p(\mathbb{N}))^*$ defined by

$$\Phi((x_n)_{n\geq 1})((y_n)_{n\geq 1}) = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ and $(y_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Moreover, by reversing the roles of p and q, we know that $(\ell_q(\mathbb{N}))^* = \ell_p(\mathbb{N})$ via the map $\Psi : \ell_p(\mathbb{N}) \to (\ell_q(\mathbb{N}))^*$ defined by

$$\Psi((y_n)_{n\geq 1})((x_n)_{n\geq 1}) = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ and $(y_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Therefore, if $x = (x_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ and $y = (y_n)_{n\geq 1} \in \ell_p(\mathbb{N})$, we see that

$$\hat{y}(\Phi(x)) = \sum_{n=1}^{\infty} x_n y_n = \Psi(y)(x)$$

so the canonical embedding is the identity map (upto identifying $(\ell_p(\mathbb{N}))^{**} = (\ell_q(\mathbb{N}))^* = \Psi(\ell_p(\mathbb{N}))$). Such spaces are some of the nicest in functional analysis and are worthy of a name.

Definition 1.6.5. A normed linear space $(\mathcal{X}, \|\cdot\|)$ is said to be *reflexive* if the canonical embedding is an isometric isomorphism of $\mathcal{X}$ onto $\mathcal{X}^{**}$.

Remark 1.6.6. In fact, using Theorem 1.5.4, Theorem 1.5.7, and Theorem 1.5.8, several canonical embeddings are isometric. Indeed recall that $c_0^* = \ell_1(\mathbb{N})$, $c^* = \ell_1(\mathbb{N})$, and $(\ell_1(\mathbb{N}))^* = \ell_\infty(\mathbb{N})$. If one considers the canonical embeddings $\mathcal{J} : c_0 \to (c_0)^{**} = \ell_\infty(\mathbb{N})$ and $\mathcal{J} : c \to c^{**} = \ell_\infty(\mathbb{N})$, by the same idea as used in Remark 1.6.4 we see that these maps are isometric.

This raises the question, "Is the canonical embedding always isometric?" Note for this to occur, we would need to show for any normed linear space $(\mathcal{X}, \|\cdot\|)$ and any $\vec{x} \in \mathcal{X}$ that

$$\|\vec{x}\| = \sup\{|\varphi(\vec{x})| \mid \varphi \in \mathcal{X}^*, \|\varphi\| \leq 1\}.$$

The main issue is, "How does one construct a *continuous* linear functional *of norm at most 1* that almost sends x to $\|\vec{x}\|$?"

For now we turn to another construction via dual spaces.

Theorem 1.6.7 (Adjoint of a Linear Map). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Define $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ by*

$$T^*(f) = f \circ T$$

for all $f \in Y^*$. Then T^* is a well-defined element of $\mathcal{B}(\mathcal{Y}^*, \mathcal{X}^*)$. Moreover $\|T^*\| \leq \|T\|$. The map T^* is called the adjoint *of T.*

Proof. To see that T^* is well-defined map note for all $f \in \mathcal{Y}^*$ that $T^*(f)$ is a bounded linear map with norm at most $\|T\| \|f\|$ being the composition of bounded linear maps by Corollary 1.4.5. As T^* is clearly linear and

$$\|T^*(f)\| \leq \|T\| \|f\|,$$

we see that $T^* \in \mathcal{B}(\mathcal{Y}^*, \mathcal{X}^*)$ with $\|T^*\| \leq \|T\|$ as desired. ∎

Remark 1.6.8. In a similar vein to Remark 1.6.6, we can ask, "Is $\|T^*\| = \|T\|$?" Indeed we know

$$\begin{aligned}\|T^*\| &= \sup\{\|f \circ T\| \mid f \in \mathcal{Y}^*, \|f\| \leq 1\}\\ &= \sup\{|f(T(x))| \mid f \in \mathcal{Y}^*, x \in \mathcal{X}, \|f\| \leq 1, \|x\|_{\mathcal{X}} \leq 1\}.\end{aligned}$$

If the answer to Remark 1.6.6 was in the affirmative, this would allow us to conclude

$$\begin{aligned}\|T^*\| &= \sup\{\|T(x))\|_{\mathcal{Y}} \mid x \in \mathcal{X}, \|x\|_{\mathcal{X}} \leq 1\}\\ &= \|T\|.\end{aligned}$$

This question will be addressed in a future Chapter.

For now, we finish with a quick study of some of the fundamental properties of the adjoint.

Proposition 1.6.9. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ be normed linear spaces, let $S, T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, and let $R \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Then*

(1) $(\alpha S + T)^* = \alpha S^* + T^*$ for all $\alpha \in \mathbb{K}$, and

(2) $(R \circ S)^* = S^* \circ R^*$.

Proof. For (1), notice for all $f \in \mathcal{Y}^*$ that

$$(\alpha S + T)^*(f) = f \circ (\alpha S + T) = \alpha(f \circ S) + (f \circ T) = \alpha S^*(f) + T^*(f)$$

as f is linear. Hence (1) follows.

For (2), notice for all $f \in \mathcal{Z}^*$ that

$$(R \circ S)^*(f) = f \circ (R \circ S) = (f \circ R) \circ S = S^*(f \circ R) = S^*(R^*(f)).$$

Hence (2) follows. ∎

Chapter 2

Banach Spaces

Not all normed linear spaces are created equal. In particular, many of the spaces examined in Chapter 1 have additional properties that are useful in functional analysis. One such property is based on convergence of sequences. As it is often difficult to verify whether or not a sequence converges as one must 'guess' the limit and show the sequence converges to the limit, the notion of a complete normed linear space where every Cauchy sequence converges bipasses these difficulties. These so called *Banach spaces* form some of the nicest objects in functional analysis and provide some of the deepest theorems.

In this chapter, we will study Banach spaces. After demonstrating that most of the spaces developed in Chapter 1 are Banach spaces, we will examine several properties of Banach spaces. Specifically we will prove some of the most important theorems for Banach spaces: the Baire Category Theorem (Theorem 2.3.1), the Open Mapping Theorem (Theorem 2.4.2), and the Principle of Uniform Boundedness (Theorem 2.5.3).

2.1 Banach Spaces

To be formal, we define the object of study in this chapter.

Definition 2.1.1. A *Banach space* is a complete normed linear space.

Of course $\mathbb{K}$ is complete by undergraduate real analysis. In addition, we have the following.

Corollary 2.1.2. *For every* $p \in [1, \infty]$ *and* $n \in \mathbb{N}$*,* $(\mathbb{K}^n, \|\cdot\|_p)$ *is a Banach space.*

Proof. If a sequence in $(\mathbb{K}^n, \|\cdot\|_p)$ is Cauchy, it is Cauchy in each entry. As $\mathbb{K}$ is complete, each entry of our sequence converges. As a sequence in $(\mathbb{K}^n, \|\cdot\|_p)$ converges if and only if it converges entrywise, the result follows. ∎

In infinite dimensions, things become more complicated as entrywise convergence does not imply convergence of the entire sequence. However, most of the normed linear spaces we studied above are Banach spaces. As we will see, the proofs of these results all follow the same pattern: take a Cauchy sequence, deduce it converges 'entrywise', stitch together the results, check the proposed limit is in the space under consideration, and check the sequences does indeed converge to the proposed limit.

Theorem 2.1.3. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces. If $\mathcal{Y}$ is a Banach space, then $(\mathcal{B}(\mathcal{X}, \mathcal{Y}), \|\cdot\|)$ is a Banach space.*

Proof. Let $(T_n)_{n\geq 1}$ be an arbitrary Cauchy sequence in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$. For each $\vec{x} \in \mathcal{X}$, notice

$$\|T_n(\vec{x}) - T_m(\vec{x})\|_{\mathcal{Y}} \leq \|T_n - T_m\| \, \|\vec{x}\|_{\mathcal{X}}$$

for all $n, m \in \mathbb{N}$. Hence it is elementary to see that $(T_n(\vec{x}))_{n\geq 1}$ is a Cauchy sequence in $\mathcal{Y}$ for all $\vec{x} \in \mathcal{X}$. Therefore, since $\mathcal{Y}$ is complete, for each $\vec{x} \in \mathcal{X}$ there exists an $T(\vec{x}) \in \mathcal{Y}$ such that $T(\vec{x}) = \lim_{n\to\infty} T_n(\vec{x})$.

To complete the proof, it suffices to verify three things: that $T : \mathcal{X} \to \mathcal{Y}$ is linear, that T is bounded, and that $\lim_{n\to\infty} \|T - T_n\| = 0$. To see that T is linear, notice for all $\vec{x}_1, \vec{x}_2 \in \mathcal{X}$ and $\alpha \in \mathbb{K}$ that

$$T(\alpha\vec{x}_1 + \vec{x}_2) = \lim_{n\to\infty} T_n(\alpha\vec{x}_1 + \vec{x}_2) = \lim_{n\to\infty} \alpha T_n(\vec{x}_1) + T_n(\vec{x}_2) = \alpha T(\vec{x}_1) + T(\vec{x}_2).$$

Hence T is linear.

To see that T is bounded, notice for all $\vec{x} \in \mathcal{X}$ with $\|\vec{x}\|_{\mathcal{X}} \leq 1$ and $m \in \mathbb{N}$ that

$$\|T(\vec{x}) - T_m(\vec{x})\|_{\mathcal{Y}} = \lim_{n\to\infty} \|T_n(\vec{x}) - T_m(\vec{x})\|_{\mathcal{Y}} \leq \limsup_{n\to\infty} \|T_n - T_m\|$$

Since $(T_n)_{n\geq 1}$ is Cauchy we know that $(T_n)_{n\geq 1}$ is bounded and therefore $\limsup_{n\to\infty} \|T_n - T_m\|$ is finite. In particular, we obtain that there exists a constant M such that

$$\|T(\vec{x})\|_{\mathcal{Y}} \leq \|T_1(\vec{x})\|_{\mathcal{Y}} + M \leq \|T_1\| + M$$

for all $\vec{x} \in \mathcal{X}$ with $\|\vec{x}\|_{\mathcal{X}} \leq 1$. Hence T is bounded with $\|T\| \leq \|T_1\| + M$.

To see that $\lim_{n\to\infty} \|T - T_n\| = 0$, let $\epsilon > 0$ be arbitrary. Since $(T_n)_{n\geq 1}$ is Cauchy, there exists an $N \in \mathbb{N}$ such that $\|T_m - T_j\| \leq \epsilon$ for all $m, j \geq N$. Hence if $j \geq N$, the above implies $\|T(\vec{x}) - T_j(\vec{x})\| \leq \epsilon$ for all $\vec{x} \in \mathcal{X}$ with $\|\vec{x}\|_{\mathcal{X}} \leq 1$ and thus

$$\|T - T_j\| \leq \epsilon$$

for all $n \geq N$. Therefore, as $\epsilon > 0$ was arbitrary, we obtain that $(T_n)_{n\geq 1}$ converges to T in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$. Hence, as $(T_n)_{n\geq 1}$ was arbitrary, $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is complete. ∎

Corollary 2.1.4. *The dual space of any normed linear space is a Banach space.*

Proof. As $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{K})$ and $\mathbb{K}$ is complete, $\mathcal{X}^*$ is complete by Theorem 2.1.3. ∎

To see other normed linear spaces consisting of functions are Banach spaces, we recall a stronger notion of convergence.

Definition 2.1.5. Let $(X, \mathcal{T})$ be a topological space and let (Y, d) be a metric space. For each $n \in \mathbb{N}$ let $f_n : X \to Y$. Given $f : X \to Y$, it is said that the sequence $(f_n)_{n\geq 1}$ *converges uniformly* to f if $(f_n)_{n\geq 1}$ converges to f with respect to the uniform metric (provided it makes sense); that is, for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $d_Y(f(x), f_n(x)) < \epsilon$ for all $n \geq N$ and for all $x \in \mathcal{X}$.

The following result is a staple of a first course in analysis (at least for continuous functions on and to $\mathbb{R}$) whose proof is a standard important argument.

Theorem 2.1.6. *Let $(X, \mathcal{T})$ be a topological space and let (Y, d) be a metric space and let $f : X \to Y$. If $(f_n)_{n\geq 1}$ is a sequence of continuous functions from X to Y that converge to f uniformly, then f is continuous.*

Proof. To see that f is continuous, let $x_0 \in X$ be arbitrary. To see that f is continuous at x_0 let $\epsilon > 0$ be arbitrary. Since $(f_n)_{n\geq 1}$ converges to f uniformly, there exists an $N \in \mathbb{N}$ such that $d_Y(f(x), f_N(x)) < \frac{\epsilon}{3}$ for all $x \in \mathcal{X}$. Since f_N is continuous at x_0, there exists an open set $U \in \mathcal{T}$ such that $x_0 \in U$ and if $x \in U$ then $d_Y(f_N(x), f_N(x_0)) < \frac{\epsilon}{3}$. Hence if $x \in X$ and $x \in U$, then, by the triangle inequality,

$$\begin{aligned} d_Y(f(x), f(x_0)) &\leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(x_0)) + d_Y(f_N(x_0), f(x_0)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Hence, as $\epsilon > 0$ was arbitrary, f is continuous at x_0. Thus, as x_0 was arbitrary, f is continuous on X. ∎

Using the above, we obtain the following result for metric spaces.

Theorem 2.1.7. *Let $(X, \mathcal{T})$ be a topological space and let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a Banach space. Then $(C_b(X, \mathcal{Y}), \|\cdot\|_\infty)$ is a Banach space.*

Proof. Let $(f_n)_{n\geq 1}$ be an arbitrary Cauchy sequence in $C_b(X, \mathcal{Y})$. For each $x \in X$, notice

$$\|f_n(x) - f_m(x)\|_{\mathcal{Y}} \leq \|f_n - f_m\|_\infty$$

for all $n, m \in \mathbb{N}$. Hence it is elementary to see that $(f_n(x))_{n\geq 1}$ is a Cauchy sequence in $\mathcal{Y}$ for all $x \in X$. Therefore, since $\mathcal{Y}$ is complete, for each

$x \in X$ there exists an $f(x) \in \mathcal{Y}$ such that $f(x) = \lim_{n\to\infty} f_n(x)$. Clearly the function $x \mapsto f(x)$ defines a function $f : X \to \mathcal{Y}$.

To complete the proof, it suffices to verify three things: that f is bounded, that $f : X \to \mathcal{Y}$ is continuous, and that $\lim_{n\to\infty} \|f - f_n\|_\infty = 0$. For the first, notice for all $x \in X$ and $m \in \mathbb{N}$ that

$$\|f(x) - f_m(x)\|_{\mathcal{Y}} = \lim_{n\to\infty} \|f_n(x) - f_m(x)\|_{\mathcal{Y}} \leq \limsup_{n\to\infty} \|f_n - f_m\|_\infty .$$

Since $(f_n)_{n\geq 1}$ is Cauchy we know that $(f_n)_{n\geq 1}$ is bounded and therefore $\limsup_{n\to\infty} \|f_n - f_m\|_\infty$ is finite. Therefore, by taking the supremum over all $x \in X$, we obtain that

$$\sup\{\|f(x) - f_m(x)\|_{\mathcal{Y}} \mid x \in X\} \leq \limsup_{n\to\infty} \|f_n - f_m\|_\infty$$

for all $m \in \mathbb{N}$. Thus, by taking $m = 1$ and using the fact that f_1 is bounded, we easily see that f is bounded.

To see that f is continuous, we will show that $(f_n)_{n\geq 1}$ converges uniformly to f using the above. Indeed let $\epsilon > 0$ be arbitrary. Since $(f_n)_{n\geq 1}$ is Cauchy in $C_b(X, \mathcal{Y})$, there exists an $N \in \mathbb{N}$ such that $\|f_j - f_m\|_\infty \leq \epsilon$ for all $m, j \geq N$. Hence if $m \geq N$, the above implies

$$\sup\{\|f(x) - f_m(x)\|_{\mathcal{Y}} \mid x \in X\} < \epsilon.$$

Thus $(f_n)_{n\geq 1}$ converges to f uniformly on X. Hence f is continuous by Theorem 2.1.6.

As the above shows that $\lim_{m\to\infty} \|f - f_m\|_\infty = 0$, $(f_n)_{n\geq 1}$ converges to f in $C_b(X, \mathcal{Y})$. Thus, as $(f_n)_{n\geq 1}$ was an arbitrary Cauchy sequence, $C_b(X, \mathcal{Y})$ is complete and thus a Banach space. ■

Corollary 2.1.8. *Let $(X, \mathcal{T})$ be a topological space. Then $(C_b(X, \mathbb{K}), \|\cdot\|_\infty)$ is a Banach space.*

Corollary 2.1.9. *Let $(X, \mathcal{T})$ be a compact Hausdorff space and let $(\mathcal{Y}, \|\cdot\|)$ be a Banach space. Then $(C(X, \mathcal{Y}), \|\cdot\|_\infty)$ is a Banach space.*

To discuss $C_0(X, \mathcal{Y})$ when X is a locally compact Hausdorff topological space and $\mathcal{Y}$ is a Banach space, it is first helpful to make a remark.

Remark 2.1.10. Given a Banach space $(\mathcal{X}, \|\cdot\|)$, if $\mathcal{Y}$ is a vector subspace of $\mathcal{X}$, then $(\mathcal{Y}, \|\cdot\| |_{\mathcal{Y}})$ is a Banach space if and only if $\mathcal{Y}$ is closed. Indeed if $\mathcal{Y}$ is closed, then it is complete (closed subsets of complete metric spaces are complete) and if $\mathcal{Y}$ is not closed there is a sequence in $\mathcal{Y}$ that converges in $\mathcal{X}$ (and thus is Cauchy in $\mathcal{X}$ and $\mathcal{Y}$) that is does not converges in $\mathcal{Y}$ thereby showing $\mathcal{Y}$ is not complete.

For this reason, we often restrict our attention to *closed subspaces* in functional analysis. We warn the reader that in main texts in functional analysis that a 'subspace' means a 'closed subspace' for just this reason.

Corollary 2.1.11. *Let $(X, \mathcal{T})$ be a locally compact Hausdorff space and let $(\mathcal{Y}, \|\cdot\|)$ be a Banach space. Then $(C_0(X, \mathcal{Y}), \|\cdot\|_\infty)$ is a Banach space. Hence $c_0 = C_0(\mathbb{N}, \mathbb{K})$ is a Banach space.*

Proof. Recall $C_0(X, \mathcal{Y})$ is a closed subspace of $C_b(X, \mathcal{Y})$ and thus a Banach space. ■

Although we know $\ell_p(\mathbb{N})$ are dual spaces by Theorem 1.5.4 (and Theorem 1.5.7 for $\ell_1(\mathbb{N})$) and thus Banach spaces, we note the spaces from Proposition 1.3.1 are Banach spaces.

Theorem 2.1.12. *For each $n \in \mathbb{N}$, let $(\mathcal{X}_n, \|\cdot\|_n)$ be a Banch space over $\mathbb{K}$. Let*

$$\mathcal{X} = \{(\vec{x}_n)_{n\geq 1} \mid \vec{x}_n \in \mathcal{X}_n \text{ for all } n \in \mathbb{N}\}.$$

Then $\bigoplus_{n\in\mathbb{N}}^p \mathcal{X}_n$ is a Banach space for all $p \in [1, \infty]$ and $c_0(\mathcal{X})$ is a Banach space.

Proof. Fix $p \in [1, \infty]$. For notational purposes, let $\mathcal{Y}_p = \bigoplus_{n\in\mathbb{N}}^p \mathcal{X}_n$. To see that $\mathcal{Y}_p$ is complete, let $(\vec{y}_k)_{k\geq 1}$ be an arbitrary Cauchy sequence in $\mathcal{Y}_p$. For each $k \in \mathbb{N}$, write $\vec{y}_k = (x_{k,n})_{n\geq 1}$ where $x_{k,n} \in \mathcal{X}_n$ for all $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Since for all $m, k, n \in \mathbb{N}$,

$$\|x_{k,n} - x_{m,n}\|_n \leq \|\vec{y}_k - \vec{y}_m\|_p\,,$$

we see that for each $n \in \mathbb{N}$ the sequence $(x_{k,n})_{k\geq 1}$ is Cauchy in $\mathcal{X}_n$. Therefore, as $\mathcal{X}_n$ is complete, $x_n = \lim_{k\to\infty} x_{k,n}$ exists in $\mathcal{X}_n$ for each $n \in \mathbb{N}$.

Let $\vec{y} = (x_n)_{n\geq 1}$. To complete the proof, it suffices to verify two things: that $\vec{y} \in \mathcal{Y}_p$, and that $\lim_{k\to\infty} \|\vec{y} - \vec{y}_k\|_p = 0$. We will only discuss the case $p \neq \infty$ and the case $p = \infty$ is similar and uses the same sorts of arguments as the proof of Theorem 2.1.7.

For $p \neq \infty$ notice for all $N, m \in \mathbb{N}$ that

$$\left(\sum_{n=1}^{N} |x_n - x_{m,n}|^p\right)^{\frac{1}{p}} = \lim_{k\to\infty} \left(\sum_{n=1}^{N} |x_{k,n} - x_{m,n}|^p\right)^{\frac{1}{p}} \leq \limsup_{k\to\infty} \|\vec{x}_k - \vec{x}_m\|_p\,.$$

Since $(\vec{x}_n)_{n\geq 1}$ is Cauchy it follows that $(\vec{x}_n)_{n\geq 1}$ is bounded and therefore $\limsup_{k\to\infty} \|\vec{x}_k - \vec{x}_m\|_p$ is finite. Therefore, by taking the limit as N tends to infinity, we obtain that

$$\left(\sum_{n=1}^{\infty} |x_n - x_{m,n}|^p\right)^{\frac{1}{p}} \leq \limsup_{k\to\infty} \|\vec{x}_k - \vec{x}_m\|_p\,.$$

By setting $m = 1$, we see that $\vec{z} = (x_n - x_{1,n})_{n\geq 1} \in \mathcal{Y}_p$. Therefore, as $\vec{y} = \vec{z} + \vec{y}_1$, we obtain that $\vec{y} \in \mathcal{Y}_p$.

To see that $\lim_{k\to\infty} \|\vec{y} - \vec{y}_k\|_p = 0$, let $\epsilon > 0$ be arbitrary. Note the above shows for all $m \in \mathbb{N}$ that

$$\|\vec{y} - \vec{y}_m\|_p \leq \limsup_{k\to\infty} \|\vec{x}_k - \vec{x}_m\|_p .$$

Since $(\vec{x}_n)_{n\geq 1}$ is Cauchy there exists an $N \in \mathbb{N}$ such that $\|\vec{y}_k - \vec{y}_m\|_p \leq \epsilon$ for all $m, k \geq N$. Hence if $m \geq N$, the above implies $\|\vec{y} - \vec{x}_m\|_p \leq \epsilon$. Therefore, as $\epsilon > 0$ was arbitrary, we obtain that $\lim_{k\to\infty} \|\vec{y} - \vec{y}_k\|_p = 0$. Hence $(\vec{y}_k)_{k\geq 1}$ converges in $\mathcal{Y}_p$ so, as $(\vec{y}_k)_{k\geq 1}$ was arbitrary, $\mathcal{Y}_p$ is complete and thus a Banach space.

As $c_0(\mathcal{X})$ is a closed vector subspace of $\bigoplus_{n\in\mathbb{N}}^{\infty} \mathcal{X}_n$, it follows that $c_0(\mathcal{X})$ is a Banach space. ∎

Corollary 2.1.13. *For all $p \in [1, \infty]$, $\ell_p(\mathbb{N})$ is a Banach space.*

Corollary 2.1.14. *The space c is a Banach space.*

Proof. As c is a closed subspace of $\ell_\infty(\mathbb{N})$, it follows that c is a Banach space. ∎

We also note that the measure-theoretic spaces mentioned above are indeed Banach spaces. For the proof that $L_p(X, \mu)$ is a Banach space, see Theorem D.2.1 and Theorem D.2.4. The proof that $M_{\mathbb{C}}(X, \mathcal{A})$ and $\mathcal{M}_{\mathbb{R}}(X, \mathcal{A})$ are Banach spaces follows from similar arguments to those used above. Moreover, these spaces provide a nice example of a normed linear space that is not a Banach space.

Example 2.1.15. Let $p \in [1, \infty)$ and consider $C[0, 1]$ as a vector subspace of $L_p([0, 1], \lambda)$ where λ is the Lebesgue measure. We claim that $C[0, 1]$ is not closed and thus not a Banach space with respect to $\|\cdot\|_p$. To see this, for each $n \in \mathbb{N}$ let $f_n \in C[0, 1]$ be defined by

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

We claim that $(f_n)_{n\geq 1}$ is a Cauchy sequence that does not converge. To see that $(f_n)_{n\geq 1}$ is Cauchy, notice if $n, m \in \mathbb{N}$ with $n > m$ then

$$\begin{aligned} \|f_n - f_m\|_p &= \left(\int_0^1 |f_n(x) - f_m(x)|^p \, dx\right)^{\frac{1}{p}} \\ &= \left(\int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{m}} |f_n(x) - f_m(x)|^p \, dx\right)^{\frac{1}{p}} \\ &\leq \left(\int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{m}} 1 \, dx\right)^{\frac{1}{p}} \leq \frac{1}{m^{\frac{1}{p}}} \end{aligned}$$

as $|f_n(x) - f_m(x)| \leq 1$ for all $x \in [0,1]$. Therefore, as $\lim_{m\to\infty} \frac{1}{m^{\frac{1}{p}}} = 0$, we obtain that $(f_n)_{n\geq 1}$ is Cauchy in $(C[0,1], \|\cdot\|_p)$.

To see that $(f_n)_{n\geq 1}$ does not have a limit in $(C[0,1], \|\cdot\|_p)$, suppose to the contrary that $f \in C[0,1]$ is a limit of $(f_n)_{n\geq 1}$. Then for all $a, b \in [0,1]$ with $a < b$, we have that

$$\begin{aligned}\limsup_{n\to\infty} \left(\int_a^b |f_n(x) - f(x)|^p \, dx\right)^{\frac{1}{p}} &\leq \limsup_{n\to\infty} \left(\int_0^1 |f_n(x) - f(x)|^p \, dx\right)^{\frac{1}{p}} \\ &= \limsup_{n\to\infty} \|f_n - f\|_p = 0\end{aligned}$$

as the integral of a positive function is positive and the function $x \mapsto x^{\frac{1}{p}}$ is increasing on $[0,\infty)$. Thus for each $a, b \in \left[0, \frac{1}{2}\right]$ with $a < b$ we obtain that

$$0 = \limsup_{n\to\infty} \left(\int_a^b |f_n(x) - f(x)|^p \, dx\right)^{\frac{1}{p}} = \left(\int_a^b |1 - f(x)|^p \, dx\right)^{\frac{1}{p}}$$

However, as f is continuous on $[0,1]$, this implies that $f(x) = 1$ for all $x \in \left[0, \frac{1}{2}\right]$. Similarly, if $\frac{1}{2} < a < b \leq 1$, we obtain by selecting n large enough so that $\frac{1}{2} + \frac{1}{n} < a$ that

$$0 = \limsup_{n\to\infty} \left(\int_a^b |f_n(x)|^p \, dx\right)^{\frac{1}{p}} = \left(\int_a^b |f(x)|^p \, dx\right)^{\frac{1}{p}}.$$

Hence, the same arguments imply that $f(x) = 0$ for all $x \in \left(\frac{1}{2}, 1\right]$. Thus, as f is continuous at $\frac{1}{2}$, we have obtained that $0 = f\left(\frac{1}{2}\right) = 1$ which is a contradiction. Thus $(f_n)_{n\geq 1}$ does not have a limit in $(C[0,1], \|\cdot\|_p)$ so $(C[0,1], \|\cdot\|_p)$ is not complete.

Of course $C[0,1]$ is dense in $L_p([0,1], \lambda)$ for $p \neq \infty$ by Theorem D.3.2 so by adding in 'a few more functions', we can turn $(C[0,1], \|\cdot\|_p)$ into a Banach space. This is the notion of a *completion* of a normed linear space. Such completions always exist as Appendix C shows.

2.2 Banach Space Properties

One example from Chapter 1 that has not been studied in the Banach space setting are the quotient spaces from Theorem 1.3.3. In particular, how does $\mathcal{V}$ being a Banach space or not relate to whether $\mathcal{V}/\mathcal{W}$ is a Banach space or not for a closed subspace $\mathcal{W}$ of $\mathcal{V}$? One method for answering this question is to look at a useful property of a Banach space which is motivated by connecting Cauchy sequences and sums as one would in undergraduate real analysis.

Definition 2.2.1. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. A series $\sum_{n=1}^\infty \vec{x}_n$ is said to be *summable* if the sequence of partial sums $(s_n)_{n\geq 1}$ converges (where $s_n = \sum_{k=1}^n \vec{x}_k$).

A series $\sum_{n=1}^\infty \vec{x}_n$ is said to be *absolutely summable* if $\sum_{n=1}^\infty \|\vec{x}_n\| < \infty$.

Theorem 2.2.2. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. Then $\mathcal{X}$ is a Banach space if and only if every absolutely summable series is summable.*

Proof. Suppose $\mathcal{X}$ is a Banach space. Let $\sum_{n=1}^\infty \vec{x}_n$ be an arbitrary absolutely summable series in $(\mathcal{X}, \|\cdot\|)$. To see that $\sum_{n=1}^\infty \vec{x}_n$ is summable, let $\epsilon > 0$ be arbitrary. Since $\sum_{n=1}^\infty \|\vec{x}_n\| < \infty$, there exists an $N \in \mathbb{N}$ such that $\sum_{n=N}^\infty \|\vec{x}_n\| < \epsilon$. Therefore, if $k, m \geq N$ and, without loss of generality, $m \geq k$, then

$$\begin{aligned}
\|s_m - s_k\| &= \left\| \sum_{n=1}^m \vec{x}_n - \sum_{n=1}^k \vec{x}_n \right\| \\
&= \left\| \sum_{n=k+1}^m \vec{x}_n \right\| \\
&\leq \sum_{n=k+1}^m \|\vec{x}_n\| \\
&\leq \sum_{n=N}^\infty \|\vec{x}_n\| < \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, the sequence of partial sums $(s_n)_{n\geq 1}$ is Cauchy. Hence $(s_n)_{n\geq 1}$ converges as $\mathcal{X}$ is complete. Thus, as $\sum_{n=1}^\infty \vec{x}_n$ was arbitrary, every absolutely summable series in $\mathcal{X}$ is summable.

For the converse, suppose every absolutely summable sequence in $\mathcal{X}$ is summable. To see that $\mathcal{X}$ is complete, let $(\vec{x}_n)_{n\geq 1}$ be an arbitrary Cauchy sequence. Since $(\vec{x}_n)_{n\geq 1}$ is Cauchy, there exists an $n_1 \in \mathbb{N}$ such that $\|\vec{x}_m - \vec{x}_j\| < \frac{1}{2}$ for all $m, j \geq n_1$. Similarly, since $(\vec{x}_n)_{n\geq 1}$ is Cauchy, there exists an $n_2 \in \mathbb{N}$ such that $n_2 > n_1$ and $\|\vec{x}_m - \vec{x}_j\| < \frac{1}{2^2}$ for all $m, j \geq n_2$. By repeating the above process, for each $k \in \mathbb{N}$ there exists an $n_k \in \mathbb{N}$ such that $n_k < n_{k+1}$ for all k and $\|\vec{x}_m - \vec{x}_j\| < \frac{1}{2^k}$ for all $m, j \geq n_k$.

For each $k \in \mathbb{N}$ let $\vec{y}_k = \vec{x}_{n_{k+1}} - \vec{x}_{n_k}$. Thus we have that

$$\sum_{k=1}^\infty \|\vec{y}_k\| \leq \sum_{k=1}^\infty \frac{1}{2^k} < \infty$$

so $\sum_{k=1}^\infty \vec{y}_k$ is an absolutely summable series in $\mathcal{X}$. Therefore, by the assumptions on $\mathcal{X}$, $\sum_{k=1}^\infty \vec{y}_k$ is summable in $\mathcal{X}$.

Let $\vec{x} = \vec{x}_{n_1} + \sum_{k=1}^\infty \vec{y}_k$. We claim that $(\vec{x}_{n_k})_{k\geq 1}$ converges to $\vec{x}$. To see this, let $\epsilon > 0$ be arbitrary. Then there exists a $M \in \mathbb{N}$ such that if $m \geq M$

then

$$\left\| \sum_{k=1}^{\infty} \vec{y}_k - \sum_{k=1}^{m} \vec{y}_k \right\| < \epsilon.$$

Therefore, if $m \geq M$,

$$\begin{aligned} \|\vec{x} - \vec{x}_{n_{m+1}}\| &\leq \left\| \sum_{k=1}^{\infty} \vec{y}_k - \sum_{k=1}^{m} \vec{y}_k \right\| + \left\| \vec{x}_{n_1} - \vec{x}_{n_{m+1}} + \sum_{k=1}^{m} \vec{y}_k \right\| \\ &< \epsilon + \left\| \vec{x}_{n_1} - \vec{x}_{n_{m+1}} + \sum_{k=1}^{m} \vec{x}_{n_{k+1}} - \vec{x}_{n_k} \right\| \\ &= \epsilon. \end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, $(\vec{x}_{n_k})_{k\geq 1}$ converges to $\vec{x}$. Thus, as $(\vec{x}_n)_{n\geq 1}$ is Cauchy, $(\vec{x}_n)_{n\geq 1}$ converges to $\vec{x}$. Therefore, as $(\vec{x}_n)_{n\geq 1}$ was an arbitrary Cauchy sequence, $\mathcal{X}$ is complete. ∎

As an immediate corollary, we obtain the following result pertaining to convergence of series of continuous functions.

Corollary 2.2.3 (Weierstrass M-Test). *Let $(X, \mathcal{T})$ be a topological space and let $(f_n)_{n\geq 1}$ be a sequence of functions from $C_b(X, \mathbb{R})$. Suppose there exists an $M \in \mathbb{R}$ such that $\sum_{n=1}^{\infty} \|f_n\|_\infty < M$. Then $\sum_{n=1}^{\infty} f_n$ converges uniformly on X to a continuous function.*

Proof. As $(f_n)_{n\geq 1}$ is absolutely summable in the Banach space $C_b(X, \mathbb{R})$, $\sum_{n=1}^{\infty} f_n$ converges in $C_b(X, \mathbb{R})$. ∎

Using the characterization of Banach spaces involving absolutely summable series, we can describe when quotients of normed linear spaces will be Banach spaces.

Theorem 2.2.4. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Then $\mathcal{V}$ is a Banach space if and only if $\mathcal{W}$ and $\mathcal{V}/\mathcal{W}$ are Banach spaces.*

Proof. Suppose $\mathcal{V}$ is a Banach space. By Remark 2.1.10 that $\mathcal{W}$ is a Banach space. To see that $\mathcal{V}/\mathcal{W}$ is a Banach space, we will appeal to Theorem 2.2.2.

Let $\sum_{n=1}^{\infty} \vec{v}_n + \mathcal{W}$ be an absolutely summable series in $\mathcal{V}/\mathcal{W}$. For each $n \in \mathbb{N}$, by the definition of the quotient norm we can find a $\vec{w}_n \in \mathcal{W}$ such that

$$\|\vec{v}_n + \vec{w}_n\| \leq \frac{1}{2^n} + \|\vec{v}_n + \mathcal{W}\|.$$

Therefore, since $\sum_{n=1}^{\infty} \vec{v}_n + \mathcal{W}$ be an absolutely summable series in $\mathcal{V}/\mathcal{W}$ we see that $\sum_{n=1}^{\infty} \vec{v}_n + \vec{w}_n$ is absolutely summable series in $\mathcal{V}$. Hence $\sum_{n=1}^{\infty} \vec{v}_n + \vec{w}_n$ converges in $\mathcal{V}$ to some vector $\vec{v} \in \mathcal{V}$.

Since the quotient map $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is continuous by Example 1.4.12, we have that

$$\begin{aligned} q(\vec{v}) &= \lim_{N\to\infty} q\left(\sum_{n=1}^{N} \vec{v}_n + \vec{w}_n\right) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} q\left(\vec{v}_n + \vec{w}_n\right) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \vec{v}_n + \mathcal{W}. \end{aligned}$$

Hence $\sum_{n=1}^{\infty} \vec{v}_n + \mathcal{W}$ converges to $q(\vec{v})$. Therefore, since $\sum_{n=1}^{\infty} \vec{v}_n + \mathcal{W}$ was arbitrary, $\mathcal{V}/\mathcal{W}$ is a Banach space by Theorem 2.2.2

Conversely, suppose that $\mathcal{W}$ and $\mathcal{V}/\mathcal{W}$ are complete. To see that $\mathcal{V}$ is complete, let $(\vec{v}_n)_{n\geq 1}$ be a Cauchy sequence in $\mathcal{V}$. By the definition of the quotient norm, we have that

$$\|(\vec{v}_n + \mathcal{W}) - (\vec{v}_m + \mathcal{W})\| = \|(\vec{v}_n - \vec{v}_m) + \mathcal{W}\| \leq \|\vec{v}_n - \vec{v}_m\|$$

for all $n, m \in \mathbb{N}$. Therefore, since $(\vec{v}_n)_{n\geq 1}$ be a Cauchy sequence in $\mathcal{V}$ we obtain that $(\vec{v}_n + \mathcal{W})_{n\geq 1}$ is a Cauchy sequence in $\mathcal{V}/\mathcal{W}$.

Since $\mathcal{V}/\mathcal{W}$ is complete, $(\vec{v}_n + \mathcal{W})_{n\geq 1}$ converges to some vector $\vec{v} + \mathcal{W} \in \mathcal{V}/\mathcal{W}$. Unfortunately, $\vec{v}$ is not the vector we want as we still need to correct it by the appropriate vector from $\mathcal{W}$. This is where the completeness of $\mathcal{W}$ will come into play.

By the definition of the quotient norm, for each $n \in \mathbb{N}$ there exists a $\vec{w}_n \in \mathcal{W}$ such that

$$\|\vec{v} - \vec{v}_n + \vec{w}_n\| \leq \frac{1}{2^n} + \|(\vec{v} - \vec{v}_n) + \mathcal{W}\|.$$

Since $\lim_{n\to\infty} \|(\vec{v} - \vec{v}_n) + \mathcal{W}\| = 0$, we see that $(\vec{v}_n - \vec{w}_n)_{n\geq 1}$ converges to $\vec{v}$ in $\mathcal{V}$ and thus is Cauchy. As $(\vec{v}_n)_{n\geq 1}$ is Cauchy and the difference of two Cauchy sequences is easily seen to be Cauchy by the triangle inequality, we obtain that $(\vec{w}_n)_{n\geq 1}$ is Cauchy in $\mathcal{V}$. However, as $\vec{w}_n \in \mathcal{W}$ for all $n \in \mathbb{N}$, we see that $(\vec{w}_n)_{n\geq 1}$ is Cauchy in $\mathcal{W}$ and thus converges to some vector $\vec{w} \in \mathcal{W}$ as $\mathcal{W}$ is complete. Since

$$\vec{v} + \vec{w} = \lim_{n\to\infty} \vec{v}_n - \vec{w}_n + \lim_{n\to\infty} \vec{w}_n = \lim_{n\to\infty} \vec{v}_n,$$

we have that $(\vec{v}_n)_{n\geq 1}$ converges to $\vec{v} + \vec{w}$ in $\mathcal{V}$. Therefore, as $(\vec{v}_n)_{n\geq 1}$ was arbitrary, $\mathcal{V}$ is complete. ∎

2.3 The Baire Category Theorem

In this section, we will prove one of the most surprisingly useful theorems pertaining to complete metric spaces (and thus Banach spaces). Although its uses will not be apparent from the statement of the theorem, we will see in the subsequent sections some of its applications.

Theorem 2.3.1 (Baire's Category Theorem). *Let (X, d) be a complete metric space. Suppose $(U_n)_{n\geq 1}$ is a sequence of open dense subsets of X. Then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

Proof. To see that $\bigcap_{n=1}^{\infty} U_n$ is dense in X, let $x \in X$ and $\epsilon > 0$ be arbitrary. We must show that there exists an element of $\bigcap_{n=1}^{\infty} U_n$ within ϵ of x. To do this, it is first useful to note that if $y \in X$ and $r > 0$ then for any $0 < r' < r$ we have that

$$B[y, r'] \subseteq B(y, r)$$

where $B[y, r']$ is the closed ball centred at y of radius r'.

Let $r_1 = \frac{1}{2}\epsilon$. Since U_1 is dense in X, there exists an element $x_1 \in U_1$ such that $d(x_1, x) < r_1$. Since U_1 is open, by the above comment there exists an $0 < r_2 < \frac{1}{4}\epsilon$ such that $B[x_1, r_2] \subseteq U_1$ (i.e. choose an open ball around x_1 contained in U_1 and then decrease the radius of the ball).

Since U_2 is dense in X, there exists an element $x_2 \in U_2$ such that $d(x_2, x_1) < r_2$. Hence $x_2 \in B(x_1, r_2)$ so $x_2 \in U_2 \cap B(x_1, r_2)$. Hence, since $U_2 \cap B(x_1, r_2)$ is open, there exists an $0 < r_2 < \frac{1}{2^3}\epsilon$ such that $B[x_2, r_3] \subseteq U_2 \cap B(x_1, r_2)$.

By recursion, for each $n \in \mathbb{N}$ there exists an $x_n \in U_n \cap B(x_{n-1}, r_n)$ and an $0 < r_{n+1} < \frac{1}{2^{n+1}}\epsilon$ such that $d(x_n, x_{n-1}) < r_n$ and $B[x_n, r_{n+1}] \subseteq U_n \cap B(x_{n-1}, r_n)$.

Notice for all $n, m \in \mathbb{N}$ with $n \geq m$ that

$$d(x_n, x_m) \leq \sum_{k=m}^{n-1} d(x_{k+1}, x_k) \leq \sum_{k=m}^{n-1} r_{k+1} \leq \frac{1}{2^m}\epsilon.$$

Therefore, as $\lim_{m\to\infty} \frac{1}{2^m}\epsilon = 0$, we obtain that $(x_n)_{n\geq 1}$ is Cauchy in (X, d) and thus converges to some element $y \in X$ as (X, d) is complete.

For each $m \in \mathbb{N}$, let $F_m = B[x_m, r_{m+1}]$. Since $F_{m+1} \subseteq F_m$ for all $m \in \mathbb{N}$, we see that $(x_n)_{n\geq m}$ is a sequence in F_m for all $m \in \mathbb{N}$. Therefore, since F_m is closed, we obtain that $y \in F_m$ for all $m \in \mathbb{N}$. Hence, as $F_m \subseteq U_m$ for all $m \in \mathbb{N}$, we obtain that $y \in \bigcap_{n=1}^{\infty} U_n$.

To see that $d(x, y) < \epsilon$, we note that $y \in F_1 = B[a_1, r_2]$ so $d(y, a_1) \leq r_2$. Hence

$$d(x, y) \leq d(x, a_1) + d(a_1, y) \leq r_1 + r_2 < \epsilon$$

by the triangle inequality. Hence the result follows. ∎

Often in functional analysis one desires to work with closed sets as convergent sequences have limits inside the set. As a set U is open and dense if and only if its complement is closed and nowhere dense (which is equivalent to having empty interior), we obtain the following implication of the Baire Category Theorem (Theorem 2.3.1).

Corollary 2.3.2. *Let (X, d) be a complete metric space. If $(F_n)_{n\geq 1}$ is a sequence of closed sets such that $X = \bigcup_{n=1}^\infty F_n$, then there exists an $N \in \mathbb{N}$ such that* $\mathrm{int}(F_N) \neq \emptyset$.

Proof. Suppose to the contrary that $\mathrm{int}(F_n) = \emptyset$ for all $n \in \mathbb{N}$. Let $U_n = X \setminus F_n$ for all $n \in \mathbb{N}$. Since $X = \bigcup_{n=1}^\infty F_n$, it follows that $\bigcap_{n=1}^\infty U_n = \emptyset$. However, since each F_n is closed with empty interior, each U_n is an open dense subset of (X, d). As the Baire Category Theorem (Theorem 2.3.1) implies $\bigcap_{n=1}^\infty U_n$ is dense and thus non-empty, we have a contradiction. ∎

There are numerous uses of the Baire Category Theorem. We conclude this section mentioning one interesting result about continuous functions on $\mathbb{R}$. First we require some notation.

Definition 2.3.3. Let (X, d) be a metric space. A subset $A \subseteq X$ is said to be G_δ if there exists a collection of open sets $\{U_n\}_{n=1}^\infty$ such that $A = \bigcap_{n=1}^\infty U_n$.

Similarly, a subset $B \subseteq \mathcal{X}$ is said to be F_σ if there exists a collection of closed sets $\{F_n\}_{n=1}^\infty$ such that $A = \bigcup_{n=1}^\infty F_n$.

Remark 2.3.4. It is not difficult to see using De Morgan's Laws that A is G_δ if and only if A^c is F_σ.

Example 2.3.5. Every closed subset of a metric space is G_δ. To see this, suppose F be a closed subset of a metric space (X, d). If $F = \emptyset$ then, as $\emptyset$ is open and as $\bigcap_{n=1}^\infty \emptyset = \emptyset$, we obtain that F is G_δ.

Otherwise, suppose F is not empty. For each $n \in \mathbb{N}$, let

$$U_n = \bigcup_{x\in F} B\left(x, \frac{1}{n}\right).$$

Clearly each U_n is an open subset such that $F \subseteq U_n$. Hence

$$F \subseteq \bigcap_{n=1}^\infty U_n.$$

For the other inclusion, suppose $x \in F^c$. Therefore $x \notin \overline{F} = F$ as F is closed. Hence $\mathrm{dist}(x, F) > 0$. Choose $n \in \mathbb{N}$ such that

$$\mathrm{dist}(x, F) \geq \frac{1}{n} > 0.$$

Hence $d(x, y) \geq \frac{1}{n}$ for all $y \in \mathcal{F}$. Thus, by the definition of U_n, $x \notin U_n$. Whence $x \notin \bigcap_{n=1}^\infty U_n$. Hence

$$F = \bigcap_{n=1}^\infty U_n$$

so F is G_δ.

For another example, we prove the following.

Proposition 2.3.6. *The rational numbers are not a G_δ subset of $\mathbb{R}$.*

Proof. Suppose to the contrary that $\mathbb{Q}$ is G_δ. Hence there exists a collection of open sets $\{U_n\}_{n=1}^\infty$ such that $\mathbb{Q} = \bigcap_{n=1}^\infty U_n$. Therefore $\mathbb{Q} \subseteq U_n$ for all n so each U_n is dense in $\mathbb{R}$. Hence each U_n^c is closed and nowhere dense.

Notice that
$$\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n=1}^\infty U_n^c.$$
Since $\mathbb{Q}$ is countable, we may write $\mathbb{Q} = \{r_n \mid n \in \mathbb{N}\}$. Thus
$$\mathbb{R} = (\mathbb{R} \setminus \mathbb{Q}) \cup \mathbb{Q} = \left(\bigcup_{n=1}^\infty U_n^c\right) \cup \left(\bigcup_{n=1}^\infty \{r_n\}\right),$$
so $\mathbb{R}$ is a countable union of nowhere dense sets. As this contradicts the Baire Category Theorem (Corollary 2.3.2), the result is complete. ∎

To use Proposition 2.3.6 to show that certain sets cannot be the discontinuities of a real-valued function, we must analyze the set of discontinuities.

Lemma 2.3.7. *Let (X, d_X) and (Y, d_Y) be metric spaces, let $f : X \to Y$, and let*
$$D(f) = \{x \in X \mid f \text{ is not continuous at } x\}.$$
For each $n \in \mathbb{N}$ let
$$D_n(f) = \left\{ x \in X \;\middle|\; \begin{array}{c} \text{for every } \delta > 0 \text{ there exists } x_1, x_2 \in X \text{ such that} \\ d_X(x, x_1) < \delta, d_X(x, x_2) < \delta, \text{ and} \\ d_Y(f(x_1), f(x_2)) \geq \frac{1}{n} \end{array} \right\}.$$
Then $D_n(f)$ is closed for all $n \in \mathbb{N}$ and $D(f) = \bigcup_{n=1}^\infty D_n(f)$. Hence $D_n(f)$ is an F_σ subset of (X, d_X).

Proof. Fix $n \in \mathbb{N}$. To see that $D_n(f)$ is closed, let $(x_n)_{n\geq 1}$ be an arbitrary sequence of elements of $D_n(f)$ that converges to some $x \in X$. To see that $x \in D_n(f)$, let $\delta > 0$ be arbitrary. Since $(x_n)_{n\geq 1}$ converges to x, there exists a $N \in \mathbb{N}$ such that $d_X(x, x_N) < \frac{1}{2}\delta$. Furthermore, since $x_N \in D_n(f)$, there exists $a_1, a_2 \in X$ such that $d_X(x_N, a_1) < \frac{1}{2}\delta$, $d_X(x_N, a_2) < \frac{1}{2}\delta$, and $d_Y(f(a_1), f(a_2)) \geq \frac{1}{n}$. As $d_X(x, a_1) < \delta$ and $d_X(x, a_2) < \delta$ by the triangle inequality, and as $d_Y(f(a_1), f(a_2)) \geq \frac{1}{n}$, we obtain that $x \in D_n(f)$ as $\delta > 0$ was arbitrary. Hence as $(x_n)_{n\geq 1}$ was arbitrary, $D_n(f)$ is closed.

To see that $D(f) = \bigcup_{n=1}^\infty D_n(f)$, first suppose $x \in \bigcup_{n=1}^\infty D_n(f)$. Hence $x \in D_n(f)$ for some $n \in \mathbb{N}$. To see that f is discontinuous at x, suppose to the contrary that f is continuous at x. Notice by the definition of $D_n(f)$ that for each $m \in \mathbb{N}$ there exists points $x_{1,m}, x_{2,m} \in X$ such that $d_X(x, x_{1,m}) < \frac{1}{m}$, $d_X(x, x_{2,m}) < \frac{1}{m}$, and $d_Y(f(x_{1,m}, f(x_{2,m})) \geq \frac{1}{n}$.

Since $(x_{1,m})_{m\geq 1}$ and $(x_{2,m})_{m\geq 1}$ converge to x, the continuity of f implies $\lim_{m\to\infty} d_Y(f(x), f(x_{1,m})) = 0 = \lim_{m\to\infty} d_Y(f(x), f(x_{1,m}))$, which, together with the triangle inequality, contradicts the fact that

$$d_Y(f(x_{1,m}), f(x_{2,m})) \geq \frac{1}{n}$$

for all $m \geq 1$. Hence we have obtained a contradiction so $x \in D(f)$. Hence $\bigcup_{n=1}^{\infty} D_n(f) \subseteq D(f)$.

For the other inclusion, notice if $x \in D(f)$ then f is discontinuous at x. Therefore there exists an $\epsilon > 0$ such that for all $\delta > 0$ there exists a $x_1 \in X$ such that $d_X(x, x_1) < \delta$ yet $d_Y(f(x), f(x_1)) \geq \epsilon$. Choose $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$. By taking $x_2 = x$ in the definition of $D_n(f)$, we see that $x \in D_n(f)$. Hence, as x was arbitrary, $D(f) \subseteq \bigcup_{n=1}^{\infty} D_n(f)$ as desired. ∎

Theorem 2.3.8. *There does not exists a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous at each point in $\mathbb{Q}$ yet discontinuous at each point in $\mathbb{R} \setminus \mathbb{Q}$.*

Proof. Let $f : \mathbb{R} \to \mathbb{R}$. By Lemma 2.3.7 the set of discontinuities of f are F_σ. Thus the points where f is continuous must be a G_δ set. As $\mathbb{Q}$ is not G_δ by Proposition 2.3.6, f cannot be continuous at each point in $\mathbb{Q}$ yet discontinuous at each point in $\mathbb{R} \setminus \mathbb{Q}$. ∎

The Baire Category Theorem (Theorem 2.3.1) has many applications. In this section, we will look at the differences between finite dimensional and infinite dimensional Banach spaces. Looking at finite dimensional normed linear spaces, our first goal is to characterize which are Banach spaces. It turns out the answer is all of them! More than that, upto consider the equivalence of norms, there is only one n-dimensional Banach space.

To begin to see this, it helps to consider bounded linear maps between an n-dimensional normed linear space and $(\mathbb{K}^n, \|\cdot\|_\infty)$.

Lemma 2.3.9. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be an n-dimensional normed linear space over $\mathbb{K}$. If $\mathbb{K}^n$ is equipped with the ∞-norm, then there exists a bijective linear map $T : \mathbb{K}^n \to \mathcal{X}$ and two numbers $0 < k_1 \leq k_2 < \infty$ such that*

$$k_1 \|\vec{z}\|_\infty \leq \|T(\vec{z})\|_{\mathcal{X}} \leq k_2 \|\vec{z}\|_\infty$$

for all $\vec{z} \in \mathbb{K}^n$.

Proof. Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis for $\mathcal{X}$. Define $T : \mathbb{K}^n \to \mathcal{X}$ by

$$T((z_1, \ldots, z_n)) = z_1\vec{v}_1 + \cdots + z_n\vec{v}_n.$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$. Clearly T is linear and bijective by construction. Furthermore for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$

$$\begin{aligned}
\|T((z_1, \ldots, z_n))\|_{\mathcal{X}} &= \|z_1\vec{v}_1 + \cdots + z_n\vec{v}_n\|_{\mathcal{X}} \\
&\leq \sum_{k=1}^{n} |z_k| \, \|\vec{v}_k\|_{\mathcal{X}} \\
&\leq \left(\sum_{k=1}^{n} \|\vec{v}_k\|_{\mathcal{X}}\right) \|(z_1, \ldots, z_n)\|_\infty .
\end{aligned}$$

Hence T is bounded with $\|T\| \leq \sum_{k=1}^{n} \|\vec{v}_k\|_{\mathcal{X}}$. Hence we may take $k_2 = \|T\|$.

To find k_1 satisfying the other inequality, let

$$S_1 = \{\vec{z} \in \mathbb{K}^n \mid \|\vec{z}\|_\infty = 1\}.$$

Clearly S_1 is a closed bounded subset of $\mathbb{K}^n$ and therefore is compact by the Heine-Borel Theorem. Hence $T(S_1)$ is a compact subset of $\mathcal{X}$ being the continuous image of a compact set. Define $f : T(S_1) \to \mathbb{R}$ by

$$f(\vec{x}) = \|\vec{x}\|_{\mathcal{X}}$$

for all $\vec{x} \in T(S_1)$. Since f is continuous and since $T(S_1)$ is compact, by the Extreme Value Theorem there exists a $\vec{x}_0 \in T(S_1)$ such that

$$k_1 = f(\vec{x}_0) \leq f(\vec{x})$$

for all $\vec{x} \in T(S_1)$. Since $\vec{x}_0 \in T(S_1)$ and since T is a bijection, $\vec{x}_0 \neq \vec{0}$ so $k_1 > 0$.

We claim that

$$k_1 \|\vec{z}\|_\infty \leq \|T(\vec{z})\|_{\mathcal{X}}$$

for all $\vec{z} \in \mathbb{K}^n$. Clearly the inequality holds when $\vec{z} = 0$. Otherwise if $\vec{z} \neq 0$ then $\frac{1}{\|\vec{z}\|_\infty}\vec{z} \in S_1$ so

$$\|T(\vec{z})\|_{\mathcal{X}} = \|\vec{z}\|_\infty \left\| T\left(\frac{1}{\|\vec{z}\|_\infty}\vec{z}\right)\right\|_{\mathcal{X}} \geq k_1 \|\vec{z}\|_\infty .$$

Thus the result follows. ∎

Note if $T : \mathbb{K}^n \to \mathcal{X}$ is as in Lemma 2.3.9, then we can identify the underlying vector space of $\mathcal{X}$ with $\mathbb{K}^n$ and we can define a norm on $\mathbb{K}^n$ via $\|\vec{z}\| = \|T(\vec{z})\|_{\mathcal{X}}$ for all $\vec{z} \in \mathbb{K}^n$. The fact that T is linear and bijective easily yields that this is indeed the norm. Hence $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is isomorphic to $(\mathbb{K}^n, \|\cdot\|)$ for some norm $\|\cdot\|$. The conclusion of Lemma 2.3.9 is that this norm is equivalent to the infinity norm. Therefore, as equivalence of norms is an equivalence relation, we automatically have the following.

Corollary 2.3.10. *Let $\mathcal{V}$ be a finite dimensional vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on $\mathcal{V}$. Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

This characterization of every norm on a finite dimensional normed linear space being equivalent to the infinity norm on $\mathbb{K}^n$ yields some more results.

Corollary 2.3.11. *Every finite dimensional normed linear space is a Banach space.*

Proof. Every finite dimensional normed linear space is isomorphic $(\mathbb{K}^n, \|\cdot\|)$ for some norm $\|\cdot\|$. As $\|\cdot\|$ is equivalent to $\|\cdot\|_\infty$ and $\|\cdot\|_\infty$ is complete, we easy see that $\|\cdot\|$ is complete. Indeed suppose $(\vec{x}_n)_{n\geq 1}$ is an arbitrary Cauchy sequence in $(\mathbb{K}^n, \|\cdot\|)$. As there exists $0 < k_1 \leq k_2 < \infty$ such that

$$k_1 \|\vec{x}\| \leq \|\vec{x}\|_\infty \leq k_2 \|\vec{x}\|$$

for all $\vec{x} \in \mathbb{K}^n$, the second inequality implies that $(\vec{x}_n)_{n\geq 1}$ is Cauchy in $(\mathbb{K}^n, \|\cdot\|_\infty)$. As $(\mathbb{K}^n, \|\cdot\|_\infty)$ is complete, $(\vec{x}_n)_{n\geq 1}$ converges to some $\vec{x} \in \mathbb{K}^n$ with respect to $\|\cdot\|_\infty$. Hence the first inequality above implies $(\vec{x}_n)_{n\geq 1}$ converges to $\vec{x} \in \mathbb{K}^n$ with respect to $\|\cdot\|$. Hence $(\mathbb{K}^n, \|\cdot\|)$ is complete. ∎

Corollary 2.3.12. *Every finite dimensional subspace of a normed linear space is closed.*

Proof. Let W be a finite dimensional subspace of a normed linear space $(\mathcal{V}, \|\cdot\|)$. As $\|\cdot\| |_W$ is a norm on W, we see that W is complete with respect to $\|\cdot\|$ by Corollary 2.3.11. Hence W is closed in $\mathcal{V}$. ∎

Corollary 2.3.13. *Every linear map from a finite dimensional normed linear space into another normed linear space is bounded.*

Proof. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a finite dimensional normed linear space of dimension n, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be another normed linear space, and let $S : \mathcal{X} \to \mathcal{Y}$ be a linear map. If $\mathbb{K}^n$ is equipped with the ∞-norm, by Lemma 2.3.9 there exists a bounded linear map $T : \mathcal{X} \to \mathbb{K}^n$ such that $T^{-1} : \mathbb{K}^n \to \mathcal{X}$ is bounded. As

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

if it can be demonstrated that $S \circ T^{-1}$ is bounded, then S is a composition of continuous functions and thus will be continuous.

Let $R = S \circ T^{-1} : \mathbb{K}^n \to \mathcal{Y}$ and for each $k \in \{1, \ldots, n\}$, let

$$\vec{y}_k = R((0, \ldots, 0, 1, 0, \ldots, 0))$$

where the 1 occurs in the k^{th} position. Thus for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$

$$\begin{aligned} \|R((z_1, \ldots, z_n))\|_{\mathcal{Y}} &= \|z_1\vec{y}_1 + \cdots + z_n\vec{y}_n\|_{\mathcal{Y}} \\ &\leq \sum_{k=1}^n |z_k| \, \|\vec{y}_k\|_{\mathcal{Y}} \\ &\leq \left(\sum_{k=1}^n \|\vec{y}_k\|_{\mathcal{Y}} \right) \|(z_1, \ldots, z_n)\|_\infty . \end{aligned}$$

Hence R is bounded with $\|R\| \leq \sum_{k=1}^{n} \|\vec{y}_k\|_{\mathcal{Y}}$ as desired. ∎

With the resolution of every norm on a finite dimensional vector space yielding a Banach space, perhaps the next natural question is to examine which finite dimensional vector spaces with countable vector space bases are Banach spaces. It turns out the answer is none.

Theorem 2.3.14. *Every vector space basis of an infinite dimensional Banach space is uncountable.*

Proof. Suppose $(\mathcal{X}, \|\cdot\|)$ is an infinite dimensional Banach space with a countable basis $\{\vec{x}_n\}_{n=1}^{\infty}$. For each $n \in \mathbb{N}$, let

$$F_n = \mathrm{span}(\{\vec{x}_1, \ldots, \vec{x}_n\}).$$

Clearly each F_n is a finite dimensional vector space and thus is closed by Corollary 2.3.12.

We claim that $\mathrm{int}(F_n) = \emptyset$ for each $n \in \mathbb{N}$. Indeed, if $\mathrm{int}(F_n) \neq \emptyset$, then there exists an element $\vec{x} \in F_n$ and an $\epsilon > 0$ such that $B(\vec{x}, \epsilon) \subseteq F_n$. However, since F_n is a subspace and closed under translation and scaling, this implies $B(\vec{0}, \epsilon) \subseteq F_n$ by translation and $B(\vec{0}, r) \subseteq F_n$ for all $r > 0$ by scaling. As the later implies $F_n = \mathcal{X}$, we would obtain $\mathcal{X}$ is finite dimensional contradicting the fact that $\mathcal{X}$ is infinite dimensional. Thus $\mathrm{int}(F_n) = \emptyset$ for each $n \in \mathbb{N}$.

The above shows each F_n is nowhere dense. Since $\{\vec{x}_n\}_{n=1}^{\infty}$ is a basis for $\mathcal{X}$ and

$$\mathcal{X} = \bigcup_{n=1}^{\infty} F_n,$$

X is a countable union of nowhere dense sets. As this contradicts the Baire Category Theorem (Corollary 2.3.2) as $\mathcal{X}$ is a Banach space, the proof is complete. ∎

To conclude our discussions on the differences between finite and infinite dimensional normed linear spaces, we note that all norms on a finite dimensional Banach space are equivalent by Corollary 2.3.10. Of course this is not the case for an infinite dimensional Banach space.

Proposition 2.3.15. *Let $(\mathcal{X}, \|\cdot\|)$ be an infinite dimensional Banach space. Then there exists another norm $\|\cdot\|_0 : \mathcal{X} \to [0, \infty)$ such that $(\mathcal{X}, \|\cdot\|_0)$ is a Banach space, yet $\|\cdot\|$ and $\|\cdot\|_0$ are not equivalent.*

Proof. Let $(\mathcal{X}, \|\cdot\|)$ be an infinite dimensional Banach space with basis $\{\vec{x}_\lambda\}_{\lambda \in \Lambda}$. By scaling if necessary, we may assume that $\|\vec{x}_\lambda\| = 1$ for all $\lambda \in \Lambda$. As Λ must be infinite, choose distinct vectors $\{\vec{x}_n\}_{n \geq 1}$ from $\{\vec{x}_\lambda\}_{\lambda \in \Lambda}$. Define a linear map $f : \mathcal{X} \to \mathbb{K}$ by defining $f(\vec{x}_n) = n$ for all $n \in \mathbb{N}$, $f(\vec{x}) = 0$ for all $\vec{x} \in \{\vec{x}_\lambda\}_{\lambda \in \Lambda} \setminus \{\vec{x}_n\}_{n \geq 1}$, and by extending the definition of f by linearity. As $|f(\vec{x}_n)| \geq n$ and $\|\vec{x}_n\| = 1$, we see that f is unbounded.

Let $\vec{y} = \vec{x}_1$ so that $f(\vec{y}) = 1$. Define $S : \mathcal{X} \to \mathcal{X}$ by

$$S(\vec{x}) = \vec{x} - 2f(\vec{x})\vec{y}$$

for all $\vec{x} \in \mathcal{X}$. Clearly S is well-defined and linear as $f : \mathcal{X} \to \mathbb{K}$ is linear.

We claim that S^2 is the identity map on $\mathcal{X}$. To see this, notice for all $\vec{x} \in \mathcal{X}$ that

$$\begin{aligned} S^2(\vec{x}) &= S(\vec{x} - 2f(\vec{x})\vec{y}) \\ &= (\vec{x} - 2f(\vec{x})\vec{y}) - 2f(\vec{x})(\vec{y} - 2f(\vec{y})\vec{y}) \\ &= (\vec{x} - 2f(\vec{x})\vec{y}) - 2f(\vec{x})(-\vec{y}) = \vec{x}. \end{aligned}$$

Therefore, as $\vec{x} \in \mathcal{X}$ was arbitrary, S^2 is the identity map on $\mathcal{X}$.

Define $\|\cdot\|_0 : \mathcal{X} \to [0, \infty)$ by

$$\|\vec{x}\|_0 = \|S(\vec{x})\|_0$$

for all $\vec{x} \in \mathcal{X}$. Since S^2 is the identity, S must be bijective which implies that $\|\cdot\|_0$ is norm on $\mathcal{X}$.

We claim that $(\mathcal{X}, \|\cdot\|_0)$ is complete. To see this, let $(\vec{x}_n)_{n\geq 1}$ be an arbitrary Cauchy in $(\mathcal{X}, \|\cdot\|_0)$. We claim that $(S(\vec{x}_n))_{n\geq 1}$ is Cauchy in $(\mathcal{X}, \|\cdot\|)$. To see this, let $\epsilon > 0$. Since $(\vec{x}_n)_{n\geq 1}$ is Cauchy in $(\mathcal{X}, \|\cdot\|_0)$, there exists an $N \in \mathbb{N}$ such that $\|\vec{x}_n - \vec{x}_m\|_0 < \epsilon$ for all $n, m \geq N$. Hence for all $n, m \geq N$ we have that

$$\|S(\vec{x}_n) - S(\vec{x}_m)\| = \|S(\vec{x}_n - \vec{x}_m)\| = \|\vec{x}_n - \vec{x}_m\|_0 < \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $(S(\vec{x}_n))_{n\geq 1}$ is Cauchy in $(\mathcal{X}, \|\cdot\|)$.

Since $(\mathcal{X}, \|\cdot\|)$ is complete, $(S(\vec{x}_n))_{n\geq 1}$ is converges in $(\mathcal{X}, \|\cdot\|)$. Hence there exists a vector $\vec{z} \in \mathcal{X}$ such that $\lim_{n\to\infty} \|S(\vec{x}_n) - \vec{z}\| = 0$. We claim that $(\vec{x}_n)_{n\geq 1}$ converges in $(\mathcal{X}, \|\cdot\|_0)$ to $S(\vec{z})$. To see this, notice that

$$\|\vec{x}_n - S(\vec{z})\|_0 = \|S(\vec{x}_n - S(\vec{z}))\| = \left\|S(\vec{x}_n) - S^2(\vec{z})\right\| = \|S(\vec{x}_n) - \vec{z}\|$$

for all $n \in \mathbb{N}$. Therefore, since $\lim_{n\to\infty} \|S(\vec{x}_n) - \vec{z}\| = 0$ we obtain that $\lim_{n\to\infty} \|\vec{x}_n - S(\vec{z})\| = 0$. Hence $(\vec{x}_n)_{n\geq 1}$ converges in $(\mathcal{X}, \|\cdot\|_0)$ to $S(\vec{z})$. Therefore, as $(\vec{x}_n)_{n\geq 1}$ was arbitrary, $(\mathcal{X}, \|\cdot\|_0)$ is complete.

Finally, we claim that $\|\cdot\|$ and $\|\cdot\|_0$ are not equivalent. To see this, we claim there exists not exist a constant $C \in \mathbb{R}$ such that

$$\|\vec{x}\|_0 \leq C \|\vec{x}\|$$

for all $\vec{x} \in \mathcal{X}$. Indeed if $\{\vec{x}_n\}_{n=1}^\infty$ are as above, then $C \|\vec{x}_n\| = C$ whereas

$$\begin{aligned} \|\vec{x}_n\|_0 &= \|S(\vec{x}_n)\| \\ &= \|\vec{x}_n - T(\vec{x}_n)\vec{y}\| \\ &= \|\vec{x}_n - n\vec{y}\| \\ &\geq n\|\vec{y}\| - \|\vec{x}_n\| \\ &\geq n - 1 \end{aligned}$$

for all $n \in \mathbb{N}$. Hence such a C does not exist. ∎

2.4 Open Mapping Theorem

The Baire Category Theorem (Theorem 2.3.1) has many applications. In this section, we will study surjective bounded linear maps between Banach spaces. In particular, since bounded linear maps are continuous, the inverse images of open sets are open. The goal of this section is to prove that surjective bounded linear maps map open sets to open sets. This enables us to prove that the inverses of bijective bounded linear maps are bounded and characterize continuous linear maps using their graphs.

To begin, we require the following odd looking result that says if an open ball is in the closure of the image of a bounded linear map of an open ball, then we can expand the later open ball to obtain strict containment.

Lemma 2.4.1. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a normed linear space, and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. If $B_{\mathcal{Y}}(\vec{0}, 1) \subseteq \overline{T(B_{\mathcal{X}}(\vec{0}, m))}$ for some $m > 0$, then $B_{\mathcal{Y}}(\vec{0}, 1) \subseteq T(B_{\mathcal{X}}(\vec{0}, 2m))$*

Proof. Let $m > 0$ be such that $B_{\mathcal{Y}}(\vec{0}, 1) \subseteq \overline{T(B_{\mathcal{X}}(\vec{0}, m))}$. Notice for all $\alpha \in \mathbb{R}$ (where for a set A, $\alpha A = \{\alpha a \mid a \in A\}$) that

$$\begin{aligned} B_{\mathcal{Y}}(\vec{0}, \alpha) = \alpha B_{\mathcal{Y}}(\vec{0}, 1) &\subseteq \alpha\overline{T(B_{\mathcal{X}}(\vec{0}, m))} \\ &= \overline{\alpha T(B_{\mathcal{X}}(\vec{0}, m))} \\ &= \overline{T(\alpha B_{\mathcal{X}}(\vec{0}, m))} = \overline{T(B_{\mathcal{X}}(\vec{0}, \alpha m))} \end{aligned}$$

by linearity and continuity of T, and by properties of the norm.

To see that $B_{\mathcal{Y}}(\vec{0}, 1) \subseteq T(B_{\mathcal{X}}(\vec{0}, 2m))$, let $\vec{y} \in B_{\mathcal{Y}}(\vec{0}, 1)$ be arbitrary. Since $\vec{y} \in \overline{T(B_{\mathcal{X}}(\vec{0}, m))}$ there exists an $\vec{x}_1 \in B_{\mathcal{X}}(\vec{0}, m)$ such that

$$\|\vec{y} - T(\vec{x}_1)\|_{\mathcal{Y}} < \frac{1}{2}.$$

Let $\vec{y}_1 = \vec{y} - T(\vec{x}_1) \in \mathcal{Y}$. Then $\vec{y}_1 \in B_{\mathcal{Y}}(\vec{0}, \frac{1}{2}) \subseteq \overline{T(B_{\mathcal{X}}(\vec{0}, \frac{1}{2}m))}$. Hence there exists an $\vec{x}_2 \in B_{\mathcal{X}}(\vec{0}, \frac{1}{2}m)$ such that

$$\|\vec{y}_1 - T(\vec{x}_2)\|_{\mathcal{Y}} < \frac{1}{2^2}.$$

Repeating this process ad nauseum, we obtain a sequence of vectors $(\vec{y}_n)_{n\geq 1}$ in $\mathcal{Y}$ and a sequence of vectors $(\vec{x}_n)_{n\geq 1}$ in $\mathcal{X}$ such that $\vec{y}_n = \vec{y}_{n-1} - T(\vec{x}_n)$, $\vec{y}_n \in B_{\mathcal{Y}}(\vec{0}, \frac{1}{2^n})$, $\vec{x}_{n+1} \in B_{\mathcal{X}}(\vec{0}, \frac{1}{2^n}m)$, and

$$\|\vec{y}_n - T(\vec{x}_{n+1})\|_{\mathcal{Y}} < \frac{1}{2^n}$$

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

Since $\mathcal{X}$ is a Banach space and since

$$\sum_{n=1}^{\infty} \|\vec{x}_n\|_{\mathcal{X}} < \sum_{n=1}^{\infty} \frac{1}{2^n} m = 2m < \infty,$$

we obtain by Theorem 2.2.2 that $\vec{x} = \sum_{n=1}^{\infty} \vec{x}_n$ exists and is an element of $B_{\mathcal{X}}(\vec{0}, 2m)$. To see that $T(\vec{x}) = \vec{y}$ thereby completing the proof, notice since T is continuous that

$$\begin{aligned}
\|\vec{y} - T(\vec{x})\|_{\mathcal{Y}} &= \lim_{n\to\infty} \left\| \vec{y} - T\left(\sum_{k=1}^{n} \vec{x}_k\right) \right\| \\
&= \lim_{n\to\infty} \left\| \vec{y} - \sum_{k=1}^{n} T(\vec{x}_k) \right\| \\
&= \lim_{n\to\infty} \left\| \vec{y}_1 - \sum_{k=2}^{n} T(\vec{x}_k) \right\| \\
&= \lim_{n\to\infty} \left\| \vec{y}_2 - \sum_{k=3}^{n} T(\vec{x}_k) \right\| \\
&\vdots \\
&= \lim_{n\to\infty} \|\vec{y}_n\| \\
&\leq \limsup_{n\to\infty} \frac{1}{2^n} = 0.
\end{aligned}$$

Hence $T(\vec{x}) = \vec{y}$. Therefore, since $\vec{y} \in B_{\mathcal{Y}}(\vec{0}, 1)$ was arbitrary, $B_{\mathcal{Y}}(\vec{0}, 1) \subseteq T(B_{\mathcal{X}}(\vec{0}, 2m))$. ∎

Combining Lemma 2.4.1 together with the Baire Category Theorem (Theorem 2.3.1), we obtain the following result.

Theorem 2.4.2 (Open Mapping Theorem). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces. If $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is surjective and $U \subseteq \mathcal{X}$ is open, then $T(U)$ is open in $\mathcal{Y}$.*

Proof. Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. First we will demonstrate that there exists an $r > 0$ such that $T(B_{\mathcal{X}}(\vec{0}, r))$ is a neighbourhood of $\vec{0}$ in $\mathcal{Y}$.

To begin, for each $m \in \mathbb{N}$ consider the set $F_m = \overline{T(B_{\mathcal{X}}(\vec{0}, m))} \subseteq \mathcal{Y}$. Clearly each F_m is a closed subset of $\mathcal{Y}$. Moreover, since T is surjective,

$$\mathcal{Y} = \bigcup_{m=1}^{\infty} F_m.$$

Therefore, since $\mathcal{Y}$ is complete, the Baire Category Theorem (Corollary 2.3.2) implies there must exists an $m_0 \in \mathbb{N}$ such that F_{m_0} is not nowhere dense.

Hence $\mathrm{int}(F_{m_0}) \neq \emptyset$. Therefore there exists an $\vec{y}_0 \in F_{m_0}$ and a $\delta > 0$ such that $B_{\mathcal{Y}}(\vec{y}_0, \delta) \subseteq F_{m_0} = \overline{T(B_{\mathcal{X}}(\vec{0}, m))}$. Since

$$\begin{aligned}
B_{\mathcal{Y}}(\vec{0}, \delta) &\subseteq \{\vec{y} - \vec{y}_0 \mid \vec{y} \in B_{\mathcal{Y}}(\vec{y}_0, \delta)\} \\
&\subseteq \{\vec{y}_1 - \vec{y}_2 \mid \vec{y}_1, \vec{y}_2 \in F_{m_0}\} \\
&= \{\vec{y}_1 + \vec{y}_2 \mid \vec{y}_1, \vec{y}_2 \in \overline{T(B_{\mathcal{X}}(\vec{0}, m))}\} \\
&\qquad \text{as } T \text{ is linear, } -B_{\mathcal{X}}(\vec{0}, m) = B_{\mathcal{X}}(\vec{0}, m) \\
&\subseteq \overline{T(B_{\mathcal{X}}(\vec{0}, 2m))} \\
&\qquad \text{by continuity, linearity, and the triangle inequality,}
\end{aligned}$$

we obtain by Lemma 2.4.1 that $B_{\mathcal{Y}}(\vec{0}, \delta) \subseteq T(B_{\mathcal{X}}(\vec{0}, 4m))$.

To complete the result, let U be an arbitrary open subset of $\mathcal{X}$. To see that $T(U)$ is open in $\mathcal{Y}$, let $\vec{y} \in T(U)$ be arbitrary. Thus there exists a $\vec{x} \in \mathcal{X}$ such that $T(\vec{x}) = \vec{y}$. Since U is open, there exists an $\epsilon > 0$ such that $B_{\mathcal{X}}(\vec{x}, \epsilon) \subseteq U$. However since

$$B_{\mathcal{Y}}\left(\vec{0}, \frac{\epsilon\delta}{4m}\right) = \frac{\epsilon}{4m} B_{\mathcal{Y}}(\vec{0}, \delta) \subseteq \frac{\epsilon}{4m} T(B_{\mathcal{X}}(\vec{0}, 4m)) = T(B_{\mathcal{X}}(\vec{0}, \epsilon))$$

we have that

$$\begin{aligned}
B_{\mathcal{Y}}\left(\vec{y}, \frac{\epsilon\delta}{4m}\right) &= \left\{\vec{y} + \vec{z} \,\middle|\, \vec{z} \in B_{\mathcal{Y}}\left(\vec{0}, \frac{\epsilon\delta}{4m}\right)\right\} \\
&\subseteq \left\{T(\vec{x}) + \vec{z} \,|\, \vec{z} \in T(B_{\mathcal{X}}(\vec{0}, \epsilon))\right\} \\
&= \left\{T(\vec{x}) + T(\vec{w}) \,|\, \vec{w} \in B_{\mathcal{X}}(\vec{0}, \epsilon)\right\} \\
&= \left\{T(\vec{x} + \vec{w}) \,|\, \vec{w} \in B_{\mathcal{X}}(\vec{0}, \epsilon)\right\} \\
&= T(B_{\mathcal{X}}(\vec{x}, \epsilon))
\end{aligned}$$

by the linearity of T. Hence $T(U)$ contains an open neighbourhood around $\vec{y}$. Therefore, since $\vec{y} \in T(U)$ was arbitrary, $T(U)$ is open in $\mathcal{Y}$. Hence since U was an arbitrary open subset of $\mathcal{X}$, the result follows. ∎

The Open Mapping Theorem has several applications.

Theorem 2.4.3 (The Inverse Mapping Theorem). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be a bijection. Then $T^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$.*

Proof. Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be a bijection. Therefore $T^{-1} : \mathcal{Y} \to \mathcal{X}$ exists. Since T is linear, clearly T^{-1} is linear. To see that T^{-1} is bounded (i.e. continuous via Theorem 1.4.6), let $U \subseteq \mathcal{X}$ be open. Then

$$(T^{-1})^{-1}(U) = T(U)$$

is open in $\mathcal{Y}$ by the Open Mapping Theorem (Theorem 2.4.2). Hence T is continuous. ∎

Using the Inverse Mapping Theorem (Theorem 2.4.3), we obtain the following property relating different norms on Banach spaces.

Corollary 2.4.4. *Let $\mathcal{X}$ be a vector space over $\mathbb{K}$ that is complete with respect to each of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If there exists a constant $c_1 \in \mathbb{R}$ such that*

$$\|\vec{x}\|_1 \leq c_1 \|\vec{x}\|_2$$

for all $\vec{x} \in \mathcal{X}$, then there exists a constant $c_2 \in \mathbb{R}$ such that

$$\|\vec{x}\|_2 \leq c_2 \|\vec{x}\|_1$$

for all $\vec{x} \in \mathcal{X}$.

Proof. Define $T : (\mathcal{X}, \|\cdot\|_2) \to (\mathcal{X}, \|\cdot\|_1)$ by $T(\vec{x}) = \vec{x}$. Clearly T is a linear map. Moreover, since

$$\|\vec{x}\|_1 \leq c_1 \|\vec{x}\|_2$$

for all $\vec{x} \in \mathcal{X}$, we see that T is a bounded linear map from $(\mathcal{X}, \|\cdot\|_2)$ to $(\mathcal{X}, \|\cdot\|_1)$. Hence, by the Inverse Mapping Theorem (Theorem 2.4.3), T^{-1} is a bounded linear map from $(\mathcal{X}, \|\cdot\|_1)$ to $(\mathcal{X}, \|\cdot\|_2)$. Since $T^{-1}(\vec{x}) = \vec{x}$ for all $\vec{x} \in \mathcal{X}$, we obtain that

$$\|\vec{x}\|_2 = \left\|T^{-1}(\vec{x})\right\|_2 \leq \left\|T^{-1}\right\| \|\vec{x}\|_1 .$$

Thus letting $c_2 = \|T^{-1}\|$ completes the proof. ∎

Another nice application of the Open Mapping Theorem (Theorem 2.4.2) is the characterization of continuous linear maps via their graphs.

Theorem 2.4.5 (The Closed Graph Theorem). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces and let $T : \mathcal{X} \to \mathcal{Y}$ be linear. The graph*

$$\mathcal{G}(T) = \{(\vec{x}, T(\vec{x})) \mid \vec{x} \in \mathcal{X}\}$$

is closed in $\mathcal{X} \oplus_1 \mathcal{Y}$ if and only if T is continuous.

Proof. To see that $\mathcal{G}(T)$ is closed when T is continuous, suppose T is continuous and let $((\vec{x}_n, T(\vec{x}_n)))_{n\geq 1}$ be an arbitrary sequence of elements of $\mathcal{G}(T)$ that converges to some element $(\vec{x}, \vec{y}) \in \mathcal{X} \oplus_1 \mathcal{Y}$. Clearly this implies $(\vec{x}_n)_{n\geq 1}$ converges to $\vec{x}$ in $\mathcal{X}$ and $(T(\vec{x}_n))_{n\geq 1}$ converges to $\vec{y} \in \mathcal{Y}$. Since T is continuous, $(\vec{x}_n)_{n\geq 1}$ converging to $\vec{x}$ in $\mathcal{X}$ implies that $(T(\vec{x}_n))_{n\geq 1}$ converges to $T(\vec{x})$. Therefore, due to uniqueness of limits, we must have that $\vec{y} = T(\vec{x})$. Hence $(\vec{x}, \vec{y}) \in \mathcal{G}(T)$ so $\mathcal{G}(T)$ is closed.

Conversely, suppose $\mathcal{G}(T)$ is closed in $\mathcal{X} \oplus_1 \mathcal{Y}$. Therefore, since $\mathcal{G}(T)$ is a vector subspace of $\mathcal{X} \oplus_1 \mathcal{Y}$ as T is linear, and since $\mathcal{X} \oplus_1 \mathcal{Y}$ is a Banach space, $\mathcal{G}(T)$ is also a Banach space.

Define $S : \mathcal{X} \to \mathcal{G}(T)$ by

$$S(\vec{x}) = (\vec{x}, T(\vec{x}))$$

for all $\vec{x} \in \mathcal{X}$. Clearly S is a linear map that is injective (by the first coordinate) and surjective. Hence S is invertible with $S^{-1} : \mathcal{G}(T) \to \mathcal{X}$ defined by

$$S^{-1}((\vec{x}, T(\vec{x}))) = \vec{x}.$$

Notice for all $(\vec{x}, T(\vec{x})) \in \mathcal{G}(T)$ that

$$\left\| S^{-1}((\vec{x}, T(\vec{x}))) \right\|_{\mathcal{X}} = \|\vec{x}\|_{\mathcal{X}} \leq \|\vec{x}\|_{\mathcal{X}} + \|T(\vec{x})\|_{\mathcal{Y}} = \|(\vec{x}, T(\vec{x}))\|_1 .$$

Therefore S^{-1} is bounded. Hence, as $\mathcal{X}$ and $\mathcal{G}(T)$ are Banach spaces, the Inverse Mapping Theorem (Theorem 2.4.3) implies that S is bounded. Therefore, since

$$\|T(\vec{x})\|_{\mathcal{Y}} \leq \|T(\vec{x})\|_{\mathcal{Y}} + \|\vec{x}\|_{\mathcal{X}} = \|S(\vec{x})\|_1 \leq \|S\| \, \|\vec{x}\|_{\mathcal{X}} ,$$

we see that T is bounded as desired. Hence T is continuous as desired. ∎

2.5 Principle of Uniform Boundedness

For our final major Banach space theorem of this chapter, we will use the Baire Category Theorem (Theorem 2.3.1) to deduce collections of objects are uniform boundedness from simply knowing they are pointwise bounded!

We begin with the following Uniform Boundness Principles for continuous functions on complete metric spaces.

Theorem 2.5.1 (Uniform Boundedness Principle). *Let $(\mathcal{X}, d_{\mathcal{X}})$ be a complete metric space, let $(\mathcal{Y}, d_{\mathcal{Y}})$ be a metric space, let $y \in \mathcal{Y}$ be a fixed element, and let $\mathcal{F} \subseteq C(\mathcal{X}, \mathcal{Y})$ be a non-empty set of functions such that for each $x \in \mathcal{X}$*

$$M_x = \sup_{f \in \mathcal{F}} d_{\mathcal{Y}}(f(x), y) < \infty.$$

Then there exists a non-empty open subset U of X and a constant $M > 0$ such that

$$d_{\mathcal{Y}}(f(x), y) \leq M$$

for all $f \in \mathcal{F}$ and $x \in U$.

Proof. For each $n \in \mathbb{N}$, let

$$F_n = \left\{ x \in \mathcal{X} \;\middle|\; \sup_{f \in \mathcal{F}} d_{\mathcal{Y}}(f(x), y) \leq n \right\}.$$

Clearly each F_n is a closed set as each element of $\mathcal{F}$ is continuous and the distance function is continuous. Furthermore, if $x \in \mathcal{X}$ then $x \in F_n$ for all $n \geq M_x$. Hence

$$\mathcal{X} = \bigcup_{n=1}^{\infty} F_n.$$

Therefore, by the Baire Category Theorem (Corollary 2.3.2), there exists an $n_0 \in \mathbb{N}$ such that F_{n_0} is not nowhere dense in $\mathcal{X}$. Therefore $\emptyset \neq \mathrm{int}(\overline{F_{n_0}}) = \mathrm{int}(F_{n_0})$ so there exists an open subset U of $\mathcal{X}$ with $U \subseteq F_{n_0}$. Hence for all $x \in U$ we have $d_{\mathcal{Y}}(f(x), y) \leq n_0$ for all $f \in \mathcal{F}$ as desired. ∎

Remark 2.5.2. It is actually possible to prove a version of Theorem 2.5.1 for continuous functions on compact Hausdorff spaces. Indeed one need only verify that every compact Hausdorff space satisfies the Baire Category Theorem. Such spaces are called *Baire spaces* in topology and behave in a very similar fashion to metric spaces. We will not present the proof that compact Hausdorff spaces are Baire here as to do so we would need to delve into the separation actions in topology.

Note Theorem 2.5.1 is most useful when $\mathcal{Y}$ is a normed linear space and $y = \vec{0}$. In this case, the assumption becomes

$$M_x = \sup_{f \in \mathcal{F}} \|f(x)\|_{\mathcal{Y}} < \infty$$

and the conclusion becomes

$$\|f(x)\|_{\mathcal{Y}} \leq M$$

for all $f \in \mathcal{F}$ and $x \in U$.

Building on the above theorem, we obtain the following Uniform Boundedness Principle for bounded linear maps between Banach spaces

Theorem 2.5.3 (Uniform Boundedness Principle - Banach space version). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ a normed linear space, and let $\mathcal{F} \subseteq \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be non-empty. Suppose for each $\vec{x} \in \mathcal{X}$ that*

$$\sup\{\|T(\vec{x})\|_{\mathcal{Y}} \mid T \in \mathcal{F}\} < \infty.$$

Then

$$\sup\{\|T\| \mid T \in \mathcal{F}\} < \infty.$$

Proof. For each $T \in \mathcal{F}$, consider the function $f_T : \mathcal{X} \to \mathbb{R}$ defined by

$$f_T(\vec{x}) = \|T(\vec{x})\|_{\mathcal{Y}}$$

for all $\vec{x} \in \mathcal{X}$. Since T and the norm are continuous functions on $\mathcal{X}$, it is elementary to see that $f_T \in C(\mathcal{X}, \mathbb{R})$ for all $T \in \mathcal{F}$.

Let
$$\mathcal{F}_0 = \{f_T \mid T \in \mathcal{F}\} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R}).$$
Since $\mathcal{X}$ is complete and since
$$\sup_{f \in \mathcal{F}_0} |f(\vec{x})| < \infty$$
for all $\vec{x} \in \mathcal{X}$, Theorem 2.5.1 implies that there exists an $M > 0$ and a non-empty open subset U of $\mathcal{X}$ such that
$$\|T(\vec{x})\| = |f_T(\vec{x})| \leq M$$
for all $\vec{x} \in U$ and $T \in \mathcal{F}$.

Since U is a non-empty open set of $\mathcal{X}$, there exists a vector $\vec{x}_0 \in U$ and an $\epsilon > 0$ so that $B_{\mathcal{X}}(\vec{x}_0, \epsilon) \subseteq U$. To obtain the conclusion, let $T \in \mathcal{F}$ be arbitrary. Notice if $\vec{x} \in B_{\mathcal{X}}(\vec{0}, \epsilon)$, then
$$\|T(\vec{x})\|_{\mathcal{Y}} \leq \|T(\vec{x} + \vec{x}_0)\|_{\mathcal{Y}} + \|-T(\vec{x}_0)\|_{\mathcal{Y}} \leq M + \|T(\vec{x}_0)\|_{\mathcal{Y}}.$$
as $\vec{x} + \vec{x}_0 \in B_{\mathcal{X}}(\vec{x}_0, \epsilon)$. Therefore, if $\vec{z} \in B_{\mathcal{X}}(\vec{0}, 1)$, then
$$\|T(\vec{z})\|_{\mathcal{Y}} = \frac{1}{\epsilon} \|T(\epsilon \vec{z})\|_{\mathcal{Y}} \leq \frac{1}{\epsilon}\left(M + \|T(\vec{x}_0)\|_{\mathcal{Y}}\right)$$
as $\epsilon \vec{z} \in B_{\mathcal{X}}(\vec{0}, \epsilon)$. Hence
$$\|T\| \leq \frac{1}{\epsilon}\left(M + \|T(\vec{x}_0)\|_{\mathcal{Y}}\right).$$
Therefore, as $T \in \mathcal{F}$ was arbitrary and as $\sup_{T \in \mathcal{F}} \|T(\vec{x}_0)\|_{\mathcal{Y}} < \infty$, the proof is complete. ∎

Of course we immediately obtain the following corollary.

Corollary 2.5.4. *Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space and let $\mathcal{F} \subseteq \mathcal{X}^*$ be non-empty. Then $\mathcal{F}$ is bounded if and only if*
$$\sup(\{f(x) \mid f \in \mathcal{F}\}) < \infty$$
for all $x \in \mathcal{X}$.

The Uniform Boundedness Principle (Theorem 2.5.3) is particularly useful to show the pointwise limit of bounded linear maps products a bounded linear map.

Theorem 2.5.5 (The Banach-Steinhaus Theorem). *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a normed linear space, and let $(T_n)_{n\geq 1}$ be a sequence of elements of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ such that for all $\vec{x} \in \mathcal{X}$*
$$\lim_{n\to\infty} T_n(\vec{x})$$
exists in $\mathcal{Y}$. Then $\sup_{n\geq 1} \|T_n\| < \infty$ and the map $T : \mathcal{X} \to \mathcal{Y}$ defined by $T(\vec{x}) = \lim_{n\to\infty} T_n(\vec{x})$ is an element of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Proof. Since for each $\vec{x} \in \mathcal{X}$ the limit $\lim_{n\to\infty} T_n(\vec{x})$ exists, the sequence $(T(\vec{x}_n))_{n\geq 1}$ is bounded. Therefore, by the Principle of Uniform Boundedness (Theorem 2.5.3), we obtain that $\sup_{n\geq 1} \|T_n\| < \infty$.

Define $T : \mathcal{X} \to \mathcal{Y}$ by $T(\vec{x}) = \lim_{n\to\infty} T_n(\vec{x})$. Clearly if $\vec{x}_1, \vec{x}_2 \in \mathcal{X}$ and $\alpha \in \mathbb{K}$ then

$$T(\vec{x}_1+\alpha\vec{x}_2) = \lim_{n\to\infty} T_n(\vec{x}_1+\alpha\vec{x}_2) = \lim_{n\to\infty} T_n(\vec{x}_1)+\alpha T_n(\vec{x}_2) = T(\vec{x}_1)+\alpha T(\vec{x}_2)$$

so T is linear. To see that T is bounded, we note for all $\vec{x} \in \mathcal{X}$ that

$$\|T(\vec{x})\| = \lim_{n\to\infty} \|T_n(\vec{x})\| \leq \limsup_{n\to\infty} \|T_n| \, \|\vec{x}\| \leq \left(\sup_{n\geq 1} \|T_n\| \right) \|\vec{x}\| .$$

Therefore, as $\sup_{n\geq 1} \|T_n\| < \infty$, T is bounded. ∎

Of course, there are many other uses of the Uniform Boundedness Principle (Theorem 2.5.3) and the Banach-Steinhaus Theorem (Theorem 2.5.5). For example, one can use the Uniform Boundedness Principle (Theorem 2.5.3) to prove that there exists a continuous function whose Fourier series does not converge pointwise. In addition, there are many more uses in functional analysis as we will see in later chapters.

Chapter 3

Topological Vector Spaces

Although Banach spaces and normed linear spaces have essential properties one wants when performing analytical computations, there are many other types of convergence that occur in analysis that do not come from norms. For example, pointwise convergence is very common in analysis. Although pointwise convergence does not behave as nice as uniform convergence, it still has its role to play. Thus we desire the appropriate structures to examine different types of convergence in analysis.

As always, we want to be working in a vector space with a natural topology making it possible to discuss convergence. To make the topology 'compatible' with the vector space structures, it is necessary that vector addition and scalar multiplication are continuous. These so called *topological vector spaces* will be the focus of this section. In particular, for these objects, we will discuss their elementary properties, a natural way to generate them, how they behave under various operations, how notions in normed linear spaces generalize, and the structures of finite dimensional and of locally convex topological vector spaces.

3.1 Introduction to Topological Vector Spaces

We begin with the central object of study of this chapter.

Definition 3.1.1. A *topological vector space* is a pair $(\mathcal{V}, \mathcal{T})$ where $\mathcal{V}$ is a vector space over $\mathbb{K}$ and $\mathcal{T}$ is a Hausdorff topology on $\mathcal{V}$ such that the maps $\sigma : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and $\rho : \mathbb{K} \times \mathcal{V} \to \mathcal{V}$ defined by

$$\sigma(x, y) = x + y \qquad \text{and} \qquad \rho(\alpha, x) = \alpha x$$

for all $x, y \in \mathcal{V}$ and $\alpha \in \mathbb{K}$ are continuous where $\mathcal{V} \times \mathcal{V}$ and $\mathbb{K} \times \mathcal{V}$ are equipped with their product topologies.

Remark 3.1.2. It is work noting that some authors do not require topological vector spaces to be Hausdorff. The rationale for why we force topological vector spaces to be Hausdorff is that we want unique limits. For those worried that we are not being general enough and in the event one encounters such a topology where Hausdorff is excluded, we note it is always possible to consider a quotient vector space that will be Hausdorff.

By Proposition 1.1.4, it is clear that every normed linear space is a topological vector space. Of course it is enough to consider sequences in metric topologies and thus Proposition 1.1.4 suffices. For other topologies that do others not from norms, one must consider nets when demonstrating addition and scalar multiplication are continuous.

Of course the discrete topology on any vector space automatically produces a topological vector space, which is quite boring as every map is continuous and nets converge if and only if they are eventually constant. Before we get to looking at non-trivial examples of topological vector spaces that are not normed linear spaces, it is useful to examine elementary properties satisfied by open sets in topological vector spaces so that we know what behaviours occur and are necessary.

Remark 3.1.3. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $x_0 \in \mathcal{V}$ be fixed. Since the map $f : \mathcal{V} \to \mathcal{V}$ defined by $f(x) = x + x_0$ is clearly seen to be a homeomorphism, we see that a subset $U \subseteq \mathcal{V}$ is a neighbourhood of a vector $y \in \mathcal{V}$ if and and only if

$$x_0 + U = \{x_0 + u \mid u \in U\}$$

is a neighbourhood of $x_0 + y$. By taking $y = \vec{0}$, we see that neighbourhood basis in $(\mathcal{V}, \mathcal{T})$ at any point is in one-to-one correspondence with the neighbourhood basis at $\vec{0}$. Thus, when considering topological matters, it often suffices to consider only neighbourhoods of $\vec{0}$ by translating the problem.

Remark 3.1.4. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let V be a neighbourhood of $\vec{0}$. Since addition is continuous, the set

$$A = \{(x, y) \in \mathcal{V}^2 \mid x + y \in V\}$$

is a neighbourhood of $(\vec{0}, \vec{0})$ in the product topology. Hence there exists sets $U_1, U_2 \in \mathcal{T}$ such that $(\vec{0}, \vec{0}) \in U_1 \times U_2 \subseteq A$. Let $U = U_1 \cap U_2$. Then $0 \in U$ and

$$U + U = \{x + y \mid x, y \in U\} \subseteq V.$$

Such open neighbourhoods are useful in many computations.

Remark 3.1.5. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $\alpha_0 \in \mathbb{K} \setminus \{0\}$ be fixed. Since the map $f : \mathcal{V} \to \mathcal{V}$ defined by $f(x) = \alpha_0 x$ is clearly seen to be a homeomorphism, we see that a subset $U \subseteq \mathcal{V}$ is a (open) neighbourhood of $\vec{0}$ if and and only if

$$\alpha_0 U = \{\alpha_0 u \mid u \in U\}$$

is a (open) neighbourhood of $\vec{0}$. Hence the neighbourhood basis in $(\mathcal{V}, \mathcal{T})$ at $\vec{0}$ is invariant under scaling.

Recall in a normed linear space $(\mathcal{V}, \|\cdot\|)$ that balls of the form $B_\epsilon(0)$ form a nice neighbourhood basis of $\vec{0}$. These balls are particularly nice when it comes to scaling. In particular, for all $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$, we know that $\alpha B_\epsilon(0) \subseteq B_\epsilon(0)$. As scaling is nice in topological vector spaces, it is natural to ask whether there are nice neighbourhoods of $\vec{0}$. To simplify these discussions, it is useful to give these types of sets a name.

Definition 3.1.6. Let $\mathcal{T}$ be a topology on a vector space $\mathcal{V}$. A neighbourhood U of $\vec{0}$ is said to be *balanced* if $\alpha U \subseteq U$ for all $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$.

Lemma 3.1.7. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. Every neighbourhood U of $\vec{0}$ contains a balanced neighbourhood V of $\vec{0}$ such that $V + V \subseteq U$.*

Proof. Let U be a neighbourhood of 0. By Remark 3.1.5 there exists a neighbourhood W of $\vec{0}$ such that $W + W \subseteq U$. To complete the proof, it suffices to find a balanced neighbourhood of $\vec{0}$ contained in W.

Since scalar multiplication is continuous, the set

$$A = \{(\alpha, x) \in \mathbb{K} \times \mathcal{V} \mid \alpha x \in W\}$$

is open in the product topology. Hence there exists an $\epsilon > 0$ and $V_0 \in \mathcal{T}$ such that $(0, 0) \in B_\epsilon(0) \times V_0 \subseteq A$. Hence

$$\{\alpha x \mid x \in V_0, 0 \leq |\alpha| < \epsilon\} \subseteq W.$$

Let

$$V = \bigcup_{0<|\alpha|<\epsilon} \alpha V_0.$$

By the above, we know that $V \subseteq W$ and, by Remark 3.1.5, we know that αV_0 is a neighbourhood of $\vec{0}$ for all $0 < |\alpha| < \epsilon$. Hence V is a union of open neighbourhoods of $\vec{0}$ and thus a neighbourhood of $\vec{0}$ contained in U. Thus it remains only to show that V is balanced.

Notice if $\alpha' \in \mathbb{K}$ is such that $|\alpha'| \leq 1$, then for all $0 < |\alpha| < \epsilon$ we have that $\alpha'\alpha V_0 \subseteq V$ as $|\alpha'\alpha| < \epsilon$ (and if $\alpha' = 0$, then $\alpha'\alpha V_0 = \{0\}$). Thus $\alpha' V \subseteq V$ for all $\alpha' \in \mathbb{K}$ such that $|\alpha'| \leq 1$. Hence V is balanced. ∎

In all of the above remarks, the Hausdorff property of topological vector spaces are not used. In particular, the above arguments allow us to prove the following that shows it suffices to show that points are closed when demonstrating the Hausdorff property for a potential topological vector space. We reminder the reader that points are closed in any Hausdorff topology.

Proposition 3.1.8. *Let $\mathcal{T}$ be a topology on a vector space $\mathcal{V}$ such that addition is continuous, scalar multiplication is continuous, and points in $\mathcal{V}$ are closed with respect to $\mathcal{T}$. Then $\mathcal{T}$ is a Hausdorff topology and thus $(\mathcal{V}, \mathcal{T})$ is a topological vector space.*

Proof. To see that $\mathcal{T}$ is Hausdorff, let $x, y \in \mathcal{V}$ be such that $x \neq y$. Since $\{x\}$ is a closed set, $U = \mathcal{V} \setminus \{x\}$ is a neighbourhood of y. Since addition is continuous, by the same proof as Remark 3.1.3 there exists a neighbourhood U_0 of $\vec{0}$ such that $U = y + U_0$. Furthermore, as addition is continuous, by the same proof as Remark 3.1.4 there exists a neighbourhood V of $\vec{0}$ such that $V + V \subseteq U_0$. Finally, as scalar multiplication is continuous, Remark 3.1.5 implies that $-V$ is a neighbourhood of $\vec{0}$. Therefore, if $W = V \cap (-V)$, then W is a neighbourhood of $\vec{0}$ such that $-W = W$ and $W + W \subseteq V + V \subseteq U_0$.

As addition is continuous, by the same proof as Remark 3.1.3 we obtain that $x + W$ and $y + W$ are neighbourhoods of x and y respectively. We claim that $(x + W) \cap (y + W) = \emptyset$. To see this, suppose to the contrary that there exists a $z \in (x + W) \cap (y + W)$. Therefore there exists $w_1, w_2 \in W$ such that $z = x + w_1 = y + w_2$. Thus

$$x = y + (w_2 - w_1) \in y + W + (-W) = y + (W + W) \in y + U_0 = U$$

which contradicts the fact that $x \notin U$. Hence $x + W$ and $y + W$ are disjoint neighbourhoods of x and y. Therefore, as x and y were arbitrary, $\mathcal{T}$ is Hausdorff. ∎

There is another property that open balls centred at $\vec{0}$ in normed linear spaces have. Indeed in a normed linear space $(\mathcal{V}, \|\cdot\|)$, if $\epsilon > 0$, then for all $x \in \mathcal{V}$ there exists a $C = \|x\| \frac{1}{\epsilon}$ such that for all $\alpha \in \mathbb{K}$ with $|\alpha| > C$ we have that $x \in \alpha B_\epsilon(0)$; that is, there is a neighbourhood basis of $\vec{0}$ that can be scaled to include every vector in the vector space. We encapsulate this property in the following definition.

Definition 3.1.9. Let $\mathcal{T}$ be a topology on a vector space $\mathcal{V}$. A neighbourhood U of $\vec{0}$ is said to be *absorbing* if for all $x \in \mathcal{V}$ there exists a $C \in \mathbb{R}$ such that for all $\alpha \in \mathbb{K}$ with $|\alpha| > C$ we have that $x \in \alpha U$.

Lemma 3.1.10. *Every neighbourhood of $\vec{0}$ in a topological vector space is absorbing.*

Proof. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. By the definition of absorbing and by Lemma 3.1.7, it suffices to show that if U is a balanced neighbourhood of $\vec{0}$, then U is absorbing.

To see this, let $x \in \mathcal{V}$ be arbitrary. Consider the function $f : \mathbb{K} \to \mathcal{V}$ defined by

$$f(\alpha) = \alpha x$$

for all $\alpha \in \mathbb{K}$. As f is continuous by the properties of a topological vector space, we know that $(f(\frac{1}{n}))_{n\geq 1}$ converges to $\vec{0}$ in $\mathcal{V}$. Hence there exists an $N \in \mathbb{N}$ such that $\frac{1}{n}x \in U$ for all $n \geq N$.

Let $\alpha \in \mathbb{K}$ be such that $|\alpha| \geq N$. To see that $x \in \alpha U$, note as $\frac{1}{N}x \in U$ and as U is balanced that $\frac{1}{\alpha}x = \frac{N}{\alpha}\left(\frac{1}{N}x\right) \in U$ as $\left|\frac{N}{\alpha}\right| \leq 1$. Thus as $x \in \alpha U$ as desired. Therefore, as x was arbitrary, U is absorbing. ∎

3.2 Generating Topological Vector Spaces

With our knowledge of the elementary properties of and requirements to have a topological vector space, we turn our attention to generating some examples of topological vector spaces. As it is not clear how to generate topologies that are not norm topologies but still have these properties, perhaps we should look at objects that are very close to being a norm.

Definition 3.2.1. Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. A *seminorm* on $\mathcal{V}$ is a function $p : \mathcal{V} \to [0, \infty)$ such that

1. $p(\alpha\vec{v}) = |\alpha|p(\vec{v})$ for all $\alpha \in \mathbb{K}$ and $\vec{v} \in \mathcal{V}$, and
2. $p(\vec{v} + \vec{w}) \leq p(\vec{v}) + p(\vec{w})$ for all $\vec{v}, \vec{w} \in \mathcal{V}$.

Remark 3.2.2. Note if p is a seminorm on a vector space $\mathcal{V}$, then the first property implies $p(\vec{0}) = p(0\vec{0}) = 0p(\vec{0}) = 0$. Thus the only difference between a norm and a seminorm is that a seminorm does not require if $p(\vec{v}) = 0$ then $\vec{v} = \vec{0}$. Of course we could mod out by all vectors in the null set of p and this would give us a norm on a quotient space and thus a normed linear space, but instead we would like to use seminorms on $\mathcal{V}$ to construct a topology on $\mathcal{V}$ that turns $\mathcal{V}$ into a topological vector space. Note we would need to mod out by the null set of p in order to ensure the Hausdorff property. As we do not want to mod out, perhaps we should look at multiple seminorms on vector spaces.

Of course, there are plenty of examples of seminorms that can be constructed using the objects discussed in previous chapters.

Example 3.2.3. Consider $\mathcal{V} = C(X)$ for some Hausdorff topological space $(X, \mathcal{T})$. For each $x \in X$, define $p_x : C(X) \to [0, \infty)$ by

$$p_x(f) = |f(x)|$$

for all $f \in C(X)$. Clearly $\{p_x \mid x \in X\}$ is a family of seminorms on $C(X)$.

Example 3.2.4. Consider $\mathcal{V} = C_0(X)$ for some locally compact Hausdorff topological space $(X, \mathcal{T})$. For each $K \subseteq X$ compact, define $p_K : C_0(X) \to [0, \infty)$ by

$$p_K(f) = \sup_{x \in K} |f(x)|$$

for all $f \in C_0(X)$. Clearly $\{p_K \mid K \subseteq X \text{ compact}\}$ is a family of seminorms on $C(X)$.

Example 3.2.5. Let $\mathcal{H} = \ell_2(\mathbb{N})$ and let $\mathcal{V} = \mathcal{B}(\mathcal{H})$. For each $h \in \mathcal{H}$, define $p_h : \mathcal{B}(\mathcal{H}) \to [0, \infty)$ by

$$p_h(T) = \|T(h)\|_2$$

for all $T \in \mathcal{B}(\mathcal{H})$. Clearly $\{p_h \mid h \in \mathcal{H}\}$ is a family of seminorms on $\mathcal{B}(\mathcal{H})$.

Example 3.2.6. Let $\mathcal{H} = \ell_2(\mathbb{N})$ and let $\mathcal{V} = \mathcal{B}(\mathcal{H})$. Define $\langle \cdot, \cdot \rangle : \mathcal{H}\times\mathcal{H} \to \mathbb{K}$ by

$$\langle (x_n)_{n\geq 1}, (y_n)_{n\geq 1}\rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

for all $(x_n)_{n\geq 1}, (y_n)_{n\geq 1} \in \mathcal{H}$. Note $\langle \cdot, \cdot \rangle$ is well-defined by Hölder's inequality.

For all $h, k \in \mathcal{H}$, define $p_{h,k} : \mathcal{B}(\mathcal{H}) \to [0, \infty)$ by

$$p_h(T) = \langle T(h), k\rangle$$

for all $T \in \mathcal{B}(\mathcal{H})$. Clearly $\{p_{h,k} \mid h, k \in \mathcal{H}\}$ is a family of seminorms on $\mathcal{B}(\mathcal{H})$.

Example 3.2.7. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. For all $f \in \mathcal{X}^*$, define $p_f : \mathcal{X} \to [0, \infty)$ by

$$p_f(x) = |f(x)|$$

for all $x \in \mathcal{X}$. Clearly $\{p_f \mid f \in \mathcal{X}^*\}$ is a family of seminorms on $\mathcal{X}$.

Example 3.2.8. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. For all $x \in \mathcal{X}$, define $p_x : \mathcal{X}^* \to [0, \infty)$ by

$$p_x(f) = |f(x)|$$

for all $f \in \mathcal{X}^*$. Clearly $\{p_x \mid x \in \mathcal{X}\}$ is a family of seminorms on $\mathcal{X}^*$.

Although all of these collections of seminorms are nice, in order to generate a topological vector space using seminorms, we will require the family of seminorm to have an additional property in order to ensure the topologies are Hausdorff.

Definition 3.2.9. A family $\mathcal{F}$ of seminorms on a vector space $\mathcal{V}$ is said to be *separating* if for all $\vec{v} \in \mathcal{V} \setminus \{\vec{0}\}$ there exists a $p \in \mathcal{F}$ such that $p(\vec{v}) \neq 0$.

Before demonstrating which of the above families of seminorms are separating, we demonstrate why separating families of seminorms yield normed linear space.

Theorem 3.2.10. *Let $\mathcal{V}$ be a vector space and let $\mathcal{F}$ be a separating family of seminorms on $\mathcal{V}$. For each $x \in \mathcal{V}$, $\epsilon > 0$, and $F \subseteq \mathcal{F}$ finite, let*

$$N(x, F, \epsilon) = \{y \in \mathcal{V} \mid p(y - x) < \epsilon \text{ for all } p \in F\}.$$

(When $F = \{p\}$ for some $p \in \mathcal{F}$, we will use $N(x, p, \epsilon)$ in place of $N(x, \{p\}, \epsilon)$).

Let

$$\mathcal{B} = \{N(x, F, \epsilon) \mid x \in \mathcal{V}, \epsilon > 0, F \subseteq \mathcal{F} \text{ finite}\}.$$

Then $\mathcal{B}$ is a basis for a topology $\mathcal{T}$ on $\mathcal{V}$ such that $(\mathcal{V}, \mathcal{T})$ is a topological vector space and every element of $\mathcal{F}$ is continuous with respect to $\mathcal{T}$. We call $\mathcal{T}$ the topology generated by the family of seminorms *$\mathcal{F}$.*

Proof. First we demonstrate that $\mathcal{B}$ is a basis for a topology on $\mathcal{V}$. To begin, clearly $\mathcal{B}$ covers $\mathcal{V}$. To demonstrate the other requirement of being a basis, let $x_1, x_2 \in \mathcal{V}$, let $\epsilon_1, \epsilon_2 > 0$, and let $F_1, F_2 \subseteq \mathcal{F}$ be finite sets such that $N(x_1, F_1, \epsilon_1) \cap N(x_2, F_2, \epsilon_2) \neq \emptyset$. Fix $y \in N(x_1, F_1, \epsilon_1) \cap N(x_2, F_2, \epsilon_2)$. Let $F = F_1 \cup F_2 \subseteq \mathcal{F}$, which is finite, and let

$$\epsilon = \min\left(\{\epsilon_1 - p(y - x_1) \mid p \in F_1\}, \{\epsilon_2 - p(y - x_2) \mid p \in F_2\}\right) > 0.$$

We claim that $N(y, F, \epsilon) \subseteq N(x_1, F_1, \epsilon_1) \cap N(x_2, F_2, \epsilon_2)$. To see this, let $z \in N(y, F, \epsilon)$ be arbitrary. Hence for all $p \in F_1$ we see that

$$p(z - x_1) \leq p(z - y) + p(y - x_1) < \epsilon + p(y - x_1) \leq \epsilon_1$$

so $z \in (x_1, F_1, \epsilon_1)$. Similarly, for all $p \in F_2$ we see that

$$p(z - x_2) \leq p(z - y) + p(y - x_2) < \epsilon + p(y - x_2) < \epsilon_2$$

so $z \in (x_2, F_2, \epsilon_2)$. Hence $N(y, F, \epsilon) \subseteq N(x_1, F_1, \epsilon_1) \cap N(x_2, F_2, \epsilon_2)$ thereby completing the proof that $\mathcal{B}$ is a basis.

To show that $(\mathcal{V}, \mathcal{T})$ is a topological vector space, we begin by showing that $\mathcal{T}$ is Hausdorff. To see this, let $x, y \in \mathcal{V}$ be such that $x \neq y$. As $\mathcal{F}$ is a separating family of seminorms, there exists a $p \in \mathcal{F}$ such that $\epsilon = \frac{1}{2}p(x - y) > 0$. Let $U = N(x, p, \epsilon)$ and let $V = N(y, p, \epsilon)$. Clearly U and V are $\mathcal{T}$-neighbourhoods of x and y by construction. To see that $U \cap V = \emptyset$, suppose to the contrary that there exists a $z \in U \cap V$. Hence $p(z - x) < \epsilon$ and $p(z - y) < \epsilon$ so

$$p(x - y) \leq p(x - z) + p(z - y) < \epsilon + \epsilon = p(x - y),$$

which is clearly a contradiction. Hence $\mathcal{T}$ is Hausdorff.

To see that addition is continuous in $(\mathcal{V}, \mathcal{T})$, let $x_0, y_0 \in \mathcal{V}$ and $U \in \mathcal{T}$ such that $x_0 + y_0 \in U$ be arbitrary. Since $\mathcal{B}$ is a basis for $\mathcal{T}$, there exists an $\epsilon > 0$ and a $F \subseteq \mathcal{F}$ finite such that $N(x_0 + y_0, F, \epsilon) \subseteq U$. Notice that $N(x_0, F, \frac{\epsilon}{2})$ and $N(y_0, F, \frac{\epsilon}{2})$ are neighbourhoods of x_0 and y_0 respectively. Moreover, notice if $x \in N(x_0, F, \frac{\epsilon}{2})$ and $y \in N(y_0, F, \frac{\epsilon}{2})$, then for all $p \in F$ we have that

$$p((x + y) - (x_0 + y_0)) \leq p(x - x_0) + p(y - y_0) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

so $x + y \in N(x_0 + y_0, F, \epsilon)$. Hence

$$N\left(x_0, F, \frac{\epsilon}{2}\right) + N\left(y_0, F, \frac{\epsilon}{2}\right) \subseteq N(x_0 + y_0, F, \epsilon) \subseteq U.$$

Therefore, as x_0, y_0, and U were arbitrary, addition is continuous in $(\mathcal{V}, \mathcal{T})$.

Finally, to see that scalar multiplication is continuous in $(\mathcal{V}, \mathcal{T})$, let $x_0 \in \mathcal{V}$, $\alpha_0 \in \mathbb{K}$, and $U \in \mathcal{T}$ such that $\alpha_0 x_0 \in U$ be arbitrary. Since $\mathcal{B}$ is a basis for

$\mathcal{T}$, there exists an $\epsilon > 0$ and a $F \subseteq \mathcal{F}$ finite such that $N(\alpha_0 x_0, F, \epsilon) \subseteq U$. Let

$$M = 1 + \max(\{p(x_0) \mid p \in F\}).$$

Notice that $B_{\mathbb{K}}\left(\alpha_0, \frac{\epsilon}{2(M+1)}\right)$ and $N\left(x_0, F, \frac{\epsilon}{2(|\alpha_0|+1)}\right)$ are neighbourhoods of α_0 and x_0 respectively. Moreover, notice if $\alpha \in B_{\mathbb{K}}\left(\alpha_0, \frac{\epsilon}{2(M+1)}\right)$ and $x \in N\left(x_0, F, \frac{\epsilon}{2(|\alpha_0|+1)}\right)$, then $|\alpha - \alpha_0| < \frac{\epsilon}{2(M+1)}$ so $|\alpha| \leq |\alpha_0| + \frac{\epsilon}{2(M+1)}$ and for all $p \in F$ we have that

$$\begin{aligned}
p(\alpha x - \alpha_0 x_0) &= p(\alpha(x - x_0) + (\alpha - \alpha_0)x_0) \\
&\leq |\alpha| p(x - x_0) + |\alpha - \alpha_0| p(x_0) \\
&< \left(|\alpha_0| + \frac{\epsilon}{2(M+1)}\right)\left(\frac{\epsilon}{2(|\alpha_0|+1)}\right) + \frac{\epsilon}{2(M+1)} M \\
&< \epsilon
\end{aligned}$$

so $\alpha x \in N(\alpha_0 x_0, F, \epsilon) \subseteq U$. Therefore, as α_0, x_0, and U were arbitrary, scalar multiplication is continuous in $(\mathcal{V}, \mathcal{T})$. ∎

Corollary 3.2.11. *Let $\mathcal{V}$ be a vector space and let $\mathcal{F}$ be a separating family of seminorms on $\mathcal{V}$. For each $x \in \mathcal{V}$, the collection*

$$\{N(x, F, \epsilon) \mid \epsilon > 0, F \subseteq \mathcal{F} \text{ finite}\}$$

is a neighbourhood basis of x.

Example 3.2.12. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. Then $\mathcal{F} = \{\|\cdot\|\}$ is a separating family of (semi)norms. If $\mathcal{T}$ is the topology generated by $\mathcal{F}$, then as

$$N(x, \|\cdot\|, \epsilon) = B(x, \epsilon)$$

for all $x \in \mathcal{X}$ and $\epsilon > 0$, we see that $\mathcal{T}$ is the norm topology as expected.

Before we examine more examples, it is useful to consider how convergence works in a topology generated by seminorms.

Proposition 3.2.13. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space generated by a separating family of seminorms $\mathcal{F}$ on $\mathcal{V}$. A net $(x_\lambda)_{\lambda \in \Lambda}$ converges to a point $x \in \mathcal{V}$ if and only if $\lim_{\lambda \in \Lambda} p(x_\lambda - x) = 0$ for all $p \in \mathcal{F}$.*

Proof. Suppose $(x_\lambda)_{\lambda \in \Lambda}$ is a net that converges to a point x in $\mathcal{V}$. Let $p \in \mathcal{F}$ and $\epsilon > 0$ be arbitrary. Since $N(x, p, \epsilon)$ is a neighbourhood of x and $(x_\lambda)_{\lambda \in \Lambda}$ converges to x, there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in N(x, p, \epsilon)$ for all $\lambda \geq \lambda_0$. By the definition of $N(x, p, \epsilon)$, this implies that $p(x_\lambda - x) < \epsilon$ for all $\lambda \geq \lambda_0$. Therefore, as ϵ and p were arbitrary, we obtain that $\lim_{\lambda \in \Lambda} p(x_\lambda - x) = 0$ for all $p \in \mathcal{F}$.

Conversely, suppose $\lim_{\lambda\in\Lambda} p(x_\lambda - x) = 0$ for all $p \in \mathcal{F}$. To see that $(x_\lambda)_{\lambda\in\Lambda}$ converges to x, let U be a neighbourhood of x. By the definition of $\mathcal{T}$, there exists an $\epsilon > 0$ and a finite subset $F \subseteq \mathcal{F}$ such that $N(x, F\epsilon) \subseteq U$. Since $\lim_{\lambda\in\Lambda} p(x_\lambda - x) = 0$ for all $p \in \mathcal{F}$, for every $p \in \mathcal{F}$ there exists a $\lambda_p \in \Lambda$ such that $p(x_\lambda - x) < \epsilon$ for all $\lambda \geq \lambda_p$. Due to the properties of directed sets, F being finite implies there exists a $\lambda_0 \in \Lambda$ such that $\lambda_0 \geq \lambda_p$ for all $p \in F$. Hence for all $\lambda \geq \lambda_0$ we obtain that $p(x_\lambda - x) < \epsilon$ for all $p \in F$ and thus $x_\lambda \in N(x, F, \epsilon) \subseteq U$. Therefore, as U was arbitrary, $(x_\lambda)_{\lambda\in\Lambda}$ converges to x. ∎

Example 3.2.14. Recall from Example 3.2.3 the family

$$\mathcal{F} = \{p_x \mid x \in X\}$$

of seminorms on $C(X)$ where $(X, \mathcal{T})$ is a Hausdorff topological space. Clearly $\mathcal{F}$ is a separating family of seminorms on $C(X)$ and thus generates a topology $\mathcal{T}_p$ on $C(X)$. Since Proposition 3.2.13 implies a net converges if and only if it converges pointwise, this is called the *pointwise convergence topology*.

Example 3.2.15. Recall from Example 3.2.4 the family

$$\mathcal{F} = \{p_K \mid K \subseteq X \text{ compact}\}$$

of seminorms on $C_0(X)$ where $(X, \mathcal{T})$ is a locally compact Hausdorff space. Clearly $\mathcal{F}$ is a separating family of seminorms on $C_0(X)$ (as singletons are compact) and thus generates a topology $\mathcal{T}_K$ on $C_0(X)$. Since Proposition 3.2.13 implies a net converges if and only if it converges uniformly on compact sets, this is called the *uniform convergence on compact sets topology*.

Example 3.2.16. With $\mathcal{H} = \mathcal{H} = \ell_2(\mathbb{N})$, recall from Example 3.2.5 the family $\mathcal{F} = \{p_h \mid h \in \mathcal{H}\}$ of seminorms on $\mathcal{B}(\mathcal{H}))$. Clearly $\mathcal{F}$ is a separating family of seminorms on $\mathcal{B}(\mathcal{H})$ and thus generates a topology $\mathcal{T}_{\mathrm{SOT}}$ on $\mathcal{B}(\mathcal{H})$. Note by Proposition 3.2.13 a net $(T_\lambda)_{\lambda\in\Lambda}$ converges to $T \in \mathcal{B}(\mathcal{H})$ if and only if $\lim_{\lambda\in\Lambda} T_\lambda(h) = T(h)$ for all $h \in \mathcal{H}$. This topology is called the *Strong Operator Topology* and is abbreviated by SOT.

Example 3.2.17. With $\mathcal{H} = \mathcal{H} = \ell_2(\mathbb{N})$ and $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{K}$ defined by

$$\langle (x_n)_{n\geq 1}, (y_n)_{n\geq 1} \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

for all $(x_n)_{n\geq 1}, (y_n)_{n\geq 1} \in \mathcal{H}$, recall from Example 3.2.6 the family $\mathcal{F} = \{p_{h,k} \mid h, k \in \mathcal{H}\}$ of seminorms on $\mathcal{B}(\mathcal{H}))$. Clearly $\mathcal{F}$ is a separating family of seminorms on $\mathcal{B}(\mathcal{H})$ (as $\langle T(h), T(h), \rangle = \|T(h)\|_2^2$ for all $h \in \mathcal{H}$) and thus generates a topology $\mathcal{T}_{\mathrm{WOT}}$ on $\mathcal{B}(\mathcal{H})$. Note by Proposition 3.2.13 a net $(T_\lambda)_{\lambda\in\Lambda}$ converges to $T \in \mathcal{B}(\mathcal{H})$ if and only if $\lim_{\lambda\in\Lambda} \langle T_\lambda(h), k\rangle = \langle T(h), k\rangle$ for all $h, k \in \mathcal{H}$. This topology is called the *Weak Operator Topology* and is abbreviated by WOT.

Example 3.2.18. For a normed linear space $(\mathcal{X}, \|\cdot\|)$, recall from Example 3.2.7 the family $\mathcal{F} = \{p_f \mid f \in \mathcal{X}^*\}$ of seminorms on $\mathcal{X}$. It is not at all clear why given an $x \in \mathcal{X} \setminus \{\vec{0}\}$ there exists an $f \in \mathcal{X}^*$ such that $f(x) \neq 0$. In fact, this is very similar to the question we encountered in Remark 1.6.6. Luckily we are approaching an answer (i.e. see Chapter 4). Thus, for now, we are unsure if $\mathcal{F}$ is a separating family of seminorms and thus generates a topology.

Example 3.2.19. For a normed linear space $(\mathcal{X}, \|\cdot\|)$, recall from Example 3.2.8 the family $\mathcal{F} = \{p_x \mid x \in \mathcal{X}\}$ of seminorms on $\mathcal{X}^*$. Clearly $\mathcal{F}$ is a separating family of seminorms on $\mathcal{X}^*$. Note by Proposition 3.2.13 a net $(f_\lambda)_{\lambda \in \Lambda}$ converges to $f \in \mathcal{X}^*$ if and only if $\lim_{\lambda \in \Lambda} f_\lambda(x) = f(x)$ for all $x \in \mathcal{X}$. This topology is called the *weak* topology* (weak because it is weaker than norm convergence in a topological sense).

Of course, we could investigate all of the individual properties each of the above distinct topologies have and their importance to functional analysis. Some of this will be done in Chapter 5 and some will be done via the assignments. For now, we focus on studying the properties that all of these topologies share. In particular, all of these topologies share a common property when it comes to the type of sets they have in their bases.

Definition 3.2.20. A subset C of a vector space $\mathcal{V}$ is said to be *convex* if for all $x, y \in C$ and $t \in [0, 1]$ we have that $tx + (1-t)y \in C$.

Indeed convex sets are well-behaved in topological vector spaces.

Lemma 3.2.21. *Let C be a convex subset of a topological vector space $(\mathcal{V}, \mathcal{T})$. Then the following hold:*

- *$x + C$ is convex for all $x \in \mathcal{V}$.*
- *$\overline{C}$ is convex.*
- *For all $r, s > 0$ we have $rC + sC = (r+s)C$.*
- *If $\mathcal{W}$ is another vector space and $T : \mathcal{V} \to \mathcal{W}$ is a linear map, then $T(C)$ is convex.*

Proof. To see that $x + C$ is convex for a fixed $x \in \mathcal{V}$, let $y, z \in x + C$ and $t \in (0, 1)$ be arbitrary. Then there exist $c_1, c_2 \in C$ such that $y = x + c_1$ and $z = x + c_2$. Hence

$$ty + (1-t)z = t(x + c_1) + (1-t)(x + c_2) = x + (tc_1 + (1-t)c_2).$$

Since $c_1, c_2 \in C$ and C is convex, $tc_1 + (1-t)c_2 \in C$ and thus $ty + (1-t)z \in x + C$. Therefore, as $y, z,$ and t were arbitrary, $x + C$ is convex.

To see that $\overline{C}$ is convex, let $x, y \in \overline{C}$. Thus there exists nets $(x_\lambda)_{\lambda\in\Lambda}$ and $(y_\lambda)_{\lambda\in\Lambda}$ of elements of C that converge to x and y respectively. Notice for all $t \in [0,1]$ that $(tx_\lambda + (1-t)y_\lambda)_{\lambda\in\Lambda}$ converges to $tx + (1-t)y$ as addition and scalar multiplication are continuous in $\mathcal{V}$. Therefore, as $tx_\lambda + (1-t)y_\lambda \in C$ for all $\lambda \in \Lambda$, we obtain that $tx + (1-t)y \in \overline{C}$ for all $t \in [0,1]$ and $x, y \in \overline{C}$. Hence $\overline{C}$ is convex.

Next let $r, s > 0$. Since $t = \frac{r}{r+s} \in (0,1)$ and $1 - t = \frac{s}{r+s}$, the convexity of C implies for all $x, y \in C$ that $\frac{r}{r+s}x + \frac{s}{r+s}y \in C$. Hence $rx + sy \in (r+s)C$ for all $x, y \in C$, so $rC + sC \subseteq (r+s)C$. As clearly $(r+s)C \subseteq rC + sC$, the second property holds.

Finally, to see that $T(C)$ is convex, let $w_1, w_2 \in T(C)$ and let $t \in [0,1]$ be arbitrary. Hence there exists $x_1, x_2 \in C$ such that $T(x_1) = w_1$ and $T(x_2) = w_2$. Since C is convex, we see that $tx_1 + (1-t)x_2 \in C$. Notice that

$$tw_1 + (1-t)w_2 = tT(x_1) + (1-t)T(x_2) = T(tx_1 + (1-t)x_2) \in T(C)$$

as T is linear. Therefore, as w_1, w_2, and t were arbitrary, $T(C)$ is convex. ∎

The commonality of all of the topologies considered in this section is the following.

Definition 3.2.22. A topological vector space $(\mathcal{V}, \mathcal{T})$ is said to be *locally convex* if there exists a basis for $\mathcal{T}$ consisting of convex sets.

As Lemma 3.2.21 shows it suffices to consider neighbourhoods of $\vec{0}$ when showing a topological vector space is locally convex, the following suffices to showing the topologies of this section are locally convex.

Theorem 3.2.23. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space generated by a separating family $\mathcal{F}$ of seminorms on $\mathcal{V}$. For each $\epsilon > 0$ and $F \subseteq \mathcal{F}$ finite, the set $N(0, F, \epsilon)$ is balanced and convex. Hence $(\mathcal{V}, \mathcal{T})$ is locally convex.*

Proof. To see that $N(0, F, \epsilon)$ is balanced, let $x \in N(0, F, \epsilon)$ and $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$ be arbitrary. Notice for all $p \in F$ that

$$p(\alpha x) = |\alpha| p(x) \leq p(x) < \epsilon$$

so $\alpha y \in N(0, F, \epsilon)$. Therefore, since x and α were arbitrary, $N(0, F, \epsilon)$ is balanced.

To complete the proof, by Corollary 3.2.11 it suffices to show that $N(x, F, \epsilon)$ is convex for all $x \in \mathcal{V}$, $\epsilon > 0$, and $F \subseteq \mathcal{F}$ finite. To see this, let $x_1, x_2 \in N(x, F, \epsilon)$ and $t \in [0,1]$ be arbitrary. Notice for all $p \in F$ that

$$\begin{aligned} p((tx_1 + (1-t)x_2) - x) &= p(t(x_1 - x) + (1-t)(x_2 - x)) \\ &\leq tp(x_1 - x) + (1-t)p(x_2 - x) \\ &< t\epsilon + (1-t)\epsilon = \epsilon \end{aligned}$$

so $tx_1 + (1-t)x_2 \in N(x, F, \epsilon)$. Therefore, since x_1, x_2, and t were arbitrary, $N(x, F, \epsilon)$ is convex. ∎

Corollary 3.2.24. *Every normed linear space is locally convex. In particular, every open ball in a normed linear space is convex.*

3.3 Constructing Topological Vector Spaces

With the above (locally convex) topological vector spaces that we have generated via seminorms given use a plethora of examples, we can turn our attention to constructing new topological vector spaces from old ones in the same ways we did for normed linear spaces: subspaces, products, and quotients.

Proposition 3.3.1. *Let $(\mathcal{V}, \mathcal{T})$ be a (locally convex) topological vector space and let $\mathcal{W}$ be a vector subspace of $\mathcal{V}$. Then $\mathcal{W}$ is a (locally convex) topological vector space with the subspace topology. Moreover $\overline{\mathcal{W}}$ is a vector subspace of $\mathcal{V}$ and thus a (locally convex) topological vector space with the subspace topology.*

Proof. As addition and scalar multiplication are continuous on $\mathcal{V}$, so too are the continuous on $\mathcal{W}$ equipped with the subspace topology. Hence, as the subspace topology of a Hausdoff topology is Hausdorff, $\mathcal{W}$ is a topological vector space with the subspace topology.

To see that $\overline{\mathcal{W}}$ is a vector subspace of $\mathcal{V}$, let $x, y \in \overline{\mathcal{W}}$ and $\alpha \in \mathbb{K}$ be arbitrary. As $x, y \in \overline{\mathcal{W}}$ there exists nets $(x_\lambda)_{\lambda\in\Lambda}$ and $(y_\lambda)_{\lambda\in\Lambda}$ in $\mathcal{W}$ that converge to x and y respectively (recall the lexicographic ordering on the product of two directed sets yields a directed set). As addition and scalar multiplication are continuous in $\mathcal{V}$, we know that $(x_\lambda + y_\lambda)_{\lambda\in\Lambda}$ and $(\alpha x_\lambda)_{\lambda\in\Lambda}$ are nets in $\mathcal{W}$ that converge to $x+y$ and αx respectively. Hence $x+y, \alpha x \in \overline{\mathcal{W}}$. Therefore, as x, y, and α were arbitrary, $\overline{\mathcal{W}}$ is a vector subspace of $\mathcal{V}$.

Finally, if $(\mathcal{V}, \mathcal{T})$ is locally convex, then so too are $\mathcal{W}$ and $\overline{\mathcal{W}}$ as the intersection of a subspace with a convex subset of $\mathcal{V}$ is convex. ∎

To discuss products, we refer the reader to the product topology constructed in Definition A.3.12.

Proposition 3.3.2. *Let I be a non-empty index set and for each $i \in I$ let $(\mathcal{V}_i, \mathcal{T}_i)$ be a (locally convex) topological vector space over $\mathbb{K}$. Then the product $\prod_{i\in I} \mathcal{V}_i$ is a (locally convex) topological vector space over $\mathbb{K}$ when equipped with the product topology.*

Proof. It is elementary to verify that $\prod_{i\in I} \mathcal{V}_i$ is a vector space over $\mathbb{K}$ with coordinate-wise addition and scalar multiplication. As the product of Hausdorff topologies is Hausdorff and as the product of continuous functions is continuous (see Theorem A.6.7), we obtain that $\prod_{i\in I} \mathcal{V}_i$ is a topological vector space. Finally, if each $(\mathcal{V}_i, \mathcal{T}_i)$ is locally convex, then as each $\mathcal{V}_i$ is a convex set and the product of convex sets is easily seen to be convex, we obtain that $\prod_{i\in I} \mathcal{V}_i$ is locally convex. ∎

In order to discuss quotients of topological vector spaces, we recall the following quotient topology, which is a simplification of Proposition A.7.8.

Definition 3.3.3. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space, let $\mathcal{W}$ be a vector subspace of $\mathcal{V}$, and let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the quotient map. The *quotient topology* on $\mathcal{V}/\mathcal{W}$ is

$$\mathcal{T}_q = \{A \subseteq \mathcal{V}/\mathcal{W} \mid q^{-1}(A) \in \mathcal{T}\}.$$

It is elementary to verify that $\mathcal{T}_q$ is indeed a topology and is the finest topology on $\mathcal{V}/\mathcal{W}$ such that $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is continuous; that is, if $U \in \mathcal{T}_q$ then $q^{-1}(U) \in \mathcal{T}$. The following shows more is true and should be compared with the Open Mapping Theorem (Theorem 2.4.2) of which we have seen several uses.

Lemma 3.3.4. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space, let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$, and let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the quotient map. If $U \in \mathcal{T}$ then $q(U)$ is open in the quotient topology.*

Proof. Let $U \in \mathcal{T}$ be arbitrary. As

$$q(U) = \{\vec{u} + \mathcal{W} \mid \vec{u} \in U\},$$

we know by the definition of the quotient topology that $q(U) \in \mathcal{T}_q$ if and only if

$$q^{-1}(q(U)) = \{\vec{u} + \vec{w} \mid \vec{u} \in U, \vec{w} \in \mathcal{W}\}$$

is open in $\mathcal{T}$. However, notice that

$$q^{-1}(q(U)) = \bigcup_{\vec{w} \in W} \vec{w} + U$$

is a union of open sets as $\vec{x} + U \in \mathcal{T}$ for all $\vec{x} \in \mathcal{V}$ by Remark 3.1.3. Hence $q^{-1}(q(U)) \in \mathcal{T}$ so $q(U)$ is open as desired. ∎

In topology, a function that maps open sets to open sets is called a *open map*. Thus the canonical vector space quotient map is open by Lemma 3.3.4. In fact, the quotient map is a quotient map in the sense of topology (see Definition A.7.18); that is, a surjective continuous map with the property that a subset of the codomain is open if and only if its inverse image is open.

Before we move on to showing the quotients of topological vector spaces are topological vector spaces, we note there is no difference when it comes to normed linear spaces.

Proposition 3.3.5. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. The quotient topology on $\mathcal{V}/\mathcal{W}$ is equal to the topology induced by the quotient norm from Theorem 1.3.3.*

Proof. Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the quotient map viewed as a map between normed linear spaces. To complete the proof, it suffices by Corollary A.7.21 to show that q is a quotient map in the topological sense.

To see that q is a quotient map, first we note that q is clearly surjective. Next let U be an arbitrary open subset of $\mathcal{V}/\mathcal{W}$. To see that $q^{-1}(U)$ is open in $\mathcal{V}$, let $\vec{v} \in q^{-1}(U)$ be arbitrary. Thus $\vec{v} + \mathcal{W} \in U$. Hence, as U is open in $\mathcal{V}/\mathcal{W}$, there exists an $\epsilon > 0$ so that if $\vec{v}_2 \in \mathcal{V}$ and $\|(\vec{v}_2 + \mathcal{W}) - (\vec{v} + \mathcal{W})\| < \epsilon$, then $\vec{v}_2 + \mathcal{W} \in U$. Therefore if $\vec{v}_2 \in \mathcal{V}$ and $\|\vec{v}_2 - \vec{v}\| < \epsilon$, then

$$\|(\vec{v}_2 + \mathcal{W}) - (\vec{v} + \mathcal{W})\| = \|(\vec{v}_2 - \vec{v}) + \mathcal{W}\| \leq \|\vec{v}_2 - \vec{v}\| < \epsilon$$

so $\vec{v}_2 + \mathcal{W} \in U$ and thus $\vec{v}_2 \in q^{-1}(U)$. Hence the ball of radius ϵ centred at $\vec{v}$ is contained in $q^{-1}(U)$. Therefore, as $\vec{v} \in q^{-1}(U)$ was arbitrary, $q^{-1}(U)$ is open in $\mathcal{V}$.

Finally, suppose U is an arbitrary subset of $\mathcal{V}/\mathcal{W}$ such that $q^{-1}(U)$ is open in $\mathcal{V}$. To see that U is open in $\mathcal{V}/\mathcal{W}$, let $\vec{u} + \mathcal{W} \in U$ be arbitrary. Hence $\vec{u} \in q^{-1}(U)$. Hence, as $q^{-1}(U)$ is open in $\mathcal{V}$ there exists an $\epsilon > 0$ so that if $\vec{v} \in \mathcal{V}$ and $\|\vec{v} - \vec{u}\| < \epsilon$ then $\vec{v} \in q^{-1}(U)$. We claim that the ball of radius ϵ centred at $\vec{u} + \mathcal{W}$ is contained in U. To see this, suppose $\vec{v} \in \mathcal{V}$ is such that

$$\|(\vec{v} - \vec{u}) + \mathcal{W}\| = \|(\vec{v} + \mathcal{W}) - (\vec{u} + \mathcal{W})\| < \epsilon.$$

Hence, by the definition of the quotient norm, there exists a $\vec{w} \in \mathcal{W}$ such that

$$\|(\vec{v} + \vec{w}) - \vec{u}\| = \|(\vec{v} - \vec{u}) + \vec{w}\| < \epsilon.$$

Hence $\vec{v} + \vec{w} \in q^{-1}(U)$ by the above computation so

$$\vec{v} + \mathcal{W} = (\vec{v} + \mathcal{W}) + (\vec{0} + \mathcal{W}) = (\vec{v} + \mathcal{W}) + (\vec{w} + \mathcal{W}) = (\vec{v} + \vec{w}) + \mathcal{W} = q(\vec{v} + \vec{w}) \in U$$

as desired. Therefore, since $\vec{u} + \mathcal{W} \in U$ was arbitrary, U is open in $\mathcal{V}/\mathcal{W}$. Hence q is a quotient map by the definition of a quotient map thereby yielding the proof. ∎

Proposition 3.3.6. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Then $\mathcal{V}/\mathcal{W}$ is a topological vector space when equipped with the quotient topology.*

Moreover, if $\mathcal{V}$ is locally convex, then $\mathcal{V}/\mathcal{W}$ is locally convex.

Proof. First, we claim that $\sigma : (\mathcal{V}/\mathcal{W}) \times (\mathcal{V}/\mathcal{W}) \to \mathcal{V}/\mathcal{W}$ defined by

$$\sigma(x + \mathcal{W}, y + \mathcal{W}) = (x + y) + \mathcal{W}$$

for all $x, y \in \mathcal{V}$ is continuous with respect to the quotient topology. To see this, let $x + \mathcal{W}, y + \mathcal{W} \in \mathcal{V}/\mathcal{W}$ be arbitrary and let U be an arbitrary neighbourhood of $(x + y) + \mathcal{W}$ in the quotient topology. By the definition of the quotient topology, $q^{-1}(U)$ is a neighbourhood of $x + y$ in $\mathcal{V}$. Therefore, as addition is continuous in $\mathcal{V}$, there exist neighbourhoods V_x and V_y of x and y respectively such that if $x' \in V_x$ and $y' \in V_y$ then $x' + y' \in q^{-1}(U)$.

Let $U_x = q(V_x)$ and $U_y = q(V_y)$ which are open neighbourhoods of $x + \mathcal{W}$ and $y + \mathcal{W}$ in the quotient topology by Lemma 3.3.4. Thus $U_x \times U_y$ is a neighbourhood of $(x + \mathcal{W}, y + \mathcal{W})$ in the product topology. Notice if $(x' + \mathcal{W}, y' + \mathcal{W}) \in U_x \times U_y$, then there exists $w_1, w_2 \in \mathcal{W}$ such that $x' + w_1 \in V_x$ and $y' + w_2 \in V_y$ and thus $(x' + w_1) + (y' + w_2) \in q^{-1}(U)$ so $(x' + \mathcal{W}) + (y' + \mathcal{W}) \in U$. Therefore σ is continuous.

Next we claim that $\rho : \mathbb{K} \times (\mathcal{V}/\mathcal{W}) \to \mathcal{V}/\mathcal{W}$ defined by

$$\rho(\alpha, x + \mathcal{W}) = \alpha x + \mathcal{W}$$

for all $x \in \mathcal{V}$ and $\alpha \in \mathbb{K}$ is continuous with respect to the quotient topology. To see this, let $\alpha \in \mathbb{K}$ and $x+\mathcal{W} \in \mathcal{V}/\mathcal{W}$ be arbitrary and let U be an arbitrary neighbourhood of $\alpha x + \mathcal{W}$ in the quotient topology. By the definition of the quotient topology, $q^{-1}(U)$ is a neighbourhood of α in $\mathcal{V}$. Therefore, as scalar multiplication is continuous in $\mathcal{V}$, there exist neighbourhoods V_α and V_x of α and x respectively such that if $\alpha' \in V_\alpha$ and $x' \in V_x$ then $\alpha' x' \in q^{-1}(U)$.

Let $U_x = q(V_x)$ which is an open neighbourhood of $x + \mathcal{W}$ in the quotient topology by Lemma 3.3.4. Thus $V_\alpha \times U_x$ is a neighbourhood of $(\alpha, x + \mathcal{W})$ in the product topology. Notice if $(\alpha', x' + \mathcal{W}) \in V_\alpha \times U_x$, then there exists a $w \in \mathcal{W}$ such that $x' + w \in V_x$ and thus $\alpha'(x' + w) \in q^{-1}(U)$ so $\alpha' x' + \mathcal{W} \in U$. Therefore ρ is continuous.

To complete the proof that $\mathcal{V}/\mathcal{W}$ is a topological vector space, note by Proposition 3.1.8 it suffices to show that points are closed in $\mathcal{V}/\mathcal{W}$. To see this, let $x \in \mathcal{V}$ be arbitrary. Notice that

$$q^{-1}(x + \mathcal{W}) = \{x + w \mid w \in \mathcal{W}\}$$

is the translation of the closed subspace $\mathcal{W}$ by x. Therefore, as translation by x is a homeomorphism, $q^{-1}(x + \mathcal{W})$ is a closed subset of $\mathcal{V}$. Hence $C = \mathcal{V} \setminus q^{-1}(x + \mathcal{W})$ is open in $\mathcal{V}$ so $q(C)$ is open in $\mathcal{V}/\mathcal{W}$ by Lemma 3.3.4. As

$$(\mathcal{V}/\mathcal{W}) \setminus q(C) = \{x + \mathcal{W}\},$$

we obtain that $\{x + \mathcal{W}\}$ is closed in $\mathcal{V}/\mathcal{W}$ as desired.

Finally, suppose that $\mathcal{V}$ is a locally convex topological vector space. To see that $\mathcal{V}/\mathcal{W}$ is locally convex, let $x + \mathcal{W} \in \mathcal{V}/\mathcal{W}$ and U a neighbourhood of $x + \mathcal{W}$ be arbitrary. Hence $q^{-1}(U)$ is a neighbourhood of x in $\mathcal{V}$. Since $\mathcal{V}$ is locally convex, there exists a convex $V \in \mathcal{T}$ such that $x \in V \subseteq q^{-1}(U)$. Therefore $x + \mathcal{W} \subseteq q(V) \subseteq U$. However, since q is both a linear and open map, $q(V)$ is convex and open by Lemma 3.2.21. Therefore, as x and U were arbitrary, $\mathcal{V}/\mathcal{W}$ is locally convex. ∎

3.4 Properties of Topological Vector Spaces

With the above constructions of topological vector spaces, our next goal is to generalize several of the properties we know for normed linear spaces to this setting. In particular, notions of completeness and continuous linear maps are essential to functional analysis. We begin with the generalization of a Cauchy sequence.

Definition 3.4.1. A net $(x_\lambda)_{\lambda\in\Lambda}$ in a topological vector space $(\mathcal{V}, \mathcal{T})$ is said to be *Cauchy* if for every neighbourhood U of $\vec{0}$ there exists a $\lambda_0 \in \Lambda$ such that $x_{\lambda_1} - x_{\lambda_2} \in U$ for all $\lambda_1, \lambda_2 \geq \lambda_0$.

Of course, like with normed linear spaces, the simplest example of Cauchy nets are convergent nets.

Example 3.4.2. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $(x_\lambda)_{\lambda\in\Lambda}$ be a net that converges to x. Then $(x_\lambda)_{\lambda\in\Lambda}$ is Cauchy. Indeed let U be any neighbourhood of $\vec{0}$. By Lemma 3.1.7 there exists a balanced neighbourhood V of $\vec{0}$ such that $V + V \subseteq U$. Since $x + V$ is a neighbourhood of x, there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in x + V$ for all $\lambda \geq \lambda_0$. Therefore, for all $\lambda_1, \lambda_2 \geq \lambda_0$ we obtain that

$$x_{\lambda_1} - x_{\lambda_2} \in (x + V) - (x + V) = V - V = V + V \subseteq U.$$

Therefore, as U was arbitrary, $(x_\lambda)_{\lambda\in\Lambda}$ is Cauchy.

The notion of a complete topological vector space now follows easily.

Definition 3.4.3. A subset A of a topological vector space $(\mathcal{V}, \mathcal{T})$ is said to be *complete* if every Cauchy net consisting of elements from A converges in $\mathcal{V}$ to an element in A.

The simplest example of complete topological vector spaces are Banach spaces. Indeed the following result shows the two notions of completeness for normed linear spaces coincide.

Proposition 3.4.4. *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space. Then $\mathcal{V}$ is complete as a normed linear space if and only if $\mathcal{V}$ is complete as a topological vector space.*

Proof. Suppose $\mathcal{V}$ is complete as a topological vector space. If $(x_n)_{n\geq 1}$ is a Cauchy sequence, then clearly $(x_n)_{n\geq 1}$ is a Cauchy net as the open balls of radius $\frac{1}{N}$ for $N \in \mathbb{N}$ form a neighbourhood basis of $\vec{0}$. Hence $(x_n)_{n\geq 1}$ converges being a Cauchy net in a complete topological vector space. Therefore, as $(x_n)_{n\geq 1}$ was arbitrary, $\mathcal{V}$ is complete as a normed linear space.

Conversely, suppose $\mathcal{V}$ is complete as a normed linear space. To see that $\mathcal{V}$ is complete as a topological vector space, let $(x_\lambda)_{\lambda\in\Lambda}$ be an arbitrary

Cauchy net. For each $n \in \mathbb{N}$, let $U_n = B_{\mathcal{V}}(0, \frac{1}{n})$. Choose $k_1 \in \Lambda$ such that if $\lambda_1, \lambda_2 \geq k_1$, then $x_{\lambda_1} - x_{\lambda_2} \in U_1$. By cofinality, there exists a $k_2 \in \Lambda$ such that $k_2 \geq k_1$ and if $\lambda_1, \lambda_2 \geq k_2$, then $x_{\lambda_1} - x_{\lambda_2} \in U_2$. By recursion, there exists $(k_n)_{n\geq 1} \subseteq \Lambda$ such that $k_n \leq k_{n+1}$ for all $n \in \mathbb{N}$ and $x_{\lambda_1} - x_{\lambda_2} \in U_n$ for all $\lambda_1, \lambda_2 \geq k_n$. These two properties together imply $(x_{k_n})_{n\geq 1}$ is a Cauchy sequence in $\mathcal{V}$ and thus converges to some $x \in \mathcal{V}$ as $\mathcal{V}$ is complete as a normed linear space.

We claim that $(x_\lambda)_{\lambda\in\Lambda}$ converges to x. To see this, let U be a neighbourhood of x and choose $N \in \mathbb{N}$ such that $B_{\mathcal{V}}(x, \frac{1}{N}) \subseteq U$. By the above construction, we know that $x_{\lambda_1} - x_{\lambda_2} \in U_{2N}$ for all $\lambda_1, \lambda_2 \geq k_{2N}$. Moreover, as $B_{\mathcal{V}}(x, \frac{1}{2N})$ is a neighbourhood of x and $(x_{k_n})_{n\geq 1}$ converges to x, there exists an $N_1 \in \mathbb{N}$ such that $x_{k_n} \in B_{\mathcal{V}}(x, \frac{1}{2N})$ for all $n \geq N_1$. Choose $N_0 \in \mathbb{N}$ such that $N_0 = \max\{2N, N_1\}$. Hence $k_{N_0} \geq k_{2N}$ and $N_0 \geq N_1$ so we obtain for all $\lambda \geq k_{N_0}$ that

$$x_\lambda - x = (x_\lambda - x_{k_{N_0}}) + (x_{k_{N_0}} - x) \in U_{2N} + U_{2N} = U_N.$$

Hence $x_\lambda \in x + U_N = B_{\mathcal{V}}(x, \frac{1}{N}) \subseteq U$ for all $\lambda \geq k_{N_0}$. Therefore, as U was arbitrary, $(x_\lambda)_{\lambda\in\Lambda}$ converges to x. Hence $\mathcal{V}$ is complete as a topological vector space. ∎

With respect to subsets, closed and completeness behave as one would expect.

Proposition 3.4.5. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $A \subseteq \mathcal{V}$. Then the following hold:*

- *If $\mathcal{V}$ is complete and A is closed in $\mathcal{V}$, then A is complete.*
- *If A is complete, then A is closed in $\mathcal{V}$.*

Proof. Suppose $\mathcal{V}$ is complete and A is closed in $\mathcal{V}$. To see that A is complete, let $(a_\lambda)_{\lambda\in\Lambda}$ be a Cauchy net in A. Hence $(a_\lambda)_{\lambda\in\Lambda}$ is a Cauchy net in $\mathcal{V}$ and thus converges to some $x \in \mathcal{V}$. However, as A is closed, we obtain that $x \in A$. Hence $(a_\lambda)_{\lambda\in\Lambda}$ converges to $x \in A$, so A is complete.

For the second part, suppose A is complete. To see that A is closed in $\mathcal{V}$, let $(a_\lambda)_{\lambda\in\Lambda}$ be a net in A that converges to some $x \in \mathcal{V}$. By Example 3.4.2 this implies $(a_\lambda)_{\lambda\in\Lambda}$ is a Cauchy net. Hence, since A is complete, $(a_\lambda)_{\lambda\in\Lambda}$ must converge to some element $a \in \mathcal{V}$. However, as $\mathcal{T}$ is Hausdorff, this implies $x = a \in A$. Hence A is closed in $\mathcal{V}$. ∎

With the above showing completeness for topological vector spaces behaves identically to normed linear spaces, we turn our attention to the stronger form of continuity: uniform continuity. It is easily seen that the following definition is a generalization of the notion of uniform continuous functions between two normed linear spaces to functions between two topological vector spaces.

Definition 3.4.6. Let $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ and $(\mathcal{W}, \mathcal{T}_{\mathcal{W}})$ be topological vector spaces and let $A \subseteq \mathcal{V}$. A function $f : A \to \mathcal{W}$ is said to be *uniformly continuous* if for all $\mathcal{T}_{\mathcal{W}}$-neighbourhoods W of $\vec{0}$ there exists a $\mathcal{T}_{\mathcal{V}}$ neighbourhood V of $\vec{0}$ such that $x, y \in A$ and $x - y \in V$ implies $f(x) - f(y) \in W$.

To see this is indeed a generalization from normed linear spaces to topological vector spaces, we note the following thereby providing us immediately with examples.

Proposition 3.4.7. *Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be normed linear spaces and let $f : \mathcal{V} \to \mathcal{W}$. Then f is a uniformly continuous function between normed linear spaces if and only if f is a uniformly continuous function between topological vector spaces.*

Proof. Suppose f is uniformly continuous as a function on normed linear spaces. To see that f is uniformly continuous as a function on topological vector spaces, let W be an arbitrary $\mathcal{T}_{\mathcal{W}}$ neighbourhood of $\vec{0}$. Thus there exists an $\epsilon > 0$ such that $B_{\mathcal{W}}(\vec{0}, \epsilon) \subseteq W$. As f is uniformly continuous as a function on normed linear spaces, there exists a $\delta > 0$ such that if $x, y \in A$ and $\|x - y\|_{\mathcal{V}} < \delta$, then $\|f(x) - f(y)\|_{\mathcal{W}} < \epsilon$. Hence, with $V = B_{\mathcal{V}}(\vec{0}, \delta)$, if $x, y \in A$ and $x - y \in V$ then $f(x) - f(y) \in W$. Therefore, as W was arbitrary, f is uniformly continuous as a function on topological vector spaces.

Conversely, suppose f is uniformly continuous as a function on topological vector spaces. To see that f is uniformly continuous as a function on normed linear spaces, let $\epsilon > 0$. Since $W = B_{\mathcal{W}}(\vec{0}, \epsilon)$ is a $\mathcal{T}_{\mathcal{W}}$-neighbourhood of $\vec{0}$ and as f is uniformly continuous as a function on topological vector spaces, there exists a $\mathcal{T}_{\mathcal{V}}$-neighbourhood V of $\vec{0}$ such that if $x, y \in A$ and $x - y \in V$ then $f(x) - f(y) \in W$. As V is a $\mathcal{T}_{\mathcal{V}}$-neighbourhood of $\vec{0}$, there exists a $\delta > 0$ such that $B_{\mathcal{V}}(\vec{0}, \delta) \subseteq V$. Hence if $x, y \in A$ and $\|x - y\|_{\mathcal{V}} < \delta$, then $\|f(x) - f(y)\|_{\mathcal{W}} < \epsilon$. Therefore, as $\epsilon > 0$ was arbitrary, f is uniformly continuous as a function on normed linear spaces. ■

Of course, for a function to be uniformly continuous, it must be continuous.

Proposition 3.4.8. *Let $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ and $(\mathcal{W}, \mathcal{T}_{\mathcal{W}})$ be topological vector spaces and let $A \subseteq \mathcal{V}$. If $f : A \to \mathcal{W}$ is uniformly continuous, then f is continuous.*

Proof. To see that f is continuous on A, fix $a \in A$ and let W be an arbitrary $\mathcal{T}_{\mathcal{W}}$-neighbourhood of $f(a)$. Since $(\mathcal{W}, \mathcal{T}_{\mathcal{W}})$ is a topological vector space, there exists a $\mathcal{T}_{\mathcal{W}}$-neighbourhood W_0 of $\vec{0}$ such that $W = f(a) + W_0$. As f is uniformly continuous, there exists a $\mathcal{T}_{\mathcal{V}}$-neighbourhood V_0 of $\vec{0}$ such that if $x, y \in A$ and $x - y \in V_0$ then $f(x) - f(y) \in W_0$. However, since $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ is a topological vector space, $V = a + V_0$ is a $\mathcal{T}_{\mathcal{V}}$-neighbourhood of a such that if $x \in V = a + V_0$ then $x - a \in V_0$ so $f(x) - f(a) \in W_0$ and thus $f(x) \in f(a) + W_0 = W$. Therefore, as a and W were arbitrary, f is continuous. ■

Of course there are functions on normed linear spaces that are continuous but not uniformly continuous, such as the function $f : (0, \infty) \to (0, \infty)$ defined by $f(x) = \frac{1}{x}$. However, for linear maps between topological vector spaces, we have the following which is our best analogue of Theorem 1.4.6 for topological vector spaces.

Theorem 3.4.9. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces and let $T : \mathcal{V} \to \mathcal{W}$ be linear. The following are equivalent:*

(1) T is uniformly continuous.

(2) there exists an $x_0 \in \mathcal{V}$ such that T is continuous at x_0.

Proof. By Proposition 3.4.8, if T is uniformly continuous, then T is continuous and hence (2) holds.

Conversely, suppose that there exists an $x_0 \in \mathcal{V}$ such that T is continuous at x_0. To see that T is uniformly continuous, let W_0 be an arbitrary $\mathcal{T}_\mathcal{W}$-neighbourhood of $\vec{0}$. Since $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ is a topological vector space, $W = T(x_0) + W_0$ is a $\mathcal{T}_\mathcal{W}$ neighbourhood of $T(x_0)$. Therefore, since f is continuous at x_0, there exists a $\mathcal{T}_\mathcal{V}$-neighbourhood V of x_0 such that if $x \in V$ then $T(x) \in W$. Since $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ is a topological vector space, there exists a $\mathcal{T}_\mathcal{V}$-neighbourhood V_0 of $\vec{0}$ such that $V = x_0 + V_0$. However, if $x, y \in \mathcal{V}$ are such that $x - y \in V_0$, then $(x - y) + x_0 \in V$ so

$$T(x) - T(y) + T(x_0) = T((x - y) + x_0) \in W = T(x_0) + W_0$$

so $T(x) - T(y) \in W_0$. Therefore, as W_0 was arbitrary, T is uniformly continuous. ∎

Recall from topology that continuous functions on compact sets are automatically uniformly continuous. For linear maps, it is convexity that replaces compactness.

Proposition 3.4.10. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces and let $T : \mathcal{V} \to \mathcal{W}$ be linear. If $C \subseteq \mathcal{V}$ is a balanced convex neighbourhood of $\vec{0}$ such that $T|_C$ is continuous at $\vec{0}$, then $T|_C$ is uniformly continuous.*

Proof. To see that $T|_C$ is uniformly continuous, let W be an arbitrary $\mathcal{T}_\mathcal{W}$-neighbourhood of $\vec{0}$. By Lemma 3.1.7, there exists a $\mathcal{T}_\mathcal{W}$-neighbourhood W_0 of $\vec{0}$ such that $W_0 + W_0 \subseteq W$. As $T|_C$ is continuous at $\vec{0}$, there exists a $\mathcal{T}_\mathcal{V}$-neighbourhood V of $\vec{0}$ such that if $x \in C \cap V$ then $T(x) \in W_0$. Again by Lemma 3.1.7 there exists a balanced $\mathcal{T}_\mathcal{V}$-neighbourhood V_0 of $\vec{0}$ such that $V_0 \subseteq V$.

Suppose $x, y \in C$ are such that $x - y \in V_0$. As C is balanced, $-y \in C$. Moreover, as C is convex, $\frac{1}{2}x + \frac{1}{2}(-y) \in C$. However, since V_0 is balanced, $\frac{1}{2}(x - y) \in V_0$. Therefore, as $\frac{1}{2}x - \frac{1}{2}y \in C \cap V_0 \subseteq C \cap V$, we obtain that

$$\frac{1}{2}T(x) - \frac{1}{2}T(y) = T\left(\frac{1}{2}x - \frac{1}{2}y\right) \in W_0.$$

Hence

$$T(x) - T(y) \in W_0 + W_0 \subseteq W.$$

Therefore, as W was arbitrary, $T|_C$ is uniformly continuous. ∎

To complete this section, by combining the notions of completeness and uniform continuity, we may demonstrate the following.

Proposition 3.4.11. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces, let $\mathcal{V}_0$ be a vector subspace of $\mathcal{V}$, and let $T_0 : \mathcal{V}_0 \to \mathcal{W}$ be a continuous linear map. If $\mathcal{W}$ is complete there exists a continuous linear map $T : \overline{\mathcal{V}_0} \to \mathcal{W}$ such that $T|_{\mathcal{V}_0} = T_0$.*

Proof. Exercise. ∎

Of course, for normed linear spaces, the operator norm is not increased.

Corollary 3.4.12. *Let $(\mathcal{V}, \|\cdot\|_\mathcal{V})$ and $(\mathcal{W}, \|\cdot\|_\mathcal{W})$ be normed linear spaces, let $\mathcal{V}_0$ be a vector subspace of $\mathcal{V}$, and let $T_0 : \mathcal{V}_0 \to \mathcal{W}$ be a bounded linear map. If $\mathcal{W}$ is complete there exists a bounded linear map $T : \overline{\mathcal{V}_0} \to \mathcal{W}$ such that $T|_{\mathcal{V}_0} = T_0$ and $\|T\| = \|T_0\|$.*

Proof. Exercise. ∎

3.5 Finite Dimensional Topological Vector Spaces

Now that we understand the similar properties objects in topological vector spaces have to normed linear spaces, we will examine two different collections of topological vector spaces to end this chapter. For the first, we consider the collection of finite dimensional topological vector spaces. In particular, although one may feel these are the simplest class of topological vector spaces, they are one of the most important due to the theoretical results of this chapter and the fact that every topological vector space contains a plethora of finite dimensional subspaces.

The main result of this section, Corollary 3.5.3, is that given any finite dimensional vector space $\mathcal{V}$, there is exactly one topology on $\mathcal{V}$ that turns $\mathcal{V}$ into a topological vector space. In particular, this will imply that all norms on $\mathcal{V}$ are equivalent! To prove this result, we begin with the one-dimensional case.

Lemma 3.5.1. *Let $(\mathcal{V}, \mathcal{T})$ be a one-dimensional topological vector space and let $\{e\}$ be a basis for $\mathcal{V}$. The map $T : \mathbb{K} \to \mathcal{V}$ defined by $T(\alpha) = \alpha e$ for all $\alpha \in \mathbb{K}$ is a homeomorphism.*

Proof. Since scalar multiplication is continuous in $(\mathcal{V}, \mathcal{T})$, we easily see that T is a bijective continuous linear map. Hence, to complete the proof, it suffices to prove that $T^{-1} : \mathcal{V} \to \mathbb{K}$ defined by

$$T^{-1}(\alpha e) = \alpha$$

for all $\alpha \in \mathbb{K}$ is continuous. As T^{-1} is clearly linear, it suffices to prove that T^{-1} is continuous at $\vec{0}$ (i.e. when $\alpha = 0$) by Theorem 3.4.9 and Proposition 3.4.8.

Let $(\alpha_\lambda)_{\lambda \in \Lambda}$ be a net in $\mathbb{K}$ such that $(\alpha_\lambda e)_{\lambda \in \Lambda}$ converges to $\vec{0}$. To see that $(\alpha_\lambda)_{\lambda \in \Lambda}$ converges to 0, let $\epsilon > 0$ be arbitrary. Since $\epsilon e \neq \vec{0}$ and since $\mathcal{V}$ is Hausdorff, there exist a neighbourhood U of $\vec{0}$ such that $\epsilon e \notin U$. By Lemma 3.1.7 there exists a balanced neighbourhood W of $\vec{0}$ such that $W \subseteq U$ and thus $\epsilon e \notin W$.

As $(\alpha_\lambda e)_{\lambda \in \Lambda}$ converges to $\vec{0}$, there exists a $\lambda_0 \in \Lambda$ such that $\alpha_\lambda e \in W$ for all $\lambda \geq \lambda_0$. Notice if there exists a $\lambda \geq \lambda_0$ such that $|\alpha_\lambda| \geq \epsilon$, then

$$\epsilon e = \left(\frac{\epsilon}{\alpha_\lambda}\right)(\alpha_\lambda e) \in W$$

as W is balanced, which is a contradiction. Hence $|\alpha_\lambda| < \epsilon$ for all $\lambda \geq \lambda_0$. Therefore, as $\epsilon > 0$ was arbitrary, $(\alpha_\lambda)_{\lambda \in \Lambda}$ converges to 0. ∎

We can upgrade Lemma 3.5.1 to all finite dimensional topological vector spaces via induction and considering quotients.

Theorem 3.5.2. *Let $\mathbb{K}^n$ be viewed as a topological vector space with topology induced by the infinity norm, let $(\mathcal{V}, \mathcal{T})$ be an n-dimensional topological vector space, and let $\{e_1, \ldots, e_n\}$ be a basis for $\mathcal{V}$. The map $T : \mathbb{K}^n \to \mathcal{V}$ defined by*

$$T((z_1, \ldots, z_n)) = \sum_{k=1}^{n} z_k e_k$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$ is a homeomorphism.

Proof. We will proceed by induction on n with the base case of $n = 1$ complete by Lemma 3.5.1. To proceed with the inductive step, suppose the result is true for all $(n-1)$-dimensional topological vector spaces and let $(\mathcal{V}, \mathcal{T})$ be an n-dimensional topological vector space and let $\{e_1, \ldots, e_n\}$ be a basis for $\mathcal{V}$. As the map $T : \mathbb{K}^n \to \mathcal{V}$ defined by

$$T((z_1, \ldots, z_n)) = \sum_{k=1}^{n} z_k e_k$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$ is a bijective linear map that is continuous as addition and scalar multiplication are continuous, it suffices to show that T^{-1} is continuous.

Let $\mathcal{W} = \mathrm{span}(\{e_1, e_2, \ldots, e_{n-1}\})$. By Proposition 3.3.1 we know that $\mathcal{W}$ is an $(n-1)$-dimensional topological vector space with the subspace topology. Thus by the inductive hypothesis, the map $T_1 : \mathcal{W} \to \mathbb{K}^{n-1}$ defined by

$$T_1\left(\sum_{k=1}^{n-1} z_k e_k\right) = (z_1, \ldots, z_{n-1})$$

for all $(z_1, \ldots, z_{n-1}) \in \mathbb{K}^{n-1}$ is a homeomorphism.

By Definition 3.4.1 it is easy to see that homeomorphisms provide a bijection between Cauchy nets and between convergence Cauchy nets. Therefore, as $\mathbb{K}^{n-1}$ is complete as a topological vector space by Proposition 3.4.4, we see that $\mathcal{W}$ is complete as a topological vector space with the subspace topology. This immediately implies that $\mathcal{W}$ is complete as a subset of $\mathcal{V}$ and thus closed by Proposition 3.4.5.

Since $\mathcal{W}$ is closed, the quotient space $\mathcal{V}/\mathcal{W}$ is a topological vector space by Proposition 3.3.6 and the map $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ defined by $q(x) = x + \mathcal{W}$ is a continuous map. As $\mathcal{V}/\mathcal{W}$ is one-dimensional with basis $\{e_n + \mathcal{W}\}$, the induction hypothesis implies there exists a continuous linear map $T_2 : \mathcal{V}/\mathcal{W} \to \mathbb{K}$ defined by

$$T_2(\alpha e_n + \mathcal{W}) = \alpha$$

is a homeomorphism.

Consider the map $S : \mathcal{V} \to \mathbb{K}^{n-1} \times \mathbb{K} = \mathbb{K}^n$ defined by

$$\begin{aligned} S\left(\left(\sum_{k=1}^{n} z_k e_k\right)\right) &= \left(T_1\left(\left(\sum_{k=1}^{n-1} z_k e_k\right)\right), (T_2 \circ q)\left(\sum_{k=1}^{n-1} z_k e_k\right)\right) \\ &= ((z_1, \ldots, z_{n-1}), z_n) = T^{-1}\left(\left(\sum_{k=1}^{n} z_k e_k\right)\right) \end{aligned}$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$. As T_1 and $T_2 \circ q$ are continuous linear maps, we obtain that S is a continuous linear map and thus T^{-1} is continuous as desired. Hence the result follows by the Principle of Mathematical Induction. ∎

Corollary 3.5.3. *Given any finite dimensional vector space $\mathcal{V}$, there is exactly one topology on $\mathcal{V}$ that makes $\mathcal{V}$ into a topological vector space.*

Proof. As the composition and inverse of homeomorphisms between topological spaces are homeomorphisms, Theorem 3.5.2 implies that any two topologies on $\mathcal{V}$ that make $\mathcal{V}$ into topological vector spaces are homeomorphic via the identity map and thus the same topology. ∎

This characterization of every finite dimensional topological vector space being homeomorphic to the infinity norm on $\mathbb{K}^n$ yields some more results.

Corollary 3.5.4. *Every finite dimensional topological vector space is complete. Hence every finite dimensional normed linear spaces is a Banach space.*

Proof. The proof of this result can be found in the inductive step of the proof of Theorem 3.5.2. ∎

Corollary 3.5.5. *Every finite dimensional subspace of a topological vector space is closed.*

Proof. Let $\mathcal{W}$ be a finite dimensional subspace of a topological vector space $(\mathcal{V}, \mathcal{T})$. By Corollary 3.5.4 $\mathcal{W}$ is complete as a topological vector space with the subspace topology. This immediately implies that $\mathcal{W}$ is complete as a subset of $\mathcal{V}$ and thus closed by Proposition 3.4.5. ∎

In fact, not only are finite dimensional subspaces closed, but when we add them to closed subspaces, the result remains a closed subspace.

Corollary 3.5.6. *Let $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ be a topological vector space. If $\mathcal{W}$ and $\mathcal{X}$ are closed subspaces of $\mathcal{V}$ such that $\mathcal{X}$ is finite dimensional, then $\mathcal{W} + \mathcal{X}$ is closed.*

Proof. Since $\mathcal{W}$ is a closed subspace of $\mathcal{V}$, Proposition 3.3.6 implies that $\mathcal{V}/\mathcal{W}$ is a topological vector space. Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the quotient map. Since $\mathcal{X}$ is a finite dimensional subspace of $\mathcal{V}$, $q(\mathcal{X})$ is a finite dimensional subspace of $\mathcal{V}/\mathcal{W}$ and thus closed by Corollary 3.5.5. By the definition of the quotient topology, the inverse image under q of a closed subset of $\mathcal{V}/\mathcal{W}$ is closed and thus

$$q^{-1}(q(\mathcal{X})) = \mathcal{X} + \mathcal{W}$$

is a closed subspace of $\mathcal{V}$. ∎

As linear maps are essential to functional analysis, we note the following corollary of Theorem 3.5.2.

Corollary 3.5.7. *Every linear map from a finite dimensional topological vector space into another topological vector space is continuous.*

Proof. Let $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ and $(\mathcal{W}, \mathcal{T}_{\mathcal{W}})$ be topological spaces with $\mathcal{V}$ finite dimensional with basis $\{e_1, \ldots, e_n\}$ and let $S : \mathcal{V} \to \mathcal{W}$ be a linear map. If $\mathbb{K}^n$ is equipped with the infinity norm, by Theorem 3.5.2 the linear map $T : \mathbb{K}^n \to \mathcal{X}$ defined by

$$T((z_1, \ldots, z_n)) = \sum_{k=1}^{n} z_k e_k$$

for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$ is a homeomorphism. Since

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

if it can be demonstrated that $S \circ T^{-1}$ is continuous, then S is a composition of continuous functions and thus will be continuous.

Let $R = S \circ T^{-1} : \mathbb{K}^n \to \mathcal{W}$ and for each $k \in \{1, \ldots, n\}$ let $y_k = R(e_k)$. Thus for all $(z_1, \ldots, z_n) \in \mathbb{K}^n$

$$R((z_1, \ldots, z_n)) = \sum_{k=1}^{n} z_k y_k.$$

As addition and scalar multiplication are continuous functions in $\mathcal{W}$, we obtain that R is continuous thereby completing the proof. ∎

In fact, finite dimensional topological spaces belong to another class of topological vector space we have already considered.

Corollary 3.5.8. *Every finite dimensional topological space is locally convex.*

Proof. Every finite dimensional topological space has a topology induced by a norm by Corollary 3.5.3 and thus is locally convex by Corollary 3.2.24. ∎

However, there is another characterization of finite dimensional topological spaces that is less apparent related to the following topological property.

Definition 3.5.9. A Hausdorff topology $\mathcal{T}$ on a set X is said to be *locally compact* if one of the following two equivalent conditions hold:

- For every $x \in X$ there exists a $U \in \mathcal{T}$ and a compact set $K \subseteq X$ such that $x \in U \subseteq K$.
- For every $x \in X$ and neighbourhood $U \in \mathcal{T}$ of x, there exists a $V \in \mathcal{T}$ such that $x \in V \subseteq \overline{V} \subseteq U$ and $\overline{V}$ is compact.

Remark 3.5.10. In respect to the definition of a locally compact topology, of course it is easier to verify the first definition as the second definition is stronger and, therefore, we will usually use the second definition when proving results. The first definition the usual definition from topology whereas the second is equivalent when we restrict to our attention to Hausdorff topologies; something we almost always do in functional analysis as we want our topologies to have unique limits. The equivalence between these two definitions can be obtained via the one-point compactification of a locally compact Hausdorff topology which places said topology as a topological subspace of a compact Hausdorff topology.

The relation between finite dimensions and local compactness in the context of topological vector spaces is quite nice.

Theorem 3.5.11. *A topological vector space $(\mathcal{V}, \mathcal{T})$ is locally compact if and only if $\mathcal{V}$ is finite dimensional.*

Proof. First suppose $\mathcal{V}$ is an n-dimensional topological vector space. Since $\mathcal{V}$ is homeomorphic to $(\mathbb{K}^n, \|\cdot\|_\infty)$ by Theorem 3.5.2 and since the notions of open and compact sets are preserved under homeomorphisms, it suffices to prove $(\mathbb{K}^n, \|\cdot\|_\infty)$ is locally compact. If $x \in \mathbb{K}^n$ and U is a neighbourhood of x, there exists an $\epsilon > 0$ such that $x \in B(x, \epsilon) \subseteq U$. Let $V = B\left(x, \frac{\epsilon}{2}\right)$. Then V is such that

$$x \in V \subseteq \overline{V} \subseteq B(x, \epsilon) \subseteq U$$

by basis metric space arguments. As $\overline{V}$ is a closed bounded set, $\overline{V}$ is compact in $\mathbb{K}^n$ by the Heine-Borel Theorem (Theorem A.8.25). Thus, as x and U were arbitrary, $\mathcal{V}$ is locally compact.

Conversely, suppose that $(\mathcal{V}, \mathcal{T})$ is a locally compact topological vector space. Let U be any neighbourhood of $\vec{0}$. Since $\mathcal{V}$ is locally compact, there exists a $V \in \mathcal{T}$ such that if $K = \overline{V}$ then K is compact and $\vec{0} \in V \subseteq K \subseteq U$. By Lemma 3.1.7, there exists a neighbourhood V_0 of $\vec{0}$ such that $V_0 + V_0 \subseteq V \subseteq K$.

Since $x + V_0$ is open in $\mathcal{V}$ for all $x \in V$, since K is compact, and since

$$K \subseteq \bigcup_{x \in \mathcal{V}} x + V_0,$$

there exists a finite number $x_1, x_2, \ldots, x_n \in \mathcal{V}$ such that

$$K \subseteq \bigcup_{k=1}^{n} x_k + V_0.$$

Let $\mathcal{W} = \mathrm{span}(\{x_1, \ldots, x_n\})$. Since $\mathcal{W}$ is a finite dimensional subspace of $\mathcal{V}$ and thus closed by Corollary 3.5.5, the quotient space $\mathcal{V}/\mathcal{W}$ is a topological vector space by Proposition 3.3.6 and the quotient map $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is continuous and open. Hence

$$q(K) \subseteq \bigcup_{k=1}^{n} q(x_k + V_0) = \bigcup_{k=1}^{n} q(V_0) = q(V_0) \subseteq q(K)$$

as $q(x_k) = 0$ for all $k \in \{1, \ldots, n\}$. Hence $q(V_0) = q(K)$. However, as $V_0 + V_0 \subseteq K$, we obtain that

$$2q(K) \subseteq q(K) + q(K) = q(V_0) + q(V_0) = q(V_0 + V_0) \subseteq q(K).$$

Hence induction implies that $2^m q(K) \subseteq q(K)$ for all $m \in \mathbb{N}$. However, since $q(V_0)$ is a neighbourhood of $\vec{0}$ as q is an open map and since every neighbourhood of $\vec{0}$ is absorbing by Lemma 3.1.10, we obtain that

$$q(K) = \bigcup_{m=1}^{\infty} 2^m q(K) = \bigcup_{m=1}^{\infty} 2^m q(V_0) = \mathcal{V}/\mathcal{W}.$$

Therefore, since K is compact and q is continuous, $\mathcal{V}/\mathcal{W}$ is compact.

Suppose $\mathcal{V}/\mathcal{W}$ is not the zero vector space. Thus there exists a vector $y \in \mathcal{V}$ such that $y + \mathcal{W} \neq \vec{0} + \mathcal{W}$. Therefore $\mathcal{X} = \mathrm{span}(\{y + \mathcal{W}\})$ is a one-dimensional subspace of $\mathcal{V}/\mathcal{W}$. Since $\mathcal{X}$ is closed in $\mathcal{V}/\mathcal{W}$ by Corollary 3.5.5 and $\mathcal{V}/\mathcal{W}$ is compact, $\mathcal{X}$ is compact. However $\mathcal{X}$ is homeomorphic to $\mathbb{K}$ by Theorem 3.5.2 thereby proving $\mathbb{K}$ is compact, a clear contradiction. Hence $\mathcal{V}/\mathcal{W}$ is the zero vector space implying that $\mathcal{V} = \mathcal{W}$ is finite dimensional as desired. ∎

Corollary 3.5.12. *The closed unit ball of a normed linear space $(\mathcal{V}, \|\cdot\|)$ is compact if and only if $\mathcal{V}$ is finite dimensional.*

Proof. First suppose $\mathcal{V}$ is finite dimensional. Thus $\mathcal{V}$ is locally compact by Theorem 3.5.11. Hence there exists a neighbourhood V of $\vec{0}$ such that $\overline{V}$ is compact. Since V is a neighbourhood of $\vec{0}$, there exists an $\epsilon > 0$ such that the closed ball $B[\vec{0}, \epsilon]$ is contained in V and thus in the compact set $\overline{V}$. Therefore $B[\vec{0}, \epsilon]$ is compact. Since $B[\vec{0}, 1] = \frac{1}{\epsilon}B[\vec{0}, \epsilon]$ and non-zero scalar multiplication is a homeomorphism, $B[\vec{0}, 1]$ is the homeomorphic image of a compact set and thus compact.

Conversely, suppose $B[\vec{0}, 1]$ is compact. As non-zero scalar multiplication and translation are homeomorphisms in topological vector spaces and as $rB[\vec{0}, 1] + x = B[x, r]$ for all $r > 0$ and $x \in \mathcal{V}$, we see that every closed ball in $\mathcal{V}$ is compact.

To see that $\mathcal{V}$ is finite dimensional, it suffices by Theorem 3.5.11 to show that $\mathcal{V}$ is locally compact. To see this, let $x \in \mathcal{V}$ and U be an arbitrary neighbourhood of x. Thus there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq U$. Note if $V = B\left(x, \frac{\epsilon}{2}\right)$, then $\overline{V} \subseteq B\left[x, \frac{\epsilon}{2}\right]$ and thus is compact being a closed subset of a compact set. Moreover

$$x \in V \subseteq \overline{V} \subseteq B\left[x, \frac{\epsilon}{2}\right] \subseteq B(x, \epsilon) \subseteq U.$$

Therefore, as x and U were arbitrary, $\mathcal{V}$ is locally compact. ∎

3.6 Locally Convex Topological Vector Spaces

We turn our attention now to the other collection of topological vector spaces we desire to study: locally convex topological vector spaces. Recall that all topological vector spaces from Section 3.2 were locally convex by Theorem 3.2.23 and all of the constructions of topological vector spaces from Section 3.3 when applied to locally convex topological vector spaces produce locally convex topological vector spaces. Thus we will begin this section by first providing a topological vector space that is not locally convex. Subsequently, our main goal of this section is to demonstrate that every locally convex topological vector spaces arises in via Theorem 3.2.10.

Before we demonstrate a topological vector space that is not locally convex, it is useful to examine be behaviour of convex sets in locally convex topological vector spaces. Thus we begin with the following method for generating convex sets.

Definition 3.6.1. The *convex hull* of a subset A of a vector space $\mathcal{V}$, denoted $\mathrm{conv}(A)$, is the set

$$\mathrm{conv}(A) = \left\{ \sum_{k=1}^{n} t_k x_k \,\middle|\, n \in \mathbb{N}, \{x_k\}_{k=1}^{n} \subseteq A, \{t_k\}_{k=1}^{n} \subseteq [0,1], \sum_{k=1}^{n} t_k = 1 \right\}$$

Example 3.6.2. In $\mathbb{R}^2$, it is not difficult to verify that if

$$A = \{(1,0), (0,1), (-1,0), (0,-1)\}$$

then $\mathrm{conv}(A)$ is the closed unit ball with respect to $\|\cdot\|_1$ and if

$$B = \{(1,1), (1,-1), (-1,1), (-1,-1)\}$$

then $\mathrm{conv}(B)$ is the closed unit ball with respect to $\|\cdot\|_\infty$.

Unsurprisingly, when dealing with convexity, the convex hull of a set is a well-behaved object.

Lemma 3.6.3. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $A \subseteq \mathcal{V}$. The following are true:*

1. $\mathrm{conv}(A)$ *is a convex set.*
2. *If A is open, then* $\mathrm{conv}(A)$ *is open.*
3. *If A is open and balanced, then* $\mathrm{conv}(A)$ *is open and balanced.*

Proof. The fact that the convex hull of a set is convex follows from a simple computation using the fact that if $\{t_k\}_{k=1}^n, \{s_j\}_{j=1}^m \subseteq [0,1]$ are such that $\sum_{k=1}^n t_k = \sum_{j=1}^m s_j = 1$, then if $\sum_{k=1}^n \sum_{j=1}^m t_k s_j = 1$.

Next suppose that A is open. To see that $\mathrm{conv}(A)$ is open, note that $\mathrm{conv}(A)$ is the union over all $n \in \mathbb{N}$, all $\{t_k\}_{k=1}^{n-1} \subseteq [0,1)$ and $t_n \in (0,1]$ such that $\sum_{k=1}^n t_k$, and all $x_1, \ldots, x_{n-1} \in A$ of the sets

$$\left(\sum_{k=1}^{n-1} t_k x_k\right) + t_n A.$$

Since addition and scalar multiplication by non-zero numbers are homeomorphisms and since A is open, each of these sets is open. Hence $\mathrm{conv}(A)$ is a union of open sets and thus is open.

Finally, to see that $\mathrm{conv}(A)$ is balanced when A is open and balanced follows easily from the description of $\mathrm{conv}(A)$ and the definition of a balanced set (i.e. for each $\sum_{k=1}^n t_k x_k$ and $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$, $\alpha \sum_{k=1}^n t_k x_k = \sum_{k=1}^n t_k(\alpha x_k)$ with $\alpha x_k \in A$ for all k). ∎

Before we get to our example of a non-locally convex topological vector space, some corollaries are in order.

Corollary 3.6.4. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. Every neighbourhood U of $\vec{0}$ contains a balanced, convex neighbourhood V of $\vec{0}$ such that $V + V \subseteq U$.*

Proof. Let U be a neighbourhood of $\vec{0}$. By Lemma 3.1.7 there exists a neighbourhood U_0 of $\vec{0}$ such that $U_0 + U_0 \subseteq U$. Since $(\mathcal{V}, \mathcal{T})$ is locally convex, there exists a convex neighbourhood W of $\vec{0}$ such that $W \subseteq U_0$. Lastly, again by Lemma 3.1.7 there exists a balanced neighbourhood W_0 of $\vec{0}$ such that $W_0 \subseteq W$.

Let $V = \mathrm{conv}(W_0)$. By Lemma 3.6.3 V is a balanced neighbourhood of $\vec{0}$. Since $W_0 \subseteq W$ and W is convex, we obtain that $V \subseteq W$ so $V \subseteq U_0 \subseteq U$ and $V + V \subseteq U_0 + U_0 \subseteq U$ thereby completing the proof. ∎

Corollary 3.6.5. *Every locally convex topological vector space admits a neighbourhood basis of $\vec{0}$ consisting of balanced, convex sets.*

Using the notion of convex hull in topological vector spaces, we can demonstrate the following topological vector space is not locally convex.

Example 3.6.6. Let $p \in (0, 1)$ and let

$$\ell_p(\mathbb{N}) = \left\{ (z_n)_{n\geq 1} \,\middle|\, \sum_{n=1}^{\infty} |z_k|^p < \infty \right\}.$$

Note $\ell_p(\mathbb{N})$ differs from our ℓ_p-normed linear spaces from previous sections as $p \in (0, 1)$.

For each $a \in [0, \infty)$, consider the function $f_a : [0, \infty) \to \mathbb{R}$ defined by

$$f_a(x) = x^p + a^p - (x + a)^p.$$

Since $f_a(0) = 0$ and f_a is differentiable with

$$f_a'(x) = px^{p-1} - p(x + a)^{p-1} = \frac{p}{x^{1-p}} - \frac{p}{(x + a)^{1-p}} > 0$$

as $1 - p > 0$, we obtain that $f_a(x) > 0$ for all $x \in [0, \infty)$. This shows that if $(z_n)_{n\geq 1}, (w_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ then $(z_n + w_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Since clearly $\alpha \in \mathbb{K}$ and $(z_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ implies $(\alpha z_n)_{n\geq 1} \in \ell_p(\mathbb{N})$, we obtain that $\ell_p(\mathbb{N})$ is a vector space with respect to coordinate-wise addition and scalar multiplication.

Define $q : \ell_p(\mathbb{N}) \to [0, \infty)$ by

$$q\left((z_n)_{n\geq 1}\right) = \sum_{n=1}^{\infty} |z_k|^p$$

for all $(z_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Notice by the above that

(1) $q\left((z_n)_{n\geq 1}\right) = 0$ if and only if $z_n = 0$ for all $n \in \mathbb{N}$,

(2) $q\left(\alpha(z_n)_{n\geq 1}\right) = |\alpha|^p q\left((z_n)_{n\geq 1}\right)$ for all $\alpha \in \mathbb{K}$ and $(z_n)_{n\geq 1} \in \ell_p(\mathbb{N})$, and

(3) $q\left((z_n + w_n)_{n\geq 1}\right) \leq q\left((z_n)_{n\geq 1}\right) + q\left((w_n)_{n\geq 1}\right)$ for all $(z_n)_{n\geq 1}, (w_n)_{n\geq 1} \in \ell_p(\mathbb{N})$.

Note there is no $\frac{1}{p}$ for otherwise the triangle inequality would fail! Thus q is not a seminorm in that (2) is not quite correct. However, the proof of Theorem 3.2.10 can easily be adapted (specifically the argument involving the fact that scalar multiplication is continuous - just replace $|\alpha|$ et al. with $|\alpha|^p$ in the appropriate places) to obtain that if for each $\vec{x} \in \ell_p(\mathbb{N})$ and $\epsilon > 0$ we let

$$N(\vec{x}, \epsilon) = \{\vec{y} \in \ell_p(\mathbb{N}) \mid q(\vec{y} - \vec{x}) < \epsilon\}$$

and

$$\mathcal{B} = \{N(\vec{x}, \epsilon) \mid \vec{x} \in \ell_p(\mathbb{N}), \epsilon > 0\},$$

then $\mathcal{B}$ is a basis for a topology $\mathcal{T}$ on $\ell_p(\mathbb{N})$ such that $(\ell_p(\mathbb{N}), \mathcal{T})$ is a topological vector space.

However, $\mathcal{T}$ is not a locally convex topology. To see this, suppose to that $(\ell_p(\mathbb{N}), \mathcal{T})$ is locally convex. Since

$$U = \{\vec{x} \in \ell_p(\mathbb{N}) \mid q(\vec{x}) < 1\}$$

is a neighbourhood of $\vec{0}$, locally convexity would imply there exists a convex neighbourhood V of $\vec{0}$ contained in U. Furthermore, as V is a neighbourhood of $\vec{0}$, this implies there exists an $\epsilon > 0$ such that

$$\{\vec{x} \in \ell_p(\mathbb{N}) \mid q(\vec{x}) < \epsilon\} \subseteq V \subseteq U.$$

Let $\delta = \left(\frac{\epsilon}{2}\right)^{\frac{1}{p}} > 0$ and for all $k \in \mathbb{N}$ let $\vec{x}_k \in \ell_p(\mathbb{N})$ be defined by

$$\vec{x}_k(n) = \begin{cases} \delta & \text{if } k = n \\ 0 & \text{if } k \neq n \end{cases}$$

(note clearly $\vec{x}_k \in \ell_p(\mathbb{N})$). As $q(\vec{x}_k) < \epsilon$, we obtain that

$$\vec{x}_k \in V \subseteq U$$

for all $k \in \mathbb{N}$ and thus

$$\mathrm{conv}(\{\vec{x}_k\}_{k \in \mathbb{N}}) \subseteq V \subseteq U$$

as V is convex. Therefore, for all $n \in \mathbb{N}$ we obtain that $\sum_{k=1}^{n} \frac{1}{n}\vec{x}_k \in V \subseteq U$ so

$$1 > q\left(\sum_{k=1}^{n} \frac{1}{n}\vec{x}_k\right) = \sum_{k=1}^{n} \frac{\delta^p}{n^p} = \delta^p n^{1-p}.$$

However, as $\delta > 0$ and $1 - p > 0$, the above inequality is impossible. Hence we have a contradiction so $(\ell_p(\mathbb{N}), \tau)$ is not locally convex.

Remark 3.6.7. Of course, given a measure space $(X, \mathcal{A}, \mu)$ and $p \in (0, 1)$, one may define $\mathcal{L}_p(X, \mu)$ to be

$$\mathcal{L}_p(X, \mu) = \left\{ f : X \to \mathbb{K} \,\middle|\, f \text{ measurable}, \int_X |f|^p \, d\mu < \infty \right\}$$

and $L_p(X, \mu)$ to be $\mathcal{L}_p(X, \mu)$ modulo the subspace of all functions equal to 0 μ-almost everywhere. By the same argument as used in Example 3.6.6, $L_p(X, \mu)$ a vector space and can be equipped with a topological vector space structure. Provided $\mathcal{A}$ contains a collection of pairwise disjoint sets with positive μ-values, a similar argument can be used to that $L_p(X, \mu)$ will not be locally convex.

With the above example of a topological vector space that is not locally convex, we turn our attention to characterizing all locally convex topological vector spaces as those constructed in Section 3.2. To do so, we need to construct seminorms based on the given locally convex topology. As a priori the functions constructed need not be seminorms and for uses in Chapter 4, we introduce the following notion.

Definition 3.6.8. Let $\mathcal{V}$ be a vector space. A function $p : \mathcal{V} \to \mathbb{R}$ is said to be a *sublinear functional* if

1. $p(x + y) \leq p(x) + p(y)$ for all $x, y \in \mathcal{V}$, and
2. $p(rx) = rp(x)$ for all $x \in \mathcal{V}$ and $r > 0$.

Example 3.6.9. Clearly every seminorm is an example of a sublinear functional whereas the identity map on $\mathbb{R}$ is a sublinear functional that is not a seminorm since it is not positive. Similarly, the map $p : \mathbb{R} \to [0, \infty)$ defined by

$$p(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

is an example of a sublinear functional that is not a seminorm (as $p(-1) = p((-1)1) \neq |-1|p(1)$).

The way locally convex topologies will give rise to sublinear functionals and, eventually, seminorms is via the following.

Definition 3.6.10. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let C be a convex neighbourhood of 0. The function $p_C : \mathcal{V} \to \mathbb{R}$ defined by

$$p_C(x) = \inf\{r \in (0, \infty) \mid x \in rC\}$$

for all $x \in \mathcal{V}$ is called the *Minkowski functional* (also called the *gauge functional*).

Remark 3.6.11. It is useful to note that although the Minkowski functional uses the word 'functional', the Minkowski functional is not a linear functional since, in particular, it never takes negative values. Moreover, it is not difficult to see that $p_C(x) < \infty$ for all $x \in \mathcal{V}$. Indeed, recall that if C is a neighbourhood of $\vec{0}$, then C is absorbing by Lemma 3.1.10. Hence for every $x \in \mathcal{V}$ there exists an $r \in (0, \infty)$ such that $x \in rC$ so $p_C(x) < \infty$.

It is also useful to note what $p_C(x)$ tells us about the values of $r \in (0, \infty)$ for which $x \in rC$. To see this, suppose $s \in (0, \infty)$ is such that $p_C(x) < s$. Thus there exists a $p_C(x) < t < s$ such that $x \in tC$. Hence there exists a $c \in C$ such that $x = tc$. However, since C is convex, for every $r \geq s$ we see that

$$x = tc = \frac{t}{r}(rc) + \left(1 - \frac{t}{r}\right)\vec{0} \in rC$$

as rC is convex (which is elementary to verify), $\vec{0}, rc \in rC$, and $0 \leq \frac{t}{r} \leq 1$. Thus if $p_C(x) < s$, then $x \in rC$ for all $r \geq s$.

The relations between the Minkowski functional, sublinear functionals, and seminorms is observed via the following result.

Proposition 3.6.12. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let C be a convex neighbourhood of 0. Then the Minkowski functional p_C is a sublinear functional on $\mathcal{V}$ such that*

$$C = \{x \in \mathcal{V} \mid p_C(x) < 1\}.$$

Moreover p_C is a seminorm whenever C is balanced.

Proof. To see that p_C is sublinear, let $x, y \in \mathcal{V}$ be arbitrary. If $r, s \in (0, \infty)$ are such that $x \in rC$ and $y \in sC$, then

$$x + y \in rC + sC \subseteq (r + s)C$$

by Lemma 3.2.21 so $p_C(x + y) \leq r + s$. As this holds for all $r, s \in (0, \infty)$ such that $x \in rC$ and $y \in sC$, we obtain that $p_C(x + y) \leq p_C(x) + p_C(y)$ thereby demonstrating the first property of being a sublinear functional.

To see the second property of being a sublinear functional, let $x \in \mathcal{V}$ and $s \in (0, \infty)$ be arbitrary. For any $r \in (0, \infty)$ we see that $sx \in sC$ if and only if $x \in C$, we obtain that $p_C(sx) = sp_C(x)$ thereby completing the proof that p_C is sublinear.

To see that $C = \{x \in \mathcal{V} \mid p_C(x) < 1\}$, first suppose $x \in \mathcal{V}$ is such that $p_C(x) < 1$. Hence there exists an $r \in (0, 1)$ such that $x \in rC$. Thus there exists a $y \in C$ such that

$$x = ry = ry + (1 - r)\vec{0}.$$

However, as $r \in (0, 1)$, $y, \vec{0} \in C$, and C is convex, we obtain that $x \in C$. Hence $C \supseteq \{x \in \mathcal{V} \mid p_C(x) < 1\}$.

To see the reverse inclusion, let $x \in C$ be arbitrary. Since C is open and scalar multiplication is continuous, there exists a $\epsilon > 0$ such that if $|t-1| < \epsilon$ then $tx \in C$. Thus $\left(1 + \frac{\epsilon}{2}\right) x \in C$ so $x \in \frac{1}{1+\frac{\epsilon}{2}} C$ and hence $p_C(x) \leq \frac{1}{1+\frac{\epsilon}{2}} < 1$. Therefore, as x was arbitrary, $C = \{x \in \mathcal{V} \mid p_C(x) < 1\}$.

To see that p_C is a seminorm when C is balanced, we note $p_C : \mathcal{V} \to [0, \infty)$ and p_C is sublinear so it suffices to examine how p_C behaves with respect to scalar multiplication. Thus let $x \in V$ and $\alpha \in \mathbb{C}$ be arbitrary. If $\alpha = 0$, then clearly $p_C(\alpha x) = p_C(\vec{0}) = 0 = \alpha p_C(x)$. Otherwise, if $\alpha \neq 0$, then

$$\begin{aligned}
p_C(\alpha x) &= \inf\{r \in (0, \infty) \mid \alpha x \in rC\} \\
&= \inf\left\{ r \in (0, \infty) \,\middle|\, x \in r\frac{1}{\alpha}C \right\} \\
&= \inf\left\{ r \in (0, \infty) \,\middle|\, x \in r\frac{1}{|\alpha|}C \right\} && \text{as } C = \frac{\alpha}{|\alpha|}C \text{ since } C \text{ is balanced} \\
&= \inf\{|\alpha| s \in (0, \infty) \mid x \in sC\} \\
&= |\alpha| \inf\{s \in (0, \infty) \mid x \in sC\} \\
&= |\alpha| p_C(x)
\end{aligned}$$

as desired. Hence p_C is a seminorm. ∎

With the above construction of seminorms in locally convex topological vector spaces complete, we need only one more ingredient in order to show the locally convex topological is generated by seminorms: we need to know which seminorms are continuous.

Proposition 3.6.13. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let p be a seminorm on $\mathcal{V}$. Then p is continuous on $\mathcal{V}$ if and only if there exists a neighbourhood U of $\vec{0}$ such that p is bounded on U.*

Proof. Suppose p is continuous on $\mathcal{V}$. Therefore $U = \{x \in \mathcal{V} \mid p(x) < 1\}$ is an open neighbourhood of $\vec{0}$ such that p is bounded by 1 on U. Thus one direction is complete.

For the converse, suppose there exists a neighbourhood U of $\vec{0}$ such that p is bounded on U by $M \in (0, \infty)$. To see that p is continuous, let $x \in \mathcal{V}$ and $\epsilon > 0$ be arbitrary. Then $V = x + \frac{\epsilon}{M} U$ is a neighbourhood of x. Moreover, if $y \in V$ then $y - x = \frac{\epsilon}{M} u$ for some $u \in U$ so

$$|p(y) - p(x)| \leq p(y - x) = p\left(\frac{\epsilon}{M} u\right) = \frac{\epsilon}{M} p(u) \leq \epsilon.$$

Therefore, as x and $\epsilon > 0$ were arbitrary, p is continuous. ∎

Corollary 3.6.14. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let C be a convex balanced neighbourhood of $\vec{0}$. Then p_C is a continuous seminorm on $(\mathcal{V}, \mathcal{T})$.*

Proof. By Proposition 3.6.12, p_C is a seminorm on $\mathcal{V}$ and

$$C = \{x \in \mathcal{V} \mid p_C(x) < 1\}.$$

As C is a neighbourhood of $\vec{0}$, the above set equality implies that p_C is bounded by 1 on a neighbourhood of $\vec{0}$ and thus is continuous by Proposition 3.6.13. ∎

Theorem 3.6.15. *If $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space, then there exists a separating family of seminorms on $\mathcal{V}$ that generated $\mathcal{T}$.*

Proof. By Corollary 3.6.5 there exists a $\mathcal{T}$-neighbourhood basis $\mathcal{C}$ of $\vec{0}$ consisting of convex balanced sets. Therefore

$$\mathcal{F} = \{p_C \mid C \in \mathcal{C}\}$$

is a collection of seminorms on $\mathcal{V}$ by Proposition 3.6.12.

To see that $\mathcal{F}$ is separating, let $x \in \mathcal{V} \setminus \{\vec{0}\}$ be arbitrary. Since $(\mathcal{V}, \mathcal{T})$ is Hausdorff, there exists a neighbourhood U of $\vec{0}$ such that $x \notin U$. By Corollary 3.6.5 there exists a $C \in \mathcal{C}$ such that $C \subseteq U$ so $x \notin C$. Hence $p_C(x) \geq 1$ by Remark 3.6.11. Therefore, as x was arbitrary, $\mathcal{C}$ is separating.

Let $\mathcal{T}_0$ be the topology on $\mathcal{V}$ generated by $\mathcal{F}$. Thus $(\mathcal{V}, \mathcal{T}_0)$ is a topological vector space by Theorem 3.2.10 and $\mathcal{T}_0$ is a locally convex topology by Theorem 3.2.23. It remains only to show that $\mathcal{T} = \mathcal{T}_0$.

Notice if $C \in \mathcal{C}$, then by Proposition 3.6.12

$$C = \{x \in \mathcal{V} \mid p_C(x) < 1\} = N(\vec{0}, p, 1) \in \mathcal{T}_0$$

by definitions. Therefore a neighbourhood basis for $\vec{0}$ from $\mathcal{T}$ is contained in $\mathcal{T}_0$. Hence, since $\mathcal{T}$ and $\mathcal{T}_0$ make $\mathcal{V}$ into a topological vector space and thus are completely defined by any neighbourhood basis of $\vec{0}$, we obtain that $\mathcal{T} \subseteq \mathcal{T}_0$.

For the reverse inclusion, note by Corollary 3.6.14 that p_C is continuous on $(\mathcal{V}, \mathcal{T})$ for all $C \in \mathcal{C}$. Hence for every $\epsilon > 0$, the set

$$N(\vec{0}, p, \epsilon) = p_C^{-1}((-\epsilon, \epsilon)) \in \mathcal{T}.$$

Hence, for all $\epsilon > 0$ and $F \subseteq \mathcal{F}$ finite, we have that

$$N(\vec{0}, F, \epsilon) = \bigcap_{p \in F} N(\vec{0}, p, \epsilon) \in \mathcal{T}.$$

Therefore a neighbourhood basis for $\vec{0}$ from $\mathcal{T}_0$ is contained in $\mathcal{T}$. Hence, since $\mathcal{T}$ and $\mathcal{T}_0$ make $\mathcal{V}$ into a topological vector space and thus are completely defined by any neighbourhood basis of $\vec{0}$, we obtain that $\mathcal{T}_0 \subseteq \mathcal{T}$ completing the proof. ∎

To conclude this section, we note that there are many collections of seminorms that could generated the same topological vector space structure. As all such seminorms must be continuous, the following result not only aids in determining which seminorms must be included, but has a few corollaries that will be of use in future chapters.

Proposition 3.6.16. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{F}$ be a separating family of seminorms that generated $\mathcal{T}$. If p is a seminorm on $\mathcal{V}$, then the following are equivalent:*

- *p is continuous on $\mathcal{V}$.*
- *There exists a constant $M > 0$ and $p_1, \ldots, p_n \in \mathcal{F}$ such that*

$$p(x) \leq M \max(\{p_1(x), \ldots, p_n(x)\}) \qquad \text{for all } x \in \mathcal{V}.$$

Proof. First suppose there exists a constant $M > 0$ and $p_1, \ldots, p_n \in \mathcal{F}$ such that

$$p(x) \leq M \max(\{p_1(x), \ldots, p_n(x)\}) \qquad \text{for all } x \in \mathcal{V}.$$

Let $U = N(\vec{0}, \{p_1, \ldots, p_n\}, 1)$. Then $U \in \mathcal{T}$ and $p_k(x) \leq 1$ for all $x \in U$ and $k \in \{1, \ldots, n\}$. Hence $p(x) \leq M$ for all $x \in U$ so p is bounded on U and thus continuous by Proposition 3.6.13.

Conversely, suppose p is continuous on $(\mathcal{V}, \mathcal{T})$. Hence $U = p^{-1}((-1, 1))$ is a neighbourhood of $\vec{0}$. Since $\mathcal{F}$ generated $\mathcal{T}$, there exists $p_1, \ldots, p_n \in \mathcal{F}$ and an $\epsilon > 0$ such that

$$N(\vec{0}, \{p_1, \ldots, p_n\}, \epsilon) \subseteq U.$$

Therefore, if $x \in \mathcal{V}$ is such that $p_k(x) < \epsilon$ for all $k \in \{1, \ldots, n\}$, then $x \in N(\vec{0}, \{p_1, \ldots, p_n\}, \epsilon) \subseteq U$ so $p(x) < 1$.

We claim if $M = \max\{1, \frac{2}{\epsilon}\}$, then

$$p(x) \leq M \max(\{p_1(x), \ldots, p_n(x)\})$$

for all $x \in \mathcal{V}$. To see this, let $x \in \mathcal{V}$ be arbitrary. If $\max(\{p_1(x), \ldots, p_n(x)\}) = 0$, then $p_k(x) = 0$ for all $k \in \{1, \ldots, n\}$ so $p_k(rx) = 0$ for all $r > 0$ thereby implying $rp(x) = p(rx) < 1$ for all $r > 0$, which implies $p(x) = 0 \leq 1 \max(\{p_1(x), \ldots, p_n(x)\})$. Otherwise, if $M_0 = \max(\{p_1(x), \ldots, p_n(x)\}) > 0$, then $y = \frac{\epsilon}{2M_0} x \in \mathcal{V}$ has the property that

$$p_k(y) \leq \frac{\epsilon}{2} < \epsilon$$

for all $k \in \{1, \ldots, n\}$. This implies $p(y) < 1$ and thus

$$p(x) < \frac{2M_0}{\epsilon} = \frac{2}{\epsilon} \max(\{p_1(x), \ldots, p_n(x)\}) \leq M \max(\{p_1(x), \ldots, p_n(x)\}).$$

Therefore, as x was arbitrary, the proof is complete. ∎

Corollary 3.6.17. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be locally convex topological vector spaces, let $\mathcal{F}_\mathcal{V}$ and $\mathcal{F}_\mathcal{W}$ be separating family of seminorms that generated $\mathcal{T}_\mathcal{V}$ and $\mathcal{T}_\mathcal{W}$ respectively, and let $T : \mathcal{V} \to \mathcal{W}$ be linear. The following are equivalent:*

- *T is continuous.*
- *For all $q \in \mathcal{F}_\mathcal{W}$ there exists $M > 0$ and $p_1, \ldots, p_n \in \mathcal{F}_\mathcal{V}$ such that*

$$q(T(x)) \leq M \max(\{p_1(x), \ldots, p_n(x)\}) \qquad \text{for all } x \in \mathcal{V}.$$

Proof. Note it is elementary to verify the composition of a seminorm with a linear map is a seminorm.

Suppose T is continuous. Since q is a continuous seminorm on $(\mathcal{W}, \mathcal{T}_\mathcal{W})$, $q \circ T$ is a continuous seminorm on $(\mathcal{V}, \mathcal{T}_\mathcal{V})$. Hence the result follows from Proposition 3.6.16.

To prove the converse, note Proposition 3.6.16 implies that $q \circ T$ is a continuous seminorm on $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ for all $q \in \mathcal{F}_\mathcal{W}$. To see that T is continuous, let U be an arbitrary neighbourhood of $\vec{0}$ in $(\mathcal{W}, \mathcal{T}_\mathcal{W})$. Hence there exists an $\epsilon > 0$ and $q_1, \ldots, q_n \in \mathcal{F}_\mathcal{W}$ such that

$$N(\vec{0}, \{q_1, \ldots, q_n\}, \epsilon) \subseteq U.$$

Since $q_k \circ T$ is a continuous seminorm on $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ for all $k \in \{1, \ldots, n\}$,

$$V = N(\vec{0}, \{q_1 \circ T, \ldots, q_n \circ T\}, \epsilon)$$

is a neighbourhood of $\vec{0}$ in $(\mathcal{V}, \mathcal{T}_\mathcal{V})$. Notice if $x \in V$ then $q_k(T(x)) < \epsilon$ for all $k \in \{1, \ldots, n\}$ so $T(x) \in N(\vec{0}, \{q_1, \ldots, q_n\}, \epsilon) \subseteq U$. Therefore, as U was arbitrary, T is continuous at $\vec{0}$ and thus continuous by Theorem 3.4.9. ∎

Corollary 3.6.18. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. A linear map $f : \mathcal{V} \to \mathbb{K}$ is continuous if and only if there exists a continuous seminorm p on $\mathcal{V}$ such that*

$$|f(x)| \leq p(x)$$

for all $x \in \mathcal{V}$.

Proof. If $f : \mathcal{V} \to \mathbb{K}$ is a continuous linear map, then the function $p : \mathcal{V} \to \mathbb{K}$ defined by

$$p(x) = |f(x)|$$

for all $x \in \mathcal{V}$ is clearly a continuous seminorm on $\mathcal{V}$. Hence one direction is complete.

Conversely, suppose there exists a continuous seminorm p on $\mathcal{V}$ such that

$$|f(x)| \leq p(x)$$

for all $x \in \mathcal{V}$. As $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space, Theorem 3.6.15 implies $\mathcal{T}$ is generated by a separating family of seminorms $\mathcal{F}$. As p is continuous, $\mathcal{T}$ is also generated by $\mathcal{F} \cup \{p\}$ and thus f is continuous by Corollary 3.6.17. ∎

One final corollary of Proposition 3.6.16 is the following which informs us how a separating family of seminorms on a locally convex topological space behaves with respect to quotients. As Proposition 3.6.16 shows that the maximum of a finite number of seminorms from a separating family of seminorms that generates a locally convex topology is a continuous seminorm and thus can be added into the family without modifying the topology, the assumption we add on the separating family in the following is moot.

Proposition 3.6.19. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space, let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$, and let $p : \mathcal{V} \to [0, \infty)$ be a seminorm on $\mathcal{V}$. If $\widetilde{p} : \mathcal{V}/\mathcal{W} \to [0, \infty)$ is defined by*

$$\widetilde{p}(v + \mathcal{W}) = \inf(\{p(v + w) \mid w \in \mathcal{W}\})$$

for all $v + \mathcal{W} \in \mathcal{V}$, then $\widetilde{p}$ is a seminorm on $\mathcal{V}/\mathcal{W}$. Moreover, if $\mathcal{F}$ is a separating family of seminorms on $\mathcal{V}$ that generate $\mathcal{T}$ and is closed under taking the maximum of finitely many elements, then $\widetilde{\mathcal{F}} = \{\widetilde{p} \mid p \in \mathcal{F}\}$ is a separating family of seminorms on $\mathcal{V}/\mathcal{W}$ that generate the quotient topology.

Proof. If p is a seminorm on $\mathcal{V}$, then by nearly identical arguments to those used in Theorem 1.3.3 it follows that $\widetilde{p}$ is a seminorm on $\mathcal{V}/\mathcal{W}$.

Let $\mathcal{F}$ be a separating family of seminorms on $\mathcal{V}$ that generate $\mathcal{T}$. To see that $\widetilde{\mathcal{F}} = \{\widetilde{p} \mid p \in \mathcal{F}\}$ is separating on $\mathcal{V}/\mathcal{W}$, suppose to the contrary that there exists a $v_0 + \mathcal{W} \in \mathcal{V}/\mathcal{W}$ such that $\widetilde{p}(v_0 + \mathcal{W}) = 0$ for all $\widetilde{p} \in \widetilde{\mathcal{F}}$. We claim that $v_0 \in \overline{\mathcal{W}}$. To see this, let U be an arbitrary neighbourhood of v_0. Thus, as $\mathcal{F}$ is a separating family of seminorms, there exists an $\epsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{F}$ such that

$$N(v_0, \{p_1, \ldots, p_n\}, \epsilon) \subseteq U.$$

Recall if $p_0 : \mathcal{V} \to [0, \infty)$ is define by

$$p_0(v) = \max(\{p_1(v), \ldots, p_n(v)\})$$

for all $v \in \mathcal{V}$, then $p_0 \in \mathcal{F}$ by assumption and

$$N(v_0, p_0, \epsilon) = N(v_0, \{p_1, \ldots, p_n\}, \epsilon) \subseteq U.$$

However, as $\widetilde{p_0}(v_0 + \mathcal{W}) = 0$, there exists a $w \in \mathcal{W}$ such that $p_0(v_0 + w) < \epsilon$ and thus $\mathcal{W} \cap U \neq \emptyset$. Therefore, as U was arbitrary, we obtain that $v_0 \in \overline{\mathcal{W}}$. However, as $\mathcal{W}$ is closed, this implies $v_0 \in \mathcal{W}$ so $v_0 + \mathcal{W} = \vec{0} + \mathcal{W}$. Hence $\widetilde{\mathcal{F}}$ is separating.

Let $\widetilde{\mathcal{T}}$ be the topology on $\mathcal{V}/\mathcal{W}$ generated by $\widetilde{\mathcal{F}}$. To show that $\widetilde{\mathcal{T}}$ is the quotient topology, we will follow the same idea as Proposition 3.3.5. Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the vector space quotient map. To complete the proof, it suffices by Corollary A.7.21 to show that q is a quotient map in the topological sense.

To see that q is a quotient map, first we note that q is clearly surjective. To show that q is continuous, note for all $v \in \mathcal{V}$, $p \in \mathcal{F}$, and $\epsilon > 0$ that

$$q^{-1}(N(v + \mathcal{W}, \widetilde{p}, \epsilon)) = \bigcup_{w \in \mathcal{W}} N(v + w, p, \epsilon).$$

Hence $q^{-1}(N(v + \mathcal{W}, \widetilde{p}, \epsilon))$ is open in $(\mathcal{V}, \mathcal{T})$ being the union of open sets. Since

$$\{N(v + \mathcal{W}, \widetilde{p}, \epsilon) \mid v \in \mathcal{V}, p \in \mathcal{F}, \epsilon > 0\}$$

form a subbasis of $\widetilde{\mathcal{T}}$, it follows that $q^{-1}(U) \in \mathcal{T}$ for all $U \in \widetilde{\mathcal{T}}$ so q is continuous.

To see that q is open, suppose U is an arbitrary subset of $\mathcal{V}/\mathcal{W}$ such that $q^{-1}(U)$ is open in $\mathcal{V}$. To see that U is open in $\mathcal{V}/\mathcal{W}$, let $x + \mathcal{W} \in U$ be arbitrary. Hence $x \in q^{-1}(U)$. Hence, as $q^{-1}(U)$ is open in $\mathcal{V}$ there exists an $\epsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{F}$ so that

$$N(x, \{p_1, \ldots, p_n\}, \epsilon) \subseteq q^{-1}(U).$$

Again, if we let $p_0 = \max(\{p_1, \ldots, p_n\})$, then $p_0 \in \mathcal{F}$ and

$$N(x, p_0, \epsilon) = N(x, \{p_1, \ldots, p_n\}, \epsilon) \subseteq q^{-1}(U).$$

We claim that $N(x + \mathcal{W}, \widetilde{p_0}, \epsilon) \subseteq U$. To see this, suppose $v \in \mathcal{V}$ is such that

$$\widetilde{p_0}((v - x) + \mathcal{W}) = \widetilde{p_0}((\vec{v} + \mathcal{W}) - (\vec{u} + \mathcal{W})) < \epsilon.$$

Hence, by the definition of $\widetilde{p_0}$, there exists a $w \in \mathcal{W}$ such that

$$p_0(v - x + w) < \epsilon.$$

Hence $v + w \in q^{-1}(U)$ by the above computation so

$$v + \mathcal{W} = q(v + w) \in U$$

as desired. Therefore, since $x + \mathcal{W} \in U$ was arbitrary, U is open in $\mathcal{V}/\mathcal{W}$. Hence q is a quotient map by the definition of a quotient map thereby yielding the proof. ∎

Chapter 4

Hahn-Banach Theorems

With our introduction to topological vector spaces and, more importantly, locally convex topological vector spaces complete, we can move on and obtain more in-depth knowledge about these spaces and their properties. As such, we will follow a similar pattern as we did for normed linear spaces; we will examine the continuous linear maps between such spaces. In particular, we will examine a specific collection of results pertaining to the continuous linear maps on a locally convex topological vector space known as the Hahn-Banach Theorems. These theorems are debatably the most important results in elementary functional analysis.

The Hahn-Banach Theorems separate into two classes of theorems: extension and separation. As motivation for the Hahn-Banach Extension Theorems, we note we still have questions relating to whether the canonical embedding of a normed linear space $\mathcal{X}$ into its second dual is an isometry (Remark 1.6.6) and whether the adjoint of a bounded linear map T between normed linear spaces has the same norm as T (Remark 1.6.8). Both of these questions can be answered via extending continuous linear functionals. Indeed by Corollary 3.5.7 we know that every linear map on a finite dimensional topological vector space is continuous so we can easily construct continuous linear functionals on finite dimensional subspaces of locally convex topological spaces. Being able to extend these linear functionals to the entire space will enable us to solve these (and many other) questions.

The Hahn-Banach Separation Theorems are more geared towards the geometric structures of locally convex topological vector spaces. Indeed the goal of the Hahn-Banach Separation Theorems is when given two disjoint sets to construct a continuous linear functional for which translates of the kernel separates the two sets. As kernels of continuous linear functionals are closed subspaces of co-dimension 1, we are effectively cutting our vector space in half with one set on either side of our space. Such geometric results are particularly useful for describing closed convex subsets of locally convex topological vector spaces using linear functionals.

We should also note that all authors use “the Hahn-Banach Theorem” to denote one of these many theorems. We will do the same, but will always provided a reference to the specific version of the theorem we are using.

4.1 Linear Functionals and Hyperplanes

Before we can attempt to prove our desired results in locally convex topological vector spaces, we must first retrace our steps and return to some elementary linear algebra that is not generally done in undergraduate courses as its uses are not seen until this point in pure mathematics. In particular, this section will focus on (not necessarily continuous) linear functionals and the subspaces their kernels define. We begin with the following.

Definition 4.1.1. Given a vector space $\mathcal{V}$ over $\mathbb{K}$, a *linear functional on* $\mathcal{V}$ is a linear map $f : \mathcal{V} \to \mathbb{K}$. The set of all linear functionals on $\mathcal{V}$ is denoted $\mathcal{V}^\sharp$ and is called the *algebraic dual of* $\mathcal{V}$.

Given a topological vector space $(\mathcal{V}, \mathcal{T})$ over $\mathbb{K}$, a *continuous linear functional on* $\mathcal{V}$ is a continuous linear map $f : \mathcal{V} \to \mathbb{K}$. The set of all continuous linear functionals on $\mathcal{V}$ is denoted $\mathcal{V}^*$ and is called the *(topological) dual of* $\mathcal{V}$.

Clearly if $\mathcal{V}$ is a topological vector space, then $\mathcal{V}^* \subseteq \mathcal{V}^\sharp$. Whereas Corollary 3.5.7 implies $\mathcal{V}^* = \mathcal{V}^\sharp$ if $\mathcal{V}$ is a finite dimensional topological vector space, it is not surprising that $\mathcal{V}^* \neq \mathcal{V}^\sharp$ is possible in the infinite dimensional setting.

Example 4.1.2. Given an infinite dimensional normed linear space $(\mathcal{X}, \|\cdot\|)$, $\mathcal{X}^* \neq \mathcal{X}^\sharp$. To see this, fix a vector space basis $\{x_\lambda\}_{\lambda\in\Lambda}$. By scaling if necessary, we may assume that $\|x_\lambda\| = 1$ for all $\lambda \in \Lambda$. As Λ must be infinite, choose distinct vectors $\{x_n\}_{n\geq 1}$ from $\{x_\lambda\}_{\lambda\in\Lambda}$. Define a linear map $f : \mathcal{X} \to \mathbb{K}$ by defining $f(x_n) = n$ for all $n \in \mathbb{N}$, $f(x) = 0$ for all $x \in \{x_\lambda\}_{\lambda\in\Lambda} \setminus \{x_n\}_{n\geq 1}$, and by extending the definition of f by linearity. As $|f(x_n)| \geq n$ and $\|x_n\| = 1$, we see that f is unbounded.

Of course, any linear functional on a normed linear space that is not bounded is not continuous and several examples of this were given in Section 1.4.

Of use in our theory will be the ability to reduce from the $\mathbb{K} = \mathbb{C}$ case to the $\mathbb{K} = \mathbb{R}$ case. The first step of doing this is the following.

Definition 4.1.3. Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. Given $f \in \mathcal{V}^\sharp$, the *real and imaginary parts of* f are the maps $\mathrm{Re}(f), \mathrm{Im}(f) : \mathcal{V} \to \mathbb{R}$ defined by

$$\mathrm{Re}(f)(\vec{v}) = \frac{f(\vec{v}) + \overline{f(\vec{v})}}{2} \qquad \text{and} \qquad \mathrm{Im}(f)(\vec{v}) = \frac{f(\vec{v}) - \overline{f(\vec{v})}}{2i}$$

for all $\vec{v} \in \mathcal{V}$.

Of course, the following is trivial to demonstrate.

Lemma 4.1.4. *Let $\mathcal{V}$ be a vector space over $\mathbb{C}$ and let $f \in \mathcal{V}^\sharp$. Then $f = \mathrm{Re}(f) + i\mathrm{Im}(f)$ and $\mathrm{Re}(f), \mathrm{Im}(f) : \mathcal{V} \to \mathbb{R}$ are $\mathbb{R}$-linear.*

However, what we are after is the following way to take $\mathbb{R}$-linear functionals and produce $\mathbb{C}$-linear functionals. In particular (3) deals with continuity in the normed linear space setting whereas (2) deals with the continuity in the locally convex topological vector space setting via Corollary 3.6.18 and knowledge of seminorms.

Lemma 4.1.5. *Let $\mathcal{V}$ be a vector space over $\mathbb{C}$ and let $f : \mathcal{V} \to \mathbb{R}$ be $\mathbb{R}$-linear. Then the following hold:*

(1) If $f_\mathbb{C} : \mathcal{V} \to \mathbb{C}$ is defined by

$$f_\mathbb{C}(\vec{v}) = f(\vec{v}) - if(i\vec{v})$$

for all $\vec{v} \in \mathcal{V}$, then $f_\mathbb{C}$ is $\mathbb{C}$-linear.

(2) If p is a $\mathbb{C}$-seminorm on $\mathcal{V}$, then $|f(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$ if and only if $|f_\mathbb{C}(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$.

(3) If $\mathcal{V}$ is a normed linear space, then $f_\mathbb{C}$ is bounded if and only if f is bounded. Moreover, if f is bounded, then $\|f_\mathbb{C}\| = \|f\|$.

(4) If $g : \mathcal{V} \to \mathbb{C}$ is $\mathbb{C}$-linear and $\mathrm{Re}(g) = f$, then $f_\mathbb{C} = g$.

Proof. For (1), we first note that $f_\mathbb{C}$ preserves addition since f and the scalar multiplication are additive. To see that $f_\mathbb{C}$ preserves $\mathbb{C}$-scalar multiplication, let $a, b \in \mathbb{R}$ and $\vec{v} \in \mathcal{V}$ be arbitrary. Then, since f is $\mathbb{R}$-linear,

$$\begin{aligned} f_\mathbb{C}((a+bi)\vec{v}) &= f((a+bi)\vec{v}) - if(i(a+bi)\vec{v}) \\ &= f(a\vec{v}) + f(ib\vec{v}) - if(ia\vec{v}) - if(-b\vec{v}) \\ &= af(\vec{v}) + bf(i\vec{v}) - aif(i\vec{v}) + bif(\vec{v}) \\ &= (a+bi)f(\vec{v}) - (a+bi)if(i\vec{v}) = (a+bi)f_\mathbb{C}(\vec{v}). \end{aligned}$$

Hence, as a, b, and $\vec{v}$ were arbitrary, $f_\mathbb{C}$ preserves $\mathbb{C}$-scalar multiplication. Hence $f_\mathbb{C}$ is $\mathbb{C}$-linear.

To see (2), note if $|f_\mathbb{C}(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$, then

$$|f(\vec{v})| \leq \sqrt{|f(\vec{v})|^2 + |if(i\vec{v})|^2} = |f(\vec{v}) - if(i\vec{v})| = |f_\mathbb{C}(\vec{v})| \leq p(\vec{v})$$

for all $\vec{v} \in \mathbb{R}$ as $f(\vec{v}), f(i\vec{v}) \in \mathbb{R}$. Conversely, suppose $|f(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$. To see that $|f_\mathbb{C}(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$, fix $\vec{v} \in \mathcal{V}$ and choose $z \in \mathbb{C}$ such that $|z| = 1$ and $zf_\mathbb{C}(\vec{v}) \in \mathbb{R}$ with $zf_\mathbb{C}(\vec{v}) \geq 0$. Thus

$$|f_\mathbb{C}(\vec{v})| = zf_\mathbb{C}(\vec{v}) = f_\mathbb{C}(z\vec{v}) = f(z\vec{v}) - if(iz\vec{v}).$$

However, as $f(z\vec{v}), f(iz\vec{v}) \in \mathbb{R}$, we obtain that $f(iz\vec{v}) = 0$ and

$$|f_{\mathbb{C}}(\vec{v})| = f(z\vec{v}) = |f(z\vec{v})| \leq p(z\vec{v}) = |z|p(\vec{v}) = p(\vec{v}).$$

Therefore, as $\vec{v} \in \mathcal{V}$ was arbitrary, (2) is complete.

To see (3), note for all $M > 0$ that the map $\vec{v} \mapsto M\|\vec{v}\|$ is a seminorm on $\mathcal{V}$. Therefore, by (2) we see that $|f(\vec{v})| \leq M\|\vec{v}\|$ for all $\vec{v} \in \mathcal{V}$ if and only if $|f_{\mathbb{C}}(\vec{v})| \leq M\|\vec{v}\|$ for all $\vec{v} \in \mathcal{V}$. Hence (3) follows.

Finally, to see (4), note that $\mathrm{Re}(g) = f$ implies that

$$\begin{aligned} f(i\vec{v}) = \mathrm{Re}(g)(i\vec{v}) &= \frac{g(i\vec{v}) + \overline{g(i\vec{v})}}{2} \\ &= \frac{ig(\vec{v}) + \overline{ig(\vec{v})}}{2} \\ &= \frac{ig(\vec{v}) - i\overline{g(\vec{v})}}{2} \\ &= -\frac{g(\vec{v}) - \overline{g(\vec{v})}}{2i} = -\mathrm{Im}(g)(\vec{v}) \end{aligned}$$

for all $\vec{v} \in \mathcal{V}$. Hence

$$f_{\mathbb{C}}(\vec{v}) = f(\vec{v}) - if(i\vec{v}) = \mathrm{Re}(g)(\vec{v}) + i\mathrm{Im}(g)(\vec{v}) = g$$

as desired. ∎

The next ingredient we require pertaining to linear functionals are their connections with the following objects.

Definition 4.1.6. Given a vector space $\mathcal{V}$ over $\mathbb{K}$, a *hyperplane in* $\mathcal{V}$ is a vector subspace $\mathcal{W}$ of $\mathcal{V}$ such that $\dim(\mathcal{V}/\mathcal{W}) = 1$.

Remark 4.1.7. Clearly if $f \in \mathcal{V}^\sharp$ is non-zero, then $\ker(f)$ is a hyperplane since if $\vec{v}_0 \in \mathcal{V}$ and $f(\vec{v}_0) \neq 0$, then $\mathcal{V}/\ker(f)$ is spanned by $\vec{v}_0 + \ker(f)$ which is non-zero. Conversely, if $\mathcal{W}$ is a hyperplane in $\mathcal{V}$, then the quotient map $q : \mathcal{V} \to \mathcal{V}/\mathcal{W} \cong \mathbb{K}$ is a linear functional with kernel equal to $\mathcal{W}$.

In fact, given multiple linear functionals, much is known about the intersection of their kernels.

Lemma 4.1.8. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$ and let $f_1, \ldots, f_n \in \mathcal{V}^\sharp$. If $\mathcal{W} = \cap_{k=1}^n \ker(f_k)$, then $\dim(\mathcal{V}/\mathcal{W}) \leq n$.*

Proof. We will proceed by induction on n. If $n = 1$, then $\mathcal{W} = \ker(f_1)$. If $f_1 = 0$ then $\mathcal{W} = \mathcal{V}$ so $\mathcal{V}/\mathcal{W}$ is the zero vector space and thus 0-dimensional. Otherwise, choose any vector $\vec{v}_1 \in \mathcal{V}$ such that $f_1(\vec{v}_1) \neq 0$. It is elementary to see that $\mathcal{V} = \mathbb{K}\vec{v}_1 + \mathcal{W}$ and hence $\mathcal{V}/\mathcal{W} = \mathrm{span}(\{\vec{v}_1 + \mathcal{W}\})$ is 1-dimensional. Hence the base case has been demonstrated.

To proceed by induction, suppose the result has been demonstrated for n linear functionals and let $f_1, \ldots, f_n, f_{n+1} \in \mathcal{V}^\sharp$ be arbitrary. Let $\mathcal{W} = \cap_{k=1}^{n+1} \ker(f_k)$ and let $\mathcal{W}_0 = \cap_{k=1}^{n} \ker(f_k)$. By the induction hypothesis, $\dim(\mathcal{V}/\mathcal{W}_0) \leq n$. Hence there exists $\vec{v}_1, \ldots, \vec{v}_n \in \mathcal{V}$ such that

$$\mathcal{V}/\mathcal{W}_0 = \operatorname{span}(\{\vec{v}_1 + \mathcal{W}_0, \ldots, \vec{v}_n + \mathcal{W}_0\})$$

and hence

$$\mathcal{V} = \mathcal{W}_0 + \operatorname{span}(\{\vec{v}_1, \ldots, \vec{v}_n\}).$$

Let $g = f_{n+1}|_{\mathcal{W}_0}$. Clearly g is linear and thus by the base case $\mathcal{W}_0/\ker(g)$ is at most one-dimensional. Thus there exists a $\vec{v}_{n+1} \in \mathcal{W}_0$ (possibly the zero vector) such that $\mathcal{W}_0/\ker(g) = \operatorname{span}(\{\vec{v}_{n+1} + \mathcal{W}_0\})$. Hence $\mathcal{W}_0 = \mathbb{K}\vec{v}_{n+1} + \ker(g)$ so

$$\mathcal{V} = \ker(g) + \operatorname{span}(\{\vec{v}_1, \ldots, \vec{v}_n, \vec{v}_{n+1}\}).$$

However, since

$$\ker(g) = \ker(f_{n+1}) \cap \mathcal{W}_0 = \mathcal{W}$$

we obtain that

$$\mathcal{V} = \mathcal{W} + \operatorname{span}(\{\vec{v}_1, \ldots, \vec{v}_n, \vec{v}_{n+1}\}).$$

Thus $\{\vec{v}_1 + \mathcal{W}, \ldots, \vec{v}_n + \mathcal{W}, \vec{v}_{n+1} + \mathcal{W}\}$ spans $\mathcal{V}/\mathcal{W}$ and hence $\dim(\mathcal{V}/\mathcal{W}) \leq n+1$ thereby completing the inductive step. ∎

Using Lemma 4.1.8 we can examine exactly when one linear functional is in the span of other linear functionals based on the kernels.

Lemma 4.1.9. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$ and let $f, f_1, f_2, \ldots, f_n \in \mathcal{V}^\sharp$. Then $\bigcap_{k=1}^n \ker(f_k) \subseteq \ker(f)$ if and only if $f \in \operatorname{span}(\{f_1, \ldots, f_n\})$.*

Proof. If $f \in \operatorname{span}(\{f_1, \ldots, f_n\})$ then it is trivial to verify using definitions that $\bigcap_{k=1}^n \ker(f_k) \subseteq \ker(f)$.

To prove the converse, first note by the previous direction we may assume without loss of generality that $\{f_1, \ldots, f_n\}$ are linearly independent (for otherwise one is a linear combination of the others and can be removed without modifying the intersection of the kernels). Let $\mathcal{W} = \cap_{k=1}^n \ker(f_k)$, which is a vector subspace of $\mathcal{V}$. For all $k \in \{1, \ldots, n\}$, define $\widetilde{f_k} \in (\mathcal{V}/\mathcal{W})^\sharp$ by

$$\widetilde{f_k}(\vec{v} + \mathcal{W}) = f_k(\vec{v})$$

for all $\vec{v} \in \mathcal{V}$. It is necessary to note that $\widetilde{f_k}$ is well-defined as if $\vec{v}_1 + \mathcal{W} = \vec{v}_2 + \mathcal{W}$, then $\vec{v}_1 - \vec{v}_2 \in \mathcal{W} \subseteq \ker(f_k)$ so $f_k(\vec{v}_1) = f_k(\vec{v}_2)$. Moreover, note if $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is the canonical quotient map, then $f_k = \widetilde{f_k} \circ q$ for all $k \in \{1, \ldots, n\}$. Similarly, if $\widetilde{f} \in (\mathcal{V}/\mathcal{W})^\sharp$ is defined by

$$\widetilde{f}(\vec{v} + \mathcal{W}) = f(\vec{v})$$

for all $\vec{v} \in \mathcal{V}$, then $\widetilde{f}$ is well-defined since $\mathcal{W} \subseteq \ker(f)$, and $f = \widetilde{f} \circ q$.

Recall from Lemma 4.1.8 that $\dim(\mathcal{V}/\mathcal{W}) \leq n$. Hence $\dim((\mathcal{V}/\mathcal{W})^\sharp) \leq n$. We claim that $\{\widetilde{f}_1, \ldots, \widetilde{f}_n\}$ is linearly independent and thus a basis for $(\mathcal{V}/\mathcal{W})^\sharp$. To see this, suppose $z_1, \ldots, z_n \in \mathbb{K}$ are such that

$$z_1\widetilde{f}_1 + \cdots + z_n\widetilde{f}_n = 0.$$

Hence, the definitions of $\widetilde{f}_k$ implies that

$$z_1 f_1(\vec{v}) + \cdots + z_n f_n(\vec{v}) = 0$$

for all $\vec{v} \in \mathcal{V}$. Therefore, since $\{f_1, \ldots, f_n\}$ are linearly independent, we obtain that $z_1 = \cdots = z_n = 0$. Hence $\{\widetilde{f}_1, \ldots, \widetilde{f}_n\}$ is linearly independent and thus a basis for $(\mathcal{V}/\mathcal{W})^\sharp$.

Since $\widetilde{f} \in (\mathcal{V}/\mathcal{W})^\sharp$, there exists $z_1, \ldots, z_n \in \mathbb{K}$ such that

$$\widetilde{f} = z_1\widetilde{f}_1 + \cdots + z_n\widetilde{f}_n.$$

By the definitions of $\widetilde{f}$ and $\widetilde{f}_k$, this implies

$$f = z_1 f_1 + \cdots + z_n f_n$$

so $f \in \operatorname{span}(\{f_1, \ldots, f_n\})$ as desired. ∎

The last ingredients we need pertaining to linear functionals are various methods to know when they are continuous. In particular, the following is the simplest method to verify a linear functional is continuous.

Proposition 4.1.10. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. If $f \in \mathcal{V}^\sharp$, then $f \in \mathcal{V}^*$ if and only if $\ker(f)$ is closed.*

Proof. If $f \in \mathcal{V}^*$, then $\ker(f) = f^{-1}(\{0\})$ is closed as f is continuous and $\{0\}$ is a closed set.

Conversely, suppose $f \in \mathcal{V}^\sharp$ and $\mathcal{W} = \ker(f)$ is closed. If $f = 0$, then clearly f is continuous. If $f \neq 0$, then note by Proposition 3.3.6 that $\mathcal{V}/\mathcal{W}$ is a topological vector space and the canonical quotient map $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is continuous. Moreover, since $f \in \mathcal{V}^\sharp$, we know that $\mathcal{V}/\mathcal{W}$ is one-dimensional.

Recall if $\widetilde{f} \in (\mathcal{V}/\mathcal{W})^\sharp$ is defined by

$$\widetilde{f}(\vec{v} + \mathcal{W}) = f(\vec{v})$$

for all $\vec{v} \in \mathcal{V}$, then $\widetilde{f}$ is well-defined and $f = \widetilde{f} \circ q$. Since q is continuous and $\widetilde{f}$ is continuous by Corollary 3.5.7 as $\mathcal{V}/\mathcal{W}$ is one-dimensional, f is continuous being the composition of continuous functions. ∎

Combining the following result, which is proved using elementary ideas, with Proposition 4.1.10 tells us the behaviour of kernels of discontinuous linear functionals.

Proposition 4.1.11. *If $(\mathcal{V}, \mathcal{T})$ is a topological vector space and $\mathcal{M} \subseteq \mathcal{V}$ is a hyperplane, then either $\mathcal{M}$ is closed in $\mathcal{V}$ or $\mathcal{M}$ is dense in $\mathcal{V}$.*

Proof. If $\mathcal{M}$ is a hyperplane, then we know that $\overline{\mathcal{M}}$ is a subspace of $\mathcal{V}$ such that $\mathcal{M} \subseteq \overline{M} \subseteq \mathcal{V}$. As $\dim(\mathcal{V}/\mathcal{M}) = 1$, this implies either $\overline{\mathcal{M}} = \mathcal{M}$ (so $\mathcal{M}$ is closed), or $\overline{\mathcal{M}} = \mathcal{V}$ (so $\mathcal{M}$ is dense in $\mathcal{V}$). ∎

Finally, we arrive at an analogue of 'bounded linear functionals are continuous' for topological vector spaces.

Proposition 4.1.12. *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $f \in \mathcal{V}^\sharp$. If there exists a neighbourhood U of $\vec{0}$ and a constant M such that $\mathrm{Re}(f(x)) \leq M$ for all $x \in U$, then f is uniformly continuous.*

Proof. By Lemma 3.1.7 there exists a balanced neighbourhood U_0 of $\vec{0}$ contained in U. Notice for any $x \in U_0$ there exists an $z_x \in \mathbb{K}$ with $|z_x| = 1$ such that

$$|f(x)| = z_x f(x) = f(z_x x) = \mathrm{Re}(f(z_x x)).$$

However, since U_0 is balanced, $z_x x \in U_0$ so this implies that $|f(x)| \leq M$ for all $x \in U_0$.

Define $p : \mathcal{V} \to [0, \infty)$ by

$$p(x) = |f(x)|$$

for all $x \in \mathcal{V}$. Clearly p is a seminorm as $f \in \mathcal{V}^\sharp$. Hence, as p is bounded on U_0, Proposition 3.6.13 implies that p is uniformly continuous on $\mathcal{V}$ and thus continuous by Proposition 3.4.8. As $f(\vec{0}) = 0 = p(\vec{0})$, this implies that f is continuous at $\vec{0}$ and thus Theorem 3.4.9 implies that f is uniformly continuous. ∎

4.2 Hahn-Banach Extension Theorems

With the preliminaries out of the way, we can proceed to show on locally convex topological vector spaces that continuous linear functional on subspaces can be extended to continuous linear functionals. This is accomplished via Proposition 3.6.16 by showing linear functionals bounded by seminorms extend to linear functionals bounded by seminorms. This is accomplished by showing linear functionals bounded by sublinear functionals extend to linear functionals bounded by sublinear functionals. This is simplified by first looking at hyperplanes in the following result, which is purely algebraic in nature.

Proposition 4.2.1. *Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ and let $p : \mathcal{V} \to \mathbb{R}$ be a sublinear functional. Suppose $\mathcal{W}$ is a hyperplane of $\mathcal{V}$ and $f : \mathcal{W} \to \mathbb{R}$ is a linear functional such that $f(\vec{w}) \leq p(\vec{w})$ for all $\vec{w} \in \mathcal{W}$. Then there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $g|_{\mathcal{W}} = f$ and $g(\vec{v}) \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$.*

Proof. Fix $\vec{v}_0 \in \mathcal{V} \setminus \mathcal{W}$. Since $\mathcal{W}$ is a hyperplane of $\mathcal{V}$, we know for all $\vec{v} \in \mathcal{V}$ there exists a unique $x \in \mathbb{R}$ and a $\vec{w} \in \mathcal{W}$ such that $\vec{v} = x\vec{v}_0 + \vec{w}$. Thus, by linearity, it suffices to extend f to have a value on $\vec{v}_0$.

To proceed, let us consider the collection of all elements h of $\mathcal{V}^\sharp$ such that $h|_\mathcal{W} = f$. Indeed, by the above decomposition of elements of $\mathcal{V}$, such h are uniquely determined by their values on $\vec{v}_0$. In particular, for each $r \in \mathbb{R}$ there exists a unique $h_r \in \mathcal{V}^\sharp$ such that $h_r(\vec{v}_0) = r$ and $h_r|_\mathcal{W} = f$; that is, $h_r : \mathcal{V} \to \mathbb{R}$ defined by

$$h_r(x\vec{v}_0 + \vec{w}) = xr + f(\vec{w})$$

for all $x \in \mathbb{R}$ and $\vec{w} \in \mathcal{W}$ is a well-defined element of $\mathcal{V}^\sharp$. The question is, "is there an element of $\{h_r\}_{r\in\mathbb{R}}$ that satisfy the conclusions of the proposition?"

To proceed, let us consider what is required of an h_r to satisfy the conclusions of proposition. Indeed h_r satisfies the conclusions of the proposition if and only if for all $x \in \mathbb{R}$ and $\vec{w} \in \mathcal{W}$ we have that

$$rx + f(\vec{w}) \leq p(x\vec{v}_0 + \vec{w}).$$

Clearly this holds when $x = 0$ by the assumptions of the proposition. For $x \neq 0$, we obtain two equivalent inequalities based on the sign on x:

- if $x > 0$ then by using $x\vec{w}$ in place of $\vec{w}$, the above inequality is equivalent to

  $$r \leq \frac{1}{x}p\left(x\vec{v}_0 + x\vec{w}\right) - \frac{1}{x}f(x\vec{w}) = p(\vec{v}_0 + \vec{w}) - f(\vec{w})$$

  for all $\vec{w} \in \mathcal{W}$.

- if $x > 0$ then by using $-x\vec{w}$ in place of $\vec{w}$, the above inequality is equivalent to

  $$r \geq \frac{1}{x}p\left(x\vec{v}_0 - x\vec{w}\right) - \frac{1}{x}f(-x\vec{w}) = f(\vec{w}) - p(-\vec{v}_0 + \vec{w})$$

  for all $\vec{w} \in \mathcal{W}$ (as p is sublinear so we can pull out $-x$).

Therefore, h_r satisfies the conclusions of the proposition if and only if

$$f(\vec{w}_1) - p(-\vec{v}_0 + \vec{w}_1) \leq r \leq p(\vec{v}_0 + \vec{w}_2) - f(\vec{w}_2)$$

for all $\vec{w}_1, \vec{w}_2 \in \mathcal{W}$. In particular, the above show that there exists an h_r that satisfies the conclusions of the proposition if and only if

$$\begin{aligned}&\sup(\{f(\vec{w}_1) - p(-\vec{v}_0 + \vec{w}_1) \mid \vec{w}_1 \in \mathcal{W}\}) \\ &\qquad \leq \inf(\{p(\vec{v}_0 + \vec{w}_2) - f(\vec{w}_2) \mid \vec{w}_2 \in \mathcal{W}\})\end{aligned}$$

(in which case we can take r to be either the supremum, the infimum, or any real number in-between).

To see that such an r exists, notice for all $\vec{w}_1, \vec{w}_2 \in \mathcal{W}$ that

$$\begin{aligned} f(\vec{w}_1) + f(\vec{w}_2) = f(\vec{w}_1 + \vec{w}_2) &\leq p(\vec{w}_1 + \vec{w}_2) \\ &= p((-\vec{v}_0 + \vec{w}_1) + (\vec{v}_0 + \vec{w}_2)) \\ &\leq p(-\vec{v}_0 + \vec{w}_1) + p(\vec{v}_0 + \vec{w}_2) \end{aligned}$$

which implies that

$$f(\vec{w}_1) - p(-\vec{v}_0 + \vec{w}_1) \leq p(\vec{v}_0 + \vec{w}_2) - f(\vec{w}_2).$$

Hence the result follows. ∎

Upgrading from hyperplanes to arbitrary subspaces is easily accomplished with a maximality argument via Zorn's Lemma.

Theorem 4.2.2 (Hahn-Banach Extension Theorem - Real Vector Spaces). *Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ and let $p : \mathcal{V} \to \mathbb{R}$ be a sublinear functional. Suppose $\mathcal{W}$ is a vector subspace of $\mathcal{V}$ and $f : \mathcal{W} \to \mathbb{R}$ is a linear functional such that $f(\vec{w}) \leq p(\vec{w})$ for all $\vec{w} \in \mathcal{W}$. Then there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $g|_{\mathcal{W}} = f$ and $g(\vec{v}) \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$.*

Proof. Let

$$\mathcal{F} = \left\{ (\mathcal{Y}, h) \,\middle|\, \begin{array}{c} \mathcal{Y} \text{ is a vector subspace of } \mathcal{V} \text{ containing } \mathcal{W}, \\ h \in \mathcal{Y}^\sharp,\ h|_{\mathcal{W}} = f, \text{ and } h(\vec{y}) \leq p(\vec{y}) \text{ for all } \vec{y} \in \mathcal{Y} \end{array} \right\}.$$

Note that $(\mathcal{W}, f) \in \mathcal{F}$ so $\mathcal{F} \neq \emptyset$.

Define a relation $\preceq$ on $\mathcal{F}$ by setting $(\mathcal{Y}_1, h_1) \preceq (\mathcal{Y}_2, h_2)$ if and only if $\mathcal{Y}_1 \subseteq \mathcal{Y}_2$ and $h_2|_{\mathcal{Y}_1} = h_1$. It is not difficult to see that $\preceq$ is a partial ordering on $\mathcal{F}$.

Let $\mathcal{C} = \{(\mathcal{Y}_i, h_i) \mid i \in I\}$ for some ordered set I be a chain in $\mathcal{F}$. It is not difficult to verify that $\mathcal{Y} = \bigcup_{i \in I} \mathcal{Y}_i$ is a subspace of $\mathcal{V}$ containing $\mathcal{W}$ and if we define $h \in \mathcal{Y}^\sharp$ by $h(\vec{y}) = h_i(\vec{y}_i)$ for all $\vec{y} \in \mathcal{Y}_i$ and $i \in I$, then h is well-defined, $(\mathcal{Y}, h) \in \mathcal{F}$, and $(\mathcal{Y}_i, h_i) \preceq (\mathcal{Y}, h)$ for all $i \in I$. Hence every chain in $\mathcal{F}$ has an upper bound.

By Zorn's Lemma there exists a maximal element $(\mathcal{W}_0, h_0)$ of $\mathcal{F}$. If $\mathcal{W}_0 \neq \mathcal{V}$, fix $\vec{v}_0 \in \mathcal{V} \setminus \mathcal{W}_0$ and let $\mathcal{V}_0 = \mathbb{R}\vec{v}_0 + \mathcal{W}_0$. As $\mathcal{W}_0$ is a hyperplane in $\mathcal{V}_0$ by construction and $p|_{\mathcal{V}_0}$ is a sublinear functional such that that $h_0(\vec{w}) \leq p|_{\mathcal{V}_0}(\vec{w})$ for all $\vec{w} \in \mathcal{W}_0$, Proposition 4.2.1 implies there exists an $h \in \mathcal{V}_0^\sharp$ such that $(\mathcal{V}_0, h) \in \mathcal{F}$, $(\mathcal{W}_0, h_0) \neq (\mathcal{V}_0, h)$, and $(\mathcal{W}_0, h_0) \preceq (\mathcal{V}_0, h)$ thereby contradicting the maximality of $(\mathcal{W}_0, h_0)$. Hence $\mathcal{W}_0 = \mathcal{V}$ and the proof is complete. ∎

To upgrade from sublinear functionals to seminorms is accomplished via linearity in the $\mathbb{K} = \mathbb{R}$ case and using the linear functionals from Lemma 4.1.5 in the $\mathbb{K} = \mathbb{C}$ case.

Theorem 4.2.3 (Hahn-Banach Extension Theorem - Seminorm). *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$ and let $p : \mathcal{V} \to \mathbb{R}$ be a seminorm. Suppose $\mathcal{W}$ is a vector subspace of $\mathcal{V}$ and $f : \mathcal{W} \to \mathbb{K}$ is a linear functional such that $|f(\vec{w})| \leq p(\vec{w})$ for all $\vec{w} \in \mathcal{W}$. Then there exists a linear functional $g : \mathcal{V} \to \mathbb{K}$ such that $g|_{\mathcal{W}} = f$ and $|g(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$.*

Proof. In the case that $\mathbb{K} = \mathbb{R}$, recall that every seminorm is a sublinear functional. Therefore, as

$$f(\vec{w}) \leq |f(\vec{w})| \leq p(\vec{w})$$

for all $\vec{w} \in \mathcal{W}$, we obtain by Theorem 4.2.2 that there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $g|_{\mathcal{W}} = f$ and $g(\vec{v}) \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$. Since

$$-g(\vec{v}) = g(-\vec{v}) \leq p(-\vec{v}) = p(\vec{v})$$

as p is a seminorm, we obtain that $|g(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$ as desired.

In the case that $\mathbb{K} = \mathbb{C}$, let $f_1 = \mathrm{Re}(f)$. By Lemma 4.1.5 (using (2) and (4)), this implies that f_1 is a $\mathbb{R}$-linear functional on $\mathcal{W}$ and $f_1(\vec{w}) \leq p(\vec{w})$ for all $\vec{w} \in \mathcal{W}$. As every vector space over $\mathbb{C}$ is a vector space over $\mathbb{R}$ and the restriction of a seminorm on a vector space over $\mathbb{C}$ to a vector space over $\mathbb{R}$ remains a seminorm, we obtain by the $\mathbb{K} = \mathbb{R}$ case that that there exists a $\mathbb{R}$-linear functional $g_0 : \mathcal{V} \to \mathbb{R}$ such that $g_0|_{\mathcal{W}} = f_1$ and $|g_0(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$. Luckily, if $g = (g_0)_{\mathbb{C}}$ as in Lemma 4.1.5, then the same lemma implies g is $\mathbb{C}$-linear, that $g|_{\mathcal{W}} = f$, and that $|g(\vec{v})| \leq p(\vec{v})$ for all $\vec{v} \in \mathcal{V}$. ■

Upgrading to continuous linear functionals on locally convex topological vector spaces is easily obtained via the connection between continuous linear functionals and continuous seminorms from Proposition 3.6.16.

Theorem 4.2.4 (Hahn-Banach Extension Theorem - Continuous Linear Functionals). *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{W}$ be a vector subspace of $\mathcal{V}$. If $f \in \mathcal{W}^*$, then there exists a $g \in \mathcal{V}^*$ such that $g|_{\mathcal{W}} = f$.*

Proof. Since $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space, there exists a separating family of seminorms $\mathcal{F}$ on $\mathcal{V}$ that generate $\mathcal{T}$ by Theorem 3.6.15. As $\mathcal{W}$ is equipped with the subspace topology from $\mathcal{V}$, it is elementary based on the definition from Theorem 3.2.10 to see that

$$\mathcal{F}_{\mathcal{W}} = \{p|_{\mathcal{W}} \mid p \in \mathcal{F}\}$$

is a separating family of seminorms on $\mathcal{W}$ that generate the subspace topology.

Since $\vec{w} \mapsto |f(\vec{w})|$ is a continuous seminorm on $\mathcal{W}$ as $f \in \mathcal{W}^*$ and since $\mathcal{F}_{\mathcal{W}}$ generates the topology on $\mathcal{W}$, Proposition 3.6.16 implies there exists a constant M and $p_1, \ldots, p_n \in \mathcal{F}$ such that

$$|f(x)| \leq M \max(\{p_1(x), \ldots, p_n(x)\})$$

for all $x \in \mathcal{W}$. Since

$$x \mapsto M \max(\{p_1(x), \dots, p_n(x)\})$$

is a seminorm on $\mathcal{V}$, Theorem 4.2.3 implies that there exists a $g \in \mathcal{V}^\sharp$ such that $g|_\mathcal{W} = f$ and

$$|g(x)| \leq M \max(\{p_1(x), \dots, p_n(x)\})$$

for all $x \in \mathcal{V}$. As another application of Proposition 3.6.16 implies $|g|$ is a continuous seminorm on $\mathcal{V}$ and thus $g \in \mathcal{V}^*$ by Corollary 3.6.18, the proof is complete. ∎

Finally, we return to the case of normed linear spaces with the additional question on norm bounds present.

Theorem 4.2.5 (Hahn-Banach Extension Theorem - Bounded Linear Functionals). *Let $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ be a normed linear space and let $\mathcal{W}$ be a closed vector subspace of $\mathcal{V}$. If $f \in \mathcal{W}^*$, then there exists a $g \in \mathcal{V}^*$ such that $g|_\mathcal{W} = f$ and $\|g\| = \|f\|$.*

Proof. Consider the map $p : \mathcal{X} \to [0, \infty)$ defined by

$$p(x) = \|f\| \, \|x\|_\mathcal{X}$$

for $x \in \mathcal{X}$. Clearly p is a seminorm on $\mathcal{X}$ such that $|f(x)| \leq p(x)$ for all $x \in \mathcal{W}$. Hence Theorem 4.2.3 implies there exists a $g \in \mathcal{V}^\sharp$ such that $g|_\mathcal{W} = f$ and $|g(x)| \leq p(x) = \|f\| \, \|x\|_\mathcal{X}$. This later inequality implies $\|g\| \leq \|f\|$ whereas $g|_\mathcal{M} = f$ implies $\|g\| \geq \|f\|$ thereby completing the proof. ∎

4.3 Corollaries of the Extension Theorems

There are many immediate corollaries of the Hahn-Banach Extension Theorems from Section 4.2, which are also often called "the Hahn-Banach Theorem". As such, we will examine such corollaries in this section. We begin with the implications for normed linear spaces by finally answering the question posed in Remark 1.6.6.

Corollary 4.3.1. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space and let $x \in \mathcal{X}$. Then*

$$\|x\| = \max\{|f(x)| \mid f \in \mathcal{X}^*, \|f\| = 1\}.$$

Proof. Fix $x \in \mathcal{X}$. Clearly the result holds when $x = \vec{0}$ so suppose $x \neq \vec{0}$. Since

$$|f(x)| \leq \|f\| \, \|x\|$$

for all $f \in \mathcal{X}^*$, we easily obtain that

$$\|x\| \leq \sup\{|f(x)| \mid f \in \mathcal{X}^*, \|f\| = 1\}.$$

Thus, to complete the proof, it suffices to show there exists an $f \in \mathcal{X}^*$ such that $\|f\| \leq 1$ and $|f(x)| = \|x\|$.

Consider $\mathcal{M} = \mathrm{span}(\{x\})$ which is a one-dimensional subspace of $\mathcal{V}$. Define $f_0 : \mathcal{M} \to \mathbb{K}$ by

$$f_0(\alpha x) = \alpha \|x\|$$

for all $\alpha \in \mathbb{K}$. Clear f_0 is linear and note for all $\alpha \in \mathbb{K}$ that

$$|f_0(\alpha x)| = |\alpha \|x\| \,| = |\alpha| \|x\| = \|\alpha x\|$$

so $f_0 \in \mathcal{M}^*$ with $\|f_0\| = 1$. Thus, by the Hahn-Banach Extension Theorem (Theorem 4.2.5) implies there exists an $f \in \mathcal{X}^*$ such that $f(x) = f_0(x) = \|x\|$ and $\|f\| = \|f_0\| = 1$, thereby completing the proof. ∎

Corollary 4.3.2. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. The canonical embedding $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ from Theorem 1.6.3 is isometric.*

Proof. This immediately follows from Theorem 1.6.3, Remark 1.6.6, and Corollary 4.3.1. ∎

Along similar lines, we have our answer to the question posed in Remark 1.6.8.

Corollary 4.3.3. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed linear spaces and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. If T^* is the adjoint of T from Theorem 1.6.7, then $\|T^*\| = \|T\|$.*

Proof. This immediately follows from Theorem 1.6.7, Remark 1.6.8, and Corollary 4.3.1. ∎

The above combined with the Uniform Boundedness Principle (Theorem 2.5.3) gives us another way to demonstrate if a subset of a Banach spaces is bounded in a similar manner to Corollary 2.5.4.

Corollary 4.3.4. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space and let $A \subseteq \mathcal{X}$ be non-empty. Then A is bounded if and only if*

$$\sup(\{|f(a)| \mid a \in A\}) < \infty$$

for all $f \in \mathcal{X}^*$.

Proof. First suppose that A is bounded. Thus there exists an $M \in \mathbb{R}$ such that $\|a\| \leq M$ for all $a \in A$. Hence for all $f \in \mathcal{X}^*$

$$|f(a)| \leq \|f\| \|a\| \leq M \|f\|$$

so

$$\sup(\{|f(a)| \mid a \in A\}) \leq M \|f\| < \infty.$$

Conversely, suppose that

$$\sup(\{|f(a)| \mid a \in A\}) < \infty$$

for all $f \in \mathcal{X}^*$. Recall for all $a \in A$ the map $\hat{a} : \mathcal{X}^* \to \mathbb{K}$ defined by

$$\hat{a}(f) = f(a)$$

for all $f \in \mathcal{X}^*$ is a bounded linear map with $\|\hat{a}\| = \|a\|$ by Corollary 4.3.2. Since the assumption of this direction of the proof implies that

$$\sup(\{|\hat{a}(f)| \mid a \in A\}) < \infty$$

for all $f \in \mathcal{X}^*$, and since $\mathcal{X}^*$ is a Banach space, the Uniform Boundedness Principle (Theorem 2.5.3) implies there exists an $M \in \mathbb{R}$ such that $\|\hat{a}\| \leq M$ for all $a \in A$. Hence$\|a\| \leq M$ for all $a \in A$, so A is bounded. ∎

Returning to the generality of locally convex topological vector spaces, we can construct continuous linear functionals mapping any finite number of vectors to any preselected scalars we choose. The reason we cannot do so with a countable number of vectors and remain continuous is contained in Example 4.1.2.

Corollary 4.3.5. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. Given a finite linearly independent set $\{x_k\}_{k=1}^n$ and constants $\{\alpha_k\}_{k=1}^n \subseteq \mathbb{K}$, there exists a $f \in \mathcal{V}^*$ such that $f(x_k) = \alpha_k$ for all $k \in \{1, \ldots, n\}$.*

Proof. Let $\mathcal{W} = \mathrm{span}(\{x_1, \ldots, x_n\})$ which is a finite dimensional vector subspace of $\mathcal{V}$. Since $\{x_k\}_{k=1}^n$ is linearly independent, elementary linear algebra implies there exists an $h \in \mathcal{W}^\sharp$ such that $h(x_k) = \alpha_k$ for all $k \in \{1, \ldots, n\}$. Since $\mathcal{W}$ is finite dimensional Corollary 3.5.7 implies that $h \in \mathcal{W}^*$. Hence the Hahn-Banach Theorem (Theorem 4.2.4) implies there exists an $f \in \mathcal{V}^*$ such that $f|_{\mathcal{W}} = h$ and thus $f(x_k) = \alpha_k$ for all $k \in \{1, \ldots, n\}$ as desired. ∎

One use of Corollary 4.3.5 is to show that certain subspaces of locally convex topological vector spaces behave in a similar way to subspaces in finite dimensional inner product spaces as follows.

Definition 4.3.6. A closed subspace $\mathcal{W}$ of a topological vector space $(\mathcal{V}, \mathcal{T})$ is said to be *topologically complemented* if there exists a closed subspace $\mathcal{Y}$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{W} \oplus \mathcal{Y}$; that is, $\mathcal{W} \cap \mathcal{Y} = \{\vec{0}\}$ and for all $\vec{v} \in \mathcal{V}$ there exists (unique) $\vec{w} \in \mathcal{W}$ and $\vec{y} \in \mathcal{Y}$ such that $\vec{v} = \vec{w} + \vec{y}$.

Of course, a Zorn's Lemma maximality argument shows that if $\mathcal{W}$ is a closed subspace of a topological vector space $(\mathcal{V}, \mathcal{T})$, then there exists a vector subspace $\mathcal{Y}$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{W} \oplus \mathcal{Y}$. Of course, the 'topological' portion of 'topologically complemented' is that $\mathcal{Y}$ is closed. It turns out that not every closed subspace of a locally convex topological vector space is topologically complemented. Indeed c_0 as a subspace of $\ell_\infty(\mathbb{N})$ is one such example, although the proof is difficult. However, having topologically complemented subspaces is useful for a wide variety of applications. For example, one can construct linear 'projections' onto each of the subspaces via the quotient maps. Consequently, the following is perhaps not surprising.

Corollary 4.3.7. *Every finite dimensional subspace of a locally convex topological vector space is topologically complemented.*

Proof. Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{W}$ be a finite dimensional subspace of $\mathcal{V}$. Let $\{x_k\}_{k=1}^n$ be a vector space basis for $\mathcal{W}$. Thus Corollary 4.3.5 implies for all $j \in \{1, \ldots, n\}$ there exist an $f_j \in \mathcal{V}^*$ such that $f_j(x_k) = \delta_{k,j}$ where $\delta_{k,j}$ is the Kronecker delta; that is

$$\delta_{k,j} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}.$$

Let $\mathcal{Y} = \bigcap_{j=1}^n \ker(f_j)$, which is a closed subspace of $\mathcal{V}$ as $f_j \in \mathcal{V}^*$ for all $j \in \{1, \ldots, n\}$. Thus, to complete the proof, it suffices to show that $\mathcal{V} = \mathcal{W} \oplus \mathcal{Y}$.

To see that $\mathcal{W} \cap \mathcal{Y} = \{\vec{0}\}$, let $x \in \mathcal{W} \cap \mathcal{Y}$ be arbitrary. As $x \in \mathcal{W}$, there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ such that

$$x = \sum_{k=1}^n \alpha_k x_k.$$

However, since $x \in \mathcal{Y}$, we obtain that $f_j(x) = 0$ so

$$0 = f_j(x) = f_j\left(\sum_{k=1}^n \alpha_k x_k\right) = \sum_{k=1}^n \alpha_k \delta_{k,j} = \alpha_j$$

for all $j \in \{1, \ldots, n\}$. Hence $x = \vec{0}$ as desired.

To see that $\mathcal{V} = \mathcal{W} + \mathcal{Y}$, let $x \in \mathcal{V}$ be arbitrary. For each $j \in \{1, \ldots, n\}$, let $\alpha_j = f_k(x)$. Clearly $w = \sum_{k=1}^n \alpha_k x_k \in \mathcal{W}$. Moreover, for all $j \in \{1, \ldots, n\}$ we see that

$$f_j(x - w) = f_j(x) - f_j(w) = \alpha_j - \sum_{k=1}^n \alpha_k \delta_{k,j} = 0$$

and thus $x - w \in \mathcal{Y}$. Hence there exists a $y \in \mathcal{Y}$ such that $x - w = y$ so $x = w + y \in \mathcal{W} + \mathcal{Y}$. Therefore, as x was arbitrary, $\mathcal{V} = \mathcal{W} + \mathcal{Y}$ thereby completing the proof. ∎

Returning to examples of constructing specific continuous linear functionals, we can further construct linear functionals that vanish on certain subspaces and take a non-zero value on a vector outside of the subspace.

Corollary 4.3.8. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. If $x \in \mathcal{V} \setminus \mathcal{W}$, then there exists a $f \in \mathcal{V}^*$ such that $f|_{\mathcal{W}} = 0$ and $f(x) \neq 0$.*

Proof. Consider $\mathcal{V}/\mathcal{W}$, which is a locally convex topological vector space by Proposition 3.3.6. Since $x \in \mathcal{V} \setminus \mathcal{W}$, $x + \mathcal{W} \neq \vec{0} + \mathcal{W}$ in $\mathcal{V}/\mathcal{W}$. Hence Corollary 4.3.5 implies there exists a $g \in (\mathcal{V}/\mathcal{W})^*$ such that $g(x + \mathcal{W}) \neq 0$.

If $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is the canonical quotient map, let $f : \mathcal{V} \to \mathbb{K}$ be defined by $f = g \circ q$. Clearly $f \in \mathcal{V}^*$ be the composition of continuous linear maps. Since $f|_{\mathcal{W}} = 0$ and $f(x) = g(x + \mathcal{W}) \neq 0$ by construction, the result is complete. ∎

Note one can improve Corollary 4.3.8 to multiple linearly independent vectors outside in $\mathcal{V}/\mathcal{W}$ via Corollary 4.3.5 provided the span of these vectors has trivial intersection with $\mathcal{W}$.

Corollary 4.3.8 allows us to revisit a family of seminorms from Chapter 3 we could not show were separating.

Example 4.3.9. Recall from Example 3.2.18 that given a normed linear space $(\mathcal{X}, \|\cdot\|)$ we have a family $\mathcal{F} = \{p_f \mid f \in \mathcal{X}^*\}$ of seminorms on $\mathcal{X}$ where

$$p_f(x) = |f(x)|$$

for all $x \in \mathcal{X}$ and $f \in \mathcal{X}^*$. By Corollary 4.3.8 $\mathcal{F}$ is a separating family of seminorms on $\mathcal{X}$. Note by Proposition 3.2.13 a net $(x_\lambda)_{\lambda \in \Lambda}$ converges to $x \in \mathcal{X}$ if and only if $\lim_{\lambda \in \Lambda} f(x_\lambda) = f(x)$ for all $f \in \mathcal{X}^*$. This topology is called the *weak topology* (weak because it is weaker than norm convergence).

Of course Example 4.3.9 can be extended to any locally convex topological space, which will be the topic of the next chapter. For now, in the case of a normed linear space, we also have control over the norms in Corollary 4.3.8.

Corollary 4.3.10. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space, let $\mathcal{Y}$ be a closed subspace of $\mathcal{X}$, and let $x \in \mathcal{X} \setminus \mathcal{Y}$. If*

$$d = \operatorname{dist}(x, \mathcal{Y}) = \|x + \mathcal{Y}\|,$$

then there exists an $f \in \mathcal{X}^*$ such that $\|f\| = 1$, $f|_{\mathcal{Y}} = 0$, and $f(x) = d$.

Proof. Consider $\mathcal{X}/\mathcal{Y}$, which is a normed linear space by Theorem 1.3.3. By Corollary 4.3.1 there exists a $g \in (\mathcal{X}/\mathcal{Y})^*$ such that $\|g\| = 1$ and $g(x+\mathcal{Y}) = d$.

If $q : \mathcal{X} \to \mathcal{X}/\mathcal{Y}$ is the canonical quotient map, let $f : \mathcal{X} \to \mathbb{K}$ be defined by $f = g \circ q$. Clearly $g \in \mathcal{X}^*$ be the composition of continuous linear maps.

Moreover $f|_{\mathcal{Y}} = 0$ and $f(x) = g(x + \mathcal{Y}) = d$ by construction. As $\|q\| \leq 1$, we see that $\|f\| \leq \|g\| \, \|q\| \leq 1$. Thus, to complete the proof, it suffice to show that $\|f\| \geq 1$.

Since $\|g\| = 1$, there exists a sequence of vectors $(x_n)_{n\geq 1}$ in $\mathcal{X}$ such that $\lim_{n\to\infty} |g(x_n + \mathcal{W})| = 1$ and $\|x_n + \mathcal{W}\| < 1$ (note we can get strictly less than one by scaling each vector by $1 - \frac{1}{n}$ if necessary). By Theorem 1.3.3 there exists $(w_n)_{n\geq 1}$ in $\mathcal{Q}$ such that $\|x_n + w_n\| < 1$. Therefore, as

$$\lim_{n\to\infty} |f(x_n + w_n)| = \lim_{n\to\infty} |g(x_n + \mathcal{W})| = 1$$

we obtain that $\|f\| \geq 1$ thereby completing the proof. ∎

Finally, we conclude with a description of every closed subspace of a locally convex topological vector space via the kernels of continuous linear maps.

Theorem 4.3.11. *If $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space and $\mathcal{W}$ is a vector subspace of $\mathcal{V}$, then*

$$\overline{\mathcal{W}} = \bigcap_{\substack{f \in \mathcal{V}^* \text{ and} \\ \mathcal{W} \subseteq \ker(f)}} \ker(f).$$

Proof. First, suppose that $f \in \mathcal{V}^*$ and $\mathcal{W} \subseteq \ker(f)$. As $\ker(f)$ is closed, this implies that $\overline{\mathcal{W}} \subseteq \ker(f)$. Hence

$$\overline{\mathcal{W}} \subseteq \bigcap_{\substack{f \in \mathcal{V}^* \text{ and} \\ \mathcal{W} \subseteq \ker(f)}} \ker(f).$$

For the reverse inclusion, let $x \in \mathcal{V} \setminus \overline{\mathcal{W}}$ be arbitrary. By Corollary 4.3.8 there exists an $f \in \mathcal{V}^*$ such that $f|_{\mathcal{W}} = 0$ and $f(x) \neq 0$. Hence $\mathcal{W} \subseteq \ker(f)$ but $x \notin \ker(f)$. Therefore, as x was arbitrary, the proof is complete. ∎

4.4 Hahn-Banach Separation Theorems

Theorem 4.3.11 demonstrates a connection between continuous linear functionals and the geometry of locally convex topological vector spaces. Thus the goal of this section is to demonstrate a different class of the Hahn-Banach theorems that emphasize this connection further. This is done via the following types of sets.

Definition 4.4.1. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space over $\mathbb{K}$. A *open half-space* is any subset $\mathcal{S}$ of $\mathcal{V}$ of the form

$$S = \{x \in \mathcal{V} \mid Re(f(x)) > \kappa\}$$

where $f \in \mathcal{V}^*$ and $\kappa \in \mathbb{R}$.

Similarly, a *closed half-space* is any subset $\mathcal{S}$ of $\mathcal{V}$ of the form

$$S = \{x \in \mathcal{V} \mid Re(f(x)) \geq \kappa\}$$

where $f \in \mathcal{V}^*$ and $\kappa \in \mathbb{R}$.

Remark 4.4.2. It is not difficult to verify one can replace $>$ with $<$ and $\geq$ with $\leq$ in the definitions of open and closed half-spaces respectively since $f \in \mathcal{V}^*$ if and only if $-f \in \mathcal{V}^*$. Moreover, as taking the real part of an open (respectively closed) subset of $\mathbb{K}$ produces an open (respectively closed) subset of $\mathbb{R}$, we see that open (respectively closed) half-spaces are open (respectively closed). Finally as if $a, b \in \mathbb{R}$ are such that $a, b > \kappa$ (respectively $\geq \kappa$) then $ta + (1-t)b > \kappa$ (respectively $\geq \kappa$) for all $t \in [0, 1]$, we see that open (respectively closed) half-spaces are convex subsets.

Of course, the terminology is motivated by the following examples.

Example 4.4.3. For all $a, b \in \mathbb{R}$, consider the $\mathbb{R}$-linear map $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f((x, y)) = ax + by$$

for all $(x, y) \in \mathbb{R}$. Thus for all $\kappa \in \mathbb{R}$, the set $\{(x, y) \mid f((x, y)) = \kappa\}$ is a line so

$$S = \{(x, y) \in \mathbb{R}^2 \mid f((x, y)) > \kappa\}$$

is the (half-)space on one side of the line.

Example 4.4.4. For all $w \in \mathbb{C}$, consider the $\mathbb{C}$-linear map $f : \mathbb{C} \to \mathbb{C}$ defined by

$$f(z) = zw$$

for all $z \in \mathbb{C}$. If $w = a + bi$ and $z = x + iy$, then

$$\mathrm{Re}(f(z)) = ax - by.$$

Thus, viewing $\mathbb{C}$ as $\mathbb{R}^2$, we see the set

$$S = \{z \in \mathbb{C} \mid \mathrm{Re}(f(z)) > \kappa\}$$

is the (half-)space on one side of a line.

As half-spaces are analogues of dividing a topological vector space in half, we make the following definitions.

Definition 4.4.5. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. A subset $\mathcal{A} \subseteq \mathcal{V}$ is said to be an *affine hyperplane* (respectively *affine vector subspace*, *affine closed subspace*) if $\mathcal{A}$ is a translate of a hyperplane (respectively vector subspace, closed subspace); that is, there exists an $x \in \mathcal{A}$ such that $\mathcal{A} - x$ is a hyperplane(respectively vector subspace, closed subspace).

Definition 4.4.6. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space over $\mathbb{R}$. Two subsets $A, B \subseteq \mathcal{V}$ are said to be *separated* if there exists closed half-planes S_A and S_B of $\mathcal{V}$ such that $A \subseteq S_A$, $B \subseteq S_B$, and $S_A \cap S_B$ is a closed affine hyperplane.

Similarly, two subsets $A, B \subseteq \mathcal{V}$ are said to be *strictly separated* if there exists disjoint open half-planes S_A and S_B of $\mathcal{V}$ such that $A \subseteq S_A$ and $B \subseteq S_B$.

Remark 4.4.7. Clearly if two subsets A and B are separated by the closed half-planes S_A and S_B, then there must exist an $f \in \mathcal{V}^*$ and a $\kappa \in \mathbb{R}$ such that

$$S_A = \{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) \geq \kappa\} \qquad \text{and} \qquad S_B = \{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) \leq \kappa\}.$$

Similarly, if two subsets A and B are strictly separated by the closed half-planes S_A and S_B, then there must exist an $f \in \mathcal{V}^*$ and a $\kappa \in \mathbb{R}$ such that

$$S_A = \{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) > \kappa\}$$

and then S_B can be replaced with $\{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) < \kappa\}$.

Example 4.4.8. In $\mathbb{R}^2$, consider the sets

$$\begin{aligned} A &= \left\{(x, y) \,\middle|\, x > 0, y \geq \frac{1}{x}\right\}, \\ B &= \left\{(x, y) \,\middle|\, x > 0, y \leq -\frac{1}{x}\right\}, \\ C &= \{(x, 0) \mid x > 0\}. \end{aligned}$$

If

$$S_1 = \{(x, y) \mid y > 0\} \quad \text{and} \quad S_2 = \{(x, y) \mid y < 0\},$$

then A and B are strictly separated by S_1 and S_2 whereas A and C are separated by $\overline{S_1}$ and $\overline{S_2}$. It can be verify that A and C cannot be strictly separated.

Notice the sets described above are convex sets. Our first step to separate such sets in any locally convex topological vector space is the following.

Proposition 4.4.9. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. If U is a non-empty, open, convex subset of $\mathcal{V}$ that does not contain $\vec{0}$, then there exists a closed hyperplane $\mathcal{W}$ such that $U \cap \mathcal{W} = \emptyset$.*

Proof. To proceed in the case that $\mathbb{K} = \mathbb{R}$, fix $x_0 \in U$. Then $C = x_0 - U$ is an open convex set containing $\vec{0}$ such that $x_0 \notin C$ (since $\vec{0} \notin G$).

Consider the Minkowski functional p_C, which is a sublinear functional by Proposition 3.6.12. Since $x_0 \notin C$, we obtain that $p_C(x_0) \geq 1$.

Let $\mathcal{M} = \mathrm{span}(\{x_0\})$ and define $f : \mathcal{M} \to \mathbb{R}$ by $f(\alpha x_0) = \alpha$ for all $\alpha \in \mathbb{R}$. Clearly f is linear. We claim that $f(x) \leq p_C(x)$ for all $x \in \mathcal{M}$. Indeed if $f(0) = 0 \leq p_C(x)$, if $\alpha > 0$ then

$$f(\alpha x_0) = \alpha \leq \alpha p_C(x_0) = p_C(\alpha x_0),$$

and if $\alpha < 0$ then

$$f(\alpha x_0) = \alpha < 0 \leq p_C(\alpha x_0)$$

as desired. By the Hahn-Banach Extension Theorem (Theorem 4.2.2), there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $g|_{\mathcal{M}} = f$ and $g(x) \leq p_C(x)$ for all $x \in \mathcal{V}$.

Notice for all $x \in C$ that

$$\mathrm{Re}(g)(x) = g(x) \leq p_C(x) \leq 1.$$

Hence Proposition 4.1.12 implies that g is uniformly continuous. Therefore $\mathcal{W} = \ker(g)$ is a hyperplane in $\mathcal{V}$.

To complete the case $\mathbb{K} = \mathbb{R}$, suppose $y \in U \cap \mathcal{W}$. As $y \in U$, $x_0 - y \in C$ so $p_C(x_0 - y) < 1$ and thus

$$g(x_0) - g(y) = g(x_0 - y) \leq p_C(x_0 - y) < 1$$

Moreover, $g(y) = 0$ as $y \in \mathcal{W}$ so the above implies $1 = f(x_0) = g(x_0) < 1$, which is a clear contradiction. Hence $U \cap \mathcal{W} = \emptyset$ as desired.

In the case that $\mathbb{K} = \mathbb{C}$, recall that $\mathcal{V}$ is also a locally convex topological vector space over $\mathbb{R}$ and thus the previous case implies there exists a continuous non-zero $\mathbb{R}$-linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $\ker(g) \cap U = \emptyset$.

Let $g_{\mathbb{C}} : \mathcal{V} \to \mathbb{C}$ be as defined in Lemma 4.1.5; that is, $g_{\mathbb{C}}(x) = g(x) - ig(ix)$ for all $x \in \mathcal{V}$. Then $g_{\mathbb{C}}$ is $\mathbb{C}$-linear so $\mathcal{W} = \ker(g_{\mathbb{C}})$ is a hyperplane in $\mathcal{V}$. As $x \in \ker(g_{\mathbb{C}})$ if and only if $x, ix \in \ker(g)$, we see that

$$\mathcal{W} = (\ker(g)) \cap (-i \ker(g))$$

is the intersection of two closed subsets (as g is continuous) and thus $\mathcal{W}$ is a closed hyperplane in $\mathcal{V}$. Since

$$\mathcal{W} \cap U \subseteq \ker(g) \cap U = \emptyset,$$

the proof is complete. ∎

Using this, we arrive at our first version of a Hahn-Banach Separation Theorem. Note Example 4.4.8 shows why 'open' cannot be removed from either set, and some basic thought about $\mathbb{R}^2$-geometry shows why 'convex' cannot be removed.

Theorem 4.4.10 (Hahn-Banach Separation Theorem - Open in Real Vector Spaces). *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space over $\mathbb{R}$. If A and B are non-empty, disjoint, open, convex subsets of $\mathcal{V}$, then there exists an $f \in \mathcal{V}^*$ and a $\kappa \in \mathbb{R}$ such that*

$$f(a) > \kappa > f(b)$$

for all $a \in A$ and $b \in B$. In particular, A and B are strictly separated.

Proof. Let

$$C = A - B = \{a - b \mid a \in A, b \in B\}.$$

Clearly $C \neq \emptyset$ as A and B are non-empty. Moreover, as $A \cap B = \emptyset$, we see that $\vec{0} \notin C$. Since $C = \bigcup_{b \in B} (-b) + A$ is a union of open sets since A is open, we obtain that C is open. Finally, we claim that C is convex. To see this, suppose $t \in [0, 1]$ and $c_1, c_2 \in C$. Hence there exist $a_1, a_2 \in A$ and $b_1, b_2 \in B$ such that $c_1 = a_1 - b_1$ and $c_2 = a_2 - b_2$. Since A and B are convex, we know that $ta_1 + (1 - t)a_2 \in A$ and $tb_1 + (1 - t)b_2 \in B$ so

$$tc_1 + (1 - t)c_2 = (ta_1 + (1 - t)a_2) - (tb_1 + (1 - t)b_2) \in A - B.$$

Therefore, as c_1, c_2, and t were arbitrary, C is a non-empty, open, convex subset of $\mathcal{V}$ that does not contain $\vec{0}$.

By Proposition 4.4.9 there exists a closed hyperplane $\mathcal{W}$ in $(\mathcal{V}, \mathcal{T})$ such that $C \cap \mathcal{W} = \emptyset$. By Remark 4.1.7 there exists an $f \in \mathcal{V}^\sharp$ such that $\mathcal{W} = \ker(f)$ and thus $f \in \mathcal{V}^*$ by Proposition 4.1.10.

Since $f \in \mathcal{V}^*$ and C is convex, $f(C)$ is a convex subset of $\mathbb{R}$ by Lemma 3.2.21. Moreover, since $C \cap \mathcal{W} = C \cap \ker(f) = \emptyset$, $0 \notin f(C)$. Hence either $f(C) \subseteq (0, \infty)$ or $f(C) \subseteq (-\infty, 0)$. By replacing f with $-f$ if necessary, we may assume that $f(C) \subseteq (0, \infty)$.

By the definition of $C = A - B$, we obtain that $f(a) - f(b) > 0$ for all $a \in A$ and $b \in B$. Hence there exists an $\kappa \in \mathbb{R}$ such that

$$\sup(\{f(b) \mid b \in B\}) \leq \kappa \leq \inf(\{f(a) \mid a \in A\}).$$

It remains only to show that $f(b) < \kappa < f(a)$ for all $b \in B$ and $a \in A$.

Recall $\mathcal{V}/\mathcal{W}$ is a locally convex topological space by Proposition 3.3.6. Moreover, since $\mathcal{W} = \ker(f)$ is a hyperplane, $\mathcal{V}/\mathcal{W}$ is one-dimensional. Let $x_0 \in \mathcal{V}$ be any vector such that $f(x_0) = 1$. Hence $x_0 \notin \mathcal{W}$ so $\{x_0 + \mathcal{W}\}$ is a basis for $\mathcal{V}/\mathcal{W}$ and the map $T : \mathcal{V}/\mathcal{W} \to \mathbb{R}$ defined by

$$T(\alpha x_0 + \mathcal{W}) = \alpha$$

for all $\alpha \in \mathbb{R}$ is a homeomorphism by Lemma 3.5.1.

Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the canonical quotient map. Notice for all $x \in \mathcal{V}$ there exists a unique $\alpha_x \in \mathbb{R}$ such that $q(x) = \alpha_x x_0 + \mathcal{W}$ and thus

$$f(x) = f(\alpha_x x_0) = \alpha_x = T(q(x)).$$

Since q is an open map by Lemma 3.3.4, $q(A)$ and $q(B)$ are open subsets of $\mathcal{V}/\ker(f)$. Therefore, since T is a homeomorphism, we obtain that $f(A) = T(q(A))$ and $f(B) = T(q(B))$ are open subsets of $\mathbb{R}$. Hence $f(b) < \kappa < f(a)$ for all $b \in B$ and $a \in A$ so if

$$S_A = \{x \in \mathcal{V} \mid f(x) > \kappa\} \quad \text{and} \quad S_B = \{x \in \mathcal{V} \mid f(x) < \kappa\},$$

then S_A and S_B are disjoint open half-spaces containing A and B respectively. Hence A and B are strictly separated. ∎

To deal with the complex case, we simply invoke our connection between real and complex linear functionals.

Theorem 4.4.11 (Hahn-Banach Separation Theorem - Open in Complex Vector Spaces). *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space over $\mathbb{C}$. If A and B are non-empty, disjoint, open, convex subsets of $\mathcal{V}$, then there exists an $f \in \mathcal{V}^*$ and a $\kappa \in \mathbb{R}$ such that*

$$\mathrm{Re}(f(a)) > \kappa > \mathrm{Re}(f(b))$$

for all $a \in A$ and $b \in B$. In particular, A and B are strictly separated.

Proof. Recall $(\mathcal{V}, \mathcal{T})$ can also be viewed as a locally convex topological vector space over $\mathbb{R}$. Hence by Theorem 4.4.10 there exists a continuous $\mathbb{R}$-linear functional $g : \mathcal{V} \to \mathbb{R}$ and a $\kappa \in \mathbb{R}$ such that

$$B \subseteq \{x \in \mathcal{V} \mid g(x) < \kappa\} \qquad \text{and} \qquad A \subseteq \{x \in \mathcal{V} \mid g(x) > \kappa\}.$$

By Lemma 4.1.5 and together with Corollary 3.6.18 implies there exists an $f \in \mathcal{V}^*$ such that $\mathrm{Re}(f) = g$. Hence the result follows. ∎

Of course, if we want to reduce the 'strictly separated' conclusion to just 'separated', we can prove the following.

Corollary 4.4.12. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. If A and B are non-empty, disjoint, convex subsets of $\mathcal{V}$ such that A is open, then there exists $f \in \mathcal{V}^*$ and a $\kappa \in \mathbb{R}$ such that*

$$A \subseteq \{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) > \kappa\} \qquad \textit{and} \qquad B \subseteq \{x \in \mathcal{V} \mid \mathrm{Re}(f(x)) \leq \kappa\}.$$

Proof. As the only part where we required B to be open in the proof of Theorem 4.4.10 was to deduce $f(B)$ was open in order to obtain strict separation, the result follows. ∎

It turns out we can prove something stronger. Indeed, being able to separate closed subsets of locally convex topological vector space will be to functional analysis as being able to separate closed sets is to topology (the topological concept known as a normal topology, which is vital in many essential results such as Urysohn's Lemma). However, by Example 4.4.8 we know it might be impossible to strictly separate closed sets. Luckily if at least one set is compact, we do not have an issue. Again, considering $\mathbb{R}^2$, it is not difficult to find examples where convexity is required.

Theorem 4.4.13 (Hahn-Banach Separation Theorem - Closed). *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. If A and B are non-empty, disjoint, closed, convex subsets of $\mathcal{V}$ such that B is compact, then there exists an $f \in \mathcal{V}^*$ and $\alpha, \beta \in \mathbb{R}$ such that*

$$\mathrm{Re}(f(a)) \geq \alpha > \beta \geq \mathrm{Re}(f(b))$$

for all $a \in A$ and $b \in B$. In particular, A and B are strictly separated.

Proof. Our goal is to invoke the previous versions of the Hahn-Banach Separation Theorem by constructing disjoint open convex sets containing A and B. This will be accomplished using local convexity and the compactness for B. In particular, it is not difficult to construct an open set containing B that is disjoint from A. By 'scaling back by a third', we can construct the open set around A.

Since $\mathcal{V} \setminus A$ is an open set containing B, we have by Corollary 3.6.4 that for each $b \in B$ there exists a balanced, convex neighbourhood U_b of $\vec{0}$ such that $b + U_b \subseteq \mathcal{V} \setminus A$. Since

$$\left\{ b + \frac{1}{3}U_b \,\middle|\, b \in B \right\}$$

is an open cover of B and since B is compact, there exists an $n \in \mathbb{N}$ and $b_1, \ldots, b_n \in B$ such that

$$B \subseteq \bigcup_{k=1}^{n} b_k + \frac{1}{3}U_{b_k}.$$

Let $U = \bigcap_{k=1}^{n} \frac{1}{3}U_{b_k}$, which is an intersection of balanced, convex neighbourhoods of $\vec{0}$ and thus a balanced, convex neighbourhood of $\vec{0}$. Furthermore, let

$$A_0 = A + U = \{a + u \mid a \in A, u \in U\}$$
$$B_0 = B + U = \{b + u \mid b \in A, u \in U\}.$$

Clearly $A \subseteq A_0$ and $B \subseteq B_0$ since $\vec{0} \in U$. Thus A_0 and B_0 are non-empty since A and B are non-empty. Moreover, since

$$A_0 = \bigcup_{a \in A} a + U \qquad \text{and} \qquad B_0 = \bigcup_{b \in B} b + U,$$

we have A_0 and B_0 are unions of open sets and thus open. Finally, we claim that A_0 and B_0 are convex. To see this, let $x_1, x_2 \in A_0$ and $t \in [0, 1]$ be arbitrary. Hence there exists $a_1, a_2 \in A$ and $u_1, u_2 \in U$ such that $x_1 = a_1 + u_1$ and $x_2 = a_2 + u_2$. Therefore, since $ta_1 + (1-t)a_2 \in A$ and $tu_1 + (1-t)u_2 \in U$ as A and U are convex, we obtain that

$$tx_1 + (1-t)x_2 = (ta_1 + (1-t)a_2) + (tu_1 + (1-t)u_2) \in A + U = A_0.$$

Therefore, since x_1, x_2, and t were arbitrary, we obtain that A_0 is convex. A nearly identical argument shows that B_0 is convex.

Finally, we claim that A_0 and B_0 are disjoint. To see this, suppose to the contrary that $A_0 \cap B_0 \neq \emptyset$. Hence, there exists $a \in A$, $b \in B$ and $u_1, u_2 \in U$ such that $a + u_1 = b + u_2$. Since $b \in B$, the above construction yields a $k \in \{1, \ldots, n\}$ such that $b \in b_k + \frac{1}{3}U_{b_k}$. Therefore,

$$\begin{aligned} a = b + (u_1 - u_2) &\in \left(b_k + \frac{1}{3}U_{b_k}\right) + \left(\frac{1}{3}U_{b_k} - \frac{1}{3}U_{b_k}\right) && \text{since } U \subseteq \frac{1}{3}U_{b_k} \\ &= b_k + \frac{1}{3}U_{b_k} + \frac{1}{3}U_{b_k} + \frac{1}{3}U_{b_k} && \text{since } U_{b_k} \text{ is balanced} \\ &= b_k + U_{b_k} && \text{by Lemma 3.2.21 since } U_{b_k} \text{ is convex.} \end{aligned}$$

However, this contradicts the fact that $b_k + U_{b_k} \subseteq \mathcal{V} \setminus A$. Hence A_0 and B_0 are disjoint.

Consequently, Theorem 4.4.10 when $\mathbb{K} = \mathbb{R}$ and Theorem 4.4.11 when $\mathbb{K} = \mathbb{C}$ implies there exists an $f \in \mathcal{V}^*$ and a $\alpha \in \mathbb{R}$ such that

$$\mathrm{Re}(f(a)) > \alpha > \mathrm{Re}(f(b))$$

for all $a \in A_0$ and $b \in B_0$. In particular, as $A \subseteq A_0$ and $B \subseteq B_0$, we obtain that

$$\mathrm{Re}(f(a)) > \alpha > \mathrm{Re}(f(b))$$

for all $a \in A$ and $b \in B$. However, as B is compact, $f(B)$ is a compact subset of $\mathbb{K}$ so $\mathrm{Re}(f(B))$ is a compact subset of $\mathbb{R}$ so if

$$\beta = \sup(\{\mathrm{Re}(f(b)) \mid b \in B\}) \in \mathbb{R},$$

then

$$\mathrm{Re}(f(a)) \geq \alpha > \beta \geq \mathrm{Re}(f(b))$$

for all $a \in A$ and $b \in B$ as desired. ∎

To conclude this section, we will use Theorem 4.4.13 to describe all closed convex subsets of a locally convex topological vector space via continuous linear functionals as a generalization on how Theorem 4.3.11 described closed subspaces via continuous linear functionals. This is done via the following useful concept in functional analysis.

Definition 4.4.14. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $A \subseteq \mathcal{V}$. The *closed convex hull of* A, denoted $\overline{\mathrm{conv}}(A)$, is the closure of the convex hull of A.

Corollary 4.4.15. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space. If $A \subseteq \mathcal{V}$ is non-empty, then $\overline{\mathrm{conv}}(A)$ is the intersection of all closed half-spaces that contain A.*

Proof. Let

$$\mathcal{I} = \{S \subseteq \mathcal{V} \mid A \subseteq S, S \text{ a closed half-space in } \mathcal{V}\}.$$

As every $S \in \mathcal{I}$ is convex by Remark 4.4.2 and contains A, we easily obtain that

$$\mathrm{conv}(A) \subseteq \bigcap_{S \in \mathcal{I}} S$$

as the intersection of convex sets is convex and thus

$$\overline{\mathrm{conv}}(A) \subseteq \bigcap_{S \in \mathcal{I}} S$$

as the intersection of closed sets is closed.

To demonstrate the reverse inclusion, let $x \in \mathcal{V} \setminus \overline{\mathrm{conv}}(A)$ be arbitrary. Since $\{x\}$ and $\overline{\mathrm{conv}}(A)$ are non-empty, disjoint, closed, convex subsets of $\mathcal{B}$ such that $\{x\}$ is compact, Theorem 4.4.13 implies there exists a closed half-space S_0 such that $A \subseteq \overline{\mathrm{conv}}(A) \subseteq S_0$ and $x \notin S_0$. Hence $S_0 \in \mathcal{I}$ so $x \notin \bigcap_{S \in \mathcal{I}} S$. Therefore, as x was arbitrary, the proof is complete. ∎

Corollary 4.4.16. *Let $\mathcal{V}$ be a vector space and let $\mathcal{T}_1$ and $\mathcal{T}_2$ be topologies on $\mathcal{V}$ such that $(\mathcal{V}, \mathcal{T}_1)$ and $(\mathcal{V}, \mathcal{T}_2)$ are locally convex topological vector spaces. If $(\mathcal{V}, \mathcal{T}_1)$ and $(\mathcal{V}, \mathcal{T}_2)$ have the same continuous linear functionals, then $(\mathcal{V}, \mathcal{T}_1)$ and $(\mathcal{V}, \mathcal{T}_2)$ have the same closed convex sets. Moreover, if $A \subseteq \mathcal{V}$, then $\overline{\mathrm{conv}(A)}^{\mathcal{T}_1} = \overline{\mathrm{conv}(A)}^{\mathcal{T}_2}$.*

Proof. This follows immediately from Corollary 4.4.15 do to the correspondence between closed half-spaces and linear functionals. ∎

Chapter 5

Dual Space Topologies

As seen in the previous chapter, the Hahn-Banach Theorems have the powerful ability to extend continuous linear functionals thereby answering some of our earlier questions and can be used to separate specific subsets via affine hyperplanes. Thus we desire to delve deeper into the functional analytical aspects dual spaces produce.

First, we recall by Chapter 3 that any separating family of seminorms on a vector space produces a locally convex topological vector space structure and Example 3.2.18 combined with the Hahn-Banach Theorem (Corollary 4.3.5) produces a nice separating family of seminorms based on the dual space. Thus our first goal of this chapter is to analyze such topologies induced by and on dual spaces. In particular, we will be able to complete determine the continuous linear functionals on such spaces.

Subsequently we will transition to proving three of the most fundamental results in elementary functional analysis. First we will examine the Banach Alaoglu Theorem(Theorem 5.3.4) which determines the closed unit ball is actually compact with respect to a certain dual topology. This is quite striking as Theorem 3.5.11 shows the only locally compact topological vector spaces are finite dimensional (for which our result does not contradict as the closed balls do not form a neighbourhood bases in this topology). This shows such topologies are easier to handle and should have many commonalities with finite dimensional spaces.

Subsequently, we will examine Goldstine's Theorem (Theorem 5.4.1) which shows the image of the canonical embedding is dense in a certain topology we can place on the double dual. Finally, we end with the Krein-Milman Theorem (Theorem 5.5.12) for which we can describe any non-empty compact convex set via the points "on the boundary".

5.1 Weak and Weak* Topologies

To begin, we shall reintroduce the weak and weak* topologies in a more general context. Basically, any separating family of linear functionals on a space can be taken and a topology can be produced. Thus we introduce such pairings as follows.

Definition 5.1.1. A *dual pair* consists of pair $(\mathcal{V}, \mathcal{L})$ where $\mathcal{V}$ is a vector space over $\mathbb{K}$ and $\mathcal{L} \subseteq \mathcal{V}^\sharp$ is a vector subspace that separates points.

Of course, if $\mathcal{L} \subseteq \mathcal{V}^\sharp$ is just a set of linear functionals that separates points, one can always take $\mathrm{span}(\mathcal{L})$ to construct a dual pair.

Remark 5.1.2. Notice if $(\mathcal{V}, \mathcal{L})$ is a dual pair, then for each $f \in \mathcal{L}$ we can define a seminorm $p_f : \mathcal{V} \to [0, \infty)$ by

$$p_f(x) = |f(x)|$$

for all $x \in \mathcal{V}$. The family $\mathcal{F} = \{p_f \mid f \in \mathcal{L}\}$ is a separating family of seminorms on $\mathcal{V}$ and thus defines a topology $\mathcal{T}$ on $\mathcal{V}$ such that $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space by Theorem 3.2.10. Recall a topological base for $\mathcal{T}$ is

$$\mathcal{B} = \{N(x, F, \epsilon) \mid x \in \mathcal{V}, \epsilon > 0, F \subseteq \mathcal{F} \text{ finite}\}$$

where

$$N(x, F, \epsilon) = \{y \in \mathcal{V} \mid |f(y) - f(x)| < \epsilon \text{ for all } f \in F\},$$

and, by Proposition 3.2.13, a net $(x_\lambda)_{\lambda \in \Lambda}$ converges to x with respect to $\mathcal{T}$ if and only if $\lim_{n\to\infty} f(x_\lambda) = f(x)$ for all $f \in \mathcal{L}$.

Definition 5.1.3. Given a dual pair $(\mathcal{V}, \mathcal{L})$, the topology $\mathcal{T}$ on $\mathcal{V}$ from Remark 5.1.2 is called the *weak topology generated by* $\mathcal{L}$ and is denoted by $\sigma(\mathcal{V}, \mathcal{L})$.

The reason this topology is called "the weak topology" is that it is the weakest topology on $\mathcal{V}$ for which the elements of $\mathcal{L}$ are continuous as the following result demonstrates.

Theorem 5.1.4. *If $(\mathcal{V}, \mathcal{L})$ is a dual pair, then*

$$\mathcal{L} = (\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))^*.$$

That is, $\mathcal{L}$ is exactly the collection of continuous linear functionals on $\mathcal{V}$ with respect to the weak topology generated by $\mathcal{L}$.

Proof. As shown above, by Proposition 3.2.13 if a net $(x_\lambda)_{\lambda\in\Lambda}$ converges to x with respect to $\mathcal{T}$ then $\lim_{n\to\infty} f(x_\lambda) = f(x)$ for all $f \in \mathcal{L}$. Thus every element of $\mathcal{L}$ is a continuous linear functional so $\mathcal{L} \subseteq (\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))^*$.

Conversely, suppose $f \in (\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))^*$. Hence $x \mapsto |f(x)|$ is a continuous seminorm on $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))$. As the absolute values of elements of $\mathcal{L}$ is a family of seminorms on $\mathcal{V}$ that generated $\sigma(\mathcal{V}, \mathcal{L})$, Proposition 3.6.16 implies there exists an $M > 0$ and $f_1, \ldots, f_n \in \mathcal{L}$ such that

$$|f(x)| \leq M \max(\{|f_1(x)|, \ldots, |f_n(x)|\})$$

for all $x \in \mathcal{V}$. Thus $\bigcap_{k=1}^n \ker(f_k) \subseteq \ker(f)$ so Lemma 4.1.9 implies that

$$f \in \mathrm{span}(\{f_1, \ldots, f_n\}) \subseteq \mathcal{L}.$$

Hence $\mathcal{L} = (\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))^*$. ∎

Of course, we can use the above to extend our notion of the weak topology on a normed linear space to any locally convex topological vector space.

Example 5.1.5. Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{L} = \mathcal{V}^*$. Then $(\mathcal{V}, \mathcal{V}^*)$ is a dual pair by Corollary 4.3.8 and thus $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ is a locally convex topological vector space by Remark 5.1.2. As this is the most common and important weak topology for a locally convex topological vector space, opposed to calling $\sigma(\mathcal{V}, \mathcal{V}^*)$ "the weak topology generated by $\mathcal{V}^*$" we refer to it as *the weak topology on* $\mathcal{V}$.

Moreover, to simplify terminology, we say that a net $(x_\lambda)_{\lambda\in\Lambda}$ *weakly* converges to an $x \in \mathcal{V}$ if $(x_\lambda)_{\lambda\in\Lambda}$ converges to x with respect to $\sigma(\mathcal{V}, \mathcal{V}^*)$; that is,

$$\lim_{\lambda\in\Lambda} f(x_\lambda) = f(x)$$

for all $f \in \mathcal{V}^*$. Notice if $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(\mathcal{V}, \mathcal{T})$, then, by continuity, for all $f \in \mathcal{V}^*$ we have that

$$\lim_{\lambda\in\Lambda} f(x_\lambda) = f(x)$$

and thus $(x_\lambda)_{\lambda\in\Lambda}$ weakly converges to x. Thus the weak topology is indeed weaker than the initial topology.

Finally, we note by Theorem 5.1.4 that the *weakly continuous linear functionals on* $\mathcal{V}$ (that is, the elements of $\mathcal{V}^\sharp$ that are continuous with respect to this topology) are exactly $\mathcal{V}^*$.

To give some examples of how the weak topology behaves, we consider the following.

Example 5.1.6. Given any finite dimensional vector space $\mathcal{V}$ and any subspace $\mathcal{L} \subseteq \mathcal{V}^\sharp$ that separates points, we have that $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{L}))$ is automatically a locally convex topological vector space and thus must be $(\mathbb{K}^n, \|\cdot\|_\infty)$ in disguise (where $n = \dim(\mathcal{V})$. That is, the weak and norm topologies coincide on any finite dimensional normed linear space.

Example 5.1.7. Consider $p \in (1, \infty)$. We claim that the weak topology on $\ell_p(\mathbb{N})$ does not agree with the norm topology. To see this, consider the sequence $(e_n)_{n\geq 1}$ where $e_n \in \ell_p(\mathbb{N})$ is the sequence with a 1 in the n^{th} position and zeros everywhere else. Clearly $e_n \in \ell_p(\mathbb{N})$.

We claim that $(e_n)_{n\geq 1}$ converges weakly to $\vec{0}$. To see this, recall by Theorem 1.5.4 that $\ell_p(\mathbb{N})^* = \ell_q(\mathbb{N})$ where $q \in (1, \infty)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$ via the map $\Phi : \ell_q(\mathbb{N}) \to \ell_p(\mathbb{N})^*$ where for all $\vec{y} = (y_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ and $\vec{x} = (x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ we have that

$$\Phi(\vec{y})(\vec{x})) = \sum_{n=1}^{\infty} x_n y_n.$$

Thus, for all $\vec{y} = (y_n)_{n\geq 1} \in \ell_q(\mathbb{N})$ and $q \neq \infty$ we see that

$$\Phi(\vec{y})(e_n) = y_n.$$

Hence, as $\vec{y} \in \ell_q(\mathbb{N})$, we obtain that

$$\lim_{n\to\infty} \Phi(\vec{y})(e_n) = 0 = \Phi(\vec{y})(\vec{0}).$$

Thus $(e_n)_{n\geq 1}$ converges weakly to $\vec{0}$

As clearly $(e_n)_{n\geq 1}$ does not converge to $\vec{0}$ in norm as $\|e_n\| = 1$ for all $n \in \mathbb{N}$, we see that the weak and norm topologies on $\ell_p(\mathbb{N})$ disagree as they have different convergent sequences.

Example 5.1.7 raises the question, "What about the weak topology $\ell_1(\mathbb{N})$?" The following sequence of results will show that the norm topology is strictly finer (meaning contains more open sets) than the weak topology, yet the norm and weak topology have the same convergent sequences! This is one of the most fundamental examples of why examining nets is required in topology.

Proposition 5.1.8. *The norm topology on $\ell_1(\mathbb{N})$ is finer than the weak topology on $\ell_1(\mathbb{N})$.*

Proof. To show that the norm topology is finer than the weak topology, it suffices to show for any $\vec{x} \in \ell_1(\mathbb{N})$ and any weak neighbourhood V of $\vec{x}$ there exists a norm neighbourhood U of $\vec{x}$ such that $U \subseteq V$. By Theorem 1.5.4 we know that $\ell_1(\mathbb{N})^* = \ell_\infty(\mathbb{N})$ and by Theorem 3.2.10 it suffices to consider the neighbourhoods

$$\begin{aligned} &N(\vec{x}, \{\vec{z}_k\}_{k=1}^m, \epsilon) \\ &\quad = \left\{ (y_n)_{n\geq 1} \in \ell_1(\mathbb{N}) \;\middle|\; \left| \sum_{n=1}^{\infty} (x_n - y_n) z_{k,n} \right| < \epsilon \text{ for all } k \in \{1, \dots, m\} \right\} \end{aligned}$$

of $\vec{x}$ where $\epsilon > 0$ and $\vec{z}_k = (z_{k,n})_{n\geq 1} \in \ell_\infty(\mathbb{R})$ for $k \in \{1, \dots, m\}$.

For ϵ and $\{\vec{z}_k\}_{k=1}^m$ as above, let

$$M = \max\{\|\vec{z}_k\|_\infty \mid k \in \{1, \ldots, m\}\} > 0.$$

We claim that

$$B\left(\vec{x}, \frac{\epsilon}{M+1}\right) \subseteq N(\vec{x}, \{\vec{z}_k\}_{k=1}^m, \epsilon).$$

Indeed if $y = (y_n)_{n\geq 1} \in B\left(\vec{x}, \frac{\epsilon}{M+1}\right)$, then

$$\sum_{n=1}^\infty |x_n - y_n| < \frac{\epsilon}{M+1}.$$

Hence for all $k \in \{1, \ldots, m\}$ we see that

$$\left|\sum_{n=1}^\infty (x_n - y_n) z_{k,n}\right| \leq \sum_{n=1}^\infty |x_n - y_n||z_{k,n}| \leq \sum_{n=1}^\infty |x_n - y_n| M < \frac{\epsilon}{M+1} M < \epsilon$$

so $\vec{y} \in N(\vec{x}, \{\vec{z}_k\}_{k=1}^m, \epsilon)$. Thus the claim is complete so the norm topology is finer than the weak topology. ∎

Although we could use Theorem 5.4.6 together with the facts that $\ell_1(\mathbb{N})^* = \ell_\infty(\mathbb{N})$ and $\ell_\infty(\mathbb{N})$ is not separable to show that the weak topology on $\ell_1(\mathbb{N})$ is indeed not induced by a norm, there is a simpler proof of this fact.

Proposition 5.1.9. *The weak topology on $\ell_1(\mathbb{N})$ is not a topology induced by a norm. Hence the weak and norm topologies on $\ell_1(\mathbb{N})$ differ.*

Proof. Suppose the weak topology is induced by a norm $\|\cdot\|_w$ and let U be the $\|\cdot\|_w$-ball of radius 1 centred at the zero vector $\vec{0}$. By the definition of the weak topology and by Theorem 1.5.4 there must exist an $m \in \mathbb{N}$, $\vec{z}_k = (z_{k,n})_{n\geq 1} \in \ell_\infty(\mathbb{N})$ for $k \in \{1, \ldots, m\}$, and an $\epsilon > 0$ such that $N(\vec{0}, \{\vec{z}_k\}_{k=1}^m, \epsilon) \subseteq U$.

For each $k \in \{1, \ldots, m\}$ let

$$\vec{v}_k = (z_{k,1}, z_{k,2}, \ldots, z_{k,m+1}) \in \mathbb{K}^{m+1}.$$

Then the set $\{\vec{v}_k \mid k \in \{1, \ldots, m\}\}$ is a set with m vectors in $\mathbb{K}^{m+1}$. Hence there exists a non-zero vector $\vec{v} = (z_1, z_2, \ldots, z_{m+1}) \in \mathbb{K}^{m+1}$ such that

$$0 = \vec{v}_k \cdot \overline{\vec{v}} = \sum_{j=1}^{m+1} z_j z_{k,j}$$

for all $k \in \{1, \ldots, m\}$.

Since $\vec{v} \in \mathbb{K}^{m+1}$ is non-zero, if we define $x_j = 0$ if $j > m+1$, then the sequence $\vec{x} = (x_n)_{n\geq 1}$ is a non-zero element of $\ell_\infty(\mathbb{N})$ such that for all $t \in \mathbb{R}$

$$\sum_{n=1}^{\infty} t x_n z_{k,n} = 0$$

for all $k \in \{1, \ldots, m\}$. Hence

$$t\vec{x} \in N(\vec{0}, \{\vec{z}_k\}_{k=1}^{m}, \epsilon) \subseteq U.$$

However, since U is the $\|\cdot\|_w$-ball of radius 1 centred at the zero vector $\vec{0}$, $t\vec{x} \in U$ for all $t \in \mathbb{R}$ implies that $|t|\,\|\vec{x}\| = \|t\vec{x}\| < 1$ for all $t \in \mathbb{R}$. However, this is impossible as $\vec{x} \neq \vec{0}$ so $\|\vec{x}\| > 0$. Thus we have a contradiction so $\mathcal{T}_w$ cannot be induced by a norm. ∎

Now onto the main result that the norm and weak topologies on $\ell_1(\mathbb{N})$ have the same convergent sequences.

Theorem 5.1.10. *In $\ell_1(\mathbb{N})$, a sequence $(\vec{x}_n)_{n\geq 1}$ converges to a point $\vec{x}$ in the norm topology if and only if it converges in the weak topology.*

Proof. Let $(\vec{x}_k)_{k\geq 1}$ be a sequence in $\ell_1(\mathbb{N})$. By Proposition 5.1.8, we know that if $(\vec{x}_n)_{n\geq 1}$ converges to a point $\vec{x}$ in the norm topology, then it converges to the same point in the weak topology.

To see the converse, notice that $(\vec{x}_k)_{k\geq 1}$ converges to a vector $\vec{x} \in \ell_1(\mathbb{N})$ in the norm topology if and only if $\lim_{k\to\infty} \|\vec{x}_k - \vec{x}\|_1 = 0$ if and only if the sequence $(\vec{x}_k - \vec{x})_{k\geq 1}$ converges to $\vec{0}$ in the norm topology. Similarly $(\vec{x}_k)_{k\geq 1}$ converges to a vector $\vec{x} \in \ell_1(\mathbb{R})$ in the weak topology if and only if the sequence $(\vec{x}_k - \vec{x})_{k\geq 1}$ converges to $\vec{0}$ in the weak topology. Thus we need only consider sequences that converge to zero in the weak topology.

Let $(\vec{x}_k)_{k\geq 1}$ be a sequence of elements in $\ell_1(\mathbb{N})$ that converges to $\vec{0}$ in the weak topology. To see that $(\vec{x}_k)_{k\geq 1}$ converges to $\vec{0}$ in the norm topology, suppose to the contrary that $(\vec{x}_k)_{k\geq 1}$ does not converge to $\vec{0}$ in the norm topology. Thus there exists an $\epsilon > 0$ and a subsequence $(\vec{x}_{k_j})_{j\geq 1}$ such that $\left\|\vec{x}_{k_j}\right\|_1 \geq \epsilon$ for all $j \in \mathbb{N}$. By replacing $(\vec{x}_k)_{k\geq 1}$ with $(\vec{x}_{k_j})_{j\geq 1}$ if necessary, we may assume that $(\vec{x}_k)_{k\geq 1}$ converges to $\vec{0}$ in the weak topology and that there exists an $\delta > 0$ such that $\|\vec{x}_k\| \geq \delta$ for all $k \in \mathbb{N}$.

Write $\vec{x}_k = (x_{k,n})_{n\geq 1}$ for all $k \in \mathbb{N}$. We claim for each $m \in \mathbb{N}$ that $\lim_{k\to\infty} x_{k,m} = 0$. Indeed fix $m \in \mathbb{N}$ and let $\vec{e}_m = (e_{m,n})_{n\geq 1} \in \ell_\infty(\mathbb{N})$ where

$$e_{m,n} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases}.$$

If $\Phi : \ell_\infty(\mathbb{N}) \to \ell_1(\mathbb{N})^*$ is the isomorphism from Theorem 1.5.4, then we know that

$$0 = \lim_{k\to\infty} \Phi(\vec{e}_m)(\vec{x}_k) = \lim_{k\to\infty} x_{k,m}$$

as desired.

Using the facts that $\|\vec{x}_k\| \geq \delta$ for all $k \in \mathbb{N}$ and that $\lim_{k\to\infty} x_{k,m} = 0$ for all $m \in \mathbb{N}$, we will obtain a contradiction to the fact that $(\vec{x}_k)_{k\geq 1}$ converges weakly to $\vec{0}$ by constructing a subsequence that does not converge weakly to $\vec{0}$. Let $k_1 = 1$ and let $n_1 \in \mathbb{N}$ be such that $\sum_{j=n_1+1}^{\infty} |x_{k_1,j}| < \frac{\delta}{6}$, which is possible since $\vec{x}_{k_1} \in \ell_1(\mathbb{N})$. As $\lim_{k\to\infty} x_{k,m} = 0$ for all $m \in \mathbb{N}$, there exists a $k_2 > k_1$ such that $\sum_{j=1}^{n_1} |x_{k,j}| < \frac{\delta}{6}$ for all $k \geq k_2$. Thus, as $\vec{x}_{k_2} \in \ell_1(\mathbb{R})$, there exists an $n_2 > n_1$ such that $\sum_{j=n_2+1}^{\infty} |x_{k_2,j}| < \frac{\delta}{6}$. By repeating the above construction inductively, we obtain increasing sequences $(k_m)_{m\geq 1}$ and $(n_m)_{m\geq 1}$ such that $\sum_{j=1}^{n_{m-1}} |x_{k,j}| < \frac{\delta}{6}$ for all $k \geq k_m$ and $\sum_{j=n_m+1}^{\infty} |x_{k_m,j}| < \frac{\delta}{6}$.

Consider the subsequence $(\vec{x}_{k_m})_{m\geq 1}$. As $(\vec{x}_k)_{k\geq 1}$ converges weakly to $\vec{0}$, $(\vec{x}_{k_m})_{m\geq 1}$ converges weakly to $\vec{0}$. Consider $\vec{y} = (y_n)_{n\geq 1} \in \ell_\infty(\mathbb{N})$ defined by

$$y_n = \begin{cases} \mathrm{sgn}(x_{1,n}) & \text{if } n \leq n_1 \\ \mathrm{sgn}(x_{k_m,n}) & \text{whenever } \quad n_{m-1}+1 \leq n \leq n_m \end{cases}$$

where

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x = 0 \\ \frac{\overline{x}}{|x|} & \text{if } x \neq 0 \end{cases}.$$

We claim that

$$\vec{x}_{k_m} \notin N\left(\vec{0}, \vec{y}, \frac{\delta}{3}\right)$$

for all $m \in \mathbb{N}$ thereby contradicting the fact that $(\vec{x}_{k_m})_{m\geq 1}$ converges weakly to $\vec{0}$. Indeed notice by construction that

$$\begin{aligned} \delta \leq \|\vec{x}_{k_m}\|_1 &= \sum_{j=1}^{n_{m-1}} |x_{k_m,j}| + \sum_{j=n_{m-1}+1}^{n_m} |x_{k_m,j}| + \sum_{j=n_m+1}^{\infty} |x_{k_m,j}| \\ &\leq \frac{\delta}{6} + \sum_{j=n_{m-1}+1}^{n_m} |x_{k_m,j}| + \frac{\delta}{6} \end{aligned}$$

so

$$\sum_{j=n_{m-1}+1}^{n_m} |x_{k_m,j}| \geq \frac{2\delta}{3}.$$

However, notice that

$$\begin{aligned} \left|\sum_{j=1}^{\infty} x_{k_m,j} y_j\right| &= \left|\sum_{j=1}^{n_{m-1}} x_{k_m,j} y_j + \sum_{j=n_{m-1}+1}^{n_m} |x_{k_m,j}| + \sum_{j=n_m+1}^{\infty} x_{k_m,j} y_j\right| \\ &\geq \sum_{j=n_{m-1}+1}^{n_m} |x_{k_m,j}| - \sum_{j=1}^{n_{m-1}} |x_{k_m,j}| - \sum_{j=n_m+1}^{\infty} |x_{k_m,j}| \\ &\geq \frac{2\delta}{3} - \frac{\delta}{6} - \frac{\delta}{6} = \frac{\delta}{3}. \end{aligned}$$

Hence $\vec{x}_{k_m} \notin N\left(\vec{0}, \vec{y}, \frac{\delta}{3}\right)$ for all $m \in \mathbb{N}$ thereby yielding a contradiction and the proof. ∎

As an immediate corollary, we have the following.

Corollary 5.1.11. *The weak topology on $\ell_1(\mathbb{N})$ is not metrizable.*

Proof. As convergent sequences completely determine a metric topology, if the weak topology on $\ell_1(\mathbb{N})$ was metrizable, then Theorem 5.1.10 would imply that the weak and norm topologies coincide, which contradicts Proposition 5.1.9. ∎

To examine other examples of convergent sequences in weak topologies, we note the following result which puts restrictions on the limits of a weakly convergent sequence.

Proposition 5.1.12. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. If $(x_n)_{n\geq 1}$ is a sequence in $\mathcal{X}$ that converges weakly to $x \in \mathcal{X}$, then*

- $\sup_{n\in\mathbb{N}} \|x_n\| < \infty$, *and*
- $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.

Proof. Let $(x_n)_{n\geq 1}$ be a sequence in $\mathcal{X}$ that converges weakly to $x \in \mathcal{X}$. Hence $f(x) = \lim_{n\to\infty} f(x_n)$ for all $f \in \mathcal{X}^*$.

Consider the canonical embedding $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ defined by $\mathcal{J}(x) = \widehat{x}$ as defined in Theorem 1.6.3, which is an isometry by the Hahn-Banach Theorem (Corollary 4.3.2). Since $(x_n)_{n\geq 1}$ converges weakly to x,

$$\widehat{x}(f) = f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \widehat{x_n}(f)$$

for all $f \in \mathcal{X}^*$. Thus the Uniform Boundedness Principle (Theorem 2.5.3) applied to the Banach space $\mathcal{X}^{**}$ implies that $(\widehat{x_n})_{n\geq 1}$ is a bounded sequence so $(x_n)_{n\geq 1}$ is bounded. Moreover, since

$$|\widehat{x}(f)| = \lim_{n\to\infty} |\widehat{x_n}(f)| \leq \liminf_{n\to\infty} \|\widehat{x}_n\| \|f\| = \liminf_{n\to\infty} \|x_n\| \|f\|$$

for all $f \in \mathcal{X}^*$, we easily obtain that

$$\|x\| = \|\widehat{x}\| \leq \liminf_{n\to\infty} \|x_n\|.$$ ∎

The following useful example in functional analysis does require some measure theory.

Corollary 5.1.13. *Let X be a compact Hausdorff space. Then a sequence $(f_n)_{n\geq 1}$ in $C(X, \mathbb{K})$ converges weakly to $f \in C(X, \mathbb{K})$ if and only if $(f_n)_{n\geq 1}$ is $\|\cdot\|_\infty$-bounded and converges to f pointwise.*

Proof. To begin, suppose $(f_n)_{n\geq 1}$ is a sequence in $C(X,\mathbb{K})$ that converges to $f \in C(X,\mathbb{K})$ weakly. By Proposition 5.1.12 we obtain that $(f_n)_{n\geq 1}$ is $\|\cdot\|_\infty$-bounded. To see that $(f_n)_{n\geq 1}$ converges to f pointwise, notice for all $x \in X$ that the map $\delta_x : C(X,\mathbb{K}) \to \mathbb{K}$ defined by

$$\delta_x(g) = g(x)$$

for all $g \in C(X,\mathbb{K})$ is a continuous linear functional. Therefore, since $(f_n)_{n\geq 1}$ converges to f weakly, we must have that

$$f(x) = \delta_x(f) = \lim_{n\to\infty} \delta_x(f_n) = \lim_{n\to\infty} f_n(x)$$

for all $x \in X$ as desired.

Conversely, suppose $(f_n)_{n\geq 1}$ is $\|\cdot\|_\infty$-bounded and converges to f pointwise. To see that $(f_n)_{n\geq 1}$ converges to f weakly, let $T \in C(X,\mathbb{K})^*$ be arbitrary. By the Riesz-Markov Theorem (Theorem D.4.9) there exists a $\mathbb{K}$-valued, finite, regular, Borel measure μ on X such that

$$T(g) = \int_X g\, d\mu$$

for all $g \in C(X,\mathbb{K})$. Since μ is finite, since $(f_n)_{n\geq 1}$ is $\|\cdot\|_\infty$-bounded, and since $(f_n)_{n\geq 1}$ converges to f pointwise, the Dominated Convergence Theorem implies that

$$T(f) = \int_X f\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu = \lim_{n\to\infty} T(f_n).$$

Hence, as $T \in C(X,\mathbb{K})^*$ was arbitrary, $(f_n)_{n\geq 1}$ converges to f weakly. ∎

With the above examples complete, it is useful to examine some properties of the weak topology. To begin, we note that not only are the continuous linear functionals weakly continuous, but all linear maps.

Lemma 5.1.14. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be locally convex topological vector spaces. If $T : \mathcal{V} \to W$ be a continuous linear map, then T is also continuous when $\mathcal{V}$ and $\mathcal{W}$ are equipped with their weak topologies.*

Proof. Let $T : \mathcal{V} \to \mathcal{W}$ be a continuous linear map. To see that T is continuous when $\mathcal{V}$ and $\mathcal{W}$ are equipped with their weak topologies, let $(v_\lambda)_{\lambda\in\Lambda}$ be an arbitrary net in $\mathcal{V}$ that converges weakly to $v \in \mathcal{V}$. To see that $(T(v_\lambda))_{\lambda\in\Lambda}$ converges weakly to $T(v)$ thereby completing the proof, let $f \in \mathcal{W}^*$ be arbitrary. Since T is continuous, $f \circ T \in \mathcal{V}^*$. Since $(v_\lambda)_{\lambda\in\Lambda}$ converges weakly to v, $f \circ T \in \mathcal{V}^*$ implies $(f(T(v_\lambda)))_{\lambda\in\Lambda}$ converges to $f(T(v))$. Therefore, as $f \in \mathcal{W}^*$ was arbitrary, $(T(v_\lambda))_{\lambda\in\Lambda}$ converges weakly to $T(v)$. Therefore, as $(v_\lambda)_{\lambda\in\Lambda}$ was arbitrary, T is continuous when $\mathcal{V}$ and $\mathcal{W}$ are equipped with their weak topologies. ∎

Perhaps what is particularly interesting about the weak topology is that it has the same closed convex sets as the initial topology.

Theorem 5.1.15. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $C \subseteq \mathcal{V}$ be convex. Then the closures of C in $(\mathcal{V}, \mathcal{T})$ and in $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ coincide.*

Proof. Recall that the closure of a set A is the smallest closed set containing A. As C is convex, the closure of C in any topological vector space structure is also convex by Lemma 3.2.21 and thus the smallest closed convex set containing C. As $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))^* = \mathcal{V}^*$ by Theorem 5.1.4, we have that $(\mathcal{V}, \mathcal{T})$ and $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ are locally convex topological spaces with the same continuous linear functionals. Hence Corollary 4.4.16 implies that $(\mathcal{V}, \mathcal{T})$ and $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ have the same closed convex sets and thus the closures of C must coincide. ∎

Corollary 5.1.16. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. If $(x_n)_{n\geq 1}$ is a sequence in $\mathcal{X}$ that converges weakly to $x \in \mathcal{X}$, then*

$$x \in \overline{\mathrm{conv}}^{\|\cdot\|}(\{x_n \mid n \in \mathbb{N}\}).$$

Proof. If $(x_n)_{n\geq 1}$ is a sequence in $\mathcal{X}$ that converges weakly to $x \in \mathcal{X}$, then clearly

$$x \in \overline{\mathrm{conv}}^{\mathrm{weak}}(\{x_n \mid n \in \mathbb{N}\}).$$

Hence $x \in \overline{\mathrm{conv}}^{\|\cdot\|}(\{x_n \mid n \in \mathbb{N}\})$ by Theorem 5.1.15. ∎

Of course, given a topological vector space $(\mathcal{V}, \mathcal{T})$, we know that $\mathcal{V}^*$ is also a vector space that can be given the weak topology induced by $\mathcal{V}^{**}$ as the generalization of canonical embedding will separate points. However, we can construct a (potentially) different and generally more useful topology to turn $\mathcal{V}^*$ into a locally convex topological vector space.

Example 5.1.17. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. Recall $\mathcal{V}^*$ is a vector space. For each $x \in \mathcal{V}$, define $\widehat{x} : \mathcal{V}^* \to \mathbb{K}$ by

$$\widehat{x}(f) = f(x)$$

for all $f \in \mathcal{V}^*$. Let $\widehat{\mathcal{V}} = \{\widehat{x} \mid x \in \mathcal{V}\}$, which clearly is a vector subspace of $(\mathcal{V}^*)^\sharp$. Notice $\widehat{\mathcal{V}}$ separates points since if $f \in \mathcal{V}^*$ is such that $\widehat{x}(f) = 0$ for all $x \in \mathcal{V}$, then $f = 0$. Hence $(\mathcal{V}^*, \widehat{\mathcal{V}})$ is a dual pair and thus $(\mathcal{V}^*, \sigma(\mathcal{V}^*, \widehat{\mathcal{V}}))$ is a locally convex topological vector space by Remark 5.1.2. As this is the most common and important weak topology on the dual of a topological vector space, opposed to calling $\sigma(\mathcal{V}^*, \widehat{\mathcal{V}})$ "the weak topology generated by $\widehat{\mathcal{V}}$" we refer to it as *the weak* topology on $\mathcal{V}^*$ generated by $\mathcal{V}$*. In addition, as $\widehat{\mathcal{V}}$ is really $\mathcal{V}$ in disguise, by convention $\sigma(\mathcal{V}^*, \widehat{\mathcal{V}})$ is usually denoted $\sigma(\mathcal{V}^*, \mathcal{V})$.

To further simplify terminology, we say that a net $(f_\lambda)_{\lambda\in\Lambda}$ is *weak* convergent* to an $f \in \mathcal{V}^*$ if $(f_\lambda)_{\lambda\in\Lambda}$ converges to f with respect to $\sigma(\mathcal{V}^*, \mathcal{V})$; that is,

$$\lim_{\lambda\in\Lambda} f_\lambda(x) = f(x)$$

for all $x \in \mathcal{V}$.

Finally, we note by Theorem 5.1.4 that the *weak* continuous linear functionals on* $\mathcal{V}^*$ (that is, the elements of $\mathcal{V}^\sharp$ that are continuous with respect to this topology) are exactly $\widehat{\mathcal{V}} = \mathcal{V}$.

It is useful to compare the examples of the weak topology given above with the following examples of the weak* topology.

Example 5.1.18. Let $\mathcal{V}$ be a finite dimensional vector space. It is elementary to verify that $\mathcal{V}^\sharp$ has the same dimension as $\mathcal{V}$ and thus is isomorphic to $\mathcal{V}$. Using this, the weak* topology turns $\mathcal{V}^\sharp$ into a locally convex topological vector space, which then must be isomorphic to $(\mathbb{K}^n, \|\cdot\|_\infty)$ where $n = \dim(\mathcal{V})$ by Theorem 3.5.2.

Example 5.1.19. Let $p \in (1, \infty)$ and consider $\ell_p(\mathbb{N})$. As $\ell_p(\mathbb{N}) = \ell_q(\mathbb{N})^*$ by Theorem 1.5.4 where $q \in (1, \infty)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$, we see that there is a weak* topology on $\ell_p(\mathbb{N})$ induced by $\ell_q(\mathbb{N})$. However, again by Theorem 1.5.4, we know that $\ell_p(\mathbb{N})^* = \ell_q(\mathbb{N})$. Therefore, as Remark 1.6.4 shows that $\ell_q(\mathbb{N})$ is reflexive in that the canonical embedding $\mathcal{J} : \ell_q(\mathbb{N}) \to \ell_q(\mathbb{N})^{**}$ is the identity map, we see that the weak and weak* topologies on $\ell_p(\mathbb{N})$ coincide.

Example 5.1.20. Consider $\ell_1(\mathbb{N})$. Recall from Theorem 1.5.7 that $c_0^* = \ell_1(\mathbb{N})$. However, recall from Theorem 1.5.8 that $c^* = \ell_1(\mathbb{N})$. Thus there are two potentially different weak* topologies that can be placed on $\ell_1(\mathbb{N})$ as a priori c_0 and c are not isomorphic.

We claim that these two weak* topologies differ. To see this, for each $n \in \mathbb{N}$ let $\vec{e}_n$ be the sequence with a 1 in the n^{th} entry and zeros everywhere else. Clearly $\vec{e}_n \in \ell_1(\mathbb{N})$ for all $n \in \mathbb{N}$. We claim that $(\vec{e}_n)_{n\geq 1}$ converges weak* to $\vec{0}$ with respect to c_0. Indeed, if $\Phi : c_0 \to \ell_1(\mathbb{N})^*$ is as defined in Theorem 1.5.7, then for all $\vec{z} = (z_n)_{n\geq 1} \in c_0$ we see that

$$\lim_{n\to\infty} \Phi(\vec{z})(\vec{e}_n) = \lim_{n\to\infty} z_n = 0 = \Phi(\vec{z})(\vec{0})$$

so $(\vec{e}_n)_{n\geq 1}$ converges weak* to $\vec{0}$ with respect to c_0. However, $(\vec{e}_n)_{n\geq 1}$ does not converges weak* to $\vec{0}$ with respect to c. Indeed if $\Phi : c \to \ell_1(\mathbb{N})^*$ is as defined in Theorem 1.5.8, then $\vec{z} = (1)_{n\geq 1} \in c$ we see that

$$\lim_{n\to\infty} \Phi(\vec{z})(\vec{e}_n) = \lim_{n\to\infty} 1 = 1 \neq 0 = \Phi(\vec{z})(\vec{0}).$$

Note this does not prove c and c_0 are not isometrically isomorphic. Indeed it is elementary to show that if c and c_0 are isomorphic then c^* and c_0^* are also isomorphic. However, the composition of the isomorphisms from Theorem 1.5.7 and Theorem 1.5.8 may not yield the identity map on $\ell_1(\mathbb{N})$ and thus the above many not yield a contradiction. We note it is true that c and c_0 are isomorphic but not isometrically isomorphic, but wait until Corollary 5.5.14.

Example 5.1.20 is why we do not refer to "the weak* topology" on a dual space; that is, to be precise, we always need to refer to the space we are taking the dual of and not just the dual space. To be formal, if a dual space has many of the following objects, then there are many possible weak* topologies.

Definition 5.1.21. Let $\mathcal{V}$ be a vector space. A *(continuous) predual of* $\mathcal{V}$ is any topological vector space $(\mathcal{W}, \mathcal{T})$ such that $\mathcal{W}^* = \mathcal{V}$ as vector spaces.

Example 5.1.22. Let $(X, \mathcal{T})$ be a compact Hausdorff topological space and let $\mathcal{V}$ denote the vector space of all $\mathbb{K}$-valued, finite, regular, Borel measures on X. By the Riesz-Markov Theorem (Theorem D.4.9), we obtain that $C(X, \mathbb{K})^* = \mathcal{V}$. Thus if $\mathcal{V}$ is equipped with the weak* topology with respect to this vector space isomorphism, we see that a net $(\mu_\lambda)_{\lambda \in \Lambda}$ in $\mathcal{V}$ converges weak* to $\mu \in \mathcal{V}$ if and only if

$$\lim_{\lambda \in \Lambda} \int_X f \, d\mu_\lambda = \int_X f \, d\mu$$

for all $f \in C(X, \mathbb{K})$.

As with the weak topology, weak* convergent sequences have some nice properties.

Lemma 5.1.23. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. If $(f_n)_{n\geq 1}$ is a sequence in $\mathcal{X}^*$ that is weak* convergent to $f \in \mathcal{X}^*$, then*

- $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$, *and*
- $\|f\| \leq \liminf_{n\to\infty} \|f_n\|$.

Proof. Suppose $(f_n)_{n\geq 1}$ is a sequence in $\mathcal{X}^*$ that weak* converges to $f \in \mathcal{X}^*$. Hence

$$f(x) = \lim_{n\to\infty} f_n(x)$$

for all $x \in \mathcal{X}$. Thus $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$ by the Uniform Boundedness Principle (Theorem 2.5.3). Moreover, since

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq \liminf_{n\to\infty} \|f_n\| \, \|x\|$$

for all $x \in \mathcal{X}$, it follows that $\|f\| \leq \liminf_{n\to\infty} \|f_n\|$. ∎

5.2 Quotients and Dual Spaces

With the above constructed topologies, we can now answer a fundamental question: What do the dual space of a quotient of topological vector spaces and the quotient of a dual space of a topological vector space look like? The answer to both of these is apparent when the weak*-topology is used and are based around the following object.

Definition 5.2.1. Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space. The *annihilator* of a subset A of $\mathcal{V}$ is the set

$$A^\perp = \{f \in \mathcal{V}^* \mid f(a) = 0 \text{ for all } a \in A\}.$$

It is elementary to see that the annihilator of a subset of a topological vector space is automatically a weak*-closed subspace of the dual space. Using annihilators, we can describe the dual of a quotient space.

Theorem 5.2.2. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space, let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$, and let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ be the canonical quotient map. Define $\Theta : (\mathcal{V}/\mathcal{W})^* \to \mathcal{W}^\perp$ by*

$$\Theta(f) = f \circ q$$

for all $f \in (\mathcal{V}/\mathcal{W})^*$. Then Θ is a well-defined bijective linear map. Furthermore, if $(\mathcal{V}/\mathcal{W})^*$ is equipped with the weak-topology induced by $\mathcal{V}/\mathcal{W}$ and $\mathcal{W}^\perp$ is equipped with the subspace weak*-topology induced by $\mathcal{V}$, then Θ is a homeomorphism. Finally in the case that $\mathcal{V}$ is a normed linear space, Θ is an isometry.*

Proof. Clearly if $f \in (\mathcal{V}/\mathcal{W})^*$, then $(f \circ q)(w) = 0$ for all $w \in \mathcal{W}$ and $f \circ q \in \mathcal{V}^*$ as the composition of continuous linear maps is continuous. Thus $\Theta(f) \in \mathcal{W}^\perp$ for all $f \in (\mathcal{V}/\mathcal{W})^*$ so Θ is well-defined. It is elementary to verify that Θ is linear.

To see that Θ is injective, suppose $f_1, f_2 \in (\mathcal{V}/\mathcal{W})^*$ are such that $\Theta(f_1) = \Theta(f_2)$. Hence $f_1 \circ q = f_2 \circ q$ so

$$f_1(v + \mathcal{W}) = f_1(q(v)) = f_2(q(v)) = f_2(v + \mathcal{W})$$

for all $v \in \mathcal{V}$ so $f_1 = f_2$. Hence Θ is injective.

To see that Θ is surjective, let $g \in \mathcal{W}^\perp$ be arbitrary. Define $\widetilde{g} : \mathcal{V}/\mathcal{W} \to \mathbb{K}$ by

$$\widetilde{g}(v + \mathcal{W}) = g(v)$$

for all $v \in \mathcal{V}$. Since $g \in \mathcal{W}^\perp$ so $g(w) = 0$ for all $w \in \mathcal{W}$, we see that $\widetilde{g}$ is well-defined. Furthermore, clearly $\widetilde{g}$ is linear since g is linear and $g = \widetilde{g} \circ q$. To see that $\widetilde{g} \in (\mathcal{V}/\mathcal{W})^*$ and thus $\Theta(\widetilde{g}) = g$, notice that

$$q^{-1}(\{v + \mathcal{W} \mid |\widetilde{g}(v + W)| < 1\}) = \{x \in \mathcal{V} \mid |g(x)| < 1\}$$

since $\mathcal{W} \subseteq \ker(g)$. Since $\{x \in \mathcal{V} \mid |g(x)| < 1\}$ is open as $g \in \mathcal{V}^*$ and since q is an open mapping by Lemma 3.3.4, we obtain that

$$\{v + \mathcal{W} \mid |\widetilde{g}(v + W)| < 1\}$$

is open in $\mathcal{V}/\mathcal{W}$. Hence Proposition 4.1.12 implies that $\widetilde{g} \in (\mathcal{V}/\mathcal{W})^*$. Thus Θ is surjective and thus bijective.

To see that Θ is a homeomorphism with respect to the designated topologies, let $(f_\lambda)_{\lambda\in\Lambda}$ be a net in $(\mathcal{V}/\mathcal{W})^*$ and let $f \in (\mathcal{V}/\mathcal{W})^*$. Then $(f_\lambda)_{\lambda\in\Lambda}$ weak* converges to f if and only if $(f_\lambda(v+\mathcal{W}))_{\lambda\in\Lambda}$ converges to $f(v+\mathcal{W})$ for all $v \in \mathcal{V}$ if and only if $(\Theta(f_\lambda)(v))_{\lambda\in\Lambda}$ converges to $\Theta(f)(v)$ for all $v \in \mathcal{V}$ if and only if $(\Theta(f_\lambda))_{\lambda\in\Lambda}$ weak* converges to $\Theta(f)$. Hence Θ is a homeomorphism.

Finally, suppose $\mathcal{V}$ is a normed linear space and let $f \in (\mathcal{V}/\mathcal{W})^*$ be arbitrary. Notice that

$$\|\Theta(f)\| = \|f \circ q\| \leq \|f\| \, \|q\| = \|f\| \, .$$

For the reverse inequality, let $\epsilon > 0$ be arbitrary. By the definition of the operator norm, there exists a $v + \mathcal{W} \in \mathcal{V}/\mathcal{W}$ such that $\|v + \mathcal{W}\| < 1$ and $|f(v + \mathcal{W})| > \|f\| - \epsilon$. By the definition of the quotient norm there exists a $w \in \mathcal{W}$ such that $\|v + w\| < 1$ and thus

$$|\Theta(f)(v + w)| = |f(v + \mathcal{W})| > \|f\| - \epsilon.$$

Since $\|v + w\| < 1$, this implies $\|\Theta(f)\| > \|f\| - \epsilon$. Therefore, as $\epsilon > 0$ was arbitrary, $\|\Theta(f)\| = \|f\|$. Hence, as $f \in (\mathcal{V}/\mathcal{W})^*$ was arbitrary, Θ is an isometry. ∎

In a similar vein, we have the following.

Theorem 5.2.3. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Define $\Theta : \mathcal{V}^*/\mathcal{W}^\perp \to \mathcal{W}^*$ by*

$$\Theta(f + \mathcal{W}^\perp) = f|_\mathcal{W}$$

for all $f \in \mathcal{V}^*$. Then Θ is a well-defined bijective linear map. Furthermore, if $\mathcal{V}^*/\mathcal{W}^\perp$ is equipped with the quotient topology when $\mathcal{V}^*$ is equipped the weak-topology induced by $\mathcal{V}$ and $\mathcal{W}^*$ is equipped with the weak*-topology induced by $\mathcal{W}$, then Θ is a homeomorphism. Finally in the case that $\mathcal{V}$ is a normed linear space, Θ is an isometry.*

Proof. Clearly if $f \in \mathcal{V}^*$, then $f|_\mathcal{W}$ is a continuous linear functional on $\mathcal{W}$. To see that Θ is well-defined, suppose $f_1 + \mathcal{W}^\perp, f_2 + \mathcal{W}^\perp \in \mathcal{V}^*/\mathcal{W}^\perp$ are such that $f_1 + \mathcal{W}^\perp = f_2 + \mathcal{W}^\perp$. Hence $f_1 - f_2 \in \mathcal{W}^\perp$ so $(f_1 - f_2)(w) = 0$ for all $w \in \mathcal{W}$ and thus $f_1|_\mathcal{W} = f_2|_\mathcal{W}$. Thus Θ is well-defined. It is elementary to verify that Θ is linear.

To see that Θ is injective, suppose $f_1 + \mathcal{W}^\perp, f_2 + \mathcal{W}^\perp \in \mathcal{V}^*/\mathcal{W}^\perp$ are such that $\Theta(f_1) = \Theta(f_2)$. Hence $f_1(w) = f_2(w)$ for all $w \in \mathcal{W}$. Thus $(f_1 - f_2)(w) = 0$ for all $w \in \mathcal{W}$ so $f_1 - f_2 \in \mathcal{W}^\perp$. Hence $f_1 + \mathcal{W}^\perp = f_2 + \mathcal{W}^\perp$ so Θ is injective.

To see that Θ is surjective, let $g \in \mathcal{W}^*$ be arbitrary. By the Hahn-Banach Theorem (Theorem 4.2.4) there exists an $f \in \mathcal{V}^*$ such that $f|_{\mathcal{W}} = g$. Hence $\Theta(f + \mathcal{W}^\perp) = g$. Therefore, since g was arbitrary, Θ is surjective and thus bijective.

To see that Θ is a homeomorphism with respect to the designated topologies, let $q : \mathcal{V}^* \to \mathcal{V}^*/\mathcal{W}^\perp$ denote the canonical quotient map and let $r : \mathcal{V}^* \to \mathcal{W}^*$ denote the restriction map. Thus $r = \Theta \circ q$. Note r is weak*-weak* continuous, and q is weak*-quotient continuous by construction.

To see that the above implies Θ is continuous, we repeat the proof of Theorem A.7.20 for convenience. Let $U \subseteq \mathcal{W}^*$ be open in the weak* topology. Hence

$$r^{-1}(U) = q^{-1}(\Theta^{-1}(U))$$

is open. However, since q is an open map by the definition of the quotient topology, we obtain that $\Theta^{-1}(U)$ is open. Therefore, as U was arbitrary, Θ is continuous.

To see that Θ^{-1} is continuous, recall by the definition of the weak* topology and the characterization of the seminorms on a quotient of a locally convex topological vector space from Proposition 3.6.19 that if for all $x \in \mathcal{V}$ we define $\widetilde{p_x} : \mathcal{V}^*/\mathcal{W}^\perp \to [0, \infty)$ by

$$\widetilde{p_x}(f + \mathcal{W}^\perp) = \inf(\{|(f+g)(x)| \mid g \in \mathcal{W}^\perp\})$$

for all $f \in \mathcal{V}^*$, then the topology on $\mathcal{V}^*/\mathcal{W}^\perp$ is generated by $\{\widetilde{p_x} \mid x \in \mathcal{V}\}$.

We claim that if $x \notin \mathcal{W}$, then $\widetilde{p_x} = 0$. To see this, suppose $x \notin \mathcal{W}$ and let $f \in \mathcal{V}^*$ be arbitrary. Let $\mathcal{L} = \operatorname{span}(\{x\})$. Thus $\mathcal{L} \cap \mathcal{W} = \emptyset$. As $\mathcal{L}$ is finite dimensional and $\mathcal{W}$ is closed, $\mathcal{L} + \mathcal{W}$ is a closed subspace of $\mathcal{V}$ by Corollary 3.5.6 and $\mathcal{L}$ and $\mathcal{W}$ are topological complements of each other in $\mathcal{L} + \mathcal{W}$. Thus, if we define $g : \mathcal{L} + \mathcal{W} \to \mathbb{K}$ by

$$g(\alpha x + w) = f(w)$$

for all $\alpha \in \mathbb{K}$ and $w \in \mathcal{W}$, then g is a well-defined linear map. Moreover, since

$$\ker(g) = \mathcal{L} + (\ker(f) \cap \mathcal{W}),$$

we obtain that $\ker(g)$ is closed by Corollary 3.5.6. Therefore $g \in (\mathcal{L} + \mathcal{W})^*$. Hence the Hahn-Banach Theorem (Theorem 4.2.4) implies there exists an $h \in \mathcal{V}^*$ such that $h|_{\mathcal{L}+\mathcal{W}} = g$. Since $h|_{\mathcal{W}} = g|_{\mathcal{W}} = f$, we have that $f - h \in \mathcal{W}^\perp$ so

$$0 \leq \widetilde{p_x}(f + \mathcal{W}^\perp) = \widetilde{p_x}(h + \mathcal{W}^\perp) \leq |h(x)| = 0.$$

Therefore, as f was arbitrary, $\widetilde{p_x} = 0$.

Returning to showing Θ^{-1} is continuous, suppose $(f_\lambda + \mathcal{W}^\perp)_{\lambda\in\Lambda}$ is a net in $\mathcal{V}^*/\mathcal{W}^\perp$ such that $(\Theta(f_\lambda + \mathcal{W}^\perp))_{\lambda\in\Lambda}$ converges weak* to $\Theta(f + \mathcal{W}^\perp)$ for some $f \in \mathcal{V}^*$. To see that $(f_\lambda + \mathcal{W}^\perp)_{\lambda\in\Lambda}$ converges weak* to $f + \mathcal{W}^\perp$, let $x \in \mathcal{V}$ be arbitrary. If $x \notin \mathcal{W}$, then

$$\widetilde{p_x}(f_\lambda + \mathcal{W}^\perp) = 0 = \widetilde{p_x}(f + \mathcal{W}^\perp)$$

for all $\lambda \in \Lambda$. However, if $x \in \mathcal{W}$, then the fact that $(\Theta(f_\lambda + \mathcal{W}^\perp))_{\lambda\in\Lambda}$ converges weak* to $\Theta(f + \mathcal{W}^\perp)$ implies that

$$\lim_{\lambda\in\Lambda} \widetilde{p_x}(f_\lambda + \mathcal{W}^\perp) = \lim_{\lambda\in\Lambda} f_\lambda(x) = f(x) = \widetilde{p_x}(f + \mathcal{W}^\perp).$$

Therefore, since $\{\widetilde{p_x} \mid x \in \mathcal{V}\}$ is a separating family of seminorms that generate the weak* topology on $\mathcal{V}/\mathcal{W}^\perp$, Proposition 3.2.13 implies that $(f_\lambda + \mathcal{W}^\perp)_{\lambda\in\Lambda}$ converges weak* to $f + \mathcal{W}^\perp$ as desired. Hence Θ is a homeomorphism.

Finally, suppose $\mathcal{V}$ is a normed linear space and let $f + \mathcal{W}^\perp \in \mathcal{V}^*/\mathcal{W}^\perp$ be arbitrary. If $\left\|f + \mathcal{W}^\perp\right\| < 1$, then there exists a $g \in \mathcal{W}^\perp$ such that $\|f + g\| < 1$. Thus

$$\left\|\Theta(f + \mathcal{W}^\perp)\right\| = \|f|_\mathcal{W}\| = \|(f + g)|_\mathcal{W}\| \leq \|f + g\| < 1.$$

Therefore, since $f + \mathcal{W}^\perp$ was arbitrary, $\|\Theta\| \leq 1$. Hence $\left\|\Theta(f + \mathcal{W}^\perp)\right\| \leq \left\|f + \mathcal{W}^\perp\right\|$ for all $f + \mathcal{W}^\perp \in \mathcal{V}^*/\mathcal{W}^\perp$.

For the other inclusion, notice let $f \in \mathcal{W}^\perp$ be arbitrary. By the Hahn-Banach Theorem (Theorem 4.2.5), there exists a $g \in \mathcal{V}^*$ such that $\|g\| = \|f\|$ and $g|_\mathcal{W} = f$. As this later condition implies that $g - f \in \mathcal{W}^\perp$, we obtain that

$$\left\|\Theta^{-1}(f)\right\| = \left\|f + \mathcal{W}^\perp\right\| = \left\|g + \mathcal{W}^\perp\right\| \leq \|g\| = \|f\|.$$

Hence, as $f \in \mathcal{W}^\perp$ was arbitrary and Θ was a bijection, Θ is an isometry. ∎

5.3 The Banach-Alaoglu Theorem

One nice property of any weak* topology is that it is characterized by convergence of functions at points. As this is very similar to the convergence in the product topology, Tychonoff's Theorem (Theorem 5.3.2) enables us to prove the Banach-Alaoglu (Theorem 5.3.4) which shows that the norm closed unit ball of any normed linear space is weak*-compact. This is quite useful in that any net in the closed unit ball has a weak*ly convergent subnet.

In order to prove the Banach-Alaoglu (Theorem 5.3.4), we recall the proof of Tychonoff's Theorem (Theorem 5.3.2) via the following lemma.

Lemma 5.3.1. *Let $(X, \mathcal{T})$ be a topological space and let $\mathcal{F} \subseteq \mathcal{P}(X)$ be a non-empty collection of sets with the finite intersection property. Then there exists an $\mathcal{M} \subseteq \mathcal{P}(X)$ such that*

(1) $\mathcal{F} \subseteq \mathcal{M}$,

(2) $\mathcal{M}$ *has the finite intersection property,*

(3) if $F \in \mathcal{P}(X) \setminus \mathcal{M}$, *then* $\mathcal{M} \cup \{F\}$ *does not have the finite intersection property,*

(4) if $\{F_k\}_{k=1}^n \subseteq \mathcal{M}$ *for some* $n \in \mathbb{N}$, *then* $\bigcap_{k=1}^n F_k \in \mathcal{M}$, *and*

(5) if $Y \subseteq X$ *and* $Y \cap M \neq \emptyset$ *for all* $M \in \mathcal{M}$, *then* $Y \in \mathcal{M}$.

Proof. Let

$$\mathcal{C} = \{\mathcal{S} \subseteq \mathcal{P}(X) \mid \mathcal{F} \subseteq \mathcal{S} \text{ and } \mathcal{S} \text{ has the finite intersection property}\}.$$

Clearly $\mathcal{C} \neq \emptyset$ since $\mathcal{F} \subseteq \mathcal{C}$. For $\mathcal{S}_1, \mathcal{S}_2 \in \mathcal{C}$, define $\mathcal{S}_1 \preceq \mathcal{S}_2$ if and only if $\mathcal{S}_1 \subseteq \mathcal{S}_2$. Clearly $(\mathcal{C}, \preceq)$ is a partially ordered set.

We claim that every chain in $(\mathcal{C}, \preceq)$ has an upper bound. To see this, suppose that $\{\mathcal{S}_\alpha\}_{\alpha \in I}$ is a chain in $(\mathcal{C}, \preceq)$. Let

$$\mathcal{S} = \bigcup_{\alpha \in I} \mathcal{S}_\alpha.$$

We claim that $\mathcal{S} \in \mathcal{C}$ from which it trivially follows that $\mathcal{S}$ is an upper bound for $\{\mathcal{S}_\alpha\}_{\alpha \in I}$. To see that $\mathcal{S} \in \mathcal{C}$, first notice since $\mathcal{F} \subseteq \mathcal{S}_\alpha$ for all $\alpha \in I$ that $\mathcal{F} \subseteq \mathcal{S}$ by construction. To see that $\mathcal{S}$ has the finite intersection property, suppose that $n \in \mathbb{N}$ and $S_1, S_2, \ldots, S_n \in \mathcal{S}$. By the properties of a chain, there exists an $\alpha_0 \in I$ such that $S_k \in \mathcal{S}_{\alpha_0}$ for all $k \in \{1, \ldots, n\}$. Therefore, since $\mathcal{S}_{\alpha_0}$ has the finite intersection property as $\mathcal{S}_{\alpha_0} \in \mathcal{C}$, we obtain that $\bigcap_{k=1}^n S_k \neq \emptyset$. Therefore, as $n \in \mathbb{N}$ and $S_1, \ldots, S_n \in \mathcal{S}$ were arbitrary, $\mathcal{S}$ has the finite intersection property and thus $\mathcal{S} \in \mathcal{C}$.

Since $(\mathcal{C}, \preceq)$ is a non-empty partially ordered set such that every chain has an upper bound, Zorn's Lemma implies that there exists an $\mathcal{M} \in \mathcal{C}$ such that if $\mathcal{S} \in \mathcal{C}$ and $\mathcal{M} \preceq \mathcal{S}$, then $\mathcal{S} = \mathcal{M}$ (i.e. $\mathcal{M}$ is maximal in $(\mathcal{C}, \preceq)$). We claim that $\mathcal{M}$ has the desired properties. Indeed $\mathcal{F} \subseteq \mathcal{M}$ and $\mathcal{M}$ has the finite intersection property since $\mathcal{M} \in \mathcal{C}$. Thus (1) and (2) hold.

To see that (3) holds, let $F \in \mathcal{P}(X) \setminus \mathcal{M}$ be arbitrary. If $\mathcal{M}_0 = \mathcal{M} \cup \{F\}$ had the finite intersection property, then since $\mathcal{F} \subseteq \mathcal{M} \subseteq \mathcal{M}_0$ we would have that $\mathcal{M}_0 \in \mathcal{C}$, $\mathcal{M}_0 \neq \mathcal{M}$ as $F \in \mathcal{M}_0 \setminus \mathcal{M}$, and $\mathcal{M} \preceq \mathcal{M}_0$ thereby contradicting the maximality of $\mathcal{M}$. Therefore (3) holds.

To see that (4) holds, let $n \in \mathbb{N}$ and $\{F_k\}_{k=1}^n \subseteq \mathcal{M}$ be arbitrary. If $F = \bigcap_{k=1}^n F_k$, then clearly $\mathcal{M} \cup \{F\}$ has the finite intersection property since $\mathcal{M}$ has the finite intersection property. Thus (3) implies that $F \in \mathcal{M}$. Therefore, as $n \in \mathbb{N}$ and $\{F_k\}_{k=1}^n \subseteq \mathcal{M}$ were arbitrary, (4) follows.

Finally, to see that (5) holds, suppose $Y \subseteq X$ is such that $Y \cap M \neq \emptyset$ for all $M \in \mathcal{M}$. Therefore (4) implies that for all $n \in \mathbb{N}$ and $\{F_k\}_{k=1}^n \subseteq \mathcal{M}$ that $Y \cap (\bigcap_{k=1}^n F_k) \neq \emptyset$. Thus $\mathcal{M} \cup \{Y\}$ has the finite intersection property so (3) implies that $Y \in \mathcal{M}$ as desired. ∎

Theorem 5.3.2 (Tychonoff's Theorem). *Let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be compact topological spaces. Then $\prod_{\alpha\in I} X_\alpha$ is a compact topological space when equipped with the product topology.*

Proof. Let $X = \prod_{\alpha\in I} X_\alpha$ and let $\mathcal{T}$ denote the product topology on X. To see that $(X, \mathcal{T})$ is compact, we will apply Theorem A.9.2 and verify that any set of closed subsets of $(X, \mathcal{T})$ with finite intersection property has non-empty intersection.

Let $\mathcal{F}$ be an arbitrary set of closed subsets of $(X, \mathcal{T})$ with the finite intersection property. Let $\mathcal{M}$ be a set with the finite intersection property containing $\mathcal{F}$ as created via Lemma 5.3.1. Since

$$\bigcap_{F\in\mathcal{F}} F \supseteq \bigcap_{A\in\mathcal{M}} \overline{A},$$

it suffices to show that $\bigcap_{A\in\mathcal{M}} \overline{A} \neq \emptyset$.

For each $\alpha \in I$, let $\pi_\alpha : X \to X_\alpha$ be the projection map from X to X_α from Example A.6.5. Since $\mathcal{M}$ has the finite intersection property, it is clear that

$$\{\pi_\alpha(A) \mid A \in \mathcal{M}\}$$

has the finite intersection property in (X_α, T_α) so

$$\left\{\overline{\pi_\alpha(A)} \mid A \in \mathcal{M}\right\}$$

is a collection of closed sets in $(X_\alpha, \mathcal{T}_\alpha)$ with the finite intersection property. Therefore, since $(X_\alpha, \mathcal{T}_\alpha)$ is compact, Theorem A.9.2 implies for all $\alpha \in I$ there exists an $x_\alpha \in X_\alpha$ such that

$$x_\alpha \in \bigcap_{A\in\mathcal{M}} \overline{\pi_\alpha(A)}.$$

Let $x = (x_\alpha)_{\alpha\in I} \in X$. We claim that $x \in \bigcap_{A\in\mathcal{M}} \overline{A}$ thereby completing the proof that $\bigcap_{A\in\mathcal{M}} \overline{A} \neq \emptyset$.

To begin, let $\alpha_0 \in I$ and $U \in \mathcal{T}_{\alpha_0}$ be such that $x_{\alpha_0} \in U$. Since $x_{\alpha_0} \in \overline{\pi_{\alpha_0}(A)}$ for all $A \in \mathcal{M}$, Theorem A.5.21 implies that $\pi_{\alpha_0}(A) \cap U \neq \emptyset$ for all $A \in \mathcal{M}$. Hence $A \cap \pi_{\alpha_0}^{-1}(U) \neq \emptyset$ for all $A \in \mathcal{M}$. Therefore, the properties of $\mathcal{M}$ from Lemma 5.3.1 imply that $\pi_{\alpha_0}^{-1}(U) \in \mathcal{M}$ for all $\alpha_0 \in I$ and $U \in \mathcal{T}_{\alpha_0}$ such that $x_{\alpha_0} \in U$.

Since $\mathcal{M}$ is closed under finite intersections from Lemma 5.3.1,

$$\left\{ \bigcap_{\alpha\in J} \pi_\alpha^{-1}(U_\alpha) \,\middle|\, \begin{array}{c} J\subseteq I \text{ finite and} \\ U_\alpha \text{ a } \mathcal{T}_\alpha\text{-neighbourhood of } x_\alpha \text{ for all } \alpha\in J \end{array} \right\}$$

is both contained in $\mathcal{M}$ and is a neighbourhood basis of x in $(X, \mathcal{T})$. Therefore, as $\mathcal{M}$ has the finite intersection property, every element of $\mathcal{M}$ has non-empty intersection with each element of a neighbourhood basis of x. Hence Theorem A.5.21 implies that $x \in \overline{A}$ for all $A \in \mathcal{M}$. Thus $x \in \bigcap_{A\in\mathcal{M}} \overline{A}$ thereby completing the proof. ∎

Of course, the proof of Tychonoff's Theorem (Theorem 5.3.2) relies on Zorn's Lemma and it is well-known that Zorn's Lemma is equivalent to the Axiom of Choice). It is perhaps surprising that Tychonoff's Theorem is equivalent to the Axiom of Choice. The proof that Tychonoff's Theorem implies the Axiom of Choice is as follows.

Theorem 5.3.3. *Suppose that Tychonoff's Theorem holds; that is, the product of compact topological spaces is compact when equipped with the product topology. Then for any non-empty set I and any set $\{X_\alpha\}_{\alpha\in I}$ of non-empty sets, the product $\prod_{\alpha\in I} X_\alpha$ is non-empty.*

Proof. Let I be an non-empty set and let $\{X_\alpha\}_{\alpha\in I}$ be a set of non-empty sets. For each $\alpha \in I$, let $Y_\alpha = X_\alpha \cup \{\infty_\alpha\}$ for some symbol ∞_α and let $Y = \prod_{\alpha\in I} Y_\alpha$. We note that Y is automatically non-empty without the use of the Axiom of Choice. Indeed we already know for all $\alpha \in I$ that $\infty_\alpha \in Y_\alpha$; that is, we do not need to choose an element of Y_α for each $\alpha \in I$ as we already know (i.e. have assigned) an element of Y_α for each $\alpha \in I$. Hence the element $\infty = (\infty_\alpha)_{\alpha\in I}$ is an element of Y without the use of the Axiom of Choice.

For each $\alpha \in I$, let $\mathcal{T}_\alpha = \{\emptyset, Y_\alpha, X_\alpha, \{\infty_\alpha\}\}$. Clearly $\mathcal{T}_\alpha$ is a topology on Y_α. Furthermore, since $\mathcal{T}_\alpha$ only has a finite number of sets, every $\mathcal{T}_\alpha$-open cover of Y_α has a finite subcover (namely the original open cover) so $(Y_\alpha, \mathcal{T}_\alpha)$ is compact. Hence Tychonoff's Theorem implies that $Y = \prod_{\alpha\in I} Y_\alpha$ is compact when equipped with the product topology.

For each $\alpha_0 \in I$, let
$$U_{\alpha_0} = \prod_{\alpha\in I} U_{\alpha_0,\alpha}$$
where
$$U_{\alpha_0,\alpha} = \begin{cases} Y_\alpha & \text{if } \alpha \neq \alpha_0 \\ \{\infty_\alpha\} & \text{if } \alpha = \alpha_0 \end{cases}.$$
Again, the construction of U_{α_0} does not require the Axiom of Choice since we do not need to choose an element of $\mathcal{T}_\alpha$ for each $\alpha \in I$ as we already know an element of $\mathcal{T}_\alpha$ for each $\alpha \in I$. Clearly U_{α_0} is open in the product topology on Y by definition.

We claim that $\mathcal{U} = \{U_\alpha\}_{\alpha\in I}$ cannot cover Y. To see this, suppose to the contrary that $\mathcal{U}$ is an open cover of Y. Since Y is compact, there exists a finite subset $J \subseteq I$ such that $Y = \bigcup_{\alpha\in J} U_\alpha$. For each $\alpha \in J$, choose $x_\alpha \in X_\alpha$. Note this does not require the Axiom of Choice since J is finite. For each $\alpha \in I \setminus J$, let $x_\alpha = \infty_\alpha$. Again, this does not require the Axiom of Choice. Thus $x = (x_\alpha)_{\alpha\in I} \in Y$ by definition. However $x \notin \bigcup_{\alpha\in J} U_\alpha$ by construction since $x_\alpha \notin \{\infty_\alpha\}$ for all $\alpha \in J$. Hence we have a contradiction so $\mathcal{U}$ is not a cover of Y.

Since $\mathcal{U}$ is not a cover of Y, there must exist an element $y = (y_\alpha)_{\alpha\in I} \in Y$ such that $y \notin U_\alpha$ for all $\alpha \in I$. Then, by the definition of U_α, we see that

$y_\alpha \notin \{\infty_\alpha\}$ for all $\alpha \in I$. Hence $y_\alpha \in X_\alpha$ for all $\alpha \in I$ so that $y \in \prod_{\alpha\in I} X_\alpha$. Hence Tychonoff's Theorem implies the Axiom of Choice. ∎

Theorem 5.3.4 (The Banach-Alaoglu Theorem). *Let $(\mathcal{V}, \mathcal{T})$ be a topological vector space and let $U \in \mathcal{T}$ be a neighbourhood of $\vec{0}$. Then the set*

$$U^\circ = \left\{ f \in \mathcal{V}^* \,\middle|\, \sup_{x\in U} |f(x)| \leq 1 \right\}$$

(which is often called the polar *of U) is compact with respect to the weak* topology $\sigma(\mathcal{V}^*, \mathcal{V})$.*

Consequently, for any normed linear space $(\mathcal{X}, \|\cdot\|)$, the closed unit ball of $\mathcal{X}^*$ is weak-compact.*

Proof. Clearly if U is the closed unit ball in a normed linear space $(\mathcal{X}, \|\cdot\|)$, then U° is the closed unit ball of $\mathcal{X}^*$ by definition. Thus it suffices to prove the first statement.

For each $x \in \mathcal{V}$ we note since U is a neighbourhood of $\vec{0}$ and thus absorbing by Lemma 3.1.10 that there exists an $r_x \in (0, \infty)$ such that $x \in r_x U$. For each $x \in \mathcal{V}$, let

$$K_x = \{z \in \mathbb{K} \mid |z| \leq r_x\}.$$

Clearly each K_x is a compact subset of $\mathbb{K}$ being closed and bounded and thus $K = \prod_{x\in\mathcal{V}} K_x$ equipped with the product topology is compact by Tychnoff's Theorem.

For each $f \in U^\circ$ and $x \in \mathcal{V}$, we know that there exists a $u_x \in U$ such that $x = r_x u_x$ and thus

$$|f(x)| = |r_x f(u_x)| = r_x |f(u_x)| \leq r_x$$

so $f(x) \in K_x$. Therefore, the map $\Phi : U^\circ \to K$ defined by

$$\Phi(f) = (f(x))_{x\in\mathcal{V}}$$

for all $f \in U^\circ$ is a well-defined map.

We claim that Φ is a weak*-homeomorphism onto $\Phi(U^\circ)$. To see this, we first note that Φ is clearly injective. Next let $(f_\lambda)_{\lambda\in\Lambda}$ be an arbitrary net in U° and let $f \in U^\circ$. Then $(f_\lambda)_{\lambda\in\Lambda}$ converges to f weak* if and only if

$$\lim_{\lambda\in\Lambda} \Phi(f_\lambda)(x) = \lim_{\lambda\in\Lambda} f_\lambda(x) = f(x) = \Phi(f)(x)$$

for all $x \in \mathcal{V}$, if and only if $(\Phi(f_\lambda))_{\lambda\in\Lambda}$ converges entrywise to $\Phi(f)$ if and only if $(\Phi(f_\lambda))_{\lambda\in\Lambda}$ converges to $\Phi(f)$ in K (see Theorem A.4.22). Hence Φ is a weak*-homeomorphism onto $\Phi(U^\circ)$. Therefore, to show that U° is weak*-compact, it suffices to show that $\Phi(U^\circ)$ is a closed subset of K (as closed subsets of compact spaces are compact).

To see that $\Phi(U^\circ)$ is closed, let $(f_\lambda)_{\lambda\in\Lambda}$ be an arbitrary net in U° such that $(\Phi(f_\lambda))_{\lambda\in\Lambda}$ converges to $f \in K$. Hence $f : \mathcal{V} \to \mathbb{K}$. To see that $f \in \mathcal{V}^\sharp$, let $x_1, x_2 \in \mathcal{V}$ and $\alpha \in \mathbb{K}$ be arbitrary. Since $(\Phi(f_\lambda))_{\lambda\in\Lambda}$ converges to f in K, we know that $(f_\lambda)_{\lambda\in\Lambda}$ converges to f pointwise so

$$\begin{aligned} f(\alpha x_1 + x_2) &= \lim_{\lambda\in\Lambda} f_\lambda(\alpha x_1 + x_2) \\ &= \lim_{\lambda\in\Lambda} \alpha f_\lambda(x_1) + f_\lambda(x_2) \\ &= \alpha f(x_1) + f(x_2). \end{aligned}$$

Therefore, as $\vec{v}_1, \vec{v}_2 \in V$ and $\alpha \in \mathbb{K}$ were arbitrary, $f \in \mathcal{V}^\sharp$.

Since $f(x) \in K_x$ for all $x \in \mathcal{V}$, we know that $|f(x)| \leq 1$ for all $x \in U$ and thus we will have $f \in U^\circ$ provided $f \in \mathcal{V}^*$. However since the seminorm $x \mapsto |f(x)|$ is bounded on U and thus continuous by Proposition 3.6.13 we obtain that $f \in \mathcal{V}^*$ so $f \in U^\circ$. Therefore, since $(f_\lambda)_{\lambda\in\Lambda}$ was arbitrary $\Phi(U^\circ)$ is a closed subset of K thereby completing the proof. ∎

Of course, there are some clear corollaries of the Banach-Alaoglu Theorem (Theorem 5.3.4).

Corollary 5.3.5. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space and let $A \subseteq \mathcal{X}^*$. If A is bounded and weak*-closed, then A is weak*-compact.*

Proof. By the Banach-Alaoglu Theorem (Theorem 5.3.4), the unit ball of $\mathcal{X}^*$ is weak*-compact. Hence, since $\mathcal{X}^*$ is a (locally convex) topological vector space so scalar multiplication and translation are homeomorphisms, every closed ball is weak*-compact. Since every bounded subset of $\mathcal{X}^*$ is contained in a closed ball of $\mathcal{X}^*$ and since every weak*-closed subset of a weak*-compact set is weak*-compact, the result follows. ∎

Corollary 5.3.6. *Every normed space $(\mathcal{X}, \|\cdot\|)$ is isometrically isomorphic to a vector subspace of $(C(X,\mathbb{K}), \|\cdot\|_\infty)$ for some compact Hausdorff space X. Moreover, if $(\mathcal{X}, \|\cdot\|)$ then the corresponding vector space is closed.*

Proof. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space and let $X = \mathcal{X}_1^*$ equipped with the weak* topology. By the Banach-Alaoglu Theorem (Theorem 5.3.4), X is a compact Hausdorff space.

Recall that the canonical embedding $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ defined by $\mathcal{J}(x) = \widehat{x}$ for all $x \in \mathcal{X}$ where

$$\widehat{x}(f) = f(x)$$

for all $f \in \mathcal{X}^*$ is an isometric isomorphism. Hence $\Phi : \mathcal{X} \to C(X, \mathbb{K})$ defined by

$$\Phi(x) = \widehat{x}|_X$$

for all $x \in \mathcal{X}$ is a well-defined contractive linear map.

To see that Φ is isometric, let $x \in \mathcal{X}$ be arbitrary. By the Hahn-Banach Theorem (Corollary 4.3.1) there exists an $f \in \mathcal{X}_1^*$ such that $|f(x)| = \|x\|$. Hence $|\Phi(x)(f)| = \|x\|$ so $\|\Phi(f)\|_\infty \geq \|x\|$. Therefore, as x was arbitrary and Φ was contractive, Φ is isometric.

Finally, if $(\mathcal{X}, \|\cdot\|)$ is a Banach space, then $\mathcal{X}$ is complete so $\Phi(\mathcal{X})$ is also complete and thus closed in $C(X, \mathbb{K})$. ∎

5.4 Goldstine's Theorem

As the Banach-Alaoglu demonstrates the closed unit ball of the dual of any normed linear space is weak*-compact, a further examination of such closed unit balls in the weak* topology is meritted. In particular, we will prove the following theorem related to the weak* topology on the double dual induced by the dual. In that which follows, given a normed linear space $(\mathcal{X}, \|\cdot\|)$ we will use $\mathcal{X}_1$ to denote the closed unit all of $\mathcal{X}$ for convenience.

Theorem 5.4.1 (Goldstine's Theorem). *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. If $\mathcal{X}^{**}$ is equipped with the weak*-topology induced by $\mathcal{X}^*$ and $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ is the canonical embedding, then $\mathcal{J}(\mathcal{X}_1)$ is weak*-dense in $\mathcal{X}_1^{**}$. Hence $\mathcal{J}(\mathcal{X})$ is weak*-dense in $\mathcal{X}^{**}$.*

Proof. Clearly the fact that $\mathcal{J}(\mathcal{X})$ is weak*-dense in $\mathcal{X}^{**}$ will follow from the fact that $\mathcal{J}(\mathcal{X}_1)$ is weak*-dense in $\mathcal{X}_1^{**}$. To see that $\mathcal{J}(\mathcal{X}_1)$ is weak*-dense in $\mathcal{X}_1^{**}$, note since $\mathcal{J}$ is a isometric linear map that $\mathcal{J}(\mathcal{X}_1)$ is a balanced convex subset of $\mathcal{X}_1^{**}$. Hence $\overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}$ is a weak*-closed, balanced, convex subset of $\mathcal{X}^{**}$.

Suppose there exists an $f_0 \in \mathcal{X}_1^{**} \setminus \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}$. As $\{f_0\}$ is weak*-compact, the Hahn-Banach Theorem (Theorem 4.4.13) implies there exists a weak*-continuous linear functional $g : \mathcal{X}^{**} \to \mathbb{K}$ and $a, b \in \mathbb{R}$ such that

$$\mathrm{Re}(g(f_0)) \geq a > b \geq \sup\left(\left\{\mathrm{Re}(g(f)) \mid f \in \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}\right\}\right).$$

Note since $\overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}$ is balanced that

$$\sup\left(\left\{\mathrm{Re}(g(f)) \mid f \in \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}\right\}\right) = \sup\left(\left\{|g(f)| \mid f \in \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}\right\}\right).$$

Moreover, since $g \in (\mathcal{X}^{**}, \sigma(\mathcal{X}^{**}, \mathcal{X}^*))^* = \mathcal{X}^*$ (by Theorem 5.1.4), there exists a $g_0 \in \mathcal{X}^*$ such that $g = \widehat{g_0}$. Hence we obtain that

$$\mathrm{Re}(f_0(g_0)) \geq a > b \geq \sup\left(\left\{|f(g_0)| \mid f \in \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}\right\}\right).$$

However, notice that

$$\begin{aligned}\sup\left(\left\{|f(g_0)| \mid f \in \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}\right\}\right) &\geq \sup\left(\{|\widehat{x}(g_0)| \mid x \in \mathcal{X}_1\}\right) \\ &= \sup\left(\{|g_0(x))| \mid x \in \mathcal{X}_1\}\right) \\ &= \|g_0\|\end{aligned}$$

whereas

$$\mathrm{Re}(f_0(g_0)) \leq |f_0(g_0)| \leq \|f_0\| \, \|g_0\| \leq \|g_0\|$$

so that

$$\|g_0\| \geq a > b \geq \|g_0\|,$$

which is a clear contradiction. Hence $\mathcal{X}_1^{**} \setminus \overline{\mathcal{J}(\mathcal{X}_1)}^{w^*}$ must be empty so the result follows. ∎

Goldstine's Theorem (Theorem 5.4.1) has several theoretical applications that relate properties of a space and properties induced by a dual topology.

Proposition 5.4.2. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. Then $\mathcal{X}$ is reflective if and only if $\mathcal{X}_1$ is weakly compact.*

Proof. Suppose $\mathcal{X}$ is reflective. Hence the canonical embedding $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ is an isometric isomorphism. Since the weak topology on $\mathcal{X}_1$ is equal to the weak*-topology on $\mathcal{J}(\mathcal{X}_1) = \mathcal{X}_1^{**}$, and since $\mathcal{X}_1^{**}$ is weak*-compact by the Banach-Alaoglu Theorem (Theorem 5.3.4), it follows that $\mathcal{X}_1$ is weakly compact.

Conversely, suppose $\mathcal{X}_1$ is weakly compact. To see that $\mathcal{X}$ is reflexive, consider the canonical embedding $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$. Since the weak topology on $\mathcal{X}_1$ is equal to the weak*-topology on $\mathcal{J}(\mathcal{X}_1)$, it follows that $\mathcal{J}$ is a weak-weak* homeomorphism onto is range so $\mathcal{J}(\mathcal{X}_1)$ is weak*-compact. Thus, as the weak*-topology on $\mathcal{X}^{**}$ is Hausdorff, $\mathcal{J}(\mathcal{X}_1)$ is weak*-closed. However, since $\mathcal{J}$ is contractive and since $\mathcal{J}(\mathcal{X}_1)$ is dense in X_1^{**} by Goldstine's Theorem (Theorem 5.4.1), it follows that $\mathcal{J}(\mathcal{X}_1) = \mathcal{X}_1^{**}$. Hence $\mathcal{J}(\mathcal{X}) = \mathcal{X}^{**}$ so $\mathcal{X}$ is reflexive. ∎

Theorem 5.4.3. *If $(\mathcal{X}, \|\cdot\|)$ is a normed linear space, then $\mathcal{X}_1^*$ is weak*-metrizable if and only if $\mathcal{X}$ is separable.*

Proof. Suppose that $\mathcal{X}$ is separable. Hence there exists a set $\{x_n\}_{n=1}^\infty \subseteq \mathcal{X}$ that is dense in $\mathcal{X}$. Define $d : \mathcal{X}_1^* \times \mathcal{X}_1^* \to [0, \infty)$ by

$$d(f, g) = \sum_{n=1}^\infty \frac{|f(x_n) - g(x_n)|}{2^{n+1}(\|x_n\| + 1)}$$

for all $f, g \in \mathcal{X}_1^*$. Clearly $d(f, g) \geq 0$ for all $f, g \in \mathcal{X}_1^*$ and since

$$|f(x_n) - g(x_n)| \leq \|f - g\| \, \|x_n\| \leq \|x_n\|,$$

it follows that $d(f, g) \leq 1$.

We claim that d is a metric on $\mathcal{X}_1^*$ the induces the weak*-topology. To see that d is a metric, clearly $d(f, f) = 0$ for all $f \in \mathcal{X}_1^*$. Next, suppose $f, g \in \mathcal{X}_1^*$ are such that $d(f, g) = 0$. Hence $f(x_n) = g(x_n)$ for all $n \in \mathbb{N}$. Therefore, since f and g are continuous and $\{x_n\}_{n=1}^\infty$ is dense in $\mathcal{X}$, it follows that

$f(x) = g(x)$ for all $x \in \mathcal{X}$. Hence $f = g$. As it is clear that $d(f, g) = d(g, f)$ for all $f, g \in \mathcal{X}_1^*$ and the triangle inequality holds, d is a metric on $\mathcal{X}_1^*$.

To see that the metric topology on $\mathcal{X}_1^*$ induced by d agrees with the weak*-topology, suppose that $(f_\lambda)_{\lambda\in\Lambda}$ is a net in $\mathcal{X}_1^*$ that weak*-converges to $f \in \mathcal{X}_1^*$. To see that $(f_\lambda)_{\lambda\in\Lambda}$ converges to f with respect to d, let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} \frac{1}{2^n} < \frac{\epsilon}{2}$$

and note for all $\lambda \in \Lambda$ that

$$\sum_{n=N+1}^{\infty} \frac{|f_\lambda(x_n) - f(x_n)|}{2^{n+1}(\|x_n\| + 1)} \leq \sum_{n=N+1}^{\infty} \frac{2\,\|x_n\|}{2^{n+1}(\|x_n\| + 1)} < \frac{\epsilon}{2}$$

by the above computations. Next, note for each $n \in \{1, \ldots, N\}$ that since $(f_\lambda(x_n))_{\lambda\in\Lambda}$ converges to $f(x_n)$ there exists an $\lambda_n \in \Lambda$ such that

$$\frac{|f_\lambda(x_n) - f(x_n)|}{2^{n+1}(\|x_n\| + 1)} < \frac{\epsilon}{2N}$$

for all $\lambda \geq \lambda_n$. Let $\lambda_0 = \max(\{\lambda_1, \ldots, \lambda_n\})$. Thus for all $\lambda \geq \lambda_0$ we have that

$$\begin{aligned}
d(f_\lambda, f) &= \sum_{n=1}^{\infty} \frac{|f_\lambda(x_n) - f(x_n)|}{2^{n+1}(\|x_n\| + 1)} \\
&\leq \frac{\epsilon}{2} + \sum_{n=1}^{N} \frac{|f_\lambda(x_n) - f(x_n)|}{2^{n+1}(\|x_n\| + 1)} \\
&< \frac{\epsilon}{2} + N\frac{\epsilon}{2N} = \epsilon.
\end{aligned}$$

Therefore, as $\epsilon > 0$ was arbitrary, we obtain that $(f_\lambda)_{\lambda\in\Lambda}$ converges to f with respect to d.

Conversely, suppose $(f_\lambda)_{\lambda\in\Lambda}$ is a net in $\mathcal{X}_1^*$ that converges to f with respect to d. To see that $(f_\lambda)_{\lambda\in\Lambda}$ weak*-converges to f, let $x \in \mathcal{X}$ and $\epsilon > 0$ be arbitrary. Since $\{x_n\}_{n=1}^\infty$ is dense in $\mathcal{X}$, there exists an $n \in \mathbb{N}$ such that $\|x - x_n\| < \frac{\epsilon}{3}$. Moreover, since

$$|f_\lambda(x_n) - f(x_n)| \leq 2^{n+1}(\|x_n\| + 1)d(f_\lambda, f)$$

so $\lim_{\lambda\in\Lambda} f_\lambda(x_n) = f(x_n)$, it follows that there exists an $\lambda_0 \in \Lambda$ such that $|f_\lambda(x_n) - f(x_n)| < \frac{\epsilon}{3}$ for all $\lambda \geq \lambda_0$. Hence for all $\lambda \geq \lambda_0$

$$\begin{aligned}
|f_\lambda(x) - f(x)| &\leq |f_\lambda(x) - f_\lambda(x_n)| + |f_\lambda(x_n) - f(x_n)| + |f(x_n) - f(x)| \\
&\leq \|f_\lambda\|\,\|x - x_n\| + \frac{\epsilon}{3} + \|f\|\,\|x_n - x\| \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Therefore, as x and ϵ were arbitrary, $(f_\lambda)_{\lambda\in\Lambda}$ weak*-converges to f. Hence the weak*-topology on $\mathcal{X}_1^*$ is metrizable.

Conversely, suppose the weak*-topology on $\mathcal{X}_1^*$ is metrizable. Thus, as the open balls of radius $\frac{1}{n}$ centred at $\vec{0}$ have intersection $\{\vec{0}\}$, there is a sequence $(U_n)_{n\geq 1}$ of weak*-open sets in $\mathcal{X}_1^*$ such that $\bigcap_{n=1}^\infty U_n = \{\vec{0}\}$. Recall for each U_n there exists an $\epsilon_n > 0$ and a finite subset $F_n \subseteq \mathcal{X}$ such that

$$\{f \in \mathcal{X}_1^* \mid |f(x)| < \epsilon_n \text{ for all } x \in F_n\} = N(\vec{0}, F_n, \epsilon_n) \subseteq U_n.$$

Hence $\bigcap_{n=1}^\infty N(\vec{0}, F_n, \epsilon_n) = \{\vec{0}\}$.

Let $A = \bigcup_{n=1}^\infty F_n$. By the above, if $f \in \mathcal{X}_1^*$ and $f(x) = 0$ for all $x \in A$, then $f \in N(\vec{0}, F_n, \epsilon_n)$ for all $n \in \mathbb{N}$ so $f = 0$.

Let $\mathcal{W} = \overline{\mathrm{span}}(A)$, which is a separable subspace of $\mathcal{X}$ (i.e. $\mathbb{K}$ is separable and A is countable so there is a countable dense subset of $\mathrm{span}(A)$). Thus, if $\mathcal{W} = \mathcal{X}$, the proof is complete.

Suppose to the contrary that there exists an $x \in \mathcal{X} \setminus \mathcal{W}$. By the Hahn-Banach Theorem (Corollary 4.3.10) there exists an $f \in \mathcal{X}_1^*$ such that $\|f\| = 1$, $|f(x)| = \|x\|$, and $f|_{\mathcal{W}} = 0$. However, the above shows $f|_{\mathcal{W}} = 0$ implies $f = 0$, which contradicts the fact that $\|f\| = 1$. Thus $\mathcal{W} = \mathcal{X}$ completing the proof. ∎

Corollary 5.4.4. *If $(\mathcal{X}, \|\cdot\|)$ is a separable normed linear space, then $\mathcal{X}_1^*$ is separable in the weak*-topology.*

Proof. By Theorem 5.4.3, the weak*-topology on $\mathcal{X}_1^*$ is metrizable. Since $\mathcal{X}_1^*$ is weak*-compact by the Banach-Alaoglu Theorem (Theorem 5.3.4), $\mathcal{X}_1^*$ must be totally bounded and therefore separable. ∎

Remark 5.4.5. Note Corollary 5.4.4 is the best that one can hope for in that there exists separable normed linear spaces $(\mathcal{X}, \|\cdot\|)$ such that $\mathcal{X}_1^*$ is not separable in the norm topology. Indeed $\ell_1(\mathbb{N})$ is readily verified to be separable whereas Theorem 1.5.4 implies that $\ell_1(\mathbb{N})^* = \ell_\infty(\mathbb{N})$ for which the closed unit ball is not separable (all the sequences of 0s and 1s cannot be finitely covered by open balls of radius $\frac{1}{2}$).

Of course, we can reverse the roles of the initial space and its dual space.

Theorem 5.4.6. *If $(\mathcal{X}, \|\cdot\|)$ is a normed linear space, then $\mathcal{X}_1$ is weakly metrizable if and only if $\mathcal{X}^*$ is separable.*

Proof. Suppose $\mathcal{X}^*$ is separable. Thus Theorem 5.4.3 implies that $\mathcal{X}_1^{**}$ is weak*-metrizable. This implies if $\mathcal{J} : \mathcal{X} \to \mathcal{X}^{**}$ is the canonical embedding then $\mathcal{J}(\mathcal{X}_1)$ is weak*-metrizable and thus $\mathcal{X}_1$ is weakly metrizable as $\mathcal{J}$ is a weak-weak* homeomorphism of $\mathcal{X}_1$ onto $\mathcal{J}(\mathcal{X}_1)$.

Conversely, suppose $\mathcal{X}_1$ is weakly metrizable. Let U_n denote the weakly open ball of radius $\frac{1}{n}$ centred as $\vec{0}$ and recall for each U_n there exists an $\epsilon_n > 0$ and a finite subset $F_n \subseteq \mathcal{X}^*$ such that

$$\{x \in \mathcal{X}_1 \mid |f(x)| < \epsilon_n \text{ for all } f \in F_n\} = N(\vec{0}, F_n, \epsilon_n) \subseteq U_n.$$

Let $F = \bigcup_{n=1}^\infty F_n$ and let $\mathcal{W} = \overline{\mathrm{span}}(F)$, which is a separable subspace of $\mathcal{X}^*$ (i.e. $\mathbb{K}$ is separable and A is countable so there is a countable dense subset of $\mathrm{span}(A)$). Thus, if $\mathcal{W} = \mathcal{X}^*$, the proof is complete.

Suppose to the contrary that there exists an $f_0 \in \mathcal{X}^* \setminus \mathcal{W}$. Let $d = \mathrm{dist}(f_0, \mathcal{W})$. By the Hahn-Banach Theorem (Corollary 4.3.10) there exists an $g \in \mathcal{X}_1^{**}$ such that $\|g\| = 1$, $g(f_0) = d$, and $g|_{\mathcal{W}} = 0$.

Let
$$V_0 = \left\{ x \in \mathcal{X}_1 \,\middle|\, |f_0(x)| < \frac{d}{2} \right\}.$$

As V_0 is a weakly open neighbourhood of $\vec{0}$, there exists an $N \in \mathbb{N}$ such that $U_N \subseteq V_0$. Let

$$V = \left\{ h \in \mathcal{X}^{**} \,\middle|\, {}^{|h(f) - g(f)| < \epsilon_N \text{ for all } f \in F_N}_{\text{and } |h(f_0) - g(f_0)| < \frac{d}{2}} \right\}$$

which is clearly a weak*-neighbourhood of g. Since Goldstine's Theorem (Theorem 5.4.1) implies there exists an $x_0 \in \mathcal{X}_1$ such that $\widehat{x}_0 \in V$. Since $F_N \subseteq \mathcal{W}$ and $g|_{\mathcal{W}} = 0$, this implies that

$$|f(x_0)| = |\widehat{x_0}(f) - g(f)| < \epsilon_N$$

for all $f \in F_N$ and thus $x_0 \in U_N \subseteq V_0$. Hence $|f_0(x_0)| < \frac{d}{2}$. However, this implies that
$$|g(f_0) - \widehat{x_0}(f_0)| = d - \widehat{x_0}(f_0) > \frac{d}{2}.$$
which contradicts the fact that $\widehat{x_0} \in V$. Therefore, as we have obtained a contradiction, $\mathcal{X}^* = \mathcal{W}$ as desired. ∎

Remark 5.4.7. Theorem 5.4.6 can also be used to show that the weak topology on $\ell_1(\mathbb{N})$ is not metrizable. Indeed as $\ell_1(\mathbb{N})^* = \ell_\infty(\mathbb{N})$ by Theorem 1.5.4 and $\ell_\infty(\mathbb{N})$ is not separable, Theorem 5.4.6 implies that the weak topology on the unit ball of $\ell_1(\mathbb{N})$ is not metrizable and thus the weak topology on $\ell_1(\mathbb{N})$ cannot be metrizable.

5.5 The Krein-Milman Theorem

To conclude this chapter, we will look at one of the fundamental results in elementary functional analysis that characterizes any non-empty, compact, convex set via a simpler collection of points. As the closed unit ball of a normed linear space is weak*-compact by the Banach-Alaoglu Theorem

(Theorem 5.3.4), this is particularly useful for simplifying the study of dual spaces. We begin with the definition of the points we are interested in using to describe such sets and examples of how these points arise in geometry.

Definition 5.5.1. Let $\mathcal{V}$ be a vector space and let $C \subseteq \mathcal{V}$ be a non-empty convex set. A point $e \in C$ is said to be an *extreme point of* C if whenever $x, y \in C$ and $t \in (0,1)$ are such that

$$e = tx + (1-t)y,$$

then $x = y = e$. The set of all extreme points of C is denoted $\mathrm{Ext}(C)$.

For some examples, we note the following where we have left the proofs to the reader as they can be easily seen via basic Euclidean geometry.

Example 5.5.2. In $\mathbb{R}^2$, consider the unit square

$$C = \{(x, y) \in \mathbb{R}^2 \mid x, y \in [-1, 1]\}.$$

It is elementary to see that

$$\mathrm{Ext}(C) = \{(1,1), (1,-1), (-1,1), (-1,-1)\}.$$

Example 5.5.3. In $\mathbb{R}^2$, consider the closed disk

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}.$$

It is elementary to see that

$$\mathrm{Ext}(D) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Example 5.5.4. In $\mathbb{R}^2$, consider the open disk

$$D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}.$$

It is elementary to see that $\mathrm{Ext}(D) = \emptyset$.

Of course, there is a generalization of the notion of an extreme point that allows for more flexibility.

Definition 5.5.5. Let $\mathcal{V}$ be a vector space and let $C \subseteq \mathcal{V}$ be a non-empty convex set. A non-empty convex subset $F \subseteq C$ is said to be an *face of* C if whenever $x, y \in C$ and $t \in (0,1)$ are such that

$$tx + (1-t)y \in F$$

then $x, y \in F$.

Clearly any non-empty convex set is a face of itself and a singleton $\{e\}$ is a face if and only if e is an extreme point. The reason for the name 'face' is the following example.

Example 5.5.6. In $\mathbb{R}^3$, consider the unit cube

$$C = \{(x, y, z) \in \mathbb{R}^3 \mid x, y, z \in [-1, 1]\}.$$

Clearly

$$\mathrm{Ext}(C) = \{(x, y, z) \mid x, y, z \in \{\pm 1\}\}$$

and each of these singleton points is a face. Furthermore, it elementary to verify that each edge of the cube is a face, and each face (from geometry) of the cube is a face (in this course). Finally, it is possible to verify that these are the only proper subsets that are faces of C.

Of course, the structure of faces of a cube satisfy the following result.

Lemma 5.5.7. *Let $\mathcal{V}$ be a vector space and let C be a non-empty convex subset of $\mathcal{V}$. If F is a face of C and let A is a face of F, then A is a face of C.*

Proof. To see that A is a face of C, let $x, y \in C$ and $t \in (0, 1)$ be such that

$$tx + (1 - t)y \in A.$$

Since A is a face of F, $A \subseteq F$ so $tx + (1 - t)y \in F$. Therefore, since F is a face of C, we obtain that $x, y \in F$. Finally, since $x, y \in F$, $tx + (1 - t)y \in A$, and A is a face of F, we obtain that $x, y \in A$. Hence A is a face of C (in your face!). ∎

Corollary 5.5.8. *Any extreme point of any face of a non-empty convex set C is an extreme point of C.*

Unsurprisingly, we can extend the definition of a face to include more than just convex combinations of two points.

Lemma 5.5.9. *Let $\mathcal{V}$ be a vector space, let C be a non-empty convex subset of $\mathcal{V}$, and let F be a face of C. If $\{x_k\}_{k=1}^n \subseteq C$ and $\{t_k\}_{k=1}^n \subseteq (0, 1)$ are such that $\sum_{k=1}^n t_k = 1$ and*

$$\sum_{k=1}^{n} t_k x_k \in F,$$

then $x_k \in F$ for all $k \in \{1, \ldots, n\}$.

Proof. We proceed by induction with the case $n = 2$ being trivial. Suppose the result holds for some $n \in \mathbb{N}$ with $n \geq 2$ and let $\{x_k\}_{k=1}^{n+1} \subseteq C$ and $\{t_k\}_{k=1}^{n+1} \subseteq (0, 1)$ be such that $\sum_{k=1}^{n+1} t_k = 1$ and

$$\sum_{k=1}^{n+1} t_k x_k \in F.$$

Let $t = t_{n+1}$ so that $t \in (0,1)$ and $1 - t = \sum_{k=1}^n t_k$, and let

$$y = \sum_{k=1}^n \frac{t_k}{1-t} x_k.$$

Since $\frac{t_k}{1-t} \in (0,1)$ for all $k \in \{1, \ldots, n\}$ and since $\sum_{k=1}^n \frac{t_k}{1-t} = 1$, we obtain that $y \in C$ as C is convex. By construction

$$tx_{n+1} + (1-t)y = \sum_{k=1}^{n+1} t_k x_k \in F.$$

Therefore, since F is a face of C, we obtain that $x_{n+1}, y \in F$. As $y \in F$, $\frac{t_k}{1-t} \in (0,1)$ for all $k \in \{1, \ldots, n\}$, and since $\sum_{k=1}^n \frac{t_k}{1-t} = 1$, the induction hypothesis implies that $x_k \in F$ for all $k \in \{1, \ldots, n\}$ thereby completing the inductive step and the proof. ∎

With the preliminaries on faces and extreme points out of the way, we can proceed to begin our examination of extreme points of non-empty, compact, convex sets in locally convex topological vector spaces. We begin with the following method of constructing faces.

Lemma 5.5.10. *Let $(\mathcal{V}, \mathcal{T})$ be a locally convex topological vector space and let K be a non-empty, convex, compact subset of $\mathcal{V}$. If $f \in \mathcal{V}^*$ and*

$$r = \sup(\{\mathrm{Re}(f(x)) \mid x \in K\}),$$

then

$$F = \{x \in K \mid \mathrm{Re}(f(x)) = r\}$$

is a compact face of K.

Proof. To begin, we note since K is compact that the Extreme Value Theorem implies that F is non-empty. Moreover, as F is clearly a closed subset of K, F is compact.

To see that F is convex, let $x, y \in F$ and $t \in [0,1]$ be arbitrary. Then $\mathrm{Re}(f(x)) = r$ and $\mathrm{Re}(f(y)) = r$ so

$$\mathrm{Re}(f(tx + (1-t)y)) = t\mathrm{Re}(f(x)) + (1-t)\mathrm{Re}(f(y)) = tr + (1-t)r = r$$

and thus $tx + (1-t)y \in F$. Hence F is convex.

To see that F is a face of K, let $x, y \in K$ and $t \in (0,1)$ be such that $tx + (1-t)y \in F$. Therefore, since $\mathrm{Re}(f(x)) \leq r$ and $\mathrm{Re}(f(y)) \leq r$,

$$r = \mathrm{Re}(f(tx + (1-t)y)) = t\mathrm{Re}(f(x)) + (1-t)\mathrm{Re}(f(y)) \leq tr + (1-t)r = r.$$

Hence equality must occur in $\mathrm{Re}(f(x)) \leq r$ and $\mathrm{Re}(f(y)) \leq r$ so $x, y \in F$ as desired. ∎

With Lemma 5.5.10 giving us a method of constructing faces using linear functionals, we can use a maximality argument together with the Hahn-Banach Theorem to obtain extreme points.

Lemma 5.5.11. *If $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space and K is a non-empty, convex, compact subset of $\mathcal{V}$, then* $\mathrm{Ext}(K) \neq \emptyset$.

Proof. Let
$$\mathcal{F} = \{F \subseteq K \mid F \text{ is a closed face of } K\}.$$
Clearly $\mathcal{F}$ is non-empty since $K \in \mathcal{F}$. Define a partial ordering $\preceq$ on $\mathcal{F}$ by reverse inclusion; that is $F_1 \preceq F_2$ if and only if $F_1 \supseteq F_2$.

Let $\mathcal{C}$ be a chain in $\mathcal{F}$ and let $F_0 = \bigcap_{F \in \mathcal{C}} F$. We claim that $F_0 \in \mathcal{C}$ and thus F_0 is an upper bound for $\mathcal{C}$. Since $\mathcal{C}$ is a chain, $\mathcal{C}$ has the finite intersection property. Therefore as any collection of closed sets with the finite intersection property in a compact Hausdorff space has non-empty intersection, we obtain that $F_0 \neq \emptyset$. Clearly F_0 is closed and convex being the intersection of closed convex sets. Finally, to see that F_0 is a face of K, suppose $x, y \in K$ and $t \in (0, 1)$ are such that $tx + (1-t)y \in F_0$. Thus $tx + (1-t)y \in F$ for all $F \in \mathcal{C}$ so $x, y \in F$ as F is a face of K. Hence $x, y \in F_0$ so F_0 is a face of K.

By Zorn's Lemma there exists a maximal closed face F_m of K. If $F_m = \{e\}$ for some $e \in K$, then e is an extreme point of K. Suppose otherwise that there exists $x, y \in F_m$ such that $x \neq y$. As $\{x\}$ and $\{y\}$ are non-empty, disjoint, compact, convex sets of $\mathcal{V}$, the Hahn-Banach Theorem (Theorem 4.4.13) implies there exists an $f \in \mathcal{V}^*$ and $a, b \in \mathbb{R}$ such that
$$\mathrm{Re}(f(x)) \geq a > b \geq \mathrm{Re}(f(y)).$$

Let
$$r = \sup(\{\mathrm{Re}(f(z)) \mid z \in F_m\}) \geq a.$$
As F_m is a closed face of K, F_m is a non-empty, convex, compact subset of $\mathcal{V}$. Therefore Lemma 5.5.10 implies that
$$\widetilde{F_m} = \{z \in F_m \mid \mathrm{Re}(f(z)) = r\}$$
is a compact face of F_m. Since $(\mathcal{V}, \mathcal{T})$ is Hausdorff, $\widetilde{F_m}$ is a closed subset of $(\mathcal{V}, \mathcal{T})$. Moreover, since $\widetilde{F_m}$ is a face of F_m and since F_m is a face of K, $\widetilde{F_m}$ is a face of K by Lemma 5.5.7. However, since
$$\mathrm{Re}(f(y)) \leq b < a \leq r,$$
we see that $y \notin \widetilde{F_m}$ so $\widetilde{F_m}$ is a strict subset of F_m. As this contradicts the maximality of F_m, we have our contradiction. Hence F_m is a singleton thereby yielding an extreme point of K. ∎

With Lemma 5.5.11 we can now prove the main result of this section. Note as the proof of Lemma 5.5.11 is the vast majority of the work, Lemma 5.5.11 is often considered as part of this theorem.

Theorem 5.5.12 (The Krein-Milman Theorem). *If $(\mathcal{V}, \mathcal{T})$ is a locally convex topological vector space and K is a non-empty, convex, compact subset of $\mathcal{V}$, then*

$$K = \overline{\mathrm{conv}}(\mathrm{Ext}(K)).$$

Proof. By Theorem 5.5.11 we know that $\mathrm{Ext}(K) \neq \emptyset$. Let $C = \overline{\mathrm{conv}}(\mathrm{Ext}(K))$ which is a non-empty, closed, convex subset of K.

Suppose there exists an $x \in K \setminus C$. As $\{x\}$ and C are non-empty, disjoint, closed, convex subsets of $\mathcal{V}$ such that $\{x\}$ is compact, the Hahn-Banach Theorem (Theorem 4.4.13) implies there exists an $f \in \mathcal{V}^*$ and $a, b \in \mathbb{R}$ such that

$$\mathrm{Re}(f(x)) \geq a > b \geq \mathrm{Re}(f(c))$$

for all $c \in C$.

Let

$$r = \sup(\{\mathrm{Re}(f(y)) \mid y \in K\}) \geq a.$$

Thus Lemma 5.5.10 implies that

$$F = \{y \in K \mid \mathrm{Re}(f(y)) = r\}$$

is a compact face of K. Hence Lemma 5.5.11 implies that there exists a extreme point e of F. Therefore, since F is a face of K, e is an extreme point of K by Corollary 5.5.8 and thus $e \in C$. Since $e \in F$, we know that $\mathrm{Re}(f(e)) = r$. However, since $e \in C$, we know that

$$\mathrm{Re}(f(e)) \leq b < a \leq r = \mathrm{Re}(f(e))$$

which is a clear contradiction. Hence $K = C$ as desired. ∎

One immediate application of the Krein-Milman Theorem (Theorem 5.5.12) is an extension of one of our previous results.

Corollary 5.5.13. *Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space and let $A \subseteq \mathcal{X}^*$. If A is bounded, convex, and weak*-closed, then A is weak*-compact and $A = \overline{\mathrm{conv}}^{w^*}(\mathrm{Ext}(A))$.*

Proof. By Corollary 5.3.5 we know that A is weak*-compact. Hence the Krein-Milman Theorem (Theorem 5.5.12) implies that $A = \overline{\mathrm{conv}}^{w^*}(\mathrm{Ext}(A))$. ∎

Of course, the Krein-Milman Theorem has other applications.

Corollary 5.5.14. *The spaces c and c_0 are isomorphic but not isometrically isomorphic as normed linear spaces.*

Proof. Exercise. ∎

Corollary 5.5.15. *The space of real-valued continuous functions* $C[0,1]$ *is not the continuous dual space of a normed linear space.*

Proof. Exercise. ∎

Chapter 6

Operator Theory

Now that we have dealt with some of the most general objects in functional analysis, we proceed in the opposite manner. In particular, instead of generalizing normed linear spaces to topological vector spaces we will become more specific and study complete normed linear spaces with norms coming from an inner product structure, which are called Hilbert spaces. By restricting our study to bounded linear operators on Hilbert spaces, we can develop a deep theory of operators.

To begin, we will study a nice collection of bounded linear operators between Banach spaces that play an introductory role in operator theory. Subsequently, we will develop the theory of Hilbert spaces. It will be shown that Hilbert spaces have orthonormal bases (i.e. maximal orthonormal sets, not orthonormal sets that are vector space bases) and Hilbert spaces can be characterized by the cardinality of their orthonormal bases. We will then begin an introduction to the theory of operators on Hilbert space culminating in a generalization of the spectral theorem for normal matrices. In particular, this entire chapter can be viewed as generalizing the results from finite dimensional inner product spaces to the infinite dimensional setting.

6.1 Compact Operators on Banach Spaces

To begin our study of operator theory, we will examine the following collection of operators, which will end up being one of the nicest ideals of the bounded linear operators on a Banach space.

Definition 6.1.1. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces. An element $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is said to be *compact* if $\overline{T(\mathcal{X}_1)}$ is compact in $\mathcal{Y}$. The set of all compact operators from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ and by $\mathcal{K}(\mathcal{X})$ in the case that $\mathcal{Y} = \mathcal{X}$.

In order to construct examples of compact operators, we note the following class of operators which will easily seen to be compact.

Definition 6.1.2. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces. An element $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is said to be of *finite rank* if $\dim(T(\mathcal{X}))$ is finite. The set of all finite rank operators from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ and by $\mathcal{F}(\mathcal{X})$ in the case that $\mathcal{Y} = \mathcal{X}$.

Proposition 6.1.3. *If $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ are Banach spaces, then $\mathcal{F}(\mathcal{X}, \mathcal{Y}) \subseteq \mathcal{K}(\mathcal{X}, \mathcal{Y})$.*

Proof. Let $T \in \mathcal{F}(\mathcal{X}, \mathcal{Y})$. Thus $\mathcal{W} = T(\mathcal{X})$ is a finite dimensional subspace of $\mathcal{Y}$. As $T(\mathcal{X}_1)$ is a bounded subset of $\mathcal{W}$ as T is bounded, $\overline{T(\mathcal{X}_1)}$ is a closed bounded subset of a finite dimensional normed linear space and thus compact by Theorem 3.5.2. Hence $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ as desired. ∎

As the following shows, except in the finite dimensional setting, there will always be operators that are not compact.

Proposition 6.1.4. *Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space. Then $\mathcal{K}(\mathcal{X}) = \mathcal{B}(\mathcal{X})$ if and only if $\mathcal{X}$ is finite dimensional.*

Proof. If $\mathcal{X}$ is finite dimensional, then $\mathcal{B}(\mathcal{X}) = \mathcal{F}(\mathcal{X}) \subseteq \mathcal{K}(\mathcal{X}) \subseteq \mathcal{B}(\mathcal{X})$ by Proposition 6.1.3 and thus $\mathcal{K}(\mathcal{X}) = \mathcal{B}(\mathcal{X})$.

Conversely, suppose $\mathcal{K}(\mathcal{X}) = \mathcal{B}(\mathcal{X})$. Thus the identity map $I_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ is a compact operator so $\overline{I_{\mathcal{X}}(\mathcal{X}_1)} = \overline{\mathcal{X}_1} = \mathcal{X}_1$ is a compact subset of $\mathcal{X}$. As this implies $(\mathcal{X}, \|\cdot\|)$ is locally compact, $\mathcal{X}$ is finite dimensional by Corollary 3.5.12. ∎

For an example of a compact operator that is not finite dimensional, we turn to the following use of some results from compact metric space theory.

Example 6.1.5. Fix $g \in C[0,1]$ and define $T_g : C[0,1] \to C[0,1]$ by

$$T_g(f)(x) = \int_0^x f(t)g(t)\, dt$$

for all $f \in C[0,1]$. It is elementary to verify that T_g is a well-defined linear operator. Furthermore, since

$$|T_g(f)(x)| \leq \int_0^x |f(t)||g(t)|\, dt \leq \int_0^x \|f\|_\infty \|g\|_\infty \, dt \leq \|f\|_\infty \|g\|_\infty$$

for all $x \in [0,1]$, we see that T_g is bounded with $\|T_g\| \leq \|g\|_\infty$.

We claim that T_g is a compact operator. To see this, it suffices to show that

$$\mathcal{F} = \left\{ h \in C[0,1] \,\middle|\, h(x) = \int_0^x f(t)g(t)\, dt \text{ for some } f \in C[0,1] \text{ with } \|f\|_\infty \leq 1 \right\}$$

has compact closure. Thus, by the Arzel'a-Ascoli Theorem, it suffices to prove that $\mathcal{F}$ is equicontinuous and pointwise bounded.

Clearly the elements of $\mathcal{F}$ are pointwise bounded by $\|g\|_\infty$ by the above computation. To see that the elements of $\mathcal{F}$ are equicontinuous, let $\epsilon > 0$ and $x \in [0,1]$ be arbitrary. Let $\delta = \frac{\epsilon}{\|g\|_\infty + 1}$. Thus for all $f \in C[0,1]$ with $\|f\|_\infty \leq 1$ and for all $y \in [0,1]$ with $|x-y| < \delta$, we see that

$$\begin{aligned}\left|\int_0^y f(t)g(t)\,dt - \int_0^x f(t)g(t)\,dt\right| &= \left|\int_y^x f(t)g(t)\,dt\right| \\ &\leq |x-y|\,\|f\|_\infty\,\|g\|_\infty \\ &< \frac{\epsilon}{\|g\|_\infty + 1}(1)\,\|g\|_\infty < \epsilon.\end{aligned}$$

Hence $\mathcal{F}$ is equicontinuous as desired. Thus T_g is a compact operator.

In order to construct additional examples of compact operators, it is useful to develop all equivalent characterizations of compact operators via the different methods of verifying a subset of a metric space is compact.

Proposition 6.1.6. *Given Banach spaces $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ and a $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, the following are equivalent:*

(1) T is compact.

(2) $\overline{T(F)}$ is compact in $\mathcal{Y}$ for all bounded subsets F of $\mathcal{X}$.

(3) If $(x_n)_{n\geq 1}$ is a bounded sequence in $\mathcal{X}$, then $(T(x_n))_{n\geq 1}$ has a convergent subsequence in $\mathcal{Y}$.

(4) $T(\mathcal{X}_1)$ is totally bounded.

Proof. Note that T is compact if and only if $\overline{T(\mathcal{X}_1)}$ is compact if and only if $T(\mathcal{X}_1)$ is totally bounded. Hence (1) and (4) are equivalent. Moreover, since the notions of compactness and sequential compactness are equivalent in metric spaces, (2) and (3) are equivalent.

Clearly (2) implies (1). Conversely, if $\overline{T(\mathcal{X}_1)}$ is compact, we obtain that $\overline{T(\mathcal{X}_r)}$ is compact for any $r > 0$. As a closed subset of a compact subset is compact in normed linear spaces, we obtain that (1) implies (2). ∎

With Proposition 6.1.6 complete, we can produce some nice non-trivial examples of compact operators.

Example 6.1.7. Fix $p \in [1,\infty]$ and $\vec{d} = (d_n)_{n\geq 1} \in \ell_\infty(\mathbb{N})$. Define $D : \ell_p(\mathbb{N}) \to \ell_p(\mathbb{N})$ by

$$D((x_n)_{n\geq 1}) = (d_n x_n)_{n\geq 1}$$

for all $(x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$. Note D is a well-defined bounded linear operator since

$$\|(d_n x_n)_{n\geq 1}\|_p \leq \left\|\vec{d}\right\|_\infty \|(x_n)_{n\geq 1}\|_p$$

for all $(x_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ by a simple computation.

We claim that D is compact if and only if $\vec{d} \in c_0$. To see this, note if $\vec{d} \notin c_0$ then there exists an $\epsilon > 0$ and an infinite increasing sequence $(n_k)_{k\geq 1}$ of natural numbers such that $|d_{n_k}| \geq \epsilon$ for all $k \in \mathbb{N}$. Thus, if $\vec{e}_{n_k}$ is the element of $\ell_p(\mathbb{N})$ with a 1 in the n_k^{th} entry and zeros everywhere else, we see that $(\vec{e}_{n_k})_{k\geq 1}$ is a bounded sequence in $\ell_p(\mathbb{N})$ such that $(D(\vec{e}_{n_k}))_{k\geq 1} = (d_{n_k}\vec{e}_{n_k})_{k\geq 1}$ has no convergent subsequences (i.e. the only possible limit point is $\vec{0}$ and $|d_{n_k}| \geq \epsilon$ for all $k \in \mathbb{N}$). Hence if $\vec{d} \notin c_0$, then D is not compact by Proposition 6.1.6.

Conversely, suppose that $\vec{d} \in c_0$. To see that D is compact, we will use Proposition 6.1.6 and show that $A = D((\ell_p(\mathbb{N}))_1)$ is totally bounded. Thus let $\epsilon > 0$ be arbitrary. Since $\vec{d} \in c_0$ there exists an $N \in \mathbb{N}$ such that $|d_n| < \frac{\epsilon}{2}$ for all $n \in \mathbb{N}$.

As $(\mathbb{K}^N, \|\cdot\|_p)$ is finite dimensional, any closed bounded set is compact and thus any bounded set is totally bounded. Since

$$A_0 = \left\{(d_1x_1, \ldots, d_Nx_N) \mid (x_1, \ldots, x_n) \in B_{\|\cdot\|_p}[\vec{0}, 1]\right\}$$

is a bounded subset, we obtain that A_0 is totally bounded. Hence there exists a finite set $\{\vec{x}_k\}_{k=1}^m$ of $B_{\|\cdot\|_p}[\vec{0}, 1]$ such that if $\vec{x}_k = (x_{k,1}, \ldots, x_{k,N})$ for all $k \in \{1, \ldots, m\}$, then

$$E_0 = \{(d_1x_{k,1}, \ldots, d_Nx_{k,N})\}_{k=1}^m$$

is an $\frac{\epsilon}{2}$-net for A_0.

For each $k \in \{1, \ldots, m\}$, let

$$\vec{z}_k = (x_{k,1}, \ldots, x_{k,N}, 0, 0, 0, \ldots).$$

Clearly $\vec{z}_k \in c_{00} \in \ell_p(\mathbb{N})$ and $\|\vec{z}_k\|_p = \|\vec{x}_k\|_p \leq 1$ for all $k \in \{1, \ldots, m\}$. We claim that $\{D(\vec{z}_k)\}_{k=1}^m$ is an ϵ-net for A. To see this, let $\vec{a} \in A$ be arbitrary. Hence there exists a $\vec{y} = (y_n)_{n\geq 1} \in \ell_p(\mathbb{N})$ such that $\|\vec{y}\|_p \leq 1$ and $D(\vec{y}) = \vec{a}$. As $(y_1, \ldots, y_N) \in B_{\|\cdot\|_p}[\vec{0}, 1]$, the fact that D_0 is an $\frac{\epsilon}{2}$-net for A_0 implies there exists a $k \in \{1, \ldots, m\}$ such that

$$\|(d_1y_1, \ldots, d_Ny_N) - (d_1x_{k,1}, \ldots, d_Nx_{k,N})\|_p < \frac{\epsilon}{2}.$$

Hence, when $p \neq \infty$,

$$\begin{aligned}
&\|\vec{a} - D(\vec{z}_k)\|_p \\
&= \|(d_n y_n)_{n\geq 1} - (d_1 x_{k,1}, \ldots, d_N x_{k,N}, 0, 0, 0, \ldots)\|_p \\
&\leq \|(d_1 y_1, \ldots, d_N y_N) - (d_1 x_{k,1}, \ldots, d_N x_{k,N})\|_p + \left(\sum_{j=N}^{\infty} |d_j y_j|^p\right)^{\frac{1}{p}} \\
&< \frac{\epsilon}{2} + \left(\sum_{j=N}^{\infty} \left(\frac{\epsilon}{2}\right)^p |y_j|^p\right)^{\frac{1}{p}} \\
&= \frac{\epsilon}{2} + \frac{\epsilon}{2}\left(\sum_{j=N}^{\infty} |y_j|^p\right)^{\frac{1}{p}} \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2}\|\vec{y}\|_p \leq \epsilon
\end{aligned}$$

(where the triangle inequality was used to obtain the first inequality). As a similar computation holds when $p = \infty$, we obtain that $\{D(\vec{z}_k)\}_{k=1}^m$ is an ϵ-net for A. Therefore, as $\epsilon > 0$ was arbitrary, A is totally bounded so D is a compact operator.

One nice corollary of Proposition 6.1.6 is the following.

Corollary 6.1.8. *If $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ are Banach spaces, then $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ is a closed subspace of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

Proof. First, to see that $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ is a vector subspace of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$, let $T, S \in \mathcal{K}(X, \mathcal{Y})$ and $\alpha \in \mathbb{K}$ be arbitrary. To see that $\alpha T + S$ is compact, let $(x_n)_{n\geq 1}$ be an arbitrary bounded sequence in $\mathcal{X}$. Since T is compact, Proposition 6.1.6 implies there exists a subsequence $(x_{n_k})_{k\geq 1}$ of $(x_n)_{n\geq 1}$ such that $(T(x_{n_k}))_{k\geq 1}$ converges. Since S is compact, Proposition 6.1.6 implies there exists a subsequence $(x_{n_{k_m}})_{m\geq 1}$ of $(x_{n_k})_{k\geq 1}$ such that $(S(x_{n_{k_m}}))_{m\geq 1}$ converges. Hence $(T(x_{n_{k_m}}))_{m\geq 1}$ also converges so $((\alpha T + S)(x_{n_{k_m}}))_{m\geq 1}$ converges. Hence, as $(x_n)_{n\geq 1}$ was arbitrary, Proposition 6.1.6 implies that $\alpha T + S$ is compact. Thus $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ is a vector subspace of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

To see that $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ is closed, suppose $(T_n)_{n\geq 1}$ is a sequence in $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ that converges to some $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. To show that T is compact, it suffices by Proposition 6.1.6 to show that $T(\mathcal{X}_1)$ is totally bounded. Thus let $\epsilon > 0$ be arbitrary. Since $(T_n)_{n\geq 1}$ converges to T, there exists an $N \in \mathbb{N}$ such that $\|T - T_N\| < \frac{\epsilon}{3}$. Moreover, since $T_N(\mathcal{X}_1)$ is totally bounded by Proposition 6.1.6, there exists a finite subset $\{x_k\}_{k=1}^m \subseteq \mathcal{X}_1$ such that $\{T_N(x_k)\}_{k=1}^m$ is an $\frac{\epsilon}{3}$-net for $T_N(\mathcal{X}_1)$.

We claim that $\{T(x_k)\}_{k=1}^m$ is an ϵ-net for $T(\mathcal{X}_1)$. To see this, let $y_0 \in \mathcal{T}(\mathcal{X}_1)$ be arbitrary. Hence there exists an $x_0 \in \mathcal{X}_1$ such that $y_0 = T(x_0)$.

Since $\{T_N(x_k)\}_{k=1}^m$ is an $\frac{\epsilon}{3}$-net for $T_N(\mathcal{X}_1)$, there exists an $j \in \{1, \ldots, m\}$ such that $\|T_N(x_0) - T_N(x_j)\| < \frac{\epsilon}{3}$. Hence

$$\begin{aligned}\|y_0 - T(x_j)\| &\leq \|T(x_0) - T_N(x_0)\| + \|T_N(x_0) - T_N(x_j)\| + \|T_N(x_j) - T(x_j)\| \\ &\leq \|T - T_N\| \, \|x_0\| + \frac{\epsilon}{3} + \|T_N - T\| \, \|x_j\| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

Hence $\{T(x_k)\}_{k=1}^m$ is an ϵ-net for $T(\mathcal{X}_1)$. Therefore, as ϵ was arbitrary, $T(\mathcal{X}_1)$ is totally bounded and thus T is compact. Hence $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ is closed. ■

More importantly, when combined with Corollary 6.1.8, the following shows that the compact operators are an ideal in $\mathcal{B}(\mathcal{X})$ for any Banach space $\mathcal{X}$.

Corollary 6.1.9. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$, $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, and $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ be Banach spaces. If $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$, $R \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, and $S \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$, then $RT \in \mathcal{K}(\mathcal{X}, \mathcal{Z})$ and $TS \in \mathcal{K}(\mathcal{Z}, \mathcal{Y})$.*

Proof. To see that RT is compact, let $(x_n)_{n\geq 1}$ be an arbitrary bounded sequence in $\mathcal{X}$. Since T is compact, Proposition 6.1.6 there exists a subsequence $(x_{n_k})_{k\geq 1}$ such that $(T(x_{n_k}))_{k\geq 1}$ converges in $\mathcal{Y}$. Therefore, since R is continuous, $(RT(x_{n_k}))_{k\geq 1}$ converges. Therefore, since $(x_n)_{n\geq 1}$ was arbitrary, Proposition 6.1.6 implies that RT is compact.

To see that TS is compact, note as S is bounded that $S(\mathcal{X}_1)$ is a bounded subset of $\mathcal{X}$. Hence Proposition 6.1.6 implies that $\overline{T(S(\mathcal{X}_1)} = \overline{TS(\mathcal{X}_1)}$ is a compact subset of $\mathcal{Y}$. Hence TS is compact. ■

Finally, compact operators behave well with respect to adjoints.

Theorem 6.1.10. *Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be Banach spaces. If $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$, then $T^* \in \mathcal{K}(\mathcal{Y}^*, \mathcal{X}^*)$.*

Proof. Let $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. To show that $T^* \in \mathcal{K}(\mathcal{Y}^*, \mathcal{X}^*)$, it suffices by Proposition 6.1.6 to show that $T^*(\mathcal{Y}_1^*)$ is totally bounded. Thus let $\epsilon > 0$ be arbitrary. Since T is compact, $T(\mathcal{X}_1)$ is totally bounded and thus there exists $\{x_k\}_{k=1}^n \subseteq \mathcal{X}_1$ such that $\{T(x_k)\}_{k=1}^n$ is an $\frac{\epsilon}{3}$-net for $T(\mathcal{X}_1)$.

Define $R : \mathcal{Y}^* \to \mathbb{K}^n$ by

$$R(f) = (f(T(x_1)), \ldots, f(T(x_n)))$$

for all $f \in \mathcal{Y}^*$. Clearly R is a well-defined linear operator. Moreover, since $\mathbb{K}^n$ is finite dimensional, $\mathcal{R}$ is a finite dimensional operator and thus compact. Hence there exists $\{f_j\}_{j=1}^m \subseteq \mathcal{Y}_1^*$ such that $\{R(f_j)\}_{j=1}^m$ is an $\frac{\epsilon}{3}$-net for $R(\mathcal{Y}_1^*)$.

We claim that $\{T^*(f_j)\}_{j=1}^m$ is an ϵ-net for $T^*(\mathcal{Y}_1^*)$. To see this, let $g \in \mathcal{Y}_1^*$ be arbitrary. Since$\{R(f_j)\}_{j=1}^m$ is an $\frac{\epsilon}{3}$-net for $R(\mathcal{Y}_1^*)$, there exists a $j_0 \in \{1, \ldots, m\}$ such that $\|R(g) - R(f_{j_0})\| < \frac{\epsilon}{3}$.

We claim that $\|T^*(g) - T^*(f_{j_0})\| < \epsilon$. To see this, let $x \in \mathcal{X}_1$ be arbitrary. Since $\{T(x_k)\}_{k=1}^n$ is an $\frac{\epsilon}{3}$-net for $T(\mathcal{X}_1)$, there exists a $k_0 \in \{1, \ldots, n\}$ such that $\|T(x) - T(x_{k_0})\| < \frac{\epsilon}{3}$. Therefore

$$\begin{aligned}
&|T^*(g)(x) - T^*(f_{j_0})(x)| \\
&= |g(T(x)) - f_{j_0}(T(x))| \\
&\leq |g(T(x)) - g(T(x_{k_0}))| + |g(T(x_{k_0})) - f_{j_0}(T(x_{k_0}))| + |f_{j_0}(T(x_{k_0})) - f_{j_0}(T(x))| \\
&\leq \|g\| \, \|T(x) - T(x_{k_0})\| + \|R(g) - R(f_{j_0})\| + \|f_{j_0}\| \, \|T(x_{k_0}) - T(x)\| \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Therefore, as x was arbitrary, we obtain that $\|T^*(g) - T^*(f_{j_0})\| \leq \epsilon$. Hence, as g was arbitrary, $\{T^*(f_j)\}_{j=1}^m$ is an ϵ-net for $T^*(\mathcal{Y}_1^*)$ thereby completing the proof. ∎

6.2 Hilbert Spaces

We now turn to the nicest example of Banach spaces. As a Banach space is a complete normed linear space, and an inner product space is a normed linear space with a norm induced by an inner product, by combining these two notions we get the following. We refer a reader unfamiliar with inner product spaces to Chapter B.

Definition 6.2.1. A *Hilbert space* is a complete inner product space.

A few of the spaces we have already seen are actually Hilbert spaces.

Example 6.2.2. The space $\ell_2(\mathbb{N})$ is a Hilbert space. Indeed if we define $\langle \cdot, \cdot \rangle : \ell_2(\mathbb{N}) \times \ell_2(\mathbb{N}) \to \mathbb{K}$ by

$$\langle (x_n)_{n\geq 1}, (y_n)_{n\geq 1} \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

for all $(x_n)_{n\geq 1}, (y_n)_{n\geq 1} \in \ell_2(\mathbb{N})$, then $\langle \cdot, \cdot \rangle$ is well-defined inner product on $\ell_2(\mathbb{N})$ (see Example B.1.10) that induces $\|\cdot\|_2$. As $\ell_2(\mathbb{N})$ is complete with respect to $\|\cdot\|_2$, $\ell_2(\mathbb{N})$ is a Hilbert space.

Example 6.2.3. For any measure space $(X, \mathcal{A}, \mu)$, the space $L_2(X, \mu)$ is a Hilbert space (and when $X = \mathbb{N}$ and μ is the counting measure, we obtain $\ell_2(\mathbb{N})$). Indeed if we define $\langle \cdot, \cdot \rangle : L_2(X, \mu) \times L_2(X, \mu) \to \mathbb{K}$ by

$$\langle f, g \rangle = \int_X f \overline{g} \, d\mu$$

for all $f, g \in L_2(X, \mu)$, then $\langle \cdot, \cdot \rangle$ is well-defined inner product on $L_2(X, \mu)$ by Hölder's inequality (Theorem D.1.7) and the fact that integration against μ doesn't distinguish between functions that are equal μ-almost everywhere. Since $\langle \cdot, \cdot \rangle$ induces $\|\cdot\|_2$ and since $L_2(X, \mu)$ is complete by the Riesz-Fisher Theorem (Theorem D.2.1), we obtain that $L_2(X, \mu)$ is a Hilbert space.

In fact, we will show shortly that the above examples are really the only examples of Hilbert spaces in existence. Before we get to that, we need to analyze why Hilbert spaces are particularly nice. As inner products give length and angles to vector spaces, it is unsurprising that Hilbert spaces have a 'nicer geometry' than Banach spaces. In particular, we have the following which will be of use in developing the notions of orthogonality in infinite dimensions.

Theorem 6.2.4. *Let $\mathcal{H}$ be a Hilbert space and let $C \subseteq \mathcal{H}$ be a non-empty, closed, convex subset of $\mathcal{H}$. For each $x \in \mathcal{H}$ there exists a unique point $y \in C$ that is closest to x; that is*

$$\|x - z\| = \mathrm{dist}(x, C)$$

if and only if $z = y$.

Proof. To begin, let $d = \mathrm{dist}(x, C)$. We will first demonstrate there exists a point $y \in C$ such that $\|x - y\| = d$. By definition of the distance, for each $n \in \mathbb{N}$ there exists $y_n \in C$ such that

$$\|x - y_n\|^2 < d^2 + \frac{1}{n}.$$

We claim that $(y_n)_{n\geq 1}$ is Cauchy in $\mathcal{H}$. To see this, notice by the Parallelogram Law (Theorem B.1.18) we have for all $n, m \in \mathbb{N}$ that

$$\begin{aligned}
\|y_n - y_m\|^2 &= \|(x - y_m) - (x - y_n)\|^2 \\
&= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - \|(x - y_m) + (x - y_n)\|^2 \\
&= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\left\|x - \frac{y_m + y_n}{2}\right\|^2 \\
&\leq 2\left(d^2 + \frac{1}{n}\right) + 2\left(d^2 + \frac{1}{m}\right) - 4d^2 \\
&= \frac{2}{n} + \frac{2}{m}
\end{aligned}$$

(where the third to fourth line follows as $\frac{y_m + y_n}{2} \in C$ since C was convex). Hence we obtain that $(y_n)_{n\geq 1}$ is Cauchy in $\mathcal{H}$. Therefore $y = \lim_{n\to\infty} y_n$ exists as $\mathcal{H}$ is complete. Since C was closed in $\mathcal{H}$, we obtain that $y \in C$. Furthermore, as

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| \leq d,$$

we obtain that $\|x - y\| = d$.

To see that y is the unique vector with this property, suppose $z \in C$ is such that $\|x - z\| = d$. A similar computation to the one above show that

$$\begin{aligned}
\|y - z\|^2 &= 2\|x - y\|^2 + 2\|x - z\|^2 - 4\left\|x - \frac{y + z}{2}\right\|^2 \\
&\leq 2d^2 + 2d^2 - 4d^2 = 0.
\end{aligned}$$

Hence $y = z$ as desired. ∎

Of course it would be nice to be able to determine the vector y in Theorem 6.2.4. In general this is a difficult task for arbitrary closed convex subsets of Hilbert spaces. However, closed vector subspaces of a Hilbert space are an abundant collection of examples of closed convex sets for which we can solve this problem!

To begin, we must use the geometry of Hilbert spaces and the following.

Definition 6.2.5. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $S \subseteq V$. The *orthogonal complement* of S in V is the set

$$S^\perp = \{x \in V \mid \langle x, z\rangle = 0 \text{ for all } z \in S\}.$$

Example 6.2.6. The orthogonal complement of the x-axis in $\mathbb{R}^2$ with respect to the standard inner product is the y-axis. Similarly, the orthogonal complement of the y-axis in $\mathbb{R}^3$ with respect to the standard inner product is the yz-plane.

Remark 6.2.7. Clearly if $S \subseteq V$, then $S^\perp$ is a closed vector subspace of V. Furthermore $S^\perp = (\operatorname{span}(S))^\perp$ and $S^\perp = (\overline{S})^\perp$. Thus the notion of the orthogonal complement is really a notion for closed vector subspaces of inner product spaces.

Returning to Theorem 6.2.4, we can obtain a description of the closed vector using orthogonal complements.

Theorem 6.2.8. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Given $x \in \mathcal{H}$ and $y \in \mathcal{K}$, $\|x - y\| = \operatorname{dist}(x, \mathcal{K})$ and and only if $x - y \in \mathcal{K}^\perp$.*

Proof. First suppose $y \in \mathcal{K}$ is such that $\|x - y\| = \operatorname{dist}(x, \mathcal{K})$. To see that $x - y \in \mathcal{K}^\perp$, suppose to the contrary that there exists a $z \in \mathcal{K}$ such that $\alpha = \langle x - y, z\rangle \neq 0$. Note this implies $z \neq \vec{0}$. By scaling z if necessary (changing the value of α), we may assume that $\|z\| = 1$.

Consider the vector $v = y + \alpha z$ which is an element of $\mathcal{K}$ as $\mathcal{K}$ is a vector subspace. Then

$$\begin{aligned}
\|x - v\|^2 &= \langle x - y - \alpha z, x - y - \alpha z\rangle \\
&= \|x - y\|^2 - \alpha\langle z, x - y\rangle - \overline{\alpha}\langle x - y, z\rangle + |\alpha|^2 \|z\|^2 \\
&= \|x - y\|^2 - |\alpha|^2 \\
&< \operatorname{dist}(x, \mathcal{K})^2,
\end{aligned}$$

which is a contradiction as $v \in \mathcal{K}$. Hence it must be the case that $x - y \in \mathcal{K}^\perp$.

Conversely, suppose $x - y \in \mathcal{K}^\perp$. Clearly $\|x - y\| \geq \operatorname{dist}(x, \mathcal{K})$ whereas for all $z \in \mathcal{K}$,

$$\begin{aligned}
\|x - z\|^2 &= \|(x - y) - (z - y)\|^2 \\
&= \|x - y\|^2 + \|z - y\|^2 \geq \|x - y\|^2
\end{aligned}$$

by the Pythagorean Theorem since $z - y \in \mathcal{K}$ (as $\mathcal{K}$ is a vector subspace) and $x - y \in \mathcal{K}^\perp$. Hence $\|x - y\| = \operatorname{dist}(x, \mathcal{K})$. ∎

Using the above, given a Hilbert space $\mathcal{H}$ and a closed subspace $\mathcal{K}$, we can decompose $\mathcal{H}$ nicely.

Theorem 6.2.9. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Then $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$; that is, every element $x \in \mathcal{H}$ can be written uniquely as a sum of elements from $\mathcal{K}$ and $\mathcal{K}^\perp$. Moreover, for all $y \in \mathcal{K}$ and $z \in \mathcal{K}^\perp$, $\|y + z\| \leq \sqrt{\|y\|^2 + \|z\|^2}$.*

Proof. Let $x \in \mathcal{H}$. By Theorems 6.2.4 and 6.2.8, there exists a unique vector $y \in \mathcal{K}$ such that $z = x - y \in \mathcal{K}^\perp$. Hence as $x = y + z$, we obtain that $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$. Furthermore, the uniqueness follows from the uniqueness of y. The norm inequality then follows from the Pythagorean Theorem. ∎

In particular, as the orthogonal complement of any set is a closed subspace of a Hilbert space, we see that any closed subspace of a Hilbert space is topologically complemented in an orthogonal way. Using the following theory of topological complements, we obtain a nice bounded idempotent linear map related to this decomposition.

Proposition 6.2.10. *Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space and let $\mathcal{Y}$ and $\mathcal{Z}$ be closed subspace of $\mathcal{X}$. Then $\mathcal{Y}$ and $\mathcal{Z}$ are topological complements if and only if there exists a bounded linear map $E : \mathcal{X} \to \mathcal{X}$ such that $E^2 = E$, $\mathrm{Im}(E) = \mathcal{Y}$, and $\ker(E) = \mathcal{Z}$.*

Proof. Suppose $\mathcal{Y}$ and $\mathcal{Z}$ are topological complements. Thus for each $x \in \mathcal{X}$ there exist unique $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$ such that $x = y + z$. Consider the map $\Theta : \mathcal{Y} \oplus_1 \mathcal{Z} \to \mathcal{X}$ defined by

$$\Theta((y, z)) = y + z$$

for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$. Clearly Θ is a well-defined bijective linear map. Moreover, since

$$\|\Theta((y, z))\|_{\mathcal{X}} = \|y + z\|_{\mathcal{X}} \leq \|y\|_{\mathcal{X}} + \|z\|_{\mathcal{X}} = \|(y, z)\|_1,$$

we see that Θ is bounded. Hence the Inverse Mapping Theorem (Theorem 2.4.3) implies that Θ is an isomorphism.

Define $\pi : \mathcal{Y} \oplus_1 \mathcal{Z} \to \mathcal{Y}$ by $\pi((y, z)) = y$ for all $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$. Clearly π is a well-defined bounded linear map due to the norm structure on $\mathcal{Y} \oplus_1 \mathcal{Z}$. Thus if we define $E : \mathcal{X} \to \mathcal{Y} \subseteq \mathcal{X}$ by $E = \pi \circ \Theta^{-1}$, then E is a well-defined bounded linear map. Moreover, notice by construction that for each $x \in \mathcal{X}$ if $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$ are the unique elements such that $x = y + z$ then $E(x) = y$. Thus it readily follows that $E^2 = E$, $\mathrm{Im}(E) = \mathcal{Y}$, and $\ker(E) = \mathcal{Z}$.

Conversely, suppose $E : \mathcal{X} \to \mathcal{X}$ is a bounded linear map such that $E^2 = E$, $\mathrm{Im}(E) = \mathcal{Y}$, and $\ker(E) = \mathcal{Z}$. Notice if $y \in \mathrm{Im}(E)$, then $y = E(x)$ for some $x \in \mathcal{X}$ and thus

$$E(y) = E^2(x) = E(x) = y.$$

Thus, as $\mathcal{Z} = \ker(E)$, we easily see that $\mathcal{Y} \cap \mathcal{Z} = \emptyset$. To see that $\mathcal{X} = \mathcal{Y} + \mathcal{Z}$, let $x \in \mathcal{X}$ be arbitrary. If $y = E(x)$ and $z = x - E(x)$, then $y \in \operatorname{Im}(E) = \mathcal{Y}$, $x = y + z$ ad

$$E(z) = E(x) - E(E(x)) = E(x) - E(x) = 0$$

so $z \in \ker(E) = \mathcal{Z}$. Hence $\mathcal{Y}$ and $\mathcal{Z}$ are topologically complemented in $\mathcal{X}$. ∎

Corollary 6.2.11. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed vector subspace of $\mathcal{H}$. There is a unique linear map $P : \mathcal{H} \to \mathcal{K} \subseteq \mathcal{H}$ such that $P(x) = x$ for all $x \in \mathcal{K}$ and $P(y) = \vec{0}$ for all $y \in \mathcal{K}^\perp$. The linear map P is called the* orthogonal projection of $\mathcal{H}$ onto $\mathcal{K}$. *Furthermore, P is bounded with $\|P\| \leq 1$ (with equality whenever $\mathcal{K} \neq \{\vec{0}\}$), $P^2 = P$, and $\|x - P(x)\| = \operatorname{dist}(x, \mathcal{K})$ for all $x \in \mathcal{H}$.*

Proof. By Theorem 6.2.9 and Proposition 6.2.10 there exists a bounded linear map $P : \mathcal{H} \to \mathcal{H}$ such that $P^2 = P$, $\operatorname{Im}(P) = \mathcal{K}$, and $\ker(P) = \mathcal{K}^\perp$. Since for each $x \in \mathcal{H}$ we may write $x = y + z$ with $y \in \mathcal{K}$ (such that $\|x - y\| = \operatorname{dist}(x, \mathcal{K})$) and $z \in \mathcal{K}^\perp$, and since $P(x) = y$, we see that $\|x - P(x)\| = \operatorname{dist}(x, \mathcal{K})$ and

$$\|P(x)\|^2 = \|y\|^2 \leq \|y\|^2 + \|z\|^2 = \|x\|^2$$

for all $x \in \mathcal{H}$ so $\|P\| \leq 1$. ∎

It is elementary to verify that if P is the orthogonal projection onto a subspace $\mathcal{K}$ and $I_\mathcal{H} : \mathcal{H} \to \mathcal{H}$ is the identity map, then $I_\mathcal{H} - P$ is the orthogonal projection on $\mathcal{K}^\perp$.

We will see how useful orthogonal projections are in the following section. For now, we can use the concept of a direct sum in Hilbert spaces to prove the following.

Corollary 6.2.12. *Let $\mathcal{H}$ be a Hilbert space and let $S \subseteq \mathcal{H}$ be non-empty. Then $(S^\perp)^\perp = \overline{\operatorname{span}(S)}$.*

Proof. To begin, let $x \in \overline{\operatorname{span}(S)}$ be arbitrary. Thus there exists a sequence $(x_n)_{n\geq 1}$ of elements of $\operatorname{span}(S)$ such that $x = \lim_{n\to\infty} x_n$. Let $y \in S^\perp$ be arbitrary. Then

$$\langle x, y\rangle = \lim_{n\to\infty} \langle x_n, y\rangle = 0$$

as $y \in S^\perp$ and $x_n \in \operatorname{span}(S)$ for all n. Therefore, as $y \in S^\perp$ was arbitrary, $x \in (S^\perp)^\perp$. Thus, as $x \in \overline{\operatorname{span}(S)}$ was arbitrary, $\overline{\operatorname{span}(S)} \subseteq (S^\perp)^\perp$.

For the other inclusion, let $x \in (S^\perp)^\perp$ be arbitrary. Since $\overline{\operatorname{span}(S)}$ is a closed vector subspace, Theorem 6.2.8 implies there exists a vector $y \in \overline{\operatorname{span}(S)}$ such that $x - y \in \overline{\operatorname{span}(S)}^\perp$. Notice for all $z \in \overline{\operatorname{span}(S)}$ that

$$\langle x - y, z\rangle = 0$$

as $x - y \in \overline{\operatorname{span}(S)}^\perp$. Similarly, if $z \in \overline{\operatorname{span}(S)}^\perp$ then $z \in S^\perp$ so

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

Therefore, as every vector in $\mathcal{H}$ can be written as the sum of elements from $\overline{\operatorname{span}(S)}$ and $\overline{\operatorname{span}(S)}^\perp$, we obtain that $\langle x - y, z\rangle = 0$ for all $z \in \mathcal{H}$. Hence by choosing $z = x - y$, we obtain that $x = y \in \overline{\operatorname{span}(S)}$. Therefore, as $x \in (S^\perp)^\perp$ was arbitrary, we obtain that $(S^\perp)^\perp \subseteq \overline{\operatorname{span}(S)}$ as desired. ∎

6.3 Isomorphisms of Hilbert Spaces

Using the theory of orthogonal projections, we can develop a notion of bases for Hilbert spaces that is far superior to taking a vector space basis. In particular, recall from Theorem 2.3.14 that any vector space basis for an infinite dimensional Banach space must be uncountable. Thus we desire 'nice' bases for Hilbert spaces that to avoid this problem and use the geometry of Hilbert spaces. These nice bases will also produce a way for us to identify all of the possible Hilbert spaces! Thus we begin with the following.

Definition 6.3.1. Let $(\mathcal{X}, \|\cdot\|)$ be a normed linear space. An element $x \in \mathcal{X}$ is said to be a *unit vector* if $\|x\| = 1$.

Definition 6.3.2. Let $\mathcal{H}$ be a Hilbert space. A subset $\{e_\alpha\}_{\alpha\in\Lambda}$ is said to be an *orthonormal set* if each e_α is a unit vector and $\langle e_\alpha, e_\beta\rangle = 0$ $\alpha, \beta \in \Lambda$ with $\alpha \neq \beta$ (i.e. an orthogonal set of unit vectors).

Remark 6.3.3. It is not difficult to see that every orthonormal set of vectors is automatically linearly independent. Indeed suppose $\{e_\alpha\}_{\alpha\in\Lambda}$ is orthonormal and there exists $\alpha_1, \ldots, \alpha_n \in \Lambda$ and $a_1, \ldots, a_n \in \mathbb{K}$ are such that

$$\sum_{k=1}^n a_k e_{\alpha_k} = \vec{0}.$$

For each $j \in \{1, \ldots, n\}$, taking the inner product with e_{α_j} produces

$$0 = \langle \vec{0}, e_{\alpha_j}\rangle = \sum_{k=1}^n a_k \langle e_{\alpha_k}, e_{\alpha_j}\rangle = a_j.$$

Hence $a_j = 0$ for all $j \in \{1, \ldots, n\}$ so $\{e_\alpha\}_{\alpha\in\Lambda}$ is linearly independent.

We desire to construct special orthonormal sets. Unfortunately, unlike with finite dimensional theory that students may have seen previously, the notion of spanning orthonormal sets is not the correct notion for infinite dimensional Hilbert spaces.

For the correct notion, given a Hilbert space $\mathcal{H}$ let $\mathcal{E}_\mathcal{H}$ denote the set of all orthonormal subsets of $\mathcal{H}$. Notice we may place a partial ordering on $\mathcal{E}_\mathcal{H}$

via inclusion. Since the union of any chain of orthonormal sets under this ordering is an upper bound for the chain (and as $\mathcal{E}_{\mathcal{H}} \neq \emptyset$), Zorn's Lemma implies there is a maximal element of $\mathcal{E}_{\mathcal{H}}$ under inclusion. These are the objects we are after.

Definition 6.3.4. Let $\mathcal{H}$ be a Hilbert space. An *orthonormal basis* of $\mathcal{H}$ is a maximal orthonormal set.

Example 6.3.5. For $n \in \mathbb{N}$ consider the vectors $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{K}^n$ where for each $j \in \{1, \ldots, n\}$

$$\vec{e}_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$$

where the unique 1 occurs in the j^{th} spot. Clearly $\mathcal{E} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ is orthonormal with respect to the standard inner product. Suppose that $\mathcal{E}$ were not a maximal orthonormal set. Then there would exist a vector $x = (x_1, \ldots, x_n) \in \mathcal{E}^\perp$ with $\|x\| = 1$. The fact that $x \in \mathcal{E}^\perp$ implies

$$0 = \langle x, \vec{e}_j \rangle = x_j$$

for all $j \in \{1, \ldots, n\}$. Thus $x = \vec{0}$, an obvious contradiction. Hence $\mathcal{E}$ is an orthonormal basis for $\mathbb{K}^n$.

Example 6.3.6. For each $n \in \mathbb{N}$ let $\vec{e}_n \in \ell_2(\mathbb{N})$ be the sequence with a 1 in the n^{th} entries and 0s everywhere else. By the same arguments as Example 6.3.5, $\mathcal{E} = \{\vec{e}_n\}_{n=1}^\infty$ is an orthonormal basis for $\ell_2(\mathbb{N})$. However, it is elementary to see that $\mathcal{E}$ does not span $\ell_2(\mathbb{N})$ (indeed the sequence $(\frac{1}{n})_{n\geq 1} \in \ell_2(\mathbb{N})$ is not a finite linear combination of elements of $\mathcal{E}$).

Remark 6.3.7. Using the argument preceding Definition 6.3.4, it is easy to see if $\mathcal{F}$ is an orthonormal subset of a Hilbert space $\mathcal{H}$ then there exists an orthonormal basis $\mathcal{E}$ for $\mathcal{H}$ containing $\mathcal{F}$ (i.e. restrict the Zorn's Lemma argument to orthonormal sets containing $\mathcal{F}$).

In the finite dimensional world, we recall we even have an algorithm for constructing orthonormal bases.

Theorem 6.3.8 (Gram-Schmidt Orthogonalization Process). *Let V be an inner product space and let $L = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a linearly independent subset of V. Then there exists an orthonormal set $O = \{\vec{e}_1, \ldots, \vec{e}_n\}$ such that* $\mathrm{span}(L) = \mathrm{span}(O)$.

Proof. As $\vec{v}_1 \neq \vec{0}$ as L is linearly independent, let $\vec{e}_1 = \frac{1}{\|v_1\|} v_1$. Then

$$\|\vec{e}_1\| = \left\| \frac{1}{\|\vec{v}_1\|} \vec{v}_1 \right\| = \frac{1}{\|\vec{v}_1\|} \|\vec{v}_1\| = 1.$$

Suppose for some $k \in \{1, \ldots, n-1\}$ we have constructed $\vec{e}_1, \ldots, \vec{e}_k$ such that $\{\vec{e}_1, \ldots, \vec{e}_k\}$ is orthonormal and $\{\vec{e}_1, \ldots, \vec{e}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is linearly independent with the same span as L. Let

$$\vec{x}_{k+1} = \vec{v}_{k+1} - \sum_{j=1}^{k} \langle \vec{v}_{k+1}, \vec{e}_j \rangle \vec{e}_j.$$

Since $\{\vec{e}_1, \ldots, \vec{e}_k\}$ is orthonormal, it is easy to see that $\vec{x}_{k+1}$ is orthogonal to $\{\vec{e}_1, \ldots, \vec{e}_k\}$. Furthermore, as $\{\vec{e}_1, \ldots, \vec{e}_k, \vec{v}_{k+1}\}$ is linearly independent, we see that $\vec{x}_k$ is non-zero and $\{\vec{e}_1, \ldots, \vec{e}_k, \vec{x}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_n\}$ is linearly independent. If we define $\vec{e}_{k+1} = \frac{1}{\|\vec{x}_{k+1}\|}\vec{x}_{k+1}$, we easily obtain that $\{\vec{e}_1, \ldots, \vec{e}_{k+1}\}$ is orthonormal and $\{\vec{e}_1, \ldots, \vec{e}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_n\}$ is linearly independent with the same span as L. The proof is then complete by recursion. ∎

Remark 6.3.9. The proof of the Gram-Schmidt Orthogonalization Process actually makes use of a formula for the orthogonal projection onto a finite subspace. Notice that if $\mathcal{K}$ is a finite dimensional vector subspace of a Hilbert space $\mathcal{H}$, $\mathcal{K}$ is closed by Corollary 3.5.5 and the Gram-Schmidt Orthogonalization Process implies $\mathcal{K}$ has a orthonormal basis which is a vector space basis, say $\{\vec{e}_1, \ldots, \vec{e}_n\}$. If P is the orthogonal projection onto $\mathcal{K}$, we claim that

$$P(x) = \sum_{k=1}^{n} \langle x, \vec{e}_k \rangle \vec{e}_k$$

for all $x \in \mathcal{H}$. Indeed if y denotes the right-hand side of the above expression, clearly $x - y$ is orthogonal to each $\vec{e}_k$ and thus $x - y \in \mathcal{K}^\perp$. As $P(x)$ is the unique vector such that $x - P(x) \in \mathcal{K}^\perp$ by Theorems 6.2.4 and 6.2.8, and by Corollary 6.2.11, we obtain that $y = P(x)$.

Although orthonormal bases for finite dimensional vector subspaces are useful for the above projection formula, as orthonormal bases need not be vector space bases in infinite dimensional Hilbert spaces, we must ask, "How close are orthonormal bases to actual vector spaces bases?" We will see that orthonormal bases are 'bases with respect to analytic conditions'. To begin, we first note the following result for countable orthonormal bases.

Theorem 6.3.10 (Bessel's Inequality, Countable). *Let $\mathcal{H}$ be a Hilbert space and let $\{e_\alpha\}_{\alpha \in \Lambda}$ be an orthonormal set with Λ countable. For each $x \in \mathcal{H}$*

$$\sum_{\alpha \in \Lambda} |\langle x, e_\alpha \rangle|^2 \leq \|x\|^2.$$

Proof. Without loss of generality $\Lambda = \mathbb{N}$ (the proof of the result for finite Λ is contained within). For each $n \in \mathbb{N}$, let $\mathcal{K}_n = \operatorname{span}(\{e_1, \ldots, e_n\})$. Then, if

P_n is the orthogonal projection onto $\mathcal{K}_n$, we obtain for all $x \in \mathcal{H}$ that

$$\begin{aligned}\|x\|^2 &\geq \|P(x)\|^2 \\ &= \left\|\sum_{k=1}^{n} \langle x, e_k\rangle e_k\right\|^2 \\ &= \sum_{k=1}^{n} |\langle x, e_\alpha\rangle|^2\end{aligned}$$

by the Pythagorean Theorem (Theorem B.1.17), Corollary 6.2.11 and Remark 6.3.9. Hence the result follows by taking the limit as n tends to infinity. ∎

Using Bessel's Inequality for countable orthonormal sets, we obtain the following important result in the case of uncountable orthonormal bases.

Lemma 6.3.11. *Let $\mathcal{H}$ be a Hilbert space and let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal set. For each $x \in \mathcal{H}$ the set $\{\alpha \in \Lambda \mid \langle x, e_\alpha\rangle \neq 0\}$ is countable.*

Proof. For each $n \in \mathbb{N}$ let

$$\mathcal{E}_n = \left\{\alpha \in \Lambda \,\middle|\, |\langle x, e_\alpha\rangle| > \frac{1}{n}\right\}.$$

We claim that each $\mathcal{E}_n$ is finite. Indeed suppose to the contrary that $\mathcal{E}_n$ is infinite. Hence there exists a collection $\{\alpha_m\}_{m\in\mathbb{N}} \subseteq \mathcal{E}_n$ such that $\alpha_m \neq \alpha_k$ whenever $k \neq m$. Therefore we obtain by Theorem 6.3.10 that

$$\|x\|^2 \geq \sum_{m\in\mathbb{N}} |\langle x, e_{\alpha_m}\rangle|^2 \geq \sum_{m\in\mathbb{N}} \frac{1}{n^2},$$

which is impossible. Hence each $\mathcal{E}_n$ must be finite.

Since

$$\{\alpha \in \Lambda \mid \langle x, e_\alpha\rangle \neq 0\} = \bigcup_{n\in\mathbb{N}} \mathcal{E}_n,$$

the set under consideration is a countable union of countable sets and thus is countable. ∎

Using the above, we immediately obtain a version of Bessel's Inequality for uncountable sets. In that which follows, we will be summing non-negative real numbers over an uncountable set via nets. However, as only countable many terms in the sum are non-zero, this sum can be thought of as a countable sum of positive real numbers.

Theorem 6.3.12 (Bessel's Inequality). *Let $\mathcal{H}$ be a Hilbert space and let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal set. For each $x \in \mathcal{H}$*

$$\sum_{\alpha\in\Lambda} |\langle x, e_\alpha\rangle|^2 \leq \|x\|^2 .$$

Corollary 6.3.13. *Let $\mathcal{H}$ be a Hilbert space and let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal set. For each $x \in \mathcal{H}$ the sum*

$$\sum_{\alpha\in\Lambda} \langle x, e_\alpha\rangle e_\alpha$$

converges.

Proof. By Lemma 6.3.11, only a countable number of coefficients are non-zero. By Theorem 6.3.12, the sum above (by which we mean sum the countable number of terms with non-zero coefficients), we obtain that the sum is absolutely summable. Hence the sum converges by Theorem 2.2.2 as $\mathcal{H}$ is complete. ∎

Finally, we obtain the characterization of an orthonormal basis that shows orthonormal bases are good analytical bases for Hilbert spaces.

Theorem 6.3.14. *Let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. The following are equivalent:*

(1) $\{e_\alpha\}_{\alpha\in\Lambda}$ *is an orthonormal basis for $\mathcal{H}$.*

(2) $\mathrm{span}(\{e_\alpha\}_{\alpha\in\Lambda})$ *is dense in $\mathcal{H}$.*

(3) For all $x \in \mathcal{H}$, $x = \sum_{\alpha\in\Lambda}\langle x, e_\alpha\rangle e_\alpha$.

(4) For all $x \in \mathcal{H}$, $\|x\|^2 = \sum_{\alpha\in\Lambda} |\langle x, e_\alpha\rangle|^2$.

Proof. To see that (1) implies (2), suppose $\{e_\alpha\}_{\alpha\in\Lambda}$ is an orthonormal basis for $\mathcal{H}$. If $\mathrm{span}(\{e_\alpha\}_{\alpha\in\Lambda})$ is not dense in $\mathcal{H}$, then $\mathcal{K} = \overline{\mathrm{span}(\{e_\alpha\}_{\alpha\in\Lambda})}$ is a closed vector subspace of $\mathcal{H}$ that is not equal to $\mathcal{H}$. Hence $\mathcal{K}^\perp \neq \emptyset$ by Theorem 6.2.9 so $\mathcal{K}^\perp$ must contain a vector x of length 1. Since x is orthogonal to each element of $\mathcal{K}$ and thus each e_α, we obtain that $\{x\} \cup \{e_\alpha\}_{\alpha\in\Lambda}$ is an orthonormal set which is larger than $\{e_\alpha\}_{\alpha\in\Lambda}$. As this contradicts the fact that $\{e_\alpha\}_{\alpha\in\Lambda}$ is a maximal orthonormal set, we have obtained a contradiction. Hence (1) implies (2).

To see that (2) implies (3), let $x \in \mathcal{H}$ be arbitrary. By Corollary 6.3.13 the vector $y = \sum_{\alpha\in\Lambda}\langle x, e_\alpha\rangle e_\alpha$ is an element of $\mathcal{H}$. Hence there exists an increasing sequence of finite subsets Λ_n of Λ such that

$$y = \lim_{n\to\infty} \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle e_\alpha.$$

Therefore, by the continuity of the inner product, we obtain that

$$\begin{aligned}
\langle x - y, e_\alpha\rangle &= \lim_{n\to\infty} \left\langle x - \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle e_\alpha, e_\beta \right\rangle \\
&= \lim_{n\to\infty} \langle x, e_\beta\rangle - \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle \langle e_\alpha, e_\beta\rangle \\
&= 0
\end{aligned}$$

for all $\beta \in \Lambda$. Hence $x - y \in (\mathrm{span}(\{e_\alpha\}_{\alpha\in\Lambda}))^\perp = \mathcal{H}^\perp = \{\vec{0}\}$. Thus $x = y$ as desired. Therefore, as $x \in \mathcal{H}$ was arbitrary, (2) implies (3).

To see that (3) implies (4), let $x \in \mathcal{H}$ be arbitrary. Notice there exists an increasing sequence of finite subsets Λ_n of Λ such that

$$x = \lim_{n\to\infty} \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle e_\alpha \qquad \text{and} \qquad \sum_{\alpha\in\Lambda} |\langle x, e_\alpha\rangle|^2.$$

Thus, by the continuity of the inner product

$$\begin{aligned} \|x\|^2 &= \lim_{n\to\infty} \left\langle \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle e_\alpha, \sum_{\alpha\in\Lambda_n} \langle x, e_\alpha\rangle e_\alpha \right\rangle \\ &= \lim_{n\to\infty} \sum_{\alpha\in\Lambda_n} |\langle x, e_\alpha\rangle|^2 \\ &= \sum_{\alpha\in\Lambda} |\langle x, e_\alpha\rangle|^2. \end{aligned}$$

Hence (3) implies (4).

Finally, to see that (4) implies (1), suppose to the contrary that $\{e_\alpha\}_{\alpha\in\Lambda}$ was not an orthonormal basis. Thus there exists a vector $x \in \mathcal{H}$ such that $\|x\|^2 = 1$ yet x is orthogonal to each e_α. However, the formula in (4) then implies $1 = 0$ which is impossible. Hence $\{e_\alpha\}_{\alpha\in\Lambda}$ is an orthonormal basis. ∎

Using the same arguments as in Remark 6.3.9, we obtain a version of the orthogonal projection formula for infinite dimensional subspaces.

Corollary 6.3.15. *Let $\mathcal{K}$ be a closed vector subspace of a Hilbert space $\mathcal{H}$. If $\{e_\alpha\}_{\alpha\in\Lambda}$ is an orthonormal basis for $\mathcal{K}$ and P is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{K}$, then for all $x \in \mathcal{H}$*

$$P(x) = \sum_{\alpha\in\Lambda} \langle x, e_\alpha\rangle e_\alpha.$$

Using orthonormal bases, we can completely characterize all Hilbert spaces in existence. The following is our first step.

Proposition 6.3.16. *If $\mathcal{H}$ is a Hilbert space then any two orthonormal basis for $\mathcal{H}$ have the same cardinality.*

Proof. If $\mathcal{H}$ has a finite orthonormal basis, then $\mathcal{H}$ is finite dimensional. Since each orthonormal basis for a finite dimensional Hilbert space is a vector space basis, the result trivial follows. Hence we will assume $\mathcal{H}$ has only infinite dimensional orthonormal bases.

Let $\{e_\alpha\}_{\alpha\in\mathcal{E}}$ and $\{f_\beta\}_{\beta\in\mathcal{F}}$ be orthonormal bases for $\mathcal{H}$. Recall for each $\alpha \in \mathcal{E}$ the set

$$\mathcal{F}_\alpha = \{\beta \in \mathcal{F} \mid \langle e_\alpha, f_\beta\rangle \neq 0\}$$

is countable by Lemma 6.3.11. Furthermore, by Theorem 6.3.14 applied to $\{e_\alpha\}_{\alpha\in\mathcal{E}}$, for each $\beta \in \mathcal{F}$ there exists an $\alpha \in \mathcal{E}$ such that $\beta \in F_\alpha$. Therefore $\mathcal{F} = \bigcup_{\alpha\in\mathcal{E}} \mathcal{F}_\alpha$ so, as $|\mathcal{F}_\alpha| \leq |\mathbb{N}|$,

$$|\mathcal{F}| \leq |\mathbb{N}||\mathcal{E}| = |\mathcal{E}|$$

by cardinality theory. By replacing the roles of $\mathcal{F}$ and $\mathcal{E}$, we obtain that $|\mathcal{E}| \leq |\mathcal{F}|$ so $|\mathcal{E}| = |\mathcal{F}|$ as desired. ∎

Because of Proposition 6.3.16, we can now make the following definition.

Definition 6.3.17. The *dimension* of a Hilbert space $\mathcal{H}$, denoted $\dim(\mathcal{H})$, is the cardinality of an orthonormal basis for $\mathcal{H}$.

Unsurprisingly, two Hilbert spaces will be "the same" if they have the same dimension. To formalize this, we must discuss the appropriate notion of isomorphism for Hilbert spaces.

Definition 6.3.18. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. A *unitary operator* from $\mathcal{H}_1$ to $\mathcal{H}_2$ is a surjective linear map $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$\langle U(x), U(y)\rangle_{\mathcal{H}_2} = \langle x, y\rangle_{\mathcal{H}_1}$$

for all $x, y \in \mathcal{H}_1$.

Note as $\|U(x)\|_{\mathcal{H}_2} = \|x\|_{\mathcal{H}_1}$ for all $x \in \mathcal{H}_1$, unitary operators are injective (and thus bijective). As clearly U^{-1} will also be a unitary operator, the following defines an equivalence relation on the class of all Hilbert spaces.

Definition 6.3.19. Two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are said to be *isomorphic* if there exists a unitary operator from $\mathcal{H}_1$ to $\mathcal{H}_2$.

Finally we can now characterize Hilbert spaces via their dimensions.

Theorem 6.3.20. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Then $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic if and only if* $\dim(\mathcal{H}_1) = \dim(\mathcal{H}_2)$.

Proof. First suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic. Therefore there exists a unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$. Let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal basis for $\mathcal{H}_1$. By the definition of a unitary, we see that $\{U(e_\alpha)\}_{\alpha\in\Lambda}$ is an orthonormal set. Furthermore, Theorem 6.3.14 implies that $\mathrm{span}(\{e_\alpha\}_{\alpha\in\Lambda})$ is dense in $\mathcal{H}_1$ so, since U is a linear map and a homeomorphism, $\mathrm{span}(\{U(e_\alpha)\}_{\alpha\in\Lambda})$ must also be dense in $\mathcal{H}_2$. Hence $\{U(e_\alpha)\}_{\alpha\in\Lambda}$ is an orthonormal basis of $\mathcal{H}_2$ by Theorem 6.3.14. Hence $\dim(\mathcal{H}_1) = |\Lambda| = \dim(\mathcal{H}_2)$ as desired.

For the converse direction, we note that since isomorphism of Hilbert spaces is an equivalence relation that it suffices to prove the following.

Corollary 6.3.21. *Let $\mathcal{H}$ be a Hilbert space and let Λ be a set such that $|\Lambda| = \dim(\mathcal{H})$. Then $\mathcal{H}$ is isomorphic to the Hilbert space*

$$\ell_2(\Lambda, \mathbb{K}) = \left\{ f : \Lambda \to \mathbb{K} \,\middle|\, {\scriptstyle \{\alpha\in\Lambda \,|\, f(\alpha)\neq 0\} \text{ is countable} \atop \scriptstyle \text{and } \sum_{\alpha\in\Lambda} |f(\alpha)|^2<\infty} \right\}$$

equipped with the inner product

$$\langle f, g\rangle_{\ell_2(\Lambda,\mathbb{K})} = \sum_{\alpha\in\Lambda} f(\alpha)\overline{g(\alpha)}.$$

Proof. First we must proof that $\ell_2(\Lambda, \mathbb{K})$ together with the inner product described is indeed a Hilbert space. The proof that $\langle f, g\rangle_{\ell_2(\Lambda,\mathbb{K})}$ is a well-defined inner product is as in Example B.1.10. The proof that $\ell_2(\Lambda, \mathbb{K})$ is a Banach space follows from the Riesz-Fisher Theorem (Theorem D.2.1) by taking the counting measure μ on Λ and equating $L_2(\Lambda, \mu)$ with $\ell_2(\Lambda, \mathbb{K})$. Hence $\ell_2(\Lambda, \mathbb{K})$ is a Hilbert space.

To complete the proof, let $\{e_\alpha\}_{\alpha\in\Lambda}$ be an orthonormal basis of $\mathcal{H}$. Define $U : \mathcal{H} \to \ell_2(\Lambda, \mathbb{K})$ by

$$U(h)(\alpha) = \langle h, e_\alpha\rangle_{\mathcal{H}}$$

for all $\alpha \in \Lambda$ and $h \in \mathcal{H}$. Note if $h \in \mathcal{H}$ then $U(h)$ is indeed an element of $\ell_2(\Lambda, \mathbb{K})$ by Bessel's inequality (Theorem 6.3.12). Hence U is a well-defined linear map that maps the orthonormal basis $\{e_\alpha\}_{\alpha\in\Lambda}$ to the orthonormal basis $\{f_\alpha\}_{\alpha\in\Lambda}$ where

$$f_\alpha(\beta) = \begin{cases} 1 & \beta = \alpha \\ 0 & \beta \neq \alpha \end{cases}.$$

Hence Theorem 6.3.14 implies that U is surjective. To see that U is a unitary (and thus injective), notice for all $x, y \in \mathcal{H}$ that by Theorem 6.3.14 and the fact that the inner product is continuous in each entry, we have

$$\begin{aligned}
\langle U(x), U(y)\rangle_{\ell_2(\Lambda,\mathbb{K})} &= \sum_{\alpha\in\Lambda} \langle x, e_\alpha\rangle_{\mathcal{H}} \overline{\langle e_\alpha, y\rangle_{\mathcal{H}}} \\
&= \sum_{\alpha\in\Lambda} \langle \langle x, e_\alpha\rangle_{\mathcal{H}} e_\alpha, \langle e_\alpha, y\rangle_{\mathcal{H}} e_\alpha\rangle_{\mathcal{H}} \\
&= \sum_{\alpha,\beta\in\Lambda} \langle \langle x, e_\alpha\rangle_{\mathcal{H}} e_\alpha, \langle e_\beta, y\rangle_{\mathcal{H}} e_\beta\rangle_{\mathcal{H}} \\
&= \left\langle \sum_{\alpha\in\Lambda} \langle x, e_\alpha\rangle_{\mathcal{H}} e_\alpha, \sum_{\beta\in\Lambda} \langle e_\beta, y\rangle_{\mathcal{H}} e_\beta \right\rangle_{\mathcal{H}} \\
&= \langle x, y\rangle_{\mathcal{H}}
\end{aligned}$$

Hence U is a unitary so $\mathcal{H}$ is isomorphic to $\ell_2(\Lambda, \mathbb{K})$. ∎

This completes the proof of Theorem 6.3.20. ∎

6.4 Bounded Linear Operators on Hilbert Spaces

With the above theory of Hilbert spaces complete, we turn our attention to bounded linear functions on Hilbert spaces. In particular, we will be able to quickly characterize the dual space of any Hilbert space. This leads us to a particular behaviour of the adjoints and a theory of the bounded linear operators on a Hilbert space. We begin with the following.

Theorem 6.4.1 (Riesz Representation Theorem). *Let $\mathcal{H}$ be a Hilbert space. For each $y \in \mathcal{H}$ define $\varphi_y : \mathcal{H} \to \mathbb{K}$ by*

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

for all $x \in \mathcal{H}$. Then $\varphi_y \in \mathcal{H}^*$ for all $y \in \mathcal{H}$.

Moreover, if we define $\Phi : \mathcal{H} \to \mathcal{H}^*$ by

$$\Phi(y) = \varphi_y$$

for all $y \in \mathcal{H}$, then Φ is a conjugate linear, isometric, and bijective.

Proof. To begin, it is elementary to see that φ_y is a linear map for all $y \in \mathcal{H}$. To see that φ is continuous, note by the Cauchy-Schwarz inequality that

$$|\varphi_y(x)| \leq \|x\| \, \|y\|$$

for all $x \in \mathcal{H}$. Hence φ is continuous and $\|\varphi\| \leq \|y\|$.

To see that $\|\varphi_y\| = \|y\|$ thereby showing Φ is isometric, notice said equality is trivial if $y = \vec{0}$. Otherwise let $z = \frac{1}{\|y\|}y$ so that z is a unit vector. Since

$$\varphi_y(z) = \left\langle \frac{1}{\|y\|}y, y\right\rangle = \|y\|,$$

the other equality follows.

It is elementary to verify that Φ is conjugate linear. Moreover, as Φ is isometric, Φ is injective. To see that Φ is surjective, let $\varphi \in \mathcal{H}^*$ be arbitrary. If $\varphi(x) = 0$ for all $x \in \mathcal{H}$, then clearly $\varphi = \varphi_{\vec{0}}$. Otherwise, suppose φ is not the zero linear functional. Hence $\ker(\varphi)$ is a closed vector subspace of $\mathcal{H}$ that does not equal $\mathcal{H}$. Thus there exists a vector $z \in \ker(\varphi)^\perp \setminus \{\vec{0}\}$. As $\varphi(z) \neq 0$, by scaling z if necessary, we may assume that $\varphi(z) = 1$.

We claim that $\mathrm{span}(\{z\}) = \ker(\varphi)^\perp$. To see this, it suffices to show that if $z_1 \in \ker(\varphi)^\perp \setminus \{\vec{0}\}$ and $\varphi(z_1) = 1$, then $z = z_1$. Indeed if z_1 has the desired properties, then $z - z_1 \in \ker(\varphi)^\perp$ and

$$\varphi(z - z_1) = 1 - 1 = 0$$

so $z - z_1 \in \ker(\varphi)$. Hence $z - z_1 \in \ker(\varphi) \cap \ker(\varphi)^\perp = \{\vec{0}\}$ so $z = z_1$ as desired. Hence $\mathrm{span}(\{z\}) = \ker(\varphi)^\perp$ and thus $\{z\}^\perp = \ker(\varphi)$ by Corollary 6.2.12.

As $z \neq \vec{0}$, let $y = \frac{1}{\|z\|^2} z$. Therefore $\{y\}^\perp = \{z\}^\perp = \ker(\varphi)$. We claim that $\varphi = \varphi_y$. To see this, we notice for all $x \in \ker(\varphi)$ that $x \in \{y\}^\perp$ so

$$\langle x, y \rangle = 0 = \varphi(x).$$

Otherwise, if $x = \beta y$ for some $\beta \in \mathbb{K}$, we see that

$$\begin{aligned}
\varphi(x) &= \beta\varphi(y) \\
&= \beta\frac{1}{\|z\|^2}\varphi(z) \\
&= \frac{\beta}{\|z\|^2} \\
&= \frac{\beta}{\|z\|^4}\langle z, z \rangle \\
&= \beta\langle y, y \rangle \\
&= \langle x, y \rangle.
\end{aligned}$$

Therefore, as $\mathcal{H} = \ker(\varphi) \oplus \ker(\varphi)^\perp = \ker(\varphi) \oplus \operatorname{span}(\{y\})$ by Theorem 6.2.9, it follows that $\varphi = \varphi_y$ as desired. ∎

Clearly the Riesz Representation Theorem (Theorem 6.4.1) shows that Hilbert spaces are reflexive and thus the weak and weak*-topologies on a Hilbert space are the same. In fact, the whole point of examining reflexive Banach spaces is that they behave more similarly to Hilbert space than other Banach spaces.

Of course, a characterization of the dual space immediately gives us some results based on our knowledge of linear functionals

Corollary 6.4.2. *Let $\mathcal{H}$ be a Hilbert space. If $x \in \mathcal{H}$ then*

$$\|x\| = \sup\{|\langle x, y \rangle| \mid y \in \mathcal{H}, \|y\| \leq 1\}.$$

Proof. This follows immediately from Corollary 4.3.1. ∎

Corollary 6.4.3. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T : \mathcal{H} \to \mathcal{K}$ be linear. Then*

$$\|T\| = \sup\{|\langle T(x), y \rangle_{\mathcal{K}}| \mid x \in \mathcal{H}, y \in \mathcal{K}, \|x\|_{\mathcal{H}}, \|y\|_{\mathcal{K}} \leq 1\}$$

(with both sides being infinity if T is not bounded).

Moreover, if $\mathcal{H}$ and $\mathcal{K}$ are Hilbert space and $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, we know the Banach space adjoint T^* of T is an element of $\mathcal{B}(\mathcal{K}^*, \mathcal{H}^*)$, which can be identified as an element of $\mathcal{B}(\mathcal{K}, \mathcal{H})$ via the Riesz Representation Theorem (Theorem 6.4.1). Due to the identification of dual spaces, the Banach space adjoint of an operator between two Hilbert spaces can be alternatively defined as follows and is known as *(Hilbert space) adjoint of T*. We include the following construction as an alternative method.

Theorem 6.4.4. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. Then there exists a unique linear map $T^* \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ such that*

$$\langle T^*(y), x\rangle_{\mathcal{H}} = \langle y, T(x)\rangle_{\mathcal{K}}$$

for all $x \in \mathcal{H}$ and $y \in \mathcal{K}$. Furthermore $\|T^*\| = \|T\|$.

Proof. Fix $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. For each $y \in \mathcal{K}$, consider the linear map $f_y : \mathcal{H} \to \mathbb{C}$ defined by

$$f_y(x) = \langle T(x), y\rangle_{\mathcal{K}}$$

for all $x \in \mathcal{H}$. Since

$$|f_y(x)| = |\langle T(x), y\rangle_{\mathcal{K}}| \leq \|T(x)\|_{\mathcal{K}} \|y\|_{\mathcal{K}} \leq \|T\| \|x\|_{\mathcal{H}} \|y\|_{\mathcal{K}}$$

via the Cauchy-Schwarz inequality, we see that f_y is a bounded linear map. Therefore, by the Riesz Representation Theorem (Theorem 6.4.1) there exists a unique vector, denoted $T_y^* \in \mathcal{H}$ such that

$$\langle T(x), y\rangle_{\mathcal{K}} = f_y(x) = \langle x, T_y^*\rangle_{\mathcal{H}}$$

for all $x \in \mathcal{H}$.

We claim that the map $T^* : \mathcal{K} \to \mathcal{H}$ defined by $T^*(y) = T_y^*$ is a bounded linear map. To see linearity, notice for all $x \in \mathcal{H}$, $y_1, y_2 \in \mathcal{K}$, and $\alpha \in \mathbb{K}$ that

$$\begin{aligned}
\langle x, T^*_{y_1+\alpha y_2}\rangle_{\mathcal{H}} &= \langle T(x), y_1 + \alpha y_2\rangle_{\mathcal{K}} \\
&= \langle T(x), y_1\rangle_{\mathcal{K}} + \overline{\alpha}\langle T(x), y_2\rangle_{\mathcal{K}} \\
&= \langle x, T^*_{y_1}\rangle_{\mathcal{H}} + \overline{\alpha}\langle x, T^*_{y_2}\rangle_{\mathcal{H}} \\
&= \langle x, T^*_{y_1} + \alpha T^*_{y_2}\rangle_{\mathcal{H}}.
\end{aligned}$$

Therefore, as the above holds for all $x \in \mathcal{H}$, we see (for example, by the uniqueness part of the Riesz Representation Theorem (Theorem 6.4.1)) that

$$T^*_{y_1+\alpha y_2} = T^*_{y_1} + \alpha T^*_{y_2}.$$

Therefore, as $y_1, y_2 \in \mathcal{K}$ and $\alpha \in \mathbb{K}$ were arbitrary, T^* is linear.

To see that T^* is bounded, we notice that

$$\begin{aligned}
&\sup\{|\langle T^*(y), x\rangle_{\mathcal{H}}| \mid x \in \mathcal{H}, y \in \mathcal{K}, \|x\|_{\mathcal{H}}, \|y\|_{\mathcal{K}} \leq 1\} \\
&= \sup\{|\langle y, T(x)\rangle_{\mathcal{K}}| \mid x \in \mathcal{H}, y \in \mathcal{K}, \|x\|_{\mathcal{H}}, \|y\|_{\mathcal{K}} \leq 1\} \\
&= \sup\{|\langle T(x), y\rangle_{\mathcal{K}}| \mid x \in \mathcal{H}, y \in \mathcal{K}, \|x\|_{\mathcal{H}}, \|y\|_{\mathcal{K}} \leq 1\} \\
&= \|T\|.
\end{aligned}$$

Thus it follows from Corollary 6.4.3 that T^* is bounded with $\|T^*\| = \|T\|$. Finally, uniqueness of T^* comes from construction and the uniqueness in the Riesz Representation Theorem. ∎

Remark 6.4.5. Note if $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces and $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, then

$$\langle x, T^*(y)\rangle_{\mathcal{H}} = \overline{\langle T^*(y), x\rangle_{\mathcal{H}}} = \overline{\langle y, T(x)\rangle_{\mathcal{K}}} = \langle T(x), y\rangle_{\mathcal{K}}$$

for all $y \in \mathcal{K}$ and $x \in \mathcal{H}$.

Of course, the Hilbert space adjoint has the same properties as the Banach space adjoint from Proposition 1.6.9. Note due to the conjugate linearity in the isomorphism between $\mathcal{H}$ and $\mathcal{H}^*$ that the Hilbert space adjoint is conjugate linear. Moreover, as $\mathcal{H}^{**} = \mathcal{H}$, the Hilbert space adjoint has square equal to the identity. We note the following are trivial to prove based on the above definition.

Lemma 6.4.6. *Let $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{L}$ be Hilbert spaces. Then*

(1) $(T + S)^* = T^* + S^*$ for all $T, S \in \mathcal{B}(\mathcal{H}, \mathcal{K})$,

(2) $(\alpha T)^* = \overline{\alpha} T^*$ for all $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $\alpha \in \mathbb{K}$,

(3) $(T^*)^* = T$ for all $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, and

(4) $(ST)^* = T^*S^*$ for all $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $S \in \mathcal{B}(\mathcal{K}, \mathcal{L})$.

At this point we can finally explain why the notation we are using for the adjoint is the same as the notation used in linear algebra for the conjugate transpose.

Proposition 6.4.7. *Let $\mathcal{H}$ be a complex separable Hilbert space and let $\mathcal{E} = \{e_n\}_{n\in I}$ be an orthonormal basis for $\mathcal{H}$ (so either I is finite or countable). If $T \in \mathcal{B}(\mathcal{H})$, then the matrix of T^* with respect to $\mathcal{E}$ is the conjugate transpose of the matrix of T with respect to $\mathcal{E}$.*

Proof. Recall the (n, m)-entry of the matrix of T is $\langle T(e_m), e_n\rangle$ whereas the (m, n)-entry of the matrix of T^* is

$$\langle T^*(e_n), e_m\rangle = \langle e_n, T(e_m)\rangle = \overline{\langle T(e_m), e_n\rangle}$$

as desired. ∎

Example 6.4.8. Let $A \in \mathcal{M}_n(\mathbb{K})$ and define $L_A : \mathbb{K}^n \to \mathbb{K}^n$ by $L_A(x) = Ax$ for all $x \in \mathbb{K}^n$ (where we write x as a column vector and use matrix multiplication). Then $(L_A)^* = L_{A^*}$ by Proposition 6.4.7.

One surprising useful result related to the norm of an adjoint is the following.

Theorem 6.4.9. *Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{B}(\mathcal{H})$. Then $\|T\|^2 = \|T^*T\|$.*

Proof. First, we note for all $x \in \mathcal{H}$ that

$$\|T^*(T(x))\| \leq \|T^*\| \, \|T(x)\| \leq \|T^*\| \, \|T\| \, \|x\| .$$

Hence $\|T^*T\| \leq \|T^*\| \, \|T\| = \|T\|^2$ as $\|T^*\| = \|T\|$ by Theorem 6.4.4.

To see the other inequality, notice that

$$\begin{aligned}
\|T\|^2 &= \sup\{\|T(x)\|^2 \mid x \in \mathcal{H}, \|x\|_{\mathcal{H}} \leq 1\} \\
&= \sup\{\langle T(x), T(x)\rangle \mid x \in \mathcal{H}, \|x\|_{\mathcal{H}} \leq 1\} \\
&= \sup\{\langle T^*T(x), x\rangle \mid x \in \mathcal{H}, \|x\|_{\mathcal{H}} \leq 1\} \\
&\leq \sup\{\langle T^*T(x), y\rangle \mid x, y \in \mathcal{H}, \|x\|_{\mathcal{H}}, \|y\|_{\mathcal{H}} \leq 1\} \\
&= \|T^*T\|
\end{aligned}$$

by Corollary 6.4.3. ∎

To further our study of bounded linear operators on Hilbert spaces, we desire a deeper analysis of the orthogonal projections. To begin, we prove the following which further relates an operator and its adjoint.

Lemma 6.4.10. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. Then $(T^*)^* = T$ and*

$$(\mathrm{Im}(T))^\perp = \ker(T^*).$$

Hence $\overline{\mathrm{Im}(T)} = \ker(T^*)^\perp$.

Proof. As $(T^*)^* = T$ by Lemma 6.4.6, clearly $\overline{\mathrm{Im}(T)} = \ker(T^*)^\perp$ will follow from $(\mathrm{Im}(T))^\perp = \ker(T^*)$ using Corollary 6.2.12.

To prove that $(\mathrm{Im}(T))^\perp = \ker(T^*)$, let $x \in \ker(T^*)$ be arbitrary. If $y \in \mathrm{Im}(T)$ is arbitrary, then there exists a vector $z \in \mathcal{H}$ such that $y = T(z)$. Hence

$$\langle y, x\rangle_{\mathcal{K}} = \langle T(z), x\rangle_{\mathcal{K}} = \langle z, T^*(x)\rangle_{\mathcal{H}} = \langle z, \vec{0}\rangle_{\mathcal{H}} = 0.$$

Therefore, as $y \in \mathrm{Im}(T)$ was arbitrary, it follows that $x \in (\mathrm{Im}(T))^\perp$. Hence, as $x \in \ker(T^*)$ was arbitrary, $\ker(T^*) \subseteq (\mathrm{Im}(T))^\perp$.

For the other direction, let $x \in (\mathrm{Im}(T))^\perp$ be arbitrary. Then for all $y \in \mathcal{H}$ we see that

$$\langle T^*(x), y\rangle_{\mathcal{H}} = \langle x, T(y)\rangle_{\mathcal{K}} = 0$$

as $T(y) \in \mathrm{Im}(T)$ and $x \in (\mathrm{Im}(T))^\perp$. Therefore, as $y \in \mathcal{H}$ was arbitrary, we see (for example, by the uniqueness part of the Riesz Representation Theorem (Theorem 6.4.1)) that $x \in \ker(T^*)$. Hence, as $x \in (\mathrm{Im}(T))^\perp$ was arbitrary, $\ker(T^*) = (\mathrm{Im}(T))^\perp$ as desired. ∎

Thus we arrive at an alternative characterization of the orthogonal projections.

Proposition 6.4.11. *Let $\mathcal{H}$ be a Hilbert space. An element $P \in \mathcal{B}(\mathcal{H})$ is an orthogonal projection onto a closed vector subspace of $\mathcal{H}$ if and only if $P^2 = P$ and $P^* = P$.*

Proof. Suppose P is the orthogonal projection onto a closed vector subspace $\mathcal{K}$ of $\mathcal{H}$. As we have previously seen that $P^2 = P$, it suffices to show that $P^* = P$. To see this, let $x, y \in \mathcal{H}$ be arbitrary. By Theorem 6.2.9 we can write $x = x_P + x_0$ and $y = y_P + y_0$ where $x_P, y_P \in \mathcal{K}$ and $x_0, y_0 \in \mathcal{K}^\perp$. Therefore we have that

$$P(x_P) = x_P, \quad P(y_P) = y_P, \quad Px_0 = \vec{0}, \quad \text{and} \quad Py_0 = \vec{0}.$$

Hence

$$\begin{aligned}
\langle P^*(x), y\rangle &= \langle x, P(y)\rangle \\
&= \langle x_P + x_0, P(y_P + y_0)\rangle \\
&= \langle x_P + x_0, y_P\rangle \\
&= \langle x_P, y_P\rangle + \langle x_0, y_P\rangle \\
&= \langle x_P, y_P\rangle \\
&= \langle x_P, y_P\rangle + \langle x_P, y_0\rangle \\
&= \langle x_P, y_P\rangle + y_0\rangle \\
&= \langle P(x_P + x_0), y_P\rangle + y_0\rangle \\
&= \langle P(x), y\rangle.
\end{aligned}$$

Therefore, as the above holds for all $y \in \mathcal{H}$, we see that $P^*(x) = P(x)$ for all $x \in \mathcal{H}$. Hence $P^* = P$ as claimed.

For the other direction, let $P \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be such that $P^2 = P = P^*$. Thus $\mathcal{K} = \ker(P)$ is a closed vector subspace. Notice by Lemma 6.4.10 that

$$\mathcal{K}^\perp = \ker(P)^\perp = \overline{\mathrm{Im}(P^*)} = \overline{\mathrm{Im}(P)}.$$

We claim that P is the orthogonal projection onto $\overline{\mathrm{Im}(P)}$. To see this, first we notice that

$$\overline{\mathrm{Im}(P)}^\perp = (\mathcal{K}^\perp)^\perp = \mathcal{K}$$

by Corollary 6.2.12. Therefore, as $P(x) = \vec{0}$ for all $x \in \mathcal{K}$, it suffices to show that P is the identity on $\overline{\mathrm{Im}(P)}$. If $x \in \mathrm{Im}(P)$, then $x = P(y)$ for some $y \in \mathcal{H}$ and thus

$$P(x) = P^2(y) = P(y) = x.$$

Therefore, P is the identity on $\mathrm{Im}(P)$. Hence P is the identity on $\overline{\mathrm{Im}(P)}$ by continuity. Thus the result follows. ∎

Orthogonal projections are examples of the following useful class of operators, which are a generalization of those studied in linear algebra.

Definition 6.4.12. Let $\mathcal{H}$ be a Hilbert space. An element $T \in \mathcal{B}(\mathcal{H})$ is said to be *self-adjoint*(or *Hermitian*) if $T^* = T$.

An alternative characterization of self-adjoint operators is as follows.

Proposition 6.4.13. *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$. Then $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint if and only if $\langle T(h), h\rangle \in \mathbb{R}$ for all $h \in \mathcal{H}$.*

Proof. Clearly if T is self-adjoint, then $T^* = T$ so

$$\overline{\langle T(h), h\rangle} = \langle h, T(h)\rangle = \langle T^*(h), h\rangle = \langle T(h), h\rangle$$

for all $h \in \mathcal{H}$. Thus $\langle T(h), h\rangle \in \mathbb{R}$ for all $h \in \mathcal{H}$.

Conversely, suppose $\langle T(h), h\rangle \in \mathbb{R}$ for all $h \in \mathcal{H}$. Hence

$$\langle T^*(h), h\rangle = \langle h, T(h)\rangle = \langle T(h), h\rangle$$

for all $h \in \mathcal{H}$. Thus the Polarization Identities (Theorem B.1.20) imply that

$$\langle T^*(h), k\rangle = \langle T(h), k\rangle$$

for all $h, k \in \mathcal{H}$ and thus $T^* = T$. ∎

In fact, we can examine isometries and thus unitaries using adjoints as follows.

Proposition 6.4.14. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $V \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. The following are equivalent:*

1. $V^*V = I_{\mathcal{H}}$.
2. $\|V(x)\|_{\mathcal{K}} = \|x\|_{\mathcal{H}}$ *for all $x \in \mathcal{H}$ (that is, V is an isometry).*
3. $\langle V(x), V(y)\rangle_{\mathcal{K}} = \langle x, y\rangle_{\mathcal{H}}$ *for all $x, y \in \mathcal{H}$.*

Proof. First, to see that (1) implies (2), suppose (1) holds. Then for all $x \in \mathcal{H}$

$$\|V(x)\|_{\mathcal{K}}^2 = \langle V(x), V(x)\rangle_{\mathcal{K}} = \langle V^*V(x), x\rangle_{\mathcal{H}} = \langle x, x\rangle_{\mathcal{H}} = \|x\|_{\mathcal{H}}^2 .$$

Hence (2) holds so (1) implies (2)

Next, to see that (2) implies (3), suppose that (2) holds. By the same proof of the Polarization Identity (Theorem B.1.20), we see that

$$\begin{aligned}\langle V(x), V(y)\rangle_{\mathcal{K}} &= \frac{1}{4}\|V(x) + V(y)\|^2 - \frac{1}{4}\|V(x) - V(y)\|^2 \\ &= \frac{1}{4}\|V(x+y)\|^2 - \frac{1}{4}\|V(x-y)\|^2 \\ &= \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 \\ &= \langle x, y\rangle_{\mathcal{H}}\end{aligned}$$

if $\mathbb{K} = \mathbb{R}$ and

$$\begin{aligned}\langle V(x), V(y)\rangle_{\mathcal{K}} &= \frac{1}{4}\sum_{k=1}^{4}\left\|V(x) + i^k V(y)\right\|^2 \\ &= \frac{1}{4}\sum_{k=1}^{4}\left\|V(x + i^k y)\right\|^2 \\ &= \frac{1}{4}\sum_{k=1}^{4}\left\|x + i^k y\right\|^2 \\ &= \langle x, y\rangle_{\mathcal{H}}\end{aligned}$$

if $\mathbb{K} = \mathbb{C}$. Hence (3) follows so (2) implies (3)

Finally, to see that (3) implies (1), suppose (3) holds. Then for all $x, y \in \mathcal{H}$

$$\langle I_{\mathcal{H}}(x), y\rangle_{\mathcal{H}} = \langle x, y\rangle_{\mathcal{H}} = \langle V(x), V(y)\rangle_{\mathcal{K}} = \langle V^*V(x), y\rangle_{\mathcal{H}}.$$

Hence it follows that $V^*V = I$ as desired. ∎

Corollary 6.4.15. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $U \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. The following are equivalent:*

1. *$U^*U = I_{\mathcal{H}}$ and $UU^* = I_{\mathcal{K}}$.*
2. *$\|U(x)\|_{\mathcal{K}} = \|x\|_{\mathcal{H}}$ for all $x \in \mathcal{H}$ and U is surjective.*
3. *$\langle U(x), U(y)\rangle_{\mathcal{K}} = \langle x, y\rangle_{\mathcal{H}}$ for all $x, yy \in \mathcal{H}$ and U is surjective (i.e. U is a unitary).*

Hence, if $U \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ is a unitary, then $\|U\| = 1$.

Proof. Clearly (1) implies (2) and (2) implies (3) by Proposition 6.4.14. Suppose (3) holds. Then $U^*U = \mathcal{I}_{\mathcal{H}}$ by Proposition 6.4.14. Since (3) holds, we obtain U is an isometry by Proposition 6.4.14. Hence U is injective and thus invertible as a linear map between vector spaces. Therefore, due to the uniqueness of the inverses, we obtain that $UU^* = I_{\mathcal{K}}$. ∎

Using all of the above (in fact, using substantially less technology), we can prove the following.

Theorem 6.4.16. *Let $A \in \mathcal{M}_n(\mathbb{K})$ and define $L_A : \mathbb{K}^n \to \mathbb{K}^n$ by $L_A(x) = Ax$ for all $x \in \mathbb{K}^n$ (where we write x as a column vector and use matrix multiplication). Then*

$$\|L_A\| = \max\left\{\sqrt{\lambda} \mid \lambda \text{ an eigenvalue for } A^*A\right\}.$$

Proof. First, consider the case $A = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ and let

$$M = \max\{|d_1|, |d_2|, \ldots, |d_n|\}.$$

To see that $\|L_A\| = M$, first notice for all $k \in \{1, \ldots, n\}$ that if $\vec{e}_k$ is the vector in $\mathbb{C}^n$ with a 1 in the k^{th} entry and 0s elsewhere, then $\|\vec{e}_k\|_2 = 1$ and

$$\|L_A(\vec{e}_k)\|_2 = \|d_k\vec{e}_k\|_2 = |d_k|.$$

Hence $\|L_A\| \geq M$.

To see the reverse inequality, notice for all $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$ such that $\|x\|_2 = \sqrt{\sum_{k=1}^n |x_k|^2} \leq 1$ that

$$\begin{aligned} \|L_A(x)\|_2 &= \|(d_1x_1, d_2x_2, \ldots, d_nx_n)\|_2 \\ &= \sqrt{\sum_{k=1}^n |d_kx_k|^2} \\ &= \sqrt{\sum_{k=1}^n |d_k|^2|x_k|^2} \\ &\leq \sqrt{\sum_{k=1}^n M^2|x_k|^2} \\ &= M\sqrt{\sum_{k=1}^n |x_k|^2} = M. \end{aligned}$$

Hence $\|L_A\| \leq M$ so $\|L_A\| = M$ as desired.

Next,l et $A \in \mathcal{M}_n(\mathbb{C})$ be arbitrary and let $U \in \mathcal{M}_n(\mathbb{C})$ be an arbitrary unitary matrix. Then $L_{U^*AU} = L_{U^*}L_AL_U = L_U^*L_AL_U$ and L_U is a unitary operator. Hence

$$\|L_{U^*AU}\| = \|L_U^*L_AL_U\| \leq \|L_U^*\|\,\|L_A\|\,\|L_U\| = \|L_A\|$$

as unitary operators have norm 1. Moreover, since $L_{U^*AU} = L_U^*L_AL_U$ implies

$$L_A = L_UL_{U^*AU}L_{U^*}$$

as $(L_U^*)^{-1} = L_U$ and $(L_U)^{-1} = L_{U^*}$, we also have that

$$\|L_A\| = \|L_UL_{U^*AU}L_U^*\| \leq \|L_U\|\,\|L_{U^*AU}\|\,\|L_U^*\| = \|L_A\|$$

Hence $\|L_{U^*AU}\| = \|L_A\|$ as desired.

Since A^*A is a self-adjoint matrix and positive semi-definite, the Spectral Theorem for Self-Adjoint Matrices implies there exists a unitary matrix $U \in \mathcal{M}_n(\mathbb{C})$ and a diagonal matrix $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ such that

$A^*A = U^*DU$ and $\lambda_1, \ldots, \lambda_n \in [0, \infty)$ are the eigenvalues of A^*A. Hence we have that

$$\begin{aligned}
\|L_A\| &= \|L_A^* L_A\|^{\frac{1}{2}} \\
&= \|L_{A^*A}\|^{\frac{1}{2}} \\
&= \|L_{U^*DU}\|^{\frac{1}{2}} \\
&= \|L_D\|^{\frac{1}{2}} \\
&= \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}^{\frac{1}{2}} \\
&= \max\left\{\sqrt{\lambda} \mid \lambda \text{ an eigenvalue for } A^*A\right\}
\end{aligned}$$

as desired. ■

Finally, returning to compact operators, we have the following nice characterization.

Theorem 6.4.17. *Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{B}(\mathcal{H})$. The following are equivalent:*

(1) T is compact.

(2) T^* is compact.

(3) There exists a sequence $(F_n)_{n\geq 1} \in \mathcal{F}(\mathcal{H})$ such that $T = \lim_{n\to\infty} F_n$.

Proof. To see that (1) implies (3), let T be a compact operator. Thus $T(\mathcal{H}_1)$ is a totally bounded subset of $\mathcal{H}$ and thus separable. Therefore $\mathcal{K} = \overline{\mathrm{span}(T(\mathcal{H}_1))}$ is a closed separable subspace of $\mathcal{H}$ and thus a separable Hilbert space. Hence there exists a countable orthonormal basis $\{e_n\}_{n=1}^\infty$ of $\mathcal{K}$.

For each $n \in \mathbb{N}$, let P_n be the orthogonal projection onto $\mathrm{span}(\{e_1, \ldots, e_n\})$ and let $F_n = P_n T$. Since P_n has finite dimensional range, $F_n \in \mathcal{F}(\mathcal{H})$ for all $n \in \mathbb{N}$. To complete the proof, it suffices to show that $T = \lim_{n\to\infty} F_n$.

Let $\epsilon > 0$ be arbitrary. Since T is compact, $T(\mathcal{H}_1)$ is totally bounded so there exist a finite set $\{x_j\}_{j=1}^m \subseteq \mathcal{H}_1$ such that $\{T(x_j)\}_{j=1}^m$ is an $\frac{\epsilon}{3}$-net for $T(\mathcal{H}_1)$. Moreover, since $\lim_{n\to\infty} P_n(T(x_j)) = T(x_j)$ for all $j \in \{1, \ldots, m\}$ by Theorem 6.3.14, there exists an $N \in \mathbb{N}$ such that

$$\|F_n(x_j) - T(x_j)\| = \|P_n(T(x_j)) - T(x_j)\| < \frac{\epsilon}{3}$$

for all $n \geq N$.

Let $x \in \mathcal{H}_1$ be arbitrary. Since $\{T(x_j)\}_{j=1}^m$ is an $\frac{\epsilon}{3}$-net for $T(\mathcal{H}_1)$ there exists a $j_0 \in \{1, \ldots, m\}$ such that $\|T(x) - T(x_{j_0})\| < \frac{\epsilon}{3}$. Therefore, for all

$n \geq N$,

$$\begin{aligned}
\|T(x) - F_n(x)\| &\leq \|T(x)_T(x_{j_0})\| + \|T(x_{j_0}) - F_n(x_{j_0})\| + \|F_n(x_{j_0}) - F_n(x)\| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \|P_n(T(x_{j_0}) - T(x))\| \\
&\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \|P_n\| \, \|T(x_{j_0}) - T(x)\| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Therefore, as $x \in \mathcal{H}_1$ was arbitrary, $\|T - F_n\| < \epsilon$ for all $n \geq N$. Hence $T = \lim_{n\to\infty} F_n$.

To see that (3) implies (2), let $(F_n)_{n\geq 1} \in \mathcal{F}(\mathcal{H})$ such that $T = \lim_{n\to\infty} F_n$. Hence $T^* = \lim_{n\to\infty} F_n^*$ by Theorem 6.4.4.

We claim that F_n^* is finite rank for all $n \in \mathbb{N}$. Indeed as F_n is finite rank, $\mathcal{Y}_n = \mathrm{Im}(F_n)$ is a finite dimensional subspace of $\mathcal{H}$. Notice if $x \in \mathrm{Im}(F_n)^\perp$, then for all $h \in \mathcal{H}$ we have that

$$\langle F_n^*(x), y\rangle = \langle x, F_n(y)\rangle = 0$$

so that $F_n^*(x) = 0$ for all $x \in \mathrm{Im}(F_n)^\perp$. Hence $F^*(\mathcal{H}) = \mathcal{F}^*(\mathrm{Im}(F_n))$, which must be finite dimensional since $\mathrm{Im}(F_n)$ is finite dimensional. Hence F_n^* is finite rank for all $n \in \mathbb{N}$ so T^* is a limit of finite rank operators and thus compact by Proposition 6.1.3 and Corollary 6.1.8.

Finally, to see that (2) implies (1), suppose T^* is compact. Hence, by (1) implies (3) implies (2), we obtain that $(T^*)^* = T$ is compact thereby completing the proof. ∎

Remark 6.4.18. For a compact operator T, note in the proof of Theorem 6.4.17 that if we replace P_n with any sequence of projections that converge to the identity in the Strong Operator Topology, then P_nT converges in norm to T. Thus, by repeating the same argument with T^* in place of T, we obtain as each orthogonal projection is self-adjoint that TP_n also converges in norm to T. Therefore, as

$$\begin{aligned}
\|P_nTP_n - T\| &\leq \|P_nTP_n - TP_n\| + \|TP_n - T\| \\
&\leq \|P_nT - T\| \, \|P_n\| + \|TP_n - T\| \\
&\leq \|P_nT - T\| + \|TP_n - T\|,
\end{aligned}$$

we obtain that P_nTP_n converges to T in norm. This is quite useful in that which follows.

Thus we can produce another interesting example of a compact operator.

Example 6.4.19. Let $\mathcal{H} = L_2([0,1], \lambda)$ where λ is the Lebesgue measure and let $K \in L_2([0,1]^2, \lambda \times \lambda)$. The *Volterra operator with kernel* K is the map $V : \mathcal{H} \to \mathcal{H}$ by

$$(Vf)(x) = \int_0^1 f(y)K(x,y)\, dy$$

for all $x \in [0,1]$ and $f \in \mathcal{H}$. Note as $(x,y) \mapsto f(y)K(x,y)$ is an element of $L_1([0,1]^2, \lambda \times \lambda)$ by Hölder's inequality (Theorem D.1.7) that Vf is indeed a measurable function by Fubini's Theorem. Moreover, since

$$\begin{aligned}
\|V(f)\|_2^2 &= \int_0^1 |(V(f))(x)|^2 \, dx \\
&= \int_0^1 \left| \int_0^1 f(y)K(x,y)\, dy \right|^2 \, dx \\
&\leq \int_0^1 \left(\int_0^1 |f(y)K(x,y)|\, dy \right)^2 \, dx \\
&\leq \int_0^1 \|f\|_2^2 \int_0^1 |K(x,y)|^2 \, dy \, dx \quad \text{by Cauchy-Schwarz} \\
&= \|f\|_2^2 \, \|K\|_2^2 .
\end{aligned}$$

Hence V does indeed map $\mathcal{H}$ into $\mathcal{H}$. As V is clearly linear, the above shows that V is bounded with $\|V\| \leq 1$.

We claim that V is compact. To see this, we claim that V is a limit of finite rank operators. To see this, let $\epsilon > 0$ be arbitrary and let

$$\mathcal{A} = \left\{ \sum_{k=1}^n f_k(x) g_k(y) \;\middle|\; n \in \mathbb{N}, \{f_k, g_k\}_{k=1}^n \in C[0,1] \right\}.$$

By the Stone-Weirstrass Theorem, $\mathcal{A}$ is dense in $C([0,1]^2)$ with respect to the infinite norm. Therefore, as $C([0,1]^2)$ is dense in $L_2([0,1]^2, \lambda \times \lambda)$ by Lusin's Theorem. Hence there exists a $K_0 \in \mathcal{A}$ such that $\|K - K_0\|_2 < \epsilon$. Hence if we define V_0 to be the Volterra operator with kernel K_0, then $V - V_0$ is the Volterra operator with kernel $K - K_0$ so

$$\|V - V_0\| \leq \|K - K_0\|_2 < \epsilon.$$

We claim that V_0 is finite rank. Indeed, as $K_0 \in \mathcal{A}$, we can write

$$K_0 = \sum_{k=1}^n f_k(x) g_k(y)$$

for some $n \in \mathbb{N}$ and $\{f_k, g_k\}_{k=1}^n \in C[0,1]$. Thus for all $f \in \mathcal{H}$

$$(V_0 f)(x) = \int_0^1 f(y) K_0(x,y) \, dy = \sum_{k=1}^n a_k f_k(x)$$

for some $\{a_k\}_{k=1}^n \in \mathbb{K}$ for all $x \in [0,1]$. Hence V_0 is finite rank. Therefore, as $\epsilon > 0$ was arbitrary, V is a compact operator.

6.5 Spectral Theorem for Compact Operators

“Thank you Mario! But our Princess is in another castle.”

To be precise, the proof of the Spectral Theorem for Compact Operators is absent from these notes as the proof will be presented to start Functional Analysis II.

Appendix A

Topological Spaces

In this section, we will briefly review the basics of topological spaces to the extent that is required for functional analysis. More advanced requirements from topology will be produced as needed throughout these notes.

A.1 Topologies

Topology, from the Greek $\tau\acute{o}\pi o\sigma$ meaning place and $\lambda\acute{o}\gamma o\sigma$ meaning study, is the study of properties of spaces and their deformations. Such a study is performed by looking at subsets that cover the entire space with certain properties.

Definition A.1.1. Let X be a set. A set $\mathcal{T} \subseteq \mathcal{P}(X)$ is said to be a *topology* on X if

(1) $\emptyset, X \in \mathcal{T}$,

(2) (closed under unions) if $\{U_\alpha\}_{\alpha\in I} \subseteq \mathcal{T}$, then $\bigcup_{\alpha\in I} U_\alpha \in \mathcal{T}$, and

(3) (closed under finite intersections) if $\{U_\alpha\}_{\alpha\in I} \subseteq \mathcal{T}$ and I is finite, then $\bigcap_{\alpha\in I} U_\alpha \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ is called a *topological space* and elements of $\mathcal{T}$ are called the *open sets* of $(X, \mathcal{T})$.

There are many topologies we can place on a given set X so by saying that $(X, \mathcal{T})$ is a topological space means we have fixed $\mathcal{T}$ to be the topology on X. Once a topology is fixed on a set, one can think of the open sets as the sets that describe how points are related to one another. In particular, open sets provide some notion of whether two points are 'close' together; that is, given two points $x, y \in X$ and a $U \in \mathcal{T}$ such that $x \in U$, then y is close to x with respect to U only if $y \in U$. Thus we can see the above definition and thoughts are motivated by undergraduate real analysis where the 'open sets' on $\mathbb{R}$ were the sets that were unions of open intervals and that two points were 'close' only if there was a 'small' open interval around open point which contained the other.

As we desire to study topological spaces, it is useful to have some examples to keep in mind. Of course the examples presented in this section are not all the examples in existence and we will continually encounter new topologies through the course.

Example A.1.2. Let X be a set. Then $\mathcal{T} = \{\emptyset, X\}$ is a topology on X known as the *trivial topology*. This name derives from the fact that the open sets do not distinguish any two elements of X and most topological results become trivial if we consider this topology. We remark that it is trivial to verify the trivial topology is a topology.

Example A.1.3. Let X be a set. Then $\mathcal{T} = \mathcal{P}(X)$ is a topology on X known as the *discrete topology*. This name derives from the fact that every set is open so singleton sets are open and thus every point is separated from the others. We remark that it is trivial to verify the discrete topology is a topology.

Example A.1.4. Let X be any set and let

$$\mathcal{T} = \{\emptyset\} \cup \{A \subseteq X \mid X \setminus A \text{ is finite}\}.$$

Then $\mathcal{T}$ is a topology on X. To see this, we note that clearly $\emptyset \in \mathcal{T}$ and that $X \in \mathcal{T}$ as $X \setminus X = \emptyset$. Next, to see that $\mathcal{T}$ is closed under unions, let $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ be arbitrary. Thus $X \setminus A_\alpha$ is finite for all $\alpha \in I$. Since

$$X \setminus \left(\bigcup_{\alpha \in I} A_\alpha \right) = \bigcap_{\alpha \in I} \left(X \setminus A_\alpha \right),$$

we see that $X \setminus \left(\bigcup_{\alpha \in I} A_\alpha \right)$ is a subset of a finite set and thus finite. Hence $\bigcup_{\alpha \in I} A_\alpha \in \mathcal{T}$ by definition. Finally, to see that $\mathcal{T}$ is closed under finite intersections, let $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ with I finite be arbitrary. Thus $X \setminus A_\alpha$ is finite for all $\alpha \in I$. Since

$$X \setminus \left(\bigcap_{\alpha \in I} A_\alpha \right) = \bigcup_{\alpha \in I} \left(X \setminus A_\alpha \right),$$

we see that $X \setminus \left(\bigcap_{\alpha \in I} A_\alpha \right)$ a finite union of finite sets and thus finite. Hence $\bigcap_{\alpha \in I} A_\alpha \in \mathcal{T}$ by definition.

The topology $\mathcal{T}$ on X is called the *cofinite topology on* X.

Example A.1.5. Let X be any set and let

$$\mathcal{T} = \{\emptyset\} \cup \{A \subseteq X \mid X \setminus A \text{ is countable}\}.$$

Then $\mathcal{T}$ is a topology on X. To see this, one need to simply repeat the proof of Example A.1.4 with 'finite' replaced with 'countable' in the appropriate places.

The topology $\mathcal{T}$ on X is called the *cocountable topology on* X.

Example A.1.6. Let X be any set and let

$$\mathcal{T} = \{A \subseteq X \mid A \text{ is finite}\}.$$

Notice if X is finite then $\mathcal{T} = \mathcal{P}(X)$ so that $\mathcal{T}$ is the discrete topology on X. However, if X is infinite then $\mathcal{T}$ is not a topology on X as $\mathcal{T}$ is not closed under unions (i.e. a countable union of finite sets is not finite).

Since we have seen that there are many possible topologies on a given set, it is useful to be able to compare the size of these topologies. The simplest way to compare topologies is based on inclusion.

Definition A.1.7. Let $\mathcal{T}$ and $\mathcal{T}'$ be two topologies on a set X. It is said that $\mathcal{T}$ is *finer* that $\mathcal{T}'$ or, equivalently, that $\mathcal{T}'$ is *coarser* than $\mathcal{T}$ if $\mathcal{T}' \subseteq \mathcal{T}$. In the case the inclusion is strict (i.e. $\mathcal{T}' \subsetneq \mathcal{T}$), it is said that $\mathcal{T}$ is *strictly finer* that $\mathcal{T}'$ or, equivalently, that $\mathcal{T}'$ is *coarser strictly* than $\mathcal{T}$. Finally, it is said that $\mathcal{T}$ and $\mathcal{T}'$ are *comparable* if $\mathcal{T} \subseteq \mathcal{T}'$ or $\mathcal{T}' \subseteq \mathcal{T}$.

The above terminology is derived from the fact that "if you have more sets in your topology, you can 'divide up your space' more finely". That is, the more pixels per square inch, the finer the image.

Example A.1.8. The discrete topology on a set is always finer than any other topology on the set and the trivial topology is always coarser than any other topology on the set. Provided the set is non-empty and not a singleton, the discrete topology is strictly finer than the trivial topology.

Example A.1.9. The cofinite topology is coarser than the cocountable topology and will be strictly coarser provided the set is infinite.

Example A.1.10. Consider the set X consisting of three distinct points $\{a, b, c\}$ and the following topologies on X:

$$\begin{aligned}\mathcal{T}_1 &= \{\emptyset, \{b\}, \{c\}, \{b, c\}, X\}\\ \mathcal{T}_2 &= \{\emptyset, \{b\}, \{a, b\}, \{b, c\}, X\}.\end{aligned}$$

As $\mathcal{T}_1 \not\subseteq \mathcal{T}_2$ and $\mathcal{T}_2 \not\subseteq \mathcal{T}_1$, $\mathcal{T}_1$ and $\mathcal{T}_2$ are not comparable topologies on X. Hence it is possible to have topologies that are not comparable.

A.2 Bases

In order to have a better understanding and control over topologies, we desire to describe smaller collections of open sets that determine the entire topology. For example, the metric topologies it is know that all open sets are unions of open balls, so provided we can understand the open balls we should be able to understand the entire topology. In particular, if we have the following sets whose properties are in analogy with the open balls in a metric space, we can form topologies with specific properties.

Theorem A.2.1. *Let X be a non-empty set and let $\mathcal{B} \subseteq \mathcal{P}(X)$ be such that*

(1) if $x \in X$ then there exists a $B \in \mathcal{B}$ such that $x \in B$, and

(2) if $x \in X$ and $B_1, B_2 \in \mathcal{B}$ are such that $x \in B_1 \cap B_2$, then there exists a $B_3 \in \mathcal{B}$ such that $x \in B_3$, $B_3 \subseteq B_1$, and $B_3 \subseteq B_2$.

Let $\mathcal{T}_\mathcal{B}$ be the set of all subsets U of X such that for all $x \in U$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$. Then $\mathcal{T}_\mathcal{B}$ is a topology on X such that $\mathcal{B} \subseteq \mathcal{T}$.

Proof. To see that $\mathcal{T}_\mathcal{B}$ is a topology, we must verify the three properties in Definition A.1.1. It is clear by definition of $\mathcal{T}_\mathcal{B}$ that $\emptyset \in \mathcal{T}_\mathcal{B}$.

To see that $X \in \mathcal{T}_\mathcal{B}$ recall by property (1) that for each $x \in X$ there exists an $B_x \in \mathcal{B}$ such that $x \in B_x$. As $B_x \subseteq X$ by definition, we obtain that $X \in \mathcal{T}_\mathcal{B}$ by the definition of $\mathcal{T}_\mathcal{B}$.

Next suppose $\{U_\alpha\}_{\alpha \in I}$ is a set of elements of $\mathcal{T}_\mathcal{B}$. To see that $\bigcup_{\alpha \in I} U_\alpha \in \mathcal{T}_\mathcal{B}$, let $x \in \bigcup_{\alpha \in I} U_\alpha$ be arbitrary. Then there must be an $\alpha_0 \in I$ such that $x \in U_{\alpha_0}$. Since $U_{\alpha_0} \in \mathcal{T}_\mathcal{B}$, there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U_{\alpha_0}$. Hence $B \subseteq U_{\alpha_0} \subseteq \bigcup_{\alpha \in I} U_\alpha$. As $x \in \bigcup_{\alpha \in I} U_\alpha$ was arbitrary, we obtain that $\bigcup_{\alpha \in I} U_\alpha \in \mathcal{T}_\mathcal{B}$ by definition.

To complete the proof that $\mathcal{T}_\mathcal{B}$ is topology, suppose $U_1, \ldots, U_n \in \mathcal{T}_\mathcal{B}$. To see that $\bigcap_{k=1}^n U_k \in \mathcal{T}_\mathcal{B}$, let $x \in \bigcap_{k=1}^n U_k$ be arbitrary. Hence $x \in U_k$ for all k so, as each $U_k \in \mathcal{T}_\mathcal{B}$, there exists a $B_k \in \mathcal{B}$ such that $x \in B_k$ and $B_k \subseteq U_k$. By applying property (2) recursively $n-1$ times, there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq B_k$ for all k. Hence $B \in \mathcal{B}$, $x \in B$, and $B \subseteq B_k \subseteq U_k$ for all k so that $B \subseteq \bigcap_{k=1}^n U_k$. Therefore, as $x \in X$ was arbitrary, $\bigcap_{k=1}^n U_k \in \mathcal{T}_\mathcal{B}$ as desired.

Finally, the fact that $\mathcal{B} \subseteq \mathcal{T}_\mathcal{B}$ follows from the definition of $\mathcal{T}_\mathcal{B}$. ∎

As subsets of the power set of a given set as described in Theorem A.2.1 are useful in constructing topologies, we define the following.

Definition A.2.2. Let X be a non-empty set. A *basis for a topology on X* is a collection of subsets $\mathcal{B} \subseteq \mathcal{P}(X)$ such that

(1) if $x \in X$ then there exists a $B \in \mathcal{B}$ such that $x \in B$, and

(2) if $x \in X$ and $B_1, B_2 \in \mathcal{B}$ are such that $x \in B_1 \cap B_2$, then there exists a $B_3 \in \mathcal{B}$ such that $B_3 \subseteq B_1$ and $B_3 \subseteq B_2$.

The topology $\mathcal{T}_\mathcal{B}$ on X from Theorem A.2.1 is called the *topology generated by the basis $\mathcal{B}$*. Note that a set $U \subseteq X$ is open with respect to $\mathcal{T}_\mathcal{B}$ if and only if for every $x \in U$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$. Consequently $\mathcal{B} \subseteq \mathcal{T}_\mathcal{B}$.

Here is one example of how we can construct topologies via bases that turns out not to be a topology we have previously seen.

Example A.2.3. Let

$$\mathcal{B} = \{[a, b) \mid a, b \in \mathbb{R}, a < b\}.$$

We claim that $\mathcal{B}$ is a basis for a topology on $\mathbb{R}$. To see this, it suffices to verify the two defining properties of being a basis from Definition A.2.2. To being, notice if $x \in \mathbb{R}$ then $x \in [x, x+1) \in \mathcal{B}$. Hence the first property is satisfied. To see the second property, let $[a_1, b_1), [a_2, b_2) \in \mathcal{B}$ and $x \in \mathbb{R}$ such that $x \in [a_1, b_1) \cap [a_2, b_2)$ be arbitrary. Let

$$a = \max(\{a_1, a_2\}) \qquad \text{and} \qquad b = \min(\{b_1, b_2\})$$

and let $B = [a, b)$. Since $x \in [a_1, b_1) \cap [a_2, b_2)$, we see that $a \leq x < b$ so $B \in \mathcal{B}$ and $x \in B$. Furthermore, by construction, $B \subseteq [a_1, b_1) \cap [a_2, b_2)$. Hence, since $[a_1, b_1), [a_2, b_2) \in \mathcal{B}$ and $x \in \mathbb{R}$ were arbitrary, $\mathcal{B}$ is a basis for a topology on $\mathbb{R}$.

The topology $\mathcal{T}_L$ on $\mathbb{R}$ generated by the basis $\mathcal{B}$ is called the *lower limit topology on* $\mathbb{R}$.

The fact that $\mathcal{T}_L$ is not the same as the canonical topology on $\mathbb{R}$ will come from material in Section A.4 where we show that 'limits' behave different in these topologies. In particular, we will see why we call $\mathcal{T}_L$ the lower limit topology. Alternatively, we know that $[a, b)$ is open in the lower limit topology, but is not open in the canonical topology.

Topologies generated by a basis are particularly nice since it is very simple to completely understand the entire topology via the basis elements based on the above and below descriptions of open sets.

Theorem A.2.4. *Let X be a non-empty set and let $\mathcal{B}$ be a basis for a topology on X. Then*

$$\mathcal{T}_{\mathcal{B}} = \left\{ \bigcup_{B \in \mathcal{B}_0} B \;\middle|\; \mathcal{B}_0 \subseteq \mathcal{B} \right\}.$$

Proof. Notice, since $\mathcal{T}_{\mathcal{B}}$ is a topology and since $\mathcal{B} \subseteq \mathcal{T}_{\mathcal{B}}$, we know that for all $\mathcal{B}_0 \subseteq \mathcal{B}$ that

$$\bigcup_{B \in \mathcal{B}_0} B$$

is a union of elements of $\mathcal{T}_{\mathcal{B}}$ and thus in $\mathcal{T}_{\mathcal{B}}$. Hence

$$\mathcal{T}_{\mathcal{B}} \supseteq \left\{ \bigcup_{B \in \mathcal{B}_0} B \;\middle|\; \mathcal{B}_0 \subseteq \mathcal{B} \right\}.$$

To see the other inclusion, let $U \in \mathcal{T}_{\mathcal{B}}$ be arbitrary. By the definition of $\mathcal{T}_{\mathcal{B}}$, for each $x \in U$ there exists a $B_x \in \mathcal{B}$ such that $x \in B_x$ and $B_x \subseteq U$. Hence we see that

$$U = \bigcup_{x \in U} B_x.$$

Therefore, as $U \in \mathcal{T}_\mathcal{B}$ was arbitrary, we obtain that

$$\mathcal{T}_\mathcal{B} = \left\{ \bigcup_{B \in \mathcal{B}_0} B \,\middle|\, \mathcal{B}_0 \subseteq \mathcal{B} \right\}$$

as claimed. ∎

As bases for a topology give us multiple nice descriptions of the open sets for that topology, it is useful when given a topology to have a basis that generates the given topology. Thus to simplify this terminology, we define the following.

Definition A.2.5. Let $(X, \mathcal{T})$ be a topological space. A set $\mathcal{B} \subseteq \mathcal{P}(X)$ is said to be a *basis for* $(X, \mathcal{T})$ if $\mathcal{B}$ is a basis for a topology on X and $\mathcal{T}_\mathcal{B} = \mathcal{T}$.

Remark A.2.6. Of course, as $\mathcal{B} \subseteq \mathcal{T}_\mathcal{B}$ for any basis of a topology $\mathcal{B}$, for a set $\mathcal{B} \subseteq \mathcal{P}(X)$ to be a basis for a topology $\mathcal{T}$, it must be the case that $\mathcal{B} \subseteq \mathcal{T}$. Furthermore, by Theorem A.2.1 and Theorem A.2.4, we see that if $\mathcal{B}$ is a basis for $(X, \mathcal{T})$ then

(1) a set $U \subseteq X$ is open if and only if for every $x \in U$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$, and

(2) the open sets in $(X, \mathcal{T})$ are exactly the union of elements of $\mathcal{B}$.

Furthermore, it is not difficult to see that every topology is generated by a basis as the following example shows.

Example A.2.7. Let $(X, \mathcal{T})$ be a topological space. Then $\mathcal{T}$ is a basis for $(X, \mathcal{T})$. Of course this is not the most useful basis as our goal is to better understand $\mathcal{T}$ by using a basis with as few elements as possible.

Of course, many of our previously discussed topologies have far nicer bases.

Example A.2.8. Let X be a non-empty set and let $\mathcal{T}$ be the discrete topology on X. Then

$$\mathcal{B} = \{\{x\} \mid x \in X\}$$

is a basis for $(X, \mathcal{T})$. Indeed clearly $\mathcal{B} \subseteq \mathcal{T}$ as $\mathcal{T}$ is the discrete topology. Next clearly the first property of Definition A.2.2 holds and the second property also clearly holds since the only way $x \in X$ and $B_1, B_2 \in \mathcal{B}$ are such that $x \in B_1 \cap B_2$ is if $B_1 = B_2 = \{x\} \in \mathcal{B}$. Hence $\mathcal{B}$ is a basis for $(X, \mathcal{T})$.

Of course, we have our motivating example.

Example A.2.9. Let (X, d) be a metric space. Then the set $\mathcal{B}$ of all open balls forms a basis for $(X, \mathcal{T}_d)$. Indeed clearly $\mathcal{B} \subseteq \mathcal{T}$ and if $x \in X$ then $B_d(x, 1) \in \mathcal{T}_d$ so the first property of Definition A.2.2 is satisfied. To see the second property of Definition A.2.2 is satisfied, let $x \in X$ and $B_1, B_2 \in \mathcal{B}$ be arbitrary such that $x \in B_1 \cap B_2$. Then there exists points $x_1, x_2 \in X$ and $r_1, r_2 > 0$ such that $B_1 = B_d(x_1, r_1)$ and $B_2 = B_d(x_2, r_2)$. Thus, as $x \in B_1 \cap B_2$, we see that

$$d(x, x_1) < r_1 \qquad \text{and} \qquad d(x, x_2) < r_2.$$

Let

$$r = \min\{r_1 - d(x, x_1), r_2 - d(x, x_2)\}.$$

Then $r > 0$. It is elementary to verify that $B_d(x, r) \subseteq B_d(x_1, r_1)$ and $B_d(x, r) \subseteq B_d(x_2, r_2)$. Hence, as $x \in X$ and $B_1, B_2 \in \mathcal{B}$ were arbitrary, the second property of Definition A.2.2 is satisfied so that $\mathcal{B}$ is a basis for $(X, \mathcal{T}_d)$.

Example A.2.10. Let (X, d) be a metric space and let $\epsilon > 0$. The set $\mathcal{B}$ of all open balls with radius at most ϵ forms a basis for $(X, \mathcal{T}_d)$. Indeed the proof is identical to that of Example A.2.9 with the additional restraint that all radii involved are at most ϵ.

Example A.2.11. Let (X, d) be a metric space. The set $\mathcal{B}$ of all open balls with radius positive rational radii forms a basis for $(X, \mathcal{T}_d)$. Indeed the proof is identical to that of Example A.2.9 with the additional restraint that all radii involved are rational. This is advantageous over Examples A.2.9 and A.2.10 as this basis only has a countable number of elements centred at each point.

There are alternate characterization for a basis for a topological space. In particular, the following is superior to Definition A.2.2 in checking that a collection of sets is a basis for a specific topology and is the converse to fact (1) in Remark A.2.6.

Proposition A.2.12. *Let $(X, \mathcal{T})$ be a topological space. Suppose $\mathcal{B} \subseteq \mathcal{T}$ has the property that for all $U \in \mathcal{T}$ and for all $x \in U$ there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$. Then $\mathcal{B}$ is a basis for $(X, \mathcal{T})$.*

Proof. To see that $\mathcal{B}$ is a basis for a topology on X, we will simply verify the two properties in Definition A.2.2. To begin, let $x \in X$ be arbitrary. Then, as $X \in \mathcal{T}$, the assumptions of the proposition imply there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq X$. Hence, as $x \in X$ was arbitrary, the first assumption of Definition A.2.2 has been verified.

To see the second property of Definition A.2.2 holds, let $x \in X$ and $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \cap B_2$ be arbitrary. As $\mathcal{B} \subseteq \mathcal{T}$, we see that $B_1, B_2 \in \mathcal{T}$ and thus $B_1 \cap B_2 \in \mathcal{T}$. Therefore, by the assumptions of the proposition there exists a $B_3 \in \mathcal{B}$ such that $x \in B_3$ and $B_3 \subseteq B_1 \cap B_2$. Hence, as $x \in X$ and $B_1, B_2 \in \mathcal{B}$ were arbitrary, the second property of Definition A.2.2 has been verified. Thus $\mathcal{B}$ is a basis for a topology on X.

To see that $\mathcal{T} = \mathcal{T}_\mathcal{B}$, we first note that as $\mathcal{B} \subseteq \mathcal{T}$ and as $\mathcal{T}$ is closed under unions, Theorem A.2.4 implies that

$$\mathcal{T}_\mathcal{B} = \left\{ \bigcup_{B \in \mathcal{B}_0} B \,\middle|\, \mathcal{B}_0 \subseteq \mathcal{B} \right\} \subseteq \mathcal{T}.$$

Conversely, if $U \in \mathcal{T}$ then the assumptions of the proposition imply that $U \in \mathcal{T}_\mathcal{B}$ by Definition A.2.2. Hence $\mathcal{T} = \mathcal{T}_\mathcal{B}$ as desired. ∎

Using the above, we obtain our final characterization of a basis for a topological space which acts as a converse to Theorem A.2.4.

Corollary A.2.13. *Let $(X, \mathcal{T})$ be a topological space. Suppose $\mathcal{B} \subseteq \mathcal{T}$ has the property that for every $U \in T$ there exists a subset $\mathcal{B}_0 \subseteq \mathcal{B}$ such that $U = \bigcup_{B \in \mathcal{B}_0} B$. Then $\mathcal{B}$ is a basis for $(X, \mathcal{T})$.*

Proof. To prove this result, we will verify that the assumption of Proposition A.2.12 holds. To see this, let $U \in \mathcal{T}$ and $x \in U$ be arbitrary. Then, by the assumptions of this corollary, there exists a subset $\mathcal{B}_0 \subseteq \mathcal{B}$ such that $U = \bigcup_{B \in \mathcal{B}_0} B$. Hence, as $x \in U$, there exists a $B_x \in \mathcal{B}_0$ such that $x \in B_x$ and $B_x \subseteq \bigcup_{B \in \mathcal{B}_0} B = U$. Therefore, as $U \in \mathcal{T}$ and $x \in U$ were arbitrary, the assumption of Proposition A.2.12 holds. Hence the result follows. ∎

It is often more useful and convenient to work with a basis for a topological space than the topology itself. For example, the following demonstrates how to use bases to determine when one topology is finer or coarser than another. In particular, we will often use the case that one of the bases for one of the topologies is the topology itself, which is valid by Example A.2.7.

Theorem A.2.14. *Let $\mathcal{T}$ and $\mathcal{T}'$ be topologies on a set X and let $\mathcal{B}$ and $\mathcal{B}'$ be bases for $\mathcal{T}$ and $\mathcal{T}'$ respectively. Then the following are equivalent:*

(i) $\mathcal{T}'$ is finer than $\mathcal{T}$.

(ii) For every $x \in X$ and $B \in \mathcal{B}$ such that $x \in B$ there exists a $B' \in \mathcal{B}'$ such that $x \in B'$ and $B' \subseteq B$.

Proof. First suppose that $\mathcal{T}'$ is finer that $\mathcal{T}$. Thus $\mathcal{T} \subseteq \mathcal{T}'$. To see that (ii) holds, let $x \in X$ and $B \in \mathcal{B}$ such that $x \in B$ be arbitrary. Since $\mathcal{B}$ is a basis for $\mathcal{T}$, $\mathcal{B} \subseteq \mathcal{T} \subseteq \mathcal{T}'$. Thus $B \in \mathcal{T}'$. Therefore, as $x \in B$, as $B \in \mathcal{T}'$, and as $\mathcal{B}'$ is a basis for $\mathcal{T}'$, we obtain that there exists a $B' \in \mathcal{B}'$ such that $x \in B'$ and $B' \subseteq B$. Therefore as $x \in X$ and $B \in \mathcal{B}$ were arbitrary, (ii) follows.

Conversely, suppose that (ii) holds. To see that T' is finer than $\mathcal{T}$, let $U \in \mathcal{T}$ be arbitrary. To see that $U \in \mathcal{T}'$, let $x \in U$ be arbitrary. As $U \in \mathcal{T}$, there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq U$. Then, by assumption (ii), there exist a $B' \in \mathcal{B}'$ such that $x \in B'$ and $B' \subseteq B \subseteq U$. Therefore $U \in \mathcal{T}'$ as $\mathcal{B}'$ is a basis for $\mathcal{T}'$. Hence, as $U \in \mathcal{T}$ was arbitrary, $\mathcal{T} \subseteq \mathcal{T}'$ so $\mathcal{T}'$ is finer than $\mathcal{T}$. ∎

Of course, to generate a topology from a basis, we need a basis. This leads to the question about how can we construct collections of sets that satisfy the assumptions of Definition A.2.2. As the second property required in Definition A.2.2 relates to the intersection of basis elements containing basis elements, one way to avoid this problem is by taking all intersections of the sets we want to use to form a basis. Thus we define the following object.

Definition A.2.15. Let $(X, \mathcal{T})$ be a topological space. A *subbasis* for $(X, \mathcal{T})$ is a collection of subsets $\mathcal{S} \subseteq \mathcal{T}$ such that the set of all finite intersections of elements of $\mathcal{S}$ is a basis for $(X, \mathcal{T})$.

Of course, for a set $\mathcal{S}$ to be a subbasis of some topology $\mathcal{T}$ on X, it is necessary that

$$X = \bigcup_{S \in \mathcal{S}} S.$$

as for each $x \in X$ there must be a basis element containing X. In fact, this is the only restriction for a collection of sets to be a subbasis for some topology on X.

Theorem A.2.16. *Let X be a non-empty set and let $\mathcal{S} \subseteq \mathcal{P}(X)$ be such that*

$$X = \bigcup_{S \in \mathcal{S}} S.$$

Let $\mathcal{B} \subseteq \mathcal{P}(X)$ be the set of all finite intersections of elements of $\mathcal{S}$. Then $\mathcal{B}$ is a basis for a topology on X for which $\mathcal{S}$ is a subbasis.

Proof. To see that $\mathcal{B}$ is a basis for a topology on X, we need only check the two conditions on a basis from Definition A.2.2. For the first, let $x \in X$ be arbitrary. Since $X = \bigcup_{S \in \mathcal{S}} S$ there exists an $S_x \in \mathcal{S}$ such that $x \in S_x$. As $\mathcal{S} \subseteq \mathcal{B}$, the first property of being a basis holds for $\mathcal{B}$.

For the second property, let $x \in X$ and let $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \cap B_2$ be arbitrary. Since $B_1, B_2 \in \mathcal{B}$, B_1 and B_2 are finite intersections of elements of $\mathcal{S}$. Hence $B_1 \cap B_2$ is a finite intersection of elements of $\mathcal{S}$ so that $B_1 \cap B_2 \in \mathcal{B}$. Therefore, as $x \in X$ and let $B_1, B_2 \in \mathcal{B}$ were arbitrary, the first property of being a basis holds for $\mathcal{B}$. Hence $\mathcal{B}$ is a basis for a topology on X. The fact that that $\mathcal{S}$ is a subbasis is a subbasis for $\mathcal{T}_{\mathcal{B}}$ is then trivial. ∎

Subbases are not as desirable as bases as the description of the entire topology is far more difficult using subbases than bases and thus make subbases far more difficult to use thereby limiting their applications. However, subbases are excellent for constructing topologies as the conditions that are required are far simpler.

A.3 Constructing Topologies

Since bases and subbases are so great for constructing topologies, let us examine how we can construct new topologies from old topologies. The first such example of this comes from restricting a topology on a set to a subset of the set.

Lemma A.3.1. *Let $(X, \mathcal{T})$ be a topological space and let $Y \subseteq X$ be non-empty. The set*

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\}$$

is a topology on Y.

Proof. To verify that $\mathcal{T}_Y$ is a topology on Y, we verify Definition A.1.1. First clearly $\mathcal{T}_Y \subseteq \mathcal{P}(Y)$ by construction. Next, as $\emptyset, X \in \mathcal{T}$, we obtain that $\emptyset = Y \cap \emptyset \in \mathcal{T}_Y$ and $Y = Y \cap X \in \mathcal{T}_Y$. The fact that $\mathcal{T}_Y$ is closed under unions and finite intersections then follows from the facts that

$$\bigcup_{\alpha \in I} (Y \cap U_\alpha) = Y \cap \left(\bigcup_{\alpha \in I} U_\alpha \right) \text{ and}$$
$$\bigcap_{\alpha \in I} (Y \cap U_\alpha) = Y \cap \left(\bigcap_{\alpha \in I} U_\alpha \right)$$

for all $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$. ■

Definition A.3.2. Let $(X, \mathcal{T})$ be a topological space and let $Y \subseteq X$ be non-empty. The *subspace topology on* Y is the topology

$$\mathcal{T}_Y = \{A \cap U \mid U \in \mathcal{T}\}.$$

In addition, the pair $(Y, \mathcal{T}_Y)$ is called a *subspace* of $(X, \mathcal{T})$.

Remark A.3.3. The subspace topology is very useful when one only wants to consider a portion of a topological space. For example, we often want to consider subspaces of $\mathbb{R}$ such as $Y = [0, 1]$ for analytical reasons. However, one should be careful as open subsets of Y need not be open subset of $\mathbb{R}$. Indeed since $[0, 1) = [0, 1] \cap (-1, 1)$ we see that $[0, 1)$ is an open subset of Y in the subspace topology but is not an open subset of $\mathbb{R}$ as $0 \in [0, 1)$ yet no open interval centred at 0 is contained in $[0, 1)$. Thus we really do need to specify the topology and space we are looking at when talking about open sets!

Of course, it is not surprising that a basis for a topological space yields a basis for any subspace.

Proposition A.3.4. *Let $(X, \mathcal{T})$ be a topological space, let $\mathcal{B}$ be a basis for $(X, \mathcal{T})$, and let $Y \subseteq X$ be non-empty. Then*

$$\mathcal{B}_Y = \{Y \cap B \mid B \in \mathcal{B}\}$$

is a basis for $(Y, \mathcal{T}_Y)$.

Proof. To prove this result, we will verify Proposition A.2.12. As such, first notice that $\mathcal{B}_Y \subseteq \mathcal{T}_Y$ by the definition of the subspace topology. Next, let $U \in \mathcal{T}_Y$ and $x \in U$ be arbitrary. Thus $x \in Y$ and, by the definition of the subspace topology, there exists a $V \in \mathcal{T}$ such that $U = Y \cap V$. Since $x \in U$ we see that $x \in V$. Therefore, as $\mathcal{B}$ is a basis for $(X, \mathcal{T})$, Remark A.2.6 implies that there exists a $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq V$. Therefore $Y \cap B \in \mathcal{B}_Y$, $x \in Y \cap B$, and $Y \cap B \subseteq Y \cap V = U$. Thus, as $U \in \mathcal{T}_Y$ and $x \in U$ were arbitrary, $\mathcal{B}_Y$ is a basis for $(T, \mathcal{T}_Y)$ by Proposition A.2.12. ∎

Unsurprisingly, we can use Proposition A.3.4 along with our knowledge of open balls in metric spaces to understand subspaces of metric spaces.

Proposition A.3.5. *Let (X, d) be a metric space and let $Y \subseteq X$ be non-empty. Then the subspace topology on Y is induced by the metric $d_Y : Y \times Y \to [0, \infty)$ defined by*

$$d_Y(y_1, y_2) = d(y_1, y_2)$$

for all $y_1, y_2 \in Y$.

Proof. Since d is a metric and restricting the domain of d will yield a metric, d_Y is a metric. Notice for all $y \in Y$ and $r > 0$ that

$$B_{d_Y}(y, r) = Y \cap B_d(y, r).$$

Therefore, since $\{Y \cap B_d(y, r) \mid y \in Y, r > 0\}$ is a basis for the subspace topology on Y induced by (X, d) by Proposition A.3.4, since $\{B_{d_Y}(y, r) \mid y \in Y, r > 0\}$ is a basis for (Y, d_Y) by definition, and since each basis completely determines the topology by Remark A.2.6, the result follows. ∎

Furthermore, a subspace of a subspace is a subspace. More accurately put, we have the following.

Proposition A.3.6. *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq B \subseteq X$ be arbitrary non-empty sets. Let $\mathcal{T}_B$ be the subspace topology on B inherited from $(X, \mathcal{T})$, let $\mathcal{T}_A$ be the subspace topology on A inherited from $(X, \mathcal{T}_X)$, and let $\mathcal{T}_{A,B}$ be the subspace topology on A inherited from $(B, \mathcal{T}_B)$. Then $\mathcal{T}_{A,B} = \mathcal{T}_A$.*

Proof. By definitions and since $B \cap A = A$, we have that

$$\begin{aligned}\mathcal{T}_{A,B} &= \{A \cap U \mid U \in \mathcal{T}_B\} \\ &= \{A \cap (B \cap V) \mid V \in \mathcal{T}_X\} \\ &= \{A \cap V \mid V \in \mathcal{T}_X\} = \mathcal{T}_A\end{aligned}$$

as desired. ∎

Since subspaces create a topology on a smaller set from a topology on a larger set, it is useful to think of the opposite; that is, can we construct topologies on larger sets from topologies on smaller sets? The simplest way to construct a larger set from two sets is to take their product. Unsurprisingly perhaps, taking the product of the topologies then yields a (basis for a) topology.

Proposition A.3.7. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces. Then*

$$\mathcal{B}_\times = \{U \times V \mid U \in \mathcal{T}, V \in \mathcal{T}'\} \subseteq \mathcal{P}(X \times Y)$$

is a basis for a topology on $X \times Y$.

Proof. To see that $\mathcal{B}_\times$ is a basis for a topology on $X \times Y$, we simply verify Definition A.2.2. Indeed clearly $\mathcal{B}_\times \subseteq \mathcal{P}(X \times Y)$. Moreover, notice for all $x \in X$ and $y \in Y$ that $x \times y \in X \times Y$ and $X \times Y \in \mathcal{B}_\times$. Hence the first condition of Definition A.2.2 holds.

To see the second condition, let $x \times y \in X \times Y$ and $B_1, B_2 \in \mathcal{B}_\times$ such that $x \times y \in B_1 \cap B_2$ be arbitrary. By the definition of $\mathcal{B}_\times$ there exists $U_1, U_2 \in \mathcal{T}$ and $V_1, V_2 \in \mathcal{T}'$ such that $B_1 = U_1 \times V_1$ and $B_2 = U_2 \times V_2$. Since $U_1 \cap U_2 \in \mathcal{T}$ and $V_1 \cap V_2 \in \mathcal{T}'$ as $\mathcal{T}$ and $\mathcal{T}'$ are topologies, and since $B_1 \cap B_2 = (U_1 \cap U_2) \times (V_1 \cap V_2)$, we see that $B_1 \cap B_2 \in \mathcal{B}_\times$ so we may take $B_3 = B_1 \cap B_2$ in Definition A.2.2. Therefore, as $x \times y \in X \times Y$ and $B_1, B_2 \in \mathcal{B}_\times$ were arbitrary, $\mathcal{B}_\times$ is a basis for a topology on $X \times Y$. ∎

Since we are taking the set of Cartesian products of the two topologies to form a topology on the product, this topology has an unsurprising name.

Definition A.3.8. Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces. The *product topology* is the topology generated by the basis

$$\{U \times V \mid U \in \mathcal{T}, V \in \mathcal{T}'\} \subseteq \mathcal{P}(X \times Y).$$

Of course, the product of $\mathbb{R}$ with itself yields a topology on $\mathbb{R}^2$ that we have seen before.

Example A.3.9. As $\mathbb{K}^2 = \mathbb{K} \times \mathbb{K}$, we can consider the product topology on $\mathbb{K}^2$ where each copy of $\mathbb{K}$ is equipped with the canonical topology. In this case, we know a basis of $\mathbb{K} \times \mathbb{K}$ consists of open sets of the form

$$U_1 \times U_2$$

where U_1 and U_2 are open subset of $\mathbb{K}$ with respect to the canonical topology. As each point in each such product contains a $\|\cdot\|_\infty$-ball in the product, and as each $\|\cdot\|_\infty$-ball is such a product, we easily obtain that the product topology on $\mathbb{K}^n$ is the same as the metric topologies by Theorem A.2.14.

Perhaps unsurprisingly, we can repeat the proof of Proposition A.3.7 to simplify the basis for the product topology.

Proposition A.3.10. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces with bases $\mathcal{B}$ and $\mathcal{B}'$ respectively. Then the set*

$$\mathcal{B}_\times = \{B \times B' \mid B \in \mathcal{B}, B' \in \mathcal{B}'\}$$

is a basis for the product topology on $X \times Y$.

Proof. To see that $\mathcal{B}_\times$ is a basis for a topology on $X \times Y$, we will apply Proposition A.2.12. To see this, let U be an arbitrary open subset of $X \times Y$ with respect to the product topology and let $x \times y \in U$ be arbitrary. By the definition of the product topology (Definition A.3.8) there exists sets $U_X \in \mathcal{T}$ and $U_Y \in \mathcal{T}'$ such that $x \times y \in U_X \times U_Y$ and $U_X \times U_Y \subseteq U$. Thus $x \in U_X$ and $y \in U_Y$. Since $\mathcal{B}$ and $\mathcal{B}'$ are bases for $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ respectively, there exists $B \in \mathcal{B}$ and $B' \in \mathcal{B}'$ such that $x \in B$, $y \in B'$, $B \subseteq U_X$ and $B' \subseteq U_Y$. Hence $x \times y \in B \times B'$ and $B \times B' \subseteq U_X \times U_Y \subseteq Y$. Therefore, as U and $x \times y$ were arbitrary, $\mathcal{B}_\times$ is a basis for the product topology on $X \times Y$ by Proposition A.2.12. ∎

Alternatively, we can consider the product topology via a subbasis.

Proposition A.3.11. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces. Then*

$$\mathcal{S} = \{U \times Y \mid U \in \mathcal{T}\} \cup \{X \times V \mid V \in \mathcal{T}'\}$$

is a subbasis for the product topology on $X \times Y$.

Proof. Since finite intersections of elements of $\mathcal{S}$ yields the set

$$\mathcal{B}_\times = \{U \times V \mid U \in \mathcal{T}, V \in \mathcal{T}'\}$$

as $\mathcal{T}$ and $\mathcal{T}'$ are topologies and thus closed under finite intersections, and since $\mathcal{B}_\times$ is a basis for the product topology on $X \times Y$ by Definition A.3.8, the result follows. ∎

It turns out that the subbasis approach to the above product topologies is far superior in functional analysis (and topology) when generalized to infinite products.

Definition A.3.12. Let I be a non-empty set and let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces. The *product topology* on $\prod_{\alpha\in I} X_\alpha$ is the topology generated by the subbasis

$$\mathcal{S} = \{\mathcal{S}_\beta \mid \beta \in I\}$$

where

$$\mathcal{S}_\beta = \left\{ \prod_{\alpha\in I} Y_\alpha \,\middle|\, Y_\alpha = X_\alpha \text{ if } \alpha \neq \beta, Y_\beta \in \mathcal{T}_\beta \right\}.$$

Of course, we should note that the set $\mathcal{S}$ described in Definition A.3.12 is actually a subbasis for some topology on $\prod_{\alpha\in I} X_\alpha$, but this simply follows from Theorem A.2.16. Furthermore, defining the subbasis immediately tells us a basis for the product topology.

Corollary A.3.13. *Let I be a non-empty set and let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces. The product topology on $\prod_{\alpha\in I} X_\alpha$ has as a basis the set of all sets of the form $\prod_{\alpha\in I} U_\alpha$ where $U_\alpha \in \mathcal{T}_\alpha$ and $U_\alpha = X_\alpha$ for all but a finite number of $\alpha \in I$.*

Proof. As the set of all finite intersections of the subbasis for the product topology described in Definition A.3.12 is exactly the sets described here as $\mathcal{T}_\alpha$ is closed under finite intersections for all $\alpha \in I$, the result follows by the definition of a subbasis (Definition A.2.15). ∎

Corollary A.3.14. *Let I be a non-empty set, let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces, and for each $\alpha \in I$ let $\mathcal{B}_\alpha$ be a basis for $(X_\alpha, \mathcal{T}_\alpha)$. Then the set of all sets of the form $\prod_{\alpha\in I} B_\alpha$ where $B_\alpha = X_\alpha$ for all but a finite number of $\alpha \in I$ and $B_\alpha \in \mathcal{B}_\alpha$ for all remaining indices is a basis for the product topology on $\prod_{\alpha\in I} X_\alpha$.*

Proof. Let $\mathcal{B}$ be the set described in the statement. To see that $\mathcal{B}$ is a basis for the product topology on $\prod_{\alpha\in I} X_\alpha$, we simply verify Definition A.2.2. Indeed clearly $\mathcal{B} \subseteq \mathcal{P}(\prod_{\alpha\in I} X_\alpha)$. Moreover, notice for all $(x_\alpha)_{\alpha\in I} \in \prod_{\alpha\in I} X_\alpha$ that $(x_\alpha)_{\alpha\in I} \in \prod_{\alpha\in I} X_\alpha \in \mathcal{B}$. Hence the first condition of Definition A.2.2 holds.

To see the second condition, let $(x_\alpha)_{\alpha\in I} \in \prod_{\alpha\in I} X_\alpha$ and $B_1, B_2 \in \mathcal{B}$ such that $(x_\alpha)_{\alpha\in I} \in B_1 \cap B_2$ be arbitrary. By the definition of $\mathcal{B}$ there exists $B_{1,\alpha}, B_{2,\alpha} \in \mathcal{B}_\alpha$ for all $\alpha \in I$ such that $B_1 = \prod_{\alpha\in I} B_{1,\alpha}$, $B_2 = \prod_{\alpha\in I} B_{2,\alpha}$, and only a finite number of $B_{1,\alpha}$ and $B_{2,\alpha}$ are not equal to X_α over all $\alpha \in I$. Thus, as $(x_\alpha)_{\alpha\in I} \in B_1 \cap B_2$, we see that $x_\alpha \in B_{1,\alpha} \cap B_{2,\alpha}$ for all $\alpha \in I$. If $B_{1,\alpha} = X_\alpha$ or $B_{2,\alpha} = X_\alpha$, let $B_{3,\alpha} = B_{1,\alpha} \cap B_{2,\alpha}$ so that either $B_{3,\alpha} = X_\alpha$ or $B_{3,\alpha} \in \mathcal{B}_\alpha$. Otherwise $B_{1,\alpha}, B_{2,\alpha} \in \mathcal{B}_\alpha$ so, since $\mathcal{B}_\alpha$ is a basis for $(X_\alpha, \mathcal{T}_\alpha)$, there exists a $B_{3,\alpha} \in \mathcal{B}_\alpha$ such that $x_\alpha \in B_{3,\alpha}$ and $B_{3,\alpha} \subseteq B_{1,\alpha} \cap B_{2,\alpha}$ for all $\alpha \in I$. Hence $B_3 = \prod_{\alpha\in I} B_{3,\alpha} \in \mathcal{B}$ as $B_\alpha \in \mathcal{B}_\alpha$ for all but a finite number of $\alpha \in I$ and $B_\alpha = X_\alpha$ for all remaining indices, $(x_\alpha)_{\alpha\in I} \in B_3$, and $B_3 \subseteq B_1 \cap B_2$. Therefore, as $(x_\alpha)_{\alpha\in I} \in \prod_{\alpha\in I} X_\alpha$ and $B_1, B_2 \in \mathcal{B}$ were arbitrary, $\mathcal{B}$ is a basis for the product topology on $\prod_{\alpha\in I} X_\alpha$. ∎

Example A.3.15. Let $n \in \mathbb{N}$ be arbitrary. Then $\mathbb{K}^n = \prod_{k\in\{1,\ldots,n\}} \mathbb{K}$ so we can consider the product topology on $\mathbb{K}^n$. In this case, we know from Corollary A.3.14 that a basis for the product topology on $\mathbb{K}^n$ is

$$I_1 \times I_2 \times \cdots \times I_n$$

where each I_k is an open ball in $\mathbb{K}$ with respect to the absolute value. As each point in each such product is contained in a $\|\cdot\|_\infty$-ball that is contained in the product, and as each $\|\cdot\|_\infty$-ball is such a product, we easily obtain that the product topology on $\mathbb{K}^n$ is the same as the metric topologies by Theorem A.2.14.

A.4 Nets and Limits

Now that we have seen several topologies and how to study them, we return to the notion that the open sets should yield some information about how close points are in topological space. In particular, we can ask what it means for a collection of points to get 'closer and closer' to a given point in a topological space.

In metric spaces, the answer is the well-known concept of convergent sequences. In particular, the ϵ-N notion of a limit of a sequence of real numbers easily generalizes to metric spaces. However, such considerations are insufficient for topologies due to the absent of a total ordering on a basis of open sets centred at each point.

Thus, in order to have a similar notion of convergence in an arbitrary topological space that is sufficient to deduce properties of the space, we need to generalize the notion of a sequence. To do this, we first need to generalize the structure and ordering on the natural numbers.

Definition A.4.1. A *directed set* is a pair $(\Lambda, \leq)$ where Λ is a non-empty set and $\leq$ is a relation on Λ such that

(1) (reflexivity) $\lambda \leq \lambda$ for all $\lambda \in \Lambda$,

(2) (transitivity) if $\lambda_1, \lambda_2, \lambda_3 \in \Lambda$ are such that $\lambda_1 \leq \lambda_2$ and $\lambda_2 \leq \lambda_3$, then $\lambda_1 \leq \lambda_3$, and

(3) (existence of upper bounds) if $\lambda_1, \lambda_2 \in \Lambda$, then there exists a $\lambda_3 \in \Lambda$ such that $\lambda_1 \leq \lambda_3$ and $\lambda_2 \leq \lambda_3$.

The relation $\leq$ is sometimes called the *direction* of Λ.

As we are generalizing the order structure of the natural numbers, our first example is no surprise.

Example A.4.2. The pair $(\mathbb{N}, \leq)$ where $\leq$ is the natural ordering on the natural numbers is easily seen to be a directed set.

Example A.4.3. The pair $(\mathbb{R}, \leq)$ where $\leq$ is the natural ordering on the real numbers is easily seen to be a directed set.

Example A.4.4. Let X be any non-empty set and let $\mathcal{F} \subseteq \mathcal{P}(X)$ be non-empty and closed under finite unions. For two sets $A, B \in \mathcal{F}$, we define $A \leq B$ if and only if $A \subseteq B$. Then $(\mathcal{F}, \leq)$ is a directed set. Indeed it is clear that $\leq$ is reflexive and transitive. Furthermore, if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$ has the property that $A \subseteq A \cup B$ so $A \leq A \cup B$, and $B \subseteq A \cup B$ so $B \leq A \cup B$. Hence $(\mathcal{F}, \leq)$ is a directed set by Definition A.4.1.

Example A.4.5. Let X be any non-empty set and let $\mathcal{F} \subseteq \mathcal{P}(X)$ be non-empty and closed under finite intersections. For two sets $A, B \in \mathcal{F}$, we define $A \leq B$ if and only if $B \subseteq A$. Then $(\mathcal{F}, \leq)$ is a directed set. Indeed it is clear that $\leq$ is reflexive and transitive. Furthermore, if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$ has the property that $A \cap B \subseteq A$ so $A \leq A \cap B$, and $A \cap B \subseteq B$ so $B \leq A \cap B$. Hence $(\mathcal{F}, \leq)$ is a directed set by Definition A.4.1.

Of course, there are many more directed sets. For notational convenience, instead of writing $(\Lambda, \leq)$ for a direct set, we will often just say that Λ is a directed set provided there is no ambiguity for the direction relation which will then be denoted by $\leq$.

With the generalization of the ordering on $\mathbb{N}$, we can describe a generalization of the notion of a sequence.

Definition A.4.6. A *net* is a function $F : \Lambda \to X$ where Λ is a direct set and X is a non-empty set. For notational convenience, we will use $(x_\lambda)_{\lambda \in \Lambda}$ to denote the net $F : \Lambda \to X$ where $F(\lambda) = x_\lambda$.

There are many examples of nets, some of which we are quite familiar with.

Example A.4.7. Every sequence is a net. Indeed a sequence $(x_n)_{n \geq 1}$ can be realized as a net by taking the directed set $(\mathbb{N}, \leq)$ (where $\leq$ is the usual ordering of the natural numbers) and defining F on $\mathbb{N}$ by $F(n) = x_n$.

Example A.4.8. Consider a closed interval $[a, b]$ and the collection $\mathcal{P}$ of all finite partitions of $[a, b]$; that is, all finite subsets $P = \{t_k\}_{k=0}^n \subseteq [a, b]$ such that

$$a = t_0 < t_1 < \cdots < t_n = b.$$

For two sets $P_1, P_2 \in \mathcal{P}$, if we define $P_1 \leq P_2$ if and only if $P_1 \subseteq P_2$, then $(\mathcal{P}, \leq)$ is a directed set by Example A.4.4 as the collection of all finite partitions is closed under finite unions.

Let $f : [a, b] \to \mathbb{R}$ be a function. For each partition $P = \{t_k\}_{k=0}^n \in \mathcal{P}$ and each $1 \leq k \leq n$, choose $c_k \in [t_{k-1}, t_k]$ and define

$$S_P = \sum_{k=1}^{n} f(c_k)(t_k - t_{k-1}).$$

Then $(S_P)_{P\in\mathcal{P}}$ is a net of Riemann sums.

The next example is motivated by trying to take sums over uncountable sets.

Example A.4.9. Let I be any infinite (and not necessarily countable) set. Let $\mathcal{F}$ be the set of all finite (non-empty) subsets of I. For two sets $F_1, F_2 \in \mathcal{F}$, if we define $F_1 \leq F_2$ if and only if $F_1 \subseteq F_2$, then $(\mathcal{F}, \leq)$ is a directed set by Example A.4.4 as finite unions of finite sets are finite.

For each $\alpha \in I$, let $x_\alpha \in \mathbb{R}$ be non-negative. For each $F \in \mathcal{F}$, define

$$S_F = \sum_{\alpha\in F} x_\alpha,$$

which is well-defined as F is finite. Then $(S_F)_{F\in\mathcal{F}}$ is a net of all finite sums of $\{x_\alpha \mid \alpha \in I\}$.

Of course, our interest does not stem from the existence of nets as generalizations of sequences, but the properties and results that the notion of the convergence of a net will yield. Thus, building on the idea of using open sets to describe convergence in metric spaces, we generalize the notion of a convergent sequence for nets in arbitrary topological spaces.

Definition A.4.10. Let $(X, \mathcal{T})$ be a topological space. A net $(x_\lambda)_{\lambda\in\Lambda}$ in X is said to *converge* to a point $x_0 \in X$ (or equivalently, x_0 is *a limit* of $(x_\lambda)_{\lambda\in\Lambda}$) if for every $U \in \mathcal{T}$ such that $x_0 \in U$ there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in U$ for all $\lambda \geq \lambda_0$.

Before we get to examples, the notion of taking a set from the topology containing a specified point is occurring in greater and greater frequency. Thus, at this point, it is about time we gave it a name.

Definition A.4.11. Let $(X, \mathcal{T})$ be a topological space. A subset $U \subseteq X$ is said to be a *neighbourhood of a point* $x \in X$ if $x \in U$ and $U \in \mathcal{T}$.

Remark A.4.12. The term 'neighbourhood' comes from the notion that an open set containing a point x contains all points that are 'geographically' close to x. However, one must be careful with the term 'neighbourhood' in topology as many authors do not require a neighbourhood of a point to be open; they just require that a neighbourhood contains an open set containing the specified point. As we will want to be working with mainly open sets in this course, our definition is preferable.

Now onto examples. Of course, this is nowhere near an exhaustive list.

Example A.4.13. It is clear that a sequence in a metric space converges to a point as a net if and only if it converges as a sequence to the same point.

Example A.4.14. Consider the net $(S_P)_{P\in\mathcal{P}}$ from Example A.4.8. If f is Riemann integrable, then $(S_P)_{P\in\mathcal{P}}$ converges and converges to $\int_a^b f(x)\,dx$. Indeed suppose f is integrable and let U neighbourhood of $\int_a^b f(x)\,dx$. Hence there exists an $\epsilon > 0$ such that

$$\left(\int_a^b f(x)\,dx - \epsilon, \int_a^b f(x)\,dx + \epsilon\right) \subseteq U.$$

By the definition of the Riemann integral, there exists a partition $P_0 \in \mathcal{P}$ such that if $U(f, P_0)$ is the upper Riemann sum of f corresponding to P_0 and $L(f, P_0)$ is the lower Riemann sum of f corresponding to P_0, then

$$L(f, P_0) \leq \int_a^b f(x)\,dx \leq U(f, P_0) < L(f, P_0) + \epsilon.$$

If $P \in \mathcal{P}$ and $P \geq P_0$, then P is a refinement of P_0 so

$$L(f, P_0) \leq L(f, P) \leq S_P \leq U(f, P) \leq U(f, P_0).$$

Hence

$$S_P \in \left(\int_a^b f(x)\,dx - \epsilon, \int_a^b f(x)\,dx + \epsilon\right) \subseteq U.$$

Therefore, as U was arbitrary, $(S_P)_{P\in\mathcal{P}}$ converges to $\int_a^b f(x)\,dx$.

Somewhat conversely, if every net from Example A.4.8 converges and converges to the same number, then f is Riemann integrable. In fact, it is only required that the net of upper Riemann sum $(U_P)_{P\in P}$ and the net of lower Riemann sums $(L_P)_{P\in P}$ converge to the same number I. To see this, suppose $(U_P)_{P\in P}$ and $(L_P)_{P\in P}$ both converge to I and let $\epsilon > 0$ be arbitrary. Since $(U_P)_{P\in P}$ converges to I, there exists a $P_1 \in \mathcal{P}$ such that

$$U_P \in \left(I - \frac{\epsilon}{2}, I + \frac{\epsilon}{2}\right)$$

for all $P \geq P_1$. Similarly, since $(L_P)_{P\in P}$ converges to I, there exists a $P_2 \in \mathcal{P}$ such that

$$L_P \in \left(I - \frac{\epsilon}{2}, I + \frac{\epsilon}{2}\right)$$

for all $P \geq P_2$. Thus, if $P_0 = P_1 \cup P_2$, then $P_0 \in \mathcal{P}$, $P_0 \geq P_1$ and $P_0 \geq P_2$ so

$$U_{P_0}, L_{P_0} \in \left(I - \frac{\epsilon}{2}, I + \frac{\epsilon}{2}\right).$$

Hence, as $L_{P_0} \leq U_{P_0}$, we obtain that $U_{P_0} - L_{P_0} < \epsilon$. Therefore, as $\epsilon > 0$ was arbitrary, f is Riemann integrable.

Example A.4.15. Consider the net $(S_F)_{F\in\mathcal{F}}$ from Example A.4.9. Then $(S_F)_{F\in\mathcal{F}}$ converges if and only if

$$L = \sup\{S_F \mid F \in \mathcal{F}\}$$

is finite, in which case $(S_F)_{F\in\mathcal{F}}$ converges to L. Indeed suppose L is finite and let U be a neighbourhood of L. Then there exists an $\epsilon > 0$ such that

$$(L-\epsilon, L+\epsilon) \subseteq U.$$

By the definition of the supremum, there exists an $F_0 \in \mathcal{F}$ such that

$$L - \epsilon < S_{F_0} \leq L.$$

Hence, as $x_\alpha \geq 0$ for all $\alpha \in I$, we see that for all $F \in \mathcal{F}$ with $F \geq F_0$ that

$$L - \epsilon < S_{F_0} \leq S_F \leq L.$$

Hence $S_F \in U$ for all $F \geq F_0$. Therefore, as U was arbitrary, $(S_F)_{F\in\mathcal{F}}$ converges to L.

Conversely suppose that $L = \infty$. Hence for any $M \in \mathbb{R}$ there exists an $F_M \in \mathcal{F}$ such that $S_{F_M} \geq M$. To proceed by contradiction, suppose $(S_F)_{F\in\mathcal{F}}$ converges to some point $K \in \mathbb{R}$. Then there exists an $F_0 \in \mathcal{F}$ such that $S_F \in (K-1, K+1)$ for all $F \geq F_0$. Hence, as $F_0 \cup F_{K+1} \in \mathcal{F}$ and $F_0 \cup F_{K+1} \geq F_0$, we must have that

$$S_{F_0 \cup F_{K+1}} \in (K-1, K+1).$$

However

$$S_{F_0 \cup F_{K+1}} \geq S_{F_{K+1}} \geq K+1$$

as $x_\alpha \geq 0$ for all $\alpha \in I$ so $S_{F_0 \cup F_{K+1}} \notin (K-1, K+1)$. Hence we have a contradiction as desired.

The above is quite useful in summing over uncountable sets. In particular, we define the *sum of* $\{x_\alpha \mid \alpha \in I\}$, denoted $\sum_{\alpha\in I} x_\alpha$, to be

$$\sum_{\alpha \in I} x_\alpha = \sup\{S_F \mid F \in \mathcal{F}\} \in [0, \infty].$$

Furthermore, if $\sum_{\alpha\in I} x_\alpha < \infty$ then for all $n \in \mathbb{N}$ we must have that $F_n = \left\{\alpha \in I \mid x_\alpha \geq \frac{1}{n}\right\}$ is finite for otherwise for each $m \in \mathbb{N}$ we can find a finite subset $F_{n,m} \subseteq F_n$ with m elements so that $S_{F_{n,m}} \geq \frac{m}{n}$ thereby yielding $\sum_{\alpha\in I} x_\alpha = \infty$. Therefore if $\sum_{\alpha\in I} x_\alpha < \infty$ then, each F_n is a finite set so

$$\bigcup_{n\geq 1} F_n = \{\alpha \in I \mid x_\alpha > 0\},$$

is countable. Thus, after removing all x_α that take the value 0, we can simply add a countable sum of non-negative numbers to determine the value of $\sum_{\alpha\in I} x_\alpha$.

Of course, as bases determine a topology, we need only check neighbourhoods of a point that come from a basis.

Lemma A.4.16. *Let $(X, \mathcal{T})$ be a topological space and let $\mathcal{B}$ be a basis for $(X, \mathcal{T})$. A net $(x_\lambda)_{\lambda\in\Lambda}$ in X converges to $x_0 \in X$ if and only if for every $B \in \mathcal{B}$ such that $x_0 \in B$ there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in B$ for all $\lambda \geq \lambda_0$.*

Proof. If $(x_\lambda)_{\lambda\in\Lambda}$ converges to x_0, then Definition A.4.10 implies that for every $B \in \mathcal{B}$ such that $x_0 \in B$ there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in B$ for all $\lambda \geq \lambda_0$ since $\mathcal{B} \subseteq \mathcal{T}$.

Conversely, suppose for every $B \in \mathcal{B}$ such that $x_0 \in B$ there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in B$ for all $\lambda \geq \lambda_0$. To see that $(x_\lambda)_{\lambda\in\Lambda}$ converges to x_0, let U be an arbitrary neighbourhood of x_0. Then, as $\mathcal{B}$ is a basis for $(X, \mathcal{T})$, there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U$. Thus, by assumption, a $\lambda_0 \in \Lambda$ such that $x_\lambda \in B \subseteq$ for all $\lambda \geq \lambda_0$. Therefore, as U was arbitrary, the proof is complete. ∎

Of course, in Lemma A.4.16, we need only information about the neighbourhoods of x_0. Consequently, we do not need to consider a basis for the entire space. In particular, we need only consider the following.

Definition A.4.17. Let $(X, \mathcal{T})$ be a topological space and let $x \in X$. A set $\mathcal{B} \subseteq \mathcal{T}$ is said to be a *neighbourhood basis of* x if $x \in B$ for all $B \in \mathcal{B}$ and for all neighbourhoods U of x there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U$.

Theorem A.4.18. *Let $(X, \mathcal{T})$ be a topological space, let $x_0 \in X$, and let $\mathcal{B}$ be a neighbourhood basis for x_0. A net $(x_\lambda)_{\lambda\in\Lambda}$ in X converges to x_0 if and only if for every $B \in \mathcal{B}$ such that $x_0 \in B$ there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in B$ for all $\lambda \geq \lambda_0$.*

Proof. The proof of this result is identical to the proof of Lemma A.4.16. ∎

Unsurprisingly, we can construct a basis from neighbourhood bases.

Proposition A.4.19. *Let $(X, \mathcal{T})$ be a topological space and for each $x \in X$ let $\mathcal{B}_x$ be a neighbourhood basis for x. Then $\bigcup_{x\in X} \mathcal{B}_x$ is a basis for $(X, \mathcal{T})$.*

Proof. This follows immediately from the definition of a neighbourhood basis and Proposition A.2.12. ∎

Returning to the notion of convergent nets, we can easily use bases to describe convergence in the lower limit, subspace, and product topologies. In particular, the following is the reason the lower limit topology has its name.

Proposition A.4.20. *Let $\mathcal{T}_L$ be the lower limit topology on $\mathbb{R}$. A net $(x_\lambda)_{\lambda\in\Lambda}$ in $\mathbb{R}$ converges to a point x in $(\mathbb{R}, \mathcal{T}_L)$ if and only if for every $\epsilon > 0$ there exists an $\lambda_0 \in \Lambda$ such that $x \leq x_\lambda < x + \epsilon$ for all $\lambda \geq \lambda_0$.*

Proof. To begin, suppose a net $(x_\lambda)_{\lambda\in\Lambda}$ in $\mathbb{R}$ converges to a point x in $(\mathbb{R}, \mathcal{T}_L)$. To see the result, let $\epsilon > 0$ be arbitrary. Since $[x, x+\epsilon)$ is a neighbourhood of x, the definition of a convergent net implies there exists an $\lambda_0 \in \Lambda$ such that $x_\lambda \in [x, x+\epsilon)$ (that is, $x \leq x_\lambda < x+\epsilon$) for all $\lambda \geq \lambda_0$. Therefore, as $\epsilon > 0$ was arbitrary, the result holds.

Conversely, suppose $(x_\lambda)_{\lambda\in\Lambda}$ is an net in $\mathbb{R}$ and $x \in \mathbb{R}$ are such that for every $\epsilon > 0$ there exists an $\lambda_0 \in \Lambda$ such that $x \leq x_\lambda < x+\epsilon$ for all $\lambda \geq \lambda_0$. To see that $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(\mathbb{R}, \mathcal{T}_L)$, let $B = [a, b) \in \mathcal{B}$ be an arbitrary element such that $x \in B$. Hence $a \leq x$ and $x < b$ so there exists an $\epsilon > 0$ such that $x < x+\epsilon < b$. Therefore, by the assumptions on $(x_\lambda)_{\lambda\in\Lambda}$, there exists a $\lambda_0 \in \Lambda$ such that $x \leq x_\lambda < x+\epsilon$ for all $\lambda \geq \lambda_0$. Hence

$$x_\lambda \in [x, x+\epsilon) \subseteq [a, b) = B \in \mathcal{B}.$$

Therefore, as $B \in \mathcal{B}$ was arbitrary, $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(\mathbb{R}, \mathcal{T}_L)$ as desired. ∎

Proposition A.4.21. *Let $(X, \mathcal{T})$ be a topological space, let $A \subseteq X$ be non-empty, let $\mathcal{T}_A$ be the subspace topology on A, let $(a_\lambda)_{\lambda\in\Lambda}$ be a net in A, and let $a \in A$. Then $(a_\lambda)_{\lambda\in\Lambda}$ converges to a in $(A, \mathcal{T}_A)$ if and only if $(a_\lambda)_{\lambda\in\Lambda}$ converges to a in $(X, \mathcal{T})$*

Proof. Since $a_\lambda \in A$ for all $\lambda \in A$, the result follows immediately by Definition A.4.10 as the neighbourhoods of a in $(A, \mathcal{T}_A)$ are precisely the neighbourhoods of a in $(X, \mathcal{T})$ intersected with A. ∎

Theorem A.4.22. *Let I be a non-empty set, let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces, let $(f_\lambda)_{\lambda\in\Lambda}$ be a net in $\prod_{\alpha\in I} X_\alpha$, and let $(x_\alpha)_{\alpha\in I} \in \prod_{\alpha\in I} X_\alpha$. Then $(f_\lambda)_{\lambda\in\Lambda}$ converges to $(x_\alpha)_{\alpha\in I}$ when $\prod_{\alpha\in I} X_\alpha$ is equipped with the product topology if and only if $(f_\lambda(\alpha))_{\lambda\in\Lambda}$ converges to x_α in $(X_\alpha, \mathcal{T}_\alpha)$ for all $\alpha \in I$.*

Proof. Suppose that $(f_\lambda)_{\lambda\in\Lambda}$ converges to $(x_\alpha)_{\alpha\in I}$ when $\prod_{\alpha\in I} X_\alpha$ is equipped with the product topology. To see that $(f_\lambda(\alpha))_{\lambda\in\Lambda}$ converges to x_α in $(X_\alpha, \mathcal{T}_\alpha)$ for all $\alpha \in I$, fix $\alpha_0 \in I$ and let U_{α_0} be an arbitrary neighbourhood of x_{α_0} in $(X_{\alpha_0}, \mathcal{T}_{\alpha_0})$. For each $\alpha \in I \setminus \{\alpha_0\}$, let $U_\alpha = X_\alpha$. As $\prod_{\alpha\in I} U_\alpha$ is an element of the subbasis for the product topology on $\prod_{\alpha\in I} X_\alpha$ by Definition A.3.12 and thus open, we easily see that $\prod_{\alpha\in I} U_\alpha$ is a neighbourhood of $(x_\alpha)_{\alpha\in I}$. Therefore, as $(f_\lambda)_{\lambda\in\Lambda}$ converges to $(x_\alpha)_{\alpha\in I}$ when $\prod_{\alpha\in I} X_\alpha$ is equipped with the product topology, there exists a $\lambda_0 \in \Lambda$ such that $f_\lambda \in \prod_{\alpha\in I} U_\alpha$ for all $\lambda \geq \lambda_0$. Hence $f_\lambda(\alpha_0) \in U_{\alpha_0}$ for all $\lambda \geq \lambda_0$. Therefore, as $\alpha_0 \in I$ and U_{α_0} where arbitrary, $(f_\lambda(\alpha))_{\lambda\in\Lambda}$ converges to x_α in $(X_\alpha, \mathcal{T}_\alpha)$ for all $\alpha \in I$.

Conversely, suppose $(f_\lambda(\alpha))_{\lambda\in\Lambda}$ converges to x_α in (X_α, T_α) for all $\alpha \in I$. Recall from Corollary A.3.13 that the product topology on $\prod_{\alpha\in I} X_\alpha$ has as a basis $\mathcal{B}$ consisting of all sets of the form $\prod_{\alpha\in I} U_\alpha$ where $U_\alpha \in \mathcal{T}_\alpha$ and $U_\alpha = X_\alpha$ for all but a finite number of $\alpha \in I$. To see that $(f_\lambda)_{\lambda\in\Lambda}$ converges to $(x_\alpha)_{\alpha\in I}$, let $\prod_{\alpha\in I} U_\alpha$ be an arbitrary element of $\mathcal{B}$ that is a neighbourhood of $(x_\alpha)_{\alpha\in I}$. Hence U_α is a neighbourhood of x_α for all $\alpha \in I$ and

$$\{\alpha \in I \mid U_\alpha \neq X_\alpha\} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$$

for some $n \in \mathbb{N}$. Since for each $k \in \{1, \ldots, n\}$ we know that $(f_\lambda(\alpha_k))_{\lambda\in\Lambda}$ converges to x_{α_k} in $(X_{\alpha_k}, T_{\alpha_k})$, there exists a $\lambda_k \in \Lambda$ such that $f_\lambda(\alpha_k) \in U_{\alpha_k}$ for all $\lambda \geq \lambda_k$. Luckily, by the properties of a direct set, there exists a $\lambda' \in \lambda$ such that $\lambda' \geq \lambda_k$ for all $k \in \{1, \ldots, n\}$. Hence $f_\lambda(\alpha_k) \in U_{\alpha_k}$ for all $\lambda \geq \lambda'$ and for all $k \in \{1, \ldots, n\}$. Since $f_\lambda(\alpha) \in X_\alpha = U_\alpha$ for all $\alpha \in I \setminus \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, we obtain that $(f_\lambda)_{\lambda\in\Lambda} \in \prod_{\alpha\in I} U_\alpha$ for all $\lambda \geq \lambda'$. Therefore, as $\prod_{\alpha\in I} U_\alpha$ was arbitrary, Lemma A.4.16 implies the result. ∎

Convergent nets are enough to completely determine topologies.

Theorem A.4.23. *Let X be a non-empty set and let $\mathcal{T}$ and $\mathcal{T}'$ be two topologies on X. Then $\mathcal{T}$ is finer than $\mathcal{T}'$ if and only if whenever $(x_\lambda)_{\lambda\in\Lambda}$ is a net that converges to x in $(X, \mathcal{T})$, then $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(X, \mathcal{T}')$.*

Consequently, if $(X, \mathcal{T})$ and $(X, \mathcal{T}')$ have exactly the same nets converging to the same points, then $\mathcal{T} = \mathcal{T}'$.

Proof. If $\mathcal{T}$ is finer than $\mathcal{T}'$, then $\mathcal{T}' \subseteq \mathcal{T}$. It is then clear that if $(x_\lambda)_{\lambda\in\Lambda}$ is a net that converges to x in $(X, \mathcal{T})$, then $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(X, \mathcal{T}')$ by the definition of a convergent net.

Conversely, suppose whenever $(x_\lambda)_{\lambda\in\Lambda}$ is a net that converges to x in $(X, \mathcal{T})$, then $(x_\lambda)_{\lambda\in\Lambda}$ converges to x in $(X, \mathcal{T}')$. To proceed by contradiction, suppose there exists a set $U \in \mathcal{T}'$ such that $U \notin \mathcal{T}$. By Theorem A.2.14 there exists an $x_0 \in U$ such that for each $\mathcal{T}$-neighbourhood V of x_0, $V \setminus U$ is non-empty.

Let

$$\Lambda = \{V \subseteq X \mid V \text{ is a } \mathcal{T}\text{-neighourhood of } x_0\}.$$

As Λ is closed under finite intersections, if for $V_1, V_2 \in \Lambda$ we define $V_1 \leq V_2$ if $V_2 \subseteq V_1$, then $(\Lambda, \leq)$ is a direct set by Example A.4.5.

For each $V \in \Lambda$, let $x_V \in V \setminus U$ (note we are using the Axiom of Choice here). We claim that $(x_V)_{V\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$ but does not converge to x_0 in $(X, \mathcal{T}')$ thereby yielding a contradiction. To see that $(x_V)_{V\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$, let V_0 be an arbitrary $\mathcal{T}$-neighbourhood x_0. Then for all $V \geq V_0$ we have that $x_V \in V \subseteq V_0$. Hence $(x_V)_{V\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$ by Definition A.4.10. To see that $(x_V)_{V\in\Lambda}$ does not converge to x_0 in $(X, \mathcal{T}')$, we simply note that U is a $\mathcal{T}'$-neighbourhood of x_0 but $x_V \notin U$ for all $V \in \Lambda$. Hence we have obtained a contradiction thereby finishing the proof. ∎

A clever observer at this point would have likely noticed that we have only defined 'a limit and not 'the limit' of a net when we defined when a net converges to a point. This is because, in a general topological space, a net can converge to multiple points so the 'the' in 'the limit' no longer make sense. This is even true if we consider sequences in general topological spaces as the following example demonstrates.

Example A.4.24. Consider the set $X = \{a, b, c\}$ and the topology

$$\mathcal{T} = \{\emptyset, \{a\}, \{b, c\}, X\}.$$

It is not difficult to see that a sequence $(x_n)_{n\geq 1}$ in X converges to a if and only if there exists an $N \in \mathbb{N}$ such that $x_n = a$ for all $n \geq N$ as $\{a\}$ is a neighbourhood of a. However, $(x_n)_{n\geq 1}$ in X converges to b if and only if there exists an $N \in \mathbb{N}$ such that $x_n \in \{b, c\}$ for all $n \geq N$ as the only open sets containing b are X and $\{b, c\}$. Similarly $(x_n)_{n\geq 1}$ in X converges to c if and only if there exists an $N \in \mathbb{N}$ such that $x_n \in \{b, c\}$ for all $n \geq N$. Thus there are several sequences in X that converge to both b and c.

The reason the above example does not yields unique limits is that there are not enough open sets to distinguish the points. The correct notion in order for there to be unique limits is the following.

Definition A.4.25. A topological space $(X, \mathcal{T})$ is said to be *Hausdorff* (equivalently $(X, \mathcal{T})$ is a *Hausdorff space*) if for all $x, y \in X$ where $x \neq y$ there exists sets $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

Example A.4.26. The trivial topology on a set with at least two points is not Hausdorff as the only open sets are the empty set and the entire set.

Example A.4.27. The discrete topology on any set is Hausdorff as every singleton is an open set.

Example A.4.28. Let X be finite. The only topology on X that is Hausdorff is the discrete topology. Indeed suppose $\mathcal{T}$ is a Hausdorff topology on X and fix a point $x \in X$. For each point $y \in X \setminus \{y\}$ there exists an open set $U_y \in \mathcal{T}$ such that $x \in U_y$ but $y \notin U_y$. Then, as $X \setminus \{x\}$ is finite, we see that

$$\{x\} = \bigcap_{y \in X\setminus\{x\}} U_y \in \mathcal{T}.$$

Therefore, as $x \in X$ was arbitrary, every singleton from X is in $\mathcal{T}$. Therefore, as X is finite, $\mathcal{T}$ must be the discrete topology.

Example A.4.29. The cofinite topology on an infinite set is not Hausdorff as the intersection of any two non-empty open sets in the cofinite topology on an infinite set must contain an infinite number of points. Similarly the cocountable topology on an uncountable set is not Hausdorff as the intersection of any two non-empty open sets in the cocountable topology on an uncountable set must contain an uncountable number of points. However, the cofinite topology on a finite set and the cocountable topology on a countable set are Hausdorff as every singleton is open (and thus the topologies are discrete in this case).

Example A.4.30. The metric topology on a metric space (X, d) is Hausdorff. Indeed given two points $x, y \in X$ with $x \neq y$, let $\delta = \frac{1}{2}d(x, y)$. Then $B_d(x, \delta)$ and $B_d(y, \delta)$ are disjoint open sets one of which contains x and the other of which contains y. Hence the topology is Hausdorff by definition. Consequently, any non-Hausdorff topology is not induced by a metric.

Example A.4.31. The lower limit topology on $\mathbb{R}$ is Hausdorff. To see this, let $a, b \in \mathbb{R}$ be such that $a < b$. Then $U = [a, b)$ and $V = [b, \infty)$ are open sets in the lower limit topology such that $a \in U$, $b \in V$, and $U \cap V = \emptyset$. Thus, as $a, b \in \mathbb{R}$ were arbitrary, the lower limit topology is Hausdorff.

Example A.4.32. A subspace of any Hausdorff space is Hausdorff. This follows directly from the definition of a Hausdorff space and the description of the open subsets in the subspace topology (i.e. the open sets are simple the intersection of open sets with the subspace).

Example A.4.33. The product topologies of Hausdorff spaces are Hausdorff. This follows directly from the description of the open sets in these topologies. To be specific, given two elements of the product $\prod_{\alpha \in I} X_\alpha$ of Hausdorff spaces, they differ at one value of α, say $\alpha_0 \in I$. Thus we can find disjoint open sets in $(X_{\alpha_0}, \mathcal{T}_{\alpha_0})$ that separate these two values and by taking the product of these open sets with X_α for all $\alpha \neq \alpha_0$, the desired open sets separating the two elements of the product have been found.

As advertised, Hausdorff spaces have unique limits.

Theorem A.4.34. *Let $(X, \mathcal{T})$ be a Hausdorff space. If a net $(x_\lambda)_{\lambda \in \Lambda}$ converges to two points $x_1, x_2 \in X$, then $x_1 = x_2$.*

Proof. Suppose to the contrary that there exists a net $(x_\lambda)_{\lambda \in \Lambda}$ that converges to two points $x_1, x_2 \in X$ where $x_1 \neq x_2$. As $(X, \mathcal{T})$ is Hausdorff, there exist $U, V \in \mathcal{T}$ such that $x_1 \in U$, $x_2 \in V$, and $U \cap V = \emptyset$. As $(x_\lambda)_{\lambda \in \Lambda}$ converges to x_1, there exists a $\lambda_1 \in \Lambda$ such that $x_\lambda \in U$ for all $\lambda \geq \lambda_1$. Similarly as $(x_\lambda)_{\lambda \in \Lambda}$ converges to x_2, there exists a $\lambda_2 \in \Lambda$ such that $x_\lambda \in V$ for all $\lambda \geq \lambda_2$. However, by the properties of directed sets, there exists a $\lambda_3 \in \Lambda$ such that $\lambda_1 \leq \lambda_3$ and $\lambda_2 \leq \lambda_3$. Hence the above yields $x_{\lambda_3} \in U \cap V$ which contradicts the fact that $U \cap V = \emptyset$. Hence the result follows. ∎

In particular, for Hausdorff spaces, we can define limits.

Definition A.4.35. Let $(X, \mathcal{T})$ be a Hausdorff space and let $(x_\lambda)_{\lambda \in \Lambda}$ be a net in X that converges in X. The unique point that $(x_\lambda)_{\lambda \in \Lambda}$ converges to in X is called the *limit of* $(x_\lambda)_{\lambda \in \Lambda}$ and is denoted $\lim_{\lambda \in \Lambda} x_\lambda$.

In fact, the only topological spaces that have unique limits for every converging net are Hausdorff spaces.

Theorem A.4.36. *Let $(X, \mathcal{T})$ be a topological space such that every convergent net in $(X, \mathcal{T})$ converges to exactly one point. Then $(X, \mathcal{T})$ is Hausdorff.*

Proof. Let $(X, \mathcal{T})$ be a topological space such that every convergent net in $(X, \mathcal{T})$ converges to exactly one point. Suppose to the contrary that that $(X, \mathcal{T})$ is not Hausdorff. Then there exist points $x, y \in X$ such that for every neighbourhood U of x and neighbourhood V of y, $U \cap V \neq \emptyset$.

Consider the set

$$\Lambda = \{(U, V) \mid U, V \in \mathcal{T} \text{ are such that } x \in U \text{ and } y \in V\}.$$

For $(U_1, V_1), (U_2, V_2) \in V$, we define $(U_1, V_1) \leq (U_2, V_2)$ if and only if $U_2 \subseteq U_1$ and $V_2 \subseteq V_1$. We claim that $(\Lambda, \leq)$ is a directed set. Indeed, clearly $\leq$ is reflexive and transitive. Furthermore, if $(U_1, V_1), (U_2, V_2) \in V$, then by taking $U_3 = U_1 \cap U_2$ and $V_3 = V_1 \cap V_2$, we easily see that $(U_3, V_3) \in \Lambda)$, $(U_1, V_1) \leq (U_3, V_3)$, and $(U_2, V_2) \leq (U_3, V_3)$. Hence $(\Lambda, \leq)$ is a directed set.

For each $(U, V) \in \Lambda$, choose a $z_{(U,V)} \in U \cap V$, which exists by assumption (note we are using the Axiom of Choice here). Hence $(z_{(U,V)})_{(U,V)\in\Lambda}$ is a net. We claim that $(z_{(U,V)})_{(U,V)\in\Lambda}$ converges to both x and y. Indeed if U is an arbitrary neighbourhood of x, then for all $(U', V') \geq (U, X)$ we see that

$$z_{(U',V')} \in U' \cap V' \subseteq U \cap X = U.$$

Hence $(z_{(U,V)})_{(U,V)\in\Lambda}$ converges to x. Similarly, if V is an arbitrary neighbourhood of x, then for all $(U', V') \geq (X, V)$ we see that

$$z_{(U',V')} \in U' \cap V' \subseteq X \cap V = V.$$

Hence $(z_{(U,V)})_{(U,V)\in\Lambda}$ converges to y. As this contradicts the fact that every convergent net in $(X, \mathcal{T})$ converges to exactly one point, the proof is complete. ∎

In general, asking that a space is Hausdorff is a very mild condition in that it is simply asking that we can separate any two distinct points with open sets. However, the fact that nets in Hausdorff spaces have unique limits is very useful for the study of spaces that are Hausdorff. In particular, trying to prove results for arbitrary topological spaces can often be difficult or impossible as they need not have enough structure. Thus, we will often impose conditions like being Hausdorff on certain topological spaces in order to be able to prove certain results, which will then only apply to certain collections of topological spaces.

To finish off this section, we recall one useful tool in undergraduate analysis is the ability to take subsequences. For nets, things are a little more delicate, but will be equally useful.

Definition A.4.37. Let X be a non-empty set, let $(\Lambda, \leq)$ and $(M, \leq_0)$ be two directed sets, and let $F : \Lambda \to X$ be a net. A *subnet* of F directed by $(M, \leq_0)$ is the composition $F \circ \varphi : M \to X$ where $\varphi : M \to \Lambda$ is such that

(1) (increasing) if $\mu_1, \mu_2 \in M$ are such that $\mu_1 \leq_0 \mu_2$, then $\varphi(\mu_1) \leq \varphi(\mu_2)$, and

(2) (cofinal) for each $\lambda \in \Lambda$ there exists a $\mu \in M$ such that $\lambda \leq \varphi(\mu)$.

Remark A.4.38. Note it is elementary to see that a subnet of a net is a net. In particular, if the net F is denoted by $(x_\lambda)_{\lambda\in\Lambda}$, we will often use $(x_{\lambda_\mu})_{\mu\in M}$ to denote a subnet where $\varphi(\mu) = \lambda_\mu$.

Subnets can be a little tricky.

Example A.4.39. A subnet of a sequence need not be a subsequence. Indeed consider the sequence $(x_n)_{n\geq1}$ and consider the directed set $(\mathbb{R}, \leq)$. Then if we define $\varphi : \mathbb{R} \to \mathbb{N}$ by

$$\varphi(x) = \begin{cases} 1 & \text{if } x < 1 \\ n & \text{if } x \in (n-1, n] \end{cases},$$

then φ is increasing and cofinal. However, clearly $(x_{n_\mu})_{\mu\in\mathbb{R}}$ is not a subsequence of $(x_n)_{n\geq1}$.

Example A.4.40. Let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in a topological space. Choose $\lambda_1 \in \Lambda$. Then there exists a $\lambda_2 \in \Lambda$ such that $\lambda_2 \geq \lambda_1$. By repetition, we can obtain a sequence $(\lambda_n)_{n\geq1}$ of elements of Λ that are increasing. However $(x_{\lambda_n})_{n\geq1}$ need not be a subnet of $(x_\lambda)_{\lambda\in\Lambda}$ as it need not be cofinal (even if $\lambda_k \neq \lambda_{k+1}$). Note this is quite different than the situation with sequences.

However, as with sequences, subnets of convergent nets still converge.

Proposition A.4.41. *Let $(X, \mathcal{T})$ be a topological space, let $(x_\lambda)_{\lambda\in\Lambda}$ be a net in X, and let $(x_{\lambda_\mu})_{\mu\in M}$ be a subnet of $(x_\lambda)_{\lambda\in\Lambda}$. If $(x_\lambda)_{\lambda\in\Lambda}$ converges to some point $x \in (X, \mathcal{T})$, then $(x_{\lambda_\mu})_{\mu\in M}$ converges to x in $(X, \mathcal{T})$.*

Proof. Suppose $(x_\lambda)_{\lambda\in\Lambda}$ converges to some point $x \in (X, \mathcal{T})$. To see that $(x_{\lambda_\mu})_{\mu\in M}$ converges to x in $(X, \mathcal{T})$, let U be an arbitrary neighbourhood of x. As $(x_\lambda)_{\lambda\in\Lambda}$ converges to x, there exists an $\lambda_0 \in \Lambda$ such that $x_\lambda \in U$ for all $\lambda \geq \lambda_0$. Due to the properties of subnets, there exists a $\mu_0 \in M$ such that $\lambda_{\mu_0} \geq \lambda_0$. Hence, by the properties of subnets, if $\mu \geq \mu_0$ then $\lambda_\mu \geq \lambda_{\mu_0} \geq \lambda_0$ and thus $x_{\lambda_\mu} \in U$. Therefore, as U was arbitrary, $(x_{\lambda_\mu})_{\mu\in M}$ converges to x in $(X, \mathcal{T})$. ∎

A.5 Sets and Points

With the completion of our basic understanding of nets, we can now turn out attention to types of points and sets inside topological spaces. These various types of points and sets occur regularly throughout topology and will be of incredible use in this course. Most of these notions are generalizations of known types of sets and points in metric spaces. In particular, the following type of sets are well known.

Definition A.5.1. Let $(X, \mathcal{T})$ be a topological space. A set $F \subseteq X$ is said to be *closed* if $X \setminus F \in \mathcal{T}$.

Example A.5.2. Every closed interval $[a, b]$ is a closed subset of $\mathbb{R}$ with its canonical topology as $\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$ is the union of two open sets and thus is open.

Example A.5.3. The set $[0, 1)$ is neither open nor closed when $\mathbb{R}$ is equipped with its canonical topology. Indeed $[0, 1)$ is not open as there is no neighbourhood of 0 that is contained in $[0, 1)$. Similarly $[0, 1)$ is not closed as $\mathbb{R} \setminus [0, 1) = (-\infty, 0) \cup [1, \infty)$ is not open as there is no neighbourhood of 1 that is contained in $\mathbb{R} \setminus [0, 1)$. Thus it is possible that sets are neither open nor closed.

Example A.5.4. Given a topological space $(X, \mathcal{T})$, the sets $\emptyset$ and X are always closed as $X \setminus \emptyset = X$ and $X \setminus X = \emptyset$ are open.

Example A.5.5. In the discrete topology, every set is closed as every set is open so the complement of every set is open.

Example A.5.6. In the cofinite topology, the closed sets are exactly the entire space and the set of finite subsets. Similarly, in the cocountable topology, the closed sets are exactly the entire space and the set of countable subsets.

Example A.5.7. Let (X, d) be a metric space. Given an $x \in X$ and an $r > 0$, the *closed d-ball of radius r centred at x*, denoted $B_d[x, r]$, is the set

$$B_d[x, r] = \{y \in X \mid d(x, y) \leq r\}.$$

Any closed ball in any metric space is closed. Indeed to see that $B_d[x, r]$ is closed, let $y \in X \setminus B_d[x, r]$ be arbitrary. Then $d(x, y) > r$. Thus $B_d(y, d(x, y) - r)$ is an open set in (X, d). Furthermore, notice if $z \in B_d(y, d(x, y) - r)$ then $d(z, y) < d(x, y) - r$ so

$$d(x, z) \geq d(x, y) - d(y, z) > d(x, y) - (d(x, y) - r) = r$$

and thus $z \notin B_d[x, r]$. Hence $B_d(y, d(x, y) - r)$ is an open set containing y that is contained in $X \setminus B_d[x, r]$. Therefore, as $y \in X \setminus B_d[x, r]$ was arbitrary, $X \setminus B_d[x, r]$ is open and thus $B_d[x, r]$ is closed.

Example A.5.8. If $(X, \mathcal{T})$ is a Hausdorff space, then every singleton is closed. Indeed let $x \in X$ be arbitrary. As $(X, \mathcal{T})$ is Hausdorff, for each $y \in Y$ there exists a $U_y \in \mathcal{T}$ such that $y \in U_y$ but $x \notin U_y$. Thus

$$X \setminus \{x\} = \bigcup_{y \in X \setminus \{x\}} U_y \in \mathcal{T}.$$

Hence $\{x\}$ is closed.

As the notion of a topological space immediately invokes properties on open sets, we immediately have the following properties of closed sets by taking complements and using De Morgan's Laws.

Proposition A.5.9. *Let $(X, \mathcal{T})$ be a topological space. Then:*

(1) $\emptyset$ and X are closed sets.

(2) If $\{F_\alpha\}_{\alpha \in I}$ are closed sets in $(X, \mathcal{T})$, then $\bigcap_{\alpha \in I} F_\alpha$ is closed in $(X, \mathcal{T})$.

(3) If $\{F_\alpha\}_{\alpha \in I}$ are closed sets in $(X, \mathcal{T})$ and I is finite, then $\bigcup_{\alpha \in I} F_\alpha$ is closed in $(X, \mathcal{T})$.

Proof. Simply apply Definition A.1.1 and De Morgan's Laws. ∎

Example A.5.10. In any Hausdorff space, any finite union of points is closed as Example A.5.8 shows singleton points are closed and Proposition A.5.9 concludes finite unions of closed sets are closed.

Example A.5.11. Let $P_0 = [0, 1]$. Construct P_1 from P_0 by removing the open interval of length $\frac{1}{3}$ from the middle of P_0 (i.e. $P_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$). Then construct P_2 from P_1 by removing the open intervals of length $\frac{1}{3^2}$ from the middle of each closed subinterval of P_1. Subsequently, having constructed P_n, construct P_{n+1} by removing the open intervals of length $\frac{1}{3^{n+1}}$ from the middle of each of the 2^n closed subintervals of P_n. Specifically, P_n is the union of the 2^n closed intervals of the form

$$\left[\sum_{k=1}^{n} \frac{a_k}{3^k}, \frac{1}{3^n} + \sum_{k=1}^{n} \frac{a_k}{3^k}\right]$$

where $a_1, \ldots, a_n \in \{0, 2\}$.

The set

$$\mathcal{C} = \bigcap_{n \geq 1} P_n$$

is known as the *Cantor set*. The Cantor set is closed by Proposition A.5.9 being the intersection of finite unions of closed sets. In fact, it can be shown that $\mathcal{C}$ is uncountable.

When we restrict to subspaces of topological spaces, the closed subsets are easy to understand.

Lemma A.5.12. *Let $(Y, \mathcal{T}_Y)$ be a subspace of a topological space $(X, \mathcal{T})$. A subset $A \subseteq Y$ is closed in $(Y, \mathcal{T}_Y)$ if and only if $A = Y \cap F$ where F is a closed set in $(X, \mathcal{T})$.*

Proof. First suppose $A = Y \cap F$ where F is a closed set in $(X, \mathcal{T})$. As F is closed in $(X, \mathcal{T})$, $V = X \setminus F \in \mathcal{T}$. Hence $U = Y \cap V$ is open in $(Y, \mathcal{T}_Y)$ so

$$Y \setminus U = \{y \in Y \mid y \notin U\} = \{y \in Y \mid y \notin V\} = A$$

is closed.

Conversely, suppose $A \subseteq Y$ is closed in $(Y, \mathcal{T}_Y)$. Then $U = Y \setminus A$ is open in $(Y, \mathcal{T}_Y)$. By the definition of the open subsets of a subspace, there exists a $V \in \mathcal{T}$ such that $U = Y \cap V$. Hence $F = X \setminus V$ is closed in $(X, \mathcal{T})$ and

$$Y \cap F = \{y \in Y \mid y \notin V\} = \{y \in Y \mid y \notin U\} = A.$$

Hence the result is complete. ∎

Example A.5.13. Consider $Y = (0, 2)$ with the subspace topology inherited from the canonical topology on $\mathbb{R}$. Then

$$(0, 1] = Y \cap [-1, 1]$$

is closed in the subspace topology even though $(0, 1]$ is not closed in $\mathbb{R}$.

The reason closed sets are awesome is due to their relations with limits of nets.

Theorem A.5.14. *Let $(X, \mathcal{T})$ be a topological space and let $F \subseteq X$. Then the following are equivalent:*

(i) F is a closed set in $(X, \mathcal{T})$.

(ii) Whenever $(x_\lambda)_{\lambda \in \Lambda}$ is a net such that $x_\lambda \in F$ for all $\lambda \in \Lambda$ that converges to a point $x_0 \in X$, then $x_0 \in F$.

Proof. To begin, suppose F is a closed set in $(X, \mathcal{T})$ and that $(x_\lambda)_{\lambda \in \Lambda}$ is a net such that $x_\lambda \in F$ for all $\lambda \in \Lambda$ that converges to a point $x_0 \in X$. Suppose to the contrary that $x_0 \notin F$. Then $x_0 \in X \setminus F$. As F is closed, $X \setminus F$ is open so $x_0 \in X \setminus F$ and the definition of a convergent net implies there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in X \setminus F$ for all $\lambda \geq \lambda_0$. As this contradicts the fact that $x_\lambda \in F$ for all $\lambda \in \Lambda$, we have a contradiction. Hence $x_0 \in F$ as desired.

Conversely, suppose that whenever $(x_\lambda)_{\lambda \in \Lambda}$ is a net such that $x_\lambda \in F$ for all $\lambda \in \Lambda$ that converges to a point $x_0 \in X$, then $x_0 \in F$. To see that F must be closed, suppose to the contrary that F is not closed. Then $X \setminus F$ is not open. Hence there exists a point $x_0 \in X \setminus F$ such that for every neighbourhood U of x_0, $U \cap F \neq \emptyset$.

Let

$$\Lambda = \{U \subseteq X \mid U \text{ is a neighourhood of } x_0\}.$$

As Λ is closed under finite intersections, if for $U_1, U_2 \in \Lambda$ we define $U_1 \leq U_2$ if $U_2 \subseteq U_1$, then $(\Lambda, \leq)$ is a direct set by Example A.4.5.

For each $U \in \Lambda$, let $x_U \in F \cap U$ (note we are using the Axiom of Choice here). We claim that $(x_U)_{U\in\Lambda}$ is a net that converges to x_0. This then leads to a contradiction as $x_U \in F$ for all $U \in \Lambda$ but $x_0 \notin F$ thereby completing the proof. To see that $(x_U)_{U\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$, let U_0 be an arbitrary neighbourhood x. Then for all $U \geq U_0$ we have that $x_U \in U \subseteq U_0$. Hence $(x_U)_{U\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$ as claimed. ∎

Example A.5.15. Let I be a non-empty set, let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces, and let F_α be a closed subset of $(X_\alpha, \mathcal{T}_\alpha)$ for all $\alpha \in I$. We claim that $\prod_{\alpha\in I} F_\alpha$ is closed in $\prod_{\alpha\in I} X_\alpha$ when equipped with the product topology. Indeed let $(f_\lambda)_{\lambda\in\Lambda}$ be an arbitrary net in $\prod_{\alpha\in I} F_\alpha$ that converges to some element $f \in \prod_{\alpha\in I} X_\alpha$. By the 'if'-direction of Theorem A.4.22, for each $\alpha \in I$ the net $(f_\lambda(\alpha))_{\lambda\in\Lambda}$ converges to $f(\alpha)$ in $(X_\alpha, \mathcal{T}_\alpha)$. Therefore, as $f_\lambda(\alpha) \in F_\alpha$ for all $\lambda \in \Lambda$ and F_α is closed in $(X_\alpha, \mathcal{T}_\alpha)$, Theorem A.5.14 implies $f(\alpha) \in F_\alpha$ for all α. Hence $f \in \prod_{\alpha\in I} F_\alpha$. Therefore, as $(f_\lambda)_{\lambda\in\Lambda}$ was arbitrary, Theorem A.5.14 implies $\prod_{\alpha\in I} F_\alpha$ is closed.

Given a subset of a topological space, there will be potentially lots of convergent nets contained in a given subset. It would be nice to find a closed set that contains all the possible points of convergence. In particular, it would be nice to find the smallest possible set with this property.

Construction A.5.16. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Note the set

$$\mathcal{F} = \{F \subseteq X \mid A \subseteq F \text{ and } F \text{ is a closed set in } (X, \mathcal{T})\}$$

is non-empty as $X \in \mathcal{F}$. Consequently, Proposition A.5.9 implies the set

$$\overline{A} = \bigcap_{F\in\mathcal{F}} F$$

is a closed set in $(X, \mathcal{T})$ that contains A. As clearly $A \subseteq \overline{A}$, we obtain that $\overline{A} \in \mathcal{F}$ and thus $\overline{A}$ is the smallest closed set in $(X, \mathcal{T})$ that contains A. This causes us to define the following.

Definition A.5.17. The *closure* of a set A in a topological space $(X, \mathcal{T})$ is the set $\overline{A}$ obtained by taking the intersection of all closed subsets of $(X, \mathcal{T})$ that contain A.

Example A.5.18. Given $\mathbb{R}$ with its canonical topology and $a, b \in \mathbb{R}$ with $a < b$, the closure of each of (a, b), $[a, b]$, $(a, b]$, and $[a, b)$ is $[a, b]$. Indeed $[a, b]$ is a closed set containing each of these sets. Furthermore, as every other close subset of $\mathbb{R}$ containing one of these sets must also contain a and b by Theorem A.5.14, $[a, b]$ is the smallest closed subset of $\mathbb{R}$ containing each of these sets and thus must be the closure of each of these sets.

Example A.5.19. Let (X, d) be a metric space with at least two points, let $x \in X$, and let $r > 0$. It is possible that $\overline{B_d(x, r)} \neq B_d[x, r]$. Indeed let d be the discrete metric on X. Then $\{x\} = B_d(x, 1)$ is closed and thus equal to its own closure. However $B_d[x, 1] = X$ which does not equal $B_d(x, 1)$.

Of course, closures of sets behave well with respect to subspaces and products.

Lemma A.5.20. *Let $(Y, \mathcal{T}_Y)$ be a subspace of a topological space $(X, \mathcal{T})$ and let $A \subseteq Y$. The closure of A in $(Y, \mathcal{T}_Y)$ is the intersection of Y and the closure of A in $(X, \mathcal{T})$.*

Proof. Let B denote the closure of A in $(Y, \mathcal{T}_Y)$ and let C denote the closure of A in $(X, \mathcal{T})$. As C is closed in $(X, \mathcal{T})$, $Y \cap C$ is closed in $(Y, \mathcal{T}_Y)$ by Lemma A.5.12. Therefore $B \subseteq Y \cap C$ by definition.

To see the other inequality, recall since B is a closed set in $(Y, \mathcal{T}_Y)$ that Lemma A.5.12 implies there exists a closed set F in $(X, \mathcal{T})$ such that $B = Y \cap F$. However as $A \subseteq B = Y \cap F \subseteq F$, and as F is a closed subset in $(X, \mathcal{T})$, the definition of the closure of a set implies $C \subseteq F$. Hence

$$Y \cap C \subseteq Y \cap F = B$$

as desired. ∎

Before we show how closures work for the product and box topologies, we demonstrate the following useful tool.

Theorem A.5.21. *Let $(X, \mathcal{T})$ be a topological space, let $A \subseteq X$, and let $x \in X$. The following are equivalent:*

(i) $x \in \overline{A}$.

(ii) There exists a net $(x_\lambda)_{\lambda \in \Lambda}$ of points in A that converges to x.

(iii) For every neighbourhood $U \in \mathcal{T}$ of x, $U \cap A \neq \emptyset$.

Furthermore, if $\mathcal{B}$ is a basis for $(X, \mathcal{T})$ or a neighbourhood basis for x, then the above are equivalent to

(iv) For every neighbourhood $U \in \mathcal{B}$ of x, $U \cap A \neq \emptyset$.

Proof. First suppose $x \in \overline{A}$. To see that (iii) holds, suppose to the contrary that there exists a neighbourhood $U \in \mathcal{T}$ of x such that $U \cap A = \emptyset$. Then $X \setminus U$ is a closed set containing A so $\overline{A} \subseteq X \setminus U$. However $x \in U$ and $x \in \overline{A}$ contradict the fact that $\overline{A} \subseteq X \setminus U$. Hence (i) implies (iii).

Next, suppose (iii) holds. To see that (ii) holds, let

$$\Lambda = \{U \subseteq X \mid U \text{ is a neighourhood of } x\}.$$

As Λ is closed under finite intersections, if for $U_1, U_2 \in \Lambda$ we define $U_1 \leq U_2$ if $U_2 \subseteq U_1$, then $(\Lambda, \leq)$ is a direct set by Example A.4.5.

For each $U \in \Lambda$, let $x_U \in A \cap U$ (note we are using the Axiom of Choice here). We claim that $(x_U)_{U\in\Lambda}$ is a net that converges to x. To see this, let U_0 be an arbitrary neighbourhood x. Then for all $U \geq U_0$ we have that $x_U \in U \subseteq U_0$. Hence $(x_U)_{U\in\Lambda}$ is a net that converges to x in $(X, \mathcal{T})$ and, as $x_U \in A$ for all $U \in \Lambda$, we have constructed an acceptable net. Hence (iii) implies (ii).

To see that (ii) implies (i), we note that if exists a net $(x_\lambda)_{\lambda\in\Lambda}$ of points in A that converges to x, then x must be in every closed subset of X containing A by Theorem A.5.14. Thus $x \in \overline{A}$ by the definition of the closure of a set. Hence (ii) implies (i).

Finally, in the case $\mathcal{B}$ is a basis for $(X, \mathcal{T})$ or a neighbourhood basis for x, i) implies iv) by identical arguments. Furthermore to see that (iv) implies (ii), consider

$$\Lambda = \{U \in \mathcal{B} \mid U \text{ is a neighourhood of } x\}.$$

Then Λ is a net with the same ordering as above by the properties of bases and neighbourhood bases. A net $(x_U)_{U\in\Lambda}$ is constructed as above and still converges to x by the properties of bases and neighbourhood bases. ∎

Using Theorem A.5.21, we can describe closure in the box and product topologies.

Proposition A.5.22. *Let I be a non-empty set, let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha\in I}$ be a non-empty indexed family of topological spaces, and let $A_\alpha \subseteq X_\alpha$ for all $\alpha \in I$. Then, when $\prod_{\alpha\in I} X_\alpha$ is equipped with the product topology,*

$$\overline{\prod_{\alpha\in I} A_\alpha} = \prod_{\alpha\in I} \overline{A_\alpha}.$$

Proof. By Example A.5.15 we see that $\prod_{\alpha\in I} \overline{A_\alpha}$ is a closed set containing $\prod_{\alpha\in I} A_\alpha$. Hence

$$\overline{\prod_{\alpha\in I} A_\alpha} \subseteq \prod_{\alpha\in I} \overline{A_\alpha}$$

by definition.

To see the other inequality we will use Theorem A.5.21. Let $x \in \prod_{\alpha\in I} \overline{A_\alpha}$ be arbitrary and write $x = (x_\alpha)_{\alpha\in I}$. To see that $x \in \overline{\prod_{\alpha\in I} A_\alpha}$, let V be a neighbourhood of x. Thus, by our knowledge of bases, we can find a set $U = \prod_{\alpha\in I} U_\alpha$ where $U_\alpha \in \mathcal{T}_\alpha$ (with $U_\alpha = X_\alpha$ for all but finitely many $\alpha \in I$ in the case we are using the product topology) such that $x \in U \subseteq V$. Since $x \in U$ we have that $x_\alpha \in U_\alpha$ for all $\alpha \in I$. Moreover, since $x \in \prod_{\alpha\in I} \overline{A_\alpha}$ we know that $x_\alpha \in \overline{A_\alpha}$ for all $\alpha \in I$ by Theorem A.5.21. Therefore, since $x_\alpha \in U_\alpha$ and since $x_\alpha \in \overline{A_\alpha}$ there exists an $a_\alpha \in A_\alpha \cap U_\alpha$ for all $\alpha \in I$ by a property of the closure. Thus

$$(a_\alpha)_{\alpha\in I} \in U \cap \left(\prod_{\alpha\in I} A_\alpha\right) \subseteq V \cap \left(\prod_{\alpha\in I} A_\alpha\right).$$

Therefore, as V was an arbitrary neighbourhood of x, we obtain that $x \in \overline{\prod_{\alpha\in I} A_\alpha}$ by Theorem A.5.21 as desired. ∎

Of course, when studying closures and closed sets via limits, for each element x in a set A there is clearly a net with elements from A that converges to x; namely a constant net where every value is x. Thus when taking a closure or asking whether a set is closed, we are more interested in nets that do not take the value of a specific point in a set. In particular, analysing the proof of Theorem A.5.21, we easily see the following.

Corollary A.5.23. *Let $(X, \mathcal{T})$ be a topological space, let $A \subseteq X$, and let $x \in X$. The following are equivalent:*

(i) There exists a net $(x_\lambda)_{\lambda\in\Lambda}$ of points in $A \setminus \{x\}$ that converges to x.

(ii) For every neighbourhood U of x, $U \cap (A \setminus \{x\}) \neq \emptyset$.

Proof. The proof that (ii) implies (i) is identical to the proof that (iii) implies (ii) in Theorem A.5.21. Conversely, the proof that (i) implies (ii) follows directly from the definition of a convergent net. ∎

As being able to determining points of convergence from non-constant nets is useful in many scenarios, we give said object a name.

Definition A.5.24. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. A point $x \in X$ is said to be a *cluster point* of A if one of the two equivalent conditions from Corollary A.5.23 hold for x, A, and $(X, \mathcal{T})$.

The set of cluster points of A is denoted $\mathrm{cluster}(A)$.

Remark A.5.25. Note some authors use the term ‘limit points’ instead of cluster points. However a disjoint set of authors use the term ‘limits points’ to mean the set of all points of convergence. Thus we endeavour to ignore this ambiguity.

Example A.5.26. Given $\mathbb{R}$ equipped with its canonical topology and $a, b \in \mathbb{R}$ with $a < b$, it is not difficult to see that the set of cluster points for $[a, b]$, (a, b), $[a, b)$, and $(a, b]$ are all $[a, b]$ as every point in $[a, b]$ is a point of convergence for some non-constant net from (a, b).

Example A.5.27. Let $A = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$ viewed as a subset of $\mathbb{R}$ with its canonical topology. Then the only cluster point of A is 0. Indeed clearly the sequence $\left(\frac{1}{n}\right)_{n\geq 1}$ converges to but never equals 0 and thus 0 is a cluster point. To see that 0 is the only cluster point of A, we first claim that $\overline{A} = A \cup \{0\}$. To see this, we note that $A \cup \{0\}$ is closed as its complement is a countable union of open intervals and thus is open. However A is not closed by Theorem A.5.14 as $\left(\frac{1}{n}\right)_{n\geq 1}$ is a sequence from A that converges to 0, which is not in A. Hence $\overline{A} = A \cup \{0\}$. Therefore, by Theorem A.5.14, the set of possible cluster points must be contained in $A \cup \{0\}$. However, it is clear that no point in A can be a cluster point of A since the distance between $\frac{1}{n}$ and any other point in A is at least $\frac{1}{n} - \frac{1}{n-1}$ so it is impossible for a net from $A \setminus \left\{\frac{1}{n}\right\}$ to converge to $\frac{1}{n}$. Hence the only cluster point of A is 0.

Example A.5.28. Let $\mathcal{C}$ be the Cantor set from Example A.5.11. Then the set of cluster points of $\mathcal{C}$ is precisely $\mathcal{C}$. To see this, note as $\mathcal{C}$ is closed that the cluster points of $\mathcal{C}$ are contained in $\mathcal{C}$. To see that every point in $\mathcal{C}$ is a cluster point of $\mathcal{C}$, let $x \in \mathcal{C}$ be arbitrary. Thus, by the definition of $\mathcal{C}$, for each $n \in \mathbb{N}$ there exists a unique closed interval I_n of the form $\left[\frac{2k_n}{3^n}, \frac{2k_n+1}{3^n}\right]$ where $k_n \in \{0, 1, \ldots, \frac{1}{2}(3^n - 1)\}$. Choose x_n to be one of the endpoints of I_n that is not equal to x (as there are two distinct endpoints, such a point exists). As it is elementary to verify that the endpoints of I_n are elements of $\mathcal{C}$, we see that $x_n \in \mathcal{C}$ and $|x - x_n| < \frac{1}{3^n}$. Hence $(x_n)_{n\geq 1}$ is a sequence in $\mathcal{C} \setminus \{x\}$ that converges to x. Hence $\mathcal{C}$ is equal to its cluster points.

Example A.5.29. Let A be a non-empty subset of a topological space $(X, \mathcal{T})$. Then the closure of A in the subspace $(A, \mathcal{T}_A)$ is A as A is closed in the subspace topology.

Perhaps unsurprisingly, the only thing that prevents a set from begin closed is it not containing its cluster points.

Theorem A.5.30. *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Then*

$$\overline{A} = A \cup \operatorname{cluster}(A).$$

Proof. First, it is clear that $A \subseteq \overline{A}$ and that $\operatorname{cluster}(A) \subseteq \overline{A}$ by Theorem A.5.21 and the definition of a cluster point. Hence

$$\overline{A} \supseteq A \cup \operatorname{cluster}(A).$$

To see the other inequality, let $x \in \overline{A}$ be arbitrary. If $x \in A$ then $x \in A \cup \text{cluster}(A)$ and there is nothing left to show. Thus we may suppose that $x \notin A$. Since $x \in \overline{A}$, Theorem A.5.21 implies that $U \cap A \neq \emptyset$ for every neighbourhood U of x. As $x \notin A$, $U \cap (A \setminus \{x\}) \neq \emptyset$ for every neighbourhood U of x. Hence Corollary A.5.23 implies that $x \in \text{cluster}(A)$. Therefore, in either case $x \in A \cup \text{cluster}(A)$. Hence, as $x \in \overline{A}$ was arbitrary, $\overline{A} = A \cup \text{cluster}(A)$ as desired. ∎

Corollary A.5.31. *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Then A is closed if and only if* $\text{cluster}(A) \subseteq A$.

Proof. If A is closed, then $A = \overline{A} = A \cup \text{cluster}(A)$ by Theorem A.5.30 and hence $\text{cluster}(A) \subseteq A$. Conversely, if $\text{cluster}(A) \subseteq A$, then Theorem A.5.30 implies that $\overline{A} = A \cup \text{cluster}(A) = A$. Therefore, as A is equal to its closure and the closure of a set is a closed set, A is closed. ∎

All of the above has been focused on closures and closed sets via describing points of convergence for nets based on a set. However, it is often useful to understand just the points inside a set. In particular, it is useful to understand the set of points in a set that are 'far away' from the complement of the set. These points are described based on the following, which is constructed in a similar fashion to how we constructed the closure of a set.

Construction A.5.32. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Note the set

$$\mathcal{U} = \{U \subseteq X \mid U \subseteq A \text{ and } U \in \mathcal{T}\}$$

is non-empty as $\emptyset \in A$. Consequently, Definition A.1.1 implies the set

$$\text{int}(A) = \bigcup_{U \in \mathcal{U}} U$$

is an open set in $(X, \mathcal{T})$ that is contained A. As clearly $\text{int}(A) \subseteq A$, we obtain that $\text{int}(A) \in \mathcal{U}$ and thus $\text{int}(A)$ is the largest open set in $(X, \mathcal{T})$ that is contained in A. This causes us to define the following.

Definition A.5.33. The *interior* of a set A in a topological space $(X, \mathcal{T})$ is the set $\text{int}(A)$ obtained by taking the union of all open subsets of $(X, \mathcal{T})$ contained A.

Example A.5.34. Given $\mathbb{R}$ equipped with its canonical topology and $a, b \in \mathbb{R}$ with $a < b$, it is not difficult to see that interior of $[a, b]$, (a, b), $[a, b)$, and $(a, b]$ is (a, b) as clearly (a, b) is open, is contained in these sets, contains all points in these sets for except possible a and b, and as the addition of a or b to (a, b) creates a set that is not open.

Example A.5.35. Let $A = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$ viewed as a subset of $\mathbb{R}$ with its canonical topology. Then the interior of A is the empty set as no open interval is contained in A.

Example A.5.36. The Cantor set $\mathcal{C}$, viewed as a subset of $\mathbb{R}$, has no interior. Indeed suppose to the contrary that $\mathrm{int}(\mathcal{C})$ is non-empty. Hence there exists an open interval $(a, b) \subseteq \mathrm{int}(\mathcal{C}) \subseteq \mathcal{C}$ by the definition of the interior. By the elementary properties of real numbers, we can choose $N \in \mathbb{N}$ such that $\frac{1}{3^N} < b - a$. This then implies that (a, b) cannot be contained in P_N as defined in Example A.5.11 as none of the separated intervals in P_N have length greater than $\frac{1}{3^N}$. Hence the Cantor set has no interior (even though it is uncountable and every point is a cluster point).

Example A.5.37. Let A be the x-axis in $\mathbb{R}^2$ equipped with its topology from the Euclidean norm. Then the interior of A is empty as A contains no open balls from $\mathbb{R}^2$.

Example A.5.38. Let A be a non-empty subset of topological space $(X, \mathcal{T})$. Then the interior of A in the subspace $(A, \mathcal{T}_A)$ is A as A is open in the subspace topology.

Although we do not have any theory here related to the interior like we did with the closure results seen above, the interior will be useful later in the course.

As we can see based on these examples, the set of interior points to a set are those that are 'far away' from the complement of the set as there is an open set containing these points that does not intersect the complement. To formalize this, we define another type of point for a given set.

Definition A.5.39. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. A point $x \in X$ is said to be a *boundary point of* A if $A \cap U \neq \emptyset$ and $(X \setminus A) \cap U \neq \emptyset$ for every neighbourhood $U \in T$ of x.

The set of boundary points of A is denoted $\mathrm{bdy}(A)$.

Before we look at examples of boundary points, we first prove two results which completely describe the set of boundary points based on objects we have previously studied.

Corollary A.5.40. *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Then* $\mathrm{bdy}(A) = \overline{A} \cap \overline{(X \setminus A)}$.

Proof. This result easily follows from Theorem A.5.21 and the definition of a boundary point. ∎

Theorem A.5.41. *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Then* $\mathrm{int}(A)$ *and* $\mathrm{bdy}(A)$ *are disjoint sets such that*

$$\overline{A} = \mathrm{int}(A) \cup \mathrm{bdy}(A).$$

Furthermore

$$\mathrm{int}(A) = A \setminus \overline{(X \setminus A)}.$$

Proof. Clearly if $x \in \text{int}(A)$, then there exists an open set U containing x (namely $\text{int}(A)$) such that $U \cap (X \setminus A) = \emptyset$ and thus $x \notin \text{bdy}(A)$. Hence $\text{int}(A)$ and $\text{bdy}(A)$ are disjoint sets.

To see that $\overline{A} \subseteq \text{int}(A) \cup \text{bdy}(A)$, let $x \in \overline{A}$ be arbitrary. Hence Theorem A.5.21 implies that for every neighbourhood U of x, $U \cap A \neq \emptyset$. If there exists a neighbourhood U of x such that $U \cap (X \setminus A) = \emptyset$, then $U \subseteq A$ and hence $x \in \text{int}(A)$. Otherwise for every neighbourhood U of x, $U \cap A \neq \emptyset$ and $U \cap (X \setminus A) \neq \emptyset$ so $x \in \text{bdy}(A)$. Therefore, as $x \in \overline{A}$ was arbitrary, $\overline{A} \subseteq \text{int}(A) \cup \text{bdy}(A)$. For the reverse inequality, we note that $\text{bdy}(A) \subseteq \overline{A}$ by the definition of a boundary point and Theorem A.5.21, and similarly $\text{int}(A) \subseteq A \subseteq \overline{A}$ trivially.

Finally, to see that $\text{int}(A) = A \setminus \overline{(X \setminus A)}$, note if $x \in \text{int}(A)$ then there exists a neighbourhood U of x contained in A and thus $x \notin \overline{(X \setminus A)}$ by Theorem A.5.21. Furthermore, as $\text{int}(A) \subseteq A$ by construction, $\text{int}(A) \subseteq A \setminus \overline{(X \setminus A)}$. To see the reverse inequality, note as

$$A \setminus \overline{(X \setminus A)} \subseteq X \setminus \overline{(X \setminus A)} \subseteq X \setminus (X \setminus A) = A,$$

and as $X \setminus \overline{(X \setminus A)}$ is the complement of an closed set and thus is open, $A \setminus \overline{(X \setminus A)} \subseteq \text{int}(A)$ by definition. Hence $\text{int}(A) = A \setminus \overline{(X \setminus A)}$ as desired. ■

Example A.5.42. Given $\mathbb{R}$ equipped with its canonical topology and $a, b \in \mathbb{R}$ with $a < b$, it is not difficult to see that boundary of $[a, b]$, (a, b), $[a, b)$, and $(a, b]$ is $\{a, b\}$. Indeed as the closure of each of these sets is $[a, b]$ by Example A.5.18 and the interior of each of these sets is (a, b) by Example A.5.34, the claim follows from Theorem A.5.41.

Example A.5.43. Given $\mathbb{R}$ equipped with its canonical topology, the boundary of $\mathbb{Q}$ is $\mathbb{R}$ as every neighbourhood of some point from $\mathbb{Q}$ contains an interval, which must contain a rational and irrational number.

Example A.5.44. Let (X, d) be a metric space with at least two points, let $x \in X$, and let $r > 0$. It is possible that $\text{bdy}(B_d(x, r))$ and $\text{bdy}(B_d[x, r])$ are not equal to

$$\{y \in Y \in\mid \ d(x, y) = r\}.$$

Indeed let d be the discrete metric on X. Then $B_d(x, 1) = \{x\}$ and $B_d[x, 1] = X$ have empty boundary sets as they are open sets and thus equal to their own interior. However the above set is X which is not equal to $\emptyset$.

The notions of points and sets observed in this section will be seen throughout the course (less so with the boundary points). As with undergraduate real analysis, the notions related to closed sets and closures of sets will be of vital importance when discussing continuous functions and compact sets; which happen to be the next two chapters.

A.6 Continuous Functions

Continuous functions between topological spaces are vital to this course. In particular, we desire both an 'open set' characterization and a 'convergent net' characterization of continuous functions.

Recall a function $f : \mathbb{R} \to \mathbb{R}$ is continuous if for each $x_0 \in \mathbb{R}$ and each $\epsilon > 0$ there exists a $\delta > 0$ such that if $x \in \mathbb{R}$ and $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. Alternatively, this can be rewritten as

$$(x_0 - \delta, x_0 + \delta) \subseteq f^{-1}((f(x_0) - \epsilon, f(x_0) + \epsilon)).$$

As open intervals form a basis for the canonical topology on $\mathbb{R}$, it is elementary to generalize the above idea of a continuous function to topological spaces.

Definition A.6.1. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A function $f : X \to Y$ is said to be *continuous* if $f^{-1}(U) \in \mathcal{T}_X$ for every $U \in \mathcal{T}_Y$; that is, the inverse image of every open set (from Y) is open (in X).

Example A.6.2. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $y_0 \in Y$. The constant function $f : X \to Y$ defined by $f(x) = y_0$ for all $x \in X$ is a continuous function. Indeed for every open set U in Y we have that

$$f^{-1}(U) = \begin{cases} X & \text{if } y_0 \in U \\ \emptyset & \text{if } y_0 \notin U \end{cases}.$$

Therefore, as $\emptyset, X \in \mathcal{T}_X$ by the definition of a topology, f is continuous.

Example A.6.3. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. If $\mathcal{T}_X$ is the discrete topology, then every function $f : X \to Y$ is continuous as $\mathcal{T}_X = \mathcal{P}(X)$ implies $f^{-1}(U) \in \mathcal{T}_X$ for every $U \in \mathcal{T}_Y$.

Example A.6.4. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. If $\mathcal{T}_Y$ is the trivial topology, then every function $f : X \to Y$ is continuous as $f^{-1}(Y) = X$, $f^{-1}(\emptyset) = \emptyset$, and $\mathcal{T}_Y = \{\emptyset, Y\}$.

Example A.6.5. Let I be a non-empty set and let $\{(X_\alpha, \mathcal{T}_\alpha)\}_{\alpha \in I}$ be a set of topological spaces. For a fixed $\alpha_0 \in I$, consider the map

$$\pi_{\alpha_0} : \prod_{\alpha \in I} X_\alpha \to X_{\alpha_0}$$

defined by

$$\pi_{\alpha_0}((x_\alpha)_{\alpha \in I}) = x_{\alpha_0}$$

for all $(x_\alpha)_{\alpha \in I} \in \prod_{\alpha \in I} X_\alpha$. The map π_{α_0} is called the *projection map onto the* α_0^{th} *coordinate*.

Every projection map is continuous when $\prod_{\alpha\in I} X_\alpha$ is equipped with either the product topology or the box topology. Indeed notice for all $U \in \mathcal{T}_{\alpha_0}$ that

$$\pi_{\alpha_0}^{-1}(U) = \prod_{\alpha\in I} V_\alpha$$

where $V_\alpha = X_\alpha$ if $\alpha \neq \alpha_0$ and $V_{\alpha_0} = U$. Thus, as $\prod_{\alpha\in I} V_\alpha$ is open in both the product and box topologies and as $U \in \mathcal{T}_{\alpha_0}$ was arbitrary, π_{α_0} is continuous.

In fact, as the collection $\{\pi_\alpha(U_\alpha) \mid \alpha \in I, U_\alpha \in \mathcal{T}_\alpha\}$ is a subbasis for the product topology, the product topology is the coarsest topology for which each projection map is continuous.

Of course, there are many ways to test whether a function on $\mathbb{R}$ is continuous. In particular, one characterization of continuous functions on $\mathbb{R}$ that is often used as the definition of continuity due to its viability is the characterization that a function is continuous if and only if it maps convergent sequences to convergent sequences. In the following result, we extend all of these characterizations to arbitrary topological spaces.

Theorem A.6.6. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$. The following are equivalent:*

(i) f is continuous.

(ii) For every $x \in X$ and every $\mathcal{T}_Y$-neighbourhood U of $f(x)$ there exists a $\mathcal{T}_X$-neighbourhood V of x such that $V \subseteq f^{-1}(U)$.

(iii) For any bases $\mathcal{B}_X$ of $(X, \mathcal{T}_X)$ and $\mathcal{B}_Y$ of $(Y, \mathcal{T}_Y)$, for every $x \in X$ and every neighbourhood $U \in \mathcal{B}_Y$ of $f(x)$ there exists a neighbourhood $V \in \mathcal{B}_X$ of x such that $V \subseteq f^{-1}(U)$.

(iv) For some bases $\mathcal{B}_X$ of $(X, \mathcal{T}_X)$ and $\mathcal{B}_Y$ of $(Y, \mathcal{T}_Y)$, for every $x \in X$ and every neighbourhood $U \in \mathcal{B}_Y$ of $f(x)$ there exists a neighbourhood $V \in \mathcal{B}_X$ of x such that $V \subseteq f^{-1}(U)$.

(v) For every net $(x_\lambda)_{\lambda\in\Lambda}$ in X that converges to some x_0 in $(X, \mathcal{T}_X)$, the net $(f(x_\lambda))_{\lambda\in\Lambda}$ converges to $f(x_0)$ in $(Y, \mathcal{T}_Y)$.

(vi) For every $A \subseteq X$, $f\left(\overline{A}\right) \subseteq \overline{f(A)}$, where the closures are taken in their appropriate spaces.

(vii) For every closed set F in $(Y, \mathcal{T}_Y)$, $f^{-1}(F)$ is closed in $(X, \mathcal{T}_X)$.

Proof. To see that (i) implies (ii), let f be continuous and let $x \in X$ and U a $\mathcal{T}_Y$-neighbourhood of $f(x)$ be arbitrary. Then, as f is continuous, $V = f^{-1}(U)$ is clearly a $\mathcal{T}_X$-neighbourhood of x such that $V \subseteq f^{-1}(U)$. Therefore, as x and U were arbitrary, (i) implies (ii).

To see that (ii) implies (iii), let $x \in X$ and $U \in \mathcal{B}_Y$ a $\mathcal{T}_Y$-neighbourhood of $f(x)$ be arbitrary. By ii) there exists a $\mathcal{T}_X$-neighbourhood V_0 of x such that $V_0 \subseteq f^{-1}(U)$. Since $\mathcal{B}_X$ is a basis for $(X, \mathcal{T}_X)$ there exists $V \in \mathcal{B}_X$ such that $x \in V \subseteq V_0$. Hence $V \in \mathcal{B}_X$ is a neighbourhood of x such that $V \subseteq V_0 \subseteq f^{-1}(U)$. Therefore, as x and U were arbitrary, (ii) implies (iii).

Note (iii) trivially implies (iv).

To see that (iv) implies (v), suppose $\mathcal{B}_X$ is a basis for $(X, \mathcal{T}_X)$ and $\mathcal{B}_Y$ is a basis for $(Y, \mathcal{T}_Y)$ such that for every $x \in X$ and every neighbourhood $U \in \mathcal{B}_Y$ of $f(x)$ there exists a neighbourhood $V \in \mathcal{B}_X$ of x such that $V \subseteq f^{-1}(U)$. Let $(x_\lambda)_{\lambda \in \Lambda}$ be an arbitrary net in X that converges to some x_0 in $(X, \mathcal{T}_X)$. To see that $(f(x_\lambda))_{\lambda \in \Lambda}$ converges to $f(x_0)$ in $(Y, \mathcal{T}_Y)$, let $U \in \mathcal{B}_Y$ such that $f(x_0) \in U$ be arbitrary. By assumption there exists a $V \in \mathcal{B}_X$ such that $x_0 \in V$ and $V \subseteq f^{-1}(U)$. Thus, as V is an open set containing x_0 and as $(x_\lambda)_{\lambda \in \Lambda}$ converges to x_0 in $(X, \mathcal{T}_X)$, there exists a $\lambda_0 \in \Lambda$ such that $x_\lambda \in V$ for all $\lambda \geq \lambda_0$. Hence $f(x_\lambda) \in f(V) \subseteq U$ for all $\lambda \geq \lambda_0$. Therefore, as U was arbitrary, $(f(x_\lambda))_{\lambda \in \Lambda}$ converges to $f(x_0)$ in $(Y, \mathcal{T}_Y)$ by Lemma A.4.16. Hence (iv) implies (v).

To see that (v) implies (vi), fix $A \subseteq X$ and let $x_0 \in \overline{A}$ be arbitrary. As $x_0 \in \overline{A}$ there exists a net $(x_\lambda)_{\lambda \in \Lambda}$ of points in A that converges to x_0 by Theorem A.5.21. Therefore, by v), $(f(x_\lambda))_{\lambda \in \Lambda}$ is a net of points in $f(A)$ that converges to $f(x_0)$ in $(Y, \mathcal{T}_Y)$. Hence Theorem A.5.21 $f(x_0) \in \overline{f(A)}$. Therefore, as $x_0 \in \overline{A}$ was arbitrary, $f\left(\overline{A}\right) \subseteq \overline{f(A)}$. Hence (v) implies (vi).

To see that (vi) implies (vii), let F be an arbitrary closed subset of $(Y, \mathcal{T}_Y)$ and let $A = f^{-1}(F \cap \overline{f(X)})$. Thus $F \cap \overline{f(X)} = f(A)$. Since $A \subseteq \overline{A}$, (vi) implies that

$$F \cap \overline{f(X)} = f(A) \subseteq f\left(\overline{A}\right) \subseteq \overline{f(A)} = \overline{F \cap \overline{f(X)}} = F \cap \overline{f(X)}$$

as $F \cap \overline{f(X)}$ is closed. Hence $f\left(\overline{A}\right) = F \cap \overline{f(X)}$ so $\overline{A} \subseteq f^{-1}(F \cap \overline{f(X)}) = A \subseteq \overline{A}$ so $A = \overline{A}$. Thus A is closed. Therefore, as F was arbitrary, (vi) implies (vii).

Finally, to see that (vii) implies (i), let $U \in \mathcal{T}_Y$ be arbitrary. Then $Y \setminus U$ is closed in $(Y, \mathcal{T}_Y)$. By assuming vii) we know that

$$f^{-1}(Y \setminus U) = f^{-1}(Y) \setminus f^{-1}(U) = X \setminus f^{-1}(U)$$

is closed in $(X, \mathcal{T}_X)$. Thus $f^{-1}(U) \in \mathcal{T}_X$. Hence, as $U \in \mathcal{T}_Y$ was arbitrary, f is continuous. Thus (vii) implies (i). ∎

Of course, alternate characterizations of continuous functions are always useful in proving results and obtaining examples of continuous functions.

Theorem A.6.7. *Let $(X, \mathcal{T})$ be a topological space, let I be a non-empty set, let $\{(Y_\alpha, \mathcal{T}_\alpha)\}_{\alpha \in I}$ be a set of topological spaces, and, for each $\alpha \in I$, let*

$f_\alpha : X \to Y_\alpha$. The function $f : X \to \prod_{\alpha \in I} Y_\alpha$ defined by

$$f(x) = (f_\alpha(x))_{\alpha \in I}$$

for all $x \in X$ is continuous when $Y = \prod_{\alpha \in I} Y_\alpha$ is equipped with the product topology if and only if f_α is continuous for all $\alpha \in I$.

Furthermore, if $\{(X_\alpha, \mathcal{T}'_\alpha)\}_{\alpha \in I}$ is a set of topological spaces, if for each $\alpha \in I$ $f_\alpha : X_\alpha \to Y_\alpha$, if $X = \prod_{\alpha \in I} X_\alpha$ is equipped with the product topology, and if $f : X \to Y$ is defined by $f((x_\alpha)_{\alpha \in I}) = (f_\alpha(x_\alpha))_{\alpha \in I}$, then f is continuous if and only if f_α is continuous for all $\alpha \in I$.

Proof. For the first part, let $(x_\lambda)_{\lambda \in \Lambda}$ be an arbitrary net in X that converges to some point x_0 in $(X, \mathcal{T})$. By Theorem A.4.22, $(f(x_\lambda))_{\lambda \in \Lambda}$ converges to $f(x_0)$ when $\prod_{\alpha \in I} X_\alpha$ is equipped with the product topology if and only if $(f_\alpha(x_\lambda))_{\lambda \in \Lambda}$ converges to $f_\alpha(x_0)$ in $(Y_\alpha, \mathcal{T}_\alpha)$ for all $\alpha \in I$. Hence the result follows from Theorem A.6.6.

Similarly, for the second part, let $(x_\lambda)_{\lambda \in \Lambda}$ be an arbitrary net in X that converges to some point x_0 in $(X, \mathcal{T})$. Hence $(x_\lambda(\alpha))_{\lambda \in \Lambda}$ converges to $x_0(\alpha)$ in $(X_\alpha, \mathcal{T}'_\alpha)$ for all $\alpha \in I$. By Theorem A.4.22, $(f(x_\lambda))_{\lambda \in \Lambda}$ converges to $f(x_0)$ when $\prod_{\alpha \in I} X_\alpha$ is equipped with the product topology if and only if $(f_\alpha(x_lambda(\alpha)))_{\lambda \in \Lambda}$ converges to $f_\alpha(x_0(\alpha))$ in $(X_\alpha, \mathcal{T}'_\alpha)$ for all $\alpha \in I$. Hence the result follows from Theorem A.6.6. ∎

Of course, in generality, we are interested in continuous functions as they will behave well with respect to any topological property we are interested in studying. On occasion, it is useful to study a more local property with respect to continuity. In particular, analyzing the proof of Theorem A.6.6 yields the following.

Theorem A.6.8. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $x_0 \in X$, and let $f : X \to Y$. The following are equivalent:*

(i) For every $\mathcal{T}_Y$-neighbourhood U of $f(x_0)$ there exists a $\mathcal{T}_X$-neighbourhood V of x_0 such that $V \subseteq f^{-1}(U)$.

(ii) For any bases $\mathcal{B}_X$ of $(X, \mathcal{T}_X)$ and $\mathcal{B}_Y$ of $(Y, \mathcal{T}_Y)$, every neighbourhood $U \in \mathcal{B}_Y$ of $f(x_0)$ there exists a neighbourhood $V \in \mathcal{B}_X$ of x_0 such that $V \subseteq f^{-1}(U)$.

(iii) For some bases $\mathcal{B}_X$ of $(X, \mathcal{T}_X)$ and $\mathcal{B}_Y$ of $(Y, \mathcal{T}_Y)$, for every neighbourhood $U \in \mathcal{B}_Y$ of $f(x_0)$ there exists a neighbourhood $V \in \mathcal{B}_X$ of x_0 such that $V \subseteq f^{-1}(U)$.

(iv) For every net $(x_\lambda)_{\lambda \in \Lambda}$ in X that converges to x_0 in $(X, \mathcal{T}_X)$, the net $(f(x_\lambda))_{\lambda \in \Lambda}$ converges to $f(x_0)$ in $(Y, \mathcal{T}_Y)$.

Proof. The fact that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) can be obtained by repeating (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) of Theorem A.6.6 verbatim.

To see that (iv) implies (i), assume that (iv) holds. Suppose to the contrary that there exists a $\mathcal{T}_Y$-neighbourhood U of $f(x_0)$ such that for every $\mathcal{T}_X$-neighbourhood V of x_0 that $V \setminus f^{-1}(U) \neq \emptyset$. Thus for every $\mathcal{T}_X$-neighbourhood V of x_0 there exists a $x_V \in V$ such that $f(x_V) \notin U$ (note we are using the Axiom of Choice here).

Let
$$\Lambda = \{V \subseteq X \mid V \text{ is a } \mathcal{T}_X\text{-neighourhood of } x_0\}.$$

As Λ is closed under finite intersections, if for $V_1, V_2 \in \Lambda$ we define $V_1 \leq V_2$ if $V_2 \subseteq V_1$, then $(\Lambda, \leq)$ is a direct set by Example A.4.5.

We claim that $(x_V)_{V\in\Lambda}$ converges to x_0 in $(X, \mathcal{T}_X)$ but $(f(x_V))_{V\in\Lambda}$ does not converge to $f(x_0)$ in $(Y, \mathcal{T}_Y)$. To see that $(x_V)_{V\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$, let V_0 be an arbitrary $\mathcal{T}$-neighbourhood x_0. Then for all $V \geq V_0$ we have that $x_V \in V \subseteq V_0$. Hence $(x_V)_{V\in\Lambda}$ is a net that converges to x_0 in $(X, \mathcal{T})$ by Definition A.4.10. Thus $(f(x_V))_{V\in\Lambda}$ does not converge to $f(x_0)$ in $(Y, \mathcal{T}_Y)$, we simply note that U is a $\mathcal{T}_Y$-neighbourhood of $f(x_0)$ but $f(x_V) \notin U$ for all $V \in \Lambda$. Hence we have obtained a contradiction thereby finishing the proof. ∎

Due to the above, we define the following.

Definition A.6.9. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $x_0 \in X$, and let $f : X \to Y$. It is said that f is *continuous at* x_0 if one of the four equivalent characterizations in Theorem A.6.8 hold

Of course, global continuity is exactly local continuity at each point.

Corollary A.6.10. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$. Then f is continuous if and only if f is continuous at each point in X.*

Proof. Combine Theorem A.6.6 and Theorem A.6.8. ∎

As mentioned earlier, it is on occasion useful to consider this local property of continuity due to all of the equivalent characterizations produced in Theorem A.6.8. Another useful ability is to be able to construct continuous functions from other continuous functions. The most well-known way to do this is the following whose proof trivially follows from the definition of continuity.

Theorem A.6.11. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$, and $(Z, \mathcal{T}_Z)$ be topological spaces. If $f : X \to Y$ and $g : Y \to Z$ are continuous functions, then $g \circ f : X \to Z$ is a continuous function.*

Proof. To see that $g \circ f$ is a continuous function, let $U \in \mathcal{T}_Z$ be arbitrary. Then $g^{-1}(U) \in \mathcal{T}_Y$ as g is continuous thus $f^{-1}(g^{-1}(U)) \in \mathcal{T}_X$ as f is continuous. Hence $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{T}_X$. Therefore, as $U \in \mathcal{T}_Z$ was arbitrary, $g \circ f$ is continuous by Definition A.6.1. ∎

One way to construct continuous function is to use inclusions and restrictions together with the subspace topology.

Lemma A.6.12. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $A \subseteq X$, and let $B \subseteq Y$. The following hold:*

(1) If A is equipped with the subspace topology, then the inclusion map $i : A \to X$ defined by $i(a) = a$ for all $a \in A$ is continuous.

(2) If A is equipped with the subspace topology and $f : X \to Y$ is continuous, then the restriction $f|_A : A \to Y$ defined by $f|_A(a) = f(a)$ for all $a \in A$ is continuous.

(3) If B is equipped with the subspace topology and $f : X \to B$ is continuous, then $f : X \to Y$ is continuous.

(4) If B is equipped with the subspace topology, $f : X \to Y$ is continuous, and $f(X) \subseteq B$, then $f : X \to B$ is continuous.

Proof. To see that (1) holds, notice for all open subsets U of X that $i^{-1}(U) = A \cap U$ is open in the subspace topology on A. Hence i is continuous by Definition A.6.1.

To see that (2) holds, notice for all open subsets U of X that $f|_A^{-1}(U) = A \cap f^{-1}(U)$ is open in the subspace topology on A as $f^{-1}(U)$ is an open subset of X since f is continuous. Hence $f|_A$ is continuous by Definition A.6.1.

To see that (3) holds, notice for all open subset V of Y that $f^{-1}(V) = f^{-1}(B \cap V)$ which must be open since $f : X \to B$ is continuous and $B \cap V$ is open in the subspace topology on B by definition. Hence $f : X \to Y$ is continuous by Definition A.6.1.

Finally, to see that (4) holds, recall that if V is an open subset of B in the subspace topology that $V = B \cap V_0$ for some open subset V_0 in Y. Therefore, since $f(X) \subseteq B$, we see that $f^{-1}(V) = f^{-1}(B \cap V_0) = f^{-1}(V_0)$ is open in X as $f : X \to Y$ is continuous and V_0 is open in Y. Hence $f : X \to B$ is continuous by Definition A.6.1. ∎

Instead of trying to restrict or compress a continuous function to obtain a continuous function, we can combine continuous functions to get continuous functions. Indeed the first of the following two results says that if we can cover a topological space with open sets and we have a function that is continuous on each of these open sets, then the function on the whole space must be continuous. The second result does the same for closed sets provided we have a finite number of closed sets with union all of X. Both of these results have uses in differential geometry.

Lemma A.6.13. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$. Suppose there exist $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ be such that $X = \bigcup_{\alpha \in I} U_\alpha$ and $f|_{U_\alpha}$ is continuous for all $\alpha \in I$, then f is continuous.*

Proof. To see that f is continuous, let $V \in \mathcal{T}_Y$ be arbitrary. Notice for all $\alpha \in I$ that $f|_{U_\alpha}^{-1}(V)$ is open in U_α equipped with the subspace topology from X as $f|_{U_\alpha}$ is continuous. Hence, by the definition of the subspace topology, there exists a $V_\alpha \in \mathcal{T}_X$ such that

$$f|_{U_\alpha}^{-1}(V) = U_\alpha \cap V_\alpha.$$

However, since $U_\alpha \in \mathcal{T}_X$, we obtain that $f|_{U_\alpha}^{-1}(V) \in \mathcal{T}_X$ being the intersection of two elements of $\mathcal{T}_X$. Therefore, since

$$f^{-1}(V) = \bigcup_{\alpha \in I} f|_{U_\alpha}^{-1}(V),$$

we obtain that $f^{-1}(V) \in \mathcal{T}_X$ as $\mathcal{T}_X$ is closed under unions. Hence, as $V \in \mathcal{T}_Y$ was arbitrary, f is continuous by Definition A.6.1. ∎

Theorem A.6.14 (The Pasting Lemma). *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be a topological spaces, let A, B be closed subsets of X such that $X = A \cup B$, and let $f : A \to Y$ and $g : B \to Y$ be continuous functions such that $f(x) = g(x)$ for all $x \in A \cap B$. Then the function $h : X \to Y$ such that $h(a) = f(a)$ for all $a \in A$ and $h(b) = g(b)$ for all $b \in B$ is continuous.*

Proof. To see that h is continuous, let F be an arbitrary closed subset of Y. Notice by construction that

$$h^{-1}(F) = f^{-1}(F) \cup g^{-1}(F).$$

However, as f and g are continuous functions, Theorem A.6.6 implies that $f^{-1}(F)$ is a closed subset of A when A is equipped with the subspace topology and $g^{-1}(F)$ is a closed subset of B when B is equipped with the subspace topology. Thus Lemma A.5.12 implies that there exist closed subsets F_1 and F_2 in X such that $f^{-1}(F) = A \cap F_1$ and $g^{-1}(F) = B \cap F_2$. Therefore, as A and B are closed in X, $f^{-1}(F)$ and $g^{-1}(F)$ are closed in X. Thus $h^{-1}(F) = f^{-1}(F) \cup g^{-1}(F)$ is closed in X. Therefore, as F was arbitrary, h is continuous by Theorem A.6.6. ∎

A.7 Homeomorphisms

With the construction of the objects and morphisms studied in this course complete, the next natural progression in mathematics is to define using ones morphisms when two objects are the same. As topological spaces are the objects in this course and continuous functions are the morphisms in this course, we study the following concept in order to determine the notion of when two topological spaces are the same.

Definition A.7.1. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A function $f : X \to Y$ is said to be a *homeomorphism* if f is bijective and both f and f^{-1} are continuous. Equivalently, a function $f : X \to Y$ is a homeomorphism if f is bijective and $U \in \mathcal{T}_X$ if and only if $f(U) \in \mathcal{T}_Y$.

Due to the above, we define the following notion.

Definition A.7.2. Two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are said to be *homeomorphic* if there is a homeomorphism from X to Y.

The reason why two homeomorphic topological spaces are 'the same' is because a bijection means the sets are the same (so $X = Y$ upto relabelling) and the continuity of the homeomorphism and its inverse implies the open sets are the same. This probably causes a modern mathematician to ask why we do not call homeomorphisms isomorphisms and why we do not call homeomorphic topological spaces isomorphic topological spaces. The only reason for this is tradition.

Of course, any notion of equality in mathematics must be an equivalence relation, we verify the following.

Proposition A.7.3. *Consider a set Φ of topological spaces and define a relation $\sim$ on Φ by $(X, \mathcal{T}_X) \sim (Y, \mathcal{T}_Y)$ if and only if $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are homeomorphic. Then $\sim$ is an equivalence relation.*

Proof. First, clearly $(X, \mathcal{T}_X) \sim (X, \mathcal{T}_X)$ via the identity map. Secondly, if $(X, \mathcal{T}_X) \sim (Y, \mathcal{T}_Y)$, then there is a homeomorphism $f : X \to Y$. As $f^{-1} : Y \to X$ is then a homeomorphism by definition, $(Y, \mathcal{T}_Y) \sim (X, \mathcal{T}_X)$.

Finally suppose $(X, \mathcal{T}_X) \sim (Y, \mathcal{T}_Y)$ and $(Y, \mathcal{T}_Y) \sim (Z, \mathcal{T}_Z)$. Thus there exists homeomorphisms $f : X \to Y$ and $g : Y \to Z$. Consider the map $h = g \circ f : X \to Z$. We claim that h is a homeomorphism. Indeed as the composition of bijections is a bijection, h is a bijection. Furthermore, by Theorem A.6.11 h is continuous being the composition of continuous functions. Finally, as $h^{-1} = f^{-1} \circ g^{-1}$, h^{-1} is the composition of continuous functions (as f and g are homeomorphisms) and thus continuous. Hence h is a homeomorphism so $(X, \mathcal{T}_X) \sim (Z, \mathcal{T}_Z)$ as desired. ∎

Now onto some examples.

Example A.7.4. Let $\mathbb{R}$ be equipped with its canonical topology and let $A = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ be equipped with the subspace topology inherited from $\mathbb{R}$. Then $\mathbb{R}$ and A are homeomorphic. Indeed consider the function $f : A \to \mathbb{R}$ defined

$$f(x) = \tan(x)$$

for all $x \in \mathbb{R}$. It is well-known that f is a continuous bijective function on A whose inverse, namely $f^{-1}(x) = \arctan(x)$ is also continuous. Hence $\mathbb{R}$ and A are homeomorphic.

As often a topological space is only homeomorphic to a subspace, we define the following.

Definition A.7.5. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. An function $f : X \to Y$ is said to be a *embedding* if $f : X \to f(X)$ is a homeomorphism when $f(X)$ is equipped with the subspace topology.

Example A.7.6. Let $\mathbb{R}^2$ and $\mathbb{R}^3$ be equipped with their Euclidean topologies, and let
$$S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$
(that is, S^2 is the boundary of the unit ball in $\mathbb{R}^3$) equipped with the subspace topology from $\mathbb{R}^3$. Since the Euclidean topologies on $\mathbb{R}^2$ and $\mathbb{R}^3$ are product topologies by Example A.3.15, Theorem A.4.22 implies a net converges in either of these spaces if and only if it converges entry-wise. Hence Proposition A.4.21 implies that a net converges in S^2 if and only if it converges entry-wise.

Consider the map $f : \mathbb{R}^2 \to S^2$ defined by
$$f(x, y) = \left(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \right)$$
for all $(x, y) \in \mathbb{R}^2$. It is not difficult to see that f is continuous by the net characterization of continuity from Theorem A.6.6. However, f is not bijective. Indeed the point $(0, 0, 1) \in S^2$ is not in the range of f.

Consider the function
$$g : S^2 \setminus \{(0, 0, 1)\} \to \mathbb{R}^2$$
defined by
$$g(x, y, z) = \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right)$$
for all $(x, y, z) \in S^2 \setminus \{(0, 0, 1)\}$. It is not difficult to see that if the codomain of f is restricted to $S^2 \setminus \{(0, 0, 1)\}$, then f and g are inverses to each other. Furthermore g is continuous by the net characterization of continuity from Theorem A.6.6. Hence f is an embedding of $\mathbb{R}^2$ into S^2 and $\mathbb{R}^2$ and $S^2 \setminus \{(0, 0, 1)\}$ are homeomorphic.

Example A.7.7. Let $\mathbb{R}$ and $\mathbb{R}^2$ be equipped with their Euclidean topologies, let $A = [0, 2\pi)$ equipped with the subspace topology induced by $\mathbb{R}$, and let
$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$
equipped with the subspace topology induced by $\mathbb{R}^2$. Since the Euclidean topologies on $\mathbb{R}$ and $\mathbb{R}^2$ are product topologies by Example A.3.15, Theorem

A.4.22 implies a net converges in either of these spaces if and only if it converges entry-wise. Hence Proposition A.4.21 implies that a net converges in S^1 if and only if it converges entry-wise.

Consider the map $f : A \to S^1$ defined by

$$f(x) = (\cos(x), \sin(x))$$

for all $x \in A$. It is elementary to see that f is a bijection. It is not difficult to see that f is continuous by the net characterization of continuity from Theorem A.6.6. However f^{-1} is not continuous. Indeed consider the set $U = [0, 1)$. Since U is an open subset of A as $U = A \cap (-\infty, 1)$, if f^{-1} were continuous, then $(f^{-1})^{-1}(U)) = f(U)$ would be open in S^1, so $S^1 \setminus f(U)$ would be closed in S^1. However, the sequence

$$\left(\left(\cos\left(2\pi - \frac{1}{n}\right), \sin\left(2\pi - \frac{1}{n}\right)\right)\right)_{n\geq 1}$$

is a net in $S^1 \setminus f(U)$ that converges to $(1, 0) \in f(U)$ thereby contradicting the fact that $S^1 \setminus f(U)$ was closed. Hence f cannot be continuous.

The reason that the map f in Example A.7.7 fails is that we have not placed the correct topology on the circle. If one wants a bijective map from a topological space to be a homeomorphism, we know exactly what topology to put on the codomain to ensure as the following result demonstrates.

Proposition A.7.8. *Let $(X, \mathcal{T}_X)$ be a topological space, let Y be a non-empty set, and let $q : X \to Y$ be a surjective map. Let*

$$\mathcal{T}_Y = \{A \subseteq Y \mid q^{-1}(A) \in \mathcal{T}_X\}.$$

Then $\mathcal{T}_Y$ is the finest topology on Y such that q is continuous. If q is bijective, then q is a homeomorphism.

Proof. First, we claim that $\mathcal{T}_Y$ is a topology. To see this, we note that $\emptyset, Y \in \mathcal{T}_Y$ since $q^{-1}(\emptyset) = \emptyset \in \mathcal{T}_X$ and $q^{-1}(Y) = X \in \mathcal{T}_X$ as q is surjective and as $\mathcal{T}_X$ is a topology. Moreover, since for all $\{U_\alpha\}_{\alpha\in I} \subseteq \mathcal{P}(Y)$ we have that

$$q^{-1}\left(\bigcup_{\alpha\in I} U_\alpha\right) = \bigcup_{\alpha\in I} q^{-1}(U_\alpha) \qquad \text{and} \qquad q^{-1}\left(\bigcap_{\alpha\in I} U_\alpha\right) = \bigcap_{\alpha\in I} q^{-1}(U_\alpha),$$

it is elementary to see that $\mathcal{T}_Y$ is closed under unions and finite intersections since $\mathcal{T}_X$ is. Hence $\mathcal{T}_Y$ is a topology.

To see that $q : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is continuous, we know that $q^{-1}(U) \in \mathcal{T}_X$ for all $U \in \mathcal{T}_Y$ by the definition of $\mathcal{T}_Y$. Hence q is continuous by definition. To see that $\mathcal{T}_Y$ is the finest topology on T so that q is continuous, suppose $\mathcal{T}$ is a topology on Y for which q is continuous. Then, by the definition of continuity, $q^{-1}(U) \in \mathcal{T}_X$ for all $U \in \mathcal{T}$. Therefore, by the definition of $\mathcal{T}_Y$, we obtain that $\mathcal{T} \subseteq \mathcal{T}_Y$. Hence $\mathcal{T}_Y$ is the finest topology on Y such that q is continuous.

Finally, to see that q is a homeomorphism when q is bijective, we need simply check that q^{-1} is continuous. To see this, let $V \in \mathcal{T}_X$ be arbitrary. Then $(q^{-1})^{-1}(V) = q(V)$ will be an element of $\mathcal{T}_Y$ by definition as $q^{-1}(q(V)) = V \in \mathcal{T}_X$. Therefore, as $V \in \mathcal{T}_X$ was arbitrary, q is a homeomorphism as desired. ∎

Due to the importance and usefulness of the above topology, we name this topology as follows.

Definition A.7.9. Let $(X, \mathcal{T}_X)$ be a topological space, let Y be a non-empty set, and let $q : X \to Y$ be a surjective map. The topology

$$\mathcal{T}_Y = \{A \subseteq Y \mid q^{-1}(A) \in \mathcal{T}_X\}$$

from Proposition A.7.8 is called the *quotient topology on Y induced by q.*

The reason we call the above the quotient topology is that one is really identifying all of the points in $q^{-1}(\{y\})$ as a single point for all $y \in Y$ and placing a topology on these collections of points based on the original topology; that is, we are taking a 'quotient' of a topological space by identify points. This idea is also motivated from geometry by 'cutting-and-pasting' to identify points to create new geometric objects. Before we formalize this and explore some examples, we first demonstrate, like with all things, how bases work in the quotient topology.

Proposition A.7.10. *Let $(X, \mathcal{T}_X)$ be a topological space, let Y be a non-empty set, let $q : X \to Y$ be a surjective map, and let $\mathcal{T}_Y$ be the quotient topology on Y induced by q. If $\mathcal{B}_X$ is a basis for $(X, \mathcal{T}_X)$, then*

$$\mathcal{B}_Y = \{A \subseteq Y \mid q^{-1}(A) \in \mathcal{B}_X\}$$

is a basis for $(Y, \mathcal{T}_Y)$.

Proof. To see that $\mathcal{B}_Y$ is a basis for $(Y, \mathcal{T}_Y)$, let $y \in Y$ and $U \in \mathcal{T}_Y$ be arbitrary. By the definition of the quotient topology, $q^{-1}(U) \in \mathcal{T}_X$. Therefore, as $q^{-1}(y) \in q^{-1}(U)$, the fact that $\mathcal{B}_X$ is a basis for $(X, \mathcal{T}_X)$ implies that there exists a $B \in \mathcal{B}_X$ such that $q^{-1}(y) \subseteq B \subseteq q^{-1}(U)$. Therefore, if $B' = q(B) \subseteq Y$, then $B' \in \mathcal{B}_Y$ by definition and $y \in q(B) = B' \subseteq U$. Therefore, as y and U were arbitrary, $\mathcal{B}_Y$ is a basis for $(Y, \mathcal{T}_Y)$. ∎

Example A.7.11. Let $A = [0, 2\pi)$ and let

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$$

equipped with the subspace topology induced by $\mathbb{R}^2$. As the intersection of open balls in $\mathbb{R}^2$ with S^1 yield open arcs on S^1, the open arcs on S^1 are a basis for the subspace topology on S^1.

Consider the map $q : S^1 \to A$ defined by

$$q\left((\cos(x), \sin(x))\right) = x$$

for all $x \in [0, 2\pi)$ and let $\mathcal{T}$ be the quotient topology on A induced by q. By the description of the basis of S^1 given above and since only arcs of arbitrarily small length around a point matter in forming a neighbourhood basis, we see for all $x \in (0, 2\pi)$ that

$$\{(x - \epsilon, x + \epsilon) \mid 0 < \epsilon < \min\{x, 2\pi - x\}\}$$

form a neighbourhood basis of x and that

$$\{[0, \epsilon) \cup (2\pi - \epsilon, 2\pi) \mid 0 < \epsilon < 2\pi\}$$

for a neighbourhood basis of 0 in the quotient topology.

Example A.7.12. Let $\mathbb{R}$ be equipped with its usual topology, let $Y = \{a, b, c\}$, and let $q : \mathbb{R} \to Y$ be defined by

$$q(x) = \begin{cases} a & \text{if } x < 0 \\ b & \text{if } x = 0 \\ c & \text{if } x > 0 \end{cases}.$$

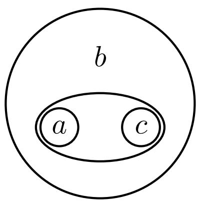
Let $\mathcal{T}$ be the quotient topology on Y induced by q. Notice for all $A \subseteq Y$ that

$$q^{-1}(A) = \begin{cases} \emptyset & \text{if } A = \emptyset \\ (-\infty, 0) & \text{if } A = \{a\} \\ \{0\} & \text{if } A = \{b\} \\ (0, \infty) & \text{if } A = \{c\} \\ (-\infty, 0] & \text{if } A = \{a, b\} \\ (-\infty, 0) \cup (0, \infty) & \text{if } A = \{a, c\} \\ [0, \infty) & \text{if } A = \{b, c\} \\ \mathbb{R} & \text{if } A = Y \end{cases}.$$

Therefore, by our knowledge of the open subsets of $\mathbb{R}$, we see that

$$\mathcal{T} = \{\emptyset, \{a\}, \{c\}, \{a, c\}, \{a, b, c\}\}.$$

Diagrammatically, the topology $\mathcal{T}$ is the following.



In Example A.7.12, one can think of a as $(-\infty, 0) \subseteq \mathbb{R}$, b as $\{0\} \subseteq \mathbb{R}$, and c as $(0, \infty) \subseteq \mathbb{R}$. That is, we can think of Y as a partition of $\mathbb{R}$ and the quotient topology on Y is then induced by this partition. To formalize this, we recall the definition a partition and how a partition produces a topological space.

Definition A.7.13. Let X be a non-empty set. A *partition of X* is a collection $\{X_\alpha\}_{\alpha\in I} \subseteq \mathcal{P}(X)$ such that $X = \bigcup_{\alpha\in I} X_\alpha$ and $X_\alpha \cap X_\beta = \emptyset$ if $\alpha, \beta \in I$ and $\alpha \neq \beta$.

Definition A.7.14. Let $(X, \mathcal{T})$ be a topological space, let X^* be a partition of X, and let $q : X \to X^*$ be the surjective map that maps each element $x \in X$ to the unique element in X^* containing x. The pair $(X^*, \mathcal{T}^*)$ where $\mathcal{T}^*$ is quotient topology on X^* induced by q is called a *quotient space*.

Example A.7.15. Let $\mathbb{R}$ be equipped with its usual topology and let

$$\mathcal{P} = \{\{x + 2\pi n \mid n \in \mathbb{Z}\} \mid x \in [0, 2\pi)\}.$$

If $(\mathbb{R}^*, \mathcal{T}^*)$ is the quotient space induced by $\mathcal{P}$, then $\mathbb{R}^*$ is in canonical bijective correspondence with $[0, 2\pi)$ by identifying $\{x + 2\pi n \mid n \in \mathbb{Z}\}$ for $x \in [0, 2\pi)$ with x. Under this identification, if $q : X \to X^*$ is the surjective map that maps each element $x \in \mathbb{R}$ to the unique element in $\mathbb{R}^*$ containing x, then we recall that

$$\mathcal{T}^* = \{A \subseteq [0, 2\pi) \mid q^{-1}(A) \text{ is open in } \mathbb{R}\}.$$

As the inverse image of every basis element exhibited in Example A.7.11 is a union of a countable number of open intervals (each of which is a translate of one fixed open interval by an integer multiple of 2π), we see that the topology from Example A.7.11 must be coarser than $\mathcal{T}^*$. Furthermore, given a subset $A \subseteq [0, 2\pi)$ we see that $q^{-1}(A)$ is 2π-periodic and will be open if and only if it is a union of open intervals and closed under 2π-periodicity. Hence $\mathcal{T}^*$ is precisely the topology on $[0, 2\pi)$ exhibited in Example A.7.11.

Example A.7.16. Let $\mathbb{R}^2$ be equipped with the Euclidean topology, let $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ equipped with the subspace topology induced by $\mathbb{R}^2$, and let

$$\mathcal{P} = \{\{(x, y)\} \mid x^2 + y^2 < 1\} \cup \{(x, y) \mid x^2 + y^2 = 1\}.$$

Consider the quotient space $(A^*, \mathcal{T}^*)$ and let $q : A \to A^*$ be the canonical surjective map. Then A^* is canonically in bijective correspondence with the shell

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\} \subseteq \mathbb{R}^3$$

via the the map $f : A^* \to S^2$ defined by

$$f(r\cos(\theta), r\sin(\theta)) = (\sin(r\pi)\cos(\theta), \sin(r\pi)\sin(\theta), \cos(r\pi))$$

for all $r \in [0,1]$ and $\theta \in [0, 2\pi)$. If S^2 is equipped with the subspace topology inherited from $\mathbb{R}^3$, then f is a homeomorphism from A^* to S^2. To see this, first notice that a subset of $\{(x,y) \mid x^2+y^2<1\}$ is open in A^* if and only if it is open in A and thus open in $\mathbb{R}^2$ by definition. Next, suppose U is an open set in A^* that contains $S^1 = \{(x,y) \mid x^2+y^2=1\}$. Thus $q^{-1}(U)$ is open in the subspace topology on A and contains S^1. Note for each point $(x,y) \in S^1$ there must exist a $\delta_{(x,y)} > 0$ such that $B((x,y), \delta_{(x,y)}) \cap A \subseteq q^{-1}(U)$. As S^1 is compact (see A.8), we may cover S^1 with a finite number of these balls in which case an $\epsilon > 0$ may be found so that $\{(x,y) \mid \epsilon < x^2+y^2 \leq 1\} \subseteq q^{-1}(U)$. Consequently, we see that $q^{-1}(U)$ is a union of a set of the form $\{(x,y) \mid \epsilon < x^2+y^2 \leq 1\}$ and a subset of $\{(x,y) \mid x^2+y^2<1\}$ that is open in A^*. It is then not difficult to see that the open sets in A^* are in bijective correspondence with those of S^2 via f. Hence f is a homeomorphism from A^* to S^2.

Example A.7.17. Let $\mathbb{R}^2$ be equipped with the Euclidean topology, let $A = [0,1]^2 \subseteq \mathbb{R}^2$ equipped with the subspace topology induced by $\mathbb{R}^2$, and let $\mathcal{P}$ be the union of

$$\begin{aligned}
&\{\{(x,y)\} \mid x,y \in (0,1)\},\\
&\{\{(x,0),(x,1)\} \mid x \in (0,1)\},\\
&\{\{(0,y),(1,y)\} \mid y \in (0,1)\}, \text{ and}\\
&\{(0,0),(1,0),(0,1),(1,1)\}.
\end{aligned}$$

Consider the quotient space $(A^*, \mathcal{T}^*)$ and let $q : A \to A^*$ be the canonical surjective map. Then A^* is canonically in bijective correspondence to a torus in $\mathbb{R}^3$ in such a way that that $\mathcal{T}^*$ corresponds to the subspace topology on the torus inherited from $\mathbb{R}^3$. The details are similar to Example A.7.16.

To better understand functions, continuous functions, and homeomorphisms on quotient spaces, we give a name to the maps under consideration when definition a quotient topology.

Definition A.7.18. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A map $q : X \to Y$ is said to be a *quotient map* if q is surjective and a set $U \subseteq Y$ is open if and only if $p^{-1}(U) \in \mathcal{T}_X$; that is, if $\mathcal{T}_Y$ is the quotient topology on Y induced by q.

Clearly quotient maps are continuous maps by the definition of a quotient map and by the definition of a continuous function. In addition, of course a quotient of a quotient is still a quotient.

Lemma A.7.19. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$, and $(Z, \mathcal{T}_Z)$ be topological spaces, let $q : X \to Y$ be a quotient map, and let $p : Y \to Z$ be a quotient map. Then $p \circ q : X \to Z$ is a quotient map.*

Proof. To see that $p \circ q$ is a quotient map, we first note that quotient maps are continuous by definition. Therefore $p \circ q$ is a composition of continuous maps and thus continuous. Hence if $U \in \mathcal{T}_Z$, then $(p \circ q)^{-1}(U)$ is open in $\mathcal{T}_X$ by continuity.

Conversely, let $U \subseteq Z$ such that $(p \circ q)^{-1}(U) = q^{-1}(p^{-1}(U))$ is open in $(X, \mathcal{T}_X)$ be arbitrary. Since q is a quotient map, $q^{-1}(p^{-1}(U))$ being open in $(X, \mathcal{T}_X)$ implies that $p^{-1}(U)$ is open in $(Y, \mathcal{T}_Y)$ by the definition of a quotient map. Therefore, since p is a quotient map, $p^{-1}(U)$ being open in $(Y, \mathcal{T}_Y)$ implies that U is open in $(Z, \mathcal{T}_Z)$ by the definition of a quotient map. Therefore, as U was arbitrary, $p \circ q$ is a quotient map. ■

One of the main reason quotient spaces are nice is that certain maps factor over quotients and preserve topological properties.

Theorem A.7.20. *Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$, and $(Z, \mathcal{T}_Z)$ be topological spaces, let $q : X \to Z$ be a quotient map, and let $g : X \to Y$ be a map that is constant on $q^{-1}(\{z\})$ for each $z \in Z$. Then there exists a unique map $f : Z \to Y$ such that $g = f \circ q$.*

$$\begin{tikzcd} X \arrow[d, "q"'] \arrow[dr, "g"] & \\ Z \arrow[r, "f"'] & Y \end{tikzcd}$$

The map f is continuous if and only if g is continuous. Furthermore, f is a quotient map if and only if g is a quotient map.

Proof. First, since g is constant on $q^{-1}(\{z\})$ for each $z \in Z$, we define $f : Z \to Y$ by setting $f(z)$ for each $z \in Z$ to be the unique value of g obtained on $q^{-1}(\{z\})$, then f is a well-defined function such that $g = f \circ q$ as desired. Furthermore, as this is clearly the only way to define f so that $g = f \circ q$ as q is surjective, uniqueness has been obtained.

Next, clearly if f is a continuous function than g is a continuous function since quotient maps are continuous and the composition of continuous functions is continuous. Conversely, suppose that g is continuous and let U be an arbitrary open set in $(Y, \mathcal{T}_Y)$. Hence $g^{-1}(U) = (f \circ q)^{-1}(U) = q^{-1}(f^{-1}(U))$ must be an open set in X. However, as q is a quotient map, $q^{-1}(f^{-1}(U)) \in \mathcal{T}_X$ implies $f^{-1}(U) \in \mathcal{T}_Z$. Hence, as U was arbitrary f is continuous as desired.

Finally, if f is a quotient map, then g is a quotient map as the composition of quotient maps is a quotient map. Conversely, suppose that g is a quotient map. Thus a set $U \subseteq Y$ is such that $U \in \mathcal{T}_Y$ if and only if $g^{-1}(U) = (f \circ q)^{-1}(U) = q^{-1}(f^{-1}(U)) \in \mathcal{T}_X$. However, as q is a quotient map, $q^{-1}(f^{-1}(U)) \in \mathcal{T}_X$ if and only if $f^{-1}(U) \in \mathcal{T}_Z$. Hence f is a quotient map by definition. ■

Using Theorem A.7.20, we obtain a better understanding of quotient spaces obtained by partitioning based on a surjective continuous linear map. In particular, every surjective continuous linear map factors through a quotient space.

Corollary A.7.21. *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, let $g : X \to Y$ be a surjective continuous linear map, let*

$$X^* = \{g^{-1}(\{y\}) \mid y \in Y\}$$

equipped with the quotient topology, and let $q : X \to X^*$ be the surjective map from Definition A.7.14 that maps each element $x \in X$ to the unique element in X^* containing x. Then there exists a unique bijective, continuous map $f : X^* \to Y$ such that $g = f \circ q$.

$$\begin{tikzcd} X \arrow[d, "q"'] \arrow[dr, "g"] & \\ X^* \arrow[r, "f"'] & Y \end{tikzcd}$$

Furthermore X^* is Hausdorff if $(Y, \mathcal{T}_Y)$ is Hausdorff. Finally f is a homeomorphism if and only if g a quotient map.

Proof. First, the fact that f exists is unique, and is continuous follows from Theorem A.7.20 as g is continuous. Furthermore, since g is surjective and $g = f \circ q$, f is surjective. To see that f is injective, suppose $x_1, x_2 \in X^*$ are such that $f(x_1) = f(x_2)$. As q is surjective, there exists $x_1', x_2' \in X$ such that $q(x_1') = x_1$ and $q(x_2') = x_2$. Hence

$$g(x_1') = f(q(x_1')) = f(x_1) = f(x_2) = f(q(x_2')) = g(x_2').$$

Therefore, by the definition of X^* we must have $x_1 = q(x_1') = q(x_2') = x_2$. Thus f is bijective.

Next, suppose $(Y, \mathcal{T}_Y)$ is Hausdorff. To see that X^* is Hausdorff, let $x_1, x_2 \in X^*$ be arbitrary points such that $x_1 \neq x_2$. Then, as f is bijective, $f(x_1) \neq f(x_2)$. Hence, as $(Y, \mathcal{T}_Y)$ is Hausdorff, there exists open sets $U_1, U_2 \in \mathcal{T}_Y$ such that $f(x_1) \in U_1$, $f(x_2) \in U_2$, and $U_1 \cap U_2 = \emptyset$. Therefore, since f is a continuous bijection, $V_1 = f^{-1}(U_1)$ and $V_2 = f^{-1}(U_2)$ are open sets in X^* such that $x_1 \in V_1$, $x_2 \in V_2$, and $V_1 \cap V_2 = \emptyset$. Therefore, as x_1 and x_2 were arbitrary, X^* is Hausdorff.

To see the last part of the statement, we note that if g is a quotient map, then f is a quotient map by Theorem A.7.20. Therefore, as f is a bijective quotient map, f is a homeomorphism by definition.

Finally, suppose f is a homeomorphism. To see that g is a quotient map, we first notice g is surjective and, since g is continuous, that if $U \in \mathcal{T}_Y$

then $g^{-1}(U) \in \mathcal{T}_X$. Thus, to complete the proof that g is a quotient map, let $U \subseteq Y$ be an arbitrary set such that $g^{-1}(U) \in \mathcal{T}_X$. Hence $g^{-1}(U) = q^{-1}(f^{-1}(U)) \in \mathcal{T}_X$. Therefore, as q is a quotient map, $f^{-1}(U)$ is open in X^*. However, since f is a homeomorphism, this implies that $U \in \mathcal{T}_Y$. Therefore, as U was arbitrary, g is a quotient map. ∎

A.8 Compact Sets

One of the most important collection of topological spaces are those that are compact. The notion of compactness follows by asking that we can extract a finite open cover of our space from any open cover we may wish to consider.

Definition A.8.1. Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$. Sets $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ are said to be an *open cover* of A if each $U_\alpha \in \mathcal{T}$ for all $\alpha \in I$ and $A \subseteq \bigcup_{\alpha \in I} U_\alpha$.

A *subcover of A from* $\{U_\alpha\}_{\alpha \in I}$ is any collection $\{U_\alpha\}_{\alpha \in J}$ where $J \subseteq I$ such that $A \subseteq \bigcup_{\alpha \in J} U_\alpha$

Definition A.8.2. A topological space $(X, \mathcal{T})$ is said to be *compact* if every open cover of $(X, \mathcal{T})$ contains a finite subcover; that is, if $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ are such that $X = \bigcup_{\alpha \in I} U_\alpha$, then there exists $J \subseteq I$ such that J is finite and $X = \bigcup_{\alpha \in J} U_\alpha$.

Of course, we have some trivial examples

Example A.8.3. Technically the empty set is compact as every open cover has a subcover consisting of one element.

Example A.8.4. The trivial topology on a set X is always compact as the only open covers of X will be $\{X\}$ and $\{X, \emptyset\}$.

Example A.8.5. Let $(X, \mathcal{T})$ be a topological space with X finite. Then $(X, \mathcal{T})$ is compact as $\mathcal{T} \subseteq \mathcal{P}(X)$ is finite.

Example A.8.6. Let X be an infinite set and let $\mathcal{T}$ be the discrete topology on X. Then $(X, \mathcal{T})$ is not compact as $\{\{x\}\}_{x \in X}$ is an open cover with no finite subcovers.

To obtain more examples of compact topological spaces, we turn our attention to subsets of $\mathbb{R}$. Of course we have the following.

Example A.8.7. If $\mathbb{R}$ is equipped with its canonical topology, then $\mathbb{R}$ is not compact. Indeed $\mathcal{U} = \{(n-1, n+1) \mid n \in \mathbb{Z}\}$ is an open cover of $\mathbb{R}$ with no finite subcovers as each element of $\mathbb{Z}$ is covered by a unique element of $\mathcal{U}$.

In order to determine which subsets of $\mathbb{R}$ are compact when equipped with the subspace topology, we note the following.

Lemma A.8.8. *Let $(X, \mathcal{T})$ be a topological space and let Y be a subspace of $(X, \mathcal{T})$. Then Y is compact if and only if every open cover of Y in $(X, \mathcal{T})$ has a finite subcover.*

Proof. For simplicity, let $\mathcal{T}_Y$ denote the subspace topology on Y inherited from $(X, \mathcal{T})$.

To begin, suppose $(Y, \mathcal{T}_Y)$ is compact. To see the result, let $\{U_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of Y in $(X, \mathcal{T})$. Hence $\{Y \cap U_\alpha\}_{\alpha \in I}$ is an open cover of Y in $(Y, \mathcal{T}_Y)$ so the fact that $(Y, \mathcal{T}_Y)$ is compact implies there exists a finite subset $J \subseteq I$ such that $\{Y \cap U_\alpha\}_{\alpha \in J}$ is an open cover of Y in $(Y, \mathcal{T}_Y)$. Hence clearly $\{U_\alpha\}_{\alpha \in J}$ is a finite open subcover of Y from $\{U_\alpha\}_{\alpha \in I}$. Therefore, as $\{U_\alpha\}_{\alpha \in I}$ was arbitrary, the claim follows.

Conversely, suppose that every open cover of Y in $(X, \mathcal{T})$ has a finite subcover. To see that $(Y, \mathcal{T}_Y)$ is compact, let $\{V_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of Y in $(Y, \mathcal{T}_Y)$. By the definition of the subspace topology there exists $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ such that $V_\alpha = Y \cap U_\alpha$ for all $\alpha \in I$. Hence $\{U_\alpha\}_{\alpha \in I}$ is an open cover of Y in $(X, \mathcal{T})$, which then must have a finite subcover $\{U_\alpha\}_{\alpha \in J}$ of Y by assumption. Hence $\{V_\alpha\}_{\alpha \in J}$ is a finite open subcover of Y from $\{V_\alpha\}_{\alpha \in I}$. Therefore, since $\{V_\alpha\}_{\alpha \in I}$ was arbitrary, $(Y, \mathcal{T}_Y)$ is compact. ∎

Example A.8.9. The subset $X = \{0\} \cup \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\} \subseteq \mathbb{R}$ is a compact subspace of $\mathbb{R}$. To see this, suppose that $\{U_\alpha\}_{\alpha \in I}$ is an open cover of X using open subsets from $\mathbb{R}$. Hence there exists an $\alpha_0 \in I$ such that $0 \in U_{\alpha_0}$. Since U_{α_0} is open, there exists an $\epsilon > 0$ such that $(-\epsilon, \epsilon) \subseteq U_{\alpha_0}$. Since X contains only a finite number of elements outside of $(-\epsilon, \epsilon)$, X contains only a finite number of element outside of U_{α_0}. Thus we can write $X \setminus U_{\alpha_0} = \{x_1, \ldots, x_m\}$ for some $m \in \mathbb{N}$. Since $\{U_\alpha\}_{\alpha \in I}$ is an open cover of X, for each $k \in \{1, \ldots, m\}$ there exists an $\alpha_k \in I$ such that $x_k \in U_{\alpha_k}$. Hence $\{U_{\alpha_k}\}_{k=0}^m$ is a finite subcover of X from $\{U_\alpha\}_{\alpha \in I}$. Hence, as $\{U_\alpha\}_{\alpha \in I}$ was arbitrary, X is compact.

Example A.8.10. The subset $X = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\} \subseteq \mathbb{R}$ of $\mathbb{R}$ is not compact. Indeed if $U_n = \left(\frac{1}{n}, 1\right)$ for all $n \in \mathbb{N}$, then $\{U_n\}_{n=1}^\infty$ is an open cover of X. However, clearly $\{U_n\}_{n=1}^\infty$ does not have a finite subcover as if $n_1, \ldots, n_m \in \mathbb{N}$ then $\bigcup_{k=1}^m U_{n_k} = \left(\frac{1}{\max\{n_1, \ldots, n_m\}}, 1\right)$ which does not contain all of X since $\lim_{n \to \infty} \frac{1}{n} = 0$.

Remark A.8.11. It is clear in the above example that the reason why X is not compact was that X was not closed. However, for a general topological space $(X, \mathcal{T})$, a subspace A of $(X, \mathcal{T})$ may still be compact even if A is not closed in $(X, \mathcal{T})$. Indeed, if X is finite, then every subspace of $(X, \mathcal{T})$ is compact as every subspace topology consists only of a finite number of sets. As there are clearly examples of topologies on finite sets such that not every set is closed, we have demonstrated our claim.

The reason why the set in Example A.8.10 is not compact because it was not closed follows from the fact that $\mathbb{R}$ is Hausdorff. In particular, mixing the notions of Hausdorff and compactness yields some powerful results. Recall that a topological space $(X, \mathcal{T})$ is Hausdorff means that $(X, \mathcal{T})$ has a lot of open sets to separate points (i.e. it is "close" to the discrete topology) whereas $(X, \mathcal{T})$ is compact means that $(X, \mathcal{T})$ does not have too many open sets as every open cover has a finite subcover (i.e. it is "close" to the trivial topology). It is this Goldilocks zone that makes compact Hausdorff topological spaces some of the nicest topological spaces to study.

In order to study compact Hausdorff topological spaces, we note the following incredibly useful lemma that lets us separate points from compact subsets.

Lemma A.8.12. *Let $(X, \mathcal{T})$ be a Hausdorff space, let Y be a compact subspace of X, and let $x_0 \in X \setminus Y$. Then there exists $U, V \in \mathcal{T}$ such that $x_0 \in U$, $Y \subseteq V$, and $U \cap V = \emptyset$.*

Proof. Since $(X, \mathcal{T})$ is Hausdorff and $x_0 \in X \setminus Y$, for each $y \in Y$ there exists $U_y, V_y \in \mathcal{T}$ such that $x_0 \in U_y$, $y \in V_y$, and $U_y \cap V_y = \emptyset$. Hence $\{V_y\}_{y \in Y}$ is an open cover of Y. Therefore, as Y is a compact subspace of X, Lemma A.8.8 implies there exists an $n \in \mathbb{N}$ and $y_1, y_2, \ldots, y_n \in Y$ such that $\{V_{y_k}\}_{k=1}^n$ is an open cover of Y. Let

$$U = \bigcap_{k=1}^n U_{y_k} \qquad \text{and} \qquad V = \bigcup_{k=1}^n V_k.$$

Clearly $Y \subseteq V$ by construction. Furthermore, as $x_0 \in U_{y_k}$ for all $k \in \{1, \ldots, n\}$, $x_0 \in U$. Finally, since $U_y \cap V_y = \emptyset$ for all $y \in Y$, we obtain that $U \cap V = \emptyset$ as desired. ∎

Using Lemma A.8.12, we can formalize the problem with Example A.8.10.

Theorem A.8.13. *Every compact subspace of a Hausdorff topological space is closed.*

Proof. Let $(X, \mathcal{T})$ be a Hausdorff topological space and let Y be a compact subspace of X. To see that Y is closed in $(X, \mathcal{T})$, it will be demonstrated that $X \setminus Y$ is open. To see that $X \setminus Y$ is open, let $x_0 \in X \setminus Y$ be arbitrary. By Lemma A.8.12, there exists open sets $U, V \in \mathcal{T}$ such that $x_0 \in U$, $Y \subseteq V$, and $U \cap V = \emptyset$. Hence U is a neighbourhood of x_0 that is contained in $X \setminus Y$. Therefore, as $x_0 \in X \setminus Y$ was arbitrary, $X \setminus Y$ is open in $(X, \mathcal{T})$. Hence Y is closed in $(X, \mathcal{T})$ as desired. ∎

In fact, Theorem A.8.13 has somewhat of a converse in compact topological spaces.

Theorem A.8.14. *Every closed subspace of a compact topological space is compact.*

Proof. Let $(X, \mathcal{T})$ be a compact topological space and let F be a closed subspace of $(X, \mathcal{T})$. To see that F is compact, we will verify the conditions of Lemma A.8.8. Thus let $\{U_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of F from $(X, \mathcal{T})$. Hence, as F is closed in $(X, \mathcal{T})$, $\{X \setminus F\} \cup \{U_\alpha\}_{\alpha \in I}$ is an open cover of $(X, \mathcal{T})$. Therefore, as $(X, \mathcal{T})$ is compact, there exists a finite subset $J \subseteq I$ such that $\{X \setminus F\} \cup \{U_\alpha\}_{\alpha \in J}$ is an open cover of $(X, \mathcal{T})$. Therefore, since $X \setminus F$ is disjoint from F, $\{U_\alpha\}_{\alpha \in I}$ is a finite subcover F from $\{U_\alpha\}_{\alpha \in I}$. Hence Lemma A.8.8 implies that F is a compact subspace of $(X, \mathcal{T})$. ∎

Combining Theorem A.8.13 and Theorem A.8.14, we can construct new compact subspaces from other compact subspaces.

Corollary A.8.15. *The arbitrary non-empty intersection of compact subspaces of a Hausdorff topological space is compact.*

Proof. Let $(X, \mathcal{T})$ be a Hausdorff topological space and let $\{K_\alpha\}_{\alpha \in I}$ be compact subspaces of $(X, \mathcal{T})$ with I non-empty. By Theorem A.8.13, K_α is closed in $(X, \mathcal{T})$ for all $\alpha \in I$. Hence $K = \bigcap_{\alpha \in I} K_\alpha$ is closed in $(X, \mathcal{T})$. As I is non-empty, K is a closed subset of K_α for all $\alpha \in I$ and thus a compact subspace K_α for all $\alpha \in I$ by Theorem A.8.14. ∎

Remark A.8.16. Note Corollary A.8.15 does not extend to (even the finite intersection) of compact subspaces of non-Hausdorff topological spaces. For such an example, let $X = \mathbb{N}$ and let

$$\mathcal{T} = \{A \mid A \subseteq \mathbb{N} \setminus \{1, 2\}\} \cup \{\mathbb{N}\} \cup \{\mathbb{N} \setminus \{1\}\} \cup \{\mathbb{N} \setminus \{2\}\}.$$

It is not difficult to verify that $\mathcal{T}$ is a topology on X. Furthermore, if $K_1 = \mathbb{N} \setminus \{1\}$ and $K_2 = \mathbb{N} \setminus \{2\}$, it is not difficult to verify that K_1 and K_2 are compact subspaces of X as any open cover of K_1 must include either K_1 or $\mathbb{N}$ (both of which are finite subcovers of K_1) and any open cover of K_2 must include either K_2 or $\mathbb{N}$ (both of which are finite subcovers of K_2). However, $K_1 \cap K_2 = \mathbb{N} \setminus \{1, 2\}$ is clearly not compact as $\{\{n\} \mid n \in \mathbb{N} \setminus \{1, 2\}\}$ is an open cover of $K_1 \cap K_2$ with no finite subcovers.

Corollary A.8.17. *The finite union of compact subspaces of a topological space is compact.*

Proof. Let $\{K_k\}_{k=1}^n$ be compact subspaces of a topological space $(X, \mathcal{T})$ and let $K = \bigcup_{k=1}^n K_k$. To see that K is compact in $(X, \mathcal{T})$, let $\mathcal{U} = \{U_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of K. Hence $\mathcal{U}$ is an open cover of K_k for all $k \in \{1, \ldots, n\}$. Therefore, since K_k is compact for all $k \in \{1, \ldots, n\}$, there exists a finite subset $J_k \subseteq I$ such that $\{U_\alpha\}_{\alpha \in J_k}$ is an open cover off K_k. Thus if $J = \bigcup_{k=1}^n J_k$, then J is a finite subset of I and $\{U_\alpha\}_{\alpha \in J}$ is an open cover of K. Therefore, as $\mathcal{U}$ was arbitrary, K is compact as desired. ∎

Remark A.8.18. Note Corollary A.8.17 does not extend to arbitrary unions of compact subspaces. Indeed clearly $X = \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$ is a union of compact subsets of $\mathbb{R}$ as every singleton in $\mathbb{R}$ is trivially compact. However X is not compact by Example A.8.10.

However, there more obstructions for a subset of $\mathbb{R}$ to be compact.

Example A.8.19. Let $X = \mathbb{Z} \subseteq \mathbb{R}$ equipped with the subspace topology inherited from the canonical topology on $\mathbb{R}$. Clearly X is a closed subset of $\mathbb{R}$ since any convergent net from X must eventually be constant as $\mathrm{dist}(n, X \setminus \{n\}) = 1$ for all $n \in \mathbb{N}$. However, we claim that X is not compact. Indeed if $U_n = (-n, n)$ for each $n \in \mathbb{N}$, then $\mathcal{U} = \{U_n\}_{n=1}^\infty$ is an open cover of X. However, $\mathcal{U}$ does not have a finite subcover since $U_n \subseteq U_{n+1}$ for all $n \in \mathbb{N}$ so that $\mathcal{U}$ is closed under unions, and since each element of $\mathcal{U}$ contains only a finite number of points in the infinite set X.

It is not difficult to see that the set in Example A.8.19 is not compact as its elements get arbitrary far away from 0. To give a name to this issue, we define the following property for metric spaces.

Definition A.8.20. A subset A of metric space (X, d) is said to be *bounded* if there exists an $M \geq 0$ such that

$$\{d(a_1, a_2) \mid a_1, a_2 \in A\} \subseteq [0, M].$$

There are many ways to characterize boundedness in a metric space.

Lemma A.8.21. *Let (X, d) be a metric space and let $A \subseteq X$ be non-empty. The following are equivalent:*

(i) A is bounded.

(ii) For each $a_0 \in A$, $A \subseteq B_d(a_0, R)$ for some $R > 0$.

(iii) For an $a_0 \in A$, $A \subseteq B_d(a_0, R)$ for some $R > 0$.

Proof. To see that (i) implies (ii), suppose A is bounded. Thus there exists an $M > 0$ such that

$$\{d(a_1, a_2) \mid a_1, a_2 \in A\} \subseteq [0, M].$$

Hence (ii) follows by taking $R = M$ for each $a_0 \in A$.

Clearly (ii) implies (iii). To see that (iii) implies (i), let $a_0 \in A$ and $R > 0$ be such that $A \subseteq B_d(a_0, R)$. Hence for all $a_1, a_2 \in A$,

$$d(a_1, a_2) \leq d(a_1, a_0) + d(a_0, a_2) \leq R + R = 2R.$$

Hence

$$\{d(a_1, a_2) \mid a_1, a_2 \in A\} \subseteq [0, 2R]$$

so A is bounded by definition. ∎

Using the same idea as Example A.8.19, we have the following.

Theorem A.8.22. *Every compact metric space is bounded.*

Proof. Let (X, d) be a compact metric space. To see that (X, d) is bounded, fix a point $x_0 \in X$. For each $n \in \mathbb{N}$, consider the open set $U_n = B_d(x_0, n)$. Since for all $x \in X$ there exists an $m \in \mathbb{N}$ such that $d(x, x_0) < m$, we see that $\bigcup_{n=1}^\infty U_n = X$. Hence $\{U_n\}_{n=1}^\infty$ is an open cover of (X, d). Therefore, since (X, d) is compact, there exists $n_1, \ldots, n_q \in \mathbb{N}$ such that $X = \bigcup_{k=1}^q U_{n_k}$. If $N = \max\{n_1, \ldots, n_q\}$, we clearly obtain that $X = B_d(x_0, N)$. Thus (X, d) is bounded by Lemma A.8.21. ∎

Combining Theorem A.8.13 and Theorem A.8.22, every compact subspace of $\mathbb{R}$ must be closed and bounded. We desire to prove the converse to this statement. To simplify notation, we define the following.

Definition A.8.23. Let (X, d) be a metric space and let $A \subseteq X$ be non-empty. The *diameter of* A, denoted $\mathrm{diam}(A)$, is defined to be

$$\mathrm{diam}(A) = \sup(\{d(a_1, a_2) \mid a_1, a_2 \in A\}) \subseteq [0, \infty].$$

Example A.8.24. In $\mathbb{R}$ equipped with its canonical metric

$$\mathrm{diam}((0, 1)) = \mathrm{diam}([0, 1]) = 1$$

whereas $\mathrm{diam}(\mathbb{R}) = \infty$.

Theorem A.8.25 (The Heine-Borel Theorem). *Let $K \subseteq \mathbb{K}^n$. Then K is compact in $(\mathbb{K}^n, \|\cdot\|_\infty)$ if and only if K is closed and bounded.*

Proof. First, suppose K is compact. As any subspace of a metric space (X, d) has topology induced by a metric that was induced from d by Proposition A.3.5, Theorem A.8.22 implies that K is bounded in $(\mathbb{K}^n, \|\cdot\|_\infty)$. Furthermore, as the subspace of any Hausdorff topological space is Hausdorff, Theorem A.8.13 implies that K is closed.

Conversely, let K be closed and bounded subspace of $(\mathbb{K}^n, \|\cdot\|_\infty)$. Suppose to the contrary that K is not compact. Hence there exists an open cover $\{U_\alpha\}_{\alpha \in I}$ of K that has no finite subcover.

Since K is bounded, there exists an $M \in \mathbb{R}$ such that

$$K \subseteq [-M, M] \times \cdots \times [-M, M]$$

when $\mathbb{K} = \mathbb{R}$, and

$$K \subseteq \{(a_1 + b_1 i, \ldots, a_n + b_n i) \mid a_i, b_j \subseteq [-M, M]\}$$

when $\mathbb{K} = \mathbb{C}$. We will proceed with the proof where $\mathbb{K} = \mathbb{R}$ as the case where $\mathbb{K} = \mathbb{C}$ follows by the same arguments using $2n$ in place of n.

Divide $[-M, M]^n$ into 2^n closed balls with side-lengths M. To be specific, for all $q_1, \ldots, q_n \in \{0, 1\}$ let

$$J_{q_1,\ldots,q_n} = [-M + Mq_1, Mq_1] \times \cdots \times [-M + Mq_n, Mq_n].$$

Clearly each $J_{q_1,\ldots,q_n}$ is closed and the union of all possible $J_{q_1,\ldots,q_n}$s contains K. Therefore, since $\{U_\alpha\}_{\alpha\in I}$ does not have a finite subcover of K, there must exist one of these $J_{q_1,\ldots,q_n}$s such that $\{U_\alpha\}_{\alpha\in I}$ does not have a finite subcover of $K \cap J_{q_1,\ldots,q_n}$ (as there are a finite number of $J_{q_1,\ldots,q_n}$s). Denote this $J_{q_1,\ldots,q_n}$ by B_1 and notice $\operatorname{diam}(B_1) = M$.

Suppose for each $k \in \mathbb{N}$ we have constructed closed balls $B_1, \ldots, B_k$ such that $B_{j+1} \subseteq B_j$, $\operatorname{diam}(B_j) = \frac{1}{2^j}M$, and $\{U_\alpha\}_{\alpha\in I}$ does not have a finite subcover of $B_j \cap K$ for all $j \in \{1, \ldots, k-1\}$. By repeating the above process on B_k, there exists a closed ball $B_{k+1} \subseteq B_k$ such that $\operatorname{diam}(B_{k+1}) = \frac{1}{2^{k+1}}M$ and such that $\{U_\alpha\}_{\alpha\in I}$ does not have a finite subcover of $B_{k+1} \cap K$. Thus, by repeating this process ad infinitum, we obtain a collection $\{B_k\}_{k=1}^\infty$ of closed balls of $\mathbb{K}^n$ such that $B_{k+1} \subseteq B_k$, $\operatorname{diam}(B_k) = \frac{1}{2^k}M$, and $\{U_\alpha\}_{\alpha\in I}$ does not have a finite subcover of $V_k \cap K$ for all $k \in \mathbb{N}$ (and thus $B_k \cap K \neq \emptyset$ for all $k \in \mathbb{N}$).

For each $k \in \mathbb{N}$, let $x_k \in B_k \cap K$. Then, as $\operatorname{diam}(B_k \cap K) \subseteq \operatorname{diam}(B_k) \leq \frac{1}{2^k}M$, we see for all $n \geq m \geq N$ that $x_n, x_m \in B_N \cap K$ so

$$d(x_n, x_m) \leq \frac{1}{2^N}M$$

Therefore, since $\lim_{N\to\infty} M\frac{1}{2^N} = 0$, we see that $(x_n)_{n\geq1}$ is a Cauchy sequence. Hence, as $\mathbb{K}^n$ is complete, there exists an $x_0 \in \mathbb{K}^n$ such that $\lim_{n\to\infty} x_n = x_0$. Moreover, since $x_n \in B_m \cap K$ for all $n \geq m$, the fact that $B_m \cap K$ is closed implies that $x_0 \in B_m \cap K$ for all $m \in \mathbb{N}$. Hence

$$Y = \bigcap_{k=1}^\infty (I_k \cap K) \neq \emptyset.$$

We claim that Y has exactly one element. Indeed if $x, y \in Y$ then $x, y \in B_k$ for all $k \in \mathbb{N}$ so $d(x, y) \leq \operatorname{diam}(B_k) = \frac{1}{2^k}M$ for all $k \in \mathbb{N}$ which implies $d(x, y) = 0$, or, equivalently, $x = y$. Hence Y contains exactly one point, say z.

By construction $z \in K$. Therefore, as $\{U_\alpha\}_{\alpha\in I}$ is an open cover of K, there exists an $\alpha_0 \in I$ such that $z \in U_{\alpha_0}$. Thus, since U_{α_0} is open, there exists an $\epsilon > 0$ such that $B(z, \epsilon) \subseteq U_{\alpha_0}$. Since $\operatorname{diam}(B_k) = \frac{1}{2^k}M$ for all $k \in \mathbb{N}$, there exists a $k_0 \in \mathbb{N}$ such that $\operatorname{diam}(B_{k_0}) < \epsilon$. Therefore, as $z \in B_{k_0}$ we obtain for all $x \in B_{k_0}$ that $d(z, x) < \epsilon$ so $x \in B(z, \epsilon) \subseteq U_{\alpha_0}$ for all $x \in B_{k_0}$. This implies $B_{k_0} \cap K \subseteq I_{k_0} \subseteq B(z, \epsilon) \subseteq U_{\alpha_0}$ which contradicts the fact that $\{U_\alpha\}_{\alpha\in I}$ did not have a finite subcover of $B_{k_0} \cap K$. As we have obtained a contradiction, it must be the case that K is compact. ∎

Now that we have the Heine-Borel Theorem (Theorem A.8.25) and thus a plethora of examples of compact topological spaces, we return to our initial motivation for compact topological spaces; namely a generalization to the Extreme Value Theorem to topological spaces. To obtain this characterization knowing that every finite closed interval in $\mathbb{R}$ is compact, we note the following exemplary property of compact topological spaces.

Theorem A.8.26 (The Extreme Value Theorem). *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and let $f : X \to Y$ be continuous. If $(X, \mathcal{T}_X)$ is compact, then $f(X)$ is a compact subspace of $(Y, \mathcal{T}_Y)$.*

Proof. To see that $f(X)$ is compact, let $\{U_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of $f(X)$ in $(Y, \mathcal{T}_Y)$. Therefore $\{f^{-1}(U_\alpha)\}_{\alpha \in I}$ is an open cover of $(X, \mathcal{T}_X)$. Hence, as $(X, \mathcal{T}_X)$ is compact, there exists a finite subset $J \subseteq I$ such that $\{f^{-1}(U_\alpha)\}_{\alpha \in J}$ is an open cover of $(X, \mathcal{T}_X)$. Therefore $f(X) \subseteq \bigcup_{\alpha \in J} U_\alpha$ so $\{U_\alpha\}_{\alpha \in J}$ is a finite subcover of $f(X)$ from $\{U_\alpha\}_{\alpha \in I}$ Therefore, as $\{U_\alpha\}_{\alpha \in I}$ was arbitrary, $f(X)$ is compact. ∎

Theorem A.8.26 has some wide-reaching implications.

Theorem A.8.27 (The Extreme Value Theorem). *Let $(X, \mathcal{T})$ be a compact topological space and let $f : X \to \mathbb{R}$ be continuous. Then there exists points $x_1, x_2 \in X$ such that $f(x_1) \leq f(x) \leq f(x_2)$ for all $x \in X$.*

Proof. Since f is continuous and X is compact, Theorem A.8.26 implies that $f(X)$ is a compact subset of $\mathbb{R}$. Hence $f(X)$ is closed and bounded by the Heine-Borel Theorem (Theorem A.8.25). Since $f(X)$ is non-empty and bounded, $\sup(f(X))$ and $\inf(f(X))$ are finite and we can construct sequences of elements of $f(X)$ converging to $\sup(f(X))$ and $\inf(f(X))$ respectively. Since $f(X)$ is also closed, this implies $\sup(f(X)), \inf(f(X)) \in f(X)$. Hence there exists $x_1, x_2 \in X$ such that $f(x_1) = \inf(f(X))$ and $f(x_2) = \sup(f(X))$ so $f(x_1) \leq f(x) \leq f(x_2)$ for all $x \in X$ as desired. ∎

Another application of Theorem A.8.26 is the following.

Corollary A.8.28. *Every topological space homeomorphic to a compact topological space is compact.*

Proof. The result easily follows from Theorem A.8.26. ∎

To finish off our preliminary study of compact topological spaces, we further note that the notions of compact and Hausdorff topological spaces intertwine nicely.

Theorem A.8.29. *Let $(X, \mathcal{T}_X)$ be a compact topological space and let $(Y, \mathcal{T}_Y)$ be a Hausdorff space. If $f : X \to Y$ is a continuous bijection, then f is a homeomorphism. Thus $(X, \mathcal{T}_X)$ is Hausdorff and $(Y, \mathcal{T}_Y)$ is compact.*

Proof. Suppose $f : X \to Y$ is a continuous bijection. To see that $f^{-1} : Y \to X$ is continuous, let $U \in \mathcal{T}_X$ be arbitrary. Then $K = X \setminus U$ is a closed subset of $(X, \mathcal{T}_X)$ and thus compact by Theorem A.8.14. Hence $(f^{-1})^{-1}(K) = f(K)$ is a compact subset of $(Y, \mathcal{T}_Y)$ by Theorem A.8.26. Therefore, since $(Y, \mathcal{T}_Y)$ is Hausdorff, $f(K)$ is closed by Theorem A.8.14. Thus

$$(f^{-1})^{-1}(U) = f(U) = f(X \setminus K) = f(X) \setminus f(K) = Y \setminus f(K)$$

is open in $(Y, \mathcal{T}_Y)$. Therefore, as $U \in \mathcal{T}_X$ was arbitrary, f^{-1} is continuous. Hence f is a homeomorphism.

The facts that $(X, \mathcal{T}_X)$ is Hausdorff and $(Y, \mathcal{T}_Y)$ is compact then follow as homeomorphisms clearly preserve these topological properties. ∎

A.9 Other Characterizations of Compactness

Now that we have some knowledge of the basics and some examples of compact topological spaces, we turn our attention to equivalent characterizations of compact topological spaces in the general topological setting. For example, in $\mathbb{R}$, the Bolzano-Weierstrass Theorem states (upto a reformulation) that a set $A \subseteq \mathbb{R}$ is closed and bounded if and only every sequence of elements of A has a convergent subsequence to an element of A. As the Heine-Borel Theorem (Theorem A.8.25) implies the closed, bounded subsets of $\mathbb{R}$ are exactly the compact subsets of $\mathbb{R}$, it is natural to ask whether 'every sequence having a convergent subsequence' is a characterization of compact topological spaces. Of course, as sequences are insufficient in general topological spaces, it is more useful to ask whether 'every net has a convergent subnet' is an equivalent characterization of compact topological spaces.

In this section, our main (and only) result will show this is indeed the case. In addition, there is another incredibly useful characterization of compact topological spaces using intersections of certain collections of subsets.

Definition A.9.1. Let $(X, \mathcal{T})$ be a topological space. A collection $\{F_\alpha\}_{\alpha \in I}$ is said to have the *finite intersection property* if $\bigcap_{\alpha \in J} F_\alpha \neq \emptyset$ for every finite subset $J \subseteq I$.

Theorem A.9.2. *Let $(X, \mathcal{T})$ be a topological. The following are equivalent:*

(i) $(X, \mathcal{T})$ *is compact.*

(ii) Whenever $\{F_\alpha\}_{\alpha \in I}$ *is a collection of closed subsets of* $(X, \mathcal{T})$ *with the finite intersection property,* $\bigcap_{\alpha \in I} F_\alpha \neq \emptyset$.

(iii) For every net $(x_\lambda)_{\lambda \in \Lambda}$*, there exists an* $x_0 \in X$ *such that*

$$x_0 \in \overline{\{x_\lambda \mid \lambda \in \Lambda \text{ such that } \lambda \geq \lambda_0\}}$$

for every $\lambda_0 \in \Lambda$ *(often* x_0 *is called a* cluster point of the net*).*

(iv) Every net in $(X, \mathcal{T})$ *has a convergent subnet.*

Proof. To see that (i) implies (ii), let $(X, \mathcal{T})$ be a compact topological space. Suppose $\{F_\alpha\}_{\alpha \in I}$ is a set of closed subsets of $(X, \mathcal{T})$ with the finite intersection property yet $\bigcap_{\alpha \in I} F_\alpha = \emptyset$. For each $\alpha \in I$, let $U_\alpha = X \setminus F_\alpha$. Hence $\{U_\alpha\}_{\alpha \in I}$ are open subsets of $(X, \mathcal{T})$ as $\{F_\alpha\}_{\alpha \in I}$ is a set of closed subsets of $(X, \mathcal{T})$. Since

$$\bigcup_{\alpha \in I} U_\alpha = \bigcup_{\alpha \in I} X \setminus F_\alpha = X \setminus \left(\bigcap_{\alpha \in I} F_\alpha \right) = X \setminus \emptyset = X,$$

we see that $\{U_\alpha\}_{\alpha \in I}$ is an open cover of X. Therefore, as $(X, \mathcal{T})$ is compact, there exists a finite subset $J \subseteq I$ such that

$$X = \bigcup_{\alpha \in J} U_\alpha.$$

Hence

$$\emptyset = X \setminus X = X \setminus \left(\bigcup_{\alpha \in J} U_\alpha \right) = \bigcap_{\alpha \in J} X \setminus U_\alpha = \bigcap_{\alpha \in J} F_\alpha$$

thereby contradicting the fact that $\{F_\alpha\}_{\alpha \in I}$ has the finite intersection property. Thus, as we have obtained our contradiction, (i) implies (ii).

To see that (ii) implies (iii), suppose (ii) holds. Let $(x_\lambda)_{\lambda \in \Lambda}$ be an arbitrary net in $(X, \mathcal{T})$. For each $\lambda \in \Lambda$, let

$$A_\lambda = \{x_{\lambda'} \mid \lambda' \geq \lambda\} \qquad \text{and} \qquad F_\lambda = \overline{A_\lambda} \subseteq X.$$

Clearly $\{F_\lambda\}_{\lambda \in \Lambda}$ are closed subsets of X. We claim that $\{F_\lambda\}_{\lambda \in \Lambda}$ has the finite intersection property. Indeed suppose $J \subseteq \Lambda$ is finite. Since J is finite, the existence of upper bounds in directed sets implies there exists a $\lambda_0 \in \Lambda$ such that $\lambda_0 \geq \lambda$ for all $\lambda \in \Lambda$. Hence

$$x_{\lambda_0} \in \bigcap_{\lambda \in J} F_\lambda \quad \text{so} \quad \bigcap_{\lambda \in J} F_\lambda \neq \emptyset.$$

Therefore, as $J \subseteq I$ was an arbitrary finite subset, $\{F_\lambda\}_{\lambda \in \Lambda}$ has the finite intersection property. By the assumption of (ii), we know that there exists an $x_0 \in \bigcap_{\lambda \in \Lambda} F_\lambda$ as desired. Hence, as $(x_\lambda)_{\lambda \in \Lambda}$ was arbitrary, (ii) implies (iii).

To see that (iii) implies (iv), let $(x_\lambda)_{\lambda \in \Lambda}$ be an arbitrary net in $(X, \mathcal{T})$. For each $\lambda \in \Lambda$, let

$$A_\lambda = \{x_{\lambda'} \mid \lambda' \geq \lambda\} \qquad \text{and} \qquad F_\lambda = \overline{A_\lambda} \subseteq X.$$

By the assumption of (iii), there exists an $x_0 \in \bigcap_{\lambda \in \Lambda} F_\lambda$.

We claim that there exists a subnet of $(x_\lambda)_{\lambda \in \Lambda}$ that converges to x_0. To see this, first notice for an arbitrary neighbourhood U of x_0 that $A_\lambda \cap U \neq \emptyset$

by Theorem A.5.21 as $x_0 \in F_\lambda$. Hence, for every neighbourhood U of x_0 and for every $\lambda \in \Lambda$ there exists a $\lambda' \in \Lambda$ such that $\lambda' \geq \lambda$ and $x_{\lambda'} \in U$.

Let

$$M = \{(U, \lambda) \mid U \text{ a neighbourhood of } x_0 \text{ and } \lambda \in \Lambda \text{ such that } x_\lambda \in U\},$$

which is non-empty by previous discussions. For two pairs $(U_1, \lambda_1), (U_2, \lambda_2) \in M$, define $(U_1, \lambda_1) \leq (U_2, \lambda_2)$ if and only if $U_1 \supseteq U_2$ and $\lambda_1 \leq \lambda_2$. We claim that $(M, \leq)$ is a directed set. Indeed it is clear $(M, \leq)$ is reflexive and transitive since reverse inclusion and the ordering on Λ are. Finally let $(U_1, \lambda_1), (U_2, \lambda_2) \in M$ be arbitrary. Let $U_3 = U_1 \cap U_2$. Clearly U_3 is a neighbourhood of x_0 as U_1 and U_2 are. As $(\Lambda, \leq)$ is a directed set, there exists a $\lambda' \in \Lambda$ such that $\lambda' \geq \lambda_1$ and $\lambda' \geq \lambda_2$. By the previous paragraph there exists a $\lambda_3 \in \Lambda$ such that $\lambda_3 \geq \lambda'$ (so $\lambda_3 \geq \lambda_1$ and $\lambda_3 \geq \lambda_2$) and $x_{\lambda_3} \in U_3$. Hence $(U_3, \lambda_3) \in M$, $(U_3, \lambda_3) \geq (U_1, \lambda_1)$, and $(U_3, \lambda_3) \geq (U_2, \lambda_2)$. Therefore, as $(U_1, \lambda_1), (U_2, \lambda_2) \in M$ were arbitrary, $(M, \leq)$ is a directed set.

We claim that $(x_\lambda)_{(U,\lambda)\in M}$ is a subnet of $(x_\lambda)_{\lambda\in\Lambda}$. To see this, define $\varphi : M \to \Lambda$ by $\varphi((U, \lambda)) = \lambda$. Clearly φ is increasing by the definition of the ordering on M. To see that φ is cofinal, let $\lambda \in \Lambda$ be arbitrary. Then clearly $(X, \lambda) \in M$ and $\varphi((X, \lambda)) = \lambda \geq \lambda$. Hence $(x_\lambda)_{(U,\lambda)\in M}$ is a subnet of $(x_\lambda)_{\lambda\in\Lambda}$ by Definition A.4.37

Finally, we claim that $(x_\lambda)_{(U,\lambda)\in M}$ converges to x_0. To see this, let U be an arbitrary neighbourhood of x_0. From previous discussions there exists a $\lambda \in \Lambda$ such that $(U, \lambda) \in M$. Thus for all $(U', \lambda') \geq (U, \lambda)$ we have that $x_{\lambda'} \in U' \subseteq U$. Therefore, as U was arbitrary, $(x_\lambda)_{(U,\lambda)\in M}$ is a subnet of $(x_\lambda)_{\lambda\in\Lambda}$ that converges to x_0. Therefore, as $(x_\lambda)_{\lambda\in\Lambda}$ was arbitrary, (iii) implies (iv).

To see that (iv) implies (i), suppose (iv) holds. To see that $(X, \mathcal{T})$ is compact, suppose to the contrary that there exists an open cover $\{U_\alpha\}_{\alpha\in I}$ of $(X, \mathcal{T})$ that has no finite subcover. We will use $\{U_\alpha\}_{\alpha\in I}$ to construct a net that has no convergent subnets.

Let

$$\Lambda = \left\{ U \subseteq X \;\middle|\; U = \bigcup_{\alpha\in J} U_\alpha \text{ for some finite subset } J \subseteq U \right\}.$$

For two sets $U_1, U_2 \in \Lambda$, define $U_1 \leq U_2$ if and only if $U_1 \subseteq U_2$. Since Λ is closed under finite unions, Example A.4.4 implies $(\Lambda, \leq)$ is a directed set.

Since $\{U_\alpha\}_{\alpha\in I}$ has no finite subcover of $(X, \mathcal{T})$, $X \setminus (\bigcup_{\alpha\in J} U_\alpha) \neq \emptyset$ for each finite set $J \subseteq I$. Hence, for each $U \in \Lambda$ we may choose a point $x_U \in X \setminus U$. Thus $(x_U)_{U\in\Lambda}$ is a net in $(X, \mathcal{T})$.

We claim that $(x_U)_{U\in\Lambda}$ has no convergent subnets thereby contradicting the assumption that (iv) holds and yielding (iv) implies (i). To see this, suppose to the contrary that $(x_U)_{U\in\Lambda}$ has a subnet $(x_{\lambda_\mu})_{\mu\in M}$ that converges to some point $x_0 \in X$. Since $\{U_\alpha\}_{\alpha\in I}$ is an open cover of $(X, \mathcal{T})$, there exists an $\alpha_0 \in I$ such that $x_0 \in U_{\alpha_0}$. Thus, as $(x_{\lambda_\mu})_{\mu\in M}$ is a subnet of $(x_U)_{U\in\Lambda}$, there exists a μ_0 such that $\lambda_{\mu_0} \geq U_{\alpha_0}$. Moreover, since $(x_{\lambda_\mu})_{\mu\in M}$ converges to x_0, there exists an $\mu_1 \in M$ such that $x_{\lambda_\mu} \in U_{\alpha_0}$ for all $\mu \geq \mu_1$. By the definition of a directed set, there exists a $\mu_2 \in M$ such that $\mu_2 \geq \mu_0$ and $\mu_2 \geq \mu_1$. Hence $\lambda_{\mu_2} \geq U_{\alpha_0}$ and $x_{\lambda_{\mu_2}} \in U_{\alpha_0}$. However, $\lambda_{\mu_2} \geq U_{\alpha_0}$ implies that $\lambda_{\mu_2} = U$ for some open set U in $(X, \mathcal{T})$ such that $U \supseteq U_{\alpha_0}$ so the definition of $x_{\lambda_{\mu_2}}$ implies that

$$x_{\lambda_{\mu_2}} \notin U \qquad \text{so} \qquad x_{\lambda_{\mu_2}} \notin U_{\alpha_0}.$$

As this contradicts the fact that $x_{\lambda_{\mu_2}} \in U_{\alpha_0}$, $(x_U)_{U\in\Lambda}$ has no convergent subnets. Hence (iv) implies (i). ∎

Appendix B

Inner Product Spaces

In this appendix chapter, we will quickly review the elementary properties of inner product spaces from linear algebra.

B.1 Inner Product Spaces

We begin with the definition of an inner product

Definition B.1.1. Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. An *inner product* on V is a map $\langle\,\cdot,\cdot\,\rangle : \mathcal{V}\times\mathcal{V}\to\mathbb{K}$ such that

1. $\langle\vec{v},\vec{v}\rangle \geq 0$ for all $\vec{v}\in\mathcal{V}$,
2. $\langle\vec{v},\vec{v}\rangle = 0$ if and only if $\vec{v}=\vec{0}$,
3. $\langle\vec{x}+\lambda\vec{y},\vec{v}\rangle = \langle\vec{x},\vec{v}\rangle + \lambda\langle\vec{y},\vec{v}\rangle$ for all $\vec{v},\vec{x},\vec{y}\in\mathcal{V}$ and $\lambda\in\mathbb{K}$ (i.e. $\langle\,\cdot,\cdot\,\rangle$) is linear in the first entry), and
4. $\overline{\langle\vec{x},\vec{y}\rangle} = \langle\vec{y},\vec{x}\rangle$ for all $\vec{x},\vec{y}\in\mathcal{V}$ (where $\overline{z}$ is the complex conjugate of z).

Remark B.1.2. Combining properties (3) and (4) in Definition B.1.1, we obtain for all $\vec{v},\vec{x},\vec{y}\in\mathcal{V}$ and $\lambda\in\mathbb{K}$ that

$$\begin{aligned}\langle\vec{v},\vec{x}+\lambda\vec{y}\rangle &= \overline{\langle\vec{x}+\lambda\vec{y},\vec{v}\rangle}\\ &= \overline{\langle\vec{x},\vec{v}\rangle + \lambda\langle\vec{y},\vec{v}\rangle}\\ &= \overline{\langle\vec{x},\vec{v}\rangle} + \overline{\lambda}\overline{\langle\vec{y},\vec{v}\rangle}\\ &= \langle\vec{v},\vec{x}\rangle + \overline{\lambda}\langle\vec{v},\vec{y}\rangle.\end{aligned}$$

That is, every inner product is *conjugate linear* in the second entry.

Remark B.1.3. Notice that if $\langle\,\cdot,\cdot\,\rangle$ is an inner product on a vector space V, the fact that $\langle\,\cdot,\cdot\,\rangle$ is linear in the first entry and conjugate linear in the second entries implies that $\langle\vec{0},\vec{v}\rangle = \langle\vec{v},\vec{0}\rangle = 0$ for all $\vec{v}\in\mathcal{V}$.

As we are interested in the vector space together with a fixed inner product, we make the following definition.

Definition B.1.4. An *inner product space* is a pair $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ where $\mathcal{V}$ is a vector space over $\mathbb{K}$ and $\langle \cdot, \cdot \rangle$ is an inner product on $\mathcal{V}$.

Remark B.1.5. Again, we will often abuse notation by said that $\mathcal{V}$ is an inner product space without specifying $\langle \cdot, \cdot \rangle$.

Example B.1.6. Let $n \in \mathbb{N}$. Define $\langle \cdot, \cdot \rangle_2 : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ by

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle_2 = \sum_{k=1}^{n} z_k \overline{w_k}$$

for all $(z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{K}^n$. It is elementary to verify that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{K}^n$. We call $\langle \cdot, \cdot \rangle$ the *standard inner product* on $\mathbb{K}^n$.

Example B.1.7. Let $n \in \mathbb{N}$ and let $\mathcal{M}_n(\mathbb{K})$ denote the set of $n \times n$ matrices with entries in $\mathbb{K}$. Define $\langle \cdot, \cdot \rangle : \mathcal{M}_n(\mathbb{K}) \times \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ by

$$\langle A, B \rangle = \mathrm{Tr}(AB^*)$$

for all $A, B \in \mathcal{M}_n(\mathbb{K})$ where B^* is the conjugate transpose of B and $\mathrm{Tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ is the trace. As the trace is linear, it is elementary to verify that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathcal{M}_n(\mathbb{K})$.

Notice if we write $A = [a_{i,j}]$ and $B = [b_{i,j}]$ then

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{i,j} \overline{b_{i,j}}.$$

Therefore, by comparing with Example B.1.6, it is elementary to see that there is an invertible linear map $\varphi : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}^{n^2}$ such that $\langle \varphi(A), \varphi(B) \rangle_{\mathbb{K}^{n^2}} = \mathrm{Tr}(AB^*)$ for all $A, B \in \mathcal{M}_n(\mathbb{K})$. In particular, $\mathcal{M}_n(\mathbb{K})$ with this inner product is really $\mathbb{K}^{n^2}$ with the standard inner product in disguise.

Example B.1.8. Let $n \in \mathbb{N}$ and let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathbb{K}^n$. It is then possible to show that there exists a matrix $A = [a_{i,j}] \in \mathcal{M}_n(\mathbb{K})$ such that A is invertible and positive definite, and

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = \sum_{i,j=1}^{n} a_{i,j} z_j \overline{w_i}$$

for all $(z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{K}^n$. We leave the proof as an exercise that will make use of the theory we will develop in this chapter and the fact that

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = \langle A(z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle_2$$

where $\langle \cdot, \cdot \rangle_2$ is the standard inner product from Example B.1.6 and where $A(z_1, \ldots, z_n)$ represents the vector obtained by matrix multiplication of A against the column vector with entries $(z_1, \ldots, z_n)$. Note if A is the identity matrix, then the standard inner product is recovered.

Example B.1.9. Define $\langle \cdot, \cdot \rangle : \mathcal{C}[0,1] \times \mathcal{C}[0,1] \to \mathbb{R}$ by

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$$

for all $f, g \in \mathcal{C}[0,1]$. It is elementary to verify that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathcal{C}[0,1]$.

Example B.1.10. Define $\langle \cdot, \cdot \rangle : \ell_2(\mathbb{N}) \times \ell_2(\mathbb{N}) \to \mathbb{K}$ by

$$\langle (z_n)_{n\geq 1}, (w_n)_{n\geq 1} \rangle = \sum_{n=1}^{\infty} z_n \overline{w_n}$$

for all $(z_n)_{n\geq 1}, (w_n)_{n\geq 1} \in \ell_2(\mathbb{N})$. It is not difficult to see that $\langle \cdot, \cdot \rangle$ will satisfy the conditions in Definition B.1.1 provided the sum under consideration actually converges in $\mathbb{K}$. Since

$$\sum_{n=1}^{\infty} |z_n \overline{w_n}| \leq \left\| (z_n)_{n\geq 1} \right\|_2 \left\| (w_n)_{n\geq 1} \right\|_2$$

by Hölder's Inequality (Theorem D.1.7), and since $\mathbb{K}$ is complete (so absolutely summable series converge by Theorem 2.2.2), the sum is finite.

We desire to show that each inner product space has a norm induced by the inner product, which happens to be the 2-norm in (almost) all of the above examples. To do this, we first prove the following very useful inequality.

Theorem B.1.11 (Cauchy-Schwarz Inequality). *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. For all $\vec{x}, \vec{y} \in \mathcal{V}$,*

$$|\langle \vec{x}, \vec{y} \rangle| \leq \langle \vec{x}, \vec{x} \rangle^{\frac{1}{2}} \langle \vec{y}, \vec{y} \rangle^{\frac{1}{2}}.$$

Furthermore, the above inequality is an equality if and only if $\{\vec{x}, \vec{y}\}$ is linearly dependent.

Proof. First notice if $\vec{x} = \vec{0}$ or $\vec{y} = \vec{0}$, then the proof is trivial by Remark B.1.3. Thus we may assume that $\vec{x}, \vec{y} \neq \vec{0}$.

Choose $\lambda \in \mathbb{K}$ with $|\lambda| = 1$ such that

$$\langle \lambda\vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle = |\langle \vec{x}, \vec{y} \rangle|,$$

and notice for all $t \in \mathbb{R}$ that

$$\begin{aligned} 0 &\leq \langle \lambda\vec{x} + t\vec{y}, \lambda\vec{x} + t\vec{y} \rangle \\ &= |\lambda|^2 \langle \vec{x}, \vec{x} \rangle + t\langle \vec{y}, \lambda\vec{x} \rangle + t\langle \lambda\vec{x}, \vec{y} \rangle + t^2 \langle \vec{y}, \vec{y} \rangle \\ &= \langle \vec{x}, \vec{x} \rangle + 2t|\langle \vec{x}, \vec{y} \rangle| + t^2 \langle \vec{y}, \vec{y} \rangle. \end{aligned}$$

By substituting

$$t_0 = -\frac{|\langle \vec{x}, \vec{y} \rangle|}{\langle \vec{y}, \vec{y} \rangle}$$

which is well-defined as $\vec{y} \neq 0$, we obtain that

$$0 \leq \langle \vec{x}, \vec{x} \rangle - 2\frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\langle \vec{y}, \vec{y} \rangle} + \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{\langle \vec{y}, \vec{y} \rangle}$$

which clearly implies the inequality.

For the additional claim, notice if $\vec{x} = \alpha\vec{y}$ for some $\alpha \in \mathbb{K}$, then

$$|\langle \vec{x}, \vec{y} \rangle| = |\alpha| \langle \vec{y}, \vec{y} \rangle = \alpha^{\frac{1}{2}} \overline{\alpha}^{\frac{1}{2}} \langle \vec{y}, \vec{y} \rangle^{\frac{1}{2}} \langle \vec{y}, \vec{y} \rangle^{\frac{1}{2}} = \langle \vec{x}, \vec{x} \rangle^{\frac{1}{2}} \langle \vec{y}, \vec{y} \rangle^{\frac{1}{2}}.$$

For the other direction, notice if the Cauchy-Schwarz inequality is an equality then the above proof shows

$$\langle \lambda\vec{x} + t_0\vec{y}, \lambda\vec{x} + t_0\vec{y} \rangle = 0.$$

Hence $\lambda\vec{x} + t_0\vec{y} = \vec{0}$ so $\{\vec{x}, \vec{y}\}$ is linearly dependent (as $\lambda \neq 0$). ■

Theorem B.1.12. *Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an inner product space. Then $\mathcal{V}$ is a normed linear space with a norm $\|\cdot\| : \mathcal{V} \to [0, \infty)$ defined by*

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

for all $\vec{v} \in V$.

Proof. It is elementary using Definition B.1.1 to see that $\|\cdot\|$ is well-defined, $\|\vec{v}\| \geq 0$ for all $\vec{v} \in \mathcal{V}$, $\|\vec{v}\| = 0$ if and only if $\vec{v} = \vec{0}$, and $\|\alpha\vec{v}\| = |\alpha| \, \|\vec{v}\|$ for all $\vec{v} \in \mathcal{V}$ and $\alpha \in \mathbb{K}$. To see that $\|\cdot\|$ satisfies the triangle inequality, notice for all $\vec{x}, \vec{y} \in \mathcal{V}$ that

$$\begin{aligned}
\|\vec{x} + \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle \\
&= \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle \\
&= \|\vec{x}\|^2 + 2\mathrm{Re}(\langle \vec{x}, \vec{y} \rangle) + \|\vec{y}\|^2 \\
&\leq \|\vec{x}\|^2 + 2|\langle \vec{x}, \vec{y} \rangle| + \|\vec{y}\|^2 \\
&\leq \|\vec{x}\|^2 + 2\,\|\vec{x}\|\,\|\vec{y}\| + \|\vec{y}\|^2 \quad \text{by Cauchy-Schwarz (Theorem B.1.11)} \\
&= (\|\vec{x}\| + \|\vec{y}\|)^2 .
\end{aligned}$$

Hence the triangle inequality follows. ■

Remark B.1.13. For the triangle inequality to be an equality, notice we require equality in the Cauchy-Schwarz inequality which implies $\vec{x}$ and $\vec{y}$ are linearly dependent. Furthermore, we notice we require $\mathrm{Re}(\langle \vec{x}, \vec{y} \rangle) = |\langle \vec{x}, \vec{y} \rangle|$ will then occur only if $\vec{x} = \alpha\vec{y}$ or $\vec{y} = \alpha\vec{x}$ for some $\alpha \in [0, \infty)$. Clearly this later condition implies equality in the triangle inequality.

Remark B.1.14. By the Cauchy-Schwarz inequality $|\langle \vec{x}, \vec{y}\rangle| \leq \|\vec{x}\| \, \|\vec{y}\|$, we see that the inner product is simultaneously continuous in its entry. Indeed if $\vec{x} = \lim_{n\to\infty} \vec{x}_n$ and $\vec{y} = \lim_{n\to\infty} \vec{y}_n$, then

$$\begin{aligned}\limsup_{n\to\infty} |\langle \vec{x}, \vec{y}\rangle - \langle \vec{x}_n, \vec{y}_n\rangle| &\leq \limsup_{n\to\infty} |\langle \vec{x}, \vec{y}\rangle - \langle \vec{x}, \vec{y}_n\rangle| + |\langle \vec{x}, \vec{y}_n\rangle - \langle \vec{x}_n, \vec{y}_n\rangle| \\ &\leq \limsup_{n\to\infty} \|\vec{x}\| \, \|\vec{y} - \vec{y}_n\| + \|\vec{x} - \vec{x}_n\| \, \|\vec{y}_n\| = 0\end{aligned}$$

as $\vec{x} = \lim_{n\to\infty} \vec{x}_n$ and $\vec{y} = \lim_{n\to\infty} \vec{y}_n$, with the later implying that $(\vec{y}_n)_{n\geq 1}$ is bounded.

Remark B.1.15. The proof of Theorem B.1.12 also enables us to develop a notion of an angle. To motivate this, recall the cosine law for a triangle which states

$$c^2 = a^2 + b^2 - 2ab\cos(\theta)$$

for a triangle with sides a, b, c and angle θ opposite to c. Thinking of a 'triangle' formed by two vectors $\vec{x}, \vec{y}$ and their difference in a real inner product space, the proof of Theorem B.1.12 demonstrates

$$\|\vec{x} - \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\langle \vec{x}, \vec{y}\rangle.$$

Thus, for a real inner product space, we would like to define the angle θ between $\vec{x}$ and $\vec{y}$ to be such that

$$\cos(\theta) = \frac{\langle \vec{x}, \vec{y}\rangle}{\|\vec{x}\| \, \|\vec{y}\|}$$

which exists by the Cauchy-Schwarz Inequality.

Using the above notion of an angle, we obtain the definition of what it means for two vectors to be perpendicular.

Definition B.1.16. Let $(\mathcal{V}, \langle\,\cdot, \cdot\,\rangle)$ be an inner product space. Two vectors $\vec{v}, \vec{w} \in \mathcal{V}$ are said to be *orthogonal* if $\langle \vec{v}, \vec{w}\rangle = 0$.

Using the properties of the inner product, it is nearly trivial to obtain the following theorems. Thus we omit the proofs.

Theorem B.1.17 (Pythagorean Theorem). *Let $(\mathcal{V}, \langle\,\cdot, \cdot\,\rangle)$ be an inner product space. If $\{\vec{v}_k\}_{k=1}^n$ is a set of orthogonal vectors, then*

$$\left\|\sum_{k=1}^n \vec{v}_k\right\|^2 = \sum_{k=1}^n \|\vec{v}_k\|^2 .$$

Theorem B.1.18 (Parallelogram Law). *Let $(\mathcal{V}, \langle\,\cdot, \cdot\,\rangle)$ be an inner product space. If $\vec{x}, \vec{y} \in \mathcal{V}$, then*

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2\|\vec{x}\|^2 + 2\|\vec{y}\|^2 .$$

Remark B.1.19. It is difficult but possible to show that any norm on any vector space over $\mathbb{K}$ that satisfies the Parallelogram Law actually comes from an inner product.

Theorem B.1.20 (Polarization Identity). *Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be an inner product space. If $\vec{x}, \vec{y} \in \mathcal{V}$, then*

- $\langle \vec{x}, \vec{y} \rangle = \frac{1}{4} \|\vec{x} + \vec{y}\|^2 - \frac{1}{4} \|\vec{x} - \vec{y}\|^2$ *if* $\mathbb{K} = \mathbb{R}$, *and*
- $\langle \vec{x}, \vec{y} \rangle = \frac{1}{4} \sum_{k=1}^{4} i^k \left\|\vec{x} + i^k \vec{y}\right\|^2$ *if* $\mathbb{K} = \mathbb{C}$.

Appendix C

Completions of Normed Linear Spaces

In this appendix chapter, we will demonstrate that given any normed linear space $(\mathcal{X}, \|\cdot\|)$ there exists a unique way to *complete* $\mathcal{X}$ to obtain a Banach space.

C.1 Completion of a Normed Linear Space

To be formal with our notion of how we are going to complete a normed linear space, we define the following.

Definition C.1.1. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a normed linear space. A *completion* of $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is a Banach space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ such that there exists an isometry $\varphi : \mathcal{X} \to \mathcal{Y}$ such that $\overline{\varphi(\mathcal{X})} = \mathcal{Y}$.

We will first prove the existence of a completion of any normed linear space followed by a proof of its uniqueness.

Theorem C.1.2. *Every normed linear space has a completion.*

Proof. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a normed linear space. Let $\mathcal{V}$ denote the set of all Cauchy sequences in $\mathcal{X}$. Note that $\mathcal{V}$ is non-empty as every constant sequence is Cauchy. Since given Cauchy sequences $(\vec{x}_n)_{n\geq 1}$ and $(\vec{y}_n)_{n\geq 1}$ and $\alpha \in \mathbb{K}$, the sequences

$$(\vec{x}_n + \vec{y}_n)_{n\geq 1} \qquad \text{and} \qquad (\alpha\vec{x}_n)_{n\geq 1}$$

are Cauchy by the properties of the norm, $\mathcal{V}$ is a vector space over $\mathbb{K}$. However, $\mathcal{V}$ is not the normed linear space we want. To construct a normed linear space, we require a quotient.

Let

$$\mathcal{W} = \left\{(\vec{x}_n)_{n\geq 1} \in \mathcal{V} \mid \lim_{n\to\infty} \vec{x}_n = \vec{0}\right\}.$$

Clearly $\mathcal{W}$ is a subspace of $\mathcal{V}$ and thus $\mathcal{V}/\mathcal{W}$ is a vector space with operations $[\vec{v}_1]+[\vec{v}_2] = [\vec{v}_1+\vec{v}_2]$ and $\alpha[\vec{v}] = [\alpha\vec{v}]$. Recall two elements $(\vec{x}_n)_{n\geq 1}, (\vec{y}_n)_{n\geq 1} \in \mathcal{V}$ produce the same element in $\mathcal{V}/\mathcal{W}$ if and only if

$$\lim_{n\to\infty} \|\vec{x}_n - \vec{y}_n\|_{\mathcal{X}} = 0$$

Define $\|\cdot\| : \mathcal{V}/\mathcal{W} \to [0,\infty)$ by

$$\|[(\vec{x}_n)_{n\geq 1}]\| = \limsup_{n\to\infty} \|\vec{x}_n\|_{\mathcal{X}}$$

and note that since $(\vec{x}_n)_{n\geq 1}$ is Cauchy and thus bounded, $\|\cdot\|$ does indeed map into $[0,\infty)$. Since we are dealing with equivalence classes, we must check that $\|\cdot\|$ is well-defined. To see this, notice if $[(\vec{x}_n)_{n\geq 1}] = [(\vec{y}_n)_{n\geq 1}]$ then

$$\lim_{n\to\infty} \|\vec{x}_n - \vec{y}_n\|_{\mathcal{X}} = 0.$$

so

$$\limsup_{n\to\infty} \|\vec{x}_n\|_{\mathcal{X}} = \limsup_{n\to\infty} \|\vec{y}_n\|_{\mathcal{X}}$$

by the reverse triangle inequality. Hence $\|\cdot\|$ is well-defined. To see that $\|\cdot\|$ is indeed a norm, note that $\|[(\vec{x}_n)_{n\geq 1}]\| = 0$ if and only if $\limsup_{n\to\infty} \|\vec{x}_n\|_{\mathcal{X}} = 0$ if and only if $(\vec{x}_n)_{n\geq 1} \in \mathcal{W}$ if and only if $[(\vec{x}_n)_{n\geq 1}] = \vec{0}_{\mathcal{V}/\mathcal{W}}$. As the other properties of a norm are trivial to verify, $(\mathcal{V}/\mathcal{W}, \|\cdot\|)$ is a normed linear space.

We will postpone the proof that $(\mathcal{V}/\mathcal{W}, \|\cdot\|)$ is complete momentarily in order to demonstrate some facts in relation to $\mathcal{X}$. Define $\varphi : \mathcal{X} \to \mathcal{V}/\mathcal{W}$ by $\varphi(\vec{x}) = [(\vec{x})_{n\geq 1}]$; that is, map each element of $\mathcal{X}$ to a constant sequence. Clearly φ is well-defined, linear, and an isometry. We claim that $\varphi(\mathcal{X})$ is dense in $\mathcal{V}/\mathcal{W}$.

To see that $\varphi(\mathcal{X})$ is dense in $\mathcal{V}/\mathcal{W}$, let $[(\vec{x}_n)_{n\geq 1}] \in \mathcal{V}/\mathcal{W}$ be arbitrary and let $\epsilon > 0$ be arbitrary. Since $(\vec{x}_n)_{n\geq 1}$ is Cauchy in $\mathcal{X}$, there exists an $N \in \mathbb{N}$ such that $\|\vec{x}_n - \vec{x}_m\|_{\mathcal{X}} < \epsilon$ for all $n, m \geq N$. Hence

$$\|\varphi(\vec{x}_N) - [(\vec{x}_n)_{n\geq 1}]\| \leq \epsilon.$$

Therefore, as $\epsilon > 0$ was arbitrary, $[(\vec{x}_n)_{n\geq 1}]$ is in the closure of $\varphi(\mathcal{X})$. Hence, as $[(\vec{x}_n)_{n\geq 1}] \in \mathcal{V}/\mathcal{W}$ was arbitrary, $\varphi(\mathcal{X})$ is dense in $\mathcal{V}/\mathcal{W}$.

To see that $(\mathcal{V}/\mathcal{W}, \|\cdot\|)$ is complete, let $(\vec{z}_n)_{n\geq 1}$ be an arbitrary Cauchy sequence in $(\mathcal{V}/\mathcal{W}, \|\cdot\|)$. Since $\varphi(\mathcal{X})$ is dense in $\mathcal{V}/\mathcal{W}$, for each $n \in \mathbb{N}$ there exists an $\vec{x}_n \in \mathcal{X}$ such that

$$\|\varphi(\vec{x}_n) - \vec{z}_n\| < \frac{1}{n}.$$

We claim that $(\vec{x}_n)_{n\geq 1}$ is a Cauchy sequence of elements of $\mathcal{X}$ and thus is an element of $\mathcal{V}$. To see this, notice for all $n, m \in \mathbb{N}$ that

$$\begin{aligned}\|\vec{x}_n - \vec{x}_m\|_{\mathcal{X}} &= \|\varphi(\vec{x}_n) - \varphi(\vec{x}_m)\| \\ &\leq \|\varphi(\vec{x}_n) - \vec{z}_n\| + \|\vec{z}_n - \vec{z}_m\| + \|\vec{z}_m - \varphi(\vec{x}_m)\| \\ &\leq \frac{1}{n} + \frac{1}{m} + \|\vec{z}_n - \vec{z}_m\|.\end{aligned}$$

Therefore, as $(\vec{z}_n)_{n\geq 1}$ is Cauchy, it is elementary to verify the above inequality implies $(\vec{x}_n)_{n\geq 1}$ is Cauchy. Finally, to see that $(\vec{z}_n)_{n\geq 1}$ converges to $\vec{z} = [(\vec{x}_n)_{n\geq 1}]$, we notice that

$$\lim_{n\to\infty} \|\varphi(\vec{x}_n) - \vec{z}\| = 0$$

as $(\vec{x}_n)_{n\geq 1}$ is Cauchy. Hence as

$$\|\vec{z}_n - \vec{z}\| \leq \|\vec{z}_n - \varphi(\vec{x}_n)\| + \|\varphi(\vec{x}_n) - \vec{z}\| \leq \frac{1}{n} + \|\varphi(\vec{x}_n) - \vec{z}\|,$$

we obtain that $(\vec{z}_n)_{n\geq 1}$ converges to $\vec{z} = [(\vec{x}_n)_{n\geq 1}]$. Therefore, as $(\vec{z}_n)_{n\geq 1}$ was an arbitrary Cauchy sequence, $\mathcal{V}/\mathcal{W}$ is complete thereby completing the proof. ■

Proposition C.1.3. *Any two completions of a metric space are isomorphic.*

Proof. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a metric space. Suppose that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ and $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ are completions of $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$. Therefore there exists isometries $\varphi_{\mathcal{Y}} : \mathcal{X} \to \mathcal{Y}$ and $\varphi_{\mathcal{Z}} : \mathcal{X} \to \mathcal{Z}$ such that $\overline{\varphi_{\mathcal{Y}}(\mathcal{X})} = \mathcal{Y}$ and $\overline{\varphi_{\mathcal{Z}}(\mathcal{X})} = \mathcal{Z}$. Our goal is to extend the identity map from $\mathcal{X} \subseteq \mathcal{Y}$ to $\mathcal{X} \subseteq \mathcal{Z}$ to obtain an isometry from $\mathcal{Y}$ to $\mathcal{Z}$.

To define an isometry $\varphi : \mathcal{Y} \to \mathcal{Z}$, let $y \in \mathcal{Y}$ be arbitrary. Hence, as $\mathcal{Y}$ is the closure of $\mathcal{X}$ there exists a sequence $(x_n)_{n\geq 1}$ of elements of $\mathcal{X}$ such that $y = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x_n)$. However, as $(\varphi_{\mathcal{Y}}(x_n))_{n\geq 1}$ converges in $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$, $(\varphi_{\mathcal{Y}}(x_n))_{n\geq 1}$ is Cauchy in $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$. Therefore, $(x_n)_{n\geq 1}$ is Cauchy in $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ as $\varphi_{\mathcal{Y}}$ is an isometry. Hence $(\varphi_{\mathcal{Z}}(x_n))_{n\geq 1}$ also must be Cauchy as $\varphi_{\mathcal{Z}}$ is an isometry. Since $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$ is complete, $(\varphi_{\mathcal{Z}}(x_n))_{n\geq 1}$ converges in $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$. Let $z_y = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x_n)$. We would like to define $\varphi : \mathcal{Y} \to \mathcal{Z}$ such that $f(y) = z_y$. There is one technical issue with this definition that we should get out of the way; that is, we desire to show that if $(x'_n)_{n\geq 1}$ is another sequence of elements of $\mathcal{X}$ such that $y = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x'_n)$, then $z_y = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x'_n)$. This will demonstrate that the sequence of elements of $\mathcal{X}$ we choose converging to $y \in \mathcal{Y}$ does not affect the limit in $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$. To see this, notice by the triangle inequality and properties of limits that

$$\begin{aligned}\lim_{n\to\infty} \|\varphi_{\mathcal{Z}}(x'_n) - \varphi_{\mathcal{Z}}(x_n)\|_{\mathcal{Z}} &= \lim_{n\to\infty} \|x'_n - x_n\|_{\mathcal{X}} \\ &= \lim_{n\to\infty} \|\varphi_{\mathcal{Y}}(x'_n) - \varphi_{\mathcal{Y}}(x_n)\|_{\mathcal{Y}} \\ &= \|y - y\|_{\mathcal{Y}} = 0.\end{aligned}$$

Hence as $z_y = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x_n)$, the above easily implies $z_y = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x_n')$. Hence the claim is complete.

Hence we may define $\varphi : \mathcal{Y} \to \mathcal{Z}$ as follows: for each $y \in \mathcal{Y}$ choose a sequence $(x_n)_{n\geq 1}$ of elements of $\mathcal{X}$ such that $y = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x_n)$ and define $\varphi(y) = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x_n)$. We claim that φ is an isometry. To see this, let $y, y' \in \mathcal{Y}$ be arbitrary. Choose sequence $(x_n)_{n\geq 1}$ and $(x_n')_{n\geq 1}$ of elements of $\mathcal{X}$ such that $y = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x_n)$ and $y' = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x_n')$. Then, by the triangle inequality and properties of limits,

$$\begin{aligned} \left\|\varphi(y) - \varphi(y')\right\|_{\mathcal{Z}} &= \lim_{n\to\infty} \left\|\varphi_{\mathcal{Z}}(x_n) - \varphi_{\mathcal{Z}}(x_n')\right\|_{\mathcal{Z}} \\ &= \lim_{n\to\infty} \left\|x_n - x_n'\right\|_{\mathcal{X}} \\ &= \lim_{n\to\infty} \left\|\varphi_{\mathcal{Y}}(x_n) - \varphi_{\mathcal{Y}}(x_n')\right\|_{\mathcal{Y}} \\ &= \left\|y - y'\right\|_{\mathcal{Y}}. \end{aligned}$$

Hence φ is an isometry (and therefore injective).

To see that φ is surjective (and thus a bijection) let $z \in \mathcal{Z}$ be arbitrary. Note as $\mathcal{Z}$ is the completion of $\varphi_{\mathcal{Z}}(\mathcal{X})$, there exists a sequence $(x_n)_{n\geq 1}$ of elements of $\mathcal{X}$ such that $z = \lim_{n\to\infty} \varphi_{\mathcal{Z}}(x_n)$. By similar arguments to those above, $y = \lim_{n\to\infty} \varphi_{\mathcal{Y}}(x_n)$ exists and thus $\varphi(y) = z$. Hence, as $z \in \mathcal{Z}$ was arbitrary, φ is surjective. Hence $\mathcal{Y}$ and $\mathcal{Z}$ are isomorphic. ∎

C.2 Completion of Inner Product Spaces

In this section, we demonstrate the following theorem that the completion of a inner product space is a Hilbert space.

Theorem C.2.1. *Let $(V, \langle\,\cdot,\cdot\,\rangle)$ be an inner product space and let $\mathcal{H}$ be the normed linear space completion of V from Theorem C.1.2. There exists an inner product $\langle\,\cdot,\cdot\,\rangle_{\mathcal{H}} : \mathcal{H}\times\mathcal{H} \to \mathbb{K}$ such that $\langle\vec{x},\vec{y}\rangle_{\mathcal{H}} = \langle\vec{x},\vec{y}\rangle$ for all $\vec{x},\vec{y} \in V$.*

Proof. To prove this result, we could proceed in one of two ways. The first way would be to complete Remark B.1.19 and show that the norm on the completion of an inner product space then satisfies the the Parallelogram Law. Instead we will use an argument similar to Proposition C.1.3 to define an inner product on the completion.

Let $(V, \langle\,\cdot,\cdot\,\rangle)$ be an inner product space and let $\mathcal{H}$ be the normed linear space completion of V from Theorem C.1.2. Define $\langle\,\cdot,\cdot\,\rangle_{\mathcal{H}} : \mathcal{H}\times\mathcal{H} \to \mathbb{K}$ as follows: Given $\vec{x},\vec{y} \in \mathcal{H}$, choose sequences $(\vec{x}_n)_{n\geq 1}$ and $(\vec{y}_n)_{n\geq 1}$ such that $\vec{x} = \lim_{n\to\infty} \vec{x}_n$ and $\vec{y} = \lim_{n\to\infty} \vec{y}_n$. We then define

$$\langle\vec{x},\vec{y}\rangle_{\mathcal{H}} = \lim_{n\to\infty} \langle\vec{x}_n,\vec{y}_n\rangle.$$

To complete the proof, we will first need to demonstrate three things: that the above limit exists, that the definition did not depend on the sequences selected, and that the resulting definition does indeed yield an inner product.

To see that the limit exists, notice for all $n, m \in \mathbb{N}$ that

$$\begin{aligned}|\langle \vec{x}_n, \vec{y}_n\rangle - \langle \vec{x}_m, \vec{y}_m\rangle| &\leq |\langle \vec{x}_n, \vec{y}_n\rangle - \langle \vec{x}_n, \vec{y}_m\rangle| + |\langle \vec{x}_n, \vec{y}_m\rangle - \langle \vec{x}_m, \vec{y}_m\rangle| \\ &= |\langle \vec{x}_n, \vec{y}_n - \vec{y}_m\rangle| + |\langle \vec{x}_n - \vec{x}_m, \vec{y}_m\rangle| \\ &\leq \|\vec{x}_n\| \|\vec{y}_n - \vec{y}_m\| + \|\vec{x}_n - \vec{x}_m\| \|\vec{y}_m\|\end{aligned}$$

with the last inequality coming from the Cauchy-Schwarz inequality. Since $(\vec{x}_n)_{n\geq 1}$ and $(\vec{y}_n)_{n\geq 1}$ converge in $\mathcal{H}$, $(\vec{x}_n)_{n\geq 1}$ and $(\vec{y}_n)_{n\geq 1}$ are bounded and Cauchy. Hence the above inequality demonstrates that $(\langle \vec{x}_n, \vec{y}_n\rangle)_{n\geq 1}$ is Cauchy in $\mathbb{K}$ and thus converges. Hence the limit exists.

Similarly, if $(\vec{x}'_n)_{n\geq 1}$ and $(\vec{y}'_n)_{n\geq 1}$ are such that $\vec{x} = \lim_{n\to\infty} \vec{x}'_n$ and $\vec{y} = \lim_{n\to\infty} \vec{y}'_n$, the above computation shows that

$$\begin{aligned}\left|\langle \vec{x}_n, \vec{y}_n\rangle - \langle \vec{x}'_n, \vec{y}'_n\rangle\right| &\leq \left|\langle \vec{x}_n, \vec{y}_n\rangle - \langle \vec{x}_n, \vec{y}'_n\rangle\right| + \left|\langle \vec{x}_n, \vec{y}'_n\rangle - \langle \vec{x}'_n, \vec{y}'_n\rangle\right| \\ &= \left|\langle \vec{x}_n, \vec{y}_n - \vec{y}'_n\rangle\right| + \left|\langle \vec{x}_n - \vec{x}'_n, \vec{y}'_n\rangle\right| \\ &\leq \|\vec{x}_n\| \left\|\vec{y}_n - \vec{y}'_n\right\| + \left\|\vec{x}_n - \vec{x}'_n\right\| \left\|\vec{y}'_n\right\|.\end{aligned}$$

Hence we see that

$$\lim_{n\to\infty} \langle \vec{x}_n, \vec{y}_n\rangle = \lim_{n\to\infty} \langle \vec{x}'_n, \vec{y}'_n\rangle.$$

Thus the definition of $\langle \vec{x}, \vec{y}\rangle_{\mathcal{H}}$ does not depend on the sequences representing $\vec{x}$ and $\vec{y}$.

To see that $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is an inner product on $\mathcal{H}$, first we notice that $\langle \vec{x}, \vec{x}\rangle_{\mathcal{H}} \geq 0$ as the limit of positive real numbers is positive. Furthermore, notice that $\langle \vec{x}, \vec{x}\rangle_{\mathcal{H}} = 0$ if and only if there exists a sequence $(\vec{x}_n)_{n\geq 1}$ such that $\vec{x} = \lim_{n\to\infty} \vec{x}_n$ and $\lim_{n\to\infty}\langle \vec{x}_n, \vec{x}_n\rangle = 0$. As the later is equivalent to $\lim_{n\to\infty} \|\vec{x}_n\| = 0$, we see that $\langle \vec{x}, \vec{x}\rangle_{\mathcal{H}} = 0$ if and only if $\vec{x} = \vec{0}$. Moreover

$$\overline{\langle \vec{x}, \vec{y}\rangle_{\mathcal{H}}} = \lim_{n\to\infty} \overline{\langle \vec{x}_n, \vec{y}_n\rangle} = \lim_{n\to\infty} \langle \vec{y}_n, \vec{x}_n\rangle = \langle \vec{y}, \vec{x}\rangle_{\mathcal{H}}.$$

Finally, we see for all $\alpha \in \mathbb{K}$ and $\vec{x}, \vec{y}, \vec{v} \in \mathcal{H}$ that $(\vec{x})_{n\geq 1}$, $(\vec{y}_n)_{n\geq 1}$, and $(\vec{v}_n)_{n\geq 1}$ are sequences in V that converge to $\vec{x}$, $\vec{y}$, and $\vec{v}$ respectively, then

$$\begin{aligned}\langle \vec{x} + \alpha\vec{y}, \vec{v}\rangle_{\mathcal{H}} &= \lim_{n\to\infty} \langle \vec{x}_n + \alpha\vec{y}_n, \vec{v}_n\rangle \\ &= \lim_{n\to\infty} \langle \vec{x}_n, \vec{v}_n\rangle + \alpha\langle \vec{y}_n, \vec{v}_n\rangle \\ &= \langle \vec{x}, \vec{v}\rangle_{\mathcal{H}} + \alpha\langle \vec{y}, \vec{v}\rangle_{\mathcal{H}}.\end{aligned}$$

Hence $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is an inner product. ∎

Appendix D

L_p Spaces

In this appendix chapter, we will look at some of the fundamental properties of L_p spaces.

D.1 Constructing the $\mathbf{L}_p$-Spaces

To begin our discuss of L_p spaces, we must begin with their definition.

Definition D.1.1. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. A measurable function $f : X \to \mathbb{C}$ is said to be *p-integrable* if

$$\int_X |f|^p \, d\mu < \infty.$$

The set of p-integrable functions on $(X, \mathcal{A}, \mu)$ is denoted $\mathcal{L}_p(X, \mu)$.

Unsurprisingly, $\mathcal{L}_p(X, \mu)$ is a vector space.

Lemma D.1.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Then $\mathcal{L}_p(X, \mu)$ is a vector space is a vector space over $\mathbb{C}$ (and thus, restricting to real-valued functions produces a vector space over $\mathbb{R}$).*

Proof. Let $f, g \in \mathcal{L}_p(X, \mu)$ and let $\alpha \in \mathbb{C}$. Then, as

$$\int_X |\alpha f|^p \, d\mu = |\alpha|^p \int_X |f|^p \, d\mu < \infty,$$

we see that $\alpha f \in \mathcal{L}_p(X, \mu)$. Moreover, since

$$\begin{aligned} |f + g|^p \leq (|f| + |g|)^p &\leq (2 \max\{|f|, |g|\})^p \\ &= 2^p \max\{|f|^p, |g|^p\} \leq 2^p (|f|^p + |g|^p), \end{aligned}$$

we see that

$$\int_X |f + g|^p \, d\mu \leq 2^p \int_X |f|^p \, d\mu + 2^p \int_X |g|^p \, d\mu < \infty$$

so $f + g \in \mathcal{L}_p(X, \mu)$. Hence $\mathcal{L}_p(X, \mu)$ is a vector space. ∎

Remark D.1.3. Of course, given $p \in [1, \infty)$, we would like to define a norm on $\mathcal{L}_p(X, \mu)$ so that we can perform analysis. In particular, given $f \in \mathcal{L}_p(X, \mu)$ we would like to define

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}$$

to be the p-norm of f. It is elementary to see that if $f \in \mathcal{L}_p(X, \mu)$, then $\|f\|_p \in [0, \infty)$. Moreover, for all $\alpha \in \mathbb{C}$ we see that

$$\begin{aligned} \|\alpha f\|_p = \left(\int_X |\alpha f|^p \, d\mu \right)^{\frac{1}{p}} &= \left(|\alpha|^p \int_X |f|^p \, d\mu \right)^{\frac{1}{p}} \\ &= |\alpha| \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}} = |\alpha| \, \|f\|_p \, . \end{aligned}$$

Furthermore, we will be able to verify the triangle inequality below. However, one problem remains. In the definition of a norm, the only vector that can have zero norm is the zero vector. However $\|f\|_p = 0$ if and only if f is zero almost everywhere. Thus it is possible there is a function f that is not the zero function (but zero almost everywhere) such that $\|f\|_p = 0$. How can we rectify this situation?

Well, as the problem is that functions that are equal almost everywhere are not equal, let's define a new notion of equality to make then equal. To begin, recall that $\mathcal{M}(X, \mathbb{C})$, the set of measurable functions from X to $\mathbb{C}$, is a vector space. Since it is elementary to verify that

$$W = \{ f \in \mathcal{M}(X, \mathbb{C}) \mid f = 0 \ \mu\text{-almost everywhere} \}$$

is a subspace of $\mathcal{M}(X, \mathbb{C})$, we can form the quotient space $\mathcal{M}(X, \mathbb{C})/W$. Given a function $f \in \mathcal{M}(X, \mathbb{C})$, we will use $[f]$ to denote the equivalence class $f + W$ in $\mathcal{M}(X, \mathbb{C})/W$. Clearly if $f, g \in \mathcal{M}(X, \mathbb{C})$, then $[f] = [g]$ if and only if $f = g$ almost everywhere. In particular, if $[f] = [g]$ then

$$\int_X |f|^p \, d\mu = \int_X |g|^p \, d\mu$$

as $|f|^p = |g|^p$ almost everywhere so $f \in \mathcal{L}_p(X, \mu)$ if and only if $g \in \mathcal{L}_p(X, \mu)$. Furthermore, since W is clearly a subspace of $\mathcal{L}_p(X, \mu)$, we can consider $\mathcal{L}_p(X, \mu)/W$

Definition D.1.4. Given a measure space $(X, \mathcal{A}, \mu)$ and a $p \in [1, \infty)$, the *L_p-space of* $(X, \mathcal{A}, \mu)$, denote $L_p(X, \mu)$, is the vector space over $\mathbb{C}$ defined by

$$L_p(X, \mu) = \{ [f] \mid f \in \mathcal{L}_p(X, \mu) \} \, .$$

Furthermore, the *p-norm* is the function $\| \cdot \|_p : L_p(X, \mu) \to [0, \infty)$ defined by

$$\|[f]\|_p = \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}$$

for all $[f] \in L_p(X, \mu)$.

Remark D.1.5. First, note that the p-norm is well-defined on $L_p(X, \mu)$. Indeed if $[f] = [g]$ then

$$\int_X |f|^p \, d\mu = \int_X |g|^p \, d\mu$$

so the value of $\|[f]\|_p$ does not depend on the representative of the equivalence class.

Due to the definition of $L_p(X, \mu)$ and Remark D.1.3, we will often not distinguish elements of $L_p(X, \mu)$ and $\mathcal{L}_p(X, \mu)$. In actuality, elements of $\mathcal{L}_p(X, \mu)$ are functions whereas elements of $L_p(X, \mu)$ are equivalence classes of functions in $\mathcal{L}_p(X, \mu)$. However, each element $\vec{v}$ of $L_p(X, \mu)$ can be represented by a function $f \in \mathcal{L}_p(X, \mu)$ and if $g \in \mathcal{L}_p(X, \mu)$ is such that $g = f$ a.e., then $\vec{v}$ can also be represented by g. Consequently, we will treat elements of $L_p(X, \mu)$ as functions that are p-integrable where we are allowed to modify the functions on a set of μ-measure zero. Thus we will often omit the notation of an equivalence class. One thing to keep in mind is that we must verify that any function defined on $L_p(X, \mu)$ respects almost everywhere equivalence.

To prove that the p-norm is a norm on $L_p(X, \mu)$, we require some inequalities.

Lemma D.1.6 (Young's Inequality). *Let $a, b \geq 0$ and let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then $ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$.*

Proof. Notice $1 = \frac{1}{p} + \frac{1}{q} = \frac{p+q}{pq}$ implies $p + q - pq = 0$. Hence $q = \frac{p}{p-1}$.

Fix $b \geq 0$. Notice if $b = 0$, the inequality easily holds. Thus we will assume $b > 0$.

Define $f : [0, \infty) \to \mathbb{R}$ by $f(x) = \frac{1}{p}x^p + \frac{1}{q}b^q - bx$. Clearly $f(0) > 0$ and $\lim_{x\to\infty} f(x) = \infty$ as $p > 1$ so x^p grows faster than x. We claim that $f(x) \geq 0$ for all $x \in [0, \infty)$ thereby proving the inequality. Notice f is differentiable on $[0, \infty)$ with

$$f'(x) = x^{p-1} - b.$$

Therefore $f'(x) = 0$ if and only if $x = b^{\frac{1}{p-1}}$. Moreover, it is elementary to see from the derivative that f has a local minimum at $b^{\frac{1}{p-1}}$ and thus f has a global minimum at $b^{\frac{1}{p-1}}$ due to the boundary conditions. Therefore, since

$$f\left(b^{\frac{1}{p-1}}\right) = \frac{1}{p}b^{\frac{p}{p-1}} + \frac{1}{q}b^q - b^{1+\frac{1}{p-1}} = \frac{1}{p}b^q + \frac{1}{q}b^q - b^q = 0,$$

we obtain that $f(x) \geq 0$ for all $x \in [0, \infty)$ as desired. ∎

Theorem D.1.7 (Hölder's Inequality). *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in \mathcal{L}_p(X, \mu)$ and $g \in \mathcal{L}_q(X, \mu)$, then $fg \in \mathcal{L}_1(X, \mu)$ and*

$$\int_X |fg| \, d\mu \leq \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}} \left(\int_X |g|^q \, d\mu \right)^{\frac{1}{q}}.$$

Proof. Let

$$\alpha = \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}} \qquad \text{and} \qquad \beta = \left(\int_X |g|^q \, d\mu \right)^{\frac{1}{q}}.$$

If $\alpha = 0$, then $|f|^p = 0$ almost everywhere. Hence $|f| = 0$ almost everywhere so $|fg| = 0$ almost everywhere and hence the inequality holds. Similarly, if $\beta = 0$ then the inequality holds. Hence we may assume that $\alpha, \beta > 0$.

Since $\alpha, \beta > 0$, we obtain that

$$\begin{aligned} \int_X |fg| \, d\mu &= \alpha\beta \int_X \frac{|f|}{\alpha} \frac{|g|}{\beta} \, d\mu \\ &\leq \alpha\beta \int_X \frac{|f|^p}{p\alpha^p} + \frac{|g|^q}{q\beta^q} \, d\mu \qquad \text{by Lemma D.1.6} \\ &= \alpha\beta \left(\frac{1}{p\alpha^p} \int_X |f|^p \, d\mu + \frac{1}{q\beta^q} \int_X |g|^q \, d\mu \right) \\ &= \alpha\beta \left(\frac{1}{p} + \frac{1}{q} \right) = \alpha\beta \end{aligned}$$

as desired. ∎

In addition to being used to prove that $\mathcal{L}_p(X, \mu)$ is a vector space, Hölder's inequality (Theorem D.1.7) also has following important corollary.

Corollary D.1.8. *Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$ and let $p \in (1, \infty)$. If $f \in \mathcal{L}_p(X, \mu)$, then $f \in \mathcal{L}_1(X, \mu)$ with*

$$\int_X |f| \, d\mu \leq \mu(X)^{\frac{1}{q}} \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

where $q \in (1, \infty)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Since $\mu(X) < \infty$, it is elementary to see that $1 \in \mathcal{L}_q(X, \mu)$; that is, the function that is one everywhere is q-integrable as

$$\int_X 1^q \, d\mu = \mu(X) < \infty.$$

Hence, by Hölder's inequality (Theorem D.1.7) $f = f1 \in \mathcal{L}_1(X, \mu)$ and

$$\int_X |f| \, d\mu \leq \mu(X)^{\frac{1}{q}} \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}.$$ ∎

Hölder's inequality (Theorem D.1.7) also enables us to show that the p-norm satisfies the triangle inequality modulo one technicality.

Theorem D.1.9 (Minkowski's Inequality). *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. If $f, g \in \mathcal{L}_p(X, \mu)$*

$$\left(\int_X |f+g|^p \, d\mu\right)^{\frac{1}{p}} \leq \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}}.$$

Proof. Let $f, g \in \mathcal{L}_p(X, \mu)$. Recall from Lemma D.1.2 that $f + g \in \mathcal{L}_p(X, \mu)$. Moreover, if $p = 1$ then

$$\int_X |f+g| \, d\mu \leq \int_X |f| + |g| \, d\mu \leq \int_X |f| \, d\mu + \int_X |g| \, d\mu$$

so the inequality holds.

Now suppose $p \in (1, \infty)$. Choose $q \in (1, \infty)$ so that $\frac{1}{p} + \frac{1}{q} = 1$. Thus $q = \frac{p}{p-1}$. Since $p \in (1, \infty)$, notice by Hölder's inequality (Theorem D.1.7) that

$$\begin{aligned}
&\int_X |f+g|^p \, d\mu \\
&= \int_X |f+g||f+g|^{p-1} \, d\mu \\
&\leq \int_X (|f| + |g|)|f+g|^{p-1} \, d\mu \\
&= \int_X |f||f+g|^{p-1} \, d\mu + \int_X |g||f+g|^{p-1} \, d\mu \\
&\leq \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} \left(\int_X (|f+g|^{p-1})^q \, d\mu\right)^{\frac{1}{q}} \\
&\qquad + \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}} \left(\int_X (|f+g|^{p-1})^q \, d\mu\right)^{\frac{1}{q}} \\
&= \left(\left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}}\right) \left(\int_X |f+g|^p \, d\mu\right)^{\frac{1}{q}}.
\end{aligned}$$

If $\int_X |f+g|^p \, d\mu = 0$, the result follows trivially. Otherwise, we may divide both sides of the equation by $(\int_X |f+g|^p \, d\mu)^{\frac{1}{q}}$ to obtain that

$$\left(\int_X |f+g|^p \, d\mu\right)^{\frac{1}{p}} = \left(\int_X |f+g|^p \, d\mu\right)^{1-\frac{1}{q}} \leq \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_X |g|^p \, d\mu\right)^{\frac{1}{p}}$$

as desired. ∎

Corollary D.1.10. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Then the p-norm is a norm on $L_p(X, \mu)$.*

Proof. To see that $\|\cdot\|_p$ is indeed a norm on $L_p(X,\mu)$, we first note by Remark D.1.5 that $\|\cdot\|_p$ is well-defined (i.e. its value does not depend on the representative of the equivalence class) and finite by the definition of $\mathcal{L}_p(X,\mu)$. Furthermore, notice that $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere if and only if $[f] = 0$. Furthermore, as clearly $\|\alpha f\|_p = |\alpha| \, \|f\|_p$ for all $\alpha \in \mathbb{C}$ and $f \in L_p(X,\mu)$, and as the triangle inequality holds by Minkowski's Inequality (Theorem D.1.9), we obtain that $\|\cdot\|_p$ is a norm on $L_p(X,\mu)$ as desired. ∎

Of course, the above did not deal with the case that $p = \infty$ as the formula for the norm does not make sense in this situation. To develop a notion of an ∞-norm for measurable functions, we define the following concept which is motivated by the fact that we don't need our functions to be bounded everywhere, just almost everywhere.

Definition D.1.11. Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f : X \to \mathbb{K}$ is said to be *essentially bounded* if there exists an $M \geq 0$ such that

$$\mu(\{x \in X \mid |f(x)| > M\}) = 0.$$

The set of essentially bounded functions on $(X, \mathcal{A}, \mu)$ is denoted $\mathcal{L}_\infty(X,\mu)$.

Of course, $\mathcal{L}_\infty(X,\mu)$ will not have a well-defined norm for the same reason that $\mathcal{L}_p(X,\mu)$ did not have a well-defined norm; we have to deal with functions that are equal almost everywhere. Notice if $f, g : X \to \mathbb{C}$ are such that f is essentially bounded and $f = g$ almost everywhere, then g is essentially bounded as the union of μ-measure zero sets has μ-measure zero. Hence we may define the following.

Definition D.1.12. Given a measure space $(X, \mathcal{A}, \mu)$, the *L_∞-space of* $(X, \mathcal{A}, \mu)$, denote $L_\infty(X,\mu)$, is

$$L_\infty(X,\mu) = \{[f] \mid f : X \to \mathbb{C} \text{ essentially bounded}\}.$$

Remark D.1.13. Given a measure space $(X, \mathcal{A}, \mu)$ and $f, g \in \mathcal{M}(X, \mathbb{C})$ such that $[f] = [g]$, we have seen that $f \in \mathcal{L}_\infty(X,\mu)$ if and only if $g \in \mathcal{L}_\infty(X,\mu)$. In particular, every representative of an equivalence class in $L_\infty(X,\mu)$ is an element of $\mathcal{L}_\infty(X,\mu)$. Because of this and to abuse notation, we will consider elements of $\mathcal{L}_\infty(X,\mu)$ as elements of $L_\infty(X,\mu)$ and drop the explicit reminder that we are dealing with an equivalence class in most (if not all) arguments.

Theorem D.1.14. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then $L_\infty(X,\mu)$ is a normed linear space with respect to the norm*

$$\|f\|_\infty = \inf\{M \geq 0 \mid \mu(\{x \mid |f(x)| > M\}) = 0\}.$$

Proof. First we claim that $L_\infty(X,\mu)$ is a subspace of $\mathcal{M}(X,\mathbb{C})/\sim$ and thus a vector space over $\mathbb{C}$. To see this, let $f,g \in L_\infty(X,\mu)$ be arbitrary. Then there exists $M_1, M_2 \geq 0$ such that

$$\mu(\{x \mid |f(x)| > M_1\}) = 0 \qquad \text{and} \qquad \mu(\{x \mid |g(x)| > M_2\}) = 0.$$

Hence as

$$\begin{aligned}&\{x \mid |f(x)+g(x)| > M_1+M_2\}\\ &\subseteq \{x \mid |f(x)|+|g(x)| > M_1+M_2\}\\ &\subseteq \{x \mid |f(x)| > M_1\} \cup \{x \mid |g(x)| > M_2\}\end{aligned}$$

we see that

$$\begin{aligned}&\mu(\{x \mid |f(x)+g(x)| > M_1+M_2\})\\ &\leq \mu(\{x \mid |f(x)| > M_1\}) + \mu(\{x \mid |g(x)| > M_2\}) = 0.\end{aligned}$$

Hence $f+g \in \mathcal{L}_\infty(X,\mu)$. Further for all $\alpha \in \mathbb{C}$

$$\mu(\{x \mid |\alpha f(x)| > |\alpha| M\}) = 0$$

so $\alpha f \in \mathcal{L}_\infty(X,\mu)$. Hence, as $0 \in \mathcal{L}_\infty(X,\mu)$, we have shown that $L_\infty(X,\mu)$ is a subspace of $\mathcal{M}(X,\mathbb{C})/\sim$ and thus a vector space over $\mathbb{C}$.

To see that $\|\cdot\|_\infty$ is a well-defined norm on $L_\infty(X,\mu)$, first notice that if $f = g$ almost everywhere and $M \geq 0$ then

$$\mu(\{x \mid |f(x)| > M\}) = 0 \quad \text{if and only if} \quad \mu(\{x \mid |g(x)| > M\}) = 0.$$

Hence $\|\cdot\|_\infty$ is well-defined. Furthermore, notice that $\|f\|_\infty < \infty$ for all $f \in \mathcal{L}_\infty(X,\mu)$ by the definition of an essentially bounded function. Next notice that $\|f\|_\infty \geq 0$ with equality if and only if

$$\mu\left(\left\{x \,\middle|\, |f(x)| > \frac{1}{n}\right\}\right) = 0$$

for all $n \in \mathbb{N}$ if and only if

$$\mu\left(\{x \mid |f(x)| > 0\}\right) = \mu\left(\bigcup_{n\geq 1}\left\{x \,\middle|\, |f(x)| > \frac{1}{n}\right\}\right) = 0$$

if and only if $f = 0$ almost everywhere if and only if $f = 0$ in $L_\infty(X,\mu)$.

Next let $\alpha \in \mathbb{C}$ and $f \in L_\infty(X,\mu)$ be arbitrary. If $\alpha = 0$, then clearly $\|\alpha f\|_\infty = 0 = |\alpha|\,\|f\|_\infty$. Otherwise, if $\alpha \neq 0$, we see that

$$\begin{aligned}\|\alpha f\|_\infty &= \inf\{M \geq 0 \mid \mu(\{x \mid |\alpha f(x)| > M\}) = 0\}\\ &= \inf\left\{M \geq 0 \,\middle|\, \mu\left(\left\{x \,\middle|\, |f(x)| > \frac{M}{|\alpha|}\right\}\right) = 0\right\}\\ &= \inf\{|\alpha| M' \geq 0 \mid \mu(\{x \mid |f(x)| > M'\}) = 0\}\\ &= |\alpha|\,\|f\|_\infty\end{aligned}$$

as desired.

Finally, to verify that $\|\cdot\|_\infty$ satisfies the triangle inequality, let $f, g \in L_\infty(X, \mu)$ be arbitrary. If $M_1, M_2 \geq 0$ are such that

$$\mu(\{x \mid |f(x)| > M_1\}) = 0 \qquad \text{and} \qquad \mu(\{x \mid |g(x)| > M_2\}) = 0,$$

the above shows that

$$\mu(\{x \mid |f(x) + g(x)| > M_1 + M_2\}) = 0.$$

Hence

$$\|f + g\|_\infty \leq M_1 + M_2.$$

Therefore, as this holds for all such M_1 and M_2, we obtain that

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

as desired. ∎

Remark D.1.15. If $f \in L_\infty(X, \mu)$, then $\mu(\{x \in X \mid |f(x)| > \|f\|_\infty\}) = 0$. To see this, for each $n \in \mathbb{N}$ let

$$A_n = \left\{ x \in X \,\middle|\, |f(x)| > \|f\|_\infty + \frac{1}{n} \right\}.$$

Then each A_n is measurable. Furthermore, by the definition of $\|f\|_\infty$, we obtain that $\mu(A_n) = 0$. Therefore, as

$$\{x \in X \mid |f(x)| > \|f\|_\infty\} = \bigcup_{n=1}^\infty A_n,$$

the claim follows by the Monotone Convergence Theorem for measures or simply the subadditivity of measures. Hence $|f(x)| \leq \|f\|_\infty$ almost everywhere.

Remark D.1.16. If $f \in C[a, b]$, then the Extreme Value Theorem implies f is bounded. Thus, as f is Lebesgue measurable, $f \in L_\infty([a, b], \lambda)$. In addition, it is not difficult to verify that two notions of the ∞-norm agree. Indeed if

$$M_0 = \sup(\{|f(x)| \mid x \in [a, b]\}) \geq 0$$

then clearly

$$\lambda(\{x \in [a, b] \mid |f(x)| > M\}) = 0.$$

Hence

$$\sup(\{|f(x)| \mid x \in [a, b]\}) \geq \inf\{M \geq 0 \mid \lambda(\{x \mid |f(x)| > M\}) = 0\}.$$

For the reverse inequality, suppose

$$0 \leq M < \sup(\{|f(x)| \mid x \in [a, b]\}).$$

By the Extreme Value Theorem, there exists an $x_0 \in [a,b]$ such that

$$|f(x_0)| = \sup(\{|f(x)| \mid x \in [a,b]\}) > M.$$

However, if $\epsilon = \frac{1}{2}(|f(x_0)| - M) > 0$, there exists a $\delta > 0$ such that if $x \in [a,b]$ and $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \epsilon$. Hence, as $(x_0 - \delta, x_0 + \delta) \cap [a,b]$ has non-zero λ-measure and

$$|f(x)| > |f(x_0)| - \epsilon = \frac{1}{2}(|f(x_0)| + M) > M$$

for all $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$, we see that

$$\lambda(\{x \in [a,b] \mid |f(x)| > M\}) > 0.$$

Thus it follows that

$$\sup(\{|f(x)| \mid x \in [a,b]\}) = \inf\{M \geq 0 \mid \lambda(\{x \mid |f(x)| > M\}) = 0\}.$$

as desired.

Clearly essentially bounded functions behave like bounded functions when it comes to integration.

Theorem D.1.17 (Hölder's Inequality). *Let $(X, \mathcal{A}, \mu)$ be a measure space. If $f \in \mathcal{L}_1(X, \mu)$ and $g \in \mathcal{L}_\infty(X, \mu)$, then $fg \in \mathcal{L}_1(X, \mu)$ and*

$$\|fg\|_1 \leq \|f\|_1 \|g\|_\infty .$$

Proof. As $|g| \leq \|g\|_\infty$ almost everywhere by Remark D.1.15, we obtain that

$$\|fg\|_1 = \int_X |f||g| \, d\mu \leq \int_X |f| \, \|g\|_\infty \, d\mu = \|f\|_1 \|g\|_\infty$$

as desired. ∎

Corollary D.1.18. *Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$ and let $p \in [1, \infty)$. If $f \in L_\infty(X, \mu)$, then $f \in L_p(X, \mu)$ and*

$$\|f\|_p \leq \|f\|_\infty \mu(X)^{\frac{1}{p}}$$

Proof. Since $\mu(X) < \infty$, it is elementary to see that

$$\begin{aligned} \left(\int_X |f|^p \, d\mu\right)^{\frac{1}{p}} &\leq \left(\int_X \|f\|_\infty^p \, d\mu\right)^{\frac{1}{p}} \\ &= (\|f\|_\infty^p \mu(X))^{\frac{1}{p}} = \|f\|_\infty \mu(X)^{\frac{1}{p}} < \infty. \end{aligned}$$

Hence $f \in L_p(X, \mu)$. ∎

D.2 L_p-Spaces are Banach Spaces

In this section, we will use measure theoretic techniques to show that all L_p spaces are Banach spaces. First we begin with the case $p \neq \infty$.

Theorem D.2.1 (Riesz-Fisher Theorem). *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. Then $L_p(X, \mu)$ is a Banach space.*

Proof. To see that $L_p(X, \mu)$ is a Banach space, let $(f_n)_{n\geq 1}$ be an arbitrary Cauchy sequence in $L_p(X, \mu)$ (of course this really means an Cauchy sequence of equivalence classes, each of which is represented by a function $f_n \in \mathcal{L}_p(X, \mu)$). Since $(f_n)_{n\geq 1}$ is Cauchy, it is not difficult to see that there exists a subsequence $(f_{k_n})_{n\geq 1}$ such that

$$\left\|f_{k_{n+1}} - f_{k_n}\right\|_p < \frac{1}{2^n}$$

for all $n \in \mathbb{N}$ (i.e. choose $k_n \in \mathbb{N}$ to be a natural number greater than k_{n-1} that works in the definition of a Cauchy sequence for $\epsilon = \frac{1}{2^n}$). As $(f_n)_{n\geq 1}$ is Cauchy, it suffices to show that $(f_{k_n})_{n\geq 1}$ converges to some element in $L_p(X, \mu)$.

Define a function $g : X \to [0, \infty]$ by

$$g(x) = |f_{k_1}(x)| + \sum_{n=1}^{\infty} |f_{k_{n+1}}(x) - f_{k_n}(x)|$$

for all $x \in X$. As the sum of measurable functions is measurable, the absolute value of measurable functions is measurable, and the pointwise limit of measurable functions is measurable, we obtain that g is a measurable function. Furthermore, as g is the pointwise limit of

$$\left(|f_{k_1}| + \sum_{n=1}^{m} |f_{k_{n+1}} - f_{k_n}|\right)_{m\geq 1},$$

we obtain by Fatou's Lemma and Minkowski's inequality (Theorem D.1.9) that

$$\begin{aligned}
\left(\int_X |g|^p\right)^{\frac{1}{p}} d\mu &\leq \liminf_{m\to\infty} \left(\int_X \left(|f_{k_1}(x)| + \sum_{n=1}^{m} |f_{k_{n+1}}(x) - f_{k_n}(x)|\right)^p d\mu\right)^{\frac{1}{p}} \\
&\leq \liminf_{m\to\infty} \left\|f_{k_1}\right\|_p + \sum_{n=1}^{m} \left\|f_{k_{n+1}} - f_{k_n}\right\|_p \\
&= \left\|f_{k_1}\right\|_p + 1 < \infty.
\end{aligned}$$

Hence $g \in L_p(X, \mu)$.

Note if $A = \{x \in X \mid g(x) = \infty\}$, then $A \in \mathcal{A}$ and $\mu(A) = 0$. By replacing each f_n with $f_n\chi_{A^c}$ (which does not affect the equivalence classes as $f_n = f_n\chi_{A^c}$ almost everywhere), we may assume that $g(x) < \infty$ for all $x \in X$.

As $g(x) < \infty$ for all $x \in X$ and as $\mathbb{C}$ is complete so every absolutely summable sequence is summable by Theorem 2.2.2, we obtain that the function $f : X \to \mathbb{C}$ defined by

$$f(x) = f_{k_1}(x) + \sum_{n=1}^{\infty} f_{k_{n+1}}(x) - f_{k_n}(x)$$

for all $x \in X$ is well-defined. Notice for all $m \in \mathbb{N}$ that

$$f_{k_1} + \sum_{n=1}^{m} f_{k_{n+1}} - f_{k_n} = f_{k_m}.$$

Hence $|f_{k_m}| \leq g$ for all $m \in \mathbb{N}$ and

$$f(x) = \lim_{n\to\infty} f_{k_n}(x)$$

for all $x \in X$. Hence f is measurable being the pointwise limit of measurable functions. Furthermore, as clearly $|f| \leq g$, we obtain that $f \in L_p(X, \mu)$.

We claim that $(f_{k_n})_{n\geq 1}$ converges to f in the p-norm. To see this, notice since $|f|^p, |f_{k_m}|^p \leq g^p$ for all $m \in \mathbb{N}$ that

$$|f - f_{k_m}|^p \leq (|f| + |f_{k_m}|)^p \leq (2|g|)^p = 2^p|g|^p.$$

Therefore, since $g \in L_p(X, \mu)$ and since $(|f - f_{k_m}|^p)_{m\geq 1}$ converges pointwise to zero, the Dominated Convergence Theorem implies that that

$$\lim_{m\to\infty} \int_X |f - f_{k_m}|^p \, d\mu = 0.$$

Hence $(f_{k_n})_{n\geq 1}$ converges to f with respect to $\|\cdot\|_p$. Therefore, as $(f_n)_{n\geq 1}$ was Cauchy, we obtain that $(f_n)_{n\geq 1}$ converges to f in $L_p(X, \mu)$. Thus, as $(f_n)_{n\geq 1}$ was an arbitrary Cauchy sequence in $L_p(X, \mu)$, we obtain that $L_p(X, \mu)$ is complete. ∎

Notice the proof of the Riesz-Fisher Theorem (Theorem D.2.1) immediately implies the following.

Corollary D.2.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space, let $p \in [1, \infty)$, and let $f \in L_p(X, \mu)$. If $(f_n)_{n\geq 1}$ is a sequence of elements of $L_p(X, \mu)$ that converge to f in $L_p(X, \mu)$, then there exists a subsequence $(f_{k_n})_{n\geq 1}$ of $(f_n)_{n\geq 1}$ that converges to f pointwise almost everywhere.*

Proof. As $(f_n)_{n\geq 1}$ converges to f in $L_p(X, \mu)$, $(f_n)_{n\geq 1}$ is Cauchy in $L_p(X, \mu)$. Therefore the proof of the Riesz-Fisher Theorem (Theorem D.2.1) implies there exists a subsequence $(f_{k_n})_{n\geq 1}$ of $(f_n)_{n\geq 1}$ that converges both pointwise almost everywhere and in $L_p(X, \mu)$ to some function h (i.e. $h(x) = f_{k_1}(x) + \sum_{n=1}^{\infty} f_{k_{n+1}}(x) - f_{k_n}(x)$ for all $x \in X$). Therefore, as limits in normed linear spaces are unique, we obtain that $h = f$ almost everywhere thereby completing the proof. ∎

For those familiar with undergraduate real analysis, it should not be surprising that the $p = 2$ case is special.

Corollary D.2.3. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then $L_2(X, \mu)$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle : L_2(X, \mu) \times L_2(X, \mu) \to \mathbb{C}$ defined by*

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu.$$

Proof. First, clearly if $g \in L_2(X, \mu)$ then $\overline{g} \in L_2(X, \mu)$ and $\|g\|_2 = \|\overline{g}\|_2$. Hence, by Hölder's inequality (Theorem D.1.7), we see that if $f, g \in L_2(X, \mu)$ then $f\overline{g} \in L_1(X, \mu)$ so

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu$$

is a well-defined element of $\mathbb{C}$. Hence, as in addition the definition of $\langle f, g\rangle$ does not depend on the representative of the equivalence classes of f and g selected, $\langle \cdot, \cdot \rangle$ is well-defined.

It is not difficult to see that $\langle f, f\rangle \geq 0$ for all $f \in L_2(X, \mu)$ with equality if and only if $f = 0$ almost everywhere, that $\langle \cdot, \cdot \rangle$ is linear in the first entry by the linearity of the integral, and that

$$\overline{\langle f, g\rangle} = \langle g, f\rangle.$$

Hence $\langle \cdot, \cdot \rangle$ is an inner product on $L_2(X, \mu)$. As $\|f\|_2 = \sqrt{\langle f, f\rangle}$ for all $f \in L_2(X, \mu)$, we obtain that $L_2(X, \mu)$ is a Hilbert space by Theorem D.2.1. ∎

Unsurprisingly, we also have the following.

Theorem D.2.4 (Riesz-Fisher Theorem). *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then $L_\infty(X, \mu)$ is a Banach space.*

Proof. To see that $L_\infty(X, \mu)$ is a Banach space, let $(f_n)_{n\geq 1}$ be an arbitrary Cauchy sequence in $L_\infty(X, \mu)$. For each $n \in \mathbb{N}$ let

$$A_n = \{x \in X \mid |f_n(x)| > \|f_n\|_\infty\}$$

and for each $n, m \in \mathbb{N}$ let

$$B_{n,m} = \{x \in X \mid |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}.$$

Hence, by Remark D.1.15, each A_n and $B_{n,m}$ are measurable for all $n, m \in \mathbb{N}$, $\mu(A_n) = 0$ for all $n \in \mathbb{N}$, and $\mu(B_{n,m}) = 0$ for all $n, m \in \mathbb{N}$. Let

$$B = \left(\bigcup_{n=1}^\infty A_n\right) \bigcup \left(\bigcup_{n,m=1}^\infty B_{n,m}\right).$$

Then B is a measurable set and $\mu(B) = 0$ as B is a countable union of μ-measure zero sets.

By replacing each f_n with $f_n \chi_{B^c}$ (which doesn't affect the equivalence classes), we may assume that $|f_n(x)| \leq \|f_n\|_\infty$ for all $x \in X$ and $n \in \mathbb{N}$, and that $|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$ for all $x \in X$ and $n, m \in \mathbb{N}$. By this assumption, for each $x \in X$ we see that $(f_n(x))_{n \geq 1}$ is a Cauchy sequence in $\mathbb{C}$ and thus converges. Hence the function $f : X \to \mathbb{C}$ defined by

$$f(x) = \lim_{n \to \infty} f_n(x)$$

for all $x \in X$ is well-defined and measurable.

We claim that $f \in \mathcal{L}_\infty(X, \mu)$ and that $(f_n)_{n \geq 1}$ converges to f in $L_\infty(X, \mu)$. To see this, notice for all $x \in X$ and $n \in \mathbb{N}$ that

$$|f(x) - f_n(x)| = \lim_{m \to \infty} |f_m(x) - f_n(x)| \leq \limsup_{m \to \infty} \|f_m - f_n\|_\infty .$$

Hence

$$\sup\{|f(x) - f_n(x)| \mid x \in X\} \leq \limsup_{m \to \infty} \|f_m - f_n\|_\infty$$

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

$$\begin{aligned}\sup\{|f(x) - f_1(x)| \mid x \in X\} &\leq \limsup_{m \to \infty} \|f_m - f_1\|_\infty \\ &\leq \limsup_{m \to \infty} \|f_m\|_\infty + \|f_1\|_\infty < \infty\end{aligned}$$

as Cauchy sequences are bounded. Hence by the definition of essentially bounded functions we see that $f - f_1 \in \mathcal{L}_\infty(X, \mu)$. Hence, as $f_1 \in \mathcal{L}_\infty(X, \mu)$ and $\mathcal{L}_\infty(X, \mu)$ is closed under addition, we see that $f \in \mathcal{L}_\infty(X, \mu)$. Thus the above shows that

$$\|f - f_n\|_\infty \leq \limsup_{m \to \infty} \|f_m - f_n\|_\infty$$

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

$$\lim_{n \to \infty} \limsup_{m \to \infty} \|f_m - f_n\|_\infty = 0$$

since $(f_n)_{n \geq 1}$ is Cauchy, we obtain that $(f_n)_{n \geq 1}$ converges to f in $L_\infty(X, \mu)$. Hence, as $(f_n)_{n \geq 1}$ was an arbitrary Cauchy sequence, we obtain that $L_\infty(X, \mu)$ is complete. ∎

D.3 Dense Subset of L_p-Spaces

To conclude this section, we note specific types of functions are dense in the L_p-spaces (well, when $p \neq \infty$).

Theorem D.3.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space. The set*

$$\mathcal{F} = \operatorname{span}\left\{\varphi : X \to [0,\infty) \,\middle|\, \substack{\varphi \text{ is simple and there exists a } A\in\mathcal{A} \\ \text{such that } \mu(A)<\infty \text{ and } \varphi|_{A^c}=0}\right\}$$

is dense in $L_p(X, \mu)$ for all $p \in [1, \infty)$.

Proof. Fix $p \in [1, \infty)$. To begin, we claim that if $\varphi : X \to [0, \infty)$ is a simple function, then $\varphi \in L_p(X, \mu)$ if and only if there exists an $A \in \mathcal{A}$ such that $\mu(A) < \infty$ and $\varphi|_{A^c} = 0$. Indeed, suppose there exists an $A \in \mathcal{A}$ such that $\mu(A) < \infty$ and $\varphi|_{A^c} = 0$. Since φ is a simple function, φ is essentially bounded. Hence the proof of Corollary D.1.18 yields

$$\|\varphi\|_p \leq \|\varphi\|_\infty^{\frac{1}{p}} \mu(A)^{\frac{1}{p}} < \infty$$

so $\varphi \in L_p(X, \mu)$. Conversely, suppose $\varphi \in L_p(X, \mu)$. Clearly if $\varphi = 0$ the result is true. Hence suppose $\varphi \neq 0$. By the definition of a simple function, there exists pairwise disjoint sets $\{A_k\}_{k=1}^n \subseteq \mathcal{A}$ and elements $\{a_k\}_{k=1}^n \subseteq (0, \infty)$ such that $\varphi = \sum_{k=1}^n a_k \chi_{A_k}$ (where we have removed the characteristic function on which φ is zero). If $c = \min\{a_1, \ldots, a_n\} > 0$, then we see that

$$c^p \mu\left(\bigcup_{k=1}^n A_k\right) = c^p \sum_{k=1}^n \mu(A_k) \leq \sum_{k=1}^n a_k^p \mu(A_k) = \int_X \varphi^p \, d\mu < \infty.$$

Therefore, if $A = \bigcup_{k=1}^n A_k \in \mathcal{A}$, then $\mu(A) < \infty$ and $\varphi|_{A^c} = 0$ as desired.

To demonstrate the theorem, we must first show that $\mathcal{F} \subseteq L_p(X, \mu)$. However, this follows from the above claim as $\mathcal{F}$ is a span of elements of $L_p(X, \mu)$ and thus is a subspace of $L_p(X, \mu)$.

Finally, to show that $\mathcal{F}$ is dense in $L_p(X, \mu)$, it suffices (as $\mathcal{F}$ and $L_p(X, \mu)$ are closed under linear combinations) to show that if $f \in L_p(X, \mu)$ and $f \geq 0$ then there exists a sequence $(\varphi_n)_{n\geq1}$ of elements of $\mathcal{F}$ such that $\lim_{n\to\infty} \|f - \varphi_n\|_p = 0$. Indeed notice it is easy to see that the positive and negative parts of the real and imaginary parts of f are smaller than $|f|$ and thus elements of $L_p(X, \mu)$. If we can approximate each of these non-negative functions in $L_p(X, \mu)$ via elements of $\mathcal{F}$, then the triangle inequality will yield the result.

Fix $f \in L_p(X, \mu)$ such that $f \geq 0$. As f is non-negative, there exists an increasing sequence of simple functions $(\varphi_n)_{n\geq1}$ that converge to f pointwise. Hence $0 \leq \varphi_n \leq f$ so

$$\int_X |\varphi_n|^p \, d\mu \leq \int_X |f|^p \, d\mu < \infty.$$

Hence $\varphi_n \in L_p(X, \mu)$ so $\varphi_n \in \mathcal{F}$ by the result at the beginning of the proof. Moreover, since $(|f - \varphi_n \chi_{A_n}|^p)_{n\geq1}$ converges to zero pointwise and since

$$|f - \varphi_n \chi_{A_n}|^p \leq |f|^p \in L_1(X, \mu),$$

we obtain by the Dominated Convergence Theorem that

$$\lim_{n\to\infty} \int_X |f - \varphi_n \chi_{A_n}|^p \, d\mu = 0.$$

Hence $\lim_{n\to\infty} \|f - \varphi_n \chi_{A_n}\|_p = 0$ as desired. ∎

Theorem D.3.2. *For all* $p \in [1, \infty)$*,*

$$\mathcal{C}_c(\mathbb{R}, \mathbb{C}) = \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, {\textit{f is continuous and there exists a compact set} \atop K \subseteq \mathbb{C} \textit{ such that } f|_{K^c} = 0} \right\}$$

is dense in $L_p(\mathbb{R}, \lambda)$.

Proof. By Theorem D.3.1 we know that

$$\mathcal{F} = \operatorname{span}\left\{ \varphi : \mathbb{R} \to [0, \infty) \,\middle|\, {\varphi \text{ is simple and there exists a } A \in \mathcal{M}(R) \atop \text{such that } \mu(A) < \infty \text{ and } \varphi|_{A^c} = 0} \right\}$$

is dense in $L_p(\mathbb{R}, \lambda)$. However, if $\varphi_n : \mathbb{R} \to [0, \infty)$ is simple and $\varphi \in L_p(\mathbb{R}, \lambda)$, then the end of the proof of Theorem D.3.1 can be used to show that $\varphi \chi_{[-n,n]}$ converges to φ in $L_p(\mathbb{R}, \lambda)$. Therefore, as Corollary D.1.18 implies that $\mathcal{C}_c(\mathbb{R}, \mathbb{C}) \subseteq L_p(\mathbb{R}, \lambda)$, to show that $\mathcal{C}_c(\mathbb{R}, \mathbb{C})$ is dense in $L_p(\mathbb{R}, \lambda)$, it suffices by the triangle inequality to show that each simple function φ such that $\varphi|_{[-n,n]^c} = 0$ for some $n \in \mathbb{N}$ can be approximated in $\|\cdot\|_p$ by an element of $\mathcal{C}_c(\mathbb{R}, \mathbb{C})$.

To see the above, let φ be an arbitrary simple function such that $\varphi|_{[-n,n]^c} = 0$ for some $n \in \mathbb{N}$ and let $\epsilon > 0$ be arbitrary. By Lusin's Theorem there exists a continuous function $f : [-n, n] \to \mathbb{C}$ such that

$$\lambda(\{x \in [-n, n] \mid f(x) \neq \varphi(x)\}) < \epsilon$$

and

$$\sup\{|f(x)| \mid x \in [-n, n]\} \leq \|\varphi\|_\infty < \infty.$$

Extend f to a continuous function $g : \mathbb{R} \to \mathbb{C}$ by defining

$$g(x) = \begin{cases} f(x) & \text{if } x \in [-n, n] \\ -\frac{f(x)}{\epsilon}(x - n) + f(x) & \text{if } x \in [n, n + \delta) \\ \frac{f(x)}{\epsilon}(x + n) + f(x) & \text{if } x \in (-n - \delta, -n] \\ 0 & \text{otherwise} \end{cases}.$$

Clearly $g \in \mathcal{C}_c(\mathbb{R}, \mathbb{C})$ and it is easy to see that $\|g\|_\infty \leq \|\varphi\|_\infty$ as we extended

f to g using linear functions connecting $f(\pm n)$ to 0. Therefore, since

$$\begin{aligned}
&\int_{\mathbb{R}} |g-\varphi|^p \, d\lambda \\
&= \int_{[-n,n]} |f-\varphi|^p \, d\lambda + \int_{[n,n+\epsilon)\cup(-n-\epsilon,-n]} |g|^p \, d\lambda \\
&= \int_{\{x\in[-n,n] \,|\, f(x)\neq\varphi(x)\}} |f-\varphi|^p \, d\lambda + \int_{[n,n+\epsilon)\cup(-n-\epsilon,-n]} |g|^p \, d\lambda \\
&\leq \int_{\{x\in[-n,n] \,|\, f(x)\neq\varphi(x)\}} (2\left\|\varphi\right\|_\infty)^p \, d\lambda + \int_{[n,n+\epsilon)\cup(-n-\epsilon,-n]} \left\|\varphi\right\|_\infty^p \, d\lambda \\
&\leq (2\left\|\varphi\right\|_\infty)^p\epsilon + 2\epsilon \left\|\varphi\right\|_\infty^p \\
&= (2^p+2)\left\|\varphi\right\|_\infty^p \epsilon
\end{aligned}$$

the proof is complete as $\left\|\varphi\right\|_\infty$ is fixed and $\epsilon > 0$ was arbitrary. ∎

By using more topological concepts, the following can be demonstrated.

Theorem D.3.3. *Let $(X, \mathcal{T})$ be a locally compact Hausdorff space, let μ be a regular measure on the Borel subets of $(X, \mathcal{T})$ such that $\mu(K) < \infty$ for all compact subsets $K \subseteq X$. For all $p \in [1, \infty)$,*

$$\mathcal{C}_c(X, \mathbb{C}) = \left\{ f : X \to \mathbb{C} \,\middle|\, \substack{f \text{ is continuous and there exists a compact set} \\ K\subseteq X \text{ such that } f|_{K^c}=0} \right\}$$

is dense in $L_p(X, \mu)$.

Proof. As Lusin's Theorem holds in this context, the proof of Theorem D.3.2 can be adapted with the use of Urysohn's Lemma and the regularity of μ. ∎

D.4 Dual Spaces from Measure Theory

Using measure theory, several examples of linear functionals and dual spaces can be discussed.

Example D.4.1. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, let $p, q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, and let $g \in L_q(X, \mu)$. Define $\Psi_g : L_p(X, \mu) \to \mathbb{C}$ by

$$\Psi_g(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$. To see that T is well-defined, notice since $g \in L_q(X, \mu)$ and $\frac{1}{p} + \frac{1}{q} = 1$ that $fg \in L_1(X, \mu)$ for all $f \in L_p(X, \mu)$ by Hölder's Inequality (Theorems D.1.7 and D.1.17). Hence Ψ_g is well-defined. Furthermore, clearly Ψ_g is linear.

To see that Ψ_g is continuous, notice for all $f \in L_p(X, \mu)$ that

$$|\Psi_g(f)| = \left|\int_X fg \, d\mu\right| \leq \int_X |fg| \, d\mu \leq \left\|f\right\|_p \left\|g\right\|_q$$

by Hölder's Inequality. Hence $\Psi_g \in L_p(X,\mu)^*$ with $\|\Psi_g\| \leq \|g\|_q$.

We claim that $\|\Psi_g\| = \|g\|_q$. To see this, we divide the discussion into three cases.

Case 1: $q = 1$. In this case $p = \infty$. Consider $f : X \to \mathbb{C}$ defined by

$$f(x) = \mathrm{sgn}(g)(x) = \begin{cases} \frac{|g(x)|}{g(x)} & \text{if } g(x) \neq 0 \\ 1 & \text{if } g(x) = 0 \end{cases}$$

for all $x \in X$. It is not difficult to see that f is measurable with $|f(x)| = 1$ for all $x \in X$ and thus $f \in L_\infty(X,\mu)$ with $\|f\|_\infty = 1$. Therefore, since

$$\Psi_g(f) = \int_X fg\, d\mu = \int_X |g|\, d\mu = \|g\|_1,$$

we see that $\|\Psi_g\| \geq \|g\|_1$ and thus $\|\Psi_g\| = \|g\|_1$ as desired.

Case 2: $1 < q < \infty$. In this case $1 < p < \infty$. Let $f = \mathrm{sgn}(g)|g|^{\frac{q}{p}}$. Clearly f is a well-defined measurable function since $1 < p, q < \infty$. We claim that $f \in L_p(X,\mu)$. To see this, notice

$$\left(\int_X |f|^p\, d\mu\right)^{\frac{1}{p}} = \left(\int_X |g|^q\, d\mu\right)^{\frac{1}{p}} = \|g\|_q^{\frac{q}{p}} < \infty$$

as $|\mathrm{sgn}(g)| = 1$ and $g \in L_q(X,\mu)$. Hence $f \in L_p(X,\mu)$ with $\|f\|_p = \|g\|_q^{\frac{q}{p}}$. Therefore, since

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \implies \quad \frac{q}{p} + 1 = q$$

we see that

$$\begin{aligned}\Psi_g(f) &= \int_X fg\, d\mu \\ &= \int_X |g|^{\frac{q}{p}+1}\, d\mu \\ &= \int_X |g|^q\, d\mu \\ &= \|g\|_q^q \\ &= \|g\|_q \|g\|_q^{\frac{q}{p}} \\ &= \|g\|_q \|f\|_p.\end{aligned}$$

If $f = 0$ then clearly $g = 0$ and the result follows. Otherwise if $h = \frac{1}{\|f\|_p} f$ then $h \in L_p(X,\mu)$, $\|h\|_p = 1$, and the above computation implies that

$$\Psi_g(h) = \|g\|_q.$$

Therefore $\|\Psi_g\| \geq \|g\|_q$ and thus $\|\Psi_g\| = \|g\|_q$ as desired.

Case 3: $q = \infty$. In this case $p = 1$. Notice the previous cases did not require μ to be σ-finite whereas we will need to use σ-finiteness here. To begin, as μ is σ-finite there exists a collection $\{X_n\}_{n=1}^\infty \subseteq \mathcal{A}$ such that $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$, $X = \bigcup_{n=1}^\infty X_n$, and $X_n \subseteq X_{n+1}$ for all $n \in \mathbb{N}$.

Let $\epsilon > 0$ be arbitrary and let

$$A_\epsilon = \{x \in X \mid |g(x)| > \|g\|_\infty - \epsilon\}.$$

Since $g \in L_\infty(X, \mu)$, we know that $\mu(A_\epsilon) > 0$. For each $n \in \mathbb{N}$ let $B_n = A_\epsilon \cap X_n$. Then clearly $A_\epsilon = \bigcup_{n=1}^\infty B_n$ and $B_n \subseteq B_{n+1}$ for all $n \in \mathbb{N}$. Therefore $\mu(A_\epsilon) = \lim_{n\to\infty} \mu(B_n)$ by the Monotone Convergence Theorem for Measures. Moreover, since $0 \leq \mu(B_n) \leq \mu(X_n) < \infty$ for all $n \in \mathbb{N}$, there exists an $N \in \mathbb{N}$ such that

$$0 < \mu(B_N) < \infty.$$

Let $f = \frac{1}{\mu(B_N)} \chi_{B_N} \mathrm{sgn}(g)$. Then f is clearly measurable with

$$\int_X |f| \, d\mu = \frac{1}{\mu(B_N)} \int_X \chi_{B_N} \, d\mu = 1.$$

Therefore, since

$$\begin{aligned}
\Psi_g(f) &= \int_X fg \, d\mu \\
&= \frac{1}{\mu(B_N)} \int_X \chi_{B_N} |g| \, d\mu \\
&\geq \frac{1}{\mu(B_N)} \int_X \chi_{B_N} (\|g\|_\infty - \epsilon) \, d\mu \\
&= \|g\|_\infty - \epsilon
\end{aligned}$$

as $B_N \subseteq A_\epsilon$, we obtain that $\|\Psi_g\| \geq \|g\|_\infty - \epsilon$. Hence, as $\epsilon > 0$ was arbitrary, the result follows.

Remark D.4.2. Notice as a direct corollary Example D.4.1 that if $(X, \mathcal{A}, \mu)$ is a σ-finite measure space and $p, q \in [1, \infty]$ are such that $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\|g\|_q = \|\Psi_g\| = \sup\left\{ \left| \int_X fg \, d\mu \right| \;\middle|\; f \in L_p(X, \mu), \|f\|_p \leq 1 \right\}$$

for all $g \in L_q(X, \mu)$. This alternative way to compute the norm can be useful on occasion.

The linear functionals from Example D.4.1 are all there are as the following theorem shows. In particular, this theorem lets us represent the complicated dual spaces of certain normed linear spaces using better understood normed linear spaces.

Theorem D.4.3 (Riesz Representation Theorem for L_p-Spaces). *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, let $1 \leq p < \infty$, and let $1 < q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. If $\Psi \in L_p(X, \mu)^*$ then there exists a unique $g \in L_q(X, \mu)$ such that*

$$\Psi(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$. Moreover $\|\Psi\| = \|g\|_q$. In particular, $L_p(X, \mu)^* = L_q(X, \mu)$.

First note that the norm estimates in the Riesz Representation Theorem for L_p-spaces (Theorem D.4.3) immediately follow for Example D.4.1. Thus it suffices to prove given a continuous linear functional on $L_p(X, \mu)$ that there is one and exactly one element of $L_q(X, \mu)$ that, via Example D.4.1, produces the continuous linear functional.

To begin, we desire to reduce to the setting that our functions are real-valued. Thus, let $L_p(X, \mu)_{\mathbb{R}}$ denote the real-valued p-integrable function and consider the following.

Lemma D.4.4. *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, let $1 \leq p < \infty$, and let $\Psi \in L_p(X, \mu)^*$. Then there exists continuous functions*

$$\psi_1, \psi_2 : L_p(X, \mu)_{\mathbb{R}} \to \mathbb{R}$$

such that ψ_1 and ψ_2 are (real-)linear and

$$\Psi(f) = \psi_1(\mathrm{Re}(f)) + i\psi_1(\mathrm{Im}(f)) + i\psi_2(\mathrm{Re}(f)) - \psi_2(\mathrm{Im}(f))$$

for all $f \in L_p(X, \mu)$.

Proof. Given a function $f \in L_p(X, \mu)$, recall the complex conjugate of f, denoted $\overline{f}$, is an element of $L_p(X, \mu)$. Define $\psi_1, \psi_2 : L_p(X, \mu)_{\mathbb{R}} \to \mathbb{R}$ by

$$\psi_1(f) = \mathrm{Re}(\Psi(f)) \qquad \text{and} \qquad \psi_2(f) = \mathrm{Im}(\Psi(f))$$

for all $f \in L_p(X, \mu)_{\mathbb{R}}$. Since Ψ is complex linear and continuous, it is elementary to see that ψ_1 and ψ_2 are real linear and continuous. Moreover, the equation

$$\Psi(f) = \psi_1(\mathrm{Re}(f)) + i\psi_1(\mathrm{Im}(f)) + i\psi_2(\mathrm{Re}(f)) - \psi_2(\mathrm{Im}(f))$$

for all $f \in L_p(X, \mu)$ is then trivial to verify. ∎

Next we require a method for verifying that a function is in $L_q(X, \mu)$ based on knowledge of its integral against elements of $L_p(X, \mu)$. This is achieved via the following two lemma (one for $p \in (1, \infty)$ and one for $p = 1$). Note this has significance outside the proof of the Riesz Representation Theorem (Theorem D.4.3) as it enables us to deduce a function is in $L_q(X, \mu)$ and obtain a bound on its norm based on integration.

Lemma D.4.5. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space, let $1 < p < \infty$, and $1 < q < \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, and let $g \in L_1(X, \mu)_{\mathbb{R}}$. If there exists an $M \in \mathbb{R}$ such that*

$$\left| \int g\varphi \, d\mu \right| \leq M \left\| \varphi \right\|_p$$

for all measurable functions $\varphi : X \to \mathbb{R}$ of finite range, then $g \in L_q(X, \mu)$ with $\left\| g \right\|_q \leq M$.

Proof. Since $|g|^q$ is a measurable function, there exists an increasing sequence $(\varphi_n)_{n\geq 1}$ of simple functions that converges to $|g|^q$ pointwise. For each $n \in \mathbb{N}$ let

$$\psi_n = \varphi_n^{\frac{1}{p}} \mathrm{sgn}(g).$$

Since $\mathrm{sgn}(g)$ obtains a finite number of values as g is real-valued, it is elementary to see that each ψ_n is a measurable function of finite range. Moreover, notice for all $n \in \mathbb{N}$ that

$$\left\| \psi_n \right\|_p = \left(\int_X |\psi_n|^p \, d\mu \right)^{\frac{1}{p}} = \left(\int_X \varphi_n \, d\mu \right)^{\frac{1}{p}}$$

and

$$0 \leq \varphi_n = \varphi_n^{\frac{1}{p}} \varphi_n^{\frac{1}{q}} \leq \varphi_n^{\frac{1}{p}} |g| = \varphi_n^{\frac{1}{p}} \mathrm{sgn}(g) g = \psi_n g.$$

Therefore, for all $n \in \mathbb{N}$

$$0 \leq \int_X \varphi_n \, d\mu \leq \int_X g\psi_n \, d\mu \leq M \left\| \psi_n \right\|_p = M \left(\int_X \varphi_n \, d\mu \right)^{\frac{1}{p}}.$$

Since all simple functions are integrable as μ is finite, we know that

$$\int_X \varphi_n \, d\mu < \infty$$

for all $n \in \mathbb{N}$. Hence the above equation implies that

$$\left(\int_X \varphi_n \, d\mu \right)^{\frac{1}{q}} = \left(\int_X \varphi_n \, d\mu \right)^{1-\frac{1}{p}} \leq M.$$

However, by the Monotone Convergence Theorem

$$\lim_{n\to\infty} \int_X \varphi_n \, d\mu = \int_X |g|^q \, d\mu$$

and thus $\left\| g \right\|_q \leq M$. Hence $g \in L_q(X, \mu)$ as desired. ∎

Lemma D.4.6. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space, and let $g \in L_1(X, \mu)_{\mathbb{R}}$. If there exists an $M \in \mathbb{R}$ such that*

$$\left| \int g\varphi \, d\mu \right| \leq M \left\| \varphi \right\|_1$$

for all measurable functions $\varphi : X \to \mathbb{R}$ of finite range, then $g \in L_\infty(X, \mu)$ with $\left\| g \right\|_\infty \leq M$.

Proof. Let $\epsilon > 0$ be arbitrary. Consider the set

$$A_\epsilon = \{x \in X \mid |g(x)| \geq M + \epsilon\}.$$

Clearly A_ϵ is measurable. Hence

$$\begin{aligned}(M+\epsilon)\mu(A_\epsilon) &\leq \int_{A_\epsilon} |g| \, d\mu \\ &= \int_X \operatorname{sgn}(g)\chi_{A_\epsilon} g \, d\mu \\ &\leq M \left\|\operatorname{sgn}(g)\chi_{A_\epsilon}\right\|_1 \\ &= M\mu(A_\epsilon)\end{aligned}$$

since $\operatorname{sgn}(g)\chi_{A_\epsilon}$ is a measurable function of finite range (as g is real-valued). Therefore $\epsilon\mu(A_\epsilon) \leq 0$ so $\mu(A_\epsilon) = 0$. Therefore, as $\epsilon > 0$ was arbitrary, we obtain that $g \in L_\infty(X, \mu)$ with $\|g\|_\infty \leq M$. ∎

Proof of the Riesz Representation Theorem for L_p-Spaces (Theorem D.4.3). Recall from Example D.4.1 that if $\Psi_g : L_p(X, \mu) \to \mathbb{C}$ is defined by

$$\Psi_g(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$, then $\Psi_g \in L_p(X, \mu)^*$ and $\|\Psi_g\| = \|g\|_q$. Furthermore, notice if $g_1, g_2 \in L_q(X, \mu)$ are such that $\Psi_{g_1} = \Psi_{g_2}$, then

$$0 = \Psi_{g_1}(f) - \Psi_{g_2}(f) = \int_X fg_1 \, d\mu - \int_X fg_2 \, d\mu = \int_X f(g_1 - g_2) \, d\mu = \Psi_{g_1 - g_2}(f)$$

for all $f \in L_p(X, \mu)$. Therefore $0 = \|\Psi_{g_1 - g_2}\| = \|g_1 - g_2\|_q$ so $g_1 = g_2$. Hence, to complete the proof, it suffices to show that if $\Psi \in L_p(X, \mu)^*$ then there exists a $g \in L_q(X, \mu)$ such that $\Psi = \Psi_g$ (as the above produces the value of the norm and uniqueness).

Fix $\Psi \in L_p(X, \mu)^*$. Recall by Lemma D.4.4 that there exists continuous real-linear functions $\psi_1, \psi_2 : L_p(X, \mu)_\mathbb{R} \to \mathbb{R}$ such that

$$\Psi(f) = \psi_1(\operatorname{Re}(f)) + i\psi_1(\operatorname{Im}(f)) + i\psi_2(\operatorname{Re}(f)) - \psi_2(\operatorname{Im}(f))$$

for all $f \in L_p(X, \mu)$. If we demonstrate that there exists $g_1, g_2 \in L_q(X, \mu)_\mathbb{R}$ such that

$$\psi_1(h) = \int_X hg_1 \, d\mu \qquad \text{and} \qquad \psi_2(h) = \int_X hg_2 \, d\mu$$

for all $h \in L_p(X, \mu)_\mathbb{R}$, then we obtain (using complex linearity) that

$$\Psi(f) = \int_X f(g_1 + ig_2) \, d\mu$$

for all $f \in L_p(X, \mu)$, which would complete the proof as $g_1 + ig_2 \in L_q(X, \mu)$. Therefore, it suffices to show that if $\psi : L_p(X, \mu)_{\mathbb{R}} \to \mathbb{R}$ is continuous and real-linear then there exists a $g \in L_q(X, \mu)_{\mathbb{R}}$ such that

$$\psi(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$.

To see the above claim, we will divide the proof into two cases.

Case 1: μ is finite. Since μ is finite, $\chi_A \in L_p(X, \mu)$ for all $A \in \mathcal{A}$. Hence define $\nu : \mathcal{A} \to \mathbb{R}$ by

$$\nu(A) = \psi(\chi_A)$$

for all $A \in \mathcal{A}$. We claim that ν is a finite signed measure that is absolutely continuous with respect to μ. To see this, first notice that

$$\nu(\emptyset) = \psi(\chi_\emptyset) = \psi(0) = 0$$

as ψ is linear. Moreover, clearly ν does not obtain the values $\pm\infty$ by definition.

To see that ν is countably additive, let $\{A_k\}_{k=1}^\infty \subseteq \mathcal{A}$ be pairwise disjoint and let $A = \bigcup_{k=1}^\infty A_k$. Since μ is a finite measure,

$$\mu(A) = \sum_{k=1}^\infty \mu(A_k) < \infty.$$

Hence

$$\lim_{n\to\infty} \sum_{k=n}^\infty \mu(A_k) = 0.$$

Therefore

$$\lim_{n\to\infty} \left\| \chi_A - \sum_{k=1}^n \chi_{A_k} \right\|_p = \lim_{n\to\infty} \left(\sum_{k=n}^\infty \mu(A_k) \right)^{\frac{1}{p}} = 0.$$

Hence $\chi_A = \sum_{k=1}^\infty \chi_{A_k}$ as a sum of vectors in $L_p(X, \mu)$. Therefore, since ψ is a continuous linear functional, we obtain that

$$\nu(A) = \psi(\chi_A) = \sum_{k=1}^\infty \psi(\chi_{A_k}) = \sum_{k=1}^\infty \nu(A_k).$$

Thus ν is countably additive. However, to show that ν is a signed measure, it is necessary to show that the sum converges absolutely. For each $n \in \mathbb{N}$ let $c_n = \operatorname{sgn}(\nu(A_n))$ and let $f_n = \sum_{k=1}^n c_k \chi_{A_k}$. Then for all $n, m \in \mathbb{N}$ with

$n \geq m$

$$\begin{aligned}\|f_n - f_m\|_p &= \left\| \sum_{k=m+1}^{n} c_k \chi_{A_k} \right\|_p \\ &= \left(\sum_{k=m+1}^{n} \mu(A_k) \right)^{\frac{1}{p}} \\ &\leq \left(\sum_{k=m}^{\infty} \mu(A_k) \right)^{\frac{1}{p}}.\end{aligned}$$

Therefore, since $\lim_{m\to\infty} \left(\sum_{k=m}^{\infty} \mu(A_k)\right)^{\frac{1}{p}} = 0$, $(f_n)_{n\geq 1}$ is Cauchy in $L_p(X, \mu)$. Since $L_p(X, \mu)$ is complete by the Riesz-Fisher Theorem (Theorems D.2.1 and D.2.4), there exists an $f \in L_p(X, \mu)$ such that $f = \lim_{n\to\infty} f_n$ in $L_p(X, \mu)$. Therefore, since ψ is a continuous linear functional, we obtain that

$$\psi(f) = \lim_{n\to\infty} \psi(f_n) = \lim_{n\to\infty} \sum_{k=1}^{n} |\nu(A_n)|.$$

Therefore, as $\psi(f) \in \mathbb{R}$, we see that the sum converges absolutely.

To see that ν is finite, notice for all $A \in \mathcal{A}$ that

$$|\nu(A)| = |\psi(\chi_A)| \leq \|\psi\| \, \|\chi_A\|_p = \|\psi\| \, \mu(A)^{\frac{1}{p}} < \infty$$

as μ is finite. Hence ν is finite. Finally, to see that ν is absolutely continuous with respect to μ, notice if $A \in \mathcal{A}$ is such that $\mu(A) = 0$, then $\chi_A = 0$ as an element of $L_p(X, \mu)$ and thus

$$\nu(A) = \psi(\chi_A) = \psi(0) = 0$$

as ψ is linear. Hence ν is a finite signed measure that is absolutely continuous with respect to μ.

By the Radon-Nikodym Theorem for signed measures there exists a real-valued function $g \in L_1(X, \mu)$ such that

$$\psi(\chi_A) = \nu(A) = \int_A g \, d\mu = \int_X g\chi_A \, d\mu$$

for all $A \in \mathcal{A}$. Using the linearity of the integral and of ψ, we obtain for any measurable function $\varphi : X \to \mathbb{R}$ with finite range that

$$\psi(\varphi) = \int_X \varphi g \, d\mu.$$

However, this implies that

$$\left| \int_X \varphi g \, d\mu \right| = |\psi(\varphi)| \leq \|\psi\| \, \|\varphi\|_p$$

for all measurable functions $\varphi : X \to \mathbb{R}$ with finite range. Hence Lemma D.4.5 or Lemma D.4.6 implies that $g \in L_q(X,\mu)_{\mathbb{R}}$.

Since

$$\psi(\varphi) = \int_X \varphi g \, d\mu$$

for all simple functions in $L_p(X,\mu)$, we obtain by linearity that

$$\psi(\varphi) = \int_X \varphi g \, d\mu$$

for all φ which are linear combinations of simple functions in $L_p(X,\mu)$. Therefore Theorem D.3.1 (along with continuity) implies that

$$\psi(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X,\mu)$ as desired.

Case 2: μ is σ-finite. Recall there exists $\{X_n\}_{n=1}^\infty \subseteq \mathcal{A}$ such that $X = \bigcup_{n=1}^\infty X_n$, $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$, and $X_n \subseteq X_{n+1}$ for all $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$, let

$$\mathcal{A}_n = \{A \cap X_n \mid A \in \mathcal{A}\}$$

and let $\mu_n = \mu|_{\mathcal{A}_n}$. It is elementary to verify that $\mathcal{A}_n$ is a σ-algebra on X_n and that μ_n is a measure on $(X_n, \mathcal{A}_n)$. Notice if $f \in L_p(X_n,\mu_n)$, we can view f as an element of $L_p(X,\mu)$ by extending f to be zero on X_n^c. Hence, for each $n \in \mathbb{N}$, we can define $\psi_n : L_p(X_n,\mu_n) \to \mathbb{R}$ by

$$\psi_n(f) = \psi(f)$$

for all $f \in L_p(X_n,\mu_n) \subseteq L_p(X,\mu)$. It is elementary to verify that ψ_n is a continuous linear functional on $L_p(X_n,\mu_n)$ with norm at most $\|\psi\|$ as the norms on $L_p(X_n,\mu_n)$ and $L_p(X,\mu)$ agree on elements of $L_p(X_n,\mu_n)$.

Since $(X_n, \mathcal{A}_n, \mu_n)$ is a finite measure space, the first case of this proof implies there exists a unique function $g_n \in L_q(X_n,\mu_n)$ such that

$$\int_{X_n} fg_n \, d\mu_n = \psi_n(f) = \psi(f)$$

for all $f \in L_p(X_n,\mu_n)$. Moreover $\|g_n\|_{L_q(X_n,\mu_n)} = \|\psi_n\| \leq \|\psi\|$.

Extend each g_n to be zero on X_n^c. Hence $g_n \in L_q(X,\mu)$ for all $n \in \mathbb{N}$, $\|g_n\|_{L_q(X_n,\mu_n)} = \|g_n\|_q$, and

$$\psi_n(f) = \int_X fg_n \, d\mu$$

for all $f \in L_p(X_n,\mu_n)$. Moreover, notice for all $n \in \mathbb{N}$ and $f \in L_p(X_n,\mu_n) \subseteq L_p(X_{n+1},\mu_{n+1})$ that

$$\int_X fg_{n+1} \, d\mu = \psi_{n+1}(f) = \psi_n(f) = \int_X fg_n \, d\mu.$$

Therefore, due to the uniqueness of g_n, we obtain that $g_{n+1}|_{X_n} = g_n$.

Define $g : X \to \mathbb{R}$ by $g(x) = g_n(x)$ whenever $x \in X_n$. As $g_{n+1}|_{X_n} = g_n$ and as $X = \bigcup_{n=1}^{\infty} X_n$, g is well-defined up to a set of measure zero and defines a measurable function (as it is the pointwise limit of $(g_n)_{n\geq 1})$. If $q = \infty$ then as $\|g_n\|_\infty \leq \|\psi\|$ for all $n \in \mathbb{N}$, we easily see that $\|g\|_\infty \leq \|\psi\| < \infty$ and thus $g \in L_\infty(X, \mu)$. Otherwise, if $q \neq \infty$, notice that as $|g_n| \leq |g_{n+1}|$ for all $n \in \mathbb{N}$ and as $(g_n)_{n\geq 1}$ converges to g pointwise almost everywhere, the Monotone Convergence Theorem implies that

$$\left(\int_X |g|^q \, d\mu\right)^{\frac{1}{q}} = \lim_{n\to\infty} \left(\int_X |g_n|^q \, d\mu\right)^{\frac{1}{q}} \leq \|\psi\| < \infty.$$

Hence $g \in L_q(X, \mu)$.

Finally, to see that

$$\psi(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$, let $f \in L_p(X, \mu)$ be arbitrary and for each $n \in \mathbb{N}$ let $f_n = f\chi_{X_n}$. Then

$$|f_n - f|^p \leq |f|^p$$

and $(|f_n - f|^p)_{n\geq 1}$ converges to zero almost everywhere. Therefore, since $|f|^p \in L_1(X, \mu)$, the Dominated Convergence Theorem implies that $\lim_{n\to\infty} \|f - f_n\|_p = 0$. As ψ is continuous

$$\psi(f) = \lim_{n\to\infty} \psi(f_n) = \lim_{n\to\infty} \psi_n(f_n) = \lim_{n\to\infty} \int_X f_n g_n \, d\mu = \lim_{n\to\infty} \int_X f_n g \, d\mu$$

since $f_n g_n = f_n g$ for all $n \in \mathbb{N}$. However, since $(f_n g)_{n\geq 1}$ converges pointwise to fg and since $|f_n g| \leq |fg| \in L_1(X, \mu)$ by Hölder's Inequality (Theorems D.1.7 and D.1.7), the Dominated Convergence Theorem implies that

$$\psi(f) = \lim_{n\to\infty} \int_X f_n g \, d\mu = \int_X fg \, d\mu$$

as desired. ∎

Using only the Riesz Representation Theorem (Theorem D.4.3)) it is possible to verify that a function is in $L_q(X, \mu)$ via only integrals against $L_p(X, \mu)$ functions.

Corollary D.4.7. *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let $p, q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and $q \neq 1$. If*

$$\sup\left\{ \left| \int_X fg \, d\mu \right| \,\middle|\, f \in L_p(X, \mu), \|f\|_p \leq 1 \right\} < \infty,$$

then $g \in L_q(X, \mu)$.

Proof. Define $\Psi : L_p(X, \mu) \to \mathbb{R}$ by

$$\Psi(f) = \int_X fg \, d\mu$$

for all $f \in L_p(X, \mu)$. By the assumptions in the statement, we easily see that Ψ is a well-defined continuous linear functional on $L_p(X, \mu)$. Therefore, by the Riesz Representation Theorem there exists an unique function $h \in L_q(X, \mu)$ such that

$$\Psi(f) = \int_X fh \, d\mu$$

for all $f \in L_p(X, \mu)$. In particular, for all $f \in L_p(X, \mu)$ and $A \in \mathcal{A}$ we see that

$$\int_A fg \, d\mu = \int_X (f\chi_A) g \, d\mu = \Psi(f\chi_A) = \int_X (f\chi_A) h \, d\mu = \int_A fh \, d\mu.$$

Therefore, as μ is σ-finite, by the Radon-Nikodym Theorem we obtain that $fg = fh$ for all $f \in L_p(X, \mu)$.

Since μ is σ-finite, there exists $\{X_n\}_{n=1}^\infty \subseteq \mathcal{A}$ such that $X = \bigcup_{n=1}^\infty X_n$, $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$, and $\{X_n\}_{n=1}^\infty$ are pairwise disjoint. Since $\mu(X_n) < \infty$, $\chi_{X_n} \in L_p(X, \mu)$ for all $n \in \mathbb{N}$. Hence the above implies that

$$g\chi_{X_n} = h\chi_{X_n}$$

for all $n \in \mathbb{N}$. Therefore, as $X = \bigcup_{n=1}^\infty X_n$, we obtain that $g = h \in L_q(X, \mu)$ as desired. ∎

Of course, there are many other other versions of the Riesz Representation Theorem we could analyze in the context of measure theory. Here are two which describe the dual spaces of two very natural collections of functions seen in this course.

Theorem D.4.8 (Riesz Representation Theorem for L_∞). *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. If $\Psi \in L_\infty(X, \mu)_{\mathbb{R}} \to \mathbb{R}$ is a continuous linear functional, then there exists a unique 'bounded, finitely additive' signed measure ν such that ν is absolutely continuous with respect to μ and*

$$\Psi(f) = \int_X f \, d\nu$$

for all $f \in L_\infty(X, \mu)$. Moreover $\|\Psi\| = |\nu|(X)$.

Theorem D.4.9 (Riesz-Markov Theorem). *Let X be a locally compact Hausdorff space and let $\mathcal{M}(X)$ denote the space of all $\mathbb{K}$-valued, finite, regular, Borel measures on X equipped with the total variation norm.*

If $\mu \in \mathcal{M}(X)$, define $T_\mu : C_0(X, \mathbb{K}) \to \mathbb{K}$ by

$$T_\mu(f) = \int_X f(x) \, d\mu$$

for all $f \in C_0(X, \mathbb{K})$*. Then* $T_\mu \in C_0(X, \mathbb{K})^*$ *for all* $\mu \in \mathcal{M}(X)$*. Moreover, if* $\Theta : \mathcal{M}(X) \to C_0(X, \mathbb{K})^*$ *is defined by*

$$\Theta(\mu) = T_\mu$$

for all $\mu \in \mathcal{M}(X)$*, then* Θ *is an isometric isomorphism.*

Index