An Introduction to Functional Analysis

Laurent W. Marcoux

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario   Canada   N2L 3G1

November 15, 2022

Preface to the Fifth Edition - November 15, 2022

Once again, I would like to thank those who caught typos in the notes, including Patrick Au, Ting Wei Liu, Arsalan Motamedi, and Shlok Ashok Nahar.

Preface to the Fourth Edition - December 15, 2018

Thanks to Adina Goldberg, Hayley Reid, Wanchun Shen, Erlang Surya, Wentao Yang and Zhenyuan Zhang for catching those typos and mistakes that I thought I had hidden really well, and for suggesting alternate proofs to some results presented in the third edition!

I've begun to enlarge the "Appendix" sections for each Chapter, and to add exercises at the end of each Chapter. The "Appendix" sections are meant to be "cultural". Students are not required to read these to understand the remainder of the text, and material appearing in the Appendix sections will not appear on the exam. The exercises are meant to be much easier than the assignment exercises, and are there to help solidify some of the definitions and concepts appearing in the notes.

The biggest change, mathematically speaking, in the notes is that I have now provided a proof of the fact that the only compact, quasinilpotent, normal operator acting on an infinite-dimensional, separable, complex Hilbert space is the zero operator. (Yes, I am aware of the fact that this holds even if N is not compact.) The usual proof of this result uses Beurling's *Spectral Radius formula*, and it took me a few hours to come up with a proof that avoided that result in the setting I needed to consider. I have nothing against Beurling nor his Spectral Radius Formula in general (in fact, you might say that I appreciate it even more now), except that I didn't have the time to cover it, and therefore I had to find a "work-around" that did not require any functional calculus.

Preface to the Third Edition - October 24, 2014

This set of notes is now undergoing its third iteration. The mathematical content outside of the appendices is mostly stabilized, and now begins the long and lonely hunt for typos, poor grammar, and awkward sentence constructions.

Please feel free to contact me if you find any mistakes – mathematical or otherwise – in these notes.

Preface to the Second Edition - December 1, 2010

This set of notes has now undergone its second incarnation. I have corrected as many typos as I have found so far, and in future instalments I will continue to add comments and to modify the appendices where appropriate. The course number for the Functional Analysis course at Waterloo has now changed to PMath 753, in case anyone is checking.

The comment in the preface to the "first edition" regarding caution and buzz saws is still *à propos*. Nevertheless, I maintain that this set of notes is worth at least twice the price[1] that I'm charging for them.

For the sake of reference: excluding the material in the appendices, and allowing for the students to study the last section on topology themselves, one should be able to cover the material in these notes in one term, which at Waterloo consists of 36 fifty-minute lectures.

My thanks to Xiao Jiang and Ian Hincks for catching a number of typos that I missed in the second revision.

Preface to the First Edition - December 1, 2008

The following is a set of class notes for the PMath 453/653 course I taught at the University of Waterloo in 2008. As mentioned on the front page, they are a work in progress, and - this being the "first edition" - they are replete with typos. A student should approach these notes with the same caution he or she would approach buzz

[1] If you were charged a single penny for the electronic version of these notes, you were robbed. You can get them for free from my website.

saws; they can be very useful, but you should be thinking the whole time you have them in your hands. Enjoy.

I would like to thank Paul Skoufranis for having pointed out to me an embarrassing number of typos. I am glad to report that he still has both hands and all of his fingers.

THE REVIEWS ARE IN!

From the moment I picked your book up until I laid it down I was convulsed with laughter. Someday I intend reading it.

Groucho Marx

This is not a novel to be tossed aside lightly. It should be thrown with great force.

Dorothy Parker

The covers of this book are too far apart.

Ambrose Bierce

I read part of it all the way through.

Samuel Goldwyn

Reading this book is like waiting for the first shoe to drop.

Ralph Novak

Thank you for sending me a copy of your book. I'll waste no time reading it.

Moses Hadas

Sometimes you just have to stop writing. Even before you begin.

Stanislaw J. Lec

Contents

1. Normed Linear Spaces

I don't like country music, but I don't mean to denigrate those who do.
And for the people who like country music, denigrate means 'put down'.

Bob Newhart

1.1. It is expected that the student of this course will have already seen the notions of a **normed linear space** and of a **Banach space**. We shall review the definitions of these spaces, as well as some of their fundamental properties. In both cases, the underlying structure is that of a vector space. For our purposes, these vector spaces will be over the field $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

1.2. Definition. *Let $\mathfrak{X}$ be a vector space over $\mathbb{K}$. A **seminorm** on $\mathfrak{X}$ is a map*

$$\nu : \mathfrak{X} \to \mathbb{R}$$

satisfying

(i) $\nu(x) \geq 0$ *for all* $x \in \mathfrak{X}$*;*
(ii) $\nu(\lambda x) = |\lambda|\ \nu(x)$ *for all* $x \in \mathfrak{X}$, $\lambda \in \mathbb{K}$*; and*
(iii) $\nu(x + y) \leq \nu(x) + \nu(y)$ *for all* $x, y \in \mathfrak{X}$.

If ν satisfies the extra condition:

(iv) $\nu(x) = 0$ *if and only if* $x = 0$,

*then we say that ν is a **norm**, and we usually denote $\nu(\cdot)$ by $\|\cdot\|$. In this case, we say that $(\mathfrak{X}, \|\cdot\|)$ (or, with a mild abuse of nomenclature, $\mathfrak{X}$) is a **normed linear space**.*

1.3. A norm on $\mathfrak{X}$ is a generalisation of the absolute value function on $\mathbb{K}$. Of course, equipped with the absolute value function on $\mathbb{K}$, one immediately defines a metric $d : \mathbb{K} \times \mathbb{K} \to \mathbb{R}$ by setting $d(x, y) = |x - y|$.

In exactly the same way, the norm $\|\cdot\|$ on a normed linear space $\mathfrak{X}$ induces a metric

$$\begin{array}{rccc} d: & \mathfrak{X} \times \mathfrak{X} & \to & \mathbb{R} \\ & (x, y) & \mapsto & \|x - y\|. \end{array}$$

The **norm topology** on $(\mathfrak{X}, \|\cdot\|)$ is the topology induced by this metric. For each $x \in \mathfrak{X}$, a neighbourhood base for this topology is given by

$$\mathcal{B}_x = \{V_\varepsilon(x) : \varepsilon > 0\},$$

where $V_\varepsilon(x) = \{y \in \mathfrak{X} : d(y, x) < \varepsilon\} = \{y \in \mathfrak{X} : \|y - x\| < \varepsilon\}$. We say that the normed linear space $(\mathfrak{X}, \|\cdot\|)$ (or informally $\mathfrak{X}$) is **complete** if the corresponding metric space $(\mathfrak{X}, d)$ is complete.

1.4. Example. Define

$$c_{00}^{\mathbb{K}}(\mathbb{N}) = \{(x_n)_{n=1}^{\infty} : x_n \in \mathbb{K}, n \geq 1, x_n = 0 \text{ for all but finitely many } n \geq 1\}.$$

For $x = (x_n)_n \in c_{00}^{\mathbb{K}}(\mathbb{N})$, set $\|x\|_\infty = \sup_{n \geq 1} |x_n|$. Then $(c_{00}^{\mathbb{K}}(\mathbb{N}), \|\cdot\|_\infty)$ is a normed linear space. It is not, however, complete.

The space

$$c_0^{\mathbb{K}}(\mathbb{N}) = \{(x_n)_{n=1}^{\infty} : x_n \in \mathbb{K}, n \geq 1, \lim_{n\to\infty} x_n = 0\},$$

equipped with the same norm $\|x\|_\infty = \sup_{n\geq 1} |x_n|$ *does* define a complete normed linear space.

1.5. Remark. We pause to make a comment about the terminology which we shall be using in these notes. A **vector subspace** of a vector space V over $\mathbb{K}$ is a non-empty subset W for which $x, y \in W$ and $k \in \mathbb{K}$ implies that $kx + y \in W$. When the vector space V does not carry a topology, there is no confusion in this terminology. When dealing with normed linear spaces $(\mathfrak{X}, \|\cdot\|)$, and more generally with the *topological vector spaces* $(\mathcal{V}, \mathcal{T})$ we shall deal with later in the text, and of which normed linear spaces are an example, one needs to distinguish between those vector subspaces which are definitely closed sets in the underlying topology from those which may or may not be closed. For this reason, we shall refer to vector subspaces of a topological vector space $(\mathcal{V}, \mathcal{T})$ which may or may not be closed as **linear manifolds** in $\mathcal{V}$, whereas **subspaces** will be used to denote *closed* linear manifolds. As a pedagogical tool, we shall also refer to these as **closed subspaces**, although strictly speaking, in our language, this is redundant.

Thus $c_{00}^{\mathbb{K}}(\mathbb{N})$ is a linear manifold in $c_0^{\mathbb{K}}(\mathbb{N})$ under the norm $\|\cdot\|_\infty$, but it is not a subspace of $c_0^{\mathbb{K}}(\mathbb{N})$, because it is not closed. In fact, it is dense in $c_0^{\mathbb{K}}(\mathbb{N})$.

1.6. Example. Consider

$$\mathcal{P}_{\mathbb{K}}([0,1]) = \{p = p_0 + p_1 z + p_2 z^2 + \cdots + p_n z^n : n \geq 1, p_i \in \mathbb{K}, 0 \leq i \leq n\}.$$

Then

$$\|p\|_\infty = \sup\{|p(z)| : z \in [0,1]\}$$

defines a norm on $\mathcal{P}_{\mathbb{K}}([0,1])$. The Stone-Weierstraß Theorem states that $\mathcal{P}_{\mathbb{K}}([0,1])$ is a dense linear manifold in the normed linear space $\mathcal{C}([0,1], \mathbb{K})$ of continuous, $\mathbb{K}$-valued functions on $[0,1]$ with the supremum norm.

If we select $x_0 \in [0,1]$ arbitrarily, then it is straightforward to check that $\nu(f) := |f(x_0)|$ defines a seminorm on $\mathcal{P}_{\mathbb{K}}([0,1])$ which is not a norm.

1.7. Example. Let $n \geq 1$ be an integer. If $1 \leq p < \infty$ is a real number, then

$$\|(x_1, x_2, ..., x_n)\|_p = \Big(\sum_{k=1}^{n} |x_k|^p\Big)^{1/p}$$

defines a norm on $\mathbb{K}^n$, called the **p-norm**. We often write ℓ_n^p for $(\mathbb{K}^n, \|\cdot\|_p)$, when the underlying field $\mathbb{K}$ is understood. We may also define

$$\|(x_1, x_2, ..., x_n)\|_\infty = \max(|x_1|, |x_2|, ..., |x_n|).$$

Observe that $(\mathbb{K}^n, \|\cdot\|_\infty)$ is a normed linear space. We abbreviate this to ℓ_n^∞ when $\mathbb{K}$ is understood.

1.8. Example. For $1 \le p < \infty$, we define

$$\ell_{\mathbb{K}}^p(\mathbb{N}) = \{(x_n)_{n=1}^\infty : x_n \in \mathbb{K}, n \ge 1 \text{ and } \sum_{n=1}^\infty |x_n|^p < \infty\}.$$

For $(x_n)_{n=1}^\infty \in \ell_{\mathbb{K}}^p(\mathbb{N})$, we set

$$\|(x_n)_n\|_p = \big(\sum_{n=1}^\infty |x_n|^p\big)^{1/p}.$$

Then $\|\cdot\|_p$ defines a norm, again called the ***p*-norm**. on $\ell_{\mathbb{K}}^p(\mathbb{N})$.

As above, we may also define

$$\ell_{\mathbb{K}}^\infty(\mathbb{N}) = \{(x_n)_{n=1}^\infty : x_n \in \mathbb{K}, n \ge 1, \ \sup_n |x_n| < \infty\}.$$

The **∞-norm** on $\ell_{\mathbb{K}}^\infty(\mathbb{N})$ is given by

$$\|(x_n)_n\|_\infty = \sup_n |x_n|.$$

In most contexts, the underlying field $\mathbb{K}$ is understood, and we shall write only $\ell^p(\mathbb{N})$, or even ℓ^p, $1 \le p \le \infty$.

The last two examples have one especially nice property not shared by $c_{00}^{\mathbb{K}}(\mathbb{N})$ and $\mathcal{P}_{\mathbb{K}}([0,1])$, namely: they are complete.

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

1.10. Example. Let $\mathcal{C}([0,1], \mathbb{K}) = \{f : [0,1] \to \mathbb{K} : f \text{ is continuous}\}$, equipped with the **uniform norm**

$$\|f\|_\infty = \max\{|f(z)| : z \in [0,1]\}.$$

Then $(\mathcal{C}([0,1], \mathbb{K}), \|\cdot\|_\infty)$ is a Banach space.

1.11. Example. Let $\mathbb{D} \subseteq \mathbb{C}$ denote the open unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, and $\mathbb{T}$ denote the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. Consider the **disc algebra**

$$\mathcal{A}(\mathbb{D}) := \{f \in \mathcal{C}(\overline{\mathbb{D}}) : f|_{\mathbb{D}} \text{ is holomorphic}\}.$$

The function

$$\|p\|_\infty := \sup_{z \in \overline{\mathbb{D}}} |p(z)|$$

is easily seen to define a norm on $\mathcal{A}(\mathbb{D})$.

It follows from elementary Complex Analysis that the space $\mathbb{C}[z]$ of polynomials (with domain restricted to $\overline{\mathbb{D}}$) is dense in $\mathcal{A}(\mathbb{D})$, and that $\mathcal{A}(\mathbb{D})$ is complete with respect to this norm. That is, $\mathcal{A}(\mathbb{D})$ is a Banach space.

Furthermore, from the Maximum Modulus Principle, we see that the map

$$\begin{array}{rccc}\Gamma: & \mathcal{A}(\mathbb{D}) & \to & \mathcal{C}(\mathbb{T}) \\ & f & \mapsto & f|_{\mathbb{T}}\end{array}$$

is isometric, and so we can (and do) identify $\mathcal{A}(\mathbb{D})$ with the algebra

$$\{f \in \mathbb{C}(\mathbb{T}) : f \text{ extends to a holomorphic function on } \mathbb{D}\},$$

equipped with the norm $\|f\|_\infty = \sup_{|z|=1} |f(z)|$.

The perspicacious reader will have noticed that the product fg of two elements f, g of $\mathcal{A}(\mathbb{D})$ is again an element of $\mathcal{A}(\mathbb{D})$, and that $\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$. This means that $\mathcal{A}(\mathbb{D})$ forms a **Banach algebra**. These important examples of Banach spaces form a topic of study in their own right, and we shall not have much to say about them here.

1.12. Definition. *Let $\mathcal{H}$ be an inner product space over $\mathbb{K}$; that is, there exists a map*

$$\langle \cdot, \cdot \rangle \to \mathbb{K}$$

which, for all $x, x_1, x_2, y \in \mathcal{H}$ and $\lambda \in \mathbb{K}$, satisfies:

(i) $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle$;
(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$;
(iii) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$;
(iv) $\langle x, x\rangle \ge 0$, *with equality holding if and only if* $x = 0$.

(Of course, when $\mathbb{K} = \mathbb{R}$, the complex conjugation in (ii) is superfluous.) Recall that the canonical norm on $\mathcal{H}$ induced by the inner product is given by

$$\|x\| = \langle x, x\rangle^{1/2}.$$

If $\mathcal{H}$ is complete with respect to the corresponding metric, then we say that $\mathcal{H}$ is a ***Hilbert space****. Thus every Hilbert space is a Banach space.*

1.13. Example. Recall that $\ell^2_{\mathbb{K}}(\mathbb{N})$ is a Hilbert space with inner product

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

More generally, let (X, μ) be a measure space. Then $\mathcal{H} = L^2(X, \mu)$ is a Hilbert space with

$$\langle f, g\rangle = \int_X f \overline{g} d\mu.$$

1.14. It is easy to see that if $\mathfrak{X}$ is a normed linear space, then the vector space operations

$$\begin{array}{rccc} \sigma: & \mathfrak{X}\times\mathfrak{X} & \to & \mathfrak{X} \\ & (x,y) & \mapsto & x+y \end{array} \qquad \text{and} \qquad \begin{array}{rccc} \mu: & \mathbb{K}\times\mathfrak{X} & \to & \mathfrak{X} \\ & (\lambda,x) & \mapsto & \lambda x \end{array}$$

of addition and scalar multiplication are continuous (from the respective product topologies on $\mathfrak{X}\times\mathfrak{X}$ and on $\mathbb{K}\times\mathfrak{X}$ to the norm topology on $\mathfrak{X}$). The proof is left as an exercise for the reader. In particular, therefore, if $0 \neq \lambda \in \mathbb{K}$, $y \in \mathfrak{X}$, then $\sigma_y : \mathfrak{X} \to \mathfrak{X}$ defined by $\sigma_y(x) = x + y$ and $\mu_\lambda : \mathfrak{X} \to \mathfrak{X}$ defined by $\mu_\lambda(x) = \lambda x$ are homeomorphisms.

As a simple corollary to this fact, a set $G \subseteq \mathfrak{X}$ is open (resp. closed) if and only if $G + y$ is open (resp. closed) for all $y \in \mathfrak{X}$, and λG is open (resp. closed) for all $0 \neq \lambda \in \mathbb{K}$. We shall return to this in a later section.

1.15. New Banach spaces from old. We now exhibit a few constructions which allow us to produce new Banach spaces from simpler building blocks.

Let $(\mathfrak{X}_n, \|\cdot\|_n)_{n=1}^\infty$ denote a countable family of Banach spaces. Let $\mathfrak{X} = \prod_n \mathfrak{X}_n$.

(a) For each $1 \leq p < \infty$, define
$$\sum_{n=1}^\infty \oplus_p \mathfrak{X}_n = \{(x_n)_n \in \mathfrak{X} : \|(x_n)_n\|_p = \big(\sum_{n=1}^\infty \|x_n\|_n^p\big)^{1/p} < \infty\}.$$
Then $\sum_{n=1}^\infty \oplus_p \mathfrak{X}_n$ is a Banach space, referred to as the ℓ^p**-direct sum** of the $(\mathfrak{X}_n)_n$.

(b) With $p = \infty$,
$$\prod_n \mathfrak{X}_n = \{(x_n)_n \in \mathfrak{X} : \|(x_n)_n\|_\infty = \sup_{n\geq 1} \|x_n\|_n < \infty\}.$$
Again, $\prod_n \mathfrak{X}_n$ is a Banach space - namely the ℓ^∞**-direct product** of the $(\mathfrak{X}_n)_n$.

(c) We may also define
$$\sum_{n=1}^\infty \oplus_\infty \mathfrak{X}_n = \{(x_n)_n \in \prod_n \mathfrak{X}_n : \lim_n x_n = 0\},$$
equipped with the norm it inherits from $\prod_n \mathfrak{X}_n$. This is then a Banach space known as the ℓ^∞**-direct sum** of the $(\mathfrak{X}_n)_n$.

1.16. Definition. *Let $\mathfrak{X}$ be a vector space equipped with two norms $\|\cdot\|$ and $|||\cdot|||$. We say that these norms are* ***equivalent*** *if there exist constants $\kappa_1, \kappa_2 > 0$ so that*
$$\kappa_1\|x\| \leq |||x||| \leq \kappa_2\|x\| \text{ for all } x \in \mathfrak{X}.$$

We remark that when this is the case,
$$\frac{1}{\kappa_2}|||x||| \leq \|x\| \leq \frac{1}{\kappa_1}|||x|||,$$

resolving the apparent lack of symmetry in the definition of equivalence of norms.

1.17. Example. Fix $n \geq 1$ an integer, and let $\mathfrak{X} = \mathbb{C}^n$. For $x = (x_1, x_2, ..., x_n) \in \mathfrak{X}$,

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| \leq \sum_{k=1}^{n} \big(\max_j |x_j| \big) = \sum_{k=1}^{n} \|x\|_\infty = n\|x\|_\infty.$$

Moreover,

$$\|x\|_\infty = \max_j |x_j| \leq \sum_{k=1}^{n} |x_k| = \|x\|_1,$$

so that

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

This proves that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are equivalent norms on $\mathfrak{X}$. As we shall later see, all norms on a finite dimensional vector space are equivalent.

1.18. Example. Let $\mathfrak{X} = \mathcal{C}([0,1], \mathbb{C})$, and consider the norms

$$\|f\|_\infty = \sup\{|f(x)| : x \in [0,1]\}$$

and

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

on $\mathfrak{X}$. If, for each $n \geq 1$, we set f_n to be the function $f_n(x) = x^n$, then $\|f_n\|_\infty = 1$, while $\|f_n\|_1 = \int_0^1 x^n dx = \frac{1}{n+1}$. Clearly $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are inequivalent norms on $\mathfrak{X}$.

1.19. Proposition. *Two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space $\mathfrak{X}$ are equivalent if and only if they generate the same metric topologies.*

Proof. Suppose first that $\|\cdot\|$ and $|||\cdot|||$ are equivalent, say $\kappa_1\|x\| \leq |||x||| \leq \kappa_2\|x\|$ for all $x \in \mathfrak{X}$, where $\kappa_1, \kappa_2 > 0$ are constants. If $x \in \mathfrak{X}$ and $(x_n)_n$ is a sequence in $\mathfrak{X}$, then it immediately follows that

$$\lim_{n\to\infty} \|x_n - x\| = 0 \qquad \text{if and only if} \qquad \lim_{n\to\infty} |||x_n - x||| = 0.$$

That is, the two notions of convergence coincide, and thus the topologies are equal.

Conversely, suppose that the metric topologies $\tau_{\|\cdot\|}$ and $\tau_{|||\cdot|||}$, induced by $\|\cdot\|$ and $|||\cdot|||$ respectively, coincide. Then $G = \{x \in \mathfrak{X} : \|x\| < 1\}$ is an open nbhd of 0 in $(\mathfrak{X}, |||\cdot|||)$, and so there exists $\delta > 0$ so that $H = \{x \in \mathfrak{X} : |||x||| < \delta\} \subseteq G$. That is, $|||x||| < \delta$ implies $\|x\| < 1$. In particular, therefore, $|||x||| \leq \delta/2$ implies $\|x\| \leq 1$, so that that $||y|| \leq (2/\delta)|||y|||$ for all $y \in \mathfrak{X}$. By symmetry, there exists a constant $\kappa_2 > 0$ so that $|||y||| \leq \kappa_2\|y\|$ for all $y \in \mathfrak{X}$.

Thus $\|\cdot\|$ and $|||\cdot|||$ are equivalent norms.

□

1.20. Corollary. *Equivalence of norms is an equivalence relation for norms on a vector space $\mathfrak{X}$.*

1.21. Definition. *Let* $(\mathfrak{X}, \|\cdot\|)$ *be a normed linear space. A series* $\sum_{n=1}^\infty x_n$ *in* $\mathfrak{X}$ *is said to be* ***absolutely summable*** *if* $\sum_{n=1}^\infty \|x_n\| < \infty$.

The following result provides a very practical tool when trying to decide whether or not a given normed linear space is complete. We remark that the second half of the proof uses the standard fact that if $(y_n)_n$ is a Cauchy sequence in a metric space (Y, d), and if $(y_n)_n$ admits a convergent subsequence with limit y_0, then the original sequence $(y_n)_n$ converges to y_0 as well.

1.22. Proposition. *Let* $(\mathfrak{X}, \|\cdot\|)$ *be a normed linear space. The following statements are equivalent:*

(a) $\mathfrak{X}$ *is complete, and hence* $\mathfrak{X}$ *is a Banach space.*
(b) *Every absolutely summable series in* $\mathfrak{X}$ *is summable.*

Proof.

(a) implies (b): Suppose that $\mathfrak{X}$ is complete, and that $\sum x_n$ is absolutely summable. For each $k \ge 1$, let $y_k = \sum_{n=1}^k x_n$. Given $\varepsilon > 0$, we can find $N > 0$ so that $m \ge N$ implies $\sum_{n=m}^\infty \|x_n\| < \varepsilon$. If $k \ge m \ge N$, then

$$\begin{aligned} \|y_k - y_m\| &= \|\sum_{n=m+1}^{k} x_n\| \\ &\le \sum_{n=m+1}^{k} \|x_n\| \\ &\le \sum_{n=m+1}^{\infty} \|x_n\| \\ &< \varepsilon, \end{aligned}$$

so that $(y_k)_k$ is Cauchy in $\mathfrak{X}$. Since $\mathfrak{X}$ is complete, $y = \lim_{k\to\infty} y_k = \lim_{k\to\infty} \sum_{n=1}^k x_n = \sum_{n=1}^\infty x_n$ exists, i.e. $\sum_{n=1}^\infty x_n$ is summable.

(b) implies (a): Next suppose that every absolutely summable series in $\mathfrak{X}$ is summable, and let $(y_j)_j$ be a Cauchy sequence in $\mathfrak{X}$. For each $n \ge 1$ there exists $N_n > 0$ so that $k, m \ge N_n$ implies $\|y_k - y_m\| < 1/2^{n+1}$. Clearly, we may assume without loss of generality that $N_1 < N_2 < N_3 < \cdots$. Let $x_1 = y_{N_1}$ and for $n \ge 2$, let $x_n = y_{N_n} - y_{N_{n-1}}$. Then $\|x_n\| < 1/2^n$ for all $n \ge 2$, so that

$$\begin{aligned} \sum_{n=1}^\infty \|x_n\| &\le \|x_1\| + \sum_{n=2}^\infty \frac{1}{2^n} \\ &\le \|x_1\| + \frac{1}{2} < \infty. \end{aligned}$$

By hypothesis, $y = \sum_{n=1}^\infty x_n = \lim_{k\to\infty} \sum_{n=1}^k x_n$ exists. But $\sum_{n=1}^k x_n = y_{N_k}$, so that $\lim_{k\to\infty} y_{N_k} = y \in \mathfrak{X}$. Recalling that $(y_j)_j$ was Cauchy, we conclude from the remark preceding the Proposition that $(y_j)_j$

also converges to y. Since every Cauchy sequence in $\mathfrak{X}$ converges, $\mathfrak{X}$ is complete.

□

1.23. Theorem. *Let $(\mathfrak{X}, \|\cdot\|)$ be a normed linear space, and let $\mathfrak{M} \subseteq \mathfrak{X}$ be a linear manifold. Then*

$$p(x+\mathfrak{M}) := \inf\{\|x+m\| : m \in \mathfrak{M}\}$$

defines a seminorm on the quotient space $\mathfrak{X}/\mathfrak{M}$.

This formula defines a norm on $\mathfrak{X}/\mathfrak{M}$ if and only if $\mathfrak{M}$ is closed.

Proof. First observe that the function p is well-defined; for if $x+\mathfrak{M} = y+\mathfrak{M}$, then $x-y \in \mathfrak{M}$ and so

$$\begin{aligned} p(y+\mathfrak{M}) &= \inf\{\|y+m\| : m \in \mathfrak{M}\} \\ &= \inf\{\|y+m+(x-y)\| = \|x+m\| : m \in \mathfrak{M}\} \\ &= p(x+\mathfrak{M}). \end{aligned}$$

Clearly $p(x+\mathfrak{M}) \geq 0$ for all $x+\mathfrak{M} \in \mathfrak{X}/\mathfrak{M}$. If $0 \neq k \in \mathbb{K}$, then $m \in \mathfrak{M}$ if and only if $\frac{1}{k}m \in \mathfrak{M}$ and so

$$\begin{aligned} p(k(x+\mathfrak{M})) &= p(kx+\mathfrak{M}) \\ &= \inf\{\|kx+m\| : m \in \mathfrak{M}\} \\ &= \inf\{\|k(x+\frac{1}{k}m)\| : m \in \mathfrak{M}\} \\ &= |k| \ \inf\{\|x+m_0\| : m_0 \in \mathfrak{M}\} \\ &= |k| p(x+\mathfrak{M}). \end{aligned}$$

If $k = 0$, then $p(0+\mathfrak{M}) = 0$, since $m = 0 \in \mathfrak{M}$.

Finally,

$$\begin{aligned} p\big((x+\mathfrak{M})+(y+\mathfrak{M})\big) &= p(x+y+\mathfrak{M}) \\ &= \inf\{\|(x+y)+m\| : m \in \mathfrak{M}\} \\ &= \inf\{\|(x+m_1)+(y+m_2)\| : m_1, m_2 \in \mathfrak{M}\} \\ &\leq \inf\{\|x+m_1\| + \|y+m_2\| : m_1, m_2 \in \mathfrak{M}\} \\ &= p(x+\mathfrak{M}) + p(y+\mathfrak{M}). \end{aligned}$$

In the case where $\mathfrak{M}$ is closed in $\mathfrak{X}$, suppose that $p(x+\mathfrak{M}) = 0$ for some $x \in \mathfrak{X}$. Then

$$\inf\{\|x+m\| : m \in \mathfrak{M}\} = 0,$$

so there exist $m_n \in \mathfrak{M}$, $n \geq 1$, so that $-x = \lim_{n\to\infty} m_n$. Since $\mathfrak{M}$ is closed, $-x \in \mathfrak{M}$ and so $x+\mathfrak{M} = x+(-x)+\mathfrak{M} = 0+\mathfrak{M}$, proving that p is a norm.

The converse statement is left as an exercise.

□

1.24. Let $\mathfrak{X}$ be a normed linear space and $\mathfrak{M}$ be a linear manifold in $\mathfrak{X}$. We shall denote the canonical quotient map from $\mathfrak{X}$ to $\mathfrak{X}/\mathfrak{M}$ by q (or $q_{\mathfrak{M}}$ if the need to be specific arises). When $\mathfrak{M}$ is closed in $\mathfrak{X}$, we shall denote the norm from Theorem 1.23 once again by $\|\cdot\|$ (or $\|\cdot\|_{\mathfrak{X}/\mathfrak{M}}$), so that

$$\|q(x)\| = \|x + \mathfrak{M}\| = \inf\{\|x + m\| : m \in \mathfrak{M}\}.$$

It is clear that $\|q(x)\| \leq \|x\|$ for all $x \in \mathfrak{X}$, and so q is continuous. Indeed, given $\varepsilon > 0$, we can take $\delta = \varepsilon$ to get $\|x - y\| < \delta$ implies $\|q(x) - q(y)\| \leq \|x - y\| < \varepsilon$. We shall see below that q is also an *open* map - i.e. it takes open sets to open sets.

1.25. Theorem. *Let $\mathfrak{X}$ be a normed linear space and $\mathfrak{M}$ be a closed subspace of $\mathfrak{X}$.*

(a) *If $\mathfrak{X}$ is complete, then so are $\mathfrak{M}$ and $\mathfrak{X}/\mathfrak{M}$.*
(b) *If $\mathfrak{M}$ and $\mathfrak{X}/\mathfrak{M}$ are complete, then so is $\mathfrak{X}$.*

Proof.

(a) Suppose that $\mathfrak{X}$ is complete. We first show that $\mathfrak{M}$ is complete.

Let $(m_n)_{n=1}^\infty$ be a Cauchy sequence in $\mathfrak{M}$. Then it is Cauchy in $\mathfrak{X}$ and $\mathfrak{X}$ is complete, so that $x = \lim_{n\to\infty} m_n \in \mathfrak{X}$. Since $\mathfrak{M}$ is closed in $\mathfrak{X}$, $x \in \mathfrak{M}$. Thus $\mathfrak{M}$ is complete.

Note that this argument shows that any closed subset of a complete metric space is complete.

Next we show that $\mathfrak{X}/\mathfrak{M}$ is also complete.

Let $\sum_n q(x_n)$ be an absolutely summable series in $\mathfrak{X}/\mathfrak{M}$. For each $n \geq 1$, choose $m_n \in \mathfrak{M}$ so that $\|x_n + m_n\| \leq \|q(x_n)\| + \frac{1}{2^n}$. Then

$$\sum_n \|x_n + m_n\| \leq \sum_n \left(\|q(x_n)\| + \frac{1}{2^n}\right) < \infty,$$

so $\sum_n (x_n + m_n)$ is summable in $\mathfrak{X}$ since $\mathfrak{X}$ is complete. Set

$$x_0 := \sum_n (x_n + m_n).$$

By the continuity of q,

$$\begin{aligned} q(x_0) &= q(\sum_n (x_n + m_n)) \\ &= \sum_n q(x_n + m_n) \\ &= \sum_n q(x_n). \end{aligned}$$

Thus every absolutely summable series in $\mathfrak{X}/\mathfrak{M}$ is summable, and so by Proposition 1.22, $\mathfrak{X}/\mathfrak{M}$ is complete.

(b) Suppose next that $\mathfrak{M}$ and $\mathfrak{X}/\mathfrak{M}$ are both complete.

Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence in $\mathfrak{X}$. Then $(q(x_n))_{n=1}^\infty$ is Cauchy in $\mathfrak{X}/\mathfrak{M}$ and thus $q(y) = \lim_{n\to\infty} q(x_n)$ exists, by the completeness of $\mathfrak{X}/\mathfrak{M}$. For $n \ge 1$, choose $m_n \in \mathfrak{M}$ so that

$$\|y - (x_n + m_n)\| < \|q(y) - q(x_n)\| + \frac{1}{2^n}.$$

Since $(x_n + m_n)_{n=1}^\infty$ converges to y in $\mathfrak{X}$, it follows that it is a Cauchy sequence. Since both $(x_n)_{n=1}^\infty$ and $(x_n + m_n)_{n=1}^\infty$ are Cauchy, it follows that $(m_n)_{n=1}^\infty$ is also Cauchy – a fact that follows easily from the observation that

$$\|m_j - m_i\| \le \|(x_j + m_j) - (x_i + m_i)\| + \|x_j - x_i\|.$$

But $\mathfrak{M}$ is complete and so $m := \lim_{n\to\infty} m_n \in \mathfrak{M}$. This yields

$$y - m = \lim_{n\to\infty} (x_n + m_n) - m = \lim_{n\to\infty} x_n,$$

so that $(x_n)_{n=1}^\infty$ converges to $y - m$ in $\mathfrak{X}$. That is, $\mathfrak{X}$ is complete.

□

1.26. Proposition. *Let $\mathfrak{X}$ be a normed linear space and $\mathfrak{M}$ be a closed subspace of $\mathfrak{X}$. Let $q : \mathfrak{X} \to \mathfrak{X}/\mathfrak{M}$ denote the canonical quotient map.*

(a) *A subset $W \subseteq \mathfrak{X}/\mathfrak{M}$ is open if and only if $q^{-1}(W)$ is open in $\mathfrak{X}$.*

(b) *The map q is an* open map *- i.e., if $G \subseteq \mathfrak{X}$ is open, then $q(G)$ is open in $\mathfrak{X}/\mathfrak{M}$.*

Proof.

(a) If $W \subseteq \mathfrak{X}/\mathfrak{M}$ is open, then $q^{-1}(W)$ is open in $\mathfrak{X}$ because q is continuous.

Suppose next that $W \subseteq \mathfrak{X}/\mathfrak{M}$ and that $q^{-1}(W)$ is open in $\mathfrak{X}$. Let $q(x) \in W$. Then $x \in q^{-1}(W)$, and so we can find $\delta > 0$ so that $V_\delta(x) \subseteq q^{-1}(W)$. If $\|q(y) - q(x)\| < \delta$, then $\|y - x + m\| < \delta$ for some $m \in \mathfrak{M}$, and thus $q(y) = q(y + m) \in q(V_\delta(x)) \subseteq W$. That is, $V_\delta(q(x)) \subseteq W$, and W is open.

(b) Let $G \subseteq \mathfrak{X}$ be an open set. Observe that $q^{-1}(q(G)) = G + \mathfrak{M} = \cup_{m\in\mathfrak{M}} G + m$ is open, being the union of open sets. By (a), $q(G)$ is open.

□

1.27. Let $\mathfrak{M}$ be a finite-dimensional linear manifold in a normed linear space $\mathfrak{X}$. Then $\mathfrak{M}$ is closed in $\mathfrak{X}$. The proof of this is left as an assignment exercise.

1.28. Proposition. *Let $\mathfrak{X}$ be a normed linear space. If $\mathfrak{M}$ and $\mathfrak{Z}$ are closed subspaces of $\mathfrak{X}$ and* $\dim \mathfrak{Z} < \infty$, *then $\mathfrak{M} + \mathfrak{Z}$ is closed in $\mathfrak{X}$.*

Proof. Let $q : \mathfrak{X} \to \mathfrak{X}/\mathfrak{M}$ denote the canonical quotient map. Since $\mathfrak{Z}$ is a finite dimensional vector space, so is $q(\mathfrak{Z})$. By the exercise preceding this Proposition, $q(\mathfrak{Z})$ is closed in $\mathfrak{X}/\mathfrak{M}$. Since q is continuous, $\mathfrak{M} + \mathfrak{Z} = q^{-1}(q(\mathfrak{Z}))$ is closed in $\mathfrak{X}$.

□

Appendix to Section 1.

1.29. This course assumes that the reader has taken at least enough Real Analysis to have seen that $(\ell^p_{\mathbb{K}}(\mathbb{N}), \|\cdot\|_p)$ is a normed linear space for each $1 \le p \le \infty$. Having said that, let us review Hölder's Inequality as well as Minkowski's Inequality in this setting, since Hölder's Inequality is also useful in studying *dual spaces* in the next Section. The reader will recall that Minkowski's Inequality is the statement that the p-norm is subadditive; that is, that the p-norm satisfies condition (iii) of Definition 1.2. We remark that both inequalities hold for more general L^p-spaces. Our decision to concentrate on ℓ^p-spaces instead of their more general counterparts is an attempt to accommodate the background of the students who took this course, as opposed to a conscious effort to avoid L^p-spaces.

Before proving Hölder's Inequality, we pause to prove the following Lemma.

1.30. Lemma. *Let a and b be positive real numbers and suppose that $1 < p, q < \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$a^{\frac{1}{p}}\, b^{\frac{1}{q}} \le \frac{a}{p} + \frac{b}{q}.$$

Proof. Let $0 < t < 1$ and consider the function

$$f(x) = x^t - tx + t - 1,$$

defined on $(0, \infty)$. Then

$$f'(x) = tx^{t-1} - t = t(x^{t-1} - 1).$$

Thus $f(1) = 0 = f'(1)$. Since $f'(x) > 0$ for $x \in (0,1)$ and $f'(x) < 0$ for $x \in (1, \infty)$, it follows that

$$f(x) < f(1) = 0 \text{ for all } x \ne 1.$$

That is, $x^t \le (1-t) + tx$ for all $x > 0$, with equality holding if and only if $x = 1$.

Letting $x = a/b$, $t = 1/p$ yields

$$\begin{aligned} a^{\frac{1}{p}} b^{\frac{1}{q}-1} &= a^{\frac{1}{p}} b^{\frac{-1}{p}} \\ &= (\frac{a}{b})^{\frac{1}{p}} \\ &\le (1 - \frac{1}{p}) + \frac{1}{p}(\frac{a}{b}) \\ &= \frac{1}{p}(\frac{a}{b}) + \frac{1}{q}. \end{aligned}$$

Multiplying both sides of the equation by b yields the desired inequality.

$\square$

1.31. Theorem. ***Hölder's Inequality***
Let $1 \le p, q \le \infty$, and suppose that $\frac{1}{p} + \frac{1}{q} = 1$. Let $x = (x_n)_n \in \ell^p$ and $y = (y_n)_n \in \ell^q$. If $z = (z_n)_n$, where $z_n = x_n y_n$ for all $n \ge 1$, then $z \in \ell^1$ and

$$\|z\|_1 \le \|x\|_p \, \|y\|_q.$$

Proof. The cases where $p = 1$ or $p = \infty$ are routine and are left to the reader.

First let us suppose that $\|x\|_p = \|y\|_q = 1$. Applying the previous Lemma to our sequences x and y yields, for each $n \ge 1$,

$$\begin{aligned} |x_n y_n| &= (|x_n|^p)^{\frac{1}{p}} \ (|y_n|^q)^{\frac{1}{q}} \\ &\le \frac{1}{p}|x_n|^p + \frac{1}{q}|y_n|^q, \end{aligned}$$

so that

$$\begin{aligned} \sum_n |z_n| &= \sum_n |x_n y_n| \\ &\le \frac{1}{p}\sum_n |x_n|^p + \frac{1}{q}\sum_n |y_n|^q \\ &= \frac{1}{p}\|x\|_p^p + \frac{1}{q}\|y\|_q^q \\ &= 1. \end{aligned}$$

In general, if $x \in \ell^p$ and $y \in \ell^q$, let $u = x/(\|x\|_p)$, $v = v/(\|y\|_q)$ so that $\|u\|_p = 1 = \|v\|_q$ and so

$$\begin{aligned} \frac{1}{\|x\|_p \, \|y\|_q} \sum_n |x_n y_n| &= \sum_n |u_n v_n| \\ &\le 1. \end{aligned}$$

Thus

$$\|z\|_1 \le \|x\|_p \, \|y\|_q.$$

□

Hölder's Inequality is the key to proving Minkowski's Inequality.

1.32. Theorem. ***Minkowski's Inequality.***
Let $1 \le p \le \infty$, and suppose that $x = (x_n)_n$ and $y = (y_n)_n$ are in ℓ^p. Then $x + y = (x_n + y_n)_n \in \ell^p$ and

$$\|x + y\|_p \le \|x\|_p + \|y\|_p.$$

Proof. Again, the cases where $p = 1$ and where $p = \infty$ are left to the reader. Suppose therefore that $1 < p < \infty$. Observe that if $a, b > 0$, then

$$\left(\frac{a+b}{2}\right)^p \le a^p + b^p,$$

so that $(a+b)^p \le 2^p(a^p+b^p)$. It follows that

$$\sum_n |x_n+y_n|^p \le 2^p \sum_n (|x_n|^p + |y_n|^p) < \infty,$$

which proves that $x+y \in \ell^p$.

By Hölder's Inequality,

$$\sum_n |x_n+y_n|^{p-1}\,|x_n| \le \|x\|_p\ \|(|x_n+y_n|^{p-1})_n\|_q,$$

and similarly

$$\sum_n |x_n+y_n|^{p-1}\,|y_n| \le \|y\|_p\ \|(|x_n+y_n|^{p-1})_n\|_q.$$

Now

$$\begin{aligned}
\|(|x_n+y_n|^{p-1})_n\|_q &= \left(\sum_n |x_n+y_n|^{(p-1)q}\right)^{\frac{1}{q}} \\
&= \left(\sum_n |x_n+y_n|^{(pq-q)}\right)^{\frac{1}{q}} \\
&= \left(\sum_n |x_n+y_n|^{p}\right)^{\frac{1}{q}} \\
&= \|(x_n+y_n)_n\|_p^{p/q}.
\end{aligned}$$

Hence

$$\begin{aligned}
\|x+y\|_p^p &= \sum_n |x_n+y_n|\ |x_n+y_n|^{p-1} \\
&\le \sum_n (|x_n|+|y_n|)\ \ |x_n+y_n|^{p-1} \\
&\le (\|x\|_p + \|y\|_p)\ \ \|(|x_n+y_n|^{p-1})_n\|_q \\
&= (\|x\|_p + \|y\|_p)\ \ \|(x_n+y_n)_n\|^{p/q},
\end{aligned}$$

from which we get

$$\|x+y\|_p = \|x+y\|_p^{p-p/q} \le \|x\|_p + \|y\|_p.$$

□

Let us now examine a couple of examples of useful Banach spaces whose definitions require a somewhat better background in Analysis than we are assuming in the main body of the text.

1.33. Example. Let $x = (x_n)_n$ be a sequence of complex (or real) numbers. The **total variation** of x is defined by

$$V(x) := \sum_{n=1}^{\infty} |x_{n+1} - x_n|.$$

If $V(x) < \infty$, we say that x has **bounded variation**. The space

$$\mathbf{bv} := \{(x_n)_n : x_n \in \mathbb{K}, n \geq 1, V(x) < \infty\}$$

is called the space of **sequences of bounded variation**. We may define a norm on $\mathbf{bv}$ as follows: for $x \in \mathbf{bv}$, we set

$$\|(x_n)_n\|_{\mathbf{bv}} := |x_1| + V(x) = |x_1| + \sum_{n=1}^{\infty} |x_{n+1} - x_n|.$$

It can be shown that $\mathbf{bv}$ is complete under this norm, and hence that $\mathbf{bv}$ is a Banach space.

If we let $\mathbf{bv}_0 = \{(x_n)_n \in \mathbf{bv} : \lim_{n\to\infty} x_n = 0\}$, then

$$\|(x_n)_n\|_{\mathbf{bv}_0} := V((x_n)_n)$$

defines a norm on $\mathbf{bv}_0$, and again, $\mathbf{bv}_0$ is a Banach space with respect to this norm.

1.34. Example. The geometric theory of real Banach spaces is an active and exciting area. For a period of time, the following question was open [**Lin71**]: *does every infinite-dimensional Banach space contain a subspace which is linearly homeomorphic to one of the spaces ℓ^p, $1 \leq p < \infty$ or c_0*? In 1974, B.S. Tsirel'son [**Tsi74**] provided a counterexample to this conjecture. In this example, we shall discuss the broad outline of the construction of the Tsirel'son space, omitting the proofs of certain technical details.

We begin by considering the space c_0 of Example 1.4. For each $n \geq 1$, let $e_n \in c_0$ denote the sequence $(0, 0, ..., 0, 1, 0, 0, ...)$, with the unique "1" occurring in the n^{th} coordinate. Given $x = (x_n)_n \in c_0$, we may write $x = \sum_{n=1}^{\infty} x_n e_n$. Let us also define the map $P_n : c_0 \to c_0$ via $P_n(x_k)_k := (0, 0, ..., 0, x_{n+1}, x_{n+2}, x_{n+3}, ...)$.

Given a finite set $\{v_1, v_2, ..., v_r\}$ of vectors in c_0, we shall say that they are **block-disjoint** for **consecutively supported** – written $v_1 < v_2 < \cdots < v_r$ – if there exist $\alpha_1, \alpha_2, ..., \alpha_r, \beta_1, \beta_2, ..., \beta_r \in \mathbb{N}$ with

$$\alpha_1 \leq \beta_1 < \alpha_2 \leq \beta_2 < \cdots < \alpha_r \leq \beta_r$$

so that $\mathrm{supp}(v_j) \subseteq [\alpha_j, \beta_j]$, $1 \leq j \leq r$. Here, for $x = (x_n)_n \in c_0$,

$$\mathrm{supp}(x) := \{j \in \mathbb{N} : x_j \neq 0\}.$$

We shall write $(v_1, v_2, ..., v_r)$ for $\sum_{j=1}^{r} v_j$ when $v_1 < v_2 < \cdots < v_r$.

For a subset $\mathcal{B} \subseteq c_0$, we consider the following set of conditions which $\mathcal{B}$ may or may not possess:

(a) $x \in \mathcal{B}$ implies that $\|x\|_\infty \leq 1$; i.e. $\mathcal{B}$ is contained in the unit ball of c_0.

(b) $\{e_n\}_{n=1}^\infty \subseteq \mathcal{B}$.
(c) If $x = \sum_{n=1}^\infty x_n e_n \in \mathcal{B}$, $y = (y_n)_n \in c_0$ and $|y_n| \le |x_n|$ for all $n \ge 1$, then $y \in \mathcal{B}$. (This is a hereditary property.)
(d) If $v_1 < v_2 < \cdots < v_r$ lie in $\mathcal{B}$, then $\frac{1}{2} P_r((v_1, v_2, ..., v_r)) \in \mathcal{B}$.
(e) For every $x \in \mathcal{B}$ there exists $n \in \mathbb{N}$ for which $2P_n(x) \in \mathcal{B}$.

Our first goal is to construct a set K which has all five of these properties.

Let $L_1 = \{re_j : -1 \le r \le 1, j \ge 1\}$ and for $n \ge 1$, set

$$L_{n+1} = L_n \cup \{\frac{1}{2} P_r((v_1, v_2, ..., v_r)) : r \ge 1, v_1 < v_2 < \cdots < v_r \in L_n\}.$$

Let K denote the pointwise closure of $\cup_{n\ge1} L_n$. It can be shown that $K \subseteq c_0$. We set $\mathcal{D} = \overline{co}(K)$ denote the closed convex hull of K (with the closure taking place in c_0).

The Tsirel'son space T is then defined as span $\mathcal{D}$. The norm on T is given by the **Minkowski functional** which we shall encounter later when studying locally convex spaces. It is given by $\|x\|_T = \inf\{r \in (0, \infty) : x \in r\mathcal{D}\}$, where $r\mathcal{D} = \{ry : y \in \mathcal{D}\}$. As we shall later see, the definition of this norm ensures that $\mathcal{D}$ is precisely the unit ball of T.

Although we shall not prove it here, $(T, \|\cdot\|_T)$ is a Banach space which does not contain any copy of c_0 or ℓ^p, $1 \le p < \infty$.

1.35. Example. Another Banach space of interest to those who study the geometry of said spaces is **James' space**.

For a sequence $(x_n)_n$ of real numbers, consider the following condition, which we shall call **condition** J: for all $k \ge 1$,

$$\sup_{n_1<n_2<\cdots<n_k} \left[(x_{n_1} - x_{n_2})^2 + (x_{n_2} - x_{n_3})^2 + \cdots + (x_{n_{k-1}} - x_{n_k})^2\right] < \infty.$$

The James' space is defined to be:

$$\mathfrak{J} = \{(x_n)_n \in c_0 : (x_n)_n \text{ satisfies condition } J\}.$$

The norm on $\mathfrak{J}$ is defined via:

$$\|(x_n)_n\|_{\mathfrak{J}} := \sup_{n_1<n_2<\cdots<n_k} \left[(x_{n_1} - x_{n_2})^2 + (x_{n_2} - x_{n_3})^2 + \cdots + (x_{n_{k-1}} - x_{n_k})^2\right]^{\frac{1}{2}}.$$

It can be shown that $\mathfrak{J}$ is a Banach space when equipped with this norm.

1.36. Example. Let X be a locally compact topological space and let $\mathcal{B}$ denote the σ–algebra of Borel subsets of X. Let μ be a positive measure on X, so that

$$\mu : \mathcal{B} \to \mathbb{R} \cup \{\infty\}$$

satisfies

(a) $\mu(\varnothing) = 0$;
(b) $\mu(B) \ge 0$ for all $B \in \mathcal{B}$;

(c) if $\{B_n\}_n$ is a *sequence* of disjoint, measurable subsets from $\mathcal{B}$, then

$$\mu(\cup_n B_n) = \sum_n \mu(B_n).$$

The measure μ is said to be **finite** if $\mu(X) < \infty$, and it is said to be **regular** if

(i) $\mu(K) < \infty$ for all compact subsets $K \in \mathcal{B}$;
(ii) $\mu(B) = \sup\{\mu(K) : K \subseteq B, K \text{ compact}\}$ for all $B \in \mathcal{B}$; and
(iii) $\mu(B) = \inf\{\mu(G) : B \subseteq G, G \text{ open}\}$ for all $B \in \mathcal{B}$.

A **complex-valued, Borel measure** on X is a function

$$\nu : \mathcal{B} \to \mathbb{C}$$

satisfying:

(a) $\nu(\varnothing) = 0$, and
(b) if $\{B_n\}_n$ is a *sequence* of disjoint, measurable subsets from $\mathcal{B}$, then

$$\nu(\cup_n B_n) = \sum_n \nu(B_n).$$

Let ν be a complex-valued Borel measure on X. For each $B \in \mathcal{B}$, a **measurable partition** of B is a finite collection $\{E_1, E_2, ..., E_k\}$ of *disjoint*, measurable sets whose union is B. We define the **variation** $|\nu|$ of ν to be the function defined as follows: for $B \in \mathcal{B}$,

$$|\nu|(B) := \sup\{\sum_{j=1}^{k} |\nu(E_j)| : \{E_j\}_{j=1}^k \text{ is a measurable partition of } B\}.$$

It is routine to verify that $|\nu|$ is then a finite, positive Borel measure on X. We say that ν is **regular** if $|\nu|$ is.

It is clear that every complex-linear combination of finite, positive, regular Borel measures yields a complex-valued, regular Borel measure on X. A standard result from measure theory known as **the Hahn-Jordan Decomposition Theorem** states that the converse holds, namely: every complex-valued, regular Borel measure can be written as a complex-linear combination of (four) finite, positive, regular Borel measures.

Let $M_{\mathbb{C}}(X)$ denote the complex vector space of all complex-valued, regular Borel measures on X. Then the map

$$\begin{array}{rccc} \|\cdot\| : & M_{\mathbb{C}}(X) & \to & [0,\infty) \\ & \nu & \mapsto & |\nu|(X) \end{array}$$

defines a norm on $M_{\mathbb{C}}(X)$, and $M_{\mathbb{C}}(X)$ is complete with respect to this norm.

1.37. In Theorem 1.25, we showed that if $\mathfrak{X}$ is a Banach space and $\mathfrak{M}$ is a closed subspace, then $\mathfrak{X}/\mathfrak{M}$ is complete. Our proof there was based upon Proposition 1.22. This result also admits a direct proof in terms of Cauchy sequences:

Theorem. *Let $\mathfrak{X}$ be a Banach space and suppose that $\mathfrak{M}$ is a closed subspace of $\mathfrak{X}$. Then $\mathfrak{X}/\mathfrak{M}$ is complete.*

Proof.

Let $(q(x_n))_{n=1}^\infty$ be a Cauchy sequence in $\mathfrak{X}/\mathfrak{M}$. For each $n \geq 1$, there exists $k_n > 1$ so that $i, j \geq k_n$ implies $\|q(x_i) - q(x_j)\| < 2^{-n}$. Without loss of generality, we may assume that $k_n > k_{n-1}$ for all $n \geq 2$.

Set $z_n := x_{k_n}$, $n \geq 1$ and let $m_1 = 0$. For $n > 1$, choose $m_n \in \mathfrak{M}$ so that

$$\|(z_{n-1} + m_{n-1}) - (z_n + m_n)\| < 2^{-(n-1)}.$$

That this is possible follows from the definition of the quotient norm along with the inequality of the second paragraph. If we now define $y_n := z_n + m_n$, $n \geq 1$, then $q(y_n) = q(z_n) = q(x_{k_n})$, and for $n_2 > n_1$,

$$\begin{aligned}\|y_{n_1} - y_{n_2}\| &\leq \sum_{j=1}^{n_2-n_1} \|y_{n_1+j} - y_{n_1+j-1}\| \\ &\leq \sum_{j=1}^{n_2-n_1} (\frac{1}{2})^{(n_1+j-1)} \\ &\leq \sum_{j=0}^{\infty} (\frac{1}{2})^{(n_1+j)} \\ &= (\frac{1}{2})^{n_1-1},\end{aligned}$$

from which it follows that $(y_n)_{n=1}^\infty$ is Cauchy in $\mathfrak{X}$. Since $\mathfrak{X}$ is complete, $y := \lim_{n\to\infty} y_n \in \mathfrak{X}$, and since the quotient norm is contractive, $q(y) = \lim_{n\to\infty} q(y_n) = \lim_{n\to\infty} q(x_{k_n})$. Since $(q(x_n))_{n=1}^\infty$ is Cauchy, $q(y) = \lim_{n\to\infty} q(x_n)$, which proves that every Cauchy sequence in $\mathfrak{X}/\mathfrak{M}$ converges - i.e. that $\mathfrak{X}/\mathfrak{M}$ is complete.

□

1.38. Conversely, in Theorem 1.25 we used Cauchy sequences to prove that if $\mathfrak{X}$ is a normed linear space, $\mathfrak{M}$ is a closed subspace of $\mathfrak{X}$ and if both $\mathfrak{X}/\mathcal{M}$ and $\mathfrak{M}$ are complete, then $\mathfrak{X}$ is complete as well. We now provide an alternate proof that uses Proposition 1.22.

Theorem. *Suppose that $\mathfrak{X}$ is a normed linear space and that $\mathfrak{M}$ is a closed subspace of $\mathfrak{X}$. Suppose furthermore that $\mathfrak{X}/\mathfrak{M}$ and that $\mathfrak{M}$ are both complete. Then $\mathfrak{X}$ is also complete.*

Proof. Suppose that $\sum_n x_n$ is an absolutely summable series in $\mathfrak{X}$, so that $\sum_n \|x_n\| < \infty$. It suffices to prove that $\sum_n x_n$ exists in $\mathfrak{X}$.

Note that the fact that $\sum_n \|x_n\|$ is finite implies that the sequence $(\sum_{n=1}^N x_n)_{N=1}^\infty$ is Cauchy. Indeed, let $\varepsilon > 0$ and choose $M > 0$ such that $\sum_{n=M}^\infty \|x_n\| < \varepsilon$. If $t \geq s \geq M$, then

$$\|\sum_{n=1}^t x_n - \sum_{n=1}^s x_n\| = \|\sum_{n=s+1}^t x_n\| \leq \sum_{n=s+1}^t \|x_n\| \leq \sum_{n=M}^\infty \|x_n\| < \varepsilon.$$

This will be useful later on.

Let $q : \mathfrak{X} \to \mathfrak{X}/\mathfrak{M}$ denote the canonical quotient map. Then q is linear and contractive, so

$$\sum_n \|q(x_n)\| \le \sum_n \|x_n\| < \infty.$$

Since $\mathfrak{X}/\mathfrak{M}$ is complete, this absolutely summable series is summable, and since q is surjective, this implies that there exists $y \in \mathfrak{X}$ so that $q(y) = \sum_n q(x_n)$. In other words,

$$\lim_{N \to \infty} \|q(y - \sum_{n=1}^N x_n)\| = 0.$$

For each $N \ge 1$, let $\delta_N := \|q(y - \sum_{n=1}^N x_n)\|$. Then $\lim_{N\to\infty} \delta_N = 0$, so we can choose a subsequence $(N_k)_k$ so that $\delta_{N_k} < \frac{1}{2^{k+1}}$, $k \ge 1$.

By definition of the quotient norm, we can choose $m_1 \in \mathfrak{M}$ so that

$$\|y - \sum_{n=1}^{N_1} x_n - m_1\| < 2\delta_{N_1} = 1/2.$$

Next, we can choose $m_2 \in \mathfrak{M}$ so that

$$\|\big(y - \sum_{n=1}^{N_2} x_n\big) - m_1 - m_2\| < 2\delta_{N_2} = 1/4.$$

More generally, for each $k \ge 1$, we can choose $m_k \in \mathfrak{M}$ so that

$$\|\big(y - \sum_{n=1}^{N_k} x_n\big) - \sum_{n=1}^{k-1} m_n - m_k\| < 2\delta_{N_k} = 1/2^k.$$

In particular, $\lim_{k\to\infty} \|\big(y - \sum_{n=1}^{N_k} x_n\big) - \sum_{n=1}^{k} m_n\| = 0$.

Observe that for any $k \ge 1$, it follows from the triangle inequality that

$$\begin{aligned}\|m_k\| &\le \|\big(y - \sum_{n=1}^{N_k} x_n\big) - \sum_{n=1}^{k} m_n\| + \|\big(y - \sum_{n=1}^{N_{k-1}} x_n\big) - \sum_{n=1}^{k-1} m_n\| + \|x_{N_k}\| \\ &< \frac{1}{2^k} + \frac{1}{2^{k-1}} + \|x_{N_k}\|.\end{aligned}$$

Thus

$$\sum_{k=1}^\infty \|m_k\| \le \sum_{k=1}^\infty \frac{1}{2^k} + \sum_{k=1}^\infty \frac{1}{2^{k-1}} + \sum_{k=1}^\infty \|x_{N_k}\| < \infty.$$

Since $\mathfrak{M}$ is assumed to be complete, it follows that $m_0 := \sum_{k=1}^{\infty} m_k \in \mathfrak{M}$ exists. Also,

$$\begin{aligned} -y + m_0 &= \lim_{k\to\infty} \big(y - \sum_{n=1}^{N_k} x_n - \sum_{n=1}^{k} m_n\big) + (-y + m_0) \\ &= \lim_{k\to\infty} -\sum_{n=1}^{N_k} x_n + (m_0 - \sum_{n=1}^{k} m_n) \\ &= \lim_{k\to\infty} -\sum_{n=1}^{N_k} x_n, \end{aligned}$$

so that $\lim_{k\to\infty} \sum_{n=1}^{N_k} x_n = m_0 - y \in \mathfrak{X}$ exists.

But then the Cauchy sequence $(\sum_{n=1}^{N} x_n)_{N=1}^{\infty}$ has a subsequence $(\sum_{n=1}^{N_k} x_n)_{k=1}^{\infty}$ which converges to $m_0 - y$, and so the original sequence must also converge to the same limit. That is,

$$\sum_{n=1}^{\infty} x_n = \lim_{N\to\infty} \sum_{n=1}^{N} x_n = \lim_{k\to\infty} \sum_{n=1}^{N_k} x_n = m_0 - y \in \mathfrak{X}.$$

Since every absolutely summable series in $\mathfrak{X}$ is summable, $\mathfrak{X}$ is complete.

□

*

I have the body of an eighteen year old. I keep it in the fridge.

Spike Milligan

Exercises for Section 1.

Question 1.

Let $\varnothing \neq X$ be a compact, Hausdorff space. Prove that for each $\varnothing \neq \Omega \subseteq X$, the function

$$\begin{array}{rccc} \nu_\Omega : & \mathcal{C}(X, \mathbb{K}) & \to & \mathbb{R} \\ & f & \mapsto & \sup_{x \in \Omega} |f(x)| \end{array}$$

defines a seminorm on $\mathcal{C}(X, \mathbb{K})$, and that it is a norm if and only if Ω is dense in X.

Question 2.

Prove that $(\mathcal{C}([0,1], \mathbb{K}), \|\cdot\|_\infty)$ is a Banach space.

Question 3.

Prove that the disc algebra $\mathcal{A}(\mathbb{D})$ defined in Example 1.11 is a Banach space.

2. An introduction to operators

Some people are afraid of heights. Not me, I'm afraid of widths.

Steven Wright

2.1. The study of mathematics is the study of mathematical objects and the relationships between them. These relationships are often measured by functions from one object to another. Of course, when both objects belong to the same category (be it the category of vector spaces, groups, rings, etc), it is to be expected that the most important maps between these objects will be morphisms from that category. In this Section we shall concern ourselves with bounded linear operators between normed linear spaces. These bounded linear maps, as we shall soon discover, coincide with those linear maps which are continuous in the norm topology. Since normed linear spaces are vector spaces equipped with a norm topology, the bounded linear operators are the natural morphisms between them.

It should be pointed out that Banach spaces can be quite complicated to analyze. For this reason, many people working in this area often study the structure and geometry of these spaces without necessarily emphasizing the study of the linear maps between them. In the next Section we shall examine the notion of a Hilbert space. These are amongst the best-behaved Banach spaces, and their structure is relatively well understood. For this reason, fewer people study Hilbert spaces alone; Hilbert space theory tends to focus on the theory of the bounded linear maps between them, as well as algebras of such bounded linear operators.

We would also be remiss if we failed to point out that not everyone on the planet restricts themselves to bounded (i.e. continuous) linear operators. Differentiation has the grave misfortune of being an unbounded linear operator, but nevertheless it is hard to avoid if one wishes to study the world around one - or around one's friends, acquaintances, enemies, and every other one. Indeed, in applied mathematics and physics, it is often the case that the unbounded linear operators are the more interesting examples. Having said that, we shall leave it to the disciples of those schools to wax poetic on these topics.

2.2. Definition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be normed linear spaces, and let $T : \mathfrak{X} \to \mathfrak{Y}$ be a linear map. We say that T is a* ***bounded operator*** *if there exists a constant $k \geq 0$ so that $\|Tx\| \leq k\|x\|$ for all $x \in \mathfrak{X}$. When T is bounded, we define*

$$\|T\| = \inf\{k \geq 0 : \|Tx\| \leq k\|x\| \text{ for all } x \in \mathfrak{X}\}.$$

We shall refer to $\|T\|$ as the ***operator norm*** *of T.*

It is, of course, understood that the norm of Tx is computed using the $\mathfrak{Y}$-norm, while the norm of x is computed using the $\mathfrak{X}$-norm. As we shall see below, the operator norm does define a *bona fide* norm on the vector space of bounded linear maps from $\mathfrak{X}$ to $\mathfrak{Y}$, thereby justifying our terminology.

Our interest in bounded operators stems from the fact that they are precisely the continuous operators from $\mathfrak{X}$ to $\mathfrak{Y}$.

2.3. Notation. Given a Banach space $\mathfrak{X}$ and a real number $r \geq 0$, we denote by $\mathfrak{X}_r$ the set

$$\mathfrak{X}_r := \{x \in \mathfrak{X} : \|x\| \leq r\},$$

and by $\mathfrak{S}_r$ the set

$$\mathfrak{S}_r := \{x \in \mathfrak{X} : \|x\| = r\}.$$

2.4. Theorem. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be normed linear spaces and $T : \mathfrak{X} \to \mathfrak{Y}$ be a linear map. The following are equivalent:*

(a) *T is continuous on $\mathfrak{X}$.*
(b) *T is continuous at 0.*
(c) *T is bounded.*
(d) $\kappa_1 := \sup\{\|Tx\| : \ x \in \mathfrak{X}, \|x\| \leq 1\} < \infty$.
(e) $\kappa_2 := \sup\{\|Tx\| : \ x \in \mathfrak{X}, \|x\| = 1\} < \infty$.
(f) $\kappa_3 := \sup\{\|Tx\|/\|x\| : \ 0 \neq x \in \mathfrak{X}\} < \infty$.

Furthermore, if any of these holds, then $\kappa_1 = \kappa_2 = \kappa_3 = \|T\|$.

Proof.

(a) implies (b): This is trivial.
(b) implies (c): Suppose that T is continuous at 0. Let $\varepsilon = 1$ and choose $\delta > 0$ so that $\|x - 0\| < \delta$ implies that $\|Tx - T0\| = \|Tx\| < \varepsilon = 1$. If $\|y\| \leq \delta/2$, then $\|Ty\| \leq 1$, and so $0 \neq x \in \mathfrak{X}$ implies that

$$\|T\big(\frac{\delta}{2\|x\|}x\big)\| \leq 1,$$

i.e. $\|Tx\| \leq \frac{2}{\delta}\|x\|$. Since $\|T0\| = 0 \leq \frac{2}{\delta}\ \|0\|$, we see that T is bounded.
(c) implies (d): This is trivial.
(d) implies (e): This is trivial.
(e) implies (f): Again, this is trivial.
(f) implies (a): Observe that for any $x \in \mathfrak{X}$, $\|Tx\| \leq \kappa_3\|x\|$. (For $x \neq 0$, this follows from the hypothesis, while the linearity of T implies that $T0 = 0$, so the inequality also holds for $x = 0$.) Thus if $\varepsilon > 0$ and $\|x - y\| < \varepsilon/(\kappa_3 + 1)$, then

$$\|Tx - Ty\| = \|T(x - y)\| \leq \kappa_3\|x - y\| < \varepsilon.$$

The proof of the final statement is left as an exercise for the reader.

□

2.5. Computing the operator norm of a given operator T is not always a simple task. For example, suppose that $\mathcal{H} = (\mathbb{C}^2, \|\cdot\|_2)$ is a two-dimensional Hilbert space with standard orthonormal basis $\{e_1 = (1,0), e_2 = (0,1)\}$. Let $T : \mathcal{H} \to \mathcal{H}$ be the map whose matrix with respect to this basis is $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, so that $T(x,y) = (x+2y, 3x+4y)$. By definition,

$$\begin{aligned}\|T\| &= \sup\{\|Tz\| : z \in \mathbb{C}^2, \|z\| \le 1\} \\ &= \sup\{\sqrt{|x+2y|^2 + |3x+4y|^2} : x,y \in \mathbb{C}, \sqrt{|x|^2+|y|^2} \le 1\},\end{aligned}$$

which involves non-linear equations. For Hilbert spaces of low dimension – say, less than dimension 5 – alternate methods exist (but won't be developed just yet). Instead, we turn our attention to special classes of operator which are simple enough to allow us to obtain interesting results.

So as to satisfy the curious reader, we mention that the norm of T is $\sqrt{15 + \sqrt{221}}$.

2.6. Example. Multiplication operators.

(a) Let $\mathfrak{X} = \big(\mathcal{C}([0,1], \mathbb{C}), \|\cdot\|_\infty\big)$, and suppose that $f \in \mathfrak{X}$. Define

$$M_f : \begin{array}{ccc} \mathfrak{X} & \to & \mathfrak{X} \\ g & \mapsto & fg \end{array}.$$

It is routine to check that M_f is linear. If $\|g\|_\infty \le 1$, then

$$\|M_f g\|_\infty = \|fg\|_\infty = \sup\{|f(x)g(x)| : x \in [0,1]\} \le \|f\|_\infty \; \|g\|_\infty.$$

Thus $\|M_f\|_\infty \le \|f\|_\infty < \infty$, and M_f is bounded.

Setting $g(x) = 1$, $x \in [0,1]$, we have $g \in \mathfrak{X}, \|g\|_\infty = 1$ and $\|M_f g\|_\infty = \|f\|_\infty$, so that $\|M_f\| \ge \|f\|_\infty$ and therefore $\|M_f\| = \|f\|_\infty$.

For (hopefully) obvious reasons, M_f is referred to as a *multiplication operator* .

(b) We now consider a similar operator acting on a Hilbert space. Let $\mathcal{H} = L^2(X, d\mu)$, where $d\mu$ is a positive, regular Borel measure. Suppose that $f \in L^\infty(X, d\mu)$ and let

$$M_f : \begin{array}{ccc} \mathcal{H} & \to & \mathcal{H} \\ g & \mapsto & fg \end{array}.$$

Once again, it is easy to check that M_f is linear, while for $g \in \mathcal{H}$,

$$\begin{aligned}\|M_f g\|_2 &= \Big(\int_X |f(x)g(x)|^2 d\mu\Big)^{\frac{1}{2}} \\ &\le \Big(\int_X \|f\|_\infty^2 |g(x)|^2 d\mu\Big)^{\frac{1}{2}} \\ &= \|f\|_\infty \; \|g\|_2,\end{aligned}$$

so that $\|M_f\| \le \|f\|_\infty$, and hence M_f is bounded. As for a lower bound on the norm of M_f, for each $n \ge 1$, let $F_n = \{x \in X : |f(x)| \ge \|f\|_\infty - 1/n\}$. Then F_n is measurable and $\mu(F_n) > 0$ by definition of $\|f\|_\infty$. Let $E_n \subseteq F_n$ be a measurable set for which $0 < \mu(E_n) < \infty$, $n \ge 1$. The existence of such sets E_n, $n \ge 1$ follows from the regularity of the measure μ. Let $g_n = \chi_{E_n}$, the characteristic function of E_n. Then $g_n \in L^2(X, \mu)$ for all $n \ge 1$ and

$$\begin{aligned} \|M_f g_n\|_2 &= \Big(\int_X |f(x) g_n(x)|^2 d\mu \Big)^{\frac{1}{2}} \\ &= \Big(\int_{E_n} |f(x)|^2 d\mu \Big)^{\frac{1}{2}} \\ &\ge \Big(\int_{E_n} (\|f\|_\infty - 1/n)^2 d\mu \Big)^{\frac{1}{2}} \\ &= \big(\|f\|_\infty - 1/n \big) \ \Big(\int_X |g_n(x)|^2 d\mu \Big)^{\frac{1}{2}} \\ &= \big(\|f\|_\infty - 1/n \big) \ \|g_n\|_2. \end{aligned}$$

From this we see that $\|M_f\| \ge \|f\|_\infty - 1/n$. Since $n \ge 1$ was arbitrary, $\|M_f\| \ge \|f\|_\infty$, and so $\|M_f\| = \|f\|_\infty$.

Observe that the computation of the norm of the operator depended very much upon the underlying norms of the spaces involved.

(c) As a special case of this phenomenon, let $X = \mathbb{N}$ and suppose that $d\mu$ is counting measure. Then $\mathcal{H} = \ell^2(\mathbb{N})$ and $f \in \ell^\infty(\mathbb{N})$. As we are wont to do when dealing with sequences, we denote by d_n the value $f(n)$ of f at $n \in \mathbb{N}$, so that $f \equiv (d_n)_{n=1}^\infty \in \ell^\infty(\mathbb{N})$. It follows that $M_f (x_n)_n = (d_n x_n)_n$ for all $(x_n)_n \in \ell^2(\mathbb{N})$. By considering the matrix $[M_f]$ of M_f with respect to the standard orthonormal basis $(e_n)_n$ for $\mathcal{H}$, we see that

$$[M_f] = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & d_3 & \\ & & & \ddots \end{bmatrix}.$$

Thus, M_f, often denoted in this circumstance as $D = \mathrm{diag}\{d_n\}_n$, is referred to as a *diagonal operator*. The above calculation shows that

$$\|D\| = \|M_f\| = \|f\|_\infty = \sup\{|f(n)| : n \ge 1\} = \sup\{|d_n| : n \ge 1\}.$$

2.7. Example. Weighted shifts. With $\mathcal{H} = \ell^2(\mathbb{N})$ and $(w_n)_n \in \ell^\infty(\mathbb{N})$, consider the map $W : \mathcal{H} \to \mathcal{H}$ defined by

$$W(x_n)_n = (0, w_1 x_1, w_2 x_2, w_3 x_3, ...) \text{ for all } (x_n)_n \in \ell^2(\mathbb{N}).$$

We leave it as an exercise for the reader to show that W is a bounded linear operator on $\mathcal{H}$, and that

$$\|W\| = \sup\{|w_n| : n \ge 1\}.$$

Such an operator is referred to as a **unilateral forward weighted shift**.

If $(v_n)_n \in \ell^\infty(\mathbb{N})$ and we define the linear map $V : \mathcal{H} \to \mathcal{H}$ via

$$V(x_n)_n = (v_1 x_2, v_2 x_3, v_3 x_4, ...) \text{ for all } (x_n)_n \in \mathcal{H},$$

then once again V is bounded, $\|V\| = \sup\{|v_n| : n \geq 1\}$, and V is referred to as a **unilateral backward weighted shift**.

Finally, consider $\mathcal{H} = \ell^2(\mathbb{Z})$, and with $(u_n)_n \in \ell^\infty(\mathbb{Z})$, define the linear map $U : \mathcal{H} \to \mathcal{H}$ via

$$U(x_n)_n = (u_{n-1}x_{n-1})_n.$$

Again, U is bounded with $\|U\| = \sup\{|u_n| : n \in \mathbb{Z}\}$, and U is referred to as a **bilateral weighted shift**. The reader should ask himself/herself why we do not refer to "forward" and "backward" bilateral shift operators.

2.8. Example. Differentiation operators. Consider the linear manifold $\mathcal{P}(\mathbb{D}) = \{p : p \text{ a polynomial}\} \subseteq (\mathcal{C}(\mathbb{D}), \|\cdot\|_\infty)$. Define the map

$$\begin{array}{rccc} D: & \mathcal{P}(\mathbb{D}) & \to & \mathcal{P}(\mathbb{D}) \\ & p & \mapsto & p', \end{array}$$

the derivative of p. Then if $p_n(z) = z^n$, $\|p_n\|_\infty = 1$ for each $n \geq 1$ and $Dp_n = np_{n-1}$, whence $\|D\| \geq n$ for all $n \geq 1$. In particular, D is not bounded. That is, differentiation is not continuous on the linear space of polynomials.

2.9. Notation. The set of bounded linear operators from the normed linear space $\mathfrak{X}$ to the normed linear space $\mathfrak{Y}$ is denoted by $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. If $\mathfrak{X} = \mathfrak{Y}$, we abbreviate this to $\mathcal{B}(\mathfrak{X})$. We now fulfil an earlier promise by proving that the map $T \mapsto \|T\|$ does indeed define a norm on $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$.

2.10. Proposition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be normed linear spaces. Then $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a vector space and the operator norm is a norm on $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$.*

Proof. Since linear combinations of continuous functions between topological spaces are continuous, $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a vector space.

As to the second assertion: for $R, T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ and $k \in \mathbb{K}$,

(i) $\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\} \geq 0$;

(ii) $\|T\| = 0$ if and only if $\|Tx\|/\|x\| = 0$ for all $x \neq 0$, which in turn happens if and only if $Tx = 0$ for all $x \in \mathfrak{X}$; i.e. if and only if $T = 0$.

(iii)

$$\begin{aligned} \|kT\| &= \sup\{\|kTx\| : \|x\| \leq 1\} \\ &= \sup\{|k|\ \|Tx\| : \|x\| \leq 1\} \\ &= |k|\ \sup\{\|Tx\| : \|x\| \leq 1\} \\ &= |k|\ \|T\|. \end{aligned}$$

(iv)

$$\begin{aligned} \|R + T\| &= \sup\{\|(R + T)x\| : \|x\| \leq 1\} \\ &\leq \sup\{\|Rx\| + \|Tx\| : \|x\| \leq 1\} \\ &\leq \sup\{\|Rx\| + \|Ty\| : \|x\|, \|y\| \leq 1\} \\ &= \|R\| + \|T\|. \end{aligned}$$

This completes the proof.

□

2.11. Theorem. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be normed linear spaces and suppose that $\mathfrak{Y}$ is complete. Then $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is complete, and as such it is a Banach space.*

Proof. Suppose that $\sum_{n=1}^{\infty} T_n$ is an absolutely summable series in $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Given $x \in \mathfrak{X}$,

$$\sum_{n=1}^{\infty} \|T_n x\| \leq \sum_{n=1}^{\infty} \|T_n\| \, \|x\| = \|x\| \big(\sum_{n=1}^{\infty} \|T_n\| \big) < \infty,$$

and thus, since $\mathfrak{Y}$ is compete, $\sum_{n=1}^{\infty} T_n x \in \mathfrak{Y}$ exists. Moreover, the linearity of each T_n implies that the map $T : \mathfrak{X} \to \mathfrak{Y}$ defined via $Tx = \sum_{n=1}^{\infty} T_n x$ is linear, while $\|x\| \leq 1$ implies from above that $\|Tx\| \leq \sum_{n=1}^{\infty} \|T_n\|$. Hence

$$\|T\| \leq \sum_{n=1}^{\infty} \|T_n\| < \infty,$$

implying that T is bounded.

Finally,

$$\begin{aligned} \|Tx - \sum_{n=1}^{N} T_n x\| &= \| \sum_{n=N+1}^{\infty} T_n x \| \\ &\leq \|x\| \sum_{n=N+1}^{\infty} \|T_n\|, \end{aligned}$$

from which it easily follows that $T = \lim_{N \to \infty} \sum_{n=1}^{N} T_n$. That is, the series $\sum_{n=1}^{\infty} T_n$ is summable. By Proposition 1.22, $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is complete.

□

As a particular case of Theorem 2.11, consider the case where $\mathfrak{Y} = \mathbb{K}$, the base field.

2.12. Definition. *Let $\mathfrak{X}$ be a normed linear space. The **dual** of $\mathfrak{X}$ is $\mathcal{B}(\mathfrak{X}, \mathbb{K})$, and it is denoted by $\mathfrak{X}^*$. The elements of $\mathfrak{X}^*$ are referred to as **continuous linear functionals** or – when no confusion is possible – as **functionals** on $\mathfrak{X}$.*

Since $\mathbb{K}$ is complete, Theorem 2.11 implies that $\mathfrak{X}^*$ is again a Banach space. As such, we may consider the dual space of $\mathfrak{X}^*$, namely $\mathfrak{X}^{(2)} = \mathfrak{X}^{**} := (\mathfrak{X}^*)^*$, known as the **double dual** of $\mathfrak{X}$, and more generally, the **$\mathbf{n^{th}}$-iterated dual spaces** $\mathfrak{X}^{(n)} = (\mathfrak{X}^{(n-1)})^*$, $n \geq 3$. All of these are Banach spaces.

Before proceeding to some examples, let us first introduce some notation and terminology which will prove useful.

2.13. Definition. *A collection $\{e_n\}_{n=1}^\infty$ in a Banach space $\mathfrak{X}$ is said to be a* ***Schauder basis*** *if every $x \in \mathfrak{X}$ can be written in a* ***unique*** *way as a norm convergent series*

$$x = \sum_{n=1}^\infty x_n e_n$$

for some choice $x_n \in \mathbb{K}$, $n \ge 1$.

2.14. Example.

(a) For each $n \ge 1$, let e_n denote the sequence $e_n = (0, 0, ..., 0, 1, 0, 0, ...) \in \mathbb{K}^{\mathbb{N}}$, where the unique "1" occurs in the n^{th} position. Then $\{e_n\}_n$ is a Schauder basis for c_0 and for ℓ^p, $1 \le p < \infty$. We shall refer to $\{e_n\}_n$ as the **standard Schauder basis** for c_0 and for ℓ^p.

Observe that it is *not* a Schauder basis for ℓ^∞. Indeed, ℓ^∞ does not admit *any* Schauder basis (why not?).

(b) It is not as obvious what one should choose as the Schauder basis for $(\mathcal{C}[0,1], \mathbb{R})$. It was Schauder [**Sch27**] who first discovered a basis for this space. The description of such a basis is non-trivial.

2.15. Example. Consider the Banach space

$$c_0 = \{(x_n)_{n=1}^\infty \in \mathbb{K}^{\mathbb{N}} : \lim_{n\to\infty} x_n = 0\},$$

equipped with the supremum norm $\|(x_n)_n\|_\infty = \sup\{|x_n| : n \ge 1\}$. We claim that c_0^* is isometrically isomorphic to $\ell^1 = \ell^1(\mathbb{N})$. To see this, consider the map

$$\begin{array}{cccc} \Theta : & (\ell^1, \|\cdot\|_1) & \to & (c_0, \|\cdot\|_\infty)^* \\ & z := (z_n)_n & \mapsto & \varphi_z \end{array},$$

where $\varphi_z((x_n)_n) = \sum_{n=1}^\infty x_n z_n$ for all $(x_n)_n \in c_0$. That Θ is linear is readily seen. That the sum converges absolutely is also easy to verify.

If $\|(x_n)_n\|_\infty \le 1$, then $|x_n| \le 1$ for all $n \ge 1$, so that

$$|\varphi_z((x_n)_n)| = |\sum_{n=1}^\infty x_n z_n| \le \sum_{n=1}^\infty |x_n z_n| \le \sum |z_n| = \|z\|_1.$$

Hence $\|\varphi_z\| \le \|z\|_1$. On the other hand, if we set $v[n] = (w_1, w_2, w_3, ..., w_n, 0, 0, 0, ...)$ for each $n \ge 1$ (where $w_j = \overline{z_j}/|z_j|$ if $z_j \ne 0$, while $w_j = 1$ if $z_j = 0$), then $v[n] \in c_0$, $\|v[n]\|_\infty = 1$ for all $n \ge 1$, and

$$\varphi_z(v[n]) = \sum_{j=1}^n |z_j|.$$

From this it follows that $\|\varphi_z\| \ge \|z\|_1$. Combining these two estimates yields $\|\varphi_z\| = \|z\|_1$.

Thus Θ is an isometric injection of ℓ^1 into c_0. There remains to prove that Θ is surjective.

To that end, suppose that $\varphi \in c_0^*$. Let $\{e_n\}_n$ denote the standard Schauder basis for c_0, and for each $n \geq 1$, let $w_n = \varphi(e_n)$. Observe that if $\beta_n := \overline{w_n}/|w_n|$ for $w_n \neq 0$, and $\beta_n := 0$ if $w_n = 0$, then

$$v[n] := \sum_{k=1}^{n} \beta_k e_k \in c_0$$

and $\|v[n]\|_\infty \leq 1$. Since, for all $n \geq 1$,

$$\begin{aligned}
\sum_{k=1}^{n} |w_k| &= \left| \sum_{k=1}^{n} \beta_k w_k \right| \\
&= \left| \sum_{k=1}^{n} \varphi(\beta_k e_k) \right| \\
&= |\varphi(v[n])| \\
&\leq \|\varphi\| \, \|v[n]\|_\infty \\
&= \|\varphi\|,
\end{aligned}$$

we see that $w := (w_1, w_2, w_3, ...) \in \ell^1$. A routine computation shows that $\varphi_w|_{c_{00}} = \varphi|_{c_{00}}$. Since φ, φ_w are both continuous and since c_{00} is dense in c_0, $\varphi = \varphi_w = \Theta(w)$. Thus Θ is onto.

2.16. Example. Let $1 \leq p < \infty$. Recall from your Real Analysis courses that there exists an isometric linear bijection $\Theta : (\ell^q, \|\cdot\|_q) \to (\ell^p, \|\cdot\|_p)^*$, where – as always – q is the Lebesgue conjugate of p satisfying $1/p + 1/q = 1$.

The map is defined via:

$$\begin{array}{rccc}
\Theta : & \ell^q & \to & (\ell^p)^* \\
& z & \mapsto & \varphi_z,
\end{array}$$

where for $z = (z_n)_n \in \ell^q$, we have $\varphi_z((x_n)_n) = \sum_n x_n z_n$.

We normally abbreviate this result by saying that *the dual of* ℓ^p *is* ℓ^q*, when* $1 \leq p < \infty$. We refer the reader to the Appendix to Section 2 for a proof of this result.

2.17. Example. The above example can be extended to more general measure spaces. Let $1 \leq p < \infty$, and suppose that μ is a σ-finite, positive, regular Borel measure on $L^p(X, \mu)$. Again, the map

$$\begin{array}{rccc}
\Theta : & L^q(X, \mu) & \to & L^p(X, \mu)^* \\
& g & \mapsto & \varphi_g,
\end{array}$$

where $\varphi_g(f) = \int_X f g d\mu$ defines a linear, isometric bijection between $L^q(X, \mu)$ and $L^p(X, \mu)^*$.

If we drop the hypothesis that μ is σ-finite, the result still holds for $1 < p < \infty$.

For reasons we shall discuss in the next Section, when $p = 2$, we often consider the related map

$$\begin{array}{rccc} \Omega: & L^2(X,\mu) & \to & L^2(X,\mu)^* \\ & g & \mapsto & \varphi_g, \end{array}$$

where $\varphi_g(f) = \int_X f\overline{g} d\mu$ defines a conjugate-linear, isometric bijection between $L^2(X,\mu)$ and $L^2(X,\mu)^*$.

2.18. Example. A function $f : [0,1] \to \mathbb{K}$ is said to be of **bounded variation** if there exists $\kappa > 0$ such that for every partition $\{0 = t_0 < t_1 < t_2 < \cdots < t_n = 1\}$ of $[0,1]$,

$$\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| \leq \kappa.$$

The infimum of all such κ's for which the above inequality holds is denoted by $\|f\|_v$, and is called the **variation** of f.

Recall from your earlier courses in Analysis that if f is a function of bounded variation, then for all $x \in (0,1]$, $f(x^-) := \lim_{t\to x^-} f(t)$ exists, and for all $x \in [0,1)$, $f(x^+) := \lim_{t\to x^+} f(t)$ exists (though they might not be equal, of course). We set $f(0^-) = f(0)$ and $f(1^+) = f(1)$. It is known that a function of bounded variation admits at most a countable number of discontinuities in the interval $[0,1]$, and for $g \in \mathcal{C}([0,1],\mathbb{K})$, the **Riemann-Stieltjes integral**

$$\int_0^1 g \, \mathrm{d}f$$

exists.

Let

$$BV[0,1] = \{f : [0,1] \to \mathbb{K} : \|f\|_v < \infty, \ f \text{ is left-continuous on } (0,1) \text{ and } f(0) = 0\}.$$

Then $(BV[0,1], \|\cdot\|_v)$ is a Banach space with norm given by the variation. Indeed, the dual of $\big(\mathcal{C}([0,1]), \|\cdot\|_\infty\big)$ is isometrically isomorphic to $BV[0,1]$. For $f \in BV[0.1]$ and $g \in \mathcal{C}([0,1])$, we define a functional $\varphi_f \in (\mathcal{C}([0,1],\mathbb{K})$ by

$$\varphi_f(g) := \int_0^1 g \, \mathrm{d}f.$$

2.19. Proposition. *Let $\mathfrak{X}$ be a normed linear space. Then there exists a contractive linear map $\mathfrak{J} : \mathfrak{X} \to \mathfrak{X}^{**}$.*

Proof. Let $z \in \mathfrak{X}$ and define a map $\widehat{z} : \mathfrak{X}^* \to \mathbb{K}$ via $\widehat{z}(x^*) = x^*(z)$. It is routine to check that $\widehat{z}$ is linear, and if $\|x^*\| \leq 1$, then $|\widehat{z}(x^*)| = |x^*(z)| \leq \|x^*\| \, \|z\|$, so that $\|\widehat{z}\| \leq \|z\|$; in particular, $\widehat{z} \in \mathfrak{X}^{**}$.

It is also easy to verify that the map

$$\begin{array}{rccc} \mathfrak{J}: & \mathfrak{X} & \to & \mathfrak{X}^{**} \\ & z & \mapsto & \widehat{z} \end{array}$$

is linear, and the first paragraph shows that $\mathfrak{J}$ is contractive.

□

2.20. The map $\mathfrak{J}$ is referred to as the **canonical embedding** of $\mathfrak{X}$ into $\mathfrak{X}^{**}$. It is not necessarily the only embedding of interest, however. Once we have proven the Hahn-Banach Theorem, we shall be in a position to show that $\mathfrak{J}$ is in fact isometric.

We point out that if $\mathfrak{J}$ is an isometric bijection from $\mathfrak{X}$ *onto* $\mathfrak{X}^{**}$, then $\mathfrak{X}$ is said to be **reflexive**. These are in some sense amongst the best behaved Banach spaces. It is worth pointing out that the notion of reflexivity of Banach spaces is stronger than merely asking that $\mathfrak{X}$ be isometrically isomorphic to $\mathfrak{X}^{**}$. James [**Jam51**] has produced an example of a non-reflexive Banach space $\mathfrak{X}$ such that $\mathfrak{X}$ is isometrically isomorphic to $\mathfrak{X}^{**}$.

We shall return to the notion of reflexivity of Banach spaces in a later section.

Appendix to Section 2.

2.21. Although norms of operators can be difficult to compute, there are cases where useful estimates can be obtained.

Consider the **Volterra operator**

$$V: \begin{array}{ccc} \mathcal{C}([0,1],\mathbb{K}) & \to & \mathcal{C}([0,1],\mathbb{K}) \\ f & \mapsto & Vf \end{array},$$

where $Vf(x) = \int_0^x f(t)dt$ for all $x \in [0,1]$. (Since all functions are continuous, it suffices to consider Riemann integration.)

Then

$$\begin{aligned} \|Vf\| &= \sup\{|Vf(x)| : x \in [0,1]\} \\ &= \sup\{|\int_0^x f(t)dt| : x \in [0,1]\} \\ &\leq \sup\{\int_0^x \|f\|_\infty dt : x \in [0,1]\} \\ &= \sup\{(x-0)\,\|f\|_\infty : x \in [0,1]\} \\ &= \|f\|_\infty. \end{aligned}$$

Thus $\|V\| \leq 1$. If $\mathbf{1}(x) := 1$ for all $x \in [0,1]$, then $\mathbf{1} \in \mathcal{C}([0,1],\mathbb{K})$, $\|\mathbf{1}\|_\infty = 1$, and $V\mathbf{1} = j$, where $j(x) = x$, $x \in [0,1]$. But then $\|V\mathbf{1}\|_\infty = \|j\|_\infty = 1$, showing that $\|V\| \geq 1$, and hence $\|V\| = 1$.

Far more interesting (and useful) is the computation of $\|V^n\|$ for $n \geq 2$.

Let us first general the construction of the operator V. We may consider the function $k : [0,1] \times [0,1] \to \mathbb{C}$ defined by

$$k(x,y) = \begin{cases} 0 & \text{if } x < y, \\ 1 & \text{if } x \geq y. \end{cases}$$

Then

$$(Vf)(x) = \int_0^x f(y)dy = \int_0^1 k(x,y)f(y)dy.$$

The function $k = k(x,y)$ is referred to as the **kernel** of the integral operator V. This should not be confused with the notion of a null space of a linear map, also referred to as its *kernel*.

Now

$$\begin{aligned}(V^2 f)(x) &= (V(Vf))(x) \\ &= \int_0^1 k(x,t)\,(Vf)(t)dt \\ &= \int_0^1 k(x,t) \int_0^1 k(t,y)\,f(y)dy\,dt \\ &= \int_0^1 f(y) \int_0^1 k(x,t)\,k(t,y)dt\,dy \\ &= \int_0^1 f(y)\,k_2(x,y)dy,\end{aligned}$$

where $k_2(x,y) = \int_0^1 k(x,t)\,k(t,y)dt$ is a new kernel for the integral operator V^2. Note that

$$\begin{aligned}|k_2(x,y)| &= |\int_0^1 k(x,t)\,k(t,y)dt| \\ &= |\int_y^x k(x,t)\,k(t,y)dt| \\ &= (x-y) \text{ for } x > y,\end{aligned}$$

while for $x < y$, $k_2(x,y) = 0$.

In general, since $x - y < 1 - 0 = 1$, we get

$$\begin{aligned}(V^n f)(x) &= \int_0^1 f(y)\,k_n(x,y)dy, \qquad \text{where} \\ k_n(x,y) &= \int_0^1 k(x,t)\,k_{n-1}(t,y)dt, \qquad \text{and where} \\ |k_n(x,y)| &\le \frac{1}{(n-1)!}(x-y)^{n-1} \le \frac{1}{(n-1)!}.\end{aligned}$$

It follows that

$$\begin{aligned}\|V^n\| &= \sup_{\|f\|=1} \|V^n f\|_\infty \\ &= \sup_{\|f\|=1} \|\int_0^1 f(y)\,k_n(x,y)dy\|_\infty \\ &\le \sup_{\|f\|=1} \|f\|_\infty\,\|k_n(x,y)\|_\infty \\ &\le 1/(n-1)!.\end{aligned}$$

A simple consequence of these computations is that

$$\lim_{n\to\infty} \|V^n\|^{\frac{1}{n}} = \lim_{n\to\infty} 1/(n-1)! = 0.$$

We shall have more to say about this in the Appendix to Section 3.

2.22. Example. Let $\{e_1, e_2, ..., e_n\}$ denote the standard basis for $\mathbb{K}^n$, so that $e_j = (0, 0, ..., 0, 1, 0, ..., 0)$, where the 1 occurs in the j^{th} position, $1 \leq j \leq n$. Suppose that $1 \leq p \leq \infty$, and that $\mathbb{K}^n$ carries the p-norm from Example 1.7.

Let $[t_{ij}] \in \mathbb{M}_n(\mathbb{K})$, and define the map

$$\begin{array}{rccc} T: & \mathbb{K}^n & \mapsto & \mathbb{K}^n \\ & x & \mapsto & [t_{ij}]x. \end{array}$$

It is instructive, while not difficult, to prove that if every row and every column of $[t_{ij}]$ has *at most* one non-zero entry, then

$$\|T\| = \max_{1 \leq i,j \leq n} |t_{ij}|.$$

We leave this as an exercise for the reader.

2.23. Example. Let $n \geq 1$ be an integer and consider $\mathbb{M}_n = \mathbb{M}_n(\mathbb{C})$. For $T \in \mathbb{M}_n$, T^*T is a hermitian matrix, and as such, has positive eigenvalues. Denote by $s_1, s_2, ..., s_n$ the square roots of these eigenvalues (counted according to multiplicity) and for $1 \leq p < \infty$, set

$$\|T\|_p = \big(\sum_{k=1}^{n} s_k^p\big)^{\frac{1}{p}}.$$

For $p = \infty$, set

$$\|T\|_\infty = \max\{s_1, s_2, ..., s_n\}.$$

The numbers $s_1, s_2, ..., s_n$ are known as the **singular values** of the matrix T.

It can be shown that $\| \cdot \|_p$ is indeed a norm on $\mathbb{M}_n$ for all $1 \leq p \leq \infty$. Let us denote the space $\mathbb{M}_n$ equipped with the norm $\| \cdot \|_p$ by $\mathcal{C}_p^n$. We shall refer to it as the *n-dimensional Schatten p-class* of operators on $\mathbb{C}^n$. Then we shall leave it as an exercise for the reader to prove that $\mathcal{C}_p^{n*} \simeq \mathcal{C}_q^n$, where q is the Lebesgue conjugate of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$.

The above identification can be realized via the map:

$$\begin{array}{rccc} \Phi: & \mathcal{C}_q^n & \to & \mathcal{C}_p^{n*} \\ & R & \mapsto & \varphi_R \end{array}$$

where $\varphi_R : \mathcal{C}_p^n \to \mathbb{C}$ is the map $\varphi_R(T) = \text{tr}(RT)$, and where $\text{tr}[x_{ij}] = \sum_{k=1}^n x_{kk}$ denotes the standard trace functional on $\mathbb{M}_n$.

The above result has a generalisation to infinite-dimensional Hilbert spaces. We refer the reader to [**Dav88**] for a more detailed treatment of this topic.

2.24. Example. We now return to the proof of the fact that the dual of ℓ^p is ℓ^q when $1 < p < \infty$, as stated in Example 2.16.

Given $z = (z_n)_n \in \ell^q$, we define

$$\begin{array}{rccc} \beta_z: & \ell^p & \to & \mathbb{K} \\ & (x_n)_n & \mapsto & \sum_n x_n z_n. \end{array}$$

Note that by Hölder's Inequality, Theorem 1.31, for $x = (x_n)_n \in \ell^p$, we have

$$\begin{aligned} |\beta_z(x)| &= \left| \sum_n x_n z_n \right| \\ &\le \sum_n |x_n z_n| \\ &= \|xz\|_1 \\ &\le \|x\|_p \, \|z\|_q, \end{aligned}$$

so that indeed $\beta_z(x) \in \mathbb{K}$. Clearly β_z is linear, and so the above argument also shows that $\|\beta_z\| \le \|z\|_q$.

Furthermore, if we set $x_n = \alpha_n z_n^{q-1} = \alpha_n z_n^{q/p}$, where $\alpha_n \in \mathbb{K}$ is chosen so that $|\alpha_n| = 1$ and $x_n z_n \ge 0$ for all $n \ge 1$, then

$$\begin{aligned} \left(\sum_n |x_n|^p \right)^{\frac{1}{p}} &= \left(\sum_n (|z_n|^{\frac{q}{p}})^p \right)^{\frac{1}{p}} \\ &= \left(\sum_n |z_n|^q \right)^{\frac{1}{p}} \\ &= \|z\|_q^{q/p}, \end{aligned}$$

so that $x \in \ell^p$, and

$$\begin{aligned} |\beta_z(x)| &= \sum_n x_n z_n \\ &= \sum_n |z_n|^q \\ &= \|z\|_q^{q/p} \; \|z\|_q \\ &= \|x\|_p \; \|z\|_q, \end{aligned}$$

so that $\|\beta_z\| \ge \|z\|_q$, and therefore $\|\beta_z\| = \|z\|_q$.

Consider the map Θ defined via:

$$\begin{array}{rccc} \Theta : & \ell^q & \to & (\ell^p)^* \\ & z & \mapsto & \beta_z, \end{array}$$

with β_z defined as above. Then Θ is easily seen to be linear, and from above, it is isometric (hence injective). There remains only to show that Θ is surjective.

To that end, let $\varphi \in (\ell^p)^*$. For each $n \ge 1$, let $z_n := \varphi(e_n)$, where $\{e_n\}_n$ is the standard Schauder basis for ℓ^p. Set $z[n] := \sum_{k=1}^n z_k e_k$, and $x[n] := \sum_{k=1}^n \alpha_k z_k^{q-1}$, where – as before – α_k is chosen so that $|\alpha_k| = 1$ and $x_k z_k \ge 0$ for all $k \ge 1$. Then $z[n] \in \ell^q$ and $x[n] \in \ell^p$ for all $n \ge 1$.

Observe that if $y = (y_n)_n \in \ell^p$, then by the continuity of φ,

$$\varphi(y) = \varphi\left(\sum_n y_n e_n\right) = \sum_n y_n \varphi(e_n) = \sum_n y_n z_n.$$

Now

$$\begin{aligned}
|\varphi(x[n])| &= \left|\sum_{k=1}^{n} \alpha_k z_k^{q-1} z_k\right| \\
&= \sum_{k=1}^{n} |z_k|^q \\
&= \|z[n]\|_q^{q-1} \; \|z[n]\|_q,
\end{aligned}$$

where

$$\begin{aligned}
\|z[n]\|_q^{q-1} &= \left(\sum_{k=1}^{n} |z_k|^q\right)^{\frac{a-1}{q}} \\
&= \left(\sum_{k=1}^{n} (|z_k|^{q/p})^p\right)^{1-\frac{1}{q}} \\
&= \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \\
&= \|x[n]\|_p.
\end{aligned}$$

Thus

$$\begin{aligned}
|\varphi(x[n])| &= \|x[n]\|_p \; \|z[n]\|_q \\
&\leq \|x[n]\|_p \; \|\varphi\| \quad \text{for all } n \geq 1.
\end{aligned}$$

It follows that $\|z[n]\|_q \leq \|\varphi\|$ for all $n \geq 1$, so that if $z := (z_n)_n$, then $z \in \ell^q$ with $\|z\|_q \leq \|\varphi\|$.

Finally, $\varphi(y) = \beta_z(y)$ for all $y \in \ell^p$, so that $\varphi = \beta_z = \Theta(z)$, proving that Θ is surjective, as required.

2.25. Example. Let X be a compact, Hausdorff space and consider the Banach space $\mathcal{C}(X, \mathbb{C})$, equipped with the norm $\|f\|_\infty := \sup_{x \in X} |f(x)|$.

The **Riesz Representation Theorem** states that the dual $\mathcal{C}(X, \mathbb{C})^*$ of $\mathcal{C}(X, \mathbb{C})$ can be identified with the set $\mathcal{M}(X)$ of all finite, regular Borel measures on X. Given a measure $\mu \in \mathcal{M}(X)$, we associate to it the linear functional φ_μ defined by

$$\varphi_\mu(f) = \int_X f(x) d\mu(x).$$

It is worth noting that the functional φ_μ sends positive functions to positive functions precisely if μ is a positive measure. Such functionals play a significant role in the theory of C^*-algebras, of which $(\mathcal{C}(X, \mathbb{C}), \|\cdot\|_\infty)$ is an example.

*

I handed in a script last year and the studio didn't change one word. The word they didn't change was on page 87.

Steve Martin

Exercises for Section 2.

Question 1.

Why does $(\ell^\infty, \|\cdot\|_\infty)$ not admit a Schauder basis?

Question 2.

Let $\mathcal{H} = \ell^2(\mathbb{N})$, and let $(w_n)_n \in \ell^\infty(\mathbb{N})$. Denote by W the unilateral forward shift determined by

$$W(x_n)_n = (0, w_1x_1, w_2x_2, w_3x_3, \ldots), \quad x = (x_n)_n \in \mathcal{H}.$$

Prove that $\|W\| = \sup_n |w_n|$.

Question 3.

Let $\mathcal{H} = \ell^2(\mathbb{N})$ and let $\{e_n\}_{n=1}^\infty$ denote the standard ONB for $\mathcal{H}$. Given $T \in \mathcal{B}(\mathcal{H})$, denote by $[T] = [t_{i,j}]$ the **matrix** for T relative to $\{e_n\}_{n=1}^\infty$, where $t_{i,j} = \langle Te_j, e_i\rangle$ for all $i, j \geq 1$.

(a) Prove that if $T \in \mathcal{B}(\mathcal{H})$, then $\sup_{i,j\geq 1} |t_{i,j}| < \infty$.
(b) Prove the converse to (a), or find a counterexample to show that it is false.
(c) Let $T \in \mathcal{B}(\mathcal{H})$. Extend the result of Question 2 and of Example 2.6 (c) by proving that if each row and each column of $[T]$ contains only one non-zero entry, then $\|T\| = \sup_{i,j\geq 1} |t_{i,j}|$.
(d) Prove that in general, given $T \in \mathcal{B}(\mathcal{H})$, $\|T\| \geq \sup_{i,j\geq 1} |t_{i,j}|$.
(e) Let $N \geq 1$ and let $\mathcal{E}_N := \{e_1, e_2, \ldots, e_N\}$ be an ONB for $\mathbb{C}^N$. Let Q_N denote the operator whose matrix relative to $\mathcal{E}_N$ is $[Q_N] = [1_{i,j}]_{1\leq i,j\leq N}$. Find $\|Q_N\|$.

Question 4.

Let $[a_{i,j}]_{1\leq i,j}$ be a matrix with finite support; i.e. suppose that the set

$$\{(i,j) : a_{i,j} \neq 0\}$$

is finite. Let $\{e_n\}_n$ denote the standard ONB for $\ell^2(\mathbb{N})$, and set $\mathcal{H}_0 := \text{span}\{e_n\}_n$. Note that we are **not** using *closed* linear spans, and so $\mathcal{H}_0$ is a dense linear submanifold of $\mathcal{H}$, consisting of finitely-supported sequences. Define a linear map $A_0 : \mathcal{H}_0 \to \mathcal{H}_0$ by setting $A_0 e_N = \sum_{i\geq 1} a_{i,N} e_i$ for all $N \geq 1$, and extending by linearity to all of $\mathcal{H}_0$. (Make sure that you know why and how you can do this!)

(a) Prove that A_0 is continuous on $\mathcal{H}_0$.
(b) Let $\mathcal{M}$ be any dense linear submanifold of $\mathcal{H}$, and suppose that $T_0 : \mathcal{M} \to \mathcal{M}$ is a continuous linear map. Show that T_0 **extends** to a **unique** continuous linear map $T \in \mathcal{B}(\mathcal{H})$ with $\|T\| = \|T_0\|$. That is, there exists (a unique) $T \in \mathcal{B}(\mathcal{H})$ with $\|T\| = \|T_0\|$ and $T|_{\mathcal{M}} = T_0$.
(c) Conclude that A_0 extends to a continuous linear map $A \in \mathcal{B}(\mathcal{H})$. What can you say about the range of A?

Question 5.

Let $\mathfrak{J} : c_0 \to \ell^\infty$ denote the canonical embedding of c_0 into $\ell^\infty = (c_0)^{**}$, as defined in Proposition 2.19. Prove that

$$\mathfrak{J}((x_n)_n) = (x_n)_n$$

for all $x = (x_n)_n \in c_0$.

Question 6.

Let $1 < p < \infty$. By Example 2.16, we see that $\ell^q \simeq (\ell^p)^*$, and thus that $\ell^p \simeq (\ell^q)^* \simeq (\ell^p)^{**}$. By identifying ℓ^p with $(\ell^p)^{**}$ under this isomorphism, we may consider the canonical embedding $\mathfrak{J} : \ell^p \to \ell^p$ of ℓ^p into its "second dual", as defined in Proposition 2.19. Prove that $\mathfrak{J}$ is the identity map.

Question 7.

Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$, $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$, and $(\mathfrak{Z}, \|\cdot\|_{\mathfrak{Z}})$ be Banach spaces. Let $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ and $R \in \mathcal{B}(\mathfrak{Y}, \mathfrak{Z})$. Prove that $RT \in \mathcal{B}(\mathfrak{X}, \mathfrak{Z})$ and that $\|RT\| \le \|R\| \, \|T\|$. Do we always have equality? Can we ever have equality?

Question 8.

Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space. An operator $T \in \mathcal{B}(\mathfrak{X})$ is said to be **invertible** if there exists $R \in \mathcal{B}(\mathfrak{X})$ such that $RT = I = TR$, where $I \in \mathcal{B}(\mathfrak{X})$ is the identity map $Ix = x$ for all $x \in \mathfrak{X}$. We say that R is the **inverse** of T, and we typically write T^{-1} instead of R to denote the inverse of T.

Prove that if $T \in \mathcal{B}(\mathfrak{X})$ is invertible, then T is **bounded below**; that is, there exists $\delta > 0$ such that $\|Tx\| \ge \delta \, \|x\|$ for all $x \in \mathfrak{X}$.

Is the converse true? Prove it or find a counterexample to show that it is false.

3. Hilbert space

You should always go to other people's funerals; otherwise, they won't come to yours.

Yogi Berra

3.1. In this brief Chapter, we shall examine a class of very well-behaved Banach spaces, namely the class of **Hilbert spaces**. Hilbert spaces are the generalizations of our familiar (two- and) three-dimensional Euclidean space. There are two basic approaches to studying Hilbert spaces. If one is interested in Banach space geometry – and many people are – then one often tries to compare other Banach spaces to Hilbert spaces. As an example of such a phenomenon, we mention the calculation of the **Banach-Mazur distance** between Banach spaces, which we define in the Appendix to this Section.

In the second approach, one decides that because Hilbert spaces are so well-behaved, they are in some sense "understood", and for this reason they are "less interesting" to study than the set of bounded linear operators acting upon them. One can then study the operators individually or in sets which have no algebraic structure – this kind of analysis belongs to *Single Operator Theory.* Alternatively, one can various *Operator Algebras*, of which there are myriads of examples. The literature dealing with operators and operator algebras is vast.

3.2. Recall that a **Hilbert space** $\mathcal{H}$ is a vector space equipped with an inner product $\langle \cdot, \cdot \rangle$ so that that the induced norm $\|x\| := \langle x, x\rangle^{1/2}$ gives rise to a complete normed linear space, i.e. a Banach space. [When the corresponding normed linear space is not complete, we refer only to **inner product spaces**.]

In any inner product space we have the **Cauchy-Schwarz Inequality**:

$$|\langle x, y\rangle|^2 \le \|x\|^2 \, \|y\|^2,$$

for all $x, y \in \mathcal{H}$. We say that x and y are **orthogonal** if $\langle x, y\rangle = 0$, and we write $x \perp y$.

3.3. Example.

(a) If (X, μ) is a measure space, then $L^2(X, \mu)$ is a Hilbert space, with the inner product given by

$$\langle f, g\rangle = \int_X f(x)\overline{g(x)} d\mu(x).$$

(b) $\ell^2 = \{(x_n)_n : x_n \in \mathbb{K}, n \ge 1 \text{ and } \sum_{n=1}^\infty |x_n|^2 < \infty\}$ is a Hilbert space, with the inner product given by

$$\langle (x_n)_n, (y_n)_n\rangle = \sum_n x_n \overline{y_n}.$$

The reader with a background in measure theory will recognize that the second example is merely a particular case of the first. While these are the canonical inner products on these spaces, they are not the only ones.

For example, if $(r_n)_n$ is any sequence of strictly positive integers, one can define a **weighted ℓ^2 space** relative to this sequence by setting

$$\ell^2_{(r_n)_n} := \{(x_n)_n \in \mathbb{K}^{\mathbb{N}} : \sum_n r_n |x_n|^2 < \infty\}$$

with inner product

$$\langle (x_n)_n, (y_n)_n \rangle = \sum_n r_n x_n \overline{y_n}.$$

3.4. Theorem. *Let $\mathcal{H}$ be a Hilbert space and suppose that $x_1, x_2, ..., x_n \in \mathcal{H}$.*

(a) [***The Pythagorean Theorem***] *If the vectors are pairwise orthogonal, then*

$$\|\sum_{j=1}^n x_j\|^2 = \sum_{j=1}^n \|x_j\|^2.$$

(b) [***The Parallelogram Law***]

$$\|x_1 + x_2\|^2 + \|x_1 - x_2\|^2 = 2\left(\|x_1\|^2 + \|x_2\|^2\right).$$

Proof. Both of these results follow immediately from the definition of the norm in terms of the inner product.

□

The Parallelogram Law is a useful tool to show that many norms are *not* Hilbert space norms.

3.5. Theorem. *Let $\mathcal{H}$ be a Hilbert space, and $K \subseteq \mathcal{H}$ be a closed, non-empty convex subset of $\mathcal{H}$. Given $x \in \mathcal{H}$, there exists a unique point $y \in K$ which is closest to x; that is,*

$$\|x - y\| = \operatorname{dist}(x, K) = \min\{\|x - z\| : z \in K\}.$$

Proof. By translating K by $-x$, it suffices to consider the case where $x = 0$.

Let $d := \operatorname{dist}(0, K)$, and choose $k_n \in K$ so that $\|0 - k_n\| < d + \frac{1}{n}$, $n \geq 1$. By the Parallelogram Law,

$$\begin{aligned} \|\frac{k_n - k_m}{2}\|^2 &= \frac{1}{2}\|k_n\|^2 + \frac{1}{2}\|k_m\|^2 - \|\frac{k_n + k_m}{2}\|^2 \\ &\leq \frac{1}{2}(d + \frac{1}{n})^2 + \frac{1}{2}(d + \frac{1}{m})^2 - d^2, \end{aligned}$$

as $\dfrac{k_n + k_m}{2} \in K$ because K is assumed to be convex.

We deduce from this that the sequence $\{k_n\}_{n=1}^\infty$ is Cauchy, and hence converges to some $k \in K$, since K is closed and $\mathcal{H}$ is complete. Clearly $\lim_{n\to\infty} k_n = k$ implies that $d = \lim_{n\to\infty} \|k_n\| = \|k\|$.

As for uniqueness, suppose that $z \in K$ and that $\|z\| = d$. Then

$$\begin{aligned} 0 \le \|\frac{k-z}{2}\|^2 &= \frac{1}{2}\|k\|^2 + \frac{1}{2}\|z\|^2 - \|\frac{k+z}{2}\|^2 \\ &\le \frac{1}{2}d^2 + \frac{1}{2}d^2 - d^2 = 0, \end{aligned}$$

and so $k = z$.

□

3.6. Theorem. *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Let $x \in \mathcal{H}$, and $m \in \mathcal{M}$. The following are equivalent:*

(a) $\|x - m\| = \operatorname{dist}(x, \mathcal{M})$;

(b) *The vector $x - m$ is orthogonal to $\mathcal{M}$, i.e., $\langle x - m, y\rangle = 0$ for all $y \in \mathcal{M}$.*

Proof.

(a) implies (b): Suppose that $\|x - m\| = \operatorname{dist}(x, \mathcal{M})$, and suppose to the contrary that there exists $y \in \mathcal{M}$ so that $k := \langle x - m, y\rangle \ne 0$. There is no loss of generality in assuming that $\|y\| = 1$. Consider $z = m + ky \in \mathcal{M}$. Then

$$\begin{aligned} \|x - z\|^2 &= \|x - m - ky\|^2 \\ &= \langle x - m - ky, x - m - ky\rangle \\ &= \|x - m\|^2 - k\langle y, x - m\rangle - \overline{k}\langle x - m, y\rangle + |k|^2 \, \|y\|^2 \\ &= \|x - m\|^2 - |k|^2 \\ &< \operatorname{dist}(x, \mathcal{M}), \end{aligned}$$

a contradiction. Hence $x - m \in \mathcal{M}^\perp$.

(b) implies (a): Suppose that $x - m \in \mathcal{M}^\perp$. If $z \in \mathcal{M}$ is arbitrary, then $y := z - m \in \mathcal{M}$, so by the Pythagorean Theorem,

$$\|x - z\|^2 = \|(x - m) - y\|^2 = \|x - m\|^2 + \|y\|^2 \ge \|x - m\|^2,$$

and thus $\operatorname{dist}(x, \mathcal{M}) \ge \|x - m\|$. Since the other inequality is obvious, (a) holds.

□

3.7. Remarks.

(a) Given any non-empty subset $\mathcal{S} \subseteq \mathcal{H}$, let

$$\mathcal{S}^\perp := \{y \in \mathcal{H} : \langle x, y\rangle = 0 \text{ for all } x \in \mathcal{S}\}.$$

It is routine to show that $\mathcal{S}^\perp$ is a norm-closed subspace of $\mathcal{H}$. In particular,

$$\left(\mathcal{S}^\perp\right)^\perp \supseteq \overline{\operatorname{span}}\, \mathcal{S},$$

the norm closure of the linear span of $\mathcal{S}$.

(b) Recall from Linear Algebra that if $\mathcal{V}$ is a vector space and $\mathcal{W}$ is a (vector) subspace of $\mathcal{V}$, then there exists a (vector) subspace $\mathcal{X} \subseteq \mathcal{V}$ such that

(i) $\mathcal{W} \cap \mathcal{X} = \{0\}$, and
(ii) $\mathcal{V} = \mathcal{W} + \mathcal{X} := \{w + x : w \in \mathcal{W}, x \in \mathcal{X}\}$.

We say that $\mathcal{W}$ is **algebraically complemented** by $\mathcal{X}$. The existence of such a $\mathcal{X}$ for each $\mathcal{W}$ says that every vector subspace of a vector space is algebraically complemented. We shall write $\mathcal{V} = \mathcal{W} \dotplus \mathcal{X}$ to denote the fact that $\mathcal{X}$ is an algebraic complement for $\mathcal{W}$ in $\mathcal{V}$.

If $\mathfrak{X}$ is a Banach space and $\mathfrak{Y}$ is a closed subspace of $\mathfrak{X}$, we say that $\mathfrak{Y}$ is **topologically complemented** if there exists a *closed* subspace $\mathfrak{Z}$ of $\mathfrak{X}$ such that $\mathfrak{Z}$ is an algebraic complement to $\mathfrak{Y}$. The issue here is that both $\mathfrak{Y}$ and $\mathfrak{Z}$ must be closed subspaces. It can be shown that the closed subspace c_0 of ℓ^∞ is *not* topologically complemented in ℓ^∞. This result is known as Phillips' Theorem (see the paper of R. Whitley [**Whi66**] for a short but elegant proof). We shall write $\mathfrak{X} = \mathfrak{Y} \oplus \mathfrak{Z}$ if $\mathfrak{Z}$ is a topological complement to $\mathfrak{Y}$ in $\mathfrak{X}$.

Now let $\mathcal{H}$ be a Hilbert space and let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace of $\mathcal{H}$. We claim that $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Indeed, if $z \in \mathcal{M} \cap \mathcal{M}^\perp$, then $\|z\|^2 = \langle z, z\rangle = 0$, so $z = 0$. Also, if $x \in \mathcal{H}$, then we may let $m_1 \in \mathcal{M}$ be the element satisfying

$$\|x - m_1\| = \mathrm{dist}(x, \mathcal{M}).$$

The existence of m_1 is guaranteed by Theorem 3.5. By Theorem 3.6, $m_2 := x - m_1$ lies in $\mathcal{M}^\perp$, and so $x \in \mathcal{M} + \mathcal{M}^\perp$. Since $\mathcal{M}$ and $\mathcal{M}^\perp$ are closed subspaces of a Banach space and they are algebraically complements, we are done.

In this case, the situation is even stronger. The space $\mathcal{M}$ may admit more than one topological complement in $\mathcal{H}$ - however, the space $\mathcal{M}^\perp$ above is unique in that it is an **orthogonal complement**. That is, as well as being a topological complement to $\mathcal{M}$, every vector in $\mathcal{M}^\perp$ is orthogonal to every vector in $\mathcal{M}$.

(c) With $\mathcal{M}$ as in (b), we have $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, so that if $x \in \mathcal{H}$, then we may write $x = m_1 + m_2$ with $m_1 \in \mathcal{M}$, $m_2 \in \mathcal{M}^\perp$ in a *unique* way. Consider the map:

$$\begin{array}{rccc} P: & \mathcal{H} & \to & \mathcal{M} \oplus \mathcal{M}^\perp \\ & x & \mapsto & m_1, \end{array}$$

relative to the above decomposition of x. It is elementary to verify that P is linear, and that P is **idempotent**, i.e., $P = P^2$. We remark in passing that $m_2 = (I - P)x$, and that $(I - P)^2 = (I - P)$ as well.

In fact, for $x \in \mathcal{H}$, $\|x\|^2 = \|m_1\|^2 + \|m_2\|^2$ by the Pythagorean Theorem, and so $\|Px\| = \|m_1\| \le \|x\|$, from which it follows that $\|P\| \le 1$. If $\mathcal{M} \ne \{0\}$, then choose $m \in \mathcal{M}$ with $\|m\| \ne 0$. Then $Pm = m$, and so $\|P\| \ge 1$. Combining these estimates, $\mathcal{M} \ne 0$ implies $\|P\| = 1$.

We refer to the map P as the **orthogonal projection** of $\mathcal{H}$ onto $\mathcal{M}$. The map $Q := (I - P)$ is the orthogonal projection onto $\mathcal{M}^\perp$, and we leave it to the reader to verify that if $\mathcal{M} \ne \mathcal{H}$, then $\|Q\| = 1$.

(d) Let $\varnothing \neq \mathcal{S} \subseteq \mathcal{H}$. We saw in (a) that $\mathcal{S}^{\perp\perp} \supseteq \overline{\operatorname{span}}\,\mathcal{S}$. In fact, if we let $\mathcal{M} = \overline{\operatorname{span}}\,\mathcal{S}$, then $\mathcal{M}$ is a closed subspace of $\mathcal{H}$, and so by (b),

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}.$$

It is routine to check that $\mathcal{S}^{\perp} = \mathcal{M}^{\perp}$. Suppose that there exists $0 \neq x \in \mathcal{S}^{\perp\perp}$, $x \notin \mathcal{M}$. Then $x \in \mathcal{H}$, and so we can write $x = m_1 + m_2$ with $m_1 \in \mathcal{M}$, and $m_2 \in \mathcal{M}^{\perp} = \mathcal{S}^{\perp}$ ($m_2 \neq 0$, otherwise $x \in \mathcal{M}$). But then $0 \neq m_2 \in \mathcal{S}^{\perp}$ and so

$$\begin{aligned} \langle m_2, x\rangle &= \langle m_2, m_1\rangle + \langle m_2, m_2\rangle \\ &= 0 + \|m_2\|^2 \\ &\neq 0. \end{aligned}$$

This contradicts the fact that $x \in \mathcal{S}^{\perp\perp}$. It follows that $\mathcal{S}^{\perp\perp} = \overline{\operatorname{span}}\,\mathcal{S}$.

(e) Suppose that $\mathcal{M}$ admits an orthonormal basis $\{e_k\}_{k=1}^n$. Let $x \in \mathcal{H}$, and let P denote the orthogonal projection onto $\mathcal{M}$. By (b), Px is the unique element of $\mathcal{M}$ so that $x - Px$ lies in $\mathcal{M}^{\perp}$. Consider the vector $w = \sum_{k=1}^n \langle x, e_k\rangle e_k$. Then

$$\begin{aligned} \langle x - w, e_j\rangle &= \langle x, e_j\rangle - \sum_{k=1}^n \langle \langle x, e_k\rangle e_k, e_j\rangle \\ &= \langle x, e_j\rangle - \sum_{k=1}^n \langle x, e_k\rangle \langle e_k, e_j\rangle \\ &= \langle x, e_j\rangle - \langle x, e_j\rangle \|e_j\|^2 \\ &= 0. \end{aligned}$$

It follows that $x - w \in \mathcal{M}^{\perp}$, and thus $w = Px$. That is, $Px = \sum_{k=1}^n \langle x, e_k\rangle e_k$.

3.8. Theorem. ***The Riesz Representation Theorem.*** *Let $\{0\} \neq \mathcal{H}$ be a Hilbert space over $\mathbb{K}$, and let $\varphi \in \mathcal{H}^*$. Then there exists a unique vector $y \in \mathcal{H}$ so that*

$$\varphi(x) = \langle x, y\rangle \qquad \text{for all } x \in \mathcal{H}.$$

Moreover, $\|\varphi\| = \|y\|$.

Proof. Given a fixed $y \in \mathcal{H}$, let us denote by β_y the map $\beta_y(x) = \langle x, y\rangle$. Our goal is to show that $\mathcal{H}^* = \{\beta_y : y \in \mathcal{H}\}$. First note that if $y \in \mathcal{H}$, then $\beta_y(kx_1 + x_2) = \langle kx_1 + x_2, y\rangle = k\langle x_1, y\rangle + \langle x_2, y\rangle = k\beta_y(x_1) + \beta_y(x_2)$, and so β_y is linear. Furthermore, for each $x \in \mathcal{H}$, $|\beta_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$ by the Cauchy-Schwarz Inequality. Thus $\|\beta_y\| \leq \|y\|$, and hence β_y is continuous - i.e. $\beta_y \in \mathcal{H}^*$.

It is not hard to verify that the map

$$\begin{aligned} \Theta : \ \mathcal{H} &\to \mathcal{H}^* \\ y &\mapsto \beta_y \end{aligned}$$

is conjugate-linear (if $\mathbb{K} = \mathbb{C}$), otherwise it is linear (if $\mathbb{K} = \mathbb{R}$). From the first paragraph, it is also contractive. But $[\Theta(y)](y) = \beta_y(y) = \langle y, y\rangle = \|y\|^2$, so that

$\|\Theta(y)\| \geq \|y\|$ for all $y \in \mathcal{H}$, and Θ is isometric as well. It immediately follows that Θ is injective, and there remains only to prove that Θ is surjective.

Let $\varphi \in \mathcal{H}^*$. If $\varphi = 0$, then $\varphi = \Theta(0)$. Otherwise, let $\mathcal{M} = \ker \varphi$, so that $\operatorname{codim} \mathcal{M} = 1 = \dim \mathcal{M}^\perp$, since $\mathcal{H}/\mathcal{M} \simeq \mathbb{K} \simeq \mathcal{M}^\perp$. Choose $e \in \mathcal{M}^\perp$ with $\|e\| = 1$.

Let P denote the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$, constructed as in Remark 3.7. Then, as $I - P$ is the orthogonal projection onto $\mathcal{M}^\perp$, and as $\{e\}$ is an orthonormal basis for $\mathcal{M}^\perp$, by Remark 3.7 (d), for all $x \in \mathcal{H}$, we have

$$x = Px + (I - P)x = Px + \langle x, e\rangle e.$$

Thus for all $x \in \mathcal{H}$,

$$\varphi(x) = \varphi(Px) + \langle x, e\rangle \varphi(e) = \langle x, \overline{\varphi(e)} e\rangle = \beta_y(x),$$

where $y := \overline{\varphi(e)} e$. Hence $\varphi = \beta_y$, and Θ is onto.

□

3.9. Remark. The fact that the map Θ defined in the proof the Riesz Representation Theorem above induces an isometric, conjugate-linear isomorphism of $\mathcal{H}$ with $\mathcal{H}^*$ is worth remembering.

3.10. Definition. *A subset $\{e_\lambda\}_{\lambda \in \Lambda}$ of a Hilbert space $\mathcal{H}$ is said to be **orthonormal** if $\|e_\lambda\| = 1$ for all λ, and $\lambda \neq \alpha$ implies that $\langle e_\lambda, e_\alpha\rangle = 0$.*

*An orthonormal set in a Hilbert space is called an **orthonormal basis** for $\mathcal{H}$ if it is maximal in the collection of all orthonormal sets of $\mathcal{H}$, partially ordered with respect to inclusion.*

If $E = \{e_\lambda\}_\lambda$ is any orthonormal set in $\mathcal{H}$, then an easy application of Zorn's Lemma implies the existence of a orthonormal basis in $\mathcal{H}$ which contains E. The reader is warned that if $\mathcal{H}$ is infinite-dimensional, then an orthonormal basis for $\mathcal{H}$ is never a vector space (i.e. a Hamel) basis for $\mathcal{H}$.

3.11. Example.

(a) If $\mathcal{H} = \ell^2$ then the standard Schauder basis $\{e_n\}_{n=1}^\infty$ for ℓ^2 as defined in Example 2.14 is an orthonormal basis for $\mathcal{H}$.

(b) If $\mathcal{H} = L^2(\mathbb{T}, dm)$, where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ and dm represents normalised Lebesgue measure, then $\{f_n\}_{n \in \mathbb{Z}}$ is an orthonormal basis for $L^2(\mathbb{T}, dm)$, where $f_n(z) = z^n$ for all $z \in \mathbb{T}$ and for all $n \in \mathbb{Z}$.

We recall from Linear Algebra:

3.12. Theorem. ***The Gram-Schmidt Orthogonalisation Process***
If $\mathcal{H}$ is a Hilbert space over $\mathbb{K}$ and $\{x_n\}_{n=1}^\infty$ is a linearly independent set in $\mathcal{H}$, then we can find an orthonormal set $\{y_n\}_{n=1}^\infty$ in $\mathcal{H}$ so that $\text{span}\{x_1, x_2, ..., x_k\} = \text{span}\,\{y_1, y_2, ..., y_k\}$ *for all $k \ge 1$.*
Proof. We leave it to the reader to verify that setting $y_1 = x_1/\|x_1\|$, and recursively defining

$$y_k := \frac{x_k - \sum_{j=1}^{k-1} \langle x_k, y_j \rangle y_j}{\|x_k - \sum_{j=1}^{k-1} \langle x_k, y_j \rangle y_j\|}, \qquad k \ge 2$$

will do.

□

3.13. Theorem. ***Bessel's Inequality***
If $\{e_n\}_{n=1}^\infty$ is an orthonormal set in a Hilbert space $\mathcal{H}$, then for each $x \in \mathcal{H}$,

$$\sum_{n=1}^\infty |\langle x, e_n \rangle|^2 \le \|x\|^2.$$

Proof. For each $k \ge 1$, let P_k denote the orthogonal projection of $\mathcal{H}$ onto $\text{span}\,\{e_n\}_{n=1}^k$. Given $x \in \mathcal{H}$, we have seen that $\|P_k\| \le 1$, and that $P_k x = \sum_{n=1}^k \langle x, e_n \rangle e_n$. Hence $\|x\|^2 \ge \|P_k x\|^2 = \sum_{n=1}^k |\langle x, e_n \rangle|^2$ for all $k \ge 1$, from which the result follows.

□

3.14. Before considering a non-separable version of the above result, we pause to define what we mean by a sum over an uncountable index set.

Given a Banach space $\mathfrak{X}$ and a set $\{x_\lambda\}_{\lambda \in \Lambda}$ of vectors in $\mathfrak{X}$, let $\mathcal{F}$ denote the collection of all finite subsets of Λ, partially ordered by inclusion. For each $F \in \mathcal{F}$, define $y_F = \sum_{\lambda \in F} x_\lambda$, so that $(y_F)_{F \in \mathcal{F}}$ is a net in $\mathfrak{X}$. If $y = \lim_{F \in \mathcal{F}} y_F$ exists, then we write $y = \sum_{\lambda \in \Lambda} x_\lambda$, and we say that $\{x_\lambda\}_\lambda$ is **unconditionally summable** to y, and that the series $\sum_{\lambda \in \Lambda} x_\lambda$ is **unconditionally convergent**.

In other words, $\sum_\lambda x_\lambda = y$ if for all $\varepsilon > 0$ there exists $F_0 \in \mathcal{F}$ so that $F \in \mathcal{F}$ and $F \supseteq F_0$ implies that $\| \sum_{\lambda \in F} x_\lambda - y\| < \varepsilon$.

3.15. Corollary. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{E} \subseteq \mathcal{H}$ be an orthonormal set.*
(a) *Given $x \in \mathcal{H}$, the set $\{e \in \mathcal{E} : \langle x, e \rangle \ne 0\}$ is countable.*
(b) *For all $x \in \mathcal{H}$, $\sum_{e \in \mathcal{E}} |\langle x, e \rangle|^2 \le \|x\|^2$.*

Proof.
(a) Fix $x \in \mathcal{H}$. For each $k \ge 1$, define $\mathcal{F}_k = \{e \in \mathcal{E} : |\langle x, e \rangle| \ge \frac{1}{k}\}$. Suppose that there exists $k_0 \ge 1$ so that $\mathcal{F}_{k_0}$ is infinite. Choose a countably infinite subset $\{e_n\}_{n=1}^\infty$ of $\mathcal{F}_{k_0}$. By Bessel's Inequality,

$$\frac{m}{k_0^2} = \sum_{n=1}^m \frac{1}{k_0^2} \le \sum_{n=1}^m |\langle x, e_n \rangle|^2 \le \|x\|^2$$

for all $m \ge 1$. This is absurd. Thus $|\mathcal{F}_k| < \infty$ for all $k \ge 1$. But then

$$\{e \in \mathcal{E} : \langle x, e\rangle \ne 0\} = \cup_{k\ge 1}\mathcal{F}_k$$

is countable.

(b) This is left as a (routine) exercise for the reader.

□

3.16. Lemma. *Let $\mathcal{H}$ be a Hilbert space, $\mathcal{E} \subseteq \mathcal{H}$ be an orthonormal set, and $x \in \mathcal{H}$. Then*

$$\sum_{e\in\mathcal{E}} \langle x, e\rangle e$$

converges in $\mathcal{H}$.

Proof. Since $\mathcal{H}$ is complete, it suffices to show that if $\mathcal{F}$ as in Paragraph 3.14 denotes the collection of finite subsets of $\mathcal{E}$, partially ordered by inclusion, and if for each $F \in \mathcal{F}$, $y_F = \sum_{e\in F}\langle x, e\rangle e$, then $(y_F)_{F\in\mathcal{F}}$ is a Cauchy net.

Let $\varepsilon > 0$. By Corollary 3.15, we can find a countable subcollection $\{e_n\}_{n=1}^\infty \subseteq \mathcal{E}$ so that $e \in \mathcal{E}\backslash\{e_n\}_{n=1}^\infty$ implies that $\langle x, e\rangle = 0$. Moreover, by Bessel's Inequality, we can find $N > 0$ so that $\sum_{k=N+1}^\infty |\langle x, e_k\rangle|^2 < \varepsilon^2$. Let $F_0 = \{e_1, e_2, ..., e_N\}$.

If $F, G \in \mathcal{F}$ and $F_0 \le F$, $F_0 \le G$, then

$$\begin{aligned}\|y_F - y_G\|^2 &= \|\sum_{e\in F\backslash G}\langle x, e\rangle e - \sum_{e\in G\backslash F}\langle x, e\rangle e\|^2 \\ &= \sum_{e\in(F\backslash G)\cup(G\backslash F)} |\langle x, e\rangle|^2 \\ &\le \sum_{k=N+1}^\infty |\langle x, e_k\rangle|^2 \\ &< \varepsilon^2.\end{aligned}$$

This shows that $(y_F)_{F\in\mathcal{F}}$ is a Cauchy net, and therefore it converges, as required. (For a proof that Cauchy nets in a complete metric space converge, see Proposition 3.29 in the Appendix to this Section.)

□

3.17. Theorem. *Let $\mathcal{E}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. The following are equivalent:*

(a) *The set $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$. (That is, $\mathcal{E}$ is a maximal orthonormal set in $\mathcal{H}$.)*
(b) *The set* $\operatorname{span}\mathcal{E}$ *is norm-dense in $\mathcal{H}$.*
(c) *For all $x \in \mathcal{H}$, $x = \sum_{e\in\mathcal{E}}\langle x, e\rangle e$.*
(d) *For all $x \in \mathcal{H}$, $\|x\|^2 = \sum_{e\in\mathcal{E}} |\langle x, e\rangle|^2$.* [***Parceval's Identity***]

Proof. Let $\mathcal{E} = \{e_\lambda\}_{\lambda\in\Lambda}$ be an orthonormal set in $\mathcal{H}$.

(a) implies (b): Let $\mathcal{M} = \overline{\operatorname{span}}\,\mathcal{E}$. If $\mathcal{M} \neq \mathcal{H}$, then $\mathcal{M}^\perp \neq \{0\}$, so we can find $z \in \mathcal{M}^\perp$, $\|z\| = 1$. But then $\mathcal{E} \cup \{z\}$ is an orthonormal set, contradicting the maximality of $\mathcal{E}$.

(b) implies (c): Let $y = \sum_{\lambda\in\Lambda}\langle x, e_\lambda\rangle e_\lambda$, which exists by Lemma 3.16. Then $\langle y - x, e_\lambda\rangle = 0$ for all $\lambda \in \Lambda$, so $y - x$ is orthogonal to $\mathcal{M} = \overline{\operatorname{span}}\,\mathcal{E} = \mathcal{H}$. But then $y - x \perp y - x$, so that $y - x = 0$, i.e. $y = x$.

(c) implies (d): $\|x\|^2 = \langle \sum_{e\in\mathcal{E}}\langle x, e\rangle e, \sum_{f\in\mathcal{E}}\langle x, f\rangle f\rangle = \sum_{e\in\mathcal{E}} |\langle x, e\rangle|^2$. [Check!]

(d) implies (a): If $e \perp e_\lambda$ for all $\lambda \in \Lambda$, then

$$\|e\|^2 = \sum_{\lambda\in\Lambda} |\langle e, e_\lambda\rangle|^2 = 0,$$

so that $\mathcal{E}$ is maximal.

□

3.18. Proposition. *If $\mathcal{H}$ is a Hilbert space, then any two orthonormal bases for $\mathcal{H}$ have the same cardinality.*

Proof. We shall only deal with the infinite-dimensional situation, since the finite-dimensional case was dealt with in linear algebra.

Let $\mathcal{E}$ and $\mathcal{F}$ be two orthonormal bases for $\mathcal{H}$. Given $e \in \mathcal{E}$, let $\mathcal{F}_e = \{f \in \mathcal{F} : \langle e, f\rangle \neq 0\}$. Then $|\mathcal{F}_e| \leq \aleph_0$. Moreover, given $f \in \mathcal{F}$, there exists $e \in \mathcal{E}$ so that $\langle e, f\rangle \neq 0$, otherwise f is orthogonal to $\overline{\operatorname{span}}\,\mathcal{E} = \mathcal{H}$, a contradiction.

Thus $\mathcal{F} = \cup_e \mathcal{F}_e$, and so $|\mathcal{F}| \leq (\sup_{e\in\mathcal{E}} |\mathcal{F}_e|)\,|\mathcal{E}| \leq \aleph_0|\mathcal{E}| = |\mathcal{E}|$. By symmetry, $|\mathcal{E}| \leq |\mathcal{F}|$, and so $|\mathcal{E}| = |\mathcal{F}|$.

□

The above result justifies the following definition:

3.19. Definition. *The **dimension** of a Hilbert space $\mathcal{H}$ is the cardinality of any orthonormal basis for $\mathcal{H}$, and it is denoted by* $\dim \mathcal{H}$.

The appropriate notion of isomorphism in the category of Hilbert spaces involve linear maps that preserve the inner product.

3.20. Definition. *Two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are said to be **isomorphic** if there exists a linear bijection $U : \mathcal{H}_1 \to \mathcal{H}_2$ so that*

$$\langle Ux, Uy\rangle = \langle x, y\rangle$$

for all $x, y \in \mathcal{H}_1$. We write $\mathcal{H}_1 \simeq \mathcal{H}_2$ to denote this isomorphism.

We also refer to the linear maps implementing the above isomorphism as **unitary operators**. Note that

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, x\rangle = \|x\|^2$$

for all $x \in \mathcal{H}_1$, so that unitary operators are isometries. Moreover, the inverse map $U^{-1} : \mathcal{H}_2 \to \mathcal{H}_1$ defined by $U^{-1}(Ux) := x$ is also linear, and

$$\langle U^{-1}(Ux) U^{-1}(Uy)\rangle = \langle x, y\rangle = \langle Ux, Uy\rangle,$$

so that U^{-1} is also a unitary operator. Furthermore, if $\mathcal{L} \subseteq \mathcal{H}_1$ is a closed subspace, then $\mathcal{L}$ is complete, whence $U\mathcal{L}$ is also complete and hence closed in $\mathcal{H}_2$.

Unlike the situation in Banach spaces, where two non-isomorphic Banach spaces can have Schauder bases of the same cardinality, the case of Hilbert spaces is as nice as one can imagine.

3.21. Theorem. *Two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ over $\mathbb{K}$ are isomorphic if and only if they have the same dimension.*

Proof. Suppose first that $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic, and let $U : \mathcal{H}_1 \to \mathcal{H}_2$ be a unitary operator. Let $\{e_\lambda\}_{\lambda\in\Lambda}$ be an orthonormal basis for $\mathcal{H}_1$. We claim that $\{Ue_\lambda\}_{\lambda\in\Lambda}$ is an orthonormal basis for $\mathcal{H}_2$.

Indeed, $\langle Ue_\alpha, Ue_\beta\rangle = \langle e_\alpha, e_\beta\rangle = \delta_{\alpha,\beta}$ (the Kronecker delta function)), $\|Ue_\alpha\| = \|e_\alpha\| = 1$ for all $\alpha \in \Lambda$, and

$$\mathcal{H}_2 = U\mathcal{H}_1 = U(\overline{\text{span}}\,\{e_\lambda\}_{\lambda\in\Lambda}) = \overline{\text{span}}\,\{Ue_\lambda\}_{\lambda\in\Lambda},$$

by the continuity of U.

Hence dim $\mathcal{H}_2 = |\{Ue_\lambda\}_{\lambda\in\Lambda}| = |\Lambda| =$ dim $\mathcal{H}_1$.

Conversely, suppose that dim $\mathcal{H}_2 =$ dim $\mathcal{H}_1$. Then we can find a set Λ and orthonormal bases $\{e_\lambda : \lambda \in \Lambda\}$ for $\mathcal{H}_1$, and $\{f_\lambda : \lambda \in \Lambda\}$ for $\mathcal{H}_2$. Consider the map

$$\begin{array}{rccc} U : & \mathcal{H}_1 & \to & \ell^2(\Lambda) \\ & x & \mapsto & (\langle x, e_\lambda\rangle)_{\lambda\in\Lambda}, \end{array}$$

where $\ell^2(\Lambda) := \left\{f : \Lambda \to \mathbb{K} : \sum_{\lambda\in\Lambda} |f(\lambda)|^2 < \infty\right\}$. This is an inner product space using the inner product $\langle f, g\rangle = \sum_{\lambda\in\Lambda} f(\lambda)\overline{g(\lambda)}$. The proof that $\ell^2(\Lambda)$ is complete is essentially the same as in the case of $\ell^2(\mathbb{N})$.

It is routine to check that U is linear. Moreover,

$$\begin{aligned} \langle Ux, Uy\rangle &= \sum_\lambda \langle x, e_\lambda\rangle\, \overline{\langle y, e_\lambda\rangle} \\ &= \sum_\lambda \langle\langle x, e_\lambda\rangle e_\lambda, \langle y, e_\lambda\rangle e_\lambda\rangle \\ &= \langle \sum_\lambda \langle x, e_\lambda\rangle e_\lambda, \sum_\gamma \langle y, e_\gamma\rangle e_\gamma\rangle \\ &= \langle x, y\rangle \end{aligned}$$

for all $x, y \in \mathcal{H}$, and so $\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, x\rangle = \|x\|^2$ for all $x \in \mathcal{H}_1$. It follows that U is isometric and therefore injective.

If $(r_\lambda)_\lambda \in \ell^2(\Lambda)$ has finite support, then $x := \sum_{\lambda\in\Lambda} r_\lambda e_\lambda \in \mathcal{H}_1$ and $Ux = (r_\lambda)_\lambda$. Thus ran U is dense. But from the comment above, $U\mathcal{H}_1$ is closed, and therefore $U\mathcal{H}_1 = \ell^2(\Lambda)$.

We have shown that U is a unitary operator implementing the isomorphism of $\mathcal{H}_1$ and $\ell^2(\Lambda)$. By symmetry once again, there exists a unitary $V : \mathcal{H}_2 \to \ell^2(\Lambda)$. But then $V^{-1}U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary, showing that $\mathcal{H}_1 \simeq \mathcal{H}_2$.

□

3.22. Corollary. *The spaces* $\ell^2(\mathbb{N})$, $\ell^2(\mathbb{Q})$, $\ell^2(\mathbb{Z})$ *and* $L^2([0,1], dx)$ *(where* dx *represents Lebesgue measure) are all isomorphic, as they are all infinite dimensional, separable Hilbert spaces.*

Appendix to Section 3.

3.23. When dealing with Hilbert space operators and operator algebras, one tends to focus upon *complex* Hilbert spaces. One reason for this is that the *spectrum* provides a terribly useful tool for analyzing operators. For $\mathfrak{X}$ a normed linear space, let I (or $I_{\mathfrak{X}}$ if we wish to emphasize the underlying space) denote the identity map $Ix = x$ for all $x \in \mathfrak{X}$.

3.24. Definition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and let $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. We say that T is* ***invertible*** *if there exists a (continuous) linear map $R \in \mathcal{B}(\mathfrak{Y}, \mathfrak{X})$ such that $RT = I_{\mathfrak{X}}$ and $TR = I_{\mathfrak{Y}}$.*

If $T \in \mathcal{B}(\mathfrak{X})$, we define the ***spectrum*** *of T to be:*

$$\sigma(T) = \{\lambda \in \mathbb{K} : (T - \lambda I) \text{ is not invertible}\}.$$

3.25. When $\mathfrak{X}$ is a finite-dimensional space, the spectrum of T coincides with the eigenvalues of T. The reader will recall from their Linear Algebra courses that eigenvalues of linear maps need not exist when the underlying field is not algebraically closed. When $\mathbb{K} = \mathbb{C}$, it can be shown that the spectrum of an operator $T \in \mathcal{B}(\mathfrak{X})$ is a non-empty, compact subset of $\mathbb{C}$, and a so-called **functional calculus** which allows one to naturally define $f(T)$ for any complex-valued function which is analytic in an open neighbourhood of $\sigma(T)$. This, however, is beyond the scope of the present notes.

If $\mathfrak{X}$ is a Banach space, an operator $Q \in \mathcal{B}(\mathfrak{X})$ is said to be **quasinilpotent** if $\sigma(Q) = \{0\}$. The argument of Paragraph 2.21 says that the Volterra operator is quasinilpotent.

A wonderful theorem of **A. Beurling**, known as **Beurling's spectral radius formula** relates the spectrum of an operator T to a limit of the kind obtained in Paragraph 2.21.

Theorem. ***Beurling's Spectral Radius Formula.***
Let $\mathfrak{X}$ be a complex Banach space and $T \in \mathcal{B}(\mathfrak{X})$. Then

$$\text{spr}(T) := \max\{|k| : k \in \sigma(T)\} = \lim_n \|T^n\|^{\frac{1}{n}}.$$

The quantity $\text{spr}(T)$ is known as the **spectral radius** of T. It is worth pointing out that an implication of Beurling's Spectral Radius Formula is that the limit on the right-hand side of the equation exists! *A priori*, this is not obvious.

3.26. As mentioned in paragraph 3.1, Hilbert spaces arise naturally in the study of Banach space geometry. In this context, much of the literature concerns *real* Hilbert spaces.

For example, for each $n \geq 1$, let $\mathcal{Q}_n$ denote the set of n-dimensional (real) Banach spaces. Given Banach spaces $\mathfrak{X}$ and $\mathfrak{Y}$ in $\mathcal{Q}_n$, we denote by $GL(\mathfrak{X}, \mathfrak{Y})$ the set of all invertible operators from $\mathfrak{X}$ to $\mathfrak{Y}$. We can define a metric δ on $\mathcal{Q}_n$ via:

$$\delta(\mathfrak{X}, \mathfrak{Y}) := \log\left(\inf\{\|T\|\,\|T^{-1}\| : T \in GL(\mathfrak{X}, \mathfrak{Y})\}\right).$$

It can be shown that $(\mathcal{Q}_n, \delta)$ is a compact metric space, known as the **Banach-Mazur compactum**. One also refers to the quantity

$$d(\mathfrak{X}, \mathfrak{Y}) = \inf\{\|T\|\,\|T^{-1}\| : T \in GL(\mathfrak{X}, \mathfrak{Y})\}$$

as the **Banach-Mazur distance** between $\mathfrak{X}$ and $\mathfrak{Y}$, and it is an important problem in Banach space geometry to calculate Banach-Mazur distances between the n-dimensional subspaces of two infinite-dimensional Banach spaces, say $\mathfrak{V}$ and $\mathfrak{W}$. Typically, one is interested in knowing something about the asymptotic behaviour of these distances as n tends to infinity.

We mention without proof two interesting facts concerning the Banach-Mazur distance:

(a) If $\mathfrak{X}$, $\mathfrak{Y}$ and $\mathfrak{Z}$ are n-dimensional Banach spaces, then

$$d(\mathfrak{X}, \mathfrak{Z}) \leq d(\mathfrak{X}, \mathfrak{Y})\; d(\mathfrak{Y}, \mathfrak{Z}).$$

(b) A Theorem of Fritz John shows that (with $\ell^2_n := (\mathbb{R}^n, \|\cdot\|_2)$),

$$d(\mathfrak{X}, \ell^2_n) \leq \sqrt{n} \text{ for all } n \geq 1.$$

It clearly follows from these two properties that $d(\mathfrak{X}, \mathfrak{Y}) \leq n$ for all $\mathfrak{X}, \mathfrak{Y} \in \mathcal{Q}_n$.

3.27. We mentioned earlier in this section that the Parallelogram Law is useful in determining that a given norm is not induced by an inner product. In fact, in can be shown that a norm on a Banach space $\mathfrak{X}$ is the norm induced by *some* inner product if and only if the norm satisfies the Parallelogram Law.

3.28. In Lemma 3.16, we invoked the claim that a Cauchy net in our Hilbert space $\mathcal{H}$ necessarily converges. Since we have defined completeness in terms of sequences instead of nets, the following result may be helpful.

3.29. Proposition. *Let (X, d) be a metric space. The following are equivalent:*

(i) *(X, d) is complete as a metric space; i.e. every Cauchy* sequence $(x_n)_{n=1}^\infty$ *converges in X.*

(ii) *Every Cauchy net $(x_\lambda)_{\lambda \in \Lambda}$ converges in X.*

Proof.

(i) implies (ii). Suppose that every Cauchy sequence in X converges, and let $(x_\lambda)_{\lambda\in\Lambda}$ be a Cauchy net in X; that is, given $\varepsilon > 0$, there exists $\lambda_0 \in \Lambda$ so that $\alpha, \beta \geq \lambda_0$ implies $d(x_\alpha, x_\beta) < \varepsilon$.

Choose $\lambda_1 \in \Lambda$ so that $\alpha, \beta \geq \lambda_1$ implies that $d(x_\alpha, x_\beta) < 1$. Choose $\lambda_2 \geq \lambda_1$ so that $\alpha, \beta \geq \lambda_2$ implies that $d(x_\alpha, x_\beta) < \frac{1}{2}$. (Note that by definition, we can always find $\gamma_2 \in \Lambda$ so that $\alpha, \beta \geq \gamma_2$ implies that $d(x_\alpha, x_\beta) < \frac{1}{2}$; by choosing $\lambda_2 \geq \lambda_1, \gamma_2$, it follows that $\alpha, \beta \geq \lambda_2$ implies $d(x_\alpha, x_\beta) < \frac{1}{2}$, as required.)

Arguing as above, for $n \geq 2$, we can find $\lambda_{n+1} \geq \lambda_n$ so that $\alpha, \beta \geq \lambda_{n+1}$ implies that $d(x_\alpha, x_\beta) < \frac{1}{n+1}$.

Consider the sequence $(x_{\lambda_n})_{n=1}^\infty$. It is easily seen that this sequence is Cauchy, and so by hypothesis, $x = \lim_{n\to\infty} x_{\lambda_n} \in X$ exists. Let $\varepsilon > 0$ and choose $N > \frac{2}{\varepsilon}$ so that $n \geq N$ implies that $d(x_{\lambda_n}, x) < \frac{\varepsilon}{2}$.

If $\lambda \geq \lambda_N$, then

$$\begin{aligned} d(x_\lambda, x) &\leq d(x_\lambda, x_{\lambda_N}) + d(x_{\lambda_N}, x) \\ &< \frac{1}{N} + \frac{\varepsilon}{2} \\ &< \varepsilon. \end{aligned}$$

That is, $\lim_\lambda x_\lambda = x$.

(ii) implies (i). Since every sequence in X is also a net, this is clear.

□

My friends tell me I have an intimacy problem. But they don't really know me.

Garry Shandling

Exercises for Section 3.

Question 1. This question builds upon Question 4 of Chapter 2.

Let $\mathcal{H}$ be an infinite-dimensional, separable Hilbert space and suppose that $T \in \mathcal{B}(\mathcal{H})$ is a **finite-rank operator**; that is, $\dim \operatorname{ran} T < \infty$. Prove that there exists a finite-dimensional subspace $\mathcal{M} \subseteq \mathcal{H}$ such that $T\mathcal{M} \subseteq \mathcal{M}$ and $T\mathcal{M}^\perp \subseteq \mathcal{M}^\perp$.

Such a subspace is said to be **reducing** for T.

Question 2.

Let $\mathcal{H}$ be a complex Hilbert space, and let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace of $\mathcal{H}$. Denote by P the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$. Given $T \in \mathcal{B}(\mathcal{H})$, we may write the **operator matrix** for T relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ as follows:

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix},$$

where

- $T_{1\,1} = PTP_{\mathcal{M}}$,
- $T_{1\,2} = PT(I-P)|_{\mathcal{M}^\perp}$,
- $T_{2\,1} = (I-P)TP|_{\mathcal{M}}$ and
- $T_{2\,2} = (I-P)T(I-P)|_{\mathcal{M}^\perp}$.

(a) Prove that the operator matrix of P relative to this decomposition is

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

(b) Show that if $R \in \mathcal{B}(\mathcal{H})$ has the decomposition $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ relative to the same decomposition of $\mathcal{H}$, then

$$RT = \begin{bmatrix} R_{11}T_{11} + R_{12}T_{21} & R_{11}T_{12} + R_{12}T_{22} \\ R_{21}T_{11} + R_{22}T_{21} & R_{21}T_{12} + R_{22}T_{22} \end{bmatrix}.$$

In other words, this behaves just like multiplication of scalar matrices.

(c) Prove that $\|T\| \ge \|T_{ij}\|$, $1 \le i, j \le 2$.

(d) Prove that

$$\|T\| \le \max(\|T_{11}\|, \|T_{22}\|) + \max(\|T_{12}\|, \|T_{21}\|).$$

(e) We say that $\mathcal{M}$ is **invariant** for T if and only if $T\mathcal{M} \subseteq \mathcal{M}$. Prove that $\mathcal{M}$ is invariant for T if and only if $T_{21} = 0$, which in turn happens if and only if $(I-P)TP = 0$.

(f) Prove that $\mathcal{M}$ is reducing for T if and only if $T_{12} = 0$ and $T_{21} = 0$, which in turn happens if and only if $TP = PT$.

(g) Suppose that $\kappa \in \mathbb{C}$ and that $RT = TR$. Prove that $\ker(T - \kappa I)$ is invariant for R. (We say that $\ker(T - \kappa I)$ is **hyper-invariant** for T. That is, it is invariant for all operators that commute with T.)

Question 3.

Suppose that $\mathcal{H}$ is a Hilbert space and that $T \in \mathcal{B}(\mathcal{H})$. Let $\{e_\alpha\}_{\alpha \in \Lambda}$ be an ONB for $\mathcal{H}$. Prove that if $R \in \mathcal{B}(\mathcal{H})$ and $Re_\alpha = Te_\alpha$ for all $\alpha \in \Lambda$, then $T = R$.

In other words, the action of a bounded linear operator on $\mathcal{H}$ is entirely determined by its action on an ONB.

Question 4.

Prove that if $F \in \mathcal{B}(\mathcal{H})$ has rank $N \geq 1$, then F is a sum of N rank-one operators.

4. Topological Vector Spaces

Someday I want to be rich. Some people get so rich they lose all respect for humanity. That's how rich I want to be.

Rita Rudner

4.1. Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. The norm topology on $\mathcal{B}(\mathcal{H})$ is but one example of an interesting topology we can place on this set. We are also interested in studying certain weak topologies on $\mathcal{B}(\mathcal{H})$ generated by a family of functions. The topologies that we shall obtain are not induced by a metric obtained from a norm. In order to gain a better understand of the nature of the topologies we shall obtain, we now turn our attention to the notion of a *topological vector space*.

4.2. Definition. *Let $\mathcal{W}$ be a vector space over the field $\mathbb{K}$, and let $\mathcal{T}$ be a topology on $\mathcal{W}$. We say that the topology $\mathcal{T}$ is* ***compatible*** *with the vector space structure on $\mathcal{W}$ if the maps*

$$\begin{array}{rccc} \sigma: & \mathcal{W}\times\mathcal{W} & \to & \mathcal{W} \\ & (x,y) & \mapsto & x+y \end{array}$$

and

$$\begin{array}{rccc} \mu: & \mathbb{K}\times\mathcal{W} & \to & \mathcal{W} \\ & (k,x) & \mapsto & kx \end{array}$$

are continuous, where $\mathbb{K}\times\mathcal{W}$ and $\mathcal{W}\times\mathcal{W}$ carry their respective product topologies.

A ***topological vector space*** *(abbreviated TVS) is a pair $(\mathcal{W},\mathcal{T})$ where $\mathcal{W}$ is a vector space with a compatible Hausdorff topology. Informally, we refer to $\mathcal{W}$ as the topological vector space.*

4.3. Remark. Not all authors require $\mathcal{T}$ to be Hausdorff in the above definition. However, for all spaces of interest to us, the topology will indeed be Hausdorff. Furthermore, one can always pass from a non-Hausdorff topology to a Hausdorff topology via a natural quotient map. (See Appendix T.)

4.4. Example. Let $(\mathfrak{X},\|\cdot\|)$ be any normed linear space, and let $\mathcal{T}$ denote the norm topology. Suppose $(x,y)\in\mathfrak{X}\times\mathfrak{X}$ and $\varepsilon>0$. Choose a net $(x_\alpha,y_\alpha)_{\alpha\in\Lambda}\in\mathfrak{X}\times\mathfrak{X}$ so that $\lim_\alpha(x_\alpha,y_\alpha)=(x,y)$. Then there exists $\alpha_0\in\Lambda$ so that $\alpha\geq\alpha_0$ implies $\|x_\alpha-x\|<\varepsilon/2$, $\|y_\alpha-y\|<\varepsilon/2$. But then $\alpha\geq\alpha_0$ implies $\|x_\alpha+y_\alpha-(x+y)\|\leq\|x_\alpha-x\|+\|y_\alpha-y\|<\varepsilon$. In other words, $\sigma(x_\alpha,y_\alpha)$ tends to $\sigma(x,y)$ and so σ is continuous.

Similarly, if $(k,x)\in\mathbb{K}\times\mathfrak{X}$, then we can choose a net $(k_\alpha,x_\alpha)_{\alpha\in\Lambda}$ so that $\lim_\alpha(k_\alpha,x_\alpha)=(k,x)$. But then we can find $\alpha_0\in\Lambda$ so that $\alpha\geq\alpha_0$ implies $|k_\alpha-k|<1$, and so $|k_\alpha|<|k|+1$. Next we can find α_1 so that $\alpha\geq\alpha_1$ implies $|k_\alpha-k|<\varepsilon/2\|x\|$, and α_2 so that $\alpha \ge \alpha_2$ implies $\|x_\alpha - x\| < \varepsilon/2(|k|+1)$. Choosing $\alpha \ge \alpha_0, \alpha_1$ and α_2 we get

$$\begin{aligned} \|\mu(k_\alpha, x_\alpha) - \mu(k,x)\| &= \|k_\alpha x_\alpha - kx\| \\ &\le \|k_\alpha x_\alpha - k_\alpha x\| + \|k_\alpha x - kx\| \\ &\le |k_\alpha| \|x_\alpha - x\| + |k_\alpha - k| \|x\| \\ &\le \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

This proves that μ is continuous. Hence $\mathfrak{X}$ is a TVS with the norm topology.

4.5. Example. As a concrete example of the situation in Example 4.4, let $n \ge 1$ be an integer and for $(x_1, x_2, ..., x_n) \in \mathbb{C}^n$, set $\|(x_1, x_2, ..., x_n)\|_\infty = \max\{|x_k| : 1 \le k \le n\}$. Then $\mathbb{C}^n$ is a TVS with the induced norm topology. Of course, in this example, the norm topology coincides with the usual topology on $\mathbb{C}^n$ coming from the Euclidean norm $\|(x_1, x_2, ..., x_n)\|_2 = \left(\sum_{k=1}^n |x_k|^2\right)^{\frac{1}{2}}$, since $\|\cdot\|_\infty$ and $\|\cdot\|_2$ are equivalent norms on $\mathbb{C}^n$.

4.6. Remark. In the assignments we shall see how to construct a TVS which is not a normed linear space. See also the discussion of Fréchet spaces in the Appendix to Section 5.

4.7. Remark. Let $(\mathcal{V}, \mathcal{T})$ be a TVS, and let $U \in \mathcal{U}_0$ be a nbhd of 0 in $\mathcal{V}$. The continuity of $\sigma : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ implies that $\sigma^{-1}(U)$ is a nbhd of $(0,0)$ in $\mathcal{V} \times \mathcal{V}$. As such, $\sigma^{-1}(U)$ contains a basic nbhd $N_1 \times N_2$ of $(0,0)$, where $N_1, N_2 \in \mathcal{U}_0$ are open (see the Appendix). But if $N = N_1 \cap N_2$, then $N \in \mathcal{U}_0$ and $N \times N \subseteq N_1 \times N_2 \subseteq \sigma^{-1}(U)$. Thus for all $U \in \mathcal{U}_0$ there exists an open set $N \in \mathcal{U}_0$ so that $\sigma(N \times N) = N + N := \{m + n : m, n \in N\} \subseteq U$.

Similarly, we can find a nbhd $V_\varepsilon(0)$ of 0 in $\mathbb{K}$ and $N \in \mathcal{U}_0^{\mathcal{V}}$ open so that $V_\varepsilon(0) \times N \subseteq \mu^{-1}(U)$, or equivalently,

$$\{kn : n \in N, 0 \le |k| < \varepsilon\} \subseteq U.$$

It is also worth observing that if $U \in \mathcal{U}_0$, then $\mathcal{V} = \cup_{n \ge 1} nU$. Indeed, if $x \in \mathcal{V}$, then consider the continuous function $f : \mathbb{R} \to \mathcal{V}$ defined by $f(t) = tx$, so that $f(0) = 0$. The continuity of f at 0 implies that there exists $\delta > 0$ so that $|t| < \delta$ forces $f(t) = tx \in U$. Choosing $n > \frac{1}{\delta}$ then yields that $\frac{1}{n} x \in U$, or in other words that $x \in nU$. Indeed, this gives us the slightly stronger and useful conclusion that $x \in nU$ for all $n > \frac{1}{\delta}$. As such, if $(k_n)_n$ is a sequence in $\mathbb{N}$ with $\lim_{n \to \infty} k_n = \infty$, then

$$\mathcal{V} = \cup_{n \ge 1} k_n U.$$

This phenomenon is often referred to by saying that *any nbhd of* 0 *in a TVS is* ***absorbing***.

4.8. Definition. *A nbhd N of* 0 *in a TVS $\mathcal{V}$ is called* ***balanced*** *if $kN \subseteq N$ for all $k \in \mathbb{K}$ satisfying $|k| \le 1$.*

4.9. Example. Let $(\mathfrak{X}, \|\cdot\|)$ be a normed linear space. For all $\delta > 0$, $V_\delta(0) = \{x \in \mathfrak{X} : \|x\| < \delta\}$ is a balanced nbhd of 0, and if $U \in \mathcal{U}_0^{\mathfrak{X}}$, then there exists $\delta > 0$ such that $V_\delta(0) \subseteq U$.

4.10. Proposition. *Let $(\mathcal{W}, \mathcal{T})$ be a TVS. Every nbhd of 0 contains a balanced, open nbhd of 0.*

Proof. By Remark 4.7, given $U \in \mathcal{U}_0$, we can find $\varepsilon > 0$ and $N \in \mathcal{U}_0$ open such that $k \in \mathbb{K}$, $0 < |k| < \varepsilon$ implies $kN \subseteq U$. Since multiplication by a non-zero scalar is a homeomorphism, each kN is open.

Let $M = \cup_{0<|k|<\varepsilon} kN$. Then $M \subseteq U$ and $M \supseteq \frac{\varepsilon}{2}N$, so $M \in \mathcal{U}_0$. A routine calculation shows that M is balanced. It is also open, being the union of open sets.

□

4.11. Suppose $(\mathcal{W}, \mathcal{T})$ is a TVS, $w_0 \in \mathcal{W}$ and $k_0 \in \mathbb{K}$. Define

$$\tau_{w_0} : \mathcal{W} \to \mathcal{W} \quad \text{via} \quad \tau_{w_0}(x) = w_0 + x.$$

By continuity of addition, we get that τ_{w_0} is continuous, and clearly τ_{w_0} is a bijection. But $\tau_{w_0}^{-1} = \tau_{-w_0}$ is also a translation, and therefore it is continuous by the above argument. That is, τ_{w_0} is a homeomorphism.

This simple observation underlies a particularly useful fact about TVS's, namely:

$$N \in \mathcal{U}_0^{\mathcal{W}} \text{ if and only if } w_0 + N \in \mathcal{U}_{w_0}^{\mathcal{W}}.$$

That is,

the nbhd system at any point in $\mathcal{W}$ is determined by the nbhd system at 0.

If $0 \neq k_0 \in \mathbb{K}$, then $\lambda_{k_0} : \mathcal{W} \to \mathcal{W}$ defined by $\lambda_{k_0}(x) = k_0 x$ is also a continuous bijection (by continuity of scalar multiplication) and has continuous inverse $\lambda_{k_0^{-1}}$. Thus

$$N \in \mathcal{U}_0^{\mathcal{W}} \text{ if and only if } k_0 N \in \mathcal{U}_0^{\mathcal{W}} \quad \text{for all} \quad 0 \neq k_0 \in \mathbb{K}.$$

The following result shows that the assumption that a TVS topology be Hausdorff may be replaced with a weaker assumption - namely that points be closed (i.e. that $\mathcal{T}$ be T_1) - from which the Hausdorff condition follows.

4.12. Proposition. *Let $\mathcal{V}$ be a vector space with a topology $\mathcal{T}$ for which*

(i) *addition is continuous;*
(ii) *scalar multiplication is continuous; and*
(iii) *points in $\mathcal{V}$ are closed in the $\mathcal{T}$-topology.*

Then $\mathcal{T}$ is a Hausdorff topology and $(\mathcal{V}, \mathcal{T})$ is a TVS.

Proof. Let $x \neq y \in \mathcal{V}$. Then $\{y\}$ is closed (i.e. $\mathcal{V}\backslash\{y\}$ is open) and so we can find an open nbhd $U \in \mathcal{U}_x$ of x so that $y \notin U$. As above, by continuity of addition, translation is a homeomorphism of $\mathcal{V}$ and so $U = x + U_0$ for some open nbhd U_0 of 0. Also by continuity of addition, there exists an open nbhd V of 0 so that $V + V \subseteq U_0$. By continuity of scalar multiplication, $-V$ is again an open nbhd of 0, and hence $W = V \cap (-V)$ is an open nbhd of 0 as well, with $W + W \subseteq V + V \subseteq U_0$.

Suppose that $(x+W) \cap (y+W) \neq \varnothing$. Then there exist $w_1, w_2 \in W$ so that $x + w_1 = y + w_2$, i.e. $x + w_1 - w_2 = y$. But $w_1 - w_2 \in W + W$, so that $y \in x + (W + W) \subseteq x + U_0 = U$, a contradiction. Hence $(x + W) \in \mathcal{U}_x$, $(y + W) \in \mathcal{U}_y$ are disjoint open nbhds of x and y respectively, and $(\mathcal{V}, \mathcal{T})$ is Hausdorff.

□

4.13. Proposition. *Let $(\mathcal{W}, \mathcal{T})$ be a TVS and $\mathcal{Y}$ be a linear manifold in $\mathcal{W}$. Then*

(a) *$\mathcal{Y}$ is a TVS with the relative topology induced by $\mathcal{T}$; and*
(b) *$\overline{\mathcal{Y}}$ is a subspace of $\mathcal{W}$.*

Proof.

(a) This is clear, since the continuity of $\sigma|_{\mathcal{Y}}$ and $\mu|_{\mathcal{Y}}$ is inherited from the continuity of σ and μ.
(b) Suppose $y, z \in \overline{\mathcal{Y}}$ and $k \in \mathbb{K}$. Choose a net $(y_\alpha, z_\alpha) \in \mathcal{Y} \times \mathcal{Y}$ so that $\lim_\alpha (y_\alpha, z_\alpha) = (y, z)$. By continuity of σ on $\mathcal{W} \times \mathcal{W}$, $y_\alpha + z_\alpha \to y + z$. But $y_\alpha + z_\alpha \in \mathcal{Y}$ for all α, and so $y + z \in \overline{\mathcal{Y}}$. Similarly, if we choose a net (k_α, y_α) in $\mathbb{K} \times \mathcal{Y}$ which converges to (k, y), then the continuity of μ implies that $k_\alpha y_\alpha \to ky$. Since $k_\alpha y_\alpha \in \mathcal{Y}$ for all α, $ky \in \overline{\mathcal{Y}}$.

□

4.14. Exercise. Let $(\mathcal{V}, \mathcal{T})$ be a TVS. Prove the following.

(a) If $C \subseteq \mathcal{V}$ is convex, then so is $\overline{C}$.
(b) If $E \subseteq \mathcal{V}$ is balanced, then so is $\overline{E}$.

4.15. Definition. *Let $(\mathcal{V}, \mathcal{T})$ be a TVS, and let $(x_\lambda)_\lambda$ be a net in $\mathcal{V}$. We say that $(x_\lambda)_\lambda$ is a* ***Cauchy net*** *if for all $U \in \mathcal{U}_0$ there exists $\lambda_0 \in \Lambda$ so that $\lambda_1, \lambda_2 \geq \lambda_0$ implies that $x_{\lambda_1} - x_{\lambda_2} \in U$.*

We say that a subset $K \subseteq \mathcal{V}$ is ***Cauchy complete*** *if every Cauchy net in K converges to some element of K.*

We pause to verify that if $(x_\lambda)_\lambda$ is a net in $\mathcal{V}$ which converges to some $x \in \mathcal{V}$, then $(x_\lambda)_\lambda$ is a Cauchy net. Indeed, let $U \in \mathcal{U}_0$ and choose a balanced, open nbhd $N \in \mathcal{U}_0$ so that $N + N = N - N \subseteq U$. Also, choose $\lambda_0 \in \Lambda$ so that $\lambda \geq \lambda_0$ implies that $x_\lambda \in x + N$. If $\lambda_1, \lambda_2 \geq \lambda_0$, then $x_{\lambda_1} - x_{\lambda_2} = (x_{\lambda_1} - x) - (x_{\lambda_2} - x) \in N - N \subseteq U$. Thus $(x_\lambda)_\lambda$ is a Cauchy net.

4.16. Example. If $(\mathfrak{X}, \|\cdot\|)$ is a normed linear space, then $\mathfrak{X}$ is Cauchy complete if and only if $\mathfrak{X}$ is complete. Indeed, the topology here being a metric topology, we need only consider sequences instead of general nets.

4.17. Lemma. *Let $\mathcal{V}$ be a TVS and $\mathcal{K} \subseteq \mathcal{V}$ be complete. Then $\mathcal{K}$ is closed in $\mathcal{V}$.*

Proof. Suppose that $z \in \overline{\mathcal{K}}$. For each $U \in \mathcal{U}_z$, there exists $y_U \in \mathcal{K}$ so that $y_U \in U$. The family $\{U : U \in \mathcal{U}_z\}$ forms a directed set under reverse-inclusion, namely: $U_1 \leq U_2$ if $U_2 \subseteq U_1$. Thus $(y_U)_{U \in \mathcal{U}_z}$ is a net in $\mathcal{K}$. By definition, this net converges to the point z, i.e. $\lim_U y_U = z$. (Since $\mathcal{V}$ is Hausdorff, this is the unique limit point of the net $(y_U)_U$.) From the comments following Definition 4.15, $(y_U)_{U \in \mathcal{U}_z}$ is a Cauchy net. Since $\mathcal{K}$ is complete, it follows that $z \in \mathcal{K}$, and hence that $\mathcal{K}$ is closed.

□

4.18. Quotient spaces. Let $(\mathcal{V}, \mathcal{T})$ be a TVS and $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Then $\mathcal{V}/\mathcal{W}$ exists as a quotient space of vector spaces. Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ denote the canonical quotient map.

We can establish a topology on $\mathcal{V}/\mathcal{W}$ using the $\mathcal{T}$ topology on $\mathcal{V}$ by defining a subset $G \subseteq \mathcal{V}/\mathcal{W}$ to be open if $q^{-1}(G)$ is open in $\mathcal{V}$.

- If $\{G_\lambda\}_\lambda \subseteq \mathcal{V}/\mathcal{W}$ and $q^{-1}(G_\lambda)$ is open in $\mathcal{V}$ for all λ, then $q^{-1}(\cup_\lambda G_\lambda) = \cup_\lambda q^{-1}(G_\lambda)$ is open, and thus $\cup_\lambda G_\lambda$ is open in $\mathcal{V}/\mathcal{W}$.
- If $G_1, G_2 \subseteq \mathcal{V}/\mathcal{W}$ and $q^{-1}(G_i)$ is open in $\mathcal{V}$, $i = 1, 2$, then $q^{-1}(G_1 \cap G_2) = q^{-1}(G_1) \cap q^{-1}(G_2)$ is open in $\mathcal{V}$, whence $G_1 \cap G_2$ is open in $\mathcal{V}/\mathcal{W}$.
- Clearly $\varnothing = q^{-1}(\varnothing)$ and $\mathcal{V} = q^{-1}(\mathcal{V}/\mathcal{W})$ are open in $\mathcal{V}$, so that $\varnothing, \mathcal{V}/\mathcal{W}$ are open in $\mathcal{V}/\mathcal{W}$ and the latter is a topological space.

We refer to this topology on $\mathcal{V}/\mathcal{W}$ as the **quotient topology**. The quotient map is continuous with respect to the quotient topology, by design. In fact, the quotient topology is the largest topology on $\mathcal{V}/\mathcal{W}$ which makes q continuous.

We begin by proving that q is an **open map**. That is, if $G \subseteq \mathcal{V}$ is open, then $q(G)$ is open in $\mathcal{V}/\mathcal{W}$.

Indeed, for each $w \in \mathcal{W}$, the set $G + w$ is open in $\mathcal{V}$, being a translate of the open set G. Hence $G + \mathcal{W} = \cup_{w \in \mathcal{W}} G + w$ is open in $\mathcal{V}$, being the union of open sets. But

$$G + \mathcal{W} = q^{-1}(q(G)),$$

so that $q(G)$ is open in $\mathcal{V}/\mathcal{W}$ by definition.

To see that addition is continuous in $\mathcal{V}/\mathcal{W}$, let $x + \mathcal{W}$, $y + \mathcal{W}$ and let E be a nbhd of $(x + y) + \mathcal{W}$ in $\mathcal{V}/\mathcal{W}$. Then $q^{-1}(E)$ is a nbhd of $x + y$ in $\mathcal{V}$. Choose open nbhds U_x of x and U_y of y in $\mathcal{V}$ so that $r \in U_x$, $s \in U_y$ implies that $r + s \in q^{-1}(E)$. Note that $x + \mathcal{W} \in q(U_x), y + \mathcal{W} \in q(U_y)$ and that $q(U_x), q(U_y)$ are open in $\mathcal{V}/\mathcal{W}$ by the argument above. If $a + \mathcal{W} \in q(U_x)$, $b + \mathcal{W} \in q(U_y)$, then $a + \mathcal{W} = g + \mathcal{W}$ and $b + \mathcal{W} = h + \mathcal{W}$ for some $g \in U_x$, $h \in U_y$. Thus

$$(a + b) + \mathcal{W} = (g + h) + \mathcal{W} \in q(q^{-1}(E)) \subseteq E.$$

Hence addition is continuous.

That scalar multiplication is continuous follows from a similar argument which is left to the reader.

Finally, to see that the resulting quotient topology is Hausdorff, it suffices (by Proposition 4.12) to show that points in $\mathcal{V}/\mathcal{W}$ are closed. Let $x + \mathcal{W} \in \mathcal{V}/\mathcal{W}$. Then

$q^{-1}(x + \mathcal{W}) = \{x + w : w \in \mathcal{W}\}$ is closed in $\mathcal{V}$, being a translation of the closed subspace $\mathcal{W}$. Hence the complement $C = \mathcal{V} \backslash \{x + w : w \in \mathcal{W}\}$ of $x + \mathcal{W}$ is open in $\mathcal{V}$. But then $q(C)$ is open, since q is an open map, and $q(C)$ is the complement of $x + \mathcal{W}$.

Finite-dimensional topological vector spaces. Our present goal is to prove that there is only one topology that one can impose upon a finite-dimensional vector space $\mathcal{V}$ to make it into a TVS. We begin with the one dimensional case.

4.19. Lemma. *Let $(\mathcal{V}, \mathcal{T})$ be a one-dimensional TVS over $\mathbb{K}$. Let $\{e\}$ be a basis for $\mathcal{V}$. Then $\mathcal{V}$ is homeomorphic to $\mathbb{K}$ via the map*

$$\begin{array}{rccc} \tau : & \mathbb{K} & \to & \mathcal{V} \\ & k & \mapsto & ke \end{array}.$$

Proof. The map τ is clearly a bijection, and the continuity of scalar multiplication in a TVS makes it continuous as well. We shall demonstrate that the inverse map $\tau^{-1}(ke) = k$ is also continuous. To do this, it suffices to show that if $\lim_\lambda k_\lambda e = 0$, then $\lim_\lambda k_\lambda = 0$. (*Why?*)

Let $\delta > 0$. Then $\delta e \neq 0$, and as $\mathcal{V}$ is Hausdorff, we can find a nbhd U of 0 so that $\delta e \notin U$. By Proposition 4.10, U contains a balanced nbhd V of 0. Obviously, $\delta e \notin V$. Since $\lim_\lambda k_\lambda e = 0$, there exists λ_0 so that $\lambda \geq \lambda_0$ implies that $k_\lambda e \in V$.

Suppose that there exists $\beta \geq \lambda_0$ with $|k_\beta| \geq \delta$. Then

$$\delta e = \left(\frac{\delta}{k_\beta}\right) k_\beta e \in V,$$

as V is balanced. This contradiction shows that $\lambda \geq \lambda_0$ implies that $|k_\lambda| < \delta$. Since $\delta > 0$ was arbitrary, we have shown that $\lim_\lambda k_\lambda = 0$.

□

4.20. Proposition. *Let $n \geq 1$ be an integer, and let $(\mathcal{V}, \mathcal{T})$ be an n-dimensional TVS over $\mathbb{K}$ with basis $\{e_1, e_2, ..., e_n\}$. The map*

$$\begin{array}{rccc} \tau : & \mathbb{K}^n & \to & \mathcal{V} \\ & (k_1, k_2, ..., k_n) & \mapsto & \sum_{j=1}^n k_j e_j \end{array}$$

is a homeomorphism.

Proof. Lemma 4.19 shows that the result is true for $n = 1$. We shall argue by induction on the dimension of $\mathcal{V}$.

It is clear that τ is a linear bijection, and furthermore, since $\mathcal{V}$ is a TVS, the continuity of addition and scalar multiplication in $\mathcal{V}$ implies that τ is continuous. There remains to show that τ^{-1} is continuous as well.

Suppose that the result is true for $1 \leq n < m$. We shall prove that it holds for $n = m$ as well. To that end, let $F = \{e_{t_1}, e_{t_2}, ..., e_{t_r}\}$ for some $1 \leq r < m$, and let $E = \{e_1, e_2, ..., e_m\} \backslash F = \{e_{p_1}, e_{p_2}, ..., e_{p_s}\}$.

Now $\mathcal{Y} = \operatorname{span} E$ is an s-dimensional space with $s < m$. By our induction hypothesis, the map $(k_1, k_2, ..., k_s) \mapsto \sum_{j=1}^s k_j e_{p_j}$ is a homeomorphism. It follows that $\mathcal{Y}$ is complete (*check!*) and therefore closed, by Lemma 4.17. By the arguments of paragraph 4.18, $\mathcal{V}/\mathcal{Y}$ is a TVS and the canonical map $q_{\mathcal{Y}} : \mathcal{V} \to \mathcal{V}/\mathcal{Y}$ is continuous. Moreover, $\{q(e_{t_1}), q(e_{t_2}), ..., q(e_{t_r})\}$ is a basis for $\mathcal{V}/\mathcal{Y}$. Since $r < m$, our induction hypothesis once again shows that the map

$$\begin{array}{rccc} \rho_{\mathcal{Y}} : & \mathcal{V}/\mathcal{Y} & \to & \mathbb{K}^r \\ & \sum_{j=1}^r k_j q(e_{t_j}) & \mapsto & (k_1, k_2, ..., k_r) \end{array}$$

is continuous. Thus

$$\begin{array}{rccc} \gamma := \rho_{\mathcal{Y}} \circ q : & \mathcal{V} & \to & \mathbb{K}^r \\ & \sum_{i=1}^n k_i e_i & \mapsto & \rho_{\mathcal{Y}}(\sum_{i=1}^n k_i q(e_i)) = (k_{t_1}, k_{t_2}, ..., k_{t_r}) \end{array}$$

is also continuous, being the composition of continuous functions.

To complete the proof, we first apply the above argument with $F = \{e_m\}$ to get that

$$\begin{array}{rccc} \gamma_1 : & \mathcal{V} & \to & \mathbb{K} \\ & \sum_{i=1}^n k_i e_i & \mapsto & k_m \end{array}$$

is continuous, and then to $F = \{e_1, e_2, ..., e_{m-1}\}$ to get that

$$\begin{array}{rccc} \gamma_2 : & \mathcal{V} & \to & \mathbb{K}^{m-1} \\ & \sum_{i=1}^n k_i e_i & \mapsto & (k_1, k_2, ..., k_{m-1}) \end{array}$$

is continuous. Since $\tau^{-1} = (\gamma_1, \gamma_2)$, it too is continuous, and we are done.

□

The previous result has a number of important corollaries:

4.21. Corollary. *Let $n \geq 1$ be an integer and $\mathcal{V}$ be an n-dimensional vector space. Then there is a unique topology $\mathcal{T}$ which makes $\mathcal{V}$ a TVS. In particular, therefore,* ***all norms on a finite dimensional vector space are equivalent****.*

Proof. Since any topology on $\mathcal{V}$ which makes it a TVS is determined completely by the product topology on $\mathbb{K}^n$, it is unique. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on $\mathcal{V}$, then they induce metric topologies which make $\mathcal{V}$ into a TVS. But these topologies coincide, from the above argument. By Proposition 1.19, the norms are equivalent.

□

4.22. Corollary. *Let $\mathcal{V}$ be a TVS and $\mathcal{W}$ be a finite-dimensional linear manifold of $\mathcal{V}$. Then $\mathcal{W}$ is closed in $\mathcal{V}$.*

Proof. This argument is embedded in the proof of Proposition 4.20; $\mathcal{W}$ is complete because of the nature of the homeomorphism between $\mathcal{W}$ and $\mathbb{K}^n$, where n is the dimension of $\mathcal{W}$. Then Lemma 4.17 implies that $\mathcal{W}$ is closed.

□

Recall from your Real Analysis course that the closed unit ball of $(\mathbb{K}^n, \|\cdot\|_2)$ is compact. Our next result essentially says that $\mathbb{K}^n$ is the only topological vector space with this property.

4.23. Definition. *A topological space $(X, \mathcal{T})$ is said to be* ***locally compact*** *if each point in X has a nbhd base consisting of compact sets.*

Suppose that X is locally compact and Hausdorff, and that $x_0 \in X$. Then for all $U \in \mathcal{U}_{x_0}$, there exists $K \in \mathcal{U}_{x_0}$ so that K is compact and $K \subseteq U$. Choose $G \in \mathcal{T}$ so that $x_0 \in G \subseteq K$. Then $\overline{G} \subseteq \overline{K} = K \subseteq U$, and so $\overline{G}$ is compact. That is, if X is Hausdorff and locally compact, then for any $U \in \mathcal{U}_{x_0}$, there exists $G \in \mathcal{T}$ so that $\overline{G}$ is compact and $x_0 \in G \subseteq \overline{G} \subseteq U$.

4.24. Example. Let $n \geq 1$ be an integer and consider $(\mathbb{K}^n, \|\cdot\|_2)$. Then for each $x \in \mathbb{K}^n$, the collection

$$\{\overline{B}_\varepsilon(x) := \{y \in \mathbb{K}^n : \|y - x\|_2 \leq \varepsilon\} : \varepsilon > 0\}$$

is a nbhd base at x consisting of compact sets, so $(\mathbb{K}^n, \|\cdot\|_2)$ is locally compact.

4.25. Theorem. *A TVS $(\mathcal{V}, \mathcal{T})$ is locally compact if and only if $\mathcal{V}$ is finite-dimensional.*

Proof. If $\dim \mathcal{V} < \infty$, then $\mathcal{V}$ is homeomorphic to $(\mathbb{K}^n, \|\cdot\|_2)$ by Proposition 4.20. Since $(\mathbb{K}^n, \|\cdot\|_2)$ is locally compact from above, so is $\mathcal{V}$.

Conversely, suppose that $\mathcal{V}$ is locally compact. Choose $K \in \mathcal{U}_0$ compact. Using Remark 4.7, we can find a nbhd $N \in \mathcal{U}_0^{\mathcal{V}}$ such that $N + N \subseteq K$. By replacing N with its interior if necessary, we may assume without loss of generality that N is open. Now $K \subseteq \cup_{x \in K}\ x + N$. Since the latter is an open cover of K, we can find $x_1, x_2, ..., x_r \in K$ so that

$$K \subseteq \cup_{i=1}^r\ x_i + N = \{x_1, x_2, ..., x_r\} + N.$$

Let $\mathcal{M} = \text{span}\{x_1, x_2, ..., x_r\}$. Consider the quotient map $q : \mathcal{V} \to \mathcal{V}/\mathcal{M}$. As we have seen, q is both continuous and open. Furthermore,

$$q(K) \subseteq \cup_{i=1}^r q(x_i + N) = \cup_{i=1}^r q(N) = q(N) \subseteq q(K),$$

as $q(x_i) = 0$ for all $1 \leq i \leq r$.

Since $N + N \subseteq K$, we see that

$$2\, q(K) \subseteq q(K) + q(K) = q(N) + q(N) \subseteq q(K).$$

By a simple induction argument, we see that $2^m q(K) \subseteq q(K)$ for all $m \geq 1$.

But $K \in \mathcal{U}_0$ implies that K is absorbing, and thus $q(\mathcal{V}) \subseteq \cup_{m \geq 1} 2^m q(K) = q(K)$. Since K is compact and q is continuous, we infer that $q(\mathcal{V})$ is compact. Suppose that $q(\mathcal{V}) \neq \{0\}$. Then $q(\mathcal{V})$ contains a one-dimensional subspace $\mathbb{K}(y + \mathcal{M})$ for some $y \in \mathcal{V} \setminus \mathcal{M}$. By Corollary 4.22, $\mathbb{K}(y + \mathcal{M})$ is closed in $q(\mathcal{V})$. Since $q(\mathcal{V})$ is compact, so is $\mathbb{K}(y + \mathcal{M})$. By Lemma 4.19, $\mathbb{K}(y + \mathcal{M})$ is homeomorphic to $\mathbb{K}$, which forces $\mathbb{K}$ to be compact as well, which is absurd.

Hence $q(\mathcal{V}) = \{0\}$, or in other words, $\mathcal{V} = \mathcal{M}$, which is finite-dimensional. This completes the proof.

□

An interesting and useful consequence of this result is the following.

4.26. Corollary. *Let $(\mathfrak{X}, \|\cdot\|)$ be a NLS. Then the closed unit ball $\mathfrak{X}_1$ of $\mathfrak{X}$ is compact if and only if $\mathfrak{X}$ is finite-dimensional.*

Proof. First suppose that $\mathfrak{X}_1$ is compact. If $U \in \mathcal{U}_0^{\mathfrak{X}}$ is any nbhd of 0, then there exists $\delta > 0$ so that $\|x\| < 2\delta$ implies that $x \in U$. But then $\mathfrak{X}_\delta \subseteq U$, and $\mathfrak{X}_\delta = \delta \mathfrak{X}_1$ is compact, being a homeomorphic image of $\mathfrak{X}_1$. By definition, $\mathfrak{X}$ is locally compact, hence finite-dimensional, by Theorem 4.25.

Conversely, suppose that $\mathfrak{X}$ is finite-dimensional. By Theorem 4.25, $\mathfrak{X}$ is locally compact. By hypothesis, there exists a compact nbhd K of 0. As above, there exists $\delta > 0$ so that $\mathfrak{X}_\delta \subseteq K$. Since $\mathfrak{X}_\delta$ is a closed subset of a compact set, it is compact. Since $\mathfrak{X}_1 = (\delta^{-1})\mathfrak{X}_\delta$ is a homeomorphic image of a compact set, it too is compact.

□

4.27. Definition. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces and suppose that $f : \mathcal{V} \to \mathcal{W}$ is a (not necessarily linear) map. We say that f is **uniformly continuous** if, given $U \in \mathcal{U}_0^{\mathcal{W}}$ there exists $N \in \mathcal{U}_0^{\mathcal{V}}$ so that $x - y \in N$ implies that $f(x) - f(y) \in U$.*

4.28. The definition of uniform continuity given here derives from the fact that the collection $\mathcal{B} = \{\mathcal{B}(U) = \{(x, y) : x - y \in U\} : U \in \mathcal{U}_0^{\mathcal{W}}\}$ defines what is known as a *uniformity* on the TVS $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ whose corresponding uniform topology coincides with the initial topology $\mathcal{T}$. The interested reader is referred to the book by Willard [**Wil70**] and to the books of Kadison and Ringrose [**KR83**] for a more complete development along these lines. We shall focus only upon that part of the theory which we require in this text.

4.29. Let us verify that in the case where $(\mathfrak{X}, \|\cdot\|_\mathfrak{X})$ and $(\mathfrak{Y}, \|\cdot\|_\mathfrak{Y})$ are normed linear spaces, our new notion of uniform continuity coincides with our metric space notion.

Observe first that if $(\mathfrak{Z}, \|\cdot\|)$ is a general normed linear space and $\delta > 0$, then $x - y \in V_\delta^{\mathfrak{Z}}(0)$ if and only if $\|x - y\| < \delta$.

Suppose $f : \mathfrak{X} \to \mathfrak{Y}$ is uniformly continuous in the sense of Definition 4.27. Given $\varepsilon > 0$, $V_\varepsilon^{\mathfrak{Y}}(0) = \{y \in \mathfrak{Y} : \|y\|_\mathfrak{Y} < \varepsilon\} \in \mathcal{U}_0^{\mathfrak{Y}}$ and so there exists $N \in \mathcal{U}_0^{\mathfrak{X}}$ such that $x - y \in N$ implies $f(x) - f(y) \in V_\varepsilon^{\mathfrak{Y}}(0)$. But $N \in \mathcal{U}_0^{\mathfrak{X}}$ implies that there exists $\delta > 0$ so that $V_\delta^{\mathfrak{X}}(0) \subseteq N$. Thus $\|x - y\|_\mathfrak{X} < \delta$ implies that $x - y \in N$, and thus $f(x) - f(y) \in V_\varepsilon^{\mathfrak{Y}}(0)$, i.e. $\|f(x) - f(y)\|_\mathfrak{Y} < \varepsilon$. Thus is the standard (metric) notion of uniform continuity in a normed space.

Conversely, suppose that $f : \mathfrak{X} \to \mathfrak{Y}$ is uniformly continuous in the standard metric sense. Let $U \in \mathcal{U}_0^{\mathfrak{Y}}$. Then there exists $\varepsilon > 0$ so that $V_\varepsilon^{\mathfrak{Y}}(0) \subseteq U$. By hypothesis, there exists $\delta > 0$ so that $\|x - y\|_\mathfrak{X} < \delta$ implies $\|f(x) - f(y)\|_\mathfrak{Y} < \varepsilon$, and hence $x - y \in V_\delta^{\mathfrak{X}}(0)$ implies that $f(x) - f(y) \in V_\epsilon^{\mathfrak{Y}}(0) \subseteq U$. That is, f is continuous in the sense of Definition 4.27.

It is useful to extend our notion of uniformly continuous functions between topological vectors spaces to functions defined only upon a subset (not necessarily a subspace) of the domain space $\mathcal{V}$.

4.30. Definition. *If $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ are topological vector spaces and $C \subseteq \mathcal{V}$, then $f : C \to \mathcal{W}$ is* ***uniformly continuous*** *if for all $U \in \mathcal{U}_0^\mathcal{W}$ there exists $N \in \mathcal{U}_0^\mathcal{V}$ such that $x, y \in C$ and $x - y \in N$ implies $f(x) - f(y) \in U$.*

4.31. Example. Now $(\mathbb{R}, |\cdot|)$ is a normed linear space, and so the comments of paragraph 4.29 apply. The function $f : (0,1) \to \mathbb{R}$ defined by $f(x) = x^2$ is therefore uniformly continuous, whereas $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$ is not.

Another way in which uniform continuity in the TVS setting extends the notion of uniform continuity in the metric setting is evinced by the following:

4.32. Proposition. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces and $f : \mathcal{V} \to \mathcal{W}$ be uniformly continuous. Then f is continuous on $\mathcal{V}$.*

Proof. Let $x_0 \in \mathcal{V}$. Let $U \in \mathcal{U}_{f(x_0)}^\mathcal{W}$. Then by paragraph 4.11, $U = f(x_0) + U_0$ where $U_0 \in \mathcal{U}_0^\mathcal{W}$. By hypothesis, there exists $N_0 \in \mathcal{U}_0^\mathcal{V}$ so that $x - x_0 \in N_0$ implies that $f(x) - f(x_0) \in U_0$. That is, $x \in x_0 + N_0$ implies $f(x) \in f(x_0) + N_0 = U$. Since $N := x_0 + N_0 \in \mathcal{U}_{x_0}^\mathcal{V}$, we see that f is continuous at x_0. But $x_0 \in \mathcal{V}$ was arbitrary, and so f is continuous on $\mathcal{V}$.

□

4.33. Theorem. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces over $\mathbb{K}$. Suppose that $T : \mathcal{V} \to \mathcal{W}$ is linear. The following are then equivalent:*

(a) *there exists $x_0 \in \mathcal{V}$ so that T is continuous at x_0; and*
(b) *T is uniformly continuous on $\mathcal{V}$.*

Proof. By Proposition 4.32, it suffices to prove that (a) implies (b). To that end, suppose that T is continuous at x_0 and let $U_0 \in \mathcal{U}_0^\mathcal{W}$. Then $U := Tx_0 + U_0 \in \mathcal{U}_{Tx_0}^\mathcal{W}$. By continuity of T at x_0, there exists $N \in \mathcal{U}_{x_0}^\mathcal{V}$ so that $T(N) \subseteq U$. But $N = x_0 + N_0$ for some $N_0 \in \mathcal{U}_0^\mathcal{V}$. Now if $z \in N_0$, then $x_0 + z \in N$, and so $T(x_0 + z) = Tx_0 + Tz \in Tx_0 + U_0$. That is, $Tz \in U_0$.

In particular, if $x - y \in N_0$, then $T(x - y) = Tx - Ty \in U_0$, and so T is uniformly continuous on $\mathcal{V}$.

□

4.34. Remark. Observe that exactly the same argument shows that if $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ are topological vector spaces over $\mathbb{C}$ and $T : \mathcal{V} \to \mathcal{W}$ is *conjugate-linear* (i.e. $T(\kappa x + y) = \overline{\kappa} Tx + Ty$ for all $\kappa \in \mathbb{C}$, $x, y \in \mathcal{V}$), then the following are then equivalent:

(a) there exists $x_0 \in \mathcal{V}$ so that T is continuous at x_0; and
(b) T is uniformly continuous on $\mathcal{V}$.

4.35. Corollary. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be topological vector spaces over $\mathbb{K}$ and suppose that* $\dim \mathcal{V} = n < \infty$. *If $T : \mathcal{V} \to \mathcal{W}$ is linear, then $\mathcal{T}$ is continuous.*

Proof. By Theorem 4.33 and Proposition 4.32, it suffices to prove that T is continuous at 0. Let $\{e_1, e_2, ..., e_n\}$ be a basis for $\mathcal{V}$, and suppose that $(x_\lambda)_{\lambda \in \Lambda}$ is a net in $\mathcal{V}$ which converges to 0. For each $\lambda \in \Lambda$ we may express x_λ as a unique linear combination of the e_j's, say

$$x_\lambda = k_{\lambda,1} e_1 + k_{\lambda,2} e_2 + \cdots + k_{\lambda,n} e_n.$$

By Proposition 4.20, $\lim_\lambda x_\lambda = 0$ implies that for $1 \le j \le n$, $\lim_\lambda k_{\lambda,j} = 0$.

Now $Tx_\lambda = \sum_{j=1}^n k_{\lambda,j} Te_j$, $\lambda \in \Lambda$. But addition and scalar multiplication in $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ are continuous, and $\lim_\lambda k_{\lambda,j} = 0$, so

$$\lim_\lambda Tx_\lambda = \sum_{j=1}^n \lim_\lambda k_{\lambda,j} Te_j = \sum_{j=1}^n 0 \; Te_j = 0 = T(\lim_\lambda x_\lambda).$$

It follows that T is continuous at 0, as was required.

□

If we restrict our attention to subsets of $\mathcal{V}$ we get:

4.36. Proposition. *Let $\mathcal{V}, \mathcal{W}$ be topological vector spaces and $T : \mathcal{V} \to \mathcal{W}$ be linear. Suppose that $0 \in C \subseteq \mathcal{V}$ is balanced and convex. If $T|_C$ is continuous at 0, then $T|_C$ is uniformly continuous.*

Proof. Our assumption is that $T|_C$ is continuous at 0, and thus for all $U \in \mathcal{U}_0^\mathcal{W}$, there exists $N \in \mathcal{U}_0^\mathcal{V}$ so that $x \in C \cap N$ implies $Tx \in \frac{1}{2}U$.

By Proposition 4.10, every nbhd $N \in \mathcal{U}_0^\mathcal{V}$ contains a balanced nbhd N_0, and so by replacing N by N_0 if necessary, we may assume that N is balanced.

Now suppose that $x, y \in C$ and that $x - y \in N$. Then C balanced implies that $-y \in C$, and C convex implies that $\frac{1}{2}x + \frac{1}{2}(-y) = \frac{1}{2}(x - y) \in C$. Since N is balanced, $\frac{1}{2}(x - y) \in N$ and so $T(\frac{1}{2}(x - y)) = \frac{1}{2}(Tx - Ty) \in \frac{1}{2}U$. Hence $x, y \in C$, $x - y \in N$ implies $Tx - Ty \in U$. That is, T is uniformly continuous on C.

□

If A, B, and C are sets with $A \subseteq B$, and if $f : A \to C$ is a map, the we say that the map $g : B \to C$ *extends* f (or that g is an *extension* of f) if $g|_A = f$.

4.37. Proposition. *Suppose that $\mathcal{V}$ and $\mathcal{W}$ are topological vector spaces and that $\mathcal{W}$ is Cauchy complete. If $\mathcal{X} \subseteq \mathcal{V}$ is a linear manifold and $T_0 : \mathcal{X} \to \mathcal{W}$ is continuous and linear, then T_0 extends to a continuous linear map $T : \overline{\mathcal{X}} \to \mathcal{W}$.*

Proof. Let $x \in \overline{\mathcal{X}}$ and choose $(x_\lambda)_{\lambda \in \Lambda_1}$ in $\mathcal{X}$ so that $\lim_\lambda x_\lambda = x$. Clearly, if T is to be continuous, we shall need $Tx = \lim_\lambda Tx_\lambda$. The issue is whether or not this limit exists and is independent of the choice of $(x_\lambda)_\lambda$.

Now $(x_\lambda)_\lambda$ is a Cauchy net. Take $U \in U_0^\mathcal{W}$. Since T_0 is continuous, there exists $N \in \mathcal{U}_0^\mathcal{X}$ such that $w \in N$ implies that $T_0 w \in U$. Hence, there exists λ_0 such that $\lambda_1, \lambda_2 \ge \lambda_0$ implies that $x_{\lambda_1} - x_{\lambda_2} \in N$ and therefore that $T_0 x_{\lambda_1} - T_0 x_{\lambda_2} = T_0(x_{\lambda_1} - x_{\lambda_2}) \in U$. Thus $(T_0 x_\lambda)_\lambda$ is also a Cauchy net. Our assumption that $\mathcal{W}$ is

Cauchy complete implies that there exists z (depending *a priori* upon $(x_\lambda)_\lambda$) such that $z = \lim_\lambda T_0 x_\lambda$.

Suppose that $(y_\beta)_{\beta\in\Lambda_2} \in X$ and that $\lim_\beta y_\beta = x$. Arguing as above, there exists $z_2 \in \mathcal{W}$ so that $z_2 = \lim_\beta T_0 y_\beta$. If we set

$$\begin{aligned} y_{(\lambda,\beta)} &:= y_\beta, \quad \lambda \in \Lambda_1 \\ x_{(\lambda,\beta)} &:= x_\lambda, \quad \beta \in \Lambda_2, \end{aligned}$$

and $\Lambda = \Lambda_1 \times \Lambda_2$ - equipped with the direction $(\lambda_1, \beta_1) \le (\lambda_2, \beta_2)$ if $\lambda_1 \le \lambda_2$ and $\beta_1 \le \beta_2$ - then

$$\lim_{(\lambda,\beta)\in\Lambda} x_{(\lambda,\beta)} = \lim_{(\lambda,\beta)} y_{(\lambda,\beta)} = x.$$

Also, $\lim_{(\lambda,\beta)} T_0 x_{(\lambda,\beta)} = z_1$, $\lim_{(\lambda,\beta)} T_0 y_{(\lambda,\beta)} = z_2$. Thus $\lim_{(\lambda,\beta)} x_{(\lambda,\beta)} - y_{(\lambda,\beta)} = 0 \in \mathcal{X}$ and so by the continuity of T_0 on $\mathcal{X}$,

$$\begin{aligned} 0 = T_0 0 &= T_0(\lim_{(\lambda,\beta)} x_{(\lambda,\beta)} - y_{(\lambda,\beta)}) \\ &= \lim_{(\lambda,\beta)} T_0(x_{(\lambda,\beta)} - y_{(\lambda,\beta)}) \\ &= z_1 - z_2. \end{aligned}$$

That is, we can set $Tx = \lim_\lambda T_0 x_\lambda$ and this is well-defined.

That T is linear on $\overline{\mathcal{X}}$ is left as an exercise.

Finally, to see that T is continuous on $\overline{\mathcal{X}}$, let $U \in \mathcal{U}_0^{\mathcal{W}}$ and choose $U_1 \in \mathcal{U}_0^{\mathcal{W}}$ so that $U_1 + U_1 \subseteq U$. Choose $N \in \mathcal{U}_0^{\mathcal{X}}$ so that $x \in N$ implies $Tx = T_0 x \in U_1$. Then $N = G \cap \mathcal{X}$ for some $G \in \mathcal{U}_0^{\mathcal{V}}$.

Let $M = G \cap \overline{\mathcal{X}}$ so that $M \in \mathcal{U}_0^{\overline{\mathcal{X}}}$. If $z \in M$, then $z = \lim_\lambda x_\lambda$ for some $x_\lambda \in N$, $\lambda \in \Lambda$. Now $Tz = \lim_\lambda T_0 x_\lambda$, so that there exists $\lambda_0 \in \Lambda$ so that $\lambda \ge \lambda_0$ implies that $Tz - T_0 x_\lambda \in U_1$.

But $T_0 x_\lambda \in U_1$ for all λ and so $Tz = (Tz - T_0 x_\lambda) + T_0 x_\lambda \in U_1 + U_1 \subseteq U$. That is, $z \in M$ implies that $Tz \in U$. Hence T is continuous at 0, and consequently T is uniformly continuous on $\overline{\mathcal{X}}$.

□

4.38. Corollary. *Suppose that $\mathfrak{X}$ and $\mathfrak{Y}$ are Banach spaces and that $\mathfrak{M} \subseteq \mathfrak{X}$ is a linear manifold. If $T_0 : \mathfrak{M} \to \mathfrak{Y}$ is bounded, then T_0 extends to a bounded linear map $T : \overline{\mathfrak{M}} \to \mathfrak{Y}$, and $\|T\| = \|T_0\|$.*

Proof. Invoking Proposition 4.37, there remains only to show that $\|T\| = \|T_0\|$. That $\|T\| \ge \|T_0\|$ is clear.

Conversely, given $x \in \overline{\mathfrak{M}}$ with $\|x\| = 1$ and $\varepsilon > 0$, there exists $y \in \mathfrak{M}$ so that $\|y\| = 1$ and $\|x - y\| < \varepsilon$. Then

$$\|T_0 y - Tx\| = \|Ty - Tx\| \le \|T\| \; \|y - x\| < \|T\|\varepsilon.$$

(Recall that T is bounded since T is continuous!) Since $\varepsilon > 0$ was arbitrary,

$$\sup\{\|T_0 y\| : y \in \mathfrak{M}, \|y\| = 1\} \ge \sup\{\|Tx\| : \|x\| = 1\},$$

so $\|T_0\| \geq \|T\|$, completing the proof.

□

Appendix to Section 4.

4.39. We have only touched upon the basics of the theory of topological vector spaces. Indeed, our interests lie much closer to the study of normed linear spaces and Banach spaces. Our reason for developing the theory of TVS's to this extent is that there are many topologies that one associates to Banach spaces that are not necessarily norm topologies, including weak and weak*-topologies which we shall study in Chapter 7 and beyond.

One can develop a theory of normed linear spaces, and deal with each of these topologies on an *ad hoc* basis, however we feel that the price we pay in introducing this more general approach is compensated by having our versions of the Hahn-Banach Theorem hold in *any* "locally convex space", a special kind of TVS whose topology, as we shall see in the next Chapter, is induced by a separating family of seminorms.

Do you know what it means to come home at night to a woman who'll give you a little love, a little affection, a little tenderness? It means you're in the wrong house, that's what it means.

Henny Youngman

Exercises for Section 4.

Question 1.

(a) Give an example of two *homeomorphic* metric spaces (X, d_X) and (Y, d_Y) such that X is complete, but Y is not complete.

(b) Why is this not an issue in Corollary 4.22? What is it about the "*nature*" of the homeomorphism between $\mathcal{W}$ and $\mathbb{K}^n$ in the proof of Proposition 4.20 that ensures that $\mathcal{W}$ is complete?

Question 2.

Let $(\mathcal{V}, \mathcal{T})$ be a TVS. Prove the following.

(a) If $C \subseteq \mathcal{V}$ is convex, then so is $\overline{C}$.

(b) If $E \subseteq \mathcal{V}$ is balanced, then so is $\overline{E}$.

Question 3.

Let $(\mathcal{V}, \mathcal{T})$ be a TVS, and let $\mathcal{W} \subseteq \mathcal{V}$ be a closed subspace. Prove that scalar multiplication is continuous in the quotient topology $\mathcal{T}_q$ on $\mathcal{V}/\mathcal{W}$.

5. Seminorms and locally convex spaces

The secret of life is honesty and fair dealing. If you can fake that, you've got it made.

Groucho Marx

5.1. Our main interest in topological vector spaces is to develop the theory of *locally convex* topological vector spaces, which appear naturally in defining certain weak topologies naturally associated with Banach spaces, including the Banach space of all bounded operators on a Hilbert space. Locally convex spaces are also the most general spaces for which (in our opinion) interesting versions of the Hahn-Banach Theorem will be shown to apply. As we shall see in this section, there is an intimate relation between locally convex topological vector spaces and separating families of seminorms on the underlying vector spaces, a phenomenon to which we now turn our attention.

5.2. Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. A **seminorm** on $\mathcal{V}$ is a map $p : \mathcal{V} \to \mathbb{R}$ satisfying*

(i) $p(x) \geq 0$ *for all* $x \in \mathcal{V}$;
(ii) $p(\lambda x) = |\lambda| p(x)$ *for all* $x \in \mathcal{V}$, $\lambda \in \mathbb{K}$;
(iii) $p(x + y) \leq p(x) + p(y)$ *for all* $x, y \in \mathcal{V}$.

It follows from this definition that a norm on $\mathcal{V}$ is simply a seminorm which satisfies the additional property that $p(x) = 0$ if and only if $x = 0$.

5.3. Remark. A few remarks are in order. If p is a seminorm on a vector space $\mathcal{V}$, then for all $x, y \in \mathcal{V}$,

$$p(x + y) \leq p(x) + p(y)$$

implies that

$$p(x + y) - p(y) \leq p(x).$$

Equivalently, with $z = x + y$, $p(z) - p(x) \leq p(z - x)$. Thus $p(x) - p(z) \leq p(x - z) = p(z - x)$. Hence

$$|p(x) - p(z)| \leq p(z - x).$$

5.4. Example. Let $\mathcal{V} = \mathcal{C}([0, 1], \mathbb{C})$. For each $x \in [0, 1]$, the map

$$p_x : \mathcal{V} \to \mathbb{R}$$

defined by setting $p_x(f) = |f(x)|$ is a seminorm on $\mathcal{V}$ which is not a norm.

5.5. Example. Let $n \geq 1$ and consider $\mathcal{V} = \mathbb{M}_n = \mathbb{M}_n(\mathbb{C})$. Fix $1 \leq k, l \leq n$. The map $\gamma_{kl} : \mathcal{V} \to \mathbb{R}$ defined by $\gamma_{kl}([x_{ij}]) = |x_{kl}|$ defines a seminorm on $\mathcal{V}$ which, once again, is not a norm.

5.6. Convexity. Recall that a subset E of a vector space $\mathcal{V}$ is said to be **convex** if $x, y \in E$ and $0 \le t \le 1$ imply $tx + (1-t)y \in E$. Geometrically, we are asking that the line segment between any two points in E must lie in E.

It is a simple but useful fact that any linear manifold of $\mathcal{V}$ is necessarily convex.

Note also that if $p_1, p_2, ..., p_m$ is a family of seminorms on a $\mathcal{V}$, $x_0 \in \mathcal{V}$, $\varepsilon > 0$ and $E = \{x \in \mathcal{V} : p_j(x - x_0) < \varepsilon, 1 \le j \le m\}$, then E is convex. Indeed, if $x, y \in E$ and $0 \le t \le 1$, then for all $1 \le j \le m$,

$$\begin{aligned} p_j(tx + (1-t)y - x_0) &= tp_j(x - x_0) + (1-t)p_j(y - x_0) \\ &< t\varepsilon + (1-t)\varepsilon = \varepsilon. \end{aligned}$$

Thus $tx + (1-t)y \in E$, and E is convex.

We leave it as an exercise for the reader to show that if $\mathcal{V}$ is a TVS and $E \subseteq \mathcal{V}$ is convex, then so is $\overline{E}$.

Another elementary but useful observation is that if $C \subseteq \mathcal{V}$ is convex and $T : \mathcal{V} \to \mathcal{W}$ is a linear map (where $\mathcal{W}$ is a second vector space), then $T(C)$ is convex as well. Finally, if $E \subseteq \mathcal{V}$ is convex, then for all $r, s > 0$, $rE + sE = (r+s)E$. Indeed, $\frac{r}{r+s}e_1 + \frac{s}{r+s}e_2 \in E$ for all $e_1, e_2 \in E$, from which the desired result easily follows.

5.7. The Minkowski functional. Let $\mathcal{V}$ be a TVS and suppose that $E \in \mathcal{U}_0^{\mathcal{V}}$ is convex. As we saw in the previous Chapter (see Remark 4.7), any nbhd of 0 in $\mathcal{V}$ is absorbing, and thus there exists $r_0 > 0$ so that $x \in r_0 E$. This allows us to define the map

$$\begin{array}{rccc} p_E : & \mathcal{V} & \to & \mathbb{R} \\ & x & \mapsto & \inf\{r \in (0, \infty) : x \in rE\}, \end{array}$$

which we call the **gauge functional** or the **Minkowski functional** for E.

Note: the name is misleading, since the map is clearly not linear - its range is contained in $[0, \infty)$. By convexity of E, if $x \in rE$ and $0 < r < s$, then $x = re$ for some $e \in E$, so $x = (1 - \frac{r}{s})0 + \frac{r}{s}(se) \in \operatorname{co}(sE) = sE$. In particular, $x \in sE$ for all $s > p_E(x)$.

5.8. Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. A function $p : \mathcal{V} \to \mathbb{R}$ is called a **sublinear functional** if it satisfies:*

(i) $p(x + y) \le p(x) + p(y)$ *for all* $x, y \in \mathcal{V}$, *and*
(ii) $p(rx) = rp(x)$ *for all* $0 < r \in \mathbb{R}$.

5.9. It is clear from the definition that every seminorm (and hence every norm) on a vector space is a sublinear functional on that space. The converse is false in general.

For example, the identity map $\kappa : \mathbb{R} \to \mathbb{R}$ is a sublinear functional on $\mathbb{R}$. It is not a seminorm since it is not even a non-negative valued function.

5.10. Proposition. *Let $\mathcal{W}$ be a TVS and $E \in \mathcal{U}_0$ be convex. Then*

(a) *The Minkowski functional p_E is a sublinear functional on $\mathcal{W}$ for E.*

(b) *If E is open, then*

$$E = \{x \in \mathcal{W} : p_E(x) < 1\}.$$

(c) *If E is balanced, then p_E is a seminorm.*

Proof.

(a) Suppose that $x, y \in E$ and that $r, s \in (0, \infty)$ with $r > p(x)$, $s > p(y)$. Then $x \in rE$, $y \in sE$ and so $x + y \in (r + s)E$. That is,

$$p(x + y) \leq r + s \text{ for all } r > p(x), s > p(y).$$

Thus $p(x + y) \leq p(x) + p(y)$.

Also, if $k > 0$, then $x \in rE$ if and only if $kx \in krE$, so that

$$\begin{aligned} p(kx) &= \inf\{s : kx \in sE\} \\ &= \inf\{kr : kx \in krE\} \\ &= k \inf\{r : x \in rE\} \\ &= kp(x). \end{aligned}$$

Thus p is a sublinear functional, as claimed.

(b) Suppose that $x \in E$ and that E is open. Since the map $f : \mathbb{R} \to \mathcal{W}$ given by $f(t) = tx$ is continuous, $1 \in f^{-1}(E)$ is open in $\mathbb{R}$ and therefore $(1 - \delta, 1 + \delta) \subseteq f^{-1}(E)$ for some $\delta > 0$. But then $(1 + \frac{\delta}{2})x \in E$, or equivalently, $x \in \frac{2}{2+\delta}E$, implying that $p(x) < \frac{2}{2+\delta} < 1$.

Conversely, suppose that $p(x) < 1$. Then $x = re$ for some $p(x) < r < 1$ and $e \in E$. But then $x = (1 - r)0 + re \in \text{co}E = E$.

(c) Suppose now that E is balanced. First observe that if $k \neq 0$, then $\frac{k}{|k|}E = E$.

Note that p is subadditive since it is a sublinear functional by (a). Also, $p(x) \geq 0$ for all $x \in \mathcal{W}$ by definition of p.

Finally, if $k = 0$, then $p(kx) = p(0)$. But $0 \in rE$ for all $r > 0$, and so

$$p(0x) = p(0) = \inf\{r > 0 : x \in rE\} = 0 = 0p(x).$$

If $k \neq 0$, then

$$\begin{aligned} p(kx) &= \inf\{r > 0 : kx \in rE\} \\ &= \inf\{s|k| > 0 : kx \in s(|k|E)\} \\ &= \inf\{s|k| > 0 : kx \in s(kE)\} \\ &= |k| \inf\{s > 0 : x \in sE\} \\ &= |k|p(x). \end{aligned}$$

Thus p is a seminorm.

□

5.11. Proposition. *Let $\mathcal{W}$ be a TVS and p be a seminorm on $\mathcal{W}$. The following are equivalent:*

(a) *p is continuous on $\mathcal{W}$;*
(b) *there exists a set $U \in \mathcal{U}_0^{\mathcal{W}}$ such that p is bounded above on U.*

Proof.

(a) implies (b): Consider the set
$$E := \{x \in \mathcal{W} : p(x) < 1\}.$$
Clearly $0 \in E$, and since p is assumed to be continuous and $E = p^{-1}(-\infty, 1)$, E is also open. Thus p is bounded above (by 1) on the open set $E \in \mathcal{U}_0^{\mathcal{W}}$.

(b) implies (a): Suppose that p is bounded above, say by $M > 0$ on an open set $U \in \mathcal{U}_0^{\mathcal{W}}$. Let $\varepsilon > 0$. If $x, y \in \mathcal{W}$ and $x - y \in (\frac{\varepsilon}{M})U$, say $x - y = \frac{\varepsilon}{M}u$ for some $u \in U$, then
$$|p(x) - p(y)| \leq p(x - y) = p(\frac{\varepsilon}{M}u) = \frac{\varepsilon}{M}p(u) < \varepsilon.$$
Thus p is (uniformly) continuous on $\mathcal{W}$.

□

5.12. Example. Recall from Example 5.4 that for each $x \in [0, 1]$,
$$\begin{array}{rccc} p_x : & \mathcal{C}([0,1],\mathbb{C}) & \to & \mathbb{R} \\ & f & \mapsto & |f(x)| \end{array}$$
is a seminorm. Now $B_1(0) := \{f \in \mathcal{C}([0,1],\mathbb{C}) : \|f\|_\infty < 1\}$ is open and $f \in B_1(0)$ implies that $p_x(f) = |f(x)| \leq \|f\|_\infty < 1$.

Thus each such p_x is continuous on $\mathcal{C}([0,1],\mathbb{C})$.

5.13. Definition. *A topology $\mathcal{T}$ on a topological vector space $\mathcal{W}$ is said to be* ***locally convex*** *if it admits a base consisting of convex sets. We shall write LCS for locally convex topological vector spaces, and for the sake of brevity, we shall refer to them as locally convex spaces.*

Since the topology on $\mathcal{W}$ is determined by the nbhds at a single point, it suffices to require that $\mathcal{W}$ admit a nbhd base at 0 consisting of convex sets; that is, given any nbhd $U \in \mathcal{U}_0$, there exists a convex nbhd $N \in \mathcal{U}_0$ so that $N \subseteq U$. In verifying that a space is a LCS, we shall often only verify this condition.

5.14. Proposition. *Let $\mathcal{W}$ be a TVS, and suppose that $U \in \mathcal{U}_0$ is convex. Then U contains a balanced, open, convex nbhd of 0.*

Proof. By Proposition 4.10, U contains a balanced, open nbhd H of 0. Set $N = \mathrm{co}(H)$. Then U convex and $H \subseteq U$ implies that $N \subseteq U$. Since H is balanced, a routine calculation shows that N is also balanced. For any choice of $t_1, t_2, ..., t_m \in [0, 1]$ with $\sum_{k=1}^m t_k = 1$, and for any $h_1, h_2, ..., h_m \in H$ we have
$$\left(\sum_{k=1}^{m-1} t_k h_k\right) + t_m H \subseteq N.$$

Since H is open, so is $\left(\sum_{k=1}^{m-1} t_k h_k\right) + t_m H$. Since $N = \text{co}(H)$, it follows that N is a union of open sets of this form, and hence N is also open.

Thus N is an open, balanced, convex nbhd of 0 contained in U, the existence of which proves our claim.

□

As an immediate consequence we obtain:

5.15. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS. Then $\mathcal{V}$ admits a nbhd base at 0 consisting of balanced, open, convex sets.*

5.16. Example. Let $(\mathfrak{X}, \|\cdot\|)$ be a normed linear space. For each $\varepsilon > 0$, the argument of paragraph 5.6 shows that $B_\varepsilon(0) = \{x \in \mathfrak{X} : \|x\| < \varepsilon\}$ is convex. Since $\{B_\varepsilon(0) : \varepsilon > 0\}$ is a nbhd base at 0 for the norm topology, $(\mathfrak{X}, \|\cdot\|)$ is a LCS.

More concretely, $(\mathbb{K}^n, \|\cdot\|_2)$ is a LCS, as is any Hilbert space $\mathcal{H}$. So is $\mathcal{B}(\mathcal{H})$.

We have already seen that the quotient of a TVS by one of its closed subspaces is a TVS. Let us first obtain the same result for locally convex spaces.

5.17. Proposition. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\mathcal{W} \subseteq \mathcal{V}$ be a closed subspace. Then $\mathcal{V}/\mathcal{W}$ is a LCS in the quotient topology.*

Proof. As mentioned above, that $\mathcal{V}/\mathcal{W}$ is a TVS follows from paragraph 4.18. There remains only to show that $\mathcal{V}/\mathcal{W}$ admits a nbhd base at 0 consisting of convex sets (see the remarks following Definition 5.13).

Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ denote the canonical quotient map, and let $U \in \mathcal{U}_0^{\mathcal{V}/\mathcal{W}}$. Then $q^{-1}(U) \in \mathcal{U}_0^{\mathcal{V}}$, as q is continuous. Since $\mathcal{V}$ is a LCS, we can find a convex nbhd $N \in \mathcal{U}_0^{\mathcal{V}}$ so that $0 \in N \subseteq q^{-1}(U)$. Let $M = q(N)$. Since q is an open map, we have $M \in \mathcal{U}_0^{\mathcal{V}/\mathcal{W}}$. Since q is linear, M is convex.

Finally, since $N \subseteq q^{-1}(U)$, $M = q(N) \subseteq U$, and we are done.

□

5.18. Definition. *A family Γ of seminorms on a vector space $\mathcal{W}$ is said to be* ***separating*** *if for all $0 \neq x \in \mathcal{W}$ there exists $p \in \Gamma$ so that $p(x) \neq 0$.*

5.19. Example. Let $\mathcal{W} = \mathcal{C}([0,1], \mathbb{C})$ and consider $\Gamma = \{p_x : x \in \mathbb{Q} \cap [0,1]\}$, where - as before - $p_x(f) = |f(x)|$ for all $f \in \mathcal{W}$.

If $0 \neq f \in \mathcal{W}$, then there exists $y \in [0,1]$ so that $f(y) \neq 0$. By continuity of f, there exists a nbhd N of y such that $f(y) \neq 0$ for all $y \in N$. Thus there exists a rational number $q \in N$ so that $0 \neq f(q)$ and hence $p_q(f) \neq 0$. Thus Γ is a separating family of seminorms.

5.20. Let Γ be a family of seminorms on a vector space $\mathcal{W}$. For $F \subseteq \Gamma$ finite, $x \in \mathcal{W}$ and $\varepsilon > 0$, set

$$N(x, F, \varepsilon) = \{y \in \mathcal{W} : p(x - y) < \varepsilon, p \in F\}.$$

Permitting ourselves a slight abuse of notation, we shall write $N(x, p, \varepsilon)$ in the case where $F = \{p\}$.

5.21. Theorem. *If Γ is a separating family of seminorms on a vector space $\mathcal{W}$, then*

$$\mathcal{B} = \{N(x, F, \varepsilon) : x \in \mathcal{W}, \varepsilon > 0, F \subseteq \Gamma \text{ finite}\}$$

is a base for a locally convex topology $\mathcal{T}$ on $\mathcal{W}$. Moreover, each $p \in \Gamma$ is $\mathcal{T}$-continuous.

Proof.

Step One: We begin by showing that $\mathcal{B}$ is a base for a Hausdorff topology $\mathcal{T}$ on $\mathcal{W}$.

- Let $x \in \mathcal{W}$ and choose $0 \neq p \in \Gamma$. (Such a p exists since Γ is assumed to be separating.) Then $x \in N(x, p, 1)$. Thus
$$\cup\{B : B \in \mathcal{B}\} \supseteq \cup\{N(x, p, 1) : x \in \mathcal{W}\} = \mathcal{W}.$$
- Next suppose that $B_1 = N(x, F_1, \varepsilon_1)$ and $B_2 = N(y, F_2, \varepsilon_2)$ lie in $\mathcal{B}$ and that $z \in B_1 \cap B_2$. We must find $B_3 \in \mathcal{B}$ so that $z \in B_3 \subseteq B_1 \cap B_2$.

  To that end, let $\varepsilon = \min\{\varepsilon_1 - p(x - z), \varepsilon_2 - q(y - z) : p \in F_1, q \in F_2\}$, so that $\varepsilon > 0$. If $w \in N(z, F_1 \cup F_2, \varepsilon)$, then
$$p(w - x) \leq p(w - z) + p(z - x) < \varepsilon + p(z - x) \leq \varepsilon_1$$
  for all $p \in F_1$, and so $w \in B_1$. An analogous argument proves that $w \in B_2$. That is, $B_3 := N(z, F_1 \cup F_2, \varepsilon)$ satisfies the required condition.

  It now follows from our work in the homework assignments that $\mathcal{B}$ is a base for a topology on $\mathcal{W}$.
- If $x, y \in \mathcal{W}$ and $x \neq y$, then our assumption that Γ is separating implies the existence of an element $p \in \Gamma$ so that $\delta := p(x - y) > 0$. But then $N(x, p, \frac{\delta}{2})$ and $N(y, p, \frac{\delta}{2})$ are disjoint nbhds of x and y respectively in the $\mathcal{T}$-topology, proving that $\mathcal{T}$ is Hausdorff.

Step Two: That $\mathcal{T}$ is locally convex follows readily from the fact that $\mathcal{B}$ is a base for $\mathcal{T}$ and each $N(x, F, \varepsilon)$ is itself convex, as is easily verified.

Step Three: Next we verify that $(\mathcal{W}, \mathcal{T})$ is a TVS; namely, that the topology $\mathcal{T}$ is compatible with the vector space operations.

- Suppose that $x_0, y_0 \in \mathcal{W}$ and let U be a nbhd of $x_0 + y_0$ in the $\mathcal{T}$-topology. Then there exists a basic nbhd $B = N(x_0 + y_0, F, \varepsilon)$ of $x_0 + y_0$ with $B \subseteq U$. Let $B_1 = N(x_0, F, \frac{\varepsilon}{2})$ and $B_2 = N(y_0, F, \frac{\varepsilon}{2})$. If $(x, y) \in B_1 \times B_2$, then
$$p\big((x + y) - (x_0 + y_0)\big) \leq p(x - x_0) + p(y - y_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$
  for all $p \in F$, and thus $\sigma(B_1 \times B_2) \subseteq B \subseteq U$. This shows that addition is continuous relative to $\mathcal{T}$.

- As for scalar multiplication, let $\lambda_0 \in \mathbb{K}$, $x_0 \in \mathcal{W}$ and U be a nbhd of $\lambda_0 x_0$ in the $\mathcal{T}$-topology. As before, choose a basic nbhd $B = N(\lambda_0 x_0, F, \varepsilon) \subseteq U$. Let $\delta > 0$. If $K := \{\lambda \in \mathbb{C} : |\lambda - \lambda_0| < \delta\}$ and $B = N(x_0, F, \delta)$, then $(\lambda, x) \in K \times B$ implies that

$$\begin{aligned} p(\lambda x - \lambda_0 x_0) &\le p(\lambda x - \lambda x_0) + p(\lambda x_0 - \lambda_0 x_0) \\ &\le |\lambda| p(x - x_0) + |\lambda - \lambda_0| p(x_0) \\ &< (|\lambda_0| + \delta)\delta + \delta p(x_0) \end{aligned}$$

  for all $p \in F$. Since F is finite, it is clear that δ can be chosen such that $p(\lambda x - \lambda_0 x_0) < \varepsilon$, $p \in F$, which proves that scalar multiplication is also continuous relative to $\mathcal{T}$.

Together, these two observations prove that $(\mathcal{W}, \mathcal{T})$ is a TVS.

Step Four: Finally, let us show that each p is continuous relative to $\mathcal{T}$.

If $p = 0$, then clearly p is continuous relative to $\mathcal{T}$.

Otherwise, let $B = N(0, p, 1)$. Then $B \in \mathcal{T}$, and for $x \in B$,

$$p(x) = p(x - 0) < 1,$$

so that p is bounded on some open set in $\mathcal{W}$. It now follows from Proposition 5.11 that p is (uniformly) continuous on $\mathcal{W}$.

□

5.22. The above result says that a separating family of seminorms on a vector space $\mathcal{W}$ gives rise to a locally convex topology on $\mathcal{W}$. Our next goal is to show that all locally convex spaces arise in this manner.

5.23. Theorem. *Suppose that $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ is a LCS. Then there exists a separating family Γ of seminorms on $\mathcal{V}$ which generate the topology $\mathcal{T}_\mathcal{V}$.*

Proof. By Corollary 5.15, $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ admits a nbhd base $\mathcal{C}_0$ at 0 consisting of balanced, open, convex sets. By Proposition 5.10, for each $E \in \mathcal{C}_0$, the Minkowski functional p_E is a seminorm and $E = \{x \in \mathcal{V} : p_E(x) < 1\}$.

Let $\Gamma = \{p_E : E \in \mathcal{C}_0\}$. We first show that Γ is separating. Indeed, suppose that $0 \ne x \in \mathcal{V}$. Since $\mathcal{T}_\mathcal{V}$ is Hausdorff by hypothesis, there exists $G \in \mathcal{C}_0$ so that $x \notin G$ (this is actually a bit weaker than the statement that $\mathcal{T}_\mathcal{V}$ is Hausdorff, but certainly implied by it). Since $G \in \mathcal{C}_0$, $p_G \in \Gamma$. But $x \notin G$ implies that $p_G(x) \ge 1$, and hence $p_G(x) \ne 0$. Thus Γ is separating. This is required before passing to the next step.

By Theorem 5.21,

$$\mathcal{B} = \{N(x, F, \varepsilon) : x \in \mathcal{V}, \varepsilon > 0, F \subseteq \Gamma \text{ finite}\}$$

is a base for a locally convex topology $\mathcal{T}_\Gamma$ on $\mathcal{V}$. Our goal, of course, is to prove that $\mathcal{T}_\mathcal{V} = \mathcal{T}_\Gamma$.

Let $E \in \mathcal{C}_0$ be a $\mathcal{T}_\mathcal{V}$-open, balanced, convex nbhd of 0. Since $E = N(0, p_E, 1) \in \mathcal{B}$, it follows that $\mathcal{T}_\Gamma$ contains a nbhd base at 0 for the topology $\mathcal{T}_\mathcal{V}$. Since both topologies are TVS-topologies, they are determined by their nbhd bases at any point (for eg., at 0), and from this it follows that $\mathcal{T}_\Gamma \supseteq \mathcal{T}_\mathcal{V}$.

On the other hand, each $p_E \in \Gamma$ is bounded above by 1 on E, and E is a $\mathcal{T}_\mathcal{V}$-open nbhd of 0. By Proposition 5.11, p_E is continuous on $(\mathcal{V}, \mathcal{T}_\mathcal{V})$. It follows that $N(0, p_E, \varepsilon) = p_E^{-1}(-\varepsilon, \varepsilon) \in \mathcal{T}_\mathcal{V}$ for all $\varepsilon > 0$. Thus $\mathcal{T}_\mathcal{V}$ contains a nbhd subbase for $\mathcal{T}_\Gamma$ at 0, and arguing as before, we get that $\mathcal{T}_\Gamma \subseteq \mathcal{T}_\mathcal{V}$.

Hence $\mathcal{T}_\mathcal{V} = \mathcal{T}_\Gamma$, and the topology $\mathcal{T}_\mathcal{V}$ is determined by the family Γ of seminorms.

□

5.24. Example. Let $(\mathfrak{X}, \|\cdot\|)$ be a normed linear space. The norm topology on $\mathfrak{X}$ is the metric topology induced by the metric $d(x, y) = \|x - y\|$. That is, a nbhd base at $x_0 \in \mathfrak{X}$ for the norm topology is

$$\begin{aligned}\mathcal{B}_{x_0} &= \{V_\varepsilon(x_0) : \varepsilon > 0\} \\ &= \{\{y \in \mathfrak{X} : \|y - x_0\| < \varepsilon\} : \varepsilon > 0\} \\ &= \{N(x_0, \|\cdot\|, \varepsilon) : \varepsilon > 0\}.\end{aligned}$$

Thus we see that the norm topology on $\mathfrak{X}$ is exactly the locally convex topology generated by $\Gamma = \{\|\cdot\|\}$. Observe that since $\|\cdot\|$ is a norm, $0 \neq x \in \mathfrak{X}$ implies that $\|x\| \neq 0$, and thus Γ is indeed separating, as required.

5.25. In Corollary 5.15, we saw that any LCS $(\mathcal{V}, \mathcal{T})$ admits a nbhd base at 0 consisting of open, balanced, convex sets.

In fact, each $N(0, \{p_1, p_2, ..., p_m\}, \varepsilon)$ is balanced, open and convex for all choices of $m \geq 1$, $p_1, p_2, ..., p_m \in \Gamma$ and $\varepsilon > 0$, where Γ is a separating family of seminorms which generate $\mathcal{T}$. It is clear from Theorems 5.21 and 5.23 that the collection of such sets is a nbhd base at 0 for $\mathcal{T}$.

Having generated a topology on a vector space using a separating family of seminorms, let us now examine what it means for a net to converge in this topology.

5.26. Proposition. *Let $\mathcal{V}$ be a vector space and Γ be a separating family of seminorms on $\mathcal{V}$. Let $\mathcal{T}$ denote the locally convex topology on $\mathcal{V}$ generated by Γ.*

A net $(x_\lambda)_\lambda$ in $\mathcal{V}$ converges to a point $x \in \mathcal{V}$ if and only if

$$\lim_\lambda p(x - x_\lambda) = 0 \text{ for all } p \in \Gamma.$$

Proof.

- Suppose first that $(x_\lambda)_\lambda$ converges to x in the $\mathcal{T}$-topology. Given $p \in \Gamma$ and $\varepsilon > 0$, the set $N(x, p, \varepsilon) \in \mathcal{T}$ and so there exists λ_0 so that $\lambda \geq \lambda_0$ implies that $x_\lambda \in N(x, p, \varepsilon)$. That is, $\lambda \geq \lambda_0$ implies that $p(x - x_\lambda) < \varepsilon$. Thus $\lim_\lambda p(x - x_\lambda) = 0$.

  Alternatively, one may argue as follows: suppose that $(x_\lambda)_\lambda$ converges to x in the $\mathcal{T}$-topology. Given $p \in \Gamma$, we know that p is continuous in the $\mathcal{T}$-topology by Theorem 5.21. Since $\lim_\lambda x - x_\lambda = 0$,

$$\lim_\lambda p(x - x_\lambda) = p(\lim_\lambda (x - x_\lambda)) = p(0) = 0.$$

- Conversely, suppose that $\lim_\lambda p(x - x_\lambda) = 0$ for all $p \in \Gamma$. Let $U \in \mathcal{U}_x$ is the $\mathcal{T}$-topology. Then there exist $p_1, p_2, ..., p_m \in \Gamma$ and $\varepsilon > 0$ so that $N(x, \{p_1, p_2, ..., p_m\}, \varepsilon) \subseteq U$. For each $1 \le j \le m$, choose λ_j so that $\lambda \ge \lambda_j$ implies that $p_j(x_\lambda - x) < \varepsilon$. Choose $\lambda_0 \ge \lambda_1, \lambda_2, ..., \lambda_m$. If $\lambda \ge \lambda_0$, then $p_j(x_\lambda - x) < \varepsilon$ for all $1 \le j \le m$ so that $x_\lambda \in N(x, \{p_1, p_2, ..., p_m\}, \varepsilon) \subseteq U$. Hence $\lim_\lambda x_\lambda = x$ in $(\mathcal{V}, \mathcal{T})$.

$\square$

5.27. Remarks. Let $\mathcal{V}$ be a vector space as above and let Γ be a separating family of seminorms on $\mathcal{V}$. Recall that if $\mathcal{T}_w$ is the weak topology on $\mathcal{V}$ induced by Γ, then $\mathcal{T}_w$ is the weakest topology for which each of the functions $p \in \Gamma$ is continuous. By Theorem 5.21, the LCS topology $\mathcal{T}$ generated by

$$\mathcal{B} = \{N(x, F, \varepsilon) : x \in \mathcal{V}, \varepsilon > 0, F \subseteq \Gamma \text{ finite}\}$$

has the property that each $p \in \Gamma$ is continuous on $(\mathcal{V}, \mathcal{T})$. It follows, therefore, that $\mathcal{T}_w \subseteq \mathcal{T}$. In other words, if $(x_\lambda)_\lambda$ is a net in $(\mathcal{V}, \mathcal{T})$ which converges to $x \in \mathcal{V}$, then $(x_\lambda)_\lambda$ converges to x in $(\mathcal{V}, \mathcal{T}_w)$; i.e. $\lim_\lambda p(x_\lambda) = p(x)$ for all $p \in \Gamma$.

That these two topologies do not, in general, coincide can be seen by examining a simple example.

Let $\mathcal{V} = \mathbb{K}$ and let $\Gamma = \{p\}$, where $p(x) = |x|$ for each $x \in \mathbb{K}$. The LCS topology on $\mathbb{K}$ generated by Γ is a TVS topology, and thus must agree with the usual topology on $\mathbb{K}$, since the latter admits a unique TVS topology, by Lemma 4.19. The weak topology $\mathcal{T}_w$ on $\mathbb{K}$ generated by Γ is the weakest topology for which p is continuous. In particular, a net $(x_\lambda)_\lambda$ converges to $x \in \mathbb{K}$ if and only if $\lim_\lambda |x_\lambda| = |x|$. For example, the sequence $(x_n)_n$, where $x_n = (-1)^n$, $n \ge 1$ converges to $x = 1$ in $(\mathbb{K}, \mathcal{T}_w)$. Since it clearly doesn't converge in $(\mathbb{K}, \mathcal{T})$, the two topologies are necessarily different, and again – by Theorem 5.21 – it follows that $(\mathbb{K}, \mathcal{T}_w)$ is not a TVS.

There is, however, a situation where we can say a bit more than this. Let $\mathcal{V}$ be a vector space and let $(\mathfrak{X}_\alpha, \|\cdot\|_\alpha)_{\alpha \in A}$ be a collection of Banach spaces (in fact, normed linear spaces will do). For each such α, suppose that $T_\alpha : \mathcal{V} \to \mathfrak{X}_\alpha$ is a linear map. Suppose furthermore that the family $\{T_\alpha\}_\alpha$ is **separating** in the sense that if $0 \ne x \in \mathcal{V}$, then there exists $\alpha \in A$ so that $0 \ne T_\alpha x \in \mathfrak{X}_\alpha$. Then each of the functions

$$\begin{array}{rccc} p_\alpha : & \mathcal{V} & \to & [0, \infty) \\ & x & \mapsto & \|T_\alpha x\|_\alpha \end{array}$$

is easily seen to be a seminorm. It is routine to verify that the fact that $\{T_\alpha\}_\alpha$ is separating implies that $\Gamma = \{p_\alpha\}_\alpha$ is a separating family of seminorms. Let $\mathcal{T}$ denote the LCS topology on $\mathcal{V}$ generated by Γ. By Proposition 5.26, a net $(x_\lambda)_\lambda$ converges to $x \in (\mathcal{V}, \mathcal{T})$ if and only if

$$\lim_\lambda p_\alpha(x - x_\lambda) = \lim_\lambda \|T_\alpha(x - x_\lambda)\|_\alpha = 0 \text{ for all } \alpha \in A.$$

That is, $\lim_\lambda x_\lambda = x$ if and only if

$$\lim_\lambda T_\alpha x_\lambda = T_\alpha x \text{ for all } \alpha \in A.$$

Since this is nothing more than the statement that each T_α is continuous, we find that in this case, the $\mathcal{T}$ topology on $\mathcal{V}$ coincides with the weak topology generated by the family $\{T_\alpha\}_{\alpha \in A}$. This is still not the same as the weak topology generated by the family Γ, however.

5.28. Example. Let $\mathcal{H} = \ell^2(\mathbb{N})$ and recall that $\mathcal{H}$ is a Hilbert space when equipped with the inner product $\langle (x_n), (y_n) \rangle = \sum_{n=1}^\infty x_n \overline{y_n}$.

Recall also that $\mathcal{B}(\mathcal{H})$ is a normed linear space with the operator norm $\|T\| := \sup\{\|Tx\| : x \in \mathcal{H}, \|x\| \le 1\}$.

From above, we see that the norm topology on $\mathcal{B}(\mathcal{H})$ admits as a nbhd base at $T \in \mathcal{B}(\mathcal{H})$ the collection

$$\{N(T, \|\cdot\|, \varepsilon) : \varepsilon > 0\} = \{V_\varepsilon(T) : \varepsilon > 0\},$$

and that this is the locally convex topology generated by the separating family $\Gamma = \{\|\cdot\|\}$ of (semi)norms.

Convergence of a net of operators $(T_\lambda)_\lambda$ to $T \in \mathcal{B}(\mathcal{H})$ in the norm topology (i.e. $\lim_\lambda \|T_\lambda - T\| = 0$) should be thought of as uniform convergence on the closed unit ball of $\mathcal{H}$.

This is certainly not the only interesting topology one can impose upon $\mathcal{B}(\mathcal{H})$. Let us first consider the topology of "pointwise convergence".

THE STRONG OPERATOR TOPOLOGY (SOT) For each $x \in \mathcal{H}$, consider

$$\begin{array}{rccc} p_x: & \mathcal{B}(\mathcal{H}) & \to & \mathbb{R} \\ & T & \mapsto & \|Tx\|. \end{array}$$

Then

(i) $p_x(T) \ge 0$ for all $T \in \mathcal{B}(\mathcal{H})$;
(ii) $p_x(\lambda T) = \|\lambda T x\| = |\lambda|\ \|Tx\| = |\lambda|\ p_x(T)$ for all $\lambda \in \mathbb{K}$;
(iii) $p_x(T_1 + T_2) = \|T_1 x + T_2 x\| \le \|T_1 x\| + \|T_2 x\| = p_x(T_1) + p_x(T_2)$,

so that p_x is a seminorm on $\mathcal{B}(\mathcal{H})$ for each $x \in \mathcal{H}$.

In general, p_x is not a norm because we can always find $T \in \mathcal{B}(\mathcal{H})$ so that $0 \ne T$ but $p_x(T) = 0$. Indeed, let $y \in \mathcal{H}$ with $0 \ne y$ and $y \perp x$. Define $T_y : \mathcal{H} \to \mathcal{H}$ via $T_y(z) = \langle z, y \rangle y$. Then $\|T_y(z)\| \le \|z\|\ \|y\|^2$ by the Cauchy-Schwarz Inequality and in particular $T_y(y) = \|y\|^2 y \ne 0$, but $T_y(x) = \langle x, y \rangle y = 0y = 0$. Thus $0 \ne T_y$ but $p_x(T_y) = 0$.

On the other hand, if $0 \ne T \in \mathcal{B}(\mathcal{H})$, then there exists $x \in \mathcal{H}$ so that $Tx \ne 0$. Thus $p_x(T) = \|Tx\| \ne 0$, proving that $\Gamma_{SOT} := \{p_x : x \in \mathcal{H}\}$ separates the points of $\mathcal{B}(\mathcal{H})$.

The locally convex topology on $\mathcal{B}(\mathcal{H})$ generated by Γ_{SOT} is called the **strong operator topology** and is denoted by SOT.

By Proposition 5.26 above, we see that a net $(T_\lambda)_\lambda \in \mathcal{B}(\mathcal{H})$ converges to $T \in \mathcal{B}(\mathcal{H})$ in the SOT if and only if

$$\lim_\lambda p_x(T_\lambda - T) = \lim_\lambda \|T_\lambda x - Tx\| = 0 \quad \text{for all } x \in \mathcal{H}.$$

Thus the SOT is the topology of pointwise convergence. That is, it is the weakest topology that makes all of the evaluation maps $T \mapsto Tx$, $x \in \mathcal{H}$ continuous.

A nbhd base for the SOT at the point $T \in \mathcal{B}(\mathcal{H})$ is given by the collection

$$\{N(T, \{x_1, x_2, ..., x_m\}, \varepsilon) : m \ge 1, x_j \in \mathcal{H}, 1 \le j \le m, \varepsilon > 0\}$$

where, for $m \ge 1$, $F := \{x_j \in \mathcal{H} : 1 \le j \le m\}$ and $\varepsilon > 0$, we have

$$N(T, F, \varepsilon) = \{R \in \mathcal{B}(\mathcal{H}) : \|Rx_j - Tx_j\| < \varepsilon,\ 1 \le j \le m\}.$$

THE WEAK OPERATOR TOPOLOGY (WOT) Next, for each pair $(x, y) \in \mathcal{H} \times \mathcal{H}$, consider the map

$$\begin{array}{rccc} q_{x,y}: & \mathcal{B}(\mathcal{H}) & \to & \mathbb{R} \\ & T & \mapsto & |\langle Tx, y\rangle|. \end{array}$$

Again, it is routine to verify that each $q_{x,y}$ is a seminorm but not a norm on $\mathcal{B}(\mathcal{H})$.

The locally convex topology on $\mathcal{B}(\mathcal{H})$ generated by $\Gamma_{WOT} := \{q_{x,y} : (x, y) \in \mathcal{H} \times \mathcal{H}\}$ is called the **weak operator topology** on $\mathcal{B}(\mathcal{H})$ and is denoted by WOT.

A net $(T_\lambda)_\lambda \in \mathcal{B}(\mathcal{H})$ converges to $T \in \mathcal{B}(\mathcal{H})$ in the WOT if and only if

$$\lim_\lambda |\langle (T_\lambda - T)x, y\rangle| = \lim_\lambda |\langle T_\lambda x, y\rangle - \langle Tx, y\rangle| = 0$$

for all $x, y \in \mathcal{H}$. In other words, the WOT is the weakest topology that makes all of the functions $T \mapsto \langle Tx, y\rangle$, $x, y \in \mathcal{H}$ continuous.

A nbhd base for the WOT at the point $T \in \mathcal{B}(\mathcal{H})$ is given by the collection

$$\{N(T, \{x_1, x_2, ..., x_m, y_1, y_2, ..., y_m\}, \varepsilon) : m \ge 1, x_j, y_j \in \mathcal{H}, 1 \le j \le m, \varepsilon > 0\},$$

where, for $m \ge 1$, $F := \{(x_j, y_j) \in \mathcal{H} \times \mathcal{H} :\ 1 \le j \le m\}$ and $\varepsilon > 0$, we have

$$N(T, F, \varepsilon) = \{R \in \mathcal{B}(\mathcal{H}) : |\langle Rx_j - Tx_j, y_j\rangle| < \varepsilon,\ 1 \le j \le\ m\}.$$

5.29. Proposition. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS, and let Γ be a separating family of seminorms on $\mathcal{V}$ which generate the locally convex topology on $\mathcal{V}$. Let p be a seminorm on $\mathcal{V}$. The following are equivalent:*

(a) *p is continuous on $\mathcal{V}$;*

(b) *there exists a constant $\kappa > 0$ and $p_1, p_2, ..., p_m \in \Gamma$ so that*

$$p(x) \le \kappa \max(p_1(x), p_2(x), ..., p_m(x)) \qquad \textit{for all } x \in \mathcal{V}.$$

Proof.

(a) implies [(b)] Suppose that p is continuous on $\mathcal{V}$. Then $M := p^{-1}((-1, 1)) = p^{-1}([0, 1))$ is a $\mathcal{T}$-open nbhd of 0, and as such, it must contain a basic nbhd $N := N(0, \{p_1, p_2, ..., p_m\}, \varepsilon)$ for some $p_1, p_2, ..., p_m \in \Gamma$ and $\varepsilon > 0$. It follows that if $p_j(x) < \varepsilon$ for $1 \le j \le m$, then $x \in N \subseteq M$, and hence $p(x) < 1$.

More generally, let $y \in \mathcal{V}$ and let $r_y = \max(p_1(y), p_2(y), ..., p_m(y))$.

- If $r_y = 0$, then for all $k > 0$, $p_j(ky) = 0 < \varepsilon$, $1 \le j \le m$, so that from above, $p(ky) = kp(y) < 1$. But then
$$p(y) = 0 \le 1 \max(p_1(y), p_2(y), ..., p_m(y)).$$
- If $r_y > 0$, then $x = \frac{\varepsilon}{2r_y} y$ satisfies $p_j(x) < \varepsilon$, $1 \le j \le m$, and so $\frac{\varepsilon}{2r_y} p(y) = p(x) < 1$. That is,
$$p(y) < \frac{2r_y}{\varepsilon} = \frac{2}{\varepsilon} \max(p_1(y), p_2(y), ..., p_m(y)).$$

We conclude that with $\kappa = \max(1, \frac{2}{\varepsilon})$,
$$p(y) \le \kappa \max(p_1(y), p_2(y), ..., p_m(y))$$
for all $y \in \mathcal{V}$.

(b) implies [(a)] Suppose that (b) holds. Now $N := N(0, \{p_1, p_2, ..., p_m\}, 1)$ is an open nbhd of 0 in the $\mathcal{T}$-topology. If $x \in N$, then $p_j(x) < 1$ for all $1 \le j \le m$, and so $p(x) \le \kappa$. But then p is bounded above on the $\mathcal{T}$-open nbhd N of 0, and hence is continuous by Proposition 5.11.

□

5.30. Proposition. *Let $(\mathcal{V}, \mathcal{T}_{\mathcal{V}})$ and $(\mathcal{W}, \mathcal{T}_{\mathcal{W}})$ be locally convex spaces. Let $\Gamma_{\mathcal{V}}$ and $\Gamma_{\mathcal{W}}$ denote separating families of seminorms which generate the corresponding locally convex topologies on $\mathcal{V}$ and $\mathcal{W}$ respectively. Finally, let $T : \mathcal{V} \to \mathcal{W}$ be a linear map.*

The following are equivalent:

(a) *T is continuous.*

(b) *For all $q \in \Gamma_{\mathcal{W}}$ there exists $\kappa > 0$ and $p_1, p_2, ..., p_m \in \Gamma_{\mathcal{V}}$ so that*
$$q(Tx) \le \kappa \max(p_1(x), p_2(x), ..., p_m(x)) \quad \text{for all } x \in \mathcal{V}.$$

Proof.

(a) implies (b): Suppose that T is continuous and that $q \in \Gamma_{\mathcal{W}}$. Clearly q is continuous as well. It is routine to verify that $q \circ T$ is a seminorm on $\mathcal{V}$. Since the composition of continuous functions is continuous, $q \circ T$ is a continuous seminorm on $\mathcal{V}$, and the result now follows from Proposition 5.29.

(b) implies (a): Conversely, suppose that for all $q \in \Gamma_{\mathcal{W}}$ there exists $\kappa > 0$ and $p_1, p_2, ..., p_m \in \Gamma_{\mathcal{V}}$ so that
$$q(Tx) \le \kappa \max(p_1(x), p_2(x), ..., p_m(x)) \quad \text{for all } x \in \mathcal{V}.$$

As before, we observe that $q \circ T$ is a seminorm on $\mathcal{V}$ for all $q \in \Gamma_{\mathcal{W}}$. Moreover, by Proposition 5.29, each such $q \circ T$ is continuous.

Let $U \in \mathcal{U}_0^{\mathcal{W}}$ and choose $q_1, q_2, ..., q_n \in \Gamma_{\mathcal{W}}$ so that $N(0, \{q_1, q_2, ..., q_n\}, \varepsilon) \subseteq U$. Since each $q_j \circ T$ is continuous on $\mathcal{V}$, we have that $N(0, \{q_1 \circ T, q_2 \circ T, ..., q_n \circ T\}, \varepsilon)$ is a nbhd of 0 in $\mathcal{V}$. Moreover,
$$x \in N(0, \{q_1 \circ T, q_2 \circ T, ..., q_n \circ T\}, \varepsilon)$$

implies that

$$Tx \in N(0, \{q_1, q_2, ..., q_n\}, \varepsilon) \subseteq U.$$

It follows that T is continuous at 0.

By Theorem 4.33 and paragraph 4.34, T is continuous on $\mathcal{V}$.

□

We shall require the following special case of the above result.

5.31. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS. A linear functional f on $\mathcal{V}$ is continuous if and only if there exists a continuous seminorm p on $\mathcal{V}$ such that*

$$|f(x)| \le p(x) \qquad \text{for all } x \in \mathcal{V}.$$

Proof. Observe that if f is a continuous linear functional on $\mathcal{V}$, then $p(x) := |f(x)|$, $x \in \mathcal{V}$ defines a continuous seminorm on $\mathcal{V}$; indeed, that p is continuous follows from the fact that f is continuous on $\mathcal{V}$ and $|\cdot|$ is continuous on $\mathbb{K}$ respectively. Obviously

$$|f(x)| \le p(x) \qquad \text{for all } x \in \mathcal{V}.$$

Conversely – and more interestingly – suppose that there exists a continuous seminorm p on $\mathcal{V}$ such that $|f(x)| \le p(x)$ for all $x \in \mathcal{V}$. As before, we may choose a separating family Γ of seminorms on $\mathcal{V}$, and without loss of generality, we may assume that $p \in \Gamma$. (Otherwise we replace Γ by $\Gamma \cup \{p\}$.). The result now follows immediately from Proposition 5.30.

□

Appendix to Section 5.

5.32. In the assignment questions we exhibited an example of a TVS which is not normable, i.e. it is not a normed linear space with respect to any norm. The technique for constructing that example can be extended to produce a large variety of such examples. The spaces we have in mind are called **Fréchet spaces**, and we define them now.

5.33. Definition. *A metric d on a vector space $\mathcal{V}$ is said to be **translation invariant** if*

$$d(x, y) = d(x + z, y + z)$$

for all $x, y, z \in \mathcal{V}$.

*We shall also say that a metric d on $\mathcal{V}$ is **complete** if $(\mathcal{V}, d)$ is a complete metric space.*

*Finally, let us say that a countable family $\{\rho_n\}_n$ of pseudo-metrics on $\mathcal{V}$ is **complete** if, whenever $(x_k)_k$ is a sequence in $\mathcal{V}$ which is Cauchy relative to each ρ_n (i.e. for all $n \ge 1$ and $\varepsilon > 0$ there exists $N = N(\varepsilon, n) > 0$ so that $j, k \ge N$ implies $\rho_n(x_j, x_k) < \varepsilon$), there exists $x \in \mathcal{V}$ so that $\lim_{k\to\infty} \rho_n(x_k, x) = 0$ for each $n \ge 1$.*

5.34. Example. Most, but certainly not all metrics we deal with are translation invariant. For example, if $d(x, y) = |x-y|$ for $x, y \in \mathbb{R}$, then d is obviously translation invariant.

On the other hand, the metric d on $\mathbb{R}$ defined via:

$$d(x, y) = |x^3 - y^3|$$

for all $x, y \in \mathbb{R}$ is not translation invariant, since $d(0, 1) = 1 \ne 7 = d(1, 2)$.

5.35. Definition. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS. If the topology $\mathcal{T}$ on $\mathcal{V}$ is induced by a translation invariant, complete metric d, then we say that $(\mathcal{V}, \mathcal{T})$ is a **Fréchet space**.*

5.36. Constructing Fréchet spaces. We know from Theorem 5.21 that if Γ is a separating family of seminorms on a vector space $\mathcal{V}$, then Γ generates a LCS topology $\mathcal{T}$ on $\mathcal{V}$. Suppose now that the family Γ possesses the following two additional properties, namely:

- the set $\Gamma = \{p_n\}_n$ is countable, and
- the family $\{\rho_n\}_n$ of pseudo-metrics defined on $\mathcal{V}$ via $\rho_n(x, y) = p_n(x - y)$ is a complete family.

Then the metric

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x - y)}{1 + p_n(x - y)} = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\rho_n(x, y)}{1 + \rho_n(x, y)}$$

is easily seen to be translation-invariant. It is not too difficult to verify that a sequence $(x_k)_k$ in $\mathcal{V}$ converges to $x \in \mathcal{V}$ relative to the metric topology induced by d if and only if $\lim_{k\to\infty} p_n(x_k - x) = 0$ for all $n \geq 1$. That is, the d-metric topology coincides with the LCS topology induced by Γ. Furthermore, observe that $(x_k)_k$ is Cauchy in the d-metric topology if and only if $(x_k)_k$ is Cauchy relative to each pseudo-metric ρ_n, $n \geq 1$. By the second item above, it follows that $(\mathcal{V}, d)$ is complete, and hence that $(\mathcal{V}, \mathcal{T})$ is a Fréchet space.

5.37. Example.

(a) Let $\mathcal{V} = \mathcal{C}^\infty(\mathbb{R})$ denote the vector space of all functions $f : \mathbb{R} \to \mathbb{R}$ which are infinitely differentiable at each point $x \in \mathbb{R}$. Let $\Gamma = \{p_{n,k}\}_{n,k\geq 0}$, where for $f \in \mathcal{V}$,

$$p_{n,k}(f) := \sup\{|f^{(n)}(x)| : x \in [-k, k]\}.$$

Let $\mathcal{T}$ denote the LCS topology on $\mathcal{V}$ generated by the separating family Γ of seminorms. Then $(\mathcal{V}, \mathcal{T})$ is a Fréchet space.

A sequence $(f_k)_k$ in $\mathcal{V}$ converges to $f \in \mathcal{V}$ if and only if

$$\lim_j \sup\{|f_j^{(n)}(x) - f^{(n)}(x)| : x \in [-k, k]\} = 0$$

for all $n \geq 0$, $k \geq 0$.

(b) If $(\mathfrak{X}, \|\cdot\|)$ is a normed linear space, then with $\Gamma = \{\|\cdot\|\}$, $\mathfrak{X}$ becomes a Fréchet space.

5.38. Many authors define a Fréchet space as a LCS with a translation-invariant metric which is complete as a **uniform topological space**. The definition of a uniform space is rather long, and instead we refer the interested reader to the book of Willard [**Wil70**] for a development of this concept.

5.39. The WOT– and SOT– topologies. In Example 5.28, we defined the weak-operator topology (WOT) and the strong-operator topology (SOT) on $\mathcal{B}(\mathcal{H})$. As we mentioned there, a net $(T_\lambda)_\lambda$ converges to $T \in \mathcal{B}(\mathcal{H})$ if and only if it converges pointwise; that is, for each $x \in \mathcal{H}$, $\lim_\lambda T_\lambda x = Tx$.

Let us now turn our attention to WOT–convergence.

Let $\{e_\alpha : \alpha \in \Omega\}$ be an ONB for an infinite-dimensional Hilbert space $\mathcal{H}$. Corresponding to any $X \in \mathcal{B}(\mathcal{H})$ is a matrix $[X] = [x_{\alpha,\beta}]_{\alpha,\beta\in\Omega}$ defined by $x_{\alpha,\beta} := \langle Xe_\beta, e_\alpha \rangle$ for all $\alpha, \beta \in \Omega$. (You will have definitely seen the finite-dimensional version of this phenomenon, where we write $[T] = [t_{i,j}] \in \mathbb{M}_n(\mathbb{C})$, with $t_{i,j} := \langle Te_j, e_i\rangle$ for some ONB $\{e_1, e_2, \ldots, e_n\}$ for $\mathcal{H} = \mathbb{C}^n$.)

There is an important difference to note when passing from finite-dimensional Hilbert spaces to infinite-dimensional Hilbert spaces, namely: in the infinite-dimensional setting, not every matrix $[x_{\alpha,\beta}]$ is the matrix of a bounded linear operator X. For example, if $\mathcal{H}$ is infinite-dimensional, there is no bounded linear map $X \in \mathcal{B}(\mathcal{H})$ such that $x_{\alpha,\beta} = 1$ for all $\alpha, \beta \in \Omega$. Even in the case where $\mathcal{H}$ is infinite-dimensional but *separable*, there are no known necessary *and* sufficient conditions to identify when a matrix is the matrix of a bounded, linear operator. Of course, certain necessary conditions are easy to state; below are two examples.

- For any fixed $\beta_0 \in \Omega$, the vector $(\langle T\beta_0, \alpha\rangle)_{\alpha\in\Omega} \in \ell^2(\Omega)$, so in particular, it is countably supported. This is because this vector is the set of Fourier coefficients of $T\beta_0$ relative to our ONB $\{e_\alpha : \alpha \in \Omega\}$.
- Given $X \in \mathcal{B}(\mathcal{H})$, for all $\alpha, \beta \in \Omega$,

$$|x_{\alpha,\beta}| = |\langle Xe_\beta, e_\alpha\rangle| \leq \|X\| \, \|e_\beta\| \, \|e_\alpha\| \leq \|X\|.$$

  Thus the entries of $[X]$ are uniformly bounded. As mentioned above, this is far from sufficient.

A sufficient (but not necessary) condition is that

$$\sum_{\alpha,\beta\in\Omega} |x_{\alpha,\beta}|^2 < \infty.$$

One thing to be particularly aware of is that even if $X \in \mathcal{B}(\mathcal{H})$ has a matrix $[X] = [x_{\alpha,\beta}]$ relative to an ONB $\{e_\alpha : \alpha \in \Omega\}$, there is no guarantee that the matrix $[|x_{\alpha,\beta}|]$ is the matrix of a bounded, linear operator. The following result was obtained independently by V.S. Sunder [**Sun78**] and by A.R. Sourour [**Sou78**].

5.40. Theorem. *Let $\mathcal{H}$ be a complex, separable, infinite-dimensional Hilbert space. Let $X \in \mathcal{B}(\mathcal{H})$. The following conditions are equivalent.*

(a) *For **every** ONB $\{e_\alpha\}_{\alpha\in\Omega}$ for $\mathcal{H}$, the matrix $[|\langle Xe_m, e_n\rangle|]_{\alpha,\beta\in\Omega}$ is the matrix of a bounded linear map in $\mathcal{B}(\mathcal{H})$;*

(b) *There exist $\lambda \in \mathbb{C}$ and $K \in \mathcal{B}(\mathcal{H})$ with the property that relative to **some** orthonormal basis $\{e_n\}_{n=1}^\infty$ for $\mathcal{H}$, $\sum_{n,m=1}^\infty |\langle Ke_n, e_m\rangle|^2 < \infty$, and*

$$X = \lambda I + K.$$

We mention in passing that an operator $K \in \mathcal{B}(\mathcal{H})$ for which there exists an ONB $\{e_n\}_{n=1}^\infty$ relative to which

$$\|K\|_2 := \left(\sum_{n,m=1}^\infty |\langle Ke_n, e_m\rangle|^2\right)^{\frac{1}{2}} < \infty$$

is said to be a **Hilbert-Schmidt operator**. It can be shown that the set $\mathcal{C}_2(\mathcal{H}) := \{K \in \mathcal{B}(\mathcal{H}) : K \text{ is Hilbert-Schmidt}\}$ is a linear manifold contained in $\mathcal{K}(\mathcal{H})$, and that $\|\cdot\|_2$ is a norm on $\mathcal{C}_2(\mathcal{H})$. Furthermore, $(\mathcal{C}_2(\mathcal{H}), \|\cdot\|_2)$ is complete, and it forms an ideal in $\mathcal{B}(\mathcal{H})$. We refer the reader to Davidson's monograph [**Dav88**] for more information on Hilbert-Schmidt operators and on other so-called **Schatten p-classes** of operators.

5.41. Returning to the issue of WOT-convergence: suppose that $(T_\lambda)_{\lambda\in\Lambda}$ is a net in $\mathcal{B}(\mathcal{H})$ and that $T \in \mathcal{B}(\mathcal{H})$. Suppose furthermore that $\{e_\alpha\}_{\alpha\in\Omega}$ is a ONB for $\mathcal{H}$.

If WOT-$\lim_\lambda T_\lambda = T$, then it is an immediate consequence of the definition of WOT-convergence that the net $([T_\lambda])_{\lambda\in\Lambda}$ of matrices for $(T_\lambda)_{\lambda\in\Lambda}$ relative to $\{e_\alpha\}_{\alpha\in\Omega}$ must converge entry-wise to the matrix $[T] = [t_{\alpha,\beta}]$ of T relative to $\{e_\alpha\}_{\alpha\in\Omega}$.

Moreover, the converse is true, *provided that the entry-wise convergence of the matrices works for any and all* ONB's. It is **not sufficient** that this work for a single, fixed ONB, as we now show.

5.42. Example. Let $\mathcal{H} = \ell^2$.

For $a, b \in \mathcal{H}$, we define the rank-one operator $a \otimes b^* \in \mathcal{B}(\mathcal{H})$ by

$$a \otimes b^*(x) = \langle x, b\rangle a, \qquad x \in \mathcal{H}.$$

Note that if $a \in \mathcal{H}$ is a norm-one vector, then $P_a := a \otimes a^*$ is the orthogonal projection of $\mathcal{H}$ onto $\mathbb{C}a$.

Now suppose that $\{e_n\}_{n=1}^\infty$ is the standard ONB for $\mathcal{H} = \ell^2$; that is, $e_N = (\delta_{n,N})_{n=1}^\infty$, where $\delta_{i,j}$ is the Dirac delta function $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$.

Consider $R_N = N^2(e_N \otimes e_N^*)$, $N \geq 1$, so that R_N is N^2 times the orthogonal projection of $\mathcal{H}$ onto $\mathbb{C}e_N$ for each N.

For any $1 \leq i, j < \infty$,

$$\lim_{N\to\infty} \langle R_N e_i, e_j\rangle = 0,$$

because for any $N \geq i$, $R_N e_i = 0$. Nevertheless, we do *not* have that $(R_N)_{N=1}^\infty$ converges in the WOT to 0.

Let $x = \sum_{n=1}^\infty \frac{1}{n} e_n \in \mathcal{H}$, since $\sum_{n=1}^\infty (\frac{1}{n})^2 < \infty$. Then for all $N \geq 1$,

$$\langle R_N x, x\rangle = \langle R_N(\sum_n \frac{1}{n} e_n), x\rangle = \langle N e_N, x\rangle = \langle N e_N, (\sum_n \frac{1}{n} e_n)\rangle = \langle N e_N, \frac{1}{N} e_N\rangle = 1.$$

Incidentally, the matrix for R_N relative to the ONB $\{e_n\}_{n=1}^\infty$ is

$$[R_N] = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & & \dots \\ 0 & 0 & 0 & \dots & 0 & & \dots \\ 0 & \dots & \ddots & 0 & 0 & & \dots \\ 0 & \dots & 0 & N^2 & 0 & & \dots \\ 0 & \dots & \dots & 0 & 0 & & \dots \\ 0 & \dots & \dots & 0 & 0 & \ddots & \dots \\ \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \end{bmatrix}.$$

That is, there is N^2 in the (N, N) entry, and 0's elsewhere. By our work in Chapter 2, $\|R_N\| = N^2$ for each $N \geq 1$.

We leave it as an exercise for the reader to show that $(R_N)_{N=1}^\infty$ does not converge to *anything* in the weak-operator topology.

5.43. Example. Let $\mathcal{H}$ denote an infinite-dimensional, separable, complex Hilbert space. By $\mathcal{B}(\mathcal{H})$ we denote the set of bounded linear operators acting on $\mathcal{H}$. The following is an example of an unbounded net of operators in $\mathcal{B}(\mathcal{H})$ which nonetheless converges to 0 in the strong operator topology, and therefore in the weak operator topology.

Applying Example 11.19 of the online notes to the nbhd system $\mathcal{U}_0^{\text{SOT}}$ of 0 in the SOT, we see that the latter forms a directed set using the relation

$$U_1 \le U_2 \text{ if } U_2 \subseteq U_1,$$

and that if we choose an element $X_U \in U$, $U \in \mathcal{U}_0^{\text{SOT}}$, then

$$\lim_{U \in \mathcal{U}_0^{\text{SOT}}} X_U = 0.$$

Recall also from your Assignments that any SOT-open nbhd of $0 \in \mathcal{B}(\mathcal{H})$ contains an infinite-dimensional SOT-closed subspace of $\mathcal{B}(\mathcal{H})$. Of course, a non-zero subspace of $\mathcal{B}(\mathcal{H})$ contains operators of any possible norm.

Each $U \in \mathcal{U}_0^{\text{SOT}}$ then contains a basic SOT-open nbhd of 0 of the form

$$N(0, F, \varepsilon) = \{T \in \mathcal{B}(\mathcal{H}) : \|Tx - 0\| < \varepsilon, x \in F\},$$

where $F = \{x_1, x_2, \ldots, x_N\} \subseteq \mathcal{H}$ is a finite set, and $\varepsilon > 0$.

Choose an operator $R_U \in N(0, F, \varepsilon) \subseteq U$ with $\|R_U\| > \frac{1}{\varepsilon}$. (Alternatively, we may ask that $\|R_U\| \ge |F|$, the cardinality of F.)

Since $R_U \in U$, $U \in \mathcal{U}_0^{\text{SOT}}$, we see from above that

$$\lim_{U \in \mathcal{U}_0^{\text{SOT}}} R_U = 0.$$

Since $\varepsilon > 0$ can be made arbitrarily small (starting with any $\delta > 0$, let $0 \ne x \in \mathcal{H}$ and consider $U = N(0, \{x\}, \delta)$), the net $(R_U)_U$ is not bounded.

We also remark that there does not exist an unbounded **sequence** $(T_n)_{n=1}^\infty$ in $\mathcal{B}(\mathcal{H})$ which converges in the WOT to 0.

Indeed, suppose otherwise; that is, suppose that $(T_n)_n$ is a sequence and that for each $x, y \in \mathcal{H}$,

$$\lim_{n \to \infty} \langle T_n x, y \rangle = 0.$$

Then, for each $x \in \mathcal{H}$, $(T_n x)_{n=1}^\infty$ is a sequence in $\mathcal{H}$ which converges weakly to 0. (This uses the Riesz Representation Theorem for Hilbert spaces to identify a functional $\beta \in \mathcal{H}^*$ with a functional $\beta_y(z) := \langle z, y \rangle \forall z \in \mathcal{H}$.) By Corollary 7.17 of the online notes, $\sup_{n \ge 1} \|T_n x\| < \infty$.

Of course, if we had a sequence $(T_n)_{n=1}^\infty$ converging in the SOT to 0, then for each $x \in \mathcal{H}$, $\lim_n \|T_n x\| = 0$, implying again that $\sup_{n \ge 1} \|T_n x\| < \infty$.

Either way (i.e. WOT or SOT convergence to 0), we can apply the Uniform Boundedness Principle: for each $x \in \mathcal{H}$,

$$\kappa_x := \sup_{n \ge 1} \|T_n x\| < \infty,$$

implying that $\sup_n \|T_n\| < \infty$.

*

I once spent a year in Philadelphia. I think it was on a Sunday.

W.C. Fields

Exercises for Section 5.

Question 1. This question is based upon Example 5.42 of the Appendix. Let $\mathcal{H} = \ell^2$ and denote by $\{e_n\}_{n=1}^\infty$ the standard ONB for $\mathcal{H}$. For each $N \geq 1$, set $R_N = N^2(e_N \otimes e_N^*)$, where for $x, y \in \mathcal{H}$, $x \otimes y^*$ denotes the rank-one operator

$$x \otimes y^*(z) = \langle z, y \rangle x, \quad z \in \mathcal{H}.$$

Prove that the sequence $(R_N)_{N=1}^\infty$ does not converge in the WOT. Conclude that it does not converge in the SOT, nor in the norm topology on $\mathcal{B}(\mathcal{H})$.

Question 2.

Let $\mathcal{H}$ be an infinite-dimensional, separable, complex Hilbert space and let $\{e_n\}_{n=1}^\infty$ be an ONB for $\mathcal{H}$. For each $N \geq 1$, let P_N denote the orthogonal projection of $\mathcal{H}$ onto span $\{e_1, e_2, \ldots, e_N\}$.

Prove that the sequence $(P_n)_{n=1}^\infty$ converges strongly to the identity; that is, that it converges to the identity operator in the SOT.

Question 3.

Let $\mathcal{H}$ be an infinite-dimensional, separable, complex Hilbert space, and denote by $\mathcal{F}(\mathcal{H})$ the set of finite-rank operators in $\mathcal{B}(\mathcal{H})$. Prove that $\mathcal{F}(\mathcal{H})$ is dense in $\mathcal{B}(\mathcal{H})$ in the WOT.

Is it dense in $\mathcal{B}(\mathcal{H})$ in the SOT?

6. The Hahn-Banach theorem

When I wake up in the morning, I just can't get started until I've had that first, piping hot pot of coffee. Oh, I've tried other enemas...

Emo Philips

6.1. It is somewhat of a misnomer to refer to *the* Hahn-Banach Theorem. In fact, there is a large number of variations on this theme. These variations fall into two groups: the separation theorems, and the extension theorems. The crucial relation between these two classes of theorems is that they all refer to linear functionals. Having said this, when one wishes to apply a version of the Hahn-Banach Theorem, one tends to say only: "by *the* Hahn-Banach Theorem...", usually leaving it to the reader to determine which version of the Theorem is being applied.

The importance of these theorems in Functional Analysis can not be overstated.

6.2. Definition. *Let $\mathcal{W}$ be a vector space over $\mathbb{K}$. A* ***linear functional*** *on $\mathcal{W}$ is a linear map $f : \mathcal{W} \to \mathbb{K}$. The vector space of all linear functionals on $\mathcal{W}$ is denoted by $\mathcal{W}^\#$ and is referred to as the* ***algebraic dual*** *of $\mathcal{W}$.*

If $\mathcal{W}$ is a TVS, the (vector) space of ***continuous*** *linear functionals is denoted by $\mathcal{W}^*$, and is referred to as the* ***(topological) dual*** *of $\mathcal{W}$. Obviously $\mathcal{W}^* \subseteq \mathcal{W}^\#$.*

6.3. Example. Let $n \geq 1$ be an integer and consider $\mathcal{W} = \mathbb{K}^n$ equipped with the norm $\|(x_1, x_2, ..., x_n)\|_\infty = \max |x_j|$. For any choice of $k_1, k_2, ..., k_n \in \mathbb{K}$, the map

$$\begin{array}{rccc} f: & \mathcal{W} & \to & \mathbb{K} \\ & (x_1, x_2, ..., x_n) & \mapsto & \sum_{i=1}^n k_i x_i \end{array}$$

is a continuous linear functional.

6.4. Remarks.

(a) Recall from basic linear algebra that every linear functional on $\mathbb{K}^n$ is of this form for some choice of $k_1, k_2, ..., k_n \in \mathbb{K}$. As such, every linear functional on $\mathbb{K}^n$ is continuous.

(b) Recall from Proposition 4.20 that if $\mathcal{V}$ is an n-dimensional TVS with basis $\{e_1, e_2, ..., e_n\}$, then $\mathcal{V}$ is homeomorphic to $\mathbb{K}^n$ via the map $\sum_{i=1}^n k_i e_i \mapsto (k_1, k_2, ..., k_n)$. Since the product topology on $\mathbb{K}^n$ is in turn equivalent to the norm topology induced by the infinity norm, it follows from (a) above that every linear functional on a finite-dimensional TVS is continuous.

6.5. Example. Let us next consider $c_{00}(\mathbb{K}) = \{(x_n)_{n=1}^\infty : x_n \in \mathbb{K}$ for all $n \geq 1$ and $x_n = 0$ for all but finitely many n's$\}$. Recall that this forms a normed linear space when equipped with the norm

$$\|(x_n)_n\|_\infty = \sup_n |x_n|.$$

Define

$$\begin{array}{rccc} f: & c_{00}(\mathbb{K}) & \to & \mathbb{K} \\ & (x_n)_n & \mapsto & \sum_{n=1}^\infty x_n \end{array}.$$

Then f is a non-continuous linear functional on $c_{00}(\mathbb{K})$. Indeed, if

$$\mathbf{y}_n = (\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}, 0, 0, 0, ...)$$

(where the $\frac{1}{n}$ term is repeated n times), then $\|\mathbf{y}_n\|_\infty = \frac{1}{n}$, and so $\lim_{n\to\infty} \mathbf{y}_n = 0$, while $f(\mathbf{y}_n) = 1$ for all n, and hence $\lim_{n\to\infty} f(\mathbf{y}_n) \neq 0 = f(\mathbf{0})$.

For a number of the results we shall obtain below, we shall assume that the underlying field is $\mathbb{R}$. In order to translate the results to the case of complex vector spaces, the following Lemma will be useful.

6.6. Lemma. *Let $\mathcal{V}$ be a vector space over $\mathbb{C}$.*

(a) *If $f : \mathcal{V} \to \mathbb{R}$ is an $\mathbb{R}$-linear functional, then the map*

$$f_{\mathbb{C}}(x) := f(x) - if(ix)$$

is a $\mathbb{C}$-linear functional on $\mathcal{V}$, and $f = \operatorname{Re} f_{\mathbb{C}}$.

(b) *If $g : \mathcal{V} \to \mathbb{C}$ is $\mathbb{C}$-linear, $f = \operatorname{Re} g$ and $f_{\mathbb{C}}$ is defined as in (a), then $g = f_{\mathbb{C}}$.*

(c) *If p is a $\mathbb{C}$-seminorm on $\mathcal{V}$ and f, $f_{\mathbb{C}}$ are as in (a) above, then $|f(x)| \leq p(x)$ for all $x \in \mathcal{V}$ if and only if $|f_{\mathbb{C}}(x)| \leq p(x)$ for all $x \in \mathcal{V}$.*

(d) *If $\mathcal{V}$ is a NLS and $f, f_{\mathbb{C}}$ are as in (a), then $\|f\| = \|f_{\mathbb{C}}\|$.*

Proof.

(a) This is routine and is left to the reader.

(b) Let $x \in \mathcal{V}$ and write $g(x) = a + ib$, where $a = \operatorname{Re} g(x) = f(x)$ and $b = \operatorname{Im} g(x)$ are real. By $\mathbb{C}$-linearity of g, $g(ix) = ig(x) = -b + ia$, and so

$$\operatorname{Im} g(x) = b = -\operatorname{Re} g(ix) = -f(ix).$$

That is, $g(x) = f(x) + i(-f(ix)) = f(x) - if(ix) = f_{\mathbb{C}}(x)$.

(c) First suppose that $|f_{\mathbb{C}}(x)| \leq p(x)$ for all $x \in \mathcal{V}$. Then $|f(x)| = |\operatorname{Re} f_{\mathbb{C}}(x)| \leq |f_{\mathbb{C}}(x)| \leq p(x)$ for all $x \in \mathcal{V}$.

Next suppose that $|f(x)| \le p(x)$ for all $x \in \mathcal{V}$. Given $x \in \mathcal{V}$, choose $\theta \in \mathbb{C}$, $|\theta| = 1$ so that $|f_{\mathbb{C}}(x)| = \theta f_{\mathbb{C}}(x)$. Then

$$\begin{aligned} |f_{\mathbb{C}}(x)| &= \theta f_{\mathbb{C}}(x) \\ &= f_{\mathbb{C}}(\theta x) \\ &= \operatorname{Re} f_{\mathbb{C}}(\theta x) \qquad \text{(as this quantity is non-negative)} \\ &= f(\theta x) \\ &\le p(\theta x) \\ &= |\theta| p(x) = p(x). \end{aligned}$$

(d) We only deal with the case where $\|f\|$ and $\|f_{\mathbb{C}}\|$ are finite, and leave the case where one of them is (and therefore both of them are) infinite to the reader.

It is routine to verify that $\|f\| \le \|f_{\mathbb{C}}\|$, and this step is left to the reader. Conversely, given $x \in \mathcal{V}$ with $\|x\| = 1$, we can find θ_x so that $|f_{\mathbb{C}}(x)| = \theta_x f_{\mathbb{C}}(x) = f_{\mathbb{C}}(\theta_x x) = \operatorname{Re} f_{\mathbb{C}}(\theta_x x) = f(\theta_x x)$. Note that $\|\theta_x x\| = 1$ because $\mathcal{V}$ is a $\mathbb{C}$-vector space and $\|\theta_x x\| = |\theta_x|\,\|x\| = \|x\| = 1$. Thus

$$\begin{aligned} \|f_{\mathbb{C}}\| &= \sup\{|f_{\mathbb{C}}(z)| : \|z\| = 1\} \\ &= \sup\{|f(\theta_z z)| : \|z\| = 1\} \\ &\le \sup\{|f(y)| : \|y\| = 1\} \\ &= \|f\|. \end{aligned}$$

□

6.7. Proposition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$ and let $f \in \mathcal{V}^{\#}$.*

(a) *If $g \in \mathcal{V}^{\#}$ and $g|_{\ker f} = 0$, then $g = kf$ for some $k \in \mathbb{K}$.*

(b) *If $g, f_1, f_2, ..., f_N \in \mathcal{V}^{\#}$ and $g(x) = 0$ for all $x \in \cap_{j=1}^N \ker f_j$, then $g \in \operatorname{span}\{f_1, f_2, ..., f_N\}$.*

Proof.

(a) If $g = 0$, then set $k = 0$ and we are done.

Otherwise, choose $z \in \mathcal{V}$ so that $g(z) \neq 0$. By hypothesis, $f(z) \neq 0$. Let $k = g(z)/f(z)$. Now $\ker f$ has codimension 1 in $\mathcal{V}$, and so if $x \in \mathcal{V}$, then $x = \alpha z + y$ for some $y \in \ker f$ and $\alpha \in \mathbb{K}$. Hence

$$\begin{aligned} g(x) = \alpha g(z) + g(y) &= \alpha g(z) + 0 \\ &= \alpha k f(z) \\ &= k(\alpha f(z) + f(y)) \\ &= k f(x). \end{aligned}$$

Since $x \in \mathcal{V}$ was arbitrary, $g = kf$.

(b) We may assume that $\{f_1, f_2, ..., f_N\}$ are linearly independent. Let $\mathcal{N} = \cap_{j=1}^N \ker f_j$. Then $\dim(\mathcal{V}/\mathcal{N}) \le N$. (This is an elementary result from Linear Algebra; there is a proof in the Appendix to this Chapter.) For

$1 \le j \le N$, define $\overline{f}_j : \mathcal{V}/\mathcal{N} \to \mathbb{K}$ via $\overline{f}_j(x+\mathcal{N}) = f_j(x)$. Since $\mathcal{N} \subseteq \ker f_j$, each $\overline{f}_j$ is well-defined, and $\overline{f}_j \in (\mathcal{V}/\mathcal{N})^{\#}$.

We claim that $\{\overline{f}_1, \overline{f}_2, ..., \overline{f}_N\}$ is also linearly independent. Otherwise, we can find $k_1, k_2, ..., k_N \in \mathbb{K}$ so that $\sum_{j=1}^N |k_j| \ne 0$, but $\sum_{j=1}^N k_j \overline{f}_j = 0$. But then $\sum_{j=1}^N k_j f_j \ne 0$, so we can find $z \in \mathcal{V}$ with $0 \ne \sum_{j=1}^N k_j f_j(z) = \sum_{j=1}^N k_j \overline{f}_j(z+\mathcal{N})$, a contradiction.

Thus $\dim(\mathcal{V}/\mathcal{N})^{\#} \ge N$. Combining this with the fact that $\dim(\mathcal{V}/\mathcal{N}) \le N$ yields $\dim(\mathcal{V}/\mathcal{N}) = \dim(\mathcal{V}/\mathcal{N})^{\#} = N$, and that $\{\overline{f}_1, \overline{f}_2, ..., \overline{f}_N\}$ is a basis for $(\mathcal{V}/\mathcal{N})^{\#}$.

Now define $\overline{g} : \mathcal{V}/\mathcal{N} \to \mathbb{K}$ via $\overline{g}(x+\mathcal{N}) = g(x)$. Again, since $\mathcal{N} \subseteq \ker g$, $\overline{g}$ is well-defined. Since $\overline{g} \in (\mathcal{V}/\mathcal{N})^{\#}$, we can write

$$\overline{g} = \sum_{j=1}^N k_j \overline{f}_j \quad \text{for some } k_1, k_2, ..., k_N \in \mathbb{K}.$$

For $x \in \mathcal{V}$,

$$\begin{aligned} 0 &= (\overline{g} - \sum_{j=1}^N k_j \overline{f}_j)(x+\mathcal{N}) \\ &= g(x) - \sum_{j=1}^N k_j f_j(x), \end{aligned}$$

so that $g = \sum_{j=1}^N k_j f_j$.

□

The first part of the above Proposition shows that if f and g are distinct linear functionals on a vector space $\mathcal{V}$, then they have the same kernel if and only if one functional is a non-zero multiple of the other.

6.8. Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. A **hyperplane** $\mathcal{M}$ in $\mathcal{V}$ is a linear manifold for which* $\dim(\mathcal{V}/\mathcal{M}) = 1$.

6.9. If $0 \ne \varphi \in \mathcal{V}^{\#}$, then from elementary linear algebra theory we see that $\mathcal{M} := \ker \varphi$ is a hyperplane in $\mathcal{V}$ and that φ induces an (algebraic) isomorphism $\overline{\varphi}$ between $\mathcal{V}/\mathcal{M}$ and $\mathbb{K}$ via

$$\overline{\varphi}(x+\mathcal{M}) := \varphi(x) \qquad \text{for all } x+\mathcal{M} \in \mathcal{V}/\mathcal{M}.$$

Conversely, if $\mathcal{M} \subseteq \mathcal{V}$ is a hyperplane, then $\mathcal{V}/\mathcal{M}$ is (algebraically) isomorphic to $\mathbb{K}$. Let $\kappa : \mathcal{V}/\mathcal{M} \to \mathbb{K}$ denote such an isomorphism. If $q : \mathcal{V} \to \mathcal{V}/\mathcal{M}$ is the canonical quotient map, then $\kappa \circ q : \mathcal{V} \to \mathbb{K}$ is a linear functional with $\ker(\kappa \circ q) = \mathcal{M}$.

Thus we have established a correspondence between linear functionals and hyperplanes. Proposition 6.7 implies that, up to a factor of a non-zero scalar multiple, this correspondence is bijective.

6.10. Proposition. *If $(\mathcal{V}, \mathcal{T})$ is a TVS and $\mathcal{M} \subseteq \mathcal{V}$ is a hyperplane, then either $\mathcal{M}$ is closed in $\mathcal{V}$, or $\mathcal{M}$ is dense.*

Proof. Since $\overline{\mathcal{M}}$ is a vector space satisfying $\mathcal{M} \subseteq \overline{\mathcal{M}} \subseteq \mathcal{V}$, and since $\dim(\mathcal{V}/\mathcal{M}) = 1$, we either have $\mathcal{M} = \overline{\mathcal{M}}$, or $\overline{\mathcal{M}} = \mathcal{V}$.

□

It is worth noting that both possibilities can occur.

6.11. Example.

(a) Let $\mathfrak{X} = (\mathcal{C}([0,1], \mathbb{C}), \|\cdot\|_\infty)$, and let $\delta_{\frac{1}{2}} : \mathfrak{X} \to \mathbb{C}$ be the map $\delta_{\frac{1}{2}}(f) := f(\frac{1}{2})$, $f \in \mathfrak{X}$. Then $\ker \delta_{\frac{1}{2}} = \{f : [0,1] \to \mathbb{C} | f \text{ is continuous and } f(\frac{1}{2}) = 0\}$. This is clearly closed.

(b) Let $\mathcal{V} = c_0(\mathbb{C})$, and let $e_k = (\delta_{1k}, \delta_{2k}, \delta_{3k}, ...)$, where $\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$. Let $z = (1, 1/2, 1/3, 1/4, ...)$. Then $\{z, e_1, e_2, e_3, ...\}$ is linearly independent in $\mathcal{V}$, and as such it can be extended to a Hamel basis (i.e. a vector space basis) for $\mathcal{V}$, say

$$\mathcal{B} = \{z, e_1, e_2, e_3, ...\} \cup \{b_\lambda : \lambda \in \Lambda\}.$$

Given $v \in \mathcal{V}$, say $v = \alpha z + \sum_{k=1}^\infty \beta_k e_k + \sum_{\lambda \in \Lambda} \gamma_\lambda b_\lambda$ for some $\alpha, \beta_k, \gamma_\lambda \in \mathbb{K}$ with only finitely many coefficients not equal to zero, define $g(v) = \alpha$.

It is clear that this defines a linear functional on $\mathcal{V}$. Since $\ker g$ is a subspace containing $e_k, k \geq 1$, and since $\text{span}\,\{e_k\}_{k=1}^\infty$ is dense in $\mathcal{V}$, $\ker g$ is dense in $\mathcal{V}$. Since $g \neq 0$, and since $\ker g$ is dense in $\mathcal{V}$, we see that $\ker g$ is not closed.

6.12. Proposition. *Let $\mathcal{V}$ be a TVS and $\rho \in \mathcal{V}^\#$. Suppose that there exists an open nbhd $U \in \mathcal{U}_0$ of 0 and a constant $\kappa > 0$ so that $\text{Re}\,\rho(x) < \kappa$ for all $x \in U$. Then ρ is uniformly continuous on $\mathcal{V}$.*

Proof. By Proposition 4.10, we can find a balanced, open nbhd N of 0 with $N \subseteq U$. Observe that for $x \in N \subseteq U$, there exists $\theta_x \in \mathbb{K}$, $|\theta_x| = 1$ so that

$$|\rho(x)| = \rho(\theta_x x) = \text{Re}\,\rho(\theta_x x).$$

But $\theta_x x \in N$ since N is balanced, and so $|\rho(x)| < \kappa$ for $x \in N$. Consider the function

$$\begin{array}{rccc} p: & \mathcal{V} & \to & \mathbb{R} \\ & x & \mapsto & |\rho(x)| \end{array}$$

which is easily seen to be a seminorm. Since p is bounded above by κ on the open nbhd N of 0, we can invoke Proposition 5.11 to conclude that p is continuous on $\mathcal{V}$. It follows that ρ is continuous at 0, and hence ρ is uniformly continuous on $\mathcal{V}$ by Theorem 4.33.

□

6.13. Corollary. *Let $\mathcal{V}$ be a TVS and $\rho \in \mathcal{V}^{\#}$. The following are equivalent:*

(a) *ρ is continuous on $\mathcal{V}$ - i.e. $\rho \in \mathcal{V}^*$;*

(b) $\ker \rho$ *is closed.*

Proof.

(a) implies (b): This is clear. If ρ is continuous, then $\ker \rho = \rho^{-1}(\{0\})$ is closed in $\mathcal{V}$ because $\{0\}$ is closed in $\mathbb{K}$.

(b) implies (a): Suppose next that $\ker \rho$ is closed. If $\rho = 0$, then ρ is obviously continuous. Suppose therefore that $\rho \neq 0$. Then $\mathcal{W} := \mathcal{V}/\ker \rho$ is a one-dimensional TVS and

$$\begin{array}{cccc} \overline{\rho}: & \mathcal{W} & \to & \mathbb{K} \\ & x + \ker \rho & \mapsto & \rho(x) \end{array}$$

is a linear functional on $\mathcal{W}$. By Remark 6.4 (b), $\overline{\rho}$ is continuous. If $q : \mathcal{V} \to \mathcal{W}$ is the canonical quotient map, then by Paragraph 4.18, q is also continuous, and thus $\rho = \overline{\rho} \circ q$ is continuous as well.

□

Recall that $\mathfrak{X}^* = \mathcal{B}(\mathfrak{X}, \mathbb{K})$ is a Banach space whenever $\mathfrak{X}$ is a NLS. Let us recall a couple of results from Measure Theory which provide us with interesting examples of classes of linear functionals.

6.14. Theorem. *Let (X, Ω, μ) be a measure space and $1 < p < \infty$. If $\frac{1}{p} + \frac{1}{q} = 1$, and if $g \in L^q(X, \Omega, \mu)$, then*

$$\beta_g(f) := \int_X fg d\mu$$

defines a continuous linear functional on $L^p(X, \Omega, \mu)$, and the map $g \mapsto \beta_g$ is an isometric linear bijection of $L^q(X, \Omega, \mu)$ onto $L^p(X, \Omega, \mu)^*$.

If (X, Ω, μ) is σ-finite, then the same conclusion holds in the case where $p = 1$ and $q = \infty$.

Recall that if X is a locally compact space, then $M_{\mathbb{K}}(X)$ denotes the space of $\mathbb{K}$-valued regular Borel measures on X with the **total variation norm**.

6.15. Theorem. *If X is locally compact and $\mu \in M_{\mathbb{K}}(X)$, then*

$$\begin{array}{cccc} \beta_\mu : & \mathcal{C}_0(X, \mathbb{K}) & \to & \mathbb{K} \\ & f & \mapsto & \int_X f d\mu \end{array}$$

defines an element of $\mathcal{C}_0(X, \mathbb{K})$, and the map $\mu \mapsto \beta_\mu$ is an isometric linear isomorphism of $M_{\mathbb{K}}(X)$ onto $\mathcal{C}_0(X, \mathbb{K})^*$.

The Extension Theorems

6.16. The Hahn-Banach Theorem is probably the most important result in Functional Analysis. It has a great many applications, and its usefulness cannot be overstated. There are two basic formulations of this result (each with a variety of consequences); the first in terms of extensions of linear functionals from linear submanifolds of a LCS to the entire LCS, and the second in terms of so-called "separation theorems", which we shall examine later.

6.17. Proposition. *Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ and $p : \mathcal{V} \to \mathbb{R}$ be a sublinear functional. Suppose that $\mathcal{M}$ is a hyperplane and that $f : \mathcal{M} \to \mathbb{R}$ is a linear functional for which $f(x) \le p(x)$ for all $x \in \mathcal{M}$. Then there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ such that $g|_{\mathcal{M}} = f$, and $g(x) \le p(x)$ for all $x \in \mathcal{V}$.*

Proof. Let $z \in \mathcal{V}\backslash\mathcal{M}$, so that $\mathcal{V} = \text{span}\,\{z, \mathcal{M}\}$. Then $v \in \mathcal{V}$ implies that $v = tz+m$ for some $t \in \mathbb{R}$, $m \in \mathcal{M}$.

For each $r \in \mathbb{R}$ we may define $h_r : \mathcal{V} \to \mathbb{R}$ by setting $h_r(z) = r$, setting $h_r(m) = f(m)$, $m \in \mathcal{M}$, and then extending h_r by linearity to all of $\mathcal{V}$. Clearly $h_r \in \mathcal{V}^{\#}$ and h_r extends f. The problem is that we do not know that $h_r(x) \le p(x)$ for all $x \in \mathcal{V}$ – in fact, this is generally not true. The question of finding a g as in the statement of the Proposition amounts to showing that for *some* $s \in \mathbb{R}$, we will have $h_s(x) \le p(x)$ for all $x \in \mathcal{V}$. To find such an s, we first examine which properties it *must* satisfy. We then demonstrate that these properties are also sufficient. Finally, the existence of s is a byproduct of reconciling these necessary and sufficient conditions.

If $h_s(x) \le p(x)$ for all $x \in \mathcal{V}$, then for all $t \in \mathbb{R}$, $m \in \mathcal{M}$ we must have

$$h_s(tz+m) = ts + f(m) \le p(tz+m).$$

- If $t > 0$, then setting $m_1 = t^{-1}m$ yields:

$$\begin{aligned} s &\le -t^{-1}f(m) + t^{-1}p(tz+m) \\ &= -f(t^{-1}m) + p(z+t^{-1}m) \quad \text{for all } m \in \mathcal{M} \\ &= -f(m_1) + p(z+m_1) \quad \text{for all } m_1 \in \mathcal{M}. \end{aligned} \tag{1}$$

- If $t < 0$, then setting $m_2 = -t^{-1}m$ yields:

$$\begin{aligned} s &\ge -t^{-1}f(m) + t^{-1}p(tz+m) \\ &= f(-t^{-1}m) - p(-z-t^{-1}m) \text{ for all } m \in \mathcal{M} \\ &= f(m_2) - p(-z+m_2) \quad \text{for all } m_2 \in \mathcal{M}. \end{aligned} \tag{2}$$

The key issue is that we can "reverse engineer" this process. Suppose that $s \in \mathbb{R}$ satisfies both (1) and (2), namely

$$f(m_2) - p(-z+m_2) \le s \le -f(m_1) + p(z+m_1) \quad \text{for all } m_1, m_2 \in \mathcal{M}. \tag{3}$$

- If $t > 0$, then
$$\begin{aligned} h_s(tz+m) &= ts + f(m) \\ &\le t(-f(m/t) + p(z + (m/t))) + f(m) \\ &= p(tz+m) \quad \text{for all } m \in \mathcal{M}, \end{aligned}$$
while
- if $t < 0$, then
$$\begin{aligned} h_s(tz+m) &= ts + f(m) \\ &\le t(f(-m/t) - p(-z - (m/t))) + f(m) \\ &= -f(m) + (-t)p(-z - (m/t)) + f(m) \\ &= p(tz+m) \quad \text{for all } m \in \mathcal{M}. \end{aligned}$$
- If $t = 0$, then $h_s(tz+m) = h_s(m) = f(m) \le p(m) = p(tz+m)$ for all $m \in \mathcal{M}$.

There remains to show, therefore, that we can find $s \in \mathbb{R}$ which satisfies (3) (or equivalently, which satisfies both (1) and (2)). Now this can be done if
$$f(m_2) + f(m_1) \le p(-z + m_2) + p(z + m_1) \quad \text{for all } m_1, m_2 \in \mathcal{M}.$$
But
$$\begin{aligned} f(m_1) + f(m_2) = f(m_1 + m_2) &\le p(m_1 + m_2) \\ &\le p(m_2 - z) + p(z + m_1) \end{aligned}$$
for all $m_1, m_2 \in \mathcal{M}$, and so we can choose
$$s_0 := \sup\{f(m_2) - p(-z + m_2) : m_2 \in \mathcal{M}\}.$$
Letting $g = h_{s_0}$ completes the proof.

□

6.18. Theorem. ***The Hahn-Banach Theorem 01***
Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ and let p be a sublinear functional on $\mathcal{V}$. If $\mathcal{M}$ is a linear manifold in $\mathcal{V}$ and $f : \mathcal{M} \to \mathbb{R}$ is a linear functional with $f(m) \le p(m)$ for all $m \in \mathcal{M}$, then there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ with $g|_{\mathcal{M}} = f$, and $g(x) \le p(x)$ for all $x \in \mathcal{V}$.

Proof. Let $\mathcal{J} = \{(\mathcal{N}, h) : \mathcal{N}$ a linear manifold in $\mathcal{V}$, $\mathcal{M} \subseteq \mathcal{N}, h \in \mathcal{N}^{\#}$, $h|_{\mathcal{M}} = f$, and $h(n) \le p(n)$ for all $n \in \mathcal{N}\}$. For $(\mathcal{N}_1, h_1), (\mathcal{N}_2, h_2) \in \mathcal{J}$, define $(\mathcal{N}_1, h_1) \preceq (\mathcal{N}_2, h_2)$ if $\mathcal{N}_1 \subseteq \mathcal{N}_2$ and $h_2|_{\mathcal{N}_1} = h_1$. Then $(\mathcal{J}, \preceq)$ is a partially ordered set with respect to $\preceq$. Moreover, $\mathcal{J} \ne \varnothing$, since $(\mathcal{M}, f) \in \mathcal{J}$.

Let $\mathcal{C} = \{(\mathcal{N}_\lambda, h_\lambda) : \lambda \in \Lambda\}$ be a chain in $\mathcal{J}$, and let $\mathcal{N} := \cup_{\lambda \in \Lambda} \mathcal{N}_\lambda$. Define $h : \mathcal{N} \to \mathbb{R}$ by setting $h(n) = h_\lambda(n)$ if $n \in \mathcal{N}_\lambda$. Then h is well-defined because $\mathcal{C}$ is a chain (check!), h is linear and $h(n) \le p(n)$ for all $n \in \mathcal{N}$. Thus $(\mathcal{N}, h) \in \mathcal{J}$ and it is an upper bound for $\mathcal{C}$. By Zorn's Lemma, $(\mathcal{J}, \preceq)$ has a maximal element $(\mathcal{Y}, g)$. Suppose that $\mathcal{Y} \ne \mathcal{V}$. Choosing $z \in \mathcal{V} \backslash \mathcal{Y}$ and letting $\mathcal{Y}_0 = \operatorname{span}\{z, \mathcal{Y}\}$, Proposition 6.17 implies the existence of a functional $g_0 : \mathcal{Y}_0 \to \mathbb{R}$ which extends g

and satisfies $g_0(y) \le p(y)$ for all $y \in \mathcal{Y}_0$. This contradicts the maximality of $(\mathcal{Y}, g)$. Hence $\mathcal{Y} = \mathcal{V}$, and g has the required properties.

□

The complex version of this theorem can now be established.

6.19. Theorem. ***The Hahn-Banach Theorem 02***
Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. Let $\mathcal{M} \subseteq \mathcal{V}$ be a linear manifold and let $p : \mathcal{V} \to \mathbb{R}$ be a seminorm on $\mathcal{V}$. If $f : \mathcal{M} \to \mathbb{K}$ is a linear functional and $|f(m)| \le p(m)$ for all $m \in \mathcal{M}$, then there exists a linear functional $g : \mathcal{V} \to \mathbb{K}$ so that $g|_{\mathcal{M}} = f$ and $|g(x)| \le p(x)$ for all $x \in \mathcal{V}$.

Proof. Suppose that $\mathbb{K} = \mathbb{R}$. Then $f(m) \le |f(m)| \le p(m)$ for all $m \in \mathcal{M}$, and p is a sublinear functional (by virtue of the fact that it is a seminorm). By the Hahn-Banach Theorem 01, there exists $g : \mathcal{V} \to \mathbb{R}$ linear so that $g|_{\mathcal{M}} = f$ and $g(x) \le p(x)$ for all $x \in \mathcal{V}$. Thus $-g(x) = g(-x) \le p(-x) = p(x)$ for all $x \in \mathcal{V}$, so that $|g(x)| \le p(x)$ for all $x \in \mathcal{V}$.

Now suppose that $\mathbb{K} = \mathbb{C}$. Let $f_1 = \mathrm{Re} f$. Then by Lemma 6.6, $|f_1(m)| \le p(m)$ for all $m \in \mathcal{M}$. By the argument of the first paragraph of this proof, there exists an $\mathbb{R}$-linear functional $g_1 : \mathcal{V} \to \mathbb{R}$ so that $g_1|_{\mathcal{M}} = f_1$ and $|g_1(m)| \le p(m)$ for all $m \in \mathcal{M}$. Let $g = (g_1)_{\mathbb{C}}$ denote the complexification of g_1, as obtained in Lemma 6.6. Then $g : \mathcal{V} \to \mathbb{C}$ is $\mathbb{C}$-linear, $g|_{\mathcal{M}} = f$, and by part (c) of that Lemma,

$$|g(x)| \le p(x) \quad \text{for all } x \in \mathcal{V}.$$

□

6.20. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\mathcal{W} \subseteq \mathcal{V}$ be a linear manifold. If $f \in \mathcal{W}^*$, then there exists $g \in \mathcal{V}^*$ so that $g|_{\mathcal{W}} = f$.*

Proof. Since $(\mathcal{V}, \mathcal{T})$ is a LCS, so is $\mathcal{W}$. Let Γ be a separating family of seminorms which generate the LCS topology on $\mathcal{V}$ (see Theorem 5.23). Then it is routine to verify that $\Gamma_{\mathcal{W}} := \{p|_{\mathcal{W}} : p \in \Gamma\}$ is a separating family of seminorms on $\mathcal{W}$ which generates the relative LCS topology on $\mathcal{W}$.

Suppose that $f \in \mathcal{W}^*$. By Proposition 5.29, there exist $\kappa > 0$ and $p_1, p_2, ..., p_m \in \Gamma$ so that

$$|f(w)| \le \kappa \max(p_1(w), p_2(w), ..., p_m(w)) \quad \text{for all } w \in \mathcal{W}.$$

Let $q(x) := \kappa \max(p_1(x), p_2(x), ..., p_m(x))$ for all $x \in \mathcal{V}$. Then, as is easily verified, q is a seminorm on $\mathcal{V}$, and q is continuous by Proposition 5.29. Moreover,

$$|f(w)| \le q(w) \quad \text{for all } w \in \mathcal{W}.$$

By the Hahn-Banach Theorem 02, we can find a linear functional $g : \mathcal{V} \to \mathbb{K}$ so that $g|_{\mathcal{W}} = f$ and

$$|g(x)| \le q(x) \quad \text{for all } x \in \mathcal{V}.$$

Another application of Corollary 5.31 shows that g is continuous, as was required.

□

The following is simply an application of Corollary 6.20 to the context of normed linear spaces. It is often the version that comes to mind when the Hahn-Banach Theorem is invoked.

6.21. Theorem. ***The Hahn-Banach Theorem 03***
Let $(\mathfrak{X}, \|\cdot\|)$ be a NLS, $\mathcal{M} \subseteq \mathfrak{X}$ be a linear manifold, and $f \in \mathcal{M}^*$ be a bounded linear functional. Then there exists $g \in \mathfrak{X}^*$ such that $g|_{\mathcal{M}} = f$ and $\|g\| = \|f\|$.
Proof. Consider the map

$$p: \begin{array}{ccc} \mathfrak{X} & \to & \mathbb{R} \\ x & \mapsto & \|f\|\,\|x\| \end{array}.$$

It is easy to check that p is a seminorm on $\mathfrak{X}$. (In fact, it is a norm unless $f = 0$.) Since $|f(m)| \le p(m)$ for all $m \in \mathcal{M}$, it follows from the Hahn-Banach Theorem 02 that there exists $g : \mathfrak{X} \to \mathbb{K}$ so that $g|_{\mathcal{M}} = f$ and $|g(x)| \le p(x) = \|f\|\,\|x\|$ for all $x \in \mathfrak{X}$. This last inequality shows that $\|g\| \le \|f\|$. That $\|g\| \ge \|f\|$ is clear, and hence $\|g\| = \|f\|$.

□

6.22. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\{x_j\}_{j=1}^m$ be a linearly independent set of vectors in $\mathcal{V}$. If $\{k_j\}_{j=1}^m \in \mathbb{K}$ are arbitrary, then there exists $g \in \mathcal{V}^*$ so that $g(x_j) = k_j$, $1 \le j \le m$.*
Proof. Let $\mathcal{M} = \text{span}\{x_j\}_{j=1}^m$, so that $\mathcal{M}$ is a finite-dimensional subspace of $\mathcal{V}$. Define $f : \mathcal{M} \to \mathbb{K}$ via

$$f(\sum_{j=1}^m a_j x_j) = \sum_{j=1}^m a_j k_j.$$

Then f is linear on $\mathcal{M}$, and thus, by Corollary 4.35, it is continuous. By Corollary 6.20, there exists $g \in \mathcal{V}^*$ so that $g|_{\mathcal{M}} = f$. Hence $g(x_j) = k_j$, $1 \le j \le m$.

□

A special case of the above Corollary which is worth pointing out is the following.

6.23. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $0 \ne y \in \mathcal{V}$. Then there exists $g \in \mathcal{V}^*$ so that $g(y) \ne 0$.*
Proof. Simply let $x_1 = y$ and $k_1 = 1$ in the previous Corollary.

□

As an application of these results, let us show that finite dimensional subspaces of locally convex spaces are topologically complemented.

6.24. Definition. *A closed subspace $\mathcal{W}$ of a LCS $(\mathcal{V}, \mathcal{T})$ is said to be* ***topologically complemented*** *if there exists a closed subspace $\mathcal{Y}$ of $\mathcal{V}$ so that $\mathcal{V} = \mathcal{Y} \oplus \mathcal{W}$. That is, $x \in \mathcal{V}$ implies that $x = y + w$ for some $y \in \mathcal{Y}$ and $w \in \mathcal{W}$, while $\mathcal{Y} \cap \mathcal{W} = \{0\}$.*

6.25. Remark. Every vector subspace W of a vector space V over $\mathbb{K}$ is algebraically complemented. If $\{w_\lambda\}_{\lambda\in\Lambda}$ is a basis for W, then it can be extended to a basis $\{w_\lambda\}_{\lambda\in\Lambda} \cup \{y_\beta\}_{\beta\in\Gamma}$ for V. Letting $Y = \text{span}\{y_\beta\}_{\beta\in\Gamma}$, we get $V = Y \oplus W$.

The key issue in the above definition is that if $\mathcal{W}$ is a **closed** subspace in the LCS $\mathcal{V}$, then we are asking that the complement $\mathcal{Y}$ of $\mathcal{W}$ also be **closed**. This is not always possible. For example, c_0 is a closed subspace of $(\ell^\infty, \|\cdot\|_\infty)$. Nevertheless, it does not possess a topological complement. The proof is omitted.

When $\mathcal{W}$ is finite-dimensional, the situation is somewhat better.

6.26. Proposition. *Let $\mathcal{W}$ be a finite-dimensional subspace of a LCS $(\mathcal{V}, \mathcal{T})$. Then $\mathcal{W}$ is topologically complemented in $\mathcal{V}$.*

Proof. First observe that $\mathcal{W}$ is closed in $\mathcal{V}$ by Corollary 4.22. Let $\{w_1, w_2, ..., w_n\}$ be a basis for $\mathcal{W}$.

By Corollary 6.22, we can find continuous linear functionals $\rho_1, \rho_2, ..., \rho_n \in \mathcal{V}^*$ so that $\rho_j(w_i) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. Let $\mathcal{Y} = \cap_{j=1}^n \ker \rho_j$. Since each ρ_j is continuous, $\ker \rho_j$ is closed for all j, and hence $\mathcal{Y}$ is also closed.

Suppose $v \in \mathcal{V}$. Let $k_j = \rho_j(v)$, $1 \le j \le n$. Then $w = \sum_{i=1}^n k_i w_i \in \mathcal{W}$. If $y := v - w$, then $\rho_j(y) = \rho_j(v) - \rho_j(w) = k_j - k_j = 0$, $1 \le j \le n$. Hence $y \in \mathcal{Y}$.

Finally, if $z \in \mathcal{Y} \cap \mathcal{W}$, then $z = \sum_{i=1}^n r_i w_i$ for some $r_i \in \mathbb{K}$, $1 \le i \le n$. But then $z \in \mathcal{Y}$, so for each $1 \le j \le n$, $0 = \rho_j(z) = r_j$. Hence $z = 0$ and $\mathcal{V} = \mathcal{Y} \oplus \mathcal{W}$.

□

6.27. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\mathcal{W} \subseteq \mathcal{V}$ be a closed subspace of $\mathcal{V}$. If $x \in \mathcal{V}$, $x \notin \mathcal{W}$, then there exists $g \in \mathcal{V}^*$ so that $g|_\mathcal{W} = 0$ but $g(x) \ne 0$.*

Proof. By Proposition 5.17, $\mathcal{V}/\mathcal{W}$ is a LCS. Also, $x \notin \mathcal{W}$ implies that $q(x) \ne 0$ in $\mathcal{V}/\mathcal{W}$, where $q : \mathcal{V} \to \mathcal{V}/\mathcal{W}$ is the canonical quotient map. We can therefore apply Corollary 6.23 to produce a functional $f \in (\mathcal{V}/\mathcal{W})^*$ so that $f(q(x)) \ne 0$. Since q is continuous, so is $g := f \circ q$, and so $g \in \mathcal{V}^*$ satisfies $g(w) = 0$ for all $w \in \mathcal{W}$, while $g(x) \ne 0$.

□

6.28. Theorem. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\mathcal{W} \subseteq \mathcal{V}$ be a linear manifold. Then*

$$\overline{\mathcal{W}} = \cap\{\ker f : f \in \mathcal{V}^* \text{ and } \mathcal{W} \subseteq \ker f\}.$$

Proof. Clearly $f \in \mathcal{V}^*$ implies that $\ker f$ is closed, so if $\mathcal{W} \subseteq \ker f$, then $\overline{\mathcal{W}} \subseteq \ker f$. Thus

$$\overline{\mathcal{W}} \subseteq \cap\{\ker f : f \in \mathcal{V}^* \text{ and } \mathcal{W} \subseteq \ker f\}.$$

Conversely, suppose that $x \in \mathcal{V}$, $x \notin \overline{\mathcal{W}}$. By Corollary 6.27, there exists $g \in \mathcal{V}^*$ so that $g|_\mathcal{W} = 0$ but $g(x) \ne 0$. This proves the reverse inclusion, and combining the two inclusions yields the desired result.

□

6.29. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\mathcal{W} \subseteq \mathcal{V}$ be a linear manifold. The following are equivalent:*

(a) *$\mathcal{W}$ is dense in $\mathcal{V}$.*
(b) *$f \in \mathcal{V}^*$ and $f|_{\mathcal{W}} = 0$ implies that $f = 0$.*

Let us now describe some quantitative versions of the above results, in the setting of normed linear spaces.

6.30. Corollary. *Let $(\mathfrak{X}, \|\cdot\|)$ be a NLS and $x \in \mathfrak{X}$. Then*

$$\|x\| = \max\{|x^*(x)| : x^* \in \mathfrak{X}^*, \|x^*\| \leq 1\}.$$

Proof. For the rest of the proof, the vector $x \in \mathfrak{X}$ is fixed.

Let $\beta := \sup\{|x^*(x)| : x^* \in \mathfrak{X}^*, \|x^*\| \leq 1\}$. Then for any $x^* \in \mathfrak{X}^*$ with $\|x^*\| \leq 1$, $|x^*(x)| \leq \|x^*\| \, \|x\| \leq \|x\|$, and so $\beta \leq \|x\|$.

Define $\mathfrak{Y} = \mathbb{K}x$, so that $\mathfrak{Y}$ is a one-dimensional normed, linear subspace of $\mathfrak{X}$. Define $f \in \mathfrak{Y}^\#$ via $f(kx) = k\|x\|$. Then $|f(kx)| = |k| \, \|x\| = \|kx\|$, and so $\|f\| = 1$. By the Hahn-Banach Theorem 03 (Theorem 6.21), there exists $y^* \in \mathfrak{X}^*$ so that $y^*|_{\mathfrak{Y}} = f$, and $\|y^*\| = \|f\| = 1$.

Thus

$$|y^*(x)| = y^*(x) = f(x) = \|x\|,$$

which proves that $\beta \geq \|x\|$, and hence that $\beta = \|x\|$. It also shows that the supremum is attained at y^*.

□

Recall from Proposition 2.19 that if $\mathfrak{X}$ is a normed linear space, then the canonical embedding $\mathfrak{J} : \mathfrak{X} \to \mathfrak{X}^{**}$ which sends $x \in \mathfrak{X}$ to $\widehat{x} \in \mathfrak{X}^{**}$, where $\widehat{x}(x^*) = x^*(x)$ for all $x^* \in \mathfrak{X}^*$ is a contractive linear mapping.

As a simple consequence of Corollary 6.30, we obtain:

6.31. Corollary. *The canonical embedding $\mathfrak{J} : \mathfrak{X} \to \mathfrak{X}^{**}$ is an isometry.*

6.32. Corollary. *Let $(\mathfrak{X}, \|\cdot\|)$ be a NLS and $\mathfrak{Y} \subseteq \mathfrak{X}$ be a closed subspace, with $z \in \mathfrak{X}$ but $z \notin \mathfrak{Y}$. Let $d := d(z, \mathfrak{Y}) = \|z + \mathfrak{Y}\|$. Then there exists $x^* \in \mathfrak{X}^*$ so that $\|x^*\| = 1$, $x^*|_{\mathfrak{Y}} = 0$, and $x^*(z) = d$.*

Proof. Let $q : \mathfrak{X} \to \mathfrak{X}/\mathfrak{Y}$ denote the canonical quotient map. Since $\mathfrak{X}/\mathfrak{Y}$ is a NLS and $\|q(z)\| = d$, Corollary 6.30 guarantees the existence of a linear functional $\xi^* \in (\mathfrak{X}/\mathfrak{Y})^*$ so that $\|\xi^*\| = 1$ and $\xi^*(q(z)) = \|q(z)\| = d$.

Let $x^* = \xi^* \circ q$. Obviously $x^*(z) = d$. Since $\|q\| \leq 1$, $\|x^*\| \leq \|\xi^*\| \, \|q\| \leq 1$. Also, for $y \in \mathfrak{Y}$, $x^*(y) = \xi^*(q(y)) = \xi^*(0) = 0$.

To see that $\|x^*\| \geq 1$, note that $\|\xi^*\| = 1$ and so we can find a sequence $(q(x_n))_{n=1}^\infty$ in $\mathfrak{X}/\mathfrak{Y}$ with $\|q(x_n)\| < 1$ for all $n \geq 1$ and $\lim_{n\to\infty} |\xi^*(q(x_n))| = 1$.

Choose $y_n \in \mathfrak{Y}$ so that $\|x_n + y_n\| < 1$ for all n. Then

$$\begin{aligned}\lim_{n\to\infty} |x^*(x_n + y_n)| &= \lim_{n\to\infty} |\xi^*(q(x_n + y_n)| \\ &= \lim_{n\to\infty} |\xi^*(q(x_n))| \\ &= 1.\end{aligned}$$

Hence $\|x^*\| \ge 1$, whence $\|x^*\| = 1$.

□

THE SEPARATION THEOREMS

6.33. Notation. The following shorthand will prove useful. Given a TVS $(\mathcal{V}, \mathcal{T})$ over $\mathbb{K}$, and an element $f \in \mathcal{V}^*$, we shall write $f_\mathbb{R}$ to denote the $\mathbb{R}$-linear function $f_\mathbb{R} := \mathrm{Re}\, f : \mathcal{V} \to \mathbb{R}$. Of course, if $\mathbb{K} = \mathbb{R}$, then $f_\mathbb{R} = f$, while if $\mathbb{K} = \mathbb{C}$, we have seen in Lemma 6.6 that f is the complexification of $f_\mathbb{R}$.

6.34. Proposition. *Let $(\mathcal{V}, \mathcal{T})$ be a TVS over $\mathbb{R}$ and let $\varnothing \ne G \subseteq \mathcal{V}$ be an open, convex subset of $\mathcal{V}$ with $0 \notin \mathcal{G}$. Then there exists a continuous linear functional $g \in \mathcal{V}^*$ such that*

$$\ker g \cap G = \varnothing.$$

Equivalently, there exists a closed hyperplane $\mathcal{M}$ in $\mathcal{V}$ such that $G \cap \mathcal{M} = \varnothing$.

Proof. Fix $x_0 \in G$ and let $H = x_0 - G$. Then $H \in \mathcal{U}_0^{\mathcal{V}}$ is open and convex. Let p_H denote the Minkowski functional on H. By Proposition 5.10,

$$H = \{x \in \mathcal{V} : p_H(x) < 1\}.$$

Observe that $0 \notin G$ implies that $x_0 \notin H$. Thus $p_H(x_0) \ge 1$. Let $\mathcal{W} = \mathbb{R}x_0$, and define $f : \mathcal{W} \to \mathbb{R}$ via $f(kx_0) = kp_H(x_0)$. Clearly $f \in \mathcal{W}^\#$. Moreover,

- if $k \ge 0$, then $f(kx_0) = kp_H(x_0) = p_H(kx_0)$, while
- if $k < 0$, then $f(kx_0) = kp_H(x_0) < 0 \le p_H(kx_0)$.

It follows from the Hahn-Banach Theorem 01 (Theorem 6.18) that there exists a linear functional $g : \mathcal{V} \to \mathbb{R}$ with $g|_\mathcal{W} = f$ and $g(x) \le p_H(x)$ for all $x \in \mathcal{V}$. Of course, the fact that $f \ne 0$ implies that $g \ne 0$. Suppose that $y \in H$. Then

$$g(y) \le p_H(y) < 1.$$

By Proposition 6.12, g is continuous on $\mathcal{V}$. Suppose $z \in G$. Then $x_0 - z \in H$, so $g(x_0) - g(z) = g(x_0 - z) \le p_H(x_0 - z) < 1$. On the other hand, $g(x_0) = f(x_0) = p_H(x_0) \ge 1$, and so

$$g(z) > g(x_0) - 1 \ge 0, \text{ and } z \notin \ker g.$$

Thus $\ker g \cap G = \varnothing$.

The last statement follows immediately by simply setting $\mathcal{M} = \ker g$.

□

6.35. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a TVS over $\mathbb{C}$ and let $\varnothing \neq G \subseteq \mathcal{V}$ be an open, convex subset of $\mathcal{V}$ with $0 \notin \mathcal{G}$. Then there exists a continuous linear functional $g \in \mathcal{V}^*$ such that*

$$\ker\, g_{\mathbb{R}} \cap G = \varnothing.$$

In particular, $\ker\, g \cap G = \varnothing$.

Proof. Since $\mathcal{V}$ is a $\mathbb{C}$-linear space, it is also an $\mathbb{R}$-linear space, and so by Proposition 6.34 above we can find a continuous $\mathbb{R}$-linear functional $h : \mathcal{V} \to \mathbb{R}$ so that $G \cap \ker\, h = \varnothing$. Let $g = h_{\mathbb{C}}$ be the complexification of h, namely $g(x) = h(x) - ih(ix)$, $x \in \mathcal{V}$. By Lemma 6.6, g is a $\mathbb{C}$-linear functional and $g_{\mathbb{R}} = h$. Clearly g is continuous, since both its real and imaginary parts are.

The last statement is obvious.

□

6.36. Definition. *An* ***affine hyperplane*** $\mathcal{M}$ *in a TVS* $(\mathcal{V}, \mathcal{T})$ *is a translate of a hyperplane; that is,* $\mathcal{M}$ *is an affine hyperplane if there exists* $x \in \mathcal{M}$ *so that* $\mathcal{M} - x$ *is a hyperplane.*

More generally, $\mathcal{L} \subseteq \mathcal{V}$ *is an* ***affine manifold*** *(resp.* ***affine subspace****) of* $\mathcal{V}$ *if there exists* $m \in \mathcal{L}$ *so that* $\mathcal{L} - m$ *is a manifold (resp. subspace) of* $\mathcal{V}$.

We remark that if there exists $m \in \mathcal{L}$ so that $\mathcal{L} - m$ is a manifold in $\mathcal{V}$, then for all $m \in \mathcal{L}$ we must have $\mathcal{L} - m$ is a manifold. The verification of this is left to the reader.

6.37. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a TVS and $\varnothing \neq G \subseteq \mathcal{V}$ be open and convex. If $\mathcal{L} \subseteq \mathcal{V}$ is an affine manifold of $\mathcal{V}$ and $\mathcal{L} \cap G = \varnothing$, then there exists a closed, affine hyperplane $\mathcal{Y} \subseteq \mathcal{V}$ so that $\mathcal{L} \subseteq \mathcal{Y}$ and $\mathcal{Y} \cap G = \varnothing$.*

Proof. Since the closure $\overline{\mathcal{L}}$ of $\mathcal{L}$ is an affine subspace and $\overline{\mathcal{L}} \cap G = \varnothing$ (as $\mathcal{L} \subseteq \mathcal{V} \setminus G$ and the latter set is closed), we may replace $\mathcal{L}$ by $\overline{\mathcal{L}}$ and assume without loss of generality that $\mathcal{L}$ is closed. Choose $m \in \mathcal{L}$ and let $\mathcal{L}_0 = \mathcal{L} - m$, so that $\mathcal{L}_0$ is a closed subspace of $\mathcal{V}$. Let $G_0 = G - m$. Since $\mathcal{L} \cap G = \varnothing$, it follows that $\mathcal{L}_0 \cap G_0 = \varnothing$. Let $q : \mathcal{V} \to \mathcal{V}/\mathcal{L}_0$ denote the canonical quotient map.

Since G is open, so is G_0. Since q is an open map (see paragraph 4.18), $q(G_0)$ is open. Furthermore, G is convex and hence so are G_0 and $q(G_0)$. Again, since $\mathcal{L}_0 \cap G_0 = \varnothing$, $0 \notin q(G_0)$. By Proposition 6.34 and Corollary 6.35, there exists a closed hyperplane $\mathcal{N}_0$ in $\mathcal{V}/\mathcal{L}_0$ so that $\mathcal{N}_0 \cap q(G_0) = \varnothing$. Let $\mathcal{Y}_0 = q^{-1}(\mathcal{N}_0)$. It is routine to check that $\mathcal{Y}_0$ is a linear manifold in $\mathcal{V}$, and $\mathcal{Y}_0$ is closed since q is continuous. Moreover,

$$\begin{aligned} \dim \mathcal{V}/\mathcal{Y}_0 &= \dim(\mathcal{V}/\mathcal{L}_0)/(\mathcal{Y}_0/\mathcal{L}_0) \\ &= \dim(\mathcal{V}/\mathcal{L}_0)/\mathcal{N}_0 \\ &= 1, \end{aligned}$$

and so $\mathcal{Y}_0$ is a closed hyperplane in $\mathcal{V}$ with $\mathcal{L}_0 \subseteq \mathcal{Y}_0$. Translating back, let $\mathcal{Y} = \mathcal{Y}_0 + m$. Then $\mathcal{Y}$ is a closed affine hyperplane of $\mathcal{V}$, $\mathcal{L} \subseteq \mathcal{Y}$ and if $z \in \mathcal{Y} \cap G$, then $q(z-m) \in q(\mathcal{Y}_0) \cap q(G_0) = \mathcal{N}_0 \cap q(G_0) = \varnothing$, a contradiction.

□

6.38. Definition. *Let $(\mathcal{V}, \mathcal{T})$ be an TVS over $\mathbb{K}$, and let $f \in \mathcal{V}^*$. We refer to a set of the form*

$$\mathcal{S} := \{x \in \mathcal{V} : f_{\mathbb{R}}(x) \geq \kappa\} \qquad (\text{resp. } \mathcal{S}^\circ := \{x \in \mathcal{V} : f_{\mathbb{R}}(x) > \kappa\})$$

*as a **closed half-space** (resp. an **open half-space**).*

Note that if $0 \neq f \in \mathcal{V}^*$ and $\kappa \in \mathbb{R}$, then $g := -f \in \mathcal{V}^*$ and

$$\{x \in \mathcal{V} : f_{\mathbb{R}}(x) < \kappa\} = \{x \in \mathcal{V} : g_{\mathbb{R}}(x) > -\kappa\},$$

and a similar statement holds for closed half-spaces.

6.39. Definition. *Let $(\mathcal{V}, \mathcal{T})$ be an TVS over the field $\mathbb{K}$. We say that two subsets A and B of $\mathcal{V}$ are **separated** if we can find $0 \neq f \in \mathcal{V}^*$ and $\kappa \in \mathbb{R}$ such that*

$$A \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) \geq \kappa\} \quad \text{and} \quad B \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) \leq \kappa\}.$$

*We say that A and B are **strictly separated** if we can find $0 \neq f \in \mathcal{V}^*$ and $\kappa \in \mathbb{R}$ such that*

$$A \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) > \kappa\} \quad \text{and} \quad B \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) < \kappa\}.$$

6.40. Example.

(a) Consider $\mathbb{R}^2$ equipped with the Euclidean norm. Let $A = \{(x,y) \in \mathbb{R}^2 : x < 0 \text{ and } y \geq 1/x^2\}$, $B = \{(x,y) \in \mathbb{R}^2 : x > 0 \text{ and } y \geq 1/x^2\}$. Then $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x$ defines a continuous linear functional on $\mathbb{R}^2$.

Then

$$A \subseteq \{(x,y) \in \mathbb{R}^2 : f(x) < 0\} \quad \text{and} \quad B \subseteq \{(x,y) \in \mathbb{R}^2 : f(x) > 0\}.$$

Hence A and B are strictly separated.

(b) With $\mathbb{R}^2$, f and A as above, set $C = \{(0,y) : y \in \mathbb{R}\}$. Then

$$A \subseteq \{(x,y) \in \mathbb{R}^2 : f(x) \leq 0\} \quad \text{and} \quad C \subseteq \{(x,y) \in \mathbb{R}^2 : f(x) \geq 0\}.$$

Thus A and C are separated. We leave it as an exercise for the reader to verify that A and C are not strictly separated.

6.41. Theorem. ***The Hahn-Banach Theorem 04***
Let $(\mathcal{V}, \mathcal{T})$ be a TVS over $\mathbb{K}$ and suppose that A and B are non-empty, disjoint, open, convex subsets of $\mathcal{V}$. Then A and B are strictly separated.
Proof. Let $G := A - B = \{a - b : a \in A, b \in B\}$. We claim that $\varnothing \neq G$ is open and convex. That $\varnothing \neq G$ is obvious as both A and B are non-empty. Since $G = \cup_{b \in B} A - b$, G is the union of open sets (each $A - b$ is a translate of the open set A), and thus G is open.

Suppose that $g_1 = a_1 - b_1$ and $g_2 = a_2 - b_2$ lie in G. Let $t \in [0, 1]$. Then

$$tg_1 + (1 - t)g_2 = [ta_1 + (1 - t)a_2] - [tb_1 + (1 - t)b_2].$$

Since A and B are convex, it follows that so is G. Observe that $A \cap B = \varnothing$ also implies that $0 \notin G$.

It now follows from Proposition 6.34 (in the case where $\mathbb{K} = \mathbb{R}$ or alternatively from Corollary 6.35 (in the case where $\mathbb{K} = \mathbb{C}$) that there exists a continuous linear functional $f \in \mathcal{V}^*$ such that

$$\ker f_{\mathbb{R}} \cap G = \varnothing.$$

Now G is convex and $f_{\mathbb{R}}$ is $\mathbb{R}$-linear, whence $f_{\mathbb{R}}(G) \subseteq \mathbb{R}$ is again convex. But $G \cap \ker f_{\mathbb{R}} = \varnothing$, so $0 \notin f_{\mathbb{R}}(G)$ and hence either $f_{\mathbb{R}}(x) > 0$ for all $x \in G$, or $f_{\mathbb{R}}(x) < 0$ for all $x \in G$. By replacing f by $-f$ if necessary, we may assume that the first condition holds. If $a \in A, b \in B$, then $c = a - b \in G$, so $f_{\mathbb{R}}(c) = f_{\mathbb{R}}(a) - f_{\mathbb{R}}(b) > 0$, i.e. $f_{\mathbb{R}}(b) < f_{\mathbb{R}}(a)$. We deduce that there exists $\kappa \in \mathbb{R}$ so that

$$\sup\{f_{\mathbb{R}}(b) : b \in B\} \leq \kappa \leq \inf\{f_{\mathbb{R}}(a) : a \in A\}.$$

Now A is open and hence $f_{\mathbb{R}}(A)$ is open (check!). Similarly, $f_{\mathbb{R}}(B)$ is open. Hence $f_{\mathbb{R}}(b) < \kappa < f_{\mathbb{R}}(a)$ for all $a \in A, b \in B$. It follows that

$$A \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) > \kappa\} \quad \text{and} \quad B \subseteq \{x \in \mathcal{V} : f_{\mathbb{R}}(x) < \kappa\},$$

which is the statement that A and B are strictly separated.

□

6.42. Remark. If A were open but not B, then $G = \cup_{b \in B} A - b$ would still be a union of open sets and hence would still be open. The conclusion would be there there exist a continuous linear functional $g \in \mathcal{V}^*$ and a constant $\kappa \in \mathbb{R}$ such that $A \subseteq \{x \in \mathcal{V} : g_{\mathbb{R}}(x) > \kappa\}$ and $B \subseteq \{x \in \mathcal{V} : g_{\mathbb{R}}(x) \leq \kappa\}$.

Our last version of the Hahn-Banach Theorem is the first separation theorem that requires $\mathcal{V}$ to be a locally convex space, rather than simply a topological vector space.

6.43. Theorem. ***The Hahn-Banach Theorem 05***
Let $(\mathcal{V}, \mathcal{T})$ be a LCS and suppose that $A, B \subseteq \mathcal{V}$ are non-empty, disjoint, closed, convex subsets of $\mathcal{V}$. Suppose furthermore that B is compact. Then there are real numbers α, β and a continuous linear functional $f \in \mathcal{V}^*$ so that

$$f_{\mathbb{R}}(a) \geq \alpha > \beta \geq f_{\mathbb{R}}(b)$$

for all $a \in A, b \in B$. In particular, A and B are strictly separated.

Proof. Observe that $\mathcal{V} \setminus A$ is open and that $b \in B$ implies that $b \in \mathcal{V} \setminus A$. It follows from Corollary 5.15 that we can find a balanced, convex, open nbhd N_b of 0 so that $b + N_b \subseteq \mathcal{V} \setminus A$. The collection $\{b + \frac{1}{3}N_b : b \in B\}$ is an open cover of B, and $B \subseteq \cup_{b\in B}\,(b + \frac{1}{3}N_b) \subseteq \cup_{b\in B}\,(b + N_b) \subseteq \mathcal{V} \setminus A$. Since B is compact, we can find a finite subcover $\{b_j + \frac{1}{3}N_{b_j}\}_{j=1}^n$ of B. Let $N = \cap_{j=1}^n \frac{1}{3}N_{b_j}$. Then N is non-empty, balanced, convex and open.

Let $A_0 = A+N = \{a+n : a \in A, n \in N\}$ and $B_0 = B+N$. Clearly $A_0 \neq \varnothing \neq B_0$. Then $A_0 = \cup_{a\in A} a + N$ is open and similarly B_0 is open. If $a_1 + n_1, a_2 + n_2 \in A_0$ and $t \in [0, 1]$, then

$$t(a_1 + n_1) + (1 - t)(a_2 + n_2) = (ta_1 + (1 - t)a_2) + (tn_1 + (1 - t)n_2) \in A + N = A_0,$$

since each of A and N is convex. Thus A_0, and similarly B_0, is convex.

Suppose $z \in A_0 \cap B_0$. Then there exists $a \in A$, $b \in B$ and $n_1, n_2 \in N$ so that $a + n_1 = b + n_2$. But $b \in B$ implies that there exists $1 \leq j \leq n$ and some $m_j \in \frac{1}{3}N_{b_j}$ so that $b = b_j + m_j$. Thus $a = b + (n_2 - n_1) = b_j + m_j + (n_2 - n_1)$. Now $n_1 \in N$ and N balanced implies that $-n_1 \in N$. Recalling that $N \subseteq \frac{1}{3}N_{b_j}$, we see that $m_j, n_2, -n_1 \in \frac{1}{3}N_{b_j}$, and thus $m_j + (n_2 - n_1) \in N_{b_j}$. This forces $a = b_j + m_j + (n_2 - n_1) \in b_j + N_{b_j}$, contradicting the fact that $b_j + N_{b_j} \subseteq \mathcal{V}\setminus A$. Hence $A_0 \cap B_0 = \varnothing$.

By Theorem 6.41 (HB04), there exists $f \in \mathcal{V}^*$ and $\alpha \in \mathbb{R}$ so that

$$f_{\mathbb{R}}(a) > \alpha > f_{\mathbb{R}}(b)$$

for all $a \in A_0$, $b \in B_0$.

But B is compact, and $f_{\mathbb{R}}$ is continuous on B, so that $\beta = \sup\{f_{\mathbb{R}}(b) : b \in B\}$ is attained at some point $b_0 \in B$. Thus

$$f_{\mathbb{R}}(a) > \alpha > \beta = f_{\mathbb{R}}(b_0) \geq f_{\mathbb{R}}(b)$$

for all $a \in A, b \in B$.

□

6.44. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS over $\mathbb{K}$ and $\varnothing \neq A \subseteq \mathcal{V}$. Then the closed, convex hull of A, $\overline{\text{co}}(A)$, is the intersection of the closed half spaces that contain A.*

Proof. Let $\Omega = \{\mathcal{S} : A \subseteq \mathcal{S}, \mathcal{S} \subseteq \mathcal{V}$ is a closed half-space$\}$. Since each $\mathcal{S} \in \Omega$ is closed and convex, $B = \cap_{\mathcal{S}\in\Omega}\mathcal{S}$ is again closed and convex. Clearly $A \subseteq B$, and so the closed convex hull of A is also a subset of B.

If $z \notin \overline{\text{co}}(A)$, then $\{z\}$ and $\overline{\text{co}}(A)$ are disjoint, non-empty, closed and convex subsets of $\mathcal{V}$. Since $\{z\}$ is compact, we can apply Theorem 6.43 (i.e. HB05) to obtain a linear functional $f \in \mathcal{V}^*$ and $\alpha, \beta \in \mathbb{R}$ such that

$$f_{\mathbb{R}}(z) \geq \alpha > \beta \geq f_{\mathbb{R}}(y)$$

for all $y \in \overline{\text{co}}(A)$. Thus if $\mathcal{S}_0 = \{x \in \mathcal{V} : f_{\mathbb{R}}(x) \leq \beta\}$, then $\mathcal{S}_0$ is a closed half-space of $\mathcal{V}$, $A \subseteq \overline{\text{co}}(A) \subseteq \mathcal{S}_0$, and $z \notin \mathcal{S}_0$. Thus $\cap_{\mathcal{S}\in\Omega}\mathcal{S} \subseteq \overline{\text{co}}(A)$, proving that

$$\overline{\text{co}}(A) = \cap_{\mathcal{S}\in\Omega}\mathcal{S}.$$

□

Appendix to Section 6.

6.45. Corollary 6.30 admits an interesting interpretation. Given $(\mathfrak{X}, \|\cdot\|)$ a NLS and $x^* \in \mathfrak{X}^*$, we do not in general expect x^* to achieve its norm. For example, if $\mathfrak{X} = c_0$, equipped with the supremum norm, and if

$$x^*((x_n)_n) := \sum_n \frac{x_n}{2^n},$$

then x^* has norm one, but there is no $x = (x_n)_n \in c_0$ with $\|x\| = 1$ and $|x^*(x)| = \|x^*\| = 1$.

Nevertheless, Corollary 6.30 allows us to conclude that we may find $x^{**} \in \mathfrak{X}^{**}$ with $\|x^{**}\| = 1$ for which $|x^{**}(x^*)| = 1$. If $\mathfrak{J} : \mathfrak{X}^* \to \mathfrak{X}^{***}$ is the canonical embedding of $\mathfrak{X}^*$ into its second dual, then $\|\widehat{x^*}\| = 1$ by Corllary 6.31 and

$$|\widehat{x^*}(x^{**})| = |x^{**}(x^*)| = 1 = \|\widehat{x^*}\|.$$

One can think of this as saying that the domain of x^* is not "large enough" to allow x^* to attain its norm, but that $\mathfrak{X}^{**}$ extends the domain of x^* enough to allow the extension $\widehat{x^*}$ of x^* to attain its norm.

6.46. The proof of Proposition 6.34 also gives us an indication of how one may try to interpret the Hahn-Banach Theorem 01, namely Theorem 6.18, geometrically.

Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ and let p be a sublinear functional on $\mathcal{V}$. Suppose that $\mathcal{M}$ is a linear manifold in $\mathcal{V}$ and $f : \mathcal{M} \to \mathbb{R}$ is a linear functional with $f(m) \le p(m)$ for all $m \in \mathcal{M}$.

Let $H = \{x \in \mathcal{V} : p(x) < 1\}$. It is routine to verify that H is convex. Since $0 \ne f$, there exists $m_0 \in \mathcal{M}$ such that $f(m_0) > 1$. This forces $p(m_0) \ge f(m_0) > 1$, and so $m_0 \notin H$. Let $K = m_0 - H = \{m_0 - h : h \in H\}$. Clearly K is also convex. Since $m_0 \notin H$, $0 \notin K$. In fact, we claim that $K \cap \ker f = \varnothing$.

Suppose otherwise: let $k \in K \cap \ker f$. Then $k \in \mathcal{M}$ and $k = m_0 - h$ for some $h \in H$, which forces $h \in H \cap \mathcal{M}$, since $k, m_0 \in \mathcal{M}$. But

$$0 = f(k) = f(m_0 - h) = f(m_0) - f(h) > 1 - p(h) > 0,$$

a contradiction. Thus $K \cap \ker f = \varnothing$. Theorem 6.18 then says that we can extend f to a linear functional $g : \mathcal{V} \to \mathbb{R}$ with $g|_{\mathcal{M}} = f$, such that

$$g(x) \le p(x) \text{ for all } x \in \mathcal{V}.$$

A similar analysis to that above shows that $K \cap \ker g = \varnothing$. In other words, one can translate the "unit ball" H of $\mathcal{V}$ (as measured by the sublinear functional p – note, p doesn't even have to be a non-negative-valued function, and hence the interpretation of H as a "unit ball" here is *very, very loose*) so that it doesn't intersect the linear manifold $\ker f$ in such a way that the manifold may be extended to a hyperplane ($\ker g$) which also doesn't intersect the translation.

This is only intended as a heuristic. Proposition 6.34 shows how to correctly use HB01, i.e. Theorem 6.18, to obtain an interesting geometric result in locally convex spaces.

Part (b) of Proposition 6.7 admits a second proof (by induction). We thank W. Shen for pointing this out and providing the following proof.

6.47. Proposition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$, $N \geq 1$ be an integer, and $g, f_1, f_2, \ldots, f_N \in \mathcal{V}^\#$ be linear functionals. Suppose that $\cap_{n=1}^N \ker f_n \subseteq \ker g$. Then $g \in \operatorname{span}\{f_1, f_2, \ldots, f_N\}$.*

Proof. The case $N = 1$ is part (a) of Proposition 6.7. Now let $M > 1$ and suppose that the result holds for $N < M$. We prove that it holds for $N = M$. To that end, we suppose that $g, f_1, f_2, \ldots, f_M \in \mathcal{V}^\#$, and that $\cap_{n=1}^M \ker f_n \subseteq \ker g$.

Let $\mathcal{W} := \cap_{n=1}^{M-1} \ker f_n$. Note that

$$\ker f_M|_{\mathcal{W}} \subseteq \ker g|_{\mathcal{W}}.$$

Since the result holds for $N = 1$, we see that $g|_{\mathcal{W}} = \kappa\, f_M|_{\mathcal{W}}$ for some $\kappa \in \mathbb{K}$. But then

$$\cap_{n=1}^{M-1} \ker f_n = \mathcal{W} \subseteq \ker (g - \kappa f_M),$$

and so by the induction hypothesis,

$$g - \kappa f_M \in \operatorname{span}\{f_1, f_2, \ldots, f_{M-1}\},$$

from which the result follows.

□

The following result was required for Proposition 6.7. For the sake of completeness, we include a proof.

6.48. Proposition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$ and $1 \leq N \in \mathbb{N}$. Suppose that $f_1, f_2, \ldots, f_N \in \mathcal{V}^\#$ are linearly independent elements, and let $\mathcal{W} := \cap_{n=1}^N \ker f_n$. Then*

$$\dim(\mathcal{V}/\mathcal{W}) \leq N.$$

Proof. To see this, we shall argue by induction on N. For $N = 1$, this is clear, as $f_1 \neq 0$ (because $\{f_1\}$ is assumed to be linearly independent), so

$$\frac{\mathcal{V}}{\ker f_1} \simeq \operatorname{ran} f_1 = \mathbb{K},$$

implying that $\dim \dfrac{\mathcal{V}}{\ker f_1} = \dim \mathbb{K} = 1$.

Another way of stating this is that if $x_1 \in \mathcal{V}$ and $x_1 \notin \ker f_1$, then $\mathcal{V} = \ker f_1 \dot{+} \mathbb{K}x_1$. (Here, $\mathcal{V} = \mathcal{V}_1 \dot{+} \mathcal{V}_2$ for vector subspaces $\mathcal{V}_1, \mathcal{V}_2$ of $\mathcal{V}$ if $\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 = \{x + y : x \in \mathcal{V}_1, y \in \mathcal{V}_2\}$ and $\mathcal{V}_1 \cap \mathcal{V}_2 = \{0\}$.)

Suppose that the result holds for $1 \le N \le M$. Consider linearly independent linear functionals $f_1, f_2, \ldots, f_M, f_{M+1}$ on $\mathcal{V}$. Let $\mathcal{W} := \cap_{n=1}^{M} \ker f_n$. Then, by the induction hypothesis,

$$\dim(\mathcal{V}/\mathcal{W}) \le M.$$

Alternatively, there exist $r \le M$ linearly independent vectors $x_1, x_2, \ldots, x_r$ such that

$$\mathcal{V} = \mathcal{W} + \operatorname{span}\{x_1, x_2, \ldots, x_r\}.$$

(One way to see this is to choose a basis $\{x_1 + \mathcal{W}, x_2 + \mathcal{W}, \ldots, x_r + \mathcal{W}\}$ for $\mathcal{V}/\mathcal{W}$, and to check that the representatives $x_1, x_2, \ldots, x_r$ of these cosets does the job.)

Let $g := f_{M+1}|_{\mathcal{W}}$. Then g is a linear functional on $\mathcal{W}$, and so $\dim(\mathcal{W}/\ker g) \le 1$. That is, there exists $y \in \mathcal{W}$ such that

$$\mathcal{W} = \ker\, g + \mathbb{K}\, y.$$

(Observe that this holds even if $\ker\, g = \mathcal{W}$.) If $w \in \ker\, g$, then $f_{M+1}(w) = g(w) = 0$, and $f_n(w) = 0$, $1 \le n \le M$ because $\ker\, g \subseteq \mathcal{W} = \cap_{n=1}^{M} \ker\, f_n$.

It follows that

$$\ker\, g \subseteq \cap_{n=1}^{M+1} \ker\, f_n,$$

and

$$\mathcal{V} = \mathcal{W} + \operatorname{span}\{x_1, x_2, \ldots, x_r\} = \ker\, g + \operatorname{span}\{x_1, x_2, \ldots, x_r, y\}.$$

In particular, $\mathcal{V}/\ker\, g$ is spanned by $\{x_1 + \ker\, g, x_2 + \ker\, g, \ldots, x_r + \ker\, g, y + \ker\, g\}$.

Hence $\dim \dfrac{\mathcal{V}}{\ker\, g} \le r + 1 \le M + 1$, which, when combined with the fact that $\ker\, g \subseteq \cap_{n=1}^{M+1} \ker\, f_n$, implies that

$$\dim \frac{\mathcal{V}}{\cap_{n=1}^{M+1} \ker\, f_n} \le M + 1.$$

□

*

There is less in this than meets the eye.

Tallulah Bankhead

Exercises for Section 6.

Question 1.

Let $\mathcal{V}$ be a vector space over $\mathbb{C}$, and let $f : \mathcal{V} \to \mathbb{R}$ be an $\mathbb{R}$-linear functional. Let $f_{\mathbb{C}}$ denote the $\mathbb{C}$-linear functional defined in Lemma 6.6, namely

$$f_{\mathbb{C}}(x) := f(x) - if(ix), \quad x \in \mathcal{V}.$$

Prove that $\|f\| = \infty$ if and only if $\|f_{\mathbb{C}}\| = \infty$.

Question 2.

For those of you who have studied Lebesgue measure: let $1 \le p < \infty$, and let q denote the Lebesgue conjugate of p, and dm denote Lebesgue measure on the real line. Let $f \in L^p([0,1], \mathbb{K})$. Prove that there exists $g \in L^q([0,1], \mathbb{K})$ such that

(i) $\|g\|_q = 1$, and

(ii)

$$\int_0^1 f\overline{g}dm = \|f\|_p.$$

This is typically one of the first things that one proves about L^p-spaces in any course on measure theory. Hopefully you can now see why.

Question 3.

Find a linear functional $\varphi \in (\ell^\infty)^*$ such that $\varphi|_{c_0} \equiv 0$. That is, describe φ explicitly.

Question 4.

Let $1 \le p \le \infty$. Consider $x = (1,2) \in (\mathbb{C}^2, \|\cdot\|_p)$. Find a linear functional $\varphi_p \in (\mathbb{C}^2, \|\cdot\|_p)^*$ such that $\|\varphi_p\| = 1$ and

$$\varphi_p(x) = \|x\|_p.$$

7. Weak topologies and dual spaces

Last week I stated that this woman was the ugliest woman I had ever seen. I have since been visited by her sister and now wish to withdraw that statement.

Mark Twain

7.1. In Remarks 5.27, we observed that if $\mathcal{V}$ is a vector space over $\mathbb{K}$, if $(\mathfrak{X}_\alpha, \|\cdot\|_\alpha)_{\alpha\in A}$ is a family of normed linear spaces, and if for each $\alpha \in A$ we have a linear map $T_\alpha : \mathcal{V} \to \mathfrak{X}_\alpha$, then each

$$\begin{array}{rccc} p_\alpha : & \mathcal{V} & \to & \mathbb{R} \\ & x & \mapsto & \|T_\alpha x\| \end{array}$$

is a seminorm. Furthermore, if $\{T_\alpha\}_{\alpha\in A}$ is separating – i.e. for each $0 \neq x \in \mathcal{V}$, there exists $\alpha_0 \in A$ so that $T_{\alpha_0} x \neq 0$ – then the family $\Gamma = \{p_\alpha\}_{\alpha\in A}$ is a separating family of seminorms. Finally, we saw there that the LCS topology on $\mathcal{V}$ generated by Γ was nothing more (nor was it anything less) than the weak topology generated by $\{T_\alpha\}_{\alpha\in A}$.

Let us now consider the following special instance of this phenomenon. Again, we begin with a vector space $\mathcal{V}$ over $\mathbb{K}$, and we assume that we are given a separating family $\Omega \subseteq \mathcal{V}^\#$. Of course, for each $\varrho \in \Omega$, we have

$$\varrho : \mathcal{V} \to \mathbb{K},$$

and $(\mathbb{K}, |\cdot|)$ is a normed linear space. Since Ω was assumed to be separating for $\mathcal{V}$, the family $\Gamma = \{\tau_\varrho : \varrho \in \Omega\}$ of functions defined by

$$\tau_\varrho(x) = |\varrho(x)|, \quad x \in \mathcal{V},$$

is a separating family of seminorms which generates a LCS topology on $\mathcal{V}$. From above, this topology coincides with the weak topology generated by Ω, and we shall denote it by $\sigma(\mathcal{V}, \Omega)$.

Thus a base for the $\sigma(\mathcal{V}, \Omega)$ topology on $\mathcal{V}$ is given by

$$\mathcal{B} = \{N(x, F, \varepsilon) : x \in \mathcal{V}, \varepsilon > 0, F \subseteq \Omega \text{ finite}\},$$

where for each x, F and $\varepsilon > 0$ as above,

$$N(x, F, \varepsilon) = \{y \in \mathcal{V} : \tau_\varrho(x - y) = |\varrho(x) - \varrho(y)| < \varepsilon, \varrho \in F\}.$$

In particular, a net $(x_\lambda)_{\lambda\in\Lambda}$ in $(\mathcal{V}, \sigma(\mathcal{V}, \Omega))$ converges to $x \in \mathcal{V}$ if and only if

$$\lim_\lambda \tau_\rho(x_\lambda - x) = \lim_\lambda |\rho(x_\lambda) - \rho(x)| = 0,$$

or equivalently,

$$\lim_\lambda \rho(x_\lambda) = \rho(x)$$

for all $\rho \in \Omega$.

7.2. Definition. *Let $\mathcal{V}$ be a vector space over $\mathbb{K}$, and suppose that $\mathcal{L} \subseteq \mathcal{V}^\#$ is both a linear manifold and a separating family of linear functionals. We say that $(\mathcal{V}, \mathcal{L})$ is a* ***dual pair****.*

7.3. Example. Suppose that $(\mathcal{V}, \mathcal{T})$ is a LCS and that $\mathcal{L} = \mathcal{V}^*$. By Corollary 6.23, $\mathcal{L}$ separates points of $\mathcal{V}$ and hence $(\mathcal{V}, \mathcal{V}^*)$ is a dual pair. The $\sigma(\mathcal{V}, \mathcal{V}^*)$ topology is sufficiently important to merit its own name, and we refer to it as **the weak topology** on $\mathcal{V}$. If $(x_\lambda)_\lambda$ is a net in $\mathcal{V}$ which converges to some x in the weak topology, we say that $(x_\lambda)_\lambda$ **converges weakly** to x.

Suppose that $(x_\lambda)_\lambda$ is a net in $\mathcal{V}$ which converges to $x \in \mathcal{V}$ in the initial topology $\mathcal{T}$. For any $f \in \mathcal{V}^*$, the fact that f is continuous implies that

$$\lim_\lambda f(x_\lambda) = f(x).$$

Thus $(x_\lambda)_\lambda$ converges to x weakly. It follows that the weak topology on $\mathcal{V}$ induced by $\mathcal{V}^*$ is weaker than the initial topology: in other words, $\sigma(\mathcal{V}, \mathcal{V}^*) \subseteq \mathcal{T}$.

Let $(V, \mathcal{L})$ be a dual pair. By Paragraph 7.1, each $\rho \in \mathcal{L}$ is continuous on V relative to the $\sigma(V, \mathcal{L})$ topology. Our present goal is to show that these are the **only** $\sigma(V, \mathcal{L})$-continuous linear functionals on V.

7.4. Theorem. *Let $(V, \mathcal{L})$ be a dual pair. Then*

$$\mathcal{L} = (V, \sigma(V, \mathcal{L}))^*.$$

Proof. That $\mathcal{L}$ is contained in $(V, \sigma(V, \mathcal{L}))^*$ was shown in Paragraph 7.1.

Suppose now that $\mu \in V^\#$ is $\sigma(V, \mathcal{L})$-continuous. Then the map $p_\mu : V \to \mathbb{R}$ satisfying $p_\mu(x) = |\mu(x)|$ defines a $\sigma(V, \mathcal{L})$-continuous seminorm on V. By Proposition 5.29, there exist $\rho_1, \rho_2, ..., \rho_n \in \mathcal{L}$ and $0 < \kappa \in \mathbb{R}$ so that

$$p_\mu(x) = |\mu(x)| \le \kappa \ \max(|\rho_1(x)|, |\rho_2(x)|, ..., |\rho_n(x)|) \text{ for all } x \in V.$$

It follows that $\ker \mu \supseteq \cap_{j=1}^n \ker \rho_j$. By Proposition 6.7, $\mu \in \operatorname{span} \{\rho_j\}_{j=1}^n \subseteq \mathcal{L}$.

□

7.5. Remark.

- We first remark that if $\Omega \subseteq \mathcal{V}^\#$ is a separating family of linear functionals but is not a linear manifold, then after setting $\mathcal{L} = \operatorname{span} \Omega$, one can verify that the $\sigma(\mathcal{V}, \mathcal{L})$-topology on $\mathcal{V}$ agrees with the $\sigma(\mathcal{V}, \Omega)$-topology, and hence that

$$\mathcal{L} = (\mathcal{V}, \sigma(V, \Omega))^*.$$

- It follows from Theorem 7.4 that the only weakly continuous linear functionals on a locally convex space $(\mathcal{V}, \mathcal{T})$ are the elements of $(\mathcal{V}, \mathcal{T})^*$.

7.6. Definition. *Suppose that $(\mathcal{V}, \mathcal{T})$ is a LCS. Then $\mathcal{V}^* \subseteq \mathcal{V}^\#$ is a vector space over $\mathbb{K}$. For each $x \in \mathcal{V}$, define*

$$\begin{array}{rccc} \widehat{x}: & \mathcal{V}^* & \to & \mathbb{K} \\ & \rho & \mapsto & \widehat{x}(\rho) := \rho(x). \end{array}$$

Then $\widehat{\mathcal{V}} := \{\widehat{x} : x \in \mathcal{V}\}$ is a linear manifold in $(\mathcal{V}^*)^\#$. If $0 \neq \rho \in \mathcal{V}^*$, then obviously there exists $x \in \mathcal{V}$ such that $\rho(x) \neq 0$. In other words, $\widehat{\mathcal{V}}$ is a separating family of linear functionals on $\mathcal{V}^*$. Hence $(\mathcal{V}^*, \widehat{\mathcal{V}})$ is a dual pair.

*By convention, the weak topology on $\mathcal{V}^*$ induced by the family $\widehat{\mathcal{V}}$ is usually denoted by $\sigma(\mathcal{V}^*, \mathcal{V})$ (as opposed to $\sigma(\mathcal{V}^*, \widehat{\mathcal{V}})$), and is referred to as the **weak*-topology** on $\mathcal{V}^*$.*

7.7. Remark. It follows that a base for the weak*-topology on $\mathcal{V}^*$ is given by

$$\mathcal{B} = \{N(\varphi, F, \varepsilon) : \varphi \in \mathcal{V}^*, \varepsilon > 0, F \subseteq \mathcal{V} \text{ finite}\},$$

where

$$N(\varphi, F, \varepsilon) = \{\rho \in \mathcal{V}^* : |\widehat{x}(\rho - \varphi)| = |\rho(x) - \varphi(x)| < \varepsilon, x \in F\}.$$

Moreover, a net $(\rho_\lambda)_\lambda$ in $\mathcal{V}^*$ converges in the weak*-topology to a functional $\rho \in \mathcal{V}^*$ if and only if $\lim_\lambda \rho_\lambda(x) = \rho(x)$ for all $x \in \mathcal{V}$. In other words, convergence in the weak*-topology on $\mathcal{V}^*$ is convergence at every point of $\mathcal{V}$.

By Theorem 7.4, a functional φ is weak*-continuous on $\mathcal{V}^*$ if and only if $\varphi = \widehat{x}$ for some $x \in \mathcal{V}$.

7.8. Proposition. *Let $(\mathcal{V}, \mathcal{T}_\mathcal{V})$ and $(\mathcal{W}, \mathcal{T}_\mathcal{W})$ be locally convex spaces, and suppose that $T : (\mathcal{V}, \mathcal{T}_\mathcal{V}) \to (\mathcal{W}, \mathcal{T}_\mathcal{W})$ is a continuous linear operator. Then T is continuous as a linear map between $\mathcal{V}$ and $\mathcal{W}$ when they are equipped with their respective weak topologies.*

Proof. Suppose that $(x_\lambda)_\lambda$ is a net in $\mathcal{V}$ which converges weakly to $x \in \mathcal{V}$. We must show that the net $(Tx_\lambda)_\lambda$ converges weakly to Tx in $\mathcal{W}$. Now, if $\rho \in \mathcal{W}^*$, then $\rho \circ T$ is continuous with respect to the $\mathcal{T}_\mathcal{V}$ topology on $\mathcal{V}$, and hence $\rho \circ T \in \mathcal{V}^*$. But the weakly continuous linear maps on $\mathcal{V}$ coincide with $\mathcal{V}^*$, and therefore $\rho \circ T$ is weakly continuous on $\mathcal{V}$, i.e. $\lim_\lambda \rho \circ T(x_\lambda) = \rho \circ T(x)$, as was to be shown.

□

When $\mathcal{V}$ is a LCS and $C \subseteq \mathcal{V}$ is convex, we get a particularly nice result concerning the weak topology.

7.9. Theorem. *Let C be a convex set in a LCS $(\mathcal{V}, \mathcal{T})$. Then the closure of C in $(\mathcal{V}, \mathcal{T})$ coincides with its weak closure in $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$.*

Proof. First observe that we can always view $(\mathcal{V}, \mathcal{T})$ and $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ as locally convex spaces over $\mathbb{R}$. Since C is assumed to be convex already, Corollary 6.44 implies that the closure of C in $(\mathcal{V}, \mathcal{T})$ (resp. in $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$) is the intersection of the $\mathcal{T}$-closed (resp. $\sigma(\mathcal{V}, \mathcal{V}^*)$-closed) half spaces which contain C.

But a closed half space in a LCS corresponds to (a constant and) a continuous linear functional on that space. Since $(\mathcal{V}, \mathcal{T})$ and $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}^*))$ share the same dual space, namely $\mathcal{V}^*$, it follows that they also share the same closed half-spaces, and hence the closure of C in these two topologies must coincide.

□

7.10. Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space. Recall from Proposition 2.19 that the canonical embedding

$$\begin{array}{rccc} \mathfrak{J}: & \mathfrak{X} & \to & \mathfrak{X}^{**} \\ & x & \mapsto & \widehat{x}, \end{array}$$

where $\widehat{x}(x^*) := x^*(x)$ for all $x^* \in \mathfrak{X}^*$, is a contractive map. In Corollary 6.31 we saw that – as a consequence of the Hahn-Banach Theorem – $\mathfrak{J}$ is in fact an isometry.

By Theorem 7.4 and Remark 7.5, $\mathfrak{J}(\mathfrak{X})$ corresponds exactly to the weak*-continuous linear functionals on $\mathfrak{X}^*$.

7.11. Proposition. *Let $\mathfrak{X}$ be a finite-dimensional Banach space. Then the norm, weak and weak*-topologies on $\mathfrak{X}$ all coincide.*

Proof. First we must decide what we mean by the weak*-topology on $\mathfrak{X}$. Observe that if $\dim \mathfrak{X} = n < \infty$, then $\dim \mathfrak{X}^* = n$ as well, and thus $\dim \mathfrak{X}^{**} = n = \dim \mathfrak{X}$. Since $\mathfrak{J} : \mathfrak{X} \to \mathfrak{X}^{**}$ is a linear isometry, it must be a bijection in this case and therefore we can identify $\mathfrak{X}$ with $\mathfrak{X}^{**} = (\mathfrak{X}^*)^*$. In this sense $\mathfrak{X} \simeq \mathfrak{J}(\mathfrak{X})$ comes equipped with a weak*-topology induced by $\mathfrak{X}^*$, namely the $\sigma(\mathfrak{J}(\mathfrak{X}), \mathfrak{X}^*)$-topology. But since we are identifying $\mathfrak{X}$ with $\mathfrak{J}(\mathfrak{X}) = \mathfrak{X}^{**}$, this is really just the $\sigma(\mathfrak{X}, \mathfrak{X}^*)$-topology, namely the weak topology on $\mathfrak{X}$.

Since the weak and the norm topologies on $\mathfrak{X}$ are both TVS topologies, and since finite-dimensional vector spaces admit a unique TVS topology, we see that all three topologies cited above must coincide.

□

We now wish to examine some of the properties of the weak and weak*-topologies in the context of normed linear spaces. We shall first require a result from Real Analysis, which we shall then adapt to the setting of normed linear spaces.

7.12. Theorem. ***The Uniform Boundedness Principle***

Let (X, d) be a complete metric space and let $H \subseteq \mathcal{C}(X, \mathbb{K})$ be a non-empty family of continuous functions on X such that for each $x \in X$,

$$M_x := \sup_{h \in H} |h(x)| < \infty.$$

Then there exists an open set $G \subseteq X$ and a constant $M > 0$ so that

$$|h(x)| \le M \text{ for all } h \in H, x \in G.$$

Proof. For each $m \ge 1$, let $E_{m,h} = \{x \in X : |h(x)| \le m\}$, and let $E_m = \cap_{h \in H} E_{m,h}$. Since each $E_{m,h}$ is closed (as $h \in H$ implies that h is continuous), so is E_m. Also, for any $x \in X$, there exists $m > M_x$, and so $x \in E_m$. Thus

$$X = \cup_{m=1}^{\infty} E_m.$$

Since X is complete, the Baire Category Theorem implies the existence of $k \geq 1$ so that the interior $\text{int}\,(E_k)$ of E_k is non-empty. This clearly leads to the desired conclusion.

□

7.13. Corollary. ***The Uniform Boundedness Principle - Banach space version***

Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ *and* $(\mathcal{Y}, \|\cdot\|_{\mathfrak{Y}})$ *be Banach spaces and let* $\mathfrak{A} \subseteq \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ *denote a family of continuous linear operators from* $\mathfrak{X}$ *to* $\mathfrak{Y}$*. Suppose that for each* $x \in \mathfrak{X}$*, we have*

$$M_x := \sup\{\|Tx\| : T \in \mathfrak{A}\} < \infty.$$

Then

$$\sup\{\|T\| : T \in \mathfrak{A}\} < \infty.$$

Proof. For each $T \in \mathfrak{A}$, let $p_T : \mathfrak{X} \to \mathbb{R}$ be the continuous seminorm given by $p_T(x) = \|Tx\|$. Since $\mathfrak{X}$ is complete, the metric space version of the Uniform Boundedness Principle (Theorem 7.12) implies that there exists an open set $\varnothing \neq G \subseteq \mathfrak{X}$ and a constant $M > 0$ so that

$$\|Tx\| \leq M \quad \text{ for all } T \in \mathfrak{A}, x \in G.$$

Now $\varnothing \neq G$ open implies that there exists $z \in G$ and $\delta > 0$ so that $V_\delta(z) = \{x \in \mathfrak{X} : \|x - z\| < \delta\} \subseteq G$. Consider $y \in V_\delta(0)$ and $T \in \mathfrak{A}$.

Then

$$\begin{aligned} \|Ty\| &\leq \|T(y+z)\| + \|-Tz\| \\ &\leq M + \|Tz\| \\ &\leq 2M, \end{aligned}$$

as z and $y + z \in G$. It follows that if $x \in \mathfrak{X}, \|x\| \leq 1$, then

$$\frac{\delta}{2}\|Tx\| = \|T(\frac{\delta}{2}x)\| \leq M + \|Tz\| \leq 2M,$$

and hence that

$$\|Tx\| \leq \frac{2}{\delta}(2M), \quad T \in \mathfrak{A}.$$

That is,

$$\sup\{\|T\| : T \in \mathfrak{A}\} \leq \frac{4M}{\delta}.$$

□

7.14. Corollary. *Let $\mathfrak{X}$ be a Banach space and $\mathcal{S} \subseteq \mathfrak{X}$. Then $\mathcal{S}$ is bounded if and only if for all $x^* \in \mathfrak{X}^*$,*

$$\sup\{|x^*(s)| : s \in \mathcal{S}\} < \infty.$$

Proof. Suppose that $\mathcal{S}$ is bounded by $M > 0$. If $x^* \in \mathfrak{X}^*$, then $|x^*(s)| \le \|x^*\| \, \|s\| \le M\|x^*\| < \infty$.

Conversely, if $\sup\{|x^*(s)| : s \in \mathcal{S}\} < \infty$, then $\sup\{|\widehat{s}(x^*)| : s \in \mathcal{S}\} < \infty$ for all $x^* \in \mathfrak{X}^*$. By the Uniform Boundedness Principle,

$$\sup\{\|\widehat{s}\| : s \in \mathcal{S}\} = \sup\{\|s\| : s \in \mathcal{S}\} < \infty.$$

□

7.15. Corollary. *Let $\mathfrak{X}$ be a Banach space and $\mathfrak{S} \subseteq \mathfrak{X}^*$. Then $\mathfrak{S}$ is bounded if and only if for all $x \in \mathfrak{X}$,*

$$\sup\{|s^*(x)| : s^* \in \mathfrak{S}\} < \infty.$$

Proof. This is an immediate consequence of the Uniform Boundedness Principle, Theorem 7.13.

□

In general, we do not expect topologies to be determined by sequential convergence. If we do have convergence of a sequence, therefore, something strong is implied.

7.16. Theorem. ***The Banach-Steinhaus Theorem***

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $\{T_n\}_{n=1}^\infty \subseteq \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a sequence which satisfies the property that for each $x \in \mathfrak{X}$, there exists $y_x \in \mathfrak{Y}$ so that

$$\lim_{n\to\infty} T_n x = y_x.$$

Then

(a) $\sup_n \|T_n\| < \infty$,
(b) *the map $T : \mathfrak{X} \to \mathfrak{Y}$ defined by $Tx = y_x$ is a bounded linear map, and*
(c) $\|T\| \le \liminf_n \|T_n\|$.

Proof. For each $x \in \mathfrak{X}$, we have that $\{T_n x\}_{n=1}^\infty$ converges to some y_x, and therefore it is bounded. That is, $\sup_{n\ge1} \|T_n x\| < \infty$ for each $x \in \mathfrak{X}$. By the Uniform Boundedness Principle, $M := \sup_{n\ge1} \|T_n\| < \infty$, which establishes part (a).

Let $Tx := \lim_n T_n x$, for each $x \in \mathfrak{X}$. Linearity of T is readily checked. Also,

$$\|Tx\| = \lim_n \|T_n x\| \le \liminf_{n\ge1} \|T_n\| \, \|x\| = (\liminf_n \|T_n\|) \, \|x\| \le M\|x\|, \quad x \in \mathfrak{X}.$$

Hence $\|T\| \le \liminf_n \|T_n\| \le M < \infty$, which completes parts (b) and (c).

□

7.17. Corollary. *Let $\mathfrak{X}$ be a Banach space.*

(a) *If $(x_n)_{n=1}^\infty$ is a **sequence** which converges weakly to $x \in \mathfrak{X}$, then*
   (i) $\sup \|x_n\| < \infty$*; and*
   (ii) $\|x\| \le \liminf \|x_n\|$.

(b) *If $(y_n^*)_{n=1}^\infty$ is a **sequence** which converges in the weak*-topology to $y^* \in \mathfrak{X}^*$, then*
   (iii) $\sup \|y_n^*\| < \infty$*; and*
   (iv) $\|y^*\| \le \liminf \|y_n^*\|$.

Proof.

(a) Consider the natural isometric embedding $\Gamma : \mathfrak{X} \to \mathfrak{X}^{**}$ sending y to $\widehat{y}$, where $\widehat{y}(x^*) = x^*(y)$ for all $x^* \in \mathfrak{X}^*$.

   Letting $\mathfrak{Y} = \mathbb{K}$, we may apply the Banach-Steinhaus Theorem above to the sequence $(\widehat{x_n})_{n=1}^\infty$, noting that convergence of $(x_n)_{n=1}^\infty$ in the weak topology is simply the statement that $\lim_{n\to\infty} \widehat{x_n}(x^*) = \widehat{x}(x^*)$ exists for all $x^* \in \mathfrak{X}^*$. Since Γ is an isometry, the result readily follows.

(b) Again, this is an immediate application of the Banach-Steinhaus Theorem, replacing $\mathfrak{Y}$ in that Theorem with $\mathbb{K}$.

□

It is worth pointing out that by Theorem 7.9, if $(x_n)_{n=1}^\infty$ converges weakly to x, then $x \in \overline{co}^{\|\cdot\|}(\{x_n\}_{n=1}^\infty)$, since the latter is a convex set in $\mathfrak{X}$, closed in the norm (and hence in the weak) topology.

Before considering our next example, let us first recall a result from Measure Theory, alternately referred to as the Riesz Representation Theorem or the Riesz-Markov Theorem.

7.18. Theorem. *Let X be a locally compact, Hausdorff topological space, and denote by $\mathcal{M}(X)$ the space of $\mathbb{K}$-valued, finite, regular, Borel measures on X, equipped with the total variation norm: $\|\mu\| = |\mu|(X)$.*

If $\mu \in \mathcal{M}(X)$, then $\beta_\mu : \mathcal{C}_0(X, \mathbb{K}) \to \mathbb{K}$ given by $\beta_\mu(f) = \int_X f d\mu$ is an element of $\mathcal{C}_0(X, \mathbb{K})^*$, and the map $\Theta : \mathcal{M}(X) \to \mathcal{C}_0(X, \mathbb{K})^*$ is an isometric linear isomorphism.

For example, if $X = \mathbb{N}$ with counting measure, then $\mathcal{C}_0(X, \mathbb{K}) = c_0(\mathbb{N}, \mathbb{K})$ and $\mathcal{M}(X) = \ell^1(\mathbb{N}, \mathbb{K})$.

When $X = [0, 1]$, we can in turn identify $\mathcal{M}([0, 1]) = \mathcal{C}([0, 1], \mathbb{K})^*$ with the space $BV[0, 1]$ of left-continuous functions of bounded variation on $[0, 1]$.

7.19. Proposition. *Let X be a compact, Hausdorff space. Then a sequence $\{f_n\}_{n=1}^\infty$ in $\mathcal{C}(X)$ converges weakly to $f \in \mathcal{C}(X)$ if and only if*

(i) $\sup_n \|f_n\| < \infty$*; and*
(ii) *For each $x \in X, (f_n(x))_{n=1}^\infty$ converges to $f(x)$.*

Proof. Suppose first that $\{f_n\}_{n=1}^\infty$ converges weakly to f. By Corollary 7.17, $\sup_n \|f_n\| < \infty$. Let $\delta_x : \mathcal{C}(X) \to \mathbb{K}$ be the evaluation functional, $\delta_x(f) = f(x)$ for all $x \in X$. Then δ_x is linear and for $f \in \mathcal{C}(X)$, $|\delta_x(f)| = |f(x)| \leq \|f\|$, so that $\|\delta_x\| \leq 1$ and $\delta_x \in \mathcal{C}(X)^*$. (In fact, δ_x corresponds to the point mass measure at x.)

Thus $\lim_{n\to\infty} \delta_x(f_n) = \delta_x(f)$, i.e. $\lim_{n\to\infty} f_n(x) = f(x)$.

Conversely, suppose that (i) and (ii) hold. If $\rho \in \mathcal{C}(X)^*$, then by the Riesz Representation Theorem above, there exists $\mu \in \mathcal{M}(X)$ with $\|\mu\| = \|\rho\|$ so that

$$\rho(f) = \int_X f d\mu,$$

for all $f \in \mathcal{C}(X)$. By the Lebesgue Dominated Convergence Theorem,

$$\rho(f) = \int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu = \lim_{n\to\infty} \rho(f_n).$$

In other words, $(f_n)_n$ converges weakly to f.

□

7.20. Theorem. ***Tychonoff's Theorem***

Suppose that $(X_\lambda, \mathcal{T}_\lambda)$ is a non-empty collection of compact, topological spaces. Then $X = \prod_\lambda X_\lambda$ is compact in the product topology.

Proof. Recall from Real Analysis that it suffices to prove that if $\mathcal{F}$ is a collection of closed subsets of X with the Finite Intersection Property (FIP), then $\cap\{F : F \in \mathcal{F}\} \neq \varnothing$. To that end, let $\mathcal{F}$ be a collection of closed subsets of X with the FIP.

Let $\mathfrak{J} = \{\mathcal{J} \subseteq \mathcal{P}(X) : \mathcal{F} \subseteq \mathcal{J}$ and $\mathcal{J}$ has the FIP$\}$, partially ordered by inclusion, so that $\mathcal{J}_1 \leq \mathcal{J}_2$ if $\mathcal{J}_1 \subseteq \mathcal{J}_2$. Since $\mathcal{F} \in \mathfrak{J}$, $\mathfrak{J} \neq \varnothing$. Suppose that $\mathcal{C} = \{\mathcal{J}_\lambda\}_\lambda$ is a chain if $\mathfrak{J}$. Clearly $\mathcal{F} \subseteq \mathcal{K} := \cup_\lambda \mathcal{J}_\lambda$, and if $H_1, H_2, ..., H_m \in \mathcal{K}$, then the fact that $\mathcal{C}$ is totally ordered implies that there exists λ_0 so that $H_1, H_2, ..., H_m \in \mathcal{J}_{\lambda_0}$. Since $\mathcal{J}_{\lambda_0}$ has the FIP, $\cap_{i=1}^m H_i \neq \varnothing$. Thus $\mathcal{K}$ has the FIP, and so $\mathcal{K} \in \mathfrak{J}$ is an upper bound for $\mathcal{C}$. By Zorn's Lemma, $\mathfrak{J}$ admits a maximal element, say $\mathcal{M}$.

We make two observations: first, if we set $\mathcal{M}_0 = \{\cap_{k=1}^r M_k : M_k \in \mathcal{M}, 1 \leq k \leq r, r \geq 1\}$, then the elements of $\mathcal{M}_0$ are finite intersections of elements of $\mathcal{M}$. It follows that $\mathcal{M}_0$ has the FIP. Moreover, $\mathcal{F} \subseteq \mathcal{M}_0$. Since $\mathcal{M} \leq \mathcal{M}_0$, the maximality of $\mathcal{M}$ implies that $\mathcal{M} = \mathcal{M}_0$. In other words, finite intersections of elements of $\mathcal{M}$ lie in $\mathcal{M}$.

Second, if $R \subseteq X$ and $R \cap M \neq \varnothing$ for all $M \in \mathcal{M}$, then $\mathcal{F} \subseteq \mathcal{M} \subseteq \mathcal{M} \cup \{R\}$ and $\mathcal{M} \cup \{R\}$ has the FIP. Again, the maximality of $\mathcal{M}$ implies that $R \in \mathcal{M}$.

Our goal now is to prove that $\cap\{\overline{M} : M \in \mathcal{M}\} \neq \varnothing$. Since $\cap\{F : F \in \mathcal{F}\} \supseteq \cap\{\overline{M} : M \in \mathcal{M}\}$, this will suffice to prove the Theorem.

For each λ, let $\pi_\lambda : X \to X_\lambda$ denote the canonical projection map. Then $\varnothing \neq \mathcal{M}_\lambda := \{\pi_\lambda(M) : M \in \mathcal{M}\}$ is a family of subsets of X_λ with the FIP. Since X_λ is compact, $\cap\{\overline{\pi_\lambda(M)}^{X_\lambda} : M \in \mathcal{M}\} \neq \varnothing$. Choose $x_\lambda \in \cap\{\overline{\pi_\lambda(M)}^{X_\lambda} : M \in \mathcal{M}\}$, and let $x = (x_\lambda)_\lambda$. We want to show that $x \in \cap\{\overline{M} : M \in \mathcal{M}\}$.

To do this, we must show that $G \in \mathcal{U}_x^X$ implies $G \cap M \neq \varnothing$ for all $M \in \mathcal{M}$. Clearly it suffices to do this when G is a basic nbhd of x, say

$$G = \cap_{j=1}^n \pi_{\lambda_j}^{-1}(U_j),$$

where $U_j \subseteq X_j$ is open, $1 \leq j \leq n$. Now for any $\lambda_0 \in \Lambda$ and $x_{\lambda_0} \in U_{\lambda_0} \in \mathcal{T}_{\lambda_0}$, $x_{\lambda_0} \in \overline{\pi_{\lambda_0}(M)}^{X_{\lambda_0}}$ for all $M \in \mathcal{M}$ implies that $U_{\lambda_0} \cap \pi_{\lambda_0}(M) \neq \varnothing$ for all $M \in \mathcal{M}$.

But then $\pi_{\lambda_0}^{-1}(U_{\lambda_0}) \cap M \neq \varnothing$ whenever $x_{\lambda_0} \in U_{\lambda_0} \subseteq X_{\lambda_0}$ is open. By maximality of $\mathcal{M}$ and the second observation above, $\pi_{\lambda_0}^{-1}(U_{\lambda_0}) \in \mathcal{M}$. Since $\mathcal{M}$ is closed under finite intersections by the first observation,

$$G = \cap_{j=1}^n \pi_{\lambda_j}^{-1}(U_{\lambda_j}) \in \mathcal{M} \text{ whenever } G \text{ is a basic nbhd of } x.$$

Thus $x \in \cap\{\overline{M} : M \in \mathcal{M}\}$, and we are done.

□

7.21. Theorem. ***The Banach-Alaoglu Theorem***

Let $\mathfrak{X}$ be a Banach space. Then the closed unit ball $\mathfrak{X}_1^* := \{x^* \in \mathfrak{X}^* : \|x^*\| \leq 1\}$ of $\mathfrak{X}^*$ is weak-compact.*

Proof. For each $x \in \mathfrak{X}$, $x^* \in \mathfrak{X}_1^*$, we have

$$|\widehat{x}(x^*)| = |x^*(x)| \leq \|x^*\| \, \|x\| \leq \|x\|.$$

Thus $\widehat{x}(\mathfrak{X}_1^*) \subseteq D_x := \{z \in \mathbb{K} : |z| \leq \|x\|\}$. Now each such D_x is compact, and so by Tychonoff's Theorem above,

$$D := \prod_{x \in \mathfrak{X}} D_x$$

is also compact in the product topology. To complete the proof, we shall show that $\mathfrak{X}_1^*$ is homeomorphic to a closed, and therefore compact, subset of D.

Define

$$\begin{array}{rccc} \Phi : & \mathfrak{X}_1^* & \to & D \\ & f & \mapsto & (\widehat{x}(f))_{x \in X} = (f(x))_{x \in X}. \end{array}$$

Clear Φ is injective. Now a net $(f_\lambda)_{\lambda \in \Lambda}$ converges weak* to f if and only if

$$\lim_\lambda f_\lambda(x) = \lim_\lambda \widehat{x}(f_\lambda) = \widehat{x}(f) = f(x) \quad \text{for all } x \in \mathfrak{X},$$

that is, if and only if $\lim_\lambda \Phi(f_\lambda) = \Phi(f)$.

Thus $\mathfrak{X}_1^*$ is homeomorphic to $\Phi(\mathfrak{X}_1^*)$. There remains to show that $\Phi(\mathfrak{X}_1^*)$ is closed in D.

Suppose that $(f_\lambda)_\lambda$ is a net in $\mathfrak{X}_1^*$, and that $(\Phi(f_\lambda))_\lambda$ converges to $d = (d_x)_{x \in \mathfrak{X}} \in D$. Then

$$\lim_\lambda f_\lambda(x) = d_x \quad \text{for all } x \in \mathfrak{X}.$$

Define $f(x) := d_x$, $x \in \mathfrak{X}$. Then f is linear since each f_λ is, and

$$|f(x)| = |d_x| \leq \|x\| \quad \text{for all } x \in \mathfrak{X},$$

so that $f \in \mathfrak{X}_1^*$. Clearly $\Phi(f) = \lim_\lambda(\Phi(f_\lambda))$, so that $\operatorname{ran}\Phi$ is closed, and we are done.

□

7.22. Corollary. *Every Banach space $\mathfrak{X}$ is isometrically isomorphic to a subspace of $(\mathcal{C}(L,\mathbb{K}), \|\cdot\|_\infty)$ for some compact, Hausdorff space L.*
Proof. Let $L := \mathfrak{X}_1^*$. Then L is weak*-compact, by the Banach-Alaoglu Theorem, and is Hausdorff since $\mathfrak{X}$ separates the points of $\mathfrak{X}_1^*$. Define

$$\begin{array}{rccc} \Delta : & \mathfrak{X} & \to & \mathcal{C}(L,\mathbb{K}) \\ & x & \mapsto & \widehat{x}|_L. \end{array}$$

Then Δ is easily seen to be linear, and $\|\widehat{x}|_L\| \le \|\widehat{x}\| = \|x\|$.

By the Hahn-Banach Theorem [Corollary 6.30], there exists $x^* \in \mathfrak{X}_1^*$ such that $|x^*(x)| = \|x\|$, and so $\|\widehat{x}|_L\| \ge |\widehat{x}(x^*)| = |x^*(x)| = \|x\|$; that is, Δ is an isometry.

□

7.23. Corollary. *Let $\mathfrak{X}$ be a Banach space and suppose that $\mathcal{A} \subseteq \mathfrak{X}^*$ is weak*-closed and bounded. Then $\mathcal{A}$ is weak*-compact.*

7.24. Theorem. ***Goldstine's Theorem***
Let $\mathfrak{X}$ be a Banach space and $\mathfrak{J} : \mathfrak{X} \to \mathfrak{X}^{**}$ denote the canonical embedding. Then $\mathfrak{J}(\mathfrak{X}_1)$ is weak-dense in $\mathfrak{X}_1^{**}$. Thus $\mathfrak{J}(\mathfrak{X})$ is weak*-dense in $\mathfrak{X}^{**}$.*
Proof. Clearly $\mathfrak{J}(\mathfrak{X}_1) = \widehat{\mathfrak{X}_1}$ is convex, since $\mathfrak{X}_1$ is. Observe that the closure of $\mathfrak{J}(\mathfrak{X}_1)$ in the weak*-topology, namely $\overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}$, is weak*-closed and convex. Being weak*-closed in the weak*-compact set $\mathfrak{X}_1^{**}$, it is also weak*-compact. Suppose that $\varphi \in \mathfrak{X}_1^{**}$ and $\varphi \notin \overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}$. Then, by the Hahn-Banach Theorem 6.43 (HB05), we can find a weak*-continuous linear functional $\widehat{x^*} \in \mathfrak{J}(\mathfrak{X}^*) \subseteq \mathfrak{X}^{***}$ so that

$$\begin{aligned} \operatorname{Re} \widehat{x^*}(\varphi) &= b \\ &> a := \sup\{\operatorname{Re} \widehat{x^*}(\xi) : \xi \in \overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}\} \\ &= \sup\{|\widehat{x^*}(\xi)| : \xi \in \overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}\}. \end{aligned}$$

(The last equality follows from the fact that $\mathfrak{X}_1$ and hence $\mathfrak{J}(\mathfrak{X}_1)$ and $\overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}$ are balanced.)

But

$$\begin{aligned} \sup\{|\widehat{x^*}(\xi)| : \xi \in \overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}\} &= \sup\{|\xi(x^*)| : \xi \in \overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*}\} \\ &\ge \sup\{|\widehat{x}(x^*)| : x \in \mathfrak{X}_1\} \\ &= \sup\{|x^*(x)| : x \in \mathfrak{X}_1\} \\ &= \|x^*\|, \end{aligned}$$

while

$$|\operatorname{Re} \widehat{x^*}(\varphi)| \le |\widehat{x^*}(\varphi)| = |\varphi(x^*)\| \le \|\varphi\| \, \|x^*\| \le \|x^*\|.$$

This contradicts our choice of $\widehat{x^*}$, and thus $\overline{\mathfrak{J}(\mathfrak{X}_1)}^{w^*} = \mathfrak{X}_1^{**}$, as claimed.

Since $\mathfrak{X}^{**} = \cup_{n\ge 1} \mathfrak{X}_n^{**}$, and since each $\mathfrak{J}(\mathfrak{X}_n)$ is weak*-dense in $\mathfrak{X}_n^{**}$ by a routine modification of the above proof, $\mathfrak{J}(\mathfrak{X})$ is weak*-dense in $\mathfrak{X}^{**}$.

□

7.25. Example. By identifying $c_0(\mathbb{K})^*$ with $\ell^1(\mathbb{K})$ and $\ell^1(\mathbb{K})^*$ with $\ell^\infty(\mathbb{K})$, we see that the unit ball $(c_0(\mathbb{K}))_1$ of $c_0(\mathbb{K})$ is weak*-dense in the closed unit ball of $\ell^\infty(\mathbb{K})$, and thus $c_0(\mathbb{K})$ is weak*-dense in $\ell^\infty(\mathbb{K})$.

Of course, $c_{00}(\mathbb{K})$ is norm dense in $c_0(\mathbb{K})$, and so $c_{00}(\mathbb{K})$ is also weak*-dense in $\ell^\infty(\mathbb{K})$.

CULTURE: Although we shall not have time to prove this, the non-commutative analogue of the above statement is that the set of finite rank operators $\mathcal{F}(\mathcal{H})$ on an infinite-dimensional Hilbert space is weak*-dense in $\mathcal{B}(\mathcal{H})$.

Let us now establish a relation between compactness and reflexivity of a Banach space.

7.26. Proposition. *Let $\mathfrak{X}$ be a Banach space. The following are equivalent.*

(a) *$\mathfrak{X}$ is reflexive.*
(b) *$\mathfrak{X}_1$ is weakly compact.*

Proof.

(a) implies (b): First suppose that $\mathfrak{X}$ is reflexive. Then $\widehat{\mathfrak{X}} = \mathfrak{X}^{**}$ and $\widehat{\mathfrak{X}_1} = \mathfrak{X}_1^{**}$ is weak*-compact the Banach-Alaoglu Theorem 7.21. But then the weak*-topology on $\widehat{\mathfrak{X}_1}$ is just the weak topology on $\mathfrak{X}_1$, so $\mathfrak{X}_1$ is weakly compact.

(b) implies (a): Next suppose that $\mathfrak{X}_1$ is weakly compact. Then $\widehat{\mathfrak{X}_1}$ is weak*-compact, and since the weak*-topology is Hausdorff, $\widehat{\mathfrak{X}_1}$ is weak*-closed. But by Goldstine's Theorem, $\widehat{\mathfrak{X}_1}$ is weak*-dense in $\mathfrak{X}_1^{**}$. Thus $\widehat{\mathfrak{X}_1} = \Gamma(\mathfrak{X}_1) = \overline{\Gamma(\mathfrak{X}_1)}^{w^*} = \mathfrak{X}_1^{**}$. This in turn implies that $\widehat{\mathfrak{X}} = \mathfrak{X}^{**}$, or in other words, that $\mathfrak{X}$ is reflexive.

□

Although in general, weak topologies are not metrizable, sometimes their restrictions to bounded sets can be:

7.27. Theorem. *Let $\mathfrak{X}$ be a Banach space. Then $\mathfrak{X}_1^*$ is weak*-metrizable if and only if $\mathfrak{X}$ is separable.*

Proof. First assume that $\mathfrak{X}$ is separable, and let $\{x_n\}_{n=1}^\infty$ be a dense subset of $\mathfrak{X}$. Define a metric d on $\mathfrak{X}_1^*$ via

$$d(x^*, y^*) = \sum_{n=1}^{\infty} \frac{|x^*(x_n) - y^*(x_n)|}{2^n \, \|x_n\|}.$$

Then a net $(x_\lambda^*)_\lambda$ in $\mathfrak{X}_1^*$ converges in the metric topology to $x^* \in \mathfrak{X}_1^*$ if and only if $(x_\lambda^*(x_n))_\lambda$ converges to $x^*(x_n)$ for all $n \geq 1$ (exercise). If $x \in \mathfrak{X}$ and $\varepsilon > 0$, we can choose $n \geq 1$ so that $\|x_n - x\| < \varepsilon/3$. Choose $\lambda_0 \in \Lambda$ so that $\lambda \geq \lambda_0$ implies that $|x_\lambda^*(x_n) - x^*(x_n)| < \varepsilon/3$. Then $\lambda \geq \lambda_0$ implies that

$$\begin{aligned} |x_\lambda^*(x) - x^*(x)| &\leq |x_\lambda^*(x) - x_\lambda^*(x_n)| + |x_\lambda^*(x_n) - x^*(x_n)| + |x^*(x_n) - x^*(x)| \\ &\leq \|x_\lambda^*\| \, \|x - x_n\| + \varepsilon/3 + \|x^*\| \, \|x_n - x\| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

Thus $x^*_\lambda(x_n)$ converges to $x^*(x_n)$ for all $n \ge 1$ if and only if $(x^*_\lambda)_\lambda$ converges in the weak*-topology to x^*. Hence the weak*-topology on $\mathfrak{X}^*_1$ is metrizable.

Next, assume that $\mathfrak{X}^*_1$ is weak*-metrizable. Then we can find a countable sequence $\{G^*_n\}_{n=1}^\infty$ of weak*-open nbhds of $0 \in \mathfrak{X}^*_1$ so that $\cap_{n=1}^\infty G^*_n = \{0\}$. There is no harm in assuming that each G^*_n is a basic weak*-open nbhd, so for each $n \ge 1$ there exists $\varepsilon_n > 0$ and a finite set $F_n \subseteq \mathfrak{X}$ so that

$$G^*_n = \{x^* \in \mathfrak{X}^*_1 : |x^*(x)| < \varepsilon_n, x \in F_n\}.$$

Let $F = \cup_{n=1}^\infty F_n$. If $x^* \in \mathfrak{X}^*_1$, $x^*(F) = 0$, then $x^* \in G^*_n$ for all $n \ge 1$, and therefore $x^* = 0$. That is, if $\mathfrak{Y} = \overline{\text{span}}^{\|\cdot\|} F$, then $\mathfrak{Y}$ is separable and $x^* \in \mathfrak{X}^*_1$, $x^*|_{\mathfrak{Y}} = 0$ implies that $x^* = 0$. By the Hahn-Banach Theorem [Corollary 6.29], $\mathfrak{Y} = \mathfrak{X}$.

□

7.28. Corollary. *Let $\mathfrak{X}$ be a separable Banach space. Then $\mathfrak{X}^*_1$ is separable in the weak*-topology.*

Proof. By the Banach-Alaoglu Theorem, $\mathfrak{X}^*_1$ is weak*-compact. By Theorem 7.27 above, $\mathfrak{X}^*_1$ is weak*-metrizable.

Since a compact metric space is always separable – see Proposition 11.10 – we see that $(\mathfrak{X}^*_1, \sigma(\mathfrak{X}^*, \mathfrak{X}))$ is separable.

□

In a similar vein, we have

7.29. Theorem. *Let $\mathfrak{X}$ be a Banach space. Then $\mathfrak{X}_1$ is weakly metrizable if and only if $\mathfrak{X}^*$ is separable.*

Proof. Assignment.

□

7.30. Definition. *Let $\mathfrak{X}$ be a Banach space and $\mathfrak{M} \subseteq \mathfrak{X}$, $\mathfrak{N} \subseteq \mathfrak{X}^*$. Then the **annihilator** of $\mathfrak{M}$ is the set*

$$\mathfrak{M}^\perp = \{x^* \in \mathfrak{X}^* : x^*(m) = 0 \textit{ for all } m \in \mathfrak{M}\},$$

*while the **pre-annihilator** of $\mathfrak{N}$ is the set*

$$^\perp\mathfrak{N} = \{x \in \mathfrak{X} : n^*(x) = 0 \textit{ for all } n^* \in \mathfrak{N}.\}$$

Observe that $\mathfrak{M}^\perp$ and $^\perp\mathfrak{N}$ are linear manifolds in their respective spaces. Moreover, both are norm-closed and hence Banach spaces in their own right.

7.31. Theorem. *Let $\mathfrak{X}$ be a Banach space, and let $\mathfrak{M} \subseteq \mathfrak{X}$ be a closed subspace. Let $q : \mathfrak{X} \to \mathfrak{X}/\mathfrak{M}$ denote the canonical quotient map. Then*

$$\begin{array}{rll} \Theta : & (\mathfrak{X}/\mathfrak{M})^* & \to \mathfrak{M}^\perp \\ & \xi & \mapsto \xi \circ q \end{array}$$

is an isometric isomorphism of Banach spaces.

Proof. Clearly Θ is linear. Let us show that Θ is injective.

If $\Theta(\xi_1) = \xi_1 \circ q = \xi_2 \circ q = \Theta(\xi_2)$, then

$$\xi_1(q(x)) = \xi_2(q(x)) \quad \text{for all } x \in \mathfrak{X},$$

and so $\xi_1 = \xi_2$.

Next we show that Θ is surjective.

Let $z^* \in \mathfrak{M}^\perp$ and define $\xi_{z^*} : \mathfrak{X}/\mathfrak{M} \to \mathbb{K}$ via $\xi_{z^*}(q(x)) = z^*(x)$. Since $\mathfrak{M} \subseteq \ker\ z^*$, the map is well-defined. Furthermore, if $x \in \mathfrak{X}$ and $\|q(x)\| < 1$, then there exists $m \in \mathfrak{M}$ so that $\|x + m\| < 1$, and

$$|\xi_{z^*}(q(x))| = |z^*(x)| = |z^*(x+m)| \leq \|z^*\|,$$

so that $\|\xi_{z^*}\| \leq \|z^*\| < \infty$. Hence $\xi_{z^*} \in (\mathfrak{X}/\mathfrak{M})^*$. Clearly $\Theta(\xi_{z^*}) = z^*$.

Thus Θ is bijective, and $\|\Theta(\xi)\| = \|\xi \circ q\| \leq \|\xi\|\,\|q\| \leq \|\xi\|$, so that $\|\Theta\| \leq 1$. Conversely, let $\varepsilon > 0$ and choose $q(x) \in \mathfrak{X}/\mathfrak{M}$ with $\|q(x)\| < 1$ so that $|\xi(q(x))| \geq \|\xi\| - \varepsilon$. Choose $m \in \mathfrak{M}$ so that $\|x + m\| < 1$. Then

$$\|\xi \circ q\| \geq |\xi \circ q(x+m)| = |\xi(q(x))| \geq \|\xi\| - \varepsilon,$$

so that $\|\Theta(\xi)\| = \|\xi \circ q\| \geq \|\xi\|$, implying that Θ is in fact isometric.

□

7.32. Theorem. *Let $\mathfrak{X}$ be a Banach space and $\mathfrak{M} \subseteq \mathfrak{X}$ be a closed linear subspace. Then the map*

$$\begin{array}{rccc} \Theta : & \mathfrak{X}^*/\mathfrak{M}^\perp & \to & \mathfrak{M}^* \\ & x^* + \mathfrak{M}^\perp & \mapsto & x^*|_{\mathfrak{M}} \end{array}$$

is an isometric isomorphism.

Proof. Note that $\mathfrak{M}^\perp$ closed implies that $\mathfrak{X}^*/\mathfrak{M}^\perp$ is a Banach space. We check that Θ is well-defined.

If $x^* + \mathfrak{M}^\perp = y^* + \mathfrak{M}^\perp$, then $x^* - y^* \in \mathfrak{M}^\perp$, so that $(x^* - y^*)|_{\mathfrak{M}} = 0$. That is, $\Theta(x^* + \mathfrak{M}^\perp) = \Theta(y^* + \mathfrak{M}^\perp)$. Working our way backwards through this argument proves that Θ is injective. That Θ is linear is easily verified.

Next suppose that $m^* \in \mathfrak{M}^*$. By the Hahn-Banach Theorem, we can find $x^* \in \mathfrak{X}^*$, $\|x^*\| = \|m^*\|$ so that $x^*|_{\mathfrak{M}} = m^*$. Then $\Theta(x^* + \mathfrak{M}^\perp) = x^*|_{\mathfrak{M}} = m^*$, so that Θ is onto. Thus Θ is bijective.

Suppose that $\|x^* + \mathfrak{M}^\perp\| < 1$. Then there exists $n^* \in \mathfrak{M}^\perp$ so that $\|x^* + n^*\| < 1$. Thus

$$\|\Theta(x^* + \mathfrak{M}^\perp)\| = \|x^*|_{\mathfrak{M}}\| = \|(x^* + n^*)|_{\mathfrak{M}}\| \leq \|x^* + n^*\| < 1.$$

It follows that $\|\Theta\| \leq 1$.

From above, given $m^* \in \mathfrak{M}^\perp$, there exists $x^* \in \mathfrak{X}^*$ with $\|x^*\| = \|m^*\|$ so that $\Theta^{-1}(m^*) = x^* + \mathfrak{M}^\perp$. Now

$$\|\Theta^{-1}(m^*)\| = \|x^* + \mathfrak{M}^\perp\| \le \|x^*\| = \|m^*\|,$$

so that Θ^{-1} is also contractive. But then Θ is isometric, and we are done.

□

7.33. It is a worthwhile exercise to think about the relationship between the annihilator $\mathfrak{M}^\perp$ of a subspace $\mathfrak{M}$ of a Hilbert space $\mathcal{H}$ and the orthogonal complement of $\mathfrak{M}$ in $\mathcal{H}$, for which we used the same notation.

In particular, one should interpret what Theorem 7.32 says in the Hilbert space setting, where $\mathfrak{M}^\perp$ refers to the orthogonal complement of $\mathfrak{M}$.

*

If you had a face like mine, you'd punch me right on the nose, and I'm just the fella to do it.

Stan Laurel

Appendix to Section 7.

The following result characterizes weak convergence of **sequences** in ℓ^p. We leave its proof as an exercise for the reader.

7.34. Proposition. *Suppose that* $1 < p < \infty$. *A sequence* $(\mathbf{x}_n)_{n=1}^\infty$ *in* $\ell^p(\mathbb{N})$ *(i.e. each* $\mathbf{x_n} = (x_{n1}, x_{n2}, x_{n3}, ...) \in \ell^p(\mathbb{N})$ *converges weakly to* $\mathbf{z} = (z_1, z_2, z_3, ...) \in \ell^p(\mathbb{N})$*) if and only if*

(i) $\sup_{n \geq 1} \|\mathbf{x}_n\| < \infty$, *and*
(ii) $\lim_{n \to \infty} x_{nk} = z_k$ *for all* $k \in \mathbb{N}$.

Exercises for Section 7.

Question 1.

Let $\mathfrak{X}$ be a Banach space and $C \subseteq \mathfrak{X}$ be convex. Prove the following.

(a) $\overline{C}^{\|\cdot\|} = \overline{C}^{\text{weak}}$.

(b) C is norm-closed if and only if C is weakly closed.

Question 2.

Let $\mathfrak{X}$ be a Banach space. Prove that $\mathfrak{X}_1$ is weakly metrizable if and only if $\mathfrak{X}^*$ is separable.

Question 3.

Give a direct proof (i.e. without appealing to Goldstine's Theorem) that c_0 is weak*-dense in ℓ^∞.

Question 4.

Let (X, τ) be a topological space. A function $f : X \to \mathbb{R}$ is said to be **lower semicontinuous** if for all $a \in \mathbb{R}$, $f^{-1}(a, \infty)$ is open in X.

(a) Show that every continuous function from X into $\mathbb{R}$ is lower semicontinuous.

(b) Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space. Prove that the function

$$\begin{array}{rccc} f: & (\mathfrak{X}, \sigma(\mathfrak{X}, \mathfrak{X}^*)) & \to & \mathbb{R} \\ & x & \mapsto & \|x\| \end{array}$$

is lower semicontinuous. That is, the norm on $\mathfrak{X}$ is lower semicontinuous for the weak topology on $\mathfrak{X}$.

(c) Let $(\mathfrak{X}, \|\cdot\|)$ be a Banach space. Prove that the function

$$\begin{array}{rccc} g: & (\mathfrak{X}^*, \sigma(\mathfrak{X}^*, \mathfrak{X})) & \to & \mathbb{R} \\ & x^* & \mapsto & \|x^*\| \end{array}$$

is lower semicontinuous. That is, the norm on $\mathfrak{X}^*$ is lower semicontinuous for the weak*- topology on $\mathfrak{X}$.

8. Extremal points

Somewhere on this globe, every ten seconds, there is a woman giving birth to a child. She must be found and stopped.

Sam Levenson

8.1. The main result of this section is the Krein-Milman Theorem, which asserts that a non-empty, compact, convex subset of a LCS has extreme points; so many, in fact, that we can generate the compact, convex set as the closed, convex hull of these extreme points.

Extreme points of convex sets appear in many different contexts in Functional Analysis. For example, it is an interesting exercise (so interesting that it may appear as an Assignment question) to calculate the extreme points of the closed unit ball $\mathcal{B}(\mathbb{C}^n)_1$ of the locally convex space $\mathcal{B}(\mathbb{C}^n)$, where $\mathbb{C}^n$ is endowed with the Euclidean norm $\|\cdot\|_2$ and $\mathcal{B}(\mathbb{C}^n)$ is given the operator norm.

Recall that a linear map $T \in \mathcal{B}(\mathbb{C}^n)$ is said to be **positive** and we write $T \geq 0$ if $\langle Tx, x\rangle \geq 0$ for all $x \in \mathbb{C}^n$. An equivalent formulation of this property says that T is positive if there exists an orthonormal basis for $\mathbb{C}^n$ with respect to which the matrix $[T]$ of T is diagonal, and all eigenvalues of T are non-negative real numbers. A linear functional $\varphi \in \mathcal{B}(\mathbb{C}^n)^*$ is said to be **positive** if $\varphi(T) \geq 0$ whenever $T \geq 0$. For example, if $\{e_1, e_2, ..., e_n\}$ is an orthonormal basis for $\mathbb{C}^n$, then the so-called **normalized trace functional**

$$\tau(T) := \frac{1}{n}\sum_{k=1}^{n}\langle Te_k, e_k\rangle$$

for $T \in \mathcal{B}(\mathbb{C}^n)$ can be shown to be a positive linear functional of norm one.

The **state space** $\mathcal{S}(\mathcal{B}(\mathbb{C}^n))$ of $\mathcal{B}(\mathbb{C}^n)$, consisting of all positive, norm-one linear functionals on $\mathcal{B}(\mathbb{C}^n)$ – called **states** – forms a non-empty, compact, convex subset of $\mathcal{B}(\mathbb{C}^n)^*$. The extreme points of the state space are called **pure states**. For example, if $x \in \mathbb{C}^n$ and $\|x\| = 1$, then the map

$$\begin{array}{cccc} \varphi_x : & \mathcal{B}(\mathbb{C}^n) & \to & \mathbb{C} \\ & T & \mapsto & \langle Tx, x\rangle \end{array}$$

defines a pure state. States on $\mathcal{B}(\mathbb{C}^n)$ (and more generally states on so-called **C^*-algebras**) are of *extreme* importance in determining the representation theory of these algebras. This, however, is beyond the scope of the present manuscript.

8.2. Definition. *Let V be a vector space and $C \subseteq V$ be a convex set. A point $e \in C$ is called an **extreme point** of C if whenever there exist $x, y \in C$ and $t \in (0, 1)$ for which*

$$e = tx + (1-t)y,$$

it follows that $x = y = e$. We denote by $\mathrm{Ext}(C)$ the (possibly empty) set of all extreme points of C.

8.3. Example.

(a) Let $V = \mathbb{C}$. Let $\mathbb{D} = \{w \in \mathbb{C} : |w| < 1\}$ denote the open disk. It is easy to see that $\mathbb{D}$ is convex. However $\mathbb{D}$ has no extreme points. If $w \in \mathbb{D}$, then $|w| < 1$, so there exists $\delta > 0$ so that $(1+\delta)|w| < 1$. Let $x = (1+\delta)w$, $y = (1-\delta)w$. Then $x, y \in \mathbb{D}$ and $w = \frac{1}{2}x + \frac{1}{2}y$.

(b) With $V = \mathbb{C}$ again, let $\overline{\mathbb{D}} = \{w \in \mathbb{C} : |w| \le 1\}$. Then every $z \in \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ is an extreme point of $\overline{\mathbb{D}}$. The proof of this is left as an exercise.

(c) Let $V = \mathbb{R}^2$, and let p_1, p_2, p_3 be three non-collinear points in V. The triangle T whose vertices are p_1, p_2, p_3 has exactly $\{p_1, p_2, p_3\}$ as its set of extreme points.

The following generalizes the concept of an extreme point.

8.4. Definition. *Let V be a vector space and let $\varnothing \ne C \subseteq V$ be convex. A non-empty convex set $F \subseteq C$ is called a **face** of C if whenever $x, y \in C$ and $t \in (0,1)$ satisfy $tx + (1-t)y \in F$, then $x, y \in F$.*

We emphasize the fact that F is convex is part of the definition of a face.

8.5. Remarks. Let V be a vector space and $C \subseteq V$ be convex.

(a) If e is an extreme point of C, then $F = \{e\}$ is a face of C. Conversely, if $F = \{z\}$ is a face of C, then $z \in \mathrm{Ext}(C)$.

(b) Let F be a face of C, and let D be a face of F. Then D is a face of C.

Indeed, let $x, y \in C$ and $t \in (0,1)$, and suppose that $tx + (1-t)y \in D$. Then $D \subseteq F$ implies that $tx + (1-t)y \in F$. Since F is a face of C, we must have $x, y \in F$. But then D is a face of F, and so it follows from $tx + (1-t)y \in D$ that $x, y \in D$.

(c) From (b), it follows that if e is an extreme point of a face F of C, then e is an extreme point of C.

8.6. Example.

(a) Let $V = \mathbb{R}^2$ and let p_1, p_2, p_3 be three non-collinear points in V. Denote by T the triangle whose vertices are p_1, p_2, p_3. Then T is a face of itself. Also, each line segment $\overline{p_i p_j}$ is a face of T. Finally, each extreme point p_j is a face of T.

(b) Let $V = \mathbb{R}^3$ and C be a cube in V, for e.g.,

$$C = \{(x, y, z) \in \mathbb{R}^3 : 0 \le x, y, z \le 1\}.$$

Then C has itself as a face. Also, the 6 (square) sides of the cube are faces. The 12 edges of the cube are also faces, as are the 8 corners. The corners are extreme points of the cube.

The definition of a face currently requires us to consider convex combinations of two elements of C. In fact, we may consider arbitrary finite convex combinations of elements of C.

8.7. Lemma. *Let $\mathcal{V}$ be a vector space, $\varnothing \neq C \subseteq \mathcal{V}$ be convex and $\varnothing \neq F \subseteq C$ be a face of C. Suppose that $\{x_j\}_{j=1}^n \subseteq C$ and that $x = \sum_{j=1}^n t_j x_j$ is a convex combination of the x_j's. If $x \in F$ and $t_j \in (0,1)$ for all $1 \leq j \leq n$, then $x_j \in F$ for all $1 \leq j \leq n$.*

Proof. We argue by induction on n. The assumption that $t_j \in (0,1)$ for all j requires that $n \geq 2$, and the case $n = 2$ is nothing more than the definition of a face. Let $k \geq 3$, and suppose that the result is true for $n < k$.

Suppose that $x = \sum_{j=1}^k t_j x_j \in F$, where $t_j \in (0,1)$ for all $1 \leq j \leq k$ and $\sum_{j=1}^k t_j = 1$. Then

$$x = (1 - t_k) \left(\sum_{j=1}^{k-1} \frac{t_j}{1 - t_k} x_j \right) + t_k x_k.$$

Since C is convex, $y := \sum_{j=1}^{k-1} \frac{t_j}{1-t_k} x_j \in C$. But then $x = (1 - t_k) y + t_k x_k \in F$, and F is a face, so that y and x_k must lie in F. Since $y \in F$, our induction hypothesis next implies that $x_j \in F$ for all $1 \leq j \leq k-1$, which completes the proof.

□

8.8. Lemma. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\varnothing \neq K \subseteq \mathcal{V}$ be a compact, convex set. Let $\rho \in \mathcal{V}^*$, and set*

$$r = \sup\{\operatorname{Re} \rho(w) : w \in K\}.$$

Then $F = \{x \in K : \operatorname{Re} \rho(x) = r\}$ is a non-empty, compact face of K.

Proof. Since $\operatorname{Re} \rho : K \to \mathbb{R}$ is continuous and K is compact, $r = \max\{\operatorname{Re} \rho(w) : w \in K\}$, and so F is non-empty. Moreover, $F = (\operatorname{Re} \circ \rho)^{-1}(\{r\})$ and $\{r\} \subseteq \mathbb{R}$ is closed, so F is closed in K, and hence F is compact.

Next, observe that if $x, y \in F \subseteq K$ and $t \in (0,1)$, then $tx + (1-t)y \in K$ as K is convex. But $\operatorname{Re} \rho(tx + (1-t)y) = t \operatorname{Re} \rho(x) + (1-t) \operatorname{Re} \rho(y) = tr + (1-t)r = r$, so that $tx + (1-t)y \in F$ and F is convex.

Suppose that $x, y \in K$, $t \in (0,1)$, and $tx + (1-t)y \in F$. As before,

$$\begin{aligned} r &= \operatorname{Re} \rho(tx + (1-t)y) \\ &= t \operatorname{Re} \rho(x) + (1-t) \operatorname{Re} \rho(y). \end{aligned}$$

But $\operatorname{Re} \rho(x) \leq r$, $\operatorname{Re} \rho(y) \leq r$, so the only way that equality can hold is if $x, y \in F$. Hence F is a face of K.

□

The following result is a crucial step in the proof of the Krein-Milman Theorem.

8.9. Lemma. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\varnothing \neq K \subseteq \mathcal{V}$ be a compact, convex set. Then* $\text{Ext}(K) \neq \varnothing$.

Proof. Let $\mathcal{J} = \{F \subseteq K : \varnothing \neq F$ is a closed face of $K\}$, and partially order $\mathcal{J}$ by reverse inclusion: i.e. $F_1 \leq F_2$ if $F_2 \subseteq F_1$. Observe that $K \in \mathcal{J}$ and so $\mathcal{J} \neq \varnothing$.

Suppose that $\mathcal{C} = \{F_\lambda\}_{\lambda \in \Lambda}$ is a chain in $\mathcal{J}$. We claim that $F = \cap_{\lambda \in \Lambda} F_\lambda$ is an upper bound for $\mathcal{C}$. Since $\{F_\lambda\}_{\lambda \in \Lambda}$ has the Finite Intersection Property and K is compact, $F \neq \varnothing$. Moreover, each F_λ is assumed to be closed and convex, and thus so is F. Suppose that $x, y \in K$, $t \in (0,1)$, and $tx + (1-t)y \in F$. Then $tx + (1-t)y \in F_\lambda$ for each λ. But F_λ is a face of K, so $x, y \in F_\lambda$ for all λ, whence $x, y \in F$ and F is face of K. Clearly it is an upper bound for $\mathcal{C}$.

By Zorn's Lemma, $\mathcal{J}$ contains a maximal element, say E. Since $E \in \mathcal{J}$, it is non-empty, convex, and closed in K, hence compact. We claim that E is a singleton set, and therefore corresponds to an extreme point of K.

Suppose to the contrary that there exist $x, y \in E$ with $x \neq y$. By the Hahn-Banach Theorem 05 (Theorem 6.43), there exists a continuous linear functional $\varphi \in \mathcal{V}^*$ so that

$$\text{Re}\,\varphi(x) > \text{Re}\,\varphi(y).$$

Since E is non-empty, convex and compact, we can apply Lemma 8.8. Let $r = \sup\{\text{Re}\,\varphi(w) : w \in E\}$, and set $H = \{x \in E : \text{Re}\,\varphi(x) = r\}$. Then H is a non-empty, compact face of E, and hence of K.

But at least one of x and y does not belong to H, and so $E < H$, contradicting the maximality of E. Thus $E = \{e\}$ is a singleton set, and $e \in \text{Ext}\,(E) \subseteq \text{Ext}(K)$, proving that the latter is non-empty.

□

8.10. Theorem. ***The Krein-Milman Theorem***

Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\varnothing \neq K \subseteq \mathcal{V}$ be a compact, convex set. Then

$$K = \overline{\text{co}}(\text{Ext}(K)),$$

the closed, convex hull of the extreme points of K.

Proof. By Lemma 8.9, $\text{Ext}(K) \neq \varnothing$. Thus $\varnothing \neq \overline{\text{co}}(\text{Ext}(K)) \subseteq K$, as K is closed and convex.

Suppose that $m \in K \backslash (\overline{\text{co}}(\text{Ext}(K)))$. By the Hahn-Banach Theorem (Theorem 6.43), there exists $\tau \in \mathcal{V}^*$ and real numbers $\alpha > \beta$ so that

$$\text{Re}\,\tau(m) \geq \alpha > \beta \geq \text{Re}\,\tau(b) \quad \text{for all } b \in \overline{\text{co}}(\text{Ext}(K)).$$

Let $s := \sup\{\text{Re}\,\tau(w) : w \in K\}$. Then $s \geq \text{Re}\,\tau(m) \geq \alpha$, and $L := \{z \in K : \text{Re}\,\tau(z) = s\}$ is a non-empty, compact face of K, by Lemma 8.8. But then $\varnothing \neq L$ is a compact, convex set in $\mathcal{V}$, and so by Lemma 8.9, $\text{Ext}(L) \neq \varnothing$. Furthermore, $\text{Ext}(L) \subseteq \text{Ext}(K)$, by virtue of the fact that L is a face of K (see Remark 8.5 (c)).

Hence there exists $e \in \text{Ext}(L) \subseteq \overline{\text{co}}(\text{Ext}(K))$ so that

$$\text{Re}\,\tau(e) = s \geq \alpha > \text{Re}\,\tau(b) \quad \text{for all } b \in \overline{\text{co}}(\text{Ext}(K)),$$

an obvious contradiction.

It follows that $K \backslash \overline{\text{co}}(\text{Ext}(K)) = \varnothing$, and thus $K = \overline{\text{co}}(\text{Ext}(K))$.

□

8.11. Corollary. *Let $(\mathcal{V}, \mathcal{T})$ be a LCS and $\varnothing \neq K \subseteq \mathcal{V}$ be a compact, convex set. If $\rho \in \mathcal{V}^*$, then there exists $e \in \mathrm{Ext}(K)$ so that*

$$\mathrm{Re}\, \rho(w) \leq \mathrm{Re}\, \rho(e) \quad \textit{for all } w \in K.$$

Proof. Let $r := \sup\{\mathrm{Re}\, \rho(w) : w \in K\}$. By Lemma 8.7, $F = \{x \in K : \mathrm{Re}\, \rho(x) = r\}$ is a non-empty, compact face of K. By Lemma 8.9, $\mathrm{Ext}(F) \neq \varnothing$. Let $e \in \mathrm{Ext}(F)$. Then $e \in \mathrm{Ext}(K)$, and

$$\mathrm{Re}\, \rho(w) \leq r = \mathrm{Re}\, \rho(e) \quad \text{for all } w \in K.$$

□

Equipped with the Krein-Milman Theorem 8.10 above, we are able to extend Corollary 7.23.

8.12. Corollary. *Let $\mathfrak{X}$ be a Banach space and suppose that $\mathcal{A} \subseteq \mathfrak{X}^*$ is weak*-closed and bounded. Then $\mathcal{A}$ is weak*-compact. If $\mathcal{A}$ is also convex, then $\mathcal{A} = \overline{\mathrm{co}}^{w^*}(\mathrm{Ext}\mathcal{A})$.*

Appendix to Section 8.

8.13. Convexity and convex hulls. It was anticipated that everyone would have seen the concept of **convexity** and of **convex hulls**. Hopefully, therefore, the following remarks will serve as a refresher for everyone.

Let $\mathcal{V}$ be a vector space over $\mathbb{K}$. Recall that a subset E of $\mathcal{V}$ is said to be **convex** if $x, y \in E$ and $0 \le t \le 1$ implies that $tx + (1-t)y \in E$. A simple finite induction argument shows that this definition is equivalent to requiring that if $n \ge 1$, $t_k \in [0, 1]$, $x_k \in E$ for all $1 \le k \le n$ and $\sum_{k=1}^n t_k = 1$, then $\sum_{k=1}^n t_k x_k \in E$.

A sum of the form
$$\sum_{k=1}^n t_k x_k$$
where $0 \le t_k \le 1$ and $\sum_{k=1}^n t_k x_k$ is called a **convex combination** of the x_k's.

If $\{E_\lambda\}_\lambda$ is a collection of convex subsets of $\mathcal{V}$, then $F := \cap_\lambda E_\lambda$ is once again convex. For if $x, y \in F$ and $0 \le t \le 1$, then for each λ, we have that $x, y \in E_\lambda$ and E_λ convex implies that $tx + (1-t)y \in E_\lambda$, whence $tx + (1-t)y \in \cap_\lambda E_\lambda = F$.

Now let $H \subseteq \mathcal{V}$ be a set. The **convex hull** $\mathrm{co}(H)$ of H is the set
$$\mathrm{co}(H) = \cap\{K \subseteq \mathcal{V} : H \subseteq K \text{ and } K \text{ is convex}\}.$$

It readily follows from the definition that $\mathrm{co}(H)$ is the smallest convex subset of $\mathcal{V}$ which contains H. Note that $\mathcal{V}$ is itself convex and clearly $H \subseteq \mathcal{V}$, so that the intersection on the right takes place over a non-empty collection.

Let us define
$$F = \{\sum_{k=1}^n t_k h_k : n \ge 1, t_k \ge 0, h_k \in H, 1 \le k \le n, \sum_{k=1}^n t_k = 1\}.$$

Observe that $h = 0h + 1h \in F$ for all $h \in H$, so that $H \subseteq F$.

If $x = \sum_{k=1}^n t_k h_k$ and $y = \sum_{j=1}^m s_j h'_j$ are elements of F (with $\sum_{k=1}^n t_k = 1 = \sum_{j=1}^m s_j$), then for $0 \le r \le 1$,
$$\begin{aligned} rx + (1-r)y &= r\left(\sum_{k=1}^n t_k h_k\right) + (1-r)\left(\sum_{j=1}^m s_j h'_j\right) \\ &= \sum_{k=1}^n (rt_k)h_k + \sum_{j=1}^m ((1-r)s_j)h'_j. \end{aligned}$$

But

$$\sum_{k=1}^{n} r t_k + \sum_{j=1}^{m} (1-r) s_j = r \left(\sum_{k=1}^{n} t_k \right) + (1-r) \left(\sum_{j=1}^{m} s_j \right) = r + (1-r) = 1,$$

and so $rx + (1-r)y \in F$. Thus $H \subseteq F$ and F is convex, so that $\mathrm{co}(H) \subseteq F$.

Furthermore, if K is any convex set which contains H, then from the first paragraph, $F \subseteq K$. Thus $F \subseteq \mathrm{co}(H)$.

In other words, a second description of the convex hull of H is:

$$\mathrm{co}(H) = \{ \sum_{k=1}^{n} t_k h_k : n \geq 1, t_k \geq 0, h_k \in H, 1 \leq k \leq n, \sum_{k=1}^{n} t_k = 1 \}.$$

That is, $\mathrm{co}(H)$ consists of all convex combinations of elements of H.

*

I want to go back to Brazil, get married, have lots of kids, and just be a couch tomato.

Ana Beatriz Barros

Exercises for Section 8.

Question 1.
Prove that there does not exist a Banach space $\mathfrak{X}$ such that $\mathfrak{X}^* = (c_0, \|\cdot\|_\infty)$.

Question 2.
Let $1 \le N$ be an integer. Find the extreme points of the closed unit ball of $(\mathcal{B}(\mathbb{C}^N), \|\cdot\|)$, where $\|\cdot\|$ denotes the operator norm.

Question 3.
Let $\mathcal{H} = \ell^2$ and let $\{e_n\}_{n=1}^\infty$ denote the standard ONB for $\mathcal{H}$. Let $S \in \mathcal{B}(\mathcal{H})$ be the unilateral forward shift defined by $Se_n = e_{n+1}$, $n \ge 1$.

Prove or disprove that S an extreme point of the closed unit ball of $\mathcal{B}(\mathcal{H})$.

Question 4.
Find the extreme points of the closed unit ball of each of the following spaces:

(a) $(\ell^1, \|\cdot\|_1)$;
(b) $(\ell^p, \|\cdot\|_p)$, $1 < p < \infty$;
(c) $(\ell^\infty, \|\cdot\|_\infty)$.

9. The chapter of named theorems

I believe that sex is one of the most beautiful, natural, wholesome things that money can buy.

Steve Martin

9.1. In general, if $f : X \to Y$ is a continuous map between topological spaces X and Y, one does *not* expect f to take open sets to open sets. Despite this, we have seen that if $\mathcal{V}$ is TVS and $\mathcal{W}$ is a closed subspace of $\mathcal{V}$, then the quotient map does just this.

The Open Mapping Theorem extends this result to surjections of Banach spaces. Many of the theorems in this Chapter are a consequence - either direct or indirect - of the Open Mapping Theorem. We begin with a Lemma which will prove crucial in the proof of the Open Mapping Theorem.

9.2. Lemma. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. If $\mathfrak{Y}_1^\circ \subseteq \overline{T\mathfrak{X}_m}$ for some $m > 0$, then $\mathfrak{Y}_1^\circ \subseteq T\mathfrak{X}_{2m}$.*

As a consequence, if $\overline{T\mathfrak{X}_r} \in \mathcal{U}_0^{\mathfrak{Y}}$ for some $r > 0$, then $T\mathfrak{X}_r \in \mathcal{U}_0^{\mathfrak{Y}}$.

Proof. Let $\delta > 0$, and observe that $\mathfrak{Y}_1^\circ \subseteq \overline{T\mathfrak{X}_m}$ if and only if $\mathfrak{Y}_\delta^\circ \subseteq \overline{T\mathfrak{X}_{\delta m}}$.

Choose $y \in \mathfrak{Y}_1^\circ$. Then there exists $x_1 \in \mathfrak{X}_m$ so that $\|y - Tx_1\| < 1/2$. Since $y - Tx_1 \in \mathfrak{Y}_{1/2}^\circ \subseteq \overline{T\mathfrak{X}_{m/2}}$, there exists $x_2 \in \mathfrak{X}_{m/2}$ so that $\|(y - Tx_1) - Tx_2\| < 1/4$, and thus

$$y - (Tx_1 + Tx_2) \in \mathcal{Y}_{1/4}^\circ.$$

Using an induction argument, for each $n \geq 1$, we can find $x_n \in \mathfrak{X}_{m/2^{n-1}}$ so that

$$\|y - \sum_{j=1}^n Tx_j\| < \frac{1}{2^n}.$$

Since $\mathfrak{X}$ is complete and $\sum_{n=1}^\infty \|x_n\| \leq \sum_{n=1}^\infty \frac{m}{2^{n-1}} = 2m$, we have $x = \sum_{n=1}^\infty x_n \in \mathfrak{X}_{2m}$. By the continuity of T,

$$Tx = T\left(\sum_{n=1}^\infty x_n\right) = \lim_{N\to\infty} T\left(\sum_{n=1}^N x_n\right) = y.$$

As for the last statement, suppose that $0 < r$ and that $\overline{T\mathfrak{X}_r} \in \mathcal{U}_0^{\mathfrak{Y}}$. Choose $\delta > 0$ such that $\mathfrak{Y}_\delta^\circ \subseteq \overline{T\mathfrak{X}_r}$. The linearity of T then implies that $\mathfrak{Y}_1^\circ \subseteq \overline{T\mathfrak{X}_{r/\delta}}$. From above,

$$\mathfrak{Y}_1^\circ \subseteq T\mathfrak{X}_{2r/\delta},$$

or equivalently,

$$\mathfrak{Y}_{\delta/2}^\circ \subseteq T\mathfrak{X}_r,$$

implying that $T\mathfrak{X}_r \in \mathcal{U}_0^{\mathfrak{Y}}$.

□

9.3. Theorem. ***The Open Mapping Theorem***
Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a surjection. Then T is an open map - i.e. if $G \subseteq \mathfrak{X}$ is open, then $TG \subseteq \mathfrak{Y}$ is open.

Proof. Step One. First let us argue that T is an open map if and only if $T\mathfrak{X}_1^\circ \in \mathcal{U}_0^{\mathfrak{Y}}$.

Indeed, if T is open, then $0 \in T\mathfrak{X}_1^\circ$, and $T\mathfrak{X}_1^\circ \subseteq \mathfrak{Y}$ is open by hypothesis. Thus $T\mathfrak{X}_1^\circ \in \mathcal{U}_0^{\mathfrak{Y}}$.

Conversely, suppose that $T\mathfrak{X}_1^\circ \in \mathcal{U}_0^{\mathfrak{Y}}$, and let $G \subseteq \mathfrak{X}$ be an open set. Note that $T\mathfrak{X}_1^\circ \in \mathcal{U}_0^{\mathfrak{Y}}$ implies that $T\mathfrak{X}_\varepsilon^\circ = \varepsilon(T\mathfrak{X}_1^\circ) \in \mathcal{U}_0^{\mathfrak{Y}}$ for all $\varepsilon > 0$. Fix $y_0 \in TG$, say $y_0 = Tx_0$, where $x_0 \in G$. Since G is open, there exists $\delta > 0$ such that $x_0 + \mathfrak{X}_\delta^\circ \subseteq G$. But then

$$y_0 + T(\mathfrak{X}_\delta^\circ) \subseteq TG,$$

and $y_0 + T(\mathfrak{X}_\delta^\circ) \in \mathcal{U}_{y_0}^{\mathfrak{Y}}$. Thus $y_0 \in \operatorname{int}(TG)$. Since $y_0 \in TG$ was arbitrary, TG is open.

Step Two. Now suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is surjective. Then

$$\mathfrak{Y} = T\mathfrak{X} = \cup_n T\mathfrak{X}_n \subseteq \cup_n \overline{T\mathfrak{X}_n}.$$

Now $\mathfrak{Y}$ is a complete metric space, and so by the Baire Category Theorem , there exists $m \geq 1$ so that $\operatorname{int}(\overline{T\mathfrak{X}_m}) \neq \varnothing$.

Let $y \in \operatorname{int}(\overline{T\mathfrak{X}_m})$, and choose $\delta > 0$ such that $y + \mathfrak{Y}_\delta^\circ(0) \subseteq \operatorname{int}(\overline{T\mathfrak{X}_m}) \subseteq \overline{T\mathfrak{X}_m}$. Then

$$\mathfrak{Y}_\delta^\circ \subseteq -y + \overline{T\mathfrak{X}_m} \subseteq \overline{T\mathfrak{X}_m} + \overline{T\mathfrak{X}_m} = \overline{T\mathfrak{X}_{2m}}.$$

(This last step uses the linearity and continuity of T.)

By Lemma 9.2, $T\mathfrak{X}_{2m} \in \mathcal{U}_0^{\mathfrak{Y}}$. But then $T\mathfrak{X}_1^\circ = \frac{1}{2m} T\mathfrak{X}_{2m} \in \mathcal{U}_0^{\mathfrak{Y}}$, and so by Step One, T is an open map.

□

9.4. Corollary. ***The Inverse Mapping Theorem***
Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is a bijection. Then T^{-1} is continuous, and so T is a homeomorphism.

Proof. If $G \subseteq \mathfrak{X}$ is open, then $(T^{-1})^{-1}(G) = TG$ is open in $\mathfrak{Y}$ by the Open Mapping Theorem above. Hence T^{-1} is continuous.

□

9.5. Corollary. ***The Closed Graph Theorem***
Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose that $T : \mathfrak{X} \to \mathfrak{Y}$ is linear. If the graph

$$\mathcal{G}(T) := \{(x, Tx) : x \in \mathfrak{X}\}$$

is closed in $\mathfrak{X} \oplus_1 \mathfrak{Y}$, then T is continuous.

Proof. The ℓ^1 norm on $\mathfrak{X} \oplus \mathfrak{Y}$ was chosen only so as to induce the product topology on $\mathfrak{X} \oplus \mathfrak{Y}$. We could have used any equivalent norm (for example, the ℓ^2 or ℓ^∞ norms).

Let $\pi_1 : \mathfrak{X} \oplus_1 \mathfrak{Y} \to \mathfrak{X}$ be the canonical projection $\pi_1(x, y) = x$, $(x, y) \in \mathfrak{X} \oplus \mathfrak{Y}$. Then π_1 is clearly linear, and

$$\|x\|_{\mathfrak{X}} = \|\pi_1(x, y)\|_{\mathfrak{X}} \leq \|x\|_{\mathfrak{X}} + \|y\|_{\mathfrak{Y}} = \|(x, y)\|_1,$$

so that $\|\pi_1\| \leq 1$. Moreover, $\mathcal{G}(T)$ is easily seen to be a linear manifold in $\mathfrak{X} \oplus_1 \mathfrak{Y}$, and by hypothesis, it is closed and hence a Banach space.

The map

$$\begin{array}{rccc} \pi_{\mathcal{G}} : & \mathcal{G}(T) & \to & \mathfrak{X} \\ & (x, Tx) & \mapsto & x \end{array}$$

is a linear bijection with $\|\pi_{\mathcal{G}}\| = \|\pi_1|_{\mathcal{G}(T)}\| \leq \|\pi_1\| \leq 1$.

By the Inverse Mapping Theorem 9.4 above, $\pi_{\mathcal{G}}^{-1}$ is also continuous, hence bounded.

Thus

$$\|Tx\|_{\mathfrak{Y}} \leq \|x\|_{\mathfrak{X}} + \|Tx\|_{\mathfrak{Y}} = \|(x, Tx)\|_1 = \|\pi_{\mathcal{G}}^{-1}(x)\|_1 \leq \|\pi_{\mathcal{G}}^{-1}\| \, \|x\|_{\mathfrak{X}}$$

for all $x \in \mathfrak{X}$, and therefore $\|T\| \leq \|\pi_{\mathcal{G}}^{-1}\| < \infty$. That is, T is continuous.

□

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. A linear map $T : \mathfrak{X} \to \mathfrak{Y}$ is continuous if and only if for all sequences $(x_n)_n$ in $\mathfrak{X}$ converging to $x \in \mathfrak{X}$, we have $\lim_n Tx_n = Tx$. Of course, given a linear map $T : \mathfrak{X} \to \mathfrak{Y}$, and given a sequence $(x_n)_n$ converging to x, there is no reason *a priori* to assume that $(Tx_n)_n$ converges to anything at all in $\mathfrak{Y}$. The following Corollary is interesting in that part (c) tells us that it in checking to see whether or not T is continuous, it suffices to *assume* that $\lim_{n\to\infty} Tx_n$ exists, and that we need only verify that the limit is the expected one, namely Tx. Linearity of T further reduces the problem to checking this condition for $x = 0$.

9.6. Corollary. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and $T : \mathfrak{X} \to \mathfrak{Y}$ be linear. The following are equivalent:*

(a) *The graph $\mathcal{G}(T)$ is closed.*
(b) *T is continuous.*
(c) *If $\lim_{n\to\infty} x_n = 0$ and $\lim_{n\to\infty} Tx_n = y$, then $y = 0$.*

Proof.

(a) implies (b): This is just the Closed Graph Theorem above.
(b) implies (c): This is clear.
(c) implies (a): Suppose that $((x_n, Tx_n))_{n=1}^\infty$ is a sequence in $\mathcal{G}(T)$ which converges to some point $(x, y) \in \mathfrak{X} \oplus_1 \mathfrak{Y}$. Then, in particular, $\lim_{n\to\infty} x_n = x$, and so $\lim_{n\to\infty}(x_n - x) = 0$. Also, $\lim_{n\to\infty} Tx_n = y$, so $\lim_{n\to\infty} T(x_n - x) = y - Tx$ exists. By our hypothesis, $y - Tx = 0$, or equivalently $y = Tx$. This in turn says that $(x, y) = (x, Tx) \in \mathcal{G}(T)$, and so the latter is closed.

□

Recall that a two closed subspaces $\mathfrak{Y}$ and $\mathfrak{Z}$ of a Banach space $\mathfrak{X}$ are said to topologically complement each other if $\mathfrak{X} = \mathfrak{Y} \oplus \mathfrak{Z}$.

9.7. Lemma. *Two closed subspaces $\mathfrak{Y}$ and $\mathfrak{Z}$ of a Banach space $\mathfrak{X}$ topologically complement each other if and only if the map*

$$\begin{array}{rccc} \iota : & \mathfrak{Y} \oplus_1 \mathfrak{Z} & \to & \mathfrak{X} \\ & (y,z) & \mapsto & y+z \end{array}$$

is a homeomorphism of Banach spaces.

Proof. First note that the norms on $\mathfrak{Y}$ and on $\mathfrak{Z}$ are nothing more than the restrictions to these spaces of the norm on $\mathfrak{X}$.

Suppose that $\mathfrak{Y}$ and $\mathfrak{Z}$ are topologically complementary subspaces of $\mathfrak{X}$. That ι is linear is clear. Moreover, since $\mathfrak{Y}$ and $\mathfrak{Z}$ are complementary subspaces, it is easy to see that ι is a bijection. Hence

$$\begin{aligned} \|\iota(y,z)\| &= \|y+z\| \\ &\le \|y\| + \|z\| \\ &= \|(y,z)\|, \end{aligned}$$

so that ι is a contraction. By the Inverse Mapping Theorem 9.4, ι^{-1} is continuous, and so ι is a homeomorphism.

Conversely, suppose that ι is a homeomorphism. Now $\operatorname{ran} \iota = \mathfrak{Y} + \mathfrak{Z} = \mathfrak{X}$, since ι is surjective, and if $w \in \mathfrak{Y} \cap \mathfrak{Z}$, then $(w, -w) \in \ker \iota = (0,0)$, so $w = 0$. Hence $\mathfrak{X}$ is the algebraic direct sum of $\mathfrak{Y}$ and $\mathfrak{Z}$. Since $\mathfrak{Y}$ and $\mathfrak{Z}$ are closed in $\mathfrak{X}$, they are also topologically complemented.

□

The next result extends our results from Section 3, where we showed that for a closed subspace $\mathcal{M}$ of a Hilbert space $\mathcal{H}$, there exists an orthogonal projection $P \in \mathcal{B}(\mathcal{H})$ whose range is $\mathcal{M}$ (see Remarks 3.7).

9.8. Proposition. *Let $\mathfrak{X}$ a Banach space and let $\mathfrak{Y}$ and $\mathfrak{Z}$ be topologically complementary subspaces of $\mathfrak{X}$. For each $x \in \mathfrak{X}$, denote by y_x and z_x the unique elements of $\mathfrak{Y}$ and $\mathfrak{Z}$ respectively such that $x = y_x + z_x$. Define $E : \mathfrak{X} \to \mathfrak{Y}$ via $Ex = E(y_x + z_x) = y_x$ for all $x \in \mathfrak{X}$. Then*

(a) *E is a continuous linear map. Moreover, $E = E^2$, $\operatorname{ran} E = \mathfrak{Y}$, and $\ker E = \mathfrak{Z}$.*
(b) *Conversely, if $E \in \mathcal{B}(\mathfrak{X})$ and $E = E^2$, then $\mathfrak{M} = \operatorname{ran} E$ and $\mathfrak{N} = \ker E$ are topologically complementary subspaces of $\mathfrak{X}$.*

Proof.

(a) From Lemma 9.7 above, we know that there exists a linear homeomorphism $\iota : \mathfrak{Y} \oplus_1 \mathfrak{Z} \to \mathfrak{X}$. Consider the map

$$\begin{array}{rccc} \pi_{\mathfrak{Y}} : & \mathfrak{Y} \oplus_1 \mathfrak{Z} & \to & \mathfrak{Y} \\ & (y,z) & \mapsto & y. \end{array}$$

It is clear that $\pi_{\mathfrak{Y}}$ is linear and contractive, and so $\pi_{\mathfrak{Y}}$ is continuous. As such, the map

$$\begin{array}{rccc} E := \pi_{\mathfrak{Y}} \circ \iota^{-1} : & \mathfrak{X} & \to & \mathfrak{Y} \\ & x & \mapsto & y_x \end{array}$$

is clearly linear (being the composition of linear functions), and

$$\|Ex\| = \|\pi_{\mathfrak{Y}} \circ \iota^{-1}\| \le \|\pi_{\mathfrak{Y}}\| \, \|\iota^{-1}\| < \infty,$$

so that E is bounded - i.e. E is continuous. That $\operatorname{ran} E = \mathfrak{Y}$ and $\ker\, E = \mathfrak{Z}$ are left as exercises.

(b) Since E is assumed to be continuous, $\mathfrak{N}$ is closed. Now $I - E$ is also continuous, and $\operatorname{ran} E = \ker(I - E)$, so $\operatorname{ran} E$ is also closed. If $z \in \operatorname{ran} E \cap \ker\, E$, then $z = Ew$ for some $w \in X$, so $z = E^2 w = Ez = 0$. Furthermore, for any $x \in \mathfrak{X}$, $x = Ex + (I - E)x \in \operatorname{ran} E + \ker\, E$. Hence $\mathfrak{M}$ and $\mathfrak{N}$ are algebraically complemented closed subspaces of $\mathfrak{X}$; i.e. they are topologically complemented.

□

9.9. Remark. A linear map $E \in \mathcal{B}(\mathfrak{X})$ is said to be **idempotent** if $E = E^2$. We point out that the term **projection** is often used in this context, although in the Hilbert space setting, the meaning of projection is slightly different.

The above Proposition says that a subspace $\mathfrak{Y}$ of a Banach space $\mathfrak{X}$ is complemented if and only if it is the range of a bounded idempotent in $\mathcal{B}(\mathfrak{X})$.

Appendix to Section 9.

9.10. As we mentioned in Chapter 3, not every closed subspace of a Banach space $\mathfrak{X}$ admits a topological complement. In particular, c_0 is not topologically complemented in ℓ^∞. Thus there does not exist a bounded idempotent $E \in \mathcal{B}(\ell^\infty)$ such that $\operatorname{ran} E = c_0$.

It is also the case that topological complements need not be unique. For example, let $\mathcal{H} = \ell^2$ and let $\{e_n\}_{n=1}^\infty$ denote the standard ONB for $\mathcal{H}$. Set $\mathcal{M} = \overline{\operatorname{span}}\{e_{2n}\}_{n=1}^\infty$, so that both $\mathcal{M}$ and $\mathcal{M}^\perp$ are infinite-dimensional subspaces of $\mathcal{H}$. Let $X \in \mathcal{B}(\mathcal{M}^\perp, \mathcal{M})$ be an arbitrary (bounded) linear operator, and consider the bounded linear operator E_X whose operator matrix relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ is

$$E_X := \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix}.$$

Then E_X is easily seen to be an idempotent, and $\operatorname{ran} E_X = \mathcal{M}$.

It follows from Proposition 9.8 that $\mathcal{M}$ is topologically complemented by $\mathcal{N}_X := \ker E_X$. Given $x \in \mathcal{H}$, write $x = y+z$, where $y \in \mathcal{M}$ and $z \in \mathcal{M}^\perp$. Then $x \in \ker E_X$ if and only if

$$E_X x = \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} y + Xz \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

i.e. if and only if $x = \begin{bmatrix} -Xz \\ z \end{bmatrix}$ for some $z \in \mathcal{M}^\perp$. That is, for each $X \in \mathcal{B}(\mathcal{M}^\perp, \mathcal{M})$, the closed subspace $\mathcal{N}_X := \left\{ \begin{bmatrix} -Xz \\ z \end{bmatrix} : z \in \mathcal{M}^\perp \right\}$ is a topological complement for $\mathcal{M}$.

Of course, $\mathcal{M}$ admits a unique **orthogonal** complement, which corresponds to the case where $X = 0$.

*

Nobody in the game of football should be called a genius. A genius is somebody like Norman Einstein.

Joe Theismann

Exercises for Section 9.

Question 1.

Let $\mathfrak{X}$ be a Banach space and $T \in \mathcal{B}(\mathfrak{X})$. Recall that a closed subspace $\mathfrak{Y}$ of $\mathfrak{X}$ is said to be **invariant** for T if $T\mathfrak{Y} \subseteq \mathfrak{Y}$.

Suppose that $\mathfrak{Y}$ and $\mathfrak{Z}$ are invariant for T, and that $\mathfrak{Z}$ is a topological complement for $\mathfrak{Y}$. Prove that there exists an idempotent $E \in \mathcal{B}(\mathfrak{X})$ such that $ET = TE$.

Question 2.

Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Prove that $E = E^2 \in \mathcal{B}(\mathcal{H})$ is an idempotent with $\operatorname{ran} E = \mathcal{M}$ if and only if there exists $X \in \mathcal{B}(\mathcal{M}^\perp, \mathcal{M})$ such that relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, we have

$$E = \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix}.$$

Question 3.

Let $\mathcal{H}$ be a Hilbert space and let $E = E^2 \in \mathcal{B}(\mathcal{H})$ be an idempotent operator. Prove that there exists an invertible operator $S \in \mathcal{B}(\mathcal{H})$ such that $P = S^{-1}ES$ is an orthogonal projection.

Question 4.

Let (X, d) be a metric space and $H \subseteq X$. We say that H is **nowhere dense** (or **meager**, or **thin**) if $G := X \setminus \overline{H}$ is dense in X. In other words, the interior of $\overline{H}$ is empty.

We say that a subset H of a metric space (X, d) is of the **first category** in (X, d) if there exists a sequence $(F_n)_{n=1}^\infty$ of closed, nowhere dense sets in X such that

$$H \subseteq \cup_{n=1}^\infty F_n.$$

Otherwise, H is said to be of the **second category**.

Prove the following alternate version of the **Banach-Steinhaus Theorem**:

Theorem. *Let $(\mathfrak{X}, \|\cdot\|_{\mathfrak{X}})$ and $(\mathfrak{Y}, \|\cdot\|_{\mathfrak{Y}})$ be Banach spaces and suppose that $\varnothing \neq \mathcal{F} \subseteq \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Let $H \subseteq \mathfrak{X}$ be a subset of the second category in $\mathfrak{X}$, and suppose that for each $x \in H$, there exists a constant $\kappa_x > 0$ such that*

$$\|Tx\|_{\mathfrak{Y}} \leq \kappa_x, \qquad T \in \mathcal{F}.$$

Then $\mathcal{F}$ is bounded; that is,

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

10. Operator Theory

I got kicked out of ballet class because I pulled a groin muscle. It wasn't mine.

Rita Rudner

10.1. Much of the work that has been done in Banach space theory has focussed on the study of the geometric structure of Banach spaces. For example, people have been interested in how "close" two Banach spaces are to being isomorphic as Banach spaces (say, in terms of the Banach-Mazur distance between them), and they have been interested in finding subspaces of a Banach space which have or are close to having a prescribed geometric structure. Other interesting questions in this area involve the study of how well finite-dimensional subspaces of one Banach space can be embedded in a second Banach space, and yet others involve the search for nice "bases" for a Banach space (for eg., the search for Schauder bases for quotients of Banach spaces), or the renorming of Banach spaces by equivalent norms. This list is anything but inclusive.

In contrast, the study of Hilbert spaces focusses very much on the structure of the bounded linear operators acting on the space. This is in part because Hilbert spaces are so well-behaved. This means that one pretends to understand the underlying space $\mathcal{H}$, and instead focusses upon understanding the more complicated structure, namely $\mathcal{B}(\mathcal{H})$. Because Hilbert spaces are so well-behaved, there is a chance of getting interesting and deep results about subsets and subalgebras of $\mathcal{B}(\mathcal{H})$.

Of course, one can also study the space $\mathcal{B}(\mathfrak{X})$ of bounded linear operators acting on a Banach space $\mathfrak{X}$. Here one is often interested in how the structure of the underlying space $\mathfrak{X}$ determines the operators in $\mathcal{B}(\mathfrak{X})$. In this Section we shall examine the notion of **compactness** of operators on a Banach space. We begin, however, with the notion of the **Banach space adjoint** of an operator.

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Let $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Given $y^* \in \mathfrak{Y}^*$, the map

$$x \mapsto y^*(Tx)$$

is a linear functional on $\mathfrak{X}$. Let us denote by $T^*y^* \in \mathfrak{X}^\#$ the functional on $\mathfrak{X}$ determined by this formula, namely:

$$T^*y^*(x) = y^*(Tx) \quad \text{for all } x \in \mathfrak{X}.$$

Observe that

$$\|T^*y^*(x)\| = \|y^*(Tx)\| \leq \|y^*\|\, \|T\|\, \|x\|,$$

and so $\|T^*y^*\| \leq \|T\|\, \|y^*\| < \infty$, implying that $T^*y^* \in \mathfrak{X}^*$. Furthermore, the map

$$\begin{array}{cccc} T^* & \mathfrak{Y}^* & \to & \mathfrak{X}^* \\ & y^* & \mapsto & T^*y^* \end{array}$$

is easily seen to be linear (by definition of linear combinations of functionals on $\mathfrak{X}$).

Moreover, the estimate

$$\|T^*y^*\| \le \|T\|\,\|y^*\|$$

for all $y^* \in \mathfrak{Y}^*$ implies that $\|T^*\| \le \|T\|$.

Conversely, let $x \in \mathfrak{X}$. By the Hahn-Banach Theorem, we can choose $y^* \in \mathfrak{Y}^*$ such that $\|y^*\| = 1$ and $y^*(Tx) = \|Tx\|$. Then

$$\begin{aligned} \|Tx\| &= y^*(Tx) \\ &= T^*y^*(x) \\ &\le \|T^*y^*\|\,\|x\| \\ &\le \|T^*\|\|x\|. \end{aligned}$$

Thus $\|T\| \le \|T^*\|$.

Combining this with the previous estimate, we have that

$$\|T^*\| = \|T\|.$$

10.2. Definition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and let $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. The map T^* defined above is called the* ***Banach space adjoint*** *of T.*

Many authors adopt the following notation for the action of a functional on a vector, namely: given $x \in \mathfrak{X}$ and $x^* \in \mathfrak{X}^*$, they write $\langle x, x^* \rangle$ to denote $x^*(x)$. In this notation, the equation $T^*y^*(x) = y^*(Tx)$ for all $x \in \mathfrak{X}$, $y^* \in \mathfrak{Y}^*$ becomes:

$$< x, T^*y^* > = < Tx, y^* >$$

for all $x \in \mathfrak{X}, y^* \in \mathfrak{Y}^*$. The reason for using this notation will become apparent when we look at Hilbert space adjoints below.

10.3. Proposition. *Let $\mathfrak{X}$, $\mathfrak{Y}$, $\mathfrak{Z}$ be Banach spaces, $S, T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$, and let $R \in \mathcal{B}(\mathfrak{Y}, \mathfrak{Z})$. Then*

(a) *for all k_1, $k_2 \in \mathbb{K}$, we have $(k_1 S + k_2 T)^* = k_1 S^* + k_2 T^*$;*
(b) $(R \circ T)^* = T^* \circ R^*$.

Proof.

(a) This is left as an exercise for the reader.
(b) Let $x \in \mathfrak{X}$, $y^* \in \mathfrak{Y}^*$, and $z^* \in \mathfrak{Z}^*$. Then

$$\begin{aligned} (R \circ T)^* z^*(x) &= z^*((R \circ T))(x) \\ &= z^*(R(Tx)) \\ &= R^* z^*(Tx) \\ &= T^* R^* z^*(x). \end{aligned}$$

Again, since this is true for all $x \in \mathfrak{X}$ and then for all $z^* \in \mathfrak{Z}^*$, we find that $(R \circ T)^* = T^* \circ R^*$.

□

10.4. Proposition. *Let $\mathfrak{X}$ be an n-dimensional Banach space over $\mathbb{K}$, and let $\mathcal{E} := \{e_1, e_2, \ldots, e_n\}$ be a (Hamel) basis for $\mathfrak{X}$. If $[A] = [a_{ij}]$ represents the matrix for A relative to $\mathcal{E}$, then the matrix of the Banach space adjoint A^* of A with respect to the dual basis $\mathcal{F}$ of $\mathcal{E}$ coincides with $[A]^t = [a_{ji}]$, the transpose of $[A]$.*

Proof. Recall that $\mathfrak{X}^* \simeq \mathfrak{X}$. We then let $\{e_i\}_{i=1}^n$ be a basis for $\mathfrak{X}$ and let $\{f_j\}_{j=1}^n$ be the corresponding dual basis for $\mathfrak{X}^*$; that is, $f_j(e_i) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. Let $x \in \mathfrak{X}$. Define $\lambda_j = f_j(x)$, $1 \le j \le n$.

Writing the matrix of $A \in \mathcal{B}(\mathfrak{X})$ as $[a_{ij}]$, we have

$$Ae_j = [a_{ij}] \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ 0 \\ 1 \\ 0 \\ . \\ . \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ . \\ . \\ a_{j-1\,j} \\ a_{jj} \\ a_{j+1\,j} \\ . \\ . \\ a_{nj} \end{bmatrix} = \sum_{k=1}^n a_{kj} e_k.$$

Thus $a_{ij} = f_i(Ae_j)$.

Now $A^* \in \mathcal{B}(\mathfrak{X}^*) \simeq \mathbb{M}_n$, and so we can also write the matrix $[\alpha_{ij}]$ for A^* with respect to $\{f_j\}_{j=1}^n$. Paralleling the above computation for $[\alpha_{ij}]$, we obtain:

$$A^* f_j = \sum_{k=1}^n \alpha_{kj} f_k,$$

and thus

$$\alpha_{ij} = (A^* f_j)(e_i) = f_j(Ae_i) = a_{ji}.$$

In particular, the matrix for A^* with respect to $\{f_j\}_{j=1}^n$ is simply the transpose of the matrix for A with respect to $\{e_j\}_{j=1}^n$.

□

The fact that a Hilbert space is isometrically isomorphic (via a conjugate-linear map) to its own dual allows us to define a separate notion of an adjoint for operators acting on these spaces. The crucial difference between the Hilbert space adjoint and the Banach space adjoint of an operator T acting on a Hilbert space $\mathcal{H}$ is that the Hilbert space adjoint will operate on the *same* Hilbert space $\mathcal{H}$, while the Banach space adjoint will be an operator in $\mathcal{B}(\mathcal{H}^*)$. While $\mathcal{H}^*$ is a Hilbert space isomorphic to $\mathcal{H}$, it is not $\mathcal{H}$ itself.

10.5. Theorem. *Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{B}(\mathcal{H})$. Then there exists a unique operator $T^* \in \mathcal{B}(\mathcal{H})$, called the* ***Hilbert space adjoint*** *of T, satisfying*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for all x, $y \in \mathcal{H}$.

Proof. Fix $y \in \mathcal{H}$. Then the map

$$\begin{array}{rccc} \phi_y : & \mathcal{H} & \to & \mathbb{C} \\ & x & \mapsto & \langle Tx, y\rangle \end{array}$$

is a linear functional and so there exists a vector $z_y \in \mathcal{H}$ such that

$$\phi_y(x) = \langle Tx, y\rangle = \langle x, z_y\rangle$$

for all $x \in \mathcal{H}$. Define a map $T^* : \mathcal{H} \to \mathcal{H}$ by $T^* y = z_y$. We leave it to the reader to verify that T^* is in fact linear, and we concentrate on showing that it is bounded.

To see that T^* is bounded, consider the following. Let $y \in \mathcal{H}$, $\|y\| = 1$. Then $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x \in \mathcal{H}$, so

$$\begin{aligned} \|T^*y\|^2 &= \langle T^*y, T^*y\rangle \\ &= \langle T\,T^*y, y\rangle \\ &\le \|T\|\,\|T^*y\|\,\|y\|. \end{aligned}$$

Thus $\|T^*y\| \le \|T\|$, and so $\|T^*\| \le \|T\|$.

Now T^* is unique, for if there exists $A \in \mathcal{B}(\mathcal{H})$ such that $\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, Ay\rangle$ for all x, $y \in \mathcal{H}$, then $\langle x, (T^*-A)y\rangle = 0$ for all x, $y \in \mathcal{H}$, and so $(T^*-A)y = 0$ for all $y \in \mathcal{H}$, i.e. $T^* = A$.

□

10.6. Corollary. *Let $T \in \mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space. Then $(T^*)^* = T$. It follows that $\|T\| = \|T^*\|$.*

Proof. For all $x, y \in \mathcal{H}$, we get

$$\begin{aligned} \langle x, (T^*)^*y\rangle &= \langle T^*x, y\rangle \\ &= \overline{\langle y, T^*x\rangle} \\ &= \overline{\langle Ty, x\rangle} \\ &= \langle x, Ty\rangle, \end{aligned}$$

and so $(T^*)^* = T$. Applying Theorem 10.5, we get

$$\|T\| = \|(T^*)^*\| \le \|T^*\| \le \|T\|,$$

and so $\|T\| = \|T^*\|$.

□

10.7. Proposition. *Let $\mathcal{H}$ be a complex, separable Hilbert space and $\mathcal{E} = \{e_n\}_n$ be a (countably infinite or finite) orthonormal basis for $\mathcal{H}$. Let $T \in \mathcal{B}(\mathcal{H})$. Then the matrix of T^* with respect to $\mathcal{E}$ is the conjugate transpose of that of T relative to $\mathcal{E}$.*

Proof. By definition, the matrix of T relative to $\mathcal{E}$ is given by $[t_{i,j}]$, where

$$t_{i,j} = \langle Te_j, e_i\rangle.$$

If we denote by $[r_{i,j}]$ the matrix of T^* relative to $\mathcal{E}$, then $r_{i,j} = \langle T^*e_j, e_i\rangle$ for all i, j.

But $t_{i,j} = \langle Te_j, e_i\rangle = \langle e_j, T^*e_i\rangle = \overline{\langle T^*e_i, e_j\rangle} = \overline{r_{j,i}}$, completing the proof.

□

10.8. Remark. For a Hilbert space $\mathcal{H}$ and $A, B \in \mathcal{B}(\mathcal{H})$, it is easy to see that we have $(AB)^* = B^* A^*$. Indeed, given $x, y \in \mathcal{H}$,

$$\langle x, (AB)^* y\rangle = \langle ABx, y\rangle = \langle Bx, A^* y\rangle = \langle x, B^* A^* y\rangle,$$

from which the result follows. The adjoint operator

$$* : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$$

is an example of an *involution* on a Banach algebra. Namely, for all α, $\beta \in \mathbb{C}$ and $A, B \in \mathcal{B}(\mathcal{H})$, we obtain

(i) $(\alpha A)^* = \overline{\alpha} A^*$;
(ii) $(A + B)^* = A^* + B^*$; and
(iii) $(AB)^* = B^* A^*$.
(iv) $(A^*)^* = A$.

A norm-closed subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ which is closed under the adjoint operation – i.e. $A \in \mathcal{A}$ implies that $A^* \in \mathcal{A}$ – is called a **(concrete)** C^***-algebra**.

10.9. Theorem. *Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Then $\|T^*T\| = \|T\|^2$.*

Proof.

- On the one hand,

$$\|T^*T\| \le \|T^*\| \, \|T\| \le \|T\| \, \|T\| = \|T\|^2.$$

- On the other hand, if $x \in \mathcal{H}$ and $\|x\| \le 1$, then (using the Cauchy-Schwarz Inequality), we find that

$$\begin{aligned} \|T^*T\| &\ge \|T^*Tx\| \\ &\ge \|T^*Tx\| \, \|x\| \\ &\ge |\langle T^*Tx, x\rangle| \\ &= \langle Tx, Tx\rangle \\ &= \|Tx\|^2. \end{aligned}$$

Since this holds for all $x \in \mathcal{H}$, $\|x\| \le 1$, it follows that $\|T^*T\| \ge \|T\|^2$.

□

The above equation is known as the C^***-equation**. While it was not terribly difficult to prove, this innocuous looking equation has amazing consequences on the structure of C^*-subalgebras of $\mathcal{B}(\mathcal{H})$.

In fact, it is a consequence of the **Gelfand-Naimark-Segal (GNS) construction** that if one starts with an involutive Banach algebra $\mathcal{A}$ for each each element $a \in \mathcal{A}$ satisfies the C^*-equation, that is $\|a^*a\| = \|a\|^2$, then there exists an isometric *-embedding of that algebra into the set of bounded linear operators on some Hilbert space. This is beyond the scope of this course, but well within the scope of the next!

10.10. Proposition. *Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Then $(\operatorname{ran} T)^{\perp} = \ker T^*$. In particular, therefore:*

(i) $\overline{\operatorname{ran} T} = (\ker T^*)^{\perp}$;
(ii) $\operatorname{ran} T$ *is not dense in* $\mathcal{H}$ *if and only if* $\ker T^* \neq 0$.

Proof. Let $y \in \mathcal{H}$. Then

$$\begin{aligned} y \in \ker T^* \quad & \text{if and only if} \quad \text{for all } x \in \mathcal{H},\ 0 = (x, T^*y), \\ & \text{if and only if} \quad \text{for all } x \in \mathcal{H},\ 0 = (Tx, y), \\ & \text{if and only if} \quad y \in (\operatorname{ran} T)^{\perp}. \end{aligned}$$

The second statement is now obvious.

□

10.11. Definition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Then T is said to be* ***compact*** *if $\overline{T(\mathfrak{X}_1)}$ is compact in $\mathfrak{Y}$. The set of compact operators from $\mathfrak{X}$ to $\mathfrak{Y}$ is denoted by $\mathcal{K}(\mathfrak{X}, \mathfrak{Y})$, and if $\mathfrak{Y} = \mathfrak{X}$, we simply write $\mathcal{K}(\mathfrak{X})$.*

Recall that a subset K of a metric space L is said to be **totally bounded** if for every $\varepsilon > 0$ there exists a finite cover $\{V_\varepsilon(y_i)\}_{i=1}^n$ of K with $y_i \in K$, $1 \leq i \leq n$, where $V_\varepsilon(y_i) = \{z \in L : \operatorname{dist}(z, y_i) < \varepsilon\}$.

We leave it as an exercise for the reader to show that if E is a subset of L and E is totally bounded, then so is $\overline{E}$.

10.12. Proposition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and $T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. The following are equivalent:*

(a) *T is compact;*
(b) *$\overline{T(F)}$ is compact in $\mathfrak{Y}$ for all bounded subsets F of $\mathfrak{X}$;*
(c) *If $(x_n)_n$ is a bounded sequence in $\mathfrak{X}$, then $(Tx_n)_n$ has a convergent subsequence in $\mathfrak{Y}$;*
(d) *$T(\mathfrak{X}_1)$ is totally bounded.*

Proof. This is left as an Assignment exercise.

□

10.13. Theorem. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Then $\mathcal{K}(\mathfrak{X}, \mathfrak{Y})$ is a closed subspace of $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$.*

Proof. This is left as an Assignment exercise.

□

10.14. Theorem. *Let $\mathfrak{W}$, $\mathfrak{X}$, $\mathfrak{Y}$, and $\mathfrak{Z}$ be Banach spaces. Suppose $R \in \mathcal{B}(\mathfrak{W}, \mathfrak{X})$, $K \in \mathcal{K}(\mathfrak{X}, \mathfrak{Y})$, and $T \in \mathcal{B}(\mathfrak{Y}, \mathfrak{Z})$. Then $TK \in \mathcal{K}(\mathfrak{X}, \mathfrak{Z})$ and $KR \in \mathcal{K}(\mathfrak{W}, \mathfrak{Y})$.*

Proof. Let $\mathfrak{X}_1$ denote the unit ball of $\mathfrak{X}$. Note that $\overline{K(\mathfrak{X}_1)}$ is compact and T is continuous, so that $T(\overline{K(\mathfrak{X}_1)})$ is compact and therefore closed.

$$\begin{aligned} \overline{T \circ K(\mathfrak{X}_1)} &= \overline{T(K(\mathfrak{X}_1))} \\ &\subseteq \overline{T(\overline{K(\mathfrak{X}_1)})} \\ &= T(\overline{K(\mathfrak{X}_1)}). \end{aligned}$$

Since $\overline{T \circ K(\mathfrak{X}_1)}$ is a closed subset of the compact set $T(\overline{K(\mathfrak{X}_1)})$, it is compact as well. Thus $TK \in \mathcal{K}(\mathfrak{X}, \mathfrak{Z})$.

Now if $\mathfrak{W}_1$ is the unit ball of $\mathfrak{W}$, then

$$\overline{KR(\mathfrak{W}_1)} = \overline{K(R(\mathfrak{W}_1))};$$

but $R(\mathfrak{W}_1)$ is bounded since R is, and so by Proposition 10.12, $\overline{KR(\mathfrak{W}_1)}$ is compact. Thus $KR \in \mathcal{K}(\mathfrak{W}, \mathfrak{Y})$.

□

10.15. Corollary. *If $\mathfrak{X}$ is a Banach space, then $\mathcal{K}(\mathfrak{X})$ is a closed, two-sided ideal of $\mathcal{B}(\mathfrak{X})$.*

10.16. Proposition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and assume that $K \in \mathcal{K}(\mathfrak{X}, \mathfrak{Y})$. Then $K(\mathfrak{X})$ is closed in $\mathfrak{Y}$ if and only if* $\dim K(\mathfrak{X})$ *is finite.*

Proof. $K(\mathfrak{X}) = \operatorname{ran} K$ is easily seen to be a submanifold of $\mathfrak{Y}$. Since finite-dimensional manifolds are always closed, we find that $\dim K(\mathfrak{X}) < \infty$ implies $K(\mathfrak{X})$ is closed.

Now assume that $K(\mathfrak{X})$ is closed. Then $K(\mathfrak{X})$ is a Banach space and the map

$$\begin{aligned} K_0 : \ \mathfrak{X} &\to K(\mathfrak{X}) \\ x &\mapsto Kx \end{aligned}$$

is a surjection. By the Open Mapping Theorem, Theorem 9.3, it is also an open map. In particular, $K_0(\operatorname{int} \mathfrak{X}_1)$ is open in $K(\mathfrak{X})$ and $0 \in K_0(\operatorname{int} \mathfrak{X}_1)$. Let G be an open ball in $K(\mathfrak{X})$ centred at 0 and contained in $K_0(\operatorname{int} \mathfrak{X}_1)$. Then $\overline{K_0(\mathfrak{X}_1)} = \overline{K(\mathfrak{X}_1)}$ is compact, hence closed, and also contains $\overline{G}$. Thus $\overline{G}$ is compact in $K(\mathfrak{X})$ and so $\dim K(\mathfrak{X})$ is finite - see Corollary 4.26.

□

10.17. Definition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Then $F \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ is said to be* ***finite rank*** *if* $\dim F(\mathfrak{X})$ *is finite. The set of finite rank operators from $\mathfrak{X}$ to $\mathfrak{Y}$ is denoted by $\mathcal{F}(\mathfrak{X}, \mathfrak{Y})$.*

10.18. Proposition. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Then $\mathcal{F}(\mathfrak{X}, \mathfrak{Y}) \subseteq \mathcal{K}(\mathfrak{X}, \mathfrak{Y})$.*

Proof. Suppose $F \in \mathcal{F}(\mathfrak{X}, \mathfrak{Y})$. Then $\overline{F\mathfrak{X}_1}$ is closed and bounded in $\operatorname{ran} F$, but $\operatorname{ran} F$ is finite dimensional in $\mathfrak{Y}$, as F is finite rank. Thus $\overline{F\mathfrak{X}_1}$ is compact in $\operatorname{ran} F$, and thus compact in $\mathfrak{Y}$ as well, showing that F is compact.

□

10.19. Proposition. *Let $\mathfrak{X}$ be a Banach space. Then $\mathcal{K}(\mathfrak{X}) = \mathcal{B}(\mathfrak{X})$ if and only if $\mathfrak{X}$ is finite dimensional.*

Proof. If $\dim \mathfrak{X} < \infty$, then $\mathcal{B}(\mathfrak{X}) = \mathcal{F}(\mathfrak{X}) \subseteq \mathcal{K}(\mathfrak{X}) \subseteq \mathcal{B}(\mathfrak{X})$, and equality follows.

If $\mathcal{K}(\mathfrak{X}) = \mathcal{B}(\mathfrak{X})$, then $I \in \mathcal{K}(\mathfrak{X})$, so $\overline{I(\mathfrak{X}_1)} = \mathfrak{X}_1$ is compact. In particular, $\mathfrak{X}$ is finite dimensional.

□

10.20. Theorem. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces and suppose $K \in \mathcal{K}(\mathfrak{X}, \mathfrak{Y})$. Then $K^* \in \mathcal{K}(\mathfrak{Y}^*, \mathfrak{X}^*)$.*

Proof. Let $\varepsilon > 0$. Then $K(\mathfrak{X}_1)$ is totally bounded, so we can find $x_1, x_2, \ldots, x_n \in \mathfrak{X}_1$ such that if $x \in \mathfrak{X}_1$, then $\|Kx - Kx_i\| < \varepsilon/3$ for some $1 \leq i \leq n$. Let

$$\begin{array}{rccc} R: & \mathfrak{Y}^* & \to & (\mathbb{C}^n, \|\cdot\|_\infty) \\ & y^* & \mapsto & (y^*(K(x_1)), y^*(K(x_2)), \ldots, y^*(K(x_n))). \end{array}$$

Then $R \in \mathcal{F}(\mathfrak{Y}^*, \mathbb{C}^n) \subseteq \mathcal{K}(\mathfrak{Y}^*, \mathbb{C}^n)$, and so $R(\mathfrak{Y}^*_1)$ is totally bounded. Thus we can find $y_1^*, y_2^*, \ldots, y_m^* \in \mathfrak{Y}_1^*$ such that if $y^* \in \mathfrak{Y}_1^*$, then $\|Ry^* - Ry_j^*\| < \varepsilon/3$ for some $1 \leq j \leq m$. Now

$$\begin{aligned} \|Ry^* - Ry_j^*\| &= \max_{1 \leq i \leq n} |y^*(K(x_i)) - y_j^*(K(x_i))| \\ &= \max_{1 \leq i \leq n} |K^*(y^*)(x_i) - K^*(y_j^*)(x_i)|. \end{aligned}$$

Suppose $x \in \mathfrak{X}_1$. Then $\|Kx - Kx_i\| < \epsilon/3$ for some $1 \leq i \leq n$, and $|K^*(y^*)(x_i) - K^*(y_j^*)(x_i)| < \varepsilon/3$ for some $1 \leq j \leq m$, so

$$\begin{aligned} |K^*(y^*)(x) - K^*(y_j^*)(x)| &\leq |K^*(y^*)(x) - K^*(y^*)(x_i)| + \\ &\qquad |K^*(y^*)(x_i) - K^*(y_j^*)(x_i)| + \\ &\qquad\quad |K^*(y_j^*)(x_i) - K^*(y_j^*)(x)| \\ &\leq \|y^*\| \, \|Kx - Kx_i\| + \varepsilon/3 + \|y_j^*\| \, \|Kx - Kx_i\| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

Thus $\|K^*y^* - K^*y_j^*\| \leq \epsilon$ and so $K^*(\mathfrak{Y}_1^*)$ is totally bounded. We conclude that $K^* \in \mathcal{K}(\mathfrak{Y}^*, \mathfrak{X}^*)$.

□

10.21. The set of compact operators acting on a Hilbert space is more tractable in general than the set of compact operators acting on an arbitrary Banach space. One of the reasons for this is the characterization given below.

10.22. Theorem. *Let $\mathcal{H}$ be a Hilbert space and let $K \in \mathcal{B}(\mathcal{H})$. The following are equivalent:*

(i) *K is compact;*
(ii) *K^* is compact;*
(iii) *There exists a sequence $\{F_n\}_{n=1}^\infty \subseteq \mathcal{F}(\mathcal{H})$ such that $K = \lim_{n\to\infty} F_n$.*

Proof.

(i) $\Rightarrow$ (iii) Let B_1 denote the unit ball of $\mathcal{H}$, and let $\epsilon > 0$. Since $\overline{K(B_1)}$ is compact, it must be separable (i.e. it is totally bounded). Thus $\mathcal{M} = \overline{\mathrm{ran}\, K}$ is a separable subspace of $\mathcal{H}$, and thus possesses an orthonormal basis $\{e_n\}_{n=1}^\infty$.

Let P_n denote the orthogonal projection of $\mathcal{H}$ onto span $\{e_k\}_{k=1}^n$. Set $F_n = P_n K$, noting that each F_n is finite rank. We now show that $K = \lim_{n\to\infty} F_n$.

Let $x \in \mathcal{H}$ and consider $y = Kx \in \mathcal{M}$, so that $\lim_{n\to\infty} \|P_n y - y\| = 0$. Thus $\lim_{n\to\infty} \|F_n x - Kx\| = \lim_{n\to\infty} \|P_n y - y\| = 0$. Since K is compact, $K(B_1)$ is totally bounded, so we can choose $\{x_k\}_{k=1}^m \subseteq B_1$ such that $K(B_1) \subseteq \cup_{k=1}^m B(Kx_k, \epsilon/3)$, where given $z \in \mathcal{H}$ and $\delta > 0$, $B(z, \delta) = \{w \in \mathcal{H} : \|w - z\| < \delta\}$.

If $\|x\| \le 1$, choose i such that $\|Kx_i - Kx\| < \epsilon/3$. Then for any $n > 0$,

$$\begin{aligned} \|Kx &- F_n x\| \\ &\le \|Kx - Kx_i\| + \|Kx_i - F_n x_i\| + \|F_n x_i - F_n x\| \\ &< \epsilon/3 + \|Kx_i - F_n x_i\| + \|P_n\|\,\|Kx_i - Kx\| \\ &< 2\epsilon/3 + \|Kx_i - F_n x_i\|. \end{aligned}$$

Choose $N > 0$ such that $\|Kx_i - F_n x_i\| < \epsilon/3$, $1 \le i \le m$ for all $n > N$. Then $\|Kx - F_n x\| \le 2\epsilon/3 + \epsilon/3 = \epsilon$. Thus $\|K - F_n\| < \epsilon$ for all $n > N$. Since $\epsilon > 0$ was arbitrary, $K = \lim_{n\to\infty} F_n$.

(iii) $\Rightarrow$ (ii) Suppose $K = \lim_{n\to\infty} F_n$, where F_n is finite rank for all $n \ge 1$. Note that F_n^* is also finite rank (why?), and that $\|K^* - F_n^*\| = \|K - F_n\|$ for all $n \ge 1$, which clearly implies that $K^* = \lim_{n\to\infty} F_n^*$, and hence that K^* is compact.

(ii) $\Rightarrow$(i) Since K compact implies K^* is compact from above, we deduce that K^* compact implies $(K^*)^* = K$ is compact, completing the proof.

□

We can restate the above Theorem more succinctly by saying that $\mathcal{K}(\mathcal{H})$ is the norm closure of the set of finite rank operators on $\mathcal{H}$. This is an extraordinarily useful result.

10.23. Remark. Contained in the above proof is the following interesting observation. If K is a compact operator acting on a separable Hilbert space $\mathcal{H}$, then for *any* sequence $\{P_n\}_{n=1}^\infty$ of finite rank projections tending strongly (i.e. pointwise) to the identity, $\|K - P_n K\|$ tends to zero. By considering adjoints, we find that $\|K - KP_n\|$ also tends to zero.

Let $\epsilon > 0$, and choose $N > 0$ such that $n \ge N$ implies $\|K - KP_n\| < \epsilon/2$ and $\|K - P_n K\| < \epsilon/2$. Then for all $n \ge N$ we get

$$\begin{aligned} \|K - P_n K P_n\| &\le \|K - KP_n\| + \|KP_n - P_n K P_n\| \\ &\le \|K - KP_n\| + \|K - P_n K\| \, \|P_n\| \\ &< \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

It follows that if $\mathcal{H}$ has an orthonormal basis indexed by the natural numbers, say $\{e_n\}_{n=1}^{\infty}$, then the matrix for K with respect to this basis comes within ϵ of the matrix for $P_N K P_N$. In other words, K "virtually lives" on the "top left-hand corner".

Alternatively, if $\mathcal{H}$ has an orthonormal basis indexed by the integers, say $\{f_n\}_{n \in \mathbb{Z}}$, and we let P_n denote the orthogonal projection onto span $\{e_k\}_{k=-n}^{n}$, then the matrix for K with respect to this basis can be arbitrarily well estimated by a sufficiently large but finite "central block".

10.24. Example. Let $\mathcal{H}$ be a separable Hilbert space with orthonormal basis $\{e_n\}_{n=1}^{\infty}$. Let $\{d_n\}_{n=1}^{\infty}$ be a bounded sequence and consider the diagonal operator $D \in \mathcal{B}(\mathcal{H})$ defined locally by $De_n = d_n e_n$ and extended to all of $\mathcal{H}$ by linearity and continuity.

Then $D \in \mathcal{K}(\mathcal{H})$ if and only if $\lim_{n \to \infty} d_n = 0$.

10.25. Example. Let $\mathcal{H} = L^2([0,1], dx)$, and consider the function $k(x,t) \in L^2([0,1] \times [0,1], dm)$, where dm represents Lebesgue planar measure. Then we define a Volterra operator

$$\begin{aligned} V: \ L^2([0,1], dx) &\to L^2([0,1], dx) \\ (Vf)(x) &= \textstyle\int_0^1 f(t)\, k(x,t)\, dt. \end{aligned}$$

(The classical Volterra operator has $k(x,t) = 1$ if $x \ge t$, and $k(x,t) = 0$ if $x < t$.) Now for $f \in L^2([0,1], dx)$ we have

$$\begin{aligned} \|Vf\|^2 &= \int_0^1 |Vf(x)|^2 dx \\ &= \int_0^1 \left| \int_0^1 f(t)\, k(x,t) dt \right|^2 dx \\ &\le \int_0^1 \left(\int_0^1 |f(t)\, k(x,t)| dt \right)^2 dx \\ &\le \int_0^1 \|f\|_2^2 \int_0^1 |k(x,t)|^2 dt dx \text{ by the Cauchy-Schwartz Inequality} \\ &= \|f\|_2^2 \, \|k\|_2^2, \end{aligned}$$

so that $\|V\| \le \|k\|_2$.

Let $\mathcal{A}$ denote the algebra of continuous functions on $[0,1] \times [0,1]$ which can be resolved as $g(x,t) = \sum_{i=1}^{n} u_i(x)\, w_i(t)$. Then $\mathcal{A}$ is an algebra which separates points, contains the constant functions, and is closed under complex conjugation. By the

Stone-Weierstraß Theorem, given $\epsilon > 0$ and $h \in \mathcal{C}([0,1] \times [0,1])$, there exists $g \in \mathcal{A}$ such that $\|h - g\|_2 \le \|h - g\|_\infty < \epsilon$. But since $\mathcal{C}([0,1] \times [0,1])$ is dense (in the L^2-topology) in $L^2([0,1] \times [0,1], dm)$, $\mathcal{A}$ must also be dense (in the L^2-topology) in $L^2([0,1] \times [0,1], dm)$.

Let $\epsilon > 0$. For k as above, choose $g \in \mathcal{A}$ such that $\|k - g\|_2 < \epsilon$. Define

$$\begin{array}{rlcl} V_0 : & L^2([0,1], dx) & \to & L^2([0,1], dx) \\ & V_0 f(x) & = & \int_0^1 f(t)\, g(x,t) dt. \end{array}$$

From above, we find that $\|V - V_0\| \le \|k - g\|_2 < \epsilon$.

To see that V_0 is finite rank, consider the following; first, $g(x,t) = \sum_{i=1}^n u_i(x)\, w_i(t)$. If we set $\mathcal{M} = \text{span}_{\,1 \le i \le n}\{u_i\}$, then $\mathcal{M}$ is a finite dimensional subspace of $L^2([0,1], dx)$. Moreover,

$$\begin{aligned} V_0 f(x) &= \int_0^1 f(t)\, g(x,t) dt \\ &= \sum_{i=1}^n \left(\int_0^1 f(t)\, w_i(t) dt \right) u_i(x), \end{aligned}$$

so that $V_0 f \in \mathcal{M}$.

Thus V can be approximated arbitrarily well by elements of the form $V_0 \in \mathcal{F}(L^2([0,1], dx)$, and so V is compact.

10.26. Definition. *Let $\mathcal{H}$ be a Hilbert space, $\mathcal{M}$ be a subspace of $\mathcal{H}$, and suppose that $T \in \mathcal{B}(\mathcal{H})$. Recall that $\mathcal{M}$ is called* ***invariant*** *for T provided that $T\mathcal{M} \subseteq \mathcal{M}$. We say that $\mathcal{M}$ is* ***reducing*** *for T if $\mathcal{M}$ is invariant both for T and for T^*.*

10.27. Exercises. Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{M}$ be a closed subspace of $\mathcal{H}$. Let P be the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$, and $T \in \mathcal{B}(\mathcal{H})$.

(a) Prove that relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, we have

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

(b) Prove that $P = P^2 = P^*$.

(c) More generally, write $T = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}$ relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Prove that

$$T^* = \begin{bmatrix} T_1^* & T_3^* \\ T_2^* & T_4^* \end{bmatrix}.$$

(d) Prove that $\mathcal{M}$ is invariant for T if and only if $(I - P)TP = 0$, and $\mathcal{M}$ is reducing for T if and only if $TP = PT$. Conclude that $\mathcal{M}$ is reducing for T if and only if both $\mathcal{M}$ and $\mathcal{M}^\perp$ are invariant for T.

10.28. Proposition. *Let $\mathcal{H}$ be a Hilbert space, $T \in \mathcal{B}(\mathcal{H})$, and $\mathcal{M}$ be a reducing subspace of $\mathcal{H}$. Then*

$$T = T_1 \oplus T_4 = \begin{bmatrix} T_1 & 0 \\ 0 & T_4 \end{bmatrix}$$

with respect to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$. Furthermore, T is compact if and only if both T_1 and T_4 are compact, and T is normal if and only if T_1 and T_4 are.

Proof. Let us denote by P the orthogonal projection of P onto $\mathcal{M}$. The matrix form for T follows directly from the matrix form for P computed in Exercise 10.27 above, combined with the equation $PT = TP$.

- If T_1 and T_4 are compact, then they are limits of finite rank operators F_n and G_n respectively, from which we conclude that T is a limit of the finite rank operators $F_n \oplus G_n$. Thus T is compact.

  If T is compact, then the compression of T to any subspace is compact, and so both T_1 and T_4 are compact.
- Observe that T is normal if and only if

  $$0 = T^*T - TT^* = \begin{bmatrix} T_1^*T_1 - T_1T_1^* & 0 \\ 0 & T_4^*T_4 - T_4T_4^* \end{bmatrix},$$

  which is equivalent to the simultaneous normality of T_1 and T_4.

□

10.29. Proposition. *Let $\mathcal{H}$ be a Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be normal.*

(a) *For all $x \in \mathcal{H}$,*

$$\|Nx\| = \|N^*x\|.$$

(b) *Let $p(x, y)$ be any polynomial in two non-commuting variables x and y. Given $\alpha \in \mathbb{C}$, $\ker(p(N, N^*) - \alpha I) = \ker(p(N, N^*) - \alpha I)^*$ is a reducing subspace for N.*

(c) *If $\alpha \neq \beta \in \mathbb{C}$, then $\ker(N - \alpha I)$ is orthogonal to $\ker(N - \beta I)$.*

Proof.

(a) Suppose that $x \in \mathcal{H}$. Then

$$\|N^*x\|^2 = \langle N^*x, N^*x\rangle = \langle NN^*x, x\rangle = \langle N^*Nx, x\rangle = \langle Nx, Nx\rangle = \|Nx\|^2.$$

As such, $\|Nx\| = \|N^*x\|$, and it is clear that $\ker N = \ker N^*$.

(b) First note that the fact that N is normal means that to calculate $p(N, N^*)$, we could just have easily taken a polynomial in two *commuting* variables.

Let $\alpha \in \mathbb{C}$. Since N is normal, so is $p(N, N^*)$. From part (a) we may conclude that $\ker(p(N, N^*) - \alpha I) = \ker(p(N, N^*) - \alpha I)^*$.

Consider $x \in \ker(p(N, N^*) - \alpha I)$. Then

$$(p(N, N^*) - \alpha I)Nx = N(p(N, N^*) - \alpha I)x = N0 = 0,$$

showing that $Nx \in \ker(p(N, N^*) - \alpha I)$, i.e that $\ker(p(N, N^*) - \alpha I)$ is invariant for N. Similarly,

$$(p(N, N^*) - \alpha I)N^*x = N^*(p(N, N^*) - \alpha I)x = N^*0 = 0,$$

showing that $N^*x \in \ker(p(N, N^*) - \alpha I)$, i.e that $\ker(p(N, N^*) - \alpha I)$ is invariant for N^*.

By Exercise 10.27, $\ker(p(N, N^*) - \alpha I)$ is reducing for N.

(c) Suppose that $\alpha \neq \beta \in \mathbb{C}$. Let $x \in \ker(N - \alpha I)$ and $y \in \ker(N - \beta I)$. Then

$$\alpha\langle x, y\rangle = \langle Nx, y\rangle = \langle x, N^*y\rangle = \langle x, \overline{\beta}y\rangle = \beta\langle x, y\rangle.$$

Since $\alpha \neq \beta$, we must have $x \perp y$.

□

10.30. Proposition. *Let $\mathcal{H}$ be a complex Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be a compact, normal operator. Then $N = 0$ if and only if $\sigma(N) = \{0\}$.*

Proof. That $N = 0$ implies that $\sigma(N) = \{0\}$ is trivial.

Conversely, suppose that $\sigma(N) = \{0\}$. First observe that if $\dim \mathcal{H} < \infty$, then the result follows easily from the fact that every normal matrix is diagonalizable (this is the Spectral Theorem for Normal Matrices). This reduces our problem to the case where $\dim \mathcal{H} = \infty$.

Case One. N is a finite-rank operator.

In this case, $\operatorname{ran} N$ is a finite-dimensional and therefore *closed* subspace of $\mathcal{H}$. Recall that

$$(\operatorname{ran} N) = \overline{\operatorname{ran} N} = (\ker N^*)^\perp = (\ker N)^\perp = \overline{(\operatorname{ran} N^*)} = \operatorname{ran} N^*.$$

(The last equality follows from the fact that $\dim(\operatorname{ran} N^*) < \infty$.) Decomposing $\mathcal{H} = (\ker N)^\perp \oplus (\ker N)$, we may write

$$N = \begin{bmatrix} N_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since $\kappa := \dim(\ker N)^\perp = \dim(\operatorname{ran} N) < \infty$, $N_1 \in \mathcal{B}(\mathbb{C}^\kappa)$ corresponds to a normal matrix, and $\sigma(N_1) \subseteq \sigma(N) = \{0\}$ implies that $\sigma(N_1) = \{0\}$ as well. As seen above, this implies that $N_1 = 0$, so that $N = 0$ as well.

Case Two. N is not a finite-rank operator.

We shall argue this case by contradiction. Suppose that $N \neq 0$. By $\|N\|^{-1} N$ if necessary, we may assume without loss of generality that $\|N\| = 1$.

Recall that since N is compact, $\mathcal{H}_0 := \overline{\operatorname{ran}} N$ is a separable subspace of $\mathcal{H}$. Let $\{e_n\}_{n=1}^\infty$ be an ONB for $\mathcal{H}_0$, and for each $n \geq 1$, set P_n to be the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_n := \operatorname{span}\{e_1, e_2, \ldots, e_n\}$. As seen in the notes, if we set $F_n = P_n N P_n$, $n \geq 1$, then each F_n has finite rank and

$$N = \lim_n F_n.$$

It follows that

$$N^*N = \lim_n F_n^* F_n.$$

But relative to $\mathcal{H} = \mathcal{H}_n \oplus \mathcal{H}_n^\perp$, we may write $F_n = \begin{bmatrix} A_n & 0 \\ 0 & 0 \end{bmatrix}$ for some $A_n \in \mathcal{B}(\mathbb{C}^n)$, whence

$$F_n^* F_n = \begin{bmatrix} A_n^* A_n & 0 \\ 0 & 0 \end{bmatrix}.$$

Now $A_n^* A_n$ is a positive semidefinite matrix, and therefore it is diagonalizable. Moreover, $\lim_n \|F_n\| = \|N\| = 1$. Thus

$$\lim_n \|A_n^* A_n\| = \lim_n \|A_n\|^2 = \lim_n \|F_n\|^2 = \|N\|^2 = 1.$$

From this we see that with $\alpha_n = \|A_n\|^2$, $\lim_n \alpha_n = 1$ and $\alpha_n \in \sigma(A_n^* A_n) \subseteq \sigma(F_n^* F_n)$ for all $n \geq 1$.

Next, recall that the set of invertible operators in $\mathcal{B}(\mathcal{H})$ is open, and note that

$$N^* N - I = \lim_n F_n^* F_n - \alpha_n I.$$

Together, these imply that $1 \in \sigma(N^* N)$. But $N^* N$ is compact, and so $1 \in \sigma_p(N^* N)$ and $\{0\} \neq \mathfrak{M} := \ker (N^* N - I)$ is finite-dimensional.

We are almost there! A routine calculation shows that $\mathfrak{M}$ is invariant for both N and N^*. (This uses the fact that N and N^* are normal, and thus they commute with $(N^* N - I)$.)

Hence we may decompose $\mathcal{H} = \mathfrak{M} \oplus \mathfrak{M}^\perp$ and write

$$N = \begin{bmatrix} N_{1,1} & 0 \\ 0 & N_{2,2} \end{bmatrix}$$

relative to this decomposition. A second routine calculation shows that $N_{1,1}$ is normal and $\sigma(N_{1,1}) \subseteq \sigma(N) = \{0\}$, from which we conclude that $\sigma(N_{1,1}) = \{0\}$.

Setting $d := \dim \mathfrak{M}$, we have that $N_{1,1} \in \mathcal{B}(\mathbb{C}^d)$, so that $N_{1,1}$ corresponds to a (normal) matrix. From above, we conclude that $N_{1,1} = 0$. But then for $0 \neq x \in \mathfrak{M}$, we obtain

$$0 = N_{1,1} x = N x = N^* N x,$$

contradicting the fact that $N^* N x = x$ for all $x \in \mathfrak{M}$.

This contradiction proves that $N = 0$.

$\square$

10.31. Let $(\mathfrak{X}, \|\cdot\|)$ be a complex Banach space and $T \in \mathcal{B}(\mathfrak{X})$. Recall that the **spectrum** of T is the set

$$\sigma(T) = \{\alpha \in \mathbb{K} : (T - \alpha I) \text{ is not invertible in } \mathcal{B}(\mathfrak{X})\}.$$

We do not yet have the machinery to be able to prove that $\sigma(T)$ is a *non-empty*, compact set contained in $\{z \in \mathbb{C} : |z| \leq \|T\|\}$. (The fact that the underlying field for $\mathfrak{X}$ is $\mathbb{C}$ is required to prove that $\sigma(T) \neq \varnothing$)

We denote the set of **eigenvalues** of T by $\sigma_p(T)$ and refer to this as the **point spectrum** of T. In general, $\sigma_p(T)$ may be empty or non-empty. Clearly $\sigma_p(T) \subseteq \sigma(T)$.

In Assignment 6, we saw that if T is a *compact* operator, then

$$\sigma(T) = \sigma_p(T) \cup \{0\},$$

and that for all $\varepsilon > 0$,

$$\sigma(T) \cap \{z \in \mathbb{C} : |z| > \varepsilon\}$$

is a finite set. We also saw in that Assignment that $\dim \ker (T - \alpha I) < \infty$ for all $\alpha \neq 0$. In other words, $\sigma(T) \setminus \{0\}$ is a sequence of eigenvalues of finite-multiplicity, and this sequence converges to 0 (in the case where it is infinite).

We shall exploit this wonderful behaviour of compact operators below.

10.32. Theorem. *Let $\mathcal{H}$ be a complex Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be a compact, normal operator. If $\sigma_p(N) = \{\alpha_n\}_{n\in\Omega}$, then $\mathcal{H} = \oplus_{n\in\Omega} \ker (N - \alpha_n I)$.*

Proof. The notation "Ω" has been introduced only so that we may deal with the cases where $\sigma_p(T)$ is finite and where it is infinite simultaneously.

Obviously $\ker (N-\alpha I) \subseteq \mathcal{H}$ for all $\alpha \in \mathbb{C}$, and more specifically for all $\alpha \in \sigma_p(N)$, and by Proposition 10.29, $\alpha \neq \beta \in \sigma_p(T)$ implies that $\ker (N - \alpha I) \perp \ker (N - \beta I)$. Thus

$$\mathfrak{M} := \oplus_{n\in\Omega} \ker (N - \alpha_n I) \subseteq \mathcal{H}.$$

Suppose that $\mathfrak{M} \neq \mathcal{H}$. Since each $\ker (N - \alpha I)$ is reducing for N, again by Proposition 10.29, we find that $\mathfrak{M}$ is reducing for N. Decomposing $\mathcal{H} = \mathfrak{M} \oplus \mathfrak{M}^\perp$, we may write

$$N = \begin{bmatrix} N_1 & 0 \\ 0 & N_4 \end{bmatrix}.$$

By Proposition 10.28, we see that N_1 and N_4 are compact and normal.

It follows that $\sigma(N_4) = \sigma_p(N_4) \cup \{0\}$. Note, however, that if $\alpha \in \sigma_p(N_4)$, then $\alpha \in \sigma_p(N)$, and therefore $\ker (N - \alpha I) \subseteq \mathfrak{M}$. This shows that $\sigma_p(N_4) = \varnothing$, from which we conclude that $\sigma(N_4) = \{0\}$.

By Proposition 10.30, $N_4 = 0$. The hypothesis that $\mathfrak{M} \neq \mathcal{H}$ then implies that $\mathfrak{M}^\perp \neq \{0\}$, which then implies that $0 \in \sigma_p(N_4)$, a contradiction.

Hence $\mathfrak{M} = \mathcal{H}$, completing the proof.

□

10.33. Theorem. The spectral theorem for compact normal operators. *Let $\mathcal{H}$ be a Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be a compact, normal operator. Suppose $\{\alpha_n\}_{n\in\Omega}$ are the distinct eigenvalues of N and that P_n is the orthogonal projection of $\mathcal{H}$ onto $\mathfrak{M}_n := \ker (N - \alpha_n I)$ for each $n \in \Omega$. Then $P_n P_m = 0$ if $n \neq m \in \Omega$, and*

$$N = \sum_{n\in\Omega} \alpha_n P_n,$$

where the series converges in the norm topology in $\mathcal{B}(\mathcal{H})$.

Proof. We remark that since Ω is either a finite or denumerable set, the above series is either just a finite sum (in which case there is no need to speak of convergence of the series), or we may as well assume that $\Omega = \mathbb{N}$, in which case the series may be written

$$N = \sum_{n=1}^{\infty} \alpha_n P_n.$$

That $n \neq m \in \Omega$ implies that $P_n P_m = 0$ is the statement that $\ker(N - \alpha_n I) \perp \ker(N - \alpha_m I)$ if $\alpha_n \neq \alpha_m$, which we proved in Proposition 10.29.

Recall from Proposition 10.29 that each $\mathfrak{M}_n$ is reducing for N and that by Theorem 10.32, $\mathcal{H} = \oplus_{n\in\Omega}\mathfrak{M}_n$. Thus

$$N = \oplus_{n\in\Omega} N_n,$$

where N_n is the compression of N to the reducing subspace $\mathfrak{M}_n$. But for $x \in \mathfrak{M}_n$, $(N - \alpha_n P_n)x = (N - \alpha_n I)x = 0$. Thus

$$N = \oplus_{n\in\Omega}\alpha_n P_n.$$

In the case where Ω is finite, this yields that $N = \sum_{n\in\Omega} \alpha_n P_n$, as there are no convergence issues.

Suppose that Ω is infinite.

Case One. $0 \notin \sigma_p(N)$.

Then $\dim \mathfrak{M}_n < \infty$ for all $n \in \Omega$, and $\mathcal{H} = \oplus_{n\in\Omega}\mathfrak{M}_n$ implies that $\mathcal{H}$ is separable. As we saw above, we may (and do) assume that $\Omega = \mathbb{N}$.

Choosing an orthonormal basis $\mathcal{B}_n = \{e_1^{(n)}, e_2^{(n)}, e_{k_n}^{(n)}\}$ for each $\mathfrak{M}_n$, $n \geq 1$, we find that $\mathcal{B} := \cup_{n=1}\mathcal{B}_n$ is an ONB for $\mathcal{H}$ and that relative to this basis we have

$$N = \mathrm{diag}(\alpha_1, \alpha_1, \ldots, \alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_2, \alpha_3, \alpha_3, \ldots),$$

where α_n is repeated exactly $k_n = \dim \mathfrak{M}_n = \mathrm{rank}\, P_n$ times.

Let $Q_M := \sum_{n=1}^{M} P_n$, so that is a finite-rank projection for each $M \geq 1$. Note that $(Q_M)_{M=1}^{\infty}$ converges in the SOT to $I \in \mathcal{B}(\mathcal{H})$. By Remark 10.23,

$$N = \lim_{M\to\infty} Q_M N Q_M = \lim_{M\to\infty} \sum_{n=1}^{M} \alpha_n P_n = \sum_n \alpha_n P_n.$$

Case Two. $0 \in \sigma_p(N)$.

Again, we may assume that $\Omega = \mathbb{N}$ and that $\alpha_1 = 0$. Since $\mathfrak{M}_0 = \ker N$ is reducing for N, we may write

$$N = \begin{bmatrix} N_1 & 0 \\ 0 & N' \end{bmatrix}$$

relative to the decomposition $\mathcal{H} = \mathfrak{M}_1 \oplus \mathfrak{M}_1^{\perp}$. Note that N' is compact, normal and $0 \notin \sigma_p(N')$. By Case One,

$$N' = \sum_{n=2}^{\infty} \alpha_n P_n.$$

But then

$$N = 0P_1 + \sum_{n=2}^{\infty} \alpha_n P_n = \sum_{n=1}^{\infty} \alpha_n P_n.$$

□

10.34. Corollary. *Let $\mathcal{H}$ be a Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be a compact, normal operator. Then there exists an orthonormal basis $\{e_\alpha\}_{\alpha \in \Lambda}$ for $\mathcal{H}$ such that each e_α is an eigenvector for N.*

Proof. Let $\{\lambda_n\}_{n=1}^{\infty}$ be the set of eigenvalues of N. (We argue the case where N has infinitely many eigenvalues and leave the case where $\sigma(N)$ is finite to the reader.) For each $n \geq 1$, choose an orthonormal basis $\{e_{(n,\beta)}\}_{\beta \in \Lambda_n}$ for $\ker(N - \lambda_n I)$. (Note that if $\lambda_n \neq 0$, then the cardinality of Λ_n is finite.) Then each $e_{(n,\beta)}$, $\beta \in \Lambda_n$, $n \geq 1$ is an eigenvector for N corresponding to λ_n, the $e_{(n,\beta)}$'s are all orthogonal since all of the $\ker(N - \lambda_n I)$'s are. Finally, $\overline{\text{span}}\{e_{(n,\beta)}\}_{\beta \in \Lambda_n, n \geq 1} = \oplus_{n=1}^{\infty} \ker(N - \lambda_n I) = \mathcal{H}$ by Theorem 10.32. Let $\{e_\alpha\}_{\alpha \in \Lambda} = \{e_{(n,\beta)}\}_{\beta \in \Lambda_n, n \geq 1}$.

□

Appendix to Section 10.

10.35. Proposition 10.30 is actually true in much greater generality. If $\mathcal{A}$ is any unital (complex) Banach algebra and $a \in \mathcal{A}$, we define the **spectral radius** of a to be

$$\mathrm{spr}(a) := \sup\{|\alpha| : \alpha \in \sigma(a)\}.$$

It can be shown that $N \in \mathcal{B}(\mathcal{H})$ normal implies that $\|N\| = \mathrm{spr}(N)$. The usual proof of this fact requires the following result.

10.36. Theorem. Beurling's Spectral Radius Formula. *If $\mathcal{A}$ is a unital Banach algebra and $a \in \mathcal{A}$, then*

$$\mathrm{spr}(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}}.$$

The proof of this result is beyond (but not much beyond) the scope of this course. (We simply didn't have the time for it.) Nevertheless, two things are worth pointing out.

Firstly, the above limit exists!! This is anything but obvious, and is interesting in its own right.

Secondly, since $\|a^n\| \le \|a\|^n$ for all $n \ge 1$, we immediately see that

$$\mathrm{spr}(a) \le \|a\|.$$

This shows that $\sigma(a)$ is always bounded. Another non-trivial fact is that $\sigma(a) \ne 0$. This relies on a Banach-space version of **Liouville's Theorem**.

Using Beurling's Spectral Radius Formula, we may easily obtain:

10.37. Proposition. *Let $\mathcal{H}$ be a complex Hilbert space and $N \in \mathcal{B}(\mathcal{H})$ be a normal operator. Then* spr $(N) = \|N\|$.

In particular, $\sigma(N) = \{0\}$ if and only if $N = 0$.

Proof. Consider first:

$$\begin{aligned}
\|N^2\| &= \sup_{\|x\|=1} \|N^2 x\| \\
&= \sup_{\|x\|=1} \|N^* N x\| \\
&\ge \sup_{\|x\|=1} |(N^* N x, x)| \\
&= \sup_{\|x\|=1} (Nx, Nx) \\
&= \sup_{\|x\|=1} \|Nx\|^2 \\
&= \|N\|^2.
\end{aligned}$$

By induction, $\|N^{2^n}\| \ge \|N\|^{2^n}$ for all $n \ge 1$. The reverse inequality follows immediately from the submultiplicativity of the norm in a Banach algebra. Thus

$\|N^{2^n}\| = \|N\|^{2^n}$ for all $n \geq 1$. By Beurling's Spectral Radius Formula, Theorem 10.36,

$$\text{spr}\,(N) = \lim_{n\to\infty} \|N^{2^n}\|^{1/2^n} = \|N\|.$$

□

This is the standard proof of this result. The proof of Proposition 10.30 given in the notes above is original, and clearly was only devised to circumvent the fact that we do not have Beurling's Spectral Radius Formula.

10.38. Definition. *Let $\mathcal{A}$ be a Banach algebra and $a \in \mathcal{A}$. If $\mathcal{A}$ is unital, then the **spectrum** of a relative to $\mathcal{A}$ is the set*

$$\sigma_{\mathcal{A}}(a) = \{\lambda \in \mathbb{C} : a - \lambda 1 \text{ is not invertible in } \mathcal{A}\}.$$

*If $\mathcal{A}$ is not unital, then $\sigma_{\mathcal{A}}(a)$ is set to be $\sigma_{\mathcal{A}^+}(a) \cup \{0\}$. When the algebra $\mathcal{A}$ is understood, we generally write $\sigma(a)$. The **resolvent** of a is the set $\rho(a) = \mathbb{C}\backslash\sigma(a)$.*

10.39. Corollary. *Let $\mathcal{A}$ be a unital Banach algebra, and let $a \in \mathcal{A}$. Then $\rho(a)$ is open and $\sigma(a)$ is compact.*

Proof. Clearly $\rho(a) = \{\lambda \in \mathbb{C} : (a - \lambda 1) \text{ is invertible}\}$ is open, since $\mathcal{A}^{-1}$ is. Indeed, if $a - \lambda_0 1$ is invertible in $\mathcal{A}$, then $\lambda \in \rho(a)$ for all $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < \|(a - \lambda_0)^{-1}\|^{-1}$. Thus $\sigma(a)$ is closed.

If $|\lambda| > \|a\|$, then $\lambda 1 - a = \lambda\,(1 - \lambda^{-1}a)$ and $\|\lambda^{-1}a\| < 1$, and so $(1 - \lambda^{-1}a)$ is invertible. This implies

$$(\lambda 1 - a)^{-1} = \lambda^{-1}\,(1 - \lambda^{-1}a)^{-1}.$$

Thus $\sigma(a)$ is contained in the disk $D_{\|a\|}(\{0\})$ of radius $\|a\|$ centred at the origin. Since it both closed and bounded, $\sigma(a)$ is compact.

□

10.40. Definition. *Let $\mathfrak{X}$ be a Banach space and $U \subseteq \mathbb{C}$ be an open set. Then a function $f : U \to \mathfrak{X}$ is said to be **weakly analytic** if the map $z \mapsto x^*(f(z))$ is analytic for all $x^* \in \mathfrak{X}^*$.*

10.41. Theorem. [**Liouville's Theorem**] *Every bounded, weakly entire function into a Banach space $\mathfrak{X}$ is constant.*

Proof. For each linear functional $x^* \in \mathfrak{X}^*$, $x^* \circ f$ is a bounded, entire function into the complex plane. By the complex-valued version of Liouville's Theorem, it must therefore be constant. Now by the Hahn-Banach Theorem, $\mathfrak{X}^*$ separates the points of $\mathfrak{X}$. So if there exist $z_1, z_2 \in \mathbb{C}$ such that $f(z_1) \neq f(z_2)$, then there must exist $x^* \in \mathfrak{X}^*$ such that $x^*(f(z_1)) \neq x^*(f(z_2))$. This contradiction implies that f is constant.

□

10.42. Definition. *Let $\mathcal{A}$ be a unital Banach algebra and let $a \in \mathcal{A}$. The map*

$$\begin{array}{rccc} R(\cdot, a): & \rho(a) & \to & \mathcal{A} \\ & \lambda & \mapsto & (\lambda 1 - a)^{-1} \end{array}$$

is called the ***resolvent function*** *of a.*

10.43. Proposition. The Common Denominator Formula. *Let $a \in \mathcal{A}$, a unital Banach algebra. Then if μ, $\lambda \in \rho(a)$, we have*

$$R(\lambda, a) - R(\mu, a) = (\mu - \lambda)\, R(\lambda, a)\, R(\mu, a).$$

Proof. The proof is transparent if we consider $t \in \mathbb{C}$ and consider the corresponding complex-valued equation:

$$\frac{1}{\lambda - t} - \frac{1}{\mu - t} = \frac{(\mu - t) - (\lambda - t)}{(\lambda - t)\,(\mu - t)} = \frac{(\mu - \lambda)}{(\lambda - t)\,(\mu - t)}.$$

In terms of Banach algebra, we have:

$$\begin{aligned} R(\lambda, a) &= R(\lambda, a)\, R(\mu, a)\, (\mu - a) \\ R(\mu, a) &= R(\mu, a)\, R(\lambda, a)\, (\lambda - a). \end{aligned}$$

Noting that $R(\lambda, a)$ and $R(\mu, a)$ clearly commute, we obtain the desired equation by simply subtracting the second equation from the first.

□

We shall return to this formula when establishing the holomorphic functional calculus in the next section.

10.44. Proposition. *If $a \in \mathcal{A}$, a unital Banach algebra, then $R(\cdot, a)$ is analytic on $\rho(a)$.*
Proof. Let $\lambda_0 \in \rho(a)$. Then

$$\begin{aligned} \lim_{\lambda \to \lambda_0} \frac{R(\lambda, a) - R(\lambda_0, a)}{\lambda - \lambda_0} &= \lim_{\lambda \to \lambda_0} \frac{(\lambda_0 - \lambda)\, R(\lambda, a)\, R(\lambda_0, a)}{\lambda - \lambda_0} \\ &= -R(\lambda_0, a)^2 \end{aligned}$$

since inversion is continuous on $\rho(a)$. Thus the limit of the Newton quotient exists, and so $R(\cdot, a)$ is analytic.

□

10.45. Corollary. [Gelfand] *If $a \in \mathcal{A}$, a Banach algebra, then $\sigma(a)$ is non-empty.*
Proof. We may assume that $\mathcal{A}$ is unital, for otherwise $0 \in \sigma(a)$ and we are done. Similarly, if $a = 0$, then $0 \in \sigma(a)$. If $\rho(a) = \mathbb{C}$, then clearly $R(\cdot, a)$ is entire. Now

for $|\lambda| > \|a\|$, we have

$$\begin{aligned}(\lambda - a)^{-1} &= \big(\lambda(1-\lambda^{-1}a)\big)^{-1} \\ &= \lambda^{-1}\sum_{n=0}^{\infty}(\lambda^{-1}a)^n \\ &= \sum_{n=0}^{\infty}\lambda^{-n-1}a^n\end{aligned}$$

so that if $|\lambda| \ge 2\|a\|$, then

$$\|(\lambda - a)^{-1}\| \le \sum_{n=0}^{\infty}\frac{\|a\|^n}{(2\|a\|)^{n+1}} \le \frac{1}{\|a\|}.$$

That is, $\|R(\lambda, a)\| \le \|a\|^{-1}$ for all $\lambda \ge 2\|a\|$.

Clearly there exists $M < \infty$ such that

$$\max_{|\lambda| \le 2\|a\|}\|R(\lambda, a)\| \le M$$

since $R(\cdot, a)$ is a continuous function on this compact set. The conclusion is that $R(\cdot, a)$ is a bounded, entire function. By Theorem 10.41, the resolvent function must be constant. This obvious contradiction implies that $\sigma(a)$ is non-empty.

□

Recall that a ***division algebra*** is an algebra in which each non-zero element is invertible.

10.46. Theorem. [Gelfand-Mazur] *If $\mathcal{A}$ is a Banach algebra and a division algebra, then there is a unique isometric isomorphism of $\mathcal{A}$ onto $\mathbb{C}$.*

Proof. If $b \in \mathcal{A}$, then $\sigma(b)$ is non-empty by Corollary 10.45. Let $\beta \in \sigma(b)$. Then $\beta 1 - b$ is not invertible, and since $\mathcal{A}$ is a division algebra, we conclude that $\beta 1 = b$; that is to say, that $\sigma(b)$ is a singleton.

Given $a \in \mathcal{A}$, $\sigma(a)$ is a singleton, say $\{\lambda_a\}$. The complex-valued map $\phi : a \mapsto \lambda_a$ is an algebra isomorphism. Moreover, $\|a\| = \|\lambda_a 1\| = |\lambda_a| = \|\phi(a)\|$, so the map is isometric as well.

If $\phi_0 : \mathcal{A} \to \mathbb{C}$ were another such map, then $\phi_0(a) \in \sigma(a)$, implying that $\phi_0(a) = \phi(a)$.

□

10.47. Definition. *Let $a \in \mathcal{A}$, a Banach algebra. The* ***spectral radius*** *of a is*

$$\operatorname{spr}(a) = \sup\{|\lambda| : \lambda \in \sigma(a)\}.$$

10.48. Lemma. The Spectral Mapping Theorem - polynomial version. *Let $a \in \mathcal{A}$, a unital Banach algebra, and suppose $p \in \mathbb{C}[z]$ is a polynomial. Then*

$$\sigma(p(a)) = p(\sigma(a)) := \{p(\lambda) : \lambda \in \sigma(a)\}.$$

Proof. Let $\alpha \in \mathbb{C}$. Then for some $\gamma \in \mathbb{C}$,

$$p(z) - \alpha = \gamma\,(z - \beta_1)\,(z - \beta_2) \cdots (z - \beta_n)$$

and so

$$p(a) - \alpha = \gamma\,(a - \beta_1)\,(a - \beta_2) \cdots (a - \beta_n).$$

Thus (as all of the terms $(a - \beta_i)$ commute),

$$\begin{aligned} \alpha \in \sigma(p(a)) \quad &\Longleftrightarrow \quad \beta_i \in \sigma(a) \text{ for some } 1 \leq i \leq n \\ &\Longleftrightarrow \quad p(z) - \alpha = 0 \text{ for some } z \in \sigma(a) \\ &\Longleftrightarrow \quad \alpha \in p(\sigma(a)). \end{aligned}$$

□

10.49. Theorem. [Beurling : The Spectral Radius Formula] *If $a \in \mathcal{A}$, a Banach algebra, then*

$$\mathrm{spr}(a) = \lim_{n \to \infty} \|a^n\|^{1/n}.$$

Proof. First observe that if $\mathcal{A}$ is not unital, then we can always embed it isometrically into a unital Banach algebra $\mathcal{A}^+$. Since both the left and right hand sides of the above equation remain unchanged when a is considered as an element of $\mathcal{A}^+$, we may (and do) assume that $\mathcal{A}$ is already unital.

Now $\sigma(a^n) = (\sigma(a))^n$, and so $\mathrm{spr}(a^n) = (\mathrm{spr}(a))^n$. Moreover, for all $b \in \mathcal{A}$, the proof of Corollary 10.39 shows that $\mathrm{spr}(b) \leq \|b\|$. Thus

$$\mathrm{spr}(a) = (\mathrm{spr}(a^n))^{1/n} \leq \|a^n\|^{1/n} \text{ for all } n \geq 1.$$

This tells us that $\mathrm{spr}(a) \leq \inf_{n \geq 1} \|a^n\|^{1/n}$.

On the other hand, $R(\cdot, a)$ is analytic on $\rho(a)$ and hence is analytic on $\{\lambda \in \mathbb{C} : |\lambda| > \mathrm{spr}(a)\}$. Furthermore, if $|\lambda| > \|a\|$, then

$$\begin{aligned} R(\lambda, a) &= (\lambda - a)^{-1} \\ &= \lambda^{-1}(1 - \lambda^{-1}a)^{-1} \\ &= \sum_{n=0}^{\infty} a^n / \lambda^{n+1}. \end{aligned}$$

Let $\phi \in \mathcal{A}^*$. Then $\phi \circ R(\cdot, a)$ is an analytic, complex-valued function,

$$[\phi \circ R(\cdot, a)](\lambda) = \sum_{n=0}^{\infty} \phi(a^n)/\lambda^{n+1}$$

and this Laurent expansion is still valid for $\{\lambda \in \mathbb{C} : |\lambda| > \|a\|\}$, since the series for $R(\cdot, a)$ is absolutely convergent on this set, and applying ϕ introduces at most a factor of $\|\phi\|$ to the absolutely convergent sum. Since $[\phi \circ R(\cdot, a)]$ is analytic on $\{\lambda \in \mathbb{C} : |\lambda| > \mathrm{spr}\,(a)\}$, the complex-valued series converges on this larger set.

From this it follows that the sequence $\{\phi(a^n)/\lambda^{n+1}\}_{n=1}^\infty$ converges to 0 as n tends to infinity for all $\phi \in \mathcal{A}^*$, so therefore is bounded for all $\phi \in \mathcal{A}^*$. It is now a consequence of the Uniform Boundedness Principle that $\{a^n/\lambda^{n+1}\}_{n=1}^\infty$ is bounded in norm, say by $M_\lambda > 0$, *for each* $\lambda \in \mathbb{C}$ satisfying $|\lambda| > \mathrm{spr}(a)$. That is:

$$\|a^n\| \le M_\lambda |\lambda^{n+1}|$$

for all $|\lambda| > \mathrm{spr}(a)$. But then, for all $|\lambda| > \mathrm{spr}(a)$,

$$\limsup_{n \ge 1} \|a^n\|^{1/n} \le \limsup_{n \ge 1} M_\lambda^{1/n} |\lambda^{n+1/n}| = |\lambda|.$$

Combining this estimate with the above yields $\mathrm{spr}(a) = \lim_{n\to\infty} \|a^n\|^{1/n}$.

□

They laughed when I said I was going to be a comedian. They're not laughing now.

Bob Monkhouse

Exercises for Section 10.

Question 1. Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces, and let $S, T \in \mathcal{B}(\mathfrak{X}, \mathfrak{Y})$. Prove that for all $k_1, k_2 \in \mathbb{K}$,

$$(k_1 S + k_2 T)^* = k_1 S^* + k_2 T^*.$$

Here, the *-map refers to the Banach space adjoint.

Question 2.

Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{M}$ be a closed subspace of $\mathcal{H}$. Let P be the orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$, and $T \in \mathcal{B}(\mathcal{H})$.

(a) Prove that relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, we have

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

(b) Prove that $P = P^2 = P^*$.

(c) More generally, write $T = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}$ relative to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Prove that

$$T^* = \begin{bmatrix} T_1^* & T_3^* \\ T_2^* & T_4^* \end{bmatrix}.$$

(d) Prove that $\mathcal{M}$ is invariant for T if and only if $(I - P)TP = 0$, and $\mathcal{M}$ is reducing for P if and only if $TP = PT$. Conclude that $\mathcal{M}$ is reducing for T if and only if both $\mathcal{M}$ and $\mathcal{M}^\perp$ are invariant for T.

Question 3.

This problem is more challenging than the others, and should rightfully be an assignment question.

Let $\mathcal{H}$ be a complex, separable Hilbert space and suppose that $K \in \mathcal{K}(\mathcal{H})$. Note that K^*K is compact and selfadjoint, hence normal. By the Spectral Theorem for Normal operators (see Corollary 10.34), there exists an ONB $\{e_n\}_{n=1}^\infty$ for $\mathcal{H}$ such that

$$K^*K = \operatorname{diag}(d_1, d_2, d_3, \ldots).$$

(a) Prove that $d_n \ge 0$ for all $n \ge 1$.

We denote by $|K|$ the diagonal operator (relative to this ONB)

$$|K| = \operatorname{diag}(s_1, s_2, s_3, \ldots),$$

where $s_n = (d_n)^{\frac{1}{2}}$, $n \ge 1$. Note that in the same way that for complex numbers we have that $|z|^2 = \overline{z}z$, we now have that $|K|^2 = K^*K$.

The sequence $(s_n)_n$ is called the sequence of **singular numbers** of K. For $1 \le p \le \infty$, we define the **Schatten p-class** $\mathcal{C}_p$ of all compact operators K such that $(s_n)_n \in \ell^p$.

It can be shown that $\mathcal{C}_p$ is a vector space, that $\|K\|_p := \|(s_n)_n\|_p$ defines a norm on $\mathcal{C}_p$, and that $(\mathcal{C}_p, \|\cdot\|_p)$ is complete, and thus a Banach space.

Let $\{f_n\}_n$ denote a second ONB for $\mathcal{H}$. Recall that for $x, y \in \mathcal{H}$, we defined the rank-one operator $x \otimes y^* \in \mathcal{B}(\mathcal{H})$ via $x \otimes y^*(z) = \langle z, y \rangle x$, $z \in \mathcal{H}$.

(b) Prove that $\sum_n s_n \, f_n \otimes e_n^*$ converges in norm in $\mathcal{B}(\mathcal{H})$.

(c) Prove that if $K \in \mathcal{K}(\mathcal{H})$ has singular numbers $(s_n)_n$, then

$$s_n = \operatorname{dist}(K, \mathcal{F}_{n-1}),$$

where $\mathcal{F}_{n-1} = \{F \in \mathcal{B}(\mathcal{H}) : \operatorname{rank} F \leq n-1\}$. That is,

$$s_n = \inf \{\|K - F\| : \operatorname{rank} F \leq n-1\}.$$

11. Appendix – topological background

A child of five could understand this. Fetch me a child of five.

Groucho Marx

11.1. At the heart of analysis is topology. A thorough study of topology is beyond the scope of this course, and we refer the reader to the excellent book [**Wil70**] *General Topology*, written by my former colleague Stephen Willard. The treatment of topology in this section borrows heavily from his book.

We shall only give the briefest of overviews of this theory - assuming that the student has some background in metric and norm topologies. We shall only cover the notions of weak topologies and nets, which are vital to the study of Functional Analysis.

11.2. Definition. *A **topology** τ on a set X is a collection of subsets of X, called **open sets**, which satisfy the following:*

(i) *$X, \varnothing \in \tau$ - i.e. the entire space and the empty set are open;*
(ii) *If $\{G_\alpha\}_\alpha \subseteq \tau$, then $\cup_\alpha G_\alpha \in \tau$ - i.e. arbitrary unions of open sets are open;*
(iii) *If $n \geq 1$ and $\{G_k\}_{k=1}^n \subseteq \tau$, then $\cap_{k=1}^n G_k \in \tau$ - i.e.* finite *intersections of open sets are open*

*A set F is called **closed** if $X \backslash F$ is open. We call (X, τ) (or more informally, we call X) a topological space.*

It is useful to observe that the intersection of a collection $\{\tau_\alpha\}_\alpha$ of topologies on X is once again a topology on X.

11.3. Example.

(i) Let X be any set. Then $\tau = \{\varnothing, X\}$ is a topology on X, called the **trivial topology** on X.
(ii) At the other extreme of the topological spectrum, if X is any non-empty set, then $\tau = \mathcal{P}(X)$, the power set of X, is a topology on X, called the **discrete topology** on X.
(iii) Let $X = \{a, b\}$, and set $\tau = \{\varnothing, \{a\}, \{a, b\}\}$. Then τ is a topology on X.
(iv) Let (X, d) be a metric space. Let

$$\tau = \{G \subseteq X : \text{ for all } g \in G \text{ there exists } \delta > 0 \text{ such that}$$
$$b_\delta(g) := \{y \in X : d(x, y) < \delta\} \subseteq G\}.$$

Then τ is a topology, called the **metric topology** on X induced by d. This is the usual topology one thinks of when dealing with metric spaces, but as we shall see, there can be many more.

(v) Let X be any non-empty set. Then

$$\tau_{cf} = \{\varnothing\} \cup \{Y \subseteq X : X \backslash Y \text{ is finite}\}$$

is a topology on X, called the **co-finite topology** on X.

11.4. Definition. *Let (X, τ) be a topological space, and $x \in X$. A set U is called a **neighbourhood** (abbreviated* nbhd*) of x if there exists $G \in \tau$ so that $x \in G \subseteq U$. The reader is cautioned that some authors require nbhds to be open* - we do not. *The **neighbourhood system at** x is $\mathcal{U}_x := \{U \subseteq X : U \text{ is a nbhd of } x\}$.*

The following result from [**Wil70**] illustrates the importance of nbhd systems.

11.5. Theorem. *Let (X, τ) be a topological space, and $x \in X$. Then:*

(a) *If $U \in \mathcal{U}_x$, then $x \in U$.*
(b) *If $U, V \in \mathcal{U}_x$, then $U \cap V \in \mathcal{U}_x$.*
(c) *If $U \in \mathcal{U}_x$, there exists $V \in \mathcal{U}_x$ such that $U \in \mathcal{U}_y$ for each $y \in V$.*
(d) *If $U \in \mathcal{U}_x$ and $U \subseteq V$, then $V \in \mathcal{U}_x$.*
(e) *$G \subseteq X$ is open if and only if G contains a nbhd of each of its points.*

Conversely, if in a set X a non-empty collections $\mathcal{U}_x$ of subsets of X is assigned to each $x \in X$ so as to satisfy conditions (a) through (d), and if we use (e) to define *the notion of an open set, the result is a topology on X in which the nbhd system at x is precisely $\mathcal{U}_x$.*

Because of this, it is clear that if we know the nbhd system of each point in X, then we know the topology of X.

There are a number of natural separation axioms that a topological space might satisfy.

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

(i) *(X, τ) is said to be $\mathbf{T_0}$ if for every $x,\ y \in X$ such that $x \neq y$, either there is a neighbourhood U_x of x with $y \notin U_x$ or there is a neighbourhood U_y of y with $x \notin U_y$.*
(ii) *(X, τ) is said to be $\mathbf{T_1}$ if for every $x,\ y \in X$ such that $x \neq y$, there are neighbourhoods U_x of x and U_y of y with $y \notin U_x$ and $x \notin U_y$.*
(iii) *(X, τ) is said to be $\mathbf{T_2}$ (or **Hausdorff**) if for every $x,\ y \in X$ such that $x \neq y$, there are neighbourhoods U_x of x and U_y of y with $U_x \cap U_y = \varnothing$.*

*We say that two subsets A and B of X can be **separated** by τ if there exist $U,\ V \in \tau$ with $A \subseteq U,\ B \subseteq V$ and $U \cap V = \varnothing$.*

(iv) *(X, τ) is said to be **regular** if whenever $F \subseteq X$ is closed and $x \notin F$, F and $\{x\}$ can be separated.*
(v) *(X, τ) is said to be **normal** if whenever $F_1,\ F_2 \subset X$ are closed and disjoint, then F_1 and F_2 can be separated.*
(vi) *(X, τ) is said to be $\mathbf{T_3}$ if it is T_1 and regular.*
(vii) *(X, τ) is said to be $\mathbf{T_4}$ if it is T_1 and normal.*

We are assuming that the next definition is a familiar one.

11.7. Definition. *Let (X, τ) be a topological space. An* ***open cover*** *of X is a collection $\mathcal{G} \subseteq \tau$ such that $X = \cup_{G \in \mathcal{G}} G$. A* ***finite subcover*** *of X relative to $\mathcal{G}$ is a finite subset $\{G_1, G_2, ..., G_n\} \subseteq \mathcal{G}$ which is again an open cover of X.*

A topological space (X, τ) is said to be ***compact*** *if every open cover of X admits a finite subcover.*

11.8. Theorem. *Let (X, d) be a metric space. Then X, equipped with the metric topology, is T_4.*

11.9. Theorem. *Let (X, τ) be a compact, Hausdorff space. Then (X, τ) is T_4.*

Recall that a topological space (X, τ) is said to be **separable** if it admits a countable dense subset.

The following result will be needed in Section 7.

11.10. Proposition. *Let (X, d) be a compact metric space. Then (X, d) is separable.*

Proof. For each $n \geq 1$, the collection $\mathcal{G}_n := \{b_{1/n}(x) : x \in X\}$ is an open cover of X. Since X is compact, we can find a finite subcover $\{b_{1/n}(x_{(j,n)}) : 1 \leq j \leq k_n\}$ of X. It is then clear that if $x \in X$, there exists $1 \leq j \leq k_n$ so that $d(x, x_{(j,n)}) < 1/n$.

As such, the collection

$$\mathcal{D} := \{x_{(j,n)} : 1 \leq j \leq k_n, 1 \leq n\}$$

is a countable, dense set in X, proving that (X, d) is separable.

□

11.11. Definition. *Let (X, τ) be a topological space. A* ***neighbourhood base*** *$\mathcal{B}_x$ at a point $x \in X$ is a collection $\mathcal{B}_x \subseteq \mathcal{U}_x$ so that $U \in \mathcal{U}_x$ implies that there exists $B \in \mathcal{B}_x$ so that $B \subseteq U$. We refer to the elements of $\mathcal{B}_x$ as* ***basic nbhds*** *of the point x.*

The importance of neighbourhood bases is that all open sets can be constructed from them, as we shall soon see.

11.12. Example. Consider (X, d) be a metric space equipped with the metric topology τ. For each $x \in X$, fix a sequence $\{r_n(x)\}_{n=1}^{\infty}$ of positive real numbers such that $\lim_{n\to\infty} r_n(x) = 0$ and consider $\mathcal{B}_x = \{V_{r_n(x)}(x) : n \geq 1\}$. (We remind the reader that for all $\varepsilon > 0$, $V_\varepsilon(x) := \{y \in X : d(x, y) < \varepsilon\}$.) Then $\mathcal{B}_x$ is a nbhd base at x for each $x \in X$.

11.13. Definition. *Let (X, τ) be a topological space. A* ***base*** *for the topology is a collection $\mathcal{B} \subseteq \tau$ so that for every $G \in \tau$ there exists $\mathcal{C} \subseteq \mathcal{B}$ so that $G = \cup\{B : B \in \mathcal{C}\}$. That is, every open set is a union of elements of $\mathcal{B}$. Note that if $\mathcal{C}$ is empty, then $\cup\{B : B \in \mathcal{C}\}$ is also empty, so we do not need to include the empty set in our base. A* ***subbase*** *for the topology is a collection $\mathcal{S} \subseteq \tau$ such that the collection $\mathcal{B}$ of all finite intersections of elements of $\mathcal{S}$ forms a base for τ.*

As we shall see in the Assignments, any collection $\mathcal{C}$ of subsets of X serves as a subbase for *some* topology on X, called **the topology generated by** $\mathcal{C}$.

11.14. Example. Let (X, τ) be a topological space, and for each $x \in X$, suppose that $\mathcal{B}_x$ is a neighbourhood base at x consisting of *open* sets. Then $\mathcal{B} := \cup_{x \in X} \mathcal{B}_x$ is a base for the topology τ on X.

11.15. Example. Consider $\mathbb{R}$ with the usual topology τ. The collection $\mathcal{B} = \{(a, b) : a, b \in \mathbb{R}, a < b\}$ is a base for τ. (You might remember from Real Analysis that every open set in $\mathbb{R}$ is a *disjoint* union of open intervals - although the fact the union is disjoint in this setting is a luxury item which we have not built into the definition of a base in general.)

The collection $\mathcal{S} = \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, \infty) : b \in \mathbb{R}\}$ is a subbase for the usual topology, but is not a base for τ.

11.16. Definition. *Let (X, τ) be a topological space. A* ***directed set*** *is a set Λ with a relation $\leq$ that satisfies:*

(i) *$\lambda \leq \lambda$ for all $\lambda \in \Lambda$;*
(ii) *if $\lambda_1 \leq \lambda_2$ and $\lambda_2 \leq \lambda_3$, then $\lambda_1 \leq \lambda_3$; and*
(iii) *if $\lambda_1, \lambda_2 \in \Lambda$, then there exists λ_3 so that $\lambda_1 \leq \lambda_3$ and $\lambda_2 \leq \lambda_3$.*

The relation $\leq$ is sometimes called a ***direction*** *on Λ.*

A ***net*** *in X is a function $P : \Lambda \to X$, where Λ is a directed set. The point $P(\lambda)$ is usually denoted by x_λ, and we often write $(x_\lambda)_{\lambda \in \Lambda}$ to denote the net.*

A ***subnet*** *of a net $P : \Lambda \to X$ is the composition $P \circ \varphi$, where $\varphi : M \to \Lambda$ is an increasing* ***cofinal*** *function from a directed set to Λ; that is,*

(a) *$\varphi(\mu_1) \leq \varphi(\mu_2)$ if $\mu_1 \leq \mu_2$ (**increasing**), and*
(b) *for each $\lambda \in \Lambda$, there exists $\mu \in M$ so that $\lambda \leq \varphi(\mu)$ (**cofinal**).*

For $\mu \in M$, we often write x_{λ_μ} for $P \circ \varphi(\mu)$, and speak of the subnet $(x_{\lambda_\mu})_\mu$.

11.17. Definition. *Let (X, τ) be a topological space. The net $(x_\lambda)_\lambda$ is said to* ***converge to*** *$x \in X$ if for every $U \in \mathcal{U}_x$ there exists $\lambda_0 \in \Lambda$ so that $\lambda \geq \lambda_0$ implies $x_\lambda \in U$.*

We write $\lim_\lambda x_\lambda = x$, or $\lim_{\lambda \in \Lambda} x_\lambda = x$.

This mimics the definition of convergence of a sequence in a metric space.

11.18. Example.

(a) Since $\mathbb{N}$ is a directed set under the usual order $\leq$, every sequence is a net. Any subsequence of a sequence is also a subnet. The converse to this is false, however. A subnet of a sequence need not be a subsequence, since its domain need not be $\mathbb{N}$ (or any countable set, for that matter).

(b) Let A be a non-empty set and Λ denote the power set of all subsets of A, partially ordered with respect to inclusion. Then Λ is a directed set, and any function from Λ to $\mathbb{R}$ is a net in $\mathbb{R}$.

(c) Let $\mathcal{P}$ denote the set of all finite partitions of $[0,1]$, partially ordered by inclusion (i.e. refinement). Let f be a continuous function on $[0,1]$; then to $P = \{0 = t_0 < t_1 < \cdots < t_n = 1\} \in \mathcal{P}$, we associate the quantity $L_P(f) = \sum_{i=1}^n f(t_{i-1})(t_i - t_{i-1})$. The map $P \mapsto L_P(f)$ is a net ($\mathcal{P}$ is a directed set), and from Calculus, $\lim_{P\in\mathcal{P}} L_P(f) = \int_0^1 f(x)dx$.

11.19. Example. Let (X, τ_X) be a topological space and $x \in X$. Let $\mathcal{U}_x$ denote the nbhd system at x. If, for $U_1, U_2 \in \mathcal{U}_x$ we define the relation $U_1 \leq U_2$ if $U_2 \subseteq U_1$, then $(\mathcal{U}_x, \leq)$ forms a directed set.

For each $U \in \mathcal{U}_x$, choose $x_U \in U$. Then $(x_U)_{U\in\mathcal{U}_x}$ forms a net in X. It is not hard to see that $\lim_{U\in\mathcal{U}_x} x_U = x$. Indeed, given $V \in \mathcal{U}_x$, we have that $x_U \in V$ for all $U \geq V$.

Observe that if (X, τ_X) is not Hausdorff, it is entirely possible that there exists $y \neq x$ in X so that $y = \lim_{U\in\mathcal{U}_x} x_U$ as well. (You should convince yourself of this by producing an example.) The property that (X, τ_X) is Hausdorff is equivalent to the condition that that limits of nets in X are unique.

11.20. Definition.
Let (X, τ_X) and (Y, τ_Y) be topological spaces. We say that a function $f : X \to Y$ is continuous if $f^{-1}(G)$ is open in X for all $G \in \tau_Y$.

That this extends our usual notion of continuity for functions between metric space is made clear by the following result:

11.21. Proposition. *If (X, d_X) and (Y, d_Y) are metric spaces with metric space topologies τ_X and τ_Y respectively, then the following are equivalent for a function $f : X \to Y$:*

(a) *f is continuous on X, i.e. $f^{-1}(G) \in \tau_X$ for all $G \in \tau_Y$.*

(b) *$\lim_n f(x_n) = f(x)$ whenever $(x_n)_{n=1}^\infty$ is a sequence in X converging to $x \in X$.*

As we shall see in the Assignments, sequences are not enough to describe convergence, nor are they enough to characterize continuity of functions between general topological spaces. On the other hand, nets are sufficient for this task, and serve as the natural replacement for sequences. (The following result also admits a *local* version, which we shall also see in the Assignments.)

11.22. Theorem. *Let (X, τ_X) and (Y, τ_Y) be topological spaces. Let $f : X \to Y$ be a function. The following are equivalent:*

(a) *f is continuous on X.*
(b) *Whenever $(x_\lambda)_{\lambda \in \Lambda}$ is a net in X which converges to $x \in X$, it follows that $(f(x_\lambda))_{\lambda\in\Lambda}$ is a net in Y which converges to $f(x)$.*

The notion of a **weak topology** on a set X generated by a family of functions $\{f_\gamma\}$ from X into topological spaces (Y_γ, τ_γ) is of crucial importance in the study of topological vector spaces and of Banach spaces. It is also vital to the understanding of the product topology on a family of topological spaces, which we shall see shortly.

11.23. Definition. *Let $\varnothing \neq X$ be a set and $\{(Y_\gamma, \tau_\gamma)\}_{\gamma\in\Gamma}$ be a family of topological spaces. Suppose that for each $\gamma \in \Gamma$ there exists a function $f_\gamma : X \to Y_\gamma$. Set $\mathcal{F} = \{f_\gamma\}_{\gamma\in\Gamma}$.*

*If $\mathcal{S} = \{f_\gamma^{-1}(G_\gamma) : G_\gamma \in \tau_\gamma, \gamma \in \Gamma\}$, then $\mathcal{S} \subseteq \mathcal{P}(X)$ and – as noted above – $\mathcal{S}$ is a subbase for a topology on X, denoted by $\sigma(X, \mathcal{F})$, and referred to as **the weak topology on X induced by $\mathcal{F}$**.*

The main and most important result concerning weak topologies induced by a family of functions is the following:

11.24. Proposition.

(a) *If τ is a topology on X and if $f_\gamma : (X, \tau) \to (Y_\gamma, \tau_\gamma)$ is continuous for all $\gamma \in \Gamma$, then $\sigma(X, \mathcal{F}) \subseteq \tau$. In other words, $\sigma(X, \mathcal{F})$ is the weakest topology on X under which each f_γ is continuous.*
(b) *Let (Z, τ_Z) be a topological space. Then $g : (Z, \tau_Z) \to (X, \sigma(X, \mathcal{F}))$ is continuous if and only if $f_\gamma \circ g : Z \to Y_\gamma$ is continuous for all $\gamma \in \Gamma$.*

11.25. Definition. *Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in\Lambda}$ be a collection of topological spaces. The **Cartesian product** of the sets X_α is*

$$\Pi_{\alpha\in\Lambda} X_\alpha = \{x : \Lambda \to \cup_\alpha X_\alpha \mid \ x(\alpha) \in X_\alpha \text{ for each } \alpha \in \Lambda\}.$$

As with sequences, we write $(x_\alpha)_\alpha$ for x.

The map $\pi_\beta : \Pi X_\alpha \to X_\beta$, $\pi_\beta(x) = x_\beta$ is called the βth projection map.

*The **product topology** on $\Pi_\alpha X_\alpha$ is the weak topology on $\Pi_\alpha X_\alpha$ induced by the family $\{\pi_\beta\}_{\beta\in\Lambda}$. As we shall see in the Assignments, this is the topology which has as a base the collection $\mathcal{B} = \{\Pi_{\alpha\in\Lambda} U_\alpha\}$, where*

(a) *$U_\alpha \in \tau_\alpha$ for all α; and*
(b) *for all but finitely many α, $U_\alpha = X_\alpha$.*

*It should be clear from the definition that in (a), it suffices to ask that we take $U_\alpha \in \mathcal{B}_\alpha$, where $\mathcal{B}_\alpha$ is a **fixed** base for τ_α, $\alpha \in \Lambda$.*

Observe that if $U_\alpha \in \tau_\alpha$ and $U_\alpha = X_\alpha$ for all α except for $\alpha_1, \alpha_2, ..., \alpha_n$, then

$$\Pi_\alpha\, U_\alpha = \pi_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \cdots \cap \pi_{\alpha_n}^{-1}(U_{\alpha_n}).$$

From this it follows that $\{\pi_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathcal{B}_\alpha, \alpha \in \Lambda\}$ *is a subbase for the product topology, where* $\mathcal{B}_\alpha$ *is a fixed base (or indeed even a subbase will do) for the topology on* X_α*.*

It is perhaps worth pointing out that it follows from the Axiom of Choice that if for all $\alpha \in \Lambda$ we have $X_\alpha \neq \varnothing$, then $X \neq \varnothing$.

We leave it to the reader to verify that the product topology on $\mathbb{R}^n = \Pi_{k=1}^n \mathbb{R}$ is just the usual topology on $\mathbb{R}^n$.

Bibliography

[Dav88] K.R. Davidson. *Nest algebras. Triangular forms for operators on a Hilbert space*, volume 191 of *Pitman Research Notes in Mathematics*. Longman Scientific and Technical, Harlow, 1988.

[Jam51] R.C. James. A non-reflexive banach space isometric with its second conjugate space. *Proc. Nat. Acad. Sci. USA*, 37:174–177, 1951.

[KR83] R.V. Kadison and J.R. Ringrose. *Fundamentals of the theory of operator algebras I: Elementary theory*. Academic Press, New York, 1983.

[Lin71] J. Lindenstrauss. The geometric theory of the classical banach spaces. *Actes du Congrès Intern. Math., 1970, Paris*, 2:365–372, 1971.

[Sch27] J. Schauder. Zur Theorie stetiger Abbildungen in Funktionalraumen. *Math. Z.*, 26:47–65, 1927.

[Sou78] A.R. Sourour. Operators with absolutely bounded matrices. *Math. Z.*, 162:183–187, 1978.

[Sun78] V.S. Sunder. Absolutely bounded matrices. *Indiana Univ. Math. J.*, 27:919–927, 1978.

[Tsi74] B.S. Tsirel'son. Not every Banach space contains an imbedding of ℓ^p or c_0. *Functional Anal. Appl.*, 8:138–141, 1974.

[Whi66] R. Whitley. Projecting m onto c_0. *The Amer. Math. Monthly*, 73:285–286, 1966.

[Wil70] Stephen Willard. *General Topology*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970.

Index