Functional Analysis
Lecture notes for 18.102

Richard Melrose

Department of Mathematics, MIT
E-mail address: rbm@math.mit.edu

$\boxed{\boxed{\text{Version 0.8D; Revised: 4-5-2010; Run: May 21, 2017}}}$.

Contents

Preface

These are notes for the course 'Introduction to Functional Analysis' – or in the MIT style, 18.102, from various years culminating in Spring 2017. There are many people who I should like to thank for comments on and corrections to the notes over the years, but for the moment I would simply like to thank, as a collective, the MIT undergraduates who have made this course a joy to teach, as a result of their interest and enthusiasm.

Introduction

This course is intended for 'well-prepared undergraduates' meaning specifically that they have a rigourous background in analysis at roughly the level of the first half of Rudin's book [**4**] – at MIT this is 18.100B. In particular the basic theory of metric spaces is used freely. Some familiarity with linear algebra is also assumed, but not at a very sophisticated level.

The main aim of the course in a mathematical sense is the presentation of the standard constructions of linear functional analysis, centred on Hilbert space and its most significant analytic realization as the Lebesgue space $L^2(\mathbb{R})$ and leading up to the spectral theory of ordinary differential operators. In a one-semester course at MIT it is only just possible to get this far. Beyond the core material I have included other topics that I believe may prove useful both in showing how to use the 'elementary' results in various directions.

Dirichlet problem. The treatment of the eigenvalue problem with potential perturbation on an interval is one of the aims of this course, so let me describe it briefly here for orientation.

Let $V : [0, 1] \longrightarrow \mathbb{R}$ be a real-valued continuous function. We are interested in 'oscillating modes' on the interval; something like this arises in quantum mechanics for instance. Namely we want to know about functions $u(x)$ – twice continuously differentiable on $[0, 1]$ so that things make sense – which satisfy the differential equation

$$(1)\qquad \begin{gathered} -\frac{d^2u}{dx^2}(x) + V(x)u(x) = \lambda u(x) \\ \text{and the boundary conditions} \\ u(0) = u(1) = 0. \end{gathered}$$

Here the eigenvalue, λ is an 'unknown' constant. More precisely we wish to know which such λ's can occur. In fact all λ's can occur with $u \equiv 0$ but this is the 'trivial solution' which will always be there for such an equation. What other solutions are there? The main result is that there is an infinite sequence of λ's for which there is a non-trivial solution of (1) $\lambda_j \in \mathbb{R}$ – they are all real, no non-real complex λ's can occur. For each of these λ_j there is at least one (and there might be more than one) non-trivial solution u_j to (1). We can say a lot more about everything here but one main aim of this course is to get at least to this point. From a Physical point of view, (1) represents a linearized oscillating string with fixed ends.

The journey to a discussion of the Dirichlet problem is rather extended and apparently wayward. The relevance of Hilbert space and the Lebesgue integral is not immediately apparent, but I hope this will become clear as we proceed. In fact in this one-dimensional setting it can be avoided, although at some cost in terms of elegance. The basic idea is that we consider a space of all 'putative' solutions to the problem at hand. In this case one might take the space of all twice continuously differentiable functions on $[0, 1]$ – we will consider such spaces at least briefly below. One of the weaknesses of such an approach is that it is not closely connected with the 'energy' invariant of a solution, which is the integral

$$(2)\qquad \int_0^1 (|\frac{du}{dx}|^2 + V(x)|u(x)|^2)dx.$$

It is the importance of such integrals which brings in the Lebesgue integral and leads to a Hilbert space structure.

In any case one of the significant properties of the equation (1) is that it is 'linear'. So we start with a brief discussion of linear spaces. What we are dealing with here can be thought of as the an eigenvalue problem for an 'infinite matrix'. This in fact is not a very good way of thinking about operators on infinite-dimensional spaces, they are not really like infinite matrices, but in this case it is justified by the appearance of *compact operators* which *are* rather more like infinite matrices. There was a matrix approach to quantum mechanics in the early days but it was replaced by the sort of 'operator' theory on Hilbert space that we will discuss below. One of the crucial distinctions between the treatment of finite dimensional matrices and an infinite dimensional setting is that in the latter *topology* is encountered. This is enshrined in the notion of a *normed linear space* which is the first important topic we shall meet.

After a brief treatment of normed and Banach spaces, the course proceeds to the construction of the Lebesgue integral and the associated spaces of 'Lebesgue interable functions' (as you will see this is by way of a universally accepted lie, but a good one). Usually I have done this in one dimension, on the line, leading to the definition of the space $L^1(\mathbb{R})$. To some extent I follow here the idea of Jan Mikusiński that one can simply define integrable functions as the almost everywhere limits of absolutely summable series of step functions and more significantly the basic properties can be deduced this way. While still using this basic approach I have dropped the step functions almost completely and instead emphasize the completion of the space of continuous functions to get the Lebesgue space. Even so, Mikusiński's approach still underlies the explicit identification of elements of the completion with Lebesgue 'functions'. This approach is followed in the book of Debnaith and Mikusiński [**1**].

After about a two-week stint of integration and then a little measure theory the course proceeds to the more gentle ground of Hilbert spaces. Here I have been most guided by the (old now) book of Simmons [**5**] which is still very much worth reading. We proceed to a short discussion of operators and the spectral theorem for compact self-adjoint operators. I have also included in the notes (but generally not in the lectures) various things that a young mathematician should know(!) such as Kuiper's Theorem. Then in the last third or so of the semester this theory is applied to the treatment of the Dirichlet eigenvalue problem, followed by a short discussion of the Fourier transform and the harmonic oscillator. Finally various loose ends are brought together, or at least that is my hope.

CHAPTER 1

Normed and Banach spaces

In this chapter we introduce the basic setting of functional analysis, in the form of normed spaces and bounded linear operators. We are particularly interested in complete, i.e. Banach, spaces and the process of completion of a normed space to a Banach space. In lectures I proceed to the next chapter, on Lebesgue integration after Section 7 and then return to the later sections of this chapter at appropriate points in the course.

There are many good references for this material and it is always a good idea to get at least a couple of different views. The treatment here, whilst quite brief, does cover what is needed later.

1. Vector spaces

You should have some familiarity with linear, or I will usually say 'vector', spaces. Should I break out the axioms? Not here I think, but they are included in Section 14 at the end of the chapter. In short it is a space V in which we can add elements and multiply by scalars with rules quite familiar to you from the the basic examples of $\mathbb{R}^n$ or $\mathbb{C}^n$. Whilst these special cases are (very) important below, this is not what we are interested in studying here. The main examples are spaces of *functions* hence the name of the course.

Note that for us the 'scalars' are either the real numbers or the complex numbers – usually the latter. To be neutral we denote by $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$, but of course consistently. Then our set V – the set of vectors with which we will deal, comes with two 'laws'. These are maps

$$+ : V \times V \longrightarrow V,\ \cdot : \mathbb{K} \times V \longrightarrow V. \tag{1.1}$$

which we denote not by $+(v, w)$ and $\cdot(s, v)$ but by $v + w$ and sv. Then we impose the *axioms of a vector space* – see Section 14 below! These are commutative group axioms for $+$, axioms for the action of $\mathbb{K}$ and the distributive law linking the two.

The basic examples:

- The field $\mathbb{K}$ which is either $\mathbb{R}$ or $\mathbb{C}$ is a vector space over itself.
- The vector spaces $\mathbb{K}^n$ consisting of ordered n-tuples of elements of $\mathbb{K}$. Addition is by components and the action of $\mathbb{K}$ is by multiplication on all components. You should be reasonably familiar with these spaces and other finite dimensional vector spaces.
- Seriously non-trivial examples such as $\mathrm{C}([0, 1])$ the space of continuous functions on $[0, 1]$ (say with complex values).

In these and many other examples we will encounter below, the 'component addition' corresponds to the addition of functions.

Lemma 1.1. *If X is a set then the spaces of all functions*

$$\mathcal{F}(X;\mathbb{R}) = \{u : X \longrightarrow \mathbb{R}\},\ \mathcal{F}(X;\mathbb{C}) = \{u : X \longrightarrow \mathbb{C}\} \tag{1.2}$$

are vector spaces over $\mathbb{R}$ and $\mathbb{C}$ respectively.

Non-Proof. Since I have not written out the axioms of a vector space it is hard to check this – and I leave it to you as the first of many important exercises. In fact, better do it more generally as in Problem 1.0 – then you can sound sophisticated by saying 'if V is a linear space then $\mathcal{F}(X;V)$ inherits a linear structure'. The main point to make sure you understand is precisely this; because we *do* know how to add and multiply in either $\mathbb{R}$ and $\mathbb{C}$, we can add functions and multiply them by constants (we can multiply functions by each other but that is not part of the definition of a vector space so we ignore it for the moment since many of the spaces of functions we consider below are *not* multiplicative in this sense):-

$$(c_1f_1 + c_2f_2)(x) = c_1f_1(x) + c_2f_2(x) \tag{1.3}$$

defines the function $c_1f_1 + c_2f_2$ if $c_1,\ c_2 \in \mathbb{K}$ and $f_1,\ f_2 \in \mathcal{F}(X;\mathbb{K})$. □

You should also be familiar with the notions of linear subspace and quotient space. These are discussed a little below and most of the linear spaces we will meet are either subspaces of these function-type spaces, or quotients of such subspaces – see Problems 1.1 and 1.2.

Although you are probably most comfortable with finite-dimensional vector spaces it is the infinite-dimensional case that is most important here. The notion of dimension is based on the concept of the linear independence of a subset of a vector space. Thus a subset $E \subset V$ is said to be *linearly independent* if for any finite collection of elements $v_i \in E$, $i = 1, \dots, N$, and any collection of 'constants' $a_i \in \mathbb{K}$, $i = 1, \dots, N$ we have the following implication

$$\sum_{i=1}^{N} a_iv_i = 0 \Longrightarrow a_i = 0\ \forall\ i. \tag{1.4}$$

That is, it is a set in which there are 'no non-trivial finite linear dependence relations between the elements'. A vector space is finite-dimensional if every linearly independent subset is finite. It follows in this case that there is a finite and maximal linearly independent subset – a basis – where maximal means that if any new element is added to the set E then it is no longer linearly independent. A basic result is that any two such 'bases' in a finite dimensional vector space have the same number of elements – an outline of the finite-dimensional theory can be found in Problem 1.3.

Still it is time to leave this secure domain behind, since we are most interested in the other case, namely infinite-dimensional vector spaces. As usual with such mysterious-sounding terms as 'infinite-dimensional' it is defined by negation.

Definition 1.1. A vector space is infinite-dimensional if it is not finite dimensional, i.e. for any $N \in \mathbb{N}$ there exist N elements with no, non-trivial, linear dependence relation between them.

As is quite typical, the idea of an infinite-dimensional space, which you may be quite keen to understand, appears just as the non-existence of something. That is, it is the 'residual' case, where there is no finite basis. This means that it is 'big'.

So, finite-dimensional vector spaces have finite bases, infinite-dimensional vector spaces do not. Make sure that you see the gap between these two cases, i.e. either a vector space has a finite-dimensional basis or else it has an infinite linearly independent set. In particular if there is a linearly independent set with N elements for any N then there is an infinite one, there is a point here – if an independent finite set has the property that there is no element of the space which can be added to it so that it remains independent then it already is a basis and any other independent set has the same or fewer elements.

The notion of a basis in an infinite-dimensional vector spaces needs to be modified to be useful analytically. Convince yourself that the vector space in Lemma 1.1 is infinite dimensional if and only if X is infinite. [1]

2. Normed spaces

We need to deal effectively with infinite-dimensional vector spaces. To do so we need the control given by a metric (or even more generally a non-metric topology, but we will only get to that very briefly near the end of this course). A norm on a vector space leads to a metric which is 'compatible' with the linear structure.

Definition 1.2. A norm on a vector space is a function, traditionally denoted

$$\|\cdot\|:V\longrightarrow[0,\infty),\tag{1.5}$$

with the following properties

(*Definiteness*)

$$v\in V,\ \|v\|=0\Longrightarrow v=0.\tag{1.6}$$

(*Absolute homogeneity*) For any $\lambda\in\mathbb{K}$ and $v\in V$,

$$\|\lambda v\|=|\lambda|\|v\|.\tag{1.7}$$

(*Triangle Inequality*) The triangle inequality holds, in the sense that for any two elements $v,\ w\in V$

$$\|v+w\|\le\|v\|+\|w\|.\tag{1.8}$$

Note that (1.7) implies that $\|0\|=0$. Thus (1.6) means that $\|v\|=0$ is equivalent to $v=0$. This definition is based on the same properties holding for the standard norm(s), $|z|$, on $\mathbb{R}$ and $\mathbb{C}$. You should make sure you understand that

$$\begin{gathered}\mathbb{R}\ni x\longrightarrow|x|=\begin{cases}x & \text{if } x\ge 0\\ -x & \text{if } x\le 0\end{cases}\in[0,\infty)\text{ is a norm as is}\\ \mathbb{C}\ni z=x+iy\longrightarrow|z|=(x^2+y^2)^{\frac12}.\end{gathered}\tag{1.9}$$

Situations do arise in which we do not have (1.6):-

Definition 1.3. A function (1.5) which satisfes (1.7) and (1.8) but possibly not (1.6) is called a *seminorm.*

[1] Hint: For each point $y\in X$ consider the function $f:X\longrightarrow\mathbb{C}$ which takes the value 1 at y and 0 at every other point. Show that if X is finite then any function $X\longrightarrow\mathbb{C}$ is a finite linear combination of these, and if X is infinite then this is an infinite set with no finite linear relations between the elements.

A metric, or distance function, on a set is a map

$$d : X \times X \longrightarrow [0, \infty) \tag{1.10}$$

satisfying three standard conditions

$$d(x, y) = 0 \Longleftrightarrow x = y, \tag{1.11}$$

$$d(x, y) = d(y, x) \ \forall \ x, y \in X \text{ and} \tag{1.12}$$

$$d(x, y) \le d(x, z) + d(z, y) \ \forall \ x, y, z \in X. \tag{1.13}$$

As you are no doubt aware, a set equipped with such a metric function is called a metric space.

If you do not know about metric spaces, then you are in trouble. I suggest that you take the appropriate course now and come back next year. You could read the first few chapters of Rudin's book [**4**] before trying to proceed much further but it will be a struggle to say the least. The point of course is

Proposition 1.1. *If $\|\cdot\|$ is a norm on V then*

$$d(v, w) = \|v - w\| \tag{1.14}$$

is a distance on V turning it into a metric space.

Proof. Clearly (1.11) corresponds to (1.6), (1.12) arises from the special case $\lambda = -1$ of (1.7) and (1.13) arises from (1.8). □

We will not use any special notation for the metric, nor usually mention it explicitly – we just subsume all of metric space theory from now on. So $\|v - w\|$ is the distance between two points in a normed space.

Now, we need to talk about a few examples; there are more in Section 7. The most basic ones are the usual finite-dimensional spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ with their Euclidean norms

$$|x| = \left(\sum_i |x_i|^2\right)^{\frac{1}{2}} \tag{1.15}$$

where it is at first confusing that we just use single bars for the norm, just as for $\mathbb{R}$ and $\mathbb{C}$, but you just need to get used to that.

There are other norms on $\mathbb{C}^n$ (I will mostly talk about the complex case, but the real case is essentially the same). The two most obvious ones are

$$\begin{gathered} |x|_\infty = \max |x_i|, \ x = (x_1, \dots, x_n) \in \mathbb{C}^n, \\ |x|_1 = \sum_i |x_i| \end{gathered} \tag{1.16}$$

but as you will see (if you do the problems) there are also the norms

$$|x|_p = (\sum_i |x_i|^p)^{\frac{1}{p}}, \ 1 \le p < \infty. \tag{1.17}$$

In fact, for $p = 1$, (1.17) reduces to the second norm in (1.16) and in a certain sense the case $p = \infty$ is consistent with the first norm there.

In lectures I usually do not discuss the notion of equivalence of norms straight away. However, two norms on the one vector space – which we can denote $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are *equivalent* if there exist constants C_1 and C_2 such that

$$\|v\|_{(1)} \le C_1 \|v\|_{(2)}, \ \|v\|_{(2)} \le C_2 \|v\|_{(1)} \ \forall \ v \in V. \tag{1.18}$$

The equivalence of the norms implies that the metrics define the same open sets – the topologies induced are the same. You might like to check that the reverse is also true, if two norms induced the same topologies (just meaning the same collection of open sets) through their associated metrics, then they are equivalent in the sense of (1.18) (there are more efficient ways of doing this if you wait a little).

Look at Problem 1.5 to see why we are not so interested in norms in the finite-dimensional case – namely any two norms on a finite-dimensional vector space are equivalent and so in that case a choice of norm does not tell us much, although it certainly has its uses.

One important class of normed spaces consists of the spaces of bounded continuous functions on a metric space X :

$$\mathcal{C}_\infty(X) = \mathcal{C}_\infty(X;\mathbb{C}) = \{u : X \longrightarrow \mathbb{C},\ \text{continuous and bounded}\}. \tag{1.19}$$

That this is a linear space follows from the (pretty obvious) result that a linear combination of bounded functions is bounded and the (less obvious) result that a linear combination of continuous functions is continuous; this we are supposed to know. The norm is the best bound

$$\|u\|_\infty = \sup_{x\in X} |u(x)|. \tag{1.20}$$

That this *is* a norm is straightforward to check. Absolute homogeneity is clear, $\|\lambda u\|_\infty = |\lambda|\|u\|_\infty$ and $\|u\|_\infty = 0$ means that $u(x) = 0$ for all $x \in X$ which is exactly what it means for a function to vanish. The triangle inequality 'is inherited from $\mathbb{C}$' since for any two functions and any point,

$$|(u+v)(x)| \le |u(x)| + |v(x)| \le \|u\|_\infty + \|v\|_\infty \tag{1.21}$$

by the definition of the norms, and taking the supremum of the left gives

$$\|u+v\|_\infty \le \|u\|_\infty + \|v\|_\infty.$$

Of course the norm (1.20) is defined even for bounded, not necessarily continuous functions on X. Note that convergence of a sequence $u_n \in \mathcal{C}_\infty(X)$ (remember this means with respect to the distance induced by the norm) is precisely *uniform convergence*

$$\|u_n - v\|_\infty \to 0 \Longleftrightarrow u_n(x) \to v(x) \text{ uniformly on } X. \tag{1.22}$$

Other examples of infinite-dimensional normed spaces are the spaces l^p, $1 \le p \le \infty$ discussed in the problems below. Of these l^2 is the most important for us. It is in fact one form of Hilbert space, with which we are primarily concerned:-

$$l^2 = \{a : \mathbb{N} \longrightarrow \mathbb{C}; \sum_j |a(j)|^2 < \infty\}. \tag{1.23}$$

It is not immediately obvious that this is a linear space, nor that

$$\|a\|_2 = \left(\sum_j |a(j)|^2\right)^{\frac{1}{2}} \tag{1.24}$$

is a norm. It is. From now on we will generally use sequential notation and think of a map from $\mathbb{N}$ to $\mathbb{C}$ as a sequence, so setting $a(j) = a_j$. Thus the 'Hilbert space' l^2 consists of the square summable sequences.

3. Banach spaces

You are supposed to remember from metric space theory that there are three crucial properties, completeness, compactness and connectedness. It turns out that normed spaces are always connected, so that is not very interesting, and they are never compact (unless you consider the trivial case $V = \{0\}$) so that is not very interesting either – in fact we will ultimately be very interested in compact subsets. So that leaves completeness. This is so important that we give it a special name in honour of Stefan Banach who first emphasized this property.

Definition 1.4. A normed space which is complete with respect to the induced metric is a *Banach* space.

Lemma 1.2. *The space $\mathcal{C}_\infty(X)$, defined in* (1.19) *for any metric space X, is a Banach space.*

Proof. This is a standard result from metric space theory – basically that the uniform limit of a sequence of (bounded) continuous functions on a metric space is continuous. However, it is worth recalling how one proves completeness at least in outline. Suppose u_n is a Cauchy sequence in $\mathcal{C}_\infty(X)$. This means that given $\delta > 0$ there exists N such that

$$n, m > N \Longrightarrow \|u_n - u_m\|_\infty = \sup_X |u_n(x) - u_m(x)| < \delta. \tag{1.25}$$

Fixing $x \in X$ this implies that the sequence $u_n(x)$ is Cauchy in $\mathbb{C}$. We know that this space is complete, so each sequence $u_n(x)$ must converge (we say the sequence of functions converges pointwise). Since the limit of $u_n(x)$ can only depend on x, we may define $u(x) = \lim_n u_n(x)$ in $\mathbb{C}$ for each $x \in X$ and so define a function $u : X \longrightarrow \mathbb{C}$. Now, we need to show that this is bounded and continuous and is the limit of u_n with respect to the norm. Any Cauchy sequence is bounded in norm – take $\delta = 1$ in (1.25) and it follows from the triangle inequality that

$$\|u_m\|_\infty \le \|u_{N+1}\|_\infty + 1,\ m > N \tag{1.26}$$

and the finite set $\|u_n\|_\infty$ for $n \le N$ is certainly bounded. Thus $\|u_n\|_\infty \le C$, but this means $|u_n(x)| \le C$ for all $x \in X$ and hence $|u(x)| \le C$ by properties of convergence in $\mathbb{C}$ and thus $\|u\|_\infty \le C$, so the limit is bounded.

The uniform convergence of u_n to u now follows from (1.25) since we may pass to the limit in the inequality to find

$$\begin{aligned} n > N \Longrightarrow |u_n(x) - u(x)| &= \lim_{m\to\infty} |u_n(x) - u_m(x)| \le \delta \\ &\Longrightarrow \|u_n - u\|_\infty \le \delta. \end{aligned} \tag{1.27}$$

The continuity of u at $x \in X$ follows from the triangle inequality in the form

$$\begin{aligned} |u(y) - u(x)| \le |u(y) - u_n(y)| + |u_n(y) - u_n(x)| + |u_n(x) - u(x)| \\ \le 2\|u - u_n\|_\infty + |u_n(x) - u_n(y)|. \end{aligned}$$

Given $\delta > 0$ the first term on the far right can be make less than $\delta/2$ by choosing n large using (1.27) and then, having chosen n, the second term can be made less than $\delta/2$ by choosing $d(x, y)$ small enough, using the continuity of u_n. □

I have written out this proof (succinctly) because this general structure arises often below – first find a candidate for the limit and then show it has the properties that are required.

There is a space of sequences which is really an example of this Lemma. Consider the space c_0 consisting of all the sequences $\{a_j\}$ (valued in $\mathbb{C}$) such that $\lim_{j\to\infty} a_j = 0$. As remarked above, sequences are just functions $\mathbb{N} \longrightarrow \mathbb{C}$. If we make $\{a_j\}$ into a function $\alpha : D = \{1, 1/2, 1/3, \dots\} \longrightarrow \mathbb{C}$ by setting $\alpha(1/j) = a_j$ then we get a function on the metric space D. Add 0 to D to get $\overline{D} = D \cup \{0\} \subset [0,1] \subset \mathbb{R}$; clearly 0 is a limit point of D and $\overline{D}$ is, as the notation dangerously indicates, the closure of D in $\mathbb{R}$. Now, you will easily check (it is really the definition) that $\alpha : D \longrightarrow \mathbb{C}$ corresponding to a sequence, extends to a *continuous* function on $\overline{D}$ vanishing at 0 if and only if $\lim_{j\to\infty} a_j = 0$, which is to say, $\{a_j\} \in c_0$. Thus it follows, with a little thought which you should give it, that c_0 is a Banach space with the norm

$$\|a\|_\infty = \sup_j \|a_j\|. \tag{1.28}$$

What is an example of a non-complete normed space, a normed space which is *not* a Banach space? These are legion of course. The simplest way to get one is to 'put the wrong norm' on a space, one which does not correspond to the definition. Consider for instance the linear space $\mathcal{T}$ of sequences $\mathbb{N} \longrightarrow \mathbb{C}$ which 'terminate', i.e. each element $\{a_j\} \in \mathcal{T}$ has $a_j = 0$ for $j > J$, where of course the J may depend on the particular sequence. Then $\mathcal{T} \subset c_0$, the norm on c_0 defines a norm on $\mathcal{T}$ but it cannot be complete, since the closure of $\mathcal{T}$ is easily seen to be all of c_0 – so there are Cauchy sequences in $\mathcal{T}$ without limit in $\mathcal{T}$. Make sure you are not lost here – you need to get used to the fact that we often need to discuss the 'convergence of sequences of convergent sequences' as here.

One result we will exploit below, and I give it now just as preparation, concerns *absolutely summable series*. Recall that a series is just a sequence where we 'think' about adding the terms. Thus if v_n is a sequence in some vector space V then there is the corresponding sequence of partial sums $w_N = \sum\limits_{i=1}^{N} v_i$. I will say that $\{v_n\}$ is a series if I am thinking about summing it.

Definition 1.5. A series $\{v_n\}$ with partial sums $\{w_N\}$ is said to be *absolutely summable* if

$$\sum_n \|v_n\|_V < \infty, \text{ i.e. } \sum_{N>1} \|w_N - w_{N-1}\|_V < \infty. \tag{1.29}$$

Proposition 1.2. *The sequence of partial sums of any absolutely summable series in a normed space is Cauchy and a normed space is complete if and only if every absolutely summable series in it converges, meaning that the sequence of partial sums converges.*

Proof. The sequence of partial sums is

$$w_n = \sum_{j=1}^{n} v_j. \tag{1.30}$$

Thus, if $m > n$ then

$$w_m - w_n = \sum_{j=n+1}^{m} v_j. \tag{1.31}$$

It follows from the triangle inequality that

$$\|w_n - w_m\|_V \le \sum_{j=n+1}^{m} \|v_j\|_V. \tag{1.32}$$

So if the series is absolutely summable then

$$\sum_{j=1}^{\infty} \|v_j\|_V < \infty \text{ and } \lim_{n\to\infty} \sum_{j=n+1}^{\infty} \|v_j\|_V = 0.$$

Thus $\{w_n\}$ is Cauchy if $\{v_j\}$ is absolutely summable. Hence if V is complete then every absolutely summable series is summable, i.e. the sequence of partial sums converges.

Conversely, suppose that every absolutely summable series converges in this sense. Then we need to show that every Cauchy sequence in V converges. Let u_n be a Cauchy sequence. It suffices to show that this has a subsequence which converges, since a Cauchy sequence with a convergent subsequence is convergent. To do so we just proceed inductively. Using the Cauchy condition we can for every k find an integer N_k such that

$$n, m > N_k \Longrightarrow \|u_n - u_m\| < 2^{-k}. \tag{1.33}$$

Now choose an increasing sequence n_k where $n_k > N_k$ and $n_k > n_{k-1}$ to make it increasing. It follows that

$$\|u_{n_k} - u_{n_{k-1}}\| \le 2^{-k+1}. \tag{1.34}$$

Denoting this subsequence as $u'_k = u_{n_k}$ it follows from (1.34) and the triangle inequality that

$$\sum_{n=1}^{\infty} \|u'_n - u'_{n-1}\| \le 4 \tag{1.35}$$

so the sequence $v_1 = u'_1$, $v_k = u'_k - u'_{k-1}$, $k > 1$, is absolutely summable. Its sequence of partial sums is $w_j = u'_j$ so the assumption is that this converges, hence the original Cauchy sequence converges and V is complete. □

Notice the idea here, of 'speeding up the convergence' of the Cauchy sequence by dropping a lot of terms. We will use this idea of absolutely summable series heavily in the discussion of Lebesgue integration.

4. Operators and functionals

The vector spaces we are most interested in are, as already remarked, spaces of functions (or something a little more general). The elements of these are the objects of primary interest but we are especially interested in the way they are related by linear maps. The sorts of maps we have in mind here are differential and integral operators. For example the indefinite Riemann integral of a continuous function $f : [0, 1] \longrightarrow \mathbb{C}$ is also a continuous function of the upper limit:

$$I(f)(x) = \int_0^x f(s)ds. \tag{1.36}$$

So, $I : \mathcal{C}([0, 1]) \longrightarrow \mathcal{C}([0, 1])$ it is an 'operator' which turns one continuous function into another. You might want to bear such an example in mind as you go through this section.

A map between two vector spaces (over the same field, for us either $\mathbb{R}$ or $\mathbb{C}$) is linear if it takes linear combinations to linear combinations:-

$$T : V \longrightarrow W,\ T(a_1v_1+a_2v_2) = a_1T(v_1)+a_2T(v_2),\ \forall\, v_1,\ v_2 \in V,\ a_1, a_2 \in \mathbb{K}. \tag{1.37}$$

In the finite-dimensional case linearity is enough to allow maps to be studied. However in the case of infinite-dimensional normed spaces we need to impose continuity. Of course it makes perfectly good sense to say, demand or conclude, that a map as in (1.37) is continuous if V and W are normed spaces since they are then metric spaces. Recall that for metric spaces there are several different equivalent conditions that ensure a map, $T : V \longrightarrow W$, is continuous:

$$v_n \to v \text{ in } V \Longrightarrow Tv_n \to Tv \text{ in } W \tag{1.38}$$

$$O \subset W \text{ open } \Longrightarrow T^{-1}(O) \subset V \text{ open} \tag{1.39}$$

$$C \subset W \text{ closed } \Longrightarrow T^{-1}(C) \subset V \text{ closed.} \tag{1.40}$$

For a linear map between normed spaces there is a direct characterization of continuity in terms of the norm.

Proposition 1.3. *A linear map* (1.37) *between normed spaces is continuous if and only if it is* bounded *in the sense that there exists a constant C such that*

$$\|Tv\|_W \le C\|v\|_V\ \ \forall\ v \in V. \tag{1.41}$$

Of course bounded for a function on a metric space already has a meaning and this is not it! The usual sense would be $\|Tv\| \le C$ but this would imply $\|T(av)\| = |a|\|Tv\| \le C$ so $Tv = 0$. Hence it is not so dangerous to use the term 'bounded' for (1.41) – it is really 'relatively bounded', i.e. takes bounded sets into bounded sets. From now on, bounded for a linear map means (1.41).

Proof. If (1.41) holds then if $v_n \to v$ in V it follows that $\|Tv - Tv_n\| = \|T(v - v_n)\| \le C\|v - v_n\| \to 0$ as $n \to \infty$ so $Tv_n \to Tv$ and continuity follows.

For the reverse implication we use the second characterization of continuity above. Denote the ball around $v \in V$ of radius $\epsilon > 0$ by

$$B_V(v, \epsilon) = \{w \in V; \|v - w\| < \epsilon\}.$$

Thus if T is continuous then the inverse image of the the unit ball around the origin, $T^{-1}(B_W(0, 1)) = \{v \in V; \|Tv\|_W < 1\}$, contains the origin in V and so, being open, must contain some $B_V(0, \epsilon)$. This means that

$$T(B_V(0, \epsilon)) \subset B_W(0, 1) \text{ so } \|v\|_V < \epsilon \Longrightarrow \|Tv\|_W \le 1. \tag{1.42}$$

Now proceed by scaling. If $0 \ne v \in V$ then $\|v'\| < \epsilon$ where $v' = \epsilon v/2\|v\|$. So (1.42) shows that $\|Tv'\| \le 1$ but this implies (1.41) with $C = 2/\epsilon$ – it is trivially true if $v = 0$. □

As a general rule we drop the distinguishing subscript for norms, since which norm we are using can be determined by what it is being applied to.

So, if $T : V \longrightarrow W$ is continous and linear between normed spaces, or from now on 'bounded', then

$$\|T\| = \sup_{\|v\|=1} \|Tv\| < \infty. \tag{1.43}$$

LEMMA 1.3. *The bounded linear maps between normed spaces V and W form a linear space $\mathcal{B}(V, W)$ on which $\|T\|$ defined by* (1.43) *or equivalently*

$$\|T\| = \inf\{C;\ (1.41)\ \textit{holds}\} \tag{1.44}$$

is a norm.

PROOF. First check that (1.43) is equivalent to (1.44). Define $\|T\|$ by (1.43). Then for any $v \in V$, $v \neq 0$,

$$\|T\| \geq \|T(\frac{v}{\|v\|})\| = \frac{\|Tv\|}{\|v\|} \Longrightarrow \|Tv\| \leq \|T\|\|v\| \tag{1.45}$$

since as always this is trivially true for $v = 0$. Thus $C = \|T\|$ is a constant for which (1.41) holds.

Conversely, from the definition of $\|T\|$, if $\epsilon > 0$ then there exists $v \in V$ with $\|v\| = 1$ such that $\|T\| - \epsilon < \|Tv\| \leq C$ for any C for which (1.41) holds. Since $\epsilon > 0$ is arbitrary, $\|T\| \leq C$ and hence $\|T\|$ is given by (1.44).

From the definition of $\|T\|$, $\|T\| = 0$ implies $Tv = 0$ for all $v \in V$ and for $\lambda \neq 0$,

$$\|\lambda T\| = \sup_{\|v\|=1} \|\lambda Tv\| = |\lambda|\|T\| \tag{1.46}$$

and this is also obvious for $\lambda = 0$. This only leaves the triangle inequality to check and for any T, $S \in \mathcal{B}(V, W)$, and $v \in V$ with $\|v\| = 1$

$$\|(T+S)v\|_W = \|Tv + Sv\|_W \leq \|Tv\|_W + \|Sv\|_W \leq \|T\| + \|S\| \tag{1.47}$$

so taking the supremum, $\|T + S\| \leq \|T\| + \|S\|$. □

Thus we see the very satisfying fact that the space of bounded linear maps between two normed spaces is itself a normed space, with the norm being the best constant in the estimate (1.41). Make sure you absorb this! Such bounded linear maps between normed spaces are often called 'operators' because we are thinking of the normed spaces as being like function spaces.

You might like to check boundedness for the example, I, of a linear operator in (1.36), namely that in terms of the supremum norm on $\mathcal{C}([0, 1])$, $\|T\| \leq 1$.

One particularly important case is when $W = \mathbb{K}$ is the field, for us usually $\mathbb{C}$. Then a simpler notation is handy and one sets $V' = \mathcal{B}(V, \mathbb{C})$ – this is called the *dual space* of V (also sometimes denoted V^*.)

PROPOSITION 1.4. *If W is a Banach space then $\mathcal{B}(V, W)$, with the norm* (1.43)*, is a Banach space.*

PROOF. We simply need to show that if W is a Banach space then every Cauchy sequence in $\mathcal{B}(V, W)$ is convergent. The first thing to do is to find the limit. To say that $T_n \in \mathcal{B}(V, W)$ is Cauchy, is just to say that given $\epsilon > 0$ there exists N such that n, $m > N$ implies $\|T_n - T_m\| < \epsilon$. By the definition of the norm, if $v \in V$ then $\|T_n v - T_m v\|_W \leq \|T_n - T_m\|\|v\|_V$ so $T_n v$ is Cauchy in W for each $v \in V$. By assumption, W is complete, so

$$T_n v \longrightarrow w \text{ in } W. \tag{1.48}$$

However, the limit can only depend on v so we can define a map $T : V \longrightarrow W$ by $Tv = w = \lim_{n\to\infty} T_n v$ as in (1.48).

This map defined from the limits is linear, since $T_n(\lambda v) = \lambda T_n v \longrightarrow \lambda T v$ and $T_n(v_1 + v_2) = T_n v_1 + T_n v_2 \longrightarrow T v_2 + T v_2 = T(v_1 + v_2)$. Moreover, $|\|T_n\| - \|T_m\|| \le \|T_n - T_m\|$ so $\|T_n\|$ is Cauchy in $[0, \infty)$ and hence converges, with limit S, and

$$\|Tv\| = \lim_{n\to\infty} \|T_n v\| \le S\|v\| \tag{1.49}$$

so $\|T\| \le S$ shows that T is bounded.

Returning to the Cauchy condition above and passing to the limit in $\|T_n v - T_m v\| \le \epsilon\|v\|$ as $m \to \infty$ shows that $\|T_n - T\| \le \epsilon$ if $n > M$ and hence $T_n \to T$ in $\mathcal{B}(V, W)$ which is therefore complete. $\square$

Note that this proof is structurally the same as that of Lemma 1.2.

One simple consequence of this is:-

Corollary 1.1. *The dual space of a normed space is always a Banach space.*

However you should be a little suspicious here since we have not shown that the dual space V' is non-trivial, meaning we have not eliminated the possibility that $V' = \{0\}$ even when $V \neq \{0\}$. The Hahn-Banach Theorem, discussed below, takes care of this.

One game you can play is 'what is the dual of that space'. Of course the dual is the dual, but you may well be able to identify the dual space of V with some other Banach space by finding a linear bijection between V' and the other space, W, which identifies the norms as well. We will play this game a bit later.

5. Subspaces and quotients

The notion of a linear subspace of a vector space is natural enough, and you are likely quite familiar with it. Namely $W \subset V$ where V is a vector space is a (linear) subspace if any linear combinations $\lambda_1 w_1 + \lambda_2 w_2 \in W$ if $\lambda_1,\ \lambda_2 \in \mathbb{K}$ and $w_1,\ w_2 \in W$. Thus W 'inherits' its linear structure from V. Since we also have a topology from the metric we will be especially interested in closed subspaces. Check that you understand the (elementary) proof of

Lemma 1.4. *A subspace of a Banach space is a Banach space in terms of the restriction of the norm if and only if it is closed.*

There is a second very important way to construct new linear spaces from old. Namely we want to make a linear space out of 'the rest' of V, given that W is a linear subspace. In finite dimensions one way to do this is to give V an inner product and then take the subspace orthogonal to W. One problem with this is that the result depends, although not in an essential way, on the inner product. Instead we adopt the usual 'myopia' approach and take an equivalence relation on V which identifies points which differ by an element of W. The equivalence classes are then 'planes parallel to W'. I am going through this construction quickly here under the assumption that it is familiar to most of you, if not you should think about it carefully since we need to do it several times later.

So, if $W \subset V$ is a linear subspace of V we define a relation on V – remember this is just a subset of $V \times V$ with certain properties – by

$$v \sim_W v' \iff v - v' \in W \iff \exists\ w \in W \text{ s.t. } v = v' + w. \tag{1.50}$$

This satisfies the three conditions for an equivalence relation:

(1) $v \sim_W v$

(2) $v \sim_W v' \Longleftrightarrow v' \sim_W v$
(3) $v \sim_W v',\ v' \sim_W v'' \Longrightarrow v \sim_W v''$

which means that we can regard it as a 'coarser notion of equality.'

Then V/W is the set of equivalence classes with respect to $\sim_W$. You can think of the elements of V/W as being of the form $v + W$ – a particular element of V plus an arbitrary element of W. Then of course $v' \in v + W$ if and only if $v' - v \in W$ meaning $v \sim_W v'$.

The crucial point here is that

$$V/W \text{ is a vector space.} \tag{1.51}$$

You should check the details – see Problem 1.2. Note that the 'is' in (1.51) should really be expanded to 'is in a natural way' since as usual the linear structure is inherited from V :

$$\lambda(v + W) = \lambda v + W,\ (v_1 + W) + (v_2 + W) = (v_1 + v_2) + W. \tag{1.52}$$

The subspace W appears as the origin in V/W.

Now, two cases of this are of special interest to us.

Proposition 1.5. *If $\|\cdot\|$ is a seminorm on V then*

$$E = \{v \in V; \|v\| = 0\} \subset V \tag{1.53}$$

is a linear subspace and

$$\|v + E\|_{V/E} = \|v\| \tag{1.54}$$

defines a norm on V/E.

Proof. That E is linear follows from the properties of a seminorm, since $\|\lambda v\| = |\lambda|\|v\|$ shows that $\lambda v \in E$ if $v \in E$ and $\lambda \in \mathbb{K}$. Similarly the triangle inequality shows that $v_1 + v_2 \in E$ if $v_1,\, v_2 \in E$.

To check that (1.54) defines a norm, first we need to check that it makes sense as a function $\|\cdot\|_{V/E} \longrightarrow [0, \infty)$. This amounts to the statement that $\|v'\|$ is the same for all elements $v' = v + e \in v + E$ for a fixed v. This however follows from the triangle inequality applied twice:

$$\|v'\| \le \|v\| + \|e\| = \|v\| \le \|v'\| + \|-e\| = \|v'\|. \tag{1.55}$$

Now, I leave you the exercise of checking that $\|\cdot\|_{V/E}$ is a norm, see Problem 1. □

The second application is more serious, but in fact we will not use it for some time so I usually do not do this in lectures at this stage.

Proposition 1.6. *If $W \subset V$ is a closed subspace of a normed space then*

$$\|v + W\|_{V/W} = \inf_{w \in W} \|v + w\|_V \tag{1.56}$$

defines a norm on V/W; if V is a Banach space then so is V/W.

For the proof see Problems 1 and 1.

6. Completion

A normed space not being complete, not being a Banach space, is considered to be a defect which we might, indeed will, wish to rectify.

Let V be a normed space with norm $\|\cdot\|_V$. A *completion* of V is a Banach space B with the following properties:-

(1) There is an injective (i.e. 1-1) linear map $I : V \longrightarrow B$
(2) The norms satisfy

$$\|I(v)\|_B = \|v\|_V \ \forall \ v \in V. \tag{1.57}$$

(3) The range $I(V) \subset B$ is dense in B.

Notice that if V is itself a Banach space then we can take $B = V$ with I the identity map.

So, the main result is:

Theorem 1.1. *Each normed space has a completion.*

There are several ways to prove this, we will come across a more sophisticated one (using the Hahn-Banach Theorem) later. In the meantime I will describe two proofs. In the first the fact that any metric space has a completion in a similar sense is recalled and then it is shown that the linear structure extends to the completion. A second, 'hands-on', proof is also outlined with the idea of motivating the construction of the Lebesgue integral – which is in our near future.

Proof 1. One of the neater proofs that any metric space has a completion is to use Lemma 1.2. Pick a point in the metric space of interest, $p \in M$, and then define a map

$$M \ni q \longmapsto f_q \in \mathcal{C}_\infty(M),\ f_q(x) = d(x, q) - d(x, p) \ \forall \ x \in M. \tag{1.58}$$

That $f_q \in \mathcal{C}_\infty(M)$ is straightforward to check. It is bounded (because of the second term) by the reverse triangle inequality

$$|f_q(x)| = |d(x, q) - d(x, p)| \le d(p, q)$$

and is continuous, as the difference of two continuous functions. Moreover the distance between two functions in the image is

$$\sup_{x\in M} |f_q(x) - f_{q'}(x)| = \sup_{x\in M} |d(x, q) - d(x, q')| = d(q, q') \tag{1.59}$$

using the reverse triangle inequality (and evaluating at $x = q$). Thus the map (1.58) is well-defined, injective and even distance-preserving. Since $\mathcal{C}^0_\infty(M)$ is complete, the closure of the image of (1.58) is a complete metric space, X, in which M can be identified as a dense subset.

Now, in case that $M = V$ is a normed space this all goes through. The disconcerting thing is that the map $q \longrightarrow f_q$ is *not* linear. Nevertheless, we can give X a linear structure so that it becomes a Banach space in which V is a dense linear subspace. Namely for any two elements $f_i \in X$, $i = 1, 2$, define

$$\lambda_1 f_1 + \lambda_2 f_2 = \lim_{n\to\infty} f_{\lambda_1 p_n + \lambda_2 q_n} \tag{1.60}$$

where p_n and q_n are sequences in V such that $f_{p_n} \to f_1$ and $f_{q_n} \to f_2$. Such sequences exist by the construction of X and the result does not depend on the choice of sequence – since if p'_n is another choice in place of p_n then $f_{p'_n} - f_{p_n} \to 0$ in X (and similarly for q_n). So the element of the left in (1.60) is well-defined. All of the properties of a linear space and normed space now follow by continuity from $V \subset X$ and it also follows that X is a Banach space (since a closed subset of a complete space is complete). Unfortunately there are quite a few annoying details to check! □

'PROOF 2' (THE LAST BIT IS LEFT TO YOU). Let V be a normed space. First we introduce the rather large space

$$\widetilde{V} = \left\{ \{u_k\}_{k=1}^{\infty}; u_k \in V \text{ and } \sum_{k=1}^{\infty} \|u_k\| < \infty \right\} \tag{1.61}$$

the elements of which, if you recall, are said to be absolutely summable. Notice that the elements of $\widetilde{V}$ are *sequences*, valued in V so two sequences are equal, are the same, only when each entry in one is equal to the corresponding entry in the other – no shifting around or anything is permitted as far as equality is concerned. We think of these as series (remember this means nothing except changing the name, a series is a sequence and a sequence is a series), the only difference is that we 'think' of taking the limit of a sequence but we 'think' of summing the elements of a series, whether we can do so or not being a different matter.

Now, each element of $\widetilde{V}$ is a Cauchy sequence – meaning the corresponding sequence of partial sums $v_N = \sum\limits_{k=1}^{N} u_k$ is Cauchy if $\{u_k\}$ is absolutely summable. As noted earlier, this is simply because if $M \geq N$ then

$$\|v_M - v_N\| = \|\sum_{j=N+1}^{M} u_j\| \leq \sum_{j=N+1}^{M} \|u_j\| \leq \sum_{j \geq N+1} \|u_j\| \tag{1.62}$$

gets small with N by the assumption that $\sum\limits_j \|u_j\| < \infty$.

Moreover, $\widetilde{V}$ is a linear space, where we add sequences, and multiply by constants, by doing the operations on each component:-

$$t_1\{u_k\} + t_2\{u'_k\} = \{t_1 u_k + t_2 u'_k\}. \tag{1.63}$$

This always gives an absolutely summable series by the triangle inequality:

$$\sum_k \|t_1 u_k + t_2 u'_k\| \leq |t_1| \sum_k \|u_k\| + |t_2| \sum_k \|u'_k\|. \tag{1.64}$$

Within $\widetilde{V}$ consider the linear subspace

$$S = \left\{ \{u_k\}; \sum_k \|u_k\| < \infty, \ \sum_k u_k = 0 \right\} \tag{1.65}$$

of those which sum to 0. As discussed in Section 5 above, we can form the quotient

$$B = \widetilde{V}/S \tag{1.66}$$

the elements of which are the 'cosets' of the form $\{u_k\} + S \subset \widetilde{V}$ where $\{u_k\} \in \widetilde{V}$. This is our completion, we proceed to check the following properties of this B.

(1) A norm on B (via a seminorm on $\tilde{V}$) is defined by

$$\|b\|_B = \lim_{n \to \infty} \|\sum_{k=1}^{n} u_k\|, \ b = \{u_k\} + S \in B. \tag{1.67}$$

(2) The original space V is imbedded in B by

$$V \ni v \longmapsto I(v) = \{u_k\} + S,\ u_1 = v,\ u_k = 0\ \forall\ k > 1 \tag{1.68}$$

and the norm satisfies (1.57).

(3) $I(V) \subset B$ is dense.

(4) B is a Banach space with the norm (1.67).

So, first that (1.67) is a norm. The limit on the right does exist since the limit of the norm of a Cauchy sequence always exists – namely the sequence of norms is itself Cauchy but now in $\mathbb{R}$. Moreover, adding an element of S to $\{u_k\}$ does not change the norm of the sequence of partial sums, since the additional term tends to zero in norm. Thus $\|b\|_B$ is well-defined for each element $b \in B$ and $\|b\|_B = 0$ means exactly that the sequence $\{u_k\}$ used to define it tends to 0 in norm, hence is in S hence $b = 0$ in B. The other two properties of norm are reasonably clear, since if b, $b' \in B$ are represented by $\{u_k\}$, $\{u'_k\}$ in $\widetilde{V}$ then tb and $b + b'$ are represented by $\{tu_k\}$ and $\{u_k + u'_k\}$ and

(1.69)

$$\lim_{n\to\infty} \|\sum_{k=1}^{n} tu_k\| = |t| \lim_{n\to\infty} \|\sum_{k=1}^{n} u_k\|, \Longrightarrow \|tb\| = |t|\|b\|$$

$$\lim_{n\to\infty} \|\sum_{k=1}^{n} (u_k + u'_k)\| = A \Longrightarrow$$

$$\text{for } \epsilon > 0\ \exists\ N \text{ s.t. }\ \forall\ n \ge N,\ A - \epsilon \le \|\sum_{k=1}^{n} (u_k + u'_k)\| \Longrightarrow$$

$$A - \epsilon \le \|\sum_{k=1}^{n} u_k\| + \|\sum_{k=1}^{n} u'_k)\|\ \forall\ n \ge N \Longrightarrow A - \epsilon \le \|b\|_B + \|b'\|_B\ \forall\ \epsilon > 0 \Longrightarrow$$

$$\|b + b'\|_B \le \|b\|_B + \|b'\|_B.$$

Now the norm of the element $I(v) = v, 0, 0, \cdots$, is the limit of the norms of the sequence of partial sums and hence is $\|v\|_V$ so $\|I(v)\|_B = \|v\|_V$ and $I(v) = 0$ therefore implies $v = 0$ and hence I is also injective.

We need to check that B is complete, and also that $I(V)$ is dense. Here is an extended discussion of the difficulty – of course maybe you can see it directly yourself (or have a better scheme). Note that I suggest that you to write out your own version of it carefully in Problem 1.

Okay, what does it mean for B to be a Banach space, as discussed above it means that every absolutely summable series in B is convergent. Such a series $\{b_n\}$ is given by $b_n = \{u_k^{(n)}\} + S$ where $\{u_k^{(n)}\} \in \widetilde{V}$ and the summability condition is that

$$\infty > \sum_n \|b_n\|_B = \sum_n \lim_{N\to\infty} \|\sum_{k=1}^{N} u_k^{(n)}\|_V. \tag{1.70}$$

So, we want to show that $\sum\limits_n b_n = b$ converges, and to do so we need to find the limit b. It is supposed to be given by an absolutely summable series. The 'problem' is that this series should look like $\sum\limits_n \sum\limits_k u_k^{(n)}$ in some sense – because it is supposed to represent the sum of the b_n's. Now, it would be very nice if we had the estimate

$$\sum_n \sum_k \|u_k^{(n)}\|_V < \infty \tag{1.71}$$

since this should allow us to break up the double sum in some nice way so as to get an absolutely summable series out of the whole thing. The trouble is that (1.71) need not hold. We know that *each* of the sums over k – for given n – converges, but not the sum of the sums. All we know here is that the sum of the 'limits of the norms' in (1.70) converges.

So, that is the problem! One way to see the solution is to note that we do not have to choose the original $\{u_k^{(n)}\}$ to 'represent' b_n – we can add to it any element of S. One idea is to rearrange the $u_k^{(n)}$ – I am thinking here of fixed n – so that it 'converges even faster.' I will not go through this in full detail but rather do it later when we need the argument for the completeness of the space of Lebesgue integrable functions. Given $\epsilon > 0$ we can choose p_1 so that for all $p \ge p_1$,

$$|\|\sum_{k\le p} u_k^{(n)}\|_V - \|b_n\|_B| \le \epsilon, \ \sum_{k\ge p} \|u_k^{(n)}\|_V \le \epsilon. \tag{1.72}$$

Then in fact we can choose successive $p_j > p_{j-1}$ (remember that little n is fixed here) so that

$$|\|\sum_{k\le p_j} u_k^{(n)}\|_V - \|b_n\|_B| \le 2^{-j}\epsilon, \ \sum_{k\ge p_j} \|u_k^{(n)}\|_V \le 2^{-j}\epsilon \ \forall \ j. \tag{1.73}$$

Now, 'resum the series' defining instead $v_1^{(n)} = \sum\limits_{k=1}^{p_1} u_k^{(n)}$, $v_j^{(n)} = \sum\limits_{k=p_{j-1}+1}^{p_j} u_k^{(n)}$ and do this setting $\epsilon = 2^{-n}$ for the nth series. Check that now

$$\sum_n \sum_k \|v_k^{(n)}\|_V < \infty. \tag{1.74}$$

Of course, you should also check that $b_n = \{v_k^{(n)}\} + S$ so that these new summable series work just as well as the old ones.

After this fiddling you can now try to find a limit for the sequence as

$$b = \{w_k\} + S, \ w_k = \sum_{l+p=k+1} v_l^{(p)} \in V. \tag{1.75}$$

So, you need to check that this $\{w_k\}$ is absolutely summable in V and that $b_n \to b$ as $n \to \infty$.

Finally then there is the question of showing that $I(V)$ is dense in B. You can do this using the same idea as above – in fact it might be better to do it first. Given an element $b \in B$ we need to find elements in V, v_k such that $\|I(v_k) - b\|_B \to 0$ as $k \to \infty$. Take an absolutely summable series u_k representing b and take $v_j = \sum\limits_{k=1}^{N_j} u_k$ where the p_j's are constructed as above and check that $I(v_j) \to b$ by computing

$$\|I(v_j) - b\|_B = \lim_{\to\infty} \|\sum_{k>p_j} u_k\|_V \le \sum_{k>p_j} \|u_k\|_V. \tag{1.76}$$

□

7. More examples

Let me collect some examples of normed and Banach spaces. Those mentioned above and in the problems include:

- c_0 the space of convergent sequences in $\mathbb{C}$ with supremum norm, a Banach space.
- l^p one space for each real number $1 \le p < \infty$; the space of p-summable series with corresponding norm; all Banach spaces. The most important of these for us is the case $p = 2$, which is (a) Hilbert space.
- l^∞ the space of bounded sequences with supremum norm, a Banach space with $c_0 \subset l^\infty$ as a closed subspace with the same norm.
- $\mathcal{C}([a, b])$ or more generally $\mathcal{C}(M)$ for any compact metric space M – the Banach space of continuous functions with supremum norm.
- $\mathcal{C}_\infty(\mathbb{R})$, or more generally $\mathcal{C}_\infty(M)$ for any metric space M – the Banach space of bounded continuous functions with supremum norm.
- $\mathcal{C}_0(\mathbb{R})$, or more generally $\mathcal{C}_0(M)$ for any metric space M – the Banach space of continuous functions which 'vanish at infinity' (see Problem 1) with supremum norm. A closed subspace, with the same norm, in $\mathcal{C}^0_\infty(M)$.
- $\mathcal{C}^k([a, b])$ the space of k times continuously differentiable (so $k \in \mathbb{N}$) functions on $[a, b]$ with norm the sum of the supremum norms on the function and its derivatives. Each is a Banach space – see Problem 1.
- The space $\mathcal{C}([0, 1])$ with norm

$$\|u\|_{L^1} = \int_0^1 |u| dx \tag{1.77}$$

  given by the Riemann integral of the absolute value. A normed space, but not a Banach space. We will construct the concrete completion, $L^1([0, 1])$ of Lebesgue integrable 'functions'.
- The space $\mathcal{R}([a, b])$ of Riemann integrable functions on $[a, b]$ with $\|u\|$ defined by (1.77). This is only a seminorm, since there are Riemann integrable functions (note that u Riemann integrable does imply that $|u|$ is Riemann integrable) with $|u|$ having vanishing Riemann integral but which are not identically zero. This cannot happen for continuous functions. So the quotient is a normed space, but it is not complete.
- The same spaces – either of continuous or of Riemann integrable functions but with the (semi- in the second case) norm

$$\|u\|_{L^p} = \left(\int_a^b |u|^p \right)^{\frac{1}{p}}. \tag{1.78}$$

  Not complete in either case even after passing to the quotient to get a norm for Riemann integrable functions. We can, and indeed will, define $L^p(a, b)$ as the completion of $\mathcal{C}([a, b])$ with respect to the L^p norm. However we will get a concrete realization of it soon.
- Suppose $0 < \alpha < 1$ and consider the subspace of $\mathcal{C}([a, b])$ consisting of the 'Hölder continuous functions' with exponent α, that is those $u : [a, b] \longrightarrow \mathbb{C}$ which satisfy

$$|u(x) - u(y)| \le C|x - y|^\alpha \text{ for some } C \ge 0. \tag{1.79}$$

Note that this already implies the continuity of u. As norm one can take the sum of the supremum norm and the 'best constant' which is the same as

$$\|u\|_{\mathcal{C}^\alpha} = \sup_{x\in[a,b]|} |u(x)| + \sup_{x\neq y\in[a,b]} \frac{|u(x)-u(y)|}{|x-y|^\alpha}; \tag{1.80}$$

it is a Banach space usually denoted $\mathcal{C}^\alpha([a,b])$.

- Note the previous example works for $\alpha = 1$ as well, then it is not denoted $\mathcal{C}^1([a,b])$, since that is the space of once continuously differentiable functions; this is the space of Lipschitz functions $\Lambda([a,b])$ – again it is a Banach space.
- We will also talk about Sobolev spaces later. These are functions with 'Lebesgue integrable derivatives'. It is perhaps not easy to see how to define these, but if one takes the norm on $\mathcal{C}^1([a,b])$

$$\|u\|_{H^1} = \left(\|u\|_{L^2}^2 + \|\frac{du}{dx}\|_{L^2}^2\right)^{\frac{1}{2}} \tag{1.81}$$

and completes it, one gets the Sobolev space $H^1([a,b])$ – it is a Banach space (and a Hilbert space). In fact it is a subspace of $\mathcal{C}([a,b]) = \mathcal{C}^0([a,b])$.

Here is an example to see that the space of continuous functions on $[0,1]$ with norm (1.77) is not complete; things are even worse than this example indicates! It is a bit harder to show that the quotient of the Riemann integrable functions is not complete, feel free to give it a try.

Take a simple non-negative continuous function on $\mathbb{R}$ for instance

$$f(x) = \begin{cases} 1-|x| & \text{if } |x| \le 1 \\ 0 & \text{if } |x| > 1. \end{cases} \tag{1.82}$$

Then $\int_{-1}^1 f(x) = 1$. Now scale it up and in by setting

$$f_N(x) = Nf(N^3x) = 0 \text{ if } |x| > N^{-3}. \tag{1.83}$$

So it vanishes outside $[-N^{-3}, N^{-3}]$ and has $\int_{-1}^1 f_N(x)dx = N^{-2}$. It follows that the sequence $\{f_N\}$ is absolutely summable with respect to the integral norm in (1.77) on $[-1,1]$. The pointwise series $\sum\limits_N f_N(x)$ converges everywhere except at $x = 0$ – since at each point $x \neq 0$, $f_N(x) = 0$ if $N^3|x| > 1$. The resulting function, even if we ignore the problem at $x = 0$, is not Riemann integrable because it is not bounded.

You might respond that the sum of the series is 'improperly Riemann integrable'. This is true but does not help much.

It is at this point that I start doing Lebesgue integration in the lectures. The following material is from later in the course but fits here quite reasonably.

8. Baire's theorem

At least once I wrote a version of the following material on the blackboard during the first mid-term test, in an an attempt to distract people. It did not work very well – its seems that MIT students have already been toughened up by this stage. Baire's theorem will be used later (it is also known as 'Baire category theory' although it has nothing to do with categories in the modern sense).

This is a theorem about complete metric spaces – it could be included in the earlier course 'Real Analysis' but the main applications are in Functional Analysis.

THEOREM 1.2 (Baire). *If M is a non-empty complete metric space and $C_n \subset M$, $n \in \mathbb{N}$, are closed subsets such that*

$$M = \bigcup_n C_n \tag{1.84}$$

then at least one of the C_n's has an interior point.

PROOF. We can assume that the first set $C_1 \neq \emptyset$ since they cannot all be empty and dropping any empty sets does no harm. Let's assume the contrary of the desired conclusion, namely that each of the C_n's has empty interior, hoping to arrive at a contradiction to (1.84) using the other properties. This means that an open ball $B(p, \epsilon)$ around a point of M (so it isn't empty) cannot be contained in any one of the C_n.

So, choose $p \in C_1$. Now, there must be a point $p_1 \in B(p, 1/3)$ which is not in C_1. Since C_1 is closed there exists $\epsilon_1 > 0$, and we can take $\epsilon_1 < 1/3$, such that $B(p_1, \epsilon_1) \cap C_1 = \emptyset$. Continue in this way, choose $p_2 \in B(p_1, \epsilon_1/3)$ which is not in C_2 and $\epsilon_2 > 0$, $\epsilon_2 < \epsilon_1/3$ such that $B(p_2, \epsilon_2) \cap C_2 = \emptyset$. Here we use both the operating hypothesis that C_2 has empty interior and the fact that it is closed. So, inductively there is a sequence p_i, $i = 1, \dots, k$ and positive numbers $0 < \epsilon_k < \epsilon_{k-1}/3 < \epsilon_{k-2}/3^2 < \cdots < \epsilon_1/3^{k-1} < 3^{-k}$ such that $p_j \in B(p_{j-1}, \epsilon_{j-1}/3)$ and $B(p_j, \epsilon_j) \cap C_j = \emptyset$. Then we can add another p_{k+1} by using the properties of C_{k+1} – it has empty interior so there is some point in $B(p_k, \epsilon_k/3)$ which is not in C_{k+1} and then $B(p_{k+1}, \epsilon_{k+1}) \cap C_{k+1} = \emptyset$ where $\epsilon_{k+1} > 0$ but $\epsilon_{k+1} < \epsilon_k/3$. Thus, we have a sequence $\{p_k\}$ in M. Since $d(p_{k+1}, p_k) < \epsilon_k/3$ this is a Cauchy sequence, in fact

$$d(p_k, p_{k+l}) < \epsilon_k/3 + \cdots + \epsilon_{k+l-1}/3 < 3^{-k}. \tag{1.85}$$

Since M is complete the sequence converges to a limit, $q \in M$. Notice however that $p_l \in B(p_k, 2\epsilon_k/3)$ for all $k > l$ so $d(p_k, q) \leq 2\epsilon_k/3$ which implies that $q \notin C_k$ for any k. This is the desired contradiction to (1.84).

Thus, at least one of the C_n must have non-empty interior. □

In applications one might get a complete metric space written as a countable union of subsets

$$M = \bigcup_n E_n, \ E_n \subset M \tag{1.86}$$

where the E_n are not necessarily closed. We can still apply Baire's theorem however, just take $C_n = \overline{E_n}$ to be the closures – then of course (1.84) holds since $E_n \subset C_n$. The conclusion from (1.86) for a complete M is

(1.87) For at least one n the closure of E_n has non-empty interior.

9. Uniform boundedness

One application of Baire's theorem is often called the *uniform boundedness principle* or Banach-Steinhaus Theorem.

Theorem 1.3 (Uniform boundedness). *Let B be a Banach space and suppose that T_n is a sequence of bounded (i.e. continuous) linear operators $T_n : B \longrightarrow V$ where V is a normed space. Suppose that for each $b \in B$ the set $\{T_n(b)\} \subset V$ is bounded (in norm of course) then* $\sup_n \|T_n\| < \infty$.

Proof. This follows from a pretty direct application of Baire's theorem to B. Consider the sets

$$S_p = \{b \in B,\ \|b\| \le 1,\ \|T_n b\|_V \le p\ \forall\ n\},\ p \in \mathbb{N}. \tag{1.88}$$

Each S_p is closed because T_n is continuous, so if $b_k \to b$ is a convergent sequence in S_p then $\|b\| \le 1$ and $\|T_n(b)\| \le p$. The union of the S_p is the whole of the closed ball of radius one around the origin in B :

$$\{b \in B; d(b,0) \le 1\} = \bigcup_p S_p \tag{1.89}$$

because of the assumption of 'pointwise boundedness' – each b with $\|b\| \le 1$ must be in one of the S_p's.

So, by Baire's theorem one of the sets S_p has non-empty interior, it therefore contains a closed ball of positive radius around some point. Thus for some p, some $v \in S_p$, and some $\delta > 0$,

$$w \in B,\ \|w\|_B \le \delta \Longrightarrow \|T_n(v+w)\|_V \le p\ \forall\ n. \tag{1.90}$$

Since $v \in S_p$ is fixed it follows that $\|T_n w\| \le \|T_n v\| + p \le 2p$ for all w with $\|w\| \le \delta$.

Moving v to $(1-\delta/2)v$ and halving δ as necessary it follows that this ball $B(v,\delta)$ is contained in the open ball around the origin of radius 1. Thus, using the triangle inequality, and the fact that $\|T_n(v)\|_V \le p$ this implies

$$w \in B,\ \|w\|_B \le \delta \Longrightarrow \|T_n(w)\|_V \le 2p \Longrightarrow \|T_n\| \le 2p/\delta. \tag{1.91}$$

The norm of the operator is $\sup\{\|Tw\|_V; \|w\|_B = 1\} = \frac{1}{\delta}\sup\{\|Tw\|_V; \|w\|_B = \delta\}$ so the norms are uniformly bounded:

$$\|T_n\| \le 2p/\delta \tag{1.92}$$

as claimed. □

10. Open mapping theorem

The second major application of Baire's theorem is to

Theorem 1.4 (Open Mapping). *If $T : B_1 \longrightarrow B_2$ is a bounded and surjective linear map between two Banach spaces then T is open:*

$$T(O) \subset B_2 \text{ is open if } O \subset B_1 \text{ is open.} \tag{1.93}$$

This is 'wrong way continuity' and as such can be used to prove the continuity of inverse maps as we shall see. The proof uses Baire's theorem pretty directly, but then another similar sort of argument is needed to complete the proof. Note however that the proof is considerably simplified if we assume that B_1 is a Hilbert space. There are more direct but more computational proofs, see Problem 1. I prefer this one because I have a reasonable chance of remembering the steps.

PROOF. What we will try to show is that the image under T of the unit open ball around the origin, $B(0,1) \subset B_1$ contains an open ball around the origin in B_2. The first part, of the proof, using Baire's theorem shows that the *closure* of the image, so in B_2, has 0 as an interior point – i.e. it contains an open ball around the origin in B_2 :

$$\overline{T(B(0,1)} \supset B(0,\delta),\ \delta > 0. \tag{1.94}$$

To see this we apply Baire's theorem to the sets

$$C_p = \mathrm{cl}_{B_2} T(B(0,p)) \tag{1.95}$$

the closure of the image of the ball in B_1 of radius p. We know that

$$B_2 = \bigcup_p T(B(0,p)) \tag{1.96}$$

since that is what surjectivity means – every point is the image of something. Thus one of the closed sets C_p has an interior point, v. Since T is surjective, $v = Tu$ for some $u \in B_1$. The sets C_p increase with p so we can take a larger p and v is still an interior point, from which it follows that $0 = v - Tu$ is an interior point as well. Thus indeed

$$C_p \supset B(0,\delta) \tag{1.97}$$

for some $\delta > 0$. Rescaling by p, using the linearity of T, it follows that with δ replaced by δ/p, we get (1.94).

If we assume that B_1 is a Hilbert space then (1.94) shows that if $v \in B_2$, $\|v\| < \delta$ there is a sequence u_n with $\|u_n\| \le 1$ and $Tu_n \to v$. As a bounded sequence u_n has a weakly convergent subsequence, $u_{n_j} \rightharpoonup u$, where we know this implies $\|u\| \le 1$ and $Au_{n_j} \rightharpoonup Au = v$ since $Au_n \to v$. This strengthens (1.94) to

$$T(B(0,1) \supset B(0,\delta/2)$$

and proves that T is an open map.

If B_1 is a Banach space but not a Hilbert space we need to work a little harder. Having applied Baire's thereom, consider now what (1.94) means. It follows that each $v \in B_2$, with $\|v\| = \delta$, is the limit of a sequence Tu_n where $\|u_n\| \le 1$. What we want to find is such a sequence, u_n, which converges. To do so we need to choose the sequence more carefully. Certainly we can stop somewhere along the way and see that

$$v \in B_2,\ \|v\| = \delta \Longrightarrow \exists\ u \in B_1,\ \|u\| \le 1,\ \|v - Tu\| \le \frac{\delta}{2} = \frac{1}{2}\|v\| \tag{1.98}$$

where of course we could replace $\frac{\delta}{2}$ by any positive constant but the point is the last inequality is now relative to the norm of v. Scaling again, if we take any $v \neq 0$ in B_2 and apply (1.98) to $v/\|v\|$ we conclude that (for $C = p/\delta$ a fixed constant)

$$v \in B_2 \Longrightarrow \exists\ u \in B_1,\ \|u\| \le C\|v\|,\ \|v - Tu\| \le \frac{1}{2}\|v\| \tag{1.99}$$

where the size of u only depends on the size of v; of course this is also true for $v = 0$ by taking $u = 0$.

Using this we construct the desired *better* approximating sequence. Given $w \in B_1$, choose $u_1 = u$ according to (1.99) for $v = w = w_1$. Thus $\|u_1\| \le C$, and $w_2 = w_1 - Tu_1$ satisfies $\|w_2\| \le \frac{1}{2}\|w\|$. Now proceed by induction, supposing that we have constructed a sequence u_j, $j < n$, in B_1 with $\|u_j\| \le C2^{-j+1}\|w\|$

and $\|w_j\| \le 2^{-j+1}\|w\|$ for $j \le n$, where $w_j = w_{j-1} - Tu_{j-1}$ – which we have for $n = 1$. Then we can choose u_n, using (1.99), so $\|u_n\| \le C\|w_n\| \le C2^{-n+1}\|w\|$ and such that $w_{n+1} = w_n - Tu_n$ has $\|w_{n+1}\| \le \frac{1}{2}\|w_n\| \le 2^{-n}\|w\|$ to extend the induction. Thus we get a sequence u_n which is absolutely summable in B_1, since $\sum\limits_n \|u_n\| \le 2C\|w\|$, and hence converges by the assumed completeness of B_1 this time. Moreover

$$w - T(\sum_{j=1}^{n} u_j) = w_1 - \sum_{j=1}^{n}(w_j - w_{j+1}) = w_{n+1} \tag{1.100}$$

so $Tu = w$ and $\|u\| \le 2C\|w\|$.

Thus finally we have shown that each $w \in B(0,1)$ in B_2 is the image of some $u \in B_1$ with $\|u\| \le 2C$. Thus $T(B(0,3C)) \supset B(0,1)$. By scaling it follows that the image of any open ball around the origin contains an open ball around the origin.

Now, the linearity of T shows that the image $T(O)$ of any open set is open, since if $w \in T(O)$ then $w = Tu$ for some $u \in O$ and hence $u + B(0,\epsilon) \subset O$ for $\epsilon > 0$ and then $w + B(0,\delta) \subset T(O)$ for $\delta > 0$ sufficiently small. □

One important corollary of this is something that seems like it should be obvious, but definitely needs completeness to be true.

Corollary 1.2. *If $T : B_1 \longrightarrow B_2$ is a bounded linear map between Banach spaces which is 1-1 and onto, i.e. is a bijection, then it is a homeomorphism – meaning its inverse, which is necessarily linear, is also bounded.*

Proof. The only confusing thing is the notation. Note that T^{-1} is generally used both for the inverse, when it exists, and also to denote the inverse map on sets even when there is no true inverse. The inverse of T, let's call it $S : B_2 \longrightarrow B_1$, is certainly linear. If $O \subset B_1$ is open then $S^{-1}(O) = T(O)$, since to say $v \in S^{-1}(O)$ means $S(v) \in O$ which is just $v \in T(O)$, is open by the Open Mapping theorem, so S is continuous. □

11. Closed graph theorem

For the next application you should check, it is one of the problems, that the product of two Banach spaces, $B_1 \times B_2$, – which is just the linear space of all pairs (u, v), $u \in B_1$ and $v \in B_2$, is a Banach space with respect to the sum of the norms

$$\|(u, v)\| = \|u\|_1 + \|v\|_2. \tag{1.101}$$

Theorem 1.5 (Closed Graph). *If $T : B_1 \longrightarrow B_2$ is a linear map between Banach spaces then it is bounded if and only if its graph*

$$\mathrm{Gr}(T) = \{(u, v) \in B_1 \times B_2; v = Tu\} \tag{1.102}$$

is a closed subset of the Banach space $B_1 \times B_2$.

Proof. Suppose first that T is bounded, i.e. continuous. A sequence $(u_n, v_n) \in B_1 \times B_2$ is in $\mathrm{Gr}(T)$ if and only if $v_n = Tu_n$. So, if it converges, then $u_n \to u$ and $v_n = Tu_n \to Tv$ by the continuity of T, so the limit is in $\mathrm{Gr}(T)$ which is therefore closed.

Conversely, suppose the graph is closed. This means that viewed as a normed space in its own right it is complete. Given the graph we can reconstruct the map it comes from (whether linear or not) in a little diagram. From $B_1 \times B_2$ consider the two projections, $\pi_1(u,v) = u$ and $\pi_2(u,v) = v$. Both of them are continuous since the norm of either u or v is less than the norm in (1.101). Restricting them to $\mathrm{Gr}(T) \subset B_1 \times B_2$ gives

$$\begin{tikzcd} & \mathrm{Gr}(T) \arrow[dl, shift left, "\pi_1"] \arrow[dr, "\pi_2"] & \\ B_1 \arrow[ur, shift left, "S"] \arrow[rr, "T"] & & B_2. \end{tikzcd} \tag{1.103}$$

This little diagram commutes. Indeed there are two ways to map a point $(u,v) \in \mathrm{Gr}(T)$ to B_2, either directly, sending it to v or first sending it to $u \in B_1$ and then to Tu. Since $v = Tu$ these are the same.

Now, as already noted, $\mathrm{Gr}(T) \subset B_1 \times B_2$ is a closed subspace, so it too is a Banach space and π_1 and π_2 remain continuous when restricted to it. The map π_1 is 1-1 and onto, because each u occurs as the first element of precisely one pair, namely $(u, Tu) \in \mathrm{Gr}(T)$. Thus the Corollary above applies to π_1 to show that its inverse, S is continuous. But then $T = \pi_2 \circ S$, from the commutativity, is also continuous proving the theorem. $\square$

12. Hahn-Banach theorem

Now, there is always a little pressure to state and prove the Hahn-Banach Theorem. This is about extension of functionals. Stately starkly, the basic question is: Does a normed space have *any* non-trivial continuous linear functionals on it? That is, is the dual space always non-trivial (of course there is always the zero linear functional but that is not very amusing). We do not really encounter this problem since for a Hilbert space, or even a pre-Hilbert space, there is always the space itself, giving continuous linear functionals through the pairing – Riesz' Theorem says that in the case of a Hilbert space that is all there is. If you are following the course then at this point you should also see that the only continuous linear functionals on a pre-Hilbert space correspond to points in the completion. I could have used the Hahn-Banach Theorem to show that any normed space has a completion, but I gave a more direct argument for this, which was in any case much more relevant for the cases of $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ for which we wanted *concrete* completions.

THEOREM 1.6 (Hahn-Banach). *If $M \subset V$ is a linear subspace of a normed space and $u : M \longrightarrow \mathbb{C}$ is a linear map such that*

$$|u(t)| \le C\|t\|_V \ \forall\ t \in M \tag{1.104}$$

then there exists a bounded linear functional $U : V \longrightarrow \mathbb{C}$ with $\|U\| \le C$ and $U\big|_M = u$.

First, by computation, we show that we can extend any continuous linear functional 'a little bit' without increasing the norm.

LEMMA 1.5. *Suppose $M \subset V$ is a subspace of a normed linear space, $x \notin M$ and $u : M \longrightarrow \mathbb{C}$ is a bounded linear functional as in (1.104) then there exists $u' : M' \longrightarrow \mathbb{C}$, where $M' = \{t' \in V; t' = t + ax,\ a \in \mathbb{C}\}$, such that*

$$u'\big|_M = u,\ |u'(t+ax)| \le C\|t+ax\|_V,\ \forall\ t \in M,\ a \in \mathbb{C}. \tag{1.105}$$

Proof. Note that the decompositon $t' = t + ax$ of a point in M' is unique, since $t + ax = \tilde{t} + \tilde{a}x$ implies $(a - \tilde{a})x \in M$ so $a = \tilde{a}$, since $x \notin M$ and hence $t = \tilde{t}$ as well. Thus

$$u'(t + ax) = u'(t) + au(x) = u(t) + \lambda a,\ \lambda = u'(x) \tag{1.106}$$

and all we have at our disposal is the choice of λ. Any choice will give a linear functional extending u, the problem of course is to arrange the continuity estimate without increasing the constant C. In fact if $C = 0$ then $u = 0$ and we can take the zero extension. So we might as well assume that $C = 1$ since dividing u by C arranges this and if u' extends u/C then Cu' extends u and the norm estimate in (1.105) follows. So we now assume that

$$|u(t)| \le \|t\|_V\ \forall\ t \in M. \tag{1.107}$$

We want to choose λ so that

$$|u(t) + a\lambda| \le \|t + ax\|_V\ \forall\ t \in M,\ a \in \mathbb{C}. \tag{1.108}$$

Certainly when $a = 0$ this represents no restriction on λ. For $a \neq 0$ we can divide through by $-a$ and (1.108) becomes

$$|a||u(-\frac{t}{a}) - \lambda| = |u(t) + a\lambda| \le \|t + ax\|_V = |a|\| - \frac{t}{a} - x\|_V \tag{1.109}$$

and since $-t/a \in M$ we only need to arrange that

$$|u(t) - \lambda| \le \|t - x\|_V\ \forall\ t \in M \tag{1.110}$$

and the general case will follow by reversing the scaling.

A complex linear functional such as u can be recovered from its real part, as we see below, so set

$$w(t) = \mathrm{Re}(u(t)),\ |w(t)| \le \|t\|_V\ \forall\ t \in M. \tag{1.111}$$

We proceed to show the real version of the Lemma, that w can be extended to a linear functional $w' : M + \mathbb{R}x \longrightarrow \mathbb{R}$ if $x \notin M$ without increasing the norm. The same argument as above shows that the only freedom is the choice of $\lambda = w'(x)$ and we need to choose $\lambda \in \mathbb{R}$ so that

$$|w(t) - \lambda| \le \|t - x\|_V\ \forall\ t \in M. \tag{1.112}$$

The norm estimate on w shows that

$$|w(t_1) - w(t_2)| \le |u(t_1) - u(t_2)| \le \|t_1 - t_2\| \le \|t_1 - x\|_V + \|t_2 - x\|_V. \tag{1.113}$$

Writing this out using the reality we find

$$\begin{gathered} w(t_1) - w(t_2) \le \|t_1 - x\|_V + \|t_2 - x\|_V \Longrightarrow \\ w(t_1) - \|t_1 - x\| \le w(t_2) + \|t_2 - x\|_V\ \forall\ t_1,\ t_2 \in M. \end{gathered} \tag{1.114}$$

We can then take the supremum on the left and the infimum on the right and choose λ in between – namely we have shown that there exists $\lambda \in \mathbb{R}$ with

$$\begin{aligned} w(t) - \|t - x\|_V \le \sup_{t_2 \in M} (w(t_1) - \|t_1 - x\|) \le \lambda \\ \le \inf_{t_2 \in M} (w(t_1) + \|t_1 - x\|) \le w(t) + \|t - x\|_V\ \forall\ t \in M. \end{aligned} \tag{1.115}$$

This in turn implies that

$$-\|t - x\|_V \le -w(t) + \lambda \le \|t - x\|_V \Longrightarrow |w(t) - \lambda| \le \|t - x\|_V\ \forall\ t \in M. \tag{1.116}$$

So we have an extension of w to a real functional $w' : M + \mathbb{R}x \longrightarrow \mathbb{R}$ with $|w'(t + ax)| \leq \|t + ax\|_V$ for all $a \in \mathbb{R}$. We can repeat this argument to obtain a further extension $w'' : M + \mathbb{C}x = M + \mathbb{R}x + \mathbb{R}(ix) \longrightarrow \mathbb{R}$ without increasing the norm.

Now we find the desired extension of u by setting

$$u'(t + cx) = w''(t + ax + b(ix)) - iw''(it - bx + a(ix)) : M + \mathbb{C}x \longrightarrow \mathbb{C}. \tag{1.117}$$

This is clearly linear over the reals and linearity over complex coefficients follows since

$$\begin{aligned} u'(it + icx) &= w''(it + -bx + a(ix))) - iw''(-t - ax - b(ix)) \\ &= i(w''(t + ax + b(ix) - iw''(it + -bx + a(ix))) = iu'(t + cx). \end{aligned} \tag{1.118}$$

The uniqueness of a complex linear functional with given real part also shows that $u'\big|_M = u$.

Finally, to estimate the norm of u' notice that for each $t \in M$ and $c \in \mathbb{C}$ there is a unique $\theta \in [0, 2\pi)$ such that

$$|u'(t + cx)| = \operatorname{Re} e^{i\theta} u'(t + cx) = w''(e^{i\theta} t + e^{i\theta} cx) \leq \|e^{i\theta} t + e^{i\theta} cx\|_V = \|t + cx\|_V. \tag{1.119}$$

This completes the proof of the Lemma. □

Proof of Hahn-Banach. This is an application of Zorn's Lemma. I am not going to get into the derivation of Zorn's Lemma from the Axiom of Choice, but if you believe the latter – and you are advised to do so, at least before lunchtime – you should believe the former.

Zorn's Lemma is a statement about partially ordered sets. A partial order on a set E is a subset of $E \times E$, so a relation, where the condition that (e, f) be in the relation is written $e \prec f$ and it must satisfy

$$e \prec e, \ e \prec f \text{ and } f \prec e \Longrightarrow e = f, \ e \prec f \text{ and } f \prec g \Longrightarrow e \prec g. \tag{1.120}$$

So, the missing ingredient between this and an order is that two elements need not be related at all, either way.

A subset of a partially ordered set inherits the partial order and such a subset is said to be a *chain* if each pair of its elements *is* related one way or the other. An *upper bound* on a subset $D \subset E$ is an element $e \in E$ such that $d \prec e$ for all $d \in D$. A *maximal* element of E is one which is not majorized, that is $e \prec f$, $f \in E$, implies $e = f$.

Lemma 1.6 (Zorn). *If every chain in a (non-empty) partially ordered set has an upper bound then the set contains at least one maximal element.*

So, we are just accepting this Lemma as axiomatic. However, make sure that you appreciate that it is true for countable sets. Namely if C is countable and has no maximal element then it must contain a chain which has no upper bound. To see this, write C as a sequence $\{c_i\}_{i \in \mathbb{N}}$. Then $x_1 = c_1$ is not maximal so there exists some c_k, $k > 1$, with $c_1 \prec c_k$ in terms of the order in C. From the properties of $\mathbb{N}$ it follows that there is a *smallest* $k = k_2$ such that c_k has this property, but $k_2 > 1$. Let this be $x_2 = c_{k_2}$ and proceed in the same way – $x_3 = c_{k_3}$ where $k_3 > k_2$ is the smallest such integer for which $c_{k_2} \prec c_{k_3}$. Assuming C is infinite in the first place this grinds out an infinite chain x_i. Now you can check that this cannot have an upper bound because every element of C is either one of these, and so cannot be an upper bound, or else it is c_j with $k_l < j < k_{l+1}$ for some l and then it is not greater than x_l.

The point of Zorn's Lemma is precisely that it applies to uncountable sets.

We are given a functional $u : M \longrightarrow \mathbb{C}$ defined on some linear subspace $M \subset V$ of a normed space where u is bounded with respect to the induced norm on M. We will apply Zorn's Lemma to the set E consisting of *all* extensions (v, N) of u with the same norm; it is generally non-countable. That is,

$$V \supset N \supset M, \ v\big|_M = u \text{ and } \|v\|_N = \|u\|_M.$$

This is certainly non-empty since it contains (u, M) and has the natural partial order that $(v_1, N_1) \prec (v_2, N_2)$ if $N_1 \subset N_2$ and $v_2\big|_{N_1} = v_1$. You should check that this is a partial order.

Let C be a chain in this set of extensions. Thus for any two elements $(v_i, N_i) \in C$, $i = 1, 2$, either $(v_1, N_1) \prec (v_2, N_2)$ or the other way around. This means that

$$\tilde{N} = \bigcup\{N; (v, N) \in C \text{ for some } v\} \subset V \tag{1.121}$$

is a linear space. Note that this union need not be countable, or anything like that, but any two elements of $\tilde{N}$ are each in one of the N's and one of these must be contained in the other by the chain condition. Thus each pair of elements of $\tilde{N}$ is actually in a common N and hence so is their linear span. Similarly we can define an extension

$$\tilde{v} : \tilde{N} \longrightarrow \mathbb{C}, \ \tilde{v}(x) = v(x) \text{ if } x \in N, \ (v, N) \in C. \tag{1.122}$$

There may be many pairs $(v, N) \in C$ satisfying $x \in N$ for a given x but the chain condition implies that $v(x)$ is the same for all of them. Thus $\tilde{v}$ is well defined, and is clearly also linear, extends u and satisfies the norm condition $|\tilde{v}(x)| \le \|u\|_M \|x\|_V$. Thus $(\tilde{v}, \tilde{N})$ is an upper bound for the chain C.

So, the set of all extension E, with the norm condition, satisfies the hypothesis of Zorn's Lemma, so must – at least in the mornings – have a maximal element $(\tilde{u}, \tilde{M})$. If $\tilde{M} = V$ then we are done. However, in the contary case there exists $x \in V \setminus \tilde{M}$. This means we can apply our little lemma and construct an extension $(u', \tilde{M}')$ of $(\tilde{u}, \tilde{M})$ which is therefore also an element of E and satisfies $(\tilde{u}, \tilde{M}) \prec (u', \tilde{M}')$. This however contradicts the condition that $(\tilde{u}, \tilde{M})$ be maximal, so is forbidden by Zorn. □

There are many applications of the Hahn-Banach Theorem. As remarked earlier, one significant one is that the dual space of a non-trivial normed space is itself non-trivial.

Proposition 1.7. *For any normed space V and element $0 \neq v \in V$ there is a continuous linear functional $f : V \longrightarrow \mathbb{C}$ with $f(v) = 1$ and $\|f\| = 1/\|v\|_V$.*

Proof. Start with the one-dimensional space, M, spanned by v and define $u(zv) = z$. This has norm $1/\|v\|_V$. Extend it using the Hahn-Banach Theorem and you will get a continuous functional f as desired. □

13. Double dual

Let me give another application of the Hahn-Banach theorem, although I have generally not covered this in lectures. If V is a normed space, we know its dual space, V', to be a Banach space. Let $V'' = (V')'$ be the dual of the dual.

PROPOSITION 1.8. *If $v \in V$ then the linear map on V' :*

$$T_v : V' \longrightarrow \mathbb{C},\ T_v(v') = v'(v) \tag{1.123}$$

is continuous and this defines an isometric linear injection $V \hookrightarrow V''$, $\|T_v\| = \|v\|$.

PROOF. The definition of T_v is 'tautologous', meaning it is almost the definition of V'. First check T_v in (1.123) is linear. Indeed, if v_1', $v_2' \in V'$ and λ_1, $\lambda_2 \in \mathbb{C}$ then $T_v(\lambda_1 v_1' + \lambda_2 v_2') = (\lambda_1 v_1' + \lambda_2 v_2')(v) = \lambda_1 v_1'(v) + \lambda_2 v_2'(v) = \lambda_1 T_v(v_1') + \lambda_2 T_v(v_2')$. That $T_v \in V''$, i.e. is bounded, follows too since $|T_v(v')| = |v'(v)| \le \|v'\|_{V'}\|v\|_V$; this also shows that $\|T_v\|_{V''} \le \|v\|$. On the other hand, by Proposition 1.7 above, if $\|v\| = 1$ then there exists $v' \in V'$ such that $v'(v) = 1$ and $\|v'\|_{V'} = 1$. Then $T_v(v') = v'(v) = 1$ shows that $\|T_v\| = 1$ so in general $\|T_v\| = \|v\|$. It also needs to be checked that $V \ni v \longmapsto T_v \in V''$ is a linear map – this is clear from the definition. It is necessarily 1-1 since $\|T_v\| = \|v\|$. □

Now, it is definitely not the case in general that $V'' = V$ in the sense that this injection is also a surjection. Since V'' is always a Banach space, one necessary condition is that V itself should be a Banach space. In fact the closure of the image of V in V'' is a completion of V. If the map to V'' is a bijection then V is said to be *reflexive*. It is pretty easy to find examples of non-reflexive Banach spaces, the most familiar is c_0 – the space of infinite sequences converging to 0. Its dual can be identified with l^1, the space of summable sequences. Its dual in turn, the bidual of c_0, is the space l^∞ of bounded sequences, into which the embedding is the obvious one, so c_0 is not reflexive. In fact l^1 is not reflexive either. There are useful characterizations of reflexive Banach spaces. You may be interested enough to look up James' Theorem:- A Banach space is reflexive if and only if every continuous linear functional on it attains its supremum on the unit ball.

14. Axioms of a vector space

In case you missed out on one of the basic linear algebra courses, or have a poor memory, here are the axioms of a vector space over a field $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$ for us).

A *vector space structure* on a set V is a pair of maps

$$+ : V \times V \longrightarrow V,\ \cdot : \mathbb{K} \times V \longrightarrow V \tag{1.124}$$

satisfying the conditions listed below. These maps are written $+(v_1, v_2) = v_1 + v_2$ and $\cdot(\lambda, v) = \lambda v$, $\lambda \in \mathbb{K}$, V, v_1, $v_2 \in V$.

additive commutativity $v_1 + v_2 = v_2 + v_1$ for all v_1, $v_2 \in V$.
additive associativity $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ for all v_1, v_2, $v_3 \in V$.
existence of zero There is an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$.
additive invertibility For each $v \in V$ there exists $w \in V$ such that $v + w = 0$.
distributivity of scalar additivity $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ for all λ_1, $\lambda_2 \in \mathbb{K}$ and $v \in V$.
multiplicativity $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$ for all λ_1, $\lambda_2 \in \mathbb{K}$ and $v \in V$.
action of multiplicative identity $1v = v$ for all $v \in V$.
distributivity of space additivity $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$ for all $\lambda \in \mathbb{K}$ v_1, $v_2 \in V$.