Functional Analysis (under construction)

Gustav Holzegel*

March 21, 2015

Contents

*Imperial College London, Department of Mathematics, South Kensington Campus, London SW7 2AZ, United Kingdom.

1 Motivation and Literature

Our basic object of study will be infinite dimensional vector spaces (I will assume you are familiar with the *finite dimensional* case from linear algebra) which carry an additional structure, for instance a norm or an inner product whose properties will allow us to do analysis on these spaces, i.e. to talk about limits, completeness, orthogonality etc.

In typical applications (and this is also how the subject developed historically) the space X will be a space of functions, for instance the space of continuous functions on an interval equipped with the maximum-norm.

We will then study linear maps between these normed spaces. An instructive example is to think of a linear partial differential equation which you would like to interpret as a linear map between two function spaces. Take for example the Laplace operator mapping C^2-functions to C^0-functions. The question "when can I solve $\Delta u = f$" for given f can then be translated into the question: When can I invert a certain linear operator between two function spaces. It is precisely with these kinds of questions in mind that the subject "functional analysis" was developed.

As prerequisites I will only assume basic linear algebra and analysis. In particular, I will try to avoid measure theory and review the necessary topological concepts as we move along.

I found the following books useful in preparing the course

1. E. Kreyszig, Introductory Functional Analysis with Applications (Wiley)
2. M. Schechter, Principles of Functional Analysis (AMS)
3. A. Friedman, Foundations of Modern Analysis (Dover)
4. E. Stein, Real Analysis (Princeton Lectures in Analysis)
5. E. Stein, Functional Analysis (Princeton Lectures in Analysis)
6. H. Brezis, Functional Analysis, Sobolev Spaces and PDEs (Springer)
7. L. Evans, Partial Differential Equations (AMS)

For the most part, the present notes follow [Kreyszig] rather closely. The more advanced material on the Fredholm alternative and PDE applications is taken from [Evans].

2 Metric Spaces

Definition 2.1. *A metric space is a pair (X, d) where X is a set and d is a* <u>*metric*</u> *on X, i.e. a function $d : X \times X \to \mathbb{R}$ such that for all $x, y, z \in X$ we have*

1. *symmetry: $d(x, y) = d(y, x)$*
2. *positivity: $d(x, y) \geq 0$ with equality if and only if $x = y$.*
3. *triangle inequality $d(x, y) \leq d(x, z) + d(z, y)$*

2.1 Examples

We give a couple of examples of metric spaces which we shall revisit at various points in the course.

(1) The most familiar example is Euclidean space $X = \mathbb{R}^2$ with distance $d(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2}$ for $x = (x^1, x^2)$ and $y = (y^1, y^2)$ being points in the plane (check triangle inequality + geometric interpretation).

(2) The space $X = l^\infty$. This space is defined as the space of all *bounded* sequences of complex numbers, i.e. every element ("point") $x \in X$ is a sequence $x = (\xi_1, \xi_2, \ldots)$, shorthand $x = (\xi_j)$, such that $|\xi_j| \leq c_x$ where c_x is independent of j but will of course depend on the particular sequence chosen. The metric is defined as

$$d(x,y) = \sup_{j \in \mathbb{N}} |\xi_j - \eta_j| \quad \text{for } x = (\xi_j) \text{ and } y = (\eta_j)$$

(3) The space $X = C[a,b]$. This is the space of continuous functions $f : [a,b] \to \mathbb{R}$ with metric

$$d_{uni}(f,g) = \max_{t \in [a,b]} |f(t) - g(t)|.$$

or alternatively, with metric

$$d_{L^1}(f,g) = \int_a^b |f(t) - g(t)| dt$$

(4) The sequence space $X = s$. This is the set of *all* (bounded or unbounded) sequences of complex numbers with distance

$$d(x,y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|}$$

Here the triangle inequality is non-trivial. We leave it as an exercise but give the following hint. Observe that the function

$$f(t) = \frac{t}{1+t}$$

is increasing and hence that $|a+b| \leq |a| + |b|$ implies $f(|a+b|) \leq f(|a| + |b|)$.

(5) The space $X = \ell^p$ for $1 \leq p < \infty$. This space consists of sequences $x = (\xi_j)$ such that

$$\sum_{j=1}^{\infty} |\xi_j|^p < \infty$$

with distance

$$d(x,y) = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{(1/p)}$$

Here it is a priori not even clear that any two points x, y have finite distance, i.e. that $x - y$ is in the space ℓ^p. Also the triangle inequality is non-trivial unless $p = 1$ when it follows from the triangle inequality for real numbers. Again we leave the details as an exercise giving only the main steps (assuming $p > 1$ by the preceding comment).

- Step 1: Prove $\alpha \cdot \beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$ for $\alpha, \beta \in \mathbb{R}^+$ and p, q such that $\frac{1}{p} + \frac{1}{q} = 1$.
- Step 2: Let $(\tilde{\xi}_j)$, $(\tilde{\eta}_j)$ be such that $\sum |\tilde{\xi}_j|^p = 1$ and $\sum |\tilde{\eta}_j|^q = 1$. Use Step 1 to conclude
$$\sum_{j=1}^{\infty} |\tilde{\xi}_j \tilde{\eta}_j| \leq 1.$$

Now pick arbitrary (ξ_j) and (η_q) such that $\sum |\xi_j|^p < \infty$ and $\sum |\eta_j|^q < \infty$ and conclude *Hölder's inequality* (recall $1/p + 1/q = 1$ and $p > 1$)

$$\sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \left(\sum_{k=1}^{\infty} |\xi_k|^p \right)^{(1/p)} \left(\sum_{k=1}^{\infty} |\eta_k|^q \right)^{(1/q)} \tag{1}$$

- Step 3: Prove Minkowski's inequality

$$\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{\frac{1}{p}} \tag{2}$$

This shows both that the sum of two element of ℓ^p is again in ℓ^p (hence in particular the distance is finite) and the triangle inequality. To prove (2) note that it is immediate for $p = 1$ and hence assume $p > 1$. Then start from

$$|\xi_j + \eta_j|^p = |\xi_j + \eta_j| |\xi_j + \eta_j|^{p-1} \leq |\xi_j| \cdot |\xi_j + \eta_j|^{p-1} + |\eta_j| \cdot |\xi_j + \eta_j|^{p-1}$$

and sum from 1 to n, apply Hölder's inequality and take the limit $n \to \infty$.

(6) The discrete metric space. In the preceding examples, the underlying spaces had a linear structure. While this will be frequently the case in applications, the definition of a metric space does not require this. For instance, we can take any set X and on it the metric

$$d(x, x) = 0 \quad \text{and} \quad d(x, y) = 1 \;\; \text{for } x \neq y$$

2.2 Topological refresher

The metric induces a topology on X. We have

$$B_r(\xi) := \{\eta \in X : d(\xi, \eta) < r\} \quad \text{the open ball of radius } r \text{ around } \xi \tag{3}$$

$$\overline{B}_r(\xi) := \{\eta \in X : d(\xi, \eta) \leq r\} \quad \text{the closed ball of radius } r \text{ around } \xi \tag{4}$$

- A subset M of a metric space (X, d) is called *open* if it contains a ball about each of its points. A subset K of X is called *closed* if $K^c = X \setminus K$ is open.
- A set $M \subset X$ is called *bounded* if there exists a ball such that $M \subset B$.
- For a subset M of X, a point $x_0 \in X$ (not necessarily in M) is called an *accumulation point* (or limit point) of M if every ball around x_0 contains at least one element $y \in M$ with $y \neq x_0$.
- For a set $M \subset X$ the set $\overline{M}$ is the set consisting of M and all of its accumulation points. The set $\overline{M}$ is called the *closure* of M. It is the smallest closed set which contains M.
- A set $M \subset X$ is called dense in X if $\overline{M} = X$. The metric space X is called *separable* if it has a countable subset which is dense in X.

You may also wish to recall the notion of *interior point* and *continuity*.

Example 2.1. *The space ℓ^p with $1 \leq p < \infty$ is separable. To see this let M be the set of all sequences y of the form*

$$y = (\eta_1, \eta_2, ..., \eta_n, 0, 0, ...) \qquad \text{with } \eta_j \in \mathcal{Q}.$$

Clearly M is countable. We show that it is dense in ℓ^p. Let $x = (\xi_j)$ be an arbitrary sequence in ℓ^p and $\epsilon > 0$ be given. Then, clearly there is a $k \in \mathbb{N}$ such that

$$\sum_k^\infty |\xi_j|^p \leq \frac{1}{2}\epsilon^p$$

as the tail of the convergent series goes to zero. On the other hand, for the first $k-1$ entries of ξ_j we can find rational numbers η_j such that

$$\sum_{j=1}^{k-1} |\xi_j - \eta_j|^p \leq \frac{1}{2}\epsilon^p \, .$$

It follows that the sequence $y = (\eta_1, ..., \eta_{j-1}, 0, ...)$ satisfies

$$d\,(x, y) < \epsilon \, ,$$

which shows that every small ball around $x \in \ell^p$ contains a $y \in M$ with $y \neq x$. Hence M is dense in ℓ^p.

Exercise 2.1. *Show that ℓ^∞ is not separable.*

2.3 Completeness

Let (X, d) be a metric space.

Definition 2.2. *A sequence (x_n) in X is said to converge if there is an $x \in X$ such that*

$$\lim_{n \to \infty} d\,(x_n, x) = 0 \, .$$

The x is called the limit of (x_n), written $\lim_{n\to\infty} x_n = x$ or $x_n \longrightarrow x$.

Lemma 2.1. *A convergent sequence in X is bounded and its limit is unique.*

Definition 2.3. *A sequence (x_n) is said to be Cauchy if for every $\epsilon > 0$ there exists an $N = N\,(\epsilon)$ such that*

$$d\,(x_m, x_n) < \epsilon \qquad \text{for all } m, n > N$$

The metric space (X, d) is called complete if every Cauchy sequence in X converges (to an element of X).

Note that every convergent sequence is a Cauchy sequence (why?) but not the other way around. Indeed one can just take $\mathbb{R} \setminus \{0\}$ and $x_n = 1/n$.

Completeness is an extremely nice property (think about approximating the solution of a problem by a sequence (of functions); completeness allows to conclude that one converges to an actual solution of the problem). We revisit our examples above.

(1) $\mathbb{R}^2$ with the Euclidean metric is complete.

(2) ℓ^∞ is complete.

(3) $C[a,b]$ is complete with the uniform metric but incomplete with respect to the L^1-metric.

(4) The space s is complete.

(5) The space ℓ^p is complete.

(6) The discrete space with discrete metric is complete.

The proof of these claims is similar in each case, essentially reducing things to completeness of $\mathbb{R}$. We present the proof for (3) and (5) and leave the other examples for the exercises.

To see the claim for (5), we consider a Cauchy sequence x_n in ℓ^p with $1 \leq p < \infty$, so $x_n = (\xi_1^n, \xi_2^n, ...)$. For every $\epsilon > 0$ there is an N such that

$$d(x_m, x_n) = \left[\sum_{j=1}^{\infty} |\xi_j^m - \xi_j^n|^p\right]^{1/p} < \epsilon\,. \tag{5}$$

This implies that for every j we must have $|\xi_j^m - \xi_j^n| < \epsilon$. So for fixed j we have a Cauchy sequence in $\mathbb{C}$ which (in view of the completeness of $\mathbb{C}$) converges. So $\xi_j^n \to \xi_j$ as $n \to \infty$. It remains to show that the sequence $x = (\xi_1, \xi_2, ...)$ composed out of these limits is in ℓ^p and that indeed $x_n \to x$. From (5) we have for any k and $m > N$

$$\left[\sum_{j=1}^{k} |\xi_j^m - \xi_j|^p\right]^{1/p} \leq \epsilon\,, \tag{6}$$

as follows from taking the limit $n \to \infty$. Now taking $k \to \infty$ we see that $x_m - x = (\xi_j^m) - \xi_j \in \ell^p$, which since $x_m \in \ell^p$ implies that x is in ℓ^p via Minkowski's inequality. The statement that $x_m \to x$ in ℓ^p is precisely (6) with $k = \infty$.

To see the first claim in (3) let f_n be a Cauchy sequence with respect to the sup-norm, i.e. for $\epsilon > 0$ we can find N_ϵ such that

$$\max_t |f_m(t) - f_n(t)| < \epsilon \qquad \text{for } m, n > N_\epsilon\,.$$

Since for every fixed $t \in [a,b]$ we have a Cauchy sequence in $\mathbb{R}$, we can define pointwise $f(t) = \lim_{n\to\infty} f_n(t)$. We have to show f is continuous and $f_n \to f$ with respect to the uniform metric. Taking $n \to \infty$, we have

$$\max_t |f_m(t) - f(t)| \leq \epsilon$$

which shows that $f_m(t)$ converges uniformly to f. Since the uniform limit of a sequence of continuous functions is continuous (cf. Problem 4 of Week 1) we have $f \in C[a,b]$ and also $f_m \to f$.

To see the incompleteness claim, consider the sequence of continuous functions

$$f_m(t) = \begin{cases} 0 & \text{if } t \in [0, 1/2], \\ \text{linear} & \text{if } t \in [1/2, 1/2 + 1/n], \\ 1 & \text{if } t \in [1/2 + 1/n, 1]. \end{cases} \tag{7}$$

This is clearly a Cauchy sequence with respect to the L^1-metric (why?). The limit, however, is the discontinuous function which is 0 on $[0, 1/2]$ and 1 on $(1/2, 1]$ (proof?).

2.3.1 Further examples of non-complete metric spaces

For instance the space of polynomials on $[a, b]$ equipped with the uniform metric. The exponential function is not a polynomial but its Taylor series converges uniformly on $[a, b]$.

2.4 Completion of metric spaces

Recall how the incomplete rational line $\mathbb{Q}$ can be completed to $\mathbb{R}$ by adding the missing points through equivalence classes of Cauchy sequences. A similar thing can be done for metric spaces, i.e. an arbitrary incomplete metric space can be completed.

Definition 2.4. *Let (X, d) and $\left(\tilde{X}, \tilde{d}\right)$ be metric spaces. Then*

- *A mapping T of X into $\tilde{X}$ is said to be an isometry if T preserves distances, i.e. if for all $x, y \in X$*

$$\tilde{d}(Tx, Ty) = d(x, y)$$

  where Tx and Ty are the images of x and y respectively.

- *The space X is said to be isometric with $\tilde{X}$ if there exists a bijective isometry of X onto $\tilde{X}$.*

Theorem 2.1. *For a metric space (X, d) there exists a complete metric space $\hat{X} = \left(\hat{X}, \hat{d}\right)$ which has a subspace W that is isometric with X and is dense in $\hat{X}$. This space $\hat{X}$ is unique modulo isometries (i.e. if there is another $\left(\tilde{X}, \tilde{d}\right)$ with dense subspace $\tilde{W}$ isometric to X, then $\tilde{X}$ and $\hat{X}$ are isometric).*

Proof. omitted, see [Kreyszig] □

3 Normed Spaces and Banach Spaces

3.1 Linear Algebra refresher

Let X be a vector space over $\mathbb{R}$ or $\mathbb{C}$. We review some definitions from linear algebra with emphasis on the infinite dimensional case.

Definition 3.1. *A set $M \subset X$ is called linearly independent if every finite subset of M is linearly independent.*[1]

[1] For a finite subset $x_1, ..., x_n$ linear independence means that $\lambda_1 x_1 + \ldots \lambda_n x_n = 0$ implies $\lambda_i = 0$ for all i.

Definition 3.2. *A set $E \subset X$ is called a Hamel basis of X if E is linearly independent and every vector $\xi \in X$ can be written uniquely as a finite linear combination of elements in E.*

Note the inherent finiteness properties built into both of these definitions.

Using Zorn's lemma (which we will discuss in the context of the Hahn-Banach theorem, cf. Section 5.1.1 and Theorem 5.4) one can prove

Theorem 3.1. *Let $M \subset X$ be linearly independent. Then X has a Hamel basis which contains M.*

In particular, every vector space has a Hamel basis. Finally, we need the notion of "dimension":

If for any integer n there exist n linearly independent elements, we say that X is infinite dimensional. If there exists an integer such that there are n linearly independent elements but not $n + 1$ we call X finite dimensional of dimension n and write $\dim X = n$.

3.2 Definition and elementary properties

Definition 3.3. *A norm on a vector space X is a real valued function $\|\cdot\| \to \mathbb{R}$ (whose value at $x \in X$ we denote by $\|x\|$) which has the properties*

1. $\|x\| > 0$ *with equality iff* $x = 0$
2. $\|\alpha x\| = |\alpha| \|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$

The definition of the norm is of course motivated by Euclidean space and the length of a vector. Note that a norm defines a metric d on X via

$$d(x, y) = \|x - y\| . \tag{8}$$

This is called the metric induced by the norm. (You should check that this is indeed a metric and you will equally easily prove the following Lemma.)

Lemma 3.1. *A metric induced by a norm satisfies*

1. $d(x + a, y + a) = d(x, y)$
2. $d(\alpha x, \alpha y) = |\alpha| \; d(x, y)$

Definition 3.4. *A normed (linear) space X is a vector space with a norm defined on it. A complete normed space is called a Banach space (completeness being understood with respect to the metric induced by the norm).*

Let us pause for a moment and recapitulate. In section 2 we studied general metric spaces. They were not required to have any linear structure (although most examples did!), i.e. it didn't a-priori make sense to add and scalar multiply elements. The normed spaces are a special kind of metric spaces, namely those which have in addition a vector space structure (i.e. it *does* make sense to add and scalar multiply elements) AND whose metric is special in that it arises from a norm. Note that the sequence space s (example (4)) is an example of a metric space which has a linear structure but whose metric does not arise from a norm (Why? − use Lemma 3.1).[2]

Recall that a metric induces a topology on the space (i.e. it defines a notion of "open set", "closed set" and "continuity"). The following Lemma shows that the marriage of linear algebra (vector space) and analysis (metric space with metric induced by a norm) inherent in Definition 3.4 is a happy one in the sense that the vector space operations of addition and scalar multiplication are continuous.

Lemma 3.2. *Let $(X, \|\cdot\|)$ be a normed space. Then*

- *The linear operations are continuous.*
- *The map $\xi \mapsto \|\xi\|$ is continuous.*

Proof. Addition is a function $f_+ : X \times X \to X$ defined by $f_+(x, y) = x + y$. Continuity at (x, y) is equivalent to $\lim_{n\to\infty} f_+(x_n, y_n) = \lim_{n\to\infty}(x_n + y_n) = x + y$ for any sequences x_n, y_n with $x_n \to x$ and $y_n \to y$ in X. But this follows from

$$\|x_n + y_n - x - y\| \le \|x_n - x\| + \|y_n - y\| .$$

Similarly for scalar multiplication, continuity follows from

$$\|\lambda x_n - \lambda x\| \le |\lambda| \|x_n - x\| .$$

For the second claim, use the reverse triangle triangle inequality

$$\Big| \|\xi\| - \|\xi_n\| \Big| \le \|\xi - \xi_n\|$$

to see that $\xi_n \to \xi$ implies $\|\xi_n\| \to \|\xi\|$, □

Remark 3.1. *The fact that the algebraic and the analytical structure are compatible in this sense is essential to obtain an interesting mathematical object. From this point of view, we could have started this section from the most abstract object satisfying this condition: a* topological vectorspace. *This is a vector space equipped with a topology (not necessarily arising from a metric) such that the vector space operations are continuous. However, for this course, normed linear spaces will be the object of study.*

We can now revisit our examples and identify (1), (2), (3a) and (5) as Banach spaces with the norms (=$\|x\| = \sqrt{(x_1)^2 + (x_2)^2}$ for (1), $\|x\| = \sup_j |\xi_j|$ for (2), $\|f\| = \max_{t\in[a,b]} |f(t)|$ for (3a) and $\|x\| = \left(\sum |\xi_j|^p\right)^{1/p}$ for (5). (What about examples (3b) and (6)?)

We also remark that incomplete normed spaces can be completed similarly to Theorem 2.1 about metric spaces. (Now one also need to show that $\hat{X}$ has a vector space structure.)

Clearly via (8) one can talk about convergence in normed spaces: $x_n \to x$ means $\lim_{n\to\infty} \|x_n - x\| = 0$ and a for $\{x_n\}$ Cauchy one can find for any $\epsilon > 0$ an N such that $\|x_m - x_n\| < \epsilon$ holds for all $m, n > N$.

[2] The space s is a so-called Frechét space, i.e. a complete metric linear space (i.e. a metric space with linear structure such that vector-operations are continuous) with translation invariant (i.e. satisfying (1) in Lemma 3.1) metric.

In a normed space we can also add elements which allows us to define series. If x_k is a sequence in X we can associate with it the sequence of partial sums

$$s_n = x_1 + \ldots + x_n$$

We have $s_n \to s$ if $\sum_{i=1}^{\infty} x_i$ converges with sum $s = \sum_{i=1}^{\infty} x_i$. If $\sum_{i=1}^{\infty} \|x_i\|$ converges, the series is said to converge absolutely. In a normed space absolute convergence implies convergence if and only if X is complete (see Problem 5 of the Week 1 Exercises).

Definition 3.5. *If a normed space X contains a sequence e_n with the property that for every $x \in X$ there is a unique sequence of scalars α_n such that*

$$\|x - \alpha_1 e_1 - \alpha_2 e_2 - \ldots \alpha_n e_n\| \to 0 \qquad \text{as } n \to \infty$$

then (e_n) is called a Schauder basis of X.

As an example, you can revisit Example 2.1 to construct a Schauder basis for ℓ^p.

Exercise 3.1. *Show that if X has a Schauder basis, then it is separable. (The converse is in general false.)*

3.3 Finite vs the Infinite dimensional normed spaces

Our next goal is to understand better in what way infinite dimensional normed spaces are different from finite dimensional ones. You may recall that for $\mathbb{R}^n$ you proved that the closed and bounded subsets are compact and vice versa ("Heine-Borel property"). Now, while the direction "compact $\implies$ closed and bounded" holds in a general metric space (so in particular for infinite dimensional normed spaces), cf. Lemma 3.4, in this section we will prove that if a normed space X is finite dimensional then "closed and bounded $\implies$ compact" and conversely that if the unit ball is compact in a normed space X, then X has to be finite dimensional.[3]

We start with a technical Lemma capturing the fact that there can be no large vectors without the scalars in a given basis being large at the same time.

Lemma 3.3. *Let $\{x_1, \ldots, x_n\}$ be a linearly independent set of vectors in a normed space (of any dimension). Then there is a number $c > 0$ such that for every choice of scalars $\alpha_1, \ldots, \alpha_n$ we have*

$$\|\alpha_1 x_1 + \ldots + \alpha_n x_n\| \geq c\,(|\alpha_1| + \ldots |\alpha_n|)$$

Proof. Let $s = |\alpha_1| + \ldots |\alpha_n|$. Wlog we assume $s > 0$ as otherwise the statement holds trivially. Dividing the desired inequality by s it suffices to prove that there is a $c > 0$ such that

$$\|\beta_1 x_1 + \ldots + \beta_n x_n\| \geq c \quad \text{for every tuple } (\beta_1, \ldots, \beta_n) \text{ with } \sum |\beta_i| = 1.$$

[3] Therefore, in an infinite dimensional normed space the closed unit ball is always non-compact. On the other hand, there exist compact (hence closed and bounded) sets in an infinite dimensional normed space as we will see in the Exercises.

Suppose this is false. Then there is a sequence

$$y_m = \beta_1^{(m)} x_1 + \ldots + \beta_n^{(m)} x_n \quad \text{with} \sum |\beta_i^{(m)}| = 1$$

with the property that $\|y_m\| \to 0$ as $m \to \infty$. Now clearly $|\beta_i^{(m)}| \leq 1$ holds for every $i = 1, ..., n$. In particular, for $i = 1$, $\beta_1^{(m)}$ is a bounded sequence in $\mathbb{R}$ (or $\mathbb{C}$) which by Bolzano-Weierstrass has a subsequence which converges to some β_1. In this way we obtain a subsequence $y_{1,m}$ which is such that the first component converges to the number β_1. Now we can turn to $\beta_2^{(1m)}$, extract a convergent subsequence such that it converges to some β_2 and obtain a subsequence $y_{2,m}$ such that the first two components converge. After n steps, we have a subsequence $y_{n,m}$ which converges to

$$y_{n,m} \to y := \sum_{j=1}^{n} \beta_j x_j \qquad \text{with} \sum |\beta_j| = 1\,.$$

Since the x_j are linearly independent and not all β_j can be zero we must have $y \neq 0$ and hence $\|y\| \neq 0$. On the other hand, we assumed that $\|y_m\|$ converges to zero. Therefore the subsequence $\|y_{n,m}\|$ also converges to zero, which implies $\|y\| = 0$. Contradiction. □

The lemma is used in the following

Theorem 3.2. *Every finite dimensional subspace Y of a normed space X is complete. In particular, every finite dimensional normed space is complete.*

Proof. Consider (y_m) a Cauchy sequence in Y. We want to show that $y_m \to y \in Y$. To do this, we will use the Lemma + the completeness of $\mathbb{R}$ (or $\mathbb{C}$) to extract a candidate for y. Let $\dim Y = n$ and pick an arbitrary basis $e_1, \ldots, e_n$ for Y. We can express the y_m uniquely as

$$y_m = \alpha_1^{(m)} e_1 + \ldots + \alpha_n^{(m)} e_n\,.$$

The Cauchy property implies that for any $\epsilon > 0$ there is an N such that

$$\epsilon > \Big\| \sum_{j=1}^{n} \left(\alpha_j^{(m)} - \alpha_j^{(k)} \right) e_j \Big\| \geq c \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j^{(k)}| \quad \text{for } m, k > N\,. \tag{9}$$

This implies that $\alpha_j^{(m)}$ is Cauchy in $\mathbb{R}$ (or $\mathbb{C}$) for fixed $j = 1, ..., n$ and by the completeness of the latter we have $\alpha_j^{(m)} \to \alpha_j$ for any j. We define

$$y = \alpha_1 e_1 + ... + \alpha_n e_n$$

which is clearly in Y. Our Cauchy sequence y_m indeed converges to this element since

$$\|y_m - y\| = \Big\| \sum_{j=1}^{n} \left(\alpha_j^{(m)} - \alpha_j \right) e_j \Big\| \leq \sum_{j=1}^{n} |\alpha_j^{(m)} - \alpha_j| \cdot \|e_j\|$$

with the right hand side converging to zero as $m \to \infty$. □

Corollary 3.1. *Every finite dimensional subspace Y of a normed space X is closed in X.*

Proof. We need to show $\overline{Y} = Y$. Let $y \in \overline{Y}$. Take a sequence $y_m \to y$ with $y_m \in Y$. Then in particular y_m is Cauchy in Y and converges to a $y \in Y$ by the completeness of Y (Theorem 3.2). □

Remark 3.2. *Infinite dimensional subspaces need not be closed. For instance, consider the space of continuous functions $X = C[0,1]$ equipped with the max-norm. Let Y be the subspace $Y = span\left(1, t, t^2, \dots\right)$ of polynomials. The exponential function is in $\overline{Y}$ but not in Y.*

As another application of our basic Lemma 3.3, we obtain that on a finite dimensional vector space any norm is equivalent to any other. Equivalence here is defined in the following way:

Definition 3.6. *Let V be a vector space and $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on V.*

- *We say that $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$ if there exists a $C > 0$ such that for all $\xi \in V$ we have $\|\xi\|_2 \leq C\|\xi\|_1$.*
- *We say that $\|\cdot\|_2$ and $\|\cdot\|_1$ are equivalent, if there exists $C, D > 0$ such that $D\|\xi\|_1 \leq \|\xi\|_2 \leq C\|\xi\|_1$.*

In particular, if two norms are equivalent, then a Cauchy sequence with respect to one of the norms is a Cauchy with respect to the other and vice versa. The two norms also induce the same topology (i.e. the notion of an open set is the same in both of the norms). As promised, we have

Theorem 3.3. *On a finite dimensional vector space V any norm $\|\cdot\|_1$ is equivalent to any other norm $\|\cdot\|_2$.*

Proof. Pick an arbitrary basis of V and represent an arbitrary $x \in V$ by $x = \alpha_1 e_1 + \ldots + \alpha_n e_n$. Then we have

$$\|x\|_1 \leq \sum_{j=1}^{n} |\alpha_j| \|e_j\|_1 \leq \max_j \|e_j\|_1 \sum_{j=1}^{n} |\alpha_j| \leq \max_j \|e_j\|_1 \cdot \frac{1}{c} \cdot \|x\|_2 \leq k\|x\|_2$$

with the first step following from the triangle inequality and the third from Lemma 3.3. This shows that $\|x\|_2$ is stronger than $\|x\|_1$ and reversing the roles of the norms in the estimate above yields the equivalence. □

Remark 3.3. *Theorem 3.3 is clearly false in infinite dimensions. We can illustrate this by revisiting $C[0,1]$ equipped either with the max-norm or the L^1-norm,*

$$\|f\|_{uni} = \max_t |f(t)| \qquad or \qquad \|f\|_{L^1} = \int_0^1 |f(t)| dt\,.$$

It is not hard to construct a sequence which has bounded L^1-norm but diverges to infinity in the maximum norm (a peak getting narrower and narrower while keeping the integral finite will do). Is $\|f\|_{uni}$ stronger than $\|f\|_{L^1}$?

As we will see, many problems in partial differential equations (the perhaps richest source of applications of functional analysis) are all about finding the "right" norm (e.g. one in which you can prove convergence of solutions within an approximation scheme for instance).

3.3.1 Compactness

We continue our investigation of the difference between finite and infinite dimensional normed spaces by studying the notion of compactness. We recall the definition for general metric spaces:

Definition 3.7. *A metric space is called (sequentially) compact if every sequence in X has a convergent subsequence. A subset $M \subset X$ is (sequentially) compact if every sequence x_n in M has a subsequence which converges in M.*

You may recall a different definition of compactness, namely "X is compact if every open cover contains a finite subcover". The point is that for metric spaces this definition is equivalent to the above (you can look up the proof in any book on metric spaces, e.g. Theorem 3.5.4 in [Friedman]). For us it will be more convenient to work with sequential formulation.

Lemma 3.4. *Let X be a metric space and $M \subset X$ be compact. Then M is closed and bounded.*

Proof. To show closed we let $x \in \overline{M}$ and choose a sequence $x_m \to x$ with $x_m \in M$. Since M is compact, this sequence has a convergent subsequence $x_{m_k} \to \tilde{x} \in M$. By the uniqueness of the limit $x = \tilde{x}$. To show bounded, assume the contrary, namely that M is not bounded. Then there exists an x_n such that $d(x_n, b) > n$ for any fixed $b \in M$. This sequence does not have a convergent subsequence. Contradiction. $\square$

As mentioned above, the converse is in general false (see the example below) but true for *finite dimensional normed spaces* (see the Theorem below).

Example 3.1. *Consider the closed unit ball $\overline{B}_1$ in the space ℓ^2. Let $e_n = (0, 0, ..., 1, 0, ...)$ (note $\|e_n\|_{\ell^2} = 1$) and (e_n) be a sequence of sequences contained in the closed unit ball $\overline{B}_1 \subset \ell^2$. This sequence does not have a convergent subsequence since $\|e_m - e_n\|_{\ell^2} = \sqrt{2}$ for $m \neq n$.*

Theorem 3.4. *Let $(X, \|\cdot\|)$ be a normed space with $dimX = n < \infty$. Then any subset $M \subset X$ is compact if and only if M is closed an bounded.*

Proof. One of the directions follows immediately from Lemma 3.4. For the other direction we take an arbitrary bounded sequence x_m in M (which is an arbitrary subset) and we would like to show that we can extract a subsequence converging in M. We start by expressing the x_m in terms of an arbitrary basis as

$$x_m = \alpha_1^{(m)} e_1 + \ldots + \alpha_n^{(m)} e_n \, .$$

The sequence is bounded, so $B > \|x_m\| \geq c \sum_{i=1}^n |\alpha_i^{(m)}|$, the last step following from Lemma 3.3. We now – just as in the proof of Lemma 3.3 – use Bolzano-Weierstrass consecutively on the bounded sequences $\alpha_1^{(m)}$, $\alpha_2^{(m)}$, ..., (each of which is a sequence in $\mathbb{R}$ or $\mathbb{C}$) to extract a convergent subsequence $x_{m_k} \to x = \alpha_1 e_1 + \ldots + \alpha_n e_n$. But M is closed and hence contains its limit points, so $x \in M$, which establishes the compactness. $\square$

Having seen that "closed + bounded $\iff$ compact" in finite dimensional normed spaces, we would like to prove that it does not hold in general in infinite dimensional normed spaces. For this we will show below that the closed unit ball in an infinite dimensional normed space is never compact. One can then ask what additional/ alternative conditions to closed and bounded of a subset are necessary to ensure compactness in an infinite dimensional normed space. If you have taken the course on metric spaces, you may remember that "complete + totally bounded $\iff$ compact". You can refresh your memory in Exercise 1 (Week 2) below. The "non-compactness of the unit ball" will follow from the following Lemma, which is due to F. Riesz.

Lemma 3.5. *Let Y and Z be subspaces of a normed space X (of any dimension) and suppose Y is closed and a proper subset of Z. Then for any real number $\theta \in (0,1)$ there is a $z \in Z$ such that*

$$\|z\| = 1 \quad \text{and} \quad \|z-y\| \geq \theta \quad \text{for all } y \in Y \tag{10}$$

Remark 3.4. *Note that we clearly need $z \in Z \setminus Y$ for this to be true. The Lemma says something about finding a point with unit norm which is also a certain distance away from the subspace Y.*

Proof. Let $v \in Z \setminus Y$ and denote the distance from Y by

$$a = \inf_{y \in Y} \|v-y\| \,. \tag{11}$$

We have $a > 0$. Indeed, if $a = 0$ then there would be a sequence y_n such that $\|y_n - v\| \to 0$ which would imply that $y_n \to v$ and $v \in Y$ since Y is closed contradicting that $v \in Z \setminus Y$. Take $\theta \in (0,1)$. By the definition of the inf we can find a $y_0 \in Y$ such that

$$a \leq \|v - y_0\| \leq \frac{a}{\theta} \,. \tag{12}$$

We claim that $z = \frac{v-y_0}{\|v-y_0\|}$ is the desired element: It clearly satisfies $\|z\| = 1$ and moreover, for any $y \in Y$ we have

$$\|z-y\| = \left\| \frac{v - (y_0 + y\|v-y_0\|)}{\|v-y_0\|} \right\| \geq \frac{a}{\|v-y_0\|} \geq \frac{a}{a/\theta} = \theta \,.$$

Here the second step follows from the fact that the expression in the round brackets is in Y and applying (11), while the third step follows from (12). □

Theorem 3.5. *If a normed space has the property that the closed unit ball $\overline{B}_1 = \{x \mid \|x\| \leq 1\}$ is compact, then X is finite dimensional.*

Proof. We assume the contrary, i.e. $\overline{B}_1$ is compact and $\dim X = \infty$. We choose x_1 with $\|x_1\| = 1$ which generates a one-dimensional subspace, which is closed (cf. Corollary 3.1). We apply the Riesz Lemma to find an x_2 (linearly independent of x_1!) with $\|x_2\| = 1$ and such that $\|x_2 - x_1\| \geq \frac{1}{2}$. In the next step $Y = span\,(x_1, x_2)$ is a two-dimensional closed subspace. Again we apply Riesz Lemma to find an x_3 such that $\|x_3\| = 1$ and $\|x_3 - y\| \geq \frac{1}{2}$ for all $y \in Y$, so in particular $\|x_3 - x_2\| \geq \frac{1}{2}$ and $\|x_3 - x_1\| \geq \frac{1}{2}$. Continuing in this fashion we construct a sequence x_n with the property that $\|x_m - x_n\| \geq \frac{1}{2}$ if $m \neq n$, which does not have a convergent subsequence. Contradiction. Hence $\dim X < \infty$. □

Exercise 3.2. *If $\dim Y < \infty$ in Riesz-Lemma, even $\theta = 1$ works (why?). Give an example of the estimate failing for $\theta = 1$ in the case that $\dim Y = \infty$.*

4 Linear Operators

We now move on to study mappings between normed spaces. Those are called "operators".

Definition 4.1. *A linear operator T is a map such that*

1. *the domain $\mathcal{D}(T)$ is a vector space and the range $\mathcal{R}(T)$ lies in a vector space over the same field ($\mathbb{R}$ or $\mathbb{C}$)*
2. *For all $x, y \in \mathcal{D}(T)$ and scalars α we have*

$$T(x+y) = Tx + Ty \quad \text{and} \quad T(\alpha x) = \alpha Tx\,. \tag{13}$$

We see that linear operators preserve the vector space operations, which makes them homomorphisms of vector spaces. The above introduced the notation $\mathcal{D}(T)$ for the domain and $\mathcal{R}(T)$ for the range to which we will stick. We also define the *null-space of* T, which is the set of all $x \in \mathcal{D}(T)$ such that $Tx = 0$, which we will denote $\mathcal{N}(T)$. The word "kernel" will also be used for the null-space.

Examples

1. The identity operator, $I_X : X \to X$, defined by $I_X x = x$ for all $x \in X$.
2. The zero operator, $0 : X \to Y$, defined by $0x = 0$ for all $x \in X$.
3. Differentiation. Let X be the set of all polynomials on $[a, b]$ and $T : X \to X$ be defined by $Tf(t) = f'(t)$.
4. Integration. Let $X = C[a, b]$ and $T : X \to X$ be defined by $Tf(t) = \int_a^t f(s)\,ds$ for $t \in [a, b]$.

Note that Definition 4.1 (and the examples) does not yet involve any norm; only the vector space structure. The following properties are easy to check and left as an exercise

- $\mathcal{R}(T)$ is a vector space
- $\mathcal{N}(T)$ is a vector space
- If $\dim \mathcal{D}(T) = n < \infty$ then $\dim \mathcal{R}(T) \leq n$.

We also recall that $T : \mathcal{D}(T) \to Y$ is called injective if $Tx_1 = Tx_2$ implies $x_1 = x_2$ which is equivalent to $Tx = 0$ implying $x = 0$. If T is injective, then there is an inverse $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ and the inverse is linear.

4.1 Bounded Linear Operators

We will be interested in linear operators on *normed* spaces.

Definition 4.2. *Let X, Y be normed spaces and $T : \mathcal{D}(T) \to Y$ be a linear operator with $\mathcal{D}(T) \subset X$. The operator T is said to be bounded if there exists a $C > 0$ such that*

$$\|Tx\| \leq C\|x\| \qquad \text{for all } x \in \mathcal{D}(T).$$

The norm of the operator is defined as

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

We have abused notation slightly already in that $\|Tx\| = \|Tx\|_Y$ is the norm in the space Y and $\|x\| = \|x\|_X$ is the norm in X. We shall almost always drop the subscripts as there is no danger of confusion. We also note the convention that if the domain consist of only the zero element, then $\|T\| = 0$.

The following two properties are easily verified (Exercise):

$$\|Tx\| \leq \|T\| \|x\| ,$$

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

and so is the following

Lemma 4.1. *The operator norm is indeed a norm.*

We now revisit our examples above:

Examples

1. The identity operator, $I_X : X \to X$ is bounded and $\|I_X\| = 1$.
2. The zero operator, $0 : X \to Y$ is bounded and $\|0\| = 0$.
3. Differentiation. Let X be the set of all polynomials on $[0, 1]$ equipped with the maximum norm and $T : X \to X$ be defined by $Tf(t) = f'(t)$. The sequence $f_n = t^n$ satisfies $\|f_n\| = 1$ (since $f_n(1) = 1$ is the maximum value assumed on $[0, 1]$) and $\|Tf_n\| = n$ which shows that the differentiation operator is *unbounded*.[4]
4. Integration. Let $X = C[0, 1]$ and $T : X \to X$ be defined by $f(t) = \int_0^t f(s)\, ds$ for $t \in [0, 1]$. We have $\|Tf\| = \max_t |\int_0^t f(s)\, ds| \leq \max_t |f| \cdot 1 \leq \|f\|$ and hence the operator is bounded.

We note is passing the following

Theorem 4.1. *If a normed space is finite dimensional then every linear operator T on X is bounded.*

Proof. Let $\dim X = n$ and pick $e_1, ..., e_n$ an arbitrary basis for X. Express $x = \sum \alpha_i e_i$ and since T is linear $Tx = \sum \alpha_i \cdot Te_i$. Clearly,

$$\|Tx\| \leq \sum |\alpha_i| \|Te_i\| \leq \max_i \|Te_i\| \sum |\alpha_i| \leq \max_i \|Te_i\| \frac{1}{c} \|x\|$$

using Lemma 3.3 in the last step. □

We now prove a fundamental fact about linear operators, namely that continuity and boundedness are equivalent concepts.

[4] The importance of this operator suggests that we should also consider *unbounded* operators. We will do so towards the end of the course.

Theorem 4.2. *Let $T : \mathcal{D}(T) \to Y$ be a linear operator where $\mathcal{D}(T) \subset X$ and X, Y are normed spaces. Then*

1. *T is continuous if and only if T is bounded*
2. *If T is continuous at a single point it is continuous everywhere*

Proof. For $T = 0$ there is nothing to show and hence we assume $T \neq 0$, i.e. $\|T\| \neq 0$. We first assume T is bounded and want to show continuity at an arbitrary $x_0 \in \mathcal{D}(T)$. For given $\epsilon > 0$ every $x \in \mathcal{D}(T)$ with $\|x - x_0\| < \frac{\epsilon}{\|T\|} =: \delta$ gets mapped to $\|Tx - Tx_0\| \leq \|T\|\|x - x_0\| = \epsilon$, which is the statement of continuity at x_0. For the other direction, we assume T is continuous at an arbitrary $x_0 \in \mathcal{D}(T)$ and conclude boundedness. Continuity at x_0 means that for any $\epsilon > 0$ one can find a $\delta > 0$ such that $\|Tx - Tx_0\| \leq \epsilon$ holds for all $x \in \mathcal{D}(T)$ with $\|x - x_0\| \leq \delta$. We now take an arbitrary $y \in \mathcal{D}(T)$, $y \neq 0$ and set

$$x = x_0 + \delta \frac{y}{|y|}$$

Clearly $\|x - x_0\| = \delta$ and hence $\|Tx - Tx_0\| = \|T\left(\delta \frac{y}{|y|}\right)\| \leq \epsilon$ which implies $\frac{\|Ty\|}{\|y\|} \leq \frac{\epsilon}{\delta}$ for any $y \neq 0$ and therefore boundedness. To prove 2., we note that the previous argument in fact gave boundedness from continuity *at a single point* x_0. But we have also already shown that boundedness implies continuity everywhere, which concludes the proof. □

Corollary 4.1. *Let T be a bounded linear operator. Then*

- *$x_n \to x$ implies $Tx_n \to Tx$ for $x_n, x \in \mathcal{D}(T)$.*
- *The null-space $\mathcal{N}(T)$ is closed.*

Proof. The first statement is immediate from $\|Tx_n - Tx\| \leq \|T\|\|x - x_n\|$. For the second statement, pick $x \in \overline{\mathcal{N}(T)} \subset \mathcal{D}(T)$ and a sequence $x_n \in \mathcal{N}(T)$ with $x_n \to x$. We have $0 = Tx_n$. Taking $n \to \infty$ and using the first statement yields $0 = Tx$ and hence $x \in \mathcal{N}(T)$. □

Remark 4.1. *The range of a bounded linear operator does not have to be closed. To see this consider the operator $T : \ell^\infty \to l^\infty$ defined by*

$$Tx = y \quad \textit{with } x = (\xi_j) \textit{ and } y = \left(\frac{\xi_j}{j}\right).$$

This operator is seen to be bounded with $\|T\| = 1$. However, its range is not closed in l^∞. To see this, consider the sequence of sequences $x_j = \left(1, \sqrt{2}, \sqrt{3}, ..., 0, 0...\right)$ with first j entries non-zero. Clearly, its image is $Tx_j = \left(1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, ...0, 0...\right)$ which converges in ℓ^∞. However, the limit does not have a preimage in ℓ^∞.

4.2 Restriction and Extension of operators

We now discuss an extension theorem for linear operators. In practice, extension will often be useful because an operator may only be defined on a dense set of a larger space and we would like to extend the operator keeping its linearity (and if applicable, boundedness) properties.

Definition 4.3. *Two operators T_1 and T_2 are said to be equal, written $T_1 = T_2$ if they have the same domain, $\mathcal{D}(T_1) = \mathcal{D}(T_2)$ and $T_1 x = T_2 x$ holds for all $x \in \mathcal{D}(T_1) = \mathcal{D}(T_2)$. The restriction of an operator $T : \mathcal{D}(T) \to Y$ to a subset $\mathcal{B} \subset \mathcal{D}(T)$ is denoted $T|_{\mathcal{B}} : \mathcal{B} \to Y$, where $T|_{\mathcal{B}} x = Tx$ for all $x \in \mathcal{B}$. An extension of T to a set $M \supset \mathcal{D}(T)$ is an operator $\tilde{T} : M \to Y$ such that $\tilde{T}|_{\mathcal{D}(T)} = T$.*

Theorem 4.3. *Let X be a normed space and Y a Banach space. Consider $T : X \supset \mathcal{D}(T) \to Y$ a bounded linear operator. Then T has an extension*

$$\tilde{T} : \overline{\mathcal{D}(T)} \to Y$$

where $\tilde{T}$ is a bounded linear operator of norm $\|\tilde{T}\| = \|T\|$.

Proof. The idea is simple. Take $x \in \overline{\mathcal{D}(T)}$ and pick a sequence $x_n \in \mathcal{D}(T)$ with $x_n \to x$. We have

$$\|Tx_m - Tx_n\| = \|T(x_m - x_n)\| \leq \|T\|\|x_m - x_n\|$$

and hence Tx_n is Cauchy, therefore $Tx_n \to y \in Y$ since Y is complete. This suggests to define

$$\tilde{T}x = y = \lim_{n\to\infty} Tx_n$$

We need to show that this is independent of the sequence x_n converging to x. Indeed, if $x_n \to x$ and $z_n \to x$ are two sequences, then the sequence $v_n = (x_1, z_1, x_2, z_2, ...)$ also converges to x and by the above argument Tv_m converges. But then any subsequence of Tv_m converges to the same value, in particular the subsequences Tx_n and Tz_n converge to the same value, which shows that $\tilde{T}$ is uniquely defined at every $x \in \overline{\mathcal{D}(T)}$. It is easy to see that $\tilde{T}$ is linear and that $\tilde{T}|_{\mathcal{D}(T)} = T$. Moreover we have $\|Tx_n\| \leq \|T\|\|x_n\|$ for any n and letting $n \to \infty$ using the continuity of the norm yields $\|\tilde{T}x\| \leq \|T\|\|x\|$, which implies $\|\tilde{T}\| \leq \|T\|$. Since the other direction is trivial, we have $\|\tilde{T}\| = \|T\|$ as claimed. □

4.3 Linear functionals

We now look at a special class of linear operators, namely those mapping a normed space X into the real or complex numbers. These functionals play a special role in the analysis as we will see in due course.

Definition 4.4. *Let X be a vector space over K. A linear functional f is a linear operator $f : X \supset \mathcal{D}(f) \to K$, where $\mathcal{D}(f)$ is a subspace of X.*

Recall that for us $K = \mathbb{R}$ or $K = \mathbb{C}$.

Definition 4.5. *A bounded linear functional f is a bounded linear operator with range in the scalar field K of the normed space X in which the domain $\mathcal{D}(f)$ lies. In particular, there exists a $C > 0$ such that*

$$|f(x)| \leq C\|x\| \qquad \text{for all } x \in \mathcal{D}(f).$$

The norm is defined as

$$\|f\| = \sup_{x\in\mathcal{D}(f), x\neq 0} \frac{|f(x)|}{\|x\|}.$$

Examples

1. The definite integral on $C[a,b]$ equipped with the max norm:

$$f(x) = \int_a^b x(t)\,dt$$

is a bounded linear functional with $\|f\| = b-a$. Indeed, we have $|f| \leq (b-a)\|x\|$ and hence $\|f\| \leq b-a$, while the constant function shows that $\|f\| \geq b-a$.

2. Fix a $t_0 \in [a,b]$ and consider on $C[a,b]$ the functional

$$g(x) = x(t_0) \ .$$

This is a bounded linear functional with $\|g\| = 1$. (Why?).

3. The space ℓ^2. Fix $a = (\alpha_j) \in \ell^2$ and define for $x = (\xi_j) \in \ell^2$ the functional

$$h(x) = \sum_{j=1}^{\infty} \xi_j \alpha_j$$

This is a bounded linear functional on ℓ^2 with $\|h\| = \|a\|_{\ell^2}$. (Why?)

Note that the space of all linear functionals on a vector space can itself be made into a vector space. This vector space is denoted $X^\star$ and called *the algebraic dual space of* X. The linear operations are defined as expected:

$$(f_1 + f_2)(x) = f_1(x) + f_2(x) \qquad \text{and} \qquad (\alpha f)(x) = \alpha f(x) \ .$$

The algebraic dual space is to be distinguished from the dual space, which we will introduce below (as the space of *bounded* (=continuous) linear functionals in the context of normed spaces.

4.4 Normed Spaces of Operators

The space of all bounded linear operators can also be turned into a vector space by defining addition and scalar multiplication as

$$(T_1 + T_2)(x) = T_1 x + T_2 x \qquad \text{and} \qquad (\alpha T)x = \alpha T x \,.$$

Since we have already shown that $\|T\|$ of (14) has the properties of a norm, we have

Theorem 4.4. *The vector space $\mathcal{B}(X,Y)$ of bounded linear operators from a normed space X into a normed space Y is itself a normed space with norm*

$$\|T\| = \sup_{x \in X, x \neq 0} \frac{\|Tx\|}{\|x\|} \ . \tag{14}$$

A more interesting observation is that the space $\mathcal{B}(X,Y)$ is complete if Y is complete (i.e. completeness of X is *not* needed).

Theorem 4.5. *Let X, Y be normed spaces. If Y is a Banach space, then $\mathcal{B}(X,Y)$ is a Banach space.*

Proof. We pick a Cauchy sequence T_n in $\mathcal{B}(X,Y)$ and have to show it converges to an element T in $\mathcal{B}(X,Y)$. We first find a suitable candidate element as follows. From

$$\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\| \tag{15}$$

and the fact that T_n is Cauchy in $\mathcal{B}(X,Y)$ we conclude that (for fixed x) $T_n x$ is Cauchy in Y. Since Y is complete, $T_n x \to y_x \in Y$, which gives our desired candidate map $T : X \to Y$, $Tx := y_x = \lim T_n x$. This map is linear (easy to check) and it remains to prove that it is bounded and that indeed $\|T_n - T\| \to 0$. For this we go back to (15) which tells us that for any $\epsilon > 0$ we can find N such that

$$\|T_n x - T_m x\| < \epsilon \|x\| \quad \text{for all } m,n > N\,.$$

Taking the limit $m \to \infty$ (for fixed x) and using that the norm is continuous we find

$$\|T_n x - Tx\| \leq \epsilon \|x\| \quad \text{for all } n > N\,.$$

Therefore $T_n - T \in \mathcal{B}(X,Y)$ and since $T = T_n - (T_n - T)$ also $T \in \mathcal{B}(X,Y)$. Finally, the last inequality implies that $\|T_n - T\| \leq \epsilon$, which means that indeed T_n converges to T in the operator norm. □

An important consequence of the last theorem is that if you consider bounded linear functionals from a normed space X into $\mathbb{R}$ or $\mathbb{C}$ (which are complete), then independently of whether the domain is complete or not, the space of bounded linear functionals will be complete. Let us formalize this observation.

Definition 4.6. *Let X be a normed space. Then the set of all bounded linear functionals on X constitutes a normed space with norm*

$$\|f\| = \sup_{x \in X, x \neq 0} \frac{|f(x)|}{\|x\|}$$

which is called the <u>*dual space*</u> *of X and denoted X'.*

Remark 4.2. *Recall that the algebraic dual space was the vector space of* <u>*all*</u> *linear functionals on X and didn't need a norm for its definition. See Exercise 7 of Week 2.*

Our observation above can now be phrased as

Theorem 4.6. *The dual space X' of a normed space X is a Banach space independently of whether X is.*

5 The dual space and the Hahn-Banach theorem

In the last section we defined the dual of a normed space X and proved that it is always a Banach space. Why should one be interested in or study the space of continuous linear functionals on a normed space X? Well, one answer is that one might learn something about the (possibly complicated) space X by studying the space of continuous maps on it (which might be easier, since it is complete!). This suggests that we should investigate the precise relation between X and X'. One thing we are going to prove below (a consequence of the famous Hahn-Banach theorem) is that X' is sufficiently large to distinguish between elements of X in the sense that if $x \neq y$ in X, then there exists a functional $f \in X'$ such that $f(x) \neq f(y)$.

Another (at this point rather vague) motivation comes from the fact that many problems in partial differential equations (regarding existence and uniqueness of solutions to PDEs) can be phrased as questions about the duals of certain spaces that the solutions live in. We will see this towards the end of the course.

For now it may be useful to remember what the linear functionals in $\mathbb{R}^n$ are. Geometrically, the zero set of a linear functional $f(v) = 0$ represents a hyperplane in $\mathbb{R}^n$ and, more generally, the set $\{v \mid f(v) = c\}$ represents a translated (affine) hyperplane. We adapt these definitions for an arbitrary (possibly ∞-dimensional) vector space V over $\mathbb{R}$, i.e. we call the set

$$H = \{v \in V \mid f(v) = c\}$$

an affine hyperplane in V. We also have

Definition 5.1. *A set $K \subset V$ is convex if $v_0, v_1 \in K$ implies that*

$$v(t) = (1-t)\, v_0 + t v_1$$

lies in K for all $0 \leq t \leq 1$.

In other words K is convex, if the straight line connecting two points lies entirely in K. The *geometric* idea of the Hahn-Banach theorem may be phrased as follows:

If K is a convex set and $v_0 \notin K$, then K and v_0 can be separated by an affine hyperplane.

More precisely, there exists a linear functional f and a real number a such that $f(v_0) \geq a$ while $f(v) < a$ if $v \in K$.

We first prove a finite dimensional version, which already contains the key idea of the general proof. Unfortunately, the general case requires the axiom of choice (which is logically equivalent to Zorn's Lemma which is what we are going to use) and we'll have to go through a bit of abstract set theory to complete the proof.

Proposition 5.1. *Let K be an open convex subset of $V = \mathbb{R}^d$ and $v_0 \notin K$. Then there exists a linear functional f and a real number a such that $f(v_0) \geq a$ while $f(v) < a$ if $v \in K$.*

Proof. Wlog we can assume K non-empty and that $0 \in K$ (otherwise we translate the problem). We now define a function which characterizes K in the sense that for each direction from the origin we associate the (inverse) distance we can go in that direction while still remaining in K:

$$p(v) = \inf_{r>0} \{r \mid \frac{v}{r} \in K\}. \tag{16}$$

This is well-defined (why?). Note that if we norm $\mathbb{R}^d$ and $K = \{v \mid \|v\| < 1\}$ is the unit ball, then $p(v) = \|v\|$.

Exercise 5.1. *Draw the function p in case that $K = (a, b) \subset \mathbb{R}^n$.*

In general, the function p characterizes K in that

$$p(v) < 1 \text{ iff } v \in K\,. \tag{17}$$

To see this, note that if $p(v) < 1$ then $v' = \frac{v}{1-\epsilon} \in K$ for some $\epsilon > 0$. This means $v = (1-\epsilon)\, v' + \epsilon \cdot 0$ which by convexity implies $v \in K$. Conversely, $v \in K$ implies $\frac{v}{1-\epsilon} \in K$ for some ϵ because K is open and therefore $p(v) \leq 1-\epsilon$.

Note also that p has the following two properties

1. $p(av) = ap(v)$ for $a \geq 0$ and $v \in V$
2. $p(v_1 + v_2) \leq p(v_1) + p(v_2)$ if $v_1, v_2 \in V$.

The first is immediate and for the second we note that if $\frac{v_1}{r_1} \in K$ and $\frac{v_2}{r_2} \in K$, then $\frac{v_1+v_2}{r_1+r_2} = \frac{r_1}{r_1+r_2}\frac{v_1}{r_1} + \frac{r_2}{r_1+r_2}\frac{v_2}{r_2} \in K$ by convexity. A functional satisfying 1. and 2. above is called a *sublinear functional.*

We now phrase our problem in terms of p. It is clear that we are done if we can construct a linear functional f such that

$$f(v_0) = 1 \qquad \text{and} \qquad f(v) \leq p(v) \quad \text{for } v \in \mathbb{R}^d \tag{18}$$

Indeed, from (17) we see that for such an f we have $f(v) < 1$ for all K. Does such a functional exist? We already have f given on $V_0 = span\{v_0\}$ and this is consistent with (18). Indeed, for $\lambda \geq 0$ we have $f(\lambda v_0) = \lambda f(v_0) \leq \lambda p(v_0) = p(\lambda v_0)$ while for $\lambda < 0$ this is immediate (why?).

The next step is to extend f from V_0 to $V_1 = span\{v_0, v_1\}$ with v_1 an arbitrary element of V which is linearly independent from v_0. Moreover, we have to make sure that (18) holds. Therefore, we need

$$af(v_1) + b = f(av_1 + bv_0) \overset{!}{\leq} p(av_1 + bv_0) \tag{19}$$

for a and b arbitrary scalars. In particular, setting $a = 1$ and $bv_0 = w$ we need

$$f(v_1) + f(w) \overset{!}{\leq} p(v_1 + w) \qquad \text{for all } w \in V_0$$

and, setting $a = -1$,

$$-f(v_1) + f(w') \overset{!}{\leq} p(-v_1 + w') \qquad \text{for all } w' \in V_0$$

You should convince yourself that conversely, having the two inequalities above is sufficient to deduce (19). To summarize, we need to find $f(v_1)$ such that for all $w, w' \in V_0$ we have

$$-p(-v_1 + w') + f(w') \leq f(v_1) \leq p(v_1 + w) - f(w) \tag{20}$$

since then we satisfy (19). But there is indeed a number between the left hand side and the right hand side since in general

$$-p(-v_1 + w') + f(w') \leq p(v_1 + w) - f(w)$$

or equivalently

$$f(w) + f(w') \leq p(v_1 + w) + p(-v_1 + w')$$

holds. (Why does this imply that we can indeed find $f(v_1)$?) To see this last inequality note that $f(w) + f(w') = f(w + w') \leq p(w + v_1 + w' - v_1)$ (using that (18) holds on V_0) and apply the sublinearity property of p.

Now that we have successfully extended the functional from V_0 to V_1, it is clear that we can continue this procedure inductively and finish the proof. □

With this proof you have already seen a key ingredient of the Hahn-Banach theorem.

Definition 5.2. *Let V be a real vector space. A functional p on V is called* <u>*sublinear*</u> *if*

1. $p(ax) = ap(x)$ *for $x \in V$ and $a > 0$ (positive homogeneous)*
2. $p(x+y) \leq p(x) + p(y)$ *for $x, y \in V$ (subadditive)*

Observe that a norm is a sublinear functional. Note also that the p in the proof of Proposition 5.1 had the additional property that it was non-negative. However, this is not needed in the following (real) version[5] of the Hahn-Banach theorem:

Theorem 5.1. *[Hahn-Banach] Let V be a real vector space and p a sublinear functional on V. Let $M \subset V$ be a linear subspace and f a linear functional on M satisfying*

$$f(v) \leq p(v) \quad \text{for all } v \in M .$$

Then f can be extended on all of V such that

$$F(v) \leq p(v) \quad \text{for all } v \in V \qquad \text{and} \qquad F(v) = f(v) \ \text{ for } v \in M .$$

Before we do the proof, let us discuss some implications of this statement. Clearly, the above does not involve any norms. However, given a bounded linear functional on a normed space we can deduce

Theorem 5.2. *Let f be a bounded linear functional on a subspace M of a normed real vectorspace X. Then there exists a bounded linear functional F on X which is an extension of f to X and has the same norm*

$$\|F\|_X = \|f\|_M$$

where

$$\|F\|_X = \sup_{x \in X, \|x\|=1} |F(x)| \qquad \text{and} \qquad \|F\|_M = \sup_{x \in M, \|x\|=1} |F(x)|$$

Proof. Wlog $M \neq \{0\}$. To apply the Hahn-Banach theorem we need to construct a suitable p from the norm. We define

$$p(x) := \|f\|_M \|x\| .$$

By the remark following Definition 5.2 this is a sublinear functional and we also have

$$|f(x)| \leq \|f\|_M \|x\| = p(x) \quad \text{for } x \in M .$$

Hahn-Banach gives us a functional F defined on all of X such that

$$F(x) \leq p(x) = \|f\|_M \|x\| \quad \text{for } x \in X$$

Now since

$$-F(x) = F(-x) \leq \|f\|_M \| - x\| ,$$

[5]For the complex version of the Hahn-Banach theorem see Theorem 11.1 in the exercise section.

we have $|F(x)| \leq \|f\|_M \|x\|$ and hence

$$\|F\|_X \leq \|f\|_M$$

Since F is an extension, the other direction is trivial and we are done. □

From this theorem we can deduce a range of interesting conclusions about constructing linear functionals.

Theorem 5.3. *Let X be a normed vector space and $x_0 \neq 0$ be a non-trivial element of X. Then there exists a bounded linear functional $F(x)$ such that*

$$\|F\| = 1 \quad , \quad F(x_0) = \|x_0\| \, .$$

Proof. We can define a linear functional f on the space $span(x_0)$ by $f(cx_0) = c\|x_0\|$. This functional is linear, has norm 1 and $F(x_0) = \|x_0\|$ as desired. Hahn-Banach extends it to all of X. □

Corollary 5.1. *The dual X' is non-trivial (contains more than the zero element) if X is.*

Corollary 5.2. *If x_1 is an element of X such that $f(x_1) = 0$ for every bounded linear functional f on X, then $x_1 = 0$.*

Proof. Assume the contrary, i.e. $x_1 \neq 0$ but nevertheless $f(x_1) = 0$ for all $f \in X'$. Then by the previous Theorem we can construct from x_1 a linear functional F with $F(x_1) = \|x_1\| \neq 0$. Contradiction. □

Corollary 5.3. *Let X be a normed space and $x_1 \neq x_2$ then there exists a bounded linear functional $f \in X'$ such that $f(x_1) \neq f(x_2)$.*

Proof. Take $x = x_1 - x_2$ in Theorem 5.3. □

The last statement may be paraphrased by saying that the dual X' is "sufficiently large" to separate points in X.

5.1 The proof

We now prove Theorem 5.1. We need Zorn's Lemma (logically equivalent to the axiom of choice) which requires some preparation.

5.1.1 Zorn's Lemma

Definition 5.3. *A partially ordered set (POS) M is a set on which there is defined a partial ordering, i.e. a binary relation "$\preceq$" satisfying*

1. *$a \preceq a$ for every $a \in M$ (reflexivity)*
2. *If $a \preceq b$ and $b \preceq a$, then $a = b$ (antisymmetry)*
3. *If $a \preceq b$ and $b \preceq c$, then $a \preceq c$ (transitivity)*

For instance, the real numbers form a partially ordered set with the usual "$\leq$" as the partial ordering. In this case any two elements can be compared with one another but the partial ordering does not require this. For instance the linear subspaces of a given vector space can be partially ordered by inclusion but many subspaces cannot be compared at all.

Definition 5.4. *A totally ordered set (TOS) is a POS such that any two elements can be compared, i.e. either $a \preceq b$ or $b \preceq a$ or both hold for any $a, b \in M$. A TOS is also called a chain.*

Definition 5.5. *An upper bound of a subset W of a partially ordered set M is an element $u \in M$ such that*

$$x \preceq u \quad \textit{holds for every } x \in W.$$

Note that u is not necessarily in W and may or may not exist (consider the real numbers vs the strictly negative real numbers...) What is important is that u is larger than any element of W.

Definition 5.6. *A maximal element of a POS M is an $m \in M$ such that*

$$m \preceq x \qquad \textit{implies } m = x.$$

Again, a maximal element may or may not exist. A maximal element need not be an upper bound as there can be elements in M that x cannot be compared with. So alternatively a maximal element is an element which is larger than any element in M that it can be compared to.

Examples.

- on $\mathbb{Z}^+$ define $a \preceq b$ if $a|b$ (a divides b). Clearly not totally ordered, but $\{1, 2, 4, 8, 16, ...\}$ would be a totally ordered subset.
- subspaces of a vector space V with $V \preceq W$ if $V \subset W$. Clearly not totally ordered but a tower of subspaces $V_1 \subset V_2 \subset V_3 \subset \ldots$ defines a totally ordered subset.
- on $\{2, 3, 4, 5, ...\}$ define $a \preceq b$ if $b|a$ (b divides a – i.e. now an element is larger than another if it is a factor of that element). Prime numbers are maximal elements.

Zorn's Lemma. *Let $M \neq \emptyset$ be a POS. Suppose that every totally ordered subset $G \subset M$ has an upper bound. Then M has at least one maximal element.*

Zorn's Lemma is equivalent to the axiom of choice.[6]

Theorem 5.4. *[cf. Theorem 3.1] Every vector space $X \neq \{0\}$ has a Hamel basis.*

Proof. The idea is to construct a basis as a maximal linearly independent set. Let M be the set of all linearly independent subsets of X. Since X is non-trivial, we have $\{x_0\} \in M$ for some $x_0 \neq 0$. Set inclusion defines a partial ordering on M. Every totally ordered subset $V_1 \subset V_2 \subset ...$ has an upper bound, namely the union $\mathbb{V} = \cup_\alpha V_\alpha$ [Indeed, $V_\alpha \subset \cup_\lambda V_\lambda$ for all α and this a linearly independent subset, i.e. any chosen $v_1, \ldots, v_n \in \mathbb{V}$ are linearly independent. Indeed, each

[6]The axiom of choice states that if $M_\alpha \neq \emptyset$ is a family of non-empty set, then the Cartesian product $\Pi_\alpha M_\alpha$ is non-empty. This means that if there is an element in each M_α then there is also a function which "chooses" from each M_α an element.

v_i is contained in some V_α, say $v_1 \in V_1$, $v_2 \in V_2$, ... But the V_i are totally ordered, so one V_i must contain all v_i which are hence linearly independent]. Zorn's Lemma implies that M has a maximal element B. We claim that B is a Hamel basis for X. Since $B \subset X$ is linearly independent we need to show $spanB = X$. Assume not, and $spanB = Y \neq X$. Then there is a $z \in X$ with $z \neq spanB$ such that $B \cup \{z\}$ is linearly independent [Assume not, then there is a k such that

$$\sum_{i=1}^{k} c_i v_i = 0$$

holds with not all c_i being zero. Moreover, one of the v_i's involved (say $v = v_k$) must be proportional z and appear with a non-zero c_k, since the elements of B are linearly independent. But then v can be expressed as a linear combination of the v_i's in B (dividing by c_k) which means that $z \in B$. Contradiction.] contradicting maximality. □

5.1.2 Proof of Theorem 5.1

The first part is as before. Wlog $M \neq V$. We pick $v_1 \neq M$ and extend the functional to a subspace M_1 spanned by M and v_1. [You may want to write the argument again in detail, as we only sketch it here. In particular, make sure that the additional positivity property of p which was available in Proposition 5.1, was actually nowhere used to construct the extension.] We need a functional f_1 satisfying $f_1(\alpha v_1 + w) = \alpha f_1(v_1) + f(w) \leq p(\alpha v_1 + w)$ for any scalar α and $w \in M$. Choose $\alpha = 1$ and $\alpha = -1$, write down the condition on $f_1(v_1)$ and conclude that such a number can be found. Then conversely show that if the condition holds on $f_1(v_1)$ you obtain the desired functional.] Now if we had

$$V = \bigcup_{k=1}^{\infty} M_k$$

we would be done, since each $x \in V$ lives in some M_k and we could define F by induction. If that's not the case[7] we need to invoke Zorn's Lemma as follows.

Consider the collection S of all linear functionals g defined on subspaces $D(g)$ such that

1. $D(g) \supset M$
2. $g(x) = f(x) \quad \text{for } x \in M$
3. $g(x) \leq p(x) \quad \text{for } x \in D(g) \quad .$

Note that S is non empty because we already extended f to M_1 above. We introduce a partial ordering on S

$$g_1 \leq g_2 \quad \text{if } D(g_1) \subset D(g_2) \quad \text{and} \quad g_1(x) = g_2(x) \text{ for } x \in D(g_1)\,.$$

[7] There are other *special* situations which can absolve one from invoking Zorn's lemma at this point. They will be explored in the exercises.

We need to show that every totally ordered subset has an upper bound. Let W be a totally ordered subset. Define h by

$$D(h) = \bigcup_{g \in W} D(g)$$
$$h(x) = g(x) \quad \text{for } g \in W,\, x \in D(g)\,. \tag{21}$$

This definition is unambiguous. Indeed, if $x \in D(g_1) \cap D(g_2)$, then because W is totally ordered we have $g_1 \leq g_2$ or $g_2 \leq g_1$, i.e. one $D(g)$ must contain the other and the functionals agree on the smaller one, hence $g_1(x) = g_2(x)$ for such x. Also $g \leq h$ for all $g \in W$, so h is an upper bound. As W was arbitrary, Zorn's Lemma implies the existence of a maximal element F of S satisfying $F(x) \leq p(x)$.

It remains to show that $D(F)$ is all of V. Suppose not. Then there is a $v \neq 0$ in $V \setminus D(F)$ and we can consider the subspace of V spanned by $D(F)$ and v. But this is the codimension 1 case we know how to treat: We can extend F to this bigger subspace and contradict the maximality of F guaranteed by Zorn's Lemma.

5.2 Examples: The dual of ℓ^1

A drawback of the Hahn-Banach theorem is that it provides no way to *construct* the functionals whose existence it promises. In some cases one can compute the dual space directly (or at least a large subset of it). We will see some examples now.

Example 5.1. *The dual of ℓ^1 is (isomorphic to) ℓ^∞, i.e. there is a bijective linear map T between the dual of ℓ^1 and ℓ^∞ which moreover preserves the norm, $\|Tx\| = \|x\|$.*

We pick a Schauder basis for ℓ^1, $e_k = \delta_{kj}$ (i.e. k^{th} entry of the sequence is 1, all others are zero). Every $x \in \ell^1$ can be written

$$x = \sum_{k=1}^{\infty} \xi_k e_k\,.$$

Consider now an arbitrary $f \in \left(\ell^1\right)'$. We have

$$f(x) = \sum_{k=1}^{\infty} \xi_k f(e_k) = \sum_{k=1}^{\infty} \xi_k \gamma_k$$

with $\gamma_k \in \mathbb{R}$. Now from

$$|\gamma_k| = |f(e_k)| \leq \|f\| \|e_k\| = \|f\| \tag{22}$$

we deduce that $(\gamma_k) \in \ell^\infty$. In other words, given $f \in \left(\ell^1\right)'$ we can associate with it a $(\gamma_k) \in \ell^\infty$. Conversely, given $b = (\beta_k) \in \ell^\infty$ we can obtain a corresponding $g \in \left(\ell^1\right)'$ as follows:

$$g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k \quad \text{for } x = (\xi_k) \in \ell^1\,. \tag{23}$$

Indeed, we have

$$|g(x)| \leq \sup_k |\beta_k| \sum_{k=1}^{\infty} |\xi_k| = \|x\| \sup_k |\beta_k|,$$

which shows that indeed g is a bounded linear functional on ℓ^1. Composition of the two maps considered above shows they are inverses of one another. Does the map $(\ell^1)' \to \ell^\infty$ preserve the norm? For the norm on the dual space we have

$$|f(x)| \leq \sup_j |\gamma_j| \|x\|$$

from which we deduce

$$\|f\|_{(\ell^1)'} \leq \sup_j |\gamma_j| = \|\gamma_j\|_{\ell^\infty} \leq \|f\|_{(\ell^1)'}$$

with the last inequality following from (22).

You will construct the dual of ℓ^p for $p > 1$ in Exercise 1 of Week 3. See also the last exercise of the previous week.

5.3 The dual of $C[a,b]$

Let us try to understand the dual space of the space of continuous functions. Naively, we know that it must contain the continuous functions themselves, since any $g \in C[a,b]$ gives rise to a continuous linear functional g'

$$g'(f) = \int_a^b f(t) \cdot g(t)\, dt \qquad \text{for } f \in C[a,b] \tag{24}$$

Exercise 5.2. *Show that* $\|g'\| = \int_a^b |g|$. *Conclude that there are elements in* $C'[a,b]$ *not of the form (24) for continuous* g. *Give an example.*

To understand the space $C'[a,b]$ we start with f an arbitrary bounded linear functional on $C[a,b]$. An idea that turns out to be fruitful is to extend the functional f to act on step functions. We recall that step functions are constant on finitely many subintervals of $[a,b]$. In other words, if we let

$$k_s(t) = \begin{cases} 1 & \text{for } a \leq t \leq s < b \\ 0 & \text{for } a < s < t \leq b \end{cases}$$

the characteristic function on the interval $[a,s]$, then an arbitrary step function $y(t)$ can be written as

$$y(t) = \sum_{i=1}^{n} \alpha_i \left(k_{t_i}(t) - k_{t_{i-1}}(t) \right)$$

where the α_i are scalars and $a = t_0 < t_1 < ... < t_n = b$ is a partition of $[a,b]$.

To extend f to this class of functions via Hahn-Banach, we note that the space $C[a,b]$ (with *sup*-norm) is contained in the space of bounded functions $\mathcal{B}$ (with *sup*-norm) and that step functions live in the latter. So indeed HB gives a functional

$$F : \mathcal{B} \to \mathbb{R} \qquad \text{with } \|F\| = \|f\|.$$

which on an arbitrary step function y acts as

$$F(y) = \sum_{i=1}^{n} \alpha_i \left(g(t_i) - g(t_{i-1})\right) \qquad \text{with } g(s) := F(k_s).$$

Now let x be an arbitrary element of $C[a,b]$. Since x is uniformly continuous (why?), we have that for any $\epsilon > 0$ we can find a δ such that

$$|t' - t| < \delta \quad \text{implies} \quad |x(t') - x(t)| < \epsilon\ .$$

Now choose a partition $a = t_0 < t_1 < ... < t_n = b$ such that

$$\eta = \max_i |t_i - t_{i-1}| < \delta$$

i.e. its largest subinterval is still smaller than the prescribed δ. Then for any $t_{i-1} \leq t_i' \leq t_i$ we have $|x(t) - x(t_i')| < \epsilon$ for all $t_{i-1} \leq t \leq t_i$ in that interval. In particular, the "step-function approximate" of x,

$$y(t) := \sum_{i=1}^{n} x(t_i') \left[k_{t_i}(t) - k_{t_{i-1}}(t)\right]$$

satisfies

$$\|x - y\| < \epsilon$$

and therefore

$$\|F(x) - F(y)\| \leq \|F\| \|x - y\| < \epsilon \|F\|$$

converges to $F(x)$ as $\epsilon \to 0$. In summary, we have that the limit

$$f(x) = F(x) = \lim_{\eta \to 0} \sum_{i=1}^{n} x(t_i') \left[g(t_i) - g(t_{i-1})\right] \tag{25}$$

exists. To recap, we have simply done the following. We extended the functional f to act on step functions and obtained its action on an arbitrary continuous function x through a limit of step functions approximating x. Why is that useful? Well, it tells us something about the general form of the functional f, which was our goal in the first place. We now know that any functional f can be written in the form (25) for some function g whose properties we still need to understand. Note, however already at this point that if $g(t) = t$, then (25) is the definition of the Riemann integral! The more general limit expression on the right hand side of (25) has a name. It is called the *Riemann-Stieltjes integral* and written

$$\int_a^b x(t)\, dg(t) := \lim_{\eta \to 0} \sum_{i=1}^{n} x(t_i') \left[g(t_i) - g(t_{i-1})\right]$$

provided the limit exists. So what are the properties of g in (25)? For an arbitrary step function

$$y(t) = \sum_{i=1}^{n} \alpha_i \left(k_{t_i}(t) - k_{t_{i-1}}(t)\right)$$

we have

$$F(y) = \sum_{i=1}^{n} \alpha_i \left(g(t_i) - g(t_{i-1})\right)$$

and hence

$$\Big| \sum_{i=1}^{n} \alpha_i \left(g(t_i) - g(t_{i-1})\right) \Big| \leq \|F\|\|y\| = \|f\| \max_i |\alpha_i| \,.$$

As this is true for any choices of α_i we have in particular (choosing α_i to be ± 1)

$$\sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| \leq \|f\| \quad \text{for any partition of } [a,b]. \tag{26}$$

Functions for which the expression on the left of (26) is bounded for any partition are called *functions of bounded variation* (the space they live in is called $BV[a,b]$). Defining the total variation of a function in $BV[a,b]$ to be

$$V(g) = \sup_{partitions} \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})|$$

we can now state our results as

Proposition 5.2. *For any bounded linear functional f on $C[a,b]$ we can find a function g of bounded variation on $[a,b]$ such that*

$$f(x) = \int_a^b x(t)\, dg(t) \quad \text{for } x \in C[a,b]$$

and $V(g) = \|f\|$.

Proof. Everything is clear except $V(g) = \|f\|$, since the above only gave $V(g) \leq \|f\|$. But the other direction follows directly from (25). □

Remark 5.1. *The converse is also true: For any function of bounded variation the Riemann-Stieltjes integral of $x \in C[a,b]$ exists.*

Remark 5.2. *The g of the above Proposition is in general non-unique but can be made unique by imposing additional conditions. See [Schechter] for the details.*

5.4 A further application

Theorem 5.5. *Let M be a subspace of a normed vector space X and suppose $x_0 \in X$ satisfies*

$$d = d(x_0, M) = \inf_{x \in M} \|x - x_0\| > 0$$

Then there is an $F \in B(X, \mathbb{R})$ such that $\|F\| = 1$, $F(x_0) = d$ and $F(x) = 0$ for $x \in M$.

Proof. See Exercise 4 of Week 3. □

Theorem 5.6. *A normed vector space V is separable if V' is separable.*

Our example of Section 5.2 shows that the converse is in general false: ℓ^1 is separable but its dual ℓ^∞ is not (cf. Example 2.1 and Exercise 2.1). In particular,

Corollary 5.4. $(\ell^\infty)' \neq \ell^1$.

Proof of Theorem 5.6. Consider $\{\lambda_k\}_{k\geq 1}$ dense in V'. For each k pick $\xi_k \in V$ with $\|\xi_k\| = 1$ such that

$$\lambda_k(\xi_k) \geq \frac{1}{2}\|\lambda_k\| .$$

Let W be the set of finite linear combinations of the ξ_k with rational coefficients – this set is countable. We claim it is dense in V. Suppose not. Then there exists a $\delta > 0$ and a $\xi \in V$ such that

$$\inf_{\eta\in W} \|\xi - \eta\| = \delta > 0 .$$

By Theorem 5.5 above, find $\lambda \in V'$ such that $\|\lambda\| = 1$, $\lambda(\xi) = \delta$, $\lambda|_W = 0$. But the λ_k are dense in V', so there exists a subsequence λ_{k_i} with $\lim_{i\to\infty} \|\lambda - \lambda_{k_i}\| = 0$. On the other hand, we have

$$\|\lambda - \lambda_{k_i}\| \geq |(\lambda - \lambda_{k_i})(\xi_{k_i})| = |\lambda_{k_i}(\xi_{k_i})| \geq \frac{1}{2}\|\lambda_{k_i}\|$$

which implies $\lambda_{k_i} \to 0$ contradicting $\|\lambda\| = 1$. □

5.5 The adjoint operator

Let X, Y be normed spaces and $T \in B(X, Y)$. The adjoint of T is a bounded linear operator from Y' to X' defined as follows. With a given $g \in Y'$ we associate a functional $f \in X'$ via

$$f(x) = g(Tx) \; . \tag{27}$$

This functional is clearly linear as g and T are both linear. It is also bounded:

$$|f(x)| = |g(Tx)| \leq \|g\|\|Tx\| \leq \|g\|\|T\|\|x\|$$

and hence

$$\|f\| \leq \|g\|\|T\| .$$

The functional

$$T^\times : Y' \ni g \mapsto f \in X'$$

with f defined by (27) is called the adjoint of T. Note that given T, (27) defines f uniquely in terms of g, since any other functional satisfying (27) would have to agree with f on all x (and hence be identical by Hahn-Banach).

Theorem 5.7. *We have* $T^\times \in B(Y', X')$ *and* $\|T^\times\| = \|T\|$.

Proof. We have

$$T^\times g(x) = g(Tx)$$

hence

$$|T^\times g(x)| \leq \|g\|\|T\|\|x\| .$$

Recall that by definition

$$\|T^{\times} g\|_{X'} = \sup_{x \neq 0} \frac{|T^{\times} g(x)|}{\|x\|} \leq \|g\| \|T\|$$

and hence

$$\|T^{\times}\|_{B(Y',X')} = \sup_{g \neq 0} \frac{\|T^{\times} g\|}{\|g\|} \leq \|T\| .$$

For the other direction, it clearly suffices to show

$$\|Tx\| \leq \|T^{\times}\| \|x\| . \tag{28}$$

From

$$|g(Tx)| \leq \|T^{\times}\| \|g\| \|x\|$$

we obtain

$$\sup_{g \neq 0} \frac{|g(Tx)|}{\|g\|} \leq \|T^{\times}\| \|x\|$$

and by Exercise 2 of Week 3, a corollary of HB, the right hand side is equal to $\|Tx\|$ establishing (28). □

The adjoint has the following properties

- $(A+B)^{\times} = A^{\times} + B^{\times}$
- $(\alpha A)^{\times} = \alpha A^{\times}$
- $(AB)^{\times} = B^{\times} A^{\times}$

which one can easily check. For the last one, note that if $A : Y \to Z$ and $B : X \to Y$, then $(AB)^{\times} : Z' \to X'$ and for $g \in Z'$ and $x \in X$

$$(AB)^{\times} g(x) = g(ABx) .$$

Since $A^{\times} g \in Y'$ and $Bx \in Y$ we also have

$$B^{\times} \left(A^{\times} g\right)(x) = A^{\times} g(Bx) = g(ABx) .$$

It is also instructive to work out the adjoint of a real matrix (it is the transpose).

Remark 5.3. *You may have already seen the notion of the Hilbert adjoint which is different from the adjoint defined above. We will relate the two at the end of Section 8.6.*

5.5.1 Why adjoints?

You may ask at this point why one considers adjoints of operators. Suppose X, Y normed spaces and $T \in B(X, Y)$. A typical problem in applications (for instance a linear PDE[8]) can be expressed as

$$Tx = y$$

[8] We're cheating a little bit here because, as we have seen, derivative operators are actually *unbounded*. The concept of the adjoint can be generalized.

and we would like to know for what y we can solve this problem, in other words, what conditions does one need to impose on y to guarantee an inverse of T. Suppose $y \in R(T)$ so there is an x such that $Tx = y$. Then, for any $g \in Y'$, we have

$$g(Tx) = g(y)$$
$$T^{\times} g(x) = g(y)$$

From this we see that if $g \in N(T^{\times})$ then necessarily $g(y) = 0$. Consequently, a necessary condition to solve $Tx = y$ is that $g(y) = 0$ for all $g \in N(T^{\times})$. Under suitable additional conditions (which we will discuss in due course) this will also be sufficient.

Exercise 5.3. *What happens for $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$?*

6 The Uniform Boundedness Principle

In the lectures of this week I followed closely the notes of A. Sokal, available under www.ucl.ac.uk/∼ucahad0/3103_handout_7.pdf. In particular, we proved a stronger version of the Baire category theorem in class than stated below. Please see the addendum to the lecture notes on my webpage.

We now turn to the uniform boundedness principle or Banach-Steinhaus theorem. The key ingredient is the Baire category theorem which we shall prove first. We will then give applications to Fourier series and “weak convergence”, which is a crucial concept in PDEs.

Definition 6.1. *A subset M of a metric space is called*

1. *nowhere dense in X if its closure $\overline{M}$ has no interior points*
2. *of the first category (or meager) if M is the union of countably many nowhere dense sets*
3. *of the second category (or non-meager) if it is not of the first category.*
4. *generic if its complement is of the first category*

The idea here is to give a (purely topological) size to sets. This concept is independent of that of Lebesgue measure (see the addendum to the lecture notes).

Theorem 6.1 (Baire category)**.** *A complete metric space is of the second category. Therefore, if $X \neq \emptyset$ is complete and*

$$X = \bigcup_{k=1}^{\infty} A_k \qquad A_k \textit{ closed}$$

then at least one A_k contains a non-empty open subset.

Proof. Assume not and X is of the first category. Then $X = \cup_{k=1}^{\infty} M_k$ and all M_k are such that $\overline{M_k}$ is nowhere dense. Now clearly $\overline{M_1} \neq X$ because $\overline{M_1}$

does not contain a non-empty open set while X does (X itself, for instance). It follows that

$$\overline{M_1}^c = X \setminus \overline{M_1}$$

is non-empty and open. We pick a point $x_1 \in \overline{M_1}^c$ and a ball $B(x_1, \epsilon_1) \subset \overline{M_1}^c$ contained in the complement, where $\epsilon_1 < 1/2$, say. This ball, in fact also the smaller ball $B(x_1, \epsilon_1/2) \subset \overline{M_1}^c$ cannot be contained in $\overline{M_2}$ because the latter contains no open balls! Hence the set

$$\overline{M_2}^c \cap B(x_1, \epsilon_1/2)$$

is non-empty and open. We hence find a point $x_2 \in \overline{M_2}^c \cap B(x_1, \epsilon_1/2)$ and a ball $B(x_2, \epsilon_2 < \epsilon_1/2) \subset \overline{M_2}^c \cap B(x_1, \epsilon_1/2)$. But this ball, in fact also the smaller ball $B(x_1, \epsilon_2/2) \subset \overline{M_2}^c \cap B(x_1, \epsilon_1/2)$ cannot be contained in $\overline{M_3}$ because M_3 is nowhere dense. By induction, we obtain a sequence of nested balls

$$B_k = B(x_k, \epsilon_k)$$

with the properties

$$B_k \cap M_k = \emptyset$$

$$B_{k+1} \subset B(x_k, 1/2\epsilon_k) \subset B_k$$

The sequence of centers is Cauchy since we have

$$d(x_{k+n}, x_k) < \frac{1}{2^{k+1}}$$

and since X is complete $x_k \to x$ for some $x \in X$. We claim that $x \in B_k$ for every k. This is clear from the triangle inequality

$$d(x, x_k) \leq d(x, x_{k+n}) + d(x_{k+n}, x_k) < d(x, x_{k+n}) + \frac{1}{2^{k+1}}$$

which holds for any k and $n \geq 0$. For $n \to \infty$ the right hand side goes to $\frac{1}{2^{k+1}} < \frac{1}{2^k} = \epsilon_k$ which shows that $x \in B_k$.

But now we are done because $x \in B_k$ for all k and on the other hand $B_k \subset \overline{M_k}^c$ hence $x \notin M_k$ for all k. So $x \notin \cup_{k=1}^{\infty} M_k$. □

Example 6.1. *The rationals $\mathbb{Q}$ are dense in $\mathbb{R}$. They are also of the first category and moreover the irrational numbers are generic in $\mathbb{R}$ (why?).*

With the help of the Baire category theorem we immediately obtain the Banach-Steinhaus theorem. This theorem allows to conclude operator boundedness from pointwise boundedness.

Theorem 6.2. *[Banach-Steinhaus] Let X be a Banach space, Y a normed space and $T_n \in B(X, Y)$ be a sequence of bounded linear operators bounded at every point $x \in X$, i.e.*

$$\|T_n x\| \leq c_x \tag{29}$$

for $c_x \geq 0$. Then the sequence of norms $\|T_n\|$ is bounded, i.e. there exists a $c > 0$ such that

$$\|T_n\| \leq c \ .$$

Proof. For every $k \in \mathbb{N}$ we define

$$A_k := \{x \in X \mid \|T_n x\| \le k \text{ for all } n\}.$$

The A_k are closed sets. Indeed, if $x \in \overline{A_k}$ we pick a sequence $x_j \in A_k$ with $x_j \to x$. For fixed n we have $k \ge \|T_n x_j\| \to \|T_n x\|$ by continuity of both the norm and T_n.

By assumption (29), every x belong to some A_k, so

$$X = \bigcup_k A_k .$$

By Baire's theorem, one of the A_k contains an open ball, say $B_0 = B(x_0, r) \subset A_{k_0}$ for some $k_0 \in \mathbb{N}$ and $r > 0$. Now let $y \in X$ be arbitrary and consider

$$x = x_0 - \frac{y}{|y|}\frac{r}{2}.$$

Since both x and x_0 are in B_0 we have

$$\|T_n\left(\frac{y}{|y|}\frac{r}{2}\right)\| = \|T(x_0 - x)\| \le 2k_0 .$$

and hence $\|T\| \le \frac{4k_0}{r}$. □

6.1 Application 1: Space of Polynomials

Consider the normed space X of all polynomials on $\mathbb{R}$ with norm

$$\|x\| = \max_j |\alpha_j|$$

where α_i are the coefficients of the polynomial x, i.e. $x(t) = \alpha_0 + \alpha_1 t + ... + \alpha_{N_x} t^{N_x}$, with N_x the degree of the polynomial x.

Claim: X is incomplete.

To prove this we will use the Banach-Steinhaus theorem. In particular, we will construct a sequence T_n of operators on X satisfying $\|T_n x\| \le c_x$ but $\|T_n\| \to \infty$. This would be impossible if X was complete! We write

$$x(t) = \sum_{j=0}^{\infty} \alpha_j t^j$$

and declare $\alpha_j = 0$ for $j \ge N_x$ (the degree of x). Our sequence of operators $T_n : X \to \mathbb{R}$ is defined by

$$T_n 0 = 0 \quad \text{and} \quad T_n x = \alpha_0 + \alpha_1 + ... + \alpha_{n-1}$$

For this sequence, we have *for fixed* x (i.e. in particular N_x fixed!)

$$\|T_n x\| \le (N_x + 1) \max_j |\alpha_j| \equiv c_x$$

and the right hand side does not depend on n. On the other hand, for

$$x_n(t) = 1 + t + t^2 + ... + t^n$$

we have $\|x\| = 1$ but $\|T_n x\| = n$ which contradicts $\|T_n\| \le c$ for any constant c.

6.2 Application 2: Fourier Series

Let x be a 2π-periodic function, $x(t) = x(t+2\pi)$. In a previous course you learned that $x(t)$ has a *Fourier series*

$$x(t) = \frac{1}{2}a_0 + \sum_{m=1}^{\infty}\left(a_m \cos(mt) + b_m \sin(mt)\right) . \tag{30}$$

where

$$a_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\cos(mt)\,dt \qquad \text{and} \qquad b_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\sin(mt)\,dt \tag{31}$$

Some time is then spent on the issue of where and in what sense the series (30) converges. You may remember that continuity of the function x at t_0 is not sufficient to guarantee that (30) converges at t_0 (differentiable is certainly sufficient to guarantee convergence).

Proposition 6.1. *There exist real valued continuous functions whose Fourier series diverge at a given point t_0.*

Proof. We let

$$X = \{f \in C(\mathbb{R}) \mid x(t) = x(t+2\pi)\} \quad \text{with } \|x\|_X = \max_t |x(t)|$$

be the Banach space of continuous functions on $\mathbb{R}$ which are 2π-periodic. Wlog we take $t_0 = 0$ as the point for which we would like the Fourier series to diverge. The idea is to apply the BS-theorem to the sequence of operators

$$T_n : X \to \mathbb{R}$$
$$T_n(x) := f_n(x) = \{\text{value of the truncated Fourier series of } x \text{ at } 0\}$$

which is

$$f_n(x) = \frac{1}{2}a_0 + \sum_{m=1}^{n} a_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\left[\frac{1}{2} + \sum_{m=1}^{n}\cos(mt)\right] dt\,,$$

as for $t=0$ the *sin*-terms in (30) drop out. Now note that

$$\begin{aligned} 2\sin\left(\frac{t}{2}\right)\sum_{m=1}^{n}\cos(mt) &= \sum_{m=1}^{n}\left[\sin\left(\left(m+\frac{1}{2}\right)t\right) - \sin\left(\left(m-\frac{1}{2}\right)t\right)\right] \\ &= \sin\left(\left(n+\frac{1}{2}\right)t\right) - \sin\left(\frac{1}{2}t\right) , \end{aligned} \tag{32}$$

the first step following from the addition formula and the second from observing that the sum is telescopic. We conclude that

$$1 + 2\sum_{m=1}^{n}\cos(mt) = \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{\sin\left(\frac{1}{2}t\right)} ,$$

which holds also in the limit $t \to 0$. Therefore, we have

$$f_n(x) = \frac{1}{2\pi}\int_0^{2\pi} x(t)\,q_n(t)\,dt \qquad \text{with } q_n(t) = \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{\sin\left(\frac{1}{2}t\right)} . \tag{33}$$

We claim that f_n is bounded (depending on n) and

$$\|f_n\| = \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| dt \tag{34}$$

We then show $\|f_n\| \to \infty$ as $n \to \infty$ which will finish the proof. Why? Well, because by BS the sequence $f_n(x)$ cannot be bounded for all x because otherwise (X is complete!) the behavior $\|f_n\| \to \infty$ as $n \to$ would be excluded!

To show the two remaining claims, we first observe

$$|f_n(x)| \leq \max |x(t)| \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| dt = \|x\| \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| dt$$

so the $\leq$ direction in (34) is simple. To show the other direction, we construct an explicit function x which satisfies the bound (arbitrarily well). We write

$$|q_n(t)| = y(t) q_n(t)$$

where $y(t) = +1$ where $q_n(t) \geq 0$ and -1 otherwise. The function y is not continuous but it can be approximated arbitrarily well by a continuous function x of norm 1 in the sense that

$$\frac{1}{2\pi} \Big| \int_0^{2\pi} [x(t) - y(t)] q_n(t) \Big| < \epsilon$$

holds for any ϵ. Therefore

$$\epsilon > \frac{1}{2\pi} \Big| \int_0^{2\pi} x(t) q_n(t) dt - \int_0^{2\pi} y(t) q_n(t) dt \Big| = \Big| f_n(x) - \frac{1}{2\pi} \int_0^{2\pi} |q_n(t)| dt \Big|$$

which establishes (34). We finally need to show

$$\int_0^{2\pi} \Big| \frac{\sin\left(\left(n + \frac{1}{2}\right) t\right)}{\sin\left(\frac{1}{2} t\right)} \Big| dt \to \infty \tag{35}$$

as $n \to \infty$. We leave this as an exercise, cf. [Kreyszig]. □

6.3 Final Remarks

We conclude with some final remarks on the proof of the uniform boundedness theorem and its application to the divergence of Fourier series. We first observe

Remark 6.1. *The conclusion of Theorem 6.2 also holds if*

$$|T_n(x)| < \infty$$

holds (only) for all x in some set E of the second category

Indeed, the proof goes through almost identically, define the $A_k \subset E$, by assumption we have $E = \cup_{k=1}^{\infty} A_k$ and one A_k must contain a ball...). This observations gives

Corollary 6.1. *The set of continuous functions whose Fourier series diverges at a point is generic (i.e. its complement is of the first category).*

By exercise 1 of Week 4 this implies that the set of continuous functions whose Fourier series diverges at a point is dense in the set of continuous functions.

Proof. Assume the complement (i.e. the set of continuous functions in X whose Fourier series converges at all points) was of the second category. Then in the notation of the proof of Proposition 6.1, we would have $|f_n(x)| < c_x$ on a set of the second category. By the above remark this would imply that $\|f_n\| \leq c$ which we have shown to be false in the previous section. □

6.4 Strong and Weak convergence

We now discuss an important application of the Uniform Boundedness theorem to weak convergence, a concept that is omnipresent in PDE.

Definition 6.2. *A sequence (x_n) in a normed space is said to be strongly convergent ("norm convergent") if there is an $x \in X$, such that*

$$\lim_{n\to\infty} \|x_n - x\| = 0 \,.$$

In this case we write $x_n \to x$, call x the strong limit of (x_n) and say (x_n) converges strongly to x.

Definition 6.3. *A sequence (x_n) in a normed space X is said to be weakly convergent if there is an $x \in X$ such that for every $f \in X'$ we have*

$$\lim_{n\to\infty} f(x_n) = f(x)$$

(convergence as a sequence of real (or complex) numbers for every $f \in X'$). We write $x_n \rightharpoonup x$, call x the weak limit of (x_n) and we say that (x_n) converges weakly to x.

Lemma 6.1. *Let (x_n) be a weakly convergent sequence in a normed space X, $x_n \rightharpoonup x$. Then*

1. *The weak limit x of (x_n) is unique.*
2. *Every subsequence of (x_n) converges weakly to x.*
3. *The sequence $\|x_n\|$ is bounded.*

Proof. For 1. say $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$. This means that

$$f(x_n) \to f(x) = f(y)$$

for every f because the limit of a converging sequence in $\mathbb{R}$ is unique. Therefore $f(x-y) = 0$ for every $f \in X'$ and by Hahn-Banach $x = y$. For 2. recall that in $\mathbb{R}$ (or $\mathbb{C}$) subsequences of converging sequences converge to the same limit. For 3. we will use the uniform boundedness theorem as follows. Since $f(x_n)$ is a converging sequence of real numbers it is bounded:

$$|f(x_n)| \leq c_f \quad \text{for all } n$$

Now use the canonical mapping (cf. Exercise 2 of Week 3) from X into its bidual,

$$C : X \to X''$$

to define $g_n(f) = g_{x_n}(f) = f(x_n)$. For all n we have

$$|g_n(f)| = |f(x_n)| \leq c_f \, .$$

Therefore, since X' is complete, by Banach Steinhaus, $\|g_n\| \leq c$. But the canonical mapping is an isometry so $\|g_n\| = \|x_n\|$ which establishes the boundedness of $\|x_n\|$. □

The distinction between weak and strong convergence is absent in finite dimensional spaces:

Theorem 6.3. *Let (x_n) be a sequence in a normed space X. Then*

1. *Strong convergence $\implies$ weak convergence with the same limit*
2. *Converse of* 1. *is not generally true*
3. *If* $\dim X < \infty$ *then weak convergence $\implies$ strong convergence*

Proof. The first part is immediate from $\|f(x_n) - f(x)\| \leq \|f\| \|x - x_n\|$. For the second consider the space ℓ^2 whose dual is also ℓ^2 is ℓ^2 (Exercise 1 of Week 3). Consider the sequence $e_n = (0, ..., 0, 1, 0...)$ in ℓ^2, i.e. the n^{th} entry of the sequence e_n is 1 while all others are zero. Then for any $\beta \in (\ell^2)' = \ell^2$ we have

$$\beta(e_n) = \sum_{j=1}^{\infty} \beta_j (e_n)_j = \beta_n$$

which goes to zero since $\sum |\beta_j|^2 < \infty$. So $e_n \rightharpoonup 0$ weakly. But clearly e_n does not converge strongly since $\|e_n - e_m\| = \sqrt{2}$ for $m \neq n$.[9]

Finally, for 3. suppose that $x_n \rightharpoonup x$ and $\dim X = k$. Pick any basis $e_1, ..., e_k$ for X and let

$$x_n = \alpha_1^{(n)} e_1 + \ldots + \alpha_k^{(n)} e_k \qquad \text{and} \qquad x = \alpha_1 e_1 + \ldots + \alpha_k e_k$$

Weak convergence means $f(x_n) \to f(x)$ for every $f \in X'$. Pick a dual basis $f_1, \ldots, f_k$ of $e_1, ..., e_k$, i.e. such that $f_j(e_i) = \delta^i_j$. Acting on the sequence x_n with every f_j we see that in particular $\alpha_j^{(n)} \to \alpha_j$ as $n \to \infty$ for every $j = 1, ...k$. But since

$$\|x_n - x\| = \left\| \sum_{j=1}^{k} \left(\alpha_j^{(n)} - \alpha_j \right) e_j \right\| \leq \sum_{j=1}^{k} |\alpha_j^{(n)} - \alpha_j| \|e_j\|$$

this implies $\|x_n - x\| \to 0$ and hence strong convergence. □

[9] For a more PDE inspired example consider the space $L^2[0, 2\pi]$ and the sequence $f_n(t) = \sin(nt)$. Show that $f_n \to 0$ weakly but not strongly.

Why should we care about weak convergence? Unfortunately, we still don't have the tools to really appreciate its PDE applications, so I will again have to be a bit vague. Suppose you have a sequence of approximate solutions (functions) x_n to an equation and you would like to extract a convergent subsequence. In some situations, establishing norm convergence can be impossible or too difficult. In such situation it is useful to first extract a weak limit (and then, if at all possible, use other techniques to show that the weak limit is actually also a strong limit).

A similar case can be made for sequences of operators. Sometimes we would like to approximate an operator T by a sequence of operators T_n which are easier to understand and we'll have to make precise it what sense our T_n converge to T:

Definition 6.4. *Let X and Y be normed spaces. A sequence $T_n \in B(X,Y)$ is said to be*

1. *uniformly operator convergent (to T) if there exists a linear operator $T : X \to Y$ such that*
$$\|T_n - T\| \to 0 \,,$$
i.e. the T_n converge in the norm of $B(X,Y)$ to a linear operator T.

2. *strongly operator convergent (to T) if there exists a linear operator $T : X \to Y$ such that*
$$\|T_n x - Tx\| \to 0 \quad \text{for all } x \in X \,,$$
i.e. the $T_n x$ converge strongly to Tx for any fixed x.

3. *weakly operator convergent (to T) if there exists a linear operator $T : X \to Y$ such that*
$$|f(T_n x) - f(Tx)| \to 0 \quad \text{for all } x \in X \text{ and all } f \in Y' \,,$$
i.e. the $T_n x$ converges weakly to Tx for any fixed x.

Lemma 6.2. *(1.) $\implies$ (2.) $\implies$ (3.).*

Proof. $|f(T_n x) - f(Tx)| \leq \|f\|\|T_n x - Tx\| \leq \|f\|\|T_n - T\|\|x\|$. □

The converse is generally false:

Example 6.2. *Consider ℓ^p for $1 < p < \infty$ with elements $x = (\xi_j)$ and the sequence of operators $T_n : \ell^p \to \ell^p$ defined by*
$$T_n x = (\xi_1, \xi_2, ..., \xi_n, 0, 0, ...) \,.$$
We have for fixed $x \in \ell^p$,
$$\|T_n x - \mathbf{1}x\| = \|(0, 0, ..., 0, \xi_{n+1}, \xi_{n+2}, ...)\| = \left[\sum_{j=n+1}^{\infty} |\xi_j|^p\right]^{(1/p)} \to 0$$

as $n \to 0$, so T_n converges strongly to the identity. On the other hand, $x = x_n$ with $(x_n)_j = \delta^{j+1}_{n+1}$ (i.e. x_n has $(n+1)^{th}$ entry 1 and all others zero) has $\|x\| = 1$ and hence

$$\|T_n - \mathbf{1}\| \geq \|T_n x_n - x_n\| = \|(0, 0, ..., 0, 1, 0, ...)\| = 1$$

while ≤ 1 is immediate, so $\|T_n - \mathbf{1}\| = 1$. It follows that T_n cannot converge uniformly in operator norm to the identity.

Example 6.3. *We let $T_n : \ell^2 \to \ell^2$ with $T_n x = (0, 0, ..., 0, \xi_1, \xi_2, ...)$. Let $b = (\beta_i) \in \left(\ell^2\right)' = \ell^2$ be arbitrary. Compute, for any x*

$$b\left(T_n x\right) = \sum_{i=1}^{\infty} \beta_i \left(T_n x\right)_i = \sum_{i=1}^{\infty} \beta_{i+n} \xi_i \leq \sqrt{\sum_{i=1}^{\infty} |\beta_{i+n}|^2} \sqrt{\sum_{i=1}^{\infty} |\xi_i|^2}$$

Now for β and x arbitrary but fixed, the right hand side goes to zero, so T_n is weakly operator convergent to the zero-operator. On the other hand, for $x = (1, 0, ...0)$ we have $\|T_m x - T_n x\| = \sqrt{2}$, so $T_n x$ cannot converge to zero.

What happens with Definition 6.4 if we restrict to functionals, i.e. $Y = \mathbb{R}$ or $Y = \mathbb{C}$? In this case 2. and 3. become equivalent concepts because $\mathbb{R}$ and $\mathbb{C}$ are finite dimensional and Theorem 6.3 applies. The remaining modes of convergence have special names:

Definition 6.5. *Let $f_n \in B(X, \mathbb{R})$ for X a normed space.*[10] *We say*

- *f_n converges strongly to f, written $f_n \xrightarrow{s} f$ if there exists an $f \in B(X, \mathbb{R})$ such that $\|f - f_n\| \to 0$. We call f the strong limit.*
- *f_n converges weak* to f, written $f_n \xrightarrow{w^\star} f$ if there exists an $f \in B(X, \mathbb{R})$ such that $f_n(x) \to f(x)$ for all $x \in X$.*

Exercise 6.1. *Of course there is also the notion of $f_n \in X'$ converging weakly to $f \in X'$, which – as we know – means that for all $g \in X''$ we have $g(f_n) \to g(f)$. Show that weak convergence implies weak*-convergence. Hint: Exercise 2 of Week 3.*

We return to the general Definition 6.4 and ask what can be said about the properties of the limiting T. In the case of uniform operator convergence $T \in B(X, Y)$ (as $T_n \in B(X, Y)$ and $T_n - T \in B(X, Y)$ by the very definition of that mode of convergence). For strong and weak operator convergence, the limiting T must not be bounded (It will still be linear, however (why?).) and even if it is, $\lim_{n\to\infty} \|T_n\| \neq \|T\|$ in general.

- Revisit Example 6.2, where $S_n := T_n - \mathbf{1}$ was strongly operator convergent to 0 but $\|S_n := T_n - \mathbf{1}\| = 1$ for all n. Conclude that if S_n is strongly operator convergent so S, then in general $\lim_{n\to\infty} \|S_n\| \neq \|S\|$.
- Consider $X \subset \ell^2$ the (incomplete) space of sequences with finitely many non-zero elements. Define

$$T_n x = (\xi_1, 2\xi_2, 3\xi_3, ..., n\xi_n, \xi_{n+1}, \xi_{n+2}, ...)$$

[10]The analogous definition can be made replacing $B(X, \mathbb{R})$ by $B(X, \mathbb{C})$).

We have $T_n \to T$ strongly in X for T defined by $Tx = (j\xi_j)$ (recall that elements in X have finitely many non-zero terms!). The operator T is unbounded.

As suggested by the last example, the failure of T to be bounded is indeed related to the incompleteness as a straightforward application of the uniform boundedness theorem gives

Lemma 6.3. *Let X be a Banach and Y a normed space, $T_n \in B(X,Y)$ a sequence of bounded linear operators. If T_n is strongly operator convergent to T, then $T \in B(X,Y)$.*

Proof. By assumption $\|T_n x\| < \infty$ for any x. Since X is complete, the Banach-Steinhaus theorem gives $\|T_n\| \leq c$ for some $c > 0$. But then

$$\|Tx\| \leq \|T_n x\| + \|Tx - T_n x\| \leq \|T_n\|\|x\| + \|Tx - T_n x\| \leq c\|x\| + \|Tx - T_n x\|$$

Taking $n \to \infty$ we find $\|Tx\| \leq c\|x\|$ hence $T \in B(X,Y)$. □

Remark 6.2. *Note that the above actually gives $\|T\| \leq \liminf \|T_n\|$. We have met an example with strict inequality above.*

6.4.1 The Banach-Alaoglu theorem

Recall that we showed that the closed unit ball is never compact in an infinite dimensional normed space. In this section we will see how weak convergence helps us to restore a version of sequential compactness for infinite dimensional normed spaces.

We first need the following criterion for strong operator convergence, which will be used in the proof of the Banach-Alaoglu theorem.

Proposition 6.2. *Let X be a normed space, Y a Banach space and $T_n \in B(X,Y)$ a sequence of operators. Then the sequence T_n is strongly operator convergent (i.e. $\|T_n x - Tx\| \to 0$ for all $x \in X$) if the sequence $\|T_n\|$ is bounded, $\|T_n\| \leq c$ AND the sequence $T_n x$ is Cauchy in Y for every x in a set M whose span is dense in X.*

Note that if X is complete, then the converse also holds from Lemma 6.3.

Proof. Let $x \in X$ be arbitrary. Given $\epsilon > 0$ we can choose a $y \in$ *span* M such that

$$\|x - y\| < \frac{\epsilon}{3c}$$

where c is such that $\|T_n\| \leq c$. Next choose N so large that

$$\|T_{n+k} y - T_n y\| < \frac{\epsilon}{3} \qquad \text{for } n \geq N \text{ and } k \geq 0\,,$$

which is possible by the Cauchy property of $T_n y$. Finally, apply the triangle inequality

$$\|T_{n+k} x - T_n x\| \leq \|T_{n+k} x - T_{n+k} y\| + \|T_{n+k} y - T_n y\| + \|T_n y - T_n x\| \leq \epsilon$$

for $n \geq N$ and $k \geq 0$ to show that $T_n x$ is also Cauchy and by the completeness converges to some $z := Tx \in Y$. □

Corollary 6.2. *A sequence f_n of bounded linear functionals on a normed linear space X is weak*-convergent if $\|f_n\| \leq c$ for all n and moreover the sequence $f_n(x)$ is Cauchy for every x in a subset $M \subset X$ whose span is dense in X.*

Theorem 6.4. *[Banach-Alaoglu] Every bounded sequence $\{f_n\}$ of functionals in the dual space X' of a separable normed linear space X contains a weak* convergent subsequence.*

Proof. Since X is separable, we can take $x_1, x_2, ..., x_n, ...$ dense in X. Since f_n is bounded the sequence $f_n(x_1)$ is bounded in $\mathbb{R}$ and we can extract by Bolzano-Weierstrass a convergent subsequence $f_n^{(1)}(x_1) \to a_1$. Next, the sequence $f_n^{(1)}(x_2)$ is a bounded sequence in $\mathbb{R}$ and we similarly extract a convergent subsequence of $f_n^{(2)}(x_2) \to a_2$. Continuing we get a system of subsequences with the property

- $f_n^{(k+1)}$ is a subsequence of $f_n^{(k)}$ for all $k = 1, 2, ...$
- $f_n^{(k)}$ converges at the points $x_1, x_2, ..., x_k$.

Taking the diagonal sequence ("Cantor's diagonal argument"), $f_1^{(1)}, f_2^{(2)}, ..., f_n^{(n)}, ...$ we see that it converges for any x_k, i.e. $f_n^{(n)}(x_k)$ converges for any x_k. By Proposition 6.2 (resp. its Corollary), this subsequence converges for all x. □

Corollary 6.3. *In a Hilbert space H every bounded sequence has a weakly convergent subsequence.*

7 The open mapping and closed graph theorem

We now turn to discuss the open mapping theorem. Next to the Hahn-Banach- and the Uniform Boundedness Theorem it is the third "big theorem" in functional analysis. It will help us to answer the question when the inverse of a bounded linear operator exists and is continuous.

Definition 7.1. *Let X and Y be metric spaces. We call*

$$T : \mathcal{D}(T) \to Y$$

with domain $\mathcal{D}(T) \subset X$ an open mapping if for every open set in $\mathcal{D}(T)$ the image is an open set in Y.

Remark 7.1. *Do not confuse this with continuity: T is continuous if and only if the pre-image of any open set is open. A continuous map is not open in general as the example $t \mapsto \sin t$, which maps $(0, 2\pi)$ to $[-1, 1]$ shows.*

Theorem 7.1. *A bounded linear operator T from a Banach space X onto (surjective!) a Banach space Y is an open mapping. Hence if T is bijective, T^{-1} is continuous and thus bounded.*

Surjectivity is crucial as you saw in Exercise 1 of Week 2 and Remark 4.1 . The proof will follow from the following

Lemma 7.1. *Let X, Y be Banach spaces and T a bounded linear operator from X onto Y. Then the image $T(B_0)$ of the open unit ball $B_0 = B(0,1) \subset X$ contains an open ball around $0 \in Y$.*

Proof. **Step 1: We show $\overline{T(B_1)} = \overline{T(B(0,1/2))}$ contains an open ball $B^\star$** (not necessarily centered around the origin)

We introduce the following notation. For $A \subset X$ we write

$$\alpha A = \{x \in X \mid x = \alpha a \ , \ a \in A\}$$
$$A + w = \{x \in X \mid x = a + w \ , \ a \in A\} \tag{36}$$

for the dilation (by α) and translation (by w) of the set A.

We consider the open ball $B_1 = B(0,1/2) \subset X$. Any $x \in X$ is contained in kB_1 for some k, hence

$$X = \bigcup_{k=1}^{\infty} kB_1$$

Since T is surjective and linear, we have

$$Y = T(X) = T\left(\bigcup_{k=1}^{\infty} kB_1\right) = \bigcup_{k=1}^{\infty} kT(B_1) = \bigcup_{k=1}^{\infty} \overline{kT(B_1)}$$

with the closure not adding any point since the left hand side is already the whole space Y. By Baire Category one of the $\overline{kT(B_1)}$ contains an open ball, and by dilating also $\overline{T(B_1)}$ contains an open ball, say $B^\star = B(y_0, \epsilon) \subset \overline{T(B_1)}$. Translating it follows that

$$B^\star - y_0 = B(0,\epsilon) \subset \overline{T(B_1)} - y_0 \tag{37}$$

Step 2. We show $B^\star - y_0 = B(0,\epsilon) \subset \overline{T(B_0)}$.

To achieve this, by (37) it suffices to show $\overline{T(B_1)} - y_0 \subset \overline{T(B_0)}$. Let $y \in \overline{T(B_1)} - y_0$. Then $y + y_0 \in \overline{T(B_1)}$ and also $y_0 \in \overline{T(B_1)}$. Take sequences $u_n \in T(B_1)$ with $u_n \to y + y_0$ and $v_n \in T(B_1)$ with $v_n \to y_0$ and find corresponding $w_n \in B_1$ and $z_n \in B_1$ such that $u_n = Tw_n$ and $v_n = Tz_n$. Since w_n and z_n are in B_1, we have

$$\|w_n - z_n\| \le \|w_n\| + \|z_n\| < \frac{1}{2} + \frac{1}{2} = 1$$

so $w_n - z_n \in B_0$. Now

$$T(w_n - z_n) = u_n - v_n \to y$$

which shows that $y \in \overline{T(B_0)}$. Note also that now that we have $B(0,\epsilon) \subset \overline{T(B_0)}$, we have by dilation

$$V_n := B\left(0, \frac{\epsilon}{2^n}\right) \subset 2^{-n}\overline{T(B_0)} = \overline{T(B_n)} \tag{38}$$

where $B_n = B(0, 2^{-n})$.

Step 3: We finally prove that $V_1 = B\left(0, \frac{1}{2}\epsilon\right) \subset T(B_0)$

Let $y \in V_1$. By (38) with $n = 1$ we have $V_1 \subset \overline{T(B_1)}$, hence $y \in \overline{T(B_1)}$. We choose $v \in T(B_1)$ such that $\|y - v\| < \frac{\epsilon}{4}$. Since $v \in T(B_1)$ we have an $x_1 \in B_1$ such that $\|y - Tx_1\| < \frac{\epsilon}{4}$. But this means $y - Tx_1 \in V_2 \subset \overline{T(B_2)}$ and hence we find an $x_2 \in B_2$ such that $\|y - Tx_1 - Tx_2\| < \frac{\epsilon}{8}$. Continuing inductively, we find

$$\|y - \sum_{k=1}^{n} Tx_k\| < \frac{\epsilon}{2^{n+1}} \tag{39}$$

Letting $z_n = x_1 + x_2 + \dots x_n$ we note that z_n is Cauchy, since for $n > m$

$$\|z_n - z_m\| = \sum_{k=m+1}^{n} \|x_k\| < \sum_{k=m+1}^{\infty} \frac{1}{2^k} \to 0$$

as $m \to \infty$ and as X is complete $z_n \to x$ for some $x \in X$. Now since T is continuous, we have $Tz_n \to Tx$ and (39) shows $y = Tx$. Finally, from

$$\|\sum_{k=1}^{\infty} x_k\| \leq \sum_{k=1}^{\infty} \|x_k\| < \frac{1}{2} + \sum_{k=2}^{\infty} \|x_k\| < 1$$

we see that $x \in B_0$, which proves that $y \in T(B_0)$. □

We are now ready to prove Theorem 7.1.

Proof. We need to show that for every open set $A \subset X$, $T(A)$ is open in Y. We show that every $y \in T(A)$ admits an open ball around it lying entirely in $T(A)$. So fix $y \in T(A)$ and let $Tx = y$. Since A is open we find an open ball around x lying entirely in A. Then $A - x$ contains an open ball $B(0, r)$, $r > 0$ around the origin. This means that $\frac{1}{r}(A - x)$ contains the unit ball $B(0, 1)$. By the Lemma,

$$T\left(\frac{1}{r}(A - x)\right)$$

contains an open ball around zero, hence $T(A - x)$ also contains an open ball around zero. But then $T(A)$ contains an open ball around $Tx = y$, which is what we needed to show.

Finally, for the continuity of the inverse we note that T^{-1} exists (T bijective) and is continuous because T is open. □

7.1 The closed graph theorem

We will now consider a new class of operators, "closed linear operators". While many operators in applications are not bounded (recall the differentiation operator), the majority of them is actually closed, which will justify our efforts. The definition may be a bit cumbersome at first but it will become more familiar once we consider a couple of examples.

Definition 7.2. *Let X, Y be normed spaces and $T : D(T) \to Y$ be a linear operator with domain $D(T) \subset X$. The operator T is called a closed linear operator if its graph*

$$\mathcal{G}(T) = \{(x, y) \mid x \in D(T) \ , \ y = Tx\}$$

is closed in the normed space $X \times Y$ equipped[11] *with the norm*

$$\|(x, y)\|_{X \times Y} = \|x\|_X + \|y\|_Y .$$

[11] The linear operations in $X \times Y$ are the obvious ones inherited from X and Y, i.e. $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\lambda(x, y) = (\lambda x, \lambda y)$.

We immediately note an alternative characterization of a closed linear operator, which is taken as the definition in many books and often easier to work with.

Theorem 7.2. *Let X, Y be normed spaces and $T : D(T) \to Y$ be a linear operator with domain $D(T) \subset X$. Then T is a closed linear operator if and only if it has the following property*

$$D(T) \ni x_n \to x \text{ and } Tx_n \to y \implies x \in D(T) \text{ and } Tx = y. \tag{40}$$

Proof. Let T be a closed linear operator and consider a sequence $x_n \in D(T)$ such that $x_n \to x$ and $Tx_n \to y$. This implies $(x_n, Tx_n) \to (x, y)$ in $X \times Y$. Since the graph $\mathcal{G}(T)$ is closed, $(x, y) \in \mathcal{G}(T)$ which implies that $x \in D(T)$ and $y = Tx$. Conversely, to show that the graph is closed assuming (40) holds, we pick $z = (x, y) \in \overline{\mathcal{G}(T)}$ and a sequence $\mathcal{G}(T) \ni z_n = (x_n, Tx_n) \to z = (x, y)$. In particular $x_n \to x$ and $Tx_n \to y$. Using property (40) we deduce that $x \in D(T)$ and $Tx = y$ hence $(x, y = Tx) \in \mathcal{G}(T)$. We have shown $\overline{\mathcal{G}(T)} = \mathcal{G}(T)$ which is the statement that the graph is closed. □

Note that if $D(T)$ is closed in X then a continuous linear operator is closed.[12] Conversely, a closed linear operator is not necessarily continuous (see below). The main result of this section says that if $D(T)$ is closed in X, then a closed linear operator is also continuous. In particular, a closed linear operator between two Banach spaces is continuous.

Theorem 7.3. *Let X and Y be Banach spaces and $T : D(T) \to Y$ a closed linear operator, where $D(T) \subset X$. Then if $D(T)$ is closed in X, the operator T is bounded.*

Proof. Note first that $X \times Y$ is complete w.r.t. the norm $\| (x, y) \| = \|x\| + \|y\|$. Indeed, if z_n is Cauchy with respect to $\| \cdot \|_{X \times Y}$ this implies x_n is Cauchy in X and y_n is Cauchy in Y. By completeness of X and Y, $x_n \to x \in X$ and $y_n \to y \in Y$. One then easily shows that $z_n \to z = (x, y) \in X \times Y$.

By assumption $\mathcal{G}(T)$ is closed in $X \times Y$ and $D(T)$ is closed in X. Being closed subspaces of complete spaces both $\mathcal{G}(T)$ and $D(T)$ are themselves complete. Define

$$P : \mathcal{G}(T) \to D(T) \quad \text{by } P(x, Tx) = x$$

Note that P is linear and bounded in view of

$$\|P(x, Tx)\|_X = \|x\|_X \leq \|x\|_X + \|Tx\|_Y = \| (x, Tx) \|_{X \times Y} .$$

It is also bijective, the inverse being $P^{-1} : D(T) \to \mathcal{G}(T)$

$$P^{-1}(x) = (x, Tx) .$$

By the bounded inverse theorem, P^{-1} is bounded, i.e. $\| (x, Tx) \|_{X \times Y} \leq c\|x\|_X$ for some $c > 0$ and this implies that T is bounded, since

$$\|Tx\|_Y \leq \|x\|_X + \|Tx\|_Y = \| (x, Tx) \|_{X \times Y} \leq c\|x\|_X$$

for all $x \in D(T)$. □

[12] If $D(T)$ is not closed in a normed space X, then a bounded linear operator does not have to be closed. You can take the identity operator defined on a proper dense subset of X. It is clearly bounded but taking an $x \in X \setminus D(T)$ and a sequence in $D(T)$ converging to x shows that it is not closed.

As the prime example of a closed linear operator which is not bounded we consider the differential operator:

Example 7.1. *We let $X = C[0,1]$ and*

$$T : D(T) \to X \quad \text{defined by } Tx = x'$$

with a prime denoting differentiation and $D(T) = C^1[0,1] \subset C[0,1]$. Let $x_n \in D(T)$ be such that $x_n \to x$ and $Tx_n = x_n' \to y$. We want to show that actually $x \in D(T)$ and $Tx = y$ (cf. Theorem 7.2). Now $x_n' \to y$ implies that in

$$\int_0^t y(\bar{t})\, d\bar{t} = \int_0^t x_n'(\bar{t})\, d\bar{t} + \int_0^t \left[y(\bar{t}) - x_n'(\bar{t})\right] d\bar{t}$$

the second term goes to zero uniformly in t so that

$$\int_0^t y(\bar{t})\, d\bar{t} = \lim_{n\to\infty} \int_0^t x_n'(\bar{t})\, d\bar{t} = x(t) - x(0) \ .$$

This implies that $x(t) = x(0) + \int_0^t y(\bar{t})\, d\bar{t}$, which is continuously differentiable, hence in $D(T)$ and $x' = y$ showing that T is closed. Note that $D(T)$ cannot be closed in X because then T would be bounded by the closed graph theorem (which we know it isn't!)!

8 Hilbert Spaces

8.1 Basic definitions

Definition 8.1. *Let H be a complex linear vector space. A complex valued function $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ is called an inner-product (or scalar-product) on H if it has the following properties for any $x, y, z \in H$ and $\lambda \in \mathbb{C}$*

1. $\langle x, x \rangle \geq 0$ *with equality iff $x = 0$ (positivity)*
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ *(linearity in first component)*
3. $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ *(linearity in first component)*
4. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ *(sesquilinearity)*

Note that the last condition implies that $\langle x, \lambda y \rangle = \overline{\langle \lambda y, x \rangle} = \overline{\lambda \langle y, x \rangle} = \overline{\lambda} \langle x, y \rangle$ so the inner-product is *conjugate linear* in the second component.

An inner-product induces a norm via

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

as you can easily check.

The inner product also allows one to talk about orthogonality: We say that "x is orthogonal to y" for $x, y \in H$ if $\langle x, y \rangle = 0$ and use the notation $x \perp y$ to indicate that x is orthogonal to y.

Definition 8.2. *A complex linear vector space equipped with an inner-product is called an inner-product space. A complete (with respect to the induced norm) inner-product space is called a Hilbert space.*

The most familiar examples of inner-product spaces are $\mathbb{R}^n$ and $\mathbb{C}^n$ with the natural scalar products (which?). They are also complete. To have an infinite dimensional example, consider the space of continuous (say real-valued) functions on the interval $[a,b]$ equipped with the inner-product

$$\langle f,g\rangle = \int_a^b f(t)\cdot g(t)\ dt$$

You can check that this indeed defines an inner-product. However, the space of continuous functions equipped with this inner-product is incomplete (why?). In can be completed (similarly to the procedure for metric spaces and normed spaces) whereby one obtains the space $L^2[a,b]$ that you have met in your measure theory course by now.

Exercise 8.1. *Check that in an inner-product space the* <u>*paralellogram identity*</u> *holds:*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \tag{41}$$

Can you give an explanation of the name?

From the above definition we see that all Hilbert spaces are Banach spaces. Is the converse true? The answer is no. In fact, we have

Proposition 8.1. *A (real or complex) Banach space H is a Hilbert space (i.e. the norm is induced from an inner product) if and only if the norm satisfies (41).*

Proof. See Theorem 6.1.5 in [Friedman]. The idea is the following. We define ("polarization identity")

$$\langle x,y\rangle := \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) + \frac{1}{4}i\left(\|x+iy\|^2 - \|x-iy\|^2\right) \tag{42}$$

and use the fact that (41) holds for the norm to prove that the above is indeed an inner-product on H. (If H is a real Banach space we drop the second round bracket.) □

Exercise 8.2. *Show that ℓ^p is a Hilbert space if and only if $p=2$.*

Exercise 8.3. *If X and Y are Hilbert spaces, is $B(X,Y)$ a Hilbert space?*

Exercise 8.4. *The following inequalities hold in an inner-product space H for $x,y\in H$.*

$$\begin{aligned} |\langle x,y\rangle| &\le \|x\|\|y\| \quad \textit{Cauchy-Schwarz inequality} \\ \|x+y\| &\le \|x\| + \|y\| \quad \textit{triangle inequality} \end{aligned} \tag{43}$$

When do you have equality in the above?

Lemma 8.1. *The inner-product is continuous.*

Proof. If $x_n \to x$ and $y_n \to y$, then

$$\langle x_n, y_n\rangle - \langle x,y\rangle = \langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle \le \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\|$$

□

Corollary 8.1. *If $y \perp x_n$ and $x_n \to x$ then $x \perp y$.*

Exercise 8.5. *Let $T : X \to X$ be a bounded linear operator on a complex inner product space X. If $\langle Tx, x\rangle = 0$ for all $x \in X$, then $T = 0$. Is this true in a real inner-product space? Hint: Rotations.*

8.2 Closed subspaces and distance

In any metric space (X, d), the distance of an element $x \in X$ to a non-empty subset $M \subset X$ is defined as

$$\delta := \inf_{\bar{y} \in M} d(x, \bar{y}) \tag{44}$$

In a normed space, this becomes

$$\delta = \inf_{\bar{y} \in M} \|x - \bar{y}\| .$$

It is a natural question under what circumstances this inf is achieved and unique. Experimentation with sets in $\mathbb{R}^2$ provide simple examples of situations when the inf is not achieved or situations when it is non-unique.

Theorem 8.1. *Let X be an inner-product space, $M \subset X$ be non-empty, convex and complete. Then for every $x \in X$ there is a unique $y \in M$ such that*

$$\delta = \inf_{\bar{y} \in M} \|x - \bar{y}\| = \|x - y\|$$

Proof. By the definition of the inf we have a sequence (y_n) such that $\|x - y_n\| = \delta_n \to \delta$. We show that y_n is Cauchy. Defining $v_n = y_n - x$ we have $\|v_n\| = \delta_n$ and

$$\|v_n + v_m\| = 2\|\frac{1}{2}(y_n + y_m) - x\| \geq 2\delta$$

by convexity. The parallelogram identity gives

$$\|y_n - y_m\|^2 = \|v_n - v_m\|^2 = -\|v_n + v_m\|^2 + 2\|v_m\|^2 + 2\|v_n\|^2$$

The right hand side goes to zero (why?), hence (y_n) is Cauchy and $y_n \to y \in M$ by completeness. Therefore $\|x - y\| \geq \delta$ and since the reverse inequality holds from $\|x - y\| \leq \|x - y_n\| + \|y_n - y\|$, we have $\|x - y\| = \delta$.

To show uniqueness assume $y_0, y_1 \in M$ with $\|x - y_0\| = \delta = \|x - y_1\|$. Then by the paralellogram identity

$$\|y_0 - y_1\|^2 = \|(y_0 - x) - (y_1 - x)\|^2 \leq -\|(y_0 - x) + (y_1 - x)\|^2 + 4\delta^2$$

and hence again by convexity

$$\|y_0 - y_1\|^2 \leq -4\|\frac{1}{2}(y_0 + y_1) - x\|^2 + 4\delta^2 \leq 0 .$$

□

Corollary 8.2. *If M is a closed subspace of a Hilbert space, then the distance of M to a given point $x \in X$ is achieved by a unique $y \in M$. Moreover, this $y \in M$ has the property that $z = x - y$ is orthogonal to M.*

Proof. Only the last statement doesn't immediately follow from the above theorem. To prove it assume there was a $\tilde{y} \in M$ with $\langle z, \tilde{y} \rangle = \beta \neq 0$. Clearly $\tilde{y} \neq 0$. Now check that

$$\|z - \frac{\bar{\beta}}{\langle \tilde{y}, \tilde{y} \rangle} \tilde{y}\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle \tilde{y}, \tilde{y} \rangle} < \delta^2$$

with the last inequality following from $\|z\| = \delta$. But this is a contradiction to the minimization since we can write the left hand side as

$$\|x - \left(y + \frac{\bar{\beta}}{\langle \tilde{y}, \tilde{y} \rangle} \tilde{y} \right) \|^2 \geq \delta$$

following from the fact that the round bracket is an element of M. □

Note that the above Corollary is very familiar from the finite dimensional setting, where you find the least distance to a subspace by taking a perpendicular.

Theorem 8.2. *Let Y be a closed subspace of a Hilbert space H. Then one has the decomposition*

$$H = Y \oplus Y^{\perp}$$

where

$$Y^{\perp} = \{ z \in H \quad | \quad z \perp Y \} \tag{45}$$

is the orthogonal component of Y.

Proof. Since Y is a closed subspace of a complete normed space Y is complete. Since Y is also convex we can apply Theorem 8.1 to find for a given $x \in H$ a $y \in Y$ such that $x = y + z$ where $z = x - y$ is in $Y^{\perp}$ by the above corollary. To show the uniqueness of the decomposition of x assume $x = y + z = y' + z'$. But then $y - y' = -z + z'$ and the left hand side is in Y while the right hand side is in $Y^{\perp}$. Since $Y \cap Y^{\perp} = \{0\}$ (why?) uniqueness follows. □

The $y \in Y$ in the decomposition $x = y + z$ ($z \in Y^{\perp}$) is called the orthogonal projection of x on Y. The decomposition promised by the theorem allows one to define the orthogonal projection as a map

$$P : H \to Y$$

which maps $P(x) = y$. P is called a projection operator. It satisfies $P^2 = P$.

Exercise 8.6. *Let $M \subset H$ be a subspace of a Hilbert space H. Show that*

- *$M^{\perp}$ is always closed (independently of whether M is) and $M^{\perp} = \overline{M}^{\perp}$.*
- *$M \subset M^{\perp\perp} = \left(M^{\perp}\right)^{\perp}$*
- *$M = M^{\perp\perp}$ if M is closed.*

Lemma 8.2. *For any non-empty subset $M \subset H$ of a Hilbert space H, the span of M is dense in H if and only if $M^{\perp} = \{0\}$.*[13]

[13] We can define $M^{\perp} := \{x \in H \mid \langle x, y \rangle = 0 \text{ for all } y \in M\}$ for an arbitrary subset $M \subset H$.

Proof. Suppose $spanM$ is dense in H and pick an arbitrary element $x \in M^\perp$. By density we have $x_n \to x$ for sequence $x_n \in spanM$. Hence $\langle x_n, x\rangle = 0$ for all n. By continuity of the inner-product, $\langle x, x\rangle = 0$ and hence $x = 0$. Conversely, let $M^\perp = \{0\}$. Setting $V = spanM$, then if $x \perp V$ we have $x \perp M$, hence $x \in M^\perp$, hence $V^\perp = \{0\}$. Since $V^\perp$ is closed and $V^\perp = \overline{V}^\perp$, we have the unique decomposition $H = \overline{V}^\perp \oplus \overline{V}^{\perp\perp} = \overline{V}$. □

8.3 Orthonormal sets and sequences [lecture by I.K.]

This section will be expanded once I revise the notes. For now it's more of a summary.

- Define orthogonal set (elements pairwise orthogonal) and orthonormal set (also have norm 1) in inner-product space H; orthogonal sequence if the elements are countable
- show that orthonormal set is linearly independent
- Discuss the example of $L^2[0, 2\pi]$. Check that

$$e_0(t) = \frac{1}{\sqrt{2\pi}} \quad , \quad e_n(t) = \frac{\cos(nt)}{\sqrt{\pi}} \quad , \quad \tilde{e}_n(t) = \frac{\sin(nt)}{\sqrt{\pi}}$$

  form an orthonormal set. ($\to$ Fourier series)

- Bessel's inequality: Let (e_k) an orthonormal sequence in an inner-product space X. Then for every $x \in X$ we have

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2$$

  The $\langle x, e_k\rangle$ are called the Fourier coefficients of x with respect to the e_k. Exercise: Prove Bessel's inequality. Idea of proof: Work with $Y_n = span\{e_1, ...e_n\}$ Given $x \in X$ define $y = \sum_{i=1}^{n} \langle x, e_k\rangle e_k$. Show that $z = x-y$ is perpendicular to y. Then Pythagoras implies $\|x\|^2 = \|y\|^2 + \|z\|^2$ and from there the result follows.

- Gram Schmidt process: Construct orthonormal sequence from linearly independent sequence

Given an orthonormal sequence (e_k) in a Hilbert space H we can consider series of the form

$$\sum_{k=1}^{\infty} \alpha_k e_k \tag{46}$$

When does the sequence of partial sums $s_n = \sum_{k=1}^{n} \alpha_k e_k$ converge to some s, i.e. when do we have $\|s_n - s\| \to 0$ as $n \to \infty$?

Theorem 8.3. *Let (e_k) be an orthonormal sequence (e_k) in an Hilbert space H. Then*

1. *The sum (46) converges if and only if* $\sum_{k=1}^{\infty} |\alpha_k|^2$ *converges*

2. *If (46) converges to some* $x \in X$*, then* $\alpha_k = \langle x, e_k \rangle$ *with* $k = 1, 2, \ldots$ *are the Fourier coefficients of* x *with respect to the sequence* e_k

3. *For any* $x \in H$*, the sum*

$$\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$$

converges.

Proof. For 1. we note that s_n Cauchy in H is the statement that $\| \sum_{j=n}^{n+k} \alpha_j e_j \|^2 \to 0$, which is equivalent to $\sum_{j=n}^{n+k} |\alpha_j|^2 \to 0$ (why?) which is the definition of the sequence $\tilde{s}_n = \sum_{j=1}^{n} |\alpha_j|^2$ being Cauchy in $\mathbb{R}$.
For 2., we note that clearly $\langle s_n, e_j \rangle = \alpha_j$ holds for $j = 1, 2, ..., k \leq n$. By continuity of the inner product, $\langle x, e_j \rangle = \alpha_j$ and this holds for any j since we let $n \to \infty$. Finally, for 3. note that by 1. the statement is equivalent to $\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$ converging. But this sum indeed converges by Bessel's inequality. □

Now suppose X is an inner product space and you have a (possibly uncountable) orthonormal family e_λ where $\lambda \in I$ for the index set I. We can again consider the Fourier coefficients $\langle x, e_\lambda \rangle$ for $x \in X, \lambda \in I$. We know that

$$\sum_{l=1}^{n} |\langle x, e_l \rangle|^2 \leq \|x\|^2 \qquad \text{for any } n$$

for any selection $e_1, ..., e_n$ from the orthonormal family by Bessel. The convergence of the sum already implies that the number of Fourier-coefficients with $|\langle x, e_\lambda \rangle| > \frac{1}{m}$ must be finite and hence

Lemma 8.3. *Let* X *be an inner-product space. Any* $x \in X$ *has at most countably many non-zero Fourier-coefficients* $\langle x, e_\lambda \rangle$ *with respect to an orthonormal family* e_λ *(for* $\lambda \in I$*) in* X*.*

So given x, you can associate with it the expression $\sum_\lambda \langle x, e_\lambda \rangle$ and rearrange the non-zero Fourier coefficients in a sequence such that $\sum_{l=1}^{\infty} \langle x, e_l \rangle$. Convergence always follows from Bessel and the only thing you need to prove is the independence of the sum from the rearrangement (see [Kreyszig]).

8.4 Total Orthonormal Sets and Sequences

A set M is a total set in a normed space X if $\overline{spanM} = X$. We can hence speak about a total orthonormal set or total orthonormal sequence in X. A total orthonormal set is sometimes called an orthornomal basis (although it is NOT a (Hamel) basis in the algebraic sense unless we are in the finite dimensional case!).

Theorem 8.4. *In every Hilbert space* $H \neq \{0\}$ *there exists a total orthonormal set.*

Sketch. In the separable case we extract from the countable dense subset of H a countable linearly independent set (how?) and via Gram Schmidt a (total) orthonormal set. In the non-separable case one has to appeal to Zorn's Lemma again (consider the set of all orthonormal subsets of H...). $\square$

One can also show that all total orthonormal sets in a given Hilbert space have the same cardinality. Let us emphasize that in the remainder of the course **we will only be dealing with separable Hilbert spaces**. In fact, in some books the "separability" is part of the definition of a Hilbert space [Stein]. In this case all total orthonormal sets are countable.

The following theorem shows that a total orthonormal system cannot be augmented.

Theorem 8.5. *Let M be a subset of an inner-product space X. Then*

- *If M is total in X, then $x \perp M$ implies $x = 0$.*
- *If X is complete, then the condition "$x \perp M$ implies $x = 0$" implies that M is total in X.*

Proof. Exercise. Use Lemma 8.2. (Completeness is actually necessary.) $\square$

Here is a useful criterion for totality of an orthonormal set:

Theorem 8.6. *An orthonormal set M in a Hilbert space H is total in H iff for all $x \in H$ we have the Parseval relation*

$$\sum_k |\langle x, e_k\rangle|^2 = \|x\|^2 \tag{47}$$

i.e. equality in Bessel's inequality.

Sketch of proof: please fill in the details! I am assuming separable here... If M is not total, then there exists an $x \neq 0$ with $x \perp M$. This violates (47). Conversely let M be total and $x \in H$ be given. Then $y = \sum_k \langle x, e_k\rangle e_k$ converges by Theorem 8.3. We need to show $x - y \perp M$ which is a direct computation. Now $x - y \in M^\perp$ and M total in H implies $M^\perp = \{0\}$ by the previous theorem. Hence $x = \sum_k \langle x, e_k\rangle e_k$ and (47) follows by direct computation. $\square$

The following result was already mentioned implicitly

Theorem 8.7. *Let H be a Hilbert space. Then*

1. *If H is separable, then every orthonormal set is countable.*
2. *If H contains a countable orthonormal set which is total in H then H is separable.*

Proof. For the first part, let H be separable and B any dense set in H. Given any orthonormal set M we have $\|x - y\| = \sqrt{2}$ for any $x \neq y$ in M. Hence, choosing balls of radius $1/4$ around each element of M, the balls are all disjoint. By density every ball needs to contain at least one element of B. If there were uncountably many elements of M, hence uncountably many balls, then B would be uncountable. This is in contradiction with the separability of H.
The second part is left as an exercise (Take finite linear combinations with rational coefficients). $\square$

We now prove a surprising result, namely that up to isomorphism there is only one separable Hilbert space.

Definition 8.3. *Let H, $\tilde{H}$ be Hilbert-spaces over the same field $\mathbb{R}$ or $\mathbb{C}$. An isomorphism between H and $\tilde{H}$ is a bijective linear map $T : H \to \tilde{H}$ such that*

$$\langle Tx, Ty \rangle = \langle x, y \rangle \qquad \textit{(isometry)}$$

If such an isomorphism exists, H and $\tilde{H}$ are called isomorphic Hilbert spaces.

Theorem 8.8. *Let H and $\tilde{H}$ be two separable infinite dimensional Hilbert spaces. Then H and $\tilde{H}$ are isomorphic.*

Proof. Let $M = \{e_1, e_2, ...\}$ and $\tilde{M} = \{\tilde{e}_1, \tilde{e}_2, ...\}$ be (countable) total orthonormal sets in H and $\tilde{H}$ respectively (cf. Theorems 8.4 and 8.7). Given x we can write (why?)

$$x = \sum_k \langle x, e_k \rangle e_k$$

We now define a map T by

$$\tilde{x} = Tx = \sum_k \langle x, e_k \rangle \tilde{e}_k$$

The series on the right hand side converges (why?). Also T is linear as the inner product is linear in the first component. The map T is also an isometry since

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle \sum_k \langle x, e_k \rangle \tilde{e}_k, \sum_j \langle x, e_j \rangle \tilde{e}_j \rangle = \sum_k |\langle x, e_k \rangle|^2 = \|x\|^2$$

The polarization identity (42) then implies that in fact $\langle Tx, Ty \rangle = \langle x, y \rangle$. Finally, T is also surjective. To see this, consider any $\tilde{x} \in \tilde{H}$, which we can write as $\tilde{x} = \sum_k \alpha_k \tilde{e}_k$ for α_k with $\sum_k |\alpha_k|^2 < \infty$. The latter means that also $\sum_k \alpha_k e_k$ converges to some x and that $\alpha_k = \langle x, e_k \rangle$. Hence $Tx = \tilde{x}$. □

Corollary 8.3. *Any separable infinite dimensional Hilbert space is isomorphic to $L^2[0, 2\pi]$ and to ℓ^2.*

8.5 Riesz Representation Theorem

Theorem 8.9. *Every bounded linear functional f on a Hilbert space H can be represented in terms of the inner product, namely*

$$f(x) = \langle x, z \rangle$$

where z depends on f (is uniquely determined by it) and has norm $\|z\| = \|f\|$.

Proof. If $f = 0$ we take $z = 0$. Let $f \neq 0$ and consider the null space $N(f)$ of f. We know it is closed and that $N(f) \neq H$ (as otherwise $f = 0$). So $N(f)^\perp$ contains a $z_0 \neq 0$. Set $v = f(x) z_0 - f(z_0) x$ for an arbitrary $x \in H$. One easily sees $f(v) = 0$ so $v \in N(f)$. Now

$$0 = \langle v, z_0 \rangle = f(x) \langle z_0, z_0 \rangle - f(z_0) \langle x, z_0 \rangle$$

and hence

$$f(x) = f(z_0)\frac{\langle x, z_0\rangle}{\langle z_0, z_0\rangle} = \langle x, \frac{\overline{f(z_0)}}{\langle z_0, z_0\rangle} z_0\rangle$$

which establishes existence of the representation. To see uniqueness, note that $f(x) = \langle x, z_1\rangle = \langle x, z_2\rangle$ for all $x \in H$ implies $\langle x, z_1 - z_2\rangle$ for all $x \in H$, in particular for $x = z_1 - z_2$ which immediately implies $z_1 - z_2 = 0$. To see the isometry, note that for $f(x) = \langle x, z\rangle$ we have

$$\|z\|^2 = \langle z, z\rangle = f(z) \leq \|f\|\|z\|$$

and by Cauchy-Schwarz, $|f(x)| \leq \|x\|\|z\|$. Combining the two inequalities yields $\|f\| = \|z\|$. □

8.6 Applications of Riesz' theorem & the Hilbert-adjoint

Our next goal is a slight generalization of the Riesz representation theorem which will allow us to define the notion of the Hilbert adjoint. We will then relate the Hilbert adjoint to the general adjoint defined in Section 5.5. Another important application of Riesz' theorem is the Lax-Milgram theorem, which you will study in Exercise 4 of Section 11.8.

Definition 8.4. *Let X and Y be vector spaces over K (which is $\mathbb{R}$ or $\mathbb{C}$). Then a sesquilinear form h on $X \times Y$ is a mapping*

$$h : X \times Y \to K$$

such that for all $x, x_1, x_2 \in X$ and $y, y_1, y_2 \in Y$ and all scalars $\alpha, \beta \in K$ we have

$$h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y) \tag{48}$$
$$h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2) \tag{49}$$
$$h(\alpha x, y) = \alpha h(x, y) \tag{50}$$
$$h(x, \beta y) = \overline{\beta} h(x, y) \tag{51}$$

So a sesquilinear form is linear in the first component and conjugate linear in the second. In the real case h is just a bilinear form.

When X and Y are normed we can talk about bounded sesquilinear forms: h is said to be bounded if there exists a $c \in \mathbb{R}$ such that

$$|h(x, y)| \leq c\|x\|\|y\|\,.$$

For a bounded sesquilinear form h we define the norm of h to be

$$\|h\| = \sup_{x \in X\setminus\{0\}, y \in Y\setminus\{0\}} \frac{|h(x, y)|}{\|x\|\|y\|} = \sup_{\|x\|=1, \|y\|=1} |h(x, y)| \tag{52}$$

The usual inner-product is clearly a bounded sesquilinear form (why?). The next theorem shows that every sesqui-linear form can be represented as an inner-product.

Theorem 8.10. *Let H_1 and H_2 be Hilbert spaces and*

$$h : H_1 \times H_2 \to K$$

be a bounded sesquilinear form. Then h has a representation

$$h(x,y) = \langle Sx, y \rangle$$

where $S : H_1 \to H_2$ is a bounded linear operator. S is uniquely determined by h and $\|h\| = \|S\|$.

Proof. Consider the map $y \mapsto \overline{h(x,y)}$ with $x \in H_1$ fixed. This is a bounded linear functional on H_2 (why?) so by Riesz representation theorem we can represent it

$$\overline{h(x,y)} = \langle y, z \rangle$$

for a unique $z \in H_2$ which of course depends on $x \in H_1$. This gives rise to a map $S : H_1 \ni x \mapsto z \in H_2$. So we have indeed

$$h(x,y) = \langle Sx, y \rangle$$

and what remains to check is that S is linear and bounded. The linearity follows from establishing that $\langle S(\alpha x_1 + \beta x_2), y \rangle = \langle \alpha S x_1 + \beta S x_2, y \rangle$ holds for all $y \in H_2$ (exercise) and the boundedness from the following two estimates:

$$\|h\| = \sup_{x \neq 0, y \neq 0} \frac{|\langle Sx, y \rangle|}{\|x\|\|y\|} \geq \sup_{x \neq 0} \frac{|\langle Sx, Sx \rangle}{\|x\|\|Sx\|} = \|S\| \tag{53}$$

$$\|h\| = \sup_{x \neq 0, y \neq 0} \frac{|\langle Sx, y \rangle|}{\|x\|\|y\|} \leq \sup_{x \neq 0} \frac{\|Sx\|\|y\|}{\|x\|\|y\|} = \|S\| \tag{54}$$

The uniqueness of S follows from the fact that $\langle Sx, y \rangle = \langle Tx, y \rangle$ for all x and y implies $Sx = Tx$ for all x and hence $S = T$. □

The previous theorem will give us existence of the Hilbert adjoint:

Definition 8.5. *Let H_1 and H_2 be Hilbert spaces and $T : H_1 \to H_2$ a bounded linear operator. The Hilbert adjoint of T is the operator $T^\star : H_2 \to H_1$ such that for all $x \in H_1$ and $y \in H_2$ we have*

$$\langle Tx, y \rangle = \langle x, T^\star y \rangle \tag{55}$$

Theorem 8.11. *The Hilbert adjoint operator $T^\star$ of T exists, is unique and satisfies $\|T^\star\| = \|T\|$.*

Proof. Define $h(y,x) = \langle y, Tx \rangle_{H_2}$. This defines a sesqui-linear form on $H_2 \times H_1$ as you can readily check (note T is linear and the inner-product is sesqui-linear in the second component!). This form is also bounded, as seen via

$$|h(y,x)| \leq \|y\|\|Tx\| \leq \|T\|\|x\|\|y\|$$

which gives $\|h\| \leq \|T\|$ and

$$\|h\| = \sup_{x \neq 0, y \neq 0} \frac{|\langle y, Tx \rangle|}{\|y\|\|x\|} \geq \sup_{x \neq 0} \frac{\|Tx\|^2}{\|Tx\|\|x\|} = \|T\|$$

which shows $\|h\| \geq \|T\|$ and hence $\|h\| = \|T\|$. Now by Theorem 8.10 we have a Riesz representation of h:

$$h(y, x) = \langle T^\star y, x\rangle \tag{56}$$

for some $T^\star : H_2 \to H_1$ which is unique and satisfies $\|T^\star\| = \|h\| = \|T\|$. Conjugating (56) yields (55). $\square$

How is the Hilbert adjoint related to the adjoint defined previously? Suppose we are given Hilbert spaces H_1 and H_2. Then the "old" adjoint $T^\times : H_2' \to H_1'$ was defined via

$$f = T^\star g \quad \text{where} \qquad f(x) = T^\times g(x) = g(Tx)$$

thereby mapping a functional $g : H_2 \to \mathbb{C}$ to a functional $f : H_1 \to \mathbb{C}$. In the Hilbert space setting that we are in, we know that both these functionals have Riesz representations, i.e. we have bijective conjugate linear (why? – check!) isometries $A_1 : H_1' \to H_1$ and $A_2 : H_2' \to H_2$ which map $f \mapsto x_0$ and $g \mapsto y_0$ such that

$$f(x) = \langle x, x_0\rangle \quad \text{for all } x \in H_1 \quad \text{and} \quad g(y) = \langle y, y_0\rangle \quad \text{for all } y \in H_2$$

We claim that the composition

$$T^\star = A_1 T^\times A_2^{-1} \qquad T^\star y_0 = x_0$$

defines the relation between $T^\star$ and $T^\times$ (draw a diagram to illustrate this). This map is linear (why?) and we only need to check that it is indeed the Hilbert adjoint, i.e. that the $T^\star$ thus defined satisfies the relation (55). But this is immediate from

$$\langle Tx, y_0\rangle = g(Tx) = f(x) = \langle x, x_0\rangle = \langle x, T^\star y_0 \,.\rangle \tag{57}$$

Q: What is the difference between $T^\star$ and $T^\times$ for matrices (finite dim. linear maps)?

The Hilbert adjoint has the following nice properties the checking of which is an exercise:

Theorem 8.12. *Let H_1 and H_2 be Hilbert spaces and $S, T : H_1 \to H_2$ be bounded linear operators, $\alpha \in \mathbb{C}$ a scalar.*

1. $\langle T^\star y, x\rangle = \langle y, Tx\rangle$
2. $(S + T)^\star = S^\star + T^\star$
3. $(\alpha T)^\star = \overline{\alpha} T^\star$ *(cf. with the adjoint: $(\alpha T)^\times = \alpha T^\times$)*
4. $(T^\star)^\star = T$
5. $\|T^\star T\| = \|TT^\star\| = \|T\|^2$ *hence $T^\star T = 0$ iff $T = 0$.*
6. $(ST)^\star = T^\star S^\star$ *(domains?)*

8.7 Self-adjoint and unitary operators

The definition of the Hilbert adjoint gives rise to the following special operators on a Hilbert space, which you have studied already in the finite dimensional context in linear algebra.

Definition 8.6. *A bounded linear operator $T : H \to H$ on a Hilbert space H is said to be*

- *self-adjoint/ Hermitean if $T^\star = T$*
- *unitary if T is bijective and $T^\star = T^{-1}$*
- *normal if $TT^\star = T^\star T$.*

Note that self-adjoint or unitary both imply normal but not conversely.

Exercise 8.7. *Investigate what the above definitions imply for matrices representing linear transformations from $\mathbb{C}^n$ to $\mathbb{C}^n$ (and $\mathbb{R}^n$ to $\mathbb{R}^n$).*

Here is a nice criterion for self-adjoint:

Theorem 8.13. *Let $T : H \to H$ be a bounded linear operator on a Hilbert space H. Then*

1. *If T is self-adjoint, then $\langle Tx, x\rangle$ is real for all $x \in H$.*
2. *Conversely, if H is complex and $\langle Tx, x\rangle$ is real for all $x \in H$, then T is self-adjoint.*

Proof. For the first, note $\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle$. For the second, note that $\overline{\langle Tx, x\rangle} = \langle Tx, x\rangle$ implies $\langle x, Tx\rangle = \langle Tx, x\rangle$ and therefore $\langle x, Tx\rangle = \langle x, T^\star x\rangle$. Now apply Exercise 3 of Section 11.7. □

A straightforward computation yields that a selfadjoint bounded linear operator $T : H \to H$ satisfies the following identity:

$$\begin{aligned} 4\langle Tx, y\rangle =& \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \\ &+ i\langle T(x+iy), x+iy\rangle - i\langle T(x-iy), x-iy\rangle \end{aligned} \tag{58}$$

This yields an alternative characterization of the operator norm of a self-adjoint operator:

Theorem 8.14. *For $T : H \to H$ a self-adjoint bounded linear operator we have*

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|$$

Proof. Let $\alpha = \sup_{\|x\|=1} |\langle Tx, x\rangle|$. Clearly $|\langle Tx, x\rangle| \leq \|T\|$ for $\|x\| = 1$ by Cauchy-Schwarz and hence $\alpha \leq \|T\|$. To see the reverse, let $u = \frac{Tz}{\lambda}$ for $\lambda > 0$ real to be determined. Then we have

$$\|Tz\|^2 = \langle T(\lambda z), u\rangle = \frac{1}{4}\left[\langle T(\lambda z + u), \lambda z + u\rangle - \langle T(\lambda z - u), \lambda z - u\rangle\right]$$

where we have used (58) (Why do only two terms appear?). Now using the definition of α we can estimate the right hand side as

$$\begin{aligned} \|Tz\|^2 &\leq \frac{1}{4}\alpha\left(\|\lambda z + u\|^2 + \|\lambda z - u\|^2\right) = \frac{1}{2}\alpha\left(\|\lambda z\|^2 + \|u\|^2\right) \\ &= \frac{1}{2}\alpha\left[\lambda^2\|z\|^2 + \lambda^{-2}\|Tz\|^2\right] \end{aligned} \tag{59}$$

Now for $z \in H$ such that $Tz \neq 0$ we can choose $\lambda^2 = \frac{\|Tz\|}{\|z\|}$ to obtain

$$\|Tz\|^2 \leq \alpha\|z\|\|Tz\| \qquad \text{or} \qquad \|Tz\| \leq \alpha\|z\|$$

which (holds trivially also for $Tz = 0$ and hence) is the statement that $\|T\| \leq \alpha$. □

The next Lemma establishes completeness of self-adjoint operators with respect to the operator norm:

Theorem 8.15. *Let (T_n) be a sequence of bounded self-adjoint operators $T_n : H \to H$ on a Hilbert space H. If $T_n \to T$ with respect to the operator norm on $B(H,H)$, i.e. $\|T_n - T\| \to 0$, then the limit T is again a bounded self-adjoint linear operator.*

Proof. The fact that the limit is a bounded linear operator is obvious from the convergence in $B(H,H)$, so we only need to show $T^\star = T$. Now

$$\|T - T^\star\| \leq \|T - T_n\| + \|T_n - T_n^\star\| + \|T_n^\star - T^\star\| \leq 2\|T_n - T\|$$

with the first inequality following from the triangle inequality and the second from $\|T_n - T_n^\star\| = 0$ and the property $\|T_n^\star - T^\star\| = \|T_n - T\|$ of the Hilbert adjoint. □

Let us collect a few useful properties of *unitary* operators. (By the way: Can you have self-adjoint and unitary without being the identity?) You will easily verify for yourself that if $U : H \to H$ is unitary then

1. U is isometric, i.e. $\|Ux\| = \|x\|$ for all $x \in H$
2. $\|U\| = 1$ (provided $H \neq \{0\}$)
3. $U^{-1} = U^\star$ is unitary

Moreover, the product of two unitary operators is again unitary. We also have

Lemma 8.4. *A bounded linear operator T on a Hilbert space H is unitary if and only if it is isometric and surjective.*

Proof. One direction is immediate from the above properties, so let's assume isometric and surjective and show unitary. Isometric implies injective, so in particular the inverse of T exists. By isometry we have for all $x, y \in H$

$$\langle T^\star T x, y\rangle = \langle Tx, Ty\rangle = \langle x, y\rangle$$

hence $\langle (T^\star T - id)x, y\rangle = 0$ for all $x, y \in H$ which implies $T^\star T = id$. Hence $T^\star = T^{-1}$ as desired. □

8.8 Projection Operators

Another special class of operators in a Hilbert space are (orthogonal) projections. Recall that we showed that if M is a closed linear subspace, then H admits a unique decomposition as a direct sum

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

that is to day any $x \in H$ can be written uniquely as the sum $x = y + z$ of an element $y \in M$ and an element $z \in M^{\perp}$ in the orthogonal complement of M. Therefore, given a closed subspace $M \subset H$ we can unambiguously define a map P_M mapping $P : x \to y$, which is called the (orthogonal) projection on M and denoted P_M.

Theorem 8.16. *An (orthogonal) projection is a self-adjoint linear operator satisfying $P^2 = P$ and $\|P\| = 1$ (unless $P = 0$).*

Proof. The linearity of P is left as an exercise. To see $P^2 = P$ note that for an arbitrary $x \in H$ we have (in the notation above) $P^2x = P(Px) = Py = y = Px$ for all $x \in H$. The Pythagorean theorem gives $\|x\|^2 = \|y\|^2 + \|z\|^2$ and hence $\|y = Px\| \leq \|x\|$, which shows $\|P\| \leq 1$. But if $P \neq 0$ then M is non-trivial and there is an $0 \neq y \in M \subset H$ with $y = Py$ and hence $\|P\| = 1$. Finally, the self-adjointness follows for arbitrary $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$ via

$$\langle Px_1, x_2\rangle = \langle y_1, x_2\rangle = \langle y_1, y_2\rangle = \langle x_1, y_2\rangle = \langle x_1, P_2x_2\rangle$$

□

Conversely, we have

Theorem 8.17. *A bounded self-adjoint operator $P : H \to H$ with $P^2 = P$ is an (orthogonal) projection.*

Proof. Let $M = P(H)$. M is a closed linear subspace. Indeed, given $z \in \overline{M}$ we have for a sequence $y_n \in M$ with $y_n \to z$ that $y_n = Px_n = P^2x_n = Py_n$. Taking $n \to \infty$ and using that P is continuous we obtain $z = Pz$ and hence $z \in M$. Next, since P is self-adjoint and $P^2 = P$, we have

$$\langle x - Px, Py\rangle = \langle Px - P^2x, y\rangle = 0 \quad \text{for all } y \in H$$

which means that $x - Px$ is in $M^{\perp}$. Therefore

$$x = Px + (x - Px)$$

is the unique decomposition of x as a sum of elements in M and $M^{\perp}$ and P is by definition the projection on M. □

Let us collect a few facts about projections.

Definition 8.7. *Two projections P_M and P_N are called orthogonal (to each other) if $P_M P_N = 0$.*

Note that $P_M P_N = 0$ if and only if $P_N P_M = 0$ as follows from $(P_M P_N)^{\star} = P_N P_M$. We have

1. Two projections P_M and P_N are orthogonal if and only if $M \perp N$.
2. The sum of two projections P_M and P_N is a projection if and only if $P_M P_N = 0$. In this case $P_M + P_N = P_{M \oplus N}$.
3. The product of two projections P_M and P_N is a projection if and only if they commute, i.e. if $P_M P_N = P_N P_M$. In this case $P_M P_N = P_{M \cap N}$.
4. If P is a projection, then $I - P$ is a projection.

The proof of these facts is again an exercise. You may also wish to illustrate the identities with familiar examples from finite dimensions.

9 Some Spectral Theory: Compact Operators

9.1 Some motivation

One of the goals that was outlined in the motivation was to link functional analysis to the theory of ODEs and PDEs. Consider the following Sturm-Liouville eigenvalue problem[14]

$$L\phi = \mu\phi \quad \text{for the operator } \ L = \frac{d^2}{dx^2} - q(x) \ , \tag{60}$$

which comes together with suitable boundary boundary conditions on the solution ψ, say $\phi(a) = \phi(b) = 0$. The task is to determine those μ's for which one can solve $L\phi = \mu\phi$ with the given boundary conditions, i.e. in other words to determine the eigenvalues ("spectrum") of L. Note how this is an infinite dimensional version of the familiar finite-dimensional eigenvalue problem from linear algebra!

The most fruitful way to study the above problem is to think about it in terms of inner-products and Hilbert spaces. Consider the space of smooth functions on $[a,b]$ vanishing at both endpoints, and equip the space with the inner-product[15]

$$\langle f, g\rangle = \int_a^b dx f(x) \overline{g(x)} \, . \tag{61}$$

The operator L above indeed acts on this space and is (formally) self-adjoint, i.e. $\langle f, Lg\rangle = \langle Lf, g\rangle$. (Check this! Can you see for which type of operators L this works?). If we can develop a spectral theory for self-adjoint operators we have a good chance of successfully resolving the eigenvalue problem. (Note that I used the word "formally" above because as we know L is not even bounded and we do not know how to define the adjoint is this case! Moreover, there are subtleties with the domain and ranges of the operators involved.)

A key element to understand the above problem will be to understand so-called compact operators. These operators form a special class of bounded-linear operators and are somewhat the closest thing to the more familiar finite

[14] A PDE problem would be to understand $\Delta u = \mu u$ for the Laplacian in a bounded connected region $\Omega \subset \mathbb{R}^n$ with suitable boundary conditions for u on $\partial\Omega$. We will consider that problem in Section **??**.

[15] Recall that this inner-product space is incomplete but that we can complete it with respect to the norm induced by the inner-product (61), thereby obtaining $L^2[a,b]$.

dimensional matrices (cf. Section 9.2 in conjunction with Exercise 6 of Week 11.8, as well as Proposition 9.2). As an important result, we will be able to generalize the spectral theorem (diagonalization of symmetric matrices) from linear algebra to compact selfadjoint operators, Theorem 9.5) below. With this at hand we shall be able to revisit our unbounded operator L above and say something about its spectrum, see Exercise 5 of Section 11.9.

9.2 Infinite diagonal matrices

Before we study compact operators, let us look at some simpler operators $T : H \to H$. We call a linear transformation diagonalized if with respect to some orthonormal basis $(\varphi_k)_{k=1}^{\infty}$ we have for all k

$$T\varphi_k = \lambda_k \varphi_k \quad \text{with } \lambda_k \in \mathbb{C}$$

In other words the φ_k are eigenvectors with eigenvalue λ_k. Therefore, if we have

$$f = \sum_{k=1}^{\infty} a_k \varphi_k \quad \text{then} \quad Tf = \sum_{k=1}^{\infty} a_k \lambda_k \phi_k \, .$$

The sequence $\{\lambda_k\}$ is called the *multiplier sequence* corresponding to T. You can easily check the following properties of diagonalized operators

1. $\|T\| = \sup_k |\lambda_k|$
   To see this, observe that $\|Tf\|^2 = \sum_k |a_k|^2 |\lambda_k|^2 \leq \sup_k |\lambda_k|^2 \|x\|^2$. What about the other direction?
2. $T^\star$ corresponds to the sequence $\overline{\lambda}_k$.
   To see this, observe $\langle \varphi_k, \overline{\lambda}\varphi_n \rangle = \langle \lambda\varphi_k, \varphi_n \rangle = \langle T\varphi_k, \varphi_n \rangle = \langle \varphi_k, T^\star \varphi_n \rangle$.
3. T is unitary if and only if $|\lambda_k| = 1$ for all k.
   Note that T unitary implies isometric, so $|\lambda_k| \|\varphi_k\| = \|\lambda\varphi_k\| = \|T\varphi_k\| = \|\varphi_k\|$ which implies $|\lambda_k| = 1$. Conversely, $|\lambda_k| = 1$ implies that T is isometric, in particular injective. It is simple to write down the inverse explicitly and check unitarity.
4. T is an orthogonal projection if and only if $\lambda_k = 0$ or 1 for all k.
   Orthogonal projection is equivalent to T selfadjoint and $T^2 = T$. The first property implies that all λ_k are real and the second $\lambda_k^2 = \lambda_k$ which only has solutions 0 and 1.

Example 9.1. *As an example, we look at $H = L^2[-\pi, \pi]$ and extend a function $f \in H$ to $\mathbb{R}$ by periodicity, i.e. $f(x + 2\pi) = f(x)$. We know that we have a Fourier expansion of f*

$$f = \sum_{k=-\infty}^{\infty} a_k e^{ikx}$$

in terms of an ONB $\varpi_k = e^{ikx}$ for $k \in \mathbb{Z}$. For fixed $h \in \mathbb{R}$ we let $U_h f(x) = f(x + h)$ to be the operator of spatial translation in physical space by h. At the Fourier level, we have

$$U_h f = \sum_{k=-\infty}^{\infty} a_k e^{ik(x+h)} = \sum_{k=-\infty}^{\infty} a_k e^{ikh} \cdot e^{ikx}$$

so U_h is a diagonalized unitary operator with multiplier sequence $\lambda_k = e^{ikh}$.

9.3 Hilbert Schmidt integral operators

Let $H = L^2(\mathbb{R}^d)$ and $T : H \to H$ be the linear map

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y)\, dy \tag{62}$$

This is an integral operator with integral kernel $K(x,y)$. Clearly the properties of the linear map (boundedness and the even nice property of compactness (see below)) will depend on the assumptions we are willing to make on the integral kernel K (see also Exercise 5 of Section 11.8). Hilbert-Schmidt integral operators are operators (62) for which the integral kernel is square integrable, i.e. for which $K(x,y) \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

Proposition 9.1. *Let T be a Hilbert Schmidt integral operator (i.e. (62) with $K(x,y) \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$). Then*

- *If $f \in L^2(\mathbb{R}^d)$, then for almost every x the function $y \mapsto K(x,y) f(y)$ is integrable.*
- *The operator T is bounded from $L^2(\mathbb{R}^d)$ into itself and*

$$\|T\| \le \|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)}$$

- *The Hilbert adjoint $T^\star$ has kernel $\overline{K(y,x)}$.*

Proof. For the first statement, note that by Fubini for almost every x the function $y \mapsto |K(x,y)|^2$ is integrable. Then apply the Cauchy-Schwarz inequality to $|K(x,y) f(y)|_{L^1(\mathbb{R}^d)} \le \|K(x,y)\|_{L^2(\mathbb{R}^d(y))} \|f\|_{L^2(\mathbb{R}^d(y))}$.
For the second statement, note that

$$|Tf(x)| \le \int_{\mathbb{R}^d} |K(x,y)||f(y)|dy \le \sqrt{\int |K(x,y)|^2 dy} \sqrt{\int |f(y)|^2 dy} \tag{63}$$

Squaring this and integrating in x yields the (square of the) L^2-norm of Tf on the left and (since the last term does not depend on x) the (square of the) $L^2(\mathbb{R}^d \times \mathbb{R}^d)$-norm of K on the right. Hence after taking square roots:

$$\|Tf\| \le \|K\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)} \|f\|_{L^2(\mathbb{R}^d)} .$$

For the final statement note that by Fubini

$$\begin{aligned}
\langle Tf, g \rangle &= \int_{\mathbb{R}^d} dx \left[\int_{\mathbb{R}^d} dy\ K(x,y) f(y) \right] \overline{g(x)} \\
&= \int_{\mathbb{R}^d} dy \left[\int_{\mathbb{R}^d} dx\ K(x,y) \overline{g(x)} \right] f(y) \\
&= \int_{\mathbb{R}^d} dx \left[\int_{\mathbb{R}^d} dy\ K(y,x) \overline{g(y)} \right] f(x) \\
&= \int_{\mathbb{R}^d} dx\ f(x) \overline{\int_{\mathbb{R}^d} dy\ \overline{K(y,x)} g(y)} = \langle f, T^\star g \rangle
\end{aligned} \tag{64}$$

From which you read off the claim. □

9.4 Compact Operators: Definition and basic properties

We finally turn to the main topic of this section, compact operators. The theory can be developed for normed spaces and therefore we'll give the definitions and basic lemmas in that more general context before restricting to the Hilbert space setting later.

Definition 9.1. *Let X, Y be normed spaces. An operator $T : X \to Y$ is called a compact*[16] *linear operator if T is linear and if for every bounded subset M of X, the image $T(M)$ is relatively compact, i.e. the closure $\overline{T(M)}$ is compact.*

Lemma 9.1. *Let X and Y be normed spaces. Then*

- *Every compact linear operator is bounded.*
- *If $\dim X = \infty$ the identity operator $I : X \to X$ (which is bounded) is not compact*

Proof. Note that the unit sphere in X, $U = \{x \in X \mid \|x\| = 1\}$ is bounded so that if T is compact, then the closure of $T(U)$ will be compact, hence in particular bounded in Y. Hence $\sup_{\|x\|=1} \|Tx\| < \infty$ which establishes that T is bounded. For the second claim remember that we showed that the closed unit-ball is always not compact in infinite dimensional normed spaces (cf. Theorem 3.5 above). □

An useful criterion for compactness of an operator is the sequential formulation of the above definition:

Theorem 9.1. *Let X, Y be normed spaces and $T : X \to Y$ be a linear operator. Then T is compact if and only if it maps every bounded sequence $\{x_n\}$ to a sequence $\{Tx_n\}$ which has a convergent subsequence.*

Proof. Let T be compact and $\{x_n\}$ an arbitrary bounded sequence. Then, by definition of T being compact, the set $\overline{\{Tx_n\}}$ is compact. By definition of sequential compactness, the sequence $\{Tx_n\}$ has a convergent subsequence, which proves one direction. For the other, assume you know that T maps every bounded sequence to a sequence which has a convergent subsequence and consider an arbitrary bounded set B. To establish that $T(B)$ is relatively compact, pick an arbitrary sequence y_n in $T(B)$, say $y_n = Tx_n$. Since $\{x_n\}$ is bounded by assumption $\{Tx_n\}$ has a convergent subsequence which is the statement that $T(B)$ is relatively compact. □

Observe that the compact linear operators from X to Y form a subspace of $B(X, Y)$. It is also complete as we will prove below in Theorem 9.3. Before we make a simple observation about the finite dimensional case

Theorem 9.2. *Let X, Y be normed spaces and $T : X \to Y$ be a linear operator. Then*

- *If T is bounded and $\dim T(X) < \infty$, then T is compact (such operators are called "finite rank operators")*
- *If $\dim X < \infty$ the operator T is compact*

[16] the name "completely continuous" is also in use.

Proof. The idea is to use that in the finite dimensional case closed and bounded implies compact. So let $\{x_n\}$ be an arbitrary bounded sequence in X. Since T is bounded and $\|Tx_n\| \le \|T\|\|x_n\|$, the set $\{Tx_n\}$ is also bounded. Since we are in finite dimensions it is also relatively compact (cf. Theorem 3.4) and therefore $\{Tx_n\}$ has a convergent subsequence, which since $\{x_n\}$ was arbitrary implies T is compact. For the second statement recall that linear maps are always bounded in finite dimensions and that the dimension of the image of such a linear map cannot be larger than the dimension of the domain (why?). Then apply the first statement. □

Exercise 9.1. *Let X be a Banach space and $T : X \to X$ be compact. Show that if T^{-1} exists and is bounded, then X has to be finite dimensional.*

Here is the completeness statement for compact operators alluded to above

Theorem 9.3. *Let X be a normed space and Y a Banach space. Consider the sequence $T_n : X \to Y$ of compact linear operators. If T_n is uniformly operator convergent to T, then T is compact.*

Proof. Let $\{x_m\}$ be a bounded sequence in X, say $\|x_m\| \le c$. We would like to extract a subsequence $\{y_m\}$ from $\{x_m\}$ such that Ty_m is Cauchy. We use a Cantor diagonal argument. Since T_1 is compact $\{x_m\}$ has a subsequence $x_m^{(1)}$ such that $T_1 x_m^{(1)}$ is Cauchy. But the subsequence $x_m^{(1)}$ is clearly also bounded and since T_2 is compact we can extract from it a subsequence $x_m^{(2)}$ such that $T_2 x_m^{(2)}$ is Cauchy. Continuing this process, we claim that the diagonal sequence $y_m = x_m^{(m)}$ has the property that $T_n y_m$ is Cauchy for any fixed n (why?). Clearly also $\|y_m\| \le c$ since it is a subsequence of $\{x_n\}$. Therefore, for given $\epsilon > 0$ we first choose N large such that

$$\|T - T_n\| < \frac{\epsilon}{3c} \quad \text{for } n \ge N$$

which is possible by the uniform convergence of the T_n. Then we choose M so large such that

$$\|T_N y_j - T_N y_k\| < \frac{\epsilon}{3} \quad \text{for } j, k \ge M$$

which is possible by the Cauchy property of the sequence $\{y_k\}$ constructed above. Hence for $j, k \ge M$ we have (why?)

$$\|Ty_j - Ty_k\| \le \|Ty_j - T_N y_j\| + \|T_N y_j - T_N y_k\| + \|T_N y_k - Ty_k\| < \epsilon$$

Therefore Ty_m is indeed Cauchy and converges since Y is complete. □

Exercise 9.2. *Can we relax uniform convergence to strong convergence in the above theorem, i.e. is it true that if $T_n \to T$ strongly (i.e. $T_n x \to Tx$ in Y for all $x \in X$) then the limit operator T is compact? Hint: Consider $H = \ell^2$ and the sequence of finite rank operators $T_n x = (\xi_1, \xi_2, ..., \xi_n, 0, 0...)$ which cuts off a given $x = (\xi_1, \xi_2, ...) \in \ell^2$ after the n^{th} component.*

Exercise 9.3. *Prove that the operator $T : \ell^2 \to \ell^2$ mapping a sequence $x \in \ell^2$ with components ξ_j to a sequence $y \in \ell^2$ with components $\eta_j = \frac{\xi_j}{j}$ is compact. Hint: Construct a sequence of finite rank operators and show convergence to T in the operator norm.*

Exercise 9.4. *Let X be a normed space and $S, T \in B(X, X)$. If T is compact then ST and TS are compact.*

Here is a very useful property of compact operators, namely that they map weakly convergent sequences into strongly convergent ones:

Theorem 9.4. *Let X, Y be normed spaces and $T : X \to Y$ be a compact linear operator. Suppose $x_n \rightharpoonup x$ weakly in X. Then Tx_n converges strongly in Y with limit $y = Tx$.*

Proof. Since T is bounded $Tx_n \rightharpoonup Tx$ (to see this, note that for an arbitrary functional $f : Y \to K$, the composition $f \circ T$ is a bounded linear functional on X, hence $f(Tx_n) \to f(Tx)$ which since f was arbitrary is the statement $Tx_n \rightharpoonup Tx$).
Suppose now that Tx_n did not converge strongly to $y := Tx$. Then there exists an $\epsilon_0 > 0$ such that for a subsequence x_{n_k} we have

$$\|Tx_{n_k} - y\| \geq \epsilon_0 \, . \tag{65}$$

Since weakly convergent subsequences are bounded (cf. Lemma 6.1), x_n and in particular x_{n_k} are bounded. By compactness of T, there is a subsequence $x_{n_{k_i}}$ such that $\{Tx_{n_{k_i}}\}$ converges to some $\tilde{y}$, $\{y_i\} := \{Tx_{n_{k_i}}\} \to \tilde{y}$. But then in particular $Tx_{n_{k_i}} \rightharpoonup \tilde{y}$ weakly and by the uniqueness of the weak limit (Lemma 6.1) we must have $\tilde{y} = y$. But $\|y_i - y\| \to 0$ is in contradiction with (65). □

The next Proposition is stated for Hilbert spaces. The second statement of it (but not the first!) remains true for X a normed space, Y a Banach space and $T \in B(X, Y)$, although the proof is a little harder. (In this case, of course the Hilbert adjoint $T^\star$ is replaced by the adjoint $T^\times$.)

Proposition 9.2. *Suppose $T : H \to H$ is a bounded linear operator on a Hilbert space H. Then*

1. *If T is compact, there is a sequence of operators of finite rank such that $\|T_n - T\| \to 0$.*
2. *T is compact if and only if $T^\star$ is compact.*

Proof. Let $\{e_k\}$ be an orthonormal basis of H and Q_n the orthogonal projection on the subspace spanned by the e_k with $k > n$, i.e.

$$Q_n(g) = \sum_{k>n} \alpha_k e_k \qquad \text{for} \quad g = \sum_{k=1}^{\infty} \alpha_k e_k \, .$$

Clearly $\|Q_n(g)\|^2$ is a decreasing sequence tending to zero as $n \to \infty$ for any g (why?). We claim that $\|Q_n T\| \to 0$ as $n \to 0$. Once this is established, the first claim will follow, because if P_n is the projection to the space spanned by $e_1, ..., e_n$, then $I = P_n + Q_n$ and $\|Q_n T\| \to 0$ means $\|T - P_n T\| \to 0$, which since $P_n T$ is manifestly finite rank implies the claim. To see that $\|Q_n T\| \to 0$, suppose not. Then there exists an $\epsilon_0 > 0$ and a subsequence such that $\|Q_{n_k} T\| \geq \epsilon_0$. Therefore, we can find a sequence $\{f_{n_k}\}$ with $\|f_{n_k}\| = 1$ such that $\|Q_{n_k} T f_{n_k}\| \geq \frac{\epsilon_0}{2}$. Since T is compact, there exists a subsequence of f_{n_k} (again denoted f_{n_k} to keep notation clean) such that $\{Tf_{n_k}\}$ converges to some $g \in H$. But then the second term in

$$Q_{n_k}(g) = Q_{n_k}Tf_{n_k} + Q_{n_k}(g - Tf_{n_k})$$

goes to zero and hence for all sufficiently large k we have

$$\|Q_{n_k}(g)\| \geq \frac{\epsilon_0}{4}$$

which is in contradiction with $\|Q_{n_k}(g)\| \to 0$ for any g observed at the beginning of the proof.

To see the second claim, note that T compact implies $\|P_nT - T\| \to 0$. Since the adjoint preserves the norm and projections are self-adjoint, we have $\|P_n - T\| = \|T^\star P_n - T^\star\| \to 0$. Hence $T^\star P_n$ is a sequence of finite rank operators with limit $T^\star$ and the latter is compact by Theorem 9.3. Running the argument backwards shows the other direction. □

9.5 The spectral theorem for compact self-adjoint operators

Now that we have some intuition for compact operators we can prove the spectral theorem. In the following, H is a separable infinite-dimensional Hilbert space. In the case of finite dimensional Hilbert spaces, the theorem reduces to the familiar theorem in linear algebra about diagonalizing symmetric matrices.

Theorem 9.5. *Suppose T is a compact self-adjoint operator $T : H \to H$ for H a (separable) infinite dimensional Hilbert space. Then there exists an orthonormal basis $\{\varphi_k\}_{k=1}^\infty$ of H consisting of eigenvectors of T:*

$$T\varphi_k = \lambda_k\varphi_k .$$

Moreover, $\lambda_k \in \mathbb{R}$ and $\lambda_k \to 0$ as $k \to \infty$.

Note that the converse follows from Exercise 6. The theorem will be proven via three Lemmas, the first of which is quite simple:

Lemma 9.2. *Let $T : H \to H$ be bounded and selfadjoint. Then*

- *eigenvalues are real: If $T\varphi = \lambda\varphi$ for some $\varphi \neq 0$, $\lambda \in \mathbb{C}$, then $\lambda \in \mathbb{R}$.*
- *orthogonality of eigenvectors belonging to distinct eigenvalues: If $\lambda_1 \neq \lambda_2$ are eigenvalues, then corresponding eigenvectors φ_1, φ_2 are orthogonal*

Proof. For the first part, observe

$$\lambda\langle\varphi,\varphi\rangle = \langle\lambda\varphi,\varphi\rangle = \langle T\varphi,\varphi\rangle = \langle\varphi,T\varphi\rangle = \langle\varphi,\lambda\varphi\rangle = \overline{\lambda}\langle\varphi,\varphi\rangle$$

which in view of $\varphi \neq 0$ implies $\lambda = \overline{\lambda}$ and hence that $\lambda \in \mathbb{R}$. For the second part, let $T\varphi_1 = \lambda_1\varphi_1$ and $T\varphi_2 = \lambda_1\varphi_2$ for some non-zero φ_1, φ_2. Then

$$\lambda_1\langle\varphi_1,\varphi_2\rangle = \langle\lambda_1\varphi_1,\varphi_2\rangle = \langle T\varphi_1,\varphi_2\rangle = \langle\varphi_1,T\varphi_2\rangle = \langle\varphi_1,\lambda_2\varphi_2\rangle = \lambda_2\langle\varphi_1,\varphi_2\rangle$$

since the λ_i are real. This identity can only hold if $\lambda_1 = \lambda_2$ or $\langle\varphi_1,\varphi_2\rangle = 0$. □

The second Lemma is crucial:

Lemma 9.3. *Suppose T is compact and $\lambda \neq 0$. Then*

- *the dimension of the null-space $N(T - \lambda \cdot id)$ is finite ((i.e. in particular any eigenspace belonging to an eigenvalue λ is finite-dimensional).*
- *the eigenvalues of T form an at most countable set $\lambda_1, ...\lambda_k, ...$ with $\lambda_k \to 0$ as $k \to \infty$.*
- *for each $\mu > 0$ the space spanned by the eigenvectors corresponding to all λ_k with $|\lambda_k| > \mu$ is finite dimensional*

Note that the third statement implies the second.

Proof. Let V_λ denote the null-space of $T - \lambda \cdot id$. Suppose V_λ was not finite-dimensional. Then there exists a sequence $\{\varphi_k\}_{k=1}^\infty$ of orthonormal vectors (Gram-Schmidt!) in V_λ. Since T is compact we can extract a subsequence such that $T\varphi_{n_k}$ converges to some $y \in H$. But since $T\varphi_{n_k} = \lambda\varphi_{n_k}$ and $\lambda \neq 0$ is fixed it means that φ_{n_k} itself converges. This is a contradiction, since $\|\varphi_{n_k} - \varphi_{n_l}\| = \sqrt{2}$ for $k \neq l$.

To prove the third statement, suppose there are infinitely many distinct eigenvalues λ with $|\lambda| \geq \mu$. Then we can take an (orthonormal[17]) sequence $\{\varphi_k\}$ belonging to a sequence λ_k of distinct eigenvalues. We have

$$T\varphi_k = \lambda_k \varphi_k$$

and since T is compact, going to a subsequence we can achieve that $T\varphi_{n_k}$ converges in H. For that subsequence $T\varphi_{n_k} = \lambda_{n_k}\varphi_{n_k}$ and hence the right hand side must also converge. However,

$$\|\lambda_{n_k}\varphi_{n_k} - \lambda_{n_l}\varphi_{n_l}\|^2 = |\lambda_{n_k}|^2 + |\lambda_{n_l}|^2 \geq 2\mu^2$$

yields a contradiction. □

Lemma 9.4. *Suppose $T \neq 0$ is compact and self-adjoint. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. Recall that in Theorem 8.14 we derived the following expression for the norm of a self-adjoint operator: $\|T\| = \sup_{\|x\|=1} |\langle x, Tx\rangle|$. Hence we either have

$$\|T\| = \sup_{\|x\|=1} \langle x, Tx\rangle \quad \text{or} \quad \|T\| = -\inf_{\|x\|=1} \langle x, Tx\rangle .$$

Suppose we are in the first case (the second will be proven entirely analogously and is left to you). We need to prove that $\lambda := \|T\|$ is an eigenvalue. By the definition of the sup we pick a sequence $\{x_n\}$ in H with $\|x_n\| = 1$ and $\langle x, Tx\rangle \to \lambda$. Since T is compact, we can extract a subsequence x_{n_k} such that Tx_{n_k} converges so some $g \in H$. We claim that this g is an eigenvector with eigenvalue λ. To prove this it suffices to show $\|Tx_{n_k} - \lambda x_{n_k}\|^2 \to 0$. Indeed if the latter was true, then since Tx_{n_k} converges to g also λx_{n_k} has to converge to g, which means that $\lambda g = \lambda \lim_{k\to\infty} T(x_{n_k}) = \lim_{k\to\infty} T(\lambda x_{n_k}) = Tg$ by continuity of T. Therefore, we have $Tg = \lambda g$ and moreover $g \neq 0$ since $g = 0$

[17]Note that by the previous Lemma the $\{\varphi_k\}$ are are indeed orthogonal to each other!

would imply that $\|Tx_{n_k}\| \to 0$ and hence that $\langle x_{n_k}, x_{n_k}\rangle \to 0$ contradicting $\|x_{n_k}\| = 1$. It remains to show $\|Tx_{n_k} - \lambda x_{n_k}\|^2 \to 0$. This follows from

$$\begin{aligned}\|Tx_{n_k} - \lambda x_{n_k}\|^2 &= \|Tx_{n_k}\|^2 - 2\lambda\langle Tx_{n_k}, x_{n_k}\rangle + \lambda^2\|x_{n_k}\|^2 \\ &\leq \|T\|^2\|x_{n_k}\|^2 + \lambda^2\|x_{n_k}\|^2 - 2\lambda\langle Tx_{n_k}, x_{n_k}\rangle \\ &\leq 2\lambda\left(\lambda - \langle Tx_{n_k}, x_{n_k}\rangle\right) \end{aligned} \tag{66}$$

and the right hand side going to zero as $k \to \infty$. □

Proof of Theorem 9.5. For the zero operator there is nothing to show so let us assume $\|T\| \neq 0$. Let S be the closure of the vector space of all eigenvectors of T. By Lemma 9.4 this space is non-empty. We need to prove that $S = H$. Suppose not. Then $S^\perp$ in the decomposition

$$H = S \oplus S^\perp$$

satisfies $S^\perp \neq \{0\}$. We will show that $S^\perp$ contains an eigenvector of T which gives a contradiction since $S \cap S^\perp = \{0\}$. To see this, note that $y \in S$ implies that $Ty \in S$ and $z \in S^\perp$ implies $Tz \in S^\perp$ (to see the latter, note $\langle Tz, y\rangle = \langle z, Ty\rangle$ for $z \in S^\perp$ and arbitrary $y \in S$). We can hence consider the operator $T_1 := T\big|_{S^\perp}$, the restriction to $S^\perp$. Now $S^\perp$ is a closed subspace of H and restricting the inner-product to $S^\perp$ makes $S^\perp$ a Hilbert space on its own. Moreover, the map $T_1 : S^\perp \to S^\perp$ is still compact and self-adjoint (why?). Hence Lemma 9.4 applies (if $\|T_1\| = 0$ then since $S^\perp$ is non-trivial, there is a non-zero eigenvector with eigenvalue zero) and we conclude that T_1 has a non-zero eigenvector in $S^\perp$, $T_1 z = \lambda z$ for $\lambda \in \mathbb{R}$ and some $z \neq 0$. But clearly z is also an eigenvalue of T itself which yields the desired contradiction. □

A few remarks are in order:

- If you drop self-adjoint or compact, T may have no eigenvectors (see Exercise 2 of Section 11.9). However, for T self-adjoint and bounded, there is a generalization of the spectral theorem, which you can look up in the literature. It involves much heavier machinery than what we used above.
- Just as for matrices, there are various generalizations of the spectral theorem (for instance to normal operators). See Exercise 1 of Section 11.9.

9.6 The Fredholm Alternative

We next prove a powerful theorem in PDEs which concerns "compact perturbations of the identity". We will not be able to fully appreciate it until we solve the Dirichlet problem in the next section and I will restrain myself from giving you extended motivation at this point. Let us just say that many linear PDEs can be phrased in terms of equations of the form

$$(id - K)\, u = f \tag{67}$$

where $K : H \to H$ is compact, the right hand side $f \in H$ is given and we want to solve for $u \in H$. The problem then is to determine for what f a unique solution of this problem exists and – in case that it doesn't exist or is non-unique – to understand the obstructions.

Theorem 9.6. *Let $K : H \to H$ be a compact linear operator for H an infinite-dimensional separable Hilbert space. Then*

- $N(id - K)$ *is finite dimensional*
- $R(id - K)$ *is closed*
- $R(id - K) = N(I - K^\star)^\perp$
- $N(id - K) = \{0\}$ *if and only if* $R(id - K) = H$.
- $\dim N(id - K) = \dim N(id - K^\star)$.

Going back to equation (67) we see that the theorem is quite helpful. It guarantees that the homogeneous equation ($f = 0$) can only have a finite dimensional solution space. Moreover, by 4., if the homogeneous equation has only the trivial solution then we can solve (67) uniquely for any given $f \in H$. Conversely, if the kernel is non-trivial (i.e. the homogeneous equation has non-trivial solutions), then obviously the solution to (67) is non-unique and by 3. we can (and will) only find A solution if f is perpendicular to the kernel of the operator $I - K^\star$. More concisely, the *Fredholm alternative* states that either

- for each $f \in H$ (67) has a unique solution OR ELSE
- the homogeneous equation $u - Ku = 0$ has non-trivial solutions $u \neq 0$. In this case, (67) can be solved if and only if $f \in N(I - K^\star)^\perp$.

Remark 9.1. *The Fredholm-Alternative also holds for compact operators in Banach spaces. See [Friedman]. You are free to use this fact in the exercises.*

Proof of Theorem 9.6. For 1. assume for contradiction that $N(id - K) = \infty$. Then we can find an orthonormal sequence $\{u_k\}_{k=1}^\infty$ in $N(id - K)$. By compactness of K, we have $Ku_{n_k} = u_{n_k}$ with the left hand side (and hence the right hand side) converging in H for a subsequence u_{n_k}. But $\|u_{n_k} - u_{n_l}\| = \sqrt{2}$ and $\{u_{n_k}\}$ cannot converge.

For the second statement we first prove that there exists a constant $\gamma > 0$ such that

$$\gamma \|u\| \leq \|u - Ku\| \quad \text{for all } u \in N(id - K)^\perp .$$

Suppose not. Then there exists a sequence $\{u_k\}$ in $N(id - K)^\perp$ such that $\|u_k\| = 1$ and $\|u_k - Ku_k\| \to 0$. By Banach-Alaoglu (Theorem 6.4) we can extract from the bounded sequence u_k a weakly convergent subsequence $u_{n_k} \rightharpoonup u$.[18] Using compactness (Theorem 9.4) we have $Ku_{k_i} \to Ku$. Now since

$$\|u_{k_i} - u_{k_j}\| \leq \|u_{k_i} - Ku_{k_i}\| + \|Ku_{k_i} - Ku_{k_j}\| + \|Ku_{k_j} - u_{k_j}\| \tag{68}$$

we conclude $u_{k_i} \to u$ strongly and hence by continuity $Ku = u$. But this means that $u \in N(id - K)$ and hence $\langle u_{k_j}, u\rangle = 0$ for every j. Letting $j \to \infty$ we conclude $\langle u, u\rangle = 0$ and hence $u = 0$ which contradicts $\|u\| = 1$.

[18] Theorem 6.4 was formulated for sequences of bounded linear functionals in the dual of a separable Banach space. Note that in a Hilbert space we can identify the dual space with itself via the Riesz representation theorem leading to a particularly simple formulation of Banach-Alaoglu in this setting, which is applied here.

With (68) at hand, the closure of the range follows immediately: Let $v \in \overline{R(id-K)}$ and pick a sequence $v_n \in R(id-K)$ with $v_n \to v$. Since $v_n \in R(id-K)$ we can find $u_n \in N(id-K)^\perp$ with $u_n - Ku_n = v_n$. Applying (68) yields

$$\|v_m - v_n\| \geq \gamma \|u_m - u_n\|$$

and since the left hand side goes to zero so does the right. By completeness $u_n \to u \in H$ and using continuity $u_n - Ku_n = v_n$ turns into $u - Ku = v$ which shows that $v \in R(id-K)$ and hence $\overline{R(id-K)} = R(id-K)$.

For 3. note first the general identity

$$R(T)^\perp = N(T^\star) \qquad \text{for } T: H \to H \text{ a bounded linear operator},$$

whose proof is an easy exercise. This is the same as $\overline{R(T)}^\perp = N(T^\star)$, cf. Exercise 8.6 and again by that exercise $\overline{R(T)} = N(T^\star)^\perp$. Since we showed in 2. that the range of $T = id - K$ was closed, the statement 3. follows.

Turning to 4., we first assume that $N(I-K) = \{0\}$ and want to conclude that the range is the whole space. Suppose for contradiction that $R(id-K) = (id-K)(H) = H_1 \subsetneq H$, so H_1 is an honest (closed by 2.) subspace of H. Next consider $H_2 = (id-K)(H_1) \subsetneq H_1$. The claim that $H_2 \subsetneq H_1$ follows from the injectivity of $id - K$.[19] Continuing in this fashion, we obtain a sequence of spaces $H_k = (id-K)^k(H)$ with

- H_k is a closed subspace of H
- $H_{k+1} \subsetneq H_k$ for $k = 1, 2, ...$

We choose a sequence $u_k \in H_k$ with $\|u_k\| = 1$ and $u_k \in H_{k+1}^\perp$. Note that for $k > l$ we then have

$$H_{k+1} \subsetneq H_k \subset H_{l+1} \subsetneq H_l \, .$$

Therefore, writing

$$\|Ku_k - Ku_l\| = \|(-u_k + Ku_k) + (u_l - Ku_l) + u_k - u_l\|$$

we see that if $k > l$ the first bracket is in H_{k+1}, the second in H_{l+1} and u_k in H_k. Now u_l being in $H_{l+1}^\perp$ is perpendicular to all of them so that computing the norm yields (the cross-term vanishes by the previous considerations and we ignore the second positive term):

$$\|Ku_k - Ku_l\| \geq 1 \, ,$$

proving that Ku_k cannot have a convergent subsequence which is in contradiction with K being compact. This shows that indeed $R(id-K) = H$ as desired.

To show the converse of 4. assume $R(id-K) = H$. By 3. we know that this implies $N(id-K^\star) = 0$ and by applying the argument of the previous paragraph to $id - K^\star$, we can conclude $R(id-K^\star) = H$. Applying again 3. shows $N(id-K)^\perp = H$ and hence $N(id-K) = \{0\}$ as desired.

[19]Indeed, suppose that $H_2 = H_1$. Then because $(id-K)(H_1) = H_1$ any element $y \in H_1$ has a preimage $x \in H_1$. But $H_1^\perp$ is non trivial by assumption and since $H_1 = (id-K)H$ must also get mapped to H_1. Hence there are elements in H_1 which have multiple preimages (one in H_1, one in $H_1^\perp$) contradicting injectivity.

We finally show 5. Note that we already know that the dimensions are finite. We first show

$$\dim N(id - K) \geq \dim R(id - K)^{\perp} \, (= \dim N(id - K^{\star})) \,. \tag{69}$$

Suppose for contradiction that $\dim N(id - K) < \dim R(id - K)^{\perp}$. Then there exists a bounded linear map

$$A : N(id - K) \to R(id - K)^{\perp} \quad \text{injective but not surjective} \tag{70}$$

We extend A to a linear map on all of H by setting $Au = 0$ for $u \in N(id - K)^{\perp}$. The map A has finite-dimensional range so A is compact, therefore $K + A$ is compact. We claim that $N(id - (K + A)) = \{0\}$. Indeed, if $u \in N(id - (K + A))$, then $u - Ku = Au$ and since the left hand side is in $R(id - K)$ and the right hand side in the orthogonal complement of that, we must have $u - Ku = 0 = Au$. Finally, since A is injective on $N(id - K)$, we conclude $u = 0$. With $N(id - (K + A)) = \{0\}$ established, we can apply 4. of the Theorem to conclude that $R(id - (K + A)) = H$. This is the desired contradiction because if $v \in R(id - K)^{\perp}$ but $v \notin R(A)$ (such a choice is possible because of (70)), then the equation

$$u - Ku - Au = v$$

does not have a solution (because $Au = v$ doesn't have one and $u - Ku \in R(id - K)$). This establishes (69). Clearly the same argument can be used replacing K by $K^{\star}$ and this gives the other direction and hence statement 5 of the theorem. □

10 PDE application: The Dirichlet problem

We have now developed enough functional analytic machinery to actually solve a real problem. The *Dirichlet problem* is a famous problem in classical physics[20] and can be stated as follows. Let $\Omega \subset \mathbb{R}^n$ be open, bounded and connected with smooth boundary $\partial\Omega$ and g be a (smooth, say) function prescribed along $\partial\Omega$. Find a function u satisfying

$$dp \begin{cases} \Delta u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \tag{71}$$

There are many physical situations which reduce to solving a problem of the above form. For instance, you have to solve DP if you are looking for the equilibrium temperature distribution assumed in Ω given that the boundary of Ω is exposed to a fixed temperature distribution g. The above problem is closely related (how?) to the following Dirichlet problem

$$DP \begin{cases} \Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{72}$$

where f is a (smooth, say) function on Ω.

Remarkably, the functional analytic techniques we already have will allow us to **prove the existence of (weak) solutions** to the above problem. This

[20] P. Dirichlet, 1805–1859

is very remarkable because I am sure so far you were only able to find solutions explicitly (using separation of variables) in simple geometries (i.e. Ω highly symmetric like a cylinder or a ball).[21]

More generally, the methods will apply not only to the Laplacian $\Delta = \partial^2_{x_1} + \ldots + \partial^2_{x_n}$ but to any suitable perturbations of the Laplacian (in fact to general, *second order elliptic* operators but let's not go that far at this point).

Finally, we will be able to show that any $f \in L^2(\Omega)$ produces a unique solution $u \in H^1_0(\Omega)$ (a space to be introduced below) of the Dirichlet problem. This means that the Laplacian has an inverse in suitable spaces. We will show that this operator is compact as a map from $L^2(\Omega)$ to $L^2(\Omega)$ and thereby obtain the **spectrum of the Laplacian**. In particular, we will conclude the existence of a discrete set of λ solving the eigenvalue problem

$$DP_\lambda \begin{cases} \Delta u + \lambda u = 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{73}$$

The corresponding eigenfunctions u_λ are related to the fundamental tones/ modes of an elastic membrane (like a drum) fixed at the boundary.

10.1 Weak solutions

We first translate our problem (DP) into the language of Hilbert spaces. We define the space

$$\tilde{C}^1_0\left(\overline{\Omega}\right) = \{\text{functions continuously differentiable on } \overline{\Omega}, \text{ vanishing on } \partial\Omega\} \tag{74}$$

and equip this space with the following inner product (why is this an inner-product?)

$$\langle u, v\rangle = \int_\Omega \sum_k u_{x_k} v_{x_k} dx = \int_\Omega \nabla u \nabla v dx\,. \tag{75}$$

with induced norm

$$\|u\|^2 = \int_\Omega |\nabla u|^2 dx\,. \tag{76}$$

Suppose we already *have* a solution $u \in C^2\left(\overline{\Omega}\right)$ of (DP) with $f \in C^0\left(\overline{\Omega}\right)$. Then it must satisfy

$$\langle u, v\rangle = -\int v\Delta u dx = \int v f dx \quad \text{for all } v \in \tilde{C}^1_0\left(\overline{\Omega}\right) \tag{77}$$

where we have used Stokes' theorem. Hence (77) is a necessary condition for u to be a solution of (DP). Now forget that (77) holds if u is a solution and look at it as an equation for an unknown function u. If you can find such a u you have at least a very good candidate for your solution of DP! More strikingly, we already know how to find such a u. Indeed, for f given, the right hand side of (77) can be interpreted as a linear functional $\phi : v \mapsto \int v f dx$. If this functional

[21]Note that in two dimensions you can use the Riemann mapping theorem to reduce the general Dirichlet problem to the Dirichlet problem on the disc thereby connecting the DP with complex analysis.

is bounded and the inner-product space above was complete, we could apply the Riesz representation theorem and infer the existence of a u satisfying (77).

Let's make this precise. First of all we need to complete the inner-product space $\tilde{C}^1_0\left(\overline{\Omega}\right)$ (why is this space incomplete?) with respect to the norm (76) in order to produce a Hilbert space (the Riesz representation theorem hinges crucially on completeness!). Let us call the resulting space $H^1_0\left(\Omega\right)$.

Definition 10.1. *We call u a weak solution of (DP) if $u \in H^1_0\left(\Omega\right)$ satisfies*

$$\langle u, v\rangle = \int v f dx \qquad \text{for all } v \in H^1_0\left(\Omega\right).$$

Remark 10.1. *Given a weak solution, one would of course like to show that the solution has more regularity, in particular it should at least be twice differentiable for the original equation to make sense. This is the topic of regularity theory, a big topic in PDEs. We will not concern ourselves with this here.*

To show that the functional $\phi : v \mapsto \int f v dx$ is a bounded linear functional it suffices to show it is a bounded linear functional on $\tilde{C}^1_0\left(\overline{\Omega}\right)$ and then argue by density. To do the former, we need to prove the Poincaré inequality for the region Ω:

Exercise 10.1. *Prove that for $v \in \tilde{C}^1_0\left(\overline{\Omega}\right)$ the inequality*

$$\int_\Omega v^2 dx \leq C_\Omega \int_\Omega |\nabla v|^2 dx \tag{78}$$

holds for a constant Ω depending only on the domain Ω.

From the exercise we clearly have

$$|\phi\left(v\right)| \leq \Big|\int v f dx\Big| \leq \sqrt{\int v^2 dx}\sqrt{\int f^2 dx} \leq \sqrt{C_\Omega}\|v\|_{H^1_0(\Omega)}\|f\|_{L^2(\Omega)}$$

which shows that ϕ is a bounded linear functional on $\tilde{C}^1_0\left(\overline{\Omega}\right)$ and by density (see Theorem 4.3) on $H^1_0\left(\Omega\right)$.

Theorem 10.1. *Given $f \in L^2\left(\Omega\right)$ there exists a unique weak solution $u \in H^1_0\left(\Omega\right)$ of (DP). The weak solution satisfies*

$$\int_\Omega |\nabla u|^2 dx \leq C \int_\Omega f^2 dx$$

for a constant C depending only on Ω.

Proof. The existence and uniqueness is immediate from the Riesz representation theorem. To prove the estimate, observe that since we have

$$\int_\Omega \nabla u \nabla v = \int_\Omega v f \ dx$$

for all $v \in H^1_0\left(\Omega\right)$ this holds in particular for $v = u$ so for any $\epsilon > 0$

$$\int_\Omega |\nabla u|^2 dx = \int_\Omega u f \ dx \leq \epsilon \int_\Omega u^2 dx + \frac{1}{\epsilon}\int_\Omega f^2 dx \leq C_\Omega \epsilon \int_\Omega |\nabla u|^2 dx + \frac{1}{\epsilon}\int_\Omega f^2 dx$$

using the Poincaré inequality. Choosing $\epsilon = \frac{1}{2C_\Omega}$ yields the result with $C = 4C_\Omega$. □

In summary, we have obtained a map

$$\Phi : L^2(\Omega) \to H_0^1(\Omega) \tag{79}$$

mapping a right hand side of (DP) to a weak solution. In the next section we will prove that the embedding (cf. again the Poincare inequality!)

$$\iota : H_0^1(\Omega) \hookrightarrow L^2(\Omega)$$

is compact (Rellich's theorem). Therefore the composition of Φ and ι is a compact linear operator from L^2 and L^2 (why?) and we'll have the spectral theorem at our disposal!

Exercise 10.2. *Show that Φ is also self-adjoint from $L^2(\Omega)$ to $L^2(\Omega)$.*

10.2 Rellich's theorem and the spectrum of the Laplacian

Theorem 10.2. *Let $\{u_m\}$ be a sequence of functions in $H_0^1(\Omega)$ such that $\|u_m\|_{H_0^1(\Omega)} \leq c$. Then there exists a subsequence $\{u_{m_k}\}$ which converges (strongly) in $L^2(\Omega)$.*

For the proof we will need two Lemmas.

Lemma 10.1. *Let Q be a cube $\{0 \leq x_i \leq \sigma\}$ in $\mathbb{R}^n$ and u a real-valued function in $C^1(\Omega)$. Then*

$$\|u\|_{L^2(Q)}^2 \leq \frac{1}{\sigma^n}\left(\int_Q u(x)\,dx\right)^2 + \frac{n}{2}\sigma^2\|u\|_{H^1(Q)}^2$$

Remark 10.2. *Compare this with the Poincaré inequality. In particular, constant functions are now allowed and accounted for by the first term.*

Proof.

$$\begin{aligned} u(x_1,...,x_n) - u(y_1,...,y_n) &= \int_{y_1}^{x_1} d\xi_1 \partial_{\xi_1} u(\xi_1, x_2, ..., x_n) \\ &+ \int_{y_2}^{x_2} d\xi_1 \partial_{\xi_2} u(y_1, \xi_2, x_3..., x_n) + \ldots + \\ &+ \int_{y_n}^{x_n} d\xi_n \partial_{\xi_2} u(y_1, y_2, ..., y_{n-1}, \xi_n)\,. \end{aligned} \tag{80}$$

Taking squares[22] and applying Cauchy-Schwarz yields

$$\begin{aligned} u^2(x) + u^2(y) - 2u(x)u(y) &\leq n\sigma \int_0^\sigma d\xi_1 \left[\partial_{\xi_1} u(\xi_1, x_2, ..., x_n)\right]^2 \\ &+ ... + n\sigma \int_0^\sigma d\xi_n \left[\partial_{\xi_n} u(y_1, y_2, ..., y_{n-1}, \xi_n)\right]^2\,. \end{aligned} \tag{81}$$

Now integrate over $x_1, ...x_n$ and $y_1, ..., y_n$ to obtain

$$2\sigma^n \int_Q [u(x)]^2\,dx - 2\left(\int_Q u(x)\,dx\right)^2 \leq n\sigma^{n+2}\sum_{i=1}^n \int_Q \left[\frac{\partial u(x)}{\partial x_i}\right]^2 dx\,. \tag{82}$$

After bringing the second term to the right and dividing by $2\sigma^n$ we obtain the desired inequality. ☐

[22] Verify and use $(A_1 + ... + A_n)^2 \leq n\left(A_1^2 + ... + A_n^2\right)$.

Lemma 10.2 (Friedrich's inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain in $\mathbb{R}^n$ as in (DP). For any $\epsilon > 0$ there exists an integer $M > 0$ and real-valued functions*

$$w_1, ..., w_M \quad \text{in } L^2(\Omega) \text{ with } \|w_j\|_{L^2(\Omega)} = 1$$

such that for any $u \in H^1_0(\Omega)$ we have

$$\|u\|^2_{L^2(\Omega)} \leq \epsilon \|u\|^2_{H^1_0(\Omega)} + \sum_{j=1}^{M} \left[\int_\Omega u \cdot w_j \right]^2 . \tag{83}$$

Note that the last term is precisely the L^2 inner-product of u and w_j.

Proof. It suffices to prove the Lemma for $u \in \tilde{C}^1_0\left(\overline{\Omega}\right)$ and then use the completion. Extend u outside Ω by zero. Put Ω in a cube of size σ_0^n with edges parallel to the coordinate axes. Divide the big cube into $M = \frac{\sigma_0^n}{\sigma^n}$ smaller cubes (choose σ so that M is an integer), which we denote $Q_1, ...Q_M$. Applying the previous Lemma in each cube we have

$$\|u\|^2_{L^2(Q_i)} \leq \frac{1}{\sigma^n} \left(\int_{Q_i} u(x)\, dx \right)^2 + \frac{n}{2} \sigma^2 \|u\|^2_{H^1(Q_i)}$$

Summing this over all cubes we find

$$\|u\|^2_{L^2(\Omega)} \leq \sum_{j=1}^{M} \left(\int_{Q_i} u \cdot w_j dx \right)^2 + \frac{n}{2} \sigma^2 \|u\|^2_{H^1(\Omega)}$$

where

$$w_j = \begin{cases} \sigma^{-n/2} & \text{in } Q_j \\ 0 & \text{outside} \end{cases} \tag{84}$$

It easy to check that $\|w_j\|_{L^2(\Omega)} = 1$ and choosing $\frac{n}{2}\sigma^2 \leq \epsilon$ produces Friedrich's inequality. □

Proof of Rellich's theorem. Let $\epsilon > 0$ be given. The sequence (u_n) is bounded in $H^1_0(\Omega)$, so by the Poincare inequality in particular bounded in $L^2(\Omega)$. Banach-Alaoglu allows us to extract a weakly convergent subsequence $u_{n_k} \rightharpoonup u$. Since u_{n_k} is of course still a sequence in $H^1_0(\Omega)$ we can apply Friedrich's inequality to the difference $u_{n_k} - u_{n_l}$, which yields

$$\|u_{n_k} - u_{n_l}\|^2_{L^2(\Omega)} \leq \epsilon \|u_{n_k} - u_{n_l}\|^2_{H^1_0(\Omega)} + \sum_{j=1}^{M} \left[\int_\Omega (u_{n_k} - u_{n_l}) \cdot w_j \right]^2 . \tag{85}$$

Since the sum has finitely many terms we can, by the weak convergence of the u_{n_k}, find K and L large such that we have

$$\|u_{n_k} - u_{n_l}\|^2_{L^2(\Omega)} \leq \epsilon \|u_{n_k} - u_{n_l}\|^2_{H^1_0(\Omega)} + \epsilon \quad \text{for } k \geq K \text{ and } l \geq L\,. \tag{86}$$

Applying the triangle inequality to the first term and using that the sequence (u_{n_k}) is bounded in $H^1_0(\Omega)$ we conclude

$$\|u_{n_k} - u_{n_l}\|^2_{L^2(\Omega)} \leq 2c \cdot \epsilon + \epsilon \quad \text{for } k \geq K \text{ and } l \geq L\,. \tag{87}$$

□

Remark 10.3. *The above is the simplest case of a (compact) Sobolev embedding theorem. You will see more of those in a PDE course.*

Exercise 10.3. *What does this tell us about the original problem (88)? What about the problem*

$$DP_\star \begin{cases} \Delta u + \lambda u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{88}$$

For what λ does there exist a unique (weak) solution?

11 Exercises and Problems

11.1 Week 1

1. Write a complete proof of Minkowski's inequality for ℓ^p.

2. Let c denote the set of all elements $(\xi_1, ...)$ in ℓ^∞ such that the sequence (ξ_j) converges and let c_0 be the set of all such elements for which $\xi_n \to 0$ as $n \to \infty$. Prove that c and c_0 are Banach spaces.

3. The sequence space s. Show that in the space s we have $x_n \to x$ if and only if $\xi_j^{(n)} \to \xi_j$ for all $j = 1, 2, ...$ where $x_n = (\xi_j^{(n)})$ and $x = (\xi_j)$. Prove that s is complete.

4. Prove that if a sequence of continuous functions $f_n : [a, b] \to \mathbb{R}$ converges uniformly on $[a, b]$, then the limit function f is continuous.

5. Let $(X, \|\cdot\|)$ be a normed space. Prove that X is complete if and only if every series $\sum x_i$ in X satisfying $\sum_{i=1}^\infty \|x_i\| < \infty$ converges to a limit in X. Give an example of a (necessarily incomplete!) space X and a series for which $\sum_{i=1}^\infty \|x_i\| < \infty$ but $\sum_{i=1}^\infty x_i$ does not converge in X. *Hint:* Consider the space of sequences with finitely many non-zero entries...

11.2 Week 2

1. This problem culminates in an example of a compact subset in an infinite dimensional normed space. Let (X, d) be a metric space. Recall that (X, d) is called totally bounded if for any $\epsilon > 0$ one can find a finite collection of open balls of radius ϵ whose union contains X.

   - Show that a bounded metric space does not have to be totally bounded (hint: discrete metric space).
   - Show that X is compact if and only if it is complete and totally bounded.
   - Consider $c_0 \subset \ell^\infty$ the space of all sequences converging to zero. Fix a sequence $x \in c_0$ and let

     $$S_x = \{y \in c_0 \mid |y_n| \leq |x_n|\}$$

     Show that S_x is a compact subset of c_0.

2. If a linear vector space is infinite dimensional, then there exist on it norms which are not equivalent.

3. A linear operator from a normed linear space X into a normed linear space Y is bounded if and only if it maps bounded sets onto bounded sets.

4. Let T be a bounded linear operator from a normed space X *onto* a normed space Y. If there exists a constant $B > 0$ such that

$$\|x\| \leq B\|Tx\| \quad \text{for all } x \in X$$

then the inverse $T^{-1} : Y \to X$ exists and is bounded.

5. Show that the inverse $T^{-1} : \mathcal{R}(T) \to X$ of a bounded linear operator $T : X \to Y$ need not be bounded. Hint: Remark 4.1.

6. If X and Y are normed vector spaces (say over $\mathbb{R}$) and $B(X, Y)$ is complete, show that Y is complete. (Converse to Theorem 4.5. Hint: Hahn-Banach!)

7. If X is a normed space with $\dim X = \infty$ show that the dual space X' is not identical with the algebraic dual space $X^\star$. Hint: Show that every infinite dimensional normed space has unbounded linear functional.

8. Show that the dual of c_0 is isomorphic to ℓ^1. (Wait for Week 3. See also first exercise of Week 3.)

11.3 Week 3

1. Show that for $p > 1$ we have $(\ell^p)' = \ell^q$ for $1/p + 1/q = 1$.

2. For every x in a normed space X we have

$$\|x\| = \sup_{f \in X', f \neq 0} \frac{|f(x)|}{\|f\|}$$

Hence for fixed x the map $X' \ni f \mapsto f(x) \in \mathbb{R}$ is a bounded linear functional on X' of norm $\|x\|$. Consider now the map

$$\begin{aligned} \phi : X &\to X'' := (X')' \\ x &\mapsto \phi_x \end{aligned} \tag{89}$$

where $\phi_x(f) = f(x)$ on functionals $f \in X'$. Show that ϕ is a bounded injective linear map with $\|\phi_x\| = \|x\|$. [The map ϕ is called the canonical mapping of X into its bidual X''. If ϕ is bijective, i.e. its range is all of X'', the normed space X is called reflexive. Can you give examples of reflexive and non-reflexive normed spaces?]

3. *Prove the Hahn-Banach theorem for complex vector spaces:

Theorem 11.1. *Let X be a real or complex vector space with a real valued functional p satisfying $p(x+y) \leq p(x) + p(y)$ and $p(\alpha x) \leq |\alpha| p(x)$ for $\alpha \in \mathbb{R}$ or $\mathbb{C}$ respectively. Let f be a linear functional defined on a subspace $M \subset X$ satisfying $|f(x)| \leq p(x)$ for all $x \in M$. Then f has a linear extension F from M to X satisfying $|F(x)| \leq p(x)$.*

You may want to consult Kreyszig (4.3-1) or Schechter (Theorem 6.26).

4. Prove Theorem 5.5. Hint: Define a linear functional f on the space of all elements of the form $z = \alpha x_0 + x$ for $x \in M$ and $\alpha \in \mathbb{R}$ and extend it. Geometric interpretation/ analogue in Euclidean space?

5. Show that for a separable (say real) normed space X we can prove the Theorem 5.2 version of the Hahn-Banach theorem without invoking Zorn's Lemma.

6. The following geometric version of the Hahn-Banach theorem is often useful. (The geometric content is that we separate an open convex set from a point by a hyperplane just as we have seen in the finite dimensional case in Proposition 5.1.).

   Theorem 11.2. *Let V be a real normed vector space, $0 \in K$ be an open convex subset of X. If $x_0 \in X$ is a point not in K, then there exists a continuous linear map $\phi : X \to \mathbb{R}$ with*

   $$\phi(x_0) = 1 \quad \text{and} \quad \phi(v) < 1 \text{ for all } v \in K$$

   Prove this theorem. One way to proceed is as in the finite dimensional case, i.e. to define

   $$p_K(x) = \inf\{\alpha > 0 \mid \frac{x}{\alpha} \in K\}$$

   and to show (just as in the finite dimensional case) that this is well-defined, sublinear and that $p_K(x) < 1$ if and only if $x \in K$. Finally, show that $0 \le p_K(x) \le c\|x\|$ for some constant c and apply the Hahn-Banach Theorem 5.2. What happens if K is convex and *closed*? Can you prove a stronger version of the theorem ("strict separation")? HINT: Find an open ball B around x not intersecting K, then separate the ball from K. To achieve the latter look at the difference $K - B = \{x \in X \mid x = k - b \text{ for } k \in K \text{ and } b \in B\}$. This set is open (why?) and convex (why?) and doesn't contain 0.

11.4 Week 4

1. In a complete metric space a generic set is dense.

2. A Banach space X with $\dim X = \infty$ cannot have a countable Hamel basis.

3. Consider the normed vector space $BV[a,b]$ with $\|f\|_{BV} = |f(a)| + V(f)$ where $V(f)$ is the variation of f on $[a,b]$ and let $\|f\|_\infty = \sup_t |f(t)|$. Show that the norm $\|f\|_\infty$ is weaker than $\|f\|_{BV}$. Show that $BV[a,b]$ with norm $\|f\|_{BV}$ is a Banach space.

4. Let V and W be normed spaces and $T \in B(V,W)$. If $T^{-1} : W \to V$ exists and is bounded, show that $\left(T^{-1}\right)^\times = \left(T^\times\right)^{-1}$.

5. *This problem describes a famous application of Baire's theorem. Consider a sequence f_n of continuous (say real) valued functions on a complete metric space X and let

   $$\lim_{n\to\infty} f_n(x) = f(x)$$

exist for every $x \in X$. Then, the set of points where f is continuous is a generic set in X. For this problem you should not hesitate to consult a book if you get stuck.

6. Let V and W be Banach spaces and

$$B : V \times W \to \mathbb{C}$$

be a bilinear functional which is continuous in each variable, i.e. $B(\xi, \cdot) : W \to \mathbb{C}$ is linear and continuous for each fixed $\xi \in V$ and similarly $B(\cdot, \eta) : V \to \mathbb{C}$ is linear and continuous for each fixed $\eta \in W$. Prove that B is continuous.

11.5 Week 5

1. Do Exercise 6.1.

2. The linear operations are continuous with respect to strong operator convergence.

3. In ℓ^1 weak and strong convergence are equivalent.
Remark: This is a hard problem. Don't hesitate to consult the literature ("Schur's Lemma", "Schur property"). Here is a sketch for a proof: Assume you have a sequence $x_n \rightharpoonup 0$ but $x_n \nrightarrow 0$. Then there is a subsequence x_{n_k} with $\|x_{n_k}\| > \epsilon$ for some $\epsilon > 0$. From that subsequence construct a bad element $y \in \ell^\infty$ that contradicts $x_{n_k} \rightharpoonup 0$. To do the latter, construct first a sub-subsequence of $x_{n_{k_j}}$ and sequence of intervals $M_1 < M_2 < M_3 < ...$ such that the energy of $x_{n_{k_j}}$ concentrates in the interval $M_{j-1} < i < M_j$ (i.e. the head and the tail of $x_{n_{k_j}}$ are small in ℓ^1). Use this sub-subsequence to construct the $y \in \ell^\infty$.

4. Do the worksheet on weak convergence (see webpage)

11.6 Week 6

1. Let X and Y be Banach spaces and $T : X \to Y$ an injective bounded linear operator. Show that $T^{-1} : R(T) \to X$ is bounded if and only if $R(T)$ is closed in Y.

2. Let X, Y be normed spaces and $T : \mathcal{D}(T) \to Y$ be a bounded linear operator with domain $\mathcal{D}(T) \subset X$. Prove the following: If T is closed and Y is complete, then $\mathcal{D}(T)$ is closed in X.

3. Let $X_1 = (X, \|\cdot\|_1)$ and $X_2 = (X, \|\cdot\|_2)$ be Banach spaces. If there exists a constant c such that $\|x\|_1 \le c\|x\|_2$ for all $x \in X$, show that there is a constant k such that $\|x\|_2 \le k\|x\|_1$ for all $x \in X$ (so the norms are equivalent).

4. Show that if the inverse T^{-1} of a closed linear operator exists, then T^{-1} is also closed.

5. Let X and Y be Banach spaces and let T be a bounded linear map from X into Y. If $T(X)$ is of the second category in Y, then $T(X) = Y$. Hint: Convince yourself that the proof of the open mapping theorem still goes through. Then conclude surjectivity.

11.7 Week 7

1. Do the exercises in the next (there are quite a few this time).
2. Show that every Hilbert space is reflexive. Cf. Exercise 2 of Week 3.
3. Let H be a *complex* Hilbert space and $T : H \to H$ a bounded linear map.
    - Show that $\langle x, Tx \rangle = 0$ for all $x \in H$ implies $T = 0$. Is this true in the real case? (Hint: Rotations.)
    - Show that if $\langle x, Tx \rangle$ is *real* for all $x \in H$, then T is self-adjoint.
4. Let H be a Hilbert space, and let $x_n \in H$ converge weakly to a limit $x \in H$. Show that the following statements are equivalent:
    (a) x_n converges strongly to x.
    (b) $\|x_n\|$ converges to $\|x\|$.

11.8 Week 8

1. The product of two bounded self-adjoint operators $S : H \to H$ and $T : H \to H$ is self-adjoint if and only if T and S commute, i.e. $ST = TS$.
2. Let (e_n) be an orthonormal basis in a separable Hilbert space. Consider the right shift operator

$$T : H \to H \qquad \text{with } \; Te_n = e_{n+1}$$

   Find the range, null space and Hilbert adjoint of T.
3. Let $S = I + T^{\star}T : H \to H$ where T is linear and bounded. Show that the inverse $S^{-1} : S(H) \to H$ exists.
4. The following theorem is the Lax-Milgram theorem, a powerful extension of Riesz theorem which is used in the existence theory for weak solutions for PDEs.

   Theorem 11.3. *Let $B(x, y)$ be a bilinear functional on $H \times H$ for a (say real) Hilbert space H. Assume that there exist constants $C > 0$ and $c > 0$ such that*

    - $|B(x, y)| \leq C\|x\|\|y\|$ *for all $x, y \in H$ (boundedness)*
    - $|B(x, x)| \geq c\|x\|^2$ *for all $x, y \in H$ (coercivity)*

   Then for any bounded linear functional $f : H \to \mathbb{R}$ there exists a unique point $x \in H$ such that

$$f(y) = B(x, y) \qquad \textit{holds for all } y \in H\,.$$

   Hints for the proof: The idea is of course to reduce it to Riesz theorem. For fixed x, the mapping $y \to B(x, y)$ is a bounded linear functional to which we can apply Riesz. This gives rise to a map $A : H \to H$ which associated to x its Riesz representation. This map is isometric (hence injective) and also surjective (here you will have to use the coercivity).

5. Let $X = C[0,1]$ the space of continuous functions on the interval $[0,1]$. For $g \in X$ consider the integral operator

$$Tf(x) = \int_0^x g(t) f(t)\, dt$$

Show that $T : C[0,1] \to C[0,1]$ is compact.

6. Prove that if T can be diagonalized with respect to some basis $\{\varphi_k\}$ of eigenvectors and corresponding eigenvalues λ_k, then T is compact if and only if $|\lambda_k| \to 0$.

7. Prove that Hilbert Schmidt operators are compact.

11.9 Week 9

1. Let H be a separable Hilbert space. Prove that
   (a) If T_1 and T_2 are two self-adjoint compact operators that commute ($T_1T_2 = T_2T_1$), then they can be diagonalized simultaneously.
   (b) Prove that if T is normal and compact then T can be diagonalized. Hint: Write $T = T_1 + iT_2$ for T_1, T_2 self-adjoint and compact.
   (c) If U is unitary and $U = \lambda \cdot id - T$ for T compact, then U can be diagonalized.

2. The following exercise illustrates that you can neither drop self-adjoint, nor compact from the assumptions of the spectral theorem.
   (a) Consider the operator $T : L^2[0,1] \to L^2[0,1]$ defined by

   $$T(f)(x) = xf(x)$$

   Prove that T is a bounded linear operator with $T = T^\star$ but that T is not compact. Show that T has no eigenvectors.
   (b) Let H be a separable Hilbert space with basis $\{\varphi_k\}_{k=1}^\infty$. Show that the operator T defined by

   $$T\varphi_k = \frac{1}{k}\varphi_{k+1}$$

   is compact but has no eigenvectors.

3. Consider Volterra's integral equation

$$f(s) = g(s) + \int_0^s K(s,t) f(t)\, dt \qquad \text{for } 0 \le s \le 1\,, \tag{90}$$

where $K(s,t)$ is continuous for $0 \le s, t \le 1$. Prove that for any continuous function g there exists a unique continuous solution f of (90).

4. *Let $\Omega \subset \mathbb{R}^n$ be a compact set of $\mathbb{R}^n$. Let $K(x,y) \in L^2(\Omega \times \Omega)$. For any given $g \in L^2(\Omega)$ consider the Fredholm integral equation (in $L^2(\Omega)$)

$$f(x) = g(x) + \lambda \int_\Omega K(x,y) f(y)\, dy\,, \tag{91}$$

with $\lambda \in \mathbb{R}$ a parameter. Prove that if $g = 0$ implies $f = 0$, then there exists a unique solution of (91).

5. *Suppose $[a, b]$ is a bounded interval, and L is defined on functions $f \in C^2[a, b]$ by

$$L(f)(x) = \frac{d^2 f}{dx^2} - q(x) f(x) \ ,$$

for $q(x)$ a continuous, non-negative[23] real-valued function on $[a, b]$. We say that $\varphi \in C^2[a, b]$ is an eigenfunction (of L) with eigenvalue μ if $L\varphi = \mu\varphi$ and $\varphi(a) = \varphi(b) = 0$. Show the following

(a) Then eigenvalues μ are strictly negative and each eigenspace is one-dimensional

(b) Eigenvectors corresponding to distinct eigenvalues are orthogonal in $L^2[a, b]$.

(c) Let $K(x, y)$ be the *Green's kernel* defined as follows. Choose $\varphi_-(x)$ to be a solution of $L(\varphi_-) = 0$ with $\varphi_-(a) = 0$ but $\varphi'_-(a) \neq 0$. Similarly, choose $\varphi_+(x)$ to be a solution of $L(\varphi_+) = 0$ with $\varphi_+(b) = 0$ but $\varphi'_+(b) \neq 0$. Let $w = \varphi'_+(x)\varphi_-(x) - \varphi'_-(x)\varphi_+(x)$ be the *Wronskian* of these solutions and note that w is a non-zero constant ($w'(x) = 0$). Set

$$K(x, y) = \begin{cases} \frac{\varphi_-(x)\varphi_+(y)}{w} & \text{if } a \leq x \leq y \leq b \\ \frac{\varphi_+(x)\varphi_-(y)}{w} & \text{if } a \leq y \leq x \leq b \end{cases} \tag{92}$$

Then the operator T defined by

$$T(f)(x) = \int_a^b K(x, y) f(y)\, dy \tag{93}$$

is a Hilbert-Schmidt operator, and hence compact (cf. Exercise 7 of Section 11.8). It is also symmetric. Moreover, whenever f is continuous on $[a, b]$, Tf is of class $C^2[a, b]$ and

$$L(Tf) = f \ .$$

(d) Conclude that each eigenvector of T with eigenvalue λ is an eigenvector of L with eigenvalue $\frac{1}{\lambda}$. By the spectral theorem (Theorem 9.5) we obtain a complete orthonormal set of eigenvectors of L.

11.10 Week 10

For mastery students: Read Chapter 6.2 of [Evans] (existence of weak solutions via Lax Milgram).

12 Acknowledgments

I thank Juan Kuntz and Pierre Vulliemin for many comments and corrections.

[23]This is for simplicity; you may want to lift that restriction and reformulate the statements after having solve the problem.