B4.2 Functional Analysis II notes

Jonah Duncan

Disclaimer: What I've written below is in no way intended to be comprehensive, but rather gives an overview of the course and stresses some of the main results. The numbering used corresponds to the lecture notes of Luc Nguyen, but any mistakes here are my own and I stress that these notes are unofficial and should be cross-checked with your own notes.

A bit of notation: $L(X, Y)$ consists of bounded (i.e. continuous) linear maps from X to Y, also written $\mathcal{B}(X, Y)$ in your notes. A 'linear functional' is assumed to be bounded, i.e. an element of X^*. Sometimes we say 'X is a Banach space' or 'X is a Hilbert space' without specifying the corresponding norm or inner product - these will be denoted $||\cdot||_X$ and $\langle \cdot, \cdot \rangle_X$ wherever they appear.

1 Chapter 1: Hilbert Spaces

Whilst some of the results in this course apply generally to Banach spaces (and you should make sure to distinguish which ones), the main goal is to introduce Hilbert spaces. These should simply be thought of as Banach spaces in which the norm is 'induced by' an inner product. Explicitly, a Hilbert space is a normed vector space $(X, ||\cdot||_X)$, complete with respect to the norm $||\cdot||_X$, for which there exists an inner product $\langle \cdot, \cdot \rangle_X : X \times X \to \mathbb{F}$ satisfying $||x||_X^2 = \langle x, x \rangle_X$ for all $x \in X$. The field $\mathbb{F}$ is typically either $\mathbb{R}$ or $\mathbb{C}$.

Perhaps it is not clear why a Hilbert space should possess so many nice properties that a general Banach space does not, but indeed this is the case and the first chapter of the course is dedicated to convincing you of this. The first main result - whose proof often comes up in past exams - is:

Theorem 1.2.4. *Let X be a Hilbert space and $K \neq \emptyset$ a closed convex subset of X. Then for every $x \in X$, there exists a unique point $y_0 \in K$ such that*

$$||x - y_0||_X = \inf_{y \in K} ||x - y||_X. \tag{1.1}$$

Remark 1.1. This result does not require K to be a sub*space* of X. Of course, if K is a subspace, then in particular it is a convex subset, and all we have to check is closure. As is evident from the proof (see your lecture notes) of the projection theorem below, if K is a subspace then a necessary and sufficient condition for $y_0 \in K$ to be the minimiser is that $x - y_0 \in K^\perp$.

The proof of Theorem 1.2.4 uses the *parallelogram law*: for $x, y \in X$,

$$||x + y||_X^2 + ||x - y||_X^2 = 2||x||_X^2 + 2||y||_X^2. \tag{1.2}$$

In fact, the identity (1.2) holds in a Banach space $(X, ||\cdot||_X)$ *if and only if* the norm $||\cdot||_X$ is induced by an inner product, that is, if and only if $(X, ||\cdot||_X)$ is actually a Hilbert space. So aside from being useful in proofs, (1.2) gives you an easy way of determining of whether or not a given Banach space is a Hilbert space. Don't forget this!

Using Theorem 1.2.4, one can then prove the projection theorem:

Theorem 1.2.5 (Projection theorem). *Let X be a Hilbert space and Y a closed subspace of X. Then $X = Y \oplus Y^\perp$.*

Remark 1.2. Whilst $Y^\perp$ is well-defined for Y just a subset of X, here we *do* require Y to be a subspace. Note that the fact we have $\oplus$ means that each element $x \in X$ can be written *uniquely* as $x = y_1 + y_2$ for $y_1 \in Y$ and $y_2 \in Y^\perp$.

As a consequence of Theorem 1.2.5, a closed subspace Y of a Hilbert space X satisfies $Y^{\perp\perp} = Y$. In general, a subset Y of a Hilbert space X only satisfies $Y \subset Y^{\perp\perp}$. See Proposition 1.2.3 for other similar properties you should know.

As is clear from the definition, orthogonality is a concept that has no meaning in general Banach spaces as we need the notion of an inner product. We can extend this notion even further (orthonormality) to gain a very useful result concerning the existence of an 'orthornormal basis'. We recall that a subset S (infinite or otherwise) of a Hilbert space X is called an *orthonormal set* (or *orthonormal sequence*) if $||x||_X = 1$ for all $x \in S$ and $\langle x, y\rangle_X = 0$ for all $x, y \in S$ with $x \neq y$. If, in addition, $\overline{\operatorname{span} S} = X$, then we call S an *orthonormal basis* (or *complete orthonormal set/sequence*). These always exist in a Hilbert space:

Theorem 1.2.10. *Every Hilbert space contains an orthonormal basis.*

Remark 1.3. The issue of defining a suitable notion of a basis for an infinite dimensional Banach space is a tricky one, dealt with partially in C4.1 Further Functional Analysis. The above definition of an orthonormal basis for a Hilbert space should be treated exactly as that: a definition. Theorem 1.2.13 below then relates this to the more usual notion of a basis you are familar with.

A direct computation verifies that for any finite orthonormal set $S = \{x_1, \ldots, x_m\}$ in a Hilbert space X, one has the Pythagorean identity

$$||x||_X^2 = \sum_{n=1}^{m} \left|\langle x, x_n\rangle_X\right|^2 + \left|\left|x - \sum_{n=1}^{m} \langle x, x_n\rangle_X \, x_n\right|\right|_X^2 \tag{1.3}$$

for all $x \in X$. An immediate consequence is Bessel's inequality, which says that for a (possibly infinite) orthonormal set $S = \{x_1, x_2, \ldots\}$, one has

$$\sum_{n=1}^{\infty} \left|\langle x, x_n\rangle_X\right|^2 \leq ||x||_X^2 \tag{1.4}$$

for all $x \in X$. This, in turn, is used to prove:

Theorem 1.2.13 (Parseval's Identity)**.** *Let X be a Hilbert space and $S = \{x_1, x_2, \ldots\}$ a (possibly infinite) orthonormal set. Then for all $x \in \overline{\operatorname{span} S}$, one can write*

$$x = \sum_{n=1}^{\infty} \langle x, x_n\rangle_X x_n \qquad \textit{in the sense that } \lim_{N\to\infty} \left|\left|x - \sum_{n=1}^{N} \langle x, x_n\rangle_X x_n\right|\right|_X = 0, \tag{1.5}$$

where the sequence $(a_i)_{i\in\mathbb{N}}$ defined by $a_i = \langle x, x_i\rangle_X$ satisfies

$$\sum_{n=1}^{\infty} |a_n|^2 = \sum_{n=1}^{\infty} \left|\langle x, x_n\rangle_X\right|^2 = ||x||_X^2. \tag{1.6}$$

In particular, $(a_i)_{i\in\mathbb{N}} \in \ell^2$.

Remark 1.4. Note that Bessel's inequality holds for all $x \in X$, whilst Parseval's identity holds only for $x \in \overline{\operatorname{span} S}$. In particular, if S is an orthonormal basis, then $\overline{\operatorname{span} S} = X$ and Theorem 1.2.13 applies to all $x \in X$. Combined with Theorem 1.2.10 it follows that we can always express an element of a Hilbert space in the form (1.5) for some orthonormal basis S (even if we don't know what S is explicitly). Are there any proofs in your notes (or problem sheets) that could also be done writing things out in this way?

The next result, the Reisz representation theorem (RRT), is one of the cornerstones of the theory of Hilbert spaces. Here is a way to think about it: given $x \in X$, we can define a canonical linear functional $\ell : X \to \mathbb{F}$ by $\ell(y) = \langle x, y\rangle_X$ for all $y \in X$. This is bounded by the Cauchy-Schwarz inequality. The RRT tells us that, conversely, *any* linear functional on a Hilbert space X has the above form, for some unique $x \in X$. Not only is this fact very useful in proofs, since we can essentially replace any linear functional (no matter how complicated its definition) with something simpler (an inner product), but this also gives us a bijective correspondence between X and X^* (in fact we can say more than this - see the remark below).

Theorem 1.3.1. (RRT)**.** *Let X be a Hilbert space and suppose $\ell \in X^*$. Then there exists a unique element $x \in X$ such that*

$$\ell(y) = \langle x, y\rangle_X \tag{1.7}$$

for all $y \in X$. Moreover, $||x||_X = ||\ell||_{X^*}$.

Remark 1.5. Recall in B4.1 Functional Analysis I we had the issue of trying to describe the dual space X^* of a given Banach space X in terms of a space we already knew well. In other words, in order to understand X^*, we would try to find a space Y isometrically isomorphic to X^* (the isometry meaning the preservation of norm, the isomorphism meaning a linear bijection), where Y was a space we already understood. For instance, for $1 \leq p < \infty$ we (partially) showed that the dual space of $L^p(\Omega)$ could be realised as the space $L^q(\Omega)$, where $\frac{1}{p} + \frac{1}{q} = 1$. Even this relatively simple example wasn't fully proved (surjectivity was left out).

The RRT implies that for a Hilbert space X we can avert this problem, as we can just realise X^* as isometrically isomorphic ($\cong$) to X itself! The linear bijection aspect is clear, and the fact that $||x||_X = ||\ell||_{X^*}$ is the isometry. Relating this to the Lebesgue spaces example: it is a fact that $L^p(\Omega)$ is Hilbert if and only if $p = 2$, with inner product

$$\langle f, g\rangle_{L^2(\Omega)} = \int_\Omega f\bar{g}, \tag{1.8}$$

but in this case $\frac{1}{p} + \frac{1}{q} = 1$ implies we also have $q = 2$. So $(L^2(\Omega))^* = (L^p(\Omega))^* \cong L^q(\Omega) = L^2(\Omega)$, in agreement with RRT.

Corollary 1.6. *A Hilbert space X is reflexive, that is $X = X^{**}$ (more strictly, X is isometrically isomorphic to X^{**}).*

Remark 1.7. Since $\mathbb{R}$ is complete with respect to $|\cdot|$, we know that X^* is Banach even when X is only a normed vector space. Similarly, X^{**} is always Banach. It follows that a reflexive normed vector space is automatically Banach, since it is isometrically isomorphic to X^{**}.

The remainder of Chapter 1 is dedicated to adjoint operators and related definitions (self-adjoint operators, unitary operators, normal operators etc.). There is not so much to say here, except that you should get used to proving results in the style of questions 1-3 in your second problem sheet, as these tend to make typical bookwork exam questions.

Remark 1.8. At least a couple of times, past exam questions have asked you to define the adjoint $T^* : Y \to X$ of a bounded operator $T : X \to Y$ between Hilbert spaces *and explain why it is well-defined as a bounded operator* (you can use RRT). It is worth mentioning that $T^* : Y^* \to X^*$ can also be defined for X, Y only Banach, but of course no identification can typically be made between X and its dual (likewise for Y). Things get much more complicated if we want to define adjoints of unbounded operators (these often arise in quantum mechanics, see Brian Hall's book...)

2 Chapter 2: Baire category theorem + consequences

The Baire category theorem comes in many different forms - the version you have to know is perhaps the simplest:

Theorem 2.1.2 (Baire category theorem)**.** *A complete metric space is never the union of a countable number of nowhere dense sets.*

Remark 2.1. The BCT is really a result in topology rather than functional analysis, but clearly has application in important functional analytic results listed below. I'd say you're more likely to be tested on these theorems rather than the proof of BCT, but then again Q4 of your second problem sheet was a similar style proof, so it's probably best to just learn it (unless you've been told otherwise).

The rest of the chapter introduces some well-known results that make us of the BCT in their proofs - note that none of these assume we are working with Hilbert spaces, and in some parts we only need normed vector spaces. Make a list of what theorem requires what! That said, these theorems may have specialised consequences in the context of Hilbert spaces (e.g. Theorems 2.2.2, 2.3.4, Example 2.4.4 in your notes).

Theorem 2.2.1 (Principle of Uniform Boundedness/ Banach-Steinhaus)**.** *Suppose X is a Banach space and Y is a normed vector space, and let $\mathcal{F} \subset L(X,Y)$ be a family of bounded linear operators. Then if for all $x \in X$ one has*

$$\sup_{T\in\mathcal{F}} ||Tx||_Y < \infty, \tag{2.1}$$

one also has

$$\sup_{T\in\mathcal{F}} ||T||_{L(X,Y)} < \infty. \tag{2.2}$$

Remark 2.2. By definition of the operator norm $||\cdot||_{L(X,Y)}$, another way of phrasing the PUB is that if for each $x \in X$ one has $\sup_{T\in\mathcal{F}} ||Tx||_Y < \infty$, one also has $\sup_{T\in\mathcal{F},||x||=1} ||Tx||_Y < \infty$. In other words, if we have boundedness of Tx over $\mathcal{F}$ for each x *separately*, then we have boundedness over $\mathcal{F}$ and x *simultaneously*. Clearly the converse holds, so PUB is really saying that we have equivalence of pointwise boundedness and uniform boundedness.

Theorem 2.3.1 (Opening mapping theorem)**.** *Let X, Y be Banach spaces and $T \in L(X,Y)$ surjective. Then T maps open sets in X to open sets in Y.*

Remark 2.3. Open mapping theorems appear in other contexts too, for instance the *invariance of domain theorem*: if $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^n$ is injective and continuous, then $f(U)$ is open and f is a homeomorphism between U and $f(U)$. This can be extended to manifolds (if you are familiar with these): if $f : M \to N$ is continuous and locally injective, then it is open and a local homeomorphism.

As a particular consequence of the Banach spaces version of the OMT: if T is *bijective* then its inverse operator is automatically bounded, i.e. $T^{-1} \in L(Y,X)$ (you can think of this corresponding to the homeomorphisms in the versions of the OMT mentioned above). This is usually stated as the *inverse mapping theorem*, and can be used to prove:

Theorem 2.4.1 (Closed graph theorem)**.** *Let X, Y be Banach spaces and T a linear operator from X to Y. Then $T \in L(X,Y)$ if and only if its graph*

$$\Gamma(T) := \{(x, Tx)\} \subset X \times Y \tag{2.3}$$

is closed in $X \times Y$ (equipped with the product topology).

Remark 2.4. The closed graph theorem is just a more complicated analogue of the result that a function $f : \mathbb{R} \to \mathbb{R}$ is continuous if and only if its graph is closed. For instance, the graph of $f(x) = x^2$ is obviously closed (any sequence of points in the graph can't have a limit outside the graph), whereas the heaviside step function h, defined by $h(0) = 0$ and $h(x) = \mathrm{sign}(x)$ for $x \neq 0$, is clearly not closed (take the sequence $(x_n, h(x_n)) = (1/n, 1)$), nor is it continuous.

If you're interested, it is worth noting that all the above results (PUB, OMT, CGT) hold under weaker assumptions, namely we can replace Banach spaces with more general 'Fréchet spaces'. Most graduate-level textbooks on functional analysis will cover this.

3 Chapter 3: Weak convergence

A normed vector space $(X, ||\cdot||_X)$ clearly induces a metric d_X on X, namely $d_X(x,y) = ||x-y||_X$ (the fact that this really is a metric follows from properties of the norm). This in turn induces a topology on X in the usual way, which we refer to as the *strong topology*. Under this topology, a set $U \subset X$ is open if and only if for every $x \in U$, there exists $r > 0$ such that the 'open ball' $B(x,r) = \{y \in X : d_X(x,y) < r\}$ is contained in U. A sequence $\{x_n\} \subset X$ is said to *converge strongly* (or sometimes just *converge*) in (X, d_X) (or in $(X, ||\cdot||_X)$, or in X with respect to $||\cdot||_X$ or d_X) if and only if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d_X(x, x_n) = ||x - x_n||_X < \epsilon$ for all $n > N$.

The strong topology on $(X, ||\cdot||_X)$ is characterised by the fact that it is the coarsest topology (the one with the 'fewest open sets') for which $||\cdot||_X$, vector addition and scalar multiplication are continuous. The *weak topology*, on the other hand, is characterised by the fact that it is the coarsest topology on

X for which every element in X^* is continous and for which vector addition/scalar multiplication are still continuous. Explicitly, a subset of X is open with respect to the weak topology on X if and only if it can be written as a (possibly infinite) union of sets, each of which is the intersection of finitely many sets of the form $\varphi^{-1}(U)$, where $\varphi \in X^*$ and $U \subset X$ is open. Don't worry about remembering this definition - we are only concerned with whether a sequence in X *converges weakly* to some $x \in X$, which (by definition for us) occurs if and only if

$$\ell(x_n) \to \ell(x) \quad \text{for all } \ell \in X^*. \tag{3.1}$$

In this case we write $x_n \rightharpoonup x$. Those who are interested in the weak topology can refer to the exam-style question at the end of these notes, although I'm of the opinion something like that is unlikely to come up in your exam (I've based it around an exercise in Brezis' book, which I highly recommend if you'd like a seamless introduction from functional analysis to PDEs).

Remark 3.1. If X is a Hilbert space, then each $\ell \in X^*$ can be written as $\ell(\cdot) = \langle \cdot, x \rangle$ for some unique $x \in X$ by RRT. Remember this when checking for weak convergence in a Hilbert space, as it will likely simplify things (the proof of Proposition 3.1.3 illustrates this very nicely).

As the names suggest, if a sequence converges strongly then it converges weakly (to the same limit), but the converse is generally false. For instance:

Proposition 3.1.3. *Let $\{x_n\}$ be an orthonormal sequence in a Hilbert space X. Then $x_n \rightharpoonup 0$ but $x_n \not\to 0$.*

Another example is Example 3.1.4 in your notes. Sometimes, however, weak convergence does imply strong convergence: for instance in ℓ^1 (see Problem Sheet 3), or in any finite dimensional normed vector space.

Morally speaking it should therefore be easier to show a sequence converges weakly, and for sometimes this may be enough to prove what we want. That said, if we *do* want strong convergence, then weak convergence can be used as a stepping stone - the only other thing we need is convergence of norms:

Example 3.1.6. *Suppose X is a Hilbert space and $\{x_n\} \subset X$ satisfies $x_n \rightharpoonup x$ and $||x_n||_X \to ||x||_X$. Then $x_n \to x$ (i.e. $||x - x_n||_X \to 0$).*

Even though weakly convergent sequences are not necessarily strongly convergent, they share some of the same properties enjoyed by the latter. For instance:

Theorem 3.2.1. *Suppose $(X, ||\cdot||_X)$ is a normed vector space and $\{x_n\} \subset X$ a weakly converging sequence. Then there exists a constant $0 \leq C < \infty$ such that*

$$||x_n||_X \leq C \quad \text{for all } n \in \mathbb{N}. \tag{3.2}$$

Remark 3.2. Note that we only assume $(X, ||\cdot||_X)$ is a normed vector space in the above. Similarly, the results following do not make any Banach/Hilbert assumptions. This is a reminder to go through the results of your course and take note of what assumptions are made, and to pick out why the proofs require these assumptions: if $(X, ||\cdot||)$ is assumed to be Banach, where do we use completeness? Is it directly applied in the proof, or perhaps the proof makes reference to some earlier result that requires the space to be Banach? If $(X, ||\cdot||)$ is assumed to be Hilbert, do we directly use the fact that $||\cdot||$ is induced by an inner product? Or (more likely) do we make use of this fact less directly, say by appealing to RRT or using the parallelogram law somewhere?

Theorem 3.2.2. *Suppose $(X, ||\cdot||_X)$ is a normed vector space and $\{x_n\} \subset X$ a sequence converging weakly to x. Then*

$$||x||_X \leq \liminf_{n\to\infty} ||x_n||_X. \tag{3.3}$$

Remark 3.3. In technical terms, the above property is referred as *weak sequential lower semicontinuity* of the norm. This is an extremely useful notion if you end up working in PDEs and/or calculus of variations. As explained in the next remark, Theorem 3.2.2 can be seen as a consequence of:

Theorem 3.2.3 (Mazur's Lemma)**.** *Let $(X, ||\cdot||_X)$ be a normed vector space and $K \subset X$ a closed convex subset. If $\{x_n\} \subset K$ converges weakly to x, then $x \in K$.*

Remark 3.4. For each $N \in \mathbb{N}$, let C_N be the minimal constant such that (3.2) holds for all $n \geq N$. Then the truncated sequence $\{x_n\}_{n\geq N}$ belongs to the closed ball of radius C_N centred at zero, and hence by Mazur's theorem we have $x \in \cap_N C_N$. But the radius of $\cap_N C_N$ is just given by $\liminf_{n\to\infty} ||x_n||_X$, so Theorem 3.2.2 follows from Mazur's lemma.

Mazur's lemma relies on the (extended) hyperplane separation theorem. This can be obtained from the Hahn-Banach theorem, or if you like it can be thought of as a geometric manifestation of the Hahn-Banach theorem.

Theorem 3.3.2. *The closed unit ball B of a reflexive Banach space is weakly sequentially compact, i.e. every sequence in B has a subsequence converging weakly to a point in B. In particular, every bounded sequence in a reflexive Banach space has a weakly convergent subsequence.*

Remark 3.5. You prove this theorem for Hilbert spaces, which are always reflexive by RRT (see Corollary 1.6). The proof is further simplified by a direct application of RRT. Remarkably, the converse of Theorem 3.3.2 also holds, which is a result due to Eberlein: if the closed unit ball of a Banach space is weakly sequentially compact, then X is reflexive. One can use Theorem 3.3.2 to prove Theorem 1.2.4 in the case that X is only a reflexive Banach space, although y_0 may no longer be unique.

4 Chapter 4: Convergence of Fourier series

Theorem 1.2.10 told us that every Hilbert space contains an orthonormal basis, although as of yet there has been no need to construct an explicit example. Of particular interest, however, are the trigonometric functions

$$\{e_n\}_{n\in\mathbb{Z}} = \left\{ \frac{1}{\sqrt{2\pi}} e^{inx} : n \in \mathbb{Z} \right\} \tag{4.1}$$

which form an orthonormal basis of $L^2(-\pi,\pi)$. Proving this fact is the content of Theorem 4.3.1 in your notes.

Remark 4.1. This chapter in the notes serves only as a brief introduction, and for those interested in reading more at a similar level, the book *Fourier Analysis* by Duoandikoetxea is very clear and concise ($\sim$200 pages). In these notes, I will instead try to motivate some of the problems studied in Fourier analysis, referring in particular to the ones in your course.

Assuming for now that (4.1) does define an orthonormal basis of $L^2(-\pi,\pi)$, then Theorem 1.2.13 tells us that any $f \in L^2(-\pi,\pi)$ can be expressed as the L^2 limit of the sequence

$$S_N f := \sum_{n=-N}^{N} \langle f, e_n \rangle_{L^2} e_n \tag{4.2}$$

as $N \to \infty$. That is, $||f - S_N f||_{L^2} \to 0$ as $N \to \infty$. As done in Theorem 1.2.13, we usually just write this as

$$f = \sum_{n=-\infty}^{\infty} \langle f, e_n \rangle_{L^2} e_n \tag{4.3}$$

where the sum on the RHS is to be implicitly understood as convergent in the L^2 sense to f, as described above. Now, we know what the e_n are and we know what the L^2 norm is, so we can write (4.3) more explicitly as

$$f(x) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\int_{-\pi}^{\pi} f(y) e^{-iny}\, dy \right) e^{inx}, \tag{4.4}$$

(we have a negative power in the integral corresponding to the complex conjugate in the definition of the L^2 inner product) or just

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} \quad \text{where } a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny}\, dy. \tag{4.5}$$

Remark 4.2. Again, this notation is deceiving as it might suggest that f is equal pointwise to the sum in (4.5), but this is NOT what we are asserting! Indeed, $f \in L^2(-\pi,\pi)$ may not even be defined pointwise. It simply means $||f - S_N f||_{L^2} \to 0$ as $N \to \infty$. This is an important distinction to make, because a large part of Fourier analysis is determining when we can really *can* say that the sum converges pointwise (or even uniformly) to continuous f, in which case the notation in (4.5) could be interpreted more literally.

You may notice that the a_n are still well-defined if we only take $f \in L^1(-\pi,\pi)$, since then

$$|a_n| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(y)| \underbrace{|e^{-iny}|}_{=1} \, dy = ||f||_{L^1(-\pi,\pi)} < \infty. \tag{4.6}$$

We still write $S_N f$ for the partial sum as in (4.2), and it follows from (4.6) that this sum is finite for each N if we at least assume $f \in L^1(-\pi,\pi)$. That said, the limiting sum as $N \to \infty$ might diverge for some x, and even if it does converge, it is not immediately clear whether it converges to f in any reasonable sense - we certainly can no longer appeal to Theorem 1.2.10. Instead, we just write

$$\mathcal{F}(f)(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} \quad \text{where } a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y) e^{-iny} \, dy \tag{4.7}$$

and call $\mathcal{F}(f)$ the *Fourier series* of the function $f \in L^1(-\pi,\pi)$, acknowleding that this sum may diverge at some points or have little to do with f where it does converge.

In summary, there are two big questions to be considered:

1. When $f \in L^2(-\pi,\pi)$, we know that $\mathcal{F}(f)$ converges to f in the L^2 sense. Do we get any other modes of convergence, say pointwise or uniformly, if we impose additional continuity assumptions on f? Or if we do not assume continuity, can we still get almost everywhere convergence to f?

2. Does the Fourier series $\mathcal{F}(f)$ of f converge in some sense when we assume something weaker than $f \in L^2(-\pi,\pi)$, say $f \in L^1(-\pi,\pi)$? If so, is this limit a reasonable representation of f?

In your notes you see that one version of Q1 can be answered in the negative if we just assume $f \in C[-\pi,\pi]$ (note this implies $f \in L^2(-\pi,\pi)$):

Theorem 4.5.1. *There exists $f \in C[-\pi,\pi]$ satisfying $f(-\pi) = f(\pi)$ and such that $\mathcal{F}(f)$ diverges at a single point.*

Remark 4.3. This result is a classical application of the principle of uniform boundedness. It follows relatively easily from Theorem 4.5.1 that we can construct a continuous function whose Fourier series diverges at any finite number of given points (see Q1b of problem sheet 4). In fact, one can construct a continuous function whose Fourier series diverges on a dense set of points in $[-\pi,\pi]$, e.g. at all rational multiples of π (see Katznelson's book). However, even if we do not assume continuity on f, a highly non-trivial result due to Carleson states that the Fourier series of an L^2 function must converge pointwise *almost everywhere* to f (so for instance, we could not construct a continuous functions whose Fourier series diverges on the *complement* of the set of all rational multiples of π).

It is worth noting that a version of Q1 *can* be answered in the affirmative if we assume slightly more than continuity, e.g. if f is Hölder continuous, or if f is of bounded variation, or if f is differentiable, then we have pointwise convergence everywhere to f.

The next result in your notes answers a version of Q2 in the affirmative, namely that if $f \in L^1(-\pi,\pi)$ is 2π-periodic and is Hölder continuous at some point x_0, then its Fourier series converges to f *at that point*:

Theorem 4.6.1. *Suppose $f \in L^1(-\pi,\pi)$ is 2π-periodic and Hölder continuous at x_0. Then*

$$\lim_{N\to\infty} S_N f(x_0) = f(x_0). \tag{4.8}$$

Remark 4.4. Following on from Carleson's theorem mentioned above, Hunt went on to prove that the Fourier series of an L^p function, where $p \in (1, \infty)$, still converges pointwise almost everywhere to f. This provides another affirmative answer in the Q2 category. However, the case $p = 1$ is distinctly different: Kolmogorov constructed (aged 19!) an example of a function in $L^1(-\pi, \pi)$ whose Fourier series diverges almost everywhere. He went on to improve this result and construct an example in L^1 diverging everywhere. So the Fourier series of a function in L^1 can behave extremely badly, but as soon as you go above $p = 1$ you get convergence almost everywhere to the original function.

5 Spectral theory

Your B4.1 course introduced spectral theory in Banach spaces, and following the theme of B4.2 we revisit this topic in the context of Hilbert spaces. In particular, we introduced the notion of the adjoint of an operator on a Hilbert space, and it is both useful and interesting to see how this notion interacts with spectral theory. (As we remarked earlier, the adjoint of an operator on a Banach space can also be defined, but things are slightly more complicated in this case as we don't have the identification between a space and its dual).

Recall that if $(X, ||\cdot||_X)$ is a normed vector space over $\mathbb{C}$ and $T \in L(X)$, then the *resolvent set* $\rho(T)$ is

$$\rho(T) := \{\lambda \in C : T - \lambda\,\mathrm{Id} \text{ is invertible}\}. \tag{5.1}$$

Its complement

$$\sigma(T) := \mathbb{C} \backslash \rho(T) = \{\lambda \in C : T - \lambda\,\mathrm{Id} \text{ is not invertible}\} \tag{5.2}$$

is then called the *spectrum* of T. We clarify that by 'invertible', we mean that $(T - \lambda\,\mathrm{Id})^{-1}$ exists as a linear operator (which occurs if and only if $T - \lambda\,\mathrm{Id}$ is both injective and surjective) *and* belongs to $L(X)$ (i.e. it is bounded). Thus there are three ways in which $T - \lambda\,\mathrm{Id}$ can fail to be invertible, or equivalently, three scenarios which could cause $\lambda \in \sigma(T)$: injectivity, surjectivity, or boundedness of the inverse (if it exists) must fail. If only the latter fails, we say that $T - \lambda\,\mathrm{Id}$ is *algebraically invertible*. You showed in B4.1 that an algebraically invertible operator T is fully invertible if and only if there exists some $\delta > 0$ for which

$$||Tx||_X \geq \delta ||x||_X \quad \text{for all } x \in X. \tag{5.3}$$

Therefore, $\lambda \in \sigma(T)$ if and only if at least one of the following holds:

1. $T - \lambda\,\mathrm{Id}$ is not injective
2. $T - \lambda\,\mathrm{Id}$ is not surjective
3. There exists no $\delta > 0$ for which $||Tx - \lambda x||_X \geq \delta ||x||_X$ for all $x \in X$.

Remark 5.1. Perhaps the condition (5.3) will become more intuitive with the following analogy: suppose $f : \mathbb{R} \to (0, \infty)$ is continuous and bounded, so that there exists a constant $C < \infty$ such that $0 < f(x) \leq C$ for all $x \in \mathbb{R}$. Since $f > 0$, the function $1/f$ is well-defined, but is it bounded? Since f is bounded above, clearly $1/f$ is bounded below (also by zero). But f may get arbitrarily close to 0, in which case $1/f$ can be arbitrarily large. However, if there exists some $\delta > 0$ such that $f(x) \geq \delta$ for all $x \in \mathbb{R}$, then $1/f \leq 1/\delta$, so $1/f$ is indeed bounded above as well. Similarly, if T is algebraically invertible, then T^{-1} exists as a linear operator but is bounded if and only if $||Tx||_X \geq \delta ||x||_X$ for all $x \in X$.

Now, $T - \lambda\,\mathrm{Id}$ is not injective if and only if its kernel is non-trivial, i.e. there exists $x \in X \backslash \{0\}$ such that $Tx - \lambda x = 0$, i.e. $Tx = \lambda x$. The elements $\lambda \in \sigma(T)$ arising in this way - that is, through lack of injectivity of $T - \lambda\,\mathrm{Id}$ - are called the *eigenvalues* of T, and together they make up the *point spectrum* $\sigma_P(T)$. Clearly

$$\sigma_P(T) \subseteq \sigma(T). \tag{5.4}$$

If $x \in X \backslash \{0\}$ and $\lambda \in \mathbb{C}$ are such that $Tx = \lambda x$, we call x an *eigenvector* corresponding to the eigenvalue λ.

The rest of $\sigma(T)$ consists of $\lambda \in \mathbb{C}$ for which $T - \lambda\,\mathrm{Id}$ *is* injective, i.e. $\ker(T - \lambda\,\mathrm{Id}) = \{0\}$. We divide these remaining elements into two further groups: those $\lambda \in \mathbb{C}$ for which $\mathrm{Im}(T - \lambda\,\mathrm{Id})$ is not dense in X, and those for which $\mathrm{Im}(T - \lambda \,\mathrm{Id})$ is dense in X. We call these the *residual spectrum* $\sigma_r(T)$ and *continuous spectrum* $\sigma_c(T)$ of T, respectively. By construction, we have the disjoint union

$$\sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T). \tag{5.5}$$

Remark 5.2. If X is a Banach space, then $T \in L(X)$ is algebraically invertible if and only if it is invertible (this is just the inverse mapping theorem, see Remark 2.3). It must therefore be the case if $\lambda \in \sigma_c(T)$, then $\mathrm{Im}(T - \lambda \,\mathrm{Id})$ must be a *proper* dense subset of X, otherwise $\mathrm{Im}(T - \lambda \,\mathrm{Id}) = X$ and $T - \lambda \,\mathrm{Id}$ would then be injective *and* surjective. This would contradict $\lambda \in \sigma(T)$, since $T - \lambda \,\mathrm{Id}$ would be algebraically invertible and thus invertible.

Clearly if 1) holds in the above then so does 3). Thus we have an inclusion

$$\sigma_p(T) \subseteq \sigma_{ap}(T) \subseteq \sigma(T) \tag{5.6}$$

where $\sigma_{ap}(T)$ is the set of $\lambda \in \mathbb{C}$ for which 3) holds, called the *approximate point spectrum*. In other words, $\lambda \in \sigma_{ap}(T)$ if and only if there exists a sequnce $\{x_n\} \subset X$ satisfying $||x_n||_X = 1$ for all n and such that $||Tx_n - \lambda x_n||_X \to 0$ as $n \to \infty$.

Lemma 5.1.2. *Let X be a Banach space and $T \in L(X)$. Then $\sigma_c(T) \subseteq \sigma_{ap}(T)$.*

Remark 5.3. When X is a Hilbert space, Lemma 5.1.2 can be proved by appealing to the weak sequential compactness of the unit ball in X (see Theorem 3.3.2).

Propositions 5.2.1 and 5.2.2 contain some useful properties relating spectra and adjoints and make standard bookwork exam questions, so make sure you learn them. They are also very useful for manipulating identities in proofs, as you may have realised in problem sheet 4. We do not list them here.

In B4.1 Theorem 8.7 you showed that if $(X, ||\cdot||_X)$ is a Banach space, $T \in L(X)$ and $T' \in L(X^*)$ is the dual operator (a.k.a. transpose operator) defined by $(T'f)(x) = f(Tx)$, then

$$\sigma(T) = \sigma_{ap}(T) \cup \sigma_p(T'). \tag{5.7}$$

Analogously, in a complex Hilbert space one has the following:

Theorem 5.2.3. *Let X be a complex Hilbert space, $T \in L(X)$ and $T^* \in L(X)$ its adjoint. Then*

$$\sigma(T) = \sigma_{ap}(T) \cup \overline{\sigma_p(T^*)} \tag{5.8}$$

where $\overline{\sigma_p(T^*)} = \{\bar{\lambda} : \lambda \in \sigma_p(T^*)\}$.

The rest of the chapter is an exposition of some spectral results for self-adjoint operators on a Hilbert space. We collect them together here:

Theorem 5.2.4-5.2.6. *If X is a complex Hilbert space and $T \in L(X)$ a self-adjoint operator, then*

1. $\sigma(T) \subset \mathbb{R}$,
2. $\sigma_r(T) = \emptyset$,
3. *If v_1, v_2 are eigenvectors corresponding to dinstinct eigenvalues, then $\langle v_1, v_2 \rangle_X = 0$,*
4. $\mathrm{rad}(\sigma(T)) = ||T||$ *where* $\mathrm{rad}(\sigma(T)) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$, *is the spectral radius of* $\sigma(T)$,
5. $\sigma(T) \subseteq [a, b] \subset \mathbb{R}$ *where* $a = \inf_{||x||=1} \langle x, Tx \rangle$ *and* $b = \sup_{||x||=1} \langle x, Tx \rangle$, *and* $a, b \in \sigma(T)$.

Proposition 5.2.7. *If X is a complex Hilbert space and $T \in L(X)$ is unitary, then $|\lambda| = 1$ for all $\lambda \in \sigma(T)$.*

6 Unofficial exam-style question

Question 1.
Let $(X, ||\cdot||)$ be a normed vector space and X^* its dual space of bounded linear functionals on X.

a) (2 marks) Define what it means for a sequence $\{x_n\} \subset X$ to converge weakly in X.

b) i) (3 marks) State (without proof) the extended hyperplane separation theorem.

ii) (3 marks) Suppose $\varphi, \varphi_1, \dots, \varphi_k \in X^*$ are such that $\varphi_i(x) = 0$ for $i = 1, \dots, k$ implies $\varphi(x) = 0$. By considering $F : X \to \mathbb{R}^{k+1}$,

$$F(x) := \big(\varphi(x), \varphi_1(x), \dots, \varphi_k(x)\big),$$

or otherwise, show that there exist constants $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that

$$\varphi = \sum_{i=1}^{k} \lambda_i \varphi_i.$$

Now suppose $(X, ||\cdot||)$ is an infinite-dimensional Banach space. We wish to show (by contradiction) that the weak topology on X is not metrizable. You may assume the following without proof:

Hint: If $\epsilon > 0$ and $f_1, \dots, f_k \in X^*$, then

$$V := \{x \in X : |f_i(x)| < \epsilon \; \forall i = 1, \dots, k\}$$

is a neighbourhood of $0 \in X$ in the weak topology on X, and such sets V form a basis of neighbourhoods of $0 \in X$ in the weak topology as we vary over $\epsilon > 0$, $k < \infty$ and $f_i \in E^*$.

c) i) (6 marks) Suppose d is a metric on X inducing the weak topology, and suppose V_k is a neighbourhood of $0 \in X$ in the weak topology such that

$$V_k \subseteq \left\{x \in X : d(x, 0) < \frac{1}{k}\right\}.$$

By considering V_k and the hint above, or otherwise, prove that there exists a sequence $\{f_n\} \subset X^*$ such that every $g \in X^*$ can be written as a finite linear combination of the f_n's.

ii) (6 marks) State and prove the Baire Category Theorem.

iii) (3 marks) Deduce from c)i) and c)ii) that X^* is finite-dimensional.

iv) (2 marks) Obtain a contradiction from c)iii) and thus conclude that the weak topology on X is not metrizable.