B4.2 Functional Analysis II

Gregory Seregin

Hilary 2023

This course is a continuation of B4.1 Functional Analysis I.

This set of lecture notes build upon and expand previous lecture notes by Hilary Priestley and Luc Nguen who taught the course in previous years. The following literature was also used (either for this set of notes, or for my predecessors'):

B.P. Rynne and M.A. Youngson, *Linear Functional Analysis*, Springer SUMS, 2nd edition, 2008.

E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley, revised edition, 1989.

N. Young, *An Introduction to Hilbert Space*, Cambridge University Press, 1988.

E.M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration & Hilbert Spaces*, Princeton Lectures in Analysis III, 2005.

M. Reed and B. Simon, *Methods of Modern Mathematical Physics Vol. I. Functional Analysis*, Academic Press, 1980.

H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer, 2011.

P.D. Lax, *Functional Analysis*, Wiley, 2002.

R.L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Dekker, 1977.

Contents

Chapter 1

Hilbert Spaces

1.1 Inner product

Definition 1.1.1. *A linear vector space X over scalar field $\mathbb{C}$ (or $\mathbb{R}$) is called Inner Product Space (ISP) if there exists a function $< \cdot, \cdot >: X \times X \to \mathbb{C}$ (or $\mathbb{R}$) having the following properties:*
(i) $< x, y >= \overline{< y, x >}$ *(or* $< x, y >= < y, x >$*)* $\forall x, y \in X$
(ii) $< \lambda x, y >= \lambda < x, y >$ $\forall \lambda \in \mathbb{C}$ *(or* $\mathbb{R}$*)* $\forall x, y \in X$
(iii) $< x + y, z >=< x, z > + < y, z >$ $\forall x, y, z \in X$
(iv) $< x, x >\in \mathbb{R}_+ := \{t \in \mathbb{R} : \ t \geq 0\}$, $< x, x >= 0 \Leftrightarrow x = 0$.

Function $< \cdot, \cdot >$ is called inner (scalar) product. Properties (i)-(iv) are called axioms of inner product.

The inner product $\langle \cdot, \cdot \rangle$ generates a norm, denoted by $\| \cdot \|$ and called induced or associated norm, as follows:

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

It should be clear that the positivity of the norm $\| \cdot \|$ follows from the positivity property (iv), and the homogeneity of $\| \cdot \|$ follows from properties (i) and (ii). To prove the triangle inequality, we need:

Theorem 1.1.2 (Cauchy-Schwarz inequality)**.** *For $x, y \in X$,*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Equality holds if and only if x and y are linearly dependent.

Proof. If $y = 0$, the conclusion is clear. Assume henceforth that $y \neq 0$. We have $\|x - ty\|^2 \geq 0$ for all $t \in \mathbb{C}$. Let us pick up t so that

$$t = \frac{< x, y >}{\|y\|^2}.$$

Then, applying the axioms of inner product, we find

$$\|x - ty\|^2 = \|x\|^2 - \frac{| < x, y > |^2}{\|y\|^2}.$$

Since the left hand side of the last identity is non-negative, we establish the required Cauchy-Scwarz inequality. Assume now that we have the identity in the Cauchy-Schwarz inequality, then from the above formula we get $\|x - ty\|^2 = 0$ and thus $x - ty = 0$.

$\square$

Using axioms of inner product and definition of the induced norm $\| \cdot \|$, it is not so difficult to derive the so-called parallelogram law:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \text{ for all } x, y \in X. \tag{1.1}$$

It is a fact that if a norm satisfies the parallelogram law (1.1), then it comes from an inner product, which can be retrieved from the norm using polarisation:

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

for real scalar field and

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

for complex scalar field.

Definition 1.1.3. *An inner product space is called a Hilbert space, if it is Banach space with respect to the induced norm.*

Given an inner product space, one can complete it with respect to the induced norm. Since the inner product is a continuous function on its factors, it can be extended to the completed space. The completed space is therefore a Hilbert space.

Example 1.1.4. *The space $\mathbb{C}^n$ or $\mathbb{R}^n$ is a Hilbert space with the standard inner product*

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k\, \bar{y}_k.$$

Example 1.1.5. *The space $\ell^2 = \{(x_1, x_2, \ldots) = (x_n) : \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ is a Hilbert space with the inner product*

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x_k\, \bar{y}_k.$$

Example 1.1.6. *The space $C[0,1]$ of continuous functions on the interval $[0,1]$ is an incomplete inner product space with the inner product*

$$\langle f, g\rangle = \int_0^1 f\, \bar{g}\, dx.$$

Example 1.1.7. *Let (E,μ) be a measure space, e.g. E is a subset of $\mathbb{R}^n$ and μ is the Lebesgue measure. The space $L^2(E,\mu)$ of all complex-valued square integrable functions is a Hilbert space with the inner product*

$$\langle f, g\rangle = \int_E f\, \bar{g}\, d\mu.$$

The completeness of $L^2(E,\mu)$ is a special case of the Riesz-Fischer theorem on the completeness of the Lebesgues space $L^p(E,\mu)$.

Example 1.1.8. *A closed subspace of a Hilbert space is a Hilbert space.*

Example 1.1.9 (Bergman space)**.** *Let $\mathbb{D}$ be the open unit disk in $\mathbb{C}$. The space $A^2(\mathbb{D})$ consists of all functions which are square integrable and holomorphic in $\mathbb{D}$ is a closed subspace of $L^2(\mathbb{D})$ and is thus a Hilbert space.*

Example 1.1.10 (Hardy space)**.** *The space $H^2(\mathbb{T})$ of all functions $f \in L^2(-\pi,\pi)$ whose Fourier series are of the form $\sum_{n\geq 0} a_n\, e^{inx}$ is a closed subspace of $L^2(-\pi,\pi)$ and is thus a Hilbert space.*

Example 1.1.11 (Sobolev space $H^1(a,b)$)**.** *We say that $u \in H^1(a,b)$ if $u \in L^2(a,b)$ and there exists a function $v \in L^2(a,b)$ such that*

$$u(x) = A + \int_a^x v(y)\, dy \tag{1.2}$$

for some constant A and for almost all $x \in (a,b)$.

Note that by (1.2), any $u \in H^1(a,b)$ has a continuous representation in $[a,b]$, since

$$|u(x) - u(\tilde{x})| = \int_x^{\tilde{x}} v(y)\, dy \le |x - \tilde{x}|^{1/2}\, \|v\|_2.$$

Also, for any given $u \in H^1(a,b)$, there is only one function v satisfying (1.2). Indeed, if there are two constants A_1, A_2 and two functions v_1, v_2 satisfying (1.2) then

$$\int_x^{\tilde{x}} [v_1(y) - v_2(y)]\, dy = A_2 - A_1 \text{ for all } x, y \in [a,b].$$

Now, since for almost all $x \in (a,b)$, it holds that

$$\lim_{\delta \to 0} \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} [v_1(y) - v_2(y)]\, dy = v_1(x) - v_2(x),$$

the above implies that $A_2 = A_1$ and $v_1 = v_2$ a.e. in (a,b).

Next, observe that

$$\lim_{\delta \to 0} \frac{u(x+\delta) - u(x)}{\delta} = \lim_{\delta \to 0} \frac{1}{\delta} \int_x^{x+\delta} v(y)\, dy = v(x) \text{ for almost all } x \in (a,b),$$

i.e. a function $u \in H^1(a,b)$ is almost everywhere differentiable in (a,b) and the derivative of u is equal the function v in (1.2) almost everywhere in (a,b). It then makes sense to call v the 'weak' or 'generalised' derivative of u and write $v = u'$. It should be clear that if u is C^1, then v is indeed the classical derivative of u. In addition, we note the integration by parts formula: if $\varphi \in C_0^1([0,1])$, then, by Fubini's theorem,

$$\int_0^1 w(x)\varphi'(x)\, dx = \int_0^1 \int_0^x v(y)\, \varphi'(x)\, dy\, dx = \int_0^1 v(y) \int_y^1 \varphi'(x)\, dx\, dy$$
$$= -\int_0^1 v(y)\, \varphi(y)\, dy.$$

Theorem 1.1.12 (Not for examination)**.** *The space $H^1(a,b)$ is a Hilbert space with the inner product*

$$\langle u, v \rangle = \int_a^b (u\bar{v} + u'\,\bar{v}')\, dx.$$

1.2 Orthogonality

Definition 1.2.1. *Two vectors x and y in an inner product space X are said to be orthogonal if $\langle x, y\rangle = 0$.*

Definition 1.2.2. *Let Y be a subset of an inner product space X. We define $Y^\perp$ as the space of all vectors $v \in X$ which are orthogonal to Y, i.e. $\langle v, y\rangle = 0$ for all $y \in Y$.*

When Y is a subspace of X, $Y^\perp$ is called the orthogonal complement of Y in X.

Proposition 1.2.3. *Let Y be a subset of an inner product space X. Then*

(i) $Y^\perp$ is a closed subspace of X.

(ii) $Y \subset Y^{\perp\perp}$.

(iii) If $Y \subset Z \subset X$, then $Z^\perp \subset Y^\perp$.

(iv) $(\overline{\mathrm{span}\, Y})^\perp = Y^\perp$.

(v) If Y and Z are subspaces of X such that $X = Y + Z$ and $Z \subset Y^\perp$, then $Y^\perp = Z$.

Proof. Exercise. □

Theorem 1.2.4 (Closest point in a closed convex subset)**.** *Let K be a non-empty closed convex subset of a Hilbert space X. Then, for every $x \in X$, there is a unique point $y \in K$ which is closer to x than any other points of K, i.e.,*

$$\|x - y\| = \inf_{z\in K} \|x - z\|.$$

Proof. Let

$$d = \inf_{z\in K} \|x - z\| \geq 0$$

and $y_n \in K$ be a minimizing sequence, i.e.

$$\lim_{n\to\infty} d_n = d, \qquad d_n = \|x - y_n\|.$$

Applying the parallelogram law (1.1) to $\frac{1}{2}(x - y_n)$ and $\frac{1}{2}(x - y_m)$ yields

$$\left\|x - \frac{1}{2}(y_n + y_m)\right\|^2 + \frac{1}{4}\|y_n - y_m\|^2 = \frac{1}{2}(d_n^2 + d_m^2).$$

Since K is convex, $\frac{1}{2}(y_n + y_m) \in K$ and so $\left\| x - \frac{1}{2}(y_n + y_m) \right\| \geq d$. This and the above implies that (y_n) is a Cauchy sequence. Let y be the limit of this sequence, which belongs to K as K is closed. We then have by the continuity of the norm that $\|x - y\| = \lim \|x - y_n\| = d$, i.e. y minimizes the distance from x.

That y is the unique minimizer follows from the same reasoning above. If y' is also a minimizer, we apply the parallelogram law to $\frac{1}{2}(x - y)$ and $\frac{1}{2}(x - y')$ to obtain

$$d^2 + \frac{1}{4}\|y - y'\|^2 \leq \left\| x - \frac{1}{2}(y + y') \right\|^2 + \frac{1}{4}\|y - y'\|^2 = \frac{1}{2}(\|x - y\|^2 + \|x - y'\|^2) = d^2.$$

This implies that $y = y'$. □

Theorem 1.2.5 (Projection theorem)**.** *If Y is a closed subspace of a Hilbert space X, then Y and $Y^\perp$ are complementary subspaces: $X = Y \oplus Y^\perp$, i.e. every $x \in X$ can be decomposed uniquely as a sum of a vector in Y and in $Y^\perp$.*

Proof. Certainly $Y \cap Y^\perp = \{0\}$. It remains to show that $X = Y + Y^\perp$.

Take any $x \in X$ and, since Y is a non-empty closed convex subset of X, there is a point $y_0 \in Y$ which is closer to x than any other points of Y by Theorem 1.2.4. To conclude, we show that $x - y_0 \in Y^\perp$. Indeed, for all $y \in Y$ and $t \in \mathbb{R}$, we have

$$\|x - y_0\|^2 \leq \|x - \underbrace{(y_0 - ty)}_{\in Y}\|^2 = \|x - y_0\|^2 + 2t \operatorname{Re} \langle x - y_0, y \rangle + t^2 \|y\|^2.$$

It follows that $2t \operatorname{Re} \langle x - y_0, y \rangle + t^2 \|y\|^2 \geq 0$ for all $t \in \mathbb{R}$. This implies $\operatorname{Re} \langle x - y_0, y \rangle = 0$. This concludes the proof if the scalar field is real.

If the scalar field is complex, we proceed as before with t replaced by it to show that $\operatorname{Im} \langle x - y_0, y \rangle = 0$. □

Caution. *It follows from Theorem 1.2.5 that every closed subspace of a Hilbert space has a closed complement. This is not true for all Banach spaces.*

Corollary 1.2.6. *If Y is a closed subspace of a Hilbert space X, then $Y = Y^{\perp\perp}$.*

Definition 1.2.7. *The closed linear span of a set S in a Hilbert space X is the smallest closed linear subspace of X containing S, i.e. the intersection of all such subspaces.*

Proposition 1.2.8. *Let S be a set in a Hilbert space X. Then the closed linear span of S is $S^{\perp\perp}$.*

Proof. Exercise. $\square$

Definition 1.2.9. *A subset S of a Hilbert space X is called an orthonormal set if $\|x\| = 1$ for all $x \in S$ and $\langle x, y\rangle = 0$ for all $x \neq y \in S$.*

S is called an orthonormal basis (or a complete orthonormal set) for X if S is an orthonormal set and its closed linear span is X.

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

Proof. We will only give a proof in the case when the Hilbert space X under consideration is separable, i.e. it contains a countable dense subset S. The proof in the more general case draws on more sophisticated arguments such as Zorn's lemma.

Label the elements of S as $y_1, y_2, \ldots$ Applying the Gram-Schmidt process[1] we obtain an orthonormal set $B = \{x_1, x_2, \ldots\}$ such that, for every n, the span of $\{x_1, \ldots, x_n\}$ contains $y_1, \ldots, y_n$. As $\bar{S} = X$, this implies that $X = \overline{\operatorname{span} B}$, and so X is the closed linear span of B. $\square$

Theorem 1.2.11 (Pythagorean theorem)**.** *Let X be a Hilbert space and $S = \{x_1, x_2, \ldots, x_m\}$ be a finite orthonormal set in X. For every $x \in X$, there holds*

$$\|x\|^2 = \sum_{n=1}^{m} |\langle x, x_n\rangle|^2 + \Big\| x - \sum_{n=1}^{m} \langle x, x_n\rangle x_n \Big\|^2.$$

The proof of this is a direct computation and is omitted. An immediate consequence is:

Lemma 1.2.12 (Bessel's inequality)**.** *Let X be a Hilbert space and $S = \{x_1, x_2, \ldots\}$ be an orthonormal sequence in X. Then, for every $x \in X$, there holds*

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

[1] The Gram-Schmidt process is usually applied to a set of finitely many linearly independent vectors yielding an orthogonal basis of the same cardinality. In our setting, we will lose the latter property as the vectors y_i's are not necessarily linearly independent.

Theorem 1.2.13. *Let X be a Hilbert space and $S = \{x_1, x_2, \ldots\}$ be an orthonormal sequence in X. Then the closed linear span of S consists of vectors of the form*

$$x = \sum_{n=1}^{\infty} a_n \, x_n \tag{1.3}$$

where the sequence of scalar $(a_1, a_2, \ldots)$ belongs to ℓ^2. The sum in (1.3) converges in the sense of the Hilbert space norm. Furthermore

$$\|x\|^2 = \sum_{n=1}^{\infty} |a_n|^2 \qquad (\textit{Parseval's identity})$$

and

$$a_n = \langle x, x_n \rangle.$$

Proof. Let Y denotes the closed linear span of S. Denote by $\bar{Y}$, the set of $x \in X$ such that

$$x = \sum_{k=1}^{\infty} a_k x_k$$

and

$$\|x - \sum_{k=1}^{n} a_k x_k\| \to 0$$

as $n \to \infty$. Let us discuss properties of $x \in \bar{Y}$. Indeed, for any $i \in \mathbb{N}$, we have

$$< x - \sum_{k=1}^{n} a_k x_k, x_i > \to 0$$

as $n \to \infty$ and thus $a_i =< x, x_i >$. Now, from Bessel's inequality, we find $a = (a_1, a_2, ..., a_n, ...) \in \ell^2$ and the Parseval's identity can be easly derived from Pythagorean theorem. Moreover, from Parseval's identity and completeness of the space ℓ^2, it follows that the set $\bar{Y}$ is closed in X (explain why). So, $Y \subseteq \bar{Y}$. Since, in our case, Y is simply the closure of span of S, we get $\bar{Y} \subseteq Y$ and thus $\bar{Y} = Y$.

$\square$

1.3 Linear functionals

If X is a Hilbert space, and $x \in X$ is fixed, then $\langle y, x\rangle = \ell(y)$ is a linear functional of y, i.e. ℓ maps X linearly into $\mathbb{R}$ or $\mathbb{C}$. Furthermore, ℓ is bounded, thanks to the Cauchy-Schwarz inequality, and so $\ell \in X^*$. It turns out that all bounded linear functionals on a Hilbert space arise this way:

Theorem 1.3.1 (Riesz representation theorem)**.** *Let X be a real (or complex) Hilbert space and $\ell : X \to \mathbb{R}$ (or $\mathbb{C}$) be a bounded linear functional. Then ℓ is of the form*

$$\ell(y) = \langle y, x\rangle \text{ for all } y \in X$$

for some $x \in X$. Furthermore, the point x is uniquely determined and $\|x\| = \|\ell\|_*$.

Remark 1.3.2. *In the case of real Hilbert spaces, the above statement means that there exists an isometric isomorphism $\pi : X \to X^*$ such that $(\pi x)(y) = \langle y, x\rangle$ for all $x, y \in X$ and $\|\pi x\|_* = \|x\|$. So the spaces X and X^* are topologically equivalent, i.e. they are the same up to isometric isomorphism. It is notated as $X^* \cong X$ or even just $X^* = X$.*

Proof. If $\ell = 0$, then $x = 0$. Assume henceforth that $\ell \not\equiv 0$. Let Y be the kernel of ℓ. Then Y is a closed subspace of X. By Theorem 1.2.5, $X = Y \oplus Y^\perp$.

Since $Y^{\perp\perp} = Y$ is a strict subspace of X (as $\ell \not\equiv 0$), $Y^\perp$ contains a non-zero element, say $y^\perp$. Note that $\ell(y^\perp) \neq 0$. Then for any $z \in X$, we have

$$z - \frac{\ell(z)}{\ell(y^\perp)} y^\perp \in Y = \operatorname{Ker} \ell$$

Taking inner product with $y^\perp$ yields

$$\langle z, y^\perp\rangle - \frac{\ell(z)}{\ell(y^\perp)} \|y^\perp\|^2 = 0 \text{ for all } z \in X.$$

In other words, x can be chosen as

$$x = \frac{\overline{\ell(y^\perp)}}{\|y^\perp\|^2} y^\perp.$$

The uniqueness is obvious.

For the last assertion, we note by the Cauchy-Schwarz inequality that $\ell(y) = \langle y, x\rangle \leq \|y\|\|x\|$ and so $\|\ell\|_* \leq \|x\|$. On the other hand, we have $\|x\|^2 = \langle x, x\rangle = \ell(x) \leq \|\ell\|_*\|x\|$ and so $\|x\| \leq \|\ell\|_*$. This completes the proof. □

By inspecting the proof, we obtain the following result which is true for more general vector spaces.

Lemma 1.3.3. *(i) The kernel of a non-trivial linear functional on a Banach space is a closed linear subspace of codimension one.*

(ii) If two linear functionals on a vector space have the same kernel space, then they are multiples of each other.

Proof. Exercise. □

1.4 Adjoint operators

Let X and Y be two Hilbert spaces and $\mathscr{B}(X, Y)$ denotes the Banach space of bounded linear operators from X to Y. If $X = Y$, we write $\mathscr{B}(X)$ in place of $\mathscr{B}(X, X)$.

Consider $A \in \mathscr{B}(X, Y)$. Then for fixed $y \in Y$, $\langle Ax, y\rangle_Y$ defines a bounded linear functional on X. Thus, by the Riesz representation theorem, there is some $A^*y \in X$ such that $\langle Ax, y\rangle_Y = \langle x, A^*y\rangle_X$. The map $y \mapsto A^*y$ from Y to X is called the adjoint operator of A.

Proposition 1.4.1. *The adjoint operator satisfies the following properties.*

(i) $\langle Ax, y\rangle_Y = \langle x, A^*y\rangle_X$.

(ii) There is a unique operator A^* satisfying (i).

(iii) $A^* \in \mathscr{B}(Y, X)$.

(iv) $\|A\|_{\mathscr{B}(X,Y)} = \|A^*\|_{\mathscr{B}(Y,X)}$.

(v) $A^{**} = A$.

(vi) If $A, B \in \mathscr{B}(X, Y)$ and $a, b \in \mathbb{C}$, then $(aA + bB)^* = \bar{a}\, A^* + \bar{b}\, B^*$.

(vii) If $T \in \mathscr{B}(X, Y)$ and $S \in \mathscr{B}(Y, Z)$, then $(ST)^* = T^*S^*$.

If $X = Y$, we also have that

(viii) $I_X^* = I_X$.

(ix) $A \in \mathscr{B}(X)$ *is invertible if and only if* A^* *is invertible.*

Proof. Exercise. $\square$

Example 1.4.2. *Let $X = \mathbb{C}^n$, $Y = \mathbb{C}^m$ and $Ax = Mx$ where M is some $m \times n$ matrix. Then A^* is given by $A^*y = M^*y$ where M^* is the conjugate transpose of M.*

Example 1.4.3. *Let $X = Y = L^2(0,1)$ and A be the integral operator*

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

where $k : (0,1)^2 \to \mathbb{R}$ is a given bounded measurable function. Then A is a linear operator of $L^2(0,1)$ into itself. The adjoint operator A^*, which is also linear operator of $L^2(0,1)$ into itself, is given by

$$(A^*g)(x) = \int_0^1 \overline{k(y,x)}g(y)\,dy.$$

This is because, by Fubini's theorem,

$$\begin{aligned}\langle Af, g\rangle &= \int_0^1 \int_0^1 k(x,y)\,f(y)\,dy\,\bar{g}(x)\,dx \\ &= \int_0^1 f(y)\overline{\int_0^1 \overline{k(x,y)}\,g(x)\,dx}\,dy = \langle f, A^*g\rangle.\end{aligned}$$

Example 1.4.4. *Let $X = Y = \ell^2$ and R be the right-shift $R((x_1, x_2, \ldots)) = (0, x_1, x_2, \ldots)$. Then R^* is the left-shift $L((x_1, x_2, \ldots)) = (x_2, x_3, \ldots)$.*

Example 1.4.5. *Let $X = Y = L^2(\mathbb{R})$ and $h : \mathbb{R} \to \mathbb{C}$ be a bounded measurable function. Define the multiplication operator M_h by $M_h f(x) = h(x)f(x)$. Then $M_h \in \mathscr{B}(X)$ and $M_h^* = M_{\bar{h}}$.*

Definition 1.4.6. *Let X be a Hilbert space. An operator $T \in \mathscr{B}(X)$ is said to be self-adjoint if $T = T^*$.*

Lemma 1.4.7. *Let X be a Hilbert space.*

(i) If $T \in \mathscr{B}(X)$, then

$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle Tx, y\rangle| : \|x\| = \|y\| = 1\}.$$

(ii) If $T \in \mathscr{B}(X)$ and T is self-adjoint, then

$$\|T\|_{\mathscr{B}(X)} = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}.$$

Proof. The first assertion follows from the definition of the operator norm and the fact that

$$\|z\| = \sup_{\|y\|=1} |\langle y, z\rangle|.$$

Let us prove (ii). Set

$$K = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\} \leq \|T\|.$$

Fix some $\varepsilon > 0$. By (i), there are vectors x, y such that $\|x\| = \|y\| = 1$ and $|\langle Tx, y\rangle| > \|T\| - \varepsilon$. Replacing y with $e^{i\theta}y$ does not change $\|y\|$ and $|\langle Tx, y\rangle|$ but one can find θ so that that $|\langle Tx, y\rangle| = \mathrm{Re}\langle Tx, y\rangle$. This implies that

$$\begin{aligned} 4(\|T\| - \varepsilon) &\leq 4\mathrm{Re}\,\langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \\ &\leq K(\|x+y\|^2 + \|x-y\|^2) = K(2\|x\|^2 + 2\|y\|^2) = 4K, \end{aligned}$$

where we have used the parallelogram law in the second-to-last identity. The conclusion follows. ☐

Noting that A^*A is self-adjoint for any $A \in \mathscr{B}(X)$, we obtain the following result.

Proposition 1.4.8. *Let X be a Hilbert space and $A \in \mathscr{B}(X)$. Then*

$$\|A^*A\|_{\mathscr{B}(X)} = \|A\|^2_{\mathscr{B}(X)}.$$

In particular, if A is self-adjoint, then $\|A^2\|_{\mathscr{B}(X)} = \|A\|^2_{\mathscr{B}(X)}$.

We have the following result on the kernel and image of adjoint operators.

Proposition 1.4.9. *Let X and Y be Hilbert spaces and $A \in \mathscr{B}(X, Y)$. Then*

(i) $Ker\, A = (Im\, A^*)^\perp$.

(ii) $(Ker\, A)^{\perp} = \overline{Im\, A^*}$.

Proof. Exercise. $\square$

Theorem 1.4.10. *Let X be a Hilbert space and Y and Z are its closed subspaces such that $X = Y \oplus Z$. Let $P : X \to Y$ be the induced direct sum projection, i.e. $P(y + z) = y$. Then the following are equivalent.*

(i) $Z = Y^{\perp}$.

(ii) $P^* = P$.

(iii) $\|P\| \leq 1$ *(and in such case* $\|P\| = 1$ *or* $P \equiv 0$*).*

Proof. Exercise. $\square$

1.5 Unitary operators

Definition 1.5.1. *A linear operator between two Hilbert spaces is called unitary if it is isometric and surjective.*

Note that the requirement of linearity can be dropped after compositions with translation in view of the following result.

Proposition 1.5.2. *Let X and Y be Hilbert spaces. If $T : X \to Y$ is an isometry and $T(0) = 0$, then T is* <u>real</u> *linear.*

Proof. It suffices to show that $T(\frac{1}{2}(x+y)) = \frac{1}{2}(T(x)+T(y))$ for all $x, y \in X$.

If $x = y$, we are done. Suppose that $x \neq y$. Write $z = \frac{1}{2}(x + y)$. Then

$$\begin{aligned} \|T(x) - T(y)\| &= \|x - y\|, \\ \|T(z) - T(x)\| &= \|z - x\| = \frac{1}{2}\|y - x\|, \\ \|T(z) - T(y)\| &= \|z - y\| = \frac{1}{2}\|y - x\|. \end{aligned}$$

So

$$\|T(x) - T(y)\| = \|T(z) - T(x)\| + \|T(z) - T(y)\|,$$

and we have a situation where the triangle inequality is saturated. In view of the equality case of Cauchy-Schwarz' inequality, this is possible only if $T(x) - T(z)$ and $T(z) - T(y)$ are linearly dependent. Without loss of generality, we assume $T(x) - T(z) = \lambda(T(z) - T(y))$ for some (real or complex) scalar λ . As $\|T(x) - T(z)\| = \|T(z) - T(y)\| \neq 0$, we have $|\lambda| = 1$. Returning to the above equation, we then have

$$|\lambda + 1| = 2,$$

which then implies that $\lambda = 1$. We deduce that $T(z) = \frac{1}{2}(T(x) + T(y))$ as desired. $\square$

Remark 1.5.3. *In the above proof, we only use the strict sub-additivity property of the norm on an inner product space: $\|a - b\| + \|b - c\| = \|a - c\|$ if and only if a, b and c are co-linear.*

We have the following characterisation of isometric and unitary operators.

Proposition 1.5.4. *Let $T, U : X \to Y$ be bounded linear operators between Hilbert spaces.*

(i) The following are equivalent:

- *(a) T is isometric.*
- *(b) $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in X$.*
- *(c) $T^*T = I_X$.*

(ii) The following are equivalent:

- *(a) U is unitary.*
- *(b) $U^*U = I_X$ and $UU^* = I_Y$.*
- *(c) Both U and U^* are isometric.*

Proof. Exercise. $\square$

There is a well-known decomposition, referred to as the Wold decomposition, which asserts that every isometry of a Hilbert space can be expressed as a (direct) sum of a unitary operator and copies of the unilateral shift. We do not pursue this in the present notes.

Example 1.5.5. *(i) The right-shift operator on ℓ^2 is isometric but not unitary. The left-shift operator on ℓ^2 is not isometric.*

(ii) *A multiplication operator M_h is unitary on $L^2(\mathbb{R})$ if and only if $|h| = 1$ a.e.*

(iii) *If g is a non-negative and measurable function on $\mathbb{R}$, then the map $f \mapsto g^{1/2} f$ is isometric from $L^2(\mathbb{R}, g\,dt)$ to $L^2(\mathbb{R})$. It is unitary if and only if $g > 0$ a.e.*

Appendix: The Radon-Nikodym theorem

Here we will consider an application, due to von Neumann, of the Riesz representation to prove the so-called Radon-Nikodym theorem. For simplicity, let m denote the Lebesgue measure and $A \subset \mathbb{R}^n$ be a set of finite Lebesgue measure. Suppose μ be a finite measure defined on the σ-algebra consisting of measurable subsets of A. We say that μ is absolutely continuous with respect to m if every set that has zero Lebesgue measure has zero μ-measure.

Theorem 1.6.1 (Radon-Nikodym)**.** *Assume that μ is absolutely continuous with respect to m. Then $d\mu = g\,dm$ where g is some non-negative integrable function with respect to m:*

$$\mu(E) = \int_E g\,dm$$

for any measurable subset E in A.

Proof. Let X be the real Hilbert space $L^2(A, \mu + m)$ with the norm $\|f\|^2 = \int_A |f|^2 d(\mu + m)$. Define

$$\ell(f) = \int_A f\,dm \text{ for } f \in X.$$

By the Cauchy-Schwarz inequality, $\ell \in X^*$. Thus, by the Riesz representation theorem, we can find some $h \in X$ such that

$$\ell(f) = \int_A f\,h\,d(\mu + m) \text{ for all } f \in X.$$

This can be rewritten as

$$\int_A f(1 - h)\,dm = \int_A f\,h\,d\mu \text{ for all } f \in X. \tag{1.4}$$

We are now tempted to define $g = \frac{1-h}{h}$ and conclude. To this end, we need to show that

$$0 < h \leq 1 \text{ except on a set of measure zero.}$$

Let $F = \{h \leq 0\}$. Choosing $f = \chi_F$ in (1.4), we get

$$m(F) \leq \int_F (1-h)\, dm = \int_F h\, d\mu \leq 0$$

This implies $m(F) = 0$.

Let $G = \{h > 1\}$. We choose $f = \chi_G$ in (1.4) and get

$$0 \geq \int_G (1-h)\, dm = \int_G h\, d\mu \geq 0,$$

where the first inequality is strict if $m(G) > 0$. This implies that $m(G) = 0$. We have thus proved that $0 < h \leq 1$ except on a set of zero Lebesgue measure. Now setting $g = \frac{1-h}{h}$ and choosing $f = \frac{1}{h}$ in (1.4), we obtain the conclusion.

The latter arguments is not quite rigorous as we do not whether such a function is admissible. To make it legal, consider function $h_n(x) = \max\{h(x) > \frac{1}{n}\}$ and let $f_n = 1/h_n$, and $g_n = (1-h)/h_n$. Then we have

$$\int\limits_A g_n dm = \int\limits_A h/h_n \mu \leq \mu(A) < \infty.$$

Let us see what happens in $n \to \infty$. Since $0 \leq g_n \leq g_{n+1}$, we can use Beppo Levi theorem and state $g_n \to g$ a.e. in A so that

$$\int\limits_A g_n dm \to \int\limits_A g dm \leq \mu(A).$$

To get opposite inequality, we use the following:

$$\int\limits_A g dm \geq \lim_{n\to\infty} \mu(A_n) = \mu(\cup_{n=1}^{\infty} A_n),$$

where $A_n = \{x \in A : h(x) > 1/n\} \subset A_{n+1}$. But

$$A = (A \setminus \cup_{n=1}^{\infty} A_n) \cup (\cup_{n=1}^{\infty} A_n).$$

The first set Lebesgue measure zero and by absolute continuity of μ it is also μ-null set. The arguments are applicable to the function $f_n = \chi_E/h_n$, where E is a measurable subset of A and imply the final formula. $\square$

Chapter 2

Bounded linear operators: The Baire category theorem and its consequences

2.1 The Baire category theorem

Definition 2.1.1. *Let S be a subset of a metric space M.*

(i) We say that S is dense in M if $\bar{S} = M$.

(ii) We say that S is nowhere dense in M if $\bar{S}$ has empty interior.

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

Proof. Suppose that M is a complete metric space and suppose, by contradiction, that $M = \cup_{n=1}^{\infty} A_n$ where each A_n is nowhere dense. We will construct a Cauchy sequence (x_n) whose limit lies out of all these A_m's, which then leads to a contradiction.

Since A_1 is nowhere dense, $\bar{A}_1 \neq M$ and so $M \setminus \bar{A}_1$ is non-empty. Pick $x_1 \in M \setminus \bar{A}_1$.

Next, since $M \setminus \bar{A}_1$ is open, there is some closed ball $\bar{B}(x_1, r_1) \subset M \setminus \bar{A}_1$ with $r_1 < 1$. Clearly $B(x_1, r_1) \cap A_1 = \emptyset$. Since A_2 is nowhere dense, $\bar{A}_2 \not\supset B(x_1, r_1)$ and so there is some $x_2 \in B(x_1, r_1) \setminus \bar{A}_2$.

We then inductively choose balls $\bar{B}(x_n, r_n) \subset B(x_{n-1}, r_{n-1}) \setminus \bar{A}_n$ with $r_n < \frac{1}{2^{n-1}}$.

Now, the sequence (x_n) is Cauchy, since if $n, m \geq N$, then $x_n, x_m \in B(x_N, r_N)$ and so $d(x_m, x_n) \leq 2\, r_N \to 0$. Since M is complete, (x_n) converges to some $x \in M$. By the above, we have that $x \in \bar{B}(x_n, r_n) \subset B(x_{n-1}, r_{n-1}) \setminus \bar{A}_n$ for all n, which implies that $x \notin A_n$ for any n. This contradicts the assumption that M is the union of the A_n's. □

2.2 Principle of uniform boundedness

Theorem 2.2.1 (Uniform Boundedness Principle)**.** *Let X be a Banach space and Y be a normed vector space. Let $\mathscr{F} \subset \mathscr{B}(X, Y)$, i.e. $\mathscr{F}$ is a family of bounded linear operators from X into Y. If it holds for each $x \in X$ that the set $\{\|Tx\|_Y : T \in \mathscr{F}\}$ is bounded, then $\{\|T\|_{\mathscr{B}(X,Y)} : T \in \mathscr{F}\}$ is bounded.*

Loosely speaking, the principle of uniform boundedness asserts that a family of bounded linear operators is bounded if and only if it is point-wise bounded.

Proof. Let $A_n = \{x \in X : \|Tx\|_Y \leq n \text{ for all } T \in \mathscr{F}\}$. Then, by hypothesis, each $x \in X$ belongs to some A_n and so $X = \cup_{n=1}^\infty A_n$. By the Baire category theorem, there is some n_0 such that $A_{n_0} = \bar{A}_{n_0}$ (since the A_n's are closed) has non-empty interior. We can thus pick a ball $B(x_0, r_0) \subset A_{n_0}$.

Now suppose that $\|x\|_X < r_0$, we proceed to bound $\|Tx\|_Y$ for all $T \in \mathscr{F}$. By triangle inequality, we have $x_0 + x \in B(x_0, r_0)$ and so, by the definition of A_{n_0},

$$\|T(x_0 + x)\|_Y \leq n_0 \text{ for all } T \in \mathscr{F}.$$

We also have $\|T(x_0)\|_Y \leq n_0$ for all $T \in \mathscr{F}$. By triangle inequality again, we thus have

$$\|Tx\|_Y \leq \|T(x_0 + x)\|_Y + \|Tx_0\|_Y \leq 2n_0 \text{ for all } T \in \mathscr{F}.$$

Since x is chosen arbitrarily in $B(0, r_0)$, we thus conclude that $\|T\|_{\mathscr{B}(X,Y)} \leq 2n_0\, r_0^{-1}$ for all $T \in \mathscr{F}$. □

The principle of uniform boundedness has far reaching consequences. We illustrate here a few such.

Theorem 2.2.2. *Let X be a Hilbert space and $\mathscr{F}$ be a subset of $\mathscr{B}(X)$ such that $\sup_{T \in \mathscr{F}} |\langle Tx, y\rangle| < \infty$ for each $x, y \in X$. Then $\{\|T\| : T \in \mathscr{F}\}$ is bounded.*

Proof. By the principle of uniform boundedness, it suffices to show that, for each fixed $x \in X$, $\{\|Tx\| : T \in \mathscr{F}\}$ is bounded.

Fix an $x \in X$. Define $K_{T,x} \in X^*$ by $K_{T,x}(y) = \langle y, Tx\rangle$. Then, for each $y \in X$, $\{|K_{T,x}(y)| : T \in \mathscr{F}\}$ is bounded. The principle of uniform boundedness implies then $\{\|K_{T,x}\|_* : T \in \mathscr{F}\}$ is bounded. As $\|K_{T,x}\|_* = \|Tx\|$, we conclude the proof. □

Theorem 2.2.3 (Banach-Steinhaus theorem)**.** *Let X and Y be Banach spaces and consider a sequence $T_n \in \mathscr{B}(X, Y)$. The following statements are equivalent.*

(i) There exists $T \in \mathscr{B}(X, Y)$ such that, for every $x \in X$, $T_n x \to Tx$ as $n \to \infty$.

(ii) For each $x \in X$, the sequence $(T_n x)$ is convergent.

(iii) There is a constant M and a dense subset Z of X such that $\|T_n\| \le M$ and the sequence $(T_n z)$ is convergent for each $z \in Z$.

Proof. It is clear that (i) ⇒ (ii). That (ii) ⇒ (iii) is a direct application of the principle of uniform boundedness. Let us prove (iii) ⇒ (i).

We claim that, for every $x \in X$, $(T_n x)$ is Cauchy, and hence convergent. To see this, fix some $x \in X$, $\epsilon > 0$, and note that, for every $z \in Z$,

$$\begin{aligned}\|T_n x - T_m x\| &\le \|T_n z - T_m z\| + \|T_n(x - z)\| + \|T_m(x - z)\| \\ &\le \|T_n z - T_m z\| + 2M\|x - z\|.\end{aligned}$$

In particular, if we choose $z \in Z$ such that $\|x - z\| \le \frac{\epsilon}{4M}$ and choose N such that $\|T_n z - T_m z\| \le \frac{\epsilon}{2}$ for $n, m \ge N$, we obtain $\|T_n x - T_m x\| \le \epsilon$ for all $n, m \ge N$. This proves the claim

For $x \in X$, define Tx as the limit of $T_n x$. It is clear that T is linear. Also, we have

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \le \limsup_{n\to\infty} \|T_n\|\|x\| \le M\|x\|.$$

Thus T is a bounded linear operator on X. We have established (i). □

Consider a simple example of an application of Banach-Steinhaus theorem related to the so-called *mollification*. Let $g : \mathbb{R} \to \mathbb{R}$ be a non-negative infinitely differentiable function, which is zero outside $[-1, 1]$. Assume also that $\int\limits_{\mathbb{R}} g(t)dt = 1$. For positive $\varepsilon > 0$, introduce $g_\varepsilon(t) = 1/\varepsilon g(t/\varepsilon)$ and notice that $\int\limits_{\mathbb{R}} g_\varepsilon(t)dt = 1$ as well.

Given $f \in L^1(\mathbb{R})$, define

$$f_\varepsilon(t) := \int\limits_{\mathbb{R}} g_\varepsilon(t-s)f(s)ds.$$

Function f_ε is infinitely many times differentiable and approximates f in the following way: $\|f_\varepsilon - f\|_{L^1(\mathbb{R})}$ as $\varepsilon \to 0$. We going to prove this fact using the above Banach-Steinhaus theorem. To this end, we introduce the operator $T_\varepsilon : X \to X$, where $T_\varepsilon f = f_\varepsilon$ and $X = L^1(\mathbb{R})$. We have

$$\|T_\varepsilon f\|_X \leq \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} g_\varepsilon(t-s)|f(s)|dsdt \leq \|f\|_X$$

by the Fubini theorem and properties of g_ε. So, $\|T_\varepsilon\| \leq 1$.

Now, let $Z = C_0^1(\mathbb{R})$ be the space of all continuously differentiable functions that vanish outside a finite segment in $\mathbb{R}$. This set is dense in X and thus we need to show $\|T_\varepsilon f - f\|_X \to 0$ as $\varepsilon \to 0$ for any $f \in Z$. Repeating the above estimate, we find

$$\|T_\varepsilon f - f\|_X = \int\limits_{\mathbb{R}} \Big| \int\limits_{\mathbb{R}} g_\varepsilon(t-s)(f(s) - f(t))ds \Big| dt \leq$$

$$\leq \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} g_\varepsilon(t-s)|f(s) - f(t)|dsdt.$$

In the last integral, one can make the change of variables $s = t - \tau$ and then the change of order of integrations:

$$\|T_\varepsilon f - f\| \leq \int\limits_{-\varepsilon}^{\varepsilon} g_\varepsilon(\tau) \int\limits_{\mathbb{R}} |f(t-\tau) - f(t)|dtd\tau.$$

Using mean value theorem, we have $|f(t-\tau) - f(t)| = \tau|f'(t-\theta\tau)|$ with $0 < \theta < 1$. Then, assuming that $\varepsilon < 1$, it is easily verified that $\|T_\varepsilon f - f\| \leq c(f)\varepsilon$ (Explain why)

2.3 The open mapping theorem

Let $f : X \to Y$, with arbitray toplogical spaces X and Y. A mapping f is open if f maps open sets onto open sets.

Lemma 2.3.1. *Let X and Y be normes space and $T : X \to Y$ be linear. The following are equivalent:*
(i) T is open;
(ii) $T(B_X)$ is open (Here, B_X is a unit ball centred at the origin of the space X);
(iii) $T(B_X)$ contains non-empty ball $B_Y(0,\varepsilon)$.

Proof. Exercise. □

Theorem 2.3.2. *(Open Mapping Theorem) Let $T \in \mathcal{B}(X,Y)$, where X and Y are Banach space. Let T be onto. Then T is open.*

Proof. **Step I** Here, we are going to show that $T(X) = Y$, X is normed, Y is Banach, then $\overline{T(B_X)} \supset B_Y(0,\varepsilon)$ for some positive ε. Indeed, since

$$T(X) = \bigcup_n nT(B_X),$$

$nT(B_X)$ is not nowhere dense for some n and thus $T(B_X)$ is not nowhere dense. So, set $\overline{T(B_X)}$ has non-empty interior. It is check that it is convex and symmetric in the following sense: if $y \in \overline{T(B_X)}$, then $-y \in \overline{T(B_X)}$. This completes the proof of the statement. (Explain why)
Step II Now, let us assume that $T : X \to Y$ is linear and continuous, X is Banach, Y is normed, and $\overline{T(B_X)} \supset B_Y(0,\varepsilon)$ for some positive ε. We are going to prove that there is $\delta > 0$ such that $T(B_X) \supset B_Y(0,\delta)$.

Let $y \in \overline{T(B_X)}$. Then we can find $y_1 \in T(B_X)$ such that $\|y - y_1\|_Y < \varepsilon/2$. Since $y - y_1 \in B_Y(0,\varepsilon/2)$, $y - y_1 \in \overline{T(\frac{1}{2}B_X)}$ and thus we can find $y_2 \in T(\frac{1}{2}B_X)$ such that $\|y - y_1 - y_2\|_Y < \varepsilon/4$. Proceeding in this way, we find sequences y_n and x_n with the following properties:

$$\|y - y_1 - y_2 - ... - y_i\|_Y < \varepsilon/2^i, \quad y_i = Tx_i, \quad x_i \in \frac{1}{2^{i-1}}B_X, \quad i = 1,2,...,n$$

for all n. Since X is Banach and $\|x_n\|_X \le \frac{1}{2^{n-1}}$, the sequence $z_n = \sum_{i=1}^n x_i$ converges to $z \in X$. Moreover, $\|z_n\|_X \le \sum_{i=1}^n \frac{1}{2^{i-1}} \le 3$. Hence, we have

$\|y - Tz_n\|_Y < \varepsilon/2^n$ and passing to the limit, we show that $y = Tz$ with $z \in 3B_X$. So, one can take $\delta = \varepsilon/3$.

Now, the result follows from Step II and Lemma 2.3.1, see (iii) there.

□

An immediate consequence is:

Theorem 2.3.3 (Inverse mapping theorem)**.** *A bounded bijective linear operator of a Banach space onto another has a bounded inverse.*

Proof. Exercise. □

Example 2.3.4. *Let X be a Banach spaces with respect to two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ and suppose that there is a constant $C > 0$ such that $\|x\|_1 \le C\|x\|_2$ for all $x \in X$. Then the two norms are equivalent, i.e. there is a constant C' such that $\|x\|_2 \le C'\|x\|_1$ for all $x \in X$.*

Another consequence of the inverse mapping theorem is:

Theorem 2.3.5. *Let $T \in \mathscr{B}(X, Y)$ be a bounded linear operators between Hilbert spaces. Then TX is closed if and only if T^*Y is closed.*

Proof. Suppose that T^*Y is closed in X. Let $N := \overline{TX}$. Define, further, $S : X \to N$ by $Sx = Tx$ for $x \in X$ and let $S^* : N \to X$ be the adjoint operator. By Proposition 1.4.9, $N = \overline{\operatorname{Im} S} = (\operatorname{Ker} S^*)^\perp$, so $\operatorname{Ker} S^* = \{0\}$, i.e. S^* is injective. Let $P : Y \to N$ be a projector. We are going to prove that $S^*(Py) = T^*y$ for $y \in Y$. Indeed,

$$< Tx, y >_Y = < Sx, Py >_Y = < x, S^*Py >_X = < x, T^*y >_X .$$

From the above, it follows that $S^*P = T^*$, $S^*N = T^*Y =: M$, and $S^*n = T^*n$ for all $n \in N$. So, we can introduce the operator $V : N \to M$ by $Vn = S^*n$ for all $n \in N$. So, now, V is a bijection and M and N are Banach spaces. By IMT, there exists continuous operator $V^{-1} : M \to N$. Just by simple properties of adjoint operators, we get $(V^*)^{-1} = (V^{-1})^* \in \mathcal{B}(N, M)$.

Now, take any $y \in Y$ and any $n \in N$. There exists a uniques $m \in M$ such that $n = V^*m$. So, we have

$$\begin{aligned} < n, y >_Y = < V^*m, y >_Y = < V^*m, Py >_Y = < m, V^{**}Py >_X = \\ = < m, VPy >_X = < m, S^*Py >_X = < Sm, Py >_Y = \\ = < Sm, y >_Y = < Tm, y >_Y \end{aligned}$$

for $y \in Y$. Hence, $n = Tm$ and $N \subseteq TX$. So, TX is closed.

Converse can be proved by observation that $T^{**} = T$.

□

2.4 The closed graph theorem

Theorem 2.4.1 (Closed graph theorem)**.** *Let X and Y be Banach spaces and T be a linear operator from X into Y. Then T is bounded if and only if its graph*

$$\Gamma(T) = \{(x, y) \in X \times Y : y = Tx\}$$

is closed in $X \times Y$.

Proof. If T is bounded, it is easy to see that $\Gamma(T)$ is closed.

Conversely, assume that $\Gamma(T)$ is closed. Since T is linear, $\Gamma(T)$ is a closed subspace of $X \times Y$. In particular, it is a Banach space with the norm induced by the norm on $X \times Y$. Consider now the continuous maps $P_1 : \Gamma(T) \to X$ and $P_2 : \Gamma(T) \to Y$ defined by

$$P_1(x, Tx) = x \text{ and } P_2(x, Tx) = Tx.$$

It is clear that P_1 is a bijection. By the inverse mapping theorem, P_1 has a continuous inverse P_1^{-1}. The conclusion follows from the fact that $T = P_2 P_1^{-1}$. □

Example 2.4.2. *Let X be a Hilbert space and $T : X \to X$ be a linear mapping. If $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in X$, then T is bounded and so self-adjoint.*

Proof. As before, we show that if $x_n \to x$ and $Tx_n \to z$, then $z = Tx$. Indeed, for any $y \in X$, we have

$$\langle Tx, y\rangle = \langle x, Ty\rangle = \lim_{n\to\infty} \langle x_n, Ty\rangle = \lim_{n\to\infty} \langle Tx_n, y\rangle = \langle z, y\rangle,$$

which implies $z = Tx$. □

Example 2.4.3. *It is clear that if $h \in L^\infty(\mathbb{R})$, then the multiplication operator $f \mapsto hf =: M_h f$ defines a (bounded) linear operator from $L^1(\mathbb{R})$ into itself. The converse of this is true: If h is some measurable function such that $M_h f \in L^1(\mathbb{R})$ for all $f \in L^1(\mathbb{R})$, then $h \in L^\infty(\mathbb{R})$.*

Proof. By hypothesis M_h maps $L^1(\mathbb{R})$ into itself. We claim that M_h is bounded. To this end, we show that if $f_n \to f$ and $M_h f_n \to g$, then $g = M_h f$. First, $f_n \to f$ in L^1, there is a subsequence, say f_{n_j}, which converges to f a.e. It follows that $M_h f_{n_j} \to M_h f$ a.e. But since $M_h f_n \to g$ in L^1, this implies that $g = M_h f$. We conclude that M_h is a bounded operator on $L^1(\mathbb{R})$. In particular,

$$\int |M_h f|\, dx \le \|M_h\| \int |f|\, dx \text{ for all } f \in L^1(\mathbb{R}). \tag{2.1}$$

We claim that

$$|h| \le \|M_h\| \text{ a.e.}$$

To this end it suffices to show that the set $Z_\epsilon := \{x : |h(x)| > \|M_h\| + \epsilon\}$ has zero measure. Fix some $n > 0$. Taking $f = \chi_{Z_\epsilon \cap [-n,n]}$ in (2.1), we obtain

$$\|M_h\| \int_{Z_\epsilon \cap [-n,n]} dx \ge \int_{Z_\epsilon \cap [-n,n]} |h|\, dx \ge (\|M_h\| + \epsilon) \int_{Z_\epsilon \cap [-n,n]} dx.$$

This is possibly only if $Z_\epsilon \cap [-n,n]$ has zero measure. Since n is arbitrary, we conclude that Z_ϵ has zero measure. □

Now, let us discuss how the conclusion of CGT can fail if not all the conditions from it are met.

1. Consider the map T defined from sequences to sequences by

$$T(x_j) = (jx_j).$$

Let $X = l^1$ and Y be a subspace of X such that

$$Y = \{x = (x_j) \in X : \sum_{j=1}^{\infty} |jx_j| < \infty\}.$$

It is easy to see that, Y is dense (consider sequences with only finitely many non-zero coordinates) and proper since $(1/j^2)$ does not belong to X. So, the Y is not closed with respect the norm of X. Moreover, T is not bounded since $T(e_n) = n$. However, $G(T)$ is closed in $X \times X$, which can be verified directly using coordinate convergence. The reason, why CGT does work here, is exactly the fact that Y is not a Banach space with l^1-norm. Our case can be described as follows: T is unbounded operator in X whose graph is closed in $X \times X$ and whose domain of definition is dense in X.

2. Define the operator $T : Y \subset X \to X$ by $Tf = f'$, where $X = C([0,1])$ and $Y = C^1([0,1])$ is the subspace of $C([0,1])$ consisting of those f for which $f' \in C([0,1])$. Y is a proper dense subset of X. The operator T is not

bounded as there exists a sequence (f_n) in X such that $\|f_n\|_\infty = 1$ and $\|f_n'\|_\infty$ is unbounded. Nonetheless, $G(T)$ is closed. Indeed, let us $g_n \to g$ and $g_n \to G$ in X. It is easy to see that limit functions satisfy the following identity

$$g(t) - g(s) = \int\limits_s^t G(\tau) d\tau$$

for all $s, t \in [0, 1]$. This implies that $G = f'$.

The two above simple examples mimic problems arising in PDE's theory and in mathematical physics. One needs to study and develop the theory of unbounded linear operators with closed graph in Banach spaces.

Chapter 3

Weak convergence

3.1 Weak convergence

Definition 3.1.1. *A sequence (x_n) in a normed vector space X is said to converges weakly to $x \in X$ if*

$$\lim_{n\to\infty} \ell(x_n) = \ell(x) \text{ for all } \ell \in X^*.$$

This relation is indicated by a half arrow:

$$x_n \rightharpoonup x.$$

This weak convergence notion should be contrasted with strong convergence in the sense of norm: y_n converges strongly to y ($y_n \to y$) if

$$\lim_{n\to\infty} \|y_n - y\| = 0.$$

It should be clear that if a sequence converges strongly to x, then it also converges weakly to x. The converse is in general not true.

Example 3.1.2. *A sequence in a finite dimensional norm vector spaces converges weakly if and only if it converges strongly.*

Proposition 3.1.3. *Let X be a Hilbert space, and (x_n) be an orthonormal sequence. Then x_n tends weakly, but not strongly, to zero.*

Proof. Pick any bounded linear functional $\ell \in X^*$. By the Riesz representation theorem, there exists $y \in X$ such that

$$\ell(x) = \langle x, y\rangle \text{ for all } x \in X.$$

We thus need to show that

$$\lim_{n\to 0}\langle x_n, y\rangle = 0,$$

but this is a consequence of Bessel's inequality:

$$\sum_{n=1}^{\infty} |\langle x_n, y\rangle|^2 \leq \|y\|^2.$$

We have thus shown that $x_n \rightharpoonup 0$.

Lastly, note that strong convergence implies convergence in norm. Hence, since $\|x_n\| = 1$, we have that $x_n \not\to 0$. ☐

Example 3.1.4. *Let $X = C[0,1]$ and*

$$x_n(t) = \begin{cases} nt & \text{for } 0 \leq t \leq \frac{1}{n}, \\ 2 - nt & \text{for } \frac{1}{n} \leq t \leq \frac{2}{n}, \\ 0 & \text{for } \frac{2}{n} \leq t \leq 1. \end{cases}$$

Then x_n converges weakly, but not strongly, to zero.

Proof. It is clear that $x_n \not\to 0$ as $\|x_n\| = 1$. Fix some $\ell \in X^*$, we will show that $\ell(x_n) \to 0$. Arguing by contradiction, assume that there are infinitely many n such that

$$\ell(x_n) > \delta \text{ for some } \delta > 0. \tag{3.1}$$

Select inductively a sequence n_k such that the above holds together with $n_1 > 2$, $n_{k+1} > 2n_k$.

Define

$$y_K = \sum_{k=1}^{K} x_{n_k}.$$

We claim that

$$0 \leq y_K \leq 3 \text{ in } [0,1].$$

We proceed by induction on K. The claim is clear for $K = 1$. Assume that the claim is true for some $K \geq 0$.

Fix some $t \in [0,1]$. If $t \geq \frac{2}{n_{K+1}}$, we have $y_{K+1}(t) = y_K(t)$, so the claim is true by induction hypothesis. Assume that $t < \frac{2}{n_{K+1}}$. Then $t < \frac{1}{n_K}$ and so

$$y_{K+1}(t) \leq x_{n_{K+1}}(t) + \sum_{k=1}^{K} \frac{n_k}{n_K} \leq 1 + \sum_{k=1}^{K} 2^{k-K} \leq 3.$$

The claim is proved.

Now by (3.1), we have

$$K\delta < \ell(y_K) \leq \|\ell\|_* \|y_K\| \leq 3\|\ell\|_*,$$

which is absurd for large K. We therefore have $x_n \rightharpoonup 0$. □

Example 3.1.5 (Schur)**.** *If a sequence (x_n) converges weakly in ℓ^1, then it converges strongly.*

Proof. Exercise. □

Example 3.1.6. *Let X be a Hilbert space. If $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$, then $x_n \to x$.*

Proof. Exercise. □

3.2 Uniform boundedness of weakly convergent sequences

Let X be a normed space and let X^* be dual of X. Here, we will denote elements of X^* by π^*. It is a Banach space (regardless whether X is complete or not) with respect to the norm

$$\|\pi^*\|_* = \|\pi^*\|_{X^*} := \sup_{\|x\| \leq 1} |\pi^*(x)|.$$

We can go further and consider the space dual to X^*, i.e., space $X^{**} = (X^*)^*$. Its elements are denoted by π^{**} and the norm in this space is:

$$\|\pi^{**}\|_{**} = \sup_{\|\pi^*\|_* \leq 1} |\pi^{**}(\pi^*)|.$$

Remember that π^{**} is a bounded linear functional on X^*.

For any $x \in X$, we can consider the special functional on X^* of the following form $\pi^{**}_x(\pi^*) = \pi^*(x)$ for any $\pi^* \in X^*$. Let us find the norm of it. First, we have $|\pi^{**}_x(\pi^*)| \leq \|\pi^*\|_* \|x\|$ and thus $\|\pi^{**}_x\|_{**} \leq \|x\|$. To get the opposite inequality, we are going to use a result in B4.1 (which is a consequence of the Hahn-Banach theorem), there is some $\pi^*_0 \in X^*$ such that $\|\pi^*_0\|_* = 1$ and $\pi^*_0(x) = \|x\|$. From this it follows that $\|\pi^{**}_x\|_{**} \geq |\pi^{**}_x(\pi^*_0)| =$

$|\pi_0^*(x)| = \|x\|$. Finally, the identity $\|\pi_x^{**}\|_{**} = \|x\|$ holds. So, we have the isometric isomorphism $\tau : X \to X^{**}$ defined $\tau x = \pi_x^{**}$.

The space X is called reflexive if $\tau(X) = X^{**}$. In this case, from topological point of view, spaces X and X^{**} are identical. So that $X^{**} \cong X$ (or just $X^{**} = X$). A good example of a reflexive space is any Hilbert space.

Theorem 3.2.1. *A weakly convergent sequence x_n in a normed vector space X is uniformly bounded in the norm.*

Proof. Note that each x_n defines a linear functional on X^*:

$$\pi_{x_n}^{**}(\pi^*) = \pi^*(x_n) \text{ for all } \pi^* \in X^*.$$

Furthermore, $\|\pi_{x_n}^{**}\|_{**} = \|x_n\|$.

Now for each $\pi^* \in X^*$, $\pi^*(x_n)$ is convergent, and hence bounded. The principle of uniform boundedness thus implies that $\|\pi_{x_n}^{**}\|$ is bounded. The conclusion follows. □

Theorem 3.2.2. *Let x_n be a sequence in a normed vector space X which converges weakly to some $x \in X$. Then*

$$\|x\| \le \liminf_{n\to\infty} \|x_n\|.$$

Proof. By the same result in B4.1 that has been mention above, there is some $\pi^* \in X^*$ such that

$$\|x\| = \pi^*(x) \text{ and } \|\pi^*\|_* = 1.$$

The conclusion follows from the inequality

$$|\pi^*(x_n)| \le \|\pi^*\|_* \|x_n\| = \|x_n\|$$

and the fact that $\pi^*(x_n) \to \pi^*(x) = \|x\|$. □

In fact, we have the following stronger statement:

Theorem 3.2.3 (Mazur)**.** *Let K be a closed convex subset of a normed vector space X, (x_n) be a sequence of points in K converging weakly to x. Then $x \in K$.*

3.3 Weak sequential compactness

Definition 3.3.1. *Let X be a normed space. We say that a sequence $\pi_m^* \in X^*$ converges to $\pi^* \in X^*$ weakly* if $\pi^*(x) \to \pi^*(x)$ for any $x \in X$. We use the following notation $\pi_m^* \xrightarrow{*} \pi^*$.*

If the space X is reflexive then the weak* convergence on X^* is simply a weak convergence on X. Recall that a Banach space is said to be reflexive if it is isometrically isomorphic to its second dual.

Definition 3.3.2. *A subset $A(or A^*)$ of a Banach space $X(or X^*)$ is called weakly (or weakly*) sequentially compact if every sequence of $A(or A^*)$ has a subsequence weakly (weakly*) convergent to a point of $A(or A^*)$.*

The following theorem is a version of the Bolzano-Weierstrass lemma in infinite dimensional setting.

Theorem 3.3.3 (Weak* sequential compactness in separable Banach spaces)**.** *Let X be a separable Banach space. Then the closed unit ball of the space X^* is sequentially weakly* compact.*

Proof. Let $\{x_j\}_{j=1}^{\infty}$ be set which is dense in X and let π_n^* so that $\|\pi_n^*\|_* \leq 1$. To construct a subsequence of π_n^* that converges weakly*, we are going to use the standard diagonal process. First, we selecting a subsequence $\{\pi_{n_1}^{*j}\}_{j=1}^{\infty}$ of the original sequence such that the sequence $\pi_{n_1}^{*j}(x_1)$ converges as $j \to \infty$. Then we selecting a subsequence $\{\pi_{n_2}^{*j}\}_{j=1}^{\infty}$ of the sequence $\{\pi_{n_1}^{*j}\}_{j=1}^{\infty}$ such that the sequence $\pi_{n_2}^{*j}(x_2)$ as $j \to \infty$.

Proceeding in the same way, we construct $\{\pi_{n_k}^{*j}\}_{j=1}^{\infty}$ such that the sequences $\pi_{n_k}^{*j}(x_m)$ convergence when $j \to \infty$ for all $m = 1, 2, ...k$.

Now, we let $l_k^* = \pi_{n_k}^{*k}$. By construction, the sequence $\pi_{n_k}^{*k}(x_m)$ is converging for all m. In order to show that it converges for all $x \in X$, we use the following estimate. Given $x \in X$,

$$|l_m^*(x) - l_n^*(x)| = |(l_m^*(x - x_j) - l_n^*(x - x_)) + l_m^*(x_j) - l_n^*(x_j)| \leq$$
$$\leq 2\|x - x_j\| + |l_m^*(x_j) - l_n^*(x_j)|.$$

We could make small the first term on the right hand side as we wish and passing to the limit as m and n tend to infinity. It remains to let

$$\pi^*(x) = \lim_{m \to \infty} l_m^*(x)$$

for $x \in X$. It is easy to see that $\|\pi^*\|_* \leq 1$. □

Corollary 3.3.4. *The unit ball of a separable reflexive Banach space is sequentially weakly compact.*

We just outline our proof as a stronger statement takes place. Indeed, in this case, $X^{**} = X$. Thus, X^{**} is separable and therefore, X^* is separable (explain why). It remains to notice that weak* convergence on X^* is equivalent to weak convergence on X.

In fact, separability can be dropped in the corollary and a more general statement holds.

Theorem 3.3.5. *Let X be a Banach space. The closed unit ball of X is sequentially weakly compact if and only if X is reflexive.*

We prove this statement partially for Hilbert space.

Theorem 3.3.6. *Let X be a Hilbert space. Then the closed unit ball of X is sequentially weakly compact.*

Proof. Let $x_n \in X$ so that $\|x_n\| \leq 1$. We can use a diagonal process and select a subsequence $\{x_{n_k}\}$ such that $< x_{n_k}, x_j >$ converges as $k \to \infty$ for any j. Therefore, it is easy to that $< x_{n_k}, y >$ is convergent for any $y \in Y = \overline{Span}(x_j)$. We know that $X = Y \oplus Y^\perp$ and thus, for any $x = y + y^\perp$,

$$< x_{n_k}, x >=< x_{n_k}, y > + < x_{n_k}, y^\perp >=< x_{n_k}, y > .$$

So, we can let $\pi^*(x) = \lim_{k\to\infty} < x, x_{n_k} >$ By RRT, there exists $x_0 \in X$ such that $\pi^*(x) =< x, x_0 >$ for all $x \in X$. It means x_{n_k} converges weakly to x_0 and $\|x_0\| \leq 1$. $\square$

As an application of Theorem 3.3.5, we obtain the following generalization of Theorem 1.2.4 for Banach spaces.

Theorem 3.3.7 (Closest point in a closed convex subset)**.** *Let K be a non-empty closed convex subset of a reflexive Banach space X. Then, for every $x \in X$, there is a point $y \in K$ such that no other point in K is which is closer to x than y.*

Note that we do not claim uniqueness; compare Theorem 1.2.4.

Proof. Exercise. $\square$

Chapter 4

Introduction to convergence of Fourier series

We have seen earlier that separable Hilbert spaces have orthonormal bases which can be obtained via the Gram-Schmidt process. The follow orthogonal bases are well known:

(a) The trigonometric functions $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\sin nx, \frac{1}{\sqrt{\pi}}\cos nx, n = 1, 2, \ldots\}$ and $\{\frac{1}{\sqrt{2\pi}}e^{inx}, n \in \mathbb{Z}\}$ in $L^2(-\pi, \pi)$.

(b) The Legendre polynomials $p_n(t)$, indexed by their degrees, in $L^2(-1, 1)$.

(c) The Laguerre polynomials $L_n(t)$ in $L^2((0, \infty); e^{-t}dt)$.

(d) The Hermite polynomials $H_n(t)$ in $L^2(\mathbb{R}; e^{-t^2}\,dt)$.

This chapter examines some introductory aspect of this in the setting of the trigonometric system.

4.1 Fourier series of an integrable periodic functions

Recall that the Fourier series of a function $f \in L^1(-\pi, \pi)$ is given by

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}, \qquad a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx.$$

(This can easily be re-expressed in the familiar trigonometric series using $e^{inx} = \cos nx + i \sin nx$.) Our questions are whether the above series converges and if yes what its sum is.

Note the system $\{e^{inx}\}_{n\in\mathbb{Z}}$ is orthogonal in the complex Hilbert space $L^2(-\pi,\pi)$. Hence, when $f \in L^2(-\pi,\pi)$, we have

$$S(f) = \sum_{n=-\infty}^{\infty} \langle f, e_n \rangle e_n$$

where $e_n = \frac{1}{\sqrt{2\pi}} e^{inx}$ and $\langle f, g \rangle = \int_{-\pi}^{\pi} f\bar{g}\,dx$, and where the infinite sum converges in the sense of L^2-norm and moreover,

$$\|S(f)\|^2 = \sum_{n=-\infty}^{\infty} |\langle f, e_n \rangle|^2.$$

We will see soon that $\{e_n\}_{n\in\mathbb{Z}}$ is in fact an orthonormal basis of $L^2(-\pi,\pi)$ and so $f = S(f)$ as L^2 functions.

A questions then arises whether the Fourier series of f converges to f in any better sense. This is a difficult question and to have a satisfactory answer to its requires knowledge which goes beyond what this course can cover. We are content instead with some brief discussion on the subject.

4.2 Term-by-term differentiation and integrations

Theorem 4.2.1 (Termwise differentiation of Fourier series). *Suppose that $f \in L^1_{loc}(\mathbb{R})$ and let F be the indefinite integral of f, i.e.*

$$F(x) = \int_a^x f(t)\,dt \text{ for some } a \in \mathbb{R}.$$

If F is 2π-periodic and $F \sim \sum c_n e^{inx}$, then $f \sim \sum in\, c_n\, e^{inx}$.

Proof. It is clear that f is 2π-periodic and its zeroth Fourier coefficient is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx = \frac{1}{2\pi}[F(\pi) - F(-\pi)] = 0.$$

For other coefficients, we integrate by parts:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,e^{-inx}\,dx = \frac{in}{2\pi}\int_{-\pi}^{\pi} F(x)\,e^{-inx}\,dx = in\,c_n.$$

This concludes the proof. □

Theorem 4.2.2 (Termwise integration of Fourier series)**.** *Suppose that $f \in L^1(-\pi,\pi)$ is 2π-periodic and let F be the indefinite integral of f, i.e.*

$$F(x) = \int_a^x f(t)\,dt \text{ for some } a \in \mathbb{R}.$$

If $f \sim \sum c_n e^{inx}$, then $F(x) - c_0\,x$ is 2π-periodic and $F(x) - c_0\,x \sim C_0 + \sum_{n\neq 0} \frac{c_n}{in}\,e^{inx}$ where C_0 is a suitable constant.

Proof. Let $G(x) = F(x) - c_0\,x$. We have

$$G(x+2\pi) - G(x) = \int_x^{x+2\pi} f(t)\,dt - 2\pi c_0 = 2\pi c_0 - 2\pi c_0 = 0,$$

and so G is 2π-periodic. By the previous theorem, the Fourier series of $f - c_0$ can be obtained by termwise differentiation of the Fourier series of G. The conclusion is readily seen. □

4.3 Convergence of Fourier series I

Theorem 4.3.1 (Completeness of the trigonometric system)**.** *Assume that $f \in L^2(-\pi,\pi)$, f is 2π-periodic. Then*

$$\lim_{N\to\infty} S_N f = f \text{ in } L^2(-\pi,\pi).$$

In other words, the system $\left\{\frac{1}{\sqrt{2\pi}}e^{inx}\right\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $L^2(-\pi,\pi)$.

Proof. Note that if we let $Sf = S(f)$ be the limit of $S_N f$, then the Fourier coefficients of $f - Sf$ are all zero. Thus, to prove the result, it suffices to show that if the Fourier coefficients of a function $f \in L^2(-\pi,\pi)$ are all zero, then $f = 0$ a.e.

We will only consider the case when f is real-valued. The complex-valued case is left as an excercise.

Suppose first that f is continuous. If $f \neq 0$, then $|f|$ attains it maximum value $M > 0$ at some point, say x_0. Replacing f by $-f$ if necessary, we may assume that $f(x_0) = M > 0$. Using a translation if necessary, we may further assume that $x_0 \in (-\pi, \pi)$. Select $\delta > 0$ such that $|f(x)| > \frac{1}{2}M$ in $(x_0 - \delta, x_0 + \delta) \subset (-\pi, \pi)$. Consider the function

$$g(x) = 1 + \cos(x - x_0) - \cos\delta.$$

Note that $g > 1$ in $(x_0 - \delta, x_0 + \delta)$ and $|g| \leq 1$ in $(-\pi, \pi) \setminus (x_0 - \delta, x_0 + \delta)$. This implies that, for any $n > 0$,

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)\, g^n(x)\, dx &\geq \int_{x_0-\delta/2}^{x_0+\delta/2} f(x)\, g^n(x)\, dx - \int_{(-\pi,\pi)\setminus(x_0-\delta,x_0+\delta)} |f(x)||g(x)|^n\, dx \\ &\geq \frac{1}{2}M\,(1 + \cos\frac{\delta}{2} - \cos\delta)^n\,\delta - 2\pi M\, 1^n \overset{n\to\infty}{\longrightarrow} \infty.\end{aligned}$$

On the other hand, since g is a trigonometric polynomial, the fact that the Fourier coefficients of f are zero implies that the $\int_{-\pi}^{\pi} f(x)\, g^n(x)\, dx = 0$ for all n, which gives a contradiction.

Let us next consider the case when f is merely square integrable. Since the zeroth Fourier coefficient of f is zero, the indefinite integral of f, say $F(x) = \int_0^x f(t)\, dt$ is periodic. Note also that F is continuous. Now, using term-by-term integration, we see that, for some suitable C_0, all the Fourier coefficients of the continuous function $F - C_0$ are zero. Therefore $F - C_0 \equiv 0$, which implies that $f = 0$ a.e. as desired. □

Remark 4.3.2. *The proof above actually shows a stronger statement: if f is an integrable function and if all Fourier coefficients of f are zero, then $f = 0$ a.e.*

Corollary 4.3.3. *Assume that $f \sim \sum c_n e^{inx} \in L^2(-\pi, \pi)$. Then we have Parseval's identity*

$$\sum_{-\infty}^{\infty} |c_n|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f|^2\, dx.$$

4.4 Partial Fourier sums

The N-th partial Fourier sum of a function f is the finite sum

$$S_N f(x) = \sum_{n=-N}^{N} a_n\, e^{inx} = \int_{-\pi}^{\pi} f(t)\, k_N(x - t)\, dt$$

where

$$k_N(x) = \frac{1}{2\pi} \sum_{n=-N}^{N} e^{inx} = \frac{1}{2\pi} \frac{\sin(N+\frac{1}{2})x}{\sin\frac{x}{2}}.$$

A simple manipulation gives also that

$$S_N f(x) = \int_0^\pi [f(x+t) + f(x-t)] k_N(t)\, dt.$$

4.5 Divergence of Fourier series

Theorem 4.5.1. *There exists a 2π-periodic continuous function whose Fourier series diverges at one point.*

Proof. The convergence of the Fourier series of a function f at $x = 0$ means that

$$\lim_{N\to\infty} S_N f(0) = \lim_{N\to\infty} \int_{-\pi}^{\pi} f(x) k_N(x)\, dx \text{ exists.}$$

Let $X = \{f \in C[-\pi, \pi] : f(\pi) = f(-\pi)\}$ and define $A_N \in X^*$ by

$$A_N(f) = \int_{-\pi}^{\pi} f(x)\, k_N(x)\, dx.$$

Assume by contradiction that the Fourier series of every continuous function converges at $x = 0$. Then $A_N(f)$ is bounded for every f. By the principle of uniform boundedness, this means that $\|A_N\|_*$ is bounded.

Now (why?)

$$\|A_N\|_* = \int_{-\pi}^{\pi} |k_N(x)|\, dx.$$

Using the formula for k_N and the inequality $\sin x \leq x$ for $x > 0$, we hence get

$$\|A_N\|_* \geq \frac{1}{\pi} \int_{-\pi}^{\pi} \Big|\sin(N + \frac{1}{2})x\Big| \frac{dx}{|x|} = \frac{2}{\pi} \int_0^{(N+\frac{1}{2})\pi} |\sin x| \frac{dx}{|x|} \geq C \ln N$$

for some positive constant C independent of N. This gives a contradiction and concludes the proof. □

Remark 4.5.2. *(i) It is clear from the proof that, for any sequence $N_j \to \infty$, there is a continuous functions f such that the subsequence $S_{N_j}(f)$ of its partial Fourier sums diverges at a point.*

(ii) *One can use the above to build a continuous function whose Fourier series diverges at any n arbitrarily given points. This is because if two functions agrees in an open interval around a point, say x_0, then their Fourier series either both converge or both diverge at x_0, which is a consequence of Theorem 4.6.1 below.*

4.6 Convergence of Fourier series II

For some $\alpha \in (0,1]$, we say that a function f is α-Hölder continuous at a point x if there is some $A > 0$ and $\delta > 0$ such that

$$|f(x+h) - f(x)| \leq A|h|^\alpha \text{ for } |h| \leq \delta.$$

When $\alpha = 1$, we say f is Lipschitz continuous at x.

Theorem 4.6.1. *Assume that $f \in L^1(-\pi,\pi)$, f is 2π-periodic and f is α-Hölder continuous at a point x_0 for some $\alpha \in (0,1]$. Then*

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

We will use:

Lemma 4.6.2 (Riemann-Lebesgue)**.** *Assume that $f \in L^1(-\pi,\pi)$. Then*

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} f(t)\, e^{ikt}\, dt = 0.$$

Proof. If $f \in L^2(-\pi,\pi)$, this is a consequence of Bessel's inequality, and the fact that $\{\frac{1}{\sqrt{2\pi}} e^{ikx}\}_{k\in\mathbb{Z}}$ is an orthonormal set in $L^2(-\pi,\pi)$.

For the general case $f \in L^1(0,\pi)$, we split $f = g + h$ where $g \in C[-\pi,\pi] \subset L^2(-\pi,\pi)$ and $\|h\|_{L^1(-\pi,\pi)} \leq \varepsilon$ where ε is some positive constant which we can choose as small as we want. Then

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} g(t)\, e^{ikt}\, dt \to 0$$

while

$$\Big| \int_{-\pi}^{\pi} h(t)\, e^{ikt}\, dt \Big| \leq \int_{-\pi}^{\pi} |h(t)|\, dt \leq \varepsilon.$$

The conclusion is readily seen. ☐

Proof of Theorem 4.6.1. The theorem holds obviously for f being a constant function. We can thus assume without loss of generality that $f(x_0) = 0$, so that $|f(x_0 + h)| \le A|h|^\alpha$ for small h.

For $\delta > 0$ to be fixed, a simple application of the Riemann-Lebesgue lemma shows that

$$\lim_{N\to\infty} \int_\delta^\pi [f(x_0 + t) + f(x_0 - t)]\, k_N(t)\, dt \to 0.$$

On the other hand, using the Hölder continuity of f at x_0, we have

$$\begin{aligned}\Big| \int_0^\delta [f(x_0 + t) + f(x_0 - t)]\, k_N(t)\, dt\Big| &\le 2A \int_0^\delta |t|^\alpha\, |k_N(t)|\, dt \\ &\le \frac{A}{\pi} \int_0^\delta \frac{|t|^\alpha}{\sin\frac{t}{2}}\, dt.\end{aligned}$$

Using the inequality $\sin\frac{t}{2} \ge \frac{t}{\pi}$ for $0 \le t \le \pi$, we see that the right hand side is bounded from above by $\frac{A}{\alpha}\delta^\alpha$. Putting everything together we obtain

$$\limsup_{N\to\infty} |S_N f(x_0)| \le \frac{A}{\alpha}\, \delta^\alpha.$$

Since δ was arbitrary, this implies $S_N f(x_0) \to 0 = f(x_0)$, as desired. $\square$

The above theorem can be used to for an alternative proof of Theorem 4.3.1 about completeness of the orthonormal sequence $\{e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}\}_{n\in\mathbb{Z}}$ in $L^2(-\pi, \pi)$. Let us outline it. Indeed, consider the sequence of linear operators $S_N : L^2(-\pi, \pi) \to L^2(-\pi, \pi)$ so that

$$S_N f = \sum_{k=-N}^{N} \langle f, e_n\rangle e_n.$$

By Bessel's inequality, $\|S_n\| \le 1$ and $S_N(f) \to f$ in $L_2(-\pi, \pi)$ for any $f \in C^1([-\pi, \pi])$ such that $f(-\pi) = f(\pi)$. Such functions are in dense in the space $X = \{f \in L^2_{loc}(\mathbb{R}) : f(x + 2\pi) = f(x)\ \forall x \in \mathbb{R}\}$ with respect to the norm of $L_2(-\pi, \pi)$. This can be verified with the help of mollification. Then the result follows from the Banach-Steinhaus Theorem.

Chapter 5

Spectral theory in Hilbert spaces

5.1 Main definitions

If T is a linear operator on $\mathbb{C}^n$, the the spectrum of T is the set of all eigenvalues of T, i.e. the complex numbers λ such that the determinant of $\lambda\, I - T$ vanishes. It consists of at most n complex numbers. If λ is not an eigenvalue of T, then $\lambda I - T$ has an inverse.

The spectral theory for operators on infinite dimensional space is far richer and of fundamental importance for an understanding the operators themselves.

Definition 5.1.1. *Let X be a complex Banach space and $T \in \mathscr{B}(X)$.*

(i) The spectrum $\sigma(T)$ of T is the set of complex numbers λ such that $\lambda I - T$ has no inverse in $\mathscr{B}(X)$.

(ii) The resolvent set $\rho(T)$ of T is the complement of $\sigma(T)$ in $\mathbb{C}$. If $\lambda \in \rho(T)$, then $R_\lambda(T) = (\lambda I - T)^{-1}$ is called the resolvent of T at λ.

(iii) The point spectrum $\sigma_p(T)$ of T is the set of complex numbers λ such that $Ker(\lambda I - T)$ is non-trivial. The elements of $\sigma_p(T)$ are called the eigenvalues of T, and, if $\lambda \in \sigma_p(T)$, the non-trivial elements of $Ker(\lambda I - T)$ are called the eigenvectors of T.

(iv) The residual spectrum $\sigma_r(T)$ of T is the set of complex numbers λ such that $Ker(\lambda I - T) = \{0\}$ and $Im(\lambda I - T)$ is not dense in X.

(v) *The continuous spectrum* $\sigma_c(T)$ *of* T *is the set of complex numbers* λ $Ker(\lambda I - T) = \{0\}$ *and* $Im(\lambda I - T)$ *is a proper dense subset of* X*.*

(vi) *The approximate point spectrum* $\sigma_{ap}(T)$ *of* T *is the set of complex numbers* λ *such that there is a sequence* $x_n \in X$ *such that* $\|x_n\| = 1$ *and* $\|Tx_n - \lambda x_n\| \to 0$*.*

It is clear that $\sigma(T)$ is the disjoint union of $\sigma_p(T)$, $\sigma_r(T)$ and $\sigma_c(T)$.

We know that $\sigma(T)$ is a non-empty closed subset of $\mathbb{C}$, and if $\lambda \in \sigma(T)$, then $|\lambda| \leq \|T\|$. We also know that

$$\sigma_p(T) \subset \sigma_{ap}(T) \subset \sigma(T) = \sigma_{ap}(T) \cup \sigma_p(T') \text{ and } \sigma_r(T) \subset \sigma_{ap}(T'),$$

where T' is the transpose of T.

Lemma 5.1.2. *Let* $T \in \mathscr{B}(X)$ *be a bounded linear operator on a Banach space* X*. Then* $\sigma_c(T) \subset \sigma_{ap}(T)$*.*

Proof. Take $\lambda \in \sigma_c(T)$. Then $\lambda I - T$ is injective and $Y := \mathrm{Im}(\lambda I - T)$ is a proper dense subspace of X. Arguing indirectly, assume that $\lambda \notin \sigma_{ap}(T)$ and so there is some $c > 0$ such that

$$\|(\lambda I - T)x\| \geq c \text{ for all } x \in X, \|x\| = 1.$$

This implies that

$$\|(\lambda I - T)x\| \geq c\|x\| \text{ for all } x \in X. \tag{5.1}$$

Note that as a map from X into Y, $\lambda I - T$ is bijective and so has an inverse, say $U : Y \to X$. It is clear that U is linear. By (5.1), we have $\|Uy\| \leq c^{-1}\|y\|$ for all $y \in Y$. Hence $U \in \mathscr{B}(Y, X)$. As Y is a dense subspace of X, U extends to a bounded linear operator on X, say $\bar{U}$.

Now, pick $p \in X \setminus Y$ and $p_n \in Y$ such that $p_n \to p$. Then $Up_n \to \bar{U}p$ and so

$$(\lambda I - T)\bar{U}p = \lim_{n\to\infty}(\lambda I - T)\bar{U}p_n = \lim_{n\to\infty} p_n = p.$$

This shows that p belongs to Y, a contradiction. $\square$

When X is a Hilbert space, Lemma 5.1.2 can be proved using Hilbert space techniques as follows.

Second proof of Lemma 5.1.2. Let $\lambda \in \sigma_c(T)$ so that $Y := \mathrm{Im}(\lambda I - T)$ is a dense proper subspace of X. Pick $p \in X \setminus Y$. Then there is some a sequence x_n such that $p_n := (\lambda I - T)x_n \to p$.

If (x_n) is bounded, then, by the weak sequential compactness property of the unit ball, we may assume without loss of generality that x_n converges weakly to some x. This implies, for $z \in X$, that

$$\langle p_n, z\rangle = \langle (\lambda I - T)x_n, z\rangle = \langle x_n, (\bar{\lambda} I - T^*)z\rangle \to \langle x, (\bar{\lambda} I - T^*)z\rangle = \langle (\lambda I - T)x, z\rangle.$$

In other words, p_n converges weakly to $(\lambda I - T)x$. By since p_n converges strongly to p, we thus obtain $p = (\lambda I - T)x$, which contradicts the choice of p. We thus deduce that (x_n) is unbounded. Replacing (x_n) by a subsequence if necessary, we may assume that $\|x_n\| \to \infty$.

Let $z_n = \|x_n\|^{-1} x_n$. We then have $\|z_n\| = 1$ and $\|(\lambda I - T)z_n\| = \|x_n\|^{-1} \|(\lambda I - T)x_n\| = \|x_n\|^{-1} \|p_n\| \to 0$ as $\|p_n\|$ is bounded sequence. Hence $\lambda \in \sigma_{ap}(T)$. □

In the rest of the chapter, we will specialise to the case where X is a Hilbert space (over $\mathbb{C}$). Note that in this case, the notions of dual operator and adjoint operator can be linked via the Riesz representation theorem.

5.2 Adjoints and spectra

We start with some simple statements.

Proposition 5.2.1. *Let X be a complex Hilbert space, $T \in \mathscr{B}(X)$ and $\lambda \in \mathbb{C}$. Then the following holds.*

(i) $(\lambda I - T)^* = \bar{\lambda} I - T^*$.

(ii) $\lambda I - T$ *is invertible if and only if* $\bar{\lambda} I - T^*$ *is invertible. In particular,* $\lambda \in \sigma(T)$ *if and only if* $\bar{\lambda} \in \sigma(T^*)$.

(iii) $Ker(\lambda I - T) = Im(\bar{\lambda} I - T^*)^{\perp}$ *and* $Ker(\lambda I - T)^{\perp} = \overline{Im(\bar{\lambda} I - T^*)}$.

Proof. Exercise. □

Proposition 5.2.2. *Let X be a complex Hilbert space, $T \in \mathscr{B}(X)$ and $\lambda \in \mathbb{C}$. Then the following holds.*

(i) If T is normal (i.e. $TT^* = T^*T$), then $Ker(\lambda I - T) = Ker(\bar{\lambda} I - T^*)$.

(ii) *If T is self-adjoint, then $\sigma_p(T) \subset \mathbb{R}$.*

Proof. (i) Assume that T is normal. Then $S := \lambda I - T$ is also normal. This implies that

$$\|Sx\|^2 = \langle Sx, Sx\rangle = \langle x, S^*Sx\rangle = \langle x, SS^*x\rangle = \langle S^*x, S^*x\rangle = \|S^*x\|^2$$

for all $x \in X$. The conclusion follows.

(ii) Assume that T is self-adjoint and $\lambda \in \sigma_p(T)$. Let x be an eigenvector of T corresponding to λ. We have

$$\lambda\|x\|^2 = \langle Tx, x\rangle = \langle x, Tx\rangle = \bar{\lambda}\|x\|^2.$$

This implies that $\lambda \in \mathbb{R}$. □

Theorem 5.2.3. *Let X be a complex Hilbert space and $T \in \mathscr{B}(X)$. Then*

$$\sigma(T) = \sigma_{ap}(T) \cup \sigma_p'(T^*)$$

where $\sigma_p'(T^*) = \{\lambda : \bar{\lambda} \in \sigma_p(T^*)\}$.

Proof. This was proved in B4.1 for Banach spaces, we recall the proof here.

In view of Proposition 5.2.1(ii), $\sigma(T) \supset \sigma_{ap}(T) \cup \sigma_p'(T^*)$. Consider the converse. Assume $\lambda \in \sigma(T) \setminus \sigma_{ap}(T)$. Then by Lemma 5.1.2, λ must lie in the residual spectrum of T. Now, by Proposition 5.2.1(iii), $\bar{\lambda}I - T^*$ has a non-trivial kernel and so $\bar{\lambda} \in \sigma_p(T^*)$ as desired. □

Theorem 5.2.4. *Let X be a complex Hilbert space and $T \in \mathscr{B}(X)$ be self-adjoint. Then*

(i) *$\sigma(T) \subset \mathbb{R}$,*

(ii) *T has no residual spectrum, i.e. $\sigma(T) = \sigma_{ap}(T) = \sigma_p(T) \cup \sigma_c(T)$,*

(iii) *and eigenvectors corresponding to different eigenvalues of T are orthogonal.*

Proof. (i) By Proposition 5.2.2, $\sigma_p(T^*) \subset \mathbb{R}$. Thus, by Theorem 5.2.3, we only need to show that $\sigma_{ap}(T) \subset \mathbb{R}$.

Let λ be an approximate eigenvalue so that there is a sequence (x_n) such that $\|x_n\| = 1$ and $\|(\lambda I - T)x_n\| \to 0$. By the Cauchy-Schwarz equality, we have

$$\lambda - \langle Tx_n, x_n\rangle = \langle(\lambda I - T)x_n, x_n\rangle \to 0.$$

In other words, $\langle Tx_n, x_n\rangle \to \lambda$. But as T is self adjoint, we have $\langle Tx_n, x_n\rangle = \langle x_n, T^*x_n\rangle = \langle x_n, Tx_n\rangle = \overline{\langle Tx_n, x_n\rangle}$ and so $\langle Tx_n, x_n\rangle \in \mathbb{R}$. Hence $\lambda \in \mathbb{R}$.

(ii) If λ is in the residual spectrum of T, then by Proposition 5.2.1(iii), $\bar{\lambda} = \lambda$ belongs to the point spectrum of $T^* = T$. But this is not possible since by definition, the point spectrum and the residual spectrum of T are disjoint.

(iii) Exercise. □

Lemma 5.2.5. *The spectral radius of a self-adjoint bounded linear operator T on a complex Hilbert space X is equal to its norm:*

$$rad(\sigma(T)) = \|T\|.$$

Proof. By Proposition 1.4.8, we have $\|T^n\| = \|T\|^n$ when $n = 2^k$, $k \in \mathbb{N}$. The conclusion then follows from Gelfand's formula (established in B4.1) which asserts that $\mathrm{rad}(\sigma(T))$ is the limit of $\|T^n\|^{1/n}$. □

We know that, for a self-adjoint operator T, acting on a complex Hilbert space, $\langle Tx, x\rangle$ is real. We also know that $|\lambda| \le \|T\|$ for $x \in \sigma(T)$ and $\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}$.

Lemma 5.2.6. *Let X be a complex Hilbert space and $\in \mathcal{B}(X)$ be self-adjoint. Assume that a unit vector x_0 satisfies the condition $\|T\| = |\langle Tx_0, x_0\rangle|$. Then x_0 is an eigenvector of T belonging to a eigenvalue λ_0, i.e., $Tx_0 = \lambda x_0$, such that $\|T\| = |\lambda_0|$.*

Proof. WLOG we can assume that $\langle Tx_0, x_0\rangle > 0$. Pick up any $y \in Y$ such that $\langle y, x_0\rangle = 0$ and consider the vector

$$x = \frac{x_0 + \alpha y}{\sqrt{1 + |\alpha|^2\|y\|^2}}$$

for any $\alpha \in \mathbb{C}$. Obviously, $\|x\| = 1$. Simple calculations show:

$$\langle Tx, x\rangle = \frac{1}{1 + |\alpha|^2\|y\|^2}(\langle Tx_0, x_0\rangle + \langle Tx_0 \alpha y\rangle + \langle x, \alpha Ty\rangle +$$

$$+|\alpha|^2\langle Ty, y\rangle) == \frac{1}{1 + |\alpha|^2\|y\|^2}(\langle Tx_0, x_0\rangle + \overline{\alpha}\langle Tx_0, y\rangle +$$

$$+\alpha\overline{\langle Tx_0, y\rangle} + |\alpha|^2\langle Ty, y\rangle) \le \langle Tx_0, x_0\rangle.$$

The latter inequality implies $\langle Tx_0, y\rangle = 0$. (Explain why?)

Now, let $Y = \mathrm{Span}\{x_0\}$ and $X = Y \oplus Y^\perp$. We have proven $Tx_0 \in (Y^\perp)^\perp = Y$. Hence, $Tx_0 = \lambda_0 x_0$ and $\langle Tx_0, x_0\rangle = \lambda_0$. □

We also can prove the following statement.

Theorem 5.2.7. *Let X be a complex Hilbert space and $T \in \mathscr{B}(X)$. If T is self-adjoint, then the spectrum of T lies in the closed interval $[a, b]$ on the real axis, where*

$$a = \inf_{\|x\|=1} \langle x, Tx \rangle \text{ and } b = \sup_{\|x\|=1} \langle x, Tx \rangle.$$

Proof. We know from Theorem 5.2.4 that $\sigma(T) \subset \mathbb{R}$ and $\sigma_r(T) = \emptyset$. The second one implies that $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) = \sigma_{ap}(T)$.

Suppose that $\lambda \in \sigma_{ap}(T)$. Then for a sequence (x_n) with $\|x_n\| = 1$ we have $\lambda x_n - Tx_n \to 0$. By the Cauchy-Schwarz inequality, we have

$$\lambda - \langle Tx_n, x_n \rangle = \langle \lambda x_n - Tx_n, x_n \rangle \to 0.$$

As $a \leq \langle Tx_n, x_n \rangle \leq b$, it follows that λ is real and $\lambda \in [a, b]$. We have thus shown that $\sigma(T) \subset [a, b]$.

□

We conclude the section with a result on spectra of unitary operators.

Proposition 5.2.8. *Let X be a complex Hilbert space and $U \in \mathscr{B}(X)$ be unitary. Then $|\lambda| = 1$ for all $\lambda \in \sigma(U)$.*

Proof. By Proposition 1.5.4, U is a surjective isometry and $U^{-1} = U^*$. It follows that $|\lambda| \leq \|U\| = 1$ for all $\lambda \in \sigma(U)$.

Assume by contradiction that there is some λ with $|\lambda| < 1$ such that $\lambda I - U$ is not invertible. It follows that $\bar{\lambda} I - U^*$ is also not invertible. Consequently, $\bar{\lambda} U - I = (\bar{\lambda} I - U^*)U$ is also not invertible (since U is invertible), and so $\bar{\lambda}^{-1} \in \sigma(U)$. This amounts to a contradiction as $|\bar{\lambda}^{-1}| > 1$. □

5.3 Examples

Example 5.3.1. *Let $X = \ell^2$ and $T((a_1, a_2, a_3, \ldots)) = (a_1, a_2/2, a_3/3, \ldots)$. Then $\sigma(T) = \sigma_{ap}(T) = \{0\} \cup \{k^{-1} : k = 1, 2, \ldots\}$, $\sigma_p(T) = \{k^{-1} : k = 1, 2, \ldots\}$, $\sigma_c(T) = \{0\}$, $\sigma_r(T) = \emptyset$.*

Example 5.3.2. *Let $X = \ell^2(\mathbb{Z})$ (i.e. the set of bi-infinite square summable sequences) and R be the right shift. Then R is unitary, $\sigma(R) = \sigma_{ap}(R) = \sigma_c(R) = \mathbb{S}^1$ and $\sigma_p(R) = \sigma_r(R) = \emptyset$. The same statement holds for the left shift.*

Example 5.3.3. *Let $X = L^2(\mathbb{R})$ and consider the multiplication operator M_h where h is real valued and belongs to $L^\infty(\mathbb{R})$. Then*

$$\begin{aligned}
\sigma(M_h) &= \sigma_{ap}(M_h) = \text{ the essential range of } h \\
&= \{\lambda \in \mathbb{R} : h^{-1}((\lambda - \epsilon, \lambda + \epsilon)) \text{ has positive measure} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for all small } \epsilon > 0\}, \\
\sigma_p(M_h) &= \{\lambda \in \mathbb{R} : \{h = \lambda\} \text{ has positive measure}\}, \\
\sigma_r(M_h) &= \emptyset, \\
\sigma_c(M_h) &= \sigma_{ap}(M_h) \setminus \sigma_p(M_h).
\end{aligned}$$

Example 5.3.4. *(not under examination) Compact self-adjoint operators.*

Definition 5.3.5. *Let X be Y be two normed spaces. $K \in \mathcal{B}(X, Y)$ is a compact operator if, for any bounded set $B \subset X$, $K(B)$ is a pre-compact set in Y.*

Simple example: if $\operatorname{Dim} Y < \infty$, then any $K \in \mathcal{B}(X, Y)$ is pre-compact. Let $K \in \mathcal{B}(X)$ be a compact operator in infinite dimensional space X. Then $0 \in \sigma(K)$. Indeed, assume that there exists $K^{-1} \in \mathcal{B}(X)$. Let B be the unit ball of X. Since $K^{-1}(B)$ is bounded, the set $B = K(K^{-1}B)$ must be pre-compact which is not true.

There are several nice properties of compact self-adjoint operators.

Proposition 5.3.6. *Let K be a compact self-adjoint operator on a complex Hilbert space X. Let non-zero $\lambda \in \sigma(K)$. Then $\lambda \in \sigma_p(K)$.*

Proof. Since K is a self-adjoint, λ is an approximate eigenvalue, i.e., there exists a sequence $x_n \in X$ such that

$$\|x_n\| = 1, \qquad \|Kx_n - \lambda x_n\| \to 0$$

as $n \to \infty$. We know that there exists a subsequence $x_{n_j} \rightharpoonup x$. Since K is compact $Kx_{n_j} \to Kx$ and thus (since $\lambda \neq 0$) $x_{n_j} \to x$. Obviously, $\|x\| = 1$, $Kx - \lambda x = 0$. ∎

Theorem 5.3.7. *Let K be a compact self-adjoint operator on a complex Hilbert space X. Then there exists at least one eigenvector of K. Moreover, this eigenvector belongs to an eigenvalue λ_0 of K satisfying $\|K\| = |\lambda_0|$.*

Proof. Without loss of generality, we may assume that $K \neq 0$. We know that for any self-adjoint operator K

$$\|K\| = \sup\{|\langle Kx, x\rangle| : \|x\| = 1\},$$

see the section about adjoint operators. Let a sequence x_n be such that $\|x_n\| = 1$ and $|\langle Kx_n, x_n\rangle| \to \|K\|$. We also can find a subsequence $x_{n_j} \rightharpoonup x_0$ with

$$1 = \liminf \|x_{n_j}\| \geq \|x_0\|,$$

and, since K is a compact operator, $Kx_{n_j} \to Kx_0$ and thus $\langle Kx_{n_j}, x_{n_j}\rangle \to \langle Kx_0, x_0\rangle = \|K\|$. (Explain why). Clearly, $x_0 \neq 0$ and in fact, $\|x_0\| = 1$. If not, let $x' = x_0/\|x_0\|$, then $\|x'\| = 1$ and

$$|\langle Kx', x'\rangle| = \frac{1}{\|x_0\|^2}\|K\| \leq \|K\|,$$

which is wrong if $\|x_0\| < 1$. The result follows from Lemma 5.2.6. $\square$

Theorem 5.3.8. *Let K be a compact self-adjoint operator on a complex Hilbert space X. Let $\delta > 0$ and let us introduce the set*

$$\sum = \mathrm{Span}\{x \in X : \|x\| = 1, \, Kx = \lambda x, \, |\lambda| \geq \delta\}.$$

Then $\dim \sum < \infty$.

Proof. Assume for contradiction that for any n there are linearly independent vectors $x_1, x_2, ..., x_n$ such that $x_i \neq 0$, $Kx_i = \lambda_i x_i$, with $|\lambda_i| \geq \delta$, $i = 1, 2, ..., n$. We let $E_n = \mathrm{Span}\{x_1, x_2, ..., x_n\}$. By construction, E_{n-1} is a proper subspace of E_n.

Now, let $y_1 = x_1/\|x_1\|$. Our aim is to show that there exists a sequence $y_2, y_3, ...$, with the following properties: $y_n \in E_n$, $\|y_n\| = 1$, and $\mathrm{dist}(y_n, E_{n-1}) \geq 1/2$, $n = 2, 3, ...$ Indeed, by assumptions $\mathrm{dist}(x_n, E_{n-1}) = \alpha > 0$. Obviously, there is $x_* \in E_{n-1}$ such that $\|x_n - x_*\| < 2\alpha$. Since $\alpha = \mathrm{dist}(x_n - x_*, E_{n-1})$, we can let $y_n = (x_n - x_*)/\|x_n - x_*\|$. We then have $\|y_n\| = 1$, $y_n \in E_n$, and $\mathrm{dist}(y_n, E_{n-1}) = \alpha/\|x_n - x_*\| > 1/2$. Notice that

$$\|y_n/\lambda_n\| = 1/\lambda_n \leq 1/\delta.$$

If we can show that sequence $K(y_n/\lambda_n)$ is not pre-compact, we get a contradiction.

So, we have

$$y_n = \sum_{i=1}^{n} \alpha_i x_i, \quad K(y_n/\lambda_n) = \sum_{i=1}^{n} \lambda_i \alpha_i x_i/\lambda_n = y_n + z_n,$$

where

$$z_n = \sum_{i=1}^{n-1} (\lambda_i/\lambda_n - 1)\alpha_i x_i \in E_{n-1}.$$

For any $n > m$, $z_n - y_m - z_m \in E_{n-1}$ and thus

$$\|K(y_n/\lambda_n) - K(y_m/\lambda_m)\| = \|y_n + z_n - y_m - z_m\| > 1/2.$$

So, there is no converging subsequence of $K(y_n/\lambda_n)$. □

Theorem 5.3.9. *Let K be a compact self-adjoint operator in a complex Hilbert space X. There exists an orthonormal sequence $\{e_n\}_{n=1}^{N}$, $N \leq \infty$, consisting of eigenvectors e_n that belong to a non-zero eigenvalue λ_n of K, with the following property. For any $x \in X$, we have a unique representation*

$$x = \sum_{n=1}^{N} c_n e_n + x'$$

where $x' \in \mathrm{Ker} K$. If $N = \infty$, then $\lambda_k \to 0$ as $k \to \infty$.

Proof. As it follows from Theorem 5.3.8, the set of non-zero eigenvalues is countable. We list them in the following order

$$|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n| \geq ....$$

Step 1 By Theorem 5.3.7, we know that

$$\|K\| = |\lambda_1| = \sup\{|\langle Kx, x\rangle| : \|x\| = 1\}$$

and there exists an eigenvector e_1 such that $\|e_1\| = 1$ and $Ke_1 = \lambda_1 e_1$.
Step 2 Now, we argue by induction, assuming that there are eigenvectors $e_1, e_2, ..., e_n$ such that $\|e_j\| = 1$, $Ke_j = \lambda_j e_j$, and $\langle e_j, e_m\rangle = 0$, by Theorem 5.2.4(iii), for $j \neq m$ and $j, m = 1, 2, ..., n$. Moreover,

$$|\lambda_j| = \sup\{|\langle Kx, x\rangle| : \|x\| = 1, \langle x, e_m\rangle = 0, m = 1, 2, ..., j-1\}.$$

Consider a closed subspace $M_n = \mathrm{Span}\{e_1, e_2, ..., e_n\}$. Then $X = M_n \oplus M_n^\perp$. Notice that $M_n^\perp$ is invariant with respect to K, i.e., $K(M_n^\perp) \subseteq M_n^\perp$. Indeed, assume that there exists $y \in K(M_n^\perp) \setminus M_n^\perp$, then $y = m + m^\perp = Kx$, with $x \in M_n^\perp$, $m \in M_n$, and $m^\perp \in M_n^\perp$. Since M_n is invariant with respect to K, $\|m\|^2 = \langle Kx, m\rangle = \langle x, Km\rangle = 0$ and thus $y \in M_n^\perp$. If we denote by $K_n : M_n^\perp \to M_n^\perp$ the restriction of K to M_n, we can repeat arguments of Step1 replacing X with M_n and K with K_n and get $e_{n+1} \in M_n^\perp$ with $\|e_{n+1}\| = 1$ and $Ke_{n+1} = K_n e_{n+1} = \lambda_{n+1} e_{n+1}$ and

$$|\lambda_{n+1}| = \sup\{|\langle K_n x, x\rangle| : \, x \in M_n^\perp, \, \|x\| = 1\} =$$

$$= \sup\{|\langle Kx, x\rangle| : \, \|x\| = 1, \, \langle x, e_m\rangle = 0, \, m = 1, 2, ..., n\}.$$

Step 3 Here, we should consider two cases. In the first case, after finite steps, we get $\langle Kx, x\rangle = 0$ for all $x \in M_{n_0}^\perp$ for some n_0. This implies that $M_{n_0}^\perp = \mathrm{Ker} K$, i.e., the set of all eigenvectors belonging to zero eigenvalue.

In the second case, $\langle Kx, x\rangle$ is not identically equal to zero on $M_n^\perp$ for all n. In this case, $\lambda_n \to 0$. If not $|\lambda_n| \geq \delta$ for all n. Contradiction follows from Theorem 5.3.8. Now, let $M = \mathrm{Span}\{e_1, e_2, ..., e_n, ...\}$ and $X = M \oplus M^\perp$. We need to show that $\mathrm{Ker} K = M^\perp$. To this end, we first notice that $M^\perp$ is invariant with respect to K. So, if $x \in \mathrm{Ker} K$, then $0 = \langle Kx, e_n\rangle = \langle x, Ke_n\rangle = \overline{\lambda}_n \langle x, e_n\rangle$ for any n, and thus $x \in M^\perp$. To show converse, we first notice that if $\langle Kx, x\rangle \equiv 0$ on $M^\perp$, then the opposite inclusion is trivially true. Otherwise, we have

$$0 < |\lambda| = \sup\{|\langle Kx, x\rangle| : \, \|x\| = 1, \, x \in M^\perp\} \leq |\lambda_n| \to 0.$$

Uniqueness is easy. □

Theorem 5.3.10. *Let K be a compact self-adjoint compact operator in a separable Hilbert space X. There exists an orthonormal basis of X consisting of eigenvalues of the operator K.*

Proof. Our first remark is that Ker K is a separable Hilbert space itself. By the Gram-Schmidt rule, there exists a orthonormal basis $\{e_n'\}$ of $\mathrm{Ker} K$ which consists of eigenvectors belonging zero eigenvalue. Now, it remains to use Theorem 5.3.9. According to it, there exists an orthonormal sequence $\{e_n''\}$ consisting of eigenvectors belonging to non-zero eigenvalues such that $\{e_n'\} \cup \{e_n''\}$ is the orthonormal basis in X. □