B4.2 Functional Analysis II

Luc Nguyen

Hilary 2018

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

This set of lecture notes build upon and expand Hilary Priestley's and Gregory Seregin's lecture notes 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. *An inner (scalar) product in a linear vector space X over $\mathbb{R}$ is a real-valued function on $X \times X$, denoted as $\langle x, y \rangle$, having the following properties:*

(i) Bilinearity. *For fixed y, $\langle x, y \rangle$ is a linear function of x, and for fixed x, $\langle x, y \rangle$ is a linear function of y.*

(ii) Symmetry. *$\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in X$.*

(iii) Posivity. *$\langle x, x \rangle > 0$ for $x \neq 0$.*

When X is a vector space over $\mathbb{C}$, $\langle x, y \rangle$ is complex-valued and properties (i) and (ii) are replaced by

(i') Sesquilinearity. *For fixed y, $\langle x, y \rangle$ is a linear function of x, and for fixed x, $\langle x, y \rangle$ is a skewlinear function of y, i.e.*

$$\langle ax, y \rangle = a\langle x, y \rangle \text{ and } \langle x, ay \rangle = \bar{a}\langle x, y \rangle \text{ for all } a \in \mathbb{C}, x, y \in X.$$

(ii') Skew symmetry. *$\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in X$.*

Caution. *In some textbooks, the sesquilinearity property is reversed: $\langle x, y \rangle$ is required instead to be skewlinear in x and linear in y.*

The inner product $\langle\cdot,\cdot\rangle$ generates a norm, denoted by $\|\cdot\|$, 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 (iii), and the homogeneity of $\|\cdot\|$ follows from the bi/sequilinearity property (i)/(i'). To prove the triangle inequality, we use:

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 $y \neq 0$. Replacing x by ax with $|a| = 1$ so that $a\langle x, y\rangle$ is real, we may assume without loss of generality that $\langle x, y\rangle$ is real.

For $t \in \mathbb{R}$, we compute using sesquilinearity and skew symmetry:

$$\|x + ty\|^2 = \langle x + ty, x + ty\rangle = \|x\|^2 + 2t\,\mathrm{Re}\,\langle x, y\rangle + t^2\|y\|^2. \tag{1.1}$$

By positivity, this quadratic polynomial in t is non-negative for all t. This implies that

$$(\mathrm{Re}\,\langle x, y\rangle)^2 - \|x\|^2\|y\|^2 \leq 0,$$

which gives the desired inequality. If equality holds, then there is some t_0 such that $x + t_0 y = 0$. The conclusion follows. □

If we set $t = \pm 1$ in (1.1) and add the resulting identities, we obtain the so-called parallelogram law:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \text{ for all } x, y \in X. \tag{1.2}$$

It is a fact that if a norm satisfies the parallelogram law (1.2), 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. *A linear vector space with an inner product is called an inner product space. If it is complete with the induced norm, it is called a Hilbert space.*

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*

$$u(x) = A + \int_a^x v(y)\,dy \tag{1.3}$$

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

Note that by (1.3), any $u \in H^1(a,b)$ has a continuous representation in $[a,b]$, since

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

Also, for any given $u \in H^1(a,b)$, there is only one function v satisfying (1.3). Indeed, if there are two constants A_1, A_2 and two functions v_1, v_2 satisfying (1.3) 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 (1.3) 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) $(\overline{\operatorname{span} Y})^\perp = Y^\perp$.

(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.*

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 (1.2) 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

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

It follows that $2t \operatorname{Re}\langle x - y_0, y\rangle + t^2 \|y\|^2 \ge 0$ 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.*

It is easy to see that the closed linear span of a set S is the closure of the linear span $\operatorname{Span} S$.

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

Proof. Exercise. □

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 $x \ne y \in S$.*

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 $X = \overline{\text{span}\, B}$, and so X is the closed linear span of B. □

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*

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

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*

$$x = \sum_{n=1}^{\infty} a_n\, x_n \tag{1.4}$$

where the sequence of scalar $(a_1, a_2, \ldots)$ belongs to ℓ^2. The sum in (1.4) *converges in the sense of the Hilbert space norm. Furthermore*

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

and

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

[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.

Proof. Let Y denotes the closed linear span of S. It is clear that, if the sequence of coefficients (a_n) belongs to ℓ^2, the the sum in (1.4) converges in the sense of the Hilbert space norm and so defines a vector in Y. Conversely, assume that $x \in Y$ and let $a_n = \langle x, x_n \rangle$. By Bessel's inequality, $(a_n) \in \ell^2$ and so the vector

$$\tilde{x} = \sum_{n=1}^{\infty} a_n x_n \in Y \subset X.$$

Now observe that $x - \tilde{x}$ is perpendicular to all x_n and thus belongs to $Y^\perp$, in view of Proposition 1.2.8. Since $Y \cap Y^\perp = \{0\}$, we deduce that $x = \tilde{x}$. This shows that x has the desired form. Parserval's identity then follows from Pythagorean theorem. □

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. □

Example 1.4.2. *Let* $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ *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

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

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 by ay for some scalar a with $\|a\| = 1$, we may assume that $|\langle Tx, y\rangle| = \langle Tx, y\rangle$. This implies that

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

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

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

Proof. Exercise. □

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. □

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.*

For normed vector spaces, the conclusion holds if one has in addition that T is surjective. This is a result due to Mazur and Ulam.

For complex Hilbert spaces, one is tempted to say that T is some sort of combinations of linear and skewlinear maps, but this is not clear. Note that there are complex spaces which are not (complex) isomorphic to their complex conjugates.

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

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

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. □

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

We have the following characterization 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. □

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 an application, due to von Neumann, of the Riesz reprensentation 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:*

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

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

$$\int_A f(1-h)\, dm = \int_A f\, h\, d\mu \text{ for all } f \in X. \tag{1.5}$$

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 (1.5), 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 (1.5) 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 $0 < h \leq$ except on a set of zero Lebesgue measure. Now setting $g = \frac{1-h}{h}$ and choosing $f = \frac{1}{h}$ in (1.5), we obtain the conclusion. □

Chapter 2

Bounded linear operators: The Baire catergory 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 (Principle of uniform boundedness; Banach-Steinhaus theorem)**.** *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 pointwise 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. *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 $\|T_n\| \leq M$ and the sequence $(T_n z)$ is convergent for each $z \in Z$.

Caution. *In the above theorem, the convergence of T_n to T is in the pointwise sense. This should not be confused with the convergence in norm, i.e. under (i), or (ii) or (iii) in the theorem, it needs not be the case that $\|T_n - T\| \to 0$. To see this consider for example $X = \ell^2$, $Y = \mathbb{R}$ and $T_n((a_1, a_2, \ldots)) = a_n$. Then, for every $x \in \ell^2$, $T_n x \to 0$, but $\|T_n\| = 1 \not\to 0$.*

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 $z \in Z$,

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

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

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

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

2.3 The open mapping theorem

Theorem 2.3.1 (Open mapping theorem)**.** *Let $T : X \to Y$ be a bounded linear operator from a Banach space X* onto *another Banach space Y. Then T is an open map, i.e. images of open sets are open.*

Proof. Let U be an open subset of X. It suffices to show that $T(U)$ is a neighborhood of Tx for all $x \in U$. Furthermore, by linearity, we have $T(U) = Tx+T(-x+U)$. Hence it suffices to show that TU is a neighborhood of the origin when U is a neighborhood of the origin. In other words, it suffices to show that, for all $r > 0$, $T(B_X(0, r))$ contains some ball $B_Y(0, r')$. But by homothety, $T(B_X(0, r)) = rT(B_X(0, 1))$, it is enough to show the above for *some* $r > 0$.

Since T is onto, we have $Y = \cup_{n=1}^{\infty} \overline{T(B_X(0, n))}$. By the Baire category theorem, there is some n_0 such that $\overline{T(B_X(0, n_0))}$ has non-empty interior. Pick $B_Y(y_0, r_0) \subset \overline{T(B_X(0, n_0))}$. Observe that, by linearity, $\overline{T(B_X(0, n_0))}$ is symmetric and convex. It follows that $B_Y(-y_0, r_0) \subset \overline{T(B_X(0, n))}$ and so $B_Y(0, r_0) \subset \overline{T(B_X(0, n_0))}$. To conclude, we show that $\overline{T(B_X(0, n_0))} \subset T(B_X(0, 3n_0))$.

Let $y \in \overline{T(B_X(0, n_0))}$. Select $x_1 \in B_X(0, n_0)$ such that

$$\|y - Tx_1\|_Y < \frac{1}{2} r_0.$$

Then $y - Tx_1 \in B_Y(0, \frac{1}{2}r_0) \subset \overline{T(B_X(0, \frac{1}{2}n_0))}$. Pick $x_2 \in B_X(0, \frac{1}{2}n_0)$ such that

$$\|y - Tx_1 - Tx_2\|_Y < \frac{1}{4} r_0.$$

By induction, we select $x_k \in B_X(0, 2^{1-k}n_0)$ such that

$$\Big\| y - \sum_{j=1}^{k} Tx_j \Big\|_Y < \frac{1}{2^k} r_0.$$

It is readily seen that $x = \sum_{k=1}^{\infty} x_k$ exists, belongs to $B_X(0, 3n_0)$ and satisfies $y = \sum_{k=1}^{\infty} Tx_k = Tx$. This shows that $y \in T(B_X(0, 3n_0))$. □

An immediate consequence is:

Theorem 2.3.2 (Inverse mapping theorem)**.** *A bounded bijective linear operator of a Banach space onto another has a bounded inverse.*

Proof. Exercise. □

Example 2.3.3. *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 $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$. Then the two norms are equivalent, i.e. there is a constant C' such that $\|x\|_2 \leq C'\|x\|_1$ for all $x \in X$.*

Another consequence of the inverse mapping theorem is:

Theorem 2.3.4. *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. It suffices to show only one direction, as $T^{**} = T$. Suppose that $W = T^*Y$ is closed in X. Let $Z = \overline{TX} \subset Y$. Then T maps X into Z. Let us rename this map S, i.e. $S \in \mathscr{B}(X, Z)$ and $Sx = Tx$ for all $x \in X$. The adjoint S^* of S is an operator from Z to X. By Proposition 1.4.9, $Z = \overline{\operatorname{Im} S} = (\operatorname{Ker} S^*)^\perp$, so $\operatorname{Ker} S^* = \{0\}$, i.e. S^* is injective.

We claim that $\operatorname{Im} S^* = W$. To this end we let P be the orthogonal projection from Y onto Z and compute, for $x \in X$ and $y \in Y$,

$$\langle Tx, y\rangle_Y = \langle Sx, Py\rangle_Y = \langle x, S^*Py\rangle_X.$$

This shows that $T^* = S^* \circ P$, and so $\operatorname{Im} S^* = W$, as claimed.

So, S^* can be regarded as a bounded bijective linear operator between between Z and W. To make the notation clearer, we rename it as $V \in \mathscr{B}(Z, W)$, $Vz = S^*z$ for all $z \in Z$. By the inverse mapping theorem, V has a bounded inverse $V^{-1} \in \mathscr{B}(W, Z)$. This implies that V^* is invertible and $(V^*)^{-1} = (V^{-1})^* \in \mathscr{B}(Z, W)$ (cf. Proposition 1.4.1).

To conclude, we show that $T \circ (V^*)^{-1} = I_Z$. This implies that $TX \supset Z$ and so $TX = Z = \overline{TX}$ which gives the conclusion. Indeed, pick an arbitrary $z \in Z$, and let $w = (V^*)^{-1}z$. We compute, for $y \in Y$:

$$\langle Tw, y\rangle_Y = \langle Sw, y\rangle_Y = \langle w, S^*y\rangle_X = \langle w, Vy\rangle_X = \langle V^*w, y\rangle_Y = \langle z, y\rangle_Y.$$

Since this holds for all $y \in Y$, we deduce that $Tw = z$ and so $T \circ (V^*)^{-1} = I_Z$ as desired. □

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 $T = P_2 \circ P_1^{-1}$. □

Remark 2.4.2. *Usually, to show that a map A from a normed vector space X to another normed vector space Y is continuous, one needs to show that if $x_n \to x$, then $A(x_n) \to A(x)$. In many situations, one struggles to prove some kind of convergence for $A(x_n)$, let alone the convergence to $A(x)$. Nevertheless, if X and Y are Banach spaces and if A is linear, by virtue of the closed graph theorem, one may assume from the beginning that $A(x_n)$ is convergent in the sense of norm!*

Example 2.4.3. *Let X be a Banach space, and Y and Z are closed subspaces of X such that $X = Y \oplus Z$. Then the direct sum projection $P : X \to Y$ from X onto the first summand Y is bounded.*

Proof. By the closed graph theorem, it suffices to show that if $x_n \to x$ and $Px_n \to y$, then $y = Px$. Let $y_n = Px_n \in Y$ and $z_n = x_n - y_n$. Since Y is closed, $y \in Y$. Also $z_n \to x - y \in Z$. This implies $x = y + z$ and $Px = y$, as desired. □

Example 2.4.4. *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 2.4.5. *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,

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

We claim that

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

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

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

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. $\square$

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

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^*:

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

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

Now for each $\ell \in X^*$, $\ell(x_n)$ is convergent, and hence bounded. The principle of uniform boundedness thus implies that $\|T_n\|$ is bounded. (Note that X^* is complete regardless whether X is complete or not.) 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*

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

Proof. By a result in B4.1 (which is a consequence of the Hahn-Banach theorem), there is some $\ell \in X^*$ such that

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

The conclusion follows from the inequality

$$|\ell(x_n)| \leq \|\ell\|_* \|x_n\| = \|x_n\|$$

and the fact that $\ell(x_n) \to \ell(x) = \|x\|$. □

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 $x \in K$.*

We assume for granted the following result, which can be obtained as a consequence of the Hahn-Banach theorem.

Theorem 3.2.4 (Extended hyperplane separation theorem)**.** *Let X be a (normed) vector space, A and B be disjoint convex subsets of X, at least one of which has an interior point. Then A and B can be separated by a hyperplane, i.e. there is a non-zero linear function ℓ and a number c such that*

$$Re\, \ell(x) \leq c \leq Re\, \ell(y) \text{ for all } x \in A, y \in B.$$

In the above theorem, there is no need to assume that X is equipped with a norm, though in such case one needs to clarify what an interior point of a set means. This will not be discussed in this set of notes. A closely related theorem is the following.

Theorem 3.2.5 (Hyperplane separation theorem)**.** *Let X be a (normed) vector space, A be a nonempty convex subsets of X such that all points of A are interior points. Then A and any other point y not in A can be separated by a hyperplane, i.e. there is a non-zero linear function ℓ and a number c (which might depend on y) such that*

$$Re\, \ell(x) < c = Re\, \ell(y) \text{ for all } x \in A.$$

Proof of Mazur's theorem. We will only consider the case when X is real. The complex case is left as an exercise.

Assume by contradiction that $x \notin K$. Since K is closed, there is some $r > 0$ such that $B(x, r) \cap K = \emptyset$. By the extended hyperplane separation theorem, there is a non-zero linear functional ℓ_0 and a number $c \in \mathbb{R}$ such that

$$\ell_0(y) \leq c \leq \ell_0(z) \text{ for all } y \in K \text{ and } z \in B(x, r). \tag{3.2}$$

Note that the second half of (3.2) implies that, for all $w \in B(0, r)$,

$$\ell_0(w) = \ell_0(x + w) - \ell_0(x) \geq c - \ell_0(x),$$

and so ℓ_0 is a bounded linear functional.

The left half of (3.2) implies that $\ell_0(x_n) \leq c$ and so by the weak convergence of (x_n) to x, we have $\ell_0(x) \leq c$. Returning to (3.2), we obtain

$$\ell_0(x) \leq \ell_0(z) \text{ for all } z \in B(x, r).$$

By linearity, this implies that

$$\ell_0(w) = \frac{1}{r}(\ell_0(x + rw) - \ell_0(x)) \geq 0 \text{ for all } w \in B(0, 1).$$

This is impossible as $\ell_0 \neq 0$. □

3.3 Weak sequential compactness

Definition 3.3.1. *A subset A of a Banach space X is called weakly sequentially compact if every sequence of A has a subsequence weakly convergent to a point of A.*

Recall that a Banach space is said to be reflexive if it is isometrically isomorphic to its second dual. The following theorem is a version of the Bolzano-Weierstrass lemma in infinite dimensional setting.

Theorem 3.3.2 (Weak sequential compactness in reflexive Banach spaces)**.** *The closed unit ball of a reflexive Banach space is weakly sequentially compact. In particular, every bounded sequence in a reflexive Banach space has a weakly convergent subsequence.*

Proof. We will only prove the theorem in the case of Hilbert spaces, which are reflexive thanks to the Riesz representation theorem.

Let (x_n) be a sequence in the unit ball of a Hilbert space X. The proof uses a diagonal process to select a subsequence (x_{n_j}) of (x_n), such that $\langle x_{n_j}, x_m\rangle$ converges for every m.

To begin with, we note that the sequence $\langle x_n, x_1\rangle$ is bounded. By the Bolzano-Weierstrass lemma, we can extract a subsequence $n_j^{(1)}$ such that $\langle x_{n_j^{(1)}}, x_1\rangle$ is convergent.

We then consider $\langle x_{n_j^{(1)}}, x_2\rangle$ and select a convergent subsequence $\langle x_{n_j^{(2)}}, x_2\rangle$. Clearly, $\langle x_{n_j^{(2)}}, x_1\rangle$ is also convergent.

Proceeding in this way, we constructed nested subsequence $(n_j^{(k)})$ such that $\langle x_{n_j^{(k)}}, x_m\rangle$ is convergent (with respect to j) for every $m \le k$.

Let $x_{n_j} = x_{n_j^{(j)}}$. Note that, for every fixed m, $(n_j)_{j\ge m}$ is a subsequence of $(n_j^{(m)})_{j\ge m}$. It follows that $\langle x_{n_j}, x_m\rangle$ is convergent for every m.

Let Y and $\bar{Y}$ be respectively the linear span and the closed linear span of the x_n's. It is clear that, for every $y \in Y$, $\langle x_{n_j}, y\rangle$ is convergent. Using the estimate

$$|\langle x_{n_j} - x_{n'_j}, \bar{y}\rangle| \le |\langle x_{n_j} - x_{n'_j}, y\rangle| + |\langle x_{n_j} - x_{n'_j}, \bar{y} - y\rangle| \le |\langle x_{n_j} - x_{n'_j}, y\rangle| + 2\|\bar{y} - y\|$$

for $\bar{y} \in \bar{Y}$ and $y \in Y$, it is readily seen that $\langle x_{n_j}, \bar{y}\rangle$ is Cauchy and thus convergent for every $\bar{y} \in \bar{Y}$.

On the other hand, it is clear that $\langle x_{n_j}, z\rangle = 0$ for all $z \in Y^\perp$. Hence, as $X = \bar{Y} \oplus Y^\perp$ by the projection theorem, we have that $\langle x_{n_j}, x\rangle$ is convergent for all $x \in X$.

Define

$$\ell(x) = \lim_{j\to\infty} \langle x, x_{n_j}\rangle, \qquad x \in X.$$

It is clear that ℓ is a bounded linear functional on X. By the Riesz representation theorem, there is some $x_* \in X$ such that $\ell(x) = \langle x, x_*\rangle$ for all $x \in X$. By the Riesz representation theorem again, this implies $x_{n_j} \rightharpoonup x_*$. □

We note that the converse of Theorem 3.3.2 is true, a result which we will not prove.

Theorem 3.3.3 (Eberlein)**.** *The closed unit ball in a Banach space X is weakly sequentially compact* only if *X is reflexive.*

As an application of Theorem 3.3.2, we obtain the following generalization of Theorem 1.2.4 for Banach spaces.

Theorem 3.3.4 (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. □

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

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

(This can easily be re-expressed in the familiar trigonometric series using $e^{inx} = \cos nx + i \sin nx$.)

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

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

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

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

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 $\tilde{f}$ be the limit of $S_N f$, then the Fourier coefficients of $f - \tilde{f}$ 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 $|g| \le 1$ in $(-\pi, \pi) \setminus (x_0 - \delta, x_0 + \delta)$. This implies that, for any $n > 0$,

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

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*

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

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 $|f(x_0 + h)| \leq A|h|^\alpha$ 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. □

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. □

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 $\|(\lambda I - T)z_n\| = \|x_n\|^{-1} \|y_n\| \to 0$. Hence $\lambda \in \sigma_{ap}(T)$. □

In the rest of the chapter, we will specialize 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, $\langle Tx_n, x_n\rangle\lambda$. 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 $\|T^n\|^{1/n}$. □

Theorem 5.2.6. *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*

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

Furthermore, the endpoints a and b belong to the spectrum of T.

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

We next show that $a, b \in \sigma(T)$. By definition of a, b, we have $|a| \le \|T\|$ and $|b| \le \|T\|$. But as $\sigma(T) \in [a, b]$, we have $\mathrm{rad}(\sigma(T)) \le \max(|a|, |b|)$. Hence at least one of a and b belongs to $\sigma(T)$. Now note that, if c is a real constant, then the spectrum of $cI + T$ is shifted by c and the "a" and "b" of $cI + T$ are also shifted by c. Applying what we just established to $cI + T$ for suitable c, we conclude that both a and b belong to $\sigma(T)$. □

Alternatively, we can show that $\sigma(T) \subset [a, b]$ as follows: It suffices to show that if λ is a real number such that $\lambda > b$ then $\lambda I - T$ is invertible. (A similar argument apply to $\lambda < a$.) We have

$$\langle x, (\lambda I - T)x \rangle = \lambda \|x\|^2 - \langle x, Tx \rangle \geq (\lambda - b)\|x\|^2.$$

It thus follows that $\langle x, y \rangle_\lambda := \langle x, (\lambda I - T)y \rangle$ defines a scalar product on X and its associated norm $\|x\|_\lambda := \langle x, (\lambda I - T)x \rangle^{1/2}$ is equivalent to $\| \cdot \|$.

For every $z \in X$, consider the linear functional

$$\ell_z(x) = \langle x, z \rangle.$$

By the Riesz representation theorem, there is some y depending on z such that

$$\ell_z(x) = \langle x, y \rangle_\lambda \text{ i.e. } \langle x, z \rangle = \langle x, (\lambda I - T)y \rangle \text{ for every } x.$$

It thus follows that $\lambda I - T$ is surjective. Since $\lambda I - T$ is self-adjoint, this implies that $\lambda I - T$ is also injective, and hence invertible.

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

Proposition 5.2.7. *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*

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