OPERATOR THEORY ON HILBERT SPACE

Class notes

John Petrovic

Contents

Chapter 1. Hilbert space 1

1.1. Definition and Properties 1

1.2. Orthogonality 3

1.3. Subspaces 7

1.4. Weak topology 9

Chapter 2. Operators on Hilbert Space 13

2.1. Definition and Examples 13

2.2. Adjoint 15

2.3. Operator topologies 17

2.4. Invariant and Reducing Subspaces 20

2.5. Finite rank operators 22

2.6. Compact Operators 23

2.7. Normal operators 27

Chapter 3. Spectrum 31

3.1. Invertibility 31

3.2. Spectrum 34

3.3. Parts of the spectrum 38

3.4. Spectrum of a compact operator 40

3.5. Spectrum of a normal operator 43

Chapter 4. Invariant subspaces 47

4.1. Compact operators 47

4.2. Line integrals 49

4.3. Invariant subspaces for compact operators 52

4.4. Normal operators 56

Chapter 5. Spectral radius algebras 64

5.1. Compact operators 64

CHAPTER 1

Hilbert space

1.1. Definition and Properties

In order to define Hilbert space $\mathcal{H}$ we need to specify several of its features. First, it is a **complex vector space** — the field of scalars is $\mathbb{C}$ (complex numbers). [See Royden, p. 217.] Second, it is an **inner product space**. This means that there is a *complex valued* function $\langle x, y\rangle$ defined on $\mathcal{H} \times \mathcal{H}$ with the properties that, for all $x, y, z \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$:

(a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$; it is linear in the first argument;

(b) $\langle x, y\rangle = \overline{\langle y, x\rangle}$; it is Hermitian symmetric;

(c) $\langle x, x\rangle \geq 0$; it is non-negative;

(d) $\langle x, x\rangle = 0$ iff $x = 0$; it is positive.

In every inner product space it is possible to define a norm as $\|x\| = \langle x, x\rangle^{1/2}$.

Exercise 1.1.1. Prove that this is indeed a norm.

Finally, Hilbert space is complete in this norm (meaning: in the topology induced by this norm).

Example 1.1.1. $\mathbb{C}^n$ is an inner product space with $\langle x, y\rangle = \sum_{k=1}^{n} x_k \overline{y_k}$ and, consequently, the norm $\|x\| = \sqrt{\sum_{k=1}^{n} |x_k|^2}$. Completeness: if $\{x^{(k)}\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{C}^n$ (here $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$) then so is $x_m^{(k)}$ for any fixed m, $1 \leq m \leq n$, and $\mathbb{C}$ is complete.

Example 1.1.2. Let $\mathcal{H}_0$ denote the collection of all complex sequences, i.e. functions $a : \mathbb{N} \to \mathbb{C}$, characterized by the fact that $a_n \neq 0$ for a finite number of positive integers n. Define the inner product on $\mathcal{H}_0$ by $\langle a, b\rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}$. The space $\mathcal{H}_0$ is not complete in the induced norm. Indeed, the sequence $\{a^{(k)}\}_{k\in\mathbb{N}}$, defined by $a_n^{(k)} = 1/2^n$ if $n \leq k$ and $a_n^{(k)} = 0$ if $n > k$ is a Cauchy sequence, but not convergent.

EXAMPLE 1.1.3. Let ℓ^2 denote the collection of all complex sequences $a = \{a_n\}_{n=1}^{\infty}$ such that $\sum_{n=1}^{\infty} |a_n|^2$ converges. Define the inner product on ℓ^2 by $\langle a, b\rangle = \sum_{n=1}^{\infty} a_n \overline{b_n}$. Suppose that $\{a^{(k)}\}_{k=1}^{\infty}$ is a Cauchy sequence in ℓ^2. Then so is $\{a_n^{(k)}\}_{k=1}^{\infty}$ for each n, hence there exists $a_n = \lim_{k\to\infty} a_n^{(k)}$. First we show that $a \in \ell^2$. Indeed, choose K so that for $k \geq K$ we have $\|a^{(k)} - a^{(K)}\| \leq 1$. Then, using Minkowski's Inequality for sequences (see Royden, p. 122), for any $N \in \mathbb{N}$,

$$\left\{\sum_{n=1}^{N} |a_n|^2\right\}^{1/2} \leq \left\{\sum_{n=1}^{N} |a_n - a_n^{(K)}|^2\right\}^{1/2} + \left\{\sum_{1=0}^{N} |a_n^{(K)}|^2\right\}^{1/2} = \lim_{k\to\infty} \left\{\sum_{n=1}^{N} |a_n^{(k)} - a_n^{(K)}|^2\right\}^{1/2} + \left\{\sum_{n=1}^{N} |a_n^{(K)}|^2\right\}^{1/2}$$

$$\leq \limsup_{k\to\infty} \|a^{(k)} - a^{(K)}\| + \|a^{(K)}\| \leq 1 + \|a^{(K)}\|.$$

Thus $a = \{a_n\} \in \ell^2$. Moreover, $\{a^{(k)}\}$ converges to a, i.e. $\lim_{k\to\infty} \|a - a^{(k)}\| = 0$. Let $\epsilon > 0$ and choose M so that $k, j \geq M$ implies that $\|a^{(k)} - a^{(j)}\| < \epsilon$. For such $k \geq M$ and any N, we have

$$\sum_{n=1}^{N} |a_n - a_n^{(k)}|^2 = \lim_{j\to\infty} \sum_{n=1}^{N} |a_n^{(j)} - a_n^{(k)}|^2 \leq \limsup_{j\to\infty} \|a^{(j)} - a^{(k)}\|^2 \leq \epsilon^2.$$

Since N is arbitrary, it follows that $\|a - a^{(k)}\| \leq \epsilon$ and, therefore, ℓ^2 is Hilbert space.

EXAMPLE 1.1.4. The space L^2 of functions $f : X \to \mathbb{C}$, such that $\int_X |f|^2 d\mu < \infty$ (where X is usually $[0, 1]$ and μ Lebesgue measure). The inner product is defined by $\langle f, g\rangle = \int_X f\overline{g}\, d\mu$ and L^2 is complete by the Riesz–Fisher Theorem (see Royden, p. 125).

EXAMPLE 1.1.5. The space H^2. Let $X = \mathbb{T}$ (the unit circle) and μ the normalized Lebesgue measure on $\mathbb{T}$. The *Hardy space* H^2 consists of those functions in $L^2(\mathbb{T})$ such that $\langle f, e^{int}\rangle = 0$ for $n = -1, -2, \ldots$.

Some important facts.

PROPOSITION 1.1.1 (Parallelogram Law). $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

PROPOSITION 1.1.2 (Polarization Identity). $4\langle x, y\rangle = \langle x + y, x + y\rangle - \langle x - y, x - y\rangle + i\langle x + iy, x + iy\rangle - i\langle x - iy, x - iy\rangle$.

EXERCISE 1.1.2. Prove Propositions 1.1.1 and 1.1.2.

PROBLEM 1. Let $\|\cdot\|$ be a norm on Banach space $\mathcal{X}$, and define $\langle x, y\rangle$ as in Polarization Identity. Assuming that the norm satisfies the Parallelogram Law, prove that $\langle x, y\rangle$ defines an inner product.

1.2. Orthogonality

In Linear Algebra a basis of a vector space is defined as a minimal spanning set. In Hilbert space such a definition is not very practical. It is hard to speak of minimality when a basis can be infinite. In fact, a basis can be uncountable, so if $\{e_i\}_{i\in I}$ is such a basis, what is the meaning of $\sum_{i\in I} x_i e_i$?

DEFINITION 1.2.1. An *orthonormal* subset of Hilbert space $\mathcal{H}$ is a set $\mathcal{E}$ such that (a) $\|e\| = 1$, for all $e \in \mathcal{E}$; (b) if $e_1, e_2 \in \mathcal{E}$ and $e_1 \neq e_2$ then $\langle e_1, e_2\rangle = 0$. An *orthonormal basis* in $\mathcal{H}$ is a maximal orthonormal set. We use abbreviations o.n.s. and o.n.b. for orthonormal set and orthonormal basis, respectively.

THEOREM 1.2.1. *Every Hilbert space has an orthonormal basis.*

PROOF. Let e be a unit vector in $\mathcal{H}$. Then $\mathcal{E} = \{e\}$ is an orthonormal set. Let $\mathcal{M}$ be the collection of all orthonormal sets in $\mathcal{H}$ that contain $\mathcal{E}$. By the Hausdorff Maximal Principle (Royden, p.25) there exists a maximal chain $\mathcal{C}$ of such orthonormal sets, partially ordered by inclusion. Let N be the union of all elements of $\mathcal{C}$. Then N is a maximal orthonormal set, hence a basis of $\mathcal{H}$. □

If the set $\{e\}$ is replaced by any orthonormal set, the same proof yields a stronger result.

THEOREM 1.2.2. *Every orthonormal set in Hilbert space can be extended to an orthonormal basis.*

EXAMPLE 1.2.1. For $k \in \mathbb{N}$, let e_k denote the sequence with only one non-zero entry, lying in the kth position and equal to 1. The set $\{e_k\}_{k\in\mathbb{N}}$ is an o.n.b. for ℓ^2. (If a vector $x \in \ell^2$ is orthogonal to all e_k, then each of its components is zero, so $x = 0$.)

EXAMPLE 1.2.2. The set $\{e_1, e_3, e_5, \dots\}$ is an orthonormal set in ℓ^2 but not a basis.

EXAMPLE 1.2.3. The set $\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\sin t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \frac{\sin 2t}{\sqrt{\pi}}, \dots\right\}$ is an o.n.b. in $L^2(-\pi, \pi)$.

EXAMPLE 1.2.4. The set $\left\{\frac{1}{\sqrt{2\pi}}\, e^{int} : n \in \mathbb{Z}\right\}$ is another o.n.b. in $L^2(-\pi,\pi)$.

In Linear Algebra, if $\{e_i\}_{i\in I}$ is an o.n.b. then every vector x can be written as $\sum_{i\in I}\langle x, e_i\rangle e_i$. In Hilbert space our first task is to make sense of this sum since the index set I need not be countable.

THEOREM 1.2.3 (Bessel's Inequality). *Let $\{e_i\}_{i=1}^k$ be an o.n.s. in $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $\sum_{i=1}^k |\langle x, e_i\rangle|^2 \leq \|x\|^2$.*

PROOF. If we write $x_i = \langle x, e_i\rangle$, then

$$0 \leq \|x - \sum_{i=1}^k x_i e_i\|^2 = \langle x - \sum_{i=1}^k x_i e_i, x - \sum_{i=1}^k x_i e_i\rangle = \|x\|^2 - 2\mathrm{Re}\langle x, \sum_{i=1}^k x_i e_i\rangle + \langle \sum_{i=1}^k x_i e_i, \sum_{j=1}^k x_j e_j\rangle$$

$$= \|x\|^2 - 2\mathrm{Re}\sum_{i=1}^k \overline{x_i}\langle x, e_i\rangle + \sum_{i=1}^k\sum_{j=1}^k x_i\overline{x_j}\langle e_i, e_j\rangle = \|x\|^2 - 2\mathrm{Re}\sum_{i=1}^k \overline{x_i}x_i + \sum_{i=1}^k x_i\overline{x_i} = \|x\|^2 - \sum_{i=1}^k |x_i|^2.$$

□

COROLLARY 1.2.4. *Let $\mathcal{E} = \{e_i\}_{i\in I}$ be an o.n.s. in $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $\langle x, e_i\rangle \neq 0$ for at most a countable number of $i \in I$.*

PROOF. Let $x \in \mathcal{H}$ be fixed and let $\mathcal{E}_n = \{e_i : |x_i| \geq 1/n\}$. If $e_{i_1}, e_{i_2}, \ldots, e_{i_k} \in \mathcal{E}_n$ then

$$\|x\|^2 \geq \sum_{j=1}^k |x_{i_j}|^2 \geq k(1/n^2).$$

So, for each $n \in \mathbb{N}$, $\mathcal{E}_n$ is a finite set, and $\mathcal{E} = \cup_n \mathcal{E}_n$. □

In view of Corollary 1.2.4 the expressions like $\sum\langle x, e_i\rangle e_i$ turn out to be the usual infinite series. Our next task is to establish their convergence. The following Lemma will be helpful in this direction.

LEMMA 1.2.5. *If $\{x_i\}_{i\in\mathbb{N}}$ is a sequence of complex numbers and $\{e_i\}_{i\in\mathbb{N}}$ is an o.n.s. in $\mathcal{H}$, then the series $\sum_{i\in\mathbb{N}} x_i e_i$ and $\sum_{i\in\mathbb{N}} |x_i|^2$ are equiconvergent.*

PROOF. Let s_m and σ_m denote the partials sums of $\sum_{i\in\mathbb{N}} x_i e_i$ and $\sum_{i\in\mathbb{N}} |x_i|^2$, respectively. Then

$$\|s_m - s_n\|^2 = \|\sum_{i=n+1}^m x_i e_i\|^2 = \langle\sum_{i=n+1}^m x_i e_i, \sum_{j=n+1}^m x_j e_j\rangle = \sum_{i=n+1}^m |x_i|^2 = |\sigma_m - \sigma_n|$$

so the series are equiconvergent. □

Now we can establish the convergence of $\sum_{i\in I}\langle x, e_i\rangle e_i$. We will use notation $x_i = \langle x, e_i\rangle$ for the Fourier coefficients of $x \in \mathcal{H}$ relative to the fixed basis $\{e_i\}_{i\in I}$.

COROLLARY 1.2.6 (Parseval's Identity). *Let $\{e_i\}_{i\in I}$ be an o.n.s. in $\mathcal{H}$, and let $x \in \mathcal{H}$. Then the series $\sum_{i\in I} x_i e_i$ and $\sum_{i\in I} |x_i|^2$ converge and $\| \sum_{i\in I} x_i e_i\|^2 = \sum_{i\in I} |x_i|^2$.*

PROOF. Since only a countable number of terms in each series is non-zero, we can rearrange them and consider the series $\sum_{i=1}^\infty x_i e_i$ and $\sum_{i=1}^\infty |x_i|^2$. The latter series converges by the Bessel's Inequality and Lemma 1.2.5 implies that the former series converges too. Moreover, their partial sums s_m and σ_m satisfy $\|s_m\| = \sigma_m$, so the last assertion of the corollary follows by letting m go to ∞. □

Now we are in the position to show that, in Hilbert space, every o.n.b. indeed spans $\mathcal{H}$. Of course, the minimality is a direct consequence of the definition.

THEOREM 1.2.7. *Let $\mathcal{E} = \{e_i\}_{i\in I}$ be an o.n.b. in $\mathcal{H}$. Then, for each $x \in \mathcal{H}$, $x = \sum_{i\in I} x_i e_i$, where $x_i = \langle x, e_i\rangle$.*

PROOF. Let $x_i = \langle x, e_i\rangle$ and $y = x - \sum_{i\in I} x_i e_i$. (Well defined since the series converges.) Then $\langle y, e_k\rangle = \langle x, e_k\rangle - \langle\sum_{i\in I} x_i e_i, e_k\rangle = 0$, for each $k \in I$, so $y \perp \mathcal{E}$. If $y \neq 0$, then $\mathcal{E} \cup \{y/\|y\|\}$ is an o.n.s., contradiciting the maximality of $\mathcal{E}$, so $y = 0$. □

The following is the analogue of a well known Linear Algebra fact. We use notation card I for the cardinal number of the set I.

THEOREM 1.2.8. *Any two orthonormal bases $\{e_i\}_{i\in I}$ and $\{f_j\}_{j\in J}$ in $\mathcal{H}$ have the same cardinal number.*

PROOF. We will assume that both cardinal numbers are infinite. If either of them is finite, one knows from Linear Algebra that the other one is finite and equal to the first. Let $j \in J$ be fixed and let $I_j = \{i \in I : \langle f_j, e_i\rangle \neq$

$0\}$. By Corollary 1.2.4, I_j is at most countable. Further, $\cup_{j\in J} I_j = I$. Indeed, if $i_0 \in I \setminus \cup_{j\in J} I_j$ then $\langle f_j, e_{i_0}\rangle = 0$ for all $j \in J$ so it would follow that $e_{i_0} = 0$. Since $\operatorname{card} I_j \leq \aleph_0$ we see that $\operatorname{card} I \leq \operatorname{card} J \cdot \aleph_0 = \operatorname{card} J$. Similarly, $\operatorname{card} J \leq \operatorname{card} I$. By Cantor–Bernstein Theorem, (see, e.g., "Proofs from the book", p.90) $\operatorname{card} I = \operatorname{card} J$. □

DEFINITION 1.2.2. The *dimension* of Hilbert space $\mathcal{H}$, denoted by $\dim \mathcal{H}$, is the cardinal number of a basis of $\mathcal{H}$.

In this course we will assume that $\dim \mathcal{H} \leq \aleph_0$.

EXERCISE 1.2.1. If $\mathcal{H}$ is an infinite dimensional Hilbert space, then $\mathcal{H}$ is separable iff $\dim \mathcal{H} = \aleph_0$. [Given a countable basis, use rational coefficients. Given a countable dense set, approximate each element of a basis close enough to exclude all other basis elements.]

Next, we want to address the question: when can we identify two Hilbert spaces? We need a vector space isomorphism (i.e., a linear bijection) that preserves the inner product.

DEFINITION 1.2.3. If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, an isomorphism is a linear surjection $U : \mathcal{H} \to \mathcal{K}$ such that, for all $x, y \in \mathcal{H}$, $\langle Ux, Uy\rangle = \langle x, y\rangle$. In this situation we say that $\mathcal{H}$ and $\mathcal{K}$ are isomorphic.

EXERCISE 1.2.2. Prove that $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all $x, y \in \mathcal{H}$ iff $\|Ux\| = \|x\|$ for all $x \in \mathcal{H}$. Conclude that a Hilbert space isomorphism is injective.

THEOREM 1.2.9. *Every separable Hilbert space of infinite dimension is isomorphic to ℓ^2. Every Hilbert space of finite dimension n is isomorphic to $\mathbb{C}^n$.*

PROOF. We will assume that $\mathcal{H}$ is an infinite dimensional Hilbert space and leave the finite dimensional case as an exercise. Since $\mathcal{H}$ is separable, there exists an o.n.b. $\{e_n\}_{n=1}^{\infty}$. For $x \in \mathcal{H}$, let $x_i = \langle x, e_i\rangle$ and $U(x) = (x_1, x_2, x_3, \dots)$. By Parseval's Identity, the series $\sum_{i=1}^{\infty} |x_i|^2$ converges, so the sequence $(x_1, x_2, x_3, \dots)$ belongs to ℓ^2. Thus U is well-defined, linear (because the inner product is linear in the first argument), and isometric:

$\|Ux\|^2 = \sum_{i=1}^\infty |x_i|^2 = \|x\|^2$. Finally, if $(y_1, y_2, y_3, \dots) \in \ell^2$ then $\sum_{i=1}^\infty |y_i|^2$ converges so, by Lemma 1.2.5, $\sum_{n=1}^\infty y_n e_n$ converges and $U(\sum_{n=1}^\infty y_n e_n) = (y_1, y_1, y_1, \dots)$. Thus, U is surjective and the theorem is proved. $\square$

EXERCISE 1.2.3. Prove that every Hilbert space of finite dimension n is isomorphic to $\mathbb{C}^n$.

PROBLEM 2. Let $\mathcal{H}$ ne a separable Hilbert space and $\mathcal{M}$ a subspace of $\mathcal{H}$. Prove that $\mathcal{M}$ is a separable Hilbert space.

PROBLEM 3. The Haar system $\{\varphi_{m,n}\}$, $m \in \mathbb{N}$, $1 \leq n \leq 2^m$, is defined as:

$$\varphi_{m,n}(x) = \begin{cases} 2^{m/2}, & \text{if } \dfrac{n-1}{2^m} \leq x \leq \dfrac{n-1/2}{2^m}, \\ -2^{m/2}, & \text{if } \dfrac{n-1/2}{2^m} \leq x \leq \dfrac{n}{2^m}, \\ 0, & \text{if } x \notin \left[\dfrac{n-1}{2^m}, \dfrac{n}{2^m}\right). \end{cases}$$

Prove that this system is an o.n.b. of $L^2[0,1]$.

1.3. Subspaces

EXAMPLE 1.3.1. Let $\mathcal{H} = L^2[0,1]$ and let G be a measurable subset of $[0,1]$. Denote by $L^2(G)$ the set of functions in L^2 that vanish outside of G. Then $L^2(G)$ is a closed subspace of $\mathcal{H}$. Further, if $f \in L^2(G)$ and $g \in L^2(G^c)$, then $\langle f, g \rangle = 0$.

DEFINITION 1.3.1. If $\mathcal{M}$ is a closed subspace of the Hilbert space $\mathcal{H}$, then the *orthogonal complement* of $\mathcal{M}$, denoted $\mathcal{M}^\perp$, is the set of vectors in $\mathcal{H}$ orthogonal to every vector in $\mathcal{M}$.

EXERCISE 1.3.1. Prove that $\mathcal{M}^\perp$ is a closed subspace of $\mathcal{H}$.

THEOREM 1.3.1. *Let $\mathcal{M}$ be a closed subspace of Hilbert space $\mathcal{H}$, and let $x \in \mathcal{H}$. Then there exist unique vectors y in $\mathcal{M}$ and z in $\mathcal{M}^\perp$ so that $x = y + z$.*

PROOF. Let $\{e_i\}_{i\in I}$ and $\{f_j\}_{j\in J}$ be orthonormal bases for $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively. Their union is an o.n.b. of $\mathcal{H}$ so $x = \sum_{i\in I}\langle x, e_i\rangle e_i + \sum_{j\in J}\langle x, f_j\rangle f_j$ and we define $y = \sum_{i\in I}\langle x, e_i\rangle e_i$, $z = \sum_{j\in J}\langle x, f_j\rangle f_j$. Then $y \in \mathcal{M}$, $z \in \mathcal{M}^\perp$, and $x = y + z$.

Suppose now that $x = y_1 + z_1 = y_2 + z_2$, where $y_1, y_2 \in \mathcal{M}$ and $z_1, z_2 \in \mathcal{M}^\perp$. Then $y_1 - y_2 = z_2 - z_1$ belongs to both $\mathcal{M}$ and $\mathcal{M}^\perp$, so $\langle y_1 - y_2, y_1 - y_2\rangle = 0$ and it follows that $y_1 = y_2$, and consequently $z_1 = z_2$. □

DEFINITION 1.3.2. In the situation described in Theorem 1.3.1 we say that $\mathcal{H}$ is the *orthogonal direct sum* of $\mathcal{M}$ and $\mathcal{M}^\perp$, and we write $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. When $z = x + y$ with $x \in \mathcal{M}$ and $y \in \mathcal{M}^\perp$ we often write $z = x \oplus y$.

Theorem 1.3.1 allows us to define a map $P : \mathcal{H} \to \mathcal{M}$ by $Px = y$. It is called the *orthogonal projection* of $\mathcal{H}$ onto $\mathcal{M}$, and it is denoted by $P_{\mathcal{M}}$. Here are some of its properties.

THEOREM 1.3.2. *Let $\mathcal{M}$ be a closed subspace of Hilbert space $\mathcal{H}$ and let P be the orthogonal projection on $\mathcal{M}$. Then:*

(a) *P is a linear transformation;*

(b) *$\|Px\| \le \|x\|$, for all $x \in \mathcal{H}$;*

(c) *$P^2 = P$;*

(d) $\operatorname{Ker} P = \mathcal{M}^\perp$ *and* $\operatorname{Ran} P = \mathcal{M}$.

PROOF. Let $\{e_i\}_{i\in I}$ and $\{f_j\}_{j\in J}$ be orthonormal bases for $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively, and let $Q = I - P$ be the orthonormal projection on $\mathcal{M}^\perp$. If $x', x'' \in \mathcal{H}$ and $\alpha', \alpha'' \in \mathbb{C}$, then $P(\alpha' x' + \alpha'' x'') = \sum_{i\in I}\langle \alpha' x' + \alpha'' x'', e_i\rangle = \alpha' Px' + \alpha'' Px''$, so (a) holds.

(b) If $x \in \mathcal{H}$, then $x = Px + Qx$ and $Px \perp Qx$. Therefore, $\|x\|^2 = \|Px\|^2 + \|Qx\|^2 \ge \|Px\|^2$.

(c) If $y \in \mathcal{M}$ then $Py = y$. Now, for any $x \in \mathcal{H}$, $Px \in \mathcal{M}$ so $P^2x = P(Px) = Px$.

(d) If $Px = 0$ then $x = Qx \in \mathcal{M}^\perp$. If $x \in \mathcal{M}^\perp$ then $Qx = x$ by (c), so $Px = 0$. The other assertion is obvious. □

PROBLEM 4. Prove that $P_{\mathcal{M}}x$ is the unique point in $\mathcal{M}$ that is nearest to x, meaning that $\|x - P_{\mathcal{M}}x\| = \inf\{\|x - h\| : h \in \mathcal{M}\}$.

PROBLEM 5. In $L^2[0,1]$ find the orthogonal complement to the subspace consisting of:

(a) all polynomials in x;

(b) all polynomials in x^2;

(c) all polynomials in x with the free term equal to 0;

(d) all polynomials in x with the sum of coefficients equal to 0.

PROBLEM 6. If $\mathcal{M}$ and $\mathcal{N}$ are subspaces of Hilbert space that are orthogonal to each other, then the sum $\mathcal{M}+\mathcal{N}=\{x+y : x\in\mathcal{M}, y\in\mathcal{N}\}$ is a subspace. Show that the theorem is not true if $\mathcal{M}$ and $\mathcal{N}$ are either: closed but not orthogonal or orthogonal but not closed.

1.4. Weak topology

Read Royden, page 236–238.

EXAMPLE 1.4.1. Consider the sequence of functions $\{\cos nt\}_{n\in\mathbb{N}}$ in $L^1[0,2\pi]$. It is easy to see that this sequence is not convergent. However, for any function $f\in L^\infty$, $\int_0^1 f(t)\cos nt\,dt\to 0$ as $n\to\infty$. Since L^∞ is the dual space of L^1, we say that $\cos nt\to 0$ weakly, and we write $w-\lim_n \cos nt = 0$.

EXAMPLE 1.4.2. Consider the sequence of functions $\{\cos nt\}_{n\in\mathbb{N}}$ in $L^\infty[0,2\pi]$. Notice that, while not a convergent sequence, if $f\in L^1$ then $\int_0^1 f(t)\cos nt\,dt\to 0$ as $n\to\infty$. Since L^∞ is the dual space of L^1, we say that $\cos nt\to 0$ in the weak* topology.

In a Banach space $\mathcal{X}$ it is useful to consider three topologies: the *norm topology*, induced by the norm; *weak topology* — the smallest topology in which all bounded linear functionals on $\mathcal{X}$ are continuous; *weak** *topology* (meaningful when $\mathcal{X}$ is the dual space of $\mathcal{Y}$ so that $\mathcal{Y}\subset\mathcal{X}^*$) — the smallest topology in which some bounded linear functionals on $\mathcal{X}$ are continuous (those that can be identified as elements of $\mathcal{Y}$). In order to dicuss these topologies (and understand their role), we need to find out what bounded linear functionals on Hilbert space $\mathcal{H}$ look like.

THEOREM 1.4.1 (Riesz Representation Theorem). *If L is a bounded linear functional on $\mathcal{H}$, then there is a unique vector $y \in \mathcal{H}$ such that $L(x) = \langle x, y \rangle$ for every $x \in \mathcal{H}$. Moreover, $\|L\| = \|y\|$.*

PROOF. Assuming that such y exists, we can write it as $y = \sum_{i \in \mathbb{N}} y_i e_i$ relative to a fixed o.n.b. $\{e_i\}_{i \in \mathbb{N}}$. Then $y_i = \langle y, e_i \rangle = \overline{\langle e_i, y \rangle} = \overline{L(e_i)}$. Therefore, we define $y = \sum_{i \in \mathbb{N}} \overline{L(e_i)}\, e_i$, and all it remains to prove is the convergence of the series. Let $s_n = \sum_{i=1}^{n} \overline{L(e_i)}\, e_i$. Then $L(s_n) = \sum_{i=1}^{n} \overline{L(e_i)}\, Le_i = \|s_n\|^2$, so $\|s_n\|^2 \leq \|L\| \|s_n\|$ from which it follows that $\|s_n\| \leq \|L\|$. Thus the series $\sum_{i=1}^{n} \overline{L(e_i)}\, e_i$ converges and the result follows from Lemma 1.2.5. □

We see that if $L \in \mathcal{H}^*$, the dual space of $\mathcal{H}$, then $L = L_y$. The mapping $\Phi : \mathcal{H} \to \mathcal{H}^*$ defined by $\Phi(y) = L_y$ is a norm preserving surjection. It is conjugate linear: $\Phi(\alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1} y_1 + \overline{\alpha_2} y_2$. Nevertheless, we identify $\mathcal{H}^*$ with $\mathcal{H}$. Consequently, $\mathcal{H}$ is reflexive (i.e., $\mathcal{H}^{**} = \mathcal{H}$) so the weak* and weak topologies on $\mathcal{H}$ coincide. Therefore, we will work with 2 topologies: weak and norm induced. The absence of a qualifier will always mean that it is the latter.

EXERCISE 1.4.1. Prove that the weak topology is weaker than the norm toplogy, i.e., if G is a weakly open set then G is an open set.

EXAMPLE 1.4.3. If $\{e_n\}_{n \in \mathbb{N}}$ is an orthonormal sequence in $\mathcal{H}$ then $w - \lim e_n = 0$ but the sequence is not convergent.

EXERCISE 1.4.2. Prove that the Hilbert space norm is continuous but not weakly continuous.

The following result shows why weak topology is important. [See Royden, p. 237]

THEOREM 1.4.2 (Banach-Alaoglu). *The unit ball $\{x \in \mathcal{H} : \|x\| \leq 1\}$ in Hilbert space $\mathcal{H}$ is weakly compact.*

REMARK 1.4.1. The unit ball B_1 of $\mathcal{H}$ is NOT compact (assuming that $\mathcal{H}$ is infinite dimensional). Reason: if $\{e_n\}_{n \in \mathbb{N}}$ is an o.n.b. then the set $\{e_1, e_2, e_3, \dots\}$ is closed but not totally bounded, hence not compact.

EXERCISE 1.4.3. Prove that if a bounded set in $\mathcal{H}$ is weakly closed then it is weakly compact.

In spite of the fact that the weak topology is weaker then the norm topology, some of the standard results remain true.

THEOREM 1.4.3. *A weakly convergent sequence is bounded.*

PROOF. Suppose that x_n is a weakly convergent sequence. Then, for any $y \in \mathcal{H}$, the sequence $\langle x_n, y\rangle$ is a convergent sequence of complex numbers, which implies that it is bounded. In other words, for any $y \in \mathcal{H}$ there exists $C = C(y) > 0$ such that $|\langle x_n, y\rangle| \leq C$. This means that, for each $n \in \mathbb{N}$, x_n can be viewed as a bounded linear functional on $\mathcal{H}$. By the Uniform Bounded Principle (Royden, p. 232), these functionals are uniformly bounded, i.e., there exists $M > 0$ such that, for all $n \in \mathbb{N}$, $\|x_n\| \leq M$. $\square$

Although weakly convergent sequence need not be convergent there are situation when it does.

THEOREM 1.4.4. *If $\{x_n\}_{n\in\mathbb{N}}$ is a weakly convergent sequence in a compact set K then it is convergent.*

PROOF. Since $\{x_n\}_{n\in\mathbb{N}} \subset K$, it has an accumulation point x' and a subsequence x'_n converging to z. If $\{x_n\}$ had another accumulation point x'', then there would be another subsequence x''_n converging to w. It would follow that $w - \lim x'_n = x'$ and $w - \lim x''_n = x''$. Since $\{x_n\}$ is weakly convergent this implies that $x' = x''$, so it has only one accumulation point, namely the limit. $\square$

By definition, the weak topology $\mathcal{W}$ is the smallest one in which every bounded linear functional L on $\mathcal{H}$ is continuous. This means that, for any such L and any open set G in the complex plane, $L^{-1}(G) \in \mathcal{W}$. Since open disks form a base of the usual topology in $\mathbb{C}$ it suffices to require that $L^{-1}(G) \in \mathcal{W}$ for each open disk G. Notice that $x \in L^{-1}(G)$ iff $L(x) \in G$, so if $G = \{z : |z - z_0| < r\}$ and $z_0 = L(x_0)$ then $x \in L^{-1}(G)$ iff $|L(x - x_0)| < r$. Now Riesz Representation Theorem implies that $L^{-1}(G) = \{x \in \mathcal{H} : |\langle x - x_0, y\rangle < r\}$ for some $y \in \mathcal{H}$. We conclude that a subbase of $\mathcal{W}$ consists of the sets $W = W(x_0; y, r) = \{x \in \mathcal{H} : |\langle x - x_0, y\rangle < r\}$.

EXERCISE 1.4.4. Prove that a bounded linear functional L is continuous in a topology $\mathcal{T}$ iff $L^{-1}(G) \in \mathcal{T}$ for every open disk G.

PROBLEM 7. Prove that a subspace of Hilbert space is closed iff it is weakly closed.

PROBLEM 8. Prove that Hilbert space is weakly complete.

PROBLEM 9. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in Hilbert space with the property that $\|x_n\| = 1$, for all n, and $\langle x_m, x_n\rangle = c$, if $m \neq n$. Prove that $\{x_n\}_{n\in\mathbb{N}}$ is weakly convergent.

PROBLEM 10. Find the weak closure of the unit sphere in Hilbert space.

CHAPTER 2

Operators on Hilbert Space

"Nobody, except topologists, is interested in problems about Hilbert space; the people who work in Hilbert space are interested in problems about operators".

Paul Halmos

2.1. Definition and Examples

Read Section 10.2 in Royden's book. Operator always means linear and bounded. The algebra of all bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$.

Example 2.1.1. Let $\mathcal{H} = \mathbb{C}^n$ and $A = [a_{ij}]$ an $n \times n$ matrix. The operator of multiplication by A is linear and bounded. Indeed, for $x = (x_1, x_2, \dots, x_n)$ and $M = \sup_{1 \le i \le n} \left(\sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}$,

$$\|Ax\| = \sup_{1 \le i \le n} |\sum_{j=1}^n a_{ij} x_j| \le \sup_{1 \le i \le n} \left(\sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2} = M\|x\|$$

so $\|A\| \le M$.

Example 2.1.2. Let $\mathcal{H} = \ell^2$ and $A = [a_{ij}]_{i,j=1}^\infty$, where $a_{ij} = c_i$ if $i = j$ and $a_{ij} = 0$ if $i \ne j$. We call such matrix *diagonal* and denote it by $\mathrm{diag}(c_1, c_2, \dots)$, or $\mathrm{diag}(c_n)$. The operator A (or, more precisely, the operator of multiplication by A) is bounded iff $c = (c_1, c_2, \dots) \in \ell^\infty$ (i.e., when c is a bounded sequence). Indeed, let $x = (x_1, x_2, \dots) \in \ell^2$, so $Ax = (c_1 x_1, c_2 x_2, \dots)$ and $\|Ax\|^2 = \sum_{i=1}^\infty |c_i x_i|^2$. If $|c_i| \le M$, $i \in \mathbb{N}$, then $\|Ax\|^2 \le M^2 \sum_{i=1}^\infty |x_i|^2 = \|x\|^2$ so A is bounded. On the other hand, if $c \notin \ell^\infty$, then for each n there exists i_n so that $|c_{i_n}| \ge n$. Then $\|Ae_{i_n}\| = \|c_{i_n} e_{i_n}\| \ge n \to \infty$ and A is unbounded.

Remark 2.1.1. It is extremely hard to decide, in general, whether an operator A is bounded just by studying its matrix $[\langle Ae_j, e_i \rangle]_{i,j=1}^\infty$.

EXAMPLE 2.1.3. Let $\mathcal{H} = \ell^2$ and let S be the *unilateral shift*, defined by $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$. Notice that $\|S(x_1, x_2, \dots)\|^2 = 0^2 + |x_1|^2 + |x_2|^2 + \cdots = \|x\|^2$ so $\|S\| = 1$. In fact, S is an isometry, hence injective, but it is not surjective!

EXAMPLE 2.1.4 (Multiplication on L^2). Let h be a measurable function and define $M_h f$, for $f \in L^2$, by $(M_h f)(t) = h(t)f(t)$. If $h \in L^\infty$ (essentially bounded functions — see Royden, p. 118), then

$$\|M_h f\|^2 = \int |hf|^2 \leq \|h\|_\infty^2 \int |f|^2 = \|h\|_\infty^2 \|f\|^2$$

so M_h is a bounded operator on L^2 and $\|M_h\| \leq \|h\|_\infty$. On the other hand, for $\epsilon > 0$, there exists a set $C \subset [0,1]$ of positive measure so that $|h(t)| \geq \|h\|_\infty - \epsilon$ for $t \in C$. If $f = \chi_C$ then

$$\|M_h f\|^2 = \int |hf|^2 = \int_C |h|^2 \geq (\|h\|_\infty - \epsilon)^2 \mu(C) = (\|h\|_\infty - \epsilon)^2 \|f\|^2,$$

and it follows that $\|M_h\| \geq \|h\|_\infty - \epsilon$. We conclude that $\|M_h\| = \|h\|_\infty$ and M_h is bounded iff $h \in L^\infty$.

EXAMPLE 2.1.5 (Integral operators on L^2). Let $K : [0,1] \times [0,1] \to \mathbb{C}$ be measurable and square integrable with respect to planar Lebesgue measure. We define the operator T_K by $(T_k f)(x) = \int_0^1 K(x,y)f(y)\,dy$. Now

$$\begin{aligned}\|T_K f\|^2 &= \int_0^1 |T_k f(x)|^2\,dx = \int_0^1 \left|\int_0^1 K(x,y)f(y)\,dy\right|^2 dx \leq \int_0^1 \left(\int_0^1 |K(x,y)f(y)|\,dy\right)^2 dx \\ &\leq \int_0^1 \left\{\int_0^1 |K(x,y)|^2\,dy\right\}\left\{\int_0^1 |f(y)|^2\,dy\right\} dx = \|f\|^2 \int_0^1 \int_0^1 |K(x,y)|^2\,dydx.\end{aligned}$$

Therefore, T_K is bounded and $\|T_K\| \leq \left\{\int_0^1 \int_0^1 |K(x,y)|^2\,dydx\right\}^{1/2}$.

EXAMPLE 2.1.6 (Weighted shifts). Let $\mathcal{H} = \ell^2$ and let $\{c_n\}_{n\in\mathbb{N}}$ be a bounded sequence of complex numbers. A *weighted shift* W on ℓ^2 is defined by $W(x_1, x_2, \dots) = (0, c_1x_1, c_2x_2, \dots)$. It can be written as $W = S\,\mathrm{diag}(c_n)$ so it is a bounded operator and $\|W\| = \|\mathrm{diag}(c_n)\|$.

In some situations it is useful to have an alternate formula for the operator norm. In what follows we will use notation B_1 for the closed unit ball of $\mathcal{H}$, i.e. $B_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}$.

PROPOSITION 2.1.1. *Let T be linear operator on Hilbert space. Then $\|T\| = \sup\{|\langle Tx, y\rangle| : x, y \in B_1\}$.*

PROOF. Let α denote the supremum above, and let us assume that $T \neq 0$ (otherwise there is nothing to prove). Clearly, for $x, y \in B_1$, $|\langle Tx, y\rangle| \leq \|T\|$, so $\alpha \leq \|T\|$. In the other direction,

$$\begin{aligned}\alpha &\geq \sup\{|\langle Tx, y\rangle| : x, y \in B_1,\, Tx \neq 0,\, y = \frac{Tx}{\|Tx\|}\}\\ &= \sup\{|\langle Tx, \frac{Tx}{\|Tx\|}\rangle| : x \in B_1,\, Tx \neq 0\}\\ &= \sup\{\|Tx\| : x \in B_1,\, Tx \neq 0\}\\ &= \|T\|,\end{aligned}$$

and the proof is complete. □

2.2. Adjoint

In Linear Algebra we learn that the column space of matrix $A = [a_{ij}]_{i,j=1}^n$ and the null space of its transpose A^T are orthogonal complements in $\mathbb{R}^n$. In $\mathbb{C}^n$, A^T needs to be replaced by $A^* = [\overline{a_{ji}}]_{i,j=1}^n$. In this situation,

$$\langle Ax, y\rangle = \langle x, A^*y\rangle. \tag{2.1}$$

EXERCISE 2.2.1. Prove that, if A is an $n \times n$ matrix and $x, y \in \mathbb{C}^n$, then $\langle Ax, y\rangle = \langle x, A^*y\rangle$.

EXAMPLE 2.2.1. Let $h \in L^\infty$ and let M_h be the operator of multiplication on L^2. Then $(M_h)^* = M_{\overline{h}}$.

The following result will show that a relation (2.1) is available for any operator.

PROPOSITION 2.2.1. *If T is an operator on $\mathcal{H}$ then there exists a unique operator S on $\mathcal{H}$ such that $\langle Tx, y\rangle = \langle x, Sy\rangle$, for all $x, y \in \mathcal{H}$.*

PROOF. Let $y \in \mathcal{H}$ be fixed. Then $\varphi(x) = \langle Tx, y\rangle$ is a bounded linear functional on $\mathcal{H}$. By Riesz Representation Theorem there exists a unique $z \in \mathcal{H}$ such that $\varphi(x) = \langle x, z\rangle$, for all $x \in \mathcal{H}$. Define $Sy = z$. Then

$\langle Tx, y\rangle = \langle x, Sy\rangle$. To show that S is linear, let $Sy_1 = z_1$, $Sy_2 = z_2$, and let $x \in \mathcal{H}$. Then

$$\langle x, S(\alpha_1 y_1 + \alpha_2 y_2)\rangle = \langle Tx, \alpha_1 y_1 + \alpha_2 y_2\rangle = \overline{\alpha_1}\langle Tx, y_1\rangle + \overline{\alpha_2}\langle Tx, y_2\rangle$$

$$= \overline{\alpha_1}\langle x, Sy_1\rangle + \overline{\alpha_2}\langle x, Sy_2\rangle = \langle x, \alpha_1 Sy_1 + \alpha_2 Sy_2\rangle.$$

By the uniqueness part of Riesz Representation Theorem S is linear. That S is unique can be deduced by contradiction: if $\langle x, Sy\rangle = \langle x, S'y\rangle$ for all $x, y \in \mathcal{H}$ then $\langle x, Sy - S'y\rangle = 0$ for all x which implies that $Sy - S'y = 0$ for all y, hence $S = S'$. Finally, S is bounded: $\|Sy\|^2 = \langle Sy, Sy\rangle = \langle TSy, y\rangle \le \|TSy\|\|y\| \le \|T\|\|Sy\|\|y\|$ so $\|Sy\| \le \|T\|\|y\|$ and $\|S\| \le \|T\|$. □

Definition 2.2.1. If $T \in \mathcal{L}(\mathcal{H})$ then the *adjoint* of T, denoted T^*, is the unique operator on $\mathcal{H}$ satisfying $\langle Tx, y\rangle = \langle x, T^*y\rangle$, for all $x, y \in \mathcal{H}$.

Here are some of the basic properties of the *involution* $T \mapsto T^*$.

Proposition 2.2.2.

(a) $I^* = I$

(b) $T^{**} = (T^*)^* = T$;

(c) $\|T^*\| = \|T\|$;

(d) $(\alpha_1 T_1 + \alpha_2 T_2)^* = \overline{\alpha_1}T_1^* + \overline{\alpha_2}T_2^*$;

(e) $(T_1 T_2)^* = T_2^* T_1^*$;

(f) *if T is invertible then so is T^* and* $(T^*)^{-1} = (T^{-1})^*$;

(g) $\|T^2\| = \|T^*T\|$.

Proof. The assertion (a) is obvious and (b) follows from $\langle x, T^{**}y\rangle = \langle T^*x, y\rangle = \overline{\langle y, T^*x\rangle} = \overline{\langle Ty, x\rangle} = \langle x, Ty\rangle$. It was shown in the proof of Proposition 2.2.1 that $\|T^*\| \le \|T\|$ so $\|T^{**}\| \le \|T^*\| \le \|T\|$ and (c) follows from (b). We leave (d) as an exercise and notice that $\langle x, (T_1T_2)^*y\rangle = \langle T_1T_2x, y\rangle = \langle T_2x, (T_1)^*y\rangle = \langle x, (T_2)^*(T_1)^*y\rangle$ establishes (e). As a consequence of (a) and (e), $T^*(T^{-1})^* = (T^{-1}T)^* = I$ and $(T^{-1})^*T^* = (TT^{-1})^* = I$ which is (f). Finally, $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$ and to prove the opposite inequality let $\epsilon > 0$ and let x be a unit vector such that $\|Tx\| \geq \|T\| - \epsilon$. Then $\|T^*T\| \geq \|T^*Tx\| \geq \langle T^*Tx, x\rangle = \|Tx\|^2 > (\|T\| - \epsilon)^2$, and (g) is proved. □

EXAMPLE 2.2.2. Let $\mathcal{H} = \ell^2$ and the let S be the unilateral shift (see Example 2.1.3). Then $S^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$. The operator S^* is called the *backward shift*.

EXAMPLE 2.2.3. Let T_K be the integral operator on L^2 (see Example 2.1.5). Then $(T_K)^* = T_{K^*}$, where $K^*(x, y) = \overline{K(y, x)}$.

We now give the Hilbert space formulation of the relation with which we have opened this section.

THEOREM 2.2.3. *If T is an operator on Hilbert space $\mathcal{H}$ then* $\operatorname{Ker} T = (\operatorname{Ran} T^*)^\perp$.

PROOF. Let $x \in \operatorname{Ker} T$ and let $y \in \operatorname{Ran} T^*$. Then there exists $z \in \mathcal{H}$ such that $y = T^*z$. Therefore $\langle x, y\rangle = \langle x, T^*z\rangle = \langle Tx, z\rangle = 0$ so $x \in (\operatorname{Ran} T^*)^\perp$. In the other direction, if $x \in (\operatorname{Ran} T^*)^\perp$ and $z \in \mathcal{H}$, then $\langle Tx, z\rangle = \langle x, T^*z\rangle = 0$. Taking $z = Tx$ we see that $Tx = 0$, and the proof is complete. □

We notice that, for $T \in \mathcal{L}(\mathcal{H})$ and $x, y \in \mathcal{H}$, the expression $\langle Tx, y\rangle$ is a form that is linear in the first and conjugate linear in the second argument. It turns out that this is sufficient for a polarization identity.

PROPOSITION 2.2.4 (Second Polarization Identity).

$$4\langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle + i\langle T(x+iy), x+iy\rangle - i\langle T(x-iy), x-iy\rangle.$$

EXERCISE 2.2.2. Prove Second Polarization Identity.

2.3. Operator topologies

In this section we take a look at the algebra $\mathcal{L}(\mathcal{H})$. It has three useful topologies which lead to 3 different types of convergence.

DEFINITION 2.3.1. A sequence of operators $T_n \in \mathcal{L}(\mathcal{H})$ *converges uniformly* (or *in norm*) to an operator T if $\|T_n - T\| \to 0$, $n \to \infty$. A sequence of operators $T_n \in \mathcal{L}(\mathcal{H})$ *converges strongly* to an operator T if $\|T_nx - Tx\| \to 0$,

$n \to \infty$, for all $x \in \mathcal{H}$. A sequence of operators $T_n \in \mathcal{L}(\mathcal{H})$ *converges weakly* to an operator T if $\langle T_n x - Tx, y\rangle \to 0$, $n \to \infty$, for any $x, y \in \mathcal{H}$.

It follows from the definition that the weak topology is the weakest of the three, while then norm topology (a.k.a. the *uniform topology*) is the strongest. Are they different?

PROPOSITION 2.3.1. *The operator norm is continuous with respect to the uniform topology but discontinuous with respect to the strong and weak topologies.*

PROOF. The first assertion is a consequence of the inequality $|\|A\| - \|B\|| \le \|A - B\|$. To prove the other two, let $\{e_n\}_{n\in\mathbb{N}}$ be an o.n.b. of $\mathcal{H}$, $\mathcal{H}_n = \vee_{k=n}^{\infty} e_k$, $P_n = P_{\mathcal{H}_n}$. Then $P_n \to 0$ strongly (hence weakly) since $\|P_n x\|^2 = \sum_{k=n+1}^{\infty} |x_k|^2 \to 0$. However, $\|P_n\| = 1$ which does not converge to 0. □

EXAMPLE 2.3.1. We say that an operator T is a *rank one* operator if there exist $u, v \in \mathcal{H}$ so that $Tx = \langle x, v\rangle u$. We use the notation $T = u \otimes v$. Let $T_n = e_n \otimes e_1$. Then $\langle T_n x, y\rangle = x_1 y_n \to 0$ while $T_n x = x_1 e_n$ is not a convergent sequence. Thus, the weak and strong topologies are different.

EXAMPLE 2.3.2. The involution $T \mapsto T^*$ is continuous in uniform topology. ($\|T_n^* - T^*\| = \|T_n - T\|$). Also, it is continuous in the weak topology, because

$$|\langle (T_n^* - T^*)x, y\rangle| = |\langle x, (T_n - T)y\rangle| = |\langle (T_n - T)y, x\rangle|.$$

However, it is not continuous in the strong topologies. Counterexample: let S be the unilateral shift, and $T_n = (S^*)^n$. Then $T_n \to 0$ strongly but $\{T_n^*\}$ is not a strongly convergent sequence. Indeed, for any $x = (x_1, x_2, \dots) \in \mathcal{H}$, $\|T_n x\|^2 = \|(x_{n+1}, x_{n+2}, \dots)\|^2 = \sum_{k=n}^{\infty} |x_k|^2 \to 0$, as $n \to \infty$. On the other hand, for $x = e_1$, $T_n^* x = S^n e_1 = e_n$, which is not a convergent sequence.

An operator $T \in \mathcal{L}(\mathcal{H})$ is a continuous mapping when $\mathcal{H}$ is given the strong topology. We will write, following Halmos, (s⟶s). One may ask about the other types of continuity.

THEOREM 2.3.2. *The three types of continuity (s⟶s), (w⟶w), and (s⟶w) are all equivalent.*

PROOF. Suppose that T is continuous, and let W be a weakly open neighborhood of Tx_0 in $\mathcal{H}$. We will show that $T^{-1}(W)$ is weakly open. It suffices to prove this assertion in the case when W belongs to the subbase of the weak topology. To that end, let $W = W(Tx_0, y, r) = \{x \in \mathcal{H} : |\langle x - Tx_0, y\rangle| < r\}$. Then $z \in T^{-1}(W) \Leftrightarrow Tz \in W \Leftrightarrow |\langle Tz - Tx_0, y\rangle| < \epsilon \Leftrightarrow |\langle z - x_0, T^*y\rangle| < \epsilon$. We see that $z \in T^{-1}(W)$ iff $z \in V(x_0, T^*y, \epsilon)$ so $T^{-1}(W) = V$ which is a weakly open set.

The implication (w$\to$w)$\Rightarrow$(s$\to$w) is trivial, so we concentrate on the implication (s$\to$w)$\Rightarrow$(s$\to$s). To that end, suppose that T is not continuous. Then it is unbounded, so there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ of unit vectors such that $\|Tx_n\| \geq n^2$, $n \in \mathbb{N}$. Clearly, $x_n/n \to 0$ and the assumption (s$\to$w) implies that Tx_n/n weakly converges to 0. By Theorem 1.4.3 the sequence $\{Tx_n/n\}$ is bounded which contradicts the fact that $\|Tx_n/n\| \geq n$. □

The fact that every operator in $\mathcal{L}(\mathcal{H})$ is weakly continuous has an interesting consequence.

COROLLARY 2.3.3. *If T is a linear operator on $\mathcal{H}$ then $T(B_1)$ is closed.*

PROOF. Banach-Alaoglu Theorem established that B_1 is weakly compact so, by Theorem 2.3.2, $T(B_1)$ is weakly compact, hence weakly closed, hence norm closed. □

EXERCISE 2.3.1. Prove that if F is a closed and bounded set in $\mathcal{H}$ then $T(F)$ is closed.

At the end of this section we consider a situation that occurs quite frequently.

THEOREM 2.3.4. *Let $\mathcal{M}$ be a linear manifold that is dense in Hilbert space $\mathcal{H}$. Every bounded linear transformation $T : \mathcal{M} \to \mathcal{H}$ can be uniquely extended to a bounded linear transformation $\hat{T} : \mathcal{H} \to \mathcal{H}$. In addition, the operator norm of T equals $\|\hat{T}\|$.*

PROOF. Let $x \in \mathcal{H}$. Then there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$ converging to x. Since $\{x_n\}_{n\in\mathbb{N}}$ is also a Cauchy sequence, for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that, $m, n \geq N \Rightarrow \|x_m - x_n\| < \epsilon/\|T\|$. It follows that, for $m, n \geq N$, $\|Tx_m - Tx_n\| < \epsilon$, so $\{Tx_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, hence convergent, and there exists $y = \lim_{n\to\infty} Tx_n$. We will define $\hat{T}x = y$, i.e., $\hat{T}(\lim x_n) = \lim Tx_n$.

First we need to establish that the definition is independent of the sequence $\{x_n\}_{n\in\mathbb{N}}$. If $\{x'_n\}_{n\in\mathbb{N}}$ is another sequence converging to x, we form the sequence $(x_1, x'_1, x_2, x'_2, \dots)$ which also converges to x. By the previous, the sequence $(Tx_1, Tx'_1, Tx_2, Tx'_2, \dots)$ must converge, and therefore, both of the subsequences $\{Tx_n\}_{n\in\mathbb{N}}$ and $\{Tx'_n\}_{n\in\mathbb{N}}$ must have the same limit.

Notice that, if $x_n \to x$, the continuity of the norm implies that $\|\hat{T}x\| = \|\lim Tx_n\| = \lim \|Tx_n\| \leq \lim \|T\|\|x_n\| = \|T\|\|x\|$ so $\|\hat{T}\| \leq \|T\|$. Since the other inequality is obvious we see that $\|\hat{T}\| = \|T\|$. In particular, $\hat{T}$ is a bounded operator. Also, $\hat{T}(\alpha x + \beta y) = \hat{T}(\alpha \lim x_n + \beta \lim y_n) = \hat{T}(\lim(\alpha x_n + \beta y_n)) = \lim T(\alpha x_n + \beta y_n) = \lim(\alpha T x_n + \beta T y_n) = \alpha \lim T x_n + \beta \lim T y_n = \alpha \hat{T} x + \beta \hat{T} y$, so $\hat{T}$ is linear.

Finally, suppose that T_1 and T_2 are two continuous extensions of T, and let $x \in \mathcal{H}$. If $x_n \to x$, the continuity implies that both $T_1 x_n \to T_1 x$ and $T_2 x_n \to T_2 x$. If $x_n \in \mathcal{M}$ then $T_1 x_n = T_2 x_n$, so $T_1 x = T_2 x$. Therefore, the extension is unique, and the proof is complete. □

`Need an example`

2.4. Invariant and Reducing Subspaces

When $\mathcal{M}$ is a closed subspace of $\mathcal{H}$, we can always write $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. Relative to this decomposition, any operator T acting on $\mathcal{H}$ can be written as a 2×2 matrix with operator entries

$$T = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}. \tag{2.2}$$

It is sometimes convenient to consider only the initial space or the target space as a direct sum. In such a situation we will use a 1×2 or 2×1 matrix. Thus $\begin{bmatrix} X & Y \end{bmatrix}$ will describe an operator $T : \mathcal{M} \oplus \mathcal{M}^\perp \to \mathcal{H}$; if $f \in \mathcal{M}$ and $g \in \mathcal{M}^\perp$ then $\left[\begin{smallmatrix} X & Y \end{smallmatrix}\right] \left[\begin{smallmatrix} f \\ g \end{smallmatrix}\right] = Xf + Yg$.

A subspace $\mathcal{M}$ is *invariant* for T if, for any $x \in \mathcal{M}$, $Tx \in \mathcal{M}$. It is *reducing* for T if both $\mathcal{M}$ and $\mathcal{M}^\perp$ are invariant for T.

EXAMPLE 2.4.1. The subspace (0) consisting of zero vector only is an invariant subspace for any operator T. Also, $\mathcal{H}$ is an invariant subspace for any operator T. Because they are invariant for every operator they are called *trivial*. A big open problem in Operator theory is whether every operator has a *non-trivial* invariant subspace.

EXAMPLE 2.4.2. If $\mathcal{M}$ is a closed subspace of $\mathcal{H}$ and T_1 is an operator on $\mathcal{M}$ with values in $\mathcal{M}$, then the operator $T = T_1 \oplus 0$, defined by $Tx = T_1x$ if $x \in \mathcal{M}$ and $Tx = 0$ if $x \in \mathcal{M}^\perp$ is an operator in $\mathcal{L}(\mathcal{H})$. However, if $\mathcal{M}$ is not invariant for T_1, the same definition ($Tx = T_1x$ for $x \in \mathcal{M}$, $Tx = 0$ for $x \in \mathcal{M}^\perp$) describes the operator $\begin{bmatrix} T_1 & 0 \end{bmatrix}$.

PROPOSITION 2.4.1. *If T is an operator on Hilbert space $\mathcal{H}$, and $P = P_\mathcal{M}$ is the projection onto the closed subspace $\mathcal{M}$, then the following are equivalent:*

(a) *$\mathcal{M}$ is invariant for T;*

(b) *$PTP = TP$;*

(c) *$Z = 0$ in (2.2).*

PROOF. It is not hard to see that the matrix for P is $\left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right]$ so $PTP - TP = \left[\begin{smallmatrix} 0 & 0 \\ -Z & 0 \end{smallmatrix}\right]$. This establishes (b) $\Leftrightarrow$ (c). Since $\left[\begin{smallmatrix} f \\ g \end{smallmatrix}\right] \in \mathcal{M}$ iff $g = 0$, we see that $T\left[\begin{smallmatrix} f \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} Xf \\ Zf \end{smallmatrix}\right] \in \mathcal{M}$ for all $x \in \mathcal{H}$ iff $Z = 0$ so (a) $\Leftrightarrow$ (c). $\square$

EXAMPLE 2.4.3. Let S be the unilateral shift, $n \in \mathbb{N}$, and $\mathcal{M} = \vee_{k \geq n} e_k$. Then $S\mathcal{M} = \vee_{k \geq n+1} e_k \subset \mathcal{M}$.

PROPOSITION 2.4.2. *If T is an operator on Hilbert space $\mathcal{H}$, and $P = P_\mathcal{M}$ then the following are equivalent:*

(a) *$\mathcal{M}$ is reducing for T;*

(b) *$PT = TP$;*

(c) *$Y, Z = 0$ in (2.2);*

(d) *$\mathcal{M}$ is invariant for T and T^*.*

PROOF. Since $PT - TP = \left[\begin{smallmatrix} 0 & Y \\ -Z & 0 \end{smallmatrix}\right]$ we see that (b) $\Leftrightarrow$ (c). Further, the matrix for T^* is $\left[\begin{smallmatrix} X^* & Z^* \\ Y^* & W^* \end{smallmatrix}\right]$ so, by Proposition 2.4.1, $\mathcal{M}$ is invariant for T and T^* iff $Z = Y^* = 0$ and (c) $\Leftrightarrow$ (d). In order to prove that (a) $\Leftrightarrow$ (d) it suffices to show that $\mathcal{M}$ is invariant for T^* iff $\mathcal{M}^\perp$ is invariant for T. By Proposition 2.4.1, $\mathcal{M}$ is invariant for T^* iff $Y^* = 0$ (iff $Y = 0$). On the other hand $T\left[\begin{smallmatrix}0\\g\end{smallmatrix}\right] = \left[\begin{smallmatrix}Yg\\Wg\end{smallmatrix}\right] \in \mathcal{M}^\perp$ iff $Yg = 0$ for all g. □

EXERCISE 2.4.1. Prove that the matrix for T^* is $\left[\begin{smallmatrix}X^* & Z^*\\Y^* & W^*\end{smallmatrix}\right]$.

EXAMPLE 2.4.4. Let $T = M_h$, let $E \subset [0,1]$, $m(E) > 0$, and let $\mathcal{M} = L^2(E)$. If $f \in \mathcal{M}$ then $Tf = hf \in \mathcal{M}$. Also, $T^* = M_{\overline{h}}$ and $T^*f = \overline{h}f \in \mathcal{M}$, so $\mathcal{M}$ is reducing for T.

EXAMPLE 2.4.5. Let S be the unilateral shift, $n \in \mathbb{N}$, and $\mathcal{M} = \vee_{k\geq n} e_k$. Then $\mathcal{M}$ is invariant for S but not reducing, since $e_n \in \mathcal{M}$ but $S^*e_n = e_{n-1} \notin \mathcal{M}$.

2.5. Finite rank operators

The closest relatives of finite matrices are the finite rank operators.

DEFINITION 2.5.1. An operator T is a *finite rank* operator if its range is finite dimensional. We denote the set of finite rank operators by $\mathbb{F}$.

EXAMPLE 2.5.1. If T is a rank one operator $u \otimes v$ (see Example 2.3.1) then the range of $u \otimes v$ is the one dimensional subspace spanned by u, so $u \otimes v \in \mathbb{F}$.

The rank one operators turn out to be the building blocks out of which finite rank operators are made.

PROPOSITION 2.5.1. *If T is a linear operator on $\mathcal{H}$ then T belongs to $\mathbb{F}$ iff there exist vectors $u_1, u_2, \dots, u_n$, and $v_1, v_2, \dots, v_n$ such that $Tx = \sum_{i=1}^n \langle x, v_i\rangle u_i$.*

PROOF. Suppose that $\operatorname{Ran} T$ is of finite dimension n, and let $e_1, e_2, \dots, e_n$ be an o.n.b. of $\operatorname{Ran} T$. Then $Tx = \sum_{i=1}^n \langle Tx, e_i\rangle e_i = \sum_{i=1}^n \langle x, T^*e_i\rangle e_i$. We leave the converse as an exercise. □

EXERCISE 2.5.1. Prove that if there exist vectors $u_1, u_2, \dots, u_n, v_1, v_2, \dots, v_n$ such that $Tx = \sum_{i=1}^n \langle x, v_i\rangle u_i$, for all $x \in \mathcal{H}$, then $\operatorname{Ran} T$ is of dimension at most n.

EXERCISE 2.5.2. Prove that if $T = \sum u_i \otimes v_i$ then $T^* = \sum v_i \otimes u_i$.

The next theorem summarizes some very important properties of the class $\mathbb{F}$.

THEOREM 2.5.2. *The set* $\mathbb{F}$ *is a minimal* **-ideal in* $\mathcal{L}(\mathcal{H})$.

Here the star means that $\mathbb{F}$ is closed under the operation of taking adjoints.

PROOF. It is obvious that $\mathbb{F}$ is a subspace of $\mathcal{L}(\mathcal{H})$. Furthermore, if $T \in \mathbb{F}$ and $A \in \mathcal{L}(\mathcal{H})$, then $\operatorname{Ran} TA \subset \operatorname{Ran} T$ so $TA \in \mathbb{F}$. Also, if T is of finite rank, then according to Proposition 2.5.1, $T = \sum_{i=1}^n u_i \otimes v_i$ so $T^* = \sum_{i=1}^n v_i \otimes u_i$. It follows that $T^* \in \mathbb{F}$ and the same is true of T^*A^*, for any $A \in \mathcal{L}(\mathcal{H})$. Consequently, AT is of finite rank, and $\mathbb{F}$ is a *-ideal. To see that it is minimal, it suffices to show that, if J is a non-zero ideal, then J contains all rank one operators. Let $T \in J$, $T \neq 0$. Then there exists vectors x, y, such that $\|y\| = 1$ and $y = Tx$. Let $u \otimes v$ be a rank one operator. Since J is an ideal, it contains the product $(u \otimes y)T(x \otimes v)$ which equals $u \otimes v$. □

A finite rank operator is a generalization of a finite matrix. What happens when we take the closure of $\mathbb{F}$ in some topology?

EXERCISE 2.5.3. Prove that the strong closure of $\mathbb{F}$ is $\mathcal{L}(\mathcal{H})$. [Hint: Prove that $P_n \to I$ strongly.] Conclude that the weak closure of $\mathbb{F}$ is also $\mathcal{L}(\mathcal{H})$.

2.6. Compact Operators

Exercise 2.5.3 established that the strong closure of $\mathbb{F}$ is $\mathcal{L}(\mathcal{H})$. Therefore, we consider the norm topology.

DEFINITION 2.6.1. An operator T in $\mathcal{L}(\mathcal{H})$ is *compact* if it is the limit of a sequence of finite rank operators. We denote the set of compact operators by $\mathbb{K}$.

EXAMPLE 2.6.1. Let $T = \operatorname{diag}(c_n)$ as in Example 2.1.2, with $\lim_{n\to\infty} c_n = 0$. Then T is compact. Reason: take $T_n = \operatorname{diag}(c_1, c_2, \dots, c_n, 0, 0, \dots)$. Then $T_n \in \mathbb{F}$ and $\|T - T_n\| = \sup\{|c_k| : k \geq n+1\} \to 0$. It follows that T is compact.

EXAMPLE 2.6.2. Let $T = T_K$ as in Example 2.1.5. If $K \in L^2([0,1] \times [0,1])$ then T_K is compact. We will point out at several different sequences in $\mathbb{F}$ that all converge to T_K

We start with a function theoretic approach: simple functions are dense in L^2 (Royden, p. 128), and a similar proof establishes that simple functions are dense in $L^2([0,1] \times [0,1])$. Since a simple function is a linear combination of the characteristic functions of rectangles $\chi_{[a,b]\times[c,d]}(x,y) = \chi_{[a,b]}(x)\chi_{[c,d]}(y)$ it follows that $K(x,y)$ is the L^2 limit of functions of the form $K_n(x,y) = \sum_{i=1}^n f_i(x)g_i(y)$, so T_K is the norm limit of T_{K_n}, which are all finite rank operators.

EXERCISE 2.6.1. Verify that $T_{K_n} \in \mathbb{F}$, if $K_n(x,y) = \sum_{i=1}^n f_i(x)g_i(y)$.

Our second approach is exploiting the fact that L^2 is Hilbert space. If $\{e_j\}_{j\in\mathbb{N}}$ is an o.n.b. of L^2 we can, for a fixed y, write $K(x,y) = \sum_{j=1}^\infty k_j(y)e_j(x)$. Now define $K_N(x,y) = \sum_{j=1}^N k_j(y)e_j(x)$ and notice that $T_{K_N} \to T_K$ as $N \to \infty$.

EXERCISE 2.6.2. Verify that $T_{K_N} \in \mathbb{F}$ and that $\lim_{N\to\infty} T_{K_N} = T_K$, if $K_N(x,y)$ is as above.

Our last method is based on the matrix for T_K. Let $k_{ij} = \langle T_K e_j, e_i\rangle$, with $\{e_n\}_{n\in\mathbb{N}}$ an o.n.b. of $L^2([0,1])$. First we notice that, for any $f \in L^2$, $\sum_k |\langle f, e_k\rangle|^2 = \|\sum_k \langle f, e_k\rangle e_k\|^2 = \|f\|^2$. Therefore,

$$\sum_{j=1}^\infty |\langle T_K e_j, e_i\rangle|^2 = \sum_{j=1}^\infty |\langle e_j, T_K^* e_i\rangle|^2 = \sum_{j=1}^\infty |\langle T_K^* e_i, e_j, \rangle|^2 = \|T_K^* e_i\|^2$$
$$= \int_0^1 \left| \int_0^1 K^*(y,x)e_i(x)\,dx \right|^2 dy = \int_0^1 \left| \int_0^1 \overline{K(x,y)}e_i(x)\,dx \right|^2 dy.$$

It follows that, for any $n \in \mathbb{N}$,

$$\sum_{i=1}^n \sum_{j=1}^\infty |k_{ij}|^2 = \sum_{i=1}^n \int_0^1 \left| \int_0^1 \overline{K(x,y)}e_i(x)\,dx \right|^2 dy = \int_0^1 \sum_{i=1}^n \left| \int_0^1 \overline{K(x,y)}e_i(x)\,dx \right|^2 dy$$
$$\le \int_0^1 \sum_{i=1}^\infty \left| \int_0^1 \overline{K(x,y)}e_i(x)\,dx \right|^2 dy = \int_0^1 \int_0^1 |K(x,y)|^2\,dxdy$$

so the series $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|k_{ij}|^2$ converges. Operators whose matrices satisfy this condition are called *Hilbert–Schmidt operators*. The Hilbert-Scmidt norm is defined as $\|T_K\|_2 = \{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|k_{ij}|^2\}^{1/2}$, and it satisfies the inequality $\|A\| \leq \|A\|_2$. Hilbert-Scmidt operators are compact because we can define T_n to be the matrix consisting of the first n rows of the matrix of T_K and having the remaining entries 0. Then each $T_n \in \mathbb{F}$ and $\|T_n - T_K\| \to 0$. Indeed, $\operatorname{Ran} T_n \subset \vee\{e_1, e_2, \dots, e_n\}$, and $\|T_K - T_n\|^2 \leq \|T_K - T_n\|_2^2 = \sum_{i=n+1}^{\infty}\sum_{j=1}^{\infty}|k_{ij}|^2 \to 0$, $n \to \infty$.

Exercise 2.6.3. Prove that the Hilbert-Scmidt norm is indeed a norm and, for any $T \in \mathcal{L}(\mathcal{H})$, $\|T\| \leq \|T\|_2$.

Next we consider some of the properties of compact operators. The first one follows directly from the definition.

Theorem 2.6.1. *The set $\mathbb{K}$ is the smallest closed *-ideal in $\mathcal{L}(\mathcal{H})$.*

The following result reveals the motivation for calling these operators compact.

Theorem 2.6.2. *An operator T in $\mathcal{L}(\mathcal{H})$ is compact iff it maps the closed unit ball of $\mathcal{H}$ into a compact set.*

Proof. Suppose that K is compact and let $\{y_n\}_{n\in\mathbb{N}}$ be a sequence in $K(B_1)$. We will show that there exists a subsequence of $\{y_n\}$ that converges to an element of $K(B_1)$. Notice that, for every $n \in \mathbb{N}$, $y_n = Kx_n$, and x_n belongs to the weakly compact set B_1. Thus, there exists a subsequence $\{x_{n_k}\}$ converging weakly to $x \in B_1$. Thus, it suffices to show that Kx_{n_k} converges to Kx. Let $\{K_n\}$ be a sequence in $\mathbb{F}$ thaty converges to K. For any $m \in \mathbb{N}$, $K_m(B_1)$ is a bounded and closed set (by Corollary 2.3.3) that is contained in a finite dimensional subspace of $\mathcal{H}$, so it is compact. By Theorem 1.4.4, $\{K_m x_{n_k}\}_{k\in\mathbb{N}}$ converges to $K_m x$. Now, let $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\|K - K_N\| < \epsilon/3$. Further, with N fixed, there exists $k_0 \in \mathbb{N}$ so that, for $k \geq k_0$, $\|K_N x_{n_k} - K_N x\| < \epsilon/3$. Therefore, for $k \geq k_0$,

$$\|Kx_{n_k} - Kx\| \leq \|(K - K_N)x_{n_k}\| + \|K_N(x_{n_k} - x)\| + \|(K_N - K)x\| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Thus, $y_{n_k} = Kx_{n_k}$ is a convergent subsequence converging to $Kx \in K(B_1)$ so $K(B_1)$ is a compact set.

Suppose now that $K(B_1)$ is compact and let $n \in \mathbb{N}$. Notice that $\cup_{y\in K(B_1)} B(y, 1/n)$ is an open covering of the compact set $K(B_1)$, so there exist vectors $x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)} \in \mathcal{H}$ so that $\cup_{i=1}^k B(Kx_i^{(n)}, 1/n)$ is a covering of $K(B_1)$. Let $\mathcal{H}_n$ be the span of $Kx_1^{(n)}, Kx_2^{(n)}, \dots, Kx_k^{(n)}$ and P_n the orthogonal projection on $\mathcal{H}_n$. Finally, let $K_n = P_nK$. Clearly, $K_n \in \mathbb{F}$. Let $\epsilon > 0$, and choose $N > 1/\epsilon$. If $n \geq N$, and $\|x\| \leq 1$, then $\|Kx - K_nx\| = \|Kx - P_nKx\|$. Since P_nKx is the point in $\mathcal{H}_n$ closest to Kx, it follows that $\|Kx - K_nx\| \leq \inf_{1\leq i\leq n} \|Kx - Kx_i^{(n)}\| < 1/n < \epsilon$. Thus $K_n \to K$ and the proof is complete. □

Remark 2.6.1. In many texts the characterization of compact operators, established in Theorem 2.6.2, is taken to be the definition of a compact operator.

Exercise 2.6.4. Prove that if F is a closed and bounded set in $\mathcal{H}$ and T is a compact operator in $\mathcal{L}(\mathcal{H})$ then $T(F)$ is a compact set.

There is another characterization of compact operators:

Proposition 2.6.3. *If T is a linear operator on $\mathcal{H}$ then T is compact iff it maps every weakly convergent sequence into a convergent sequence. In this situation, if $w - \lim x_n = x$ then $\lim Tx_n = Tx$.*

Proof. Suppose first that T is compact and let $w - \lim x_n = x$. By Proposition 1.4.3, there exists $M > 0$ such that, for all $n \in \mathbb{N}$, $\|x_n\| \leq M$. Therefore, $Tx_n/M \in T(B_1)$, which is compact by Theorem 2.6.2. Now Theorem 1.4.4 implies that $\lim x_n = x$.

In order to establish the converse, we will demonstrate that $T(B_1)$ is compact by showing that every sequence in $T(B_1)$ has a convergent subsequence. Let $\{y_n\}_{n\in\mathbb{N}} \subset T(B_1)$. Then $y_n = Tx_n$, for $x_n \in B_1$, so the Banach–Alaoglu Theorem implies that $\{x_n\}$ has a weakly convergent subsequence $\{x_{n_k}\}$ and, by assumption, $\{Tx_{n_k}\}$ is a (strongly) convergent subsequence of $\{Tx_n\}$. □

Example 2.6.3. We have seen in Example 2.6.1 that if $T - \text{diag}(c_n)$ and $c_n \to 0$, then T is compact. The converse is also true: if $\{e_n\}$ is the o.n.b. which makes T diagonal, then $Te_n \to 0$ (because $w - \lim e_n = 0$ and T is compact) so $\|c_ne_n\| \to 0$.

It is useful to know that compactness is inherited by the parts of an operator.

THEOREM 2.6.4. *Suppose that T is a compact operator on Hilbert space $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ and that, relative to this decomposition, $T = \left[\begin{smallmatrix} X & Y \\ Z & W \end{smallmatrix}\right]$. Then each of the operators X, Y, Z, W is compact.*

PROOF. Let $\{T_n\}$ be a sequence of finite rank operators that converges to T. Write, for each $n \in \mathbb{N}$, $T_n = \left[\begin{smallmatrix} X_n & Y_n \\ Z_n & W_n \end{smallmatrix}\right]$. Then all the operators $X_n, Y_n, Z_n, W_n \in \mathbb{F}$ and they converge to X, Y, Z, W, respectively. □

EXERCISE 2.6.5. Prove that $X_n, Y_n, Z_n, W_n \in \mathbb{F}$ and that they converge to X, Y, Z, W, respectively. [Consider the projections $P_1 = P_{\mathcal{M}}$ and $P_2 = P_{\mathcal{M}^\perp}$ and notice that, for example $P_1 T P_2 = \left[\begin{smallmatrix} 0 & Y \\ 0 & 0 \end{smallmatrix}\right]$, so $\|Y_n - Y\| \leq \|T_n - T\|$ and $\operatorname{Ran} Y_n \subset \operatorname{Ran} P_1 T_n P_2$ the later being finite dimensional.]

2.7. Normal operators

DEFINITION 2.7.1. If T is an operator on Hilbert space $\mathcal{H}$ then:

(a) T is *normal* if $TT^* = T^*T$;

(b) T is *self-adjoint* (or *Hermitian*) if $T = T^*$;

(c) T is *positive* if $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$;

(d) T is *unitary* if $TT^* = T^*T = I$.

EXAMPLE 2.7.1. Let $T = \operatorname{diag}(c_n)$. Then $T^* = \operatorname{diag}(\overline{c_n})$ so T is normal. Also, $T = T^*$ iff $c_n \in \mathbb{R}$, $n \in \mathbb{N}$, and T is positive iff $c_n \geq 0$, $n \in \mathbb{N}$. Finally, $T^*T = \operatorname{diag}(|c_n|^2)$ so T is unitary iff $|c_n| = 1$, $n \in \mathbb{N}$.

EXERCISE 2.7.1. Let $T = M_h$ on L^2. Prove that T is normal and that it is: self-adjoint iff $h(x) \in \mathbb{R}$, a.e.; positive iff $h(x) \geq 0$ a.e.; unitary iff $|h(x)| = 1$ a.e..

The relationship between T and T^* that defines each of these classes allows us to establish some of their significant properties.

PROPOSITION 2.7.1. *An operator T on Hilbert space $\mathcal{H}$ is self-adjoint iff $\langle Tx, x\rangle$ is real for any $x \in \mathcal{H}$.*

Proof. If $T = T^*$ then $\langle Tx, x\rangle = \langle x, T^*x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$ so $\langle Tx, x\rangle \in \mathbb{R}$. On the other hand, if $\langle Tx, x\rangle$ is real for any $x \in \mathcal{H}$ then Second Polarization Identity implies that $\langle Tx, y\rangle = \overline{\langle Ty, x\rangle}$ so $T = T^*$. □

Exercise 2.7.2. Prove that $\langle Tx, x\rangle \in \mathbb{R}$ implies that $\langle Tx, y\rangle = \overline{\langle Ty, x\rangle}$.

Corollary 2.7.2. *If P is a positive operator on Hilbert space $\mathcal{H}$ then P is self-adjoint.*

Example 2.7.2. If P is the orthogonal projection on a subspace $\mathcal{M}$ of Hilbert space $\mathcal{H}$, then P is a positive operator. Indeed, if $z \in \mathcal{H}$ write $z = x + y$ relative to $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$. By Theorem 1.3.2, $Pz = x$ and $Py = 0$, so $\langle Pz, z\rangle = \langle x, x + y\rangle = \|x\|^2 \geq 0$.

Combining Theorem 1.3.2 and Example 2.7.2 we see that every projection is a positive idempotent. In fact, the converse is also true.

Theorem 2.7.3. *If T is an idempotent self-adjoint operator then T is a projection on $\mathcal{M} = \{x \in \mathcal{H} : Tx = x\}$.*

Proof. Let $z \in \mathcal{H}$ and write it as $z = Tz + (z - Tz)$. Now $T(Tz) = Tz$ so $Tz \in \mathcal{M}$. Also, $z - Tz \in \mathcal{M}^\perp$. Indeed, if $x \in \mathcal{M}$, then $\langle x, z - Tz\rangle = \langle x, z\rangle - \langle x, Tz\rangle = \langle x, z\rangle - \langle Tx, z\rangle = 0$. □

By Proposition 2.1.1, the norm of every operator T in $\mathcal{L}(\mathcal{H})$ can be computed by considering the supremum of the values of its bilinear form $\langle Tx, y\rangle$. The next result shows that, when T is self adjoint, it suffices to consider only some pairs of $x, y \in B_1$.

Proposition 2.7.4. *If T is a self-adjoint operator on Hilbert space $\mathcal{H}$ then $\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}$.*

Proof. Clearly, $|\langle Tx, x\rangle| \leq \|T\|\|x\|^2$, so if we denote by α the supremum above, we have that $\alpha \leq \|T\|$. To prove that $\alpha = \|T\|$, we use the Second Polarization Identity, and we notice that, in view of the assumption $T = T^*$ and Proposition 2.7.1, $4\mathrm{Re}\langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle$. Moreover, using Parallelogram Law, and assuming that x and y are unit vectors, we obtain that $4\mathrm{Re}\langle Tx, y\rangle \leq \alpha\|x+y\|^2 + \alpha\|x-y\|^2 = \alpha(2\|x\|^2 + 2\|y\|^2) = 4\alpha$. When x is selected so that $\|Tx\| \neq 0$, $y = Tx/\|Tx\|$ we obtain $\mathrm{Re}\|Tx\| \leq \alpha$ so $\|T\| \leq \alpha$. □

Exercise 2.7.3. Prove that two product of two self-adjoint operators is self-adjoint iff the operators commute.

Remark 2.7.1. If we write $A = (T + T^*)/2$ and $B = (T - T^*)/2i$ then the operators A, B are self-adjoint and $T = A + iB$. We call them the *real part* and the *imaginary part* of T.

Proposition 2.7.5. *If T is an operator on Hilbert space $\mathcal{H}$ then the following are equivalent.*

(a) *T is a normal operator;*

(b) *$\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{H}$;*

(c) *the real and imaginary part of T commute.*

Proof. Notice that $\|Tx\|^2 - \|T^*x\|^2 = \langle (T^*T - TT^*)x, x\rangle$. If T is normal then the right side is 0, so (a) implies (b). If (b) is true, then the left side is 0, for all x. Since $T^*T - TT^*$ is self-adjoint, Proposition 2.7.4 implies that its norm is 0, so (b) implies (a). A calculation shows that, if A and B are the real and imaginary part of T, resp., then $AB - BA = (T^*T - TT^*)/2i$ so (a) is equivalent to (c). □

In Definition 1.2.3 we have introduced the concept of the Hilbert space isomorphsim. Since it preserves the inner product ($\langle Ux, Uy\rangle = \langle x, y\rangle$), it preserves the norm, and hence both weak and strong toplogies. Therefore, if $U : \mathcal{H} \to \mathcal{K}$, we do not distinguish between an operator $T \in \mathcal{L}(\mathcal{H})$ and $UTU^{-1} \in \mathcal{L}(\mathcal{H})$, and we say that they are *unitarily equivalent*. Since, by Definition 2.7.1, an operator T is unitary iff $TT^* = T^*T = I$, we should check that $UU^* = U^*U = I$.

Exercise 2.7.4. Verify that $UU^* = I_{\mathcal{K}}$ and $U^*U = I_{\mathcal{H}}$.

Notice that both equalities need to be verified, because it is quite possible for one to hold but not the other. Example: the unilateral shift S satisfies $S^*S = I \neq SS^*$.

Exercise 2.7.5. Prove that T is an isometry iff $T^*T = I$.

Exercise 2.7.3 asserts that the product of two self-adjoint operators is itself self-adjoint iff the operators commute. What if *self-adjoint* is replaced by *normal*? If M, N are commuting normal operators, their product is normal if MN commutes with N^*M^* and it looks like we need the additional assumption that M commutes with N^* (which also gives that M^* commutes with N). When an operator T commutes with both N and N^* we say that T *doubly commutes* with N. When N is normal we can establish even a stronger result.

THEOREM 2.7.6 (Fuglede–Putnam Theorem). *Suppose that M, N are normal operators and $T \in \mathcal{L}(\mathcal{H})$ intertwines M and N: $MT = TN$. Then $M^*T = TN^*$.*

PROOF. Let λ be a complex number, and denote $A = \overline{\lambda} M$, $B = \overline{\lambda} N$. Notice that $AT = TB$, so $A^2T = A(AT) = A(TB) = (AT)B = (TB)B = TB^2$, and inductively $A^kT = TB^K$, for $k \in \mathbb{N}$. It follows that, if we denote the exponential function by $\exp(z)$, $\exp(A)T = T\exp(B)$. It is not hard to see that $\exp(-A)\exp(A) = I$ so

$$T = \exp(-A)T\exp(B).$$

If we denote by $U_1 = \exp(A^* - A)$, $U_2 = \exp(B - B^*)$, then both U_1, U_2 are unitary operators. Indeed, $U_1^* = [\sum(A^* - A)^n/n!]^* = \sum(A - A^*)^n/n! = \exp(A - A^*) = U_1^{-1}$, and similarly for U_2. Now we have that $\exp(A^*)T\exp(-B^*) = U_1TU_2$ and $\|\exp(A^*)T\exp(-B^*)\| = \|T\|$. We conclude that

$$\|\exp(\lambda M^*)T\exp(-\lambda N^*)\| = \|T\|$$

for all $\lambda \in \mathbb{C}$. Now $f(\lambda) = \exp(\lambda M^*)T\exp(-\lambda N^*)$ is an entire bounded function, hence a constant. Therefore, $f'(0) = 0$. On the other hand, $f'(\lambda) = M^*\exp(\lambda M^*)T\exp(-\lambda N^*) + \exp(\lambda M^*)T\exp(-\lambda N^*)(-N^*)$ so $f'(0) = M^*T - TN^*$, and the theorem is proved. □

EXERCISE 2.7.6. Prove that $\exp(-T)\exp(T) = I$ for any operator $T \in \mathcal{L}(\mathcal{H})$.

COROLLARY 2.7.7. *The product of two normal operators is itself normal iff the operators commute.*

EXERCISE 2.7.7. Prove Corollary 2.7.7.

CHAPTER 3

Spectrum

3.1. Invertibility

In Linear Algebra we learn that each the properties of being invertible, injective, or surjective implies the other two. Things are very different in infinite dimesional Hilbert space.

Example 3.1.1. Let $T = \operatorname{diag}(1/n)$. It is easy to see that $\operatorname{Ker} T = (0)$ so T is injective. However, it is not surjective, because its range does not contain the sequence $(1, 1/2, 1/3, \dots) \in \ell^2$.

Exercise 3.1.1. Prove that $T = \operatorname{diag}(1/n)$ is injective but $(1, 1/2, 1/3, \dots) \notin \operatorname{Ran} T$.

Example 3.1.2. The backward shift S^* (see Example 2.2.2) is surjective: given $(y_1, y_2, \dots) \in \ell^2$ we have that $S^*(0, y_1, y_2, \dots) = (y_1, y_2, \dots)$. On the other hand $S^* e_1 = 0$ so S^* is not injective. Also, $S^* S(x_1, x_2, x_3, \dots) = S^*(0, x_1, x_2, \dots) = (x_1, x_2, x_3, \dots)$, so $S^* S = I$. However, $SS^*(x_1, x_2, x_3, \dots) = S(x_2, x_3, \dots) = (0, x_2, x_3, \dots)$ so $SS^* \neq I$.

We say that an operator T is *left invertible* if there exists an operator $L \in \mathcal{L}(\mathcal{H})$ such that $LT = I$. It is *right invertible* if there exists an operator R such that $TR = I$. Therefore, the unilateral shift S is left invertible, while S^* is right invertible. Since S is injective, it is tempting to jump to the conclusion that an operator is injective iff it is left invertible.

Example 3.1.3. The Volterra integral operator V is defined on L^2 by $Vf(x) = \int_0^x f(t)\,dt$. Since this is an integral operator T_K with $K = \chi_E(x, y)$ where $E = \{(x, y) : y \leq x\}$ and $\chi_E \in L^2$, V is a compact operator so it cannot be left invertible. Yet, V is injective since $Vf = 0$ implies that $f = 0$.

Exercise 3.1.2. Prove that the Volterra integral operator V is injective.

EXERCISE 3.1.3. Prove that the range of the Volterra integral operator V is a dense linear manifold in $\mathcal{H}$.

Instead of injectivity, another condition plays a major role in the questions about invertibility.

DEFINITION 3.1.1. An operator $T \in \mathcal{L}(\mathcal{H})$ is *bounded below* if there exists $\alpha > 0$ such that $\|Tx\| \geq \alpha\|x\|$, for all $x \in \mathcal{H}$.

EXAMPLE 3.1.4. Let $T = \operatorname{diag}(c_n)$. Then T is bounded below iff $|c_n| \geq \alpha > 0$, $n \in \mathbb{N}$.

An immediate consequence of this property concerns the range of the operator.

THEOREM 3.1.1. *If an operator T on Hilbert space $\mathcal{H}$ is bounded below then its range is a closed subset of $\mathcal{H}$.*

PROOF. Let y_n be a sequence of vectors in $\operatorname{Ran} T$ converging to y. Then $y_n = Tx_n$ for some $x_n \in \mathcal{H}$, so $\|y_n - y_m\| = \|Tx_n - Tx_m\| \geq \alpha\|x_n - x_m\|$. Since $\{y_n\}$ is a Cauchy sequence, the same is true of $\{x_n\}$. Let $x = \lim x_n$. Then $Tx_n \to Tx$, i.e. $y_n \to Tx$. Thus $y = Tx \in \operatorname{Ran} T$, and $\operatorname{Ran} T$ is closed. □

Example 3.1.3 shows that the injectivity is not sufficient to guarantee the left invertibility. The next result gives the correct necessary and sufficient conditions.

THEOREM 3.1.2. *Let T be an operator in $\mathcal{L}(\mathcal{H})$. The following are equivalent:*

(a) *T is left invertible;*

(b) $\operatorname{Ker} T = (0)$ *and* $\operatorname{Ran} T$ *is closed;*

(c) *T is bounded below.*

PROOF. If $LT = I$ then $\|x\| = \|LTx\| \leq \|L\|\|Tx\|$, so T is bounded below with $\alpha = 1/\|L\|$, and (a) ⇒ (c). Clearly, if T is bounded below it must be injective, and the fact that its range is closed is Theorem 3.1.1, so (c) implies (b). If (b) is true then, by the Open Mapping Theorem (Royden, p.230), there exists a bounded linear operator $L_1 : \operatorname{Ran} T \to \mathcal{H}$, such that $L_1 T = I$. If we define $L = \begin{bmatrix} L_1 & 0 \end{bmatrix}$ relative to $\mathcal{H} = \operatorname{Ran} T \oplus (\operatorname{Ran} T)^\perp$ (see Example 2.4.2), then $L \in \mathcal{L}(\mathcal{H})$ and $LT = I$. □

EXERCISE 3.1.4. Prove that $T = \left[\begin{smallmatrix} I & A \\ 0 & I \end{smallmatrix}\right]$ is bounded below for any operator A.

A similar characterization is available for surjectivity. The most efficient approach seems to be based on the observation that T is right invertible iff T^* is left invertible. In order to continue in this direction we need the following result, which is significant on its own.

THEOREM 3.1.3. *The operator T has closed range iff the range of T^* is closed.*

PROOF. Since $T^{**} = T$ it suffices to prove one of the two implications. To that end, let $\operatorname{Ran} T$ be closed, and let x_n be a sequence of vectors such that $T^* x_n$ converges to y. We will show that $y \in \operatorname{Ran} T^*$. Since $\operatorname{Ran} T$ is closed we can write $\mathcal{H} = \operatorname{Ran} T \oplus \operatorname{Ker} T^*$. If relative to this decomposition $x_n = x_n' \oplus x_n''$, then $T^* x_n = T^* x_n'$ so, without loss of generality, we may assume that the sequence x_n belongs to $\operatorname{Ran} T$. The convergence of $T^* x_n$ implies the weak convergence so, for any $z \in \mathcal{H}$, $\langle T^* x_n, z\rangle \to \langle y, z\rangle$. It follows that $\langle x_n, Tz\rangle \to \langle y, z\rangle$ and, moreover, that $\langle x_n, w\rangle$ converges for any $w \in \mathcal{H}$. Indeed, if we write $w = w_1 \oplus w_2$, where $w_1 \in \operatorname{Ran} T$ (so $w_1 = Tz_1$) and $w_2 \in \operatorname{Ker} T^*$, (so $\langle x_n, w_2\rangle = 0$), we see that $\{x_n\}$ is a weakly convergent sequence. If $w - \lim x_n = x$ then $w - \lim T^* x_n = T^* x$. On the other hand, $T^* x_n$ converges to y, so $y = T^* x \in \operatorname{Ran} T^*$. □

Now we can deliver the promised characterizations of surjectivity.

THEOREM 3.1.4. *Let T be an operator in $\mathcal{L}(\mathcal{H})$. The following are equivalent:*

(a) *T is right invertible;*

(b) *T^* is bounded below.*

(c) *T is surjective.*

PROOF. The equivalence of (a) and (b) follows from Theorem 3.1.2 applied to T^*. Further, $TR = I$ implies that TR is surjective. Since $\operatorname{Ran} TR \subset \operatorname{Ran} T$, T is surjective and (a) implies (c). Finally, let T be surjective. This implies that $\operatorname{Ker} T^* = (0)$ and also, via Theorem 3.1.3, that $\operatorname{Ran} T^*$ is closed. Applying Theorem 3.1.2 we see that T^* is left invertible and the result follows by taking adjoints. □

We close this section with a sufficient condition for invertibility that is of quite a different nature.

THEOREM 3.1.5. *If T is an operator on Hilbert space $\mathcal{H}$ and $\|I - T\| < 1$ then T is invertible.*

PROOF. Let $\alpha = 1 - \|I - T\| \in (0, 1]$. If $x \in \mathcal{H}$, then $\|Tx\| = \|x - (1 - T)x\| \geq \|x\| - \|1 - T\|\|x\| = \alpha\|x\|$ so T is bounded below. Suppose now that the range of T is not dense in $\mathcal{H}$. Then there exists $y \in \mathcal{H}$ such that $d = \inf\{\|y - x\| : x \in \operatorname{Ran} T\} > 0$. It follows that there exists $x \in \operatorname{Ran} T$ such that $(1 - \alpha)\|y - x\| < d$. (Obvious if $\alpha = 1$, otherwise $\beta = 1/(1 - \alpha) > 1$ so there exists x such that $\|y - x\| < \beta d$.) Notice that $x + T(y - x) \in \operatorname{Ran} T$ so $d \leq \|y - x - T(y - x)\| \leq \|I - T\|\|y - x\| < d$, which is a contradiction, so T has dense range. □

SECOND PROOF: The series $I + T + T^2 + T^3 + \dots$ converges in the operator norm, and it is easy to verify that $(I - T)(I + T + T^2 + T^3 + \dots) = I$. □

EXERCISE 3.1.5. Prove that, if $\|T\| < 1$, the series $\sum_{n=0}^{\infty} T^n$ converges uniformly.

EXERCISE 3.1.6. Verify that, if $\|T\| < 1$, $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$.

3.2. Spectrum

A complex number λ belongs to the *spectrum* of an operator T (notation: $\lambda \in \sigma(T)$) if $T - \lambda I$ is not invertible. The complement of $\sigma(T)$ is called the *resolvent set* of T and is denoted by $\rho(T)$. The *spectral radius* of T, $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$. While it is more pedantic to write λI, it is customary to omit the identity and write just λ for the operator λI. As usual, the interest in the spectrum of a linear operator T is motivated by the finite dimensional case. In that situation, $\lambda \in \sigma(T)$ iff λ is an eigenvalue of T, and eigenvalues play an essential role in the structure theory via the Jordan form. As we will see, the situation is quite different in the infinite dimensional Hilbert space.

EXAMPLE 3.2.1. Let $T = \operatorname{diag}(c_n)$. If $\lambda = c_n$ for some n, then $T - \lambda$ has non-trivial kernel (containing e_n) so the spectrum contains the whole diagonal. Is there more? If $T = \operatorname{diag}(1/n)$ then T is not invertible so 0 belongs to the spectrum of T, although it is not one of the diagonal entries and not an eigenvalue. What about the sequence $\{c_n\} = (1/2, 1/3, 2/3, 1/4, 3/4, 1/5, 4/5, \dots)$? The operator $T = \operatorname{diag}(c_n)$ is not invertible, but neither is $T-1$, so both 0 and 1 belong to the spectrum of T. Should we include limit points of the sequence as well? The truth is, we cannot address the problem before we establish some essential properties of the spectrum.

PROPOSITION 3.2.1. *If T is an operator on Hilbert space $\mathcal{H}$ then $r(T) \leq \|T\|$.*

PROOF. If $|\lambda| > \|T\|$ then $\|T/\lambda\| < 1$. By Theorem 3.1.5, the operator $I - T/\lambda$ is invertible, so $\lambda \notin \sigma(T)$. $\square$

EXAMPLE 3.2.2. Let S^* be the backward shift on ℓ^2 (see Example 2.2.2). If $|\lambda| < 1$ then the sequence $u = \{\lambda^n\}$ is in ℓ^2 and it is an eigenvector of S^*, i.e., $S^*u = \lambda u$ so $S^* - \lambda$ has non-trivial kernel and is not invertible. Consequently, the spectrum of S^* contains the open unit disk. On the other hand, $\|S^*\| = \|S\| = 1$ so, by Proposition 3.2.1, $\sigma(S^*)$ is contained in the closed unit disk.

Example 3.2.2 raises once again the question whether the spectrum must contain its boundary points.

THEOREM 3.2.2. *If T is an operator on Hilbert space $\mathcal{H}$ then $\sigma(T)$ is a non-empty compact set.*

PROOF. Proposition 3.2.1 shows that the spectrum of T is bounded. To show that it is closed, we will show that $\rho(T)$ is open. Let $\lambda_0 \in \rho(T)$ so that $T - \lambda_0$ is invertible. Since

$$\|1 - (T-\lambda_0)^{-1}(T-\lambda)\| = \|(T-\lambda_0)^{-1}\left[T - \lambda_0 - (T-\lambda)\right]\| = \|(T-\lambda_0)^{-1}\||\lambda - \lambda_0|$$

we see that $\|1-(T-\lambda_0)^{-1}(T-\lambda)\| < 1$ if $|\lambda - \lambda_0|$ is sufficiently small. By Theorem 3.1.5, for such λ the operator $(T-\lambda_0)^{-1}(T-\lambda)$ is invertible so the same is true of $T-\lambda$. Consequently $\rho(T)$ is open. $\square$

Our next goal is to show that the spectrum of a bounded operator cannot be empty. In order to do that, let $x, y \in \mathcal{H}$, and consider the complex-valued function $F(\lambda) = \langle (T-\lambda)^{-1}x, y\rangle$ defined for $\lambda \in \rho(T)$.

PROPOSITION 3.2.3. *The function F is analytic in $\rho(T) \cup \{\infty\}$.*

PROOF. Let $\lambda_0 \in \rho(T)$. Write

$$T - \lambda = (T-\lambda_0) - (\lambda - \lambda_0) = (T-\lambda_0)\left[1 - (T-\lambda_0)^{-1}(\lambda-\lambda_0)\right]$$

and notice that if $|\lambda - \lambda_0|$ is sufficiently small, then $\|(T-\lambda_0)^{-1}(\lambda-\lambda_0)\| < 1$. By Exercise 3.1.6, we can write

$$(T-\lambda)^{-1} = (T-\lambda_0)^{-1}\sum_{n=0}^{\infty}(T-\lambda_0)^{-n}(\lambda-\lambda_0)^n.$$

Therefore, the function $F(\lambda) = \sum_{n=0}^{\infty}\langle (T-\lambda_0)^{-n-1}x, y\rangle(\lambda-\lambda_0)^n$ is analytic in a neighborhood of λ_0. As for $\lambda_0 = \infty$, we consider the function

$$G(\lambda) = F(1/\lambda) = \langle (T-1/\lambda)^{-1}x, y\rangle \tag{3.1}$$

at $\lambda = 0$. Since $T - 1/\lambda = -(1-\lambda T)/\lambda$, for $\lambda \neq 0$, Theorem 3.1.5 and Exercise 3.1.6 show that, for λ sufficiently small (but different from 0), the operator $T - 1/\lambda$ is invertible and $G(\lambda) = -\lambda\sum_{n=0}^{\infty}\langle T^n x, y\rangle\lambda^n$ is analytic at 0. Furthermore, $F(\infty) = G(0) = 0$. If the spectrum of T were empty, F would be an entire function that is bounded, hence by Liouville's Theorem, a constant. Since $F(\infty) = 0$ it would follow that F is a zero function for any $x, y \in \mathcal{H}$, which is impossible. (Take $x = (T-\lambda I)y$, $y \neq 0$.) Thus $\sigma(T)$ is non-empty. □

Now we can return to Example 3.2.2 and conclude that the spectrum of S^* is the closed unit disk. What about $\sigma(S)$?

Exercise 3.2.1. A complex number λ belongs to $\sigma(T)$ iff $\overline{\lambda} \in \sigma(T^*)$.

Exercise 3.2.2. Given a non-empty compact set $F \subset \mathbb{C}$, show that there exists an operator $T \in \mathcal{L}(\mathcal{H})$ such that $\sigma(T) = F$.

Example 3.2.3. The spectrum of the unilateral shift S is the closed unit disk. However, S has no eigenvalues.

Theorem 3.2.4 (Spectral mapping theorem). *Let $T \in \mathcal{L}(\mathcal{H})$ and let p be a polynomial. Then $\sigma(p(T)) = p(\sigma(T))$.*

Proof. Suppose that $\lambda_0 \in \sigma(T)$, and write $p(\lambda) - p(\lambda_0) = (\lambda-\lambda_0)q(\lambda)$. Then $p(T) - p(\lambda_0) = (T-\lambda_0)q(T)$ and it is not hard to see that the operator $A = p(T) - p(\lambda_0)$ cannot be invertible. Otherwise, we would have that

$T - \lambda_0$ has both the left inverse $A^{-1}q(T)$ and the right inverse $q(T)A^{-1}$. Thus $p(\lambda_0) \in \sigma(p(T))$, and we obtain that $p(\sigma(T)) \subset \sigma(p(T))$.

To prove the converse, let $\lambda_0 \in \sigma(p(T))$, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the roots of $p(\lambda) = \lambda_0$. Then $p(T) - \lambda_0 = \alpha(T - \lambda_1)(T - \lambda_2)\ldots(T - \lambda_n)$ for some non-zero complex number α. Since $p(T) - \lambda_0$ is not invertible there exists j, $1 \leq j \leq n$, such that $T - \lambda_j$ is not invertible. For this j, $\lambda_j \in \sigma(T)$ and $p(\lambda_j) = \lambda_0$ so $\lambda_0 \in (\sigma(T))$. Consequently, $\sigma(p(T)) \subset p(\sigma(T))$ and the theorem is proved. □

Exercise 3.2.3. Let X and T be operators in $\mathcal{L}(\mathcal{H})$, and suppose that X is invertible. Then $\sigma(X^{-1}TX) = \sigma(T)$.

In many instances it is quite hard to determine the spectrum of an operator. However, it may be possible to determine its spectral radius, using the next result.

Theorem 3.2.5 (Spectral Radius Formula). *Let $T \in \mathcal{L}(\mathcal{H})$. Then $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$.*

Proof. By the Spectral mapping theorem, $\sigma(T^n) = [\sigma(T)]^n$ so $[r(A)]^n = r(A^n) \leq \|A^n\|$. Thus, $r(A) \leq \|A^n\|^{1/n}$ and $r(A) \leq \liminf_{n\to\infty} \|A^n\|^{1/n}$. In order to prove the converse we consider the function $G(\lambda)$ defined by (3.1) for $\lambda \neq 0$ and $1/\lambda \in \rho(T)$. For such λ, G is analytic by Proposition 3.2.3 and it can be represented by the convergent series $-\lambda \sum_{n=0}^{\infty} \lambda^n \langle T^n x, y\rangle$. Thus, the sequence $\lambda^n \langle T^n x, y\rangle$ must be bounded. That means that for each y, the sequence of bounded linear functionals $\{\lambda^n T^n x\}$ is bounded at y, i.e., there exists $C(y)$ such that $|\langle \lambda^n T^n x, y\rangle| \leq C(y)$. By the Uniform Boundedness Principle, the sequence $\{\lambda^n T^n x\}$ is uniformly bounded. This means that, for each x, there exists $C(x)$, such that $\|\lambda^n T^n x\| \leq C(x)$. Applying the Uniform Boundedness Principle once again, we obtain $M > 0$ such that $|\lambda|^n \|T^n\| \leq M$, $n \in \mathbb{N}$. It follows that $|\lambda| \|T^n\|^{1/n} \leq M^{1/n}$ and $|\lambda| \limsup_{n\to\infty} \|T^n\|^{1/n} \leq 1$. Since this is true for any λ such that $1/\lambda \in \rho(T)$ it holds all the more whenever $1/|\lambda| > r(T)$. It follows that $\limsup_{n\to\infty} \|T^n\|^{1/n} \leq r(T)$ and the theorem is proved. □

3.3. Parts of the spectrum

A combination of Theorems 3.1.2 and 3.1.4 established that an operator is invertible iff it is bounded below and has closed range.

DEFINITION 3.3.1. A complex number λ belongs to the *approximate point spectrum* $\sigma_{app}(T)$ of a linear operator T if $T-\lambda$ is not bounded below. It belongs to the *compression spectrum* $\sigma_{comp}(T)$ of T if the closure of $\text{Ran}\,(T-\lambda)$ is a proper subspace of $\mathcal{H}$. Finally, it belongs to $\sigma_p(T)$ — the *point spectrum* of T, if it is an eigenvalue of T.

REMARK 3.3.1. There is more than one classification of the parts of the spectrum. The *residual spectrum* is $\sigma_{comp}(T)-\sigma_p(T)$, and the *continuous spectrum* is $\sigma(T)-(\sigma_{comp(T)}\cup\sigma_p(T))$. The *left spectrum* consists of those complex numbers λ such that $T-\lambda$ is not left invertible, and similarly for the *right spectrum*.

EXAMPLE 3.3.1. Let $T=\text{diag}(c_n)$. First we notice that T is invertible iff the sequence $\{c_n\}$ is invertible. Indeed, if $c_nd_n=1$, and $d_n\in\ell^\infty$, define $T^{-1}=\text{diag}(d_n)$. If T is invertible, then $T^{-1}e_n=e_n/c_n$ so $1/|c_n|=\|T^{-1}e_n\|\leq\|T^{-1}\|$ shows that $1/c_n\in\ell^\infty$. Therefore, $\lambda\in\sigma(T)$ iff $c_n-\lambda$ is not invertible, which is true iff there exists a subsequence $\{c_{n_k}\}$ such that $c_{n_k}-\lambda\to 0$. In other words, if and only if λ is an accumulation point of $\{c_n\}$. Thus $\sigma(T)$ is the closure of the diagonal.

What are the parts of the spectrum of $\text{diag}(c_n)$? Suppose that $\|(T-\lambda)x\|\geq\alpha\|x\|$. Then $\sum_{i=1}^\infty|(c_n-\lambda)x_n|^2\geq\alpha\sum_{i=1}^\infty|x_n|^2$. By taking $x=e_n$ we obtain that $|c_n-\lambda|\geq\sqrt{\alpha}$, for all $n\in\mathbb{N}$, which means that $c_n-\lambda$ is invertible and, hence, $\lambda\notin\sigma(T)$. This shows that $\sigma(T)\subset\sigma_{app}(T)$ and therefore $\sigma(T)=\sigma_{app}(T)$.

The previous example is a special case of a more general result.

THEOREM 3.3.1. *If T is a normal operator then $\sigma(T)=\sigma_{app}(T)$.*

PROOF. By Proposition 2.7.5, taking into account that $T-\lambda$ is normal, for any $x\in\mathcal{H}$, $\|(T-\lambda)x\|=\|(T^*-\overline{\lambda})x\|$ so $\sigma_p(T)=\overline{\sigma_p(T^*)}$. Also, $\lambda\in\sigma_p(T^*)\Leftrightarrow\text{Ker}\,(T^*-\lambda)\neq(0)\Leftrightarrow\text{Ran}\,(T^*-\lambda)^*$ is not dense $\Leftrightarrow\text{Ran}\,(T-\overline{\lambda})$

is not dense $\Leftrightarrow \overline{\lambda} \in \sigma_{comp}(T)$. Conclusion: $\sigma_{comp}(T) \subset \sigma_p(T) \subset \sigma_{app}(T)$. Since $\sigma(T) = \sigma_{app}(T) \cup \sigma_{comp}(T)$ the result follows. □

Remark 3.3.2. The proof of Theorem 3.3.1 established that a complex number λ belongs to $\sigma_p(T^*)$ iff $\overline{\lambda} \in \sigma_{comp}(T)$.

Exercise 3.3.1. If $T \in \mathcal{L}(\mathcal{H})$ then $\sigma(T) = \sigma_{app}(T) \cup \sigma_{comp}(T)$.

Since the spectrum is the union of two parts, it is interesting that its boundary is always in the same one.

Theorem 3.3.2. *The boundary of the spectrum is included in the approximate point spectrum.*

Proof. Let $\lambda \in \partial\sigma(T)$. The spectrum of T is closed so $\lambda \in \sigma(T)$, which means that either $\lambda \in \sigma_{app}(T)$ (in which case there is nothing to prove) or $\lambda \in \sigma_{comp}(T)$. In the latter case there exists a non-zero vector x orthogonal to $\operatorname{Ran}(T - \lambda)$. Let $\{\lambda_n\} \subset \rho(T)$ such that $\lambda_n \to \lambda$. Since $T - \lambda_n$ is invertible, we can define unit vectors $f_n = (T - \lambda_n)^{-1} f / \|(T - \lambda_n)^{-1} f\|$. Now

$$\|(T-\lambda)f_n\|^2 \leq \|(T-\lambda)f_n\|^2 + \|(T-\lambda_n)f_n\|^2 = \|(\lambda-\lambda_n)f_n\|^2 = |\lambda-\lambda_n| \to 0$$

where we have used the fact that $(T - \lambda)f_n$ is a multiple of f, hence orthogonal to $(T - \lambda_n)f_n$. Consequently, $\lambda \in \sigma_{app}(T)$. □

Example 3.3.2. We have seen in Example 3.2.3 that the spectrum of the unilateral shift S is $\mathbb{D}^-$. By Exercise 3.2.1 the same is true of $\sigma(S^*)$. Since S is an isometry, 0 canot be an eigenvalue of S. ($Sx = 0$ imnplies $\|x\| = \|Sx\| = 0$.) If $\lambda \neq 0$ then $S(x_1, x_2, \dots) = \lambda(x_1, x_2, \dots)$ leads to $0 = \lambda x_1$ and $x_n = \lambda x_{n+1}$, $n \in \mathbb{N}$, and we see that $x = 0$. Therefore, $\sigma_p(S)$ is empty and, by Exercise 3.3.2, so is $\sigma_{comp}(S^*)$.

The equation $S^*x = \lambda x$ leads to $x_{n+1} = \lambda x_n$, $n \in \mathbb{N}$, and thus to $x = x_1(1, \lambda, \lambda^2, \dots)$. Therefore, x is a non-zero vector in ℓ^2 iff $|\lambda| < 1$. Consequently, $\sigma_p(S^*) = \sigma_{comp}(S) = \mathbb{D}$.

By Theorem 3.3.2, the approximate point spectra of S and S^* include the unit circle $\mathbb{T}$. For S that is all because, if $|\lambda| < 1$ then $\|Sx - \lambda x\| \geq |\|Sx\| - \|\lambda x\|| = (1 - |\lambda|)\|x\|$ so $S - \lambda$ is bounded below. On the other hand, the approximate point spectrum always includes the eigenvalues, so $\sigma_{app}(S^*) = \mathbb{D}^-$.

THEOREM 3.3.3. *Suppose that $\mathcal{M}$ is a closed subspace of Hilbert space $\mathcal{H}$, and that, relative to $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$, $T = \left[\begin{smallmatrix} T_1 & 0 \\ 0 & T_2 \end{smallmatrix}\right]$. Then $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$.*

PROOF. If $T - \lambda$ is not invertible then $T_1 - \lambda$ and $T_2 - \lambda$ cannot both be invertible, so $\sigma(T) \subset \sigma(T_1) \cup \sigma(T_2)$. On the other hand, if either T_1 or T_2 is not bounded below, say $\|T_1 x_n\| \to 0$, then $\|T(x_n \oplus 0)\| \to 0$, so $\sigma_{app}(T_1) \cup \sigma_{app}(T_2) \subset \sigma(T)$. The corresponding inclusion for the compression spectra can be obtained by switching to the adjoints and using Exercise 3.3.2. □

PROBLEM 11. Suppose that $\mathcal{H} = \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \dots$ and that relative to this decomposition $T = \operatorname{diag}(T_n)$ is a diagonal matrix with operator entries $T_1, T_2, \dots$. Is it true that $\sigma(T) = (\cup \sigma(T_n))^-$?

3.4. Spectrum of a compact operator

In this section we take a more detailed look at compact operators and their spectra.

THEOREM 3.4.1. *Let T be a compact operator, let λ be a non-zero complex number, and suppose that $T - \lambda$ is not bounded below. Then $\lambda \in \sigma_p(T)$.*

PROOF. Let $\{x_n\}$ be a sequence of unit vectors such that $\|(T - \lambda)x_n\| \to 0$, $n \to \infty$. Since B_1 is weakly compact, $\{x_n\}$ has a weakly convergent subsequence $\{x_{n_k}\}$, so the compactness of T implies that $\{T x_{n_k}\}$ is a convergent sequence. Let $x = \lim_k T x_{n_k}$. Notice that $\|x\| \geq \|\lambda x_{n_k}\| - \|(T - \lambda)x_{n_k}\| \to |\lambda|$ so x is a non-zero vector. Moreover, $\|(T - \lambda)x\| \leq \|(T - \lambda)(T x_{n_k} - x)\| + \|(T - \lambda)T x_{n_k}\| \to 0$ so $\lambda \in \sigma_p(T)$. □

Theorem 3.4.1 established that the non-zero points in the approximate point spectrum are eigenvalues. Our goal is to prove a similar inclusion for the compression spectrum. We start with the following result.

THEOREM 3.4.2. *Let T be a compact operator and let λ be a non-zero complex number. Then* $\operatorname{Ran}(T - \lambda)$ *is closed.*

PROOF. First we show that, if $\operatorname{Ran} T$ is closed, it must be finite dimensional. Indeed, if we denote by T_1 the restriction of T to its initial space $(\operatorname{Ker} T)^\perp$, then T_1 is an injective linear transformation from $(\operatorname{Ker} T)^\perp$ onto

$\operatorname{Ran} T$, hence invertible. Let B be the intersection of the closed ball of radius $\|T_1^{-1}\|$ and $\operatorname{Ran} T$. Now, if $y \in B$ then $y = T_1 x$, for some $x \in (\operatorname{Ker} T)^\perp$, so $x = T_1^{-1} y$. Since $\|y\| \leq \|T_1^{-1}\|$ it follows that $x \in B_1 \cap (\operatorname{Ker} T)^\perp$. We conclude that B is contained in the compact set $T(B_1 \cap (\operatorname{Ker} T)^\perp)$ so B must be compact, hence finite dimensional.

Next we observe that $\operatorname{Ker}(T - \lambda)$ must be finite dimensional. Reason: $\operatorname{Ker}(T - \lambda)$ is invariant for T and the restriction of T to $\operatorname{Ker}(T - \lambda)$ is a compact operator with range $\operatorname{Ker}(T - \lambda)$. (If $x \in \operatorname{Ker}(T - \lambda)$ write $x = \frac{1}{\lambda}[Tx - (T - \lambda)x] = T(x/\lambda) \in T(\operatorname{Ker}(T - \lambda))$.)

Finally, we prove the theorem. Let S be the restriction of $T - \lambda$ to $\operatorname{Ker}(T - \lambda)^\perp$. Notice that $\operatorname{Ran} S = \operatorname{Ran}(T - \lambda)$ so it suffices to show that $\operatorname{Ran} S$ is closed. By Theorem 3.1.2 we will accomplish this goal by establishing that S is bounded below. However, if S is not bounded below then Theorem 3.4.1 shows that $(T - \lambda)x = 0$ for some nonzero vector x in $\operatorname{Ker}(T - \lambda)^\perp$. This is impossible, so $\operatorname{Ran} S$ is closed and the proof is complete. $\square$

Before we can proceed we need this technical result.

LEMMA 3.4.3. *Let T be a compact operator and let $\{\lambda_n\}$ be a sequence of complex numbers. Suppose that there exists a nested sequence of distinct subspaces $\mathcal{M}_1 \subsetneq \mathcal{M}_2 \subsetneq \mathcal{M}_3 \subsetneq \dots$ such that $(T - \lambda_n)\mathcal{M}_{n+1} \subset \mathcal{M}_n$. Then λ_n converges to 0.*

PROOF. Let $\{e_n\}$ be an sequence of unit vectors such that $e_1 \in \mathcal{M}_1$ and $e_{n+1} \in \mathcal{M}_{n+1} \ominus \mathcal{M}_n$. Clearly, this is an orthonormal system. Moreover, for $n \geq 2$, $\langle (T - \lambda_n)e_n, e_n \rangle = 0$ which implies that $\|Te_n\| \geq |\langle Te_n, e_n \rangle| = |\langle (T - \lambda_n)e_n, e_n \rangle + \langle \lambda_n e_n, e_n \rangle| = |\lambda_n|$. Since T is compact and $w - \lim e_n = 0$ it follows that $\lim_n Te_n = 0$ so $\lim_n \lambda_n = 0$. $\square$

Theorem 3.4.1 shows that if $\lambda \in \sigma(T)$ then either $\lambda = 0$, or $\lambda \in \sigma_p(T)$, or $T - \lambda$ is bounded below (hence injective) but not surjective. By Theorem 3.1.4, $T - \lambda$ not being surjective is the same as $(T - \lambda)^*$ not being bounded below. Since T^* is also compact, another application of Theorem 3.4.1 allows us to conclude that $\overline{\lambda} \in \sigma_p(T^*)$. The next result shows that there is even less variation in the spectrum of a compact operator.

Theorem 3.4.4. *Let T be a compact operator and let λ be a non-zero complex number. Then $\lambda \in \sigma_p(T)$ iff $\overline{\lambda} \in \sigma_p(T^*)$.*

Proof. Clearly, it suffices to prove either direction. Suppose that $\lambda \in \sigma_p(T)$. By Theorem 3.4.2, the range of $T - \lambda$ is closed. We will show that it must be a proper subspace of $\mathcal{H}$. Suppose to the contrary that $T - \lambda$ is surjective, and denote $\mathcal{M}_n = \operatorname{Ker}(T - \lambda)^n$. Since λ is an eigenvalue of T we can inductively define a sequence $\{x_n\}$ of nonzero vectors such that $(T - \lambda)x_n = x_{n-1}$, with $x_0 = 0$. Clearly x_n belongs to $\mathcal{M}_n$ but not to $\mathcal{M}_{n-1}$, and $(T - \lambda)\mathcal{M}_{n+1} \subset \mathcal{M}_n$, so Lemma 3.4.3 implies that the constant sequence $\lambda, \lambda, \lambda, \dots$ converges to 0, which contradicts the assumption that $\lambda \neq 0$. Therefore, $\operatorname{Ran}(T - \lambda)$ (which coincides with $\operatorname{Ker}(T - \lambda)^*$) is a proper subspace of $\mathcal{H}$ and $\overline{\lambda} \in \sigma_p(T^*)$. □

To summarize, the spectrum of a compact operator consists of the point spectrum and, possibly, 0. On the infinite dimensional Hilbert space, 0 must be in the spectrum because if a compact operator T were invertible, then so would be the identity (a product of TT^{-1}), contradicting the conclusions of Example 2.6.1. Thus we have a corollary.

Corollary 3.4.5. *The spectrum of a compact operator consists of* 0 *and its eigenvalues.*

It is reasonable to ask about the location of the eigenvalues.

Theorem 3.4.6. *For any $C > 0$ there is a finite number of linearly independent eigenvectors of a compact operator corresponding to eigenvalues λ such that $|\lambda| \geq C$.*

Proof. Suppose to the contrary that there is an infinite sequence $\{x_n\}$ of unit vectors, and a sequence of eigenvalues λ_n of T, $|\lambda_n| \geq C$, so that $Tx_n = \lambda_n x_n$. Let $\mathcal{M}_n = \vee_{k=1}^n x_k$. If $x \in \mathcal{M}_n$ then $x = \sum_{k=1}^n c_k x_k$ so $(T - \lambda_n)x = (T - \lambda_n)\sum_{k=1}^n c_k x_k = \sum_{k=1}^n c_k (T - \lambda_n) x_k = \sum_{k=1}^n c_k(\lambda_k - \lambda_n)x_k \in \mathcal{M}_{n-1}$. Applying Lemma 3.4.3 we obtain that $\lambda_n \to 0$, which contradicts $|\lambda_n| \geq C$. □

Corollary 3.4.7. *If λ is a non-zero eigenvalue of a compact operator T, then the nullspace of $T - \lambda$ is a finite dimensional subspace.*

COROLLARY 3.4.8. *The spectrum of a compact operator T is at most countable, and the only accumulation point of it can be zero.*

REMARK 3.4.1. If $T = \mathrm{diag}(c_n)$ where $c_1 = 1$ and $c_n = 0$ for $n \geq 2$, then T is compact, and $\sigma(T) = \{0, 1\}$ so it has no accumulation points.

Last remark raises a question: can a compact operator have a one-point spectrum? Since compact operators are never invertible, the single point is necessarily 0, so the question can be reformulated as: are there compact *quasinilpotent* operators? (An operator T is quasinilpotent if $\sigma(T) = \{0\}$.) In finite dimensions, a quasinilpotent operator is nilpotent, i.e. there exists a positive integer N such that $T^N = 0$. This need not be the case in infinite dimensional Hilbert space.

EXAMPLE 3.4.1. Let T be a weighted shift (see Example 2.1.6) with weight sequence $\{1/n\}_{n\in\mathbb{N}}$. It is compact following Example 2.6.1. Since $We_n = (1/n)e_{n+1}$ it follows that $W^k e_n = \frac{1}{n(n+1)\ldots(n+k-1)} e_{n+k}$. This shows that W^k is a product of S^k and a $\mathrm{diag}(\frac{1}{n(n+1)\ldots(n+k-1)})$. Since S^k is an isometry, $\|W^k\| = \sup_n \{\frac{1}{n(n+1)\ldots(n+k-1)}\} = 1/k!$. Now $r(W) = \lim_k \|W^k\|^{1/k} = \lim_k (1/k!)^{1/k} = 0$. Therefore, W is a compact quasinilpotent operator.

3.5. Spectrum of a normal operator

On the first glance, normal operators appear to be too diverse to fit one description. Before we can correct this misconception, we will need to make a thorough study of this class, and some of its prominent subclasses.

THEOREM 3.5.1. *(a) If T is a unitary operator then $\sigma(T)$ is a subset of the unit circle. (b) If T is a self-adjoint operator then $\sigma(T)$ is a subset of the real axis. (c) If T is a positive operator then $\sigma(T)$ is a subset of the non-negative real axis. (d) If T is a non-trivial projection then $\sigma(T) = \{0, 1\}$.*

PROOF. All operators listed are normal, so by Theorem 3.3.1, it suffices to prove assertions (a) – (d) with $\sigma_{app}(T)$ instead of $\sigma(T)$. To that end, we will prove that, if λ does not belong to the appropriate set, then $T - \lambda$ is bounded below.

(a) If T is unitary and $|\lambda| \neq 1$, then $\|Tx - \lambda x\| \geq |\|Tx\| - \|\lambda x\|| = |(1 - \lambda)| \, \|x\|$ so T is bounded below.

(b) Let $\lambda = \alpha + i\beta$. Then $\|Tx - \lambda x\|^2 = \|Tx - \alpha x\|^2 - 2\mathrm{Re}\langle Tx - \alpha x, i\beta x\rangle + \|i\beta x\|^2$. If α, β are real numbers and $T = T^*$ we have that $\langle T - \alpha x, x\rangle \in \mathbb{R}$ by Proposition 2.7.1, and it follows that $\mathrm{Re}\langle Tx - \alpha x, i\beta x\rangle = 0$. Therefore, $\|Tx - \lambda x\|^2 \geq |\beta|^2 \|x\|^2$, so $\beta \neq 0$ implies that $T - \lambda$ is bounded below.

(c) If $T \geq 0$ then T is self-adjoint, so $\sigma(T) \subset \mathbb{R}$. Notice that $\|Tx - \lambda x\|^2 = \|Tx\|^2 - 2\mathrm{Re}\langle Tx, \lambda x\rangle + \|\lambda x\|^2$. If $\lambda < 0$ then $\langle Tx, \lambda x\rangle < 0$ (by definition of a positive operator) so $\|Tx - \lambda x\|^2 \geq |\lambda|^2 \|x\|^2$ and $T - \lambda$ is bounded below.

(d) If T is a non-trivial projection then neither T nor $I - T$ (the projection on the orthogonal complement of the range of T) can be invertible, so $\{0, 1\} \subset \sigma(T)$. If $\lambda \notin \{0, 1\}$, a calculation shows that $\frac{1}{\lambda(1-\lambda)} T - \frac{1}{\lambda}$ is the inverse of T. □

Exercise 2.7.1 asserts that the operator of multiplication by an L^∞ function is a normal operator. In addition, it showed that M_h belongs to one of the important subclasses iff its (essential) range belonged to a specific subset of the complex plane. On the other hand, Theorem 3.5.1 showed that for a general normal operator, a membership in each of the mentioned subclasses implies the analogous behavior of its spectrum. This is no coincidence. First we need a proposition.

Proposition 3.5.2. *Let $T = M_h$ on L^2. Then the following are equivalent:*

(a) $\mathrm{Ran}\, T$ *is dense;*

(b) $h(x) \neq 0$ *a.e.;*

(c) T *is injective;*

(d) T^* *is injective.*

Proof. Let $A = \{x : h(x) = 0\}$. Suppose that $\mu(A) \neq 0$ and let $f = \chi_A$. For any $g \in L^2$, $\langle Tg, f\rangle = \int hg\overline{f} = \int_A hg = 0$ so f is a non-zero function that is orthogonal to $\mathrm{Ran}\, T$. Thus (a) implies (b). Next, if $Tf = 0$ then $h(x)f(x) = 0$ a.e., so assuming (b) we see that $f = 0$, and (c) follows. Notice that if $T^* f = 0$ then $\overline{h(x)} f(x) = 0$ so $T\overline{f} = 0$ and (c) implies (d). Finally, the implication (d) $\Rightarrow$ (a) is a direct consequence of Theorem 2.2.3. □

Recall that the essential range of a function $h \in L^\infty(X, \mu)$ is the set of all complex numbers z such that the measure of $E_\epsilon = \{x \in X : |h(x) - z| < \epsilon\}$ is different from zero for all $\epsilon > 0$.

THEOREM 3.5.3. *Let $T = M_h$ on L^2. Then $\sigma(T)$ is the essential range of h.*

PROOF. Notice that $M_h - \lambda$ is a multiplication by $h - \lambda$. Let us denote by $A = \{x : h(x) \neq \lambda\}$, $B = \{x : h(x) = \lambda\}$, and define a function $g(x) = 1/(h(x) - \lambda)$ if $x \in A$ and $g(x) = 0$ if $x \in B$.

Suppose first that $\lambda \in \rho(T)$. By Proposition 3.5.2, $\mu(B) = 0$. Thus, $g(x) = 1/(h(x) - \lambda)$ a.e. and $M_g M_{h-\lambda} = M_{h-\lambda} M_g = I$. Since the assumption is that $M_{h-\lambda}$ is invertible, the operator M_g is bounded, and by Example 2.1.4, $g \in L^\infty$. The estimate $|g(x)| \leq M$ a.e. implies that $|h(x) = \lambda| \geq 1/M$ a.e. so $\mu(E_{1/M}) = 0$ and λ is not in the essential range of h.

Conversely, if λ is not in the essential range of h, then there exists $\epsilon_0 > 0$ such that $\mu(E_{\epsilon_0}) = 0$. Consequently, $|h(x) - \lambda| \geq \epsilon_0$ a.e., whence $|g(x)| \leq 1/\epsilon_0$ a.e., and M_g is a bounded operator. This shows that $M_{h-\lambda}$ is invertible and the proof is complete. □

Proposition 3.2.1 established that $r(T) \leq \|T\|$. For normal operators more can be said, and the following result paves the way to that goal.

PROPOSITION 3.5.4. *If T is a normal operator then $\|T^n\| = \|T\|^n$, $n \in \mathbb{N}$.*

PROOF. First we notice that, in view of Proposition 2.7.5, for $n \in \mathbb{N}$,

$$\|T^n x\|^2 = \langle T^n x, T^n x\rangle = \langle T^* T^n x, T^{n-1} x\rangle \leq \|T^* T^n x\| \|T^{n-1} x\| = \|T^{n+1} x\| \|T^{n-1} x\| \leq \|T^{n+1}\| \|T^{n-1}\| \|x\|^2$$

so $\|T^n\|^2 \leq \|T^{n+1}\| \|T^{n-1}\|$.

Now we prove the assertion of the proposition using induction. We will assume that $\|T\| \neq 0$, otherwise the theorem is trivially correct. It is easy to see that the statement is valid for $n = 0$ and $n = 1$. Suppose that it is true for n. Then

$$\|T\|^{2n} = (\|T\|^n)^2 = \|T^n\|^2 \leq \|T^{n+1}\| \|T^{n-1}\| \leq \|T^{n+1}\| \|T\|^{n-1}$$

and, dividing both sides by $\|T\|^{n-1}$, it follows that $\|T\|^{n+1} \leq \|T^{n+1}\|$. Since the opposite inequality is obvious, the theorem is proved. □

COROLLARY 3.5.5. *If T is a normal operator then $r(T) = \|T\|$.*

PROOF. By Theorems 3.2.5 and 3.5.4, $\|T\| = \sqrt[n]{\|T^n\|} \to r(T)$. □

CHAPTER 4

Invariant subspaces

4.1. Compact operators

We have seen that the spectrum of a compact operator consists of the eigenvalues and 0 which may be but is not necessarily an eigenvalue. Furthermore, each of the eigenspaces $E(\lambda) = \mathrm{Ker}\,(T - \lambda)$, corresponding to $\lambda \neq 0$, is finite dimensional. The situation is especially pleasant when T is self-adjoint, in addition to being compact. One of the benefits of this additional hypothesis concerns the eigenspaces.

Proposition 4.1.1. *If T is a compact, self-adjoint operator on Hilbert space, and if λ, μ are two different eigenvalues of T, then the corresponding eigenspaces $E(\lambda), E(\mu)$ are mutually orthogonal.*

Proof. If $Tx = \lambda x$ and $Ty = \mu y$, then $\lambda\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \mu\langle x, y\rangle$, since $\mu \in \mathbb{R}$. Given that $\lambda \neq \mu$ it follows that $\langle x, y\rangle = 0$. □

Proposition 4.1.1 shows that $\mathcal{H}$ can be written as a direct sum $\mathcal{M} \oplus \mathcal{M}^\perp$, where $\mathcal{M} = \oplus_{n\in\mathbb{N}} E(\lambda_n)$, the orthogonal direct sum of all eigenspaces. When T is self-adjoint, the subspace $\mathcal{M}^\perp$ is just a mirage.

Theorem 4.1.2. *If T is a compact, self-adjoint operator on $\mathcal{H}$ space, and $\sigma_p(T) = \{\lambda_i\}_{i\in I}$, then $\mathcal{H} = \oplus_{i\in I} E(\lambda_i)$.*

Proof. Let $\mathcal{M} = \oplus_{i\in I} E(\lambda_i)$ and suppose that $\mathcal{M} \neq \mathcal{H}$. Notice that $\mathcal{M}$ is invariant for $T = T^*$, so $\mathcal{M}^\perp$ is also reducing for T. Let T_1 be the restriction of T to $\mathcal{M}^\perp$. Then $\sigma(T_1) \subset \sigma(T)$ by Theorem 3.3.3. Since T_1 is compact, if $\lambda \neq 0$ is in its spectrum it must be an eigenvalue. However, the corresponding eigenvectors would also be eigenvectors of T and, as such, would belong to $\mathcal{M}$. It follows that T_1 must be quasinilpotent. On the other hand T_1 is normal which would necessitate that its norm and spectral radius are equal, so $T_1 = 0$ which means that $\mathcal{M}^\perp \subset E(0) \subset \mathcal{M}$. The obtained contradiction shows that $\mathcal{H} = \oplus_{i\in I} E(\lambda_i)$. □

REMARK 4.1.1. Each eigenspace $E(\lambda)$ is reducing for a self-adjoint operator so, relative to the decomposition $\mathcal{H} = \oplus_{i \in I} E(\lambda_i)$, T can be represented as $\operatorname{diag}(T_i)$, where T_i is an operator mapping $E(\lambda_i)$ into itself, and $\sigma(T_i)$ is a singleton $\{\lambda_i\}$. In addition, regardless of whether T is self-adjoint or not, each eigenspace is *hyperinvariant* for T. This means that it is invariant for any operator that commutes with T. Indeed, if A commutes with T, then $T - \lambda$ annihilates Ax together with x.

When T is not self-adjoint, the situation is much more complicated. The eigenspaces need not be mutually orthogonal any more. The eigenvectors do not necessarily span $\mathcal{H}$. In fact, there are compact operators without eigenvalues, (so they are necessarily quasinilpotent). Still, we can see some of the structure remaining. The eigenspaces are hyperinvariant (if there are any), although they need not be reducing. Since all operators on $\mathbb{C}^n$ are compact, it is instructive to look at finite matrices.

EXAMPLE 4.1.1. Let $T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ acting on $\mathbb{C}^2$. Then $\sigma(T) = \{1\}$ and $E(1) = \mathbb{C} \oplus (0)$ which is neither invariant for T^*, nor is the span of eigenvectors of T equal to $\mathbb{C}^2$.

EXAMPLE 4.1.2. Let $T = \left[\begin{smallmatrix} 2 & 1 \\ 0 & 3 \end{smallmatrix}\right]$ acting on $\mathbb{C}^2$. The eigenvalues of T are 2 and 3, with corresponding eigenvectors $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, and they are not mutually orthogonal.

When T has eigenvalues, it must have a non-trivial invariant subspace. What about the case of a compact quasinilpotent operator?

EXAMPLE 4.1.3. Let T be the Volterra-type integral operator with kernel K, i.e., $Tf(x) = \int_0^x K(x,y) f(y)\, dy$. It is compact (Example 2.6.2) and has no eigenvalues different from 0. Indeed, let $\lambda \in \sigma(T)$, $\lambda \neq 0$ and let $f \in L^2$ be the appropriate eigenfunction. Define $g(x) = \int_0^x |f(y)|^2\, dy$. Clearly, g is a monotone differentiable function and $g'(x) = |f(x)|^2$ a.e. Let $a = \sup\{x \in [0,1] : g(x) = 0\}$. (Since $g(0) = 0$ such a number exists.) Now, for a.e. x,

$$|\lambda f(x)|^2 = |Tf(x)|^2 = \left| \int_0^x K(x,y) f(y)\, dy \right|^2 \leq \int_0^x |K(x,y)|^2\, dy \int_0^x |f(y)|^2\, dy,$$

so $|\lambda|^2 g'(x)/g(x) \le \int_0^x |K(x,y)|^2\,dy$ for a.e. $x \in (a,1)$. By integrating the last inequality we obtain

$$|\lambda|^2 \ln g(x)\,\Big|_a^1 \le \int_a^1 \int_0^x |K(x,y)|^2\,dy \le \|T\|^2$$

which is a contradiction since $\ln g(1) = \ln \|f\|^2$ and $\|T\|$ are finite, but $\ln g(a)$ is not.

This example shows that there are many compact quasinilpotent operators. For the Volterra-type integral operators we can exhibit some invariant subspaces.

THEOREM 4.1.3. *Let T be a Volterra-type integral operator with kernel K, let $a \in [0,1]$, and let $M_a = \{f \in L^2 : f(x) = 0$ when $x \le a\}$. Then M_a is a subspace of L^2 that is invariant for T.*

EXERCISE 4.1.1. Prove Theorem 4.1.3.

A deep result in the theory of integral operators is that every compact quasinilpotent operator is unitarily equivalent to an operator of the form as in Example 4.1.3. Consequently every compact operator (quasinilpotent or not) has an invariant subspace. As we will demonstrate, there is a way to prove an even stronger theorem. (See Theorem 4.3.2 below.)

4.2. Line integrals

In this section we make a brief detour, by considering line integrals of functions of a complex variable with values in $\mathcal{L}(\mathcal{H})$.

EXAMPLE 4.2.1. Let $T \in \mathcal{L}(\mathcal{H})$ and consider the function $\rho(\lambda) = (T-\lambda)^{-1}$ defined for $\lambda \in \rho(T)$. This function is known as the *resolvent* of T.

Let C be a curve in the complex plane. We will assume that it is parametrized by a continuous function $\gamma : [0,1] \to \mathbb{C}$ and that it is *rectifiable*, which means that γ is a function of bounded variation. Suppose that S is a function defined and continuous on C, with values in $\mathcal{L}(\mathcal{H})$. Let P be a partition of $[0,1]$: $0 = t_0 < t_1 < t_2 <$

$\cdots < t_n = 1$ and, for $1 \leq k \leq n$ let $t_k^* \in [t_{k-1}, t_k]$. Then we have a partition of C with points $\gamma_i = \gamma(t_i)$ and intermediate points $\gamma_i^* = \gamma(t_i^*)$. Let us denote $\Delta\gamma_i = \gamma_i - \gamma_{i-1}$ and consider the sum

$$\sum_{k=1}^{n} S(\gamma_k^*)\,\Delta\gamma_k.$$

It can be shown that these sums converge to a unique operator S which we denote as $S = \int_C S(\gamma)\,d\gamma$. Moreover, if T is an operator that commutes with each $S(\gamma)$, then T commutes with S.

EXAMPLE 4.2.2. Let $T \in \mathcal{L}(\mathcal{H})$, and let C be a curve in $\rho(T)$ defined by $\gamma = \gamma(t)$. For every $\lambda \in \rho(T)$, the function $\rho(\lambda)$ is a continuous function (in the uniform topology), so we can consider $\int_C \rho(\gamma)\,d\gamma$.

What happens when the curve C is replaced by a curve C' that is not far from C?

THEOREM 4.2.1. *Let C_0 be a rectifiable curve in the resolvent set of T, and let C_1 be a curve homotopic to C_0. Then $\int_{C_0} \rho(\gamma)\,d\gamma = \int_{C_1} \rho(\gamma)\,d\gamma$.*

REMARK 4.2.1. All these facts can be established following the same procedures as in the case when the integrand is a complex-valued function. [See Conway.]

Now we turn to operators. Example 4.2.2 showed that the operator $\int_C \rho(\gamma)\,d\gamma$ is well defined. It turns out that this operator has some interesting properties.

THEOREM 4.2.2. *Let C be a simple closed rectifiable curve in $\rho(T)$. Then the operator*

$$P = -\frac{1}{2\pi i}\int\limits_C \rho(\lambda)\,d\lambda \tag{4.1}$$

is a projection (not necessarily orthogonal) that commutes with every operator that commutes with T. Consequently, the subspaces $\operatorname{Ran} P$ *and* $\operatorname{Ker} P$ *are both invariant for T.*

PROOF. Let C' be a simple closed rectifiable curve in $\rho(T)$ that lies inside C and is homotopic to C. Then

$$(2\pi i)^2 P^2 = \int_C \rho(\gamma)\,d\gamma \int_{C'} \rho(\lambda)\,d\lambda = \int_C \int_{C'} \rho(\gamma)\rho(\lambda)\,d\gamma d\lambda.$$

A calculation shows that $\rho(\gamma)\rho(\lambda) = [\rho(\gamma) - \rho(\lambda)](\gamma - \lambda)^{-1}$. Thus we have that

$$(2\pi i)^2 P^2 = \int_{C'} \rho(\gamma) \int_C (\gamma - \lambda)^{-1}\, d\lambda d\gamma - \int_C \rho(\lambda) \int_{C'} (\gamma - \lambda)^{-1}\, d\gamma d\lambda = -2\pi i \int_{C'} \rho(\gamma)\, d\gamma - 0 = (2\pi i)^2 P.$$

So, $P^2 = P$, and it follows from the definition of the integral and $\rho(\lambda)$, that if A commutes with T then A commutes with P.

Finally, if $y \in \operatorname{Ran} P$, then $Ty = TPy = PTy$ so $Ty \in \operatorname{Ran} P$. Similarly, if $x \in \operatorname{Ker} P$, then $0 = TPx = PTx$ so $Tx \in \operatorname{Ker} P$. □

Exercise 4.2.1. Verify that $\rho(\gamma)\rho(\lambda) = [\rho(\gamma) - \rho(\lambda)](\gamma - \lambda)^{-1}$.

Theorem 4.2.2 required that the closed curve C lies in $\rho(T)$, but made no reference to the spectrum of T. Consequently, we may have a part of the spectrum inside C and a part outside. In that case we obtain a decomposition of T.

Theorem 4.2.3. *Let T be an operator in $\mathcal{L}(\mathcal{H})$, let C be a simple closed rectifiable curve in $\rho(T)$, let P be the projection defined in (4.1), and let T' and T'' be the restrictions of T to* $\operatorname{Ran} P$ *and* $\operatorname{Ker} P$*, respectively. Then $T = T' + T''$, the spectrum of T' is precisely the subset of $\sigma(T)$ inside C, and the spectrum of T'' is precisely the subset of $\sigma(T)$ outside C.*

Proof. Since $\rho(\lambda)$ commutes with P, for any $\lambda \in \rho(T)$, the subspaces $\operatorname{Ran} P$ and $\operatorname{Ker} P$ are invariant for $\rho(\lambda)$. Let $\rho'(\lambda)$ and $\rho''(\lambda)$ denote the restrictions of $\rho(\lambda)$ to these subspaces. If we denote by I' and I'' the identity operators on these subspaces, then $\rho'(\lambda)(\lambda I' - T') = I'$ and $\rho''(\lambda)(\lambda I'' - T'') = I''$. Therefore, if $\lambda \in \rho(T)$ then λ must belong to both $\rho(T')$ and $\rho(T'')$. In the other direction, if $\lambda \in \rho(T') \cap \rho(T'')$ then there exist operators A' and A'' such that $A'(\lambda I' - T') = I'$ and $A''(\lambda I'' - T'') = I''$. Now we can define, for any $x \in \mathcal{H}$, $Ax = A'Px + A''(I - P)x$. It is not hard to see that the restricitons of A to $\operatorname{Ran} P$ and $\operatorname{Ker} P$ are precisely A' and A'', and that $A(\lambda I - T)x = x$ when x belongs to either $\operatorname{Ran} P$ or $\operatorname{Ker} P$. It follows that $A(\lambda I - T)x = x$ holds for all $x \in \mathcal{H}$, so $\lambda \in \rho(T)$. We conclude that $\lambda \in \sigma(T)$ iff $\lambda \in \sigma(T')$ or $\lambda \in \sigma(T'')$.

Suppose now that λ lies outside of C. We will show that $\lambda \in \rho(T')$, which is true iff there exists an operator A' acting on $\operatorname{Ran} P$ and satsifying $A'(\lambda I' - T') = I'$. Actually, we will show that there exists an operator $A \in \mathcal{L}(\mathcal{H})$ that commutes with T and $A(\lambda I - T) = P$. To that end, we notice that

$$(T-\lambda I)\rho(\gamma) = (T-\lambda I)(T-\gamma I)^{-1} = (T-\gamma I)(T-\gamma I)^{-1} + (\gamma-\lambda)(T-\gamma I)^{-1} = I + (\gamma-\lambda)(T-\gamma I)^{-1}.$$

Therefore,

$$(4.2) \qquad (T-\lambda I)\frac{1}{2\pi i}\int_C \rho(\gamma)(\gamma-\lambda)^{-1}\,d\gamma = \frac{1}{2\pi i}\int_C (\gamma-\lambda)^{-1}\,d\gamma\, I + \frac{1}{2\pi i}\int_C \rho(\gamma)\,d\gamma = 0 - P = -P.$$

On the other hand, if λ lies inside of C, then the integral in (4.2) equals $I - P$, so the restriction to $\operatorname{Ker} P$ yields I''. Once again, this shows that $\lambda I'' - T''$ is invertible. □

4.3. Invariant subspaces for compact operators

In Section 4.1 we have discovered that every compact operator on Hilbert space has an invariant subspace. What more is there to say? For one thing, if λ is an eigenvalue of T, then $E(\lambda)$ is hyperinvariant. Thus, it is natural to ask whether a compact quasinilpotent operator always has a hyperinvariant subspace.

Before we address this question, let us take a look at the set of all operators that commute with T. It is called the *commutant* of T, it is denoted by $\{T\}'$, and it is an *algebra*. The last statement means that $\{T\}'$ is closed under sums, products, and multiplication by scalars.

Exercise 4.3.1. Prove that $\{T\}'$ is an algebra.

Definition 4.3.1. A subalgebra of $\mathcal{L}(\mathcal{H})$ is *transitive* if it is weakly closed, *unital* (containing the identity operator), and has only the trivial invariant subspaces.

Example 4.3.1. The algebra $\mathcal{L}(\mathcal{H})$ is transitive. It is clearly weakly closed and unital. If $\mathcal{L}(\mathcal{H})$ had a non-trivial invariant subspace $\mathcal{M}$, then we could pick non-zero vectors $x \in \mathcal{M}^\perp$ and $y \in \mathcal{M}$, and consider the rank one operator $T = x \otimes y$. This would lead to a contradiciton, since $y \in \mathcal{M}$ but $Ty = (x \otimes y)y = \langle y, y\rangle x \in \mathcal{M}^\perp$.

A big open problem in operator theory is whether $\mathcal{L}(\mathcal{H})$ is the only transitive algebra. This is true when $\mathcal{H}$ is finite dimensional.

THEOREM 4.3.1 (Burnside's Theorem). *Let $\mathcal{H}$ be a finite dimensional vector space of dimension larger than 1. If $\mathcal{A}$ is a transitive algebra of linear transformations on $\mathcal{H}$, then $\mathcal{A} = \mathcal{L}(\mathcal{H})$.*

PROOF. We will show that $\mathcal{A}$ contains a rank one operator. Let T_0 be an operator with minimal non-zero rank d. If $d > 1$, choose x_1 and x_2 so that vectors T_0x_1, T_0x_2 are linearly independent, and then choose $A \in \mathcal{A}$ so that $AT_0x_1 = x_2$. (Such an operator A exists, otherwise $\{AT_0x_1 : A \in \mathcal{A}\}$ would be a subspace of $\mathcal{H}$, invariant for $\mathcal{A}$.) Then $T_0AT_0x_1$ $(= T_0x_2)$ and T_0x_1 are linearly independent, and $T_0AT_0 - \lambda T_0$ is not a zero transformation for any $\lambda \in \mathbb{C}$. On the other hand, there exists a complex number λ_0 such that the restriction of $T_0A - \lambda_0$ to $\operatorname{Ran} T_0$ is not invertible. Therefore, $T_0AT_0 - \lambda_0T_0$ has rank less than d and greater than 0, contradicitng the minimality of d. Hence $d = 1$.

If $T_0 = x \otimes y$, we will show that $\mathcal{A}$ contains all rank one operators. Let $u \otimes v$ be a rank one operator. Once again, there must be an operator $A_1 \in \mathcal{A}$ such that $A_1x = u$. Notice that the algebra $\mathcal{A}^* = \{A^* : A \in \mathcal{A}\}$ is also transitive. Therefore, there exists an operator $A_2 \in \mathcal{A}$ such that $A_2^*y = v$. Then $A_1T_0A_2 = u \otimes v$ so $\mathcal{A}$ contains all rank one operators and, hence, all finite rank operators, i.e. $\mathcal{L}(\mathcal{H})$. $\square$

EXERCISE 4.3.2. Prove that if $\mathcal{A}$ is a subalgebra of $\mathcal{L}(\mathcal{H})$ and $x \in \mathcal{H}$, then $\mathcal{A}x = \{Ax : A \in \mathcal{A}\}$ is a subspace of $\mathcal{H}$, invariant for $\mathcal{A}$.

EXERCISE 4.3.3. Prove that $\mathcal{A}$ is transitive iff $\mathcal{A}^*$ is transitive.

THEOREM 4.3.2 (Lomonosov's Theorem). *Let A be a non-scalar operator on Hilbert space that commutes with a compact operator. Then A has a nontrivial hyperinvariant subspace.*

The proof of this result uses a fixed point theorem.

THEOREM 4.3.3. *Let F be a compact and convex subset of Hilbert space $\mathcal{H}$, and let T be a linear operator in $\mathcal{L}(\mathcal{H})$ with the property that $T(F) \subset F$. Then there exists $p \in \mathcal{H}$ such that $Tp = p$.*

PROOF. For every $n \in \mathbb{N}$, let $T_n = (1 + T + T^2 + \cdots + T^{n-1})/n$. The set $T_n(F)$ is convex, (Exercise 4.3.4), and compact, as the image of a compact set under a continuous map. Also, $T_n(F) \subset F$, because if $x \in F$ then $T^k x \in F$, $0 \le k \le n-1$, and K is convex. Further, for any $m, n \in \mathbb{N}$, $T_m(F)T_n(F) \subset T_m(F) \cap T_n(F)$ which shows that the family $\{T_n(F)\}_{n \in \mathbb{N}}$ has a finite intersection property. Since they are all subset of a compact set F, they all have a non-empty intersection, i.e., there exists $p \in \cap\{T_n(F) : n \in \mathbb{N}\}$. We will show that $Tp = p$.

Suppose, to the contrary, that $Tp \neq p$. Then there exists $\alpha > 0$ such that $\|Tp - p\| \ge \alpha$. Since F is a bounded set, there exists $M > 0$ such that $\|x\| \le M$, for $x \in F$. Let n be a positive integer satisfying $n > 2M/\alpha$. Since $p \in T_n(F)$, there exists $x_n \in F$ such that $p = T_n x_n$ and, therefore,

$$Tp - p = (T-1)T_n x_n = (T-1)\,\frac{1 + T + T^2 + \cdots + T^{n-1}}{n}\, x_n = \frac{T^n - 1}{n}\, x_n.$$

Then $\alpha \le \|Tp - p\| = \|(T^n - 1)/n x_n\| \le (\|T^n x_n\| + \|x_n\|)/n \le 2M/n$ which contradicts the choice of n. □

EXERCISE 4.3.4. Prove that if C is a convex set in Hilbert space $\mathcal{H}$ and $T \in \mathcal{L}(\mathcal{H})$, then $T(C)$ is a convex set.

Now we can prove the result which is frequently referred to as the Lomonosov's Lemma.

THEOREM 4.3.4. *If $\mathcal{A}$ is a transitive subalgebra of $\mathcal{L}(\mathcal{H})$ and if K is a non-zero compact operator in $\mathcal{L}(\mathcal{H})$, then there exists an operator $A \in \mathcal{A}$ and a non-zero vector $x \in \mathcal{H}$ such that $AKx = x$.*

PROOF. Without loss of generality we will assume that $\|K\| = 1$. As we have already noticed, it suffices to consider the case when K is quasinilpotent. Let x_0 be a vector in $\mathcal{H}$ such that $\|Kx_0\| > 1$ and notice that this implies that $\|x_0\| > 1$, so the closed ball $B(x_0, 1)$ does not contain 0. Let D be the image under K of the closed ball $B(x_0, 1)$. By Exercise 2.6.4, D is a compact set. In addition, it is convex, by Exercise 4.3.4 and it does not contain 0. Indeed, for any $x \in B(x_0, 1)$, $\|Kx\| \ge \|Kx_0\| - \|K(x - x_0)\| > 1 - \|x - x_0\| \ge 0$.

For an operator $T \in \mathcal{A}$, consider the set $U_T = \{y \in \mathcal{H} : \|Ty - x_0\| < 1\}$. Notice that $U_T = T^{-1}(\{z : \|z - x_0\| < 1\}$ so it is an open set. Moreover, every non-zero vector y belongs to U_T, for some $T \in \mathcal{A}$. Indeed, $\mathcal{A}$ is transitive so the linear manifold $\{Ty : T \in \mathcal{A}\}$ must be dense in $\mathcal{H}$ and, hence, there exists $T \in \mathcal{A}$ such that $\|Ty - x_0\| < 1$, which means that $y \in U_T$. Thus, $\cup_{T \in \mathcal{A}} U_T$ is a covering of $\mathcal{H} - \{0\}$, and all the more of D. As established earlier, D is a compact set, so there exist operators $T_1, T_2, \dots, T_n \in \mathcal{A}$ such that $D \subset \cup_{i=1}^n U_{T_i}$. This means that, for any $y \in D$ there exists T_i, $1 \le i \le n$, such that $\|T_i y - x_0,\| < 1$.

Now, for each j, $1 \le j \le n$, and $y \in D$, we define $\alpha_j(y) = \max\{0,\, 1 - \|T_j y - x_0\|\}$. Notice that each α_j is continuous on D, $0 \le \alpha_j \le 1$, and $\sum_{j=1}^n \alpha_j(y) > 0$, for all $y \in D$. Furthermore, $\alpha_j(y) \ne 0$ iff $\|T_j y - x_0\| < 1$. Define, for $y \in D$ and $1 \le j \le n$,

$$\beta_j(y) = \frac{\alpha_j(y)}{\sum\limits_{i=1}^n \alpha_i(y)},$$

and notice that each β_j is continuous on D, $0 \le \beta_j \le 1$, and $\sum_{j=1}^n \beta_j(y) = 1$, for all $y \in D$. Also, $\beta_j(y) \ne 0$ iff $\alpha_j(y) \ne 0$ iff $\|T_j y - x_0\| < 1$. Finally, let $\Psi : D \to \mathcal{H}$ be defined by $\Psi(y) = \sum_{j=1}^n \beta_j(y) T_j y$. It is easy to see that Ψ is continuous on D. We will show that $\Psi(D) \subset B(x_0, 1)$. Let $y \in D$. Then

$$\|\Psi(y) - x_0\| = \|\sum_{j=1}^n \beta_j(y) T_j y - \sum_{j=1}^n \beta_j(y) x_0\| \le \sum_{j=1}^n |\beta_j(y)| \|T_j y - x_0\| \le 1$$

so $\Psi(y) \in B(x_0, 1)$ and $\Psi(D) \subset B(x_0, 1)$. If we define $\Phi : B(x_0, 1) \to \mathcal{H}$ by $\Phi(y) = \Psi(Ky)$, then Φ is a continuous map of $B(x_0, 1)$ into itself. Since $B(x_0, 1)$ is a compact, convex set, Theorem 4.3.3 shows that Φ has a fixed point $p \in B(x_0, 1)$, hence non-zero. Now we define the operator $A = \sum\limits_{j=1}^n \beta_j(Kp) T_j$ which is in $\mathcal{A}$. Finally, $AKp = \sum\limits_{j=1}^n \beta_j(Kp) T_j Kp = \Psi(Kp) = \Phi(p) = p$. □

Now we can prove Lomonosov's theorem.

Proof of Lomonosov's Theorem. Let $\mathcal{A} = \{A\}'$ and suppose, to the contrary, that $\mathcal{A}$ is transitive. By Theorem 4.3.4, there exists an operator $T \in \{A\}'$ such that $TKx = x$. In other words, a compact operators AK has 1 as an eigenvalue. Let $E(1)$ denote the appropriate eigenspace which is finite dimensional. Since A commutes with TK, the subspace $E(1)$ is invariant for A as well. The restriction of A to $E(1)$ must have an eigenvalue λ and, since $E(1)$ is invariant for A, we see that λ is an eigenvalue for A (not just the restriction). Let $\mathcal{M}$ denote the eigenspace of A corresponding to λ, i.e., $\mathcal{M} = \{x \in \mathcal{H} : Ax = \lambda x\}$. Being an eigenspace, it is hyperinvariant for A. It is not (0), so it remains to notice that it is not $\mathcal{H}$ because $A \ne \lambda$. □

4.4. Normal operators

We have seen in Exercise 2.7.1 that a multiplication operator M_h on L^2 is a normal operator. In this section we will show that, in a sense, every normal operator is a multiplication by an essentially bounded function.

Example 4.4.1. Let $T = \left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$, with $a, b \in \mathbb{C}$. Then $TT^* = T^*T$. Let $X = \{1, 2\}$ and let μ be a counting measure on X. Notice that $L^2(X, \mu)$ is the collection of all functions $f : X \to \mathbb{C}$ with norm $\left(\int_X |f|^2 \, d\mu\right)^{1/2} = \left(|f(1)|^2 + |f(2)|^2\right)^{1/2}$. Since this is the Euclidean norm, we see that $L^2(X, \mu)$ is just $\mathcal{L}(\mathbb{C}^2)$. Finally, let h be a function on X, $h(1) = a$, $h(2) = b$. Then T can be identified with M_h on $\mathcal{L}(\mathbb{C}^2)$.

Remark 4.4.1. A similar construction can be made for the case when T is an $n \times n$ diagonal matrix, $T = \operatorname{diag}(c_n)$.

Example 4.4.2. Let $T = \operatorname{diag}(c_n)$, with $c_n \in \mathbb{C}$ for all $n \in \mathbb{N}$. Let $X = \mathbb{N}$ and $\mu(\{n\}) = 1/2^n$. Then (X, μ) is a *finite* measure space. Further, let $h : X \to \mathbb{C}$ be defined by $h(n) = c_n$. Then T can be identified with the operator M_h on $L^2(X, \mu)$.

The last example shows the danger of going through the motions. What does it mean "can be identified"? While it is easy to see that $Tf = M_h f$ for any sequence f, their domains are not the same. Namely, T acts on ℓ^2 but M_h acts on $L^2(X, \mu)$, and these 2 spaces are not the same. For example, the sequence $(1, 1, 1, \dots)$ belongs to $L^2(X, \mu)$ but not to ℓ^2. However, these two spaces are isomorphic. Let $U : L^2(X, \mu) \to \ell^2$ be defined by $U(f) = (f(1)/\sqrt{2}, f(2)/\sqrt{2^2}, f(3)/\sqrt{2^3}, \dots)$. It is easy to verify that U is injective and surjective linear map so, by the Open Mapping Principle, it is an isomorphism. Moreover, if $f \in L^2(X, \mu)$, then

$$U^{-1}TU(f) = U^{-1}T\Big(\frac{f(1)}{\sqrt{2}}, \frac{f(2)}{\sqrt{2^2}}, \frac{f(3)}{\sqrt{2^3}}, \dots\Big) = U^{-1}\Big(\frac{c_1 f(1)}{\sqrt{2}}, \frac{c_2 f(2)}{\sqrt{2^2}}, \frac{c_3 f(3)}{\sqrt{2^3}}, \dots\Big)$$
$$= U^{-1}\Big(\frac{h(1)f(1)}{\sqrt{2}}, \frac{h(2)f(2)}{\sqrt{2^2}}, \frac{h(3)f(3)}{\sqrt{2^3}}, \dots\Big) = hf,$$

so T is unitarily equivalent to M_h.

EXERCISE 4.4.1. Prove that the map $U : L^2(X, \mu) \to \ell^2$, constructed in Example 4.4.2, is an isometric isomorphism.

Notice that in Examples 4.4.1 and 4.4.2 the measure was defined on each of the pieces. What happens if pieces are not that obvious? How do we define a piece?

DEFINITION 4.4.1. A vector ξ is *cyclic* for an operator T if the set $\{p(T)\xi : p \text{ is a polynomial}\}$ is dense in $\mathcal{H}$. An operator T is *cyclic* if it has a cyclic vector.

EXAMPLE 4.4.3. Let $T = S$, the unilateral shift. The vector $\xi = e_1$ is cyclic for S. If $x \in \ell^2$, $x = (x_1, x_2, \dots)$ then x can be approximated by truncated sequences $(x_1, x_2, \dots, x_n, 0, 0, \dots) = \sum_{k=1}^n x_k e_k = \sum_{k=1}^n T^k e_1$.

EXAMPLE 4.4.4. Let $\{\dots, e_{-2}, e_{-1}, e_0, e_1, e_2, \dots\}$ be an o.n.b. of $\mathcal{H}$, and let T be the *bilateral shift*: $Te_n = e_{n+1}$, $n \in \mathbb{Z}$. Then $\xi = e_0$ is not a cyclic vector for T, because $\{p(T)e_0\}^- = \vee_{k=0}^\infty e_k$. However, $T^*e_n = e_{n-1}$, $n \in \mathbb{Z}$, so we need to replace polynomials in T by polynomials in T and T^*, i.e., $f(T) = \sum_{i,j=1}^n T^i T^{*j}$. If the set $\{f(T)\xi : f \text{ is a polynomial in } T, T^*\}$ is dense in $\mathcal{H}$, we say that e_0 is a *star-cyclic vector* for T.

Before we proceed, we revisit the Stone–Weierstrass Theorem [Bartle, p. 184]. Although it is proved under the assumption that K is a compact subset of $\mathbb{R}^p$, the same proof is valid when K is a compact set in $\mathbb{C}$. Also, we will rephrase it using the following terinology. We will say that an algebra $\mathcal{A}$ of functions *separates points on* K if, for any two distinct points $x, y \in K$ there is a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. If for each $x \in K$ there is a function $g \in \mathcal{A}$ such that $g(x) \neq 0$, we say that $\mathcal{A}$ *vanishes at no point of* K.

THEOREM 4.4.1 (Stone–Weierstrass Theorem). *Let $\mathcal{A}$ be an algebra of continuous, real-valued functions on a compact set K in $\mathbb{C}$. If $\mathcal{A}$ separates points on K and if $\mathcal{A}$ vanishes at no point of K, then the uniform closure of $\mathcal{B}$ of $\mathcal{A}$ consists of all real-valued continuous functions on K.*

The Stone–Weierstrass Theorem deals only with real-valued functions of complex variable. Now we extend it to complex-valued functions. We will require that $\mathcal{A}$ be *self-adjoint*, meaning that if $f \in \mathcal{A}$ the $\overline{f} \in \mathcal{A}$.

Theorem 4.4.2. *Let $\mathcal{A}$ be a self-adjoint algebra of continuous, complex functions on a compact set K in $\mathbb{C}$. If $\mathcal{A}$ separates points on K and if $\mathcal{A}$ vanishes at no point of K, then the uniform closure of $\mathcal{B}$ of $\mathcal{A}$ consists of all complex continuous functions on K.*

Proof. Let $f = u+iv$ be a continuous function on K, and let $\mathcal{A}_{\mathbb{R}}$ denote the set of all real-valued functions in $\mathcal{A}$. Since u, v are continuous real-valued continuous function on K, it suffice to show that every such function lies in the closure of $\mathcal{A}_{\mathbb{R}}$. Since $\mathcal{A}_{\mathbb{R}}$ is clearly an algebra, the result will follow from the Stone–Weierstrass Theorem, once we show that $\mathcal{A}_{\mathbb{R}}$ separates points on K and vanishes at no point of K.

Suppose that z_1, z_2 are distinct points in K. By assumption, $\mathcal{A}$ separates points on K so it contains a function f such that $f(z_1) \neq f(z_2)$. Also, $\mathcal{A}$ vanishes at no point of K, so it contains two functions g, h such that $g(z_1) \neq 0$, $h(z_2) \neq 0$. Then, the function

$$F(z) = \frac{f(z)g(z) - f(z_2)g(z)}{f(z_1)g(z_1) - f(z_2)g(z_1)}$$

belongs to $\mathcal{A}$ and has the property that $F(z_1) = 1$, $F(z_2) = 0$. Notice that, if $F = u + iv \in \mathcal{A}$, then $\overline{F} \in \mathcal{A}$ and $u = (F + \overline{F})/2 \in \mathcal{A}_{\mathbb{R}}$. Clearly, $u(z_1) = 1$, $u(z_2) = 0$ so $\mathcal{A}_{\mathbb{R}}$ separates points on K.

Let $z_0 \in K$. Then there exists a function $G \in \mathcal{A}$ such that $G(z_0) \neq 0$. Let λ be a complex number such that $\lambda G(z_0) > 0$ and notice that $H = \operatorname{Re}(\lambda G)$ is a function in $\mathcal{A}_{\mathbb{R}}$ such that $H(z_0) > 0$. Thus, $\mathcal{A}_{\mathbb{R}}$ vanishes at no point of K and the proof is complete. $\square$

Now we are ready to establish a stronger connection between normal operators and operators of multiplication.

Theorem 4.4.3. *Let T be a normal operator in $\mathcal{L}(\mathcal{H})$ with a star-cyclic vector ξ. Then there exists a finite measure μ on $\sigma(T)$, a bounded function $h : \sigma(T) \to \mathbb{R}$, and an isomporphism $U : L^2(\sigma(T), \mu) \to \mathcal{H}$ such that $U^{-1}TUf(x) = h(x)f(x)$ for a.e. $x \in \sigma(T)$ and all $f \in L^2(\sigma(T), \mu)$.*

Proof. Let $\mathcal{A}$ be the algebra of complex-valued polynomials in $z, \overline{z}$. For $f \in \mathcal{A}$ we define $L(f) = \langle f(T)\xi, \xi\rangle$. Clearly, L is a linear functional and it is bounded on $\mathcal{A}$. Indeed, $|L(f)| = |\langle f(T)\xi, \xi\rangle| \leq \|f(T)\xi\|\|\xi\| \leq \|f(T)\|\|\xi\|^2$. Further, T is normal, so $f(T)$ is also normal and, by Corollary 3.5.5, $\|f(T)\| = r(f(T)) = \sup\{|\lambda| : \lambda \in \sigma(f(T))\}$.

Finally, by the Spectral Mapping Theorem, $\lambda \in \sigma(f(T))$ iff $\lambda = f(\mu)$, for some $\mu \in \sigma(T)$. Thus, $\|f(T)\| = \sup\{|f(\mu)| : \mu \in \sigma(T)\} = \|f\|_\infty$. We conclude that $|L(f)| \leq \|f\|_\infty \|\xi\|^2$, so L is bounded on $\mathcal{A}$. By Theorem 4.4.2, $\mathcal{A}$ is dense in $C(\sigma(T))$ so we can extend L to a bounded linear functional on $C(\sigma(T))$. If f is a non-negative function in $C(\sigma(T))$, then so is $\sqrt{f}$ and it can be approximated by a sequence $f_n \in \mathcal{A}$. It follows that f can be approximated by the sequence $\overline{f_n} f_n$ and, by the continuity of L, $L(f)f = \lim L(\overline{f_n} f_n) = \langle \overline{f_n}(T) f_n(T)\xi, \xi\rangle = \|f_n(T)\xi\|^2 \geq 0$. Thus, L is positive, and by Riesz Representation Theorem [Royden, p. 352] there exists a finite positive measure μ on $\sigma(T)$ such that $\langle f(T)\xi, \xi\rangle = \int f\, d\mu$. Now define the operator U on $\mathcal{A}$ by $U(f) = f(T)\xi$. Since $|f|^2 = \overline{f} f$ we have that $\int |f|^2\, d\mu = \langle \overline{f}(T) f(T)\xi, \xi\rangle = \|f(T)\xi\|^2 = \|U(f)\|^2$. That way, U is an isometry on $\mathcal{A}$. Further, $\mathcal{A}$ is dense in $L^2(\mu)$ because it is dense in $C(\sigma(T))$, and the latter set is dense in L^2 ([Rudin, Theorem 3.14]). Therefore, by Theorem 2.3.4, U can be extended to an isometry $U : L^2(\sigma(T), \mu) \to \mathcal{H}$. Since ξ is star-cyclic, the set $\{f(T)\xi : f \in \mathcal{A}\}$ is dense in $\mathcal{H}$ so the range of U is dense. Since U is bounded below its range is closed so U is surjective.

Finally, if we denote by $\tilde{f}(z)$ the function $zf(z)$, then $U^{-1}TU(f) = U^{-1}Tf(T)\xi = U^{-1}\tilde{f}(T)\xi = \tilde{f}$ so T can be identified with M_z on $L^2(\sigma(T), \mu)$. $\square$

What if T does not have a star-cyclic vector?

THEOREM 4.4.4. *Let T be a normal operator in $\mathcal{L}(\mathcal{H})$. Then there exists a compact set X, a finite measure μ on X, a bounded function $h : X \to \mathbb{R}$, and an isomporphism $U : L^2(X, \mu) \to \mathcal{H}$ such that $U^{-1}TUf(x) = h(x)f(x)$ for a.e. $x \in X$ and all $f \in L^2(X, \mu)$.*

PROOF. Let x_1 be a non-zero vector and let $\mathcal{M}_1$ be the closed linear span of $\{f(T)x_1 : f \in \mathcal{A}\}$. If $\mathcal{M}_1 = \mathcal{H}$ then x_1 is a star-cyclic vector for T and Theorem 4.4.3 applies. If $\mathcal{M}_1 \neq \mathcal{H}$ there exists a non-zero vector $x_2 \in \mathcal{M}_1^\perp$. Notice that $\mathcal{M}_1$ is invariant (hence reducing) for T and T^*, so the same is true of $\mathcal{M}_1^\perp$. Now, either the closed linear span of $\{f(T)x_2 : f \in \mathcal{A}\}$ equals $\mathcal{M}_1^\perp$, in which case $T = T_1 \oplus T_2$ and both T_1 and T_2 are star-cyclic, or we continue the process. Applying the Hausdorff Maximal Principle, we obtain a decomposition of $\mathcal{H}$ relative to which $T = \operatorname{diag}(T_i)$ and each of the operators on the diagonal is star-cyclic. By Theorem 4.4.3, for each i there exists a finite measure space (X_i, μ_i), a function $h_i \in L^2(X_i, \mu_i)$, and unitary operator $U_i : L^2(X_i, \mu_i) \to \mathcal{M}_i$, such that $U_i^{-1} T_i U_i = M_{h_i}$. Next we define X to be the union of X_i and μ a measure on X so that $\mu = \mu_i$ on X_i. Finally, we define a function h so that $h = h_i$ on X_i and a unitary operator $U = \operatorname{diag}(U_i)$. Then T can be identified with M_h on $L^2(X, \mu)$, i.e., $U^{-1} T U = M_h$. □

We will now introduce a very important concept.

Definition 4.4.2. If X is a set, Ω a σ-algebra of subsets of X, and $\mathcal{H}$ is Hilbert space, a *spectral measure* for $(X, \Omega, \mathcal{H})$ is a function $E : \Omega \to \mathcal{L}(\mathcal{H})$ such that

(a) for each Δ in Ω, $E(\Delta)$ is a projection;

(b) $E(\emptyset) = 0$ and $E(X) = 1$;

(c) $E(\Delta_1 \cap \Delta_2) = E(\Delta_1) E(\Delta_2)$;

(d) if $\{\Delta_i\}_{i \in I}$ are pairwise disjoint sets in Ω, then $E(\cup_{i \in I} \Delta_i) = \sum_{i \in I} E(\Delta_i)$.

Example 4.4.5. Let $X = \mathbb{N}$, let Ω be the set of all subsets of $\mathbb{N}$, and let $\{e_n\}_{n \in \mathbb{N}}$ be an o.n.b. of $\mathcal{H}$. For $\Delta \subset \mathbb{N}$, define $E(\Delta)$ to be the projection onto the span $\vee_{n \in \Delta} e_n$. Properties (a) and (b) of Definition 4.4.2 are obvious. Since $E(\Delta) e_i$ is either e_i or 0, depending on whether i belongs to Δ or not, we see that $E(\Delta_1) E(\Delta_2) e_i = 0$ unless $i \in \Delta_1 \cap \Delta_2$, in which case it equals e_i. Thus, for $x = \sum x_i e_i$, $E(\Delta_1) E(\Delta_2) x = \sum_{i \in \Delta_1 \cap \Delta_2} x_i e_i = E(\Delta_1 \cap \Delta_2) x$, and (c) holds as well. Finally, if $\{\Delta_i\}_{i \in I}$ are pairwise disjoint sets in Ω, and $\Delta = \cup_{i \in I} \Delta_i$, writing $x = \sum_{n \in \mathbb{N}} x_n e_n$, we have that $E(\Delta) x = \sum_{i \in \Delta_1} x_i e_i + \sum_{i \in \Delta_2} x_i e_i + \cdots = E(\Delta_1) x + E(\Delta_2) x + \dots$.

Example 4.4.6. If X is a set, Ω a σ-algebra of subsets of X, and μ a measure on Ω, let $\mathcal{H} = L^2(X, \mu)$, and define, for $\Delta \in \Omega$ and $f \in L^2$, $E(\Delta) f = \chi_\Delta f$. Then, E is a spectral measure.

Exercise 4.4.2. Verify that E in Example 4.4.6 is a spectral measure.

We will now show that the equality $U^{-1} T U = M_h$, established in Theorem 4.4.3, can be extended in the following manner. Suppose that F is a bounded function on $\sigma(T)$. Then we can define $F(T) = U M_{F \circ h} U^{-1}$ since, for $x \in \mathcal{H}$, $U^{-1} x \in L^2$ and $M_{F \circ h} U^{-1} x$ is also in L^2, so $U M_{F \circ h} U^{-1} x$ is well defined.

THEOREM 4.4.5. *Let T be a bounded linear operator on Hilbert space $\mathcal{H}$. The mapping $F \mapsto F(T)$ is an algebra homomorphism from $L^\infty(\sigma(T), \mu)$ to $\mathcal{L}(\mathcal{H})$.*

EXERCISE 4.4.3. Prove Theorem 4.4.5.

REMARK 4.4.2. The homomorphism $F \mapsto F(T)$ is called a *functional calculus.*

Example 4.4.6 shows that a spectral measure can be defined using multiplication by characteristic functions. We present a variation on this theme.

THEOREM 4.4.6. *If T is a normal operator on Hilbert space, Δ is a measurable subset of $\sigma(T)$, and $F = \chi_\Delta$, then the mapping E defined by $E(\Delta) = F(T) = UM_{F\circ h}U^{-1}$ is a spectral measure.*

EXERCISE 4.4.4. Prove Theorem 4.4.6.

EXERCISE 4.4.5. What is E when $T = \text{diag}(c_n)$?

Let $x, y \in \mathcal{H}$ and denote by $f = U^{-1}x$ and $g = U^{-1}y$. Since U is a surjective isometry, $U^{-1} = U^*$ so $f = U^*x$ and $g = U^*y$. If $F = \chi_\Delta$ then, by definition, $\langle E(\Delta)x, y\rangle = \langle F(T)x, y\rangle = \langle UM_{F\circ h}U^{-1}x, y\rangle = \langle M_{F\circ h}f, g\rangle = \int F \circ h\, f\overline{g}\, d\mu$. On the other hand, E is the spectral measure of T, so $\langle E(\Delta)x, y\rangle$ also defines a measure $\nu(\Delta)$. It is often called the *scalar spectral measure* of T.

EXERCISE 4.4.6. Verify that ν is a measure.

Now, $\langle E(\Delta)x, y\rangle$ is equal to $\int \chi_\Delta\, d\nu$ as well as to $\int F \circ h f\overline{g}\, d\mu$, so we have the equality

$$\int F \circ h f\overline{g}\, d\mu = \int F\, d\nu \tag{4.3}$$

whenever F is a characteristic function. Since every simple function is a linear combination of characteristic functions, it is not hard to see that (4.3) remains true when F is a simple function. Further, every bounded function can be approximated by simple functions so, by relying on Lebesgue Dominated Convergence Theorem, we obtain that (4.3) holds for any bounded function F. In particular, if $F(\lambda) = \lambda$, we obtain that $\langle Tx, y\rangle = \int \lambda\, d\nu = \int \lambda\, d\langle E(\lambda)x, y\rangle$. Since this is true for all $x, y \in \mathcal{H}$, we can write $T = \int \lambda\, dE(\lambda)$ or

$$T = \int \lambda\, dE. \tag{4.4}$$

More generally, since (4.3) holds for any bounded function F, it follows that, for any such function,

$$F(T) = \int F(\lambda)\, dE. \tag{4.5}$$

Theorem 4.4.6 established that to every normal operator there corresponds a spectral measure. The following result shows how essential this measure is for the operator.

THEOREM 4.4.7. *If T is a normal operator and E the associated spectral measure, then an operator A commutes with T iff A commutes with $E(\Delta)$ for every Borel set $\Delta \subset \sigma(T)$.*

PROOF. Let $x, y \in \mathcal{H}$, and let F be a bounded function on $\sigma(T)$. Then

$$\langle AF(T)x, y\rangle = \langle F(T)x, A^*y\rangle = \int F(\lambda)\, d\langle E(\lambda)x, A^*y\rangle, \quad \text{and}$$
$$\langle F(T)Ax, y\rangle = \int F(\lambda)\, d\langle E(\lambda)Ax, y\rangle.$$

If A and T commute, Fuglede–Putnam Theorem implies that A commutes with T^*, hence with $F(T)$, for any bounded function F. In particular, by taking $F = \chi_\Delta$, we obtain that $\langle E(\Delta)x, A^*y\rangle = \langle E(\Delta)Ax, y\rangle$ or, equivalently that $\langle AE(\Delta)x, y\rangle = \langle E(\Delta)Ax, y\rangle$. Since this holds for all $x, y \in \mathcal{H}$ it follows that A commutes with $E(\Delta)$.

Conversely, if A commutes with $E(\Delta)$, then $\langle E(\Delta)x, A^*y\rangle = \langle AE(\Delta)x, y\rangle = \langle E(\Delta)Ax, y\rangle$. Since $\langle ATx, y\rangle = \int \lambda\, d\langle E(\lambda)x, A^*y\rangle$ and $\langle TAx, y\rangle = \int \lambda\, d\langle E(\lambda)Ax, y\rangle$, we obtain that $\langle ATx, y\rangle = \langle TAx, y\rangle$ for all $x, y \in \mathcal{H}$. Thus $AT = TA$ and the proof is complete. $\square$

Theorem 4.4.7 has an important consequence that concerns the existence of hyperinvariant subspaces.

Corollary 4.4.8. *If T is a normal operator in $\mathcal{L}(\mathcal{H})$, and E is its spectral measure, then $E(\Delta)$ is a hyperinvariant subspace for T, for any Borel set $\Delta \subset \sigma(T)$. Consequently, if T is not a scalar multiple of the identity, then T has a non-trivial hyperinvariant subspace.*

Exercise 4.4.7. Prove Corollary 4.4.8.

CHAPTER 5

Spectral radius algebras

5.1. Compact operators

In Section 4.3 we have shown that every compact operator is contained in an algebra, namely its commutant, that is not transitive. Are there other algebras that would contain a given operator and still have an invariant subspace? We will show that the answer is affirmative. Let us denote the class of quasinilpotent operators as $\mathcal{Q}$. The following is a direct consequence of Theorem 4.3.4.

PROPOSITION 5.1.1. *Let $\mathcal{A}$ be a unital subalgebra of $\mathcal{L}(\mathcal{H})$ and let K be a compact operator in $\mathcal{L}(\mathcal{H})$. If $AK \in \mathcal{Q}$ for each $A \in \mathcal{A}$, then $\mathcal{A}$ has a n. i. s.*

PROOF. If $\mathcal{A}$ is transitive, by Theorem 4.3.4 there exists $A \in \mathcal{A}$ such that $1 \in \sigma_p(AK)$, so $AK \notin \mathcal{Q}$. $\square$

Our goal is to find an algebra $\mathcal{A}$ with the property stated in Proposition 5.1.1. Let $A \in \mathcal{L}(\mathcal{H})$. For $m \in \mathbb{N}$, define

$$d_m = \frac{m}{1 + mr(A)}, \quad \text{and } R_m = \left(\sum_{n=0}^{\infty} d_m^{2n} A^{*n} A^n\right)^{1/2}. \tag{5.1}$$

EXERCISE 5.1.1. Prove that the series in (5.1) converges uniformly and, for each $m \in \mathbb{N}$, R_m is invertible with $||R_m^{-1}|| \leq 1$.

If A is an operator in $\mathcal{L}(\mathcal{H})$ and R_m is as in (5.1), we associate with A the collection

$$\mathcal{B}_A = \left\{ T \in \mathcal{L}(\mathcal{H}) : \sup_m ||R_m A R_m^{-1}|| < \infty \right\}.$$

EXERCISE 5.1.2. Show that $\mathcal{B}_A$ is an algebra.

We will show that $\mathcal{B}_A$ contains all operators that commute with A. In fact, we can prove a stronger result.

Proposition 5.1.2. *Suppose A is a nonzero operator, B is a power bounded operator commuting with A, and T is an operator for which $AT = BTA$. Then $T \in \mathcal{B}_A$.*

An operator T is *power bounded* if there exists $C > 0$ such that $\|T^n\| \leq C$, for all $n \in \mathbb{N}$. For example, if $\|T\| \leq 1$, then T is power bounded.

Proof. It is easy to verify that $A^2T = B^2TA^2$. Using induction one can prove that $A^nT = B^nTA^n$, for every $n \in \mathbb{N}$. The operator B is power bounded so there is a constant C such that $\|B^n\| \leq C$, for each $n \in \mathbb{N}$. For any vector $x \in \mathcal{H}$ and any positive integer m, we have that

$$\|R_m x\|^2 = \langle R_m x, R_m x\rangle = \langle R_m^2 x, x\rangle = \sum_{n=0}^{\infty} d_m^{2n} \langle A^{*n} A^n x, x\rangle = \sum_{n=0}^{\infty} d_m^{2n} \langle A^n x, A^n x\rangle = \sum_{n=0}^{\infty} d_m^{2n} \|A^n x\|^2. \tag{5.2}$$

On the other hand, $\|A^n T R_m^{-1} x\| = \|B^n T A^n R_m^{-1} x\| \leq C\|T\|\|A^n R_m^{-1} x\|$ so we obtain that

$$\begin{aligned}
\|R_m T R_m^{-1} x\|^2 &= \sum_{n=0}^{\infty} d_m^{2n} \, \|A^n T R_m^{-1} x\|^2 \\
&\leq C^2 \|T\|^2 \sum_{n=0}^{\infty} d_m^{2n} \, \|A^n R_m^{-1} x\|^2 \\
&= C^2 \|T\|^2 \|R_m R_m^{-1} x\|^2 \\
&= C^2 \|T\|^2 \|x\|^2.
\end{aligned}$$

Thus $T \in \mathcal{B}_A$. □

From this we deduce an easy consequence.

Corollary 5.1.3. *Let T be an operator such that $AT = \lambda TA$ for some complex number λ with $|\lambda| \leq 1$. Then $T \in \mathcal{B}_A$. In particular $\mathcal{B}_A$ contains the commutant of A.*

Example 5.1.1. If u and v are unit vectors then $\mathcal{B}_{u \otimes v} = \{T \in \mathcal{L}(\mathcal{H}) : v \text{ is an eigenvector for } T^*\}$. Let $A = u \otimes v$ be a rank one operator, with u and v are unit vectors. One knows that $r(u \otimes v) = |\langle u, v\rangle|$. A calculation shows that, for $n \in \mathbb{N}$, $A^n = \langle u, v\rangle^{n-1}\, u \otimes v$ and $A^{*n}A^n = r^{2n-2}\, v \otimes v$. Therefore,

$$R_m^2 = I + \left(\sum_{n=1}^{\infty} d_m^{2n}\, r^{2n-2}\right) v \otimes v = I + \frac{d_m^2}{1 - d_m^2 r^2}\, v \otimes v.$$

Let $\lambda_m = \sqrt{1 + d_m^2/(1 - d_m^2 r^2)}$ for every $m \in \mathbb{N}$. Notice that $\lambda_m \to \infty$ as $m \to \infty$. Indeed, either $d_m \to 1/r$ or, if A is quasinilpotent, $\lambda_m = \sqrt{1 + m^2}$. If we denote by $\mathcal{M}$ the one dimensional space spanned by v then, relative to $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$, the matrix of R_m is $R_m = \left[\begin{smallmatrix} \lambda_m & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $R_m^{-1} = \left[\begin{smallmatrix} 1/\lambda_m & 0 \\ 0 & 1 \end{smallmatrix}\right]$. If T is an arbitrary operator, say $T = \left[\begin{smallmatrix} X & Y \\ Z & W \end{smallmatrix}\right]$, then

$$R_m T R_m^{-1} = \begin{bmatrix} \lambda_m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} \begin{bmatrix} 1/\lambda_m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X & Y\lambda_m \\ Z/\lambda_m & W \end{bmatrix}$$

and it is easy to see that $\sup_m \|R_m T R_m^{-1}\| < \infty$ if and only if $Y = 0$. This means that $\mathcal{M}^{\perp}$ is invariant for T or, equivalently, that $\mathcal{M}$ is invariant for T^*, and this is true iff v is an eigenvector for T^*.

Exercise 5.1.3. Prove that $r(u \otimes v) = |\langle u, v\rangle|$.

Now we define $\mathcal{Q}_A = \{T \in \mathcal{L}(\mathcal{H}) : \|R_m T R_m^{-1}\| \to 0\}$.

Theorem 5.1.4. *$\mathcal{Q}_A$ is a two sided ideal in $\mathcal{B}_A$ and every operator in $\mathcal{Q}_A$ is quasinilpotent. Furthermore, if A is quasinilpotent, then $A \in \mathcal{Q}_A$.*

Proof. Let $T \in \mathcal{Q}_A$ and let $X \in \mathcal{B}_A$. Then $\|R_m T X R_m^{-1}\| \le \|R_m T R_m^{-1}\|\, \|R_m X R_m^{-1}\| \to 0$ so $\mathcal{Q}_A$ is a right ideal. Since the same estimate holds for XT we see that $\mathcal{Q}_A$ is a two sided ideal in $\mathcal{B}_A$. On the other hand $r(T) = r(R_m T R_m^{-1}) \le \|R_m T R_m^{-1}\|$ which shows that if $T \in \mathcal{Q}_A$ then it must be quasinilpotent. Finally, if $A \in \mathcal{Q}$ then $r(A) = 0$ and $d_m = m$. Using (5.2) we see that

$$\begin{aligned} \|R_m A R_m^{-1} x\|^2 &= \sum_{n=0}^{\infty} m^{2n} \|A^{n+1} R_m^{-1} x\|^2 = \frac{1}{m^2} \sum_{n=0}^{\infty} m^{2n+2} \|A^{n+1} R_m^{-1} x\|^2 \\ &= \frac{1}{m^2}\left[-\|R_m^{-1} x\|^2 + \sum_{n=0}^{\infty} m^{2n} \|A^n R_m^{-1} x\|^2\right] = \frac{1}{m^2}\left[\|x\|^2 - \|R_m^{-1} x\|^2\right] \le \frac{\|x\|^2}{m^2} \end{aligned}$$

from which it follows that $\|R_m A R_m^{-1}\| \le 1/m \to 0$, $m \to \infty$. ∎

REMARK 5.1.1. The ideal $\mathcal{Q}_A$ need not contain every quasinilpotent operator in $\mathcal{B}_A$. Indeed, if A is the unilateral forward shift a calculation shows that $R_m^2 = 1/(1-d_m^2)$. Since every operator commutes with a scalar multiple of the identity it follows that $\mathcal{B}_A = \mathcal{L}(\mathcal{H})$. On the other hand, $\|R_m T R_m^{-1}\| = \|T\|$ for any T in $\mathcal{L}(\mathcal{H})$, so $\mathcal{Q}_A = (0)$.

The following result justifies our interest in $\mathcal{Q}_A$.

THEOREM 5.1.5. *If $\mathcal{Q}_A \neq (0)$ and there exists a nonzero compact operator in $\mathcal{B}_A$, then $\mathcal{B}_A$ has a n. i. s.*

PROOF. Let K be a nonzero compact operator in $\mathcal{B}_A$. Without loss of generality we may assume that $QK = 0$ for every $Q \in \mathcal{Q}_A$. Indeed, if $QK \neq 0$ for some $Q \in \mathcal{Q}_A$, then QK is a compact quasinilpotent operator with the property that $\mathcal{B}_A QK \subset \mathcal{Q}$ and the result follows from Proposition 5.1.1.

Let Q be a fixed nonzero operator in $\mathcal{Q}_A$ and let T be an arbitrary operator in $\mathcal{B}_A$. Then $QT \in \mathcal{Q}_A$ and, hence, $QTK = 0$. Since $K \neq 0$ there is a nonzero vector z in the range of K. Clearly, $QTz = 0$ so $Tz \in \ker Q$ for all $T \in \mathcal{B}_A$. Naturally, the closure of the subspace $\{Tz : T \in \mathcal{B}_A\}$ is an invariant subspace for $\mathcal{B}_A$. It is nonzero since $z \neq 0$ and the identity operator is in $\mathcal{B}_A$. Finally, it is not $\mathcal{H}$ since it is contained in the kernel of a nonzero operator Q. □

From Theorem 5.1.5 we deduce some easy consequences.

COROLLARY 5.1.6. *Suppose that A is a quasinilpotent operator, B is a power bounded operator commuting with A, and K is a nonzero compact operator satisfying $AK = BKA$. Then $\mathcal{B}_A$ has a n. i. s.*

PROOF. By Proposition 5.1.2, K is in $\mathcal{B}_A$. Since $A \in \mathcal{Q}$, Theorem 5.1.4 shows that $A \in \mathcal{Q}_A$. The result then follows from Theorem 5.1.5. □

COROLLARY 5.1.7. *Suppose that A is a quasinilpotent operator, λ is a complex number, and K is a nonzero compact operator satisfying $AK = \lambda KA$. Then either $\mathcal{B}_A$ or $\mathcal{B}_{A^*}$ has a n. i. s. In any case, A has a proper hyperinvariant subspace.*

Proof. If $|\lambda| \le 1$ Corollary 5.1.6 implies that $\mathcal{B}_A$ has a n. i. s. For $|\lambda| > 1$, we have $A^*K^* = (1/\overline{\lambda})K^*A^*$ so the same argument shows that $\mathcal{B}_{A^*}$ has a n. i. s. If $\mathcal{M}$ is such a subspace then it is hyperinvariant for A^*. It follows that $\mathcal{M}^\perp$ is a proper hyperinvariant subspace for A. □

Now we arrive to the main result of this section.

Theorem 5.1.8. *Let K be a nonzero compact operator on the separable, infinite dimensional Hilbert space $\mathcal{H}$. Then $\mathcal{B}_K$ has a n. i. s.*

Proof. We will show that $\mathcal{Q}_K \neq (0)$. The result will then follow from Theorem 5.1.5. Of course, if K is quasinilpotent, Theorem 5.1.4 shows that $K \in \mathcal{Q}_K$. Therefore, for the rest of the proof, we will assume that $r(K) > 0$.

Notice that $x \otimes y \in \mathcal{Q}_A$ iff $\|R_m(x \otimes y)R_m^{-1}\| \to 0$. However, $\|R_m(x \otimes y)R_m^{-1}\| = \|R_m x\|\|R_m^{-1}y\|$ so it suffices to exhibit a rank one operator $x \otimes y$ with $\sup_m \|R_m x\| < \infty$ and $\lim_m \|R_m^{-1}y\| = 0$. A vector y with the desired property is supplied by the following lemma.

Lemma 5.1.9. *Suppose that K is a compact operator and $r(K) > 0$. Then there exists a unit vector v such that $\left\|R_m^{-1}v\right\| \to 0$, $m \to \infty$.*

Proof. Let λ be a complex number in $\sigma(K)$ such that $|\lambda| = r(K)$. Then $\overline{\lambda} \in \sigma(K^*)$ so there are unit vectors u and v for which $Ku = \lambda u$ and $K^*v = \overline{\lambda}v$. An easy calculation shows that $K\,(u \otimes v) = (u \otimes v)\,K$ so that $u \otimes v \in \{K\}' \subset \mathcal{B}_A$. It then follows that $\sup_m ||R_m u||\,||R_m^{-1}v|| < \infty$. On the other hand, a straightforward calculation shows that $||R_m u|| \to \infty$, $m \to \infty$. Since $\sup_m ||R_m u||\,||R_m^{-1}v|| < \infty$ it must follow that $||R_m^{-1}v|| \to 0$. □

Exercise 5.1.4. Prove that $\|R_m u\| \to \infty$, $m \to \infty$.

So it remains to provide a nonzero vector x with the property that

$$\sup_m \|R_m x\| < \infty. \tag{5.3}$$

To that end, it suffices for x to satisfy

$$\limsup_n \|K^n x\|^{1/n} < r(K). \tag{5.4}$$

Indeed, (5.4) implies that the power series $\sum_{n=0}^{\infty} ||K^n x||^2 \; z^n$ has radius of convergence bigger than $1/r^2$ and, consequently, the series $\sum_n \|K^n x\|^2/r^{2n}$ converges. Since

$$||R_m x||^2 = \sum_{n=0}^{\infty} \left(\frac{m}{1+mr} \right)^{2n} ||K^n x||^2$$

and $\{m/(1+mr)\}$ is an increasing sequence converging to $1/r$, we see that (5.4) implies (5.3).

It is not hard to see that, if K has an eigenvalue λ with the property that $|\lambda| < r(K)$, then any eigenvector corresponding to λ satisfies (5.4). Thus we may assume that 0 is an isolated point of $\sigma(K)$. Let Γ be a positively oriented circle around the origin such that 0 is the only element of $\sigma(K)$ inside the circle, and let

$$P = -\frac{1}{2\pi i} \int_{\Gamma} (K - \lambda I)^{-1} \, d\lambda.$$

By Theorem 4.2.2, P is a projection that commutes with K, and the restriction K_0 of K to the invariant subspace $\operatorname{Ran} P$ is quasinilpotent. It follows that, if x is a unit vector in $\operatorname{Ran} P$, then $||K^n x||^{1/n} = ||K_0^n x||^{1/n} \le \|K_0^n\|^{1/n} \to 0$. This completes the proof of the theorem. □

EXERCISE 5.1.5. Prove that if $u \otimes v$ is a rank one operator, then $\|u \otimes v\| = \|u\|\|v\|$.

As mentioned earlier, the presence of proper invariant subspaces for $\mathcal{B}_K$ (K compact) is an advancement in invariant subspace theory only if $\mathcal{B}_K$ differs from $\{K\}'$. We do not know at the present time if $\mathcal{B}_K$ can equal $\{K\}'$ for a compact nonzero operator K on an infinite dimensional space. We do know that the answer is *no* if K has positive spectral radius.

PROPOSITION 5.1.10. *Let K be a compact operator on an infinte dimensional Hilbert space such that $r(K) > 0$. Then $\mathcal{B}_K \neq \{K\}'$.*

Proof. Notice that the vectors x and y obtained in the proof of Theorem 5.1.8 satisfy (5.3) and $K^*y = \bar{\lambda}y$, with $|\lambda| = r(K)$. Since it was established that $x \otimes y \in \mathcal{B}_K$ it suffices to prove that $K(u \otimes v) \neq (u \otimes v)K$. This follows from the fact that $Kx \neq \lambda x$ which is a simple consequence of (5.3). $\square$