SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The spectral theorem for normal operators and applications

av

Marko Kocic

2020 - No K6

The spectral theorem for normal operators and applications

Marko Kocic

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Sven Raum

2020

Abstract

This paper aims to present the spectral theorem for normal operators and then describe the bilateral shift operator on a sequence space using the spectral theorem. To do so, the notion of Hilbert spaces needs to be defined first, and then we study bounded operators on a Hilbert space. The structures of bounded operators which are presented are the adjoint, the inverse, and the spectrum of bounded operators. Further the spectral theorem for self-adjoint operators is presented, which then gets extended to the spectral theorem for normal operators. Finally we apply the spectral theorem for normal operators to the bilateral shift operator.

Sammanfattning

Målet med arbetet är presentera spektralsatsen för normala operatorer och sen beskriva dubbelsidiga skiftoperatorn på ett sekvensrum. För att göra detta behövs begreppet Hilbertrum definieras först, och sen studeras bundna operatorer på ett Hilbertrum. Strukturer av bundna operatorer som presenteras är adjunkter, inverser och spektrum av bundna operatorer. Vidare visas spektralsatsen för Hermitiska operatorer, som sedan utvidgas till spektralsatsen för normala operatorer. I slutet använder vi spektralsatsen för normala operatorer på dubbelsidiga skiftoperatorn.

Acknowledgements

Special thanks go to my supervisor Sven Raum for his immense patience and willingness to guide and help me in this work. I would also like to thank Mikica Kocic for intellectual support and helping me understand the applications of this problem in physics. Finally I want to thank the rest of my family and friends for supporting directly and indirectly throughout the work.

Contents

Introduction

Studying contemporary mathematics, science and engineering is not possible without the concept of higher dimensional vector spaces. The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are typical examples, extending the notion of $\mathbb{R}$ and $\mathbb{C}$ into higher dimensions. Notwithstanding, these two vector spaces are finite dimensional. In mathematics (e.g., in Fourier analysis) and physics (quantum mechanics) one needs to work with function spaces, wherein the finite dimensional spaces show their limitations. Here the notion of a *Hilbert space* comes in place. It is defined as a vector space with an inner product, where the space is possibly infinite-dimensional but complete with respect to the metric induced by that inner product. This allows us, for instance, to view $\mathbb{C}^n$ as a Hilbert space with the friendly property of being finite dimensional.

The first section of this work is solely dedicated to setup the environment that will be relevant for the succeeding sections. It introduces Hilbert spaces and their properties. By extending the notion of matrices to Hilbert spaces, we arrive at the concept of *bounded operators* on Hilbert spaces, which are bounded linear transformations from a Hilbert space $\mathcal{H}$ to itself. Many of the special types of matrices in $\mathbb{C}^n$ have similar definitions in a Hilbert space, such as *self-adjoint* (or Hermitian), normal, and unitary matrices/operators.

The focus of this work is on the *spectral decomposition* of a bounded normal operator in a general Hilbert space. In $\mathbb{C}^n$, the spectral decomposition of a normal matrix $M \in M_n(\mathbb{C}^n)$ is the same as its eigenvalue decomposition, that is, to write M in the form $U^*MU = D$ where U is a unitary matrix and D is a diagonal matrix whose entries are the eigenvalues of M. Here, an eigenvalue of M is a number $\lambda \in \mathbb{C}$ for which there exists a non-zero vector x satisfying $Mx = \lambda x$. In general Hilbert spaces, however, there may exist normal operators T that have no eigenvalues. The goal is to present a theory that is applicable to the set of λ such that $T - \lambda I$ is not invertible, where the set is called the *spectrum* of T. In a finite-dimensional Hilbert space the set of eigenvalues coincides with the spectrum, which does not hold in an infinite-dimensional space. Our aim is to show the existence of the spectral theorem for normal operators in a general Hilbert space, using the concepts of the approximate eigenvalues and the spectrum.

For the most of theorems and their proofs, this work follows *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* by Paul R. Halmos [1]. Some proofs come from other sources if the proof in Halmos' book is unclear, and some complementary theory not explained clearly in Halmos' book are also included. For further readings on this topic, one can consult *A course in Functional Analysis* by John B. Conway [2] and *Quantum theory for Mathematicians* by Brian C. Hall [3]. The latter covers the applications of the spectral theorem in quantum mechanics.

1 Hilbert Space

1.1 Inner product

In this section we will introduce the notion of *symmetric sesquilinear form*, which is how we will define an inner product. Vaguely one can describe a symmetric sesquilinear form as a map $\mathbb{V} \times \overline{\mathbb{V}} \to \mathbb{C}$ with some additional properties, where $\mathbb{V}$ denotes a complex vector space. What we will do is to define a sesquilinear form, what it means for a sesquilinear form to be symmetric, and finally impose some requirements for a symmetric sesquilinear form to be an inner product.

Definition 1.1 Given two complex vector spaces $\mathbb{V}$ and $\mathbb{W}$, a linear transformation from $\mathbb{V}$ to $\mathbb{W}$ is a map $T : \mathbb{V} \to \mathbb{W}$ such that for all $x, y \in \mathbb{V}$ and for all $\alpha, \beta \in \mathbb{C}$ we have $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$.

Definition 1.2 Let $\mathbb{V}$ be a complex vector space. A linear functional on $\mathbb{V}$ is a linear map $\xi : \mathbb{V} \to \mathbb{C}$. A conjugate linear functional on $\mathbb{V}$ is a map $\xi : \mathbb{V} \to \mathbb{C}$ such that $^{-} \circ \xi$ is a linear functional on $\mathbb{V}$.

Remark 1.3 While the definition of a conjugate linear functional has been given in terms of linear functionals, it is possible to provide a direct description. Let $\xi : \mathbb{V} \to \mathbb{C}$ be a map from a complex vector space $\mathbb{V}$ to $\mathbb{C}$. Then ξ is a conjugate linear functional if and only if

$$\forall x, y \in \mathbb{V}\ \forall \alpha, \beta \in \mathbb{C} : \xi(\alpha x + \beta y) = \overline{\alpha}\xi(x) + \overline{\beta}\xi(y).$$

Indeed, if ξ is a conjugate linear functional, then

$$\overline{\xi(\alpha x + \beta y)} = \overline{\overline{\alpha}\xi(x) + \overline{\beta}\xi(y)} = \alpha\overline{\xi(x)} + \beta\overline{\xi(y)}.$$

Hence, $^{-} \circ \xi$ is a linear functional on $\mathbb{V}$. Conversely, if $^{-} \circ \xi$ is a linear functional on $\mathbb{V}$, then

$$\xi(\alpha x + \beta y) = \overline{\overline{\xi(\alpha x + \beta y)}} = \overline{\alpha\overline{\xi(x)} + \beta\overline{\xi(y)}} = \overline{\alpha}\xi(x) + \overline{\beta}\xi(y).$$

Hence, ξ is a conjugate linear functional.

Definition 1.4 A sesquilinear form on a complex vector space $\mathbb{V}$ is a map $\phi : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ such that

(i) for all $y \in \mathbb{V}$, the map $x \mapsto \phi(x, y)$ is a linear functional, and

(ii) for all $x \in \mathbb{V}$, the map $y \mapsto \phi(x, y)$ is a conjugate linear functional.

Sesquilinear forms as described in Definition 1.4 can be characterised as those maps $\phi : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ satisfying

$$\phi(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1\phi(x_1, y) + \alpha_2\phi(x_2, y) \text{ and}$$
$$\phi(x, \beta_1 y_1 + \beta_2 y_2) = \overline{\beta_1}\phi(x, y_1) + \overline{\beta_2}\phi(x, y_2)$$

for all $x, y, x_1, x_2, y_1, y_2 \in \mathbb{V}$ and for all $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$. This is shown by applying Definition 1.1 and Remark 1.3.

Example 1.5 One example of a sesquilinear form is the standard form on $\mathbb{C}^n$,

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle \stackrel{\text{def}}{=} \sum_{i=1}^{n} x_i \overline{y_i}.$$

We come back to this point in Example 1.9 after introducing the inner product.

Further examples of sesquilinear forms on $\mathbb{C}^n$ can be obtained by the formula $\langle x, y \rangle_M = \langle Mx, y \rangle$, where $M \in M_n(\mathbb{C})$ is any fixed matrix.

Definition 1.6 A sesquilinear form ϕ on a complex vector space $\mathbb{V}$ is symmetric if $\phi(x, y) = \overline{\phi(y, x)}$, for all $x, y \in \mathbb{V}$.

Definition 1.7 An inner product, denoted by $\langle \cdot, \cdot \rangle$, is a symmetric sesquilinear form that is positive definite; i.e. $\langle x, x \rangle > 0$ when $x \neq 0$ and $\langle x, x \rangle = 0$ when $x = 0$. An inner product space is a complex vector space $\mathbb{V}$ with an inner product, denoted by the pair $(\mathbb{V}, \langle \cdot, \cdot \rangle)$.

Now that we know what an inner product is, let us state a useful criterion to check equality of vectors.

Lemma 1.8 *Let $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ be an inner product space and $x \in \mathbb{V}$. Then a necessary and sufficient condition that $x = 0$ is that $\langle x, y \rangle = 0$ for all $y \in \mathbb{V}$.*

Proof. If $\langle x, y \rangle = 0$ for all y, then it follows that $\langle x, x \rangle = 0$. The inner product is strictly positive, hence $x = 0$. Conversely, if $x = 0$, then $\langle x, y \rangle = \langle 0x, y \rangle = 0 \langle x, y \rangle = 0$ for all y. □

Example 1.9 One example of an inner product was already mentioned in Example 1.5 which is the standard form on $\mathbb{C}^n$. The fact that it is symmetric can be verified by applying Definition 1.6 on the standard form, and checking that it is positive definite is done by a computation.

The other sesquilinear form mentioned in Example 1.4, that is the map $(x, y) \mapsto \langle Mx, y \rangle$, is symmetric if and only if $M = M^*$ where M^* is the conjugate transpose of M. In particular, taking $M = I_n$ gives the standard form on $\mathbb{C}^n$ which is an inner product, which follows right away from the definition of the standard form mentioned in Example 1.5. Let us now argue that $M = M^*$ is necessary and sufficient condition for $\langle Mx, y \rangle$ to be symmetric on $\mathbb{C}^n$. If $M = M^*$, then

$$\langle x, y \rangle_M = \langle Mx, y \rangle = \langle x, My \rangle = \overline{\langle My, x \rangle} = \overline{\langle y, x \rangle}_M$$

shows that $\langle \cdot, \cdot \rangle_M$ is symmetric. Conversely, if $\langle \cdot, \cdot \rangle_M$ is symmetric then we have that

$$\langle x, y \rangle_M = \overline{\langle y, x \rangle}_M = \overline{\langle My, x \rangle} = \langle x, My \rangle = \langle M^* x, y \rangle ,$$

which shows that $\langle Mx, y\rangle = \langle M^*x, y\rangle$ for all $x, y \in \mathbb{C}^n$. Hence, $\langle (M - M^*)x, y\rangle = 0$ for all $x, y \in \mathbb{C}^n$. By Lemma 1.8 we have that $M - M^* = 0$ and hence $M = M^*$.

To conclude this example we will show the following statement, $(\mathbb{C}^n, \langle\cdot,\cdot\rangle_M)$ is an inner product space if and only if M is positive definite. Assume that M is positive definite. We note that M is positive definite if $z^*Mz > 0$ for all non-zero $z \in \mathbb{C}^n$. This translates exactly to $\langle Mz, z\rangle = \langle z, z\rangle_M > 0$ so we have our desired result. Conversely, suppose $\langle\cdot,\cdot\rangle_M$ is an inner product. Then by definition of an inner product, $\langle z, z\rangle_M = \langle Mz, z\rangle \geq 0$ with $\langle Mz, z\rangle = 0$ only when $z = 0$. Hence, M is positive definite.

Before we end this section, let us state an important inequality that follows from an inner product.

Theorem 1.10 (The Cauchy-Schwarz inequality) *Let* $(\mathbb{V}, \langle\cdot,\cdot\rangle)$ *be an inner product space. Then the inequality* $|\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle$ *holds for all* $x, y \in \mathbb{V}$.

Proof. If either x or y is 0, then the inequality is clear. Let $x \neq 0$ and $y \neq 0$. Define the complex number λ to be $\lambda = \langle x, y\rangle / \langle y, y\rangle$, then

$$\begin{aligned}
0 &\leq \langle x - \lambda\cdot y, x - \lambda\cdot y\rangle \\
&= \langle x, x\rangle - \langle x, \lambda\cdot y\rangle - \langle \lambda\cdot y, x\rangle + \langle \lambda\cdot y, \lambda\cdot y\rangle \\
&= \langle x, x\rangle - \overline{\lambda}\,\langle x, y\rangle - \lambda\,\langle y, x\rangle + \lambda\overline{\lambda}\,\langle y, y\rangle \\
&= \langle x, x\rangle - \overline{\lambda}\,\langle x, y\rangle - \lambda\overline{\langle x, y\rangle} + \lambda\overline{\lambda}\,\langle y, y\rangle \\
&= \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} = \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}.
\end{aligned}$$

Hence,

$$0 \leq \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}$$

which is equivalent to

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle\,. \qquad \square$$

This variation of the proof can be found on Wikipedia's article on Cauchy-Schwarz inequality [4]. While the source is usually criticised for its inaccuracy, one can verify that this is indeed a correct proof of the Cauchy-Schwarz inequality.

1.2 Metric and norm

While it is easier to think of Hilbert spaces as a generalisation of Euclidean spaces, the preferred way to go is to impose stricter properties on abstract sets. To make it simpler for us, we will skip topological spaces entirely and start directly on the definition of a metric space.

Definition 1.11 A metric (sometimes called a distance function) on a set X is a function $d : X \times X \to \mathbb{R}_{\geq 0}$ such that the following properties holds

(i) for all $x, y \in X$ we have $d(x, y) > 0$ if $x \neq y$ and $d(x, x) = 0$.

(ii) For all $x, y \in X$ we have $d(x, y) = d(y, x)$.

(iii) For all $x, y, z \in X$ we have $d(x, y) \leq d(x, z) + d(z, y)$.

A metric space is a pair (X, d) where X is a set and d is a metric on X.

What we want now is the case when a metric space is said to be *complete*. This is going to be our next few definitions.

Definition 1.12 Let (X, d) be a metric space. A sequence $(x_n)_n$ is said to converge if there exists a point $x \in X$ such that for all $\epsilon > 0$ there exists an integer N such that $n \geq N$ implies that $d(x_n, x) < \epsilon$. A sequence $(x_n)_n$ is said to be a *Cauchy sequence* if for all $\epsilon > 0$ there exists an integer N such that $n \geq N$ and $m \geq N$ implies that $d(x_n, x_m) < \epsilon$.

Definition 1.13 Let (X, d) be a metric space. If every Cauchy sequence in X converges to some point in X, then we say that (X, d) is complete.

The next step is to define the notion of a *normed vector space*.

Definition 1.14 Let $\mathbb{V}$ be a complex vector space. A norm on $\mathbb{V}$ is a map $\|\cdot\| : \mathbb{V} \to \mathbb{R}_{\geq 0}$ such that

(i) for all $x \in \mathbb{V}$ we have $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$ (strict positiveness).

(ii) For all $x \in \mathbb{V}$ and all $\alpha \in \mathbb{C}$ we have $\|\alpha x\| = |\alpha| \cdot \|x\|)$ (positively homogeneous).

(iii) For all $x, y \in \mathbb{V}$ we have $\|x + y\| \leq \|x\| + \|y\|$ (subadditivity).

A normed vector space is a pair $(\mathbb{V}, \|\cdot\|)$ where $\mathbb{V}$ is a vector space and $\|\cdot\|$ is a norm on $\mathbb{V}$.

Following this definition, we want to define a metric with the norm of a normed vector space. This leads to the following theorem.

Theorem 1.15 *Let* $(\mathbb{V}, \|\cdot\|)$ *be a normed vector space. Then* $d : \mathbb{V} \times \mathbb{V} \to \mathbb{R}_{\geq 0}$ *defined to be* $d(x, y) = \|x - y\|$ *is a metric.*

Proof [5]. A metric needs to satisfy the properties of Definition 1.11, so let us show precisely that. Then *(i)* follows from the strict positiveness of the norm in Definition 1.14, and $\|x - y\| = 0$ if and only if $x = y$. The *(ii)* point follows from the positive homogeneity of the norm in conjunction with the identity $(x - y) = -1(y - x)$, by the calculation

$$\|x - y\| = \|(-1)(y - x)\| = |-1| \, \|y - x\| \, .$$

The triangle inequality *(iii)* follows from the subadditivity of the norm and the fact that $x - y = (x - z) + (z - y)$, by the calculation

$$\|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| \, . \qquad \square$$

With this theorem we have proved that normed vector spaces are also metric spaces, where a metric is induced by the norm. This also leads to a special case of a normed vector space, in particular when it is complete. A complete normed vector space is called a *Banach space.*

1.3 Hilbert space and examples

While Banach spaces have possible topics of research where a norm is good enough, we want to impose one more structure to a Banach space which we do with an inner product.

Theorem 1.16 *Let* $(\mathbb{V}, \langle\cdot,\cdot\rangle)$ *be an inner product space. Then the formula*

$$\|x\| \overset{\text{def}}{=} \sqrt{\langle x, x\rangle}$$

defines a norm on $\mathbb{V}$*.*

Proof. Strict positiveness of the norm is a straight forward consequence of the strict positiveness of the inner product. The positive homogeneity of the norm is a consequence of the identity

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\,\langle x, x\rangle = |\alpha|^2 \cdot \|x\|^2\,.$$

The subadditivity of the norm follows from the Cauchy-Schwarz inequality, that is

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle &= \langle x, x\rangle + \langle x, y\rangle\,\langle y, x\rangle + \langle y, y\rangle \\ &\le \|x\|^2 + |\langle x, y\rangle| \cdot |\langle y, x\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\,\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

□

With this theorem we conclude that a norm is induced by the inner product of an inner product space, which also induces a metric on the same space. Thus we are allowed to talk about complete inner product spaces, which finally lets us define a Hilbert space.

Definition 1.17 A Hilbert space is an inner product space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ that is complete with respect to its metric. In this paper, Hilbert spaces are denoted by $\mathcal{H}$.

Example 1.18 In the Examples 1.5 and 1.9 we talked about the space $\mathbb{C}^n$, which we now can verify that it is indeed a Hilbert space. The inner product for $\mathbb{C}^n$ is the standard form, which is defined to be

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i\overline{y_i}.$$

A Hilbert space similar to $\mathbb{C}^n$ is the real Hilbert space $\mathbb{R}^n$. By appropriately defining an analogue notion of a real Hilbert space over $\mathbb{R}$, then $\mathbb{R}^n$ also fits in this framework. Note however that both $\mathbb{R}^n$ and $\mathbb{C}^n$ are finite dimensional, yet Hilbert spaces do not place any requirements about the dimension. Hilbert spaces can possibly be infinite dimensional, if not stated otherwise.

Example 1.19 Another example of a Hilbert space is the Lebesgue space $L^2(X, \mu)$, the space of all equivalence classes of complex, measurable, and square-integrable functions on a measure space X with the measure μ, which we will introduce now. First we consider the space of all such functions

$$\mathcal{L}^2(X, \mu) = \left\{ f : X \to \mathbb{C} \middle| \int_X |f(t)|^2 \mathrm{d}\mu(t) < \infty \right\}.$$

Then a symmetric sesquilinear form can be defined on $\mathcal{L}^2(X, \mu)$ by

$$\langle f, g \rangle = \int_X f(t)\overline{g(t)} \, \mathrm{d}\mu(t).$$

Note however that there exists a possibly non-trivial subspace

$$N = \left\{ f \in \mathcal{L}^2(X, \mu) \middle| \langle f, f \rangle = 0 \right\} \subseteq \mathcal{L}^2(X, \mu),$$

which would contradict the inner product by Definition 1.7. We hence have the quotient space

$$L^2(X, \mu) = \mathcal{L}^2(X, \mu)/N$$

and obtain a well-defined inner product

$$\langle f + N, g + N \rangle \stackrel{\text{def}}{=} \langle f, g \rangle .$$

One can check by means of Cauchy-Schwarz that the inner product is indeed well-defined, and that the space is complete.

Example 1.20 The final example of a Hilbert space in this paper will be the Hardy space H^2. The Hardy space H^2 is the space

$$H^2(\mathbb{D}) = \left\{ f : \mathbb{D} \to \mathbb{C} \middle| \begin{array}{c} f \text{ holmorphic} \\ f(z) = \sum\limits_{n=0}^{\infty} a_n z^n \text{ around } 0 \\ \sum\limits_{n=0}^{\infty} |a_n|^2 < \infty \end{array} \right\}.$$

The inner product of $f(z) = \sum\limits_{n=0}^{\infty} a_n z^n$ with $g(z) = \sum\limits_{n=0}^{\infty} b_n z^n$ is defined as

$$\langle f, g \rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

Then one can show that this is well-defined and that the norm can be characterised as

$$\|f\|^2 = \sup_{0<r<1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 \mathrm{d}\theta < \infty.$$

One can also check that this space is indeed complete.

This is an example of a Hilbert space whose elements are actual functions and not just equivalence classes of functions. Such Hilbert spaces, assuming additionally certain compatibility condition with point evaluation, are called reproducing kernel Hilbert spaces.

2 The structure of bounded operators on a Hilbert space

2.1 Fréchet-Riesz representation theorem

The theorem we will state here will be very useful later on. However, we need to state the next definition to make sense of the theorem.

Definition 2.1 Let $\mathcal{H}$ a Hilbert space and $\xi : \mathcal{H} \to \mathbb{C}$ be a linear functional. We say that ξ is bounded if there exists a positive real number α such that $\|\xi(x)\| \leq \alpha \|x\|$ for all $x \in \mathcal{H}$. The norm of ξ, denoted by $\|\xi\|$, is the infimum of all such α.

Example 2.2 One example of such bounded linear functional $\xi : \mathcal{H} \to \mathbb{C}$ is the map $x \mapsto \langle x, y \rangle$ for some $y \in \mathcal{H}$. The norm of this functional is given by $\|y\|$, which follows from Cauchy-Schwarz inequality.

Theorem 2.3 (Fréchet-Riesz representation theorem) *Let ξ be a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there exists a unique vector $y \in \mathcal{H}$ such that $\xi(x) = \langle x, y \rangle$ for all $x \in \mathcal{H}$.*

Proof. Let $\mathcal{K} = \ker(\xi)$. Then $\mathcal{K}$ is a subspace of $\mathcal{H}$. If $\mathcal{K} = \mathcal{H}$, then $\xi(x) = \langle x, 0 \rangle$ for all $x \in \mathcal{H}$. Otherwise, take $z \in \mathcal{K}^\perp$ satisfying $\xi(z) = 1$. Note that such z exists, since if there are two non-zero $z_1, z_2 \in \mathcal{K}^\perp$ then there exists a $\lambda \in \mathbb{R}$ such that $\lambda\xi(z_1) = \xi(z_2)$. Observing the fact that $\lambda z_1 - z_2 \in \mathcal{K}^\perp$ and $\xi(\lambda z_1 - z_2) = 0$ we obtain $\lambda z_1 - z_2 \in \mathcal{K}$, i.e. $\lambda z_1 - z_2 = 0$. Then given $x \in \mathcal{H}$, we have

$$\xi(x - \xi(x)z) = \xi(x) - \xi(x)\xi(z) = 0.$$

So $x - \xi(x)z \in \mathcal{K}$. Hence

$$x = (x - \xi(x)z) + \xi(x)z \in \mathcal{K} + \mathbb{C}z,$$

which shows that $\mathcal{H} = \mathcal{K} + \mathbb{C}z$.

Let $y = z / \|z\|^2$. Then

$$\xi(z) = 1 = \langle z, y \rangle$$

and

$$\xi(x) = \langle x, y \rangle = 0$$

for all $x \in \mathcal{K}$. So $\xi = \langle \cdot, y \rangle$ on generators of $\mathcal{H} = \mathcal{K} + \mathbb{C}z$. □

2.2 Operators, adjoint and self-adjoint

Now is an appropriate time to talk about operators on a Hilbert space $\mathcal{H}$. For the purpose of this paper, we are concerned with bounded operators. However, one can make the jump to unbounded operators in fields where those are required, such as quantum mechanics, but that will not be covered here.

Definition 2.4 Let $\mathcal{H}$ a Hilbert space and $T : \mathcal{H} \to \mathcal{H}$ be a linear transformation. Then we say that T is an *operator*. Further, we say that T is a *bounded operator* if there exists a positive real number α such that $\|Tx\| \leq \alpha \|x\|$ for all $x \in \mathcal{H}$. The set of bounded operators will be denoted by $B(\mathcal{H})$.

We have already seen examples of the next definitions before in Examples 1.5 and 1.9. The examples were however in $\mathbb{C}^n$, where the given M is a matrix in $M_n(\mathbb{C})$, so M^* becomes the transpose conjugate of M. It is now convenient for us to properly define the adjoint, and more importantly self-adjointness, of a bounded operator and see that Examples 1.5 and 1.9 are the special cases when the Hilbert space is $\mathbb{C}^n$.

Theorem 2.5 *Let $\mathcal{H}$ be a Hilbert space. The map from bounded operators on $\mathcal{H}$ to bounded sesquilinear forms on $\mathcal{H}$, i.e. the map $\psi : B(\mathcal{H}) \to BS(\mathcal{H})$ satisfying $T \mapsto ((x, y) \mapsto \langle x, Ty\rangle)$, is a well-defined conjugate linear isomorphism of vector spaces.*

Proof. This proof is split up into four parts.

Well-definedness: Let $T \in B(\mathcal{H})$ and $x, y \in \mathcal{H}$. Then boundedness of $\psi(T)$ follows from

$$\|\langle x, Ty\rangle\| \leq \|x\| \cdot \|Ty\| \leq \|x\| \cdot \|T\| \cdot \|y\| .$$

Sesquilinearity of $\psi(T)$ follows from the sesquilinearity of the inner product.

Conjugate linearity: Let $S, T \in B(\mathcal{H})$ and $x, y \in \mathcal{H}$. Then

$$\langle x, \lambda(S + T)y\rangle = \overline{\lambda}(\langle x, Sy\rangle + \langle x, Ty\rangle) = \overline{\lambda} \langle x, Sy\rangle + \overline{\lambda} \langle x, Ty\rangle .$$

Injectivity: By linearity, it suffices to check that $\ker \psi = 0$. If $\langle x, Ty\rangle = 0$ for all x and y, then by Lemma 1.8 we get $Ty = 0$ for all $y \in \mathcal{H}$. Hence $T = 0$.

Surjectivity: We need to show that for all $F \in (\mathcal{H} \otimes \overline{\mathcal{H}})^*$ there exists a $T \in B(\mathcal{H})$ such that for all $x, y \in \mathcal{H}$ we have $F(x, y) = \langle x, Ty\rangle$. We start by defining a map $T : \mathcal{H} \to \mathcal{H}$. Given $y \in \mathcal{H}$, denote by Ty the unique vector satisfying $F(x, y) = \langle x, Ty\rangle$. This vector indeed exists by Fréchet-Riesz representation theorem. We will show that T is a bounded linear operator.

To show T is linear, we need to show that for all vectors y_1 and y_2, and for all $\lambda \in \mathbb{C}$ that we have $T(\lambda(y_1 + y_2)) = \lambda(Ty_1 + Ty_2)$. For all vectors x, we have

$$\begin{aligned}
&\langle x, T(\lambda(y_1 + y_2)\rangle \\
&\quad = F(x, \lambda(y_1 + y_2)) \\
&\quad = \overline{\lambda}(F(x, y_1) + F(x, y_2)) \\
&\quad = \overline{\lambda}(\langle x, Ty_1\rangle + \langle x, Ty_2\rangle) \\
&\quad = \langle x, \lambda(Ty_1 + Ty_2)\rangle .
\end{aligned}$$

Further, for $y \in \mathcal{H}$ we have

$$\|Ty\| = \sup_{x,\|x\|\le 1} \langle x, Ty\rangle = \sup_{x,\|x\|\le 1} F(x,y) \le \sup_{x,\|x\|\le 1} \|F\| \cdot \|x\| \cdot \|y\| \le \|F\| \cdot \|y\| .$$

This shows that T is bounded and that $\|T\| \le \|F\|$, so $T \in B(\mathcal{H})$. □

Proposition 2.6 *Let $T \in B(\mathcal{H})$. Then there exists a unique operator S such that $\langle Sx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{H}$.*

Proof. Both existence and uniqueness follows from Theorem 2.5. Let $T \in B(\mathcal{H})$. Then $(x, y) \mapsto \langle x, Ty\rangle$ is a bounded sesquilinear form. Hence, there is a unique $S \in B(\mathcal{H})$ such that $\langle Sx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{H}$. □

Definition 2.7 Let $T \in B(\mathcal{H})$. The *adjoint* of T is denoted by T^* and satisfies the equality $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in \mathcal{H}$. If $T = T^*$, then we say that the operator T is *self-adjoint*. The set of bounded self-adjoint operators will be denoted by $B(\mathcal{H})_{sa}$.

Proposition 2.8 *Let $T \in B(\mathcal{H})$. Then T is self-adjoint if and only if $\psi(T)$ defined in Theorem 2.5 is symmetric.*

Proof. Let $T \in B(\mathcal{H})_{sa}$. Then for arbitrary $x, y \in \mathcal{H}$ we have

$$\psi(T)(x,y) = \langle x, Ty\rangle = \langle Tx, y\rangle = \overline{\langle y, Tx\rangle} = \overline{\psi(T)(y,x)}.$$

So $\psi(T)$ is symmetric. Conversely, if $\langle x, Ty\rangle = F(x,y)$ is symmetric, then

$$\langle x, Ty\rangle = F(x,y) = \overline{F(y,x)} = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle$$

for all $x, y \in \mathcal{H}$. Hence, T is self-adjoint. □

Before we end this section, let us characterise some properties of adjoints.

Proposition 2.9 *Let $T, S \in B(\mathcal{H})$ and let $\lambda \in \mathbb{C}$. Then*

(i) $(T^*)^* = T$,

(ii) $(\lambda T)^* = \overline{\lambda} T^*$,

(iii) $(T+S)^* = T^* + S^*$, *and*

(iv) $(TS)^* = S^*T^*$.

Proof. All of these are implied by the following identities,

(i) $\langle (T^*)^*x, y\rangle = \langle x, T^*y\rangle = \overline{\langle T^*y, x\rangle} = \overline{\langle y, Tx\rangle} = \langle Tx, y\rangle$,

(ii) $\langle (\lambda T)^*x, y\rangle = \langle x, \lambda Ty\rangle = \overline{\lambda}\,\langle x, Ty\rangle = \overline{\lambda}\,\langle T^*x, y\rangle = \left\langle \overline{\lambda} T^*x, y\right\rangle$,

(iii) $\langle (T+S)^*x, y\rangle = \langle x, (T+S)y\rangle = \langle x, Ty\rangle + \langle x, Sy\rangle = \langle T^*x, y\rangle + \langle S^*x, y\rangle = \langle (T^* + S^*)x, y\rangle$, and

(iv) $\langle (TS)^*x, y\rangle = \langle x, TSy\rangle = \langle T^*x, Sy\rangle = \langle S^*T^*x, y\rangle$. □

2.3 Inverse of an operator

Definition 2.10 Let $T \in B(\mathcal{H})$. Then T is *invertible* if there exists an operator $S \in B(\mathcal{H})$ such that $TS = ST = I$, where I denotes the identity operator. The inverse of T will be denoted by T^{-1}.

Lemma 2.11 *Let $T \in B(\mathcal{H})$ and α be some positive real number. If $\|Tx\| \geq \alpha \|x\|$ for all $x \in \mathcal{H}$, then the image of T is closed.*

Proof. Let $(x_n)_n$ be a sequence such that $y_n = Tx_n$ defines a convergent sequence $y_n \to y \in \mathcal{H}$. Then we have

$$\|y_n - y_m\| = \|Tx_n - Tx_m\| \geq \alpha \|x_n - x_m\|$$

for all $n, m \in \mathbb{N}$, which shows $(x_n)_n$ is a Cauchy sequence and hence that there exists some vector x such that $x_n \to x$. Since

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq \|T\| \cdot \|x_n - x\| \xrightarrow{n\to\infty} 0,$$

we get that

$$y = \lim_{n\to\infty} y_n = \lim_{n\to\infty} Tx_n = Tx.$$

Hence $y \in \operatorname{Img} T$. □

Theorem 2.12 *Let $T \in B(\mathcal{H})$. If the image of T is dense in $\mathcal{H}$ and there exists a positive real number α such that $\|Tx\| \geq \alpha \|x\|$ for every vector x, then T is invertible.*

Proof. First we show that T is bijective. By Lemma 2.11, the image of T is all of $\mathcal{H}$. Further, if $Tx_1 = Tx_2$ for some vectors $x_1, x_2 \in \mathcal{H}$ then

$$0 = \|Tx_1 - Tx_2\| \geq \alpha \|x_1 - x_2\|.$$

Hence, $x_1 = x_2$, which shows injectivity of T. So for every $y \in \mathcal{H}$ there exists exactly one vector $x \in \mathcal{H}$ such that $y = Tx$. This gives a transformation $S : \mathcal{H} \to \mathcal{H}$ such that $Sy = STx = x$.

What is left to show is that S is a bounded linear transformation, i.e. a bounded operator. For any $y_1, y_2 \in \mathcal{H}$ such that $Sy_1 = x_1$ and $Sy_2 = x_2$, and $\alpha, \beta \in \mathbb{C}$, linearity is given by

$$S(\alpha y_1 + \beta y_2) = S(\alpha Tx_1 + \beta Tx_2) = ST(\alpha x_1 + \beta x_2) = \alpha x_1 + \beta x_2 = \alpha Sy_1 + \beta Sy_2.$$

Boundedness is given by

$$\|y\| = \|Tx\| \geq \alpha \|x\| = \alpha \|Sy\|,$$

so S is indeed a bounded operator. Since $TSy = Tx = x$ and since $STx = Sy = y$ for all $x, y \in \mathcal{H}$, we have that $TS = ST = I$ which gives our desired result. □

Theorem 2.13 *Let $T \in B(\mathcal{H})$ and I be the identity operator. If $\|I - T\| < 1$, then T is invertible.*

Proof. By Theorem 2.12, it suffices to show that the image of T is dense in $\mathcal{H}$ and that $\|Tx\| \geq \alpha \|x\|$. Let $1 - \|I - T\| = \alpha$ for $0 \leq \alpha < 1$. Then we have

$$\|Tx\| = \|x - (x - Tx)\| \geq \|x\| - \|(I - T)x\| \geq \|x\| - \|I - T\| \|x\| = \alpha \|x\|$$

for all $x \in \mathcal{H}$.

To show density, let $y \in \mathcal{H}$. Then for all $n \in \mathbb{N}$ we have $(I - T)^n y = y - Tx$ for some $x \in \mathcal{H}$. Further, we have that

$$\|(I - T)^n y\| \leq \|(I - T)^n\| \cdot \|y\| = \|I - T\|^n \cdot \|y\| \xrightarrow{n \to \infty} 0.$$

Hence, for all $\epsilon > 0$ there is some $x \in \mathcal{H}$ such that $\|y - Tx\| < \epsilon$. □

2.4 Spectrum of an operator

Recall the definition of eigenvalues. We say that $\lambda \in \mathbb{C}$ is an *eigenvalue* to $T \in B(\mathcal{H})$ if there exists a vector $x \in \mathcal{H}$ such that $Tx = \lambda x$. The idea of this section is to generalise this notion, since in the finite dimensional $\mathbb{C}^n$ eigenvalues of T directly correspond to the set of values when $(T - \lambda I)$ is not invertible. Let us formally introduce this notion.

Definition 2.14 Let $T \in B(\mathcal{H})$. The *spectrum* of T is the set of all those complex numbers λ for which $T - \lambda I$ is not invertible. The spectrum is denoted by $\sigma(T)$.

Example 2.15 Let $\mathcal{H}$ be the Hilbert space $\mathbb{C}^n$ and $M \in M_n(\mathbb{C}^n)$. Then λI is given by the matrix

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Since we are in a finite-dimensional environment, the determinant of $M - \lambda I$ is defined. In fact, because a matrix is not invertible if and only if its determinant is zero, we have that $M - \lambda I$ is not invertible if and only if $\det(M - \lambda I) = 0$. This gives us a polynomial of n-th degree, whose roots are in $\sigma(M)$. Hence, the spectrum of $M \in \mathbb{C}^n$ is given by the roots of the polynomial $\det(M - \lambda I) = 0$. Since the Hilbert space is $\mathbb{C}^n$, these are exactly the eigenvalues of M.

In finite-dimensional Hilbert spaces, eigenvalues of T correspond directly to $\sigma(T)$. This may not be the case in infinite-dimensional Hilbert spaces. We will see an example of a bounded operator in Section 4 that has no eigenvalues, but has a non-empty spectrum. However, there is a similar notion to eigenvalues called *approximate eigenvalues* which we define now.

Definition 2.16 Let $T \in B(\mathcal{H})$. A complex number λ is called an *approximate eigenvalue* of T if for every positive ϵ there exists a unit vector x such that $\|Tx - \lambda x\| < \epsilon$. The *approximate point spectrum* of T is the set of approximate eigenvalues and will be denoted by $\sigma_{ap}(T)$.

Lemma 2.17 *Let $T \in B(H)$. Then $\sigma_{ap}(T) \subseteq \sigma(T)$.*

Proof. Let $\lambda \notin \sigma(T)$. Since $T - \lambda I$ is invertible, we have for all $x \in \mathcal{H}$ that

$$\begin{aligned} \|x\| &= \left\|(T-\lambda I)^{-1}(T-\lambda I)x\right\| \\ &= \left\|(T-\lambda I)^{-1}(Tx-\lambda x)\right\| \\ &\leq \left\|(T-\lambda I)^{-1}\right\| \cdot \|Tx-\lambda x\| . \end{aligned}$$

This implies that $\|Tx - \lambda x\| \geq \epsilon \|x\|$, with $\epsilon = 1/\left\|(T-\lambda I)^{-1}\right\|$, for all $x \in \mathcal{H}$. Hence, $\sigma_{ap}(T) \subseteq \sigma(T)$. □

Proposition 2.18 *Let $T \in B(\mathcal{H})$ such that T is invertible. Then $\sigma(T^{-1}) = (\sigma(T))^{-1}$.*

Proof. Let $\lambda \notin \sigma(T)$. Since

$$(\lambda I - T)\lambda^{-1}T^{-1} = T^{-1} - \lambda^{-1}I,$$

and the left hand side is invertible, so is the right hand side. In other words $\lambda^{-1} \notin \sigma(T^{-1})$ so $\sigma(T^{-1}) \subset (\sigma(T))^{-1}$. Since $(T^{-1})^{-1} = T$, the reverse inclusion follows from the same proof applied to T^{-1} instead of T. □

Before we state the next proposition, we need a corollary from Proposition 2.9

Corollary 2.19 *Let $T \in B(\mathcal{H})$ be invertible. Then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.*

Proof. Apply (iv) to $S = T^{-1}$. Then we have

$$(T^{-1})^*T^* = (T^{-1}T)^* \qquad \text{and} \qquad T^*(T^{-1})^* = (TT^{-1})^*. \qquad \square$$

Proposition 2.20 *Let $T \in B(\mathcal{H})$. Then $\sigma(T^*) = \{\overline{\lambda} | \lambda \in \sigma(T)\}$.*

Proof. Let $\lambda \notin \sigma(T)$ so $T - \lambda I$ is invertible. By Corollary 2.19 and Proposition 2.9 we have that $(T - \lambda I)^* = T^* - \overline{\lambda} I$ is invertible and hence $\overline{\lambda} \notin \sigma(T^*)$. Since $(T^*)^* = T$ by Proposition 2.9, the reverse inclusion follows from the same proof applied to T^* instead of T. □

Proposition 2.20 applied to a self-adjoint operator gives the result that the spectrum of a self-adjoint operator is symmetric with respect to the real axis. However, the spectrum of a self-adjoint operator is actually a subset of the real numbers. Before we prove that however, we need the following lemma.

Lemma 2.21 *Let $T \in B(\mathcal{H})$. Then $(\operatorname{Img} T)^{\perp} = \ker T^*$.*

Proof. The statement follows from the following chain of equivalences

$$\begin{aligned} y \in (\operatorname{Img} T)^{\perp} &\iff \langle Tx, y\rangle = 0 \text{ for all } x \in \mathcal{H} \\ &\iff \langle x, T^{*}y\rangle = 0 \text{ for all } x \in \mathcal{H} \\ &\iff T^{*}y = 0 \\ &\iff y \in \ker T^{*}. \end{aligned}$$

□

Proposition 2.22 *Let $T \in B(\mathcal{H})_{sa}$. Then $\sigma(T)$ is a subset of $\mathbb{R}$.*

Proof. Let $\lambda = a + ib \in \mathbb{C}$ with $b \neq 0$ and let $x \in \mathcal{H}$. Then we have that

$$\begin{aligned} &\langle (T-(a+ib)I)x, (T-(a+ib)I)x\rangle \\ &\quad = \langle (T-aI)x, (T-aI)x\rangle + ib\,\langle x, (T-aI)\rangle \\ &\qquad - ib\,\langle (T-aI)x, x\rangle + b^{2}\,\langle x, x\rangle\,. \end{aligned}$$

Since $(T-aI)$ is self-adjoint, $\langle (T-aI)x, (T-aI)x\rangle$ is real and the imaginary terms cancel each other. So we are left with

$$\langle (T-\lambda I)x, (T-\lambda I)x\rangle \geq b^{2}\,\langle x, x\rangle$$

and hence $T-\lambda I$ is injective. Further, we have

$$(\operatorname{Img}\,(T-\lambda I))^{\perp} = \ker(T-\overline{\lambda} I)$$

by Lemma 2.21. Since $T-\overline{\lambda} I$ also has a non-zero imaginary part, it also is injective by the first part of the proof. So the image of $T-\lambda I$ is dense in $\mathcal{H}$. By Theorem 2.12, since $T-\lambda I$ is bounded from below and the range is dense in $\mathcal{H}$ we conclude that $T-\lambda I$ is invertible for $\lambda \in \mathbb{C}\setminus\mathbb{R}$. □

Theorem 2.23 *Let $T \in B(\mathcal{H})$. Then $\sigma(T)$ is a compact subset of the complex plane. Further, if $\lambda \in \sigma(T)$, then $|\lambda| \leq \|T\|$.*

Proof. First we show that $\mathbb{C}\setminus\sigma(T)$ is open. Let $\lambda_0 \notin \sigma(T)$ so that $T-\lambda_0 I$ is invertible. Then

$$\begin{aligned} &\left\| I-(T-\lambda_0 I)^{-1}(T-\lambda I)\right\| \\ &\quad = \left\|(T-\lambda_0 I)^{-1}((T-\lambda_0 I)-(T-\lambda I))\right\| \\ &\quad \leq \left\|(T-\lambda_0 I)^{-1}\right\| \cdot |\lambda-\lambda_0|. \end{aligned}$$

This shows that $\|I-(T-\lambda_0 I)^{-1}(T-\lambda I)\| < 1$ if $|\lambda-\lambda_0|$ is small enough. So by Theorem 2.13, $T-\lambda I$ is invertible if $|\lambda-\lambda_0|$ is small enough. Hence, $\sigma(T)^{C}$ is an open subset of the complex plane.

If $|\lambda| > \|T\|$, then $\|\lambda^{-1}T\| < 1$ and hence $(I-\lambda^{-1}T)$ is invertible. Hence $\lambda \notin \sigma(T)$, so if $\lambda \in \sigma(T)$ then $|\lambda| \leq \|T\|$. □

Theorem 2.24 *Let $T \in B(\mathcal{H})$ and p be a polynomial. Then*

$$\sigma(p(T)) = p(\sigma(T)) = \{p(\lambda) | \lambda \in \sigma(T)\}.$$

Proof. We first show that $p(\sigma(T)) \subseteq \sigma(p(T))$. Let λ_0 be any complex number. Then there exists a polynomial q such that

$$p(\lambda) - p(\lambda_0) = (\lambda - \lambda_0)q(\lambda)$$

identically in λ, and it follows that

$$p(T) - p(\lambda_0 I) = (T - \lambda_0 I)q(T).$$

Now we need to show that if $\lambda_0 \in \sigma(T)$, then $S = (T - \lambda_0)q(T)$ is not invertible. Assume for a contradiction, then S is invertible and

$$(T - \lambda_0 I)q(T)S^{-1} = SS^{-1} = I = S^{-1}S = S^{-1}(T - \lambda_0 I)q(T) = S^{-1}q(T)(T - \lambda_0 I).$$

This shows that $T - \lambda_0 I$ is also invertible, which gives us a contradiction. So we conclude that $p(T) - p(\lambda_0 I)$ is not invertible and hence $p(\lambda_0) \in \sigma(p(T))$ and $p(\sigma(T)) \subseteq \sigma(p(T))$.

Conversely, let $\lambda_0 \in \sigma(p(T))$ and let λ_j be the roots of $p(\lambda) - \lambda_0 = 0$ for $1 \leq j \leq n$, $j, n \in \mathbb{N}$. Then it follows that

$$p(T) - \lambda_0 I = \alpha \prod_{j=1}^{n} (T - \lambda_j I)$$

for some $\alpha \in \mathbb{C}$. Hence, $T - \lambda_j I$ is not invertible in at least one value of j. Then, since $\lambda_j \in \sigma(T)$ and $p(\lambda_j) = \lambda_0$, we have that $\lambda_0 \in p(\sigma(T))$ and hence $\sigma(p(T)) \subseteq p(\sigma(T))$. □

Theorem 2.25 *Let $T \in B(\mathcal{H})_{sa}$. Then* $\|T\| = \sup\left\{|\lambda| \middle| \lambda \in \sigma(T)\right\}$.

Proof. Let $\alpha = \sup\left\{|\lambda| \middle| \lambda \in \sigma(T)\right\}$. By Theorem 2.23, we know that $\alpha \leq \|T\|$. So we need to show that $\alpha \geq \|T\|$, which we do by showing that $\|T\|^2 \in \sigma_{ap}(T^2)$. By combining Theorem 2.24 and Lemma 2.17, we then have that $\|T\| \in \sigma(T)$ or $-\|T\| \in \sigma(T)$. We note that

$$\left\|T^2 x - \lambda^2 x\right\|^2 = \left\|T^2 x\right\|^2 - 2\lambda^2 \|Tx\|^2 + \lambda^4 \|x\|$$

for all $\lambda \in \mathbb{R}$ and for all $x \in \mathcal{H}$. Let $(x_n)_n$ be a sequence of unit vectors such that $\|Tx_n\| \to \|T\|$, and let $\lambda = \|T\|$. Then

$$\left\|T^2 x_n - \lambda^2 x_n\right\|^2 \leq (\|T\| \cdot \|Tx_n\|)^2 - 2\lambda^2 \|Tx_n\|^2 + \lambda^4 = \lambda^4 - \lambda^2 \|Tx_n\|^2 \to 0.$$

Hence, $\|T\|^2 \in \sigma_{ap}(T^2)$. □

Corollary 2.26 *Let* $T \in B(\mathcal{H})_{sa}$ *and* p *be a real polynomial. Then*

$$\|p(T)\| = \sup \left\{ |p(\lambda)| \middle| \lambda \in \sigma(T) \right\}.$$

Proof. Since p is a real polynomial, then $p(T)$ is also self-adjoint. Hence applying Theorem 2.25 to $p(T)$ and using Theorem 2.24 we get

$$\begin{aligned} \|p(T)\| &= \sup \left\{ |\lambda| \middle| \lambda \in \sigma(p(T)) \right\} \\ &= \sup \left\{ |\lambda| \middle| \lambda \in p(\sigma(T)) \right\} \\ &= \sup \left\{ |p(\lambda)| \middle| \lambda \in \sigma(T) \right\}. \end{aligned}$$

□

3 Spectral theorem

Let us briefly look at the eigenvalues and eigenvectors of a self-adjoint matrix M on $\mathbb{C}^n$. Suppose we have an orthonormal set of eigenvectors $\{x_i | \, i \in \{1, \ldots, n\}\}$ of a self-adjoint matrix M where $Mx_i = \lambda_i x_i$. Define a projection operator by

$$\mathrm{proj}_u v = P_u v \stackrel{\text{def}}{=} \frac{\langle u, v \rangle}{\langle u, u \rangle} u,$$

which projects v on u. Further define the projection the normalised eigenvector x_i,

$$P_i v \stackrel{\text{def}}{=} P_{x_i} v = \langle x_i, v \rangle \, x_i.$$

The set of projection operator P_i satisfies,

$$\text{(i)} \qquad \begin{cases} P_i P_i = P_i, \\ P_i P_j = 0 \quad \text{for } i \neq j, \end{cases}$$

$$\text{(ii)} \qquad \sum_{i=1}^{n} P_i = I,$$

and we can express M as a sum of scaled projection operators by

$$\text{(iii)} \qquad M = \sum_{i=1}^{n} \lambda_i P_i.$$

Such an eigenvalue decomposition of a matrix is the main motivation and template when going into a general Hilbert space. In particular, the sum (iii) shall converge when $n \to \infty$ and the sum (ii), which states that the decomposition is complete, shall also hold.

3.1 Spectral theorem for self-adjoint operators

The goal is to prove the spectral theorem for normal operators. However it easier to prove the spectral theorem for bounded self-adjoint operators first and then use this theory for bounded normal operators. All theory about normal operators will be in Section 3.2.

Definition 3.1 Let (X, Σ) be a measurable space, where Σ is a specified σ-algebra of subsets of X. A *spectral measure* in X is a function E whose values are idempotent, self-adjoint operators (projections) on $\mathcal{H}$ such that $E(X) = I$ and such that $E(\cup_n M_n) = \sum_n E(M_n)$ when M_n are countably many disjoint sets in Σ.

Theorem 3.2 *Let E be a projection-valued function on the measurable space (X, Σ). If $E(X) = I$, and for each pair of vectors x and y, the complex-valued set function $\mu_{x,y}$ defined for every M in Σ by $\mu_{x,y}(M) = \langle E(M)x, y \rangle$ is countably additive, then E is a spectral measure.*

Remark 3.3 Before we state the proof of this theorem, there are two remarks. Firstly, $\mu_{x,y}(M)$ is a complex measure. Secondly Theorem 3.2 is an if and only if statement, so if E is a spectral measure then

(i) $E(X) = I$, and

(ii) for each pair of vectors x and y, the complex-valued set function μ defined for every M in Σ by $\mu(M) = \langle E(M)x, y\rangle$ is countably additive.

We will not prove the "only if" part since it is not necessary for the proof of the spectral theorem. Like most of the proofs in this paper, it is in Halmos [1].

Proof of 3.2. Let M and N be two disjoint measurable sets. Then we have

$$\langle E(M \cup N)x, y\rangle = \langle E(M)x, y\rangle + \langle E(N)x, y\rangle = \langle (E(M) + E(N))x, y\rangle .$$

Since this shows that $E(M \cup N) = E(M) + E(N)$, we get that E is finitely additive. Further, let $(M_n)_n$ is a disjoint sequence of measurable sets such that $\cup_n M_n = M$. Then multiplicativity of E implies that $(E(M_n))_n$ is an orthogonal sequence of projections and hence that $(E(M_n)x)_n$ is an orthogonal sequence of vectors for all $x \in X$. Since

$$\sum_n \|E(M_n)x\|^2 = \sum_n \langle E(M_n)x, x\rangle = \langle E(M)x, x\rangle = \|E(M)x\|^2 ,$$

we have that the sequence $(E(M_n)x)_n$ is summable. Now we have that

$$\langle E(M)x, y\rangle = \sum_n \langle E(M_n)x, y\rangle = \left\langle \left(\sum_n E(M_n)\right) x, y\right\rangle$$

for all x and y, and hence that $E(M) = \sum_n E(M_n)$. □

Before we state the big theorem which this subsection is named after, we need the following useful theorems which will be stated without proof.

Theorem 3.4 (Riesz-Markov representation theorem) *Let X be a compact separable space. To each bounded linear functional ϕ on $C(X)$, there exists a unique Borel measure μ on X such that*

$$\phi(f) = \int_X f(x)\mathrm{d}\mu(x)$$

for all $f \in C(X)$.

The proof of this theorem can be found in *Real and Complex analysis* by Rudin [6]. Note however that Rudin covers a more general variant where X is a locally compact Hausdorff space.

Theorem 3.5 (Stone-Weierstrass theorem) *Let X be a compact set and let $\mathcal{A}$ be an algebra of real (or complex) continuous functions on X. If $\mathcal{A}$ separates points on X (and $\mathcal{A}$ is self-adjoint) and contains constant functions, then the uniform closure of $\mathcal{A}$ is equal to the algebra of all continuous functions on X.*

While a lot of words need proper definitions to make sense of the theorem, we will simplify the theorem for our purposes. The theorem states that polynomials on a compact set X are dense in the set of all continuous function on $X \subseteq \mathbb{R}$ (or $X \subseteq \mathbb{C}$), which will be useful for us in the following theorems. The proof for this theorem for both the real and complex variants can be found in *Principles of Mathematical Analysis* by Rudin [5].

Further, we need some explanation to the notation used in the theorem. Let $E : X \to B(\mathcal{H})$ be a spectral measure. Then

$$\int_X \lambda \mathrm{d}E(\lambda)$$

is the unique operator $T \in B(\mathcal{H})$ satisfying

$$\langle Tx, y\rangle = \int_X \lambda \mathrm{d}\mu_{x,y}(\lambda)$$

for all $x, y \in \mathcal{H}$, where $\mu_{x,y}$ is the unique measure on X satisfying $\mu_{x,y}(M) = \langle E(M)x, y\rangle$ for all measurable subsets $M \subseteq X$.

One can show that such T always exists and that it is unique.

Theorem 3.6 (Spectral theorem for self-adjoint operators) *Let $T \in B(\mathcal{H})_{sa}$. Then there exists a spectral measure E on $\sigma(T)$, called the spectral measure of T, such that*

$$T = \int_{\sigma(T)} \lambda \mathrm{d}E(\lambda).$$

Proof. Let x and y be two fixed vectors and write

$$L(p) = \langle p(T)x, y\rangle$$

for every polynomial p. From Corollary 2.26 and by Cauchy-Schwarz inequality it follows that

$$|L(p)| \leq \|p(T)x\| \cdot \|y\| \leq \|p(T)\| \cdot \|x\| \cdot \|y\| = \sup\left\{|p(\lambda)|\Big|\lambda \in \sigma(T)\right\}$$

and hence that L extends to a unique bounded linear functional on $C(\sigma(T))$ by the Stone-Weierstrass theorem. By the Riesz-Markov representation theorem, there exists a unique complex measure $\mu_{x,y}$ on the compact set $\sigma(T)$ such that

$$\langle p(T)x, y\rangle = \int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{x,y}(\lambda)$$

for every real polynomial p. We also have that $|\mu_{x,y}(M)| \leq \|x\| \cdot \|y\|$ for every Borel set M.

Now we can prove with computations that the map $(x, y) \mapsto \mu_{x,y}(M)$ is a bounded symmetric sesquilinear form for each Borel set M. Given vectors x_1, x_2, and y and a polynomial function p, additivity follows from

$$\begin{aligned}\int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{x_1+x_2,y}(\lambda) &= \langle p(T)(x_1+x_2), y\rangle = \langle p(T)x_1, y\rangle + \langle p(T)x_2, y\rangle \\ &= \int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{x_1,y}(\lambda) + \int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{x_2,y}(\lambda) = \int_{\sigma(T)} p(\lambda)\mathrm{d}(\mu_{x_1,y}+\mu_{x_2,y})(\lambda).\end{aligned}$$

Given vectors x and y, a polynomial function p, and a complex number α, homogeneity follows from

$$\int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{\alpha x,y}(\lambda) = \langle p(T)\alpha x, y\rangle = \alpha \langle p(T)x, y\rangle = \alpha \int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{x,y}(\lambda).$$

Symmetricity follows from the fact that given a real polynomial function p, then $p(A)$ is self-adjoint. Hence, we get

$$\int_{\sigma(T)} p(\lambda)\mathrm{d}\mu_{y,x}(\lambda) = \langle p(T)y, x\rangle = \langle y, p(T)x\rangle = \overline{\langle p(T)x, y\rangle} = \int_{\sigma(T)} p(\lambda)\mathrm{d}\overline{\mu_{x,y}}(\lambda).$$

Since $\mu_{x,y}(M)$ is bounded, by the Fréchet-Riesz representation theorem and by Proposition 2.8 there exists a unique self-adjoint operator $E(M)$ such that $\mu_{x,y}(M) = \langle E(M)x, y\rangle$ for all x and y. Considering first the polynomial $p_0(\lambda) = 1$, this gives us

$$\langle E(X)x, y\rangle = \int_{\sigma(T)} \mathrm{d}\mu_{x,y}(\lambda) = \langle x, y\rangle$$

for all x and y, and hence $E(X) = I$. Considering the polynomial $p_1(\lambda) = \lambda$, this gives us

$$\int_{\sigma(T)} \lambda \mathrm{d}\mu_{x,y}(\lambda) = \langle Tx, y\rangle$$

for all x and y, which translates to

$$T = \int_{\sigma(T)} \lambda \mathrm{d}E(\lambda).$$

The only thing left to do is to show that E is actually a spectral measure. We do this by showing that E is projection-valued using Theorem 3.2, which is done by showing that E satisfies $E(M \cap N) = E(M) \cdot E(N)$. First define

$$\nu(M) = \int_M q(\lambda)\mathrm{d}\mu_{x,y}(\lambda)$$

for some fixed real polynomial q and for fixed x and y. Then for any real polynomial

p we have

$$\int p(\lambda)\mathrm{d}\nu(\lambda) = \int p(\lambda)q(\lambda)\mathrm{d}\mu_{x,y}(\lambda) = \int (p\cdot q)(\lambda)\mathrm{d}\mu_{x,y}(\lambda)$$
$$= \langle (p\cdot q)(T)x, y\rangle = \langle p(T)x, q(T)y\rangle = \int p(\lambda)\mathrm{d}\mu_{x,q(T)y}(\lambda).$$

So we get $\nu = \mu_{x,q(T)y}$. Further, applying this on $\nu(M)$ we get

$$\nu(M) = \mu_{x,q(T)y}(M) = \langle E(M)x, q(T)y\rangle = \langle q(T)E(M)x, y\rangle .$$

Since q is chosen arbitrarily, we find that

$$\int_M q(\lambda)\mathrm{d}\mu_{x,y}(\lambda) = \langle q(T)E(M)x, y\rangle = \int q(\lambda)\mathrm{d}\mu_{E(M)x,y}(\lambda)$$

and hence $\mu_{x,y}|_M = \mu_{E(M)x,y}$. Finally we need to show $E(M\cap N) = E(M)\cdot E(N)$ for all Borel sets M and N. Given M and N Borel sets, we have

$$\langle E(M\cap N)x, y\rangle = \mu_{x,y}(M\cap N) = \int \mathbb{1}_N(\lambda)\mathrm{d}\mu_{x,y}|_M(\lambda)$$
$$= \int \mathbb{1}_N(\lambda)\mathrm{d}\mu_{E(M)x,y}(\lambda) = \mu_{E(M)x,y}(N) = \langle E(N)E(M)x, y\rangle .$$

Hence, $E(M\cap N) = E(M)\cdot E(N)$ which shows that E is a spectral measure. $\square$

3.2 Spectral theorem for normal operators

Before talking about bounded normal operators, the following proposition will be useful for the definition.

Proposition 3.7 *Let $T \in B(\mathcal{H})$. Then there exist two unique self-adjoint operators S and R such that $T = S + iR$.*

Proof. Uniqueness of these operators is given by the fact that if $T = S + iR$ for $S, R \in B(\mathcal{H})_{sa}$, then $T^* = S^* - iR^* = S - iR$. Hence we get that

$$S = \frac{1}{2}(T + T^*) \qquad \text{and} \qquad R = \frac{1}{2i}(T - T^*).$$

Note that this also shows existence of S and R, so we are done. $\square$

In some sense, this gives an analogy related to complex numbers with $S = \operatorname{Re} T$ and $R = \operatorname{Im} T$. It is not often in this form $\operatorname{Re} T$ and $\operatorname{Im} T$ commute. However, the special case when they do commute gives the following definition and proposition. Note that we will denote commutativity by $[\cdot,\cdot] = 0$.

Definition 3.8 Let $T \in B(\mathcal{H})$. If $[T, T^*] = 0$, in other words $TT^* = T^*T$, then we say that T is *normal.*

Proposition 3.9 *Let $T \in B(\mathcal{H})$. Then T is normal if and only if* $[\operatorname{Re} S, \operatorname{Im} R] = 0$.

Proof. This is proven by a straight forward computation. Let $S = \operatorname{Re} T$ and $R = \operatorname{Im} T$. Then we have the following computation,

$$\begin{aligned} TT^* &= (S+iR)(S^*-iR^*) = (S+iR)(S-iR) = S^2 - iSR + iRS + R^2 \\ T^*T &= (S^*-iR^*)(S+iR) = (S-iR)(S+iR) = S^2 + iSR - iRS + R^2. \end{aligned}$$

If T is normal, i.e. $TT^* = T^*T$, then we get $iRS - iSR = iSR - iRS$ which implies $SR = RS$.

Conversely, if $[S, R] = 0$, then with the same computation we get that $TT^* = T^*T$. $\square$

Since bounded normal operators consists of two commuting self-adjoint operators, the idea is to use the two self-adjoint operators to show the spectral theorem for normal operators as well. Before stating the theorem and proving it, we need some additional theory.

Lemma 3.10 *Let $T, S \in B(\mathcal{H})_{sa}$ with spectral measure E_T and E_S respectively. If $[T, S] = 0$, then $[E_T(M), E_S(N)] = 0$ for all Borel sets M and N.*

Proof. First, let us show that $[E_T(M), S] = 0$ for all Borel sets M, which we are going to do by showing that $\langle SE_T x, y\rangle = \langle E_T S x, y\rangle$ for all x and y. Firstly we have

$$\langle SE_T(M)x, y\rangle = \langle E_T(M)x, Sy\rangle = \mu_{x,Sy}(M)$$

for all $x, y \in \mathcal{H}$. Secondly, we have

$$\langle E_T(M)Sx, y\rangle = \langle Sx, E_T(M)y\rangle = \overline{\langle E_T(M)y, Sx\rangle} = \overline{\mu_{y,Sx}}(M)$$

for all $x, y \in \mathcal{H}$. What is left to be shown is that $\mu_{x,Sy} = \overline{\mu_{y,Sx}}$. By the Stone-Weierstrass theorem it suffices to do this by checking on polynomials. If p is any (real) polynomial then

$$\int p(t)\mathrm{d}\overline{\mu_{y,Sx}}(t) = \overline{\langle p(T)y, Sx\rangle} = \langle Sx, p(T)y\rangle.$$

Since $[T, S] = 0$, then also $[p(T), S] = 0$ and hence

$$\langle Sx, p(T)y\rangle = \langle p(T)x, Sy\rangle = \int p(t)\mathrm{d}\mu_{Sx,y}(t).$$

Now that we know $[E_T(M), S] = 0$ for all Borel sets M, we need to show that $[E_T(M), E_S(N)] = 0$ for all Borel sets M and N. By fixing M, we apply the same proof to show that $[E_T(M), E_S(N)] = 0$. But since this is true for any Borel set M, we get that $[E_T(M), E_S(N)] = 0$ for all Borel sets M and N. $\square$

The next theorem and proposition will be stated without proof. The proof for these can be extracted from Halmos' proof of the spectral theorem for normal operators [1].

Theorem 3.11 *Let $E_1 : X_1 \to B(\mathcal{H})$ and $E_2 : X_2 \to B(\mathcal{H})$ be two spectral measures such that $[E_1(M_1), E_2(M_2)] = 0$ for all measurable sets $X_i \subset X, i \in \{1, 2\}$. Then there exists a unique spectral measure $E : X_1 \times X_2 \to B(\mathcal{H})$ such that $E(M_1 \times M_2) = E_1(M_1)E_2(M_2)$ for all measurable subsets $M_i \subset X_i, i \in \{1, 2\}$.*

Proposition 3.12 *Let $T, S \in B(\mathcal{H})_{sa}$ with spectral measures T_A and S_B such that $[T, S] = 0$. Then the product measure $E = E_T \times E_S$ satisfies*

$$\int p(s, t)\mathrm{d}E(s, t) = p(T, S)$$

for all complex polynomials in two variables p.

Theorem 3.13 (Spectral theorem for normal operators) *Let $T \in B(\mathcal{H})$ be a normal operator. Then there exists a spectral measure E on $\sigma(T)$ such that*

$$\int_{\sigma(T)} p(z, \overline{z})\mathrm{d}E(z) = p(T, T^*)$$

for all complex polynomials p in two variables.

Proof. Since $T = \operatorname{Re} T + i \operatorname{Im} T$ with $\operatorname{Re} T$ and $\operatorname{Im} T$ self-adjoint, we apply the spectral theorem for self-adjoint operators on these two to get the spectral measures E_1 and E_2 respectively. By Lemma 3.10, the two spectral measures commute, and hence we can construct the product measure $F = E_1 \times E_2$ with Theorem 3.11. This product measure satisfies

$$\int p(s, t)\mathrm{d}F(s, t) = p(\operatorname{Re} T, \operatorname{Im} T)$$

for all complex polynomials p in two variables.

Consider now the isomorphism $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ given by $\phi(s, t) = s + it$. Then $M \mapsto F(\phi^{-1}(M)) = E(M)$ defines a spectral measure on $\mathbb{C}$. For a polynomial p in two variables, by Proposition 3.12 it satisfies

$$\begin{aligned} \int p(z, \overline{z})\mathrm{d}E(z) &= \int p(s + it, s - it)\mathrm{d}F(s, t) \\ &= p(\operatorname{Re} T + i \operatorname{Im} T, \operatorname{Re} T - i \operatorname{Im} T) = p(T, T^*), \end{aligned}$$

since $(s, t) \mapsto p(s + it, s - it)$ is a polynomial in two variables. Note that we skip the proof that F is actually supported on $\sigma(T)$, but this is also part of the proof in Halmos. $\square$

4 The bilateral shift

Consider the Hilbert space $\ell^2(\mathbb{Z})$, the space of square-summable sequences indexed by $\mathbb{Z}$. The natural orthonormal basis of $\ell^2(\mathbb{Z})$ is given by

$$\delta_n(m) = \begin{cases} 1 \text{ if } n = m \\ 0 \text{ otherwise} \end{cases}$$

for $n, m \in \mathbb{Z}$.

One operator on $\mathcal{H} = \ell^2(\mathbb{Z})$ is the *bilateral shift* operator. The bilateral shift operator, which from now on will be denoted by S, takes one vector of the natural orthonormal basis δ_n and maps it to δ_{n+1}. In other words $S\delta_n = \delta_{n+1}$ for all $n \in \mathbb{Z}$.

Boundedness of the bilateral shift is verified as follows. Let

$$c = \sum_{n=-\infty}^{\infty} c_n \delta_n$$

be a finite sum where $(\delta_n)_n$ is the natural orthonormal basis. Then

$$\|Sc\|^2 = \left\| \sum_{n=-\infty}^{\infty} c_n S\delta_n \right\|^2 = \sum_{n=-\infty}^{\infty} \|c_n S\delta_n\|^2 = \sum_{n=-\infty}^{\infty} |c_n|^2 = \|c\|^2$$

The adjoint of S satisfies

$$\langle S^*\delta_n, \delta_m \rangle = \langle \delta_n, S\delta_m \rangle = \langle \delta_n, \delta_{m+1} \rangle = \begin{cases} 1 \text{ if } n - 1 = m \\ 0 \text{ otherwise} \end{cases} .$$

Hence S^* is the backward shift of S, in other words $S^* = S^{-1}$. We end up with $S^*S = I = SS^*$ which gives us that the shift operator S is normal. In fact, it is a special class of normal operators called *unitary* operators. Let us define properly what it means for a bounded operator to be unitary.

Definition 4.1 Let $T \in B(\mathcal{H})$ be a normal operator. If $TT^* = I = T^*T$, then we say that T is a *unitary* operator.

The bilateral shift is one example of an operator that has no eigenvalues. Indeed, let $\lambda \in \mathbb{C}$ and let $x \in \ell^2(\mathbb{Z})$ be a sequence such that $Sx = \lambda x$. Then we have that $x_n = \lambda^n x_0$ holds for $n \in \mathbb{Z}$, so

$$\|x\|^2 = \sum_{n=-\infty}^{\infty} x_n^2 = \sum_{n=-\infty}^{\infty} \lambda^{2n} x_0^2 = x_0^2 \left(\sum_{n=1}^{\infty} \lambda^{2n} + \sum_{n=0}^{-\infty} \lambda^{2n} \right).$$

The left sum diverges for $\lambda \geq 1$ while the right sum diverges for $\lambda \leq 1$. Hence we get that the only vector satisfying the equality $x_n = \lambda^n x_0$ is the zero vector, which cannot be an eigenvector.

The spectrum of the bilateral shift is non-empty, however. The next step will be characterising the spectrum of the bilateral shift.

Lemma 4.2 *Let $T \in B(\mathcal{H})$ be a unitary operator. Then $\sigma(T) \subseteq \mathrm{S}^1$.*

Proof. First we show that $\sigma(T) \subseteq \mathrm{S}^1$. The norm of T can be easily computed to 1, since

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle = \langle x, x \rangle = \|x\|^2 .$$

Hence by Theorem 2.23, $\sigma(T) \subseteq \mathbb{D}^2$. Note that $0 \notin \lambda(S)$ since T is invertible. Further, by Proposition 2.18 we have that $(\sigma(T))^{-1} = \sigma(T^{-1}) = \sigma(T^*)$. Since $\|T^*\| = 1$ as well, we find that $\sigma(T) \subseteq \mathrm{S}^1$.

Proposition 4.3 *The bilateral shift S satisfies $\sigma(S) = \mathrm{S}^1$.*

Since S is a unitary operator, by Lemma 4.2 we know $\sigma(S) \subseteq \mathrm{S}^1$, so we need to show the converse inclusion. Let $\lambda \in \mathrm{S}^1 \subseteq \mathbb{C}$ and define

$$x_n = \frac{1}{\sqrt{2n+1}} \sum_{i=-n}^{n} \lambda^i \delta_i.$$

Then we have that

$$\|x_n\|^2 = \frac{1}{2n+1} \sum_{i=-n}^{n} |\lambda^i|^2 = 1.$$

Further, we have that

$$Sx_n = \frac{1}{\sqrt{2n+1}} \sum_{i=-n}^{n} \lambda^i \delta_{i+1} \qquad \text{and} \qquad \lambda x_n = \frac{1}{\sqrt{2n+1}} \sum_{i=-n}^{n} \lambda^{i-1} \delta_i.$$

Hence, we have that

$$\|(S-\lambda)x_n\|^2 = \frac{1}{2n+1} \left(|\lambda^{-n-1}|^2 + |\lambda^n|^2 \right) = \frac{2}{2n+1} \xrightarrow{n\to\infty} 0.$$

Now assume for a contradiction that $(S - \lambda)$ is invertible. Let $T \in B(\mathcal{H})$ be the inverse of $(S-\lambda)$. If T were such an operator satisfying $T(S-\lambda) = I$, then we would have

$$1 = \|x_n\| = \|T(S-\lambda)x_n\| \leq \|T\| \cdot \|(S-\lambda)x_n\| \to 0,$$

which contradicts $T = (S-\lambda)^{-1}$. □

Theorem 4.4 *Let $T \in B(\mathcal{H})$ be a normal operator. Let E be the spectral measure for T satisfying for all two variable polynomials p that*

$$\langle p(T, T^*)x, x \rangle = \int_{\sigma(T)} p(z, \overline{z}) \mathrm{d}\mu_{x,x}(z)$$

where the unique complex measure $\mu_{x,x}$ is defined by

$$\mu_{x,x}(M) = \langle E(M)x, x \rangle .$$

Further let $K = \overline{\mathrm{span}}\{p(T,T^*)x \,|\, p\,\mathrm{polynomial}\}$*. Then the map* $U : K \to L^2(\sigma(T),\mu)$ *for* $\mu = \mu_{x,x}$ *satisfying* $p(T,T^*)x \mapsto p(z,\overline{z})$ *is a well-defined unitary such that*

$$UTU^* = m$$

where $m : L^2(\sigma(T),\mu) \to L^2(\sigma(T),\mu)$ *is the multiplication operator satisfying* $(m\,f)(z) = zf(z)$ *for all* $f \in C(\sigma(T))$*.*

Proof. The fact that U is a well-defined unitary follows from

$$\begin{aligned}
&\langle U\left(p(T,T^*)x\right), U\left(q(T,T^*)x\right)\rangle \\
&\quad = \langle p(z,\overline{z}), q(z,\overline{z})\rangle \\
&\quad = \int_{\sigma(T)} (p\cdot\overline{q})(z,\overline{z})\mathrm{d}\mu(z) \\
&\quad = \langle p\cdot\overline{q}(T,T^*)x, x\rangle \\
&\quad = \langle p(T,T^*)x, q(T,T^*)x\rangle\,.
\end{aligned}$$

What is left to show is that $UTU^* = m$, or equivalently $UT = mU$, which can checked on $p(T,T^*)x$ where p runs through complex polynomials. Fix such p and let $q(a,b) \overset{\mathrm{def}}{=} a\cdot p(a,b)$. Then we get

$$UTp(T,T^*)x = Uq(T,T^*)x = q(z,\overline{z}) = z\cdot p(z,\overline{z}) = m\cdot Up(T,T^*)x$$

in $L^2(\sigma(T),\mu)$. $\square$

The next example can be put in a more general theory, known as Fourier theory. Fourier theory provides a link between $\mathbb{Z}$ and S^1 as two groups which are dual to each other. In fact, the example can be identified as a special case of the Plancherel theorem [7], which provides a natural isomorphism of L^2-spaces over two groups that are dual to each other. We will not go in-depth into either Fourier theory or Plancherel theorem however, and will only apply Theorem 4.4 to the bilateral shift and view the results.

Example 4.5 Let $S \in B(\ell^2(\mathbb{Z}))$ be the bilateral shift operator. From Proposition 4.3, we know that $\sigma(S) = \mathrm{S}^1$. Further, let E be the spectral measure for S, let $x = \delta_0$, and

$$\mu(M) \overset{\mathrm{def}}{=} \langle E(M)x, x\rangle\,.$$

Then we have that

$$\begin{aligned}
K &= \overline{\mathrm{span}}\{p(S,S^*) \,|\, p\,\text{complex polynomial}\} \\
&= \overline{\mathrm{span}}\{\delta_n \in \ell^2(\mathbb{Z}) | n \in \mathbb{Z}\} = \ell^2(\mathbb{Z}).
\end{aligned}$$

Hence $U : \ell^2(\mathbb{Z}) \xrightarrow{\cong} L^2(\mathrm{S}^1,\mu)$ intertwines S and the multiplication operator as

$$USU^* = m_{\mathrm{S}^1}.$$

What is left to do is to identify the measure μ. We will show that μ is the unique measure satisfying

$$\int_{\mathrm{S}^1} p(z,\overline{z})\mathrm{d}\mu(z) = \frac{1}{2\pi}\int_0^{2\pi} p(\cos t + i\sin t, \cos t - i\sin t)\mathrm{d}t.$$

Note that it suffices to check monomials $p(z,\overline{z}) = z^n$ for $n \in \mathbb{Z}$. Further, we note that E is defined by

$$\langle p(S,S^*)x, x\rangle = \int_{\mathrm{S}^1} p(z,\overline{z})\mathrm{d}\mu(z)$$

with $\langle p(S,S^*)x, x\rangle = \langle p(S,S^*)\delta_0, \delta_0\rangle$. Since S is unitary, we only have to check the values

$$\begin{aligned}
&\langle S^n\delta_0, \delta_0\rangle = \langle \delta_n, \delta_0\rangle = \delta_{n,0}\\
&\langle (S^*)^n\delta_0, \delta_0\rangle = \delta_{n,0}.
\end{aligned}$$

We are hence looking for the probability measure μ on S^1 satisfying

$$\int_{\mathrm{S}^1} z^n\mathrm{d}\mu(z) = \delta_{n,0} = \int_{\mathrm{S}^1} \overline{z}^{\,n}\mathrm{d}\mu(z)$$

for all $n \in \mathbb{Z}$. We identify $\mathrm{S}^1 = [0, 2\pi)$, so the normalised Lebesgue integral can be calculated by Riemann integration using the formula

$$\int_{\mathrm{S}^1} f(z)\,\mathrm{d}\mu(z) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\mathrm{d}t,$$

which holds for all continuous functions $f : \mathrm{S}^1 \to \mathbb{C}$. Since z corresponds to the map $t \mapsto e^{it}$ and hence $t \mapsto (\cos t, \sin t)$, in general z^n corresponds to $(\cos nt, \sin nt)$. Since complex conjugation is compatible with integration for Riemann integrals, it suffices to show

$$\frac{1}{2\pi}\int_0^{2\pi} \cos nt\,\mathrm{d}t = 0 = \frac{1}{2\pi}\int_0^{2\pi} \sin nt\,\mathrm{d}t$$

for all non-zero integers n. Note that if $n = 0$, then indeed we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} 1\,\mathrm{d}t = 1.$$

If $n \neq 0$, both the integrals are 0 since the anti-derivatives of $\cos nt$ and $\sin nt$ are

$$\frac{\sin nt}{n} \qquad \text{and} \qquad \frac{-\cos nt}{n}$$

respectively. Hence, we find that μ is the Lebesgue measure on S^1.

References

[1] Paul R. Halmos. *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*. Chelsea, 1951.

[2] John B. Conway. *A course in Functional Analysis*. Springer, 1990.

[3] Brian C. Hall. *Quantum theory for Mathematicians*. Springer, 2013.

[4] Wikipedia contributors. Cauchy-Schwarz inequality — Wikipedia, the free encyclopedia, 2019. [Online; accessed 10-January-2020].

[5] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 1976.

[6] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, 1987.

[7] Kosaku Yosida. *Functional Analysis*. Springer-Verlag, 1980.