Texts and Readings in Mathematics 71

V.S. Sunder

Operators on Hilbert Space

HINDUSTAN BOOK AGENCY

Springer

Texts and Readings in Mathematics

Volume 71

The *Texts and Readings in Mathematics* series publishes high-quality textbooks, research-level monographs, lecture notes and contributed volumes. Undergraduate and graduate students of mathematics, research scholars, and teachers would find this book series useful. The volumes are carefully written as teaching aids and highlight characteristic features of the theory. The books in this series are co-published with Hindustan Book Agency, New Delhi, India.

More information about this series at http://www.springer.com/series/15141

V.S. Sunder

Operators on Hilbert Space

HINDUSTAN
BOOK AGENCY

Springer

V.S. Sunder
Department of Mathematics
Institute of Mathematical Sciences
Chennai, Tamil Nadu
India

This work is a co-publication with Hindustan Book Agency, New Delhi, licensed for sale in all countries in electronic form only. Sold and distributed in print across the world by Hindustan Book Agency, P-19 Green Park Extension, New Delhi 110016, India. ISBN: 978-93-80250-74-8

ISSN 2366-8725 (electronic)
Texts and Readings in Mathematics
ISBN 978-981-10-1816-9 (eBook)
DOI 10.1007/978-981-10-1816-9

Library of Congress Control Number: 2016943790

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Preface

This book was born out of a desire to have a *brief* introduction to operator theory – the spectral theorem (arguably the most important theorem in Hilbert space theory), polar decomposition, compact operators, trace-class operators, etc., which would involve a minimum of initial spadework (avoiding such digressions as, for example, the Gelfand theory of commutative Banach algebras), and which only needed simple facts from a first semester graduate course on Functional Analysis. I believe the cleanest formulation of the spectral theorem is as a statement of the existence and uniqueness of appropriately homeomorphic (continuous and measurable) functional calculi of one or more pairwise commuting self-adjoint, and more generally normal, operators on a separable Hilbert space, rather than one about spectral measures.

This book may be thought of as a re-take of my earlier book ([Sun]) on Functional Analysis, but with so many variations as to not really look like a 'second edition': the operator algebraic point of view is minimised drastically, resulting in an *essentially operator-theoretic* proof of the spectral theorem – first for self-adjoint, and later for normal, operators. What is probably new here is what I call the joint spectrum of a family of commuting self-adjoint operators, a new proof of the Fuglede theorem on the commutant of a normal operator being *-closed, and the extension of the spectral theorem to a family of commuting normal operators. The third chapter contains, in addition to everything in the fourth chapter of [Sun], a section about Hilbert-Schmidt and trace-class operators, and the duality results involving compact operators, trace-class operators and all bounded operators.

This book is fondly dedicated to the memory of Paul Halmos.

Acknowledgements

I wish to record my appreciation of the positive encouragement of Rajendra Bhatia (the Managing Editor of the series in which [Sun] appeared) to consider coming up with a second edition but with enough work put in rather than a sloppy cut-and-paste mish-mash. Even though I have lifted fairly large chunks from [Sun], I believe there is enough new material here to merit this book having a different name rather than be thought of as the second edition of the older book. I will be remiss if I did not record my gratitude to (i) one of the referees who displayed tremendous patience in wading through the manuscript in spite of instances of my sloppiness being everywhere dense in it, and painstakingly prepared a report that pointed out many howlers which were fortunately caught before this appeared in print, and (ii) my long lasting colleagues and friends, Kesavan and Vijay, for having kindly allowed me to subject them to proof-reading stints with my typo-prone manuscript.

Finally, it is a pleasure to thank the Institute of Mathematical Sciences for the wonderful infrastructural facilities and for the congenial atmosphere it has provided me all these years. I must also thank the DST (and its former Secretary Dr. Ramasami, in particular) for interpreting their rules to permit my use of the JC Bose Fellowship in a wonderful example of 'reasonable accommodation' which allowed me to continue being a 'working mathematician' in spite of my not-so-great health.

Contents

About the Author

V.S. Sunder is professor of mathematics at the Institute of Mathematical Sciences (commonly known as MATSCIENCE), Chennai, India. He specialises in subfactors, operator algebras and functional analysis in general. In 1996, he was awarded the Shanti Swarup Bhatnagar Prize for Science and Technology, the highest science award in India, in the mathematical sciences category. He is one of the first Indian operator algebraists. In addition to publishing over 60 papers, he has written six books including at least three monographs at the graduate level or higher on von Neumann algebras. One of the books was co-authored with Vaughan Jones, an operator algebraist, who has received the Fields Medal.

Chapter 1

Hilbert space

1.1 Introduction

This book is about (bounded, linear) operators on (**always separable and complex**) Hilbert spaces, usually denoted by $\mathcal{H}, \mathcal{K}, \mathcal{M}$ and variants thereof. Vectors in Hilbert spaces will usually be denoted by symbols such as x, y, z and their variants, such as y_n, x'. The collection of all bounded complex-linear operators on $\mathcal{H}$ will be denoted by $B(\mathcal{H})$, whose elements will usually be denoted by symbols such as $A, B, E, F, P, Q, T, U, V, X, Y, Z$.

The only prerequisites needed for reading this book are: a nodding acquaintance with the basics of Hilbert space theory (eg: the definitions of orthonormal basis, orthogonal projection, unitary operator, etc., all of which are briefly discussed in this chapter); a first course in Functional Analysis – the spectral radius formula, the Open Mapping Theorem and the Uniform Boundedness Principle, the Riesz Representation Theorem (briefly mentioned in Appendix) which identifies $C(\Sigma)^*$ with the space $M(\Sigma)$ of finite complex measures on the compact Hausdorff space Σ, and outer and inner regularity of finite positive measures on Σ; some basic measure theory, such as the Bounded Convergence Theorem, and the not so basic Lusin's theorem (also briefly discussed in Appendix) which leads to the conclusion – see Lemma A2 in the Appendix – that any bounded measurable function on Σ is the pointwise a.e. limit of a sequence of continuous functions on Σ, and also – see Lemma A1 in the Appendix – that $C(\Sigma)$ 'is' dense in $L^2(\Sigma, \mu)$. Also, in the section on **von Neumann-Schatten ideals**, basic facts concerning the Banach sequence spaces c_0, ℓ^p and the duality relations among them will be needed/used. All the above facts may be found in [Hal], [Hal1], [Sun] and [AthSun]. Although these standard facts may also be found in other classical texts written by distinguished mathematicians, the references are limited to a very small number of books, because the author knows precisely where which fact can be found in the union of the four books mentioned above.

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9_1

1.2 Inner Product spaces

While normed spaces permit us to study 'geometry of vector spaces', we are constrained to discussing those aspects which depend only upon the notion of 'distance between two points'. If we wish to discuss notions that depend upon the angles between two lines, we need something more – and that something more is the notion of an *inner product.*

The basic notion is best illustrated in the example of the space $\mathbb{R}^2$ that we are most familiar with, where the most natural norm is what we call $||\cdot||_2$. The basic fact from plane geometry that we need is the so-called *cosine law* which states that if A, B, C are the vertices of a triangle and if θ is the angle at the vertex C, then

$$2(AC)(BC)\cos\theta \;=\; (AC)^2 + (BC)^2 - (AB)^2 \;.$$

If we apply this to the case where the points A, B and C are represented by the vectors $x = (x_1, x_2),\ y = (y_1, y_2)$ and $(0, 0)$ respectively, we find that

$$\begin{aligned} 2||x||\cdot||y||\cdot\cos\theta &= ||x||^2 + ||y||^2 - ||x-y||^2 \\ &= 2\ (x_1 y_1 \ + \ x_2 y_2\ ). \end{aligned}$$

Thus, we find that the function of two (vector) variables given by

$$\langle x, y\rangle = x_1 y_1 \ + \ x_2 y_2 \tag{1.2.1}$$

simultaneously encodes the notion of angle as well as distance (and has the explicit interpretation $\langle x, y\rangle \;=\; ||x||\;||y||\cos\theta$). This is because the norm can be recovered from the inner product by the equation

$$||x|| \;=\; \langle x, x\rangle^{\frac{1}{2}} \;. \tag{1.2.2}$$

The notion of an inner product is the proper abstraction of this function of two variables.

DEFINITION 1.2.1. *(a) An* **inner product** *on a (complex) vector space V is a mapping $V\times V \ni (x, y) \;\mapsto\; \langle x, y\rangle \in \mathbb{C}$ which satisfies the following conditions, for all $x, y, z \in V$ and $\alpha \in \mathbb{C}$:*
(i) (positive definiteness) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0 \;\Leftrightarrow\; x = 0$;
(ii) (Hermitian symmetry) $\langle x, y\rangle \;=\; \overline{\langle y, x\rangle}$;
(iii) (linearity in first variable) $\langle \alpha x + \beta z, y\rangle \;=\; \alpha\langle x, y\rangle + \beta\langle z, y\rangle$.

An **inner product space** *is a vector space equipped with a (distinguished) inner product.*

(b) An inner product space which is complete in the norm coming from the inner product (as in Equation (1.2.2)*) is called a* **Hilbert space***. In this book, however, we shall only be concerned with Hilbert spaces which are separable when viewed as metric spaces, with the metric coming from the norm induced by the inner-product – see Proposition* 1.2.4 *and Corollary* 1.2.5.

EXAMPLE 1.2.2. (1) If $z = (z_1, \ldots, z_n), w = (w_1, \ldots, w_n) \in \mathbb{C}^n$, define

$$\langle z, w \rangle \;=\; \sum_{i=1}^{n} z_i \overline{w}_i \;; \tag{1.2.3}$$

it is easily verified that this defines an inner product on $\mathbb{C}^n$.

(2) The equation

$$\langle f, g \rangle \;=\; \int_{[0,1]} f(x)\overline{g(x)}\; dx \tag{1.2.4}$$

is easily verified to define an inner product on $C[0,1]$. □

As in the (real) case discussed earlier of $\mathbb{R}^2$, it is generally true that any inner product gives rise to a norm on the underlying space via equation (1.2.2). Before verifying this fact, we digress with an exercise that states some easy consequences of the definitions.

EXERCISE 1.2.3. *Suppose we are given an inner product space V; for $x \in V$, define $||x||$ as in equation* (1.2.2)*, and verify the following identities, for all $x, y, z \in V,\ \alpha \in \mathbb{C}$:*

(1) $\langle x, y + \alpha z \rangle \;=\; \langle x, y \rangle \;+\; \overline{\alpha}\langle x, z \rangle$;

(2) $||x + y||^2 \;=\; ||x||^2 \;+\; ||y||^2 \;+\; 2\,\mathrm{Re}\,\langle x, y \rangle$;

(3) *two vectors in an inner product space are said to be* **orthogonal** *if their inner product is* 0*; deduce from* (2) *above and an easy induction argument that if $\{x_1, x_2, \ldots, x_n\}$ is a set of pairwise orthogonal vectors, then* $||\sum_{i=1}^n x_i||^2 \;=\; \sum_{i=1}^n ||x_i||^2$.

(4) $||x + y||^2 \;+\; ||x - y||^2 \;=\; 2\;(||x||^2 + ||y||^2)$; *draw some diagrams and convince yourself as to why this identity is called the* **parallelogram identity**.

(5) *(Polarisation identity)* $4\langle x, y \rangle \;=\; \sum_{k=0}^3 \; i^k \langle x + i^k y, x + i^k y \rangle$*, where, of course,* $i = \sqrt{-1}$.

The first (and very important) step towards establishing that any inner product defines a norm via equation (1.2.2) is the following celebrated inequality.

PROPOSITION 1.2.4. (**Cauchy-Schwarz inequality**)

If x, y are arbitrary vectors in an inner product space V, then

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

Further, this inequality is an equality if and only if the vectors x and y are linearly dependent.

Proof. If $y = 0$, there is nothing to prove; so we may, without loss of generality, assume that $||y|| = 1$ (since the statement of the proposition is unaffected upon scaling y by a constant).

Notice now that, for arbitrary $\alpha \in \mathbb{C}$,

$$\begin{aligned} 0 &\leq ||x - \alpha y||^2 \\ &= ||x||^2 + |\alpha|^2 - 2\,\mathrm{Re}\,(\alpha\langle y, x\rangle)\ . \end{aligned}$$

A little exercise in the calculus shows that this last expression is minimised for the choice $\alpha_0 = \langle x, y\rangle$, for which choice we find, after some minor algebra, that

$$0 \ \leq\ ||x - \alpha_0 y||^2 \ = \ ||x||^2 - |\langle x, y\rangle|^2\ ,$$

thereby establishing the desired inequality.

The above reasoning shows that the inequality becomes an equality only if $x = \alpha_0 y$, and the proof is complete. □

COROLLARY 1.2.5. *Any inner product gives rise to a norm*[1] *via Equation* (1.2.2).

Proof. Positive-definiteness and homogeneity with respect to scalar multiplication are obvious; as for the triangle inequality,

$$\begin{aligned} ||x + y||^2 &= ||x||^2 \ + \ ||y||^2 \ + \ 2\ \mathrm{Re}\ \langle x, y\rangle \\ &\leq ||x||^2 \ + \ ||y||^2 \ + \ 2||x|| \cdot ||y||\ , \end{aligned}$$

and the proof is complete. □

EXERCISE 1.2.6. (1) *Show that*

$$\Big|\sum_{i=1}^{n} z_i\overline{w_i}\Big|^2 \ \leq\ \Big(\sum_{i=1}^{n} |z_i|^2\Big)\ \Big(\sum_{i=1}^{n} |w_i|^2\Big)\ ,\ \forall\ z, w \in \mathbb{C}^n\ .$$

(In view of the notation used in (2) *below, we shall write* ℓ^2_n *for* $\mathbb{C}^n$ *with the 'standard inner product' defined above.)*

(2) *Deduce from* (1) *that the series* $\sum_{i=1}^{\infty} \alpha_i\overline{\beta_i}$ *converges, for any* $\alpha, \beta \in \ell^2 = \{\gamma = (\gamma_1, \ldots, \gamma_n, \ldots) \in \mathbb{C}^{\mathbb{N}} : \sum_n |\gamma_n|^2 < \infty\}$, *and that*

$$\Big|\sum_{i=1}^{\infty} \alpha_i\overline{\beta_i}\Big|^2 \ \leq\ \Big(\sum_{i=1}^{\infty} |\alpha_i|^2\Big)\ \Big(\sum_{i=1}^{\infty} |\beta_i|^2\Big)\ ,\ \forall\ \alpha, \beta \in \ell^2\ ;$$

deduce that ℓ^2 *is indeed (a vector space, and in fact) an inner product space, with respect to inner product defined by*

$$\langle \alpha, \beta\rangle \ = \ \sum_{i=1}^{\infty} \alpha_i\overline{\beta}_i\ . \qquad (1.2.5)$$

[1]Recall that (a) a norm on a vector space V is a function $V \ni x \mapsto \|x\| \in [0, \infty)$ which satisfies (i) (*positive-definiteness*) $\|x\| = 0 \Leftrightarrow x = 0$; (ii) (*homogeneity*) $\|\alpha x\| = |\alpha|\|x\|$ and (iii) (*triangle inequality*) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ and $\alpha \in \mathbb{C}$; (b) a vector space equipped with a norm is a **normed space**; and (c) a normed space which is complete with respect to the norm is called a **Banach space**.

(3) *Write down what the Cauchy-Schwarz inequality translates into in Example* 1.2.2 (2).

(4) *Show that the inner product is continuous as a mapping from* $V \times V$ *into* $\mathbb{C}$. *(In view of Corollary* 1.2.5, *this makes sense.)*

1.3 Hilbert spaces : examples

Our first step is to arm ourselves with a reasonably adequate supply of examples of Hilbert spaces.

EXAMPLE 1.3.1. (1) $\mathbb{C}^n$ is an example of a finite-dimensional Hilbert space, and we shall soon see that these are essentially the only such examples.

(2) ℓ^2 is an infinite-dimensional Hilbert space – see Exercise 1.2.6(2). Nevertheless, this Hilbert space is not 'too big', since it is at least equipped with the pleasant feature of being a **separable** Hilbert space – i.e., it is separable as a metric space, meaning that it has a countable dense set. (Verify this assertion!)

(3) More generally, let S be an arbitrary set, and define

$$\ell^2(S) \;=\; \{x \;=\; ((x_s))_{s \in S} \in \mathbb{C}^S : \textstyle\sum_{s \in S} |x_s|^2 < \infty\} \;.$$

(The possibly uncountable sum might be interpreted as follows: a typical element of $\ell^2(S)$ is a family $x = ((x_s))$ of complex numbers which is indexed by the set S, and which has the property that $x_s = 0$ except for s coming from some countable subset of S (which depends on the element x) and which is such that the possibly non-zero x_s's, when written out as a sequence in any (equivalently, some) way, constitute a norm-square-summable sequence.)

Verify that $\ell^2(S)$, in a natural fashion, is a Hilbert space.

(4) This example will make sense to the reader who is already familiar with the theory of measure and Lebesgue integration; the reader who is not, may safely skip this example; the subsequent exercise will effectively recapture this example, at least in all cases of interest.

Suppose $(X, \mathcal{B}, \mu)$ is a measure space. Let $\mathcal{L}^2(X, \mathcal{B}, \mu)$ denote the space of $\mathcal{B}$-measurable complex-valued functions f on X such that $\int_X |f|^2 d\mu < \infty$. Note that $|f+g|^2 \le 2(|f|^2 + |g|^2)$, and deduce that $\mathcal{L}^2(X, \mathcal{B}, \mu)$ is a vector space. Note next that $|f\overline{g}| \le \frac{1}{2}(|f|^2 + |g|^2)$, and so the right-hand side of the following equation makes sense, if $f, g \in \mathcal{L}^2(X, \mathcal{B}, \mu)$:

$$\langle f, g \rangle \;=\; \int_X f \overline{g} \, d\mu \;. \tag{1.3.6}$$

It is easily verified that the above equation satisfies all the requirements of an inner product with the solitary possible exception of the positive-definiteness axiom: if $\langle f, f \rangle = 0$, it can only be concluded that $f = 0$ *a.e* – meaning that $\{x : f(x) \neq 0\}$ is a set of μ-measure 0 (which might very well be non-empty).

Observe, however, that the set $N = \{f \in \mathcal{L}^2(X, \mathcal{B}, \mu) : f = 0 \ a.e.\}$ is a vector subspace of $\mathcal{L}^2(X, \mathcal{B}, \mu)$; and a typical element of the quotient space $L^2(X, \mathcal{B}, \mu) = \mathcal{L}^2(X, \mathcal{B}, \mu)/N$ is just an equivalence class of square-integrable functions, where two functions are considered to be equivalent if they agree outside a set of μ-measure 0.

For simplicity of notation, we shall just write $L^2(X)$ or $L^2(\mu)$ for $L^2(X, \mathcal{B}, \mu)$, and we shall denote an element of $L^2(\mu)$ simply by such symbols as f, g, etc., and think of these as actual functions with the understanding that we shall identify two functions which agree μ-almost everywhere. The point of this exercise is that equation (1.3.6) now does define a genuine inner product on $L^2(X)$; most importantly, it is true that $L^2(X)$ is complete and is thus a Hilbert space. □

EXERCISE 1.3.2. (1) *Suppose X is an inner product space. Let $\overline{X}$ be a completion of X regarded as a normed space. Show that $\overline{X}$ is actually a Hilbert space. (Thus, every inner product space has a Hilbert space completion.)*

(2) *Let $X = C[0,1]$ and define*

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\, dx .$$

Verify that this defines a genuine, i.e., positive-definite, inner product on $C[0,1]$. The completion of this inner product space is a Hilbert space – see (1) *above – which may be identified with what was called $L^2([0,1], \mathcal{B}, m)$ in Example* 1.3.1(4)*, where ($\mathcal{B}$ is the σ-algebra of Borel sets in $[0,1]$ and) m denotes the so-called Lebesgue measure on $[0,1]$.*

1.4 Orthonormal bases

In the sequel, N will always denote a (possibly empty, finite or infinite) countable set.

DEFINITION 1.4.1. *A collection $\{x_n : n \in N\}$ in an inner product space is said to be* **orthonormal** *if*

$$\langle x_m, x_n \rangle = \delta_{mn} := \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \quad \forall\, m, n \in N .$$

Thus, an orthonormal set is nothing but a set of unit vectors which are pairwise orthogonal; we shall write $x \perp y$ if two vectors x, y in an inner product space are orthogonal, i.e., satisfy $\langle x, y \rangle = 0$.

EXAMPLE 1.4.2. (1) In ℓ^2_n, for $1 \leq i \leq n$, let e_i be the element whose i-th co-ordinate is 1 and all other co-ordinates are 0; then $\{e_1, \ldots, e_n\}$ is an orthonormal set in ℓ^2_n.

(2) In ℓ^2, let e_n be the element whose n-th co-ordinate is 1 and all other co-ordinates are 0, for $1 \leq n < \infty$; then $\{e_n : n = 1, 2, \ldots\}$ is an orthonormal set in ℓ^2.

(3) In the inner product space $C[0,1]$ – with inner product as described in Exercise 1.3.2 – consider the family $\{e_n : n \in \mathbb{Z}\}$ defined by $e_n(x) = \exp(2\pi i n x)$, and show that this is an orthonormal set; hence this is also an orthonormal set when regarded as a subset of $L^2([0,1], m)$ – see Exercise 1.3.2(2). $\square$

PROPOSITION 1.4.3. *Let $\{e_1, e_2, \ldots, e_n\}$ be an orthonormal set in an inner product space X, and let $x \in X$ be arbitrary. Then,*

(i) if $x = \sum_{i=1}^n \alpha_i e_i, \ \alpha_i \in \mathbb{C}$, then $\alpha_i = \langle x, e_i \rangle \ \forall i$;

(ii) $(x - \sum_{i=1}^n \langle x, e_i \rangle e_i\ ) \perp e_j \ \forall 1 \leq j \leq n$;

(iii) (**Bessel's inequality**) $\sum_{i=1}^n |\langle x, e_i \rangle|^2 \leq ||x||^2$.

Proof. (i) If x is a linear combination of the e_j's as indicated, compute $\langle x, e_i \rangle$, and use the assumed orthonormality of the e_j's, to deduce that $\alpha_i = \langle x, e_i \rangle$.

(ii) This is an immediate consequence of (i).

(iii) Write $y = \sum_{i=1}^n \langle x, e_i \rangle e_i, \ z = x - y$, and deduce from (two applications of) Exercise 1.2.3(3) that

$$\begin{aligned} ||x||^2 &= ||y||^2 + ||z||^2 \\ &\geq ||y||^2 \\ &= \sum_{i=1}^n |\langle x, e_i \rangle|^2 \ . \end{aligned}$$

$\square$

We wish to remark on a few consequences of this proposition; for one thing, (i) implies that an arbitrary orthonormal set is linearly independent; for another, if we write $\bigvee\{e_n : n \in N\}$ for the vector subspace *spanned* by $\{e_n : n \in N\}$ – this is the set of linear combinations of the e_n's, and is the smallest vector subspace containing $\{e_n : n \in N\}$ – it follows from (i) that we know how to write any element of $\bigvee\{e_n : n \in N\}$ as a linear combination of the e_n's.

We shall find the following notation convenient in the sequel: if $\mathcal{S}$ is a subset of an inner product space X, let $\bigvee S$ (reps., $[\mathcal{S}]$) denote the smallest subspace (resp. closed subspace) containing $\mathcal{S}$; it should be clear that this could be described in either of the following equivalent ways: (a) $[\mathcal{S}]$ is the intersection of all closed subspaces of X which contain $\mathcal{S}$, and (b) $[\mathcal{S}] = \overline{\bigvee \mathcal{S}}$. (Verify that (a) and (b) describe the same set.)

LEMMA 1.4.4. *Suppose $\{e_n : n \in N\}$ is a countable orthonormal set in a Hilbert space $\mathcal{H}$. Then the following conditions on an arbitrary family $\{\alpha_n : n \in N\}$ of complex numbers are equivalent:*

(i) the sum $\sum_{n \in N} \alpha_n e_n$ makes sense as a finite sum in case N is a finite set, and as an 'unconditionally' norm-convergent series in $\mathcal{H}$ if N is

infinite, meaning: if $\phi : \mathbb{N} \to N$ is any bijection, and if we define $x(\phi)_k = \sum_{n=1}^{k} \alpha_{\phi(n)} e_{\phi(n)}$, then the sequence $\{x(\phi)_k : k \in \mathbb{N}\}$ is norm-convergent and the limit of this sum is independent of the bijection ϕ used; the symbol of a sum of elements of a Hilbert space, and $\mathbb{C}$, in particular, which is indexed by arbitrary countably infinite sets (other than $\mathbb{N}$, when, of course, series may be conditionally convergent) will always be used to denote only such 'unconditionally convergent series'.

(ii) $\sum_{n\in N} |\alpha_n|^2 < \infty$.

(iii) there is a vector $x \in [\{e_n : n \in N\}]$ *such that* $\langle x, e_n\rangle = \alpha_n \ \forall n \in N$.

Proof. If N is finite, the first two assertions are obvious, while the third is seen by choosing $x = \sum_{n\in N} \alpha_n e_n$.

So suppose that N is infinite, and that $\phi, x(\phi)_k$ are as above.

$(i) \Rightarrow (iii)$: Fix a bijection ϕ as in (i). Condition (i) says that $\|x(\phi)_k - x(\phi)\| \to 0$ for some $x(\phi) \in \mathcal{H}$. As $\langle x(\phi)_k, e_n\rangle = \langle x(\phi)_\ell, e_n\rangle = \alpha_n \forall k, \ell \geq \phi^{-1}(n)$, we find that $\langle x(\phi), e_n\rangle = \alpha_n \ \forall n$. Since each $x(\phi)_k \in [\{e_n : n \in N\}]$, it is clear that also $x(\phi) \in [\{e_n : n \in N\}]$.

$(iii) \Rightarrow (ii)$ is an immediate consequence of Bessel's inequality.

$(ii) \Rightarrow (i)$: Condition (ii) is seen to imply that $\{x(\phi)_k : k \in \mathbb{N}\}$ is a Cauchy sequence and hence convergent in $\mathcal{H}$. The argument given in the proof of $(i) \Rightarrow (iii)$ applies with ϕ replaced by any other bijection ψ. And $x(\phi) - x(\psi)$ would be an element of $[\{e_n : n \in N\}]$ which would be orthogonal to each e_n and hence to a dense subspce of $[\{e_n : n \in N\}]$, thereby forcing the equality $x(\phi) = x(\psi)$, as asserted. $\square$

We are now ready to establish the fundamental proposition concerning *orthonormal bases* in a Hilbert space.

PROPOSITION 1.4.5. *The following conditions on a countable orthonormal set $\{e_n : n \in N\}$ in a Hilbert space $\mathcal{H}$ are equivalent: (in items (ii), (iii) and (iv), the sums indexed by the set N are to be understood as indicated in Lemma* 1.4.4(i)).

(i) $\{e_n : n \in N\}$ is a maximal orthonormal set, meaning that it is not strictly contained in any other orthonormal set;

(ii) $x \in \mathcal{H} \Rightarrow x = \sum_{n\in N} \langle x, e_n\rangle e_n$;

(iii) $x, y \in \mathcal{H} \Rightarrow \langle x, y\rangle = \sum_{n\in N} \langle x, e_n\rangle \langle e_n, y\rangle$;

(iv) $x \in \mathcal{H} \Rightarrow ||x||^2 = \sum_{n\in N} |\langle x, e_n\rangle|^2$.

Such an orthonormal set is called an **orthonormal basis** *of* $\mathcal{H}$.

Proof. $(i) \Rightarrow (ii)$: It is a consequence of Bessel's inequality which states that $\sum_{n\in N} |\langle x, e_n\rangle|^2 < \infty$ and (the implication $(ii) \Leftrightarrow (i)$ of) the last lemma that there exists a vector, call it $x_0 \in \mathcal{H}$, such that $x_0 = \sum_{n\in N} \langle x, e_n\rangle e_n$. If $x \neq x_0$, and if we set $e = \frac{1}{||x-x_0||}(x - x_0)$, then it is easy to see that $\{e_n : n \in N\} \cup \{e\}$ is an orthonormal set which contradicts the assumed maximality of the given orthonormal set.

$(ii) \Rightarrow (iii)$: This is obvious if N is finite, so assume without loss of generality that $N = \mathbb{N}$ For $n \in \mathbb{N}$, let $x_n = \sum_{i=1}^n \langle x, e_i \rangle e_i$ and $y_n = \sum_{i=1}^n \langle y, e_i \rangle e_i$, and note that, by the assumption (ii), continuity of the inner-product, and the assumed orthonormality of the e_i's, we have

$$\begin{aligned}
\langle x, y \rangle &= \lim_{n\to\infty} \langle x_n, y_n \rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, e_i \rangle \overline{\langle y, e_i \rangle} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, e_i \rangle \langle e_i, y \rangle \\
&= \sum_{i=1}^{\infty} \langle x, e_i \rangle \langle e_i, y \rangle \ .
\end{aligned}$$

$(iii) \Rightarrow (iv)$: Put $y = x$.

$(iv) \Rightarrow (i)$: Suppose $\{e_i : i \in I \cup J\}$ is an orthonormal set with J a non-empty index set disjoint from I; then for $j \in J$, we find, in view of (iv), that

$$1 = ||e_j||^2 \ = \ \sum_{i \in I} |\langle e_j, e_i \rangle|^2 \ = \ 0 \ ;$$

hence it must be that J is empty – i.e., the maximality assertion of (i) is indeed implied by (iv). □

The reason for only considering countable orthonormal sets lies in the following proposition.

PROPOSITION 1.4.6. *The following conditions on a Hilbert space $\mathcal{H}$ are equivalent:*

(i) $\mathcal{H}$ is separable;

(ii) Any orthonormal set in $\mathcal{H}$ is countable.

Proof. $(i) \Rightarrow (ii)$: Suppose D is a countable dense set in $\mathcal{H}$ and suppose $\{e_i : i \in I\}$ is an orthonormal set in $\mathcal{H}$. Notice that

$$i \neq j \ \Rightarrow \ ||e_i - e_j||^2 = 2 \ . \tag{1.4.7}$$

Since D is dense in $\mathcal{H}$, we can, for each $i \in I$, find a vector $x_i \in D$ such that $||x_i - e_i|| < \frac{\sqrt{2}}{2}$. The identity (1.4.7) shows that the map $I \ni i \mapsto x_i \in D$ is necessarily 1-1; since D is countable, we may conclude that so is I.

$(ii) \Rightarrow (i)$: If I is a countable (finite or infinite) set and if $\{e_i : i \in I\}$ is an orthonormal basis for $\mathcal{H}$, let D be the set whose typical element is of the form $\sum_{j \in J} \alpha_j e_j$, where J is a finite subset of I and α_j are complex numbers whose real and imaginary parts are both rational numbers; it can then be seen that D is a countable dense set in $\mathcal{H}$. □

COROLLARY 1.4.7. *Any orthonormal set in a Hilbert space is contained in an orthonormal basis– meaning that if* $\{e_i : i \in I\}$ *is any orthonormal set in a Hilbert space* $\mathcal{H}$*, then there exists an orthonormal set*$\{e_i : i \in J\}$ *such that* $I \cap J = \emptyset$ *and* $\{e_i : i \in I \cup J\}$ *is an orthonormal basis for* $\mathcal{H}$*. (If* $\{e_i : i \in I\}$ *is already an orthonormal basis, then* $J = \emptyset$*.)*

In particular, every Hilbert space admits an orthonormal basis.

Proof. This is an easy consequence of Zorn's lemma. □

REMARK 1.4.8. (1) Although we have formally defined an orthonormal basis only in separable Hilbert spaces, Proposition 1.4.5 is true verbatim without the countability hypothesis. The details of this generalisation – which necessitates a digression into what is meant by sums of families of vectors indexed by arbitrary, possibly uncountable, sets – may be found in [Sun], for instance. Thus, non-separable Hilbert spaces are those whose orthonormal bases are uncountable. It is probably fair to say that any true statement about a general non-separable Hilbert space can be established as soon as one knows that the statement is valid for separable Hilbert spaces; it is probably also fair to say that almost all useful Hilbert spaces are separable. So, the reader may safely assume that all Hilbert spaces in the sequel are separable; among these, the finite-dimensional ones are, in a sense, 'trivial', and one only need really worry about infinite-dimensional separable Hilbert spaces.

(2) Every separable non-zero Hilbert space is isometrically isomorphic to exactly one of the family $\{\ell^2_n : n \in \mathbb{N}\} \cup \{\ell^2\}$, where $\mathbb{N} = \{1, 2, \ldots\}$. Thus the cardinality of an orthonormal basis is a complete invariant 'up to isometric isomorphism'. It is clear this is an invariant. For finite-dimensional spaces, the cardinality of an orthonormal basis is the usual vector space dimension, and vector spaces of differing finite dimension are not isomorphic. Also, no finite-dimensional Hilbert space can be isometrically isomorphic to ℓ^2 as the unit ball of ℓ^2 is not compact. (*Reason:* the orthonormal basis$\{e_n : n \in \mathbb{N}\}$ can have no Cauchy subsequence as $\|e_n - e_m\| = \sqrt{2}$ if $m \neq n$.)

REMARK 1.4.9. (1) It follows from Proposition 1.4.5 (ii) that if $\{e_i : i \in I\}$ is an orthonormal basis for a Hilbert space $\mathcal{H}$, then $\mathcal{H} = [\{e_i : i \in I\}]$; conversely, it is true – see Corollary 1.4.14 – that if an orthonormal set is **total** (meaning that the vector subspace spanned by the set is dense in the Hilbert space), then such an orthonormal set is necessarily an orthonormal basis. (*Reason:* apply Theorem 1.4.13(ii), with $\mathcal{M}$ as the closed subspace spanned by the orthonormal set.)

(2) Each of the three examples of an orthonormal set that is given in Example 1.4.2, is in fact an orthonormal basis for the underlying Hilbert space. This is obvious in cases (1) and (2). As for (3), it is a consequence of the Stone-Weierstrass theorem that the vector subspace of finite linear combinations of the exponential functions $\{\exp(2\pi inx) : n \in \mathbb{Z}\}$ (usually called the set

of trigonometric polynomials) is dense in $\{f \in C[0,1] : f(0) = f(1)\}$ (with respect to the uniform norm – i.e., with respect to $||\cdot||_\infty$); in view of Exercise 1.3.2 (2), it is not hard to conclude that this orthonormal set is total in $L^2([0,1],m)$ and hence, by remark (1) above, this is an orthonormal basis for the Hilbert space in question.

Since $\exp(\pm 2\pi i n x) = \cos(2\pi n x) \pm i \sin(2\pi n x)$, and since it is easily verified that $\cos(2\pi m x) \perp \sin(2\pi n x)\ \forall m, n = 1, 2, \ldots$, we find easily that

$$\{1 = e_0\} \cup \{\sqrt{2}\cos(2\pi n x),\ \sqrt{2}\sin(2\pi n x) : n = 1, 2, \ldots\}$$

is also an orthonormal basis for $L^2([0,1],m)$. (Reason: this is orthonormal, and this sequence spans the same vector subspace as is spanned by the exponential basis.) (Also, note that these are real-valued functions, and that the inner product of two real-valued functions is clearly real.) It follows, in particular, that if f is any (real-valued) continuous function defined on [0,1], then such a function admits the following **Fourier series** (with real coefficients):

$$f(x) \;=\; a_0 + \sum_{n=1}^{\infty} (a_n \cos(2\pi n x) + b_n \sin(2\pi n x)\,)$$

where the meaning of this series is that we have convergence of the sequence of the partial sums to the function f with respect to the norm in $L^2[0,1]$. Of course, the coefficients a_n, b_n are given by

$$\begin{aligned} a_0 &= \int_0^1 f(x)\,dx \\ a_n &= 2\int_0^1 f(x)\cos(2\pi n x\,)\,dx\ ,\ \forall\ n > 0, \\ b_n &= 2\int_0^1 f(x)\sin(2\pi n x)\,dx\ ,\ \forall\ n > 0 \end{aligned}$$

The theory of Fourier series was the precursor to most of modern functional analysis; it is for this reason that if $\{e_i : i \in I\}$ is any orthonormal basis of any Hilbert space, it is customary to refer to the numbers $\langle x, e_i\rangle$ as the **Fourier coefficients** of the vector x with respect to the orthonormal basis $\{e_i : i \in I\}$. □

It is a fact that any two orthonormal bases for a Hilbert space have the same cardinality, and this common cardinal number is called the **dimension** of the Hilbert space; the proof of this statement, in its full generality, requires facility with infinite cardinal numbers and arguments of a transfinite nature, and may be found in [Sun]; our interest will be confined to separable Hilbert spaces; the proof in that case of the dimension being an invariant has been outlined in Remark 1.4.9.

We next establish a lemma which will lead to the important result which is sometimes referred to as 'the projection theorem'.

LEMMA 1.4.10. *Let $\mathcal{M}$ be a closed subspace of a Hilbert space $\mathcal{H}$; (thus $\mathcal{M}$ may be regarded as a Hilbert space in its own right;) let $\{e_i : i \in I\}$ be any orthonormal basis for $\mathcal{M}$, and let $\{e_j : j \in J\}$ be any orthonormal set such that $\{e_i : i \in I \cup J\}$ is an orthonormal basis for $\mathcal{H}$, where we assume that the index sets I and J are disjoint. Then, the following conditions on a vector $x \in \mathcal{H}$ are equivalent:*

(i) $x \perp y \ \forall \ y \in \mathcal{M}$;

(ii) $x \ = \ \sum_{j \in J} \langle x, e_j \rangle e_j$.

Proof. The implication $(ii) \Rightarrow (i)$ is obvious. Conversely, it follows easily from Lemma 1.4.4 and Bessel's inequality that the 'series' $\sum_{i \in I} \langle x, e_i \rangle e_i$ and $\sum_{j \in J} \langle x, e_j \rangle e_j$ converge in $\mathcal{H}$. Let the sums of these 'series' be denoted by y and z respectively. Further, since $\{e_i : i \in I \cup J\}$ is an orthonormal basis for $\mathcal{H}$, it should be clear that $x = y + z$. Now, if x satisfies condition (i) of the lemma, it should be clear that $y = 0$ and that hence, $x = z$, thereby completing the proof of the lemma. □

We now come to the basic notion of **orthogonal complement**.

DEFINITION 1.4.11. *The orthogonal complement $S^\perp$ of a subset S of a Hilbert space is defined by*

$$S^\perp \ = \ \{x \in \mathcal{H} : x \perp y \ \forall \ y \in S\} \ .$$

EXERCISE 1.4.12. *If $S_0 \subset S \subset \mathcal{H}$ are arbitrary subsets, show that*

$$S_0^\perp \ \supset S^\perp \ = \ \left(\bigvee S\right)^\perp \ = \ ([S])^\perp \ .$$

Also show that $S^\perp$ is always a closed subspace of $\mathcal{H}$.

We are now ready for the basic fact concerning orthogonal complements of closed subspaces.

THEOREM 1.4.13. *Let $\mathcal{M}$ be a closed subspace of a Hilbert space $\mathcal{H}$. Then,*

(i) $\mathcal{M}^\perp$ *is also a closed subspace;*

(ii) $\left(\mathcal{M}^\perp\right)^\perp \ = \ \mathcal{M}$;

(iii) any vector $x \in \mathcal{H}$ can be uniquely expressed in the form $x = y + z$, where $y \in \mathcal{M}, \ z \in \mathcal{M}^\perp$;

(iv) if x, y, z are as in (3) above, then the equation $Px \ = \ y$ defines a bounded operator $P \in B(\mathcal{H})$ with the property that

$$||Px||^2 = \langle Px, x \rangle = ||x||^2 - ||x - Px||^2 \ , \ \forall x \in \mathcal{H} \ .$$

Proof. (i) This is easy – see Exercise 1.4.12.

(ii) Let $I, J, \{e_i : i \in I \cup J\}$ be as in Lemma 1.4.10. We assert, to start with, that in this case, $\{e_j : j \in J\}$ is an orthonormal basis for $\mathcal{M}^\perp$. Suppose this is not true; since this is clearly an orthonormal set in $\mathcal{M}^\perp$, this means that $\{e_j : j \in J\}$ is not a maximal orthonormal set in $\mathcal{M}^\perp$, which implies the existence of a unit vector $x \in \mathcal{M}^\perp$ such that $\langle x, e_j\rangle = 0 \; \forall \; j \in J$; such an x will satisfy condition (i) of Lemma 1.4.10, but not condition (ii).

If we now reverse the roles of $\mathcal{M}, \{e_i : i \in I\}$ and $\mathcal{M}^\perp, \{e_j : j \in J\}$, we find from the conclusion of the preceding paragraph that $\{e_i : i \in I\}$ is an orthonormal basis for $\left(\mathcal{M}^\perp\right)^\perp$, from which we may conclude the validity of (ii) of this theorem.

(iii) The existence of y and z was demonstrated in the proof of Lemma 1.4.10; as for uniqueness, note that if $x = y_1 + z_1$ is another such decomposition, then we would have

$$y - y_1 \;=\; z_1 - z \;\in\; \mathcal{M} \cap \mathcal{M}^\perp \;;$$

but $w \in \mathcal{M} \cap \mathcal{M}^\perp \;\Rightarrow\; w \perp w \;\Rightarrow ||w||^2 = 0 \;\Rightarrow\; w = 0$.

(iv) The uniqueness of the decomposition in (iii) is easily seen to imply that P is a linear mapping of $\mathcal{H}$ into itself; further, in the notation of (iii), we find (since $y \perp z$) that

$$||x||^2 \;=\; ||y||^2 + ||z||^2 \;=\; ||Px||^2 + ||x - Px||^2 \;;$$

this implies that $||Px|| \leq ||x|| \; \forall \; x \in \mathcal{H}$, and hence $P \in B(\mathcal{H})$.

Also, since $y \perp z$, we find that

$$||Px||^2 \;=\; ||y||^2 \;=\; \langle y, y + z\rangle \;=\; \langle Px, x\rangle \;,$$

thereby completing the proof of the theorem. □

The following corollary to the above theorem justifies the final assertion made in Remark 1.4.9 (1).

COROLLARY 1.4.14. *The following two conditions on an orthonormal set* $\{e_i : i \in I\}$ *in a Hilbert space* $\mathcal{H}$ *are equivalent:*

(i) $\{e_i : i \in I\}$ *is an orthonormal basis for* $\mathcal{H}$*;*

(ii) $\{e_i : i \in I\}$ *is total in* $\mathcal{H}$ *– meaning, of course, that* $\mathcal{H} = [\ \{e_i : i \in I\}\ ]$.

Proof. As has already been observed in Remark 1.4.9 (1), the implication $(i) \Rightarrow (ii)$ follows from Proposition 1.4.5(ii).

Conversely, suppose (i) is not satisfied; then $\{e_i : i \in I\}$ is not a maximal orthonormal set in $\mathcal{H}$; hence there exists a unit vector x such that $x \perp e_i \; \forall i \in I$; if we write $\mathcal{M} = [\ \{e_i : i \in I\}\ ]$, it follows easily that $x \in \mathcal{M}^\perp$, whence $\mathcal{M}^\perp \neq \{0\}$; then, we may deduce from Theorem 1.4.13(2) that $\mathcal{M} \neq \mathcal{H}$ - i.e., (ii) is also not satisfied. □

A standard and easily proved fact is that the following conditions on a linear map $T : \mathcal{H} \rightarrow \mathcal{K}$ between Hilbert spaces are equivalent:

(1) T is continuous; i.e. $\|x_n - x\| \to 0 \Rightarrow \|Tx_n - Tx\| \to 0$;

(2) T is continuous at 0; i.e. $\|x_n\| \to 0 \Rightarrow \|Tx_n\| \to 0$;

(3) $\sup\{\|Tx\| : \|x\| \leq 1\} = \inf\{C > 0 : \|Tx\| \leq C\|x\| \ \forall x \in \mathcal{H}\} < \infty$.

On account of (3) above, such continuous linear maps are called **bounded operators** and we write $B(\mathcal{H}, \mathcal{K})$ for the vector space of all bounded operators from $\mathcal{H}$ to $\mathcal{K}$. It is a standard fact that $B(\mathcal{H}, \mathcal{K})$ is a Banach space if $\|T\|$ is defined as the common value of the two expressions in item (3) above.

We write $B(\mathcal{H})$ for $B(\mathcal{H}, \mathcal{H})$, and note that $B(\mathcal{H})$ is a **Banach algebra** when equipped with composition product $AB = A \circ B$.

It is customary to write $\mathcal{H}^* = B(\mathcal{H}, \mathbb{C})$. We begin by identifying this Banach dual space $\mathcal{H}^*$.

THEOREM 1.4.15. (**Riesz lemma**)
Let $\mathcal{H}$ be a Hilbert space.
(a) If $y \in \mathcal{H}$, the equation

$$\phi_y(x) = \langle x, y \rangle \tag{1.4.8}$$

defines a bounded linear functional $\phi_y \in \mathcal{H}^*$; and further, $||\phi_y||_{\mathcal{H}^*} = ||y||_{\mathcal{H}}$.
(b) Conversely, if $\phi \in \mathcal{H}^*$, there exists a unique element $y \in \mathcal{H}$ such that $\phi = \phi_y$ as in (a) above.

Proof. (a) Linearity of the map ϕ_y is obvious, while the Cauchy-Schwarz inequality shows that ϕ_y is bounded and that $||\phi_y|| \leq ||y||$. Since $\phi_y(y) = ||y||^2$, it easily follows that we actually have equality in the preceding inequality.

(b) Suppose conversely that $\phi \in \mathcal{H}^*$. Since $||\phi_{y_1} - \phi_{y_2}|| = ||y_1 - y_2||$ for all $y_1, y_2 \in \mathcal{H}$, the uniqueness assertion is obvious; we only have to prove existence. Let $\mathcal{M} = \ker \phi$. Since existence is clear if $\phi = 0$, we may assume that $\phi \neq 0$, i.e., that $\mathcal{M} \neq \mathcal{H}$, or equivalently that $\mathcal{M}^\perp \neq 0$.

Notice that the map ϕ is 1-1 from $\mathcal{M}^\perp$ into $\mathbb{C}$; since $\mathcal{M}^\perp \neq 0$, it follows that $\mathcal{M}^\perp$ is one-dimensional. Let z be a unit vector in $\mathcal{M}^\perp$. The y that we seek – assuming it exists – must clearly be an element of $\mathcal{M}^\perp$ (since $\phi(x) = 0 \ \forall x \in \mathcal{M}$). Thus, we must have $y = \alpha z$ for some uniquely determined scalar $0 \neq \alpha \in \mathbb{C}$. With y defined thus, we find that $\phi_y(z) = \overline{\alpha}$; hence we must have $\alpha = \overline{\phi(z)}$. Since any element in $\mathcal{H}$ is uniquely expressible in the form $x + \gamma z$ for some $x \in \mathcal{M}$, and scalar $\gamma \in \mathbb{C}$, we find easily that we do indeed have $\phi = \phi_{\overline{\phi(z)}z}$. $\square$

It must be noted that the mapping $y \mapsto \phi_y$ is not quite an isometric isomorphism of Banach spaces; it is not a linear map, since $\phi_{\alpha y} = \overline{\alpha}\phi_y$; it is only 'conjugate-linear'. The dual (*à priori Banach*) space $\mathcal{H}^*$ is actually a Hilbert space if we define

$$\langle \phi_y, \phi_z \rangle = \langle z, y \rangle ;$$

that this equation satisfies the requirements of an inner product are an easy consequence of the Riesz lemma (and the conjugate-linearity of the mapping

$y \mapsto \phi_y$ already stated); that this inner product actually gives rise to the norm on $\mathcal{H}^*$ is a consequence of the fact that $||y|| \;=\; ||\phi_y||$.

EXERCISE 1.4.16. *(1) Where is the completeness of $\mathcal{H}$ used in the proof of the Riesz lemma; more precisely, what can you say about X^* if you only know that X is an (not necessarily complete) inner product space? (Hint: Consider the completion of X.)*

(2) If $T \in B(\mathcal{H}, \mathcal{K})$, where $\mathcal{H}, \mathcal{K}$ are Hilbert spaces, prove that

$$||T|| \;=\; \sup\{|\langle Tx, y\rangle| : x \in \mathcal{H}, y \in \mathcal{K},\; ||x|| \leq 1, ||y|| \leq 1\}\ .$$

A mapping $B : \mathcal{H} \times \mathcal{K} \to \mathbb{C}$ is called a **bounded sesquilinear form** if

$$B\Big(\sum_{i=1}^{m} \alpha_i x_i, \sum_{j=1}^{n} \beta_j y_j\Big) = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i \bar{\beta}_j B(x_i, y_j),\ \forall \alpha_i, \beta_j \in \mathbb{C}, x_i \in \mathcal{H}, y_j \in \mathcal{K} \tag{1.4.9}$$

and

$$\|B\| := \sup\{|B(x,y)| : \|x\|, \|y\| \leq 1\} < \infty\ . \tag{1.4.10}$$

The following is an easy consequence of the Riesz lemma (see Theorem 1.4.15)) and so its proof is omitted.

PROPOSITION 1.4.17. *(1) $B : \mathcal{H} \times \mathcal{K} \to \mathbb{C}$ is a bounded sesquilinear form if and only if there exists a unique bounded operator $T \in B(\mathcal{H}, \mathcal{K})$ such that $B(x,y) = \langle Tx, y\rangle\ \forall x, y$; furthermore, $\|T\| = \|B\|$.*

(2) Every sesquilinear form defined on $\mathcal{H} \times \mathcal{H}$ satisfies the **polarisation identity***:*

$$4B(x,y) = \sum_{j=0}^{3} (\sqrt{-1}^{\;j} B(x + \sqrt{-1}^{\;j} y, x + \sqrt{-1}^{\;j} y)$$

It is a consequence of the open mapping theorem that the following conditions on a $T \in B(\mathcal{H}, \mathcal{K})$ are equivalent:

(1) There exists an $S \in B(\mathcal{K}, \mathcal{H})$ such that $ST = id, TS = id_{\mathcal{K}}$.

(2) T is a set-theoretic bijection, i.e., both 1-1 and onto.

We call such an operator T invertible, and write $S = T^{-1}$. It is a fact (see [Sun]) that the collection $GL(\mathcal{H}, \mathcal{K})$ of such invertible operators is open in the norm-topology of $B(\mathcal{H}, \mathcal{K})$, and that the mapping $T \mapsto T^{-1}$ is a norm-continuous map of $GL(\mathcal{H}, \mathcal{K})$ onto $GL(\mathcal{K}, \mathcal{H})$.

Recall that the **spectrum** of a $T \in B(\mathcal{H})$ is defined to be $\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda \notin GL(\mathcal{H})\}$. It follows from the previous paragraph that $\sigma(T)$ is a closed set. It is also true that $\sigma(T)$ is a non-empty compact set for any $T \in B(\mathcal{H})$.

An elementary fact about spectra that will be needed later is a special case of a more general **spectral mapping theorem**.

Proposition 1.4.18. *If $p \in \mathbb{C}[t]$ is any polynomial with complex coefficients, and if $T \in B(\mathcal{H})$, then $\sigma(p(T)) = p(\sigma(T))$.*

Proof. Fix a $\lambda \in \mathbb{C}$. If p is a constant, the proposition is obvious, so we assume p is a polynomial of degree $n \geq 1$. Then the algebraic closedness of $\mathbb{C}$ permits a factorisation of the form $p(t) - \lambda = \alpha_n \prod_{i=1}^n (t - \mu_i)$. Clearly, then $p(T) - \lambda = \alpha_n \prod_{i=1}^n (T - \mu_i)$ (where the order of the product is immaterial as the factors commute pairwise). We need the fairly easy fact that if $T_1, \ldots, T_n$ are n pairwise commuting operators, then their product $T_1 \ldots T_n$ is invertible if and only if each T_i is invertible. (Verify this!) Hence conclude that

$$\lambda \notin \sigma(p(T)) \Leftrightarrow \mu_i \notin \sigma(T) \forall i$$

or equivalently, that $\lambda \in \sigma(p(T))$ if and only if there exists some i such that $\mu_i \in \sigma(T)$. This is equivalent to saying that $\lambda \in p(\sigma(T))$; and thus, indeed $\sigma(p(T)) = p(\sigma(T))$. □

It is a fact that $\lambda \in \sigma(T) \Rightarrow |\lambda| \leq \|T\|$ and that the spectrum is always compact. The non-emptiness is a more non-trivial fact. (This statement for all finite-dimensional $\mathcal{H}$ is equivalent to the fact that $\mathbb{C}$ is algebraically closed, i.e., that every complex polynomial is a product of linear factors.)

Another proof that simultaneously establishes the fact that $\sigma(T)$ is non-empty and compact is the (not surprisingly complex analytic) proof of the so-called **spectral radius formula**:

$$\operatorname{spr}(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} . \tag{1.4.11}$$

This says two things: (i) that the indicated limit exists, and (ii) that the value of the limit is as asserted. Part (ii) shows that the spectral radius is non-negative, and hence that spectrum is always non-empty. We will shortly be using part (i) to establish that $\operatorname{spr}(T) = \|T\|$ if T is 'normal', which is a key ingredient in the proof of the spectral theorem.

Most of this required background material can be found in the initial chapters of most standard books (such as [Sun]) covering the material of a first course in Functional Analysis.

1.5 Adjoints

An immediate consequence of the Riesz lemma (Lemma 1.4.15) is:

Proposition 1.5.1. *If $T \in B(\mathcal{H}, \mathcal{K})$, there exists a unique operator $T^* \in B(\mathcal{K}, \mathcal{H})$ – called the* **adjoint** *of the operator T – such that*

$$\langle T^* y, x\rangle = \langle y, Tx\rangle \ \forall x \in \mathcal{H}, y \in \mathcal{K}.$$

Proof. Notice that the right side of the displayed equation above defines a bounded sesquilinear form on $\mathcal{K} \times \mathcal{H}$, and appeal to Proposition 1.4.17 to lay hands on the desired operator T^*. □

We list below some simple properties of this process of taking adjoints.

PROPOSITION 1.5.2. (1) *For all* $\alpha \in \mathbb{C}, S, S_1, S_2 \in B(\mathcal{H},\mathcal{K}), T \in B(\mathcal{M},\mathcal{H})$, *we have:*

$$\begin{aligned} (\alpha S_1 + S_2)^* &= \bar{\alpha} S_1^* + S_2^* \ ; \\ (S^*)^* &= S \ ; \\ (ST)^* &= T^* S^* \ ; \\ id_{\mathcal{H}}^* &= id_{\mathcal{H}} \ . \end{aligned}$$

(2) $\|T\|^2 = \|T^*T\|$ *and hence, also* $\|T^*\| = \|T\|$;

(3) $\ker(T^*) = \mathrm{ran}^\perp(T) := (\mathrm{ran}(T))^\perp$; *equivalently,* $\ker^\perp(T^*) = \overline{\mathrm{ran}(T)}$.

Proof. (1) Most of these identities follow from the fact that the adjoint is characterised by the equation it satisfies. Thus, for instance,

$$\begin{aligned} \langle (\alpha S_1 + S_2)^* y, x\rangle &= \langle y, (\alpha S_1 + S_2)x\rangle \\ &= \bar{\alpha}\langle y, S_1 x\rangle + \langle y, S_2 x\rangle \\ &= \bar{\alpha}\langle S_1^* y, x\rangle + \langle S_2^* y, x\rangle \\ &= \langle (\bar{\alpha} S_1^* + S_2^*)y, x\rangle. \end{aligned}$$

The other three statements are even more straight-forward to verify.

(2) On the one hand,

$$\begin{aligned} \|T\|^2 &= \sup\{\|Tx\|^2 : \|x\| \leq 1\} \\ &= \sup\{\langle T^*Tx, x\rangle : \|x\| \leq 1\} \\ &\leq \|T^*T\| \ , \end{aligned}$$

while on the other,

$$\begin{aligned} \|T^*T\| &= \sup\{|\langle T^*Tx_1, x_2\rangle| : \|x_1\||, \|x_2\| \leq 1\} \\ &\leq \sup\{\|Tx_1\|\|Tx_2\| : \|x_1\|, \|x_2\| \leq 1\} \\ &\leq \|T\|^2 \ . \end{aligned}$$

(Observe that the Cauchy-Schwarz inequality $|\langle x, y\rangle| \leq \|x\| \ \|y\|$ has been used in the proofs of both inequalities above – in the third line of the first, and in the second line of the second.) The desired equality follows, and the sub-multiplicativity of the norm then implies that $\|T^*\| \leq \|T\|$. By interchanging the roles of T and T^*, we find that, indeed $\|T^*\| = \|T\|$.

(3)

$$\begin{aligned} y \in \ker(T^*) &\Leftrightarrow T^*y = 0 \\ &\Leftrightarrow \langle T^*y, x\rangle = 0 \ \forall x \\ &\Leftrightarrow \langle y, Tx\rangle = 0 \ \forall x \\ &\Leftrightarrow y \in \mathrm{ran}^\perp(T) \ . \end{aligned}$$

□

The polarisation identity has the following immediate corollaries:

COROLLARY 1.5.3. (1) *If $T \in B(\mathcal{H})$, then*

$$T = 0 \ \Leftrightarrow \ \langle Tx, x\rangle = 0 \ \forall x \in \mathcal{H}.$$

(2)

$$T = T^* \ \Leftrightarrow \ \langle Tx, x\rangle \in \mathbb{R} \ \forall x \in \mathcal{H} \ .$$

This corolllary leads to the definition of an important class of operators:

DEFINITION 1.5.4. *An operator $T \in B(\mathcal{H})$ is said to be* **self-adjoint** *(or* **Hermitian***) if $T = T^*$.*

A slightly larger class of operators, which is *the correct* class of operators for the purposes of the spectral theorem, is dealt with in our next definition.

DEFINITION 1.5.5. *An operator $Z \in B(\mathcal{H})$ is said to be* **normal** *if $Z^*Z = ZZ^*$.*

PROPOSITION 1.5.6. *Let $Z \in B(\mathcal{H})$.*

(1) *Z is normal if and only if $\|Zx\| = \|Z^*x\| \ \forall x \in \mathcal{H}$.*

(2) *If Z is normal, then $\|Z^2\| = \|Z^*Z\| = \|Z\|^2$; more generally, $\|Z^{2^n}\| = \|Z\|^{2^n}$ and consequently $\mathrm{spr}(Z) = \|Z\|$.*

Proof. (1)

$$\begin{aligned} Z^*Z = ZZ^* \quad &\Leftrightarrow \quad \langle Z^*Zx, x\rangle = \langle ZZ^*x, x\rangle \ \forall x \in \mathcal{H} \\ &\Leftrightarrow \quad \|Zx\|^2 = \|Z^*x\|^2 \ \forall x \in \mathcal{H}. \end{aligned}$$

(2) Suppose Z is normal. Then,

$$\begin{aligned} \|Z^2\| \quad &= \quad \sup\{\|Z^2x\| : \|x\| = 1\} \\ &= \quad \sup\{\|Z^*Zx\| : \|x\| = 1\} \qquad \text{(by part (1) above)} \\ &= \quad \|Z^*Z\| \\ &= \quad \|Z\|^2 \end{aligned}$$

where we have used Proposition 1.5.2(2) in the last step; an easy induction argument now yields the statement about 2^n, which implies that $\|Z\| = \lim_{n\to\infty} \|Z^{2^n}\|^{\frac{1}{2^n}} = \mathrm{spr}(Z)$.

$\square$

We now have the tools at hand to prove a key identity.

PROPOSITION 1.5.7. *If $X \in B(\mathcal{H})$ is self-adjoint, and $p \in \mathbb{C}[t]$ is any polynomial with complex coefficients, then*

$$\|p(X)\| = \|p\|_{\sigma(X)} := \sup\{|p(t)| : t \in \sigma(X)\}. \qquad (1.5.12)$$

Proof. Notice that $q = |p|^2 = \bar{p}p$ is a polynomial with real coefficients, and hence $q(X)$ is self-adjoint. Deduce from Proposition 1.5.6 (2) and the spectral mapping theorem (Proposition 1.4.18) that

$$\begin{aligned} \|p(X)\|^2 &= \|p(X)^*p(X)\| \\ &= \|\bar{p}(X)p(X)\| \\ &= \|q(X)\| \\ &= \sup\{|\lambda| : \lambda \in \sigma(q(X))\} \\ &= \sup\{|q(t)| : t \in \sigma(X)\} \\ &= \|q\|_{\sigma(X)} \\ &= \|p\|^2_{\sigma(X)} \ , \end{aligned}$$

as desired. □

Just as every complex number has a unique decomposition into real and imaginary parts, it is seen that each $Z \in \mathcal{B}(H)$ has a unique **Cartesian decomposition** $Z = X + iY$, with X and Y being self-adjoint (these being necessarily given by $X = \frac{1}{2}(Z + Z^*)$ and $Y = \frac{1}{2i}(Z - Z^*)$, so that, in fact, $\langle Xx, x\rangle = \text{Re}\ \langle Zx, x\rangle$ and $\langle Yx, x\rangle = \text{Im}\ \langle Zx, x\rangle$). For this reason, we sometimes write $X = \text{Re}\ Z, Y = \text{Im}\ Z$.

For future reference, we make some observations on the Cartesian decomposition of a normal operator.

PROPOSITION 1.5.8. *Let $Z = X + iY$ be the Cartesian decomposition of an operator. Then, the following conditions are equivalent:*

(1) *Z is normal.*

(2) *$\|Zx\|^2 = \|Xx\|^2 + \|Yx\|^2\ \forall x \in \mathcal{H}$.*

(3) *$XY = YX$.*

Proof. First notice that for $Z = X + iY$, we have

$$\begin{aligned} \|Zx\|^2 &= \|Xx + iYx\|^2 \\ &= \|Xx\|^2 + \|Yx\|^2 - 2\text{Re}\ (i\langle Xx, Yx\rangle) \end{aligned}$$

while

$$\begin{aligned} \|Z^*x\|^2 &= \|Xx - iYx\|^2 \\ &= \|Xx\|^2 + \|Yx\|^2 + 2\text{Re}\ (i\langle Xx, Yx\rangle), \end{aligned}$$

so that

$$\|Z^*x\|^2 = \|Zx\|^2 \Leftrightarrow \text{Re}\ (i\langle Xx, Yx\rangle) = 0 \Leftrightarrow \|Zx\|^2 = \|Xx\|^2 + \|Yx\|^2.$$

Notice finally that

$$\text{Re}\ i\langle Xx, Yx\rangle = 0 \Leftrightarrow \langle Xx, Yx\rangle \in \mathbb{R}$$

and that (since X, Y are self-adjoint)

$$\langle Xx, Yx\rangle \in \mathbb{R}\ \forall x \in \mathcal{H} \Leftrightarrow XY = (XY)^* = YX.$$

The truth of the lemma is evident now. □

1.6 Approximate eigenvalues

DEFINITION 1.6.1. *A scalar $\lambda \in \mathbb{C}$ is said to be an* **approximate eigenvalue** *of an operator $Z \in B(\mathcal{H})$ if there exists a sequence $\{x_n : n \in \mathbb{N}\} \subset S(\mathcal{H})$ such that* $\lim_{n\to\infty} \|(Z-\lambda)x_n\| = 0$. *Here and in the sequel, we shall employ the symbol $S(\mathcal{H})$ to denote the unit sphere of $\mathcal{H}$; thus, $S(\mathcal{H}) := \{x \in \mathcal{H} : \|x\| = 1\}$.*

The importance – as emerges from [Hal] – of this notion in the context of the spectral theorem (equivalently, the study of self-adjoint or normal operators) lies in the following result:

THEOREM 1.6.2. *Suppose $Z \in B(\mathcal{H})$ is normal. Then:*

(1) $Z \in GL(\mathcal{H}) \Leftrightarrow Z$ is bounded below; i.e., there is an $\epsilon > 0$ such that $\|Zx\| \geq \epsilon\|x\|\ \forall x \in \mathcal{H}$, equivalently, $\inf\{\|Zx\| : x \in S(\mathcal{H})\} \geq \epsilon > 0$ *(assuming $\mathcal{H} \neq 0$).*

(2) $\lambda \in \sigma(Z)$ if and only if λ is an approximate eigenvalue of Z.

Proof. (1) If Z is invertible, then note that

$$\|x\| = \|Z^{-1}Zx\| \leq \|Z^{-1}\|\|Zx\|\ \forall x$$

which shows that $\|Zx\| \geq \|Z^{-1}\|^{-1}\|x\|\ \forall x$ and that Z is indeed bounded below.

If, conversely, Z is bounded below, deduce two consequences, *viz.*,

(a) Z^* is also bounded below (by part (1) of Proposition 1.5.6)) and hence $\ker(Z^*)(= \ker(Z)) = \{0\}$ so that $\mathrm{ran}(Z)$ is dense in $\mathcal{H}$ (by part (3) of Proposition1.5.2).

(b) Z has a closed range (*Reason*: If $Zx_n \to y$ then $\{Zx_n : n \in \mathbb{N}\}$, and consequently also $\{x_n : n \in \mathbb{N}\}$, must be a Cauchy sequence, forcing $y = Z(\lim_{n\to\infty} x_n)$.)

It follows from (a) and (b) above that Z is a bijective linear map of $\mathcal{H}$ onto itself and hence invertible.

(2) Note first that $(Z-\lambda)$ inherits normality from Z, then deduce from (1) above that $\lambda \in \sigma(Z)$ if and only if there exists a sequence $x_n \in S(\mathcal{H})$ such that $\|(Z-\lambda)x_n\| < \frac{1}{n}\ \forall n$, i.e., λ is an approximate eigenvalue of z, as desired.

□

Corollary 1.6.3.

$$X = X^* \ \Rightarrow \ \sigma(X) \subset \mathbb{R}.$$

Proof. If there exists a sequence $\{x_n : n \in \mathbb{N}\} \subset S(\mathcal{H})$ satisfying the condition $\|(X - \lambda)x_n\| \to 0$, then also $\langle (X - \lambda)x_n, x_n \rangle \to 0$ and hence

$$\lambda = \lim_{n\to\infty} \langle \lambda x_n, x_n \rangle = \lim_{n\to\infty} \langle X x_n, x_n \rangle \in \mathbb{R}$$

(by Corollary 1.5.3 (2)). □

For later reference, we record an immediate consequence of Theorem 1.6.2 (2) and Proposition 1.5.8 (2).

Corollary 1.6.4. *Suppose $\lambda = \alpha + i\beta$ and $Z = X + iY$ are the Cartesian decompositions of a scalar λ and a normal operator Z respectively. Then the following conditions are equivalent:*

(1) $\lambda \in \sigma(Z)$;

(2) *There exists a sequence $\{x_n : n \in \mathbb{N}\}$ such that $\|(X - \alpha)x_n\| \to 0$ and $\|(Y - \beta)x_n\| \to 0$.*

1.7 Important classes of operators

1.7.1 Projections

Remark 1.7.1. The operator $P \in B(\mathcal{H})$ constructed in Theorem 1.4.13(4) is referred to as the **orthogonal projection** onto the closed subspace $\mathcal{M}$. When it is necessary to indicate the relation between the subspace $\mathcal{M}$ and the projection P, we will write $P = P_{\mathcal{M}}$ and $\mathcal{M} = \operatorname{ran} P$ (note that $\mathcal{M}$ is indeed the range of the operator P); some other facts about closed subspaces and projections are spelt out in the following exercises. □

Exercise 1.7.2. (1) *Show that $\left(S^\perp\right)^\perp = [S]$, for any subset $S \subset \mathcal{H}$.*

(2) *Let $\mathcal{M}$ be a closed subspace of $\mathcal{H}$, and let $P = P_{\mathcal{M}}$;*

(a) Show that $P_{\mathcal{M}^\perp} = 1 - P_{\mathcal{M}}$;

(b) Let $x \in \mathcal{H}$; the following conditions are equivalent:

(i) $x \in \mathcal{M}$;

(ii) $x \in \operatorname{ran} P \ (:= P\mathcal{H})$;

(iii) $Px = x$;

(iv) $||Px|| = ||x||$.

(c) Show that $\mathcal{M}^\perp = \ker P = \{x \in \mathcal{H} : Px = 0\}$.

(3) *Let* $\mathcal{M}$ *and* $\mathcal{N}$ *be closed subspaces of* $\mathcal{H}$, *and let* $P = P_{\mathcal{M}}, Q = P_{\mathcal{N}}$; *show that the following conditions are equivalent:*

(i) $\mathcal{N} \subset \mathcal{M}$;

(ii) $PQ = Q$;

(i)′ $\mathcal{M}^{\perp} \subset \mathcal{N}^{\perp}$;

(ii)′ $(1 - Q)(1 - P) = 1 - P$;

(iii) $QP = Q$.

(4) *With* $\mathcal{M}, \mathcal{N}, P, Q$ *as in* (3) *above, show that the following conditions are equivalent:*

(i) $\mathcal{M} \perp \mathcal{N}$ – *i.e.*, $\mathcal{N} \subset \mathcal{M}^{\perp}$;

(ii) $PQ = 0$;

(iii) $QP = 0$.

(5) When the equivalent conditions of (4) are met, show that:

(a) $[\mathcal{M} \cup \mathcal{N}] = \mathcal{M} + \mathcal{N} = \{x + y : x \in \mathcal{M}, y \in \mathcal{N}\}$; *and that*

(c) $(P + Q)$ *is the projection onto the subspace* $\mathcal{M} + \mathcal{N}$.

(6) *Show, more generally, that*

(a) *if* $\{\mathcal{M}_i : 1 \leq i \leq n\}$ *is a family of closed subspaces of* $\mathcal{H}$ *which are pairwise orthogonal, then their 'vector sum' defined by* $\sum_{i=1}^{n} \mathcal{M}_i = \{\sum_{i=1}^{n} x_i : x_i \in \mathcal{M}_i \ \forall i\}$ *is a closed subspace and the projection onto this subspace is given by* $\sum_{i=1}^{n} P_{\mathcal{M}_i}$; *and that*

(b) *if* $\{\mathcal{M}_n : n \in \mathbb{N}\}$ *is a family of closed subspaces of* $\mathcal{H}$ *which are pairwise orthogonal, and if* $\mathcal{M} = [\bigcup_{n \in \mathbb{N}} \mathcal{M}_n]$, *then* $P_{\mathcal{M}}$ *is given by the sum of the series* $\sum_{n \in \mathbb{N}} P_{\mathcal{M}_n}$ *which is interpreted in the SOT-sense (see Definition* 2.2.4*): meaning that* $(\sum_{n \in \mathbb{N}} P_{\mathcal{M}_n})x = \sum_{n \in \mathbb{N}} P_{\mathcal{M}_n} x$, *with the series on the right side converging in the norm for each* $x \in \mathcal{H}$.

Self-adjoint operators are the building blocks of all operators, and they are by far the most important subclass of all bounded operators on a Hilbert space. However, in order to see their structure and usefulness, we will have to wait until after we have proved the fundamental spectral theorem. This will allow us to handle self-adjoint operators with exactly the same facility with which we handle real-valued functions.

Nevertheless, we have already seen one important special class of self-adjoint operators as shown by the next result.

PROPOSITION 1.7.3. *Let $P \in B(\mathcal{H})$. Then the following two conditions are equivalent:*
(i) $P = P_{\mathcal{M}}$ is the orthogonal projection onto some closed subspace $\mathcal{M} \subset \mathcal{H}$;
(ii) $P = P^2 = P^*$.

Proof. $(i) \Rightarrow (ii)$: If $P = P_{\mathcal{M}}$, the definition of an orthogonal projection shows that $P = P^2$; the self-adjointness of P follows from Theorem 1.4.13 (4) and Corollary 1.5.3 (2).

$(ii) \Rightarrow (i)$: Suppose (ii) is satisfied; let $\mathcal{M} = \operatorname{ran} P$, and note that

$$\begin{aligned} x \in \mathcal{M} \quad &\Rightarrow \quad \exists y \in \mathcal{H} \;\text{ such that }\; x = Py \\ &\Rightarrow \quad Px = P^2 y = Py = x \;; \end{aligned} \tag{1.7.13}$$

on the other hand, note that

$$\begin{aligned} y \in \mathcal{M}^{\perp} \quad &\Leftrightarrow \quad \langle y, Pz \rangle = 0 \;\forall z \in \mathcal{H} \\ &\Leftrightarrow \quad \langle Py, z \rangle = 0 \;\forall z \in \mathcal{H} \quad (\text{since } P = P^*) \\ &\Leftrightarrow \quad Py \;=\; 0 \;; \end{aligned} \tag{1.7.14}$$

hence, if $z \in \mathcal{H}$ and $x = P_{\mathcal{M}} z,\; y = P_{\mathcal{M}^{\perp}} z$, we find from equations (1.7.13) and (1.7.14) that $Pz = Px + Py = x = P_{\mathcal{M}} z$. □

Direct Sums and Operator Matrices

If $\{\mathcal{M}_n : n \in \mathbb{N}\}$ are pairwise orthogonal closed subspaces – see Exercise 1.7.2 (5)(d) – and if $\mathcal{M} = [\bigcup_{n \in \mathbb{N}} \mathcal{M}_n]$ we say that $\mathcal{M}$ is the **direct sum** of the closed subspaces $\mathcal{M}_i, 1 \leq i \leq n$, and we write

$$\mathcal{M} = \bigoplus_{n=1}^{\infty} \mathcal{M}_i \;; \tag{1.7.15}$$

conversely, whenever we use the above symbol, it will always be tacitly assumed that the $\mathcal{M}_i$'s are closed subspaces which are pairwise orthogonal and that $\mathcal{M}$ is the (closed) subspace spanned by them.

To clarify matters, let us first consider the direct sum of two subspaces. (We are going to try and mimic the success of operators on $\mathbb{C}^2$ being identifiable with the operation of matrices acting on column vectors by multiplication.)

So suppose $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. We shall think of a typical element $x \in \mathcal{H}$ as a column vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, with $x_i \in \mathcal{H}_i$. Let $P_i = P_{\mathcal{H}_i}$ so $P_i x = x_i$ in the above notation. If we think of P_i as being an element of $B(\mathcal{H}, \mathcal{H}_i)$, then it is easily seen that its adjoint is the isometric element V_i of $B(\mathcal{H}_i, \mathcal{H})$ described thus:

$$V_1 x_1 = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \text{ and } V_2 x_2 = \begin{bmatrix} 0 \\ x_2 \end{bmatrix}.$$

Given a $T \in B(\mathcal{H})$, define $T_{ij} = P_i T V_j \in B(\mathcal{H}_j, \mathcal{H}_i)$ and observe that we have

$$Tx = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

If we refer to $((T_{ij}))$ as the matrix corresponding to T, then the matrices corresponding to P_1 and P_2 are seen to be

$$\begin{bmatrix} \mathrm{id}_{\mathcal{H}_1} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{id}_{\mathcal{H}_2} \end{bmatrix}.$$

More generally, if $\mathcal{H} = \bigoplus_{j \in \mathbb{N}} \mathcal{H}_j, \mathcal{K} = \bigoplus_{i \in \mathbb{N}} \mathcal{K}_i$, there exists a unique matrix $((T_{ij}))$ with $T_{ij} \in B(\mathcal{H}_j, \mathcal{K}_i)$ such that whenever $\xi_j \in \mathcal{H}_j$ satisfy $\sum_{j \in \mathbb{N}} \|\xi_j\|^2 < \infty$ (so that the series $\sum_{j \in \mathbb{N}} \xi_j$ converges in $\mathcal{H}$ (to ξ, say), then $T\xi = \sum_{i \in \mathbb{N}} \left(\sum_{j \in \mathbb{N}} T_{ij} \xi_j \right)$ – with the inner series converging in $\mathcal{K}_i$ for each $i \in \mathbb{N}$ to η_i, say, with $\sum_{i \in \mathbb{N}} \|\eta_i\|^2 < \infty$ and $T\xi = \sum_{i \in \mathbb{N}} \eta_i$. In the special case when each $\mathcal{H}_j$ and $\mathcal{K}_i$ is one-dimensional, this reduces to saying that if $T \in B(\mathcal{H}, \mathcal{K})$ and if $\{x_j : j \in \mathbb{N}\}$ (resp., $\{y_i : i \in \mathbb{N}\}$) is an orthonormal basis in $\mathcal{H}$ (resp., $\mathcal{K}$), then the operator T can be described by matrix multiplication in the following sense: if the vector $x \in \mathcal{H}$ (resp., $y \in \mathcal{K}$) is thought of as the countably infinite column matrix $[x] = [\beta_j]$ with $\beta_j = \langle x, x_j \rangle$ (resp., $[y] = [\alpha_i]$ with $\alpha_i = \langle y, y_i \rangle$), and if $[T]$ is the matrix $((t_{ij}))$ with countably infinitely many rows and columns with $t_{ij} = \langle Tx_j, y_i \rangle$, then $Tx = y \Leftrightarrow \alpha_i = \sum_j t_{ij} \beta_j \ \forall i$.

Exercise 1.7.4. *(1) Verify the assertions of the previous paragraphs. (*Hint: *The computation in the case of finite direct sums will show what needs to be done in the infinite case.)*

(2) With the notation of the paragraph preceding this exercise, verify that the familiar E_{ij} matrix whose only non-zero entry is a 1 in the (i, j)-th spot is the matrix of the operator denoted by $(\bar{x}_j \otimes y_i)$ in Exercise 3.2.11, *and defined in the paragraph preceding that exercise.*

(3) Verify the following fundamental rules concerning the **system** $\{E_{ij}\}$*:*

(1) $E_{ij}^* = E_{ji}$;

(2) $E_{ij} E_{kl} = \delta_{jk} E_{il}$

where the Kronecker symbol is defined by

$$\delta_{pq} = \begin{cases} 1 & if\ p = q \\ 0 & otherwise \end{cases}.$$

1.7.2 Isometric versus Unitary

The two propositions given below identify two important classes of operators between Hilbert spaces.

PROPOSITION 1.7.5. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces; the following conditions on an operator $U \in B(\mathcal{H}, \mathcal{K})$ are equivalent:*

(*i*) *if $\{e_i : i \in I\}$ is any orthonormal set in $\mathcal{H}$, then also $\{Ue_i : i \in I\}$ is an orthonormal set in $\mathcal{K}$;*

(*ii*) *there is an orthonormal basis $\{e_i : i \in I\}$ for $\mathcal{H}$ such that $\{Ue_i : i \in I\}$ is an orthonormal set in $\mathcal{K}$;*

(*iii*) $\langle Ux, Uy\rangle \ = \ \langle x, y\rangle \ \forall \ x, y \in \mathcal{H}$*;*

(*iv*) $||Ux|| \ = \ ||x|| \ \forall \ x \in \mathcal{H}$*;*

(*v*) $U^*U \ = \ 1_{\mathcal{H}}$.

An operator satisfying these equivalent conditions is called an **isometry**.

Proof. $(i) \Rightarrow (ii)$: There exists an orthonormal basis for $\mathcal{H}$.

$(ii) \Rightarrow (iii)$: If $x, y \in \mathcal{H}$ and if $\{e_i : i \in I\}$ is as in (ii), then

$$\begin{aligned}
\langle Ux, Uy\rangle &= \left\langle U\Big(\sum_{i\in I}\langle x, e_i\rangle e_i\Big) \ , \ U\Big(\sum_{j\in I}\langle y, e_j\rangle e_j\Big) \right\rangle \\
&= \sum_{i,j\in I}\langle x, e_i\rangle\langle e_j, y\rangle\langle Ue_i, Ue_j\rangle \\
&= \sum_{i\in I}\langle x, e_i\rangle\langle e_i, y\rangle \\
&= \langle x, y\rangle \ .
\end{aligned}$$

$(iii) \Rightarrow (iv)$: Put $y = x$.

$(iv) \Rightarrow (v)$: If $x \in \mathcal{H}$, note that

$$\langle U^*Ux, x\rangle \ = \ ||Ux||^2 \ = \ ||x||^2 \ = \ \langle 1_{\mathcal{H}}x, x\rangle \ ,$$

and appeal to the fact that a bounded operator is determined by its quadratic form – see Corollary 1.5.3.

$(v) \Rightarrow (i)$: If $\{e_i : i \in I\}$ is any orthonormal set in $\mathcal{H}$, then

$$\langle Ue_i, Ue_j\rangle \ = \ \langle U^*Ue_i, e_j\rangle \ = \ \langle e_i, e_j\rangle \ = \ \delta_{ij} \ .$$

□

PROPOSITION 1.7.6. *The following conditions on an isometry $U \in B(\mathcal{H}, \mathcal{K})$ are equivalent:*

(*i*) *if $\{e_i : i \in I\}$ is any orthonormal basis for $\mathcal{H}$, then $\{Ue_i : i \in I\}$ is an orthonormal basis for $\mathcal{K}$;*

(*ii*) *there is an orthonormal set $\{e_i : i \in I\}$ in $\mathcal{H}$ such that $\{Ue_i : i \in I\}$ is an orthonormal basis for $\mathcal{K}$;*

(*iii*) $UU^* \ = \ 1_{\mathcal{K}}$*;*

(*iv*) *U is invertible;*

(v) U maps $\mathcal{H}$ onto $\mathcal{K}$.

An isometry which satisfies the above equivalent conditions is said to be **unitary**.

Proof. $(i) \Rightarrow (ii)$: Obvious.

$(ii) \Rightarrow (iii)$: If $\{e_i : i \in I\}$ is as in (ii), and if $x \in \mathcal{K}$, observe that

$$\begin{aligned} UU^*x &= UU^*\Big(\sum_{i \in I} \langle x, Ue_i\rangle Ue_i\Big) \\ &= \sum_{i \in I} \langle x, Ue_i\rangle UU^*Ue_i \\ &= \sum_{i \in I} \langle x, Ue_i\rangle Ue_i \quad \text{(since } U \text{ is an isometry)} \\ &= x \ . \end{aligned}$$

$(iii) \Rightarrow (iv)$: The assumption that U is an isometry, in conjunction with the hypothesis (iii), says that $U^* = U^{-1}$.

$(iv) \Rightarrow (v)$: Obvious.

$(v) \Rightarrow (i)$: If $\{e_i : i \in I\}$ is an orthonormal basis for $\mathcal{H}$, then $\{Ue_i : i \in I\}$ is an orthonormal set in $\mathcal{H}$, since U is isometric. Now, if $z \in \mathcal{K}$, pick $x \in \mathcal{H}$ such that $z = Ux$, and observe that

$$\begin{aligned} ||z||^2 &= ||Ux||^2 \\ &= ||x||^2 \\ &= \sum_{i \in I} |\langle x, e_i\rangle|^2 \\ &= \sum_{i \in I} |\langle z, Ue_i\rangle|^2 \ , \end{aligned}$$

and since z was arbitrary, this shows that $\{Ue_i : i \in I\}$ is an orthonormal basis for $\mathcal{K}$. □

Thus, unitary operators are the natural isomorphisms in the context of Hilbert spaces. The collection of unitary operators from $\mathcal{H}$ to $\mathcal{K}$ will be denoted by $\mathcal{U}(\mathcal{H}, \mathcal{K})$; when $\mathcal{H} = \mathcal{K}$, we shall write $\mathcal{U}(\mathcal{H}) = \mathcal{U}(\mathcal{H}, \mathcal{H})$. We list some elementary properties of unitary and isometric operators in the next exercise.

EXERCISE 1.7.7. (1) *Suppose that $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces and suppose $\{e_i : i \in I\}$ (resp., $\{f_i : i \in I\}$) is an orthonormal basis (resp., orthonormal set) in $\mathcal{H}$ (resp., $\mathcal{K}$), for some index set I. Show that:*

(a) $\dim \mathcal{H} \leq \dim \mathcal{K}$*; and*

(b) there exists a unique isometry $U \in B(\mathcal{H}, \mathcal{K})$ such that $Ue_i = f_i \ \forall i \in I$.

(2) *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Show that:*

(a) there exists an isometry $U \in B(\mathcal{H}, \mathcal{K})$ if and only if dim $\mathcal{H} \leq \dim \mathcal{K}$;

(b) there exists a unitary $U \in B(\mathcal{H}, \mathcal{K})$ if and only if dim $\mathcal{H}$ = dim $\mathcal{K}$.

(3) *Show that $\mathcal{U}(\mathcal{H})$ is a group under multiplication, which is a (norm-) closed subset of the Banach space $B(\mathcal{H})$.*

(4) *Suppose $U \in \mathcal{U}(\mathcal{H}, \mathcal{K})$; show that the association*

$$B(\mathcal{H}) \ni T \overset{ad\,U}{\mapsto} UTU^* \in B(\mathcal{K}) \tag{1.7.16}$$

*defines a mapping $(ad\,U) : B(\mathcal{H}) \to B(\mathcal{K})$ which is an 'isometric isomorphism of Banach *-algebras', meaning that:*

(a) $ad\,U$ is an isometric isomorphism of Banach spaces: i.e., $ad\,U$ is a linear mapping which is 1-1, onto, and is norm-preserving; (Hint: verify that it is linear and preserves norm and that its inverse is given by $ad\,U^*$.)

(b) $ad\,U$ is a product-preserving map between Banach algebras; i.e., $(ad\,U)(T_1T_2) = \big((ad\,U)\,(T_1)\big)\,\big(ad\,U)(T_2)\big)$, for all $T_1, T_2 \in B(\mathcal{H})$;

*(c) $ad\,U$ is a *-preserving map between *-algebras; i.e.,*

$$\big((ad\,U)(T)\big)^* = (ad\,U)(T^*)\ \forall T \in B(\mathcal{H}).$$

(5) *Show that the map $U \mapsto (ad\,U)$ is a homomorphism from the group $\mathcal{U}(\mathcal{H})$ into the group $Aut\,B(\mathcal{H})$ of all automorphisms (= isometric isomorphisms of the Banach *-algebra $B(\mathcal{H})$ onto itself); further, verify that if $U_n \to U$ in $\mathcal{U}(\mathcal{H}, \mathcal{K})$, then $(ad\,U_n)(T) \to (ad\,U)(T)$ in $B(\mathcal{K})$ for all $T \in B(\mathcal{H})$.*

A unitary operator between Hilbert spaces should be viewed as 'implementing an inessential variation'; thus, if $U \in \mathcal{U}(\mathcal{H}, \mathcal{K})$ and if $T \in B(\mathcal{H})$, then the operator $UTU^* \in B(\mathcal{K})$ should be thought of as being 'essentially the same as T', except that it is probably being viewed from a different observer's perspective. All this is made precise in the following definition.

DEFINITION 1.7.8. *Two operators $T \in B(\mathcal{H})$ and $S \in B(\mathcal{K})$ (on two possibly different Hilbert spaces) are said to be* **unitarily equivalent** *if there exists a unitary operator $U \in \mathcal{U}(\mathcal{H}, \mathcal{K})$ such that $S = UTU^*$.*

We conclude this section with a discussion of some examples of isometric operators, which will illustrate the preceding notions quite nicely.

EXAMPLE 1.7.9. To start with, notice that if $\mathcal{H}$ is a finite-dimensional Hilbert space, then an isometry $U \in B(\mathcal{H})$ is necessarily unitary. (Prove this!) Hence, the notion of non-unitary isometries of a Hilbert space into itself makes sense only in infinite-dimensional Hilbert spaces. We discuss some examples of a non-unitary isometry in a separable Hilbert space.

(1) Let $\mathcal{H} = \ell^2$ $(= \ell^2(\mathbb{N})$). Let $\{e_n : n \in \mathbb{N}\}$ denote the standard orthonormal basis of $\mathcal{H}$ (consisting of sequences with a 1 in one co-ordinate and 0 in all other co-ordinates). In view of Exercise 1.7.7(1)(b), there exists a unique isometry $S \in B(H)$ such that $Se_n = e_{n+1}\ \forall n \in \mathbb{N}$; equivalently, we have

$$S(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots).$$

For obvious reasons, this operator is referred to as a 'shift' operator; in order to distinguish it from a near relative, we shall refer to it as the **unilateral shift**. It should be clear that S is an isometry whose range is the proper subspace $\mathcal{M} = \{e_1\}^\perp$, and consequently, S is not unitary.

A minor computation shows that the adjoint S^* is the 'backward shift':

$$S^*(\alpha_1, \alpha_2, \ldots) \;=\; (\alpha_2, \alpha_3, \ldots)$$

and that $SS^* \;=\; P_{\mathcal{M}}$ (which is another way of seeing that S is not unitary). Thus S^* is a left-inverse, but not a right-inverse, for S. (This, of course, is typical of a non-unitary isometry.)

Further – as is true for any non-unitary isometry – each power $S^n, n \geq 1$, is a non-unitary isometry.

(2) The 'near-relative' of the unilateral shift, which was referred to earlier, is the so-called **bilateral shift**, which is defined as follows: consider the Hilbert space $\mathcal{H} = \ell^2(\mathbb{Z})$ with its standard basis $\{e_n : n \in \mathbb{Z}\}$ for $\mathcal{H}$. The bilateral shift is the unique isometry B on $\mathcal{H}$ such that $Be_n \;=\; e_{n+1}\ \forall n \in \mathbb{Z}$. This time, however, since B maps the standard basis onto itself, we find that B is unitary. The reason for the terminology 'bilateral shift' is this: denote a typical element of $\mathcal{H}$ as a 'bilateral' sequence (or a sequence extending to infinity in both directions); in order to keep things straight, let us underline the 0-th co-ordinate of such a sequence; thus, if $x = \sum_{n=-\infty}^{\infty} \alpha_n e_n$, then we write $x = (\ldots, \alpha_{-1}, \underline{\alpha_0}, \alpha_1, \ldots)$; we then find that

$$B(\ldots, \alpha_{-1}, \underline{\alpha_0}, \alpha_1, \ldots) \;=\; (\ldots, \alpha_{-2}, \underline{\alpha_{-1}}, \alpha_0, \ldots)\ .$$

(3) Consider the Hilbert space $\mathcal{H} = L^2([0,1], m)$ (where, of course, m denotes 'Lebesgue measure') – see Remark 1.4.9(2) – and let $\{e_n : n \in \mathbb{Z}\}$ denote the exponential basis of this Hilbert space. Notice that $|e_n(x)|$ is identically equal to 1, and conclude that the operator defined by

$$(Wf)(x) \;=\; e_1(x)f(x)\ \forall f \in \mathcal{H}$$

is necessarily isometric; it should be clear that this is actually unitary, since its inverse is given by the operator of multiplication by e_{-1}.

It is easily seen that $We_n = e_{n+1}\ \forall n \in \mathbb{Z}$. If $U : \ell^2(\mathbb{Z}) \to \mathcal{H}$ is the unique unitary operator such that U maps the n-th standard basis vector to e_n, for each $n \in \mathbb{Z}$, it follows easily that $W = UBU^*$. Thus, the operator W of this example is unitarily equivalent to the bilateral shift (of the previous example).

More is true; let $\mathcal{M}$ denote the closed subspace $\mathcal{M} = [\{e_n : n \geq 1\}]$; then $\mathcal{M}$ is invariant under W – meaning that $W(\mathcal{M}) \subset \mathcal{M}$; and it should be clear that the restricted operator $W|_{\mathcal{M}} \in B(\mathcal{M})$ is unitarily equivalent to the unilateral shift.

(4) More generally, if $(X, \mathcal{B}, \mu)$ is any measure space and if $\phi : X \to \mathbb{C}$ is any measurable function such that $|\phi| = 1$ μ-a.e., then the equation

$$M_\phi f \;=\; \phi f \;, f \in L^2(X, \mathcal{B}, \mu)$$

defines a unitary operator on $L^2(X, \mathcal{B}, \mu)$ (with inverse given by $M_{\bar{\phi}}$). $\square$

Chapter 2

The Spectral Theorem

2.1 C^*-algebras

It will be convenient, indeed desirable, to use the language of C^*-algebras.

Definition 2.1.1. *A C^*-algebra is a Banach algebra $\mathcal{A}$ equipped with an adjoint operation $\mathcal{A} \ni S \mapsto S^* \in \mathcal{A}$ which satisfies the following conditions for all $S, T \in \mathcal{A}$*

$$\begin{aligned} (\alpha S_1 + S_2)^* &= \bar{\alpha} S_1^* + S_2^* \\ (S^*)^* &= S \\ (ST)^* &= T^* S^* \\ \|T\|^2 &= \|T^*T\| \ (\mathbf{C}^* - identity). \end{aligned}$$

All our C^*-algebras will be assumed to have a multiplicative identity, which is necessarily self-adjoint (as 1^* is also a multiplicative identity), and has norm one – thanks to the C^*-identity ($\|1\|^2 = \|1^*1\| = \|1\|$). (We ignore the trivial possibility $1 = 0$, i.e., $\mathcal{A} = \{0\}$.)

Example 2.1.2. (1) $B(\mathcal{H})$ is a C^*-algebra, and in particular $M_n(\mathbb{C})$ $\forall n$, so also $\mathbb{C} = M_1(\mathbb{C})$.

(2) Any norm-closed unital *-subalgebra of a C^*-algebra is also a C^*-algebra with the induced structure from the ambient C^*-algebra.

(3) For any subset S of a C^*-algebra, there is a smallest C^*-subalgebra of $\mathcal{A}$, denoted by $C^*(S)$, which contains S. (*Reason:* $C^*(S)$ may be defined somewhat uninformatively as the intersection of all C^*-subalgebras that contain S, and described more constructively as the norm-closure of the linear span of all 'words' in the alphabet $\{1\} \cup S \cup S^* := \{1\} \cup \{x : x \in S \text{ or } x^* \in S\}$.) The latter description in the previous sentence shows that $C^*(\{x\})$ is a *commutative* 'singly generated' C^*-subalgebra if and only if x satisfies $x^*x = xx^*$; such an element of a C^*-algebra, which commutes with its adjoint, is said to be **normal**.

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9_2

(4) If Σ is any compact space, then $C(\Sigma)$ is a commutative C^*-algebra – with respect to pointwise algebraic operations, $f^* = \bar{f}$ and $\|f\| = \sup\{|f(x)| : x \in \Sigma\}$. If $\Sigma \subset \mathbb{R}$ (resp., $\mathbb{C}$), then the Weierstrass polynomial approximation theorem (resp., the Stone-Weierstrass theorem) shows that $C(\Sigma)$ is a commutative unital C^*-algebra which is singly generated – with generator given by $f_0(t) = t \ \forall t \in \Sigma$.

DEFINITION 2.1.3. *A* **representation** *of a C^*-algebra $\mathcal{A}$ on a Hilbert space $\mathcal{H}$ is just a *-preserving unital algebra homomorphism of $\mathcal{A}$ into $B(\mathcal{H})$.*

Representations $\pi_i : \mathcal{A} \to B(\mathcal{H}_i)$, $i = 1, 2$, are said to be **equivalent** *if there exists a unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $\pi_2(a) = U\pi_1(a)U^* \ \forall a \in \mathcal{A}$.*

REMARK 2.1.4. It is true that any representation – and more generally, any unital *-algebra homomorphism between C^*-algebras – is contractive. This is essentially a consequence of (a) the C^*-identity, which shows that it suffices to check that $\|\pi(x)\| \leq \|x\| \ \forall x = x^*$ (b) the fact that the norm of a self-adjoint operator is its spectral radius, (see the last part of Proposition 1.5.6 (2)), and (c) the obvious fact that a unital homomorphism preserves invertibility and hence 'shrinks spectra'. Thus,

$$\begin{aligned} \|\pi(x)\|^2 &= \|\pi(x)^*\pi(x)\| = \|\pi(x^*x)\| \\ &= \operatorname{spr}(\pi(x^*x)) \leq \operatorname{spr}(x^*x) \leq \|x^*x\| = \|x\|^2. \end{aligned}$$

But we will not need this fact in this generality, so we shall say no more about it.

The observation that sets the ball rolling for us is Proposition 1.5.7.

PROPOSITION 2.1.5. *Let $\Sigma \subset \mathbb{R}$ be a compact set and let $f_0 \in C(\Sigma)$ be given by $f_0(t) = t \ \forall t \in \Sigma$. If $X \in B(\mathcal{H})$ is a self-adjoint operator such that $\sigma(X) \subset \Sigma$, then there exists a unique representation $\pi : C(\Sigma) \to B(\mathcal{H})$ such that $\pi(f_0) = X$. Conversely given any representation $\pi : C(\Sigma) \to B(\mathcal{H})$, it is the case that $\pi(f_0)$ is a self-adjoint operator X satisfying $\sigma(X) \subset \Sigma$.*

Proof. To begin with, if $X \in B(\mathcal{H})$ is a self-adjoint operator such that $\sigma(X) \subset \Sigma$, then it follows from the inequality (1.5.12) that $\|p(X)\|_{B(\mathcal{H})} \leq \|p\|_{C(\sigma(X))} \leq \|p\|_{C(\Sigma)}$ for any polynomial p. It is easily deduced now, from Weierstrass' theorem, that this mapping $\mathbb{C}[t] \ni p \mapsto p(X) \in B(\mathcal{H})$ extends uniquely to the desired *-homomorphism from $C(\Sigma)$ to $B(\mathcal{H})$.

Conversely, it is easily seen that $f_0 - \lambda$ is not invertible in $C(\Sigma)$ if and only if $\lambda \in \Sigma$ and as π preserves invertibility, we find that

$$\sigma(X) = \sigma(\pi(f_0)) \subset \sigma(f_0) = \Sigma$$

as desired. (Strictly speaking, we have only defined spectra of operators, while we are here talking of the spectra of elements of unital Banach algebras – $C(\Sigma)$, to be precise – but the definition is more or less the same.) $\square$

Remark 2.1.6. Representations $\pi_i : C(\Sigma) \to B(\mathcal{H}_i)$ are equivalent if and only if the operators $\pi_i(f_0), i = 1, 2$ are unitarily equivalent. This is because a representation of a singly generated C^*-algebra is uniquely determined by the image of the generator.

2.2 Cyclic representations and measures

Assume, for the rest of this book, that Σ is a separable compact metric space. Suppose $\pi : C(\Sigma) \to B(\mathcal{H})$ is a representation of $C(\Sigma)$ on a separable Hilbert space.

Definition 2.2.1. *A representation $\pi : C(\Sigma) \to B(\mathcal{H})$ is said to be* **cyclic** *if there exits a vector $x \in \mathcal{H}$ such that $\pi(C(\Sigma))x$ is a dense subspace of $\mathcal{H}$. In such a case, the vector x is called a* **cyclic vector** *for the representation. If such a vector exists, one can always find a unit vector which is cyclic for the representation.*

Before proceeding, it will be wise to spell out a trivial, but nevertheless very useful, observation.

Lemma 2.2.2. *If $\mathcal{S}_i = \{x_j^{(i)} : j \in \Lambda\}$ is a set which linearly spans a dense subspace of a Hilbert space $\mathcal{H}_i$ for $i = 1, 2$, and if $\langle x_j^{(1)}, x_k^{(1)}\rangle = \langle x_j^{(2)}, x_k^{(2)}\rangle$ for all $j, k \in \Lambda$, then there exists a unique unitary operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ such that $Ux_j^{(1)} = x_j^{(2)}\ \forall j \in \Lambda$.*

Proof. The hypotheses guarantee that the equation

$$U_0\Big(\sum_{\ell=1}^{n} \alpha_\ell x_{j_\ell}^{(1)}\Big) = \sum_{\ell=1}^{n} \alpha_\ell x_{j_\ell}^{(2)}$$

unambiguously defines a linear bijection U_0 between dense linear subspaces of the two Hilbert spaces preserving inner product, and hence extends uniquely to a unitary operator U with the desired property. Uniqueness of such a U follows from the fact that the difference between two such U's would have a dense linear subspace in its kernel. □

Proposition 2.2.3. *(1) If μ is a finite positive measure defined on the Borel subsets of Σ, then the equation*

$$(\pi_\mu(f))\,(g) = fg\ \forall f \in C(\Sigma),\ g \in L^2(\Sigma, \mu)$$

defines a cyclic representation π_μ of $C(\Sigma)$ with cyclic vector $g_0 \equiv 1$.

(2) Conversely, if $\pi : C(\Sigma) \to B(\mathcal{H})$ is a representation with a cyclic vector x, then there exists a finite positive measure μ defined on the Borel subsets of Σ and a unitary operator $U : \mathcal{H} \to L^2(\Sigma, \mu)$ such that $Ux = g_0$ and $U\pi(f)U^* = \pi_\mu(f)\ \forall f \in C(\Sigma)$.

(3) *In the setting of (1) above, there exists a unique representation* $\widetilde{\pi_\mu} : L^\infty(\Sigma, \mu) \to B(L^2(\Sigma, \mu))$ *such that (i)* $\widetilde{\pi_\mu}|_{C(\Sigma)} = \pi_\mu$*, and (ii) if* $\{f_n : n \in \mathbb{N}\}$ *is such that* $\sup_n \|f_n\|_{L^\infty(\Sigma,\mu)} < \infty$ *and* $f_n \to f \quad \mu$*-a.e., then* $\|\widetilde{\pi_\mu}(f_n)g - \widetilde{\pi_\mu}(f)g\|_{L^2(\Sigma,\mu)} \to 0 \ \forall g \in L^2(\Sigma, \mu)$.

Further, the measure μ is a probability measure precisely when the cyclic vector x is a unit vector.

Proof. (1) It is fairly clear that $C(\Sigma) \ni f \mapsto \pi_\mu(f) \in B(\mathcal{H})$ is a representation of $C(\Sigma)$ and $\|\pi_\mu(f)\|_{B(L^2(\Sigma,\mu))} \leq \|f\|_{L^\infty(\Sigma,\mu)}$. Clearly each $\pi_\mu(f)$ is normal, and it follows from Theorem 1.6.2 (2) that

$$\lambda \in \mathrm{sp}\,(\pi_\mu(f)) \ \Leftrightarrow \ \mu(\{w \in \Sigma : |f(w) - \lambda| < \epsilon\}) > 0 \ \forall \epsilon > 0,$$

and in particular, $\|\pi_\mu(f)\| = spr(\pi_\mu(f)) = \|f\|_{L^\infty(\Sigma,\mu)} \leq \|f\|_{C(\Sigma)}$. it is a basic fact from measure theory – see Lemma A1 in the Appendix – that g_0 is indeed a cyclic vector for the representation π_μ.

(2) Consider the functional $\phi : C(\Sigma) \to \mathbb{C}$ defined by $\phi(f) = \langle \pi(f)x, x\rangle$. It is clear that if $f \in C(\Sigma)$ is non-negative, then also $f^{\frac{1}{2}} \in C(\Sigma)$ is non-negative and, in particular, real-valued, and hence

$$\phi(f) = \langle \pi(f^{\frac{1}{2}})x, \pi(f^{\frac{1}{2}})x\rangle \geq 0.$$

Thus ϕ is a positive – and clearly bounded – linear functional on $C(\Sigma)$, and the **Riesz representation theorem** – which identifies the dual space of $C(\boldsymbol{\Sigma})$ with the set $M(\Sigma)$ of finite complex measures – guarantees the existence of a positive measure μ defined on the Borel sets of Σ such that $\phi(f) = \int f \, d\mu$. It follows that for arbitrary $f, g \in C(\Sigma)$, we have

$$\begin{aligned} \langle \pi(f)x, \pi(g)x\rangle &= \langle \pi(\bar{g}f)x, x\rangle \\ &= \phi(\bar{g}f) \\ &= \int \bar{g}f \, d\mu \\ &= \langle \pi_\mu(f)g_0, \pi_\mu(g)g_0\rangle \ . \end{aligned}$$

An appeal to Lemma 2.2.2 now shows that there exists a unitary operator $U : \mathcal{H} \to L^2(\Sigma, \mu)$ such that $U\pi(f)x = \pi_\mu(f)g_0 \ \forall f \in C(\Sigma)$. Setting $f = 1$, we find that $Ux = g_0$. And for all $g \in C(\Sigma)$, we see that $U\pi(f)U^*\pi_\mu(g)g_0 = U\pi(f)\pi(g)x = \pi_\mu(f)\pi_\mu(g)x_0$ with the result that, indeed, $U\pi(f)U^* = \pi_\mu(f)$, completing the proof of (2).

(3) Simply define $\widetilde{\pi_\mu}(\phi)g = \phi g \ \forall g \in L^2(\Sigma, \mu)$. Then (i) is clearly true, while (ii) is just a restatement of the bounded convergence theorem of measure theory. The uniqueness assertion regarding $\widetilde{\pi_\mu}$ follows from the demanded (i) and Lemma A2 in the Appendix.

$\square$

It would make sense to introduce a definition and a notation for a notion that has already been encountered more than once.

Definition 2.2.4. *A sequence $\{X_n : n \in \mathbb{N}\}$ in $B(\mathcal{H})$ is said to converge in the* **strong operator topology** *– henceforth abbreviated to SOT – if $\{X_n x : n \in \mathbb{N}\}$ converges in the norm of $\mathcal{H}$ for every $x \in \mathcal{H}$. It is a consequence of the 'uniform boundedness principle' that in this case, the equation*

$$Xx = \lim_{n\to\infty} X_n x$$

defines a bounded operator $X \in B(\mathcal{H})$. We shall abbreviate all this by writing $X_n \stackrel{SOT}{\longrightarrow} X$.

We record a couple of simple but very useful facts concerning SOT convergence. But first, recall that a set $\mathcal{S} \subset \mathcal{H}$ is said to be **total** if the linear subspace spanned by $\mathcal{S}$ is dense in $\mathcal{H}$. (eg: any orthonormal basis (onb) is total.)

Lemma 2.2.5. *(1) The following conditions on a sequence $\{X_n : n \in \mathbb{N}\} \subset B(\mathcal{H})$ are equivalent:*

(a) $X_n \stackrel{SOT}{\longrightarrow} X$ for some $X \in B(\mathcal{H})$;

(b) $\sup_n \|X_n\| < \infty$, and there exists some total set $\mathcal{S} \subset \mathcal{H}$ such that $X_n x$ converges for all $x \in \mathcal{S}$;

(c) $\sup_n \|X_n\| < \infty$, and there exists a dense subspace $\mathcal{M} \subset \mathcal{H}$ such that $X_n x$ converges for all $x \in \mathcal{M}$.

(2) If sequences $X_n \stackrel{SOT}{\longrightarrow} X$ and $Y_n \stackrel{SOT}{\longrightarrow} Y$ in $B(\mathcal{H})$, then also $X_n Y_n \stackrel{SOT}{\longrightarrow} XY$.

Proof. (1) The implication $(a) \Rightarrow (b)$ follows from the uniform boundedness principle, while $(b) \Rightarrow (c)$ is seen on setting $\mathcal{M} = \bigvee \mathcal{S}$, the vector subspace spanned by $\mathcal{S}$. As for $(c) \Rightarrow (a)$, if $\sup_n \|X_n\| < K (> 0)$, note that the equation $Xx = \lim_n X_n x$ defines a linear map $X : \mathcal{M} \to \mathcal{H}$ with $\|Xx\| \leq K\|x\| \; \forall x \in \mathcal{M}$; the assumed density of $\mathcal{M}$ ensures that X admits a unique extension to an element of $B(\mathcal{H})$, also denoted by X, with $\|X\| \leq K$ and $X_n x \to Xx \; \forall x \in \mathcal{M}$. Now, if $x \in \mathcal{H}$ and $\epsilon > 0$, choose $x' \in \mathcal{M}$ such that $\|x - x'\| < \epsilon/3K$, then choose an $n_0 \in \mathbb{N}$ such that $\|(X_n - X)x'\| < \epsilon/3 \; \forall n \geq n_0$ and compute thus, for $n \geq n_0$:

$$\begin{aligned} \|(X_n - X)x\| &\leq \|(X_n - X)(x - x')\| + \|(X_n - X)x'\| \\ &< (2K)\frac{\epsilon}{3K} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned}$$

(2) Begin by deducing from the uniform boundedness principle that there exists a constant $K > 0$ such that $\|X_n\| \leq K$ and $\|Y_n\| \leq K$ for all n. Fix $x \in \mathcal{H}$ and an $\epsilon > 0$. Under the hypotheses, we can find an $n_0 \in \mathbb{N}$ such

that $\|(Y_n - Y)x\| < \epsilon/2K$ and $\|(X_n - X)Yx\| < \epsilon/2$ for all $n \geq n_0$. We then see that for every $n \geq n_0$

$$\begin{aligned} \|(X_nY_n - XY)x\| &= \|(X_nY_n - X_nY + X_nY - XY)x\| \\ &\leq \|X_n(Y_n - Y)x\| + \|(X_n - X)Yx\| \\ &< \epsilon, \end{aligned}$$

thus proving that indeed $X_nY_n \stackrel{SOT}{\longrightarrow} XY$.

□

The following important consequence of Proposition 2.2.3 is 'one half' of the celebrated Hahn-Hellinger classification of separable representations of $C(\Sigma)$. (See Remark 2.3.3.)

THEOREM 2.2.6. *If $\pi : C(\Sigma) \to B(\mathcal{H})$ is a representation on a separable Hilbert space $\mathcal{H}$, there exists a countable collection $\{\mu_n : n \in N\}$ (for some countable set N) of probability measures defined on the Borel-σ-algebra $\mathcal{B}_\Sigma$ such that π is (unitarily) equivalent to $\oplus\pi_{\mu_n} : C(\Sigma) \to B\big(\bigoplus L^2(\Sigma, \mu_n)\big)$.*

Proof. Note that $\mathcal{H}$ is separable, as is the Hilbert space underlying any cyclic representation of $C(\Sigma)$ (since the latter is separable). Also observe that $\pi(C(\Sigma))$ is closed under adjoints, as a consequence of which, if a subspace of $\mathcal{M} \subset \mathcal{H}$ is left invariant by the entire *-algebra $\pi(C(\Sigma))$, then so is $\mathcal{M}^\perp$. It follows from the previous sentence and a simple use of Zorn's lemma, that there exists a countable (possibly finite) collection $\{x_n : n \in N\}$ (for some countable set N) of unit vectors such that $\mathcal{H} = \bigoplus_{n\in N} \overline{(\pi(C(\Sigma))x_n)}$. Clearly each $\mathcal{M}_n = \overline{(\pi(C(\Sigma))x_n)}$ is a closed subspace that is invariant under the algebra $\pi(C(\Sigma))$ and yields a cyclic subrepresentation $\pi_n(\cdot) = \pi(\cdot)|_{\mathcal{M}_n}$. It follows from Proposition 2.2.3 (2) that

$$\pi = \bigoplus\nolimits_{n\in N} \pi_n \sim \bigoplus \pi_{\mu_n} \ ,$$

for the probability measures given by

$$\int_\Sigma f \, d\mu_n = \langle \pi(f)x_n, x_n \rangle \ .$$

□

LEMMA 2.2.7. *In the notation of Proposition 2.2.3(3), the following conditions on a bounded sequence $\{f_n\}$ in $L^\infty(\mu)$ are equivalent:*

(1) *the sequence $\{f_n\}$ converges in (μ-) measure to 0;*

(2) $\widetilde{\pi_\mu}(f_n) \stackrel{SOT}{\longrightarrow} 0$.

Proof. (1) ⇒ (2): This is an immediate consequence of a version of the dominated convergence theorem.

(2) $\Rightarrow$ (1): Since the constant function $g_0 \equiv 1$ belongs to $L^2(\Sigma, \mu)$, it follows from the inequality

$$\begin{aligned} \mu(\{|f_n - f| \geq \epsilon\}) &\leq \epsilon^{-2} \int_{\{|f_n - f| \geq \epsilon\}} |f_n - f|^2 d\mu \\ &\leq \epsilon^{-2} \int |f_n - f|^2 d\mu \end{aligned}$$

that indeed $\mu(\{|f_n - f| \geq \epsilon\}) \to 0 \ \forall \epsilon > 0$. □

Theorem 2.2.8. *Let $\pi : C(\Sigma) \to B(\mathcal{H})$ and $\{\mu_n : n \in N\}$ be as given in Theorem 2.2.6. Choose some set $\{\epsilon_n : n \in N\}$ of strictly positive numbers such that $\sum_{n \in N} \epsilon_n = 1$, and define the probability measure μ on $(\Sigma, \mathcal{B}_\Sigma)$ by $\mu = \sum_{n \in N} \epsilon_n \mu_n$. Then, we have:*

(1) For $E \in \mathcal{B}_\Sigma$ we have $\mu(E) = 0 \Leftrightarrow \mu_n(E) = 0 \ \forall n \in N$. Further, $\phi \in L^\infty(\Sigma, \mu) \Rightarrow \phi \in L^\infty(\Sigma, \mu_n) \ \forall n \in N$ and $\sup_n \|\phi\|_{L^\infty(\mu_n)} = \|\phi\|_{L^\infty(\mu)}$.

(2) The equation $\tilde{\pi} = \bigoplus_{m \in N} \widetilde{\pi_{\mu_m}}$ defines an isometric representation $\tilde{\pi} : L^\infty(\Sigma, \mu) \to B(\mathcal{H})$ such that the following conditions on a uniformly norm-bounded sequence $\{\phi_n : n \in \mathbb{N}\}$ in $L^\infty(\mu)$ are equivalent:

(a) $\phi_n \to 0$ in measure w.r.t. μ.

(b) $\phi_n \to 0$ in measure w.r.t. μ_m for all m.

(c) $\tilde{\pi}(\phi_n) \xrightarrow{SOT} 0$.

Proof. Before proceeding with the proof, we wish to underline the (so far unwritten) convention that we use throughout this book: we treat elements of different L^p-spaces as if they were functions (rather than equivalence classes of functions agreeing almost everywhere.).

(1) Since $\epsilon_n > 0 \ \forall n \in N$, it follows that $\mu(E) = 0 \Leftrightarrow \mu_n(E) = 0 \ \forall n \in N$.

Since a countable union of null sets is also a null set, it is clear that if $\phi \in L^\infty(\mu)$, we may find a μ-null set F which will satisfy the condition $\|\phi\|_{L^\infty(\mu)} = \sup\{|\phi(\lambda)| : \lambda \in \Sigma \setminus F\}$. For $E \in \mathcal{B}_\Sigma$ we have $\mu(E) = 0 \Leftrightarrow \mu_n(E) = 0 \ \forall n \in N$. Let $F_n = F \cup \{\frac{d\mu_n}{d\mu} = 0\}$. Clearly,

$$\mu_n(\textstyle\bigcap_{m \in N} F_m) \leq \mu_n(F_n) = 0 \ \forall n$$

so, also $\mu(\bigcap_{m \in N} F_m) = 0$. Since $(\bigcap_{m \in N} F_m) \supset F$, we see thus that

$$\begin{aligned} \|\phi\|_{L^\infty(\mu)} &= \sup\left\{|\phi(\lambda)| : \lambda \in \Sigma \setminus \textstyle\bigcap_{m \in N} F_m\right\} \\ &= \sup\left\{|\phi(\lambda)| : \lambda \in \textstyle\bigcup_{m \in N} (\Sigma \setminus F_m)\right\} \\ &= \sup_m \ \sup\{|\phi(\lambda)| : \lambda \in \Sigma \setminus F_m\} \\ &= \sup_m \|\phi\|_{L^\infty(\mu_m)} . \end{aligned}$$

(2) If $\phi \in L^\infty(\mu)$, then

$$\begin{aligned}\|\tilde{\pi}(\phi)\| &= \sup_{m\in N} \|\widetilde{\pi_{\mu_m}}(\phi)\| \\ &= \sup_{m\in N} \|\phi\|_{L^\infty(\mu_m)} \\ &= \|\phi\|_{L^\infty(\mu)} \qquad \text{by part (1) of this Theorem}\end{aligned}$$

so $\tilde{\pi}$ is indeed an isometry.

Suppose $\sup_{n\in\mathbb{N}} \|\phi_n\|_{L^\infty(\mu)} \leq C < \infty$.

$(a) \Rightarrow (b)$: This follows immediately from $\mu_m \leq \epsilon_m^{-1}\mu$.

$(b) \Rightarrow (a)$: Let $\delta, \epsilon > 0$. We assume, for this proof, that the index set N is the whole of $\mathbb{N}$; the case of finite N is trivially proved. First choose $N' \in N$ such that $\sum_{m=N'+1}^\infty \epsilon_m < \epsilon/2$. Then choose an n_0 so large that $n \geq n_0 \Rightarrow \mu_m(\{|\phi_n| > \delta\}) < \epsilon/2N'\epsilon_m$; and conclude that for an $n \geq n_0$, we have

$$\begin{aligned}\mu(\{|\phi_n| > \delta\}) &\leq \sum_{m=1}^{N'} \epsilon_m \mu_m(\{|\phi_n| > \delta\}) + \sum_{m=N'+1}^{\infty} \epsilon_m \\ &< \sum_{m=1}^{N'} \epsilon_m \frac{\epsilon}{2N'\epsilon_m} + \frac{\epsilon}{2} \\ &= \epsilon.\end{aligned}$$

$(b) \Rightarrow (c)$: As $\|\tilde{\pi}(\phi_n)\| \leq C \; \forall n$, using Lemma 2.2.5 it is enough to prove that $\lim_{n\to\infty} \tilde{\pi}(\phi_n)x = 0$ whenever $x = ((x_m)) \in \bigoplus_{m=1}^\infty L^2(\mu_m)$ is such that $x_m = 0 \; \forall \; m \neq k$ for some one k. By Lemma 2.2.7, the condition (b) is seen to imply that $\|\widetilde{\pi_{\mu_k}}(\phi_n)x_k\| \to 0$; but $\|\tilde{\pi}(\phi_n)x\| = \|\widetilde{\pi_{\mu_k}}(\phi_n)x_k\|$ and we are done.

$(c) \Rightarrow (b)$: If condition (c) holds, it can be seen by restricting to the subspace $L^2(\mu_m)$ that $\widetilde{\pi_{\mu_m}}(\phi_n) \xrightarrow{SOT} 0$, and it now follows by applying Lemma 2.2.7 that the sequence $\phi_n \to 0$ in measure with respect to μ_m for each $m \in \mathbb{N}$.

□

2.3 Spectral Theorem for self-adjoint operators

Throughout this section, we shall assume that $X \in B(\mathcal{H})$ is a self-adjoint operator and that $\Sigma = \sigma(X)$. In the interest of minimising parentheses, we shall simply write $C^*(X)$ rather than $C^*(\{X\})$ for the (unital) C^*-algebra generated by X. As advertised in the preface, we shall prove the following formulation of what we would like to think of as the spectral theorem.

Theorem 2.3.1. *[Spectral theorem for self-adjoint operators]*

(1) **(Continuous Functional Calculus)** *There exists a unique isometric *-algebra isomorphism*

$$C(\Sigma) \ni f \mapsto f(X) \in C^*(X)$$

of $C(\Sigma)$ *onto* $C^*(X)$ *such that* $f_0(X) = X$.[1]

(2) **(Measurable Functional Calculus)** *There exists a measure* μ *defined on* $\mathcal{B}_\Sigma$ *and a unique isometric *-algebra homomorphism*

$$L^\infty(\Sigma, \mu) \ni f \mapsto f(X) \in B(\mathcal{H})$$

of $L^\infty(\Sigma, \mu)$ *into* $B(\mathcal{H})$ *such that (i)* $f_0(X) = X$, *and (ii) a norm-bounded sequence* $\{f_n : n \in \mathbb{N}\}$ *in* $L^\infty(\Sigma, \mu)$ *converges in measure w.r.t.* μ *(to* f, *say) if and only if the sequence* $\{f_n(X) : n \in \mathbb{N}\}$ *SOT converges (to* $f(X)$*).*

Proof. (1) It follows from Proposition 2.1.5 that there exists a unique representation $\pi : C(\Sigma) \to B(\mathcal{H})$ such that $\pi(f_0) = X$. As for the 'isometry' assertion, observe that for any $p \in \mathbb{C}[t]$, the spectral mapping theorem ensures that

$$\|\pi(p)\|_{B(\mathcal{H})} = \mathrm{spr}(p(X)) = \|p\|_\Sigma = \|p\|_{C(\Sigma)}$$

and the Weierstrass approximation theorem now guarantees that

$$\|\pi(f)\|_{B(\mathcal{H})} = \|f\|_{C(\Sigma)} \ \forall f \in C(\Sigma)$$

as desired.

(2) As $f \mapsto f(X)$ is a representation, say π, of $C(\Sigma)$, if μ and $\tilde{\pi}$ are as in Theorem 2.2.8 (2), the equation $\tilde{\pi}(\phi) = \phi(X)$ defines a measurable functional calculus with the desired properties. Thanks to Lemma A2 in the Appendix, there can be at most one isometric (unital) *-homomorphism of $L^\infty(\Sigma, \mu)$ into $B(\mathcal{H})$, i.e., a measurable functional calculus, which (i) extends the continuous functional calculus π and (ii) maps uniformly bounded sequences converging in measure (w.r.t. μ) to SOT convergent sequences.' So we see that there indeed exists a unique *-homomorphism from $L^\infty(\Sigma, \mu)$ into $B(\mathcal{H})$ with the desired property.

□

Corollary 2.3.2. *If* $\mu_i, i = 1, 2$ *are two probability measures satisfying the conditions imposed on* μ *in Theorem* 2.3.1, *then* μ_1 *and* μ_2 *are mutually absolutely continuous. In particular, the Banach algebra* $L^\infty(\boldsymbol{\Sigma}, \mu)$ *featuring in Theorem* 2.3.1 (2) *is uniquely determined by the operator* X, *even if* μ *itself is not.*

[1] Recall that $\Sigma \subset \mathbb{R}$ – see Corollary 1.6.3 – and that f_0 denotes the function $f_0 : \Sigma \to \mathbb{R}$ defined by $f_0(t) = t$.

Proof. Suppose $\pi_i : L^\infty(\Sigma, \mu_i) \to B(\mathcal{H}), i = 1, 2$ are isometric *-isomorphisms which (i) extend the continuous functional calculus (call it $\pi : C(\Sigma) \to C^*(\{X\})$), and (ii) satisfy the convergence in measure – sequential SOT convergence homeomorphism property as in part (2) of Theorem 2.3.1. Define $\nu = (\mu_1 + \mu_2)/2$. Then convergence in measure w.r.t ν implies convergence in measure w.r.t. μ_i for $1 = 1, 2$ since $\mu_i \leq 2\nu$.

Suppose $\mu_1(E) = 0$ for some $E \in \mathcal{B}_{\mathbb{C}}$.

Then appeal to Lemma A2 of the Appendix to find a sequence $\{f_n : n \in \mathbb{N}\} \subset C(\Sigma)$ such that $\|f_n\| \leq 1$ ν-a.e. and such that $f_n \to 1_E$ in measure w.r.t. ν. Then also $f_n \to 1_E$ in measure w.r.t. $\mu_i, 1 = 1, 2$. Then the assumptions imply that

$$\begin{aligned} \pi_2(1_E) &= SOT - \lim_n \pi_2(f_n) \\ &= SOT - \lim_n \pi_1(f_n) \\ &= \pi_1(1_E) \\ &= 0 \end{aligned}$$

and hence $\mu_2(E) = 0$. By toggling the roles of 1 and 2, we find that μ_1 and μ_2 are mutually absolutely continuous, thereby proving the corollary.

The last assertion is an off-shoot of the statement that the 'identity map' is an isometric isomorphism between L^∞ spaces of mutually absolutely continuous probability measures. □

Remark 2.3.3. (1) Our proof of the spectral theorem, for self-adjoint operators, actually shows that if Σ is a compact metric space and $\pi : C(\Sigma) \to B(\mathcal{H})$ is a representation, i.e., a unital *-homomorphism, *on a separable Hilbert space*, there exists a probability measure μ defined on $\mathcal{B}_\Sigma$ – which is unique up to mutual absolute continuity – and a representation $\tilde{\pi} : L^\infty(\mu) \to B(\mathcal{H})$ which is uniquely determined by (i) $\tilde{\pi}$ 'extends' π, and (ii) a norm-bounded sequence $\{f_n : n \in \mathbb{N}\} \subset L^\infty(\mu)$ converges to 0 in (μ) measure if and only if $\tilde{\pi}(f_n)$ SOT-converges to 0.

(2) Further, if π is isometric, so is $\tilde{\pi}$ and in particular, if U is a non-empty open set in Σ, then $\mu(U) \neq 0$, or equivalently $\tilde{\pi}(1_U) \neq 0$.

(3) All this is part of the celebrated **Hahn-Hellinger theorem** which says: the representation π is determined up to unitary equivalence by the measure class (w.r.t. mutual absolute continuity) of μ and a measurable **spectral multiplicity function** $m : \Sigma \to (\{0\} \cup \bar{\mathbb{N}}) := \{0, 1, 2, \cdots, \aleph_0\}$, which is determined uniquely up to sets of μ measure zero; in fact if $E_n = m^{-1}(n),\ n \in \{0\} \cup \bar{\mathbb{N}}$, then π is unitarily equivalent to the representation on $\bigoplus_{n \in \bar{\mathbb{N}}} L^2(E_n, \mu|_{E_n}) \otimes \mathcal{H}_n$ given by $\bigoplus_{n \in \bar{\mathbb{N}}} \pi_{\mu|_{E_n}} \otimes \mathrm{id}_{\mathcal{H}_n}$, where $\mathcal{H}_n$ is some (multiplicity) Hilbert space of dimension n.

2.4 The spectral subspace for an interval

This section is devoted to a pretty and useful characterisation, from [Hal], of the spectral subspace for the unit interval. We first list some simple facts concerning spectral subspaces (= ranges of spectral projections). We use the following notation below: for a self-adjoint operator X, let $\mathcal{M}_X(E) = \operatorname{ran} 1_E(X)$. We also use the as yet undefined notions (but only the definition) of order and positivity – see Proposition 2.8.12, especially part (b) and the final paragraph in it – in the following Proposition.

PROPOSITION 2.4.1. *Let $X \in B(\mathcal{H})$ be self-adjoint. Then,*

(1) $a\|x\|^2 \leq \langle Xx, x\rangle \leq b\|x\|^2 \ \forall x \in \mathcal{M}_X([a,b])$;

(2) $X1_{[0,\infty)}(X) \geq 0$;

(3) $\epsilon > 0, x \in \mathcal{M}_X(\mathbb{R} \setminus (t_0 - \epsilon, t_0 + \epsilon)) \Rightarrow \|(X - t_0)x\| \geq \epsilon\|x\|$;

(4) $t_0 \in \sigma(X) \Leftrightarrow \mathcal{M}_X((t_0 - \epsilon, t_0 + \epsilon)) \neq \{0\} \ \forall \epsilon > 0$; *and*

(5) $\mathcal{M}_X(\{t_0\}) = \ker(X - t_0)$.

Proof. (1) Notice first that *-homomomorphisms of C^*-algebras are *order-preserving* since

$$\begin{aligned} x \leq y \quad &\Rightarrow \quad y - x \geq 0 \ \ (i.e., \exists z \text{ such that } y - x = z^*z) \\ &\Rightarrow \quad \pi(y) - \pi(x) = \pi(y - x) = \pi(z)^*\pi(z) \geq 0 \\ &\Rightarrow \quad \pi(x) \leq \pi(y) \ . \end{aligned}$$

Hence

$$a1_{[a,b]}(t) \leq t1_{[a,b]}(t) \leq b1_{[a,b]}(t) \Rightarrow a1_{[a,b]}(X) \leq X1_{[a,b]}(X) \leq b1_{[a,b]}(X)$$

and the desired result follows from the fact that $1_{[a,b]}(X)x = x \ \forall x \in \mathcal{M}_X([a,b])$.

(2) This follows from (1) since $1_{[0,\infty)}(X) = 1_{[0,\|X\|]}(X)$.

(3) It follows from (1) that if $x \in \mathcal{M}_X(\mathbb{R} \setminus (t_0 - \epsilon, t_0 + \epsilon) = \mathcal{M}_{(X-t_0)^2}([\epsilon^2, \infty))$ (by the spectral mapping theorem), then $\epsilon^2\|x\|^2 \leq \langle (X - t_0)^2 x, x\rangle = \|(X - t_0)x\|^2$.

(4) If μ is as in Theorem 2.3.1 (2), observe that

$$\begin{aligned} t_0 \notin \sigma(X) \quad &\Leftrightarrow \quad (X - t_0) \in GL(\mathcal{H}) \\ &\Leftrightarrow \quad (f_0 - t_0) \text{ is invertible in } L^\infty(\sigma(X), \mu) \\ &\Leftrightarrow \quad \exists \epsilon > 0 \text{ such that } |f_0 - t_0| \geq \epsilon \ \mu - a.e. \\ &\Leftrightarrow \quad \exists \epsilon > 0 \text{ such that } \mu((t_0 - \epsilon, t_0 + \epsilon)) = 0 \\ &\Leftrightarrow \quad \exists \epsilon > 0 \text{ such that } \mathcal{M}_X((t_0 - \epsilon, t_0 + \epsilon)) = \{0\} \ . \end{aligned}$$

(5) Clearly X commutes with $1_E(X)$ $\forall X$ and hence the subspace $\mathcal{M}_X(E)$ is invariant under X for all E, X. It follows from (1) above that $\langle X_0 x, x\rangle = t_0\|x\|^2$ $\forall x \in \mathcal{M}_X\{t_0\}$ where $X_0 = X|_{\mathcal{M}_X(\{t_0\})}$ and hence $\ker(X - t_0) \supset \mathcal{M}_X(\{t_0\})$. Conversely, if $x \in$ ker $(X - t_0)$, then for any $\epsilon > 0$, we have $(X - t_0)1_{\mathbb{R}\setminus(t_0-\epsilon,t_0+\epsilon)}(X)(x) = 1_{\mathbb{R}\setminus(t_0-\epsilon,t_0+\epsilon)}(X)(X - t_0)(x) = 0$, whence $1_{\mathbb{R}\setminus(t_0-\epsilon,t_0+\epsilon)}(X)x = 0$ by (3) above. So $x \in M_X(t_0 - \epsilon, t_0 + \epsilon)$; since ϵ was arbitrary, we have $x \in \bigcap_{\epsilon>0} M_X(t_0 - \epsilon, t_0 + \epsilon) = M_X(\{t_0\})$, so indeed ker $(X - t_0) = M_X(\{t_0\})$.

□

Now we come to the much advertised pretty description by Halmos of $\mathcal{M}_X([-1, 1])$.

PROPOSITION 2.4.2. *Let $X = X^*$ be as above, and let $x \in \mathcal{H}$. The following conditions are equivalent:*

(1) $x \in \mathcal{M}_X([-1, 1])$.

(2) $\|X^n x\| \leq \|x\|$ $\forall n \in \mathbb{N}$.

(3) $\{\|X^n x\| : n \in \mathbb{N}\}$ *is a bounded set.*

Proof. (1) $\Rightarrow$ (2): The operator X leaves the subspace $\mathcal{M}_X([-1, 1])$ invariant, and its restriction X_1 to this spectral subspace satisfies $-1 \leq X_1 \leq 1$ (by Proposition 2.4.1(1) and hence $\|X_1\| = \mathrm{spr}(X_1) \leq 1$ whence also $\|X_1^n\| \leq 1$, as desired.

(2) $\Rightarrow$ (3) is obvious.

(3) $\Rightarrow$ (1): If we let $x_1 = 1_{[-1,1]}(X)x$, we need to show that $x = x_1$; for this, note that

$$\begin{aligned} x - x_1 &= (1 - 1_{[-1,1]}(X))x \\ &= 1_{\mathbb{R}\setminus[-1,1]}(X)x \\ &= \lim_{n\to\infty} 1_{\mathbb{R}\setminus I_n}(X)x \end{aligned}$$

where we write the symbol I_n to denote the interval $(-1 - \frac{1}{n}, 1 + \frac{1}{n})$; so it suffices to show that $1_{\mathbb{R}\setminus I_n}(X)x = 0$ $\forall n$. Indeed, if there exists some n such that $y_n = 1_{\mathbb{R}\setminus I_n}(X)x \neq 0$, it would follow from Proposition 2.4.1 (3) that $\|Xy_n\| \geq (1+\frac{1}{n})\|y_n\|$ and that hence $\|X^m x\| \geq \|1_{\mathbb{R}\setminus I_n}(X)X^m x\| = \|X^m y_n\| \geq (1 + \frac{1}{n})^m\|y_n\|$. So the sequence $\{\|X^m x\| : m \in \mathbb{N}\}$ is not a bounded set if any $y_n \neq 0$. □

COROLLARY 2.4.3. *(1) $x \in \mathcal{M}_X([t_0 - \epsilon, t_0 + \epsilon]) \Leftrightarrow \{\left(\frac{X-t_0}{\epsilon}\right)^n x : n \in \mathbb{N}\}$ is bounded.*

(2) If a $Y \in B(\mathcal{H})$ commutes with X, ie., $YX = XY$, then Y leaves $\mathcal{M}_X(I)$ invariant for every bounded interval I.

Proof. (1) This follows by applying Proposition 2.4.2 to $(X - t_0)/\epsilon$ rather than to the operator X.

(2) If $YX = XY$ and if I is a compact interval (which can always be written in the form $[t_0 - \epsilon, t_0 + \epsilon]$), it follows from (1) above that

$$\begin{aligned} x \in \mathcal{M}_X([t_0 - \epsilon, t_0 + \epsilon]) &\Rightarrow \left\{ \left(\frac{X - t_0}{\epsilon} \right)^n x : n \in \mathbb{N} \right\} \text{ is bounded} \\ &\Rightarrow \left\{ Y \left(\frac{X - t_0}{\epsilon} \right)^n x : n \in \mathbb{N} \right\} \text{ is bounded} \\ &\Rightarrow \left\{ \left(\frac{X - t_0}{\epsilon} \right)^n Yx : n \in \mathbb{N} \right\} \text{ is bounded} \\ &\Rightarrow Yx \in \mathcal{M}_X([t_0 - \epsilon, t_0 + \epsilon]), \end{aligned}$$

so Y leaves the spectral subspaces corresponding to compact intervals invariant.

If I is an open interval, there exist an increasing sequence $\{I_n : n \in \mathbb{N}\}$ of compact intervals such that $I = \bigcup_{n\in\mathbb{N}} I_n$. But then $1_I(X) = SOT - \lim_{n\to\infty} 1_{I_n}(X)$ and $\mathcal{M}_X(I) = \overline{(\bigcup_n \mathcal{M}_X(I_n))}$. The previous paragraph shows that Y leaves each $\mathcal{M}_X(I_n)$, and hence also $\mathcal{M}(I)$, invariant.

Similar approximation arguments can be conjured up if I is of the form $[a, b)$ or $(a, b]$. (For example, $[a, b - \frac{1}{n}] \uparrow [a, b)$ and $[a + \frac{1}{n}, b] \uparrow (a, b]$.) $\square$

2.5 Finitely many commuting self-adjoint operators

We assume in the rest of this chapter that $X_1, \ldots, X_n, \ldots$ are commuting self-adjoint operators on $\mathcal{H}$.

DEFINITION 2.5.1. *Consider the set $\mathbf{\Sigma_k} = \Sigma(X_1, \ldots, X_k)$ consisting of those $(\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$ for which there exists a sequence $\{x_n : n \in \mathbb{N}\}$ of unit vectors in $\mathcal{H}$ such that $\lim_{n\to\infty} \|(X_i - \lambda_i)x_n\| = 0$ for $1 \le i \le k$. Thus $\mathbf{\Sigma_k}$ consists of k-tuples of scalars which admit a sequence of 'simultaneous approximate eigenvectors' of the X_i's, and will be referred to simply as the* **joint spectrum of** *$X_1, \ldots, X_k$.*

If $(\lambda_1, \ldots, \lambda_k) \in \mathbf{\Sigma_k}$, it is clear that $\lambda_i \in \sigma(X_i)$ for $1 \le i \le k$, and in particular $\mathbf{\Sigma_k} \subset \prod_{i=1}^k \sigma(X_i)$ and is hence bounded.

LEMMA 2.5.2. *(1) $\mathbf{\Sigma_k}$ is a compact set for $k > 0$; and*

(2) If $k > 0$ then $\mathrm{pr}_k(\mathbf{\Sigma_k}) = \Sigma(X_k)$, where $\mathrm{pr}_k : \mathbb{R}^k \to \mathbb{R}$ denotes the projection onto the k-th coordinate; in particular, $\mathbf{\Sigma_k} \ne \emptyset$.

Proof. (1) We have already seen above that $\mathbf{\Sigma_k}$ is bounded, so we only need to prove that it is closed. So suppose $(\lambda_1^{(n)}, \ldots, \lambda_k^{(n)}) \in \mathbf{\Sigma_k}$ for each $n \in \mathbb{N}$ and $\lambda_j^{(n)} \to \lambda_j$ for each $1 \le j \le k$. Pick any $\epsilon > 0$. Then $(\lambda_1^{(n)}, \ldots, \lambda_k^{(n)}) \in$

$\mathbf{\Sigma_k} \Rightarrow \exists x \in S(\mathcal{H})$ such that $\|(X_j - \lambda_j^{(n)})x\| < \epsilon/2$ for $1 \leq j \leq k$ (and for all n). Next, $\lambda_j^{(n)} \to \lambda_j \Rightarrow \exists n$ such that $|\lambda_j^{(n)} - \lambda_j| < \epsilon/2$. Thus, for any $\epsilon > 0$, we have shown that $\exists x \in S(\mathcal{H})$ such that

$$\|(X_j - \lambda_j)x\| \leq \|(X_j - \lambda_j^{(n)})x\| + |\lambda_j^{(n)} - \lambda_j| < \epsilon \text{ for } 1 \leq j \leq k$$

and indeed $(\lambda_1, \ldots, \lambda_k) \in \mathbf{\Sigma_k}$ and $\mathbf{\Sigma_k}$ is closed.

(2) We shall prove the result by induction on k. For $k = 1$, assertion (2) follows from Theorem 1.6.2 (2) and the non-emptiness of $\sigma(X_1)$.

Suppose now that the Theorem is valid for k, and suppose we are given commuting self-adjoint operators $X_1, \ldots, X_k, X_{k+1}$. Let us prove that $\lambda_{k+1} \in \sigma(X_{k+1})$ implies that there exists $(\lambda_1, \ldots, \lambda_k) \in \Sigma(X_1, \ldots, X_k)$ such that $(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}) \in \Sigma(X_1, \cdots, X_k, X_{k+1})$.

For each $n \in \mathbb{N}$, let $\mathcal{M}_n = \mathcal{M}_{X_{k+1}}(\lambda_{k+1} - \frac{1}{n}, \lambda_{k+1} + \frac{1}{n})$, where we continue to use the notation $\mathcal{M}_X(E) := 1_E(X)$ of the last section. By Proposition 2.4.1 (4), we see that $\mathcal{M}_n \neq \{0\}\ \forall n$. By Corollary 2.4.3 (2), each X_i leaves $\mathcal{M}_n$ invariant. Define $X_i(n) = X_i|_{\mathcal{M}_n}\ \forall 1 \leq i \leq k, n \in \mathbb{N}$. Deduce by induction hypothesis that $\mathbf{\Sigma_k}(n) := \mathbf{\Sigma_k}(X_1(n), \ldots, X_k(n)) \neq \emptyset\ \forall n$. Since $\{\mathcal{M}_n : n \in \mathbb{N}\}$ is a decreasing sequence of subspaces, it is clear that also $\{\mathbf{\Sigma_k}(n) : n \in \mathbb{N}\}$ is a decreasing sequence of non-empty compact sets. The finite intersection property then assures us that we can find a $(\lambda_1, \ldots, \lambda_k)$ in the non-empty set $\bigcap_{n\in\mathbb{N}} \mathbf{\Sigma_k}(n)$. Hence, by definition of the joint spectrum of commuting self-adjoint operators, we can find unit vectors $x_n \in \mathcal{M}_n$ such that $\|(X_i - \lambda_i)x_n\| = \|(X_i(n) - \lambda_i)x_n\| < \frac{1}{n}$ for $1 \leq i \leq k$, and $n \in \mathbb{N}$. On the other hand, it follows from the definition of $\mathcal{M}_n$ that $\|(X_{k+1} - \lambda_{k+1})x_n\| < \frac{1}{n}$. Thus, $\|(X_i - \lambda_i)x_n\| < \frac{1}{n}\ \forall 1 \leq i \leq k+1$ for every $n \in \mathbb{N}$; in other words, $(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}) \in \Sigma(X_1, \ldots, X_k, X_{k+1})$. Since $\Sigma(X_{k+1}) = \sigma(X_{k+1}) \neq \emptyset$ the proof is complete.

□

Proposition 2.5.3. *For any polynomial $p \in \mathbb{C}[t_1, \ldots, t_k]$, the operator $Z = p(X_1, \ldots, X_k)$ is normal, and*

(1) $\sigma(Z) = p(\mathbf{\Sigma_k})$; and

(2) $\|p(X_1, \ldots, X_k)\| = \|p\|_{\mathbf{\Sigma_k}}$, where the p on the right is the evaluation function on $\mathbf{\Sigma_k}$ given by the polynomial p.

Proof. (1) Let $q = \frac{1}{2}(p + \bar{p}), r = \frac{1}{2i}(p - \bar{p})$ and $X_{k+1} = q(X_1, \ldots, X_k), Y_{k+1} = r(X_1, \ldots, X_k)$. Then clearly $q, r \in \mathbb{R}[t_1, \ldots, t_k]$, so that X_{k+1} and Y_{k+1} are self-adjoint operators commuting with $X_1, \ldots, X_k$ and with each other as well (so Z is indeed normal). Since it follows from Corollary 1.6.4 that $\lambda = \alpha + i\beta \in \sigma(Z) \Leftrightarrow \alpha \in \sigma(X_{k+1})$ and $\beta \in \sigma(Y_{k+1})$, we see that it suffices to prove the case when $p = q$ is real-valued and $Z = X_{k+1}$ is a self-adjoint operator which is a real polynomial in $X_1, \ldots, X_k$ (and hence commutes with each X_i).

Suppose $\lambda_{k+1} \in \sigma(X_{k+1})$. It then follows from Lemma 2.5.2 that there exists $(\lambda_1, \ldots, \lambda_k) \in \mathbf{\Sigma_k}$ such that $(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}) \in \Sigma(X_1, \ldots, X_k, X_{k+1})$. Thus there exists a sequence $\{x_n : n \in \mathbb{N}\}$ of unit vectors in $\mathcal{H}$ such that $\|(X_i - \lambda_i)x_n\| \to 0 \; \forall 1 \leq i \leq k+1$. It follows easily from this requirement for the first k i's that then, necessarily, we must have $\|[p(X_1, \cdots, X_k) - p(\lambda_1, \cdots, \lambda_k)]x_n\| \to 0$ while also $\|(X_{k+1} - \lambda_{k+1})x_n\| \to 0$, which forces $\lambda_{k+1} = p(\lambda_1, \cdots, \lambda_k)$; in view of the arbitrariness of λ_{k+1}, this shows that $\sigma(X_{k+1}) \subset p(\mathbf{\Sigma_k})$. Conversely, it must be clear that if $(\lambda_1, \ldots, \lambda_k) \in \mathbf{\Sigma_k}$, then $p((\lambda_1, \ldots, \lambda_k))$ is an approximate eigenvalue of $p(X_1, \cdots, X_k)$ and thus, indeed, $\sigma(p(X_1, \cdots, X_k)) = p(\Sigma(X_1, \cdots, X_k))$.

(2) This follows immediately from (1) above and Proposition 1.5.6 (2).

□

Corollary 2.5.4. *With the notation of Proposition 2.5.3, we have:*

*(1) The 'polynomial functional calculus' extends uniquely to a isometric *-algebra isomorphism*

$$C(\Sigma) \ni f \overset{\pi}{\mapsto} f(X_1, \cdots, X_k) \in C^*(\{X_1, \ldots, X_k\}) \; ;$$

*(2) There exists a probability measure μ on $\mathcal{B}_\Sigma$ and an isometric *-algebra monomorphism $\tilde{\pi} : L^\infty(\mu) \to B(\mathcal{H})$ such that (i) $\tilde{\pi}$ 'extends' π, and (ii) a norm-bounded sequence $\{f_n : n \in \mathbb{N}\}$ in $L^\infty(\mu)$ converges to the constant function 0 in (μ) measure if and only if $\tilde{\pi}(f_n)$ SOT-converges to 0.*

Proof. (1) This follows from Proposition 2.5.3(2) and a routine application of the Stone-Weierstrass theorem, to show that the collection of complex polynomial functions on a compact subset $\mathbf{\Sigma}$ of $\mathbb{R}^k$, by virtue of being a self-adjoint unital subalgebra of functions which separates points of $\mathbf{\Sigma}$, is dense in $C(\mathbf{\Sigma})$.

(2) This is a consequence of item (1) above and Remark 2.3.3.

□

2.6 The Spectral Theorem for a normal operator

We are now ready to generalise Theorem 2.3.1 to the case of a normal operator. This is essentially just the specialisation of Corollary 2.5.4 for $k = 2$.

Thus, assume that $Z = X + iY \in B(\mathcal{H})$ is the Cartesian decomposition of a normal operator and that $\Sigma = \sigma(Z)$. In view of Proposition 2.5.3 (1), we see that $\Sigma = \{s + it : (s,t) \in \Sigma(X,Y)\}$, and we may and will identify $\Sigma \subset \mathbb{C}$ with $\Sigma(X,Y) \subset \mathbb{R}^2$.

In the following formulation of the spectral theorem for the normal operator Z (as above), the functions $f_i, i = 1, 2$ denote the functions $f_i : \Sigma \to \mathbb{R}$ defined by $f_1(z) = \mathrm{Re}\, z$, $f_2(z) = \mathrm{Im}\, z$. We omit the proof as it is just Corollary 2.5.4 for $k = 2$.

THEOREM 2.6.1. *(1)* **(Continuous Functional Calculus)** *There exists a unique isometric *-algebra isomorphism*

$$C(\Sigma) \ni f \mapsto f(Z) \in C^*(Z)$$

of $C(\Sigma)$ onto $C^*(Z)$ such that $f_1(Z) = X, f_2(Z) = Y$.

(2) **(Measurable Functional Calculus)** *There exists a measure μ defined on $\mathcal{B}_\Sigma$ and a unique isometric *-algebra homomorphism*

$$L^\infty(\Sigma, \mu) \ni f \mapsto f(Z) \in B(\mathcal{H})$$

of $L^\infty(\Sigma, \mu)$ into $B(\mathcal{H})$ such that (i) $f_1(Z) = X, f_2(Z) = Y$, and (ii) a norm-bounded sequence $\{f_n : n \in \mathbb{N}\}$ in $L^\infty(\Sigma, \mu)$ converges in (μ)-measure to f if and only if the sequence $\{f_n(Z) : n \in \mathbb{N}\}$ SOT-converges to $f(Z)$.

Now we proceed to the conventional formulation of the spectral theorem in terms of spectral or projection-valued measures $P : \mathcal{B}_\mathbb{C} \to B(\mathcal{H})$.

THEOREM 2.6.2. *Let N be a normal operator on a separable Hilbert space $\mathcal{H}$. Then there exists a unique mapping $P := P_N : \mathcal{B}_\mathbb{C} \to B(\mathcal{H})$ such that:*

(1) $P(E)$ is an orthogonal projection for all $E \in \mathcal{B}_\mathbb{C}$;

(2) $E \mapsto P(E)$ is a projection-valued measure; i.e., whenever $\{E_n : n \in \mathbb{N}\} \subset \mathcal{B}_\mathbb{C}$ is a sequence of pairwise disjoint Borel sets, and $E = \coprod_{n\in\mathbb{N}} E_n$, then $P(E) = \sum_{n\in\mathbb{N}} P(E_n)$, the series being interpreted as the SOT-limit of the sequence of partial sums;

(3) for $x \in \mathcal{H}$, the equation $P_{x,x}(E) = \langle P(E)x, x\rangle$ defines a finite positive scalar measure with $P_{x,x}(\mathbb{C}) = \|x\|^2$;

(4) for $x, y \in \mathcal{H}$, the equation $P_{x,y}(E) = \langle P(E)x, y\rangle$ defines a finite complex measure, with the property that

$$\langle Nx, y\rangle = \int_\mathbb{C} \lambda \, dP_{x,y}(\lambda) \; ; \tag{2.6.1}$$

more generally for any bounded measurable function $f : \mathbb{C} \to \mathbb{C}$, we have

$$\langle f(N)x, y\rangle = \int_\mathbb{C} f(\lambda) \, dP_{x,y}(\lambda) \; ; \tag{2.6.2}$$

(5) the **spectral measure** *P is 'supported' on the spectrum of N in the sense that $P(U) \neq 0$ for all open sets U that have non-empty intersection with $\Sigma := \sigma(N)$ – or equivalently Σ is the smallest closed set with $P(\Sigma) = 1$.*

Proof. Existence: Use the measurable functional calculus to define $P(E) = 1_N(E)$. As $1_E = \overline{1_E} = 1_E^2$, we see immediately that $P(E) = P(E)^* = P(E)^2$, and hence (1) is proved. As for (2), note that the pairwise disjointness assumption ensures that $1_{\coprod_{k=1}^n E_k} = \sum_{k=1}^n 1_{E_k}$, while $\coprod_{k=1}^n E_k \uparrow \coprod_{k\in\mathbb{N}} E_k$ implies $P(\coprod_{k=1}^\infty E_k) = SOT\text{-}\lim_{n\to\infty} P(\coprod_{k=1}^n E_k)$, thus establishing (2).

Since $\langle Qx, x\rangle = \|Qx\|^2 \geq 0$ for any projection Q, item (3) follows immediately from item (2). The polarisation identity and the definitions show that $P_{x,y} = \frac{1}{4}\sum_{j=0}^3 i^j P_{x+i^j y, x+i^j y}$, thereby demonstrating that $P_{x,y}$ is a complex linear combination of four finite positive measures, and is hence a finite complex measure. To complete the proof of item (4), it suffices to prove equation (2.6.2) since equation (2.6.1) is a special case (with $f(z) = 1_\Sigma(z)z$). Equation (2.6.2) is, by definition, valid when f is of the form 1_E, and hence by linearity, also valid for any simple function. For a general bounded measurable function f, and an $\epsilon > 0$, choose a simple function s such that $\|s - f\| < \epsilon$ uniformly. Then,

$$|\langle f(N)x, y\rangle - \langle s(N)x, y\rangle| \leq \epsilon\|x\|\ \|y\|$$

and

$$\left|\int f\, dP_{x,y} - \int s\, dP_{x,y}\right| \leq \epsilon\|P_{x,y}\|$$

so

$$\left|\langle f(N)x, y\rangle - \int f\, dP_{x,y}\right| \leq \epsilon(\|x\|\ \|y\| + \|P_{x,y}\|)\ .$$

As ϵ was arbitrary, we find that equation (2.6.2) indeed holds for any bounded measurable f.

As for (5), suppose $P(U) = 0$ for some open U, and $z_0 \in U$. Pick $\epsilon > 0$ such that $D = \{z \in \mathbb{C} : |z - z_0| < \epsilon\} \subset U$. Then $P(U) = 0 \Rightarrow P(D) = 0 \Rightarrow \|1_D\|_{L^\infty(\mu)} = 0 \Rightarrow \mu(D) = 0 \Rightarrow \frac{1}{f_0 - z_0} \in L^\infty(\mu) \Rightarrow z_0 \notin \sigma(N)$, so, indeed $P(U) = 0, U$ open $\Rightarrow U \cap \Sigma = \emptyset$.

Uniqueness: If, conversely $\tilde{P}$ is another such spectral measure satisfying the conditions (1)–(5) of the theorem, it follows from equation (2.6.2) that

$$\int z^m \bar{z}^n\, d\tilde{P}_{x,y}(z) = \langle N^m N^{*n} x, y\rangle = \int z^m \bar{z}^n\, dP_{x,y}(z)\ \forall m, n \in \mathbb{Z}_+\ .$$

Since functions of the form $z \mapsto z^m\bar{z}^n$ span a dense subspace of $C(\Sigma)$, thanks to the Stone-Weierstrass theorem, it now follows from the Riesz representation theorem that $\tilde{P}_{x,y} = P_{x,y}$. The validity of this equality for all $x, y \in \mathcal{H}$ shows, finally, that indeed $\tilde{P} = P$, as desired. □

Remark 2.6.3. Now that we have the uniqueness assertion of Theorem 2.6.2, we can re-connect with a way to produce probability measures in the measure class of the mysterious μ appearing in the measurable functional calculus. If P denotes **the** spectral measure of X, the following conditions on an $E \in \mathcal{B}_\Sigma$ are equivalent:

(1) $1_E(X)(= P(E)) = 0$.

(2) $\mu(E) = 0$.

(3) $P_{x,x}(E) = 0$ for all x in a total set $\mathcal{S} \subset \mathcal{H}$.

Hence, a possible choice for μ is $\sum_{n\in\mathbb{N}} 2^{-n} P_{e_n,e_n}$ where $\{e_n : n \in \mathbb{N}\}$ is an orthonormal basis for $\mathcal{H}$.

Incidentally, a measure of the form $P_{x,x}$ is sometimes called a **scalar spectral measure** for N.

Reason: (1) $\Leftrightarrow$ (2) This is because $L^\infty(\mu) \ni f \mapsto f(X) \in B(\mathcal{H})$ is isometric by Theorem 2.3.1 (2).

(1) $\Leftrightarrow$ (3) This is because (i) for a projection P – in this case, $P(E)$ – $\langle Px, x\rangle = 0 \Leftrightarrow Px = 0$, and (ii) a bounded operator is the zero operator if and only if its kernel contains a total set.

Remark 2.6.4. To tie a loose-end, we wish to observe that $\|P_{x,y}\| \leq \|x\|\ \|y\|$. This is because

$$\|P_{x,y}\| = \inf\{K > 0 : \left|\int f\, dP_{x,y}\right| \leq K\|f\|_{C(\Sigma)}\ \forall f \in C(\Sigma)\}$$

and

$$\begin{aligned} \left|\int f\, dP_{x,y}\right| &= |\langle f(N)x, y\rangle| \\ &\leq \|f(N)\|\ \|x\|\ \|y\| \\ &\leq \|f\|_{C(\Sigma)}\ \|x\|\ \|y\|. \end{aligned}$$

Remark 2.6.5. This final remark is an advertising pitch for my formulation of the spectral theorem in terms of functional calculi, in comparison with the conventional version in terms of spectral measures: the difference is between having some statement for all bounded measurable functions and only having it for indicator functions and having to go through the exercise of integration every time one wants to get to the former situation!

Exercise 2.6.6. *Let $\pi_\mu : L^\infty(\mu) \to B(L^2(\mu))$ be the 'multiplication representation' as in Proposition* 2.2.3. *Can you identify the spectral measure P_N where $N = \pi_\mu(f)$?* (Hint: *Consider the cases $\Sigma = \{z \in \mathbb{C} : |z| = 1\}$ and $f(z) = z^n$ with $n = 1, 2, \ldots$, in increasing order of difficulty as n varies.*)

2.7 Several commuting normal operators

2.7.1 The Fuglede Theorem

Theorem 2.7.1. *[Fuglede] If an operator T commutes with a normal operator N, then it necessarily also commutes with N^*.*

Proof. When $\mathcal{H}$ is finite-dimensional, the spectral theorem says that N admits the decomposition $N = \sum_{i=1}^{k} \lambda_i P_i$ where $\sigma(N) = \{\lambda_1, \ldots, \lambda_k\}$ and $P_i = 1_{\{\lambda_i\}}(N)$; observe that $P_i = p_i(N)$ for appropriate polynomials $p_1, \ldots, p_k$, and deduce that T commutes with each P_i and hence also with $f(N)$ for any function $f : \sigma(N) \to \mathbb{C}$, and in particular with $N^* = \bar{f}_0$ where $f_0(z) = z$.

We shall similarly prove that T commutes with each spectral projection $1_E(N), E \in \mathcal{B}_{\mathbb{C}}$ and hence also with $f(N)$ for each (simple, and hence each) bounded measurable function f, and in particular, for $f(z) = 1_{\sigma(N)}(z)\bar{z}$. Note that T commutes with a projection P if and only if T leaves both $\mathcal{M}$ and $\mathcal{M}^\perp$ invariant, where $\mathcal{M} = \operatorname{ran}(P)$.

We shall write $\mathcal{M}(E) = \operatorname{ran} 1_E(N)$. Since $\mathcal{M}(E)^\perp = \mathcal{M}(E')$ (where we write $E' = \mathbb{C} \setminus E$), we see from the previous paragraph that Fuglede's theorem is equivalent to the assertion that if T commutes with a normal N, then T leaves each $\mathcal{M}(E)$ invariant – which is what we shall accomplish in a sequence of simple steps:

Define $\mathcal{F} = \{E \in \mathcal{B}_{\mathbb{C}} : T \text{ leaves } \mathcal{M}(E) \text{ invariant}\}$, so we need to prove that $\mathcal{F} = \mathcal{B}_{\mathbb{C}}$.

(1) Write $D(z_0, r) = \{z \in \mathbb{C} : |z - z_0| < r\}$ and simply $\mathbb{D} = D(0,1)$, so the closure $\bar{\mathbb{D}} = \{z \in \mathbb{C} : |z| \le 1\}$. We shall need the following analogue of Proposition 2.4.2 for normal operators: The following conditions on an $x \in \mathcal{H}$ are equivalent:

   (1) $x \in \mathcal{M}(\bar{\mathbb{D}})$.

   (2) $\|N^n x\| \le \|x\| \ \forall n \in \mathbb{N}$.

   (3) $\{\|N^n x\| : n \in \mathbb{N}\}$ is a bounded set.

   Reason: $(a) \Rightarrow (b)$:

$$\bar{z}z1_{\overline{\mathbb{D}}}(z) \le 1 \Rightarrow N^*N1_{\overline{\mathbb{D}}}(N) \le id_{\mathcal{H}} \Rightarrow \|Nx\|^2 \le 1 \ \forall x \in \mathcal{M}(\bar{\mathbb{D}}) \ .$$

   $(b) \Rightarrow (c)$ is obvious.

   $(c) \Rightarrow (a)$ Let $x_m := 1_{\{z:|z|\ge 1+\frac{1}{m}\}}(N)x \ \forall m \in \mathbb{N}$; then, for all $n \in \mathbb{N}$, we have, by Proposition 2.4.2,

$$\begin{aligned} \|N^{2n}x\| = \|(N^*N)^n x\| &\ge \|1_{\{[(1+\frac{1}{m})^2,\infty)}(N * N)(N^*N)^n x\| \\ &\ge \left(1 + \frac{1}{m}\right)^{2n} \|x_m\| \ ; \end{aligned}$$

   and now, the assumed boundedness condition (c) implies that we must have $x_m = 0 \ \forall m$ and hence that $x = x - \lim_{m\to\infty} x_m \in \mathcal{M}(\bar{\mathbb{D}})$; and the proof of the normal analogue of Proposition 2.4.2 is complete.

   Since

$$\|N^n Tx\| = \|TN^n x\| \le \|T\|\|N^n x\|,$$

   condition (c) above implies that if $x \in \mathcal{M}(\bar{\mathbb{D}})$, then also $Tx \in \mathcal{M}(\bar{\mathbb{D}})$; so $\bar{\mathbb{D}} \in \mathcal{F}$.

(2) $D(z,r) \in \mathcal{F}\ \forall z \in \mathbb{C}, r > 0$.

Reason: This follows by applying item (1) above to $\left(\frac{N-z}{r}\right)$.

(3) $\mathcal{F}$ is closed under countable monotone limits, and is hence a 'monotone class'.

Reason: If $E_n \in \mathcal{F}\ \forall n$ and if $E_n \uparrow E$ (resp., $E_n \downarrow E$), then $1_{E_n}(N) \stackrel{SOT}{\rightarrow} 1_E(N)$ so that $\mathcal{M}(E) = \overline{(\bigcup \mathcal{M}(E_n))}$ (resp., $\mathcal{M}(E) = \bigcap \mathcal{M}(E_n))$ whence also $E \in \mathcal{F}$.

(4) $\mathcal{F}$ contains all (open or closed) discs.

Reason: The assertion regarding closed discs is item (2) above, and open discs are increasing unions of closed discs.

(5) $\mathcal{F}$ contains all (open or closed) half-planes.

Reason: This is because (i) every open half-plane is an increasing union of closed discs (for example, $R_a := \{z \in \mathbb{C} : \mathrm{Re}\, z > a\} = \bigcup_{n=1}^{\infty}\{z \in \mathbb{C} : |z-(a+n)| \leq n\}$); and (ii) every closed half-plane is a decreasing intersection of open half-planes (eg: $\{\mathrm{Re}\, z > a - \frac{1}{n}\} \downarrow \{\mathrm{Re}\, z \geq a\}$.)

However, we will only need this fact for the special half-planes $R_a, L_b = \{z \in \mathbb{C} : \mathrm{Re}\, z \leq b\}, U_c = \{z \in \mathbb{C} : \mathrm{Im}\, z > c\}, D_d = \{z \in \mathbb{C} : \mathrm{Im}\, z \leq d\}$.

(6) $\mathcal{F}$ is closed under finite intersections and countable disjoint unions.

Reason: $1_{\cap_{i=1}^n E_i} = \prod_{i=1}^n 1_{E_i} \Rightarrow \mathcal{M}(\bigcap_{i=1}^n E_i) = \bigcap_{i=1}^n \mathcal{M}(E_i)$ so if $E_1, \ldots E_n \in \mathcal{F}$, and $x \in \mathcal{M}(\bigcap_{i=1}^n E_i)$, then $x \in \mathcal{M}(E_i)\ \forall i$ and $Tx \in \mathcal{M}(\cap_{i=1}^n E_i)$, so $\bigcap_{i=1}^n E_i \in \mathcal{F}$. Similarly $\mathcal{M}(\coprod_{n=1}^{\infty} E_n) = [\bigcup_{n=1}^{\infty} \mathcal{M}(E_n)]$ implies that $\mathcal{F}$ is closed under countable disjoint unions.

(7) $\mathcal{F} = \mathcal{B}_{\mathbb{C}}$.

Reason: It follows from items (5) and (6) above that $\mathcal{F}$ contains $(a,b] \times (c,d] = R_a \cap L_b \cap U_c \cap D_d$ and the collection $\mathcal{A}$ of all finite disjoint unions of such rectangles. Since $\mathcal{A} \cup \{\emptyset, \mathbb{C}\}$ is an algebra of sets which generates $\mathcal{B}_{\mathbb{C}}$ as a σ-algebra, and since $\mathcal{F}$ is a monotone class containing $\mathcal{A} \cup \{\emptyset, \mathbb{C}\}$, the desired conclusion is a consequence of the monotone class theorem.

$\square$

REMARK 2.7.2. Putnam proved – see [Put] – this extension to Fuglede's theorem: if $N_i, i = 1, 2$ is a normal operator on $\mathcal{H}_i$ and if $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ satisfies $TN_1 = N_2 T$, then, we also necessarily have $TN_1^* = N_2^* T$. (A cute 2×2 matrix proof of this – see [Hal2] – applies Fuglede's theorem to the operators on $\mathcal{H}_1 \oplus \mathcal{H}_2$ given by the operator matrices $\begin{bmatrix} 0 & 0 \\ T & 0 \end{bmatrix}$ and $\begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$.)

2.7.2 Functional calculus for several commuting normal operators

This section addresses the analogue of the statement that a family of commuting normal operators on a finite-dimensional Hilbert space can be simultaneously diagonalised, equivalently, that an arbitrary family $\{N_j : j \in I\}$ of pairwise commuting normal operators admits a joint functional calculus – i.e., an appropriate continuous and measurable 'joint functional calculus' identifying (algebraically and topologically) appropriate closures of the *-algebras generated by the family $\{N_j; j \in I\}$.

Suppose $\{X_i : i \in I\}$ is a (possibly infinite, maybe even uncountable) family of self-adjoint operators on $\mathcal{H}$. For each finite set $F \subset I$, let Σ_F be the joint spectrum of $\{X_j : j \in F\}$. Recall that $\mathbf{\Sigma}_F \subset \prod_{i \in F} \sigma(X_i)$. Let $\mathrm{pr}_F : \prod_{i \in I} \sigma(X_i) \to \prod_{i \in F} \sigma(X_i)$ denote the natural projection.

We start with a mild generalisation of Lemma 2.5.2(2).

Lemma 2.7.3. *If $F \subset E \subset I$ are finite sets, and if $\mathrm{pr}^E_F : \prod_{i \in E} \sigma(X_i) \to \prod_{i \in F} \sigma(X_i)$ is the natural projection, then $\Sigma_F = \mathrm{pr}^E_F(\Sigma_E)$.*

Proof. This assertion is easily seen to follow by induction on $|E \setminus F|$ from the special case of the Lemma when $|E \setminus F| = 1$. (*Reason:* If the result is known for $F_n \subset E, |E \setminus F_n| = n$ and if $|E \setminus F| = n + 1$, we can find F_n such that $F \subset F_n \subset E$, $|E \setminus F_k| = k$, and observe that $\pi^E_F = \pi^{F_n}_F \circ \pi^E_{F_n}$, and deduce the truth of the assertion for $n + 1$ from that of n and 1, thus: $\Sigma_F = \mathrm{pr}^{F_n}_F(\Sigma_{F_n}) = \mathrm{pr}^{F_n}_F(\mathrm{pr}^E_{F_n}(\Sigma_E))$.)

So suppose $E = \{1, 2, \ldots, k + 1\}$ and $F = \{1, 2, \ldots, k\}$. Suppose $(\lambda_1, \ldots, \lambda_k) \in \Sigma_F$. If $\epsilon > 0$, it is seen from Corollary 2.5.4 and Remark 2.3.3(2) that $\mathcal{M}(\epsilon) := \tilde{\pi}(1_{\{(t_1,\ldots,t_k) \in \Sigma_F : |t_i - \lambda_i| < \epsilon \ \forall i \in F\}}) \neq 0$ and is invariant under each $X_i, 1 \le i \le k + 1$. If $X_{k+1}(\epsilon) = X_{k+1}|_{\mathcal{M}(\epsilon)}$ and $\lambda_{k+1} \in \sigma(X_{k+1}(\epsilon))$, it is seen that $\exists x(\epsilon) \in S(\mathcal{M}(\epsilon))$ such that $\|(X_{k+1} - \lambda_{k+1})x(\epsilon)\| < \epsilon$. Since $\|(X_i - \lambda_i)x\| < \epsilon \ \forall x \in S(\mathcal{M}(\epsilon))$, we see that $\{x(\frac{1}{n})\}$ is a sequence of unit vectors such that $\|(X_i - \lambda_i)x(\frac{1}{n})\| < \frac{1}{n} \ \forall n$, and indeed $(\lambda_1, \ldots \lambda_{k+1}) \in \mathbf{\Sigma}_E$ so $\mathbf{\Sigma}_E \subset \mathrm{pr}^E_F(\mathbf{\Sigma}_E)$. The reverse inclusion is obvious, and the proof is complete. □

For each finite $F \subset I$, let $\mathbf{\Sigma}(F) = \mathrm{pr}^{-1}_F(\mathbf{\Sigma}_F)$ and let $\mathbf{\Sigma} = \bigcap_F \mathbf{\Sigma}(F)$.

Theorem 2.7.4. *With the foregoing notation, we have:*

(1) $\mathbf{\Sigma}$ is a non-empty compact set, which we shall refer to as the joint spectrum of $\{X_j : j \in I\}$.

(2) There exists a unique isomorphism $\pi : C(\Sigma) \to C^*(\{X_j : j \in I\})$ such that $\pi(\mathrm{pr}_{\{j\}}) = X_j \ \forall j \in I$.

*(3) There exists a probability measure μ defined on $\mathcal{B}_\Sigma$, unique up to mutual absolute continuity, such that the continuous functional calculus π above 'extends' to an isometric *-algebra monomorphism $\tilde{\pi}$ of $L^\infty(\Sigma, \mathcal{B}_\Sigma, \mu) \to B(\mathcal{H})$ with the property that a norm-bounded sequence $\{f_n : n \in \mathbb{N}\} \subset$*

$L^\infty(\Sigma, \mathcal{B}_\Sigma, \mu)$ converges in (μ) measure if and only if the image of this sequence under this 'joint measurable functional calculus' is SOT*-convergent.*

Proof. (1) It is clear that $\mathbf{\Sigma}$ is the closed subset of $\mathbb{R}^I$ consisting of those tuples $((\lambda_i))_{i\in I}$ such that for any finite $F \subset I$, it is possible to find a sequence of unit vectors $x_n^F, n \in \mathbb{N}$ such that $\|(X_i - \lambda_i)x_n^F\| \to 0\ \forall i \in F$ so that, in particular $\mathbf{\Sigma}$ is a closed subset of $\prod_{i\in I}\sigma(X_i)$ and hence compact. It is not hard to see (from Lemma 2.7.3 and Lemma 2.5.2) that $\{\mathbf{\Sigma}(F) : F$ a finite subset of $I\}$ is a family of non-empty compact sets with the finite intersection property, and that hence, their intersection, i.e., $\mathbf{\Sigma}$, is also non-empty and compact.

(2) On the one hand, the family $\{\mathrm{pr}_F : F$ a finite subset of $I\}$ linearly spans a self-adjoint subalgebra of functions which separates points of $\mathbf{\Sigma}$, which is dense in $C(\mathbf{\Sigma})$. It then follows from Proposition 2.5.3 (2) that there is a unique isometric *-algebra isomorphism $\pi : C(\mathbf{\Sigma}) \to C^*(\{X_i : i \in I\})$ such that $\pi(\mathrm{pr}_{\{j\}}) = X_j$.

(3) This follows immediately from Remark 2.3.3.

□

Suppose now that $N_j = A_j + iB_j$ (resp., $\lambda_j = \alpha_j + i\beta_j$) is the Cartesian decomposition of N_j as in the last paragraph (resp., $\lambda_j \in \sigma(N_j)$), and denote their **joint spectrum** by the set $\Sigma = \{\lambda = ((\lambda_j))_{j\in I} \in \mathbb{C}^I$ – or alternatively $\{((\alpha_j, \beta_j))_{j\in I} \in (\mathbb{R}^2)^I\}$ – of those tuples for which it is possible to find a sequence $\{x_n : n \in \mathbb{N}\}$ of unit vectors such that

$$\lim_{n\to\infty} \|(N_j - \lambda_j)x_n\|^2 = \lim_{n\to\infty} \left(\|(A_j - \alpha_j)x_n\|^2 + \|(B_j - \beta_j)x_n\|^2\right) = 0\ \forall j \in I.$$

In view of Fuglede's theorem, we see that commutativity of the family $\{N_j : j \in I\}$ of normal operators is equivalent to that of the family $\{A_j, B_j : j \in I\}$ of self-adjoint operators. It must be clear that $\{((\alpha_j + i\beta_j)) \in \mathbb{C}^I : (((\alpha_j, \beta_j))) \in \mathbf{\Sigma}(\{A_j, B_j : j \in I\}$ may be defined as the joint spectrum of the family $\{N_j : j \in I\}$ of normal operators, and the exact counterpart of Theorem 2.7.4 (with mild modifications, usually involving changing $\mathbb{R}$ to $\mathbb{C}$ and self-adjoint to normal) for a family of commuting normal operators is valid.

EXERCISE 2.7.5. *(1) Formulate and prove the precise statement of the 'normal version' of Theorem* 2.7.4.

(2) Also state and prove a formulation of the 'joint spectral theorem' for a family of commuting normal operators in terms of projection-valued measures.

2.8 Typical uses of the spectral theorem

We now list some simple consequences of the spectral theorem (i.e., the functional calculi) for a normal operator.

Proposition 2.8.1. *1. Let $T \in B(\mathcal{H})$ be a normal operator. Then*

(a) *T is self-adjoint if and only if $\sigma(T) \subset \mathbb{R}$.*

(b) *T is a projection if and only if $\sigma(T) \subset \{0,1\}$.*

(c) *T is unitary if and only if $\sigma(T) \subset \{z \in \mathbb{C} : |z| = 1\}$.*

2. *The following conditions on an operator $A \in B(\mathcal{H})$ are equivalent:*

(a) *There exists some Hilbert space $\mathcal{K}$ and an operator $T \in B(\mathcal{H}, \mathcal{K})$ such that $A = T^*T$.*

(b) *$\langle Ax, x \rangle \geq 0 \ \forall x \in \mathcal{H}$.*

(c) *A is self-adjoint and $\sigma(A) \subset [0, \infty)$.*

(d) *A is normal and $\sigma(A) \subset [0, \infty)$.*

(e) *There exists a self-adjoint operator $B \in B(\mathcal{H})$ such that $A = B^2$.*

Such an operator A is said to be **positive**, *and we write $A \geq 0$, and more generally, we shall write $A \geq C$ if and only if A, C are self-adjoint operators satisfying $A - C \geq 0$.*

3. *If $A \geq 0$, there exists a unique $B \geq 0$ such that $A = B^2$, and we denote this unique positive square root of A by $A^{\frac{1}{2}}$.*

4. *Let $U \in B(\mathcal{H})$ be a unitary operator. Then there exists a self-adjoint operator $A \in B(\mathcal{H})$ such that $U = e^{iA}$, where the right hand side is interpreted as the result of the continuous functional calculus for A; further, given any $a \in \mathbb{R}$, we may choose A to satisfy $\sigma(A) \subset [a, a + 2\pi]$.*

5. *If $T \in B(\mathcal{H})$ is a normal operator, and if $n \in \mathbb{N}$, then there exists a normal operator $A \in B(\mathcal{H})$ such that $T = A^n$.*

6. *Any self-adjoint operator T admits a unique decomposition $T = T_+ - T_-$, where $T_\pm \geq 0$ and $T_+T_- = 0 = T_-T_+$*

7. *Any self-adjoint contraction (i.e., an operator T satisfying $T = T^*$ and $\|T\| \leq 1$ is expressible as the average of at most two unitary operators, and hence any operator is expressible as a linear combination of at most four unitary operators.*

Proof. (1) A normal operator T is self-adjoint (resp., a projection, resp., unitary precisely when it satisfies $T = T^*$, or $T = T^* = T^2$, or $T^*T = 1$ respectively. while the function $f_0 \in C(\Sigma)$, for $\Sigma \subset \mathbb{C}$, defined by $f_0(z) = z$ satisfies $f_0 = \overline{f_0}$ (resp., $f_0 = \overline{f_0} = f_0^2$, resp., $f_0\overline{f_0} = 1$) precisely when $\Sigma \subset \mathbb{R}$ (resp., $\Sigma \subset \{0,1\}$, resp., $\Sigma \subset \{z : |z| = 1\}$).

(2) The implications $(e) \Rightarrow (a) \Rightarrow (b)$ and $(c) \Rightarrow (d)$ are obvious. As for $(d) \Rightarrow (e)$, note that (d) implies that A is self-adjoint by $1(a)$. If the function defined on $[0, \infty)$, by $f(t) = t^{\frac{1}{2}}$, denotes the positive square-root, then the condition (c) implies that $f \in C(\sigma(A))$, and we see that $B = f(A)$ works.

(Notice that $B \in C^*(A)$ by construction). As for $(b) \Rightarrow (c)$, the self-adjointness of A follows from Corollary 1.5.3 (2), and the positivity of elements of $\sigma(A)$ follows then from Theorem 1.6.2(2).

(3) Suppose B_1 is another prospective positive square root of A. Since $B \in C^*(A) \subset C^*(B_1) \cong C(\sigma(B_1))$, there must be a non-negative $g \in C(\sigma(B_1))$ such that $B = g(B_1)$. As $B^2 = A = B_1^2$, we must have $g(t)^2 = t^2 \ \forall t \in \sigma(B_1)$, and we must have $g(t) = t$ so $B = B_1$.

(4) Let $\phi : \mathbb{C} \setminus \{0\} \to \{z \in \mathbb{C} : \operatorname{Im} z \in [a, a + 2\pi)\}$ be any (measurable) branch of the logarithm - for instance, we might set $\phi(z) = log|z| + i\theta$, if $z = |z|e^{i\theta},\ a \leq \theta < a + 2\pi$. Setting $A = \phi(U)$, we find - since $e^{\phi(z)} = z$ - that $U = e^{iA}$.

(5) This is proved like 4 above, by taking some measurable branch of the logarithm defined everywhere in $\mathbb{C} \setminus \{0\}$ and choosing the $z^{\frac{1}{n}}$ as the exponential of $\frac{1}{n}$ times this choice of logarithm.

(6) Define $T_\pm = f_\pm(T)$ where $f_\pm$ are the obviously continuous functions $f_\pm : \mathbb{R} \to \mathbb{R}$ defined by $f_\pm = (|f_0| \pm f_))/2$. Then indeed

$$f_0 = f_+ - f_-, f_\pm \geq 0 \quad \text{and} \quad f_+ f_-(= f_- f_+) = 0$$

and hence

$$T_0 = T_+ - T_-, T_\pm \geq 0 \quad \text{and} \quad T_+ T_- = 0 = (T_+ T_-)^* = T_- T_+.$$

As for uniqueness, if $T = A_+ - A_-$ with $A_\pm \geq 0, A_+ A_- = 0$, note first that

$$A_+ A_- = 0 \Rightarrow A_- A_+ = (A_+ A_-)^* = 0$$

and hence that

$$(A_+ + A_-)^2 = A_+^2 + A_-^2 = (A_+ - A_-)^2 = T^2 = |T|^2$$

where $|T|$ represents the image, under the functional calculus for T, of the function $f(t) = |t|$; and we may deduce from the uniqueness of the positive square root of a positive operator that $(A_+ + A_-) = |T|$ and hence we must have

$$A_\pm = \tfrac{1}{2}(|T| \pm T) = T_\pm \ ,$$

as desired.

(7) Consider $v_\pm \in C([-1, 1])$ defined by $v_\pm(t) = t \pm i\sqrt{1 - t^2}$. Note that $t = \frac{1}{2}(v_+(t) + v_-(t))$ and $|v_\pm(t)| = 1$ for $t \in [-1, 1]$. Define $U_\pm = v_\pm(T)$.

As $U_\pm$ are unitary with average T, it follows, by scaling, that every self-adjoint operator is a linear combination of at most[2] two unitary operators, and the Cartesian decomposition completes the proof of the proposition. □

[2] The reason for the 'at most' is that T might have already been self-adjoint and unitary (i.e., satisfying $T^2 = 1$)

Chapter 3

Beyond normal operators

3.1 Polar decomposition

In this section, we establish the very useful **polar decomposition** for bounded operators on Hilbert space. We begin with a few simple observations and then introduce the crucial notion of a **partial isometry**.

LEMMA 3.1.1. *Let $T \in B(\mathcal{H}, \mathcal{K})$. Then,*

$$\ker T \;=\; \ker\, (T^*T) \;=\; \ker\, (T^*T)^{\frac{1}{2}} \;=\; \mathrm{ran}^{\perp} T^* \;. \tag{3.1.1}$$

In particular, also

$$\ker^{\perp} T \;=\; \overline{\mathrm{ran}\ T^*} \;.$$

(In the equations above, we have used the notation $\mathrm{ran}^{\perp} T^*$ *and* $\ker^{\perp} T$*, for* $(\mathrm{ran}\ T^*)^{\perp}$ *and* $(\ker\ T)^{\perp}$*, respectively.)*

Proof : First observe that, for arbitrary $x \in \mathcal{H}$, we have

$$||Tx||^2 \;=\; \langle T^*Tx, x\rangle \;=\; \langle (T^*T)^{\frac{1}{2}}x, (T^*T)^{\frac{1}{2}}x\rangle \;=\; ||(T^*T)^{\frac{1}{2}}x||^2 \;, \tag{3.1.2}$$

whence it follows that $\ker T \;=\; \ker(T^*T)^{\frac{1}{2}}$.

Notice next that

$$\begin{aligned} x \in \mathrm{ran}^{\perp} T^* \quad &\Leftrightarrow \quad \langle x, T^*y\rangle = 0 \;\forall\; y \in \mathcal{K} \\ &\Leftrightarrow \quad \langle Tx, y\rangle = 0 \;\forall\; y \in \mathcal{K} \\ &\Leftrightarrow \quad Tx = 0 \end{aligned}$$

and hence $\mathrm{ran}^{\perp} T^* \;=\; \ker\ T$. 'Taking perps' once again, we find – because of the fact that $V^{\perp\perp} = \overline{V}$ for any linear subspace $V \subset \mathcal{K}$ – that the last statement of the Lemma is indeed valid.

Finally, if $\{p_n\}_n$ is any sequence of polynomials with the property that $p_n(0) = 0$ for all n and such that $\{p_n(t)\}$ converges uniformly to $\sqrt{t}$ on $\sigma(T^*T)$, it follows that $||p_n(T^*T) - (T^*T)^{\frac{1}{2}}|| \to 0$, and hence,

$$x \in\ \ker(T^*T) \;\Rightarrow\; p_n(T^*T)x = 0 \;\forall n \;\Rightarrow (T^*T)^{\frac{1}{2}}x = 0$$

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9_3

and hence we see that also $\ker(T^*T) \subset \ker(T^*T)^{\frac{1}{2}}$; since the reverse inclusion is clear, the proof of the lemma is complete. $\square$

PROPOSITION 3.1.2. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces; then the following conditions on an operator $U \in B(\mathcal{H}, \mathcal{K})$ are equivalent:*

(i) $U = UU^*U$;

(ii) $P = U^*U$ is a projection;

(iii) $U|_{\ker^\perp U}$ is an isometry.

An operator which satisfies the equivalent conditions (i)–(iii) is called a **partial isometry**.

Proof. $(i) \Rightarrow (ii)$: The assumption (i) clearly implies that $P^2 = U^*UU^*U = U^*U = P$, while P is clearly self-adjoint.

$(ii) \Rightarrow (iii)$: Let $\mathcal{M} = \text{ran } P$. Then notice that, for arbitrary $x \in \mathcal{H}$, we have: $||Px||^2 = \langle Px, x\rangle = \langle U^*Ux, x\rangle = ||Ux||^2$; this clearly implies that $\ker U = \ker P = \mathcal{M}^\perp$, and that U is isometric on $\mathcal{M}$ (since P is identity on $\mathcal{M}$).

$(iii) \Rightarrow (ii)$: Let $\mathcal{M} = \ker^\perp U$. For $i = 1, 2$, suppose $z_i \in \mathcal{H}$, and $x_i \in \mathcal{M}, y_i \in \mathcal{M}^\perp$ are such that $z_i = x_i + y_i$; then note that

$$\begin{aligned} \langle U^*Uz_1, z_2\rangle &= \langle Uz_1, Uz_2\rangle \\ &= \langle Ux_1, Ux_2\rangle \\ &= \langle x_1, x_2\rangle \quad \text{since } U|_{\mathcal{M}} \text{ is isometric)} \\ &= \langle x_1, z_2\rangle \qquad (\text{since } \langle x_1, y_2\rangle = 0), \end{aligned}$$

and hence U^*U is the projection onto $\mathcal{M}$.

$(ii) \Rightarrow (i)$: Let $\mathcal{M} = \text{ran } U^*U$; then (by Lemma 3.1.1) $\mathcal{M}^\perp = \ker U^*U = \ker U$, and so, if $x \in \mathcal{M}, y \in \mathcal{M}^\perp$, are arbitrary, and if $z = x + y$, then observe that $Uz = Ux + Uy = Ux = U(U^*Uz)$. $\square$

REMARK 3.1.3. Suppose $U \in B(\mathcal{H}, \mathcal{K})$ is a partial isometry. Setting $\mathcal{M} = \ker^\perp U$ and $\mathcal{N} = \text{ran } U (= \overline{\text{ran } U})$, we find that U is identically 0 on $\mathcal{M}^\perp$, and U maps $\mathcal{M}$ isometrically onto $\mathcal{N}$. It is customary to refer to $\mathcal{M}$ as the **initial space**, and to $\mathcal{N}$ as the **final space**, of the partial isometry U.

On the other hand, upon taking adjoints in condition (ii) of Proposition 3.1.2, it is seen that $U^* \in B(\mathcal{K}, \mathcal{H})$ is also a partial isometry. In view of the preceding lemma, we find that $\ker U^* = \mathcal{N}^\perp$ and that $\text{ran } U^* = \mathcal{M}$; thus $\mathcal{N}$ is the inital space of U^* and $\mathcal{M}$ is the final space of U^*.

Finally, it follows from Proposition 3.1.2(ii) (and the proof of that proposition) that U^*U is the projection (of $\mathcal{H}$) onto $\mathcal{M}$ while UU^* is the projection (of $\mathcal{K}$) onto $\mathcal{N}$. $\square$

EXERCISE 3.1.4. *If $U \in B(\mathcal{H}, \mathcal{K})$ is a partial isometry with initial space $\mathcal{M}$ and final space $\mathcal{N}$, show that if $y \in \mathcal{N}$, then U^*y is the unique element $x \in \mathcal{M}$ such that $Ux = y$.*

Before stating the polar decomposition theorem, we introduce a convenient bit of notation: if $T \in B(\mathcal{H}, \mathcal{K})$ is a bounded operator between Hilbert spaces, we shall always use the symbol $|T|$ to denote the unique positive square root of the positive operator $|T|^2 = T^*T \in B(\mathcal{H})$; thus, $|T| = (T^*T)^{\frac{1}{2}}$. (If T is self-adjoint – in fact, even normal – for which, of course, we would need $\mathcal{H} = \mathcal{K}$, this notation/definition is consistent with that yielded by the continuous functional calculus.)

THEOREM 3.1.5. **(Polar Decomposition)**

(a) Any operator $T \in B(\mathcal{H}, \mathcal{K})$ *admits a decomposition* $T = UA$ *where*
(i) $U \in B(H, \mathcal{K})$ *is a partial isometry;*
(ii) $A \in B(\mathcal{H})$ *is a positive operator; and*
(iii) $\ker T = \ker U = \ker A$.

(b) Further, if $T = VB$ *is another decomposition of* T *as a product of a partial isometry* V *and a positive operator* B *such that* $\ker V = \ker B$, *then necessarily* $U = V$ *and* $B = A = |T|$. *This unique decomposition is called the polar decomposition of* T.

(c) If $T = U|T|$ *is the polar decomposition of* T, *then* $|T| = U^*T$.

Proof. (a) If $x, y \in \mathcal{H}$ are arbitrary, then,

$$\langle Tx, Ty\rangle \;=\; \langle T^*Tx, y\rangle \;=\; \langle |T|^2x, y\rangle \;= \langle |T|x, |T|y\rangle \;,$$

whence it follows – see Exercise 2.2.2 – that there exists a unique unitary operator $U_0 : \overline{\text{ran }|T|} \to \overline{\text{ran } T}$ such that $U_0(|T|x) = Tx \;\forall\; x \in \mathcal{H}$. Let $\mathcal{M} = \overline{\text{ran }|T|}$ and let $P = P_{\mathcal{M}}$ denote the orthogonal projection onto $\mathcal{M}$. Then the operator $U = U_0P$ clearly defines a partial isometry with initial space $\mathcal{M}$ and final space $\mathcal{N} = \overline{\text{ran } T}$ which further satisfies $T = U|T|$ (by definition). It follows from Lemma 3.1.1 that $\ker U = \ker|T| = \ker T$.

(b) Suppose $T = VB$ as in (b). Then V^*V is the projection onto $\ker^\perp V = \ker^\perp B = \overline{\text{ran } B}$, which clearly implies that $B = V^*VB$; hence, we see that $T^*T = BV^*VB = B^2$; thus B is a, and hence the, positive square root of $|T|^2$, i.e., $B = |T|$. It then follows that $V(|T|x) = Tx = U(|T|x)\;\forall x$; by continuity, we see that V agrees with U on $\overline{\text{ran }|T|}$, but since this is precisely the initial space of both partial isometries U and V, we see that we must have $U = V$.

(c) This is an immediate consequence of the definition of U and Exercise 3.1.4. □

EXERCISE 3.1.6. *(1) Prove the 'dual' polar decomposition theorem; i.e., each* $T \in B(\mathcal{H}, \mathcal{K})$ *can be uniquely expressed in the form* $T = BV$ *where* $V \in B(\mathcal{H}, \mathcal{K})$ *is a partial isometry,* $B \in B(\mathcal{K})$ *is a positive operator and* $\ker B = \ker V^* = \ker T^*$. *(Hint: Consider the usual polar decomposition of* T^*, *and take adjoints.)*

(2) Show that if $T = U|T|$ *is the (usual) polar decomposition of* T, *then* $U|_{\ker^\perp T}$ *implements a unitary equivalence between* $|T|\,|_{\ker^\perp|T|}$ *and* $|T^*|\,|_{\ker^\perp|T^*|}$.

(Hint: Write $\mathcal{M} = \ker^{\perp} T$, $\mathcal{N} = \ker^{\perp} T^*$, $W = U|_{\mathcal{M}}$*; then* $W \in B(\mathcal{M}, \mathcal{N})$ *is unitary; further* $|T^*|^2 = TT^* = U|T|^2U^*$*; deduce that if* A *(resp.,* B*) denotes the restriction of* $|T|$ *(resp.,* $|T^*|$*) to* $\mathcal{M}$ *(resp.,* $\mathcal{N}$*), then* $B^2 = WA^2W^*$*; now deduce, from the uniqueness of the positive square root, that* $B = WAW^*$*.)*

(3) Apply (2) above to the case when $\mathcal{H}$ *and* $\mathcal{K}$ *are finite-dimensional, and prove that if* $T \in L(V, W)$ *is a linear map of vector spaces (over* $\mathbb{C}$*), then* $dim\ V = rank(T) + nullity(T)$*, where* $rank(T)$ *and* $nullity(T)$ *denote the dimensions of the range and kernel (or null-space), respectively, of the map* T*.*

(4) Show that an operator $T \in B(\mathcal{H}, \mathcal{K})$ *can be expressed in the form* $T = WA$*, where* $A \in B(\mathcal{H})$ *is a positive operator and* $W \in B(\mathcal{H}, \mathcal{K})$ *is unitary if and only if* $dim(\ker T) = dim(\ker T^*)$*. (Hint: In order for such a decomposition to exist, show that it must be the case that* $A = |T|$ *and that the* W *should agree, on* $\ker^{\perp} T$*, with the* U *of the polar decomposition, so that* W *must map* $\ker T$ *isometrically onto* $\ker T^*$*.)*

(5) In particular, deduce from (4) that in case $\mathcal{H}$ *is a finite-dimensional inner product space, then any operator* $T \in B(\mathcal{H})$ *admits a decomposition as the product of a unitary operator and a positive operator. (In view of Proposition 2.8.1 1(c) and 2(c), note that when* $\mathcal{H} = \mathbb{C}$*, this boils down to the usual polar decomposition of a complex number.)*

Several problems concerning a general bounded operator between Hilbert spaces can be solved in two steps: in the first step, the problem is 'reduced', using the polar decomposition theorem, to a problem concerning positive operators on a Hilbert space; and in the next step, the positive case is settled using the spectral theorem. This is illustrated, for instance, in Exercise 3.1.7(2).

EXERCISE 3.1.7. *(1) Recall that a subset* Δ *of a (real or complex) vector space* V *is said to be* **convex** *if it contains the 'line segment joining any two of its points'; i.e.,* Δ *is convex if* $x, y \in \Delta, 0 \leq t \leq 1 \ \Rightarrow tx + (1-t)y \in \Delta$*.*

(a) If V *is a normed (or simply a topological) vector space, and if* Δ *is a closed subset of* V*, show that* Δ *is convex if and only if it contains the mid-point of any two of its points – i.e.,* Δ *is convex if and only if* $x, y \in \Delta \ \Rightarrow \ \frac{1}{2}(x+y) \in \Delta$*. (Hint: The set of dyadic rationals, i.e., numbers of the form* $\frac{k}{2^n}$ *is dense in* $\mathbb{R}$*.)*

(b) If $\mathcal{S} \subset V$ *is a subset of a vector space, show that there exists a smallest convex subset of* V *which contains* $\mathcal{S}$*; this set is called the* **convex hull** *of the set* $\mathcal{S}$ *and we shall denote it by the symbol* $co(\mathcal{S})$*. Show that* $co(\mathcal{S}) = \{\sum_{i=1}^n \theta_i x_i : n \in \mathbb{N}, x_i \in \mathcal{S}, \theta_i \geq 0, \sum_{i=1}^n \theta_i = 1\}$*.*

(c) Let Δ *be a convex subset of a vector space; show that the following conditions on a point* $x \in \Delta$ *are equivalent:*

(i) $x = \frac{1}{2}(y+z), \ y, z \in \Delta \ \Rightarrow \ x = y = z$*;*

(ii) $x = ty + (1-t)z, 0 < t < 1, \ y, z \in \Delta \ \Rightarrow x = y = z$*.*

The point x *is called an* **extreme point** *of a convex set* Δ *if* $x \in \Delta$ *and if* x *satisfies the equivalent conditions (i) and (ii) above.*

(d) It is a fact, called the Krein-Milman theorem – see [Yos], for instance – that if K *is a compact convex subset of a Banach space (or more generally, of*

a locally convex topological vector space which satisfies appropriate 'completeness conditions'), then $K = \overline{co(\partial_e K)}$, *where* $\partial_e K$ *denotes the set of extreme points of* K. *Note that the above equality can hold even without the compactness assumption as in the case* $K = ball(\mathcal{H}) = \{x \in \mathcal{H} : ||x|| \leq 1\}$, *where* $\mathcal{H}$ *is a Hilbert space, by showing that* $\partial_e(ball\ \mathcal{H}) = \{x \in \mathcal{H} : ||x|| = 1\}$. *(Hint: Use the parallelogram law – see Exercise* 1.2.3(4).)

(e) Show that $\partial_e(ball\ X) \neq \{x \in X : \ ||x|| = 1\}$, *when* $X = \ell^1_n,\ n > 1$. *(Thus, not every point on the unit sphere of a normed space need be an extreme point of the unit ball.)*

(2) Let $\mathcal{H}$ *and* $\mathcal{K}$ *denote (separable) Hilbert spaces, and let* $\mathbb{B} = \{A \in B(\mathcal{H}, \mathcal{K}) : ||A|| \leq 1\}$ *denote the unit ball of* $B(\mathcal{H}, \mathcal{K})$. *The aim of the following exercise is to show that an operator* $T \in \mathbb{B}$ *is an extreme point of* $\mathbb{B}$ *if and only if either* T *or* T^* *is an isometry. (See (1)(c) above, for the definition of an extreme point.)*

(a) Let $\mathbb{B}_+ = \{T \in B(\mathcal{H}) : T \geq 0,\ ||T|| \leq 1\}$. *Show that* $T \in \partial_e \mathbb{B}_+ \ \Leftrightarrow T$ *is a projection. (Hint: suppose* P *is a projection and* $P = \frac{1}{2}(A+B), A, B \in \mathbb{B}_+$; *then for arbitrary* $x \in ball(\mathcal{H})$, *note that* $0 \leq \frac{1}{2}(\langle Ax, x\rangle + \langle Bx, x\rangle) \leq 1$; *since* $\partial_e[0,1] = \{0,1\}$, *deduce that* $\langle Ax, x\rangle = \langle Bx, x\rangle = \langle Px, x\rangle\ \forall\ x \in (\ker P\ \cup$ ran $P)$; *but* $A \geq 0$ *and* $\ker P \subset\ \ker A$ *imply that* $A(\text{ran } P) \subset\ \text{ran } P$; *similarly also* $B(\text{ran } P) \subset\ \text{ran } P$; *conclude (from Exercise 1.4.16) that* $A = B = P$. *Conversely, if* $T \in \mathbb{B}_+$ *and* T *is not a projection, then it must be the case – see Proposition* 2.8.1(1)*(b) – that there exists* $\lambda \in \sigma(T)$ *such that* $0 < \lambda < 1$; *fix* $\epsilon > 0$ *such that* $(\lambda - 2\epsilon, \lambda + 2\epsilon) \subset (0,1)$; *since* $\lambda \in \sigma(T)$, *deduce that* $P \neq 0$ *where* $P = 1_{(\lambda-\epsilon,\lambda+\epsilon)}(T)$; *notice now that if we set* $A = T - \epsilon P, B = T + \epsilon P$, *then the choices ensure that* $A, B \in \mathbb{B}_+, T = \frac{1}{2}(A+B)$, *but* $A \neq T \neq B$, *whence* $T \notin \partial_e \mathbb{B}_+$.)

(b) Show that the only extreme point of ball $B(\mathcal{H}) = \{T \in B(\mathcal{H}) : ||T|| \leq 1\}$ *which is a positive operator is* 1, *the identity operator on* $\mathcal{H}$. *(Hint: Prove that 1 is an extreme point of ball* $B(\mathcal{H})$ *by using the fact that 1 is an extreme point of the unit disc in the complex plane; for the other implication, by (a) above, it is enough to show that if* P *is a projection which is not equal to 1, then* P *is not an extreme point in ball* $B(\mathcal{H})$; *if* $P \neq 1$, *note that* $P = \frac{1}{2}(U_+ + U_-)$, *where* $U_\pm = P \pm (1 - P)$.)

(c) Suppose $T \in \partial_e \mathbb{B}$; *if* $T = U|T|$ *is the polar decomposition of* T, *show that* $|T|\ |_{\mathcal{M}}$ *is an extreme point of the set* $\{A \in B(\mathcal{M}) : ||A|| \leq 1\}$, *where* $\mathcal{M} = \ker^{\perp}|T|$, *and hence deduce, from (b) above, that* $T = U$. *(Hint: if* $|T| = \frac{1}{2}(C + D)$, *with* $C, D \in ball\ B(\mathcal{M})$ *and* $C \neq |T| \neq D$, *note that* $T = \frac{1}{2}(A + B)$, *where* $A = UC, B = UD$, *and* $A \neq T \neq B$.)

(d) Show that $T \in \partial_e \mathbb{B}$ *if and only if* T *or* T^* *is an isometry. (Hint: suppose* T *is an isometry; suppose* $T = \frac{1}{2}(A + B)$, *with* $A, B \in \mathbb{B}$; *deduce from (1)(d) that* $Tx = Ax = Bx\ \forall x \in \mathcal{H}$; *thus* $T \in \partial_e \mathbb{B}$; *similarly, if* T^* *is an isometry, then* $T^* \in \partial_e \mathbb{B}$. *Conversely, if* $T \in \partial_e \mathbb{B}$, *deduce from (c) that* T *is a partial isometry; suppose it is possible to find unit vectors* $x \in \ker T,\ \ y \in \ker T^*$;

define $U_\pm z = Tz \pm \langle z, x\rangle y$, and note that $U_\pm$ are partial isometries which are distinct from T and that $T = \frac{1}{2}(U_+ + U_-)$.)

3.2 Compact operators

DEFINITION 3.2.1. *A linear map $T : X \to Y$ between Banach spaces is said to be* **compact** *if it satisfies the following condition: for every bounded sequence $\{x_n\}_n \subset X$, the sequence $\{Tx_n\}_n$ has a subsequence which converges with respect to the norm in Y.*

The collection of compact operators from X to Y is denoted by $B_0(X, Y)$ (or simply $B_0(X)$ if $X = Y$).

Thus, a linear map is compact precisely when it maps the unit ball of X into a set whose closure is compact – or equivalently, if it maps bounded sets into totally bounded sets[1]; in particular, every compact operator is bounded.

Although we have given the definition of a compact operator in the context of general Banach spaces, we shall really only be interested in the case of Hilbert spaces. Nevertheless, we state our first result for general Banach spaces, after which we shall specialise to the case of Hilbert spaces.

PROPOSITION 3.2.2. *Let X, Y, Z denote Banach spaces.*

(a) $B_0(X, Y)$ is a norm-closed subspace of $B(X, Y)$.

(b) If $A \in B(Y, Z), B \in B(X, Y)$, and if either A or B is compact, then AB is also compact.

(c) In particular, $B_0(X)$ is a closed two-sided ideal in the Banach algebra $B(X)$.

Proof. (a) Suppose $A, B \in B_0(X, Y)$ and $\alpha \in \mathbb{C}$, and suppose $\{x_n\}$ is a bounded sequence in X; since A is compact, there exists a subsequence – call it $\{y_n\}$ of $\{x_n\}$ – such that $\{Ay_n\}$ is a norm-convergent sequence; since $\{y_n\}$ is a bounded sequence and B is compact, we may extract a further subsequence - call it $\{z_n\}$ - with the property that $\{Bz_n\}$ is norm-convergent. It is clear then that $\{(\alpha A + B)z_n\}$ is a norm-convergent sequence; thus $(\alpha A + B)$ is compact; in other words, $B_0(X, Y)$ is a subspace of $B(X, Y)$.

Suppose now that $\{A_n\}$ is a sequence in $B_0(X, Y)$ and that $A \in B(X, Y)$ is such that $||A_n - A|| \to 0$. We wish to prove that A is compact. We will do this by a typical instance of the so-called 'diagonal argument'. Thus, suppose $S_0 = \{x_n\}$ is a bounded sequence in X. Since A_1 is compact, we can extract a subsequence $S_1 = \{x_n^{(1)}\}$ of S_0 such that $\{A_1 x_n^{(1)}\}$ is convergent in Y. Since A_2 is compact, we can extract a subsequence $S_2 = \{x_n^{(2)}\}$ of S_1 such that $\{A_2 x_n^{(2)}\}$ is convergent in Y. Proceeding in this fashion, we can find a sequence $\{S_k\}$ such that $S_k = \{x_n^{(k)}\}$ is a subsequence of S_{k-1} and $\{A_k x_n^{(k)}\}$ is convergent in

[1] Recall that a subset F of a metric space is said to be totally bounded if for every $\epsilon > 0$, it is possible to find a finite subset S such that $dist(x, S) < \epsilon \; \forall x \in F$; and that a subset of a metric space is compact if and only if it is complete and totally bounded.

Y, for each $k \geq 1$. Let us write $z_n = x_n^{(n)}$; since $\{z_n : n \geq k\}$ is a subsequence of S_k, note that $\{A_k z_n\}$ is a convergent sequence in Y, for every $k \geq 1$.

The proof of (a) will be completed once we establish that $\{Az_n\}$ is a Cauchy sequence in Y. Indeed, suppose $\epsilon > 0$ is given; let $K = 1 + \sup_n ||z_n||$; first pick an integer N such that $||A_N - A|| < \epsilon/3K$; next, choose an integer n_0 such that $||A_N z_n - A_N z_m|| < \epsilon/3 \ \forall \ n, m \geq n_0$; then observe that if $n, m \geq n_0$, we have:

$$\begin{aligned} ||Az_n - Az_m|| &\leq ||(A - A_N)z_n|| + ||A_N z_n - A_N z_m|| \\ &\quad + ||(A_N - A)z_m|| \\ &\leq \frac{\epsilon}{3K}K + \frac{\epsilon}{3} + \frac{\epsilon}{3K}K \\ &= \epsilon \ . \end{aligned}$$

(b) Let $\mathbb{B}$ denote the unit ball in X; we need to show that $(AB)(\mathbb{B})$ is totally bounded; this is true in case (i) A is compact, since then $B(\mathbb{B})$ is bounded, and A maps bounded sets to totally bounded sets, and (ii) B is compact, since then $B(\mathbb{B})$ is totally bounded, and A (being bounded and linear) maps totally bounded sets to totally bounded sets. □

Corollary 3.2.3. *Let $T \in B(\mathcal{H}_1, \mathcal{H}_2)$, where $\mathcal{H}_i$ are Hilbert spaces. Then*

(a) T is compact if and only if $|T|$ $(= \ (T^*T)^{\frac{1}{2}}\ )$ is compact; and consequently,

(b) T is compact if and only if T^* is compact.

Proof. If $T = U|T|$ is the polar decomposition of T, then also $U^*T = |T|$ – see Theorem 3.1.5; so each of T and $|T|$ is a multiple of the other. Now appeal to Proposition 3.2.2(b) to deduce (a) above. Also, since $T^* = |T|U^*$, we see that the compactness of T implies that of T^*; and (b) follows from the fact that we may interchange the roles of T and T^*. □

Exercise 3.2.4. *(1) Let X be a metric space; if $x, x_1, x_2, \ldots \in X$, show that the following conditions are equivalent:*

(i) the sequence $\{x_n\}$ converges to x;

(ii) every subsequence of $\{x_n\}$ has a further subsequence which converges to x.

(Hint: for the non-trivial implication, note that if the sequence $\{x_n\}$ does not converge to x, then there must exist a subsequence whose members are 'bounded away from x'.)

(2) Show that the following conditions on an operator $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ are equivalent:

(i) T is compact;

(ii) if $\{x_n\}$ is a sequence in $\mathcal{H}_1$ which converges weakly to 0 – i.e, $\langle x, x_n\rangle \to 0 \ \forall \ x \in \mathcal{H}_1$ – then $||Tx_n|| \to 0$.

(iii) if $\{e_n\}$ is any infinite orthonormal sequence in $\mathcal{H}_1$, then $||Te_n|| \to 0$.

(Hint: for (i) $\Rightarrow$ (ii), suppose $\{y_n\}$ is a subsequence of $\{x_n\}$; by compactness, there is a further subsequence $\{z_n\}$ of $\{y_n\}$ such that $\{Tz_n\}$ converges, to z, say; since $z_n \to 0$ weakly, deduce that $Tz_n \to 0$ weakly; this means $z = 0$, since strong convergence implies weak convergence; by (1) above, this proves (ii). The implication (ii) $\Rightarrow$ (iii) follows form the fact that any orthonormal sequence converges weakly to 0. For (iii) $\Rightarrow$ (i), deduce from Proposition 3.2.7(c) that if T is not compact, there exists an $\epsilon > 0$ such that $\mathcal{M}_\epsilon = \text{ran } 1_{[\epsilon,\infty)}(|T|)$ is infinite-dimensional; then any infinite orthonormal set $\{e_n : n \in \mathbb{N}\}$ in $\mathcal{M}_\epsilon$ would violate condition (iii).)

Recall that if $T \in B(\mathcal{H}, \mathcal{K})$ and if $\mathcal{M}$ is a subspace of $\mathcal{H}$, then T is said to be 'bounded below' on $\mathcal{M}$ if there exists an $\epsilon > 0$ such that $||Tx|| \geq \epsilon ||x|| \; \forall \, x \in \mathcal{M}$.

LEMMA 3.2.5. *If $T \in B_0(\mathcal{H}_1, \mathcal{H}_2)$ and if T is bounded below on a subspace $\mathcal{M}$ of $\mathcal{H}_1$, then $\mathcal{M}$ is finite-dimensional.*

In particular, if $\mathcal{N}$ is a closed subspace of $\mathcal{H}_2$ such that $\mathcal{N}$ is contained in the range of T, then $\mathcal{N}$ is finite-dimensional.

Proof. If T is bounded below on $\mathcal{M}$, then T is also bounded below (by the same constant) on $\overline{\mathcal{M}}$; we may therefore assume, without loss of generality, that $\mathcal{M}$ is closed. If $\mathcal{M}$ contains an infinite orthonormal set, say $\{e_n : n \in \mathbb{N}\}$, and if T is bounded below by ϵ on $\mathcal{M}$, then note that $||Te_n - Te_m|| \geq \epsilon\sqrt{2} \; \forall n \neq m$; then $\{e_n\}$ would be a bounded sequence in $\mathcal{H}$ such that $\{Te_n\}$ had no Cauchy subsequence, thus contradicting the assumed compactness of T; hence $\mathcal{M}$ must be finite-dimensional.

As for the second assertion, let $\mathcal{M} = T^{-1}(\mathcal{N}) \cap (\ker^\perp T)$; note that T maps $\mathcal{M}$ 1-1 onto $\mathcal{N}$; by the open mapping theorem, T must be bounded below on $\mathcal{M}$; hence by the first assertion of this Lemma, $\mathcal{M}$ is finite-dimensional, and so also is $\mathcal{N}$. □

The purpose of the next exercise is to convince the reader of the fact that compactness is an essentially 'separable phenomenon', so that our restricting ourselves to separable Hilbert spaces is essentially of no real loss of generality, as far as compact operators are concerned.

EXERCISE 3.2.6. *(In the following problem and in the sequel, while discussing the continuous functional calculus of self-adjoint operators, we shall blur the distinction between bounded continuous functions defined on all of $\mathbb{R}$ or only on $\sigma(T)$ for some self-adjoint T; this should lead to no confusion, since $f(T)$ depends only on $f|_{\sigma(T)}$.)*

(a) Let $T \in B(\mathcal{H})$ be a positive operator on a (possibly non-separable) Hilbert space $\mathcal{H}$. Let $\epsilon > 0$ and let $\mathcal{S}_\epsilon = \{f(T)x : f \in C(\sigma(T)), \; f(t) = 0 \; \forall t \in [0, \epsilon]\}$. If $\mathcal{M}_\epsilon = [\mathcal{S}_\epsilon]$ denotes the closed subspace generated by $\mathcal{S}_\epsilon$, then show that $\mathcal{M}_\epsilon \subset \text{ ran } T$. (Hint: let $g \in C(\sigma(T))$ be any continuous function such that

$g(t) = t^{-1}$ *whenever* $t \geq \epsilon/2$*; for instance, you could take*

$$g(t) = \begin{cases} \frac{1}{t} & \text{if } t \geq \epsilon/2 \\ \frac{4t}{\epsilon^2} & \text{if } 0 \leq t \leq \epsilon/2 \end{cases} ;$$

then notice that if $f \in C(\sigma(T))$ *satisfies* $f(t) = 0 \ \forall t \leq \epsilon$, *then* $f(t) = tg(t)f(t) \ \forall t$*; deduce that* $\mathcal{S}_\epsilon$ *is a subset of* $\mathcal{N} = \{z \in \mathcal{H} : z = Tg(T)z\}$*; but* $\mathcal{N}$ *is a closed subspace of* $\mathcal{H}$ *which is contained in* ran T*.)*

(b) Let $T \in B_0(\mathcal{H}_1, \mathcal{H}_2)$*, where* $\mathcal{H}_1, \mathcal{H}_2$ *are arbitrary (possibly non-separable) Hilbert spaces. Show that* $\ker^\perp T$ *and* $\overline{\text{ran } T}$ *are separable Hilbert spaces. (Hint: Let* $T = U|T|$ *be the polar decomposition, and let* $\mathcal{M}_\epsilon$ *be associated to* $|T|$ *as in (a) above; show that* $U(\mathcal{M}_\epsilon)$ *is a closed subspace of* ran T *and deduce from Lemma* 3.2.5 *that* $\mathcal{M}_\epsilon$ *is finite-dimensional; note that* $\ker^\perp T = \ker^\perp |T|$ *is the closure of* $\bigcup_{n=1}^\infty \mathcal{M}_{\frac{1}{n}}$*, and that* $\overline{\text{ran } T} = U(\ker^\perp T)$*.)*

We now return to our standing assumption that all Hilbert spaces are separable.

Proposition 3.2.7. *The following conditions on an operator* $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ *are equivalent:*

(a) T *is compact;*

(b) $|T|$ *is compact;*

(c) ran $1_{[\epsilon,\infty)}(|T|)$ *is finite-dimensional, for every* $\epsilon > 0$*;*

(d) there is a sequence $\{T_n\}_{n=1}^\infty \subset B(\mathcal{H}_1, \mathcal{H}_2)$ *such that (i)* $||T_n - T|| \to 0$*, and (ii)* ran T_n *is finite-dimensional, for each* n*;*

(e) ran T *does not contain any infinite-dimensional closed subspace of* $\mathcal{H}_2$*.*

Proof. For $\epsilon > 0$, let us use the notation $1_\epsilon = 1_{[\epsilon,\infty)}$ and $P_\epsilon = 1_\epsilon(|T|)$.

$(a) \Rightarrow (b)$: See Corollary 3.2.3.

$(b) \Rightarrow (c)$: Since $t \geq \epsilon 1_\epsilon(t) \ \forall \ t \geq 0$, we find easily that $|T|$ is bounded below (by ϵ) on ran P_ϵ, and (c) follows from Lemma 3.2.5.

$(c) \Rightarrow (d)$: Define $T_n = TP_{\frac{1}{n}}$; notice that $0 \leq t(1 - 1_{\frac{1}{n}}(t)) \leq \frac{1}{n} \ \forall \ t \geq 0$; conclude that $|| \ |T|(1 - 1_{\frac{1}{n}}(|T|))|| \leq \frac{1}{n}$; if $T = U|T|$ is the polar decomposition of T, deduce that $||T - T_n|| \leq \frac{1}{n}$; finally, the condition (c) clearly implies that each ($P_{\frac{1}{n}}$ and consequently) T_n has finite-dimensional range.

$(d) \Rightarrow (a)$: In view of Proposition 3.2.2(a), it suffices to show that each T_n is a compact operator; but any bounded operator with finite-dimensional range is necessarily compact, since any bounded set in a finite-dimensional space is totally bounded.

$(a) \Rightarrow (e)$: See Lemma 3.2.5.

$(e) \Rightarrow (c)$: Pick any bounded measurable function g such that $g(t) = \frac{1}{t}$, $\forall t \geq \epsilon$; then $tg(t) = 1 \ \forall t \geq \epsilon$, whence $tg(t)1_\epsilon(t) = 1_\epsilon(t) \ \forall t$, and $|T|g(|T|)P_\epsilon = P_\epsilon$; hence ran P_ϵ is a closed subspace of (ran $|T|$, and consequently of) the initial space of the partial isometry U; deduce that U(ran P_ϵ) is a closed subspace of ran T; by condition (e), this implies that $U(\text{ran } P_\epsilon)$ is finite-dimensional. As U is isometric on ran P_ϵ, we see that ran P_ϵ is finite-dimensional, as desired. □

We now discuss normal compact operators.

PROPOSITION 3.2.8. *Let $T \in B_0(\mathcal{H})$ be a normal (compact) operator on a separable Hilbert space, and let $E \mapsto P(E) = 1_E(T)$ be the associated spectral measure.*

(a) If $\epsilon > 0$, let $P_\epsilon = P(\{\lambda \in \mathbb{C} : |\lambda| \geq \epsilon\})$ denote the spectral projection associated to the complement of the ϵ-neighbourhood of 0. Then ran P_ϵ *is finite-dimensional.*

(b) If $\Sigma^* = \sigma(T) \setminus \{0\}$, then

(i) $\lambda \in \Sigma^* \Rightarrow \lambda$ is an eigenvalue of finite multiplicity; that is, $0 < \dim\ker(T-\lambda) < \infty$;

(ii) With $P(\lambda) = 1_T(\{\lambda\})$, the subspaces $\{\text{ran } P(\lambda) : \lambda \in \Sigma^*\}$ are pairwise orthogonal finite-dimensional subspaces, so Σ^* is a countable set and ran $\left(\sum_{\lambda\in\Sigma^*} P(\{\lambda\})\right) = \ker^\perp(T)$;

(iii) the only possible accumulation point of Σ^* is 0; and

(iv) there exist a countable set N and scalars $\lambda_n \in \Sigma^*, n \in N$ such that

$$Tx \;=\; \sum_{n\in N} \lambda_n \; P(\{\lambda_n\}) \; .$$

Proof. (a) Note that the function defined by the equation

$$g(\lambda) \;=\; \begin{cases} \frac{1}{\lambda} & \text{if } |\lambda| \geq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

is a bounded measurable function on $\sigma(T)$ such that $g(\lambda)\lambda = 1_{F_\epsilon}(\lambda)$, where $F_\epsilon = \{z \in \mathbb{C} : |z| \geq \epsilon\}$. It follows that $g(T)T = Tg(T) = P_\epsilon$. Hence ran P_ϵ is contained in ran T, and the desired conclusion follows from Proposition 3.2.7(e).

(b) (i) By Theorem 1.6.2 (2), if $\lambda \in \Sigma^*$, there exists a sequence of unit vectors x_n such that $\lim_{n\to\infty} \|Ax_n - \lambda x_n\| = 0$. By compactness, there is some subsequence $\{x_{n_k}\}$ such that $y = \lim_{k\to\infty} Ax_{n_k}$ exists. Since $\|y\| = \lim_{k\to\infty} \|Ax_{n_k}\| = |\lambda| \neq 0$ it follows that $x_{n_k} \to x$ (so $\|x\| = 1$) and that $Ax = \lambda x$. So λ is an eigenvalue of T. Also, it follows from Proposition 2.4.1(5) that ran $P(\lambda) = \ker(T-\lambda)$. Since A is clearly bounded below (by $|\lambda|$) on $\ker(A-\lambda)$, it is seen from Lemma 3.2.5 that λ is an eigenvalue of finite multiplicity.

(ii) If $\lambda \in \Sigma^*$, and $x \in \ker(A-\lambda)$, note that also

$$\|(A^* - \bar{\lambda})x\| = \|(A-\lambda)^* x\| = 0$$

since $(A-\lambda)$ inherits normality from A. Hence if $\lambda, \mu \in \Sigma^*$, if $\lambda \neq \mu$ and if x and y are eigenvectors of A corresponding to λ and μ respectively, then

$$\lambda\langle x, y\rangle = \langle Ax, y\rangle = \langle x, A^*y\rangle = \mu\langle x, y\rangle$$

and we find that $\{\ker(A - \lambda) = \text{ran } P(\{\lambda\}) : \lambda \in \Sigma^*\}$ is a set of pairwise orthogonal non-zero subspaces of the separable space $\mathcal{H}$. Hence Σ^* must be countable. It follows that

$$\begin{aligned}\Big(\sum_{\lambda\in\Sigma^*} P(\{\lambda\})\Big) &= P(\Sigma^*) \\ &= P(\mathbb{C}\setminus\{0\}) \\ &= \text{ projection onto } \ker^\perp(T) \ .\end{aligned}$$

(iii) For any $\epsilon > 0$, Lemma 3.2.5 implies that ran P_ϵ is finite-dimensional. But clearly $P_\epsilon = \sum_{\lambda\in\Sigma^*,|\lambda|>\epsilon} \text{ran } P(\{\lambda\})$; so it must be that $\{\lambda \in \Sigma^* : |\lambda| > \epsilon\}$ is finite, thereby establishing (iii).

(iv) $T = TP(\mathbb{C}) = TP(\Sigma^*) = \sum_{\lambda\in\Sigma^*} TP(\{\lambda\}) = \sum_{\lambda\in\Sigma^*} \lambda P(\{\lambda\})$. □

Exercise 3.2.9. *Let X be a compact Hausdorff space and let $\mathcal{B}_X \ni E \mapsto P(E)$ be a spectral measure; let μ be a measure which is 'mutually absolutely continuous' with respect to P – thus, for instance, we may* (*see Remark* 2.6.3) *take $\mu(E) = \sum ||P(E)e_n||^2$, where $\{e_n\}$ is some orthonormal basis for the underlying Hilbert space $\mathcal{H}$ – and let $\pi : C(X) \to B(\mathcal{H})$ be the associated representation.*

(a) Show that the following conditions are equivalent:

(i) $\mathcal{H}$ is finite-dimensional;

(ii) there exists a finite set $F \subset X$ such that $\mu = \mu|_F$, and such that ran $P(\{x\})$ *is finite-dimensional, for each $x \in F$.*

(b) If $x_0 \in X$, show that the following conditions on a vector $x \in \mathcal{H}$ are equivalent:

(i) $\pi(f)x = f(x_0)x \ \forall\ f \in C(X)$;

(ii) $x \in$ ran $P(\{x_0\})$.

(*Hint: See Corollary* 2.4.3.)

By piecing together our description in Proposition 3.2.8 – applied to $|T|$ – and the polar decomposition, we arrive at the useful 'singular value decomposition' of a general compact operator.

Proposition 3.2.10. *Suppose $T \in B_0(\mathcal{H}_1, \mathcal{H}_2)$ has polar decomposition $T = U|T|$. Then, with $P(\{\lambda\}) = 1_{|T|}(\{\lambda\})$ (as in Proposition* 3.2.8 *applied to $|T|$), we have*

(1) $\sigma(|T|) = \Sigma^* \coprod\{0\}$ where Σ^* admits an enumeration $\Sigma^* = \{\lambda_n : n \in N\}$ for some countable set N; if Σ^* is finite, then T has finite rank; and if N is infinite, we may assume without loss of generality that $N = \mathbb{N}$ and that $\lambda_1 > \lambda_2 > \cdots > \lambda_n > \lambda_{n+1} > \cdots$ and $\lambda_n \downarrow 0$

(2) $T = \sum_{n\in N} \lambda_n UP(\{\lambda_n\})$; more explicitly, if $\{x_k^{(n)} : k \in I_n\}$ is an orthonormal basis for ran $P(\{\lambda_n\})$ *and $y_k^{(n)} = Ux_k^{(n)}$, then*

$$Tx = \sum_{n\in N} \lambda_n \Big(\sum_{k\in I_n} \langle x, x_k^{(n)}\rangle y_k^{(n)}\Big) . \tag{3.2.3}$$

Here, and in the rest of this chapter the symbol N will always denote one of the sets $\{1, 2, \ldots, n\}$ [2]

The sequence $\{s_n = s_n(T) : n < \dim(\mathcal{H}_1) + 1\}$ defined by

$$s_n \;=\; \begin{cases} \lambda_1 & \text{if } 0 < n \;\leq\; \text{card}\,(I_1) \\ \lambda_2 & \text{if card}\,(I_1) < n \leq (\text{card}\,(I_1) \;+\; \text{card}\,(I_2)) \\ \cdots & \\ \lambda_m & \text{if } \sum_{1 \leq k < m} \text{card}\,(I_k) < n \leq \sum_{1 \leq k \leq m} \text{card}\,(I_k) \\ 0 & \text{if } \sum_{k \in N} \text{card}\,(I_k) < n \end{cases} \tag{3.2.4}$$

$\{s_n(T)\}_n$ is the non-increasing arrangement (counting multiplicity) of the eigenvalues of $|T|$, and the s_n's are called the **singular values** *of the compact operator T.*

The proof is essentially spelt out in the statement of the Proposition itself, and is left as an exercise to the reader. A less verbose – and slightly less informative – way to rephrase the content of Proposition 3.2.10 uses the following notation: for $x \in \mathcal{H}, y \in \mathcal{K}$, write $(\bar{x} \otimes y)$ for the rank one operator in $B(\mathcal{H}, \mathcal{K})$ given by $(\bar{x} \otimes y)x' = \langle x', x \rangle y$.

EXERCISE 3.2.11. *If $x \in \mathcal{H}, y, y' \in \mathcal{K}, z \in \mathcal{M}$, then verify that $(\bar{y'} \otimes z)(\bar{x} \otimes y) = \langle y, y' \rangle (\bar{x} \otimes z)$.*

Singular value decomposition: *If $s_1, s_2, \ldots$ are the singular values of a compact operator T, then T admits a so-called* **singular value decomposition** *(sometimes abbreviated to SVD)*

$$T = \sum_{n \in N} s_n(T) \; (\bar{x}_n \otimes y_n) \tag{3.2.5}$$

where $\{x_n : n \in N\}$ (resp., $\{y_n : n \in N\}$) is an orthonormal basis for $\ker^\perp(T)$, (resp., $\ker^\perp(T^*)$). (Note that even if $\dim \operatorname{ran} T < \dim(\mathcal{H})$, the singular values s_n are defined for $n < \dim(\mathcal{H}) + 1$.)

Note that while the singular values are uniquely determined, this is no longer true for the SVD, since, for instance, $\ker(|T| - \lambda)$ might have dimension more than one for some λ. Some more useful properties of singular values are listed in the following exercises.

EXERCISE 3.2.12. *(1) Let $T \in B(\mathcal{H})$ be a positive compact operator on a Hilbert space. In this case, we write $\lambda_n = s_n(T)$, since ($T = |T|$ and consequently) each λ_n is then an eigenvalue of T. The purpose of this set of exercises is to prove and study some consequences of*

$$\lambda_n \;=\; \max_{\dim \mathcal{M} \leq n} \min\{\langle Tx, x \rangle : x \in \mathcal{M}, ||x|| = 1\} \;, \tag{3.2.6}$$

[2] Note that $n = 0 \Leftrightarrow N = \emptyset \Leftrightarrow T = 0$.

where the the maximum is to be interpreted as a supremum, over the collection of all subspaces $\mathcal{M} \subset \mathcal{H}$ with appropriate dimension, and part of the assertion of the exercise is that this supremum is actually attained (and is consequently a maximum); in a similar fashion, the minimum is to be interpreted as an infimum which is attained. (This identity is called the **max-min principle** *and is also referred to as the* **Rayleigh-Ritz principle**.*)*

(2) *Define $\mathcal{M}_n = [\{x_j : 1 \leq j \leq n\}]$, where $\{x_k : k \in N\}$ are as in Equation* (3.2.5)*; observe that $\lambda_n = \min\{\langle Tx, x\rangle : x \in \mathcal{M}_n, ||x|| = 1\}$; this proves the inequality $\leq$ in* (3.2.6). *Conversely, if $\dim \mathcal{M} \leq n$, argue that there must exist a unit vector $x_0 \in \mathcal{M} \cap \mathcal{M}_{n-1}^{\perp}$ (since the projection onto $\mathcal{M}_{n-1}$ cannot be injective on $\mathcal{M}$), to conclude that $\min\{\langle Tx, x\rangle : x \in \mathcal{M}, ||x|| = 1\} \leq \lambda_n$.*

(3) *If $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a compact operator between Hilbert spaces, show that*

$$s_n(T) = \max_{dim \mathcal{M} \leq n} \min\{||Tx|| : x \in \mathcal{M}, ||x|| = 1\} . \tag{3.2.7}$$

(Hint: Note that $||Tx||^2 = \langle |T|^2 x, x\rangle$, apply equation (3.2.6) *above to $|T|^2$, and note that $s_n(|T|^2) = s_n(|T|)^2$.)*

(4) $T \in B_0(\mathcal{H}) \Rightarrow s_1(T) = \|T\|$. *(*Hint: *This is the case $n = 1$ of Equation* (3.2.7) *above.)*

(5) *If T is compact, show that*

$$s_n(T) = \min\{\|T - F\| : \dim(\mathrm{ran}\,(F)) < n\} \ \forall n \geq 1 .$$

(Hint: If $s_n(T) = 0$, then $\dim(\mathrm{ran}\,(T)) < n$ and this equality is obvious. So assume $s_n(T) > 0$.

Prove two inequalities. Let Equation (3.2.5) *be the SVD of T. If $F_n = \sum_{k=1}^{n-1} s_k \bar{x}_k \otimes y_k$, then $\dim(\mathrm{ran}\,(F_n)) < n$ and $\|T - F_n\| = s_n(T)$ since $\{s_n(T)\}$ is a non-increasing sequence; so indeed*

$$s_n(T) \leq \inf\{\|T - F\| : dim(\mathrm{ran}\,(F)) < n\} \ \forall n \geq 1\} .$$

Conversely suppose F is an operator whose range (call it $\mathcal{M}$) has dimension less than n. Let $\mathcal{N} = \ker(F)$. Note first that $\mathrm{ran}(F)$ and $\mathrm{ran}\,(F^*)$ have the same dimension (by polar decomposition) and hence $\dim(\mathcal{N}^{\perp}) = \dim(\mathrm{ran}(F^*)) < n$. Let $\mathcal{M}_n = [\{x_1, \ldots, x_n\}]$. The assumption $s_n(T) > 0$ implies that $\dim(\mathcal{M}_n) = n$. If P denotes the projection onto $\mathcal{N}^{\perp}$, it follows that $P|_{\mathcal{M}_n}$ cannot be injective; since $\ker(P) = \mathcal{N}$, we can find a unit vector $x \in \mathcal{M}_n \cap \mathcal{N}$. Then,

$$\begin{aligned} \|(T - F)x\| &= \|Tx\| \\ &\geq \min\{\|Tz\| : z \in S(\mathcal{M}_n)\} \\ &= s_n(T) , \end{aligned}$$

thus yielding the reverse inequality.

Since $\lim_{n\to\infty} s_n(T) = 0$*, this exercise also gives another proof of Proposition* 3.2.7 (*d*).)

(6) *If* $T \in B_0(\mathcal{H}_1, \mathcal{H}_2)$ *and if* $s_n(T) > 0$*, then show that* $n \leq dim(\mathcal{H}_2)$ *(so that* $s_n(T^*)$ *is defined) and* $s_n(T^*) = s_n(T)$*. (Hint: Use Exercise* 3.1.6 (2) *or the polar decomposition; in fact if* $T = \sum s_n(\bar{x}_n \otimes y_n)$ *is an SVD of* T*, then* $T^* = \sum s_n(\bar{y}_n \otimes x_n)$ *is an SVD of* T^* *.)*

3.3 von Neumann-Schatten ideals

We begin with a brief survey of this section which might help to motivate the notation and development of this section. This section may be regarded as containing non-commutative analogues of their 'commutative counterparts' of facts regarding classical sequence spaces $c_0, \ell^2, \ell^1, \ell^\infty$ and finally also $\ell^p, 1 < p < \infty$ as well as the duality relations among them. Specifically, we shall establish the following facts in this section:

(1) The following conditions on an operator $T \in B(\mathcal{H}, \mathcal{K})$ are equivalent, where we assume that $\mathcal{H}, \mathcal{K}$ are infinite-dimensional and separable:

  (1) There exists an orthonormal basis $\{x_n : n \in N\}$ (resp., $\{y_n : n \in N\}$) for $\ker^\perp(T)$ (resp., $\ker^\perp(T^*)$) such that $Tx_n = s_n y_n \ \forall n \in N$ for a sequence $s(T) = \{s_n : n \in \mathbb{N}\}$ of positive scalars such that $s(T) \in c_0$ (resp., $s(T) \in \ell^p$ where $p \in [1, \infty)$).

  (2) $T \in B_0(\mathcal{H}, \mathcal{K})$ (resp., $T \in B^p(\mathcal{H}, \mathcal{K})$) [3]

(2) With the notation of item (1) above, the sets $B^p(\mathcal{H}, \mathcal{K})$ are Banach spaces – called the **von Neumann-Schatten classes** – when equipped with the norms given by $\|T\|_p = (\sum s_n(T)^p)^{\frac{1}{p}}$.

(3) $B_0(\mathcal{H})^* = B^1(\mathcal{H})$ (where $B_0(\mathcal{H})$ is viewed as a subspace of $B(\mathcal{H})$), and there exist natural identifications $B^p(\mathcal{H})^* = B^q(\mathcal{H})$ for $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, and $B^1(\mathcal{H})^* = B(\mathcal{H})$.

(4) Each $B^p(\mathcal{H})$ is a (non-closed) two-sided ideal in $B(\mathcal{H})$ and the *Schatten p-norm* $\|\cdot\|_p$ is unitarily invariant in the sense that $\|UTV\|_p = \|T\|_p$ whenever U and V are unitary.

(5) The space $B_{00}(\mathcal{H})$ of finite rank operators on $\mathcal{H}$ is a dense linear subspace of each Banach space $B^p(\mathcal{H}), 1 \leq p < \infty$(all of which are, in turn, dense subspaces of the Banach space $B_0(\mathcal{H})$.

[3] Thus $B^p(\mathcal{H}, \mathcal{K}) = \{T \in B_0(\mathcal{H}, \mathcal{K})\}; s(T) \in \ell^p\}$ where $p \in [1, \infty)$.

3.3.1 Hilbert-Schmidt operators

LEMMA 3.3.1. *The following conditions on a linear operator $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ are equivalent:*

(i) $\sum_n ||Te_n||^2 < \infty$, *for some orthonormal basis* $\{e_n\}$ *of* $\mathcal{H}_1$*;*

(ii) $\sum_m ||T^* f_m||^2 < \infty$, *for every orthonormal basis* $\{f_m\}$ *of* $\mathcal{H}_2$.

(iii) $\sum_n ||Te_n||^2 < \infty$, *for every orthonormal basis* $\{e_n\}$ *of* $\mathcal{H}_1$, *with this sum being independent of the chosen orthonormal basis.*

If these equivalent conditions are satisfied, then the sums of the series in (ii) and (iii) are independent of the choice of the orthonormal bases and are all equal to one another.

Proof. If $\{e_n\}$ (resp., $\{f_m\}$) is any orthonormal basis for $\mathcal{H}_1$ (resp., $\mathcal{H}_2$), then note that

$$\begin{aligned}\sum_n ||Te_n||^2 &= \sum_n \sum_m |\langle Te_n, f_m \rangle|^2 \\ &= \sum_m \sum_n |\langle T^* f_m, e_n \rangle|^2 \\ &= \sum_m ||T^* f_m||^2 ,\end{aligned}$$

and all the assertions of the proposition are seen to follow. □

DEFINITION 3.3.2. *An operator $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ is said to be a* **Hilbert-Schmidt** *operator if it satisfies the equivalent conditions of Lemma* 3.3.1*, and the* **Hilbert-Schmidt norm** *of such an operator is defined by*

$$||T||_2 := \Big(\sum_n ||Te_n||^2 \Big)^{\frac{1}{2}} , \tag{3.3.8}$$

where $\{e_n\}$ is any [4] *orthonormal basis for $\mathcal{H}_1$. The collection of all Hilbert-Schmidt operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ will be denoted by $B^2(\mathcal{H}_1, \mathcal{H}_2)$.*

Some elementary properties of the class of Hilbert-Schmidt operators are contained in the following proposition.

PROPOSITION 3.3.3. *Suppose $T \in B(\mathcal{H}_1, \mathcal{H}_2)$, $S \in B(\mathcal{H}_2, \mathcal{H}_3)$, where $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ are Hilbert spaces.*

(a) $T \in B^2(\mathcal{H}_1, \mathcal{H}_2) \Rightarrow T^* \in B^2(\mathcal{H}_2, \mathcal{H}_1)$*; moreover,* $||T^*||_2 = ||T||_2 \geq ||T||_\infty$, *where we write* $||\cdot||_\infty$ *to denote the usual operator norm;*

(b) if either S or T is a Hilbert-Schmidt operator, so is ST, and

$$||ST||_2 \leq \begin{cases} ||S||_2 ||T||_\infty & \text{if } S \in B^2(\mathcal{H}_2, \mathcal{H}_3) \\ ||S||_\infty ||T||_2 & \text{if } T \in B^2(\mathcal{H}_1, \mathcal{H}_2) \ ; \end{cases} \tag{3.3.9}$$

(c) $B^2(\mathcal{H}_1, \mathcal{H}_2) \subset B_0(\mathcal{H}_1, \mathcal{H}_2)$*;*

[4] This adjective is justified by Proposition 3.3.1(*iii*).

(d) if $T \in B_0(\mathcal{H}_1, \mathcal{H}_2)$, then T is a Hilbert-Schmidt operator if and only if $\sum_n s_n(T)^2 < \infty$; in fact,

$$||T||_2^2 = \sum_n s_n(T)^2 .$$

Proof. (a) The equality $||T||_2 = ||T^*||_2$ was proved in Lemma 3.3.1. If x is any unit vector in $\mathcal{H}_1$, pick an orthonormal basis $\{e_n\}$ for $\mathcal{H}_1$ such that $e_1 = x$, and note that

$$||T||_2 = \left(\sum_n ||Te_n||^2 \right)^{\frac{1}{2}} \geq ||Tx|| ;$$

since x was an arbitrary unit vector in $\mathcal{H}_1$, deduce that $||T||_2 \geq ||T||_\infty$, as desired.

(b) Suppose T is a Hilbert-Schmidt operator; then, for an arbitrary orthonormal basis $\{e_n\}$ of $\mathcal{H}_1$, we find that

$$\sum_n ||STe_n||^2 \leq ||S||_\infty^2 \sum_n ||Te_n||^2 ,$$

whence ST is also a Hilbert-Schmidt operator and that $||ST||_2 \leq ||S||_\infty ||T||_2$; if T is a Hilbert-Schmidt operator, then, so is T^*, and by the already proved case, also S^*T^* is a Hilbert-Schmidt operator, and

$$||TS||_2 = ||(TS)^*||_2 \leq ||S^*||_\infty ||T^*||_2 = ||S||_\infty ||T||_2 .$$

(c) Let $\mathcal{M}_\epsilon = \operatorname{ran} 1_{[\epsilon,\infty)}(|T|)$; then $\mathcal{M}_\epsilon$ is a closed subspace of $\mathcal{H}_1$ on which T is bounded below, by ϵ; so, if $\{e_1, \ldots, e_N\}$ is any orthonormal set in $\mathcal{M}_\epsilon$, we find that $N\epsilon^2 \leq \sum_{n=1}^N ||Te_n||^2 \leq ||T||_2^2$, which clearly implies that *dim* $\mathcal{M}_\epsilon$ is finite (and can not be greater than $(||T||_2/\epsilon)^2$). We may now infer from Proposition 3.2.7 that T is necessarily compact.

(d) Let $Tx = \sum_n s_n(T)\langle x, x_n\rangle y_n$ for all $x \in \mathcal{H}_1$, as in Equation (3.2.5), for an appropriate orthonormal (finite or infinite) sequence $\{x_n\}$ (resp., $\{y_n\}$) in $\mathcal{H}_1$ (resp., in $\mathcal{H}_2$). Then notice that $||Tx_n|| = s_n(T)$ and that $Tx = 0$ if $x \perp x_n \ \forall n$. If we compute the Hilbert-Schmidt norm of T with respect to an orthonormal basis obtained by extending the orthonormal set $\{x_n\}$, we find that $||T||_2^2 = \sum_n s_n(T)^2$, as desired. □

REMARK 3.3.4. (1) $B^2(\mathcal{H})$ is a two-sided ideal in $B_0(\mathcal{H})$. (This follows from Proposition 3.3.3 (b), (d) and the fact that $\ell^2(\mathbb{N})$ is a vector space.)

(2) The set $B_{00}(\mathcal{H})$ of all finite rank operators is the smallest non-zero (two-sided) ideal of $B(\mathcal{H})$. (*Reason:* If $\mathcal{I}$ is any non-zero ideal in $B(\mathcal{H})$, there exists a $T \in \mathcal{I}$ and $x_0, y_0 \in \mathcal{H} \setminus \{0\}$ such that $Tx_0 = y_0$. Then, for any $0 \neq x, y \in \mathcal{H}$, we have

$$\bar{x} \otimes y = \|y_0\|^{-2}(\bar{y_0} \otimes y)T(\bar{x} \otimes x_0) \in \mathcal{I}$$

and we are done, because $B_{00}(\mathcal{H})$ is linearly spanned by operators of the form $(\bar{x} \otimes y)$.)

(3) $B^2(\mathcal{H},\mathcal{K})$ is a Hilbert space with respect to the inner product given by

$$\langle S, T\rangle_\epsilon = \sum \langle Se_n, Te_n\rangle$$

for an orthonormal basis $\epsilon = \{e_n\}$ of $\mathcal{H}$; and this definition is independent of the orthonormal basis ϵ. (*Reason:* For any orthonormal basis ϵ of $\mathcal{H}$, this series is convergent (by two applications of the Cauchy-Schwarz inequality, once in $\mathcal{K}$ and then again in ℓ^2) and is easily seen to define a sesquilinear form B_ϵ on $B^2(\mathcal{H},\mathcal{K})$ with associated quadratic form q_ϵ being independent of the orthonormal basis ϵ. It follows from the polarisation identity that B_ϵ is also independent of ϵ. As $q_\epsilon(T) \geq \|T\|_\infty$, any q_ϵ-Cauchy sequence $\{T_n : n \in \mathbb{N}\}$ in $B^2(\mathcal{H},\mathcal{K})$ is also a Cauchy sequence in $B(\mathcal{H},\mathcal{K})$. If $T \in B(\mathcal{H},\mathcal{K})$ and $\|T_n - T\| \to 0$, it follows from the boundedness of $\{q_\epsilon(T_n) : n \in \mathbb{N}\}$ that also $T \in B^2(\mathcal{H},\mathcal{K})$ and that $q_\epsilon(T_n - T) \to 0$. If $\bar{\mathcal{H}}$ denotes the 'conjugate Hilbert space' of $\mathcal{H}$ - with an anti-unitary operator $\mathcal{H} \ni f \mapsto \bar{f}$ – it is not hard to show that $\{(\bar{e}_i \otimes f_j) : i, j \in \mathbb{N}\}$ is an orthonormal basis for $B^2(\mathcal{H},\mathcal{K})$ whenever $\{e_n : n \in \mathbb{N}\}$ (resp., $\{f_n : n \in \mathbb{N}\}$) is an orthonormal basis for $\mathcal{H}$ (resp., $\mathcal{K}$), and hence we have a natural identification $B^2(\mathcal{H},\mathcal{K}) \cong \bar{\mathcal{H}} \otimes \mathcal{K}$.)

(4) $B_{00}(\mathcal{H})$ is dense in the Hilbert space $B^2(\mathcal{H})$ as well as in the Banach space $B_0(\mathcal{H})$. (The SVD of a compact operator is the fastest way to see this.)

(5) $B_0(\mathcal{H})$ is the only non-trivial closed ideal in $B(\mathcal{H})$. (*Reason:* If $\mathcal{I}$ is any non-zero closed ideal in $B(\mathcal{H})$, it follows from items (2) and (4) above that $B_0(\mathcal{H}) \subset \mathcal{I}$. It suffices to show that $B_0(\mathcal{H})$ is the largest ideal in $B(\mathcal{H})$. (This requires separability of $\mathcal{H}$.) Suppose $\mathcal{I}$ is a two-sided ideal containing a non-compact operator T. Then, by Proposition 3.2.7 (e), we can find an infinite-dimensional closed subspace $\mathcal{M}$ contained in ran(T). Let $\mathcal{N} = T^{-1}(\mathcal{M}) \cap \ker^\perp(T)$. Then T is a bijective bounded operator of $\mathcal{N}$ onto $\mathcal{M}$ and hence there exists an $S \in B(\mathcal{M},\mathcal{N})$ such that $TS = id_{\mathcal{M}}$. Since $\mathcal{M}$ is infinite-dimensional, there exists an isometry $V \in B(\mathcal{H})$ such that ran$(V) = \mathcal{M}$. It is seen that $id_{\mathcal{H}} = V^*TSV \in \mathcal{I}$, whence the ideal $\mathcal{I}$ must be all of $B(\mathcal{H})$.)

Probably the most useful fact regarding Hilbert-Schmidt operators is their connection with integral operators. (Recall that a measure space $(Z, \mathcal{B}_Z, \lambda)$ is said to be σ-finite if there exists a partition $Z = \coprod_{n=1}^\infty E_n$, such that $E_n \in \mathcal{B}_Z, \mu(E_n) < \infty \, \forall n$. The reason for our restricting ourselves to σ-finite measure spaces is that it is only in the presence of some such hypothesis that Fubini's theorem is valid.)

Proposition 3.3.5. *Let $(X, \mathcal{B}_X, \mu)$ and $(Y, \mathcal{B}_Y, \nu)$ be σ-finite measure spaces. Let $\mathcal{H} = L^2(X, \mathcal{B}_X, \mu)$ and $\mathcal{K} = L^2(Y, \mathcal{B}_Y, \nu)$. Then the following conditions on an operator $T \in B(\mathcal{H},\mathcal{K})$ are equivalent:*

(i) $T \in B^2(\mathcal{K},\mathcal{H})$;

(ii) there exists $k \in L^2(X \times Y, \mathcal{B}_X \otimes \mathcal{B}_Y, \mu \times \nu)$ such that

$$(Tg)(x) \;=\; \int_Y k(x,y)g(y)\; d\nu(y)\;\; \nu - a.e.\; \forall\; g \in \mathcal{K}\;. \tag{3.3.10}$$

If these equivalent conditions are satisfied, then,

$$||T||_{B^2(\mathcal{K},\mathcal{H})} \;=\; ||k||_{L^2(\mu\times\nu)}\;.$$

Proof. $(ii) \Rightarrow (i)$: Suppose $k \in L^2(\mu \times \nu)$; then, by Tonelli's theorem, we can find a set $A \in \mathcal{B}_X$ such that $\mu(A) = 0$ and such that $x \notin A \;\Rightarrow\; k^x\; (= k(x,\cdot)\,) \in L^2(\nu)$, and further,

$$||k||^2_{L^2(\mu\times\nu)} \;=\; \int_{X\setminus A} ||k^x||^2_{L^2(\nu)} d\mu(x)\;.$$

It follows from the Cauchy-Schwarz inequality that if $g \in L^2(\nu)$, then $k^x g \in L^1(\nu)\; \forall\; x \notin A$; hence equation (3.3.10) does indeed meaningfully define a function Tg on $X \setminus A$, so that Tg is defined almost everywhere; another application of the Cauchy-Schwarz inequality shows that

$$\begin{aligned} ||Tg||^2_{L^2(\mu)} \;&=\; \int_X \left|\int_Y k(x,y)g(y)\; d\nu(y)\right|^2 d\mu(x) \\ &=\; \int_{X-A} |\langle k^x, \overline{g}\rangle_{\mathcal{K}}|^2\; d\mu(x) \\ &\leq\; \int_{X-A} ||k^x||^2_{L^2(\nu)} ||g||^2_{L^2(\nu)}\; d\mu(x) \\ &=\; ||k||^2_{L^2(\mu\times\nu)} ||g||^2_{L^2(\nu)}\;, \end{aligned}$$

and we thus find that equation (3.3.10) indeed defines a bounded operator $T \in B(\mathcal{K}, \mathcal{H})$.

Before proceeding further, note that if $g \in \mathcal{K}$ and $f \in \mathcal{H}$ are arbitrary, then, (by Fubini's theorem), we find that

$$\begin{aligned} \langle Tg, f\rangle \;&=\; \int_X (Tg)(x)\overline{f(x)}\, d\mu(x) \\ &=\; \int_X \left(\int_Y k(x,y)g(y)\, d\nu(y) \right) \overline{f(x)} d\mu(x) \\ &=\; \langle k, f \otimes \overline{g}\rangle_{L^2(\mu\times\nu)}\;, \end{aligned} \tag{3.3.11}$$

where we have used the notation $(f \otimes \overline{g})$ to denote the function on $X \times Y$ defined by $(f \otimes \overline{g})(x,y) \;=\; f(x)\overline{g(y)}$.

Suppose now that $\{e_n : n \in N\}$ and $\{g_m : m \in M\}$ are orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$ respectively; then, notice that also $\{\overline{g}_m : m \in M\}$ is an orthonormal basis for $\mathcal{K}$; deduce from equation (3.3.11) above that

$$\begin{aligned} \sum_{m\in M, n\in N} |\langle Tg_m, e_n\rangle_{\mathcal{H}}|^2 \;&=\; \sum_{m\in M, n\in N} |\langle k, e_n \otimes \overline{g}_m\rangle_{L^2(\mu\times\nu)}|^2 \\ &=\; ||k||^2_{L^2(\mu\times\nu)}\;; \end{aligned}$$

thus T is a Hilbert-Schmidt operator with Hilbert-Schmidt norm agreeing with the norm of k as an element of $L^2(\mu \times \nu)$.

$(i) \Rightarrow (ii)$: If $T : \mathcal{K} \to \mathcal{H}$ is a Hilbert-Schmidt operator, then, in particular – see Proposition 3.3.3(c) – T is compact; let

$$Tg \;=\; \sum_n \lambda_n \langle g, g_n \rangle f_n$$

be the singular value decomposition of T (see Proposition 3.2.10). Thus $\{g_n\}$ (resp., $\{f_n\}$) is an orthonormal sequence in $\mathcal{K}$ (resp., $\mathcal{H}$) and $\lambda_n = s_n(T)$. It follows from Proposition 3.3.3 (d) that $\sum_n \lambda_n^2 < \infty$, and hence we find that the equation

$$k \;=\; \sum_n \lambda_n \; f_n \otimes \overline{g}$$

defines a unique element $k \in L^2(\mu \times \nu)$; if $\tilde{T}$ denotes the 'integral operator' associated to the 'kernel function' k as in Equation (3.3.10), we find from equation (3.3.11) that for arbitrary $g \in \mathcal{K}, f \in \mathcal{H}$, we have

$$\begin{aligned}
\langle \tilde{T} g, f \rangle_{\mathcal{H}} &= \langle k, f \otimes \overline{g} \rangle_{L^2(\mu \times \nu)} \\
&= \sum_n \lambda_n \langle f_n \otimes \overline{g}_n, f \otimes \overline{g} \rangle_{L^2(\mu \times \nu)} \\
&= \sum_n \lambda_n \langle f_n, f \rangle_{\mathcal{H}} \langle \overline{g}_n, \overline{g} \rangle_{\mathcal{K}} \\
&= \sum_n \lambda_n \langle f_n, f \rangle_{\mathcal{H}} \langle g, g_n \rangle_{\mathcal{K}} \\
&= \langle Tg, f \rangle_{\mathcal{H}} \; ,
\end{aligned}$$

whence we find that $T = \tilde{T}$ and so T is, indeed, the integral operator induced by the kernel function k. □

EXERCISE 3.3.6. *If T and k are related as in Equation* (3.3.10), *we say that T is the integral operator induced by the kernel k, and we shall write $T = Int\ k$.*

For $i = 1, 2, 3$, let $\mathcal{H}_i = L^2(X_i, \mathcal{B}_i, \mu_i)$, where $(X_i, \mathcal{B}_i, \mu_i)$ is a σ-finite measure space. Let $h \in L^2(X_2 \times X_3, \mathcal{B}_2 \otimes \mathcal{B}_3, \mu_2 \times \mu_3), k, k_1 \in L^2(X_1 \times X_2, \mathcal{B}_1 \otimes \mathcal{B}_2, \mu_1 \times \mu_2)$, and let $S = Int\ h \in B^2(\mathcal{H}_3, \mathcal{H}_2)$, $T = Int\ k, T_1 = Int\ k_1 \in B^2(\mathcal{H}_2, \mathcal{H}_1)$; show that

(i) if $\alpha \in \mathbb{C}$, then $T + \alpha T_1 = Int\ (k + \alpha k_1)$;

(ii) if we define $k^*(x_2, x_1) = \overline{k(x_1, x_2)}$, then $k^* \in L^2(X_2 \times X_1, \mathcal{B}_2 \otimes \mathcal{B}_1, \mu_2 \times \mu_1)$ and $T^* = Int\ k^*$;

(iii) $TS \in B^2(\mathcal{H}_3, \mathcal{H}_1)$ and $TS = Int\ (k * h)$, where

$$(k * h)(x_1, x_3) \;=\; \int_{X_2} k(x_1, x_2) h(x_2, x_3) \; d\mu_2(x_2)$$

for $(\mu_1 \times \mu_3)$-almost all $(x_1, x_3) \in X \times X$.

(Hint: for (ii), note that k^* is a square-integrable kernel, and use equation (3.3.11) *to show that* $\text{Int } k^* = (\text{Int } k)^*$*; for (iii), note that* $|(k*h)(x_1,x_3)| \leq ||k^{x_1}||_{L^2(\mu_2)}||h_{x_3}||_{L^2(\mu_2)}$ *to conclude that* $k*h \in L^2(\mu_1 \times \mu_3)$*; again use Fubini's theorem to justify interchanging the orders of integration in the verification that* $\text{Int } (k*h) = (\text{Int } k)(\text{Int } h)$*.)*

3.3.2 Trace-class operators

PROPOSITION 3.3.7. *1. The following conditions on an operator $T \in B(\mathcal{H})$ are equivalent:*

(a) T is compact and $\sum_{n \in N} s_n(T) < \infty$.

(b) There exist Hilbert-Schmidt operators H_1, H_2 such that $T = H_1 H_2$.

(c) $\sum |\langle Tx_n, y_n\rangle| < \infty$ for any pair $\{x_n\}$ and $\{y_n\}$ of orthonormal sets in $\mathcal{H}$.

The collection of operators satisfying the three equivalent conditions above is denoted by $B^1(\mathcal{H})$.

2. $B^1(\mathcal{H})$ is a self-adjoint two-sided ideal in $B(\mathcal{H})$.

Due to the next property the class $B^1(\mathcal{H})$ is called the **Trace Class** *and its elements are said to be* **trace class operators**.

3. If T is of trace class, the sum $\sum\langle Tx_n, x_n\rangle$ is absolutely convergent for any orthonormal basis $\{x_n\}_n$ of $\mathcal{H}$; this sum is independent of the orthonormal basis $\{x_n\}_n$, is called the trace of T, and is denoted by $\text{Tr}\,(T)$.

4. If $T \in B^1(\mathcal{H}), A \in B(\mathcal{H})$, then $\text{Tr}\,(AT) = \text{Tr}\,(TA)$. *(Note that both sides of this equation make sense in view of (3) above.) And, in particular* $\text{Tr}\,(UTU^*) = \text{Tr}\,(T)$ *for any unitary U.*

Proof. (1) $(a) \Rightarrow (b)$: Let $T = \sum_{n\in N} s_n(T)\bar{x}_n \otimes y_n$ be the SVD of T. Let $H_1 = \sum_{n\in N} s_n(T)^{\frac{1}{2}}(\bar{y}_n \otimes y_n)$ and $H_2 = \sum_{n\in N} s_n(T)^{\frac{1}{2}}(\bar{x}_n \otimes y_n)$. Then the H_i's are Hilbert-Schmidt operators (these defining equations are, in fact, the SVD's of the H_i's, and hence $s_n(H_i) = s_n(T)^{\frac{1}{2}}, i = 1, 2$ and so, $\sum s_n(H_i)^2 = \sum s_n(T) < \infty$) and clearly $T = H_1 H_2$.

$(b) \Rightarrow (c)$: It follows from (two applications of) the Cauchy-Schwarz inequality (once in $\mathcal{H}_2$ and once in ℓ^2) and Proposition 3.3.3 that

$$\begin{aligned} \sum |\langle Tx_n, y_n\rangle| &\leq \sum |\langle H_2 x_n, H_1^* y_n\rangle| \\ &\leq \sum \|H_2 x_n\| \cdot \|H_1^* y_n\| \\ &\leq \|H_2\|_2 \|H_1^*\|_2 \\ &< \infty\ , \end{aligned}$$

as desired.

$(c) \Rightarrow (a)$: We first wish to show that the assumption (c) implies that T is compact. For this, it suffices to prove that $|T|$ is compact, or equivalently that $\mathrm{ran}(1_{[\epsilon,\infty)}(|T|)$ is finite-dimensional, for any $\epsilon > 0$.

Assertion: Let $\mathcal{M}_\epsilon = \mathrm{ran}(1_{[\epsilon,\infty)}(|T|)$. Suppose $\mathcal{M}_\epsilon$ is infinite-dimensional for some $\epsilon > 0$. Then there exist orthonormal sets $\{x_n\}$ in $\mathcal{M}_\epsilon$ and $\{y_n\}$ in $T(\mathcal{M}_\epsilon)$ such that $\langle Tx_n, y_n\rangle = \|Tx_n\| \; \forall n$.

Reason: Pick a unit vector $x_1 \in \mathcal{M}_\epsilon$. Then $\|Tx_1\| = \||T|x_1\| \geq \epsilon$, so $|T|x_1 \neq 0$. Let $z_1 = \frac{1}{\|Tx_1\|}|T|x_1$. Note that $z_1 \in \mathcal{M}_\epsilon$ since $1_E(|T|)$ commutes with $|T|$ for any Borel set $E \subset \mathbb{R}$. Let $V_1 = [\{x_1, z_1\}]$. As V_1 is a finite-dimensional subspace of $\mathcal{M}_\epsilon$, we may find a unit vector $x_2 \in \mathcal{M}_\epsilon \cap V_1^\perp$. As before, let $z_2 = \frac{1}{\|Tx_2\|}|T|x_2$. Under the assumed infinite dimensionality of $\mathcal{M}_\epsilon$, we may keep repeating this process to find an infinite orthonormal set $\{x_n\} \subset \mathcal{M}_\epsilon$ such that the subspaces $V_n = [x_n, z_n]$ are pairwise orthogonal subspaces of $\mathcal{M}_\epsilon$, where $z_n = \frac{1}{\|Tx_n\|}|T|x_n$. Define $y_n = Uz_n$ where $T = U|T|$ is the polar decomposition of T. Since $\{z_n\}$ is an orthonormal set in $\mathcal{M}_\epsilon$ and hence in the initial space of U, it follows that $\{y_n\}$ is an orthonormal set. The construction implies that $\langle Tx_n, y_n\rangle = \langle |T|x_n, U^*y_n\rangle = \langle |T|x_n, z_n\rangle = \||T|x_n\| \geq \epsilon$, so there is no way the infinite series $\sum |\langle Tx_n, y_n\rangle|$ can converge; hence the assumed infinite-dimensionality of $\mathcal{M}_\epsilon$ is untenable.

So $|T|$, and hence T, must be compact.

As T is compact, let $\{x_n\}, \{y_n\}$ be as in Equation (3.2.5). The desired result follows by applying condition (3) to this choice of of $\{x_n\}, \{y_n\}$.

(2) That it is a vector space and is a two-sided ideal follow from 1(c) and 1(b) of this Proposition, and from Proposition 3.3.3.

(3) The first assertion on absolute convergence of the series follows at once from 1(c) of this Proposition. For the other assertions in (3), we consider three cases of increasing levels of generality:

Case (i) $T \geq 0$: In this case, it follows from Proposition 3.3.7 (1) and Proposition 3.3.3 (c) that $T^{\frac{1}{2}} \in B^2(\mathcal{H})$, and the desired conclusions are consequences of Lemma 3.3.1.

Case (ii) $T = T^*$: Observe that $T_+ = 1_{[0,\infty)}(T)T$ and $T_- = -1_{(-\infty,0]}(T)T$ are also trace-class operators, by (2) above. It follows from the already established Case (i) that both $T_\pm$, and consequently also T, have a well-defined trace independent of the orthonormal basis.

Case (iii): T arbitrary: This follows from Case (ii) and the Cartesian decomposition (since part (2) of this proposition shows that $B^1(\mathcal{H})$ is closed under taking real and imaginary parts.

(4) We prove this also in three stages like (3) above.

Case (i) $T \geq 0$: In this case, the SVD decomposition or the spectral theorem will yield decomposition $T = \sum s_m(T)(\bar{x}_m \otimes x_m)$ for some orthonormal basis $\{x_m\}$ for $\ker^{\perp} T$. Extend this orthonormal set $\{x_m\}$ to an orthonormal basis$\{e_n\}$ for $\mathcal{H}$, and note that $e_n \notin \{x_m\} \Rightarrow e_n \in \ker(T) \Rightarrow Te_n = 0$; we then find that

$$\begin{aligned} \operatorname{Tr}(AT) &= \sum_n \langle ATe_n, e_n \rangle \\ &= \sum_m \langle ATx_m, x_m \rangle \\ &= \sum_m s_m(T) \langle Ax_m, x_m \rangle \end{aligned}$$

while also

$$\begin{aligned} \operatorname{Tr}(TA) &= \sum \langle TAe_n, e_n \rangle \\ &= \sum \langle Ae_n, Te_n \rangle \\ &= \sum s_m(T) \langle Ax_m, x_m \rangle \ . \end{aligned}$$

Case(ii) $T = T^*$: Apply the already proved Case (i) for $T_{\pm}$ and use $T = T_+ - T_-$.

Case(iii) T arbitrary: Use the Cartesian decomposition.

□

The next Exercise outlines an alternative proof of part 4 of Proposition 3.3.7.

Exercise 3.3.8. *(1) If U is unitary and $T \in B^1(\mathcal{H})$, deduce from part* 3 *of Proposition* 3.3.7 *that* $\operatorname{Tr}(UTU^*) = \operatorname{Tr}(T)$.

(2) Show that for unitary U and arbitrary $S \in B^1(\mathcal{H})$, we have $\operatorname{Tr}(US) = \operatorname{Tr}(SU)$. *(*Hint: *Put $T = SU$ in part 1 above.)*

(3) Show that for $A \in B(\mathcal{H})$ and arbitrary $S \in B^1(\mathcal{H})$, we have $\operatorname{Tr}(AS) = \operatorname{Tr}(SA)$. *(*Hint: *Use Proposition* 2.8.1 (7) *and part* 2 *of this Exercise.)*

3.3.3 Duality results

In this section, we shall establish the non-commutative analogues of $\ell^1 \cong (c_0)^*$ and $\ell^\infty \cong (\ell^1)^*$.

For $T \in B^1(\mathcal{H})$, define its **trace norm** $\|T\|_1$ by

$$\|T\|_1 = \sum s_n(T) \ . \tag{3.3.12}$$

PROPOSITION 3.3.9. *There exists a linear bijection $T \leftrightarrow \phi_T$ of $\mathbb{B}^1(\mathcal{H})$ onto $B_0(\mathcal{H})^*$ such that $\|T\|_1 = \|\phi_T\|_{B_0(\mathcal{H})^*}$, and hence $B^1(\mathcal{H})$ is a Banach space with respect to $\|\cdot\|_1$ such that $B^1(\mathcal{H}) \cong B_0(\mathcal{H})^*$.*

Similarly, there exists a bijection $A \leftrightarrow \psi(A)$ of $B(\mathcal{H})$ onto $B^1(\mathcal{H})^*$ such that $\|A\|_{B(\mathcal{H})} = \|\psi_A\|_{B^1(\mathcal{H})^*}$, and hence $B(\mathcal{H}) \cong B^1(\mathcal{H})^*$.

Proof. For $T \in B^1(\mathcal{H})$, and $A \in B_0(\mathcal{H})$ (resp., $A \in B(\mathcal{H})$), define $\phi_T(A) = Tr(AT)$ (resp., $\psi_A(T) = Tr(AT)$).

Suppose $T = U|T|$ is the polar decomposition of T and $T = \sum s_m(\bar{x}_m \otimes y_m)$ is an SVD of T. Let $\{e_n\}$ be a completion of $\{x_n\}$ to an orthonormal basis of $\mathcal{H}$. Then,

$$\begin{aligned} |\operatorname{Tr}(TA)| &= |\operatorname{Tr}(AT)| \\ &= |\sum \langle ATe_n, e_n\rangle| \\ &\leq \sum s_m |\langle Ay_m, x_m\rangle| \\ &\leq \|T\|_1 \|A\|_\infty . \end{aligned}$$

Conclude that (i) $\phi_T \in (B_0(\mathcal{H}))^*$ and $\|\phi_T\| \leq \|T\|_1$ and (ii) $\psi_A \in \left(B^1(\mathcal{H})\right)^*$ and $\|\psi_A\| \leq \|A\|_\infty$.

For T as above, define $W_n = \sum_{k=1}^n \bar{y}_k \otimes x_k$. Then W is a partial isometry with finite-dimensional range, so W_n is compact and $\|W_n\| \leq 1$. Clearly

$$\|\phi_T\| \geq \sup_n |\phi_T(W_n)| = \sup_n |Tr(W_nT))| = \sup_n \sum_{k=1}^n s_k = \|T\|_1$$

so indeed $\|\phi_T\| = \|T\|_1$. And if $A \in B(\mathcal{H})$, then

$$\begin{aligned} \|A\| &= \sup\{|\langle Ax, y\rangle| : x, y \in S(\mathcal{H})\} \\ &= \sup\{|\psi_A(\bar{y} \otimes x))| : x, y \in S(\mathcal{H})\} \\ &\leq \|\psi_A\| , \end{aligned}$$

whence indeed $\|\psi_A\| = \|A\|$.

On the other hand, if $\phi \in (B_0(\mathcal{H}))^*$, (resp., $\psi \in \left(B^1(\mathcal{H})\right)^*$, notice that the equation $B_\phi(x, y) = \phi(\bar{y} \otimes x)$ (resp., $B_\psi(x, y) = \psi(\bar{y} \otimes x)$) defines a sesquilinear form on $\mathcal{H}$ and that $|B_\phi(x, y)| \leq \|\phi\| \, \|x\| \, \|y\|$ (resp., $|B_\psi(x, y)| \leq \|\psi\| \, \|x\| \, \|y\|$) – since $\|(\bar{y} \otimes x)\| = \|(\bar{y} \otimes x)\|_1 = 1$. Deduce the existence of a bounded operator T (resp., A) such that $B_\phi(x, y) = \langle Tx, y\rangle$ (resp., $B_\psi(x, y) = \langle Ax, y\rangle$).

It follows that $\phi(F) = \operatorname{Tr}(TF)$ and $\psi(F) = \operatorname{Tr}(AF)$ whenever $F \in B_{00}(\mathcal{H})$. It follows easily from the SVD that $B_{00}(\mathcal{H})$ is dense in $B^1(\mathcal{H})$, which then shows that $\psi = \psi_A$. To complete the proof, we should show that $T \in \mathbb{B}^1(\mathcal{H})$ and that $\phi = \phi_T$. For this, suppose $\{x_n\}$ and $\{y_n\}$ are a pair of orthonormal sets. Define

$$\alpha_k = \begin{cases} \dfrac{|\langle Tx_k, y_k\rangle|}{\langle Tx_k, y_k\rangle} & \text{if } \langle Tx_k, y_k\rangle \neq 0 \\ 1 & \text{otherwise} \end{cases} .$$

and $F_n = \sum_{k=1}^n \alpha_k(\bar{y_k} \otimes x_k)$. Observe now that for each n, the operator F_n is a partial isometry of finite rank n (and operator norm at most 1) and that

$$\begin{aligned}\sum_{k=1}^n |\langle Tx_k, y_k\rangle| &= \sum_{k=1}^n \alpha_k \langle Tx_k, y_k\rangle \\ &= Tr(TF_n) \\ &= \phi(F_n) \\ &\leq \|\phi\| ,\end{aligned}$$

so $\sum_{k=1}^\infty |\langle Tx_k, y_k\rangle| < \infty$.

In particular, we may deduce from Proposition 3.3.7 (1c) that $T \in B^1(\mathcal{H})$. Since ϕ and ϕ_T agree on the dense subspace $B_{00}(\mathcal{H})$ of $B_0(\mathcal{H})$, we see that $\phi = \phi_T$, as desired. □

Corollary 3.3.10. *If $T \in B^1(\mathcal{H})$, then*

$$\|T\|_1 = \sup\big\{ \textstyle\sum |\langle Tu_n, v_n\rangle| : \{u_n\} \text{ and } \{v_n\} \text{ are orthonormal sets}\big\}.$$

Proof. $\geq$: If $T = \sum s_n(T)\bar{x}_n \otimes y_n$ is a SVD of T, then

$$\begin{aligned}\|T\|_1 &= \sum s_n(T) \\ &= \sum \langle Tx_n, y_n\rangle\end{aligned}$$

so indeed the inequality $\geq$ is valid.

$\leq$: Suppose $\{u_n\}$ and $\{v_n\}$ are orthonormal sets. It is clearly enough to prove the desired inequality under the added assumption that both these orthonormal sets are infinite. Let γ_n be a number of unit modulus such that $|\langle Tu_n, v_n\rangle| = \gamma_n \langle Tu_n, v_n\rangle$. If $\{e_n\}$ is an orthonormal basis for $\mathcal{H}$, define partial isometries U, V by $Ue_n = \gamma_n u_n,\ Ve_n = v_n$. Then observe that

$$\begin{aligned}\sum \langle |Tu_n, v_n\rangle| &= \sum \gamma_n \langle Tu_n, v_n\rangle \\ &= \sum \langle TUe_n, Ve_n\rangle \\ &= \mathrm{Tr}\,(V^*TU) \\ &= \mathrm{Tr}\,(TUV^*) \\ &\leq \|T\|_1,\end{aligned}$$

since $\|UV^*\| \leq 1$. □

Remark 3.3.11. The so-called **von Neumann-Schatten p-class** is the (non-closed) two-sided self-adjoint ideal of $B_0(\mathcal{H})$ defined by $B^p(\mathcal{H}) = \{T \in B_0(\mathcal{H}) : ((s_n(T))) \in \ell^p\}, 1 \leq p < \infty$. (To see that $B^p(\mathcal{H})$ are ideals in $B(\mathcal{H})$, notice that $s_n(UTV) = s_n(T)$ so that $T \in B^p(\mathcal{H}) \Leftrightarrow UTV \in B^p(\mathcal{H})$ whenever U, V are unitary, then appeal to Proposition 2.8.1 (7).) These are Banach spaces w.r.t. $\|T\|_p = \|((s_n(T)))\|_{\ell^p}$, and the expected duality statement $(B^p(\mathcal{H}))^* = B^q(\mathcal{H})$ where $q = \frac{p}{p-1}$ is the conjugate index to p. (In thee previous line, if p had been

1, q should be interpreted as ∞.) These may be proved by using the classical fact $(\ell^p)^* = \ell^q$ and imitating, with obvious modifications, our proof above of $B^1(\mathcal{H}) = (B_0(\mathcal{H}))^*$.

3.4 Fredholm operators

Recall (see Remark 3.3.4 (5)) that $B_0(\mathcal{H})$ is the unique closed two-sided ideal in $B(\mathcal{H})$. This section is devoted to invertibility modulo this ideal.

PROPOSITION 3.4.1. **(Atkinson's theorem)** *If $T \in B(\mathcal{H}_1, \mathcal{H}_2)$, then the following conditions are equivalent:*

(a) there exist operators $S_1, S_2 \in B(\mathcal{H}_2, \mathcal{H}_1)$ and compact operators $K_i \in B(\mathcal{H}_i), i = 1, 2,$ such that

$$S_1 T = 1_{\mathcal{H}_1} + K_1 \quad and \quad TS_2 = 1_{\mathcal{H}_2} + K_2 \ .$$

(b) T satisfies the following conditions:

(i) ran T *is closed; and*

(ii) ker T *and* ker T^* *are both finite-dimensional.*

(c) There exists $S \in B(\mathcal{H}_2, \mathcal{H}_1)$ such that both $\mathrm{id}_{\mathcal{H}_1} - ST$ *and* $\mathrm{id}_{\mathcal{H}_2} - TS$ *are projections with finite-dimensional range.*

Proof. $(a) \Rightarrow (b)$: Begin by fixing a finite-rank operator F such that $||K_1 - F|| < \frac{1}{2}$ (see Proposition 3.2.7(d)); set $\mathcal{M} = \ker F$ and note that if $x \in \mathcal{M}$, then

$$||S_1|| \cdot ||Tx|| \geq ||S_1 Tx|| = ||x + K_1 x|| = ||x + (K_1 - F)x|| \geq \frac{1}{2}||x||,$$

which shows that T is bounded below on $\mathcal{M}$; it follows that $T(\mathcal{M})$ is a closed subspace of $\mathcal{H}_2$; note, however, that $\mathcal{M}^\perp$ is finite-dimensional (since F maps this space injectively onto its finite-dimensional range). It is a fact – see [Sun] Exercise A.6.5 (3) – that the vector sum of a closed subspace and a finite-dimensional subspace (in any Banach space, in fact) is always closed; and hence T satisfies condition (i) thanks to the obvious identity ran $T = T(\mathcal{M}) + T(\mathcal{M}^\perp)$.

As for (ii), since $S_1 T = 1_{\mathcal{H}_1} + K_1$, note that $K_1 x = -x$ for all $x \in \ker T$; this means that ker T is a closed subspace which is contained in ran K_1 and the compactness of K_1 now demands the finite-dimensionality of ker T. Similarly, ker $T^* \subset$ ran K_2^* and condition (ii) is verified.

$(b) \Rightarrow (c)$: Let $\mathcal{N}_1 = \ker T$, $\mathcal{N}_2 = \ker T^*$ $(= \mathrm{ran}^\perp T)$; thus T maps $\mathcal{N}_1^\perp$ 1-1 onto ran T; the condition (b) and the open mapping theorem imply the existence of a bounded operator $S_0 \in B(\mathcal{N}_2^\perp, \mathcal{N}_1^\perp)$ such that S_0 is the inverse of the restricted operator $T|_{\mathcal{N}_1^\perp}$; if we set $S = S_0 P_{\mathcal{N}_2^\perp}$, then $S \in B(\mathcal{H}_2, \mathcal{H}_1)$ and by definition, we have $ST = 1_{\mathcal{H}_1} - P_{\mathcal{N}_1}$ and $TS = 1_{\mathcal{H}_2} - P_{\mathcal{N}_2}$; by condition (ii), both subspaces $\mathcal{N}_i$ are finite-dimensional.

$(c) \Rightarrow (a)$: Obvious. □

Remark 3.4.2. (1) An operator which satisfies the equivalent conditions of Atkinson's theorem is called a **Fredholm operator**, and the collection of Fredholm operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ is denoted by $\mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$, and as usual, we shall write $\mathcal{F}(\mathcal{H}) = \mathcal{F}(\mathcal{H}, \mathcal{H})$. It must be observed – as a consequence of Atkinson's theorem, for instance – that a necessary and sufficient condition for $\mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ to be non-empty is that either (i) $\mathcal{H}_1$ and $\mathcal{H}_2$ are both finite-dimensional, in which case $B(\mathcal{H}_1, \mathcal{H}_2) = \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$, or (ii) both $\mathcal{H}_1$ and $\mathcal{H}_2$ are infinite-dimensional (under the standing assumption of only considering separable Hilbert spaces in the discussion).

(2) The quotient $\mathcal{Q}(\mathcal{H}) = B(\mathcal{H})/B_0(\mathcal{H})$ (of $B(\mathcal{H})$ by the ideal $B_0(\mathcal{H})$) is a Banach algebra, which is called the **Calkin algebra**. If we write $\pi_{B_0} : B(\mathcal{H}) \to \mathcal{Q}(\mathcal{H})$ for the quotient mapping, then we find that an operator $T \in B(\mathcal{H})$ is a Fredholm operator precisely when $\pi_{B_0}(T)$ is invertible in the Calkin algebra; thus, $\mathcal{F}(\mathcal{H}) = \pi_{B_0}^{-1}(\mathcal{G}(\mathcal{Q}(\mathcal{H})))$ – where the symbol $\mathcal{G}(\mathcal{A})$ stands for 'group of invertible elements of the unital algebra $\mathcal{A}$'. (It is a fact, which we shall not need and consequently do not go into here, that the Calkin algebra is a C^*-algebra – as is the quotient of any C^*-algebra by a norm-closed $*$-ideal.)

(3) It is customary to use the adjective 'essential' to describe a property of an operator $T \in B(\mathcal{H})$ which is actually a property of the corresponding element $\pi_{B_0}(T)$ of the Calkin algebra, thus, for instance, the **essential spectrum** of T is defined to be

$$\sigma_{\mathrm{ess}}(T) = \sigma_{\mathcal{Q}(\mathcal{H})}(\pi_{B_0}(T)) = \{\lambda \in \mathbb{C} : (T - \lambda) \notin \mathcal{F}(\mathcal{H})\} . \tag{3.4.13}$$

□

The next exercise is devoted to illustrating the notions of Fredholm operator and essential spectrum at least in the case of normal operators.

Exercise 3.4.3. *(1) Let $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ have polar decomposition $T = U|T|$. Then show that*

(a) $T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2) \Leftrightarrow U \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ and $|T| \in \mathcal{F}(\mathcal{H}_1)$.

(b) A partial isometry is a Fredholm operator if and only if both its initial and final spaces have finite co-dimension (i.e., have finite-dimensional orthogonal complements).

(Hint: for both parts, use the characterisation of a Fredholm operator which is given by Proposition 3.4.1(b).)

(2) If $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$, consider the following conditions on an operator $T \in B(\mathcal{H})$:

(i) T is normal;

(ii) U and $|T|$ commute.

Show that (i) $\Rightarrow$ (ii), and find an example to show that the reverse implication is not valid in general.

(Hint: if T is normal, then note that

$$|T|^2 U = T^*TU = TT^*U = U|T|^2U^*U = U|T|^2 ;$$

thus U commutes with $|T|^2$; deduce that in the decomposition $\mathcal{H} = \ker T \oplus \ker^\perp T$, we have $U = 0 \oplus U_0$, $|T| = 0 \oplus A$, where U_0 (resp., A) is a unitary (resp., positive injective) operator of $\ker^\perp T$ onto (resp., into) itself; and infer that U_0 and A^2 commute; since U_0 is unitary, deduce from the uniqueness of positive square roots that U_0 commutes with A, and finally that U and $|T|$ commute; for the 'reverse implication', let T denote the unilateral shift, and note that $U = T$ and $|T| = 1$.)

(3) Suppose $T = U|T|$ is a normal operator as in (2) above. Then show that the following conditions on T are equivalent:

(i) T is a Fredholm operator;

(ii) there exists an orthogonal direct-sum decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$, where dim $\mathcal{N} < \infty$, with respect to which T has the form $T = T_1 \oplus 0$, where T_1 is an invertible normal operator on $\mathcal{M}$;

(iii) there exists an $\epsilon > 0$ such that $1_{\mathbb{D}_\epsilon}(T) = 1_{\{0\}}(T) = P_0$, where (a) $E \mapsto 1_E(T)$ denotes the measurable functional calculus for T, (b) $\mathbb{D}_\epsilon = \{z \in \mathbb{C} : |z| < \epsilon\}$ is the ϵ-disc around the origin, and (c) P_0 is some finite-rank projection.

(Hint: For (i) $\Rightarrow$ (ii), note, as in the hint for exercise (2) above, that we have decompositions $U = U_0 \oplus 0$, $|T| = A \oplus 0$ – with respect to $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$, with $\mathcal{M} = \ker^\perp T$ and $\mathcal{N} = \ker T$ finite-dimensional under the assumption (i) – where U_0 is unitary, A is 1-1 and positive, and U_0 and A commute; deduce from the Fredholm condition that $\mathcal{N}$ is finite-dimensional and that A is invertible; conclude that in this decomposition, $T = U_0 A \oplus 0$ and $U_0 A$ is normal and invertible. For (ii) $\Rightarrow$ (iii), if $T = T_1 \oplus 0$ has polar decomposition $T = U|T|$, then $|T| = |T_1| \oplus 0$ and $U = U_0 \oplus 0$ with U_0 unitary and $|T_1|$ positive and invertible; then if $\epsilon > 0$ is such that T_1 is bounded below by ϵ, then argue that $1_{\mathbb{D}_\epsilon}(T) = 1_{[0,\epsilon)}(|T|) = 1_{\{0\}}(|T|) = 1_{\{0\}}(T) = P_{\mathcal{N}}$.)

(4) Let $T \in B(\mathcal{H})$ be normal; prove that the following conditions on a complex number λ are equivalent:

(i) $\lambda \notin \sigma_{\mathrm{ess}}(T)$;

(ii) there exists an orthogonal direct-sum decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$, where dim $\mathcal{N} < \infty$, with respect to which T has the form $T = T_1 \oplus \lambda$, where $(T_1 - \lambda)$ is an invertible normal operator on $\mathcal{M}$;

(iii) there exists $\epsilon > 0$ such that $1_{\mathbb{D}_\epsilon + \lambda}(T) = 1_{\{\lambda\}}(T) = P_\lambda$, where $\mathbb{D}_\epsilon + \lambda$ denotes the ϵ-disc around the point λ, and P_λ is some finite-rank projection. (Hint: apply (3) above to $T - \lambda$.)

We now come to an important definition.

DEFINITION 3.4.4. *If $T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ is a Fredholm operator, its* **(Fredholm) index** *is the integer defined by*

$$\operatorname{ind} T = \dim(\ker T) - \dim(\ker T^*).$$

Several elementary properties of the index are discussed in the following remark.

REMARK 3.4.5. (1) The index of a normal Fredholm operator is always 0. (Reason: If $T \in B(\mathcal{H})$ is a normal operator, then $|T|^2 = |T^*|^2$, and the uniqueness of the square root implies that $|T| = |T^*|$; it follows that $\ker T = \ker |T| = \ker T^*$.)

(2) It should be clear from the definitions that if $T = U|T|$ is the polar decomposition of a Fredholm operator, then $\operatorname{ind} T = \operatorname{ind} U$.

(3) If $\mathcal{H}_1$ and $\mathcal{H}_2$ are finite-dimensional, then $B(\mathcal{H}_1, \mathcal{H}_2) = \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ and $\operatorname{ind} T = \dim \mathcal{H}_1 - \dim \mathcal{H}_2 \ \forall \ T \in B(\mathcal{H}_1, \mathcal{H}_2)$; in particular, the index is independent of the operator in this case. (Reason: let us write $\rho = \dim(\text{ran } T)$ (resp., $\rho^* = \dim(\text{ran } T^*)$) and $\nu = \dim(\ker T)$ (resp., $\nu^* = \dim(\ker T^*)$) for the rank and nullity of T (resp., T^*); on the one hand, deduce from Exercise 3.1.6(3) that if $\dim \mathcal{H}_i = n_i$, then $\rho = n_1 - \nu$ and $\rho^* = n_2 - \nu^*$; on the other hand, by Exercise 3.1.6(2), we find that $\rho = \rho^*$; hence,

$$\operatorname{ind} T = \nu - \nu^* = (n_1 - \rho) - (n_2 - \rho) = n_1 - n_2 \ .)$$

(4) If $S = UTV$, where $S \in B(\mathcal{H}_1, \mathcal{H}_4), U \in B(\mathcal{H}_3, \mathcal{H}_4), T \in B(\mathcal{H}_2, \mathcal{H}_3)$, $V \in B(\mathcal{H}_1, \mathcal{H}_2)$, and if U and V are invertible (i.e., are 1-1 and onto), then S is a Fredholm operator if and only if T is, in which case, $\operatorname{ind} S = \operatorname{ind} T$. (This should be clear from Atkinson's theorem and the definition of the index.)

(5) Suppose $\mathcal{H}_i = \mathcal{N}_i \oplus \mathcal{M}_i$ and $\dim \mathcal{N}_i < \infty$, for $i = 1, 2$; suppose $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ is such that T maps $\mathcal{N}_1$ into $\mathcal{N}_2$, and such that T maps $\mathcal{M}_1$ 1-1 onto $\mathcal{M}_2$. Thus, with respect to these decompositions, T has the matrix decomposition

$$T = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} ,$$

where D is invertible; then it follows from Atkinson's theorem that T is a Fredholm operator, and the assumed invertibility of D implies that $\operatorname{ind} T = \operatorname{ind} A = \dim \mathcal{N}_1 - \dim \mathcal{N}_2$, see (3) above. $\square$

LEMMA 3.4.6. *Suppose $\mathcal{H}_i = \mathcal{N}_i \oplus \mathcal{M}_i$, for $i = 1, 2$; suppose $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ has the associated matrix decomposition*

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} ,$$

where $A \in B(\mathcal{N}_1, \mathcal{N}_2), B \in B(\mathcal{M}_1, \mathcal{N}_2), C \in B(\mathcal{N}_1, \mathcal{M}_2)$, and $D \in B(\mathcal{M}_1, \mathcal{M}_2)$; assume that D is invertible – i.e, D maps $\mathcal{M}_1$ 1-1 onto $\mathcal{M}_2$. Then

$$T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2) \ \Leftrightarrow \ (A - BD^{-1}C) \in \mathcal{F}(\mathcal{N}_1, \mathcal{N}_2) \ ,$$

and $\operatorname{ind} T = \operatorname{ind} (A - BD^{-1}C)$*; in particular, if* $\dim \ \mathcal{N}_i < \infty, \ i = 1, 2$*, then T is necessarily a Fredholm operator and* $\operatorname{ind} T = \dim \mathcal{N}_1 - \dim \mathcal{N}_2$.

Proof. Let $U \in B(\mathcal{H}_2)$ (resp., $V \in B(\mathcal{H}_1)$) be the operator which has the matrix decomposition

$$U = \left[\begin{array}{cc} 1_{\mathcal{N}_2} & -BD^{-1} \\ 0 & 1_{\mathcal{M}_2} \end{array} \right], \quad (\text{resp., } V = \left[\begin{array}{cc} 1_{\mathcal{N}_1} & 0 \\ -D^{-1}C & 1_{\mathcal{M}_1} \end{array} \right])$$

with respect to $\mathcal{H}_2 = \mathcal{N}_2 \oplus \mathcal{M}_2$ (resp., $\mathcal{H}_1 = \mathcal{N}_1 \oplus \mathcal{M}_1$).

Note that U and V are invertible operators, and that

$$UTV = \left[\begin{array}{cc} A - BD^{-1}C & 0 \\ 0 & D \end{array} \right] ;$$

since D is invertible, we see that $\ker(UTV) = \ker(A - BD^{-1}C)$ and that $\ker(UTV)^* = \ker(A - BD^{-1}C)^*$; also, it should be clear that UTV has closed range if and only if $(A - BD^{-1}C)$ has closed range; we thus see that T is a Fredholm operator precisely when $(A - BD^{-1}C)$ is Fredholm, and that $\operatorname{ind} T = \operatorname{ind}(A - BD^{-1}C)$ in that case. For the final assertion of the lemma (concerning finite-dimensional $\mathcal{N}_i$'s), appeal now to Remark 3.4.5(5). □

We now state some simple facts in an exercise, before proceeding to establish the main facts concerning the index of Fredholm operators.

EXERCISE 3.4.7. *(1) Suppose $D_0 \in B(\mathcal{H}_1, \mathcal{H}_2)$ is an invertible operator; show that there exists $\epsilon > 0$ such that if $D \in B(\mathcal{H}_1, \mathcal{H}_2)$ satisfies $||D - D_0|| < \epsilon$, then D is invertible. (Hint: let $D_0 = U_0|D_0|$ be the polar decomposition; write $D = U_0(U_0^* D)$, note that $||D - D_0|| = ||(U_0^* D - |D_0|)||$, and that D is invertible if and only if $U_0^* D$ is invertible, and use the fact that the set of invertible elements in any Banach algebra ($B(\mathcal{H}_1)$ in this case) form an open set.)*

(2) Show that a function $\phi : [0, 1] \to \mathbb{Z}$ which is locally constant, is necessarily globally constant.

(3) Suppose $\mathcal{H}_i = \mathcal{N}_i \oplus \mathcal{M}_i, i = 1, 2,$ are orthogonal direct sum decompositions of Hilbert spaces.

(a) Suppose $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ is represented by the operator matrix

$$T = \left[\begin{array}{cc} A & 0 \\ C & D \end{array} \right],$$

where A and D are invertible operators; show, then, that T is also invertible and that T^{-1} is represented by the operator matrix

$$T^{-1} = \left[\begin{array}{cc} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{array} \right].$$

(b) Suppose $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ is represented by the operator matrix

$$T = \left[\begin{array}{cc} 0 & B \\ C & 0 \end{array} \right],$$

where B is an invertible operator; show that $T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ if and only if $C \in \mathcal{F}(\mathcal{N}_1, \mathcal{M}_2)$, and that if this happens, then $\text{ind}\, T = \text{ind}\, C$.

THEOREM 3.4.8. *(a) $\mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ is an open set in $B(\mathcal{H}_1, \mathcal{H}_2)$ and the function* $\text{ind} : \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2) \to \mathbb{C}$ *is 'locally constant'; i.e., if $T_0 \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$, then there exists $\delta > 0$ such that whenever $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ satisfies $||T - T_0|| < \delta$, it is then the case that $T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ and* $\text{ind}\, T = \text{ind}\, T_0$.

(b) $T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2), K \in B_0(\mathcal{H}_1, \mathcal{H}_2) \Rightarrow (T+K) \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$ and $\text{ind}\, (T+K) = \text{ind}\, T$.

(c) $S \in \mathcal{F}(\mathcal{H}_2, \mathcal{H}_3), T \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2) \Rightarrow ST \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_3)$ and $\text{ind}\, (ST) = \text{ind}\, S + \text{ind}\, T$.

Proof. (a) Suppose $T_0 \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_2)$. Set $\mathcal{N}_1 = \ker T_0$ and $\mathcal{N}_2 = \ker T_0^*$, so that $\mathcal{N}_i, i = 1, 2$, are finite-dimensional spaces and we have the orthogonal decompositions $\mathcal{H}_i = \mathcal{N}_i \oplus \mathcal{M}_i, i = 1, 2$, where $\mathcal{M}_1 = \text{ran}\ T_0^*$ and $\mathcal{M}_2 = \text{ran}\ T_0$. With respect to these decompositions of $\mathcal{H}_1$ and $\mathcal{H}_2$, it is clear that the matrix of T_0 has the form

$$T_0 = \begin{bmatrix} 0 & 0 \\ 0 & D_0 \end{bmatrix},$$

where the operator $D_0 : \mathcal{M}_1 \to \mathcal{M}_2$ is (a bounded bijection, and hence) invertible.

Since D_0 is invertible, it follows – see Exercise 3.4.7(1) – that there exists a $\delta > 0$ such that $D \in B(\mathcal{M}_1, \mathcal{M}_2),\ ||D - D_0|| < \delta \ \Rightarrow\ D$ is invertible. Suppose now that $T \in B(\mathcal{H}_1, \mathcal{H}_2)$ and $||T - T_0|| < \delta$; let

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

be the matrix decomposition associated to T; then note that $||D - D_0|| < \delta$ and consequently D is an invertible operator. Conclude from Lemma 3.4.6 that T is a Fredholm operator and that

$$\text{ind}\, T \ = \text{ind}\, (A - BD^{-1}C) = \dim \mathcal{N}_1 - \dim \mathcal{N}_2 = \text{ind}\, T_0 \ .$$

(b) If T is a Fredholm operator and K is compact, as in (b), define $T_t = T + tK$, for $0 \leq t \leq 1$. It follows from Proposition 3.4.1 that each T_t is a Fredholm operator; further, it is a consequence of (a) above that the function $[0, 1] \ni t \mapsto$ *ind* T_t is a locally constant function on the interval $[0, 1]$; the desired conclusion follows easily – see Exercise 3.4.7(2).

(c) Let us write $\mathcal{K}_1 = \mathcal{H}_1 \oplus \mathcal{H}_2$ and $\mathcal{K}_2 = \mathcal{H}_2 \oplus \mathcal{H}_3$, and consider the operators $U \in B(\mathcal{K}_2), R \in B(\mathcal{K}_1, \mathcal{K}_2)$ and $V \in B(\mathcal{K}_1)$ defined, by their matrices with respect to the afore-mentioned direct-sum decompositions of these spaces, as follows:

$$U = \begin{bmatrix} 1_{\mathcal{H}_2} & 0 \\ -\epsilon^{-1}S & 1_{\mathcal{H}_3} \end{bmatrix}, \ R = \begin{bmatrix} T & \epsilon 1_{\mathcal{H}_2} \\ 0 & S \end{bmatrix},$$

$$V = \begin{bmatrix} -\epsilon 1_{\mathcal{H}_1} & 0 \\ T & \epsilon^{-1} 1_{\mathcal{H}_2} \end{bmatrix},$$

where we first choose $\epsilon > 0$ to be so small as to ensure that R is a Fredholm operator with index equal to $ind\ T + ind\ S$; this is possible by (a) above, since the operator R_0, which is defined by modifying the definition of R so that the 'off-diagonal' terms are zero and the diagonal terms are unaffected, is clearly a Fredholm operator with index equal to the sum of the indices of S and T.

It is easy to see that U and V are invertible operators – see Exercise 3.4.7(3)(a) – and that the matrix decomposition of the product $URV \in \mathcal{F}(\mathcal{K}_1, \mathcal{K}_2)$ is given by:

$$URV = \begin{bmatrix} 0 & 1_{\mathcal{H}_2} \\ ST & 0 \end{bmatrix},$$

which is seen – see Exercise 3.4.7(3)(b) – to imply that $ST \in \mathcal{F}(\mathcal{H}_1, \mathcal{H}_3)$ and that $\mathrm{ind}\ (ST) = \mathrm{ind}\ R = \mathrm{ind}\ S + \mathrm{ind}\ T$, as desired. □

EXAMPLE 3.4.9. Fix a separable infinite-dimensional Hilbert space $\mathcal{H}$; for the sake of definiteness, we assume that $\mathcal{H} = \ell^2$. Let $S \in B(\mathcal{H})$ denote the unilateral shift – see Example 1.7.9(1). Then, S is a Fredholm operator with $\mathrm{ind}\ S = -1$, and $\mathrm{ind}\ S^* = 1$; hence Theorem 3.4.8 (c) implies that if $n \in \mathbb{N}$, then $S^n \in \mathcal{F}(\mathcal{H})$ and $\mathrm{ind}\ (S^n) = -n$ and $\mathrm{ind}\ (S^*)^n = n$; in particular, there exist operators with all possible indices.

Let us write $\mathcal{F}_n = \{T \in \mathcal{F}(\mathcal{H}) : \mathrm{ind}\ T = n\}$, for each $n \in \mathbb{Z}$.

First consider the case $n = 0$. Suppose $T \in \mathcal{F}_0$; then it is possible to find a partial isometry U_0 with initial space equal to $\ker T$ and final space equal to $\ker T^*$); then define $T_t = T + tU_0$. Observe that $t \neq 0 \Rightarrow T_t$ is invertible; and hence, the map $[0,1] \ni t \mapsto T_t \in B(\mathcal{H})$ (which is clearly norm-continuous) is seen to define a path – see Exercise 3.4.10(1) – which is contained in $\mathcal{F}_0$ and connects T_0 to an invertible operator; on the other hand, the set of invertible operators is a path-connected subset of $\mathcal{F}_0$; it follows that $\mathcal{F}_0$ is path-connected.

Next consider the case $n > 0$. Suppose $T \in \mathcal{F}_n, n < 0$. Then note that $T(S^*)^n \in \mathcal{F}_0$ (by Theorem 3.4.8(c)) and since $(S^*)^n S^n = 1$, we find that $T = T(S^*)^n S^n \in \mathcal{F}_0 S^n$; conversely since Theorem 3.4.8(c) implies that $\mathcal{F}_0 S^n \subset \mathcal{F}_n$, we thus find that $\mathcal{F}_n = \mathcal{F}_0 S^n$.

For $n > 0$, we find, by taking adjoints, that $\mathcal{F}_n = \mathcal{F}^*_{-n} = (S^*)^n \mathcal{F}_0$.

We conclude that for all $n \in \mathbb{Z}$, the set $\mathcal{F}_n$ is path-connected; on the other hand, since the index is 'locally constant', we can conclude that $\{\mathcal{F}_n : n \in \mathbb{Z}\}$ is precisely the collection of 'path-components' (= maximal path-connected subsets) of $\mathcal{F}(\mathcal{H})$. □

EXERCISE 3.4.10. (1) *A* **path** *in a topological space X is a continuous function $f : [0,1] \to X$; if $f(0) = x, f(1) = y$, then f is called a path joining (or connecting) x to y. Define a relation $\sim$ on X by stipulating that $x \sim y$ if and only if there exists a path joining x to y.*

Show that $\sim$ is an equivalence relation on X.

The equivalence classes associated to the relation $\sim$ are called the **path-components** *of X; the space X is said to be* **path-connected** *if X is itself a path component.*

(2) *Let $\mathcal{H}$ be a separable Hilbert space. In this exercise, we regard $B(\mathcal{H})$ as being topologised by the operator norm.*

(a) Show that the set $B_{sa}(\mathcal{H})$ of self-adjoint operators on $\mathcal{H}$ is path-connected. (Hint: Consider $t \mapsto tT$.)

(b) Show that the set $B_+(\mathcal{H})$ of positive operators on $\mathcal{H}$ is path-connected. (Hint: Note that if $T \geq 0, t \in [0,1]$, then $tT \geq 0$.)

(c) Show that the set $GL_+(\mathcal{H})$ of invertible positive operators on $\mathcal{H}$ form a connected set. (Hint: If $T \in GL_+(\mathcal{H})$, use straight line segments to first connect T to $||T|| \cdot 1$, and then $||T|| \cdot 1$ to 1.)

(d) Show that the set $\mathcal{U}(\mathcal{H})$ of unitary operators on $\mathcal{H}$ is path-connected. (Hint: If $U \in \mathcal{U}(\mathcal{H})$, find a self-adjoint A such that $U = e^{iA}$ – see Proposition 2.8.1 (4) *– and look at $U_t = e^{itA}$.)*

We would like to conclude this section with the so-called 'spectral theorem for a general n'. As a preamble, we start with an exercise which is devoted to 'algebraic (possibly non-orthogonal) direct sums' and associated non-self-adjoint projections.

EXERCISE 3.4.11. (1) *Let $\mathcal{H}$ be a Hilbert space, and let $\mathcal{M}$ and $\mathcal{N}$ denote closed subspaces of $\mathcal{H}$. Show that the following conditions are equivalent:*

(a) $\mathcal{H} = \mathcal{M} + \mathcal{N}$ and $\mathcal{M} \cap \mathcal{N} = \{0\}$;

(b) every vector $z \in \mathcal{H}$ is uniquely expressible in the form $z = x + y$ with $x \in \mathcal{M}, y \in \mathcal{N}$.

(2) If the equivalent conditions of (1) above are satisfied, show that there exists a unique $E \in B(\mathcal{H})$ such that $Ez = x$, whenever z and x are as in (b) above. (Hint: note that $z = Ez + (z - Ez)$ and use the closed graph theorem to establish the boundedness of E.)

(3) If E is as in (2) above, then show that

(a) $E = E^2$;

(b) the following conditions on a vector $x \in \mathcal{H}$ are equivalent:

(i) $x \in$ ran E;

(ii) $Ex = x$.

(c) ker $E = \mathcal{N}$.

The operator E is said to be the 'projection on $\mathcal{M}$ along $\mathcal{N}$'.

(4) Show that the following conditions on an operator $E \in B(\mathcal{H})$ are equivalent:

(i) $E = E^2$;

(ii) there exists a closed subspace $\mathcal{M} \subset \mathcal{H}$ such that E has an operator-matrix (with respect to the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$) of the form:

$$E = \left[\begin{array}{cc} 1_{\mathcal{M}} & B \\ 0 & 0 \end{array} \right] ;$$

(iii) there exists a closed subspace $\mathcal{N} \subset \mathcal{H}$ *such that* E *has an operator-matrix (with respect to the decomposition* $\mathcal{H} = \mathcal{N}^\perp \oplus \mathcal{N}$*) of the form:*

$$E = \begin{bmatrix} 1_{\mathcal{N}^\perp} & 0 \\ C & 0 \end{bmatrix} ;$$

(iv) there exist closed subspaces $\mathcal{M}, \mathcal{N}$ *satisfying the equivalent conditions of (1) such that* E *is the projection on* $\mathcal{M}$ *along* $\mathcal{N}$.
(Hint: $(i) \Rightarrow (ii)$: $\mathcal{M} = \text{ran } E \ (= \ker(1-E))$ *is a closed subspace and* $Ex = x \ \forall \ x \in \mathcal{M}$*; since* $\mathcal{M} = \text{ran } E$, *(ii) follows. The implication* $(ii) \Rightarrow (i)$ *is verified by easy matrix-multiplication. Finally, if we let* $(i)^*$ *(resp.,* $(ii)^*$*) denote the condition obtained by replacing* E *by* E^* *in condition* (i) *(resp.,* (ii)*), then* $(i) \Leftrightarrow (i)^* \Leftrightarrow (ii)^*$*; take adjoints to find that* $(ii)^* \Leftrightarrow (iii)$*. The implication* $(i) \Leftrightarrow (iv)$ *is clear.)*

(5) Show that the following conditions on an idempotent operator $E \in B(\mathcal{H})$*, i.e.,* $E^2 = E$*, are equivalent:*

(i) $E = E^*$*;*

(ii) $||E|| = 1$.

(Hint: Assume E *is represented in matrix form, as in (4)(iii) above; notice that* $x \in \mathcal{N}^\perp \Rightarrow ||Ex||^2 = ||x||^2 + ||Cx||^2$*; conclude that* $||E|| = 1 \ \Leftrightarrow \ C = 0$*.)*

(6) If E *is the projection onto* $\mathcal{M}$ *along* $\mathcal{N}$ *– as above – show that there exists an invertible operator* $S \in B(\mathcal{H})$ *such that* $SES^{-1} = P_{\mathcal{M}}$*. (Hint: Assume* E *and* B *are related as in (4)(ii) above; define*

$$S = \begin{bmatrix} 1_{\mathcal{M}} & B \\ 0 & 1_{\mathcal{M}^\perp} \end{bmatrix} ;$$

deduce from (a transposed version of) Exercise 3.4.7 *that* S *is invertible, and that*

$$\begin{aligned} SES^{-1} &= \begin{bmatrix} 1_{\mathcal{M}} & B \\ 0 & 1_{\mathcal{M}^\perp} \end{bmatrix} \begin{bmatrix} 1_{\mathcal{M}} & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1_{\mathcal{M}} & -B \\ 0 & 1_{\mathcal{M}^\perp} \end{bmatrix} \\ &= \begin{bmatrix} 1_{\mathcal{M}} & 0 \\ 0 & 0 \end{bmatrix} . \end{aligned}$$

(7) Show that the following conditions on an operator $T \in B(\mathcal{H})$ *are equivalent:*

(a) there exists closed subspaces $\mathcal{M}, \mathcal{N}$ *as in (1) above such that*

(i) $T(\mathcal{M}) \subset \mathcal{M}$ *and* $T|_{\mathcal{M}} = A$*; and*

(ii) $T(\mathcal{N}) \subset \mathcal{N}$ *and* $T|_{\mathcal{N}} = B$*;*

(b) there exists an invertible operator $S \in B(\mathcal{H}, \mathcal{M} \oplus \mathcal{N})$ *– where the direct sum considered is an 'external direct sum' – such that* $STS^{-1} = A \oplus B$.

We will find the following bit of terminology convenient. Call operators $T_i \in B(\mathcal{H}_i), i = 1, 2$, **similar** if there exists an invertible operator $S \in B(\mathcal{H}_1, \mathcal{H}_2)$ such that $T_2 = ST_1S^{-1}$.

LEMMA 3.4.12. *The following conditions on an operator $T \in B(\mathcal{H})$ are equivalent:*

(a) T is similar to an operator of the form $T_0 \oplus Q \in B(\mathcal{M} \oplus \mathcal{N})$, where
(i) $\mathcal{N}$ is finite-dimensional;
(ii) T_0 is invertible, and Q is nilpotent.

(b) $T \in \mathcal{F}(\mathcal{H}), \text{ind}\ (T) = 0$ and there exists a positive integer n such that $\ker\ T^n = \ker\ T^m\ \forall\ m \geq n$.

Proof. $(a) \Rightarrow (b)$: If $STS^{-1} = T_0 \oplus Q$, then it is obvious that $ST^nS^{-1} = T_0^n \oplus Q^n$, which implies – because of the assumed invertibility of T_0 – that $\ker T^n = S^{-1}(\{0\} \oplus \ker Q^n)$, and hence, if $n = \dim \mathcal{N}$, then for any $m \geq n$, we see that $\ker T^m = S^{-1}(\{0\} \oplus \mathcal{N})$.

In particular, $\ker T$ is finite-dimensional; similarly $\ker\ T^*$ is also finite-dimensional, since $(S^*)^{-1}T^*S^* = T_0^* \oplus Q^*$; further,

$$\text{ran}\ T = S^{-1}(\text{ran}\ (T_0 \oplus Q)) = S^{-1}(\mathcal{M} \oplus (\text{ran}\ Q))\ ,$$

which is closed since S^{-1} is a homeomorphism, and since the sum of the closed subspace $\mathcal{M} \oplus \{0\}$ and the finite-dimensional space $(\{0\} \oplus\ \text{ran}\ Q)$ is closed in $\mathcal{M} \oplus \mathcal{N}$ – see the remark at the end of the first paragraph of the proof of $(a) \Rightarrow (b)$ of Atkinson's theorem. Hence T is a Fredholm operator.

Finally,

$$\text{ind}\ (T) = \text{ind}\ (STS^{-1}) = \text{ind}\,(T_0 \oplus Q) = \text{ind}\ (Q) = 0.$$

$(b) \Rightarrow (a)$: Let us write $\mathcal{M}_k = \text{ran}i\ T^k$ and $\mathcal{N}_k = \ker\ T^k$ for all $k \in \mathbb{N}$; then, clearly,

$$\mathcal{N}_1 \subset \mathcal{N}_2 \subset \cdots\ ;\ \mathcal{M}_1 \supset \mathcal{M}_2 \supset \cdots\ .$$

We are told that $\mathcal{N}_n = \mathcal{N}_m\ \forall\ m \geq n$. The assumption *ind* $T = 0$ implies that $\text{ind}\ T^m = 0\ \forall\ m$, and hence, we find that $\dim(\ker\ T^{*m}) = \dim(\ker\ T^m) = \dim(\ker\ T^n) = \dim(\ker\ T^{*n}) < \infty$ for all $m \geq n$. But since $\ker\ T^{*m} = \mathcal{M}_m^\perp$, we find that $\mathcal{M}_m^\perp \subset \mathcal{M}_n^\perp$, from which we may conclude that $\mathcal{M}_m = \mathcal{M}_n\ \forall\ m \geq n$.

Let $\mathcal{N} = \mathcal{N}_n,\ \mathcal{M} = \mathcal{M}_n$, so that we have

$$\mathcal{N} = \ker T^m \text{ and } \mathcal{M} = \text{ran}\, T^m\ \ \forall\ m \geq n\ . \tag{3.4.14}$$

The definitions clearly imply that $T(\mathcal{M}) \subset \mathcal{M}$ and $T(\mathcal{N}) \subset \mathcal{N}$ (since $\mathcal{M}$ and $\mathcal{N}$ are actually invariant under any operator which commutes with T^n).

We assert that $\mathcal{M}$ and $\mathcal{N}$ yield an algebraic direct sum decomposition of $\mathcal{H}$ (in the sense of Exercise 3.4.11(1)). Firstly, if $z \in \mathcal{H}$, then $T^nz \in \mathcal{M}_n = \mathcal{M}_{2n}$, and hence we can find $v \in \mathcal{H}$ such that $T^nz = T^{2n}v$; thus $z - T^nv \in \ker\ T^n$; i.e., if $x = T^nv$ and $y = z - x$, then $x \in \mathcal{M}, y \in \mathcal{N}$ and $z = x + y$; thus, indeed $\mathcal{H} = \mathcal{M} + \mathcal{N}$. Notice that T (and hence also T^n) maps $\mathcal{M}$ onto itself; in particular, if $z \in \mathcal{M} \cap \mathcal{N}$, we can find an $x \in \mathcal{M}$ such that $z = T^nx$; the assumption $z \in \mathcal{N}$ implies that $0 = T^nz = T^{2n}x$; this means that $x \in \mathcal{N}_{2n} = \mathcal{N}_n$, whence

$z = T^n x = 0$; since z was arbitrary, we have shown that $\mathcal{N} \cap \mathcal{M} = \{0\}$, and our assertion has been substantiated.

If $T_0 = T|_{\mathcal{M}}$ and $Q = T|_{\mathcal{N}}$, the (already proved) fact that $\mathcal{M} \cap \mathcal{N} = \{0\}$ implies that T^n is 1-1 on $\mathcal{M}$; thus T_0^n is 1-1; hence T_0 is 1-1; it has already been noted that T_0 maps $\mathcal{M}$ onto $\mathcal{M}$; hence T_0 is indeed invertible; on the other hand, it is obvious that Q^n is the zero operator on $\mathcal{N}$. □

Corollary 3.4.13. *Let $K \in B_0(\mathcal{H})$; assume $0 \neq \lambda \in \sigma(K)$; then K is similar to an operator of the form $K_1 \oplus A \in B(\mathcal{M} \oplus \mathcal{N})$, where*

(a) $K_1 \in B_0(\mathcal{M})$ and $\lambda \notin \sigma(K_1)$; and

(b) $\mathcal{N}$ is a finite-dimensional space, and $\sigma(A) = \{\lambda\}$.

Proof. Put $T = K - \lambda$; then, the hypothesis and Theorem 3.4.8 ensure that T is a Fredholm operator with $\operatorname{ind}(T) = 0$. Consider the non-decreasing sequence

$$\ker T \subset \ker T^2 \subset \cdots \subset \ker T^n \subset \cdots . \tag{3.4.15}$$

Suppose $\ker T^n \neq \ker T^{n+1}\ \forall\ n$; then we can pick a unit vector $x_n \in (\ker T^{n+1}) \cap (\ker T^n)^\perp$ for each n. Clearly the sequence $\{x_n\}_{n=1}^\infty$ is an orthonormal set. Hence, $\lim_n ||Kx_n|| = 0$ (by Exercise 3.2.4(3)).

On the other hand,

$$\begin{aligned} x_n \in \ker T^{n+1} &\Rightarrow\ Tx_n \in \ker T^n \\ &\Rightarrow\ \langle Tx_n, x_n \rangle = 0 \\ &\Rightarrow\ \langle Kx_n, x_n \rangle = \lambda \end{aligned}$$

contradicting the hypothesis that $\lambda \neq 0$ and the already drawn conclusion that $Kx_n \to 0$.

Hence, it must be the case that $\ker T^n = \ker T^{n+1}$ for some $n \in \mathbb{N}$; it follows easily from this that $\ker T^n = \ker T^m\ \forall\ m \geq n$.

Thus, we may conclude from Lemma 3.4.12 that there exists an invertible operator $S \in B(\mathcal{H}, \mathcal{M} \oplus \mathcal{N})$ – where $\mathcal{N}$ is finite-dimensional – such that $STS^{-1} = T_0 \oplus Q$, where T_0 is invertible and $\sigma(Q) = \{0\}$; since $K = T + \lambda$, conclude that $SKS^{-1} = (T_0 + \lambda) \oplus (Q + \lambda)$; set $K_1 = T_0 + \lambda$, $A = Q + \lambda$, and conclude that indeed K_1 is compact, $\lambda \notin \sigma(K_1)$ and $\sigma(A) = \{\lambda\}$.

□

We are finally ready to state the spectral theorem for a compact operator.

Theorem 3.4.14. *Let $K \in B_0(\mathcal{H})$ be a compact operator on a Hilbert space $\mathcal{H}$. Then,*

(a) $\lambda \in \sigma(K) \setminus \{0\} \ \Rightarrow\ \lambda$ is an eigenvalue of K and λ is 'isolated' in the sense that there exists $\epsilon > 0$ such that $0 < |z - \lambda| < \epsilon \ \Rightarrow\ z \notin \sigma(K)$;

(b) if $\lambda \in \sigma(K) \setminus \{0\}$, then λ is an eigenvalue with 'finite algebraic multiplicity' in the strong sense described by Corollary 3.4.13;

(c) $\sigma(K)$ is a countable set, and the only possible accumulation point of $\sigma(K)$ is 0.

Proof. Assertions (a) and (b) are immediate consequences of Corollary 3.4.13, while (c) follows immediately from (a). □

Appendix

Some measure theory

We briefly recall here the two non-trivial theorems from measure theory that we used in this book. They are the **Riesz representation theorem** and **Lusin's theoem**. We shall only consider compactly supported probability measures defined in $\mathcal{B}_{\mathbb{C}}$ here.

The former identifies positivity-preserving linear functionals on $C(\Sigma)$. with Σ a compact Hausdorff space, as being given by integration against positive regular measures defined on the Borel σ-algebra $\mathcal{B}_\Sigma$. Recall that a finite positive measure μ defined on $\mathcal{B}_{\mathbb{C}}$ is said to be regular if it is both **inner** and **outer** regular in the sense that for any $E \in \mathcal{B}_\Sigma$ and any $\epsilon > 0$, there exists a compact set K and and open set U such that $K \subset E \subset U$ and $\mu(U \setminus K) < \epsilon$. We spell out a consequence of this regularity below.

LEMMA A1. *$C(\Sigma)$ is dense in $L^p(\Sigma, \mu)$ for each $p \in [1, \infty)$.*[5]

Proof. Since simple functions are dense in L^p, it is enough to show that functions of the form $1_E, E \in \mathcal{B}_{\mathbb{C}}$ are in the L^p-closure of $C(\mathbb{C})$. If $E \in \mathcal{B}_{\mathbb{C}}$ and $\epsilon > 0$, pick a compact K and open U as in the paragraph preceding the lemma. Next invoke Urysohn's lemma to find an $f \in C(\mathbb{C})$ such that $1_K \leq f \leq 1_U$. Observe that $\{z \in \mathbb{C} : f(z) = 1_E(z)\} \supset K \cup (\mathbb{C} \setminus U)$ and hence $\{z \in \mathbb{C} : f(z) \neq 1_E(z)\} \subset (U \setminus K)$, and since $0 \leq 1_E, f \leq 1$, we see that

$$\int |f - 1_E|^p d\mu \leq \mu(U \setminus K) < \epsilon ,$$

as desired. □

Lusin's theorem says that if ϕ is any bounded Borel measurable function, and if $\epsilon > 0$ is arbitrary, then there exists an $f \in C(\mathbb{C})$ such that $\mu(\{x \in \mathbb{C} : \phi(x) \neq f(x)\}) < \epsilon$. We shall need the following consequence:

[5]Strictly speaking, $C(\Sigma)$ is not contained in $L^p(\Sigma, \mu)$, but there is a natural mapping - sending an element of $C(\Sigma)$ to the equivalence class (mod μ) that it defines, which satisfies $\|f\|_{L^p(\Sigma,\mu)} \leq \|f\|_{C(\Sigma)}$ - from $C(\Sigma)$ onto a dense subspace of $L^p(\Sigma, \mu)$.

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9

Lemma A2. *For any function $\phi \in L^\infty(\mathbb{C}, \mathcal{B}_\mathbb{C}, \mu)$, there exists a sequence $\{f_n : n \in \mathbb{N}\} \subset C(\mathbb{C})$ such that $\sup\{|f_n(z)| : z \in \mathbb{C}\} \le \|\phi\|_{L^\infty(\mu)}$ for all n and $f_n \to \phi$ in (μ-) measure.*

Proof. By Lusin's theorem, we may, for each $n \in \mathbb{N}$, find an $f_n \in C(\mathbb{C})$ such that $\mu(\{z \in \mathbb{C} : f_n(z) \neq \phi(z)\}) < \frac{1}{n}$. Let $r : \mathbb{C} \to \{z \in \mathbb{C} : |z| \le \|\phi\|_{L^\infty(\mu)}\}$ denote the radial retraction defined by

$$r(z) = \begin{cases} z & \text{if } |z| \le \|\phi\|_{L^\infty(\mu)} \\ \left(\dfrac{\|\phi\|_{L^\infty(\mu)}}{|z|}\right) z & \text{otherwise} \end{cases}$$

and set $g_n = r \circ f_n$. Then notice that $|g_n(z)| \le \|\phi\|_{L^\infty(\mu)}$ for all $z \in \mathbb{C}$ and that $\{|\phi| \le \|\phi\|_{L^\infty(\mu)}\} \cap \{f_n = \phi\} \subset \{g_n = \phi\}$; or equivalently, $\{g_n \neq \phi\} \subset \{|\phi| > \|\phi\|_{L^\infty(\mu)}\} \cup \{f_n \neq \phi\}$ so that $\mu(\{g_n \neq \phi\}) < \frac{1}{n}$ and so indeed the continuous functions $\{g_n : n \in \mathbb{N}\}$ are uniformly bounded by $\|\phi\|_{L^\infty(\mu)}$ and converge in (μ) measure to ϕ. □

Some pedagogical subtleties

I believe the natural stage to discuss the measurable functional calculus is in the language of von Neumann algebras. The more symmetric formulation of the spectral theorem is to say that the continuous (resp., measurable) functional calculus is an isomorphism of $C(\sigma(T))$ (resp., $L^\infty(\sigma(T), \mu)$ for appropriate μ) onto $C^*(T)$ (resp., $W^*(T)$) in the category of C^*-algebras (resp., W^*-algebras). This is what was done in [Sun], but that approach necessitates a digression into C^*-algebras and W^*(= von Neumann) algebras.

But my goal here was to convey the essence of the spectral theorem in purely 'operator-theoreric' terms (not making too many demands of a graduate student just getting introduced to functional analysis). This is made possible thanks to the considerations described in the next paragraphs.

One of many equivalent definitions of a von Neumann algebra is as a unital *-subalgebra of $B(\mathcal{H})$ which is closed in the SOT. In fact, I never even defined the acronym SOT properly. In fact the **strong operator topology** is the smallest topology on $B(\mathcal{H})$ for which $B(\mathcal{H}) \ni T \mapsto Tx \in \mathcal{H}$ is continuous for each $x \in \mathcal{H}$. More formally, the collection of sets of the form $\{T \in B(\mathcal{H}) : \|Tx - T_0x\| < \epsilon\}$, with (T_0, x, ϵ) ranging over $B(\mathcal{H}) \times \mathcal{H} \times (0, \infty)$, yields a sub-base for this topology. If one carried this formal process just a little further, one finds, fairly quickly, various unpleasant pathologies (even when $\mathcal{H}$ is separable, but infinite-dimensional) such as: (i) this topological space does not satisfy 'the first axiom of countability', as a result of which sequential convergence is generally insufficient to describe the possible nastiness that this topological algebra is capable of exhibiting, and one needs to deal with nets or filters instead; (ii) the product mapping $B(\mathcal{H}) \times B(\mathcal{H}) \ni (S, T) \mapsto ST \in B(\mathcal{H})$ is not continuous, contrary to what Lemma 2.2.5 (2) might lead one to expect; and (iii) the adjoint mapping $B(\mathcal{H}) \ni T \mapsto T^* \in B(\mathcal{H})$ is not continuous.

Fortunately, it is possible to not have to deal with the unpleasant features of the SOT that were advertised above, thanks to the useful **Kaplansky Density Theorem** which ensures that if a *-algebra $\mathcal{A}$ is SOT-dense in a von Neumann algebra $\mathcal{M}$, then it is **sequentially dense**; more precisely, the theorem says that if $T \in \mathcal{M}$, then one can find a sequence $\{T_n : n \in \mathbb{N}\} \subset \mathcal{A}$ such that $T_n \stackrel{SOT}{\longrightarrow} T$; and such an approximating sequence can be found so that, in addition, $\|T_n\| \leq \|T\| \; \forall n \in \mathbb{N}$. Our interest here lies naturally in the case $\mathcal{A} = C^*(T)$ and $\mathcal{M} = W^*(T)$, for a normal T.

Bibliography

[Hal] P. R. Halmos, *Introduction to Hilbert Space and Spectral Multiplicity*, Chelsea Publishing House, 1951.

[Hal1] P. R. Halmos, *Measure Theory*, Springer-Verlag, New York, 1974.

[Hal2] P. R. Halmos, *A Hilbert space Problem Book*, Springer-Verlag, New York, 1974.

[Put] C.R. Putnam, On normal operators in Hilbert space, Am. J. Math. 73 (1951) 35-36.

[Fug] B. Fuglede, A commutativity theorem for normal operators, *Proc. of Nat. Acad. Sci.*, 36, 1950, 36-40.

[Sun] V.S. Sunder, *Functional Analysis: Spectral Theory*, TRIM Series, No. 13, Hindustan Book Agency, Delhi, 1997, and Birkhauser Advanced Texts, Basel, 1997.

[AthSun] S.R. Athreya and V.S. Sunder, *Measure and Probability*, Universities Press, Hyderabad, 2008.

[Yos] K. Yosida, *Functional Analysis*, Springer, Berlin, 1965.

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9

Index

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9

Texts and Readings in Mathematics

1. R. B. Bapat: Linear Algebra and Linear Models (Third Edition)
2. Rajendra Bhatia: Fourier Series (Second Edition)
3. C. Musili: Representations of Finite Groups
4. H. Helson: Linear Algebra (Second Edition)
5. D. Sarason: Complex Function Theory (Second Edition)
6. M. G. Nadkarni: Basic Ergodic Theory (Third Edition)
7. H. Helson: Harmonic Analysis (Second Edition)
8. K. Chandrasekharan: A Course on Integration Theory
9. K. Chandrasekharan: A Course on Topological Groups
10. R. Bhatia (ed.): Analysis, Geometry and Probability
11. K. R. Davidson: C*-Algebras by Example (Reprint)
12. M. Bhattacharjee et al.: Notes on Infinite Permutation Groups
13. V. S. Sunder: Functional Analysis – Spectral Theory
14. V. S. Varadarajan: Algebra in Ancient and Modern Times
15. M. G. Nadkarni: Spectral Theory of Dynamical Systems
16. A. Borel: Semisimple Groups and Riemannian Symmetric Spaces
17. M. Marcolli: Seiberg-Witten Gauge Theory
18. A. Bottcher and S. M. Grudsky: Toeplitz Matrices, Asymptotic Linear Algebra and Functional Analysis
19. A. R. Rao and P. Bhimasankaram: Linear Algebra (Second Edition)
20. C. Musili: Algebraic Geometry for Beginners
21. A. R. Rajwade: Convex Polyhedra with Regularity Conditions and Hilbert's Third Problem
22. S. Kumaresan: A Course in Differential Geometry and Lie Groups
23. Stef Tijs: Introduction to Game Theory
24. B. Sury: The Congruence Subgroup Problem
25. R. Bhatia (ed.): Connected at Infinity
26. K. Mukherjea: Differential Calculus in Normed Linear Spaces (Second Edition)
27. Satya Deo: Algebraic Topology: A Primer (Corrected Reprint)
28. S. Kesavan: Nonlinear Functional Analysis: A First Course
29. S. Szabó: Topics in Factorization of Abelian Groups
30. S. Kumaresan and G. Santhanam: An Expedition to Geometry
31. D. Mumford: Lectures on Curves on an Algebraic Surface (Reprint)
32. J. W. Milnor and J. D. Stasheff: Characteristic Classes (Reprint)
33. K. R. Parthasarathy: Introduction to Probability and Measure (Corrected Reprint)
34. Amiya Mukherjee: Topics in Differential Topology
35. K. R. Parthasarathy: Mathematical Foundations of Quantum Mechanics
36. K. B. Athreya and S. N. Lahiri: Measure Theory
37. Terence Tao: Analysis I (Third Edition)

V.S. Sunder, *Operators on Hilbert Space*, Texts and Readings in Mathematics 71,
DOI 10.1007/978-981-10-1816-9

38. Terence Tao: Analysis II (Third Edition)
39. W. Decker and C. Lossen: Computing in Algebraic Geometry
40. A. Goswami and B. V. Rao: A Course in Applied Stochastic Processes
41. K. B. Athreya and S. N. Lahiri: Probability Theory
42. A. R. Rajwade and A. K. Bhandari: Surprises and Counterexamples in Real Function Theory
43. G. H. Golub and C. F. Van Loan: Matrix Computations (Reprint of the Third Edition)
44. Rajendra Bhatia: Positive Definite Matrices
45. K. R. Parthasarathy: Coding Theorems of Classical and Quantum Information Theory (Second Edition)
46. C. S. Seshadri: Introduction to the Theory of Standard Monomials (Second Edition)
47. Alain Connes and Matilde Marcolli: Noncommutative Geometry, Quantum Fields and Motives
48. Vivek S. Borkar: Stochastic Approximation: A Dynamical Systems Viewpoint
49. B. J. Venkatachala: Inequalities: An Approach Through Problems
50. Rajendra Bhatia: Notes on Functional Analysis
51. A. Clebsch (ed.): Jacobi's Lectures on Dynamics (Second Revised Edition)
52. S. Kesavan: Functional Analysis
53. V. Lakshmibai and Justin Brown: Flag Varieties: An Interplay of Geometry, Combinatorics, and Representation Theory
54. S. Ramasubramanian: Lectures on Insurance Models
55. Sebastian M. Cioaba and M. Ram Murty: A First Course in Graph Theory and Combinatorics
56. Bamdad R. Yahaghi: Iranian Mathematics Competitions, 1973-2007
57. Aloke Dey: Incomplete Block Designs
58. R. B. Bapat: Graphs and Matrices (Second Edition)
59. Hermann Weyl: Algebraic Theory of Numbers (Reprint)
60. Carl Ludwig Siegel: Transcendental Numbers (Reprint)
61. Steven J. Miller and Ramin Takloo-Bighash: An Invitation to Number Theory (Reprint)
62. John Milnor: Dynamics in One Complex Variable (Reprint)
63. R. P. Pakshirajan: Probability Theory: A Foundational Course
64. Sharad S. Sane: Combinatorial Techniques
65. Hermann Weyl: The Classical Groups: Their Invariants and Representations (Reprint)
66. John Milnor: Morse Theory (Reprint)
67. R. Bhatia (ed.): Connected at Infinity II
68. Donald Passman: A Course in Ring Theory (Reprint)
69. Amiya Mukherjee: Atiyah-Singer Index Theorem: An Introduction
70. Fumio Hiai and Dénes Petz: Introduction to Matrix Analysis and Applications