OPERATOR ALGEBRAS AND NON-COMMUTATVE ANALYSIS:
An introductory course with application in quantum mechanics

Eric A. Carlen[1]

Department of Mathematics, Hill Center,
Rutgers University, 110 Frelinghuysen Road Piscataway NJ 08854-8019 USA

May 4, 2016

Contents

Abstract

We give an elementary introduction to the subject of operator algebras and non-commutative analysis with the emphasis on material related to a number of open problems arising from quantum mechanics.

1 Introduction

1.1 Basic definitions and notation

1.1 DEFINITION (Banach algebra)**.** A Banach algebra is an algebra $\mathscr{A}$ over the complex numbers equipped with a norm $\|\cdot\|$ under which it is complete as a metric space such that

$$\|ab\| \leq \|a\|\|b\| \quad \text{for all} \quad a, b \in \mathscr{A} \ . \tag{1.1}$$

1.2 EXAMPLE. Let X be a locally compact Hausdorff space, and let $\mathscr{C}_0(X)$ denote the set of continuous complex valued functions on X that vanish at infinity, and equip it with the supremum norm. Then with the usual algebraic structure of pointwise addition and multiplication, $\mathscr{A} = \mathscr{C}_0(X)$ is a Banach algebra. This is the canonical example of a commutative Banach algebra. There is a multiplicative identity if and only if X is compact.

1.3 EXAMPLE. Let $\mathscr{A} = L^1(\mathbb{R}^n)$ equipped with the convolution product

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\mathrm{d}y \ .$$

Let the norm be the L^1 norm. Then $\mathscr{A}$ is a commutative Banach algebra that does not have an identity.

1.4 EXAMPLE. Let $\mathscr{H}$ be a Hilbert space, and let $\mathscr{A} = \mathscr{B}(\mathscr{H})$, the set of all continuous linear mapping from $\mathscr{H}$ to $\mathscr{H}$, equipped with the composition product and the operator norm

$$\begin{aligned} \|a\| &= \sup\{ \ \|a\psi\|_{\mathscr{H}} \ : \ \psi \in \mathscr{H}, \|\psi\|_{\mathscr{H}} = 1 \ \} \\ &= \sup\{ \ \Re(\langle \varphi, a\psi \rangle_{\mathscr{H}}) \ : \ \varphi, \psi \in \mathscr{H}, \|\varphi\|_{\mathscr{H}}, \|\psi\|_{\mathscr{H}} = 1 \ \} \ , \end{aligned} \tag{1.2}$$

where $\|\cdot\|_{\mathscr{H}}$ is the norm on $\mathscr{H}$, and $\langle \cdot, \cdot \rangle_{\mathscr{H}}$ is the inner product in $\mathscr{H}$. This is the canonical example of a non-commutative Banach algebra.

1.5 EXAMPLE. Let $\mathscr{A}$ be the algebra of $n \times n$ matrices. The *Frobenius*, or *Hilbert-Schmidt* norm on $\mathscr{A}$ is the norm $\|\cdot\|_2$ given by

$$\|a\|_2 = \left(\sum_{i,j=1}^n |a_{i,j}|^2 \right)^{1/2}$$

where $a_{i,j}$ denotes the i,jth entry of a. By the Cauchy-Schwarz inequality, for all $a, b \in \mathscr{A}$,

$$\|ab\|_2 = \left(\sum_{i,j=1}^n \left| \sum_{k=1}^n a_{i,k} b_{k,j} \right|^2 \right)^{1/2} \le \left(\sum_{i,j=1}^n \left(\sum_{k=1}^n |a_{i,k}|^2 \right) \left(\sum_{k=1}^n |b_{k,j}|^2 \right) \right)^{1/2} = \|a\|_2 \|b\|_2 \ ,$$

and thus (1.1) is satisfied. Note that the algebra of $n \times n$ matrices with the operator norm is the special case of Example 1.4 in which $\mathscr{H} = \mathbb{C}^n$.

1.6 DEFINITION (C^*-algebra)**.** A C^* *algebra* is a Banach algebra equipped with a conjugate linear map $* : \mathscr{A} \to \mathscr{A}$, the action of which is written as $a \mapsto a^*$, and which satisfies the properties

(i) The $*$ map is an involution; for all $a \in \mathscr{A}$, $a^{**} = a$.

(Ii) For all $a, b \in \mathscr{A}$, $(ab)^* = b^* a^*$.

(iii) For all $a \in \mathscr{A}$,

$$\|aa^*\| = \|a\|^2 \ . \tag{1.3}$$

When discussing a C^* algebra it is convenient and standard to refer to the map $a \mapsto a^*$ as *the involution* in $\mathscr{A}$.

In a C^* algebra, the involution is always an isometry. This is because

$$\|a\|^2 = \|aa^*\| \le \|a\|\|a^*\| \ ,$$

where we used (1.3) and (1.1) in succession. Then for $a \neq 0$, we have $\|a\| \le \|a^*\|$, and then $\|a^*\| \le \|a^{**}\| = \|a\|$, so that $\|a^*\| = \|a\|$ for all $a \in \mathscr{A}$. The condition (1.3) is much stronger than the condition that $a \mapsto a^*$ be an isometry.

1.7 EXAMPLE. In Example 1.2, take the involution to be pointwise complex conjugation of the functions that constitute the algebra. The conditions *(i)*, *(ii)* and *(iii)* are all clearly satisfied in this case. Thus, $\mathscr{C}_0(X)$ equipped with this structure is a commutative C^*-algebra.

Similarly, in Example 1.4, take define the involution by taking a^* to be the Hermitian conjugate of a. That is, for all $\varphi, \psi \in \mathscr{H}$,

$$\langle a^*\varphi, \psi\rangle_{\mathscr{H}} = \langle \varphi, a\psi\rangle_{\mathscr{H}} \ .$$

It is immediate from this that $a \mapsto a^*$ is conjugate linear and an involution satisfying *(ii)*. Also, It is immediate from this and (1.2) that $\|a^*\| = \|a\|$ for all a. Moreover, for all $\psi \in \mathscr{H}$,

$$\langle \psi, aa^*\psi\rangle_{\mathscr{H}} = \langle a^*\psi, a^*\psi\rangle_{\mathscr{H}} = \|a^*\psi\|^2 \ ,$$

and hence

$$\begin{aligned}\|aa^*\| &= \sup\{\ \Re(\langle \varphi, aa^*\psi\rangle_{\mathscr{H}})\ :\ \varphi,\psi \in \mathscr{H}, \|\varphi\|_{\mathscr{H}}, \|\psi\|_{\mathscr{H}} = 1\ \}\\ &\ge (\sup\{\ \|a^*\psi\|_{\mathscr{H}}\ :\ \psi \in \mathscr{H}, \|\psi\|_{\mathscr{H}} = 1\ \})^2 = \|a^*\|^2 = \|a\|^2 \ .\end{aligned}$$

1.8 EXAMPLE. If we equip the convolution algebra of Example 1.3 with the convolution given by pointwise complex conjugation, this innvolution is an isometry since for any $f \in L^1(\mathbb{R})$, $\|f\|_{L^1(\mathbb{R})} = \|f^*\|_{L^1(\mathbb{R})}$. However, the stronger property (1.3) does not hold in general in this algebra. To see this, let ρ and σ be two non-negative functions in $L^1(\mathbb{R}^n)$. Fix $\lambda \in \mathbb{R}$ and define functions f and g in $L^1(\mathbb{R}^n)$ by

$$f(x) = \rho(x)e^{i\lambda x} \quad \text{and} \quad g(x) = \sigma(x)e^{i\lambda x} \ .$$

Then

$$f * g(x) = e^{i\lambda x}\rho \star \sigma(x) \qquad \text{so that} \qquad \|f * g\|_{L^1(\mathbb{R})} = \|\rho * \sigma\|_{L^1(\mathbb{R})} \ .$$

However,

$$f * g^*(x) = e^{i\lambda x}\int_{\mathbb{R}} \rho(x)\sigma(x-y)e^{-2i\lambda y}\mathrm{d}y \ ,$$

and now a simple argument using the Riemann-Lebesgue Lemma and the Dominated Convergence Theorem shows that $\|f * g^*\|_{L^1(\mathbb{R})}$ converges to zero as λ is taken to infinity. Hence (1.3) fails in this Banach algebra.

Now consider Example 1.5, and again define a^* to be the Hermitian conjugate of a. This is a conjugate linear involution, and as above, $(ab)^* = b^*a^*$ for all a, b. This involution is even an isometry in the Frobenius norm since

$$\|a\|_2^2 = \sum_{i,j=1}^n |a_{i,j}|^2 = \sum_{i,j=1}^n |a^*_{j,i}|^2 = \|a^*\|_2^2 \ .$$

However, the property (1.3) fails. Recall that this algebra has a multiplicative identity e, the $n \times n$ identity matrix. Evidently, $e^* = e$, and so were (1.3) to hold, we would have $\|e\|_2 = \|ee^*\|_2 = \|e\|_2^2$, but for $n > 1$ this is false since $\|e\|_2 = \sqrt{n}$.

The condition (1.3) is very strong. As we shall soon see, given an algebra $\mathscr{A}$ with a conjugate linear involution $*$ satisfying $(ab)^* = b^*a^*$ for all $a, b \in \mathscr{A}$, there is at most one norm on $\mathscr{A}$ that makes it a C^* algebra. We shall also see that our two examples of C^*-algebras, given in Example 1.7 are universal. In particular, a theorem of Gelfand and Naimark says that every C^* algebra is isomorphic to a sub-algebra of $\mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$ that one constructs from the algebra itself using the *Galfand-Naimark-Segal* construction, to which we sahll come.

An *operator C^* algebra* is an operator norm closed subalgebra of $\mathscr{B}(\mathscr{H})$ that is closed under Hermitian conjugation and closed in the operator norm. The term C^*-algebra was first applied in this context by Irving Segal in 1947. The theorem of Galfand and Naimark shows that "abstract" C^*-algebras, as defined here, are essentially the same thing.

There are other important topologies in $\mathscr{B}(\mathscr{H})$ that are weaker than the topology given by the operator norm. In particular, there is the *weak operator topology* on $\mathscr{B}(\mathscr{H})$ which is the weakest topology under which the maps

$$a \mapsto \langle \varphi, a\psi \rangle_{\mathscr{H}}$$

are continuous for all $\varphi, \psi \in \mathscr{H}$. Every subset (and hence subalgebra) of $\mathscr{B}(\mathscr{H})$ that is weakly closed is norm closed, but not *vice-versa* in the case that $\mathscr{H}$ is infinite dimensional. A *von Neumann algebra* is a subalgebra of $\mathscr{B}(\mathscr{H})$ that is closed under Hermitian conjugation and closed in the weak operator topology. Thus, von Neumann algebras are a special type of C^* operator algebras, and these shall be important to us also.

We shall be especially interested in von Neumann algebras, but every von Neumann algebra is a C^* algebra, and every C^* algebra is a Banach algebra, and it is natural to begin developing the theory at this general level, which we do in the next subsection.

1.2 The spectrum and the resolvent set

Let X be a locally compact Hausdorff space that is not compact. Then $\mathscr{A} = \mathscr{C}_0(X)$ equipped with the usual structures is a Banach algebra without an identity. Let $\widetilde{\mathscr{A}}$ be the larger algebra obtained by adjoining to $\mathscr{A}$ the constant functions, $\lambda 1$, $\lambda \in \mathbb{C}$. Then every $\widetilde{a} \in \widetilde{\mathscr{A}}$ has the form $\widetilde{a}(x) = \lambda + a(x)$ where $a \in \mathscr{C}_0(X)$. Then for $\lambda + a$ and $\mu + b$ in $\widetilde{\mathscr{A}}$,

$$(\lambda + a)(\mu + b) = \lambda\mu + (\lambda b + \mu a + ab) \ .$$

The constant function 1 is the multiplicative identity in $\widetilde{\mathscr{A}}$.

The procedure can be done in general. Let $\mathscr{A}$ be any Banach algebra, with or without a unit. Define $\widetilde{\mathscr{A}}$ to be $\mathbb{C} \oplus \mathscr{A}$ with the multiplication

$$(\lambda, a)(\mu, b) = (\lambda\mu, \lambda b + \mu a + ab) \ , \tag{1.4}$$

and the norm

$$\|(\lambda a)\| = |\lambda| + \|a\| \ . \tag{1.5}$$

By the definitions,

$$\begin{aligned} \|(\lambda, a)(\mu, b)\| = \|(\lambda\mu, \lambda b + \mu a + ab)\| &= |\lambda\mu| + \|\lambda b + \mu a + ab\| \\ &\le |\lambda||\mu| + |\lambda|\|b\| + |\mu|\|a\| + \|a\|\|b\| \\ &= (|\lambda| + \|a\|)(|\mu| + \|b\|) = \|(\lambda, a)\|\|(\mu, b)\| \ . \end{aligned}$$

This shows that (1.1) is satisfied, and hence that $\widetilde{\mathscr{A}}$ is a Banach algebra. Now define $e = (1, 0) \in \widetilde{\mathscr{A}}$. Then $(1, 0)(\lambda, a) = (\lambda, a)(1, 0) = (\lambda, a)$ so that e is the identity in $\widetilde{\mathscr{A}}$.

The original algebra $\mathscr{A}$ is embedded in $\widetilde{\mathscr{A}}$ as the subalgebra consisting of elements of the form $(0, a)$. None of these elements are invertible even when $\mathscr{A}$ itself has an identity. Indeed, if (λ, a) has an inverse (μ, b), then

$$(1, 0) = (\lambda, a)(\mu, b) = (\lambda\mu, \lambda b + \mu a + ab) \ ,$$

and this is impossible if $\lambda = 0$. However, it will be important in what follows that if $\mathscr{A}$ has a unit 1, then $1 - a$ is invertible in $\mathscr{A}$ if and only if $(1, -a)$ is invertible in $\widetilde{\mathscr{A}}$.

1.9 PROPOSITION. *Let $\mathscr{A}$ be a Banach algebra with unit 1. Then $1 - a$ is invertible if and only if there exists $b \in \mathscr{A}$ such that*

$$ab = ba = b - a \ . \tag{1.6}$$

consequently, $1 - a$ is invertible in $\mathscr{A}$ if and only if $e - a$ is invertible in $\widetilde{\mathscr{A}}$.

Proof. Suppose that $1 - a$ is invertible. Define $b = (1 - a)^{-1} - 1$. Then $(1 - a)b = 1 - (1 - a) = a$, and hence $ab = b - a$. The proof of $ba = b - a$ is similar.

Now suppose that there exists $b \in \mathscr{A}$ such that (1.6) is true. Then

$$(1 + b)(1 - a) = 1 + b - a - ab = 1 \quad \text{and} \quad (1 - a)(1 + b) = 1 - b + a - ba = 1 \ .$$

This proves the first part.

For the second part, suppose that $1 - a$ is invertible in $\mathscr{A}$. Then there exists $b \in \mathscr{A}$ such that (1.6) is satisfied. Regarding a and b as elements of $\widetilde{\mathscr{A}}$, (1.6) is satisfied also in $\widetilde{\mathscr{A}}$, and hence $(1, -a)$ is invertible in $\widetilde{\mathscr{A}}$.

Finally, suppose that $(1, -a)$ is invertible in $\widetilde{\mathscr{A}}$, and let (λ, b) be the inverse. Then

$$(1, 0) = (1, -a)(\lambda, b) = (\lambda, b - \lambda a - ab) \ .$$

Evidently $\lambda = 1$, and then $b - a - ab = 0$. A similar argument shows that $b - a - ba = 0$, and now the first part implies that $1 - a$ is invertible in $\mathscr{A}$. □

1.10 DEFINITION (Spectrum and resolvent set)**.** Let $\mathscr{A}$ be a Banach algebra, and let $a \in \mathscr{A}$. If $\mathscr{A}$ has a unit, the *spectrum of a in $\mathscr{A}$*, $\sigma_{\mathscr{A}}(a)$ is defined to be the set of all $\lambda \in \mathbb{C}$ such that $\lambda 1 - a$ is not invertible. If $\mathscr{A}$ does not have a unit, then $\sigma_{\mathscr{A}}(a)$ is defined to be the spectrum of $(0, a) \in \widetilde{\mathscr{A}}$. The resolvent set of a in $\mathscr{A}$, $\rho_{\mathscr{A}}(a)$ is defined to be the complement of $\sigma_{\mathscr{A}}(a)$.

Let $\mathscr{A}$ be a Banach algebra with a identity 1. Then we can still carry out the process of adjoining an identity to form $\widetilde{\mathscr{A}}$, and can regard each $a \in \mathscr{A}$ also as an element of $\widetilde{\mathscr{A}}$. Since no element of $\mathscr{A}$ is invertible in $\widetilde{\mathscr{A}}$, $0 \in \sigma_{\widetilde{\mathscr{A}}}(a)$ for all $a \in \mathscr{A}$. However, for $\lambda \neq 0$, $\lambda 1 - a$ is invertible if and only if $1 - a/\lambda$ is invertible. Likewise, $(\lambda - a)$ is invertible if and only if $(1, -a/\lambda)$ is invertible. Then by Proposition 1.9, $\lambda 1 - a$ is invertible in $\mathscr{A}$ if and only if $(1, 0) - (0, a/\lambda)$ is invertible in $\widetilde{\mathscr{A}}$. This shows that for $\lambda \neq 0$, $\lambda \in \sigma_{\mathscr{A}}(a) \iff \lambda \in \sigma_{\widetilde{\mathscr{A}}}((0, a))$. We summarize:

$$\{0\} \cup \sigma_{\mathscr{A}}(a) = \sigma_{\widetilde{\mathscr{A}}}((0, a)) \ . \tag{1.7}$$

1.11 LEMMA (Spectral Mapping Lemma). *Let $\mathscr{A}$ be a Banach algebra, and let p be a polynomial. In case $\mathscr{A}$ has no identity, we suppose that p has no constant term. Then*

$$p(\sigma_{\mathscr{A}}(a)) = \sigma_{\mathscr{A}}(p(a)) \ .$$

Proof. We may suppose that p is not identically constant. We first suppose that $\mathscr{A}$ has an identity. Fix $\lambda \in \sigma_{\mathscr{A}}(a)$. We shall show that $p(\lambda)1 - p(a)$ is not invertible. The polynomial $p(\lambda) - p(z)$ has a root at $z = \lambda$, and hence

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

for some polynomial $q(z)$. Replacing z by a,

$$p(\lambda)1 - p(a) = (\lambda - a)q(a) \ .$$

Were $p(\lambda)1 - p(a)$ invertible, we would have $1 = (\lambda - a)[q(a)(p(\lambda) - p(a))^{-1}]$, and then since polynomials in a commute, $1 = [q(a)(p(\lambda) - p(a))^{-1}](\lambda - a)$. This would mean that $\lambda 1 - a$ is invertible, with contradicts our hypothesis that $\lambda \in \sigma_{\mathscr{A}}(a)$. Hence $p(\lambda) - p(a)$ is not invertible, and hence $p(\lambda) \in \sigma_{\mathscr{A}}(p(a))$. this shows that $p(\sigma_{\mathscr{A}}(a)) \subset \sigma_{\mathscr{A}}(p(a))$.

Next, fix $\mu \in \sigma_{\mathscr{A}}(p(a))$, and factor

$$\mu - p(z) = \alpha(\lambda_1 - a) \cdots (\lambda_n - z)$$

where $\alpha \neq 0$ and $n \geq 1$. For each j, $\mu = p(\lambda_j)$. We have

$$\mu 1 - p(a) = \alpha(\lambda_1 1 - a) \cdots (\lambda_n 1 - a)$$

and if each $\lambda_j 1 - a$ were invertible, then $\mu 1 - p(a)$ would be invertible, but this is not the case. Hence Hence for some j, $\lambda_j \in \sigma_{\mathscr{A}}(a)$, and $\mu = p(\lambda_j) \in \sigma_{\mathscr{A}}(p(a))$. This shows that $\sigma_{\mathscr{A}}(p(a)) \subset p(\sigma_{\mathscr{A}}(a))$, and completes th proof when $\mathscr{A}$ has an identity. The general case now follows by adjoining an identity and then appealing to (1.7). □

1.3 Properties of the inverse function

Now let $\mathscr{A}$ be a Banach algebra with an identity 1. Let $a \in \mathscr{A}$ be such that $\|1 - a\| = r < 1$. Then by the defining property (1.1), $\|(1 - a)^n\| \leq r^n$ for all $n \in \mathbb{N}$. For all $n \in \mathbb{N}$, define

$$s_n = \sum_{j=1}^{n} (1 - a)^j$$

where, as usual, we interpret $(1 - a)^0 = 1$. Then for all $n > m$, by the triangle inequality and (1.1),

$$\|s_n - s_m\| \leq \sum_{j=m+1}^{n} \|(1 - a)^j\| \leq \sum_{j=m+1}^{n} r^j = \frac{r^m - r^n}{r - 1} \ .$$

Hence $\{s_n\}_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathscr{A}$. Now for the first time we use the metric completeness of $\mathscr{A}$: There exists $b \in \mathscr{A}$ such that $\lim_{n \to \infty} \|b - s_n\| = 0$. But then

$$ba = \lim_{n \to \infty} s_n a = \lim_{n \to \infty} s_n(1 - (1 - a)) = \lim_{n \to \infty} (1 - (1 - a)^{n+1}) = 1 \ .$$

The same reasoning shows that $ab = 1$, and so a is invertible. Let Ω denote the set of invertible elements in $\mathscr{A}$. This brings us to:

1.12 LEMMA. *Let $\mathscr{A}$ be a Banach algebra with a unit. Let Ω be the set of invertible elements of $\mathscr{A}$. Then Ω contains every $a \in \mathscr{A}$ such that $\|1-a\| < 1$, and in this case a^{-1} is given by the convergent series*

$$a^{-1} = \sum_{j=0}^{\infty}(1-a)^j \ .$$

Moreover, if $|\lambda| > \|a\|$, then $\lambda 1 - a$ is invertible, with

$$\|(\lambda 1 - a)^{-1}\| \le \frac{1}{|\lambda| - \|a\|} \ . \tag{1.8}$$

In particular, $\sigma_{\mathscr{A}}(a)$ is contained in the closed disk of radius $\|a\| \in \mathbb{C}$.

Proof. It remains to prove the final part. If $|\lambda| > \|a\|$, the $\lambda 1 - a = \lambda(1-\lambda^{-1}a)$ and $\|1-(1-\lambda^{-1}a)\| = |\lambda|^{-1}\|a\| < 1$, so that $(1-\lambda^{-1}a)$ is invertible. □

At this point, we do not know in general that $\sigma_{\mathscr{A}}(a)$ is not empty, but we do know this of $\rho_{\mathscr{A}}(a)$. We now claim that Ω is open. This has the immediate consequence that $\rho_{\mathscr{A}}(a)$ is open, and hence that $\sigma_{\mathscr{A}}(a)$ is closed, though at this point the possibility that $\sigma_{\mathscr{A}}(a) = \emptyset$ has not yet ben eliminated.

Let $a_0 \in \Omega$ and $a \in \mathscr{A}$. Then $\|1 - aa_0^{-1}\| = \|(a_0 - a)a_0^{-1}\| \le \|a - a_0\|\|a_0^{-1}\|$ Therefore, for any $r \in (0,1)$,

$$\|a - a_0\| \le r\|a_0^{-1}\|^{-1} \quad \Rightarrow \quad \|1 - aa_0^{-1}\| \le r \quad \Rightarrow \quad aa_0^{-1} \in \Omega \ .$$

Since Ω is closed under multiplication, $a = (aa_0^{-1})a_0 \in \Omega$. This shows that for all $a_0 \in \Omega$, the open ball of radius $\|a_0^{-1}\|^{-1}$ an center a_0 is contained in Ω. In particular, Ω is open.

Now recall that a function F from a Banach space X to itself is *Frechet differentiable* at $x_0 \in X$ in case there is a continuous linear transformation L from X to itself such that for all $x \in X$,

$$\|F(x_0 + x) - F(x_0) - Lx\| = o(\|x\|) \ ,$$

and in this case, L is unique and is the *Frechet derivative* of F at x_0. We now show that the inverse function $a \mapsto a^{-1}$ is Frechet differentiable at every $a_0 \in \mathscr{A}$, and that the derivative is the linear transformation

$$a \mapsto -a_0^{-1}aa_0^{-1} \ .$$

This is a simple consequence of an important identity that we record in a lemma:

1.13 LEMMA (First resolvent identity)**.** *Let $\mathscr{A}$ be a Banach algebra with an identity* 1*. Let Ω be the set of invertible elements. For all $a, b \in \Omega$,*

$$a^{-1} - b^{-1} = a^{-1}(b-a)b^{-1} \ . \tag{1.9}$$

Proof. Simply expand the right hand side. □

Now suppose that $a_0, a_0 + a \in \Omega$. Then

$$(a_0 + a)^{-1} - a_0^{-1} = -(a_0 + a)^{-1}aa_0^{-1} = -a_0^{-1}aa_0^{-1} + [a_0^{-1} - (a_0 + a)^{-1}]aa_0^{-1} \ .$$

By (1.1) once more and the continuity proved above,

$$\|[a_0^{-1} - (a_0 + a)^{-1}]aa_0^{-1}\| \le \|a_0^{-1}\|\|a_0^{-1} - (a_0 + a)^{-1}\|\|a\| = o(\|a\|) \ .$$

We are now ready to show that for all a in any Banach algebra, $\sigma_{\mathscr{A}}(a) \neq \emptyset$. Let φ be any continuous linear functional on $\mathscr{A}$, regarded as a Banach space. Such functionals exist (and are plentiful) by the Hahn-Banach Theorem. Define a complex valued function f in the resolvent set $\rho_{\mathscr{A}}(a)$ by

$$f(\zeta) = \varphi((\zeta 1 - a)^{-1}) \ .$$

Note that the resolvent set includes $\{\zeta \ : \ |\zeta| > \|a\|\}$, and that by (1.8),

$$\lim_{\zeta \to \infty} f(\zeta) = 0 \ . \tag{1.10}$$

Next, by the identity (1.9),

$$f(\zeta + \eta) - f(\zeta) = \eta\varphi[(((\zeta + \eta)1 - a)^{-1})(\zeta 1 - a)^{-1}] \ .$$

From this identity and the continuity of the inverse function, it follows that

$$\lim_{\eta \to 0} \frac{f(\zeta + \eta) - f(\zeta)}{\eta} = \varphi[(\zeta 1 - a)^{-2}] \ ,$$

which shows that f is an analytic function on $\rho_{\mathscr{A}}(a)$.

If the resolvent set $\rho_{\mathscr{A}}(a)$ were all of $\mathbb{C}$, f would be an entire analytic function, and on account of (1.10), f would also be bounded. By Liouville's Theorem it would then be constant, and by (1.10), the constant would have to be zero. In particular, we would have $f(0) = 0$. Therefore, for every continuous linear functional φ on $\mathscr{A}$, it would be the case that $\varphi(a^{-1}) = 0$. This contradicts the Hahn-Banach Theorem. We summarize:

1.14 THEOREM. *Let $\mathscr{A}$ be any Banach algebra with an identity* 1*. Then for all $a \in \mathscr{A}$, $\sigma_{\mathscr{A}}(a)$ is a nonempty closed set contained in the closed disc of radius $\|a\|$ centered at* 0 *in $\mathbb{C}$.*

It is now a simple matter to prove:

1.15 THEOREM (Gelfand-Mazur Theorem)**.** *Let $\mathscr{A}$ be a Banach algebra with identity* 1*. If $\mathscr{A}$ is a division algebra, then $\mathscr{A}$ is isomorphic to $\mathbb{C}$. More specifically, each element a of $\mathscr{A}$ satisfies $a = \lambda 1$ for some necessarily unique $\lambda \in \mathbb{C}$, and $a \mapsto \lambda$ is an isomorphism with $\mathbb{C}$.*

Proof. Suppose that $\mathscr{A}$ is a division algebra. By Theorem 1.14, there exists $\lambda \in \sigma_{\mathscr{A}}(a)$. Thus $\lambda 1 - a$ is not invertible. Since the only non-invertible element in a division algebra is 0, $a = \lambda 1$. □

1.16 DEFINITION (Spectral radius)**.** The spectral radius of an element a of a Banach algebra $\mathscr{A}$ is

$$\nu(a) = \max\{ \ |\lambda| \ : \ \lambda \in \sigma_{\mathscr{A}}(a) \ \} \ . \tag{1.11}$$

1.17 THEOREM (Gelfand's formula)**.** *The spectral radius of an element a of a Banach algebra $\mathscr{A}$ is given by*

$$\nu(a) = \lim_{n \to \infty} \|a^n\|^{1/n} \ . \tag{1.12}$$

In particular, the limit exists.

Proof. Suppose that $\lambda \in \sigma_{\mathscr{A}}(a)$. Then by the Spectral Mapping Lemma, $\lambda^n \in \sigma_{\mathscr{A}}(a^n)$, and then by Theorem 1.14 and (1.7) in case $\mathscr{A}$ lacks an identity, $|\lambda|^n \leq \|a^n\|$. Taking the nth root, we obtain $\nu(a) \leq \|a^n\|^{1/n}$ for all $n \in \mathbb{N}$.

It remains to show that

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq \nu(a) \ . \tag{1.13}$$

To this end, pick λ with $|\lambda| > \nu(a)$ so that $\lambda \in \rho_{\mathscr{A}}(a)$. Let φ be any continuous linear functional on $\mathscr{A}$. Then, as in the proof of Theorem 1.14, the function f defined by $f(\lambda) = \varphi((\lambda 1 - a)^{-1})$ is analytic on $\rho_{\mathscr{A}}(a)$. Define $\zeta = 1/\lambda$, $g(\zeta) = f(1/\lambda) = \zeta\varphi((1-\zeta a)^{-1})$, which is analytic on the open disc about 0 with radius $1/\nu(a)$.

For ζ with $|\zeta| < \|a\|^{-1}$, $(1-\zeta a)^{-1}$ has the convergent power series $(1-\zeta a)^{-1} = \sum_{n=0}^{\infty} \zeta^n a^n$. Therefore, by the uniqueness of the power series representation, $g(z) = \sum_{n=0}^{\infty} \zeta^{n+1}\varphi(a^n)$ is a convergent power series for all ζ with $|\zeta| \leq 1/\nu(a)$. It follows that for all such ζ, $\lim_{n\to\infty} \zeta^{n+1}\varphi(a^n) = 0$. In particular, there exists a finite constant C_φ such that

$$|\zeta^{n+1}\varphi(a^n)| \leq C_\varphi \quad \text{for all} \quad n \in \mathbb{N} \ . \tag{1.14}$$

Now, for each $n \in \mathbb{N}$ define a linear functional Λ_n on $\mathscr{A}^*$, the Banach space dual of $\mathscr{A}$, by

$$\Lambda_n(\varphi) = \zeta^{n+1}\varphi(a^n) \ .$$

Then (1.15) says that

$$\sup_{n\in\mathbb{N}}\{\ |\Lambda_n(\varphi)|\} \leq C_\varphi \ . \tag{1.15}$$

The Uniform Boundedness Principle then implies that there exists a finite constant M such that $\|\Lambda_n\| \leq M$ for all n, and hence for all $\varphi \in \mathscr{A}^*$ with $\|\varphi\| = 1$,

$$|\zeta|^{n+1}|\varphi(a^n)| \leq M \quad \text{for all} \quad n \in \mathbb{N}$$

. The Hahn-Banach Theorem provides $\varphi \in \mathscr{A}^*$ with $\|\varphi\| = 1$ such that $\varphi(a^n) = \|a^n\|$. Hence we have $|\zeta|^{n+1}\|a^n\| \leq M$. Taming the nth root of both sides,

$$|\zeta|\|a^n\|^{1/n} \leq \left(\frac{M}{|\zeta|}\right)^{1/n} \ .$$

This proves that $|\zeta| \limsup_{n\to\infty} \|a^n\|^{1/n} \leq 1$. However, ζ was any complex number with $|\zeta| < 1/\nu(a)$, this proves (1.13). □

We close this section with the following results that is trivial for commutative Banach algebras, and familiar for the algebra of $n \times n$ matrices.

1.18 THEOREM (Spectrum of ab and ba)**.** *If $\mathscr{A}$ is a Banach algebra, then for all $a, b \in \mathscr{A}$,*

$$\{0\} \cup \sigma_{\mathscr{A}}(ab) = \{0\} \cup \sigma_{\mathscr{A}}(ba) \ . \tag{1.16}$$

Proof. By passing to $\widetilde{\mathscr{A}}$, we may suppose that $\mathscr{A}$ has an identity. For each $\lambda \neq 0$, we must show that $(\lambda 1 - ab)$ is invertible if and only iff $(\lambda 1 - ba)$ is invertible. Dividing through by λ, we may take $\lambda = 1$. Therefore, suppose that $(1 - ab)$ is invertible, and let $z = (1 - ab)^{-1}$. Then

$$\begin{aligned}(1 - ba)(1 + bza) &= 1 - ba + bza - babza \\ &= 1 - ba + b(1 - ab)za = 1 - ba + ba = 1 \ .\end{aligned}$$

Likewise, $(1 + bza)(1 - ba) = 1$, and so $(1 - ba)$ is invertible with inverse $(1 + bza)$. □

1.19 THEOREM (Spectral Contraction Theorem)**.** *Let $\mathscr{A}$ and $\mathscr{B}$ be Banach algebras, and let $\pi : \mathscr{A} \to \mathscr{B}$ be a homomorpism. Then for all $a \in \mathscr{A}$,*

$$\sigma_{\mathscr{B}}(\pi(a)) \subset \{0\} \cup \sigma_{\mathscr{A}}(a) \ . \tag{1.17}$$

Proof. Adjoin identities to $\mathscr{A}$ and $\mathscr{B}$, and define $\widetilde{\pi} : \widetilde{\mathscr{A}} \to \widetilde{\mathscr{B}}$ by $\widetilde{\pi}((1, a)) = (1, \pi(a))$. This is a homomorphism, and takes the identity in $\widetilde{\mathscr{A}}$ to the identity in $\widetilde{\mathscr{B}}$. Since adjoinng an identity had no effect on non-zero spectrum, we may assume that $\mathscr{A}$ and $\mathscr{B}$ have identities $1_{\mathscr{A}}$ and $1_{\mathscr{B}}$ respectively, and that $\pi(1_{\mathscr{A}}) = 1_{\mathscr{B}}$.

Now suppose that $\lambda \in \rho_{\mathscr{A}}(a)$. Then $1_{\mathscr{A}} = (\lambda 1_{\mathscr{A}} - a)(\lambda 1_{\mathscr{A}} - a)^{-1}$. Since π is a homomorphism,

$$1_{\mathscr{B}} = \pi(1_{\mathscr{A}}) = (\lambda 1_{\mathscr{B}} - \pi(a))\pi((\lambda 1_{\mathscr{A}} - a)^{-1}) \ .$$

Thus $\pi((\lambda 1_{\mathscr{A}} - a)^{-1})$ is a right inverse of $\lambda 1_{\mathscr{B}} - \pi(a)$, and the same reasoning shows it is also a left inverse. Hence $\lambda \in \rho_{\mathscr{B}}(\pi(a))$. This shows that $\rho_{\mathscr{A}}(a) \subset \rho_{\mathscr{B}}(\pi(a)$, which is equivalent to the statement $\sigma_{\mathscr{B}}(\pi(a)) \subset \sigma_{\mathscr{A}}(a)$, and even shows that when $\mathscr{A}$ and $\mathscr{B}$ have identities and π takes the identity in $\mathscr{A}$ to that in $\mathscr{B}$, it is not necessary to adjoin $\{0\}$ on the right side in (1.17) □

1.4 Characters and the Gelfand Transform

1.20 DEFINITION (Characters)**.** A *character* of a Banach algebra $\mathscr{A}$ is a non-zero algebraic homomorphism from $\mathscr{A}$ to $\mathbb{C}$. The set of characters of $\mathscr{A}$ is denoted $\Delta(\mathscr{A})$, and the set $\{0\} \cup \Delta(\mathscr{A})$ is denoted $\Delta'(\mathscr{A})$.

Though characters are defined with respect to the algebraic structure alone, they are necessarily continuous:

1.21 LEMMA. *If $\mathscr{A}$ is a Banach algebra and φ is a character of $\mathscr{A}$, then $\varphi(a) \in \sigma_{\mathscr{A}}(a)$, and*

$$|\varphi(a)| \leq \|a\| \tag{1.18}$$

for all $a \in \mathscr{A}$. That is φ is a contraction from $\mathscr{A}$ to $\mathbb{C}$. Moreover, if $\mathscr{A}$ has an identity 1*, then $\varphi(1) = 1$.*

Proof. Suppose first that $\mathscr{A}$ contains an identity 1. We first prove the final statement. Since $\varphi(1) = \varphi(1^2) = (\varphi(1))^2$, $\varphi(1)$ solves $\zeta - \zeta^2 = 0$, so either $\varphi(1) = 0$ or $\varphi(1) = 1$. But if $\varphi(1) = 0$, then for all $a \in \mathscr{A}$, $\varphi(a) = \varphi(1a) = \varphi(1)\varphi(a) = 0$, and this is excluded by the definition. Hence $\varphi(1) = 1$.

Next, for any $a \in \mathscr{A}$, $\varphi(a)1 - a$ is not invertible, and hence $\varphi(a) \in \sigma_{\mathscr{A}}(a)$. To see this, note that $\varphi(\varphi(a)1 - a) = 0$, but if $\varphi(a)1 - a$ had even a right inverse b, we would have

$$1 = \varphi(1) = \varphi((\varphi(a)1 - a)b) = 0\varphi(b) = 0 \ .$$

Then since $\sigma_{\mathscr{A}}(a)$ is contained in the closed centered disc of radius $\|a\|$, (1.18) is proved.

Now suppose that $\mathscr{A}$ lacks a unit. Let $\widetilde{\mathscr{A}}$ be the algebra obtained by adjoining an identity, and let $\widetilde{\varphi}$ be the character on $\widetilde{\mathscr{A}}$ given by

$$\widetilde{\varphi}((\lambda, a)) = \lambda + \varphi(a) \ ,$$

which is easily seen to be a character. Since $\sigma_{\mathscr{A}}(a)) = \sigma_{\widetilde{\mathscr{A}}}((0,a))$ be definition, and $\widetilde{\varphi}((0,a)) = \varphi(a)$, it follows from the above that $\varphi(a) \in \sigma(a)$, and then that $\|\varphi(a)\| \leq \|(0,a)\| = \|a\|$. □

Note that if $\varphi \in \Delta(\mathscr{A})$, then for all $a, b \in \mathscr{A}$,

$$\varphi(ab) = \varphi(a)\varphi(b) = \varphi(b)\varphi(a) = \varphi(ba) \ .$$

Consequently, $\varphi(ab - ba) = 0$ for all a, b. When the algebra $\mathscr{A}$ is not commutative, this can be a stringent constraint, and *there may not exist any characters at all.*

1.22 EXAMPLE. Let $\mathscr{A}$ be the algebra of 2×2 matrices. The Pauli matrices are

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ , \qquad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \text{and} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \ .$$

Then with $[a, b]$ denoting the commutator $ab - ba$,

$$[\sigma_1, \sigma_2] = i2\sigma_3 \ , \qquad [\sigma_2, \sigma_3] = i2\sigma_1 \quad \text{and} \quad [\sigma_3, \sigma_1] = i2\sigma_2 \ .$$

It follows that for any homomorphism φ of $\mathscr{A}$ into $\mathbb{C}$, $\varphi(\sigma_j) = 0$ for $j = 1, 2, 3$. Next, the identity matrix I satisfies $I = \sigma_1^2$, and so $\varphi(I) = (\varphi(\sigma_1))^2 = 0$. Thus, for all $(z_0, z_1, z_2, z_3) \in \mathbb{C}^4$,

$$\varphi(z_0 I + z_1\sigma_1 + z_2\sigma_2 + z_3\sigma_3) = 0 \ .$$

Since every evidently $\{I, \sigma_1, \sigma_2, \sigma_3\}$ is linearly independent and $\mathscr{A}$ is 4 dimensional, $\mathscr{A}$ is the span of $\{I, \sigma_1, \sigma_2, \sigma_3\}$, and hence φ vanishes identically on $\mathscr{A}$. Thus, if $\mathscr{A}$ is the algebra of 2×2 matrices, $\Delta(\mathscr{A}) = \emptyset$ and $\Delta'(\mathscr{A})$ is the one-point space $\{0\}$.

Even when $\mathscr{A}$ is commutative, there may be no non-trivial characters. However, as we shall see in the next chapter, when $\mathscr{A}$ is a commutative C^* algebra, characters are plentiful enough to justify our present considerations. In the rest of this chapter, commutativity of the algebras will not play any role in the proofs, and so we shall state the results without making any reference to commutativity. However, one should keep in mind that without commutativity, and even with commutativity alone, $\Delta(\mathscr{A})$ may be empty and $\Delta'(\mathscr{A})$ may be a one-point space, as in the previous example.

1.23 DEFINITION (Gelfand topology)**.** For a Banach algebra $\mathscr{A}$, the *Gelfand topology* on $\Delta'(\mathscr{A})$ is the relative weak-$*$ topology on $\Delta'(\mathscr{A})$ considered as a subset of $\mathscr{A}^*$, the Banach space dual to $\mathscr{A}$. That is, the Gelfand topology is the weakest topology on $\Delta'(\mathscr{A})$ that makes the functions $\varphi \mapsto \varphi(a)$ continuous for all $a \in \mathscr{A}$.

1.24 LEMMA. *Let $\mathscr{A}$ be a Banach algebra. Then $\Delta'(\mathscr{A})$, equipped with the Gelfand topology is a compact Hausdorff space. If $\mathscr{A}$ does not have an identity, then with the Gelfand topology, $\Delta(\mathscr{A})$ is a locally compact Hausdorff space, and $\Delta'(\mathscr{A})$ is its one-point compactification. If $\mathscr{A}$ has an identity, $\Delta(\mathscr{A})$ itself is compact and 0 is an isolated point in $\Delta'(\mathscr{A})$.*

Proof. Equip $\mathscr{A}^*$ with the weak-$*$ topology; i.e., the weakest topology making all of functions $\varphi \mapsto \varphi(a)$ continuous for all $a \in \mathscr{A}$. The Banach-Alaoglu Theorem asserts that the unit ball in $\mathscr{A}^*$ is compact in the weak-$*$ topology. For each $a, b \in \mathscr{A}$, define a function $f_{a,b}$ on $\mathscr{A}^*$ by

$$f_{a,b}(\varphi) = \varphi(ab) - \varphi(a)\varphi(b) \ .$$

This is evidently continuous for the weak-$*$ topology. Now note that

$$\Delta'(\mathscr{A}) = \bigcap_{a,b\in\mathscr{A}} \{\varphi \in \mathscr{A}^* \ : \ f_{a,b}(\varphi) = 0 \ \} \ .$$

This displays $\Delta'(\mathscr{A})$ as an intersection of closed sets. Hence $\Delta'(\mathscr{A})$ is a closed subset of the unit ball in $\mathscr{A}^*$, and hence is compact.

For $\varphi_1, \varphi_2 \in \Delta'(\mathscr{A})$ with $\varphi_1 \neq \varphi_2$, there exists $a \in \mathscr{A}$ such that $\varphi_1(a) \neq \varphi_2(a)$. Let U_1 and U_2 be disjoint open sets in $\mathbb{C}$ that contain $\varphi_1(a)$ and $\varphi_2(a)$ respectively. Then $\{\psi \in \Delta'(\mathscr{A}) \ : \ \psi(a) \in U_1 \ \}$ and $\{\psi \in \Delta'(\mathscr{A}) \ : \ \psi(a) \in U_2 \ \}$ are disjoint open sets in $\Delta'(\mathscr{A})$ that contain φ_1 and φ_2 respectively. In particular, for each $\varphi \in \Delta(\mathscr{A})$, there disjoint open neighborhoods V_1 of φ and V_2 of 0, and then since $V_1 \subset V_2^c$, V_2^c is a compact neighborhood of φ. Thus, $\Delta(\mathscr{A})$ is locally compact. If $\mathscr{A}$ has an identity 1, $\varphi(1) = 1$ for all $\varphi \in \Delta(\mathscr{A})$, while $0(1) = 0$. Consequently, the zero homomorphism is an isolated point of $\Delta'(\mathscr{A})$ in this case. □

1.25 DEFINITION (Gelfand transform)**.** Let $\mathscr{A}$ be a Banach algebra. The *Gelfand transform* is the map γ from $\mathscr{A}$ to $\mathscr{C}(\Delta'(\mathscr{A}))$ given by

$$(\gamma(a))[\varphi] = \varphi(a) \ . \tag{1.19}$$

That is, $\gamma(a)$ is the function of evaluation at a, and it is continuous by the definition of the Gelfand topology.

1.26 THEOREM. *Let $\mathscr{A}$ be a Banach algebra. The Gelfand transform is a norm reducing homomorphism from $\mathscr{A}$ to $\mathscr{C}((\Delta'(\mathscr{A}))$. That is, the Gelfand transform is a homomorphism of algebras and for all $a \in \mathscr{A}$,*

$$\|\gamma(a)\|_{\mathscr{C}(\Delta'(\mathscr{A}))} \leq \|a\| \ .$$

Proof. The homomorphism property is evident since for all $a, b \in \mathscr{A}$ and all $\varphi \in \Delta'(\mathscr{A})$,

$$(\gamma(ab))[\varphi] = \varphi(ab) = \varphi(a)\varphi(b) = (\gamma(a))[\varphi](\gamma(b))[\varphi] \ .$$

Next, suppose that $\mathscr{A}$ has an identity 1. If $\varphi \in \Delta(\mathscr{A})$ and $a \in \mathscr{A}$, then

$$\varphi(\varphi(a)1 - a) = \varphi(a) - \varphi(a) = 0 \ ,$$

and so $\varphi(a)1 - a$ is not invertible. This means that $\varphi(a) \in \sigma_{\mathscr{A}}(a)$, and this is contained in the closed centered disc of radius $\nu(a) \leq \|a\|$.

If $\mathscr{A}$ lacks an identity, adjoin an identity to form $\widetilde{\mathscr{A}}$. For $\varphi \in \Delta(\mathscr{A})$, define $\widetilde{\varphi}$ on $\widetilde{\mathscr{A}}$ by

$$\widetilde{\varphi}(\lambda, a) = \lambda + \varphi(a) \ .$$

It is easy to check that $\widetilde{\varphi} \in \Delta(\widetilde{\mathscr{A}})$. Let $e = (1,0)$ denote the identity in $\widetilde{\mathscr{A}}$. Then for all $a \in \mathscr{A}$,

$$\widetilde{\varphi}(\varphi(a)e - (0,a)) = \varphi(a) - \varphi(a) = 0 \ ,$$

so that once again, we have that $\varphi(a) \in \sigma_{\mathscr{A}}(a)$. □

This result, as it stands, does not take us far at all. The problem is that at this level of generality, there may be no characters at all, and the transform may be a trivial homomorphism into a trivial algebra. As indicated above, characters can only be expected to be plentiful for commutative algebras. Even then, there may be too few characters for the Gelfand transform to be of interest. However, afundamental theorem of Gelfand and Naimark says that for commutative C^*-algebras, the Gelfand transform is an *isometric isomorphism.* This is explained in the next chapter. We close this chapter with some examples, and then an important theorem on characters in a *commutative* Banach algebra.

1.27 EXAMPLE. Let a_0 be the $n \times n$ matrix, $n > 1$, with

$$a_{i,j} = \begin{cases} 1 & j = i+1 \\ 0 & j \neq i+1 \end{cases} \ .$$

That is a_0 is the $n \times n$ matrix with 1 in every entry just above the diagonal, and zero elsewhere. It is easy to see that $a_0^n = 0$,

Let $\mathscr{A}$ denote that subalgebra of the $n \times n$ matrices that are polynomials in a_0. That is, every $a \in \mathscr{A}$ has the form

$$a = \sum_{j=0}^{n-1} p_j a_0^j \ , \tag{1.20}$$

where higher order terms are zero. This is a commutative algebra with an identity. Let $\varphi \in \Delta'(\mathscr{A})$. Then $0 = \varphi(a_0^n) = (\varphi(a_0))^n$ so that $\varphi(a_0) = n$. Then for a given by (1.20), $\varphi(a) = p_0\varphi(I) = p_0$. Thus, the only candidate for a character on $\mathscr{A}$ is the map φ_0 given by

$$\varphi_0\left(\sum_{j=0}^{n-1} p_j a_0^j\right) = p_0 \ ,$$

It is readily checked that this is indeed a homomorphism and it is non-zero. Hence $\Delta(\mathscr{A}) = \{\varphi_0\}$ and $\Delta'(\mathscr{A}) = \{\varphi_0\} \cup \{0\}$. Since $\Delta'(\mathscr{A})$ consists of two isolated points, we may identify $\mathscr{C}(\Delta'(\mathscr{A}))$ with $\mathbb{C}^2$ in the usual way, and then we may write the Gelfand transform as

$$\gamma\left(\sum_{j=0}^{n-1} p_j a_0^j\right) = (p_0, 0) \ ,$$

which is indeed a norm reducing homomorphism, but not very interesting.

Before leaving this example, we note that for elements of $\mathscr{A}$, the spectrum is as trivial as Theorem 1.14 allows: For all $a \in \mathscr{A}$, $\sigma_{\mathscr{A}}(a)$ consists of a single point: $\sigma_{\mathscr{A}}(a) = \{\varphi_0(a)\}$. This is true since when a is given by (1.20), then a is invertible if and only if $p_0 \neq 0$.

finally, note that while the algebra of all $n \times n$ matrices equipped with the usual norm is a C^* algebra, this subalgebra is not closed under the Hermitian adjoint, and hence is not a C^* algebra.

1.28 EXAMPLE. This example illustrates not what can go wrong when $\mathscr{A}$ is not a commutative C^* algebra, but what the utility of these considerations might be when $\mathscr{A}$ is a commutative C^* algebra.

Let a_0 be any $n \times n$ normal matrix; i.e., $a_0 a_0^* = a_0^* a_0$. Let $\mathscr{A}$ be the algebra of all polynomials in a_0 and a_0^*. (We may unambiguously evaluate a polynomial $p(\eta, \zeta)$ in two the variables η, ζ at $\eta = a_0$ and $\zeta = a_0^*$ precisely because a_0 and a_0^* commute.) This is a commutative algebra with an identity.

The Spectral Theorem for $n \times n$ matrices says that there exists a orthonormal basis $\{\phi_1, \dots, \phi_n\}$ of $\mathbb{C}^n$ such that each ϕ_j is an eigenvector of a_0. Let λ_j be the corresponding eigenvalue. That is, $a_0 \phi_j = \lambda_j \phi_j$. Then $a_0^* \phi_j = \lambda_j^* \phi_j$, and especially, for any polynomial p,

$$p(a_0, a_0^*)\phi_j = p(\lambda_j, \lambda_j^*)\phi_j \ .$$

For each $j = 1, \dots, n$ define a linear functional φ_j on $\mathscr{A}$ by

$$\varphi_j(a) = \langle \phi_j, a\phi_j \rangle \ .$$

By what we have noted above, for any polynomial

$$\varphi_j(p(a_0, a_0^*) = p(\lambda_j, \lambda_j^*) \ .$$

It is evident that each φ_j is a character, and that if $\lambda_j \neq \lambda_k$ than $\varphi_j \neq \varphi_k$.

In this case we have plenty of characters. We shall see in the next chapter that there are no other characters besides these. Granted that, $\Delta(\mathscr{A})$ can be identified with the set $\{\mu_1, \dots, \mu_m\}$ of distinct eigenvalues of a_0, and the Gelfand transform identities $p(a_0, a_0^*)$ with the function on $\{\mu_1, \dots, \mu_m\}$ given by $\mu \mapsto p(\mu, \mu^*)$. Since the operator norm of a normal matrix is the maximum of the absolute values of its eigenvalues, it is evident that the Gelfand transform is an isometry in this case.

1.5 Characters and spectrum in commutative Banach algebras

The Hahn-Banach Theorem, which provides the existence of continuous linear functionals on a Banach space, may be viewed as a theorem asserting the existence of maximal closed subspaces containing a given subspace. In the Banach algebra setting, the kernel of a homomorphism of a Banach algebra $\mathscr{A}$ to $\mathbb{C}$ is not only a closed subspace, it is a closed *ideal*, as we now explain, and consideration of *maximal ideals* leads to a Banach algebra version of the Hahn-Banach Theorem for commutative Banach algebras. Much of what is introduced here is also useful without assuming the $\mathscr{A}$ is commutative. We therefore start in general, and shall be clear about the key point when commutativity enters.

1.29 DEFINITION. Let $\mathscr{A}$ be a Banach algebra. An *ideal* of $\mathscr{A}$ is a subspace of $\mathscr{A}$ such that for all $b \in \mathscr{J}$ and $a \in \mathscr{A}$, $ba \in \mathscr{J}$ and $ab \in \mathscr{J}$. An ideal of $\mathscr{A}$ is *proper* in case it is not equal to $\mathscr{A}$ itself. An ideal of $\mathscr{A}$ is a *closed* in case it is topologically closed as a subset of $\mathscr{A}$. If $\mathscr{J}$ is an ideal, an element u of $\mathscr{A}$ is called a *unit mod* $\mathscr{J}$ in case

$$au - a \in \mathscr{J} \quad \text{and} \quad ua - a \in \mathscr{J} \quad \text{for all} \quad a \in \mathscr{A} \ . \tag{1.21}$$

An ideal $\mathscr{J}$ is called a *modular ideal* in case there exists a unit mod $\mathscr{J}$.

Evidently if $\mathscr{J}$ is an ideal in $\mathscr{A}$, and $\overline{\mathscr{J}}$ is the norm closure of $\mathscr{J}$, then $\overline{\mathscr{J}}$ is also an ideal in $\mathscr{A}$.

Given a Banach algebra $\mathscr{A}$ and an ideal $\mathscr{J}$, there is a natural equivalence relation $\sim$ on $\mathscr{A}$ given by

$$a \sim b \iff a - b \in \mathscr{J} \ .$$

Let $\{a\}$ and $\{b\}$ denote the equivalence classes of a and b respectively. Let $\tilde{a}$ and $\tilde{b}$ be any other representative of $\{a\}$ and $\{b\}$ respectively. Then for some $x, y \in \mathscr{J}$, $\tilde{a} = a + x$ and $\tilde{b} = b + y$. Then

$$\tilde{a}\tilde{b} = (a + x)(b + y) = ab + (ay + xb + xy) \sim ab \ .$$

Even more simply one sees that $\tilde{a} + \tilde{b} \sim a + b$ and for all $\lambda \in \mathbb{C}$, $\lambda\tilde{a} \sim \lambda a$. Hence $\mathscr{A}/\mathscr{J}$, the set of equivalence classes in $\mathscr{A}$, equipped with the operations

$$\{a\}\{b\} = \{ab\} \quad \text{and} \quad \{a\} + \{b\} = \{a + b\} \quad \text{and} \quad \lambda\{a\} = \{\lambda a\}$$

is an algebra, and $a \mapsto \{a\}$ is a homomorphism of $\mathscr{A}$ onto $\mathscr{A}/\mathscr{J}$.

Now introduce a norm on $\mathscr{A}/\mathscr{J}$ by

$$\|\{a\}\| = \inf\{\ \|\tilde{a}\|\ :\ \tilde{a} \sim a\ \} = \inf\{\ \|a - b\|\ :\ b \in \mathscr{J}\ \} \ .$$

Note that $\|\{a\}\| \le \|a\|$. To see that

$$\|\{a\}\{b\}\| \le \|\{a\}\|\|\{b\}\| \tag{1.22}$$

for all $\{a\}, \{b\} \in \mathscr{A}/\mathscr{J}$, let $0 < \epsilon < \min\{\|\{a\}\|, \|\{b\}\|\}$, and pick $\tilde{a} \in \{a\}$ and $\tilde{b} \in \{b\}$ so that $\{a\} > \|\tilde{a}\| - \epsilon$ and $\{b\} > \|\tilde{b}\| - \epsilon$. Then

$$\|\{a\}\{b\}\| = \|\{\tilde{a}\tilde{b}\}\| \le \|\tilde{a}\tilde{b}\| \le \|\tilde{a}\|\|\tilde{b}\| \le (\|\{a\}\| + \epsilon)(\|\{b\}\| + \epsilon) \ .$$

Since ϵ can be taken arbitrarily small, (1.22) is proved.

Therefore, $\mathscr{A}/\mathscr{J}$ will be a Banach algebra with this norm provided it is complete in this norm. Consider a Cauchy sequence $\{\{a\}_n\}_{n\in\mathbb{N}}$ in $\mathscr{A}/\mathscr{J}$. A standard argument shows that this sequence always has a limit if $\mathscr{J}$ is closed. Thus, when $\mathscr{J}$ is a closed ideal, $\mathscr{A}/\mathscr{J}$ is a Banach algebra, and the map $a \mapsto \{a\}$ is a contractive homomorphism of $\mathscr{A}$ onto $\mathscr{A}/\mathscr{J}$. This homomorphism is called the *natural homomorphism of* $\mathscr{A}$ *onto* $\mathscr{A}/\mathscr{J}$.

It is possible for $\mathscr{A}/\mathscr{J}$ to have an identity evan when $\mathscr{A}$ does not. Suppose that $\mathscr{J}$ is modular, and that u is a unit mod $\mathscr{J}$. Then for all $a \in \mathscr{A}$, $\{u\}\{a\} = \{ua\} = \{a\}$ and $\{a\}\{u\} = \{au\} = \{a\}$. Thus, $\{u\}$ is a multiplicative identity in $\mathscr{A}/\mathscr{J}$. Clearly if $\mathscr{A}$ has an identity 1, 1 is a unit mod $\mathscr{J}$.

There is a close connections between closed ideals and kernels of continuous homomorphisms of Banach algebras.

1.30 PROPOSITION. *Let $\mathscr{A}$ and $\mathscr{B}$ be Banach algebras, and let $\pi : \mathscr{A} \to \mathscr{B}$ be a continuous homomorphism. Then $\mathscr{J} = \ker(\mathscr{A})$ is a closed ideal in $\mathscr{A}$. Conversely, if $\mathscr{J}$ is a closed ideal in $\mathscr{A}$, then the map $a \mapsto \{a\}_{\mathscr{J}}$, sending a to its equivalence class mod $\mathscr{J}$, is a homomorphism of $\mathscr{A}$ onto $\mathscr{A}/\mathscr{J}$.*

Proof. Suppose that $\pi : \mathscr{A} \to \mathscr{B}$ be a continuous homomorphism. Then evidently $\mathscr{J} = \ker(\mathscr{A})$ is a closed by the continuity of π, and it is a subspace by the linearity of π. Next, for all $x \in \mathscr{J}$ and $a, b \in \mathscr{A}$, $\pi(axb) = \pi(a)\pi(x)\pi(b) = \pi(a)0\pi(b) = 0$. Hence $axb \in \ker(\pi)$, and so $\mathscr{J}$ is an ideal. The converse is clear from the construction of $\mathscr{A}/\mathscr{J}$ described above. □

Now consider a *commutative* Banach algebra $\mathscr{A}$. For any $x_0 \in \mathscr{A}$, we can define $\mathscr{J}(x_0)$ to be the subset of $\mathscr{A}$ given by

$$\mathscr{J}(x_0) = \{\ x_0 y\ :\ y \in \mathscr{A}\ \}\ . \tag{1.23}$$

Then for all $yx_0 \in \mathscr{J}(x_0)$ and all $a, b \in \mathscr{A}$, $ayx_0b = (ayb)x_0 \in \mathscr{J}(x_0)$, and evidently $\mathscr{A}$ is a subspace of $\mathscr{A}$. Hence $\mathscr{J}(x_0)$ is an ideal, and it is called *the ideal generated by* x_0.

In the non-commutative setting, one could consider the set $\{\ yx_0z\ :\ y, z \in \mathscr{A}\ \}$ which is closed under left and right multiplication by elements of $\mathscr{A}$. However, without some additional hypothesis on x_0, such as that x_0 commutes with all elements of $\mathscr{A}$, it need not be a subspace, and the clousre of its span might be all of $\mathscr{A}$.

Let $\mathscr{A}$ be a commutative Banach algebra with an identity 1. Let x_0 be a non-invertible element of $\mathscr{A}$. Let $\mathscr{J}(x_0)$ be the ideal generated by x_0. Then no element of $\mathscr{J}(x_0)$ is invertible. Indeed, if x_0y were invertible, there would exist $z \in \mathscr{A}$ such that $(x_0y)z = x_0(yz) = 1$, and then (since $\mathscr{A}$ is commutative), yz would be an inverse of x_0, which is not possible. Hence, for all non-invertible x_0, $\mathscr{J}(x_0)$ consists entirely of non-invertible elements. Since the open unit ball about the identity consists of invertible elements, $\mathscr{J}(x_0)$ does not intersect the open unit ball about the identity 1. In particular, 1 does not belong to $\overline{\mathscr{J}(x_0)}$, the closure of $\mathscr{J}(x_0)$.

Now we come to a crucial construction of characters in a commutative Banach algebra:

1.31 THEOREM. *Let $\mathscr{A}$ be a commutative Banach algebra with identity* 1*. Then for all non-invertible $x_0 \in \mathscr{A}$, there exists a character $\varphi \in \Delta(\mathscr{A})$ such that $\varphi(x_0) = 0$.*

Proof. Since x_0 is not invertible, $\mathscr{J}(x_0)$ is a proper ideal in $\mathscr{A}$, and in fact, as explained above, the open unit ball about 1 does not intersect $\mathscr{J}(x_0)$. Now consider any chain of proper ideals in $\mathscr{A}$, ordered by inclusion. Since no proper ideal contains the identity, the union of this chain is again a proper ideal. Hence by Zorn's Lemma, there exists a maximal proper ideal $\mathscr{M}$ containing $\mathscr{J}(x_0)$. Since no proper ideal can contain any invertible elements, this ideal does not intersect the open unit ball about 1. Hence its closure $\overline{\mathscr{M}}$ also contains $\mathscr{J}(x_0)$ and is proper. Since $\mathscr{M}$ is maximal among such ideals, $\mathscr{M} = \overline{\mathscr{M}}$. Hence in a commutative Banach algebra $\mathscr{A}$ with identity 1, for each non-invertible $x_0 \in \mathscr{A}$, there exists a closed proper ideal $\mathscr{M}$ that contains any ideal in $\mathscr{A}$ that contains $\mathscr{J}(x_0)$.

We now claim that the Banach algebra $\mathscr{B} = \mathscr{A}/\mathscr{M}$ is a division algebra. Suppose not. Then it contains a non-zero, non-invertible element $\{y_0\}_{\mathscr{M}}$. Let $\mathscr{N}$ be the closure of the ideal in $\mathscr{B}$ generated by $\{y_0\}_{\mathscr{M}}$. Let π_1 be the natural homomorphism of $\mathscr{A}$ onto $\mathscr{B}$, and let π_2 be the natural homomorphism of $\mathscr{B}$ onto $\mathscr{B}/\mathscr{N}$. Then $\pi_2 \circ \pi_1$ is a homomorphism of $\mathscr{A}$ onto $\mathscr{B}/\mathscr{N}$. By Proposition 1.30, $\ker(\pi_2 \circ \pi_1)$ is a closed ideal that contains $\mathscr{M} = \ker(\pi_1)$. The containment is proper since $\pi_2 \circ \pi_1(y_0) = 0$, but $y_0 \notin \mathscr{M}$ since $\{y_0\}_{\mathscr{M}} \neq 0$. Finally, $1 \notin \ker(\pi_1 \circ \pi_1)$ since $\{1\}_{\mathscr{M}}$ is a unit $\mathscr{B} = \mathscr{A}/\mathscr{M}$, and $\mathscr{N}$ does not contain any invertible elements, so $\pi_2(\{1\}_{\mathscr{M}}) = \pi_2(\pi_1(1)) \neq 0$. Thus, $\ker(\pi_2 \circ \pi_1)$ is a closed proper ideal that strictly contains $\mathscr{M}$, which is impossible. Hence the hypothesis that $\mathscr{B} = \mathscr{A}/\mathscr{M}$ contains a non-zero, non-invertible element is false. This shows that $\mathscr{B} = \mathscr{A}/\mathscr{M}$ is a division algebra, and then the Gelfand-Mazur Theorem tells us that $\mathscr{B} = \mathscr{A}/\mathscr{M}$ is canonically isomorphic to $\mathbb{C}$. Hence π_1 may be regarded as a character of $\mathscr{A}$, and by construction $x_0 \in \mathscr{J}(x_0) \subset \mathscr{M} = \ker(\pi_1)$. □

This theorem has the following important consequence:

1.32 COROLLARY. *Let $a \in \mathscr{A}$, where $\mathscr{A}$ is a commutative Banach algebra. Let $\lambda \in \sigma_{\mathscr{A}}(a)$. Then there exists $\varphi \in \Delta(\mathscr{A})$ such that $\varphi(a) = \lambda$. In particular, the spectral radius $\nu(a)$ of a is given by*

$$\nu(a) = \sup\{\ |\varphi(a)|\ :\ \varphi \in \Delta(\mathscr{A})\ \}\ . \tag{1.24}$$

Proof. Adjoining an identity if need be has no effect on the spectral radius, so we may assume that $\mathscr{A}$ has an identity 1. We have already seen that for all $\varphi \in \Delta(\mathscr{A})$, $\varphi(a) \in \sigma_{\mathscr{A}}(a)$. We now show that for *every* $\lambda \in \sigma_{\mathscr{A}}(a)$, there exists $\varphi \in \Delta(\mathscr{A})$ with $\varphi(a) = \lambda$.

Since $\lambda 1 - a$ is not invertible, by Theorem 1.31, there exists $\varphi \in \Delta(\mathscr{A})$ such that $\varphi(\lambda 1 - a) = 0$. But $\varphi(\lambda 1 - a) = \lambda\varphi(1) - \varphi(a) = \lambda - \varphi(a)$. □

2 The Spectral Theorem for C^* Algebras

Let $\mathscr{A}$ be a C^* algebra. The involution $*$ allows us to to define certain classes of elements in $\mathscr{A}$:

2.1 DEFINITION (Self-adjoint, normal and unitary elements of a C^* algebra)**.** Let $\mathscr{A}$ be a C^* algebra. Then:

(i) $a \in \mathscr{A}$ is *self-adjoint* in case $a = a^*$.

(ii) $a \in \mathscr{A}$ is *normal* in case $aa^* = a^*a$.

(iii) In case $\mathscr{A}$ has an identity a, $a \in \mathscr{A}$ is *unitary* in case $a^*a = a^*a = 1$.

This definition generalizes these notions from the basic example in which $\mathscr{A}$ is the algebra of $n \times n$ matrices or the bounded operators on some Hilbert space $\mathscr{H}$.

2.2 LEMMA. *Let $\mathscr{A}$ be a C^* algebra with an identity* 1*. Then* 1 *is self adjoint and* $\|1\| = 1$*. Moreover, for any unitary $u \in \mathscr{A}$, $\|u\| = 1$.*

Proof. $1^* = 1^*1$. Applying the involution $1^* = 1^*1$, showing that 1 is self adjoint. The next two parts use the strong condition on the norm in a C^* algebra, which is that for all $a \in \mathscr{A}$, $\|a^*a\| = \|a\|^2$. We use this first in

$$\|1\| = \|1^2\| = \|1^*1\| = \|1\|^2\ ,$$

where the second equality is true since $1 = 1^*$. Thus $\|1\| = 1$ or $\|1\| = 0$. The latter is impossible. Finally, if u is unitary, $1 = \|1\| = \|u^*u\| = \|u\|^2$, so that $\|u\| = 1$. □

2.3 THEOREM (In a C^* algebra, self-adjointness implies real spectrum)**.** *Let $\mathscr{A}$ be a C^* algebra, and let $a \in \mathscr{A}$ then if $a = a^*$, $\sigma_{\mathscr{A}}(a) \subset (\mathbb{R})$.*

Proof. It is no loss of generality to assume that $\mathscr{A}$ has an identity since we may adjoin one if need be with out any effect on the spectrum apart from possibly adjoining 0 to it. Therefore, suppose that $\mathscr{A}$ has an identity but contains some self adjoint element a with some $\lambda \in \sigma_{\mathscr{A}}(a)$ such that $\lambda \notin \mathbb{R}$. Then taking a appropriate real multiple of a (so the the multiple is still self adjoint), we may suppose that $e^{i\lambda} = 2$ for some $\lambda \in \sigma_{\mathscr{A}}(a)$.

For each $n \in \mathbb{N}$, define the polynomial $p_n(\zeta) = \sum_{j=0}^{n}(i\zeta)^j/j!$. By the Spectral Mapping Lemma, for each n, $p_n(\lambda) \in p_n(a)$. For $n > m$,

$$\|p_n(a) - p_m(a)\| \le \sum_{j=m+1}^{n} \|a\|^j/j! \ ,$$

and hence by standard estimates on the exponential power series for numbers, $\{p_n(a)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathscr{A}$. Therefore, there exist $u \in \mathscr{A}$ with $u = \lim_{n\to\infty} p_n(a)$. Evidently, for all n, $(p_n(a))^* = p_n(-a)$, so that once again, by standard estimates for the exponential power series, $u^*u = uu^* = 1$; that is, u is unitary, and by Lemma 2.2, $\|u\| = 1$. Therefore, for all μ with $|\mu| > 1$, $\mu 1 - u$ is invertible. But

$$\lim_{n\to\infty}(p_n(\lambda)1 - p_n(a)) = e^{i\lambda}1 - u \ .$$

Since $|e^{i\lambda}1| = 2$, $e^{i\lambda}1 - u$ is invertible. However, $p_n(\lambda)1 - p_n(a)$ is non-invertible for each n. Since the set of invertible elements is open, it cannot be that a sequence of non-invertible elements converges to an invertible element. Thus, the hypothesis that $\mathscr{A}$ contains some self adjoint element a with some $\lambda \in \sigma_{\mathscr{A}}(a)$ such that $\lambda \notin \mathbb{R}$ leads to contradiction. □

2.4 DEFINITION (Hermitian character)**.** Let $\mathscr{A}$ be a C^* algebra. A character φ of $\mathscr{A}$ is *Hermitian* in case for all $a \in \mathscr{A}$,

$$\varphi(a^*) = (\varphi(a))^* \ .$$

2.5 LEMMA. *All characters of a C^* algebra are Hermitian.*

Proof. For any $a \in \mathscr{A}$ define $x = \dfrac{1}{2}(a + a^*)$ and $y = \dfrac{1}{2i}(a - a^*)$. Then x and y are self-adjoint, and $a = x + iy$. For any character φ of $\mathscr{A}$,

$$\varphi(a) = \varphi(x + iy) = \varphi(x) + i\varphi(y) \quad \text{and} \quad \varphi(a^*) = \varphi(x - iy) = \varphi(x) - i\varphi(y) \ .$$

By Theorem 2.3, $\varphi(x)$ and $\varphi(y)$ are real, and hence $\varphi(a^*) = (\varphi(a))^*$. □

The next theorem again makes use of the strong condition on the norm in a C^* algebra, which is that for all $a \in \mathscr{A}$, $\|a^*a\| = \|a\|^2$.

2.6 THEOREM (Norm and spectral radius in a C^* algebra)**.** *Let $\mathscr{A}$ be a C^* algebra. Then for all $a \in \mathscr{A}$,*

$$\|a\|^2 = \nu(a^*a) \tag{2.1}$$

and if a is normal,

$$\|a\| = \nu(a) \ . \tag{2.2}$$

Proof. Suppose first that a is normal. Then $(a^*a)*(a^*a) = a^*aa^*a = (a^2)^*(a^2)$. Then by the C^*-algebra identitiy $\|b^*b\| = \|b\|^2$ applied twice, and the isometry propert of the involuton,

$$\|a\|^2 = \|a\|^2\|a^*\|^2 = \|a^*a\|^2 = \|(a^*a)*(a^*a)\| = \|(a^2)^*(a^2)\| = \|a^2\|^2 \ .$$

Therefore, $\|a^2\| = \|a\|^2$, and by an obvious induction, for all $m \in \mathbb{N}$, $\|a^{2m}\| = \|a\|^{2m}$. Then by Gelfand's Formula,

$$\nu(a) = \lim_{m\to\infty}(\|a^{2m}\|)^{1/2m} = \|a\|.$$

This proves (2.1). Next, for any $a \in \mathscr{A}$, a^*a is self adjoint and so $\nu(a^*a) = \|a^*a\|$. Then since $\mathscr{A}$ is a C^* algebra, $\|a^*a\| = \|a\|^2$, and this proves (2.2).

□

2.7 THEOREM (Commutative Gelfand-Naimark Theorem)**.** *Let $\mathscr{A}$ be a commutative C^*-algebra. Then the Gelfand transform is an isometric isomorphism of $\mathscr{A}$ onto $\mathscr{C}_0(\Delta(\mathscr{A}))$.*

Proof. By Lemma 2.5, for all $a \in \mathscr{A}$ and all $\varphi \in \Delta(\mathscr{A})$,

$$\gamma(a^*)[\varphi] = \varphi(a^*) = (\varphi(a))^* = \gamma(a)^*[\varphi]$$

since the involution in $\mathscr{C}_0(\Delta(\mathscr{A}))$ is pointwise complex conjugation.

Next, $|\gamma(a)[\varphi]| = |\varphi(a)|$. By the easy Lemma 1.21, $\sup_{\varphi\in\Delta(\mathscr{A})}\{\ |\varphi(a)|\ \} \leq \nu(a)$. By the deeper Corollary 1.32 of Theorem 1.31, $\sup_{\varphi\in\Delta(\mathscr{A})}\{\ |\varphi(a)|\ \} \geq \nu(a)$. Combining these two inequalities with Theorem 2.6, and noting that in a commutative C^* algebra, *every* element is normal,

$$\sup_{\varphi\in\Delta(\mathscr{A})}\{\ |\gamma(a)[\varphi]|\ \} = \nu(a) = \|a\| \ ,$$

which proves that the Gelfand transform is an isometry, and hence is injective onto a subalgebra of $\gamma(\mathscr{A})$ of $\mathscr{C}_0(\Delta(\mathscr{A}))$. However, $\gamma(\mathscr{A})$ separates points, and does not vanish at any $\varphi \in \Delta(\mathscr{A})$, and is closed under complex conjugation. Hence by the Stone-Wierstrass Theorem, and the closure of $\gamma(\mathscr{A})$, $\gamma(\mathscr{A}) = \mathscr{C}_0(\Delta(\mathscr{A}))$. □

2.1 Spectral invariance and the Abstract Spectral Theorem

Let $\mathscr{A}$ be a Banach algebra with identity, and let $\mathscr{B}$ be a Banach subalgebra. It can happen that some $b \in \mathscr{B}$ is not invertible in $\mathscr{B}$, but is invertible in $\mathscr{A}$.

2.8 EXAMPLE. Let D denote the closed unit disc in $\mathbb{C}$, and let C denote its boundary, the unit circle. Let $\mathscr{A} = \mathscr{C}(C)$, the algebra of continuous functions on C. Let $\mathscr{B}$ denote the algebra of continuous functions on D that are holomorphic in the interior of D. These functions are determined by their values on C, and their maximum absolute value is attained on C. Therefore, restriction to C is an isometric embedding of $\mathscr{B}$ in $\mathscr{A}$, so we may regard $\mathscr{B}$ as a subalgebra of $\mathscr{A}$.

Let b denote the function $f(e^{i\theta}) = e^{i\theta}$, the identity function on C, which evidently belongs to $\mathscr{B}$. Then $1\lambda - b$ is invertible in $\mathscr{A}$ if and only if $\lambda \notin C$, in which case the inverse is the function $g(e^{i\theta}) = (\lambda - e^{i\theta})^{-1}$. However, for λ in the interior of D, $\zeta \mapsto (\lambda - \zeta)^{-1}$ is not holomorphic in the interior of D, and so the inverse of b in $\mathscr{A}$ does not belong to $\mathscr{B}$. That is $\sigma_{\mathscr{A}}(b) = C$, but $\sigma_{\mathscr{B}}(b) = D$.

Now we specialize to C^* algebras, first introducing certain minimal subalgebras:

2.9 DEFINITION. Let $\mathscr{A}$ be a C^* algebra with unit 1. For all $b \in \mathscr{A}$, $C(b)$ is the smallest C^* subalgebra of $\mathscr{A}$ that contains b and 1.

2.10 THEOREM. *Let $\mathscr{A}$ be a C^* algebra with unit 1, and let $b \in \mathscr{A}$. if b is invertible in $\mathscr{A}$, $b^{-1} \in C(b)$, and hence b is invertible in every C^* subalgebra of $\mathscr{A}$ that contians 1 and b. In particular, if $\mathscr{B}$ is a C^* subalgebra of $\mathscr{A}$ that contains 1 and b, then*

$$\sigma_{\mathscr{B}}(b) = \sigma_{\mathscr{A}}(b) \ .$$

Proof. Suppose first that b is self adjoint and invertible in $\mathscr{A}$. By Theorem 2.3, $\sigma_{C(b)}(b) \subset \mathbb{R}$, and consequently, for all $n \in \mathbb{N}$, $b-(i/n)1$ is invertible in $C(b)$. Since $\lim_{n\to\infty}(b-(i/n)1) = b$ in $\mathscr{A}$, and since the inverse is continuous, $\lim_{n\to\infty}(b-(i/n)1)^{-1} = b^{-1}$ in $\mathscr{A}$. But since $(b-(i/n)1)^{-1} \in C(b)$ for all n, and since $C(b)$ is closed,

$$b^{-1} = \lim_{n\to\infty}(b-(i/n)1)^{-1} \in C(b) \ .$$

Hence, b is invertible within $C(b)$.

Now let b be any invertible element of $\mathscr{A}$. Then b^* and b^*b are invertible in $\mathscr{A}$, and also belong to $C(b)$. Since b^*b is self adjoint, what we have proved above says that $(b^*b)^{-1} \in C(b)$. Define $x = (b^*b)^{-1}b^* \in C(b)$. Evidently $xb = 1$. Thus, b has a left inverse in $C(b)$. The same argument shows that $y = b^*(bb^*)^{-1}$ is a well defined right inverse of b in $C(b)$, and then $x = x(by) = (xb)y = y$ so $x = y$ is the inverse of b in $C(b)$. In particular, for all $\lambda \in C$, $\lambda 1 = b$ is invertible in $C(b)$ if and only if it is invertible in $\mathscr{A}$. Thus, $\lambda 1 - b$ is invertible in $\mathscr{A}$ if and only if it is invertible in $C(b)$, and this proves the final statement. □

2.11 LEMMA. *Let $\mathscr{A}$ be a C^* algebra with identity 1, and let $a \in \mathscr{A}$ be normal. Then the map $\varphi \mapsto \varphi(a)$ is a homeomorphism of $\Delta(C(a))$ onto $\sigma_{\mathscr{A}}(a)$.*

Proof. Since a and a^* commute, the closure of the linear span of $\{\ a^m(a^*)^n \ : m,n \geq 0\ \}$ is a C^* algebra that contains 1 and a. Evidently, it is $C(a)$. Hence if $\varphi \in \Delta(C(a))$, φ is determined by its values on a and a^*. In fact, since φ is necessarily Hermitian, φ is determined by its value on a. That is, for any $\varphi, \psi \in \Delta(C(a))$,

$$\varphi = \psi \qquad \Longleftrightarrow \qquad \varphi(a) = \psi(a) \ .$$

We have also seen that for all $\varphi \in \Delta(C(a))$, $\varphi(a) \in \sigma_{C(a)}(a) = \sigma_{\mathscr{A}}(a)$, and for all $\lambda \in \sigma_{\mathscr{A}}(a) = \sigma_{C(a)}(a)$, there is a $\varphi_\lambda \in \Delta(C(a))$ such that $\varphi_\lambda(a) = \lambda$. This shows that the map $\varphi \mapsto \varphi(a)$ is a one-to-one map of $\Delta(C(a))$ onto $\sigma_{\mathscr{A}}(a)$. This map is also continuous by the definition of the Gelfand topology, and continuous bijections between compact spaces are homeomorphisms. □

We now come to the Abstract Spectral Theorem:

2.12 THEOREM (Abstract Spectral Theorem)**.** *Let $\mathscr{A}$ be a C^* algebra with identity 1, and let $a \in \mathscr{A}$ be normal. Then identifying $\mathscr{C}_{\Delta(C(a))}$ and $\mathscr{C}(\sigma_{\mathscr{A}}(a))$ through the homeomorphism provided by Lemma 2.11, we may regard the Gelfand transform as a homomorphism of $C(a)$ into $\mathscr{C}(\sigma_{\mathscr{A}}(a))$. Then the Gelfand transform γ is an isometric isomorphism of $C(a)$ onto $\mathscr{C}(\sigma_{\mathscr{A}}(a))$. For all non-negative integers m,n, $\gamma(a^m(a^*)^n)$ is the function on $\sigma_{\mathscr{A}}(a)$ given by*

$$\lambda \mapsto \lambda^m(\lambda^*)^n \ . \tag{2.3}$$

Proof. The Commutative Gelfand-Naimark Theorem says that γ is an isometric isomorphism, and if $\varphi \in \Delta(C(a))$,

$$\gamma(a^m(a^*)^m)[\varphi] = \varphi(a)^m((\varphi(a))^*)^n = \lambda^m(\lambda^*)^n$$

for $\lambda = \varphi(a)$ so that under the identification provided by Lemma 2.11, $\gamma(a^m(a^*)^m)$ is indeed given by (2.3). □

2.13 DEFINITION. For $\mathscr{A}$ a C^* algebra with identity 1, a a normal element of $\mathscr{A}$, and $f \in \mathscr{C}(\sigma_{\mathscr{A}}(a))$, $f(a)$ is defined by $\gamma^{-1}(f)$; i.e., $f(a)$ is the inverse image of f under the isometric isomorphism of $C(a)$ onto $\mathscr{C}(\sigma_{\mathscr{A}}(a))$ that is provided by the Commutative Gelfand Naimark Theorem.

2.14 THEOREM (Spectral Mapping Theorem)**.** *Let $\mathscr{A}$ be a C^* algebra with identity* 1*, a a normal element of $\mathscr{A}$, and $f \in \mathscr{C}(\sigma_{\mathscr{A}}(a))$. Then*

$$\sigma_{\mathscr{A}}(f(a)) = f(\sigma_{\mathscr{A}}(a)) \ .$$

Proof. For all $\mu \in \mathbb{C}$, the function $\lambda \mapsto \mu - f(\lambda)$ is invertible in $\mathscr{C}(\sigma_{\mathscr{A}}(a))$ if and only if μ does not belong to the range of f, which is $f(\sigma_{\mathscr{A}}(a))$. Then, using the isomorphism provided by the Commutative Gelfand Naimark Theorem, we see that $\mu 1 - f(a)$ is invertible in $C(a)$ if and only if $\mu \notin f(\sigma_{\mathscr{A}}(a))$, and hence $\sigma_{C(a)}(f(a)) = \sigma_{\mathscr{A}}(a)$. Finally, by Theorem 2.10, the spectrum of $f(a)$ is the same in all C^* subalgebras of $\mathscr{A}$ that contain $f(a)$ and 1. In particular, $\sigma_{C(a)}(f(a)) = \sigma_{cA}(f(a))$. □

2.2 Continuity of the spectrum and the functional calculus

2.15 THEOREM (Newburgh's Theorem)**.** *Let $\mathscr{A}$ be a Banach algebra and $a \in \mathscr{A}$. Let U be an open subset of $\mathbb{C}$ with $\sigma_{\mathscr{A}}(a) \subset U$. Then there exists a $\delta > 0$ such that if $\|b - a\| \leq \delta$,*

$$\sigma_{\mathscr{A}}(b) \subset U \ .$$

Proof. First note that for all $b \in \mathscr{A}$ with $\|b - a\| < 1$, $\|b\| < \|a\| + 1$, and hance for all $\lambda \in \mathbb{C}$ with $\lambda \geq \|a\| + 1$, $\lambda \in \rho_{\mathscr{A}}(b)$. Hence when $\|b - a\| \leq 1$, $\sigma_{\mathscr{A}}(b)$ is contained in the closed centered disc of radius $\|a\| + 1$.

Let $K = U^c \cap \{ \ \lambda \ : \ |\lambda| \leq \|a\| + 1 \ \}$ which is a compact subset of $\rho_{\mathscr{A}}(a)$. It suffices to show that there is an $r > 0$ so that for all $\mu \in K$, $(\mu 1 - b)$ is invertible whenever $\|b - a\| < r$.

Let $\lambda \in K$. Then $\lambda \in \rho_{\mathscr{A}}(a)$, and for all $b \in \mathscr{A}$ and $\mu \in \mathbb{C}$ with

$$|\mu - \lambda| + \|b - a\| < \|(\lambda 1 - a)^{-1}\|^{-1} \quad \Rightarrow \quad \|(\mu 1 - b) - (\lambda 1 - a)\| \leq \|(\lambda 1 - a)^{-1}\|^{-1} \ ,$$

and hence $\mu \in \rho_{\mathscr{A}}(b)$. For each $\lambda \in K$, define $U_\lambda = \{\mu \ : |\mu - \lambda| < \frac{1}{2}\|(\lambda 1 - a)^{-1}\|^{-1} \ \}$. Since K is compact, there exists a finite subcover $\{U_{\lambda_1}, \ldots, U_{\lambda_n}\}$. Define

$$r = \min\{\frac{1}{2}\|(\lambda_1 1 - a)^{-1}\|^{-1} \ , \ldots , \ \frac{1}{2}\|(\lambda_n 1 - a)^{-1}\|^{-1} \ \} \ .$$

Then for any b with $\|b - a\| < r$ and any $\mu \in K$, $\mu \in U_{\lambda_j}$ for some $j = 1, \ldots, n$, and then

$$\|(\mu 1 - b) - (\lambda_j 1 - a)\| \leq |\mu - \lambda_j| + \|b - a\| < \|(\lambda_j 1 - a)^{-1}\|^{-1} \ .$$

Therefore, $(1\mu - b)$ is invertible. Thus, for all $\mu \in K$, whenever $\|b - a\| < r$, $\mu \in \rho_{\mathscr{A}}(b)$. □

Now $\mathscr{A}$ be an Banach algebra, and let $x, y \in \mathscr{A}$, and $t \in \mathbb{R}$. Then for all $n \in \mathbb{N}$, by the telescoping sum identitiy

$$(x+y)^n - x^n = \sum_{j=0}^{n-1}((x+y)^{n-j}x^j - (x+y)^{n-j-1}x^{j+1}$$
$$= \sum_{j=0}^{n-1}(x+y)^{n-j-1}yx^j$$

Therefore,

$$\|(x+y)^n - x^n\| \le \left(\sum_{j=0}^{n-1}\|(x+y)^{n-j-1}\|\|x^j\|\right)\|y\| \le n(\|x\|+\|y\|)^n\|y\| \ .$$

Therfore, when $\|y\| < \delta$, $\|(x+y)^n - x^n\| < n(\|x\|+\delta)^n\delta$, and this proves that $x \mapsto x^n$ is continuous. In a C^* algebra we can say more.

2.16 THEOREM. *Let $\mathscr{A}$ be a C^* algebra with identity 1. Let $U \subset \mathbb{C}$ be open with $\overline{U}$ compact. Let $\mathscr{N}_U$ be given by*

$$\mathscr{N}_U = \{\ a \in \mathscr{A}\ \ : aa^* = a^*a \quad \text{and} \quad \sigma_{\mathscr{A}}(a) \subset U\ \}\ . \tag{2.4}$$

Then $\mathscr{N}_U$ is an open subset of the normal elements of $\mathscr{A}$. Moreover, let f be a continuous complex vauled function on $\overline{U}$, and for all $a \in \mathscr{N}_U$, define $f(a) \in \mathscr{A}$ using the Gelfand-Naimark isomorphism. Then the map $a \mapsto f(a)$ is continous on $\mathscr{N}_U$.

Proof. The first assertion is an immediate consequence of Newburg's Theorem. For the second, consider any sequence $\{p_n\}$ of polynomials converging uniformly to f on $\overline{U}$. Then for all $a \in \mathscr{N}_U$,

$$\|p_n(a) = f(a)\| \le \sup_{\lambda\in\overline{U}}\{p_n(\lambda) - f(\lambda)|\}.$$

That is,

$$\lim_{n\to\infty}\left(\sup_{a\in\mathscr{N}_U}\{\ \|p_n(a) - f(a)\|\ \}\right) = 0\ .$$

Thus, the function $a \mapsto f(a)$ is the *uniform* limit of the continuous functions $a \mapsto p_n(a)$, □

For normal elements of a C^* algebra, there is a quantitative version of Newburgh's Theorem.

2.17 THEOREM. *Let $\mathscr{A}$ be a C^* algebra, and let $a, x \in \mathscr{A}$ be normal. Then*

$$\sigma_{\mathscr{A}}(a+x) \subset \{\ \lambda\ :\ \mathrm{dist}(\lambda, \sigma_{\mathscr{A}}(a)) \le \|x\|\ \}\ .$$

Proof. Let $\lambda \in \rho_{\mathscr{A}}(a)$. By the Gelfand-Naimark isomorphism,

$$\|(\lambda 1 - a)^{-1}\| = \sup\{|\lambda - \mu|^{-1}\ :\ \mu \in \sigma_{\mathscr{A}}(a)\ \} = \frac{1}{\mathrm{dist}(\lambda, \sigma_{\mathscr{A}}(a))}\ .$$

Let $x \in \mathscr{A}$ be normal, Then $[\lambda 1 - (a+x)] - [\lambda 1 - a] = -x$. Therefore, as long as $\|x\| < \mathrm{dist}(\lambda, \sigma_{\mathscr{A}}(a))$, $[\lambda 1 - (a+x)]$ is invertible □

2.3 Positivity in C^* algebras

2.18 DEFINITION. Let $\mathscr{A}$ be a C^* algebra. Then a self adjoint element a in $\mathscr{A}$ is *positive* in case $\sigma_{\mathscr{A}}(a) \subset [0,\infty)$. The set of all positive elements of $\mathscr{A}$ is denoted $\mathscr{A}^+$.

If $a \in \mathscr{A}^+$, we may use the Abstract Spectral Theorem to define $\sqrt{a}$, and then $a = (\sqrt{a})^2 = (\sqrt{a})^*(\sqrt{a})$. It is also true that in any C^* algebra, every element of the form b^*b is positive. This was not known to Gelfand and Naimark when they wrote their 1943 paper, in which they raised the question as to whether it was true or not. They included an extra hypothesis in their paper, namely that for all b in $\mathscr{A}$, $1 \notin \sigma_{\mathscr{A}}(b^*b)$.

The fact that for all b in a $\mathbb{C}^*$ algebra $\mathscr{A}$, $b^*b \in \mathscr{A}^+$ was finally proved in 1952 and 1953 through the contributions of Fukamiya and Kaplansky. The history is interesting: Kaplansky had managed to prove that *if* the sum of two positive elements is necessarily positive, then b^*b is necessarily positive. However, he was unable to show that $\mathscr{A}^+$ was closed under sum. He published nothing, but showed his proof to many people. When Fukamiya proved the closure of $\mathscr{A}^+$ in 1952, Kaplansky communicated his proof to the reviewer of Fukamiya's paper for Math Reviews, and the proof was published there.

2.19 THEOREM (Fukamiya's Theorem)**.** *Let $\mathscr{A}$ be a C^* algebra. Then $\mathscr{A}^+$ is a pointed convex cone. That is:*

(1) For all $\lambda \in \mathbb{R}^+$, and all $a \in \mathscr{A}^+$, $\lambda a \in \mathscr{A}^+$, and for all $a, b \in \mathscr{A}^+$, $a + b \in \mathscr{A}^+$.

(2) $-\mathscr{A}^+ \cap \mathscr{A}^+ = \{0\}$.

(The first part says that $\mathscr{A}^+$ is a convex cone; the second part says that this cone is pointed.)

Proof. We may suppose that $\mathscr{A}$ has an identity 1 since otherwise we may adjoin an identity without affecting positivity.

Let $B_{\mathscr{A}}$ denote the closed unit ball in $\mathscr{A}$. Fukamiya observed that that $\mathscr{A}^+ \cap B_{\mathscr{A}}$ consists precisely of the self-adjoint elements a with both a and $1-a$ are in $B_{\mathscr{A}}$. To see this suppose that $a \in \mathscr{A}^+ \cap B_{\mathscr{A}}$. Then since a is self adjoint, $\nu(a) = \|a\| \leq 1$, and so $\sigma_{\mathscr{A}}(a) \subset [0,1]$. By (an easy case of) the Spectral Mapping Lemma, $\sigma_{\mathscr{A}}(1-a) \subset [0,1]$, and hence $\|1-a\| = \nu(1-a) \leq 1$.

Conversely, suppose a is self-adjoint and both a and $1-a$ are in $B_{\mathscr{A}}$. Since a is self adjoint and $\|a\| \leq 1$, $\sigma_{\mathscr{A}}(a) \subset [-1,1]$. Then by the Spectral Mapping Lemma, $\sigma_{\mathscr{A}}(1-a) \subset [0,2]$. However, if $\|1-a\| \leq 1$, then $\sigma_{\mathscr{A}}(1-a) \subset [-1,1]$, and altogether we have that $\sigma_{\mathscr{A}}(1-a) \subset [0,1]$, and then by the identity $a = 1-(1-a)$ and the Spectral Mapping Lemma once more, $\sigma_{\mathscr{A}}(a) \subset [0,1]$, so that $a \in \mathscr{A}^+$. That is,

$$\mathscr{A}^+ \cap B_{\mathscr{A}} = \{\ a \in \mathscr{A}\ :\ a = a^* \quad \text{and} \quad a \in B_{\mathscr{A}} \cap (1 - B_{\mathscr{A}})\ \}\ . \tag{2.5}$$

Now let $a, b \in B_{\mathscr{A}}$. Then by Minkowski's inequality $\|(a+b)/2\| \in B_{\mathscr{A}}$ and

$$\left\| 1 - \frac{a+b}{2} \right\| \leq \frac{1}{2}(\|1-a\| + \|1-b\|)\ . \tag{2.6}$$

If furthermore, $a, b \in \mathscr{A}^+$, then we also have that $\|1-a\| \leq 1$ and $\|1-b\| \leq 1$, and then from (2.6), $\|1-(a+b)/2\| \leq 1$. Thus, $(a+b)/2$ is self adjoint and belongs to both $B_{\mathscr{A}}$ and $1 - B_{\mathscr{A}}$. , and hence $(a+b)/2 \in \mathscr{A}^+$.

Since the closure of $\mathscr{A}^+$ under positive multiples is clear, it then clear that $\mathscr{A}^+$ is closed under sums. Finally, if $a \in \mathscr{A}^+$ and $-a \in \mathscr{A}^+$, then $\sigma_{\mathscr{A}}(a) \subset (-\infty, 0] \cap [0, \infty) = \{0\}$, so that $\|a\| = \nu(a) = 0$, and hence $a = 0$. □

2.20 THEOREM (Fukamiya-Kaplansky Theorem)**.** *For all $a \in \mathscr{A}$, a C^* algebra, $a^*a \in \mathscr{A}^+$.*

Proof. We first show that if $a^*a \in -\mathscr{A}^+$, then $a^*a = 0$. Since a^*a and aa^* have the same spectrum, Fukamiya's Theorem says that $a^*a + aa^* \in -\mathscr{A}^+$. However, writing $a = x + iy$ with x and y self adjoint,

$$a^*a + aa^* = 2(x^2 + y^2) \in \mathscr{A}^+$$

where once again we have used Fukamiya's Theorem, and the Spectral Mapping Lemma. Since $\mathscr{A}^+$ is a pointed cone, his means that $a^*a + aa^* = 0$. But then $a^*a = (a^*a + aa^*) - aa^* = -aa^* \in \mathscr{A}^+$. Again since $\mathscr{A}^+$ is pointed, this means that $a^*a = 0$, as claimed.

Now suppose that $x = b^*b$ for some $b \in \mathscr{A}$, Define continuous functions $f, g : \mathbb{R} \to \mathbb{R}$ by $f(t) = \max\{t, 0\}$ and $g(t) = t - f(t)$. Note that $f(t)g(t) = 0$ for all t. By the Abstract Spectral Theorem, if we define $y = f(x)$ and $z = g(x)$, then $yz = 0$, and $y + z = x = b^*b$. Now define $w = bz$, Then

$$w^*w = zb^*bz = z(y + z)z = z^3 \ .$$

Since $\sigma_{\mathscr{A}}(z) \subset (-\infty, 0]$, $z^3 \in -\mathscr{A}^+$, and the first part of the proof says that $w^*w = 0$. Therefore, $z = 0$, and so $b^*b = f(b^*b) \in \mathscr{A}^+$. □

2.4 Homeomorphisms of C^* algebras

2.21 THEOREM. *Let $\mathscr{A}$ be a C^* algebra, and let $\mathscr{J}$ be a norm-closed ideal in $\mathscr{A}$. Then $\mathscr{J}$ is closed under the involution.*

The heart of the proof is the following approximation lemma:

2.22 LEMMA. *Let $\mathscr{A}$ be a C^* algebra, and let $\mathscr{J}$ be a norm-closed ideal in $\mathscr{A}$. Then for every $a \in \mathscr{J}$, there is a sequence $\{u_n\}_{n \in \mathbb{N}}$ of positive elements of $\mathscr{J}$ with $\|u_n\| \leq 1$ for all n such that*

$$\lim_{n\to\infty} \|au_n - a\| = 0 \ .$$

Proof. Consider the sequence of continuous functions $f_n : \mathbb{R}_+ \to \mathbb{R}_+$ given by $f_n(t) = \min\{nt, 1\}$. Note that

$$t(1 - f_n(t))^2 = \begin{cases} t(1 - nt)^2 & t \leq 1/n \\ 0 & t > 1/n \ . \end{cases}$$

Evidently $\sup_{t \geq 0} |\{t(1 - f_n(t))^2\}| \leq 1/n$. Consequently, by the Abstract Spectral Theorem applied to a^*a for any $a \in \mathscr{A}$, $\|(f_n(a^*a) - 1)a^*a(f_n(a^*a) - 1)\| \leq 1/n$. Note that

$$\|(f_n(a^*a) - 1)a^*a(f_n(a^*a) - 1)\| = \|af_n(a^*a) - a\|^2 \ .$$

Thus, $\lim_{n\to\infty} \|af_n(a^*a) - a\| = 0$. By the Abstract Spectral Theorem, $\|f_n(a^*a)\| \leq 1$ and $f_n(a^*a) \in \mathscr{A}^+$ for all n. It remains to show that when $a \in \mathscr{J}$, then $f_n(a^*a) \in \mathscr{J}$ for all n. Clearly when $a \in \mathscr{J}$, a^*a and all polynomials in a^*a belong to $\mathscr{J}$. But then since f_n may be uniformly approximated by polynomials, and since $\mathscr{J}$ is norm closed, $f_n(a^*a) \in \mathscr{J}$. □

Proof of Theorem 2.21. Let $a \in \mathscr{J}$, and let $\{u_n\}_{n\in\mathbb{N}}$ be a sequence of positive elements of $\mathscr{J}$ such that $\lim_{n\to\infty} \|au_n - a\| = 0$. Since $\|u_n a^* - a^*\| = \|au_n - a\|$, and $u_n a^* \in \mathscr{J}$, $\lim_{n\to\infty} \|u_n a^* - a^*\|0$. Then since $\mathscr{J}$ is closed, $a^* \in \mathscr{J}$. □

Now let $\mathscr{A}$ be a C^* algebra, and let $\mathscr{J}$ be a closed ideal in $\mathscr{J}$. As usual, let $\{a\}$ denote equivalence class of a mod $\mathscr{J}$, and let $\|\{a\}\||$ denote the quotient norm of $\{a\}$; that is,

$$\|\{a\}\|| = \inf\{\ \|a - b\|\ :\ b \in \mathscr{J}\ \}\ .$$

Then $\mathscr{A}/\mathscr{J}$ is a Banach algebra with the quotient space norm. By Theorem 2.21, $a - b \in \mathscr{J} \iff a^* - b^* \in \mathscr{J}$, and therefore we may define an involution on on $\mathscr{J}/\mathscr{A}$ by

$$\{a\}^* = \{a^*\}\ .$$

Evidently this involution is an isometry, and so for all $a \in \mathscr{A}$, $\|\{a\}^*\{a\}\| \le \|\{a\}\|^2$. To show that $\mathscr{A}/\mathscr{J}$ is a C^* algebra with this involution, we need only show that for all $a \in \mathscr{A}$,

$$\|\{a\}\|^2 \le \|\{a\}^*\{a\}\|\ . \tag{2.7}$$

We shall use the following lemma:

2.23 LEMMA. *Let $\mathscr{A}$ be a C^* algebra, and let $\mathscr{J}$ be a closed ideal in $\mathscr{A}$. For all $a \in \mathscr{A}$, the quotient norm of $\{a\}$ is given by*

$$\|\{a\}\| = \inf\{\ \|a - au\|\ :\ u^* = u \in \mathscr{J} \quad \text{and} \quad u \in \mathscr{A}^+ \cap B_{\mathscr{A}}\ \}\ . \tag{2.8}$$

Proof. Whenever $u \in \mathscr{J}$, $a \sim (a - au)$ so that $\|\{a\}\|$ is no greater than the right hand side of (2.8). To prove the equality, pick $\epsilon > 0$ and $b \in \mathscr{J}$ so that $\|\{a\}\| \ge \|a - b\| - \epsilon$. Then by (2.5), $\|1 - u\| \le 1$, and so

$$\|a - b\| \ge \|a - b\|\|1 - u\| \ge \|(a - b)(1 - u)\| = \|(a - au) - (b - bu)\| \ge \|(a - au)\| - \|(b - bu)\|\ .$$

By Lemma 2.22 we can choose u so that $\|b - bu\| < \epsilon$. We then have $\|\{a\}\| \ge \|a - au\| - 2\epsilon$, and since $\epsilon > 0$ is arbitrary, (2.8) is proved. □

Now to prove (2.7), pick $\epsilon > 0$ and $u = u^* \in \mathscr{J}$ with $u \in \mathscr{A}^+ \cap B_{\mathscr{A}}$ so that

$$\|a^*a(1 - u)\| \le \|\{a\}^*\{a\}\| + \epsilon = \|\{a\}\|^2 + \epsilon\ .$$

Then

$$\|\{a\}\|^2 \le \|a(1 - u)\|^2 = \|(1 - u)a^*a(1 - u)\| \le \|(1 - u)\|\|a^*a(1 - u)\| \le \|a^*a(1 - u)\|\ .$$

where the last in equality is valid since by (2.5), $\|1 - u\| \le 1$. Altogether, $\|\{a\}\|^2 \le \|\{a\}^*\{a\}\| + \epsilon$, and since $\epsilon > 0$ is arbitrary, (2.7) is proved. We have shown:

2.24 THEOREM. *Let $\mathscr{A}$ be a C^* algebra, and let $\mathscr{J}$ be a norm closed ideal in $\mathscr{A}$, then $\mathscr{J}$ is closed under the involution, and the definition $\{a\}^* = \{a^*\}$ defines an involution on $\mathscr{A}/\mathscr{J}$ so that, equipped with the quotient norm, $\mathscr{A}/\mathscr{J}$ is a C^* algebra.*

2.25 LEMMA. *Let $\mathscr{A}$ and $\mathscr{B}$ be C^* algebras, and let $\pi : \mathscr{A} \to \mathscr{B}$ be a $*$-homomorphism. Then π is a contraction; i.e., $\|\pi(a)\| \le \|a\|$ for all $a \in \mathscr{A}$. If moreover π is one-to-one, π is an isometry.*

Proof. For all $a \in \mathscr{A}$, by the Spectral Contraction Theorem, $\nu(\pi(a)^*\pi(a)) = \nu(\pi(a^*a) \le \nu(a^*a)$. Then since for self adjoint elements of a C^* algebra, the norm is the spectral radius, $\|\pi(a)^*\pi(a)\| \le \|a^*a\|$. Then by the crucial defining property of a C^* algebra relating the norm and the involution, $\|\pi(a)\|^3 \le \|a\|^2$, and this proves that π is a contraction.

Notice that if $\nu(\pi(a^*a) = \nu(a^*a)$, the argument gives $\|\pi(a)\| = \|a\|$. Hence it remains to show that if π is one-to-one, π cannot decrease the spectral radius of any self adjoint element of $\mathscr{A}$.

Indeed, let $a = a^* \in \mathscr{A}$, and suppose that $\nu(\pi(a)) < \nu(a)$. Then there is a non-zero continuous bounded function f supported on $[-\nu(a), \nu(a)]$ that vanishes identically on $[-\nu(\pi(a)), \nu(\pi(a))]$. since f may be approximated by polynomials, $\pi(f(a)) = f(\pi(a))$. However, since f vanishes identically on the spectrum of $\pi(a)$, $f(\pi(a)) = 0$. Thus, $f(a)$ is in the kernel of π, which is a contradiction. Hence, when π is one-to-one, it preserves the spectral radius of self adjoint elements. □

We summarize with the following theorem:

2.26 THEOREM (Homomorphisms of C^* algebras)**.** *Let $\mathscr{A}$ and $\mathscr{B}$ be C^* algebras, and let $\pi : \mathscr{A} \to \mathscr{B}$ be a $*$-homomorphism. Then π is a contraction, $\pi(\mathscr{A})$ is a C^*-subalgebra of $\mathscr{B}$, and π induces an isometric isomorphism of $\mathscr{A}/\ker(\pi)$ onto $\pi(\mathscr{A})$.*

2.5 Projections in C^* algebras

2.27 DEFINITION. Let $\mathscr{A}$ be a C^* algebra. A self adjoint element e of $\mathscr{A}$ is a *projection* in case $e^2 = e$. A projection e is a *central projection* in case e commutes with every element of $\mathscr{A}$.

Note that 0 is a projection, as is 1 when $\mathscr{A}$ has an identity. Any other projections, should they exist, are *non-trivial* projections. Suppose that e is a non-trivial projection in $\mathscr{A}$. Then $1 - e$ is also a non-trivial projection in $\mathscr{A}$.

Associated to e are the two subalgebras, namely $e\mathscr{A}e$ and $(1-e)\mathscr{A}(1-e)$, where $e\mathscr{A}e$ consists of all elements of $\mathscr{A}$ of the form eae, $a \in \mathscr{A}$, and likewise $(1-e)\mathscr{A}(1-e)$ consists of all elements of $\mathscr{A}$ of the form $(1-e)a(1-e)$, $a \in \mathscr{A}$. Evidently, these are both C^* subalgebras of $\mathscr{A}$. Note that e is the identity in $e\mathscr{A}e$, and $(1-e)$ is the identity in $(1-e)\mathscr{A}(1-e)$

The name "corner algebra" come from the case in which $\mathscr{A} = M_n(\mathbb{C})$, the algebra of $n \times n$ complex matrices, and for some $1 \le m \le m-1$, e is the orthogonal projection onto the subspace of $\mathbb{C}^n$ consisting of vectors $(\eta_1, \dots, \eta_n)$ such that $\eta_j = 0$ for $j \ge m+1$. The general element of $a \in \mathscr{A}$ can then be written in "block form" as $a = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$ where $x \in M_m(\mathbb{C})$, $w \in M_{n-m}(\mathbb{C})$ and y and z^T are $m \times (n-m)$ matrices. Then

$$eae = \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad (1-e)a(1-e) = \begin{bmatrix} 0 & 0 \\ 0 & y \end{bmatrix} .$$

Continuing with this example, note that if $a = eae + (1-e)a(1-e)$ so that $a = \begin{bmatrix} x & 0 \\ 0 & w \end{bmatrix}$, then a is invertible if and only if both x and z are invertible, and so for such matrices a,

$$\sigma_{M_n(\mathbb{C})}(a) = \sigma_{M_m(\mathbb{C})}(x) \cup \sigma_{M_{n-m}(\mathbb{C})}(w) \ .$$

We shall be especially interested in the case in which e is central. Note that this is not the case in the example we just considered.

If e commutes with every element of cA, then $e\mathscr{A}e = \{ \ ea : a \in \mathscr{A} \} = \{ \ ae \ : \ a \in \mathscr{A} \ \}$, and $e\mathscr{A}e$ is then evidently an ideal in $\mathscr{A}$, as is $(1-e)\mathscr{A}(1-e)$. In this case,

$$\mathscr{A} = e\mathscr{A}e \oplus (1-e)\mathscr{A}(1-e)$$

since for all $a \in \mathscr{A}$, $a = ae + a(1-e)$ and if $a \in e\mathscr{A}e \cap (1-e)\mathscr{A}(1-e)$, $a = ea$ and $a = (1-e)a$ so that $a = e(1-e)a = 0$.

For all $a \in \mathscr{A}$, a is invertible if and only if ea is invertible in $e\mathscr{A}e$ and $(1-e)a$ is invertible in $(1-e)\mathscr{A}(1-e)$. To see this, suppose that a is invertible in $\mathscr{A}$. Then $e = eaa^{-1} = e^2aa^{-1} = (ea)(ea^{-1}) = (ea^{-1})(ea)$, and so the inverse of ea in $e\mathscr{A}e$ is ea^{-1}. The same reasoning shows that the inverse of $(1-e)a$ in $(1-e)\mathscr{A}(1-e)$ is $(1-e)a^{-1}$. For the converse, suppose that ea has the inverse ex in $e\mathscr{A}e$ and that $(1-e)a$ has the inverse $(1-e)y$ in $(1-e)\mathscr{A}(1-e)$. Then

$$\begin{aligned} a(ex + (1-e)y) &= (ae + a(1-e))((ex + (1-e)y) = ae^2x + a(1-e)^2y \\ &= (ea)(ex) + (1-e)a(1-e)y = e + (1-e) = 1 \ , \end{aligned}$$

thus showing that $(ex + (1-e)y)$ is a right inverse of a. A similar computation shows that it is a left inverse. This leads to the following result:

2.28 THEOREM. *Let $\mathscr{A}$ be a C^* algebra with identity 1, and let e be a central projection in $\mathscr{A}$. For any $a \in \mathscr{A}$, let $a = ea + (1-e)a$ be the unique decomposition of a corresponding to $\mathscr{A} = e\mathscr{A}e \oplus (1-e)\mathscr{A}(1-e)$. Then*

$$\sigma_{\mathscr{A}}(a) = \sigma_{e\mathscr{A}e}(ea) \cup \sigma_{(1-e)\mathscr{A}(1-e)}(a) \ .$$

Proof. For all $\lambda \in \mathbb{C}$, $\lambda 1 - a = (\lambda e - ea) + (\lambda(1-e) - (-1e)a)$. By the remarks above the theorem, $\lambda \in \rho_{\mathscr{A}}(a)$ if and only if $\lambda \in \rho_{e\mathscr{A}e}(ea) \cap \rho_{(1-e)\cap(1-e)}((1-e)a)$. □

3 Lin's Theorem

3.1 Almost commuting and nearly commuting

In 1995, Huaxin Lin proved a theorem that settled an old conjecture arising from the work of John von Neumann on quantum mechanics. His theorem concerns the C^* algebra $M_n(\mathbb{C})$ of complex $n \times n$ matrices:

3.1 THEOREM (Lin's Theorem)**.** *For every $\epsilon > 0$, there is a $\delta > 0$ such that for any $n \in \mathbb{N}$ and any pair of self-adjoint $a, b \in M_n(\mathbb{C})$ with $\|a\|, \|b\| \leq 1$ and*

$$\|ab - ba\| \leq \delta \ , \tag{3.1}$$

there exists a commuting pair of self adjoint $\tilde{a}, \tilde{b} \in M_n(\mathbb{C})$ *such that*

$$\|a - \tilde{a}\| + \|b - \tilde{b}\| \leq \epsilon \ . \tag{3.2}$$

When a pair $a, b \in M_n(\mathbb{C})$ satisfies (3.1) for small δ, we may say that they *almost* commute. When a pair $a, b \in M_n(\mathbb{C})$ is such that there exists a commuting pair $\tilde{a}, \tilde{b} \in M_n(\mathbb{C})$ such that (3.2) is satisfied for small ϵ, we may say that a and b are *nearly* commuting – they are near to matrices that exactly commute.

The theorem may be rephrased as a theorem about "almost normal" and "nearly normal" matrices. Let $a, b \in M_n(\mathbb{C})$ be self adjoint $n \times n$ matrices Let $x = a + ib$. Then $x^*x - xx^* = 2i(ab - ba)$ so that

$$\|x^*x - xx^*\| = 2\|ab - ba\| \tag{3.3}$$

so that $\|x^*x - xx^*\| \leq 2\delta$.

Let $\tilde{a}$ and $\tilde{b}$ be another pair of self adjoint $n \times n$ matrices, and define $\tilde{x} = \tilde{a} - \tilde{b}$. Note that $x - \tilde{x} = (a - \tilde{a}) + i(b - \tilde{b})$, so that $\|x - \tilde{x}\| \leq \|a - \tilde{a}\| + \|b - \tilde{b}\|$. However,

$$a - \tilde{a} = \frac{(x - \tilde{x}) + (x^* - \tilde{x}^*)}{2}$$

so that $\|a - \tilde{a}\| \leq \|x - \tilde{x}\|$. Likewise, we have $\|b - \tilde{b}\| \leq \|x - \tilde{x}\|$. Altogether,

$$\|x - \tilde{x}\| \leq \|a - \tilde{a}\| + \|b - \tilde{b}\| \leq 2\|x - \tilde{x}\| \ . \tag{3.4}$$

Combining (3.3) and (3.4), we arrive at an alternate formulation of Lin's Theorem:

3.2 THEOREM (Lin's Theorem, Alternate Formulation)**.** *For every* $\epsilon > 0$*, there is a* $\delta > 0$ *such that for any* $n \in \mathbb{N}$ *and every* $x \in M_n(\mathbb{C})$ *with* $\|x\| \leq 1$*, and*

$$\|x^*x - xx^*\| \leq \delta \ , \tag{3.5}$$

there exists a normal $\tilde{x} \in M_n(\mathbb{C})$ *such that*

$$\|x - \tilde{x}\| \leq \epsilon \ . \tag{3.6}$$

The crucial feature of Lin's Theorem is that δ depends only on ϵ and not on n. Without the requirement that δ be independent of n, the result is trivial. It then suffices to show that for fixed n, and fixed $\epsilon > 0$, there does not exists a sequence $\{x_j\}_{j \in \mathbb{N}}$ of $n \times n$ matrices with $\|x_j\| \leq 1$ for all j and such that $\lim_{j \to \infty} \|x_j^* x_j - x_j x_j^*\| = 0$ but $\|x_j - x\| > \epsilon$ for all normal x and all j

Suppose such a sequence exists. By the compactness of the unit ball in the space of $n \times n$ matrices, there is a subsequence $\{x_{j_k}\}_{k \in \mathbb{N}}$ and an x with $\|x\| \leq 1$ such that $\lim_{k \to \infty} \|x_{j_k} - x\| = 0$. Evidently

$$x^*x - xx^* = \lim_{k \to \infty} (x_{j_k}^* x_{j_k} - x_{j_k} x_{j_k}^*) = 0 \ .$$

Therefore, x is normal but $\|x_{j_k} - x\| < \epsilon$ for all sufficiently large k.

3.2 The finite spectrum problem

Lin's proof of his theorem turns on the analysis of two C^* algebras that we now define: For any sequence $\{n_j\}_{j\in\mathbb{N}}$ of natural numbers, define two sets of sequences of matrices as follows:

$$\mathscr{A} = \{\ \{a_j\}_{j\in\mathbb{N}}\ :\ a_j \in M_{n_j}(\mathbb{C}) \quad \text{and} \quad \sup_{j\in\mathbb{N}} \|a_j\| < \infty\ \} \tag{3.7}$$

and

$$\mathscr{J} = \{\ \{b_j\}_{j\in\mathbb{N}}\ :\ b_j \in M_{n_j}(\mathbb{C}) \quad \text{and} \quad \lim_{j\to\infty} \|b_j\| = 0\ \} \tag{3.8}$$

Obviously $\mathscr{J} \subset \mathscr{A}$. Equip $\mathscr{A}$ with the operations of term by term addition and multiplication and the norm

$$\|\{a_j\}_{j\in\mathbb{N}}\| = \sup_{j\in\mathbb{N}} \|a_j\|\ .$$

This makes $\mathscr{A}$ a Banach algebra. Equip $\mathscr{A}$ with the involution consisting of term by term Hermitian conjugation. This make $\mathscr{A}$ a C^* algebra, and $\mathscr{J}$ a closed ideal in $\mathscr{A}$. Let $\mathscr{B}$ denote the quotient algebra $\mathscr{A}/\mathscr{J}$, and let π denote the natural homomorphism of $\mathscr{A}$ onto $\mathscr{B}$. This notation will be used throughout this section.

The relevance of this construction is as follows: If Theorem 3.2 were false, there would exist a sequence $\{n_j\}_{j\in\mathbb{N}}$ of natural numbers, and a sequence $\{x_j\}_{j\in\mathbb{N}}$ with each $x_j \in M_{n_j}(\mathbb{C})$ such that for some $\epsilon > 0$,

$$\|x_j - \tilde{x}_j\| \geq \epsilon \quad \text{for all } j \in \mathbb{N} \quad \text{and all normal} \quad \tilde{x}_j \in M_{n_j}(\mathbb{C})\ , \tag{3.9}$$

and

$$\lim_{j\to\infty} \|x_j^* x_j - x_j x_j^*\| = 0\ . \tag{3.10}$$

Let us write x to denote $\{x_j\}_{j\in\mathbb{N}}$ considered as an element of $\mathscr{A}$, and let us write y to denote $\pi(x) \in \mathscr{B}$. By (3.10), which says that $\{x_j^* x_j - x_j x_j^*\}_{j\in\mathbb{N}} \in \mathscr{J}$,

$$y^*y - yy^* = \pi(x^*)\pi(x) - \pi(x)\pi(x^*) = \pi(x_j^* x_j - x_j x_j^*) = 0\ .$$

Thus, for any $x = \{x_j\}_{j\in\mathbb{N}}$ satisfying (3.10), $y = \pi(x)$ is normal in $\mathscr{B}$.

We say that an element of a Banach algebra has *finite spectrum* if its spectrum is a finite subset of $\mathbb{C}$. There are two parts to Lin's proof. One is to show that every normal $y \in \mathscr{B}$ can be approximated arbitrarily well in norm by a normal element $\tilde{y}$ that has finite spectrum. The other is to show that for any $x = \{x_j\}_{j\in\mathbb{N}} \in \mathscr{A}$, if $\pi(x)$ is normal with finite spectrum, then (3.9) is impossible. We begin with the latter point.

3.3 LEMMA. *Let $x = \{x_j\}_{j\in\mathbb{N}} \in \mathscr{A}$ and suppose that $y = \pi(x)$ is normal and has finite spectrum. Then there exists a normal $\tilde{x} = \{\tilde{x}_j\}_{j\in\mathbb{N}} \in \mathscr{A}$ such that $\pi(\tilde{x}) = y$, and, consequently, such that*

$$\lim_{j\to\infty} \|x_j - \tilde{x}_j\| = 0\ . \tag{3.11}$$

In other words, normal equivalence classes in $\mathscr{A}/\mathscr{J}$ that have finite spectrum have a normal representative.

Proof. Let $\{\lambda_1, \ldots, \lambda_m\}$ be the points in the spectrum of y. Let p and q be complex polynomials with

$$p(\lambda_j) = j \quad \text{and} \quad q(j) = \lambda_j \quad \text{for} \quad j = 1, \ldots, n \ .$$

Notice that $q \circ p(\lambda) = \lambda$ on $\sigma_{\mathscr{B}}(y)$ so that $q(p(y)) = y$. Since $p(\lambda) \in \mathbb{R}$ for all $\lambda \in \sigma_{\mathscr{B}}(y)$, $p(y)$ is self adjoint.

Let z be any element of $\mathscr{A}$ with $\pi(z) = p(y)$. Then $\pi(z)$ is self adjoint, and so $\pi((z^* + z)/2) = p(y)$. Then

$$q(\pi((z^* + z)/2) = \pi(q((z^* + z)/2)) = q(p(y)) = y \ ,$$

Then since $(z^* + z)/2$ is self adjoint, $q((z + z^*)/2)$ is normal, and thus the equivalence class of y contains a normal representative, namely $q((z^* + z)/2)$, that we denote by $\tilde{x}$.

By the definition of the norm in the quotient algebra, for all $\epsilon > 0$, there exists $b = \{b_j\}_{j \in \mathbb{N}} \in \mathscr{J}$ such that $\|x - \tilde{x} - b\| \leq \epsilon$. This means that

$$\|x_j - \tilde{x}_j\| \leq \epsilon + \|b_j\| \ .$$

Then since $\epsilon > 0$ is arbitrary and $\lim_{j\to\infty} \|b_j\| = 0$, (3.11) is proved. □

3.3 Approximation of normals by normals with finite spectrum

It remains to show that every normal element of $\mathscr{B}$ can be well-approximated by normal elements with finite spectrum. To prepare for this, we make several observations about the algebra $\mathscr{A}$. Consider $x \in M_n(\mathbb{C})$ for some $n \in \mathbb{N}$. Then x has a *singular value decomposition*

$$x = \tilde{u}s\tilde{v}^*$$

where s is a diagonal matrix with non-negative entries, and u and v are unitary matrices. Since $\tilde{u}^*\tilde{u} = 1$, $x = \tilde{v}\tilde{u}^*(\tilde{u}s\tilde{u}^*)$. We define $u = \tilde{v}\tilde{u}^*$ and $|x| = \tilde{u}s\tilde{u}^*$. Then we have

$$x = u|x| \quad \text{and} \quad |x| = \sqrt{x^*x} \ ,$$

where the square root is defined by the functional calculus. Because u is unitary, this is called a *unitary polar decomposition* of x.

Now we observe that every element $x = \{x_j\}_{j\in\mathbb{N}} \in \mathscr{A}$ has a unitary polar decomposition $x = u|x|$: Simply choose such a decomposition $x_j = u_j|x_j|$ for each j, and then $u = \{u_j\}_{j\in\mathbb{N}}$ and $|x| = \{|x|_j\}_{j\in\mathbb{N}}$.

Next consider any $y \in \mathscr{B}$, and any $x \in \mathscr{A}$ such that $\pi(x) = y$. Let $u|x|$ be a unitary polar decomposition of x. Then

$$y = \pi(u)\pi(|x|) = \pi(u)\pi(\sqrt{x^*x}) = \pi(u)\sqrt{y^*y} = \pi(u)|y| \ ,$$

and $\pi(u)$ is unitary. Therefore, each element of $\mathscr{B}$ has a unitary polar decomposition.

Essentially the same argument shows that every unitary $v \in \mathscr{B}$, has a unitary representative in $\mathscr{A}$; i.e, there exists a unitary $u \in \mathscr{A}$ such that $\pi(u) = v$. To see this, consider any $x \in \mathscr{A}$ such that $\pi(x) = y$, and and let $x = u|x|$ be a unitary polar decomposition of x. Then $y = \pi(u)\pi(|x|) = \pi(u)|\pi(x)| = \pi(u)|y| = \pi(u)$. While we have to do significant work to obtain even an approcximate normal representative for normal $y \in \mathscr{B}$, for unitary $v \in \mathscr{B}$, things are much simpler: THere is always an *exact* unitiary representative in $\mathscr{A}$. This will be used below.

3.4 LEMMA. *Let any $\epsilon > 0$ and any countable subset F of $\mathbb{C}$ be given. Then for all normal $y \in \mathscr{B}$, there exists a normal $\tilde{y} \in \mathscr{B}$ such that $\|y - \tilde{y}\| \le \epsilon$ and $F \cap \sigma_{\mathscr{B}}(\tilde{y}) = \emptyset$.*

Proof. The set of invertible normal elements is dense in the set of normal elements of $\mathscr{B}$. To see this let, $y \in \mathscr{B}$ be normal and let $y = v|y|$ be a unitary polar decomposition. Then $|y|^2 = y^*y = yy^* = v|y|^2v^*$, which mean that $|y|^2v = v|y|^2$ so that v computes with $|y|^2$, and hence any polynomial in $|y|^2$, and hence any continuous function of $|y|^2$. In particular, v commutes with $|y|$. Evidently, $v(|y| + 1\epsilon)$ is invertible and normal since v is unitary and commutes with $|y|$. Clearly $\|y - v(|y| + 1\epsilon)\| \le \epsilon$ and this justifies the claim that the set of invertible normal elements is dense in the set of normal elements of $\mathscr{B}$.

It follows that for each $\lambda \in F$, the set of normal elements $z \in \mathscr{B}$ such that $\lambda 1 - z$ is invertible is dense and open in the relative topology. By Baire's Theorem, the intersection of these sets over all $\lambda \in F$ is dense in the normal elements of $\mathscr{B}$. □

This lemma shall be applied to approximate an arbitrary normal $y \in \mathscr{B}$ by another normal $\tilde{y} \in \mathscr{B}$ where $\sigma_{\mathscr{B}}(\tilde{y})$ lies on the ϵ grid $\Gamma_\epsilon \subset \mathbb{C}$, where for $\epsilon > 0$,

$$\Gamma_\epsilon = \{s + it \in \mathbb{C} \ : \ s \in \epsilon\mathbb{Z} \quad \text{or} \quad t \in \epsilon\mathbb{Z} \ \} \ . \tag{3.12}$$

To do this, fix $\epsilon > 0$, and let F be the set of the centers of the squares in Γ_ϵ. That is, the set

$$F_\epsilon := \{ \ s + it \in \mathbb{C} \ : \ s+ \in \epsilon(\mathbb{Z} + 1/2) \quad \text{and} \quad t \in \epsilon(\mathbb{Z} + 1/2) \ \} \ .$$

Let f be the obvious continuous contraction from $\mathbb{C} \backslash F_\epsilon$ onto Γ_ϵ, such that for all $\lambda \in \mathbb{C} \backslash F_\epsilon$,

$$|f(\lambda) - \lambda| \le \epsilon/\sqrt{2} \ . \tag{3.13}$$

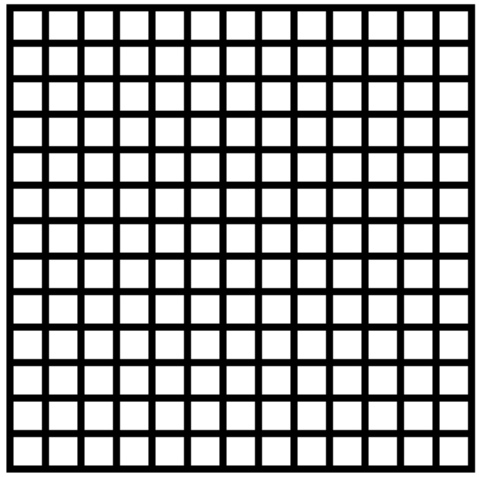
Define $\tilde{y} = f(y)$. Then $\tilde{y}$ is normal and by the Spectral Mapping Theorem, $\sigma_{\mathscr{B}}(\tilde{y}) \in \Gamma_\epsilon$. By (3.13), $\|\tilde{y} - y\| < \epsilon/\sqrt{2}$. This proves:

3.5 LEMMA. *For all normal $y \in \mathscr{B}$ there and all $\epsilon > 0$, there exists a normal $\tilde{y} \in \mathscr{B}$ such that $\sigma_{\mathscr{B}}(\tilde{y}) \subset \Gamma_\epsilon$ and $\|y - \tilde{y}\| < \epsilon$.*

Now fix $\epsilon > 0$ and consider any normal $y \in \mathscr{B}$ such that $\sigma_{\mathscr{B}}(\tilde{y}) \subset \Gamma_\epsilon$. Then since $\sigma_{\mathscr{B}}(\tilde{y})$ is a closed subset of $\mathbb{C}$ contained in $D_{\|y\|}$, the closed centered disc of radius $\|y\|$ in $\mathbb{C}$,

$$\sigma_{\mathscr{B}}(\tilde{y}) \subset \Gamma_\epsilon \cap D_{\|y\|} \ .$$

At this point we have that the spectrum of y lies in a subset of $\mathbb{C}$ that looks something like the following:



Now consider two the sets

$$\Lambda_\epsilon = \{ \ s+it \in \mathbb{C} \ : \ s \in \epsilon\mathbb{Z} \text{ and } t \in \epsilon(\mathbb{Z}+1/2) \quad \text{or} \quad s \in \epsilon\mathbb{Z} \text{ and } t \in \epsilon(\mathbb{Z}+1/2) \ \} \tag{3.14}$$

and

$$\tilde{\Lambda}_\epsilon = \Lambda_\epsilon = \{ \ s+it \in \mathbb{C} \ : \ s \in \epsilon\mathbb{Z} \text{ and } t \in \epsilon\mathbb{Z} \ \} \ . \tag{3.15}$$

Note that Λ_ϵ is the set of midpoints on the elementary segments of the grid Γ_ϵ, and $\tilde{\Lambda}_\epsilon$ is the set of intersection points of the grid Γ_ϵ.

There is an obvious continuous g retraction of $\Gamma_\epsilon\backslash\Lambda_\epsilon$ onto $\tilde{\Lambda}_\epsilon$ such that for all $\lambda \in \Gamma_\epsilon$,

$$|g(\lambda) - \lambda| \le \epsilon/2 \ .$$

Therefore, if y is any normal element of $\mathscr{B}$ with spectrum in $\Gamma_\epsilon\backslash\Lambda_\epsilon$, $g(y)$ is a normal element with $\|y - g(y)\| \le \epsilon/2$, and $\sigma_{\mathscr{B}}(g(y)) \subset \tilde{\Lambda}_\epsilon \cap D_{\|y\|}$, a finite set whose cardinality depends only on ϵ and $\|y\|$.

We now turn to the lemma that will enable us to remove, one at a time, the finitely many points of $\Lambda_\epsilon \cap D_{\|y\|}$ from the spectrum of our normal element y, (whose spectrum lies in $\Gamma_\epsilon \cap D_{\|y\|}$). This will give us the approximation by elements of finite spectrum that we seek.

3.6 LEMMA. *Let $y \in \mathscr{B}$ be normal. Let V be an open subset in $\mathbb{C}$ such that $V \cap \sigma_{\mathscr{B}}(y)$ is contained in a subset X of $\mathbb{C}$ that is homeomorphic to the open unit interval. Let $y_0 \in V \cap \sigma_{\mathscr{B}}(y)$ and suppose that λ_0 is not an isolated point of $\sigma_{\mathscr{B}}(y)$. Then for each $\epsilon > 0$, there exists a normal $\tilde{y} \in \mathscr{B}$ such that*

$$\sigma_{\mathscr{B}}(\tilde{y}) \subset \sigma_{\mathscr{B}}(y)\backslash\{\lambda_0\} \quad \text{and} \quad \|y - \tilde{y}\| < \epsilon \ .$$

We preface the proof with remarks on the strategy. Suppose we can find a commutative subalgebra $\mathscr{C}$ of $\mathscr{B}$ that contains $C(y)$ and a projection e, necessarily central, such that for some small neighborhood U of λ_0

$$\sigma_{e\mathscr{C}e}(ey) \subset \overline{U} \quad \text{and} \quad \sigma_{(1-e)\mathscr{C}(1-e)}((1-e)y) \subset \sigma_{\mathscr{B}}(y)\backslash U \ . \tag{3.16}$$

Pick any $\lambda_1 \neq \lambda_0 \in \sigma_{\mathscr{A}}(y) \cap U$. The function $\lambda \mapsto |\lambda_1 - \lambda|$ is bounded by $\text{diam}(U)$ on $\sigma_{e\mathscr{C}e}(ey)$. Therefore, by the Gelfand-Naimark Theorem, $\|\lambda_1 e - ey\| \le \text{diam}(U)$.

Define $\tilde{y} = \lambda_1 e + (1-e)y$ Then $\tilde{y}_1$ is normal, and $\|\tilde{y} - y\| = \|\lambda_1 e - ey\| \leq \mathrm{diam}(U)$. Finally, by Theorem 2.28 and (3.16), $\sigma_{\mathscr{C}}(\tilde{y}) = \{\lambda_1\} \cup \sigma_{(1-e)\mathscr{C}(1-e)}((1-e)y) \subset \{\lambda_1\} \cup \sigma_{\mathscr{B}}(y)\backslash U$. Then by Theorem 2.10,

$$\sigma_{\mathscr{B}}(\tilde{y}) \subset \{\lambda_1\} \cup \sigma_{\mathscr{B}}(y)\backslash U \subset \{\lambda_0\} \cup \sigma_{\mathscr{B}}(y)\backslash\{\lambda_0\} \ .$$

The construction of e requires some ingenuity: If we could apply the characteristic function 1_U to y, we would readily obtain e. However, 1_U need not be continuous on $\sigma_{\mathscr{B}}(y)$, and so the Abstract Spectral Theorem is not available. If y had finite spectrum, then of course there would be a continuous function f agreeing with 1_U on $\sigma_{\mathscr{B}}(y)$, and then we could define $e = f(y)$. However, $\sigma_{\mathscr{B}}(y)$ need not have any isolated points, and then there will be no such continuous function.

We will use the fact that y is an equivalence class of sequences of matrices, represented by some $x = \{x_j\}_{j\in\mathbb{N}}$. If each x_j were normal, we could apply the spectral theorem to define $1_u(x_j)$ for each j, and this would provide us with a projection of the sort we seek. However, we do not know that in general that normal $y \in \mathscr{B}$ have normal representatives – except in the special case that y is not only normal, but unitary. Therefore, we use the continuous functional calculus to convert y into a unitary, and then we work with a unitary representative of this in $\mathscr{A}$ to produce our desired projection.

Proof. Choose a relatively open set $U \subset X$ with

$$\lambda_0 \in U \subset \overline{U} \subset X \quad \text{and} \quad \mathrm{diam}(U) < \epsilon \ .$$

Let f_0 be a homeomorphism of X onto $\mathbb{T}\backslash\{-1\}$ where $\mathbb{T}$ is the unit circle in $\mathbb{C}$. Extend f_0 to a continuous function $f : \sigma_{\mathscr{B}}(y) \to \mathbb{C}$ by

$$f(\lambda) = \begin{cases} f_0(\lambda) & \lambda \in X \\ -1 & \lambda \in \sigma_{\mathscr{B}}(y) \cap X^c \ . \end{cases}$$

Set $v = f(y)$. Observe that v is unitary. Let u be any unitary in $\mathscr{A}$ with $\pi(u) = v$. Since f_0 is a homeomorphism of X onto $\mathbb{T}\backslash\{-1\}$, and since U is open in X, $W = f_0(U)$ is open in $\mathbb{T}$. Let 1_W denote the characteristic function of W.

We now use the Spectral Theorem for $n \times n$ matrices to define $1_W(u_j)$ for each $j \in \mathbb{N}$. For each $j \in \mathbb{N}$, $1_W(u_j)$ is a projection in $M_{n_j}(\mathbb{C})$. Therefore, $e = \pi\left(\{1_W(u_j)\}_{j\in\mathbb{N}}\right)$ is a projection in $\mathscr{B}$. For each $j \in \mathbb{N}$, $u_j 1_W(u_j) = 1_W(u_j)u_j$, and hence, $ue = eu$.

Let $\hat{\varphi}$ be any continuous function on $\mathbb{T}$. Then $\hat{\varphi}(u) \in C(u)$, and since e commutes with u, e commutes with $\hat{\varphi}(u)$. Now let $\varphi : \sigma_{\mathscr{B}}(y) \to \mathbb{C}$ be any continuous function with $\varphi(\lambda) = 0$ on $\sigma_{\mathscr{B}}(y)\backslash V$. Define a function $\hat{\varphi} : \mathbb{T} \to \mathbb{C}$ by

$$\hat{\varphi}(\lambda) = \begin{cases} \varphi(f_0^{-1}(\lambda)) & \lambda \in \mathbb{T}\backslash\{-1\} \\ 0 & \lambda = -1 \ . \end{cases}$$

Then $\varphi(y) = \hat{\varphi}(u)$ so that e commutes with $\varphi(y)$.

Suppose that $\varphi = 1$ everywhere on U. Then $\hat{\varphi} = 1$ everywhere on W. Hence, for each $j \in \mathbb{N}$,

$$1_W(u_j)\hat{\varphi}(u_j) = \hat{\varphi}(u_j)1_W(u_j) = 1_W(u_j) \ , \tag{3.17}$$

and hence $e\varphi(y) = \varphi(y)e = e$. Finally, if $\varphi|_{X\backslash U} = 0$, then $\hat{\varphi}|_{\mathbb{T}\backslash W} = 0$, and then

$$1_W(u_j)\hat{\varphi}(u_j) = \hat{\varphi}(u_j)1_W(u_j) = \hat{\varphi}(u_j)\ , \tag{3.18}$$

with the conseqeunce that $\varphi(y)e = e\varphi(y) = \varphi(y)$.

Summarizing the last two paragraphs, when φ is continuous on $\sigma_{\mathscr{B}}(y)$, then:

$$\varphi|_{\sigma_{\mathscr{B}}(y)\backslash V} = 0 \quad \Rightarrow \quad e\varphi(y) = \varphi(y)e\ , \tag{3.19}$$

$$\varphi|_{\sigma_{\mathscr{B}}(y)\backslash V} = 0 \quad \text{and} \quad \varphi|_{\overline{U}} = 1 \quad \Rightarrow \quad e\varphi(y) = \varphi(y)e = e\ , \tag{3.20}$$

and

$$\varphi|_{\sigma_{\mathscr{B}}(y)\backslash U} = 0 \quad \Rightarrow \quad e\varphi(y) = \varphi(y)e = \varphi(y)\ . \tag{3.21}$$

Now let $h : X \to [0,1]$ be a continuous function such that $h|_{\overline{U}} = 1$ and $h|_{X\backslash V} = 0$. Then using (3.20), (3.19), the commutativity of $C(y)$, and then (3.20) once more,

$$ye = yh(y)e = eyh(y) = eh(y)y = ey\ .$$

Thus, e commutes with y.

Let $\mathscr{C}$ be the smallest C^* algebra containing y, e and 1. Evidently $\mathscr{C}$ is commutative, and e is a central projection. We now claim that $\sigma_{e\mathscr{C}e}(ey) \subset \overline{U}$. To see this, it suffices to show that whenever ψ is continuous on $\sigma_{\mathscr{B}}(y)$ with $\psi|_{\overline{U}} = 0$, then $\psi(ey) = 0$. In fact, it suffices to do this for continuous ψ such that $\psi|_{\sigma_{\mathscr{B}}(y)\backslash V} = 1$, since by the Gelfand-Naimark Theorem, $\psi(ey) \neq 0$ whenever ψ takes on any non-zero value anywhere on the spectrum of ey. Then since y and e commute and e is a projection, $\psi(ey) = e\psi(y)$. But $(1-\psi)|_{\sigma_{\mathscr{B}}(y)\backslash V} = 0$ and $(1-\psi)|_{\overline{U}} = 1$, and so by (3.19), $e = e(1-\psi(y))$. Altogether,

$$\psi(ey) = e\psi(y) = e(1-(1-\psi(y)) = e - e = 0\ .$$

More simply, let ψ be continuous on $\sigma_{\mathscr{B}}(y)$ with $\psi|_{\sigma_{\mathscr{B}}(y)\backslash U} = 0$. Then by (3.21), $(1-e)\psi(y) = 0$, but as above $\psi((1-e)y) = (1-e)\psi(y)$. Hence there is no spectrum of $\psi((1-e)y)$ outside $\sigma_{\mathscr{B}}(y)\backslash U$.

By Theorem 2.10 and the Spectral Invariance Theorem, putting $\tilde{y} = \lambda_1 e + (1-e)y$, we have $\sigma_{\mathscr{B}}(\tilde{y}) \subset \{\lambda_1\} \cup (\sigma_{\mathscr{B}}(y)\backslash U)$. Finally, $\|\tilde{y} - y\| = \|\lambda_1 e - ey\| \leq \epsilon$. □

Proof of Lin's Theorem. Let $\epsilon > 0$, and let y be normal in $\mathscr{B}$, and suppose that $\sigma_{\mathscr{B}}(y) \subset \Gamma_\epsilon$. Since $\sigma_{\mathscr{B}}(y)$ is contained in the disc of radius $\|y\|$, there are at most $2\|y\|(2\|y\|+1)/\epsilon^2$ edges of the elementary squares in Γ_ϵ whose midpoints intersect $\sigma_{\mathscr{B}}(y)$. That is, with Λ_ϵ defined as in (3.14), there are at most $2\|y\|(2\|y\|+1)/\epsilon^2$ points of Λ_ϵ within $\sigma_{\mathscr{B}}(y)$.

We claim that there is a normal $\tilde{y}$ such that $\|\tilde{y}\| \leq \|y\|$, $\|\tilde{y} - y\| \leq \epsilon$ and

$$\sigma_{\mathscr{B}}(\tilde{y}) \subset \Gamma_\epsilon \backslash \Lambda_\epsilon\ .$$

This is true because if $\lambda_0 \in \sigma_{\mathscr{B}}(y) \cap \Lambda_\epsilon$, we have the following alternative: Either λ_0 is an isolated point of $\sigma_{\mathscr{B}}(y)$, or it is not.

If λ_0 is an isolated point of $\sigma_{\mathscr{B}}(y)$, then we can find a continuous function $f : \Gamma_\epsilon \to \Gamma_\epsilon$ such that $f(\lambda) = \lambda$ except on a small neighborhood of λ_0, and such that λ_0 is not in the range of f. We may choose the neighborhood small enough that

$$\sup_{\lambda\in\Gamma_\epsilon}\{|f(\lambda) - \lambda|\} \leq \epsilon^3/(2\|y\|(2\|y\|+1)\ .$$

Moreover, we can always arrange that applying f does not increase the spectral radius. Then $\|f(y)\| \le \|y\|$, Then $\|f(y) - y\| \le \epsilon^3/(2\|y\|(2\|y\|+1)$, $f(y)$ is normal, $\sigma_{\mathscr{B}}(f(y)) \subset \Gamma_\epsilon \backslash \{\lambda_0\}$. In this way, we remove all points in λ_ϵ that are isolated points of the spectrum.

Now we apply Lemma 3.6 to remove all points in λ_ϵ that are isolated points of the spectrum, noting that we only affect the spectrum near the each such point of Λ_ϵ at each step. We may arrange that the shift in y at each step has norm no more than $\epsilon^3/(2\|y\|(2\|y\|+1)$.

At the end of at most $2\|y\|(2\|y\|+1)/\epsilon^2$ operations, each of which shifted y by at most $\epsilon^3/(2\|y\|(2\|y\|+1)$ in norm, we arrive at $\tilde{y}$ which is normal and has $\|\tilde{y} - y\| \le \epsilon$ and

$$\sigma_{\mathscr{B}}(\tilde{y}) \subset \Gamma_\epsilon \backslash \Lambda_\epsilon .$$

The set $\Gamma_\epsilon \backslash \Lambda_\epsilon$ is a disjoint union of open crosses, and there is a continuous function f on $\Gamma_\epsilon \backslash \Lambda_\epsilon$ that retracts each cross onto its center. That is, there is a continuous function $f : \Gamma_\epsilon \backslash \Lambda_\epsilon \to \tilde{\Lambda}_\epsilon$ with

$$\sup_{\lambda \in \Gamma_\epsilon \backslash \Lambda_\epsilon} \{|f(\lambda) - \lambda|\} \le \epsilon/2 \ .$$

Then $f(\tilde{y})$ is normal with finite spectrum and $\|f(\tilde{y}) - \tilde{y}\| \le \epsilon/2$. Consequently, $\|f(\tilde{y}) - y\| \le 3\epsilon/2$.

Combining this with Lemma 3.5, we see that for every normal element of $\mathscr{B}$, there is a normal element with finite spectrum arbitrarily close in norm, which is what we had to show. □

3.4 The Bott invariant and obstructions to commutativity

One might hope that one could extend Lin's Theorem to three or more matrices. That is, one might conjecture that for all $\epsilon > 0$, there is a $\delta > 0$ such that if $\{h_1, h_2, h_3\}$ is a set of $n \times n$ self adjoint matrices such that

$$\|[h_1, h_2]\| + \|[h_2, h_3]\| + \|[h_3, h_1]\| \le \delta \ ,$$

then there exists a set of three self adjoint commuting matrices $\{k_1, k_2, k_3\}$ such that

$$\|h_1 - k_1\| + \|h_2 - k_2\| + \|h_3 - k_3\| \le \epsilon \ .$$

This is false. Hastings and Loring, building on previous work, have shown the following:

3.7 THEOREM. *For all $j \in \frac{1}{2}\mathbb{N}$ there exists a set $\{h_1, h_2, h_3\}$ a set of self adjoint $(2j+1) \times (2j+1)$ matrices such that*

$$\|[h_1, h_2]\| + \|[h_2, h_3]\| + \|[h_3, h_1]\| \le \frac{1}{\sqrt{j(j+1)}} \ ,$$

and such that if $\{k_1, k_2, k_3\}$ is any set of commuting self adjoint $(2j+1) \times (2j+1)$ matrices, then

$$\|h_1 - k_1\| + \|h_2 - k_2\| + \|h_3 - k_3\| \ge \sqrt{1 - 4/\sqrt{j(j+1)}} \ .$$

3.8 DEFINITION. Let $\delta > 0$ be given. A δ *representation of the sphere* in $M_n(\mathbb{C})$ is a set $\{h_1, h_2, h_3\}$ of self adjoint $n \times n$ matrices such that

$$\|[h_1, h_2]\| \le \delta \ , \quad \|[h_2, h_3]\| \le \delta \quad \text{and} \quad \|[h_3, h_1]\| \le \delta \ , \tag{3.22}$$

and

$$\|1 - (h_1^2 + h_2^2 + h_3^2)\| \le \delta \ . \tag{3.23}$$

3.9 EXAMPLE. A set of three self-adjoint $n \times n$ matrices $\{s_1, s_2, s_3\}$ that satisfy

$$[s_1, s_2] = is_3 \ , \quad [s_2, s_3] = is_1 \quad \text{and} \quad [s_3, s_1] = is_2 \ . \tag{3.24}$$

is an n dimensional representation of the Lie algebra $\mathfrak{su}(2)$. The representation is *irreducible* in case there is no subspace of $\mathbb{C}^n$ that is invariant under each of s_1, s_2 and s_3. The simplest example is provided by the Pauli matrices, multiplied by 1/2:

$$s_1 = \frac{1}{2}\sigma_1 = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ , \qquad s_2 = \frac{1}{2}\sigma_2 = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \text{and} \quad s_3 = \frac{1}{2}\sigma_3 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \ .$$

For each $j \in \frac{1}{2}\mathbb{N}$ there is an irreducible representation by $(2j+1) \times (2j+1)$ matrices $\{s_1, s_2, s_3\}$ and these matrices satisfy $s_1^2 + s_2^2 + s_3^2 = j(j+1)1$, and they each have the same spectrum consisting of $\{-j, -j+1, \dots, j-1, j\}$. It is then easy to see that defining $h_j = (j(j+1))^{-1/2}s_j$, $j = 1, 2, 3$, we obtain a $(j(j+1))^{-1/2}$ representation of the sphere.

We now explain the representation theory on which this construction depends, partly for completeness, and partly because it will be useful for some calculations that follow.

Notice that for *any* representation, $s_3 s_1^2 = s_1 s_3 s_1 + i[s_3, s_1]s_1 = s_1^2 s_3 + i[s_3, s_1]s_1 + is_1[s_3, s_1]$. Using (3.24), this reduces to $[s_3, s_1^2] = -s_2 s_1 - s_1 s_2$. A similar calculation shows that $[s_3, s_2^2] = s_1 s_2 + s_2 s_1$. Altogether, $[s_3, (s_1^2 + s_2^2 + s_3^2)] = 0$, and by symmetry, $s_1^2 + s_2^2 + s_3^2$ commutes with s_1 and s_2 as well. In summary, defining the positive matrix s by

$$s^2 := s_1^2 + s_2^2 + s_3^2 \ , \tag{3.25}$$

$$[s_1, s^2] = [s_2, s^2] = [s_3, s^2] = 0 \ . \tag{3.26}$$

Suppose that $\{s_1, s_2, s_3\}$ is an irreducible n-dimnsional representation of $\mathfrak{su}(2)$. By (3.25), the eigenspaces of s^2 are invariant under each of s_1, s_2 and s_3. Since the representation is irreducible, it must be that s^2 is a multiple of the identitiy. Let μ temporarily denote this multiple, so that $s^2 = \mu 1$.

Define operators s_+ and s_- by

$$s_+ = s_1 + is_2 \quad \text{and} \quad s_- = s_1 - is_2 \ . \tag{3.27}$$

We compute $s_3(s_1 + is_2) = (s_1 + is_2)s_3 + ([s_3, s_1] + i[s_3, s_2]) = (s_1 + is_2)s_3 + (is_2 + s_1)$. That is $[s_3, s_+] = s_+$. Taking the adjoint, $[s_3, s_-] = -s_-$, and we have

$$[s_3, s_+] = s_+ \quad \text{and} \quad [s_3, s_-] = -s_- \ . \tag{3.28}$$

Therefore, if ζ is an eigenvector of s_3 with $s_3\zeta = \lambda s_3$,

$$s_3(s_+\zeta) = s_+(s_3\zeta) + s_+\zeta = (\lambda + 1)s_+\zeta \ .$$

That is, either $\lambda + 1$ is an eigenvlaue of s_3, or $s_+\zeta = 0$. In the same way we see that either $\lambda - 1$ is an eigenvalue of s_3 or else $s_-\zeta = 0$.

Now let ζ_1 be an eigenvector of s_3 with minimal eigenvalue. (This is a *least weight vector* in the language of representation theory.) Then $s_-\zeta_1 = 0$. Define vectors $\zeta_k = (s_+)^{k-1}\zeta$. Suppose that for some $m \in \mathbb{N}$, no vector in $\{\zeta_1, \dots, \zeta_m\}$ is zero. By what we have noted above, each is an eigenvector of s_3, and the successive eigenvlaues are all different, so that this set is orthogonal. Evidently, there is some least $m \in \mathbb{N}$ for which $(s_+)^{m+1} = 0$. Let m be this integer.

By construction, $s_+\zeta_k = \zeta_{k+1}$ for $k < m$, and $s_+\zeta_m = 0$. Next, we compute that $s_+s_- = s_1^2 + s_2^2 + s_3$ and $s_-s_+ = s_1^2 + s_2^2 - s_3$. Adding and subtracitng s_3^2,

$$s_+s_- = s^2 - s_3^2 + s_3 \quad \text{and} \quad s_-s_+ = s^2 - s_3^2 - s_3 \ . \tag{3.29}$$

Since each vector in $\{\zeta_1, \ldots, \zeta_m\}$ is an eigenvector of $s^2 - s_3^2 + s_3$, it is an eigenvector of s_-s_+ and of s_+s_-. For each $k = 2, \ldots, m$, ζ_k is a multiple of $s_+\zeta_{k-1}$. Hence $s_-\zeta_k$ is a mutliple of $s_-s_+\zeta_{k-1}$ which, by the above, is a multiple ζ_{k-1}. For $k = 1$, $s_-\zeta_k = 0$ since otherwise it would be an eigenvector of s_3 with an eigenvalue lower by one than the least eigenvalue. Hence the span of $\{\zeta_1, \ldots, \zeta_m\}$ is invariant under s_- as well as s_+ and s_3. Hence it is invariant under eauch of s_1, s_2 and s_3. Since the representation is irreducible, $m = n$ and the span is all of $\mathbb{C}^n$.

Now let λ be the least eigenvalue of s_3, and recall that μ denotes the single eienvalue of s^2. By construction, ζ_n is an eigenvector of s_3 with eigenvalue $\lambda + n - 1$. Since $s_-\zeta_1 = 0$ and $s_+\zeta_n = 0$, (3.29) gives us

$$0 = \mu - \lambda^2 + \lambda \quad \text{and} \quad 0 = \mu - (\lambda + n - 1)^2 - (\lambda + n - 1) \ .$$

Thus, $\lambda^2 + 2\lambda(n-1) + (n-1)^2 + \lambda + (n-1) = \lambda^2 - \lambda$ so that

$$\lambda 2n = -(n-1) - (n-1)^2 = -n(n-1) \ .$$

We obtain

$$\lambda = -\frac{n-1}{2} \quad \text{and} \quad \mu = \frac{n^2-1}{2} \ . \tag{3.30}$$

At this point it is traditional to introduce $j \in \frac{1}{2}\mathbb{N}$ by $j = \frac{n-1}{2}$, so that $n = 2j+1$, and then $s^2 = j(j+1)1$. The eigenvalues of s_3 are then given, in increasing order, by $\{-j-(j-1), \ldots, j-1, j\}$, and by symmetry, s_1 and s_2 have the same spectrum.

So far we have seen that if for some $j \in \frac{1}{2}\mathbb{N}$ there is a $2j+1$ dimensional irreducible representation of $\mathfrak{su}(2)$, then there is an orthonormal basis $\{eta_{-j}, \ldots, \eta_j\}$ of $\mathbb{C}^{2j+1}$ such that $s_3\eta_k = k\eta_k$ for each $k = -2j-1, \ldots, 2j+1$. This gives us the (diagonal) form of the matrix for s_3 in this basis.

Moreover, we have seen that for all $k = -2j-1, \ldots, 2j$, $a_+\eta_k = t_k\eta_k$ for some poisitive mutiple t, while $a_+\eta_{2j+1} = 0$. We compute

$$t_k^2 = \langle \eta_k, s_-s_+\eta_k \rangle = \langle \eta_k, (s^2 - s_3^2 - s_3)\eta_k \rangle = j(j+1) - k(k+1) \ .$$

That is,

$$s_+\eta_k = \sqrt{j(j+1) - k(k+1)}\eta_{+1} \qquad \text{for all} \qquad k = -2j-1, \ldots, 2j+1 \ . \tag{3.31}$$

This gives us the form of the matrix representing s_+ in this basis, and taking the hermitain conjugate we get the matrix that represents s_-.

Finally, it is easy to check that the matrices determine a triple $\{s_1, s_2, s_3\}$ of self adjoint $(2j+1) \times (2j+1)$ matrices that satisfy (3.24). Hence, for each j, there is a representation of $\mathfrak{su}(2)$ by $(2j+1) \times (2j+1)$ matrices, and any two such representations are unitarily equivalent. Any such representation is called a *spin j representation.*

Each such representation gives rise to a natural example of a $(j(j+1))^{-1/2}$ representation of the sphere: Since

$$s_1^2 + s_2^2 + s_3^2 = j(j+1)1 \ ,$$

we define $h_j = (j(j+1))^{-1/2}s_j$, $j = 1,2,3$, $\{h_1, h_2, h_3\}$, and this provides a $(j(j+1))^{-1/2}$ representation of the sphere.

We will show, following Hastings and Loring, that if $\{k_1, k_2, k_3\}$ is a set of three commting self adjoint $(2j+1) \times (2j+1)$ matrices, then for the spin j representation of $\mathfrak{su}(2)$,

$$\|h_1 - k_k\| + \|h_2 - k_2\| + \|h_3 - k_3\| \geq \sqrt{1 - 2/j}\ .$$

The method involves a topological invariant, the Bott invariant that we now define.

3.10 DEFINITION (Bott invariant)**.** For any set of three $n \times n$ hermitian matrices h_1, h_2 and h_3, define the $2n \times 2n$ matrix $b(h_1, h_2, h_3)$ by

$$b(h_1, h_2, h_3) = \sum_{j=1}^{3} \sigma_j \otimes h_j = \left[\begin{array}{cc} h_3 & h_1 - ih_2 \\ h_1 + ih_2 & -h_3 \end{array}\right] .$$

Note that $b(h_1, h_2, h_3)$ is self adjoint so that all of its eigenvalues are real. Let $N_+(h_1, h_2, h_3)$ be the number of strictly positive eigenvalues of $b(h_1, h_2, h_3)$, and let $N_-(h_1, h_2, h_3)$ be the number of strictly negative eigenvalues of $b(h_1, h_2, h_3)$. Suppose that 0 is not an eigenvalue of $b(h_1, h_2, h_3)$, so that $N_+(h_1, h_2, h_3) + N_-(h_1, h_2, h_3) = 2n$. Then $N_+(h_1, h_2, h_3) - N_-(h_1, h_2, h_3) = 2n - 2N_-(h_1, h_2, h_3)$ is an even intteger, so that

$$\mathrm{bott}(h_1, h_2, h_3) := \frac{1}{2}[N_+(h_1, h_2, h_3) - N_-(h_1, h_2, h_3)]$$

is an integer. This integer is the *Bott invariant* of $\{h_1, h_2, h_3\}$. Note that $\mathrm{bott}(h_1, h_2, h_3)$ is only defined when 0 is not an eigenvalue of $b(h_1, h_2, h_3)$.

Note that the three matrices σ_1, σ_2 and σ_3 are have the same spectrum, namely $\{-1, 1\}$. Then for any self adjoint $h \in M_n(\mathbb{C})$, $\sigma_1 \otimes h$, $\sigma_2 \otimes h$ and $\sigma_3 \otimes h$ all have the same eigenvalues, namely $\{\pm\lambda_1, \dots, \pm\lambda_n\}$ where $\{\lambda_1, \dots, \lambda_n\}$ is the set of eigenvalues of h. Hence

$$\|\sigma_1 \otimes h\| = \|\sigma_2 \otimes h\| = \|\sigma_3 \otimes h\| = \|h\| .$$

It follows that for any self adjoint triple $\{h_1, h_2, h_3\}$,

$$\|b(h_1, h_2, h_3)\| \leq \|h_1\| + \|h_2\| + \|h_3\| . \tag{3.32}$$

Suppose that h_1, h_2 and h_3 are three commuting Hermitian $n \times n$ matrices. Then there is a unitary $n \times n$ matrix u such that $k_j := u^* h_j u$ is diagonal for each $j = 1, 2, 3$. Evidently

$$\left[\begin{array}{cc} u^* & 0 \\ 0 & u^* \end{array}\right] b(h_1, h_2, h_3) \left[\begin{array}{cc} u & 0 \\ 0 & u \end{array}\right] = b(k_1, k_2, k_3) .$$

Therefore,

$$\mathrm{bott}(h_1, h_2, h_3) = \mathrm{bott}(k_1, k_2, k_3) .$$

Let $\alpha_1, \dots, \alpha_n$, $\beta_1, \dots, \beta_n$, and $\gamma_1, \dots, \gamma_n$ be the diagonal entries of k_1, k_2 and k_3 respectively. Then $b(k_1, k_2, k_3)$ is unitarily equivalent to

$$\bigoplus_{\ell=1}^{n} \left[\begin{array}{cc} \gamma_\ell & \alpha_\ell - i\beta_\ell \\ \alpha_\ell + i\beta_\ell & -\gamma_\ell \end{array}\right] . \tag{3.33}$$

A simple computation shows that $\begin{bmatrix} \gamma_\ell & \alpha_\ell - i\beta_\ell \\ \alpha_\ell + i\beta_\ell & -\gamma_\ell \end{bmatrix}^2 = (\alpha_\ell^2 + \beta_\ell^2 + \gamma_\ell^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Since the trace of $\begin{bmatrix} \gamma_\ell & \alpha_\ell - i\beta_\ell \\ \alpha_\ell + i\beta_\ell & -\gamma_\ell \end{bmatrix}$ is zero, it then follows that the eigenvalues are $\pm\sqrt{\alpha_\ell^2 + \beta_\ell^2 + \gamma_\ell^2}$. Thus, as long no non-zero vector is in the null space of each of h_1, h_2 and h_3, each of the blocks in (3.33) has one strictly positive eigenvalue and one strictly negative eigenvalue. In particular, this is the case if each of h_1, h_2 and h_3 are invertible. It follows that in this case, $\mathrm{bott}(h_1, h_2, h_3) = 0$. We have proved:

3.11 LEMMA. *Let h_1, h_2 and h_3 be three commuting invertible Hermitian $n \times n$ matrices. Then* $\mathrm{bott}(h_1, h_2, h_3) = 0$.

3.12 EXAMPLE. Let $\{h_1, h_2, h_3\}$ be the $(j(j+1))^{-1/2}$ representation of the sphere provided by a spin j representation of $\mathfrak{su}(2)$. Let $\{\eta_{-2j-1}, \ldots, \eta_{2j+1}\}$ be the corresponding sequence of eigenvectors of s_3 and hence h_3. Since, using the notation of the previous example,

$$b(h_1, h_2, h_3) = \frac{1}{\sqrt{j(j+1)}} \begin{bmatrix} s_3 & s_- \\ s_+ & -s_3 \end{bmatrix} ,$$

is is easy to see that the vectors of the form $(\eta_k, \pm\eta_{k+1})$, $-2j-1, \ldots, 2j$, togehter with $(0, \eta-2j-1)$ and $(\eta_{2j+1}, 0)$. are a set of $2n$ orthonormal eigenvalues of $b(h_1, h_2, h_3)$. A simple calculaution show that the pairs $(\eta_k, \pm\eta_{k+1})$ contribute a positive and a negative eigenvalue each, while the two special case eigenvectors have positive eigenvalues. Hence

$$\mathrm{bott}(h_1, h_2, h_3) = 1$$

for all j.

Next, we show that the Bott invariant is defined for all δ representations of the sphere with $\delta < 1/4$.

3.13 LEMMA. *Let $\{h_1, h_2, h_3\}$ be a δ-representation of the sphere in $M_n(\mathbb{C})$ with $\delta < 1/4$. Then*

$$\sigma(b(h_1, h_2, h_3)) \subset [-\sqrt{1+4\delta}, -\sqrt{1-4\delta}] \cup [\sqrt{1-4\delta}, \sqrt{1+4\delta}] \ . \tag{3.34}$$

Moreover, if $\{k_1, k_2, k_3\}$ is any triple of self adjoint operators with

$$\gamma = \|h_1 - k_1\| + \|h_2 - k_2\| + \|h_3 - k_3\| < \sqrt{1-4\delta} \ , \tag{3.35}$$

then for all $t \in [0, 1]$,

$$\sigma(s((1-t)h_1 + tk_1, (1-t)h_2 + tk_2, (1-t)h_3 + tk_3)) \subset \\ [-\gamma - \sqrt{1+4\delta}, \gamma - \sqrt{1-4\delta}] \cup [\sqrt{1-4\delta} - \gamma, \sqrt{1+4\delta} + \gamma] \ . \tag{3.36}$$

and consequently, $\mathrm{bott}((1-t)h_1 + tk_1, (1-t)h_2 + tk_2, (1-t)h_3 + tk_3))$ *is well-defined for all $t \in [0, 1]$.*

Proof. We compute that

$$(s(h_1,h_2,h_3))^2 = 1\otimes(h_1^2+h_2^2+h_3^2)+\sigma_3\otimes i[h_1,h_2]+\sigma_1\otimes i[h_2,h_3]+\sigma_2\otimes i[h_3,h_1]\ .$$

Therefore, $\|s(h_1,h_2,h_3))^2-1\|\le 4\delta$, and then by the Spectral Mapping Lemma,

$$\sigma(s(h_1,h_2,h_3))\subset\{\ t\in\mathbb{R}\ :\ |t^2-1|<4\delta\ \}\ .$$

Next, note that by (3.32),

$$\|b((1-t)h_1+tk_1,(1-t)h_2+tk_2,(1-t)h_3+tk_3))-b(h_1,h_2,h_3))\| = \\ t\|b(k_1-h_1,k_2-h_2,k_3-h_3))\|\le t\gamma\ . \quad (3.37)$$

Then (3.34) and Theorem 2.17 yield (3.39), and $\mathrm{bott}((1-t)h_1+tk_1,(1-t)h_2+tk_2,(1-t)h_3+tk_3))$ is well-defined for all $t\in[0,1]$. □

3.5 The Bott invarinat as a trace function

Recall that for an $n\times n$ matrix a, the *trace* of a, $\mathrm{Tr}[a]$, is defined by

$$\mathrm{Tr}[a]=\sum_{j=1}^n a_{i,i} \quad (3.38)$$

where $a_{i,j}$ denotes the i,j entry of a.

A simple computation shows that for any $a\in M_n(\mathbb{C})$, and any invertible $b\in M_n(\mathbb{C})$, $\mathrm{Tr}[b^{-1}ab]=\mathrm{Tr}[a]$. Let $\{\eta_1,\dots,\eta_n\}$ be any orthonormal basis of $\mathbb{C}^n$, and let $\{\chi_1,\dots,\chi_n\}$ be the standard basis. Let u be the unitary matrix with $u\chi_j=\eta_j$ for $j=1,\dots,n$. Then

$$\mathrm{Tr}[a]=\mathrm{Tr}[u^*au]=\sum_{j=1}^n\langle\chi_i,u^*au\chi_i\rangle=\sum_{j=1}^n\langle\eta_j,b^{-1}ab\eta_j\rangle\ ,$$

showing that the trace may be computed at the sum of the diagonal elements in any ortonormal basis. If a is self adjoint, there is an orthonormal basis consisting of eigenvectors of a; $a\eta_j=\lambda_j\eta_j$ for each $j=1,\dots,n$. Then evidently $\mathrm{Tr}[a]-\sum_{j=1}^n\lambda_j$. The function $a\mapsto\mathrm{Tr}[a]$ is evidently continuous.

Now let $\epsilon>0$ be given, and let f_ϵ be any continuous function form $\mathbb{R}$ to $[-1,1]$ such that $f(t)=-1$ for $t\le-\epsilon$ and $f(t)=1$ for $t\ge\epsilon$.

Let $a\in M_n(\mathbb{C})$ be self adjoint and such that $(-\epsilon,\epsilon)\cap\sigma_{\mathscr{A}}(a)=\emptyset$. Let $\{\eta_1,\dots,\eta_n\}$ be an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of a with $a\eta_j=\lambda_j\eta_j$ for each $j=1,\dots,n$. Define

$$N_+(a)=\sum_{j=1} n1_{(0,\infty)}(\lambda_j)\quad\text{and}\quad N_-(a)=\sum_{j=1} n1_{(-\infty,0)}(\lambda_j)\ .$$

Then by considering a sequence of polynomial approximations of f_ϵ on $[-\|a\|,\|a\|]$, we have that

$$\mathrm{Tr}[f(a)]=\sum_{j=1}^n\langle\eta_j,f(a)\eta_j\rangle=\sum_{j=1}^n f_\epsilon(\lambda_j)=N_+(a)-N_-(a)\ .$$

It follows that when $\{h_1,h_2,h_3\}$ is a δ representation of the sphere in $M_n(\mathbb{C})$ with $\delta<1/4$,

$$\mathrm{bott}(\{h_1,h_2,h_3\})=\mathrm{Tr}[f_{1-4\delta}(s(h_1,h_2,h_3))]\ .$$

Proof of Theorem 3.7. Consider the $1/\sqrt{j(j+1)}$ representation of the sphere associated to the spin j representation of $\mathfrak{su}(2)$. By Lemma 3.13, if $\{k_1,k_2,k_3\}$ is any triple of self adjoint $(2j+1)\times(2j+1)$ matrices with

$$\gamma = \|h_1-k_1\|+\|h_2-k_2\|+\|h_3-k_3\| < \sqrt{1-4/\sqrt{j(j+1)}}\ , \tag{3.39}$$

then for all $t\in[0,1]$,

$$\begin{aligned}&\sigma(s((1-t)h_1+tk_1,(1-t)h_2+tk_2,(1-t)h_3+tk_3))\subset\\&[-\gamma-\sqrt{1+4\sqrt{j(j+1)}},\gamma-\sqrt{1-4\sqrt{j(j+1)}}]\cup[\sqrt{1-4\sqrt{j(j+1)}}-\gamma,\sqrt{1+4\sqrt{j(j+1)}}+\gamma]\ .\end{aligned} \tag{3.40}$$

and consequently, for any $\epsilon < \sqrt{1-4\sqrt{j(j+1)}}-\gamma$,

$$\mathrm{bott}((1-t)h_1+tk_1,(1-t)h_2+tk_2,(1-t)h_3+tk_3))-\mathrm{Tr}[f_\epsilon((1-t)h_1+tk_1,(1-t)h_2+tk_2,(1-t)h_3+tk_3))\ ,$$

with f_ϵ defined as in the paragraphs above. Then the right hand side is a continuous integer valued function of t, and so

$$\mathrm{bott}(k_1,k_2,k_3)=\mathrm{bott}(h_1,h_2,h_3)=1\ .$$

Therefore, $\{k_1,k_2,k_3\}$ cannot be a commuting triple.

□

4 Operators on Hilbert space

4.1 Topologies on $\mathscr{B}(\mathscr{H})$

Let $\mathscr{H}$ be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_{\mathscr{H}}$ and norm $\|\cdot\|_{\mathscr{H}}$, and let $\mathscr{B}(\mathscr{H})$, as usual, denote the C^*-algebra of bounded linear operators on $\mathscr{H}$. There are two important non-metric topologies in $\mathscr{B}(\mathscr{H})$, waker than the norm topology, that are essential to what follows.

4.1 DEFINITION (Strong and weak operator topologies)**.** The *strong operator topology* on $\mathscr{B}(\mathscr{H})$ is the weakest topology such that for each $\xi\in\mathscr{H}$, the function $a\mapsto a\xi$ from $\mathscr{B}(\mathscr{H})$ to $\mathscr{H}$ is continuous with the usual norm topology on $\mathscr{H}$. The *weak operator topology* on $\mathscr{B}(\mathscr{H})$ is the weakest topology such that for each $\xi,\zeta\in\mathscr{H}$, that function $a\mapsto\langle\zeta,a\xi\rangle_{\mathscr{H}}$ is continuous from $\mathscr{B}(\mathscr{H})$ to $\mathbb{C}$.

It follows from the definitions that a basic set of neighborhoods of 0 for the strong operator topology is given by the sets

$$U_{\epsilon,\xi_1,\dots,\xi_n}=\{a\in\mathscr{B}(\mathscr{H})\ :\ \|a\xi_j\|_{\mathscr{H}}<\epsilon\quad\text{for}\quad j=1,\dots,n\ \} \tag{4.1}$$

where $\epsilon>0$ and $\xi_1,\dots,x_n\in\mathscr{H}$. Likewise, it follows that a basic set of neighborhoods of 0 for the weak operator topology is given by the sets

$$V_{\epsilon,\zeta_1,\dots,\zeta_n,\xi_1,\dots,\xi_n}=\{a\in\mathscr{B}(\mathscr{H})\ :\ |\langle\zeta_j,a\xi_j\rangle_{\mathscr{H}}|<\epsilon\quad\text{for}\quad j=1,\dots,n\ \} \tag{4.2}$$

$\epsilon > 0$ and $\zeta_1, \ldots, \zeta_n, \xi_1, \ldots, x_n \in \mathscr{H}$. Note that both topologies are evidently Hausdorff.

It is clear that for each $\xi \in \mathscr{H}$, $a \mapsto a\xi$ is continuous in the norm topology on $\mathscr{B}(\mathscr{H})$, so that the norm topology is stronger than the strong operator topology. Furthermore, since for all $\zeta \in \mathscr{H}$, $\xi \mapsto \langle \zeta, \xi \rangle_{\mathscr{H}}$ is continuous on $\mathscr{H}$, the function $a \mapsto \langle \zeta, a\xi \rangle_{\mathscr{H}}$ is continuous in the strong operator topology on $\mathscr{B}(\mathscr{H})$, being the composition of continuous functions, and hence the strong operator topology is stronger than the weak operator topology.

The following proposition shows that the norm topology is *strictly stronger* than the strong operator topology, which is in turn *strictly stronger* than the weak operator topology.

4.2 PROPOSITION (Continuity of the norm and adjoint)**.** *Let $\mathscr{H}$ be an infinite dimensional Hilbert space. Then:*

(1) The function $a \mapsto \|a\|$ from $\mathscr{B}(\mathscr{H})$ to $\mathbb{R}_+$ is continuous in the norm topology, but is only lower semicontinuous in the strong ad weak operator topologies.

(2) The function $a \mapsto a^*$ is continuous from $\mathscr{B}(\mathscr{H})$ to $\mathscr{B}(\mathscr{H})$ in the norm and the weak operator topologies, but not in the strong operator topology.

Proof. Let $\{\zeta_j\}$ be an orthonormal sequence in $\mathscr{H}$. For each $n \in \mathbb{N}$, let p_n denote the orthogonal projection onto the span of $\{\zeta_1, \ldots, \zeta_n\}$. Then for all $\xi \in \mathscr{H}$, $\lim_{n\to\infty} \|p_n\xi\| = 0$ by Bessel's inequality, so that $\lim_{n\to\infty} p_n = 0$ in the strong operator topology. However, for $n \neq m$, $\|p_n - p - m\| = 1$, so that the sequence $\{p_n\}$ is not even Cauchy in the norm topology. Hence the norm is discontinuous in the strong operator topology, and hence also in the weak operator topology.

To see that the norm is lower semicontinuous in these topologies, it suffices to show that the sub-level sets $\{\ a \in \mathscr{B}(\mathscr{H})\ :\ \|a\| \leq t\ \}$ are closed for each $t > 0$. Fix $t > 0$ and b in the closure of $\{\ a \in \mathscr{B}(\mathscr{H})\ :\ \|a\| \leq t\ \}$. Then for each unit vector $\xi \in \mathscr{H}$, and each $n \in \mathbb{N}$ there is an $a_n \in \{\ a \in \mathscr{B}(\mathscr{H})\ :\ \|a\| \leq t\ \}$ such that $b - a_n \in U_{1/n,\xi}$, which means that $\|(b - a_n)\xi\| < 1/n$. This means that $\|b\xi\| \leq \|a_n\| + 1/n \leq t + 1/n$. Since n is arbitrary, $\|b\xi\| \leq t$. Then since ξ is an arbitrary unit vector in $\mathscr{H}$, $\|b\| \leq t$. This proves the closure in the strong operator topology, and a very similar argument proves the closure for the weak operator topology.

For the second part, since every infinite dimensional Hilbert space contains a copy of ℓ_2, the Hilbert space of all square summable functions from $\mathbb{N}$ to $\mathbb{C}$, we may suppose without loss of generality that $\mathscr{H} = \ell_2$. Define the shift operator $a \in \mathscr{B}(\mathscr{H})$ by

$$(a\zeta)_j = \begin{cases} \zeta_{j-1} & j \geq 2 \\ 0 & j = 1 \end{cases}$$

Evidently, for all ζ, $\|a\zeta\|_{\mathscr{H}} = \|\zeta\|_{\mathscr{H}}$. The adjoint is given by $(a^*\zeta)_j = \zeta_{j+1}$ for all $j \in \mathbb{N}$. Therefore, $\|a^*\zeta\|^2_{\mathscr{H}} = \sum_{j=2}^{\infty} |\zeta_j|^2 = \|\zeta\|^2 - |\zeta_1|^2$. It follows that for all $\zeta \in \mathscr{H}$,

$$\lim_{n\to\infty} \|(a^n)^*\zeta\|_{\mathscr{H}} = 0 \quad \text{while} \quad \|a^n\zeta\|_{\mathscr{H}} = \|\zeta\|_{\mathscr{H}}\ .$$

Hence the sequence $\{(a^n)^*\}$ converges to zero in the strong operator topology, but the sequence $\{a^n\}$ does not. Since $\{a^n\} = \{(a^n)^{**}\}$ this shows that the involution is not continuous in the strong operator topology.

The continuity of the involution is obvious in the norm topology since the involution is an isometry, and in the weak operator topology it follows from the fact that

$$(V_{\epsilon,\zeta_1,\ldots,\zeta_n,\xi_1,\ldots,\xi_n,})^* = V_{\epsilon,\xi_1,\ldots,\xi_n,\zeta_1,\ldots,\zeta_n}\ .$$

□

As far as sequences are concerned, a sequence $\{a_n\}$ in $\mathscr{B}(\mathscr{H})$ converges to $a \in \mathscr{B}(\mathscr{H})$ in the strong operator topology if and only if for all $\xi \in \mathscr{H}$, $\lim_{n\to\infty} a_n\xi = a\xi$, and likewise, converges to $a \in \mathscr{B}(\mathscr{H})$ in the weak operator topology if and only if for all $\zeta, \xi \in \mathscr{H}$, $\lim_{n\to\infty}\langle \zeta, a_n\xi\rangle_{\mathscr{H}} = \langle a\xi\rangle_{\mathscr{H}}$. A sequence $\{a_n\}$ in $\mathscr{B}(\mathscr{H})$ is a *Cauchy sequence for the weak operator topology* in case for every basic open neighborhood $V_{\epsilon,\zeta_1,\dots,\zeta_n,\xi_1,\dots,\xi_n}$ of 0, $a_m - a_\ell \in V_{\epsilon,\zeta_1,\dots,\zeta_n,\xi_1,\dots,\xi_n}$ for all but finitely many ℓ, m. Cauchy sequences for the strong operator topology are defined analogously.

4.3 THEOREM. *Let $\{a_n\}$ be a Cauchy sequence for the weak operator topology. Then $\{\|a_n\|\}$ is a bounded sequence, and there exists an $a \in \mathscr{B}(\mathscr{H})$ with $\|a\| \le sup_{n\in\mathbb{N}}\{\|a_n\|\}$ and such that $\lim_{n\to\infty} a_n = a$ in the weak operator topology. Moreover, the analogous statement for the strong operator topology is also true.*

Proof. Let $\{a_n\}$ be a Cauchy sequence for the weak operator topology. We first show that $\{\|a_n\|\}$ is a bounded sequence. To see this, note that for each $\zeta, \xi \in \mathscr{H}$, $\{\langle \zeta, a_n\xi\rangle_{\mathscr{H}}\}$ is a Cauchy sequence in $\mathbb{C}$, and hence convergent and bounded. Thus, if we define the sets $C_m \subset \mathscr{H} \times \mathscr{H}$ by

$$C_m = \{(\zeta, \xi \in \mathscr{H} \times \mathscr{H} \ : \ \sup_{n\in\mathbb{N}} |\langle \zeta, a_n\xi\rangle_{\mathscr{H}}| \le m \ \}$$

we have that $\cup_{m\in\mathbb{N}} C_m = \mathscr{H} \times \mathscr{H}$. If $\{(\zeta_k, \xi_k)\}$ is a convergent sequence in C_m with limit (ζ, ξ), then for all n,

$$|\langle \zeta, a_n\xi\rangle_{\mathscr{H}}| = \lim_{k\to\infty} |\langle \zeta_k, a_n\xi_k\rangle_{\mathscr{H}}| \le m$$

so that C_m is closed. Since $\mathscr{H} \times \mathscr{H}$ with the product metric is a complete metric space, by Baire's Theorem, for at least one $m \in \mathbb{N}$, C_m contains an open set, and then it is clear that $\{\|a_n\|\}$ is a bounded sequence.

Now let $L = \sup_{n\in\mathbb{N}}\{\|a_n\|\}$, and for all $\zeta, \xi \in \mathscr{H}$, define $q(\zeta, \xi) = \lim_{n\to\infty}\langle \zeta, a_n\xi\rangle_{\mathscr{H}}$, which exists since the sequence on the right is Cauchy in $\mathbb{C}$. It is easy to see that $\zeta, \xi \mapsto q(\zeta, \xi)$ is a sesquilinear form on $\mathscr{H}$, with

$$|q(\zeta, \xi)| \le L\|\zeta\|_{\mathscr{H}}\|\xi\|_{\mathscr{H}} \ .$$

For each $\xi \in \mathscr{H}$, the map $\zeta \mapsto q(\zeta, \xi)$ is a conjugate linear functional on $\mathscr{H}$, and hence by the Riesz Representation Theorem, there is a uniquely determined vector $\eta_\xi \in \mathscr{H}$ such that $q(\zeta, \xi) = \langle \zeta, \eta_\xi\rangle_{\mathscr{H}}$ for all $\zeta \in \mathscr{H}$, and $\|\eta_\xi\|_{\mathscr{H}} \le r\|\xi\|_{\mathscr{H}}$. Since q is sesquilinear, the map $\xi \mapsto \eta_\xi$ is linear, and thus there exists $a \in \mathscr{B}(\mathscr{H})$ such that $\|a\| \le L$ and $\eta_\xi = a\xi$ for all $\xi \in \mathscr{H}$. It now follows that for each $\zeta, \xi \in \mathscr{H}$, $\lim_{n\to\infty}\langle \zeta, a_n\xi\rangle_{\mathscr{H}} = \langle \zeta, a\xi\rangle_{\mathscr{H}}$, and hence that $\lim_{n\to\infty} a_n = a$ in the weak operator topology. The corresponding proof for the strong operator topology is easier, and is left as an exercise.

□

We next claim that the strong operator topology is not metrizable when $\mathscr{H}$ is infinite dimensional. The basic open set $U_{\epsilon,\xi_1,\dots,\xi_n}$ contains all $a \in \mathscr{B}(\mathscr{H})$ with $a\xi_j = 0$ for each $j = 1 \dots, n$. If $\mathscr{H}$ is infinite dimensional, then there is a non-trivial subspace of $\mathscr{B}(\mathscr{H})$ contained in every $U_{\epsilon,\xi_1,\dots,\xi_n}$, and hence in every open set about the origin.

For each $n \in \mathbb{N}$, the set $C_n := \{ \ a \in \mathscr{B}(\mathscr{H}) \ : \ \|a\| \le n\}$ is closed in the strong operator topology by the lower semicontinuity of the norm. By what we have just said, each C_n is nowhere dense, since no ball can contain a non-trivial subspace. Since evidently $\mathscr{B}(\mathscr{H}) = \cup_{n=1}^{\infty} C_n$, it follows that $\mathscr{B}(\mathscr{H})$ is a countable union of closed, nowhere dense sets in the strong operator topology. Suppose the strong operator topology were metrizable. Then by Theorem refwscom, $\mathscr{B}(\mathscr{H})$ equipped with this topology would be a complete metric space. Baire's Theorem says that a complete metric space is never the countable union of closed nowhere dense sets, so the strong topology on $\mathscr{B}(\mathscr{H})$ cannot be metrized. A similar argument applies to the weak operator topology. However, as we show next, *The relative weak and strong operator topologies on bounded subsets of* $\mathscr{B}(\mathscr{H})$ *are metrizable when* $\mathscr{H}$ *is separable.*

4.4 THEOREM. *For* $r > 0$*, let* $\overline{B_r}$ *denote the closed unit ball of radius* r *in* $\mathscr{B}(\mathscr{H})$*. That is,* $\overline{B_r} = \{\ a \in \mathscr{B}(\mathscr{H})\ :\ \|a\| \leq r\ \}$*. Then there are metrics* ρ_w *and* ρ_s *on* $\overline{B_r}$ *such that the metric topologies are equivalent to the relative weak and strong operator topologies respectively, and such that* $(\overline{B_r}, \rho_w)$ *and* $(\overline{B_r}, \rho_s)$ *and are complete metric spaces.*

Proof. Let $\{\eta_j\}$ be any sequence of unit vectors that is dense in the unit sphere of $\mathscr{H}$, For all $a, b \in \mathscr{B}(\mathscr{H})$, define

$$\rho_s(a,b) = \sum_{j=1}^{\infty} 2^{-j} \|(a-b)\eta_j\|_{\mathscr{H}} \quad \text{and} \quad \rho_w(a,b) = \sum_{j,k=1}^{\infty} 2^{-j-k} |\langle \eta_k (a-b)\eta_j \rangle_{\mathscr{H}}| \ . \tag{4.3}$$

It is easy to verify that these are indeed metrics.

We now show that the relative strong operator topology on $\overline{B}_r$ coincides with the metric topology on $\overline{B_r}$ induced by the metric ρ_s. First, we first show that for every $t > 0$, $\{\ a\ :\ \rho_s(a,0) < t\ \}$ contains a neighborhood of 0 in the relative strong operator topology . Choose n so that $r2^{-n} < t/2$. Then for $b \in U_{t/2,\eta_1,\dots,\eta_n} \cap \overline{B_r}$,

$$\rho_s(b,0) = \sum_{j=1}^{\infty} 2^{-t} \|b\eta_j\| \leq \sum_{j=1}^{n} 2^{-j} \frac{t}{2} + \sum_{j=n+1}^{\infty} r \leq t$$

and consequently, $U_{t/2,\eta_1,\dots,\eta_n} \cap \overline{B_r} \subset \{\ a\ :\ \rho_s(a,0) < t\ \}$.

We next show that every basic strong operator topology neighborhood $U_{\epsilon,\xi_1,\dots,\xi_m}$ contains an open ball about 0 in the relative metric topology. By decreasing epsilon as necessary, we may suppose that ξ_j is a unit vector for each j. Choose $\{\eta_{j_1}, \dots, \eta_{j_m}\}$ such that $\|\eta_{j_k} - \xi_k\| < \epsilon/2$ for $k = 1, \dots, m$. Let $M = \max\{j_1, \dots, j_M\}$. Then for $b \in \overline{B_r} \cap \{\ a\ :\ \rho_s(a,0) < 2^{-M}\epsilon\ \}$, $\|b\eta_j\| \leq \epsilon$ for each $j = 1, \dots, m$, and consequently $b \in U_{\epsilon,\xi_1,\dots,\xi_m}$. A similar argument shows that on each $\overline{B_r}$, the relative weak operator topology is metrizable. □

We shall be especially concerned with bounded subsets of the self adjoint elements of $\mathscr{B}(\mathscr{H})$, for which there is an even simpler description of the relative weak operator topology, for which there is an even simpler criterion for weak convergence:

$$\lim_{n\to\infty} \langle \eta_j a_n \eta_j \rangle_{\mathscr{H}} = \langle \eta_j a \eta_j \rangle_{\mathscr{H}} \quad \text{for all} \quad j \in \mathbb{N} \ .$$

4.5 LEMMA (Polarization identify)**.** *Let* $a \in \mathscr{B}(\mathscr{H})$ *be self adjoint. Then for all* ζ *and* ξ *in* $\mathscr{H}$*,*

$$\begin{aligned} \langle \zeta, a\xi \rangle_{\mathscr{H}} = &\frac{1}{4} \left[\langle (\zeta+\xi), a(\zeta+\xi) \rangle_{\mathscr{H}} - \langle (\zeta-\xi), a(\zeta-\xi) \rangle_{\mathscr{H}} \right] \\ &- \frac{i}{4} \left[\langle (\zeta+i\xi), a(\zeta+i\xi) \rangle_{\mathscr{H}} - \langle (\zeta-i\xi), a(\zeta-i\xi) \rangle_{\mathscr{H}} \right] \ . \end{aligned}$$

Proof. This is a direct computation. □

4.6 REMARK. *It follows that for the relative weak operator topology on the self adjoint elements of $\mathscr{B}(\mathscr{H})$, a basic set of neighborhoods at the origin is given by the sets*

$$V_{\epsilon,\xi_1,\dots,\xi_n,} = \{a \in \mathscr{B}(\mathscr{H}) \ : \ |\langle \xi_j, a\xi_j\rangle_{\mathscr{H}}| < \epsilon \quad \text{for} \quad j = 1,\dots,n\ \} \tag{4.4}$$

$\epsilon > 0$ and $\xi_1, \dots, x_n \in \mathscr{H}$.

4.7 THEOREM (Continuous linear functions for the strong operator topology)**.** *Let $\mathscr{H}$ be a Hilbert space, and let φ be a linear functional on $\mathscr{B}(\mathscr{H})$ that is continuous in the strong operator topology. Then there exists $n \in \mathbb{N}$ and two sets of vectors $\{\zeta_1, \dots, \zeta_n\}$ and $\{\xi_1, \dots, \xi_n\}$ such that for all $a \in \mathscr{B}(\mathscr{H})$,*

$$\varphi(a) = \sum_{j=1}^{n} \langle \zeta_j, a\xi_j\rangle_{\mathscr{H}} \ . \tag{4.5}$$

Evidently, every such linear functional is weakly continuous, and hence every strongly continuous linear functional is weakly continuous. Consequently, a convex subset of $\mathscr{B}(\mathscr{H})$ is strongly closed if and only if it is weakly closed.

Proof. If φ is strongly continuous, then $\varphi^{-1}(\{\lambda \ : \ |\lambda| < 1\})$ contain a neighborhood of 0 in $\mathscr{B}(\mathscr{H})$. Thus, there exists an $\epsilon > 0$ and a set of n vectors $\xi_1, \dots, \xi_n$, which without loss of generality we may assume to be orthonormal, such that if $\|a\xi_j\| < \epsilon$ for $j = 1, \dots, n$, $|\varphi(a)| < 1$. Note that if $a\xi_j = 0$ for $j = 1, \dots, n$, then $t > 0$, $\|ta\xi_j\| < \epsilon$ for $j = 1, \dots, n$, and consequently $t|\varphi(a)| < 1$. It follows that

$$a\xi_j = 0 \quad \text{for} \quad j = 1,\dots,n \quad \Rightarrow \quad \varphi(a) = 0\ . \tag{4.6}$$

For any $a \in \mathscr{B}(\mathscr{H})$, define $\widehat{a}$ by $\widehat{a} = \sum_{j=1}^{n} a\xi_j\langle \xi_j, \cdot\rangle_{\mathscr{H}}$. Evidently $(a - \widehat{a})\xi_j = 0$ for $j = 1, \dots, n$, and hence by (4.6),

$$\varphi(a) = \varphi(\widehat{a}) = \sum_{j=1}^{n} \varphi[(a\xi_j)\langle \xi_j, \cdot\rangle_{\mathscr{H}}] \ . \tag{4.7}$$

For each fixed j, and any $\eta \in \mathscr{H}$, consider the rank-one operator $\eta\langle \xi_j, \cdot\rangle_{\mathscr{H}}$. Then $\eta \mapsto \varphi(\eta\langle \xi_j, \cdot\rangle_{\mathscr{H}})$ is a bounded linear functional on $\mathscr{H}$, and therefore by the Riesz Representation Theorem, there is a vector $\zeta_j \in \mathscr{H}$ such that $\langle \zeta_j, \eta\rangle_{\mathscr{H}} = \varphi(\eta\langle \xi_j, \cdot\rangle_{\mathscr{H}})$ for all $\eta \in \mathscr{H}$. Combining this with (4.7) yields (4.5). The final statement is a standard application of the Hahn-Banach Theorem. □

4.2 The measurable functional calculus

Let $a \in \mathscr{B}(\mathscr{H})$ be self adjoint, and for brevity let $\sigma(a)$ denote $\sigma_{\mathscr{B}(\mathscr{H})}(a)$, Let $\eta \in \mathscr{H}$, and define a linear functional μ_η on $\mathscr{C}(\sigma(a))$ through

$$\mu_\eta(f) = \langle \eta, f(a)\eta\rangle_{\mathscr{H}} \ . \tag{4.8}$$

Then μ is evidently a positive linear functional with $\mu_\eta(1) = \|\eta\|^2_{\mathscr{H}}$. By the Reisz-Markoff Theorem, there is a positive Borel measure of total mass $\|\eta\|^2_{\mathscr{H}}$, also denoted by μ_η, so that for all $f \in \mathscr{C}(\sigma(a))$,

$$\mu_\eta(f) = \int_{\sigma(a)} f \mathrm{d}\mu_\eta \ . \tag{4.9}$$

Combining (4.8) and (4.9), we conclude that for all $f \in \mathscr{C}(\sigma(a))$, $\langle \eta, f(a)\eta\rangle_{\mathscr{H}} = \displaystyle\int_{\sigma(a)} f \mathrm{d}\mu_\eta$

Now let $\{\eta_j\}$ be any dense sequence in the unit sphere of $\mathscr{H}$, and define the probability measure ν on $\sigma(a)$ by

$$\nu = \sum_{j=1}^{\infty} 2^{-j}\mu_{\eta_j} \ . \tag{4.10}$$

By (4.8) and (4.10), for all $f, g \in \mathscr{C}(\sigma(a))$, and all k,

$$2^{-k}\|(f(a) - g(a))\eta_k\|^2_{\mathscr{H}} \le \sum_{j=1}^{\infty} 2^{-j}\langle \eta_j, |f(a) - g(a)|^2\eta_j\rangle_{\mathscr{H}} \le \int_{\sigma(a)} |f - g|^2 \mathrm{d}\nu \ . \tag{4.11}$$

Recall that the continuous functions on $\sigma(a)$ are dense in $L^1(\sigma(a), \mu)$ for any Borel measure μ, and that from any sequence that converges in $L^1(\sigma(a), \mu)$, one can extract a subsequence that converges a.e. to f. It follows that if f is any bounded Borel function of $\sigma(a)$, there exists a sequence $\{f_n\}$ of uniformly bounded continuous functions on $\sigma(a)$ with $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ for ν a.e. λ, then by the Lebesgue Dominated Convergence Theorem, $\lim_{n\to\infty} \int_{\sigma(a)} |f_n - f|^2 \mathrm{d}\nu = 0$. Consequently for all $\epsilon > 0$, $\int_{\sigma(a)} |f_n - f_m|^2 \mathrm{d}\nu < \epsilon$ for all but finitely many m and n. Then by (4.10), $\{f_n(a)\}$ is a Cauchy sequence for the strong operator topology, and hence $\lim_{n\to\infty} f_n(a) = b$ exists for this topology. In particular, for all $\xi \in \mathscr{H}$,

$$\lim_{n\to\infty} f_n(a)\xi = b\xi \ . \tag{4.12}$$

We would like to define $f(a) = b$, but at this point, one might suppose that the definition depends on the approximating sequence of continuous functions, or on the choice of the dense sequence $\{\eta_j\}$ in the unit sphere of $\mathscr{H}$. In fact, it does not.

First, let f be a bounded Borel function on $\sigma(a)$, and let $\{f_n\}$ and $\{\tilde{f}_n\}$ be two sequences of continuous functions that converge ν a.e. to f, where ν is defined by (4.10) for some choice of a dense sequence $\{\eta_j\}$ in the unit sphere of $\mathscr{H}$. Define the "interlaced" sequence $\{g_n\}$ by $g_{2n-1} = f_n$ and $g_{2n} = \tilde{f}_n$. Then evidently $\{g_n\}$ converges ν a.e. to f, and so $b = \lim_{n\to\infty} g_n(a)$ exists in the weak operator topology. Since subsequences of convergent sequences converge to the same limit, we have

$$b = \lim_{n\to\infty} f_n(a) = \lim_{n\to\infty} \tilde{f}_n(a) \ .$$

Next, let $\{\eta_j\}$ and $\{\tilde{\eta}_j\}$ be two dense sequences in the unit sphere of $\mathscr{H}$, and define ν and $\tilde{\nu}$ in terms of them as in (4.10). Then ν and $\tilde{\nu}$ are equivalent measures, meaning that a Borel set $E \subset \sigma(a)$ is a null set for one if and only if it is also a null set for the other. To see this, suppose on the contrary that E is a Borel subset of $\sigma(a)$ with $\nu(E) > 0$ but $\tilde{\nu}(E) = 0$. Let f be the indicator function of E. Let $\mu = \nu + \tilde{\nu}$, and let $\{f_n\}$ be a sequence of continuous non-negative functions that converges μ a.e. to f, and hence converges both ν and $\tilde{\nu}$ a.e. Since each $f_n(a)$ is self adjoint and nonnegative, so it the weak limit b. Therefore

$$\lim_{n\to\infty}\int_{\sigma(a)} f_n(\lambda)\mathrm{d}\nu = \nu(E) > 0 \quad \text{while} \quad \lim_{n\to\infty}\int_{\sigma(a)} f_n(\lambda)\mathrm{d}\tilde{\nu} = \tilde{\nu}(E) = 0 \ .$$

This would imply that

$$\sum_{\jmath=1}^{\infty} 2^{-j}\langle \eta_j b \eta_j\rangle_{\mathscr{H}} > 0 \quad \text{while} \quad \sum_{\jmath=1}^{\infty} 2^{-j}\langle \tilde{\eta}_j b \tilde{\eta}_j\rangle_{\mathscr{H}} = 0 \ ,$$

and then since each term in the sum on the right is non-negative, $\langle \tilde{\eta}_j b \tilde{\eta}_j\rangle_{\mathscr{H}} = 0$ for all j, while for at least one j_0, $\langle \eta_{j_0} b \eta_{j_0}\rangle_{\mathscr{H}} > 0$. But since $\{\tilde{\eta}_j\}$ is dense in the unit sphere, there is a subsequence $\{\tilde{\eta}_{j_k}\}$ with $\lim_{k\to\infty}\tilde{\eta}_{j_k} = \eta_{j_0}$, and this then forces $\langle \eta_{j_0} b \eta_{j_0}\rangle_{\mathscr{H}} = 0$. The contradiction shows there is no such Borel set E and hence the two measures are equivalent.

In other words, the self adjoint operator a determines a class of mutually equivalent Borel measures, and if E is a null set for this class then $\mu_\eta(E) = 0$ for all $\eta \in \mathscr{H}$ since for $\eta \neq 0$, we may include $\|\eta\|_{\mathscr{H}}^{-1}\eta$ is any dense sequence in the unit sphere. Thus, when discussing *a.e.* convergence of functions on the spectrum of a, we shall always mean almost everywhere with respect to any one of these equivalent measures, and then for any Bounded Borel function f on $\sigma(a)$, we define

$$f(a) = \lim_{n\to\infty} f_n(a) \tag{4.13}$$

where $\{f_n\}$ is any sequence of continuous functions on $\sigma(a)$ that converges almost everywhere to f in this sense. By what we have noted above, such sequences always exist, the limit always exists, and the limit is independent of the approximating sequence $\{f_n\}$ and of the particular reference measure used in the construction. We have prepared the way for an easy proof of the following theorem:

4.8 THEOREM (Functional Calculus For Bounded Self-Adjoint Operators)**.** *Let $\mathscr{H}$ be a separable Hilbert space, and let a be a self-adjoint element of $\mathscr{B}(\mathscr{H})$. Let $\mathscr{B}(\sigma(a))$ denote the bounded Borel functions on $\sigma(a)$. Then for $f \in \mathscr{B}(\sigma(a))$, $f(a)$ is defined through (4.13) as described above.*

The function $f \mapsto f(a)$ is a norm-reducing $*$-isomorphism from $\mathscr{B}(\sigma(a))$ into $\mathscr{B}(\mathscr{H})$. Moreover:

(1) Its restriction to the continuous functions $\mathscr{B}(\sigma(a))$ agrees with the function $f \mapsto f(a)$ given by that Abstract Spectral Theorem. In particular, a is the image of $\lambda \mapsto \lambda$ and the identity is the image of $\lambda \mapsto 1$.

(2) Let $\{f_n\}$ be a sequence in $\mathscr{B}(\sigma(a))$ $\sup_{n\in\mathbb{N}}\{\|f_n\|_\infty\} < \infty$, and such that $\lim_{n\to\infty} f_n(\lambda) = f(\lambda)$ for all $\lambda \in \sigma(a)$. Then $f(a) = \lim_{n\to\infty} f_n(a)$ in the strong operator topology.

(3) The function $f \mapsto f(a)$ preserves order: If $g \geq f$ in $\mathscr{B}(\sigma(a))$, then $g(a) - f(a)$ is non-negative.

Proof. Let $f \in \mathscr{B}(\sigma(a))$ and let $\{f_n\}$ be a bounded sequence in $\mathscr{C}(\sigma(a))$ converging *a.e.* to f. Then for all $z\eta, ci \in \mathscr{H}$,

$$\langle \zeta, (f(a))^*\xi\rangle_{\mathscr{H}} = \langle f(a)\zeta, \xi\rangle_{\mathscr{H}} = \lim_{n\to\infty}\langle f_n(a)\zeta, \xi\rangle_{\mathscr{H}} = \lim_{n\to\infty}\langle \zeta, f_n^*(a)\xi\rangle_{\mathscr{H}} = \langle \zeta, f^*(a)\xi\rangle_{\mathscr{H}} \ .$$

The map $f \mapsto f(a)$ is evidently linear, and we conclude the proof that it is a 8-homomorphism by showing that for $f, g \in \mathscr{B}(\sigma(a))$, $fg(a) = f(a)g(a)$. Let $\{f_n\}$ and $\{g_n\}$ be bounded sequences in $\mathscr{C}(\sigma(a))$ converging *a.e.* to f and g respectively. For any $\zeta, \xi \in \mathscr{H}$,

$$\langle \zeta, f(a)g(a)\xi\rangle_{\mathscr{H}} = \langle f^*(a)\zeta, g(a)\xi\rangle_{\mathscr{H}} = \lim_{n\to\infty} \langle f_n^*(a)\zeta, g_n(a)\xi\rangle_{\mathscr{H}} =$$
$$\lim_{n\to\infty} \langle \zeta, f_n g_n(a)\xi\rangle_{\mathscr{H}} = \langle \zeta, fg(a)\xi\rangle_{\mathscr{H}} \ .$$

Next, for all $f \in \mathscr{B}(\sigma(a))$, and all $\xi \in \mathscr{H}$ with $\|\xi\|_{\mathscr{H}} = 1$

$$\|f(a)\xi\|^2_{\mathscr{H}} = \int_{\sigma(a)} |f(\lambda)|^2 \mathrm{d}\mu_\xi \leq \|f\|^2_\infty$$

since μ_ξ is a probability measure. This completes the proof that $f \mapsto f(a)$ is a nor reducing $*$-homomorphism from $\mathscr{B}(\sigma(a))$ to $\mathscr{B}(\mathscr{H})$.

Properties *(1)* and *(2)* have been proved above, and to prove *(3)* write $g - f = h^2$ and use the $*$-homomorphism property. □

The $*$-homomorphism provided by Theorem 4.8 need not be an isomorphism. The following example is useful elsewhere: Let $\lambda_0 \in \sigma(a)$m and consider the function 1_{λ_0} given by $1_{\lambda_0}(\lambda) = 1$ for $\lambda = \lambda_0$ and zero otherwise. Then for all λ, $\lambda_0 1_{\lambda_0}(\lambda) = \lambda 1_{\lambda_0}(\lambda)$. By the $*$-homomorphism property,

$$\lambda_0 1_{\lambda_0}(a) = a 1_{\lambda_0}(a) \ .$$

It follows that any non-zero vector in the range of $1_{\lambda_0}(a)$ is an eigenvector of a with eigenvalue λ_0, and conversely any such eigenvector ξ is in the range of $1_{\lambda_0}(a)$ as one sees by considering a continuous approximation $\{f_n\}$ to 1_{λ_0}: We suppose that $\|\xi\| = 1$, and note that

$$\langle \xi, 1_{\lambda_0}(a)\xi\rangle_{\mathscr{H}} = \lim_{n\to\infty} \langle \xi, f_n(a)\xi\rangle_{\mathscr{H}} = \lim_{n\to\infty} f_n(\lambda_0) = 1 \ .$$

By the conditions for equality in the Cauchy-Schwarz inequality, and the nor reduction property, $1_{\lambda_0}(a)\xi = \xi$.

In fact, for any Borel set $E \subset \mathbb{R}$, the indicator function 1_E is a projector in $\mathscr{B}(\sigma(a))$, and hence, by the $*$-homomorphism property, $1_E(a)$ is a projector in $\mathscr{B}(\mathscr{H})$. Taking $E = (s, t]$ for $s < t$ yields a useful family of projectors that we shall encounter later on.

4.3 The polar decomposition

4.9 DEFINITION (Operator absolute value)**.** . Let $\mathscr{H}$ be a Hilbert space and let $a \in \mathscr{B}(\mathscr{H})$. Then the operator absolute value of a is the operator $|a|$ defined by

$$|a| = \sqrt{a^*a} \ , \tag{4.14}$$

where the square root is taken using the Abstract Spectral Theorem.

4.10 REMARK. *One should not be misled by the notation: It is not in general true that* $|ab| = |a||b|$*, or that* $|a^*| = |a|$ *or even that* $|a + b| \leq |a| + |b|$*.*

Next, or each $t > 0$, define the operator

$$u_t := a(t1 + |a|)^{-1}.$$

This does not require the Abstract Spectral Theorem; since $|a| \geq 0$, $(1t + |a|)$ is invertible. Now note that for $s, t > 0$, by the Resolvent Identity

$$u_t - u_s = (s - t)a[(t1 - |a|)^{-1}(s1 - |a|)^{-1}] \ .$$

Hence for any $\xi \in \mathscr{H}$, and $0 < s < t$,

$$\begin{aligned}
\|(u_t - u_s)\xi\|^2 &= (s - t)^2 \langle \xi, |a|^2 (t1 - |a|)^{-2}(s1 - |a|)^{-2}\xi\rangle_{\mathscr{H}} \\
&= (t - s)^2 \int_{\sigma(|a|)} \frac{\lambda^2}{(t + \lambda)^2 (s + \lambda)^2} \mathrm{d}\mu_\xi \\
&\leq \int_{\sigma(|a|)\backslash\{0\}} \frac{t^2}{(t + \lambda)^2} \mathrm{d}\mu_\xi
\end{aligned}$$

Since $0 \leq t^2/(t + \lambda)^2 \leq 1$ for all $\lambda > 0$, and since $\lim_{t\to 0} t^2/(t + \lambda)^2 = 0$ for all $\lambda > 0$, the Lebesgue Dominated Convergence Theorem yields

$$\lim_{t\to 0} \left(\sup_{s<t} \{\|(u_t - u_s)\xi\|^2\} \right) = 0 \ .$$

Thus, the strong limit $u = \lim_{t\to 0} u_t$ exists. Note that $u|a| = \lim_{t\to 0} u_t|a| = x \lim_{t\to 0} f_t(|a|)$ where $f_t(\lambda) = \lambda/(t + \lambda)$. since $\lim_{t\to 0} f_t(\lambda) = 1_{(0,\infty)}\lambda$ for all $\lambda \geq 0$, it follows from Theorem 4.8 that $\lim_{t\to 0} f_t(|a|) = 1_{(0,\infty)}(|a|) = 1 - 1_{\{0\}}(|a|)$. Since $1_{\{0\}}(|a|)$ is the projector onto the null space of $|a|$, which is the null space of a, $a1_{\{0\}}(|a|) = 0$, and hence

$$u|a| = a \ . \tag{4.15}$$

Next note that $u^*u = \lim_{t\to\infty} f_t^2(|a|)$ with $f_t(\lambda) = \lambda/(t + \lambda)$ once more. It follows that

$$u^*u = 1_{(0,\infty)}(|a|) \tag{4.16}$$

which is the projector onto $\ker(a)^\perp$. It follows from (4.15) that $\mathrm{ran}(u) = \mathrm{ran}(a)$, and hence u is a partial isometry from $\ker(a)^\perp$ onto $\mathrm{ran}(a)$.

Taking the adjoint of (4.15), we obtain $a^* = |a|u^*$ and hence $aa^* = ua^*au^*$. Squaring both sides and observing that $au^*u = a$ follows from (4.16), we obtain $(aa^*)^2 = u(a^*a)^2u^*$. An induction now yields $(aa^*)^n = u(a^*a)^n u^*$ for all n, and then taking a polynomial approximation to the square root, we conclude that

$$u|a|u^* = |a^*| \ , \tag{4.17}$$

and then since $a^* = |a|u^* = u^*u|a|u^*$,

$$a^* = u^*(u|a|u^*) \tag{4.18}$$

is the polar decomposition of a^*.

4.4 Compact operators

Let $\mathscr{H}$ be a Hilbert space. An operator $a \in \mathscr{B}(\mathscr{H})$ has finite rank if $\mathrm{ran}(a)$ is a finite dimensional subspace of $\mathscr{H}$, or equivalently, $\ker(a)^\perp$ is finite dimensional. If a is finite rank and $\{\eta_1, \dots, \eta_m\}$ is an orthonormal basis for $\ker(a)^\perp$, then we may write a in the form

$$a = \sum_{j=1}^{m} |\xi_j\rangle\langle\eta_j| \ , \tag{4.19}$$

where for each j, $\xi_j = a\eta_j$. Conversely, every operator of the form (4.19), even without the assumption that $\{\eta_1, \dots, \eta_m\}$ is orthonormal, is evidently finite rank. It is clear that the set of finite rank operators on $\mathscr{B}(\mathscr{H})$ are a two-sided $*$-ideal in $\mathscr{B}(\mathscr{H})$, but it is not closed in $\mathscr{B}(\mathscr{H})$, which brings us to the following definition.

4.11 DEFINITION. An operator $a \in \mathscr{B}(\mathscr{H})$ is *compact* in case it is the norm limit of finite rank operators. That is, the set $\mathscr{C}(\mathscr{H})$ of all compact operators on $\mathscr{H}$ is the norm closure of the set of finite rank operators on $\mathscr{H}$.

The closure of a 2-sided $*$-ideal is a 2-sided $*$-ideal, and hence $\mathscr{C}(\mathscr{H})$ is a 2-sided $*$-ideal in $\mathscr{B}(\mathscr{H})$ and a C^* subalgebra of $\mathscr{B}(\mathscr{H})$.

4.12 EXAMPLE. Let $\{\lambda_j\}$ be a sequence of complex numbers such that $\lim_{j\to\infty} \lambda_j = 0$. Let $\{\eta_j\}$ and $\{\xi_j\}$ be any two orthonormal bases of $\mathscr{H}$, and define a sequence of operators a_n by $a_n = \sum_{j=1}^{n} \lambda_j |\eta_j\rangle\langle\xi_j|$. Then for any $\zeta, \zeta' \in \mathscr{H}$ and any $m < n \in \mathbb{N}$,

$$\begin{aligned} |\langle\zeta, (a_n - a_m)\zeta'\rangle_{\mathscr{H}}| &\leq \sum_{j=m+1}^{n} |\lambda_j| |\langle\zeta, \eta_j\rangle_{\mathscr{H}}|^{1/2} |\langle\zeta', \eta_j\rangle_{\mathscr{H}}|^{1/2} \\ &\leq \left(\max_{j\geq m+1}\{|\lambda_j|\}\right) \sum_{j=m+1}^{n} |\langle\zeta, \eta_j\rangle_{\mathscr{H}}|^{1/2} |\langle\zeta', \eta_j\rangle_{\mathscr{H}}|^{1/2} \\ &\leq \left(\max_{j\geq m+1}\{|\lambda_j|\}\right) \|\zeta\|_{\mathscr{H}} \|\zeta'\|_{\mathscr{H}} \ , \end{aligned}$$

where we have used Cauchy-Schwarz in the last line. This shows that $\|a_n - a_m\| \leq \max_{j\geq m+1}\{|\lambda_j|\}$, and hence that $\{a_n\}$ is Cauchy in the operator norm. The limit is a compact operator that is finite rank if and only if $\lambda_j \neq 0$ for only finitely many j.

4.13 THEOREM. *An operator $a \in \mathscr{B}(\mathscr{H})$ is compact if and only if the image of the unit ball in $\mathscr{H}$ under a has compact closure.*

Proof. Suppose that a is compact. Choose $\epsilon > 0$, and a finite rank operator $b = \sum_{j=1}^{m} |\xi_j\rangle\langle\eta_j|$ such that $\|b - a\| < \epsilon/2$. Let $b(B)$ denote the image of the unit ball under b, which is evidently isometric to a bounded subset of $\mathbb{C}^m$, and hence it may be covered by a finite collection of balls of radius $\epsilon/2$. Let $\{\xi_1, \dots, \xi_n\}$ denote the centers of these balls. For any η in the unit ball of $\mathscr{H}$, $\|b\eta - \xi_j\|_{\mathscr{H}} < \epsilon/2$ for some j. Then $\|a\eta - \xi_j\|_{\mathscr{H}} \leq \|(a-b)\eta\|_{\mathscr{H}} + \|b\eta - \xi_j\|_{\mathscr{H}} < \epsilon$, and hence the image of the unit ball B under a, $a(B)$,is covered by finitely many balls of radius ϵ. Since $\epsilon > 0$ is arbitrary, $a(B)$ has compact closure.

Conversely, suppose that $\overline{a(B)}$ has compact closure. Let $a = u|a|$ be the polar decomposition of a. Since u is an isometry from $\overline{\text{ran}(|a|)}$ onto $\overline{\text{ran}(a)}$, $|a|(B)$ has compact closure. For any $\epsilon > 0$, consider the spectral projection $p_\epsilon = 1_{(\epsilon,\|a\|]}(a)$. By the Spectral Theorem, $p_\epsilon^\perp|a|$ has spectrum in $(0,\epsilon]$ and hence $\|p_\epsilon^\perp|a|\| \leq \epsilon$. Since $|a| = p_\epsilon|a| + p_\epsilon^\perp|a|$,

$$\||a| - p_\epsilon|a|\| \leq \|p_\epsilon^\perp|a|\| \leq \epsilon \ .$$

Hence to show that $|a|$ is compact, it suffices to show that $p_\epsilon|a|$ is finite rank for all $\epsilon > 0$.

For any η in the range of p_ϵ, $\||a|\eta\|_{\mathscr{H}} \geq \epsilon\|\eta\|_{\mathscr{H}}$. Therefore, if $\{\eta_1,\dots,\eta_m\}$ is any orthonormal set in the range of p_ϵ, $\||a|\eta_i - |a|\eta_j\|_{\mathscr{H}} \geq \sqrt{2}\epsilon$ for all $i \neq j$. Hence the range of p_ϵ cannot be infinite dimensional when $|a|(B)$ has compact closure. This shows that $|a|$ is compact, and then since $\mathscr{C}(\mathscr{H})$ is an ideal, $a = u|a|$ is compact as well. □

More useful information can be gleaned from the proof of Theorem 4.13. Let $a \in \mathscr{C}(\mathscr{H})$, and let $a = u|a|$ be its polar decomposition. For $\epsilon > 0$, consider once more the spectral projection $p_\epsilon = 1_{(\epsilon,\|a\|]}(a)$. Then as we have seen, $p_\epsilon|a| = p_\epsilon|a|p_\epsilon$ is finite rank and is self adjoint. Hence it has discrete spectrum, and then by the Spectral Theorem, we may write

$$p_\epsilon|a| = \sum_{j=1}^{m} \sigma_j|\eta_j\rangle\langle\eta_j| \tag{4.20}$$

for some set of m numbers $\sigma_1,\dots,\sigma_m$ in the interval $(\epsilon,\|a\|]$ and some orthonormal set $\{eta_1,\dots,\eta_m\}$. We may assume without loss of generality that the indexing is such that $\sigma_j \leq \sigma_i$ for $j > i$. If all of the σ_j are distinct, then the representation in (4.20) is then uniquely determined, and the vectors η_j are determined up to a complex multiple of unit modulus. As noted above,

$$\||a| - p_\epsilon|a|\| \leq \epsilon \ .$$

Thus, $\lim_{\epsilon\to 0} p_\epsilon|a| = |a|$. As ϵ decreases, the only effect on (4.20) is the addition of more terms on the right, since for $0 < \epsilon' < \epsilon$ and $p_{\epsilon',\epsilon} := 1_{(\epsilon',\epsilon]}(a)$, $p_{\epsilon'}|a| = p_{\epsilon',\epsilon}|a| + p_\epsilon|a|$, and the ranges of the two operators on the right are orthogonal. This reasoning leads to:

4.14 THEOREM. *Let $a \in \mathscr{C}(\mathscr{H})$. Then there exist two orthonormal sets $\{\eta_j\}$ and $\{\xi_j\}$, not necessarily complete, and a monotone non-increasing sequence of positive numbers $\{\sigma_j\}$ such that a has the norm convergent expansion*

$$a = \sum_{j=1}^{\infty} \sigma_j|\xi_j\rangle\langle\eta_j| \ . \tag{4.21}$$

Proof. We have already seen that this is the case when $a = |a|$, and then we may take $\xi_j = \eta_j$ for each j. In general, let $a = u|a|$ be the polar decomposition of a, and define $\xi_j = u\eta_j$. □

4.5 Trace class operators

Let a be a positive operator on a separable Hilbert space $\mathscr{H}$, and let $\{\eta_j\}$ and $\{\xi_k\}$ be two orthonormal bases for $\mathscr{H}$. Then

$$\sum_{j=1}^{\infty}\langle\eta_j, a\eta_j\rangle_{\mathscr{H}} = \sum_{j=1}^{\infty}\|a^{1/2}\eta_j\|_{\mathscr{H}} = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\left|\langle\xi_k, a^{1/2}\eta_j\rangle_{\mathscr{H}}\right|^2 \ .$$

Since infinite series of non-negative terms may be summed in any order, the right hand side is actually symmetric in $\{\eta_j\}$ and $\{\xi_k\}$. Therefore, by symmetry in the two orthonormal bases,

$$\sum_{j=1}^{\infty}\langle \eta_j, a\eta_j\rangle_{\mathscr{H}} = \sum_{k=1}^{\infty}\langle \xi_k, a\xi_k\rangle_{\mathscr{H}} \ ,$$

showing that $\sum_{j=1}^{\infty}\langle \eta_j, a\eta_j\rangle_{\mathscr{H}}$ depends only on a, and not the particular orthonormal basis $\{\eta_j\}$.

4.15 DEFINITION (Trace class)**.** Let $\mathscr{H}$ be a separable Hilbert space. For positive $a \in \mathscr{B}(\mathscr{H})$, we define

$$\mathrm{Tr}[a] = \sum_{j=1}^{\infty}\langle \eta_j, a\eta_j\rangle_{\mathscr{H}}$$

where $\{\eta_j\}$ is any orthonormal basis of $\mathscr{H}$. If $\mathrm{Tr}[a] < \infty$, we say that a is *trace class.* More generally, $\mathscr{T}(\mathscr{H})$, the set of *trace class operators* on $\mathscr{H}$ is defined by

$$\mathscr{T}(\mathscr{H}) = \{ \ a \in \mathscr{B}(\mathscr{H}) \ : \ \mathrm{Tr}[|a|] < \infty \ \} \ . \tag{4.22}$$

Note that $\mathscr{T}(\mathscr{H}) \subset \mathscr{C}(\mathscr{H})$: To see this, let $a \in \mathscr{T}(\mathscr{H})$, and for $\epsilon > 0$, let $p_\epsilon = 1_{(\epsilon,\|a\|]}(a)$. Then if $\{\zeta_1, \dots, \zeta_n\}$ is an orthonormal subset of $\mathrm{ran}(|a|)$,

$$\sum_{j=1}^{n}\langle \zeta_j, |a|\zeta_j\rangle_{\mathscr{H}} \geq n\epsilon \ ,$$

and since the sum cannot exceed $\mathrm{Tr}[|a|] < \infty$, $n \leq \mathrm{Tr}[|a|]/\epsilon$. Hence $|a| = p_\epsilon|a| + p_\epsilon^\perp|a|$ where $p_\epsilon|a|$ is finite rank, and where $\|p_\epsilon^\perp|a|\| \leq \epsilon$. However, not all compact operators are trace class:

4.16 EXAMPLE. We have seen in Example 4.12 that if $\{\lambda_j\}$ is a sequence of complex numbers such that $\lim_{j\to\infty}\lambda_j = 0$, and $\{\eta_j\}$ and $\{\xi_j\}$ are any two orthonormal sequences of $\mathscr{H}$, $a = \sum_{j=1}^{\infty}\lambda_j|\eta_j\rangle\langle\eta_j|$ is norn convergent. It is easy to see that $|a| = \sum_{j=1}^{\infty}|\lambda_j||\eta_j\rangle\langle\eta_j|$, and thus $a \in \mathscr{T}(\mathscr{H})$ if and only if $\sum_{j=1}^{\infty}|\lambda_j| < \infty$. Thus, there exist compact operators that are not trace class.

It turns out that the set of trace class operators in $\mathscr{B}(\mathscr{H})$ is an ideal in $\mathscr{B}(\mathscr{H})$. The following lemma is useful for showing this:

4.17 LEMMA. *Let $\mathscr{A}$ be a unital C^* algebra, Then every element a of $\mathscr{A}$ can be written as a linear combination of four unitaries, each of which belong to $C^*(a)$, the smallest C^* subalgebra of $\mathscr{A}$ that contains a and 1.*

Proof. Let $y \in \mathscr{A}$ be self-adjoint with $\|y\| \leq 1$. Define $z = y + i\sqrt{1-y^2}$, which clearly belongs to $C^*(y) \subset \mathscr{A}$. Then $z^*z = zz^* = 1$ so that z is unitary, and $y = (z+z^*)/2$ then displays y as a linear combination of unitaries in $C^*(y)$. Assume $a \neq 0$, and let $t = (2\|a\|)^{-1}$. Then

$$a = \frac{1}{2t}x + i\frac{1}{2t}y \quad \text{where} \quad x = t(a+a^*) \quad \text{and} \quad y = -it(a-a^*) \ .$$

displays a as a linear combination of two self adjoint contractions in $C^*(a)$. □

4.18 THEOREM. *For any separable Hilbert space, $\mathscr{T}(\mathscr{H})$ is a (non-closed) $*$-ideal in $\mathscr{B}(\mathscr{H})$, and for all $a, b \in \mathscr{T}(\mathscr{H})$,*

$$\mathrm{Tr}[|a+b|] \leq \mathrm{Tr}[|a|] + \mathrm{Tr}[|b|] \ . \tag{4.23}$$

4.19 REMARK. *It is* not *the case that for all $a, b \in \mathscr{B}(\mathscr{H})$, or even in $M_2(\mathbb{C})$, that $|a+b| \leq |a| + |b|$. Nonetheless, (4.23) is true in general.*

Proof. We first show that $\mathscr{T}(\mathscr{H})$ is closed under addition. Let $a, b \in \mathscr{T}(\mathscr{H})$, and let $a+b = u|a+b|$ be the polar decomposition of $a+b$. Then $|a+b| = u^*(a+b)$ and so for any orthonormal basis $\{\eta_j\}$,

$$\mathrm{Tr}[|a+b|] = \sum_{j=1}^{\infty} \langle \eta_j u^* a \eta_j \rangle_{\mathscr{H}} + \sum_{j=1}^{\infty} \langle \eta_j u^* b \eta_j \rangle_{\mathscr{H}} \ . \tag{4.24}$$

Hence it suffice to show that for all $a \in \mathscr{T}(\mathscr{H})$ all partial isometries u, and all orthonormal basis $\{eta_j\}$, the series $\sum_{j=1}^{\infty} \langle \eta_j u^* a \eta_j \rangle_{\mathscr{H}}$ is absolutely summable. To this end, let $a = v|a|$ be the polar decomposition of a, and note that by Cauchy-Schwarz,

$$|\langle \eta_j u^* a \eta_j \rangle_{\mathscr{H}}| = |\langle |a|^{1/2} v^* u \eta_j |a|^{1/2} \eta_j \rangle_{\mathscr{H}}| \leq \langle \xi_j |a| \xi_j \rangle_{\mathscr{H}}^{1/2} \leq \langle \xi_j |a| \xi_j \rangle_{\mathscr{H}}^{1/2}$$

where $\xi_j = v^* u \eta_j$. Since $\{\xi_j\}$ is orthonormal in $\mathscr{H}$ (but not necessarily complete), one more application of Cauchy-Schwarz yields

$$\sum_{j=1}^{\infty} |\langle \eta_j u^* a \eta_j \rangle_{\mathscr{H}}| \leq \left(\sum_{j=1}^{\infty} \langle \xi_j, |a| \xi_j \rangle_{\mathscr{H}} \right)^{1/2} \left(\sum_{j=1}^{\infty} \langle \eta_j, |a| \eta_j \rangle_{\mathscr{H}} \right)^{1/2} \leq \mathrm{Tr}[|a|] < \infty \ .$$

Using this twice in (4.24) shows that $\mathrm{Tr}[|a+b|] \leq \mathrm{Tr}[|a|] + \mathrm{Tr}[|b|] < \infty$, which is (4.23). This proves that $\mathscr{T}(\mathscr{H})$ is closed under addition, and it is now clear that $\mathscr{T}(\mathscr{H})$ is a vector subspace of $\mathscr{B}(\mathscr{H})$.

Next, recall from (4.18) that if $a = u|a|$ is the polar decomposition of a, $|a^*| = u|a|u^*$, and hence for any orthonormal basis $\{\eta_j\}$,

$$\sum_{j=1}^{\infty} \langle \eta_j, |a^*| \eta_j \rangle_{\mathscr{H}} = \sum_{j=1}^{\infty} \langle u^* \eta_j |a| u^* \eta_j \rangle_{\mathscr{H}} \leq \mathrm{Tr}[|a|] < \infty$$

since $\{u^* \eta_j\}$ is orthonormal (though not necessarily complete).

Because $(ba)^* = a^* b^*$, if $\mathscr{T}(\mathscr{H})$ is a left or right ideal it is a two-sided ideal. Hence it suffices to show that for all $a \in \mathscr{T}(\mathscr{H})$ and all $b \in \mathscr{B}(\mathscr{H})$, $ba \in \mathscr{T}(\mathscr{H})$. By Lemma 4.17, it suffices to do this for b unitary in $\mathscr{B}(\mathscr{H})$. But is b is unitary, $|ba| = |a|$, and so it is evident that $ba \in \mathscr{T}(\mathscr{H})$ when $a \in \mathscr{T}(\mathscr{H})$ and b is unitary.

Example 4.16 shows that every finite rank operator belongs to $\mathscr{T}(\mathscr{H})$, but there are operators in the norm closure of the set of finite rank operators that are not in $\mathscr{T}(\mathscr{H})$. Hence $\mathscr{T}(\mathscr{H})$ is not closed in $\mathscr{B}(\mathscr{H})$. □

Now let x be a self adjoint operator in $\mathscr{B}(\mathscr{H})$. Then let x_+ and x_- be the positive an negative parts of x, defined vis the spectral calculus, so that $x = x_+ - x_-$. Then $|x| = x_+ + x_-$. Applying the previous theorem, we conclude that $x \in \mathscr{T}(\mathscr{H})$ if and only if both x_+ and x_- belong to $\mathscr{T}(\mathscr{H})$. Now consider any $a \in \mathscr{B}(\mathscr{H})$ and write $a = x + iy$ where $x = (a+a^*)/2$ and $y = (a-a^*)/2i$.

Then, applying the previous theorem once more, $a \in \mathscr{T}(\mathscr{H})$ if and only if each of x_+, x_-, y_+ and y_- belong to $\mathscr{T}(\mathscr{H})$. It follows that for $a \in \mathscr{T}(\mathscr{H})$ and any two orthonormal bases $\{\eta_j\}$ and $\{\xi_j\}$ of $\mathscr{H}$,

$$\sum_{j=1}^{\infty} \langle \eta_j a \eta_j \rangle_{\mathscr{H}} = \sum_{j=1}^{\infty} \langle \xi_j a \xi_j \rangle_{\mathscr{H}} \ ,$$

and both sums are absolutely convergent.

4.20 DEFINITION (Trace)**.** For all $a \in \mathscr{T}(\mathscr{H})$, the *trace* of a, $\mathrm{Tr}[a]$, is defined by

$$\mathrm{Tr}[a] = \sum_{j=1}^{\infty} \langle \eta_j a \eta_j \rangle_{\mathscr{H}}$$

where $\{\eta_j\}$ is any orthonormal basis of $\mathscr{H}$. The *trace norm* is the norm $\|\cdot\|_1$ on $\mathscr{T}(\mathscr{H})$ by

$$\|a\|_1 = \mathrm{Tr}[|a|] \ , \tag{4.25}$$

which is a norm since it obviously satisfies $\|\lambda a\|_1 = |\lambda|\|a\|_1$, and for $a, b \in \mathscr{T}(\mathscr{H})$, and it satisfies the triangle inequality on account of (4.23).

4.21 THEOREM (Properties of the trace)**.** *The functional* $a \mapsto \mathrm{Tr}[a]$ *is linear on* $\mathscr{T}(\mathscr{H})$ *and* $\mathrm{Tr}[a^*] = \mathrm{Tr}[a]^*$ *for all* $a \in \mathscr{T}(\mathscr{H})$*. Moreover:*
(i) For all $a \in \mathscr{T}(\mathscr{H})$ *and all* $b \in \mathscr{B}(\mathscr{H})$

$$\mathrm{Tr}[ab] = \mathrm{Tr}[ba] \tag{4.26}$$

(ii) For all $\zeta, \xi \in \mathscr{H}$*, and all* $b \in \mathscr{B}(\mathscr{H})$*,*

$$\mathrm{Tr}\big[|\zeta\rangle\langle\xi|a\big] = \langle \xi, b\zeta \rangle_{\mathscr{H}} \ . \tag{4.27}$$

(iii) For all $a \in \mathscr{T}(\mathscr{H})$ *and all* $b \in \mathscr{B}(\mathscr{H})$*,*

$$|\mathrm{Tr}[ab]| \le \|a\|_1 \|b\| \ . \tag{4.28}$$

Proof. The linearity is evident, and since $\langle \eta, a\eta \rangle_{\mathscr{H}} = \langle \eta, a^*\eta \rangle_{\mathscr{H}}^*$, it follows that $\mathrm{Tr}[a^*] = \mathrm{Tr}[a]^*$. In view of the linearity and Lemma 4.17, to prove (4.26) it suffices to show that when u is unitary and $a \in \mathscr{T}(\mathscr{H})$, $\mathrm{Tr}[au] = \mathrm{Tr}[ua]$. To prove this, let $\{\eta_j\}$ be any orthonormal basis. Then $\{\xi_j\} = \{u\eta_j\}$ is another, and

$$\mathrm{Tr}[au] = \sum_{j=1}^{\infty} \langle \eta_j, au\eta_j \rangle_{\mathscr{H}} = \sum_{j=1}^{\infty} \langle u^*\xi_j, a\xi_j \rangle_{\mathscr{H}} = \sum_{j=1}^{\infty} \langle \xi_j, ua\xi_j \rangle_{\mathscr{H}} = \mathrm{Tr}[ua] \ .$$

Next, we prove (4.27). For any $\zeta, \xi \in \mathscr{H}$, $|\zeta\rangle\langle\xi|$ is rank-one and hence trace class. Let $a \in \mathscr{B}(\mathscr{H})$. Then by Theorem 4.18, $|\zeta\rangle\langle\xi|a \in \mathscr{T}(\mathscr{H})$. Since (4.27) is trivially ture if $\zeta = 0$, we may suppose $\zeta \neq 0$. Choose any orthonormal basis $\{\eta_j\}$ of $\mathscr{H}$ with $\eta_1 = \|\zeta\|_{\mathscr{H}}^{-1}\zeta$. Then

$$\mathrm{Tr}\big[|\zeta\rangle\langle\xi|a\big] = \sum_{j=1}^{\infty} \langle \eta_j, |\zeta\rangle\langle\xi|a\eta_j \rangle_{\mathscr{H}} = \langle \eta_1, \zeta \rangle \langle \xi|a\eta_1 \rangle_{\mathscr{H}} = \|\zeta\|_{\mathscr{H}} \langle \xi|a\eta_1 \rangle_{\mathscr{H}} = \langle \xi, a\zeta \rangle_{\mathscr{H}} \ ,$$

and this proves (4.27).

To prove (4.28) note that

$$(ba)^*(ba) = a^*b^*ba \leq \|b\|^2 a^*a \ .$$

By the operator monotonicity of the square root function, $|ba| \leq \|b\||a|$, and hence (4.28) follows from the linearity of the trace. □

4.22 THEOREM. *Let $a \in \mathscr{C}(\mathscr{H})$ have the norm convergence expansion*

$$a = \sum_{j=1}^{\infty} \sigma_j |\zeta_j\rangle\langle\zeta_j'|$$

where $\{\sigma_j\}$ is a monotone non-increasing sequence of positive numbers and $\{\zeta_j\}$ and $\{\zeta_j'\}$ are two orthonormal sequences in $\mathscr{H}$. Then for each $k \in \mathbb{N}$, let $\mathcal{U}_k$ denote the set of rank-k partial isometries on $\mathscr{H}$. Then

$$\sum_{j=1}^{k} \sigma_j = \max\{\Re(\mathrm{Tr}[ua]) \ : \ u \in \mathcal{U}_k \ \} \ .$$

In particular, the sequence $\{\sigma_j\}$ is uniquely determined by a.

Proof. The general element u of $\mathcal{U}_k$ has the form $u = \sum_{\ell=1}^{k} |\xi_\ell\rangle\langle\eta_\ell|$ where $\{\xi_1, \dots, \xi_\ell\}$ and $\{\eta_1, \dots, \eta_k\}$ are two orthonormal sets in $\mathscr{H}$. Then by (4.27),

$$\begin{aligned}
\Re\left(\mathrm{Tr}[ua]\right) &= \Re\left(\sum_{\ell=1}^{k} \langle\xi_\ell, a\eta_\ell\rangle_{\mathscr{H}}\right) \\
&= \Re\left(\sum_{\ell=1}^{k}\sum_{j=1}^{\infty} \sigma_j \langle\xi_\ell, \zeta_j\rangle_{\mathscr{H}} \langle\zeta_j', \eta_\ell\rangle_{\mathscr{H}}\right) \\
&\leq \sum_{\ell=1}^{k}\sum_{j=1}^{\infty} \left(\sqrt{\sigma_j}|\langle\xi_\ell, \zeta_j\rangle_{\mathscr{H}}|\right)\left(\sqrt{\sigma_j}|\langle\zeta_j', \eta_\ell\rangle_{\mathscr{H}}|\right) \\
&\leq \left(\sum_{j=1}^{\infty} \sigma_j \sum_{\ell=1}^{k} |\langle\xi_\ell, \zeta_j\rangle_{\mathscr{H}}|^2\right)^{1/2} \left(\sum_{j=1}^{\infty} \sigma_j \sum_{\ell=1}^{k} |\langle\eta_\ell, \zeta_j'\rangle_{\mathscr{H}}|^2\right)^{1/2}
\end{aligned}$$

Define the numbers p_j, p_j', $j \in \mathbb{N}$ by

$$p_j = \sum_{\ell=1}^{k} |\langle\xi_\ell, \zeta_j\rangle_{\mathscr{H}}|^2 \quad \text{and} \quad p_j' = \sum_{\ell=1}^{k} |\langle\eta_\ell, \zeta_j'\rangle_{\mathscr{H}}|^2 \ .$$

By Bessel's inequality, for each j, $p_j \leq \|\zeta_j\|^2 \leq 1$, and $p_j' \leq \|\zeta_j'\|^2 \leq 1$. For the same reason, $\sum_{j=1}^{\infty} p_j \leq \sum \ell = 1^k \|\xi_\ell\|_{\mathscr{H}}^2 \leq k$ and likewise, $\sum_{j=1}^{\infty} p_j' \leq k$.

It is a classical inequality of Hardy, Littlewood and Polya that under the conditions that $\{\sigma_j\}$ is a non-increasing sequence and $\{p_j\}$ is a sequence of non negative numbers such that $p_j \leq 1$ for all j, and $\sum_{j=1}^{\infty} p_j \leq k$,

$$\sum_{\jmath=1}^{\infty} \sigma_j p_j \leq \sum_{j=1}^{k} \sigma_k \ ,$$

and there is equality in case $p_j = 1$ for $j \leq k$ and $p_j = 0$ for $j > k$. There is equality *only* for this case if $\sigma_{k+1} < \sigma_k$. Applying this inequality to the two sums obtained above yields the result. □

4.23 DEFINITION (Singular values)**.** Let $a \in \mathscr{C}(\mathscr{H})$ have the expansion

$$a = \sum_{j=1}^{\infty} \sigma_j |\zeta_j\rangle\langle\zeta_j'| \tag{4.29}$$

where $\{\sigma_j\}$ is a monotone non-increasing sequence of positive numbers and $\{\zeta_j\}$ and $\{\zeta_j'\}$ are two orthonormal sequences in $\mathscr{H}$. The sequence $\{\sigma_j\}$ is called the sequence of *singular values* of a. When more than one operator is under consideration, we write $\sigma_j(a)$ to denote the jth singular vlaue of $a \in \mathscr{C}(\mathscr{H})$.

Evidently, when a has the form (4.29),

$$|a| = \sum_{j=1}^{\infty} \sigma_j |\zeta_j'\rangle\langle\zeta_j'| \tag{4.30}$$

and hence $\mathrm{Tr}[|a|] = \sum_{j=1}^{\infty} \sigma_j$. That is, $a \in \mathscr{C}(\mathscr{H})$ is trace class if and only if its sequence of singular values is summable.

Theorem 4.25 has a useful corralary:

4.24 COROLLARY (Corroloary of Theorem 4.25)**.** *Let $a, b \in \mathscr{C}(\mathscr{H})$. Then for all $k \in \mathbb{N}$,*

$$\left| \sum_{j=1}^{k} \sigma_j(a) - \sum_{j=1}^{k} \sigma_j(b) \right| \leq \sqrt{k}\|a - b\| \ . \tag{4.31}$$

Proof. Let u be a rank-k partial isometry. Then $|\Re\mathrm{Tr}[au] - \Re\mathrm{Tr}[bu]| \leq \|a - b\|\|u\|_1 = k^{1/2}\|a - b\|$. By Theorem 4.22, there is a choice of u so that $\Re\mathrm{Tr}[au] = \sum_{j=1}^{k} \sigma_j(a)$, and then by Theorem 4.22 again, for this u, $\Re\mathrm{Tr}[bu] \leq \sum_{j=1}^{k} \sigma_j(b)$. Therefore,

$$\sum_{j=1}^{k} \sigma_j(a) - \sum_{j=1}^{k} \sigma_j(b) \leq k^{1/2}\|a - b\| \ .$$

By the symmetry in a and b, (4.31) follows. □

4.25 THEOREM. *$\mathscr{T}(\mathscr{H})$ is a Banach space in the metric given by the trace norm Moreover:*
(i) For every $a \in \mathscr{T}(\mathscr{H})$ define a linear functional ϕ_a on $\mathscr{C}(\mathscr{H})$ by

$$\phi_a(x) = \mathrm{Tr}[ax] \quad \text{for all} \quad x \in \mathscr{C}(\mathscr{H}) \ . \tag{4.32}$$

The mapping $a \mapsto \phi_a$ is an isometric isomorphism of $\mathscr{T}(\mathscr{H})$ onto $\mathscr{C}(\mathscr{H})^*$.
(i) For every $b \in \mathscr{B}(\mathscr{H})$ define a linear functional ψ_b on $\mathscr{T}(\mathscr{H})$ by

$$\phi_b(x) = \mathrm{Tr}[bx] \quad \text{for all} \quad x \in \mathscr{T}(\mathscr{H}) \ . \tag{4.33}$$

The mapping $b \mapsto \psi_b$ is an isometric isomorphism of $\mathscr{B}(\mathscr{H})$ onto $\mathscr{T}(\mathscr{H})^*$.

Proof. Once we have proved *(i)*, the completeness of $\mathscr{T}(\mathscr{H})$ follows since the dual of a Banach space is always complete. The proofs of *(i)* and *(ii)* are almost the same:

For any $\zeta, \zeta' \in \mathscr{H}$, consider the rank-one operator $|\zeta\rangle\langle\zeta'|$. Then

$$\||\zeta\rangle\langle\zeta'|\| = \||\zeta\rangle\langle\zeta'|\|_1 = \|\zeta\|_{\mathscr{H}}\|\zeta'\|_{\mathscr{H}} \ . \tag{4.34}$$

Hence, if $\phi \in \mathscr{C}(\mathscr{H})^*$,

$$|\phi(|\zeta\rangle\langle\zeta'|)| \leq \|\phi\|\||\zeta\rangle\langle\zeta'|\| = \|\phi\|\|\zeta\|_{\mathscr{H}}\|\zeta'\|_{\mathscr{H}} \ .$$

Define the sesquilinear form $q_\phi(\zeta, \zeta')$ by $q_\phi(\zeta, \zeta') = \phi(|\zeta\rangle\langle\zeta'|)$. We have seen that this is bounded, and so by the Riesz Lemma, there is an $a \in \mathscr{B}(\mathscr{H})$ with $\|a\| = \|\phi\|$ such that for all $\zeta, \zeta' \in \mathscr{H}$, $q_\phi(\zeta, \zeta') = \langle\zeta, a\zeta'\rangle_{\mathscr{H}}$. But then for $\zeta, \zeta' \in \mathscr{H}$,

$$\phi(|\zeta\rangle\langle\zeta'|) = \langle\zeta, a\zeta'\rangle_{\mathscr{H}} = \mathrm{Tr}\big[a|\zeta\rangle\langle\zeta'|\big] \ ,$$

and then by linearity, $\phi(x) = \mathrm{Tr}[ax]$ for all finite rank x. Since finite rank operators are dense in $\mathscr{C}(\mathscr{H})$, this is valid for all $x \in \mathscr{C}(\mathscr{H})$.

We now show that $a \in \mathscr{T}(\mathscr{H})$. Let $a = u|a|$ be the polar decomposition of a. Let $\{\eta_1, \dots, \eta_n\}$ be any set of n orthonormal vectors in $\mathscr{H}$, and define the finite rank partial isometry $v = \sum_{j=1}^n |\eta_j\rangle\langle u\eta_j|$. Then

$$\sum_{j=1}^n \langle\eta_j, |a|\eta_j\rangle_{\mathscr{H}} \mathrm{Tr}[av] = \phi(v) \leq \|\phi\|\|v\| = \|\phi\| \ .$$

Since n is arbitrary, $a \in \mathscr{T}(\mathscr{H})$. This shows that $a \mapsto \phi_a$ is an isomorphism of $\mathscr{T}(\mathscr{H})$ onto $\mathscr{C}(\mathscr{H})^*$. By (4.28), $\|\phi_a\| \leq \|a\|_1$. By Theorem 4.22, $\|\phi_a\| \geq \|a\|_1$. Hence $\|\phi_a\| = \|a\|_1$, and the isomorphism is isometric.

Next, let $\psi \in \mathscr{T}(\mathscr{H})^*$. As above, define a sequilinear form q_ψ on $\mathscr{H}\times\mathscr{H}$ by $q_\phi(\zeta, \zeta') = \phi(|\zeta\rangle\langle\zeta'|)$ for all $\zeta, \zeta' \in \mathscr{H}$. By (4.34,

$$|\psi(|\zeta\rangle\langle\zeta'|)| \leq \|\psi\|\||\zeta\rangle\langle\zeta'|\|_1 = \|\psi\|\|\zeta|_{\mathscr{H}}\|\zeta'\|_{\mathscr{H}} \ .$$

By the Riesz Lemma, there is an $b \in \mathscr{B}(\mathscr{H})$ with $\|b\| \leq \|\phi\|$ such that for all $\zeta, \zeta' \in \mathscr{H}$, $q_\psi(\zeta, \zeta') = \langle\zeta, b\zeta'\rangle_{\mathscr{H}}$. But then for $\zeta, \zeta' \in \mathscr{H}$,

$$\psi(|\zeta\rangle\langle\zeta'|) = \langle\zeta, b\zeta'\rangle_{\mathscr{H}} = \mathrm{Tr}[b|\zeta\rangle\langle\zeta'|] \ ,$$

and then by linearity, $\psi(x) = \mathrm{Tr}[bx]$ for all finite rank x. Since finite rank operators are dense in $\mathscr{T}(\mathscr{H})$ in the trace norm, this is valid for all $x \in \mathscr{T}(\mathscr{H})$. This shows that $a \mapsto \phi_a$ is an isomorphism of $\mathscr{B}(\mathscr{H})$ onto $\mathscr{T}(\mathscr{H})^*$. By (4.28), $\|\phi_b\| \leq \|b\|$. Since $\|b\| = \sup\{\ \langle\zeta, b\zeta'\rangle_{\mathscr{H}}\ :\ \|\zeta\|_{\mathscr{H}}, \|\zeta'\|_{\mathscr{H}} = 1\ \}$ and since

$$|\langle\zeta, b\zeta'\rangle_{\mathscr{H}}| = |\mathrm{Tr}\big[a|\zeta\rangle\langle\zeta'|\big]| = |\psi_a(|\zeta\rangle\langle\zeta'|)| \leq \|\psi_a\|\|\zeta\|_{\mathscr{H}}\|\zeta'\|_{\mathscr{H}} \ ,$$

we also have $\|b\| \leq \|\psi_a$. Hence $\|\psi_b\| = \|b\|_1$, and the isomorphism is isometric. $\square$

4.6 Hilbert-Schmidt operators

Let $a \in \mathscr{B}(\mathscr{H})$ and let $a = u|a|$ be the polar decomposition of a. Then $aa^* = ua^*au^*$ so that whenever $a^*a \in \mathscr{T}(\mathscr{H})$, $aa^* \in \mathscr{T}(\mathscr{H})$. By symmetry, $a^*a \in \mathscr{T}(\mathscr{H})$ if and only if $aa^* \in \mathscr{T}(\mathscr{H})$.

4.26 DEFINITION (Hilbert-Schmidt operators)**.** An operator $a \in \mathscr{B}(\mathscr{H})$ is *Hilbert-Schmidt* in case $a^*a \in \mathscr{T}(\mathscr{H})$, or equivalently, in case $aa^* \in \mathscr{T}(\mathscr{H})$. The set of all Hilbert-Schmidt operators on $\mathscr{H}$ is denoted $\mathscr{C}_2(\mathscr{H})$.

It is evident that $\mathscr{C}_2(\mathscr{H})$ is self adjoint; i.e., $a \in \mathscr{C}_2(\mathscr{H})$ if and only if $a^* \in \mathscr{C}_2(\mathscr{H})$. We now show that $\mathscr{C}_2(\mathscr{H})$ is a two-sided $*$-ideal in $\mathscr{B}(\mathscr{H})$. Suppose that $a \in \mathscr{C}_2(\mathscr{H})$ and $b \in \mathscr{B}(\mathscr{H})$. Then

$$(ba)^*(ba) = a^*(b^*b)a \leq \|b\|^2 a^*a \in \mathscr{T}(\mathscr{H}) \ , \tag{4.35}$$

and hence $ba \in \mathscr{C}_2(\mathscr{H})$ for all $a \in \mathscr{C}_2(\mathscr{H})$ and all $b \in \mathscr{B}(\mathscr{H})$. Since $\mathscr{C}_2(\mathscr{H})$ is self adjoint, $ab \in \mathscr{C}_2(\mathscr{H})$ for all $a \in \mathscr{C}_2(\mathscr{H})$ and all $b \in \mathscr{B}(\mathscr{H})$. To see that $\mathscr{C}_2$ is closed under addition, note that

$$(a+b)^*(a+b) + (a-b)^*(a-b) = 2a^*a + 2b^*b \ .$$

When $a, b \in \mathscr{C}_2(\mathscr{H})$, the right hand side is in $\mathscr{T}(\mathscr{H})$ by definitionm and then $a+b \in \mathscr{C}_2(\mathscr{H})$. This proves that $\mathscr{C}_2(\mathscr{H})$ is a two-sided $*$-ideal in $\mathscr{B}(\mathscr{H})$.

If $a \in \mathscr{C}_2(\mathscr{H})$, then $|a|^2 \in \mathscr{T}(\mathscr{H})$, and hence $|a|^2$ is compact. It follows from the Spectral Theorem that for all $\epsilon > 0$,

$$\| |a| 1_{[0,\epsilon^2]}(|a|^2) - |a| \| \leq \epsilon$$

and $1_{[0,\epsilon^2]}(|a|^2)$ is finite rank. Hence $|a|$ and then a belong to $\mathscr{C}(\mathscr{H})$. Let $a = \sum_{j=1}^\infty \sigma_j |\zeta_j\rangle\langle\zeta'_j|$ be a singular value decomposition of a. Then $a^*a = \sum_{j=1}^\infty \sigma_j^2 |\zeta'_j\rangle\langle\zeta'_j|$ and hence

$$\mathrm{Tr}[a^*a] = \sum_{j=1}^\infty \sigma_j^2 \ . \tag{4.36}$$

This proves that $a \in \mathscr{C}_2(\mathscr{H})$ is and only if $\{\sigma_j\} \in \ell_2$, and since $\sigma_1(a) = \|a\|$, it proves that for all $a \in \mathscr{C}_2(\mathscr{H})$,

$$\|a\| \leq \|a\|_2 \ . \tag{4.37}$$

Note that for all $a, b \in \mathscr{C}_2(\mathscr{H})$, $(a+ib)^*(a+ib) = a^*a + b^*b + i(a^*b - b^*a)$ and hence that $a^*b - b^*a \in \mathscr{T}(\mathscr{H})$. Also, $(a+b)^*(a+b) = a^*a + b^*b + (a^*b + b^*a)$ and hence that $a^*b + b^*a \in \mathscr{T}(\mathscr{H})$. It follows that for all $a, b \in \mathscr{C}_2(\mathscr{H})$, $a^*b \in \mathscr{T}(\mathscr{H})$. Therefore, we may define a sesquilinear form $\langle \cdot, \cdot \rangle_2$ on $\mathscr{C}_2(\mathscr{H})$, called the *Hilbert-Schmidt inner product* and the associated *Hilbert-Schmidt inner norm* by

$$\langle a, b \rangle_2 = \mathrm{Tr}[a^*b] \quad \text{and} \quad \|a\|_2^2 = \langle a, a \rangle_2 \ . \tag{4.38}$$

By (4.35) for all $a \in \mathscr{T}(\mathscr{H})$ and all $b \in \mathscr{B}(\mathscr{H})$,

$$\|ab\|_2 \leq \|a\|_2 \|b\| \ . \tag{4.39}$$

4.27 THEOREM. *The space $\mathscr{C}_2(\mathscr{H})$, equipped with the Hilbert-Schmidt inner product, is a Hilbert space.*

Proof. We need only show the completeness. Suppose that $\{a_n\}$ is a Cauchy sequence in $\mathscr{C}_2(\mathscr{H})$ for the Hilbert-Schmidt norm. By (4.37), $\{a_n\}$ is a Cauchy sequence in the operator norm. Since $\mathscr{B}(\mathscr{H})$ is complete, there exists $a \in \mathscr{B}(\mathscr{H})$ such that $\lim_{n\to\infty} \|a_n - a\| = 0$. Since each a_n is compact, a is compact.

By Corollary 4.24 of Theorem 4.22, for each $k \in \mathbb{N}$, $\lim_{n\to\infty} \sigma_k(a_n) = \sigma_k(a)$. Consequently,

$$\sum_{k=1}^{\infty} \sigma_k(a)^2 \leq \liminf_{n\to\infty} \left(\sum_{k=1}^{\infty} \sigma_k(a_n)^2 \right) = \liminf_{n\to\infty} \|a_n\|_2^2 < \infty \ .$$

Thus, $a \in \mathscr{C}_2(\mathscr{H})$.

Now choose $\epsilon > 0$ and $n \in \mathbb{N}$ such that for all $j, k \geq n$, $\|a_j - a_k\|_2 \leq \epsilon$. Then for any $j \geq n$ and any finite rank projection p,

$$\|(a_j - a)p\|_2 \leq \|(a_j - a_k)p\|_2 + \|(a_k - a)p\|_2 \ .$$

By (4.39), $\|(a_k - a)p\|_2 \leq \|a_k - a\|\|p\|_2$ so that $\lim_{k\to\infty} \|(a_k - a)p\|_2 = 0$. Hence $\|(a_j - a)p\|_2 \leq \epsilon$. Note that

$$\|(a_j - a)p\|_2^2 = \mathrm{Tr}[p|a_j - a|^2 p]$$

and we may choose p to make the right hand side arbitrarily close to $\mathrm{Tr}[|a_j - a|^2]$. Hence $\|a_j - a\|_2 \leq \epsilon$, and this shows that $\lim_{j\to\infty} a_j = a$ in the Hilbert-Schmidt norm. □

Positive linear functionals on $\mathscr{C}(\mathscr{H})$ may be written in terms of the Hilbert-Schmidt inner product in a useful way. By Theorem 4.25, if ϕ is any conitnuous linear functional on $\mathscr{C}(\mathscr{H})$, then there is an operator $a \in \mathscr{T}(\mathscr{H})$ such that $\phi(x) = \mathrm{Tr}[ax]$ for all $x \in \mathscr{C}(\mathscr{H})$. Taking $x = |\eta\rangle\langle\eta|$ for $\eta\ in \mathscr{H}$, we see that the positivity of ϕ implies the positivity of a. Let $b = \sqrt{a}$, and observe that $b \in \mathscr{C}_2(\mathscr{H})$. Then $\mathrm{Tr}[ax] = \mathrm{Tr}[bbx] = \mathrm{Tr}[b^*bx] = \langle b, xb\rangle_2$, so that for all $x \in \mathscr{C}(\mathscr{H})$,

$$\phi(x) = \langle b, xb\rangle_2 \ .$$

4.7 The σ-weak topology

We introduce one more topology on $\mathscr{B}(\mathscr{H})$, namely the weak-$*$ toplogy. By Theorem 4.25, this is the weakest topology making all of the maps $x \mapsto \mathrm{Tr}[ax]$, $x \in \mathscr{T}(\mathscr{H})$, continuous. By Theorem 4.7, this topology is stronger than the strong operator topology. It is often called the *σ-weak topology on $\mathscr{B}(\mathscr{H})$* or the *ultraweak topology on $\mathscr{B}(\mathscr{H})$*, though this last name is somewhat ambiguous: The weak-$*$ topology lies "beyond" the weak operator topology in that it is a *stronger* topology, not weaker.

A basic set of neighborhoods of the origin for the σ-weak topology is given by the sets

$$W_{a_1,\ldots,a_n,\epsilon} = \{\ x \in \mathscr{B}(\mathscr{H}) \ : \ |\mathrm{Tr}[a_j x]| \leq \epsilon \ , \quad j = 1, \ldots, n\} \tag{4.40}$$

where $a_1, \ldots, a_n \in \mathscr{T}(\mathscr{H})$ and $\epsilon > 0$.

4.28 THEOREM. *Every linear functional ϕ on $\mathscr{B}(\mathscr{H})$ that is σ-weakly continuous is of the form $\phi(x) = \mathrm{Tr}[bx]$ for some $b \in \mathscr{T}(\mathscr{H})$. That is, linear functionals on $\mathscr{B}(\mathscr{H})$ that are continuous with respect to the σ-weak topology belong to the predual of $\mathscr{B}(\mathscr{H})$.*

Proof. Let ϕ be a non-zero linear functional on $\mathscr{B}(\mathscr{H})$ that is σ-weakly continuous. Let U be the open disk of radius 1/2 in $\mathbb{C}$ centered at 1. Let $x \in \mathscr{B}(\mathscr{H})$ be such that $\phi(x) = 1$. Then $\phi^{-1}(U)$ contains $W_{a_1,\dots,a_n,\epsilon}$ for some $a_1, \dots, a_n \in \mathscr{T}(\mathscr{H})$ and $\epsilon > 0$. Since $\mathscr{T}(\mathscr{H}) \subset \mathscr{C}_2(\mathscr{H})$ we may apply the Gram-Schmidt algorithm in $\mathscr{C}_2(\mathscr{H})$, and may therefore assume without loss of generality that $\{a_1, \dots, a_n\}$ is orthonormal in $\mathscr{C}_2(\mathscr{H})$.

Now observe that for all $x \in \mathscr{B}(\mathscr{H})$, if $\mathrm{Tr}[xa_j] = 0$ for $j = 1, \dots, n$, then every multiple of x belongs to $W_{a_1,\dots,a_n,\epsilon}$, and hence $\phi(x) = 0$. For all $x \in \mathscr{B}(\mathscr{H})$, define $y = \sum_{j=1}^n \phi(a_j)\mathrm{Tr}[xa_j]$ and $z = x - y$. Note that $\mathrm{Tr}[za_j] = 0$ for $j = 1, \dots, n$,and hence $\phi(z) = 0$. Therefore, for all $x \in \mathscr{B}(\mathscr{H})$, $\phi(x) = \displaystyle\sum_{j=1}^n \phi(a_j)\mathrm{Tr}[xa_j] = \mathrm{Tr}[bx]$ where $b = \displaystyle\sum_{j=1}^n \phi(a_j)a_j$. □

4.29 DEFINITION. A positive linear functional ϕ on $\mathscr{B}(\mathscr{H})$ is *completely additive* in case whenever $\{p_j\}$ is a familay of mutually orthogonal projections with $\sum_{j=1}^\infty = 1$, then

$$\sum_{j=1}^\infty \phi(p_j) = \phi(1) \ . \tag{4.41}$$

4.30 THEOREM. *A positive linear functionla ϕ on $\mathscr{B}(\mathscr{H})$ is completely additive if and only if it is σ-weakly continuous.*

Proof. Let $\{\phi_j\}$ be any set of mutually orthogonal projections such that $\sum_{j=1}^\infty p_j = 1$, and let $q_n = \sum_{j=1}^n p_j$. Suppose that ϕ is σ-weakly continuous. Then for some positive $b \in \mathscr{T}(\mathscr{H})$ and all $x \in \mathscr{B}(\mathscr{H})$, $\phi(x) = \mathrm{Tr}[bx]$. Then each q_n is an orthonormal projection and $\lim_{n\to\infty} q_n = 1$ in the strong operator topology. It follows easily that

$$\lim_{n\to\infty} \phi(q_n) = \lim_{n\to\infty} \mathrm{Tr}[bq_n] = \mathrm{Tr}[b] = \phi(1) \ ,$$

and hence ϕ is completely additive.

Suppose next that ϕ is completely additive. Let $\{p_j\}$ be any set of mutually orthogonal *finite rank* projections such that $\sum_{j=1}^\infty p_j = 1$, and let $q_n = \sum_{j=1}^n p_j$ as above. Let $x \in \mathscr{B}(\mathscr{H})$ be positive. Then $\phi(x) = \phi(xq_n) + \phi(x^{1/2}x^{1/2}q_n^\perp)$, and by the Cauchy-Schwarz inequality, $|\phi(x^{1/2}x^{1/2}q_n^\perp)| \leq \phi(x)^{1/2}\phi(q_n^\perp x q_n^\perp)^{1/2}$. But

$$\phi(q_n^\perp x q_n^\perp) \leq \|x\|\phi(q_n^\perp) = \|x\|(\phi(1) - \phi(q_n) \ .$$

and hence $\lim_{n\to\infty} \phi(xq_n) = \phi(x)$.

By Theorem 4.25, the restriction of ϕ to $\mathscr{C}(\mathscr{H})$ has the form $\phi(x) = \mathrm{Tr}[bx]$ where $b \in \mathscr{T}(\mathscr{H})$. Since for all $x \in \mathscr{B}(\mathscr{H})$ and all n, $xq_n \in \mathscr{C}(\mathscr{H})$, $\phi(xq_n) = \mathrm{Tr}[bxq_n]$. It now follows that $\phi(x) = \lim_{n\to\infty} \phi(xq_n) = \mathrm{Tr}[bx]$. □

5 Representations of C^* algebras

5.1 Irreducible representations

5.1 DEFINITION. A *representation* of a C^*-algebra $\mathscr{A}$ is a $*$-homomorphism π from $\mathscr{A}$ into $\mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$. For any subspace $\mathscr{K}$ of $\mathscr{H}$, we define

$$\pi(\mathscr{A})\mathscr{K} = \{\ \pi(a)\eta\ :\ a \in \mathscr{A}\ ,\ \eta \in \mathscr{K}\ \}\ .$$

A subspace $\mathscr{K}$ of $\mathscr{H}$ is *invariant under* π in case $\pi(\mathscr{A})\mathscr{K} \subset \mathscr{K}$. The representation π is *irreducible* in case no non-trivial subspace $\mathscr{K}$ of $\mathscr{H}$ is invariant under π. The representation π is *non-degenerate* in case $\overline{\pi(\mathscr{A})\mathscr{H}} = \mathscr{H}$. Let π_1 and π_2 be two representations of $\mathscr{A}$ on Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively. Then π_1 and π_2 are *equivalent representations* of $\mathscr{A}$ in case there exists a unitary transformation from u from $\mathscr{H}_1$ onto $\mathscr{H}_2$ such that for all $a \in \mathscr{A}$,

$$\pi_2(a)u = u\pi_1(a)\ .$$

The notion of the *commutant* of a subset $\mathcal{S} \subset \mathscr{B}(\mathscr{H})$ plays a crucial role in the study of irreducibility.

5.2 DEFINITION. Let $\mathscr{H}$ be a Hilbert space, and $\mathcal{S} \subset \mathscr{B}(\mathscr{H})$. The *commutant* $\mathcal{S}'$ of $\mathcal{S}$ is the subset of $\mathscr{B}(\mathscr{H})$ given by

$$\mathcal{S}' = \{\ a \in \mathscr{B}(\mathscr{H})\ :\ ab - ba = 0 \quad \text{for all} \quad b \in \mathcal{S}\ \}\ .$$

5.3 LEMMA. *Let $\mathscr{H}$ be a Hilbert space, and $\mathcal{S} \subset \mathscr{B}(\mathscr{H})$. The* commutant *$\mathcal{S}'$ of $\mathcal{S}$ has the following properties:*

(1) $\mathcal{S}'$ is a closed in the weak operator topology on $\mathscr{B}(\mathscr{H})$, and contains the identity 1*.*

(2) $\mathcal{S}'$ is a subalgebra of $\mathscr{B}(\mathscr{H})$.

(3) If $\mathcal{S}$ is closed under the involution, then so is $\mathcal{S}'$, so that $\mathcal{S}'$ is a weakly closed $*$-subalgebra of $\mathscr{B}(\mathscr{H})$ that contains the identity.

Proof. It is evident that for all $\mathcal{S}$, $1 \in \mathcal{S}'$. Moreover, for any $\zeta, \xi \in \mathscr{H}$ and $b \in \mathcal{S}$, define the linear functional $\varphi_{\zeta,\xi,b}$ on $\mathscr{B}(\mathscr{H})$ by

$$\varphi_{\zeta,\xi,b}(a) = \langle \zeta, (ab - ba)\xi \rangle_{\mathscr{H}} = \langle \zeta, a(b\xi) \rangle_{\mathscr{H}} - \langle (b^*\zeta), a\xi \rangle_{\mathscr{H}}\ .$$

Since $\varphi_{\zeta,\xi,b}$ is weakly continuous, $\varphi^{-1}_{\zeta,\xi,b}(\{0\})$ is weakly closed. Then since

$$\mathcal{S}' = \bigcap \{\ \varphi^{-1}_{\zeta,\xi,b}(\{0\})\ :\ \zeta, \xi \in \mathscr{H}\ ,\ b \in \mathcal{S}\ \}\ ,$$

(1) is proved. *(2)* is evident, and the *(3)* follows from the fact that $(ab - ba)^* = (b^*a^* - a^*b^*)$ together with *(1)* and *(2)*. □

5.4 DEFINITION. A *von Neumann algebra* is a $*$-subalgebra $\mathscr{M}$ of $\mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$ such that $1 \in \mathscr{M}$ and such that $\mathscr{M}$ is a weakly closed subset of $\mathscr{B}(\mathscr{H})$

By Lemma 5.2, the commutant of any $*$-subalgebra of $\mathscr{B}(\mathscr{H})$ is a von Neumann algebra. Note that every von Neumann algebra $\mathscr{M}$ is generated by the projections it contains. Indeed, $\mathscr{M}$ is generated by is self adjoint elements, and by the Spectral Theorem, each self adjoint $a \in \mathscr{M}$ is the strong limit os a sequence of finite linear combinations of the spectral projections of a, which themselves belong to $\mathscr{M}$, being strong limits of polynomials in a.

5.5 LEMMA. *Let $\mathscr{A}$ be a C^* algebra, and let π be a non-zero representation of it as an algebra of operators on some Hilbert space $\mathscr{H}$. Then a closed subspace $\mathscr{K}$ of $\mathscr{H}$ is invariant under $\pi(\mathscr{A})$ if and only if the orthogonal projection of $\mathscr{H}$ onto $\mathscr{K}$ belongs to $(\pi(\mathscr{A}))'$.*

Proof. First, $\mathscr{K}$ is invariant under $\pi(\mathscr{A})$ if and only if $\mathscr{K}^\perp$ is invariant under $\pi(\mathscr{A})$. To see this, let $\zeta \in \mathscr{K}^\perp$ and $\xi \in \mathscr{K}$, and $a \in \mathscr{A}$. If $\mathscr{K}$ is in variant, $\pi(a^*)\xi \in \mathscr{K}$, and hence

$$\langle \pi(a)\zeta, \xi\rangle_{\mathscr{H}} = \langle \zeta, \pi(a^*)\xi\rangle_{\mathscr{H}} = 0 \ .$$

Thus the invariance of $\mathscr{K}$ implies the invariance of $\mathscr{K}^\perp$, and then by symmetry, the reverse implication is valid as well.

Now let p be the orthogonal projection onto $\mathscr{K}$. Then when $\mathscr{K}$ is invariant, for all $a \in \mathscr{A}$,

$$0 = p\pi(a)(1-p) = p\pi(a) - p\pi(a)p = p\pi(a) - \pi(a)p$$

where the last equality is true since the range of $\pi(a)p$ lies in $\mathscr{K}$. Therefore, $p \in (\pi(\mathscr{A}))'$. Conversely, if $p \in (\pi(\mathscr{A}))'$ and $\xi \in \mathscr{K}$, then for all $a \in \mathscr{A}$,

$$\pi(a)\xi = \pi(a)p\xi = p\pi(a)\xi \in \mathscr{K} \ ,$$

which shows the invariance of $\mathscr{K}$. □

Lemma 5.5 permits us to make the following definition:

5.6 DEFINITION. For a representation π of a $\mathbb{C}^*$ algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$, and a non-zero projector $p \in (\pi(\mathscr{A}))'$, π_p is the subrepresntation obtained by restricting π to $\mathrm{ran}(p)$.

5.7 THEOREM. *Let $\mathscr{A}$ be a C^* algebra, and let π be a non-zero representation of it as an algebra of operators on some Hilbert space $\mathscr{H}$. Then π is irreducible if and only if $(\pi(\mathscr{A}))'$ consists of scalar multiples of the identity.*

Proof. If $(\pi(\mathscr{A}))'$ consists of scalar multiples of the identity, then $(\pi(\mathscr{A}))'$ contains no non-trivial orthogonal projections, and hence by Lemma 5.5, π is irreducible. On the other hand, if $(\pi(\mathscr{A}))'$ contains some operator that is not a multiple of the identity, then it contains a self adjoint operator a that is not a multiple of the identity. Any such $a \in (\pi(\mathscr{A}))'$ has a non-trivial spectral projection that is also in $(\pi(\mathscr{A}))'$ since $(\pi(\mathscr{A}))'$ is a von Neumann algebra containing a. □

5.8 THEOREM (von Neumann Double Commutant Theorem)**.** *Let $\mathscr{A}$ be a $*$-subalgebra of $\mathscr{B}(\mathscr{H})$ that contains the identity. Then $\mathscr{A}''$ is the weak operator topology closure of $\mathscr{A}$.*

Proof. Since $\mathscr{A}$ is convex, the weak and strong operator topology closes of $\mathscr{A}$ coincide. Hence is suffices to show that for all $a \in \mathscr{A}''$, every strong neighborhood of a contains some $b \in \mathscr{A}$. That is, it suffices to show that for all $n \in \mathbb{N}$ and all $\{\eta_1, \ldots, \eta_n\} \subset \mathscr{H}$, and all $\epsilon > 0$, there is some $b \in \mathscr{A}$ such that $\|(b-a)\eta_j\| < \epsilon$ for all $j = 1, \ldots, n$.

Let $\hat{\mathscr{H}} = \mathscr{H} \oplus \cdots \oplus \mathscr{H}$, the direct sum of n copies of $\mathscr{H}$. The elements of $\mathscr{B}(\hat{\mathscr{H}})$ are $n \times n$ matrices $[b_{i,j}]$ with entries in $\mathscr{B}(\mathscr{H})$.

Let $\hat{\mathscr{A}}$ be the algebra of all operators on $\hat{\mathscr{H}}$ the form $[a\delta_{i,j}]$ with $a \in \mathscr{A}$. Evidently, its commutator $\hat{\mathscr{A}}'$ consists of all $[b_{i,j}]$ with each $b_{i,j} \in \mathscr{A}'$. Thus, for all $a \in \mathscr{A}''$, $[a\delta_{i,j}] \in \hat{\mathscr{A}}''$.

Let $\eta = \eta_1 \oplus \cdots \oplus \eta_n$, and define $\mathscr{K} = \overline{\hat{\mathscr{A}}\eta}$ which is a closed subspace of $\hat{\mathscr{H}}$ that is invariant under $\hat{\mathscr{A}}$. By Lemma 5.5, the orthogonal projection p of $\mathscr{H}$ onto $\mathscr{K}$ belongs to $\hat{\mathscr{A}}'$, and hence to $\hat{\mathscr{A}}'''$. Then by Lemma 5.5 again, for $a \in \mathscr{A}''$, $\mathscr{K}$ is invariant under $\hat{\mathscr{A}}''$. In particular, for all $a \in \mathscr{A}''$, $\mathscr{K}$ is invariant under $[a\delta_{i,j}]$.

Since $\mathscr{A}$ contains the identity, $\eta \in \mathscr{K}$, so that $a\eta_1 \oplus \cdots \oplus a\eta_n \in \mathscr{K}$. Therefore, for all $\epsilon > 0$, there exists $b \in \mathscr{A}$ such that

$$\|b\eta_1 \oplus \cdots \oplus b\eta_n - a\eta_1 \oplus \cdots \oplus a\eta_n\|^2_{\mathscr{H}} \le \epsilon^2 \ .$$

□

In particular, the weak operator topological closure of any $*$-subalgebra of $\mathscr{B}(\mathscr{H})$ containing the identity is again a $*$-algebra containing the identity, and hence is a von Neumann algebra, though this can be seen directly.

For any self-adjoint operator a in $\mathscr{B}(\mathscr{H})$, note that $\{a\}' = (C(a))'$ where $C(a)$ is the C^* algebra generated by a. Hence $\{a\}''$ is the smallest von Neumann algebra that contains a. That is, $\{a\}''$ is *the von Neumann algebra generated by a*.

5.9 THEOREM. *On a separable Hilbert space $\mathscr{H}$, every abelian von Neumann algebra $\mathscr{Z}$ is generated by a single self adjoint operator; i.e, for some self adjoint $a \in \mathscr{Z}$, $\mathscr{Z} = \{a\}''$.*

Proof. Recall that subsets of separable spaces are separable. Let $\{p_n\}$ be a sequence of projections in $\mathscr{Z}$ that is dense for the strong operator topology in the set of all projections in $\mathscr{Z}$. Define

$$a = \sum_{j=1}^{\infty} 3^{-j} p_j \ .$$

The sum converges in operator norm, and hence belongs to $\mathscr{Z}$. Note also that $\|\sum_{j=2}^{\infty} 3^{-j}p_j\| \le 2/9$, and hence $\|a - 3^{-1}p_1\| \le 2/9$.

Pick $\lambda_0 \in (2/9, 1/3)$ and let $q = 1_{(\lambda_0,1)}(a)$. Then q and p_1 are commutng projections, and hence $q^\perp p_1$ and $p_1 q^\perp$ are projections.

If $qp_1^\perp \neq 0$, there is a unit vector η with $q\eta = \eta$ and $p_1^\perp\eta = \eta$. Then since $qaq \ge \lambda_0 q$,

$$\lambda_0 \le \langle \eta qaq\eta\rangle_{\mathscr{H}} = \langle \eta a\eta\rangle_{\mathscr{H}} = \langle \eta p_1^\perp a p_1^\perp \eta\rangle_{\mathscr{H}} \le \langle \eta(a - 3^{-1}p_1)\eta\rangle_{\mathscr{H}} \le 2/9 \ .$$

This is a contradiction, and so $qp_1^\perp = 0$.

If $q^\perp p_1 \neq 0$, there is a unit vector η with $q^\perp\eta = \eta$ and $p_1\eta = \eta$. Then since $\lambda_0 q^\perp \ge q^\perp a q^\perp$,

$$\lambda_0 \ge \langle \eta q^\perp a q^\perp \eta\rangle_{\mathscr{H}} = \langle \eta a\eta\rangle_{\mathscr{H}} = \langle \eta p_1 a p_1 \eta\rangle_{\mathscr{H}} \ge \langle \eta(3^{-1}p_1)\eta\rangle_{\mathscr{H}} \ge 1/3 \ .$$

This is a contradiction, and so $q^\perp p_1 = 0$.

Then since $q^\perp p_1 = qp_1^\perp = 0$, $q = qp_1 + qp_1^\perp = qp_1 = qp_1 + q^\perp p_1 = p_1$. This shows that the spectral projection of a for the interval $(\lambda_0, 1)$ is p_1. Inductively, one finds that each p_j is a spectral projection for a, and hence belongs to $\{a\}''$. □

5.2 Central covers

Let π be a non-degenerate representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$. If q is any projection in the center of $(\pi(\mathscr{A}))'$, which is $(\pi(\mathscr{A}))' \cap (\pi(\mathscr{A}))''$, the range of q is invariant under both $\pi(\mathscr{A})$ and $(\pi(\mathscr{A}))'$ and thus the restriction of π_q is a subrepresntation of π.

5.10 DEFINITION (Central projection for a representation)**.** Let π be a non-degenerate representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$. A *central projection* for π is a projection in the center of $(\pi(\mathscr{A}))'$

5.11 LEMMA. *Let π be a non-degenerate representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$, and p be a projection in $(\pi(\mathscr{A}))'$. There exits a central projection $\overline{p}$ that is dominated by every central projection that dominates p. Moreover, $\overline{p}$ is the projection onto*

$$\overline{(\pi(\mathscr{A}))'\mathrm{ran}(p)}\ .$$

Finally, if π_p is irreducible, $\overline{p}$ is the central cover of every projection $q \in (\pi(\mathscr{A}))'$ dominated by $\overline{p}$.

Proof. If q is a projection in the center of $(\pi(\mathscr{A}))$ that dominates p, then the range of q contains $\mathscr{K}$, and since q commutes with $(\pi(\mathscr{A}))'$, $(\pi(\mathscr{A}))'\mathscr{K}$ is contained in the range of q.

For the final part, suppose that $q \in (\pi(\mathscr{A}))'$ is dominated by $\overline{p}$. Since $\overline{p}$ is a central projection that dominates q, $\overline{q} \leq \overline{p}$. It suffices to show that $p \leq \overline{q}$.

Note that $p\overline{q}$ and $p\overline{q}^\perp$ are projections in $(\pi(\mathscr{A}))'$ that are dominated by p. Since π_p is irreducible, one must be zero, and the other must be p. If $p\overline{q}^\perp = p$, then $\overline{p\overline{q}^\perp}$ is a central projection dominating p, and hence $\overline{p\overline{q}^\perp} = \overline{p}$. This is impossible since $q \leq \overline{p}$. Hence it must be the case that $p\overline{q} = p$, which is what we needed to show. □

5.12 DEFINITION (Central cover and central subrepresentations)**.** Let π be a non-degenerate representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$, and let σ be a subrepresentation of π on a subspace $\mathscr{K}$ of $\mathscr{H}$. Let p be the projector onto $\mathscr{K}$. Then the *central cover of* σ is the representation $\overline{\sigma} = \pi_{\overline{\sigma}}$. A subrepresntation σ of π is a *central subrepresentation* in case $\sigma = \overline{\sigma}$.

Unless σ is already a central representation, its central cover is a strictly larger subrepresentation of π. The precise sense in which it is larger makes notion of central representations fundamentally important in the study of the structure of representations.

5.13 LEMMA. *Let π be a non-degenerate representation of a C^* algebra on a Hilbert space $\mathscr{H}$. Let p and q be two projections in $(\pi(\mathscr{A}))'$ and suppose that they have the same central cover; i.e., $\overline{p} = \overline{q}$. Then there is a partial isometry $u \in (\pi(\mathscr{A}))'$ such that $uu^* \leq q$ and $u^*u \leq p$.*

Proof. Let $\mathscr{K}_1$ and $\mathscr{K}_2$ denote the ranges of p and q respectively. Since $\overline{(\pi(\mathscr{A}))'\mathscr{K}_1} = \overline{(\pi(\mathscr{A}))'\mathscr{K}_2}$, there is an $a \in (\pi(\mathscr{A}))'$ and vectors η_1 and η_2 in $\mathscr{K}_1$ and $\mathscr{K}_2$ respectively so that $\langle \eta_2, a\eta_2 \rangle_{\mathscr{H}} \neq 0$. This means that $z = qap$ is a non-zero element of $\pi(\mathscr{A}))'$. Let $z = u|z|$ be its polar decomposition. Then u is a partial isometry in $\pi(\mathscr{A}))'$, such that $uu^* \leq q$ and $u^*u \leq p$. Thus, π_{u^*u} is a subrepresentation of π_p that is equivalent to π_{uu^*}, a subrepresentation of q. □

Lemma 5.13 has the following consequence: Since for any partial isometry u u^*u and uu^* are, respectively, the projectors onto the final and initial spaces of u, π_{u^*u} and π_{uu^*} are non-zero equivalent subrepresentations of the representation π_p and π_q discussed in the lemma. In particular, if π_p is irreducible, there is a partial isometry $u \in (\pi(\mathscr{A}))'$ such that $u^*u = p$ and $uu^* \leq q$, so that π_p is equivalent to a subrepresentation of π_q. The fact that the equivalence is due to a partial isometry in $\pi(\mathscr{A}))'$ is important in what follows.

5.3 The structure of type I factors of von Neumann algebras

5.14 DEFINITION. Let $\mathscr{H}$ be a Hilbert space, and let $\mathscr{M}$ be a von Neumann algebra on $\mathscr{H}$. $\mathscr{M}$ is a *factor* in case $\mathscr{M}$ has a trivial center.

Note that the center of $\mathscr{M}$ is $\mathscr{M} \cap \mathscr{M}'$, which by the Double Commutant Theorem, is the same as $\mathscr{M}' \cap \mathscr{M}''$, so that $\mathscr{M}$ and $\mathscr{M}'$ have the same center.

The center $\mathscr{Z}$ of $\mathscr{M}$ is evidently an abelian von Neumann algebra, and therefore, by Theorem 5.9, when $\mathscr{M}$ is a von Neumann algebra on a separable Hilbert space $\mathscr{H}$, its center $\mathscr{Z}$ is generated by a single self-adjoint $a \in \mathscr{Z}$, and every spectral projection of a is a central projection for the identity representations of $\mathscr{M}$ and $\mathscr{M}'$.

Now suppose that this operator a happens to have finite spectrum, as it must in case $\mathscr{H}$ is finite dimensional. Then there is a finite set $\{p_1, \ldots, p_n\}$ of central projections. For each $j = 1, \ldots, n$, let $\mathscr{K}_j$ denote the range of p_j. Then $\mathscr{H} = \bigoplus_{j=1}^n \mathscr{K}_j$, and defining $\mathscr{M}_j = \mathscr{M} p_j$,

$$\mathscr{M} = \bigoplus_{j=1}^{n} \mathscr{M}_j \ . \tag{5.1}$$

Evidently, each $\mathscr{M}_j$ has a trivial center; its center is spanned by the identity on $\mathscr{K}_j$. The Spectral Theorem can be used to give a "direct integral" decomposition on $\mathscr{M}$ without making any assumption on the spectrum of $\mathscr{A}$, as was shown by von Neumann. This line of reasoning reduces the investigation of the structure of von Neumann algebras on a separable Hilbert space that that of von Neumann algebras with trivial center, which motivates the following definition:

5.15 DEFINITION (Factor)**.** A *factor* $\mathscr{M}$ is a von Neumann algebra $\mathscr{M}$ with a trivial center. A factor is *type I* in case it contains a non-zero minimal projection; i.e., a non-zero projection p such that the only projection in $\mathscr{M}$ that is dominated by p is the zero projection.

Evidently every factor on a finite dimensional Hilbert space contains a minimal projection – any projection whose range has minimal dimension – and so every factor on a finite dimensional Hilbert space is type I. This is not true for infinite dimensional Hilbert spaces, and we shall return to a classification of types of factors and investigate their structure later. For the rect of this subsection, we focus on the structure of type I factors.

Looking at (5.1), one might think "summand" would be better terminology than "factor", but the following theorem justifies the terminology:

5.16 THEOREM. *Let $\mathscr{K}$ be a separable Hilbert space, and let $\mathscr{M}$ be a type I factor on $\mathscr{K}$. Then there exist Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ and a unitary $u : \mathscr{K} \to \mathscr{H}_1 \otimes \mathscr{H}_2$ such that*

$$u\mathscr{M}u^* = \mathscr{B}(\mathscr{H}_1) \otimes 1_{\mathscr{H}_2} \ . \tag{5.2}$$

5.17 REMARK. *The commutant of $\mathscr{M}'$ of $\mathscr{M} = \mathscr{B}(\mathscr{H}_1) \otimes 1_{\mathscr{H}_2}$ in $\mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$ is evidently $1_{\mathscr{H}_1} \otimes \mathscr{B}(\mathscr{H}_2)$, and hence $\mathscr{M} \cap \mathscr{M}'$ consists of mutiples of the idenetity. Hence $\mathscr{B}(\mathscr{H}_1) \otimes 1_{\mathscr{H}_2}$ is a factor in $\mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$. The theorem says that all factors are of this type, and when $\mathscr{M}$ is a factor on a separable Hilbert space $\mathscr{K}$, $\mathscr{B}(\mathscr{K})$ is the closed span of elements of the form ab where $a \in \mathscr{M}$ and $b \in \mathscr{M}'$, which may be viewed as a kind of factorization of $\mathscr{B}(\mathscr{K})$.*

Proof of Theorem 5.16. If $\mathscr{M}$ consists of multiples of the identity, then we may take $\mathscr{H}_1 = \mathbb{C}$, and the conclusion is obvious. Therefore, let us us assume that $\mathscr{M}$ does not consist of multiples of the identity, or, what is the same thing by Theorem 5.7: the identity representation of $\mathscr{M}'$ is reducible.

Let π denote the identity representation of $\mathscr{M}'$, whose commutant is $\mathscr{M}$. Let p_1 be a minimal projection in $\mathscr{M}$. We now apply Lemma 5.13 in this setting. Since the center of $\mathscr{M}$ is trivial, the central covers of both p_1 and ${p_1}^\perp$ are the identity in $\mathscr{M}$, and since p_1 is minimal, π_{p_1} is an irreducible representation of $\mathscr{M}'$. Hence by Lemma 5.13, there is a partial isometry $u \in \mathscr{M}$ such that $u^*u = p_1$ and $uu^* \le p_1^\perp$. Since $p_1 = p_1^2 = u^*(uu^*)u$, uu^* is also minimal.

Define $\mathscr{K}_j = p_j\mathscr{K}$ for $j = 1, 2$. If $\mathscr{K} = \mathscr{H}_1 \oplus \mathscr{H}_2$, we have decomposed $\mathscr{K}$ as a direct sum of subspaces on which $\mathscr{M}'$ acts irreducibly and equivalently. If not, repeat the argument made above with $p_1^\perp$ replaced by $1 - p_1 - p_2$, thus producing a minimal projection p_3 in $\mathscr{M}$ with $p_3p_j = 0$ for $j = 1, 2$, and u_3, an isometry in $\mathscr{M}$ that maps $\mathscr{H}_1$ onto the range of p_3. If after some finite number n of such steps, $\mathscr{K}$ is exhausted, we have produced a set $\{p_1, \dots, p_n\}$ of minimal projections in $\mathscr{M}$ with $p_ip_j = 0$ for $i \ne j$, and a set $\{u_1, \dots, u_m\}$ of partial isometries in $\mathscr{M}$ where u_j maps $\mathscr{H}_1$ onto $\mathscr{K}_j$, the range of p_j. (Note that u_1 is p_1 itself.)

Consider any $a \in \mathscr{M}$. We claim that there is a matrix $[a] \in M_n(\mathbb{C})$ such that

$$a = \sum_{i,j=1}^n [a]_{i,j} u_j u_i^* \ . \tag{5.3}$$

To see this observe that

$$a = \sum_{i,j=1}^n p_i a p_j = \sum_{i,j=1}^n u_i(u_i^* a u_j)u_j^* \ .$$

However, for each i, j, $u_i^* a u_j \in p_1\mathscr{M}p_1$, and since p_1 is minimal, $p_1\mathscr{M}p_1 = \mathbb{C}p_1$. Hence for some $\lambda_{i,j} \in \mathbb{C}$, $u_i^* a u_j = \lambda_{i,j}p_1$. Define $[a]_{i,j} = \lambda_{i,j}$ to obtain (5.3).

Let $\{\zeta_1, \dots, \zeta_n\}$ denote the standard basis of $\mathbb{C}^n$. Define a linear transformation u from $\mathbb{C}^n \otimes \mathscr{H}_1$ to $\mathscr{K}$ by

$$u\left(\sum_{j=1}^n \zeta_j \otimes \eta_j\right) = \sum_{j=1}^n u_j\eta_j \ .$$

It is evident that this map is unitary. Moreover, for any $a \in \mathscr{M}$, using (5.3), we have

$$au\left(\sum_{j=1}^n \zeta_j \otimes \eta_j\right) = \sum_{i,j,\ell=1}^n [a]_{i,j} u_j u_i^* u_\ell \eta_\ell =$$

$$\sum_{i,j=1}^n [a]_{i,j} u_j \eta_i = u\left(\sum_{j=1}^n \left(\sum_{i=1}^n [a]_{i,j}\zeta_i\right) \otimes \eta_j\right) = u\left(\sum_{j=1}^n [a]\zeta_j \otimes \eta_j\right)$$

That is, with $\mathscr{H}_1 = \mathbb{C}^n$ and $\mathscr{H}_2 = \mathscr{H}_1$,

$$uau^* = [a] \otimes 1_{\mathscr{H}_2} \ .$$

This proves (5.2) in case the procedure for producing a sequence of orthogonal minimal projections terminates in finitely many steps.

When this process does not terminate, a simple application of Zorn's Lemma shows that there is a sequence $\{p_n\}$ of minimal projections in $\mathscr{M}$ such that $p_m p_n = 0$ for $m \neq n$ and $\mathscr{K} = \bigoplus_{n=1}^{\infty} p_n \mathscr{K}$, and moreover, there is a sequence $\{u_n\}$ of partial isometries in $\mathscr{M}$ such that u_n maps $\mathscr{H}_1$ onto $\mathscr{K}_n$. Then the strong closure of the set of operators of the form (5.3) for some $n \in \mathbb{N}$ is easily seen to be dense in $\mathscr{M}$, and then with $\mathscr{H}_1 = \ell_2$, we obtain (5.2) in this case as well. □

Theorem 5.16 has an important corroalary:

5.18 COROLLARY. *Let $\mathscr{H}$ and $\mathscr{K}$ be separable Hilbert spaces, and let π be an injective represetnation of $\mathscr{B}(\mathscr{H})$ on $\mathscr{K}$. Then there exists a Hilbert space $\mathscr{H}'$ and a unitary $u : \mathscr{K} \to \mathscr{H} \otimes \mathscr{H}'$ such that for all $a \in \mathscr{B}(\mathscr{H})$,*

$$u\pi(a)u^* = a \otimes 1_{\mathscr{H}'} \ . \tag{5.4}$$

Proof. Since π is injective, by Theorem 2.26 π is an ismetric $*$-isomorphism from $\mathscr{B}(\mathscr{H})$ onto $\pi(\mathscr{B}(\mathscr{H}))$. Then evidently $\pi(\mathscr{B}(\mathscr{H}))$ is a type I factor in $\mathscr{B}(\mathscr{K})$, and Theorem 5.16 applies.

In the setting of Corollary 5.18, we can be somewhat more explicit about the construction used in the proof of Theorem 5.16, and this leads to (5.4). Let $\{\eta_j\}$ be any orthonormal basis for $\mathscr{H}$, and for each j, let q_j be the orthognal projection onto the span of η_j, For each j, define $p_j = \pi(q_j)$. This gives us a family of ortogonal projections in $\pi(\mathscr{B}(\mathscr{H}))$ that are minimal and satisfy $p_i p_j = 0$ for all i, j. A simple maximality argument shows that $\sum_{j=1} p_j = 1_{\mathscr{K}}$.

Let $\mathscr{H}'$ be the range of p_1. By Lemma 5.13 there are partial isometries $u_j \in \pi(\mathscr{B}(\mathscr{H}))$ such that u_j has $\mathscr{H}'$ as its initial space and $\mathrm{ran}(p_j)$ as its final space, and $p_j = u_j u_j^*$. Now pick any orthonormal basis $\{\zeta_k\}$ of $\mathscr{H}'$. Note that $\{u_j \zeta_k\}$ is an orthonormal basis of $\mathscr{K}$. Define the unitary $u : \mathscr{K} \to \mathscr{H} \otimes \mathscr{H}'$ by $u(\eta_j \otimes \zeta_k) = u_j \zeta_k$. Then for all j, k, ℓ, m,

$$\langle u^*(\eta_\ell \otimes \zeta_m), u\pi(a)u^*(\eta_j \otimes \zeta_k)\rangle_{\mathscr{K}} = \langle \zeta_m, (u_\ell \pi(a) u_j^*)\zeta_k\rangle_{\mathscr{H}'}$$

As noted in the proof of Theorem 5.16, $u_\ell \pi(a) u_j^* \in p_(\pi(\mathscr{B}(\mathscr{H}))p_1$, and then since p_1 is minimal, it acts trivially on $\mathscr{H}'$. Therefore, if we define $[a]_{\ell,j} \in \mathbb{C}$ by $u_\ell \pi(a) u_j^* = [a]_{\ell,j} 1_{\mathscr{H}'}$, we have

$$\langle (\eta_\ell \otimes \zeta_m), u\pi(a)u^*(\eta_j \otimes \zeta_k)\rangle_{\mathscr{K}} = [a]_{\ell,j}\delta_{m,k} \ . \tag{5.5}$$

Now for each j let $\tilde{u}_j = \pi^{-1}(u_j)$. Then evidently $\tilde{u}_j \tilde{u}_j^* = q_j$ and $\tilde{u}_j^* \tilde{u}_j = q_1$. Hence, up to a complex multiple of modulus 1, $\tilde{u}_j$ is the rank one transformation sending η_1 into η_j, and we can absorb this multiple into the definitition of our basis $\{\eta_j\}$. Hence with this choice,

$$u_\ell \pi(a) u_j^* = \pi(\tilde{u}_\ell)\pi(a)\pi(\tilde{u}_j) = \pi(\tilde{u}_\ell a \tilde{u}_j) = \pi(\langle \eta_\ell, a\eta_j\rangle_{\mathscr{H}} q_i) = \langle \eta_\ell, a\eta_j\rangle_{\mathscr{H}} p_1 \ .$$

Going back to (5.6), we have that $[a]_{\ell,j} = \langle \eta_\ell, a\eta_j\rangle_{\mathscr{H}}$, and thus we may rewrite (5.6) as

$$\langle (\eta_\ell \otimes \zeta_m), u\pi(a)u^*(\eta_j \otimes \zeta_k)\rangle_{\mathscr{K}} = \langle (\eta_\ell \otimes \zeta_m) a \otimes 1_{\mathscr{H}'}(\eta_j \otimes \zeta_k)\rangle_{\mathscr{K}} \ , \tag{5.6}$$

and this proves (5.4). □

5.4 States on a C^* algebra

5.19 DEFINITION. Let $\mathscr{A}$ be a C^* algebra. A linear functional φ on $\mathscr{A}^*$, the Banach space dual to $\mathscr{A}$ regarded as a Banach space, is *positive* in case $\varphi(a) \geq 0$ for all $a \geq 0$. If $\mathscr{A}$ has an identity 1, a *state* on $\mathscr{A}$ is a positive linear functional φ such that $\varphi(1) = 1$. We denote the set positive linear functionals by $\mathscr{A}^*_+$ and the states by $\mathscr{A}^*_{+,1}$. A state $\varphi \in \mathscr{A}^*_{+,1}$ is *faithful* in case

$$\varphi(a^*a) = 0 \quad \Rightarrow \quad a = 0 \ . \tag{5.7}$$

Evidently, for all $\varphi \in \mathscr{A}^*_+$, the map

$$(a, b) \mapsto \varphi(a^*b) = \langle a, b\rangle_\varphi$$

defines a (possibly degenerate) inner product on $\mathscr{A}$; this inner product is non-degenerate if and only if φ is faithful. In any case, the fact that $\langle a, a\rangle_\varphi \geq 0$ for all $a \in \mathscr{A}$ yields the Cauchy-Schwarz inequality:

$$|\langle a, b\rangle_\varphi| \leq \langle a, a\rangle_\varphi^{1/2}\langle b, b\rangle_\varphi^{1/2} \ . \tag{5.8}$$

5.20 THEOREM (Positivity and continuity)**.** *Let $\mathscr{A}$ be a C^* algebra with identity* 1*. Then:*

(1) Every $\varphi \in \mathscr{A}^*_+$ is bounded, and $\|\varphi\| = \varphi(1)$.

(2) Every bounded linear functional φ such that $\|\varphi\| = \varphi(1)$ is positive.

Proof. Let $\varphi \in \mathscr{A}^*_+$. For all $a \in \mathscr{A}$,

$$|\varphi(a)| = |\varphi(1a)| = |\langle 1, a\rangle_\varphi| \leq \langle 1, 1\rangle_\varphi^{1/2}\langle a, a\rangle_\varphi^{1/2} = \varphi(1)^{1/2}\varphi(a^*a)^{1/2} \ .$$

Since $\sigma_{\mathscr{A}}(a^*a) \subset [0, \|a\|^2]$, $\|a\|^2 1 - a^*a \geq 0$, and hence $\varphi(a^*a)^{1/2} \leq \|a\|\varphi(1)^{1/2}$. Combining these inequalities, we have $|\varphi(a)| \leq \varphi(1)\|a\|$ which proves *(1)*.

For the second part, suppose that $\varphi \in \mathscr{A}^*$ and $\varphi(1) = \|\varphi\|$. If $\varphi - 0$, it is positive. If $\varphi \neq 0$, we may divide by $\|\varphi\|$ and thus may suppose that $\|\varphi\| = \varphi(1) = 1$.

We claim that for all $\varphi \in \mathscr{A}^*$ such that $\varphi(1) = \|\varphi\|$, $\varphi(a)$ belongs to the convex hull of $\sigma_{\mathscr{A}}(a)$ for all $a \geq 0$ in $\mathscr{A}$. To see this suppose that the closed disc of radius r centered on λ contains $\sigma_{\mathscr{A}}(a)$. Then $\lambda 1 - a$ is normal, and its spectrum is continued in $\{\lambda - t \ : \ t \in \sigma_{\mathscr{A}}(a)\}$, and hence the spectral radius of $\lambda 1 - a$ is at most r. Since $\lambda 1 - a$ is normal, $\|\lambda 1 - a\| \leq r$. Therefore,

$$|\lambda - \varphi(a)| = |\varphi(\lambda 1 - a)| \leq \|\lambda 1 - a\| \leq r \ .$$

Thus for all $r > 0$ and $\lambda \in \mathbb{C}$, $\varphi(a)$ is contained in the closed disc of radius r centered on λ contains $\sigma_{\mathscr{A}}(a)$. The intersection over all such discs is the convex hull of $\sigma_{\mathscr{A}}(a)$. □

5.21 LEMMA. *Let $\mathscr{A}$ be a C^* algebra with identity* 1*. For all self adjoint $a \in \mathscr{A}$, there exists a state φ such that $|\varphi(a)| = \|a\|$.*

Proof. Consider the C^* algebra $C(a)$ generated by a and 1. This is a commutative C^* algebra, and so there is a character φ_0 of $C(a)$ such that $|\varphi_0(a)| = \|a\|$, and since φ_0 is a character $\varphi(1) = 1$. Then by Theorem 5.20, $\varphi_0 \in \mathscr{A}^*_+$, and so φ is a state on $C(a)$.

By the Hahn-Banach Theorem, there is a norm preserving extension φ of φ_0 (as a linear functional) to $\mathscr{A}$. Then $\varphi(1) = \varphi_0(1) = 1$, and hence by Theorem 5.20, φ is a state, and since φ extends φ_0, $\varphi(x) = \|x\|$. □

Lemma 5.21 says, in particular, that $\mathscr{A}^*_{+,1}$ is not empty. It is evidently a closed subset of the unit ball in $\mathscr{A}^*$ in the weak-$*$ topology, and hence is compact. $\mathscr{A}^*_{+,1}$ is also evidently convex. The Krein-Milman Theorem says that every non-empty convex set in $\mathscr{A}^*$ that is compact in the weak-$*$ topology is the convex hull of its extreme points. Hence there exist extreme points in $\mathscr{A}^*_{+,1}$.

5.22 DEFINITION (Pure state)**.** Let $\mathscr{A}$ be a C^* algebra with identity 1. A *pure state* is an extreme point of $\mathscr{A}^*_{+,1}$.

5.23 THEOREM. *Let $\mathscr{A}$ be a C^* algebra with identity* 1*. For all self adjoint $a \in \mathscr{A}$, there exists a pure state φ such that $|\varphi(a)| = \|a\|$.*

Proof. By Lemma 5.21, the set $\mathcal{S}$ of states φ such that $\varphi(a) = \|a\|$ is non-empty, and evidently it is convex and closed in the weak-$*$ topology. By the Krein-Milman, $\mathcal{S}$ has at least one extreme point ψ. We now show that ψ is extreme in $\mathscr{A}^*_{+,1}$ as well as in $\mathcal{S}$.

Suppose that $\psi_1, \psi_2 \in \mathscr{A}^*_{+,1}$ and that $\psi = t\psi_1 + (1-t)\psi_2$ for some $t \in (0,1)$. Evaluating both sides at a,

$$\|a\| = \psi(a) = t\psi_1(a) + (1-t)\psi_2(a) \leq t\|a\| + (1-t)\|a\| = \|a\| \ .$$

Hence $\psi_1, \psi_2 \in \mathcal{S}$, and so $\psi_1 = \psi_2 = \psi$. □

5.24 DEFINITION. Let π be a representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$. A vector $\eta \in \mathscr{H}$ is cyclic for π in case $a \mapsto \pi(a)\eta$ has dense range, and is a *separating* vector for π in case $a \mapsto \pi(a)\eta$ is injective. If a cyclic vector exists, then π is a *cyclic representation.*

For any representation π of $\mathscr{A}$ on $\mathscr{H}$, and any *unit vector* $\eta \in \mathscr{H}$, the functional $\eta_\eta \in \mathscr{A}^*$ defined by

$$\varphi_\eta(a) = \langle \eta, \pi(a)\eta \rangle_{\mathscr{H}} \tag{5.9}$$

is a state. Evidently, $\varphi_\eta(a^*a) = \langle \eta, \pi(a^*a)\eta \rangle_{\mathscr{H}} = \|\pi(a)\eta\|^2_{\mathscr{H}}$, and hence η is separating for π is and ony if φ_η is faithful. The next theorem links gives an cyclicty an irreducibility.

5.25 THEOREM. *Let π be a representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$, and let η be a cyclic unit vector for π. Then with φ_η denoting the state defined in (5.9). Then π is irreducibly iff and only if φ_η is pure.*

The heart of the matter is the following lemma:

5.26 LEMMA. *Let π be a representation of a C^* algebra $\mathscr{A}$ on a Hilbert space $\mathscr{H}$, and let η be a cyclic unit vector for π. Then with φ_η denoting the state defined in (5.9). Suppose that $\psi \in \mathscr{A}^*_{+,1}$, and that for some $r \in (0,\infty)$,*

$$\psi(a) \leq r\varphi_\eta(a) \quad \text{for all} \quad a \in \mathscr{A} \ .$$

Then there is a positive operator $x \in (\pi(\mathscr{A}))'$ such that $\|x\| \leq r$ and for all $a, b \in \mathscr{A}$,

$$\psi(a^*b) = \langle \eta, \pi(a), x\pi(b)\eta \rangle_{\mathscr{H}} \ . \tag{5.10}$$

Proof. Define a sesquilinear form q on $\pi(\mathscr{A})\eta$ by $q(\pi(a)\eta, \pi(b)\eta) = \psi(a^*b)$. We have

$$|q(\pi(a)\eta, \pi(b)\eta)| \leq r|\langle \pi(a)\eta, \pi(b)\eta\rangle_{\mathscr{H}}| \leq r\|\pi(a)\|_{\mathscr{H}}\|\pi(b)\|_{\mathscr{H}} \ .$$

Since η is cyclic, q is densely defined on $\mathscr{H}$ and extends to a sesquilinear form on all of $\mathscr{H}$, still denoted by q, that satisfies $|q(\zeta, \xi)| \leq r\|\zeta\|_{\mathscr{H}}\|\xi\|_{\mathscr{H}}$ for all $\zeta, \xi \in \mathscr{H}$. By Reisz's Lemma, there exists a self adjoint operator $x \in \mathscr{B}(\mathscr{H})$ such that $q(\zeta, \xi) = \langle \zeta, x\xi\rangle_{\mathscr{H}}$ for all $\zeta, \xi \in \mathscr{H}$, and $\|x\| \leq r$. Since $q(\zeta, \zeta) \geq 0$ for ζ in the dense set $\pi(\mathscr{A})\eta$, x is positive.

Finally, note that for all $a, b, c \in \mathscr{A}$, $a^*(bc) = (b^*a)^*c$, and hence $\psi(a^*(bc)) = \psi((b^*a)^*c)$. This means that $q(\pi(a)\eta, \pi(b)\pi(c)\eta) = q(\pi(b^*)\pi(a)\eta, \pi(c)\eta)$ which is the same as

$$\langle \pi(a)\eta, x\pi(b)\pi(c)\eta\rangle_{\mathscr{H}} = \langle \pi(a)\eta, \pi(b)x\pi(c)\eta\rangle_{\mathscr{H}} \ .$$

Thus for all ζ, ξ in a dense subset of $\mathscr{H}$, $\langle \zeta, x\pi(b)\xi\rangle_{\mathscr{H}} = \langle \zeta, \pi(b)x\xi\rangle_{\mathscr{H}}$ and this shows that x commutes with $\pi(b)$ for arbitrary $b \in \mathscr{A}$. □

Proof of Theorem 5.25. Suppose that π is irreducible. Let ψ_1, ψ_2 be two states such that $\varphi_\eta = t\psi_1 + (1-t)\psi_2$ for some $t \in (0, 1)$. By Lemma 5.26, applied to ψ_1, which satisfies $\psi_1 \leq t^{-1}\varphi_\eta$, there is a positive $x \in (\pi(\mathscr{A}))'$

$$\psi_1(a^*b) = \langle \eta, \pi(a), x\pi(b)\eta\rangle_{\mathscr{H}} \quad \text{for all} \quad a, b \in \mathscr{A} \ . \tag{5.11}$$

Since π is irreducible, x must be a scalar multiple of the identity. Since ψ_1 is a state, taking $a = b = 1$ in (5.11), $1 = \psi(1) = \langle \eta, x\eta\rangle_{\mathscr{H}}$, which shows that $x = 1$. Then taking $a = 1$ in (5.11) shows that $\psi(b) = \varphi_\eta(b)$ so that $\psi_1 = \varphi_\eta$. By symmetry, $\psi_2 = \varphi_\eta$ as well, and this proves φ_η is extreme.

For the converse, suppose that π is not irreducible. Then there exists a projection $p \in (\pi(\mathscr{A})')$ such that neither p nor $p^\perp$ is zero. Suppose that $p\eta = 0$. Then for all $a \in \mathscr{A}$, $\pi(q)p\eta = p(\pi(a)\eta) = 0$ and this would mean that p vanishes on a sense subspace, which is not the case. Hence $\|p\eta\|_{\mathscr{H}} > 0$, and the same reasoning shows that $\|p^\perp\eta\|_{\mathscr{H}} > 0$. Define $\eta_1 = \|p\eta\|_{\mathscr{H}}^{-1}p\eta$ and $\eta_2 = \|p^\perp\eta\|_{\mathscr{H}}^{-1}p^\perp\eta$. For all $a \in \mathscr{A}$,

$$\langle \eta_1, \pi(a)\eta_2\rangle_{\mathscr{H}} = \langle p\eta_1, \pi(a)p^\perp\eta_2\rangle_{\mathscr{H}} = \langle \eta_1, pp^\perp\pi(a)\eta_2\rangle_{\mathscr{H}} = 0 \ .$$

Define $t \in (0, 1)$ by $t = \|p\eta\|_{\mathscr{H}}^2$. Since $\|p\eta\|_{\mathscr{H}}^2 + \|p^\perp\eta\|_{\mathscr{H}}^2 = 1$, $\|p^\perp\eta\|_{\mathscr{H}}^2 = 1 - t$. Then by the orthogonality proved just above, for all $a \in \mathscr{A}$,

$$\begin{aligned}\varphi_\eta(a) &= \langle [\sqrt{t}\eta_1 + \sqrt{1-t}\eta_2], \pi(a)[\sqrt{t}\eta_1 + \sqrt{1-t}\eta_2]\rangle_{\mathscr{H}} \\ &= t\langle \eta_1, \pi(a)\eta_1\rangle_{\mathscr{H}} + (1-t)\langle \eta_2, \pi(a)\eta_2\rangle_{\mathscr{H}} \ ,\end{aligned}$$

and this displays φ_η as a non-trivial convex combination of states. Hence φ_η is not extreme. □

5.5 The GNS construction

A construction due to Gelfand, Naimark and Segal, known as the GNS construction, associates to every state φ on an C^* algebra $\mathscr{A}$ a representation π of $\mathscr{A}$ on a Hilbert space built out of $\mathscr{A}$ itself and the state φ.

5.27 THEOREM (The GNS construction). *Let $\mathscr{A}$ be a C^* algebra with identity* 1*, and let φ be a state on $\mathscr{A}$. Then there exists a Hilbert space $\mathscr{H}$ and a cyclic representation π of $\mathscr{A}$ on $\mathscr{H}$ with a distinguished cyclic unit vector η such that for all $a \in \mathscr{A}$,*

$$\varphi(a) = \langle \eta, \pi(a)\eta \rangle_{\mathscr{H}} \ . \tag{5.12}$$

The representation π is irreducible if and only if φ is a pure state.

Proof. Let $\langle a, b\rangle_\varphi$ be the possibly degenerate inner product on $\mathscr{A}$ defined by $\langle a, b\rangle_\varphi = \varphi(a^*b)$. Define

$$\mathscr{N} := \{ \ a \in \mathscr{A} \ : \ \langle a, a\rangle_\varphi = 0 \ \} \ .$$

Since φ is continuous, $\mathscr{N}$ is closed. In fact, $\mathscr{N}$ is a closed left ideal. To see this, consider $b \in \mathscr{A}$ and $a \in \mathscr{N}$. Then

$$\langle ba, ba\rangle_\varphi = \varphi(a^*b^*ba) = \langle a, b^*ba\rangle_\varphi \leq \langle a, a\rangle_\varphi^{1/2} \langle b^*ba, b^*ba\rangle_\varphi^{1/2} = 0 \ .$$

A similar but simpler argument shows that $\mathscr{N}$ is a subspace.

Now consider the vector space $\mathscr{A}/\mathscr{N}$. With $\sim$ denoting equivalence mod $\mathscr{N}$, we have

$$a \sim a' \quad \text{and} \quad b \sim b' \quad \Rightarrow \quad \langle a, b\rangle_\varphi = \langle a', b'\rangle_\varphi \ ,$$

and hence we may define a *non-degenerate* inner product on $\mathscr{A}/\mathscr{N}$ by $\langle \{a\}, \{b\}\rangle = \langle a, b\rangle_\varphi$. Let $\mathscr{H}$ be the completion of $\mathscr{A}/\mathscr{N}$ in the corresponding Hilbertian norm, and let $\langle \cdot, \cdot\rangle_{\mathscr{H}}$ denote the resulting inner product on $\mathscr{H}$.

For $a \in \mathscr{A}$, let $\pi(a)$ denote the linear operator on $\mathscr{A}/\mathscr{N}$ defined by $\pi(a)\{b\} = \{ab\}$ which is well-defined since $\mathscr{N}$ is a left ideal. Next note that since $b^*a^*ab = \|a\|^2 b^*b + b^*(\|a^*a\|1 - a^*a)b$, and $b^*(\|a^*a\|1 - a^*a)b$ is positive,

$$\|\pi(a)\{b\}\|_{\mathscr{H}}^2 = \varphi(b^*a^*ab) \leq \|a\|^2 \varphi(b^*b) = \|a\|^2 \|\{b\}\|_{\mathscr{H}}^2 \ .$$

Since $\mathscr{A}/\mathscr{N}$ is dense in $\mathscr{H}$, $\pi(a)$ extends to a bounded operator on $\mathscr{H}$ with $\|\pi(a)\| \leq \|a\|$. It is evident that π is a homomorphism of $\mathscr{A}$ into $\mathscr{B}(\mathscr{H})$, and note that for all $x, y \in \mathscr{A}$,

$$\langle \{x\}, \pi(a)\{y\}\rangle_{\mathscr{H}} = \varphi(x^*ay) = \varphi((a^*x)^*y = \langle \pi(a)\{x\}, \{y\}\rangle_{\mathscr{H}} \ ,$$

showing that $\pi(a^*) = \pi(a)^*$, and thus π is a $*$-homomorphism.

The representation π is cyclic since for all $a \in \mathscr{A}$, $\{a\} = \{a1\} = \pi(a)\{1\}$, showing that $\eta := \{1\}$ is a cyclic vector for π. Finally, note that $\langle \eta, \pi(a)\eta\rangle_{\mathscr{H}} = \varphi(1^*a1) = \varphi(a)$, and this proves (5.12). The final statement now follows from Theorem 5.25. □

5.28 COROLLARY. *Let $\mathscr{A}$ be a C^* algebra with identity* 1*. For every non-zero $a \in \mathscr{A}$, there is a representation π of $\mathscr{A}$ such that $\|\pi(a)\| = \|a\|$.*

Proof. By Lemma 5.21, there exists $\varphi \in \mathscr{A}_{+,1}^*$ such that $|\varphi(a^*a)| = \|a\|^2$. Let π be the GNS representation of $\mathscr{A}$ associated to φ, and η the associated distinguished cyclic unit vector. Then

$$\|\pi(a)\eta\|_{\mathscr{H}}^2 = \langle \eta \pi(a^*a)\eta\rangle_{\mathscr{H}} = \varphi(a^*a) = \|a\|^2 \ ,$$

showing that $\|\pi(a)\| \geq \|a\|$, and since it is automatic that $\|\pi(a)\| \leq \|a\|$, $\|\pi(a)\| = \|a\|$. □

We now arrive at the Non-Commutative Gelfand-Naimark Theorem:

5.29 THEOREM (Non-Commutative Gelfand-Naimark Theorem). *Every C^* algebra $\mathscr{A}$ with an identity is isometrically $*$-isomorphic to a C^* algebra of operators.*

Proof. For each $a \in \mathscr{A}$ choose an irreducible representation π of $\mathscr{A}$ such that $\|\pi(a)\| = \|a\|$. Now form the direct sum of all of these representations. □

5.6 The GNS construction for $M_n(\mathbb{C})$ and the normalized trace

Fis $n \in \mathbb{N}$, $n \geq 2$, and let $\mathscr{A} = M_n(\mathbb{C})$. Define $\varphi_{\rm tr} \in \mathscr{A}^*$ by

$$\varphi_{\rm tr}(a) = \frac{1}{n}{\rm Tr}(a) \ .$$

Since $\varphi_{\rm tr}(a^*a) = \sum_{i,j=1}^n |a_{i,j}|^2$, it is evident that $\varphi_{\rm tr}(a)$ is a state, called the *normalized trace*. It is also evident from the same computation that the normalized trace is faithful. It has one more important property: For all $a, b \in \mathscr{A}$,

$$\varphi_{\rm tr}(ab) = \varphi_{\rm tr}(ba) \ , \tag{5.13}$$

as one readily verifies.

Since $\varphi_{\rm tr}$ is faithful, the left ideal $\mathscr{N}$ that arose in the GNS construction is simply $\{0\}$, and so the Hilbert space $\mathscr{H}$ is simply $\mathscr{A}$ itself equipped with the inner product

$$\langle a, b\rangle_{\mathscr{H}} = \frac{1}{n}{\rm Tr}[a^*b] \ , \tag{5.14}$$

which is a normalized form of the Hilbert-Schmidt inner product.

In this finite dimensional setting, no completion is needed; $\mathscr{H}$ is simply $M_n(\mathbb{C})$ itself, with the inner product (5.14). For all $x \in \mathscr{A}$, let ζ_x denote x regarded as an element of $\mathscr{H}$.

Let π denote the GNS representation of $\mathscr{A}$ determined by $\varphi_{\rm tr}$. Then $\pi(a)\eta_x = ax$, so that if we define the operator L_a on $\mathscr{H}$ by $L_a\zeta_x = \zeta_{ax}$, then $\pi(a) = L_a$, the operation of left multiplication by a. Since $\varphi_{\rm tr}$ is faithful, $\ker(\pi) = \{0\}$, and then by Theorem 2.26, $\pi(\mathscr{A})$ is a C^* algebra, and $\pi : \mathscr{A} \to \pi(\mathscr{A})$ is an isometric $*$-isomorphism.

In this finite dimensional setting $\pi(\mathscr{A})$ is not only a C^* algebra, but also a von Neumann algebra. Let us use $\mathscr{M}$ to denote $\pi(\mathscr{A})$. Now observe that since the center of $\mathscr{A}$ is trivial, and since π is a isomorphism of $\mathscr{A}$ onto $\mathscr{M}$, the center of $\mathscr{M}$ is trivial, so that $\mathscr{M}$ is a factor. By Theorem 5.16, there is a unitary u from $\mathscr{B}(\mathscr{H})$ onto $\mathbb{C}^n \otimes \mathbb{C}^n$ such that $u\mathscr{M}u^* = M_n(\mathbb{C}) \otimes 1$. (We are also using the fact that $M_m(\mathbb{C})$ and $M_n(\mathbb{C})$ are not isomorphic for $m \neq n$.) The commutant $\mathscr{M}'$ of $\mathscr{M}$ then consists of all elements of $\mathscr{B}(\mathscr{H})$ of the form $u(1 \otimes b)u^*$, with $b \in \mathscr{A}$.

We can make this more explicit as follows. For $\zeta \otimes \xi \in \mathbb{C}^n \otimes \mathbb{C}^n$, define $v(\zeta \otimes \xi)$ to be the $n \times n$ matrix $\sqrt{n}[\zeta_i\xi_j^*]$ which we regard as an element of $\mathscr{H}$. It is evident that

$$\|v(\zeta \otimes \xi)\|_{\mathscr{H}} = \|\zeta\|_{\mathbb{C}^n}\|\xi\|_{\mathbb{C}^n} = \|\zeta \otimes \xi\|_{\mathbb{C}^n\otimes\mathbb{C}^n} \ .$$

Extending v by linearity, we obtain an isometry from $\mathbb{C}^n \otimes \mathbb{C}^n$ into $\mathscr{H}$, which is necessarily unitary since the dimensions of the domain and range are equal. We again denote the extension by v. Now observe that for all $a \in \mathscr{A}$, $L_a v(\zeta \otimes \xi) = v(a\zeta \otimes \xi)$, or what is the same thing,

$$v^*L_a v = a \otimes 1_{\mathbb{C}^n} \ .$$

In particular, $v\mathscr{M}v^* = M_n(\mathbb{C}) \otimes 1_{\mathbb{C}^n}$, and thus $u = v^*$ is one choice of the unitary provided by Theorem 5.16. In the same way we see that

$$v^* R_a v = 1_{\mathbb{C}^n} \otimes a^* \ .$$

Since $(M_n(\mathbb{C}) \otimes 1_{\mathbb{C}^n})' = 1_{\mathbb{C}^n} \otimes M_n(\mathbb{C})$, it follows that $v^*\mathscr{M}v = \mathscr{M}'$.

Define a conjugate linear transformation J from $\mathscr{H}$ to itself by

$$J\zeta_a = \zeta_{a^*}$$

for all $a \in \mathscr{A}$. Note that $\|J\zeta_a\|^2_{\mathscr{H}} = \frac{1}{n}\mathrm{Tr}[aa^*] = \frac{1}{n}\mathrm{Tr}[a^*a] = \|\zeta_a\|_{\mathscr{H}}$. That is, because of (5.13), J is an isometry. Moreover, $J^2 = 1$ and so $J = J^* = J^{-1}$.

For each $b \in \mathscr{A}$ define the operator R_b on $\mathscr{H}$ by

$$R_b\zeta_x = \zeta_{xb}$$

for all $c \in \mathscr{A}$. That is, R_b is the operator of right multiplication by b. Now observe that for all $a, x \in \mathscr{A}$, $J(L_a\zeta_x) = J\zeta_{ax} = \zeta_{x^*a^*} = R_{a^*}J\zeta_x$. In short,

$$JL_aJ = R_{a^*} \ , \tag{5.15}$$

and hence $J\mathscr{M}J = \mathscr{M}'$.

To bring out the symmetry between $\mathscr{M}$ and $\mathscr{M}'$, let us introduce $\mathscr{H}^*$ to be the Hilbert space that is the same set as $\mathscr{H}$ with the same law of vector addition, but with the scalar multiplication $(\lambda, \zeta) \mapsto \lambda^*\zeta$ and the inner product $(\zeta, \xi) \mapsto \langle \xi, \zeta\rangle^*_{\mathscr{H}} =: \langle \zeta, \xi\rangle_{\mathscr{H}^*}$. Note that $\mathscr{B}(\mathscr{H}^*) = \mathscr{B}(\mathscr{H})$, and so we may regard each R_{a^*} as an element of $\mathscr{B}(\mathscr{H}^*)$. Then it is easy to check that the map $a \mapsto R_{a^*} =: \pi'(a)$ is a representation of $\mathscr{A}$ on $\mathscr{H}^*$. The map J is then unitary from $\mathscr{H}$ to $\mathscr{H}^*$ (though no longer self adjoint), and now (5.15) can be written as

$$J\pi J^* = \pi' \ , \tag{5.16}$$

and we have that $(\pi(\mathscr{A}))' = \pi'(\mathscr{A})$. The fact that the GNS construction in this simple case yields not one, but two commuting isometric representations of $\mathscr{A}$ will turn out to be very useful later on.

6 Completely positive maps

6.1 Some important isomorphisms

Let $\mathscr{H}$ and $\mathscr{K}$ be two Hilbert spaces with $\mathscr{K}$ separable. Let $\{\eta_j\}$ be an orthonormal basis for $\mathscr{K}$. Then the general element ξ of the Hilbert space $\mathscr{H} \otimes \mathscr{K}$ has the form

$$\xi = \sum_{j=1}^{\dim\mathscr{K}} \zeta_j \otimes \eta_j \qquad \text{and} \qquad \|\xi\|^2_{\mathscr{H}\otimes\mathscr{K}} = \sum_{j=1}^{\dim\mathscr{K}} \|\zeta_j\|^2_{\mathscr{H}} \ . \tag{6.1}$$

For $n \in \mathbb{N}$, let $\mathscr{H}_n$ denote the direct sum of n copies of $\mathscr{H}$,

$$\mathscr{H}_n := \underbrace{\mathscr{H} \oplus \cdots \oplus \mathscr{H}}_{n \text{ times}} \ ,$$

Let $\mathscr{H}_{\aleph}$ denote the direct sum of countably infinitely many copies of $\mathscr{H}$. We write $\{\zeta_j\}_{1\le j\le n}$ to denote the general element of $\mathscr{H}_n$, and we write $\{\zeta_j\}_{j\in\mathbb{N}}$ to denote the general element of $\mathscr{H}_{\aleph}$.

When $\dim(\mathscr{K}) = n < \infty$, we define an isomorphism from $\mathscr{H}_n$ onto $\mathscr{H}\otimes\mathscr{K}$ by choosing an orthonormal basis $\{\eta_1,\dots,\eta_n\}$ of $\mathscr{K}$ and then define the map

$$\{\zeta_j\}_{j\in\mathbb{N}} \mapsto \sum_{j=1}^{n} \zeta_j\otimes\eta_j \ . \tag{6.2}$$

This gives a unitary map from $\mathscr{H}_n$ onto $\mathscr{H}\otimes\mathscr{K}$. When $\mathscr{K}$ is infinite dimensional and $\{\eta_j\}$ is an orthonormal basis of $\mathscr{K}$, the map given in (6.2) is unitary from $\mathscr{H}_{\aleph}$ onto $\mathscr{K}\otimes\mathscr{K}$.

6.1 DEFINITION. Let $\mathscr{H}$ be any Hilbert space and let $\mathscr{K}$ be a separable Hilbert space. Let $\{\eta_j\}$ be an orthonormal basis of $\mathscr{K}$. Define $V_j : \mathscr{K}\to\mathscr{H}\otimes\mathscr{K}$ by

$$V_j\zeta = \zeta\otimes\eta_j \ . \tag{6.3}$$

Note that V_j is an isometry from $\mathscr{K}$ into $\mathscr{H}\otimes\mathscr{K}$, and that

$$\sum_{\text{j}=1}^{\dim(\mathscr{H})} V_j^*V_j = \mathbb{1}_{\mathscr{H}} \ . \tag{6.4}$$

Next, for every $\zeta\in\mathscr{H}$ and $\eta\in\mathscr{K}$ we define the rank-one operator $|\zeta\rangle\langle\eta|$ from $\mathscr{H}$ to $\mathscr{K}$ by

$$|\zeta\rangle\langle\eta|\xi = (\langle\eta,\xi\rangle_{\mathscr{K}})\,\zeta \ .$$

Let $\{\eta_j\}$ be an orthonormal basis for $\mathscr{K}$. Define the sesquilinear map

$$\sum_{j=1}^{\dim\mathscr{K}} \zeta_j\otimes\eta_j \quad\mapsto\quad \sum_{\text{j}=1}^{\dim\mathscr{K}} |\zeta_j\rangle\langle\eta_j| \ .$$

On the right we have the general element of $\mathscr{C}_2(\mathscr{K},\mathscr{H})$, the Hilbert space of Hilbert-Schmidt linear maps from $\mathscr{K}$ to $\mathscr{H}$; that is, the space of linear maps $x:\mathscr{K}\to\mathscr{H}$ such that $\mathrm{Tr}[x^*x]<\infty$. Moreover, this map is easily seen to be an isometry; i.e.,

$$\left\|\sum_{j=1}^{\dim\mathscr{K}} \zeta_j\otimes\eta_j\right\|_{\mathscr{H}\otimes\mathscr{K}}^2 = \sum_{j=1}^{\infty}\|\zeta_j\|_{\mathscr{H}}^2 = \left\|\sum_{\text{j}=1}^{\dim\mathscr{K}} |\zeta_j\rangle\langle\eta_j|\right\|_{\mathscr{C}_2(\mathscr{K},\mathscr{H})}^2 \ .$$

Next, still under the assumption that $\mathscr{K}$ is separable, consider the algebraic tensor product $\mathscr{B}(\mathscr{H})\otimes\mathscr{B}(\mathscr{K})$. The general element of $\mathscr{B}(\mathscr{H})\otimes\mathscr{B}(\mathscr{K})$ is a linear combination of elements of the form $x\otimes y$. We may regard these as operators on $\mathscr{H}\otimes\mathscr{K}$ through

$$(x\otimes y)\zeta\otimes\eta = (x\zeta)\otimes(y\eta) \ . \tag{6.5}$$

This gives us a natural embedding of $\mathscr{B}(\mathscr{H})\otimes\mathscr{B}(\mathscr{K})$ into $\mathscr{B}(\mathscr{H}\otimes\mathscr{K})$

6.2 LEMMA. *For Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$, and any $\zeta_1,\zeta_2\in\mathscr{H}$, and any $\eta_1,\eta_2\in\mathscr{K}$,*

$$|\zeta_1\otimes\eta_1\rangle\langle\zeta_2\otimes\eta_2| = |\zeta_1\rangle\langle\eta_1|\otimes|\zeta_2\rangle\langle\eta_2| \ , \tag{6.6}$$

where the right hand side is regarded as an element of $\mathscr{B}(\mathscr{H}\otimes\mathscr{K})$ through the natural embedding of $\mathscr{B}(\mathscr{H})\otimes\mathscr{B}(\mathscr{K})$ into $\mathscr{B}(\mathscr{H}\otimes\mathscr{K})$

Proof. It suffices to check that for all $\zeta_3 \in \mathscr{H}$ and $\eta_3 \in \mathscr{K}$, both sides have the same action on $\zeta_3 \otimes \eta_3$. By the definitions,

$$\begin{aligned} |\zeta_1 \otimes \eta_1\rangle\langle\zeta_2 \otimes \eta_2|\zeta_3 \otimes \eta_3 &= [\langle\zeta_2 \otimes \eta_2, \zeta_3 \otimes \eta_3\rangle_{\mathscr{H}\otimes\mathscr{K}}]\, \zeta_1 \otimes \eta_1 \\ &= [\langle\zeta_2, \zeta_3\rangle_{\mathscr{H}}\langle\eta_2, \eta_3\rangle_{\mathscr{K}}]\, \zeta_1 \otimes \eta_1 \\ &= [\langle\zeta_2, \zeta_3\rangle_{\mathscr{H}}\langle\zeta_1] \otimes [\langle\eta_2, \eta_3\rangle_{\mathscr{K}}\eta_1] \\ &= (|\zeta_1\rangle\langle\eta_1|\zeta_3) \otimes (|\zeta_2\rangle\langle\eta_2|\eta_3)\ . \end{aligned}$$

□

Now we specialize to a case that will be important in what follows. Let $\mathscr{K}$ be finite dimensional, and identify it with $\mathbb{C}^n$ for $n = \dim(\mathscr{K})$. We may then identify $\mathscr{B}(\mathscr{K})$ with $M_n(\mathbb{C})$. Let $\{\eta_1, \dots, \eta_n\}$ be any orthonormal basis of $\mathbb{C}^n$. Then $\{|\eta_i\rangle\langle\eta_j|\ :\ 1 \le i, j \le n\}$ is a basis for for $M_n(\mathbb{C})$. It is easy to check that

$$|\eta_i\rangle\langle\eta_j||\eta_k\rangle\langle\eta_\ell| = \delta_{j,k}|\eta_i\rangle\langle\eta_\ell|\ . \tag{6.7}$$

For any $a_1 \otimes m_1, a_2 \otimes m_2 \in \mathscr{B}(\mathscr{H}) \times \mathscr{B}(\mathscr{K})$, we define the product $(a_1 \otimes m_1)(a_2 \otimes m_2)$ by

$$(a_1 \otimes m_1)(a_2 \otimes m_2) = a_1a_2 \otimes m_1m_2,$$

and extend this by linearity. Likewise, for any $a \in \mathscr{B}(\mathscr{H})$ and any $m \in M_n(\mathbb{C})$, the involution $(a \otimes m)^* = a^* \otimes m^*$, and also extend this by linearity.

Since $\{\ |\eta_i\rangle\langle\eta_j|\ :\ 1 \le i, j \le n\}$ is a basis of $M_n(\mathbb{C})$, the general element $\widetilde{a}$ of $\mathscr{B}(\mathscr{H}) \otimes M_n(\mathbb{C})$ can be written as $\widetilde{a} = \sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j|$. By (6.7),

$$\left(\sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j|\right)\left(\sum_{k,\ell=1}^{n} b_{k,\ell} \otimes |\eta_k\rangle\langle\eta_\ell|\right) = \sum_{i,\ell=1}^{n}\left(\sum_{j=1}^{n} a_{i,j}b_{j,\ell}\right) \otimes |\eta_i\rangle\langle\eta_\ell|\ . \tag{6.8}$$

We define $M_n(\mathscr{B}(\mathscr{H}))$ to be the set of $n \times n$ matrices with entries in $\mathscr{B}(\mathscr{H})$. Let $[a_{i,j}]$ denote the element of $M_c(\mathscr{B}(\mathscr{H}))$ with i, j entry is $a_{i,j}$. Define a linear transformation from $\mathscr{B}(\mathscr{H}) \otimes M_n(\mathbb{C})$ onto $M_n(\mathscr{B}(\mathscr{H}))$ by

$$\widetilde{a} = \sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j| \mapsto [a_{i,j}]\ . \tag{6.9}$$

The transformation in (6.9) is evidently injective, and hence is a vector space isomorphism. In fact, the inverse map is simply given by

$$[a_{i,j}] \mapsto \sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j|\ . \tag{6.10}$$

By (6.8), vector space isomorphism is also a algebra isomorphism where $M_n(\mathscr{B}(\mathscr{H}))$ is given the natural product, and one easily checks that $[a_{i,j}]^* = [a^*_{j,i}]$, so that it is a $*$-isomorphism.

Another identification will be useful in what follows: There is a natural $*$-isomorphism of $\mathscr{B}(\mathscr{H} \otimes \mathbb{C}^n)$ with $M_n(\mathscr{B}(\mathscr{H}))$. Let $\{\eta_1, \dots, \eta_n\}$ be an orthonormal basis of $\mathbb{C}^n$. For $j = 1, \dots, n$, let V_j be the isometry from $\mathscr{H}$ into into $\mathscr{H}\otimes\mathbb{C}^n$ given by (6.3). Define a linear transformation from $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ to $M_n(\mathscr{B}(H))$ by

$$\widehat{a}\mapsto[V_i^*\widehat{a}V_j]\ . \tag{6.11}$$

for all $\widehat{a}\in\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$.

This too is also a vector space isomorphism. To see this, let $[a_{i,j}]\in M_n(\mathscr{B}(H))$, and $\{\eta_1,\dots,\eta_n\}$ an orthonormal basis of $\mathbb{C}^n$, consider $\sum_{i,j=1}^n[a_{i,j}]\otimes|\eta_i\rangle\langle\eta_j|$ as an element of $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ through the natural embedding of $\mathscr{B}(\mathscr{H})\otimes M_n(\mathbb{C})$ into $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ that is given by (6.5). Then for all $\widehat{a}\in\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$, and all $\zeta\in\mathscr{H}$ and all $\eta\in\mathbb{C}^n$,

$$\begin{aligned}\left(\sum_{i,j=1}^n[V_i^*\widehat{a}V_j]\otimes|\eta_i\rangle\langle\eta_j|\right)\zeta\otimes\eta &= \sum_{i,j=1}^n\langle\eta_j,\eta\rangle[V_i^*\widehat{a}V_j\zeta]\otimes\eta_i\\ &= \sum_{i,j=1}^n\langle\eta_j,\eta\rangle[V_i^*\widehat{a}(\zeta\otimes\eta_j)]\otimes\eta_i\\ &= \sum_{i,j=1}^n[V_i^*\widehat{a}(\zeta\otimes\eta)]\otimes\eta_i=\widehat{a}(\zeta\otimes\eta)\end{aligned}$$

That is, the transformation (6.10), followed by the natural embedding of $\mathscr{B}(\mathscr{H})\otimes M_n(\mathbb{C})$ into $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$, is inverse to the map defined in (6.11). Finally, using (6.4), one readily checks that the map in (6.11) is a $*$-algebra isomorphism as well as a vector space isomorphism. We have proved:

6.3 THEOREM. *For any Hilbert space $\mathscr{H}$ and any $n\in\mathbb{N}$, $\mathscr{B}(H)\otimes M_n(\mathbb{C})$, $M_n(\mathscr{B}(\mathscr{H}))$ and $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$, equipped with their natural $*$-algebra structures, are all $*$-algebra isomorphic. Moreover, for any orthonormal basis $\{\eta_1,\dots,\eta_n\}$ of $\mathbb{C}^n$, and:*

(1) The map in (6.9) is a $*$-isomorphism of $\mathscr{B}(\mathscr{H})\otimes M_n(\mathbb{C})$ onto $M_n(\mathscr{B}(\mathscr{H}))$, and the map in (6.10) is its inverse.

(2) The map in (6.11) is a $*$-isomorphism of $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ onto $M_n(\mathscr{B}(H))$, and the map in (6.10), followed by the natural embedding of $\mathscr{B}(\mathscr{H})\otimes M_n(\mathbb{C})$ into $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$, is its inverse.

Since $\mathscr{H}\otimes\mathbb{C}^n$ is a Hilbert space, $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ is a C^*-algebra. We may use the $*$-algebra isometries provided by Theorem 6.3 to transfer the operator norm in $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$ to $\mathscr{B}(\mathscr{H})\otimes M_n(\mathbb{C})$ and to $M_n(\mathscr{B}(\mathscr{H})$, thus making them C^*-algebras, isomorphic to $\mathscr{B}(\mathscr{H}\otimes\mathbb{C}^n)$.

The isomorphism is basis dependent, but the norm is not since the norm on a C^* algebra is unique.

6.2 The C^* algebras $\mathscr{A}\otimes M_n(\mathbb{C})$.

There is a natural family of C^* algebras associated to every C^* algebra $\mathscr{A}$, namely the C^* algebras $\mathscr{A}\otimes M_n(\mathbb{C})$ for each $n\in\mathbb{N}$, which we now define. As a vector space, $\mathscr{A}\otimes M_n(\mathbb{C})$ is the algebraic tensor product of the vector spaces $\mathscr{A}$ and $M_n(\mathbb{C})$. Let $\{\eta_1,\dots,\eta_n\}$ be any orthonormal basis for

$\mathbb{C}^n$. Since the $\{|\eta_i\rangle\langle\eta_j| \ : \ 1 \le i,j \le n\}$ is a basis for $M_n(\mathbb{C})$, the general element $\widetilde{a}$ of $\mathscr{A} \otimes M_n(\mathbb{C})$ has the form

$$\widetilde{a} = \sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j| \ , \tag{6.12}$$

where for each i,j, $a_{i,j} \in \mathscr{A}$.

We give it the natural algebraic structure by defining $(a_1 \otimes m_1)(a_2 \otimes m_2) = (a_1a_2 \otimes m_1m_2)$, and then extending this by linearity. If $\widetilde{a} = \sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j|$ and $\widetilde{b} = \sum_{i,j=1}^{n} b_{i,j} \otimes |\eta_i\rangle\langle\eta_j|$ as in (6.12), then one checks as above that

$$\widetilde{a}\widetilde{b} = \left(\sum_{j=1}^{n} a_{i,j}b_{j,\ell}\right) \otimes |\eta_i\rangle\langle\eta_\ell| \ . \tag{6.13}$$

Defining $(a \otimes m)^* = (a^* \otimes b^*)$, makes $\mathscr{A} \otimes M_n(\mathbb{C})$ a $*$-algebra.

Let $M_n(\mathscr{A})$ denote the set of $n \times n$ matrices with entries in $\mathscr{A}$. We write $[a_{i,j}]$ to denote the element of $M_n(\mathscr{A})$ whose i,j entry is $a_{i,j}$. $M_n(\mathscr{A})$ is a $*$-algebra with the obvious operations. Then by (6.13), the map

$$\sum_{i,j=1}^{n} a_{i,j} \otimes e^{(i,j)} = \widetilde{a} \ \mapsto [a_{i,j}]$$

is a $*$ isomorphism of $\mathscr{A} \otimes M_n(\mathbb{C})$ *onto* $M_n(\mathscr{A})$, and we use this isomorphism to identify $\mathscr{A} \otimes M_n(\mathbb{C})$ with $M_n(\mathscr{A})$.

We now claim that there is a norm on $\mathscr{A} \otimes M_n(\mathbb{C})$ that makes it a C^* algebra. To see this easily, suppose first that $\mathscr{A}$ is a C^* sub algebra of $\mathscr{H}$ for some Hilbert space $\mathscr{H}$. Then $M_n(\mathscr{A})$ is evidently a closed subspace of $M_n(\mathscr{B}(\mathscr{H}))$ on which we have a natural C^*-algebra norm through the identification of $M_n(\mathscr{B}(\mathscr{H}))$ with $\mathscr{B}(\mathscr{H}_n)$ as explained in the previously. Thus, $M_n(\mathscr{A})$ is a C^*-subalgebra of $M_n(\mathscr{B}(\mathscr{H}))$.

In general, we can always use the Gelfand-Naimark Theorem to identify $\mathscr{A}$ with a C^*-algebra of operators, and thus the above discussion applies to the general case.

6.4 DEFINITION. Let $\mathscr{A}$ and $\mathscr{B}$ be C^*-algebras, and let Φ be a bounded linear map from $\mathscr{A}$ to $\mathscr{B}$. Then for all $n \in \mathbb{N}$, define $\Phi_n := \Phi \otimes 1_{\mathbb{C}^n}$, so that

$$\Phi_n(a \otimes m) = \Phi(a) \otimes m$$

for all $a \in \mathscr{A}$ and all $m \in M_n(\mathbb{C})$. In particular, for any orthonormal basis $\{\eta_1,\dots,\eta_n\}$ of $\mathbb{C}^n$,

$$\Phi_n\left(\sum_{i,j=1}^{n} a_{i,j} \otimes |\eta_i\rangle\langle\eta_j|\right) = \sum_{i,j=1}^{n} \Phi(a_{i,j}) \otimes |\eta_i\rangle\langle\eta_j| \ .$$

so that for $\widetilde{a}$ given by (6.12), $\Phi_n(\widetilde{a})$, considered as an element of $M_n(\mathscr{A})$, is given by

$$[\Phi_n(\widetilde{a})_{i,j}] = [\Phi(a_{i,j})] \ .$$

That is, the action of Φ_n on $[a_{i,j}] \in M_n(\mathscr{A})$ is given by the action of Φ on each entry of $[a_{i,j}]$.

6.3 Positive and completely positive maps

6.5 DEFINITION. Let $\mathscr{A}$ and $\mathscr{B}$ be C^* algebras. A linear map $\Phi : \mathscr{A} \to \mathscr{B}$ is *positive* in case $\Phi(a^*a) \geq 0$ for all $a \in \mathscr{A}$. If $\mathscr{A}$ and $\mathscr{B}$ have identities $1_{\mathscr{A}}$ and $1_{\mathscr{B}}$ respectively, then Φ is *unital* in case $\Phi(1_{\mathscr{A}}) = 1_{\mathscr{B}}$.

In any C^* algebra $\mathscr{A}$, we can write the general element a as $a = (a_1 - a_2) + i(b_1 - b_2)$ where $a_1, a_2, b_1, b_2 \in \mathscr{A}_+$. Then

$$\Phi(a^*) = \Phi((a_1 - a_2) - i(b_1 - b_2)) = (\Phi(a_1) - \Phi(a_2)) - i(\Phi(b_1) - \Phi(b_2)) = \Phi(a)^* \ .$$

That is, Φ automatically respects the involutions on $\mathscr{A}$ and $\mathscr{B}$.

If Φ is any $*$-homomorphism of $\mathscr{A}$ into $\mathscr{B}$, then for all $a \in \mathscr{A}$, $\Phi(a^*a) = \Phi(a)^*\Phi(a) \geq 0$, and since every element of $\mathscr{A}_+$ if is the form a^*a, it follows that every $*$-homomorphism is positive.

Here is another important example: Let $\mathscr{A} = \mathscr{B} = M_n(\mathbb{C})$, and for $a \in \mathscr{A}$, let $\Phi(a) = a^T$, the transpose of a. Then evidently $\Phi : a \to a^T$ is positive. Since for all $a_1, a_2 \in \mathscr{A}$, $(a_1a_2)^T = a_2^T a_2^T$, Φ is not a $*$-homomorphism.

6.6 DEFINITION (n-positive and completely positive)**.** Let $\mathscr{A}$ and $\mathscr{B}$ be C^* algebras. A linear map $\Phi : \mathscr{A} \to \mathscr{B}$ is *n-positive* in case $\Phi_n : \mathscr{A} \otimes \mathbb{C}^n \to \mathscr{B} \otimes \mathbb{C}^n$ is positive. A linear map $\Phi : \mathscr{A} \to \mathscr{B}$ is *completely positive* in case Φ_n is positive for all $n \in \mathbb{N}$.

6.7 THEOREM. *Let $\mathscr{A}$ be a C^*-subalgebra of $\mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$. Let $\mathscr{K}$ be a second Hilbert space, and let $\Phi : \mathscr{A} \to \mathscr{B}(\mathscr{K})$ be given by*

$$\Phi(a) = \sum_{j=1}^{m} W_j^* a W_j \tag{6.14}$$

where for each $j = 1, \ldots, m$, W_j is a bounded linear transformation from $\mathscr{K}$ to $\mathscr{H}$. Then Φ is completely positive.

Proof. Since a sum of completely positive maps is evidently completely positive, it suffices to consider the case $\Phi(a) = W^*aW$ for a bounded linear transformation W from $\mathscr{K}$ to $\mathscr{H}$. But for any $n \in \mathbb{N}$, if $\widetilde{a} = [a_{i,j}]$ is any element of $M_n(\mathscr{A})$,

$$\Phi_n(\widetilde{a}) = [W^*a_{i,j}W] = \left(\sum_{i=1}^{n} W^* \otimes |\eta_i\rangle\langle\eta_i|\right)\left(\sum_{k,\ell} a_{i,j} \otimes |\eta_k\rangle\langle\eta_\ell|\right)\left(\sum_{j=1}^{n} W \otimes |\eta_j\rangle\langle\eta_j|\right)^*$$

which is clearly positive. □

We will see later that this is essentially the only example: All completely positive maps have such a form, at least when the Hilbert spaces are finite dimensional.

6.8 LEMMA. *Let $\mathscr{A}$ be a C^* algebra with identity 1. Then for all $a, b \in \mathscr{A}$, $a^*a \leq b$ if and only if $\begin{bmatrix} 1 & a \\ a^* & b \end{bmatrix}$ is positive in $M_2(\mathscr{A})$.*

Proof. By the Gelfand-Naimark Theorem, we may suppose that $\mathscr{A}$ is a C^* subalgebra of $\mathscr{B}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$.

Suppose $b \geq a^*a$, and define $c = b - a^*a$. Then

$$\begin{bmatrix} 1 & a \\ a^* & b \end{bmatrix} = \begin{bmatrix} 1 & a \\ a^* & a^*a \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix} = \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}^* \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix} ,$$

which displays the left hand side as a sum of positive operators.

For the converse, supose that $\begin{bmatrix} 1 & a \\ a^* & b \end{bmatrix}$ is positive. Then for all $\xi, \eta \in \mathscr{H}$,

$$\left\langle \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \begin{bmatrix} 1 & a \\ a^* & b \end{bmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle_{\mathscr{H}\oplus\mathscr{H}} = \|\xi\|^2_{\mathscr{H}} + 2\Re\langle \xi, a\eta\rangle_{\mathscr{H}} + \langle \eta, b\eta\rangle_{\mathscr{H}} \geq 0 \ .$$

For $\xi = -a\eta$, this becomes $\langle \eta, [b - a^*a]\eta\rangle_{\mathscr{H}} \geq 0$, and this shows that $b \geq a^*a$. □

6.9 THEOREM (Kadison's Inequality)**.** *Let $\mathscr{A}$ and $\mathscr{B}$ be C^* algebras with identities $1_{\mathscr{A}}$ and $1_{\mathscr{B}}$ respectively. Let Φ be a unital 2-positive map from $\mathscr{A}$ to $\mathscr{B}$. Then for all $a \in \mathscr{A}$,*

$$\Phi(a)^*\Phi(a) \leq \Phi(a^*a) \ . \tag{6.15}$$

Proof. Since $\begin{bmatrix} 1 & a \\ a^* & a^*a \end{bmatrix} \geq 0$ in $M_2(\mathscr{A})$,

$$\Phi_2\left(\begin{bmatrix} 1 & a \\ a^* & a^*a \end{bmatrix} \right) = \begin{bmatrix} 1 & \Phi(a) \\ \Phi(a)^* & \Phi(a^*a) \end{bmatrix} \geq 0 \ ,$$

and by Lemma 6.8, this implies (6.15). □

Associated to every completely positive map Φ from a C^* algebra to $\mathscr{B}(\mathscr{H})$ for a Hilbert space $\mathscr{H}$ is a non-negative sesquilinear form on $\mathscr{A} \otimes \mathscr{H}$ that we now describe.

6.10 DEFINITION (The Stinespring inner product)**.** Let $\mathscr{A}$ be a C^*-algebra and let $\mathscr{H}$ be a Hilbert space. Let $\Phi : \mathscr{A} \to \mathscr{B}(\mathscr{H})$ be completely positive. Consider two arbitrary elements of $\mathscr{A} \otimes \mathscr{H}$ which we may take to be of the form

$$\sum_{j=1}^n a_j \otimes \eta_j \quad \text{and} \quad \sum_{j=1}^n b_j \otimes \xi_j$$

for some common value of $n \in \mathbb{N}$ by allowing some terms to be zero.

Then the sesquilinear form $\langle \cdot, \cdot \rangle_\Phi$ on $\mathscr{A} \otimes \mathscr{H}$ given by

$$\left\langle \sum_{j=1}^n a_j \otimes \eta_j, \sum_{j=1}^n b_j \otimes \xi_j \right\rangle_\Phi = \sum_{i,j=1}^n \langle \eta_i, \Phi(a_i^* b_j)\xi_j\rangle_{\mathscr{H}} \tag{6.16}$$

is the *Stinespring inner product on $\mathscr{A} \otimes \mathscr{H}$*.

The Stinespring inner product is non-negative (as the name suggests): Let $\sum_{j=1}^n a_j \otimes \eta_j \in \mathscr{A} \otimes \mathscr{H}$, and define $\widetilde{a} = [a_i^* a_j] \in M_n(\mathscr{A})$. Also define $\widetilde{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \in \mathscr{H}_n$, the direct sum of n copies of $\mathscr{H}$.

Note that

$$[a_i^* a_j] = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}^* \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \geq 0$$

Therefore, $\Phi_n([a_i^* a_j]) = [\Phi(a_i^* a_j)]$ is positive in $M_n(\mathscr{B}(\mathscr{H}))$, and

$$\sum_{i,j=1}^n \langle \eta_i, \Phi(a_i^* a_j)\eta_j \rangle_{\mathscr{H}} = \langle \widetilde{\eta}, \Phi_n([a_i^* a_j])\widetilde{\eta} \rangle_{\mathscr{H}_n} \geq 0 \ .$$

6.4 The partial trace

Recall that every linear functional ϕ on $M_n(\mathbb{C})$ is of the form $\phi(x) = \mathrm{Tr}[xz]$ for some uniquely determined $z \in M_n(\mathbb{C})$. This is because if we equip $M_n(\mathbb{C})$ with the Hilbert-Schmidt inner product $\langle x, y \rangle = \mathrm{Tr}[x^* y]$, it is a (finite dimensional) Hilbert space, and then we may apply the Riesz Lemma.

Now suppose that $a \in \mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^p)$. Define a linear functional ϕ_a on $M_n(\mathbb{C})$ by

$$x \mapsto \mathrm{Tr}[(x \otimes \mathbb{1})a] =: \phi_a(x)$$

Since this is a linear functional on $M_n(\mathbb{C})$, it has the form $\phi_a(x) = \mathrm{Tr}[bx]$ for some uniquely determined $b \in M_n(\mathbb{C})$. This brings us to the following definition:

6.11 DEFINITION. Let $a \in \mathscr{B}(\mathbb{C}^m \otimes \mathbb{C}^n)$. The partial traces of a on $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively are the operators $\mathrm{Tr}_2[a] \in M_m(\mathbb{C})$ and $\mathrm{Tr}_1[a] \in M_n(\mathbb{C})$ such that for all $x \in M_m(\mathbb{C})$ and all $y \in M_n(\mathbb{C})$,

$$\mathrm{Tr}[x\mathrm{Tr}_2[a]] = \mathrm{Tr}[(x \otimes \mathbb{1})a] \quad \text{and} \quad \mathrm{Tr}[y\mathrm{Tr}_1[a]] = \mathrm{Tr}[(\mathbb{1} \otimes y)a] \ . \tag{6.17}$$

By using an orthonormal basis of $\mathbb{C}^m \otimes \mathbb{C}^m$ of the form $\{\eta_i \otimes \zeta_j \ : \ 1 \leq i \leq n \ , 1 \leq j \leq m\}$, it is evident that for all $y \in M_n(\mathbb{C})$ and all $z \in M_m(\mathbb{C})$,

$$\mathrm{Tr}[y \otimes z] = \mathrm{Tr}[y]\mathrm{Tr}[z] \ . \tag{6.18}$$

Therefore, when $a = y \otimes z$ with $y \in M_m(\mathbb{C})$ and $z \in M_n(\mathbb{C})$,

$$\mathrm{Tr}[(x \otimes \mathbb{1})a] = \mathrm{Tr}[xy \otimes z] = \mathrm{Tr}[xy]\mathrm{Tr}[z]$$

where on the right the traces are taken in $M_m(\mathbb{C})$ and $M_n(\mathbb{C})$ respectively. Thus, $\mathrm{Tr}_2[y \otimes z] = \mathrm{Tr}[z]y$. Likewise, $\mathrm{Tr}_1[y \otimes z] = \mathrm{Tr}[y]z$.

By Theorem 6.3, the general element a of $\mathscr{B}(\mathbb{C}^m \otimes \mathbb{C}^n)$ has the form

$$a = \sum_{i,j=1}^n V_i^* a V_j \otimes |\eta_i\rangle\langle\eta_j|$$

where $\{\eta_1, \dots, \eta_n\}$ is an orthonormal basis of $\mathbb{C}^n$, the the isometries V_j, $j = 1, \dots, n$, are defined in (6.3). Then by linearity of the trace and the obvious identity $\mathrm{Tr}[\eta_i\rangle\langle\eta_j|] = \delta_{i,j}$,

$$\mathrm{Tr}_2[a] = \sum_{j=1}^{n} V_i^* a V_j \ . \tag{6.19}$$

By symmetry, the same reasoning applies to $a \mapsto \mathrm{Tr}_1[a]$. Let $\{\zeta_1, \dots, \zeta_m\}$ be an orthonormal basis of $\mathbb{C}^m$. For $j = 1, \dots, m$, define $W_j : \mathbb{C}^n \to \mathbb{C}^m \otimes \mathbb{C}^n$ by $W_j \eta = \zeta_j \otimes \eta$, then we have $\mathrm{Tr}_1[a] = \sum_{j=1}^{n} W_j^* a W_j$. By Theorem 6.7, the maps $a \mapsto \mathrm{Tr}_1[a]$ and $a \mapsto \mathrm{Tr}_2[a]$ are both completely positive.

6.12 EXAMPLE (The partial transpose)**.** As our terminology suggests, not every positive map is completely positive. Here is an important example. Let $\mathscr{A} = M_2(\mathbb{C})$, and let Ψ be the transpose map $\Psi(a) = a^T$. Then Ψ is positive, but Ψ_2 is not. Indeed, identifying $\mathscr{A} \otimes M_2(\mathbb{C}) = M_2(\mathbb{C}) \otimes M_2(\mathbb{C})$ with $M_2(M_2())$ ad above, we have that for any $a, b, c, d \in M_2(\mathbb{C})$,

$$\Psi_2\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} \Psi(a) & \Psi(b) \\ \Psi(c) & \Psi(d) \end{bmatrix} = \begin{bmatrix} a^T & b^T \\ c^T & d^T \end{bmatrix}$$

need not be positive. To see this, consider the choice $a = e^{(1,1)}$, $b = e^{(1,2)}$, $c = e^{(2,1)}$and $d = e^{(2,2)}$. Then $\begin{bmatrix} e^{(1,1)} & e^{(1,2)} \\ e^{(2,1)} & e^{(2,2)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ is positive, but

$$\Psi_2\left(\begin{bmatrix} e^{(1,1)} & e^{(1,2)} \\ e^{(2,1)} & e^{(2,2)} \end{bmatrix}\right) = \begin{bmatrix} e^{(1,1)} & e^{(2,1)} \\ e^{(1,2)} & e^{(2,2)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is not positive.

The fact that $\begin{bmatrix} e^{(1,1)} & e^{(1,2)} \\ e^{(2,1)} & e^{(2,2)} \end{bmatrix} = \sum_{i,j=1}^{2} e^{(i,j)} \otimes e^{(1,j)}$ revealed the failure of Ψ to be 2-positive is no accident, as we explain in the the next section.

6.5 Choi's Theorem

Choi's Theorem gives a complete description of completely positive maps in finite dimensions. Let $\mathscr{L}(M_n(\mathbb{C}), M_p(\mathbb{C}))$ denote the space of linear transformations from $M_n(\mathbb{C})$ to $M_p(\mathbb{C})$. Let $\Phi \in \mathscr{L}(M_n(\mathbb{C}), M_p(\mathbb{C}))$ for some $n, p \in \mathbb{N}$, and then for any $m \in \mathbb{N}$, let $\Phi_m : M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \to M_p(\mathbb{C}) \otimes M_m(\mathbb{C})$ be defined as above.

We may identify $M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$ with $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$ as in Theorem 6.3, and thus may regard Φ_m as a linear map from $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$ to $M_p(\mathbb{C}) \otimes M_m(\mathbb{C})$. To show that Φ_m is positive, one has to show that for all positive $a \in \mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$, $\Phi_m(a)$ is positive in $M_p(\mathbb{C}) \otimes M_n(\mathbb{C})$. By the Spectral Theorem, every positive element a of $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$ may be decomposed as a sum of rank-one projections, and thus Φ_m is positive if and only if it positive on every rank-one projection in $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$. Let $|\xi\rangle\langle\xi|$ be such a rank-one projection.

The general element of ξ of $\mathbb{C}^n \otimes \mathbb{C}^m$ has the form

$$\xi = \sum_{j=1}^{n} \eta_j \otimes \zeta_j \tag{6.20}$$

for some set $\{\zeta_1, \ldots, \zeta_n\}$ of n vectors in $\mathbb{C}^m$.

Now define a linear transformation a from $\mathbb{C}^m$ to $\mathbb{C}^m$ by $a\eta_j = \sqrt{n}\zeta_j$ for $j = 1, \ldots, n$, and define the vector $\omega \in \mathbb{C}^n \otimes \mathbb{C}^n$ by

$$\omega = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \eta_j \otimes \eta_j \ . \tag{6.21}$$

Note that ω is a unit vector in $\mathbb{C}^n \otimes \mathbb{C}^n$ and that

$$\xi = (\mathbb{1}_n \otimes a)\omega \ . \tag{6.22}$$

We now claim that

$$\Phi_m(|\xi\rangle\langle\xi|) = (\mathbb{1}_n \otimes a)\Phi_n(|\omega\rangle\langle\omega|)(\mathbb{1}_n \otimes a)^* \ . \tag{6.23}$$

This is true since Φ_m acts on the first factor of $M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$, while multiplication by $(\mathbb{1}_n \otimes a)$ on the left and by $(\mathbb{1}_n \otimes a)^*$ on the right act on the second factor of $M_n(\mathbb{C}) \otimes M_m(\mathbb{C})$. To write it out explicitly, note that by (6.21), $|\xi\rangle\langle\xi| = |(\mathbb{1}_n \otimes a)\omega\rangle\langle(\mathbb{1}_n \otimes a)\omega| = (\mathbb{1}_n \otimes a)|\omega\rangle\langle\omega||(\mathbb{1}_n \otimes a)^*$. By Lemma 6.2, (6.20) and (6.21), it follows that

$$|\xi\rangle\langle\xi| = \sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_j| \otimes |\zeta_i\rangle\langle\zeta_j| \quad \text{and} \quad |\omega\rangle\langle\omega| = \frac{1}{n}\sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_j| \otimes |\eta_i\rangle\langle\eta_j| \ .$$

Then

$$\begin{aligned} \Phi_m(|\xi\rangle\langle\xi|) &= \sum_{i,j=1}^{n} \Phi(|\eta_i\rangle\langle\eta_j|) \otimes |\zeta_i\rangle\langle\zeta_j| \\ &= \frac{1}{n}\sum_{i,j=1}^{n} \Phi(|\eta_i\rangle\langle\eta_j|) \otimes |a\eta_i\rangle\langle a\eta_j| \\ &= (\mathbb{1}_n \otimes a)\left(\frac{1}{n}\sum_{i,j=1}^{n} \Phi(|\eta_i\rangle\langle\eta_j|) \otimes |\eta_i\rangle\langle\eta_j|\right)(\mathbb{1}_n \otimes a) \\ &= (\mathbb{1}_n \otimes a)\Phi_n(|\omega\rangle\langle\omega|)(\mathbb{1}_n \otimes a)^* \ , \end{aligned}$$

which proves (6.23). By (6.23), whenever $\Phi_n(|\omega\rangle\langle\omega|) \geq 0$, then for all $\xi \in \mathbb{C}^n \otimes \mathbb{C}^m$, $\Phi_m(|\xi\rangle\langle\xi|) \geq 0$.

Note that ω is a unit vector in $\mathbb{C}^n \otimes \mathbb{C}^n$, and so $\displaystyle\frac{1}{n}\sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_j| \otimes |\eta_i\rangle\langle\eta_j| = |\omega\rangle\langle\omega|$ is a rank-one projector, and in particular, is positive. .

We can use any orthonormal basis for form ω and the projector onto its span, but at this level of generality, we might as well use the standard basis:

6.13 DEFINITION (Choi projector and Choi matrix)**.** Let $\{\eta_1, \ldots, \eta_n\}$ be the standard basis of $\mathbb{C}^n$ so that $|\eta_i\rangle\langle\eta_j|$ is the i,jth *matrix unit*; i.e., the element of $M_n(\mathbb{C})$ with 1 in the i,j place and

0 elsewhere. The Choi projector in $M_n(\mathbb{C}^n) \otimes M_n(\mathbb{C}^n)$ is the element

$$p_C = \frac{1}{n} \sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_j| \otimes |\eta_i\rangle\langle\eta_j| \ . \tag{6.24}$$

As we have seen above, it is the orthogonal projection $|\omega\rangle\langle\omega|$ onto the span of the unit vector ω given in (6.21).

Let Φ be any linear transformation of $M_n(\mathbb{C})$ to $M_p(\mathbb{C})$. The *Choi matrix* of Φ is the the element of $M_n(M_p(\mathbb{C}))$ given by

$$\Phi_n(p_C) = \frac{1}{n} \sum_{i,j=1}^{n} \Phi(|\eta_i\rangle\langle\eta_j|) \otimes |\eta_i\rangle\langle\eta_j| \ . \tag{6.25}$$

We can now restate the conclusion that whenever $\Phi_n(|\omega\rangle\langle\omega|) \geq 0$, then for all $\xi \in \mathbb{C}^n \otimes \mathbb{C}^m$, $\Phi_m(|\xi\rangle\langle\xi|) \geq 0$:

6.14 THEOREM (Choi's Theorem)**.** *Let Φ be a linear transformation from $M_n(\mathbb{C})$ to $M_p(\mathbb{C})$. Then Φ is completely positive if and only if $\Phi_n(p_C)$ is positive where p_C is the Choi projector given by (6.24).*

Choi's Theorem says that $\Phi \in \mathscr{L}(M_n(\mathbb{C}), M_p(\mathbb{C}))$ is completely positive if and only if Φ_n is positive, but it says much more than that: To check this, we need only look at Φ_n applied to the single positive element p_C.

This proof turns on two essential points: First the extreme points of the unit ball of $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$ are the rank one projectors in $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^m)$, and thus it suffices to consider $\Phi_m(|\xi\rangle\langle\xi|)$ for such a projector. Next, *every* vector ξ in $\mathbb{C}^n \otimes \mathbb{C}^m$ can be written in the form $\xi = (\mathbb{1}_n \otimes a)\omega$ for some $a \in \mathscr{L}(\mathbb{C}^n, \mathbb{C}^m)$. That is, acting *only on the right factor*, one can "steer" ω into a general position in $\mathbb{C}^n \otimes \mathbb{C}^m$. This fact has important consequences in quantum mechanics to which we shall return, but it for $m = n$ it is something familiar to us: We know that the representation π of $M_n(\mathbb{C})$ on $\mathbb{C}^n \otimes \mathbb{C}^n$ given by $\pi(a) = \mathbb{1}_n \otimes a$ is irreducible, and ω is a cyclic vector for it. We now give a second proof of Choi's Theorem that provides additional information.

Second proof of Theorem 6.14. Suppose that a linear map Φ from $M_n(\mathbb{C})$ to $M_p(\mathbb{C})$ is such that $\Phi(p_C)$ is positive. We then apply the Spectral Theorem to write

$$\Phi_n(p_C) = \sum_{j=1}^{np} \lambda_j |\zeta_j\rangle\langle\zeta_j|$$

where the λ_j are the (non-negative) eigenvalues of $\Phi(e_{(c)})$, and the ζ_j are the eigenvectors.

Each ζ_j has an expansion $\zeta_j = \sum_{k=1}^{n} \zeta_{j,k} \otimes \eta_j$ for vectors $\zeta_{j,k} \in \mathbb{C}^p$ and where $\{\eta_k\}$ is the standard basis of $\mathbb{C}^n$. Then by Lemma 6.2,

$$|\zeta_j\rangle\langle\zeta_j| = \sum_{k,\ell=1}^{n} |\zeta_{j,k}\rangle\langle\zeta_{j,\ell}| \otimes |\eta_k\rangle\langle\eta_\ell| \ .$$

Therefore, $\Phi_n(p_C) = \sum_{k,\ell=1}^{n}\sum_{j=1}^{np} \lambda_j |\zeta_{j,k}\rangle\langle\zeta_{j,\ell}| \otimes |\eta_k\rangle\langle\eta_\ell|$. It then follows from (6.25) that

$$\Phi(|\eta_k\rangle\langle\eta_\ell|) = n\sum_{j=1}^{np} \lambda_j |\zeta_{j,k}\rangle\langle\zeta_{j,\ell}| \tag{6.26}$$

Now define V_j to be the $p \times n$ matrix whose k column is $\zeta_{j,k}$, so that $V_j\eta_k = \zeta_{j,k}$. Then $|\zeta_{j,k}\rangle\langle\zeta_{j,\ell}| = V_j|\eta_k\rangle\langle\eta_\ell|V_j^*$. Therefore, if we define $W_j = \sqrt{n\lambda_j}V_j^*$ for each j, we can rewrite (6.26) as $\Phi(|\eta_k\rangle\langle\eta_\ell|) = \sum_{j=1}^{np} W_j^*|\eta_k\rangle\langle\eta_\ell|W_j$. But then by linearity, for all $a \in M_n(\mathbb{C})$,

$$\Phi(a) = \sum_{j=1}^{np} W_j^* a W_j \ . \tag{6.27}$$

such maps are completely positive by Theorem 6.7. □

The second proof has also yielded another result of Choi:

6.15 THEOREM. *Let $\Phi \in \mathscr{L}(M_n(\mathbb{C}), M_p(\mathbb{C}))$ be completely positive. Then there is a set $\{W_1, \dots, W_{np}\}$ of $n \times p$ matrices such that for all $a \in M_n(\mathbb{C})$, $\Phi(a)$ is given by (6.27).*

6.16 DEFINITION (Krauss operators)**.** If $\Phi \in \mathscr{L}(M_n(\mathbb{C}), M_p(\mathbb{C}))$ is completely positive, then a set of of $n \times p$ matrices $\{W_1, \dots, W_m\}$ such that

$$\Phi(a) = \sum_{j=1}^{m} W_j^* a W_j \tag{6.28}$$

are called a set of *Krauss operator* for Φ, and (6.28) is a *Krauss representation* of Φ.

We shall discuss minimality of Krauss representations in the next section. Note that if the completely positive map Φ given by (6.28) is unital; i.e., $\Phi(\mathbb{1}_m) = \mathbb{1}_p$ if and only if

$$\sum_{j=1}^{m} W_j^* W_j = \mathbb{1}_p \tag{6.29}$$

6.17 EXAMPLE (The partial transpose in $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^n)$)**.** Let Ψ be the transpose map on $M_n(\mathbb{C})$, and let Ψ_n be its extension to $M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$ which, upon identifying $M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$ with $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^n)$, we refer to as the *partial transpose* on $\mathscr{B}(\mathbb{C}^n \otimes \mathbb{C}^n)$. We compute

$$\Psi_n(p_C) = \frac{1}{n}\sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_j| \otimes |\eta_j\rangle\langle\eta_i| \ . \tag{6.30}$$

We now show that $\Psi_n(p_C)$, while self-adjoint, is not positive. Here is an easy way to see this: Using (6.7) and (6.30), we compute that

$$(\Psi_n(p_C))^2 = \frac{1}{n^2}\sum_{i,j=1}^{n} |\eta_i\rangle\langle\eta_i| \otimes |\eta_j\rangle\langle\eta_j| = \frac{1}{n^2}\mathbb{1}_n \ .$$

Hence all of the eigenvalues of $\Psi_n(p_C)$ are all $\pm 1/n$. By (6.18) and (6.30),

$$\mathrm{Tr}[\Psi_n(p_C)] = \frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Tr}[|\eta_i\rangle\langle\eta_j|]^2 = 1 \ .$$

Hence $1/n$ is an eigenvalue of multiplicity $n(n+1)/2$ and $-1/n$ is an eigenvalue of multiplicity $n(n-1)/2$.

6.6 Stinespring's Theorem

6.18 THEOREM. *Let $\mathscr{A}$ be a C^* algebra with identity* 1*, and let $\Phi : \mathscr{A} \to \mathscr{B}(\mathscr{H})$ be a completely positive map. Then there exists a Hilbert space $\mathscr{K}$ and a unital $*$-homomorphism $\pi : \mathscr{A} \to \mathscr{B}(\mathscr{K})$ and a bounded operator $V : \mathscr{H} \to cK$ such that $\|V\|^2 = \|\Phi(1)\|$ and for all $a \in \mathscr{A}$,*

$$\Phi(a) = V^*\pi(a)V \quad \text{and} \quad \|\pi(a)\| \leq \|a\| \ . \tag{6.31}$$

When Φ is unital, V is an isometry, and

$$\|\pi(a)\|^2 \geq \Phi(a^*a) \ . \tag{6.32}$$

Proof. Equip the vector space $\mathscr{A} \otimes \mathscr{H}$ with the Stinespring inner product $\langle\cdot,\cdot\rangle_\Phi$. Define

$$\mathscr{N} = \{\xi \in \mathscr{A} \otimes \mathscr{H} \ : \ \langle\xi,\xi\rangle_\Phi = 0\} \ .$$

Then $\mathscr{N}$ is a subspace of $\mathscr{A} \otimes \mathscr{H}$, and for all $\xi, \xi' \in \mathscr{A} \otimes \mathscr{H}$ we write $\xi \sim \xi'$ in case $\xi - \xi' \in \mathscr{N}$. We define an inner product, again denoted $\langle\cdot,\cdot\rangle_\Phi$ on the quotient space $\mathscr{A} \otimes \mathscr{H}/\mathscr{N}$ by

$$\langle\{\xi\},\{\zeta\}\rangle_\Phi = \langle\xi,\zeta\rangle_\Phi$$

for all $\xi, \zeta \in \mathscr{A} \otimes \mathscr{H}/\mathscr{N}$ and note that by the Cauchy-Schwarz inequality, the inner product is independent of the choice of representatives. Let $\mathscr{K}$ be the Hilbert space completion of $\mathscr{A} \otimes \mathscr{H}/\mathscr{N}$ in the metric associated to this inner product.

The general element ξ of $\mathscr{A} \otimes \mathscr{H}$ has the form $\xi = \displaystyle\sum_{j=1}^{n} b_j \otimes \zeta_j$ where $\{\zeta_1, \dots, \zeta_n\} \subset \mathscr{H}$. Let $a \in \mathscr{A}$, and define

$$\pi(a)\xi = \sum_{j=1}^{n} ab_j \otimes \zeta_j \ .$$

Then $\|\pi(a)\xi\|^2_{\mathscr{K}} = \langle\zeta_i\Phi(b_ia^*ab_j)\zeta_j\rangle_{\mathscr{H}}$. Since $\{\zeta_1, \dots, \zeta_n\}$ is orthonormal, $[b_ia^*ab_j] \in M_n(\mathscr{A})$ is given by

$$[b_i^*a^*ab_j] = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}^* \begin{bmatrix} a^*a & 0 & \cdots & b_n \\ 0 & a^*a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a^*a \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \geq 0 \tag{6.33}$$

Since

$$\|a\|^2 \begin{bmatrix} 1_{\mathscr{A}} & 0 & \cdots & b_n \\ 0 & 1_{\mathscr{A}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1_{\mathscr{A}} \end{bmatrix} \geq \begin{bmatrix} a^*a & 0 & \cdots & b_n \\ 0 & a^*a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a^*a \end{bmatrix} ,$$

the factorization (6.33) implies that $\|a\|^2[b_i^*b_j] \geq [b_i^*a^*ab_j]$, and then since Φ is completely positive,

$$\|a\|^2\Phi_n([b_i^*b_j]) \geq \Phi_n([b_i^*a^*ab_j]) \ .$$

Therefore

$$\|a\|^2\langle \xi, \xi\rangle_\Phi = \sum_{i,j=1}^n \|a\|^2\langle \zeta_i \Phi(b_i^*b_j)\zeta_j\rangle_{\mathscr{H}} \geq \sum_{i,j=1}^n \langle \zeta_i \Phi(b_i^*a^*ab_j)\zeta_j\rangle_{\mathscr{H}} = \langle \pi(a)\xi, \pi(a)\xi\rangle_\Phi \ . \tag{6.34}$$

In particular, whenever $\xi \in \mathscr{N}$, then $\pi(a)\xi \in \mathscr{N}$, and conseqeuntly for all $\xi, \xi' \in \mathscr{A} \otimes \mathscr{H}$,

$$\xi \sim \xi' \Rightarrow \pi(a)\xi \sim \pi(a)\xi' \ .$$

Therefore, $\pi(a)$ induces a linear transformation on $\mathscr{A} \otimes \mathscr{H}/\mathscr{N}$ through the definition

$$\pi(a)\{\xi\} = \{\pi(a)\xi\} \ ,$$

and as a further conseqeunce of (6.34), this linear transformation on $\mathscr{A} \otimes \mathscr{H}/\mathscr{N}$ is bounded with norm no greater than $\|a\|$. It therefore extends to an element of $\mathscr{B}(\mathscr{K})$, still denoted $\pi(a)$, that has the same norm. It is easy to show, mimicking the corresponding argument form the GNS construction, that $a \mapsto \pi(a)$ is a $*$-homomorphism from $\mathscr{A}$ to $\mathscr{B}(\mathscr{H})$, and we have proved the inequality on the right in (6.31). To prove the inequlaity in (6.32), suppose also that Φ is unital, and let $\zeta \in \mathscr{H}$. Then $\|\{1_{\mathscr{A}} \otimes \zeta\}\|_{\mathscr{K}}^2 = \langle \zeta, \Phi(1_{\mathscr{A}})\zeta\rangle_{\mathscr{H}} = \|\zeta\|_{\mathscr{H}}^2$ and $\|\pi(a)\{1_{\mathscr{A}} \otimes \zeta\}\|_{\mathscr{K}}^2 = \langle \zeta, \Phi(a^*a)\zeta\rangle_{\mathscr{H}}$. Hence for all non-zero $\zeta \in \mathscr{H}$,

$$\frac{\|\pi(a)\{1_{\mathscr{A}} \otimes \zeta\}\|_{\mathscr{K}}^2}{\|\{1_{\mathscr{A}} \otimes \zeta\}\|_{\mathscr{K}}^2} = \frac{\langle \zeta, \Phi(a^*a)\zeta\rangle_{\mathscr{H}}}{\|\zeta\|_{\mathscr{H}}^2} \ ,$$

and this proves (6.32).

Next, define $V : \mathscr{H} \to \mathscr{K}$ by

$$V\zeta = \{1_{\mathscr{A}} \otimes \zeta\} \in \mathscr{A} \otimes \mathscr{H}/\mathscr{N} \ .$$

Then $\|V\zeta\|_{\mathscr{K}}^2 = \langle \zeta, \Phi(1_{\mathscr{A}})\zeta\rangle_{\mathscr{H}}$ so that

$$\|V\| = \sup_{\|\zeta\|_{\mathscr{H}}=1} \{ \ \|V\zeta\|_{\mathscr{K}}^2\} = \sup_{\|\zeta\|_{\mathscr{H}}=1} \{ \ \langle \zeta, \Phi(1_{\mathscr{A}})\zeta\rangle_{\mathscr{H}}\} = \|\Phi(1_{\mathscr{A}})\| \ .$$

Next, for all $\zeta_1, \zeta_2 \in \mathscr{H}$

$$\langle \zeta_1, V^*\pi(a)V\zeta_2\rangle_{\mathscr{H}} = \langle \{1_{\mathscr{A}} \otimes \zeta_1\}, \pi(a)\{1_{\mathscr{A}} \otimes \zeta\}\rangle_\Phi = \langle \zeta_1, \Phi(a)\zeta_2\rangle_{\mathscr{H}} \ ,$$

and this proves (6.31). $\square$

Now let us specialize to the case in which that $\mathscr{A} = \mathscr{B}(\mathscr{H})$ for a separable Hilbert space $\mathscr{H}$. and in which Φ a completely positive map from $\mathscr{B}(\mathscr{H})$ to $\mathscr{B}(\mathscr{H})$ that is *faithful*, meaning that $\Phi(a^*a) = 0$ only for $a = 0$. We also suppose that Φ is unital.

Then by Stinespring's Theorem, there exists a Hilbert space $\mathscr{K}$, and an isomerty $V : \mathscr{H} \to \mathscr{H}$, and a represetnation π of $\mathscr{B}(\mathscr{H})$ on $\mathscr{B}(\mathscr{K})$. By (6.32), when Φ is faithful, π is injective. Therefore, by Theorem 2.26, π is an isometric isomprphism of $\mathscr{B}(\mathscr{H})$ otno its image in $\mathscr{B}(\mathscr{K})$.

Let $\mathscr{B}$ denote $\pi(\mathscr{B}(\mathscr{H}))$, the image of $\mathscr{B}(\mathscr{H})$ usnder π. We may apply Corollary 5.18 to show that there is a Hilbert space $\mathscr{H}'$ such and a unitary $W : \mathscr{K} \to \mathscr{H} \otimes \mathscr{H}'$ such that $\pi(a) = W^*(a \otimes 1_{\mathscr{H}'})W$, and then

$$\Phi(a) = (WV)^*(a \otimes 1_{\mathscr{H}'})WV \ .$$

Moreover, since V is an isometry, the range of WV may be identified with $\mathscr{H}$. Then choosing any unit vector $\zeta \in \mathscr{H}'$, the map $V_\eta : \mathscr{H} \to \mathscr{H} \otimes \mathscr{K}$ is an isometry, and its range may also be identified with $\mathscr{H}$. Let U be any unitary extnding the map $\eta \otimes \zeta \mapsto WV\eta$, and then we have that $UV_\eta = WV$ so that finally we obtain $\Phi(a) = V_\eta^* U^*(a \otimes 1_{\mathscr{H}'})UV_\eta$. We have proved:

6.19 THEOREM. *Let $\mathscr{H}$ be a separable Hilbert space and let Φ be a fiathful unital completely positive map from $\mathscr{B}(\mathscr{H})$ to $\mathscr{B}(\mathscr{H})$. Then there is a Hilbert space $\mathscr{H}'$, a unit vector $\zeta \in \mathscr{H}'$ and a unitary U on $\mathscr{H} \otimes \mathscr{H}'$ such that for all $a \in \mathscr{B}(\mathscr{H})$,*

$$\Phi(a) = V_\eta^* U^*(a \otimes 1_{\mathscr{H}'})UV_\eta \tag{6.35}$$

where V_ζ is the isomentry from $\mathscr{H}$ to $\mathscr{H} \otimes \mathscr{H}'$ given by $V_\zeta \eta = \eta \otimes \zeta$.

Now let ρ be a density matrix on $\mathscr{H}$; i.e., a postive trace class operator with $\mathrm{Tr}[\rho] = 1$. Define a linear functional on $\mathscr{B}(\mathscr{H})$ by

$$a \mapsto \mathrm{Tr}_{\mathscr{H}}[\rho\Phi(a)] \ .$$

By Theorem 6.19, we can write

$$\begin{aligned}
\mathrm{Tr}_{\mathscr{H}}[\rho\Phi(a)] &= \mathrm{Tr}_{\mathscr{H}}[\rho V_\eta^* U^*(a \otimes 1_{\mathscr{H}'})UV_\eta] \\
&= \mathrm{Tr}_{\mathscr{H}\otimes\mathscr{H}'}[(\rho \otimes |\zeta\rangle\langle\zeta|)U^*(a \otimes 1_{\mathscr{H}'})U] \\
&= \mathrm{Tr}_{\mathscr{H}\otimes\mathscr{H}'}[(U(\rho \otimes |\zeta\rangle\langle\zeta|)U^*)(a \otimes 1_{\mathscr{H}'})] \\
&= \mathrm{Tr}_{\mathscr{H}}[\mathrm{Tr}_{\mathscr{H}'}[U(\rho \otimes |\zeta\rangle\langle\zeta|)U^*]a]
\end{aligned}$$

Therefore, we may define $\Phi^*(\rho)$ by

$$\Phi^*(\rho) = \mathrm{Tr}_{\mathscr{H}'}[U(\rho \otimes |\zeta\rangle\langle\zeta|)U^*] \ . \tag{6.36}$$

Then we have

$$\mathrm{Tr}[\Phi^*(\rho)a] = \mathrm{Tr}[\rho\Phi(a)]] \tag{6.37}$$

for all $a \in \mathscr{B}(\mathscr{H})$. Recalling the Krauss representation for the partial trace $\mathrm{Tr}_{\mathscr{H}'}$ that is associated to any orthonormal basis $\{\zeta_j\}$ of $\mathscr{H}'$, we have the Krauss representation of $\Phi^*(\rho)$:

$$\Phi^*(\rho) = \sum_j V_{\eta_j}^*[U(\rho \otimes |\zeta\rangle\langle\zeta|)U^*]V_{\eta_j} = \sum_j (U^*V_{\eta_j})^*(V_\zeta \rho V_\zeta^*)(U^*V_{\eta_j}) \ . \tag{6.38}$$

That is,

$$\mathrm{Tr}[\Phi^*(\rho)a] = \sum_j A_j^*\rho A_j \quad \text{where} \quad A_j = V_\zeta^* U^* V_{\eta_j} \ . \tag{6.39}$$

6.7 Fixed points

6.20 LEMMA. *Let $\mathscr{A}$ be a C^* algebra. Then* $\begin{bmatrix} 0 & a \\ a^* & b \end{bmatrix} \geq 0$ *in $M_2(\mathscr{A})$ if and only if $a = 0$ and $b \geq 0$. Likewise,* $\begin{bmatrix} b & a \\ a^* & 0 \end{bmatrix} \geq 0$ *in $M_2(\mathscr{A})$ if and only if $a = 0$ and $b \geq 0$.*

Proof. We compute

$$\left\langle \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \begin{bmatrix} 0 & a \\ a^* & b \end{bmatrix} \begin{pmatrix} \eta \\ \xi \end{pmatrix} \right\rangle_{\mathscr{H}_2} = 2\Re\langle \eta, a\xi\rangle_{\mathscr{H}} + \langle \xi, b\xi\rangle_{\mathscr{H}} \ .$$

If $a \neq 0$, choose ξ so that $a\xi \neq 0$, and then, for $t > 0$, $\eta = -ta\xi$. Then

$$2\Re\langle \eta, a\xi\rangle_{\mathscr{H}} + \langle \xi, b\xi\rangle_{\mathscr{H}} = -2t\|a\xi\|^2 + \langle \xi, b\xi\rangle_{\mathscr{H}} .$$

For sufficiently large t, the irght hand side is negative. Hence positivity of $\begin{bmatrix} 0 & a \\ a^* & b \end{bmatrix}$ implies that $a = 0$, and then it is clear that $b \geq 0$. The converse is evident, and the statement for the matrix with 0 in the lower right position follows in the same way.

□

Let $\mathscr{A}$ and $\mathscr{B}$ be unital C^* algebras, and let $\Phi : \mathscr{A} \to \mathscr{B}$ be completely positive and unital. Since Φ_2 is 2-positive, Kadison's inequality applied to Φ_2 at $\begin{bmatrix} a & b^* \\ 0 & 0 \end{bmatrix}$ yields

$$\left(\Phi_2\left(\begin{bmatrix} a & b^* \\ 0 & 0 \end{bmatrix}\right)\right)^* \left(\Phi_2\left(\begin{bmatrix} a & b^* \\ 0 & 0 \end{bmatrix}\right)\right) \leq \Phi_2\left(\begin{bmatrix} a^*a & a^*b^* \\ ba & bb^* \end{bmatrix}\right) ,$$

and this is

$$\begin{bmatrix} \Phi(a^*a) - \Phi(a)^*\Phi(a) & \Phi(a^*b^*) - \Phi(a)^*\Phi(b)^* \\ \Phi(ba) - \Phi(b)\Phi(a) & \Phi(bb^*) - \Phi(b)\Phi(b)^* \end{bmatrix} .$$

By Lemma 6.20, if either $\Phi(a^*a) = \Phi(a)^*\Phi(a)$ or $\Phi(bb^*) = \Phi(b)\Phi(b)^*$, then $\Phi(ba) = \Phi(b)\Phi(a)$. Conversely, if for some $a \in \mathscr{A}$, $\Phi(ba) = \Phi(b)\Phi(a)$ for all b, then taking $b = a^*$, we have $\Phi(a^*a) = \Phi(a)^*\Phi(a)$, and if for some $b \in \mathscr{A}$, $\Phi(ba) = \Phi(b)\Phi(a)$ for all a, then taking $a = b^*$, $\Phi(bb^*) = \Phi(b)\Phi(b)^*$. We obtain:

6.21 THEOREM. *Let $\mathscr{A}$ and $\mathscr{B}$ be unital C^* algebras, and let $\Phi : \mathscr{A} \to \mathscr{B}$ be completely positive and unital. Then*

(i) $\{\ a \in \mathscr{A}\ :\ \Phi(a^*a) = \Phi(a)^*\Phi(a)\ \} = \{a \in \mathscr{A}\ :\ \Phi(ba) = \Phi(b)\Phi(a) \text{ for all } b \in \mathscr{A}\ \}$, *and this set is a subalgebra of $\mathscr{A}$, and Φ is a homomorphism when restricted to this set.*

(ii) $\{\ a \in \mathscr{A}\ :\ \Phi(aa^*) = \Phi(a)\Phi(a)^*\ \} = \{a \in \mathscr{A}\ :\ \Phi(ab) = \Phi(a)\Phi(b) \text{ for all } b \in \mathscr{A}\ \}$, *and this set is a subalgebra of $\mathscr{A}$, and Φ is a homomorphism when restricted to this set.*

(iii) the set

$$\{\ a \in \mathscr{A}\ :\ \Phi(a^*a) = \Phi(a)^*\Phi(a)\} \cap \{\ a \in \mathscr{A}\ :\ \Phi(aa^*) = \Phi(a)\Phi(a)^*\ \} \tag{6.40}$$

is a C^* subalgebra of $\mathscr{A}$, and Φ is a $*$-homomorphism when restricted to this set.

6.22 DEFINITION (Multiplicative domain)**.** The set in *(i)* of Theorem 6.21 is the *left multiplicative domain of* Φ, and the set *(ii)* of Theorem 6.21 is the *right multiplicative domain of* Φ. Their intersection is the *multiplicative domain of* Φ.

6.23 DEFINITION (Invariant states)**.** Let $\mathscr{A}$ be a unital C^* algebras and let $\Phi : \mathscr{A} \to \mathscr{B}$ be completely positive and unital. A state $\phi \in \mathscr{A}_{+,1}$ is *invariant under* Φ in case for all $a \in \mathscr{A}$,

$$\phi(\Phi(a)) = \phi(a) \ . \tag{6.41}$$

6.24 THEOREM. *Let $\mathscr{A}$ be a unital C^* algebras and let $\Phi : \mathscr{A} \to \mathscr{B}$ be completely positive and unital, and suppose that ϕ is a faithful state on $\mathscr{A}$ that is invariant under Φ. Define*

$$\mathscr{C} = \{ \ c \in \mathscr{A} \ : \Phi(c) = c \ \} \ .$$

Then $\mathscr{C}$ is a C^* algebra of $\mathscr{A}$, and for all $a, b \in \mathscr{A}$ and all $c \in \mathscr{C}$,

$$\Phi(acb) = \Phi(a)c\Phi(b) \ . \tag{6.42}$$

Proof. By Kadison's inequality, for $c \in \mathscr{C}$,

$$\Phi(c^*c) - c^*c \geq 0 \ .$$

Applying the faithful invariant state ϕ yields

$$\phi(\Phi(c^*c) - c^*c) = \phi(c^*c) - \phi(c^*c) = 0 \ .$$

Since ϕ is faithful, $\Phi(c^*c) = c^*c$, and so $c^*c \in \mathscr{C}$. We now have that $\Phi(c^*c) = \Phi(c)^*\Phi(c)$. Since $\mathscr{C}$ is closed under the involution, the same applies with c replaced by c^*. Thus $\mathscr{C}$ is in the multiplicative domain of Φ, and is a C^* subalgebra of $\mathscr{A}$ by Theorem 6.20, which then also yields (6.42).

□

7 Quantum Measurement

7.1 Measurement in the early days of the Schrödinger equation

In Quantum Mechanics, the state of a system is given by a positive trace class operator ρ on a Hilbert space $\mathscr{H}$ such that $\mathrm{Tr}[\rho] = 1$. The state is a *pure state* in case ρ is rank-one.

This is more or less consistent with the terminology that we have previously introduced since we may regard ρ as a linear functional on $\mathscr{B}(\mathscr{H})$ through the identifincation of ρ with that linear functional $a \mapsto \mathrm{Tr}[\rho a]$. Not all states on $\mathscr{B}(\mathscr{H})$ are of this form, but all σ-weakly continuous states are: The totaliy of these may be identified with the set

$$\{ \ \rho \in \mathscr{T}(\mathscr{H}) \ : \ \rho \geq 0 \quad \text{and} \quad \mathrm{Tr}[\rho] = 1 \ \} \ .$$

By the Spectral Theorem, the extreme points of this set is precisely the set of rank-one projections.

The time evolution of states in quantum mechanics is given by the Schroödinger equation: There is a one parameter unitary group u_t on $\mathscr{H}$ of the from $u_t = e^{-ith}$ where h is a self adjoint operator on $\mathscr{H}$. If at time $t = 0$ the system is in the state ρ, then at time t it is in the state $u_t\rho u_t^*$.

The evolution of states is completely deterministic. However, the outcomes of experiemnts are intrinsically random. The observable of a quantum system; i.e., the quantifiable properties of the system that may be measured in the laboratory correspond to self adjoint operators a on $\mathscr{H}$. The totality of these observables generate a subalgebra of $\mathscr{B}(\mathscr{H})$ called the *algebra of observables.*

von Neumann gave the first mathematical treatment of quantum measurement in his 1927 paper [25]. This followed the 1926 paper of Born [3] which contains in a footnote the "Born interpretation" of the meaning Schrödinger's wave function: Born studied scattering in the new quantum mechanical framework and concluded that the only possible interpretation of Schöedinger's wave function $\psi(x,t)$ was that the outcomes of quantum measurements were *inherently probabilistic*, and that with $\psi(x,t)$ normalized so that $\int_{\mathbb{R}^3} |\psi(x,t)|^2 \mathrm{d}x = 1$, the probability of finding the position of the particle described by this wave function ψ in a (Borel measurable) set $B \subset \mathbb{R}^3$ is $\int_B |\psi(x,t)|^2 \mathrm{d}x$, and likewise, if $\widehat{\psi}(k,t)$ is the Fourier transfrom of $\psi(x,t)$, then $\int_B |\widehat{\psi}(k,t)|^2 \mathrm{d}k$ is the probability of finding the momentum of the particle described by this wave function ψ in a (Borel measurable) set $B \subset \mathbb{R}^3$ is In 1952, Max Born was awarded the Nobel Prize in Physics for the content of this footnote among other contributions.

In 1927, von Neumann [25] developed this statistical interpretation further. In von Neumann's theory of measurement, observables correspond to self adjoint operators a on a Hilbert space $\mathscr{H}$, and a measurement of such an observable a always yeilds a value in the spectrum of a. One of the triumphs of Schrödinger's work on his equation was that he was able to solve the eigenvalue problem

$$(-\Delta - |x|^{-1})\psi(x) = \lambda\psi(x)$$

for the energy of a hydrogen atom; here Δ denote the Laplaician and $x \in \mathbb{R}^3$. The differences between the eigenvalues that he found corresponded precisely to the energies of spctral lines observed in the light emitted from hydrogen atoms in scattering experiment: In the "old quantum mechanics", the possible energy levels of a hydrogen were quantized by fiat. In the "new quantum mechanics", they were quantized because the energy was reprentend by a self adjoint operator, not a function on "phase space", and the possible vlaues one coul observe were precisely the eigenvalues.

Extrapoloating form this and other early experiments, and Born's interpretation, von Neumann proceeded to a more general formulation.

The probability that the experiment measuring the observable represented by a self-adjoint operator a, yields a value in a Borel set $B \subset \sigma(a)$, is givin by

$$\langle \psi, 1_B(a)\psi \rangle_{\mathscr{H}}$$

if the system is in the state given by the orthogonal projection onto the normalized vector $\psi \in \mathscr{H}$. von Neumann's theorey was only deleoped in the context of observales (self adjoint operators) with discrete spectrum. However, it went further than the Born interpretation in an important way: von Neumann went on to describe the state of the system after a measurement, and the matter of repeated measurements.

von Neumann's discussion of measurement applied to self adjoint observable with discrete spectrum. Let a be such an observable, and let

$$a = \sum_j \lambda_j e_j \tag{7.1}$$

be its spectral decomposition so that e_j is the orthoongnal projector onto the jth eigenspace.

During the measurement process, the state ρ is transfromed into

$$\Phi^*(\rho) = \sum_j p_j \rho_j$$

where

$$p_j = \mathrm{Tr}[e_j \rho_j] \quad \text{and} \quad \rho_j = \begin{cases} p_j^{-1} r_j \rho e_j & p_j \neq 0 \\ 0 & p_j = 0 \end{cases} .$$

By the cyclicity of the trace, $p_j = \mathrm{Tr}[e_j \rho e_j] \geq 0$ since $e_j \rho e_j \geq 0$. Moreover,

$$\sum_j p_j = \mathrm{Tr}\left[\sum_j \rho e_j\right] = \mathrm{Tr}[\rho] = 1 .$$

Thus, the p_j specify a discrete probability distribution. In agreement with the Born interpretation, for each j, p_j is the probability that the measurement of a yields the value λ_j when the sytem is prepared in the state ρ.

For each j with $p_j > 0$, $\rho_j = \dfrac{1}{\mathrm{Tr}[\rho e_j]} e_j \rho e_j$ is a density matrix, and it gives the state of the system after the measurement process in the case that λ_j is the observed value of a.

The map Φ^* is the predual of the map Φ defined by

$$\Phi(x) = \sum_j e_j x e_j .$$

Notice that Φ is completely positive and unital. Moreover, since $e_j e_k = \delta_{j,k} e_j$,

$$\Phi(\Phi(x)) = \sum_{j,k} e_k e_j x e_j e_k = \sum_j e_j x e_j = \Phi(x)) ,$$

Φ is idempotent, and likewise, Φ^* is idempotent.

This is a direct reflection of von Neumann's *repeatibitly hypothesis*: If a measurement is repeated a second time, the same result is obtained both times. Though this hypothesis fit well with the experiments done in the early days when von Neumann made his proposal, efforts to extend it his measurement theory to observables with continuous spectrum indicated that this hypothesis might not be applicable in general, and Wigner [27] gave physical arguments agiaits its general validity.

7.2 Quanum instruments, operations and channels

Important progress was made by Davies and Lewis [5] who introduced the notion of a *quantum instrument*, dispensed completely with the repeatibility hypothesis, and treated measurement of variables with continuous spectrum. The following version of their definition is taken from Ozawa [20].

7.1 DEFINITION (Quantum instrument)**.** Let X be a complete, separable metric space, and let $\mathscr{B}(X)$ denote its Borel σ-algebra. Let $\mathscr{M}$ be a von Neumann algebra on a separable Hilbert space $\mathscr{H}$, and let $\mathscr{M}_*$ be its predual. Let $\mathscr{P}(\mathscr{M}_*)$ denote the set of all positve linear transformation of

$\mathscr{M}_*$ into itslef. Let $\langle \cdot, \cdot \rangle$ denote the dual pairing between $\mathscr{M}_*$ and $\mathscr{M}$. Then a *quantum instrument* is a map $\mathscr{I} : \mathscr{B}(X) \to \mathscr{P}(\mathscr{M}_*)$ such that

(i) For each $\rho \in \mathscr{M}_*$, $\langle \mathscr{I}(X)\rho, 1_{\mathscr{M}} \rangle = \langle \rho, 1_{\mathscr{M}} \rangle$.

(ii) For each disjoint sequence $\{B_j\}$ in $\mathscr{B}(X)$,

$$\mathscr{I}(\cup_{j=1}^{\infty} B_j) = \sum_{j=1}^{\infty} \mathscr{I}(B_j)$$

where the sum is convergent in the strong operator topology on $\mathscr{P}(\mathscr{M}_*)$.

A quantum instrument is called a *completely positive quantum instrument* in case

(iii) for each $B \in \mathscr{B}(X)$, $\mathscr{I}(B)$ is completely positive.

Every von Neuman measurement is given a completely positive quantum instrument. Indeed, for a self adjoint operator a with discrete spectrum and spectral decomposition (7.1), let X be the spectrum of a; i.e., $X = \{\lambda_j\}$. For each j, define Φ_j^* by

$$\Phi_j^*(\rho) = e_j \rho e_j \ .$$

Finally, for any $B \subset X$, define

$$\mathscr{I}(\rho) = \sum_{j:\lambda_j \in B} \Phi_j^*(\rho) \ .$$

Two other definition are central to the theory of quantum measurement:

7.2 DEFINITION. Let $\mathscr{M}$ be a von Neumann algebra. A *quantum operation* is the predual of a completely positve map Φ from $\mathscr{M}$ to $\mathscr{M}$ such that $\Phi(1) \leq 1$. A *quantum channel* is a unital quantum operation.

Note that if $\mathscr{I}$ is a quantum instrument for the metric space X, the $\mathscr{I}(X)$ is a quantum channel.

7.3 The Mean Ergodic Theorem

In the next section we shall need von Neumann's 1931 Mean Ergodic Theorem. We begin with some preliminaries.

Recall that for any Hilbert space $\mathscr{H}$ and any $a \in \mathscr{B}(\mathscr{H})$,

$$\ker(a^*)^\perp = \overline{\operatorname{ran}(a)} \ . \tag{7.2}$$

To see this, note that for any $\zeta \in \operatorname{ran}(a)^\perp$, and any $\eta \in \mathscr{H}$, $\langle a^*\zeta, \eta \rangle_{\mathscr{H}} = \langle \zeta, a\eta \rangle_{\mathscr{H}} = 0$, which implies that $\operatorname{ran}(a)^\perp \subset \ker(a^*)$, and therefore $\ker(a^*)^\perp \subset \operatorname{ran}(a)^{\perp\perp} = \overline{\operatorname{ran}(a)}$.

On the other hand, for any $\zeta \in \ker(a^*)$, and any $\eta \in \mathscr{H}$, $0 = \langle a^*\zeta, \eta \rangle_{\mathscr{H}} = \langle \zeta, a\eta \rangle_{\mathscr{H}} = 0$, which implies that $\ker(a^*) \subset \operatorname{ran}(a)^\perp$, and therefore $\overline{\operatorname{ran}(a)} = \operatorname{ran}(a)^{\perp\perp} \subset \ker(a^*)^\perp$.

Now let u be unitary on $\mathscr{H}$. Then $\ker(u-1)$ is precisely the sets of vectors in $\mathscr{H}$ that are invariant under u. Notice that

$$\eta \in \ker(u-1) \iff u\eta = \eta \iff \eta - u^*\eta \iff \eta \in \ker((u-1)^*) \ .$$

That is, $\ker(u-1) = \ker((u-1)^*)$, and then from (7.2), for any unitary u,

$$\ker(u-1)^\perp = \overline{\operatorname{ran}(u-1)} \ . \tag{7.3}$$

7.3 THEOREM (von Neumann's Mean Ergodic Theorem)**.** *Let u be any unitary on a Hilbert space $\mathscr{H}$, and let p be the orthogonal projector onto* $\ker(u-1)$*. Then*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} u^n = p$$

where the convergence is in the strong operator toplogy.

Proof. We must show that for all $\eta \in \mathscr{H}$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} u^n\eta = p\eta \ . \tag{7.4}$$

Write $\eta - \eta_1 + \eta_2$ where $\eta_1 = p\eta$ and $\eta_2 = p^\perp\eta$. By (7.3), $\eta_2 \in \overline{\text{ran}}(u-1)$. Since $u^n\eta_1 = \eta_1 = p\eta_1$ for all n, (7.4) is trivially true with η_1 in place of η.

Since the operator $\displaystyle\frac{1}{N}\sum_{n=1}^{N} u^n$ is a contraction, it converges to zero strongly on $\overline{\text{ran}(u-1)} = \ker(u-1)^\perp$ if and only if it converges to zero strongly on $\text{ran}(u-1)$. Let $(u-1)\zeta \in \text{ran}(u-1)$, By the telescoping sum identity, for each N,

$$\frac{1}{N}\sum_{n=1}^{N} u^n(u-1)\zeta = \frac{1}{N}(u^{N+1}-u)\zeta \ ,$$

and this tends to zero as N tends to infinity. Since $p\eta_2 = 0$, it follows that (7.4) is true with η_2 in place of η. □

We need a more general result that is obtained by combining Theorem 7.3 with a *dilation theorem* of Sz.-Nagy. We introduce the concept of a dilation theorem with the sinplest example:

Let $\mathscr{H}$ be a Hilbert space and let v be a partial on $\mathscr{H}$. Let $p = (vv^*)^\perp = 1_{\mathscr{H}} - vv^*$ be the projector onto $\text{ran}(v)^\perp$. Define an operator u on $\mathscr{H} \oplus \mathscr{H}$ by

$$u = \begin{bmatrix} v & p \\ 0 & v^* \end{bmatrix} \ . \tag{7.5}$$

Then since $v^*v = 1_{\mathscr{H}}$ and $pv = 0$, $uu^* = \begin{bmatrix} v & p \\ 0 & v^* \end{bmatrix}\begin{bmatrix} v^* & 0 \\ p & v \end{bmatrix} = \begin{bmatrix} 1_{\mathscr{H}} & 0 \\ 0 & 1_{\mathscr{H}} \end{bmatrix} = 1_{\mathscr{H}\oplus\mathscr{H}}$. This shows that u is unitary on $\mathscr{H} \oplus \mathscr{H}$. Also, $u^2 = \begin{bmatrix} v^2 & 0 \\ 0 & (v^*)^2 \end{bmatrix}$, so that for all $m \in \mathbb{N}$, $u^{2m} = \begin{bmatrix} v^{2m} & 0 \\ 0 & (v^*)^{2m} \end{bmatrix}$. Therefore, of we define $U : \mathscr{H} \to \mathscr{H} \oplus \mathscr{H}$ by $U\eta = (\eta, 0)$, we have that for all $n \in \mathbb{N}$,

$$v^n = U^*u^nU \ . \tag{7.6}$$

Note that U is one of the obvious embeddings of $\mathscr{H}$ into $\mathscr{H} \oplus \mathscr{H}$, and U^* is the correpsonding projection back onto $\mathscr{H}$. When u and v are related in this way, we say that the unitary u is the *dilation* of the isometry v, and that v is the *compression* of the unitary u.

We now explain how to dilate a contraction into an isometry. Let $a \in \mathscr{B}(\mathscr{H})$ be an arbitrary contraction; i.e., $\|a\| \leq 1$. The Hilbert space $\ell_2 \otimes \mathscr{H}$ may be identifed with the space of all sequences $\{\eta_j\}$, $\eta_j \in \mathscr{H}$ for $j \in \mathbb{N}$ such that $\sum j = 1^{\infty} \|\eta_j\|^2_{\mathscr{H}} < \infty$, and equipped with the inner product

$$\langle \{\eta_j\}, \{\zeta_j\} \rangle_{\ell_2 \otimes \mathscr{H}} = \sum_{j=1}^{\infty} \langle \eta_j, \zeta_j \rangle_{\mathscr{H}} \ .$$

Define $w : \ell_2 \otimes \mathscr{H} \to \ell_2 \otimes \mathscr{H}$ by

$$w(\eta_1, \eta_2, \eta_3, \dots) = (a\eta_1, (1 - a^* a)^{1/2} \eta_1, \eta_2, \eta_3, \dots) \ .$$

Then w is evidently an isometry, and if we define $W : \mathscr{H} \to \mathscr{K}$ by $W\eta = (\eta, 0, 0, \dots)$, we have that for all $n \in \mathbb{N}$,

$$a^n = W^* w^n W \ . \tag{7.7}$$

Note that W is one of the obvious embeddings of $\mathscr{H}$ into $\ell_2 \otimes \mathscr{H}$, and W^* is the correpsonding projection back onto $\mathscr{H}$.

7.4 THEOREM (Sz.-Nagy's Dilation Theorem)**.** *Let $\mathscr{H}$ be a Hilbert space, and let a be a contraction on $\mathscr{H}$. Then there is a Hilbert space $\mathscr{K}$ and an isometry $V : \mathscr{H} \to \mathscr{K}$, and a unitary u on $\mathscr{K}$ such that for all $n \in \mathbb{N}$, $n \in \mathbb{N}$,*

$$a^n = V^* u^n V \ . \tag{7.8}$$

We now combine Theorem 7.3 and Theorem 7.4: Let $\mathscr{H}$ be a Hilbert space, and let a be a contraction on $\mathscr{H}$. Then with V and u as in Theorem 7.4,

$$\frac{1}{N} \sum_{n=1}^{N} a^n = V^* \left(\frac{1}{N} \sum_{n=1}^{N} u^n \right) V \ .$$

By Theorem 7.3, $\lim_{N \to \infty} \left(\frac{1}{N} \sum_{n=1}^{N} a^n \right)$ exists and is equal to $V^* p V$ where p is the projection onto $\ker(u - 1)$. Clearly,

$$\lim_{N \to \infty} a \left(\frac{1}{N} \sum_{n=1}^{N} a^n \right) = \lim_{N \to \infty} \left(\frac{1}{N} \sum_{n=1}^{N} a^n \right) a = \lim_{N \to \infty} \left(\frac{1}{N} \sum_{n=1}^{N} a^n \right) \ .$$

It follows that $aV^*pV = V^*pVa = V^*pV$, and then for all $n \in \mathbb{N}$,

$$a^n V^* p V = V^* p V a^n = V^* p V \ .$$

Averaging over $n = 1, \dots, N$ and taking the limit $N \to \infty$, we obtain $(V^*pV)^2 = V^*pV$, and hence V^*pV is an orthogonal projection. If $\eta \in \ker(a - 1_{\mathscr{H}})$, then $a^n \eta = \eta$ for all n, and hence $V^*pV\eta = \eta$. That is, $\ker(a - 1_{\mathscr{H}}) \subset \operatorname{ran}(V^*pV)$. Next, using $aV^*pV = V^*pV$ once more, $V^*pV(a - 1_{\mathscr{H}})^*(a - 1_{\mathscr{H}})V^*pV = 0$, and hence $\operatorname{ran}(V^*pV) \subset \ker(a - 1_{\mathscr{H}})$. This shows that V^*pV is the orthogonal projection onto $\ker(a - 1_{\mathscr{H}})$. We have proved:

7.5 THEOREM (Mean Ergodic Theorem for Contractions)**.** *Let $\mathscr{H}$ be a Hilbert space. Let a be a contraction on $\mathscr{H}$. Let q be the orthogonal projection onto* $\ker(a - 1_{\mathscr{H}})$*. Then*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} a^n = q$$

where the convergence is in the strong operator topology.

7.4 Lindblad's No Cloning Theorem

In this section, following Lindblad, we consider *quantum copying* in a finite dimensional quantum system: The Hilbert space $\mathscr{H}$ is simply $\mathbb{C}^n$ for some finite n, and the algebra of observables $\mathscr{M}$ is a sub-algebra of $M_n(\mathbb{C})$. In fact, following Lindblad, we make the further assumption that $\mathscr{M} = M_n(\mathbb{C})$, so that all self adjoint $n \times n$ matrices are observable.

Let $\Phi : \mathscr{M} \to \mathscr{M}$ be a quantum channel. Since $\mathscr{M} = M_n(\mathbb{C})$ we may apply Choi's Theorem, and can write

$$\Phi(a) = \sum_{j=1}^{m} V_j^* a V_j \quad \text{where} \quad \sum_{j=1}^{m} V_j^* V_j = 1 \ , \tag{7.9}$$

where $m \leq n^2$.

We are also interested in the predual action of Φ on states: For a state ρ, define $\Phi_*(\rho)$ by $\Phi_*(\rho)(a) = \rho(\Phi(a))$ for all a.

More genrally, Φ_* is defined on the set of all self adjoint linear functionals on $\mathscr{M}$, which we regard as a real Banach space X. Then the positive linear functionals are a convex cone K in X and $X = K - K$. Since K is invariant under Φ_*, it follows from the Krein-Rutman Theorem that the spectral radius of Φ_* is an eigenvalue λ of Φ_* with an eigenvector in K. That is, there exists $\rho_0 \in K$ such that $\Phi_*(\rho_0) = \lambda\rho_0$. Since ρ_0 is not zero, we may normalize it so that $\rho_0(1) = 1$.

Since Φ is unital, for all n, $\Phi(1) = 1$, and so

$$1 = \rho_0(1) = \rho_0(\Phi 1) = \lambda\rho(1) \ .$$

Therefore, $\lambda = 1$. Thus, ρ_0 is an invariant state for Φ.

Let $\mathcal{S}_\Phi$ denote the set of invariant states:

$$\mathcal{S}_\Phi = \{\rho \in \mathscr{M}_*^+ \ : \ \rho(1) = 1 \quad \text{and} \quad \Phi_*\rho = \rho \ \}$$

We have just seen that $\mathcal{S}_\Phi$ is not empty, and clearly it is a compact, convex set in $\mathscr{M}_*$.

For each $\rho \in \mathcal{S}_\Psi$, let p_ρ denote the *support* of ρ; i.e., the smallest orthogonal projector such that $\rho(p_\rho) = 1$, and it has the property that doe all $a \in \mathscr{M}$,

$$\rho(a) = \rho(p_\rho a p_\rho) \ .$$

In fact, identifying ρ with the density matrix such that $\rho(a) = \mathrm{Tr}[\rho a]$, p_ρ is the projection onto the range of the density matrix rho, in which case $\rho = p_\rho \rho p_\rho$.

By taking convex combinations, we find a $\rho_0 \in \mathcal{S}_\Psi$ with maximal support p_Ψ. Then the range of p_Ψ contains the range of ρ, considered as a density matrix, so that

$$p_\Psi \rho p_\Psi = \rho \quad \text{for all} \quad \rho \in \mathcal{S}_\Phi \ . \tag{7.10}$$

By Kadison's inequality, $\Psi(p_\Psi) = \Psi(p_\Psi^2) \geq \Psi(p_\Psi)^2$, and hence $\Psi(p_\Psi)$ is a contraction. Next,

$$1 = \rho_0(p_\Psi) = \rho_0(\Psi(p_\Psi))$$

and hence $p_\Psi \leq \Psi(p_\psi)$.

Now consider the map $\widetilde{\Psi}(a) = p_\Psi(\Psi(p_\Psi(a)p_\Psi))p_\Psi$. This is completely positive and $\widetilde{\Psi}(p_\Psi) = p_\Psi$. By (7.10), for all a, and all $\rho \in \mathcal{S}_\Phi$,

$$\rho(\widetilde{\Psi}(a)) = \rho(p_\Psi \Psi(p_\Psi a p_\Psi)p_\Psi) = \rho(\Psi(p_\Psi a p_\Psi)) = \rho(p_\Psi a p_\Psi) \ . \tag{7.11}$$

Note that p_Ψ is the identity in the algebra $\widetilde{\mathscr{M}} := p_\Psi \mathscr{M} p_\Psi$, and if we regard all operators in this algebra as opaerators on $\mathscr{H}_\Phi = \mathrm{ran}(p_\Phi)$, it is a von Neumann algebra and $\widetilde{\Psi}$ is a unital completely positive map on $\widetilde{\mathscr{M}}$. By (7.11), for all $a \in \widetilde{\mathscr{M}}$, $\rho(\widetilde{\Psi}(a)) = \rho(a)$, and thus every state in $\mathcal{S}_\Phi$ is invariant for $\widetilde{\Psi}$.

In the other direction, suppose that $\widetilde{\rho}$ is invariant for $\widetilde{\Psi}$ on $\widetilde{\mathscr{M}}$. Define a state ρ on $\mathscr{M}$ by

$$\rho(a) = \widetilde{\rho}(p_\Phi a p_\Phi) \ ,$$

and note that for all $a \in \mathscr{M}$, $\rho(a) = \rho(p_\Phi a p_\Phi)$. Then for all $a \in \mathscr{M}$,

$$\rho(a) = \widetilde{\rho}(p_\Phi a p_\Phi) = \widetilde{\rho}(\widetilde{\Psi}(p_\Phi a p_\Phi)) = \widetilde{\rho}(p_\Psi \Psi(p_\Phi a p_\Phi)p_\Psi) = \rho(\Psi(p_\Phi a p_\Phi))$$

Then ρ is an inavariant state for the completely positive but non-unital map $a \mapsto \Psi(p_\Phi a p_\Phi$.

In summary, every invariant state for Φ is an invariant state for another completely positive unital map $\widetilde{\Phi}$ on smaller algebra for which there is a fiathful invariant state ρ_0.

Now assume that we have made this reduction and the $\mathcal{S}_\Phi$ contains a faithful state ρ_0.

7.6 LEMMA. *Let $\mathscr{C}$ be the fixed point algebra of Φ. Then $\mathscr{C}$ is the commutant of $\{V_1, \dots, V_m\}$.*

Proof. Since $\sum_{j=1}^m V_j^* V_j = 1$, if $a \in \{V_1, \dots, V_m\}'$ than evidently $a \in \mathscr{C}$. That is, $\{V_1, \dots, V_m\}' \subset \mathscr{C}$.

The other containment more work to prove and makes use of the faithful invariant st ρ_0. For $a \in M_n(\mathbb{C})$, define $\rho(a) = n^{-1}\mathrm{Tr}[a]$. By Theorem 6.24, the existence of ρ_0 ensures that $\mathscr{C}$ is a C^* algebra. Next, for all $a \in M_n(\mathbb{C})$,

$$\begin{aligned}
\sum_{j=1}^n [a, V_j]^*[a, V_j] &= \sum_{j=1}^n (V_j^* a^* - a^* V_j^*)(a V_j - V_j a) \\
&= \sum_{j=1}^n (V_j^* a^* a V_j - V_j^* a^* V_j a - a^* V_j^* a V_j + a^* V_j^* V_j a) \\
&= \Phi(a^* a) - \Phi(a)^* a - a^* \Phi(a) - a^* a
\end{aligned}$$

Because $\mathscr{C}$ is a C^* algebra, When $a \in \mathscr{C}$, $\Phi(a^*a) = a\Phi(a)^* = a^*\Phi(a) = a^*a$. Therefore, when $a \in \mathscr{C}$ the calculuation we have made above yeilds $\sum_{j=1}^n [a, V_j]^*[a, V_j] = 0$, and this shows $\mathscr{C} \subset \{V_1, \dots, V_m\}'$. □

For our second use of the faithful invariant state ρ_0, we equip $\mathscr{M}$ with the inner product

$$\langle a, b \rangle_{\rho_0} = \rho_0(a^* b) \ ,$$

which makes it a Hilbert space, since ρ_0 is faithful. Now note that by Kadison's inequality,

$$\rho_0(a^* a) = \rho_0(\Phi(a^* a)) \geq \rho_0(\Phi(a)^* \Phi(a)) \ ,$$

which shows that Φ is a contraction in this Hilbert space.

7.7 LEMMA. *Let Φ be unital completely positive map from $M_n(\mathbb{C})$ to $M_n(\mathbb{C})$ with faithful invariant state ρ_0. . Then*

$$\lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Phi^n =: E_\Phi$$

exists and is a unital completely positive projection from $M_n(\mathbb{C})$ to $M_n(\mathbb{C})$.

Proof. Since Φ is a contraction in the Hilbert space asociated to any faithful invariant state, this is a direct consequence of Theorem 7.5. □

It is obvious that for all $a \in M_n(\mathbb{C})$, $\Phi(E_\Phi(a)) = E_\Phi(a)$, so that the range of E_Φ is contianed in $\mathscr{C}$, the fixed point algebra of Φ. But obviously if $a \in \mathscr{C}$, then $E_\Phi(a) = a$, and hence the range of E_Φ is exactly $\mathscr{C}$. Again by Theorem 6.24, for all $c \in \mathscr{C}$ and all $a, b \in M_n(\mathbb{C})$,

$$E_\Phi(acb) = E_\Phi(a)cE_\Phi(b) \quad \text{and} \quad E_\Phi = \Phi \circ E_\Phi = E_\Phi \circ \Phi = E_\Phi \circ E_\Phi \ . \tag{7.12}$$

We refer to E_Φ as the *conditional expectation* in $M_n(\mathbb{C})$ onto the subalgebra $\mathscr{C}$.

7.8 LEMMA. *Let Φ be a unital completely positive map on $\mathscr{M}$ with a faithful invariant invariant state ρ_0. Let E_Φ be the conditional expectation onto the C^* algebra of fixed points of Φ. Then $\mathcal{S}_\Phi$ is the image of the set of all states on $\mathscr{M}$ under $(E_\Phi)_*$.*

Proof. Let ρ be any state on $\mathscr{M}$, and consider the state $(E_\Phi)_*\rho$. Then for all $a \in \mathscr{M}$,

$$E_(\Phi)_*\rho(\Phi a)) = \rho(E_\Phi(\Phi(a)) = \rho(\Phi E_\Phi(a)) = \rho(E_\Phi(a)) = (E_\Phi)_*\rho(a) \ .$$

Thus, $(E_\Phi)_*\rho$ is an invariant state. On the other hand if ρ is an invariant state, $(E_\Phi)_*\rho = \rho$. □

We now apply our knowledge of the structure of $\mathscr{C}$: It is evidently a finite sum of type I factors. Thus, our Hilbert space $\mathscr{H} = \mathbb{C}^n$ is the direct sum of finitely many subspaces

$$\mathscr{H} = \bigoplus_{n=1}^{N} \mathscr{H}_n \ ,$$

each of which is invariant under $\mathscr{C}$, and such that for each n, the center of the restriction of $\mathscr{C}$ to $\mathscr{H}_n$ is trivial. Let p_n be the projection in $\mathscr{H}$ onto $\mathscr{H}_n$, and define $\mathscr{C}_n = \mathscr{C}p_n = p_n\mathscr{C}p_n$.

Since each $\mathscr{C}_n$ is a type I factor, each $\mathscr{H}_n$ can be factored as

$$\mathscr{H}_n = \mathscr{H}_{n,\ell} \otimes \mathscr{H}_{n,r}$$

and we have that

$$\mathscr{C}_n = \mathscr{B}(\mathscr{H}_{n,\ell}) \otimes 1_{\mathscr{H}_{n,r}} \ .$$

7.9 THEOREM. *Let Φ be a unital completely positive map on $\mathscr{M}$ with a faithful invariant invariant state ρ_0. Let E_Φ be the conditional expectation onto the C^* algebra of fixed points of Φ. Then there are density matrices $\{\sigma_1, \dots, \sigma_N\}$, where σ_n acts on $\mathscr{H}_{n,r}$, such that for all $y \in \mathscr{B}(\mathscr{H})$,*

$$E_\Phi(y) = \sum_{n=1}^{N} \mathrm{Tr}_{\mathscr{H}_{n,r}}[(1_{\mathscr{H}_{n,\ell}} \otimes \sigma_n)p_n y p_n] \otimes 1_{\mathscr{H}_{n,r}} \ .$$

Proof. Define $\Psi_0 : \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H})$ by $\Psi_0(a) = \sum_{n=1}^N p_n a p_n$ which is unital and completely positive. Since each $\mathscr{C}_n$ is invariant under Φ, it is clear that

$$E_\Phi = E_\Phi \circ \Psi_0 = \Psi_0 \circ E_\Phi \ .$$

Therefore, E_Φ has the form

$$E_\Phi(a) = \sum_{n=1}^N (E_\Phi)_n(p_n a p_n)$$

For each n, let u_n be the partial isometry embedding $\mathscr{H}_n$ into $\mathscr{H}$ so that u_n^* is the projection p_n considered as a map from $\mathscr{H}$ to $\mathscr{H}_n$, and not as an elelment of $\mathscr{B}(\mathscr{H})$. Then for any n and any $a \in \mathscr{B}(\mathscr{H}_{n,\ell})$, $u_n(a \otimes 1_{\mathscr{H}_{n,r}})u_n^* \in \mathscr{C}_n \subset \mathscr{C}$, and hence $E_\Phi(u_n(a \otimes 1_{\mathscr{H}_{n,r}})u_n^*) = u_n(a \otimes 1_{\mathscr{H}_{n,r}})u_n^*$, which we can write more simply as

$$(E_\Phi)_n(a \otimes 1_{\mathscr{H}_{n,r}}) = a \otimes 1_{\mathscr{H}_{n,r}} \ .$$

Since E_Φ has the properties listed in (7.12), so does each $(E_\Phi)_n$, and by what we have remarked just above, since

$$a \otimes b = (a \otimes 1_{\mathscr{H}_{n,r}})(1_{\mathscr{H}_{n,\ell}} \otimes b) = (1_{\mathscr{H}_{n,\ell}} \otimes b)(a \otimes 1_{\mathscr{H}_{n,r}}) \ ,$$

$$(E_\Phi)_n(a \otimes b) = (a \otimes 1_{\mathscr{H}_{n,r}})(E_\Phi)_n(1_{\mathscr{H}_{n,\ell}} \otimes b) = (E_\Phi)_n(1_{\mathscr{H}_{n,\ell}} \otimes b)a \otimes 1_{\mathscr{H}_{n,r}} \ .$$

Thus $(E_\Phi)_n(1_{\mathscr{H}_{n,\ell}} \otimes b)$ belongs to $\mathscr{C}_n = \mathscr{B}(\mathscr{H}_{n,\ell}) \otimes 1_{\mathscr{H}_{n,r}}$ and commutes with every element of $\mathscr{B}(\mathscr{H}_{n,\ell}) \otimes 1_{\mathscr{H}_{n,r}}$. It follows that $(E_\Phi)_n(1_{\mathscr{H}_{n,\ell}} \otimes b)$ is a multiple of the identity on $\mathscr{H}_n$; i.e.,

$$(E_\Phi)_n(1_{\mathscr{H}_{n,\ell}} \otimes b) =: \lambda(b) 1_{\mathscr{H}_n} \ .$$

Since $(E_\Phi)_n$ is completely positive and unital, the map sending b to $\lambda(b)$ is evidently a state, and thus there is a positve matrix σ_n on $\mathscr{H}_{n,r}$ with $\mathrm{Tr}[\sigma_n] = 1$ such that $\lambda(b) = \mathrm{Tr}[\sigma_n b]$ for all b. Therefore,

$$(E_\Phi)_n(a \otimes b) = \mathrm{Tr}[\sigma_n b] a \otimes 1_{\mathscr{H}_{n,r}} \ .$$

This in turn shows that for all $x \in \mathscr{B}(\mathscr{H}_n)$,

$$(E_\Phi)_n(x) = \mathrm{Tr}_{\mathscr{H}_{n,r}}[(1_{\mathscr{H}_{n,\ell}} \otimes \sigma_n)x] \otimes 1_{\mathscr{H}_{n,r}} \ .$$

□

7.10 THEOREM. *Let Φ be a unital completely positive map on $\mathscr{M}$ with a faithful invariant invariant state ρ_0. Then there are density matrices $\{\sigma_1, \dots, \sigma_N\}$, where σ_n acts on $\mathscr{H}_{n,r}$, such that for all whenver $\rho \in \mathcal{S}_\Phi$, there are positive numbers $\{w_1, \dots, w_N\}$ summing to 1 and density matrices $\{\rho_1, \dots, \rho_N\}$, where ρ_n acts on $\mathscr{H}_{n,\ell}$ such that*

$$\rho = \sum_{n=1}^N w_n \rho_n \otimes \sigma_n \ .$$

Proof. Let ρ be any state on $\mathscr{M}$. Then for all y,

$$(E_\Phi)_* \rho(y) = \rho\left(\sum_{n=1}^N \mathrm{Tr}_{\mathscr{H}_{n,r}}[(1_{\mathscr{H}_{n,\ell}} \otimes \sigma_n)p_n y p_n] \otimes 1_{\mathscr{H}_{n,r}}\right) \ .$$

Considering ρ as a density matrix, we may wrte this is

$$\begin{aligned}(E_\Phi)_*\rho(y) &= \sum_{n=1}^N \mathrm{Tr}[\rho \mathrm{Tr}_{\mathscr{H}_{n,r}}[(1_{\mathscr{H}_{n,\ell}} \otimes \sigma_n)p_n y p_n] \otimes 1_{\mathscr{H}_{n,r}}] \\ &= \sum_{n=1}^N \mathrm{Tr}[(\mathrm{Tr}_{\mathscr{H}_{n,r}}[p_n\rho p_n] \otimes \sigma_n)p_n y p_n] \ .\end{aligned}$$

For each $n = 1, \ldots, N$, define $w_n = \mathrm{Tr}[p_n \rho p_n]$. Then if $w_n = 0$, define ρ_n to be an arbitrary density matrix on $\mathscr{H}_{n.\ell}$, and in case $w_n > 0$, define

$$\rho_n = w_n^{-1}\mathrm{Tr}_{\mathscr{H}_{n,r}}[p_n \rho p_n] \ .$$

Then we have $(E_\Phi)_*\rho(y) = \displaystyle\sum_{n=1}^N w_n \mathrm{Tr}[(\rho_n \otimes \sigma_n)y]$, and if $\rho \in \mathcal{S}_\Phi$, $\rho = (E_\Phi)_*\rho$. □

A finite dimensional *quantum copying machine* is a particular sort of completely positive unital map Ψ from $\mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$ to $\mathscr{B}(\mathscr{H}_1)$. The predual Ψ_* of Ψ maps states on $\mathscr{B}(\mathscr{H}_1)$ to states on $\mathscr{B}(\mathscr{H}_1 \otimes \mathscr{H}_2)$.

Associated to Ψ are the two maps Ψ_1 and Ψ_2 given by

$$\Psi_1(a) = \Psi(a \otimes 1_{\mathscr{H}_2}) \quad \text{and} \quad \Psi_2(b) = \Psi(1_{\mathscr{H}_1} \otimes b) \ .$$

For $j = 1, 2$ let $\Psi_{*,j}$ denote the predual of Ψ_j. Then $\Psi_{*,j}$ maps states on $\mathscr{B}(\mathscr{H}_1)$ to states on $\mathscr{B}(\mathscr{H}_j)$.

7.11 THEOREM. *In the notation established above, let $\mathscr{C}$ be the C^* subalgebra of $\mathscr{B}(\mathscr{H}_1)$ consisting of elements that are invariant under Ψ_1. Then for all $b \in \mathscr{B}(\mathscr{H}_2)$,*

$$\Psi_2(b) \in \mathscr{C}' \ .$$

Proof. Let $\Psi(x) = \sum_j W_j^* x W_j$ be the Krauss representation of Ψ. Let $\{\eta_k\}$ and $\{\xi_\ell\}$ be orthonormal bases of $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively. Define $A_\ell : \mathscr{H}_1 \to \mathscr{H}_1 \otimes \mathscr{H}_2$ by $A_\ell \zeta = \zeta \otimes \xi_\ell$, and define $B_k : \mathscr{H}_2 \to \mathscr{H}_1 \otimes \mathscr{H}_2$ by $B_k\zeta = \eta_k \otimes \zeta$.

Then for all $a \in \mathscr{B}(\mathscr{H}_1)$, $a \otimes 1_{\mathscr{H}_2} = \sum_\ell A_\ell^* a A_\ell$ and for all $b \in \mathscr{B}(\mathscr{H}_2)$, $1_{\mathscr{H}_1} \otimes b = \sum_k B_k^* a B_k$. Therefore, $\Psi_1(a) = \sum_{j,\ell} W_j^* A_\ell^* a A_\ell W_j$, and then by Lemma 7.6,

$$\mathscr{C} = \{A_\ell W_j\}'.$$

Hence, for $a \in \mathscr{C}$,

$$a A_\ell W_j = A_\ell W_j a \quad \text{and} \quad a W_j^* A_\ell^* = W_j^* A_\ell^* a \ .$$

Therefore, for all $a \in \mathscr{C}$ and all $b \in \mathscr{B}(\mathscr{H}_2)$,

$$
\begin{aligned}
\Psi(a \otimes b) &= \Psi((a \otimes 1_{\mathscr{H}_2})(1_{\mathscr{H}_1} \otimes b)) \\
&= \sum_\ell \Psi((A_\ell^* a A_\ell(1_{\mathscr{H}_1} \otimes b)) \\
&= \sum_{j,\ell} V_j^* A_\ell^* a A_\ell (1_{\mathscr{H}_1} \otimes b) V_j \\
&= \sum_{j,\ell} a V_j^* A_\ell^* A_\ell (1_{\mathscr{H}_1} \otimes b) V_j \\
&= \sum_j a V_j^* (1_{\mathscr{H}_1} \otimes b) V_j = a \Psi_2(b)
\end{aligned}
$$

Since it is also true that $a \otimes b = (1_{\mathscr{H}_1} \otimes b)(a \otimes 1_{\mathscr{H}_2})$, a similar computation shows that $\Psi(A \otimes b) = \Psi_2(b)a$. Altogether, we have $a\Psi_2(b) = \Psi_2(b)a$. □

Now we specialze to the case in which $\mathscr{H}_1 = \mathscr{H}_2 = \mathscr{H}$, and consider a unital completely postive map Ψ from $\mathscr{B}(\mathscr{H}) \otimes \mathscr{B}(\mathscr{H})$ to $\mathscr{B}(\mathscr{H})$. Then the predual Ψ_* maps states on $\mathscr{B}(\mathscr{H})$ to states on $\mathscr{B}(\mathscr{H}) \otimes \mathscr{B}(\mathscr{H})$. We say that Ψ_* *successfully clones* a state ρ on $\mathscr{B}(\mathscr{H})$ in case

$$
\mathrm{Tr}_{\mathscr{H}_2}[\Psi_*(\rho)] = \rho \quad \text{and} \quad \mathrm{Tr}_{\mathscr{H}_1}[\Psi_*(\rho)] = \rho \ .
$$

This is the same as $\Psi_{*,j}\rho = \rho$ for $j = 1, 2$. For $j = 1, 2$, let $\mathscr{C}_j$ be the fixed point algebra of Ψ_j. By Theorem 7.11, the image of Ψ_2 is contained in $\mathscr{C}_1'$, and hence $\mathscr{C}_2 \subset \mathscr{C}_1'$. Likdewise, $\mathscr{C}_1 \subset \mathscr{C}_2'$. The larger the invariant algebra, the larger the set of invariant states, and hence the set of invariant states is larges when $\mathscr{C}_1 = \mathscr{C}_2'$. But in this case, every state ρ that is invariant under both $\Psi_{1,*}$ and $\Psi_{2,*}$ has the form

$$
\rho = \sum_{n=1}^N w_k \rho_n \otimes \sigma_n
$$

for two fixed sets $\{\rho_1, \dots, \rho_N\}$ and $\{\sigma_1, \dots, \sigma_N\}$ of sensity matrices determined by Ψ_1 and Ψ_2 respectively. This is a comuting set of density matries. Thus, a quantum copying machine is strictly limited in what it can successfully copy.

7.5 Majorization

Let a be a self-adjoint $n \times n$ matrix. Let $(\lambda_1, \dots, \lambda_n)$ be the eigenvalues of a, repeated according to multiplicity, and arranged into a vector in $\mathbb{R}^n$. Let $\{\eta_1, \dots, \eta_n\}$ be an orthonormal basis of $\mathbb{C}^n$, and define the vector $(\alpha_1, \dots, \alpha_n)$ where for each j, $\alpha_j = \langle \eta_j, a\eta_j \rangle$. (This is the *diagonal sequence* of a in the basis $\{\eta_1, \dots, \eta_n\}$; i.e., the sequence of diagonal entries of the matrix representing q in this basis.) For any vector $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$, let x^* be the vector obtained from x be rearranging its entries in non-decreasing order: $x_1^* \geq x_2^* \geq \cdots \geq x_n^*$.

By the variational principle for the eigenvalues of a, for each $k = 1, \dots, n$,

$$
\sum_{j=1}^k \alpha_j^* \leq \sum_{j=1}^k \lambda_j^* \ , \tag{7.13}
$$

and since the traces may be computed in any orthonormal basis,

$$\sum_{j=1}^{n} \alpha_j^* = \sum_{j=1}^{k} \lambda_j^* \ . \tag{7.14}$$

7.12 DEFINITION (Majorization in $\mathbb{R}^n$)**.** Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ be two vectors in $\mathbb{R}^n$ Then λ *weakly majorizes* α in case (7.13) is valid for all $k = 1, \ldots, n$, and in this case we write $\alpha \prec_w \lambda$. We say that λ *majorizes* α in case $\alpha \prec \lambda$ and moreover, (7.14) is satisfied.

Note that the decreasing rearrangement of $-a$ is given by $(-\alpha)^* = (-\alpha_n^*, \cdots -\alpha_1^*)$, and therefore $\alpha \prec \lambda$ if and only if $-\alpha \prec -\lambda$.

It is easy to see that for if p is any *probability vector* in $\mathbb{R}^n$, i.e., a vector in $\mathbb{R}^n$ with non-negative entries that sum to 1, and we define

$$p_{\min} = (1/n, 1/n, \ldots, 1/n) \quad \text{and} \quad p_{\max} = (1, 0, \ldots, 0) \ . \tag{7.15}$$

then

$$p_{\min} \prec p \prec p_{\max} \ . \tag{7.16}$$

As we shall see, majorization provides a useful measure of how "disordered" a probability vector is, and by considering sequences of eigenvalues associated to density matrices, a useful measure of how "disordered" a quantum state is.

Some caution is required in the passage from finite to infinite dimensions. Let $\mathscr{H}$ be a separable Hilbert space, and let a be a compact self adjoint operator on $\mathscr{H}$. Suppose that a has infinitely many positive eigenvalues and but also some negative eigenvalues. If we arrange the eigenvalues in non-increasing order, then all of the negative eigenvalues, along with any that are zero, are pushed infinitely far out in the sequence and are lost. Thus we cannot use rearrangement to define majorization for infinite sequences in which both signs are present.

Let $c_0(\mathbb{N}, \mathbb{R})$ denote the set of real sequences α indexed by $\mathbb{N}$ such that $\lim_{n\to\infty} \alpha_n = 0$. Note that if a is a self adjoint compact operator, then both its diagonal sequence and its eigenvalue sequence belong to c_0. For $J \subset \mathbb{N}$, let $|J|$ denote the cardinality of J. Let $\ell_1(\mathbb{N}, \mathbb{R})$ denote the subspace of $c_0(\mathbb{N}, \mathbb{R})$ consisting of absolutely summable sequences.

7.13 DEFINITION (Majorization in $c_0(\mathbb{N}, \mathbb{R})$)**.** Let $\alpha, \lambda \in c_0(\mathbb{N}, \mathbb{R})$. Then λ *weakly majorizes* α in case for all $k \in \mathbb{N}$

$$\sup \left\{ \sum_{j \in J} \lambda_j \ : \ |J| = k \right\} \geq \sup \left\{ \sum_{j \in J} \alpha_j \ : \ |J| = k \right\} \tag{7.17}$$

and in this case we write $\alpha \prec_w \lambda$.

In case $\alpha, \lambda \in \ell_1(\mathbb{N}, \mathbb{R})$, we say that λ *majorizes* α in case both $\alpha \prec_w \lambda$ and $-\alpha \prec_w -\lambda$ and moreover,

$$\sum_{j=1}^{\infty} \lambda_j = \sum_{j=1}^{\infty} \alpha_j \ . \tag{7.18}$$

In this case we write $\alpha \prec \lambda$.

In contrast to the the finite dimensional case where in the presence of (7.14) either of $\alpha \prec_w \lambda$ or $-\alpha \prec_w -\lambda$ implies the other, this is not the case for infinite sequences. The definition is such that it is still the case that $\alpha \prec \lambda$ if and only if $-\alpha \prec -\lambda$, a basic symmetry that will be referenced frequently.

The suprema in (7.17) need not be maxima. The simplest case is that in which all term in λ have a single sing; e.g., they are all non-negative. Then of course we can form the non-increasing rearrangement λ^*, and

$$\sup\left\{\sum_{j\in J}\lambda_j \ :\ |J| = k\right\} = \sum_{j=1}^{k}\lambda_j^* \ . \tag{7.19}$$

More generally, if we let λ_j^* denote the jth non-negative term in λ and there are m non-negative terms in λ, then (7.19) is valid for all $k \le m$.

However, if there are only m non-negative terms in λ, with $m < \infty$, then for all $k > m$,

$$\sup\left\{\sum_{j\in J}\lambda_j \ :\ |J| = k\right\} = \sum_{j=1}^{\infty}(\lambda_j)_+ \ , \tag{7.20}$$

and the supremum is not obtained: After using all of the non-negative terms, one must start using negative terms, and these can be chosen arbitrarily close to zero sine $\lambda \in c_0(\mathbb{N},\mathbb{R})$, but not equal to zero.

Note that if α has no positive terms, then for all k, $\sup\left\{\sum_{j\in J}\alpha_j \ :\ |J| = k\right\} = 0$. More generally, if α has only m positive terms, then for all $k > m$

$$\sup\left\{\sum_{j\in J}\alpha_j \ :\ |J| = k\right\} = 0 = \sup\left\{\sum_{j\in J}\alpha_j \ :\ |J| = m\right\} = 0 \ ,$$

and thus (7.17) is valid for all k if it is valid for $k \le m$.

The following characterization of weak majorization is essential in what follows.

7.14 THEOREM. *Let $\alpha, \lambda \in c_0(\mathbb{N},\mathbb{R})$. Then $\alpha \prec_w \lambda$ if and only if for all $t > 0$,*

$$\sum_{j=1}^{\infty}(\alpha_j - t)_+ \le \sum_{j=1}^{\infty}(\lambda_j - t)_+ \ . \tag{7.21}$$

Proof. Suppose that $\alpha \prec_w \lambda$. If α is non-positive, then (7.21) is trivially true. If α has positive terms, for any $t > 0$, it can have only finitely many satisfying $\alpha_j \ge t$. Fix $t > 0$, and let k be the cardinality of $\{\ j\ :\alpha_j \ge t\ \}$. Then the identity is (7.19) is applicable and

$$\sum_{j=1}^{\infty}(\alpha_j - t)_+ = \sup\left\{\sum_{j\in J}\alpha_j \ :\ |J| = k\right\} - kt \ . \tag{7.22}$$

Now fix $\epsilon > 0$, and pick $K \subset \mathbb{N}$ with $|K| = k$ such that $\displaystyle\sum_{j\in K}\lambda_j \ge \sup\left\{\sum_{j\in J}\lambda_j \ :\ |J| = k\right\} - \epsilon$. Then

$$\sum_{j=1}^{\infty}(\lambda_j - t)_+ \ge \sum_{j\in K}(\lambda_j - t)_+ \ge \sum_{j\in K}\lambda_j - kt \ge \sup\left\{\sum_{j\in J}\lambda_j \ :\ |J| = k\right\} - \epsilon - kt \ .$$

Combining this with (7.22) and $\alpha \prec_w \lambda$ yields (7.21) for this, and hence all, $t > 0$.

Conversely, suppose that (7.21) is valid for all $t > 0$. Taking the limit $t \downarrow 0$, we see that

$$\sum_{j=1}^{\infty} (\alpha_j)_+ \leq \sum_{j=1}^{\infty} (\lambda_j)_+ \ . \tag{7.23}$$

Suppose that λ has only m positive terms, where m may be finite or infinite. For any $k \leq m$, let t be the kth smallest positive term in λ. Then using the identity (7.19) again,

$$\sum_{j=1}^{\infty} (\lambda_j - t)_+ = \sup \left\{ \sum_{j \in J} \lambda_j \ : \ |J| = k \right\} - tk$$

However, for any set $J \subset \mathbb{N}$ with $|J| = k$, $\sum_{j=1}^{\infty} (\alpha_j - t)_+ \geq \sum_{j \in J} (\alpha_j - t)_+ \geq \sum_{j \in J} \alpha_j - tk$. Hence

$$\sum_{j \in J} \alpha_j \leq \sup \left\{ \sum_{j \in J} \lambda_j \ : \ |J| = k \right\}$$

for all such J and all $k < m$. For $k \geq m$, (if m is finite) $\sup \left\{ \sum_{j \in J} \lambda_j \ : \ |J| = k \right\} = \sum_{j=1}^{\infty} (\lambda_j)_+$, and for all k,

$$\sup \left\{ \sum_{j \in J} \alpha_j \ : \ |J| = k \right\} \leq \sum_{j=1}^{\infty} (\alpha_j)_+ \ .$$

Hence the general conclusion follows from (7.23). □

7.15 COROLLARY. *Let $\alpha, \lambda \in \ell_1(\mathbb{N}, \mathbb{R})$ with $\alpha \prec \lambda$. Then*

$$\sum_{j=1}^{\infty} (\alpha_j)_+ \leq \sum_{j=1}^{\infty} (\lambda_j)_+ \quad \text{and} \quad \sum_{j=1}^{\infty} (-\alpha_j)_+ \leq \sum_{j=1}^{\infty} (-\lambda_j)_+ \ .$$

Proof. This follows from (7.21) in the limit $t \downarrow 0$. □

7.16 LEMMA. *Let $\alpha, \lambda \in \ell_1(\mathbb{N}, \mathbb{R})$ be non-zero and suppose that $\alpha \prec \lambda$, and that all of the entries of λ are non-negative. Then all of the entries of α are non-negative, and there is a sequence $\tilde{\lambda} \in \ell_1(\mathbb{N}, \mathbb{R})$ differing from λ in at most two terms such that*

$$\alpha \prec \tilde{\lambda} \prec \lambda$$

and such that $\max\{\alpha_j\}$ is equal to a term in $\tilde{\lambda}$. More specifically, there exist terms λ_m and λ_n in λ such that $\lambda_m \leq \max\{\alpha_j\} \leq \lambda_n$ and such that with

$$\tilde{\lambda}_m := \max\{\alpha_j\} \quad \text{and} \quad \tilde{\lambda}_n := \lambda_m + \lambda_n - \max\{\alpha_j\} \ ,$$

and $\tilde{\lambda}_j := \lambda_j$ for $j \neq m, n$, then $\alpha \prec \tilde{\lambda} \prec \lambda$.

Proof. Since λ is non-negative and $-\alpha \prec_w -\lambda$, it follows that α is non-negative as well. Consider the set of terms in λ that are greater than or equal to $\max\{\alpha_j\}$. This set in finite since $\lambda \in \ell_1(\mathbb{N}, \mathbb{R})$ and non-empty since $\alpha \prec \lambda$. Let m be its cardinality. Let $\{\lambda_{n_1}, \ldots, \lambda_{n_m}\}$ be this set arranged in non-increasing order.

Let $\lambda_{n_{m+1}}$ be a term in λ with the next largest value which is necessarily non-negative but less than $\max\{\alpha_j\}$. We then have

$$\lambda_{n_1} \geq \cdots \geq \lambda_{n_m} \geq \max\{\alpha_j\} > \lambda_{n_{m+1}} \ .$$

If $\lambda_{n_m} = \max\{\alpha_j\}$ we are already done, so suppose that this is not the case. Note that all terms in λ that are not included in $\{\lambda_{n_1}, \ldots, \lambda_{n_{m+1}}\}$ have a value that is no greater than $\lambda_{n_{m+1}}$

Now define

$$\begin{aligned} \tilde{\lambda}_{n_m} &= \max\{\max\{\alpha_j\} \ , \ \lambda_{n_m} + \lambda_{n_{m+1}} - \max\{\alpha_j\}\} \\ \tilde{\lambda}_{n_{m+1}} &= \min\{\max\{\alpha_j\} \ , \ \lambda_{n_m} + \lambda_{n_{m+1}} - \max\{\alpha_j\}\} \ , \end{aligned} \tag{7.24}$$

and for all $j \neq k, k+1$, define $\tilde{\lambda}_j = \lambda_j$. Note that $\sum_{j=1}^{\infty} \tilde{\lambda}_j = \sum_{j=1}^{\infty} \lambda_j = \sum_{j=1}^{\infty} \alpha_j$.

Observe that for $k < m$, or $k = m+1$,

$$\sup\left\{\sum_{j\in J} \lambda_j : |J| = k\right\} = \sum_{j=1}^{k} \lambda_{n_j} = \sum_{j=1}^{k} \tilde{\lambda}_{n_j} = \sup\left\{\sum_{j\in J} \tilde{\lambda}_j : |J| = k\right\} \ , \tag{7.25}$$

and for all $k > m$, $\sup\left\{\sum_{j\in J} \lambda_j : |J| = k\right\} = \sup\left\{\sum_{j\in J} \tilde{\lambda}_j : |J| = k\right\}$. This shows that that for all k expect possibly $k = m$.

$$\sup\left\{\sum_{j\in J} \alpha_j : |J| = k\right\} \leq \sup\left\{\sum_{j\in J} \tilde{\lambda}_j : |J| = k\right\} \ .$$

To deal with the final case,

$$\begin{aligned} \sup\left\{\sum_{j\in J} \alpha_j : |J| = m\right\} &\leq \sup\left\{\sum_{j\in J} \alpha_j : |J| = m-1\right\} + \alpha_1 \\ &\leq \sup\left\{\sum_{j\in J} \lambda_j : |J| = m-1\right\} + \alpha_1 = \sup\left\{\sum_{j\in J} \tilde{\lambda}_j : |J| = m-1\right\} + \alpha_1 \\ &= \sum_{j=1}^{m-1} \tilde{\lambda}_{n_j} + \alpha_1 \leq \sum_{j=1}^{m} \tilde{\lambda}_{n_j} \ . \end{aligned} \tag{7.26}$$

This proves that $\alpha \prec_w \lambda$. Since none of the negative terms of λ have been changed, the fact that $-\alpha \prec_w -\tilde{\lambda}$ follows directly from $-\alpha \prec_w -\lambda$.

To see that $\tilde{\lambda} \prec \lambda$, note that

$$\{\tilde{\lambda}_{m_n}, \tilde{\lambda}_{n_{m+1}}, 0, 0 \ldots\} \prec \{\lambda_{m_n}, \lambda_{n_{m+1}}, 0, 0 \ldots\}$$

and then by Theorem 7.14, for all $t > 0$,

$$(\tilde{\lambda}_{m_n} - t)_+ + (\tilde{\lambda}_{n_{m+1}} - t)_+ \leq (\lambda_{m_n} - t)_+ + (\lambda_{n_{m+1}} - t)_+ \ .$$

Since all other terms in the two sequences are equal,

$$\sum_{j=1}^{\infty}(\tilde{\lambda}_j - t)_+ \leq \sum_{j=1}^{\infty}(\lambda_j - t)_+$$

for all $t > 0$, and then $\tilde{\lambda} \prec \lambda$ by Theorem 7.14 once more. □

7.17 THEOREM (Horn-Gohberg-Marcus Theorem). *Let h be a positive semidefinite trace class operator on a separable Hilbert space $\mathscr{H}$. Let $\{\eta_j\}_{j\in\mathbb{N}}$ be any orthonormal basis of $\mathscr{H}$. Define $\lambda \in \ell_1(\mathbb{N}, \mathbb{R})$ by*

$$\lambda_j = \langle \eta_j, h\eta_j \rangle_{\mathscr{H}} \qquad j \in \mathbb{N} \ . \tag{7.27}$$

Let $\alpha \in \ell_1(\mathbb{N}, \mathbb{R})$ be non-negative and non-increasing and such that $\alpha \prec \lambda$.

Then there exists a closed subspace $\mathscr{H}_1 \subset \mathscr{H}$ and an orthonormal basis $\{\zeta_j\}$ for $\mathscr{H}_1$ such that

$$\alpha_j = \langle \zeta_j, h\zeta_j \rangle_{\mathscr{H}} \quad \text{for all } j \in \mathbb{N} \ . \tag{7.28}$$

If $\mathscr{H}_1$ is a proper subspace of $\mathscr{H}$, its orthogonal complement lies in $\ker(h)$.

Proof. We build the orthonormal basis recursively using a method of Gohberg and Marcus. Since α is non-increasing, $\max\{\alpha_j\} = \alpha_1$. By Lemma 7.16 there exist terms λ_m and λ_n in λ such that $\lambda_m \leq \alpha_1 \leq \lambda_n$ and such that with

$$\tilde{\lambda}_m := \alpha_1 \quad \text{and} \quad \tilde{\lambda}_n := \lambda_m + \lambda_n - \max\{\alpha_j\} \ ,$$

and $\tilde{\lambda}_j := \lambda_j$ for $j \neq m, n$, then $\alpha \prec \tilde{\lambda}$.

Consider the parameterized unit circle $\eta(t) := \cos t\eta_n + \sin t\eta_m$ in the real span of $\{\eta_m, \eta_n\}$. Since $\langle \eta(t), h\eta(t) \rangle_{\mathscr{H}}$ is equal to λ_n for $t = 0$ and to λ_m for $t = \pi/2$, continuity ensures that there is some t_0 such that $\langle \eta(t), h\eta(t_0) \rangle_{\mathscr{H}} = \max\{\alpha_j\}$. Define $\zeta_1 = \eta(t_0)$, and $\xi_1 = \eta(t_0 + \pi/2)$. Then $\{\zeta_1, \xi_1\}$ is orthonormal with the same span as $\{\eta_m, \eta_m\}$. Replace $\mathscr{H}$ with the orthogonal complement of ζ_1, and now let $\{\eta_j\}$ denote the the orthonormal basis for this space in which ξ_1 is the first element, and the rest is given by the sequence $\{\eta_j\}$ with the terms η_m and η_n deleted.

Replace α with the sequence obtained by deleting α_1 and shifting the other terms up. Replace λ by the sequence $\tilde{\lambda}_j = \langle \eta_j, h\eta_j \rangle_{\mathscr{H}}$ using the new orthonormal basis. By Lemma 7.16, the new α and λ again satisfy $\alpha \prec \lambda$.

We may now repeat the procedure to find, in the orthogonal complement of ζ_1, a unit vector ζ_2 such that $\langle \zeta_2, h\zeta_2 \rangle_{\mathscr{H}} = \alpha_2$, and again get an orthonormal set of vectors such that the corresponding sequence of diagonal entries of h majorizes the sequence αm now with α_1 and α_2 split off, and the other terms moved up.

Evidently, this operation may be repeated indefinitely, thus producing the orthonormal set $\{\zeta_j\}$ such that (7.28) is satisfied. Let $\mathscr{H}_1$ be the span of the orthonormal set $\{\zeta_j\}$. If $\mathscr{H}_1 = \mathscr{H}$ we are done. It remains to show that if $\mathscr{H}_1$ is a proper subspace of $\mathscr{H}$, its orthogonal complement lies in $\ker(h)$.

Let $\{\xi_k\}_{k\in K}$ be an orthonormal basis for $\mathscr{H}_2 := \mathscr{H}_1^{\perp}$. Define the compact positive operator a by

$$\begin{aligned} a &= \sum_{j=1}^{\infty}\langle\zeta_j, h\zeta_j\rangle_{\mathscr{H}}|\zeta_j\rangle\langle\zeta_j| + \sum_{k\in K}\langle\xi_k, h\xi_k\rangle_{\mathscr{H}}|\xi_k\rangle\langle\xi_k| \\ &= \sum_{j=1}^{\infty}\alpha_j|\zeta_j\rangle\langle\zeta_j| + \sum_{k\in K}\langle\xi_k, h\xi_k\rangle_{\mathscr{H}}|\xi_k\rangle\langle\xi_k| \end{aligned}$$

Since a has the same diagonal elements as h in the orthonormal basis $\{\zeta_j\}\cup\{\xi_k\}$, and since we may compute the trace in any orthonormal basis we choose,

$$\mathrm{Tr}[h] = \mathrm{Tr}[a] = \sum_{j=1}^{\infty}\alpha_j + \sum_{k\in K}\langle\xi_k, h\xi_k\rangle_{\mathscr{H}} \ .$$

Computing the trace of h in the original $\{\eta_j\}$ basis, using using the fact that $\alpha \prec \lambda$, we find

$$\mathrm{Tr}[h] = \sum_{j=1}^{\infty}\lambda_j = \sum_{j=1}^{\infty}\alpha_j \ .$$

It follows that $\langle\xi_k, h\xi_k\rangle_{\mathscr{H}} = 0$ for all $k\in K$, and then since h is positive, this means that $h\xi_k = 0$ for all $k\in K$. Thus, h vanishes on $\mathscr{H}_2$. □

7.18 COROLLARY. *Let $\mathscr{H}$ be a separable Hilbert space, and let h be a positive semi-definite trace class operator on $\mathscr{H}$. Let $\lambda\in\ell_1(\mathbb{N},\mathbb{R})$ be the sequence of eigenvalues of h repeated according to multiplicity and arranged in non-increasing order. Let $\alpha\in\ell_1(\mathbb{N},\mathbb{R})$ be arranged in non-increasing order. Then an necessary and sufficient condition for there to exists an orthonormal basis $\{\zeta_j\}$ of $\mathscr{H}$ and a unitary u on $\mathscr{H}$ such that for all $j\in\mathbb{N}$,*

$$\alpha_j = \langle\zeta_j, (u^*hu)\zeta_j\rangle_{\mathscr{H}} \quad \text{and} \quad \lambda_j = \langle\zeta_j, h\zeta_j\rangle_{\mathscr{H}} \ . \tag{7.29}$$

is that $\alpha \prec \lambda$.

Proof. Because u^*hu and h have the same eigenvalues with the same multiplicities, when (7.29) and $\{\eta_j\}$ is an orthonormal basis of $\mathscr{H}$ consisting of eigenvectors of h, then $\alpha\prec\lambda$ by the variational principle for sums of eigenvalues. In the other direction, let $\{eta_j\}$ and $\{\zeta_j\}$ be the two bases constructed in the theorem, and let u be the unitary such that $u\eta_j = \zeta_j$ for all j. □

7.19 COROLLARY. *A necessary and sufficient condition on two non-negative and non-increasing sequences α,λ in $\ell_1(\mathbb{N},\mathbb{R})$ for $\alpha\prec\lambda$ is that there exist non-negative numbers*

$$\{w_{j,k} \ : \ j,k\in\mathbb{N}\}$$

such that

$$\sum_{j=1}^{\infty} w_{j,k} = 1 \text{ for all } k \quad \text{and} \quad \sum_{k=1}^{\infty} w_{j,k} = 1 \text{ for all } j \ . \tag{7.30}$$

and

$$\alpha_j = \sum_{k=1}^{\infty} w_{j,k}\lambda_k \quad \text{for all } j \ . \tag{7.31}$$

Proof. Suppose $\alpha \prec \lambda$. Let $\{\eta_j\}$ be any orthonormal basis for $\mathscr{H}$, and define the operator

$$h = \sum_{j=1}^{\infty} \lambda_j |\eta_j\rangle\langle\eta_j| \ .$$

Then λ is the eigenvalue sequence of h. By Corollary 7.18, there to exists an orthonormal basis $\{\zeta_j\}$ of $\mathscr{H}$ and a unitary u on $\mathscr{H}$ such that for all $j \in \mathbb{N}$, such that for each j, $h = \sum_{\ell=1}^{\infty} \lambda_m |\zeta_m\rangle\langle\zeta_m|$ and

$$\alpha_j = \langle \zeta_j, u^* h u \zeta_j\rangle_{\mathscr{H}} = \sum_{\ell=1}^{\infty} \lambda_\ell \langle \zeta_j u^*, \zeta_\ell\rangle_{\mathscr{H}} \langle \zeta_\ell, u\zeta_j\rangle_{\mathscr{H}} = \sum_{\ell=1}^{\infty} \lambda_\ell |\langle \zeta_\ell, u\zeta_j\rangle_{\mathscr{H}}|^2 \ .$$

Define $w_{j,\ell} = |\langle \zeta_\ell, u\zeta_j\rangle_{\mathscr{H}}|^2$. Note that since u is unitary, both (7.30) and (7.31) are satisfied.

Conversely, suppose that both (7.30) and (7.31) are satisfied. Then for each $k \in \mathbb{N}$,

$$\begin{aligned}
\sum_{j=1}^{k} \alpha_j &= \sum_{j=1}^{k} \left(\sum_{\ell=1}^{k} w_{j,\ell}\lambda_\ell + \sum_{\ell=k+1}^{\infty} w_{j,\ell}\lambda_\ell \right) \\
&\geq \sum_{\ell=1}^{k} \left(\sum_{j=1}^{k} w_{j,\ell} \right) \lambda_\ell + \sum_{j=1}^{k} \left(\sum_{\ell=k+1}^{\infty} w_{j,\ell} \right) \lambda_{k+1} \\
&= \sum_{\ell=1}^{k} \left(1 - \sum_{j=k+1}^{\infty} w_{j,\ell} \right) \lambda_\ell + \sum_{j=1}^{k} \left(\sum_{\ell=k+1}^{\infty} w_{j,\ell} \right) \lambda_{k+1} \\
&= \sum_{\ell=1}^{k} \lambda_\ell + \sum_{j=1}^{k} \left(\sum_{\ell=k+1}^{\infty} w_{j,\ell} \right) (\lambda_\ell - \lambda_{k+1}) \geq \sum_{\ell=1}^{k} \lambda_\ell \ .
\end{aligned}$$

This proves that $\alpha \prec_w \lambda$, and $-\alpha \prec_w -\lambda$ is trivially true. It follows immediately from (7.30) and (7.31) that $\sum_{j=1}^{\infty} \alpha_j = \sum_{j=1}^{\infty} \lambda_j$, and thus $\alpha \prec \lambda$. □

7.20 THEOREM. *Let f be a concave function on $[0,\infty)$. Let $\alpha, \lambda \in \ell_1(\mathbb{N},\mathbb{R})$ be non-negative and non-increasing. If $\alpha \prec \lambda$, then*

$$\sum_{j=1}^{\infty} f(\alpha_j) \geq \sum_{j=1}^{\infty} f(\lambda_j) \ . \tag{7.32}$$

If furthermore f is strictly concave, then there is equality in (7.32) is and only if $\alpha_j = \lambda_j$ for all j.

Finally, if f is not only concave, but also nondecreasing, then for all $k \in \mathbb{N}$,

$$\sum_{j=k}^{\infty} f(\alpha_j) \geq \sum_{j=k}^{\infty} f(\lambda_j) \ . \tag{7.33}$$

Proof. Note first of all that since f is concave on $[0,\infty)$, f has a definite sign $(0,\delta)$ for some $\delta > 0$, and hence all but at most finitely many terms in the sequences $\{f(\alpha_j)\}$ and $\{f(\lambda_j)\}$ have the same sign, and this both sums are well defined.

By Corollary 7.19, there exist non-negative numbers

$$\{w_{j,k} \ : \ j,k \in \mathbb{N}\}$$

such that (7.30) and (7.31) are satisfied.

Then for each j

$$f(\alpha_j) = f\left(\sum_{\ell=1}^{\infty} w_{j,\ell}\lambda_\ell\right) \geq \sum_{\ell=1}^{\infty} w_{j,\ell} f(\lambda_\ell) \ . \tag{7.34}$$

Therefore,

$$\sum_{j=1}^{\infty} f(\alpha_j) \geq \sum_{j,\ell=1}^{\infty} w_{j,\ell} f(\lambda_\ell) = \sum_{\ell=1}^{\infty} f(\lambda_\ell) \ .$$

If there is equality, then there must be equality in (**??**) for each j. In case f is strictly convex, this means that for each j, there exists j_ℓ such that $w_{j,\ell} = 1$ for $j = j_\ell$ and $w_{j,\ell} = 0$ for $j \neq j_\ell$. Since α and λ are both non increasing, this means that $\alpha = \lambda$.

Finally suppose that f is non-decreasing as well as concave. For each $k \in \mathbb{N}$ define

$$g(x) = \min\{\ f(x)\ ,\ f(\lambda_k)\ \} \ .$$

Then g is concave, and by the first part.$\sum_{j=1}^{\infty} g(\alpha_j) \geq \sum_{j=1}^{\infty} g(\lambda_j)$. Therefore, since for $j < k$, $g(\lambda_j) = g(\lambda_{k-1}) \geq g(x)$ for all x, and since for $j \geq k$, $g(\lambda_j) = f(\lambda_j)$,

$$\sum_{j=k}^{\infty} g(\alpha_j) = \sum_{j=k}^{\infty} g(\lambda_j) + \left(\sum_{j=1}^{k-1}[g(\lambda_{k-1}) - g(\alpha_j)]\right) \geq \sum_{j=1K}^{\infty} f(\lambda_j) \ .$$

Then since $g(x) \geq f(x)$ for all x, (7.33) is proved. □

7.6 Some open problems discussed in class

The Lindblad No-Cloning Theorem [15] is only proved in a finite dimensional setting. A crucial part of the argument was a structure theorem for algebras of observables that are left invariant by a quantum operation, the associated conditional expectation, and the associated set of invariant states. Conditional expectations are studied in Arveson [2]; this paper discussions some aspects of the construction in a more general setting.

Lindblad's argument is central to another focus of current research activity that grew out of fundamental work of Lieb and Ruskai [10, 11, 12] on *Strong subadditivity of quantum entropy*, as discussed in class. As we have seen with Kadison's inequality, knowing the cases of equality can also be important. This is certainly the case with strong subadditivity. The cases of equality *in the finite dimensional case* have been determined in a paper of Hayden, Jozsa, Petz and Winter, [7], and the connection with Lindblad's No-Cloning Theorem is made explicit. Again, it would be good to go beyond the finite dimensional case. A new and interesting approach to these inequalities can be found in the very recent paper of Sutter, Berta and Tomamichael [23].

With regard to Lin's Theorem discussed earlier in the course, there is a very interesting recent paper of Ogata [18] who proves an old conjecture of von Neumann, discussed in class, concerning approximation of almost commuting operators. This is non-quantitative and non-constructive. The very recent quantitative version of Lin's Theorem due to Kachkovskiy and Safarov [8] may provide some clues as to how to prove a quantitative and constructive version.

Also, Neilsen's Theorem [16] which characterizes the pure states that can be reached form a given pure state using only LOCC operations is a beautiful application of the theory of majorization, but there is no such result yet for mixed states.

The proceedings volume *Entropy and the Quantum* by Simms and Ueltschi [22] contains a number of introductory lectures to open problems in the area. The lectures in this volume by Luc Rey-Belet on quantum large deviations are particularly interesting in terms of the challenges posed. There are many, many other interesting open problems, but this is a selection of problems and references discussing problems that might be particularly timely.

References

[1] H. Araki and M.M. Yanase, *Measurement of quantum mechanical operators*, Phys. Rev. **120** (1960), 622-626.

[2] W.B. Arveson, *Analyticity in operator algebras*, Amer. J. Math. **89** (1967), 578-642.

[3] M. Born: *Zur Quantenmechanik der Stoṿorgänge* Zeitschrift fr Physik, **37**, no. 12 1926, pp. 863-867 (German); English translation, *On the quantum mechanics of collisions, in Quantum theory and measurement*, J. A. Wheeler and W. H. Zurek, eds., Princeton, New Jersey: Princeton University Press, 1983

[4] E.B. Davies, *Quantum theory of open systems.* London, Academic Press, 1976

[5] E.B. Davies and J.T. Lewis, *An operational approach to quantum probability.* Commun. Math. Phys. **17** (1970), 239-260.

[6] P. Friis and M. Rordam: *Almost commuting self-adjoint matrices – a short proof of Huaxin Lin's theorem.* J. Reine Angew. Math. **479**, 121 (1996)

[7] P. Hayden, R. Jozsa, D. Petz, A. Winter, *Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality*, Comm. Math. Physics **246**, no. 2 359-374 (2004)

[8] I. Kachkovskiy and Y. Safarov, *Distance to normal elements in* C^**-algebras of real rank zero*, J. Amer. Math. Soc. **29** 61-80 (2016)

[9] K. Kraus, *General state changes in quantum theory.* Ann. Phys. **64** (1971), 311-335.

[10] E.H. Lieb: *Convex trace functions and the Wigner-Yanase-Dyson Conjecture*, Adv. Math. **11** 267-288 (1973)

[11] E.H. Lieb: *Some convexity and subadditivity properties of entropy*, Bull. Amer. Math. Soc. **81**, 1975, pp. 1-13.

[12] E.H. Lieb and M.B. Ruskai: *Proof of the strong subadditivity of quantum-mechanical entropy*, J. Math. Phys. **14**, 1973, pp. 1938-1941.

[13] E.H. Lieb and M.B. Ruskai: *Some operator inequalities of the Schwarz type*, Adv. in Math. **12**, 269-273 (1974).

[14] H. Lin: *Almost commuting self-adjoint matrices and applications.* Fields. Inst. Commun. **13**, 193 (1995)

[15] G. Lindblad *The general no-cloning theorem* Lett. Math. Phys., **47**, no. 2, 189-196 (1999)

[16] M. Nielsen, *Conditions for a Class of Entanglement Transformations*, Phys. Rev. Lett., **83** no. 2, 436-439, 1999.

[17] J. Newburgh: *The variation of spectra.* Duke Math. Journal. **18**, 165-176 (1951)

[18] Y. Ogata, *Approximating macroscopic observables in quantum spin systems with commuting matrices*, Journal of Functional Analysis **264**, no. 6, 2005-2033 (2013)

[19] M. Ozawa, *Quantum measuring processes of continuous observables*, J. Math. Phys. **25** (1984), 79-87.

[20] M. Ozawa, *Conditional Probability and A Posteriori States in Quantum Mechanics*, Publ. RIMS, Kyoto Univ. **21** (1985), 279-295

[21] L. Schwartz, *Lectures on disintegration of measures*. Bombay, Tata Institute of Fundamental Research, 1975.

[22] R. Sims and D. Ueltschi, eds., *Entropy and the quantun*, Contemporary Mathematics, A.M.S., Providence RI. 2009.

[23] D. Sutter, M. Berta, M. Tomamichael, Multivariate Trace Inequaities, arXiv preprint 2016.

[24] J. von Neumann, *Mathematische Begrndung der Quantenmechanik*, Nachr. Ges. Wiss. Göttingen (1927) pp. 1-57.

[25] J. von Neumann, *Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik*, Nachr. Ges. Wiss. Göttingen (1927) pp. 245-272.

[26] J. von Neumann *Mathematische Grundlagen der Quantenmechanik*, Berlin, J. Springer, 1932.

[27] E.P. Wigner, *Die Messung Quantenmechanischer Operatoren*, Z. Phys. **133** (1952), 101-108.