E/Ea Thompson (They/Them)

# Math 617: Functional Analysis

## – In Pursuit of Abstract Nonsence –

Monday $6^{th}$ February, 2023

Springer Nature

# Preface

This is a collection of notes associated with Math 617 (Functional Analysis) taken at the University of Calgary.

University of Calgary,

*E/Ea Thompson (They/Them)*
Monday 6th February, 2023

# Contents

# Notation

List of common notations used in these notes.

| | |
|---|---|
| $\mathbb{N}$ | Natural numbers |
| $\mathbb{Z}$ | Integers |
| $\mathbb{Q}$ | Rational numbers |
| $\mathbb{R}$ | Real numbers |
| $\mathbb{C}$ | Complex numbers |
| $\mathbb{F}$ | Field of real or complex numbers |
| $\langle \cdot, \cdot \rangle$ | Inner product |

# Chapter 1

# Hilbert Spaces

**Abstract** Summary of material in chapter (to be completed after chapter)

## 1.1 Elementary Properties

**Definition 1.1.1 (Semi-inner Product)** If $V$ is a vector space over $\mathbb{F}$, a **semi-inner product** on $V$ is a mapping $u : V \times V \to F$ such that for all $x, y, z \in V$ and $\alpha, \beta \in F$

(a) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$ (linearity in first component)
(b) $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$ (conjugate-linearity in second component)
(c) $u(x, x) \geq 0$ (real non-negativity)
(d) $u(x, y) = \overline{u(y, x)}$ (conjugate-symmetry)

We note a simple properties of semi-inner products; $u(x, 0) = u(0, y) = 0$ for any $x, y \in V$.

An **inner product** on $V$ is a semi-inner product that is also positive. That is it satisfies the following

(e) If $u(x, x) = 0$, then $x = 0$.

As indicated in the notation, we use $\langle x, y \rangle$ to denote an inner product.

Example:
Let $S = \{\alpha : \mathbb{N} \to \mathbb{F} : \operatorname{supp} \alpha \text{ is finite}\}$. If addition and scalar multiplication are defined component-wise then $S$ is a vector space over $\mathbb{F}$. If we define a map

$$\langle \alpha, \beta \rangle = \sum_{n=1}^{\infty} \alpha_{2n} \overline{\beta}_{2n}$$

then $\langle \cdot, \cdot \rangle$ is a semi-inner product that is not an inner product. Note there are no convergence issues in the sum on the right since only finitely many terms are non-zero. Further, if $\alpha : \mathbb{N} \to \mathbb{F}$ is given by $\alpha_1 = 1$ and $\alpha_n = 0$ for $n > 1$, then $\langle \alpha, \alpha \rangle = 0$. On the other hand,

$$\langle \alpha, \beta \rangle = \sum_{n=1}^{\infty} \alpha_n \overline{\beta}_n$$

$$\langle \alpha, \beta \rangle = \sum_{n=1}^{\infty} \frac{1}{n} \alpha_n \overline{\beta}_n$$

$$\langle \alpha, \beta \rangle = \sum_{n=1}^{\infty} n^5 \alpha_n \overline{\beta}_n$$

are all inner products on $V$.

Example:

Let $(X, \Omega, \mu)$ be a measure space consisting of a set $X$, a $\sigma$-algebra $\Omega$, and a countably additive measure $\mu$ defined on $\Omega$ with values in the non-negative extended reals. If $f, g \in L^2(\mu)$, then Hölder's inequality implies $f\overline{g} \in L^1(\mu)$, with

$$\int_X f\overline{g} d\mu \leq ||f||_{L^2} \, ||\overline{g}||_{L^2} < \infty$$

Taking the left hand side as the value of $\langle f, g \rangle$, this defines an inner product on $L^2(\mu)$.

Note that for positivity we must be in the quotient space $L^2(\mu) = \mathcal{L}^2(\mu)/\mathcal{N}$, where $\mathcal{N}$ is the subspace of $\mathcal{L}^2(\mu)$ functions on $X$ which are 0 almost everywhere with respect to $\mu$.

Hölder's inequality in the special case used in this example is a general inequality satisfied by semi-inner products:

**Theorem 1.1.2 (Cauchy-Bunyakowsky-Schwarz Inequality)** If $\langle \cdot, \cdot \rangle$ is a semi-inner product on $V$, then

$$| \langle x, y \rangle |^2 \leq \langle x, x \rangle \langle y, y \rangle$$

for any $x, y \in V$. Moreover, equality occurs if and only if there are scalars $\alpha, \beta \in F^{\times}$, such that $\langle \beta x + \alpha y, \beta x + \alpha y, = \rangle\, 0$.

***Proof*** If $\alpha \in \mathbb{F}$ and $x, y \in V$, then

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

Suppose $\langle y, x \rangle = be^{i\theta}, b \geq 0$, and let $\alpha = e^{-i\theta}t$, $t \in \mathbb{R}$. The above inequality becomes

$$0 \leq \langle x, x \rangle - 2tb + t^2 \langle y, y \rangle = c - 2bt + at^2 =: q(t)$$

where $c = \langle x, x \rangle$ and $a = \langle y, y \rangle$. $q(t)$ is a quadratic polynomial in the real variable $t$, and $q(t) \geq 0$ for all $t$. This implies that $q(t) = 0$ has at most one real solution $t$. From the quadratic formula it follows that $4b^2 - 4ac \leq 0$. Hence

$$0 \geq b^2 - ac = |\langle x, y\rangle|^2 - \langle x, x\rangle \langle y, y\rangle$$

proving the inequality.
First we show sufficiency of equality. Hence suppose $\alpha, \beta \in \mathbb{F}^\times$ such that $\langle \beta x + \alpha y, \beta x + \alpha y\rangle = 0$. As $\beta \neq 0$, we may without loss of generality replace $\beta = 1$ (multiply both sides by $1/|\beta|^2$), so $\langle x + \alpha y, x + \alpha y\rangle = 0$. Let $\alpha = te^{i\varphi}$, $t > 0$, and $\langle y, x\rangle = be^{i\theta}$, as before. Then

$$0 = \langle x, x\rangle + te^{i(\varphi+\theta)}b + te^{-i(\varphi+\theta)}b + t^2 \langle y, y\rangle = \langle x, x\rangle + 2t\cos(\varphi + \theta)b + t^2 \langle y, y\rangle$$

From before the discriminant of this polynomial in $t$ must be $\leq 0$. As this expression implies it the polynomial has a real root, we must have that the discriminant $= 0$, so

$$4b^2\cos^2(\varphi + \theta) - 4\langle x, x\rangle \langle y, y\rangle = 0$$

But $\cos^2(\varphi + \theta) \leq 1$, so

$$b^2 \geq \langle x, x\rangle \langle y, y\rangle \geq |\langle x, y\rangle|^2$$

so we have equality.
To show necessity suppose $|\langle x, y\rangle|^2 = \langle x, x\rangle \langle y, y\rangle$. Then in particular $|\langle \beta x, \alpha y\rangle|^2 = \langle \beta x, \beta x\rangle \langle \alpha y, \alpha y\rangle$ for any $\alpha, \beta \in \mathbb{F}$. If both $\langle x, x\rangle = \langle y, y\rangle = 0$, we may take $\alpha = \beta = 1$. Otherwise, suppose without loss of generality that $\langle y, y\rangle \neq 0$. Take $\beta = 1$ and $\alpha = -te^{-i\theta}$, $t \in \mathbb{R}$, as before, so

$$\langle \beta x + \alpha y, \beta x + \alpha y\rangle = \langle x, x\rangle - 2bt + t^2 \langle y, y\rangle$$

Since $\langle y, y\rangle \neq 0$, we may set $t = \sqrt{\frac{\langle x,x\rangle}{\langle y,y\rangle}}$. Then the equation becomes

$$\langle \beta x + \alpha y, \beta x + \alpha y\rangle = \langle x, x\rangle - 2\sqrt{\langle x, x\rangle \langle y, y\rangle}\sqrt{\frac{\langle x, x\rangle}{\langle y, y\rangle}} + \frac{\langle x, x\rangle}{\langle y, y\rangle}\langle y, y\rangle = 0$$

completing the proof. □

This theorem provides us with the triangle inequality.

**Corollary 1.1.3** If $\langle \cdot, \cdot\rangle$ is a semi-inner product on $V$, and $||x|| := \sqrt{\langle x, x\rangle}$, then for all $x, y \in V$ and $\alpha \in \mathbb{F}$,

(a) $||x + y|| \leq ||x|| + ||y||$ (triangle inequality)
(b) $||\alpha x|| = |\alpha|\,||x||$ (absolute homegeneity)

If $\langle \cdot, \cdot\rangle$ is an inner product, then

(c) $||x|| = 0$ implies $x = 0$

***Proof*** To see (a), let $x, y \in V$. Then

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

$$\leq ||x||^2 + 2|\langle x, y\rangle| + ||y||^2$$
$$\leq ||x||^2 + 2\,||x||\,||y|| + ||y||^2 \qquad \text{(by CBS 1.1)}$$
$$= (||x|| + ||y||)^2$$

The inequality follows by taking square roots. (b) and (c) are immediate from definitions. □

The identity

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

in the previous proof is called the **polar identity**.

The quantity $||x|| = \sqrt{\langle x, x\rangle}$ is the special case of a mapping called a **norm**. Norms induce metrics on a space, through the formula $d(x, y) = ||x - y||$, giving a metric space structure to $V$. Note that, essentially by construction, with this metric and its induced topology, the inner rpoduct becomes a continuous mapping from $V \times V$ to $\mathbb{F}$.

**Definition 1.1.4 (Hilbert Space)** A **Hilbert space** is a vector space $\mathscr{H}$ over $\mathbb{F}$ together with an inner product $\langle \cdot, \cdot\rangle$ such that relative to the metric $d(x, y) = ||x - y||$ induced by the norm, $\mathscr{H}$ is a complete metric space.

$\mathscr{H} = L^2(\mu)$ with the standard inner product is an example of a Hilbert space, as is $\mathbb{F}^d$ with its usual inner product.

Example:

Let $I$ be any set and let $l^2(I)$ denote the set of functions $x : I \to \mathbb{F}$ with countable suppose and $\sum_{i\in I} |x(i)|^2 < \infty$. For $x, y \in l^2(I)$, define

$$\langle x, y\rangle = \sum_i x(i)\overline{y(i)}$$

Then $l^2(I)$ is a Hilbert space.

In fact, $l^2(I)$ is a special case of $L^2(\mu)$ with the counting measure $\mu$ on $I$.

We recall a notion from analysis: A function $f : I \to \mathbb{R}$, where $I$ is an interval in $\mathbb{R}$, is said to be **absolutely continuous** on $I$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that whenever a finite sequence of pairwise disjoint sub-intervals $(x_k, y_k)$ of $I$ satisfies

$$\sum_k (y_k - x_k) < \delta$$

then

$$\sum_k |f(y_k) - f(x_k)| < \epsilon$$

Note this is stronger than uniform continuity. We have the following characterization of such functions.

**Theorem 1.1.5** A function $F : [a,b] \to \mathbb{R}$ is **absolutely continuous** if and only if there is a function $f : [a,b] \to \mathbb{R}$ such that $f$ is Lebesgue measurable and integrable with respect to $m$ (the Lebesgue measure on the line) and such that

$$F(x) = F(a) + \int_{[a,x]} f\,dx, \forall a \leq x \leq b$$

***Proof*** First consider the if part. Suppose the equation holds. Since $\int_{[a,b]} |f| dm < \infty$, for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $m(A) < \delta$

$$\int_A |f| dm < \epsilon$$

Hence if $(a_j, b_j) \subset [a,b]$, $j = 1, ..., k$ are such that $\sum_{j=1}^k (b_j - a_j) < \delta$, then

$$\sum_{j=1}^k |F(b_j) - F(a_j)| \leq \int_{\bigcup_{j=1}^k (a_j, b_j)} |f| dm < \epsilon$$

since $m(\bigcup_{j=1}^k (a_j, b_j)) \leq \sum_{j=1}^k (b_j - a_j) < \delta$. Thus $F$ is absolutely continuous.
Now consider the only if part. (To be completed later). □

Example:

Let $\mathscr{H}$ = the collection of all absolutely continuous functions on $[0,1]$ such that $f(0) = 0$ and $f' \in L^2((0,1))$. If $\langle f, g \rangle = \int_{[0,1]} f'(t)\overline{g'(t)}dt$ for $f, g \in \mathscr{H}$, then $\mathscr{H}$ is a Hilbert space.

If $V$ is a vector space with an inner product, what happens if it is not complete with respect to the induced metric?

**Proposition 1.1.6** If $V$ is an inner product space with inner product $\langle \cdot, \cdot \rangle_V$, and $\mathscr{H}$ is the completion of $V$ w.r.t the metric induced by the norm on $V$, then there is an inner product $\langle \cdot, \cdot \rangle_{\mathscr{H}}$ on $\mathscr{H}$ which agrees on $V$ with its inner product. That is, the completion of $V$ is a Hilbert space.

This says we can embed an incomplete inner product space into a Hilbert space. We close with an example of a Hilbert space from analytic function theory.

**Definition 1.1.7** If $G$ is an open subset of the complex plane $\mathbb{C}$, then $L_a^2(G)$ denotes the collection of all analytic functions $f : G \to \mathbb{C}$ such that

$$\int \int_G |f|^2 d\mu < \infty$$

$L_a^2(G)$ is called the **Bergman space** for $G$.

Note that $L_a^2(G) \subseteq L^2(\mu)$, so that $L_a^2(G)$ has a natural inner product and norm from $L^2(\mu)$.

**Lemma 1.1.8** If $F$ is analytic in a neighborhood of $\overline{B}(a;r)$, then

$$f(a) = \frac{1}{\pi r^2} \int \int_{B(a;r)} f$$

***Proof*** By the mean value property, if $0 < t \leq r$, $f(a) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(a + te^{i\theta}) d\theta$. Hence

$$\begin{aligned} \frac{1}{\pi r^2} \int \int_{B(a;r)} f &= \frac{1}{\pi r^2} \int_0^r t \left[ \int_{-\pi}^{\pi} \right] dt \\ &= \frac{2}{r^2} \int_0^r t f(a) dt = f(a) \end{aligned}$$

where we use Fubini's theorem in the first line. □

**Corollary 1.1.9** If $f \in L_a^2(G)$, $a \in G$, and $0 < r < \text{dist}(a, \partial G)$, then

$$|f(a)| \leq \frac{1}{r\sqrt{\pi}} ||f||_2$$

***Proof*** Since $\overline{B}(a;r) \subseteq G$, the preceding lemma and the CBS inequality imply

$$\begin{aligned} |f(a)| &= \frac{1}{\pi r^2} \left| \int \int_{B(a;r)} f \cdot 1 \right| \\ &\leq \frac{1}{\pi r^2} ||f||_2 \sqrt{\text{Area}(B(a;r))} \\ &= \frac{1}{r\sqrt{\pi}} ||f||_2 \end{aligned}$$

**Proposition 1.1.10** $L_a^2(G)$ is a Hilbert space.

***Proof*** If $\mu$ denotes the Lebesgue measure induced on $G$, then $L^2(\mu)$ is a Hilbert space and $L_a^2(G) \subseteq L^2(\mu)$. So it sufficies to show $L_a^2(G)$ is closed in $L^2(\mu)$. Let $\{f_n\}$ be a sequence in $L_a^2(G)$ and let $f \in L^2(\mu)$ such that $||f_n - f||_2 \to 0$ as $n \to \infty$.
Suppose $\overline{B}(a;r) \subseteq G$ and let $0 < \rho < \text{dist}(B(a;r), \partial G)$. By the preceding corollary there exists $C$ such that $|f_n(z) - f_m(z)| \leq C ||f_n - f_m||_2$ for all $n, m$ and $|z - a| \leq \rho$. Thus $\{f_n\}$ is uniformly Cauchy on any closed disk in $G$. By Montel's Theorem or Morera's Theorem, there is an analytic function $g$ on $G$ such that $f_n(z) \to g(z)$ uniformly on compact subsets of $G$. But since $||f_n - f||_2 \to 0$, a result of Riesz implies there is a subsequence $\{f_{n_k}\}$ such that $f_{n_k}(z) \to f(z)$ almost everywhere w.r.t $\mu$. Thus $f = g$ a.e. $[\mu]$, and so $f \in L_a^2(G)$. □

## 1.2 Orthogonality

One of the greatest advantages to a Hilbert space is its underlying concept of orthogonality.

**Definition 1.2.1** If $\mathscr{H}$ is a Hilbert space and $f, g \in \mathscr{H}$, then $f$ and $g$ are said to be **orthogonal** if $\langle f, g\rangle = 0$. In symbols $f \perp g$. If $A, B \subseteq \mathscr{H}$, then $A \perp B$ means $f \perp g$ for all $f \in A$ and $g \in B$.

This allows us to generalize the notion of Pythagorean's theorem.

**Theorem 1.2.2 (The Pythagorean Theorem)** If $f_1, ..., f_n$ are pairwise orthogonal vectors in $\mathscr{H}$, then

$$||f_1 + f_2 + \cdots + f_n||^2 = ||f_1||^2 + ||f_2||^2 + \cdots + ||f_n||^2$$

***Proof*** If $f_1 \perp f_2$, then

$$||f_1 + f_2||^2 = ||f_1||^2 + 2\text{Re}\,\langle f_1, f_2, +\rangle\, ||f_2||^2$$

by the polar identity. As $f_1 \perp f_2$, this implies the result for $n = 2$. The general case follows by induction with the observeration $f_1 + \cdots + f_n \perp f_{n+1}$ if $f_i \perp f_{n+1}$ for all $1 \le i \le n$. □

We also have a general law for norms.

**Theorem 1.2.3 (Parallelogram Law)** If $\mathscr{H}$ is a Hilbert space and $f, g \in \mathscr{H}$, then

$$||f + g||^2 + ||f - g||^2 = 2(||f||^2 + ||g||^2)$$

***Proof*** For any $f, g \in \mathscr{H}$, the polar identity implies

$$||f + g||^2 = ||f||^2 + 2\text{Re}\,\langle f, g\rangle + ||g||^2$$
$$||f - g||^2 = ||f||^2 - 2\text{Re}\,\langle f, g\rangle + ||g||^2$$

Adding gives the result. □

Before moving to the next concept we recall the notion of convexity.

**Definition 1.2.4** If $V$ is a vector space over $\mathbb{F}$ and $A \subseteq V$, then $A$ is **convex** if for any $x, y \in A$, and $t \in [0, 1]$, $tx + (1 - t)y \in A$.

Note $\{tx + (1 - t)y : t \in [0, 1]\}$ is the straight-line segment joining $x$ and $y$. Thus convex sets contain straight lines between points in the set. A common example is given by any linear subspace of $V$. Further, any singleton set or intersection of convex sets is again convex. In particular, if $V$ is an inner product space, then every open ball $B(f; r) = \{g \in V : ||f - g|| < r\}$ is convex, as is every closed ball.

**Theorem 1.2.5** If $\mathscr{H}$ is a Hilbert space, $K$ is a closed convex nonempty subset of $\mathscr{H}$, and $h \in \mathscr{H}$, then there is a unique point $k_0 \in K$ such that

$$||h - k_0|| = \mathrm{dist}(h, K) := \inf\{||h - k|| : k \in K\}$$

***Proof*** Consider $K - h$, the translate of $K$ by $-h$, so it suffices to assume that $h = 0$. Indeed if $K$ is closed and convex, so is $K - h$. Hence we want to show there is a unique vector $k_0$ in $K$ such that

$$||k_0|| = \mathrm{dist}(0, K)$$

Let $d = \mathrm{dist}(0, K)$. Then there is a sequence $\{k_n\}$ in $K$ such that $||k_n|| \to d$. The Parallelogram Law implies that

$$\left\|\frac{k_n - k_m}{2}\right\|^2 = \frac{1}{2}(||k_n||^2 + ||k_m||^2) - \left\|\frac{k_n + k_m}{2}\right\|^2$$

Since $K$ is convex $\frac{1}{2}(k_n + k_m) \in K$. Hence $||(k_n + k_m)/2||^2 \geq d^2$. If $\epsilon > 0$, choose $N$ such that $n \geq N$ implies $||k_n||^2 < d^2 + \epsilon^2/4$. By the equation above, if $n, m \geq N$, then

$$\left\|\frac{k_n - k_m}{2}\right\|^2 < \frac{1}{2}(2d^2 + \epsilon^2/2) - d^2 = \epsilon^2/4$$

Thus $||k_n - k_m|| < \epsilon$ for $n, m \geq N$, and $\{k_n\}$ is a Cauchy sequence. Since $\mathscr{H}$ is complete and $K$ is closed, there is a $k_0 \in K$ such that $||k_n - k_0|| \to 0$. Also for all $k_n$,

$$d \leq ||k_0|| \leq ||k_0 - k_n|| + ||k_n|| \to d$$

so $||k_0|| = d$.
To prove uniqueness suppose $h_0 \in K$ such that $||h_0|| = d$. By convexity $(k_0 + h_0)/2 \in K$, so

$$d \leq ||(h_0 + k_0)/2|| \leq (||h_0|| + ||k_0||)/2 = d$$

So $||(h_0 + k_0)/2|| = d$, and the Parallelogram Law implies

$$d^2 = d^2 - \left\|\frac{h_0 - k_0}{2}\right\|^2$$

so $h_0 = k_0$. □

If the general convex set is replaced by a closed linear subspace, more can be said.

**Theorem 1.2.6** If $W$ is a closed linear subspace of $\mathscr{H}$, $h \in \mathscr{H}$, and $f_0$ is the unique element of $M$ such that $||h - f_0|| = \mathrm{dist}(h, M)$, then $h - f_0 \perp M$. Conversely, if $f_0 \in M$ such that $h - f_0 \perp M$, then $||h - f_0|| = \mathrm{dist}(h, M)$.

***Proof*** Suppose $f_0 \in M$ is as described. Let $f \in M$ so $f_0 + f \in M$, and hence

$$||h - f_0||^2 \leq ||h - (f_0 + f)||^2 = ||h - f_0||^2 - 2\mathrm{Re}\,\langle h - f_0, f, +\rangle\, ||f||^2$$

Thus

$$2\mathrm{Re}\,\langle h - f_0, f, \leq\rangle\, ||f||^2$$

for any $f$ Id $M$. Fix $f \in M$ and substitute $te^{i\theta}$ for $f$ in the preceeding inequality, where $\langle h - f_0, f, =\rangle\, re^{i\theta}$, $r \geq 0$. This yields

$$2tr \leq t^2\, ||f||^2$$

Letting $t \longrightarrow 0$, we see that $r = 0$, so $h - f_0 \perp f$.
For the converse suppose $f_0 \in M$ such that $h - f_0 \perp M$. If $f \in M$, then $h - f_0 \perp f_0 - f$, so

$$||h - f||^2 = ||h - f_0||^2 + ||f_0 - f||^2 \geq ||h - f_0||^2$$

with equality only if $f_0 = f$. Thus $||h - f_0|| = \mathrm{dist}(h, M)$. □

If $A \subseteq \mathscr{H}$ we write $A^\perp := \{f \in \mathscr{H} : f \perp A\}$. Note that for any $A$, $A^\perp$ is a closed linear subspace of $\mathscr{H}$.

If $M$ is a closed linear subspace of $\mathscr{H}$, let $P : \mathscr{H} \longrightarrow M$ be the function defined by $Ph = f_0 <$ where $f_0 \in M$ is the unique element such that $h - f_0 \in M^\perp$.

**Theorem 1.2.7** If $M$ is a closed linear subspace of $\mathscr{H}$, then

(a) $P$ is a linear transformation on $\mathscr{H}$
(b) $||Ph|| \leq ||h||$ for all $h \in \mathscr{H}$,
(c) $P^2 = P$
(d) $\ker P = M^\perp$ and $\mathrm{ran} P = M$.

***Proof*** For (a), if $h_1, h_2 \in \mathscr{H}$, and $\alpha \in \mathbb{F}$, then for any $f \in M$,

$$\langle(\alpha h_1 + h_2) - (\alpha Ph_1 + Ph_2), f\rangle = \alpha_1\, \langle h_1 - Ph_1, f, +\rangle\, \alpha_2\, \langle h_2 - Ph_2, f, =\rangle\, 0$$

so by uniqueness $P(\alpha h_1 + h_2) = \alpha Ph_1 + Ph_2$.
For (b), if $h \in \mathscr{H}$, then

$$||h||^2 = ||h - Ph||^2 + ||Ph||^2 \geq ||Ph||^2$$

as $Ph \in M$ and $h - Ph \in M^\perp$.
For (c), if $f \in M$, then $Pf = f$. Hence, for any $h \in \mathscr{H}$, as $Ph \in M$, $P^2h = P(Ph) = Ph$, so $P^2 = P$.
For (d), $Ph = 0$ if and only if $h \in M^\perp$. Further, $\mathrm{ran} P \subseteq M$ and $Pf = f$ for any $f \in M$, so $\mathrm{ran} P = M$. □

**Definition 1.2.8** If $M$ is a closed linear subspace of $\mathscr{H}$, then the map $P$ in the previous theorem is called the **orthogonal projection** of $\mathscr{H}$ onto $M$.

We write $M \leq \mathscr{H}$ to signify $M$ is a closed linear subspace of $\mathscr{H}$.

**Corollary 1.2.9** If $M \leq \mathscr{H}$, then $(M^\perp)^\perp = M$.

***Proof*** Note that $\mathrm{id}_{\mathscr{H}} - P_M = P_{M^\perp}$. Thus $(M^\perp)^\perp = \ker(\mathrm{id}_{\mathscr{H}} - P_M) = \mathrm{ran} P_M = M$ since $P_M$ is idempotent. □

**Corollary 1.2.10** If $A \subseteq \mathscr{H}$, then $(A^{\perp})^{\perp}$ is the closed linear span of $A$ in $\mathscr{H}$.

***Proof*** Indeed if $A' = \langle A \rangle$ denotes the closed linear span of $A$, then $(A^{\perp})^{\perp} \subseteq (A'^{\perp})^{\perp} = A'$ from the previous result. Conversely, $(A^{\perp})^{\perp}$ is a closed linear span containing $A$, so it must coincide with $A'$. □

**Corollary 1.2.11** If $Y$ is a linear subspace of $\mathscr{H}$, then $Y$ is dense in $\mathscr{H}$ if and only if $Y^{\perp} = (0)$.

## 1.3 The Riesz Representation Theorem

The present section deals with the representation of certain linaer functionals on Hilbert space, although there is another Riesz representation theorem which shall be discussed later.

**Proposition 1.3.1** Let $\mathscr{H}$ be a Hilbert space and $L : \mathscr{H} \to \mathbb{F}$ a linear functional. The following are equivalent:

(a) $L$ is continuous
(b) $L$ is continuous at 0
(c) $L$ is continuous at some point
(d) There is a constant $c > 0$ such that $|L(h)| \leq c\,||h||$ for every $h \in \mathscr{H}$.

***Proof*** Note that $(a) \implies (b) \implies (c)$ and $(d) \implies (b)$. Hence it is sufficient to show $(c) \implies (a)$ and $(b) \implies (d)$.
First, suppose $(c)$, and let $x \in \mathscr{H}$ be the point at which $L$ is continuous. Then fix $y \in \mathscr{H}$ and $\epsilon > 0$. Then there exists $\delta > 0$ such that if $||x - h|| < \delta$, $|L(x) - L(h)| < \epsilon$. Now let $z \in \mathscr{H}$ such that $||y - z|| < \delta$. Then setting $h = (x - y + z)$, $x - h = y - z$, so $||y - z|| < \delta$ implies $||x - h|| < \delta$, and consequently

$$|L(y) - L(z)| = |L(y - z)| = |L(x - h)| = |L(x) - L(h)| < \epsilon$$

so $L$ is continuous.
Finally, to show $(b) \implies (d)$ consider $\epsilon = 1$. Then there exists $\delta > 0$ such that if $||h|| < \delta$, $|L(h)| < 1$. Setting $c = \frac{2}{\delta}$ we have that $|L(\delta h / 2\,||h||)| < 1$, and so by linearity

$$|L(h)| < c\,||h||$$

completing the proof. □

Note that the previous result did not rely on the inner product of the Hilbert space but rather just its induced norm, and that it didn't require completeness of the domain or codomain. Hence, the previous result applies more generally for any linear map between normed spaces.

**Definition 1.3.2** A **bounded linear functional** $L$ on $\mathscr{H}$ is a linear functional for which there is a constant $c > 0$ such that $|L(h)| \leq c\,||h||$ for all $h \in \mathscr{H}$. That is a bounded linear functional is precisely a continuous linear functional.

For a bounded linear function $L : \mathscr{H} \to \mathbb{F}$, define

$$||L|| = \sup\{|L(h)| : ||h|| \leq 1\}$$

Note that $||L|| < \infty$, and in particular a linear functional is bounded if and only if this value is finite. $||L||$ is called the supremum norm of $L$. Note again this may be generalized to linear maps between normed spaces.

**Proposition 1.3.3** If $L$ is a bounded linear functional, then

$$\begin{aligned} ||L|| &= \sup\{|L(h)| : ||h|| = 1\} \\ &= \sup\{|L(h)|/||h|| : h \in \mathscr{H}, h \neq 0\} \\ &= \inf\{c > 0 : |L(h)| \leq c\,||h||\,, h \in \mathscr{H} \end{aligned}$$

Also $|L(h)| \leq ||L||\,||h||$ for all $h \in \mathscr{H}$.

***Proof*** Note the first two notions are supremums over the same sets, and hence agree. Further, if $0 < ||h|| < 1$, $|L(h)| = ||h||\,|L(h/||h||)| \leq |L(h/||h||)|$, so the supremum must occur on the boundary of the unit circle. Thus it is sufficient to consider the last equality.
First, observe that if $c > 0$ and $|L(h)| \leq c\,||h||$ for all $h$, then in particular $|L(h)| \leq c$ for $||h|| = 1$, so $||L|| \leq c$ by definition of the supremum. Thus $||L|| \leq \inf\{c > 0 : |L(h)| \leq c\,||h||\,, h \in \mathscr{H}\}$. Further, as $|L(h)| \leq ||L||$ for all $h$, $||h|| = 1$, $|L(h)| \leq ||L||\,||h||$ by linearity, so we have equality. □

Observe that for any $h_0 \in \mathscr{H}$ we have a linear functional $L : \mathscr{H} \to \mathbb{F}$ given by $L(h) = \langle h, h_0\rangle$. Further, the CBS inequality implies that $|L(h)| \leq ||h||\,||h_0||$, so $L$ is bounded and $||L|| \leq ||h_0||$. In fact, $L(h_0/||h_0||) = ||h_0||$, so we have equality. We now give a converse to these observations.

**Theorem 1.3.4 (The Riesz Representation Theorem)** If $L : \mathscr{H} \to \mathbb{F}$ is a bounded linear functional on a Hilbert space, then there is a unique $h_0 \in \mathscr{H}$ for which $L(h) = \langle h, h_0\rangle$ for every $h \in \mathscr{H}$.

***Proof*** Let $M = \ker L$. Because $L$ is continuous $M$ is a closed linear subspace of $\mathscr{H}$. Since we may assume that $M \neq \mathscr{H}$, $M^\perp \neq (0)$. Hence there exists $f_0 \in M^\perp$ such that $L(f_0) = 1$. Now if $h \in \mathscr{H}$ and $\alpha = L(h)$, then $L(h - \alpha f_0) = L(h) - \alpha = 0$, so $h - L(h)f_0 \in M$. Thus

$$0 = \langle h - L(h)f_0, f_0\rangle = \langle h, f_0\rangle - L(h)\,||f_0||^2$$

So if $h_0 = ||f_0||^{-2}\, f_0$, $L(h) = \langle h, h_0\rangle$ for all $h \in \mathscr{H}$.
For uniqueness observe that if $h_0' \in \mathscr{H}$ also satisfies the claim, then $h_0 - h_0' \perp \mathscr{H}$, so in particular $h_0 - h_0' \perp h_0 - h_0'$, so $h_0' = h_0$ by positive defniteness. □

**Corollary 1.3.5** If $(X, \Omega, \mu)$ is a measure space and $F : L^2(\mu) \to \mathbb{F}$ is a bounded linear functional, then there is a unique $h_0$ in $L^2(\mu)$ such that

$$F(h) = \int_X h\overline{h_0} d\mu$$

for all $h \in L^2(\mu)$.

This is a special case of a more general result on bounded linear functionals on $L^p(\mu)$, $1 \leq p < \infty$, but it is interesting to note that it is only a consequence of the result for Hilpert spaces and the fact that $L^2(\mu)$ has a natural Hilbert space structure.

## 1.4 Orthonormal Sets and Bases

As in Euclidean space, each Hilbert space can be coordinatized in a suitable sense. By this we mean we can introduce a notion of "orthonormal basis" on the space.

**Definition 1.4.1** An **orthonormal** subset of a Hilbert space $\mathscr{H}$ is a subset $\mathscr{O}$ having the properties:

(a) for $e \in \mathscr{O}$, $||e|| = 1$
(b) if $e_1, e_2 \in \mathscr{O}$, and $e_1 \neq e_2$, then $e_1 \perp e_2$.

An **orthonormal basis** for $\mathscr{H}$ is a maximal orthonormal set.

Note that in general an orthonormal basis need not be a vector space basis for $\mathscr{H}$. Indeed, if $\mathscr{H}$ is infinite-dimensional then an orthonormal basis is never a vector space basis.

**Proposition 1.4.2** If $\mathscr{O}$ is an orthonormal set in $\mathscr{H}$, then there is an orthonormal basis for $\mathscr{H}$ containing $\mathscr{O}$.

This is a straightforward application of Zorn's Lemma.

Example:

Let $\mathscr{H} = L^2_{\mathbb{C}}([0, 2\pi])$ and for $n \in \mathbb{Z}$ define $e_n$ in $\mathscr{H}$ by $e_n t = (2\pi)^{-1/2} \exp(int)$. Then $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal set in $\mathscr{H}$.

It happens that the set in this example is in fact a basis. First we need a bit of theory to make this easier though.

Example:

If $\mathscr{H} = \mathbb{F}^d$, and $1 \leq k \leq d$, $e_k$ = the $d$-tuple with 1 in the $k$th place and zeros elsewhere, then $\{e_1, ..., e_d\}$ is a basis for $\mathscr{H}$.

Example:

Let $\mathscr{H} = l^2(I)$. For each $i \in I$, define $e_i$ in $\mathscr{H}$ to be the indicator function. Then $\{e_i : i \in I\}$ is a basis.

**Theorem 1.4.3 (The Gram-Schmidt Orthogonalization Process)** If $\mathscr{H}$ is a Hilbert space and $\{h_n : n \in \mathbb{N}\}$ is a linearly independent subset of $\mathscr{H}$, then there is an orthonormal set $\{e_n : n \in \mathbb{N}\}$ such that for every $n$, the linear space of $\{e_1, ..., e_n\}$ equals the linear span of $\{h_1, ..., h_n\}$.

***Proof*** We shall proceed inductively. If $n = 1$, set $e_1 = h_n/||h_n||$. Now suppose $e_1, ..., e_k$ have been constructed for some $k \geq 1$, such that for all $m \leq k$, $\text{span}(e_1, ..., e_m) = \text{span}(h_1, ..., h_m)$, and $\{e_1, ..., e_k\}$ is an orthonormal set. Then define $e_{k+1}$ by setting

$$e'_{k+1} = h_{k+1} - \sum_{i=1}^{m} \langle h_{k+1}, e_i \rangle \, e_i$$

and then setting $e_{k+1} = e'_{k+1}/\left\|e'_{k+1}\right\|$, where $e'_{k+1}$ is non-zero as $\{h_1, ..., h_k, h_{k+1}\}$ is linearly independent by assumption. Thus $\{e_1, ..., e_k, e_{k+1}\}$ satisfies the claim, so by induction we have the desired orthonormal set $\{e_n : n \in \mathbb{N}\}$. □

For $A \subseteq \mathscr{H}$, we write $\bigvee A$ for the closed linear span of $A$.

**Proposition 1.4.4** Let $\{e_1, ..., e_n\}$ be an orthonormal set in $\mathscr{H}$ and let $M = \bigvee\{e_1, ..., e_n\}$. If $P$ is the orthogonal projection of $\mathscr{H}$ onto $M$, then

$$Ph = \sum_{k=1}^{n} \langle h, e_k, e \rangle_k$$

for all $h \in \mathscr{H}$.

Follows from uniqueness of the vector $h_0 \in M$ for $h \in \mathscr{H}$ such that $h - h_0 \perp M$, and the definition on the right of the proposed equation.

**Theorem 1.4.5 (Bessel's Inequality)** If $\{e_n : n \in \mathbb{N}\}$ is an orthonormal set and $h \in \mathscr{H}$, then

$$\sum_{n=1}^{\infty} | \langle h, e_n \rangle |^2 \leq ||h||^2$$

***Proof*** Let $h_n = h - \sum_{k=1}^{n} \langle h, e_k \rangle e_k$. Then $h_n \perp e_k$ for $1 \le k \le n$. By the Pythagorean Theorem

$$\begin{aligned} ||h||^2 &= ||h_n||^2 + \left\| \sum_{k=1}^{n} \langle h, e_k \rangle e_k \right\|^2 \\ &= ||h_n||^2 + \sum_{k=1}^{n} | \langle h, e_k \rangle |^2 \\ &\ge \sum_{k=1}^{n} | \langle h, e_k \rangle |^2 \end{aligned}$$

As $n$ was arbitrary the sum exists and the result follows. □

**Corollary 1.4.6** If $\mathcal{O}$ is an orthonormal set in $\mathcal{H}$ and $h \in \mathcal{H}$, then $\langle h, e \rangle \neq 0$ for at most a countable number of vectors $e \in \mathcal{O}$.

***Proof*** For each $n \ge 1$, let $\mathcal{O}_n = \{e \in \mathcal{O} : | \langle h, e \rangle | \ge 1/n\}$. By Bessel's inequality, $\mathcal{O}_n$ must be finite. Thus, as the desired collection is a countable union of finite sets, it is countable. □

**Corollary 1.4.7** If $\mathcal{O}$ is an orthonormal set and $h \in \mathcal{H}$, then

$$\sum_{e \in \mathcal{O}} | \langle h, e \rangle |^2 \le ||h||^2$$

We make precise the notion of an infinite sum now. First we recall the notion of a **net**.

**Definition 1.4.8** A function whose domain is a **directed set** is called a **net**, where a directed set is a nonempty set $A$ together with a reflexive and transitive binary relation $\le$ (i.e. a pre-order), such that for any $a, b \in A$, there exists $c \in A$ such that $a \le c$ and $b \le c$.
A net $f : (A, \le) \to X$ in a topological space $X$ is said to be **eventually in** $S$, where $S$ is a subset of $X$, if there exists $a \in A$ such that for all $b \in A$ with $b \ge a$, $f(b) \in S$. A point $x \in X$ is called a **limit** of the net $f$ if and only if for every open neighborhood $U$ of $x$, then net $f$ is eventually in $U$.

Let $\mathcal{F}$ be the collection of all finite subsets of $I$ and order $\mathcal{F}$ by inclusion. For each $F \in \mathcal{F}$, define

$$h_F = \sum \{h_i : i \in F\}$$

Since this is a finite sum, $h_F$ is a well-defined element of $\mathcal{H}$ (by commutativity and associativity of addition). Now $\{h_F : F \in \mathcal{F}\}$ is a net in $\mathcal{H}$.

**Definition 1.4.9** With notation above, the sum $\sum\{h_i : i \in I\}$ converges if the net $\{h_F : F \in \mathcal{F}\}$ converges; the value of the sum is the limit of the net.

Note that if $I$ is countable this definition of convergent sum is not the usual one. That is, if $\{h_n\}$ is a sequence in $\mathcal{H}$, then the convergence of $\sum\{h_n : n \in \mathbb{N}\}$ is not equivalent to the convergence of $\sum_{n=1}^{\infty} h_n$.

Even if $\mathcal{H} = \mathbb{F}$, these concepts do not coincide. If, however, $\sum\{h_n : n \in \mathbb{N}\}$ converges, then $\sum_{n=1}^{\infty} h_n$ converges.

**Lemma 1.4.10** If $\mathcal{O}$ is an orthonormal set and $h \in \mathcal{H}$, then

$$\sum\{\langle h, e\rangle\, e : e \in \mathcal{O}\}$$

converges in $\mathcal{H}$.

***Proof*** By our previous work there are $e_1, e_2, \ldots$ in $\mathcal{O}$ such that $\{e \in \mathcal{O} : \langle h, e\rangle \neq 0\} = \{e_1, e_2, \ldots\}$. We also know that $\sum_{n=1}^{\infty} |\langle h, e_n\rangle|^2 \leq ||h||^2 < \infty$. So if $\epsilon > 0$, there is an $N$ such that $\sum_{n=N}^{\infty} |\langle h, e_n\rangle|^2 < \epsilon^2$. Let $F_0 = \{e_1, \ldots, e_{N-1}\}$, and let $\mathcal{F}$ be the collection of all finite subsets of $\mathcal{O}$. For $F$ in $\mathcal{F}$ define $h_F := \sum\{\langle h, e\rangle\, e : e \in F\}$. If $F, G \in \mathcal{F}$, and both contain $F_0$, then

$$\begin{aligned} ||h_F - h_G||^2 &= \sum\{|\langle h, e\rangle|^2 : e \in (F\backslash G) \cup (G\backslash F)\} \\ &\leq \sum_{n=N}^{\infty} |\langle h, e_n\rangle|^2 < \epsilon^2 \end{aligned}$$

So $\{h_F : F \in \mathcal{F}\}$ is a Cauchy net in $\mathcal{H}$. Because $\mathcal{H}$ is complete this net converges. In fact, it converges to $\sum_{n=1}^{\infty} \langle h, e_n\rangle\, e_n$. □

**Theorem 1.4.11** If $\mathcal{O}$ is an orthonormal set in $\mathcal{H}$, then the following statements are equivalent.

(a) $\mathcal{O}$ is a basis for $\mathcal{H}$
(b) If $h \in \mathcal{H}$ and $h \perp \mathcal{O}$, then $h = 0$
(c) $\bigvee \mathcal{O} = \mathcal{H}$
(d) If $h \in \mathcal{H}$, then $h = \sum\{\langle h, e\rangle\, e : e \in \mathcal{O}\}$
(e) If $g, h \in \mathcal{H}$, then

$$\langle g, h\rangle = \sum\{\langle g, e\rangle \langle e, h\rangle : e \in \mathcal{O}\}$$

(f) If $h \in \mathcal{H}$, then $||h||^2 = \sum\{|\langle h, e\rangle|^2 : e \in \mathcal{O}\}$ (**Parseval's identity**)

***Proof*** $(a) \implies (b)$ follows from maximality, $(b) \iff (c)$ was shown previously, $(e) \implies (f)$ is immediate. For $(b) \implies (d)$, if $h \in \mathcal{H}$ then $f = h - \sum\{\langle h, e\rangle\, e : e \in \mathcal{O}\}$ is a well-defined vector by the previous Lemma. If $e_1 \in \mathcal{O}$, then

$$\langle f, e_1\rangle = \langle h, e_1\rangle - \sum\{\langle h, e\rangle \langle e, e_1\rangle : e \in \mathcal{O}\} = \langle h, e_1\rangle - \langle h, e_1\rangle = 0$$

That is $f \in \mathcal{O}^{\perp}$, so $f = 0$.
$(f) \implies (a)$ follows by contradiction. If $\mathcal{O}$ is not a basis then there is a unit vector $e_0$ in $\mathcal{H}$ such that $e_0 \perp \mathcal{O}$. Hence $0 = \sum\{|\langle e_0, e\rangle|^2 : e \in \mathcal{O}\}$, contrary to $(f)$.
Finally, for $(d) \implies (e)$, consider $f, g \in \mathcal{H}$. Then $f = \sum\{\langle f, e\rangle\, e : e \in \mathcal{O}\}$ and $g = \sum\{\langle g, e\rangle\, e : e \in \mathcal{O}\}$. Then we have

$$\langle g, f\rangle = \left\langle \sum\{\langle g, e\rangle\, e : e \in \mathcal{O}\}, f\right\rangle = \sum\{\langle g, e\rangle \langle e, f\rangle : e \in \mathcal{O}$$

using continuity and linearity of the inner product in its first component. □

Just as in finite dimensional spaces, we can use an orthonormal basis in Hilbert space to define a concept of dimension.

**Proposition 1.4.12** If $\mathscr{H}$ is a Hilbert space, any two orthonormal basis have the same cardinality.

***Proof*** Let $\mathscr{O}$ and $\mathscr{F}$ be two orthonormal bases for $\mathscr{H}$, and put $\kappa$ for the cardinality of $\mathscr{O}$ and $\eta$ for the cardinality of $\mathscr{F}$. If $\kappa$ or $\eta$ is finite, then $\mathscr{H}$ is finite dimensional so $\kappa = \dim \mathscr{H} = \eta$. Suppose both $\kappa$ and $\eta$ are infinite. For $e \in \mathscr{O}$, let $\mathscr{F}_e := \{f \in \mathscr{F} : \langle e, f\rangle \neq 0\}$; so $\mathscr{F}_e$ is countable. By part (b) of the previous theorem, each $f \in \mathscr{F}$ belongs to at leasat one set $\mathscr{F}_e$, $e \in \mathscr{O}$. That is, $\mathscr{F} = \bigcup\{\mathscr{F}_e : e \in \mathscr{O}\}$. Hence $\eta \leq \kappa\aleph_0 = \kappa$ since $\kappa$ is an infinite cardinal. Similarly $\kappa \leq \eta$, so $\kappa = \eta$. □

**Definition 1.4.13** The **dimension** of a Hilbert space is the cardinality of abasis and is denoted by $\dim \mathscr{H}$.

Recall that if $(X, d)$ is a separable metric space and $\{B_i = B(x_i; \epsilon_i) : i \in I\}$ is a collection of pairwise disjoint open balls in $X$, then $I$ must be countable. Indeed, if $D$ is a countable dense subset, we can obtain an inclusion of $I$ into $D$ by taking a point $y_i$ in each $B_i \cap D \neq \emptyset$. Thus $I$ must be countable.

**Proposition 1.4.14** If $\mathscr{H}$ is an infinite dimensional Hilbert space, then $\mathscr{H}$ is separable if and only if $\dim \mathscr{H} = \aleph_0$.

***Proof*** Let $\mathscr{O}$ be a basis for $\mathscr{H}$. If $e_1, e_2 \in \mathscr{O}$, then $||e_1 - e_2||^2 = ||e_1||^2 + ||e_2||^2 = 2$. Hence $\{B(e : 1/\sqrt{2}) : e \in \mathscr{O}\}$ is a collection of pairwise disjoint open balls in $\mathscr{H}$. From the previous discussion, the assumption that $\mathscr{H}$ is separable implies $\mathscr{O}$ is countable.
Conversely, if $\mathscr{O}$ is countable, then first recall $\mathbb{Q} + \mathbb{Q}i$ is countable and dense in $\mathbb{C}$, and $\mathbb{Q}$ is dense and countable in $\mathbb{R}$. Let $S$ be the appropriate countable dense subset in $\mathbb{F}$. Then

$$A := \{\sum_{i=1}^{n} s_i e_i : n \geq 1, s_i \in S\}$$

corresponds to the collection of finite sequences with terms in $S$, and hence is countable as $S$ is. We claim this is our dense subset of $\mathscr{H}$.
Let $x \in H$, so

$$x = \sum_{i=1}^{\infty} \langle x, e_i\rangle\, e_i$$

from our previous Theorem. Since this is convergent in the norm of $\mathscr{H}$, fixing $\epsilon > 0$ we can find $N$ such that

$$\left\| \sum_{n=N+1}^{\infty} c_n e_n \right\| < \epsilon/2$$

Also, since $S$ is a dense subset of $\mathbb{F}$, for every $i \leq N$, we can find $s_i \in S$ such that

$$|c_i - s_i| < \epsilon/2^{i+1}$$

Consider the element $x_N = \sum_{i=1}^{N} s_i e_i \in A$. Then

$$\begin{aligned}
||x - x_N|| &= \left\| \sum_{n\geq 1} c_n e_n - \sum_{i=1}^{N} s_i e_i \right\| \\
&\leq \left\| \sum_{n\geq N+1} c_n e_n \right\| + \left\| \sum_{i=1}^{N} (c_i - s_i) e_i \right\| \\
&\leq \epsilon/2 + \sum_{i=1}^{N} |c_i - s_i| \\
&\leq \sum_{i=1}^{N} \frac{\epsilon}{2^{i+1}} + \epsilon/2 \\
&\leq \sum_{n\geq 1} \frac{\epsilon}{2^{n+1}} + \epsilon/2 \\
&= \epsilon
\end{aligned}$$

completing the proof. □

## 1.5 Isomorphic Hilbert Spaces

We now define maps which preserve the structure of our Hilbert spaces to obtain a category.

**Definition 1.5.1** If $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces, an **isomorphism** between $\mathscr{H}$ and $\mathscr{K}$ is a linear isomorphism $U : \mathscr{H} \to \mathscr{K}$ such that

$$\langle Uh, Ug \rangle = \langle h, g \rangle$$

for all $h, g \in \mathscr{H}$.

Note if $U$ is an isomorphism so is $U^{-1}$ so this is the correct notion. Recall an **isometry** between metric spaces is a map that preserves distances.

**Proposition 1.5.2** If $V : \mathscr{H} \to \mathscr{K}$ is a linear map between Hilbert spaces, then $V$ is an isometry if and only if $\langle Vh, Vg \rangle = \langle h, g \rangle$ for all $h, g \in \mathscr{H}$.

***Proof*** Assume $\langle Vh, Vg \rangle = \langle h, g \rangle$ for all $h, g \in \mathscr{H}$. Then $||Vh||^2 = \langle Vh, Vh \rangle = \langle h, h \rangle = ||h||^2$, so $V$ is an isometry.
Now assume that $V$ is an isometry. If $h, g \in \mathscr{H}$ and $\lambda \in \mathbb{F}$, then $||h + \lambda g||^2 = ||Vh + \lambda Vg||^2$. Using the polar identity on both sides of this equation gives

$$||h||^2 + 2\mathrm{Re}\bar{\lambda} \langle h, g \rangle + |\lambda|^2 ||g||^2 = ||Vh||^2 + 2\mathrm{Re}\bar{\lambda} \langle Vh, Vg \rangle + |\lambda|^2 ||Vg||^2$$

But $||Vh|| = ||h||$ and $||Vg|| = ||g||$, so this equation becomes

$$\mathrm{Re}\overline{\lambda}\langle h, g\rangle = \mathrm{Re}\overline{\lambda}\langle Vh, Vg\rangle$$

for any $\lambda \in \mathbb{F}$. If $\mathbb{F} = \mathbb{R}$, take $\lambda = 1$, and if $\mathbb{F} = \mathbb{C}$, first take $\lambda = 1$, and then take $\lambda = i$, to find that $\langle h, g\rangle$ and $\langle Vh, Vg\rangle$ have the same real and imaginary components. □

Note this implies that isomorphisms preserve completeness, so if an inner product space is isomorphic to a Hilbert space, then it must be complete (i.e. a Hilbert space in its own right).

Example:

Define $S : l^2 \to l^2$ by $S(\alpha_1, \alpha_2, \dots) = (0, \alpha_1, \alpha_2, \dots)$. Then $S$ is an isometry that is not surjective.

**Theorem 1.5.3** Two Hilbert spaces are isomorphic if and only if they have the same dimension.

***Proof*** If $U : \mathcal{H} \to \mathcal{K}$ is an isomorphism and $\mathcal{O}$ is a basis for $\mathcal{H}$, then $U\mathcal{O}$ is a basis for $\mathcal{K}$. Hence $\dim \mathcal{H} = \dim \mathcal{K}$.
Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{O}$ be a basis for $\mathcal{H}$. Consider the Hilbert space $l^2(\mathcal{O})$. If $h \in \mathcal{H}$, define $\hat{h} : \mathcal{O} \to \mathbb{F}$ by $\hat{h}(e) = \langle h, e\rangle$. By Parseval's Identity $\hat{h} \in l^2(\mathcal{O})$ and $||h|| = \left\|\hat{h}\right\|$. Define $U : \mathcal{H} \to l^2(\mathcal{O})$ by $Uh = \hat{h}$. Thus $U$ is linear and an isometry. The range of $U$ contains all the functions $f$ in $l^2(\mathcal{O})$ such that $f(e) = 0$ for all but a finite number of $e$; that is, the range is dense. But $U$, being an isometry, must have a closed range. Hence $U : \mathcal{H} \to l^2(\mathcal{O})$ is an isomorphism.
If $\mathcal{K}$ is a Hilbert space with a basis $\mathcal{J}$, $\mathcal{K}$ is isomorphic to $l^2(\mathcal{J})$. If $\dim \mathcal{H} = \dim \mathcal{K}$, $\mathcal{O}$ and $\mathcal{J}$ have the same cardinality; it follows that $l^2(\mathcal{O})$ is isomorphic to $l^2(\mathcal{J})$. Therefore $\mathcal{H}$ and $\mathcal{K}$ are isomorphic. □

**Corollary 1.5.4** All separable infinite dimensional Hilbert spaces are isomorphic.

Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

**Theorem 1.5.5** If $f : \partial\mathbb{D} \to \mathbb{C}$ is a continuous function, then there is a sequence $\{p_n(z, \overline{z})\}$ of polynomials in $z$ and $\overline{z}$ such that $p_n(z, \overline{z}) \to f(z)$ uniformly on $\partial\mathbb{D}$.

Note that if $z \in \partial\mathbb{D}, \overline{z} = z^{-1}$. Thus a polynomial in $z$ and $\overline{z}$ on $\partial\mathbb{D}$ becomes a Laurent polynomial

$$\sum_{k=-m}^{n} a_k z^k$$

If we put $z = e^{i\theta}$, this becomes

$$\sum_{k=-m}^{n} a_k e^{ik\theta}$$

Such functions are called **trigonometric polynomials**.

**Theorem 1.5.6** If for each $n \in \mathbb{Z}$, $e_n(t) \equiv (2\pi)^{-1}\exp(int)$, then $\{e_n : n \in \mathbb{Z}\}$ is a basis for $L^2_{\mathbb{C}}[0, 2\pi]$.

***Proof*** Let $\mathscr{T} = \left\{\sum_{k=-n}^{n} a_k e_k : a_k \in \mathbb{C}, n \geq 0\right\}$. Then $\mathscr{T}$ is a subalgebra of $C_{\mathbb{C}}[0, 2\pi]$, the algebra of all continuous $\mathbb{C}$-valued functions on $[0, 2\pi]$. Note that if $f \in \mathscr{T}$, $f(0) = f(2\pi)$. We want to show the uniform closure of $\mathscr{T}$ is $\mathscr{C} \equiv \{f \in C_{\mathbb{C}}[0, 2\pi] : f(0) = f(2\pi)\}$. To do this let $f \in \mathscr{C}$ and define $F : \partial\mathbb{D} \to \mathbb{C}$ by $F(e^{it}) = f(t)$. $F$ is continuous. By the previous result there is a sequence of polynomials in $z$ and $\overline{z}$, $p_n$, such that $p_n(z, \overline{z}) \to F(z)$ uniformly on $\partial\mathbb{D}$. Thus $p_n(e^{it}, e^{-it}) \to f(t)$ uniformly on $[0, 2\pi]$. But $p_n(e^{it}, e^{-it}) \in \mathscr{T}$.
Now the closure of $\mathscr{C}$ in $L^2_{\mathbb{C}}[0, 2\pi]$ is all of $L^2_{\mathbb{C}}[0, 2\pi]$. Hence, the closed span of $\{e_n : n \in \mathbb{Z}\}$ is $L^2_{\mathbb{C}}[0, 2\pi]$ and $\{e_n\}$ is thus a basis. □

Define $e_n(t) = \exp(int)$. Hence $\{e_n : n \in \mathbb{Z}\}$ is a basis for $L^2_{\mathbb{C}}([0, 2\pi], (2\pi)^{-1}dt)$. If $f \in \mathscr{H}$, then

$$\hat{f}(n) := \langle f, e_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt$$

is called the $n$**th Fourier coefficient** of $f$, $n \in \mathbb{Z}$. By the previous theorem

$$f = \sum_{n=-\infty}^{\infty} \hat{f}(n) e_n$$

where the convergence occurs in the norm induced by our inner product on the space. This is called the **Fourier series** of $f$.

If $\mathscr{H}$ is any Hilbert space and $\mathscr{O}$ is a basis, the scalars $\{\langle h, e \rangle : e \in \mathscr{O}\}$ are called the **Fourier coefficients** of $h$ relative to $\mathscr{O}$, and the series is called the **Fourier expansion of** $h$.

**Theorem 1.5.7** The Riemann-Lebesgue Lemma If $f \in L^2_{\mathbb{C}}[0, 2\pi]$, then $|int_0^{2\pi} f(t) e^{-int} dt \to 0$ as $n \to \pm\infty$.

For $f \in L^2_{\mathbb{C}}[0, 2\pi]$, the function $\hat{f} : \mathbb{Z} \to \mathbb{C}$ is called the **Fourier transform of** $f$. The map $U : L^2_{\mathbb{C}}[0, 2\pi] \to l^2(\mathbb{Z})$ defined by $Uf = \hat{f}$ is the **Fourier transform**.

**Theorem 1.5.8** The Fourier transform is a linear isometry from $L^2_{\mathbb{C}}[0, 2\pi]$ onto $l^2(\mathbb{Z})$.

## 1.6 Direct Sum of Hilbert Spaces

Throughout let $\mathscr{H}$ and $\mathscr{K}$ be Hilbert spaces.

**Definition 1.6.1** Define an inner product on $\mathscr{H} \oplus \mathscr{K}$ by

$$\langle h_1 \oplus k_1, h_2 \oplus k_2 \rangle := \langle h_1, h_2 \rangle + \langle k_1, k_2 \rangle$$

This is again a Hilbert space. However, we can lose completeness when switching to infinite direct sums.

**Proposition 1.6.2** If $\mathscr{H}_1, \mathscr{H}_2, \ldots$ are Hilbert spaces, let $\mathscr{H}$ be the subspace of their direct sum consisting of sequences $(h_n)$ such that $\sum_{n\geq 1} ||h_n|| < \infty$. For $h = (h_n)$ and $g = (g_n)$ in $\mathscr{H}$, define

$$\langle h, g\rangle = \sum_{n\geq 1} \langle h_n, g_n\rangle$$

Then $\langle \cdot, \cdot\rangle$ is an inner product on $\mathscr{H}$ and the norm relative to this inner product is the usual. WIth this inner product $\mathscr{H}$ is a Hilbert space.

***Proof*** If $h = (h_n), g = (g_n) \in \mathscr{H}$, then the CBS inequality implies

$$\sum | \langle h_n, g_n\rangle | \leq \sum ||h_n|| \, ||g_n|| \leq \left(\sum ||h_n||^2\right)^{1/2} \left(\sum ||g_n||^2\right)^{1/2} < \infty$$

Hence the series in the statement converges absolutely. □

**Definition 1.6.3** If $\mathscr{H}_1, \ldots$ are Hilbert spaces, the space $\mathscr{H}$ of the previous proposition is called the **direct sum** of $\mathscr{H}_1, \ldots$.

## Problems

**1.1** A given problem or Excercise is described here. The problem is described here. The problem is described here.

# Chapter 2

# Operators on Hilbert Spaces

**Abstract** Summary of material in chapter (to be completed after chapter)

## 2.1 Elementary Properties

**Proposition 2.1.1** Let $\mathscr{H}$ and $\mathscr{K}$ be Hilbert spaces and $A : \mathscr{H} \to \mathscr{K}$ a linear operator. The following are equivalent:

1. $A$ is continuous
2. $A$ is continuous at 0
3. $A$ is continuous at some point
4. There is a constant $c > 0$ such that $||Ah|| \leq c\,||h||$ for all $h \in \mathscr{H}$

The proof of this proposition is identical to the one for linear functionals seen in the previous chapter. We recall the equivalent definitions of the operator norm:

$$\begin{aligned} ||A|| &= \sup\{||Ah|| : h \in \mathscr{H}, ||h|| \leq 1\} \\ &= \sup\{||Ah|| : ||h|| = 1\} \\ &= \sup\{||Ah|| / ||h|| : h \neq 0\} \\ &= \inf\{c > 0 : ||Ah|| \leq c\,||h||\,, h \in \mathscr{H}\} \end{aligned}$$

Also, $||Ah|| \leq ||A||\,||h||$. Let $\mathscr{B}(\mathscr{H}, \mathscr{K})$ denote the set of bounded linear transformations from $\mathscr{H}$ into $\mathscr{K}$.

**Proposition 2.1.2**

(a) If $A, B \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then $A + B \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, and $||A + B|| \leq ||A|| + ||B||$
(b) If $\alpha \in \mathbb{F}$ and $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then $\alpha A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $||\alpha A|| = |\alpha|\,||A||$
(c) If $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $B \in \mathscr{B}(\mathscr{K}, \mathscr{L})$, then $BA \in \mathscr{B}(\mathscr{H}, \mathscr{L})$, and $||BA|| \leq ||B||\,||A||$.

***Proof*** For (a), observe that $||(A+B)h|| = ||Ah+Bh|| \leq ||Ah||+||Bh|| \leq ||A||\,||h||+||B||\,||h||$. As $A, B$ are bounded, $||A||\,, ||B|| < \infty$, so $||A||+||B|| < \infty$. Thus $A+B$ is bounded, and $||A+B|| \leq ||A||+||B||$ by definition of the infimum.
For (b), $||\alpha Ah|| = |\alpha|\,||Ah|| \leq |\alpha|\,||A||\,||h||$. Thus $\alpha A$ is bounded and $||\alpha A|| \leq |\alpha|\,||A||$. If $\alpha = 0$, $\alpha A = 0$ so $||\alpha A|| = 0 = |\alpha|\,||A||$. Otherwise, we can write $||A|| \leq \frac{1}{|\alpha|}\,||\alpha A||$, so $|\alpha|\,||A|| \leq ||\alpha A||$. Hence $||\alpha A|| = |\alpha|\,||A||$.
Finally, for (c) we have $||BAh|| \leq ||B||\,||Ah|| \leq ||B||\,||A||\,||h||$, so as $B$ and $A$ are both bounded, $||B||\,||A|| < \infty$, and so $BA$ is bounded with $||BA|| \leq ||B||\,||A||$. □

Thus the operator norm is indeed a norm on the vector space of bounded linear operators, and so we can define a metric $d(A,B) = ||A-B||$.

Example:

If $\dim \mathscr{H} = n < \infty$ and $\dim \mathscr{K} = m < \infty$, let $\{e_1, ..., e_n\}$ be an orthonormal basis for $\mathscr{H}$ and let $\{\epsilon_1, ..., \epsilon_m\}$ be an orthonormal basis for $\mathscr{K}$. Let $A$ be a linear transformation between these spaces, with associated matrix $(a_{i,j})$, with $a := \max\{|a_{i,j}|\}$. Then we have for $x = \sum_i x_i e_i$,

$$\begin{aligned}
||Ax||^2 &= \left\|\sum_i \sum_j a_{i,j} x_i \epsilon_j\right\|^2 \\
&= \left\|\sum_j \left(\sum_i a_{i,j} x_i\right) \epsilon_j\right\|^2 \\
&= \sum_j \left|\left(\sum_i a_{i,j} x_i\right)\right| ||\epsilon_j||^2 \\
&\leq \sum_j \sum_i |a_{i,j}|^2 |x_i|^2 \left\|\epsilon_j\right\|^2 \\
&\leq \sum_j \sum_i a^2 |x_i|^2 c^2 && (c = \max\{\left\|\epsilon_j\right\|\}) \\
&= mac^2 \sum_i |x_i|^2 \\
&\leq mac^2 \sum_i |x_i|^2 \frac{||e_i||^2}{d^2} && (d = \min\{||e_i||\}) \\
&= \frac{mac^2}{d} \sum_i ||x_i e_i||^2 = \frac{mac^2}{d}\, ||x||
\end{aligned}$$

Thus $A$ is bounded. Observe that $a_{i,j} = \langle Ae_j, \epsilon_i \rangle$.

Example:

let $l^2 = l^2(\mathbb{N})$ and let $e_1, e_2, ...$ be its usual basis. If $A \in \mathscr{B}(l^2)$, form $\alpha_{i,j} = \langle Ae_j, e_i \rangle$. The infinite matrix $(\alpha_{i,j})$ represents $A$ as finite matrices represent operators on finite dimensional spaces. However, this representation has limited value unless the matrix has a special form.

**Theorem 2.1.3** Let $(X, \Omega, \mu)$ be a $\sigma$-finite measure space and put $\mathscr{H} = L^2(X, \Omega, \mu) = L^2(\mu)$. If $\phi \in L^\infty(\mu)$, define $M_\phi : L^2(\mu) \to L^2(\mu)$ by $M_\phi f = \phi f$. Then $M_\phi \in \mathscr{B}(L^2(\mu))$, and $\left\|M_\phi\right\| = ||\phi||_\infty$.

***Proof*** Here $||\phi||_\infty$ is the $\mu$-essential supremum norm. Recall

$$\begin{aligned} ||\phi||_\infty &:= \inf\{\sup\{|\phi(x)| : x \in N\} : N \in \Omega : \mu(N) = 0\} \\ &= \inf\{c > 0 : \mu(\{x \in X : |\phi(x)| > c\}) = 0\} \end{aligned}$$

Thus $||\phi||_\infty$ is the infimum of all $c > 0$ such that $|\phi(x)| \leq c$ almost everywhere $[\mu]$ and, moreover, $|\phi(x)| \leq ||\phi||_\infty$ almost everywhere $[\mu]$. As $L^\infty(\mu)$ are equivalence classes of functions, we can assume that $\phi$ is a bounded measurable function and $|\phi(x)| \leq ||\phi||_\infty$ for all $x$. So if $f \in L^2(\mu)$, then

$$\int_X |\phi f|^2 d\mu \leq ||\phi||_\infty^2 \int_X |f|^2 d\mu$$

That is $M_\phi \in \mathscr{B}(L^2(\mu))$, and $||M||_\phi \leq ||\phi||_\infty$. If $\epsilon > 0$, the $\sigma$-finiteness of the measure space implies that there is a set $\Delta$ in $\Omega$, $0 < \mu(\Delta) < \infty$, such that $|\phi(x)| \leq ||\phi||_\infty - \epsilon$ on $\Delta$. If $f = (\mu(\delta))^{-1/2}\chi_\Delta$, then $f \in L^2(\mu)$ and $||f||_2 = 1$. So

$$\left\|M_\phi\right\|^2 \geq ||\phi f||_2^2 = (\mu(\Delta))^{-1} \int_\delta |\phi|^2 d\mu \geq (||\phi||_\infty - \epsilon)^2$$

Letting $\epsilon \to 0$, we get that $\left\|M_\phi\right\| \geq ||\phi||_\infty$. □

The operator $M_\phi$ is called a **multiplication operator**.

If the measure space $(X, \Omega, \mu)$ is not $\sigma$-finite, then the conclusion of the theorem need not hold.

**Theorem 2.1.4** Let $(X, \Omega, \mu)$ be a measure space and suppose $k : X \times X \to \mathbb{F}$ is an $\Omega \times \Omega$-measurable function for which there are constants $c_1$ and $c_2$ such that

$$\begin{aligned} \int_X |k(x, y)| d\mu(y) &\leq c_1, \text{ a.e.}[\mu] \\ \int_X |k(x, y)| d\mu(x) &\leq c_2, \text{ a.e.}[\mu] \end{aligned}$$

If $K : L^2(\mu) \to L^2(\mu)$ is defined by

$$(Kf)(x) = \int_X k(x, y) f(y) d\mu(y)$$

then $K$ is a bounded linear operator and $||K|| \leq (c_1 c_2)^{1/2}$.

***Proof*** If $f \in L^2(\mu)$, observe that

$$
\begin{aligned}
|Kf(x)| &\leq \int_X |k(x,y)||f(y)|d\mu(y) \\
&= \int_X |k(x,y)|^{1/2}|k(x,y)|^{1/2}|f(y)|d\mu(y) \\
&\leq \left[\int_X |k(x,y)|d\mu(y)\right]^{1/2}\left[\int_X |k(x,y)||f(y)|^2 d\mu(y)\right]^{1/2} && \text{(by Hölder's inequality)} \\
&\leq c_1^{1/2}\left[\int_X |k(x,y)||f(y)|^2 d\mu(y)\right]^{1/2}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\int_X |Kf(x)|^2 d\mu(x) &\leq c_1 \int_X \int_X |k(x,y)||f(y)|^2 d\mu(y) d\mu(x) \\
&= c_1 \int_X |f(y)|^2 \int_X |k(x,y)| d\mu(x) d\mu(y) && \text{(by Fubini's Theorem)} \\
&\leq c_1 c_2 \, ||f||^2 < \infty
\end{aligned}
$$

This shows that $Kf \in L^2(\mu)$, the formula defining $Kf$ is finite a.e. $[\mu]$, and $||Kf||^2 \leq c_1 c_2 \, ||f||^2$, so $||K|| \leq \sqrt{c_1 c_2}$. □

The operator described above is called an **integral operator** and the function $k$ is called its **kernel**.

Example:

Let $k : [0,1] \times [0,1] \to \mathbb{R}$ be the characteristic function of $\{(x,y) : y < x\}$. The corresponding operator $V : L^2(0,1) \to L^2(0,1)$ defined by $Vf(x) = \int_0^1 k(x,y)f(y)dy$ is called the **Volterra operator**. Note that

$$Vf(x) = \int_0^x f(y)dy$$

Note that any isometry is a bounded operator with norm 1.

## 2.2 The Adjoint of an Operator

**Definition 2.2.1** If $\mathscr{H}$ and $\mathscr{K}$ are Hilbert spaces, a function $u : \mathscr{H} \times \mathscr{K} \to \mathbb{F}$ is a **sesquilinear form** if for $h, g \in \mathscr{H}$, $k, f \in \mathscr{K}$, and $\alpha, \beta \in \mathbb{F}$,

(a) $u(\alpha h + \beta g, k) = \alpha u(h,k) + \beta u(g,k)$
(b) $u(h, \alpha k + \beta f) = \overline{\alpha} u(h,k) + \overline{\beta} u(h,f)$

The prefix “sesqui" is used becuase the function is linear in one variable but only conjugate linear in the other.

A sesquilinear form is **bounded** if there is a constant $M$ such that $|u(h,k)| \leq M\,||h||\,||k||$ for all $h \in \mathscr{H}$ and $k \in \mathscr{K}$. The constant $M$ is called a **bound for** $u$.

If $A \in \mathscr{B}(\mathscr{H}, \mathscr{K}$ and $B \in \mathscr{B}(\mathscr{K}(\mathscr{H})$, then both $\langle Ah, k\rangle$ and $\langle h, Bk\rangle$ are bounded sesquilinear forms.

**Theorem 2.2.2** If $u : \mathscr{H} \times \mathscr{K} \to \mathbb{F}$ is a bounded sesquilinear form with bound $M$, then there are unique operators $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ such that

$$u(h,k) = \langle Ah, k\rangle = \langle h, Bk\rangle$$

for all $h \in \mathscr{H}, k \in \mathscr{K}$, and $||A||\,, ||B|| \leq M$.

***Proof*** For each $h \in \mathscr{H}$, define $L_h : \mathscr{K} \to \mathbb{F}$ by $L_h(k) = \overline{u(h,k)}$. Then $L_h$ is linear and $L_h(k) \leq M\,||h||\,||k||$. By the Riesz Representation THeorem there is a unique vector $f \in \mathscr{K}$ such that $\langle k, f\rangle = \overline{L_h(k)} = u(h,k)$ and $||f|| \leq M\,||h||$. Let $Ah = f$. By the uniqueness part of the Riesz Theorem $A$ is linear. Also, $\langle Ah, k\rangle = \overline{\langle k, Ah\rangle} = \overline{k f} = u(h,k)$.
The proof for $B$ is similar. If $A_1 \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, and $u(h,k) = \langle A_1 h, k\rangle$, then $\langle Ah - A_1 h, k\rangle = 0$ for all $k$, so $Ah - A_1 h$ for all $h$. Thus, $A$ is unique. □

**Definition 2.2.3** If $A \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then the unique operator $B \in \mathscr{B}(\mathscr{K}, \mathscr{H})$ satisfying the equality $\langle Ah, k\rangle = \langle h, Bk\rangle$ for all $h, k$ is called the **adjoint** of $A$, and is denoted by $B = A^*$.

**Proposition 2.2.4** If $U \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, then $U$ is an isomorphism if and only if $U$ is invertible and $U^{-1} = U^*$.

***Proof*** If $U$ is an isomorphism we know that $U$ is invertible and $\langle Uh, Uh'\rangle = \langle h, h'\rangle$ for all $h, h' \in \mathscr{H}$. Then we have that $\langle h, U^*Uh'\rangle = \langle h, h'\rangle$ for all $h, h'$, so $U^*U = \mathrm{id}_{\mathscr{H}}$. As $U$ is invertible $U^*$ must be its inverse by uniqueness.
The reverse implication is immediate from the observation $\langle Uh, Uh'\rangle = \langle h, U^*Uh'\rangle = \langle h, h'\rangle$. □

**Proposition 2.2.5** If $A, B \in \mathscr{B}(\mathscr{H})$, and $\alpha \in \mathbb{F}$, then:

(a) $(\alpha A + B)^* = \overline{\alpha} A^* + B^*$
(b) $(AB)^* = B^* A^*$
(c) $A^{**} = (A^*)^* = A$.
(d) If $A$ is invertible in $\mathscr{B}(\mathscr{H})$ and $A^{-1}$ is its inverse, then $A^*$ is invertible and $(A^*)^{-1} = (A^{-1})^*$.

***Proof*** For (a), observe that

$$\langle(\alpha A + B)h, k\rangle = \alpha\,\langle Ah, k\rangle + \langle Bh, k\rangle = \alpha\,\langle h, A^*k\rangle + \langle h, B^*k\rangle = \langle h, (\overline{\alpha} A^* + B^*)k\rangle$$

By uniqueness $(\alpha A + B)^* = \overline{\alpha} A^* + B^*$.
For (b), observe

$$\langle ABh, k\rangle = \langle Bh, A^*k\rangle = \langle h, B^*A^*k\rangle$$

so by uniqueness $(AB)^* = B^*A^*$.

For (c),

$$\langle A^*h, k\rangle = \overline{\langle k, A^*h\rangle} = \overline{\langle Ak, h\rangle} = \langle h, Ak\rangle$$

so $(A^*)^* = A$ by uniqueness.
Finally, for (d) we have

$$\left\langle (A^{-1})^* A^* h, k\right\rangle = \left\langle A^* h, A^{-1}k\right\rangle = \left\langle h, AA^{-1}k\right\rangle = \langle h, k\rangle$$

so by uniqueness $(A^{-1})^* A^* = \mathrm{id}_{\mathscr{H}}$. Similarly, we have

$$\left\langle A^*(A^{-1})^* h, k\right\rangle = \left\langle (A^{-1})^* h, Ak\right\rangle = \left\langle h, A^{-1}Ak\right\rangle = \langle h, k\rangle$$

so $A^*(A^{-1})^* = \mathrm{id}_{\mathscr{H}}$. Thus $A^*$ is invertible and $(A^*)^{-1} = (A^{-1})^*$. □

**Proposition 2.2.6** If $A \in \mathscr{B}(\mathscr{H})$, $||A|| = ||A^*|| = ||A^*A||^{1/2}$.

***Proof*** For $h \in \mathscr{H}$, $||h|| \leq 1$,

$$||Ah||^2 = \langle Ah, Ah\rangle = \langle A^*Ah, h, \leq\rangle\, ||A^*Ah||\, ||h|| \leq ||A^*A|| \leq ||A^*||\, ||A||$$

Hence $||A||^2 \leq ||A^*A|| \leq ||A^*||\, ||A||$. Using the two ends of this string of inequalities $||A|| \leq ||A^*||$. But $A = (A^*)^*$, and so if $A^*$ is substituted for $A$ we get $||A|| = ||A^*||$. Thus the string of inequalities becomes a string of equalities and the proof is complete. □

Example:
Let $(X, \Omega, \mu)$ be a $\sigma$-finite measure space and let $M_\phi$ be the multiplication operator with symbol $\phi$. Then $M_\phi^*$ is $M_{\overline{\phi}}$, the multiplication operator with symbol $\overline{\phi}$.

If an operator on $\mathbb{F}^d$ is presented by a matrix, then its adjoint is represented by the conjugate transpose of the matrix.

Example:
If $K$ is the integral operator with kernel $k$, then $K^*$ is the integral operator with kernel $k^*(x, y) = \overline{k(x, y)}$.

**Proposition 2.2.7** IF $S : l^2 \to l^2$ is defined by $S(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots)$, then $S$ is an isomotry and $S^*(\alpha_1, \alpha_2, \ldots) = (\alpha_2, \alpha_3, \ldots)$.

***Proof*** We have already shown that $S$ is an isometry. For $(\alpha_n), (\beta_n) \in l^2$,

$$\langle S^*(\alpha_n), (\beta_n)\rangle = \langle (\alpha_n), S(\beta_n), =\rangle\, \alpha_2\overline{\beta}_1 + \alpha_3\overline{\beta}_2 + \cdots = \langle (\alpha_2, \alpha_3, \ldots), (\beta_n)\rangle$$

and so the result follows by uniqueness. □

The operator $S$ is called the **unilateral shift** and the operator $S^*$ is called the **backward shift**.

**Definition 2.2.8** If $A \in \mathscr{B}(\mathscr{H})$, then

(a) is **hermitian** or **self-adjoint** if $A^* = A$;
(b) $A$ is **normal** if $A^*A = AA^*$.

Notice that hermitian and unitary operators are normal. In light of the previous examples, every multiplication operator $M_\phi$ is normal; $M_\phi$ is hermitian if and only if $\phi$ is real-valued; $M_\phi$ is unitary if and only if $|\phi| \equiv 1$ almost everywhere $[\mu]$. Additionally, an integral operator $K$ with kernel $k$ is hermitian if and only if $k(x, y) = \overline{k(y, x)}$ almost everywhere $[\mu \times \mu]$.

**Proposition 2.2.9** If $\mathscr{H}$ is a $\mathbb{C}$-Hilbert space and $A \in \mathscr{B}(\mathscr{H})$, then $A$ is hermitian if and only if $\langle Ah, h\rangle \in \mathbb{R}$ for all $h \in \mathscr{H}$.

***Proof*** If $A = A^*$, then $\langle Ah, h\rangle = \langle h, Ah\rangle = \overline{\langle Ah, h\rangle}$; hence $\langle Ah, h\rangle \in \mathbb{R}$.
For the converse assume $\langle Ah, h\rangle$ is real for every $h \in \mathscr{H}$. If $\alpha \in \mathbb{C}$ and $h, g \in \mathscr{H}$, then

$$\langle A(h + \alpha g), h + \alpha g\rangle = \langle Ah, h, +\rangle\, \overline{\alpha}\, \langle Ah, g\rangle + \alpha \langle Ag, h\rangle + |\alpha|^2 \langle Ag, g\rangle \in \mathbb{R}$$

So this expression equals its complex conjugate. Using the fact that $\langle Ah, h\rangle$ and $\langle Ag, g\rangle \in \mathbb{R}$ yields

$$\begin{aligned} \alpha \langle Ag, h\rangle + \overline{\alpha} \langle Ah, g\rangle &= \overline{\alpha} \langle h, Ag\rangle + \alpha \langle g, Ah\rangle \\ &= \overline{\alpha} \langle A^*h, g\rangle + \alpha \langle A^*g, h\rangle \end{aligned}$$

By first taking $\alpha = 1$ and then $\alpha = i$, we obtain

$$\begin{aligned} \langle Ag, h\rangle + \langle Ah, g\rangle &= \langle A^*h, g\rangle + \langle A^*g, h\rangle \\ i \langle Ag, h\rangle - i \langle Ah, g\rangle &= -i \langle A^*h, g\rangle + i \langle A^*g, h\rangle \end{aligned}$$

Multiplying the second equation by $i$ and adding the two we obtain

$$2 \langle Ah, g\rangle = 2 \langle A^*h, g\rangle$$

As this holds for all $h, g$, we find $A = A^*$. □

The preceding proposition is false if it is only assumed that $\mathscr{H}$ is an $\mathbb{R}$-Hilbert space.

**Proposition 2.2.10** If $A = A^*$ then

$$||A|| = \sup\{|\langle Ah, h\rangle| : ||h|| = 1\}$$

***Proof*** Put $M = \sup\{|\langle Ah, h\rangle| : ||h|| = 1\}$. If $||h|| = 1$, then $|\langle Ah, h\rangle| \le ||Ah||\,||h|| \le ||A||$; hence $M \le ||A||$. On the other hand, if $||h|| = ||g|| = 1$, then

$$\begin{aligned} \langle A(h \pm g), h \pm g\rangle &= \langle Ah, h\rangle \pm \langle Ah, g\rangle \pm \langle Ag, h\rangle + \langle Ag, g\rangle \\ &= \langle Ah, h\rangle \pm \langle Ah, g\rangle \pm \langle g, Ah\rangle + \langle Ag, g\rangle \qquad \text{(since } A = A^*\text{)} \\ &= \langle Ah, h\rangle \pm 2\mathrm{Re} \langle Ah, g\rangle + \langle Ag, g\rangle \end{aligned}$$

Subtracting one of these equations from the other gives

$$4\mathrm{Re}\,\langle Ah, g\rangle = \langle A(h+g), h+g\rangle - \langle A(h-g), h-g\rangle$$

Now it is follows by linearity that $|\langle Af, f\rangle| \le M\,||f||^2$ for any $f \in \mathscr{H}$. Hence, using the parallelogram law we get

$$\begin{aligned} 4\mathrm{Re}\,\langle Ah, g\rangle &\le M(||h+g||^2 + ||h-g||^2) \\ &= 2M(||h||^2 + ||g||^2) \\ &= 4M \end{aligned}$$

since $h$ and $g$ are unit vectors. Now suppose $\langle Ah, g\rangle = e^{i\theta}|\langle Ah, g\rangle|$. Replacing $h$ in the inequality above with $e^{-i\theta}h$ gives $|\langle Ah, g\rangle| \le M$ if $||h|| = ||g|| = 1$. Taking the supremum over all $g$ gives $||Ah|| \le M$ when $||h|| = 1$. Thus $||A|| \le M$. □

**Corollary 2.2.11** If $A = A^*$ and $\langle Ah, h\rangle = 0$ for all $h$, then $A = 0$.

The preceding corollary is not true unless $A = A^*$. However, if a complex Hilbert space is present, this hypothesis can be deleted.

**Proposition 2.2.12** If $\mathscr{H}$ is a $\mathbb{C}$-Hilbert space and $A \in \mathscr{B}(\mathscr{H})$, such that $\langle Ah, h\rangle = 0$ for all $h \in \mathscr{H}$, then $A = 0$.

***Proof*** Suppose $\langle Ah, h\rangle = 0$ for all $h$. Then for $h, g$,

$$0 = \langle A(h \pm ig), h \pm ig\rangle = \langle Ah, h\rangle \pm i\,\langle Ag, h\rangle \mp i\,\langle Ah, g\rangle - \langle Ag, g\rangle = \pm i\,\langle Ag, h\rangle \mp i\,\langle Ah, g\rangle$$

Thus $\pm i\,\langle Ag, h\rangle = \pm i\,\langle Ah, g\rangle$. Hence $\langle Ag, h\rangle = \langle Ah, g\rangle$. Further,

$$0 = \langle A(h \pm g), h \pm g\rangle = \pm\langle Ag, h\rangle \pm \langle Ah, g\rangle$$

so $\langle Ag, h\rangle = -\langle Ah, g\rangle = -\langle Ag, h\rangle$, so $\langle Ag, h\rangle = 0$. As this holds for all $g, h$, $A = 0$. □

If $\mathscr{H}$ is a $\mathbb{C}$-Hilbert space and $A \in \mathscr{B}(\mathscr{H})$, then $B = (A + A^*)/2$ and $C = (A - A^*)/2i$ are self-adjoint and $A = B + iC$. The operators $B$ and $C$ are called the **real and imaginary parts of** $A$.

**Proposition 2.2.13** If $A \in \mathscr{B}(\mathscr{H})$, the following statements are equivalent.

(a) $A$ is normal.
(b) $||Ah|| = ||A^*h||$ for all $h$.
(c) If $\mathscr{H}$ is also a $\mathbb{C}$-Hilbert space, these are equivalent to the real and imaginary parts of $A$ commuting.

***Proof*** If $h \in \mathscr{H}$, then

$$||Ah||^2 - ||A^*h||^2 = \langle Ah, Ah\rangle - \langle A^*h, A^*h\rangle = \langle (A^*A - AA^*)h, h\rangle$$

Since $A^*A - AA^*$ is hermitian, the equivalence of (a) and (b) follow from the corollary.
If $B, C$ are the real and imaginary parts of $A$, then a calculation yields

$$A^*A = B^2 - iCB + iBC + C^2$$
$$AA^* = B^2 + iCB - iBC + C^2$$

Hence $A^*A = AA^*$ if and only if $2iCB = 2iBC$, if and only if $CB = BC$. □

**Proposition 2.2.14** If $A \in \mathscr{B}(\mathscr{H})$, the following statements are equivalent.

(a) $A$ is an isometry.
(b) $A^*A = I$
(c) $\langle Ah, Ag\rangle = \langle h, g\rangle$ for all $h, g \in \mathscr{H}$

***Proof*** The equivalence of (a) and (c) was shown in Chapter 1. Note that if $h, g \in \mathscr{H}$, then $\langle A^*Ah, g\rangle = \langle Ah, Ag\rangle$. Hence (b) and (c) are equivalent. □

**Proposition 2.2.15** If $A \in \mathscr{B}(\mathscr{H})$, then the following statements are equivalent.

(a) $A^*A = AA^* = I$
(b) $A$ is unitary
(c) $A$ is a normal isometry

***Proof*** (a) implies $A$ is invertible and an isometry by the previous result, so we have (b). For (b) implies (c) observe that $A^*A = I$ since $A$ is an isometry. But, $A$ is invertible so by uniqueness of the inverse $A^* = A^{-1}$ and $A^*A = AA^* = I$. Thus $A$ is normal.
Finally, for (c) implies (a), $A^*A = I$ since $A$ is an isometry, and as $A$ is also normal $AA^* = A^*A = I$. □

**Theorem 2.2.16** If $A \in \mathscr{B}(\mathscr{H})$, then $\ker A = (\operatorname{Im} A^*)^\perp$.

***Proof*** If $h \in \ker A$ and $g \in \mathscr{H}$, then $\langle h, A^*g\rangle = \langle Ah, g\rangle = 0$, so $\ker A \subseteq (\operatorname{Im} A^*)^\perp$. On the other hand, if $h \perp \operatorname{Im} A^*$ and $g \in \mathscr{H}$, then

$$\langle Ah, g\rangle = \langle h, A^*g\rangle = 0$$

so $(\operatorname{Im} A^*)^\perp \subseteq \ker A$. □

Since $A^{**} = A$, it also holds that $\ker A^* = (\operatorname{Im} A)^\perp$. Second, $(\ker A)^\perp \neq \operatorname{Im} A^*$ in general, since $\operatorname{Im} A^*$ may not be closed. All that can be said is $(\ker A)^\perp = \operatorname{cl}(\operatorname{Im} A^*)$ and $(\ker A^*)^\perp = \operatorname{cl}(\operatorname{Im} A)$.

## 2.3 Projections and Idempotents

**Definition 2.3.1** An **idempotent** on $\mathscr{H}$ is a bounded linear operator $E$ on $\mathscr{H}$ such that $E^2 = E$. A **projection** is an idempotent $P$ such that $\ker P = (\operatorname{Im} P)^{\perp}$.

If $\mathcal{M} \leq \mathscr{H}$, then $P_{\mathcal{M}}$ is a projection.

Let $E$ be any idempotent and set $\mathcal{M} = \operatorname{Im} E$, and $\mathcal{N} = \ker E$. Since $E$ is continuous, $\mathcal{N}$ is a closed subspace of $\mathscr{H}$. Notice that $(I - E)^2 = I - 2E + E^2 = I - 2E + E = I - E$; thus $I - E$ is also an idempotent. Also $0 = (I - E)h = h - Eh$ if and only if $Eh = h$, so $\operatorname{Im} E \supseteq \ker(I - E)$. On the other hand, if $h \in \operatorname{Im} E$, $h = Eg$, and so $Eh = E^2 g = Eg = h$; hence $\operatorname{Im} E = \ker(I - E)$. Similarly, $\operatorname{Im}(I - E) = \ker E$.

**Proposition 2.3.2**

(a) $E$ is an idempotent if and only if $I - E$ is an idempotent
(b) $\operatorname{Im} E = \ker(I - E)$, $\ker E = \operatorname{Im}(I - E)$, and both $\operatorname{Im} E$ and $\ker E$ are closed linear subspaces of $\mathscr{H}$.
(c) If $\mathcal{M} = \operatorname{Im} E$ and $\mathcal{N} = \ker E$, then $\mathcal{M} \cap \mathcal{N} = (0)$ and $\mathcal{M} + \mathcal{N} = \mathscr{H}$.

If $\mathcal{M}, \mathcal{N} \leq \mathscr{H}$, $\mathcal{M} \cap \mathcal{N} = (0)$, and $\mathcal{M} + \mathcal{N} = \mathscr{H}$, then there is an idempotent $E$ such that $\mathcal{M} = \operatorname{Im} E$ and $\mathcal{N} = \ker E$; moreover, $E$ is unique.

**Proposition 2.3.3** If $E$ is an idempotent on $\mathscr{H}$ and $E \neq 0$, the following statements are equivalent.

(a) $E$ is a projection
(b) $E$ is the orthogonal projection of $\mathscr{H}$ onto $\operatorname{Im} E$
(c) $||E|| = 1$
(d) $E$ is hermitian
(e) $E$ is normal
(f) $\langle Eh, h\rangle \geq 0$ for all $h \in \mathscr{H}$.

***Proof*** For (a) implies (b) let $\mathcal{M} = \operatorname{Im} E$ and $P = P_{\mathcal{M}}$. If $h \in \mathscr{H}$, $Ph =$ the unique vector in $\mathcal{M}$ such that $h - Ph \in \mathcal{M}^{\perp} = (\operatorname{Im} E)^{\perp} = \ker E$ by (a). But $h - Eh = (I - E)h \in \ker E$.Hence $Eh = Ph$ by uniqueness.
For (b) implies (c) we have $||E|| \leq 1$. But $Eh = h$ for $h \in \operatorname{Im} E$, so $||E|| = 1$.
For (c) implies (a) let $h \in (\ker E)^{\perp}$. Now $\operatorname{Im}(I - E) = \ker E$, so $h - Eh \in \ker E$. Hence $0 = \langle h - Eh, h\rangle = ||h||^2 - \langle Eh, h\rangle$. Hence $||h||^2 = \langle Eh, h\rangle \leq ||Eh||\,||h|| \leq ||h||^2$. So for $h \in (\ker E)^{\perp}$, $||Eh|| = ||h|| = \langle Eh, h\rangle^{1/2}$. But then for $h \in (\ker E)^{\perp}$,

$$||h - Eh||^2 = ||h||^2 - 2\operatorname{Re}\langle Eh, h\rangle + ||Eh||^2 = 0$$

That is $(\ker E)^{\perp} \subseteq \ker(I - E) = \operatorname{Im} E$. On the other hand, if $g \in \operatorname{Im} E$, $g = g_1 + g_2$, where $g_1 \in \ker E$ and $g_2 \in (\ker E)^{\perp}$. Thus $g = Eg = Eg_2 = g_2$; that is, $\operatorname{Im} E \subseteq (\ker E)^{\perp}$. Therefore $\operatorname{Im} E = (\ker E)^{\perp}$ and $E$ is a projection.
For (b) implies (f) If $h \in \mathscr{H}$, write $h = h_1 + h_2$, $h_1 \in \operatorname{Im} E$ and $h_2 \in \ker E = (\operatorname{Im} E)^{\perp}$. Hence

$$\langle Eh, h\rangle = \langle E(h_1 + h_2), h_1 + h_2\rangle = \langle Eh_1, h_1\rangle = \langle h_1, h_1\rangle = ||h_1||^2 \geq 0$$

.
For (f) implies (a) let $h_1 \in \operatorname{Im} E$ and $h_2 \in \ker E$. Then

$$0 \leq \langle E(h_1 + h_2), h_1 + h_2 \rangle = \langle h_1, h_1 \rangle + \langle h_1, h_2 \rangle$$

Hence $-\|h_1\|^2 \leq \langle h_1, h_2 \rangle$. If there are such $h_1, h_2$ with $\langle h_1, h_2 \rangle = \overline{\alpha} \neq 0$, then substituting $k_2 = -2\alpha^{-1} \|h_1\|^2 h_2$ for $h_2$ in this inequality, we obtain $-\|h_1\|^2 \leq -2\|h_1\|^2$, a contradiction. Hence $\langle h_1, h_2 \rangle = 0$ whenever $h_1 \in \operatorname{Im} E$ and $h_2 \in \ker E$. That is, $E$ is a projection.
For (a) implies (d) let $h, g \in \mathscr{H}$ and put $h = h_1 + h_2$ and $g = g_1 + g_2$, where $h_1, h_2 \in \operatorname{Im} E$ and $h_2, g_2 \in \ker E = (\operatorname{Im} E)^\perp$. Hence $\langle Eh, g \rangle = \langle h_1, g_1 \rangle$. Also, $\langle E^* h, g \rangle = \langle h, Eg \rangle = \langle h_1, g_1, = \rangle \langle Eh, g \rangle$. Thus $E = E^*$.
(d) implies (e) is always true for a hermitian operator.
Finally, for (e) implies (a), recall that $E$ being normal implies $\|Eh\| = \|E^* h\|$ for all $h$. Hence $\ker E = \ker E^*$. But $\ker E^* = (\operatorname{Im} E)^\perp$, so $E$ is a projection. □

**Definition 2.3.4** If $\{\mathscr{M}_i\}$ is a collection of pairwise orthogonal subspaces of $\mathscr{H}$, then

$$\bigoplus_i \mathscr{M}_i := \wedge_i \mathscr{M}_i$$

If $\mathscr{M}, \mathscr{N}$ are two closed linear subspaces of $\mathscr{H}$, then

$$\mathscr{M} \ominus \mathscr{N} := \mathscr{M} \cap \mathscr{N}^\perp$$

THis is called the **orthogonal difference** of $\mathscr{M}$ and $\mathscr{N}$.

Note that if $\mathscr{M}, \mathscr{N} \leq \mathscr{H}$ and $\mathscr{M} \perp \mathscr{N}$, then $\mathscr{M} + \mathscr{N}$ is closed.

**Definition 2.3.5** If $A \in \mathscr{B}(\mathscr{H})$ and $\mathscr{M} \leq \mathscr{H}$, say that $\mathscr{M}$ is an **invariant subspace** for $A$ if $Ah \in \mathscr{M}$ whenever $h \in \mathscr{M}$. In other words, if $A\mathscr{M} \subseteq \mathscr{M}$. Say that $\mathscr{M}$ is a **reducing subspace** for $A$ if $A\mathscr{M} \subseteq \mathscr{M}$ and $A\mathscr{M}^\perp \subseteq \mathscr{M}^\perp$.

If $\mathscr{M} \leq \mathscr{H}$ then $\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^\perp$. If $A \in \mathscr{B}(\mathscr{H})$, then $A$ can be written as a $2 \times 2$ matrix with operator entries

$$A = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$$

where $W \in \mathscr{B}(\mathscr{M}), X \in \mathscr{B}(\mathscr{M}^\perp, \mathscr{M}), Y \in \mathscr{B}(\mathscr{M}, \mathscr{M}^\perp)$, and $Z \in \mathscr{B}(\mathscr{M}^\perp)$.

**Proposition 2.3.6** If $A \in \mathscr{B}(\mathscr{H}), \mathscr{M} \leq \mathscr{H}$ and $P = P_{\mathscr{M}}$, then statements (a) though (c) are equivalent.

(a) $\mathscr{M}$ is invariant for $A$
(b) $PAP = AP$
(c) In the above matrix $Y = 0$.

Also, statements (d) through (g) are equivalent.

(d) $\mathscr{M}$ reduces $A$

(e) $PA = AP$.
(f) In the above matrix $Y$ and $X$ are 0.
(g) $\mathscr{M}$ is invariant for both $A$ and $A^*$.

***Proof*** For (a) implies (b), if $h \in \mathscr{H}$, $Ph \in \mathscr{M}$. So $APh \in \mathscr{M}$. Hence $P(APh) = APh$. That is, $PAP = AP$. For (b) implies (c), if $P$ is represented as a $2 \times 2$ matrix relative to $\mathscr{H} = \mathscr{M} \oplus \mathscr{M}^\perp$, then

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

Hence,

$$PAP = \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix} = AP = \begin{bmatrix} W & 0 \\ Y & 0 \end{bmatrix}$$

So $Y = 0$.
For (c) implies (a), if $Y = 0$ and $h \in \mathscr{M}$, then

$$Ah = \begin{bmatrix} W & X \\ 0 & Z \end{bmatrix} \begin{bmatrix} h \\ 0 \end{bmatrix} = \begin{bmatrix} Wh \\ 0 \end{bmatrix} \in \mathscr{M}$$

Next, for (d) implies (e) since both $\mathscr{M}$ and $\mathscr{M}^\perp$ are invariant for $A$, (b) implies that $AP = PAP$ and $A(I-P) = (I-P)A(I-P)$. Multiplying the second equation gives $A - AP = A - AP - PA + PAP$. Thus $PA = PAP = AP$.
for (e) implies (f) we have again the presentation of $P$ as above, so

$$PA = \begin{bmatrix} W & X \\ 0 & 0 \end{bmatrix} = AP = \begin{bmatrix} W & 0 \\ Y & 0 \end{bmatrix}$$

so $X$ and $Y$ are 0.
for (f) implies (g), if $X$ and $Y$ are 0, then

$$A = \begin{bmatrix} W & 0 \\ 0 & Z \end{bmatrix} \quad \text{and} \quad A^* = \begin{bmatrix} W^* & 0 \\ 0 & Z^* \end{bmatrix}$$

By (c), $\mathscr{M}$ is invariant for both $A$ and $A^*$.
For (g) implies (d), if $h \in \mathscr{M}^\perp$ and $g \in \mathscr{M}$, then $\langle g, Ah \rangle = \langle A^*g, h \rangle = 0$, sicne $A^*g \in \mathscr{M}$. Since $g$ was an arbitrary vector of $\mathscr{M}$, $Ah \in \mathscr{M}^\perp$. That is, $A\mathscr{M}^\perp \subseteq \mathscr{M}^\perp$. □

If $\mathscr{M}$ reduces $A$, then $X$ and $Y$ are 0 in our matrix decomposition. This says that a study of $A$ is reduced to the study of the smaller operators $W$ and $Z$. If $A \in \mathscr{B}(\mathscr{H})$ and $\mathscr{M}$ is an invariant subspace, then $||A|_{\mathscr{M}}|| \leq ||A||$, so the restriction is also bounded.

## 2.4 Compact Operators

Let $B_{\mathscr{H}}$ denote the closed unit ball in $\mathscr{H}$.

**Definition 2.4.1** A linera transformation $T : \mathscr{H} \to \mathscr{K}$ is **compact** if $T(B_{\mathscr{H}})$ has compact closure in $\mathscr{K}$. The set of compact operators from $\mathscr{H}$ into $\mathscr{K}$ is denoted by $\mathscr{B}_0(\mathscr{H}, \mathscr{K})$, and $\mathscr{B}_0(\mathscr{H}) = \mathscr{B}_0(\mathscr{H}, \mathscr{H})$.

**Proposition 2.4.2**

(a) $\mathscr{B}_0(\mathscr{H}, \mathscr{K}) \subseteq \mathscr{B}(\mathscr{H}, \mathscr{K})$

(b) $\mathscr{B}_0(\mathscr{H}, \mathscr{K})$ is a linear space and if $\{T_n\} \subseteq \mathscr{B}_0(\mathscr{H}, \mathscr{K})$ and $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ such that $||T_n - T|| \to 0$, then $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.

(c) If $A \in \mathscr{B}(\mathscr{H})$, $B \in \mathscr{B}(\mathscr{K})$, and $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, then $TA$ and $BT \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.

***Proof*** (a) If $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, then $\mathrm{cl}[T(B_{\mathscr{H}})]$ is compact in $\mathscr{K}$. Hence, there is a constant $C > 0$ such that $T(B_{\mathscr{H}}) \subseteq \{k \in \mathscr{K} : ||k|| \leq C\}$. Thus $||T|| \leq C$.
(b) Let $T, S \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$ and $\alpha \in \mathbb{F}$. If $\alpha = 0$ then $\alpha T = 0$, which is trivially compact. Otherwise, if $\alpha \neq 0$, multiplication by $\alpha$ is a homeomorphism on $\mathscr{K}$. Further, if $h \in \mathrm{cl}[\alpha T(B_{\mathscr{H}})]$, then for all $\epsilon > 0$ there exists $k \in B_{\mathscr{H}}$ such that $||h - \alpha T(k)|| < \epsilon$. This implies $h/\alpha \in \mathrm{cl}[T(B_{\mathscr{H}})]$, and so $h \in \alpha\mathrm{cl}[T(B_{\mathscr{H}})]$. By symmetry $\mathrm{cl}[\alpha T(B_{\mathscr{H}})] = \alpha\mathrm{cl}[T(B_{\mathscr{H}})]$, we have that the two spaces are homeomorphic, and so as one is compact so must the other be. Hence $\alpha T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.
Next, as $\mathscr{K}$ is complete it is sufficient to show that $(S + T)(B_{\mathscr{H}})$ is totally bounded. For $\epsilon > 0$, there exist $s_1, ..., s_n, t_1, ..., t_m$ such that $S(B_{\mathscr{H}}) \subseteq \bigcup_{i=1}^n B_{\epsilon/2}(s_i)$ and $T(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^m B_{\epsilon/2}(t_j)$. I claim $(S + T)(B_{\mathscr{H}}) \subseteq \bigcup_{i,j} B_\epsilon(s_i + t_j)$. Let $b \in B_{\mathscr{H}}$. Then there exist $i, j$ such that $S(b) \in B_{\epsilon/2}(s_i)$ and $T(b) \in B_{\epsilon/2}(t_j)$. Then

$$\left\|S(b) + T(b) - s_i - t_j\right\| \leq ||S(b) - s_i|| + \left\|T(b) - t_j\right\| < \epsilon$$

as desired. Therefore the closure of $(S + T)(B_{\mathscr{H}})$ is compact, so $S + T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.
Next, suppose $T_n$ and $T$ are as in (b). We again show totally boundedness for $T(B_{\mathscr{H}})$. Let $\epsilon > 0$ and choose $n$ such that $||T - T_n|| < \epsilon/3$. Since $T_n$ is compact, there are vectors $h_1, ..., h_m$ in $B_{\mathscr{H}}$ such that $T_n(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^m B_{\epsilon/3}(T_n h_j)$. So if $||h|| \leq 1$, there is an $h_j$ with $\left\|T_n h_j - T_n h\right\| < \epsilon/3$. Thus

$$\begin{aligned}\left\|T h_j - T h\right\| &\leq \left\|T h_j - T_n h_j\right\| + \left\|T_n h_j - T_n h\right\| + ||T_n h - T h|| \\ &< 2\,||T - T_n|| + \epsilon/3 \\ &< \epsilon\end{aligned}$$

Hence $T(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^m B_\epsilon(T h_j)$, completing the claim.
For (c), suppose $A, B$, and $T$ are as in the claim. As $A$ is bounded $A(B_{\mathscr{H}}) \subseteq ||A||\, B_{\mathscr{H}}$. But then $T(||A||\, B_{\mathscr{H}}) = ||A||\, T(B_{\mathscr{H}})$, which has compact closure from part (b). Thus as $\mathrm{cl}[TA(B_{\mathscr{H}})] \subseteq \mathrm{cl}[||A||\, T(B_{\mathscr{H}})]$, it follows that $\mathrm{cl}[TA(B_{\mathscr{H}})]$ is compact, so $TA$ is compact.
Finally, it is sufficient to show $BT(B_{\mathscr{H}})$ is totally bounded. If $B = 0$ then this is trivial. Otherwise, let $\epsilon > 0$, so there exist $b_1, ..., b_n \in B_{\mathscr{H}}$ such that $T(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^n B_{\epsilon/||B||}(b_j)$. Now

$$BT(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^n B(B_{\epsilon/||B||}(b_j)) \subseteq \bigcup_{j=1}^n ||B||\, B_{\epsilon/||B||}(b_j) = \bigcup_{j=1}^n B_\epsilon(b_j)$$

This completes the proof. □

**Definition 2.4.3** An operator $T$ on $\mathscr{H}$ has **finite rank** if $\operatorname{Im} T$ is finite dimensional. The set of continuous finite rank operators is denoted by $\mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$.

It is clear that $\mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$ is a linear space. It is also true that $\mathscr{B}_{00}(\mathscr{H}, \mathscr{K}) \subseteq \mathscr{B}_0(\mathscr{H}, \mathscr{K})$ since **TBC**

**Theorem 2.4.4** If $T \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, the following statements are equivalent.

(a) $T$ is compact
(b) $T^*$ is compact
(c) There is a sequence $\{T_n\}$ of operators of finite rank such that $||T - T_n|| \longrightarrow 0$.

***Proof*** (c) implies (a) follows from the previous result and the fact that $\mathscr{B}_{00}(\mathscr{H}, \mathscr{K}) \subseteq \mathscr{B}_0(\mathscr{H}, \mathscr{K})$.
For (a) implies (c), since $\mathrm{cl}[T(B_{\mathscr{H}})]$ is compact, it is separable. Therefore, $\mathrm{cl}(\operatorname{Im} T) = \mathscr{L}$ is a separable subspace of $\mathscr{K}$. Let $\{e_1, e_2, \ldots\}$ be a basis for $\mathscr{L}$ and let $P_n$ be the orthogonal projection of $\mathscr{K}$ onto $\wedge\{e_j : 1 \leq j \leq n\}$. Put $T_n = P_n T$; note that each $T_n$ has finite rank. It will be shown that $||T_n - T|| \longrightarrow 0$.
**Claim.** If $h \in \mathscr{H}$, $||T_n h - Th|| \longrightarrow 0$.
In fact, $k = Th \in \mathscr{L}$, so $||P_n k - k|| \longrightarrow 0$ from our previous results, and so $||P_n Th - Th|| \longrightarrow 0$ and the claim is proved.
Since $T$ is compact, if $\epsilon > 0$, there are vectors $h_1, \ldots, h_m \in B_{\mathscr{H}}$ such that $T(B_{\mathscr{H}}) \subseteq \bigcup_{j=1}^m B_{\epsilon/3}(Th_j)$. So if $||h|| \leq 1$, choose $h_j$ with $\left\|Th - Th_j\right\| < \epsilon/3$. Thus for any integer $n$,

$$\begin{aligned}||Th - T_n h|| &\leq \left\|Th - Th_j\right\| + \left\|Th_j - T_n h_j\right\| + \left\|P_n(Th_j - Th)\right\| \\ &\leq 2\left\|Th - Th_j\right\| + \left\|Th_j - T_n h_j\right\| \\ &\leq 2\epsilon/3 + \left\|Th_j - T_n h_j\right\|\end{aligned}$$

Using the claim we can find $n_0$ such that $\left\|Th_j - T_n h_j\right\| < \epsilon/3$ for $1 \leq j \leq m$ and $n \geq n_0$. So $||Th - T_n h|| < \epsilon$ uniformly for $h \in B_{\mathscr{H}}$. Therefore, $||T - T_n|| < \epsilon$ for $n \geq n_0$.
For (c) implies (b), if $\{T_n\}$ is a sequence in $\mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$ such that $||T_n - T|| \longrightarrow 0$, then $\left\|T_n^* - T^*\right\| = ||T_n - T|| \longrightarrow 0$. But $T_n^* \in \mathscr{B}_{00}(\mathscr{H}, \mathscr{K})$ (TBC). Since (c) implies (a), $T^*$ is compact.
(b) implies (a) (TBC). □

**Corollary 2.4.5** If $T \in \mathscr{B}_0(\mathscr{H}, \mathscr{K})$, then $\mathrm{cl}(\operatorname{Im} T)$ is separable and if $\{e_n\}$ is a basis for $\mathrm{cl}(\operatorname{Im} T)$ and $P_n$ is the projection of $\mathscr{K}$ onto $\wedge\{e_j : 1 \leq j \leq n\}$, then $||P_n T - T|| \longrightarrow 0$.

**Proposition 2.4.6** Let $\mathscr{H}$ be a separable Hilbert space with basis $\{e_n\}$; let $\{\alpha_n\} \subseteq \mathbb{F}$ with $M = \sup\{|\alpha_n| : n \geq 1\} < \infty$. If $Ae_n = \alpha_n e_n$ for all $n$, then $A$ extends by linearity to a bounded operator on $\mathscr{H}$ with $||A|| = M$. The operator $A$ is compact if and only if $\alpha_n \longrightarrow 0$ as $n \longrightarrow \infty$.

***Proof*** By definition we have that $||A|| \geq M$. Further, if $h \in \mathscr{H}$ with $||h|| = 1$ we have that $h = \sum_n \langle h, e_n \rangle e_n$ and $1 = ||h||^2 = \sum_n |\langle h, e_n \rangle|^2$. Then $Ah_k = \sum_{n=1}^k \alpha_n \langle h, e_n \rangle e_n$. As $\sum_n |\alpha_n \langle h, e_n \rangle|^2 \leq M^2 \sum_n |\langle h, e_n \rangle|^2 < \infty$ converges, we have from previous results that $Ah_k$ converges to $\sum_n \alpha_n \langle h, e_n \rangle e_n$, which we define to be $Ah$. Then $||Ah||^2 = \sum_n |\alpha_n \langle h, e_n \rangle|^2 \leq M^2 ||h||^2 = M^2$, so $||Ah|| \leq M$ and so $||A|| \leq M$ taking supremums. Thus $A$ is bounded and $||A|| = M < \infty$.

Let $P_n$ be the projection of $\mathscr{H}$ onto $\wedge\{e_1, ..., e_n\}$. Then $A_n = A - AP_n$ is seen to satisfy $A_n e_j = \alpha_j e_j$ if $j > n$ and $A_n e_j = 0$ if $j \leq n$. So $AP_n \in \mathscr{B}_{00}(\mathscr{H})$ and $||A_n|| = \sup\{|\alpha_j| : j > n\}$ from our initial work. If $\alpha_n \to 0$ then $||A_n|| \to 0$ and so $A$ is compact since it is the limit of a sequence of finite-rank operators. Conversely, if $A$ is compact, then the previous Corollary implies $||A_n|| \to 0$; hence $\alpha_n \to 0$. □

**Proposition 2.4.7** If $(X, \Omega, \mu)$ is a measure space and $k \in L^2(X \times X, \Omega \times \Omega, \mu \times \mu)$, then

$$(Kf)(x) = \int k(x,y) f(y) d\mu(y)$$

is a compact operator and $||K|| \leq ||k||_2$.

We require the following lemma.

**Lemma 2.4.8** If $\{e_i : i \in I\}$ is a basis for $L^2(X, \Omega, \mu)$ and

$$\phi_{ij}(x,y) = e_j(x)\overline{e_i(y)}$$

for $i, j \in I$, $x, y \in X$, then $\{\phi_{ij} : i, j \in I\}$ is an orthonormal set in $L^2(X \times X, \Omega \times \Omega, \mu \times \mu)$. If $k$ and $K$ are as in the preceding proposition, then $\left\langle k, \phi_{ij} \right\rangle = \left\langle Ke_j, e_i \right\rangle$.

We can now prove the proposition.

***Proof*** First we show that $K$ defines a bounded operator. In fact, if $f \in L^2(\mu)$,

$$||Kf||^2 = \int \left| \int k(x,y) f(y) d\mu(y) \right|^2 d\mu(x) \leq \int \left( \int |k(x,y)|^2 d\mu(y) \right) \left( \int |f(y)|^2 d\mu(y) \right) d\mu(x) = ||k||^2 \, ||f||^2$$

Hence $K$ is bounded and $||K|| \leq ||k||_2$.
Now let $\{e_i\}$ be a basis for $L^2(\mu)$ and define $\phi_{ij}$ as in the statement of the lemma. Thus

$$||k||^2 \geq \sum_{i,j} | \left\langle k, \phi_{ij} \right\rangle |^2 = \sum_{i,j} | \left\langle Ke_j, e_i \right\rangle |^2$$

Since $k \in L^2(\mu \times \mu)$, there are at most a countable number of $i$ and $j$ such that $\left\langle k, \phi_{ij} \right\rangle \neq 0$; denote these by $\{\psi_{km} : 1 \leq k, m < \infty\}$. Note that $\left\langle Ke_j, e_i \right\rangle = 0$ unless $\phi_{ij} \in \{\psi_{km}\}$. Let $P_n$ be the orthogonal projection onto $\wedge\{e_k : 1 \leq k \leq n\}$, and put $K_n = KP_n + P_n K - P_n K P_n$; so $K_n$ is a finite rank operator. Let $f \in L^2(\mu)$ with $||f||^2 \leq 1$; so $f = \sum_j \alpha_j e_j$. Hence

$$\begin{aligned} ||Kf - K_n f||^2 &= \sum_i | \langle Kf - K_n f, e_i \rangle |^2 \\ &= \sum_i \left| \sum_j \alpha_j \left\langle (K - K_n) e_j, e_i \right\rangle \right|^2 \\ &= \sum_k \left| \sum_m \alpha_m \langle (K - K_n) e_m, e_k \rangle \right|^2 \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_k \left[\sum_m |\alpha_m|^2\right]\left[\sum_m |\left\langle (K-K_n)e_m, e_k\right\rangle|^2\right] \\
&\leq ||f||^2 \sum_k \sum_m |\left\langle Ke_m, e_k, -\right\rangle \left\langle KP_ne_m, P_ne_k\right\rangle \\
&\quad - \left\langle KP_ne_m, P_ne_k\right\rangle + \left\langle KP_ne_m, P_ne_k\right\rangle|^2 \\
&\leq \sum_{k=n+1}^{\infty} \sum_{m=n+1}^{\infty} |\left\langle Ke_m, e_k\right\rangle|^2 \\
&= \sum_{k=n+1}^{\infty} \sum_{m=n+1}^{\infty} |\left\langle k, \psi_{km}\right\rangle|^2
\end{aligned}
$$

SInce $\sum_{k,m} |\langle k, \psi_{km}\rangle|^2 < \infty$, $n$ can be chosen sufficiently large such that for any $\epsilon > 0$ this last sum will be smaller than $\epsilon^2$. Thus $||K - K_n|| \to 0$. □

We now begin the spectral theory for our operators.

**Definition 2.4.9** If $A \in \mathscr{B}(\mathscr{H})$, a scalar $\alpha \in \mathbb{F}$ is an **eigenvalue** of $A$ if $\ker(A-\alpha) \neq (0)$. If $h \in \ker(A-\alpha)\backslash(0)$, $h$ is called an **eigenvector** for $\alpha$; thus $Ah = \alpha h$. Let $\sigma_p(A)$ denote the set of eigenvalues of $A$.

Example:

Let $A$ be the diagonalizable operator (i.e. we have a diagonal basis) in the proposition above. Then $\sigma_p(A) = \{\alpha_1, \alpha_2, ...\}$. If $\alpha \in \sigma_p(A)$, let $J_\alpha = \{j ßn\mathbb{N} : \alpha_j = \alpha\}$. Then $h$ isan eigenvector for $\alpha$ if and only if $h \in \wedge\{e_j : j \in J_\alpha\}$.

Example:

The Volterra operator has no eigenvalues.

Example:

Let $h \in \mathscr{H} = L^2_{\mathbb{C}}(-\pi, \pi)$ and define $K : \mathscr{H} \to \mathscr{H}$ by

$$(Kf)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} h(x-y)f(y)dy$$

If $\lambda_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} h(x)\exp(-inx)dx = \hat{h}(n)$, the $n$th Fourier coefficient of $h$, then $Ke_n = \lambda_n e_n$, where $e_n(x) = \frac{1}{\sqrt{2\pi}}\exp(-inx)$.

Operators on finite dimensional spaces over $\mathbb{C}$ always have eigenvalues. As the Volterra operator illustrates, the analogy between operators on finite dimensional spacs and compact operators breaks down here.

**Proposition 2.4.10** If $T \in \mathscr{B}_0(\mathscr{H})$, $\lambda \in \sigma_p(T)$, and $\lambda \neq 0$, then the eigenspace $\ker(T-\lambda)$ is finite dimensional.

***Proof*** Suppose there is an infinite orthonormal sequence $\{e_n\}$ in $\ker(T-\lambda)$. Since $T$ is compact there is a subsequence $\{e_{n_k}\}$ such that $\{Te_{n_k}\}$ converges. Thus $\{Te_{n_k}\}$ is a Cauchy sequence. But for $n_k \neq n_j$, $\left\|Te_{n_k} - Te_{n_j}\right\|^2 = \left\|\lambda e_{n_k} - \lambda e_{n_j}\right\|^2 = 2|\lambda|^2 > 0$, sicne $\lambda \neq 0$. This contradiction shows that $\ker(T-\lambda)$ must be finite dimensional. □

**Proposition 2.4.11** If $T$ is a compact operator on $\mathscr{H}$, $\lambda \neq 0$, and $\inf\{||(T-\lambda)h|| : ||h|| = 1\} = 0$, then $\lambda \in \sigma_p(T)$.

***Proof*** By hypothesis, there is a sequence of unit vectors $\{h_n\}$ such that $||(T-\lambda)h_n|| \to 0$. Since $T$ is compact, there is a vector $f$ in $\mathscr{H}$ and a subsequence $\{h_{n_k}\}$ such that $\left\|Th_{n_k} - f\right\| \to 0$. But $h_{n_k} = \lambda^{-1}[(\lambda - T)h_{n_k} + Th_{n_k}] \to \lambda^{-1}f$. So $1 = \left\|\lambda^{-1}f\right\| = |\lambda|^{-1}\,||f||$ and $f \neq 0$. Also, it must be that $Th_{n_k} \to \lambda^{-1}Tf$. Since $Th_{n_k} \to f$, $f = \lambda^{-1}Tf$, or $Tf = \lambda f$. That is, $f \in \ker(T-\lambda)$ and $f \neq 0$, so $\lambda \in \sigma_p(T)$. □

**Corollary 2.4.12** If $T$ is a compact operator on $\mathscr{H}$, $\lambda \neq 0$, $\lambda \notin \sigma_p(T)$, and $\overline{\lambda} \notin \sigma_p(T^*)$, then $\mathrm{Im}(T-\lambda) = \mathscr{H}$ and $(T-\lambda)^{-1}$ is a bounded operator on $\mathscr{H}$.

***Proof*** Since $\lambda \notin \sigma_p(T)$, the preceding proposition implies there exists $c > 0$ such that $||(T-\lambda)h|| \geq c\,||h||$ for all $h \in \mathscr{H}$. If $f \in \mathrm{cl}[\mathrm{Im}(T-\lambda)]$, then there is a sequence $\{h_n\}$ in $\mathscr{H}$ such that $(T-\lambda)h_n \to f$. Thus $||h_n - h_m|| \leq c^{-1}\,||(T-\lambda)h_n - (T-\lambda)h_m||$, and so $\{h_n\}$ is a Cauchy sequence. Hence $h_n \to h$ for some $h \in \mathscr{H}$. THus $(T-\lambda)h = f$. So $\mathrm{Im}(T-\lambda)$ is closed, and hence $\mathrm{Im}(T-\lambda) = [\ker(T-\lambda)^*]^\perp = \mathscr{H}$, by hypothesis.
So for $f \in \mathscr{H}$ let $Af =$ the unique $h$ such that $(T-\lambda)h = f$. Thus $(T-\lambda)Af = f$ for all $f \in \mathscr{H}$. From the inequality above,

$$c\,||Af|| \leq ||(T-\lambda)Af|| = ||f||$$

so $||Af|| \leq c^{-1}\,||f||$ and $A$ is bounded. Also,

$$(T-\lambda)A(T-\lambda)h = (T-\lambda)h$$

so $0 = (T-\lambda)[A(T-\lambda)h - h]$. Since $\lambda \notin \sigma_p(T)$, $A(T-\lambda)h = h$. That is, $A = (T-\lambda)^{-1}$. □

## 2.5 Diagonalization of Compact Self-Adjoint Operators

The main result we aim to prove in this section is the following.

**Theorem 2.5.1** If $T$ is a compact self-adjoint operator on $\mathscr{H}$, then $T$ has only a countable number of distinct eigenvalues. If $\{\lambda_1, \lambda_2, \ldots\}$ are the distinct nonzero eigenvalues of $T$, and $P_n$ is the projection of $\mathscr{H}$ onto $\ker(T - \lambda_n)$, then $P_n P_m = P_m P_n = 0$ if $n \neq m$, eahc $\lambda_n$ is real, and

$$T = \sum_{n=1}^{\infty} \lambda_n P_n$$

where the series converges to $T$ in the metric defined by the norm of $\mathscr{B}(\mathscr{H})$.

We first collect some preliminary results after providing some consequences of this theorem.

**Corollary 2.5.2** With the notation as in the theorem,

(a) $\ker T = [\wedge\{P_n\mathscr{H} : n \geq 1\}]^{\perp} = (\operatorname{Im} T)^{\perp}$;
(b) each $P_n$ has finite rank;
(c) $||T|| = \sup\{|\lambda_n| : n \geq 1\}$ and $\lambda_n \longrightarrow 0$ as $n \longrightarrow \infty$.

***Proof*** Since $P_n \perp P_m$ for $n \neq m$, if $h \in \mathscr{H}$, then

$$||Th||^2 = \sum_{n=1}^{\infty} ||\lambda_n P_n h||^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \, ||P_n h||^2$$

Hence $Th = 0$ if and only if $P_n h = 0$ for all $n$. That is $h \in \ker T$ if and only if $h \perp P_n\mathscr{H}$ for all $n$, whence (a).
Part (b) follows from a proposition in the previous section on eigenspaces of nonzero eigenvalues.
For part (c), if $\mathscr{L} = \operatorname{cl}[\operatorname{Im} T]$, $\mathscr{L}$ is invariant for $T$. Since $T = T^*$, $\mathscr{L} = (\ker T)^{\perp}$ and $\mathscr{L}$ reduces $T$. So we can consider the restriction of $T$ to $\mathscr{L}$, $T|_{\mathscr{L}}$. Now $\mathscr{L} = \wedge\{P_n\mathscr{H} : n \geq 1\}$ by (a). Let $\{e_j^{(n)} : 1 \leq j \leq N_n\}$ be a basis for $P_n\mathscr{H} = \ker(T - \lambda_n)$, so $Te_j^{(n)} = \lambda_n e_j^{(n)}$ for $1 \leq j \leq N_n$. Thus $\{e_j^{(n)} : 1 \leq j \leq N_n, n \geq 1\}$ is a basis for $\mathscr{L}$ and $T|_{\mathscr{L}}$ is diagonalizable with respect to this basis. Part (c) now follows from a result in the previous section. □

**Corollary 2.5.3** If $T$ is a compact self-adjoint operator, then there is a sequence $\{\mu_n\}$ of real numbers and an orthonormal basis $\{e_n\}$ for $(\ker T)^{\perp}$ such that for all $h$,

$$Th = \sum_{n=1}^{\infty} \mu_n \langle h, e_n \rangle \, e_n$$

Note that there may be repitions in the sequence $\{\mu_n\}$.

**Corollary 2.5.4** If $T \in \mathscr{B}_0(\mathscr{H})$, $T = T^*$, and $\ker T = (0)$, then $\mathscr{H}$ is separable.

Now we begin collecting our results for the proof of our main theroem.

**Proposition 2.5.5** If $A$ is a normal operator and $\lambda \in \mathbb{F}$, then $\ker(A-\lambda) = \ker(A-\lambda)^*$ and $\ker(A-\lambda)$ is a reducing subspace for $A$.

***Proof*** Since $A$ is normal, so is $A - \lambda$. Hence $||(A-\lambda)h|| = ||(A-\lambda)^*h||$, so $\ker(A-\lambda) = \ker(A-\lambda)^*$. If $h \in \ker(A-\lambda)$, $Ah = \lambda h \in \ker(A-\lambda)$. Also, $A^*h = \overline{\lambda}h \in \ker(A-\lambda)$. Therefore $\ker(A-\lambda)$ reduces $A$. □

**Proposition 2.5.6** If $A$ is a normal operator and $\lambda, \mu$ are distinct eigenvalues of $A$,then $\ker(A-\lambda) \perp \ker(A-\mu)$.

***Proof*** If $h \in \ker(A-\lambda)$ and $g \in \ker(A-\mu)$, then the previous proposition implies $A^*g = \overline{\mu}g$ and $\lambda \langle h, g\rangle = \mu \langle h, g\rangle$. Thus $(\lambda-\mu)\langle h, g\rangle = 0$. Since $\lambda - \mu \neq 0$, $h \perp g$. □

**Proposition 2.5.7** If $A = A^*$ and $\lambda \in \sigma_p(A)$, then $\lambda$ is a real number.

***Proof*** If $Ah = \lambda h$, then $Ah = A^*h = \overline{\lambda}h$, so $(\lambda - \overline{\lambda})h = 0$. Since $h$ can be chosen different from 0, $\lambda = \overline{\lambda}$.□

**Lemma 2.5.8** If $T$ is a compact self-adjoint operator, then either $\pm ||T||$ is an eigenvalue of $T$.

***Proof*** If $T = 0$, the result is immediate. So suppose $T \neq 0$. As $T$ is self-adjoint there is a sequence $\{h_n\}$ of unit vectors such that $|\langle Th_n, h_n\rangle| \to ||T||$. By passing to a subsequence if necessary, we may assume that $\langle Th_n, h_n\rangle \to \lambda$, where $|\lambda| = ||T||$. It will be shown that $\lambda \in \sigma_p(T)$. Since $|\lambda| = ||T||$,

$$0 \leq ||(T-\lambda)h_n||^2 = ||Th_n||^2 - 2\lambda\langle Th_n, h_n\rangle + \lambda^2 \leq 2\lambda^2 - 2\lambda\langle Th_n, h_n\rangle \to 0$$

Hence $||(T-\lambda)h_n|| \to 0$. Thus $\lambda \in \sigma_p(T)$. □

We now proceed to the proof of the main theorem.

***Proof*** By the Lemma there is a real number $\lambda_1$ in $\sigma_p(T)$ with $|\lambda_1| = ||T||$. Let $\mathscr{E}_1 = \ker(T-\lambda_1)$, $P_1 =$ the projection onto $\mathscr{E}_1$, $\mathscr{H}_2 = \mathscr{E}_1^{\perp}$. Then $\mathscr{E}_1$ reduces $T$, so $\mathscr{H}_2$ reduces $T$. Let $T_2 = T|_{\mathscr{H}_2}$; then $T_2$ is a self-adjoint compact operator on $\mathscr{H}_2$.
By the Lemma there is an eigenvalue $\lambda_2$ for $T_2$ such that $|\lambda_2| = ||T_2||$. Let $\mathscr{E}_2 = \ker(T_2 - \lambda_2)$. Note that $(0) \neq \mathscr{E}_2 \subseteq \ker(T-\lambda_2)$. Since $\mathscr{E}_1 \perp \mathscr{E}_2$, $\lambda_1 \neq \lambda_2$. Let $P_2$ be the projection of $\mathscr{H}$ onto $\mathscr{E}_2$ and put $\mathscr{H}_3 = (\mathscr{E}_1 \oplus \mathscr{E}_2)^{\perp}$. Note that $||T_2|| \leq ||T||$, so $|\lambda_2| \leq |\lambda_1|$.
Using induction we obtain a sequence $\{\lambda_n\}$ of real eigenvalues of $T$ such that

(i) $|\lambda_1| \geq |\lambda_2| \geq \cdots$;
(ii) If $\mathscr{E}_n = \ker(T-\lambda_n)$, $|\lambda_{n+1}| = \left\|T|_{(\mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_n)^{\perp}}\right\|$

By (i) there is a nonnegative number $\alpha$ such that $|\lambda_n| \to \alpha$.
**Claim.** $\alpha = 0$.
In fact, let $e_n \in \mathscr{E}_n$, $||e_n|| = 1$. Since $T$ is compact there is an $h \in \mathscr{H}$ and a subsequence $\{e_{n_j}\}$ such that $\left\|Te_{n_j} - h\right\| \to 0$. But $e_n \perp e_m$ for $n \neq m$ and $Te_{n_j} = \lambda_{n_j}e_{n_j}$. Hence $\left\|Te_{n_j} - Te_{n_i}\right\|^2 = \lambda_{n_j}^2 + \lambda_{n_i}^2 \geq 2\alpha^2$. Since $\{Te_{n_j}\}$ is a Cauchy sequence, $\alpha = 0$.

Now put $P_n$ = the projection of $\mathscr{H}$ onto $\mathscr{E}_n$ and examine $T - \sum_{j=1}^n \lambda_j P_j$. If $h \in \mathscr{E}_k$, $1 \le k \le n$, then

$$\left(T - \sum_{j=1}^n \lambda_j P_j\right) h = T_h - \lambda_k h = 0$$

Hence $\mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_n \subseteq \ker\left(T - \sum_{j=1}^n \lambda_j P_j\right)$. If $h \in (\mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_n)^\perp$, then $P_j h = 0$ for $1 \le j \le n$; so $\left(T - \sum_{j=1}^n \lambda_j P_j\right) h = Th$. These two statements together with the fact that $(\mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_n)^\perp$ reduces $T$, imply that

$$\left\| T - \sum_{j=1}^n \lambda_j P_j \right\| = \left\| T|_{(\mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_n)^\perp} \right\| = |\lambda_{n+1}| \to 0$$

Therefore the series $\sum_{n=1}^\infty \lambda_n P_n$ converges in the metric of $\mathscr{B}(\mathscr{H})$ to $T$. □

This is called the **Spectral Theorem** for compact self-adjoint operators.

## 2.6 The Spectral Theorem and Functional Calculus for Compact Normal Operators

**Proposition 2.6.1** Let $\{P_i : i \in I\}$ be a family of pairwise orthogonal projections in $\mathscr{B}(\mathscr{H})$. (That is, $P_i P_j = P_j P_i = 0$ for $i \neq j$.) If $h \in \mathscr{H}$, then $\sum_i \{P_i h : i \in I\}$ converges in $\mathscr{H}$ to $Ph$, where $P$ is the projection of $\mathscr{H}$ onto $\bigvee \{P_i \mathscr{H} : i \in I\}$.

***Proof*** (TBD) □

If $\{P_i : i \in I\}$ is as in the preceding proposition and $\mathscr{M}_i = P_i \mathscr{H}$, then $P$ is the projection of $\mathscr{H}$ onto $\bigoplus_i \mathscr{M}_i$. Write $P = \sum_i P_i$. Note that $Ph = \sum_i P_i h$ where the convergence is in the norm of $\mathscr{H}$, but $\sum_i P_i$ does not converge to $P$ in the norm of $\mathscr{B}(\mathscr{H})$ unless $I$ is finite.

**Definition 2.6.2** A **partition of the identity** on $\mathscr{H}$ is a family $\{P_i : i \in I\}$ of pairwise orthogonal projections on $\mathscr{H}$ such that $\bigvee_i P_i \mathscr{H} = \mathscr{H}$. This might be indicated by $1 = \sum_i P_i$.

**Definition 2.6.3** An operator $A$ on $\mathscr{H}$ is **diagonalizable** if there is a partition of the identity on $\mathscr{H}$, $\{P_i : i \in I\}$, and a family of scalars $\{\alpha_i : i \in I\}$ such that $\sup_i |\alpha|_i < \infty$ and $Ah = \alpha_i h$ whenever $h \in \operatorname{Im} P_i$.

In this case $||A|| = \sup_i |\alpha_i|$.

To denote a diagonalizable operator satisfying the conditions of this definition, write

$$A = \sum_i \alpha_i P_i$$

Note it was note assumed that the scalars $\alpha_i$ are distinct. However, if $\alpha_i = \alpha_j$ then we can replace $P_i$ and $P_j$ with $P_i + P_j$.

**Proposition 2.6.4** An operator $A$ on $\mathscr{H}$ is diagonalizable if and only if there is an orthonormal basis for $\mathscr{H}$ consisting of eigenvectors for $A$.

***Proof*** (TBD) □

Also note that if $A = \sum_i \alpha_i P_i$, then $A^* = \sum_i \overline{\alpha}_i P_i$ and $A$ is normal.

**Theorem 2.6.5** If $A = \sum_i \alpha_i P_i$ is diagonalizable and all the $\alpha_i$ are distinct, then an operator $B \in \mathscr{B}(\mathscr{H})$ satisfies $AB = BA$ if and only if for each $i$, $\operatorname{Im} P_i$ reduces $B$.

***Proof*** If all the $\alpha_i$ are distinct, then $\operatorname{Im} P_i = \ker(A - \alpha_i)$. If $AB = BA$ and $Ah = \alpha_i h$, then $ABh = BAh = B(\alpha_i h) = \alpha_i Bh$; hence $Bh \in \operatorname{Im} P_i$ whenever $h \in \operatorname{Im} P_i$. Thus $\operatorname{Im} P_i$ is left invaraint by $B$. Therefore $B$ leaves $\bigvee\{\operatorname{Im} P_j : j \neq i\} = \mathcal{N}_i$ invariant. But since $\sum_i P_i = 1$, $\mathcal{N}_i = (\operatorname{Im} P_i)^\perp$. Thus $\operatorname{Im} P_i$ reduces $B$.
Now assume that $B$ is reduced by each $\operatorname{Im} P_i$. Thus $BP_i = P_i B$ for all $i$. If $h \in \mathscr{H}$, then $Ah = \sum_i \alpha_i P_i h$. Hence $BAh = \sum_i \alpha_i BP_i h = \sum_i \alpha_i P_i Bh = ABh$. □

Using the notation of the preceding theorem, if $AB = BA$, let $B_i = B|_{\operatorname{Im} P_i}$. THen it is appropriate to write $B = \sum_i B_i$ on $\mathscr{H} = \sum_i P_i \mathscr{H}$. One might paraphrase the theorem by saying that $B$ commutes with a diagonalizable operator if and only if $B$ can be “diagonalized with operator entries."

**Theorem 2.6.6 (Spectral Theorem for Compact Normal Operators)** If $T$ is a compact normal operator on the complex Hilbert space $\mathscr{H}$, then $T$ has only a countable number of distinct eigenvalues. If $\{\lambda_1, \lambda_2, \ldots\}$ are the distinct nonzero eigenvalues of $T$, and $P_n$ is the projection of $\mathscr{H}$ onto $\ker(T - \lambda_n)$, then $P_n P_m = P_m P_n = 0$ if $n \neq m$ and

$$T = \sum_{n=1}^{\infty} \lambda_n P_n \tag{2.6.1}$$

where this series converges to $T$ in the metric defined by the norm on $\mathscr{B}(\mathscr{H})$.

***Proof*** Let $A = (T + T^*)/2, B = (T - T^*)/2i$. So $A, B$ are compact self-adjoint operators, $T = A + iB$, and $AB = BA$ since $T$ is normal. By Theorem 2.5, $A = \sum_n \alpha_n E_n$, where $\alpha_n \in \mathbb{R}$, $\alpha_n \neq \alpha_m$ if $n \neq m$, and $E_n$ is the projection of $\mathscr{H}$ onto $\ker(A - \alpha_n)$. Since $AB = BA$ we can diagonalize $A$ and $B$ simultaneously.
For each $n$, $E_n \mathscr{H} = \mathcal{L}_n$ reduces $B$ since $BA = AB$. Let $B_n = B|_{\mathcal{L}_n}$. Then $B_n = B_n^*$ and $\dim \mathcal{L}_n < \infty$. Applying the spectral theorem for finite dimensional linear operators to $B_n$, there is a basis $\{e^{(n)} : 1 \leq j \leq d_n\}$ for $\mathcal{L}_n$ and real numbers $\{\beta_j^{(n)} : 1 \leq j \leq d_n\}$ such that $B_n e_j^{(n)} = \beta_j^{(n)} e_j^{(n)}$. Thus $Te_j^{(n)} = Ae_j^{(n)} + iBe_j^{(n)} = (\alpha_n + i\beta_j^{(n)}) e_j^{(n)}$.
Therefore, $\{e_j^{(n)} : 1 \leq j \leq d_n, n \geq 1\}$ is a basis for $\operatorname{cl}(\operatorname{Im} A)$ consisting of eigenvectors for $T$. It may be that $\operatorname{cl}(\operatorname{Im} A) \neq \operatorname{cl}(\operatorname{Im} T)$. Since $B$ is reduced by $\ker A = (\operatorname{Im} A)^\perp$ and $B_0 = B|_{\ker A}$ is a compact

self-adjoint operator there is an orthonormal basis $\{e_j^{(0)} : j \geq 1\}$ for $\mathrm{cl}(\mathrm{Im}\, B_0)$ and scalars $\{\beta_j^{(0)} : j \geq 1\}$ such that $Be_j^{(0)} = \beta_j^{(0)} e_j^{(0)}$. It follows that $Te_j^{(0)} = i\beta_j^{(0)} e_j^{(0)}$. Moreover, $\ker T^* \supseteq \ker A \cap \ker B_0$, so $\mathrm{cl}(\mathrm{Im}\, T) \subseteq \mathrm{cl}(\mathrm{Im}\, A) \oplus \mathrm{cl}(\mathrm{Im}\, B_0)$.
Thus, $T$ has a countable number of distinct eigenvalues **TO BE CONTINUED** □

**Corollary 2.6.7** With notation as in the theorem

(a) $\ker T = [\bigvee\{P_n \mathscr{H} : n \geq 1\}]^{\perp}$;
(b) each $P_n$ has finite rank;
(c) $||T|| = \sup\{|\lambda_n| : n \geq 1\}$ and either $\{\lambda_n\}$ is finite or $\lambda_n \to 0$ as $n \to \infty$.

**Corollary 2.6.8** If $T$ is a compact operator on a complex Hilbert space, then $T$ is normal if and only if $T$ is diagonalizable.

We now develop a functional calculus for normal operators $T$. That is an operator $\phi(T)$ will be defined for every bounded Borel function $\phi$ on $\mathbb{C}$ and certain properties of the map $\phi \mapsto \phi(T)$ will be deduced. A projection-valued measure will then be obtained by letting $\mu(\Delta) = \chi_\Delta(T)$, where $\chi_\Delta$ is the characteristic function of the set $\Delta$. We restrict to the case of compact normal operators for now.

**Definition 2.6.9** Denote by $l^\infty(\mathbb{C})$ all the bounded functions $\phi : \mathbb{C} \to \mathbb{C}$. If $T$ is a compact normal operator which has description as in the theorem, define $\phi(T) : \mathscr{H} \to \mathscr{H}$ by

$$\phi(T) = \sum_{n=1}^{\infty} \phi(\lambda_n) P_n + \phi(0) P_0$$

where $P_0$ = the projection of $\mathscr{H}$ onto $\ker T$.

Note that $\phi(T)$ is a diagonalizable operator and $||\phi(T)|| = \sup_i \{|\phi(0)|, |\phi(\lambda_i)|\}$.

**Theorem 2.6.10 (Functional Calculus for Compact Normal Operators)** If $T$ is a compact normal operator on a $\mathbb{C}$-Hilbert space $\mathscr{H}$, then the map $\phi \mapsto \phi(T)$ of $l^\infty(\mathbb{C}) \to \mathscr{B}(\mathscr{H})$ has the following properties:

(a) $\phi \mapsto \phi(T)$ is a multiplicative linear map of $l^\infty(\mathbb{C})$ into $\mathscr{B}(\mathscr{H})$. If $\phi \equiv 1$, $\phi(T) = 1$; if $\phi(z) = z$on $\sigma_p(T) \cup \{0\}$, then $\phi(T) = T$.
(b) $||\phi(T)|| = \sup\{|\phi(\lambda)| : \lambda \in \sigma_p(T)\}$
(c) $\phi(T)^* = \phi^*(T)$ where $\phi^*$ is the function defined by $\phi^*(z) = \overline{\phi(z)}$.
(d) If $A \in \mathscr{B}(\mathscr{H})$ and $AT = TA$, then $A\phi(T) = \phi(T)A$ for all $\phi \in l^\infty(\mathbb{C})$.

***Proof*** For (a), if $\phi, \psi \in l^\infty(\mathbb{C})$, then $(\phi\psi)(z) = \phi(z)\psi(z)$ for $z \in \mathbb{C}$. Also,

$$\phi(T)\psi(T)h = \left[\phi(0)P_0 + \sum_n \phi(\lambda_n)P_n\right]\left[\psi(0)P_0 + \sum_m \psi(\lambda_m)P_m\right]h = \left[\phi(0)P_0 + \sum_n \phi(\lambda_n)P_n\right]\left[\psi(0)P_0 h + \sum_m \psi(\lambda\right.$$

Since $P_n P_m = 0$ whenever $n \neq m$, this gives that

$$\phi(T)\psi(T)h = \phi(0)\psi(0)P_0 h + \sum_n \phi(\lambda_n)\psi(\lambda_n)P_n h = (\phi\psi)(T)h$$

THus $\phi \mapsto \phi(T)$ is multiplicative. The linearity of the map is immediate. If $\phi(z) = 1$, then

$$\phi(T) = 1(T) = P_0 + \sum_n P_n = 1$$

since $\{P_0, P_1, ...\}$ is a partition of the identity. If $\phi(z) = z$, $\phi(\lambda_n) = \lambda_n$ and so $\phi(T) = T$.
Part (c) is clear and part (b) follows from previous work.
For part (d) if $AT = TA$ then by Theorem 2.6 $P_0\mathscr{H}, P_1\mathscr{H}, \dots$ all reduce $A$. Fix $h_n \in P_n\mathscr{H}$, $n \geq 0$. If $\phi \in l^\infty(\mathbb{C})$, then $Ah_n \in P_n\mathscr{H}$ and so

$$\phi(T)Ah_N = \phi(\lambda_n)Ah_n = A(\phi(\lambda_n)h_n) = A\phi(T)h_n$$

If $h \in \mathscr{H}$, then $\sum_{n\geq 0} h_n$, where $h_n \in P_n$. Hence

$$\phi(T)Ah = \sum_{n\geq 0} \phi(T)Ah_n = \sum_{n\geq 0} A\phi(T)h_n = A\phi(T)h$$

Question?

Which operators on $\mathscr{H}$ can be expressed as $\phi(T)$ for some $\phi$ in $l^\infty(\mathbb{C})$?

**Theorem 2.6.11** If $T$ is a compact normal operator on a $\mathbb{C}$-Hilbert space, then $\{\phi(T) : \phi \in l^\infty(\mathbb{C})\}$ is equal to

$$\{B \in \mathscr{B}(\mathscr{H}) : BA = AB \text{ whenever } AT = TA\}$$

***Proof*** Half of the desired equality is obtained from part (d) of the previous theorem. So let $B \in \mathscr{B}(\mathscr{H})$ and assume that $AB = BA$ whenever $AT = TA$. Thus $B$ must commute with $T$ itself. By Theorem 2.6 $B$ is reduced by each $P_n\mathscr{H} =: \mathscr{H}_n, n \geq 0$. But $B_n = B|_{\mathscr{H}_n}$. Fix $n \geq 0$ and let $A_n$ be any bounded operator on $\mathscr{B}(\mathscr{H}_n)$. Define $Ah = A_n h$ if $h \in \mathscr{H}_n$ and $A_n h = 0$ if $h \in \mathscr{H}_m, m \neq n$, and extend $A$ to $\mathscr{H}$ by linearity; so $A = \sum_{m\geq 0} A_m$ where $A_m = 0$ if $m \neq n$. By Theorem 2.6 $AT = TA$; hence $AB = BA$. THis implies that $B_n A_n = A_n B_n$. Since $A_n$ was arbitrarily chosen from $\mathscr{B}(\mathscr{H}_n)$, $B_n = \beta_n$ for some $\beta_n$. If $\phi : \mathbb{C} \to \mathbb{C}$ is defined by $\phi(0) = \beta_0$ and $\phi(\lambda_n) = \beta_n$ for $n \geq 1$, then $B = \phi(T)$. □

**Definition 2.6.12** If $A \in \mathscr{B}(\mathscr{H})$, then $A$ is **positive** if $\langle Ah, h\rangle \geq 0$ for all $h \in \mathscr{H}$.

Recall that every positive operator on a complex Hilbert space is self-adjoint.

**Proposition 2.6.13** If $T$ is a compact normal operator, then $T$ is positive if and only if all its eigenvalues are non-negative real numbers.

***Proof*** Let $T = \sum_n \lambda_n P_n$. If $T \geq 0$ and $h \in P_n\mathscr{H}$ with $||h|| = 1$, then $Th = \lambda_n h$. Hence $\lambda_n = \langle Th, h\rangle \geq 0$. Conversely, assume each $\lambda_n \geq 0$. If $h \in \mathscr{H}$, $h = h_0 + \sum_n h_n$, where $h_0 \in \ker T$ and $h_n \in P_n\mathscr{H}$ for $n \geq 1$. Then $Th = \sum_n \lambda_n h_n$. Hence

$$\langle Th, h\rangle = \left\langle \sum_n \lambda_n h_n, h_0 + \sum_m h_m \right\rangle = \sum_n \sum_m \lambda_n \langle h_n, h_m\rangle = \sum_n \lambda_n ||h_n||^2 \geq 0$$

since $\langle h_n, h_m\rangle = 0$ when $n \neq m$. □

**Theorem 2.6.14** If $T$ is a compact self-adjoint operator, then there are unique positive compact operators $A, B$ such that $T = A - B$ and $AB = BA = 0$.

***Proof*** Let $T = \sum_n \lambda_n P_n$ as in the Spectral theorem. Define $\phi, \psi : \mathbb{C} \to \mathbb{C}$ by $\phi(\lambda_n) = \lambda_n$ if $\lambda_n > 0$, $\phi(z) = 0$ otherwise; $\psi(\lambda_n) = -\lambda_n$ if $\lambda_n < 0$, $\psi(z) = 0$ otherwise. Put $A = \phi(T)$ and $B = \psi(T)$. Then $A = \sum\{\lambda_n P_n : \lambda_n > 0\}$ and $B = \sum\{-\lambda_n P_n : \lambda < 0\}$. Thus $T = A - B$. Since $\phi\psi = 0$ $AB = BA = 0$. Since $\phi, \psi \geq 0$, $A, B \geq 0$ by the preceding proposition. It remains to show uniqueness.
Suppose $T = C - D$ where $C, D$ are compact positive operators and $CD = DC = 0$. Then $C$ and $D$ commute with $T$. Put $\lambda_0 = 0$ and $P_0 =$ the projection of $\mathscr{H}$ onto $\ker T$. Thus $C$ and $D$ are reduced by $P_n\mathscr{H} =: \mathscr{H}_n$ for all $n \geq 0$. Let $C_n = C|_{\mathscr{H}_n}$ and $D_n = D|_{\mathscr{H}_n}$. So $C_n D_n = D_n C_n = 0$, $\lambda_n P_n = T|_{\mathscr{H}_n} = C_n - D_n$, and $C_n, D_n$ are positive. Suppose $\lambda_n > 0$ and let $h \in \mathscr{H}_n$. Since $C_n D_n = 0$,

$$\ker C_n \supseteq \mathrm{cl}[\mathrm{Im}\, D_n] = (\ker D_n)^{\perp}$$

So if $h \in (\ker D_n)^{\perp}$, then $\lambda_n h = -D_n h$. Hence $\lambda_n ||h||^2 = -\langle D_n h, h\rangle \leq 0$. Thus $h = 0$ since $\lambda_n > 0$. That is $|kerD_n = \mathscr{H}_n$. Thus $D_n = 0 = B|_{\mathscr{H}_n}$ and $C_n = \lambda_n P_n = A|_{\mathscr{H}_n}$. Similarly, if $\lambda_n < 0$, $C_n = 0 = A|_{\mathscr{H}_n}$ and $D_n = -\lambda_n P_n = B|_{\mathscr{H}_n}$. On $\mathscr{H}_0$, $T|_{\mathscr{H}_0} = 0 = C_0 - D_0$. Thus $C_0 = D_0$. But $0 = C_0 D_0 = C_0^2$. Thus $0 = \left\langle C_0^2 h, h\right\rangle = ||C_0 h||^2$, so $C_0 = A|_{\mathscr{H}_0}$ and $D_0 = 0 = B|_{\mathscr{H}_0}$. Therefore $C = A$ and $D = B$. □

**Theorem 2.6.15** If $T$ is a positive compact operator, then there is a unique positive compact operator $A$ such that $A^2 = T$.

***Proof*** Let $T = \sum_n \lambda_n P_n$ as in the Spectral theorem. Since $T \geq 0$, $\lambda_n > 0$ for all $n$. Let $\phi(\lambda_n) = \sqrt{\lambda_n}$ and $\phi(z) = 0$ otherwise; put $A = \phi(T)$. It is easy to check that $A \geq 0$; $A = \sum_n \sqrt{\lambda_n} P_n$ so that $A$ is compact; and $A^t = T$.
Uniqueness TBC □

## 2.7 Unitary Equivalence for Compact Normal Operators

The isomorphism between Hilbert spaces induces an equivalence relation between the operators on the spaces.

**Definition 2.7.1** If $A, B$ are bounded operators on Hilbert spaces $\mathscr{H}, \mathscr{K}$, then $A$ and $B$ are **unitarily equivalent** if there is an isomorphism $U : \mathscr{H} \to \mathscr{K}$ such that $UAU^{-1} = B$ (note $U^{-1} = U^*$).

We now give necessary and sufficient conditions that two compact normal operators be unitarily equivalent.

**Definition 2.7.2** If $T$ is a compact operator, the **multiplicity function** for $T$ is the cardinal number valued function $m_T$ defined for every complex number $\lambda$ by $m_T(\lambda) = \dim \ker(T - \lambda)$.

Note $m_T(\lambda) > 0$ if and only if $\lambda$ is an eigenvalue for $T$. Note $m_T(\lambda) < \omega$ if $\lambda \neq 0$.

If $T, S$ are compact operators on Hilbert spaces and $U : \mathscr{H} \to \mathscr{K}$ is an isomorphism with $UTU^{-1} = S$, then $U \ker(T - \lambda) = \ker(S - \lambda)$ for every $\lambda$ in $\mathbb{C}$. In particular, it must be that $m_T = m_S$.

**Theorem 2.7.3** Two compact normal operators are unitarily equivalent if and only if they have the same multiplicity function.

***Proof*** Let $T, S$ be compact normal operators on $\mathscr{H}, \mathscr{K}$. If $T \equiv S$ then it has already been shown that $m_T = m_S$. Conversely, suppose $m_T = m_S$.
Let $T = \sum_n \lambda_n P_n$ and let $S = \sum_n \mu_n Q_n$ as in the Spectral Theorem. So if $n \neq m$ then $\lambda_n \neq \lambda_m$ and $\mu_n \neq \mu_m$, and each of the projections $P_n$ and $Q_n$ has finite rank. Let $P_0, Q_0$ be the projections onto the kernels, so

$$P_0 = \left(\sum_n P_n\right)^\perp, \text{ and } Q_0 = \left(\sum_n Q_n\right)^\perp$$

Put $\lambda_0 = \mu_0 = 0$.
Since $m_T = m_S$, $0 < m_T(\lambda_n) = m_S(\lambda_n)$. Hence there is a unique $\mu_j$ such that $\mu_j = \lambda_n$. Define $\pi : \mathbb{N} \to \mathbb{N}$ by letting $\mu_{\pi(n)} = \lambda_n$. Let $\pi(0) = 0$. Note that $\pi$ is one-to-one. Also, since $0 < m_S(\mu_n) = m_T(\mu_n)$, for every $n$ there is a $j$ such that $\pi(j) = n$. Thus $\pi$ is a bijection. Since $\dim P_n = m_T(\lambda_n) = m_S(\mu_{\pi(n)}) = \dim Q_{\pi(n)}$, there is an isomorphism $U_n : P_n\mathscr{H} \to Q_{\pi(n)}\mathscr{K}$ for $n \neq 0$. Define $U : \mathscr{H} \to \mathscr{K}$ by letting $U = U_n$ on $P_n\mathscr{H}$ and extending by linearity. Hence $U = \sum_{n \geq 0} U_n$. (TBC) it can be checked that $U$ is an isomorphism.
Also if $h \in P_n\mathscr{H}$, $n \geq 0$, then

$$UTh = \lambda_n Uh = \mu_{\pi(n)} Uh = SUh$$

Hence $UTU^{-1} = S$. □

If $V$ is the Volterra operator, then $m_V \equiv 0$ and $V$ and the zero operator are definitely not unitarily equivalent, so the preceding theorem only applied to compact normal operators.

## Problems

**2.1** A given problem or Excercise is described here. The problem is described here. The problem is described here.

# Index