Operations in Hilbert space

Dennis Michael Gumaer

Operators in Hilbert Space

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Mathematics

by

Dennis Michael Gumaer

September 2008

OPERATORS IN HILBERT SPACE

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Dennis Michael Gumaer

September 2008

Approved by:

[redacted] 7-22-08

Dr. Yuichiro Kakihara, Committee Chair — Date

[redacted]

Dr. Charles Stanton, Committee Member

[redacted]

Dr. Belisario Ventura, Committee Member

[redacted]

Dr. Peter Williams, Chair,
Department of Mathematics

Dr. Joseph Chavez
Graduate Coordinator,
Department of Mathematics

ABSTRACT

This review of some of the major topics in elementary Hilbert space theory starts with a study of basic operations on a Hilbert space. The theory of operators is developed by providing details regarding several types of operators, in particular compact operators.

Compact operators on a Hilbert space are operators which map the closed unit ball to a set whose closure is compact. This study of compact operators is the start of the refinement of bounded linear operators to those which are also members of the Schatten p-class operators – the final goal of this discussion.

Acknowledgements

First, I would like to extend my many thanks to my advisor and committee chair, Dr. Yuichiro Kakihara. I appreciate the considerable time and effort you spent commenting on the multiple versions of this text. Your explanations of the finer points of the theory were instrumental to my understanding of this topic and my growth as a mathematician.

I would like to thank the Department of Mathematics, and specifically Dr. Joseph Chavez. The flexibility you demonstrated allowed the time so this thesis could be properly completed. I would also thank Dr. Charles Stanton and Dr. Belisario Ventura for the detailed reading and suggestions you both put forth.

I offer my unending gratitude to my family. To my parents for your love and support; you instilled in me the belief I could do anything I set my mind to. To my mother-in-law and father-in-law for the bets that kept me on track. I thank you all for helping with the kids after a long day of work.

Most importantly, I credit my loving wife Danaca. Thank you for setting your expectations high and helping me meet them. You're the only one who spent as much time on this thesis as I did. Every moment I spent writing you spent taking care of our family. I dedicate this thesis to you.

Table of Contents

Chapter 1

Introduction

Deeper levels of abstraction arise when there is a general understanding of underlying theory. Hilbert Space is one of those abstractions. When $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{C}$ became less and less mysterious, mathematicians looked to further dimensions. David Hilbert worked with the extreme case and asked questions regarding spaces with arbitrary dimension, including infinite dimensions. There are some basics that would be desired in this arbitrary dimensional space to help coincide with $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{C}$. A portion of the thesis will deal with these desired properties and their consequences.

Hilbert Space H is a complete normed linear space with an inner product. First off, Hilbert Space is a complete space. In any normed linear space X, if every Cauchy sequence approaches a limit that is contained in X, then that space is said to be complete. So sequences which have a limit can actually attain that limit within the given space. Completeness is an essential component of the definition of Hilbert Space. It brings structure to the space allowing further revelations, not to mention keeping properties of Hilbert Space in line with those of $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{C}$. This property will be used throughout the thesis. However, since it is part of the definition, it is usually not specifically pointed out. This is not to belittle its importance.

The primary example of a Hilbert Space for the purpose of the thesis is a linear space called ℓ_2. The space ℓ_2 is the space of all square summable sequences (a_n) of complex numbers.

$$\ell_2 = \left\{ (a_n) \middle| \sum_{n=1}^{\infty} |a_n|^2 < +\infty \right\}.$$

Properties of this space will follow shortly.

The norm is the arbitrary dimensional analogue of the absolute value. In higher dimensions it is denoted $\|\cdot\|$. There are many ways for the norm to be defined. In the case of the space ℓ_2, the norm is defined as

$$\|(a_n)\| = \left(\sum_{n=1}^{\infty} |a_n|^2\right)^{\frac{1}{2}}, \qquad \forall \text{ sequences } (a_n) \in \ell_2.$$

This norm defines a metric on the space allowing further structure as it narrows down the spaces which can be called a Hilbert Space.

An inner product is the arbitrary dimensional analogue of the dot product. (The convention used in my primary research resource, *Hilbert Space: Compact Operators and the Trace Theorem* by J.R. Retherford[Ret93], is that inner products are denoted by $(\cdot,\cdot)$. This convention will be followed throughout the thesis.) Many different inner products are possible. However each must satisfy a list of properties – one property of particular importance will follow shortly. For our primary example ℓ_2 of the thesis, the inner product is defined over ℓ_2 as

$$\forall (a_n),(b_n) \in \ell_2, \qquad \Big((a_n),(b_n)\Big) = \sum_{n=1}^{\infty} a_n \overline{b_n}$$

with the bar denoting complex conjugation.

Hilbert Space is called a complete normed linear space *with an inner product* because it has a special connection between its norm and the given inner product. The special connection can be shown to be

$$(x,x)^{\frac{1}{2}} = \|x\|, \qquad \forall x \in H.$$

(Note that ℓ_2 clearly meets this criteria.)

Initially, many of the problems discussed in the thesis will use the relationship between norms and inner products. This work pervades all later work and is necessary as complexity progresses. Several detailed examples will be used to allow the reader sufficient background for this later work.

The remainder of the thesis will deal with operators on Hilbert Space. Most, if not all, of the operators to be studied are called bounded operators. A linear transformation T is called a bounded operator if

$$\exists M \in \mathbb{R}^+ \text{ such that } \forall x \in H, \qquad \|Tx\| \leq M\,\|x\|.$$

Among bounded operators, there are many further designations: unitary, normal, self-adjoint, compact, Hilbert-Schmidt, Trace and Schatten class operators. Each of these operators have specific properties and uses and each will be touched on in the remainder of thesis. Particular attention will be focused on the latter portion of the list.

As an example of what will be covered regarding these operators, a brief look into trace operators is warranted. The linear transformation A from $\mathbb{R}^n$ to $\mathbb{R}^n$ can be described by an $n \times n$ matrix (a_{ij}). The trace of that $n \times n$ matrix is the sum of the diagonal entries, or

$$Tr(A) = \sum_{i=1}^{n} a_{ii}.$$

This trace is something that is useful for Hilbert Space as well. Yet going from an infinite dimensional space to another infinite dimensional space would not admit a traditional matrix to sum the diagonal terms. So, as with the abstraction of norms and inner products, an abstraction of the trace is necessary to glean the desired information.

The means of the abstraction of operators with finite trace will make judicious use of the Schmidt Decomposition Theorem. That is, for all $x \in H$ a compact operator T can be written as

$$Tx = \sum_{n=1}^{\infty} \alpha_n (x, x_n) y_n$$

for some orthonormal sets (x_n) and (y_n), where α_n is an eigenvalue of $[T^*T]^{1/2}$. This statement ties together a significant portion of the work that will be covered toward the end of this thesis. The Schmidt Decomposition Theorem gives a form for the infinite dimension by infinite dimension matrix that would be impossible to describe in the traditional matrix form.

The formulation and properties of the operators listed above are the goal of the text. Each operator will be defined and put into context. Then examples using each operator will be explored. While practical applications abound, that is not the intent of the thesis and none will be offered. The intent of the text is to familiarize the reader with the theoretical background necessary for such applications.

Chapter 2

The Basics

Planar geometry was the origin of ideas which mathematicians later extended from $\mathbb{R}^2$ to arbitrary dimensional spaces. The familiar absolute value was generalized to the norm denoted $\|\cdot\|$ which gives the length of the vector in the defined space. The dot product used to find angles between vectors in $\mathbb{R}^2$ or in $\mathbb{R}^3$ became the more general inner product denoted $(\cdot,\cdot)$ in arbitrary complex linear spaces. The concept of perpendicular lines was extended via one of its planar properties. In higher dimensions orthogonal lines are defined by the inner product.

In the Euclidean Space $\mathbb{R}^n$, repeated use of the Pythagorean Theorem gives the formula for the 2-norm:

$$\forall x = (x_1, x_2, x_3, \ldots, x_n) \in \mathbb{R}^n,\ \|x\|_2 = \Big(\sum_{i=1}^{n} x_i^2\Big)^{\frac{1}{2}}.$$

But the norm could be defined in a different way such as

$$\forall x = (x_1, x_2, x_3, \ldots, x_n) \in X,\ \|x\|_3 = \Big(\sum_{i=1}^{n} |x_i|^3\Big)^{\frac{1}{3}},$$

or even

$$\forall x = (x_1, x_2, x_3, \ldots, x_n) \in X,\ \|x\|_\infty = max(|x_1|, \ldots, |x_n|).$$

These last two norms are obviously different from the usual norm in $\mathbb{R}^2$, but these norms are not coming from the usual $\mathbb{R}^2$.

2.1 Norm and Inner Product

Definition 2.1. Let X be a complex linear space. A *norm* is a function from X into the non-negative reals $\mathbb{R}^+$ satisfying:

1. $\|x\| = 0$ iff $x = 0$;
2. for each $x \in X$ and each scalar $\alpha \in \mathbb{C}$, $\|\alpha x\| = |\alpha|\|x\|$;
3. the *triangle inequality* $\|x + y\| \leq \|x\| + \|y\|$ holds for all $x, y \in X$.

$(X, \|\cdot\|)$ or X with a norm $\|\cdot\|$ is called a *normed space.*

Definition 2.2. Let (x_n) be a sequence in a normed space X. Then (x_n) is called a *Cauchy* sequence if

$$\forall \epsilon > 0 \ \exists \ N \in \mathbb{Z}^+ \ni m, n \geq N \Rightarrow \|x_n - x_m\| < \epsilon.$$

Definition 2.3. If every Cauchy sequence (x_n) in a normed space X converges in X then X is called a *complete* space.

Definition 2.4. Let X be a linear space over $\mathbb{C}$. An *inner product*, $(\cdot, \cdot)$, is a function defined on $X \times X$, the set of all pairs of elements in X, satisfying:

1. $(x, x) \geq 0$ with equality iff $x = 0$;
2. $(x, y) = \overline{(y, x)}$ where the bar denotes the complex conjugate;
3. $\forall \alpha, \beta \in \mathbb{C}$ and $\forall x_1, x_2, y \in X$; $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$.

$(X, (\cdot, \cdot))$ or X with an *inner product* $(\cdot, \cdot)$ is called an inner product space.

A complete normed linear space is called a Banach space, named after the Polish mathematician Stefan Banach. Hilbert Space – the focus of this paper – is a specific type of those spaces. Hilbert Space is a Banach Space with the added restriction that it has an inner product derived from the norm of the space, more specifically $\forall x \in H$ $\|x\|^2 = (x, x)$.

An important and recurring example of Hilbert Space is ℓ_2. A complex sequence (a_n) is an element of ℓ_2 provided the sum of the magnitude squared of each term is finite. Mathematically

$$\ell_2 = \left\{ (a_n) \middle| \sum_{n=1}^{\infty} |a_n|^2 < +\infty \right\}.$$

When a space is said to be a Hilbert Space, it must be complete and a norm must be given. The space ℓ_2 is well known to be complete. (A proof of a stronger result is detailed in Proposition A.3.) For (a_n), $(b_n) \in \ell_2$ the norm and the induced inner product are given respectively by

$$\|(a_n)\| = \left(\sum_{n=1}^{\infty} |a_n|^2\right)^{\frac{1}{2}}, \quad ((a_n),(b_n)) = \sum_{n=1}^{\infty} a_n \overline{b_n}.$$

2.2 Orthogonality

The idea of perpendicular lines is something that mathematicians would like to have extended to other spaces even when there may not be a ninety degree angle to visualize. It so happens that there is a way to transfer the usefulness of perpendicular lines to any space with an inner product. The usual inner product from $\mathbb{R}^2$ (the dot product), on two perpendicular lines is zero. It is this property of perpendicular lines that will persist regardless of the inner product space. But the word perpendicular is usually reserved for $\mathbb{R}^2$, $\mathbb{C}$ and other spaces where actual ninety degree angles exist. So when dealing with spaces where there aren't angles to measure, the word orthogonal is used. Two different objects in an inner product space are called orthogonal when their inner product is equal to zero.

In ℓ_2 there is an orthogonal basis with unit length, or orthonormal basis, denoted (e_n). This sequence has a zero as each of its terms except for the n^{th} term which is a one. Mathematically,

$$e_i = \underbrace{(0, 0, \ldots, 1, \ldots)}_{\text{one in the } i^{th} \text{ position}} \quad \text{for } i = 1, 2, \ldots$$

It should be noted that $(e_i, e_j) = \delta_{ij}$ and $\|e_i\| = 1$.

One common equality in Hilbert Space theory is Parseval's Equality, named after the French mathematician Antoine Parseval. It is offered here, without proof, in a form slightly modified from the general equality for our uses.

Theorem 2.5. *(Parseval's Equality). For a complete orthonormal set (y_n) of Hilbert Space ℓ_2 and $\forall x \in \ell_2$ it holds that*

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x, y_n)|^2.$$

While this theorem holds for every complete orthonormal set, it is mostly commonly used with the canonical orthonormal basis e_i for ℓ_2.

2.3 Applications

After several definitions it is generally a good idea to give some properties and applications using those definitions. These properties and applications follow.

Proposition 2.6. *In an inner product space X, $(x,u) = (x,v)$ for all x implies that $u = v$.*

Proof. Observe that,

$$\begin{aligned}(x,u) = (x,v) \quad &\Rightarrow \quad (x,u) - (x,v) = 0 \\ &\Rightarrow \quad (x, u-v) = 0.\end{aligned}$$

This has to be true for all x, so in particular for $x = u - v$. This implies $(u-v, u-v) = 0$. By Def. 2.4 part 1, $u - v = 0$ and $u = v$. □

In Hilbert Space a sequence can converge strongly or weakly. As the names imply, a strongly convergent sequence meets stricter requirements than a weakly convergent sequence. So a strongly convergent sequence is automatically weakly convergent. This leads to the question, when does weak convergence imply strong convergence? But first, the definitions of these terms are given.

Definition 2.7. For Hilbert Space H, a sequence $(x_n) \subset H$ is said to *converge strongly* to the point $x \in H$ if $\|x_n - x\| \to 0$ as $n \to \infty$. This type of convergence is also known as norm convergence or convergence in the norm.

Definition 2.8. For Hilbert Space H, a sequence $(x_n) \subset H$ is said to *converge weakly* to $x \in H$ if, for all $y \in H$, $(x_n, y) \to (x, y)$ as $n \to \infty$.

Under an additional condition weak convergence implies strong convergence as shown below.

Proposition 2.9. *In Hilbert Space H let a sequence $(x_n) \subset H$ converge weakly to $x \in H$. If the sequence $(\|x_n\|)$ converges to $\|x\|$ then (x_n) converges strongly to x.*

Proof. Start by using a property of norms

$$\begin{aligned}
\lim_{n\to\infty} \|x_n - x\| &= \lim_{n\to\infty} \sqrt{(x_n - x, x_n - x)} \\
&= \lim_{n\to\infty} \sqrt{(x_n, x_n) - (x_n, x) - \overline{(x_n, x)} + (x, x)} \\
&= \lim_{n\to\infty} \sqrt{\|x_n\|^2 - (x_n, x) - \overline{(x_n, x)} + \|x\|^2} \\
&= \sqrt{\|x\|^2 - (x, x) - \overline{(x, x)} + \|x\|^2} \\
&= 0.
\end{aligned}$$

Thus as n approaches infinity $\|x_n - x\| = 0$ as desired. □

Proposition 2.10. *For x, y in Hilbert Space H, $x \perp y$ iff $\forall \alpha \in \mathbb{C}, \|x + \alpha y\| \geq \|x\|$.*

Proof. Assume $x \perp y$. Then $\forall \alpha \in \mathbb{C}$,

$$\begin{aligned}
(x, x) &= (x, x) + \overline{\alpha}(x, y) + \alpha(y, x) \\
&\leq (x, x) + (x, \alpha y) + (\alpha y, x) + \alpha\overline{\alpha}(y, y) \\
&= (x + \alpha y, x + \alpha y)
\end{aligned}$$

So $\|x\|^2 \leq \|x + \alpha y\|^2$ and $\|x\| \leq \|x + \alpha y\|$.

Assume $\forall \alpha \in \mathbb{C}$, $\|x\| \leq \|x + \alpha y\|$. Then,

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

leaving

$$0 \leq \alpha(y, x) + \overline{\alpha}(x, y) + \alpha\overline{\alpha}(y, y). \tag{2.1}$$

Let $\alpha = \frac{1}{n}$ with $n > 0$. Substituting α into (2.1) above leaves part of the inequality needed. So

$$0 \leq \frac{1}{n}(y, x) + \frac{1}{n}(x, y) + \frac{1}{n^2}(y, y)$$

and

$$-\frac{1}{n^2}(y, y) \leq \frac{1}{n}(y, x) + \frac{1}{n}(x, y). \tag{2.2}$$

Now let $\alpha = -\frac{1}{n}$ with $n > 0$. Substituting α into (2.1) above solves for the remainder of the required inequality.

$$0 \leq -\frac{1}{n}(y,x) + -\frac{1}{n}(x,y) + \frac{1}{n^2}(y,y),$$

$$\frac{1}{n}(y,x) + \frac{1}{n}(x,y) \leq \frac{1}{n^2}(y,y). \tag{2.3}$$

Combining (2.2) and (2.3) above and multiplying through by n yields

$$-\frac{1}{n}(y,y) \leq (y,x) + (x,y) \leq \frac{1}{n}(y,y).$$

Now take the limit as n approaches infinity, so that

$$\lim_{n\to\infty}\left[-\frac{1}{n}(y,y) \leq (y,x) + (x,y) \leq \frac{1}{n}(y,y)\right].$$

Hence, $0 \leq (y,x)+(x,y) \leq 0$ so $0 = (y,x)+(x,y)$. Then letting $(x,y) = \beta+i\gamma$ for $\beta, \gamma \in \mathbb{R}$, it also follows that $(y,x) = \beta - i\gamma$. So $0 = (y,x)+(x,y) \Rightarrow 0 = (\beta+i\gamma)+(\beta-i\gamma) \Rightarrow \beta = 0$. Once it is shown that γ is also zero the proof will be complete.

Return to (2.1) and solve when $\alpha = \frac{i}{n}$ and $\alpha = \frac{-i}{n}$ simultaneously. This leads to a similar inequality

$$-\frac{1}{n}(y,y) \leq i(x,y) - i(y,x) \leq \frac{1}{n}(y,y).$$

Again take the limit as n approaches zero to get

$$0 \leq i(x,y) - i(y,x) \leq 0$$

and

$$i(x,y) - i(y,x) = 0.$$

Since (x,y) has been defined as $\beta + i\gamma$ the equation turns into

$$i(\beta+i\gamma) - i(\beta - i\gamma) = 0 \Rightarrow i\beta - \gamma - i\beta - \gamma = 0 \Rightarrow \gamma = 0.$$

Since both β and γ are both zero, $(x,y) = 0$ and $x \perp y$. $\square$

Chapter 3

Introduction to Operators

Hilbert Space has been carefully created to mimic Euclidean Space. As a result, research in the area doesn't need to focus on the space itself. Interest is usually concentrated on the functions – or more precisely operators – that act on the space and the consequences of that action.

Functions are defined as a mapping from one set to another. In general terms, a function is simply a set of rules which describes how elements in one set are mapped to elements in another set. Restrictions are usually placed on these functions depending on the focus of study. For the purpose of Hilbert Space, the functions are required to be linear. In any linear space, function f is linear when for x, y in the domain and scalars a, b it holds that

$$f(ax + by) = af(x) + bf(y).$$

Linear functions from Hilbert Space H to Hilbert Space K are called linear operators.

Definition 3.1. In Hilbert Spaces H and K an operator $T : H \to K$ is called a *bounded operator* if

$$\exists M \in \mathbb{R}^+ \text{ such that } \forall x \in H, \qquad \|Tx\| \leq M \|x\| .$$

The set of all linear bounded operators T on Hilbert Space H to K is denoted $L(H, K)$. That is,

$$L(H, K) = \{T : H \to K \mid T \text{ is linear and bounded}\}.$$

Operators are treated as objects in their own right. As an object, it is necessary to know what the norm of that object is.

Definition 3.2. For $T \in L(H)$, the *norm of* T is defined by

$$\|T\| = \sup\{\|Tx\| \mid \|x\| \leq 1\}.$$

It is routine to verify that $\|\cdot\|$ is a norm on $L(H)$. Also, it should be noted that if $T \notin L(H)$ the above definition does not define a norm on a Hilbert Space. It is boundedness that forces $\|Tx\|$ to be finite for all $x \in H$ with $\|x\| \leq 1$.

Another common property of operators defined on a space is that of continuity. The following definition focuses on Hilbert Space operators, but it could easily be modified to work in a general Banach Space.

Definition 3.3. An operator T defined on Hilbert Space H is *continuous* when every sequence (x_n) which converges strongly to x implies that the sequence (Tx_n) converges strongly to Tx.

Theorem 3.4. *For Hilbert Spaces H, K and linear operator T from H to K the following are equivalent:*

1. *T is bounded,*

2. *T is continuous at* 0*,*

3. *T is continuous over H.*

Theorem 3.4 brings together two notable properties of operators, boundedness and continuity. An operator being continuous is an important property. However, to minimize redundancy, continuity is not usually listed as a property of an operator. The phrase bounded operator is sufficient.

A few quick examples will help to solidify this theory.

Example 3.5. *For the following examples, let the Hilbert Space $H = \ell_2$ and let the sequence (e_i) be a countable orthonormal basis in ℓ_2.*

1. Let $Te_i = e_{i+1}$ for $i \geq 1$. This operator is known as a shift operator. Then $\|T\| = 1$.

   Proof. Let $x = \sum_{i=1}^{\infty}(x, e_i)e_i$ which implies $Tx = \sum_{i=1}^{\infty}(x, e_i)e_{i+1}$.

Parseval's equation (Theorem 2.5) states that

$$\|Tx\|^2 = \sum_{i=1}^{\infty} |(x, e_i)|^2 = \|x\|^2$$

or $\|Tx\| = \|x\|$. Thus $\|T\| = 1$. □

2. Let $Te_i = e_i + e_{i+1}$. Then $\|T\| = 2$.

Proof. It is clear that $\|e_i\| = 1$ and $\|Te_i\| = \|e_i + e_{i+1}\| = \sqrt{2}\ \forall i$. For $x \in \ell_2$,

$$x = \sum_{i=1}^{\infty}(x, e_i)e_i \quad \text{and} \quad Tx = \sum_{i=1}^{\infty}(x, e_i)(e_i + e_{i+1}).$$

Hence

$$\begin{aligned}
\|Tx\|^2 &= \left\|(x, e_1)e_1 + \sum_{i=2}^{\infty}[(x, e_i) + (x, e_{i-1})]e_i\right\|^2 \\
&= |(x, e_1)|^2 + \sum_{i=2}^{\infty}|(x, e_i) + (x, e_{i-1})|^2 \\
&\leq |(x, e_1)|^2 + 2\sum_{i=2}^{\infty}[|(x, e_i)|^2 + |(x, e_{i-1})|^2] \\
&\leq 4\|x\|^2
\end{aligned}$$

Thus $\|Tx\| \leq 2\|x\|$ and $\|T\| \leq 2$. Now it must be shown that $\|T\| \to 2$. Let $f_n = \frac{1}{\sqrt{n}}(e_1 + e_2 + \cdots e_n)$. Then, $\|f_n\| = 1$ and $Tf_n = \frac{1}{\sqrt{n}}(e_1 + 2e_2 + \cdots 2e_n + e_{n+1})$. So $\|Tf_n\| = \frac{1}{\sqrt{n}}\sqrt{1 + 4(n-1) + 1} = \sqrt{4 - 2/n}$. Hence, $\|Tf_n\| \to 2$ as $n \to \infty$. Since the norm is the supremum of all possible values, $\|T\| = 2$. □

Definition 3.6. The *dual space* H^* of Hilbert Space H is the set of all linear bounded operators, in this case linear functionals, T such that T maps elements of H to the scalar field $\mathbb{C}$.

The dual space H^* of H is identified with H itself as seen from the following.

Theorem 3.7. *(Riesz Representation Theorem). Let H be a Hilbert Space.*

A) Let $y \in H$. Define $f_y(x) = (x, y)$. Then $f_y \in H^*$ and $\|f_y\| = \|y\|$.

B) Let $f \in H^*$. Then there is a unique $y \in H$ such that $f(x) = (x, y)\ \forall x \in H$. Moreover, $\|f\| = \|y\|$.

Chapter 4

Real Versus Complex Hilbert Space

Recall that the inner product has, as part of its definition, a property $(x, y) = \overline{(y, x)}$. Calculations concerning inner products of complex terms can quickly become more involved than the same calculation with only real terms. However, as it will be shown, the real case may be too simple and important information can be lost. A few examples are in order.

4.1 Properties of Real Hilbert Spaces

In a real Hilbert Space H, $(x, y) = \overline{(y, x)}$ can be written as $(x, y) = (y, x)$. This leads to simplifications that would be impossible if H is a complex Hilbert Space.

Proposition 4.1. *In a real inner product space H, if $x, y \in H$ and $\|x\| = \|y\|$ then $\|ax + by\| = \|bx + ay\|$ $\forall a, b \in \mathbb{R}$.*

Proof. Observe that

$$\begin{aligned}
\|ax + by\|^2 &= (ax + by, ax + by) \\
&= (ax + by, ax) + (ax + by, by) \\
&= (ax, ax) + (by, ax) + (ax, by) + (by, by) \\
&= a^2\|x\|^2 + ba(y, x) + ab(x, y) + b^2\|y\|^2 \text{ (given } \|x\| = \|y\|.)
\end{aligned}$$

$$
\begin{aligned}
&= a^2\|y\|^2 + ab(y,x) + ba(x,y) + b^2\|x\|^2 \\
&= (ay,ay) + (ay,bx) + (bx,ay) + (bx,bx) \\
&\qquad\qquad \text{(taking this back to a single inner product)} \\
&= (ay+bx, ay+bx) \\
&= \|ay+bx\|^2.
\end{aligned}
$$

Finally, taking the square root of both sides leaves $\|ax+by\| = \|bx+ay\|$. □

The definition of an isosceles triangle is well known. Plainly, it's a triangle that has two sides of equal length. This idea is generalized to a real Hilbert Space in the following theorem.

Theorem 4.2. *(The Isosceles Triangle Property) Let H be a real Hilbert Space. If $x, y, z \in H$, $x+y+z=0$ and $\|x\| = \|y\|$, then $\|x-z\| = \|y-z\|$.*

Proof. It holds that

$$
\begin{aligned}
\|x-z\|^2 &= \|2x+y\|^2 \text{ (since } x+y+z=0) \\
&= (2x+y, 2x+y) \\
&= (2x,2x) + (2x,y) + (y,2x) + (y,y) \\
&= 4(x,x) + (2x,y) + (y,2x) + (y,y) \\
&= 4\|x\|^2 + 2(x,y) + 2(y,x) + \|y\|^2 \text{ (since } \|x\| = \|y\|) \\
&= 4\|y\|^2 + (x,2y) + (2y,x) + \|x\|^2 \\
&= 4(y,y) + (x,2y) + (2y,x) + (x,x) \\
&= (2y,2y) + (x,2y) + (2y,x) + (x,x) \\
&= (2y+x, 2y+x) \\
&= \|2y+x\|^2 \\
&= \|y-z\|^2 \text{ (since } x+y=-z).
\end{aligned}
$$

Then taking the square root of both sides gives the desired equality. □

4.2 Properties of Complex Hilbert Spaces

Having $(x,y) = (y,x)$ can obviously be beneficial in some cases. But sometimes this property can allow terms to cancel even when they are necessary for the completion of a proof.

Proposition 4.3. *For a real Hilbert Space H, $x, y \in H$ and $\|x\| = \|y\|$ imply that $(x+y, x-y) = 0$. If H is complex, this assertion is not true.*

Proof. 1. For a real Hilbert Space H,

$$\begin{aligned}(x+y, x-y) &= (x, x-y) + (y, x-y) \\ &= (x,x) - (x,y) + (y,x) - (y,y) \\ &= \|x\|^2 - (x,y) + \overline{(x,y)} - \|y\|^2 \\ &= \|x\|^2 - \|x\|^2 - (x,y) + (x,y) = 0.\end{aligned}$$

2. For the complex case, a counterexample will suffice.
Let H be $\mathbb{C}$ and let $x = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$ and $y = 1$. Clearly $\|x\| = \|y\|$. But

$$(x+y, x-y) = \left(\frac{\sqrt{2}+2}{2} + i\frac{\sqrt{2}}{2}, \frac{\sqrt{2}-2}{2} + i\frac{\sqrt{2}}{2}\right).$$

Taking the inner product

$$\left(\frac{\sqrt{2}+2}{2} + i\frac{\sqrt{2}}{2}\right)\left(\overline{\frac{\sqrt{2}-2}{2} + i\frac{\sqrt{2}}{2}}\right) = -\sqrt{2}i \neq 0.$$

So for the x and y given above $\|x\| = \|y\|$ but $(x+y, x-y) \neq 0$. □

Proposition 4.4. *In complex Hilbert Space H, $T \in L(H)$ and $(Tx, x) = 0$ $\forall x \in H$ implies $T = 0$. This is false if H is real.*

Proof. 1. If H is complex:
$\big(T(x+y), x+y\big) = 0$ by assumption. Now,

$$\begin{aligned}\big(T(x+y), x+y\big) &= (Tx + Ty, x+y) \\ &= (Tx, x) + (Tx, y) + (Ty, x) + (Ty, y) \\ &\qquad \text{(given the } 1^{st} \text{ and } 4^{th} \text{ terms are both zero.)} \\ &= (Tx, y) + (Ty, x).\end{aligned}$$

So $0 = \big(T(x+y), x+y\big) = (Tx, y) + (Ty, x)$
Similarly with $\big(T(ix+y), ix+y\big) = 0$ we have

$$\big(T(ix+y), ix+y\big) = (Tix + Ty, ix+y)$$

$$
\begin{aligned}
&= (Tix, ix) + (Tix, y) + (Ty, ix) + (Ty, y) \\
&\qquad\qquad \text{(given the } 1^{st} \text{ and } 4^{th} \text{ terms are both zero.)} \\
&= (Tix, y) + (Ty, ix) \\
&= i(Tx, y) + \bar{i}(Ty, x) \\
&= i[(Tx, y) - (Ty, x)].
\end{aligned}
$$

Once again $0 = \big(T(ix + y), ix + y\big) = (Tx, y) - (Ty, x)$.
This gives two equations and two unknowns. Namely,

$$
\begin{aligned}
(Tx, y) + (Ty, x) &= 0 \\
(Tx, y) - (Ty, x) &= 0
\end{aligned}
$$

So $(Tx, y) = (Ty, x) = 0\ \forall x, y \in H$. And therefore $T = 0$.

2. If H is real:

Here is a counterexample. Let $T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and let $x \in \mathbb{R}^2$ with $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Then

$$
Tx = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.
$$

Set $y = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$. Does $(Tx, x) = 0, \forall x \in \mathbb{R}^2$ imply $T = 0$? Well, $(Tx, x) = (y, x)$ and the inner product in $\mathbb{R}^2$ is the dot product. So (y, x) with x and y defined as above is $(-x_2)(x_1) + (x_2)(x_1) = -x_2x_1 + x_2x_1 = 0$. Thus $\exists\ T \in L(H)$ such that $(Tx, x) = 0\ \forall x \in \mathbb{R}^2$ but $T \neq 0$. $\square$

Chapter 5

Unitary Operators

The next notion to be covered is the extremely useful adjoint of operators. The adjoint of an operator is actually a direct result of the Riesz Representation Theorem.

Definition 5.1. Let $T \in L(H)$ and let $y \in H$. Define $S \in H^*$ such that

$$S(x) = (Tx, y).$$

The Riesz Representation Theorem (3.7) states $\exists z \in H$ such that $\forall x \; S(x) = (x, z)$. So $(Tx, y) = (x, z)$. Then there is an operator that maps y onto z. This operator is T^* and is called the *adjoint* of T. T^* is a bounded linear operator and $\|T^*\| = \|T\|$.

A unitary operator is one which is defined by special properties involving its adjoint. The definition of this operator follows.

Definition 5.2. *Unitary Operators.* Let $T \in L(H)$. If it is given that $T^{-1} \in L(H)$ and if $T^* = T^{-1}$ then T is said to be *unitary.*

Before a theorem regarding unitary operators, a brief lemma is in order. Let ϕ be a Hermitian bilinear form on $H \times H$. That is, ϕ is linear in the first term and conjugate linear in the second. Let $\psi(x) = \phi(x, x)$. Note that ψ is called a quadradic form associated with ϕ.

Lemma 5.3. *Given ϕ and ψ as defined above, it follows that*

$$\psi\left(\frac{x+y}{2}\right) - \psi\left(\frac{x-y}{2}\right) + i\psi\left(\frac{x+iy}{2}\right) - i\psi\left(\frac{x-iy}{2}\right) = \phi(x, y).$$

Proof. Observe that

$$\begin{aligned}\psi\left(\frac{x+y}{2}\right) &= \phi\left(\frac{x+y}{2}, \frac{x+y}{2}\right)\\ &= \frac{1}{4}\phi(x+y, x+y)\\ &= \frac{1}{4}[\phi(x,x)+\phi(x,y)+\phi(y,x)+\phi(y,y)],\end{aligned}$$

$$\begin{aligned}\psi\left(\frac{x-y}{2}\right) &= \phi\left(\frac{x-y}{2}, \frac{x-y}{2}\right)\\ &= \frac{1}{4}\phi(x-y, x-y)\\ &= \frac{1}{4}[\phi(x,x)-\phi(x,y)-\phi(y,x)+\phi(y,y)].\end{aligned}$$

Combining the first two terms consolidates to

$$\psi\left(\frac{x+y}{2}\right) - \psi\left(\frac{x-y}{2}\right) = \frac{1}{2}[\phi(x,y)+\phi(y,x)].$$

Similar calculations concerning the second pair of terms shows

$$i\psi\left(\frac{x+iy}{2}\right) - i\psi\left(\frac{x-iy}{2}\right) = \frac{1}{2}[\phi(x,y)-\phi(y,x)].$$

Lastly, combining the two intermediate results leaves

$$\psi\left(\frac{x+y}{2}\right) - \psi\left(\frac{x-y}{2}\right) + i\psi\left(\frac{x+iy}{2}\right) - i\psi\left(\frac{x-iy}{2}\right) = \phi(x,y)$$

as desired. □

It is important to note that the inner product is also a Hermitian form, so the above lemma applies to the following theorem.

Theorem 5.4. *$T \in L(H)$ is unitary iff T is an isometric isomorphism.*

Proof. Given T is unitary, $\exists\ T^{-1} \in L(H)$ such that $T^* = T^{-1}$. $L(H)$ is defined to be the set of continuous operators so T and T^{-1} are continuous. The condition $T^* = T^{-1}$ implies that $TT^{-1} = T^{-1}T$ and therefore T is one-to-one and onto. So if T is unitary then T is an isomorphism. Then to show T is an isometry:

$$\|Tx\|^2 = (Tx, Tx)$$

$$
\begin{aligned}
&= (T^*Tx, x) \\
&= (T^{-1}Tx, x) \qquad \text{since } T \text{ is unitary} \\
&= (x, x) \\
&= \|x\|^2
\end{aligned}
$$

So if T is unitary then T is an isometry and therefore T is an isometric isomorphism.

Given T is an isometric isomorphism, it follows that $\forall x, y \in H$

$$
\begin{aligned}
(T^*Tx, y) &= (Tx, Ty) \\
&= \frac{1}{4}\left[\|Tx + Ty\|^2 - \|Tx - Ty\|^2 + i\|Tx + iTy\|^2 - i\|Tx - iTy\|^2\right] \\
&\qquad\qquad \text{(by Lemma 5.3.)} \\
&= \frac{1}{4}\left[\|T(x + y)\|^2 - \|T(x - y)\|^2 + i\|T(x + iy)\|^2 - i\|T(x - iy)\|^2\right] \\
&= \frac{1}{4}\left[\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2\right] \\
&= (x, y).
\end{aligned}
$$

So $(T^*Tx, y) = (x, y)\ \forall x, y \in H \implies T^*T = I \implies T^* = T^{-1} \implies T$ is unitary. $\square$

Chapter 6

Compact Operators

In analysis, a significant amount of time is spent working on compactness of sets and spaces. This is for several good reasons. In operator theory, it would be nice to have operators that take bounded sets to compact sets. The following definition formalizes this idea.

Definition 6.1. Given the unit ball $A = \{x \in H : \|x\| \leq 1\}$ a *compact operator* is an operator $T \in L(H)$ such that $\overline{T(A)}$ is compact. The set of all compact operators is denoted $K(H)$.

When it is given that an operator is compact, it would be convenient to know if related operators are also compact.

Proposition 6.2. *Let H be a Hilbert space.*

1. $T \in K(H) \iff T^* \in K(H)$;
2. *If $T \in K(H), S \in L(H)$ then both TS and ST are elements of $K(H)$;*
3. $T \in K(H) \iff T^*T \in K(H)$;
4. $T \in K(H) \iff TT^* \in K(H)$;
5. *$T \in K(H)$ and $T^{-1} \in L(H)$ imply that H is finite dimensional;*
6. *If $T \in L(H)$ has finite rank, then $T \in K(H)$.*

In $\mathbb{R}^2$ and $\mathbb{C}$ there is a well known result that states every bounded sequence has a convergent subsequence. For Hilbert space this isn't the case. However, given a few more restrictions a similar result can be obtained.

First, a quick definition. A Hilbert space is said to be *separable* if it contains a countable, dense subset.

Theorem 6.3. *Every bounded sequence (x_n) in a separable Hilbert Space H has a weakly convergent subsequence.*

Proof. Space H is separable so $\exists (y_i) \in H$ – a countable dense set which is treated as a sequence. Define

$$A = \{(x_1, y_1), (x_2, y_1), (x_3, y_1), \ldots\}.$$

A is a sequence of complex numbers, as consequently contains a convergent subsequence A_1. That is, there is a subsequence $(x_{1,k})$ of (x_n) such that the subsequence $A_1 = \{(x_{1,k}, y_1)\}$ converges. This particular subsequence of (x_n) will referred to as the first subsequence. So the first convergent sequence, which converges to a_1, is

$$A_1 = \{(x_{1,1}, y_1), (x_{1,2}, y_1), (x_{1,3}, y_1), \ldots\}, \quad (x_{1,k}, y_1) \to a_1.$$

After choosing y_2 from (y_i) there is a second subsequence $(x_{2,k})$ of (x_n) such that $(x_{2,k}) \subseteq (x_{1,k})$ and

$$A_2 = \{(x_{2,1}, y_2), (x_{2,2}, y_2), (x_{2,3}, y_2), \ldots\}, \quad (x_{2,k}, y_2) \to a_2.$$

Continuing this process – the underlined parts would be explained later – gives the following sequences:

$$\begin{aligned}
A_1 &= \{\underline{(x_{1,1}, y_1)}, (x_{1,2}, y_1), (x_{1,3}, y_1), \ldots\}, \quad (x_{1,k}, y_1) \to a_1 \\
A_2 &= \{(x_{2,1}, y_2), \underline{(x_{2,2}, y_2)}, (x_{2,3}, y_2), \ldots\}, \quad (x_{2,k}, y_2) \to a_2 \\
A_3 &= \{(x_{3,1}, y_3), (x_{3,2}, y_3), \underline{(x_{3,3}, y_3)}, \ldots\}, \quad (x_{3,k}, y_3) \to a_3 \\
&\vdots \\
A_m &= \{(x_{m,1}, y_m), (x_{m,2}, y_m), (x_{m,3}, y_m), \ldots, \underline{(x_{m,m}, y_m)}, \ldots\}, \quad (x_{m,k}, y_m) \to a_m \\
&\vdots
\end{aligned}$$

It also follows that $(x_n) \supseteq (x_{1,k}) \supseteq (x_{2,k}) \supseteq \ldots \supseteq (x_{m,m}) \supseteq \ldots$

A new sequence is created using the underlined terms in the above sequences:

$$B = \{(x_{1,1}, y_1), (x_{2,2}, y_2), (x_{3,3}, y_3), \dots, (x_{m,m}, y_m), \dots\}.$$

And $\forall i$, as m approaches infinity $(x_{m,m}, y_i) \to a_i$. This is a direct result of each sequence being a subsequence of the preceeding sequence.

As a matter of notation $x_{m,m}$ will be written as z_m.

Let M be a bound for (x_n). Let $w \in H$ and $\epsilon > 0$ be arbitrary. Then $\exists y_i$ such that $\|w - y_i\| < \frac{\epsilon}{3M}$. Since $\big((z_m, y_i)\big)$ is a Cauchy sequence, $\exists N \geq 1$ such that $|(z_\ell - z_j, y_i)| < \frac{\epsilon}{3}$, $\forall \ell, j \geq N$. Thus we have that

$$\begin{aligned}
|(z_\ell, w) - (z_j, w)| &= |(z_\ell, w) - (z_\ell, y_i) + (z_\ell, y_i) - (z_j, y_i) + (z_j, y_i) - (z_j, w)| \\
&\leq |(z_\ell, w) - (z_\ell, y_i)| + |(z_\ell, y_i) - (z_j, y_i)| + |(z_j, y_i) - (z_j, w)| \\
&= |(z_\ell, w - y_i)| + |(z_\ell - z_j, y_i)| + |(z_j, y_i - w)| \\
&\leq \|z_\ell\| \cdot \|w - y_i\| + |(z_\ell - z_j, y_i)| + \|z_j\| \cdot \|y_i - w\| \\
&< M\left(\frac{\epsilon}{3M}\right) + \left(\frac{\epsilon}{3}\right) + M\left(\frac{\epsilon}{3M}\right) \\
&= \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

Hence $\big((z_m, w)\big)$ is convergent. So $(x_{m,m})$ is weaky Cauchy and is therefore weakly convergent. □

Riesz worked with operators which he called completely continuous. According to Riesz, an operator $T \in L(H)$ is completely continuous if T maps weakly convergent sequences to norm convergent sequences–these terms are defined in Definitions 2.7 and 2.8. While the phrase completely continuous has fallen out of favor, the work that Riesz did with this type of operator is worthwhile. It turns out that the operator called completely continuous coincides perfectly with the compact operator defined in this chapter. While the definitions of these operators are different, they are logically equivalent. This, of course, must be shown. But first, a brief lemma.

Lemma 6.4. *For weakly convergent $(x_n) \in H$, (x_n) is bounded.*

Proposition 6.5. *Let H be a separable Hilbert space. $T \in K(H)$ iff T is completely continuous in the Riesz sense.*

Proof. Let $T \in K(H)$. If T were not completely continuous then there exists a weakly convergent sequence (x_n) that converges weakly to x such that

$$\lim_{n\to\infty} \|Tx_n - Tx\| \neq 0.$$

Then there is an $\epsilon > 0$ and a subsequence $(Tx_{n,1})$ of (Tx_n) such that

$$\|Tx_{n,1} - Tx\| \geq \epsilon, \ \forall n \geq 1.$$

Using the weak convergence of (x_n) and Lemma 6.4, (x_n) has a bound. Thus $(Tx_{n,1})$ has a convergent subsequence $(Tx_{n,2})$. Define m such that $(Tx_{n,2})$ converges to m. It is given that $(Tx_{n,2})$ converges weakly to Tx so $m = Tx$. So $\|Tx_{n,2} - Tx\|$ approaches 0, a contradiction. T is therefore completely continuous.

For the completely continuous operator T to also be compact, it must map the unit ball $\{x \in H : \|x\| \leq 1\}$ into a set whose closure is compact.

Let $A = \{x \in H : \|x\| \leq 1\}$ and let $T(A) = B$. If it can be shown that every sequence in B has a subsequence converging to some member of $\overline{B}$, then B is relatively compact. So T would then map the unit ball to a relatively compact set and $T \in K(H)$ as desired.

Let $\{y_k\} \subset B$. There exists a sequence $\{x_k\} \subset A$ such that $Tx_k = y_k \ \forall k$. Note that $\{x_k\}$ is bounded because $\|x\| \leq 1$. So by Theorem 6.3 $\exists \{z_n\} \subset \{x_k\}$ such that $\{z_n\}$ is weakly convergent. Also (Tz_n) is norm convergent since T is completely continuous. So B is relatively compact. Therefore T is compact. □

Compact operators are one of the main focuses throughout the rest of the text and are worthy of further investigation.

While this next proposition is instructive, it serves another purpose as well. It offers insight into a topic that will be covered in the next chapter. Keep this proposition in mind when reading Chapter 7.

Proposition 6.6. *Let $T \in K(H)$. If $\lambda \neq 0$ then* $\ker(\lambda I - T)$ *has finite dimension.*

Proof. Assume $\ker(\lambda I - T)$ has infinite dimension. Then $\exists \{x_n\} \subseteq \ker(\lambda I - T)$ such that $\{x_n\}$ is an infinite orthonormal set. (Otherwise $\ker(\lambda I - T)$ would have a finite orthonormal set which would serve as a finite basis for the space. For the moment, this is assumed not to be the case.)

Since T is compact and $\{x_n\}$ is weakly convergent to 0 there exists a convergent subsequence of $\{x_n\} = \{\frac{1}{\lambda}Tx_n\}$. But $\forall x_i, x_j \in \{x_n\}, i \neq j \Rightarrow \|x_i - x_j\| = \sqrt{2}$. So there is no such convergent subsequence of $\{x_n\}$. A contradiction. Therefore $\ker(\lambda I - T)$ has finite dimension. $\square$

Theorem 6.7. *An operator $T \in K(H)$ iff there are finite rank operators A_n such that*

$$\lim_{n\to\infty} \|T - A_n\| = 0.$$

An outline of the proof follows.

Assume there is such an A_n. Choose a bounded sequence (x_n) and use the compactness of A_n to find a subsequence $(x_{n,k})$ of (x_n) such that $(A_n(x_{n_m}))$ converges. This sequence is Cauchy. This is used to show that $(T(x_{n_m}))$ is Cauchy and therefore T is compact.

Then assuming $T \in K(H)$, let (u_i) be a complete orthonormal set. The inner product expansion of T and of the orthogonal projection P_n would be

$$Tx = \sum_{i=1}^{\infty}(Tx, u_i)u_i \qquad \text{and} \qquad P_n x = \sum_{i=1}^{n}(x, u_i)u_i$$

respectively. It can be shown that the required A_n is given by the composition $A_n = P_nT$. Thus

$$\begin{aligned} \|T - A_n\| &= \|T - P_nT\| \\ &= \sup_{\|x\|\leq 1} \|(T - P_nT)x\| \\ &= \sup_{\|x\|\leq 1} \left\| \sum_{i=n+1}^{\infty} (Tx, u_i)u_i \right\|. \end{aligned}$$

Since $\|x\| \leq 1$, the closure of Tx is compact. The above sum tends to zero for large enough n when Tx is in a compact set. So $\lim_{n\to\infty} \|T - A_n\| = 0$. $\square$

Chapter 7

The Spectrum

It is a straightforward calculation to find the eigenvalues of a transformation defined by a matrix. This procedure is taught in an elementary linear algebra course. Unfortunately, this procedure fails when the transformation is an operator in Hilbert Space. The obvious difficulty is that there is no matrix to calculate the eigenvalues from, but the heart of the matter remains. If H is a Hilbert Space, is there a way to find out when $Tx - \lambda x$ is not invertible?

The following should seem familiar, as it is the same reasoning that is given when explaining the algorithm for finding eigenvalues in basic linear algebra. $Tx = \lambda x$ implies

$$\begin{aligned} Tx - \lambda x &= 0, \\ (T - \lambda I)x &= 0, \qquad I \text{ being the identity matrix.} \end{aligned}$$

This has a non-trivial solution when $T - \lambda I$ is not invertible. It is here that the explanation for the purposes of this paper diverges from the explanation given in linear algebra. In linear algebra the procedure concluded with writing down the matrix that corresponded with T and subtracting λI. Then to find out where $T - \lambda I$ is not invertible, this procedure dictates setting the determinant of the above difference to zero and solving for λ.

In the Hilbert Space version, there is no matrix. Since no calculations can be done, the generalization stops at $(Tx - \lambda I)$ being not invertible. In the Hilbert Space version, the result of this process is called the spectrum.

Definition 7.1. Let $T \in L(H)$. The *spectrum* of T, denoted $\sigma(T)$, is defined to be exactly what it should be to match up with the equations above. Namely,

$$\sigma(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ is not invertible}\}.$$

What of the elements of $\mathbb{C}$ that are not in $\sigma(T)$?

Definition 7.2. Let $T \in L(H)$. The *resolvent* of T, or $\rho(T)$, is the set of all complex numbers not in $\sigma(T)$. That is,

$$\rho(T) = \mathbb{C} \backslash \sigma(T).$$

The spectrum of arbitrary operators in Hilbert Space can be complicated. Later work with $\sigma(T)$ will be restricted to operators that have more structure than most arbitrary operators– compact operators being a notable example. This eases some of the complication. This chapter, however, deals with the spectrum of arbitrary operators.

For an arbitrary operator T, $\sigma(T)$ can be partitioned into three pairwise disjoint sets:

$$\sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T).$$

The properties that distinguish these sets are as follows, for $\lambda \in \sigma(T)$: $\lambda \in \sigma_p(T)$– or the point spectrum– when $\lambda I - T$ is not one to one; $\lambda \in \sigma_r(T)$– or the residual spectrum– when $\lambda I - T$ is one to one but the range of $\lambda I - T$ is not dense in H; $\lambda \in \sigma_c(T)$–or the continuous spectrum– when $\lambda I - T$ is one to one and has a dense range. It is important to note that the elements of $\sigma_p(T)$ are the eigenvalues of T.

An exercise in partitioning the spectrum will be instructive. But first, a theorem to aid in that exercise. The proof of the theorem leads off topic and is omitted.

Theorem 7.3. *The set $\sigma(T)$ is closed and bounded for $T \in L(H)$.*

Example 7.4. *Let T be an operator in $L(\ell_2)$ such that*

$$Tx = \sum_{n=1}^{\infty} (x, e_{n+1}) e_n,$$

where (e_n) is the standard basis of ℓ_2. Show that if $|\lambda| \leq 1$ then $\lambda \in \sigma(T)$. Partition the spectrum.

Solution. First, find the actual eigenvalues of T. Let $(\lambda I - T)(x) = \lambda x - Tx = 0$, i.e., $\lambda \in \sigma_p$. Then

$$\lambda \sum_{n=1}^{\infty} (x, e_n) e_n - \sum_{n=1}^{\infty} (x, e_{n+1}) e_n = \sum_{n=1}^{\infty} \left[\lambda(x, e_n) - (x, e_{n+1})\right] e_n = 0.$$

Since $\{e_i\}$ is linearly independant, we have

$$\begin{aligned}
\lambda(x,e_1)-(x,e_2)=0 &\;\Rightarrow\; \lambda(x,e_1)=(x,e_2)\\
\lambda(x,e_2)-(x,e_3)=0 &\;\Rightarrow\; \lambda^2(x,e_1)=(x,e_3)\\
&\;\vdots\\
\lambda(x,e_n)-(x,e_{n+1})=0 &\;\Rightarrow\; \lambda^n(x,e_1)=(x,e_{n+1}),
\end{aligned}$$

$$\sum_{n=1}^{\infty}\lambda^{n-1}(x,e_1)e_n=\sum_{n=1}^{\infty}(x,e_n)e_n=x.$$

So

$$x=\sum_{n=1}^{\infty}\lambda^{n-1}(x,e_1)e_n.$$

This must be true for all $x\in\ell_2$. When $(x,e_1)\neq 0$, we have $0<|\lambda|<1$. This also implies that if $|\lambda|>1$ then $\lambda\in\rho(T)$.

Secondly, if $\lambda=0$ and $Tx=0$ then $0(x,e_1)-(x,e_2)=0$ and $(x,e_2)=0$. This leads to $x=(x_1,0,0,\ldots)$ but $Tx=0$. So $\lambda I-T$ is not one to one and $\lambda=0$ is an eigenvalue.

Since $\mathbb{C}_0=\{z\in\mathbb{C}\mid |z|<1\}\subset\sigma(T)$ and $\sigma(T)$ is closed, we have that $\lambda_1\in\mathbb{C}_1=\{z\in\mathbb{C}\mid |z|=1\}\subset\sigma(T)$. Then $\lambda_1\notin\sigma_r(T)$ because $\lambda\in\sigma_r(T)$ implies that $\lambda I-T$ has a dense range. Also $\lambda_1\notin\sigma_p(T)$ because $\lambda\in\sigma_p(T)$ implies that $\lambda I-T$ is one to one. Thus $\lambda_1\in\sigma_c(T)$ since $\sigma_p(T)$, $\sigma_r(T)$ and $\sigma_c(T)$ are disjoint and their union is $\sigma(T)$.

This completes the partition. $\square$

Chapter 8

Self-Adjoint Operators

General operators with no restrictions are sometimes difficult to work with. So much so, the very first topic regarding on operators in a Hilbert Space was the restriction to bounded and linear operators. The introduction of the adjoint of operators allowed work with inner products that had previously been impossible. Even further refinements were given in Chapter 5. Many interesting results arose when the general theory of operators was narrowed to unitary operators. Another such refinement of the adjoint of operators will continue this specialization of general Hilbert Space theory.

Definition 8.1. A *self-adjoint operator* of a Hilbert Space H is an operator $T \in L(H)$ such that $T = T^*$. In other words,

$$(Tx, y) = (x, Ty), \quad \forall x, y \in H.$$

This definition has some major consequences. Given Hilbert Space H, for self-adjoint operator T and $\forall x \in H$, $(Tx, x) = (x, Tx)$. (It is important to note that, by definition, T being self-adjoint implies that $T \in L(H)$.) But part of the definition of inner product states that $(Tx, x) = \overline{(x, Tx)}$. From these equalities, it is obvious that (Tx, x) is real for all $x \in H$ – a significant restriction.

One important property that results from working with strictly real terms is defined next. This definition revolves around the fact that the real line is well ordered. It follows that some sort of ordering can be imposed on self-adjoint operators.

Definition 8.2. In Hilbert Space H, a self-adjoint operator T is said to be *nonnegative*, denoted $T \geq 0$, if

$$\forall x \in H, (Tx, x) \geq 0.$$

Similarly, for self-adjoint operators A and B,

$$A \geq B \text{ iff } A - B \geq 0.$$

There are some basic properties of self-adjoint operators that will prove useful in later work.

Proposition 8.3. *In Hilbert Space H, for operator $T \in L(H)$ and for self-adjoint operators $A, B, C \in L(H)$:*

1. $A \geq 0$ *and* $B \geq 0 \Rightarrow A + B \geq 0$*;*
2. $A \geq B$ *and* $B \geq A \Rightarrow A = B$*;*
3. $A \geq B$ *and* $B \geq C \Rightarrow A \geq C$*;*
4. $T^*T \geq 0$ *and* $TT^* \geq 0$.

The first three statements in the above proposition are rather trivial. However, the last deserves a second look as it will be used frequently in later work.

Proof. In Hilbert Space H, let $T \in L(H)$. Then $\forall x \in H$, $(Tx, Tx) \geq 0$ hence $(T^*Tx, x) \geq 0$. Then by Definition 8.2 $T^*T \geq 0$. For TT^*, the same process with (T^*x, T^*x) will suffice. $\square$

This partial ordering of the self-adjoint operators of $L(H)$ allows some interesting results with the spectrum from the previous chapter.

Theorem 8.4. *If T is self-adjoint then $\sigma(T)$ has only real elements. Also, if $T \geq 0$, then $\sigma(T) \subseteq [0, +\infty)$.*

Theorem 8.5. *For self-adjoint T and Hilbert Space H, if $(Tx, x) = 0$ $\forall x \in H$ then $T = 0$.*

Before the next proposition, a definition is required.

Definition 8.6. *Normal Operators.* If $T^*T = TT^*$ then operator T is said to be *normal.*

Proposition 8.7. *The operator $T \in L(H)$ is normal iff $\|Tx\| = \|T^*x\|$ $\forall x \in H$.*

Proof. Given T is normal. Then we have for $x \in H$ that

$$\begin{aligned} \|Tx\|^2 &= (Tx, Tx) \\ &= (T^*Tx, x) \\ &= (TT^*x, x) \\ &= (T^*x, T^*x) \\ &= \|T^*x\|^2, \end{aligned}$$

so $\|Tx\| = \|T^*x\|$.

Conversely when $\|Tx\| = \|T^*x\|$ $\forall x \in H$, it follows that

$$\begin{aligned} (T^*Tx, x) &= (Tx, Tx) \\ &= \|Tx\|^2 \\ &= \|T^*x\|^2 \\ &= (T^*x, T^*x) \\ &= (TT^*x, x), \end{aligned}$$

so $(T^*Tx, x) = (TT^*x, x)$. Working with this equality,

$$\begin{aligned} (T^*Tx, x) - (TT^*x, x) &= 0 \\ (T^*Tx - TT^*x, x) &= 0 \\ ((T^*T - TT^*)x, x) &= 0. \end{aligned}$$

Then let $S = T^*T - TT^*$. S is shown to be self-adjoint since

$$\begin{aligned} S^* &= (T^*T - TT^*)^* \\ &= (T^*T)^* - (TT^*)^* \\ &= T^*(T^*)^* - (T^*)^*T^* \\ &= T^*T - TT^* = S. \end{aligned}$$

Theorem 8.5 states that for self-adjoint S,

$$\forall x \in H,\ (Sx, x) = 0 \Rightarrow S = 0.$$

Then $T^*T = TT^*$ and T is normal as desired. $\square$

Chapter 9

The Polar Decomposition Theorem

The Polar Decomposition Theorem defines a method for finding two operators whose composition is equivalent to a given operator. This theorem is one of the last pieces to the puzzle needed to prove the main classes of operators in this thesis.

9.1 The Schur Representation Theorem

There is a technique to systematically find eigenvalues of T when T is a compact, self-adjoint operator. Theorem 9.1 will be the starting point for finding these eigenvalues.

Theorem 9.1. *If $T \neq 0$ is both a compact and self-adjoint operator, then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Given this eigenvalue, an algorithm finding eigenvalues of an operator can be described. This procedure is given on page 21 of R. Schatten's book "Norm Ideals of Completely Continuous Operators."[Sch60] The proof provided is quite laborious and only slightly instructive. An outline of the details will suffice.

Each dimension in space has an associated eigenvalue. Start with the eigenvalue given by Theorem 9.1. The operator T is then restricted to the space orthogonal to the eigenvector associated to that eigenvalue. This effectively reduces the dimension of the space by one. The restricted operator is also compact and self-adjoint. If it is also non-zero, then use of Theorem 9.1 will give a new eigenvalue and a new associated eigenvector.

The restricted operator maybe zero, which is how this procedure is completed in finite dimensional spaces. This algorithm is continued ad finitum if the original space is infinite dimensional. Beginning with the eigenvalue from the first iteration,

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$$

terminating in the finite dimensional case. Also, the sequence (λ_n) converges to zero by Theorem 6.7.

Theorem 9.2. *(Schur's Theorem.) Let H be a Hilbert Space and let $T \in L(H)$ be a compact and self-adjoint operator. For a sequence $\{\lambda_n\}$ of nonzero eigenvalues of T, the associated orthonormal sequence of eigenvectors (x_n) and $\forall x \in H$ the following equality holds:*

$$Tx = \sum_{n=1}^{\infty} \lambda_n (x, x_n) x_n, \quad x \in H.$$

This is the Schur Representation of the operator T. Also $\sigma(T) = \{\lambda_n\} \cup \{0\}$.

9.2 Square Roots

Positive self-adjoint operators have enough structure that they have many analogies to real numbers. There is a partial ordering of these operators (as was discussed in Proposition 8.3), similar to the ordering of the real numbers. Another analogy to real numbers is the existence of a unique square root. Both existence and uniqueness must be proven and proof is available in a variety of texts on Hilbert Space. The consequences of this square root are numerous and useful in the remainder of this work.

Definition 9.3. Let $T \in L(H)$ be positive and self-adjoint operator. If there is a self-adjoint $A \in L(H)$ such that $A^2 = T$ then A is called a *square root* of T. If A is also positive then A is called the (meaning unique) square root of T and denoted $A = T^{1/2}$.

Proposition 9.4. *For a positive self-adjoint operator $T \in K(H)$ there is a positive operator A with $A^2 = T$. A is the unique square root of T.*

Proof. The Schur Representation of positive self-adjoint $T \in K(H)$ is

$$Tx = \sum_{n=1}^{\infty} \lambda_n (x, x_n) x_n.$$

(Schur's Theorem requires that $T \neq 0$. If $T = 0$ the theorem, and hence the above representation, does not apply. However, in this case $A = 0$ fills all the requirements to be the unique positive square root.) Theorem 8.4 states that $\{\lambda_i\}$ are real numbers (which are known to have unique square roots). Define $\mu_i = \lambda_i^{1/2}$.

There exists a positive self-adjoint operator A such that

$$Ax = \sum_{n=1}^{\infty} \mu_n (x, x_n) x_n, \quad x \in H.$$

Then A is self-adjoint and positive.

It follows that A is unique by the uniqueness of $\lambda_i^{1/2}$. To show $A^2 = T$ a few calculations are necessary. Start with the Schur Representation of A and apply the operator A again. So, for $x \in H$ it holds that

$$\begin{aligned} A^2 x = AAx &= \sum_{m=1}^{\infty} \mu_m \left(\sum_{n=1}^{\infty} \mu_n (x, x_n)(x_n, x_m) \right) x_m \\ &= \sum_{n=1}^{\infty} \mu_n^2 (x, x_n) x_n \\ &= \sum_{n=1}^{\infty} \lambda_n (x, x_n) x_n \\ &= Tx. \end{aligned}$$

Leaving the operator $A^2 = T$. □

9.3 Polar Decomposition

There are two more details that need attention before the main topic of this chapter. The polar decomposition, defined below, provides a connection between the operators T, T^*T, and its square root A. Note: Proposition 8.3 states $T^*T \geq 0$. The operator T^*T then has a unique positive square root.

Theorem 9.5. *(The Polar Decomposition Theorem.) Let $T \in L(H)$ and $A = (T^*T)^{1/2}$. There is an operator $U \in L(H)$ such that*

1. *$T = UA$;*

2. *$\|Ux\| = \|x\|$ for $x \in \overline{R(A)}$;*

3. $Ux = 0$ *for* $x \in \overline{R(A)}^{\perp}$.

Proof. Since $A^2 = T^*T = A^*A$, for all $x \in H$ $\|Ax\|^2 = (Ax, Ax) = (A^*Ax, x) = (T^*Tx, x) = \|Tx\|^2$, and $\|Ax\| = \|Tx\|$. If we let $U : R(A) \to R(T)$ be given by $U(Ax) = Tx$ for $x \in H$, then U is a well-defined bounded linear operator on $R(A)$ and $\|Ux\| = \|Tx\|$ for $x \in R(A)$. Hence, U can be extended onto $\overline{R(A)}$ and, moreover, we can define $U = 0$ on $\overline{R(A)}^{\perp}$. Then, $U \in L(H)$ and it satisfies all the above statements.

□

This operator U has the added property of taking the set of eigenvectors of A (which are orthonormal) to another set which is orthonormal as well.

Theorem 9.6. *Let $T \in K(H)$ and let U be the operator defined by Theorem 9.5. Let (x_n) be the orthonormal eigenvectors of $A = (T^*T)^{1/2}$. Then the set (Ux_n) is also orthonormal.*

Proof. Given (x_n) is orthonormal, the second part of Theorem 9.5 has $\|Ux_n\| = \|x_n\| = 1$ for all n. Then for $n \neq m$,

$$\|Ux_n - Ux_m\|^2 = \|U(x_n - x_m)\|^2 = \|(x_n - x_m)\|^2 = 2.$$

But

$$\begin{aligned} \|Ux_n - Ux_m\|^2 &= (Ux_n - Ux_m, Ux_n - Ux_m) \\ &= (Ux_n, Ux_n) - (Ux_n, Ux_m) - (Ux_m, Ux_n) + (Ux_m, Ux_m) \\ &= 1 - (Ux_n, Ux_m) - (Ux_m, Ux_n) + 1. \end{aligned}$$

Combining the two equations, $-(Ux_n, Ux_m) = \overline{(Ux_n, Ux_m)}$. This implies that the real part of (Ux_n, Ux_m) must be zero.

A similar calculation for $x_n + ix_m$ results in $\|Ux_n + Uix_m\|^2 = 2$ and thus $(Ux_n, Ux_m) = \overline{(Ux_n, Ux_m)}$. So the imaginary part is also zero. Therefore $(Ux_n, Ux_m) = 0$ for all $n \neq m$ and the set (Ux_n) is orthonormal. □

Chapter 10

The Schmidt Representation

There has now been enough background in operators to move onto the major topics of this paper. Many of the previous topics will be refined and combined into a much more focused underlying theory.

Schur's Representation is helpful in calculations involving compact, self-adjoint operators. A similar result would be useful for operators that don't meet the strict restriction of being self-adjoint. Combining topics from the previous chapters gives such a result.

To ease notation, a brief lemma.

Lemma 10.1. *Given $T \in K(H)$, let $A = (T^*T)^{1/2}$. The eigenvalues $\sigma_n(T)$ of A found by the "procedure" – outlined on pages 31,32 – given by*

$$\sigma_n(T) = \inf\{\|T - B\| : B \text{ is a finite rank operator with rank } B < n\}.$$

These $\sigma_n(T)$ are also called the singular numbers of T. Some examples using singular numbers will follow shortly.

Finally, the Schmidt Representation Theorem, due to the German mathematician Erhard Schmidt.

Theorem 10.2. *(The Schmidt Representation Theorem.) Let $T \in K(H)$ and $A = (T^*T)^{1/2}$. Also let $(\sigma_n(T))$ be the eigenvalues of A and (x_n) the associated orthonormal eigenvectors. Then there is an orthonormal set (y_n) such that*

$$Tx = \sum_{n=1}^{\infty} \sigma_n(T)(x, x_n)y_n, \quad x \in H.$$

Proof. For $T \in K(H)$, A is compact and self-adjoint. A then has a Schur Representation

$$Ax = \sum_{n=1}^{\infty} \sigma_n(T)(x, x_n)x_n$$

for the eigenvalues, $\sigma_n(T)$, of A and associated eigenvectors (x_n). Then Theorem 9.6 says that $y_n = Ux_n$ is an orthonormal set. Thus using the polar decomposition of the operator T leaves

$$Tx = UAx = U\sum_{n=1}^{\infty} \sigma_n(T)(x, x_n)x_n = \sum_{n=1}^{\infty} \sigma_n(T)(x, x_n)y_n$$

as desired. □

Calculations using the Schmidt Representation deal with infinite sums and inner products which are relatively easy to work with. It is the singular numbers $\sigma_n(T)$ which may take some work to gain intuition of their properties. Some examples will assist in that intuition.

Example 10.3. *Show* $\|T\| = \sigma_1(T) \geq \sigma_2(T) \geq \ldots$ *for* $T \in K(H)$.

Proof. The definition of $\sigma_n(T)$ states that $\sigma_1(T) = \inf\{\|T\|\}$. (Note that rank $A < 1$ implies that $A = 0$.) So $\sigma_1(T) = \|T\|$.

Since $\sigma_n(T) = \inf\{\|T - A\| :$ rank $A < n\}$, $\sigma_{n+1}(T) = \inf\{\|T - A\| :$ rank $A < n + 1\}$, and

$\{A : A$ is of finite rank with rank $A < n+1\} \subset \{A : A$ is of finite rank with rank $A < n\}$

we see that $\sigma_n(T) \geq \sigma_{n+1}(T)$.

□

Next, an example demonstrating how to break up, or combine, the eigenvalues of two operators.

Example 10.4. *Show for* $S, T \in K(H)$, $\sigma_{m+n-1}(S + T) \leq \sigma_n(S) + \sigma_m(T)$.

Proof. By definition, $\sigma_{m+n-1}(S + T) = \inf\{\|S + T - A\| :$ rank $A < m + n - 1\}$. Then for $A \in L(H)$ with rank $A < m+n-1$, $\exists B, C \in L(H)$ such that $A = B+C$, rank $B \leq n$, rank $C < m$ and $B \cap$ range $C = \{0\}$. (The only element of both ranges is the element zero.) So rank $A =$ rank $(B + C)$.

Starting with one side

$$\begin{aligned}
\sigma_{m+n-1}(S+T) &= \inf\{\|S+T-A\| : \text{ rank } A < m+n-1\} \\
&= \inf\{\|S+T-B-C\| : \text{ rank } (B+C) < m+n-1\} \\
&\leq \inf\{\|S-B\| + \|T-C\| : \text{ rank } (B+C) < m+n-1\} \\
&\leq \inf\{\|S-B\| \text{ rank } B \leq n-1\} + \inf\{\|T-C\| \text{ rank } C < m\} \\
&\leq \inf\{\|S-B\| \text{ rank } B < n\} + \inf\{\|T-C\| \text{ rank } C < m\} \\
&= \sigma_n(S) + \sigma_m(T).
\end{aligned}$$

So $\sigma_{m+n-1}(S+T) \leq \sigma_n(S) + \sigma_m(T)$ as desired. □

Lastly, an example establishing a rule for the norm multiple operators and their corresponding eigenvalues.

Example 10.5. *Show* $\|R\| \cdot \|T\| \sigma_n(S) \geq \sigma_n(RST)$ *for operators* $R, S, T \in K(H)$.

Proof. Starting with the left hand side

$$\begin{aligned}
\|R\| \cdot \|T\| \sigma_n(S) &= \|R\| \sigma_n(S) \|T\|, \\
&= \|R\| \cdot \inf\{\|S-A\| : \text{ rank } A < n\} \cdot \|T\|, \\
&= \inf\{\|R\| \|S-A\| \|T\| : \text{ rank } A < n\}, \\
&\geq \inf\{\|R(S-A)T\| : \text{ rank } A < n\}, \\
&= \inf\{\|RST - RAT\| : \text{ rank } A < n\} \\
&\geq \inf\{\|RST - B\| : \text{ rank } B < n\} = \sigma_n(RST)
\end{aligned}$$

since rank $A < n$ implies rank $RAT < n$. □

Chapter 11

The Schatten Class Operators

This is where the work in the previous chapters comes together. This culmination begins with the work done by Robert Schatten. The Schatten p-class operators or S_p operators are defined below.

Definition 11.1. The *Schatten p-class operators* on Hilbert space H are given by

$$S_p(H) = \{T \in K(H) | (\sigma_n(T)) \in \ell_p\}, \ 1 \leq p < \infty.$$

This definition states that a compact operator T is in $S_p(H)$ if the sequence of eigenvalues of $[T^*T]^{1/2}$ are p-summable. Soon p will be fixed for certain situations, but for the following examples p will be unspecified.

Example 11.2. *For Hilbert space H, if $T \in S_p(H)$ then T can be factored into $T = BDA$ as follows:*

$$\begin{tikzcd} H \arrow[r, "T"] \arrow[d, "A"'] & H \\ \ell_\infty \arrow[r, "D"'] & \ell_p \arrow[u, "B"'] \end{tikzcd}$$

where $D\big((a_n)\big) = \big(\sigma_n(T)a_n\big)$ for all $a_n \in \ell_\infty$.

Proof. Let $x \in H$. Then x can be expanded and written as

$$x = \sum_{n=1}^{\infty}(x, x_n)x_n$$

for the complete orthonormal basis x_n created in Schur's Theorem, Theorem 9.2. Let $Ax = \big((x, x_n)\big) = (a_n)$. Then $(a_n) \in \ell_\infty$, because $\sum_{n=1}^{\infty} |a_n|^2 < \infty$. Let $D\big((a_n)\big) = \big(\sigma_n(T)a_n\big)$. But is $\big(\sigma_n(T)a_n\big) \in \ell_p$? Well

$$\sum_{n=1}^{\infty} |\sigma_n(T)a_n|^p \leq \sum_{n=1}^{\infty} \|(a_n)\|_\infty^p \cdot |\sigma_n(T)|^p = \|(a_n)\|_\infty^p \cdot \sum_{n=1}^{\infty} |\sigma_n(T)|^p < \infty.$$

So $\big(\sigma_n(T)a_n\big) \in \ell_p$.

Now let $B\big(\sigma_n(T)a_n\big) = \sum_{n=1}^{\infty} \sigma_n(T)(a_n)y_n$, where $T(x_n) = y_n$ for $n \geq 1$ and (y_n) is the orthonormal set found by using Theorem 9.6, on the orthonormal set (x_n). Finally,

$$\begin{aligned} BDA(x) &= BDA\sum_{n=1}^{\infty}(x, x_n)x_n, \\ &= BD\sum_{n=1}^{\infty}(a_n)x_n, \\ &= B\sum_{n=1}^{\infty}\big(\sigma_n(T)a_n\big)x_n, \\ &= \sum_{n=1}^{\infty}\sigma_n(T)(a_n)y_n, \\ &= \sum_{n=1}^{\infty}\sigma_n(T)(x, x_n)y_n, \\ &= Tx. \end{aligned}$$

The last equality being the Schmidt Representation of Tx. □

As with any object defined so far, the Schatten class operators must have a norm associated with it.

Definition 11.3. Given Hilbert space H and $T \in S_p(H)$ the *Schatten p-norm*, written $s_p(H)$ is

$$s_p(T) = \left(\sum_{n=1}^{\infty} \sigma_n(T)^p\right)^{1/p}.$$

Theorem 11.4. *Let $T \in K(H)$. If for all $U, V \in L(\ell_2, H)$ $\big((TUe_n, Ve_n)\big) \in \ell_p$, then $T \in S_p(H)$*

Proof. Given $T \in K(H)$ so T has Schmidt representation

$$Tx = \sum_{n=1}^{\infty} \sigma_n(T)(x, x_n)y_n$$

for some orthonormal sets $(x_n), (y_n)$. Let U, V be operators which change the orthonormal set (e_i) as follows:

$$Ue_i = x_i \text{ and } Ve_i = y_i, \quad i \geq 1.$$

The statement of the theorem says that for every, and as such this particular, U, V that $\big((TUe_n, Ve_n)\big) \in \ell_p$. So $\big((Tx_i, y_i)\big) \in \ell_p$. On the other hand,

$$\begin{aligned} (Tx_i, y_i) &= \left(\sum_{n=1}^{\infty} \sigma_n(T)(x_i, x_n)y_n, y_i\right) \\ &= (\sigma_i(T)y_i, y_i) \\ &= \sigma_i(T). \end{aligned} \tag{11.1}$$

Therefore $\big(\sigma_i(T)\big) \in \ell_p$ and $T \in S_p(H)$ as desired. $\square$

Chapter 12

Trace Class and Hilbert-Schmidt Operators

To complete this paper two final classes of operators will be discussed. Both of these operators predate the Schatten p-class operators. It will be shown that the trace class and Hilbert-Schmidt class operators are actually specific types of the Schatten p-class operators.

Definition 12.1. An operator $T \in L(H)$ is in the *trace class*, denoted $T \in TC(H)$, if

$$Tx = \sum_{n=1}^{\infty} (x, y_n) x_n$$

where $\sum \|x_n\| \|y_n\| < +\infty$. The norm of this operator is defined by

$$\tau(T) = \inf \left\{ \sum \|y_n\| \|x_n\| \right\}$$

with the infimum taken over all possible representations of $T \in TC(H)$.

Theorem 12.2. *The trace class is identical with the Schatten* 1*-class. That is,*

$$TC(H) = S_1(H)$$

and

$$s_1(T) = \tau(T) \quad \textit{for} \quad T \in TC(H).$$

Proof. If $T \in TC(H)$, then $T \in K(H)$ and T has a Schmidt representation. Let $\epsilon > 0$ be given. Since $T \in TC(H)$ there is some representation of T such that for all $x \in H$ there is a (w_n) and (z_n) where

$$Tx = \sum_{n=1}^{\infty}(x, w_n)z_n \quad \text{and} \quad \sum_{n=1}^{\infty} \|w_n\|\|z_n\| < (1+\epsilon)\tau(T).$$

This gives a representation that admits $\sum_{n=1}^{\infty} \|w_n\|\|z_n\|$ arbitrarily close to – but larger than – $\tau(T)$. For T to be in $S_1(H)$, $\sum \sigma_n(T)$ must be finite. This can be verified by finding an upper bound for $\sum \sigma_n(T)$.

Let (x_n), (y_n) be orthonormal sets. Then, beginning with a calculation similar to (11.1) in Example 11.4,

$$\begin{aligned}
\sum_{n=1}^{\infty} \sigma_n(T) &= \sum_{n=1}^{\infty}(Tx_n, y_n) \\
&= \sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty}(x_n, w_m)z_m, y_n\right) \\
&= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(x_n, w_m)(z_m, y_n) \\
&\leq \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}|(x_n, w_m)|^2\right)^{1/2}\left(\sum_{n=1}^{\infty}|(z_m, y_n)|^2\right)^{1/2} \\
&\leq \sum_{m=1}^{\infty}\|w_m\|\|z_m\| \\
&< (1+\epsilon)\tau(T) < \infty.
\end{aligned}$$

So $TC(H) \subset S_1(H)$ and $s_1(T) \leq \tau(T)$.

Then let $T \in S_1(H)$. So T has a Schmidt representation,

$$Tx = \sum_{n=1}^{\infty}\sigma_n(T)(x, z_n)x_n.$$

Letting $\sigma_n(T)z_n = y_n$,

$$Tx = \sum_{n=1}^{\infty}(x, y_n)x_n$$

which means T is also in $TC(H)$. And $\sum_{n=1}^{\infty}\|y_n\|\|x_n\| = \sum_{n=1}^{\infty}\sigma_n(T)$. So $\tau(T) \leq s_1(T)$. Therefore $S_1(H)$ is identical to $TC(H)$. $\square$

Definition 12.3. An operator $T \in L(H)$, is in the *Hilbert-Schmidt class*, denoted $T \in HS(H)$, if there is a complete orthonormal set (y_n) for H with

$$\sum_{n=1}^{\infty} \|Ty_n\|^2 < +\infty,$$

where the *Hilbert-Schmidt norm* of T operator is defined by,

$$hs(T) = \left(\sum_{n=1}^{\infty} \|Ty_n\|^2\right)^{1/2}$$

for any complete orthonormal set (y_n).

It should be noted that $hs(T)$ is independent of the choice of the complete orthonormal set of (y_n).

The Hilbert-Schmidt operators were studied at length before Schatten began work with the Schatten p-classes. It would be nice to bring the solutions found using Hilbert-Schimdt operators into the framework of the Schatten p-class operators. The following theorem will classify this connection.

Theorem 12.4. *The Hilbert-Schmidt class, $HS(H)$, is identical with the Schatten 2-class, $S_2(H)$. Their norms also coincide.*

Proof. Let $T \in HS(H)$ and let (x_n) be a complete orthonormal set in H. Then for $\epsilon > 0$ there is an N such that

$$\left(\sum_{n=N+1}^{\infty} \|Tx_n\|^2\right)^{1/2} < \epsilon.$$

Define the orthogonal projection $Px = \sum_{i=1}^{N}(x, x_i)x_i$. Thus for $x \in H$,

$$\begin{aligned}
\|(T - TP)x\| &= \left\|\sum_{n=N+1}^{\infty} (x, x_i)Tx_i\right\| \\
&\leq \left(\sum_{n=N+1}^{\infty} |(x, x_i)|^2\right)^{1/2} \left(\sum_{n=N+1}^{\infty} \|Tx_i\|^2\right)^{1/2} \\
&< \|x\|\epsilon.
\end{aligned}$$

This implies that $\|T - TP\| < \epsilon$ hence $T \in K(H)$. So T has a Schmidt representation

$$Tx = \sum_{k=1}^{\infty} \sigma_k(T)(x, u_k)v_k$$

for orthonormal sets (u_k) and (v_k). Using the representation of T,

$$\begin{aligned}\sum_{n=1}^{\infty}\|Tx_n\|^2 &= \sum_{n=1}^{\infty}(Tx_n, Tx_n)\\ &= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\sigma_k(T)^2|(x_n, u_k)|^2\\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\sigma_k(T)^2|(x_n, u_k)|^2\\ &= \sum_{k=1}^{\infty}\sigma_k(T)^2.\end{aligned}$$

By definition $T \in S_2(H)$ and $hs(T) = s_2(T)$.

Assuming $T \in S_2(H)$, the above proof works in reverse to show $T \in HS(T)$. So these two classes are identical. □

Definition 12.5. For a Banach space X, a sequence $(y_n) \subset X$ is said to be *weakly square summable* provided

$$\sum_{n=1}^{\infty}|f(y_n)|^2 < +\infty$$

for all $f \in X^*$ and square summable if

$$\sum_{n=1}^{\infty}\|y_n\|^2 < +\infty.$$

Definition 12.6. Let X and Y be Banach spaces. An operator $T \in L(X,Y)$ is called *absolutely* 2*-summing* if T maps weakly square summable sequences in X into square summable sequences in Y.

The set of absolutely 2-summing operators is denoted $\prod_2(X,Y)$. Any operator T in this set is called a $\prod_2$ operator.

Once again, the new operator should be put into context with respect to S_p.

Theorem 12.7. *The Schatten* 2*-class operators are identical to the absolutely* 2*-summing operators.*

Proof. As a reminder, $S_2 = \{T \in K(H) | (\sigma_n(T)) \in \ell_2\}$ and

$$\prod_2(H) = \left\{T \in L(H) \middle| \sum_{n=1}^{\infty}|(x, y_n)|^2 < +\infty \;\; \forall x \in H \Rightarrow \sum_{n=1}^{\infty}\|Ty_n\|^2 < +\infty\right\}.$$

The proof is by double inclusion. First, it will be shown that $\prod_2 \subset S_2$. Note: Theorem 12.4 states $S_2(H)$ is equivalent to $HS(H)$.

Let $T \in \prod_2(H)$ and let $(y_n) \subset H$ be a complete orthonormal set. Then (y_n) is a weakly square summable sequence. Then $T \in \prod_2(H)$ implies that (Ty_n) is square summable, ie. $\sum \|Ty_n\|^2 < \infty$. Thus $T \in HS(H)$ and therefore $T \in S_2(H)$. So $\prod_2 \subset S_2$.

Next choose $T \in S_2$ and let (y_n) be weakly square summable. Theorem 11.4 states that $T \in K(H)$ implies $\forall U, V \in L(\ell_2, H), \big((TUe_n, Ve_n)\big) \in \ell_2$. This is true for all U, V, so in particular, the U where $Ue_n = y_n$ and V where $V = TU$. Then $\big((Ty_n, Ty_n)\big) \in \ell_2$ and $\big(\|Ty_n\|^2\big) \in \ell_2$, which implies

$$\sum_{n=1}^{\infty} \|Ty_n\|^2 < +\infty.$$

Thus $T \in \prod_2$. Therefore $\prod_2(H) = S_2(H)$. $\square$

Chapter 13

Conclusion

In an effort to extend information from $\mathbb{R}^n$ to an infinite dimensional space many abstractions have been explored. Among those abstractions were the familiar notions of length, dot product, and eigenvalues which were generalized to norm, inner product, and the spectrum respectively. An appropriate framework for considering these abstractions was given along with meaningful examples.

The push for an appropriate analogue of the trace for an operator in a Hilbert space has led to the classification of compact operators, self-adjoint operators, the square root of operators and finally the trace class operators. In fact, the trace class operator was found to be just one type of a wide range of interesting operators, the Schatten p-class operators.

Appendix A

Banach Space

A complete normed linear space is called a Banach space. This space was touched on very briefly in Section 2.1. While the focus of this paper is on Hilbert space, there are undeniable links between these two types of spaces. One such link is studied when dealing with the Schatten p-class operators, so a short primer on Banach spaces is warranted [Roy88].

The Banach spaces of interest are called ℓ_p spaces. The familiar space ℓ_2 is one of these spaces. (A Hilbert space is after all a Banach space with an inner product.) The ℓ_p spaces are defined

$$\ell_p = \left\{ (a_n) \middle| \sum_{n=1}^{\infty} |a_n|^p < +\infty \right\}.$$

Again, a norm is required and it is defined to be

$$\|(a_n)\|_p = \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}.$$

This space is also defined for $p = \infty$. In that case,

$$\|(a_n)\|_\infty = \sup_{n \geq 1} |a_n|.$$

There are some fundamental theorems in Banach space theory, such as Hölder's Inequality and Minkowski's Inequality, which are given here in a form conducive for work in Hilbert space. These inequalities are focused on Banach space which is a secondary goal. Consequently, only brief outlines of the proofs are given. Complete proofs are readily available.

Theorem A.1. *(Hölder's Inequality). Let $a_1, \ldots, a_n, b_1, \ldots, b_n \subset \mathbb{C}$ and let $p,\ q > 0$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For $p > 1$ it holds that*

$$\sum_{i=1}^{n} |a_i b_i| \leq \left[\sum_{i=1}^{n} |a_i|^p\right]^{\frac{1}{p}} \left[\sum_{i=1}^{n} |b_i|^q\right]^{\frac{1}{q}}.$$

This inequality is also obtained for $p = 1$:

$$\sum_{i=1}^{n} |a_i b_i| \leq \left[\sum_{i=1}^{n} |a_i|\right] max\{|b_i| : 1 \leq i \leq n\}$$

Proof. The proof of Hölder's Inequality comes easily using Young's Inequality, $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$, and a clever substitution for a and b. □

Minkowski's Inequality provides the triangle inequality for ℓ_p spaces.

Theorem A.2. *(Minkowski's Inequality). Let $a_1, \ldots, a_n, b_1, \ldots, b_n \subset \mathbb{C}$ and let $p,\ q > 0$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For $p \geq 1$,*

$$\left[\sum_{i=1}^{n} |a_i + b_i|^p\right]^{\frac{1}{p}} \leq \left[\sum_{i=1}^{n} |a_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{n} |b_i|^q\right]^{\frac{1}{q}}.$$

Rewriting using the definition of the norm,

$$\|a + b\|_p \leq \|a\|_p + \|b\|_p.$$

Proof. The proof follows from Hölder's Inequality by factoring out a term of $|a + b|$ from the sum notation. Note $p = p/q + 1$. □

In Section 2.1 ℓ_2 was not shown to be complete. Instead, it will be shown here that ℓ_p is complete.

Proposition A.3. *The space ℓ_p is complete for $1 \leq p < \infty$.*

Proof. First, as a matter of notation, build a sequence of sequences in the following manner. Let $a^{(i)} \in \ell_p$ and $\{a^{(i)}\}_{i=1}^{\infty}$ be a Cauchy sequence $a^{(i)} = (a_k^{(i)})_{k=1}^{\infty}$. Given the sequence is Cauchy says $\forall \epsilon > 0, \exists N$ such that $i, j > N \Rightarrow \|a^{(i)} - a^{(j)}\|_p < \epsilon$ and by definition,

$$\left(\sum_{k=1}^{\infty} \left|a_k^{(i)} - a_k^{(j)}\right|^p\right)^{1/p} < \epsilon.$$

For any k we have

$$\left|a_k^{(i)} - a_k^{(j)}\right|^p \leq \sum_{k=o}^{\infty} \left|a_k^{(i)} - a_k^{(j)}\right|^p < \epsilon^p.$$

For this to be true for all ϵ, $\left|a_k^{(i)} - a_k^{(j)}\right|$ must approach zero as $i, j \to \infty$. Thus $(a_k^{(i)})$ is Cauchy in $\mathbb{R}$. So $\lim_{i\to\infty} a_k^{(i)} = a_k$ exists. Let $a = (a_k)_{k=1}^{\infty}$. All that is left to show is $\|a^{(i)} - a\|_p \to 0$ and $a \in \ell_p$.

For $i, j > N$,

$$\left(\sum_{k=1}^{\infty} \left|a_k^{(i)} - a_k^{(j)}\right|^p\right)^{1/p} < \epsilon$$

and taking the limit as j goes to infinity

$$\lim_{j\to\infty} \left(\sum_{k=1}^{\infty} \left|a_k^{(i)} - a_k^{(j)}\right|^p\right)^{1/p} \leq \epsilon.$$

Taking the limit

$$\left(\sum_{k=1}^{\infty} \left|a_k^{(i)} - a_k\right|^p\right)^{1/p} \leq \epsilon$$

and $\|a^{(i)} - a\|_p < \epsilon$. Hence $a^{(i)} \to a$ as $i \to \infty$.

Finally, pick an i such that, using Minkowski's inequality, $\|a\|_p \leq \|a - a^{(i)}\|_p + \|a^{(i)}\|_p$. Since $\|a^{(i)} - a\|_p \leq \epsilon$, $\|a\|_p \leq \epsilon + \|a^{(i)}\|_p$ which is finite as it is in ℓ_p. Therefore $\|a\|_p < \infty$ and $a \in \ell_p$. $\square$

While all the work in this section holds for $p = \infty$, the above results are sufficient for present purposes.

Bibliography

[Ret93] J. R. Retherford. *Hilbert Space: Compact Operators and the Trace Theorem*, volume 27 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1993.

[Roy88] H. L. Royden. *Real Analysis*. Macmillan Publishing Company, New York, third edition, 1988.

[Sch60] Robert Schatten. *Norm Ideals of Completely Continuous Operators*. Ergebnisse der Mathematik und ihrer Grenzgebiete. N. F., Heft 27. Springer-Verlag, Berlin, 1960.