Module structure of a Hilbert space

Ralph Daniel Leon

Follow this and additional works at: https://scholarworks.lib.csusb.edu/etd-project

Recommended Citation

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.
MODULE STRUCTURE OF A HILBERT SPACE

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree

Masters of Arts
in
Mathematics

by
Ralph Daniel Leon

September 2003
MODULE STRUCTURE OF A HILBERT SPACE

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by

Ralph Daniel Leon

September 2003

Approved by:

Yuichiro Kakihara, Committee Chair, 7-23-03
Charles Stanton, Committee Member
Belisario Ventura, Committee Member

Peter Williams, Chair
Department of Mathematics, Chair

Terri Hallet
Graduate Coordinator
Department of Mathematics
This paper demonstrates the properties of a Hilbert structure. In order to have a Hilbert structure it is necessary to satisfy certain properties or axioms. The main body of the paper is centered on six questions that develop these ideas (Appendix A). This paper demonstrates the Hilbert structure on $\mathbb{C}^n$ ($n$-dimensional complex Euclidean space), the gramian (Hilbert $\mathbb{M}_p(\mathbb{C})$-module), and the reproducing kernel Hilbert $\mathbb{M}_p(\mathbb{C})$-module.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLE</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF SYMBOLS</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER ONE: PRELIMINARY</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Euclidean Space</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Module</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1 Ring</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 Field</td>
<td>7</td>
</tr>
<tr>
<td>1.2.3 Vector Space</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Map</td>
<td>12</td>
</tr>
<tr>
<td>1.4 Dual Space: Hom(V,W)</td>
<td>13</td>
</tr>
<tr>
<td>1.4.1 Homomorphism</td>
<td>13</td>
</tr>
<tr>
<td>1.4.2 Developing a Basis Structure</td>
<td>15</td>
</tr>
<tr>
<td>1.4.2a Notation</td>
<td>17</td>
</tr>
<tr>
<td>1.4.3 The Dimension Perspective</td>
<td>17</td>
</tr>
<tr>
<td>1.5 Linear Functional</td>
<td>18</td>
</tr>
<tr>
<td>1.5.1 Notation</td>
<td>19</td>
</tr>
<tr>
<td>1.5.2 The Notation X^f: (Vector Space X^f)</td>
<td>21</td>
</tr>
<tr>
<td>1.6 Bilinear Functional/The Inner Product</td>
<td>21</td>
</tr>
<tr>
<td>1.7 Normed Space</td>
<td>25</td>
</tr>
</tbody>
</table>
CHAPTER TWO: $\mathbb{C}^n$, THE N-DIMENSIONAL COMPLEX EUCLIDEAN SPACE

2.1 Question One ........................................ 29
  2.1.1 Inner Product/Pre-Hilbert Space ................. 30
    2.1.1a Positive Definite Property .................. 31
    2.1.1b Conjugate Symmetry Property ................. 32
    2.1.1c Linear Property ................................ 32
    2.1.1d Conjugate Symmetry Property (revisited) .... 33
  2.1.2 Hilbert Space ..................................... 35
    2.1.2a Completion ...................................... 37

CHAPTER THREE: INTRODUCTION: ORTHOGONALITY

3.1 Question Two .......................................... 39
  3.1.1 Part I ............................................. 40
  3.1.2 Part II ............................................ 42

CHAPTER FOUR: ORTHOGONALITY

4.1 Question Three ......................................... 43
  4.2 Properties of Orthogonality: Inner Product Perspective ................. 43
  4.3 Projection: An Orthogonal Application .............. 46
    4.3.1 Notation ......................................... 47
    4.3.2 Bessel's Inequality ............................. 48
  4.4 Gram-Schmidt Orthogonalization Process ............ 49
    4.4.1 Orthogonal Basis ................................ 50
4.5 Orthogonal Complement .................................. 51
  4.5.1 Notation: (Orthogonal Complement) ........ 52

CHAPTER FIVE: GRAMIAN, THE HILBERT M_p(C)MODULE

  5.1 Question Four ........................................... 53
    5.1.1 Module .............................................. 53
    5.1.2 Gramian ............................................. 53
    5.1.3 (Normal) Pre-Hilbert M_n(C)-Module .......... 54
      5.1.3a Positive Definite Property .......... 54
      5.1.3b Linear Property ......................... 55
      5.1.3c Adjoint (Symmetry Property) .... 57

  5.2 Trace ................................................ 58
    5.2.1 Definition: (Trace-Class) .............. 59

CHAPTER SIX: UNITARY SPACE .............................. 60

  6.1 Matrix Perspective ..................................... 61
    6.1.1 Linear Map ....................................... 61
    6.1.2 Vector Space: (M_n(C),+,·(C)) .......... 62
    6.1.3 Algebra:(M_n(C),+,·(C),*) ............. 62

  6.2 Unitary Matrix ........................................ 62

  6.3 Similarity Transformations ....................... 63

  6.4 Theorems Related by a Similarity Transformation ........ 64

  6.5 Group Relation ....................................... 66
    6.5.1 Group ............................................ 66

vi
LIST OF TABLES

1.1 : Operational Closure for Number Systems . . 2
1.2 : Euclidean vs. Affine Transformations . . 5
1.3 : Additive Properties for the Ring . . . 6
1.4 : Multiplicative Properties for a Ring . . 7
1.5 : Properties of a Field . . . . . . . 8
1.6 : Additive Structure (Abelian Group): (V, +) . 10
1.7 : Multiplicative (Scalar) Structure: (V, .) . 11
1.8 : Notation for a Module . . . . . . . 12
1.9 : Two Sets for a Homomorphism . . . . 14
1.10: Distributive Property . . . . . . . 18
1.11: Linear Functional . . . . . . . . . 20
1.12: Bilinear Functional (Bilinear Form) . . 23
1.13: Norm of a Vector in a Linear Space . . 27
2.1 : Ordered Quadruple (X,+,,,) . . . . 30
2.2 : Cauchy-Schwartz Inequality . . . . . 36
4.1 : Question Three . . . . . . . . . 43
6.1 : Motion from A to B, but Different Basis . 63
7.1 : Range Comparison of the Kernel . . . . 72
7.2: The Reproducing Property
LIST OF FIGURES

1.1 : Pathway to a Hilbert Space ............... 1
1.2 : Euclidean Geometries ⊆ Affine Geometries . 5
4.1 : Conjugate Vectors ................. 44
4.2 : Projection of Vector y onto Vector x . . 46
4.3 : Process for $z_j \perp z_k$, $k \neq j = 1, 2, \ldots, n$ . 49
LIST OF SYMBOLS

Abelian group with operation addition over $\mathbb{R}$ .... $(\mathbb{R},+)$

Conjugate of $x$ ........ $\overline{x}$

Contradiction ........ $\rightarrow\leftarrow$

Element $x$ of $V$ ........ $x \in V$

For all ........ $\forall$

Homomorphisms from $V$ into $W$ .... $\text{Hom}(V,W)$

If and only if ........ $\iff$

Implies ........ $\Rightarrow$

Inner product ........ $\langle , \rangle$

Inner product in some books .... $\{ | \}$

N-dimensional complex Euclidean space .... $\mathbb{C}^n$

Negation of $x$ ........ $-x$

Norm on a linear space ........ $\| \|$ 

Sequence of elements ........ $(x_k)_{1 \leq k \leq n}$

Set ........ $\{,\ldots,\}$

Set of all the linear functionals on $X$ .... $X^f$ or $\hat{X}$

Set of linear transformations from $V$ into $W$ over the ring $F$ .... $\mathcal{L}_F(V,W)$

Same as above if it is understood to be over $F$ ........ $\mathcal{L}(V,W)$
Summation sign \[= \alpha_1 + \alpha_2 + \cdots + \alpha_n\]
from \[k = 1\] to \[n\] \[\sum_{k=1}^{n} \alpha_k\]
There exists \[\exists\]
Therefore \[\therefore\]
Transpose of a row matrix equals a column matrix \[(x_1, x_2, \ldots, x_n)^T\]
Vector space \(V\) over some field \(F\), or a module \(V\) over some ring \(F\); a \(F\)-module \(V\) with Abelian group \((V, +)\) and is a function from \(F \times V\) onto \(V\) \[(V, +, .)\]
CHAPTER ONE
PRELIMINARY

Figure 1.1: Pathway to a Hilbert Space ([Shi] pg. 2).

Question one asks to show that $\mathbb{C}^n$, the $n$-dimensional complex Euclidean space, is a Hilbert space (see Appendix A for all the questions). The goal of this chapter is to introduce mathematical structure/scaffolding that will allow us to see how a Hilbert space is constructed. We must start somewhere, so let's start at the number systems.

The number systems we normally think of are natural number $\{\mathbb{N}\}$, whole numbers $\{\mathbb{W}\}$, integers $\{\mathbb{I} \text{ or } \mathbb{Z}\}$, real numbers $\{\mathbb{R}\}$ (i.e., rational $\{\mathbb{Q}\}$ and irrational numbers) and complex numbers $\{\mathbb{C}\}$. Each number system has their permissible operations that define the idea of closure. We will look at closure for the operations of addition,
subtraction multiplication and division. Define the set of naturals $N = \{1, 2, 3, \ldots\}$, then the operation $1 - 3 = -2$, but $-2$ is not in the set of the natural numbers. Therefore, the operation of subtraction is not closed with the set $N$ and we do not use it. The table below represents the operations permissible for their respective number systems.

**Table 1.1: Operational Closure for Number Systems.**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>Set</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Natural</td>
<td>${1, 2, 3, \ldots}$</td>
<td>$+, \times$</td>
</tr>
<tr>
<td>W</td>
<td>Whole</td>
<td>${0, 1, 2, 3, \ldots}$</td>
<td>$+, \times$</td>
</tr>
<tr>
<td>I</td>
<td>Integers</td>
<td>${\ldots, -2, -1, 0, 1, 2, \ldots}$</td>
<td>$+, -, \times$</td>
</tr>
<tr>
<td>R</td>
<td>Reals</td>
<td>Rational and irrational</td>
<td>$+, -, \times, \div$</td>
</tr>
<tr>
<td>C</td>
<td>Complex</td>
<td>$a+bi \ a, b \in R, i = \sqrt{-1}$</td>
<td>$+, -, \times, \div$</td>
</tr>
</tbody>
</table>

The table represents why when working with computers a variable defined as a natural number (2 operations) would take less memory than a number defined as a real number (4 operations), so distinction of the number systems are important.
You may have noticed we have not mentioned the word “length” even though many intuitively think \( 2 + 3 = 5 \). Length is a concept that is drilled into our minds during high school geometry (Euclidean geometry). The number systems above are just blank tools that may represent many things, therefore we call the numbers “abstract.” When we put the label “length” on a number we are structurally moving to another concept and subtle movements like “length” will bring us closer to the idea of a Hilbert space. The next section fine-tunes high school geometry.

1.1 Euclidean Space

Question one asks to show that a particular Euclidean space is a Hilbert space. So, we are given a Euclidean space; a Euclidean geometry. In Euclidean geometry or high school geometry segments have “lengths” and lines cross at “angles” one can measure. What does that mean? If, we compare Euclidean geometry with the affine geometry our Euclidean perspective becomes clearer.

Affine geometry speaks of ratio of lengths along parallel lines and all triangles are (affine) congruent,
but this means they do not retain the same size. Since ratio of lengths along parallel lines does not preserve length it does not make sense to speak of distance in affine geometry. In affine geometry transformations have the form

\[ t(x) = Ax + a \]

where A is an invertible matrix and \( a \in \mathbb{R}^2 \). However, Euclidean geometry has transformations that preserve length. These are transformations of the form

\[ t(x) = Ux + a \]

where U is an orthogonal matrix and \( a \in \mathbb{R}^2 \). Therefore, distance and orthogonal matrices are important characteristics of a Euclidean geometry.

In both geometries, the allowable transformations form a group (A group will be discussed in further detail later in Chapter 6, question 5). Moreover the group of Euclidean transformations is a subgroup of the group of affine transformations. Both geometries concern the same space and each geometry involve a group of transformations that acts on this space. The Euclidean group is a subgroup of the affine group. In the sense Euclidean geometry is a sub-geometry of affine geometry.
Figure 1.2: Euclidean Geometries ⊂ Affine Geometries.

Table 1.2: Euclidean vs. Affine Transformations

<table>
<thead>
<tr>
<th>Geometry</th>
<th>Space</th>
<th>Transformation</th>
<th>Invariants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affine</td>
<td>$\mathbb{R}^n$</td>
<td>$t:x \mapsto Ax + b$, where $A$ is a $n \times n$ invertible Matrix</td>
<td>Ratio along Parallel lines</td>
</tr>
<tr>
<td>Euclidean</td>
<td>$\mathbb{R}^n$</td>
<td>$t:x \mapsto Ux + b$, where $A$ is a $n \times n$ orthogonal Matrix</td>
<td>Length, angle</td>
</tr>
</tbody>
</table>

([BrEG] pg. 360).

Length and orthogonal matrices are main characteristics of interest. A module structure is another characteristic of concern.

1.2 Module

When analyzing a mathematical event we can study the functions that create the event or we can study the space created by the event. Presently, we are interested in a structure similar to a space called a module. Roughly speaking, a module is a vector space over a ring
instead of a field ([War] pg.253). So it is necessary to define a ring, field, and vector space so the concept of "module" can be understood.

1.2.1 Ring

A ring $R$ is an algebraic system that is split into two sets, an additive structure and a multiplicative structure. The algebraic structure consists of $x, y, z \in R$ on which two operations $+$ and $\cdot$ (addition and multiplication) are defined below.

Table 1.3: Additive Properties for the Ring.

<table>
<thead>
<tr>
<th>1) Closure:</th>
<th>If vectors $x, y \in R$, then $x+y \in R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2) Associative:</td>
<td>$(x + y) + z = x + (y + z)$</td>
</tr>
<tr>
<td>3) Additive Identity:</td>
<td>$x + 0 = 0 + x = x$</td>
</tr>
<tr>
<td>4) Inverse:</td>
<td>For each vector $x \in R$, $\exists (-x) \in R$ such that $x + (-x) = 0$</td>
</tr>
</tbody>
</table>

A group is represented above with the binary operation of addition. Sometimes $\ast, \cdot, \circ$, or other suitable symbol (besides $+$), or no symbol at all is used.

4a) Commutative: $x + y = y + x$
Including the commutative property represents an Abelian group \((R, +)\) under addition.

Table 1.4: Multiplicative Properties for a Ring.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Closure: ( \text{If } x, y \in R, \text{ then } xy \in R )</td>
</tr>
<tr>
<td>2</td>
<td>Associative: ( (xy)z = x(yz), \forall x, y, z \in R )</td>
</tr>
<tr>
<td>3</td>
<td>Left Distributive: ( x(y + z) = xy + xz, \forall x, y, z \in R )</td>
</tr>
<tr>
<td>4</td>
<td>Right Distributive: ( (y + z)x = yx + zx, \forall x, y, z \in R )</td>
</tr>
</tbody>
</table>

The above properties characterize a ring. Rings are patterned after and are generalizations of the algebraic aspect of ordinary integers ([Her] pg. 120, [Blo] pg. 249).

The particular ring defined next will be used to define a field. A ring \( R \) with identity is said to be a division ring if its nonzero elements \( R - \{0\} \) have the group characteristics under multiplication. That means every nonzero element has a multiplicative inverse ([Fan] pg. 66, 131, [Her] pg. 126, 171 [War] pg. 211, 252).

1.2.2 Field

Just like the ring \( R \), a field \( F \) is an algebraic system that is split into two sets, an additive structure
and a multiplicative structure. In fact, a ring is a field \( F \) if it forms an Abelian group under both its additive and multiplicative structure. That is, a set \( F \) is a field if it satisfies the following properties:

**Table 1.5: Properties of a Field.**

<table>
<thead>
<tr>
<th>No.</th>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>Ring:</td>
<td>Properties for a ring.</td>
</tr>
<tr>
<td></td>
<td>(If we add a few more properties, then we also have a group over multiplication.)</td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td>Identity:</td>
<td>( ex = xe = x, \forall x \in F )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( (1 \cdot x = x \cdot 1 = x) )</td>
</tr>
<tr>
<td>3)</td>
<td>Inverse:</td>
<td>( x \cdot x^{-1} = x^{-1} \cdot x = e, \exists! x^{-1} \forall x \in F )</td>
</tr>
<tr>
<td>4)</td>
<td>Commutative:</td>
<td>( x \cdot y = y \cdot x, \forall x, y \in F )</td>
</tr>
</tbody>
</table>

A field is a commutative ring in which we can divide by any nonzero element ([Her] pg. 207, [Fan] pg. 159). That is, a field is a commutative division ring. And every nonzero element of the division ring is cancelable for multiplication;

\[
ab = ac \\
\Rightarrow a^{-1}ab = a^{-1}ac \\
\Rightarrow b = c.
\]
Thus, a division ring contains no zero-divisor. A commutative ring with no zero-divisors is called an integral domain. In other words, a ring R is an integral domain providing that R is commutative, has unity, and has no divisors of zero ([Blo] pg. 258). Thus, a field is an integral domain ([War] pg. 211, [Her] pg. 126).

An easier way to see a full concept with just a few symbols is by using mathematical notation. Notation will be used to define a Hilbert space and is introduced below with a vector space.

1.2.3 Vector Space
A vector space V, over a field F or a F-vector space is an ordered triple \((V, +, \cdot)\) that is split into two sets.

Vector Space: \((V, +, \cdot)\)

- Additive Structure (Abelian group): \((V, +)\)
- Multiplicative Structure (scalar multiplication): \((V, \cdot)\)

The first four properties (axioms) are concerned with the additive structure of elements of V.
Table 1.6: Additive Structure (Abelian Group): $(V, +)$.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Closure:</td>
<td>$x, y, z \in V, \rightarrow x + y \in V$</td>
</tr>
<tr>
<td>2) Associative:</td>
<td>$(x + y) + z = x + (y + z)$</td>
</tr>
<tr>
<td>3) Null (zero) vector:</td>
<td>$0: x + 0 = x$</td>
</tr>
<tr>
<td>4) Inverse:</td>
<td>$\forall x \in V, \exists (-x) \in V$</td>
</tr>
<tr>
<td>(The Group Structure)</td>
<td>such that $x + (-x) = 0$</td>
</tr>
</tbody>
</table>

The abstract concept of a group has its origins in the set of mappings, or permutations, of sets onto itself.

4a) Commutative:

$x + y = y + x$

Including the commutative property shows that a vector space is the Abelian group $(V, +)$ under addition. Note, that the remaining axioms are concerned with the action of the scalars $F$ on $V$ (i.e., $:F x V \rightarrow V$).
Table 1.7: Multiplicative (Scalar) Structure: $(V, \cdot)$.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For any scalar $\alpha \in F$ and any vectors $x, y \in V$, $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$</td>
</tr>
<tr>
<td>2</td>
<td>For any scalar $\alpha, \beta \in F$ and any vector $x \in V$, $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$</td>
</tr>
<tr>
<td>3</td>
<td>For any scalar $\alpha, \beta \in F$ and any vector $x \in V$, $(\alpha \beta) \cdot x = \alpha \cdot (\beta x)$</td>
</tr>
<tr>
<td>4</td>
<td>For the unit scalar $1 \in F$ and any vector $x \in V$, $1 \cdot x = x$.</td>
</tr>
</tbody>
</table>

Since a vector space necessarily over a field has been defined, we can now define a module. We look at the additive structure for the ring, field or vector space and define a module. The Abelian group under addition is sometimes called a module. Note that a module has been defined at a low mathematical structural level. At a higher structure level we find that, roughly speaking, a module is a vector space over a ring instead of a field ([War] pg. 253).

Notation

Let $R$ be a ring then the module over $R$ (or an $R$-module) is:
Table 1.8: Notation for a Module.

<table>
<thead>
<tr>
<th>R-module</th>
<th>(ordered triple)</th>
<th>(V, +, ·)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive Structure:</td>
<td>(Abelian Group)</td>
<td>(V, +)</td>
</tr>
<tr>
<td>Multiplicative Structure:</td>
<td>(R x V onto V)</td>
<td>(V, ·)</td>
</tr>
</tbody>
</table>

([Her] pg. 171, [Fan] pg. 66, [War] pg. 252). If in addition R is a ring with identity and if 1·x = x for all x ∈ V, then (V, +, ·) is called a unitary R-module (since it contains the unit 1). Elements of R are called scalars and R is called a scalar ring ([War] pg. 253).

The next section defines a class of actions from one set to another. The names are abundant and have their own subtly. It may seem that we are peeling down too many layers of an onion just to answer the six questions of this thesis, but grouping these terms may actually show semantic clarity. Besides many of the proofs included in this thesis use these terms, therefore they should be defined.

1.3 Map

Getting from one set to another set involves a process of a map. Other words for map are: function,
functional, arrow, transformation, operator and morphism. A map $f$ involves three things:

1. a set $A$, called the domain (input) of the map $f$;
2. a set $B$, called the co-domain (output) of the set $f$;
3. And the rule (or process) for $f$, assigning each element of $A$ exactly one element of $B$

({LawS} pg. 14, 22).

1.4 Dual Space: $\text{Hom}(V,W)$

Recall, orthogonal matrices are functions in a Euclidean space that preserve length. We are going to be interested in the sets of functions that operate on vector spaces and how to classify those sets. Therefore, we need a larger vocabulary palette that continues below.

1.4.1 Homomorphism

A mapping $\phi$ from the ring $R$ into the ring $R'$ is said to be a homomorphism if
Table 1.9: Two Sets for a Homomorphism.

1). \(\varphi(a + b) = \varphi(a) + \varphi(b)\); That is, the image of the sum is equal to the sum of the images.

2). \(\varphi(ab) = \varphi(a)\varphi(b)\); That is, the image of the product is equal to the product of the images ([Her] pg. 131).

Let \(R\) be a ring. A homomorphism from a \(R\)-algebraic structure \((V, +, \cdot)\) with one composition \(\cdot\), into another \((W, +, \cdot)\) is a function \(f\) from \(V\) into \(W\). The function \(f\) is a homomorphism \((f \in \text{Hom}(V,W))\) from the algebraic structure \((V, +)\) into \((W, +)\) (additive structure) and satisfies \(f(\alpha \cdot x) = \alpha \cdot f(x)\) (multiplicatively structure) for all \(x \in V\) and all \(\alpha \in R\) ([War] pg. 274).

Similarly instead of ring \(R\), if we are given any two vector spaces \(V\) and \(W\) over a field \(F\), we defined \(\text{Hom}(V,W)\) to be the set of all vector space homomorphisms of \(V\) into \(W\) (i.e., \(\text{Hom}(\text{domain} \ V, \text{Range} \ W))\). The question that comes to mind is whether or not \(\text{Hom}(V,W)\) is a vector space over \(F\), that is \((\text{Hom}(V,W), +, \cdot)\). Do we have a vector space when our set of elements is functions that form homomorphisms from vector space \(V\) to \(W\)? The
answer is "yes", Hom(V,W) does form a homomorphism (see Appendix B for the proof).

Hom(V,W) is a vector space and the homomorphism f from one module V into another W, i.e. f ∈ Hom(V,W), is usually called a linear transformation. And a linear transformation from one module W into itself is called a linear operator on W ([War] pg. 275).

Also, if V is a vector space; (V,+,.) over F, then its dual space is Hom(V,F) ([Her] pg. 187). Note that the homomorphism goes from the vector space V to the field F.

The notation V* (sometimes Vf, V*, or Hom(V,F)) is used for the dual space of V.

Question two and three are about orthogonality and the next section has some of the tools that will be used in that section.

1.4.2 Developing a Basis Structure

Let (xk), 1 ≤ k ≤ n be a sequence of elements of a R-module V (or vector space). An element v of V is a linear combination of (xk), 1 ≤ k ≤ n if there exists a sequence of scalar elements αk, 1 ≤ k ≤ n of the ring R (or αk ∈ F; field), such that
$$v = \sum_{k=1}^{n} \alpha_k x_k = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

([War] pg. 261, [Her] pg. 177). Now, if \( W \) is a nonempty subset of the vector space \( V \), then \( L(W) \), the linear span of \( W \), is the set of all linear combinations of \( W \) ([Her] pg. 177). And a subset \( W \) of a vector space \( V \) is called a basis of \( V \) if \( W \) consists of linearly independent elements and \( V = L(W) \) ([Her] pg. 180).

Another way of defining a basis, that is an ordered basis of a unitary \( K \)-module \( X \), is a linearly independent sequence \( (x_j), 1 \leq j \leq n \) of elements of \( X \) such that \( \{x_1, x_2, \ldots, x_n\} \) is a set of generators for \( X \) ([War] pg. 263). Note that for a basis of \( n \) elements determine \( n! \) ordered bases. However, there is one standard basis that will be defined below.

Let \( R \) be a ring with identity, and let \( n > 0 \) be an integer, and for each \( k \in [1, n] \) let \( e_k \) be the ordered \( n \)-tuple of elements of \( R \) whose \( k^{th} \) entry is 1 and all of whose other entries are 0. Then \( (e_k), 1 \leq k \leq n \) is an ordered basis of the \( R \)-module \( R^n \) since

$$\sum_{j}^{n} \lambda_j e_j = (\lambda_1, 0, 0, \ldots, 0) + (0, \lambda_2, 0, \ldots, 0) + (0, 0, \ldots, \lambda_n)$$

$$= (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{n-1}, \lambda_n).$$
This ordered basis is called the standard ordered basis of $\mathbb{R}^n$, and the corresponding set \{e_1, e_2, \ldots, e_n\} is called the standard basis of $\mathbb{R}^n$ ([War] pg. 263).

1.4.2a Notation. A frequently used practice in mathematical notation is the Kronecker Delta; if $W$ is some (index) set and if $R$ is a ring with identity occurring in a given context, then for every $(j, k) \in W \times W$, $\delta_{jk}$ is defined by

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$$

where "1" and "0" denote respectively the multiplicative and additive neutral elements of $R$ ([War] pg. 281). However, suppose $\{v_j\}$ is the set of vector in $V$. We call this set an orthonormal set if

$$\langle x_j, x_k \rangle = \begin{cases} 1 & \text{if } j = k \text{ (length), i.e., } \langle x_j, x_j \rangle = 1 \\ 0 & \text{for } j \neq k, \text{ i.e., } \langle x_j, x_k \rangle = 0 \end{cases}$$

([Her] pg.196). And if $V$ is a $n$-dimensional inner product space (inner product space will be defined later), then the orthonormal set is a basis (i.e., orthonormal basis) ([Her] pg. 196).

1.4.3 The Dimension Perspective

The dimensional perspective becomes applicable with Chapter Six question five when we replace the geometry
of symmetry operations with the algebra of matrices. The matrices we will be interested are those of \( M_n(C) \). That is, \( n \times n \) matrices with elements from \( C \) (\( C \), complex numbers). This section presents a few theorems/corollaries with reference to dimension.

**Theorem:** If \( V \) and \( W \) are of dimension \( m \) and \( n \), respectively, over a field \( F \) (i.e., or \( V \) and \( W \) are \( F \)-vector spaces), then \( \text{Hom}(V,W) \) is of dimension \( mn \) over \( F \).

(see Appendix B for the proof)

**Corollary 1:** If \( \dim_F W = n \), then \( \dim_F \text{Hom}(W,W) = n^2 \).

Proof. In the theorem above put \( V = W \), so \( n = m \), hence \( nm = n^2 \) ([Her] pg. 186).

**Corollary 2:** If \( \dim_F V = n \) (over the field \( F \)), then \( \dim_F \text{Hom}(V,F) = n \) ([Her] pg. 186).

(See Appendix B for the Proof.)

1.5 Linear Functional

**Table 1.10:** Distributive Property.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A \cdot (B + C) )</td>
<td>( AB + AC )</td>
</tr>
<tr>
<td>( (B + C) \cdot A )</td>
<td>( BA + BC )</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Since defining the multiplicative structure for a ring, we are familiar with the concept of left-distributive and right-distributive properties, which in appearance are similar to the practice that makes use of the word "linear" when dealing with functions. Recall, the notation \( \hat{V} \) (sometimes \( V^f, V^*, \) or \( \text{Hom}(V,F) \)) is used for the dual space of \( V \). An element of \( \hat{V} \) will be called a linear functional on \( V \) into \( F \) ([Her] pg. 187). A linear functional may also be called a linear map, linear mapping, linear operator, or linear transformation depending on the situation. Therefore, the adjective linear will make all the functions have a distributive characteristic in appearance.

1.5.1 Notation

This section repeats the description for a vector space, yet it is done with new notation found in other books and it is done for a linear vector space. If \( f \) is evaluated at a point \( x \in X \), for \( x' = f \) we may write \( f(x) = x'(x) = x'x \). Frequently we will also find the notations \( f(x) \overset{\wedge}{=} (x,f) = (x,x') \). Letting \( x' = f \) will let us mentally transition from a vector space to a function
space. So, if $f$ is a linear functional then it follows that:

Table: 1.11: Linear Functional

<table>
<thead>
<tr>
<th></th>
<th>Additive Structure</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>$f(x_1 + x_2) = (x_1 + x_2, x') = (x_1, x') + (x_2, x')$</td>
<td>([Mich] pg.110)</td>
</tr>
<tr>
<td></td>
<td>$= f(x_1) + f(x_2)$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Multiplicative Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>2)</td>
<td>$f(\alpha x) = (\alpha x, x') = \alpha(x, x') = \alpha f(x)$.</td>
</tr>
</tbody>
</table>

Recall, when we talk of fields we have transformations between vector spaces, and when we talk of rings we have transformations between modules. Along this line, we denote the set of all linear transformations from $V$ into $W$ by $L_R(V,W)$ or just $L(V,W)$ if it is clear that the scalar ring is $R$ ([War] pg. 281). And for $V$, a module over a commutative ring $R$, a linear form on $V$ is a linear transformation $L_R(V,R)$ from $V$ into the $R$-module $R$. The $R$-module $L(V,R)$ of all linear forms on $V$ is usually denoted by $V^*$ and is called the
algebraic dual of $V$. Linear forms are also called linear functionals ([War] pg. 282).

1.5.2 The Notation $X^f$: (Vector Space $X^f$)

Theorem: Let $X$ be a linear space and let $X^f$ (or $\hat{X}$) denote the set of all linear functionals on $X$. The space $X^f$ with vector addition and multiplication of vectors by scalars is a vector space over $F,(X^f,+,\cdot)$ (see Appendix B for the proof).

Definition: The basis $\{e'_1, \ldots, e'_n\}$ of $X^f$ is called the dual basis of $\{e_1, \ldots, e_n\}$ ([MicH] pg. 112).

Definition: Now let $X$ be a linear space and let $X^f$ denote the set of all linear functionals on $X$. The linear space $X^f$ is called the algebraic conjugate of $X$ ([MicH] pg. 110).

1.6 Bilinear Functional/The Inner Product

A Hilbert space that has a function with two terms, the second term has a conjugate form. Below we start to introduces this concept, however, it has conjugate form on both terms. Let $X$ be a vector space over $\mathbb{C}$ (Complex), $(X,+,\cdot)$. A mapping $f: X \to \mathbb{C}$ is said to be a conjugate linear functional if
\[ f(\alpha x + \beta y) = \overline{\alpha} f(x) + \overline{\beta} f(y) \]

(Distributive form, but only with functions i.e., "linear")

for all \( x, y \in X \), and for \( \alpha, \beta \in \mathbb{C} \), where \( \overline{\alpha} \) denotes the complex conjugate of \( \alpha \) and \( \overline{\beta} \) denotes the complex conjugate of \( \beta \) ([Mich] pg. 114).

Remember that getting from one set to another set involves a process of a map. Next, we start to develop the structure for the inner product; we start with two vectors (i.e. the prefix "bi-"). Therefore, the next definition shows characteristics of the domain (input set) that makes the functional gain the grammatical adjectives 'bilinear' or 'conjugate linear'. That is, we have a distributive structure with the first term and a conjugate distributive structure with the second term. However, since we are working with functionals we do not use the word "distributive," we use the word "linear."

For example, let \( X \) be a vector space over \( \mathbb{C} \). A mapping \( f: X \times X \rightarrow \mathbb{C} \) is called a \textit{bilinear functional} or a \textit{bilinear form} if

22
Table 1.12: Bilinear Functional (Bilinear Form).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i) For each fixed ( y ), ( f(x, y) ) is a linear functional in ( x )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>( f(\alpha x_1 + \beta x_2, y) = \alpha f(x_1, y) + \beta f(x_2, y) ), and</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ii) For each fixed ( x ), ( f(x, y) ) is a conjugate linear</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>functional in ( y ) (( y ) is the second term)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>( f(x, \alpha y_1 + \beta y_2) = \overline{\alpha} f(x, y_1) + \overline{\beta} f(x, y_2) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>for all ( x_j, y_j \in X, j = 1, 2, ..., n )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>and for ( \alpha, \beta \in \mathbb{C} ) (complex ( \mathbb{C} )) ([MicH] pg. 114-115).</td>
<td></td>
</tr>
</tbody>
</table>

Eventually, we want to show that \( \mathbb{C}^n \) is a Hilbert space. So, we must consider the complex linear space. So let \( X \) be a complex linear space, \( (X, +, \cdot) \). For all \( x, y \in X \) a bilinear functional \( f \) is said to be

a) **Symmetric**; \( \text{if } f(x, y) = \overline{f(y, x)} \)

b) **Positive**; \( \text{if } f(x, x) \geq 0 \text{ for all } x \in X \),

c) **Strictly positive**; \( \text{if } f(x, x) > 0 \text{ for all } x \neq 0. \)

([MicH] pg.115)
Definition: Let $X$ be a complex vector space, and let $f$ be a bilinear functional. We call the function $\tilde{f}: X \to \mathbb{C}$ defined by $\tilde{f}(x) = f(x, x)$ for all $x \in X$, the quadratic form induced by $f$ (we frequently omit the phrase "induced by $f$") ([MicH] pg. 115).

Theorem: A bilinear function $f$ on a complex vector space $X$ is symmetric if and only if $\tilde{f}$ is real (i.e., $\tilde{f}(x)$ is real for all $x \in X$) (see Appendix B for the Proof).

The inner product is an important characteristic for a Hilbert space and is the subject of the first major question that initiated this paper. We say a strictly positive, symmetric bilinear functional $f$ on a complex linear space $X$ is called an inner product. The inner product of two vectors is not a vector. ([MicH] pg. 115-117, [Wre] pg. 57). From Question 1 the inner product is defined for:

$$x = (x_1, x_2, \ldots, x_n)^T, \; y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n,""T"" \text{ for transpose}$$

$$\langle x, y \rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k\overline{y_k} : \text{the inner product.}$$
1.7 Normed Space

Recall, that the concept length and transformation of orthogonal matrices were emphasized when we discussed a Euclidean geometry. However, when we discussed vector spaces the concept of length was not discussed. A norm in a linear space follows the concept of length and lets us consider the concept of convergence, which lets us consider completion. It will be necessary to show completion of a pre-Hilbert space to demonstrate a Hilbert space.

Question one defines a norm using the inner product:

\[
\begin{align*}
\langle x, y \rangle &= \sum_{k=1}^{n} x_k \overline{y_k} : \text{the inner product} \\
\|x\| &= \langle x, x \rangle^{\frac{1}{2}} : \text{the norm}
\end{align*}
\]

Since a complete normed linear space is a Banach space ([MicH] pg. 345), we can take the inner product perspective and discuss a Hilbert space and skip the analysis of Banach spaces. This can be done since we will be using the tool of the inner product in the Hilbert space to represent the norm. However, completion
is still necessary to show a Hilbert space so the concept of convergence must be considered. That is, a Hilbert space is a Banach space but not vice versa. So, we will not be able to dispense with the norm altogether.

Introducing a norm in a linear space is one of the methods for introducing convergence. The fundamental concepts we study in mathematical analysis are the various types of convergence and the limit. Derivatives, integrals, series expansion etc are based on these notions.

A norm of a vector in a linear space X is an abstraction of the absolute value of a vector in the geometric vector space. The absolute value (modulus) of a vector in the geometric vector space has the following properties. Denoting the absolute value of the vectors $x, y$ in the usual manner as $\| x \|, \| y \|$ we have:
Table 1.13 Norm of a Vector in a Linear Space

1) Additive Structure
(i) \( \| x \| \geq 0 \) and \( \| x \| = 0 \) if and only if \( x = 0 \) (Positive)
(ii) \( \| x + y \| \leq \| x \| + \| y \| \), \( x, y \in X \) (Δ Inequality)

2) Multiplicative Structure
(iii) \( \| \lambda x \| = |\lambda| \| x \| \), \( x \in X \), \( \lambda \in \mathbb{C} \) (complex scalars)

The linear space \( X \) is called a normed linear space if there exists a mapping \( x \rightarrow \| x \| \) from \( X \) into the set of non-negative numbers. The mapping is defined for every \( x \in X \) for which the properties (i) - (iii) are satisfied. The non-negative number \( \| x \| \) is called the norm of \( x \) ([Mat] pg. 4). A sequence \( \{x_n\} \) of vectors in a normed space \( X \) is called convergent if there exists \( x \in X \) such that the sequence \( \| x_n - x \| \rightarrow 0 \). In Notation this is written:

\[
\lim_{n \to \infty} x_n = x \text{ or } x_n \rightarrow x.
\]

In this case, \( x \) is called the limit of the convergent sequence \( \{x_n\} \) ([Mat] pg. 4).
This concludes the Preliminary chapter and the next chapter starts with the first of six questions that make up this thesis.
CHAPTER TWO

C^n, THE N-DIMENSIONAL COMPLEX
EUCLIDEAN SPACE

2.1 Question One

Show that C^n is a Hilbert space with inner product given below.

Let

\[
C^n = \{(x_1, x_2, \ldots, x_n)^T: x_k \in \mathbb{C}, 1 \leq k \leq n\}; \\
C^n = n - \text{dimensional complex Euclidean space}
\]

For \( x = (x_1, x_2, \ldots, x_n)^T \), \( y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n \)

(“T” for transpose)

\[
\begin{align*}
\langle x, y \rangle &= \sum_{k=1}^{n} x_k \overline{y}_k : \text{the inner product} \\
\| x \| &= \langle x, x \rangle^{1/2} : \text{the norm}
\end{align*}
\]

We say that C^n is a Hilbert space, if C^n is a pre-Hilbert space (inner product space) that has completion. Showing that this mathematical object C^n is a pre-Hilbert space, we will have to show it has the positive definite property, the linear property, and some type of symmetry property. C^n will be shown to be complete by showing all the limit points are in C^n. Therefore, C^n is a Hilbert space.
2.1.1 Inner Product/Pre-Hilbert Space

An inner product space is an ordered quadruple:

\((X, +, ., \langle \cdot, \cdot \rangle)\), \((X, +, ., \cdot)\), \((X, +, ., \{\cdot\})\) or similar where

<table>
<thead>
<tr>
<th>Table 2.1: Ordered Quadruple ((X, +, ., \cdot)).</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) ((X, +, .)) is a vector space over (\mathbb{C}) (or (\mathbb{C})-vector space).</td>
</tr>
<tr>
<td>2) (\langle \cdot, \cdot \rangle), (\cdot), or ({\cdot}) is an inner product on the vector space (X) (or (\mathbb{C})-vector space) ([War], pg. 712).</td>
</tr>
</tbody>
</table>

Ordered Quadruple, the Inner Product Space.

And for every pair of vectors \(x, y \in X\), \(\langle x, y \rangle\) defines a complex number called the scalar product (inner product) of \(x\) and \(y\) satisfying the following conditions:

(a) 2.1.1a Positive Definite Property

\[ \langle x, x \rangle > 0 \text{ if } x \neq 0, \quad \langle 0, 0 \rangle = 0; \]

(b) 2.1.1b Conjugate Symmetry Property

\[ \langle x, y \rangle = \overline{\langle y, x \rangle} \text{ for every } x, y \in H; \]

(c) 2.1.1c Linear Property

i. \(\langle \alpha x, y \rangle = \alpha \langle x, y \rangle\) for every \(x, y \in H\)

and every complex number \(\alpha\);

ii. \(\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle\) for
2.1.2 Hilbert Space: (including the above properties)

(d) Completion

The format for the next few pages to define the pre-Hilbert space, then the Hilbert space is shown below.

2.1.1a Positive Definite Property. \( \langle x, x \rangle > 0 \) if \( x \neq 0 \),

\[ \langle 0, 0 \rangle = 0; \] From question 1 \( x = (x_1, x_2, \ldots x_n)^T \),

\[ y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n \text{ with} \]

\[
\begin{align*}
\langle x, y \rangle &= \sum_{k=1}^{n} x_k \overline{y_k} : \text{the inner product} \\
\| x \| &= \langle x, x \rangle^{\frac{1}{2}} : \text{the norm}
\end{align*}
\]

For \( C = \) the complex number field \( x \in C \) where \( a, b \in \mathbb{R} \),

For \( \begin{cases} x = a + bi \\ \overline{x} = a - bi \end{cases} \) (conjugate) \( \Rightarrow \begin{cases} x\overline{x} = a^2 + b^2 \\ \| x \| = \sqrt{a^2 + b^2} \text{ (pos. real)} \end{cases} \)

([Lip] pg. 51).

Therefore, replacing the second term \( y \) with \( x \) gives us

\[
\langle x, x \rangle = \sum_{k=1}^{n} x_k \overline{x_k} = x_1 \overline{x_1} + \ldots + x_n \overline{x_n}
\]

\[
= |x_1|^2 + \ldots + |x_n|^2 \geq 0
\]

and

\[
= |x_1|^2 + \ldots + |x_n|^2 = 0 \text{ if and only if each } x_i = 0.
\]
2.1.1b Conjugate Symmetry Property. \( \langle x, y \rangle = \overline{\langle y, x \rangle} \)

for every \( x, y \in H \);

For \( \left\{ \begin{array}{l}
x = a + bi \\
\overline{x} = a - bi \text{ (conjugate)}
\end{array} \right. \Rightarrow \overline{x} = x \) ([Lip] pg. 51).

Now for \( x = (x_1, x_2, \ldots, x_n)^T \), \( y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n \) we define the inner product:

\[
\langle x, y \rangle = \sum_{k=1}^{n} \overline{x_k} y_k = \overline{x_1 y_1} + \ldots + \overline{x_n y_n}
\]

\[
= y_1 \overline{x_1} + \ldots + y_n \overline{x_n} \quad \text{(i.e.,} \overline{x} = x) \]

\[
= \overline{\langle y, x \rangle}. \]

Another approach for the conjugate symmetry property will be given after linear property is presented since it applies the linear property.

2.1.1c Linear Property. (i). \( \langle ax, y \rangle = \alpha \langle x, y \rangle \) for every \( x, y \in H \) and every complex number \( \alpha \);

\[
\langle ax, y \rangle = \sum_{k=1}^{n} \alpha x_k \overline{y_k} = \alpha x_1 \overline{y_1} + \ldots + \alpha x_n \overline{y_n}
\]

\[
= \alpha (x_1 \overline{y_1} + \ldots + x_n \overline{y_n}) = \alpha \langle x, y \rangle.
\]

and similarly

\[
\langle x, \alpha y \rangle = \sum_{k=1}^{n} x_k \overline{\alpha y_k} = x_1 \overline{\alpha y_1} + \ldots + x_n \overline{\alpha y_n}
\]

\[
= \overline{\alpha (x_1 y_1 + \ldots + x_n y_n)} = \overline{\alpha} \langle x, y \rangle.
\]
If \( x = (\alpha_1 x_1, \alpha_2 x_2, \ldots, \alpha_n x_n) \) and \( y = (\beta_1 y_1, \beta_2 y_2, \ldots, \beta_n y_n) \) then the inner product is defined:

\[
\langle x, y \rangle = \left\langle \sum_{j=1}^{n} \alpha_j x_j, \sum_{k=1}^{n} \beta_k y_k \right\rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \beta_k x_j y_k.
\]

(ii) \( \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \) for every \( x, y, z \in H \).

\[
\langle x + y, z \rangle = \sum_{k=1}^{n} (x_k + y_k)z_k = (x_1 + y_1)z_1 + \ldots + (x_n + y_n)z_n
\]

\[
= (x_1 z_1 + \ldots + x_n z_n) + (y_1 z_1 + \ldots + y_n z_n)
\]

\[
= \langle x, z \rangle + \langle y, z \rangle.
\]

2.1.1d **Conjugate Symmetry Property (revisited)** \( \langle x, y \rangle = \langle y, x \rangle \) (linear property perspective)

\[
\langle x + y, x + y \rangle = \| x + y \|^2
\]

\[
\langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle
\]

\[
= \| x \|^2 + \langle x, y \rangle + \langle y, x \rangle + \| y \|^2
\]

\[
\| x + y \|^2 = \| x \|^2 + 2 \text{Re}\langle x, y \rangle + \| y \|^2
\]

\[
2 \text{Re}\langle x, y \rangle = \| x + y \|^2 - \| x \|^2 - \| y \|^2
\]

The arguments given above show that \( \text{Re}\langle x, y \rangle \) has all the properties of a real scalar product. Moreover, if \( \langle x, y \rangle \) is a complex scalar product, then

\[
\text{Im}\langle x, y \rangle = \text{Re}\langle x, iy \rangle = \text{Re}\bar{\langle x, y \rangle} = \text{Re}(-i)\langle x, y \rangle = -\text{Re}(ix, y)
\]
And $\text{Re}(ix, iy) = \text{Re}(i\bar{i})(x, y) = \text{Re}(x, y)$

therefore

$$\langle x, y \rangle = \text{Re}(x, y) + \text{Im}(x, y) = \text{Re}(x, y) - i\text{Re}(ix, y)$$

This gives

$$= \text{Re}(y, x) - ii\text{Re}(x, y)$$
$$= \text{Re}(y, x) - ii\text{Re}(y, x) \quad \text{(since } \text{Re}(x, y) = \text{Re}(y, x))$$
$$= \text{Re}(y, x) + i\text{Re}(y, ix)$$
$$= \langle y, x \rangle \quad \text{([Sche] pg. 245-246).}$$

Before looking at the completion of the inner product space, we define some terms for this space.

Definition: Let $(X, (\langle \cdot | \cdot \rangle_1))$ and $(W, (\langle \cdot | \cdot \rangle_2))$ be inner product spaces over the same field (complex numbers). A function $f: X \to W$ is an inner product space isomorphism if $f$ is an isomorphism (1-1 homomorphism, $1$-1 Hom$(X, W)$) from the vector space $X$ onto the vector space $W$ and if

$$(f(x)|f(y))_2 = (x|y)_1 \text{ for all } x, y \in X \quad ([War] \text{ pg. 713}).$$

Definition: An automorphism of an inner product space $X$ is an inner product space isomorphism from $X$ onto itself. The automorphism is also called a unitary linear operator on $X \quad ([War] \text{ pg. 714}).$
2.1.2 Hilbert Space

We need to go from a pre-Hilbert space to a Hilbert space and this procedure is called a "completion" of the original space (pre-Hilbert space) ([Ber] Pg. 26). To define a complete space lets review convergence of a sequence.

A sequence of vectors \( x_n \) \((n = 1, 2, \ldots)\) in a normed space \( X \) is said to converge to the vector \( x \) in \( X \) if

\[
\lim_{n \to \infty} \| x_n - x \| = 0,
\]

i.e., for any \( \varepsilon > 0 \) there is an integer \( n(\varepsilon) \) such that \( \| x_n - x \| < \varepsilon \) for \( n > n(\varepsilon) \) (uniform convergence).

A sequence of vectors \( x_n \) is said to be a fundamental (Cauchy) sequence if \( \lim_{m,n \to \infty} \| x_m - x_n \| = 0 \). The advantage of having a Cauchy sequence (form) is that it is not necessary to know the limit to show convergence. And the space \( X \) is said to be complete if every fundamental sequence converges to a vector in the vector space \( X \), i.e., if \( \lim_{m,n \to \infty} \| x_m - x_n \| = 0 \) implies the existence of a vector \( x \in X \) for which \( \lim_{n \to \infty} \| x_n - x \| = 0 \) holds ([Mat] pg. 47).
Thus, we say a vector space that has a scalar product (inner product) and is complete with respect to the induced norm (abstractly an absolute value/length) is a Hilbert space ([Sche] pg. 11). Succinctly, a complete inner product space is called a Hilbert space.

Related theorems that demonstrate convergence are presented next.

Table 2.2: Cauchy-Schwartz Inequality.

| Theorem | In a Hilbert space \( H \), the norm of the inner product \( |(x, y)| \leq \|x\|\|y\| \) \( x, y \in H \), where \( \|x\| := (x, x)^{1/2} \) and \( \|y\| := (y, y)^{1/2} \) (see Appendix B for the proof). |

Theorem: If \( x_n \rightarrow x \) and \( y_n \rightarrow y \), then \( \langle x_n, y_n \rangle \rightarrow \langle x, y \rangle \). (see Appendix B for the proof).

Theorem: Let \( \{x_n\} \) be a sequence in \( X \) such that \( x_n \rightarrow x \), where \( x \in X \), and let \( \{y_n\} \) be a sequence in \( X \). Then

(i) \( (z, x_n) \rightarrow (z, x) \) for all \( z \in X \);

(ii) \( (x_n, z) \rightarrow (x, z) \) for all \( z \in X \);

(iii) \( \|x_n\| \rightarrow \|x\| \); and
if \( \sum_{n=1}^\infty y_n \) is convergent in \( X \), then \( \left( \sum_{n=1}^\infty y_n, z \right) = \sum_{n=1}^\infty (y_n, z) \) for all \( z \in X \).

2.1.2 Completion. This section shows the completion of the pre-Hilbert space. Given two Cauchy sequences of vectors in \( \mathbb{C}^n \) where the difference of the norm converges to zero implies that each Cauchy sequence converges in \( \mathbb{C}^n \). The limit that the sequence \( x_m \) converges to below is \( x_0 \). Proof:

For \( \mathbb{C}^n = \{ (x_1, \ldots, x_n) \} : \| (x_1, \ldots, x_n) \| = \left( \sum_{i=1}^n |x_i|^2 \right)^{1/2} \)

we have \( \{ x_m = (x_{m1}, \ldots, x_{mn}) \}_{m=1}^\infty \) is Cauchy. So for

(sequence) \( x_m = (x_{m1}, x_{m2}, \ldots x_{mn}) = (x_{m,1}) \) and

(sequence) \( x_k = (x_{k,1}, x_{k,2}, \ldots x_{k,n}) = (x_{k,1}) \)

we have \( \| x_m - x_k \| \to 0 \) as \( m, k \to \infty \) (Def. Cauchy Seq.)

But \( 0 < \| x_m - x_k \| \geq |x_{m,1} - x_{k,1}| \) for \( 1 \leq i \leq n \)

\( \therefore \{ x_{m,1} \} \) is a Cauchy sequence in \( \mathbb{C} \) (i = 1, ..., n)

The Limit

\[ \therefore \exists x_0^i \in \mathbb{C} : x_{m,1} \to x_0^i \text{ as } m \to \infty \]

Let \( x_0 = (x_0^1, x_0^2, \ldots, x_0^n) \in \mathbb{C}^n \), \( \| x_m - x_0 \| \to 0 \) as \( m \to \infty \)
\[ C^n \text{ is complete.} \]

We conclude \( C^n \) is a Hilbert space. That is, the \( n \)-dimensional complex Euclidean space is a Hilbert space.
CHAPTER THREE

INTRODUCTION: ORTHOGONALITY

Recall we are given a Euclidean space. In Euclidean geometry we admit only the transformations that preserve length. These are transformations of the form \( t(x) = Ux + a \) where \( U \) is an orthogonal matrix and \( a \in \mathbb{R}^n \).

Questions 2 and 3 in this chapter and Chapter Four respectively, will be developing the concept of orthogonality. Orthogonality has a certain structure in space and orthogonality can be demonstrated using mathematical notation. Recall a strictly positive, symmetric bilinear functional \( f \) on a complex linear space \( X \) is called an inner product (i.e., \( \{ \} \)). We will be taking an inner product perspective to orthogonality. ([Mich] pg. 115-117).

3.1 Question Two

For \( x, y \in \mathbb{C}^n \) show the following:

3.1.1 Part I: Show \( \langle x, y \rangle \geq 0 \iff \| x + \alpha y \| \leq \| x + y \| \)

for \( \forall \alpha \in \mathbb{C} \) s. t. \( |\alpha| = 1 \)
3.1.2 Part II: Show \( \langle x, y \rangle = 0 \) (i.e., \( x \) and \( y \) are orthogonal)

\[ \iff \| x + \alpha y \| \leq \| x + y \| \text{ for } \forall \alpha \in \mathbb{C} \text{ s.t. } |\alpha| = 1 \]

Note if we just concentrate on the Part II, Part I is just a subsection to show the orthogonality of part II.

3.1.1 Part I

(Solution) \( \Rightarrow \)

Suppose \( \langle x, y \rangle \geq 0 \Rightarrow \) then we want our conclusion to be

\[ \| x + \alpha y \| \leq \| x + y \| \text{ for } \alpha \in \mathbb{C} \text{ s.t. } |\alpha| = 1. \]

That is

\[ \| x + \alpha y \|^2 \leq \| x + y \|^2 \]

Now

\[ \| x + y \|^2 - \| x + \alpha y \|^2 \]

\[ = \langle x + y, x + y \rangle - \langle x + \alpha y, x + \alpha y \rangle \]

\[ = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \langle x, x \rangle + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle + \langle \alpha y, \alpha y \rangle \]

\[ = \| x \|^2 + \langle x, y \rangle + \langle y, x \rangle + \| y \|^2 - \| x \|^2 + \langle x, \alpha y \rangle + \langle \alpha y, x \rangle + \alpha \overline{\alpha} \langle y, y \rangle \]

\[ = \| x \|^2 + \langle x, y \rangle + \langle y, x \rangle + \| y \|^2 - \| x \|^2 - \langle x, \alpha y \rangle - \langle \alpha y, x \rangle - \| y \|^2 \]

\[ = \langle x, y \rangle + \langle y, x \rangle - [\langle x, \alpha y \rangle + \langle \alpha y, x \rangle] \]

\[ = \langle x, y \rangle + \overline{\langle x, y \rangle} - [\langle x, \alpha y \rangle + \overline{\langle x, \alpha y \rangle}] \]

\[ = 2 \text{Re} \{ \langle x, y \rangle \} - \overline{\alpha} \langle x, y \rangle \]

\[ = 2 \text{Re} \{ (1 - \overline{\alpha}) \langle x, y \rangle \}, \]
= \langle x, y \rangle \Re \left[ (1 - \bar{\alpha}) \right].

We are given \( \langle x, y \rangle \geq 0 \), and \( \Re \left[ (1 - \bar{\alpha}) \right] \geq 0 \) for
\( |\alpha| = 1 \). Thus,

\[
\| x + y \|^2 - \| x + \alpha y \|^2 = \langle x, y \rangle \Re \left[ (1 - \bar{\alpha}) \right] \geq 0
\]

and hence,

\[
\| x + \alpha y \| \leq \| x + y \| \text{ for all } \alpha \in C \text{ s.t. } |\alpha| = 1.
\]

Part I: Converse \( \Leftarrow \) (Proof by contradiction)

Suppose \( \langle x, y \rangle \neq 0 \)

and \( \| x + \alpha y \| \leq \| x + y \| \) for all \( \alpha \in C \text{ s.t. } |\alpha| = 1 \).

Then we shall show by contradiction that \( \langle x, y \rangle \geq 0 \).

Let \( \alpha = \frac{\langle x, y \rangle}{\| x, y \|} \neq 0 \). It follows from the first part we

have \( \Re \left[ (1 - \bar{\alpha}) \langle x, y \rangle \right] \geq 0 \) and since \( \bar{\alpha} = \frac{\langle x, y \rangle}{\| x, y \|} \)

\[
\Re \left[ (1 - \bar{\alpha}) \langle x, y \rangle \right] = \Re \left[ (1 - \frac{\langle x, y \rangle}{\langle x, y \rangle}) \langle x, y \rangle \right],
\]

\[
= \Re \left[ \langle x, y \rangle - \frac{\langle x, y \rangle^2}{\langle x, y \rangle} \right],
\]

\[
= \Re \left[ \langle x, y \rangle - \langle x, y \rangle \right],
\]

\[
= \Re \left[ \langle x, y \rangle - \langle x, y \rangle \right] < 0
\]
since \( (x, y) \neq 0 \) \( \implies \) a contradiction since by hypothesis \( 2 \Re[(1 - \bar{\alpha})(x, y)] \)
\[ \geq 0 \quad : \quad (x, y) \geq 0. \]

3.1.2 Part II

\( (x, y) = 0 \iff \| x + \alpha y \| \leq \| x + y \| \) for \( \forall \alpha \in \mathbb{C} \) s.t. \( |\alpha| = 1 \)
i.e., \( x, y \) are orthogonal. Applying Part I we need only check the following

\( (x, y) \geq 0 \iff \| x + \alpha y \| \leq \| x - y \| \) for \( \forall \alpha \in \mathbb{C} \) s.t. \( |\alpha| = 1 \)

For \( (x, y) = 0 \) both conditions above must be satisfied,

that is \( (x, y) = 0 \)

\[
\begin{align*}
(x, y) \geq 0 & \iff \| x + \alpha y \| \leq \| x + y \| \quad \text{(Part I)} \\
\text{and} & \\
(x, y) \leq 0 & \iff (x, -y) \geq 0 \\
& \iff \| x + \alpha(-y) \| \leq \| x - y \|, \text{ for } \forall \alpha \in \mathbb{C} \text{ s.t. } |\alpha| = 1
\end{align*}
\]

(true since \( \alpha \) can be \( \pm \))

For \( (x, y) = 0 \) both conditions above must be satisfied,
CHAPTER FOUR
ORTHOGONALITY

4.1 Question Three

Table 4.1: Question Three.
Collect known equivalence conditions for orthogonality
for two vectors in a Hilbert space (or in \( \mathbb{C}^n \)).

When we think of orthogonality the idea that
normally comes to mind is an angle of 90° or
perpendicular vectors. But how do we measure an n-
dimensional angle? We don’t, but orthogonality is a
concept of n-dimensional space.

4.2 Properties of Orthogonality:
Inner Product Perspective

Let \( X \) be a Hilbert space and let \( \langle \cdot, \cdot \rangle \) be a function
defined on a subset of \( X \times X \). We call this an
orthogonality. We say the orthogonality is:

1. Symmetric: \( x \perp y \Rightarrow y \perp x \);

inner product form: \( \langle x, y \rangle = \langle y, x \rangle = 0 \) (Real Symmetry)

One dimensional perspective shows that if \( x \perp y \) then \( \overline{x} \perp \overline{y} \).
(2) Homogeneous: \( x \parallel y \Rightarrow \forall a, b \) (scalars) \( ax \parallel by \)

inner product form: \( \langle ax, by \rangle = ab \langle x, y \rangle = ab \cdot 0 = 0 \)

(Conjugate Linear)

(3) Additive (linear) on the right:

\( x \perp y \) and \( x \perp z \Rightarrow x \perp (y + z) \) and

inner product form: \( \langle x, y + z \rangle = 0 = \langle x, y \rangle + \langle x, z \rangle \)

Additive (linear) on the left:

\( x \perp y \) and \( z \perp y \Rightarrow (x + z) \perp y \)

inner product form: \( \langle x, y \rangle + \langle z, y \rangle = 0 + 0 = \langle x + z, y \rangle = 0 \)

(4) Nontrivial: \( x, y \in X \Rightarrow \exists k, a scalar \) s.t. \( x \perp (kx + y) \).

If \( x \) is the zero vector, then it's true \( 0 \perp (k0 + y) \)

If \( k = 0 \), then \( x \perp (kx + y) \)

\[ = x \perp (0x + y) \]

\[ \Rightarrow x \perp y \quad \text{([Ist] pg. 171).} \]
Theorem: If $x$ is orthogonal to each of $y_1, y_2, \ldots, y_n$ (i.e., $x \perp y_k$), then $x$ is orthogonal to every linear combination of the $y_k$.

Proof.

If $x \perp y_k \ \forall k$, and $y = \sum_{k=1}^{n} \lambda_k y_k$, then inner product form: $\langle x, y \rangle = \sum_{k=1}^{n} \langle x, \lambda_k y_k \rangle = \sum_{k=1}^{n} \overline{\lambda_k} \langle x, y_k \rangle = \sum_{k=1}^{n} \overline{\lambda_k} 0 = 0$

([Ber] pg. 43).

Definition: Let $x \in H$ (Hilbert space) and $S$ be a subset of $H$. If $x \perp s$ for all $s \in S$ then $x \perp S$. Since $\langle x, s \rangle = 0$ for all $s \in S$, the set of all such vectors $x$ is called the annihilator of $S$. This set is denoted $S^\perp$. Thus:

$S^\perp = \{ x \in H: \langle x, s \rangle = 0 \ \text{for all}s \in S \}$

([Ber] pg. 59).

Definition: A set $S$ of vectors is said to be orthogonal if whenever distinct vectors $x, y \in S$ we have $x \perp y$. A sequence (finite or infinite) of vectors $x_n$ is called an orthogonal sequence if $x_j \perp x_k$ whenever $j \neq k$

([Ber] pg 43).
4.3 Projection: An Orthogonal Application

Figure 4.2 Projection of Vector \( y \) onto Vector \( x \).

**Theorem:** Let \( x \) be a nonzero vector. Then for any other nonzero vector \( y \), the vector \( z = y - \left( \frac{\langle y, x \rangle}{\|x\|^2} \right)x \) is orthogonal to \( x \) \hspace{1cm} ([Gro] pg. 17, 18)

**Proof.** \( z \cdot x = \)

\[
\left( y - \frac{\langle y, x \rangle x}{\|x\|^2} , x \right) = \langle y, x \rangle - \left( \frac{\langle y, x \rangle x}{\|x\|^2} , x \right) = \langle y, x \rangle - \frac{\langle y, x \rangle \|x\|^2}{\|x\|^2} = 0.
\]
4.3.1 Notation

Let \( y \) and \( x \) be nonzero vectors, then the projection of \( y \) on \( x \) is a vector, denoted by \( \text{Proj}_x y \), which is defined by

\[
\text{Proj}_x y = \frac{\langle y, x \rangle}{\langle x, x \rangle} x = \frac{\langle y, x \rangle}{\|x\|^2} x = \alpha \frac{x}{\|x\|}.
\]

The component of \( y \) in the direction of \( x \) is \( \alpha = \frac{\langle y, x \rangle}{\|x\|} \).

The scalar \( \alpha \) is also called the Fourier coefficient of \( y \) with respect to the component of \( y \) along \( x \). Note that \( \frac{x}{\|x\|} \) is a unit vector in the direction of \( x \)

([Gro] pg. 17, 18).

In application, let \( y \) be a function and \( x_k, k = 1, 2, 3, \ldots \) the standard ordered basis. Then, \( f \) can be thought of as an electronic signal (example: video, audio, etc.) and \( \alpha_j \) represents the Fourier coefficients or wavelet coefficients. That is, \( \alpha_j = \langle f, e_j \rangle \) where \( e_j, j = 1, 2, 3, \ldots \) forms an orthonormal basis which is constructed in the section under the Gram-Schmidt orthogonalization process coming up. Wavelets mentioned above can be thought as the next generation of Fourier
analysis. Orthogonality makes the role in finding coefficients easier.

4.3.2 Bessel's Inequality

Suppose \{e_1, e_2, \ldots, e_n\} is an orthonormal set of vectors in X. For any vector \(y \in X\) and \(\alpha_j\) the Fourier coefficients of \(y\) with respect to \(x_j\). Then

\[
\sum_{k=1}^{n} \alpha_k^2 \leq \|y\|^2.
\]

Note that \(\alpha_j = \langle y, e_j \rangle\) since \(\|e_j\| = 1\). Then

\[
\langle e_i, e_j \rangle = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
\]

and summing from \(k = 1\) to \(n\), we get

\[
0 \leq \left\langle y - \sum_{k=1}^{n} \alpha_k e_k, y - \sum_{k=1}^{n} \alpha_k e_k \right\rangle = \langle y, y \rangle - 2\left\langle y, \sum_{k=1}^{n} \alpha_k e_k \right\rangle + \sum_{k=1}^{n} \alpha_k^2
\]

\[
= \langle y, y \rangle - \sum_{k=1}^{n} 2\alpha_k \langle y, e_k \rangle + \sum_{k=1}^{n} \alpha_k^2
\]

\[
= \langle y, y \rangle - \sum_{k=1}^{n} 2\alpha_k^2 + \sum_{k=1}^{n} \alpha_k^2
\]

\[
= \langle y, y \rangle - \sum_{k=1}^{n} \alpha_k^2
\]

\[
0 \leq \langle y, y \rangle - \sum_{k=1}^{n} \alpha_k^2 \quad \Rightarrow \quad \sum_{k=1}^{n} \alpha_k^2 \leq \langle y, y \rangle
\]
Bessel's Inequality ([Lip] pg.212).

4.4 Gram-Schmidt Orthogonalization Process

This section considers a basis and constructs an orthogonal basis by the process represented in projections, except the process is repeated. The subspace spanned (generated, enveloped) by one basis will be orthogonal to the subspace spanned (generated, enveloped) by the other basis, which leads to defining orthogonal complement.

Suppose \( \{y_1, \ldots, y_n\} \) is a basis of an inner product space \( \mathcal{H} \). An orthonormal basis \( \{z_1, \ldots, z_n\} \) can be obtained for \( \mathcal{H} \) as follows. Set \( z_1 = y_1 \) and \( z_2 \perp z_1 \) so by the process above

\[
z_2 = y_2 - \frac{\langle y_2, z_1 \rangle}{\|z_1\|^2} z_1.
\]

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.3.png}
\caption{Process for \( z_j \perp z_k, \ k \neq j = 1, 2, \ldots, n \) (Space \( \mathcal{H} \))}
\end{figure}
$z_3 \perp z_1$ and $z_3 \perp z_2$ or $z_3 \perp (z_1 + z_2)$; our next basis vector

\[ z_3 = y_3 - \left\{ \frac{\langle y_3, z_1 \rangle}{\|z_1\|^2} z_1 + \frac{\langle y_3, z_2 \rangle}{\|z_2\|^2} z_2 \right\} \]

\[ = y_3 - \frac{\langle y_3, z_1 \rangle}{\|z_1\|^2} z_1 - \frac{\langle y_3, z_2 \rangle}{\|z_2\|^2} z_2 \]

\[ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \]

\[ z_n = y_n - \left\{ \frac{\langle y_n, z_1 \rangle}{\|z_1\|^2} z_1 + \frac{\langle y_n, z_2 \rangle}{\|z_2\|^2} z_2 + \ldots + \frac{\langle y_n, z_{n-1} \rangle}{\|z_{n-1}\|^2} z_{n-1} \right\} \]

\[ z_n = y_n - \frac{\langle y_n, z_1 \rangle}{\|z_1\|^2} z_1 - \frac{\langle y_n, z_2 \rangle}{\|z_2\|^2} z_2 - \ldots - \frac{\langle y_n, z_{n-1} \rangle}{\|z_{n-1}\|^2} z_{n-1} \]

By the Gram–Schmidt Orthogonalization Process \{z_1, \ldots, z_n\}
is an orthonormal basis to the given basis \{y_1, \ldots, y_n\}

{[Lip] pg. 212-213}.

4.4.1 Orthogonal Basis

Hilbert space $H$. Then for any $y \in H$,

\[ y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n \]

where

\[ \alpha_1 = \frac{\langle y, x_1 \rangle}{\|x_1\|^2}, \quad \alpha_2 = \frac{\langle y, x_2 \rangle}{\|x_2\|^2}, \quad \ldots, \quad \alpha_n = \frac{\langle y, x_n \rangle}{\|x_n\|^2} \]

that is,

\[ y = \frac{\langle y, x_1 \rangle}{\|x_1\|^2} x_1 + \frac{\langle y, x_2 \rangle}{\|x_2\|^2} x_2 + \ldots + \frac{\langle y, x_n \rangle}{\|x_n\|^2} x_n \]

The above scalar,
\[ \alpha_j \equiv \frac{\langle y, x_j \rangle}{\| x_j \|^2} = \frac{\langle y, x_j \rangle}{\| x_j \|^2} \]

is called the Fourier coefficient of \( y \) with respect to \( x_j \). Note that vector \( y \) is not orthogonal to \( x_j \), \( x_j \) are basis elements ([Lip] page 218, 210).

4.5 Orthogonal Complement

Vector \( (z - y) \perp Y \) (Subspace). Suppose \( Y \) is a subspace of \( H \) with basis \( \{y_1, ..., y_n\} \). For \( z \in H \) there corresponds a unique \( y \in Y \) such that

\[ \| z - y \| = \inf_{y' \in Y} \| z - y' \|. \]

That is, \( y \) is the point of \( Y \) nearest to point \( z \). Then \( z - y \) is orthogonal to every \( y_j \in Y \) and therefore orthogonal to the subspace \( Y \).

Let's define \( x = z - y \), since \( (z - y) \perp Y \rightarrow x \perp Y \), then it follows that \( x = z - y \rightarrow z = x + y \) where \( y \in Y \) and \( x \perp Y \). The set of vectors \( X \) orthogonal to the subspace \( Y \) is a subspace.

Let \( x_j \in X \) (1 ≤ j) and \( x_j \rightarrow x \).

Then \( \langle x_j, y \rangle = 0 \) and \( \langle x, y \rangle = \langle x - x_j, y \rangle \) and in absolute value
\[\|x - x_j, y\| \leq \|x - x_j\| \cdot \|y\| \text{ which converges to zero}\]

(See Bessel' inequality in this section). Hence, \(\langle x, y \rangle = 0\).

The subspace \(X\) is called the orthogonal complement of \(Y\) and is expressed

\[X = Z \ominus Y\]

It is easily seen that \(Y = Z \ominus X\).

And \(Z = Y \oplus X\).

\(Z\) is called direct sum of the subspaces \(X\) and \(Y\), in the given case, the orthogonal sum ([AkG] pg. 10).

4.5.1 Notation: (Orthogonal Complement)

The orthogonal complement of a subset \(S\) of Hilbert space \(H\) is denoted by \(H \ominus S\) or \(S^\perp\). The set

\[S^\perp = \{x \in H: \langle x, s \rangle = 0 \text{ for all } s \in S\}\]

([Lip] pg. 218, 210).
CHAPTER FIVE
GRAMIAN, THE HILBERT
$M_n(C)$-MODULE

This chapter demonstrates the properties of the Hilbert structure on the gramian.

5.1 Question Four
Given the information below:

5.1.1 Module
Let $M_n$ = the set of all $n \times n$ matrices with complex entries. $C^n$ becomes a left $M_n(C)$-module by defining the module action "•" by

$$A \cdot x = Ax = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Therefore $\cdot : M_n(C) \times C^n \rightarrow C^n$

(note: output is $C^n$: Hilbert space).

5.1.2 Gramian
Let $M_n$ = the set of all $n \times n$ matrices with complex entries. In this sense $C^n$ is a Hilbert $M_n(C)$-module, where the $M_n(C)$-valued inner product $[\cdot, \cdot] : C^n \times C^n \rightarrow M_n(C)$ (note: output is the matrix $M_n(C)$) is defined by
\[ [x, y] = x \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{bmatrix} \in M_n(C) \]

is a (normal) pre-Hilbert \( M_n(C) \)-module which satisfies the properties of question 4 ([Kak] pg. 17).

5.1.3 (Normal) Pre-Hilbert \( M_n(C) \)-Module

Check the following properties: for \( x, y, z \in C^n \)

1. 5.1.3a Positive Definite Property
   \[ [x, x ] \geq 0 \]
   \[ [x, x ] = 0 \iff x = 0 \]

2. 5.1.3b Linear Property
   \[ [ x + y, z ] = [ x, z ] + [ y, z ] \]

3. 5.1.3c Adjoint (Symmetry Property)
   \[ [ y, x ] = [ x, y ]^* \quad (\text{the adjoint matrix}) \]

4. 5.2 Trace. \( \text{Tr} [ x, y ] = ( x, y ) \), that is, \( \text{Tr}[\cdot] \) equals the original inner product in \( C^n \), and hence \( C^n \) is a normal Hilbert \( M_n(C) \)-module.

Solution:

5.1.3a Positive Definite Property. (1) \( A = [x, x] \geq 0 \).

Suppose \( w = (w_1, \ldots, w_n) \in C^n \),
Then \( A \cdot w = A^t w = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} a_{11} w_1 & \cdots & a_{1n} w_n \\ \vdots & \ddots & \vdots \\ a_{n1} w_1 & \cdots & a_{nn} w_n \end{bmatrix}. \)

If \( A = [x, x] \), then by definition

\[
A = [x, x] = x^t x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} x_1^2 & \cdots & x_n^2 \\ \vdots & \ddots & \vdots \\ x_1^2 & \cdots & x_n^2 \end{bmatrix}.
\]

Now \( Aw = \)

\[
\begin{bmatrix} x_1 x_1 & \cdots & x_1 x_n \\ \vdots & \ddots & \vdots \\ x_n x_1 & \cdots & x_n x_n \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} x_1 w_1 + x_2 w_2 + \cdots + x_n w_n \\ \vdots \\ x_n w_1 + x_2 w_2 + \cdots + x_n w_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n x_j w_j \\ \vdots \\ \sum_{j=1}^n x_j w_j \end{bmatrix}.
\]

Therefore, the inner product

\[
(Aw, w) = (Aw)^t w = \begin{bmatrix} \sum_{j=1}^n x_j w_j \\ \vdots \\ \sum_{j=1}^n x_j w_j \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j w_i w_j = \left( \sum_{i=1}^n x_i^2 w_i \right) \left( \sum_{j=1}^n x_j w_j \right) = \left( \sum_{i=1}^n x_i^2 w_i \right)^2 \geq 0.
\]

5.1.3b Linear Property. (2) \([x+y, z] = [x, z] + [y, z]\)

\[
[x + y, z] = (x + y)^t z = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.
\]
$$\begin{bmatrix}
(x_1 + y_1)_{\overline{z_1}} & \cdots & (x_n + y_n)_{\overline{z_n}} \\
\vdots & \ddots & \vdots \\
(x_n + y_n)_{\overline{z_1}} & \cdots & (x_n + y_n)_{\overline{z_n}}
\end{bmatrix}$$

Regroup

$$\begin{bmatrix}
x_1_{\overline{z_1}} + y_1_{\overline{z_1}} & \cdots & x_1_{\overline{z_n}} + y_1_{\overline{z_n}} \\
\vdots & \ddots & \vdots \\
x_n_{\overline{z_1}} + y_n_{\overline{z_1}} & \cdots & x_n_{\overline{z_n}} + y_n_{\overline{z_n}}
\end{bmatrix}$$

\((x \text{ part}) \quad (y \text{ part})\)

$$\begin{bmatrix}
x_1_{\overline{z_1}} & \cdots & x_n_{\overline{z_n}} \\
\vdots & \ddots & \vdots \\
x_n_{\overline{z_1}} & \cdots & x_n_{\overline{z_n}}
\end{bmatrix} + \begin{bmatrix}
y_1_{\overline{z_1}} & \cdots & y_1_{\overline{z_n}} \\
\vdots & \ddots & \vdots \\
y_n_{\overline{z_1}} & \cdots & y_n_{\overline{z_n}}
\end{bmatrix}$$

$$\begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} \begin{bmatrix}
[\overline{z_1}] & \cdots & [\overline{z_n}]
\end{bmatrix} + \begin{bmatrix}
y_1 \\
\vdots \\
y_n
\end{bmatrix} \begin{bmatrix}
[\overline{z_1}] & \cdots & [\overline{z_n}]
\end{bmatrix}$$

\[= x \overline{z}^t + y \overline{z}^t\]

\[= [x , z] + [y , z].\]

(3) 
\[\{Ax, y\} = A \{x, y\}\text{ for }A \in M_n(C)\]

\[\{Ax, y\} = (Ax) \overline{y} = \begin{bmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{bmatrix} \begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} \begin{bmatrix}
y_1 & \cdots & y_n
\end{bmatrix}
\]

\[= \begin{bmatrix}
\sum_{j=1}^{n} a_{1j}x_j \\
\vdots \\
\sum_{j=1}^{n} a_{nj}x_j
\end{bmatrix} \begin{bmatrix}
y_1 & \cdots & y_n
\end{bmatrix}\]
Start from the other side we get:

\[
A \begin{bmatrix} x \ v \ y \end{bmatrix} = A(x^t y) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 y_1 \vdots \vdots \vdots \vdots \vdots \vdots x_n y_n \end{bmatrix} = \begin{bmatrix} a_{11}(x_1 y_1) + a_{12}(x_2 y_2) + \cdots + a_{1n}(x_n y_n) \\ \vdots \\ a_{n1}(x_1 y_1) + a_{n2}(x_2 y_2) + \cdots + a_{nn}(x_n y_n) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i} x_i y_i \\ \vdots \\ \sum_{i=1}^n a_{ni} x_i y_i \end{bmatrix}
\]

:. \( [Ax, y] = A[x, y] \).

5.1.3c Adjoint (Symmetry Property).

\[ [y, x] = y^t x = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \begin{bmatrix} x_1 \cdots x_n \end{bmatrix} = \begin{bmatrix} Y_1 x_1 \cdots Y_1 x_n \\ \vdots \\ Y_n x_1 \cdots Y_n x_n \end{bmatrix} = \begin{bmatrix} \overline{Y_1} & \cdots & \overline{Y_n} \\ \vdots & \ddots & \vdots \\ \overline{Y_n} & \cdots & \overline{Y_n} \end{bmatrix} \]

For \( [x, y] \) we get:

\[ [x, y] = x^t y = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 \cdots y_n \end{bmatrix} = \begin{bmatrix} \overline{x_1} & \cdots & \overline{x_n} \\ \vdots & \ddots & \vdots \\ \overline{x_n} & \cdots & \overline{x_n} \end{bmatrix} = \begin{bmatrix} \overline{x_1} y_1 \cdots \overline{x_n} y_n \\ \vdots \\ \overline{x_n} y_1 \cdots \overline{x_n} y_n \end{bmatrix} \]

Therefore, the adjoint of \( [x, y] \), that is, the conjugate transpose, is represented below:

\[ [x, y]^* = \]
Conjugate $= \begin{bmatrix} x_1y_1 & \cdots & x_1y_n \\ \vdots & \ddots & \vdots \\ x_ny_1 & \cdots & x_ny_n \end{bmatrix} = \begin{bmatrix} y_1x_1 & \cdots & y_1x_n \\ \vdots & \ddots & \vdots \\ y_nx_1 & \cdots & y_nx_n \end{bmatrix}$, and the conjugate transpose $= \begin{bmatrix} y_1x_1 & \cdots & y_1x_n \\ \vdots & \ddots & \vdots \\ y_nx_1 & \cdots & y_nx_n \end{bmatrix} = [y, x]$.

5.2 Trace

(5) $\text{Tr} [x, y] = (x, y)$

$[x, y] = x \overline{y}^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} \overline{y}_1 & \cdots & \overline{y}_n \end{bmatrix} = \begin{bmatrix} x_1\overline{y}_1 & \cdots & x_1\overline{y}_n \\ \vdots & \ddots & \vdots \\ x_n\overline{y}_1 & \cdots & x_n\overline{y}_n \end{bmatrix}$

$\text{Tr} [x, y] = \text{sum of the diagonal of } [x, y]$. Therefore,

$\text{Tr}[x, y] = \sum_{k=1}^{n} x_k\overline{y}_k = (x, y) = (x, y)_{C^n}$, the original inner product in $C^n$. The pre-Hilbert space (inner product space) in $C^n$ was shown to be complete in question one. $C^n$ is a Hilbert space, hence $C^n$ is a normal Hilbert $M_n(C)$-module.

This concludes question four, however to answer question five, we must add to our vocabulary. We have already looked at positive definite functions (operators) and a finite dimensional space $C^n$. It so happens that, the operators $A$ for which the summation $\sigma(A) < +\infty$ form
the *E. Schmidt-class* \((\sigma_c)\). Therefore, *Schmidt-class* \((\sigma_c)\) is a finite Hilbert space ([Scha] pg. 31, 34).

So for operators \(A, B\) belonging to \((\sigma_c)\), let \(\{\phi_j\}\) be a basis in \(C^n\) (complex Hilbert Space), then

\[
\left\{ \sum_{j=1}^{\infty} (A\phi_j, B\phi_j) : \text{inner product in } (\sigma_c), (A, A)^{1/2} = \sigma(A) : \text{norm} \right\}
\]

Thus, \((\sigma_c)\) is a Hilbert space that is independent on the chosen basis \(\{\phi_j\}\) ([Scha] pg. 31, 34).

5.2.1 **Definition:** (Trace-Class)

The product of two operators in \((\sigma_c)\) form the trace-class \((\tau c)\). For \(A \in (\tau c)\) the finite number

\[
t(A) = \sum_{j=1}^{\infty} (A\phi_j, \phi_j) : \text{trace of } A \text{ (sum of the diagonal elements)} \text{ and } \tau(A) = t((A^*A)^{1/2}), \tau(A) = \sum_{j=1}^{\infty} \left( (A^*A)^{1/2} \phi_j, \phi_j \right) : \text{trace norm} \text{ ([Scha] pg. 4, 36, 37)}.
\]
A complex vector space with inner product is called a complex inner product space or a unitary space. Conversely we say that a complete unitary space is called a Hilbert space. The unitary spaces of finite dimension are necessarily complete ([Ham] pg. 94).

Question 5 deals with a unitary operator/matrix so this chapter looks into the concept of being unitary before answering question 5.

In three-dimensional Euclidean space the simplest operation after that of projection is rotation of the space, which changes neither the length of the vector nor the angles between pairs of them. Unitary operators are analogous to this in a Hilbert space ([AkG] pg. 72). The symmetry operations of an object (or point set, Platonic or Archimedian polyhedra, polytope, molecule, etc.) have a matrix representation. We will be replacing the geometry of the symmetry operations with the algebra of matrices.
6.1 Matrix Perspective

Let $k \in C$ and $k^n$ be the set of all ordered $n$-tuples of elements of $k$. When working with $n^2$-tuples there is nothing to gain by writing $n^2$ elements in a line instead of an $n \times n$ array. So, for $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ on $k^n$ define:

1. Addition
   
   $x + y = (x_1 + y_1, ..., x_n + y_n)$,

   Identity:
   
   $0 = (0, ..., 0)$, and

2. For scalar $c \in k$
   
   $cx = (cx_1, ..., cx_n)$.

Thus makes $k^n$ into a vector space over $k$ ([Cur] pg. 12).

6.1.1 Linear Map

$M_n(C)$ is the set of all $n \times n$ matrices with elements from $C$ (C, the complex numbers). If $M \in M_n(k)$, $M = (m_{jk})$

$(m_{jk} \in k)$, then a linear map $\phi(M)$ is defined by

$\phi(M)(x_1, ..., x_n) = (x_1, ..., x_n)(m_{jk})$

and this is seen to be linear by ([Cur] pg. 13)

1. $\phi(M)(x + y) = (x + y)(m_{jk})$

   $= (x_1, ..., x_n)(m_{jk}) + (y_1, ..., y_n)(m_{jk})$

2. $\phi(M)(cx + dy) = (cx + dy)(m_{jk})$

   $= c(x_1, ..., x_n)(m_{jk}) + d(y_1, ..., y_n)(m_{jk})$
6.1.2 Vector Space: \((M_n(C), +, \cdot(C))\)

The set \(M_n(C)\) is a vector space and is demonstrated below:

Additive Structure: \((M_n(C), +)\)

(i) If \(A = (a_{jk})\) and \(B = (b_{jk})\), then

\[ A + B = (a_{jk} + b_{jk}); \]

Multiplicative (Scalar) Structure: \((M_n(C), \cdot(C))\)

(ii) If \(A = (a_{jk})\) and \(c \in k\), then

\[ cA = (ca_{jk}) \]


6.1.3 Algebra: \((M_n(C), +, \cdot(C), \cdot)\)

1) \((M_n(C), +, \cdot(C))\) is a vector space

2) \(\cdot\) is Associative (i.e., multiplication of matrices)

3) a. \(A(B+C) = AB + AC\); left distributive

\[ A(B+C) = AB + AC; \]

b. \((B+C)A = BA + CA\); right distributive

([Cur] pg. 15).

6.2 Unitary matrix

A matrix is unitary if its adjoint is equal to its inverse ([Ham] pg. 91, [War] pg. 737).

\[ \overline{U}^T = U^{-1} \]

\[ \rightarrow \overline{U}^T U = U^{-1}U = E \] (T stands for transpose)

\[ U^{-1}U = E \] and \[ UU^{-1} = E \]
6.3 Similarity Transformations

A matrix $B$ is similar to a matrix $A$ if there exists a nonsingular matrix $U$ such that $B = U^{-1}AU$. We say that $A$ and $B$ are similar matrices and are related by a similar transformation. If $U$ is a unitary matrix, then $A$ and $B$ are related by a unitary transformation ([Bis] pg. 57).

Table 6.1: Motion from $A$ to $B$, but Different Basis.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi(A)$</td>
<td>$\phi(A)$</td>
<td>$\gamma(A)$</td>
</tr>
<tr>
<td>$B$</td>
<td>$B$</td>
<td>$B$</td>
</tr>
<tr>
<td>$X_{3a}$</td>
<td>$X_{3b}$</td>
<td>$X_{3c}$</td>
</tr>
<tr>
<td>$X_{2a}$</td>
<td>$X_{2b}$</td>
<td>$X_{2c}$</td>
</tr>
<tr>
<td>$X_{1a}$</td>
<td>$X_{1b}$</td>
<td>$X_{1c}$</td>
</tr>
</tbody>
</table>

$A = \sum_{i=1}^{3} \alpha_j x_{ja}$
$\phi A = \phi (\sum_{i=1}^{3} \alpha_j x_{ja}) = B$

$A = \sum_{i=1}^{3} \beta_j x_{jb}$
$\phi A = \phi (\sum_{i=1}^{3} \beta_j x_{jb}) = B$

$A = \sum_{i=1}^{3} \mu_j x_{jc}$
$\gamma A = \gamma (\sum_{i=1}^{3} \mu_j x_{jc}) = B$

(The bases in the figure are orthonormal.)

There are many different matrix representations by the same transformation operators and the same function space but with different choices of basis functions describing that space that are equivalent (see Table 6.1.
above for simple example). Any pair of such equivalent representations has corresponding matrices that are linked by similarity transformations ([Bis] pg.103).

6.4 Theorems Related by a Similarity Transformation

(See Appendix B for the proofs of the theorems below.)

([Bis] pg. 65-67)

1). Theorem: If \( U^1A U = B \), then \( \det(A) = \det(B) \)

(Their Determinants are Equal).

2). Theorem: If \( U^1AU = B \), then the eigenvalues of \( A \) and \( B \)

are identical (Their Eigenvalues are Equal).

3). Theorem: If \( A' = U^1AU \), \( B' = U^1BU \), \( C' = U^1CU \) etc.,

then any relationship between \( A \), \( B \), \( C \), etc. is also

satisfied by \( A' \), \( B' \), \( C' \) etc.

4). Theorem: If \( A \) and \( B \) are two matrices which can be

diagonalized by the same matrix \( U \) then \( A \) and \( B \)

commute.

5). Theorem: If \( X \) is the matrix formed from the

eigenvectors of \( A \), then \( X^1AX \) is a diagonal matrix \( A \)

composed of the eigenvalues of \( A \).

6). Theorem: If \( U^1AU = B \), then \( \text{trace}(A) = \text{trace}(B) \)

(Their Traces are Equal).
Theorem: A unitary transformation leaves a unitary matrix unitary.

(See Appendix B for the proofs of the theorems.)

It is always possible to find a set of basis functions that produce unitary matrices by the Gram-Schmidt orthogonalization process. If we choose our basis functions for a particular function space to be orthonormal (orthogonal and normalized), then since the transformation operators are unitary, the representation created will consist of unitary matrices (a unitary representation). That is, a transformation operator $U$ that will leave the scalar product of two functions of the function space unchanged (i.e., $\langle Uf_1, Uf_2 \rangle = \langle f_1, f_2 \rangle$).

And convenience dictates that we do this since unitary matrices are much easier to handle and manipulate than non-unitary matrices ([Bis] pg. 108-109).

Definition: Elements of a group such that $U^{-1}AU = B$ makes $A$ and $B$ conjugate to each other. Elements of a group that are congruent to each other are said to form a class (i.e., class of rotation, class of reflections, etc.), ([Bis] pg. 32).
6.5 Group Relation

Recall $\mathbb{C}^n$ is a Hilbert $\mathcal{M}_n(\mathbb{C})$-module, where the $\mathcal{M}_n(\mathbb{C})$-valued inner product $[\cdot, \cdot]: \mathbb{C}^n \times \mathbb{C}^n \to \mathcal{M}_n(\mathbb{C})$ is called a Gramian. $\mathcal{M}_n(\mathbb{C})$, the output (co-domain), is also a linear space and an algebra. This section is concerned with the group elements of $\mathcal{M}_n(\mathbb{C})$.

6.5.1 Group

A group is a set $G$ of elements under a binary operation satisfying the axioms of:

1) Closure: $a, b \in G \to a \cdot b \in G$
    (or more simply $ab \in G$).

2) Associatively: $a(bc) = (ab)c$ for $a, b, c \in G$

3) Identity: $e \in G$ such that $ea = ae = a$, for every $a \in G$.

4) Inverse: $a^{-1} \in G$ such that $a^{-1}a = aa^{-1} = e$, for every $a \in G$ ([Fan] pg. 65).

Definition: For $\mathcal{M}_n(\mathbb{C})$ an algebra, $A \in \mathcal{M}_n(\mathbb{C})$ is a unit if there exist some $B \in \mathcal{M}_n(\mathbb{C})$ such that $AB = I = BA$, i.e., if it is a multiplicative inverse ([Cur] pg. 15).

Proposition: If $\mathcal{M}_n(\mathbb{C})$ is an algebra with an associative multiplication, $A \in \mathcal{M}_n(\mathbb{C})$ is a unit if there exist some $B \in \mathcal{M}_n(\mathbb{C})$ such that $AB = I = BA$, i.e., if it is a
multiplicative inverse ([Cur] pg. 15). And if $U \subseteq M_n(C)$ is the set of units in $M_n(C)$, then $U$ is a group under multiplication.

Now, the group of units has a special notation. The group of units in the algebra $M_n(C)$ is denoted by $GL(n,C)$ where $C$ represents the complex numbers and

$$GL(n,C) = \{ A \in M_n(C) : \det A \neq 0 \}.$$ 

Definition: The orthogonal group is defined:

$$O(n,k) = \{ A \in M_n(k) : \langle Ax, Ay \rangle = \langle x, y \rangle \text{ for all } x, y \in k^n \}$$

for $k = \text{reals numbers, complex numbers, or the quaternions } (k \in \{ \mathbb{R}, \mathbb{C}, \mathbb{H} \}, \text{ respectively})$. For $k = \text{complex numbers}$ we write $O(n, C)$ as $U(n)$ and call it a unitary group ([Cur] pg. 26, 27).

6.6 Unitary Group

When $k$ is the complex number field we write $O(n,C)$ as $U(n)$ and it is called the unitary group, but the group part still has to be proven

(See Appendix B for the Proof).

Let $A \in M_n(C)$. Then the following conditions are equivalent:

(i) $A \in U(n)$, (unitary group).
(ii) \( \langle e_i A, e_j A \rangle = \delta_{i,j} \)

(iii) \( A \) sends orthonormal bases to orthonormal bases

(iv) The rows of \( A \) form an orthonormal basis

(v) The columns of \( A \) form an orthonormal basis

(vi) \( A^* = A^{-1} \)

(See Appendix B for the Proof, [Cur] pg 27).

Question 5 will be answered by applying the next theorem.

Theorem: Let \( A \in \{\text{tc}\} \), the trace-class. Then for

\( \text{R}_t(A) \) - the range of \( t(A) \) and

\( \text{R}_t(UA) \) - the range of \( t(UA) \),

we have

\[ \text{R}_t(A) \geq \text{R}_t(UA) \text{ for all unitary } U \text{ if and only if } A \geq 0. \]

(see [Scha] pg. 43).

(This will be used to prove question five)

6.7 Question Five

Find the necessary and sufficient condition for \([x, y] \geq 0\)

\([x, y] \geq 0 \iff \|x + Uy\| \leq \|x + y\| \text{ for } \forall U: \text{unitary} \)

6.5.1a Proof. \[ \Rightarrow \]

\([x, y] \geq 0 \Rightarrow \|x + y\|^2 - \|x + Uy\|^2 \]

\[ = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \{\langle x, x \rangle + \langle x, Uy \rangle + \langle Uy, x \rangle + \langle Uy, Uy \rangle\} \]

\[ = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \langle x, x \rangle - \langle x, Uy \rangle - \langle Uy, x \rangle - \langle y, y \rangle \]
\[(\text{since } \langle U^*_y, U^*_y \rangle = \langle y, y \rangle)\]

\[= \langle x, y \rangle + \langle y, x \rangle - \langle x, U^*_y \rangle - \langle U^*_y, x \rangle\]

\[= \langle x, y \rangle + \langle y, x \rangle - \langle u^*_x, u^*_y \rangle - \langle U^*_x, U^*_y \rangle\]

\[= \langle x, y \rangle + \langle y, x \rangle - \langle u^*_x, y \rangle - \langle U^*_x, y \rangle\]

\[= 2 \text{Re} \langle x, y \rangle - 2 \text{Re} \langle U^*_x, y \rangle\]

\[= 2 \text{Re}[\langle (I - U^*)x, y \rangle] \quad \text{(Question 4.5 we have } \text{Tr}[x, y] = \langle x, y \rangle)\]

\[= 2 \text{Re}[\langle (I - U^*)x, y \rangle] = 2 \text{Re} \text{Tr}((I - U^*)[x, y]) \quad \text{(A)}\]

---

\[\text{Since } A \geq 0 \text{ if and only if } \text{Rt}(A) \geq \text{Rt}(UA) \text{ for all unitary } U.\]

\[\Rightarrow \quad \text{Since } [x, y] \geq 0 \text{ we have } \text{Tr}[x, y] \geq \text{Tr}[U^*x, y]\]

\[\Rightarrow \quad 2 \text{Re} \text{Tr}((I - U^*)[x, y]) \geq 0 \quad \text{(A)}\]

\[\Rightarrow \quad \| x + U y \| \leq \| x + y \| \text{ for } \forall U \text{ unitary.}\]

\[\text{Proof of the Converse. } \Leftrightarrow \]

\[\| x + y \| \geq \| x + U y \| \text{ for } \forall U \text{ unitary}\]

\[\Rightarrow \quad \| x + y \|^2 - \| x + U y \|^2 \geq \| x + U y \|^2 - \| x + U y \|^2\]

\[\Rightarrow 0 \leq \| x + y \|^2 - \| x + U y \|^2\]

\[= \| x, x \| + \| x, y \| + \| y, x \| + \| y, y \| - (\langle x, x \rangle + \langle x, U y \rangle + \langle U y, x \rangle + \langle U y, U y \rangle)\]

\[= \| x, x \| + \| y, y \| \quad \text{(from part 1 above)}\]
\[= \quad 2\text{Re}(x, y) - 2\text{Re}(U^*x, y)\]

\[= 2\text{Re}\{(I - U^*)x, y\}\]

(Question 4.5 we have \(\text{Tr}[x, y] = (x, y)\))

\[= 2\text{Re}\{(I - U^*)x, y\} = 2\text{Re} \text{Tr}((I - U^*)[x, y])\]

\[\text{Since } \text{Re}(A) \geq \text{Re}(UA) \text{ for all unitary } U \text{ if and only if } A \geq 0.\]

(From the previous proof)

\[2\text{Re} \text{Tr}((I - U^*)[x, y]) \geq 0\]

\[\Rightarrow \text{Tr}[x, y] \geq \text{Tr}[U^*x, y]\]

and since \(\text{Tr}[x, y] \geq \text{Tr}[U^*x, y]\) we have \([x, y] \geq 0\).
CHAPTER SEVEN

REPRODUCING KERNEL HILBERT

$\mathcal{M}_p(C)$-MODULE

This chapter demonstrates the Hilbert structure on the reproducing kernel Hilbert $\mathcal{M}_p(C)$-module. Just as for $\mathbb{C}^p$ (the $p$-dimensional complex Euclidean space), and the gramian, we will have to show positive definite property, a linear property and some type of symmetry property.

First, we consider the reproducing kernel Hilbert space (RKHS) and use it as an analogy to interpret and define the reproducing kernel Hilbert $\mathcal{M}_p(C)$-module (RKH-module). That is, for the kernel $\gamma(x,y)$ there corresponds a class $F$ of functions $f(x)$, in respect to which $\gamma$ possess the "reproducing" property. On the other hand, to a class of functions $F$, there may correspond a kernel $\gamma$ with the "reproducing" property ([Ar] pg. 338).

There are two trends in the consideration of these kernels. Those following the first trend consider a given kernel $\gamma$ and study it in itself, or eventually apply it in various domains (as integral equations, theory of groups, general metric geometry, etc). The
class F corresponding to \( \gamma \) may be used as a tool of research for the kernel. In the second trend, one is interested primarily in the class of functions F, and the corresponding kernel \( \gamma \) is used essentially as a tool in the study of the functions of this class ([Ar] pg. 338).

7.1 Reproducing Kernel Hilbert Space and Reproducing Kernel Hilbert-\( M_p(C) \)-Module

An obvious difference between the RKHS and RKH-module is shown in the table below.

<table>
<thead>
<tr>
<th>Table 7.1: Range Comparison of the Kernel.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kernel</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>RKHS</td>
</tr>
<tr>
<td>RKH-module</td>
</tr>
</tbody>
</table>

Table 7.2: The Reproducing Property.

\[
f(\omega) = (f(\bullet), \gamma(\omega, \bullet)) \\

\text{for } \forall f \in \text{Hilbert Space}, \forall \omega \in \Omega; \text{ RKHS.}
\]

\[
x(\omega) = [x(\bullet), \Gamma(\omega, \bullet)]
\]

\text{for } \forall x \in \text{Hilbert } M_p(C)-\text{module}, \forall \omega \in \Omega; \text{ RKH-module.}

72
7.1.1 Reproducing Kernel Hilbert Space

The reproducing kernel Hilbert space (RKHS) is defined below. For $\forall \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}$, $\forall \omega_1, \ldots, \omega_n \in \Omega \ \forall n \geq 1$

$$\begin{cases}
X \text{ be a set of functions on } \Omega \\
\gamma: \Omega \times \Omega \rightarrow \mathbb{C}, \text{ positive definite } \Rightarrow X \text{ is a RKHS of } \gamma \\
\gamma: \text{pos. def. } \Leftrightarrow \sum_{j,k} \alpha_j \alpha_k \gamma(\omega_j, \omega_k) \geq 0
\end{cases}$$

(When we consider the kernel $\gamma$ by itself it is called a positive matrix or positive definite ([Ar] pg. 341)). The conditional statement above together implies:

- $X$: a Hilbert space with an inner product $(\cdot, \cdot)$
  - (1) $\gamma(\omega, \cdot) \in X$ for $\forall \omega \in \Omega$
  - (2) $f(\omega) = (f(\cdot), \gamma(\omega, \cdot))$ for $\forall f \in X, \forall \omega \in \Omega$

(When we show that the kernel corresponding to a class of functions it is called the reproducing kernel of the class ([Ar] pg. 341)).

Now, we repeat for a module.

7.1.2 Reproducing Kernel Hilbert $M_p(\mathbb{C})$-Module

For $\forall A_1, \ldots, A_n \in M_p(\mathbb{C}), \forall \omega_1, \ldots, \omega_n \in \Omega \ , \forall n \geq 1$, let

$$\begin{cases}
X \text{ be a set of functions on } \Omega \\
\Gamma: \Omega \times \Omega \rightarrow M_p(\mathbb{C}), \text{ pos. definite } \Rightarrow \begin{cases}X \text{ is a RKH } M_p(\mathbb{C}) \text{ }-\text{module of } \Gamma.\end{cases} \\
\Gamma: \text{pos. def. } \Leftrightarrow \sum_{j,k=1}^{n} A_j \Gamma(\omega_j, \omega_k) A^*_k \geq 0
\end{cases}$$
7.2 Question Six

Construct a reproducing kernel Hilbert $M_p(C)$-module $X_r$ from $\Gamma: \Omega \times \Omega \rightarrow M_p(C)$, pos. def. in analogy to RKHS.

Let $X_0$ be a pre-Hilbert $M_p(C)$-module such that,

$$X_0 = \left\{ \sum_{j=1}^{n} A_j \Gamma(\omega_j, \cdot): A_j \in M_p(C), \; \omega_j \in \Omega, \; 1 \leq j \leq n, \; n = 1, 2, \ldots \right\}$$

and define the gramian:

$$[\Gamma(\omega, \cdot), \Gamma(\omega', \cdot)] = \Gamma(\omega, \omega') \text{ for } \omega, \omega' \in \Omega.$$

Therefore, we must show

7.2.1 Reproducing Kernel Pre-Hilbert $M_p(C)$-Module

(1) 7.2.1a Positive definite Property

$$[f, f] \geq 0$$

(2) 7.2.1b Linear Property

$$[f + g, h] = [f, h] + [g, h]$$

(3) $$[Af(\cdot), g(\cdot)] = A[f(\cdot), g(\cdot)]$$

(4) 7.2.1c Adjoint (Symmetry Property)

$$[f, g]^* = [g, f]$$

and
7.2.2 Reproducing Kernel Hilbert $M_p(C)$-Module:

Functional Completion. (above included)

(5) Functional Completion.

This demonstrates the construction of the RKH $M_p(C)$-module.

7.2.1 Reproducing Kernel Pre-Hilbert $M_p(C)$-Module

7.2.1a Positive Definite Property. \([f, f] \geq 0.\)

For \(f(*) = \sum_{j=1}^{n} A_j \Gamma(\omega_j, *)\in X_0\) we have,

\[
[f, f] = \left[ \sum_{j=1}^{n} A_j \Gamma(\omega_j, *), \sum_{j=1}^{n} A_j \Gamma(\omega_j, *) \right]
\]

\[
= \sum_{j,k=1}^{n} A_j \Gamma(\omega_j, *) \Gamma(\omega_k, *) A_k^* \]

\[
\geq 0 \text{ since } \Gamma \text{ is positive definite.}
\]

7.2.1b Linear Property. \([f+g, h] = [f, h] + [g, h]\)

For \(f(*) = \sum_{j=1}^{n} A_j \Gamma(\omega_j, *), \ g(*) = \sum_{k=1}^{n} B_k \Gamma(\omega_k, *), \) and

\(h(*) = \sum_{i=1}^{l} C_i \Gamma(\omega_i, *) \in X_0.\)

\[
[f(*) + h(*), g(*)] = \left[ \sum_{j=1}^{n} A_j \Gamma(\omega_j, *) + \sum_{i=1}^{l} C_i \Gamma(\omega_i, *), \sum_{k=1}^{n} B_k \Gamma(\omega_k, *) \right]
\]
\[
\sum_{j=1}^{n} \sum_{k=1}^{m} A_j \left[ \Gamma(\omega_j, \bullet), \Gamma(\omega'_k, \bullet) \right] B_k^* + \sum_{i=1}^{d} \sum_{k=1}^{m} C_i \left[ \Gamma(\omega'_i, \bullet), \Gamma(\omega'_k, \bullet) \right] B_k^*
\]

\[
= \sum_{j=1}^{n} \sum_{k=1}^{m} A_j \Gamma(\omega_j, \omega'_k) B_k^* + \sum_{i=1}^{d} \sum_{k=1}^{m} C_i \Gamma(\omega'_i, \omega'_k) B_k^*
\]

\[
[f(\bullet), g(\bullet)] + [h(\bullet), g(\bullet)]
\]

\[
[A f(\bullet), g(\bullet)] = A [f(\bullet), g(\bullet)]
\]

For \(f(\bullet) = \sum_{j=1}^{n} A_j \Gamma(\omega_j, \bullet) \in X_0\) and \(g(\bullet) = \sum_{k=1}^{n} B_k \Gamma(\omega'_k, \bullet) \in X_0\)

\[
[A f(\bullet), g(\bullet)] = \left[ A \sum_{j=1}^{n} A_j \Gamma(\omega_j, \bullet), \sum_{k=1}^{m} B_k \Gamma(\omega'_k, \bullet) \right]
\]

\[
A \sum_{j=1}^{n} \sum_{k=1}^{m} A_j \left[ \Gamma(\omega_j, \bullet), \Gamma(\omega'_k, \bullet) \right] B_k^*
\]

\[
= A \sum_{j=1}^{n} \sum_{k=1}^{m} A_j \Gamma(\omega_j, \omega'_k) B_k^*
\]

\[
= A \left[ f(\bullet), g(\bullet) \right]
\]

7.2.1c Adjoint. \([f, g]^* = [g, f]\).

For \(f(\bullet) = \sum_{j=1}^{n} A_j \Gamma(\omega_j, \bullet) \in X_0\) and \(g(\bullet) = \sum_{k=1}^{n} B_k \Gamma(\omega'_k, \bullet) \in X_0\)

\[
[f(\bullet), g(\bullet)]^* = \left[ \sum_{j=1}^{n} A_j \Gamma(\omega_j, \bullet), \sum_{k=1}^{m} B_k \Gamma(\omega'_k, \bullet) \right]^*
\]

\[
= \left[ \sum_{j=1}^{n} A_j \Gamma(\omega_j, \omega'_k) B_k^* \right]^*
\]

\[
= \sum_{j=1}^{n} \sum_{k=1}^{m} B_k^* \Gamma(\omega_j, \omega'_k) A_j^*
\]
\[ n \times m = Z2X \]

\[ \langle \phi(\cdot), f(\cdot) \rangle. \]

7.2.1d Positive Definite Property: \( f = 0 \). [\( f, f \geq 0 \).]

We see that \( [\cdot, \cdot]_{x_0} \) satisfies all the properties of a gramian except that \( \langle f, f \rangle_{x_0} = 0 \) implies \( f = 0 \) (i.e., \( f(\omega) = 0 \) for \( \omega \in \Omega \)).

Define \( \| f \|_{x_0} = \left[ \langle f, f \rangle_{x_0} \right]^{\frac{1}{2}} \) (trace norm) for \( f \in X_0 \), then by the reproducing property we have:

\[ \langle f(\cdot), \Gamma(\omega, \cdot) \rangle_{x_0} = \left[ \sum_{j=1}^{n} A_j \Gamma(\omega_j, \cdot), \Gamma(\omega, \cdot) \right] = \sum_{j=1}^{n} A_j \Gamma(\omega_j, \omega) = f(\omega). \]

Suppose that \( f \in X_0 \) and \( \langle f, f \rangle_{x_0} = 0 \), then the reproducing property \( f(\omega) = \langle f(\cdot), \Gamma(\cdot, \omega) \rangle_{x_0} \)

\[ \Rightarrow \| f(\omega) \|_{x_0} \leq \| f \|_{x_0} \| \Gamma(\omega, \cdot) \|_{x_0} = 0 \]

(Cauchy-Schwartz inequality)

(or Cauchy-Buniakowsky inequality [ Ist] pg. 50).

for every \( \omega \in \Omega \). That is, \( f = 0 \) and hence \( [\cdot, \cdot]_{x_0} \) defines a gramian in \( X_0 \). Thus, \( X_0 \) is a pre-Hilbert \( M_p(C) \)-module, in fact, it is a normal pre-Hilbert \( M_p(C) \)-module (i.e., the trace norm) ([Kak] pg. 37-38).
7.2.2 Reproducing Kernel Hilbert $M_p(C)$-Module: Functional Completion

7.2.2a Definition: Functional Completion. Let be $X_0$ a normal pre-Hilbert $M_p(C)$-module such that,

$$X_0 = \left\{ \sum_{j=1}^{n} A_j \Gamma(\omega_j, \cdot) : A_j \in M_p(C), \ \omega_j \in \Omega, \ 1 \leq j \leq n, \ n = 1, 2, \ldots \right\}$$

consisting of $T(C)$-valued functions on $\Omega$ ($T(C)$ is the set of all trace class operators on $C$). Then the functional completion of $X_0$ is to obtain a normal Hilbert module $X_\tau$ by adding $T(C)$-valued functions on $\Omega$ in such a way that the value of a function $f \in X_\tau$ at $\omega \in \Omega$ depends continuously on $f \in X_\tau$ and that $X_0$ is a dense sub-module of $X_\tau$. In this case the resulting space $X_\tau$ is called the functional completion of $X_0$ ([Kak] pg.11, 37).

7.2.2b The Limit. Let $\{f_n\}_{n=1}^{\infty} \subseteq X_0$ be a Cauchy sequence. It follows from the reproducing property that

$$\| f_n(\omega) - f_m(\omega) \|_\tau = \| \left[ f_n(\cdot), \Gamma(\omega, \cdot) \right]_{X_0} - \left[ f_m(\cdot), \Gamma(\omega, \cdot) \right]_{X_0} \|_\tau$$

$$= \| \left[ f_n(\cdot) - f_m(\cdot), \Gamma(\omega, \cdot) \right]_{X_0} \|_\tau$$

$$\leq \| f_n - f_m \|_{X_0} \| \Gamma(\omega, \cdot) \|_{X_0} \rightarrow 0$$

as $n, m \rightarrow \infty$ for $\omega \in \Omega$. (Cauchy)
Hence there exists a $T(C^p)$-valued function $f$ (the limit) on $\Omega$ such that $\| f_n(\omega) - f(\omega) \|_t \to 0$ for $\omega \in \Omega$. Denote by $X_r$ the set of all functions $f$ obtained in this way.

Thus, $X_0 \subseteq X_r$ (Kak) pg. 38).

7.2.2c Well Defined. Define for $f, g \in X_r$, $[f, g]_{X_r} = \lim_{n \to \infty} [f_n, g_n]_{X_0}$ (in the trace norm), where $\{f_n\}_{n=1}^\infty$, $\{g_n\}_{n=1}^\infty \subseteq X_0$ determine $f$ and $g$. If $\{f'_n\}_{n=1}^\infty$, $\{g'_n\}_{n=1}^\infty \subseteq X_0$ are also determining sequences of $f$ & $g$, then

(1) $\| f'_n(\omega) - f_n(\omega) \|_{X_r}, \| g'_n(\omega) - g_n(\omega) \|_{X_r} \to 0$ for $\omega \in \Omega$. 

(Since $\{f'_n - f_n\}_{n=1}^\infty$ and $\{g'_n - g_n\}_{n=1}^\infty$ are Cauchy in $X_0$)

(2) And the limit functions are zero.

We conclude that $\| f'_n - f_n \|_{X_0}, \| g'_n - g_n \|_{X_0} \to 0$.

Hence

\[
\| [f'_n, g'_n]_{X_0} - [f_n, g_n]_{X_0} \|_{X_r} \leq \| [f'_n - f_n, g'_n - g_n]_{X_0} \|_{X_r} + \| [f'_n, g'_n - g_n]_{X_0} \|_{X_r} + \| [f'_n, g'_n - g_n]_{X_0} \|_{X_r} \to 0.
\]

(Cauchy)

Thus, $[f, g]_{X_r}$ is independent of the choice of determining sequences of $f$ and $g$, and is well defined.

Now it is easily seen that $[\bullet, \bullet]_{X_r}$ is a gramian on $X_r$. 

79
Let \( f \in X_R \) be arbitrary \( \Rightarrow \| f \|_{x_o} = \| f \|_{x_R} \) for \( f \in X_0 \).

That is,

**If** \( \{ f_n \}_{n=1}^{\infty} \subseteq X_0 \) is a determining sequence of \( f \in X_R \),

Then \( \lim_{n \to \infty} \| f_n - f \|_{x_R} = \lim_{n \to \infty} \| f_n - f_n \|_{x_0} = 0 \)

Furthermore, to see that \( X_R \) is complete,

\[
\begin{cases}
X_0 \text{ is dense in } X_R \\
\{ f_n \}_{n=1}^{\infty} \subseteq X_R \text{ (Cauchy seq.)}
\end{cases} \Rightarrow \begin{cases}
\exists \{ f'_n \}_{n=1}^{\infty} \subseteq X_0 \text{ (Cauchy seq.)} \\
s. t. \| f'_n - f_n \|_{x_R} \to 0
\end{cases}
\]

This sequence determines the element \( f \in X_R \) such that

\[
\| f'_n(\omega) - f_n(\omega) \|_{x} \to 0 \text{ for } \omega \in \Omega, \text{ and } \| f'_n - f \|_{x_R} \to 0.
\]

Therefore \( \| f_n - f \|_{x_R} \to 0 \) ([Kak] pg. 38-39).

7.2.2d **Uniqueness of** \( X_R \). Let \( X_k \) be a normal Hilbert \( M_p(C) \)-module with gramian \([\cdot, \cdot]\) which consists of \( T(C^p) \)-valued functions on \( \Omega \) and admits \( \Gamma \) as a r.k. Since \( \Gamma(\omega, \cdot) \in X_k \) for \( \omega \in \Omega \) by definition, we have \( X_0 \subseteq X_k \)

Moreover, it holds that

\[
[f, g]_{x_0} = [f, g] \text{ for } g, g \in X_0. \quad (A)
\]

Furthermore, \( X_0 \) is dense in \( X_k \) since \( \{ \Gamma(\omega, \cdot) : \omega \in \Omega \} \)

is complete in \( X_k \). Thus we must have \( X_R = X_k \). For any \( f, g \in X_k \) we can choose sequences \( \{ f_n \}_{n=1}^{\infty}, \{ g_n \}_{n=1}^{\infty} \subseteq X_0 \) which
converge to \( f \) and \( g \) in \( X_k \), respectively. Then we have by (A)

\[
[f, g] = \lim_{n \to \infty} [f_n, g_n] = \lim_{n \to \infty} [f_n, g_n] = [f, g]_{X_k}.
\]

Therefore \( X_k \) and \( X_r \) are identical Hilbert \( M_p(C) \)-modules ([Kak] pg.39). The reproducing kernel pre-Hilbert \( M_p(C) \)-module has been embedded to form functional completion, thus a reproducing kernel Hilbert \( M_p(C) \)-module.
APPENDIX A

THE SIX QUESTIONS
Let $C^n = \{(x_1, x_2, \ldots, x_n)^T: x_k \in \mathbb{C}, 1 \leq k \leq n\}$. For $n$-dimensional complex Euclidean space $x = (x_1, x_2, \ldots, x_n)^T, y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{C}^n, ("^T" \text{ for transpose})$

$$
\begin{aligned}
\langle x, y \rangle &= \sum_{k=1}^{n} x_k \overline{y_k} : \text{the inner product} \\
\| x \| &= \langle x, x \rangle^{\frac{1}{2}} : \text{the norm}
\end{aligned}
$$

Question One

Show that $\mathbb{C}^n$ is a Hilbert space with inner product given above.

Question Two

For $x, y \in \mathbb{C}^n$ show the following

Part I

$\langle x, y \rangle \geq 0 \iff \| x + \alpha y \| \leq \| x + y \|$, for $\forall \alpha \in \mathbb{C}$ s. t. $|\alpha| = 1$

Part II

$\langle x, y \rangle = 0 \iff \| x + \alpha y \| \leq \| x + y \|$

(i.e. $x, y$ are orthogonal).
Question Three

Collect known equivalence conditions for orthogonality for two vectors in a Hilbert space (or in \( \mathbb{C}^n \)).

Introduction to Question Four

Let \( M_n = \) the set of all \( n \times n \) matrices with complex entries. \( \mathbb{C}^n \) becomes a left \( M_n(\mathbb{C}) \) module by defining the module action "\( \cdot \)" by

\[
A \cdot x^T = Ax^T = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},
\]

Therefore, \( \mathbb{M}_n(\mathbb{C}) \times \mathbb{C}^n \to \mathbb{C}^n \)

In this sense \( \mathbb{C}^n \) is a Hilbert \( M_n(\mathbb{C}) \)-module, where the \( M_n(\mathbb{C}) \)-valued inner product \( [\cdot, \cdot] : \mathbb{C}^n \times \mathbb{C}^n \to M_n(\mathbb{C}) \) is defined by

\[
[x, y] = x^T \begin{bmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_n \end{bmatrix} = \begin{bmatrix} x_1 \overline{y}_1 & \cdots & x_1 \overline{y}_n \\ \vdots & \ddots & \vdots \\ x_n \overline{y}_1 & \cdots & x_n \overline{y}_n \end{bmatrix} \in M_n(\mathbb{C})
\]

also called the Gramian of \( x \) and \( y \).

\( A \in M_n(\mathbb{C}) \): positive \( \iff (Ax, x) \geq 0, \forall x \in \mathbb{C}^n, x \neq 0 \)

\( \{A \geq 0\} \) (nonnegative) \( \{ (Ax, x) \geq 0, \forall x \in \mathbb{C}^n \} \)
Question Four

Check the following properties: for \( x, y, z \in \mathbb{C}^n \)

1. \([x, x] \geq 0\)
   \[ [x, x] = 0 \iff x = 0 \]

2. \([x + y, z] = [x, z] + [y, z]\)

3. \([Ax, y] = A [x, y]\) for \(A \in \mathbb{M}_n(\mathbb{C})\)

4. \([y, x] = [x, y]^\ast\) \(\text{ (the adjoint matrix)}\)

5. \(\text{Tr}[x, y] = (x, y)\) \(\text{(Tr\{\cdot\} = the trace of \{\cdot\})}\)

Question Five

Find the necessary and sufficient condition for \([x, y] \geq 0\)

(in analogy to Question 2, Part I and Part II).

Introduction to Question Six

Let \(\Omega\) be a set and \(\gamma: \Omega \times \Omega \to \mathbb{C}\).

\(\gamma\) is positive definite if \(\sum_{j,k} \alpha_j \overline{\alpha_k} \gamma(\omega_j, \omega_k) \geq 0\)

for \(\forall \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}, \forall \{\omega_1, \ldots, \omega_n\} \subset \Omega\) \(\forall n \geq 1\)

Let \(X\) be a set of functions on \(\Omega\)

\(\gamma: \Omega \times \Omega \to \mathbb{C}\), positive definite \(\Rightarrow X_\gamma\) (RKHS) of \(\gamma\)

\(\left\{\begin{array}{l}
X: \text{a Hilbert space with an inner product \((\cdot, \cdot)\)} \\
1) \gamma(\cdot, \omega) \in X \text{ for } \forall \omega \in \Omega \\
2) \mathcal{f}(\omega) = (\mathcal{f}(\cdot), \gamma(\cdot, \omega)) \text{ for } \forall \mathcal{f} \in X, \forall \omega \in \Omega
\end{array}\right\} \)
It is known that for every positive definite $\gamma: \Omega \times \Omega \to C$, there exists a Reproducing Kernel Hilbert Space of $\gamma$.

Now $\Gamma: \Omega \times \Omega \to M_n(C)$, where $\Gamma$ is positive definite.

Now $\Gamma: \Omega \times \Omega \to M_n(C)$, where $\Gamma$ is positive definite.

Question Six

Construct a reproducing kernel Hilbert $M_n(C)$-module $X_\Gamma$ from $\Gamma: \Omega \times \Omega \to M_n(C)$, positive definite in analogy to RKHS.

Let $X$ be the set of functions on $\Omega$ and $\Gamma: \Omega \times \Omega \to M_n(C)$, positive definite $\Rightarrow X$ is a RKHS $M_n(C)$-module of $\Gamma$.

$$X: \text{a Hilbert } M_n(C) \text{-module with a Gramian } [\bullet, \bullet] \text{ s.t.}$$

\[\begin{align*}
(1) & \Gamma(\bullet, \omega) \in X \text{ for } \forall \omega \in \Omega \\
(2) & x(\omega) = [x(\bullet), \Gamma(\bullet, \omega)] \text{ for } \forall x \in X, \forall \omega \in \Omega
\end{align*}\]

Let $\Gamma: \Omega \times \Omega \to M_n(C)$, positive definite.

$\Rightarrow \gamma: \Omega \times \Omega \to C$, positive definite where $\gamma(\bullet, \bullet) = \text{Tr}(\Gamma(\bullet, \bullet))$.

$\Rightarrow \exists X_\gamma$: the RKHS of $\gamma$. 

86
APPENDIX B

PROOF OF THEOREMS
Chapter 1, Page 15

**Hom(V,W)** is a Vector Space

There is an additive and multiplicative structure.

First, the additive structure:

1). **Closure**: \( f_1, f_2 \in \text{Hom}(V,W) \), \( f_1 + f_2 \in \text{Hom}(V,W) \)

\( f_1, f_2 \) are both vector space homomorphisms of \( V \) into \( W \).

By definition we must have:

\[
\text{for } f_1 \in \text{Hom}(V,W) \quad (x_1 + x_2)f_1 = x_1f_1 + x_2f_1 \\
\text{and} \quad (\alpha x_1)f_1 = \alpha(x_1f_1)
\]

for all \( x_1, x_2 \in V \) and all \( \alpha \in F \). (The same condition also holds for \( f_2 \)). Therefore, for \( f_1 + f_2 \), if \( x_1, x_2 \in V \) and since \( (x_1 + x_2)f_1 = x_1f_1 + x_2f_1 \) and \( (x_1 + x_2)f_2 = x_1f_2 + x_2f_2 \) and addition in \( W \) is commutative, we get

a) **Additive Structure**

\[
(x_1 + x_2)(f_1 + f_2) = (x_1 + x_2)f_1 + (x_1 + x_2)f_2 = x_1f_1 + x_2f_1 + x_1f_2 + x_2f_2 \\
= x_1f_1 + x_1f_2 + x_2f_1 + x_2f_2 = x_1(f_1 + f_2) + x_2(f_1 + f_2)
\]

and

b) **Multiplicatative Structure**

\[
\alpha(x_1 + x_2)(f_1 + f_2) = (\alpha x_1 + \alpha x_2)(f_1 + f_2) = (\alpha x_1 + \alpha x_2)f_1 + (\alpha x_1 + \alpha x_2)f_2 \\
= \alpha x_1f_1 + \alpha x_1f_2 + \alpha x_2f_1 + \alpha x_2f_2 = \alpha x_1(f_1 + f_2) + \alpha x_2(f_1 + f_2).
\]
The Zero Map: Let \( 0 \) be that homomorphism of \( V \) into \( W \) that sends every element of \( V \) onto the zero element of \( W \).

Inverse: For \( f \in \text{Hom}(V,W) \) let \(-f\) be defined by \( x(-f) = -x(f) \).

Commutative: From the closure property above we can conclude that \( \text{Hom}(V,W) \) is an Abelian group under addition.

Now, we check the multiplicative structure: Let \( f \in \text{Hom}(V,W) \), for \( \lambda, \alpha \in F \) and \( x, y \in V \) we have

\[
(\lambda f)(x + y) = \lambda f(x + y) \\
= \lambda(f(x) + f(y)) \\
= \lambda f(x) + \lambda f(y) \\
= (\lambda f)(x) + (\lambda f)(y).
\]

And

\[
(\alpha f)(\alpha x) = \lambda f(\alpha x) \\
= \lambda \alpha f(x) \\
= \alpha \lambda f(x) \\
= \alpha(\lambda f)(x).
\]

\( \text{Hom}(V,W) \) is a vector space over \( F \) ([Her] pg. 185, [War] pg. 281).
Theorem. If \( V \) and \( W \) are of dimension \( m \) and \( n \), respectively, over a field \( F \) (i.e., or \( V \) and \( W \) are \( F \)-vector spaces), then \( \text{Hom}(V,W) \) is of dimension \( mn \) over \( F \).

Proof. We shall prove the theorem by explicitly exhibiting a basis of \( \text{Hom}(V,W) \) over \( F \) consisting of \( mn \) elements such that

\[
x_1, x_2, ..., x_m \text{ is a basis of } V \text{ over } F \text{ and } y_1, y_2, ..., y_n \text{ is a basis of } W \text{ over } F.
\]

If \( x \in V \) where \( \alpha_1, ..., \alpha_n \) are uniquely defined elements of \( F \);
then \( x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \).

Define \( f_{ij}: V \to W \) by \( xf_{ij} = \lambda_i y_j \).

From the viewpoint of the bases we are letting

\[
x_{k}f_{ij} = \begin{cases} 
0 & \text{if } i \neq k \\
y_j & \text{if } i = k 
\end{cases}
\]

Since \( j \) can be any of \( 1, 2, ..., m \) and \( k \) any of \( 1, 2, ..., n \) there are \( mn \) such \( f_{ij} \)'s. The next claim is that these \( mn \) elements constitute a basis of \( \text{Hom}(V,W) \) over \( F \).

We must show: 1) it spans the space and,

2) it is linearly independent.
Let $f_2 \in \text{Hom}(V,W) \rightarrow x_1 f_2 \in W$

Since any element in $W$ (F-vector space) is a linear combination over $F$ of $y_1, y_2, ..., y_n$ for some $\lambda_{11}, \lambda_{12}, \ldots$, $\lambda_{ln} \in F$ is given as

$$x_1 f_2 = \lambda_{11} y_1 + \lambda_{12} y_2 + \ldots + \lambda_{ln} y_n,$$

In fact, $x_k f_2 = \lambda_{k1} y_1 + \lambda_{k2} y_2 + \ldots + \lambda_{kn} y_n$ for $k = 1, \ldots, m$.

Consider $f_0 = \lambda_{11} f_{11} + \lambda_{12} f_{12} + \ldots + \lambda_{ln} f_{ln} + \lambda_{21} f_{21} + \ldots + \lambda_{2n} f_{2n}$

$$+ \ldots + \lambda_{j1} f_{j1} + \ldots + \lambda_{jn} f_{jn} + \lambda_{m1} f_{m1} + \ldots + \lambda_{mn} f_{mn}. $$

Then

$$x_k f_0 = x_k (\lambda_{11} f_{11} + \ldots + \lambda_{mn} f_{mn})$$

$$= \lambda_{11} (x_k f_{11}) + \ldots + \lambda_{m1} (x_k f_{m1}) + \ldots + \lambda_{mn} (x_k f_{mn}).$$

Since $x_k f_{ij} = \begin{cases} 0 & \text{if } i \neq k \\ y_j & \text{if } i = k \end{cases}$ the sum above reduces to

$$\Rightarrow x_k f_0 = \lambda_{k1} y_1 + \lambda_{k2} y_2 + \ldots + \lambda_{kn} y_n = x_k f_2$$

Recall $x_k f_2 = \lambda_{k1} y_1 + \lambda_{k2} y_2 + \ldots + \lambda_{kn} y_n$ for $k = 1, \ldots, m$.

$$\Rightarrow x_k f_0 = x_k f_2.$$

Thus the homomorphism $f_0$ and $f_2$ agree on the basis of $V$; that is, they are both linear combinations of $x_1 f_{jk}$. We have shown that the mn elements $f_{11}, f_{12}, \ldots, f_{ln}, \ldots, f_{m1}, \ldots, f_{mn}$ span $\text{Hom}(V,W)$ over $F$. 

91
Linear Independence

Suppose that
\[\beta_{11}f_{11} + \beta_{12}f_{12} + \cdots + \beta_{1n}f_{1n} + \cdots + \beta_{j1}f_{j1} + \cdots + \beta_{jn}f_{jn} + \cdots + \beta_{ml}f_{ml} + \cdots + \beta_{mn}f_{mn} = 0 \text{ with } \beta_{jk} \text{ all in } F.\]

Then applying this to \(x_k\) we get
\[0 = x_k(\beta_{11}f_{11} + \cdots + \beta_{jk}f_{jk} + \cdots + \beta_{mn}f_{mn}) \]
\[= \beta_{k1}y_1 + \beta_{k2}y_2 + \cdots + \beta_{kn}y_n \text{ since } x_kf_{ij} = \begin{cases} 0 & \text{if } i \neq k \\ y_j & \text{if } i = k \end{cases}.\]

However, \(y_1, y_2, \ldots, y_n\) are linearly independent over \(F\), hence they form a basis of \(\text{Hom}(V,W)\) over \(F\) ([Her] pg. 186).

Chapter 1, Page 17

Corollary 2: If \(\dim_F V = n\) (over the field \(F\)), then \(\dim_F \text{Hom}(V,F) = n\) ([Her] pg. 186).

Proof. A vector space \(V\) over a field \(F\) or a \(F\)-vector space is of dimension 1 over \(F\), so \(\text{Hom}(V,F) = 1, n = n\) ([Her] pg. 186).

Chapter 1, Page 21

Theorem: Let \(X\) be a linear space and let \(X^f\) (or \(\hat{X}\)) denote the set of all linear functionals on \(X\). The space \(X^f\) with vector addition and multiplication of vectors by scalars is a vector space over \(F, (X^f, +, \cdot)\).
ADDITION AND SCALAR MULTIPLICATION

Now let \( f_1, f_2 \in X^f \) and let \( \alpha \in F \). Let us define \( f_1 + f_2 \) and \( \alpha f \) by

\( (f_1 + f_2)(x) = \langle x, x'_1 + x'_2 \rangle = \langle x, x'_1 \rangle + \langle x, x'_2 \rangle = f_1(x) + f_2(x), \) and

1) Additive Structure

2) Multiplicative Structure

\[ f(\alpha x) = \langle \alpha x, x' \rangle = \alpha \langle x, x' \rangle = \alpha f(x) \] (\([\text{Mich}] \) pg110).

Chapter 1, Page 24

Theorem: A bilinear function \( f \) on a complex vector space \( X \) is symmetric if and only if \( \hat{f} \) is real (i.e., \( \hat{f}(x) \) is real for all \( x \in X \)).

Proof. Suppose that \( f \) is

Symmetric: \( f(x, y) = \overline{f(y, x)} \) for all \( x, y \in X \).

Setting \( x = y \) we obtain \( \hat{f}(x) = f(x, x) = \overline{f(x, x)} = \overline{\hat{f}(x)} \)

for all \( x \in X \). But this implies that \( \hat{f} \) is real.
Conversely, if \( f(x) \) is real for all \( x \in X \), then for

\[
g(x, y) = \overline{f(y, x)} \Rightarrow \overline{g(x)} = \overline{f(x, x)} = \bar{f}(x).
\]

Since \( \bar{f} = \bar{g} \) it follows that \( f = g \), and thus

\[
f(x, y) = g(x, y) = \overline{f(y, x)} \quad ([\text{Mich}] \text{ pg. 117})
\]

Chapter 2, Page 36

2.1.2a Cauchy-Schwartz Inequality. Before we prove the next theorem it is necessary to prove the Cauchy-Schwartz inequality.

Theorem: In a Hilbert space \( H \), the norm of the inner product \( \langle x, y \rangle \) \( \leq \| x \| \| y \| \) \( x, y \in H \)

where \( \| x \| := \langle x, x \rangle^{1/2} \) and \( \| y \| := \langle y, y \rangle^{1/2} \).

Proof. If \( x \) or \( y \) is the zero vector, then obviously the equality holds. That is for \( x \) equal the zero vector,

\[
\| 0, y \| = 0 \quad \| y \| \quad x, y \in H
\]

If \( x \) or \( y \) does not equal the zero vector then for any scalar \( \lambda \) we have

\[
0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle + \langle y, y \rangle - \langle \lambda y, x \rangle - \langle x, \lambda y \rangle
\]

\[
= \langle x, x \rangle + \frac{1}{2} \langle \lambda y, x \rangle + \frac{1}{2} \langle x, \lambda y \rangle - [(\lambda y, x) + \overline{(\lambda y, x)}]
\]

94
(note: $a + bi + a - bi = 2a$, just the real part)

$$= (x, x) + \lambda \int (y, y) - 2\text{Re}(\lambda(y, x))$$

For $\lambda = \frac{(x, y)}{(y, y)}$ we have

$$= (x, x) + \frac{|(x, y)|^2}{(y, y)} (y, y) - 2\text{Re}\left(\frac{(x, y)}{(y, y)} (y, x)\right)$$

$$= (x, x) + \frac{|(x, y)|^2}{(y, y)} (y, y) - 2\text{Re}\frac{(y, x)}{(y, y)} (y, x)$$

$$= (x, x) + \frac{|(x, y)|^2}{(y, y)} - 2\text{Re}\frac{|(x, y)|^2}{(y, y)}$$

$$= (x, x) - \frac{|(x, y)|^2}{(y, y)} \text{, so}$$

$$0 \leq (x, x) - \frac{|(x, y)|^2}{(y, y)}$$

and by rearranging the terms the Cauchy-Schwarz inequality is obtained.

$$\frac{|(x, y)|^2}{(y, y)} \leq (x, x) \Rightarrow |(x, y)|^2 \leq (x, x)(y, y)$$

$$\Rightarrow |(x, y)|^2 \leq \|x\| \|y\|^2$$

$$\Rightarrow |(x, y)| \leq \|x\| \|y\|.$$
But why did we choose $\lambda = \frac{(x, y)}{(y, y)}$? The motivation for the proof above is the solution of the following optimization problem. Choose $\lambda$ for fixed $x, y \in H$ in the real scalar product space $H$ so that the distance $\|x - \lambda y\|$ is minimal (the closest approximation to $y$; see figure below). Solution:

$$\|x - \lambda y\|^2 = (x - \lambda y, x - \lambda y)$$

$$= (x, x) - 2\lambda(x, y) + \lambda^2(y, y)$$

since $H$ is a real scalar product space.

Figure. Closest Approximation of Vector $x$ to vector $y$.

Hence

$$\frac{d}{d\lambda}[(x - \lambda y, x - \lambda y) = \frac{d}{d\lambda} [(x, x) - 2\lambda(x, y) + \lambda^2(y, y)]$$

$$= 2\lambda(y, y) - 2(x, y)$$
and

\[ \frac{d}{d\lambda}(x - \lambda y, x - \lambda y) = 0 \Rightarrow 2\lambda(y, y) - 2(x, y) = 0\]

\[ \Rightarrow 2\lambda(y, y) = 2(x, y) \]

\[ \Rightarrow \lambda(y, y) = (x, y) \]

\[ \Rightarrow \lambda = \frac{(x, y)}{(y, y)} \]

([Mat] pg. 46).

Chapter 2, Page 36

Theorem. If \( x_n \to x \) and \( y_n \to y \), then \( (x_n, y_n) \to (x, y) \).

Proof.

\[ |(x_n, y_n) - (x, y)| = |(x_n, y_n) - (x, y) + (x, y) - (x, y)| \]

(A inequality) \[ \leq |(x_n, y_n) - (x, y)| + |(x_n, y) - (x, y)| \]

\[ = |(x_n, y_n - y)| + |(x_n - x, y)| \]

(Cauchy-Schwartz) \[ \leq \| x_n \| \| y_n - x \| + \| x_n - x \| \| y \| \]

\[ \leq K (\| y_n - x \| + \| x_n - x \|) \to 0, \]

so that \( (x_n, y_n) \to (x, y) \) as \( n \to \infty \), where \( K \) is a common upper bound for \( \| y \| \) of the Cauchy sequence \( \{x_n\} \) and \( \{y_n\} \) ([Mat] pg. 47, 48).
6.4 Theorems Related by a Similarity Transformation

1). Their Determinants are Equal.

Theorem. If $U^{-1}AU = B$, then $\det(A) = \det(B)$.

Proof. Since $\det(AB) = \det(A)\det(B)$ we have

\[
\det(B) = \det(U^{-1}AU)
\]
\[
= \det(U^{-1})\det(AU)
\]
\[
= \det(U^{-1})\det(A)\det(U)
\]
\[
= \det(U^{-1})\det(U)\det(A)
\]
\[
= \det(U^{-1}U)\det(A)
\]
\[
= \det(A).
\]

2). Their Eigenvalues are Equal.

Theorem. If $U^{-1}AU = B$, then the eigenvalues of $A$ and $B$ are identical.

Proof. Since $(B - \lambda E) = (U^{-1}AU - \lambda E)$

\[
= U^{-1}(A - \lambda E)U
\]

taking the determinant of both sides we get

\[
\det(B - \lambda E) = \det(U^{-1})\det(A - \lambda E)\det(U)
\]
\[
= \det(U^{-1}U)\det(A - \lambda E)
\]
\[
= \det(A - \lambda E).
\]
The roots of \( \det(A - \lambda E) = 0 \) and \( \det(B - \lambda E) = 0 \) must be identical, since the equations are identical.

3). Theorem. If \( A' = U^{-1}AU \), \( B' = U^{-1}BU \), \( C' = U^{-1}CU \) etc., then any relationship between \( A \), \( B \), \( C \), etc. is also satisfied by \( A', B', C' \) etc.

Proof. Consider, as an example,

\[ D = ABC \]

Then,

\[ D' = U^{-1}DU \]

\[ D' = U^{-1}ABCU \]

\[ = U^{-1}A \, U^{-1}B \, U U^{-1}CU \]

\[ = A'B'C'. \]

4). Theorem. If \( A \) and \( B \) are two matrices which can be diagonalized by the same matrix \( U \) then \( A \) and \( B \) commute.

Proof. Since,

\[ U^{-1}AU = D_a \]

\[ U^{-1}BU = D_b \]

Where \( D_a \) and \( D_b \) are diagonal matrices, then

\[ U^{-1}(AB)U = U^{-1}(A \, UU^{-1}B)U \]

\[ = D_a D_b \]

Chapter 6, Page 67-68

\[ = D_b D_a \quad \text{(diagonal matrices)} \]

\[ = (U^{-1}AU)(U^{-1}BU) \]
and therefore AB = BA and A and B commute.

5). Theorem. If X is the matrix formed from the eigenvectors of A, then $X^{-1}AX$ is a diagonal matrix A composed of the eigenvalues of A.

Proof. A is the diagonal matrix composed of the eigenvalues of A.

\[ AX =XA \]

\[ X^{-1}AX = X^{-1}XA \]

\[ X^{-1}AX = A. \]

If A is Hermitian (A = $A^H$ (or $X^*$) where H stands for the conjugate transpose), then X will be unitary and the transformation will be unitary. If A is symmetric, X and the transformation will be orthogonal.

Chapter 6, Page 67-68


Proof. Since $X = U^{-1}AU$, where A and U are unitary, we have:

\[ X^{-1} = (AU)^{-1}U \quad \text{and} \quad X^T = (\overline{AU})^T(U^{-1})^T \quad \left(U^T = U^{-1}, \quad \overline{A}^T = A^{-1}\right) \]
\[ \begin{align*}
  &= U^{-1}A^{-1}U \\
  &= U^T\overline{A^T}(U^{-1})^T \\
  &= U^{-1}A^{-1}U \\
  &= U^T A^T (U^{-1})^T
\end{align*} \]

\((\overline{U^T})^T = (U^{-1})^{-1} = U)\]

\[ X^{-1} = X^T \text{ and } X \text{ is unitary.} \]

Likewise, if \( A \) and \( U \) are orthogonal (real as well as unitary) then so is \( X \).

Their Traces are Equal:

7). Theorem. If \( U^T A U = B \), then \( \text{trace}(A) = \text{trace}(B) \)

Proof. \[
\text{Trace}(B) = \sum_i B_{ii} \\
= \sum_i \sum_j \sum_k (U^{-1})_{ik} A_{kj} U_{ji} \\
= \sum_j \sum_k A_{kj} \sum_i U_{ji} (U^{-1})_{ik} \\
= \sum_j \sum_k A_{kj} (UU^{-1})_{ik} \\
= \sum_j \sum_k A_{kj} \delta_{ik} \\
= \sum_k A_{kk} \\
= \text{trace}(A)
\]

(Notation) \( t(A) \) or \( tr(A) \) ([Bis] pg. 65-67).

Chapter 6, Page 67-68

Proposition. \( U(n) \) is a group.

Proof.

1). Closure

101
If $A, B \in M_n(k)$

then \((xA, yB) = \langle xA, yA \rangle = \langle x, y \rangle \rightarrow AB \in M_n(k)\).

2). Identity:

Clearly the identity matrix $I$ is in $U(n)$.

If $A \in U(n)$ we have

\[
\langle e_jA, e_kA \rangle = \langle e_j, e_k \rangle = \delta_{jk} = \begin{cases} 
1 & \text{if } j = k \\
0 & \text{if } j \neq k 
\end{cases}
\]

3). Inverse:

$e_j$ is just the $j^{th}$ row of $A$ and $\langle e_jA, e_kA \rangle$ is just the $jk^{th}$ entry in the product $A\overline{A}^T$.

Thus $A\overline{A}^T = I$.

Now for any $x, y \in C^n$ and $A \in M_n(C)$ we have

\[
\langle x, yA \rangle = \sum_{j=1}^{n} x_j \cdot \overline{(yA)_j} \\
= \sum_{j=1}^{n} x_j \cdot \left( \sum_{k=1}^{n} \overline{y_k} \overline{a_{kj}} \right) \\
= \sum_{j=1}^{n} \left( x_j \cdot \overline{a_{kj}} \right) \sum_{k=1}^{n} \overline{y_k} \\
= \sum_{k=1}^{n} \left( x \cdot \overline{A^T}_k \right) \overline{y_k} \\
= \langle x\overline{A^T}, y \rangle.
\]

So $\langle x\overline{A^T}, y \rangle = \langle x, yA \rangle \rightarrow \langle e_j\overline{A^T}, e_k\overline{A^T} \rangle = \langle e_j, e_k \rangle = \langle e_j, e_k\overline{A^T}A \rangle$
Thus $\overline{A}^T = A^{-1}$, is a right and left inverse of $A$.

Finally $\{xA^{-1}, yA^{-1}\} = \{x\overline{A}^{-1}A, y\overline{A}^{-1}A\} = \{x, y\} \rightarrow A^{-1} \in U(n)$.

$U(n)$ is a group ([MLC] pg. 26).

Chapter 6, Page 71

Question 5 will be answered by applying the next theorem.

Theorem. Let $A \in (\tau_c)$, the trace-class. Then for

$Rt(A) -$ the real part of $t(A)$ and

$Rt(UA) -$ the real part of $t(UA)$, we have

$Rt(A) \geq Rt(UA)$ for all unitary $U$ if and only if $A \geq 0$.

(The above will be inserted in the proof of question five.)

Proving one direction is shown below:

Assume $A \geq 0$.

Then $A = (A^*A)^{1/2} \Rightarrow t(A) = t((A^*A)^{1/2}) = \tau(A)$.

Thus,

$Rt(UA) \leq |t(UA)| \leq \tau(UA) \leq |U|\tau(A) = \tau(A) = t(A) = Rt(A)$

(see [Scha] pg. 43).

Chapter 6, Page 71

Let $A \in M_n(C)$. Then the following conditions are equivalent:

(vi) $A \in U(n)$, (unitary group).
(vii) \[ \langle e_i A, e_j A \rangle = \delta_{ij} \]

(viii) A sends orthonormal bases to orthonormal bases

(ix) The rows of A form an orthonormal basis

(x) The columns of A form an orthonormal basis

(vi) \[ A^T = A^{-1} \]

([Cur] pg 27).

Note below that the arrows emphasizing the variable represents a vector and \( \{ \hat{e}_k \} \) represents the standard ordered basis.

(i) \( \Rightarrow \) (ii)

Here we are assuming that \( A \in U(n) \), which means that A preserves the inner product,

\[ \langle \bar{u} A, \bar{v} A \rangle = \langle \bar{u}, \bar{v} \rangle. \]

We want to show that \[ \langle \hat{e}_i A, \hat{e}_j A \rangle = \delta_{ij}. \]

\[ \langle \hat{e}_i A, \hat{e}_j A \rangle = \langle \hat{e}_i, \hat{e}_j \rangle \] because \( A \in U(n) \)

\[ = \delta_{ij} \] because \( \{ \hat{e}_i \} \) forms an orthonormal basis.

(ii) \( \Rightarrow \) (iii)

We are assuming that \[ \langle \hat{e}_i A, \hat{e}_j A \rangle = \delta_{ij}. \]

We want to prove that if \( \{ \overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_n} \} \) is an orthonormal basis, then \( \{ \overrightarrow{v_1 A}, \overrightarrow{v_2 A}, ..., \overrightarrow{v_n A} \} \) is also an orthonormal basis.
Let \( \vec{v}_i = \sum_{k=1}^{n} a_{jk} \hat{e}_k \) where \( \{ \hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n \} \) is the standard basis.

Then

\[
\{ \vec{v}_i, \vec{v}_j \} = \left( \sum_{k=1}^{n} a_{ik} \hat{e}_k, \sum_{l=1}^{n} a_{lj} \hat{e}_l \right)
\]

\[
= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} (\hat{e}_k, \hat{e}_l) = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} \delta_{k,l}
\]

\[
= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} \quad \text{(note:} \quad (\hat{e}_i, \hat{e}_j) = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} = \delta_{ij})\]

\[= \delta_{ij}\]

We wanted to show that \( \{ \vec{v}_1A, \vec{v}_2A, \ldots, \vec{v}_nA \} \) is also an orthonormal basis.

Consider \( \{ \vec{v}_iA, \vec{v}_jA \} = \left( \sum_{k=1}^{n} a_{ik} \hat{e}_kA, \sum_{l=1}^{n} a_{lj} \hat{e}_lA \right) \)

\[
= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} (\hat{e}_kA, \hat{e}_lA)
\]

\[
= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} \quad (\hat{e}_k, \hat{e}_l) = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} \delta_{k,l}
\]

\[
= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} \overline{a_{jl}} = \delta_{ij}
\]

Thus \( \{ \vec{v}_1A, \vec{v}_2A, \ldots, \vec{v}_nA \} \) forms an orthonormal basis.
We want to show that the rows of $A$ form an orthonormal basis. We are assuming that $A$ sends orthonormal basis to orthonormal basis. Now the rows of $A$ are \{$\hat{e}_i A, \ldots, \hat{e}_n A$\}. Since \{$\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n$\} is an orthonormal basis, (iii) will give us \{$\hat{e}_i A, \ldots, \hat{e}_n A$\} is also an orthonormal basis, and therefore, the rows form an orthonormal basis.

(iv) $\Rightarrow$ (v)

Note that if $A \in U(n)$ then $A^* = \overline{A}^T$, actually shows that the rows of $A$ form an orthonormal basis if and only if $AA^T = I$. Indeed, the rows of $A$ form an orthonormal basis iff $\langle \hat{e}_i A, \hat{e}_j A \rangle = \delta_{i,j}$ iff $\langle \hat{e}_i AA^T, \hat{e}_j AA^T \rangle = \langle \hat{e}_i AA^T, \hat{e}_j \rangle = \delta_{i,j}$. But $\langle \hat{e}_i AA^T, \hat{e}_j \rangle$ is precisely the $(i, j)$ entry of $AA^T$, so the last equivalence is saying that the rows of $A$ are orthonormal iff $AA^T = I$. In a similar fashion we can prove the columns of $A$ form an orthonormal basis iff $\overline{A}^T A = I$. Here, we can combine the appropriate parts of these equivalences to build our proof.
Assume that the rows form an orthonormal basis. Then 
\[ \langle \mathbf{e}_i A, \mathbf{e}_j A \rangle = \delta_{i,j}. \] Hence, \[ \langle \mathbf{e}_i A A^T, \mathbf{e}_j \rangle = \delta_{i,j}, \] that is \[ A A^T = I. \] But then, we also have that \[ A^T A = I. \] Thus 
\[ \langle \mathbf{e}_i A A^T, \mathbf{e}_j \rangle = \delta_{i,j}, \] and therefore \[ \langle \mathbf{e}_i A, \mathbf{e}_j A^T \rangle = \delta_{i,j}, \] which is saying that the columns of \( A \) form an orthonormal basis.

\((v) \Rightarrow (vi)\)

This is now just a repetition of the last part of the argument given in the proof of the previous implication. If the columns of \( A \) are orthonormal, then 
\[ \langle \mathbf{e}_i \mathbf{A}^T, \mathbf{e}_j \mathbf{A}^T \rangle = \delta_{i,j}. \] Thus 
\[ \langle \mathbf{e}_i A A^T, \mathbf{e}_j \rangle = \delta_{i,j}, \] and we conclude that \( \mathbf{A}^T A = I \), and therefore that \( A^{-1} = \mathbf{A}^T \).

\((vi) \Rightarrow (i)\)

If we have that \( A^{-1} = \mathbf{A}^T \), then \( A A^T = I \), and hence, for any two vectors \( u \) and \( v \)

\[ \langle u, v \rangle = \langle u A A^T, v \rangle = \langle u (A^T A), v A \rangle = \langle u A, v A \rangle \]

which shows that \( A \in \text{U}(n) \).
APPENDIX C

INDEX
Adjoint Property, 57, 76-77
Affine Geometry, 3, 4
Algebra, 62, 67
Algebraic
  Conjugate, 21
  Dual, 20, 21
Annihilator, 45
Automorphism, 34
Banach Space, 25, 26
Basis, 16
  Ordered, 16
    Standard Ordered, 17
Bessel's Inequality, 48-49
Bilinear Functional, 21-23, 93
Bilinear Form, 22
Cauchy-Schwartz Inequality, 36, 94-97
Class, 65
Completion, 25, 29, 34, 37
  Functional, 78
Complex
  Conjugate, 32
    Linear space, 23
Conjugate
  Group Elements, 65
    Linear Functional, 21-22
    Symmetry Property, 32
Convergent
  Sequence, 27
    Uniform, 35
Direct Sum, 52
Dimension, 17, 18
Dual
  Basis, 21
    Space, 15
Euclidean Geometry, 3
Field, 7, 8
Fourier Coefficient, 47, 48, 51
Gram-Schmidt Orthogonalization, 49-50
Gramian, 53, 74, 77
Group, 6, 7, 10
  Abelian, 7, 9, 10
Hermitian, 100
Hilbert
  Module, 53
  (normal) Pre-Hilbert, 54, 58
  Space, 29, 36-38, 60
Homomorphism, 13, 14
  Vector Space, 88
Inner product, 24
  Space, 29, 30
  Space Isomorphism, 34
Integral Domain, 9
Kronecker Delta, 17
Linear
  Combination, 15
  Form, 20
  Functional, 18-21
  Independent, 16, 92
  Map, 61
  Operator, 15
  Property, 32, 55-56, 75-76
  Span, 16
  Transformation, 15, 20
Maps, 12, 13
Module, 5, 6, 11, 12, 53
Norm, 25-27
Normed Space, 25
  Linear, 27
Orthogonality, 39, 43
  Properties, 43
Orthogonal
  Basis, 49, 50
Complement, 51-52
Group, 67
Sequence, 45
Set of Vectors, 45
Orthonormal
  Set, 17
  Basis, 17, 49, 50, 104
Positive Definite Property, 31, 54-55, 75
Pre-Hilbert
  (normal) $\mathbb{M}_n(\mathbb{C})$-module, 54, 78
  Space, 29, 30
  Reproducing Kernel Module, 75
Projections, 46
Quadratic Form, 24
Reproducing Kernel Hilbert
  Module, 73, 75, 78
  Space, RKHS, 72, 73
Ring, 6
  Division, 7
  Scalar, 12, 20
Sequence, 27
  Cauchy, 35, 78-79
  Fundamental, 35
Similar Transformations, 63
  Theorems, 64-65, 98-101
Schmidt-Class ($\Theta_C$), 58-59
Span, 91
Trace, 54, 58
Trace-Class ($\Theta_{C}$), 59, 68, 103
  Norm, 59
  Set of, 78
Uniform Convergence, 35
Unit, 66
  Group of, 67
Unitary
  Group, $\mathbb{U}(n)$, 67
Linear Operator, 34
Matrix, 60, 63, 65
Operators, 60
R-module, 12
Representation, 65
Space, 60
Transformation, 65
Vector space, 9, 21, 62
Homomorphism, 88, 89


