# HILBERT SPACE

by

JOHN DAVID PERINE

B. S., Washburn University, 1964

---

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1965

Approved by:

Neal E. Foland
Major Professor

TABLE OF CONTENTS

## INTRODUCTION

The theory of Hilbert space grew out of an infinite dimensional vector space which was first extensively studied by the German mathematician David Hilbert in his work on the theory of integral equations. This space is the classical Hilbert space $l^2$, which is described in Example 7 below. It was abstracted by J. von Neumann to the variety of linear vector space known as Hilbert space.

The infinite-dimensionality of classical Hilbert space has brought about the convention of reserving the term Hilbert space for infinite-dimensional spaces satisfying certain conditions, while finite-dimensional spaces satisfying the same conditions are termed Euclidian spaces. However, most of the definitions and theorems contained in this report apply to either type of space; therefore this convention will not be employed. That is, Euclidian space will be termed finite-dimensional Hilbert space.

In the following pages, the structure of Hilbert space will be developed as a special type of linear vector space, and some of the general properties of this space will be presented.

## VECTOR SPACES

A vector space $\mathcal{V}(\mathcal{F})$ over a field $\mathcal{F}$ is a set of elements x, y, z, ... satisfying the following axions:

(1) $\mathcal{V}$ forms an additive Abelian group, where the operation, denoted "+", is from $\mathcal{V} \times \mathcal{V}$ into $\mathcal{V}$. That is, for $x, y, z \in \mathcal{V}$,

$A_1$ $\quad x + y = y + x$ ,

$A_2$ $\quad x + (y + z) = (x + y) + z$ ,

$A_3$ there is a $\theta \in \mathcal{V}$ such that $x + \theta = x$ for all $x \in \mathcal{V}$,

$A_4$ if $x \in \mathcal{V}$, then there is a $-x \in \mathcal{V}$ such that $x + (-x) = \theta$ .

(2) There is an operation from $\mathcal{F} \times \mathcal{V}$ into $\mathcal{V}$, denoted by juxtaposition and called <u>scalar multiplication</u>, such that for $x,y \in \mathcal{V}$ and $\lambda, \mu \in \mathcal{F}$,

$M_1$ $\lambda(x + y) = \lambda x + \lambda y$ ,

$M_2$ $(\lambda \oplus \mu)x = \lambda x + \mu x$ ,

$M_3$ $\lambda(\mu x) = (\lambda \cdot \mu)x$ ,

$M_4$ $1x = x$ , where 1 is the multiplicative identity of $\mathcal{F}$.

In the above axions, $\oplus$ and $\cdot$ denote addition and multiplication, respectively, in the field $\mathcal{F}$ . A vector space can be defined over any algebraic field $\mathcal{F}$ , and is often described by the particular field. Thus $\mathcal{V}(\mathcal{R})$, where $\mathcal{R}$ is the field of real numbers, is called a real vector space, and $\mathcal{V}(\mathcal{C})$, where $\mathcal{C}$ is the field of complex numbers, is called a complex vector space. However, in the sequel, only complex fields will be used, and the terms vector space and complex vector space will be equivalent. The elements of the field are called <u>scalars</u>. Thus in the vector spaces to be considered here, the term scalar will refer to the complex numbers. By reserving the lower case letters, x,y,z,..., to denote the elements of $\mathcal{V}$, called <u>vectors</u>, and using lower case Greek letters for scalars, no confusion will arise from the use of + for $\oplus$ , and juxtaposition for $\cdot$; therefore this convention will be adopted in the sequel.

EXAMPLE 1. For n a fixed positive integer, let $\mathcal{V}$ be the set of all n-tuples $x = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, where the $\lambda i$ are scalars. If $y = (\mu_1, \mu_2, \ldots, \mu n)$, then define $x = y$ iff $\lambda i = \mu i$ for all $i = 1,2,\ldots,n$. $\lambda i$ is called the i-th component of x. Let $\theta = (0,0,\ldots,0)$ with a total of n components, $-x = (-\lambda_1, -\lambda_2, \ldots, -\lambda n)$, $x + y = (\lambda_1 + \mu_1, \lambda_2 + \mu_2, \ldots, \lambda_n + \mu_n)$, and $\lambda x = (\lambda\lambda_1, \lambda\lambda_2, \ldots, \lambda\lambda n)$. Then $\mathcal{V}$ satisfies axions $A_{1-4}$ and $M_{1-4}$ and hence is a vector space, referred to as the vector space of n-tuples.

EXAMPLE 2. Let $\mathcal{V}$ consist of the set of all scalar-valued functions x which are defined and continuous on a closed interval $[a,b]$, together with the definitions: for all $t \in [a,b]$, $x = y$ iff $x(t) = y(t)$, $\theta$ is the constant function $\theta(t) \equiv 0$, $-x = -x(t)$, $(x + y)(t) = x(t) + y(t)$, and $(\lambda x)(t) = \lambda x(t)$. Since the sum of two continuous functions and the product of two continuous functions result in continuous functions, these definitions are suitable for defining a vector space.

EXAMPLE 3. Let $\mathcal{V}$ consist of the set of all sequences of scalars $x = (\lambda_k)$, $k = 1, 2, \ldots$, all of whose terms are zero from some index m on out. That is, $x \in \mathcal{V}$ implies that there is an integer m such that $\lambda_i = 0$ for $i \geq m$. If $y = (\mu_k)$, define $x = y$ iff $\lambda_k = \mu_k$ for all k, $\theta = (0)$, $-(\lambda_k) = (-\lambda_k)$, and $\lambda(\lambda_k) = (\lambda\lambda_k)$. Again $\mathcal{V}$ with these operations satisfies the axions for a vector space. This space is called the vector space of finitely nonzero sequences.

A few properties of vector spaces which follow from the axions are listed in the theorem below.

THEOREM 1. If $\mathcal{V}$ is a vector space and if $x, y, z \in \mathcal{V}$, then

(1) the equation $x + y = z$ has a unique solution y for every $x, z \in \mathcal{V}$;

(2) if $z + z = z$, then $z = \theta$;

(3) $\lambda\theta = \theta$ for all scalars $\lambda$;

(4) $0x = \theta$ for all $x \in \mathcal{V}$;

(5) if $\lambda x = \theta$, then either $\lambda = 0$ or $x = \theta$ .

Proof:

(1) Let $y = (-x) + z$ . Then $x + y = x + ((-x) + z) = (x + (-x)) + z = \theta + z = z$ . If also $x + y' = z$, then $x + y = x + y'$, from which $y = y'$ by adding $-x$ to each side.

(2) Since $z + \theta = z$, this follows immediately from (1).

(3) $\lambda(\theta) = \lambda(\theta + \theta) = \lambda\theta + \lambda\theta$ , thus by (2), $\lambda\theta = \theta$.

(4) $0\,x = (0 + 0)x = 0\,x + 0\,x$. Thus by (2), $0x = \theta$.

(5) if $\lambda \neq 0$ , then $\lambda^{-1}(\lambda x) = (\lambda^{-1}\lambda)x = 1\,x = x = \theta$ .

Note that $0\,x = (1 + (-1))x = 1\,x + (-1)x = x + (-1)x = \theta$ , hence $(-1)x = -x$ , since the negative of x is unique by (1) above.

Let $\sum_1^n x_1 = x_1 + x_2 + x_3 + \ldots + x_n$, where $x_i \in \mathcal{V}$, a vector space, and i ranges over $1,2,\ldots,n$. The associative law, $A_2$, may be generalized to a sum of n terms, so that any grouping of the summands yields the same sum. The commutative law, $A_1$, can also be generalized, so that in $\sum_1^n x_i$ , any permutation of $1,2,\ldots,n$ gives the same sum. Thus $\sum_1^n x_1$ , is uniquely defined. If for $x \in \mathcal{V}$, $x = \sum_1^n \lambda_i x_i$, where $\lambda_i$ are scalars, then x is said to be a linear combination of the $x_i$ . For example, in the vector space of n-tuples, let $x_i$ be the vector all of whose components are zero except the i-th component, which is one. Then $x = (\lambda_1, \lambda_2, \ldots, \lambda_n) = \sum_1^n \lambda_i x_i$ . If $\mathcal{V}$ is a vector space, a linear subspace of $\mathcal{V}$ is defined to be a subset $\mathcal{N}$ of $\mathcal{V}$ such that (i) $\theta \in \mathcal{N}$, (ii) $x \in \mathcal{N}$ and $y \in \mathcal{N}$ imply that $x + y \in \mathcal{N}$, and (iii) $x \in \mathcal{N}$ and $\lambda$ a scalar imply that $\lambda x \in \mathcal{N}$. It is evident that $\mathcal{N}$ is then a vector space in its own right; axions $A_{1-3}$ and $M_{1-4}$ must be satisfied, while

$x \in \mathcal{N}$ and condition (iii) give $(-1)x = -x \in \mathcal{N}$. Thus $A_4$ is satisfied. An immediate consequence of this definition is that if $\mathcal{N}$ is a linear subspace of $\mathcal{V}$, then every linear combination of vectors in $\mathcal{N}$ belongs to $\mathcal{N}$. Conversely, if $\mathcal{N}$ is a subset of $\mathcal{V}$ and if every linear combination of vectors in $\mathcal{N}$ belongs to $\mathcal{N}$, then $\mathcal{N}$ automatically satisfies conditions (ii) and (iii) of the definition of a linear subspace of $\mathcal{V}$. Since $\Theta = \sum_1^n 0x_i$, where $x_i \in \mathcal{N}$, condition (i) is also satisfied, and $\mathcal{N}$ is therefore a linear subspace of $\mathcal{V}$. Thus if $\mathcal{S}$ is any subset of $\mathcal{V}$, the set which consists of all linear combinations of vectors in $\mathcal{S}$, denoted $[\mathcal{S}]$, is a linear subspace of $\mathcal{V}$, and is said to be the linear subspace <u>generated</u> by $\mathcal{S}$. If a vector x can be expressed as a linear combination of the elements of $\mathcal{S}$, then x is said to be <u>linearly dependent</u> on $\mathcal{S}$. Similarly, a vector x is said to be <u>linearly independent</u> of $\mathcal{S}$ if x cannot be expressed as a linear combination of the vectors in $\mathcal{S}$. If $\mathcal{S}$ is a set of vectors such that no vector in $\mathcal{S}$ is a linear combination of the other vectors in $\mathcal{S}$, then $\mathcal{S}$ is said to be a <u>linearly independent</u> set of vectors. For brevity, the term <u>linear</u> is usually not written. Thus in the vector space $\mathcal{V}$ of 3-tuples (Example 1), $\mathcal{S} = \{(0,0,1),(0,1,0),(1,0,0)\}$ is an independent set. Furthermore, if $x \in \mathcal{V}$, $x = (\lambda_1, \lambda_2, \lambda_3)$, then $x = \sum_1^3 \lambda_i x_i$ where $x_i$ is the element of $\mathcal{S}$ whose i-th component is 1 and all other components zero. Thus $\mathcal{S}$ generates $\mathcal{V}$, and every vector in $\mathcal{V}$ is dependent on $\mathcal{S}$.

A vector space $\mathcal{V}$ is said to be finitely generated if there exists a finite set $x_1, x_2, \ldots, x_n$ of vectors which generates $\mathcal{V}$. A set of vectors $\mathcal{T}$ is a basis for a vector space $\mathcal{V}$ provided (i) $\mathcal{T}$ generates $\mathcal{V}$, and

(ii) $\mathcal{T}$ is independent. It can be shown that every finitely generated vector space has a basis, and that the number n of vectors in any basis is unique for $\mathcal{V}$. This number is called the dimension of $\mathcal{V}$, and $\mathcal{V}$ is said to be finite-dimensional. If $\mathcal{V}$ is not finitely generated, then $\mathcal{V}$ is said to be infinite-dimensional. Thus the vector space of 3-tuples has dimension three. On the other hand, the vector space of finitely nonzero sequences, (Example 3), is infinite-dimensional.

A complex vector space $\mathcal{P}$ is said to be a pre-Hilbert space provided there is an operation from $\mathcal{P} \times \mathcal{P}$ into $\mathcal{C}$, the complex field, denoted by $(x|y)$, there $x, y \in \mathcal{P}$, and such that this operation satisfies the conditions:

$P_1$ $(y|x) = \overline{(x|y)}$, where $\overline{\lambda}$ denotes the conjugate of $\lambda$,

$P_2$ $(x + y|z) = (x|z) + (y|z)$,

$P_3$ $(\lambda x|y) = \lambda(x|y)$, and

$P_4$ $(x|x) > 0$ when $x \neq \theta$.

This operation is called an inner product or a scalar product, the latter leading to a convenient verbalization of $(x|y)$: "x scalar y". Before considering the properties possessed by the inner product, examples of inner products defined on the vector spaces already considered will be given.

EXAMPLE 4. Let $\mathcal{P}$ be the vector space of n-tuples, and $x = (\lambda_1, \ldots, \lambda_n)$, $y = (\mu_1, \ldots, \mu_n)$. Define $(x|y) = \sum_1^n \lambda_i \overline{\mu_i}$. Then $(y|x) = \sum_1^n \mu_i \overline{\lambda_i}$ $= \sum_1^n \overline{\lambda_i \overline{\mu_i}} = \overline{(x|y)}$, thus $P_1$ is satisfied. Conditions $P_2$ ans $P_3$ follow by properties of the summation, and $P_4$ by the identity $\lambda \overline{\lambda} = |\lambda|^2$, $\lambda \in \mathcal{C}$.

Thus $\mathcal{P}$ is a pre-Hilbert space, known as n-dimensional unitary space, and denoted $\mathbb{C}^n$. It should be noted that $\mathbb{C}^1$ is essentially the set of complex numbers considered as a vector space, with $(x|y) = x\bar{y}$ , $x,y \in \mathbb{C}$.

EXAMPLE 5. Let $\mathcal{P}$ be the vector space of continuous functions on the closed interval $[a,b]$ , $a < b$ . Then for $x,y \in \mathcal{P}$, define $(x|y) = \int_a^b x(t)\overline{y(t)}dt$ . It will now be shown that $(x|y)$ thus defined is an inner product.

To show that $P_1$ is satisfied, let $\sum_1^n x(t_i)\overline{y(t_i)}(t_i - t_{i-1})$ be one of the approximating sums of the integral which defines $(x|y)$. Then $(y|x) = \int_a^b \overline{x(t)}y(t)dt$ has the corresponding approximating sum $\sum_1^n \overline{x(t_i)}y(t_i)(t_i - t_{i-1})$. Now $t_i - t_{i-1}$ , being a real number, is its own conjugate. Thus

$$\overline{\left( \sum_1^n \overline{x(t_i)}y(t_i)(t_i - t_{i-1}) \right)} = \sum_1^n x(t_i)\overline{y(t_i)}(t_i - t_{i-1}).$$

Since the defining integral is composed of the limit of such sums, it follows that $\overline{(y|x)} = (x|y)$ , or equivalently, $(y|x) = \overline{(x|y)}$. $P_2$ follows by the linearity of the Riemann integral with respect to its integrand, and $P_3$ from the property that $\int_a^b \lambda f(z)dz = \lambda \int_a^b f(z)dz$ . If $x(t) \not\equiv 0$ on $[a,b]$ , then there is an $\varepsilon > 0$ such that for some subinterval $[a_1, b_1]$, $a_1 < b_1$ , of $[a,b]$ , $t \in [a_1,b_1]$ implies that $|x(t)| > \sqrt{\varepsilon}$ . Thus $\int_a^b |x(t)|^2dt \geq \int_a^b \varepsilon^2 dt = \varepsilon^2(b_1 - a_1) > 0$. This proves that $(x|x) = \int_a^b |x(t)|^2 dt > 0$ if $x \neq \theta$, and $P_4$ is satisfied. Thus the vector space $\mathcal{P}$ forms a pre-Hilbert space.

EXAMPLE 6. Let $\mathcal{P}$ be the vector space of finitely nonzero sequences. If $x = (\lambda_k)$ and $y = (\mu_k)$, define $(x|y) = \sum_1^\infty \lambda_k \overline{\mu_k}$ . Since the terms of this sum are all zero from some term on, this sum is finite and well-defined. The finiteness of the sum establishes $P_{1-3}$ , and $P_4$ follows from the fact that $\lambda_k \overline{\lambda_k} = |\lambda_k|^2 > 0$ for some k when $x = (\lambda_k) \neq \theta$ .

PRE-HILBERT SPACES

Some properties of a pre-Hilbert space implied by the inner product are listed in the following theorem.

THEOREM 2. In a pre-Hilbert space $\mathcal{P}$, if $x,y,z \in \mathcal{P}$, then

(1) $(x|y + z) = (x|y) + (x|z)$ ;

(2) $(x|\lambda y) = \overline{\lambda}(x|y)$ ;

(3) $(\theta|y) = (x|\theta) = 0$ ;

(4) $(x - y|z) = (x|z) - (y|z)$ and $(x|y - z) = (x|y) - (x|z)$ ;

(5) if $(x|z) = (y|z)$ for all z, then $x = y$ .

Proof:

(1) $$\begin{aligned}(x|y + z) &= \overline{(y + z|x)} \text{ by } P_1 ,\\ &= \overline{[(y|x) + (z|x)]} \text{ by } P_2 ,\\ &= \overline{(y|x)} + \overline{(z|x)}\\ &= (x|y) + (x|z) \text{ by } P_1 .\end{aligned}$$

(2) $$\begin{aligned}(x|\lambda y) = \overline{(\lambda y|x)} &= \overline{[\lambda(y|x)]} \text{ by } P_1 \text{ and } P_3,\\ &= \overline{\lambda}\,\overline{(y|x)}\\ &= \overline{\lambda}(x|y) \text{ by } P_1 .\end{aligned}$$

(3) $(\theta|y) = (\theta + \theta|y) = (\theta|y) + (\theta|y)$ , which implies that $(\theta|y) = 0$ . Similarly, $(x|\theta) = (x|\theta + \theta) = (x|\theta) + (x|\theta)$, and $(x|\theta) = 0$ follows.

(4) Consider $(x + \lambda_1 y|z)$ and $(x|y + \lambda_1 z)$, where $\lambda_1 = \overline{\lambda_1} = -1$ . Then apply $P_1$ and (1).

(5) $(x|z) = (y|z)$ for all z implies that $(x - y|z) = (x|z) - (y|z) = 0$ for all z. Thus $(x - y|x - y) = 0$ ; hence $x - y = \theta$ by $P_4$.

One of the special classes of vector spaces is the class of normed vector spaces. This class has a function defined on each vector $x \in \mathcal{V}$ , called the norm of x and denoted $\|x\|$ , which has the properties:

$n_1$ $\|x\|$ is strictly positive, that is $\|x\| \geq 0$ with equality only if $x = \theta$ ,

$n_2$ $\|\lambda x\| = |\lambda| \, \|x\|$ , and

$n_3$ $\|x + y\| \leq \|x\| + \|y\|$ .

It is evident that the definition $\|x\| = \sqrt{(x|x)}$ , where x is a vector in a pre-Hilbert space $\mathcal{P}$ , satisfies $n_1$ and $n_2$ . That this definition also satisfies $n_3$ is shown in Theorem 6. Note that in Example 4, the pre-Hilbert space of n-tuples, $\|x\| = ( \sum_1^n |\lambda_k|^2)^{\frac{1}{2}}$, where $x = (\lambda_1,\ldots,\lambda_n)$ . In Example 5, the pre-Hilbert space of functions continuous on $[a,b]$ , $\|x\| = ( \int_a^b |x(t)|^2 dt)^{\frac{1}{2}}$ , and in Example 6, the pre-Hilbert space of finitely nonzero sequences, if $x = (\lambda_k)$ , then $\|x\| = ( \sum_1^\infty |\lambda_k|^2)^{\frac{1}{2}}$ . The following theorems give some of the basic properties of $\|x\|$ as defined for a pre-Hilbert space $\mathcal{P}$ .

THEOREM 3. (The Parallelogram Law.) If $x,y \in \mathcal{P}$ , then

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 .$$

Proof:

First $\|x + y\|^2 = (x + y|x + y) = (x + y|x) + (x + y|y)$

$$= (x|x) + (y|x) + (x|y) + (y|y).$$

$$= \|x\|^2 + \|y\|^2 + (y|x) + (x|y) .$$

Also $\|x - y\|^2 = (x|x) - (y|x) - (x|y) + (y|y)$

$$= \|x\|^2 + \|y\|^2 - (y|x) - (x|y) .$$

The result follows.

THEOREM 4. (Polarization.) If $x,y \in \mathcal{P}$, then

$$(x|y) = 1/4 \left[\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2\right] .$$

Proof:

Using the first half of the proof of Theorem 3, with y replaced by -y, iy, and -iy successively,

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

$$\|x + iy\|^2 = \|x\|^2 + \|y\|^2 + i(y|x) - i(x|y) , \text{ and}$$

$$\|x - iy\|^2 = \|x\|^2 + \|y\|^2 - i(y|x) + i(x|y) .$$

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

$$i\|x + iy\|^2 = i\|x\|^2 + i\|y\|^2 - (y|x) + (x|y) ,$$

$$-i\|x - iy\|^2 = -i\|x\|^2 - i\|y\|^2 - (y|x) + (x|y) ,$$

But $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + (y|x) + (x|y) .$

By summing the last four identities,

$$4(x|y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 ,$$

as desired.

THEOREM 5. (The Cauchy-Schwarz inequality.) If $x, y \in \mathcal{P}$, then $|(x|y)| \leq \|x\|\,\|y\|$ .

(Since this theorem is a special case of Theorem 9, it will be proved after Theorem 9 is established.)

THEOREM 6. (The Triangle inequality.) If $x,y \in \mathcal{P}$, then

$$\|x + y\| \leq \|x\| + \|y\| \quad .$$

Proof:

Again $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + (x|y) + \overline{(x|y)}$

$$= \|x\|^2 + \|y\|^2 + 2\,\mathcal{R}_e[(x|y)]$$

$$\leq \|x\|^2 + \|y\|^2 + 2|(x|y)|$$ , where $\mathcal{R}_e(\lambda)$

denotes the real part of $\lambda$ . Now applying the Caucy-Schwarz inequality,

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2 \; .$$

Theorem 6 is actually the statement $n_3$ of the third property of a norm and justifies writing $\|x\|$ for $(x|x)^{\frac{1}{2}}$ . Thus every pre-Hilbert space becomes a normed vector space with $\|x\|$ thus defined. Consider now the pre-Hilbert space of complex numbers, which is essentially Example 5 with n = 1. For <u>vectors</u> $\lambda, \mu \in C$,

$$\|\lambda - \mu\| = (|\lambda - \mu|^2)^{\frac{1}{2}} = |\lambda - \mu| ,$$

which also represents the distance from the point $\lambda$ to the point $\mu$ in the complex plane. This provides the motivation, and the following theorem the justification, for defining $\|x - y\|$ , where x and y are vectors in a pre-Hilbert space $\mathcal{P}$ , to be the <u>distance</u> between x and y.

THEOREM 7. If $x,y,z \in \mathcal{P}$, then

(1) $\|x - y\| \geq 0$ , with equality iff $x = y$ ,

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

(3) $\|x - z\| \leq \|x - y\| + \|y - z\|$ .

Proof:

(1) $\|x\| > 0$ if $x \neq \theta$ , and $(\theta/\theta) = 0$ shows that $\|\theta\| = 0$ .

(2) $[x - y] = - [y - x]$ .

(3) Use Theorem 6, with $(x - z) = (x - y) + (y - z)$ .

Any nonempty set $\mathcal{X}$ having a distance function defined on every pair of elements, such that d is strictly positive, symmetric, and satisfies the triangle inequality, is said to be a metric space. These properties are precisely those of $\|x - y\|$ given by Theorem 7. Therefore every pre-Hilbert space $\mathcal{P}$ is a metric space, with $d(x,y) = \|x - y\|$ , for all $x,y \in \mathcal{P}$ . Given a metric space $\mathcal{X}$ , a sequence $\{x_n\}$ is said to be a Cauchy sequence if $d(x_m,x_n) \to 0$ as $m,n \to \infty$ . If $\{x_n\}$ being a Cauchy sequence implies that there is an $x \in \mathcal{X}$ such that $d(x,x_n) \to 0$ as $n \to \infty$ , then $\mathcal{X}$ is said to be complete. Two of the given examples of pre-Hilbert spaces fail to be complete. Example 5, the pre-Hilbert space of functions continuous on $[a,b]$ , is incomplete. The sequence of functions

$$x_n(t) = \begin{cases} 0 \text{ if } -1 \leq t \leq 0 , \\ nt \text{ if } 0 < t < \frac{1}{n} , \\ 1 \text{ if } \frac{1}{n} \leq t \leq 1 , \end{cases}$$

defined and continuous on $[-1,1]$ , is a Cauchy sequence, but it fails to converge to a continuous function defined on $[-1,1]$ . Again, Example 6, the pre-Hilbert space of finitely nonzero sequences, is incomplete. Consider the sequence $\{x_n\} = \{(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, 0, 0, \ldots)\}$ . Then

$$\|x_{n+p} - x_n\|^2 = \|(0,\ldots,0, \frac{1}{n+1}, \frac{1}{n+2}, \ldots, \frac{1}{n+p}, 0, \ldots)\|^2 = \sum_{n+1}^{n+p} \frac{1}{k^2} .$$

The series $\sum_1^\infty \frac{1}{k^2}$ is known to be convergent, thus $\sum_{n+1}^{n+p} \frac{1}{k^2} \to 0$ as $n \to \infty$. Hence $d(x_{n+p}, x_n) = \left(\sum_{n+1}^{n+p} \frac{1}{k^2}\right)^{\frac{1}{2}} \to 0$ as $n \to \infty$.

Thus $\{x_n\}$ is a Cauchy sequence. Assume $\{x_n\} \to x$, $x \in \mathcal{P}$. Let $x = (\lambda_1, \lambda_2, \ldots, \lambda_N, 0, \ldots)$. Then choosing $n \geq N$, $\|x_n - x\|^2 =$

$$\sum_1^n \left|\frac{1}{k} - \lambda_k\right|^2 + \sum_{n+1}^\infty |\lambda_k|^2 = \sum_1^n \left|\frac{1}{k} - \lambda_k\right|^2 .$$

Letting $n \to \infty$, $\sum_1^\infty \left|\frac{1}{k} - \lambda_k\right|^2 = 0$. This requires that $\lambda_k = \frac{1}{k}$ for all k, which contradicts the hypothesis that x is finitely nonzero. Thus no such x exists, and $\{x_n\}$ does not converge to a vector in $\mathcal{P}$.

A Hilbert space is defined to be a complete pre-Hilbert space. Two examples have been given of pre-Hilbert spaces which fail to be Hilbert spaces; the following is a complete pre-Hilbert space.

EXAMPLE 7. (The Hilbert Space $1^2$.) Let $\mathcal{H}$ be the set of all absolutely square-summable sequences $(\lambda_k)$; that is, $\sum_1^\infty |\lambda_k|^2 < \infty$, $\lambda_k \in C$.

LEMMA 1. If $x = (\lambda_k)$ and $y = (\mu_k)$ are in $\mathcal{H}$, and if $\lambda \in C$, then $x + y$ and $\lambda x$ are in $\mathcal{H}$.

Proof:

Applying Theorem 3 to the complex numbers

$$|\lambda_k + \mu_k|^2 + |\lambda_k - \mu_k|^2 = 2|\lambda_k|^2 + 2|\mu_k|^2 .$$

Hence
$$\sum_1^\infty |\lambda_k + \mu_k|^2 \leq 2\sum_1^\infty |\lambda_k|^2 + 2\sum_1^\infty |\mu_k|^2 < \infty .$$

Also
$$\sum_1^\infty |\lambda \lambda_k|^2 = |\lambda|^2 \sum_1^\infty |\lambda_k|^2 < \infty .$$

Therefore the sum of two elements of $\mathcal{H}$, and any scalar multiple of an element of $\mathcal{H}$, is an element of $\mathcal{H}$. $\mathcal{H}$ forms a vector space under these operations, with $\theta$ the sequence all of whose terms are zero.

LEMMA 2. If $x = (\lambda_k)$ and $y = (\mu_k)$ belong to $\mathcal{H}$, then the series $\sum_1^\infty \lambda_k \bar{\mu}_k$ is absolutely convergent.

Proof:

For all real numbers a and b, $(a - b)^2 \geq 0$, hence $ab < \frac{1}{2}(a^2 + b^2)$.

Thus $$|\lambda_k \bar{\mu}_k| = |\lambda_k||\mu_k| \leq \tfrac{1}{2}\left(|\lambda_k|^2 + |\mu_k|^2\right).$$

Hence $$\sum_1^\infty |\lambda_k \bar{\mu}_k| \leq \tfrac{1}{2}\left(\sum_1^\infty |\lambda_k|^2 + \sum_1^\infty |\mu_k|^2\right) < \infty.$$

With this lemma as justification, let $(x|y) = \sum_1^\infty \lambda_k \bar{\mu}_k$, since this series converges. Then $(x|y)$ satisfies $P_{1-4}$. Thus $\mathcal{H}$ is a pre-Hilbert space. To complete the proof that $\mathcal{H}$ is a Hilbert space, let $\{x^n\}$ be a Cauchy sequence of vectors in H. That is, $\| x^m - x^n \| \to 0$ as $m, n \to \infty$. Let $x^n = (\lambda_k^{(n)})$. Now for each k,

$$|\lambda_k^{(m)} - \lambda_k^{(n)}|^2 \leq \sum_{k=1}^\infty |\lambda_k^{(m)} - \lambda_k^{(n)}|^2 = \|X^m - X^n\|^2 .$$

Thus the sequence $(\lambda_k^{(n)})$, where n is the running index, that is, the sequence of k-th components of $\{x^n\}$, is a Cauchy sequence. Since the complex numbers are complete $(\lambda_k^{(n)}) \to \lambda_k$ as $n \to \infty$, where $\lambda_k \in C$. It will now be shown that $\{x_k\} \to (\lambda_k)$; that is, $\mathcal{H}$ is complete.

LEMMA 3. $\sum_1^\infty |\lambda_k|^2 < \infty$, and $\{x^n\} \to x$, as $n \to \infty$, where $x = (\lambda_k)$.

Proof:

Choose $\varepsilon > 0$ and a positive integer p such that $\|x^m - x^n\|^2 \leq \varepsilon$ whenever $m,n \geq p$. Now for any positive integer r,

$$\sum_{k=1}^r |\lambda_k^{(m)} - \lambda_k^{(n)}| \leq \|X^m - X^n\|^2 \leq \varepsilon ,$$

provided that $m,n \geq p$ . Let $m \to \infty$; then $\sum_{k=1}^{r} | \lambda_k - \lambda_k^{(n)} |^2 \leq \varepsilon$ provided $n \geq p$ . Hence $\sum_{k=1}^{\infty} | \lambda_k - \lambda_k^{(p)} | \leq \varepsilon$ , which implies that the sequence $(\lambda_k - \lambda_k^{(p)})$ is a vector in $\mathcal{H}$ . Adding this to the vector $x_p = (\lambda_k^{(p)})$ , $x = (\lambda_k) \in \mathcal{H}$ by Lemma 1. Now from $\sum_1^{\infty} |\lambda_k - \lambda_k^{(n)}|^2 \leq \varepsilon$, it follows that $\| x - x^n \|^2 \leq \varepsilon$ whenever $n \geq p$ . Thus $\{x^n\} \to x$ .

The Hilbert space $\mathcal{H}$ of absolutely square-summable sequences is usually denoted $l^2$ .

## ORTHOGONAL VECTORS

If x and y are vectors in a pre-Hilbert space $\mathcal{P}$ , and if $(x|y) = 0$ , then x is said to be orthogonal to y. This is denoted $x \perp y$ . Since $(x|y) = \overline{(y|x)}$ and $(x|x) = 0$ iff $x = \theta$ , it follows that $y \perp x$ if $x \perp y$ , and that $x \perp x$ iff $x = \theta$ . Again, $x \perp \theta$ holds for all $x \in \mathcal{P}$ . A set $\mathcal{S}$ of vectors is said to be orthogonal if $x \perp y$ for any distinct $x,y \in \mathcal{S}$ . A sequence $\{x_n\}$ of vectors is defined to be an orthogonal sequence if $x_i \perp x_j$ whenever $i \neq j$ .

LEMMA 4. If x is orthogonal to each of $y_1, y_2, \ldots, y_n$, then x is orthogonal to every linear combination of the $y_k$ .

Proof:

If $y = \sum_1^n \lambda_k y_k$ ,then $(x|y) = \sum_1^n \bar{\lambda}_k (x|y_k) = \sum_1^n \bar{\lambda}_k 0 = 0$ .

THEOREM 8. (The Pythagorean Relation.) If $x \perp y$, then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ . In general, if $x_1, \ldots, x_n$ form an orthogonal set, then $\| \sum_1^n x_k \|^2 = \sum_1^n \|x_k\|^2$ .

Proof:

$$\|x + y\|^2 = (x|x) + (x|y) + (y|x) + (y|y)$$

$$= \|x\|^2 + 0 + 0 + \|y\|^2 .$$

Assume inductively that $\|\sum_1^{n-1} x_k\|^2 = \sum_1^{n-1}\|x_k\|^2$ . Then for $x = \sum_1^{n-1} x_k$ and $y = x_n$ , $x \perp y$ by Lemma 4. Thus

$$\|\sum_1^n x_k\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2$$

$$= \|\sum_1^{n-1} x_k\|^2 + \|x_n\|^2$$

$$= \sum_1^{n-1}\|x_k\|^2 + \|x_n\|^2$$

by hypothesis. But this last sum is $\sum_1^n \|x_k\|^2$ , which completes the induction.

COROLLARY. If $\{x_n\}$ is an orthogonal sequence of nonzero vectors, then the $x_k$ are linearly independent.

Proof:

The sequence $\{\lambda_n x_n\}$ is orthogonal. Hence if $\sum_1^n \lambda_k x_k = 0$, then by Theorem 8, $\|\sum_1^n \lambda_k x_k\|^2 = 0 = \sum_1^n \|\lambda_k x_k\|^2 = \sum_1^n |\lambda_k|^2 \|x_k\|^2$ .

But $\|x_k\| > 0$ for all k, hence $\lambda_k = 0$ for all k.

An orthogonal sequence $\{x_n\}$ such that $\|x_n\| = 1$ for all n is said to be an <u>orthonormal</u> <u>sequence</u>. This may be expressed by writing $(x_j|x_k) = \delta_{jk}$ , where $\delta_{jk}$ denotes the Knonecker delta function.

EXAMPLE 8. In the Hilbert space $l^2$ , let $\{e_n\}$ be the sequence whose n-th term is the sequence whose n-th term is one, and all others zero: $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$ , where n-1 zeros precede the 1 . (It is clear that the absolute square-sum of each $e_k$ is one, hence $e_n \in l^2$ for all k.) Again, $(e_j|e_k) = \delta_{jk}$ . If $x = (\lambda_k)$ , then $(x|e_j) = \lambda_j$ . Thus for every vector x, $\sum_1^\infty |(x|e_k)|^2 < \infty$ . That a similar result holds for any orthonormal sequence of vectors is shown by the following theorem.

THEOREM 9. (Bessel's Equality and Inequality.) Let $x_1, \ldots, x_n$ be an orthonormal sequence in a pre-Hilbert space $\mathcal{P}$ . Then for every vector x ,

(1) $\|x - \sum_1^n (x|x_k)x_k\|^2 = \|x\|^2 - \sum_1^n |(x|x_k)|^2$ , and

(2) $\sum_1^n |(x|x_k)|^2 \leq \|x\|^2$ .

Proof:

By Theorem 8, if $\lambda_k \in \mathcal{C}$ for $k \leq n$ , then

$$\begin{aligned}\Big\|\sum_1^n \lambda_k x_k\Big\|^2 &= \sum_1^n \|\lambda_k x_k\|^2 \\ &= \sum_1^n |\lambda_k|^2 \|x_k\|^2 \\ &= \sum_1^n |\lambda_k|^2 \quad .\end{aligned}$$

By expanding the norm:

$$\begin{aligned}\Big\|x - \sum_1^n \lambda_k x_k\Big\|^2 &= \|x\|^2 - \Big(\sum_1^n \lambda_k x_k \Big| x\Big) - \Big(x \Big| \sum_1^n \lambda_k x_k\Big) + \sum_1^n |\lambda_k|^2 \\ &= \|x\|^2 - \sum_1^n \lambda_k \overline{(x|x_k)} - \sum_1^n (x|x_k)\overline{\lambda}_k + \sum_1^n \lambda_k \overline{\lambda}_k \\ &= \|x\|^2 - \sum_1^n |(x|x_k)|^2 + \sum_1^n \Big[(x|x_k)\overline{(x|x_k)} - \\ &\quad \overline{\lambda}_k (x|x_k) - \lambda_k \overline{(x|x_k)} + \lambda_k \overline{\lambda}_k\Big]\end{aligned}$$

$$= \|x\|^2 - \sum_1^n |(x|x_k)|^2 + \sum_1^n [(x|x_k) - \lambda_k] \overline{[(x|x_k) - \lambda_k]}$$

$$= \|x\|^2 - \sum_1^n |(x|x_k)|^2 + \sum_1^n |(x|x_k) - \lambda_k|^2 \quad .$$

Setting $\lambda_k = (x|x_k)$ gives the result (1) . Since $\|x - \sum_1^n (x|x_k)x_k\|^2 \geq 0$, result (2) follows immediately.

COROLLARY. If $\{x_n\}$ is an infinite orthonormal sequence in a pre-Hilbert space, then $\sum_1^\infty |(x|x_k)|^2 \leq \|x\|^2$ . Consequently, $(x|x_k) \to 0$ as $n \to \infty$.

Proof:

This is an immediate consequence of Bessel's inequality, since this result must hold for each n.

Since $\|x\|$ is strictly positive, equality will hold in (2) iff x is a linear combination of the $x_k$ . That this is sufficient in order to have equality follows from the implication: if $x = \sum_1^n \lambda_k x_k$ , then $\lambda_k = (x|x_k)$ . A generalization of this result is given by the last step in the proof of (1), which shows that the choice $\lambda_k = (x|x_k)$ minimizes $\|x - \sum_1^n \lambda_k x_k\|$ . Therefore this choice of the $\lambda_k$ provides the best approximation of x by a linear combination of the $x_k$ .

Theorem 5, the Cauchy-Schwarz inequality $|(x|y)| \leq \|x\| \|y\|$ , will now be proved.

Proof of Theorem 5. If either x or y is $\theta$ , then both sides are zero. Assume $y \neq \theta$, and let $y_0 = y\|y\|^{-1}$, so that $\|y_0\| = 1$. Then $y_0$ forms an orthonormal sequence of one term. Hence by part (2) of Theorem 9 above, $|(x|y_0)|^2 \leq \|x\|^2$. Replacing $y_0$ by $y\|y\|^{-1}$ , the theorem follows.

It has been shown that every orthonormal sequence is linearly independent. A construction process will now be described by which an orthonormal sequence may be derived from a given linearly independent sequence.

LEMMA 5. In the notation of Theorem 9, let $y = \sum_1^n (x|x_k)x_k$ , and let $z = x - y$ . Then $(z|x_k) = 0$ for all $k \leq n$ .

Proof:

$$\begin{aligned}(z|x_k) &= (x|x_k) - (y|x_k) \\ &= (x|x_k) - (\textstyle\sum_1^n (x|x_j)x_j | x_k) \\ &= (x|x_k) - (x|x_k)(x_k|x_k) = 0,\end{aligned}$$

where the sum of the scalar products given by $(y|x_k)$ reduces to the one term by virtue of the fact that $\{x_k\}$ is orthonormal.

THEOREM 10. (The Gram-Schmidt Orthonormalization Process.) If $y_1, y_2, \ldots,$ is a sequence of linearly independent vectors in a pre-Hilbert space $\mathcal{P}$, then there exists an orthonormal sequence $x_1, x_2, \ldots$ in $\mathcal{P}$ such that $[y_1, \ldots, y_n] = [x_1, \ldots, x_n]$ for all n (the $y_k$ and the $x_k$ generate the same subspaces).

Proof:

Define $x_1 = y_1 \, \|y_1\|^{-1}$. Clearly $x_1$ and $y_1$ generate the same subspace. Assume inductively that orthonormal vectors $x_1, x_2, \ldots, x_{n-1}$ are given, such that $[y_1, \ldots, y_{n-1}] = [x_1, \ldots, x_{n-1}]$ . Let $z =$ $y_n - \sum_1^{n-1} (y_n|x_k)x_k$ . By the lemma, $z \perp x_k$ for all $k \leq n$ . If $z = \theta$ , then $y_n$ is a linear combination of $x_1, \ldots, x_{n-1}$ , hence also of $y_1, \ldots, y_{n-1}$ , which contradicts the independence of $\{y_n\}$. Thus $z \neq \theta$, and it is permissible to define $x_n = z \, \|z\|^{-1}$. Now

$$\sum_{k=1}^{n} \lambda_k x_k = \sum_{k=1}^{n-1} \lambda_k x_k + \lambda_n y_n - \lambda_n \sum_{k=1}^{n-1} (y_n | x_k) x_k$$

$$= \sum_{k=1}^{n-1} \left[\lambda_k - \lambda_n (y_n | x_k)\right] x_k + \lambda_n y_n$$

$$= \sum_{1}^{n-1} \mu_k y_k + \lambda_n y_n \quad \text{by the induction hypothesis.}$$

Thus $[x_1, \ldots, x_n] = [y_1, \ldots, y_n]$, which completes the induction.

Since every finite dimensional pre-Hilbert space has a basis, an immediate consequence of Theorem 10 is that every finite-dimensional pre-Hilbert space has a basis which consists of orthonormal vectors. This property is essential for the proof of the following theorem.

THEOREM 11. Every finite-dimensional pre-Hilbert space $\mathcal{P}$ is complete, hence is a Hilbert space.

Proof:

Let $\mathcal{P}$ be of dimension n, and let $x_1, \ldots, x_n$ be a basis for $\mathcal{P}$ consisting of orthonormal vectors. If $x = \sum_1^n \lambda_k x_k$, then $\|x\|^2 =$

$\sum_1^n (\lambda_k x_k | \lambda_k x_k) = \sum_1^n |\lambda_k|^2$ by Theorem 8. Thus if $\{y_m\}$ is a Cauchy sequence in $\mathcal{P}$, where $y_m = \sum_{k=1}^{n} \lambda_k^{(m)} x_k$, then $\| y_m - y_p \|^2 =$ $\sum_1^n | \lambda_k^{(m)} - \lambda_k^{(p)} |^2 \to 0$ as $m, p \to \infty$. In particular,

$| \lambda_k^{(m)} - \lambda_k^{(p)} |^2 \leq \|y_m - y_p\|^2$ for each fixed k. Thus the sequence $(\lambda_k^{(m)})$, where m in the running index, is a Cauchy sequence. The complex numbers are complete, therefore $(\lambda_k^{(m)}) \to (\lambda_k)$ as $m \to \infty$. For a fixed k and any $\varepsilon > 0$ , there is an $M_k(\varepsilon)$ such that $|\lambda_k^{(m)} - \lambda_k| < \sqrt{\varepsilon}$ for all $m > M_k(\varepsilon)$ . Let $M(\varepsilon) = \max\{M_k(\varepsilon)\}$ . Thus if $y = \sum_1^n \lambda_k x_k$ , then $\|y_m - y\|^2 = \sum_1^n |\lambda_k^{(m)} - \lambda_k|^2 < \sum_1^n \varepsilon = n\varepsilon$ for $m > M(\varepsilon)$. Thus $\{y_m\} \to y$ . Clearly $y \in \mathcal{P}$ . Therefore $\mathcal{P}$ is complete.

## HILBERT SPACE

A few general properties of Hilbert space will now be discussed. The first of these is concerned with infinite sums of vectors in an infinite-dimensional Hilbert space. If $x_1, x_2, x_3, \ldots$ is a sequence of vectors such that the sequence $y_n = \sum_1^n x_k$ converges to a limit x, then this limit is denoted $x = \sum_1^\infty x_k$ .

THEOREM 12. If $\{x_n\}$ is an infinite orthonormal sequence of vectors in a Hilbert space $\mathcal{H}$ and $\{\lambda_n\}$ a sequence of scalars such that $\sum_1^\infty |\lambda_k|^2 < \infty$ , then the sequence $\{y_n\}$ , where $y_n = \sum_1^n \lambda_k x_k$, converges to a limit $x = \sum_1^\infty \lambda_k x_k$.

Proof:

$$\|y_{n+p} - y_n\|^2 = \sum_{n+1}^{n+p} \|\lambda_k x_k\|^2$$

$$= \sum_{n+1}^{n+p} |\lambda_k|^2 \longrightarrow 0$$

as $n \longrightarrow \infty$ . Thus $\{y_n\}$ is a Cauchy sequence, and $\{y_n\} \to x \in \mathcal{H}$ since $\mathcal{H}$ is complete.

LEMMA 6. In a pre-Hilbert space ,

(1) if $\{x_n\} \to x$ and $\{y_n\} \to y$, then $\{(x_n|y_n)\} \to (x|y)$;

(2) if $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences of vectors, then $\{(x_n|y_n)\}$ is a Cauchy sequence of scalars, hence is convergent.

Proof:

(1) For all n, $(x_n|y_n) - (x|y) = (x_n - x|y_n - y) + (x|y_n - y) + (x_n - x|y)$

Hence

$$|(x_n|y_n) - (x|y)| \leq |(x_n - x|y_n - y)| + |(x|y_n - y)| + |(x_n - x|y)|$$

$$\|x_n - x\| \cdot \|y_n - y\| + \|x\| \cdot \|y_n - y\|$$

$$+ \|x_n - x\| \cdot \|y\|$$

by the Cauchy-Schwarz inequality. Since the right side of this inequality goes to zero as $n \to \infty$, it follows that $\{(x_n|y_n)\} \to (x|y)$

(2) $|(x_n|y_n) - (x_m|y_m)| \leq \|x_n - x_m\| \cdot \|y_n - y_m\|$

$$+ \|x_m\| \cdot \|y_n - y_m\| + \|x_n - x_m\| \cdot \|y_m\| .$$

It will be shown that every Cauchy sequence is bounded, thus the right side of this inequality must go to zero as $n,m \to \infty$. Choose N such that $\|x_m - x_n\| \leq 1$ whenever $m,n \geq N$. Then

$$\|x_n\| = \|x_n - x_N + x_N\| \leq \|x_n - x_N\| + \|x_N\| \leq 1 + \|x_N\|$$

for all $n \geq N$. Choose $M = \max \{1 - \|x_N\|, \|x_1\|, \ldots, \|x_{N-1}\|\}$ .
Then $\|x_n\| \leq M$ for all n.

THEOREM 13. Let $\{x_n\}$ be an orthonormal sequence of vectors in a Hilbert space $\mathcal{H}$ . If $x = \sum_1^\infty \lambda_k x_k$ and $y = \sum_1^\infty \mu_k x_k$ in the notation of Theorem 12, then

(1) $(x|y) = \sum_1^\infty \lambda_k \bar{\mu}_k$ , where this series converges absolutely;

(2) $(x|x_k) = \lambda_k$ ;

(3) $\|x\|^2 = \sum_1^\infty |\lambda_k|^2 = \sum_1^\infty |(x|x_k)|^2$ .

Proof:

(1) Let $s_n = \sum_1^n \lambda_k x_k$ and $t_n = \sum_1^n \mu_k x_k$ , so that $\{s_n\} \to x$ and $\{t_n\} \to y$ by definition. Then $(s_n|t_n) \to (x|y)$ by Lemma 6. Thus

$$(s_n|t_n) = \left(\sum_{k=1}^n \lambda_k x_k \,\Big|\, \sum_{j=1}^n \mu_j x_j\right)$$

$$= \sum_{k=1}^n \left(\sum_{j=1}^n \lambda_k \bar{\mu}_j (x_k|x_j)\right)$$

$$= \sum_{k=1}^n \lambda_k \bar{\mu}_k \ .$$

Thus $(x|y) = \sum_1^\infty \lambda_k \mu_k$ . That the sequence is absolutely convergent follows by Lemma 2.

(2) In (1), let $\mu_k = 1$ and $\mu_j = 0$ , $j \neq k$ .

(3) In (1), let $x = y$, then let $\lambda_k = (x|x_k)$ by (2).

The sequence $\{x_n\}$ need not be orthogonal in order to give $\sum_1^\infty x_n$ meaning. For example, if $\{x_n\}$ is a sequence in a Hilbert space $\mathcal{H}$ such that $\sum_1^\infty \|x_k\| < \infty$ , then the sequence $\{y_n\} = \{\sum_1^n x_k\}$ is a Cauchy sequence: $\|y_{n+p} - y_n\| = \|\sum_{n+1}^{n+p} x_k\| \leq \sum_{n+1}^{n+p} \|x_k\|$ by Theorem 6. Thus $\|y_{n+p} - y_n\| \to 0$ as $n \to \infty$ . Since $\mathcal{H}$ is complete, it follows that $\{y_n\} \to x \in \mathcal{H}$ , where $x = \sum_1^\infty x_k$ .

If $\{x_n\}$ is an infinite orthonormal sequence in a Hilbert space $\mathcal{H}$ , then by the Corollary to Theorem 9, the scallars $\lambda_k = (x|x_k)$ satisfy

$\sum_1^\infty |\lambda_k|^2 < \infty$ for any $x \in \mathcal{H}$. Thus by Theorem 12, there is a $y \in \mathcal{H}$ such that $y = \sum_1^\infty \lambda_k x_k$, and by Theorem 13, $(y|x_k) = \lambda_k$ for each k. Thus $(y - x|x_k) = (y|x_k) - (x|x_k) = \lambda_k - \lambda_k = 0$ for each k. However, this would imply that $x = y$ iff $z \in \mathcal{H}$ and $z \perp x_k$ for each k implies that $z = \theta$. This leads to the following definition: a set of vectors in a pre-Hilbert space $\mathcal{P}$ is said to be <u>total</u> if the only vector in $\mathcal{P}$ which is orthogonal to every vector in $\mathcal{S}$ is the vector $\theta$. Similarly, a sequence $\{x_n\}$ in $\mathcal{P}$ is a <u>total sequence</u> if $z \perp x_k$ for all k implies that $z = \theta$.

EXAMPLE 9. The set of all vectors in a pre-Hilbert space $\mathcal{P}$ is a total set, since $z \perp x$ for all $x \in \mathcal{P}$ implies in particular that $z \perp z$, hence $z = \theta$.

EXAMPLE 10. In the Hilbert space $l^2$, the sequence whose n-th term is $e_n = (0,0,..,0,1,0...)$, where the 1 is the n-th component, is a total sequence of vectors. This sequence is also a total sequence in the pre-Hilbert space of finitely nonzero sequences.

A sequence of vectors $\{x_n\}$, either finite or infinite, in a Hilbert space $\mathcal{H}$ is said to be an <u>orthonormal basis</u> for $\mathcal{H}$ provided (i) $\{x_n\}$ is orthonormal, and (ii) $\{x_n\}$ is total. The sequence described in Example 10 is an orthonormal basis for $l^2$. It is called the canonical <u>orthonormal basis</u> of $l^2$.

If $\mathcal{P}$ is a pre-Hilbert space possessing a finite sequence $x_1,\dots,x_n$ which is orthonormal and total, then by Lemma 5, $x - \sum_1^n (x|x_k)x_k$ is orthogonal to each $x_k$; thus $x = \sum_1^n (x|x_k)x_k$. It follows that the $x_k$ form a basis for $\mathcal{P}$, and $\mathcal{P}$ has dimension n. The following theorem shows that a similar result holds for infinite-dimensional spaces.

THEOREM 14. If $\{x_n\}$ is an infinite orthonormal sequence in a Hilbert space $\mathcal{H}$, then the following are equivalent:

(1) $\{x_n\}$ forms an orthonormal basis of $\mathcal{H}$.

(2) $\sum_1^\infty |(x|x_k)|^2 = \|x\|^2$ for each $x \in \mathcal{H}$.

(3) $\sum_1^\infty (x|x_k)x_k = x$ for each $x \in \mathcal{H}$.

Proof:

Statement (1) implies statement (3) by Theorems 12 and 13. Statement (3) implies statement (2) by Theorem 13, part (3). Statement (2) implies statement (1), since if $(x|x_k) = 0$ for all k, then $\|x\| = 0$, hence $x = \theta$.

Not every Hilbert space has an orthonormal basis. Those that do certainly contain a total sequence, and such Hilbert spaces are said to be separable. This condition is not only necessary in order that a Hilbert space have an orthonormal basis, but it is also sufficient.

THEOREM 15. The following statements about a Hilbert space $\mathcal{H}$ are equivalent:

(1) $\mathcal{H}$ is separable.

(2) $\mathcal{H}$ has an orthonormal basis $\{x_n\}$.

Proof:

(2) implies (1) by the definitions of an orthonormal basis and of separable space. (1) implies (2) : Let $z_1, z_2, z_3, \ldots$ be a total sequence in $\mathcal{H}$. Then there is a linearly independent subsequence $y_1, y_2, y_3, \ldots$ of $\{z_k\}$ which generates the same linear subspace as $\{z_k\}$. This follows by a

constructive argument: if $z_k = \theta$ for all k, then let the sequence be the empty sequence. This by convention is taken to be an independent sequence which generates the subspace $\theta$ . If not $z_k = \theta$ for all k, then let $k_1$ be the smallest index such that $z_{k_1} \neq \theta$ . Then for $k < k_1$ , $z_k = \theta = 0z_{k_1}$ . If for $k > k_1$ each $z_k$ is a scalar multiple of $z_{k_1}$ , then let the sequence $\{y_n\}$ consist of the single term $z_{k_1}$ . If this is not the case, let $k_2$ be the least positive integer greater than $k_1$ such that $z_{k_2}$ is not a multiple of $z_{k_1}$ . If each $z_k$ for $k > k_2$ is expressible as a linear combination of $z_{k_1}$ , $z_{k_2}$ , then let $\{y_n\}$ consist of the vectors $z_{k_1}$ and $z_{k_2}$ . If this is not the case, let $z_{k_3}$ denote the next vector not expressible as a linear combination of $z_{k_1}$ , $z_{k_2}$ , and continue the process. If $\{z_k\}$ is finite, then the process will exhaust the sequence, yielding a finite independent subsequence $\{z_{k_j}\} = \{y_j\}$ . If $\{z_k\}$ is not finite, the process continues inductively. The sequence (possibly finite) which the process yields is also total: if a vector z is orthogonal to every $y_j$ , then it is orthogonal to every linear combination of the $y_j$; therefore it is orthogonal to each $z_k$ .

By Theorem 10, there is an orthonormal sequence $\{x_j\}$ generating the same subspace as $\{y_j\}$ . An argument similar to the one used to show that $\{y_j\}$ is a total sequence shows that $\{x_j\}$ is also total; which completes the proof.

Two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ are said to be <u>isomorphic</u> if there is a one-one mapping T of $\mathcal{H}$ onto $\mathcal{K}$ such that for $x,y \in \mathcal{H}$ ,

(i) $(x + y)T = xT + yT$ ;

(ii) $(\lambda x)T = \lambda(xT)$ for any $\lambda \in \mathcal{C}$;

(iii) $(xT|yT) = (x|y)$ .

The function T is called a Hilbert space isomorphism of $\mathcal{H}$ onto $\mathcal{K}$.

Two types of separable Hilbert spaces are given in the following examples. The next theorem shows that up to isomorphism, these are the only separable Hilbert spaces.

EXAMPLE 11. If $\mathcal{H}$ is a finite-dimensional Hilbert space, then every basis is total, thus $\mathcal{H}$ is separable. In particular, for any positive integer n, the unitary space $\mathcal{C}^n$ is an n-dimensional separable Hilbert space.

EXAMPLE 12. The Hilbert space $l^2$ is separable, as is shown by Example 10.

THEOREM 16. If $\mathcal{H}$ is a separable Hilbert space, then $\mathcal{H}$ is isomorphic to $\mathcal{C}^n$ or to $l^2$ , according as $\mathcal{H}$ has finite dimension n or has infinite dimension.

Proof:

First assume that $\mathcal{H}$ has dimension n. Then by Theorem 15, $\mathcal{H}$ has an orthonormal basis $x_1,\ldots,x_n$ . By Lemma 5, if $x \in \mathcal{H}$ , then $(x - \sum_1^n (x|x_k)x_k) \perp x_k$ for all k. The sequence $x_n$ is total; therefore $x = \sum_1^n (x|x_k)x_k$ . Let $xT = ((x|x_k))$ ; that is, $xT = (\lambda_n) \in \mathcal{C}^n$ such that $\lambda_i = (x|x_i)$. Then if $x,y \in \mathcal{H}$ such that $x \neq y$, then $x - y \neq \theta$ . Hence $(x - y|x_k) \neq 0$ for some k, say k = j. Thus $(x|x_j) - (y|x_j) \neq 0$, from which $((x|x_k)) \neq ((y|x_k))$ . If $(\lambda_n) \in \mathcal{C}^n$, then $x = \sum_1^n \lambda_k x_k \in \mathcal{H}$ .

By the remark above, this implies that $\lambda_k = (x|x_k)$ . Therefore T is one-one from $\mathcal{H}$ onto $\mathcal{C}^n$ .

(i) $x + y = \sum_1^n [(x|x_k)\, x_k + (y|x_k)x_k]$

$$= \sum_1^n [(x|x_k) + (y|x_k)]\, x_k$$

$$= \sum_1^n (x + y|x_k)x_k \quad .$$

Thus $(x + y)T = ((x + y|x_k)) = ((x|x_k) + (y|x_k))$

$$= ((x|x_k)) + ((y|x_k))$$

$$= xT + yT \quad .$$

(ii) $\lambda x = \lambda \sum_1^n (x|x_k)x_k = \sum_1^n (\lambda x|x_k)x_k$ .

Thus $(\lambda x)T = ((\lambda x|x_k)) = (\lambda(x|x_k))$

$$= \lambda((x|x_k)) = \lambda(xT) \quad .$$

(iii) $(x|y) = \left( \sum_{k=1}^n (x|x_k)x_k \;\middle|\; \sum_{j=1}^n (y|x_j)x_j \right)$

$$= \sum_{k=1}^n \sum_{j=1}^n (x|x_k)\overline{(y|x_j)}(x_k|x_j)$$

$$= \sum_{k=1}^n (x|x_k)\overline{(y|x_k)}$$

$$= (xT|yT) \quad .$$

If on the other hand $\mathcal{H}$ has infinite dimension, then $\mathcal{H}$ is isomorphic to $l^2$ . Again an orthonormal basis $\{x_n\}$ is given by Theorem 15. By the Corollary of Theorem 9, $\sum_1^\infty |(x|x_k)|^2 < \infty$ ; thus it is meaningful to define $xT = ((x|x_k))$ . If $x \neq y$, then $(x - y|x_j) \neq 0$ for some j ; thus

$xT \neq yT$ . Let $(\lambda_k) \in l^2$ , and define $x = \sum_1^\infty \lambda_k x_k$ . By Theorem 13, $\lambda_j = (x|x_j)$ . Thus $xT = (\lambda_k)$ , and T is one-one from $\mathcal{H}$ onto $l^2$ . If $x,y \in \mathcal{H}$ , then

(i) $(x + y|x_k) = (x|x_k) + (y|x_k)$ for all k, hence $(x + y)T = xT + yT$ .

(ii) $(\lambda x|x_k) = \lambda(x|x_k)$ , thus $(\lambda x)T = \lambda(xT)$ .

(iii) By Theorem 14, $x = \sum_1^\infty (x|x_k)x_k$ and $y = \sum_1^\infty (y|x_k)x_k$ . Thus

$$(x|y) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} (x|x_k)\overline{(y|x_j)}(x_k|x_j)$$

$$= \sum_1^\infty (x|x_k)\overline{(y|x_k)}$$

$$= (xT|yT) .$$

LEMMA 7. If a pre-Hilbert space $\mathcal{P}$ is isomorphic to a Hilbert space $\mathcal{H}$ in the sense that there exists a mapping T of $\mathcal{P}$ onto $\mathcal{H}$ which satisfies the conditions for a Hilbert space isomorphism, then $\mathcal{P}$ is itself a Hilbert space.

Proof:

Let the sequence $\{x_n\}$ be a Cauchy sequence in $\mathcal{P}$ . Then by condition (iii) for an isomorphism, $\|x_mT - x_nT\|^2 = \|x_m - x_n\|^2 < \mathcal{E}$, $n,m > N(\mathcal{E})$ . Thus $\{x_mT\}$ is a Cauchy sequence in $\mathcal{H}$ , and $\{x_mT\} \rightarrow y \in \mathcal{H}$ . But there is an $x \in \mathcal{P}$ such that $xT = y$. Hence $\|x_n - x\|^2 = \|x_nT - xT\|^2 < \mathcal{E}$ for $n > N$, and $\{x_n\} \rightarrow x$ .

The term pre-Hilbert space has been used to designate an inner-product linear vector space which is not necessarily complete in the metric defined by this inner product. If the space $\mathcal{P}$ is not complete, then the set of vectors in $\mathcal{P}$ may be enlarged by the ordinary process of metric space completion to a new set $\mathcal{P}^*$ in such a manner that $\mathcal{P}^*$ forms a complete metric space. Furthermore, this may be done in such a way that $\mathcal{P}^*$ forms an inner-product space which contains $\mathcal{P}$ (isomorphically) as a subspace.[1] That is, any pre-Hilbert space can be extended to a Hilbert space. This justifies the use of the term <u>pre-Hilbert space</u> and explains the importance of these spaces in the theory of Hilbert space.

The only infinite-dimensional Hilbert space which has been considered is $l^2$ . The following theorem describes a more general Hilbert space which includes both $l^2$ and the unitary spaces as special cases.

THEOREM 17. (The Hilbert Space $l^2[Q]$ .) Let Q be any nonempty set of elements. Let $l^2[Q]$ denote the class of scalar-valued functions x defined on Q, such that the set of all $q \in Q$ for which $x(q) \neq 0$ is either finite or countably infinite, and having the property that

$\sum_{q \in Q} |x(q)|^2 < \infty$ . Define the sum of elements of $l^2[Q]$ and scalar multiplication as in Example 2, and for $x,y \in l^2[Q]$ , $(x|y) =$

$\sum_{q \in Q} x(q)\overline{y(q)}$ . Then $l^2[Q]$ is a Hilbert space.

Proof:

Since the union of two countable sets is itself a countable set, the sum of two elements of $l^2[Q]$ is an element of $l^2[Q]$ . Clearly $l^2[Q]$ is a linear vector space which forms a pre-Hilbert space with the

---

[1] See Angus E. Taylor, Introduction to Functional Analysis, pp. 74-75, 98-99, 119.

inner product as defined. Thus it is necessary only to show that $l^2[Q]$ is complete.

If Q is itself finite or countably infinite, then $l^2[Q]$ is isomorphic either to one of the spaces $\mathcal{C}^n$ or to $l^2$, and is therefore complete by Lemma 7. If on the other hand Q is not countable, then it must be shown that a given Cauchy sequence $\{x_n\}$ in $l^2[Q]$ converges to an element of $l^2[Q]$. By hypothesis, for $\mathcal{E} > 0$ there is an $N(\mathcal{E}) > 0$ such that

$$\|x_m - x_n\|^2 = (x_m - x_n | x_m - x_n) = \sum_{q \in Q} |x_m(q) - x_n(q)|^2 < \mathcal{E}$$

whenever $m,n > N$. In particular, for a fixed $q \in Q$, then

$$|x_m(q) - x_n(q)|^2 \leq \sum_{q \in Q} |x_m(q) - x_n(q)|^2 < \mathcal{E} \text{ whenever } n,m > N.$$

Thus the sequence of complex numbers $\{x_m(q)\}$ is a Cauchy sequence, and must therefore converge to an element of $\mathcal{C}$, which may be denoted $x(q)$. Now letting q range over Q, the scalar-valued function $x(q)$ is defined for all $q \in Q$. Let $P_i = \{q \in Q: x_i(q) \neq 0\}$, and $P = \{q \in Q: x(q) \neq 0\}$. If $q' \in P$, and if $x_n(q') = 0$ for an infinite number of n, then necessarily $x(q') = 0$, which contradicts the choice of $q'$. Thus $x_k(q') \neq 0$ for some k. Hence $P \subset \bigcup P_i$. Since the countable union of countable sets is itself countable, it follows that P is countable.

From the hypothesis $\sum_{q \in P_i} |x_i(q)|^2 < \infty$, where $P_i$ is countable, it follows that each $x_i(q)$ generates an absolutely square-summable sequence, an element of $l^2$. The completeness of $l^2$ assures that the sequence

generated by the limit $x(q)$, where $q \in P$, must also be an element of $l^2$.
Thus

$$\sum_{q \in P} |x(q)|^2 = \sum_{q \in Q} |x(q)|^2 < \infty \quad , \text{ and } x(q) \in l^2 [Q] .$$

Therefore $l^2 [Q]$ is complete and is a Hilbert space.

It is interesting to note that the set of functions $\{x_p(q): p \in Q\}$ defined by $x_p(q) = 0$ if $q \neq p$, $x_p(p) = 1$, is an orthonormal set having the same cardinal number as the set Q itself. Thus there exists a Hilbert space containing an orthonormal set whose cardinal number is any given infinite cardinal. If Q is an uncountably infinite set, then it can be shown that $l^2 [Q]$ is a Hilbert space which is not separable.

REFERENCES

Aronszajn, N. Introduction to the Theory of Hilbert Spaces, Vol. I. Stillwater, Oklahoma: Research Foundation, 1950

Berberbian, Sterling K. Introduction to Hilbert Space. New York: Oxford University Press, 1961.

Day, Mahlon M. Normed Linear Spaces. Berlin: Springer-Verlag, 1958.

Halmos, Paul R. Introduction to Hilbert Space. New York: Chelsea Publishing Company, 1951.

Kolomogorov, A. N., and S. V. Fomin. Measure, Lebesgue Integrals, and Hilbert Space. New York: Academic Press, 1961.

Taylor, Angus E. Introduction to Functional Analysis. New York: John Wiley and Sons, 1958.

ACKNOWLEDGMENT

The author wishes to express his sincere thanks and appreciation to Dr. Neal E. Foland for his assistance in the preparation of this report.

HILBERT SPACE

by

JOHN DAVID PERINE

B. S., Washburn University, 1964

---

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1965

The purpose of this report is to study the variety of linear vector space known as Hilbert space. The original example of this type of space was first studied extensively by the German mathematician David Hilbert, in connection with his work with integral equations. This space is known as classical Hilbert space. It was abstracted by J. von Neumann to the variety of linear vector space called Hilbert space.

The initial sections of the report contain basic properties of linear vector spaces including a particular type of complex-valued function, called an inner product, on pairs of elements of these spaces. The norm of each element is defined in terms of this function, and a metric on the space is defined in terms of this norm. As metric spaces, the vector spaces may or may not possess the topological property of completeness, in the sense that every Cauchy sequence of vectors converges in the metric to a vector of the space. The vector spaces with an inner product which possess this property of completeness are known as Hilbert spaces.

One of the most important concepts in the theory of inner-product spaces (called pre-Hilbert spaces), is that of orthogonality. Two vectors are said to be orthogonal if their inner product is zero. A sequence of vectors is said to be total provided that each vector has norm one, that each pair of distinct vectors is orthogonal, and that the only vector which is orthogonal to every vector of the sequence is the null vector. A Hilbert space containing such a sequence is said to be separable. Theorem 16 states that every finite-dimensional separable Hilbert space is isomorphic to the unitary space, described in Example 4, of the same dimension, and every infinite-dimensional separable Hilbert space is isomorphic to classical Hilbert space. The final theorem describes a class of Hilbert spaces defined on arbitrary sets Q. For particular Q, this space may be separable and either finite or infinite, or infinite and not separable.