A BRIEF INTRODUCTION TO HILBERT SPACE AND QUANTUM LOGIC

JOEL KLIPFEL

"We must know—we will know!"

-David Hilbert [5]

1. Introduction

Among his many contributions to the development of mathematics, the German mathematician David Hilbert (1862 - 1943) is known for his pioneering work in the field of functional analysis [6]. One of the cornerstones of functional analysis, the notion of a Hilbert space, emerged from Hilbert's efforts to generalize the concept of Euclidean space to an infinite dimensional space [7]. The theory of Hilbert space that Hilbert and others developed has not only greatly enriched the world of mathematics but has proven extremely useful in the development of scientific theories, particularly quantum mechanics. For example, the ability to treat functions as vectors in a Hilbert space, as permitted by Hilbert space theory, has enabled quantum physicists to solve difficult differential and integral equations by using mere algebra. What is more, the theory and notation of Hilbert space has become so ingrained in the world of quantum mechanics that it is commonly used to describe many interesting phenomenon, including the EPR paradox (entanglement), quantum teleportation, and quantum telecloning.

Unfortunately, much of the deep understanding behind Hilbert space theory is often lost in the translation from the mathematical world to the world of physicists. Given the importance of Hilbert space theory to quantum mechanics, a thorough mathematical understanding of the Hilbert space theory that underpins much of quantum mechanics will likely aid in the future development of quantum theory. As such, we explore some of the fundamentals of Hilbert space theory from the perspective of a mathematician and use our insights gained to begin an investigation of one mathematical formulation of quantum mechanics called quantum logic.

2. Prerequisites

As we begin our exploration of Hilbert space, the reader is assumed to have some background in linear algebra and real analysis. Nonetheless, for the sake of clarity, we begin with a discussion of three notions that are fundamental to the field of functional analysis, namely metric spaces, normed linear spaces, and inner product spaces.[1]

Few definitions are as fundamental to analysis as that of the metric space. In essence, a metric space is simply a collection of objects (e.g. numbers, matrices, pineapple flavored Bon Bons covered with flax seeds) with an associated rule, or function, that determines "distance" between two objects in the space. Such a function is termed a metric. Perhaps the most intuitive example of a metric space is the real number line with the associated metric $|x-y|$, for $x, y \in \mathbb{R}$. In general, though, a metric need only satisfy four basic criteria. More formally:

Definition 2.1 (**Metric Space).**
A *metric space* (X, d) is a set X together with an assigned metric function $d : X \times X \to \mathbb{R}$ that has the following properties:

Positive: $d(x, y) \geq 0$ for all $x, y, z \in X$,
Nondegenerate: $d(x, y) = 0$ if and only if $x = y$,
Symmetric: $d(x, y) = d(y, x)$ for all $x, y, z \in X$,
Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Whenever there is little risk for ambiguity, we denote a metric space (X, d) as simply X.

As we show in Theorems 2.1 and 2.2, the next two spaces for which we provide definitions, normed linear spaces and inner product spaces, are simply special cases of metric spaces. Moreover, an inner product space is a form of a normed linear space.

Definition 2.2 (**Normed Linear Space).**

A (complex) *normed linear space* $(L, \|\cdot\|)$ is a linear (vector) space with a function $\|\cdot\| : L \to \mathbb{R}$ called a norm that satisfies the properties:

Positive: $\|v\| \geq 0$ for all $v \in L$,
Nondegenerate: $\|v\| = 0$ if and only if $v = 0$.
Multiplicative: $\|\lambda v\| = |\lambda| \, \|v\|$ for all $v \in L$ and $\lambda \in \mathbb{C}$,
Triangle Inequality: $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in L$.

Recall that a complex conjugate of $a \in \mathbb{C}$ is often denoted as $\overline{a}$. We use this notation throughout the remainder of this paper.

Definition 2.3 (**Inner Product Space).**
If V is a linear space, then a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ is said to be an *inner product* provided that

Positive: $\langle v, v \rangle \geq 0$, for all $v \in V$
Nondegenerate: $\langle v, v \rangle = 0$ if and only if $v = 0$.

[1] The study of Hilbert space theory is a subset of the field of functional analysis.

Multiplicative: $\langle \lambda u, v\rangle = \lambda \langle u, v\rangle$, for all $u, v \in V$ and $\lambda \in \mathbb{C}$,
Symmetric: $\langle u, v\rangle = \overline{\langle v, u\rangle}$, whenever $u, v \in V$,
Distributive: $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$, for all $u, w, v \in V$.

A linear space V is defined to be the *inner product space* $(V, \langle\cdot,\cdot\rangle)$ if it has an inner product defined on it.

The symmetry criterion in Definition 2.3 is sometimes referred to as *Hermitian* symmetry.

For the sake of expediency, a normed linear space $(L, \|\cdot\|)$ is often denoted as L. Likewise, an inner product space $(V, \langle\cdot,\cdot\rangle)$ is commonly denoted V.

In the following two theorems, we formalize our assertions about the relationships between metric spaces, normed linear spaces, and inner product spaces.

Theorem 2.1.
A normed linear space $(L, \|\cdot\|)$ is a metric space with metric d given by

$$d(v, w) = \|v - w\|,$$

where $v, w \in L$.

Proof. That d satisfies the positive and nondegeneracy requirements of metrics, can be seen as an immediate consequence of the positive and nondegeneracy properties of norms. The multiplicative property of normed spaces allows us to make the following simple calculation:

$$d(v, w) = \|v - w\| = \|(-1)(w - v)\| = |-1| \cdot \|w - v\| = \|w - v\| = d(w, v),$$

which demonstrates that d is symmetric.

Finally, we need to show that d satisfies the triangle inequality. To do so, we choose any three $x, y, z \in L$. Since the vectors $x - z$ and $z - y$ are in L, the triangle inequality of norms allows us to see that

$$\begin{aligned} d(x, y) &= \|x - y\| \\ &= \|(x - z) + (z - y)\| \\ &\leq \|x - z\| + \|z - y\| \\ &= d(x, z) + d(z, y) \end{aligned}$$

Hence, since d is positive, nondegenerate, symmetric and satisfies the triangle inequality, we conclude that it is indeed a metric. □

Theorem 2.2.
An inner product space $(V, \langle\cdot,\cdot\rangle)$ is a normed linear space with norm

$$\|v\| = \sqrt{\langle v, v\rangle},$$

for all $v \in V$.

Proof. The positive and nondegenerate properties of inner products guarantees that $\|\cdot\|$ also has these properties.

A simple consequence of *Hermitian* symmetry ($\langle v, w\rangle = \overline{\langle w, v\rangle}$) and the multiplicative property of inner products is the fact that $\langle v, \lambda w\rangle = \overline{\lambda}\langle v, w\rangle$ whenever $v, w \in V$ and $\lambda \in \mathbb{C}$. Using this equality, we see that

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda\overline{\lambda}}\sqrt{\langle v, v\rangle} = |\lambda| \cdot \|v\|.$$

To show that $\|\cdot\|$ satisfies the triangle inequality criterion of norms, we utilize the Cauchy-Schwartz inequality, which states that if $v, w \in V$, then $|\langle v, w\rangle| \leq \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}$. Since both Karen Saxe [9] and Carol Schumacher [10] provide elegant proofs of the Cauchy-Schwarz inequality, we will use the Cauchy-Schwarz inequality without providing a proof in this paper.

Using the distributive property of inner products, we see that for $v, w \in V$,

$$\|v + w\|^2 = \langle v, v\rangle + \langle w, w\rangle + \langle v, w\rangle + \langle w, v\rangle .$$

According to the Cauchy-Schwarz inequality then,

$$\begin{aligned} \|v + w\|^2 &\leq \|v\|^2 + \|w\|^2 + \|v\|\|w\| + \|w\|\|v\| \\ &= (\|v\| + \|w\|)^2 . \end{aligned}$$

So, $\|v + w\| \leq \|v\| + \|w\|$. We conclude that $\|\cdot\|$ is a norm on V. $\square$

Before we delve too far into our exploration of inner product spaces and normed linear spaces, it is useful to pause to examine a couple of interesting examples.

Example 2.1. In this first example we explore the set of all real, bounded sequences, often termed ℓ^∞, and show that ℓ^∞ is a normed linear space with norm

$$\|(x_n)\|_\infty = \sup\{|x_n| : n \in \mathbb{N}\}.$$

Since the absolute value function maps from $\mathbb{R}$ to $\mathbb{R}_+ \cup \{0\}$ (the set of all non-negative real numbers), the function $\|\cdot\|_\infty$ is always positive. Further, if $\|(x_n)\|_\infty = 0$ for some $(x_n) \in \ell^\infty$, then each term in (x_n) must equal zero, as the least upper bound of the set $\{|x_n| : n \in N\}$ is zero. Similarly, if a sequence $(x_n) \in \ell^\infty$ is the zero sequence, then, by definition, the supremum of the set $\{x_n\}$ must equal zero, which implies that $\|\cdot\|_\infty$ is also non-degenerate.

The basic properties of the supremum and absolute value allow us to further see that for sequences $(x_n), (y_n) \in \ell^\infty$ and any complex number λ,

$$\begin{aligned} \|(\lambda x_n)\|_\infty &= \sup\{|\lambda x_n| : n \in \mathbb{N}\} \\ &= \sup\{|\lambda||x_n| : n \in \mathbb{N}\} \\ &= |\lambda| \sup\{|x_n| : n \in \mathbb{N}\} \\ &= |\lambda| \, \|(x_n)\|_\infty , \end{aligned}$$

and

$$\begin{aligned} \|(x_n + y_n)\|_\infty &= \sup\{|x_n + y_n| : n \in \mathbb{N}\} \\ &\leq \sup\{|x_n| + |y_n| : n \in \mathbb{N}\} \\ &\leq \sup\{|x_n| : n \in \mathbb{N}\} + \sup\{|y_m| : m \in \mathbb{N}\} \\ &= \|(x_n)\| + \|(x_n)\| . \end{aligned}$$

Since $\|(x_n)\|_\infty$, satisfies all criteria for a norm, ℓ^∞ is a normed linear space.

Example 2.2. Another important example of a normed linear space is the collection of all continuous functions on a closed interval $[a, b]$, denoted $\mathcal{C}[a, b]$, with the supremum norm

$$\|f\|_\infty = \sup\{|f(x)| : x \in [a, b]\}.$$

An analogous argument to the one given above for ℓ^∞ demonstrates that $\mathcal{C}[a, b]$ with norm $\|f\|_\infty$ is indeed a normed linear space.

3. Introduction to the ℓ^p spaces

A natural question to ask at this point is: "Are the definitions of of metric space, normed linear space, and inner product space equivalent?" That is, now that we have shown

$$\text{inner product space} \Rightarrow \text{normed linear space} \Rightarrow \text{metric space},$$

can we reverse the direction of these implications? In short, no, we cannot do so. Consider, for example, the set $T = \mathbb{R}/\{0\}$. While this set is certainly a metric space with metric $d(j, k) = |j - k|$, for $j, k \in T$, it lacks an additive unity (i.e. $0 \notin T$), and therefore is not even a linear space, let alone a normed linear space or an inner product space. That said, we still need to address the question: does normed linear space imply inner product space? Again, the answer is a very definitive no. In order to verify this claim, we will study the historically significant set of l^p spaces (pronounced "little ell pea") in Section 4. However, our study of the ℓ^p spaces requires the use of the Parallelogram Law, to which we now turn our attention.

3.1. **Parallelogram Law.** The proof of the Parallelogram Law for complex normed linear spaces proceeds in nearly an identical fashion as in the real case−only messier. For this reason, we omit the more general proof of the Parallelogram Law for complex normed linear spaces and instead provide the more instructive proof of the Parallelogram Law for real linear vector spaces.

Lemma 3.1.
Let L be a normed linear space. The norm $\|\cdot\| : L \to \mathbb{R}$ is continuous.

Proof. Let $\varepsilon > 0$ and choose $\delta = \varepsilon$. Now pick a point a of L and any $x \in L$ such that $d(x, a) = \|x - a\|$. The reverse triangle inequality implies that

$$d(\|x\|, \|a\|) = \big|\|x\| - \|a\|\big| \leq \|x - a\| < \varepsilon.$$

Thus $\|\cdot\|$ is continuous on L. □

Theorem 3.1 (**Parallelogram Law**).
A normed linear space L is an inner product space if and only if its associated norm ($\|\cdot\|$) satisfies the parallelogram equality

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2,$$

for every $u, v \in L$.

Proof. Suppose L is an inner product space with associated norm $\|w\| = \sqrt{\langle w, w\rangle}$, for all $w \in L$. If $u, v \in L$, then

$$\begin{aligned}
\|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v\rangle + \langle u - v, u - v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&\quad + \langle u, u\rangle + \langle u, -v\rangle + \langle -v, u\rangle + \langle -v, -v\rangle \\
&= 2\langle u, u\rangle + 2\langle v, v\rangle + (\langle u, v\rangle - \langle u, v\rangle) \\
&\quad + (\langle v, u\rangle - \langle v, u\rangle) \\
&= 2\|u\|^2 + 2\|v\|^2.
\end{aligned}$$

Thus $\|\cdot\|$ satisfies the parallelogram equality.

Now suppose that L is a normed space whose norm satisfies the parallelogram equality. Define a function $\langle\cdot,\cdot\rangle : L \times L \to \mathbb{R}$ by

$$\langle u, v\rangle = \frac{1}{4}\left(\|u + v\|^2 - \|u - v\|^2\right),$$

for all $u, v \in L$. We wish to show that $\langle\cdot,\cdot\rangle$ is an inner product. We begin by observing that $\langle u, u\rangle = \frac{1}{4}\left(\|2u\|^2 - \|0\|^2\right)$, for all $u \in L$. So $\langle\cdot,\cdot\rangle$ is non-negative and non-degenerate. Further, since

$$\langle u, v\rangle = \frac{1}{4}\left(\|u + v\|^2 - \|u - v\|^2\right) = \frac{1}{4}\left(\|v + u\|^2 - \|v - u\|^2\right) = \langle v, u\rangle,$$

the function $\langle \cdot, \cdot \rangle$ is also symmetric.

To see that $\langle \cdot, \cdot \rangle$ satisfies the multiplicative property, notice that for any $u, v, w \in L$, the parallelogram equality implies

$$2\left\|v+u\right\|^2 + 2\left\|w\right\|^2 = \left\|(v+u)+w\right\|^2 + \left\|(v+u)-w\right\|^2, \tag{1}$$

and

$$2\left\|v-u\right\|^2 + 2\left\|w\right\|^2 = \left\|(v-u)+w\right\|^2 + \left\|(v-u)-w\right\|^2. \tag{2}$$

By subtracting equation (2) from equation (1), we see that

$$\begin{aligned} 2\left(\left\|v+u\right\|^2 - \left\|v-u\right\|^2\right) &= \left\|(v+w)+u\right\|^2 - \left\|(v+w)-u\right\|^2 \\ &\quad + \left\|(v-w)+u\right\|^2 - \left\|(v-w)-u\right\|^2, \end{aligned}$$

which can be rewritten as

$$2\langle v, u\rangle = \langle v+w, u\rangle + \langle v-w, u\rangle. \tag{3}$$

If $w = v$, then $2\langle v, u\rangle = \langle 2v, u\rangle$, as

$$\langle 0, v\rangle = \langle u \cdot 0, v, =\rangle \frac{1}{4}\left(\left\|0+v\right\| - \left\|0-v\right\|\right) = 0.$$

We now use induction to demonstrate that the multiplicative property holds for all positive integers. Suppose there exists some integer n so that if $n > 2$, then for every i, where $1 \leq i \leq n$, the equality $i\langle v, u\rangle = \langle iv, u\rangle$ holds. So, by an immediate consequence of Equation (3),

$$\begin{aligned} \langle (n+1)v, u\rangle &= \langle nv+v, u\rangle \\ &= 2\langle nv, u\rangle - \langle nv-v, u\rangle \\ &= 2n\langle v, u\rangle - (n-1)\langle v, u\rangle \\ &= (n+1)\langle v, u\rangle. \end{aligned}$$

Hence $\langle \cdot, \cdot \rangle$ satisfies the multiplicative criteria for all positive integers. Since

$$\begin{aligned} \langle -u, v\rangle &= \frac{1}{4}\left(\left\|-u+v\right\| - \left\|-u-v\right\|\right) \\ &= \frac{1}{4}\left(\left\|u-v\right\| - \left\|u+v\right\|\right) \\ &= -1 \cdot \langle u, v\rangle, \end{aligned}$$

the function $\langle \cdot, \cdot \rangle$ satisfies the multiplicative property for all integers.

Suppose that r is any rational number. Then there exist integers n and m such that $r = \frac{m}{n}$. Since $n\frac{m}{n} = m$, then $n\frac{m}{n}\langle u, v\rangle = \left\langle n\frac{m}{n}u, v\right\rangle = n\left\langle \frac{m}{n}u, v\right\rangle$, which implies that $\frac{m}{n}\langle u, v\rangle = \left\langle \frac{m}{n}u, v\right\rangle$, or, equivalently, $r\langle u, v\rangle = \langle ru, v\rangle$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for any x in $\mathbb{R}$, we can find a sequence (r_n) of rational numbers that converges to x. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ so that $|x_n - x| < \frac{\varepsilon}{|\langle u,v\rangle|+1}$ for all $n > N$. So, since the co-domain of $\langle \cdot, \cdot \rangle$ is $\mathbb{R}$ (as the co-domain of any norm is $\mathbb{R}$), we see that for any $u, v \in L$,

$$\begin{aligned} d(\langle x_n u, v\rangle, x\langle u, v\rangle) &= |\langle x_n u, v\rangle - x\langle u, v\rangle| \\ &= |x_n \langle u, v\rangle - x\langle u, v\rangle| \\ &= \langle u, v\rangle |x_n - x| \\ &< \varepsilon. \end{aligned}$$

Lemma 3.1 implies that $\lim_{n\to\infty} \langle x_n u, v\rangle = \langle xu, v\rangle$. Therefore, by the uniqueness of limits, we conclude that $\langle xu, v\rangle = x\langle u, v\rangle$.

The final step in this proof is showing that $\langle \cdot, \cdot \rangle$ also satisfies the distributive property. Choose any three members u, v, w of L, and pick $x, y \in L$ given by

$$x = \frac{u+w}{2} \qquad \text{and} \qquad y = \frac{u-w}{2}.$$

Recall that Equation (3) implies $\langle 2x, v\rangle = \langle x+y, v\rangle + \langle x-y, v\rangle$. So,

$$\begin{aligned} \langle u+w, v\rangle &= \langle 2x, v\rangle \\ &= \langle x+y, v\rangle + \langle x-y, v\rangle \\ &= \langle u, v\rangle + \langle w, v\rangle. \end{aligned}$$

Given $\langle \cdot, \cdot \rangle$ satisfies all five criteria of an inner product space, we conclude that L is an inner product space. □

On page 248 of his book *Introduction to Topology and Modern Analysis*, George Simmons proves the Parallelogram Law for a complex normed linear space L by using the inner product $\langle u, v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2 + i\|u+v\|^2 - i\|u-v\|^2\right)$ [11]. Hence, for the remainder of this paper, we assume that the Parallelogram Law is valid for complex linear normed vector spaces.

3.2. **Introduction to the ℓ^p spaces.** The ℓ^p spaces are a class of normed linear spaces that seem to hold a special place in the hearts of many functional analysts. For each $p \in \mathbb{R}$, where $p > 1$, ℓ^p is the set of all sequences (x_n) such that the series

$$\sum_{n=1}^{\infty} |x_n|^p$$

is convergent. The standard norm $\|\cdot\|_p$ for ℓ^p is given by

$$\|(x_n)\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}},$$

for each $(x_n) \in \ell^p$ [9]. We begin our exploration of the ℓ^p spaces by first considering two specific examples, namely the ℓ^1 and ℓ^2 spaces.

Example 3.1. In this example, we show that the linear space ℓ^1 is a normed linear space but not an inner product space. We begin by showing that ℓ^1 is a normed linear space. The properties of the absolute value function and summations ensure us that $\|\cdot\|_1$ is both nonnegative and nondegenerate. To see that $\|\cdot\|_1$ is multiplicative, notice that for any $\lambda \in \mathbb{C}$ and $(x_n) \in \ell^1$,

$$\|(\lambda x_n)\|_1 = \sum_{n=1}^{\infty} |\lambda x_n| = \sum_{n=1}^{\infty} |\lambda|\ |x_n| = |\lambda| \sum_{n=1}^{\infty} |x_n| = |\lambda|\, \|(x_n)\|_1 .$$

Finally, the triangle inequality of the absolute value function allows us to make the calculation

$$\|(x_n + y_n)\|_1 = \sum_{n=1}^{\infty} |x_n + y_n| \leq \sum_{n=1}^{\infty} |x_n| + \sum_{n=1}^{\infty} |y_n| = \|(x_n)\|_1 + \|(y_n)\|_1 ,$$

for any $(x_n), (y_n) \in \ell^1$. Hence, $\|\cdot\|_1$ also satisfies the triangle inequality. Since $\|\cdot\|_1$ satisfies all four criteria of a norm, ℓ^1 is a normed linear space.

To see that ℓ^1 is not an inner product space, we will examine the counterexample of $(x_n) = \left(\frac{1}{2^{n-1}}\right)$ and $(y_n) = \left(\frac{2}{3^{n-1}}\right)$. Since both (x_n) and (y_n) are geometric, a simple set of calculations shows that $\|(x_n)\|_1 = 2$, $\|(y_n)\|_1 = 3$, $\|(x_n - y_n)\|_1 = \frac{4}{3}$, and $\|(x_n + y_n)\|_1 = 5$. Further arithmetic calculations lead to the conclusion that $2\,\|(x_n)\|_1^2 + 2\,\|(y_n)\|_1^2 = 22$, but $\|(x_n + y_n)\|_1^2 + \|(x_n - y_n)\|_1^2 = \frac{239}{9}$, in violation of the parallelogram equality. We thusly conclude that ℓ^1 is not an inner product space.

We now have verification that an inner product space is a special case of normed space and that the term “inner product space” is not simply an synonym for “normed space.”

Example 3.2. Here we show that the linear space ℓ^2 is an inner product space as well as a normed linear space. Since ℓ^2 is an ℓ^p space, we assume that ℓ^2 is a normed space and then use the parallelogram equality to show that ℓ^2 is an inner product space as well.

Suppose $(a_n), (b_n) \in \ell^2$. Then

$$\begin{aligned}
\|(a_n+b_n)\|_2^2 + \|(a_n-b_n)\|_2^2 &= \sum_{n=1}^{\infty} |a_n+b_n|^2 + \sum_{n=1}^{\infty} |a_n-b_n|^2 \\
&= \sum_{n=1}^{\infty} \left(|a_n+b_n|^2 + |a_n-b_n|^2\right) \\
&= \sum_{n=1}^{\infty} \left(a_n^2 + 2a_nb_n + b_n^2 + a_n^2 - 2a_nb_n + b_n^2\right) \\
&= \sum_{n=1}^{\infty} \left(2a_n^2 + 2b_n^2\right) \\
&= 2\sum_{n=1}^{\infty} |a_n|^2 + 2\sum_{n=1}^{\infty} |b_n|^2 \\
&= 2\,\|(a_n)\|_2^2 + 2\,\|(b_n)\|_2^2 .
\end{aligned}$$

Since $\|\cdot\|_2$ satisfies the Parallelogram Law, ℓ^2 is an inner product space.

At this point, it is worth pausing to note the significance of the ℓ^2 space. As we demonstrate in the next section ℓ^2 is the only one of the ℓ^p spaces that happens to be an inner product space. Moreover, as Karen Saxe notes in her textbook *Beginning Functional Analysis*, this "space" of sequence played an important role in the naming of Hilbert space. The notion of the ℓ^p space comes from none other than David Hilbert, and Hilbert was known to discuss the ℓ^p spaces during some of his lectures. Owing to the ℓ^2 space's uniqueness from its peers, Hilbert's students often referred to it as Hilbert's space, from which we have the term Hilbert space [9].

4. Properties of the ℓ^p Spaces

4.1. **Minkowski's Inequality.** In the previous section, we claimed that ℓ^2 is the only ℓ^p space that is an inner product space. Since this is such an interesting property, we formalize this claim by turning it into a theorem in Subsection 4.2. Before doing so, however, we examine one property that is common to all ℓ^p spaces: they are all normed linear spaces. The proof of this claim relies on Minkowski's Inequality. The method we use to prove Minkowski's Inequality involves concave functions, and is based on Lemma 4.1. The reader may recall that:

Definition 4.1 (Concave Function).
A function $g : \mathbb{R}_+ \to \mathbb{R}$ is said to be *concave* if it satisfies the inequality

$$\lambda g(x) + (1-\lambda)g(y) \leq g(\lambda x + (1-\lambda)y),$$

for all $x, y \in \mathbb{R}_+$ and all λ that satisfy $0 \leq \lambda \leq 1$.

The following figure demonstrates a geometric interpretation of a concave function.

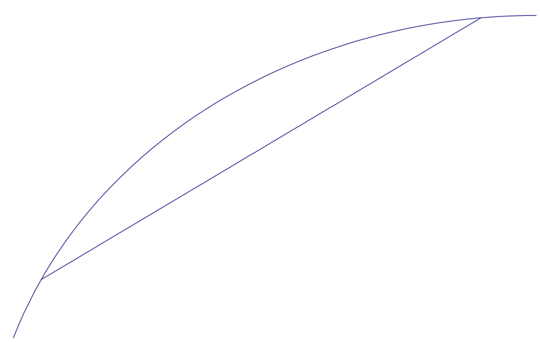

FIGURE 1. A geometric interpretation of a concave function. The curve represents $\lambda g(x)+(1-\lambda)g(y)$ and the line segment represents $g(\lambda x+(1-\lambda)y$.

Lemma 4.1.
Let $g : \mathbb{R}_+ \to \mathbb{R}$ be a concave function and define a function $f : \mathbb{R}_+^2 \to \mathbb{R}$ by

$$f(x, y) = y\ g(x/y).$$

Then the inequality

$$\sum_{i=1}^{n} f(x_i, y_i) \le f\left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i\right)$$

is valid for all positive real numbers $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.

Proof. We proceed by induction. Let $\lambda = y_1/(y_1 + y_2)$. Then the base case is as follows:

$$\begin{aligned} f(x_1, y_1) + f(x_2, y_2) &= y_1 g(x_1/y_1) + y_2 g(x_2/y_2) \\ &= (y_1 + y_2)\left(\frac{y_1}{y_1 + y_2} g(x_1/y_1) + \frac{y_2}{y_1 + y_2} g(x_2/y_2)\right) \\ &\le (y_1 + y_2) g\big((x_1 + x_2)/(y_1 + y_2)\big) \\ &= f(x_1 + x_2, y_1 + y_2), \end{aligned}$$

as $1 - \lambda = \frac{y_2}{y_1+y_2}$. Suppose that the inequality

$$\sum_{i=1}^{k} f(x_i, y_i) \le f\left(\sum_{i=1}^{k} x_i, \sum_{i=1}^{k} y_i\right)$$

holds for all positive integers k strictly less than some positive integer n. Then

$$\begin{aligned}
\sum_{i=1}^{n} f(x_i, y_i) &= \sum_{i=1}^{n-1} f(x_i, y_i) + f(x_n, y_n) \\
&\leq f\left(\sum_{i=1}^{n-1} x_i, \sum_{i=1}^{n-1} y_i\right) + f(x_n, y_n) \\
&\leq f\left(x_n + \sum_{i=1}^{n-1} x_i, y_n + \sum_{i=1}^{n-1} y_i\right) \\
&= f\left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i\right),
\end{aligned}$$

precisely as we wished. □

We are now ready to tackle Minkowski's Inequality:

Theorem 4.1 (**Minkowski's Inequality**).
Let $(x_i), (y_i)$ be points in ℓ^p, where $p > 1$. Then the following inequality holds:

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{\frac{1}{p}}.$$

Proof. We break this proof into three cases:

(1) (x_i) and (y_i) contain only positive terms,
(2) (x_i) and (y_i) contain no zero terms, and
(3) (x_i) and (y_i) contain positive, negative, and zero terms.

Suppose sequences (x_i) and (y_i) contain only strictly positive terms. Define new sequences (s_i) and (t_i) by $s_i = x_i^p$ and $t_i = y_i^p$ for all $i \in \mathbb{N}$. Consider the function $g(s) = (1 + s^{1/p})^p$. Since $g''(s) < 0$, we know that g is concave.[2] Now define a new function $f : \mathbb{R}_+^2 \to \mathbb{R}$ by $f(s,t) = y\, g(s/t)$. With some algebraic manipulation we see that $f(s,t) = \left(s^{1/p} + t^{1/p}\right)^p$. So, if we apply Lemma 4.1, we then see that

$$\sum_{i=1}^{n} \left(s_i^{1/p} + t_i^{1/p}\right)^p \leq \left(\left(\sum_{i=i}^{n} s_i\right)^{\frac{1}{p}} + \left(\sum_{i=i}^{n} t_i\right)^{\frac{1}{p}}\right)^p. \tag{4}$$

By taking the p^{th} root of each side of the inequality (4) and substituting (x_i) and (y_i) into (4), we obtain the inequality

$$\left(\sum_{i=1}^{n} (x_i + y_i)^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=i}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=i}^{n} y_i^p\right)^{\frac{1}{p}}. \tag{5}$$

[2]See http://en.wikipedia.org/wiki/Concave_functions for more information [4].

Since (x_i) and (y_i) are strictly positive sequences and (5) is valid for all $n \in \mathbb{N}$, the inequality (5) is precisely the Minkowski inequality for this case:

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=i}^{\infty} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=i}^{\infty} |y_i|^p\right)^{\frac{1}{p}}. \tag{6}$$

Now suppose that (x_i) and (y_i) contain no zero terms. The triangle inequality implies that for all $i \in \mathbb{N}$, $\big|x_i+y_i\big| \leq \big||x_i|+|y_i|\big|$. This, in turn, implies that $\big|x_i+y_i\big|^p \leq \big||x_i|+|y_i|\big|^p$. So

$$\left(\sum_{i=1}^{\infty} \big|x_i + y_i\big|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} \big||x_i| + |y_i|\big|^p\right)^{\frac{1}{p}}.$$

Since

$$\left(\sum_{i=1}^{\infty} \big||x_i| + |y_i|\big|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=i}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p}$$

by Inequality (6),

$$\left(\sum_{i=1}^{\infty} \big|x_i + y_i\big|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{\frac{1}{p}}. \tag{7}$$

Finally, suppose that (x_i) and (y_i) contain zero terms. Then we will need to construct new sequences $(x_{i,m})$ and $(y_{i,m})$ as follows:

$$\begin{aligned} x_{i,m} &= \begin{cases} x_i & \text{if } x_i \neq 0 \\ \frac{1}{m} \cdot \frac{1}{i} & \text{if } x_i = 0 \end{cases} \\ y_{i,m} &= \begin{cases} y_i & \text{if } y_i \neq 0 \\ \frac{1}{m} \cdot \frac{1}{i} & \text{if } y_i = 0 \end{cases}. \end{aligned}$$

Notice that $(x_{i,m}), (y_{i,m}) \in \ell^p$, as $\left(\frac{1}{i}\right) \in \ell^p$ for all $p > 1$. Now consider the sequences

$$(\sigma_m) = \left(\left(\sum_{i=1}^{\infty} \big|x_{i,m} + y_{i,m}\big|^p\right)^{\frac{1}{p}}\right)$$

and

$$(\tau_m) = \left(\left(\sum_{i=1}^{\infty} |x_{i,m}|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_{i,m}|^p\right)^{\frac{1}{p}}\right).$$

Inequality (7) implies that $\sigma_m \leq \tau_m$ for all $m \in \mathbb{N}$. Hence

$$\lim_{m\to\infty} \sigma_m \leq \lim_{m\to\infty} \tau_m.$$

Fortunately,

$$\lim_{m\to\infty} \sigma_m = \left(\sum_{i=1}^{\infty} \left| x_i + y_i \right|^p \right)^{\frac{1}{p}}$$

and

$$\lim_{m\to\infty} \tau_m = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_i|^p \right)^{\frac{1}{p}}.$$

We therefore conclude that the Minkowski inequality holds for all $(x_i), (y_i) \in \ell^p$. $\square$

When verifying that the standard norm for ℓ^p is, in fact, a norm, the bulk of the work is done simply proving Minkowski's inequality. Minkowski's inequality tells us that $\|\cdot\|_p$ satisfies the triangle inequality (indeed, Minkowski's inequality can be thought of as the triangle inequality for ℓ^p). As with the ℓ^1 norm, the properties of the absolute value function guarantees that $\|\cdot\|_p$ is both nonnegative and non-degenerate. So, we now need only to show that $\|\cdot\|_p$ satisfies the multiplicative property of norms. Choose any $\lambda \in \mathbb{C}$ and any sequence $(x_n) \in \ell^p$. Then

$$\begin{aligned} \|(\lambda x_n)\|_p &= \left(\sum_{n=1}^{\infty} |\lambda x_n|^p \right)^{\frac{1}{p}} \\ &= \left(\sum_{n=1}^{\infty} |\lambda|^p |x_n|^p \right)^{\frac{1}{p}} \\ &= \left(|\lambda|^p \sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} \\ &= |\lambda| \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} \\ &= |\lambda| \cdot \|(x_n)\|_p, \end{aligned}$$

which is sufficient justification to conclude that all ℓ^p spaces are normed linear spaces.

4.2. **Properties of the ℓ^p Spaces.** As promised, we now prove that ℓ^2 is the only ℓ^p space that is an inner product space.

Theorem 4.2.
If p is any real number strictly greater than 1, then ℓ^p is an inner product space if and only if $p = 2$.

Proof. We already know from Example 3.2 that ℓ^2 is an inner product space. So, we now show that $p = 2$ is the only value of p for which ℓ^p is an inner product. Consider the sequences

$$(x_n) = \{1, 0, 0, 0, \ldots, 0, 0, \ldots\} \qquad \text{and} \qquad (y_n) = \{0, 1, 0, 0, \ldots, 0, 0, \ldots\},$$

where both sequences contain exactly one term equal to 1 with the remaining terms equal to zero. Observe that $(x_n), (y_n) \in \ell^p$ and $\sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} |y_n|^p = 1$ for all p. So, by plugging (x_n) and (y_n) into the parallelogram equality, we see that

$$\begin{aligned} 2\left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{2}{p}} + 2\left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{2}{p}} &= \left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{\frac{2}{p}} + \left(\sum_{n=1}^{\infty} |x_n - y_n|^p\right)^{\frac{2}{p}} \\ 2 \cdot 1^{2/p} + 2 \cdot 1^{2/p} &= 2^{2/p} + 2^{2/p} \\ 2 &= 2^{2/p}. \end{aligned}$$

Since $2 = 2^{2/p}$ only when $p = 2$, we conclude by the Parallelogram Law that ℓ^2 is the only ℓ^p space that is an inner product space. □

We now know a little about the ℓ^p spaces. So, are there any more interesting properties that we have yet to uncover about them? As is often the case with this sort of question in mathematics, the answer is a very emphatic yes. All ℓ^p spaces share a property that we have not yet discussed. Specifically, all ℓ^p spaces are complete. The reader may recall from an introductorty real analysis course that Cauchy sequences and complete metric spaces are defined in the following manner:

Definition 4.2 (Cauchy Sequences).
Let (a_n) be a sequence in a metric space (X, d). The sequence (a_n) is said to be *Cauchy* if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(a_n, a_m) < \varepsilon$ whenever $n, m > N$.

Definition 4.3 (Complete Metric Space).
If every Cauchy sequence in a metric space X converges to an element of X, then X is said to be *complete*.

Proving that all ℓ^p spaces are complete can be a little tricky, as we are essentially proving that a Cauchy sequence of sequences converges to a sequence that is in ℓ^p, which means that if we are not careful, sequence notation in such a proof can quickly become overwhelming. In that light, we now provide a brief discussion on the sequence notation that we use during the proof of Theorem 4.3. Consider a Cauchy sequence (a_n) in ℓ^p. In order to help dissipate possible confusion, we will denote the sequence (a_n) as $\left(\alpha_i^{(n)}\right)$ and use the index n to refer to each member sequence in (a_n) and the index i to refer to the individual elements in a given member sequence. In other words, for a fixed m, the sequence $\left(\alpha_i^{(m)}\right)$ is an element of ℓ^p and, for a fixed k, the number $\alpha_k^{(m)}$ is an element of $\mathbb{C}$.

Theorem 4.3.
All ℓ^p spaces are complete.

Proof. Suppose $(a_n) = \left(\alpha_i^{(n)}\right)$ is a Cauchy sequence in ℓ^p. Let $\varepsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that for all $n, m > N$,

$$\left(\sum_{i=1}^{\infty} \left|\alpha_i^{(n)} - \alpha_i^{(m)}\right|^p\right)^{\frac{1}{p}} < \varepsilon. \tag{8}$$

It follows from Equation (8) that for all fixed $k \in \mathbb{N}$, the sequence $\left(\alpha_k^{(n)}\right)_{n=1}^{\infty}$ is Cauchy. Since each $\left(\alpha_k^{(n)}\right)$ is a Cauchy sequence of complex numbers, it converges to a point in $\mathbb{C}$. Define a new sequence (b_i) such that b_1 is the limit of $\left(\alpha_1^{(n)}\right)$, the scalar b_2 is the limit of $\left(\alpha_2^{(n)}\right)$, and so forth.

In order to show that $(b_i) \in \ell^p$, first observe that Equation (8) implies that for all $j \in \mathbb{N}$,

$$\sum_{i=1}^{j} \left|\alpha_i^{(n)} - a_i^{(m)}\right|^p < \varepsilon^p.$$

So

$$\lim_{m\to\infty} \sum_{i=1}^{j} \left|\alpha_i^{(n)} - \alpha_i^{(m)}\right|^p = \sum_{i=1}^{j} \left|\alpha_i^{(n)} - b_i\right|^p < \varepsilon^p.$$

Since $\sum_{i=1}^{j} \left|\alpha_i^{(n)} - b_i\right|^p < \varepsilon^p$ for all $j \in \mathbb{N}$, we see that

$$\lim_{j\to\infty} \sum_{i=1}^{j} \left|\alpha_i^{(n)} - b_i\right|^p < \varepsilon^p.$$

This means that for any given $\varepsilon > 0$, we can find an $N \in \mathbb{N}$ so that if $n > N$, then $\|a_n - b\|_p < \varepsilon$, where a_n denotes $\left(\alpha_i^{(n)}\right)_{i=1}^{\infty}$ and b denotes (b_i). Therefore, $a_n \to b$. To show that $b \in \ell^p$, fix $L \in \mathbb{N}$ so that $L > N$. Then $a_L - b \in \ell^p$, as

$$\sum_{i=1}^{\infty} \left|a_i^{(L)} - b\right|^p$$

converges. Since ℓ^p is a vector space, we know that $a_L - (a_L - b)$ is also an element of ℓ^p. Hence $b \in \ell^p$. $\square$

If the definition of (b_i) in the proof of Theorem 4.3 seems confusing, consider the following diagram.

$$\begin{array}{cccccccc}
\alpha_1^{(1)} & \alpha_1^{(2)} & \alpha_1^{(3)} & \alpha_1^{(4)} & \cdots & \to & b_1 \\
\alpha_2^{(1)} & \alpha_2^{(2)} & \alpha_2^{(3)} & \alpha_2^{(4)} & \cdots & \to & b_2 \\
\alpha_3^{(1)} & \alpha_3^{(2)} & \alpha_3^{(3)} & \alpha_3^{(4)} & \cdots & \to & b_3 \\
\alpha_4^{(1)} & \alpha_4^{(2)} & \alpha_4^{(3)} & \alpha_4^{(4)} & \cdots & \to & b_4 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \to & \vdots
\end{array}$$

The first column in the above diagram represents the first element in the Cauchy sequence (a_n), the second column represents the second element in (a_n), etc. In other words, the elements of the Cauchy sequence (a_n) are represented vertically. However, when we define (b_i), we are interested in the horizontal sequences displayed in the above diagram which, unlike their vertical counterparts, need not in general be elements of ℓ^p. So, b_1, for example, is defined as the limit of the sequence $\left(\alpha_i^{(n)}\right) = \{\alpha_1^{(1)}, \alpha_1^{(2)}, \alpha_1^{(3)}, \alpha_1^{(4)}, \ldots\}$, the number b_2 is the limit of the sequence $\left(\alpha_2^{(n)}\right) = \{\alpha_2^{(1)}, \alpha_2^{(2)}, \alpha_2^{(3)}, \alpha_2^{(4)}, \ldots\}$, and so on.

In the last line of the proof of Theorem 4.3, we invoke the fact that the ℓ^p spaces are vector spaces, a point which we have not yet explicitly mentioned; though we tacitly assume it in earlier proofs. The reason for this omission is the trivial nature of the proof that all ℓ^p spaces are vector spaces.

5. Hilbert Space and Quantum Mechanics

5.1. **Hilbert Space.** Not all normed or inner product spaces are complete. For this reason, we want to be able to distinguish complete normed or inner product spaces from those that are not, which is the motivation behind the definition of Banach and Hilbert spaces.

Definition 5.1 (Banach Space).
A *Banach space* is a complete normed linear space.

Definition 5.2 (Hilbert Space).
A *Hilbert space* is a complete inner product space.

We know from our discussion of the ℓ^p spaces that not all Banach spaces are Hilbert spaces. Specifically, while all ℓ^p are are Banach spaces, only ℓ^2 is a Hilbert space. However, is there an inner product space that is not a Hilbert space? That is, are all inner product spaces necessarily complete? In the first sentence of this section we claimed that answer to this question is no. To see why this is so, one need only consider the set $\mathbb{Q}$ of rational numbers. Normal multiplication on rational numbers is a function that maps from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q} \subset \mathbb{R}$ and satisfies all four of the criterion of Definition 2.3. However, the sequence (s_n) of rational numbers, whose n term is given by

$$s_n = \sum_{i=1}^{n} \frac{1}{n^2},$$

which, thanks to Leonhard Euler, is known to converge to an irrational number, specifically $\frac{\pi^2}{6}$. Since (s_n) is convergent, and therefore Cauchy, we can safely conclude that while $\mathbb{Q}$ is an inner product space, it is not a Hilbert space. In the following example, we explore an even more interesting example of a normed linear space that is not complete.

Example 5.1. Let c_{00} denote the set of all finitely nonzero sequences. That is, a sequence (x_n) is a member of c_{00} if and only if there exists an $N \in \mathbb{N}$ such that $x_n = 0$ for all $n > N$. Consider the function $\|\cdot\| : c_{00} \to \mathbb{R}$ given by

$$\|(x_n)\| = \sup\{|x_n| : n \in \mathbb{N}\},$$

for all $(x_n) \in c_{00}$. We know from Example 2.1 that $\|\cdot\|$ is a norm. So, it follows that c_{00} is a normed linear space.

Proving that c_{00} is not complete, again places us in the mildly disconcerting position of needing to construct a Cauchy sequence of sequences. Using the notation in the proof of Theorem 4.3, for a fixed $n \in \mathbb{N}$, we define the sequence $a_n = \left(\alpha_i^{(n)}\right) \in c_{00}$ to be the sequence

$$a_n = \left(\alpha_i^{(n)}\right) = \Bigg\{\underbrace{\frac{1}{1}, \frac{1}{2}, \ldots, \frac{1}{i}, \ldots, \frac{1}{n}}_{n \text{ nonzero terms}}, 0, 0, \ldots, 0, \ldots\Bigg\}.$$

Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ so that $\frac{1}{N} < \varepsilon$. If $n, m > N$, where $n < m$, then

$$\begin{aligned} d(a_n, a_m) &= \sup\left\{\left|\alpha_i^{(n)} - \alpha_i^{(m)}\right| : i \in \mathbb{N}\right\} \\ &= \frac{1}{n+1} \\ &< \varepsilon. \end{aligned}$$

So, (a_n) is certainly Cauchy. However, (a_n) converges to the sequence

$$\left\{\frac{1}{1}, \frac{1}{2}, \ldots, \frac{1}{i}, \ldots,\right\},$$

which contains no zero terms and is therefore not a member of c_{00}.

5.2. **Introduction to Quantum Mechanics.** Early in the history of quantum mechanics (1922, to be precise), German physicists Otto Stern and Walter Gerlach conducted an experiment designed to detect the intrinsic angular momentum of silver atoms.[3] The silver atoms used in the Stern-Gerlach experiment, as this 1922 experiment is now commonly called, did in fact behave as though they had a sort of intrinsic angular momentum. The property responsible for the behavior of these silver atoms is now called spin.[4] Moreover, while classical mechanics predicts that the silver atoms should be allowed to have a continuous range of possible spin values, Stern and Gerlach measured the silver atoms used in their experiment to have one of two possible spin values−often referred to as "spin up" and "spin down"−a fact that holds true for an entire class of particles.

[3]For a more detailed discussion of the Stern-Gerlach experiment and its ramifications, one can turn to John Townsend's *A Modern Approach to Quantum Mechanics* [12].

[4]While term "spin" seems to indicate that the silver atoms are rotating, or spinning, in space, this is actually not the case. Instead, spin is simply an intrinsic property of the electron.

In general, the measurement of the spin a particle must be made with respect to a specified direction. So, suppose we set the z-axis of a standard Euclidean three dimensional coordinate system to point upwards (i.e. away from the center of the Earth). Then, in conventional quantum mechanics notation, if we measure a particle, say an electron, to have "spin up" with respect to the positive z-axis, we denote it as $|+\mathbf{z}\rangle$. Likewise, if we measure an electron to have "spin down" with respect to the z-direction, then we notate our find as $|-\mathbf{z}\rangle$. The $|\pm\mathbf{z}\rangle$ notation is commonly called a ket vector. Townsend reports that if we send a stream of electrons all with spin $|+\mathbf{z}\rangle$ through an apparatus that determines the spin of each electron with respect to the x-axis, we will find that half of the electrons have spin $|+\mathbf{x}\rangle$ and the other half have spin $|-\mathbf{x}\rangle$, indicating a potential relationship between $|+\mathbf{z}\rangle$ and $|\pm\mathbf{x}\rangle$ [12]. As it happens, the relationships between $|\pm\mathbf{z}\rangle$ and $|\pm\mathbf{x}\rangle$ are given by

$$|+\mathbf{z}\rangle = \frac{1}{\sqrt{2}}\,|+\mathbf{x}\rangle + \frac{1}{\sqrt{2}}\,|-\mathbf{x}\rangle\,, \tag{9}$$

and

$$|-\mathbf{z}\rangle = \frac{1}{\sqrt{2}}\,|+\mathbf{x}\rangle - \frac{1}{\sqrt{2}}\,|-\mathbf{x}\rangle\,. \tag{10}$$

In the language of quantum mechanics, Equations (9) and (10) imply that if an electron is in spin state $|+\mathbf{z}\rangle$, then we say that it is in a superposition of spin states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$. That is, an electron in spin state $|+\mathbf{z}\rangle$ is also simultaneously in both spin states $|+\mathbf{x}\rangle$ and $|-\mathbf{x}\rangle$. Since, a particle can only be measured in either state $|+\mathbf{x}\rangle$ or state $|-\mathbf{x}\rangle$ (but not both), the $|\pm\mathbf{x}\rangle$ spin states form an orthonormal basis for the $|+\mathbf{z}\rangle$ spin state.[5] More generally, in three dimensional Euclidean space, there are exactly three orthonormal bases for the spin states of an electron, namely,

$$\{|+\mathbf{x}\rangle\,, |-\mathbf{x}\rangle\}, \qquad \{|+\mathbf{y}\rangle\,, |-\mathbf{y}\rangle\}, \qquad \text{and} \qquad \{|+\mathbf{z}\rangle\,, |-\mathbf{z}\rangle\}.$$

As such, a generic spin state $|\psi\rangle$ of an electron can be written in the $|\pm\mathbf{u}\rangle$ basis, where u denotes a some specified direction in space, as a linear combination of $|+\mathbf{u}\rangle$ and $|-\mathbf{u}\rangle$.

The general spin state $|\psi\rangle$ for an electron is often called its state function. In the $|\pm\mathbf{z}\rangle$ basis, $|\psi\rangle$ may be written as

$$|\psi\rangle = a\,|+\mathbf{z}\rangle + b\,|-\mathbf{z}\rangle\,,$$

for some a, and b in $\mathbb{C}$. Equivalently, we may write $|\psi\rangle$ in the $|\pm\mathbf{z}\rangle$ basis as

$$|\psi\rangle \quad = \quad \begin{pmatrix} a \\ b \end{pmatrix}_{|\pm\mathbf{z}\rangle}\,,$$

[5]That is, if a particle is found to be in the $|+\mathbf{x}\rangle$ spin state, then a second spin measurement with respect to the x-axis will result has an one-hundred percent chance of finding the particle in $|+\mathbf{x}\rangle$ and a zero percent chance of finding the particle in $|-\mathbf{x}\rangle$.

where, of course, $|+\mathbf{z}\rangle = (1,0)$ and $|-\mathbf{z}\rangle = (0,1)$. Related to the ket vector $|\psi\rangle$ is the bra vector $\langle\psi|$, which is defined as $\langle\psi| = \overline{|\psi\rangle}^T$, where $\overline{|\psi\rangle}^T$ denotes the (complex) vector transpose of $|\psi\rangle$. As in linear algebra, the right (matrix) multiplication of $|\psi\rangle$ by $\langle\psi|$ is an inner product. Due to the ingenuousness of physicists in developing new terms, the inner product of $\langle\psi|$ with $|\psi\rangle$ (denoted $\langle\psi|\psi\rangle$) is called a bracket (or, bra-ket). More explicitly this is

$$\langle\psi|\psi\rangle \quad = \quad \begin{pmatrix}\overline{a} & \overline{b}\end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 + |b|^2.$$

One should note that, in general, the coefficients a and b may be functions of both position and time. In other words, $|\psi\rangle$ itself may vary with respect to position and time. However, there is one caveat that any good state function $|\psi\rangle$ must satisfy, namely, $\langle\psi|\psi\rangle = 1$.[6] So, if we want to know the probability of finding $|\psi\rangle$ in, say, the state $|+\mathbf{z}\rangle$, we need only find the projection of $|\psi\rangle$ onto $|+\mathbf{z}\rangle$. More explicitly, the probability that when we measure the spin of $|\psi\rangle$ we will find it to be spin-up in the z-direction is $\langle+\mathbf{z}|\psi\rangle$.

Now, if we wish to somehow modify the spin of an electron, we can mathematically represent the effect of this modification as a matrix. In fact, as we have already seen, whenever we measure the spin of an electron, we invariably modify its spin. If the electron was in some arbitrary spin state $|\psi\rangle$ before a measurement, then as soon as we measure the spin of the electron with respect to some spacial direction $\mathbf{u}$, then the spin state of the electron becomes either $|+\mathbf{u}\rangle$ or $|-\mathbf{u}\rangle$, depending on the outcome of the measurement. In other words, a measurement of spin may be represented as a matrix whose eigenvectors are the only possible outcomes of a given measurement.

As we have seen, the function that represents the spin state of an electron is a function that maps from $\mathbb{C}^n$−if the spin state function is a function of n variables−to $\mathbb{C}^2$. In other words, the Hilbert space in which electron spin is represented is $\mathbb{C}^2$. As one might imagine, quantum physicists select the inner product space in which they work based on the sort of problem with which they are currently contending. For example, the set $\mathcal{C}[a,b]$ of all continuous functions on the closed bounded interval $[a,b]$ with inner product

$$\langle f,g\rangle = \int_a^b \overline{f}g$$

is an inner product space that is frequently used in quantum mechanics.

Now that we very cursory understanding of quantum mechanics, we turn our attention to quantum logic. Quantum logic was first introduced by Garrett Birkhoff and John von Neumann in 1936 as a possible system of logic based on the bizarre nature of quantum mechanics [2].

[6] Given any ket vector $|\phi\rangle$, we can define a state function $|\psi\rangle$ by $|\psi\rangle = |\phi\rangle / \sqrt{\langle\phi|\phi\rangle}$. Observe that $\langle\phi|\phi\rangle = 1$.

6. The Manual: The Foundation of Quantum Logic

David Cohen, in his book *An Introduction to Hilbert Space and Quantum Logic*, outlines a system of logic for quantum mechanics based on Hilbert space theory called quantum logic [3]. While quantum logic is certainly one interesting vantage point from which to conduct a thorough mathematical study of quantum mechanics, such a study falls outside the discourse of this paper. However, we provide an introduction to quantum logic with the goal of providing the reader with a foundation from which to conduct further research.

6.1. **The Manual.** Central to quantum logic is the concept of a manual. A manual is a special example of something Cohen calls a quasimanual, which is simply a collection of sets. So, for the sake of completeness, we begin our exploration of quantum logic by giving a definition for a quasimanual and associated terminology.

Definition 6.1 (Quasimanual).
A *quasimanual* $\mathscr{Q}$ is a nonempty collection of nonempty sets called *experiments* [3].

Definition 6.2 (Outcomes).
The members of an experiment are called *outcomes*. The set of all possible outcomes in a quasimanual $\mathscr{Q}$ is denoted $X_{\mathscr{Q}}$ [3].

Definition 6.3 (Event).
An *event* in a quasimanual $\mathscr{Q}$ is a subset of an experiment in $\mathscr{Q}$ [3].

Cohen uses the terminology "test for an event A" to mean "performing an experiment that contains A" [3]. Accordingly, Cohen says that if such a test returns an element of A, then "event A has occurred" [3]. We, in this paper, follow suit.

Admittedly, these definitions can seem a bit convoluted, especially at first glance. Accordingly, the reader may find it helpful to bear in mind that Cohen's motivation for constructing said definitions is that humanity's understanding of specific aspects of natural world comes from data collected during any number of related experiments [3]. Accordingly, it seems natural to compile the results of experiments that are interconnected in some manner into a single set, which we call a manual. Before we get ahead of ourselves and trot onto the definition of a manual, we pause to consider a simple example of a quasimanual in order to help build intuition for the world of quantum logic.

Example 6.1. Consider the set $\mathscr{Q} = \{A, B, C\}$, where $A = \{0, 1, 2\}, B = \{2, 3, 4\}$, and $C = \{0, -1, -2\}$. In the terminology set forth above, since $\mathscr{Q}$ is a collection of sets, it is a *quasimanual*. Further, A, B, and C are each experiments, the subset $\{0, -1\}$ of C is an *event*, and $X_{\mathscr{Q}} = \{-2, -1, 0, 1, 2, 3, 4\}$.

In constructing the definition for a manual, we need to think about how we want our experiments within one to be related. The next two definitions will help us in this regard.

Definition 6.4 (**Orthogonal**).
Two events A, B in a quasimanual $\mathscr{Q}$ are said to be *orthogonal*, denoted $A \perp B$, if they are disjoint subsets of a single experiment in $\mathscr{Q}$. For outcomes x and y of $\mathscr{Q}$ we write $x \perp y$ to mean $\{x\} \perp \{y\}$ [3]

Definition 6.5 (**Orthogonal Complements**).
If A, B are orthogonal events in $\mathscr{Q}$ and $A \cup B$ is an experiment in $\mathscr{Q}$, then we say that A and B are *orthogonal complements* in $\mathscr{Q}$. We denote this by A oc B [3].

We now have everything we need to define a manual. In the next subsection, we examine a specific example of a manual. In doing so, the motivation for the manual definition should become more apparent.

Definition 6.6 (**Manual**).
A *manual* is a quasimanual $\mathfrak{M}$ which satisfies the following:

(i) If A, B, C, D are events in $\mathfrak{M}$ with A oc B, B oc C, and C oc D, then $A \perp D$.
(ii) If $E, F \in \mathfrak{M}$ and $E \subseteq F$, then $E = F$.
(iii) If x, y, z are outcomes in $\mathfrak{M}$ with $x \perp y, y \perp z, z \perp x$, then $\{x, y, z\}$ is an event in $\mathfrak{M}$.

6.2. **The Bow-Tie Manual.** To help us better understand the definitions from the preceding subsection, let's start by exploring an example which Cohen refers to as the "bow-tie manual" [3]. Suppose we have clear plastic box containing a firefly, and then drawn a vertical line down the center of the front of the box and another vertical line down the center of a side adjacent to the front. (Please see Figure 2.) If we look through the front of the

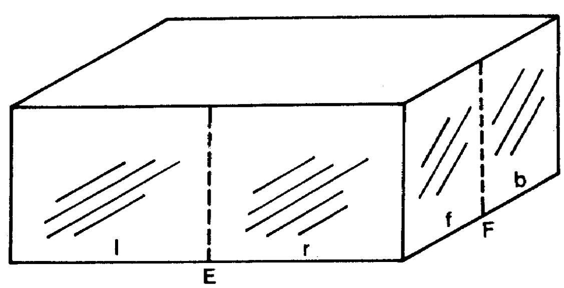

Figure 2. Experimental set-up for the so-called bow-tie manual. The box contains a firefly. Figure courtesy of Cohen [3].

box, we either see a light on the right side of the box (outcome r), a light on the left side of the box (outcome l), or no light at all (outcome n). Similarly, if we look through the side of the box, we either see a light in the front of the box (f), a light in the back of the box (b), or, again, no light at all (n). So, in essence, each time we look at the box, we perform an experiment whose outcome depends on the position of the firefly within the

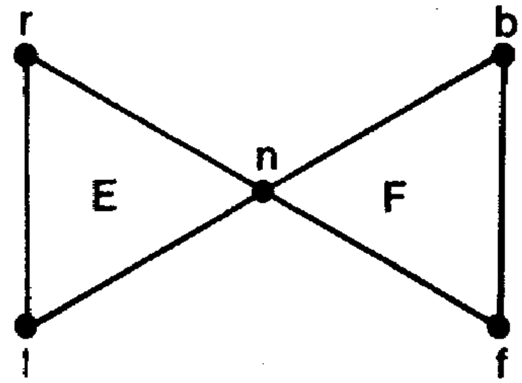

FIGURE 3. Orthogonality diagram for the bow-tie manual. Courtesy of [3].

box and the firefly's willingness to emit light for our viewing pleasure. For simplicity, we will refer to the separate acts of looking into the front and looking into the side of the box as experiments E and F, respectively.

In order to verify that the bow-tie manual is indeed a manual as purported, we find it useful to first diagram the orthogonality relations within it (see Figure 3). The bow-tie manual is certainly a quasimanual, as it is a collection of two sets. Further, since neither experiments E nor F are subsets of the other, the bow-tie manual automatically satisfies property (ii) of Definition 6.6.

We now proceed to demonstrate that the bowtie manual satisfies the criterion (i) of Definition 6.6. Consider the event $\{r\}$ (recall that an event is simply a subset of an experiment). Then $\{r\}$ is an orthogonal complement to $\{l, n\}$. However, the only orthogonal complement to $\{l, n\}$ is $\{r\}$, giving us the chain

$$\{r\} \text{ oc } \{l, n\} \text{ oc } \{r\} \text{ oc } \{l, n\}.$$

The event $\{l, n\}$ is certainly orthogonal to $\{r\}$. Of course, if we start with $\{l\}$ instead of $\{r\}$, we obtain the nearly identical orthogonal complement chain

$$\{l\} \text{ oc } \{r, n\} \text{ oc } \{l\} \text{ oc } \{r, n\}.$$

So, now consider the event $\{r, l\}$. The only orthogonal compliment of $\{r, l\}$ is $\{n\}$, which is an orthogonal complement to $\{f, b\}$. However, since the event $\{n\}$ is the only orthogonal complement to $\{f, b\}$ and we've already used it in our sequence of orthogonal complements, we are done. Analogous arguments hold if we start by selecting an event in F.

In order for two events or outcomes to be orthogonal, they must be from the same experiment. So, the bow-tie manual trivially satisfies property (iii), and is therefore a manual as claimed.

6.3. **Further Examples.**

Example 6.2. Consider a set X with at least two members. The collection $\mathscr{Q}_X$ of all nonempty subsets of X is certainly a quasimanual. Suppose x, y are elements of X. Then $\{x\}, \{x, y\} \in \mathscr{Q}_X$. Since $\{x\}$ is a proper subset of $\{x, y\}$, $\mathscr{Q}_X$ fails criteria (ii) of definition 6.6, and is therefore not a manual.

Example 6.3 (Partitional Manual on $\mathbb{R}$)**.** In this example, we show the that so-called *partition manual on* $\mathbb{R}$ is a manual. The partition manual of $\mathbb{R}$ is the collection $\mathbb{B}$ of all countable partitions of the set $\mathbb{R}$ of real numbers that satisfy

(i) $\mathbb{I} \subseteq \mathbb{B}$
(ii) if $B \in \mathbb{B}$, then $\mathbb{R} \backslash B \in \mathbb{B}$;
(iii) $\mathbb{B}$ is closed under countable unions,

where $\mathbb{I} = \{(a, b] : a, b \in \mathbb{R}\}$. Suppose $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ are events in $\mathbb{B}$ with $\mathcal{A}$ oc $\mathcal{B}$, $\mathcal{B}$ oc $\mathcal{C}$, and $\mathcal{C}$ oc $\mathcal{D}$. Since $\mathcal{A}$ is countably large, we can index the sets contained within $\mathcal{A}$ so that $\bigcup_{n\in\mathbb{N}} A_n$ denotes the union of all sets contained contained within $\mathcal{A}$. We can also do the same for $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$. Then

$$\bigcup_{n\in\mathbb{N}} A_n \cap \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} B_n \cap \bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} C_n \cap \bigcup_{n\in\mathbb{N}} D_n = \{\},$$

and

$$\bigcup_{n\in\mathbb{N}} A_n \cup \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} B_n \cup \bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} C_n \cup \bigcup_{n\in\mathbb{N}} D_n = \mathbb{R}.$$

It follows then that $\bigcup_{n\in\mathbb{N}} A_n$ is disjoint to $\bigcup_{n\in\mathbb{N}} D_n$. Hence $\mathcal{A}$ is perpendicular to $\mathcal{D}$. Since $\mathbb{B}$ automatically satisfies properties (ii) and (iii) of Definition 6.6, we conclude that $\mathbb{B}$ is a manual as claimed.

7. Building up a System of Logic

Our next task is to construct a system of logic. The key to doing so is understanding the orthogonal complement relationships between events of a manual. As such, the following definition of *operationally perspective* proves useful.

Definition 7.1 (**Operationally Perspective**)**.**
If $\mathfrak{M}$ is a manual, A, B and C are events in $\mathfrak{M}$, and A oc B and B oc C, then we say that A and C are operationally perspective, which we denote by A op C.

To better understand what it means for two events to be operationally perspective, consider the bow-tie manual. In the bow-tie manual, the event $\{r, l\}$ is an orthogonal complement to the event $\{n\}$, which, in turn, is an orthogonal complement to the even $\{b, f\}$. Hence $\{r, l\}$ and $\{b, f\}$ are operationally perspective. In symbols this is $\{r, l\}$ op $\{b, f\}$.

A reasonable definition for implication follows nicely from Definition 7.1. As we see in Lemma 7.1, such an implication has the rather nice property that it is a partial ordering on the collection of events in a manual.

Definition 7.2 (**Implies**).
If $\mathfrak{M}$ is a manual and A and B are events in $\mathfrak{M}$, then we say A *implies* B, denoted $A \leqq B$, if and only if there is an event C with $C \perp A$ and $(C \cup A)$ op B.

Definition 7.3 (**Logically Equivalent**).
If $\mathfrak{M}$ is a manual and A and B are events in $\mathfrak{M}$, then we say that A is *logically equivalent* to B, denoted $A \leftrightarrow B$, if and only if $A \leqq B$ and $B \leqq A$.

Lemma 7.1. *Let $\mathfrak{M}$ be a manual. Then implication is a partial ordering on the collection of events in $\mathfrak{M}$.*

Proof. Recall that a partial ordering of a set is a binary relation $\leq$ that is reflexive $(a \leq a)$, antisymmetric (if $a \leq b$ and $b \leq a$, then $a = b$), and transitive (if $a \leq b$ and $b \leq c$, then $a \leq c$).

Since $\leqq$ is almost trivially reflexive (consider $A \cup \{\}$, for some event A) and antisymmetry is taken care of by the logical equivalent definition, we need only show that implication is transitive. So, suppose A, B, C are events in $\mathfrak{M}$ with $A \leqq B$ and $B \leqq C$. Then exist events D, E, F, G such that $D \perp A$, $E \perp B$, $A \cup D$ oc F, F oc B, $B \cup E$ oc G, and G oc C. Since $B \perp E$, B oc $G \cup E$. By the first property of manuals, $A \cup D \perp G \cup E$, which implies that $A \cup D \perp G$. So there exists some event Γ such that $\Gamma \perp A$ and $A \cup D \cup \Gamma$ oc G. Hence $A \cup D \cup \Gamma$ op C. Since $A \perp D \cup \Gamma$, $A \leqq C$. □

At this point, a brief recess in our orgy of gratuitously convoluted definitions is in order. One technique that greatly aids in understanding many of the tortuous relationships found in quantum logic is diagraming the orthogonal complements. So, we'll begin by diagraming the operationally perspective definition. Suppose that $\mathfrak{M}$ is a manual, Γ and Λ are experiments in $\mathfrak{M}$, and there exist $A \subseteq \Gamma$ and $C \subseteq \Lambda$ such that A op C. Then there exists an event B in both Γ and Λ such that A oc B and B oc C. Figure 4 presents a diagram of these orthogonal relationships. In essence, if A happens, then B cannot happen, so C must happen. Conversely, if B happens, then neither A nor C can happen. In other words, if two events are operationally perspective, if one occurs, then so must the other. It is important to note that two operationally perspective events need not come from the same experiment.

Now let's look at Definition 7.2. One can certainly conceive of a situation when an event B happens whenever an event A happens but not necessarily *vise versa*. In the bow-tie manual, for example, if event $\{r\}$ happens, then the event $\{b, f\}$ must also happen. However, if $\{b, f\}$ occurs, then $\{r\}$ need not occur, as $\{l\}$ may occur instead. For the

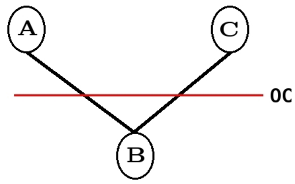

FIGURE 4. OC diagram for two operationally perspective events A and C. The solid black lines indicate oc relationships.

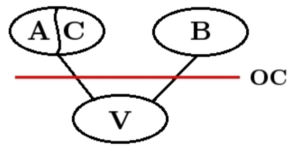

FIGURE 5. OC diagram for $A \leqq B$.

sake of constructing an OC diagram for Definition 7.2, suppose $A \leqq B$. Then there exist events C and V such that $A \perp B$ and $A \cup B$ oc V and V oc B (see Figure 5). Notice that A op B only when C is the empty set. This observation begs the question, "what exactly is the difference between logically equivalence and operationally perspective?" As we see in Lemma 7.2, there isn't any difference between operationally perspective and logically equivalent.

Lemma 7.2.
If A, B are events in a manual $\mathfrak{M}$, then A op B if and only if $A \leftrightarrow B$.

Proof. Suppose A op B. So, $A \cup \{\}$ op B and $B \cup \{\}$ op A. Hence $A \leqq B$ and $B \leqq A$, as $A \perp \{\}$ and $B \perp \{\}$.

Now suppose that $A \leftrightarrow B$. Then there exist events C, D, E, and F with the properties that $A \perp C$, $B \perp E$, $A \cup C$ oc D, D oc B, $B \cup E$ oc F, and F oc A. So, B oc $E \cup F$, which implies that $A \cup C \perp E \cup F$, by the first property of manuals. As such, there exists some event Γ such that $A \cup C \cup E \cup F \cup \Gamma$ is an experiment. Since A oc F, $A \cup F$ is an experiment. So, by the second property of manuals, $C = E = \Gamma = \{\}$. Hence A op B. $\square$

While seemingly trivial in nature, this next set of lemmas are interesting nonetheless.

Lemma 7.3.
If E and F are experiments in a manual $\mathfrak{M}$, then $E \leftrightarrow F$.

Proof. Since E oc $\{\}$ and $\{\}$ oc F, we see that E op F. We conclude by Lemma 7.2 that E and F are logically equivalent. $\square$

Lemma 7.4.
If A, B, C, D are events in $\mathfrak{M}$ and $A \leftrightarrow C$ and $B \leftrightarrow D$, then $A \leqq B$ if and only if $C \leqq D$.

Proof. Suppose $A \leqq B$. Since $C \leqq A$ and $B \leqq D$, we see that $C \leqq A \leqq B \leqq D$. So $C \leqq D$.

Similarly, if $C \leqq D$, then $A \leqq C \leqq D \leqq B$. Hence $A \leqq B$. □

Lemma 7.5.
If A is an event in $\mathfrak{M}$ and $E, F \in \mathfrak{M}$, with $A \subseteq E$ and $A \subseteq F$, then $F \backslash A \leftrightarrow E \backslash A$.

Proof. Since A oc $E\backslash A$ and A oc $F\backslash A$, then $E\backslash A$ op $F\backslash A$. Thus $F\backslash A \leftrightarrow E\backslash A$. □

Definition 7.4 (Lattice).
Let P be a set with a partial ordering $\leqq$. Then $(P, \leqq)$ is called a *lattice* if for all $p, q \in P$ the set $\{p, q\}$ has a greatest lower bound and a least upper bound in P.

We call $\inf\{p, q\}$ the *meet* of p and q, and we denote it by $p \wedge q$.

We call $\sup\{p, q\}$ the *join* of p and q, and we denote it by $p \vee q$.

A couple of notes on this definition: If $\inf\{p, q\} = s$, then s has the property that $s \leqq p, q$ (that is, s implies both p and q) and for all r that implies both p and q, we have $r \leqq s$.

Definition 7.5 (Unit and Zero of a Lattice).
If $(L, \leqq)$ is a lattice, then a *unit* 1_L for L is a member of L such that $p \leqq 1_L$ for all $p \in L$. A *zero* 0_L for L is a member of L such that $0_L \leqq p$ for all $p \in L$.

It is standard practice to drop the subscript notation from both the zero and unit if it is clear from context to which lattice they belong.

Example 7.1. Let X be a nonempty set and let $\mathrm{Sub}(X)$ be the collection of all subsets of X. Then $\mathrm{Sub}(X)$, partially ordered by inclusion, is a lattice with a unit and a zero. To see that this is so, consider any $K_1, K_2 \in \mathrm{Sub}(X)$. Then $K_1 \wedge K_2 = K_1 \cap K_2$, and $K_1 \vee K_2 = K_1 \cup K_2$. So, $\mathrm{Sub}(X)$ is a lattice. Further, since the empty set $\{\}$ is a subset of all elements of $\mathrm{Sub}(X)$ and all elements of $\mathrm{Sub}(X)$ are subsets of X, $\mathrm{Sub}(X)$ has zero $0 = \{\}$ and unit $1 = X$.

Example 7.2. Let H be a finite dimensional Hilbert Space, and let L be the collection of all subspaces of H. Define a partial order on L as follows: for $R, S \in L$, $R \leqq S$ if and only if R is a subspace of S. Then $(L, \leqq)$ is a lattice with unit and zero.

Proof. Consider any $T, Q \in L$. If $x, y \in T \cap Q$ and α, β are scalars, then, since $\alpha x + \beta y \in T$ and $\alpha x + \beta y \in Q$, we have $\alpha x + \beta y \in T \cap Q$. So $T \wedge Q = T \cap Q$, as $T \cap Q$ is the largest possible subspace of H that is a subspace of both T and Q. Now consider the set $\cap\mathbb{I}$ that is the intersection of all subspaces of H that contain both T and Q. If $p, q \in \cap\mathbb{I}$ and a, b are scalars, then $ap + bq$ is in $\cap\mathbb{I}$. So $T \vee Q = \cap\mathbb{I}$. Hence L is a lattice with zero $\{\}$ and unit H. □

Definition 7.6 (Logic)**.**
A *logic* $(L, \leqq, ')$ is a lattice $(L, \leqq)$ with unit and zero, together with an operation $' : L \to L$, called an *orthocomplemetation*, that satisfies:

(i) for all $p \in L$, $p'' = p$ and $p \wedge p' = 0_L$;
(ii) for all $p, q \in L$, if $p \leqq q$, then $q' \leqq p'$;
(iii) for $p, q \in L$, if $p \leqq q$, then $q = p \vee (p' \wedge q)$.

A logic $(L, \leqq, ')$ is often denoted L if there is no fear of ambiguity. The members of a logic are referred to as propositions.

Definition 7.7 (Orthogonal)**.**
If L is a logic and $p, q \in L$, we say that p is *orthogal* to q, denoted $p \perp q$, if and only if $p \leqq q'$.

Theorem 7.1 (DeMorgan's Law)**.**
If L is a logic and $p, q \in L$, then

$$(p \vee q)' = p' \wedge q' \qquad \textit{and} \qquad (p \wedge q)' = p' \vee q'.$$

Proof. Choose any $p, q \in L$. Since $p \leqq p \vee q$ and $q \leqq p \vee q$, property (ii) from definition 7.6 tells us that $(p \vee q)' \leqq p'$ and $(p \vee q)' \leqq q'$. It follows that $(p \vee q)' \leqq p' \wedge q'$. Conversely, $p' \wedge q' \leqq p'$ and $p' \wedge q' \leqq q'$. So $p \leqq (p' \wedge q')' \leqq p'$ and $p \leqq (p' \wedge q')' \leqq q'$. Hence, $p \vee q \leqq (p' \wedge q')'$. So, by applying property (ii) from definition 7.6 another time, we see that $p' \wedge q' \leqq (p \vee q)'$. Therefore $(p \vee q)' = p' \wedge q'$.

We will use the first equality to prove the second. Since $(p' \wedge q')' = (p \vee q)'' = p \vee q$, it follows that $(p \wedge q)' = p' \vee q'$. □

8. Quantum Logic and Quantum Mechanics

While, as previously stated, a thorough discussion of quantum mechanics through the lens of quantum logic is outside the scope of this paper, the purpose of this final section is to provide the reader with a brief taste for how such a discussion might progress. Consequently, we do not provide a comprehensive descant; we merely provide a few definitions and concepts that are relevant to the topic. Nonetheless, the interested reader can seek Cohen's *An Introduction to Hilbert Space and Quantum Logic* for further information [3].

Definition 8.1 (Compatible Propositions).
Two propositions p and q in a logic L are called *compatible* if there exist $u, v, w \in L$ such that if $\{u, v, w\}$ is a pairwise orthogonal set in L, then $u \vee v = p$ and $v \vee w = q$.

The set $\{u, v, w\}$ is called a *compatibility decomposition* for p and q.

Definition 8.2 (Quantum Logic).
A *quantum logic* is a logic with at least two propositions that are not compatible.

Definition 8.3 (**Classical Logic**).
A *classical logic* is a logic in which every pair of propositions is a compatible pair.

Definition 8.4 (**State Function**).
Suppose L is a logic. Then a *state function* $s : L \to [0, 1]$, or *state*, for short is a function with the properties that

(i) for $p, q \in L$, if $p \perp q$, the $s(p \vee q) = s(p) + s(q)$, and
(ii) $s(1_L) = 1$.

Definition 8.5 (**Classical States**).
A *classical state* s on a logic L is a state function with the property that for all $p \in L$, either $s(p) = 0$ or $s(p) = 1$).

Definition 8.6 (**Pure State**).
A *pure state* on a logic L is one that cannot be written as a nontrivial mixture of other states on L.

Lemma 8.1. *If $\mathcal{B}$ is a basis for a finite Hilbert space H and A, B are disjoint subsets of $\mathcal{B}$ with the property that $A \cup B = \mathcal{B}$, then A and B span different subsets of H. That is, and span$(A) \cap$span$(B) = \{0\}$.*

Proof. Suppose x is in the span of both A and B. Then, there exist scalars $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_k$ such that

$$x = \alpha_1 \mathbf{a_1} + \alpha_2 \mathbf{a_2} + \cdots + \alpha_n \mathbf{a_n} \qquad \text{and} \qquad x = \beta_1 \mathbf{b_1} + \beta_2 \mathbf{b_2} + \cdots + \beta_k \mathbf{b_k}$$

where $A = \{\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}\}$ and $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_k}\}$. So, it follows that

$$\alpha_1 \mathbf{a_1} + \alpha_2 \mathbf{a_2} + \cdots + \alpha_n \mathbf{a_n} + (-\beta_1)\mathbf{b_1} + (-\beta_2)\mathbf{b_2} + \cdots + (-\beta_k)\mathbf{b_k} = 0. \tag{11}$$

Since the only solution to (11) is the trivial solution, we have $x \in \{0\}$. Since 0 is in both the span of A and the span of B, the intersection of the span of A with the span of B is the set $\{0\}$. $\square$

Interesting Fact 8.1.
If H is a finite dimensional Hilbert space and $\mathscr{F}(H)$ is the collection of all orthonormal bases for H, then $\mathscr{F}(H)$ is a manual called the *frame manual.*

Proof. We'll begin by verifying property (i) of definition 6.6. Consider any bases A, B, C in $\mathscr{F}(H)$ where A oc B, B oc C, and C oc D. Then there exist $E, F, G \in \mathscr{F}(H)$ such that $A \cup B = E$, $B \cup C = F$ and $C \cup D = G$. Lemma 11 tells us that A and B span different subspaces of H and B and C span different subspaces of H. Suppose a is in the span of A and $a \neq 0$. Then a is not in the span of B. However, since a is in the span of $F = B \cup C$, then it must be in the span of C. An analogous argument demonstrates that the span of C is also a subset of the span of A. Hence A and C span the same subspaces of H. Through the use of similar reasoning, we can see that that B and D span the same subspaces. The proof of Lemma 8.1 implies that the elements of A and D are linearly independent, and so form a basis.

If $I, J \in \mathscr{F}(H)$ and $I \subseteq J$, then, since $|I| = |J| = \dim(H)$, $I = J$, where $|\cdot|$ denotes the order (size) of a set. Since property (iii) of definition 6.6 is satisfied trivially, we conclude that $\mathscr{F}(H)$ is a manual. $\square$

9. Conclusion

In Section 5.2, we mentioned in passing that any external modification of the spin state of an electron may be represented as a matrix. This fact applies more generally to any state function that is in a Hilbert space that is isomorphic to $\mathbb{R}^n$. That is, if the state of some particle is given by $|\phi\rangle$, where the co-domain of $|\phi\rangle$ is $\mathbb{R}^n$, then any external modification of $|\phi\rangle$ may be represented by some matrix A. Matrix A is actually a special example of a particular type of creature, known as an operator, that lives in the seedy underground of the mathematical representation of quantum mechanics. Operators play a particularly important part in quantum mechanics, and a truly thorough understanding of mathematics that underpins quantum mechanics requires a study of the theory that governs them. While, operator and spectral theory are both outside the realm of this paper, Sterling Berberian [1], David Cohen [3] and J.R. Retherford [8] all provide approachable discourses on both spectral and operator theory. As such, all readers interested in the mathematical theory behind quantum mechanics are highly encouraged to seek these three sources.

References

[1] Berberian, Sterling. *Introduction to Hilbert Space*. New York: Oxford University Press. 1961.

[2] Birkhoff, Garrett and John von Neumann. *The Logic of Quantum Mechanics*. Annals of Mathematics 37, 823843. 1936.

[3] Cohen, David. *An Introduction to Hilbert Space and Quantum Logic*. New York: Springer-Verlag. 1989.

[4] Wikipedia. *Concave function*. *http://en.wikipedia.org/wiki/Concave_functions*. 21 Oct 2008.

[5] "David Hilbert." Wikiquote. *http://en.wikiquote.org/wiki/David_Hilbert*. Jan 28, 2009.

[6] "David Hilbert." Wikipedia. *http://en.wikipedia.org/wiki/David_Hilbert*. Jan 28, 2009.

[7] "Hilbert Space." Wikipedia. *http://en.wikipedia.org/wiki/Hilbert_space*. Jan 28, 2009.

[8] Retherford, J.R. *Hilbert Space: Compact Operators and the Trace Theorem*. London: Cambridge University Press. 1993.

[9] Saxe, Karen. *Beginning Functional Analysis*. New York: Springer. 2002.

[10] Schumacher, Carol. *Closer and Closer: Introducing Real Analysis*. Boston: Jones and Bartlett Publishers. 2008.

[11] Simmons, George. *Introduction to Topology and Modern Analysis*. San Francisco: McGraw-Hill Book Company, Inc. 1963.

[12] Townsend, John. *A Modern Approach to Quantum Mechanics*. Sausalito, California: University Science Books. 2000.