Rudiments of Hilbert Space Theory

Tom Potter

Contents

Preface

This work started as part of an expository master's thesis, and was later rejected as the thesis changed direction. I learned most of this material from Conway [4] and Folland [6], and the reader familiar with Conway's book will no doubt see his influence. I struggled with the material presented there, however, and it wasn't until I wrote out my own version of the material that I felt I really understood it and knew it. So these notes are essentially for myself: I wrote them to get a better grasp on the material, and for the exercise of writing math. I regard them as a work in progress. I have found several mistakes over the years, some quite embarrassing, and I expect many still exist. I will be grateful to learn of any mistakes, so that I can fix them.

The first section may seem eccentric; it is largely and essentially borrowed from a math.stackexchange.com thread: [1] and Tom Apostol's book [2]. I use results from this section only twice in the remainder of these notes. The fact that the iterated limsup is bounded by the double limsup is used in the proof that an infinite direct sum of Hilbert spaces is complete (this is the proof given in Dunford & Schwartz); it isn't necessary to use this fact here, but I've seen it used elsewhere in the literature (for example, the proof given in Pedersen's *Analysis Now* that $\mathscr{B}(\mathscr{X}, \mathscr{Y})$—the space of bounded linear transformations from $\mathscr{X}$ to $\mathscr{Y}$—is complete whenever $\mathscr{Y}$ is), so I decided it was worth recording.

Since I have chosen to use i as an index throughout, I have used the symbol $\imath$—the "\imath" symbol—to denote the imaginary unit in the complex numbers.

1 Preliminaries: Limits, Sup, Limsup

1.1 Double Limits; Iterated Limits

The material in this subsection is borrowed largely and essentially from a discussion found on math.stackexchange.com: [1], as well as on material from Tom Apostol's beautiful book *Mathematical Analysis*.

Definition 1.1. A *double sequence* is a function f whose domain is $\mathbb{N} \times \mathbb{N}$.

Definition 1.2. Let (X, d) be a metric space. We say that a double sequence $f(m, n)$ converges to $a \in X$ if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(f(m, n), a) < \varepsilon$ whenever m and n are both greater than or equal to N. When this is the case we write $\lim_{m,n\to\infty} f(m, n) = a$, and we call a the *double limit* of the sequence f.

If $\lim_{n\to\infty} f(m, n)$ exists for each m, we can consider the *iterated limit*

$$\lim_{m\to\infty} \lim_{n\to\infty} f(m, n) \equiv \lim_{m\to\infty} \Big(\lim_{n\to\infty} f(m, n)\Big);$$

similarly, if $\lim_{n\to\infty} f(m, n)$ exists for each n we can consider the iterated limit

$$\lim_{n\to\infty} \lim_{m\to\infty} f(m, n) \equiv \lim_{n\to\infty} \Big(\lim_{m\to\infty} f(m, n)\Big).$$

In general it is *not* the case that the iterated limits exist when the double limit exists, and vice versa. Indeed, it may be the case that the double limit exists but only one, or neither of the iterated limits exist; or it may be the case that one or both of the iterated limits exist but the double limit fails to exist—see Apostol's book for examples.

Theorem 1.3. If the double sequence f converges to $a \in X$ and, if, for each m, the limit $\lim_{n\to\infty} f(m, n)$ exists, then the iterated limit

$$\lim_{m\to\infty} \lim_{n\to\infty} f(m, n)$$

exists and is equal to a.

Proof. Let $F(m) = \lim_{n\to\infty} f(m, n)$. Let $\varepsilon > 0$. Since the double limit of f is equal to a, there exists an $N \in \mathbb{N}$ such that

$$d\big(f(m, n), a\big) < \frac{\varepsilon}{2} \quad \text{for all } m, n \geq N.$$

Let $m \geq N$, and choose $n \geq N$, (which will depend on m as well as on ε), such that

$$d\big(f(m,n), F(m)\big) < \frac{\varepsilon}{2};$$

then by the triangle inequality,

$$d\big(F(m), a\big) < \varepsilon.$$

We've shown that for each $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $d(F(m), a) < \varepsilon$ whenever $m \geq N$, which is what was to be proven. $\square$

We note that switching the roles of m and n yields the analogous result for the other iterated limit. Thus we have the following corollary:

Corollary 1.4. When the double limit of f exists, and when it makes sense to talk about the iterated limits—i.e., when $\lim_{n\to\infty} f(m,n)$ exists for each m and $\lim_{m\to\infty} f(m,n)$ exists for each n—the iterated limits both exist and are both equal to the double limit of f.

1.2 Infinite Double Limits

When X is taken to be the real numbers, we can speak of a sequence $f : \mathbb{N} \to \mathbb{R}$ approaching infinity.

Definition 1.5. If for every $M \in \mathbb{R}$ there exists an $N \in \mathbb{N}$ such that $f(n) \geq M$ whenever $n \geq N$, we say that f tends to infinity, and write $\lim_{n\to\infty} f(n) = \infty$.

Definition 1.6. Given a double sequence f, we say that f tends to infinity in the double limit if, for each $M \in \mathbb{R}$, there exists an N such that $f(m,n) \geq M$ whenever m and n are both greater than or equal to N; in this case we write $\lim_{m,n\to\infty} f(m,n) = \infty$. Just as for finite limits, we can consider iterated limits and explore the relationship between the double limit and the iterated limits.

Theorem 1.7. Suppose that the double sequence $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ tends to infinity in the double limit, and assume that for each m, $f(m,n)$ converges to a number $F(m) \in \mathbb{R}$. Then

$$\lim_{m\to\infty} \lim_{n\to\infty} f(m,n) \equiv \lim_{m\to\infty} F(m) = \infty.$$

Proof. Let $M > 0$. There exists $N \in \mathbb{N}$ such that $f(m,n) \geq M$ whenever $m, n \geq N$. Let $m \geq N$; since $f(m,n) \geq M$ for each $n \geq N$, it follows that $F(m) \geq M$. We've shown that for each $M > 0$, there exists N such that $F(m) \geq M$ for all $m \geq N$; that is, F tends to infinity, as was to be shown. $\square$

This is an analogue of Theorem 1.3. In fact, it is a corollary of Theorem 1.3: for consider $X = \overline{\mathbb{R}}$, the extended real numbers, endowed with the metric

$$\rho(x, y) = |\arctan(x) - \arctan(y)|,$$

where we define $\arctan(\infty) = \frac{\pi}{2}$ and $\arctan(-\infty) = \frac{-\pi}{2}$. The reader can easily verify that $\lim_{n\to\infty} f(n) = \infty$ in this metric space if, and only if, for each $M > 0$ there exists an N such that $f(n) \geq M$ for all $n \geq N$; thus the notion of convergence to infinity that we discussed for real-valued functions agrees with the metric space notion of convergence to the point $\infty \in \overline{\mathbb{R}}$.

Just as for finite limits, the double limit can fail to exist even though one of the iterated limits exist—the standard example here is $f(m, n) = \frac{mn}{m^2+n^2}$, for which both iterated limits are 0, but the double limit does not exist.

The double limit can fail to tend to infinity even though both of the iterated limits tend to infinity.

Example 1.8. Let f be the double sequence given by

$$f(m, n) = \frac{m(n-m)}{m^2+n} + \frac{n(m-n)}{m+n^2}.$$

It is easy to verify that for each fixed m, $\lim_{n\to\infty} f(m, n) = m - 1$, whence

$$\lim_{m\to\infty} \lim_{n\to\infty} f(m, n) = \infty.$$

Similarly,

$$\lim_{n\to\infty} \lim_{m\to\infty} f(m, n) = \infty.$$

But $f(m, m) = 0$ for all $m \in \mathbb{N}$, from which it follows that the double limit does not tend to infinity.

The reader can easily verify that if even *one* of the iterated limits of a double sequence f exists, then f cannot tend to infinity in the double limit.

1.3 Double and Iterated Limits Superior

We've discussed double and iterated limits for double sequences; what about double and iterated limits superior? The situation is different for lim sup because it always exists in the extended reals, and is finite for bounded sequences.

Definition 1.9. Given a sequence $f : \mathbb{N} \to \overline{\mathbb{R}}$, we define its *limit superior* to be

$$\inf_{l \in \mathbb{N}} \sup_{n \geq l} f(n) = \lim_{l \to \infty} \sup_{n \geq l} f(n),$$

which we denote by

$$\limsup f(n) \quad \text{or} \quad \limsup_{n} f(n).$$

It is easy to see that lim sup always exists in $\overline{\mathbb{R}}$, because by definition it is the limit of a monotone sequence.

Definition 1.10. We say that the *double limit superior* of the double sequence $f : \mathbb{N} \times \mathbb{N} \to \overline{\mathbb{R}}$ is $a \in \overline{\mathbb{R}}$ if

$$\lim_{N \to \infty} \Big(\sup_{m,n \geq N} f(m,n) \Big) = a.$$

Theorem 1.11. Suppose that f is a bounded double sequence taking its values in the extended real numbers. The double limit superior

$$\limsup_{m,n \to \infty} f(m,n) = a$$

exists and the iterated limits superior

$$\limsup_{m} \Big(\limsup_{n} f(m,n) \Big) \quad \text{and} \quad \limsup_{n} \Big(\limsup_{m} f(m,n) \Big)$$

are both bounded above by a.

Proof. Notice that $\sup_{m,n \geq N} f(m,n)$ is a decreasing sequence in N, so its limit exists in $\overline{\mathbb{R}}$; hence the existence of the double limit superior. If $\limsup_{m,n} f(m,n) = \infty$, then the second statement in the theorem holds automatically; so suppose that $\limsup_{m,n} f(m,n) = a \in \mathbb{R}$, and let $\varepsilon > 0$. There exists $N(\varepsilon) \in \mathbb{N}$ such that

$$\sup_{m,n \geq N(\varepsilon)} f(m,n) < a + \varepsilon.$$

It follows that for fixed $n \geq N(\varepsilon)$,

$$\limsup_{m} f(m,n) < a + \varepsilon.$$

From this it follows that

$$\sup_{n \geq N(\varepsilon)} \Big(\limsup_{m} f(m,n) \Big) \leq a + \varepsilon,$$

whence

$$\limsup_n \left(\limsup_m f(m,n) \right) \le a + \varepsilon.$$

Since ε was arbitrary, we have

$$\limsup_n \left(\limsup_m f(m,n) \right) \le a.$$

A symmetric argument shows that

$$\limsup_m \left(\limsup_n f(m,n) \right) \le a.$$

For the case when $\limsup_{m,n} f(m,n) = -\infty$, the argument is essentially the same: for each $M \in \mathbb{R}$, there exists $N(M)$ such that

$$\sup_{m,n \ge N(M)} f(m,n) < M;$$

it follows that

$$\limsup_m f(m,n) < M \quad \text{for each fixed} \quad n \ge N(M),$$

whence

$$\sup_{n \ge N(M)} \left(\limsup_m f(m,n) \right) \le M,$$

whence

$$\limsup_n \left(\limsup_m f(m,n) \right) \le M.$$

Since M was arbitrary, we conclude that

$$\limsup_n \left(\limsup_m f(m,n) \right) = -\infty,$$

and since the argument is symmetric in m and n, we have that the other iterated limit equals $-\infty$ too. This concludes the proof. $\square$

Example 1.12. It may be that the double limit superior is strictly greater than the iterated limits superior; the sequence

$$f(m,n) = \begin{cases} m & \text{if } m = n \\ 0 & \text{if } m \ne n. \end{cases}$$

illustrates this; its double limit superior is ∞, but its iterated limits superior are 0.

Example 1.13. The iterated limits superior need not be equal. Consider, for example, the sequence

$$f(m,n) = \begin{cases} m & \text{if } n = km \text{ for some } k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

The double limit superior

$$\limsup_{m,n} f(m,n)$$

is ∞, since for every positive integer N, $\sup_{m,n \geq N} f(m,n) = \infty$. The iterated limit superior

$$\limsup_{m} \Big(\limsup_{n} f(m,n) \Big)$$

is also ∞, because for each m,

$$\limsup_{n} f(m,n) = m.$$

Notice, however, that

$$\limsup_{n} \Big(\limsup_{m} f(m,n) \Big) = 0$$

because, for each $n \in \mathbb{N}$, we have $f(m,n) = 0$ for all $m > n$; and hence, for each fixed n,

$$\limsup_{m} f(m,n) = 0.$$

The reader can verify that similar results can be stated and proven for the limits inferior.

1.4 Swapping Supremums

We've previously considered double sequences and determined a sufficient condition for when one can switch the order in which the iterated limit is taken for such sequences. One may in certain situations wish to swap the order in an iterated supremum of a double sequence; that is, one may wish to write

$$\sup_{m} \Big(\sup_{n} f(m,n) \Big) = \sup_{n} \Big(\sup_{m} f(m,n) \Big).$$

Of course, one can always do this. We prove a more general result, as it is useful to have the more general result, and the order structure $\leq$ on $\mathbb{N}$ is not needed for the proof.

Theorem 1.14. Let X and Y be sets, and let $f : X \times Y \to \mathbb{R}$ be a function. Then

$$\sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right) = \sup_{y \in Y} \left(\sup_{x \in X} f(x, y) \right) = \sup_{(x,y) \in X \times Y} f(x, y).$$

Proof. Suppose that

$$\sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right) < \sup_{(x,y) \in X \times Y} f(x, y).$$

Then since supremum, by definition, means *least* upper bound,

$$\sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right)$$

is not an upper bound for

$$\{f(x, y) : x \in X, y \in Y\};$$

so there exists some (x_0, y_0) such that

$$\sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right) < f(x_0, y_0).$$

But

$$f(x_0, y_0) \leq \sup_{y \in Y} f(x_0, y) \leq \sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right),$$

a contradiction. Now suppose that

$$\sup_{x \in X} \left(\sup_{y \in Y} f(x, y) \right) > \sup_{(x,y) \in X \times Y} f(x, y).$$

Then $\sup_{(x,y) \in X \times Y} f(x, y)$ is not an upper bound for

$$\left\{ \sup_{y \in Y} f(x, y) : x \in X \right\},$$

so there exists some $x_0 \in X$ such that

$$\sup_{y \in Y} f(x_0, y) > \sup_{(x,y) \in X \times Y} f(x, y).$$

Now this equations implies that $\sup_{(x,y) \in X \times Y} f(x, y)$ is not an upper bound for

$$\{f(x_0, y) : y \in Y\},$$

and hence there exists $y_0 \in Y$ such that

$$f(x_0, y_0) > \sup_{(x,y)\in X\times Y} f(x,y),$$

which is a contradiction. Assuming either quantity is strictly larger than the other brings about a contradiction, so we must have

$$\sup_{x\in X}\left(\sup_{y\in Y} f(x,y)\right) = \sup_{(x,y)\in X\times Y} f(x,y).$$

In precisely the same way one shows that

$$\sup_{y\in Y}\left(\sup_{x\in X} f(x,y)\right) = \sup_{(x,y)\in X\times Y} f(x,y).$$

This concludes the proof. $\square$

The reader will see that an analogous result can be stated and proven for infimums. In general one cannot, however, transpose the order in which an infimum and a supremum are taken.

1.5 Swapping Supremums and Limits

When working with double sequences, one may wish to alternate the order in which one is taking a limit and a supremum with respect to each of the two variables.

Theorem 1.15. Let f be a double sequence, and let $A \subset \mathbb{N}$. Suppose that, for each fixed $m \in A$,

$$\lim_{n\to\infty} f(m,n)$$

exists or tends to infinity. Suppose further that

$$\lim_{n\to\infty}\left(\sup_{m\in A} f(m,n)\right)$$

exists or tends to infinity. Then

$$\sup_{m\in A}\left(\lim_{n\to\infty} f(m,n)\right) \le \lim_{n\to\infty}\left(\sup_{m\in A} f(m,n)\right). \tag{1}$$

Proof. Clearly, for each fixed $(m, n) \in A \times \mathbb{N}$,

$$f(m, n) \leq \sup_{m \in A} f(m, n),$$

and so

$$\lim_{n \to \infty} f(m, n) \leq \lim_{n \to \infty} \left(\sup_{m \in A} f(m, n) \right). \tag{2}$$

Since $m \in A$ in the left hand side of equation (2) is arbitrary, we have

$$\sup_{m \in A} \left(\lim_{n \to \infty} f(m, n) \right) \leq \lim_{n \to \infty} \left(\sup_{m \in A} f(m, n) \right),$$

establishing (1). □

Corollary 1.16. Let f be a double sequence, $A \subset \mathbb{N}$. Suppose that, for each fixed $m \in A$, f is increasing in n. Then the limits in (1) exist or tend to infinity, and we have equality in (1).

Proof. If, for each fixed $m \in A$, f is increasing in n, then for each $(m, n) \in A \times \mathbb{N}$ we have

$$f(m, n) \leq \lim_{n \to \infty} f(m, n),$$

whence

$$\sup_{m \in A} f(m, n) \leq \sup_{m \in A} \left(\lim_{n \to \infty} f(m, n) \right). \tag{3}$$

Since the left hand side of (3) is increasing in n, its limit as n approaches infinity exists or tends to infinity, and

$$\lim_{n \to \infty} \left(\sup_{m \in A} f(m, n) \right) \leq \sup_{m \in A} \left(\lim_{n \to \infty} f(m, n) \right).$$

□

We remark that similar results can be stated and proven for swapping limits and infimums.

2 Banach Spaces

2.1 Norms

Definition 2.1. Given a vector space $\mathscr{X}$ over a field $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, a *norm* on $\mathscr{X}$ is a function $\|\cdot\| : \mathscr{X} \to [0, \infty)$ which satisfies

(i) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathscr{X}$ (the *triangle inequality*).

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{F}$ and $x \in \mathscr{X}$ (scaling).

(iii) $\|x\| = 0$ implies $x = 0$ (non-degeneracy).

If we relax the third condition then we have a *semi-norm*.

Usually we denote semi-norms by p, or $\{p_\alpha\}_{\alpha \in \mathscr{A}}$ for an indexed family of semi-norms, and reserve the $\|\cdot\|$ notation for norms. In a normed space $\mathscr{X}$ the norm gives rise to a metric, called the *norm metric*, defined by $d(x, y) = \|x - y\|$ for all $x, y \in \mathscr{X}$. Therefore metric space notions apply in a normed space. In particular, a normed space $\mathscr{X}$ has the metric topology, i.e., the open sets are the sets which are open with respect to the norm metric.

Definition 2.2. A normed space which is complete with respect to this metric is called a *Banach space*. We will usually assume our vector spaces to be over $\mathbb{C}$, noting where discrepancies arise.

Theorem 2.3. The function $x \mapsto \|x\|$ is continuous from $\mathscr{X}$ to $[0, \infty)$.

Proof. If $x_n \to x$ in $\mathscr{X}$, then

$$\big|\|x_n\| - \|x\|\big| \leq \|x_n - x\| \to 0,$$

i.e., $\|x_n\| \to \|x\|$. □

If $\mathscr{X}$ and $\mathscr{Y}$ are two normed spaces, then we can endow $\mathscr{X} \times \mathscr{Y}$ with the *product norm*, defined by $\|(x, y)\| = \max\{\|x\|, \|y\|\}$ for all pairs $(x, y) \in \mathscr{X} \times \mathscr{Y}$. We note that the equality

$$\|(x_n, y_n) - (x, y)\| = \|(x_n - x, y_n - y)\| = \max\{\|x_n - x\|, \|y_n - y\|\}$$

implies that $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ and $y_n \to y$.

Theorem 2.4. Addition $+ : \mathscr{X} \times \mathscr{X} \to \mathscr{X}$ and scalar multiplication $\cdot : \mathbb{C} \times \mathscr{X} \to \mathscr{X}$ are continuous.

Proof. If $(x_n, y_n) \to (x, y)$ in $\mathscr{X} \times \mathscr{X}$ then

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0,$$

which shows that addition is continuous. Now suppose that $(\alpha_n, x_n) \to (\alpha, x)$ in $\mathbb{C} \times \mathscr{X}$. Then $x_n \to x$, and hence x_n is bounded, say by M. Now

$$\begin{aligned} \|\alpha_n x_n - \alpha x\| &\leq \|\alpha_n x_n - \alpha x_n\| + \|\alpha x_n - \alpha x\| \\ &= |\alpha_n - \alpha| \|x_n\| + |\alpha| \|x_n - x\| \\ &\leq |\alpha_n - \alpha| M + |\alpha| \|x_n - x\| \\ &\to 0. \end{aligned}$$

Thus, scalar multiplication is continuous. □

2.2 Linear Transformations

We call a linear map between vector spaces a *linear operator* or a *linear transformation*, or simply a linear map. If $\mathscr{X}$ and $\mathscr{Y}$ are normed vector spaces, the set of continuous linear maps between them is denoted $\mathscr{B}(\mathscr{X}, \mathscr{Y})$. With addition and scalar multiplication defined pointwise, $\mathscr{B}(\mathscr{X}, \mathscr{Y})$ is a vector space. We say that a linear map $T : \mathscr{X} \to \mathscr{Y}$ is *bounded* if there exists a positive constant $c > 0$ such that $\|Tx\| \leq c\|x\|$ for all $x \in \mathscr{X}$.

Definition 2.5. If $\mathscr{X}$ and $\mathscr{Y}$ are normed vector spaces and T is a linear map between them, we define

$$\|T\| \equiv \inf\{c > 0 : \|Tx\| \leq c\|x\| \text{ for } x \in \mathscr{X}\}.$$

If there is no c such that $\|Tx\| < c\|x\|$ for all $x \in \mathscr{X}$, then $\|T\| = \inf \emptyset = \infty$. Thus T is bounded if and only if $\|T\| < \infty$. As the notation suggests, and as we verify presently, $T \mapsto \|T\|$ defines a norm on $\mathscr{B}(\mathscr{X}, \mathscr{Y})$, called the *operator norm*. The operator norm is sometimes denoted by $\| \cdot \|_{op}$.

Theorem 2.6. $T \mapsto \|T\|$ defines a norm on $\mathscr{B}(\mathscr{X}, \mathscr{Y})$.

Proof. Let $S, T \in \mathscr{B}(\mathscr{X}, \mathscr{Y})$. For each $x \in \mathscr{X}$ we have

$$\|(S + T)(x)\| \leq \|Sx\| + \|Tx\| \leq (\|S\| + \|T\|)\|x\|,$$

and hence $\|S+T\| \leq \|S\| + \|T\|$.

Now note that for $\alpha \neq 0$ and $c > 0$, we have $\|\alpha Tx\| \leq c\|x\|$ if and only if $\|Tx\| \leq \frac{c}{|\alpha|}\|x\|$. Thus

$$\begin{aligned} \|\alpha T\| &= \inf\{c > 0 : \|\alpha Tx\| \leq c\|x\|\} \\ &= \inf\left\{c > 0 : \|Tx\| \leq \tfrac{c}{|\alpha|}\|x\|\right\} \\ &= |\alpha|\|T\|. \end{aligned}$$

Lastly, if $\|T\| = 0$, then it is clear from the definition that $\|Tx\| \leq c\|x\|$ for all $c > 0$. If $x = 0$ then it is clear that $Tx = 0$. If $x \neq 0$ then

$$\|Tx\| \leq (c\|x\|^{-1})\|x\| = c$$

for all $c > 0$; this implies that $\|Tx\| = 0$, which in turn implies that $Tx = 0$. Thus we see that $\|T\| = 0$ implies $T = 0$. □

Proposition 2.7. If $\mathscr{X}$ and $\mathscr{Y}$ are normed vector spaces and T is a linear map between them, then the following are equivalent:

(i) $T \in \mathscr{B}(\mathscr{X}, \mathscr{Y})$.

(ii) T is continuous at 0.

(iii) T is continuous at some point.

(iv) T is bounded.

Proof. If (i) holds, i.e., if T is continuous, then (ii) and (iii) clearly hold. Suppose that T is continuous at some point $x_0 \in \mathscr{X}$. Then if $x \in \mathscr{X}$ and $x_n \to x$, we have $x_n - x + x_0 \to x_0$, and hence $T(x_n - x + x_0) \to T(x_0)$. Then since T is linear we have

$$T(x_n) - T(x) + T(x_0) \to T(x_0),$$

which means $T(x_n) \to T(x)$. We have thus shown that if T is continuous at some point then it is continuous at every point, i.e., (iii) implies (i).

Suppose now that T is continuous at 0; let $B(x,r)$ be our notation for the disk of radius r centered at x. Then there exists a $\delta > 0$ such that

$$T(B_{\mathscr{X}}(0,\delta)) \subset B_{\mathscr{Y}}(T(0),1) = B_{\mathscr{Y}}(0,1),$$

whence

$$T(\overline{B_{\mathscr{X}}(0,\delta)}) \subset \overline{T(B_{\mathscr{X}}(0,\delta))} \subset \overline{B_{\mathscr{Y}}(0,1)}.$$

In other words, $\|Tx\| \leq 1$ whenever $\|x\| \leq \delta$. Now for any $x \neq 0$ we have

$$\big\|\delta\|x\|^{-1}x\big\| = \delta,$$

and so

$$\big\|T(\delta\|x\|^{-1}x)\big\| \leq 1.$$

Therefore $\|T(x)\| \leq \delta^{-1}\|x\|$ for all $x \in \mathscr{X}$, and thus (ii) implies (iv).

Now suppose that T is bounded, i.e. there exists $C > 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in \mathscr{X}$. Then if $\|x_1 - x_2\| < C^{-1}\varepsilon$ we have

$$\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \leq C\|x_1 - x_2\| < \varepsilon,$$

so that T is continuous (in fact, it is uniformly continuous), and thus (iv) implies (i). $\square$

Proposition 2.8. For $T : \mathscr{X} \to \mathscr{Y}$ a linear map between normed vector spaces, we have

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| : \|x\| \leq 1\} \\ &= \sup\{\|Tx\| : \|x\| = 1\} \\ &= \sup\left\{\frac{\|Tx\|}{\|x\|} : \|x\| \neq 0\right\}.\end{aligned}$$

Proof. Let $\alpha = \sup\{\|Tx\| : \|x\| \leq 1\}$. We will first show that $\|T\| = \alpha$. For any $\varepsilon > 0$ and any $x \in \mathscr{X}$ we have, by definition of α,

$$\big\|T((\|x\| + \varepsilon)^{-1}x)\big\| \leq \alpha,$$

whence $\|Tx\| \leq \alpha(\|x\| + \varepsilon)$. Letting $\varepsilon \to 0$ we get $\|Tx\| \leq \alpha\|x\|$ for all $x \in \mathscr{X}$. Thus, $\|T\| \leq \alpha$. On the other hand, if $\|Tx\| \leq c\|x\|$ for all x, then for $\|x\| \leq 1$ we have $\|Tx\| \leq c\|x\| \leq c$, and hence $\alpha \leq c$. This means that α is a lower bound for

$$B = \{c > 0 : \|Tx\| \leq c\|x\| \text{ for } x \in \mathscr{X}\}.$$

But $\|T\|$ is the infimum of B, and hence $\alpha \leq \|T\|$. Therefore, $\alpha = \|T\|$.

Now to show the other equalities. Note that if $0 < \|x\| \leq 1$ then $\|x\|\|Tx\| \leq \|Tx\|$, and therefore

$$\|Tx\| \leq \|T(x/\|x\|)\|.$$

Since $\|x/\|x\|\| = 1$ for $x \neq 0$, it follows that

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

Lastly, since

$$\{x \in \mathscr{X} : \|x\| = 1\} = \left\{\frac{x}{\|x\|} : x \in \mathscr{X}, x \neq 0\right\},$$

we have

$$\sup\{\|Tx\| : \|x\| = 1\} = \sup\ \left\{\|T(\|x\|^{-1}x)\| : x \neq 0\right\} = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\},$$

as desired. □

We note that a useful fact comes out of the proof, namely that for any $T \in \mathscr{B}(\mathscr{X}, \mathscr{Y})$, we have $\|Tx\| \leq \|T\|\|x\|$ for all $x \in \mathscr{X}$.

Proposition 2.9. If $\mathscr{Y}$ is complete, so is $\mathscr{B}(\mathscr{X}, \mathscr{Y})$.

Proof. Let $\{T_n\}$ be a Cauchy sequence in $\mathscr{B}(\mathscr{X}, \mathscr{Y})$. Given $x \in \mathscr{X}$,

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|,$$

and hence $\{T_n x\}$ is Cauchy in $\mathscr{Y}$. Since $\mathscr{Y}$ is complete, $\{T_n x\}$ converges, and we can define a map $T : \mathscr{X} \to \mathscr{Y}$ by $Tx = \lim_{n\to\infty} T_n x$. It follows from the linearity of each T_n and the continuity of addition and scalar multiplication that T is linear:

$$T(x+y) = \lim_n T_n(x+y) = \lim_n (T_n x + T_n y) = Tx + Ty,$$

and

$$T(\alpha x) = \lim_n T_n(\alpha x) = \alpha \lim_n T_n x = \alpha T x.$$

Since $\{T_n\}$ is Cauchy, it is bounded, by M say. Then for each $x \in \mathscr{X}$, we have

$$\|Tx\| = \lim_n \|T_n x\| \leq \|T_n\|\|x\| \leq M\|x\|.$$

Thus T is bounded. Let $\varepsilon > 0$. There exists and N such that $\|T_m - T_n\| < \varepsilon$ whenever $m, n \geq N$. So for any $x \in \mathscr{X}$ and $n \geq N$ fixed,

$$\|T_n x - Tx\| = \lim_m \|T_n x - T_m x\| \leq \lim_m \|T_n - T_m\|\|x\| \leq \varepsilon\|x\|.$$

Therefore,

$$\|T_n - T\| \leq \varepsilon$$

for all $n \geq N$, proving that $T_n \to T$. □

3 Hilbert Spaces

3.1 Inner Products

Definition 3.1. Let $\mathscr{X}$ be a complex vector space. A *semi-inner product* on $\mathscr{X}$ is a map $u : \mathscr{X} \times \mathscr{X} \to \mathbb{C}$ such that:

(i) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$ for all $x, y, z \in \mathscr{X}$ and $\alpha, \beta \in \mathbb{C}$.

(ii) $u(y, x) = \overline{u(x, y)}$ for all $x, y \in \mathscr{X}$.

(iii) $u(x, x) \in [0, \infty)$ for all $x \in \mathscr{X}$.

Condition (i) may equivalently be expressed by the two conditions

(i′) $u(x + y, z) = u(x, z) + u(y, z)$ or all $x, y, z \in \mathscr{X}$, and

(i″) $u(\alpha x, y) = \alpha u(x, y)$ for all $x, y \in \mathscr{X}$, $\alpha \in \mathbb{C}$.

Notice that (i) and (ii) imply $u(x, \alpha y + \beta z) = \bar{\alpha} u(x, y) + \bar{\beta} u(x, z)$. We note in passing that if u is a semi-inner product, then p defined by $p(x) = \sqrt{u(x, x)}$ for all $x \in \mathscr{X}$ is a semi-norm, but proving this requires a generalized version of the Cauchy-Schwarz inequality for semi-inner products, which we shall omit.

Definition 3.2. A semi-inner product for which $x = 0$ whenever $u(x, x) = 0$ is called an *inner product*, and for these we write $\langle x, y\rangle$ instead of $u(x, y)$. A vector space equipped with an inner product is called a *pre-Hilbert space*, or more commonly, an *inner product space*.

Definition 3.3. If $\mathscr{X}$ is an inner product space, we define

$$\|x\| = \sqrt{\langle x, x\rangle}; \qquad x \in \mathscr{X}.$$

As the notation suggests, this defines a norm on $\mathscr{X}$, as we now proceed to show. If $x \in \mathscr{X}$ and $\alpha \in \mathbb{C}$ then

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = |\alpha|^2 \|x\|^2,$$

so that $\|\alpha x\| = |\alpha| \|x\|$. Moreover, if $\|x\| = 0$ then $\|x\|^2 = \langle x, x\rangle = 0$, which by definition of the inner product implies $x = 0$. To see that $\|x\| = \sqrt{\langle x, x\rangle}$ satisfies the triangle inequality we need the following important result:

Proposition 3.4 (The Cauchy-Schwarz Inequality)**.** Given an inner product $\langle\cdot,\cdot\rangle$, we have $|\langle x,y\rangle| \leq \|x\|\|y\|$ for all $x,y \in \mathscr{X}$, with equality if and only if x and y are linearly dependent.

Proof. The result is clear for $\langle x,y\rangle = 0$, so assume $\langle x,y\rangle \neq 0$. Let $\xi = \operatorname{sgn}\langle x,y\rangle$. Then for $t \in \mathbb{R}$ we have

$$\begin{aligned} 0 \leq \langle x - t\xi y, x - t\xi y\rangle &= \|x\|^2 - t\bar{\xi}\langle x,y\rangle - t\xi\langle y,x\rangle + t^2|\xi|^2\|y\|^2 \\ &= \|x\|^2 - 2t|\langle x,y\rangle| + t^2\|y\|^2 \\ &\equiv q(t). \end{aligned}$$

Since $q(t) \geq 0$ for all t, the discriminant of q must be less than or equal to 0, i.e.:

$$4|\langle x,y\rangle|^2 - 4\|y\|^2\|x\|^2 \leq 0,$$

i.e.:

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

as desired. Moreover, q will have a root only if this holds with equality. In that case, evaluating q at its root t_0 gives

$$0 = q(t_0) = \|x - t_0\xi y\|^2,$$

which implies $x = t_0\xi y$, i.e., that x and y are linearly dependent. Suppose, conversely that x and y are linearly dependent, i.e., that $y = \alpha x$ for some $\alpha \in \mathbb{C}$. Then

$$|\langle x,y\rangle| = |\langle x,\alpha x\rangle| = |\alpha|\|x\|^2 = \|x\||\alpha|\|x\| = \|x\|\|\alpha x\| = \|x\|\|y\|.$$

□

Now to see that $\|x\| \equiv \sqrt{\langle x,x\rangle}$ satisfies the triangle inequality, note that

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

so by the Cauchy-Schwarz inequality,

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

as desired. We can also prove that the inner product is continuous:

Proposition 3.5 (Continuity of the Inner Product)**.** If $x_n \to x$ and $y_n \to y$ then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.

Proof.

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\ &\leq |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \\ &\leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\|. \end{aligned}$$

The last quantity tends to zero because $y_n \to y$ implies that $\|y_n\|$ is bounded. □

The following identities, known as the *polarization identities*, are verified by straightforward computation. In a *real* inner product space $\mathscr{X}$ we have the identity

$$\langle x, y\rangle = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2\Big)$$

for all $x, y \in \mathscr{X}$. In a *complex* inner product space we have

$$\langle x, y\rangle = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2 + \imath\|x+\imath y\|^2 - \imath\|x-\imath y\|^2\Big) \tag{4}$$

for all $x, y \in \mathscr{X}$.

The *Parallelogram Law* (stated below) characterizes inner product spaces, in the sense that any normed vector space $(\mathscr{X}, \|\cdot\|)$ which satisfies it has an inner product – namely, as given by the relevant polarization identity. This characterization is known as the Fréchet – von Neumann – Jordan Theorem; we shall not prove it here, but refer the reader to [8], pp. 160–162 for the case of real normed vector spaces.

Proposition 3.6 (The Parallelogram Law)**.** For all $x, y \in \mathscr{X}$,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{5}$$

The name of this equality comes from the following interpretation: let x and y represent the two different sides of a parallelogram; then the equality says that the sum of the squares of the diagonals of the parallelogram is equal to the sum of the squares of the four sides. To obtain (5) in a complex inner product space, simply add the two equations

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

Definition 3.7. An inner product space that is complete with respect to the norm $\|x\| = \sqrt{\langle x, x\rangle}$ is called a *Hilbert space*. In what follows we use $\mathscr{H}$ to denote a Hilbert space.

We give a couple examples of Hilbert spaces, skipping over the details in verifying that they are indeed Hilbert spaces.

3.2 Examples

(a) $L^2(\mu)$. Let $(X, \mathscr{M}, \mu)$ be a measure space, and consider the measurable functions $f : X \to \mathbb{C}$ such that $\int |f|^2 \, d\mu < \infty$. The relation $f \sim g$ if $f = g$ a.e. is an equivalence relation on these functions; we use $L^2(\mu)$ or simply L^2 to denote the space of equivalence classes. We write $f \in L^2$ to indicate that f is a representative function in an equivalence class. It follows from the inequality $ab \leq \frac{1}{2}(a^2 + b^2)$, valid for all $a, b \geq 0$, that if $f, g \in L^2(\mu)$ then $|f\bar{g}| \leq \frac{1}{2}(|f|^2 + |g|^2)$, and hence $\int |f\bar{g}| \, \mathrm{d}\mu < \infty$. It is easy to see that

$$\langle f, g \rangle \equiv \int f\bar{g} \, \mathrm{d}\mu$$

defines an inner product on $L^2(\mu)$. In fact, $L^2(\mu)$ is a Hilbert space for any measure μ. This is proven in any standard measure and integration theory text.

(b) $l^2(I)$. If we consider the special case of the above example that we get by taking μ to be the counting measure on $(I, \mathscr{P}(I))$, where I is any nonempty set and $\mathscr{P}(I)$ is its power set, then we get the Hilbert space known as *little* l^2 *of* I. This is the space of functions $f : I \to \mathbb{C}$ such that the sum $\sum_{i \in I} |f(i)|^2$ is finite (we will discuss the meaning of this sum for I possibly uncountable presently). The inner product in this space is given by

$$\langle f, g \rangle \equiv \sum_{i \in I} f(i)\overline{g(i)}.$$

To shed a bit of light on this example, we need a definition and a lemma:

Definition 3.8. If I is an arbitrary set, and $f : I \to [0, \infty]$, we define $\sum_{i \in I} f(i)$ to be the supremum of its partial sums:

$$\sum_{i \in I} f(i) = \sup \Big\{ \sum_{i \in F} f(i) : F \subset I, \ F \text{ finite} \Big\}.$$

We say that the family $\{f(i) : i \in I\}$ is *summable*, or that the sum $\sum_{i \in I} f(i)$ *converges to a finite number*, if $\sum_{i \in I} f(i) < \infty$.

Lemma 3.9. If f is a nonnegative function defined on a set $E \subset I$ and μ is the counting measure, then $\int_E f \, \mathrm{d}\mu = \sum_{i \in E} f(i)$.

Proof. First note that if ϕ is a nonnegative simple function on X with standard representation

$$\phi = \sum_{j=1}^{n} \beta_j \chi_{\phi^{-1}(\beta_j)},$$

where $\{\beta_1, \ldots, \beta_n\}$ are the distinct nonzero elements in the range of ϕ, then

$$\begin{aligned}
\int_E \phi \, \mathrm{d}\mu &= \sum_1^n \beta_j \mu(\phi^{-1}(\beta_j) \cap E) \\
&= \sum_1^n \beta_j \sum_{i \in \phi^{-1}(\beta_j) \cap E} 1 \\
&= \sum_1^n \sum_{i \in \phi^{-1}(\beta_j) \cap E} \beta_j \\
&= \sum_1^n \sum_{i \in \phi^{-1}(\beta_j) \cap E} \phi(i) \\
&= \sum_{i \in \bigcup_1^n \phi^{-1}(\beta_j) \cap E} \phi(i) \\
&= \sum_{i \in E} \phi(i),
\end{aligned}$$

where the last two equalities are using the obvious fact that

$$E = \bigcup_1^n \phi^{-1}(\beta_j) \cap E$$

is a partition of $E \setminus \phi^{-1}(0)$. Thus,

$$\begin{aligned}
\int_E f \, \mathrm{d}\mu &= \sup_{\substack{0 \le \phi \le f \\ \phi \text{ simple}}} \int_E \phi \, \mathrm{d}\mu \\
&= \sup_{\substack{0 \le \phi \le f \\ \phi \text{ simple}}} \sum_{i \in E} \phi(i) \\
&= \sup_{\substack{0 \le \phi \le f \\ \phi \text{ simple}}} \sup_{\substack{F \subset E \\ F \text{ finite}}} \sum_{i \in F} \phi(i).
\end{aligned}$$

Now for any finite set F, we have that $f\chi_F$ is a simple function which is maximal among simple functions which are less than or equal to f on F. Hence, for F finite,

$$\sup_{\substack{0\leq\phi\leq f\\ \phi \text{ simple}}} \sum_{i\in F} \phi(i) = \sum_{i\in F} f(i).$$

Thus, using Theorem 1.14,

$$\int_E f \,\mathrm{d}\mu = \sup_{\substack{0\leq\phi\leq f\\ \phi \text{ simple}}} \sup_{\substack{F\subset E\\ F \text{ finite}}} \sum_{i\in F} \phi(i) = \sup_{\substack{F\subset E\\ F \text{ finite}}} \sum_{i\in F} f(i) = \sum_{i\in E} f(i).$$

□

Applying this lemma to our example shows that for $f \in L^2(\mu)$,

$$\|f\|^2 = \int_I |f|^2 = \sum_{i\in I} |f(i)|^2,$$

justifying our claim that $l^2(I)$—the set of functions $f : I \to \mathbb{C}$ such that $\sum_{i\in I} |f(i)|^2 < \infty$—is in fact $L^2(\mu)$ in the special case where μ is the counting measure on $(I, \mathscr{P}(I))$.

3.3 More on Possibly Uncountable Sums

In our last example we discussed infinite sums where the index set may be uncountable, so let us give a more general definition of the sum

$$\sum_{i\in I} f(i),$$

where f lands in some topological space X, and I may be uncountable. To do this, we introduce the concepts of *directed sets* and *nets*.

Definition 3.10. A directed set is a partially ordered set $(I, \leqslant)$ such that for any pair $i_1, i_2 \in I$ there exists $i_3 \in I$ with $i_3 \geqslant i_1$ and $i_3 \geqslant i_2$. Given a topological space X, a net in X is a pair $((I, \leqslant), f)$, where $(I, \leqslant)$ is a directed set and f is a function from I into X. We say that a net $((I, \leqslant), f)$ converges to $x \in X$, and that x is a *limit* of the net, if for every open neighbourhood $\mathscr{U}$ of x there exists an element $i_0 \in I$ such that $f(i) \in \mathscr{U}$ whenever $i \geq i_0$.

Definition 3.11. If I is a directed set and $f : I \to X$ is a function taking its values in a metric space (X, d), we say that the net $((I, \leqslant), f)$ is Cauchy if, for each $\varepsilon > 0$, there exists an element $i_0(\varepsilon) \in I$ such that $d(f(i), f(j)) < \varepsilon$ whenever $i, j \geq i_0(\varepsilon)$. A definition of completeness is as follows: a metric space (X, d) is complete if every Cauchy net converges in X. We will assume that all our Hilbert and Banach spaces are complete in this sense.

Exercise 3.12. A metric space X is complete if and only if every Cauchy sequence converges in X.

If we let $\mathscr{F}$ be the set of all finite subsets of I, then $\mathscr{F}$ is a directed set ordered by inclusion, i.e., $F \leqslant G$ if $F \subset G$. For each $F \in \mathscr{F}$, we define

$$f(F) = \sum_{i \in F} f(i),$$

where $f(i)$ are elements of X, as above; $(\mathscr{F}, f)$ is a net in X.

Definition 3.13. We say that the collection $\{f(i) : i \in I\} \subset X$ is *summable*, with sum x, if the net $\big((\mathscr{F}, \subset), f\big)$ converges to x; in this case we write $\sum_{i \in I} f(i) = x$. By an abuse of language, we sometimes speak of the net $\sum_{i \in I} f(i)$.

Remark 3.14. If $\{x_n\}$ is a sequence in X, the convergence of the net $\sum_{n \in \mathbb{N}} x_n$ with respect to set inclusion implies convergence of the sequence $\sum_{n=1}^{\infty} x_n$ in the usual sense. To see this, suppose that $x = \sum_{n \in \mathbb{N}} x_n$, and let $\varepsilon > 0$. There exists a finite subset $F(\varepsilon) \subset \mathbb{N}$ such that $\| \sum_{i \in F(\varepsilon)} x_i - x \| < \varepsilon$ for every finite $F \supset F(\varepsilon)$. Let $N \in \mathbb{N}$ be such that $F(\varepsilon) \subset \{1, 2, \dots, N\}$. Then for each $n \geq N$, $F(\varepsilon) \subset \{1, 2, \dots, n\}$, and so $\| \sum_{i=1}^{n} x_i - x \| < \varepsilon$. Thus $x = \sum_{n=1}^{\infty} x_n$.

Exercise 3.15. When does the converse hold?

Remark 3.16. It is not hard to see that our previous definition of convergence of a (possibly uncountable) infinite sum agrees with this more general definition in special case where $X = [0, \infty]$. For suppose that $f : I \to [0, \infty]$ is such that the net $\sum_{i \in I} f(i)$ converges to $x \in [0, \infty)$. Then for each $\varepsilon > 0$ there exists a finite set $F(\varepsilon)$ such that for any finite set $F \supset F(\varepsilon)$,

$$\Big| \sum_{i \in F} f(i) - x \Big| < \varepsilon.$$

So we can choose F to make $\sum_{i \in F} f(i)$ as close to x as we like. Since f is nonnegative, it follows that $\sum_{i \in F} f(i) \leq x$ for all finite $F \subset I$, and therefore

$$\sup\Big\{\sum_{i\in F} f(i) : F\subset I \text{ and } F \text{ finite}\Big\} = x.$$

In other words, $\sum_{i\in I} f(i) = x$ in the sense of Definition 3.8.

Conversely, suppose that

$$\sup\Big\{\sum_{i\in F} f(i) : F\subset I \text{ and } F \text{ finite}\Big\} = x < \infty.$$

Then for each $\varepsilon > 0$ we can choose a finite set $F(\varepsilon)$ such that

$$0 \le x - \sum_{i\in F(\varepsilon)} f(i) < \varepsilon.$$

For any $F \supset F(\varepsilon)$ we have $\sum_{i\in F(\varepsilon)} f(i) \le \sum_{i\in F} f(i) \le x$, and hence

$$\Big|x - \sum_{i\in F} f(i)\Big| \le \Big|x - \sum_{i\in F(\varepsilon)} f(i)\Big| < \varepsilon.$$

So $\sum_{i\in I} f(i) = x$ in the sense of Definition 3.13.

Proposition 3.17. If $\{x_i\}$ and $\{y_i\}$ are families of vectors in a Hilbert space $\mathscr{H}$, indexed by the same set I, such that $\sum_{i\in I} x_i = x$ and $\sum_{i\in I} y_i = y$, and if $\alpha \in \mathbb{C}$, then

(i) $\sum_{i\in I} \alpha x_i = \alpha x$,

(ii) $\sum_{i\in I}(x_i + y_i) = x + y$, and

(iii) $\sum_{i\in I}\langle x_i, z\rangle = \langle x, z\rangle$ and $\sum_{i\in I}\langle z, x_i\rangle = \langle z, x\rangle$ for every vector $z \in \mathscr{H}$.

Proof. The result follows trivially from the following three relations, which are valid for any finite $F \subset I$:

(i) $\|\alpha x - \sum_{i\in F} \alpha x_i\| = |\alpha|\|x - \sum_{i\in F} x_i\|$,

(ii) $\|(x + y) - \sum_{i\in F}(x_i + y_i)\| \le \|x - \sum_{i\in F} x_i\| + \|y - \sum_{i\in F} y_i\|$,

(iii) $|\langle x, z\rangle - \sum_{i\in F}\langle x_i, z\rangle| = |\langle x - \sum_{i\in F} x_i, z\rangle| \le \|x - \sum_{i\in F} x_i\|\|z\|$.

$\square$

Remark 3.18. Using Proposition 3.17 and looking at positive and negative, real and imaginary parts, one sees that Lemma 3.9 holds for all complex-valued functions.

3.4 The Pythagorean Theorem

We say that x is *orthogonal* to y if $\langle x, y\rangle = 0$. We now present the famous Pythagorean theorem.

Theorem 3.19 (The Pythagorean Theorem)**.** If $x_1, x_2, \ldots, x_n$ are pairwise orthogonal then

$$\Big\| \sum_1^n x_i \Big\|^2 = \sum_1^n \|x_i\|^2.$$

Proof. $\left\| \sum x_i \right\|^2 = \left\langle \sum x_i, \sum x_i \right\rangle = \sum_{i,j} \langle x_i, x_j\rangle = \sum \|x_i\|^2$, where the last equality is because $\langle x_i, x_j\rangle = 0$ whenever $i \neq j$. □

The continuity of the norm allows us to extend the Pythagorean Theorem to include infinite sums:

$$\Big\| \sum_{i=1}^{\infty} x_i \Big\|^2 = \Big\| \lim_{n\to\infty} \sum_{i=1}^{n} x_i \Big\|^2 = \lim_{n\to\infty} \Big\| \sum_{i=1}^{n} x_i \Big\|^2 = \lim_{n\to\infty} \sum_{i=1}^{n} \|x_i\|^2 = \sum_{i=1}^{\infty} \|x_i\|^2.$$

In fact, the theorem holds for uncountable sums. If $x = \sum_{i\in I} x_i$, where the x_i's are pairwise orthogonal, then by Lemma 2.5 (iii) we have

$$\|x\|^2 = \langle x, x\rangle = \sum_i \langle x_i, x\rangle = \sum_i \sum_j \langle x_i, x_j\rangle = \sum_i \langle x_i, x_i\rangle = \sum_i \|x_i\|^2.$$

3.5 Orthogonal Complements

For $E \subset \mathscr{H}$ we define $E^\perp$, the *orthogonal complement* of E in $\mathscr{H}$, to be the set

$$E^\perp = \{x \in \mathscr{H} : \langle x, y\rangle = 0 \text{ for all } y \in E\}.$$

By the linearity and continuity of the inner product in the first variable, $E^\perp$ is a closed subspace of $\mathscr{H}$. Also, for any $E \subset \mathscr{H}$, $E \cap E^\perp = 0$, for it is easy to see that $0 \in E \cap E^\perp$, and if $x \in E \cap E^\perp$, then $\|x\|^2 = \langle x, x\rangle = 0$ implies that $x = 0$. We present a few basic facts about orthogonal complements:

Theorem 3.20. If $E \subset \mathscr{H}$, then $E \subset E^{\perp\perp}$ (where $E^{\perp\perp}$ is the orthogonal complement of $E^\perp$).

Proof. If $x \in E$ and $y \in E^\perp$, then $x \perp y$, so $x \in E^{\perp\perp}$. Thus $E \subset E^{\perp\perp}$ □

Theorem 3.21. If $E \subset F$ then $F^\perp \subset E^\perp$.

Proof. If $x \in F^\perp$ then $x \perp y$ for all $y \in F$, and in particular $x \perp y$ for all $y \in E$, and hence $x \in E^\perp$. $\square$

Theorem 3.22. If $E \subset \mathscr{H}$, $E^\perp = E^{\perp\perp\perp}$.

Proof. Applying Theorem 3.20 with $E^\perp$ in place of E yields $E^\perp \subset E^{\perp\perp\perp}$, and applying Theorem 3.21 to the containment $E \subset E^{\perp\perp}$ yields the reverse containment. $\square$

3.6 Orthonormal Sets

We say that a subset $\{u_i\}_{i\in I}$ of $\mathscr{H}$ is *orthonormal* if the u_i's are pairwise orthogonal and of unit length. We now introduce the *Gram-Schmidt process*, a way of converting a sequence of linearly independent vectors $\{x_n\}_{n=1}^\infty$ into a sequence $\{u_n\}_{n=1}^\infty$ of orthonormal vectors such that the linear span of $\{x_n\}_1^N$ coincides with the linear span of $\{u_n\}_1^N$ for each N. The process is this: we take $u_1 = x_1/\|x_1\|$. Then we define each u_N in terms of the previous u_n's as follows: for each $N > 1$, we define v_N by $v_N = x_N - \sum_{n=1}^{N-1}\langle x_N, u_n\rangle u_n$, and take $u_N = v_N/\|v_N\|$. Since each u_n is a linear combination of $x_1, \dots x_n$, we have that x_N is not in the span of $u_1, \dots, u_{N-1}$, and hence $v_N \neq 0$; so we don't have to worry about division by zero in our definition of u_N. Moreover, the previously defined $u_1, \dots, u_{N-1}$ are pairwise orthogonal unit vectors, and hence for $m < N$ we have

$$\begin{aligned}
\langle v_N, u_m\rangle &= \left\langle x_N - \sum_{n=1}^{N-1}\langle x_N, u_n\rangle u_n, u_m\right\rangle \\
&= \langle x_N, u_m\rangle - \sum_{n=1}^{N-1}\langle x_N, u_n\rangle\langle u_n, u_m\rangle \\
&= \langle x_N, u_m\rangle - \langle x_N, u_m\rangle\|u_m\|^2 \\
&= 0.
\end{aligned}$$

Thus u_N is orthogonal to $u_1, \dots, u_{N-1}$. Since, as we've mentioned before, each u_n is a linear combination of $x_1, \dots, x_n$, we have

$$\operatorname{span}\{u_1, \dots, u_N\} \subset \operatorname{span}\{x_1, \dots, x_N\}$$

for each N. Also, for each n,

$$x_n = \|v_n\|u_n + \sum_{m=1}^{n-1}\langle x_n, u_m\rangle u_m \in \operatorname{span}\{u_1, \dots, u_n\},$$

and hence $\operatorname{span}\{x_1, \dots, x_N\} \subset \operatorname{span}\{u_1, \dots, u_N\}$.

3.7 Bessel's Inequality

Proposition 3.23. Given a set I and a function $f : I \to [0, \infty]$, if

$$\sum_{i \in I} f(i) < \infty,$$

then $f(i) \neq 0$ for at most a countable number of points.

Proof. Assume that $\sum_{i \in I} f(i) < \infty$ and let $I_n = \{i \in I : f(i) > \frac{1}{n}\}$, so that

$$\{i \in I : f(i) \neq 0\} = \bigcup_{n=1}^{\infty} I_n.$$

Then for any n and any finite $F \subset I_n$,

$$\sum_{i \in F} f(i) > \frac{|F|}{n},$$

where $|F|$ is the cardinality of F. Thus, if I_n is infinite, we have

$$\sum_{i \in I} f(i) \geq \sum_{i \in I_n} f(i) = \sup_{\substack{F \subset I_n \\ F \text{ finite}}} \sum_{i \in F} f(i) \geq \frac{1}{n} \sup_{\substack{F \subset I_n \\ F \text{ finite}}} |F| = \infty,$$

so each I_n must be finite. Hence, $\{i \in I : f(i) \neq 0\}$ is countable. $\square$

Proposition 3.24 (Bessel's Inequality)**.** If $\{u_i\}_{i \in I}$ is an orthonormal set in $\mathscr{H}$, then for any $x \in \mathscr{H}$,

$$\sum_{i \in I} |\langle x, u_i \rangle|^2 \leq \|x\|^2.$$

À propos of the preceding proposition, $\{i \in I : \langle x, u_i \rangle \neq 0\}$ is countable.

Proof. For any finite subset $F \subset I$ we have

$$\begin{aligned} 0 &\leq \Big\| x - \sum_{i \in F} \langle x, u_i \rangle u_i \Big\|^2 \\ &= \|x\|^2 - 2\mathrm{Re}\, \Big\langle x, \sum_{i \in F} \langle x, u_i \rangle u_i \Big\rangle + \Big\| \sum_{i \in F} \langle x, u_i \rangle u_i \Big\|^2 \\ &= \|x\|^2 - 2\mathrm{Re} \sum_{i \in F} \overline{\langle x, u_i \rangle} \langle x, u_i \rangle + \sum_{i \in F} |\langle x, u_i \rangle|^2 \\ &= \|x\|^2 - 2 \sum_{i \in F} |\langle x, u_i \rangle|^2 + \sum_{i \in F} |\langle x, u_i \rangle|^2 \\ &= \|x\|^2 - \sum_{i \in F} |\langle x, u_i \rangle|^2, \end{aligned}$$

where the Pythagorean theorem was used in the third line. So

$$\sum_{i \in F} |\langle x, u_i \rangle|^2 \leq \|x\|^2$$

for any finite $F \subset I$, and taking the supremum over the set of all such F's yields the result. $\square$

3.8 Orthogonal Projections

The following is essentially Theorem 5.24 in Folland, and we refer the reader to his book for a proof.

Theorem 3.25 (The Projection Theorem)**.** If $\mathscr{M}$ is a closed subspace of $\mathscr{H}$, then for each $x \in \mathscr{H}$ there exists a unique $y_0 \in \mathscr{M}$ such that

$$\|x - y_0\| = \operatorname{dist}(x, \mathscr{M}) \equiv \inf\{\|x - y\| : y \in \mathscr{M}\}.$$

Moreover, $x - y_0 \in \mathscr{M}^\perp$.

Corollary 3.26. Let $\mathscr{M} \leqslant \mathscr{H}$. For each $x \in \mathscr{H}$, there are unique elements $y_0 \in \mathscr{M}$ and $z_0 \in \mathscr{M}^\perp$ such that $x = y_0 + z_0$.

Proof. By the projection theorem there is a $y_0 \in \mathscr{M}$ such that $x - y_0 \in \mathscr{M}^\perp$. Let $z_0 = x - y_0$; then $x = y_0 + z_0$. If $x = y_1 + z_1$ with $y_1 \in \mathscr{M}$ and $z_1 \in \mathscr{M}^\perp$, then

$$(y_0 - y_1) + (z_0 - z_1) = 0,$$

whence, by the Pythagorean theorem,

$$\|y_0 - y_1\|^2 + \|z_0 - z_1\|^2 = 0,$$

whence $y_0 = y_1$ and $z_0 = z_1$. $\square$

Corollary 3.27. For each $x \in \mathscr{H}$, there is a unique element $y_0 \in \mathscr{M}$ such that $x - y_0 \in \mathscr{M}^\perp$.

Definition 3.28. Given $x \in \mathscr{H}$ and $\mathscr{M} \leqslant \mathscr{H}$, Corollary 3.26 determines a well-defined map $P_{\mathscr{M}} : \mathscr{H} \to \mathscr{M}$, given by $P_{\mathscr{M}} x = y_0$. We call $P_{\mathscr{M}}$ the *orthogonal projection of $\mathscr{H}$ onto $\mathscr{M}$*. Notice that $P_{\mathscr{M}} x$ is the unique element described in Corollary 3.27.

Theorem 3.29. For any $\mathscr{M} \leqslant \mathscr{H}$, $P_{\mathscr{M}}$ is a bounded linear transformation.

Proof. Let $x_1, x_2 \in \mathscr{H}$, and let $\alpha \in \mathbb{C}$. By Corollary 3.26 there exist $y_1, y_2 \in \mathscr{M}$ and $z_1, z_2 \in \mathscr{M}^\perp$ such that $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$. Therefore,

$$\begin{aligned} x_1 + \alpha x_2 &= y_1 + z_1 + \alpha(y_2 + z_2) \\ &= (y_1 + \alpha y_2) + (z_1 + \alpha z_2), \end{aligned}$$

whence

$$P_{\mathscr{M}}(x_1 + \alpha x_2) = y_1 + \alpha y_2 = P_{\mathscr{M}} x_1 + \alpha P_{\mathscr{M}} x_2.$$

To establish boundedness, we employ the Pythagorean theorem:

$$\|P_{\mathscr{M}} x_1\|^2 = \|y_1\|^2 \leq \|y_1\|^2 + \|z_1\|^2 = \|y_1 + z_1\|^2 = \|x_1\|^2.$$

In fact, we have $\|P_{\mathscr{M}}\|_{op} = 1$, since for any $0 \neq x \in \mathscr{M}$ we have $P_{\mathscr{M}} x = x$. □

Theorem 3.30. Given $P_{\mathscr{M}}$, we have

(i) $\ker P_{\mathscr{M}} = \mathscr{M}^\perp$; and

(ii) $\operatorname{ran} P_{\mathscr{M}} = \{x \in \mathscr{H} : P_{\mathscr{M}} x = x\} = \mathscr{M}$.

Proof. (i): If $x \in \ker P_{\mathscr{M}}$, then $P_{\mathscr{M}} x = 0$, so $x = x - P_{\mathscr{M}} x \in \mathscr{M}^\perp$. Conversely, if $x \in \mathscr{M}^\perp$, then $x - 0 \in \mathscr{M}^\perp$, and hence by uniqueness $P_{\mathscr{M}} x = 0$.

(ii): Clearly,

$$\{x \in \mathscr{H} : P_{\mathscr{M}} x = x\} \subset \operatorname{ran} P_{\mathscr{M}},$$

and by definition of $P_{\mathscr{M}}$ we have $\operatorname{ran} P_{\mathscr{M}} \subset \mathscr{M}$. Moreover, if $x \in \mathscr{M}$, then $x - x = 0 \in \mathscr{M}^\perp$, which by uniqueness implies $x = P_{\mathscr{M}} x$. Thus

$$\mathscr{M} \subset \{x \in \mathscr{H} : P_{\mathscr{M}} x = x\},$$

and we therefore have a circular chain of inclusions that gives the result. □

3.9 Vector Sums; Internal Direct Sums

As a particular case of the definition given in Section 3.3, we say that a family $\{x_i\}_{i \in I}$ of vectors in $\mathscr{H}$ is *summable*, with sum x, if the net $(\mathscr{F}, f)$—where $\mathscr{F}$ is the set of all finite subsets $F \subset I$ directed under inclusion, and $f : \mathscr{F} \to \mathscr{H}$ is the map $F \mapsto \sum_{i \in F} x_i$—converges to x. In this case we write $\sum_{i \in I} x_i = x$, and we say that the infinite vector sum $\sum_{i \in I} x_i$ converges in $\mathscr{H}$.

Theorem 3.31. Let $\{x_i\}$ be an orthogonal family of vectors in $\mathscr{H}$; that is, $x_i \perp x_j$ whenever $i \neq j$. Then $\{x_i\}$ is summable if and only if $\sum_{i \in I} \|x_i\|^2 < \infty$.

Proof. [1] The "only if" direction follows immediately from the Pythagorean theorem: if $x = \sum_{i \in I} x_i$, then

$$\sum_{i \in I} \|x_i\|^2 = \Big\| \sum_{i \in I} x_i \Big\|^2 = \|x\|^2 < \infty.$$

For the other direction, suppose that $\sum_{i \in I} \|x_i\|^2 < \infty$. By Proposition 3.23, the set $\{i \in I : x_i \neq 0\}$ is countable, so we can enumerate its elements: $\{x_1, x_2, \dots\}$. For any $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that

$$\sum_{n=N}^{\infty} \|x_i\|^2 < \varepsilon^2.$$

If we let $\mathscr{F}$ be the set of all finite subsets of I (directed by inclusion), and if for each $F \in \mathscr{F}$ we let $f(F) = \sum_{i \in F} x_i$, then $(\mathscr{F}, f)$ is a net in $\mathscr{H}$. Let $F_0 = \{1, \dots, N-1\}$ and suppose that $F, G \in \mathscr{F}$ are such that F and G both contain F_0. We will show that $\|f(F) - f(G)\| < \varepsilon$, so that the net $(\mathscr{F}, x)$ is Cauchy. By the Pythagorean theorem we have that

$$\begin{aligned}
\|f(F) - f(G)\|^2 &= \Big\| \sum_{i \in F} x_i - \sum_{i \in G} x_i \Big\|^2 \\
&= \Big\| \sum_{i \in F \setminus G} x_i - \sum_{i \in G \setminus F} x_i \Big\|^2 \\
&= \sum_{i \in F \triangle G} \|x_i\|^2,
\end{aligned}$$

where $F \triangle G$ denotes the symmetric difference $(F \setminus G) \cup (G \setminus F)$. Now $F \triangle G$ is finite and excludes F_0, so it follows that there is a natural number $M > N$ such that

$$\{i \in F \triangle G : x_i \neq 0\} \subset [N, M].$$

Hence,

$$\sum_{i \in F \triangle G} \|x_i\|^2 \leq \sum_{n=N}^{M} \|x_i\|^2 < \varepsilon^2.$$

So $(\mathscr{F}, x)$ is indeed a Cauchy net in $\mathscr{H}$, and since $\mathscr{H}$ is complete, this net converges. □

[1] for a different proof, see Halmos, §8

Definition 3.32. An arbitrary intersection of closed subspaces is still a closed subspace, so given a subset $E \subset \mathscr{H}$ we can take the intersection of all closed subspaces containing E to get a minimal closed subspace containing E. We denote this closed subspace by $\vee E$ and call it the *closed linear span of* E. We sometimes write $\mathscr{M} \leqslant \mathscr{H}$ to express that $\mathscr{M}$ is a closed subspace of $\mathscr{H}$.

Proposition 3.33. Given a nonempty subset $E \subset \mathscr{H}$, $\vee E \leqslant \mathscr{H}$, and $\vee E$ is the smallest closed subspace of $\mathscr{H}$ containing E. Moreover, $\vee E = \mathrm{cl}[\operatorname{span} E]$.

Proof. We know from topology that arbitrary intersections of closed sets are closed. It's also clear that vector subspaces are closed under intersection, so $\vee E$ is indeed a closed subspace. To see that $\vee E$ is the *smallest* closed subspace containing E, note that any closed subspace containing E is trivially contained in

$$\bigcap_{\substack{\mathscr{M} \supset E \\ \mathscr{M} \text{ a closed subspace}}} \mathscr{M} = \vee E.$$

Now to prove the second part of the proposition. Since $\vee E$ is a vector subspace containing E, it is clear that $\operatorname{span} E \subset \vee E$. But $\vee E$ is closed, so $\mathrm{cl}[\operatorname{span} E] \subset \vee E$. For the reverse inclusion, it suffices to show that $\mathrm{cl}[\operatorname{span} E]$ is a vector subspace, since $\mathrm{cl}[\operatorname{span} E] \supset E$ and $\vee E$ is the smallest closed subspace containing E. Let $x, y \in \mathrm{cl}[\operatorname{span} E]$ and $\alpha \in \mathbb{C}$. There are sequences $\{x_n\}, \{y_n\}$ in $\operatorname{span} E$ such that $x_n \to x$ and $y_n \to y$ in $\mathscr{H}$, and hence

$$\|x + y - (x_n + y_n)\| \leq \|x - x_n\| + \|y - y_n\| \to 0,$$

and

$$\|\alpha x - \alpha x_n\| = |\alpha| \|x - x_n\| \to 0.$$

So $\mathrm{cl}[\operatorname{span} E]$ is closed under addition and scalar multiplication. Since $\mathrm{cl}[\operatorname{span} E]$ is clearly nonempty, we have shown that $\mathrm{cl}[\operatorname{span} E]$ is a vector subspace, which completes the proof. We have also shown that the closure of *any* vector subspace is a (closed) vector subspace. □

Corollary 3.34. If $\mathscr{Y} \subset \mathscr{H}$ is a vector subspace, then $\vee \mathscr{Y} = \overline{\mathscr{Y}}$.

Definition 3.35. Given a family $\{\mathscr{M}_i\}_{i \in I}$ of closed subspaces, we define their vector sum, denoted $\sum_{i \in I} \mathscr{M}_i$, to be the collection of all vectors of the form $\sum_{i \in I} x_i$, where $\{x_i\}_{i \in I}$ is summable, with $x_i \in \mathscr{M}_i$ for each i. It is easy to verify that $\sum_{i \in I} \mathscr{M}_i$ is a vector subspace of $\mathscr{H}$.

Definition 3.36. Given a family $\{\mathscr{M}_i\}_{i\in I}$ of closed subspaces, we define their closed linear span, denoted $\bigvee_{i\in I}\mathscr{M}_i$, to be the closed linear span of $\bigcup_{i\in I}\mathscr{M}_i$ in $\mathscr{H}$.

Theorem 3.37. For any family $\{\mathscr{M}_i\}_{i\in I}$ of closed subspaces, $\bigvee_{i\in I}\mathscr{M}_i = \mathrm{cl}\big[\sum_{i\in I}\mathscr{M}_i\big]$.

Proof. Any element of $\mathrm{span}\{\bigcup_{i\in I}\mathscr{M}_i\}$ can be written as $\sum_{i\in I} x_i$ with $x_i \in \mathscr{M}_i$ for each i, and $x_i = 0$ for all but finitely many i; hence $\mathrm{span}\{\bigcup_{i\in I}\mathscr{M}_i\} \subset \sum_{i\in I}\mathscr{M}_i$, and taking closures gives $\bigvee_{i\in I}\mathscr{M}_i \subset \mathrm{cl}\big[\sum_{i\in I}\mathscr{M}_i\big]$.

On the other hand, it follows from the definition of infinite vector sums that $\mathrm{span}\{\bigcup_{i\in I}\mathscr{M}_i\}$ is dense in $\sum_{i\in I}\mathscr{M}_i$, i.e.,

$$\sum_{i\in I}\mathscr{M}_i \subset \mathrm{cl}\big[\mathrm{span}\{\bigcup_{i\in I}\mathscr{M}_i\}\big],$$

whence $\mathrm{cl}\big[\sum_{i\in I}\mathscr{M}_i\big] \subset \bigvee_{i\in I}\mathscr{M}_i$. $\square$

Definition 3.38. If $\{\mathscr{M}_i\}_{i\in I}$ is a family of closed subspaces such that $\mathscr{M}_i \perp \mathscr{M}_j$ whenever $i \neq j$, then we say that $\{\mathscr{M}_i\}_{i\in I}$ is an *orthogonal family* of closed subspaces.

The following is Theorem 13.2 in Halmos, and we essentially reproduce his proof here.

Theorem 3.39. If $\{\mathscr{M}_i\}_{i\in I}$ is an orthogonal family of closed subspaces, then $\bigvee_{i\in I}\mathscr{M}_i = \sum_{i\in I}\mathscr{M}_i$. Moreover, each element of $\sum_{i\in I}\mathscr{M}_i$ can be expressed uniquely in the form $\sum_{i\in I} x_i$ with $x_i \in \mathscr{M}_i$ for each i.

Proof. To prove the first part of the theorem is suffices to show that $\bigvee_{i\in I}\mathscr{M}_i \subset \sum_{i\in I}\mathscr{M}_i$; Theorem 3.37 gives the other inclusion. Let $x \in \bigvee_{i\in I}\mathscr{M}_i$. By Corollary 3.26, there exist, for each $i \in I$, $y_i \in \mathscr{M}_i$ and $z_i \in \mathscr{M}_i^{\perp}$ such that $x = y_i + z_i$. For each i such that $y_i \neq 0$ we have $\langle x, y_i/\|y_i\|\rangle = \|y_i\|$. It follows from Bessel's inequality that $\sum_{i\in I}\|y_i\|^2 < \infty$, whence by 3.31, $\{y_i\}$ is summable; call its sum x_0. If $\xi \in \mathscr{M}_j$ for some j, then by Proposition 3.17,

$$\begin{aligned}\langle x - x_0, \xi\rangle &= \langle y_j, \xi\rangle + \langle z_j, \xi\rangle - \langle x_0, \xi\rangle \\ &= \langle y_j, \xi\rangle + \langle z_j, \xi\rangle - \sum_{i\in I}\langle y_i, \xi\rangle \\ &= 0.\end{aligned}$$

Thus $x - x_0 \in \mathscr{M}_j^{\perp}$ for all j. It follows that $x - x_0 \perp \sum_{i\in I}\mathscr{M}_i$; by Theorem 3.37 and the continuity of the inner product, we have $x - x_0 \perp \bigvee_{i\in I}\mathscr{M}_i$. But $x - x_0 \in \bigvee_{i\in I}\mathscr{M}_i$, therefore $x - x_0 = 0$, proving that $x \in \sum_{i\in I}\mathscr{M}_i$.

For second part of the theorem, note that if $\sum_{i\in I} x_i = 0$, with $x_i \in \mathscr{M}_i$ for each i, then by the Pythagorean theorem, each $x_i = 0$. It follows that element of $\sum_{i\in I}\mathscr{M}_i$ can be expressed uniquely in the form $\sum_{i\in I} x_i$ with $x_i \in \mathscr{M}_i$ for each i. $\square$

Definition 3.40. If the subspaces $\{\mathscr{M}_i\}_{i\in I}$ of $\mathscr{H}$ are mutually orthogonal, then we say that $\mathscr{M} = \sum_{i\in I} \mathscr{M}_i$ is their *internal direct sum*; in this case, we sometimes write $\mathscr{M} = \sum_{i\in I} \oplus \mathscr{M}_i$.

3.10 More on Orthogonal Projections

Proposition 3.41. If $\mathscr{M} \leqslant \mathscr{H}$, then $I - P_{\mathscr{M}} = P_{\mathscr{M}^\perp}$.

Proof. Given $x \in \mathscr{H}$,

$$(I - P_{\mathscr{M}})x = x - P_{\mathscr{M}}x \in \mathscr{M}^\perp$$

by definition of $P_{\mathscr{M}}$. Moreover,

$$x - (I - P_{\mathscr{M}})x = P_{\mathscr{M}}x \in \mathscr{M} \subset \mathscr{M}^{\perp\perp}.$$

So $(I - P_{\mathscr{M}})x$ is the unique element of $\mathscr{M}^\perp$ such that $x - (I - P_{\mathscr{M}})x \in \mathscr{M}^{\perp\perp}$, which means that $I - P_{\mathscr{M}}$ is the projection of $\mathscr{H}$ onto $\mathscr{M}^\perp$. □

Corollary 3.42. If $\mathscr{M} \leqslant \mathscr{H}$, then $\mathscr{M}^{\perp\perp} = \mathscr{M}$.

Proof. By Theorem 3.30 (ii) we have

$$\mathscr{M} = \operatorname{ran} P_{\mathscr{M}} = \{x \in \mathscr{H} : P_{\mathscr{M}}x = x\}.$$

But $P_{\mathscr{M}}x = x$ if and only if $(I - P_{\mathscr{M}})x = 0$, so

$$\ker(I - P_{\mathscr{M}}) = \operatorname{ran} P_{\mathscr{M}}.$$

Now by Proposition 3.41 we have $I - P_{\mathscr{M}} = P_{\mathscr{M}^\perp}$, and by Theorem 3.30 (i) this means

$$\ker(I - P_{\mathscr{M}}) = \mathscr{M}^{\perp\perp}.$$

Altogether,

$$\mathscr{M}^{\perp\perp} = \ker(I - P_{\mathscr{M}}) = \operatorname{ran} P_{\mathscr{M}} = \mathscr{M}.$$

□

Corollary 3.43. If $E \subset \mathscr{H}$ then $E^{\perp\perp} = \vee E$.

Proof. $E \subset E^{\perp\perp}$, and $E^{\perp\perp}$ is a closed subspace of $\mathscr{H}$. Therefore, since $\vee E$ is the smallest closed subspace containing E, we have $\vee E \subset E^{\perp\perp}$. For the reverse inclusion, note that $E \subset \vee E$, so applying Theorem 3.21 yields $(\vee E)^\perp \subset E^\perp$, and applying Theorem 3.21 again to this, together with Corollary 3.42, yields $E^{\perp\perp} \subset (\vee E)^{\perp\perp} = \vee E$. □

Corollary 3.44. If $\mathscr{Y}$ is a vector subspace in $\mathscr{H}$, then $\mathscr{Y}$ is dense in $\mathscr{H}$ if and only if $\mathscr{Y}^\perp = 0$.

Proof. If $\mathscr{Y}^\perp = 0$, then

$$\mathscr{H} = 0^\perp = \mathscr{Y}^{\perp\perp} = \vee\mathscr{Y} = \overline{\mathscr{Y}}.$$

Conversely, if $\overline{\mathscr{Y}} = \mathscr{H}$, then $\mathscr{H} = \overline{\mathscr{Y}} = \vee\mathscr{Y} = \mathscr{Y}^{\perp\perp}$, and therefore

$$0 = \mathscr{H}^\perp = \mathscr{Y}^{\perp\perp\perp} = \mathscr{Y}^\perp.$$

□

3.11 Isomorphisms in Hilbert Space

A linear map $U : \mathscr{X} \to \mathscr{Y}$ is an *isomorphism* of normed vector spaces if it is a continuous bijection with continuous inverse. When such a linear map exists, $\mathscr{X}$ and $\mathscr{Y}$ are said to be *isomorphic*. If, in addition, U is an isometry—that is, a norm-preserving map—then U is an *isometric isomorphism*. When $\mathscr{X}$ and $\mathscr{Y}$ are Hilbert spaces, it turns out that requiring an isomorphism to be isometric ensures that it respects the inner product structure:

Proposition 3.45. For a linear map $U : \mathscr{H} \to \mathscr{K}$, the condition that

$$\langle Ux, Uy\rangle = \langle x, y\rangle \tag{6}$$

for all x, y is equivalent to U being an isometry.

Proof. If $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all x, y, then

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, x\rangle = \|x\|^2$$

so U is an isometry. Suppose, conversely, that U is an isometry. Then if $x, y \in \mathscr{H}$ and $\lambda \in \mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we have

$$\|x + \lambda y\|^2 = \|Ux + \lambda Uy\|^2.$$

Expanding both sides according to the identity $\|x\|^2 = \langle x, x\rangle$ yields

$$\|x\|^2 + 2\mathrm{Re}\,\bar{\lambda}\langle x, y\rangle + |\lambda|^2\|y\|^2 = \|Ux\|^2 + 2\mathrm{Re}\,\bar{\lambda}\langle Ux, Uy\rangle + |\lambda|^2\|Uy\|^2. \tag{7}$$

But since U is an isometry, we have $\|Ux\| = \|x\|$ and $\|Uy\| = \|y\|$, so that (7) becomes

$$\mathrm{Re}\,\bar{\lambda}\langle x, y\rangle = \mathrm{Re}\,\bar{\lambda}\langle Ux, Uy\rangle$$

for any $\lambda \in \mathbb{F}$. If $\mathbb{F} = \mathbb{R}$, take $\lambda = 1$. If $\mathbb{F} = \mathbb{C}$, take $\lambda = 1$ to see that $\mathrm{Re}\,\langle x, y\rangle = \mathrm{Re}\,\langle Ux, Uy\rangle$; since for any $z \in \mathbb{C}$ we have $\mathrm{Re}(-\imath z) = \mathrm{Im}(z)$, taking $\lambda = \imath$ shows that $\mathrm{Im}\,\langle x, y\rangle = \mathrm{Im}\,\langle Ux, Uy\rangle$. □

If U is a surjective linear map which satisfies (6), then it follows from the non-degeneracy property of the norm that U is injective, and it follows from Proposition 2.7 and the preceding proposition that U is a homeomorphism (that is, a continuous map with a continuous inverse). Thus any surjective linear map satisfying (6)—i.e., any surjective linear isometry—is an isometric isomorphism. Potentially confusing matters is the fact that some authors, such as Conway in [4], define an isomorphism between two Hilbert spaces to be any surjective map satisfying (6). This is understandable, because isometric isomorphisms are the structure-preserving maps between Hilbert spaces—the "true" isomorphisms, as Folland puts it. Some authors call these maps unitary maps, or unitary transformations, while others reserve the term "unitary" for isometric isomorphisms from a Hilbert space to itself. In linear algebra, equation (6) with inner product taken to be the usual scalar product on $\mathbb{C}^n$, is one of the equivalent conditions for a complex square matrix U to be called unitary.

Consider the map $f_y : \mathscr{H} \to \mathbb{C}$ given by $f_y(x) = \langle x, y\rangle$. The Cauchy-Schwarz inequality gives

$$|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|,$$

so f_y is bounded and $\|f_y\| \leq \|y\|$. But since

$$f_y(y/\|y\|) = \|y\|^2/\|y\| = \|y\|,$$

we have

$$\|f_y\| = \sup\{|f_y(x)| : \|x\| = 1\} \geq \|y\|$$

and hence $\|f_y\| = \|y\|$. So the map $y \mapsto f_y$ is an antilinear (or *conjugate linear*) isometry from $\mathscr{H}$ into $\mathscr{H}^*$, where $\mathscr{H}^*$ denotes the *dual space of* $\mathscr{H}$, i.e., the space of bounded linear functionals on $\mathscr{H}$ with the operator norm. It can be shown (see Conway [4], pp. 74–75, for example) that for any Banach space $\mathscr{B}$, the dual space $\mathscr{B}^*$ is a Banach space. Moreover, for a Hilbert space $\mathscr{H}$ the operator norm on $\mathscr{H}^*$ can be shown to satisfy the parallelogram law, making $\mathscr{H}^*$ a Hilbert space. This next theorem shows that this map is in fact surjective, and therefore the antilinear map $y \mapsto f_y$ is an anti-isomorphism from $\mathscr{H}$ to its dual.

Theorem 3.46 (The Riesz Representation Theorem)**.** Given any $\xi \in \mathscr{H}^*$, there exists a unique $y \in \mathscr{H}$ such that $\xi(x) = \langle x, y\rangle$ for all $x \in \mathscr{H}$, i.e. $\xi = f_y$.

Proof. The uniqueness part is easy: if $\langle x, y\rangle = \langle x, y'\rangle$ for all $x \in \mathscr{H}$, then taking $x = y - y'$ gives $\|y - y'\|^2 = 0$, and hence $y = y'$. To show existence, let $\mathscr{M} = \ker \xi$. If ξ is the zero functional then we simply take $y = 0$, so we may assume $\mathscr{M} \neq \mathscr{H}$.

Then $\mathscr{M}^\perp \neq 0$; for if $\mathscr{M}^\perp$ were 0 then since it is easy to show that $\mathscr{M} \leqslant \mathscr{H}$, we would have

$$\mathscr{M} = \mathscr{M}^{\perp\perp} = 0^\perp = \mathscr{H}.$$

Now, there exists $y_0 \in \mathscr{M}^\perp$ with $\|y_0\| = 1$. For $x \in \mathscr{H}$, if $u = \xi(x)y_0 - \xi(y_0)x$ then

$$\xi(u) = \xi(x)\xi(y_0) - \xi(y_0)\xi(x) = 0,$$

so $u \in \mathscr{M}$. Hence,

$$0 = \langle u, y_0 \rangle = \xi(x)\|y_0\|^2 - \xi(y_0)\langle x, y_0 \rangle = \xi(x) - \langle x, \overline{\xi(y_0)}y_0 \rangle.$$

Thus $\xi(x) = \langle x, y \rangle$, where $y = \overline{\xi(y_0)}y_0$. □

3.12 Basis in Hilbert Space

While Banach spaces inherit the notions of basis and dimension from their vector space structure (Hamel basis and dimension), the structure of the inner product on a Hilbert space $\mathscr{H}$ gives it another notion of basis: *orthonormal basis.*

Definition 3.47. An *orthonormal basis* for a Hilbert space $\mathscr{H}$ is a maximal orthonormal subset of $\mathscr{H}$.

Given any Hilbert space $\mathscr{H}$, the set of orthonormal subsets of $\mathscr{H}$ is a partially ordered set, ordered by inclusion. An application of Zorn's lemma shows that $\mathscr{H}$ has an orthonormal basis.

A remark on the distinction between the concepts of orthonormal basis and Hamel basis: a Hamel basis for a vector space V is a maximal linearly independent subset of V, or equivalently, a linearly independent subset which spans V. A finite orthonormal basis is a Hamel basis (as we shall later see), but it can be shown that an infinite orthonormal basis for a Hilbert space is never a Hamel basis (see [4], p. 19).

Theorem 3.48 (Basis Theorem)**.** If $\{u_i\}_{i\in I}$ is an orthonormal set in $\mathscr{H}$, then the following are equivalent:

(i) $\{u_i\}_{i\in I}$ is an orthonormal basis for $\mathscr{H}$.

(ii) $\{u_i : i \in I\}^\perp = 0$.

(iii) $\bigvee_{i\in I} u_i = \mathscr{H}$.

(iv) If $x \in \mathscr{H}$, then $x = \sum_{i\in I}\langle x, u_i\rangle u_i$.

Proof. (i) $\implies$ (ii): If $x \perp \{u_i\}_{i\in I}$ and $x \neq 0$, then $\{u_i\}_{i\in I} \cup \left\{\frac{x}{\|x\|}\right\}$ is an orthornormal set which properly contains $\{u_i\}_{i\in I}$, contradicting maximality.

(ii) $\iff$ (iii): It follows from Corollary 3.44 that $\operatorname{cl}\left[\operatorname{span}\{u_i\}_{i\in I}\right] = \mathscr{H}$ if and only if $\left(\operatorname{span}\{u_i\}_{i\in I}\right)^{\perp} = 0$. It is also easy to see that $\left(\operatorname{span}\{u_i\}_{i\in I}\right)^{\perp} = 0$ if and only if $\{u_i : i \ \in I\}^{\perp} = 0$. This, together with the fact that

$$\operatorname{cl}\left[\operatorname{span}\{u_i\}_{i\in I}\right] = \bigvee_{i\in I} u_i$$

gives the equivalence of (ii) and (iii).

(ii) $\implies$ (iv): If $x \in \mathscr{H}$, let $y = x - \sum_{i\in I}\langle x, u_i\rangle u_i$; y is well-defined by Proposition 3.31. For any $u_j \in \{u_i\}_{i\in I}$ we have

$$\begin{aligned}
\langle y, u_j\rangle &= \left\langle x - \sum_{i\in I}\langle x, u_i\rangle u_i, u_j\right\rangle \\
&= \langle x, u_j\rangle - \sum_{i\in I}\langle x, u_i\rangle\langle u_i, u_j\rangle \quad \text{by Proposition 3.17} \\
&= \langle x, u_j\rangle - \langle x, u_j\rangle \\
&= 0.
\end{aligned}$$

Since this holds for any $u_j \in \{u_i\}_{i\in I}$, we have $y \in \{u_i : i \ \in I\}^{\perp}$, and hence $y = 0$.

(iv) $\implies$ (i): Suppose that (iv) holds, but $\{u_i\}_{i\in I}$ is not a maximal orthonormal set in $\mathscr{H}$. Then there exists a unit vector $\nu \in \{u_i : i \ \in I\}^{\perp}$. Writing $\nu = \sum_{i\in I}\langle \nu, u_i\rangle u_i$ and applying the Pythagorean theorem gives

$$\|\nu\|^2 = \sum_{i\in I} |\langle \nu, u_i\rangle|^2 = 0,$$

a contradiction. □

The coefficients $\langle x, u_i\rangle$ in (iv) are called the *Fourier coefficients* of x relative to the basis $\{u_i\}_{i\in I}$, and the series representation in (iv) is called the *Fourier expansion* or *Fourier series* for x.

Theorem 3.49 (Parseval's Identity)**.** The family $\{u_i\}_{i\in I}$ is an orthonormal basis for $\mathscr{H}$ if and only if

$$\|x\|^2 = \sum_{i\in I} |\langle x, u_i\rangle|^2 \quad \text{for each} \quad x \in \mathscr{H}. \tag{8}$$

Proof. If $\{u_i\}_{i\in I}$ is an orthonormal basis, then applying the Pythagorean theorem to the Fourier series for x gives (8). Conversely, if $\{u_i\}_{i\in I}$ is not an orthonormal basis, then there exists a unit vector ν such that $\nu \perp u_i$ for all i; then

$$\sum_{i\in I} |\langle \nu, u_i\rangle|^2 = 0,$$

and hence (8) does not hold. □

Remark 3.50. The Fourier series $\sum_{i\in I}\langle x, u_i\rangle u_i$ for x converges whenever $\{u_i\}_{i\in I}$ is an orthonormal family—this follows from Theorem 3.31 together with Bessel's inequality. Parseval's identity tells us that when $\{u_i\}_{i\in I}$ is an orthonormal basis, Bessel's inequality holds *with equality*; and *vice-versa*.

3.13 The Dimension of a Hilbert Space

Proposition 3.51. Every finite-dimensional vector subspace of a Hilbert space is closed.

Proof. [2] Let $\mathscr{M} \subset \mathscr{H}$ be a linear subspace with Hamel basis $\{x_1, x_2, \ldots, x_n\}$. Let's assume that $\mathscr{M}$ is not closed, and let $y \in \overline{\mathscr{M}} \setminus \mathscr{M}$. Then $x_1, x_2, \ldots, x_n, y$ are linearly independent, and hence, by Gram-Schmidt, there exist $u_1, u_2, \ldots, u_{n+1}$ such that u_i are pairwise orthonormal,

$$\operatorname{span}\{u_1, \ldots, u_i\} = \operatorname{span}\{x_1, \ldots, x_i\} \tag{9}$$

for $i = 1, \ldots, n$, and

$$\operatorname{span}\{u_1, \ldots, u_n, u_{n+1}\} = \operatorname{span}\{x_1, \ldots, x_n, y\}. \tag{10}$$

Then (9) implies that $\overline{\mathscr{M}} = \operatorname{cl}\big[\operatorname{span}\{u_1, \ldots, u_n\}\big]$, i.e.,

$$\overline{\mathscr{M}} = \bigvee\{u_1, \ldots, u_n\},$$

and therefore $\{u_1, \ldots, u_n\}$ is an orthonormal basis for $\overline{\mathscr{M}}$. But (10) says that

$$\{u_1, \ldots, u_{n+1}\} \subset \overline{\mathscr{M}},$$

contradicting the maximality of $\{u_1, \ldots, u_n\}$ as an orthonormal set in $\overline{\mathscr{M}}$. Hence $\overline{\mathscr{M}} \setminus \mathscr{M} = \emptyset$, so that $\mathscr{M} = \overline{\mathscr{M}}$ is closed.

□

[2] It is true more generally that every finite-dimensional vector subspace of a normed vector space is closed—see Loomis & Sternberg [9], p. 209; their proof makes implicit use of $\mathbf{AC}_{\boldsymbol{\omega}}$, the axiom of countable choice.

Proposition 3.52. Any two orthonormal bases of a Hilbert space $\mathscr{H}$ have the same cardinality.

Proof. Let $\{u_i\}_{i\in I}$ and $\{v_j\}_{j\in J}$ be two bases for $\mathscr{H}$. Let $|I|$ denote the cardinality of I, and $|J|$ the cardinality of J. Suppose that I is finite and $|I| < |J|$. The u_i are linearly independent, for if

$$\alpha_1 u_1, \ldots, \alpha_{|I|} u_{|I|} = 0,$$

then for $i \in \{1, 2, \ldots, |I|\}$ we have

$$0 = \langle \alpha_1 u_1, \ldots, \alpha_{|I|} u_{|I|}, u_i \rangle = \alpha_i \|u_i\|^2 = \alpha_i.$$

Moreover, we have from (iv) of Theorem 3.48 that $\mathscr{H} = \operatorname{span}\{u_1, \ldots, u_{|I|}\}$, so $\{u_i\}_{i\in I}$ is in fact a Hamel basis for $\mathscr{H}$. Choose any $|I| + 1$ vectors in $\mathscr{F}$. By the same argument used above, these vectors are linearly independent. Moreover, they are in $\mathscr{H} = \operatorname{span}\{u_1, \ldots, u_{|I|}\}$, which is a contradiction because, as we know from basic linear algebra, we cannot have $|I| + 1$ linearly independent vectors in a linear subspace spanned by $|I|$ vectors. Thus if at least one of the two bases is finite, they must have the same cardinality.

Suppose that both bases are infinite. For each $i \in I$, set

$$J_i \equiv \{j \in J : \langle u_i, v_j \rangle \neq 0\}.$$

Since $\{v_j\}_{j\in J}$ is an orthonormal set, we have that J_i is countable for each i. Now by (ii) of Theorem 3.48, each $j \in J$ must belong to at least one J_i, and hence $J = \bigcup_{i\in I} J_i$. Therefore, $|J| \leq |I| \cdot \aleph_0 = |I|$. Similarly, $|I| \leq |J|$. $\square$

Definition 3.53. The *dimension* of a Hilbert space $\mathscr{H}$ is the cardinality of any basis for $\mathscr{H}$, and is denoted by $\dim \mathscr{H}$.

Lemma 3.54. If (X, d) is a separable metric space and $\{B_i\}_{i\in I}$ is a collection of disjoint open balls in X, then I must be countable.

Proof. Let $\mathscr{D}$ be a countable dense subset of X. Then $B_i \cap \mathscr{D} \neq \emptyset$ for each $i \in I$. Thus there is an element $x_i \in B_i \cap \mathscr{D}$ for each i, and since the B_i are disjoint,

$$x_i \leftrightarrow B_i$$

is a 1 to 1 correspondence. Thus[3] there is collection of points $\{x_i : i \in I\} \subset \mathscr{D}$ with the cardinality of I, which proves that I is countable. $\square$

[3]by $\mathbf{AC}_{\boldsymbol{\omega}}$, the axiom of countable choice

Proposition 3.55. If $\mathscr{H}$ is an infinite dimensional Hilbert space, then $\mathscr{H}$ is separable if and only if $\dim \mathscr{H} = \aleph_0$.

Proof. We prove this for $\mathscr{H}$ a complex Hilbert space—when $\mathscr{H}$ is a Hilbert space over the reals, the proof is even simpler. Suppose that $\dim \mathscr{H} = \aleph_0$ and let $\{u_i\}_{i\in I}$ be a basis for $\mathscr{H}$. Since $\{u_i\}_{i\in I}$ is countable we can enumerate its elements: $\{u_1, u_2, \dots\}$.

Let
$$\mathbb{Q} + \imath\mathbb{Q} = \{q + r\imath : q, r \in \mathbb{Q}\};$$
it is easy to show that $\mathbb{Q} + \imath\mathbb{Q}$ is dense in $\mathbb{C}$. Let
$$\mathscr{D}_0 = \Big\{ \sum_{j=1}^{n} \beta_j u_j : \beta_j \in \mathbb{Q} + \imath\mathbb{Q}, n \in \mathbb{N} \Big\},$$
and let
$$\mathscr{D} = \Big\{ \sum_{j=1}^{n} \alpha_j u_j : \alpha_j \in \mathbb{C}, n \in \mathbb{N} \Big\}.$$
It is easy to verify that $\mathscr{D}_0$ is dense in $\mathscr{D}$. Now let $x \in \mathscr{H}$; by (iv) of Theorem 3.48, $x = \sum_{i=1}^{\infty} \langle x, u_i \rangle u_i$, and the partial sums of this infinite series are in $\mathscr{D}$; so $\mathscr{D}$ is dense in $\mathscr{H}$. Together with the fact that $\mathscr{D} \subset \overline{\mathscr{D}_0}$, we have $\overline{\mathscr{D}_0} = \overline{\mathscr{D}} = \mathscr{H}$.

Conversely, suppose that $\mathscr{H}$ is separable and let $\{u_i\}_{i\in I}$ be a basis. Given $u_i, u_j \in \{u_i\}_{i\in I}$ with $i \neq j$, we have
$$\|u_i - u_j\|^2 = \|u_i\|^2 + \|u_j\|^2 = 2,$$
and hence $\big\{B(u_i, \frac{1}{\sqrt{2}}) : i \in I\big\}$ is a collection of disjoint open balls in $\mathscr{H}$. By Lemma 3.54, $\{u_i\}_{i\in I}$ is countable; therefore $\dim \mathscr{H} = \aleph_0$. □

Proposition 3.56. Consider the Hilbert space $l^2(I)$. The collection $\{\delta_i\}_{i\in I}$ given by
$$\delta_i(k) = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k, \end{cases}$$
is an orthonormal basis. Therefore the dimension of $l^2(I)$ is $|I|$.

Proof. For any $i \in I$ we have $\|\delta_i\|^2 = \sum_{k\in I} |\delta_i(k)|^2 = 1$, so each δ_i is a unit vector. If $i \neq j$ then
$$\langle \delta_i, \delta_j \rangle = \sum_{k\in I} \delta_i(k)\overline{\delta_j(k)} = 0,$$

so the δ_i are pairwise orthogonal. Thus, $\{\delta_i\}_{i\in I}$ is an orthonormal set in $l^2(I)$. It is easy to see that any $f \in l^2(I)$ may be written as $f = \sum_{i\in I} f(i)\delta_i$, and for any $i \in I$ we have

$$\langle f, \delta_i\rangle \equiv \sum_{k\in I} f(k)\overline{\delta_i(k)} = f(i).$$

Thus $f = \sum_{i\in I}\langle f, \delta_i\rangle\delta_i$, which, by Theorem 3.48, proves that $\{\delta_i\}_{i\in I}$ is a basis for $l^2(I)$. □

Theorem 3.57. If $\mathscr{H}$ is a Hilbert space with orthonormal basis $\{u_i\}_{i\in I}$, then $\mathscr{H}$ is isometrically isomorphic to $l^2(I)$. It follows that two Hilbert spaces are isometrically isomorphic if and only if they have the same dimension.

Proof. For each $x \in \mathscr{H}$, define $\mathbf{U}(x) : I \to \mathbb{C}$ by $\mathbf{U}(x)(i) = \langle x, u_i\rangle$. Observe that $\mathbf{U}(x) \in l^2(I)$, since by Bessel's inequality, $\sum_{i\in I} |\mathbf{U}(x)(i)|^2 \le \|x\|^2$. Moreover, $\mathbf{U}$ is linear because the inner product is linear in the first variable, and $\mathbf{U}$ is an isometry because, by the (iv) of Basis Theorem together with the Pythagorean theorem, we have

$$\|\mathbf{U}x\|^2 = \sum_{i\in I} |\langle x, u_i\rangle|^2 = \|x\|^2.$$

Now let $f \in l^2(I)$; then

$$\sum_{i\in I} \|f(i)u_i\|^2 = \sum_{i\in I} |f(i)|^2 < \infty,$$

so $\sum_{i\in I} f(i)u_i$ converges in $\mathscr{H}$. Since $\mathbf{U}$ is linear and continuous, we may write

$$f = \sum_{i\in I} f(i)\mathbf{U}(i) = \mathbf{U}\big(\sum_{i\in I} f(i)u_i\big);$$

therefore $\mathbf{U}$ is surjective, and is therefore an isometric isomorphism.

We now turn to the second statement in the theorem. Suppose that $U : \mathscr{H} \to \mathscr{K}$ is an isomorphism of Hilbert spaces. Let $\{u_i\}_{i\in I}$ be a basis for $\mathscr{H}$, and let $v_i = U(u_i)$ for each i. We will show that $\{v_i : i \in I\}$ is a basis for $\mathscr{K}$, thereby showing that $\mathscr{H}$ and $\mathscr{K}$ have the same dimension. It's clear that v_i is a unit vector for each i, as U is an isometry. That the v_i are mutually orthogonal comes from the identity $\langle Ux, Uy\rangle = \langle x, y\rangle$. Lastly, the maximality of $\{v_i\}_{i\in I}$ comes from the maximality of $\{u_i\}_{i\in I}$; for if there existed a unit vector $\eta \perp \{v_i : i \in I\}$, then $\eta = U\xi$ for some unit vector $\xi \in \mathscr{H}$, $\xi \notin \{u_i\}_{i\in I}$; but then

$$\langle \xi, u_i\rangle = \langle \eta, v_i\rangle = 0,$$

contradicting the maximality of $\{u_i\}_{i\in I}$ as an orthonormal set in $\mathscr{H}$.

On the other hand, if $\mathscr{H}$ and $\mathscr{K}$ have the same dimension, we can index their bases with the same set I, and it follows from the first statement in the theorem that they are isometrically isomorphic. $\square$

3.14 Example: The Fourier Transform on the Circle.

We wish to give an important example of the isomorphism described in Theorem 3.57. To this end, we cover some material from complex analysis, including the Weierstrass Approximation Theorem for the circle in $\mathbb{C}$. Generally, we follow Conway's books [4],[5], filling in details and working through what he leaves as exercise.

Consider the space $L^2[0,2\pi]$; this is a Hilbert space with inner product

$$\langle f,g\rangle_{L^2[0,2\pi]} = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\overline{g(\theta)}\,\mathrm{d}\theta.$$

Let D denote the open disk of radius 1 in $\mathbb{C}$, and let ∂D denote its boundary. Let $\gamma : [0,2\pi]\to\partial D$ be the curve defined by $\gamma(\theta)=\mathrm{e}^{\imath\theta}$. We endow $L^2[\partial D]$ with the inner product

$$\begin{aligned}
\langle f,g\rangle_{L^2[\partial D]} &= \frac{1}{2\pi}\int_\gamma f(z)\overline{g(z)}|\,\mathrm{d}z| \\
&\equiv \frac{1}{2\pi}\int_0^{2\pi} f(\gamma(\theta))\overline{g(\gamma(\theta))}|\gamma'(\theta)|\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\mathrm{e}^{\imath\theta})\overline{g(\mathrm{e}^{\imath\theta})}|\imath\mathrm{e}^{\imath\theta}|\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(\mathrm{e}^{\imath\theta})\overline{g(\mathrm{e}^{\imath\theta})}\,\mathrm{d}\theta \\
&= \langle f\circ\gamma, g\circ\gamma\rangle_{L^2[0,2\pi]}.
\end{aligned}$$

Every $f\in L^2[\partial D]$ maps to an element $f\in L^2[0,2\pi]$ via $f\mapsto f\circ\gamma$. This map is surjective because γ is a bijection from $[0,2\pi)$ onto ∂D, and elements of L^2 are not determined by values on sets of measure zero. Thus $f\mapsto f\circ\gamma$ is an isometric isomorphism from $L^2[\partial D]$ onto $L^2[0,2\pi]$.

We record a basic fact from complex analysis:

Proposition 3.58 (Lemma 1.0 in [10])**.** Suppose $(c_n)_{n=0}^\infty$ is a sequence of complex numbers, and define $R\in[0,\infty]$ by

$$R=\sup\{\rho\geq 0 : \text{the sequence } (c_n\rho^n) \text{ is bounded}\}$$

Then the series $\sum_{n=0}^{\infty} |c_n||z - z_0|^n$ converges uniformly on every compact subset of the ball $B(z_0, R)$, and the power series $\sum_{n=0}^{\infty} c_n(z - z_0)^n$ diverges at every point z such that $|z - z_0| > R$.

Theorem 3.59 (The Weierstrass Approximation Theorem for ∂D)**.** Let $C[\partial D]$ be the set of continuous, complex-valued functions on ∂D. The polynomials in $z, \overline{z}$ are uniformly dense in $C[\partial D]$. In other words, for each $f \in C[\partial D]$, there exists a sequence of polynomials $p_N(z, \overline{z})$ that converges to f uniformly on ∂D.

Proof. This is part of exercise 3 in [5], pp. 262–263. First we show that if $g : \overline{D} \to \mathbb{C}$ is a continuous function and, for $0 \leq r < 1$, g_r is defined on ∂D by $g_r(z) = g(rz)$, then $g_r \to g$ uniformly on ∂D as $r \to 1^-$. To see this, first note that since g is continuous and $\overline{D}$ is compact, g is uniformly continuous on $\overline{D}$. For all $z \in \partial D$, we have

$$|z - rz| = (1 - r)||z| = 1 - r,$$

which tends to 0 as $r \to 1^-$; thus, by the uniform continuity of g on $\overline{D}$, we have $\|g - g_r\|_{\text{unif}} \to 0$ as $r \to 1^-$.

Now, if $f : \partial D \to \mathbb{C}$ is a continuous function, we can define a function $\tilde{f} : \overline{D} \to \mathbb{C}$ by $\tilde{f}(z) = f(z)$ for $z \in \partial D$, and

$$\tilde{f}(re^{\imath\theta}) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{\imath t}) P_r(\theta - t)\, dt,$$

for $0 \leq r < 1$, where P_r is the *Poisson kernel*, defined by

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{\imath\theta n},$$

for $0 \leq r < 1$ and $-\infty < \theta < \infty$. An application of Proposition 3.58 shows that the Poisson kernel is well-defined. We note that $\tilde{f}$ is continuous on $\overline{D}$, by Theorem 2.4 on page 257 of [5].

We define $\tilde{f}_r(z) = \tilde{f}(rz)$ for each $0 \leq r < 1$, and we show that for each r, there exists a sequence of polynomials $p_N(z, \overline{z})$ that converges uniformly to $\tilde{f}_r$ on ∂D. Let

$$p_N(\theta) = \sum_{n=-N}^{N} \Big(\frac{1}{2\pi} \int_0^{2\pi} f(e^{\imath t}) e^{-\imath t n}\, dt \Big) r^{|n|} e^{\imath\theta n};$$

letting $z = \mathrm{e}^{\imath\theta}$ and γ as above, we see that p_N may be regarded as a polynomial in z and $\overline{z}$; explicitly,

$$\begin{aligned}
p_N(z, \overline{z}) &= \sum_{n=-N}^{N} \Big(\frac{1}{2\pi}\int_\gamma f(u)\overline{u}^n \,\mathrm{d}u\Big) r^{|n|} z^n \\
&= \frac{1}{2\pi}\int_\gamma f(u)\,\mathrm{d}u + \sum_{n=1}^{N}\Big(\frac{1}{2\pi}\int_\gamma f(u)\overline{u}^n\,\mathrm{d}u\Big) r^n z^n + \sum_{n=1}^{N}\Big(\frac{1}{2\pi}\int_\gamma f(u)\overline{u}^n\,\mathrm{d}u\Big) r^n \overline{z}^n.
\end{aligned}$$

Now since f is continuous on ∂D, f is bounded there, by M, say. Let

$$q(\theta, t) = \sum_{n=-\infty}^{\infty} r^{|n|}\mathrm{e}^{\imath(\theta - t)n},$$

and let

$$q_N(\theta, t) = \sum_{n=-N}^{N} r^{|n|}\mathrm{e}^{\imath(\theta - t)n}$$

denote the N-th symmetric partial sum of q. It follows from Proposition 3.58 that q_N converges to q absolutely and uniformly on $[0, 2\pi] \times [0, 2\pi]$. Thus for each $\varepsilon > 0$ there exists an $N(\varepsilon, r)$ depending on ε and r such that

$$\|q_N - q\|_{\text{unif}} < \frac{\varepsilon}{M}$$

for all $N \geq N(\varepsilon, r)$. Hence for any $z = \mathrm{e}^{i\theta} \in \partial D$ and any $N \geq N(\varepsilon, r)$,

$$\begin{aligned}
&|p_N(z, \overline{z}) - \tilde{f}_r(z)| \\
&= |p_N(\theta) - \tilde{f}(r\mathrm{e}^{\imath\theta})| \\
&= \Big|\sum_{n=-N}^{N}\Big[\Big(\frac{1}{2\pi}\int_0^{2\pi} f(\mathrm{e}^{\imath t})\mathrm{e}^{-\imath t n}\,\mathrm{d}t\Big) r^{|n|}\mathrm{e}^{\imath\theta n}\Big] - \frac{1}{2\pi}\int_0^{2\pi} f(\mathrm{e}^{\imath t})\sum_{n=-\infty}^{\infty} r^{|n|}\mathrm{e}^{\imath(\theta - t)n}\,\mathrm{d}t\Big| \\
&= \frac{1}{2\pi}\Big|\int_0^{2\pi} f(\mathrm{e}^{\imath t})\Big[\sum_{n=-N}^{N} r^{|n|}\mathrm{e}^{\imath(\theta - t)n} - \sum_{n=-\infty}^{\infty} r^{|n|}\mathrm{e}^{\imath(\theta - t)n}\Big]\,\mathrm{d}t\Big| \\
&\leq \frac{1}{2\pi}\int_0^{2\pi} |f(\mathrm{e}^{\imath t})||q_N(\theta, t) - q(\theta, t)|\,\mathrm{d}t \\
&\leq \varepsilon.
\end{aligned}$$

Therefore,

$$\|p_N(z, \overline{z}) - \tilde{f}_r(z)\|_{\text{unif}} \leq \varepsilon,$$

proving that $p_N(z,\overline{z})$ converges uniformly to $\tilde{f}_r(z)$ on ∂D.

Finally, for $z \in \partial D$ we have

$$\|p_N(z,\overline{z}) - f(z)\|_{\text{unif}} \leq \|p_N(z,\overline{z}) - \tilde{f}_r(z)\|_{\text{unif}} + \|\tilde{f}_r(z) - f(z)\|_{\text{unif}}\,;$$

choosing r sufficiently close to 1, we can make the second term on the right hand side as small as we like; then, choosing N—which will depend on r—sufficiently large, we can make the first term on the right hand side as small as we like. Thus p_N converges uniformly to f on ∂D. $\square$

We will need the following lemma:

Lemma 3.60 (Theorem 7.1 in [3])**.** Suppose that (X,μ) is a finite measure space, and $\{f_n\}$ is a sequence in $L^p(X,\mu)$ that converges uniformly on X to f. Then $f \in L^p(X,\mu)$ and the sequence $\{f_n\}$ converges in L^p to f.

Theorem 3.61. The functions $\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}$ form an orthonormal basis for $L^2[0,2\pi]$.

Proof. A routine calculation shows that $\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}$ is an orthonormal set, so by the Basis Theorem it's just a matter of showing that $\bigvee\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\} = L^2[0,2\pi]$. Let f be a continuous function on $[0,2\pi]$. By the Weierstrass Approximation Theorem, the function $f \circ \gamma^{-1}$ is the uniform limit of a sequence of polynomials $p_N(z,\overline{z})$. As in the proof of that theorem, by letting $z = \mathrm{e}^{\imath\theta} = \gamma(\theta)$, we may regard the p_N as polynomials in θ; then f is a uniform limit of polynomials in θ. But by definition, $p_N(\theta) \in \operatorname{span}\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}$, proving that the set of continuous functions on $[0,2\pi]$ is the uniform closure of

$$\operatorname{span}\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}.$$

It follows from Lemma 3.60 that $\operatorname{span}\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}$ is L^2-dense in $C[0,2\pi]$, so that we have

$$C[0,2\pi] \subset \bigvee\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\} \subset L^2[0,2\pi].$$

One can show—though it involves some work—that $C[0,2\pi]$ is dense in $L^2[0,2\pi]$. The conclusion follows immediately. $\square$

Remark 3.62. That $C[0,2\pi]$ is the uniform closure of $\operatorname{span}\{\mathrm{e}^{\imath n\theta} : n \in \mathbb{Z}\}$ can be proven much more concisely using the Stone-Weierstrass theorem, which is a generalization of the Weierstrass Approximation theorem.

Put another way, Theorem 3.61 shows that $\{\gamma^n : n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2[0, 2\pi]$, where γ is the curve $\gamma(\theta) = \mathrm{e}^{\imath\theta}$, as before. Theorem 3.57 now tells us that $L^2[0, 2\pi] \cong l^2(\mathbb{Z})$; in the notation of that theorem, we have

$$\hat{f}(n) := \mathbf{U}(f)(n) = \langle f, \gamma^n \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\mathrm{e}^{-\imath n\theta}\,\mathrm{d}\theta$$

for each $f \in L^2[0, 2\pi]$; we call this the *n-th Fourier coefficient of f*, and we call the series representation

$$f = \sum_{n \in \mathbb{Z}} \hat{f}(n)\gamma^n \tag{11}$$

the *Fourier series* for f. Applying the Pythagorean theorem to (11) shows that

$$\sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2 = \lim_{N \to \infty} \sum_{n=-N}^{N} |\hat{f}(n)|^2 < \infty,$$

which proves a the Riemann-Lebesgue lemma:

Corollary 3.63 (Riemann-Lebesgue)**.** If $f \in L^2[0, 2\pi]$, then

$$\int_0^{2\pi} f(\theta)\mathrm{e}^{-\imath n\theta}\,\mathrm{d}\theta \to 0$$

as $n \to \pm\infty$.

As discussed at the beginning of this section, $L^2[0, 2\pi]$ is, up to isomorphism, $L^2[\partial D]$; so the map $U : L^2[0, 2\pi] \to l^2(\mathbb{Z})$ of Theorem 3.57, which maps f to its sequence $\hat{f}$ of Fourier coefficients, induces an isometric isomorphism from $L^2[\partial D]$ onto $l^2(\mathbb{Z})$; explicitly,

$$\mathscr{F} : L^2[\partial D] \xrightarrow{\cong} l^2(\mathbb{Z})$$

is given by $f \mapsto \mathbf{U}(f \circ \gamma)$. That is,

$$\begin{aligned} \mathscr{F}(f)(n) &:= \mathbf{U}(f \circ \gamma)(n) \\ &= \langle f \circ \gamma, \gamma^n \rangle \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(\gamma(\theta))[\gamma(\theta)]^{-n}\,\mathrm{d}\theta \qquad (12) \\ &= \frac{1}{2\pi} \int_{\partial D} f(z) z^{-n}\,\mathrm{d}z, \end{aligned}$$

for each $f \in L^2[\partial D]$. The map $\mathscr{F}$ is called the *Fourier transform on the circle.* It is not the form of the Fourier transform with which students of analysis are often first made aquainted. This is because this version of the Fourier transform is usually encountered in abstract harmonic analysis, not classical real analysis. In abstract harmonic analysis one often works with *locally compact abelian groups* (LCA groups)—topological groups for which the underlying topology is locally compact and Hausdorff. Given an LCA group G, we define its dual group $\widehat{G}$ to be the group of continuous homomorphisms from G into the circle group $\mathbb{T}$; these homomorphisms are called *characters.* For any $f \in L^1(G)$, we can define the Fourier transform

$$(\mathscr{F} f)(\xi) := \int_G f(x)\overline{\xi(x)}\, \mathrm{d}\mu(x), \tag{13}$$

where μ is a measure on the group G, called the *Haar measure.* If one takes G to be $\mathbb{R}$ with Lebesgue measure, then this abstract Fourier transform is just the usual Fourier transform on $L^1(\mathbb{R})$; so this is indeed a true generalization. Now what if we take $G = \mathbb{T}$? The characters of $\mathbb{T}$ are of the form $x \mapsto x^n$ for $n \in \mathbb{Z}$, and the Haar measure on $\mathbb{T}$ is

$$\mu(S) = \frac{1}{2\pi}\big[m\big(\gamma^{-1}(S)\big)\big],$$

where m is Lebesgue measure. In light of these observations, we see that when we take $G = \mathbb{T}$, definition (13) is equivalent to (12). Thus, the Fourier transform on the circle and the usual Fourier transform on $\mathbb{R}$ are just two instances of the same object.

3.15 External Direct Sums

Given a collection $\{\mathscr{H}_i\}_{i\in I}$ of Hilbert spaces, we can consider their Cartesian product $\prod_{i\in I}\mathscr{H}_i$. Defining addition and scalar multiplication coordinatewise on $\prod_{i\in I}\mathscr{H}_i$—that is, by the formulas

$$\alpha\{x_i\} = \{\alpha x_i\}, \quad \{x_i\} + \{y_i\} = \{x_i + y_i\}$$

—makes $\prod_{i\in I}\mathscr{H}_i$ into a vector space. We define

$$\|\{x_i\}\|_2 = \left[\sum_{i\in I}\|x_i\|^2_{\mathscr{H}_i}\right]^{\frac{1}{2}}, \tag{14}$$

and we define the *external direct sum* of the $\mathscr{H}_i$'s, denoted $\bigoplus_{i\in I}\mathscr{H}_i$, to be

$$\Big\{\{x_i\} \in \prod_{i\in I}\mathscr{H}_i : \|\{x_i\}\|_2 < \infty\Big\};$$

one can verify that this is a vector subspace of $\prod_{i\in I} \mathscr{H}_i$—indeed, it's a matter of showing that (14) satisfies the triangle inequality, which is standard L^p theory. We can define an inner product on $\bigoplus_{i\in I} \mathscr{H}_i$ by

$$\Big\langle \{x_i\}, \{y_i\} \Big\rangle = \sum_{i\in I} \langle x_i, y_i \rangle_{\mathscr{H}_i}. \tag{15}$$

The sum on the right hand side of equation (15) converges absolutely, since by the Cauchy-Schwarz inequality and Hölder's inequality,

$$\begin{aligned} \sum_{i\in I} |\langle x_i, y_i \rangle_{\mathscr{H}_i}| &\le \sum_{i\in I} \|x_i\| \|y_i\| \\ &\le \|\{x_i\}\|_2 \|\{y_i\}\|_2. \end{aligned}$$

Thus (15) gives a well-defined function, and it is not hard to verify that it satisfies the criteria for being an inner product. The norm determined by this inner product is of course given by (14).

Theorem 3.64. With the inner product defined in (15), $\bigoplus_{i\in I} \mathscr{H}_i$ is a Hilbert space.

Proof. It remains only to show that $\bigoplus_{i\in I} \mathscr{H}_i$ is complete with respect to the norm determined by the inner product. Let $\{x_i^n\}$, $n = 1, 2, \ldots$ be a Cauchy sequence in $\bigoplus_{i\in I} \mathscr{H}_i$. Then it is easy to see that for each fixed i, the sequence $\{x_i^n\}_{n=1}^\infty$ is a Cauchy sequence in $\mathscr{H}_i$, which therefore converges to some $x_i \in \mathscr{H}_i$. For any finite subset $F \subset I$ and any fixed $n \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{i\in F} \|x_i^n - x_i\|^2 &= \lim_{m\to\infty} \sum_{i\in F} \|x_i^n - x_i^m\|^2 \\ &= \limsup_m \sum_{i\in F} \|x_i^n - x_i^m\|^2 \\ &\le \limsup_m \|\{x_i^n\} - \{x_i^m\}\|_2^2; \end{aligned} \tag{16}$$

the right hand side of (16) is finite because $\{x_i^m\}_{m=1}^\infty$, being a Cauchy sequence, is bounded. Taking the supremum over all finite $F \subset I$ in (16) yields

$$\|\{x_i^n\} - \{x_i\}\|_2^2 \le \limsup_m \|\{x_i^n\} - \{x_i^m\}\|_2^2, \tag{17}$$

which shows that $\{x_i^n - x_i\} \in \bigoplus_{i\in I} \mathscr{H}_i$. It follows that $\{x_i\} \in \bigoplus_{i\in I} \mathscr{H}_i$, for

$$\{x_i\} = \{x_i^n\} - \{x_i^n - x_i\} \in \bigoplus_{i\in I} \mathscr{H}_i.$$

Taking the limit superior over n in (17) yields

$$\begin{aligned}\limsup_n \|\{x_i^n\} - \{x_i\}\|_2^2 &\leq \limsup_n \Big(\limsup_m \|\{x_i^n\} - \{x_i^m\}\|_2^2 \Big) \\ &\leq \limsup_{m,n} \|\{x_i^n\} - \{x_i^m\}\|_2^2 \qquad (18) \\ &\to 0,\end{aligned}$$

since $\{x_i^n\}$ is Cauchy. Note that we've invoked Theorem 1.11 in (18). We've shown that $\{x_i^n\}$ converges to $\{x_i\}$, which completes the proof. $\square$

Recall that the internal direct sum $\sum_{i\in I} \oplus \mathscr{M}_i$ of a collection $\{\mathscr{M}_i\}_{i\in I}$ of mutually orthogonal subspaces of $\mathscr{H}$ was defined to be the subspace of *summable* elements of $\sum_{i\in I} \mathscr{M}_i$. Recall, as well, from 3.31 that $\{x_i\}$ is summable if and only if $\sum_{i\in I} \|x_i\|^2 < \infty$.

Theorem 3.65. Let $\{\mathscr{M}_i\}_{i\in I}$ be an orthogonal family of closed subspaces of $\mathscr{H}$, so that each $\mathscr{M}_i$ may be regarded as a Hilbert space with the inner product inherited from $\mathscr{H}$. If $\{\mathscr{M}_i\}_{i\in I}$ is an orthogonal family of subspaces of $\mathscr{H}$, then

$$\bigoplus_{i\in I} \mathscr{M}_i \cong \sum_{i\in I} \oplus \mathscr{M}_i. \qquad (19)$$

Proof. Assume that $\{\mathscr{M}_i\}_{i\in I}$ is an orthogonal family and let $\phi : \bigoplus_{i\in I} \mathscr{M}_i \to \sum_{i\in I} \oplus \mathscr{M}_i$ be the map

$$\{x_i\}_{i\in I} \mapsto \sum_{i\in I} x_i.$$

By the Pythagorean theorem we have

$$\|\phi(\{x_i\})\|^2 = \Big\| \sum_{i\in I} x_i \Big\|^2 = \sum_{i\in I} \|x_i\|^2 = \|\{x_i\}\|_2^2,$$

so that ϕ is an isometry. Now let $x \in \sum_{i\in I} \oplus \mathscr{M}_i$. By Theorem 3.39, $\{\mathscr{M}_i\}_{i\in I}$ being an orthogonal family means that x can be written uniquely in the form $\sum_{i\in I} x_i$, with $x_i \in \mathscr{M}_i$. Moreover, $\{x_i\}$ is summable to x by definition, and hence $\{x_i\} \in \bigoplus_{i\in I} \mathscr{M}_i$, by Theorem 3.31. Thus we have $x = \phi(\{x_i\})$, and therefore ϕ is surjective. $\square$

References

[1] https://math.stackexchange.com/questions/15240/when-can-you-switch-the-order-of-limits

[2] Tom M. Apostol. *Mathematical Analysis, 2nd ed.*. Addison-Wesley, 1974.

[3] Robert G. Bartle, *The Elements of Integration and Lebesgue Measure.* John Wiley & Sons, 1966; Wiley Classics Library Edition Published 1995.

[4] John B. Conway. *A Course in Functional Analysis, 2nd ed.*. Springer, 1990.

[5] John B. Conway. *Functions of One Complex Variable I, 2nd ed.*. Springer, 1978.

[6] Gerald B. Folland. *Real Analysis.* John Wiley & Sons, 1999.

[7] Paul R. Halmos. *Introduction to Hilbert Space, 2nd ed.*. Chelsea (1957); reprinted by the American Mathematical Society, 2000.

[8] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis. Revised English Edition Translated and Edited by Richard A. Silverman.* Dover, 1975.

[9] Lynn H. Loomis and Shlomo Sternberg. *Advanced Calculus.* Addison-Wesley, 1968; reprinted 1980.

[10] David C. Ullrich. *Complex Made Simple.* AMS Graduate Studies in Mathematics volume 97, 2008.