7 Hilbert Spaces

In MAT203A we have considered different kinds of topological spaces. In increasing order of specialization we studied arbitrary toplogical spaces, metric spaces, normed linear spaces, and Banach spaces. Hilbert spaces are Banach spaces with a norm that is derived from an inner product. Thus, again, they have an extra feature as compared to arbitrary Banach spaces, which makes them more special. On the other hand, the extra structure will lead to new interesting properties.

The material of this sections elementary introduction to Hilbert space is covered in Chapter 4 of
W. Rudin, *Real and Complex Analysis*, 3rd ed, McGraw-Hill, New York 1986.

P.R. Halmos wrote a well-known study book on Hilbert space, which consists of 250 problems with hints and solutions:
P.R. Halmos, *A Hilbert Space Problem Book*, 2nd ed, Springer Verlag, New York 1982.

The new feature of Hilbert spaces is the *inner product*, which, in Section 8, we will use to introduce the notion of orthonormal basis of a Hilbert space.

Hilbert spaces are Banach spaces with a norm that is derived from an inner product.

7.1 Inner Products

Definition 7.1 (Inner Product) *An inner product on a linear space X over $\mathbb{C}$ (or $\mathbb{R}$) is a map $(\cdot\,,\cdot):X\times X\to\mathbb{C}$ (or $\mathbb{R}$), with the following properties: for all $x,y,z\in X$ and $\lambda\in\mathbb{C}$ (or $\mathbb{R}$) one has*
i) $(x,y+\lambda z)=(x,y)+\lambda(x,z)$ (linear in the second argument)
ii) $(y,x)=\overline{(x,y)}$, where $\overline{\cdot}$ denotes complex conjugation ("hermitian symmetric")
iii) $(x,x)\geq 0$ (non-negative, or positive semi-definite)
iv) $(x,x)=0\Rightarrow x=0$ (positive definite).

A linear space with an inner product is often referred to as an *inner product space.*

From i) and ii) it follows that $(\cdot\,,\cdot)$ is *antilinear* in the first argument: $(y+\lambda z,x)=(y,x)+\overline{\lambda}(z,x)$, for all $x,y,z\in X$ and $\lambda\in\mathbb{C}$ (or $\mathbb{R}$). In the case of complex X, this means that $(\cdot\,,\cdot)$ is not bilinear. It is called *sesquilinear*, which literally means $1\frac{1}{2}$ linear.

Note that there are two conventions for the linearity of the inner product in the literature: in most of the mathematics oriented literature $(\cdot\,,\cdot)$ is linear in the first argument. We will adopt the convention that the inner product is linear in the second argument, which is the most common in applied mathematics and physics oriented works.

The following very useful inequality for inner products on a real or complex linear spaces is a consequence of the non-negative definiteness:

$$|(x,y)|^2 \leq (x,x)(y,y), \quad \text{for all } x, y \in X \tag{7.1}$$

It is usually called the *Cauchy-Schwarz inequality* .

Exercise 7.1 *Prove the Cauchy-Schwarz inequality.*

For simplicity of notation we will henceforth assume that X is a linear space over $\mathbb{C}$. Many statements carry over to the case of real linear spaces.

Let X be a linear space equipped with an inner product $(\cdot, \cdot)$. We can define a norm on X by

$$\|x\| = \sqrt{(x,x)}$$

Check that this indeed defines a norm.

Examples:

1. The standard inner product on $\mathbb{C}^n$ is given by

$$(x,y) = \sum_{i=1}^{n} \overline{x_i} y_i, \quad \text{for all } x, y \in \mathbb{C}^n$$

2. Let $C([a,b],\mathbb{C})$ denote the space of all complex-valued continuous functions defined on the interval $[a,b]$. An inner product on this space can be defined by

$$(f,g) = \int_a^b \overline{f(x)} g(x)\, dx, \quad \text{for all } f, g \in C([a,b],\mathbb{C})$$

By continuity this can be extended to $L_2([a,b],\mathbb{C})$, defined as the completion of $C([a,b],\mathbb{C})$ for the norm defined by $\|f\| = (f,f)^{1/2}$.

3. Define l_2 by

$$l_2 = \{(z_n)_{n=1}^{\infty} \mid \sum_{n=1}^{\infty} |z_n|^2 < \infty\}$$

l_2 is a complex linear space with the obvious operations of addition and multiplication by a scalar, and a natural inner product on it is given by

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

4. Let M_n denote the space of all $n \times n$ matrices with complex entries. One can define a inner product on M_n by

$$(A, B) = \mathrm{Tr} A^* B$$

where Tr denotes the trace and * is hermitian conjugation of matrices. Check that this is the natural inner product one obtains by identifying M_n with $\mathbb{C}^{n^2}$.

Theorem 7.2 *A normed linear space X is an inner product space with the norm derived from the inner product iff*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad \textit{for all } x, y \in X \quad . \tag{7.2}$$

When this condition is satisfied, the inner product is necessarily given by the formula

$$(x, y) = \frac{1}{4}\left\{\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right\}, \quad \textit{for all } x, y \in X \quad . \tag{7.3}$$

Proof: The proof of this theorem is homework. ∎

The relation (7.2) has a geometric content and is called the *Parallelogram Law.* It says that, if one considers the parallelogram spanned by the vectors x and y in the two dimensional subspace ($\equiv$ plane) generated by x and y, the sum of the squares of two adjacent sides of the parallelogram is equal to half the sum of the squares of the diagonals. Equation (7.3) is called the polarization formula. A useful consequence of this formula is that a sequilinear form is uniquely determined by its values on the diagonal, i.e., its values for equal first and second arguments.

Let $(X, (\cdot\,,\,\cdot)_X)$ and $(Y, (\cdot\,,\,\cdot)_Y)$ be two inner product spaces. Then $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$ has a natural inner product defined by

$$((x_1, y_1), (x_2, y_2))_{X \times Y} = (x_1, x_2)_X + (y_1, y_2)_Y$$

which in turn leads to a natural norm on $X \times Y$:

$$\|(x, y)\| = \sqrt{\|x\|^2 + \|y\|^2}$$

Unless otherwise stated, this will be the norm considered on a Cartesian product of two inner product spaces.

Theorem 7.3 *Let X be an inner product space, considered as a normed space with the norm the norm derived from the inner product. Then, the inner product is continuous as a map from $X \times X \to \mathbb{C}$.*

Proof: For all $x, y \in X$ one has the bound

$$|(x, y)| \leq \|x\| \, \|y\|$$

which is nothing else than the Cauchy-Schwartz inequality. Just as for linear maps, continuity follows from this kind of boundedness:

$$|(x_1, y_1) - (x_2, y_2)| = |(x_1 - x_2, y_1) + (x_2, y_1 - y_2)| \leq \|x_1 - x_2\| \, \|y_1\| + \|x_2\| \, \|y_1 - y_2\|$$

■

7.2 Hilbert Spaces

Definition 7.4 *A Hilbert space is an inner product space X that is complete for the norm derived from the inner product.*

In particular, a Hilbert space is a Banach space.

Examples:
Except for $C([a, b], \mathbb{C})$, the examples of inner product spaces given in the previous section are all complete and, therefore, are Hilbert spaces.

Exercise 7.2 *Let $\mathcal{H}_0$ be a linear subspace of a Hilbert spaec $\mathcal{H}$. Prove that $\mathcal{H}_0$, equipped with the same inner product as $\mathcal{H}$, is a Hilbert space iff $\mathcal{H}_0$ is a closed subspace of $\mathcal{H}$.*

The inner product structure of a Hilbert space allows us to introduce the concept of orthogonality of vectors. Let $\mathcal{H}$ be a Hilbert space. We say that x and $y \in \mathcal{H}$ are *orthogonal* iff $(x, y) = 0$. Shorthand for "x and y are orthogonal" is $x \perp y$. For two linear subspaces $\mathcal{H}_0$ and $\mathcal{H}_1$ of $\mathcal{H}$, we say that $\mathcal{H}_0$ and $\mathcal{H}_1$ are orthogonal if $x \perp y$ for all $x \in \mathcal{H}_0$ and $y \in \mathcal{H}_1$, in which case one writes $\mathcal{H}_0 \perp \mathcal{H}_1$. Similarly, $x \perp \mathcal{H}_0$ means $x \perp y$, for all $y \in \mathcal{H}_0$. The *orthogonal complement* $\mathcal{H}_0^\perp$ of a linear subspace $\mathcal{H}_0$ of $\mathcal{H}$ is defined by

$$\mathcal{H}_0^\perp = \{x \in \mathcal{H} \mid x \perp y, \text{for all } y \in \mathcal{H}_0\}$$

The notion of orthogonality makes it possible to think about vectors and linear subspaces of a Hilbert space in a geometric way. This is sometimes very helpful, much in the same way as our geometric intuition for the Euclidean spaces $\mathbb{R}^n$ helps us to understand linear transformations of $\mathbb{R}^n$. It should be mentioned that the geometry of Banach spaces can be surprisingly complicated and quite different from what naive extrapolations of the finite-dimensional situation would suggest. The following two results, and their applications rely very much on the notion of orthogonality.

Theorem 7.5 *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}_0$ a subset of $\mathcal{H}$. Then, $\mathcal{H}_0^\perp$ is a closed linear subspace of $\mathcal{H}$.*

Proof: Let us first prove that $\mathcal{H}_0^\perp$ is indeed a linear subspace. For this we must show that for all $x, y \in \mathcal{H}_0^\perp$, and $\lambda \in \mathbb{C}$, $x + \lambda y \in \mathcal{H}_0^\perp$. But this follows immediately form the linearity of the inner product:

$$(x_0, x + \lambda y) = (x_0, x) + \lambda(x_0, y) = 0 + 0, \quad \text{for all } x_0 \in \mathcal{H}_0$$

Next, we show that $\mathcal{H}_0^\perp$ is closed. A Hilbert space is a metric space, and therefore, it suffices to show that for any convergent sequence of elements $x_n \in \mathcal{H}_0^\perp$, the limit x also belongs to $\mathcal{H}_0^\perp$. Let $x_0 \in \mathcal{H}_0$. From Theorem 7.3 we have that the inner product is continuous and therefore, we can compute (x_0, x) as follows:

$$(x_0, x) = (x_0, \lim_n x_n) = \lim_n (x_0, x_n) = \lim_n 0 = 0$$

Hence, $x \in \mathcal{H}_0^\perp$. ■

Exercise 7.3 *Prove that for any set $A \subset \mathcal{H}$, one has*

$$(\overline{A})^\perp = A^\perp$$

The following theorem expresses a deep property of Hilbert spaces. Beware that there is no analogous theorem for arbitrary Banach spaces.

Theorem 7.6 *Let $\mathcal{H}_0$ a closed linear subspace of a Hilbert space $\mathcal{H}$.*
i) For each $x \in \mathcal{H}$ there is a unique $x_0 \in \mathcal{H}_0$ such that

$$\|x - x_0\| = \min_{y \in \mathcal{H}_0} \|x - y\| \tag{7.4}$$

ii) The point $x_0 \in \mathcal{H}_0$, closest to $x \in \mathcal{H}$, is the unique element of $\mathcal{H}_0$ with the property that $x - x_0 \perp \mathcal{H}_0$.

Proof:
Step 1. First, we prove that there exists $x_0 \in \mathcal{H}_0$ such that

$$\|x - x_0\| = d := \inf\{\|x - y\| \mid y \in \mathcal{H}_0\} \tag{7.5}$$

From the definition of d it follows that there is a sequence of elements $y_n \in \mathcal{H}_0$, such that $\lim_n \|x - y_n\| = d$. So, for all $\epsilon > 0$, there exists N such that for all $n \geq N$, $\|x - y_n\| \leq d + \epsilon$. Let us show that the sequence (y_n) is Cauchy. By the Parallelogram Law we have

$$\|y_m - y_n\|^2 + \|2x - y_m - y_n\|^2 = 2\|x - y_m\|^2 + 2\|x - y_n\|^2$$

As $(y_m + y_n)/2 \in \mathcal{H}_0$, we have

$$\|x - (y_m + y_n)/2\| \geq d$$

By combining these relations we obtain:

$$\begin{aligned} \|y_m - y_n\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - \|2x - y_m - y_n\|^2 \\ &\leq 4(d+\epsilon)^2 - 4d^2 \\ &\leq 8\epsilon(d+\epsilon) \end{aligned}$$

for all $m, n \geq N$. This proves that (y_n) is Cauchy. As a Hilbert space is complete by definition, there is y such that $y_n \to y$, and as $\mathcal{H}_0$ is closed by assumption, we have that $y \in \mathcal{H}_0$, and we can take $x_0 = y$ by continuity of the distance.
Step 2. Next, we prove the uniqueness of the vector $x_0 \in \mathcal{H}_0$ minimizing $\|x - x_0\|$. Suppose x_0 and x_0' both minimize the distance to x, i.e., $\|x - x_0\| = \|x - x_0'\| = d$. Again, by the Parallelogram Law, we have

$$2\|x - x_0\|^2 + 2\|x - x_0'\|^2 = \|2x - x_0 - x_0'\|^2 + \|x_0 - x_0'\|^2$$

and hence, as $x_0 + x_0' \in \mathcal{H}_0$,

$$\|x_0 - x_0'\|^2 = 4d^2 - \|2x - x_0 - x_0'\|^2 \leq 0$$

Therefore, $\|x_0 - x_0'\| = 0$, and we must conclude that $x_0 = x_0'$. So there can be only one $x_0 \in \mathcal{H}_0$ such that $\|x - x_0\| = d$.
Step 3. The unique x_0 found in Steps 1 and 2, must satisfy $(z, x - x_0) = 0$, for all $z \in \mathcal{H}_0$. This follows from the inequality

$$\|x - x_0\|^2 \leq \|x - (x_0 - \lambda z)\|$$

for any $z \in H_0, \lambda \in \mathbb{C}$. Indeed, by expanding the RHS and using $\|x - x_0\|^2 = d^2$, we obtain

$$d^2 \leq d^2 + |\lambda|^2\|z\|^2 - 2\mathrm{Re}\,\lambda(x - x_0, z)$$

which simplifies to

$$|\lambda|^2\|z\|^2 \geq 2\mathrm{Re}\,\lambda(x - x_0, z)$$

Take λ of the form $\epsilon e^{i\varphi}$, with $\epsilon > 0$ and $\varphi \in \mathbb{R}$, divide by ϵ, and take the limit $\epsilon \downarrow 0$, and you obtain

$$2\mathrm{Re}\,e^{i\varphi}(x - x_0, z) \leq 0, \quad \text{for all } \varphi \in \mathbb{R}$$

It follows that $(x - x_0, z) = 0$ (consider this inequality for $\varphi = 0, \pi/2, \pi$, and $3\pi/2$).
Step 4. Finally, we show that x_0 is the only element in $\mathcal{H}_0$ such that $x - x_0 \perp \mathcal{H}_0$. Suppose x_0' is another such element in $\mathcal{H}_0$. Then $x_0 - x_0' \in \mathcal{H}_0$, and, for any $y \in \mathcal{H}_0$, we have

$$(y, x_0 - x_0') = (y, x - x_0') - (y, x - x_0) = 0 - 0 = 0$$

In particular, we can take $y = x_0 - x_0'$, and therefore we must have $x_0 = x_0'$. ■

Note that for the proof of Theorem 7.6, part i), it is sufficient that $\mathcal{H}_0$ is a closed convex subset of $\mathcal{H}$.

Exercise 7.4 *For any linear subspace $\mathcal{H}_0 \subset \mathcal{H}$, prove that*
i) $\mathcal{H}_0^{\perp\perp} := (\mathcal{H}_0^{\perp})^{\perp} = \overline{\mathcal{H}_0}$
ii) $\mathcal{H}_0^{\perp\perp\perp} = \mathcal{H}_0^{\perp}$.
iii) How do the relation i) and ii) have to be changed if $\mathcal{H}_0$ is just a subset and not a linear subspace?

Exercise 7.5 *Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be two linear subspaces of a Hilbert space $\mathcal{H}$, such that $\mathcal{H}_0 \perp \mathcal{H}_1$.*
i) Prove that $\mathcal{H}_0^{\perp\perp} \perp \mathcal{H}_1^{\perp\perp}$.
ii) Is it true, in general, that $\mathcal{H}_0^{\perp\perp\perp} \perp \mathcal{H}_1^{\perp\perp\perp}$?

8 Orthonormal Bases

Let $\mathcal{H}$ be a Hilbert space ans denote its inner product by $\langle\cdot,\cdot\rangle$. This is another common notation for inner products, which is generally reserved for Hilbert spaces.

A set $U \subset \mathcal{H}$, $0 \notin U$, is called *orthogonal* if any pair of distinct elements (vectors) $u, v \in U$ are mutually orthogonal, i.e., $\langle u, v\rangle = 0$. A set of vectors U is called *orthonormal* if it is orthogonal and $\|u\| = 1$ for all $u \in U$, in which case the vectors u are said to be *normalized.*

Exercise 8.1 *Prove that if U is an orthogonal set, U is a set of linearly independent vectors.*

For the finite-dimensional Hilbert spaces $\mathbb{C}^n$ we have the notion of *orthonormal basis*, often abbreviated as o.n. basis. A set of vectors $\{e_1, \ldots, e_n\} \subset \mathbb{C}^n$ is an o.n. basis if it has the following properties:
i) $\langle e_i, e_j\rangle = \delta_{ij}$ for $1 \leq i, j \leq n$, and
ii) for all $x \in \mathbb{C}^n$ there are unique coordinates $x_i \in \mathbb{C}$ such that

$$x = \sum_{i=1}^{n} x_i e_i \tag{8.1}$$

Our goal is now to generalize the notion of o.n. basis to arbitrary Hilbert spaces. Of course, we expect that, in general, an o.n. basis will be an infinite set of vectors. In the case of *separable* Hilbert spaces (i.e., containing a countable dense set), we will find that an o.n. basis is at most countably infinite. In this case the sum in (8.1) simply becomes an infinite series in a Banach space, for which we have a natural notion of convergence. But how do we interpret a sum with an uncountable number of terms? We address this question first.

8.1 Unordered infinite sums

Let I be a (not necessarily countable) index set, and consider a set of elements of linear topological space X indexed by $I : \{x_\alpha \mid \alpha \in I\} \subset X$. We would like to give sense to

$$\sum_{\alpha\in I} x_\alpha \tag{8.2}$$

For every finite subset $J \subset I$ define a partial sum S_J by

$$S_J = \sum_{\alpha\in J} x_\alpha$$

We say that $\sum_{\alpha\in I} x_\alpha$ converges to $x \in X$ if for any neighborhood V of x there is a finite set J_0 such that for all finite sets J containing J_0 one has $S_J \in V$. This is a special case of convergence of nets. Note that (8.2) is an 'unordered sum', and that it converges iff any permutation of its terms converges. The following exercise shows that, in the case of normed linear spaces, this notion of convergence is not very different from the usual notion of (absolutely) convergent series with a countable number of terms.

Exercise 8.2 *Prove the following:*
i) If X is a normed linear space and $\sum_{\alpha\in I} x_\alpha \to x$, then there exists a countable subset $\{\alpha_1, \alpha_2, \ldots\} \subset I$ such that the series $\sum_{i=1}^{\infty} x_{\alpha_i}$ converges to x.
ii) Under the same conditions as in i) at most a countable number of the x_α are non-zero.

This exercise also shows that in the case of $X = \mathbb{C}$ (or $\mathbb{R}$), $\sum_{\alpha\in I} x_\alpha \to x$ iff $\sum_{\alpha\in I} |x_\alpha| < \infty$, for absolute convergence and unconditional convergence coincide for numerical series.

Exercise 8.3 *Let $\{x_\alpha\}_{\alpha\in I}$ be a family of non-negative real numbers. Prove that*

$$\sum_{\alpha\in I} x_\alpha = \sup\{\sum_{\alpha\in J} x_\alpha \mid J \subset I, J \text{ finite } \}$$

It is quite obvious how to generalize the Cauchy criterion for convergence of unordered sum in a Banach space (such as, e.g., a Hilbert space).

Proposition 8.1 *If X is a Banach space we have that $\sum_{\alpha\in I} x_\alpha$ converges iff for every $\epsilon > 0$ there exists a finite set $J_0 \subset I$ such that $\|S_J\| < \epsilon$ for every $J \subset I \setminus J_0$.*

8.2 Bessel's inequality etc.

In this section the sums $\sum_{\alpha\in I}$ have to be interpreted as unordered infinite sums in the sense of above. Although in the future we will mostly consider separable Hilbert spaces (and countable sums) it is worth postponing this restriction for another while. This will take the mystery out of non-separable Hilbert spaces.

Lemma 8.2 *Let $\mathcal{H}$ be a Hilberts space and $U = \{u_\alpha \mid \alpha \in I\} \subset \mathcal{H}$ an orthogonal set. Then, the sum $\sum_{\alpha\in I} u_\alpha$ converges iff $\sum_{\alpha\in I} \|u_\alpha\|^2 < \infty$. In this case one has*

$$\|\sum_{\alpha\in I} u_\alpha\|^2 = \sum_{\alpha\in I} \|u_\alpha\|^2 \tag{8.3}$$

If U is orthonormal one has

$$\|\sum_{\alpha\in I} c_\alpha u_\alpha\|^2 = \sum_{\alpha\in I} |c_\alpha|^2 \tag{8.4}$$

for arbitrary $c_\alpha \in \mathbb{C}$ for which the RHS is finite.

Proof: For any finite set J we have

$$\|\sum_{\alpha\in J} u_\alpha\|^2 = \sum_{\alpha,\beta\in J}\langle u_\alpha, u_\beta\rangle = \sum_{\alpha\in J}\langle u_\alpha, u_\alpha\rangle = \sum_{\alpha\in J}\|u_\alpha\|^2$$

It follows that the Cauchy criterion is fulfilled for $\sum_{\alpha\in I} u_\alpha$ iff it is for $\sum_{\alpha\in I}\|u_\alpha\|^2$, and, therefore, each of the sums converges iff the other does. (8.3) follows because the norm is a continuous function.

(8.4) is a simple application of (8.3): If U is orthonormal, the set $\{c_\alpha u_\alpha\}$ is orthogonal, and we have

$$\sum_{\alpha\in I}\|c_\alpha u_\alpha\|^2 = \sum_{\alpha\in I}|c_\alpha|^2$$

■

Theorem 8.3 *Let $U = \{u_\alpha \mid \alpha\in I\}\subset\mathcal{H}$ be an orthonormal set and $x\in\mathcal{H}$. Then*
i) $\sum_{\alpha\in I}|\langle u_\alpha, x\rangle|^2 \le \|x\|^2$,
ii) $x_U := \sum_{\alpha\in I}\langle u_\alpha, x\rangle u_\alpha$ is a convergent sum, and
iii) $x - x_U \in U^\perp$

Proof: We begin by computing $\|x - \sum_{\alpha\in J}\langle u_\alpha, x\rangle u_\alpha\|$ for any finite subset $J\subset I$:

$$\begin{aligned}
\|x - \sum_{\alpha\in J}\langle u_\alpha, x\rangle u_\alpha\|^2 &= \left\langle x - \sum_{\alpha\in J}\langle u_\alpha, x\rangle u_\alpha, x - \sum_{\beta\in J}\langle u_\beta, x\rangle u_\beta\right\rangle \\
&= \langle x, x\rangle - \sum_{\beta\in J}\langle u_\beta, x\rangle\langle x, u_\beta\rangle - \sum_{\alpha\in J}\overline{\langle u_\alpha, x\rangle}\langle u_\alpha, x\rangle \\
&\quad + \sum_{\alpha,\beta\in J}\overline{\langle u_\alpha, x\rangle}\langle u_\beta, x\rangle\langle u_\alpha, u_\beta\rangle \\
&= \|x\|^2 - \sum_{\alpha\in J}|\langle u_\alpha, x\rangle|^2
\end{aligned}$$

Hence

$$\sum_{\alpha\in J}|\langle u_\alpha, x\rangle|^2 = \|x\|^2 - \|x - \sum_{\alpha\in J}\langle u_\alpha, x\rangle u_\alpha\|^2 \le \|x\|^2$$

and this proves i) by 'taking the limit $J\uparrow I$'.

The convergence claimed in ii) now follows by application of Lemma 8.2.

In order to prove iii) consider any $u_{\alpha_0}\in U$ and relying on the orthonormality of U and the continuity of the inner product we compute:

$$\begin{aligned}
\langle x - \sum_{\alpha\in I}\langle u_\alpha, x\rangle u_\alpha, u_{\alpha_0}\rangle &= \langle x, u_{\alpha_0}\rangle - \sum_{\alpha\in I}\overline{\langle u_\alpha, x\rangle}\langle u_\alpha, u_{\alpha_0}\rangle \\
&= \langle x, u_{\alpha_0}\rangle - \langle x, u_{\alpha_0}\rangle = 0
\end{aligned}$$

Hence, $x - \sum_{\alpha \in I} \langle u_\alpha, x \rangle u_\alpha \in U^\perp$. ∎

The inequality i) is called *Bessel's inequality.*

Lemma 8.4 *For any orthonormal set $U = \{u_\alpha \mid \alpha \in I\}$ in $\mathcal{H}$, the set of vectors of the form*

$$\sum_{\alpha \in I} c_\alpha u_\alpha$$

with $c_\alpha \in \mathbb{C}$, such that the sum converges is a closed linear subspace of $\mathcal{H}$.

Exercise 8.4 *Prove Lemma 8.4.*

Let $[U]$ denote the smallest closed linear subspace of $\mathcal{H}$ that contains U, which is described in Lemma 8.4. By Lemma 8.2, we have

$$[U] = \{\sum_{\alpha \in I} c_\alpha u_\alpha \mid c_\alpha \in \mathbb{C} \text{ such that } \sum_{\alpha \in I} |c_\alpha|^2 < \infty\}$$

By combining Theorem 7.6 and Theorem 8.3 we see that x_U, defined in Theorem 8.3 ii), is the unique element of $[U]$ satisfying

$$\|x - x_U\| = \min_{u \in [U]} \|x - u\|$$

In particular, it follows that $x_U = x$ if $[U] = \mathcal{H}$. In this case U is called a *orthonormal (o.n.) basis* of $\mathcal{H}$.

Here are some equivalent conditions for this property, often referred to as the *completeness* of the set U.

Theorem 8.5 *If $U = \{u_\alpha \mid \alpha \in I\} \subset \mathcal{H}$ is orthonormal, the following properties are equivalent:*
i) $\langle u_\alpha, x \rangle = 0$ for all $\alpha \in I$ implies $x = 0$, i.e., $U^\perp = \{0\}$.
ii) $x = \sum_{\alpha \in I} \langle u_\alpha, x \rangle u_\alpha$, for all $x \in \mathcal{H}$.
iii) $\|x\|^2 = \sum_{\alpha \in I} |\langle u_\alpha, x \rangle|^2$, for all $x \in \mathcal{H}$.
iv) $[U] = \mathcal{H}$.
v) U is maximal in the sense that for any $0 \neq v \in \mathcal{H} \setminus U$, the set $U \cup \{v\}$ is no longer orthogonal.

Proof: We prove the string of implications *i)* $\Rightarrow$ *ii)* $\Rightarrow$ *iii)* $\Rightarrow$ *iv)* $\Rightarrow$ *v)* $\Rightarrow$ *i)*.
i) $\Rightarrow$ *ii)*: By Theorem 8.3 iii), $U^\perp = \{0\}$ implies ii).
ii) $\Rightarrow$ *iii)*: This is Lemma 8.2 (8.4).
iii) $\Rightarrow$ *iv)*: iii) implies that $U^\perp = 0$ which is equivalent to iv).
iv) $\Rightarrow$ *v)*: iv) implies that any v is of the form $v = \sum_{\alpha \in I} c_\alpha u_\alpha$. Therefore, $\langle u_\alpha, v \rangle =$

c_α, and hence $v \perp U$ implies $c_\alpha = 0$ for all α, and hence $v = 0$.
v) ⇒ i): i) is a reformulation of v). ■

i) is often the easiest condition to verify. ii) is the property that is most often used explicitly. iii) is called *Parseval's identity*, iv) simply expresses completeness of the basis, and v) will be used in the proof of existence of an o.n. basis in a arbitrary Hilbert space (Theorem 8.6).

Our next result is that every Hilbert space has a o.n. basis. It also says that any orthonormal set can be extended to an orthonormal basis.

Theorem 8.6 *Every Hilbert space $\mathcal{H}$ has a o.n. basis, and if U is an orthonormal set, then $\mathcal{H}$ has an o.n. basis containing U.*

Proof: The proof is a standard application of Zorn's Lemma (see, e.g., W. Rudin, *Real and Complex Analysis*)). If $\mathcal{H} = \{0\}$, the statement is trivially true with $U = \emptyset$. Suppose $\mathcal{H}_0 \neq \{0\}$. Then, the family of all orthonormal sets of $\mathcal{H}$ is non-empty and has a partial order given by the inclusion of sets. If $\{U_\alpha\}$ is a totally ordered family of orthonormal sets (i.e., a chain for the order relation), then $\bigcup_\alpha U_\alpha$ is an orthonormal set and is an upper bound (in the sense of inclusion) for the family $\{U_\alpha\}$. Zorn's Lemma then implies that that the family of all orthonormal sets in $\mathcal{H}$ has a maximal element, which then satisfies the equivalent conditions of Theorem 8.5.

To prove that any orthonormal set U can be extended to a o.n. basis of $\mathcal{H}$, repeat the above argument starting with the family of all orthonormal sets containing U. ■

Remark: Do not confuse the notion of o.n. basis with that of Hamel basis of a linear space. U is a Hamel basis if for all $x \in \mathcal{H}$, x can be uniquely written as *finite* linear combination of elements in U.

8.3 The Gram-Schmidt orthogonalization procedure

From the previous Theorem we know that every Hilbert space has an o.n. basis (and, hence, many o.n. bases). This would not be very useful in practical applications if we would not have a means of constructing one. The Gram-Schmidt procedure is a recipe for constructing an o.n. basis starting from a countable linearly independent set with the property that its linear span is dense in $\mathcal{H}$.

Let V be a countable set of linearly independent vectors in a Hilbert space $\mathcal{H}$. The Gram-Schmidt orthogonalization procedure is a method to construct an o.n. set U such that $[U] = [V]$, where for any subset $A \subset \mathcal{H}$, $[A]$ denotes the set of all convergent sums of the form

$$\sum_{v \in A} c_v v$$

with $c_v \in \mathbb{C}$.

Denote the elements of V by v_n. The o.n. set $U = \{u_n\}$ is then constructed inductively by putting $u_1 = v_1/\|v_1\|$ and

$$u_{n+1} = C\left(v_{n+1} - \sum_{k=1}^{n} \langle u_k, v_{n+1}\rangle u_k\right)$$

for all $n \geq 1$, and where C is chosen such that $\|u_{n+1}\| = 1$.

It is straightforward to check that $[\{v_1, \ldots, v_n\}] = [\{u_1, \ldots, u_n\}]$ for all $n \geq 1$, and hence

$$[V] = \overline{\bigcup_n [\{v_1, \ldots, v_n\}]} = \overline{\bigcup_n [\{u_1, \ldots, u_n\}]} = [U]$$

Exercise 8.5 *As a metric space a Hilbert space is called separable if it contains a countable dense subset. Prove that a Hilbert space is separable iff it has a finite or countably infinite o.n. basis.*

Exercise 8.6 *Prove that if H_0 is a dense linear subspace of a Hilbert space $\mathcal{H}$, $\mathcal{H}$ has a o.n. basis consisting of elements in $\mathcal{H}_0$. Does the same hold for arbitrary dense subsets of $\mathcal{H}$?*

Consider $\mathcal{H} = L^2([-1,1])$, and let it be given that the set $V = \{1, x, x^2, \ldots\}$ of all monic monomials is linearly independent and one has $[V] = L^2([-1,1])$. In the following excercise you are asked to compute the first few elements of a particular o.n. basis of $[V] = L^2([-1,1])$, called the Legendre polynomials.

Exercise 8.7 *Calculate the first 3 (or more) elements of the o.n. basis obtained from V, with the given ordering, by applying the Gram-Schmidt procedure.*

This example can be generalized as follows. Let $[a,b] \subset \mathbb{R}$ be a finite or infinite interval and $w : [a,b] \to \mathbb{R}$ is a continuous function, $w(x) > 0$ except for at most finitely many x, and such that

$$\int_a^b w(x)\, dx < +\infty \quad .$$

For any such weight function w, one can define an inner product on $C_b([a,b]) = \{f : [a,b] \to \mathbb{C} \mid f \text{ continuous and } |f| \text{ bounded}\}$, by

$$(f,g) = \int_a^b \overline{f(x)} g(x) w(x)\, dx, \qquad \text{for all } f, g \in C_b([a,b]) \quad .$$

Let $L^2([a,b], w)$ be the Hilbert space obtained by completion of $C_b([a,b])$ for the metric derived from this inner product. One can then use the Gram-Schmidt procedure to obtain an o.n. basis of polynomials for this Hilbert space.