(*August 30, 2005*)

Hilbert spaces

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

Hilbert spaces are possibly-infinite-dimensional analogues of the finite-dimensional Euclidean spaces familiar to us. In particular, Hilbert spaces have *inner products*, so notions of *perpendicularity* (or *orthogonality*), and *orthogonal projection* are available. Reasonably enough, in the infinite-dimensional case we must be careful not to extrapolate too far based only on the finite-dimensional case.

Perhaps strangely, few naturally-occurring spaces of functions are Hilbert spaces. Given the intuitive geometry of Hilbert spaces, this fact is a little disappointing, as it suggests that our physical intuition is a little distant from the behavior of spaces of functions, for example. [1]

- Pre-Hilbert spaces: definition
- Cauchy-Schwarz-Bunyakowski inequality
- Example: spaces ℓ^2
- Triangle inequality, associated metric, continuity issues
- Hilbert spaces, completions, infinite sums
- Minimum principle
- Orthogonal projections to closed subspaces
- Orthogonal complements $W^\perp$
- Riesz-Fischer theorem on linear functionals
- Orthonormal sets, separability
- Parseval equality, Bessel inequality
- Riemann-Lebesgue lemma
- The Gram-Schmidt process

---

1. Pre-Hilbert spaces: definition

Let V be a complex vector space. A complex-valued function

$$\langle,\rangle : V \times V \to \mathbf{C}$$

of two variables on V is a **(hermitian) inner product** if

$$\begin{cases} \langle x,y\rangle & = & \overline{\langle y,x\rangle} & \text{(the } \textit{hermitian-symmetric} \text{ property)} \\ \langle x+x',y\rangle & = & \langle x,y\rangle + \langle x',y\rangle & (\textit{additivity in} \text{ first argument}) \\ \langle x,y+y'\rangle & = & \langle x,y\rangle + \langle x,y'\rangle & (\textit{additivity} \text{ in second argument}) \\ \langle x,x\rangle & \geq & 0 & \text{(and equality only for } x=0\text{: } \textit{positivity}) \\ \langle \alpha x,y\rangle & = & \alpha\langle x,y\rangle & (\textit{linearity} \text{ in first argument}) \\ \langle x,\alpha y\rangle & = & \bar{\alpha}\langle x,y\rangle & (\textit{conjugate-linearity} \text{ in second argument}) \end{cases}$$

Then V equipped with such a $\langle,\rangle$ is a **pre-Hilbert space**. Among other easy consequences of these requirements, for all $x,y \in V$

$$\langle x,0\rangle = \langle 0,y\rangle = 0$$

where inside the angle-brackets the 0 is the zero-*vector*, and outside it is the zero-*scalar*.

---

[1] However, a little later we will see that suitable *families* of Hilbert spaces *may* capture what we want. Such ideas originate with Sobolev in the 1930's. Sobolev's ideas were not widely known in the West when Schwartz formulated his notions of *distributions*, so were not directly incorporated. Certainly Sobolev's ideas fit into Schwartz' general scheme, but they do also offer some useful specifics, as we will see.

The **associated norm** $|\ |$ on V is defined by

$$|x| = \langle x, x\rangle^{1/2}$$

with the non-negative square-root. Even though we use the same notation for the norm on V as for the usual complex value $||$, context should always make clear which is meant.

Sometimes such spaces V with $\langle,\rangle$ are called *inner product spaces* or *hermitian inner product spaces.*

For two vectors v, w in a pre-Hilbert space, if $\langle v, w\rangle = 0$ then v, w are **orthogonal** or **perpendicular**, sometimes written $v \perp w$. A vector v is a **unit vector** if $|v| = 1$.

There are several essential algebraic identities, variously and ambiguously called **polarization identities**. First, there is

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

which is obtained simply by expanding the left-hand side and cancelling where opposite signs appear. In a similar vein,

$$|x + y|^2 - |x - y|^2 = \langle x, y\rangle + \langle y, x\rangle$$

Therefore,

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

These and closely-related identites are of frequent use.

---

2. Cauchy-Schwarz-Bunyakowski inequality

This inequality is fundamental. It is necessary to prove that the triangle inequality holds for the norm, from which we get the associated *metric*, as indicated below.

The **Cauchy-Schwarz-Bunyakowsky inequality** in a pre-Hilbert space asserts that

$$|\langle x, y\rangle| \le |x| \cdot |y|$$

with strict inequality unless x, y are collinear, i.e., unless one of x, y is a multiple of the other.

Proof: Suppose that x is not a scalar multiple of y, and that neither x nor y is 0. Then $x - \alpha y$ is not 0 for any complex α. Consider

$$0 < |x - \alpha y|^2$$

We know that the inequality is indeed *strict* for all α since x is not a multiple of y. Expanding this,

$$0 < |x|^2 - \alpha\langle x, y\rangle - \bar{\alpha}\langle x, y\rangle + \alpha\bar{\alpha}|y|^2$$

Let

$$\alpha = \mu t$$

with real t and with $|\mu| = 1$ so that

$$\mu\langle x, y\rangle = |\langle x, y\rangle|$$

Then

$$0 < |x|^2 - 2t|\langle x, y\rangle| + |t^2|y|^2$$

The *minimum* of the right-hand side, viewed as a function of the real variable t, occurs when the derivative vanishes, i.e., when

$$0 = -2|\langle x, y\rangle| + 2t|y|^2$$

Using this minimization as a *mnemonic* for the value of t to substitute, we indeed substitute

$$t = \frac{|\langle x, y\rangle|}{|y|^2}$$

into the inequality to obtain

$$0 < |x|^2 + \left(\frac{|\langle x, y\rangle|}{|y|^2}\right)^2 \cdot |y|^2 - 2\frac{|\langle x, y\rangle|}{|y|^2} \cdot |\langle x, y\rangle|$$

which simplifies to

$$|\langle x, y\rangle|^2 < |x|^2 \cdot |y|^2$$

as desired. ///

---

3. Example: spaces ℓ^2

Before any further abstract discussion of Hilbert spaces, we can note that, up to isomorphism,[2] there is just one *infinite-dimensional* Hilbert space occurring in practice,[3] namely the space ℓ^2 constructed as follows. Proof that *most* Hilbert spaces are isomorphic to this one will be given later.

Let ℓ^2 be the collection of sequences $a = \{a_i : 1 \le i < \infty\}$ of complex numbers meeting the constraint

$$\sum_{i=1}^{\infty} |a_i|^2 < +\infty$$

For two such sequences $a = \{a_i\}$ and $b = \{b_i\}$, the *inner product* is

$$\langle a, b\rangle = \sum_i a_i \overline{b_i}$$

The associated *norm* is[4]

$$|a| = \langle a, a\rangle^{1/2}$$

We can immediately generalize this construction in one fashion by replacing the countable set $\{1, 2, 3, \ldots\}$ by an arbitrary set A.[5] Let A be an arbitrary index set, and let $\ell^2(A)$ be the collection of complex-valued functions f on A such that

$$\sum_\alpha |f(\alpha)|^2 < +\infty$$

[2] And, lest anyone be fooled, often the description of such isomorphisms is where any subtlety lies.

[3] Most infinite-dimensional Hilbert spaces occurring in practice have a countable dense subset, and this itself is because the Hilbert spaces are completions of spaces of continuous functions on topological spaces with a countable basis to the topology. This will be amplified subsequently.

[4] That the triangle inequality holds is not immediate, needing the Cauchy-Schwarz-Bunyakowsky inequality. We will give the proof shortly.

[5] Replacement of $\{1, 2, \ldots\}$ by an arbitrary set A is mildly pointless except as an exercise in technique, since, as noted already, in practice we will rarely encounter Hilbert spaces not isomorphic to ℓ^2, if not already isomorphic to the finite-dimensional spaces $\mathbf{C}^n$.

For two such functions f, φ, define an inner product by

$$\langle f, \varphi \rangle = \sum_\alpha f(\alpha)\overline{\varphi(\alpha)}$$

For any *finite* subset A_o of A_o we can apply the *Cauchy-Schwarz-Bunyakowsky inequality* to obtain

$$|\sum_{\alpha \in A_o} f(\alpha)\overline{\varphi(\alpha)}| \le (\sum_{\alpha \in A_o} |f(\alpha)|^2)^{1/2} \, (\sum_{\alpha \in A_o} |\varphi(\alpha)|^2)^{1/2}$$

Thus, the net of all partial sums $\sum_{\alpha \in A_o} f(\alpha)\overline{\varphi(\alpha)}$ has a limit, so is necessarily *Cauchy*.

In fact, the sum $\langle f, \varphi \rangle = \sum_\alpha f(\alpha)\overline{\varphi(\alpha)}$ is *absolutely convergent:* for each α let μ_α be a complex number of absolute value 1 so that

$$\mu_\alpha f(\alpha)\varphi(\alpha) = |f(\alpha)\varphi(\alpha)|$$

and let

$$F(\alpha) = \mu_\alpha f(\alpha)$$

Then F is still in $\ell^2(A)$, and

$$\langle F, \varphi \rangle = \sum_\alpha |f(\alpha)\overline{\varphi(\alpha)}|$$

We just saw that the partial sums of the latter infinite sum form a Cauchy net, so we have the asserted absolute convergence.

Remark: The more general spaces $L^2(X, \mu)$ for abstract measure spaces X, μ have a similar treatment, but need somewhat greater preparation in terms of *integration theory*.

---

4. Triangle inequality, associated metric, continuity issues

As corollary of the Cauchy-Schwarz-Bunyakowsky inequality, we have a *norm* and associated *metric* topology on a pre-Hilbert space:

Again, the **(associated) norm** on a pre-Hilbert space V is

$$|x - y| = \langle x - y, x - y \rangle^{1/2}$$

and the **associated metric** is

$$d(x, y) = |x - y|$$

The reflexivity, symmetry, and positivity of this alleged distance function are clear from the definitional properties of $\langle , \rangle$, but the **triangle inequality**

$$d(x, z) \le d(x, y) + d(y, z)$$

needs proof. That is, we want

$$|x - z| \le |x - y| + |y - z|$$

Assuming for the moment that this triangle inequality holds, we *do* have a metric on the pre-Hilbert space V, and we can show that the map

$$\langle , \rangle : V \times V \to \mathbf{C}$$

is *continuous* as a function of two variables. Indeed, suppose that $|x - x'| < \varepsilon$ and $|y - y'| < \varepsilon$ for $x, x', y, y' \in V$. Then

$$\langle x, y \rangle - \langle x', y' \rangle = \langle x - x', y \rangle + \langle x', y \rangle - \langle x', y' \rangle = \langle x - x', y \rangle + \langle x', y - y' \rangle$$

Using the triangle inequality for the ordinary absolute value, and then the Cauchy-Schwarz-Bunyakowsky inequality, we obtain

$$|\langle x,y\rangle - \langle x',y'\rangle| \le |\langle x-x',y\rangle| + |\langle x', y-y'\rangle| \le |x-x'||y| + |x'||y-y'|$$

$$< \varepsilon(|y|+|x'|)$$

This proves the continuity of the inner product.

Further, scalar multiplication and vector addition are readily seen to be continuous. In particular, it is easy to check that for any fixed $y \in V$ and for any fixed $\lambda \in \mathbf{C}^\times$ both maps

$$x \to x + y$$

$$x \to \lambda x$$

are *homeomorphisms* of V to itself.

Now we prove the desired inequality

$$|x-z| \le |x-y| + |y-z|$$

which is equivalent to the triangle inequality for the alleged metric $d(x,y) = |x-y|$. The appearance of this can be simplified a bit. Replacing x, z by $x+y, z+y$ in this, we see that we want

$$|x-z| \le |x| + |z|$$

We have

$$|x-z|^2 = |x|^2 - \langle x,z\rangle - \langle z,x\rangle + |z|^2 \le |x|^2 + 2|x||z| + |z|^2$$

by the Cauchy-Schwarz-Bunyakowsky inequality. The right-hand side is the square of $|x|+|z|$, as desired. *Done.*

---

5. Hilbert spaces, completions, infinite sums

If a pre-Hilbert space is *complete* with respect to the metric arising from its inner product (and norm), then it is called a **Hilbert space**.

An arbitrary pre-Hilbert space can be *completed* as metric space. Since metric spaces have *countable local bases* for their topology (e.g., open balls of radii $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$) all points in the completion are limits of Cauchy *sequences* (rather than being limits of more complicated Cauchy *nets*). The completion inherits an inner product defined by a limiting process

$$\langle \lim_m x_m, \lim_n y_n\rangle = \lim_{m,n} \langle x_m, y_n\rangle$$

It is not hard to verify that the indicated limit exists (for Cauchy sequences $\{x_m\}, \{y_n\}$), and gives a hermitian inner product on the completion. The completion process *does nothing* to a space which is already complete.

In a Hilbert space, we can consider *infinite* sums

$$\sum_{\alpha\in A} v_\alpha$$

for sets $\{v_\alpha : \alpha \in A\}$ of vectors in V. Not wishing to have a notation that only treats sums indexed by $1, 2, 3, \ldots$, we must consider the **directed system** $\mathcal{A}$ of all finite subsets of A. Consider the **net** of *finite partial sums* of $\sum v_\alpha$ indexed by $\mathcal{A}$ by

$$s(A_o) = \sum_{\alpha\in A_o} v_\alpha$$

where $A_o \in \mathcal{A}$. This is a **Cauchy net** if, given $\varepsilon > 0$, there is a finite subset A_o of A so that for any two finite subsets A_1, A_2 of A both containing A_o we have

$$|s(A_1) - s(A_2)| < \varepsilon$$

If the net is Cauchy, then by the *completeness* there is a unique $v \in V$, the **limit of the Cauchy net**, so that for all $\varepsilon > 0$ there is a finite subset A_o of A so that for any finite subset A_1 of A containing A_o we have

$$|s(A_1) - v| < \varepsilon$$

6. Minimum principle

This fundamental minimum principle asserts that *a non-empty closed convex set in a Hilbert space has a unique element of least norm.* This is essential in the sequel.

Proof: Recall that a set C in a vector space is *convex* if, for all $x, y \in C$ and $0 \le t \le 1$,

$$tx + (1-t)y \in C$$

Let x, y be two elements in a closed convex set C inside a Hilbert space V so that both $|x|$ and $|y|$ are within $\varepsilon > 0$ of the infimum μ of the norms of elements of C. Then

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

Since C is convex,

$$\frac{x+y}{2} \in C$$

Thus,

$$|x+y|^2 = 4|\frac{x+y}{2}|^2 \ge 4\mu^2$$

Thus,

$$|x-y|^2 = 2|x|^2 + 2|y|^2 - |x+y|^2 \le 2(\mu+\varepsilon)^2 + 2(\mu+\varepsilon)^2 - 4\mu^2 = \varepsilon \cdot (8\mu + 4\varepsilon)$$

Thus, any sequence (or *net*) in C whose norms approach the infimum must be a Cauchy sequence (*net*). Since C is closed, such a sequence must converge to an element of C. Further, the inequality shows that any two Cauchy sequences (or *nets*) converging to points minimizing the norm on C must have the same limit. Thus, the minimizing point is *unique*, as claimed.

7. Orthogonal projections to closed subspaces

The next essential ingredient makes use of the minimization principle above:

Let W be a complex vector subspace of a pre-Hilbert space V. If W is *closed* in the topology on V then, reasonably enough, we say that W is a **closed subspace**. For an arbitrary complex vector subspace W of a pre-Hilbert space V, the topological closure $\bar{W}$ is readily checked to be a complex vector subspace of V, so is a *closed subspace*. Because it is necessarily complete, *a closed subspace of a Hilbert space is a Hilbert space in its own right.*

Let W be a closed subspace of a Hilbert space V. Let $v \in V$. We have seen that the closed convex subset

$$v + W = \{v + w : w \in W\}$$

of V has a unique element v_o of least norm. The **orthogonal projection** pv of v to W is

$$pv = v - v_o$$

We *claim* that pv is the unique element in W so that

$$\langle v - pv, w\rangle = 0$$

for all $w \in W$. And then v_1 is called the **orthogonal projection** of v to W.

If there were two vectors $v_1, v_2 \in W$ so that

$$\langle v - v_i, w\rangle = 0$$

for both $i = 1, 2$ and for all $w \in W$, then, by subtracting, we would have

$$\langle v_1 - v_2, w\rangle = 0$$

for all $w \in W$. In particular, we could take $w = v_1 - v_2$, so we see that necessarily $v_1 - v_2 = 0$. This proves *uniqueness*.

Now let v_o be the unique element of $v + W$ of least norm. For any $w \in W$, the vector $v_o + w$ is still in $v + W$, so

$$\langle v_o, v_o\rangle \le \langle v_o + w, v_o + w\rangle$$

from which it follows that

$$0 \le \langle v_o, w\rangle + \langle w, v_o\rangle + |w|^2$$

Replacing w by μw with μ a complex number with $|\mu| = 1$ and

$$\langle v_o, \mu w\rangle = -|\langle v_o, w\rangle|$$

we have

$$0 \le 2|\langle v_o, w\rangle| + |w|^2$$

Replacing w by tw with $t > 0$, this is

$$0 \le 2t|\langle v_o, w\rangle| + t^2|w|^2$$

Dividing by t and letting $t \to 0^+$, this gives

$$\langle v_o, w\rangle = 0$$

Then

$$\langle v - pv, w\rangle = \langle v - (v - v_o), w\rangle = \langle v_o, w\rangle = o$$

as required. *Done.*

---

8. Orthogonal complements $W^\perp$

Let W be a complex vector subspace of a pre-Hilbert space V. Define the **orthogonal complement** $W^\perp$ of W by

$$W^\perp = \{v \in V : \langle v, w\rangle = 0, \quad \forall w \in W\}$$

It is easy to check that $W^\perp$ is a complex vector subspace of V. Since for each $w \in W$ the set

$$w^\perp = \{v \in V : \langle v, w\rangle = 0\}$$

is the inverse image of the closed set $\{0\}$ of $\mathbf{C}$ under the continuous map

$$v \to \langle v, w \rangle$$

it is *closed*. Thus, the orthogonal complement $W^\perp$ is the intersection of a family of closed sets, so is *closed*.

One point here is that *if the topological closure $\bar{W}$ of W in a Hilbert space V is properly smaller than V then $W^\perp \neq \{0\}$.* Indeed, if $\bar{W} \neq V$ then we can find $y \not\in \bar{W}$. Let py be the orthogonal projection of y to $\bar{W}$. Then $y_o = y - py$ is non-zero and is orthogonal to W, so is orthogonal to $\bar{W}$, by continuity of the inner product. Thus, as claimed, $W^\perp \neq \{0\}$.

As a corollary, for any complex vector subspace W of the Hilbert space V, *the topological closure of W in V is the subspace*

$$\bar{W} = W^{\perp\perp}$$

One direction of containment, namely that

$$\bar{W} \subset W^{\perp\perp}$$

is easy: it is immediate that $W \subset W^{\perp\perp}$, and then since the latter is closed we get the asserted containment. If $W^{\perp\perp}$ were strictly larger than $\bar{W}$, then there would be y in it not lying in $\bar{W}$. Now $W^{\perp\perp}$ is a Hilbert space in its own right, in which $\bar{W}$ is a closed subspace, so the orthogonal complement of $\bar{W}$ in $W^{\perp\perp}$ contains a non-zero element z, from above. But then $z \in W^\perp$, and certainly

$$W^\perp \cap (W^\perp)^\perp = \{0\}$$

contradiction. *Done.*

---

9. Riesz-Fischer theorem on linear functionals

A **(linear) functional** λ on a pre-Hilbert space V is a complex-valued function λ on V so that for $\alpha \in \mathbf{C}$ and $x, y \in V$

$$\lambda(x + y) = \lambda(x) + \lambda(y) \qquad \text{(additivity)}$$

$$\lambda(\alpha x) = \alpha\lambda(x) \qquad \text{(linearity)}$$

The **kernel** or **nullspace** of a linear functional λ is

$$\ker \lambda = \{v \in V : \lambda(v) = 0\}$$

A functional is **continuous** if it is continuous in the topology on V with the usual topology on $\mathbf{C}$. A functional is **bounded** if there is a finite real constant C so that, for all $x \in V$,

$$|\lambda(x)| < C|x|$$

The collection of all continuous linear functionals on a pre-Hilbert space V is denoted by

$$V^*$$

We claim that *continuity of the functional λ is equivalent to boundedness*. Indeed, continuity *at zero* is the assertion that for all $\varepsilon > 0$ there is an open ball $B = \{x \in V : |x| < \delta\}$ (with $\delta > 0$) such that $|\lambda(x)| < \varepsilon$ for $x \in B$. In particular, take $\delta > 0$ so that for $|x| < \delta$ we have

$$|\lambda(x)| < 1$$

Then for arbitrary $0 \neq x \in V$ we have

$$|\frac{\delta}{2|x|} \cdot x| < \delta$$

Therefore,

$$|\lambda(\frac{\delta}{2|x|} \cdot x)| < 1$$

That is, using the linearity of λ,

$$|\lambda(x)| < \frac{2|x|}{\delta}$$

That is, we see that continuity implies boundedness.

On the other hand, suppose that there is a finite real constant C so that, for all $x \in V$,

$$|\lambda(x)| < C|x|$$

Then for $|x - y| < \varepsilon/C$ we have

$$|\lambda(x) - \lambda(y)| = |\lambda(x - y)| < C|x - y| < C \cdot \frac{\varepsilon}{C} = \varepsilon$$

showing that boundedness implies continuity. Thus, we have proven that *boundedness and continuity are equivalent.*

For a pre-Hilbert space V with completion $\bar{V}$, *a continuous linear functional λ on V has a unique extension to a continuous linear functional on $\bar{V}$*, defined by

$$\bar{\lambda}(\lim_n x_n) = \lim_n \lambda(x_n)$$

It is not difficult to check that this formula gives a well-defined function (due to the continuity of the original λ), and is additive and linear.

Now we prove the **Riesz-Fischer theorem:** *every continuous linear functional λ on a pre-Hilbert space V is of the form*

$$\lambda(x) = \langle x, y \rangle$$

for a uniquely-determined y in the completion $\bar{V}$ of V.

To prove this, we may as well suppose at the outset that V is *complete*, i.e., is a Hilbert space, since every continuous linear functional extends to the completion anyway. Let λ be a non-zero continuous linear functional on V. Then the kernel N of λ is a proper closed subspace. From above, there is a non-zero element $z \in N^\perp$. Let

$$y = \frac{z}{\lambda(z)}$$

so that $\lambda(y) = 1$. Then, for any $v \in V$,

$$\lambda(v - \lambda(v)y) = \lambda(v) - \lambda(v) \cdot 1 = 0$$

so $v - \lambda(v)y \in N$. Therefore,

$$0 = \langle v - \lambda(v)y, y \rangle$$

Thus,

$$\langle v, y \rangle = \lambda(v)\langle y, y \rangle$$

so that

$$\langle v, \frac{y}{\langle y, y \rangle} \rangle = \lambda(v)$$

as desired. *Done.*

10. Orthonormal sets, separability

A set $\{e_\alpha : \alpha \in A\}$ in a pre-Hilbert space V is **orthogonal** if

$$\langle e_\alpha, e_\beta \rangle = 0$$

for all $\alpha \neq \beta$. If, further,

$$|e_\alpha| = 1$$

for all indices then the set is **orthonormal**. An orthogonal set of non-zero vectors can easily be turned into an ortho*normal* one by replacing each e_α by $e_\alpha/|e_\alpha|$.

We claim that not only are the elements e_α in an orthonormal set *linearly independent* in the usual purely algebraic sense, but, further, if we have a convergent infinite sum $\Sigma_{\alpha \in A} c_\alpha e_\alpha$ with complex c_α and if

$$\Sigma_\alpha c_\alpha e_\alpha = 0$$

then all coefficients c_α are 0. Indeed, for given $\varepsilon > 0$ let A_o be a large-enough finite subset of A so that for any finite subset $A_1 \supset A_o$

$$|\Sigma_{\alpha \in A_1} c_\alpha e_\alpha| < \varepsilon$$

Then for any particular index β we may as well enlarge A_1 to include β, and

$$|\langle \Sigma_{\alpha \in A_1} c_\alpha e_\alpha, e_\beta \rangle| \leq |\Sigma_{\alpha \in A_1} c_\alpha e_\alpha| \cdot |e_\beta| < \varepsilon \cdot |e_\beta| = \varepsilon$$

by the Cauchy-Schwarz-Bunyakowsky inequality. On the other hand, using the orthonormality,

$$|\langle \Sigma_{\alpha \in A_1} c_\alpha e_\alpha, e_\beta \rangle| = |c_\beta| \cdot |e_\beta|^2 = |c_\beta|$$

Together, this gives $|c_\beta| < \varepsilon$. Since this is true for all $\varepsilon > 0$, it must be that $c_\beta = 0$. This holds for all indices β. *Done.*

A *maximal* orthonormal set inside a pre-Hilbert space is called an **orthonormal basis**. The property of maximality of an orthonormal set $\{e_\alpha : \alpha \in A\}$ is the natural one, that there be no *other* unit vector e perpendicular to all the e_α. (It is important to note that the phrase 'orthonormal basis' has a different sense in other contexts).

Let $\{e_\alpha : \alpha \in A\}$ be an orthonormal set in a Hilbert space V. Let W_o be the complex vector space of all finite linear combinations of vectors in $\{e_\alpha : \alpha \in A\}$. Then we claim that $\{e_\alpha : \alpha \in A\}$ *is an orthonormal basis if and only if* W_o *is dense in* V. Indeed, if the closure W of W_o were a proper subspace of V, then it would have a non-trivial orthogonal complement, so we could make a further unit vector, so $\{e_\alpha : \alpha \in A\}$ could not have been maximal. On the other hand, if $\{e_\alpha : \alpha \in A\}$ is not maximal, let e be a unit vector orthogonal to all the e_α. Then e is orthogonal to all finite linear combinations of the e_α, so is orthogonal to W_o, and thus to W by continuity. That is, W_o cannot be dense. ///

Next, we show that *any orthonormal set can be enlarged to be an orthonormal basis*. To prove this requires invocation of an equivalent of the Axiom of Choice. Specifically, we want to *order* the collection X of orthonormal sets (containing the given one) by inclusion, and note that any *totally ordered* collection of orthonormal sets in X has a supremum, namely the union of all. Thus, we are entitled to conclude that there *are* maximal orthonormal sets containing the given one. If such a maximal one were not an orthonormal basis, then (as observed just above) we could find a further unit vector orthogonal to all vectors in the orthonormal set, contradicting the maximality within X. ///

If a Hilbert space has a *countable* orthonormal basis, then it is called **separable**. Most Hilbert spaces of practical interest are separable, but at the same time most elementary results do not make any essential use of separability so there is no compulsion to worry about this at the moment.

---

11. Bessel inequality, Parseval isomorphism

Let $\{e_\alpha : \alpha \in A\}$ be an orthonormal basis in a Hilbert space V. *Granting* for the moment that $v \in V$ has an expression

$$v = \sum_\alpha c_\alpha e_\alpha$$

we can determine the coefficients c_α, as follows. By the continuity of the inner product, this equality yields

$$\langle v, e_\beta \rangle = \langle \sum_\alpha c_\alpha e_\alpha, e_\beta \rangle = \sum_\alpha c_\alpha \langle e_\alpha, e_\beta \rangle = c_\alpha$$

An expression

$$v = \sum_\alpha c_\alpha e_\alpha$$

is an **abstract Fourier expansion**

$$v = \sum_\alpha \langle v, e_\alpha \rangle e_\alpha$$

The coefficients are the (abstract) **Fourier coefficients** in terms of the orthonormal basis.

Remark: We have not quite proven that every vector *has* such an expression. We do so after proving a necessary preparatory result.

Proposition: *(Bessel's inequality)* Let $\{e_\beta : \beta \in B\}$ be an orthonormal set in a Hilbert space V. Then

$$\langle v, v \rangle \geq \sum_{\beta \in B} |\langle v, e_\beta \rangle|^2$$

Proof: Just using the positivity (and continuity) and orthonormality

$$0 \leq |v - \sum_{\beta \in B} \langle v, e_\beta \rangle \, e_\beta|^2 = |v|^2 - \sum_{\beta \in B} \langle v, e_\beta \rangle \overline{\langle v, e_\beta \rangle} - \sum_{\beta \in B} \overline{\langle v, e_\beta \rangle} \langle v, e_\beta \rangle + \sum_{\beta \in B} |\langle v, e_\beta \rangle|^2$$

$$= |v|^2 - \sum_{\beta \in B} |\langle v, e_\beta \rangle|^2$$

This gives the desired inequality. ///

Proposition: Every vector $v \in V$ has a unique expression as

$$v = \sum_{\alpha \in A} c_\alpha e_\alpha$$

More precisely, for $v \in V$ and for each finite subset B of A let

$$v_B = \text{projection of } v \text{ to } \textstyle\sum_{\alpha \in B} \mathbf{C} \cdot e_\alpha = v - \displaystyle\sum_{\alpha \in B} \langle v, e_\alpha \rangle \, e_\alpha$$

Then the net

$$\{v_B : B \text{ finite}, B \subset A\}$$

is Cauchy and has limit v.

Proof: Uniqueness follows from the previous discussion of the density of the subspace V_o of finite linear combinations of the e_α.

Bessel's inequality

$$|v|^2 \geq \sum_{\alpha \in B} |\langle v, e_\alpha \rangle|^2$$

implies that the net is Cauchy, since the tails of a convergent sum must go to 0. Let w be the limit of this net. Given $\varepsilon > 0$, let B be a large enough finite subset of A such that for finite subset $C \supset B$ $|w - v_C| < \varepsilon$. Given $\alpha \in A$ enlarge B if necessary so that $\alpha \in B$. Then

$$|\langle v - w, e_\alpha \rangle| \leq |\langle v - v_B, e_\alpha \rangle| + |\langle w - v_B, e_\alpha \rangle| \leq |\langle v - v_B, e_\alpha \rangle| + |\langle w - v_B, e_\alpha \rangle| \leq 0 + |w - v_B| < \varepsilon$$

since $\langle v - v_B, e_\alpha \rangle = 0$ for $\alpha \in B$. Thus, if $v \neq w$, we can construct a further vector of length 1 orthogonal to all the e_α, namely a unit vector in the direction of $v - w$. This would contradict the maximality of the collection of e_α. ///

Remark: If V were only a pre-Hilbert space, that is, were not complete, then a maximal collection of mutually orthogonal vectors of length 1 may not have the property of the theorem. That is, the collection of (finite) linear combinations may fail to be dense. This is visible in the proof above, wherein we needed to be able to take the limit that yielded the auxiliary vector w. For example, inside the standard ell^2 let $e_1, e_2, \ldots$ be the usual

$$e_1 = (1, 0, 0, \ldots),\ e_1 = (0, 1, 0, \ldots),\ \text{(etc.)}$$

and let

$$v_1 = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots)$$

Let V be pre-Hilbert space which is the (algebraic) span of

$$v_1, e_2, e_3, \ldots$$

Certainly

$$B = \{e_2, e_3, \ldots\}$$

is an orthonormal set. In fact, this collection is *maximal*, but that v_1 is *not* in the closure of the span of B.

For $v \in V$, write

$$\hat{v} = \langle v, e_\alpha \rangle$$

Corollary: *(Parseval isomorphism)* The map

$$v \to \hat{v}$$

is *an isomorphism of Hilbert spaces*

$$V \to \ell^2(A)$$

That is, the map is an isomorphism of complex vector spaces, is a homeomorphism of topological spaces, and

$$\langle v, w \rangle = \langle \hat{v}, \hat{w} \rangle \qquad |v|^2 = |\hat{v}|^2_{\ell^2(A)}$$

(where the inner product on the left is that in V and the inner product on the right is that in $\ell^2(A)$.)

Proof: Expand any vector v in terms of the given orthonormal basis as

$$v = \sum_\alpha \hat{v}(\alpha) v$$

The assertion that

$$\langle v, w\rangle = \langle \hat{v}, \hat{w}\rangle$$

is a consequence of the expansion in terms of the orthonormal basis, together with continuity. That $\hat{v}$ lies in $\ell^2(A)$, and in fact has norm equal to that of v, is the assertion of Parseval.

The only thing of any note is the point that any $\{c_\alpha\} \in \ell^2(A)$ can actually occur as the (abstract) Fourier coefficients of some vector in V. That is, for $f \in \ell^2(A)$, we want to show that the net of finite sums

$$\sum_{\alpha\in A_o} f(\alpha)e_\alpha$$

(for A_o a finite subset of A) is *Cauchy*. Since $f \in \ell^2(A)$, for given $\varepsilon > 0$ there is large-enough finite A_o so that

$$\left(\sum_{\alpha\in A-A_o} |f(\alpha)|^2\right)^{1/2} = |\sum_{\alpha\in A-A_o} f(\alpha)e_\alpha| < \varepsilon$$

(using the orthonormality). Then for A_1, A_2 both containing A_o,

$$|\sum_{\alpha\in A_1} f(\alpha)e_\alpha - \sum_{\alpha\in A_2} f(\alpha)e_\alpha|^2 = \sum_{\alpha\in (A_1\cup A_2)-A_o} |f(\alpha)e_\alpha|^2 \le \sum_{\alpha\in A-A_o} |f(\alpha)|^2 < \varepsilon^2$$

From this the Cauchy property follows. *Done.*

---

12. Riemann-Lebesgue lemma

The result of this section is an essentially trivial consequence of previous observations, and is certainly much simpler to prove than the genuine Riemann-Lebesgue lemma for Fourier *transforms*.

Let $\{e_\alpha : \alpha \in A\}$ be an orthonormal basis for a Hilbert space V. For $v \in V$, write

$$\hat{v} = \langle v, e_\alpha\rangle$$

Then the version of a **Riemann-Lebesgue lemma** relevant here is that

$$\lim_\alpha |\hat{v}(\alpha)| = 0$$

More explicitly, this means that for given $\varepsilon > 0$ there is a finite subset A_o of A so that for $\alpha \notin A_o$ we have

$$|\hat{v}(\alpha)| < \varepsilon$$

This follows from the fact that the infinite sum

$$\sum_\alpha |\hat{v}(\alpha)|^2$$

is *convergent.*

---

13. The Gram-Schmidt process

Let $S = \{v_n : n = 1, 2, 3, \ldots\}$ be a well-ordered set of vectors in a pre-Hilbert space V. For simplicity, we are also assuming that S is *countable.* Let V_o be the collection of all finite linear combinations of S, and suppose that S is *dense* in V. Then we can obtain an *orthonormal basis* from S by the following procedure, called the **Gram-Schmidt process:**

Let v_{n_1} be the first of the v_i which is non-zero, and put

$$e_1 = \frac{v_{n_1}}{|v_{n_1}|}$$

Let v_{n_2} be the first of the v_i which is *not* a multiple of e_1. Put

$$f_2 = v_{n_2} - \langle v_{n_2}, e_1 \rangle e_1$$

and

$$e_2 = \frac{f_2}{|f_2|}$$

Inductively, suppose we have chosen $e_1, \ldots, e_k$ which form an orthonormal set. Let $v_{n_{k+1}}$ be the first of the v_i not expressible as a linear combination of $e_1, \ldots, e_k$. Put

$$f_{k+1} = v_{n_{k+1}} - \sum_{1 \le i \le k} \langle v_{n_{k+1}}, e_i \rangle e_i$$

and

$$e_{k+1} = \frac{f_{k+1}}{|f_{k+1}|}$$

Then induction on k proves that the collection of all finite linear combinations of $e_1, \ldots, e_k$ is the same as the collection of all finite linear combinations of $v_1, v_2, v_3, \ldots, v_{n_k}$. Thus, the collection of all finite linear combinations of the orthonormal set $e_1, 3_2, \ldots$ is *dense* in V, so this is an orthonormal basis.