(*October 29, 2016*)

Hilbert spaces

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

[This document is http://www.math.umn.edu/˜garrett/m/fun/notes_2016-17/03_hsp.pdf]

Hilbert spaces are possibly-infinite-dimensional analogues of the familiar finite-dimensional Euclidean spaces. 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.

Unfortunately, few naturally-occurring spaces of functions are Hilbert spaces. Given the intuitive geometry of Hilbert spaces, this is disappointing, suggesting that physical intuition is a little distant from the behavior of natural spaces of functions. However, a little later we will see that suitable *families* of Hilbert spaces *can* capture what we want. Such ideas were developed by Beppo Levi (1906), Frobenius (1907), and Sobolev (1930's). These ideas do fit into Schwartz' (c. 1950) formulation of his notion of *distributions*, but it seems that they were not explicitly incorporated, or perhaps were viewed as completely obvious at that point. We will see that Levi-Sobolev ideas offer some useful specifics in addition to Schwartz' over-arching ideas.

Most of the geometric results on Hilbert spaces are corollaries of the *minimum principle*.

Most of what is done here applies to vector spaces over either $\mathbb{R}$ or $\mathbb{C}$.

---

1. Cauchy-Schwarz-Bunyakowsky inequality

A complex vector space V with a complex-valued function

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

of two variables on V is a *(hermitian) inner product space* or *pre-Hilbert space*, and $\langle , \rangle$ is a *(hermitian) inner product*, when we have the usual conditions

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

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*.

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 will make clear which is meant. The *metric* on a Hilbert space is $d(v,w) = |v-w|$: the triangle inequality follows from the *Cauchy-Schwarz-Bunyakowsky inequality* just below.

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,

$$|x+y|^2 - |x-y|^2 \;=\; 2\langle x, y\rangle + 2\langle y, x\rangle \;=\; 4\mathrm{Re}\,\langle x, y\rangle$$

Therefore,

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

These and closely-related identites are of frequent use.

[1.1] Theorem: *(Cauchy-Schwarz-Bunyakowsky inequality)*

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

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,

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

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. ///

[1.2] Corollary: *(Triangle inequality)* For v, w in a Hilbert space V, we have $|v+w| \le |v| + |w|$. Thus, with distance function $d(v,w) = |v - w|$, we have the triangle inequality

$$d(x,z) \;=\; |x-z| \;=\; |(x-y)+(y-z)| \;\le\; |x-y| + |y-z| \;=\; d(x,y) + d(y,z)$$

Proof: Squaring and expanding, noting that $\langle v, w\rangle + \langle w, v\rangle = 2\mathrm{Re}\,\langle v, w\rangle$,

$$(|v|+|w|)^2 - |v+w|^2 \;=\; \Big(|v|^2 + 2|v|\cdot|w| + |w|^2\Big) - \Big(|v|^2 + \langle v,w\rangle + \langle w,v\rangle + |w|^2\Big) \;\ge\; 2|v|\cdot|w| - 2|\langle v,w\rangle| \;\ge\; 0$$

giving the asserted inequality. ///

An inner product space *complete* with respect to the metric arising from its inner product (and norm) is a *Hilbert space*.

[1.3] Continuity issues

The map

$$\langle , \rangle : V \times V \longrightarrow \mathbb{C}$$

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 $\lambda \in \mathbb{C}^\times$ both maps

$$x \to x + y$$

$$x \to \lambda x$$

are *homeomorphisms* of V to itself.

2. *Example:* ℓ^2

Before further abstract discussion, we note that, up to isomorphism, there is essentially just one *infinite-dimensional* Hilbert space occurring in practice, namely the space ℓ^2 constructed as follows. Most infinite-dimensional Hilbert spaces occurring in practice have a countable dense subset, because the Hilbert spaces are completions of spaces of continuous functions on topological spaces with a countably-based topology.

Lest anyone be fooled, often subtlety is in the description of the isomorphisms and *mappings* among such Hilbert spaces.

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

$$\sum_{i=1}^{\infty} |f(i)|^2 < +\infty$$

For two such sequences f and g, the *inner product* is

$$\langle f, g\rangle = \sum_i f(i)\overline{g(i)}$$

[2.1] Claim: ℓ^2 is a vector space. The sum defining the inner product on ℓ^2 is absolutely convergent.

Proof: That ℓ^2 is closed under scalar multiplication is clear. For $f, g \in \ell^2$, by Cauchy-Schwarz-Bunyakowsky,

$$\Big|\sum_{n\le N} f(i)\cdot\overline{g(i)}\Big| \le \sum_{n\le N} |f(i)|\cdot|g(i)| \le \Big|\sum_{n\le N}|f(i)|^2\Big|^{\frac{1}{2}}\cdot\Big|\sum_{n\le N}|g(i)|^2\Big|^{\frac{1}{2}} \le |f|_{\ell^2}\cdot|g|_{\ell^2}$$

giving the absolute convergence of the infinite sum for $\langle f, g\rangle$. Then, expanding,

$$\sum_{n\le N}|f(i)+g(i)|^2 \le \sum_{n\le N}|f(i)|^2 + 2|f(i)|\cdot|g(i)| + |g(i)|^2 < +\infty$$

by the previous. ///

[2.2] Claim: ℓ^2 is *complete.*

Proof: Let $\{f_n\}$ be a Cauchy sequence of elements in ℓ^2. For every $i \in \{1, 2, 3, \ldots\}$,

$$|f_m(i) - f_n(i)|^2 \le \sum_{i\ge 1}|f_m(i) - f_n(i)|^2 = |f_m - f_n|^2_{\ell^2}$$

so $f(i) = \lim_n f_n(i)$ exists for every i, and is the obvious candidate for the limit in ℓ^2. It remains to see that this limit is indeed in ℓ^2. This will follow from an easy case of Fatou's lemma:

$$\sum_i |f(i)|^2 = \sum_i |\lim_n f_n(i)| = \sum_i |\liminf_n f_n(i)| \le \liminf_n \sum_i |f_n(i)| = \liminf_n |f_n|^2_{\ell^2}$$

Since $\{f_n\}$ is a Cauchy sequence, certainly $\lim_n |f_n|^2_{\ell^2}$ exists. ///

[2.3] Remark: A similar result holds for $L^2(X, \mu)$ for general measure spaces X, μ, but requires more preparation.

3. *Completions, infinite sums*

An arbitrary pre-Hilbert space can be *completed* as metric space, giving a *Hilbert 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 can 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

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

4. *Minimum principle, orthogonality*

This fundamental minimum principle, that *a non-empty closed convex*[1] *set in a Hilbert space has a unique element of least norm*, is essential in the sequel. It generally *fails* in more general types of topological vector spaces.

[4.1] Theorem: A non-empty closed convex subset of a HIlbert space has a unique element of least norm.

Proof: For two elements x, y in a closed convex set C inside a Hilbert space with both $|x|$ and $|y|$ within $\varepsilon > 0$ of the infimum μ of the norms of elements of C,

$$|x-y|^2 \;=\; 2|x|^2+2|y|^2-|x+y|^2 \;=\; 2|x|^2+2|y|^2-4\Big(\frac{|x+y|}{2}\Big)^2 \;\le\; 2(\mu+\varepsilon)^2+2(\mu+\varepsilon)^2-4\mu^2 \;=\; \varepsilon\cdot(8\mu+4\varepsilon)$$

since $\frac{x+y}{2} \in C$ by convexity of C. Thus, any sequence (or *net*) in C with norms approaching the inf is a Cauchy sequence (*net*). Since C is closed, such a sequence converges to an element of C. Further, the

[1] Recall that a set C in a vector space is *convex* when $tx + (1-t)y \in C$ for all $x, y \in C$ and for all $0 \le t \le 1$.

inequality shows that any two Cauchy sequences (or *nets*) converging to points minimizing the norm on C have the same limit. Thus, the minimizing point is *unique*. ///

[4.2] Corollary: Given a closed, convex, non-empty subset E of a Hilbert space V, and a point $v \in V$ not in E, there is a unique point $w \in E$ closest to v.

Proof: Since $v \notin E$, the set $E - v$ does not contain 0. The map $x \to x - v$ is a homeomorphism, because non-zero scalar multiplication and vector addition are continuous, and have continuous inverses. Thus, $E-v$ is closed. It is also still convex. Thus, there is a unique element $x_o - v \in E - v$ of smallest norm. That is, $|x_o - v| < |x - v|$ for all $x \neq x_o$ in E. That is, the distance from x_o is the minimum. ///

[4.3] Orthogonal projections to closed subspaces Existence of orthogonal projections makes essential use of the minimization principle. Let W be a complex vector subspace of a 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 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 $v \in V$. From the corollary just above, the closed convex subset W contains a unique element w_o closest to v.

[4.4] Claim: The element w_o is the *orthogonal projection* of v to W, in the sense that $w_o \in W$ is the unique element in W such that $\langle v - w_o, w\rangle = 0$ for all $w \in W$.

Proof: For two vectors $w_1, w_2 \in W$ so that

$$\langle v - w_i, w\rangle \;=\; 0 \qquad\qquad \text{(for both } i = 1, 2 \text{ and for all } w \in W)$$

by subtracting, we would have

$$\langle w_1 - w_2, w\rangle \;=\; 0$$

for all $w \in W$. In particular, with $w = w_1 - w_2$, necessarily $w_1 - w_2 = 0$, proving *uniqueness* of the orthogonal projection.

With w_o the unique element of W closest to v, for any $w \in W$, since $w_o + w$ is still in W,

$$|v - w_o|^2 < |v - (w_o + w)|^2$$

Expanded slightly, this is

$$|v - w_o|^2 \;\le\; |v - w_o|^2 - \langle v - w_o, w\rangle - \langle w, v - w_o\rangle + |w|^2$$

which gives

$$\langle v - w_o, w\rangle + \langle w, v - w_o\rangle \;\le\; |w|^2$$

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

$$\langle v - w_o, \mu w\rangle \;=\; |\langle v - w_o, w\rangle|$$

this gives

$$2|\langle v - w_o, w\rangle| \;\le\; |w|^2 \qquad\qquad (\text{for } w \neq 0)$$

Replacing w by tw with $t > 0$ gives

$$2t|\langle v - w_o, w\rangle| \;\le\; t^2|w|^2$$

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

$$|\langle v - w_o, w\rangle| \;\le\; 0$$

as required. ///

[4.5] Orthogonal complements $W^\perp$

Let W be a vector subspace of a Hilbert space V. The *orthogonal complement* $W^\perp$ of W is

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

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 $\mathbb{C}$ 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 $y \notin \bar{W}$. 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. ///

[4.6] Orthonormal sets, separability

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

$$\langle e_\alpha, e_\beta\rangle \;=\; 0 \qquad\qquad \text{(for all } \alpha \neq \beta)$$

When

$$|e_\alpha| \;=\; 1$$

for all indices the set is *orthonormal.* An orthogonal set of non-zero vectors is 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, in a convergent infinite sum $\sum_{\alpha\in A} c_\alpha e_\alpha$ with complex c_α with

$$\sum_\alpha c_\alpha e_\alpha \;=\; 0$$

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

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

For any particular index β we may as well enlarge A_1 to include β, and by Cauchy-Schwarz-Bunyakowsky.

$$\Big|\langle \sum_{\alpha\in A_1} c_\alpha e_\alpha, e_\beta\rangle\Big| \;\le\; \Big|\sum_{\alpha\in A_1} c_\alpha e_\alpha\Big| \cdot |e_\beta| \;<\; \varepsilon \cdot |e_\beta| \;=\; \varepsilon$$

On the other hand, using the orthonormality,

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

Together, $|c_\beta| \;<\; \varepsilon$. This holds for all $\varepsilon > 0$, so $c_\beta = 0$. This holds for all indices β. ///

A *maximal* orthonormal set in 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.

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.

---

5. *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

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

An expression

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

is an *abstract Fourier expansion.* The coefficients $c_\alpha = \langle v, e_\alpha\rangle$ are the (abstract) *Fourier coefficients* in terms of the orthonormal basis. When the orthonormal basis $\{e_\alpha : \alpha \in A\}$ is understood, we may write $\widehat{v}(\alpha)$ for $\langle v, e_\alpha\rangle$.

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

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

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

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. ///

[5.3] Claim: 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

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

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

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

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. ///

[5.4] 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 ℓ^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 v_1 is *not* in the closure of the span of B.

For $v \in V$, write

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

[5.5] Corollary: *(Parseval isomorphism)* With orthonormal basis $\{e_\alpha : \alpha \in A\}$, 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 $\ell^2(A)$.) That is,

$$|v|^2 \; = \; \sum_{\alpha \in A} |\langle v, e_\alpha \rangle|^2$$

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

$$v \; = \; \sum_\alpha \hat{v}(\alpha) \; e_\alpha \; = \; \sum_\alpha \langle v, e_\alpha \rangle \; e_\alpha$$

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

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

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

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

From this the Cauchy property follows. ///

6. *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

$$\widehat{v}(\alpha) \;=\; \langle v, e_\alpha \rangle$$

The *Riemann-Lebesgue lemma* relevant here is

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

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

$$|\widehat{v}(\alpha)| \;<\; \varepsilon$$

This follows from the fact that the infinite sum

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

is *convergent.*

7. *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 V_o 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 $v_{n_1}, v_{n_2}, v_{n_3}, \ldots, v_{n_k}$. Thus, the collection of all finite linear combinations of the orthonormal set $e_1, e_2, \ldots$ is *dense* in V, so this is an orthonormal basis.

---

8. *Linear maps, linear functionals, Riesz-Fréchet theorem*

We consider maps $T : V \to W$ from one Hilbert space to another which are not only *linear*, but also *continuous*. The linearity is

$$T(av + bw) \;=\; a \cdot Tv + b \cdot Tw \qquad\qquad \text{(for scalars } a, b \text{ and } v, w \in V\text{)}$$

and the continuity is as expected for a map from one metric space to another: given $v \in V$ and given $\varepsilon > 0$, there is small enough $\delta > 0$ such that for $v' \in V$ with $|v' - v|_V < \delta$, we have $|Tv - Tv'|_W < \varepsilon$.

A *(continuous, linear) functional* λ on a Hilbert space V is a continuous linear map $\lambda : V \to \mathbb{C}$.

The *kernel* or *nullspace* of a linear map T is

$$\ker T \;=\; \{v \in V : Tv = 0\}$$

A linear map $T : V \to W$ is *bounded* when there is a finite real constant C so that, for all $v \in V$,

$$|Tv|_W \;<\; C|v|_V \qquad\qquad \text{(for all } v \in V\text{)}$$

The collection of all continuous linear functionals on a Hilbert space V is denoted by V^*.

[8.1] Claim: *Continuity* of a linear map $T : V \to W$ is equivalent to *boundedness.*

Proof: Continuity *at zero* is the assertion that for all $\varepsilon > 0$ there is an open ball $B = \{v \in V : |v|_V < \delta\}$ (with $\delta > 0$) such that $|Tv|_W < \varepsilon$ for $v \in B$. In particular, take $\delta > 0$ so that for $|v| < \delta$ we have

$$|Tv| \;<\; 1$$

For arbitrary $0 \neq x \in V$ we have

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

Therefore,

$$\Big|T\Big(\frac{\delta}{2|x|} \cdot x\Big)\Big|_W \;<\; 1$$

By the linearity of T,

$$|Tx|_W \;<\; \frac{2}{\delta} \cdot |x|_V$$

That is, continuity implies boundedness.

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

$$|Tx| \;<\; C|x|$$

For $|x - y| < \varepsilon/C$

$$|Tx - Ty|_W \;=\; |T(x-y)|_W \;<\; C|x-y|_V \;<\; C \cdot \frac{\varepsilon}{C} \;=\; \varepsilon$$

showing that boundedness implies continuity. Thus, *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.

The dual V^* has a natural *norm*

$$|\lambda|_{V^*} \;=\; \sup_{v\in V:|v|\le 1} |\lambda(v)| \qquad\qquad (\text{for } \lambda \in V^*)$$

By the *minimum principle*, the sup is attained.

[8.2] Theorem: *(Riesz-Fréchet)* Every continuous linear functional λ on a Hilbert space V is of the form

$$\lambda(x) \;=\; \langle x, y_\lambda\rangle$$

for a uniquely-determined y_λ in V. Further, $|y_\lambda|_V = |\lambda|_{V^*}$. Thus, the map $V \to V^*$ by $v \to \lambda_v$ defined by $\lambda_v(w) = \langle w, v\rangle_V$ is a *conjugate-linear* isomorphism $V \to V^*$: it preserves vector addition and preserves the metric, but scalar multiplication is conjugated: $y_{a\lambda} = \bar{a}\cdot y_\lambda$ for $a \in \mathbb{C}$.

Proof: The kernel $\ker\lambda$ of a non-zero continuous linear functional λ is a proper closed subspace. From above, there is a non-zero element $z \in (\ker\lambda)^\perp$. Replace z by $z/\lambda(z)$ so that $\lambda(z) = 1$ without loss of generality. For any $v \in V$,

$$\lambda(v - \lambda(v)z) \;=\; \lambda(v) - \lambda(v)\cdot 1 \;=\; 0$$

so $v - \lambda(v)z \in \ker\lambda$. Therefore,

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

Thus,

$$\langle v, z\rangle \;=\; \lambda(v)\cdot\langle z, z\rangle$$

so that

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

proving *existence*. For *uniqueness*, when $\langle x, z\rangle = \langle x, z'\rangle$ for specific z, z' and for all x, then $\langle x, z - z'\rangle = 0$ for all x gives $z = z'$, giving uniqueness.

Of course, every $y \in V$ gives a continuous linear functional by $x \to \langle x, y\rangle_V$. This is the inverse map to $\lambda \to y_\lambda$, so both are bijections. Addition is preserved:

$$\langle v, y_{\lambda+\mu}\rangle_V \;=\; (\lambda+\mu)(v) \;=\; \lambda v + \mu v \;=\; \langle v, y_\lambda\rangle_V + \langle v, y_\mu\rangle_V \;=\; \langle v, y_\lambda + y_\mu\rangle_V \qquad (\text{for all } v)$$

The conjugation of scalars follows similarly, from the hermitian-ness of $\langle,\rangle_V$:

$$\langle v, y_{a\lambda}\rangle_V \;=\; (a\lambda)(v) \;=\; a\cdot\lambda v \;=\; a\langle v, y_\lambda\rangle_V \;=\; \langle v, \bar{a}\cdot y_\lambda\rangle_V$$

as claimed. ///

[8.3] Corollary: The dual V^* has a natural Hilbert space structure, given by

$$\langle\lambda, \mu\rangle_{V^*} \;=\; \overline{\langle y_\lambda, y_\mu\rangle_V} \qquad\qquad (\text{where } \lambda(v) = \langle v, y_\lambda\rangle_V \text{ and } \mu(v) = \langle v, y_\mu\rangle_V, \text{ for all } v \in V)$$

Proof: Checking the pre-Hilbert space properties is straightforward. Completeness follows from the property $|y_\lambda|_V = |\lambda|_{V^*}$. ///

[8.4] Corollary: $V \approx (V^*)^*$ as Hilbert spaces, given by map $\varphi : V \to V^{**}$ by $\varphi(v)(\lambda) = \lambda v$.

Proof: One checks directly that φ gives a continuous, injective, complex-linear map $V \to V^{**}$. We claim that it is composite of the two conjugate-linear isomorphisms $V \to V^*$ and $V^* \to (V^*)^*$. Let $v \to \lambda_v = \langle -, v\rangle$ be the map $V \to V^*$, and $\mu \to \Lambda_m u = \langle -, \mu\rangle_{V^*}$ the map $V^* \to (V^*)^*$. For $v, w \in V$,

$$\varphi(v)(\lambda_w) \;=\; \lambda_w(v) \;=\; \langle v, w\rangle_V \;=\; \langle v, w\rangle_V \;=\; \langle \lambda_w, \lambda_v\rangle_{V^*} \;=\; \Lambda_{\lambda_v}(\lambda_w)$$

By Riesz-Fréchet, the vectors λ_w fill out V^* for $w \in V$. Thus, $\varphi(v) = \Lambda_{\lambda_v}$, as claimed. ///

---

9. *Adjoints*

[9.1] Claim: Given a continuous linear map $T : V \to W$ of Hilbert spaces, there is a unique continuous linear $T^* : W^* \to V^*$ characterized by

$$(T^*\mu)(v) \;=\; \mu(Tv) \qquad\qquad (\text{for } \mu \in W^*, \text{ for all } v \in V)$$

Proof: The map $V \to \mathbb{C}$ by $v \to Tv \to \mu(Tv)$ is a composite of continuous functions, so is continuous. It is linear for the same reason. Call it $T^*\mu \in V^*$. To show that $\mu \to T^*\mu$ is continuous, it is convenient to look at *bounds*: since T is continuous, it is bounded, so there is C such that $|Tv|_W \le C \cdot |v|_V$, and then

$$|(T^*\mu)(v)| \;=\; |\mu(Tv)| \;\le\; |\mu|_{W^*} \cdot |Tv|_W \;\le\; |\mu|_{W^*} \cdot C \cdot |v|_V$$

Thus, $|T^*\mu|_{V^*} \le |\mu|_{W^*} \cdot C < \infty$, so T^* is continuous. ///

[9.2] Remark: Somewhat surprisingly, for most continuous linear maps $T : V \to W$ of Hilbert spaces, the Riesz-Fréchet conjugate-linear isomorphisms $\alpha_V : V \to V^*$ and $\alpha_W : W \to W^*$ are not compatible with adjoints. That is, it is rare that the following square *commutes*:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ {\scriptstyle \alpha_V}\downarrow & & \downarrow{\scriptstyle \alpha_W} \\ V^* & \xleftarrow{T^*} & W^* \end{array}$$

In fact, the only situation in which such a square commutes is when T is an isometry to its image.
