Math 212a Lecture 3.

Basic facts about Hilbert spaces.

David Hilbert

Born: **23 Jan 1862 in Königsberg, Prussia (now Kaliningrad, Russia)**
Died: **14 Feb 1943 in Göttingen, Germany**

Review: Scalar products.

V is a complex vector space. A rule assigning to every pair of vectors $f, g \in V$ a complex number (f, g) is called a **semi-scalar product** if

1. (f, g) is linear in f when g is held fixed.
2. $(g, f) = \overline{(f, g)}$. This implies that (f, g) is anti-linear in g when f is held fixed. In other words. $(f, ag + bh) = \overline{a}(f, g) + \overline{b}(f, h)$. It also implies that (f, f) is real.
3. $(f, f) \geq 0$ for all $f \in V$.

If 3. is replaced by the stronger condition

4. $(f, f) > 0$ for all non-zero $f \in V$

then we say that $(\ , \ )$ is a **scalar product**.

Review: The Cauchy-Schwarz inequality.

This says that if (,) is a semi-scalar product then

$$|(f,g)| \leq (f,f)^{\frac{1}{2}}(g,g)^{\frac{1}{2}}. \tag{1}$$

For any semiscalar product define

$$\|f\| := (f,f)^{\frac{1}{2}}$$

so we can write the Cauchy-Schwarz inequality as

$$|(f,g)| \leq \|f\|\|g\|.$$

Review: The triangle inequality, norms.

The **triangle inequality** says that

$$\|f+g\| \leq \|f\| + \|g\|. \tag{3}$$

We also proved that

$$\|cf\| = |c|\|f\| \tag{4}$$

Let V be a complex vector space and let $\|\cdot\|$ be a map which assigns to any $f \in V$ a non-negative real $\|f\|$ number such that $\|f\| > 0$ for all non-zero f. If $\|\cdot\|$ satisfies the triangle inequality (3) and equation (4) it is called a **norm**. A vector space endowed with a norm is called a normed space. The pre-Hilbert spaces can be characterized among all normed spaces by the parallelogram law as we will discuss below.

The Pythagorean theorem.

Let V be a pre-Hilbert space. We have

$$\|f+g\|^2 = \|f\|^2 + 2\mathrm{Re}(f,g) + \|g\|^2.$$

So

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2 \quad \Leftrightarrow \quad \mathrm{Re}\ (f,g) = 0. \tag{5}$$

We make the definition

$$f \perp g \ \Leftrightarrow \ (f,g) = 0$$

and say that f is perpendicular to g or that f is orthogonal to g. Notice that this is a stronger condition than the condition for the Pythagorean theorem, the right hand condition in (5). For example $\|f+if\|^2 = 2\|f\|^2$ but $(f,if) = -i\|f\|^2 \neq 0$ if $\|f\| \neq 0$.

Orthogonality and independence.

If u_i is some finite collection of mutually orthogonal vectors, then so are $z_i u_i$ where the z_i are any complex numbers. So if

$$u = \sum_i z_i u_i$$

then by the Pythagorean theorem

$$\|u\|^2 = \sum_i |z_i|^2 \|u_i\|^2.$$

In particular, if the $u_i \neq 0$, then $u = 0 \Rightarrow z_i = 0$ for all i. This shows that any set of mutually orthogonal (non-zero) vectors is linearly independent.

The theorem of Apollonius.

Adding the equations

$$\begin{aligned} \|f+g\|^2 &= \|f\|^2 + 2\mathrm{Re}\,(f,g) + \|g\|^2 && (6) \\ \|f-g\|^2 &= \|f\|^2 - 2\mathrm{Re}\,(f,g) + \|g\|^2 && (7) \end{aligned}$$

gives

$$\|f+g\|^2 + \|f-g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right). \qquad (8)$$

This is known as the **parallelogram law**. It is the algebraic expression of the theorem of Apollonius which asserts that the sum of the areas of the squares on the sides of a parallelogram equals the sum of the areas of the squares on the diagonals.

If we subtract (7) from (6) we get

$$\mathrm{Re}(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2\right). \qquad (9)$$

Completing a pre-Hilbert space.

$$\mathrm{Re}(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2\right). \tag{9}$$

Now $(if,g) = i(f,g)$ and $\mathrm{Re}\{i(f,g)\} = -\mathrm{Im}(f,g)$ so

$$\mathrm{Im}(f,g) = -\mathrm{Re}(if,g) = \mathrm{Re}(f,ig)$$

so

$$(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\right). \tag{10}$$

If we now complete a pre-Hilbert space, the right hand side of this equation is defined on the completion, and is a continuous function there. It therefore follows that the scalar product extends to the completion, and, by continuity, satisfies all the axioms for a scalar product, plus the completeness condition for the associated norm. In other words, the completion of a pre-Hilbert space is a Hilbert space.

$$\|f+g\|^2+\|f-g\|^2 = 2\left(\|f\|^2+\|g\|^2\right). \tag{8}$$

$$(f,g) = \frac{1}{4}\left(\|f+g\|^2-\|f-g\|^2+i\|f+ig\|^2-i\|f-ig\|^2\right). \tag{10}$$

The theorem of Jordan and von Neumann.

This is essentially a converse to the theorem of Apollonius. It says that if $\|\cdot\|$ is a norm on a (complex) vector space V which satisfies (8), then V is in fact a pre-Hilbert space with $\|f\|^2 = (f,f)$. If the theorem is true, then the scalar product must be given by (10). So we must prove that if we take (10) as the definition, then all the axioms on a scalar product hold. The easiest axiom to verify is

$$(g,f) = \overline{(f,g)}.$$

Indeed, the real part of the right hand side of (10) is unchanged under the interchange of f and g (since $g-f = -(f-g)$ and $\|-h\| = \|h\|$ for any h is one of the properties of a norm). Also $g+if = i(f-ig)$ and $\|ih\| = \|h\|$ so the last two terms on the right of (10) get interchanged, proving that $(g,f) = \overline{(f,g)}$.

$$(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\right). \quad (10)$$

It is just as easy to prove that

$$(if,g) = i(f,g).$$

Indeed replacing f by if sends $\|f+ig\|^2$ into $\|if+ig\|^2 = \|f+g\|^2$ and sends $\|f+g\|^2$ into $\|if+g\|^2 = \|i(f-ig)\|^2 = \|f-ig\|^2 = i(-i\|f-ig\|^2)$ so has the effect of multiplying the sum of the first and fourth terms by i, and similarly for the sum of the second and third terms on the right hand side of (10).

$$\|f+g\|^2+\|f-g\|^2=2\left(\|f\|^2+\|g\|^2\right). \tag{8}$$

$$\operatorname{Re}(f,g)=\frac{1}{4}\left(\|f+g\|^2-\|f-g\|^2\right). \tag{9}$$

$$(f,g)=\frac{1}{4}\left(\|f+g\|^2-\|f-g\|^2+i\|f+ig\|^2-i\|f-ig\|^2\right). \tag{10}$$

Now (10) implies (9). Suppose we replace f, g in (8) by f_1+g, f_2 and by f_1-g, f_2 and subtract the second equation from the first. We get

$$\|f_1+f_2+g\|^2-\|f_1+f_2-g\|^2+\|f_1-f_2+g\|^2-\|f_1-f_2-g\|^2$$

$$=2\left(\|f_1+g\|^2-\|f_1-g\|^2\right).$$

In view of (9) we can write this as

$$\operatorname{Re}\,(f_1+f_2,g)+\operatorname{Re}\,(f_1-f_2,g)=2\operatorname{Re}\,(f_1,g). \tag{11}$$

Now the right hand side of (9) vanishes when $f=0$ since $\|g\|=\|-g\|$. So if we take $f_1=f_2=f$ in (11) we get

$$\operatorname{Re}\,(2f,g)=2\operatorname{Re}\,(f,g).$$

$$\mathrm{Re}\ (f_1 + f_2, g) + \mathrm{Re}\ (f_1 - f_2, g) = 2\mathrm{Re}\ (f_1, g). \tag{11}$$

$$\mathrm{Re}\ (2f, g) = 2\mathrm{Re}\ (f, g).$$

We can thus write (11) as

$$\mathrm{Re}\ (f_1 + f_2, g) + \mathrm{Re}\ (f_1 - f_2, g) = \mathrm{Re}\ (2f_1, g).$$

In this equation make the substitutions

$$f_1 \mapsto \frac{1}{2}(f_1 + f_2), \quad f_2 \mapsto \frac{1}{2}(f_1 - f_2).$$

This yields

$$\mathrm{Re}\ (f_1 + f_2, g) = \mathrm{Re}\ (f_1, g) + \mathrm{Re}\ (f_2, g).$$

$$\mathrm{Re}(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2\right). \tag{9}$$

$$(f,g) = \frac{1}{4}\left(\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2\right). \tag{10}$$

(10) implies (9).

We have shown that

$$\mathrm{Re}\ (f_1+f_2,g) = \mathrm{Re}\ (f_1,g) + \mathrm{Re}\ (f_2,g).$$

Since it follows from (10) and (9) that

$$(f,g) = \mathrm{Re}\ (f,g) - i\mathrm{Re}\ (if,g)$$

we conclude that

$$(f_1+f_2,g) = (f_1,g) + (f_2,g).$$

Taking $f_1 = -f_2$ shows that

$$(-f,g) = -(f,g).$$

Consider the collection $\mathcal{C}$ of complex numbers α which satisfy

$$(\alpha f, g) = \alpha(f, g)$$

(for all f, g). We know from $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$ that

$$\alpha, \beta \in \mathcal{C} \Rightarrow \alpha + \beta \in \mathcal{C}.$$

So $\mathcal{C}$ contains all integers. If $0 \neq \beta \in \mathcal{C}$ then

$$(f, g) = (\beta \cdot (1/\beta) f, g) = \beta((1/\beta) f, g)$$

so $\beta^{-1} \in \mathcal{C}$. Thus $\mathcal{C}$ contains all (complex) rational numbers. The theorem will be proved if we can prove that $(\alpha f, g)$ is continuous in α.

The triangle inequality

$$\|f+g\| \leq \|f\| + \|g\|$$

applied to $f = f_2, g = f_1 - f_2$ implies that

$$\|f_1\| \leq \|f_1 - f_2\| + \|f_2\|$$

or

$$\|f_1\| - \|f_2\| \leq \|f_1 - f_2\|.$$

Interchanging the role of f_1 and f_2 gives

$$|\ \|f_1\| - \|f_2\|\ | \leq \|f_1 - f_2\|.$$

Therefore

$$|\ \|\alpha f \pm g\| - \|\beta f \pm g\|\ | \leq \|(\alpha - \beta) f\|.$$

Since $\|(\alpha-\beta)f\| \to 0$ as $\alpha \to \beta$ this shows that the right hand side of (10) when applied to αf and g is a continuous function of α. Thus $\mathcal{C} = \mathbf{C}$. We have proved

$$\|f+g\|^2 + \|f-g\|^2 = 2\left(\|f\|^2 + \|g\|^2\right). \tag{8}$$

Theorem 1 [P. Jordan and J. von Neumann] *If V is a normed space whose norm satisfies (8) then V is a pre-Hilbert space.*

Notice that the condition (8) involves only two vectors at a time. So we conclude as an immediate consequence of this theorem that

Corollary 1 *A normed vector space is pre-Hilbert space if and only if every two dimensional subspace is a Hilbert space in the induced norm.*

Actually, a weaker version of this corollary, with two replaced by three had been proved by Fréchet, *Annals of Mathematics*, July 1935, who raised the problem of giving an abstract characterization of those norms on vector spaces which come from scalar products. In the immediately following paper Jordan and von Neumann proved the theorem above leading to the stronger corollary that two dimensions suffice.

Orthogonal subsets.

We continue with the assumption that V is pre-Hilbert space. If A and B are two subsets of V, we write $A \perp B$ if $u \in A$ and $v \in B \Rightarrow u \perp v$, in other words if every element of A is perpendicular to every element of B. Similarly, we will write $v \perp A$ if the element v is perpendicular to all elements of A. Finally, we will write $A^{\perp}$ for the set of all v which satisfy $v \perp A$. Notice that $A^{\perp}$ is always a linear subspace of V, for any A.

Orthogonal projection.

Now let M be a (linear) subspace of V. Let v be some element of V, not necessarily belonging to M. We want to investigate the problem of finding a $w \in M$ such that $(v - w) \perp M$. Of course, if $v \in M$ then the only choice is to take $w = v$. So the interesting problem is when $v \notin M$. Suppose that such a w exists, and let x be any (other) point of M. Then by the Pythagorean theorem,

$$\|v - x\|^2 = \|(v - w) + (w - x)\|^2 = \|v - w\|^2 + \|w - x\|^2$$

since $(v - w) \perp M$ and $(w - x) \in M$. So

$$\|v - w\| \leq \|v - x\|$$

and this inequality is strict if $x \neq w$. In words: if we can find a $w \in M$ such that $(v - w) \perp M$ then w is the unique solution of the problem of finding the point in M which is closest to v.

if we can find a $w \in M$
such that $(v - w) \perp M$ then w is the unique solution of the problem of finding the point in M which is closest to v. Conversely, suppose we found a $w \in M$ which has this minimization property, and let x be any element of M. Then for any real number t we have

$$\|v - w\|^2 \leq \|(v - w) + tx\|^2 = \|v - w\|^2 + 2t\text{Re } (v - w, x) + t^2\|x\|^2.$$

Since the minimum of this quadratic polynomial in t occurring on the right is achieved at $t = 0$, we conclude (by differentiating with respect to t and setting $t = 0$, for example) that

$$\text{Re } (v - w, x) = 0.$$

By our usual trick of replacing x by $e^{i\theta}x$ we conclude that

$$(v - w, x) = 0.$$

Since this holds for all $x \in N$, we conclude that $(v - w) \perp M$. So to find w we search for the minimum of $\|v - x\|, \; x \in M$.

Approaching the minimum.

Now $\|v - x\| \geq 0$ and is some finite number for any $x \in M$. So there will be some real number m such that $m \leq \|v - x\|$ for $x \in M$, and such that no strictly larger real number will have this property. (m is known as the "greatest lower bound" of the values $\|v - x\|,\ x \in M$.) So we can find a sequence of vectors $x_n \in M$ such that

$$\|v - x_n\| \to m.$$

We claim that the x_n form a Cauchy sequence. Indeed,

$$\|x_i - x_j\|^2 = \|(v - x_j) - (v - x_i)\|^2$$

and by the parallelogram law this equals

$$2\left(\|v - x_i\|^2 + \|v - x_j\|^2\right) - \|2v - (x_i + x_j)\|^2.$$

Now the expression in parenthesis converges to $2m^2$. The last term on the right is

$$-\|2(v - \frac{1}{2}(x_i + x_j)\|^2.$$

Since $\frac{1}{2}(x_i + x_j) \in M$, we conclude that

$$\|2v - (x_i + x_j)\|^2 \geq 4m^2$$

so

$$\|x_i - x_j\|^2 \leq 4(m + \epsilon)^2 - 4m^2$$

for i and j large enough that $\|v - x_i\| \leq m + \epsilon$ and $\|v - x_j\| \leq m + \epsilon$. This proves that the sequence x_n is Cauchy.

Using completeness.

Here is the crux of the matter: If M is complete, then we can conclude that the x_n converge to a limit w which is then the unique element in M such that $(v - w) \perp M$. It is at this point that completeness plays such an important role.

Put another way, we can say that if M is a subspace of V which is complete (under the scalar product $(\ ,\ )$ restricted to M) then we have the orthogonal direct sum decomposition

$$V = M \oplus M^{\perp},$$

which says that every element of V can be uniquely decomposed into the sum of an element of M and a vector perpendicular to M.

Fourier coefficients.

For example, if M is the one dimensional subspace consisting of all (complex) multiples of a non-zero vector y, then M is complete, since $\mathbf{C}$ is complete. So w exists. Since all elements of M are of the form ay, we can write $w = ay$ for some complex number a. Then $(v - ay, y) = 0$ or

$$(v, y) = a\|y\|^2$$

so

$$a = \frac{(v, y)}{\|y\|^2}.$$

We call a the **Fourier coefficient** of v with respect to y. Particularly useful is the case where $\|y\| = 1$ and we can write

$$a = (v, y). \tag{12}$$

Getting back to the general case, if $V = M \oplus M^\perp$ holds, so that every v corresponds to a unique $w \in M$ satisfying $(v - w) \in M^\perp$ the map $v \mapsto w$ is called orthogonal projection of V onto M and will be denoted by π_M.

Linear and anti-linear maps.

Let V and W be two complex vector spaces. A map

$$T : V \to W$$

is called **linear** if

$$T(\lambda x + \mu y) = \lambda T(x) + \mu T(Y) \quad \forall\ x, y \in V, \quad \lambda, \mu \in \mathbf{C}$$

and is called **anti-linear** if

$$T(\lambda x + \mu y) = \overline{\lambda} T(x) + \overline{\mu} T(Y) \quad \forall\ x, y \in V \quad \lambda, \mu \in \mathbf{C}.$$

Linear functions.

If $\ell : V \to \mathbf{C}$ is a linear map, (also known as a linear function) then

$$\ker\, \ell := \{x \in V |\, \ell(x) = 0\}$$

has codimension one (unless $\ell \equiv 0$). Indeed, if

$$\ell(y) \neq 0$$

then

$$\ell(x) = 1 \quad \text{where } x = \frac{1}{\ell(y)} y$$

and for any $z \in V$,

$$z - \ell(z)x \in \ker\, \ell.$$

If V is a normed space and ℓ is continuous, then ker (ℓ) is a closed subspace. The space of continuous linear functions is denoted by V^*.

The dual space.

If V is a normed space and ℓ is continuous, then ker (ℓ) is a closed subspace. The space of continuous linear functions is denoted by V^*. It has its own norm defined by

$$\|\ell\| := \sup_{x \in V, \|x\| \neq 0} |\ell(x)|/\|x\|.$$

Suppose that H is a pre-hilbert space. Then we have an antilinear map

$$\phi : H \to H^*, \quad (\phi(g))(f) := (f, g).$$

The Cauchy-Schwarz inequality implies that

$$\|\phi(g)\| \leq \|g\|$$

and in fact

$$(g, g) = \|g\|^2$$

shows that

$$\|\phi(g)\| = \|g\|.$$

In particular the map ϕ is injective.

The Riesz representation theorem.

Suppose that H is a pre-hilbert space. Then we have an antilinear map

$$\phi : H \to H^*, \quad (\phi(g))(f) := (f, g).$$

$$\|\phi(g)\| \leq \|g\|$$

In particular the map ϕ is injective.

The Riesz representation theorem says that if H is a Hilbert space, then this map is surjective:

Theorem 2 *Every continuous linear function on H is given by scalar product by some element of H.*

The proof is a consequence of the theorem about projections applied to

$$N := \ker \ell :$$

If $\ell = 0$ there is nothing to prove. If $\ell \neq 0$ then N is a closed subspace of codimension one. Choose $v \notin N$. Then there is an $x \in N$ with $(v - x) \perp N$. Let

$$y := \frac{1}{\|v - x\|}(v - x).$$

$$y \perp N \quad \text{and} \quad \|y\| = 1.$$

For any $f \in H$,

$$[f - (f, y)y] \perp y$$

so

$$f - (f, y)y \in N$$

or

$$\ell(f) = (f, y)\ell(y),$$

so if we set

$$g := \overline{\ell(y)}y$$

then

$$(f, g) = \ell(f)$$

for all $f \in H$. QED

Frigyes Riesz

Born: **22 Jan 1880 in Györ, Austria-Hungary (now Hungary)**
Died: **28 Feb 1956 in Budapest, Hungary**

What is $L_2(\mathbf{T})$?

We have defined the space $L^2(\mathbf{T})$ to be the completion of the space $\mathcal{C}(\mathbf{T})$ under the L_2 norm $\|f\|_2 = (f, f)^{\frac{1}{2}}$. In particular, every linear function on $\mathcal{C}(\mathbf{T})$ which is continuous with respect to the this L_2 norm extends to a unique continuous linear function on $L_2(\mathbf{T}$. By the Riesz representation theorem we know that every such continuous linear function is give by scalar product by an element of $L_2(T)$. Thus we may think of the elements of $L_2(\mathbf{T})$ as being the linear functions on $\mathcal{C}(\mathbf{T})$ which are continuous with respect to the L_2 norm. An element of $L_2(\mathbf{T})$ should not be thought of as a function, but rather as a linear function on the space of continuous functions relative to a special norm - the L_2 norm.