1 Orthonormal sets in Hilbert space

Let $S \subseteq H$. We denote by $[S]$ the span of S, i.e., the set of all linear combinations of elements from S. A set $\{u_\alpha : \alpha \in A\}$ is called orthonormal, if $\langle u_\alpha, u_\beta \rangle = 0$ for all $\alpha \neq \beta$ and $\|u_\alpha\| = 1$ for all α. (Here A is some index set.)

For every $x \in H$ we define a transform $\widehat{x} : A \to \mathbb{C}$ by $\widehat{x}(\alpha) = \langle x, u_\alpha \rangle$ and call these the Fourier coefficients of x with respect to $\{u_\alpha : \alpha \in A\}$.

Let $F \subseteq A$ be finite and set $M_F = [\{u_\alpha : \alpha \in F\}]$. We observe the following facts.

1. If $\varphi : A \to \mathbb{C}$ with $\varphi(\alpha) = 0$ for $\alpha \notin F$, then $y \in M_F$ defined by

$$y = \sum_{\alpha \in F} \varphi(\alpha) u_\alpha$$

   satisfies $\widehat{y}(\alpha) = \varphi(\alpha)$. Also,

$$\|y\|^2 = \sum_{\alpha \in F} |\varphi(\alpha)|^2.$$

2. If $x \in H$ and s_F is defined by

$$s_{F,x} = \sum_{\alpha \in F} \widehat{x}(\alpha) u_\alpha,$$

   then

$$\|x - s_{F,x}\| < \|x - s\|$$

   for every $s \in M_F$ with $s \neq s_F$. Moreover,

$$\sum_{\alpha \in F} |\widehat{x}(\alpha)|^2 \leq \|x\|^2.$$

The first statement of these follows immediately from the orthogonality conditions. For the second part note that $s_{F,x}$ and x have the same Fourier coefficient for $\alpha \in F$, i.e., $x - s_{F,x} \perp M_F$. Since $s_{F,x} \in M_F$, we obtain

$$x - s_{F,x} \perp s - s_{F,x}$$

for all $s \in M_F$. Hence for $s \in M_F$

$$\|x - s\|^2 = \|x - s_{F,x}\|^2 + \|s_{F,x} - s\|^2$$

and the second term on the right is zero only if $s = s_{F,x}$. The choice $s = 0$ gives the last inequality. (This means in particular that $s_{F,x}$ is the unique best approximation to x in M_F with respect to $\|.\|$)

Example. Rewrite these statements if $H = L^2([0,1])$ and the orthonormal system is given by the exponentials $u_n(t) = e^{2\pi itn}$ where $n \in \mathbb{Z}$.

We would like to remove the finiteness condition from the previous statements. Let A be an arbitrary index set and $0 \leq \varphi(\alpha) \leq \infty$ for every $\alpha \in A$. Then

$$\sum_{\alpha \in A} \varphi(\alpha)$$

is short notation for the supremum of the set of all finite sums $\varphi(\alpha_1) + ... + \varphi(\alpha_n)$ with $\alpha_i \in A$. (In Math 750 terms: the series is the Lebesgue integral of φ with respect to counting measure on A.)

We write $\ell^2(A)$ to indicate the class of functions φ with

$$\sum_{\alpha \in A} |\varphi(\alpha)|^2 < \infty.$$

We note that this is also a Hilbert space with scalar product

$$\langle \varphi, \psi \rangle = \sum_{\alpha \in A} \varphi(\alpha)\overline{\psi(\alpha)}.$$

We note that the simple functions are dense in every L^p space. In particular, the set of functions φ that are zero on all but finitely many elements of A is dense in $\ell^2(A)$. For completeness we include a proof of the density statement.

Theorem 1. *Let μ be a Borel measure, and let S be the class of complex, measurable, simple functions on X so that*

$$\mu(\{x : s(x) \neq 0\}) < \infty.$$

If $1 \leq p < \infty$, then S is dense in $L^p(\mu)$.

Proof. Evidently $S \subseteq L^p(\mu)$. For the other direction, suppose first $f \geq 0$ in $L^p(\mu)$. Let s_n be a sequence of simple functions approximating f from below. Then $s_n \in L^p$ and hence in S. Since $|f - s_n|^p \leq f^p$, dominated convergence shows that the p-norm of the difference goes to zero, and the complex case follows by taking real and imaginary parts, followed by taking positive and negative parts for each. □

Lemma 1. *If $\varphi \in \ell^2(A)$, then $\{\alpha \in A : \varphi(\alpha) = 0\}$ is at most countable.*

Proof. Let $A_n = \{\alpha \in A : |\varphi(\alpha)| \geq 1/n\}$. Then

$$\sum_{\alpha \in A_n} 1 \leq \sum_{\alpha \in A_n} |n\varphi(\alpha)|^2 \leq n^2 \sum_{\alpha \in A} |\varphi(\alpha)|^2$$

and the right side is finite. Hence A_n is a finite set, and the set of values where φ is nonzero is a countable union of finite sets. □

Definition 1. Let (X_1, d_1) and (X_2, d_2) be two metric spaces. A map $F : X_1 \to X_2$ is called an isometry, if

$$d_2(F(a), F(b)) = d_1(a, b)$$

for all $a, b \in X_1$.

The next goal is to show that the map $\mathcal{F} : H \to \ell^2(A)$ defined by $\mathcal{F}(x) = \widehat{x}$ is an isometry from the span of linear combinations of an orthonormal basis $\{u_\alpha\}$ onto $\ell^2(A)$.

Theorem 2. *Let X, Y be two metric spaces where X is complete. Assume*

1. *$f : X \to Y$ is continuous,*
2. *X has a dense subspace X_0 on which f is an isometry,*
3. *$f(X_0)$ is dense in Y.*

Then f is an isometry of X onto Y.

Proof. From continuity of f it is immediately clear that f is an isometry on X. Let $y \in Y$. Let $(x_n) \subseteq X_0$ be a sequence with $f(x_n) \to y$. The sequence $f(x_n)$ is therefore Cauchy, and since f is an isometry on X_0, (x_n) is Cauchy and by completeness of x has a limit. Continuity of f implies $f(x) = y$. □

Theorem 3. *Let $U = \{u_\alpha : \alpha \in A\}$ be an orthonormal set in H, and let P be the space of finite linear combinations of U. Then for every $x \in H$,*

$$\sum_{\alpha \in A} |\widehat{x}(\alpha))|^2 \leq \|x\|^2,$$

and $\mathcal{F} : H \to \ell^2(A)$ defined by $\mathcal{F}(x) = \widehat{x}$ is a continuous linear mapping whose restriction to $\overline{P}$ is an isometry onto $\ell^2(A)$.

Proof. We had seen that the inequality holds for every finite set $F \subseteq A$. Theorem 1 implies that it holds for all $x \in H$. (This is also called Bessel's inequality.)

It follows from this inequality that $\mathcal{F}$ maps H into $\ell^2(A)$. Evidently $\mathcal{F}$ is linear, and an application of Bessel to $x - y$ shows that $\mathcal{F}$ is continuous.

We had seen before that $\mathcal{F}$ is an isometry of P onto the subspace of all elements in $\ell^2(A)$ with finite support. This subspace is dense in $\ell^(A)$ (Theorem 1 again). From Theorem 2 it follows that $\mathcal{F}$ is an isometry of P onto $\ell^2(A)$. (This is also called the Riesz-Fischer theorem.) $\square$

Theorem 4. *Each of the following four conditions on an orthonormal set u_α implies the other three.*

1. *$\{u_\alpha\}$ is a maximal orthonormal set in H,*

2. *The set P of all finite linear combinations of elements from $\{u_\alpha\}$ is dense in H,*

3. *The equality*
$$\sum_{\alpha \in A} |\widehat{x}(\alpha)|^2 = \|x\|^2$$
holds for all $x \in H$,

4. *The equality*
$$\sum_{\alpha \in A} \widehat{x}(\alpha)\overline{\widehat{y}(\alpha)} = \langle x, y \rangle$$
holds for all $x, y \in H$.

By defining the scalar product

$$\langle a, b \rangle_{\ell^2(A)} = \sum_{\alpha \in A} a(\alpha)\overline{b(\alpha)},$$

the last identity can be written as $\langle \widehat{x}, \widehat{y} \rangle_{\ell^2(A)} = \langle x, y \rangle_H$. Maximal orthonormal sets are also called orthonormal bases.

Proof. To say that $\{u_\alpha\}$ is maximal means that no vector from H can be added to this set in such a way that the resulting set is still orthonormal. (See also the current homework.)

Assume that p is not dense in H. Then there exists $x \in H \backslash \overline{P}$. By the theorem about closed subspaces, there exists $y \in \overline{P}^\perp$ of norm 1. This can be added to $\{u_\alpha\}$ to yield a larger orthonormal set.

It follows that (1) impliies (2).

The previous theorem showed that the Fourier transform is an isometry on $\overline{P}$. If this is all of H, then (3) follows.

Polarization:

$$4\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 3 - i\|x-iy\|^2.$$

Hence every norm identity yields a corresponding scalar product identity. In particular, (3) implies (4).

Finally, if (1) does not hold, then there exists $u \in H$ with $\langle u, u_\alpha\rangle = 0$ for all α and $\|u\| = 1$. The evidently every Fourier coefficient of u is zero, hence so is the left side of (4), but the right side with $x = y = u$ is 1. □

2 Application to the triogonmetric system

From the previous theorem we know that in order to prove that the system $\{u_n\}_{n\in\mathbb{Z}}$ with

$$u_n(t) = e^{2\pi int}$$

is an orthonormal basis of $H = L^2([0,1])$, we need to prove that it is dense in H. Recall that X can be the real line or the unit circle. We denote by $C_c(x)$ the continuous functions with compact support in X. One more tool from Math 750:

Theorem 5. *For $1 \le p < \infty$, $C_c(X)$ is dense in $L^p(\mu)$ where μ is a Borel measure on X.*

Proof. Recall that S is the set of complex, measurable, simple functions on X. Lusin's Theorem (Folland, Section 2.4, Exercise 44):

For every $s \in S$ and $\varepsilon > 0$ there exists $g \in C_c(X)$ such that $g = s$ except on a set of measure $< \varepsilon$, and $|g| \le \|s\|_\infty$.

Putting this together leads to

$$\|g - s\|_p \le 2\varepsilon^{1/p}\|s\|_\infty.$$

We had seen that S is dense in $L^p(\mu)$, hence any $f \in L^p(\mu)$ can be approximated by functions in $C_c(X)$. □

For $X = [0,1]$ this means that the continuous functions are dense in $L^2([0,1])$. Thus we need to prove that every continous function can be approximated in arbitarily close in L^2 by trigonometric polynomials. There is a useful connection between $\|.\|_2$ and $\|.\|_\infty$ on compact sets:

$$\left(\int |g|^2 d\mu\right)^{1/2} \leq \|g\|_\infty,$$

hence it suffices to approximate continuous functions by trigonometric polynomials in L^∞ norm.

Goal. Let f be continuous, and let $\varepsilon > 0$. Show that there exists a finite sequence of coefficients $a_n = a_n(\varepsilon)$ so that f_ε defined by

$$f_\varepsilon(t) = \sum_{n=-N}^{N} a_n e^{2\pi i n t}$$

satisfies

$$\sup_{x\in[0,1]} |f(x) - f_\varepsilon(x)| < \varepsilon.$$

We construct f_ε via convolution. Some preliminary ideas: Let $\varphi \in L^2([0,1])$ with $\int \varphi = 1$. Define

$$f_\varphi(x) = \int_0^1 f(x-u)\varphi(u)du.$$

Two crucial identities: First,

$$\widehat{f_\varphi}(n) = \widehat{f}(n)\widehat{\varphi}(n),$$

and secondly

$$f_\varphi(x) - f(x) = \int_0^1 \varphi(u)(f(x) - f(x-u))du.$$

This means that if φ is a trigonometric polynomial of degree N, then automatically f_φ is as well, and if we want to estimate the difference, it is sufficient to estimate the differences under the integral sign!

Assume that we can find a family $\{Q_k\}$ of trigonometric polynomials with the following properties.

(i) $Q_k(x) \geq 0$ for all $x \in \mathbb{R}$,

(ii) $\int_{-1/2}^{1/2} Q_k(x)dx = 1$ for all k,

(iii) If $0 < \delta < 1/2$, then $Q_k(x) \to 0$ uniformly for all $\delta \leq |x| \leq 1/2$ as $k \to \infty$.

We show first that the existence of such a family implies the desired density statement. Let $\varepsilon > 0$. Let $\delta > 0$ so that

$$|x - t| < \delta \text{ implies } |f(x) - f(t)| < \varepsilon \tag{1}$$

for all x and t. (Note that our assumptions imply uniform continuity of f.)

We obtain from property (iii) that there exists k_0 such that for all $k \geq k_0$ and $\delta \leq |x| \leq 1/2$ we have

$$Q_k(x) \leq \frac{\varepsilon}{2\|f\|_\infty}. \tag{2}$$

Define

$$f_k(x) = \int_0^1 f(u) Q_k(x - u) du = \int_{-1/2}^{1/2} f(x - u) Q_k(u) du.$$

We note that f_k is a trigonometric polynomial since Q_k is a finite linear combination of exponentials; plug the corresponding representation of Q_k into the first integral and change summation and integration. (The two representations can be shown to be equal with a change of variable.) Property (ii) implies that

$$f_k(x) - f(x) = \int_{-1/2}^{1/2} (f(x - u) - f(x)) Q_k(u) du.$$

Break the integral into two pieces, one over $|u| \leq \delta$ and the other over $\delta \leq |u| \leq 1/2$. Note that (1) and property (ii) imply

$$\left| \int_{|u| \leq \delta} (f(x - u) - f(x)) Q_k(u) du \right| \leq \varepsilon$$

and that (2) implies

$$\left| \int_{\delta \leq |u| \leq 1/2} (f(x - u) - f(x)) Q_k(u) du \right| \leq 2\|f\|_\infty \int_{\delta \leq |u| \leq 1/2} Q_k(u) du \leq \varepsilon.$$

It remains to show that a family $\{Q_k\}$ with the stated properties exists. We define

$$Q_k(x) = c_k \left(\frac{1 + \cos 2\pi x}{2} \right)^k$$

where c_k is chosen so that $\int_0^1 Q_k(x)dx = 1$. If you have seen Gamma functions,

$$c_k = \frac{k!\sqrt{\pi}}{2\Gamma[k+\frac{1}{2}]},$$

but this is not necessary to know.

Expanding Q_k using the binomial theorem implies that Q_k is a trigonometric polynomial. Evidently $Q_k \geq 0$, and $\int Q_k = 1$ by construction. It remains to show that Q_k goes to zero uniformly away from the origin. We note first that Q_k is decreasing on $[0, 1/2]$. Hence for $\delta > 0$ and $\delta \leq x \leq 1/2$ we have

$$Q_k(x) \leq Q_k(\delta),$$

and if we can show that this value goes to zero, the uniform convergence follows. First, an inequality for c_k: Since Q_k is even, we have

$$1 = 2c_k \int_0^{1/2} \left(\frac{1+\cos(2\pi t)}{2}\right)^k dt > 2c_k \int_0^{1/2} \left(\frac{1+\cos(2\pi t)}{2}\right)^k \sin(2\pi t)dt,$$

note that $\sin(2\pi t) \geq 0$ on $[0, 1/2]$. The integral on the right hand side can be evaluated to $(\pi(k+1))^{-1}$. Hence

$$c_k \leq \frac{\pi(k+1)}{2},$$

and we obtain

$$Q_k(\delta) \leq \frac{\pi(k+1)}{2} \left(\frac{1+\cos(2\pi\delta)}{2}\right)^k.$$

This goes to zero for fixed $0 < \delta < 1/2$ as $k \to \infty$, since it is of the form $C(k+1)\eta^k$ with fixed $C > 0$ and $0 < \eta < 1$.