Integration and Fourier Theory

Lecture 13

Morten Grud Rasmussen

March 22, 2013

1 Orthonormal systems

Definition 1.1 (Linear combination, linear independence, span)**.** Let V be a vector space and $x_1, \ldots, x_k \in V$. Then, for any scalars $c_1, \ldots, c_k$, we call $c_1x_1 + \cdots + c_kx_k$ a *linear combination*. If $c_1x_1 + \cdots + c_kx_k = 0$ implies $c_1 = \cdots = c_k = 0$ (i.e. the only linear combination which equals 0 is the trivial one), then $x_1, \ldots, x_k$ are said to be *(linearly) independent*. Let $S \subset V$ be any set. Then S is said to be (linearly) independent if all finite subsets of S are independent. We let $[S] = \operatorname{span} S$ denote the set of all (finite) linear combinations of members of S.

Remark 1.2. Note that $\operatorname{span} S$ is a subspace. In fact, it is the smallest subspace containing S.

Definition 1.3 (Orthonormal system)**.** Let H be a Hilbert space and A an index set. A set of vectors $\{u_\alpha\}_{\alpha\in A}$ is said to be *orthonormal* (or an *orthonormal system*) if

$$(u_\alpha, u_\beta) = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}.$$

Remark 1.4. The 0 when $\alpha \neq \beta$ is the background for the "ortho-", and the 1 when $\alpha = \beta$ the background for the "-normal" part of the name.

If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal set, we associate to each $x \in H$ a function $\hat{x}\colon A \to \mathbb{C}$ by

$$\hat{x}(\alpha) = (x, u_\alpha).$$

The numbers $\hat{x}(\alpha)$ are sometimes referred to as *Fourier coefficients*.

Theorem 1.5. *Assume that $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal system in the Hilbert space H and that $F \subset A$ is a finite set. Write $M_F = \operatorname{span}\{u_\alpha \mid \alpha \in F\}$.*

(a) *Let $\varphi\colon A \to \mathbb{C}$ have support in F. Then*

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

satisfies that $\hat{y} = \varphi$, i.e. $\hat{y}(\alpha) = \varphi(\alpha)$ for all $\alpha \in A$, and that

$$\|y\|^2 = \sum_{\alpha \in F} |\varphi(\alpha)|^2. \tag{1}$$

(b) *Let $x \in H$ and*

$$s_F(x) = \sum_{\alpha \in F} \hat{x}(\alpha) u_\alpha \,. \tag{2}$$

Then

$$\|x - s_F(x)\| < \|x - s\| \tag{3}$$

for every $s \in M_F$ except $s_F(x)$ and

$$\|s_F(x)\|^2 = \sum_{\alpha \in F} |\hat{x}(\alpha)|^2 \leqslant \|x\|^2 \,. \tag{4}$$

Remark 1.6. The right-hand side of (1) is the ℓ^2-norm of φ. Coincidence? Hardly... The identity (1) is a particular instance of Parseval's identity and the inequality (4) is a special case of the Bessel inequality. The inequality (3) states that the "partial sum" $s_F(x)$ of the "Fourier series" $\sum \hat{x}(\alpha) u_\alpha$ is the best approximation to x in M_F relative to the metric defined by the Hilbert space norm.

Proof of Theorem 1.5. Part (a) follows trivially from the assumptions:

$$\|y\|^2 = (y, y) = \left(\sum_{\alpha \in F} \varphi(\alpha) u_\alpha, \sum_{\alpha \in F} \varphi(\alpha) u_\alpha \right) = \sum_{\alpha \in F} |\varphi(\alpha)|^2 \,. \tag{5}$$

To see (b), write s_F for $s_F(x)$ and note that $\hat{s}_F(\alpha) = \hat{x}(\alpha)$ for all $\alpha \in F$. Hence $s_F - x \perp u_\alpha$ for all $\alpha \in F$ which implies that $s_F - x \perp s$ for all $s \in M_F$. Therefore

$$\|x - s\|^2 = \|(x - s_F) - (s - s_F)\|^2 = \|x - s_F\|^2 + \|s - s_F\|^2 \,, \tag{6}$$

which gives (3) and, for $s = 0$, (4). $\square$

2 The Bessel inequality and Parceval's identity

Given Remark 1.6, it should not come as a surprise that the topic of this section is to extend Theorem 1.5. In fact, we will show that Theorem 1.5 also holds, even for uncountable F.

Obviously, this means that we need to make sense of sums such as $\sum_{\alpha \in A} \varphi(\alpha)$ for any set A:

Definition 2.1. Let $\varphi\colon A \to [0, \infty]$. Then we define

$$\sum_{\alpha \in A} \varphi(\alpha) = \sup\{\varphi(\alpha_1) + \cdots + \varphi(\alpha_k) \mid k \in \mathbb{N}, \alpha_i \in A, i = 1, \ldots, k, \alpha_i \neq \alpha_j \text{ for } i \neq j\} \,, \tag{7}$$

i.e. the symbol $\sum_{\alpha \in A} \varphi(\alpha)$ denotes the supremum of all finite sums $\varphi(\alpha_1) + \cdots + \varphi(\alpha_k)$, where k runs over all natural numbers and α_i, $i = 1, \ldots, k$, all distinct values in A.

Remark 2.2. Note that (7) is in fact just the Lebesgue integral of $\varphi\colon A \to [0,\infty]$ relative to the counting measure on A. [If this is not clear to you, it is worthwhile spending some time on your own with paper and a pencil *making* it clear to yourself.] This means that $\sum_{\alpha\in A}\varphi(\alpha) < \infty$ exactly when $\varphi \in \ell^1(A)$, or, more generally, $\sum_{\alpha\in A}|\varphi(\alpha)|^p < \infty$ if and only if $\varphi \in \ell^p(A)$, $0 < p < \infty$.

In today's self study session, you saw that $L^2(\mu)$ is a Hilbert space for any measure μ when the inner product is given by $(f,g) = \int f\bar{g}\,\mathrm{d}\mu$ ([Rudin, Example 4.5(b)]). In particular, $\ell^2(A)$ is a Hilbert space with the inner product

$$(\varphi,\psi) = \sum_{\alpha\in A}\varphi(\alpha)\bar{\psi}(\alpha)\,.$$

Note that $\varphi\bar{\psi} \in \ell^1(A)$ when $\varphi,\psi \in \ell^2(A)$ by Hölder's inequality.

We recall the following theorem from this morning's self study session ([Rudin, Theorem 3.13]):

Theorem 2.3. *Let S be the class of all complex, measurable, simple functions on X such that*

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

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

In the present context, this means that *the functions that are zero except on some finite subset of A are dense in $\ell^p(A)$*, in particular in $\ell^2(A)$. Moreover, if $\varphi \in \ell^p(A)$, then the set $\{\alpha \in A \mid \varphi(\alpha) \neq 0\}$ *is at most countable*. To see this, note that if $A_n = \{\alpha \in A \mid |\varphi(\alpha)| > \frac{1}{n}\}$, then the number of elements of A_n is

$$\#A_n = \sum_{\alpha\in A_n} 1 < \sum_{\alpha\in A_n} n^p|\varphi(\alpha)|^p \leqslant n^p\sum_{\alpha\in A}|\varphi(\alpha)|^p < \infty\,,$$

i.e. there are only finitely many α for which $|\varphi(\alpha)| > \frac{1}{n}$. As the set of α with nonzero $\varphi(\alpha)$ is

$$\{\alpha \in A \mid \varphi(\alpha) \neq 0\} = \bigcup_{n=1}^{\infty} A_n\,,$$

a countable union of finite sets, hence it is at most countable.

Remark 2.4. The situation is quite different in $\ell^\infty(A)$: If A is uncountable (e.g. $A = \mathbb{R}$ equipped with the counting measure), then $\varphi\colon A \to \mathbb{C}$ given by $\varphi(\alpha) = 1$ for all $\alpha \in A$ is measurable and (essentially) bounded, so $\varphi \in \ell^\infty(A)$, but $\{\alpha \in A \mid \varphi(\alpha) \neq 0\} = A$, which was assumed to be uncountable.

Before proceeding to the proof of the Bessel inequality and Parseval's identity, we prove a lemma which will make it easy to pass from the finite to the infinite setup.

Lemma 2.5. *Assume that the following holds:*

(a) X and Y are metric spaces, X is complete,

(b) $f\colon X \to Y$ is continuous,

(c) X has a dense subset X_0 on which f is an isometry, and

(d) $f(X_0)$ is dense in Y.

Then f is an isometry of X onto Y.

Recall that an isometry f between two metric spaces X and Y is a function which satisfies $d_X(x_1, x_2) = d_Y(f(x_1), f(x_2))$, where d_X and d_Y are the metrics on X and Y, respectively, i.e. an isometry is a function which preserves distances.

Remark 2.6. The most important part of the conclusion of the lemma is the fact that f is *onto* Y. Note that the conclusion also implies that Y is complete, meaning that assumptions (a)–(d) cannot all be satisfied if Y is not complete.

Proof of Lemma 2.5. As X_0 is dense in X and f is an isometry on X_0 and continuous everywhere, it is an isometry everywhere: $d_X(x_n, y_n) = d_Y(f(x_n), f(y_n))$ for all $x_n, y_n \in X_0$ and continuity implies that $d_X(x, y) = d_Y(f(x), f(y))$ for all $x, y \in X$.

Let $y \in Y$. Since $f(X_0)$ is dense in Y, one can find a sequence $\{x_n\}_{n=1}^{\infty}$ in X_0 such that $f(x_n) \to y$ when $n \to \infty$. Hence $\{f(x_n)\}_{n=1}^{\infty}$ is a Cauchy sequence in Y, and since f is an isometry on X_0, $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in X. X is complete, so $x_n \to x$ for some $x \in X$. But by continuity of f, this implies that $f(x) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = y$. □

Theorem 2.7. *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal set in a Hilbert space H and let $P = \operatorname{span}\{u_\alpha\}_{\alpha\in A}$, the set of all (finite) linear combinations of elements of $\{u_\alpha\}_{\alpha\in A}$.*

The inequality

$$\sum_{\alpha\in A} |\hat{x}(\alpha)|^2 \leqslant \|x\|^2 \tag{8}$$

holds for every $x \in H$. The mapping $H \ni x \mapsto \hat{x} \in \ell^2(A)$ is continuous, linear and onto, and its restriction to $\bar{P}$, the closure of P, is an isometry of $\bar{P}$ onto $\ell^2(A)$.

Remark 2.8. The inequality (8) is known as the *Bessel inequality*, and the "onto" part of the theorem is known as the *Riesz-Fischer theorem*.

Proof of Theorem 2.7. The Bessel inequality (8) follows from (4) and the definition of infinite sums.

In the following, the function $H \ni x \mapsto \hat{x}$ is called f, i.e. $f(x) = \hat{x}$. Then the Bessel inequality shows that f maps into $\ell^2(A)$. The linearity of f is obvious. If we apply (8) to $x - y$ then

$$\|f(x) - f(y)\|_2 = \|\hat{x} - \hat{y}\|_2 \leqslant \|x - y\|,$$

which shows that f is continuous. Theorem 1.5(a) shows that f is an isometry from P onto the dense subset of $\ell^2(A)$ consisting of those functions whose support is a finite set $F \subset A$. Since H is complete, $\bar{P}$ is complete, and we conclude the proof by applying Lemma 2.5 with $X = \bar{P}$, $X_0 = P$, $Y = \ell^2(A)$ and $f = f$. □

Definition 2.9. An orthonormal set $\{u_\alpha\}_{\alpha\in A}$ is said to be a *maximal* orthonormal set, a *complete* orthonormal set or an orthonormal *basis* if one cannot add a vector to the set $\{u_\alpha\}_{\alpha\in A}$ and still have orthonormality.

Theorem 2.10. *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal set. The following are equivalent:*

(i) The set $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal basis.

(ii) The span of the set $\{u_\alpha\}_{\alpha\in A}$, $P = \operatorname{span}\{u_\alpha\}_{\alpha\in A}$, is dense in H.

(iii) The equality

$$\sum_{\alpha \in A} \|\hat{x}(\alpha)\|^2 = \|x\|^2 \tag{9}$$

holds for every $x \in H$.

(iv) The equality

$$\sum_{\alpha \in A} \hat{x}(\alpha)\overline{\hat{y}(\alpha)} = (x, y) \tag{10}$$

holds for all $x, y \in H$.

Remark 2.11. Both (9) and (10) are referred to as Parseval's identity. As we will soon see, whether we choose one or the other is not very important.

Proof. We will prove the equivalence through the following cycle: (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) ⇒ (i).

Assume (i). If P is not dense, then $\bar{P}$ is not equal to H and the corollary of [Rudin, Theorem 4.11] implies that there exists a nonzero vector in $P^{\perp}$. But this violates (i), so P must be dense.

Assume (ii). Then Theorem 2.7 tell us that $f(x) = \hat{x}$ is an isometry from $\bar{P} = H$ onto $\ell^2(A)$. The left-hand side of (9) is exactly the ℓ^2-norm of $f(x)$ which equals the right-hand side because f is an isometry.

Assume (iii). The left-hand side of (10) is the inner product in $\ell^2(A)$, while the right-hand side is the inner product in H. This means that if we can express inner products in terms of Hilbert space norms, (10) follows from (9). But it is easy to verify that

$$4(x, y) = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 . \tag{11}$$

Assume (iv). If (i) fails, then there exists a nonzero $u \in H$ with $(u, u_\alpha) = 0$ for all α. Put $x = y = u$. Then the left-hand side of (10) is 0, while the right-hand side is $\|u\|^2 \neq 0$. This contradicts (iv). □

Remark 2.12. Note that the polarization identity (11) in fact tells us that we only need to know the norm of a Hilbert space to know the inner product, or, in other words, when you know the inner product on the "diagonal", you know it everywhere. This important fact relies heavily on the fact that we are working in a *complex* Hilbert space. One can also define *real* Hilbert spaces, but in real Hilbert spaces, no such identity holds true.