Functional analysis

Lecture 9: The best approximation property of Fourier coefficients; Parseval's identity; the Riesz–Fischer theorem; Variation of complex measures

Proposition 5.11. *Every separable infinite-dimensional Hilbert space $\mathcal{H}$ has an orthonormal basis $(e_n)_{n=1}^{\infty}$. Each $x \in \mathcal{H}$ can be expressed as a norm convergent series*

$$x = \sum_{n=1}^{\infty} \widehat{x}(n) e_n, \tag{5.1}$$

where $\widehat{x}(n) = (x, e_n)$ is the n^{th} Fourier coefficient of x. Moreover, for each $N \in \mathbb{N}$ the Fourier coefficients give the best approximation of x by an element of $\operatorname{lin}\{e_1, \dots, e_N\}$, that is, we have

$$\operatorname{dist}\big(x, \operatorname{lin}\{e_1, \dots, e_N\}\big) = \Big\| x - \sum_{j=1}^{N} \widehat{x}(j) e_j \Big\|. \tag{5.2}$$

Proof. By Proposition 5.9, $\mathcal{H}$ has an orthonormal basis which must be infinite as $\mathcal{H}$ is infinite-dimensional. If $\{u_\alpha \colon \alpha \in A\} \subset \mathcal{H}$ is an orthonormal set, then $\|u_\alpha - u_\beta\| = \sqrt{2}$ for all $\alpha \neq \beta$. It follows that every orthonormal set in $\mathcal{H}$ must be countable, as otherwise we would have an uncountable separated subsets which contradicts $\mathcal{H}$ being separable. Hence, every orthonormal basis forms a sequence of elements of $\mathcal{H}$, and we denote one of them as $(e_n)_{n=1}^{\infty}$.

Now, we want to find linear combinations which solve the 'best approximation problem' in $\operatorname{lin}\{e_1, \dots, e_N\}$. Fix an arbitrary scalar sequence $(c_j)_{j=1}^{N}$. Notice that for each $1 \leq j \leq N$ we have

$$|c_j - (x, e_j)|^2 = (c_j - (x, e_j))\overline{(c_j - (x, e_j))} = |c_j|^2 - c_j(x, e_j) - c_j(e_j, x) + |(x, e_j)|^2$$

and hence

$$\begin{aligned} \Big\| x - \sum_{j=1}^{N} c_j e_j \Big\|^2 &= \Big(x - \sum_{j=1}^{N} c_j e_j, x - \sum_{j=1}^{N} c_j e_j \Big) \\ &= \|x\|^2 - \sum_{j=1}^{N} c_j(x, e_j) - \sum_{j=1}^{N} c_j(e_j, x) + \sum_{j=1}^{N} |c_j|^2 \qquad (5.3) \\ &= \|x\|^2 + \sum_{j=1}^{N} |c_j - (x, e_j)|^2 - \sum_{j=1}^{N} |(x, e_j)|^2. \end{aligned}$$

Since the first and last summands are constant, the distance between x and $\sum_{j=1}^{N} c_j e_j$ is minimized exactly when the middle sum of squares vanishes, i.e. $c_j = (x, e_j) = \widehat{x}(j)$ for every $1 \leq j \leq N$. This proves formula (5.2).

To see that (5.1) holds true, observe first that

$$\overline{\operatorname{lin}}\{e_n \colon n \in \mathbb{N}\} = \mathcal{H},$$

as otherwise, by Corollary 5.6, the orthogonal complement of the closed linear subspace at the left-hand side would be nonempty and then $(e_n)_{n=1}^{\infty}$ would not be a maximal orthonormal set. Fix any $\varepsilon > 0$ and, using the preceding observation, pick $n_0 \in \mathbb{N}$ such that

$$\operatorname{dist}\big(x, \operatorname{lin}\{e_1, \ldots, e_{n_0}\}\big) < \varepsilon.$$

Then, for every $n \geq n_0$ we have

$$\begin{aligned} \Big\| x - \sum_{j=1}^{n} \widehat{x}(j) e_j \Big\| &= \operatorname{dist}\big(x, \operatorname{lin}\{e_1, \ldots, e_n\}\big) \\ &\leq \operatorname{dist}\big(x, \operatorname{lin}\{e_1, \ldots, e_{n_0}\}\big) < \varepsilon. \end{aligned}$$

□

Remark 5.12. **(a)** It is worth pointing out that Fourier coefficients satisfy the following consistency phenomenon. Namely, if $(e_j)_{j=1}^{N}$ is an orthonormal set and $n < N$, then for any $x \in \mathcal{H}$ the Fourier coefficients of x with respect to $\operatorname{lin}\{e_1, \ldots, e_N\}$ coincide with the first n Fourier coefficients corresponding to $\operatorname{lin}\{e_1, \ldots, e_N\}$. This is absolutely obvious from the defintion of $\widehat{x}(j)$, but it means that the solution of the best approximation problem for the subspace $\operatorname{lin}\{e_1, \ldots, e_N\}$ is an 'extension' of the linear combination that gives the solution for any smaller subspace $\operatorname{lin}\{e_1, \ldots, e_n\}$. In other words, the optimal coefficients for initial indices do not change when passing to a large subspace. From this point of view, such a phenomenon is not obvious at all.

(b) Suppose $\{u_\alpha \colon \alpha \in A\}$ is any orthonormal set in $\mathcal{H}$ and let $\{\alpha_1, \ldots, \alpha_k\}$ be any finite subset of A. Define $\delta = \operatorname{dist}\big(x, \operatorname{lin}\{u_{\alpha_1}, \ldots, u_{\alpha_k}\}\big)$ which, as we know from formula (5.2), equals $\|x - \sum_{j=1}^{k} \widehat{x}(\alpha_j) u_{\alpha_j}\|$. Calculation (5.3) shows that

$$\sum_{j=1}^{k} |\widehat{x}(\alpha_j)|^2 = \|x\|^2 - \delta^2.$$

(c) In view of the above formula, we infer that for every orthonormal set $\{u_\alpha \colon \alpha \in A\}$ and each $x \in \mathcal{H}$ we have

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

the so-called *Bessel inequality.* It follows that the sum at the left-hand side is always finite, thus there are at most countably many nonzero Fourier coefficients of x, no matter how large the given orthonormal set is.

The following important theorem shows that equality in Bessel's inequality is one of criterions for an orthonormal set to be a basis of a Hilbert space.

Theorem 5.13. *Let $\mathcal{H}$ be a Hilbert space and $\{u_\alpha \colon \alpha \in A\}$ be an orthonormal set. Then, the following assertions are equivalent:*

(i) *$\{u_\alpha \colon \alpha \in A\}$ is maximal, i.e. it is an orthonormal basis;*

(ii) *$\overline{\operatorname{lin}}\{u_\alpha \colon \alpha \in A\} = \mathcal{H}$, i.e. $\{u_\alpha \colon \alpha \in A\}$ is linearly dense;*

(iii) *for every $x \in \mathcal{H}$ we have $\|x\|^2 = \sum_{\alpha \in A} |\widehat{x}(\alpha)|^2$;*

(iv) *for all $x, y \in \mathcal{H}$ we have $(x, y) = \sum_{\alpha \in A} \widehat{x}(\alpha)\overline{\widehat{y}(\alpha)}$ (the Parseval identity).*

Proof. (i) ⇒ (ii): As we already observed, if $\{u_\alpha \colon \alpha \in A\}$ was not linearly dense, then its orthogonal complement would be nonempty due to Corollary 5.6. But then it could not be a basis of $\mathcal{H}$.

(ii) ⇒ (iii): In view of Remark 5.12(b), for arbitrarily small $\delta > 0$ we can find a finite set $\{\alpha_1, \dots, \alpha_k\}$ with

$$0 \leq \|x\|^2 - \sum_{j=1}^{k} |\widehat{x}(\alpha_j)|^2 < \delta$$

which shows that equality in (iii) holds true.

(iii) ⇒ (iv): Note that, by Bessel's inequality, $\widehat{x} \in \ell_2(A)$ for every $x \in \mathcal{H}$. Then, equality in (iii) can be rewritten in the form $(x, x) = (\widehat{x}, \widehat{x})$, where at the right-hand side we have the inner product in $\ell_2(A)$. Fix $x, y \in \mathcal{H}$ and consider any $\lambda \in \mathbb{K}$. Note that $(x + \lambda y, x + \lambda y) = (\widehat{x} + \lambda\widehat{y}, \widehat{x} + \lambda\widehat{y})$, which yields

$$\lambda\overline{(x, y)} + \overline{\lambda}(x, y) = \lambda\overline{(\widehat{x}, \widehat{y})} + \overline{\lambda}(\widehat{x}, \widehat{y}).$$

Putting $\lambda = 1$ and $\lambda = \mathrm{i}$ we get that the real and imaginary parts of (x, y) and $(\widehat{x}, \widehat{y})$ are equal, which proves the Parseval identity.

(iv) ⇒ (i): If (i) fails to hold, we can pick a nonzero vector $x \perp \{u_\alpha \colon \alpha \in A\}$. Taking $y = x$ in the Parseval identity we obtain $(x, x) = 0$; a contradiction. □

The next theorem is the last step towards a fundamental result on classification of Hilbert spaces.

Theorem 5.14 (Riesz–Fischer theorem). *Let $\mathcal{H}$ be a Hilbert space and $\{u_\alpha \colon \alpha \in A\}$ be an orthonormal set. Then, for every $\varphi \in \ell_2(A)$ there exists $x \in \mathcal{H}$ such that $\widehat{x} = \varphi$.*

Proof. Given any $\varphi \in \ell_2(A)$, consider the sets $A_n = \{\alpha \in A \colon |\varphi(\alpha)| > \frac{1}{n}\}$ $(n \in \mathbb{N})$. Observe that each A_n is finite, hence it corresponds to a vector

$$x_n = \sum_{\alpha \in A_n} \varphi(\alpha) u_\alpha \quad \text{such that} \quad \widehat{x}_n = \varphi \cdot \mathbb{1}_{A_n}.$$

Notice that:

- $\lim_{n\to\infty} \widehat{x}_n(\alpha) = \varphi(\alpha)$ for each $\alpha \in A$;
- $\|\widehat{x}_n - \varphi\|_2 \leq \|\varphi\|_2$ for each $n \in \mathbb{N}$.

Therefore, applying the Lebesgue theorem on dominated convergence (to the set A with the counting measure) we obtain $\lim_{n\to\infty} \|\widehat{x}_n - \varphi\|_2 = 0$. Hence, $(\widehat{x}_n)_{n=1}^\infty$ is a Cauchy sequence in $\ell_2(A)$. Since for all $m, n \in \mathbb{N}$ we have $\|x_m - x_n\| = \|\widehat{x}_m - \widehat{x}_n\|_2$, the sequence $(x_n)_{n=1}^\infty$ is a Cauchy sequence in $\mathcal{H}$. So, let $x = \lim_{n\to\infty} x_n$ and note that for every $\alpha \in A$ we have

$$\widehat{x}(\alpha) = (x, u_\alpha) = \lim_{n\to\infty} (x_n, u_\alpha) = \lim_{n\to\infty} \widehat{x}_n(\alpha) = \varphi(\alpha),$$

which shows that $\widehat{x} = \varphi$. □

Now, we can summarize what follows from Proposition 5.9 and Theorems 5.13, 5.14. Namely, let $\mathcal{H}$ be a Hilbert space and $\{u_\alpha\colon \alpha\in A\}$ be an orthonormal basis, so the cardinality of A equals the density of $\mathcal{H}$. Consider a map $\Phi\colon \mathcal{H}\to\ell_2(A)$ defined by $\Phi(x)=\widehat{x}$. It is obviously linear and by Parseval's identity, Φ is an isometry, so in particular, it is one-to-one. In view of the Riesz–Fischer theorem, Φ is surjective, and hence it is an isometric isomorphism. Consequently, we have

$$\mathcal{H}\cong\ell_2(A).$$

6 Complex measures and the Radon–Nikodym theorem

Henceforth, by a complex measure we understand a σ-additive set function μ defined on a σ-algebra and taking values in $\mathbb{C}$. Recall that we defined the variation of μ by the formula

$$|\mu|(E)=\sup\Big\{\sum_{i=1}^n|\mu(E_i)|\colon (E_1,\ldots,E_n)\in\Pi(E)\Big\},$$

where $\Pi(E)$ stands for the family of all measurable finite partitions of E (see Definition 3.20). That $|\mu|$ is a σ-additive measure is not excessively surprising, however, that this measure is finite is a much more interesting fact.

Theorem 6.1. *Let $\mu\colon\mathfrak{M}\to\mathbb{C}$ be a complex measure defined on a σ-algebra $\mathfrak{M}$ of subsets of X. Then:*

(a) *$|\mu|$ is σ-additive;*
(b) *$|\mu|(X)<\infty$.*

It means that $|\mu|$ is a finite positive measure.

Proof of (a). Fix a sequence $(E_n)_{n=1}^\infty$ of pairwise disjoint sets from $\mathfrak{M}$, let $E=\bigcup_{n=1}^\infty E_n$ and let $(A_i)_{i=1}^m\in\Pi(E)$. We have

$$\begin{aligned}\sum_{i=1}^m|\mu(A_i)|&=\sum_{i=1}^m\Big|\sum_{n=1}^\infty\mu(A_i\cap E_n)\Big|\leq\sum_{i=1}^m\sum_{n=1}^\infty|\mu(A_i\cap E_n)|\\&=\sum_{n=1}^\infty\sum_{i=1}^m|\mu(A_i\cap E_n)|\leq\sum_{n=1}^\infty|\mu|(E_n),\end{aligned}$$

where the last inequality follows from the fact that $(A_i\cap E_n)_{i=1}^m\in\Pi(E_n)$ for each $n\in\mathbb{N}$. Since the partition $(A_i)_{i=1}^m$ of E was arbitrary, we infer that

$$|\mu|(E)\leq\sum_{n=1}^\infty|\mu|(E_n).$$

For the reverse inequality, fix any numbers $t_n<|\mu|(E_n)$, so that there exist partitions

$$(A_{n,i})_{i=1}^{m_n}\in\Pi(E_n)\quad\text{satisfying}\quad\sum_{i=1}^{m_n}|\mu(A_{n,i})|>t_n\quad\text{for each } n\in\mathbb{N}.$$

Since for every $N \in \mathbb{N}$ the family $(A_{n,i}\colon 1 \le n \le N,\ 1 \le i \le m_n)$ is a measurable partition of a subset of E, we have

$$\begin{aligned}\sum_{n=1}^{\infty} t_n &\le \sum_{n=1}^{\infty}\sum_{i=1}^{m_n} |\mu(A_{n,i})| \\ &= \lim_{N\to\infty} \sum_{n=1}^{N}\sum_{i=1}^{m_n} |\mu(A_{n,i})| \le |\mu|(E).\end{aligned}$$

Passing to supremum with t_n's, we obtain

$$|\mu|(E) \ge \sum_{n=1}^{\infty} |\mu|(E_n)$$

which completes the proof of assertion (a). □

In order to prove assertion (b) we need the following geometric lemma.

Lemma 6.2. *Let $n \in \mathbb{N}$ and $z_1, \dots, z_n \in \mathbb{C}$. There exists a subset $S \subseteq \{1, \dots, n\}$ such that*

$$\Big|\sum_{j\in S} z_j\Big| \ge \frac{1}{6}\sum_{j=1}^{n} |z_j|.$$

Proof. Define $w = \sum_{j=1}^{n} |z_j|$. The complex plane can be decomposed into four quadrants bounded by half-lines given by the equation $\operatorname{Im} z = \pm \operatorname{Re} z$. There exists at least one quadrant Q such that

$$\sum_{\{j\colon z_j\in Q\}} |z_j| \ge \frac{w}{4}$$

With no loss of generality, we can assume that Q is the one given by $\operatorname{Re} z \ge |\operatorname{Im} z|$ and let $S = \{1 \le j \le n\colon z_j \in Q\}$. Notice that every complex number $z \in Q$ satisfies $\operatorname{Re} z \ge |z|/\sqrt{2}$. Therefore,

$$\Big|\sum_{j\in S} z_j\Big| \ge \sum_{j\in S} \operatorname{Re} z_j \ge \frac{1}{\sqrt{2}} \sum_{j\in S} |z_j| \ge \frac{w}{4\sqrt{2}} \ge \frac{w}{6}.$$ □