Instytut Matematyki i Informatyki, Politechnika Wrocławska, 50-370 Wrocław
Tomasz Downarowicz

Wrocław, 18 maja 2011

Analiza Funkcjonalna WPPT IIr.

WYKŁAD 7: Przestrzeń Sprzężona

Zaczynamy od przypomnienia definicji

Definicja. Zbiór wszystkich funcjonałów na przestrzeni unormowanej V stanowi przestrzeń linową unormowaną (normą operatorową). Ponieważ $\mathbb{R}$ i $\mathbb{C}$ są zupełne, przestrzeń funkcjonałów jest przestrzenią Banacha (niezależnie od tego, czy V jest zupełna czy nie). Nazywamy ją *przestrzenią dualną do* V i oznaczmy V^*.

Twierdzenie (Riesza). Jeśli H jest przestrzenią Hilberta, to $H^* \approx H$, gdzie $\approx$ oznacza izometryczny izomorfizm.

Dowód: Najpierw wskażemy, że pod $\approx$ można wstawić izometryczny antyizomorfizm, to znaczy bijektywną izometrię, która jest addytywna i antyjednorodna, tzn. spełnia $\Phi(\alpha x) = \bar{\alpha}\Phi(x)$. W przestrzeniach RZECZYWISTYCH, to załatwia sprawę, gdyż wtedy $\bar{\alpha} = \alpha$ i Φ jest izomorfizmem. Nad przypadkiem zespolonym trzeba będzie popracować trochę więcej.

Nasze Φ będzie określone jako bijekcja z H do H^* wzorem $\Phi(x) = \langle\cdot|x\rangle$ Wyrażenie po prawej (jako funkcja "kropki" pod którą ukryta jest zmienna przebiegająca H) jest oczywiście funkcjonałem na H, co wynika z własności iloczynu skalarnego traktowanego jako funkcja pierwszej zmiennej, i jego norma wynosi $\|x\|$, co wynika z nierówności Schwarza: $|\langle y|x\rangle| \le \|y\|\|x\|$ oraz z tego, że na $y = x$ funcjonał ten osiąga taką właśnie normę: $|\langle x|x\rangle| = \|x\|\|x\|$.

To, że $\Phi(x + y) = \Phi(x) + \Phi(y)$ oraz $\Phi(\alpha x) = \bar{\alpha}\Phi(x)$ wynika natychmiast z własności iloczynu skalarnego, traktowanego jako funkcja drugiej zmiennej.

Czyli Φ jest injekcją, izometrią i antyizomorfizmem. Pozostało sprawdzić, że jest surjekcją. Innymi słowy, mamy pokazać, że dowolny funkcjonał ciągły f jest postaci $\Phi(x)$ dla pewnego $x \in H$.

W tym celu rozważmy jądro H_0 funkcjonału f. Jest to podprzestrzeń liniowa domknięta. Jeśli $H_0 = H$, to $f \equiv 0$ i wtedy $f = \Phi(0)$. W przeciwnym wypadku przestrzeń $H_0^\perp$ zawiera chociaż jeden wektor niezerowy x_0. Najpierw pokażemy, że $H_0^\perp$ nie zawiera dwóch wektorów niezależnych liniowo. Gdyby tak było (x_0 i x_1 byłyby takimi wektorami) to $f(x_0) \ne 0$ i $f(x_1) \ne 0$, więc można napisać kombinację liniową $y = \frac{1}{f(x_0)}x_0 - \frac{1}{f(x_1)}x_1$. Wektor ten należy do $H_0^\perp$. Z drugiej strony $f(y) = 0$, co oznacza, że $y \in H_0$, a to oznacza, że $y = 0$. Ale to przeczy niezależności liniowej x_0 i x_1, bo współczynniki kombinacji definiującej y są niezerowe.

Udowodniliśmy, że $H_0^\perp$ jest jednowymiarowa, czyli postaci $\{\alpha x_0 : \alpha \in \mathbb{K}\}$. Niech $x = \frac{\overline{f(x_0)}}{\|x_0\|^2}x_0$. Sprawdźmy, czy $f(y) = \langle y|x\rangle$ dla każdego $y \in H$. Taki y rozkłada się na $y_0 + y_1$, gdzie $y_0 \in H_0$ oraz $y_1 \in H_0^\perp$ (a wtedy $y_1 = \alpha x_0$). Wtedy mamy

$$f(y) = f(y_0) + f(y_1) = 0 + \alpha f(x_0).$$

Z drugiej strony,

$$\langle y|x\rangle = \langle y_0|x\rangle + \langle y_1|x\rangle = 0 + \alpha\frac{f(x_0)}{\|x_0\|^2}\langle x_0|x_0\rangle = \alpha f(x_0), \text{ czyli to samo!} \quad \square$$

W przypadku zespolonym można zadowolić się antyizomorfizmem, ale jeśli komuś bardzo zależy na izomorfiźmie, to trzeba zmienić odwzorowanie Φ. Zrobimy to w przypadku ośrodkowym, (a w nieośrodkowym jako dodatek nieobowiązkowy).

W ośrodkowej przestrzeni Hilberta istnieje baza ortonormalna $\{e_1, e_2, \dots\}$. Każdy element rozwija się w szereg $x = \sum_{n=1}^{\infty} x_n e_n$ (dla przypomnienia, $x_n = \langle x|e_n\rangle$ są współczynnikami Fouriera, ale nie będziemy z tego korzystać). Do każdego elementu x możemy teraz zdefiniować "element sprzężony" $\bar{x}$: jeśli $x = \sum_{n=1}^{\infty} x_n e_n$ to $\bar{x} = \sum_{n=1}^{\infty} \bar{x}_n e_n$ (definicja takiego "sprzężenia" zależy od wyboru bazy, dlatego nie jest to procedura jednoznaczna).

Teraz możemy zdefiniować $\Psi(x) = \Phi(\bar{x})$, to znaczy Ψ przyporządkowuje elementowi x funkcjonał $\langle\cdot|\bar{x}\rangle$. Sprawdzenie, że to jest izometryczna bijekcja addytywna jest natychmiastowe: takie jest Φ i takie jest nasze "sprzężenie" $x \mapsto \bar{x}$ (a Ψ jest złożeniem tych dwóch). Łatwo sprawdzić, że "sprzężenie" jest antyjdnorodne: jeśli $x = \sum_{n=1}^{\infty} x_n e_n$ to $\alpha x = \sum_{n=1}^{\infty} \alpha x_n e_n$ i wtedy $\overline{\alpha x} = \sum_{n=1}^{\infty} \overline{\alpha x_n} e_n = \bar{\alpha}\bar{x}$. Teraz Ψ, jako złożenie dwóch funkcji antyjednorodnych jest jednorodne. Koniec.

A teraz nieobowiązkowa część o nieośrodkowych przestrzeniach Hilberta. Otóż w takiej przestrzeni H też istnieje "baza", tylko jest ona nieprzeliczalna: $\{e_\iota : \iota \in J\}$ (gdzie J jest dużym zbiorem indeksów). Bazę taką konstruuje się następująco: Zbiory ortonormnalne wektorów (tzn. takie, że wszystkie wektory są unormowane i parami ortogonalne) spełniają założenia Lematu K–Z, stąd istnieje maksymalny zbór ortonormalny. To właśnie jest nasza baza $\{e_\iota : \iota \in J\}$. Teraz każdy wektor x ma swoje "współczynniki Fouriera" $x_\iota = \langle x|e_\iota\rangle$. Dla każdego podzbioru skończonego $F \subset J$ rzut ortogonalny x na podprzestrzeń H_F generowaną przez $\{e_\iota : \iota \in F\}$ jest równy $\sum_{\iota\in F} x_\iota e_\iota$ oraz norma tego rzutu wynosi $\sqrt{\sum_{\iota\in F} |x_\iota|^2}$ i jest NIE WIĘKSZA niż $\|x\|$. Z tego wynika, że (przy ustalonym $x \in H$) co najwyżej przeliczalnie wiele współczynników Fouriera jest niezerowych (gdyby ich było nieprzeliczalnie wiele, to sumy skończone byłyby nieograniczone – to jest standardowy fakt o nieprzeliczalnych zbiorach liczb dodatnich). Następnie sprawdzamy jednoznaczność przedstawienia x w bazie: Jeśli $x = \sum x'_\iota e_\iota$ (to znowu co najwyżej przeliczalnie wiele współczynników może być niezerowych), to dla każdego ι, z ortonormalności bazy, $\langle x|e_\iota\rangle = x'_\iota$, stąd x'_ι jest współczynnikiem Fouriera.

To tyle uzupełnienia wiedzy o nieośrodkowych przestrzeniach Hilberta. Mając bazę $\{e_\iota : \iota \in J\}$ definiujemy "sprzężenie" elementu x dokładnie tak samo, jak w przestrzeniach ośrodkowych, tak samo definiujemy Ψ i tak samo sprawdzamy, że to jest izometryczny izomorfizm. Koniec ostateczny dowodu Twierdzenia Riesza. $\square$

Twierdzenie. Przestrzeń sprzężona do ℓ^1 jest izometrycznie izomorficzna z ℓ^∞.

Dowód: Niech $y = (y_n) \in \ell^\infty$. Zdefiniujmy funkcjonał f_y na ℓ^1 wzorem

$$f_y(x) = \sum_{n=1}^{\infty} x_n y_n,$$

gdzie $x = (x_n) \in \ell^1$. Ten szereg jest bezwzględnie sumowalny i suma modułów jest nie większa niż $\|x\|_1 \, \|y\|_\infty$. Elementarnie sprawdza się, że f_y jest liniowy (przy sprawdzaniu addytywności $f(x + x') = f(x) + f(x')$ istotna jest *bezwzględna* sumowalność szeregów składowych definiujących $f(x)$ i $f(x')$, zwykła sumowalność nie wystarczyłaby). Powyższe oszacowanie sumy modułów daje, że norma f_y nie przekracza $\|y\|_\infty$. Zatem określiliśmy $\Phi : \ell^\infty \to (\ell^1)^*$, które nie zwiększa normy. Linowość Φ sprawdza się elementarnie (znów ingeruje bezwzględna sumowlność).

Idźmy dalej. W ℓ^1 jest baza $\{e_n : n \geq 1\}$, gdzie $e_n = (0, 0, 0, \ldots, 0, 1, 0, 0, 0, \ldots)$ (jedynka na pozycji n). Każdy element $x = (x_n)$ jest szeregiem (zbieżnym w ℓ^1)

$$x = \sum_{n=1}^{\infty} x_n e_n.$$

Zatem każdy funkcjonał liniowy i ciągły f na ℓ^1 działa tak:

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n).$$

Definiujemy $y = (y_n)$, gdzie $y_n = f(e_n)$. Czy jest to ciąg ograniczony? Tak, bo e_n mają normę 1, a funkcjonał jest ograniczony. Co więcej, dostajemy, że $\|y\|_\infty \leq \|f\|$. Ostatni wystrzelony wzór pokazuje $f = f_y$, czyli pokazaliśmy, że Φ^{-1} jest określone na całym $(\ell^1)^*$ i też nie zwiększa normy. To daje, że Φ jest bijekcją i izometrią, co kończy dowód. $\square$

Twierdzenie (Landaua). Dla $p \in (1, \infty)$ przestrzeń sprzężona do ℓ^p jest izometrycznie izomorficzna z ℓ^q, gdzie q jest liczbą "dualną" do p (tzn spełnia $\frac{1}{p} + \frac{1}{q} = 1$).

Uwaga: Zauważmy, że to twierdzenie uogólnia Twierdzenie Riesza dla ośrodkowych przestrzeni Hilberta, bowiem każda ośrodkowa przestrzeń Hilberta jest izometrycznie izomorficzna z ℓ^2 oraz liczbą dualną do $p = 2$ jest $q = 2$.

Lemat Niech (z_n) będzie ciągiem liczb nieujemnych ograniczonym i niesumowalnym. Wtedy istnieje ciąg nieujemny (a_n) taki, że $(a_n z_n)$ jest niesumowalny, a $(a_n^p z_n)$ – sumowalny dla dowolnego $p > 1$.

Dowód: Niech M oznacza ograniczenie na $|z_n|$. Łatwo widać, że można pogrupować wyrazy ciągu z_n w "pęczki"

$$g_k = z_{n_k} + z_{n_k+1} + z_{n_k+2} + \ldots z_{n_{k+1}-1}$$

(gdzie n_k jest podciągiem startującym od $n_1 = 1$), tak aby $1 \leq g_k \leq 2M$ (trzeba sumować tak długo, aż suma "pęczka" po raz pierwszy przekroczy M). Wtedy definiuejmy a_n wzorem $a_n = \frac{1}{k}$ dla $n \in [n_k, n_{k+1} - 1]$. Sprawdzenie, że taki ciąg spełnia tezę jest natychmiastowy:

$$\sum_{n=1}^{\infty} a_n z_n = \sum_{k=1}^{\infty} \frac{1}{k}\, g_k \geq \sum_{k=1}^{\infty} \frac{1}{k} = \infty,$$
$$\sum_{n=1}^{\infty} a_n^p z_n = \sum_{k=1}^{\infty} \frac{1}{k^p}\, g_k \leq 2M \sum_{k=1}^{\infty} \frac{1}{k^p} < \infty. \quad \square$$

Dowód Twierdzenia Landaua: W ℓ^p mamy bazę $\{e_n : n \geq 1\}$, gdzie $e_n = (0,0,0,\dots,0,1,0,0,0,\dots)$ (z jedynką na miejscu n-tym). Każdy $x \in \ell^p$ zapisuje się jednoznacznie jako szereg (zbieżny w ℓ^p)

$$x = (x_n) = \sum_{n=1}^{\infty} x_n e_n.$$

Zatem każdy funkcjonał liniowy ciągły f na ℓ^p działa następująco:

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n).$$

Ciąg $x' = (x'_n)$, gdzie $x'_n = \frac{|x_n f(e_n)|}{f(e_n)}$ ma takie same moduły jak (x_n) więc też należy do ℓ^p, zatem nasz funkcjonał musi na nim dawać wartość skończoną, czyli

$$f(x') = \sum_{n=1}^{\infty} \frac{|x_n f(e_n)|}{f(e_n)} f(e_n) = \sum_{n=1}^{\infty} |x_n||f(e_n)| < \infty.$$

Innymi słowy, ciąg $y = (y_n)$, gdzie $y_n = f(e_n)$ ma własność *bezwzględnej sumowalności z każdym ciągiem z* ℓ^q. Pokażemy teraz, że ciąg (y_n) ma tą własność wtedy i tylko wtedy, gdy należy do ℓ^q.

W jedną stronę wynika to wprost z nierówności Höldera: każdy ciąg $y \in \ell^q$ ma powyższą własność (suma szeregu modułów iloczynów $x_n y_n$ w nie przekracza iloczynu norm $\|y\|_q \, \|x\|_p$).

Zabieramy się za drugą implikację. Przypuśćmy, że (y_n) ma powyższą własność, ale nie należy do ℓ^q. Gdyby (y_n) był nieograniczony, to miałby podciąg (y_{n_k}) o modułach większych niż 2^k, a wtedy nie byłby bezwzględnie sumowalny z ciągiem (x_n) zdefiniowanym jako $x_{n_k} = 2^{-k}$ (i zera poza tym podciągiem). Taki ciąg (x_k) należy do ℓ^1, więc i do ℓ^p. To dowodzi, że (y_n) jest ograniczony. Wtedy ograniczony jest też ciąg $z_n = |y_n|^q$, ale skoro $(y_n) \notin \ell^q$, to (z_n) jest jest niesumowalny. No to bierzemy ciąg nieujemny (a_n) z lematu. Wtedy sumowalny jest ciąg

$$a_n^p z_n = \left(a_n |y_n|^{\frac{q}{p}}\right)^p,$$

co oznacza, że ciąg $x_n = a_n |y_n|^{\frac{q}{p}}$ należy do ℓ^p. Ale (y_n) jest z założenia bezwzględnie sumowalny z ciągiem (x_n), czyli sumowalny jest ciąg

$$|x_n y_n| = a_n |y_n|^{\frac{q}{p}} |y_n| = a_n |y_n|^{\frac{q}{p}+1} = a_n |y_n|^q = a_n z_n$$

(skorzystaliśmy z równości $\frac{q}{p}+1 = q$, która jest prostym przekształceniem zależności $\frac{1}{p} + \frac{1}{q} = 1$). To jest jednak ciąg niesumowalny, bo tak był dobrany ciąg (a_n). Ta sprzeczość dowodzi, że jednak $(y_n) \in \ell^p$.

Wykazaliśmy, że odwzorowanie $\Phi : \ell^q \to (\ell^p)^*$ przyporządkowujące elementowi $y = (y_n) \in \ell^q$ funkcjonał f_y zadany wzorem

$$f_y(x) = \sum_{n=1}^{\infty} x_n y_n,$$

jest bijekcją. Jej liniowość jest banalna (trzeba jednak skorzystać z bezwzględnej sumowalności – zwykła sumowalność nie wystarczyłaby). Pozostało sprawdzić, że jest izometrią. W jedną stronę jest to całkiem trywialne i w zasadzie zostało napisane: z nierówności Höldera $|f_y(x)| \leq \|x\|_p \, \|y\|_q$, czyli $\|f_y\| \leq \|y\|_q$. Na odwrót: mając $y = (y_n) \in \ell^q$ weźmy $x = (x_n)$, gdzie

$$x_n = \frac{|y_n|^q}{y_n}$$

(jeśli $y_n = 0$ to kładziemy $x_n = 0$). Teraz $(x_n) \in \ell^p$, bo $|x_n|^p = |y_n|^{(q-1)p} = |y_n|^q$, a to jest sumowalne (znowu $(q-1)p = q$ to oczywista przeróbka zależności $\frac{1}{p} + \frac{1}{q} = 1$). Norma p-ta z x wynosi

$$(\sum |x_n^p|)^{\frac{1}{p}} = (\sum |y_n|^q)^{\frac{1}{p}} = \|y\|_q^{\frac{q}{p}}.$$

Natomiast

$$|f_y(x)| = |\sum x_n y_n| = \left|\sum \frac{|y_n|^q}{y_n} y_n\right| = \sum |y_n|^q = \|y\|_q{}^q = \|y\|_q{}^{\frac{q}{p}+1} =$$

$$\|y\|_q{}^{\frac{q}{p}} \, \|y\|_q = \|x\|_p \, \|y\|_q,$$

czyli, że $\|f_y\| \geq \|y\|_q$. To kończy dowód. $\square$

Uwaga. Z powyższego twierdzenia wynika, że $((\ell^p)^*)^* = \ell^p$. Przestrzenie o tej własności V^{**} nazywamy *refleksywnymi.* (Każda przestrzeń unormowana V spełnia inkluzję $V \subset V^{**}$ – to jest na ćwiczeniach – do refleksywności istotna jest więc tylko przeciwna inkluzja). W naszych twierdzeniach nie identyfikujemy przestrzeni $(\ell^\infty)^*$. Okazuje się, że jest ona istotnie większa od ℓ^1 (jej opis pomijamy, gdyż wymaga on bardziej zaawansowanej teorii miary), dlatego przestrzenie ℓ^1 i ℓ^∞ nie są refleksywne.