Andrzej Sołtysiak

WSTĘP DO TEORII SPEKTRALNEJ

WYDAWNICTWO NAUKOWE UAM

WSTĘP DO TEORII SPEKTRALNEJ

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU

Andrzej Sołtysiak

WSTĘP DO TEORII SPEKTRALNEJ

WYDAWNICTWO NAUKOWE

POZNAŃ 2016

Recenzent: prof. dr hab. Krzysztof Rudol

Publikacja dofinansowana przez Wydział Matematyki i Informatyki Uniwersytetu im. Adama Mickiewicza w Poznaniu

Projekt okładki: Ewa Wąsowska
Redaktor: Anna Rąbalska

ISBN 978-83-232-2970-4

WYDAWNICTWO NAUKOWE
UNIWERSYTETU IM. ADAMA MICKIEWICZA W POZNANIU
61-701 POZNAŃ, UL. ALEKSANDRA FREDRY 10
www.press.amu.edu.pl
Sekretariat: tel. 61 829 46 46, faks 61 829 46 47, e-mail: wydnauk@amu.edu.pl
Dział sprzedaży: tel. 61 829 46 40, e-mail: press@amu.edu.pl
Wydanie I. Ark. wyd. 12,00. Ark. druk. 12,25
DRUK I OPRAWA: EXPOL, WŁOCŁAWEK, UL. BRZESKA 4

Spis treści

Przedmowa

Skrypt ten powstał na podstawie wykładów z teorii spektralnej i analizy funkcjonalnej, jakie autor prowadził dla studentów matematyki Uniwersytetu im. Adama Mickiewicza w Poznaniu w ciągu minionych dwudziestu lat. Zawiera on wykład z podstaw teorii spektralnej operatorów na przestrzeniach Banacha i Hilberta, poprzedzony trzema rozdziałami wstępnymi, w których są wprowadzone podstawowe definicje i twierdzenia z analizy funkcjonalnej.

Mała liczba godzin przeznaczonych na wykłady z teorii spektralnej zmusiła autora do dość zwięzłej formy wykładów i w konsekwencji również zaważyła na treści tego podręcznika.

Głównym celem, jaki sobie postawił autor, prowadząc wykłady z teorii spektralnej, było przedstawienie dowodu twierdzenia spektralnego dla ograniczonych operatorów samosprzężonych na przestrzeni Hilberta. Dowód przedstawiony w skrypcie pochodzi od Frigyesa Riesza i jest, zdaniem autora, najprostszym dowodem tego twierdzenia występującym w literaturze, niewymagającym budowy zaawansowanego aparatu analitycznego do jego przeprowadzenia.

Autor zakłada, że Czytelnik zna materiał wykładany na zajęciach z analizy matematycznej i algebry liniowej. Powołując się w tekście na fakty z tych przedmiotów, autor odwołuje się do napisanych przez siebie skryptów z analizy matematycznej i algebry liniowej.

Większość twierdzeń w tym skrypcie jest zamieszczona z pełnymi dowodami. Koniec dowodu tradycyjnie jest oznaczany symbolem $\blacksquare$. Na końcu każdego rozdziału znajduje się zestaw ćwiczeń, które mają ułatwić Czytelnikowi lepsze zrozumienie i opanowanie zawartych w skrypcie pojęć. Należy jednak zaznaczyć, że ich liczba i zakres nie są wystarczające do nabrania niezbędnej biegłości w posługiwaniu się podstawowymi pojęciami analizy funkcjonalnej. W tym celu należałoby, zdaniem autora, sięgnąć po jakiś zbiór zadań (np. z tych zamieszczonych w spisie literatury) i rozwiązać odpowiednią liczbę ćwiczeń.

Literatura dotycząca zagadnień przedstawionych w skrypcie jest niezwykle bogata. Jej wybór zamieszczony na jego końcu jest subiektywny. Zawiera zarówno pozycje, z których autor korzystał przy pisaniu tego skryptu, jak również takie, które, zdaniem autora, pomogą Czytelnikowi rozszerzyć swoją wiedzę.

Autor ma nadzieję, że książka ta okaże się użyteczna dla studentów matematyki, i będzie wdzięczny za wszelkie uwagi, które pozwolą mu usunąć błędy i niedoskonałości jej tekstu.

Autor dziękuje Panu doktorowi Maciejowi Łuczakowi za pomoc w przygotowaniu tekstu do druku z pomocą programu LaTeX. Serdeczne podziękowania należą się również Panu doktorowi habilitowanemu Krzysztofowi Rudolowi za wnikliwą recenzję i wiele cennych uwag, które pozwoliły ulepszyć tekst tej książki.

Skrypt ten powstał dzięki życzliwości władz Wydziału Matematyki i Informatyki Uniwersytetu im. Adama Mickiewicza, za co autor wyraża im serdeczne podziękowanie.

Poznań, październik 2015

Rozdział 1

Przestrzenie Banacha i operatory liniowe ograniczone

W tym rozdziale przedstawimy podstawowe pojęcia analizy funkcjonalnej, takie jak norma, przestrzeń unormowana, przestrzeń Banacha i operator liniowy ograniczony.

Symbole: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{N}$ i $\mathbb{R}_+$ będą w dalszym ciągu oznaczały odpowiednio zbiory: liczb rzeczywistych, zespolonych, naturalnych i liczb rzeczywistych nieujemnych.

1.1. Definicja i przykłady przestrzeni Banacha

Definicja 1.1. Niech X będzie przestrzenią liniową nad ciałem $\mathbb{K}$ liczb rzeczywistych lub zespolonych. *Normą* na przestrzeni X nazywamy funkcję $\|\cdot\|\colon X \to \mathbb{R}_+$, która ma następujące własności:

(1) $\|x\| = 0$ wtedy i tylko wtedy, gdy $x = 0$;

(2) $\|\lambda x\| = |\lambda|\|x\| \quad (\lambda \in \mathbb{K}, x \in X)$;

(3) $\|x + y\| \leqslant \|x\| + \|y\| \quad (x, y \in X)$.

Przestrzeń liniową X z normą $\|\cdot\|$ nazywamy *przestrzenią liniową unormowaną* lub krótko *przestrzenią unormowaną*.

Przykłady 1.1. (a) Przestrzeń euklidesowa skończenie wymiarowa $\mathbb{R}^k$ (lub $\mathbb{C}^k$) z normą $|\cdot|$ określoną wzorem $|x| = \sqrt{(x^1)^2 + (x^2)^2 + \ldots + (x^k)^2}$, gdzie $x = (x^1, x^2, \ldots, x^k)$, (w przypadku $\mathbb{C}^k$ wzór ten ma postać $|x| = \sqrt{|x^1|^2 + |x^2|^2 + \ldots + |x^k|^2}$) jest przestrzenią unormowaną. Innym przykładem normy w przestrzeni $\mathbb{R}^k$ (lub $\mathbb{C}^k$) jest funkcja $\|x\| = \max\limits_{1 \leqslant j \leqslant k} |x^j|$.

(b) Jeżeli K jest zwartą przestrzenią metryczną, to przestrzeń $\mathcal{C}(K)$ (funkcji ciągłych na przestrzeni K) z normą zbieżności jednostajnej $\|f\|_\infty = \sup\limits_{x \in K} |f(x)|$ jest przestrzenią unormowaną.

Każda przestrzeń unormowana X z normą $\|\cdot\|$ jest przestrzenią metryczną. Mianowicie, wzór

$$d(x,y) = \|x-y\| \qquad (x, y \in X)$$

określa metrykę w przestrzeni X. Wszystkie pojęcia z teorii przestrzeni metrycznych odnoszą się również do przestrzeni unormowanych. W szczególności zbieżność ciągu punktów (x_n) przestrzeni unormowanej X do punktu $x \in X$ oznacza, że $\lim\limits_{n\to\infty} \|x_n - x\| = 0$. Tak określoną zbieżność nazywamy *zbieżnością według normy*.

Twierdzenie 1.1 (własności normy i zbieżności według normy). *Niech X będzie przestrzenią unormowaną z normą $\|\cdot\|$. Wówczas:*

(a) $\|x_1 + x_2 + \ldots + x_n\| \leqslant \|x_1\| + \|x_2\| + \ldots + \|x_n\| \quad (x_1, x_2, \ldots, x_n \in X)$.

(b) $\big|\|x\| - \|y\|\big| \leqslant \|x - y\| \quad (x, y \in X)$.

(c) *Norma jest funkcją ciągłą, tzn. jeżeli $x_n \to x$, to $\|x_n\| \to \|x\|$.*

(d) *Działania algebraiczne w przestrzeni unormowanej są ciągłe, tzn.*
— jeżeli $x_n \to x$ i $y_n \to y$, to $x_n + y_n \to x + y$;
— jeżeli $\alpha_n \to \alpha$ i $x_n \to x$, to $\alpha_n x_n \to \alpha x$.

D o w ó d. Własność (a) wynika natychmiast poprzez indukcję z własności (3) normy.

Aby otrzymać (b), zauważmy, że

$$\|x\| = \|(x-y) + y\| \leqslant \|x - y\| + \|y\|.$$

Zatem

$$\|x\| - \|y\| \leqslant \|x - y\|.$$

Zamieniając miejscami elementy x i y, otrzymujemy nierówność

$$\|y\| - \|x\| \leqslant \|y - x\| = \|x - y\|.$$

Z obu powyższych nierówności wynika teza.

Własność (c) jest natychmiastową konsekwencją (b).

W końcu, aby udowodnić (d), zauważmy, że

$$\begin{aligned}
\|(x_n + y_n) - (x + y)\| &= \|(x_n - x) + (y_n - y)\| \leqslant \|x_n - x\| + \|y_n - y\|, \\
\|\alpha_n x_n - \alpha x\| &= \|\alpha_n(x_n - x) + (\alpha_n - \alpha)x\| \leqslant \\
&\leqslant \|\alpha_n(x_n - x)\| + \|(\alpha_n - \alpha)x\| = \\
&= |\alpha_n|\|x_n - x\| + |\alpha_n - \alpha|\|x\|. \qquad \blacksquare
\end{aligned}$$

Definicja 1.2. Przestrzeń unormowaną zupełną (w metryce wyznaczonej przez normę) nazywamy *przestrzenią Banacha.*

Przestrzeniami Banacha są przestrzenie $\mathbb{R}^k$ ($\mathbb{C}^k$) i $\mathcal{C}(K)$ z naturalnymi normami (por. [27], tw. 3.12 z części I i tw. 7.7 z części II). Poznamy teraz inne przykłady przestrzeni Banacha.

Przykłady 1.2. (a) *Przestrzeń* ℓ_∞

Jest to przestrzeń utworzona ze wszystkich nieskończonych ciągów liczbowych (ξ_k), które są ograniczone. Działania liniowe są określone w niej w naturalny sposób, tzn.

$$(\xi_k) + (\eta_k) = (\xi_k + \eta_k), \qquad \alpha(\xi_k) = (\alpha\xi_k). \tag{1.1}$$

Jest oczywiste, że suma dwóch ciągów ograniczonych jest ciągiem ograniczonym i że iloczyn ciągu ograniczonego przez skalar jest również ciągiem ograniczonym. Ponadto jest również jasne, że z tak określonymi działaniami ℓ_∞ jest przestrzenią liniową. Zauważmy, że elementem zerowym w tej przestrzeni jest ciąg (ξ_k), w którym $\xi_k = 0$ dla $k = 1, 2, \ldots$

Normę ciągu $x = (\xi_k) \in \ell_\infty$ określamy wzorem

$$\|x\|_\infty = \sup_k |\xi_k|. \tag{1.2}$$

Sprawdzimy, że istotnie jest to norma.

Aksjomat (1) jest oczywiście spełniony. Aby sprawdzić (2), zauważmy, że dla ciągu $x = (\xi_k) \in \ell_\infty$ i skalara α mamy

$$|\alpha\xi_k| = |\alpha||\xi_k| \leqslant |\alpha| \sup_k |\xi_k| = |\alpha|\|x\|_\infty$$

dla $k = 1, 2, \ldots$ i w konsekwencji

$$\|\alpha x\|_\infty = \sup_k |\alpha\xi_k| \leqslant |\alpha|\|x\|_\infty. \tag{1.3}$$

Stąd zakładając, że $\alpha \neq 0$, otrzymujemy

$$\|x\|_\infty = \left\|\frac{1}{\alpha}\alpha x\right\|_\infty \leqslant \frac{1}{|\alpha|}\|\alpha x\|_\infty,$$

czyli nierówność

$$|\alpha|\|x\|_\infty \leqslant \|\alpha x\|_\infty. \tag{1.4}$$

Nierówność ta jest również prawdziwa, gdy $\alpha = 0$. Z nierówności (1.3) i (1.4) wynika równość (2) z definicji normy.

W końcu, aby sprawdzić, że spełniony jest aksjomat (3), weźmy dowolne ciągi $x = (\xi_k)$ i $y = (\eta_k) \in \ell_\infty$. Wówczas mamy

$$|\xi_k + \eta_k| \leqslant |\xi_k| + |\eta_k| \leqslant \sup_k |\xi_k| + \sup_k |\eta_k| \leqslant \|x\|_\infty + \|y\|_\infty$$

dla $k = 1, 2, \ldots$ Stąd

$$\|x + y\|_\infty \leqslant \|x\|_\infty + \|y\|_\infty.$$

Sprawdzimy teraz, że przestrzeń ℓ_∞ jest zupełna. Niech $(x_n) = \Big(\big(\xi_k^{(n)}\big)\Big)$ będzie ciągiem z przestrzeni ℓ_∞ spełniającym warunek Cauchy'ego. Zatem dla dowolnej liczby $\varepsilon > 0$ istnieje taki wskaźnik n_0, że dla n, $m \geqslant n_0$ zachodzi nierówność:

$$\|x_n - x_m\|_\infty = \sup_k \big|\xi_k^{(n)} - \xi_k^{(m)}\big| < \varepsilon. \tag{1.5}$$

Udowodnimy, że ciąg (x_n) jest zbieżny do pewnego punktu $x \in \ell_\infty$. Z nierówności (1.5) otrzymujemy

$$\big|\xi_k^{(n)} - \xi_k^{(m)}\big| < \varepsilon \quad \text{dla} \quad n,\, m \geqslant n_0;\ k = 1, 2, \ldots \tag{1.6}$$

Wynika stąd, że każdy z ciągów liczbowych $\big(\xi_k^{(n)}\big)$ $(k = 1, 2, \ldots)$ jest zbieżny. Niech $\lim\limits_{n\to\infty} \xi_k^{(n)} = \xi_k$. Przechodząc teraz z m do granicy, $m \to \infty$, w nierówności (1.6) otrzymujemy

$$\big|\xi_k^{(n)} - \xi_k\big| \leqslant \varepsilon \quad \text{dla} \quad n \geqslant n_0;\ k = 1, 2, \ldots \tag{1.7}$$

Zatem ciąg $\big(\xi_k^{(n_0)} - \xi_k\big)$ jest ograniczony. Ponieważ

$$\xi_k = \xi_k^{(n_0)} - \big(\xi_k^{(n_0)} - \xi_k\big),$$

więc ciąg (ξ_k) jest również ograniczony, tzn. jest on elementem przestrzeni ℓ_∞. Oznaczmy więc $x = (\xi_k)$. Z nierówności (1.7) otrzymujemy

$$\|x_n - x\|_\infty = \sup_k \big|\xi_k^{(n)} - \xi_k\big| \leqslant \varepsilon, \quad \text{gdy} \quad n \geqslant n_0.$$

Wobec dowolności liczby ε oznacza to, że $\|x_n - x\|_\infty \to 0$, gdy $n \to \infty$. Pokazaliśmy więc, że przestrzeń ciągów ograniczonych ℓ_∞ jest przestrzenią Banacha.

(b) *Przestrzeń* c_0

Jest to przestrzeń utworzona ze wszystkich ciągów zbieżnych do zera z działaniami określonymi wzorami (1.1). Jest oczywiste, że jest to podprzestrzeń przestrzeni ℓ_∞. Norma w tej przestrzeni jest również określona wzorem (1.2).

Pokażemy, że c_0 jest przestrzenią Banacha. Ponieważ domknięty podzbiór przestrzeni zupełnej jest przestrzenią zupełną (por. [27], wn. 3.13 z części I), wystarczy więc pokazać, że c_0 jest domkniętą podprzestrzenią przestrzeni ℓ_∞.

Niech $(x_n) = \left(\left(\xi_k^{(n)}\right)\right)$ będzie ciągiem elementów przestrzeni c_0 zbieżnym według normy (1.2) do ciągu $x = (\xi_k)$. Pokażemy, że ciąg x jest zbieżny do zera. Mamy

$$|\xi_k| \leqslant \left|\xi_k - \xi_k^{(n)}\right| + \left|\xi_k^{(n)}\right| \leqslant \|x - x_n\|_\infty + \left|\xi_k^{(n)}\right|. \tag{1.8}$$

Wybieramy dowolną liczbę $\varepsilon > 0$ i ustalamy wskaźnik n tak duży, aby $\|x - x_n\|_\infty < \frac{1}{2}\varepsilon$. Ponieważ ciąg $x_n = \left(\xi_k^{(n)}\right)$ jest zbieżny do zera, więc istnieje wskaźnik k_0 taki, że $\left|\xi_k^{(n)}\right| < \frac{1}{2}\varepsilon$ dla wszystkich $k \geqslant k_0$. Zatem z (1.8) wynika, że dla $k \geqslant k_0$ mamy $|\xi_k| < \varepsilon$. Oznacza to, że $x = (\xi_k)$ jest ciągiem zbieżnym do zera. Pokazaliśmy więc, że c_0 jest domkniętą podprzestrzenią przestrzeni ℓ_∞, czyli jest ona przestrzenią Banacha.

(c) *Przestrzeń* $\mathcal{C}^{(1)}[a,b]$

Jest to przestrzeń funkcji klasy $\mathcal{C}^{(1)}$ na przedziale $[a,b]$ (o wartościach rzeczywistych lub zespolonych), tzn. przestrzeń funkcji mających ciągłą pochodną na tym przedziale, przy czym na końcach przedziału pochodne są rozumiane jako odpowiednie pochodne jednostronne. Działania liniowe określone są w zwykły sposób, tzn.

$$(f+g)(x) = f(x) + g(x), \qquad (\alpha f)(x) = \alpha f(x).$$

Z własności operacji różniczkowania (por. [27], tw. 5.3 z części I) i funkcji ciągłych (por. [27], tw. 4.7 z części I) wynika, że $\mathcal{C}^{(1)}[a,b]$ z tak określonymi działaniami jest przestrzenią liniową. Normę w tej przestrzeni określimy wzorem

$$\|f\| = |f(a)| + \sup_{a \leqslant x \leqslant b} |f'(x)|. \tag{1.9}$$

Jest oczywiste, że tak określony funkcjonał jest skończony dla każdej funkcji f klasy $\mathcal{C}^{(1)}$. Ponadto jeżeli $\|f\| = 0$, to funkcja f ma pochodną stale równą zeru w przedziale $[a,b]$ i $f(a) = 0$. Ponieważ f jest wówczas funkcją stałą (por. [27], wn. 5.9 z części I), więc $f = 0$. Pozostałe warunki z definicji normy sprawdza się podobnie jak w przypadku normy w przestrzeni funkcji ciągłych. Zauważmy, że jeżeli (f_n) jest ciągiem Cauchy'ego w przestrzeni $\mathcal{C}^{(1)}[a,b]$, to oznacza, że ciąg pochodnych (f_n') spełnia warunek Cauchy'ego ze względu na zbieżność jednostajną oraz ciąg liczbowy $(f_n(a))$ jest ciągiem Cauchy'ego. Zatem ciąg (f_n') jest zbieżny jednostajnie na przedziale $[a,b]$ do pewnej funkcji g (por. [27], tw. 7.1 z części II) i funkcja ta jest funkcją ciągłą (por. [27], tw. 7.5 z części II). Ponadto ciąg $\big(f_n(a)\big)$ jest zbieżny do $f(a)$. Z twierdzenia o różniczkowaniu ciągów funkcyjnych wynika, że ciąg (f_n) jest zbieżny jednostajnie na przedziale $[a,b]$ do funkcji różniczkowalnej f i $f' = g$ (por. [27], tw. 7.10 z części II). Oznacza to, że funkcja f należy do przestrzeni $\mathcal{C}^{(1)}[a,b]$ i $\|f_n - f\| \to 0$, gdy $n \to \infty$, czyli przestrzeń $\mathcal{C}^{(1)}[a,b]$ jest przestrzenią Banacha.

Ważnymi przykładami przestrzeni Banacha są przestrzenie związane z teorią Lebesgue'a. Będzie im poświęcony następny podrozdział.

1.2. Przestrzenie L_p

Niech $(X, \mathfrak{M}, \mu)$ będzie przestrzenią z miarą. Zatem $\mathfrak{M}$ jest σ-algebrą podzbiorów zbioru X, a μ jest miarą na tej σ-algebrze (por. [27], część II, rozdz. 11 lub [22], rozdz. 11). Symbolem $\mathcal{L}_p(X, \mathfrak{M}, \mu)$ $(p \geqslant 1)$ lub krótko $\mathcal{L}_p$ oznaczamy zbiór funkcji określonych na X o wartościach rzeczywistych lub zespolonych, mierzalnych względem σ-algebry $\mathfrak{M}$ i takich, że funkcja $x \mapsto |f(x)|^p$ jest całkowalna na X.

Lemat 1.2. *Zbiór $\mathcal{L}_p$ jest przestrzenią liniową z działaniami określonymi w naturalny sposób.*

D o w ó d. Wystarczy wykazać, że $\mathcal{L}_p$ jest podprzestrzenią liniową przestrzeni $\mathcal{F}(X, \mathbb{K})$ wszystkich funkcji określonych na zbiorze X o wartościach skalarnych, tzn. wystarczy pokazać, że jeżeli f, $g \in \mathcal{L}_p$, to $f + g \in \mathcal{L}_p$ oraz $\alpha f \in \mathcal{L}_p$ dla każdego $\alpha \in \mathbb{K}$.

Jest oczywiste, że jeżeli $f \in \mathcal{L}_p$, to $\alpha f \in \mathcal{L}_p$ dla każdej liczby α, bo

$$\int\limits_X |\alpha f|^p \, d\mu = |\alpha|^p \int\limits_X |f|^p \, d\mu < \infty.$$

Niech teraz f, $g \in \mathcal{L}_p$. Aby pokazać, że $f + g \in \mathcal{L}_p$, zauważmy najpierw, że dla dowolnych liczb a, $b \geqslant 0$ mamy

$$\begin{aligned}(a+b)^p &\leqslant \big(2\max\{a,b\}\big)^p = 2^p\big(\max\{a,b\}\big)^p = \\ &= 2^p \max\{a^p, b^p\} \leqslant 2^p(a^p + b^p).\end{aligned}$$

Wynika stąd, że

$$|f(x) + g(x)|^p \leqslant \big(|f(x)| + |g(x)|\big)^p \leqslant 2^p\big(|f(x)|^p + |g(x)|^p\big)$$

i w konsekwencji

$$\int\limits_X |f+g|^p \, d\mu \leqslant 2^p\bigg(\int\limits_X |f|^p \, d\mu + \int\limits_X |g|^p \, d\mu\bigg) < \infty,$$

co oznacza, że $f + g \in \mathcal{L}_p$. ■

Niech $\mathcal{N}$ będzie podzbiorem przestrzeni $\mathcal{L}_p$ określonym następująco: $\mathcal{N} = \{f \in \mathcal{L}_p\colon\ \mu\{x\colon\ f(x) \neq 0\} = 0\}$, czyli $f \in \mathcal{N}$, jeżeli $f(x) = 0$ prawie wszędzie na X. Zauważmy, że tak określony zbiór $\mathcal{N}$ jest podprzestrzenią liniową przestrzeni $\mathcal{L}_p$. Wynika to z następujących oczywistych inkluzji:

$$\begin{aligned}\{x\colon\ (f+g)(x) \neq 0\} &\subset \{x\colon\ f(x) \neq 0\} \cup \{x\colon\ g(x) \neq 0\}, \\ \{x\colon\ (\alpha f)(x) \neq 0\} &\subset \{x\colon\ f(x) \neq 0\}.\end{aligned}$$

Zatem możemy utworzyć przestrzeń ilorazową $\mathcal{L}_p/\mathcal{N}$. Jak wiemy, składa się ona z klas abstrakcji względem relacji równoważności ([26], s. 71–72):

$$f \sim g \quad \text{wtedy i tylko wtedy, gdy} \quad f - g \in \mathcal{N}.$$

Niech $[f]$ będzie klasą abstrakcji funkcji f względem tej relacji. Przestrzeń $\mathcal{L}_p/\mathcal{N}$ oznaczamy symbolem L_p lub dokładniej $L_p(X, \mathfrak{M}, \mu)$, jeżeli chcemy zaznaczyć na jakiej przestrzeni z miarą się znajdujemy. Przestrzeń L_p jest przestrzenią liniową z działaniami określonymi w naturalny sposób:

$$\begin{aligned}[f] + [g] &= [f+g], \\ \alpha[f] &= [\alpha f].\end{aligned}$$

Zauważmy, że $[f] = [g]$ wtedy i tylko wtedy, gdy $f = g$ prawie wszędzie na X. W dalszym ciągu dla uproszczenia zapisu elementy przestrzeni L_p będziemy oznaczać tak jak funkcje z przestrzeni $\mathcal{L}_p$. Należy przy tym pamiętać, że równość dwóch „funkcji” z przestrzeni L_p oznacza ich równość prawie wszędzie.

Twierdzenie 1.3 (nierówność Höldera). *Załóżmy, że liczby p, $q > 1$ są takie, że $\frac{1}{p} + \frac{1}{q} = 1$. Wówczas dla funkcji mierzalnych f i g zachodzi nierówność*

$$\int_X |fg|\,d\mu \leqslant \Big(\int_X |f|^p\,d\mu\Big)^{1/p}\Big(\int_X |g|^q\,d\mu\Big)^{1/q}. \tag{1.10}$$

D o w ó d. Wiemy, że (por. [27], przykład 5.2 z części I) dla a, $b \geqslant 0$ i p, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ prawdziwa jest *nierówność Younga*, tzn.

$$a^{1/p}b^{1/q} \leqslant \frac{a}{p} + \frac{b}{q}.$$

Jeżeli jedna z całek po prawej stronie nierówności (1.10) jest równa zeru, to wówczas iloczyn fg jest funkcją równą zeru prawie wszędzie (bo jedna z funkcji f, g jest równa zeru prawie wszędzie (por. [27], lem. 11.36 z części II)) i nierówność Höldera jest prawdziwa. Gdy jedna z całek po prawej stronie jest nieskończona, to jest oczywiste, że ta nierówność zachodzi. Zatem możemy założyć, że obydwie całki po prawej stronie są skończone i większe od zera. Ponadto możemy założyć, że funkcje f i g przyjmują wszędzie wartości skończone (por. [27], tw. 11.34 z części II).

Przyjmijmy

$$a = \frac{|f(x)|^p}{\int_X |f|^p\,d\mu}, \qquad b = \frac{|g(x)|^q}{\int_X |g|^q\,d\mu}.$$

Dla tak określonych a i b stosujemy nierówność Younga. Mamy

$$\frac{|f(x)g(x)|}{\Big(\int_X |f|^p\,d\mu\Big)^{1/p}\Big(\int_X |g|^q\,d\mu\Big)^{1/q}} \leqslant \frac{1}{p}\,\frac{|f(x)|^p}{\int_X |f|^p\,d\mu} + \frac{1}{q}\,\frac{|g(x)|^q}{\int_X |g|^q\,d\mu}.$$

Z monotoniczności całki (por. [27], tw. 11.31 (b) z części II) otrzymujemy

$$\frac{\int_X |fg|\,d\mu}{\Big(\int_X |f|^p\,d\mu\Big)^{1/p}\Big(\int_X |g|^q\,d\mu\Big)^{1/q}} \leqslant \frac{1}{p}\,\frac{\int_X |f|^p\,d\mu}{\int_X |f|^p\,d\mu} + \frac{1}{q}\,\frac{\int_X |g|^q\,d\mu}{\int_X |g|^q\,d\mu} =$$

$$= \frac{1}{p} + \frac{1}{q} = 1. \qquad \blacksquare$$

U w a g i. 1. W przypadku, gdy $p = q = 2$, nierówność Höldera nosi również nazwę *nierówności Schwarza.*

2. Niech $\Omega \subset \mathbb{R}^k$ będzie zbiorem mierzalnym w sensie Lebesgue'a i niech $\mathfrak{M}$ będzie σ-algebrą podzbiorów zbioru Ω mierzalnych w sensie Lebesgue'a, a m miarą Lebesgue'a na Ω. Wówczas nierówność Höldera przyjmuje postać

$$\int\limits_{\Omega} |f(x)g(x)|\,dx \leqslant \left(\int\limits_{\Omega} |f(x)|^p\,dx\right)^{1/p} \left(\int\limits_{\Omega} |g(x)|^q\,dx\right)^{1/q}.$$

Zwyczajowo całkę względem miary Lebesgue'a $\int\limits_{\Omega} f\,dm$ będziemy w dalszym ciągu oznaczać symbolem $\int\limits_{\Omega} f\,dx$ lub $\int\limits_{\Omega} f(x)\,dx$.

3. Jeżeli weźmiemy przestrzeń mierzalną $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, gdzie $\mathcal{P}(\mathbb{N})$ jest σ-algebrą wszystkich podzbiorów zbioru $\mathbb{N}$, a μ jest miarą liczącą, to funkcjami mierzalnymi są ciągi, a całkami sumy szeregów. Zatem całkowalność funkcji $f = (\xi_n)$ oznacza, że szereg $\sum\limits_{n=1}^{\infty} \xi_n$ jest bezwzględnie zbieżny i wówczas zachodzi równość $\int\limits_{\mathbb{N}} f\,d\mu = \sum\limits_{n=1}^{\infty} \xi_n$. Nierówność Höldera przyjmuje w tym przypadku postać

$$\sum_{n=1}^{\infty} |\xi_n \eta_n| \leqslant \left(\sum_{n=1}^{\infty} |\xi_n|^p\right)^{1/p} \left(\sum_{n=1}^{\infty} |\eta_n|^q\right)^{1/q}.$$

W szczególności wynika stąd nierówność dla sum skończonych

$$\sum_{k=1}^{n} |\xi_k \eta_k| \leqslant \left(\sum_{k=1}^{n} |\xi_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |\eta_k|^q\right)^{1/q}.$$

Wniosek 1.4 (nierówność Minkowskiego). *Dla liczby $p \geqslant 1$ oraz funkcji mierzalnych f i g zachodzi nierówność*

$$\left(\int\limits_X |f+g|^p\,d\mu\right)^{1/p} \leqslant \left(\int\limits_X |f|^p\,d\mu\right)^{1/p} + \left(\int\limits_X |g|^p\,d\mu\right)^{1/p}.$$

D o w ó d. Jeżeli całka po lewej stronie tej nierówności jest równa zeru lub jeżeli jedna z całek po prawej stronie jest nieskończona, to nierówność jest spełniona w sposób oczywisty. Jest również jasne, że nierówność ta zachodzi w przypadku, gdy $p = 1$. W dalszym ciągu możemy więc założyć, że całka z funkcji $|f+g|^p$ jest dodatnia, f, $g \in \mathcal{L}_p$ i $p > 1$. Ponadto możemy przyjąć, że funkcje f i g przyjmują wartości skończone.

Mamy

$$\begin{aligned}
\int\limits_X |f+g|^p \, d\mu &= \int\limits_X |f+g|\,|f+g|^{p-1} \, d\mu \leqslant \int\limits_X \Big(|f|+|g|\Big)\,|f+g|^{p-1} \, d\mu = \\
&= \int\limits_X |f|\,|f+g|^{p-1} \, d\mu + \int\limits_X |g|\,|f+g|^{p-1} \, d\mu \leqslant \\
&\leqslant \Big(\int\limits_X |f|^p \, d\mu\Big)^{1/p} \Big(\int\limits_X |f+g|^{(p-1)q} \, d\mu\Big)^{1/q} + \\
&\quad + \Big(\int\limits_X |g|^p \, d\mu\Big)^{1/p} \Big(\int\limits_X |f+g|^{(p-1)q} \, d\mu\Big)^{1/q} = \\
&= \Big(\int\limits_X |f+g|^p \, d\mu\Big)^{1/q} \left\{\Big(\int\limits_X |f|^p \, d\mu\Big)^{1/p} + \Big(\int\limits_X |g|^p \, d\mu\Big)^{1/p}\right\}.
\end{aligned}$$

Stąd

$$\Big(\int\limits_X |f+g|^p \, d\mu\Big)^{1-1/q} \leqslant \Big(\int\limits_X |f|^p \, d\mu\Big)^{1/p} + \Big(\int\limits_X |g|^p \, d\mu\Big)^{1/p},$$

czyli

$$\Big(\int\limits_X |f+g|^p \, d\mu\Big)^{1/p} \leqslant \Big(\int\limits_X |f|^p \, d\mu\Big)^{1/p} + \Big(\int\limits_X |g|^p \, d\mu\Big)^{1/p}. \qquad \blacksquare$$

U w a g i. 1. W przypadku, gdy $\Omega \subset \mathbb{R}^k$ jest zbiorem mierzalnym w sensie Lebesgue'a, nierówność Minkowskiego przyjmuje postać

$$\Big(\int\limits_\Omega |f(x)+g(x)|^p \, dx\Big)^{1/p} \leqslant \Big(\int\limits_\Omega |f(x)|^p \, dx\Big)^{1/p} + \Big(\int\limits_\Omega |g(x)|^p \, dx\Big)^{1/p}.$$

2. Jeżeli przestrzenią mierzalną jest przestrzeń $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, to nierówność Minkowskiego wygląda następująco:

$$\Big(\sum_{n=1}^{\infty} |\xi_n + \eta_n|^p\Big)^{1/p} \leqslant \Big(\sum_{n=1}^{\infty} |\xi_n|^p\Big)^{1/p} + \Big(\sum_{n=1}^{\infty} |\eta_n|^p\Big)^{1/p}.$$

Zauważmy, że w szczególności wynika stąd nierówność dla sum skończonych

$$\left(\sum_{k=1}^{n}|\xi_k+\eta_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n}|\xi_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n}|\eta_k|^p\right)^{1/p}.$$

Twierdzenie 1.5. *Niech $(X,\mathfrak{M},\mu)$ będzie przestrzenią mierzalną. Przestrzeń $L_p = L_p(X,\mathfrak{M},\mu)$ $(p \geqslant 1)$, w której norma jest określona wzorem*

$$\|f\|_p = \left(\int\limits_X |f|^p\,d\mu\right)^{1/p}, \tag{1.11}$$

jest przestrzenią Banacha.

D o w ó d. Warunki (1) i (2) z definicji normy są oczywiste. Natomiast warunek (3) wynika z nierówności Minkowskiego. Pozostaje do wykazania zupełność tej przestrzeni.

Niech ciąg (f_n) spełnia warunek Cauchy'ego w L_p. Dla dowolnej liczby $\varepsilon > 0$ istnieje wskaźnik n_0 taki, że dla wszystkich n, $m \geqslant n_0$ zachodzi nierówność

$$\int\limits_X |f_n - f_m|^p\,d\mu < \varepsilon^p.$$

Biorąc kolejno $\varepsilon = \dfrac{1}{3^{k/p}}$ dla $k = 1,2,\dots$, wybieramy rosnący ciąg wskaźników (n_k) taki, że

$$\int\limits_X |f_n - f_{n_k}|^p\,d\mu < \frac{1}{3^k} \qquad \text{dla} \quad n \geqslant n_k,\ k = 1,2,\dots$$

W szczególności

$$\int\limits_X |f_{n_{k+1}} - f_{n_k}|^p\,d\mu < \frac{1}{3^k} \qquad \text{dla} \quad k = 1,2,\dots$$

Niech

$$E_k = \left\{x\colon\ |f_{n_{k+1}}(x) - f_{n_k}(x)| > \frac{1}{2^{k/p}}\right\} \qquad (k = 1,2,\dots).$$

Wówczas $|f_{n_{k+1}}(x) - f_{n_k}(x)|^p > \dfrac{1}{2^k}$ dla $x \in E_k$. Stąd otrzymujemy

$$\frac{1}{3^k} > \int\limits_X |f_{n_{k+1}} - f_{n_k}|^p\,d\mu \geqslant \int\limits_{E_k} |f_{n_{k+1}} - f_{n_k}|^p\,d\mu \geqslant \frac{1}{2^k}\mu(E_k),$$

czyli $\mu(E_k) < (\frac{2}{3})^k$. Oznaczmy $E = \bigcap\limits_{j=1}^{\infty} \bigcup\limits_{k=j}^{\infty} E_k$. Jest to zbiór mierzalny. Pokażemy, że $\mu(E) = 0$. Mamy

$$\mu(E) = \mu\Big(\bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} E_k\Big) \leqslant \mu\Big(\bigcup_{k=j}^{\infty} E_k\Big) \leqslant$$

$$\leqslant \sum_{k=j}^{\infty} \mu(E_k) \leqslant \sum_{k=j}^{\infty} \left(\frac{2}{3}\right)^k = \frac{\left(\frac{2}{3}\right)^j}{1-\frac{2}{3}} = 3\left(\frac{2}{3}\right)^j$$

dla każdego $j = 1, 2, \ldots$ Przechodząc w tej nierówności z j do granicy, $j \to \infty$, otrzymujemy $\mu(E) = 0$.

Teraz wykażemy, że dla $x \notin E$ szereg $\sum\limits_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|$ jest zbieżny. Zauważmy, że na mocy praw de Morgana (por. [27], tw. 1.1 z części I) mamy

$$E^c = \Big(\bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} E_k\Big)^c = \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} E_k^c.$$

Zatem $x \notin E$ oznacza, że istnieje taki wskaźnik j, że dla wszystkich $k \geqslant j$ mamy $x \notin E_k$. Tak więc $x \notin E$ wtedy i tylko wtedy, gdy istnieje takie j, że dla każdego $k \geqslant j$ zachodzi nierówność

$$|f_{n_{k+1}}(x) - f_{n_k}(x)| \leqslant \frac{1}{2^{k/p}}.$$

Ponieważ szereg liczbowy $\sum\limits_{k=1}^{\infty} \dfrac{1}{2^{k/p}}$ jest zbieżny, więc na podstawie kryterium porównawczego szereg $\sum\limits_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$ jest zbieżny dla każdego $x \in$ $\in E^c$ (co więcej, na podstawie kryterium Weierstrassa szereg ten jest zbieżny jednostajnie i bezwzględnie na tym zbiorze). Wynika stąd, że szereg ten jest zbieżny prawie wszędzie na zbiorze X. Oznaczmy

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} \big(f_{n_{k+1}}(x) - f_{n_k}(x)\big).$$

Funkcja f jest określona prawie wszędzie na X. Ponadto $f_{n_k}(x) \to f(x)$ prawie wszędzie.

Wykażemy, że $f \in L_p$ oraz $\|f_{n_k} - f\|_p \to 0$, gdy $k \to \infty$. Rozważmy, przy ustalonym k i zmiennym m $(m \geqslant k)$, ciąg $\big(|f_{n_m}(x) - f_{n_k}(x)|^p\big)$. Wówczas

$$\lim_{m\to\infty} |f_{n_m}(x) - f_{n_k}(x)|^p = |f(x) - f_{n_k}(x)|^p.$$

Do tego ciągu stosujemy lemat Fatou. Mamy więc

$$\int\limits_X |f(x) - f_{n_k}(x)|^p \, d\mu \leqslant \liminf_{m\to\infty} \int\limits_X |f_{n_m}(x) - f_{n_k}(x)|^p \, d\mu \leqslant \frac{1}{3^k}.$$

Stąd

$$\Bigg(\int\limits_X |f(x) - f_{n_k}(x)|^p \, d\mu\Bigg)^{1/p} \leqslant \frac{1}{3^{k/p}} \qquad \text{dla} \quad k = 1, 2, \ldots$$

Oznacza to, że funkcja $f - f_{n_k} \in L_p$. Ponieważ $f_{n_k} \in L_p$, więc stąd otrzymujemy $f = (f - f_{n_k}) - f_{n_k} \in L_p$. Ponadto

$$\|f - f_{n_k}\|_p = \Bigg(\int\limits_X |f(x) - f_{n_k}(x)|^p \, d\mu\Bigg)^{1/p} \leqslant \frac{1}{3^{k/p}} \to 0, \qquad \text{gdy} \quad k \to \infty.$$

Zauważmy, że stąd wynika, że $\|f - f_n\|_p \to 0$, bo mamy

$$\|f - f_n\|_p \leqslant \|f - f_{n_k}\|_p + \|f_{n_k} - f_n\|_p \leqslant \frac{2}{3^{k/p}} \qquad \text{dla} \quad n \geqslant n_k.$$

Zatem L_p z normą określoną wzorem (1.11) jest przestrzenią zupełną. ■

U w a g i. 1. Wymienimy poniżej szczególne przypadki przestrzeni L_p, które występują najczęściej w różnych zagadnieniach:

(a) Jeżeli Ω jest podzbiorem mierzalnym w sensie Lebesgue'a przestrzeni $\mathbb{R}^k$, to przestrzeń $L_p(\Omega, \mathfrak{M}, m)$ jest przestrzenią Banacha względem normy

$$\|f\|_p = \Bigg(\int\limits_\Omega |f(x)|^p \, dx\Bigg)^{1/p}.$$

Przestrzeń tę będziemy w dalszym ciągu oznaczać symbolem $L_p(\Omega)$.

(b) Biorąc jako przestrzeń mierzalną X zbiór liczb naturalnych $\mathbb{N}$ z miarą liczącą, otrzymujemy przestrzeń ℓ_p ciągów liczbowych $x = (\xi_n)$, dla których szereg $\sum\limits_{n=1}^{\infty} |\xi_n|^p$ jest zbieżny. Jest ona przestrzenią Banacha z normą

$$\|x\|_p = \Bigg(\sum_{n=1}^{\infty} |\xi_n|^p\Bigg)^{1/p}.$$

2. W przypadku $p = 2$ przestrzeń L_p ma dodatkową strukturę. Mianowicie można w niej zdefiniować *iloczyn skalarny* wzorem*

$$\langle f, g\rangle = \int_X f\overline{g}\, d\mu.$$

Przestrzeń L_2 z tak zdefiniowanym iloczynem skalarnym jest *przestrzenią Hilberta*. Przestrzenie te odgrywają istotną rolę w wielu działach matematyki i fizyki teoretycznej. Przestrzeniom Hilberta będzie poświęcony następny rozdział.

1.3. Szeregi elementów przestrzeni unormowanej

Definicja szeregu zbieżnego, którego wyrazy są wektorami w przestrzeni unormowanej X, jest identyczna z tą dla szeregów o wyrazach liczbowych. Mianowicie szereg

$$\sum_{n=1}^{\infty} x_n \qquad (x_n \in X \ \text{ dla } \ n = 1, 2, \dots) \tag{1.12}$$

jest *zbieżny*, jeżeli zbieżny jest jego ciąg sum częściowych

$$s_n = x_1 + x_2 + \ldots + x_n \qquad (n = 1, 2, \dots).$$

Granica ciągu sum częściowych (s_n) (jeżeli istnieje) jest sumą szeregu (1.12). Jeżeli szereg (1.12) jest zbieżny, to jego sumy częściowe spełniają warunek Cauchy'ego, tzn. *dla dowolnego* $\varepsilon > 0$ *istnieje* n_0 *takie, że dla dowolnych* $n \geqslant m \geqslant n_0$ *zachodzi nierówność*

$$\|x_m + x_{m+1} + \ldots + x_n\| < \varepsilon.$$

Zauważmy, że biorąc $n = m$ w tym warunku, otrzymujemy, że jeżeli szereg (1.12) jest zbieżny, to $\lim\limits_{n\to\infty} x_n = 0$.

Jeżeli X jest przestrzenią Banacha, to warunek Cauchy'ego jest również warunkiem dostatecznym zbieżności szeregu (1.12) w przestrzeni X. Szereg (1.12) nazywamy *bezwzględnie zbieżnym*, jeżeli zbieżny jest szereg

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

* Kreska nad g oznacza operację brania liczby sprzężonej. Jeżeli rozważamy przestrzenie nad ciałem liczb rzeczywistych, to jest ona zbyteczna.

Twierdzenie 1.6. *Jeżeli X jest przestrzenią Banacha, to każdy szereg elementów tej przestrzeni $\sum\limits_{n=1}^{\infty} x_n$, który jest bezwzględnie zbieżny, jest zbieżny oraz zachodzi nierówność*

$$\left\| \sum_{n=1}^{\infty} x_n \right\| \leqslant \sum_{n=1}^{\infty} \|x_n\|. \tag{1.13}$$

D o w ó d. Szereg $\sum\limits_{n=1}^{\infty} x_n$ jest zbieżny, bo mamy

$$\|x_m + x_{m+1} + \ldots + x_n\| \leqslant \|x_m\| + \|x_{m+1}\| + \ldots + \|x_n\|,$$

a więc ciąg sum częściowych $s_n = x_1 + x_2 + \ldots + x_n$ spełnia warunek Cauchy'ego i z zupełności przestrzeni X wynika, że jest on zbieżny. Przechodząc w nierówności

$$\|s_n\| \leqslant \|x_1\| + \ldots + \|x_n\|$$

do granicy, $n \to \infty$, otrzymujemy drugą część tezy. ■

U w a g i. 1. Prawdziwe jest również twierdzenie odwrotne, tzn. prawdą jest, że *jeżeli w przestrzeni unormowanej X dowolny szereg bezwzględnie zbieżny jest zbieżny, to przestrzeń ta jest przestrzenią Banacha.*

2. Większość własności szeregów liczbowych przenosi się bez zmian na przypadek szeregów o wyrazach w przestrzeni Banacha, dlatego pominiemy tu ich sformułowania.

1.4. Przestrzenie ośrodkowe

Definicja 1.3. Przestrzeń metryczna X jest *ośrodkowa*, jeżeli istnieje zbiór $Z \subset X$ co najwyżej przeliczalny i gęsty w X.

Można udowodnić następujące (patrz np. [14], 15.7):

Twierdzenie 1.7. *Podzbiór przestrzeni metrycznej ośrodkowej jest przestrzenią ośrodkową.*

Przestrzenie skończenie wymiarowe $\mathbb{R}^k$ i $\mathbb{C}^k$ są ośrodkowe. Z twierdzenia aproksymacyjnego Weierstrassa ([27], wn. 7.29 z części II) wynika, że przestrzeń $\mathcal{C}[a,b]$ jest również ośrodkowa. Jest łatwo sprawdzić, że zbiór wszystkich wielomianów o współczynnikach wymiernych jest przeliczalny i gęsty w tej przestrzeni (jeżeli przestrzeń $\mathcal{C}[a,b]$ rozważamy nad ciałem liczb zespolonych, to bierzemy wielomiany, których współczynniki mają części rzeczywiste i urojone, będące liczbami wymiernymi). Proponujemy Czytelnikowi sprawdzenie, że przestrzeń c_0 ciągów zbieżnych do zera jest przestrzenią ośrodkową.

Przykład 1.3. Przestrzeń ℓ_∞ nie jest ośrodkowa.

Przypuśćmy, że tak nie jest, tzn. istnieje przeliczalny podzbiór Z gęsty w ℓ_∞. Wówczas zbiór wszystkich kul $B_{1/2}(z)$, $z \in Z$, pokrywa całą przestrzeń ℓ_∞. Niech Z_0 będzie zbiorem wszystkich ciągów $x = (\xi_k)$, dla których $\xi_k = 0$ lub $\xi_k = 1$ dla każdego k. Zbiór Z_0 jest nieprzeliczalny, a zbiór kul $B_{1/2}(z)$ jest przeliczalny, więc istnieją przynajmniej dwa różne punkty x_1 i x_2 ze zbioru Z_0, które leżą w jednej kuli $B_{1/2}(z)$. Wówczas

$$\|x_1 - x_2\|_\infty \leqslant \|x_1 - z\|_\infty + \|z - x_2\|_\infty < \frac{1}{2} + \frac{1}{2} = 1.$$

Natomiast z definicji normy w przestrzeni ℓ_∞ otrzymujemy $\|x_1 - x_2\|_\infty = 1$. Otrzymana sprzeczność dowodzi nieośrodkowości przestrzeni ℓ_∞.

Przykład 1.4. Przestrzeń ℓ_p $(p \geqslant 1)$ jest ośrodkowa.

Rozważmy zbiór Z złożony z ciągów postaci $z = (\zeta_1, \zeta_2, \dots, \zeta_n, 0, 0, \dots)$, gdzie $n = 1, 2, \dots$, a ζ_j są liczbami wymiernymi (dla przestrzeni zespolonej części rzeczywista i urojona tej liczby są wymierne). Wykażemy, że Z jest przeliczalnym i gęstym podzbiorem przestrzeni ℓ_p.

Przeliczalność zbioru Z jest oczywista. Wykażemy, że jest on gęstym podzbiorem tej przestrzeni. Niech $x = (\xi_k) \in \ell_p$ i $\varepsilon > 0$ będą dowolne. Wybieramy liczbę naturalną n, dla której

$$\sum_{k=n+1}^{\infty} |\xi_k|^p < \frac{1}{2}\varepsilon^p.$$

Następnie wybieramy takie liczby wymierne $\zeta_1, \zeta_2, \dots, \zeta_n$, aby

$$\sum_{k=1}^{n} |\zeta_k - \xi_k|^p < \frac{1}{2}\varepsilon^p.$$

Biorąc $z = (\zeta_1, \zeta_2, \dots, \zeta_n, 0, 0, \dots)$, otrzymujemy

$$\|z - x\|_p^p = \sum_{k=1}^{n} |\zeta_k - \xi_k|^p + \sum_{k=n+1}^{\infty} |\xi_k|^p < \varepsilon^p,$$

skąd $\|z - x\|_p < \varepsilon$, a to oznacza, że zbiór Z jest gęsty w przestrzeni ℓ_p.

Zbadamy teraz ośrodkowość przestrzeni funkcji całkowalnych z p-tą potęgą ($p \geqslant 1$) względem miary Lebesgue'a na przedziale $[a, b]$. Przestrzeń tę w skrócie będziemy oznaczać symbolem $L_p[a, b]$. Najpierw wykażemy następujące twierdzenie:

Twierdzenie 1.8. *Zbiór funkcji ciągłych na przedziale* $[a, b]$ *jest gęstym podzbiorem przestrzeni* $L_p[a, b]$.

D o w ó d. Teza twierdzenia oznacza, że dla dowolnej funkcji f całkowalnej z p-tą potęgą (względem miary Lebesgue'a) na przedziale $[a, b]$ i dowolnego $\varepsilon > 0$ istnieje funkcja ciągła g na tym przedziale, dla której

$$\|f - g\|_p = \left(\int_a^b |f(x) - g(x)|^p dx \right)^{1/p} < \varepsilon. \tag{1.14}$$

Inaczej mówiąc, funkcje całkowalne z p-tą potęgą można przybliżać funkcjami ciągłymi w normie przestrzeni $L_p[a, b]$.

Zauważmy, że w dowodzie możemy ograniczyć się do funkcji przyjmujących wartości rzeczywiste. W przypadku funkcji zespolonej f rozpatrujemy jej części rzeczywistą u i urojoną v. Są one funkcjami całkowalnymi z p-tą potęgą, bo $|u| \leqslant |f|$, $|v| \leqslant |f|$ i jeżeli one mogą być przybliżane poprzez funkcje ciągłe g_u i g_v, to $g = g_u + ig_v$ jest również funkcją ciągłą i przybliża ona funkcję f, bo mamy

$$\begin{aligned} |f(x) - g(x)|^p &= \left((u(x) - g_u(x))^2 + (v(x) - g_v(x))^2 \right)^{p/2} \leqslant \\ &\leqslant (|u(x) - g_u(x)| + |v(x) - g_v(x)|)^p \,. \end{aligned}$$

Stąd na mocy nierówności Minkowskiego otrzymujemy

$$\begin{aligned} \|f - g\|_p &= \left(\int_a^b |f(x) - g(x)|^p dx \right)^{1/p} \leqslant \\ &\leqslant \left(\int_a^b (|u(x) - g_u(x)| + |v(x) - g_v(x)|)^p \, dx \right)^{1/p} \leqslant \\ &\leqslant \left(\int_a^b |u(x) - g_u(x)|^p dx \right)^{1/p} + \left(\int_a^b |v(x) - g_v(x)|^p dx \right)^{1/p} = \\ &= \|u - g_u\|_p + \|v - g_v\|_p. \end{aligned}$$

Niech F będzie domkniętym podzbiorem przedziału $[a, b]$ i niech χ_F oznacza funkcję charakterystyczną tego zbioru. Rozważny funkcję pomocniczą $t(x) = \operatorname{dist}(x; F) = \inf\{|x - y| \colon\ y \in F\}$. Z nierówności

$$|t(x_1) - t(x_2)| \leqslant |x_1 - x_2|$$

wynika, że t jest funkcją ciągłą na przedziale $[a, b]$. Ponadto $t(x) = 0$ wtedy i tylko wtedy, gdy $x \in F$. Niech

$$g_n(x) = \frac{1}{1 + nt(x)} \quad (n = 1, 2, \ldots).$$

Wówczas funkcje g_n są ciągłe na przedziale $[a, b]$, $g_n(x) \leqslant 1$ dla dowolnego $x \in [a, b]$, $g_n(x) = 1$ dla $x \in F$ oraz $g_n(x) \to 0$ dla $x \in E = [a, b] \setminus F$. Z twierdzenia Lebesgue'a o zbieżności ograniczonej ([27], tw. 11.44 z części II) otrzymujemy, że

$$\|\chi_F - g_n\|_p = \Big(\int\limits_E g_n^p(x)\, dx\Big)^{1/p} \to 0.$$

Biorąc odpowiednio duże n, otrzymujemy nierówność (1.14) dla funkcji charakterystycznej dowolnego zbioru domkniętego zawartego w przedziale $[a, b]$.

Jeżeli $A \subset [a, b]$ jest zbiorem mierzalnym w sensie Lebesgue'a, to dla dowolnego $\varepsilon > 0$ istnieje zbiór domknięty $F \subset A$ taki, że $m(A \setminus F) < \varepsilon^p$ (patrz [27], tw. 11.20 z części II). Wówczas

$$\|\chi_A - \chi_F\|_p = \Big(\int\limits_{A \setminus F} dx\Big)^{1/p} = m(A \setminus F)^{1/p} < \varepsilon.$$

Stąd wynika, że funkcje charakterystyczne zbiorów mierzalnych w sensie Lebesgue'a można przybliżać funkcjami ciągłymi w normie przestrzeni $L_p[a, b]$. Ponieważ dowolna mierzalna funkcja prosta jest kombinacją liniową funkcji charakterystycznych zbiorów mierzalnych, więc również mierzalne funkcje proste mogą być przybliżane funkcjami ciągłymi w tej przestrzeni.

Jeżeli $f \in L_p[a, b]$ jest dowolną funkcją nieujemną, to istnieje niemalejący ciąg funkcji prostych mierzalnych s_n taki, że $s_n(x) \to f(x)$ dla $x \in [a, b]$ (patrz [27], tw. 11.30 z części II). Ponieważ $(f(x) - s_n(x))^p \leqslant f^p(x)$ dla dowolnego $x \in [a, b]$, więc z twierdzenia Lebesgue'a o zbieżności ograniczonej wynika, że wówczas $\|f - s_n\|_p \to 0$, gdy $n \to \infty$. Stąd wnioskujemy, że dowolną nieujemną funkcję całkowalną z p-tą potęgą można przybliżać funkcjami ciągłymi.

Aby zakończyć dowód, wystarczy zauważyć, że dowolną rzeczywistą funkcję całkowalną z p-tą potęgą można przedstawić w postaci różnicy dwóch funkcji nieujemnych również całkowalnych z p-tą potęgą. ■

Wniosek 1.9. *$L_p[a,b]$ $(p \geqslant 1)$ jest przestrzenią ośrodkową.*

D o w ó d. Wiemy, że zbiór wielomianów o współczynnikach wymiernych jest zbiorem przeliczalnym i gęstym w przestrzeni $\mathcal{C}[a,b]$. Jeżeli f jest dowolną funkcją całkowalną z p-tą potęgą na przedziale $[a,b]$, to na mocy twierdzenia 1.8 istnieje funkcja ciągła g na tym przedziale, dla której $\|f-g\|_p < \frac{1}{3}\varepsilon$. Na mocy twierdzenia aproksymacyjnego Weierstrassa ([27], wn. 7.29 z części II) istnieje wielomian w taki, że

$$\|g-w\|_\infty = \sup\{|g(x)-w(x)|\colon\ x \in [a,b]\} < \frac{\varepsilon}{3(b-a)^{1/p}}.$$

Z kolei istnieje wielomian v o współczynnikach wymiernych, dla którego

$$\|w-v\|_\infty < \frac{\varepsilon}{3(b-a)^{1/p}}.$$

Ostatecznie więc mamy

$$\begin{aligned}\|f-v\|_p &\leqslant \|f-g\|_p + \|g-w\|_p + \|w-v\|_p \leqslant \\ &\leqslant \frac{\varepsilon}{3} + 2\frac{\varepsilon}{3(b-a)^{1/p}}\left(\int_a^b dx\right)^{1/p} = \varepsilon. \qquad ■\end{aligned}$$

Wniosek 1.10. *Przestrzenie $L_p(\mathbb{R})$ i $L_p(\mathbb{R}_+)$ $(p \geqslant 1)$ są przestrzeniami ośrodkowymi.*

D o w ó d. Ponieważ $\mathbb{R} = \bigcup_{n=1}^{\infty} [-n,n]$, więc dowolna funkcja f całkowalna z p-tą potęgą na prostej $\mathbb{R}$ jest granicą ciągu funkcji $f_n = f\chi_n$, gdzie χ_n jest funkcją charakterystyczną przedziału $[-n,n]$. Z twierdzenia Lebesgue'a o zbieżności ograniczonej mamy

$$\|f-f_n\|_p = \left(\int_{-\infty}^{+\infty} |f(x)-f_n(x)|^p dx\right)^{1/p} \to 0.$$

Zatem dla dowolnego $\varepsilon > 0$ istnieje n takie, że $\|f-f_n\|_p < \frac{1}{2}\varepsilon$. Następnie na mocy wniosku 1.9 wybieramy wielomian w o współczynnikach wymiernych, dla którego

$$\left(\int_{-n}^{n} |f_n(x)-w(x)|^p dx\right)^{1/p} = \|f_n - w\chi_n\|_p < \frac{1}{2}\varepsilon.$$

Otrzymujemy więc

$$\|f - w\chi_n\|_p \leqslant \|f - f_n\|_p + \|f_n - w\chi_n\|_p < \varepsilon. \tag{1.15}$$

Ponieważ zbiór funkcji postaci $w\chi_n$, gdzie w jest wielomianem o współczynnikach wymiernych, a n jest dowolną liczbą naturalną, jest przeliczalny, nierówność (1.15) oznacza, że zbiór ten jest gęsty w przestrzeni $L_p(\mathbb{R})$. Analogicznie udowadnia się ośrodkowość przestrzeni $L_p(\mathbb{R}_+)$. ■

Uwaga. Można również udowodnić, że przestrzeń $L_p(\Omega)$ ($p \geqslant 1$), gdzie $\Omega \subset \mathbb{R}^k$ jest zbiorem mierzalnym w sensie Lebesgue'a, jest przestrzenią ośrodkową (patrz [14], 23.4).

1.5. Operatory liniowe ograniczone

Najważniejsze odwzorowania rozważane w analizie, tj. branie granicy ciągu czy też funkcji, różniczkowanie czy całkowanie, są przekształceniami liniowymi, a więc funkcjami A, dla których

$$\begin{aligned} A(x_1 + x_2) &= Ax_1 + Ax_2 \\ A(\lambda x) &= \lambda Ax, \end{aligned}$$

gdzie x_1, x_2, x są wektorami z dziedziny przekształcenia A, a λ jest skalarem. Gdy będziemy rozważali odwzorowania liniowe, będziemy na ogół pisać Ax zamiast $A(x)$.

Z punktu widzenia analizy najistotniejsze są te przekształcenia liniowe, które są ciągłe. W tym podrozdziale będą rozważane podstawowe własności takich przekształceń.

Definicja 1.4. Niech X i Y będą przestrzeniami unormowanymi (nad tym samym ciałem skalarów $\mathbb{K}$, przy czym $\mathbb{K}$ będzie zawsze oznaczać albo ciało liczb rzeczywistych $\mathbb{R}$, albo ciało liczb zespolonych $\mathbb{C}$). Normy w tych przestrzeniach będziemy oznaczali odpowiednio symbolami $\|\cdot\|_X$ i $\|\cdot\|_Y$ lub krótko symbolem $\|\cdot\|$, jeżeli to nie będzie prowadziło do nieporozumień. Operator liniowy $A\colon X \to Y$ jest *ograniczony*, jeżeli istnieje stała $M > 0$ taka, że

$$\|Ax\| \leqslant M\|x\| \quad \text{dla dowolnego} \quad x \in X. \tag{1.16}$$

Jest oczywiste, że warunek ten jest równoważny ograniczoności operatora A na kuli jednostkowej $\{x \in X\colon \|x\| \leqslant 1\}$ (lub na dowolnej kuli).

Jeżeli $Y = \mathbb{K}$, to operator liniowy $f\colon X \to \mathbb{K}$ nazywamy zazwyczaj *funkcjonałem liniowym*. Jeżeli spełnia on warunek (1.16), to nazywamy go *funkcjonałem liniowym ograniczonym*.

Twierdzenie 1.11. *Następujące warunki są równoważne*:

(a) *operator A jest ograniczony*;

(b) *operator A jest ciągły*;

(c) *operator A jest ciągły w jednym punkcie przestrzeni X*;

(d) *operator A jest ciągły w zerze.*

Dowód. Jeżeli warunek (a) jest spełniony, to

$$\|Ax - Ax_0\| = \|A(x - x_0)\| \leqslant M\|x - x_0\|,$$

a to oznacza ciągłość operatora A w dowolnym punkcie przestrzeni X.

Warunek (b) w sposób oczywisty implikuje (c). Załóżmy teraz, że spełniony jest warunek (c). Wybierzmy dowolne $\varepsilon > 0$. Wówczas istnieje $\delta > 0$ takie, że $\|Ax - Ax_0\| < \varepsilon$ dla wszystkich $\|x - x_0\| < \delta$. Jeżeli $\|y\| < \delta$, to $\|(y + x_0) - x_0\| < \delta$ i w konsekwencji $\|Ay\| = \|A(y + x_0) - Ax_0\| < \varepsilon$, co oznacza ciągłość operatora A w zerze.

W końcu pokażemy, że (a) jest konsekwencją (d). Warunek (d) w szczególności implikuje, że istnieje liczba $\eta > 0$ taka, że $\|Ax\| < 1$ dla $\|x\| < \eta$. Weźmy dowolne $x \neq 0$ i niech $y = \dfrac{\eta}{2\|x\|}\,x$. Wówczas $\|y\| = \frac{1}{2}\eta < \eta$, a zatem $\|Ay\| < 1$. Stąd $\|Ax\| \leqslant \dfrac{2}{\eta}\|x\|$ (zauważmy, że nierówność ta jest w oczywisty sposób prawdziwa dla $x = 0$). ■

Przykłady 1.5. (a) Pokażemy, że dowolne odwzorowanie liniowe $A\colon \mathbb{K}^n \to \mathbb{K}^m$ jest ograniczone (por. [27], lem. 8.2 z części II). Zakładamy, że przestrzenie $\mathbb{K}^n$ i $\mathbb{K}^m$ wyposażone są w normy euklidesowe. Niech $e_1, e_2, \ldots, e_n$ i $f_1, f_2, \ldots, f_m$ będą standardowymi bazami w tych przestrzeniach i niech (α_j^k) będzie macierzą operatora A względem tych baz. Zatem mamy

$$x = (\xi^1, \xi^2, \ldots, \xi^n) = \sum_{j=1}^{n} \xi^j e_j,$$

$$Ae_j = \sum_{k=1}^{m} \alpha_j^k f_k \qquad (j = 1, 2, \ldots, n).$$

Wówczas

$$|Ax|^2 = \Big|\sum_{j=1}^{n} \xi^j Ae_j\Big|^2 = \Big|\sum_{j=1}^{n} \xi^j \sum_{k=1}^{m} \alpha_j^k f_k\Big|^2 = \Big|\sum_{k=1}^{m}\Big(\sum_{j=1}^{n} \xi^j \alpha_j^k\Big) f_k\Big|^2 =$$
$$= \sum_{k=1}^{m}\Big|\sum_{j=1}^{n} \xi^j \alpha_j^k\Big|^2 \leqslant \sum_{k=1}^{m}\Big(\sum_{j=1}^{n} |\xi^j|^2 \sum_{j=1}^{n} |\alpha_j^k|^2\Big) = \sum_{j=1}^{n}\sum_{k=1}^{m} |\alpha_j^k|^2 \sum_{j=1}^{n} |\xi^j|^2.$$

Ostatecznie

$$|Ax| \leqslant M|x|, \quad \text{gdzie} \quad M = \Big(\sum_{j=1}^{n}\sum_{k=1}^{m} |\alpha_j^k|^2\Big)^{1/2}.$$

(b) Niech $X = Y = \ell_2$ i niech przekształcenie $S\colon\ \ell_2 \to \ell_2$ będzie zdefiniowane wzorem

$$S(\xi_1, \xi_2, \dots) = (0, \xi_1, \xi_2, \dots) \qquad (x = (\xi_k) \in \ell_2).$$

Jest to operator *jednostronnego przesunięcia*. Jest oczywiste, że

$$\|Sx\|_2 = \|x\|_2 \qquad (x \in \ell_2). \tag{1.17}$$

Zatem operator jednostronnego przesunięcia jest operatorem liniowym ograniczonym, co więcej — jest on izometrią.

(c) Nie każdy operator liniowy musi być ograniczony. Niech $X = \mathcal{C}^{(1)}[0,1]$, $Y = \mathcal{C}[0,1]$ i operator $A\colon\ X \to Y$ będzie określony wzorem $Ax(t) = x'(t)$. Wówczas A jest operatorem liniowym i ograniczonym. Mamy bowiem

$$\|x\|_X = |x(0)| + \sup_{0\leqslant t\leqslant 1} |x'(t)|,$$
$$\|Ax\|_Y = \sup_{0\leqslant t\leqslant 1} |x'(t)| \leqslant |x(0)| + \sup_{0\leqslant t\leqslant 1} |x'(t)| = \|x\|_X.$$

Jeżeli jednak rozważymy ten sam operator jako odwzorowanie z podprzestrzeni $X_1 \subset Y = \mathcal{C}[0,1]$, składającej się z wszystkich funkcji mających ciągłą pochodną na przedziale $[0,1]$, to jest on operatorem liniowym nieograniczonym. Norma w przestrzeni X_1 jest oczywiście normą z przestrzeni $\mathcal{C}[0,1]$. Jeżeli weźmiemy ciąg funkcji $x_n(t) = t^n$ $(n = 1, 2, \dots)$, to mamy

$$\|x_n\|_{X_1} = \|x_n\|_Y = \sup_{0\leqslant t\leqslant 1} |t^n| = 1,$$
$$\|Ax_n\|_Y = \sup_{0\leqslant t\leqslant 1} |nt^{n-1}| = n \quad (n = 1, 2, \dots).$$

Zatem $\|Ax_n\|_Y \to +\infty$, gdy $n \to \infty$.

Definicja 1.5. Niech X i Y będą przestrzeniami unormowanymi. Symbolem $\mathcal{B}(X,Y)$ oznaczymy zbiór wszystkich operatorów liniowych ograniczonych odwzorowujących przestrzeń X w przestrzeń Y. Jest oczywiste, że jest to przestrzeń liniowa (z naturalnymi działaniami). W przypadku, gdy $X = Y$, będziemy krótko pisać $\mathcal{B}(X)$ zamiast $\mathcal{B}(X,X)$. Natomiast $\mathcal{B}(X,\mathbb{K})$ będziemy oznaczać symbolem X' i nazywać *przestrzenią sprzężoną* z przestrzenią X lub *przestrzenią dualną* do przestrzeni X. Przestrzeń sprzężona jest również często oznaczana symbolem X^*. W dalszym ciągu przekonamy się, że w przestrzeni $\mathcal{B}(X,Y)$ można w naturalny sposób zdefiniować strukturę przestrzeni unormowanej.

Definicja 1.6. Niech X i Y będą przestrzeniami unormowanymi i $A\colon X \to Y$ operatorem liniowym ograniczonym. *Normą operatora* A nazywamy liczbę

$$\|A\| = \inf\{M > 0\colon \|Ax\| \leqslant M\|x\| \quad \text{dla każdego} \quad x \in X\}. \tag{1.18}$$

U w a g a. Zauważmy, że norma operatora S w przykładzie 1.5 (b) jest równa jeden. Istotnie z równości (1.17) mamy $\|S\| \leqslant 1$. Ponadto, jeżeli weźmiemy $e_n = (\delta_n^k)$, gdzie δ_n^k jest *deltą Kroneckera* ($\delta_n^k = 1$ dla $n = k$ i $\delta_n^k = 0$ w pozostałych przypadkach), to $\|e_n\|_2 = 1$ i $\|Se_n\|_2 = \|e_{n+1}\|_2 = 1$, czyli nierówność $\|Sx\|_2 < M\|x\|_2$ nie jest spełniona dla wszystkich ciągów $x \in \ell_2$ przy żadnym $M < 1$. Zatem $\|S\| = 1$.

Analogicznie można sprawdzić, że norma operatora $A\colon \mathcal{C}^{(1)}[0,1] \to \mathcal{C}[0,1]$ w przykładzie 1.5 (c) jest równa 1 (pozostawiamy to jako łatwe ćwiczenie dla Czytelnika). Natomiast dla operatora z przykładu 1.5 (a) możemy podać tylko oszacowanie normy z góry ($\|A\| \leqslant M$). Obliczenie normy jest w tym przypadku możliwe, gdy znamy macierz (α_j^k) przekształcenia A.

Obliczanie normy operatora ułatwia następujące twierdzenie:

Twierdzenie 1.12 (o normie operatora). *Niech X i Y będą przestrzeniami unormowanymi, a $A\colon X \to Y$ — operatorem liniowym ograniczonym. Wówczas*

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|\leqslant 1} \|Ax\|, \tag{1.19}$$

$$\|Ax\| \leqslant \|A\|\|x\| \quad \textit{dla każdego} \quad x \in X. \tag{1.20}$$

Ponadto zbiór $\mathcal{B}(X,Y)$ z normą operatorową określoną wzorem (1.18) *jest przestrzenią unormowaną.*

U w a g a. Aby równość $\|A\| = \sup\limits_{\|x\|=1} \|Ax\|$ była prawdziwa, przestrzeń X musi zawierać przynajmniej jeden niezerowy wektor. Natomiast równość $\|A\| = \sup\limits_{\|x\|\leqslant 1} \|Ax\|$ zachodzi bez tego dodatkowego założenia.

D o w ó d. Weźmy dowolne $M > \|A\|$. Wówczas $\|Ax\| \leqslant M\|x\|$ dla każdego $x \in X$. Stąd

$$\sup_{\|x\|=1} \|Ax\| \leqslant \sup_{\|x\|\leqslant 1} \|Ax\| \leqslant M$$

i wobec dowolności $M > \|A\|$ otrzymujemy

$$\sup_{\|x\|=1} \|Ax\| \leqslant \sup_{\|x\|\leqslant 1} \|Ax\| \leqslant \|A\|.$$

Jeżeli $\|A\| = 0$, to stąd otrzymujemy równości (1.19). Niech $\|A\| > 0$ i niech $0 < M < \|A\|$ będzie dowolne. Istnieje wówczas wektor $x_0 \in X$ taki, że $\|Ax_0\| > M\|x_0\|$. Stąd w szczególności wynika, że $x_0 \neq 0$. Jeżeli weźmiemy $x_1 = \frac{1}{\|x_0\|}x_0$, to $\|x_1\| = 1$ oraz

$$\|Ax_1\| = \left\| \frac{1}{\|x_0\|} Ax_0 \right\| = \frac{\|Ax_0\|}{\|x_0\|} > M.$$

Zatem

$$\sup_{\|x\|=1} \|Ax\| \geqslant \|Ax_1\| > M$$

i wobec dowolności $M < \|A\|$ mamy

$$\sup_{\|x\|=1} \|Ax\| \geqslant \|Ax_1\| \geqslant \|A\|.$$

Dowód (1.19) jest więc zakończony.

Zauważmy, że (1.20) jest prawdziwe, gdy x jest wektorem zerowym. Załóżmy więc, że $x \neq 0$ i niech $y = \frac{1}{\|x\|}x$. Wówczas $\|y\| = 1$ i mamy

$$\|Ay\| \leqslant \sup_{\|z\|=1} \|Az\| = \|A\|,$$

czyli

$$\left\| A\left(\frac{1}{\|x\|}x\right) \right\| \leqslant \|A\|$$

i w konsekwencji

$$\|Ax\| \leqslant \|A\|\|x\|.$$

Teraz udowodnimy, że norma operatorowa jest normą na przestrzeni $\mathcal{B}(X,Y)$. Jest oczywiste, że $\|A\|$ jest zawsze liczbą nieujemną oraz że $A=0$ implikuje $\|A\|=0$. Na odwrót, jeżeli $\|A\|=0$, to z (1.20) wynika, że $\|Ax\|=0$ dla dowolnego $x\in X$. Zatem $Ax=0$ dla dowolnego x, co oznacza, że $A=0$. Pozostaje do wykazania subaddytywność i jednorodność normy operatorowej. Ponieważ dla A i $B\in\mathcal{B}(X,Y)$ mamy

$$\|(A+B)x\|\leqslant\|Ax\|+\|Bx\|\leqslant(\|A\|+\|B\|)\,\|x\|,$$

więc

$$\|A+B\|\leqslant\|A\|+\|B\|.$$

W końcu dla $A\in\mathcal{B}(X,Y)$ i $\lambda\in\mathbb{K}$ mamy

$$\|(\lambda A)x\|=\|\lambda Ax\|=|\lambda|\|Ax\|\leqslant|\lambda|\|A\|\|x\|.$$

Zatem

$$\|\lambda A\|\leqslant|\lambda|\|A\|.$$

Jeżeli $\lambda\neq 0$, to

$$\|A\|=\left\|\frac{1}{\lambda}\lambda A\right\|\leqslant\frac{1}{|\lambda|}\|\lambda A\|,$$

czyli

$$|\lambda|\|A\|\leqslant\|\lambda A\|$$

i ostatecznie mamy

$$\|\lambda A\|=|\lambda|\|A\|.$$

Jest oczywiste, że ta równość jest prawdziwa również dla $\lambda=0$. ■

Przykłady 1.6. (a) Wyznaczymy normę operatora całkowego zdefiniowanego w następujący sposób. Niech Ω będzie niepustym zbiorem i niech $(t,\omega)\mapsto D(t,\omega)$ będzie funkcją określoną na produkcie $[a,b]\times\Omega$ o wartościach rzeczywistych taką, że funkcja $t\mapsto D(t,\omega)$ jest ciągła na przedziale $[a,b]$ przy każdym ustalonym $\omega\in\Omega$. Ponadto załóżmy, że liczba

$$D=\sup_{\omega\in\Omega}\int_a^b|D(t,\omega)|\,dt$$

jest skończona. Dla dowolnej funkcji ciągłej x na przedziale $[a,b]$ definiujemy funkcję y zmiennej ω wzorem:

$$y(\omega)=\int_a^b D(t,\omega)x(t)\,dt. \tag{1.21}$$

Mamy

$$
\begin{aligned}
(1.22)\qquad |y(\omega)| &= \left|\int_a^b D(t,\omega)x(t)\,dt\right| \leqslant \\
&\leqslant \sup_{a\leqslant t\leqslant b} |x(t)| \int_a^b |D(t,\omega)|\,dt \leqslant D\|x\|_\infty .
\end{aligned}
$$

Zatem operator A określony wzorem $Ax = y$, gdzie y jest funkcją zdefiniowaną za pomocą (1.21), przekształca przestrzeń funkcji ciągłych $\mathcal{C}[a,b]$ w przestrzeń funkcji ograniczonych $B(\Omega)$. Przestrzeń $B(\Omega)$ z naturalnymi działaniami algebraicznymi i normą $\|y\|_\infty = \sup\{|y(\omega)|\colon\ \omega \in \Omega\}$ jest przestrzenią Banacha. Z liniowości całki wynika liniowość operatora A, a jego ograniczoność jest konsekwencją nierówności (1.22). Z tej nierówności wynika również $\|A\| \leqslant D$. Pokażemy, że w istocie zachodzi tu równość, a więc pokażemy, że

$$
(1.23)\qquad \|A\| = \sup_{\omega\in\Omega} \int_a^b |D(t,\omega)|\,dt.
$$

Rozważmy funkcje u_n $(n = 1, 2, \dots)$ określone wzorem

$$
u_n(s) = \begin{cases} -1, & \text{gdy } s \leqslant -\dfrac{1}{n}, \\ ns & \text{gdy } -\dfrac{1}{n} < s < \dfrac{1}{n}, \\ 1, & \text{gdy } s \geqslant \dfrac{1}{n}. \end{cases}
$$

Funkcje $x_{n,\omega}(t) = u_n\left(D(t,\omega)\right)$ $(n = 1, 2, \dots)$ są funkcjami ciągłymi zmiennej t dla dowolnego $\omega \in \Omega$, a ponadto iloczyn $D(t,\omega)x_{n,\omega}(t)$ jest funkcją nieujemną równą $|D(t,\omega)|$, gdy $|D(t,\omega)| \geqslant \dfrac{1}{n}$ i mniejszą bądź równą 1 w pozostałych punktach. Niech

$$
A_{n,\omega} = \left\{t \in [a,b]\colon\ |D(t,\omega)| \geqslant \frac{1}{n}\right\}, \quad B_{n,\omega} = \left\{t \in [a,b]\colon\ |D(t,\omega)| < \frac{1}{n}\right\}.
$$

Wówczas $[a, b] = A_{n,\omega} \cup B_{n,\omega}$ oraz

$$\begin{aligned}\int_a^b |D(t,\omega)|\,dt &= \int_{A_{n,\omega}} |D(t,\omega)|\,dt + \int_{B_{n,\omega}} |D(t,\omega)|\,dt \leqslant \\ &\leqslant \int_{A_{n,\omega}} |D(t,\omega)|\,dt + \frac{1}{n}(b-a).\end{aligned}$$

Mamy więc

$$\begin{aligned}Ax_{n,\omega}(\omega) &= \int_a^b D(t,\omega)x_{n,\omega}(t)\,dt \geqslant \int_{A_{n,\omega}} |D(t,\omega)|\,dt \geqslant \\ &\geqslant \int_a^b |D(t,\omega)|\,dt - \frac{1}{n}(b-a).\end{aligned}$$

Ponieważ $\|x_{n,\omega}\|_\infty = \sup_{a \leqslant t \leqslant b} |u_n(D(t,\omega)| \leqslant 1$ dla dowolnego $\omega \in \Omega$, więc

$$\|A\| = \sup_{\|x\|_\infty \leqslant 1} \|Ax\|_\infty \geqslant \sup_{n,\omega} |Ax_{n,\omega}(\omega)]| = \sup_{\omega} \int_a^b |D(t,\omega)|\,dt.$$

Stąd łącznie z nierównością (1.22) otrzymujemy (1.23).

(b) Niech D będzie funkcją ciągłą na przedziale $[a, b]$. Wówczas wzór

$$f(x) = \int_a^b D(t)x(t)\,dt \tag{1.24}$$

definiuje funkcjonał liniowy ograniczony na przestrzeni $\mathcal{C}[a, b]$, który może być rozważany jako szczególny przypadek poprzedniego przykładu, gdy zbiór parametrów jest jednoelementowy. Stosując otrzymany wynik, stwierdzamy, że norma funkcjonału f dana jest wzorem

$$\|f\| = \int_a^b |D(t)|\,dt, \tag{1.25}$$

Twierdzenie 1.13. *Jeżeli X jest przestrzenią unormowaną, a Y przestrzenią Banacha, to przestrzeń $\mathcal{B}(X, Y)$ z normą operatorową jest również przestrzenią Banacha.*

D o w ó d. Musimy udowodnić, że przestrzeń $\mathcal{B}(X,Y)$ jest zupełna. Niech (A_n) będzie ciągiem Cauchy'ego elementów przestrzeni $\mathcal{B}(X,Y)$. Obierzmy dowolne $\varepsilon > 0$. Istnieje wskaźnik n_0 taki, że

$$\|A_n - A_m\| < \varepsilon \quad \text{dla} \quad n,m \geqslant n_0.$$

Stąd otrzymujemy

$$\|A_n x - A_m x\| \leqslant \|A_n - A_m\|\|x\| < \varepsilon\|x\|. \tag{1.26}$$

To oznacza, że ciąg $(A_n x)$ jest ciągiem Cauchy'ego w przestrzeni Y. Ponieważ ta przestrzeń jest zupełna, więc ciąg ten jest zbieżny. Oznaczmy $Ax = \lim\limits_{n\to\infty} A_n x$. Ponieważ operatory A_n są liniowe, więc również odwzorowanie $A\colon X \to Y$ jest liniowe. Wykażemy, że jest ono ograniczone. Przechodząc w (1.26) do granicy, $m \to \infty$, otrzymujemy

$$\|(A_n - A)x\| = \|A_n x - Ax\| \leqslant \varepsilon\|x\| \quad \text{dla} \quad x \in X,\ n \geqslant n_0. \tag{1.27}$$

Stąd w szczególności wynika, że operator $A_n - A$ (dla odpowiednio dużego n) jest ograniczony. Ponieważ $A = A_n - (A_n - A)$, więc i operator A jest ograniczony, czyli $A \in \mathcal{B}(X,Y)$. Z (1.27) otrzymujemy $\|A_n - A\| \leqslant \varepsilon$ dla $n \geqslant n_0$, co oznacza, że $A_n \to A$, gdy $n \to \infty$, w przestrzeni $\mathcal{B}(X,Y)$. ■

Ponieważ ciało skalarów $\mathbb{K}$ jest przestrzenią zupełną (por. [27], tw. 3.12 z części I), więc z twierdzenia 1.13 otrzymujemy następujący fakt:

Wniosek 1.14. *Przestrzeń sprzężona X' z przestrzenią unormowaną X jest przestrzenią Banacha z normą*

$$\|f\| = \sup_{\|x\|=1} |f(x)|.$$

Z twierdzenia 1.13 wynika również, że jeżeli X jest przestrzenią Banacha, to zbiór wszystkich operatorów liniowych i ograniczonych $\mathcal{B}(X)$ na przestrzeni X jest przestrzenią Banacha (z normą operatorową). Zauważmy, że przestrzeń ta ma bogatszą strukturę algebraiczną. Mianowicie, jeżeli zdefiniujemy *iloczyn operatorów* jako ich złożenie, tzn.

$$(AB)x = (A \circ B)x = A(Bx) \qquad (x \in X),$$

to otrzymamy operator liniowy AB na przestrzeni X. Ponadto z nierówności

$$\|(AB)x\| \leqslant \|A(BX)\| \leqslant \|A\|\|Bx\| \leqslant \|A\|\|B\|\|x\|$$

wynika, że

$$\|AB\| \leqslant \|A\|\|B\|,$$

a więc operator AB jest ograniczony. Zatem iloczyn operatorów jest działaniem w przestrzeni $\mathcal{B}(X)$. Łatwo się sprawdza, że ma on następujące własności:

$$A(BC) = (AB)C; \tag{1.28}$$

$$A(B+C) = AB + AC, \quad (B+C)A = BA + CA; \tag{1.29}$$

$$\lambda(AB) = (\lambda A)B = A(\lambda B), \tag{1.30}$$

gdzie A, B i $C \in \mathcal{B}(X)$ oraz $\lambda \in \mathbb{K}$. Przestrzeń liniowa, w której określone jest mnożenie spełniające powyższe warunki, nosi nazwę *algebry*. Podkreślmy, że jeżeli $\dim X > 1$, to algebra $\mathcal{B}(X)$ jest nieprzemienna. Elementem neutralnym względem mnożenia, czyli *jedynką*, jest operator identycznościowy I_X określony wzorem $I_X(x) = x$. Oczywiście $\|I_X\| = 1$. Jeżeli nie będzie to prowadziło do nieporozumienia, to operator I_X będziemy krótko oznaczali symbolem I.

Twierdzenie 1.15. *Mnożenie w algebrze $\mathcal{B}(X)$ jest ciągłe, tzn. jeżeli $A_n \to A$ i $B_n \to B$, gdy $n \to \infty$, to $A_nB_n \to AB$.*

D o w ó d. Mamy

$$\begin{aligned} \|A_nB_n - AB\| &= \|(A_n - A)B - A_n(B - B_n)\| \leqslant \\ &\leqslant \|A_n - A\|\|B\| + \|A_n\|\|B - B_n\|. \end{aligned} \tag{1.31}$$

Ponieważ $A_n \to A$ implikuje $\|A_n\| \to \|A\|$, więc z (1.31) otrzymujemy tezę. ■

Jeżeli X jest przestrzenią Banacha, to algebra $\mathcal{B}(X)$ jest przykładem *algebry Banacha*, tj. przestrzeni Banacha, w której określone jest ciągłe mnożenie spełniające warunki (1.28)–(1.30).

1.6. Przestrzenie skończenie wymiarowe

Definicja 1.7. Niech X i Y będą przestrzeniami unormowanymi (nad tym samym ciałem $\mathbb{K}$). Liniowa bijekcja $T\colon X \to Y$ taka, że T i T^{-1} są ciągłe, nazywa się *izomorfizmem przestrzeni unormowanych*.

Twierdzenie 1.16. *Dowolne dwie n-wymiarowe przestrzenie unormowane są ze sobą izomorficzne.*

D o w ó d. Jest oczywiste, że relacja izomorfizmu przestrzeni unormowanych jest relacją równoważności. Wystarczy więc udowodnić, że dowolna przestrzeń unormowana n-wymiarowa (nad ciałem $\mathbb{K}$) jest izomorficzna z przestrzenią $\mathbb{K}^n$ z normą euklidesową, tzn.

$$|x| = \left(\sum_{j=1}^{n} |x^j|^2\right)^{1/2}$$

dla $x = (x^1, x^2, \ldots, x^n) \in \mathbb{K}^n$ (w przypadku przestrzeni $\mathbb{R}^n$ moduły pod znakiem sumy są zbędne). Niech X będzie n-wymiarową przestrzenią unormowaną z normą $\|\cdot\|$ i niech $e_1, e_2, \ldots, e_n$ będzie jej bazą. Wówczas dowolny element $x \in X$ można w jednoznaczny sposób przedstawić w postaci

$$x = \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n \qquad (\xi^1, \xi^2, \ldots, \xi^n \in \mathbb{K}).$$

Odwzorowanie $T\colon\ \mathbb{K}^n \to X$ określamy wzorem

$$T\xi = T(\xi^1, \xi^2, \ldots, \xi^n) = x = \sum_{j=1}^{n} \xi^j e_j.$$

Jest oczywiste, że przekształcenie T jest liniową bijekcją przestrzeni $\mathbb{K}^n$ na przestrzeń X (por. [26], tw. 5.5). Należy udowodnić, że T i T^{-1} są odwzorowaniami ciągłymi lub — co na jedno wychodzi — ograniczonymi. Dla dowolnego $\xi \in \mathbb{K}^n$ mamy

$$\begin{aligned}\|T\xi\| = \|x\| &= \left\|\sum_{j=1}^{n} \xi^j e_j\right\| \leqslant \sum_{j=1}^{n} |\xi^j| \|e_j\| \leqslant \\ &\leqslant \left(\sum_{j=1}^{n} \|e_j\|^2\right)^{1/2} \left(\sum_{j=1}^{n} |\xi^j|^2\right)^{1/2} = M\left(\sum_{j=1}^{n} |\xi^j|^2\right)^{1/2} = M|\xi|,\end{aligned}$$

gdzie $M = \left(\sum\limits_{j=1}^{n} \|e_j\|^2\right)^{1/2}$ jest stałą. Stąd wynika, że odwzorowanie T jest ograniczone. Teraz wykażemy ograniczoność operatora T^{-1}. Niech $S = \{\xi \in \mathbb{K}^n\colon\ |\xi| = 1\}$ będzie sferą jednostkową w $\mathbb{K}^n$. Z twierdzenia Heinego-Borela ([27], tw. 2.15 z części I) wynika, że S jest zbiorem zwartym. Zdefiniujemy funkcję $f\colon\ S \to \mathbb{R}$ wzorem

$$f(\xi) = f(\xi^1, \xi^2, \ldots, \xi^n) = \|T\xi\| = \left\|\sum_{j=1}^{n} \xi^j e_j\right\|.$$

Ponieważ przekształcenia $y \mapsto \|y\|$ i T są ciągłe, więc f jest ciągłe jako złożenie funkcji ciągłych. Niech $m = \inf\{f(\xi)\colon \xi \in S\}$. Wówczas $m \geqslant 0$. Z twierdzenia Weierstrassa ([27], tw. 4.11 z części I) wynika, że istnieje $\xi_0 \in S$ takie, że $f(\xi_0) = m$. Gdyby $m = 0$, to $\|T\xi_0\| = f(\xi_0) = 0$. Stąd $T\xi_0 = 0$. Z wzajemnej jednoznaczności odwzorowania T wynikałoby, że $\xi_0 = 0$. To jest niemożliwe, bo $|\xi_0| = 1$. Zatem $m > 0$. Ponieważ $\|T\xi\| \geqslant m$ dla $\xi \in S$, więc $\|T\xi\| \geqslant m|\xi|$ dla dowolnego $\xi \in \mathbb{K}^n$. Stąd, jeżeli $x = T\xi$, to $\xi = T^{-1}x$ i mamy $|T^{-1}x| \leqslant \dfrac{1}{m}\|x\|$, co oznacza, że odwzorowanie T^{-1} jest ograniczone.

■

Definicja 1.8. Dwie normy $\|\cdot\|_1$ i $\|\cdot\|_2$ na przestrzeni liniowej X są równoważne, jeżeli dla dowolnego ciągu (x_n) elementów tej przestrzeni $\|x_n\|_1 \to 0$ wtedy i tylko wtedy, gdy $\|x_n\|_2 \to 0$.

Wniosek 1.17. *Dowolne dwie normy na przestrzeni liniowej skończenie wymiarowej są równoważne.*

Wniosek 1.18. *Każda skończenie wymiarowa przestrzeń unormowana jest zupełna.*

D o w ó d. Niech X i $T\colon \mathbb{K}^n \to X$ będą takie same jak w twierdzeniu 1.16. Wówczas dla dowolnych x i $y \in X$ mamy $x = \sum\limits_{j=1}^{n} \xi^j e_j$, $y = \sum\limits_{j=1}^{n} \eta^j e_j$. Niech $\xi = (\xi^1, \xi^2, \dots, \xi^m)$, $\eta = (\eta^1, \eta^2, \dots, \eta^n)$. Z własności odwzorowania T otrzymujemy

$$\begin{aligned} m|\xi - \eta| &\leqslant \|T\xi - T\eta\| = \|x - y\| = \\ &= \left\| \sum_{j=1}^{n} (\xi^j - \eta^j) e_j \right\| \leqslant M|\xi - \eta|. \end{aligned} \tag{1.32}$$

Z tych nierówności wynika, że jeżeli (x_m) jest ciągiem Cauchy'ego w przestrzeni X, to ciąg $(\xi_m) = (T^{-1}x_m)$ jest ciągiem Cauchy'ego w przestrzeni $\mathbb{K}^n$. Ponieważ przestrzeń $\mathbb{K}^n$ jest zupełna ([27], tw. 3.12 z części I), więc ciąg (ξ_m) jest zbieżny do $\xi_0 = (\xi_0^1, \xi_0^2, \dots, \xi_0^n)$. Z nierówności (1.32) wnioskujemy, że ciąg (x_m) jest zbieżny do $x_0 = \sum\limits_{j=1}^{n} \xi_0^j e_j$. To dowodzi zupełności przestrzeni X. ■

Wniosek 1.19. *Skończenie wymiarowa podprzestrzeń przestrzeni unormowanej jest domknięta.*

Dowód. Teza wynika z tego, że zupełny podzbiór przestrzeni metrycznej jest domknięty. ■

Wniosek 1.20. *Ograniczony i domknięty podzbiór przestrzeni unormowanej skończenie wymiarowej jest zwarty.*

Dowód. Niech X i $T\colon \mathbb{K}^n \to X$ będą takie same jak w twierdzeniu 1.16. Niech E będzie ograniczonym i domkniętym podzbiorem przestrzeni X. Z nierówności (1.32) wynika, że $T^{-1}(E)$ jest ograniczonym i domkniętym podzbiorem przestrzeni $\mathbb{K}^n$. Z twierdzenia Heinego-Borela ([27], tw. 2.15 z części I) otrzymujemy zwartość zbioru $T^{-1}(E)$. Ponieważ ciągły obraz zbioru zwartego jest zwarty ([27], tw. 4.9 z części I), więc zbiór $E = T(T^{-1}(E))$ jest zwarty. ■

Ćwiczenia

1. Udowodnić, że jeżeli podprzestrzeń liniowa X_0 przestrzeni unormowanej X zawiera pewną kulę, to $X_0 = X$.

2. Niech X i Y będą przestrzeniami unormowanymi z normami odpowiednio $\|\cdot\|_X$ i $\|\cdot\|_Y$. W zbiorze $X \times Y$ wprowadzamy działania dodawania wektorów i mnożenia wektora przez skalar w naturalny sposób. Wykazać, że funkcjonał zdefiniowany wzorem $\|(x,y)\| = \|x\|_X + \|y\|_Y$ jest normą w przestrzeni $X \times Y$. Ponadto przestrzeń ta jest zupełna, jeżeli zupełne są przestrzenie X i Y.

3. Wykazać bezpośrednio (nie korzystając z tw. 1.5), że przestrzeń ℓ_p $(1 \leqslant p < \infty)$ z naturalną normą jest przestrzenią Banacha.

4. Wykazać, że przestrzeń c_{00} wszystkich ciągów liczbowych $x = (\xi_n)$, dla których tylko skończenie wiele wyrazów jest różnych od zera, z normą $\|x\|_1 = \sum\limits_{n=1}^{\infty} |\xi_n|$ nie jest przestrzenią zupełną.

5. Wykazać, że przestrzeń funkcji ograniczonych $B(\Omega)$ na niepustym zbiorze Ω z normą $\|x\|_\infty = \sup\{|x(\omega)|\colon \omega \in \Omega\}$ jest przestrzenią Banacha.

6. Niech $\Omega \subset \mathbb{R}^k$ będzie zbiorem mierzalnym w sensie Lebesgue'a i niech $\mathcal{L}_\infty(\Omega)$ oznacza zbiór wszystkich funkcji mierzalnych w sensie Lebesgue'a na Ω, które są prawie wszędzie ograniczone, tzn.

$$(*) \qquad |f(x)| \leqslant M$$

dla pewnego $M > 0$ zależnego od funkcji f i dla wszystkich $x \in \Omega$ poza zbiorem miary zero. Zbiór $\mathcal{L}_\infty(\Omega)$ jest przestrzenią liniową z działaniami określonymi w naturalny sposób. Niech $\sim$ oznacza relację równości prawie wszędzie. Zbiór klas abstrakcji $\mathcal{L}_\infty(\Omega)/\sim$ oznaczymy przez $L_\infty(\Omega)$. Podobnie jak w przypadku przestrzeni L_p będziemy utożsamiali funkcję z jej klasą abstrakcji. Norma $\|f\|_\infty$ jest kresem dolnym liczb występujących w nierówności $(*)$. Sprawdzić, że istotnie jest to norma i że przestrzeń $L_\infty(\Omega)$ z tą normą jest przestrzenią Banacha.

7. Wykazać, że przestrzeń $\mathcal{C}_{2\pi}$ funkcji ciągłych i 2π-okresowych określonych na prostej z normą $\|f\|_\infty = \sup\{|f(x)|\colon\ x \in \mathbb{R}\}$ jest przestrzenią Banacha.

8. Wykazać, że na to, aby dwie normy $\|\cdot\|_1$ i $\|\cdot\|_2$ na przestrzeni liniowej X były równoważne, potrzeba i wystarcza, aby istniały takie stałe dodatnie m i M, że

$$m\|x\|_1 \leqslant \|x\|_2 \leqslant M\|x\|_1$$

dla dowolnego $x \in X$.

9. Udowodnić, że jeżeli każdy bezwzględnie zbieżny szereg elementów przestrzeni unormowanej X jest zbieżny, to X jest przestrzenią Banacha.

10. Mówimy, że szereg elementów przestrzeni unormowanej jest *bezwarunkowo zbieżny*, jeżeli dla każdej permutacji σ zbioru liczb naturalnych zbieżny jest szereg $\sum\limits_{n=1}^{\infty} x_{\sigma(n)}$ i $\sum\limits_{n=1}^{\infty} x_{\sigma(n)} = \sum\limits_{n=1}^{\infty} x_n$. Dowieść, że

(a) każdy bezwzględnie zbieżny szereg elementów przestrzeni Banacha jest bezwarunkowo zbieżny;

(b) szereg elementów $\sum\limits_{n=1}^{\infty} x_n$ przestrzeni ℓ_∞, gdzie $x_n = \big(\xi_k^{(n)}\big)$ oraz $\xi_k^{(n)} = \dfrac{1}{n}$, gdy $n = k$ i $\xi_k^{(n)} = 0$, gdy $n \neq k$, jest bezwarunkowo zbieżny, ale nie jest zbieżny bezwzględnie.

11. Wykazać, że przestrzeń $\mathcal{C}^{(1)}[a,b]$ jest przestrzenią ośrodkową.

12. Obliczyć normę funkcjonału f na przestrzeni $\mathcal{C}[0,1]$ danego wzorem

$$f(x) = \alpha\, x(0) + \beta\, x(1),$$

gdzie α i β są ustalonymi liczbami.

13. Obliczyć normę funkcjonału F na przestrzeni $L_1[0,1]$ danego wzorem

$$F(x) = \int\limits_0^1 x(t)a(t)\,dt,$$

gdzie $t \mapsto a(t)$ jest funkcją należącą do przestrzeni $L_\infty[0,1]$.

14. Niech (λ_n) będzie ciągiem liczbowym ograniczonym. Sprawdzić, że wzór

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

określa funkcjonał liniowy ograniczony na przestrzeni ℓ_1, którego norma wyraża się wzorem $\|f\| = \sup\{|\lambda_n|\colon\ n \in \mathbb{N}\}$.

Rozdział 2

Przestrzenie Hilberta

W tym rozdziale będziemy zajmowali się przestrzeniami liniowymi wyposażonymi w dodatkową strukturę zwaną iloczynem skalarnym. Przestrzeń otrzymana w ten sposób jest przestrzenią unormowaną, a w przypadku, gdy jest ona zupełna, nosi nazwę przestrzeni Hilberta. Przestrzenie te stanowią niezwykle ważną klasę przestrzeni Banacha.

2.1. Definicja i przykłady przestrzeni Hilberta

Definicja 2.1. Niech X będzie przestrzenią liniową nad ciałem liczb zespolonych $\mathbb{C}$. *Iloczynem skalarnym* na przestrzeni X nazywamy funkcjonał $\langle \cdot\,, \cdot \rangle \colon\ X \times X \to \mathbb{C}$, który spełnia następujące warunki:*

(a) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ dla $x,\ y \in X$;

(b) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ dla $x,\ y,\ z \in X$;

(c) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ dla $x,\ y \in X$, $\alpha \in \mathbb{C}$;

(d) $\langle x, x \rangle > 0$ dla każdego wektora $x \in X$, $x \neq 0$.

Przestrzeń liniową X wyposażoną w iloczyn skalarny $\langle \cdot\,, \cdot \rangle$ będziemy nazywać *przestrzenią unitarną*.

Uwagi. 1. Z własności (c) wynika, że $\langle 0, y \rangle = 0$ dla każdego $y \in X$. W szczególności $\langle 0, 0 \rangle = 0$.

2. Własności (b) i (c) mówią, że funkcjonał $x \mapsto \langle x, y \rangle$, dla dowolnego ustalonego $y \in X$, jest funkcjonałem liniowym na X.

3. Z (a) i (c) wynika, że $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$.

4. Z (a) i (b) wynika następujący warunek: $\langle z, x + y \rangle = \langle z, x \rangle + \langle z, y \rangle$.

5. Można również rozważać przestrzenie liniowe nad ciałem liczb rzeczywistych z iloczynem skalarnym. Ponieważ nie odgrywają one istotnej roli w teorii spektralnej, nie będziemy się nimi zajmować.

---

* Symbol $\overline{\alpha}$ oznacza liczbę (zespoloną) sprzężoną do liczby α.

W dalszym ciągu przez X będziemy oznaczać ustaloną przestrzeń unitarną, a przez $\langle \cdot\,, \cdot \rangle$ — iloczyn skalarny określony w tej przestrzeni.

Twierdzenie 2.1 (nierówność Schwarza). *Jeżeli x, y są wektorami w przestrzeni unitarnej X, to*

$$|\langle x, y\rangle|^2 \leqslant \langle x, x\rangle\langle y, y\rangle. \tag{2.1}$$

D o w ó d. Dla $y = 0$ nierówność (2.1) jest oczywista. Niech więc $y \neq 0$. Dla dowolnej liczby $\lambda \in \mathbb{C}$ mamy

$$\langle x + \lambda y, x + \lambda y\rangle \geqslant 0,$$

czyli

$$\langle x, x\rangle + \overline{\lambda}\langle x, y\rangle + \lambda\overline{\langle x, y\rangle} + |\lambda|^2\langle y, y\rangle \geqslant 0. \tag{2.2}$$

Podstawiając do (2.2)

$$\lambda = -\frac{\langle x, y\rangle}{\langle y, y\rangle},$$

otrzymujemy

$$\langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} \geqslant 0,$$

skąd wynika (2.1). ■

Definicja 2.2. Ponieważ $\langle x, x\rangle \geqslant 0$ dla każdego wektora $x \in X$, więc wyrażenie $\sqrt{\langle x, x\rangle}$ ma sens. Liczbę tę oznaczamy przez $\|x\|$ i nazywamy *normą* lub *długością* wektora x.

Twierdzenie 2.2 (własności normy wektora). *W przestrzeni unitarnej X norma wektora $\|\cdot\|$ ma następujące własności:*

(a) $\|x\| > 0$ *dla* $x \in X$, $x \neq 0$, $\|0\| = 0$;

(b) $\|\lambda x\| = |\lambda|\|x\|$ *dla* $x \in X$, $\lambda \in \mathbb{C}$;

(c) $\|x + y\| \leqslant \|x\| + \|y\|$ *dla* x, $y \in X$;

(d) $\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$ *dla* x, $y \in X$.

D o w ó d. Własności (a) i (b) są oczywiste. Aby udowodnić (c), skorzystamy z nierówności Schwarza. Mamy bowiem

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \overline{\langle x, y\rangle} + \langle y, y\rangle = \\ &= \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2.\end{aligned}$$

Ponieważ
$$\operatorname{Re}\langle x, y\rangle \leqslant |\langle x, y\rangle|,$$
więc na mocy nierówności Schwarza
$$\|x+y\|^2 \leqslant \|x\|^2+2|\langle x, y\rangle|+\|w\|^2 \leqslant \|x\|^2+2\|x\|\,\|y\|+\|y\|^2 = (\|x\|+\|y\|)^2 .$$

Natomiast (d) wynika z następujących przekształceń:
$$\begin{aligned}\|x+y\|^2+\|x-y\|^2 &= \langle x+y, x+y\rangle+\langle x-y, x-y\rangle = \\ &= 2\left(\langle x, x\rangle+\langle y, y\rangle\right)=2\left(\|x\|^2+\|y\|^2\right). \qquad \blacksquare\end{aligned}$$

U w a g i. 1. Z własności (a)–(c) wynika, że istotnie przestrzeń unitarna X z normą określoną wzorem $\|x\|=\sqrt{\langle x, x\rangle}$ jest przestrzenią unormowaną.

2. Własność (c) nosi nazwę *nierówności trójkąta*, a własność (d) — *tożsamości równoległoboku*. Rozważmy płaszczyznę $\mathbb{R}^2$ jako przestrzeń liniową wektorów swobodnych z iloczynem skalarnym wektorów. Wówczas otrzymamy przestrzeń euklidesową (unitarną nad ciałem liczb rzeczywistych). Niech x i y będą niezerowymi wektorami w tej przestrzeni. Zbudujmy z nich trójkąt. Wówczas trzeci bok tego trójkąta możemy interpretować jako wektor $x+y$, a długości boków tego trójkąta odpowiadają liczbom $\|x\|$, $\|y\|$ i $\|x+y\|$. Wówczas nierówność trójkąta dla normy wyraża znaną własność geometryczną, która mówi, że *długość boku w trójkącie jest mniejsza niż długość sumy dwóch pozostałych boków tego trójkąta*. Proponujemy Czytelnikowi jako ćwiczenie podanie geometrycznej interpretacji tożsamości równoległoboku na płaszczyźnie.

Przykłady 2.1. (a) W przestrzeni $\mathbb{C}^n$ przykładem iloczynu skalarnego jest standardowy iloczyn skalarny określony wzorem
$$\langle x, y\rangle=\sum_{j=1}^n \xi^j \overline{\eta^j},$$
gdzie $x=(\xi^1, \xi^2, \ldots, \xi^n)$ i $y=(\eta^1, \eta^2, \ldots, \eta^n)$. Zatem przestrzeń $\mathbb{C}^n$ z naturalnym iloczynem skalarnym jest przestrzenią unitarną. Zauważmy ponadto, że baza standardowa tej przestrzeni e_1, e_2, …, e_n, gdzie $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, („1” znajduje się na j-tym miejscu) ma następującą własność:
$$\langle e_j, e_k\rangle=\delta_{jk} \quad \text{(delta Kroneckera)}.$$

Nierówność Schwarza ma w tym przypadku postać

$$\Big|\sum_{j=1}^{n} \xi^j \overline{\eta^j}\Big| \leqslant \Big(\sum_{j=1}^{n} |\xi^j|^2\Big)^{1/2} \Big(\sum_{j=1}^{n} |\eta^j|^2\Big)^{1/2}.$$

(b) Przykładem nieskończenie wymiarowej przestrzeni unitarnej jest przestrzeń funkcji ciągłych $\mathcal{C}[-1,1]$, w której iloczyn skalarny jest określony za pomocą wzoru

$$\langle x, y\rangle = \int_{-1}^{1} x(t)\overline{y(t)}\,dt. \tag{2.3}$$

(c) Niech $(X, \mathfrak{M}, \mu)$ będzie przestrzenią mierzalną. Zbiór $L_2 = L_2(X, \mathfrak{M}, \mu)$ „funkcji" całkowalnych z kwadratem, czyli takich, że

$$\int_X |f(x)|^2\,d\mu(x) < \infty,$$

staje się przestrzenią unitarną, jeżeli iloczyn skalarny w tej przestrzeni określimy wzorem

$$\langle f, g\rangle = \int_X f(x)\overline{g(x)}\,d\mu(x).$$

Nierówność Schwarza ma wówczas postać

$$\Big|\int_X f(x)\overline{g(x)}\,d\mu(x)\Big| \leqslant \Big(\int_X |f(x)|^2\,d\mu(x)\Big)^{1/2} \Big(\int_X |g(x)|^2\,d\mu(x)\Big)^{1/2}.$$

(d) Niech ℓ_2 oznacza zbiór wszystkich ciągów $x = (\xi_n)$, dla których

$$\sum_{n=1}^{\infty} |\xi_n|^2 < \infty.$$

Jest to szczególny przypadek przykładu 2.1 (c), w którym μ oznacza miarę liczącą na zbiorze liczb naturalnych $\mathbb{N}$. Iloczyn skalarny określony jest w ℓ_2 wzorem

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \xi_n \overline{\eta_n}.$$

Z tak określonym iloczynem skalarnym przestrzeń ℓ_2 jest przestrzenią unitarną. Nierówność Schwarza ma w tej przestrzeni postać

$$\Big|\sum_{n=1}^{\infty} \xi_n \overline{\eta_n}\Big| \leqslant \Big(\sum_{n=1}^{\infty} |\xi_n|^2\Big)^{1/2} \Big(\sum_{n=1}^{\infty} |\eta_n|^2\Big)^{1/2}.$$

Twierdzenie 2.3. *Iloczyn skalarny w przestrzeni unitarnej X jest funkcjonałem ciągłym na produkcie $X \times X$, tzn. jeżeli $x_n \to x$ i $y_n \to y$, to $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

D o w ó d. Teza wynika z następujących nierówności:

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &\leqslant |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| = \\ &= |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \leqslant \\ &\leqslant \|x_n\|\,\|y_n - y\| + \|x_n - x\|\,\|y\|. \qquad \blacksquare\end{aligned}$$

Definicja 2.3. Przestrzeń unitarna X, która jest przestrzenią Banacha w normie $\|x\| = \sqrt{\langle x, x\rangle}$, nazywa się *przestrzenią Hilberta*.

Zauważmy, że przestrzenie z przykładów 2.1 (a), (c), (d) są przestrzeniami Hilberta, natomiast przestrzeń z przykładu 2.1 (b) nie jest taką przestrzenią.

2.2. Twierdzenie o rzucie ortogonalnym

Definicja 2.4. Mówimy, że dwa wektory x i y z przestrzeni unitarnej X są *ortogonalne* lub *prostopadłe* i piszemy $x \perp y$, jeżeli $\langle x, y\rangle = 0$. Jeżeli natomiast E jest podzbiorem przestrzeni X i wektor x jest ortogonalny do każdego $y \in E$, to mówimy, że x *jest ortogonalny do zbioru* E i piszemy $x \perp E$.

Twierdzenie 2.4 (twierdzenie Pitagorasa). *Jeżeli wektory x_1, x_2, ..., $x_n \in X$ są wzajemnie ortogonalne, to*

$$\|x_1\|^2 + \|x_2\|^2 + \ldots + \|x_n\|^2 = \|x_1 + x_2 + \ldots + x_n\|^2.$$

D o w ó d. Dowód prowadzimy przez indukcję względem liczby wektorów. Dla $n = 1$ twierdzenie jest oczywiste. Załóżmy więc, że teza zachodzi dla $n - 1$ i niech wektory x_1, x_2, ..., $x_n \in V$ będą wzajemnie ortogonalne. Oznaczmy przez y wektor $x_1 + x_2 + \ldots + x_{n-1}$. Wówczas mamy

$$\begin{aligned}\|y + x_n\|^2 &= \langle y + x_n, y + x_n\rangle = \langle y, y\rangle + 2\mathrm{Re}\,\langle y, x_n\rangle + \langle x_n, x_n\rangle = \\ &= \|x_1\|^2 + \|x_2\|^2 + \ldots + \|x_{n-1}\|^2 + \|x_n\|^2,\end{aligned}$$

ponieważ z założenia indukcyjnego

$$\|y\|^2 = \|x_1\|^2 + \|x_2\|^2 + \ldots + \|x_{n-1}\|^2$$

i ponadto

$$\langle y, x_n\rangle = \langle x_1, x_n\rangle + \langle x_2, x_n\rangle + \ldots + \langle x_{n-1}, x_n\rangle = 0. \qquad \blacksquare$$

Twierdzenie 2.5 (o rzucie ortogonalnym). *Niech M będzie domkniętą podprzestrzenią przestrzeni Hilberta H. Każdy wektor $x \in H$ ma przedstawienie w postaci*

$$x = x_0 + z, \quad \textit{gdzie} \;\; x_0 \in M \;\; \textit{i} \;\; z \perp M, \tag{2.4}$$

przy czym rozkład ten jest jednoznaczny.

Wektor x_0 nazywamy rzutem ortogonalnym *wektora x na podprzestrzeń M.*

D o w ó d. Jeżeli $x \in M$, to wystarczy przyjąć $x_0 = x$ i $z = 0$. Niech więc $x \in H \setminus M$. Oznaczmy

$$\rho = \inf\{\|x - y\| \colon \; y \in M\}.$$

Wówczas $\rho > 0$, bo podprzestrzeń M jest domknięta. Wybierzmy taki ciąg (y_n) wektorów z podprzestrzeni M, aby $\|x - y_n\| \to \rho$, gdy $n \to \infty$. Niech $\rho_n = \|x - y_n\|$. Wtedy $\rho_n \geqslant \rho$ dla $n = 1, 2, \ldots$ i $\rho_n \to \rho$, gdy $n \to \infty$. Dla dowolnego wektora $y \in M$ i dowolnej liczby λ mamy $y_n + \lambda y \in M$, więc

$$\|x - (y_n + \lambda y)\| \geqslant \rho \quad (n = 1, 2, \ldots). \tag{2.5}$$

Ustalmy wskaźnik n i przyjmijmy

$$\lambda = \frac{\langle x - y_n, y\rangle}{\|y\|^2}.$$

Zakładamy przy tym, że $y \neq 0$. Wówczas dla dowolnego n mamy

$$\begin{aligned} \|x - (y_n + \lambda y)\|^2 &= \langle (x - y_n) - \lambda y, (x - y_n) - \lambda y\rangle = \\ &= \|x - y_n\|^2 - \lambda\overline{\langle x - y_n, y\rangle} - \overline{\lambda}\langle x - y_n, y\rangle + |\lambda|^2\|y\|^2 = \\ &= \|x - y_n\|^2 - \frac{|\langle x - y_n, y\rangle|^2}{\|y\|^2}. \end{aligned}$$

Wobec (2.5)

$$\|x - y_n\|^2 - \frac{|\langle x - y_n, y\rangle|^2}{\|y\|^2} \geqslant \rho^2 \quad (n = 1, 2, \ldots),$$

a ponieważ $\|x - y_n\| = \rho_n$, więc mamy

$$|\langle x - y_n, y\rangle| \leqslant \sqrt{\rho_n^2 - \rho^2}\, \|y\| \quad (n = 1, 2, \ldots) \tag{2.6}$$

dla dowolnego $y \in M$. Zauważmy, że wzór ten jest również prawdziwy dla $y = 0$. Stąd

$$\begin{aligned}|\langle y_n - y_m, y\rangle| &\leqslant |\langle y_n - x, y\rangle| + |\langle x - y_m, y\rangle| \leqslant \\ &\leqslant \left(\sqrt{\rho_n^2 - \rho^2} + \sqrt{\rho_m^2 - \rho^2}\right) \|y\|\end{aligned}$$

dla n, $m = 1, 2, \ldots$ i dowolnego $y \in M$. Podstawiając $y = y_n - y_m$, otrzymujemy

$$\|y_n - y_m\| \leqslant \sqrt{\rho_n^2 - \rho^2} + \sqrt{\rho_m^2 - \rho^2} \quad (n,\, m = 1, 2, \ldots).$$

Ponieważ $\rho_n \to \rho$, więc ciąg (y_n) jest ciągiem Cauchy'ego. Wobec zupełności przestrzeni H jest on zbieżny. Niech $x_0 = \lim\limits_{n\to\infty} y_n$. Ponieważ podprzestrzeń M jest domknięta, więc $x_0 \in M$. Przechodząc do granicy, $n \to \infty$, w nierówności (2.6) otrzymujemy $\langle x - x_0, y\rangle = 0$ dla dowolnego $y \in M$. Przyjmując $z = x - x_0$, mamy $x = x_0 + z$, przy czym $x_0 \in M$, a $z \perp M$, czyli żądane przedstawienie wektora x. Pozostaje udowodnić jednoznaczność tego rozkładu. Niech $x = x_0' + z' = x_0'' + z''$, gdzie x_0', $x_0'' \in M$, z', $z'' \perp M$. Wówczas $x_0' - x_0'' = z'' - z'$ i $x_0' - x_0'' \in M$, $z'' - z' \perp M$. Stąd

$$\|x_0' - x_0''\|^2 = \langle x_0' - x_0'', z'' - z'\rangle = 0,$$

co oznacza, że $x_0' = x_0''$, a w konsekwencji także $z' = z''$. ■

Definicja 2.5. Jeżeli E jest podzbiorem przestrzeni unitarnej X, to zbiór

$$E^\perp = \{x \in X\colon\ x \perp y \text{ dla wszystkich } y \in E\}$$

nazywamy *dopełnieniem ortogonalnym* zbioru E. W przypadku, gdy $E = \{x\}$, będziemy pisać $x^\perp$ zamiast $\{x\}^\perp$.

Lemat 2.6. *$E^\perp$ jest domkniętą podprzestrzenią liniową przestrzeni X.*

Dowód. Zbiór $E^\perp$ jest niepusty, ponieważ $0 \in E^\perp$ (wektor 0 jest ortogonalny do każdego wektora). Natomiast fakt, że $\lambda x_1 + \mu x_2 \in E^\perp$, jeżeli tylko x_1 i $x_2 \in E^\perp$, wynika z własności (b) i (c) iloczynu skalarnego. Zatem $E^\perp$ jest podprzestrzenią przestrzeni X. Zbiór $x^\perp$ jest podprzestrzenią domkniętą dla dowolnego $x \in X$, bo jest on przeciwobrazem zbioru domkniętego $\{0\}$ poprzez funkcję ciągłą $y \mapsto \langle y, x\rangle$. Ponieważ $E^\perp = \bigcap\limits_{x\in E} x^\perp$, więc podprzestrzeń $E^\perp$ jest domknięta. ■

Twierdzenie o rzucie ortogonalnym mówi nam, że przestrzeń Hilberta H jest sumą prostą podprzestrzeni M i $M^\perp$, tzn.

$$X = M \oplus M^\perp.$$

Ponieważ podprzestrzenie M i $M^\perp$ są ortogonalne, więc ten przypadek sumy prostej nazywamy *ortogonalną sumą prostą*. W dalszym ciągu zapis $M_1 \oplus M_2$ będzie zawsze oznaczał sumę prostą tego rodzaju.

Podamy teraz wnioski z twierdzenia o rzucie ortogonalnym.

Wniosek 2.7. *Jeżeli x_0 jest rzutem ortogonalnym elementu x przestrzeni Hilberta H na podprzestrzeń M, to*

$$\|x - x_0\| \leqslant \|x - y\|$$

dla każdego $y \in M$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $y = x_0$.

D o w ó d. Niech y będzie dowolnym wektorem z podprzestrzeni M. Ponieważ $x - x_0 \perp M$ oraz $x_0 - y \in M$, więc $x - x_0 \perp x_0 - y$. Ponieważ $x - y = (x - x_0) + (x_0 - y)$, więc na mocy twierdzenia Pitagorasa zachodzi równość

$$\|x - y\|^2 = \|x - x_0\|^2 + \|x_0 - y\|^2,$$

z której wynika teza. ■

Wniosek 2.8. *Podprzestrzeń M przestrzeni Hilberta H jest gęsta w H wtedy i tylko wtedy, gdy $M^\perp = \{0\}$.*

D o w ó d. Jeżeli $\overline{M} = H$, to dla dowolnego $x \in H$ istnieje ciąg (y_n) wektorów z M taki, że $x = \lim\limits_{n\to\infty} y_n$. Niech $x \perp M$. Ponieważ $\langle x, y_n\rangle = 0$ dla wszystkich n, więc z ciągłości iloczynu skalarnego wynika, że $\|x\|^2 = \langle x, x\rangle = 0$. Zatem $x = 0$.

Na odwrót, jeżeli $\overline{M} \neq H$, to istnieje $x \in H \setminus \overline{M}$. Na mocy twierdzenia o rzucie ortogonalnym wektor x ma przedstawienie w postaci $x = x_0 + z$, gdzie $x_0 \in \overline{M}$ i $z \perp \overline{M}$. Zatem $z \in M^\perp$ i $z \neq 0$, czyli $M^\perp \neq \{0\}$. ■

Zauważmy, że wektor $x_0 \in M$ taki, że $z = x - x_0 \perp M$, jest wyznaczony jednoznacznie, zatem określona jest funkcja $P\colon\ H \to M$, $P(x) = x_0$. Funkcję tę nazywamy *rzutowaniem* (*ortogonalnym*) *na podprzestrzeń* M.

Wniosek 2.9 (własności rzutowania na podprzestrzeń). *Dla rzutowania P na* (*domkniętą*) *podprzestrzeń M przestrzeni Hilberta H i x, $y \in H$ oraz α, $\beta \in \mathbb{C}$ spełnione są następujące warunki*:

(a) $\langle Px, y\rangle = \langle Px, Py\rangle = \langle x, Py\rangle$;

(b) $\langle P(Px), y\rangle = \langle Px, y\rangle$;

(c) $\langle Px, x\rangle = \|Px\|^2$;

(d) $\|Px\| \leqslant \|x\|$;

(e) $\|x\|^2 = \|x - Px\|^2 + \|Px\|^2$;

(f) $M = \{x \in H\colon\ Px = x\} = \{x \in H\colon\ \|Px\| = \|x\|\}$;

(g) $Px = 0$ *wtedy i tylko wtedy, gdy* $x \perp M$;

(h) $P(\alpha x + \beta y) = \alpha Px + \beta Py$;

(i) $P(H) = M$.

Łatwe dowody powyższych własności pozostawiamy jako ćwiczenie dla Czytelnika.

Z liniowości iloczynu skalarnego ze względu na pierwszy argument wynika, że dla ustalonego wektora a przestrzeni unitarnej X funkcjonał $f\colon X \to \mathbb{C}$, określony wzorem $f(x) = \langle x, a\rangle$, jest funkcjonałem liniowym na przestrzeni X. Ponadto z nierówności Schwarza wynika, że funkcjonał ten jest ograniczony. Mamy bowiem

$$|f(x)| = |\langle x, a\rangle| \leqslant \|a\|\,\|x\|. \tag{2.7}$$

Zauważmy ponadto, że wówczas $\|f\| = \|a\|$. Dla $a = 0$ jest to oczywiste. Jeżeli $a \neq 0$, to z nierówności (2.7) wynika, że $\|f\| \leqslant \|a\|$, a ponadto $f\left(\frac{1}{\|a\|}\,a\right) = \|a\|$. Okazuje się, że każdy funkcjonał liniowy ograniczony na przestrzeni Hilberta jest takiej postaci.

Twierdzenie 2.10 (Riesza).** *Jeżeli f jest funkcjonałem liniowym ograniczonym na przestrzeni Hilberta H, to istnieje dokładnie jeden wektor $a \in H$ taki, że*

$$f(x) = \langle x, a\rangle \quad \textit{dla każdego } x \in H. \tag{2.8}$$

Ponadto

$$\|f\| = \|a\|.$$

D o w ó d. Aby otrzymać jednoznaczność wektora a w reprezentacji (2.8), przypuśćmy, że $f(x) = \langle x, \widetilde{a}\rangle$ dla każdego wektora $x \in H$. Wówczas $0 = \langle x, a - \widetilde{a}\rangle$ i biorąc $x = a - \widetilde{a}$, otrzymujemy $\|a - \widetilde{a}\| = 0$, czyli $a = \widetilde{a}$.

** Autor chciałby podkreślić, że wszystkie twierdzenia zamieszczone w skrypcie i opatrzone nazwiskiem Riesz są autorstwa Frigyesa Riesza.

Jeżeli $f(x) = 0$ dla każdego wektora $x \in H$, to wystarczy przyjąć $a = 0$. Przypuśćmy zatem, że $f(y) \neq 0$ dla pewnego wektora $y \in H$. Ponieważ f jest funkcjonałem liniowym ograniczonym, więc $M = \ker f = f^{-1}(\{0\})$ jest domkniętą podprzestrzenią przestrzeni H. Z twierdzenia o rzucie ortogonalnym otrzymujemy rozkład wektora y w postaci

$$y = y_0 + y_1, \quad \text{gdzie} \;\; y_0 \in M, \;\; y_1 \perp M.$$

Mamy przy tym $f(y_1) \neq 0$. Oznaczając $z = \dfrac{1}{f(y_1)}\, y_1$, otrzymujemy $f(z) = 1$. Wówczas dla dowolnego wektora $x \in H$ mamy $f(x - f(x)\, z) = 0$, czyli $x - f(x)\, z \in M$. Stąd wynika, że wektory $x - f(x)\, z$ i z są ortogonalne. Zatem

$$\langle x, z\rangle = \langle x - f(x)\, z, z\rangle + f(x)\,\langle z, z\rangle = f(x)\, \|z\|^2.$$

Przyjmując $a = \dfrac{1}{\|z\|^2}\, z$, otrzymujemy $f(x) = \langle x, a\rangle$. ■

2.3. Układy ortonormalne

Definicja 2.6. Niech H będzie przestrzenią Hilberta. Podzbiór $E \subset H$ nazywamy *zbiorem ortonormalnym*, jeżeli

— $\|e\| = 1$ dla każdego $e \in E$;
— $\langle e, f\rangle = 0$ dla dowolnych $e,\ f \in E$, $e \neq f$.

Tak więc zbiór jest ortonormalny, jeżeli składa się z wektorów jednostkowych (tzn. wektorów o długości jeden) parami ortogonalnych. Zamiast terminu *zbiór ortonormalny* będziemy również używać określenia *układ ortonormalny*. Jeżeli zbiór E jest skończony, tzn. $E = \{e_1, \ldots, e_n\}$, to będziemy krótko mówić, że *wektory* $e_1, \ldots, e_n$ *są ortonormalne.*

Przykłady 2.2. (a) W przestrzeni $\mathbb{C}^n$ wektory $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, $j = 1, 2, \ldots, n$, tworzą układ ortonormalny.

(b) W przestrzeni ℓ_2 ciąg elementów $e_n = (\delta_{nk})_{k=1}^{\infty}$, $n = 1, 2, \ldots$, jest układem ortonormalnym.

(c) W przestrzeni $L_2[-\pi, \pi]$ ciąg funkcji

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}} \cos nx, \quad \frac{1}{\sqrt{\pi}} \sin nx \qquad (n = 1, 2, \ldots)$$

jest układem ortonormalnym.

(d) Innym przykładem zbioru ortonormalnego w przestrzeni $L_2[-\pi, \pi]$ jest układ funkcji

$$\frac{1}{\sqrt{2\pi}} e^{inx} \qquad (n = \pm 0, \pm 1, \pm 2, \ldots).$$

Jest oczywiste, że każdy zbiór ortonormalny składa się z wektorów liniowo niezależnych. Istotnie, jeżeli $e_1, \ldots, e_n \in E$, $\alpha_1 e_1 + \ldots + \alpha_n e_n = 0$, to

$$\alpha_j = \langle \alpha_1 e_1 + \ldots + \alpha_n e_n, e_j \rangle = 0 \quad \text{dla} \quad j = 1, 2, \ldots, n.$$

Twierdzenie 2.11 (ortogonalizacja Grama-Schmidta). *Jeżeli H jest przestrzenią Hilberta i $\{x_n\colon n = 1, 2, \ldots\}$ jest podzbiorem liniowo niezależnym, to istnieje układ ortonormalny $\{e_n\colon n = 1, 2, \ldots\}$ w przestrzeni H taki, że dla każdego n*

$$\operatorname{span}\{e_1, e_2, \ldots, e_n\} = \operatorname{span}\{x_1, x_2, \ldots, x_n\}.$$

D o w ó d. Definiujemy

$$e_1 = \frac{x_1}{\|x_1\|}, \quad e_2 = \frac{y_2}{\|y_2\|}, \quad \text{gdzie } \ y_2 = x_2 - \langle x_2, e_1 \rangle e_1.$$

Z liniowej niezależności wektorów x_1 i x_2 wynika, że $x_1 \neq 0$ i $y_2 \neq 0$. Mamy $\|e_1\| = \|e_2\| = 1$ i $\langle e_1, e_2 \rangle = 0$.

Jeżeli już zostały zdefiniowane wektory $e_1, e_2, \ldots, e_n$, to przyjmujemy

$$e_{n+1} = \frac{y_{n+1}}{\|y_{n+1}\|}, \quad \text{gdzie } \ y_{n+1} = x_{n+1} - \sum_{k=1}^{n} \langle x_{n+1}, e_k \rangle e_k.$$

Zauważmy, że $y_{n+1} \neq 0$, bo w przeciwnym razie wektory $x_1, x_2, \ldots, x_{n+1}$ byłyby liniowo zależne. Ponadto mamy $\|e_{n+1}\| = 1$ i $\langle e_k, e_{n+1} \rangle = 0$ dla $k = 1, 2, \ldots, n$. Ponieważ $e_n \in \operatorname{span}\{x_1, x_2, \ldots, x_n\}$ i $x_n \in \operatorname{span}\{e_1, e_2, \ldots, e_n\}$ dla dowolnego n, więc stąd wynika teza. ■

Definicja 2.7. Iloczyn skalarny $\langle x, e \rangle$ nazywamy *współczynnikiem Fouriera* wektora x względem elementu e.

Twierdzenie 2.12 (nierówność Bessela). *Jeżeli wektory $e_1, \ldots, e_n$ są ortonormalne, to*

$$\sum_{j=1}^{n} |\langle x, e_j \rangle|^2 \leqslant \|x\|^2.$$

D o w ó d. Oznaczmy $y = x - \sum\limits_{j=1}^{n} \langle x, e_j \rangle e_j$. Wówczas $\langle y, e_k \rangle = 0$ dla $k = 1, 2, \ldots, n$. Zatem wektory $\sum\limits_{j=1}^{n} \langle x, e_j \rangle e_j$ i y są ortogonalne. Na mocy twierdzenia Pitagorasa mamy

$$\|x\|^2 = \|y\|^2 + \left\| \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\|^2 = \|y\|^2 + \sum_{j=1}^{n} \|\langle x, e_j \rangle e_j\|^2 = \\ = \|y\|^2 + \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 .$$

Stąd

$$\|x\|^2 \geqslant \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 . \qquad \blacksquare$$

Z nierówności Bessela otrzymujemy:

Wniosek 2.13. *Każdy wektor x ma co najwyżej przeliczalnie wiele niezerowych współczynników Fouriera względem ustalonego układu ortonormalnego.*

D o w ó d. Niech E będzie układem ortonormalnym oraz niech

$$E_n = \left\{ e \in E\colon\ |\langle x, e \rangle| \geqslant \frac{1}{n} \right\}.$$

Z nierówności Bessela wynika, że E_n jest zbiorem skończonym. Ale

$$\bigcup_{n=1}^{\infty} E_n = \{ e \in E\colon\ \langle x, e \rangle \neq 0 \}. \qquad \blacksquare$$

Twierdzenie 2.14. *Jeżeli (e_n) jest przeliczalnym układem ortonormalnym w przestrzeni Hilberta H, to szereg*

$$\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$

jest zbieżny w H i różnica $x - \sum\limits_{n=1}^{\infty} \langle x, e_n \rangle e_n$ jest wektorem ortogonalnym do wszystkich wektorów e_j, $j = 1, 2, \ldots$

D o w ó d. Z nierówności Bessela wynika, że szereg $\sum\limits_{n=1}^{\infty} |\langle x, e_n \rangle|^2$ jest zbieżny. Ponieważ

$$\left\| \sum_{k=n}^{m} \langle x, e_k \rangle e_k \right\|^2 = \sum_{k=n}^{m} |\langle x, e_k \rangle|^2 ,$$

więc stąd wynika, że i szereg $\sum\limits_{n=1}^{\infty} \langle x, e_n\rangle e_n$ jest zbieżny, bo spełnia warunek Cauchy'ego. Ponadto

$$\left\langle x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k, e_j \right\rangle = 0 \quad \text{dla} \quad n > j.$$

Z ciągłości iloczynu skalarnego otrzymujemy

$$\left\langle x - \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k, e_j \right\rangle = 0 \quad \text{dla} \quad j = 1, 2, \ldots \qquad \blacksquare$$

Aby rozważać ortogonalne rodziny wektorów dowolnej mocy, w teorii przestrzeni Hilberta wprowadza się pojęcie sumy dowolnej niekoniecznie skończonej czy też przeliczalnej rodziny wektorów.

Definicja 2.8. Niech $\{x_\alpha\colon \alpha \in A\}$ będzie rodziną wektorów z przestrzeni Hilberta H indeksowaną przez dowolny niepusty zbiór A. Mówimy, że rodzina $\{x_\alpha\colon \alpha \in A\}$ jest *sumowalna*, jeżeli w H istnieje wektor x taki, że dla dowolnego $\varepsilon > 0$ istnieje zbiór skończony $J_0 \subset A$ o tej własności, że

$$\left\| x - \sum_{\alpha \in J} x_\alpha \right\| < \varepsilon$$

dla dowolnego skończonego podzbioru J zbioru A, dla którego $J_0 \subset J$. Wówczas wektor x nazywamy *sumą* rodziny $\{x_\alpha\colon \alpha \in A\}$ i piszemy $\sum\limits_{\alpha \in A} x_\alpha = x$.

Uwagi. 1. Jeżeli rodzina A jest skończona, $A = \{\alpha_1, \ldots, \alpha_n\}$, to jest ona zawsze sumowalna i jej suma jest „zwykłą” sumą wektorów $x_{\alpha_1} + \ldots + x_{\alpha_n}$.

2. Definicja sumowalności ma sens również wtedy, gdy rozważamy zbiory złożone z liczb zespolonych. W szczególności, gdy zbiór A jest przeliczalny, np. $A = \mathbb{N}$, to zachodzi następujące twierdzenie:

Twierdzenie 2.15. *Rodzina* $\{\xi_n\colon \xi_n \in \mathbb{C},\ n \in \mathbb{N}\}$ *jest sumowalna i jej suma jest równa* ξ *wtedy i tylko wtedy, gdy szereg* $\sum\limits_{n=1}^{\infty} \xi_n$ *jest bezwzględnie zbieżny i jego suma jest równa* ξ.

Proponujemy, aby Czytelnik przeprowadził dowód tego twierdzenia jako ćwiczenie.

3. W przypadku dowolnej przestrzeni Hilberta H i zbioru $A = \mathbb{N}$ zbieżność szeregu $\sum\limits_{n=1}^{\infty} \|x_n\|$ zapewnia sumowalność rodziny $\{x_n\colon n \in \mathbb{N}\}$, ale nie na odwrót.

W dalszym ciągu dla skrócenia oznaczeń będziemy pisać $\{x_\alpha\}$ i $\sum\limits_\alpha x_\alpha$ zamiast $\{x_\alpha \in A\}$ oraz $\sum\limits_{\alpha\in A} x_\alpha$.

Uogólnione sumowanie związane jest z operacjami liniowymi i iloczynem skalarnym w następujący sposób:

Twierdzenie 2.16. *Niech $\{x_\alpha\}$ i $\{y_\alpha\}$ będą dwiema rodzinami wektorów z przestrzeni Hilberta H (indeksowanymi przez ten sam zbiór) i niech λ będzie dowolną liczbą zespoloną. Wówczas*:

(a) *jeżeli $\sum\limits_\alpha x_\alpha = x$, to $\sum\limits_\alpha \lambda x_\alpha = \lambda x$;*

(b) *jeżeli $\sum\limits_\alpha x_\alpha = x$ i $\sum\limits_\alpha y_\alpha = y$, to $\sum\limits_\alpha (x_\alpha + y_\alpha) = x + y$;*

(c) *jeżeli $\sum\limits_\alpha x_\alpha = x$, to $\sum\limits_\alpha \langle x_\alpha, y\rangle = \langle x, y\rangle$ oraz $\sum\limits_\alpha \langle y, x_\alpha\rangle = \langle y, x\rangle$ dla dowolnego wektora $y \in H$.*

D o w ó d. Teza wynika z następujących nierówności prawdziwych dla dowolnego skończonego zbioru indeksów J:

$$\Big\|\lambda x - \sum_{\alpha\in J} \lambda x_\alpha\Big\| = |\lambda| \Big\|x - \sum_{\alpha\in J} x_\alpha\Big\|,$$

$$\Big\|(x+y) - \sum_{\alpha\in J}(x_\alpha + y_\alpha)\Big\| \leqslant \Big\|x - \sum_{\alpha\in J} x_\alpha\Big\| + \Big\|y - \sum_{\alpha\in J} y_\alpha\Big\|,$$

$$\Big\|\langle x, y\rangle - \sum_{\alpha\in J}\langle x_\alpha, y\rangle\| = \Big|\Big\langle x - \sum_{\alpha\in J} x_\alpha, y\Big\rangle\Big| \leqslant \Big\|x - \sum_{\alpha\in J} x_\alpha\Big\| \, \|y\|. \qquad \blacksquare$$

Badanie sumowalności ułatwia następujące twierdzenie:

Twierdzenie 2.17 (kryterium sumowalności). *Rodzina $\{x_\alpha\}$ wektorów z przestrzeni Hilberta jest sumowalna wtedy i tylko wtedy, gdy dla dowolnego $\varepsilon > 0$ istnieje zbiór skończony indeksów J_0 taki, że $\Big\|\sum\limits_{\alpha\in J} x_\alpha\Big\| < \varepsilon$ dla dowolnego skończonego zbioru indeksów J rozłącznego z J_0.*

D o w ó d. Załóżmy, że rodzina $\{x_\alpha\}$ jest sumowalna i $\sum\limits_\alpha x_\alpha = x$. Wówczas dla dowolnego $\varepsilon > 0$ istnieje zbiór skończony J_0 taki, że $\Big\|x - \sum\limits_{\alpha\in J} x_\alpha\Big\| < \frac{1}{2}\,\varepsilon$

dla dowolnego skończonego zbioru indeksów $J \supset J_0$. Zatem, jeżeli $J \cap J_0 = \emptyset$, to

$$\Big\|\sum_{\alpha\in J} x_\alpha\Big\| = \Big\|\sum_{\alpha\in J\cup J_0} x_\alpha - \sum_{\alpha\in J_0} x_\alpha\Big\| \leqslant$$
$$\leqslant \Big\|x - \sum_{\alpha\in J\cup J_0} x_\alpha\Big\| + \Big\|x - \sum_{\alpha\in J_0} x_\alpha\Big\| < \varepsilon.$$

Na odwrót, jeżeli warunek z twierdzenia jest spełniony, to dla dowolnego naturalnego n istnieje zbiór skończony J_n taki, że $\Big\|\sum\limits_{\alpha\in J} x_\alpha\Big\| < \dfrac{1}{n}$ dla dowolnego zbioru skończonego J takiego, że $J \cap J_n = \emptyset$. Biorąc $K_n = J_1 \cup J_2 \cup \ldots \cup J_n$, otrzymujemy wstępującą rodzinę zbiorów skończonych $\{K_n\}$ o tej samej własności, jaką miała rodzina $\{J_n\}$. Zauważmy teraz, że jeżeli $n < m$, to

$$\Big\|\sum_{\alpha\in K_m} x_\alpha - \sum_{\alpha\in K_n} x_\alpha\Big\| = \Big\|\sum_{\alpha\in K_m\setminus K_n} x_\alpha\Big\| < \frac{1}{n},$$

bo $(K_m \setminus K_n) \cap K_n = \emptyset$. Zatem ciąg $\Big(\sum\limits_{\alpha\in K_n} x_\alpha\Big)$ jest ciągiem Cauchy'ego. Z zupełności przestrzeni Hilberta istnieje wektor x taki, że $\Big\|\sum\limits_{\alpha\in K_n} x_\alpha - x\Big\| \to 0$, gdy $n \to \infty$. Jeżeli J jest dowolnym skończonym zbiorem indeksów zawierającym zbiór K_m, to

$$\Big\|x - \sum_{\alpha\in J} x_\alpha\Big\| \leqslant \Big\|x - \sum_{\alpha\in K_m} x_\alpha\Big\| + \Big\|\sum_{\alpha\in J\setminus K_m} x_\alpha\Big\|$$

i stąd wynika, że $\{x_\alpha\}$ jest sumowalne oraz $x = \sum\limits_{\alpha} x_\alpha$. ■

Wniosek 2.18. *Jeżeli rodzina $\{x_\alpha\}$ jest sumowalna, to zbiór $\{\alpha\colon\ x_\alpha \neq 0\}$ jest co najwyżej przeliczalny.*

D o w ó d. Wybieramy wstępującą rodzinę zbiorów $\{K_n\}$ taką jak w dowodzie poprzedniego twierdzenia. Wówczas zbiór $K = \bigcup\limits_{n=1}^{\infty} K_n$ jest co najwyżej przeliczalny. Jeżeli $\alpha \notin K$, to $\|x_\alpha\| < \dfrac{1}{n}$ dla dowolnego naturalnego n, a więc $x_\alpha = 0$. ■

Twierdzenie 2.19. *Ortogonalna rodzina wektorów* $\{x_\alpha\}$ *z przestrzeni Hilberta jest sumowalna wtedy i tylko wtedy, gdy sumowalna jest rodzina liczb* $\{\|x_\alpha\|^2\}$. *Jeżeli* $x = \sum_\alpha x_\alpha$, *to*

$$\|x\|^2 = \sum_\alpha \|x_\alpha\|^2. \tag{2.9}$$

D o w ó d. Jeżeli rodzina $\{x_\alpha\}$ jest sumowalna, to na mocy kryterium sumowalności dla dowolnego $\varepsilon > 0$ istnieje skończony zbiór indeksów J_0 taki, że $\left\|\sum_{\alpha\in J} x_\alpha\right\| < \varepsilon$ dla każdego skończonego zbioru indeksów J takiego, że $J \cap J_0 = \emptyset$. Stąd na mocy twierdzenia Pitagorasa otrzymujemy

$$\sum_{\alpha\in J} \|x_\alpha\|^2 = \left\|\sum_{\alpha\in J} x_\alpha\right\|^2 < \varepsilon^2 \tag{2.10}$$

i ponownie z kryterium sumowalności wynika, że rodzina $\{\|x_\alpha\|^2\}$ jest sumowalna.

Na odwrót, z sumowalności rodziny $\{\|x_\alpha\|^2\}$ dla dowolnego $\varepsilon > 0$ otrzymujemy skończony zbiór indeksów J_0 taki, że $\sum_{\alpha\in J} \|x_\alpha\|^2 < \varepsilon^2$, jeżeli tylko $J \cap J_0 = \emptyset$. Wówczas z (2.10) wynika, że $\left\|\sum_{\alpha\in J} x_\alpha\right\| < \varepsilon$, co na mocy twierdzenia 2.17 oznacza sumowalność rodziny $\{x_\alpha\}$.

Ostatnia część tezy wynika z twierdzenia 2.16 (c) oraz równości

$$\begin{aligned}\|x\|^2 &= \langle x, x\rangle = \left\langle \sum_\alpha x_\alpha, x\right\rangle = \sum_\alpha \langle x_\alpha, x\rangle = \sum_\alpha \left\langle x_\alpha, \sum_\beta x_\beta \right\rangle = \\ &= \sum_\alpha \sum_\beta \langle x_\alpha, x_\beta\rangle = \sum_\alpha \langle x_\alpha, x_\alpha\rangle = \sum_\alpha \|x_\alpha\|^2. \qquad \blacksquare\end{aligned}$$

U w a g a. Równość (2.9) możemy traktować jako uogólnienie twierdzenia Pitagorasa na nieskończone rodziny zbiorów ortogonalnych.

Definicja 2.9. Układ ortonormalny jest *zupełny*, jeżeli nie istnieje niezerowy wektor, który jest ortogonalny do wszystkich wektorów tego ukladu. Innymi słowy, układ jest zupełny, jeżeli nie może być rozszerzony do większego układu ortonormalnego.

Twierdzenie 2.20. *Układ ortonormalny jest zupełny wtedy i tylko wtedy, gdy podprzestrzeń przez niego generowana jest gęsta.*

D o w ó d. Jeżeli $\overline{\operatorname{span}\{e_\alpha\colon \alpha \in A\}} = H$, to $\langle x, e_\alpha\rangle = 0$ dla każdego $\alpha \in A$ implikuje $x = 0$, a więc układ jest zupełny.

Na odwrót, jeżeli $H_0 = \overline{\operatorname{span}\{e_\alpha\colon \alpha \in A\}}$ jest właściwą podprzestrzenią przestrzeni H, to istnieje wektor $x \neq 0$ taki, że $x \perp H_0$, a więc układ $\{e_\alpha\colon \alpha \in A\}$ nie jest zupełny. ■

U w a g a. Układ ortonormalny zupełny nazywa się również *bazą* (*ortonormalną*) *przestrzeni Hilberta*. W przypadku, gdy $\dim H < \infty$ (wymiar przestrzeni liniowej), to baza ortonormalna jest również bazą Hamela. Gdy $\dim H = \infty$, to baza ortonormalna nie musi być bazą Hamela.

Twierdzenie 2.21. *Każda niezerowa przestrzeń Hilberta ma bazę.*

D o w ó d. Rozważymy rodzinę $\mathcal{E}$ wszystkich układów ortonormalnych w przestrzeni Hilberta $H \neq \{0\}$. Oczywiście rodzina ta jest niepusta. Jest ona częściowo uporządkowana przez inkluzję. Jeżeli (E_γ) jest łańcuchem w $\mathcal{E}$, to $E = \bigcup\limits_\gamma E_\gamma$ jest układem ortonormalnym górującym nad tym łańcuchem. Zatem na mocy lematu Kuratowskiego-Zorna rodzina $\mathcal{E}$ ma elementy maksymalne. Taki element maksymalny jest oczywiście układem ortonormalnym zupełnym. ■

W przypadku przestrzeni ośrodkowej możemy się obejść bez pewnika wyboru. Prawdziwe jest następujące twierdzenie:

Twierdzenie 2.22. *W ośrodkowej przestrzeni Hilberta istnieje co najwyżej przeliczalny układ ortonormalny zupełny.*

D o w ó d. Niech (x_n) będzie gęstym ciągiem w przestrzeni H. Z tego ciągu wyrzucamy każdy wektor x_n, który jest kombinacją liniową poprzedzających go wektorów $x_1, x_2, \ldots, x_{n-1}$. Otrzymujemy w ten sposób nowy skończony lub nieskończony ciąg, który oznaczymy przez (y_n), o tej własności, że dowolny wektor y_n nie jest kombinacją liniową wektorów go poprzedzających i $\operatorname{span}\{y_n\colon n \in \mathbb{N}\} \supset \{x_n\colon n \in \mathbb{N}\}$. Zatem $\overline{\operatorname{span}\{y_n\colon n \in \mathbb{N}\}} = H$.

Niech $M_n = \operatorname{span}\{y_1, y_2, \ldots, y_n\}$. Oznaczmy przez y'_{n+1} rzut ortogonalny wektora y_{n+1} na M_n i ponadto niech

$$e_1 = \frac{1}{\|y_1\|}\, y_1, \quad e_{n+1} = \frac{1}{\|y_{n+1} - y'_{n+1}\|}\,(y_{n+1} - y'_{n+1}), \quad n = 1, 2, \ldots$$

Wówczas wektory $\{e_n\}$ tworzą skończony lub nieskończony układ ortonormalny, którego powłoka liniowa pokrywa się z powłoką liniową wektorów $\{y_n\}$, a zatem jest gęsta w H. Na mocy twierdzenia 2.20 układ $\{e_n\}$ jest zupełny. ■

Prawdziwe jest również twierdzenie odwrotne, tzn. *jeżeli istnieje skończony lub przeliczalny układ ortonormalny zupełny* $\{e_n\}$ *w przestrzeni Hilberta* H*, to* H *jest przestrzenią ośrodkową.*

Istotnie, kombinacje liniowe wektorów e_n o współczynnikach wymiernych tworzą układ przeliczalny, który jest gęsty w H.

Twierdzenie 2.23. *Niech* $\{e_\alpha\colon \alpha \in A\}$ *będzie układem ortonormalnym w przestrzeni Hilberta* H*. Następujące warunki są równoważne*:

(a) *układ* $\{e_\alpha\colon \alpha \in A\}$ *jest zupełny w* H;

(b) *każdy wektor* $x \in H$ *ma przedstawienie w postaci*

$$x = \sum_\alpha \langle x, e_\alpha\rangle e_\alpha;$$

(c) *dla dowolnych wektorów* x, $y \in H$ *zachodzi równość*

$$\langle x, y\rangle = \sum_\alpha \langle x, e_\alpha\rangle\langle e_\alpha, y\rangle;$$

(d) *dla każdego* $x \in H$ *mamy*

$$\|x\|^2 = \sum_\alpha |\langle x, e_\alpha\rangle|^2 \quad (\textit{tożsamość Parsevala}).$$

D o w ó d. Pokażemy, że (b) jest konsekwencją (a). Niech $\{e_\alpha\colon \alpha \in A\}$ będzie układem zupełnym. Wówczas wektor $x - \sum_\alpha \langle x, e_\alpha\rangle e_\alpha$ jest ortogonalny do każdego e_α, a zatem jest równy zeru.

Następnie udowodnimy, że z warunku (b) wynika (c). Załóżmy, że zachodzi (b). Wówczas dla x, $y \in H$ na mocy ciągłości iloczynu skalarnego mamy

$$\begin{aligned}\langle x, y\rangle &= \left\langle \sum_\alpha \langle x, e_\alpha\rangle e_\alpha, \sum_\beta \langle y, e_\beta\rangle e_\beta \right\rangle = \\ &= \sum_\alpha \langle x, e_\alpha\rangle \overline{\langle y, e_\alpha\rangle}.\end{aligned}$$

Teraz pokażemy, że warunek (c) implikuje (d). Jeżeli zachodzi (c), to podstawiając $x = y$ do równości

$$\langle x, y\rangle = \sum_\alpha \langle x, e_\alpha\rangle\langle e_\alpha, y\rangle,$$

otrzymujemy tożsamość Parsevala.

W końcu udowodnimy, że warunek (a) wynika z warunku (d). Przypuśćmy, że (d) zachodzi. Niech x będzie wektorem ortogonalnym do wszystkich e_α. Wówczas $\langle x, e_\alpha\rangle = 0$ dla $\alpha \in A$ i z równości $\|x\|^2 = \sum_\alpha |\langle x, e_\alpha\rangle|^2$ wynika, że $x = 0$, czyli układ $\{e_\alpha\colon\ \alpha \in A\}$ jest zupełny. ■

U w a g a. Jeżeli układ ortonormalny $\{e_\alpha\}$ jest przeliczalny, to uogólnione sumy w warunkach (b), (c) i (d) stają się sumami szeregów zbieżnych (por. tw. 2.14).

Można udowodnić następujące twierdzenie:

Twierdzenie 2.24. *Wszystkie układy ortonormalne zupełne w przestrzeni Hilberta mają tę samą moc.*

Definicja 2.10. Moc układu ortonormalnego zupełnego nazywamy *wymiarem przestrzeni Hilberta*.

Definicja 2.11. Niech H i K będą przestrzeniami Hilberta. *Izomorfizmem przestrzeni Hilberta* nazywamy przekształcenie $U\colon\ H \to K$, które ma następujące własności:

— U jest suriekcją,
— $\langle Ux, Uy\rangle = \langle x, y\rangle$ dla dowolnych x, $y \in H$.

Prawdziwe jest następujące twierdzenie:

Twierdzenie 2.25. *Dwie przestrzenie Hilberta są izomorficzne wtedy i tylko wtedy, gdy mają ten sam wymiar.*

Z twierdzeń 2.22 i 2.23 wynika jako bezpośredni wniosek następujące twierdzenie:

Twierdzenie 2.26 (Riesza-Fischera). *Nieskończenie wymiarowa ośrodkowa przestrzeń Hilberta jest izomorficzna z przestrzenią ℓ_2.*

D o w ó d. Niech H będzie nieskończenie wymiarową i ośrodkową przestrzenią Hilberta i niech $\{e_n\colon\ n = 1, 2, \dots\}$ będzie jej bazą ortonormalną. Funkcję $U\colon\ H \to \ell_2$ określamy za pomocą wzoru $Ux = (\langle x, e_n\rangle)_{n=1}^{\infty}$. Z twierdzenia 2.23 (c) mamy

$$\langle Ux, Uy\rangle = \sum_{n=1}^{\infty} \langle x, e_n\rangle \overline{\langle y, e_n\rangle} = \sum_{n=1}^{\infty} \langle x, e_n\rangle \langle e_n, y\rangle = \langle x, y\rangle.$$

Pozostaje wykazać, że odwzorowanie U jest suriekcją. Niech $(c_n) \in \ell_2$ i niech $x_n = c_1 e_1 + \ldots + c_n e_n$ dla $n = 1, 2, \ldots$ Jeżeli $n > m$, to mamy

$$\begin{aligned}\|x_n - x_m\|^2 &= \|c_{m+1}e_{m+1} + \ldots + c_n e_n\|^2 = \\ &= \langle c_{m+1}e_{m+1} + \ldots + c_n e_n, c_{m+1}e_{m+1} + \ldots + c_n e_n\rangle = \\ &= |c_{m+1}|^2 + \ldots + |c_n|^2.\end{aligned}$$

Ze zbieżności szeregu $\sum\limits_{n=1}^{\infty} |c_n|^2$ wynika, że ciąg (x_n) jest ciągiem Cauchy'ego w przestrzeni H. Zatem istnieje wektor $x \in H$ taki, że $\lim\limits_{n\to\infty} \|x - x_n\| = 0$. Ustalmy dowolne $k \in \mathbb{N}$ i niech $n > k$. Wówczas mamy

$$\begin{aligned}\langle x, e_k\rangle &= \langle x - x_n, e_k\rangle + \langle x_n, e_k\rangle = \\ &= \langle x - x_n, e_k\rangle + \langle c_1 e_1 + \ldots + c_n e_n, e_k\rangle = \langle x - x_n, e_k\rangle + c_k.\end{aligned}$$

Stąd

$$|\langle x, e_k\rangle - c_k| \leqslant \|x - x_n\|$$

i po przejściu do granicy, $n \to \infty$, otrzymujemy równość $\langle x, e_k\rangle = c_k$, która oznacza suriektywność odwzorowania U. ■

Jest oczywiste, że układy ortonormalne z przykładów 2.2 (a) i (b) są bazami ortonormalnymi. Teraz pokażemy, że układ ortonormalny z przykładu 2.2 (c), czyli ciąg funkcji

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}} \cos nx, \quad \frac{1}{\sqrt{\pi}} \sin nx \qquad (n = 1, 2, \ldots), \tag{2.11}$$

jest układem ortonormalnym zupełnym w przestrzeni $L_2[-\pi, \pi]$.

Z twierdzenia 1.8 wynika, że funkcje ciągłe na przedziale $[-\pi, \pi]$ tworzą gęsty podzbiór przestrzeni $L_2[-\pi, \pi]$. Zauważmy, że dowolną funkcję ciągłą f na tym przedziale można przybliżać (w normie przestrzeni funkcji całkowalnych z kwadratem) funkcjami ciągłymi g, dla których $g(-\pi) = g(\pi)$. Z kolei na mocy twierdzenia Fejéra (patrz [27], tw. 7.25 z części II) funkcja g jest granicą jednostajnie zbieżnego ciągu wielomianów trygonometrycznych (a więc kombinacji liniowych funkcji tworzących układ (2.11)). Zatem podprzestrzeń liniowa generowana przez układ (2.11) jest zbiorem gęstym w przestrzeni $L_2[-\pi, \pi]$, a więc na mocy twierdzenia 2.20 stanowi jej bazę ortonormalną.

Ze wzorów Eulera (patrz wzory (7.14), [27], część II) wynika, że układ funkcji

$$\frac{1}{\sqrt{2\pi}} e^{inx} \qquad (n = \pm 0, \pm 1, \pm 2, \ldots)$$

jest również bazą ortonormalną przestrzeni $L_2[-\pi, \pi]$.

Przykłady 2.3. (a) Kombinacje liniowe funkcji

$$1, x, x^2, \dots, x^n, \dots \tag{2.12}$$

tworzą zbiór wszystkich wielomianów. Ponadto są one liniowo niezależne w dowolnej przestrzeni $L_2[a,b]$, gdzie $[a,b]$ jest dowolnym skończonym przedziałem. Ortonormalizując ten układ na przedziale $[-1,1]$, otrzymujemy układ ortonormalny

$$Q_0(x), Q_1(x), Q_2(x), \dots, Q_n(x), \dots$$

Zauważmy, że każda z funkcji Q_n jest wielomianem stopnia n. Układ ten jest zupełny, bo zbiór wszystkich wielomianów jest gęsty w przestrzeni $L_2[-1,1]$, a każdy wielomian jest liniową kombinacją funkcji (2.12), a te z kolei są kombinacjami liniowymi wielomianów Q_n. Można udowodnić (patrz np. [13] lub [11]), że wielomiany Q_n są z dokładnością do znaku równe wielomianom

$$\sqrt{n+\tfrac{1}{2}}P_n(x), \quad n = 0, 1, 2, \dots,$$

gdzie P_n jest *wielomianem Legendre'a*, określonym wzorem

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n.$$

(b) Rozważmy teraz przestrzeń $L_2(\mathbb{R})$ funkcji całkowalnych z kwadratem na całej prostej. Nie można w niej zbudować układu ortogonalnego ani z wielomianów, ani z wielomianów trygonometrycznych, bo ani jedne, ani drugie nie należą do tej przestrzeni. „Materiału” do zbudowania bazy w przestrzeni $L_2(\mathbb{R})$ należy szukać wśród funkcji dostatecznie szybko malejących w nieskończoności. W szczególności taką bazę można otrzymać, ortonormalizując ciąg

$$e^{-x^2/2}, xe^{-x^2/2}, x^2e^{-x^2/2}, \dots, x^ne^{-x^2/2}, \dots \tag{2.13}$$

Jest oczywiste, że każda z funkcji postaci $P(x)e^{-x^2/2}$, gdzie P jest wielomianem, należy do przestrzeni $L_2(\mathbb{R})$. Po zastosowaniu do funkcji (2.13) procesu ortonormalizacji otrzymamy układ funkcji postaci

$$\varphi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}}H_n(x)e^{-x^2/2}, \qquad n = 0, 1, 2, \dots,$$

gdzie H_n jest *wielomianem Hermite'a*, określonym wzorem

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}.$$

Można pokazać (patrz np. [13] lub [11]), że układ funkcji φ_n $(n = 0, 1, 2, \dots)$ jest bazą ortonormalną przestrzeni $L_2(\mathbb{R})$.

(c) Jeżeli rozważymy przestrzeń $L_2(\mathbb{R}_+)$ funkcji całkowalnych z kwadratem na półprostej $[0, +\infty)$, to w niej bazę ortonormalną można otrzymać, ortonormalizując ciąg funkcji

$$e^{-x/2}, xe^{-x/2}, x^2e^{-x/2}, \dots, x^ne^{-x/2}, \dots$$

Otrzymujemy wówczas układ funkcji postaci

$$\psi_n(x) = \frac{1}{n!}e^{-x/2}L_n(x), \qquad n = 0, 1, 2, \dots,$$

gdzie L_n jest *wielomianem Laguerre'a*, określonym wzorem

$$L_n(x) = e^x \frac{d^n}{dx^n}(x^n e^{-x}).$$

Podobnie jak poprzednio, można udowodnić (patrz np. [13], [11] lub [19]), że układ ten jest bazą ortonormalną w przestrzeni $L_2(\mathbb{R}_+)$.

Na zakończenie rozważymy jeszcze jeden, bardziej ogólny, przykład bazy ortonormalnej w przestrzeni funkcji całkowalnych z kwadratem.

Niech $\Omega \subset \mathbb{R}^k$ będzie zbiorem mierzalnym (w sensie Lebesgue'a) o mierze dodatniej. Przestrzeń $L_2(\Omega) = L_2(\Omega, \mathfrak{M}, m)$ (m oznacza miarę Lebesgue'a w zbiorze $\Omega \subset \mathbb{R}^k$) jest ośrodkową przestrzenią Hilberta (wielomiany o współczynnikach wymiernych tworzą zbiór przeliczalny gęsty), a więc ma przeliczalną bazę. W dalszym ciągu będziemy potrzebowali następującego twierdzenia:

Twierdzenie 2.27. *Jeżeli układy* $\left(\varphi_n^{(1)}\right)_{n=1}^{\infty}$, $\left(\varphi_n^{(2)}\right)_{n=1}^{\infty}$ *są dwiema bazami ortonormalnymi przestrzeni* $L_2(\Omega)$, *to ciąg funkcji*

$$\varphi_{n,p}(x, y) = \varphi_n^{(1)}(x)\varphi_p^{(2)}(y) \qquad n, p = 1, 2, \dots; x, y \in \Omega$$

jest bazą ortonormalną przestrzeni $L_2(\Omega \times \Omega)$.

D o w ó d. Z twierdzenia Fubiniego otrzymujemy

$$\iint\limits_{\Omega\times\Omega} \varphi_{n,p}(x, y)\overline{\varphi_{i,j}(x, y)}\, dxdy = \int\limits_{\Omega} \varphi_n^{(1)}(x)\overline{\varphi_i^{(1)}(x)}\, dx \int\limits_{\Omega} \varphi_p^{(2)}(y)\overline{\varphi_j^{(2)}(y)}\, dy.$$

Stąd wynika, że każda spośród funkcji $\varphi_{n,p}$ należy do przestrzeni $L_2(\Omega \times \Omega)$ i że stanowią one układ ortonormalny.

Udowodnimy, że jest on bazą ortonormalną przestrzeni $L_2(\Omega\times\Omega)$. Niech $f\in L_2(\Omega\times\Omega)$ będzie funkcją ortogonalną do wszystkich funkcji $\varphi_{n,p}$, tzn.

$$\iint\limits_{\Omega\times\Omega} f(x,y)\overline{\varphi_n^{(1)}(x)\varphi_p^{(2)}(y)}\,dxdy=0 \qquad (n,p=1,2,\dots).$$

Zgodnie z twierdzeniem Fubiniego (patrz [27], tw. 11.50 z części II) oznacza to, że

$$\int\limits_{\Omega}\left(\int\limits_{\Omega} f(x,y)\overline{\varphi_p^{(2)}(y)}\,dy\right)\overline{\varphi_n^{(1)}(x)}\,dx=0 \qquad (n,p=1,2,\dots). \tag{2.14}$$

Każda z funkcji $\psi_p(x)=\int\limits_{\Omega} f(x,y)\overline{\varphi_p^{(2)}(y)}\,dy$ $(p=1,2,\dots)$ należy do przestrzeni $L_2(\Omega)$, bo mamy

$$\begin{aligned}\int\limits_{\Omega}\left|\int\limits_{\Omega} f(x,y)\overline{\varphi_p^{(2)}(y)}\,dy\right|^2 dx &\leqslant \int\limits_{\Omega}\left(\int\limits_{\Omega}|f(x,y)|^2\,dy\int\limits_{\Omega}|\varphi_p^{(2)}(y)|^2\,dy\right)dx\leqslant\\ &\leqslant\int\limits_{\Omega}|\varphi_p^{(2)}(y)|^2\,dy\int\limits_{\Omega}\left(\int\limits_{\Omega}|f(x,y)|^2\,dx\right)dy<\infty.\end{aligned}$$

Jeżeli ustalimy wskaźnik p, to wobec zupełności układu $\left(\varphi_n^{(1)}\right)$ z (2.14) wynika, że $\psi_p(x)=0$ prawie wszędzie na Ω. Zatem każdy ze zbiorów $\Omega_p=\{x\in\Omega\colon\ \psi_p(x)\neq 0\}$ ma miarę zero. W konsekwencji zbiór $\Omega_0=\bigcup\limits_{p=1}^{\infty}\Omega_p$ ma miarę zero. Niech $x\in\Omega\setminus\Omega_0$. Wówczas

$$\int\limits_{\Omega} f(x,y)\overline{\varphi_p^{(2)}(y)}\,dy=0 \qquad (p=1,2,\dots)$$

i z zupełności układu $\left(\varphi_p^{(2)}\right)$ otrzymujemy, że $f(x,y)=0$ dla prawie wszystkich $y\in\Omega$. Jeżeli więc $\Omega(x)=\{y\in\Omega\colon\ f(x,y)\neq 0\}$, to miara zbioru $\Omega(x)$ jest równa zeru dla $x\in\Omega\setminus\Omega_0$. Niech $Z=\{(x,y)\in\Omega\times\Omega\colon\ f(x,y)\neq 0\}$. Mamy

$$m(Z)=\int\limits_{\Omega} m\left(\Omega(x)\right)\,dx=\int\limits_{\Omega\setminus\Omega_0} m\left(\Omega(x)\right)\,dx=0,$$

czyli $f(x,y)=0$ prawie wszędzie na $\Omega\times\Omega$. ■

Ćwiczenia

1. Udowodnić, że nierówność Schwarza przechodzi w równość wtedy i tylko wtedy, gdy wektory x i y są liniowo zależne.

2. Udowodnić, że jeżeli $\operatorname{Re}\langle x,y\rangle = \|x\|\,\|y\|$, to wektory x i y mają jednakowy kierunek, tzn. $x=\alpha y$ dla $\alpha \geqslant 0$ lub $y=\beta x$ dla $\beta \geqslant 0$.

3. Podać interpretację geometryczną tożsamości równoległoboku (tw. 2.2 (d)).

4. Wykazać, że zachodzą następujące równości zwane *tożsamościami polaryzacyjnymi*:

(a) w rzeczywistej przestrzeni unitarnej

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right);$$

(b) w zespolonej przestrzeni unitarnej

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right).$$

5. Wykazać, że jeżeli w rzeczywistej przestrzeni unormowanej norma $\|\cdot\|$ spełnia tożsamość równoległoboku, to można w niej wprowadzić taki iloczyn skalarny $\langle\cdot,\cdot\rangle$, że $\|x\| = \sqrt{\langle x,x\rangle}$.

6. Wykazać, że dla wektorów x, y, z przestrzeni unitarnej X równość

$$\|x-z\| = \|x-y\| + \|y-z\|$$

zachodzi wtedy i tylko wtedy, gdy $y=\alpha x + (1-\alpha)z$ dla pewnego $\alpha \in [0,1]$.

7. Wykazać, że jeżeli wektory $x \neq 0$ i $y \neq 0$ przestrzeni unitarnej spełniają równość $\|x+y\| = \|x\| + \|y\|$, to $x=\alpha y$ dla pewnego $\alpha > 0$.

8. Wykazać, że w przestrzeni unitarnej z warunków

$$\|x_n\| = \|y_n\| = 1, \qquad \lim_{n\to\infty} \|x_n + y_n\| = 2$$

wynika, że $\lim\limits_{n\to\infty} \|x_n - y_n\| = 0$.

9. Wykazać, że w przestrzeni unitarnej, jeżeli $\|x_n\| \to \|x_0\|$ i $\langle x_n, x_0\rangle \to \|x_0\|^2$, to $x_n \to x_0$.

10. Wykazać, że przestrzeń $\mathcal{C}[-1,1]$ z iloczynem skalarnym (2.3) nie jest przestrzenią Hilberta.

11. Niech I będzie niepustym zbiorem i niech $\ell_2(I)$ oznacza zbiór wszystkich funkcji $x\colon I \to \mathbb{C}$ takich, że $\sum\limits_{i\in I} |x(i)|^2 < \infty$. Dla $x,y \in \ell_2(I)$ definiujemy $\langle x,y\rangle = \sum\limits_{i\in I} x(i)\overline{y(i)}$. Udowodnić, że $\ell_2(I)$ jest przestrzenią Hilberta.

12. Wykazać, że jeżeli $\operatorname{card} I = \operatorname{card} J$, to przestrzenie $\ell_2(I)$ i $\ell_2(J)$ są izomorficzne.

13. Udowodnić, że jeżeli M jest domkniętą podprzestrzenią przestrzeni Hilberta H, to $M = (M^\perp)^\perp$. Czy równość ta pozostaje prawdziwa dla podprzestrzeni, które nie są domknięte?

14. Niech X będzie przestrzenią unitarną (niezupełną) utworzoną przez wszystkie ciągi postaci $x = (\xi_n)$, gdzie $\xi_k = 0$ dla wszystkich $k \geqslant n(x)$, z działaniami i iloczynem skalarnym zdefiniowanym tak jak w przestrzeni ℓ_2. Niech

$$M = \left\{ x = (\xi_n) \in X \colon \ \sum_{n=1}^{\infty} \frac{1}{n}\, \xi_n = 0 \right\}.$$

Wykazać, że M jest domkniętą podprzestrzenią liniową przestrzeni H, $M \neq H$ oraz że nie istnieje w H element różny od zera i ortogonalny do M.

15. Udowodnić własności rzutowania na podprzestrzeń (wn. 2.9 (a)–(i)).

16. W przestrzeni ℓ_2 przy ustalonym n_0 zbiór

$$M = \{x = (\xi_n) \in \ell_2 \colon \ \xi_{n_0+k} = 0 \quad \text{dla} \ \ k \geqslant 0\}$$

jest domkniętą podprzestrzenią. Wykazać, że rzutowanie P na podprzestrzeń M jest funkcją $Px = (\xi_1, \dots, \xi_{n_0-1}, 0, 0, \dots)$.

17. Wykazać, że jeżeli dla rzutowania P i wektora $x \neq 0$ zachodzi równość $Px = \alpha x$, to albo $\alpha = 0$, albo $\alpha = 1$.

18. Udowodnić, że jeżeli $M = \{x \in H \colon \ f(x) = 0\}$, gdzie f jest funkcjonałem liniowym ograniczonym na przestrzeni Hilberta H i $M \neq H$, to $M^\perp$ jest przestrzenią jednowymiarową.

19. Wykazać, że ciągi funkcji

$$\sqrt{\frac{2}{\pi}} \sin nx \quad (n = 1, 2, \dots),$$
$$\frac{1}{\sqrt{\pi}}, \ \sqrt{\frac{2}{\pi}} \cos nx \quad (n = 1, 2, \dots)$$

są układami ortonormalnymi zupełnymi w przestrzeni $L_2[0, \pi]$.

20. Niech (e_n) będzie układem ortonormalnym w przestrzeni Hilberta H, a (λ_n) — ciągiem liczbowym ograniczonym. Sprawdzić, że

(a) wzór $Ax = \sum\limits_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n$ określa operator A odwzorowujący przestrzeń H w siebie, tzn. szereg ten jest zbieżny dla każdego $x \in H$;

(b) A jest operatorem liniowym ograniczonym;

(c) $\|A\| = \sup_n |\lambda_n|$.

Rozdział 3

Trzy zasady analizy funkcjonalnej

W tym rozdziale przedstawimy trzy twierdzenia, które mają fundamentalne znaczenie w teorii operatorów liniowych na przestrzeniach Banacha. Są to twierdzenia: Banacha-Steinhausa, Banacha o odwzorowaniu otwartym i Hahna-Banacha.

3.1. Twierdzenie Baire'a

Definicja 3.1. Podzbiór Z przestrzeni metrycznej X (z metryką d) jest *gęsty w kuli* $B_r(x) = \{y \in X\colon\ d(x,y) < r\}$, jeżeli $B_r(x) \subset \overline{Z}$. (Przypomnijmy, że symbol $\overline{Z}$ oznacza domknięcie zbioru Z). Zbiór Z, który nie jest gęsty w żadnej kuli, nazywa się *nigdziegęstym*. Jeżeli zbiór Z da się przedstawić w postaci

$$Z = \bigcup_{n=1}^{\infty} Z_n,$$

gdzie każdy ze zbiorów Z_n jest nigdziegęsty, to mówimy, że Z jest *zbiorem pierwszej kategorii*. Zbiór, który nie jest pierwszej kategorii, nazywa się *zbiorem drugiej kategorii*.

Twierdzenie 3.1 (Baire'a). *Przestrzeń metryczna zupełna niepusta jest zbiorem drugiej kategorii.*

D o w ó d. Przypuśćmy, że niepusta przestrzeń zupełna X jest zbiorem pierwszej kategorii. Zatem

$$X = \bigcup_{n=1}^{\infty} Z_n, \tag{3.1}$$

gdzie każdy ze zbiorów Z_n jest nigdziegęsty.

Ponieważ zbiór Z_1 jest nigdziegęsty, więc istnieje kula domknięta $\overline{B}_{r_1}(x_1)$ niezawierająca punktów zbioru Z_1. Z kolei z tego, że zbiór Z_2 jest nigdziegęsty, wynika, że w kuli $\overline{B}_{r_1}(x_1)$ zawiera się kula domknięta $\overline{B}_{r_2}(x_2)$, niezawierająca punktów zbioru Z_2, przy czym możemy założyć, że $r_2 \leqslant \frac{1}{2}r_1$.

Kontynuując to postępowanie, otrzymamy ciąg kul domkniętych $\big(\overline{B}_{r_n}(x_n)\big)$ takich, że

$$r_{n+1} \leqslant \frac{1}{2^n} r_1, \tag{3.2}$$

$$\overline{B}_{r_{n+1}}(x_{n+1}) \subset \overline{B}_{r_n}(x_n), \tag{3.3}$$

$$\overline{B}_{r_n}(x_n) \cap Z_n = \emptyset. \tag{3.4}$$

Z warunków (3.2) i (3.3) na mocy twierdzenia Cantora ([27], tw. 12.11 z części III) wynika, że przekrój wszystkich kul $\overline{B}_{r_n}(x_n)$ jest niepusty, a więc zawiera punkt x. Na mocy (3.4) punkt x nie należy do żadnego ze zbiorów Z_n i w konsekwencji otrzymujemy sprzeczność z równością (3.1). ■

3.2. Zasada jednostajnej ograniczoności

Twierdzenie 3.2 (Banacha-Steinhausa). *Niech (A_n) będzie ciągiem operatorów liniowych ograniczonych przekształcających przestrzeń Banacha X w przestrzeń unormowaną Y. Jeżeli dla dowolnego punktu $x \in X$ ciąg $(A_n x)$ jest ograniczony, to ciąg norm $(\|A_n\|)$ jest również ograniczony.*

Powyższe twierdzenie bywa również nazywane *zasadą jednostajnej ograniczoności*.

Dowód. Niech $Z_{nk} = \{x\colon \|A_n x\| \leqslant k\}$ dla $n = 1, 2, \dots,$ $k = 1, 2 \dots$ Zauważmy, że każdy ze zbiorów Z_{nk} jest domknięty. Wynika to z ciągłości operatorów A_n i ciągłości normy. Oznaczmy $Z_k = \bigcap\limits_{n=1}^{\infty} Z_{nk}$. Zbiór Z_k jest domknięty, ponieważ przekrój zbiorów domkniętych jest zbiorem domkniętym ([27] tw. 2.6 z części I). Z założenia $X = \bigcup\limits_{k=1}^{\infty} Z_k$. Ponieważ przestrzeń X jest zupełna, więc na mocy twierdzenia Baire'a istnieje wskaźnik k_0 taki, że zbiór Z_{k_0} jest gęsty w pewnej kuli $B_r(x_0)$. Ponieważ zbiór Z_{k_0} jest domknięty, więc $\overline{B}_r(x_0) \subset Z_{k_0}$. Zatem jeżeli $\|x - x_0\| \leqslant r$, to $\|A_n x\| \leqslant k_0$ dla $n = 1, 2, \dots$ W szczególności

$$\|A_n x_0\| \leqslant k_0 \quad \text{dla} \;\; n = 1, 2, \dots \tag{3.5}$$

Niech $x \neq 0$ będzie dowolnym punktem przestrzeni X. Z równości

$$\left\| \left(\frac{r}{\|x\|} x + x_0 \right) - x_0 \right\| = r$$

otrzymujemy

$$\left\| A_n \left(\frac{r}{\|x\|} x + x_0 \right) \right\| \leqslant k_0 \quad \text{dla} \ \ n = 1, 2, \ldots \tag{3.6}$$

Ponieważ

$$\begin{aligned} \|A_n x\| &= \left\| A_n \left(\frac{r}{\|x\|} x + x_0 \right) - A_n x_0 \right\| \frac{\|x\|}{r} \leqslant \\ &\leqslant \left(\left\| A_n \left(\frac{r}{\|x\|} x + x_0 \right) \right\| + \|A_n x_0\| \right) \frac{\|x\|}{r}, \end{aligned}$$

więc z nierówności (3.5) i (3.6) wynika, że dla każdego wskaźnika n i dowolnego $x \in X$ (również dla $x = 0$) zachodzi nierówność

$$\|A_n x\| \leqslant \frac{2k_0}{r} \|x\|.$$

Stąd otrzymujemy

$$\|A_n\| \leqslant \frac{2k_0}{r} \quad \text{dla} \ \ n = 1, 2, \ldots \qquad \blacksquare$$

Twierdzenie 3.3. *Jeżeli (A_n) jest ciągiem operatorów liniowych ograniczonych przekształcających przestrzeń Banacha X w przestrzeń unormowaną Y takim, że ciąg $(A_n x)$ jest zbieżny dla każdego $x \in X$, to operator A zdefiniowany wzorem*

$$Ax = \lim_{n \to \infty} A_n x$$

jest operatorem liniowym ograniczonym.

D o w ó d. Liniowość operatora A wynika z liniowości operatorów A_n i ciągłości działań algebraicznych w przestrzeni Y. Pozostaje do udowodnienia, że jest on ograniczony. Mamy

$$\|A_n x\| \leqslant \|A_n\| \|x\| \quad \text{dla} \ \ x \in X, \ \ n = 1, 2, \ldots$$

Ciąg $(A_n x)$ jest zbieżny, a więc ograniczony dla dowolnego $x \in X$. Na mocy twierdzenia Banacha-Steinhausa istnieje stała $M > 0$ taka, że $\|A_n\| \leqslant M$ dla $n = 1, 2, \ldots$ Wobec tego

$$\|A_n x\| \leqslant M \|x\| \quad \text{dla} \ \ x \in X, \ n = 1, 2, \ldots$$

Ustalając $x \in X$ i przechodząc w tej nierówności do granicy, $n \to \infty$, otrzymujemy

$$\|Ax\| \leqslant M \|x\|,$$

a to oznacza, że operator A jest ograniczony. $\blacksquare$

Definicja 3.2. Podzbiór Z przestrzeni unormowanej X jest *liniowo gęsty* w X, jeżeli powłoka liniowa zbioru Z, $\operatorname{span} Z$, jest gęsta w X. Oznacza to, że dla dowolnego wektora $x \in X$ i dowolnego $\varepsilon > 0$ istnieją wektory $x_1, x_2, \ldots,$ $x_n \in Z$ i skalary $\alpha_1, \alpha_2, \ldots, \alpha_n$ takie, że $\|x-(\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_n x_n)\| < \varepsilon$.

Twierdzenie 3.4. *Niech (A_n) będzie ciągiem operatorów liniowych ograniczonych przekształcających przestrzeń unormowaną X w przestrzeń Banacha Y. Niech ciąg norm $(\|A_n\|)$ będzie ograniczony. Jeżeli ciąg $(A_n x)$ jest zbieżny dla każdego x należącego do zbioru Z liniowo gęstego w X, to jest on zbieżny dla każdego $x \in X$.*

D o w ó d. Ponieważ operatory A_n są liniowe i działania algebraiczne w przestrzeni Y są ciągłe, więc dla dowolnego $x \in \operatorname{span} Z$ ciąg $(A_n x)$ jest zbieżny. Z założenia wynika, że istnieje liczba $M > 0$ taka, że $\|A_n\| \leqslant M$ dla $n = 1, 2, \ldots$ Niech będzie dane $\varepsilon > 0$ i wektor $x \in X$. Wybierzmy taki wektor $z \in \operatorname{span} Z$, aby

$$\|x - z\| < \frac{\varepsilon}{4M}. \tag{3.7}$$

Ciąg $(A_n z)$ jest zbieżny, więc spełnia warunek Cauchy'ego. Zatem istnieje wskaźnik n_0 taki, że

$$\|A_n z - A_m z\| < \frac{1}{2}\varepsilon \quad \text{dla} \;\; n,\, m \geqslant n_0. \tag{3.8}$$

Ponieważ

$$\begin{aligned} \|A_n x - A_m x\| &= \|A_n(x - z) + (A_n z - A_m z) + A_m(z - x)\| \leqslant \\ &\leqslant \|A_n\|\|x - z\| + \|A_n z - A_m z\| + \|A_m\|\|z - x\| \leqslant \\ &\leqslant 2M\|x - z\| + \|A_n z - A_m z\|, \end{aligned}$$

więc z (3.7) i (3.8) wynika, że

$$\|A_n x - A_m x\| < \varepsilon \quad \text{dla} \;\; n,\, m \geqslant n_0.$$

Zatem ciąg $(A_n x)$ spełnia warunek Cauchy'ego i wobec zupełności przestrzeni Y jest zbieżny. ■

Pokażemy teraz przykład ilustrujący zastosowanie twierdzenia Banacha-Steinhausa do rozstrzygania problemów klasycznej analizy.

Przykład 3.1. Udowodnimy, że szereg Fouriera funkcji ciągłej i 2π-okresowej nie musi być do niej (punktowo) zbieżny. Przestrzeń $\mathcal{C}_{2\pi}$ funkcji ciągłych i 2π-okresowych określonych na prostej jest przestrzenią Banacha z normą $\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|$ (ćwiczenie 7 w rozdz. 1).

Wiemy, że n-ta suma częściowa szeregu Fouriera funkcji f ma przedstawienie całkowe postaci:

$$s_n(f;x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\mathcal{D}_n(t)\,dt \qquad (n = 0, 1, 2, \dots),$$

gdzie $\mathcal{D}_n(t)$ jest *jądrem Dirichleta*, tzn.

$$\mathcal{D}_n(t) = \frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}$$

(patrz [27], wzór (7.21) s. 38 w części II). Zatem badanie zbieżności szeregu Fouriera funkcji f w punkcie x oznacza sprawdzenie, czy zachodzi równość

$$\lim_{n\to\infty} s_n(f;x) = f(x).$$

Przy ustalonym x funkcja $f \mapsto s_n(f;x)$ jest funkcjonałem liniowym ograniczonym na przestrzeni $\mathcal{C}_{2\pi}$. Liniowość jest oczywista, a ograniczoność wynika z nierówności

$$|s_n(f;x)| \leqslant \|f\|_\infty \frac{1}{\pi}\int_{-\pi}^{\pi} |\mathcal{D}_n(t)|\, dt.$$

Aby uprościć oznaczenia, przyjmijmy $x = 0$ i $\varphi_n(f) = s_n(f;0)$. Wówczas φ_n jest ciągłym i ograniczonym funkcjonałem liniowym na przestrzeni $\mathcal{C}_{2\pi}$, którego norma jest na mocy przykładu 1.6 (b) (wzór (1.25)) równa

$$\|\varphi_n\| = \frac{1}{\pi}\int_{-\pi}^{\pi} |\mathcal{D}_n(t)|\,dt.$$

Pokażemy, że

$$\sup_n \int_{-\pi}^{\pi} |\mathcal{D}_n(t)|\,dt = +\infty. \tag{3.9}$$

Jeżeli zastosujemy nierówność $\sin\frac{1}{2}t \leqslant \frac{1}{2}t$ dla $0 \leqslant t \leqslant \pi$, to otrzymamy

$$\int_{-\pi}^{\pi} |\mathcal{D}_n(t)|\,dt = \frac{1}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac{1}{2})t|}{\sin\frac{1}{2}t}\,dt \geqslant \frac{2}{\pi}\int_0^{\pi} \frac{|\sin(n+\frac{1}{2})t|}{t}\,dt.$$

Po wykonaniu zamiany zmiennej $(n+\frac{1}{2})t = u$ otrzymamy

$$\int\limits_{-\pi}^{\pi} |\mathcal{D}_n(t)|\, dt \geqslant \frac{2}{\pi} \int\limits_{0}^{(n+1/2)\pi} \frac{|\sin u|}{u}\, du.$$

Ostatnia całka dąży do nieskończoności, gdy $n \to \infty$, bo całka niewłaściwa $\int\limits_0^{+\infty} \frac{\sin x}{x}\, dx$ nie jest bezwzględnie zbieżna (por. [27], przykład 6.11 z części I). To dowodzi prawdziwości relacji (3.9). Na mocy twierdzenia Banacha-Steinhausa istnieje funkcja ciągła 2π-okresowa f_0, dla której

$$\sup_n |\varphi_n(f_0)| = \sup_n |s_n(f_0; 0)| = +\infty.$$

Zatem szereg Fouriera funkcji f_0 nie jest zbieżny w punkcie $x = 0$. Oczywiście tę konstrukcję można wykonać w dowolnym punkcie. Co więcej, można pokazać, że zbiór funkcji ciągłych 2π-okresowych, których szereg Fouriera jest rozbieżny w danym punkcie x, jest nieprzeliczalny (patrz [23]).

Udowodniliśmy więc klasyczne twierdzenie o szeregach Fouriera.

Twierdzenie 3.5 (du Bois-Reymonda). *Dla dowolnego punktu x z przedziału $[-\pi, \pi]$ istnieje funkcja ciągła i 2π-okresowa, której szereg Fouriera ma w tym punkcie nieograniczone sumy częściowe.*

Inne zastosowania twierdzenia Banacha-Steinhausa do dowodów twierdzeń z klasycznej analizy może Czytelnik znaleźć w [2], [20], [5].

3.3. Twierdzenie Banacha o odwzorowaniu otwartym

Twierdzenie 3.6 (Banacha o odwzorowaniu otwartym). *Jeżeli A jest operatorem liniowym ograniczonym przekształcającym przestrzeń Banacha X na przestrzeń Banacha Y, to obraz $A(G)$ dowolnego zbioru otwartego G w przestrzeni X jest zbiorem otwartym w Y, a więc A jest odwzorowaniem otwartym.*

D o w ó d. Niech U i V będą kulami jednostkowymi odpowiednio w przestrzeniach X i Y. Wówczas dla dowolnego $r > 0$, $rU = \{x \in X \colon \|x\| < r\}$ i odpowiednio $rV = \{y \in Y \colon \|y\| < r\}$.

Z suriektywności operatora A mamy

$$Y = A(X) = A\Bigl(\bigcup_{k=1}^{\infty} kU\Bigr) = \bigcup_{k=1}^{\infty} A(kU).$$

Stąd

$$Y = \bigcup_{k=1}^{\infty} \overline{A(kU)}.$$

Ponieważ Y jest przestrzenią zupełną, więc z twierdzenia Baire'a wynika, że istnieją wskaźnik k_0 i $r > 0$ takie, że $B_r(y_0) \subset \overline{A(k_0U)}$, tzn. jeżeli $\|y - y_0\| < r$, to istnieje ciąg (x_n) taki, że $\|x_n\| < k_0$ dla $n = 1, 2, \dots$ i $y = \lim_{n\to\infty} Ax_n$.

Weźmy dowolne $y \in Y$, $\|y\| < r$. Istnieją takie ciągi (x'_n), $(x''_n) \in k_0U$, że $y_0 = \lim_{n\to\infty} Ax'_n$ i $y + y_0 = \lim_{n\to\infty} Ax''$. Przyjmując $x_n = x''_n - x'_n$, otrzymujemy $\|x_n\| \leqslant \|x''_n\| + \|x'_n\| < 2k_0$ oraz $\lim_{n\to\infty} Ax_n = y$. Zatem wykazaliśmy, że jeżeli $\|y\| < r$, to $y \in \overline{A(2k_0U)}$, czyli prawdziwa jest inkluzja

$$rV \subset \overline{A(2k_0U)}.$$

Stąd natychmiast wynika, że dla dowolnego $\rho > 0$ zachodzi

$$\rho V \subset \overline{A\Big(\frac{2k_0\rho}{r}U\Big)}. \tag{3.10}$$

W szczególności

$$\frac{r}{2k_0}V \subset \overline{A(U)}.$$

W następnym kroku wykażemy, że obraz kuli jednostkowej $A(U)$ zawiera kulę o środku w zerze, dokładniej udowodnimy, że zachodzi inkluzja

$$\frac{r}{4k_0}V \subset A(U). \tag{3.11}$$

Weźmy dowolne $y \in \dfrac{r}{4k_0}V$. Ponieważ na mocy (3.10) prawdziwa jest inkluzja

$$\frac{r}{4k_0}V \subset \overline{A\Big(\frac{1}{2}U\Big)},$$

więc istnieje punkt $y_1 \in A\left(\frac{1}{2}U\right)$ leżący dowolnie blisko punktu y. Niech y_1 będzie takim punktem, że $\|y - y_1\| < \dfrac{r}{8k_0}$.

Z kolei, ponieważ

$$\frac{r}{8k_0}V \subset \overline{A\Big(\frac{1}{4}U\Big)},$$

więc istnieje $y_2 \in A\left(\frac{1}{4}U\right)$ takie, że $\|y - y_1 - y_2\| < \dfrac{r}{16k_0}$. Kontynuując to postępowanie, otrzymamy ciąg punktów (y_n), $y_n \in A\left(2^{-n}U\right)$, dla którego

$$\|y - y_1 - y_2 - \ldots - y_n\| < \frac{r}{2^{n+2}k_0}.$$

Zatem $y = \sum\limits_{n=1}^{\infty} y_n$. Ponadto $y_n = Ax_n$, gdzie $x_n \in 2^{-n}U$, czyli $\|x_n\| < \frac{1}{2^n}$ dla $n = 1, 2, \ldots$ Ponieważ X jest przestrzenią Banacha, więc na mocy twierdzenia 1.6 szereg $\sum\limits_{n=1}^{\infty} x_n$ jest zbieżny. Ponadto jeżeli $x = \sum\limits_{n=1}^{\infty} x_n$, to wówczas

$$\|x\| \leqslant \sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1$$

oraz na mocy ciągłości operatora A

$$Ax = A\Big(\sum_{n=1}^{\infty} x_n\Big) = \sum_{n=1}^{\infty} Ax_n = \sum_{n=1}^{\infty} y_n = y.$$

Zatem wykazaliśmy, że zachodzi (3.11).

Natychmiastową konsekwencją (3.11) jest następująca inkluzja

$$\frac{r\delta}{4k_0} V \subset A\,(\delta U)\,, \tag{3.12}$$

prawdziwa dla dowolnego $\delta > 0$.

Niech G będzie dowolnym zbiorem otwartym w X i niech $y_0 \in A(G)$. Wówczas $y_0 = Ax_0$ dla pewnego $x_0 \in G$. Ponieważ x_0 jest punktem wewnętrznym zbioru G, więc istnieje takie $\delta > 0$, że $B_\delta(x_0) \subset G$. Niech $\varepsilon = \frac{r\delta}{4k_0}$.

Zauważmy, że $B_\varepsilon(y_0) = \{y \in Y\colon\ \|y - y_0\| < \varepsilon\} = y_0 + \varepsilon V$ (symbol $y_0 + \varepsilon V$ oznacza zbiór $\{y \in Y\colon\ y = y_0 + y_1,\ y_1 \in \varepsilon V\}$). Mamy również $B_\delta(x_0) = \{x \in X\colon\ \|x - x_0\| < \delta\} = x_0 + \delta U$. Z (3.12) otrzymujemy

$$\begin{aligned} B_\varepsilon(y_0) &= y_0 + \varepsilon V \subset y_0 + A(\delta U) = Ax_0 + A(\delta U) = \\ &= A(x_0 + \delta U) = A(B_\delta(x_0)) \subset A(G). \end{aligned}$$

Otrzymana inkluzja oznacza, że y_0 jest punktem wewnętrznym zbioru $A(G)$. ■

U w a g a. Inkluzja (3.11) oznacza, że dla liniowej ograniczonej suriekcji A istnieje $\gamma > 0$ takie, że

$$\gamma V \subset A(U). \tag{3.13}$$

3.4. Operatory odwracalne

Niech X i Y będą przestrzeniami unormowanymi i niech $A\colon X \to Y$ będzie przekształceniem liniowym. Przypomnijmy, że jądro i obraz przekształcenia A są zbiorami zdefiniowanymi w następujący sposób:

$$\begin{aligned}\ker A &= A^{-1}\{0\} = \{x \in X\colon\ Ax = 0\},\\ \operatorname{im} A &= A(X) = \{y \in Y\colon\ y = Ax \ \text{ dla pewnego } x \in X\}.\end{aligned}$$

Są one podprzestrzeniami odpowiednio przestrzeni X i Y (por. [26], lem. 5.1). Ponadto wiemy, że A jest wzajemnie jednoznaczne wtedy i tylko wtedy, gdy $\ker A = \{0\}$.

Zauważmy, że jeżeli przekształcenie A jest ograniczone, to $\ker A$ jest podprzestrzenią domkniętą jako przeciwobraz poprzez funkcję ciągłą zbioru domkniętego $\{0\}$, natomiast obraz $\operatorname{im} A$ nie musi być domknięty.

Wiadomo, że przekształcenie odwrotne A^{-1} istnieje wtedy i tylko wtedy, gdy A jest bijekcją i jest ono zdefiniowane wzorem:

$$A^{-1}y = x \quad \text{wtedy i tylko wtedy, gdy} \quad Ax = y \tag{3.14}$$

lub, co na jedno wychodzi,

$$AA^{-1} = I_Y, \qquad A^{-1}A = I_X.$$

Można łatwo wykazać, że przekształcenie odwrotne $A^{-1}\colon Y \to X$ do przekształcenia liniowego (jeżeli ono istnieje) jest również liniowe ([26], tw. 5.4).

Jeżeli przekształcenie $A\colon X \to Y$ jest wzajemnie jednoznaczne, ale nie jest na, to możemy je rozpatrywać jako funkcję $A\colon X \to \operatorname{im} A$. Wówczas jest ona bijekcją i istnieje do niej przekształcenie odwrotne $A^{-1}\colon \operatorname{im} A \to X$. W dalszym ciągu symbol A^{-1} będzie oznaczał tak określoną funkcję. Będziemy wówczas mówili, że operator A jest *odwracalny*.

Jeżeli przekształcenie A jest ograniczone i odwracalne, to operator odwrotny A^{-1} nie musi być ograniczony. Wynika to z następującego przykładu:

Przykład 3.2. Niech $X = Y = \mathcal{C}[0,1]$ i niech operator $A\colon X \to Y$ będzie określony wzorem:

$$Ax(t) = \int_0^t x(s)\,ds, \qquad (t \in [0,1]). \tag{3.15}$$

Operator A jest oczywiście liniowy i ograniczony, bo $\|Ax\|_\infty \leqslant \|x\|_\infty$ (co więcej, $\|A\| = 1$). Ponadto jest to odwzorowanie wzajemnie jednoznaczne.

Jeżeli bowiem $Ax(t) = 0$, to po zróżniczkowaniu obu stron (3.15), na mocy twierdzenia 6.16 z [27], część I, mamy $x(t) = 0$ dla dowolnego $t \in [0,1]$, a więc $x = 0$. Operator odwrotny do operatora A dany jest wzorem

$$A^{-1}y(t) = y'(t) \qquad (t \in [0,1]).$$

Jest on określony na podprzestrzeni X_0, składającej się z wszystkich funkcji mających ciągłą pochodną na przedziale [0,1] i przyjmujących wartość zero w zerze. Operator ten nie jest ograniczony (por. przykład 1.5 (c)).

Pokażemy teraz, że taka sytuacja nie może mieć miejsca, gdy operator A przekształca przestrzeń Banacha na przestrzeń Banacha.

Twierdzenie 3.7 (Banacha o operatorze odwrotnym). *Jeżeli A jest wzajemnie jednoznacznym przekształceniem liniowym ograniczonym przestrzeni Banacha X na przestrzeń Banacha Y, to istnieje liczba $\gamma > 0$ taka, że*

$$\|Ax\| \geqslant \gamma\|x\| \qquad (x \in X). \tag{3.16}$$

Warunek (3.16) *oznacza, że przekształcenie odwrotne $A^{-1}\colon Y \to X$ jest ograniczone.*

D o w ó d. Z (3.13) (uwaga po twierdzeniu Banacha o odwzorowaniu otwartym) wiemy, że istnieje takie $\gamma > 0$, że jeżeli $\|y\| < \gamma$, to $y = Ax$ dla pewnego x takiego, że $\|x\| < 1$. Ponieważ operator A jest wzajemnie jednoznaczny, więc to oznacza, że warunek $\|Ax\| < \gamma$ implikuje $\|x\| < 1$. Zatem, jeżeli $\|x\| \geqslant 1$, to $\|Ax\| \geqslant \gamma$ i w konsekwencji otrzymujemy nierówność (3.16). Stąd i z (3.14) wynika, że $\|A^{-1}\| \leqslant \dfrac{1}{\gamma}$. ■

Wniosek 3.8. *Niech na przestrzeni liniowej X dane będą dwie normy $\|\cdot\|_1$ i $\|\cdot\|_2$. Jeżeli X jest przestrzenią Banacha przy każdej z tych norm i dla dowolnego ciągu (x_n) elementów przestrzeni X spełniony jest warunek:*

$$\|x_n\|_1 \to 0 \quad \textit{implikuje} \quad \|x_n\|_2 \to 0,$$

to i na odwrót

$$\|x_n\|_2 \to 0 \quad \textit{implikuje} \quad \|x_n\|_1 \to 0.$$

Zatem normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są równoważne.

D o w ó d. Identyczność $I_X\colon (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ jest operatorem liniowym ograniczonym, więc odwzorowanie odwrotne $I_X\colon (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ jest również ograniczone. ■

W przypadku, gdy operator A nie jest suriekcją, ograniczoność operatora odwrotnego związana jest z domkniętością obrazu $\operatorname{im} A$.

Twierdzenie 3.9. *Jeżeli A jest wzajemnie jednoznacznym przekształceniem liniowym ograniczonym przestrzeni Banacha X w przestrzeń Banacha Y, to następujące warunki są równoważne*:

(a) *operator odwrotny A^{-1} jest ograniczony*;

(b) *istnieje liczba $\gamma > 0$ taka, że $\|Ax\| \geqslant \gamma\|x\|$ dla dowolnego $x \in X$*;

(c) *obraz operatora A jest domknięty.*

D o w ó d. Najpierw pokażemy, że warunek (b) jest konsekwencją (a). Mamy

$$\|x\| = \|A^{-1}Ax\| \leqslant \|A^{-1}\|\|Ax\|.$$

Zatem $\gamma = \dfrac{1}{\|A^{-1}\|}$.

Z kolei wykazujemy, że z (b) wynika (c). Niech $y_n = Ax_n$ dla $n = 1, 2, \ldots$ oraz $y_n \to y_0$. Musimy znaleźć x_0 takie, że $y_0 = Ax_0$. Mamy

$$\|y_n - y_m\| = \|A(x_n - x_m)\| \geqslant \gamma\|x_n - x_m\| \quad (n,\, m = 1, 2, \ldots).$$

Ponieważ ciąg (y_n) jest zbieżny, więc stąd wynika, że ciąg (x_n) spełnia warunek Cauchy'ego. Wobec zupełności przestrzeni X istnieje punkt x_0 taki, że $x_n \to x_0$. Ponieważ operator A jest ciągły, więc także $y_n = Ax_n \to$ $\to Ax_0$. Stąd $y_0 = Ax_0$.

W końcu udowodnimy, że (c) implikuje (a). Obraz operatora A jest domkniętą podprzestrzenią przestrzeni Banacha Y. Jest więc on przestrzenią Banacha. Z twierdzenia Banacha wynika, że operator odwrotny A^{-1} jest ograniczony. ■

U w a g a. Z dowodu wynika, że warunki (a) i (b) są równoważne bez założenia, że przestrzenie X i Y są zupełne.

Na zakończenie tego podrozdziału poznamy jeszcze jedno twierdzenie będące konsekwencją twierdzenia Banacha o odwzorowaniu odwrotnym. Niech $f\colon X \to Y$ będzie funkcją odwzorowującą zbiór X w zbiór Y. Wówczas zbiór

$$\operatorname{graph} f = \{(x, f(x))\colon\ x \in X\} \subset X \times Y$$

nazywamy *wykresem odwzorowania f*. W dalszym ciągu ograniczymy się do rozpatrywania wykresów odwzorowań liniowych $A\colon X \to Y$ pomiędzy przestrzeniami Banacha. Łatwo sprawdzić, że wówczas $\operatorname{graph} A$ jest podprzestrzenią liniową przestrzeni $X \times Y$ wyznaczoną jednoznacznie przez operator A. Jeżeli ponadto operator A jest ograniczony, to podprzestrzeń $\operatorname{graph} A$ jest domknięta, a więc jest przestrzenią Banacha względem normy $\|(x, Ax)\| = \|x\|_X + \|Ax\|_Y$. Istotnie, niech (x_0, y_0) będzie punktem skupienia zbioru $\operatorname{graph} A$ i niech (x_n, Ax_n) będzie ciągiem zbieżnym do (x_0, y_0). To oznacza, że $x_n \to x_0$ i $Ax_n \to y_0$, gdy $n \to \infty$. Z ciągłości operatora A wynika, że $Ax_n \to Ax_0$. Zatem $y_0 = Ax_0$ i w konsekwencji $(x_0, y_0) = (x_0, Ax_0) \in \operatorname{graph} A$.

Zauważmy, że w tym dowodzie wykorzystywaliśmy jedynie ciągłość operatora A, a więc udowodniliśmy, że wykres funkcji ciągłej (pomiędzy przestrzeniami metrycznymi X i Y) jest zbiorem domkniętym w produkcie $X \times Y$. W przypadku dowolnych przekształceń twierdzenie odwrotne nie jest prawdziwe, ale okazuje się, że jest ono prawdziwe w przypadku, gdy A jest operatorem liniowym pomiędzy przestrzeniami Banacha.

Twierdzenie 3.10 (Banacha o domkniętym wykresie). *Jeżeli operator liniowy A przekształcający przestrzeń Banacha X w przestrzeń Banacha Y ma domknięty wykres, to jest on operatorem ograniczonym.*

D o w ó d. Na podprzestrzeni $\operatorname{graph} A \subset X \times Y$, rozpatrywanej jako przestrzeń Banacha, zdefiniujemy operator rzutowania $P_1\colon \operatorname{graph} A \to X$ określony wzorem $P_1(x, Ax) = x$. Jest on liniowy, ograniczony i odwzorowuje w sposób wzajemnie jednoznaczny przestrzeń $\operatorname{graph} A$ na przestrzeń X. Z twierdzenia Banacha o operatorze odwrotnym istnieje przekształcenie odwrotne $P_1^{-1}\colon X \to \operatorname{graph} A$, które jest liniowe i ograniczone. Ponadto operator $P_2(x, Ax) = Ax$ działający z przestrzeni $\operatorname{graph} A$ w przestrzeń Y jest także liniowy i ograniczony. Ponieważ $A = P_2 \circ P_1^{-1}$, więc A jest operatorem ograniczonym. ∎

3.5. Twierdzenie Hahna-Banacha

Niech X będzie przestrzenią unormowaną i M jej podprzestrzenią (niekoniecznie domkniętą). Jeżeli F jest funkcjonałem liniowym ograniczonym na X, to jego obcięcie do podprzestrzeni $f = F\,\big|_M$ jest również funkcjonałem liniowym ograniczonym na M, bo mamy

$$|f(x)| = |F(x)| \quad \text{dla dowolnego} \;\; x \in M.$$

Stąd

$$\begin{aligned}\|f\|_M &= \sup\{|f(x)|\colon\, x \in M, \|x\| \leqslant 1\} \leqslant \\ &\leqslant \sup\{|F(x)|\colon\, x \in X, \|x\| \leqslant 1\} = \|F\|_X.\end{aligned}$$

Powstaje pytanie, czy każdy funkcjonał liniowy ograniczony na podprzestrzeni M jest obcięciem do M pewnego funkcjonału liniowego ograniczonego na całej przestrzeni X. Odpowiedź pozytywną na to pytanie daje następujące twierdzenie:

Twierdzenie 3.11 (Hahna-Banacha). *Jeżeli f jest ograniczonym funkcjonałem liniowym na podprzestrzeni M przestrzeni unormowanej X, to istnieje na X taki funkcjonał liniowy ograniczony F, że*

$$\begin{aligned} f(x) &= F(x) \quad \text{dla} \;\; x \in M, \\ \|f\|_M &= \|F\|_X. \end{aligned}$$

U w a g a. Twierdzenie to jest prawdziwe zarówno dla rzeczywistego, jak i zespolonego ciała skalarów. Jednakże jest ono przede wszystkim twierdzeniem „rzeczywistym". Udowodnimy je najpierw dla rzeczywistego ciała skalarów, a potem dzięki odpowiedniemu „trickowi" uzyskamy jego zespoloną wersję. W dalszym ciągu przestrzeń liniową nad ciałem liczb zespolonych będziemy krótko nazywać *zespoloną przestrzenią liniową*, a przestrzeń nad ciałem liczb rzeczywistych — *rzeczywistą przestrzenią liniową*.

Zauważmy, że każdą zespoloną przestrzeń liniową możemy traktować jako rzeczywistą przestrzeń liniową, jeżeli zawęzimy działanie mnożenia przez skalar do ciała liczb rzeczywistych. Przyjmijmy więc następującą terminologię (por. [27], s. 170–171 w części III): Niech V będzie zespoloną przestrzenią liniową. Funkcja $f\colon V \to \mathbb{C}$ jest *funkcjonałem $\mathbb{C}$-liniowym*, jeżeli jest ona liniowa ze względu na skalary zespolone, natomiast funkcja $f\colon V \to \mathbb{R}$ jest *funkcjonałem $\mathbb{R}$-liniowym*, jeżeli jest ona liniowa ze względu na skalary rzeczywiste. Zauważmy ponadto, że część rzeczywista u funkcjonału $\mathbb{C}$-liniowego f jest funkcjonałem $\mathbb{R}$-liniowym, gdyż

$$\alpha u(x) = \alpha \mathrm{Re} f(x) = \mathrm{Re}(\alpha f(x)) = \mathrm{Re} f(\alpha x) = u(\alpha x)$$

dla dowolnej liczby rzeczywistej α. Związki, jakie zachodzą pomiędzy f i u, opisuje następujący lemat:

Lemat 3.12. *Niech V będzie zespoloną przestrzenią liniową.*

(a) *Jeżeli u jest częścią rzeczywistą funkcjonału $\mathbb{C}$-liniowego na przestrzeni V, to*

$$f(x) = u(x) - iu(ix) \qquad (x \in V). \tag{3.17}$$

(b) *Jeżeli u jest funkcjonałem $\mathbb{R}$-liniowym na V, to funkcja f określona wzorem* (3.17) *jest funkcjonałem $\mathbb{C}$-liniowym na tej przestrzeni.*

(c) *Jeżeli V jest przestrzenią unormowaną i między funkcjonałami f oraz u zachodzi równość* (3.17), *to* $\|f\| = \|u\|$.

D o w ó d. Aby udowodnić część (a), zauważmy, że

$$\mathrm{Re}(if(x)) = \mathrm{Re}(f(ix)) = u(ix).$$

Jeżeli oznaczymy $z = f(x)$, to $u(x) = \mathrm{Re} z$ i (3.17) wynika z równości

$$z = \mathrm{Re}\, z - i\mathrm{Re}(iz)$$

prawdziwej dla dowolnej liczby zespolonej z.

Jeżeli u jest funkcjonałem $\mathbb{R}$-liniowym i f jest dane wzorem (3.17), to oczywiście f jest $\mathbb{R}$-liniowe. Ponadto mamy

$$f(ix) = u(ix) - iu(-x) = u(ix) + iu(x) = if(x),$$

co oznacza, że f jest $\mathbb{C}$-liniowe i kończy dowód części (b).

Ponieważ $|u(x)| \leqslant |f(x)|$ dla dowolnego $x \in V$, więc $\|u\| \leqslant \|f\|$. Aby zakończyć dowód, wystarczy pokazać, że zachodzi nierówność przeciwna. Niech $x \in V$ będzie dowolnym wektorem. Istnieje liczba zespolona α taka, że $|\alpha| = 1$ i $\alpha f(x) = |f(x)|$. Wówczas

$$|f(x)| = \alpha f(x) = f(\alpha x) = u(\alpha x) \leqslant \|u\|\|\alpha x\| = \|u\|\|x\|$$

i w konsekwencji $\|f\| \leqslant \|u\|$. ■

D o w ó d t w i e r d z e n i a H a h n a - B a n a c h a. Niech X będzie rzeczywistą przestrzenią liniową, f funkcjonałem liniowym i ograniczonym na podprzestrzeni M. Jeżeli $\|f\| = 0$, to tezę otrzymujemy, biorąc funkcjonał zerowy $F = 0$ na przestrzeni X.

W przypadku, gdy $\|f\| > 0$, możemy założyć, że $\|f\| = 1$. (Gdyby tak nie było, rozważylibyśmy funkcjonał αf, gdzie α jest odpowiednią liczbą rzeczywistą). Wybierzmy wektor $x_0 \in X \setminus M$. Niech $M_1 = \mathrm{span}\,(M \cup \{x_0\})$. Wówczas każdy wektor z przestrzeni M_1 ma postać $x + \lambda x_0$, gdzie $x \in M$ i $\lambda \in \mathbb{R}$, i to przedstawienie jest jednoznaczne. Jeżeli zdefiniujemy funkcję $f_1 \colon M_1 \to \mathbb{R}$ wzorem $f_1(x + \lambda x_0) = f(x) + \lambda\alpha$, gdzie α jest ustaloną liczbą rzeczywistą, to otrzymamy funkcjonał liniowy, który jest rozszerzeniem na M_1 funkcjonału f, tzn. $f_1(x) = f(x)$ dla $x \in M$. Teraz wykażemy, że liczbę α można tak wybrać, aby norma funkcjonału f_1 była równa 1. Aby tak było, wystarcza, aby dla $x \in M$ i $\lambda \in \mathbb{R}$ zachodziła nierówność

$$|f(x) + \lambda\alpha| \leqslant \|x + \lambda x_0\|.$$

Wstawiając wektor $-\lambda x$ zamiast x do tej nierówności, a następnie dzieląc obie strony przez $|\lambda|$, otrzymamy

$$|f(x)-\alpha| \leqslant \|x-x_0\| \qquad (x \in M)$$

lub, co jest równoważne,

$$f(x)-\|x-x_0\| \leqslant \alpha \leqslant f(x)+\|x-x_0\|.$$

Oznaczając lewą stronę tej nierówności przez a_x, a prawą — przez b_x, stwierdzamy, że szukana przez nas liczba α istnieje, gdy wszystkie przedziały $[a_x, b_x]$, $x \in M$, mają przynajmniej jeden punkt wspólny, a to ma miejsce wtedy i tylko wtedy, gdy

$$a_x \leqslant b_y$$

dla dowolnych x, $y \in M$. Mamy

$$\begin{aligned} a_x &= f(x)-\|x-x_0\| = f(x-y)+f(y)-\|(x-y)+(y-x_0)\| \leqslant \\ &\leqslant \|x-y\|+f(y)-\|x-y\|+\|y-x_0\| = b_y. \end{aligned}$$

Udowodniliśmy więc, że przekrój rodziny przedziałów $[a_x, b_x]$, $x \in M$, jest niepusty, a to oznacza, że istnieje rozszerzenie f_1 funkcjonału f o normie jeden.

Oznaczmy przez $\mathcal{P}$ zbiór wszystkich par uporządkowanych (M', f'), gdzie M' jest podprzestrzenią przestrzeni X zawierającą M oraz f' jest rozszerzeniem liniowym funkcjonału f takim, że $\|f'\| = 1$. W zbiorze $\mathcal{P}$ wprowadzamy relację częściowego porządku wzorem

$$\begin{gathered} (M', f') \prec (M'', f'') \quad \text{wtedy i tylko wtedy, gdy} \\ M' \subset M'' \ \text{ i } \ f'(x) = f''(x) \ \text{ dla } \ x \in M'. \end{gathered}$$

Jest oczywiste, że $\prec$ jest relacją częściowego porządku w zbiorze $\mathcal{P}$. Ponadto rodzina $\mathcal{P}$ jest niepusta, bo zawiera parę (M, f). Niech $\{(M_\alpha, f_\alpha)\}_\alpha$ będzie łańcuchem. Wówczas $M_0 = \bigcup\limits_\alpha M_\alpha$ jest podprzestrzenią przestrzeni X zawierającą M. Funkcjonał $f_0\colon M_0 \to \mathbb{R}$ zdefiniowany wzorem $f_0(x) = f_\alpha(x)$, gdy $x \in M_\alpha$ jest poprawnie określony. Zatem para (M_0, f_0) jest elementem górującym nad łańcuchem $\{(M_\alpha, f_\alpha)\}_\alpha$. Z lematu Kuratowskiego-Zorna wynika, że istnieje element maksymalny $(\widetilde{M}, \widetilde{f})$ w zbiorze $\mathcal{P}$. Zauważmy, że $\widetilde{M} = X$, gdyż w przeciwnym przypadku istniałoby $x_0 \in X \setminus \widetilde{M}$ i na mocy poprzedniej części dowodu istniałoby dalsze rozszerzenie funkcjonału $\widetilde{f}$, a to przeczyłoby maksymalności elementu $(\widetilde{M}, \widetilde{f})$. Zatem $\widetilde{M} = X$ i dowód w przypadku rzeczywistego ciała skalarów jest zakończony.

Niech f będzie funkcjonałem $\mathbb{C}$-liniowym na podprzestrzeni M zespolonej przestrzeni unormowanej X i niech u będzie jego częścią rzeczywistą. Wówczas na mocy lematu 3.12 wiemy, że $\|u\| = \|f\|$. Z poprzedniej części dowodu wynika, że istnieje $\mathbb{R}$-liniowe rozszerzenie U funkcjonału u na całą przestrzeń X takie, że $\|U\| = \|u\|$. Jeżeli zdefiniujemy funkcjonał $F\colon X \to \mathbb{C}$ wzorem $F(x) = U(x) - iU(ix)$, to na mocy lematu 3.12, F jest funkcjonałem $\mathbb{C}$-liniowym, $F(x) = u(x) - iu(ix) = f(x)$ dla $x \in M$ oraz $\|F\| = \|U\| = \|u\| = \|f\|$, czyli F jest żądanym rozszerzeniem funkcjonału f. ■

U w a g a. Z dowodu twierdzenia Hahna-Banacha wynika, że rozszerzenie F funkcjonału f nie jest na ogół jednoznaczne.

Twierdzenie 3.13. *Niech M będzie podprzestrzenią przestrzeni unormowanej X i niech $x_0 \in X$. Wektor x_0 należy do domknięcia $\overline{M}$ podprzestrzeni M wtedy i tylko wtedy, gdy nie istnieje funkcjonał liniowy ograniczony f na X taki, że $f(x_0) \neq 0$ i $f(x) = 0$ dla dowolnego $x \in M$.*

D o w ó d. Jeżeli $x_0 \in \overline{M}$ i f jest funkcjonałem liniowym ograniczonym na X takim, że $f(x) = 0$ dla dowolnego $x \in M$, to z ciągłości funkcjonału f otrzymujemy $f(x_0) = 0$.

Na odwrót, jeżeli $x_0 \notin \overline{M}$, to istnieje $\delta > 0$ takie, że $B_\delta(x_0) \cap M = \emptyset$, tzn. jeżeli $x \in M$, to $\|x - x_0\| \geqslant \delta$. Niech $M_1 = \operatorname{span}(M \cup \{x_0\})$. Wówczas dowolny wektor z podprzestrzeni M_1 ma jednoznaczne przedstawienie w postaci $x + \lambda x_0$, gdzie $x \in M$ i $\lambda \in \mathbb{K}$. Definiujemy funkcjonał $f\colon M_1 \to \mathbb{K}$ wzorem $f(x + \lambda x_0) = \lambda$. Ponieważ

$$\delta|f(x + \lambda x_0)| = \delta|\lambda| \leqslant |\lambda| \left\| x_0 + \frac{1}{\lambda}x \right\| = \|x + \lambda x_0\|,$$

więc f jest funkcjonałem liniowym ograniczonym na M_1 o normie mniejszej bądź równej $\frac{1}{\delta}$. Ponadto $f(x_0) = 1$ i $f(x) = 0$ dla dowolnego $x \in M$. Na mocy twierdzenia Hahna-Banacha istnieje rozszerzenie funkcjonału f na całą przestrzeń. ■

Twierdzenie 3.14 (o wydobywaniu normy za pomocą funkcjonałów). *Jeżeli x_0 jest niezerowym wektorem przestrzeni unormowanej X, to istnieje funkcjonał liniowy ograniczony f na X taki, że $\|f\| = 1$ i $f(x_0) = \|x_0\|$.*

D o w ó d. Niech $M = \operatorname{span}\{x_0\}$ i $f(\lambda x_0) = \lambda\|x_0\|$. Wówczas f jest funkcjonałem liniowym ograniczonym na M o normie równej jeden. Na mocy twierdzenia Hahna-Banacha funkcjonał ten ma rozszerzenie na całą przestrzeń. ■

Wniosek 3.15. *Jeżeli $x \neq 0$, to istnieje funkcjonał liniowy ograniczony f, dla którego $f(x) \neq 0$. Innymi słowy, funkcjonały liniowe ograniczone rozdzielają punkty przestrzeni unormowanej.*

Uwagi. 1. Wiemy, że przestrzeń sprzężona X' z przestrzenią unormowaną X jest przestrzenią Banacha z normą $\|f\| = \sup\{|f(x)|\colon \|x\| \leqslant 1\}$. Jeżeli przestrzeń X jest niezerowa, to z wniosku 3.15 wynika, że również przestrzeń X' jest nietrywialna.

Zauważmy również, że z twierdzenia 3.14 wynika, że dla dowolnego $x \in \in X$ prawdziwa jest równość $\|x\| = \sup\{|f(x)|\colon f \in X', \|f\| = 1\}$. Zatem jeżeli ustalimy wektor $x \in X$, to odwzorowanie $f \mapsto f(x)$ jest funkcjonałem liniowym ograniczonym na przestrzeni X', którego norma jest równa $\|x\|$. Ten związek pomiędzy przestrzenią X a przestrzenią z nią sprzężoną X' jest podstawą tzw. teorii dualności, której omawianie wykracza poza ramy tego podręcznika. Zainteresowany Czytelnik może zajrzeć do jakiejś książki z analizy funkcjonalnej, np. do [2], [20], [24] lub [5].

2. W przypadku, gdy przestrzeń X jest ośrodkowa, można udowodnić twierdzenie Hahna-Banacha, nie stosując pewnika wyboru. Przestrzeń X zawiera wówczas zbiór przeliczalny gęsty $Z = \{x_1, x_2, \dots\}$. Niech $M_1 = = \operatorname{span}(M \cup \{x_1\})$, $M_2 = \operatorname{span}(M_1 \cup \{x_2\})$, itd. Mamy więc

$$M_n = \operatorname{span}(M_{n-1} \cup \{x_n\}) \quad \text{dla} \quad n = 2, 3, \dots$$

Rozszerzamy funkcjonał f z zachowaniem normy najpierw na podprzestrzeń M_1, potem na M_2, itd. Niech (f_n) oznacza ciąg otrzymanych w ten sposób rozszerzeń funkcjonału f. Zauważmy, że zbiór $M_0 = \bigcup\limits_{n=1}^{\infty} M_n$ jest gęstą podprzestrzenią przestrzeni X, gdyż z konstrukcji wynika, że $M_n \subset M_m$ dla $n \leqslant m$ oraz $Z \subset M_0$. Definiujemy funkcjonał $f_0\colon M_0 \to \mathbb{K}$ wzorem $f_0(x) = f_n(x)$ dla $x \in M_n$. Ponieważ $\|f_n\| = \|f\|$ dla dowolnego n, więc f_0 jest funkcjonałem liniowym ograniczonym, będącym rozszerzeniem funkcjonału f na podprzestrzeń M_0, dla którego $\|f_0\| = \|f\|$. Jeżeli $x \in X$, to istnieje ciąg $x_n \in M_0$ taki, że $x_n \to x$, gdy $n \to \infty$. Z nierówności

$$|f_0(x_n) - f_0(x_m)| \leqslant \|f_0\|\|x_n - x_m\| \qquad (n, m = 1, 2, \dots)$$

wynika, że ciąg $(f_0(x_n))$ spełnia warunek Cauchy'ego. Wobec zupełności ciała skalarów jest on zbieżny. Niech

$$F(x) = \lim_{n\to\infty} f_0(x_n).$$

Element $F(x)$ jest wyznaczony jednoznacznie, tzn. nie zależy od ciągu (x_n). Albowiem jeżeli (x'_n) jest innym ciągiem wektorów z podprzestrzeni M_0 zbieżnym do x, to

$$|f_0(x_n) - f_0(x'_n)| \leqslant \|f_0\|\|x_n - x'_n\| \qquad (n = 1, 2, \dots).$$

Przechodząc w tej nierówności do granicy, otrzymujemy

$$\lim_{n\to\infty} f_0(x_n) = \lim_{n\to\infty} f_0(x'_n).$$

Jest łatwo sprawdzić, że funkcjonał F jest liniowy. Jeżeli przejdziemy do granicy, $n \to \infty$, w nierówności

$$|f_0(x_n)| \leqslant \|f_0\|\|x_n\|,$$

to otrzymamy nierówność

$$|F(x)\| \leqslant \|f_0\|\|x\|,$$

która oznacza, że funkcjonał F jest ograniczony oraz

$$\|F\| \leqslant \|f_0\| = \|f\|.$$

Zauważmy ponadto, że dla $x \in M_0$ mamy

$$|f_0(x)| = |F(x)| \leqslant \|F\|\|x\|,$$

czyli

$$\|f_0\| \leqslant \|F\|.$$

Ostatecznie więc $\|F\| = \|f_0\|$.

Ćwiczenia

1. Korzystając z twierdzenia Baire'a, wykazać, że każda nieskończenie wymiarowa przestrzeń Banacha ma nieprzeliczalną bazę Hamela.

2. Niech (α_n) będzie takim ciągiem liczbowym, że szereg $\sum\limits_{n=1}^{\infty} \alpha_n \xi_n$ jest zbieżny dla każdego ciągu $(\xi_n) \in \ell_1$. Wykazać, że (α_n) jest ciągiem ograniczonym.

3. Niech (α_n) będzie takim ciągiem liczbowym, że szereg $\sum\limits_{n=1}^{\infty} \alpha_n \xi_n$ jest zbieżny dla każdego ciągu $(\xi_n) \in c_0$. Wykazać, że ciąg $(\alpha_n) \in \ell_1$.

4. Mówimy, że normy $\|\cdot\|_1$ i $\|\cdot\|_2$ na przestrzeni liniowej X są *zgodne*, jeżeli dla dowolnego ciągu (x_n) wektorów z przestrzeni X z warunków $\|x_n - x\|_1 \to 0$, $\|x_n - \widetilde{x}\|_2 \to 0$ wynika, że $x = \widetilde{x}$. Wykazać, że jeżeli normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są zgodne oraz przestrzeń X jest zupełna z każdą z nich, to normy te są równoważne.

5. Niech X będzie przestrzenią liniową złożoną z ciągów liczbowych z działaniami algebraicznymi określonymi w naturalny sposób. Wykazać, że jeżeli normy $\|\cdot\|_1$ i $\|\cdot\|_2$ na przestrzeni liniowej X spełniają warunek:

$(*)$ *jeżeli* $x_n = \big(\xi_k^{(n)}\big) \in X$ $(n = 1, 2, \dots)$ *i* $\|x_n\|_j \to 0$, *to* $\xi_k^{(n)} \to 0$ *dla* $k = 1, 2, \dots$ *oraz* $j = 1, 2$

i przestrzeń X jest zupełna z każdą z tych norm, to normy te są równoważne.

6. Niech X i Y będą przestrzeniami Banacha złożonymi z ciągów liczbowych z działaniami algebraicznymi określonymi w naturalny sposób oraz niech normy w obu przestrzeniach spełniają warunek $(*)$ z poprzedniego ćwiczenia. Niech $\big(\alpha_k^j\big)_{j,k=1}^{\infty}$ będzie nieskończoną macierzą liczbową o tej własności, że dla każdego ciągu $x = (\xi_k) \in X$ szeregi $\sum\limits_{k=1}^{\infty} \alpha_k^j \xi_k$ $(j = 1, 2, \dots)$ są zbieżne i ciąg $Ax = \Big(\sum\limits_{k=1}^{\infty} \alpha_k^j \xi_k\Big)_{j=1}^{\infty}$ należy do przestrzeni Y. Wykazać, że A jest operatorem liniowym ograniczonym.

7. Niech M będzie podprzestrzenią liniową przestrzeni unormowanej X oraz niech $A\colon M \to \ell_\infty$ będzie operatorem liniowym ograniczonym. Wykazać, że A da się rozszerzyć bez zmiany normy do operatora liniowego ograniczonego na całej przestrzeni X.

8. Niech X będzie przestrzenią unormowaną, a Y przestrzenią Banacha. Ponadto niech M będzie gęstą podprzestrzenią liniową przestrzeni X. Wykazać, że dla każdego operatora liniowego ograniczonego $A\colon M \to Y$ istnieje operator liniowy ograniczony $\widetilde{A}\colon X \to Y$ taki, że $Ax = \widetilde{A}x$ dla każdego $x \in M$ oraz $\|A\| = \|\widetilde{A}\|$.

9. Niech f będzie funkcjonałem liniowym ograniczonym na podprzestrzeni liniowej M przestrzeni Hilberta H. Wykazać, że istnieje dokładnie jeden funkcjonał liniowy ograniczony F na przestrzeni H, który jest rozszerzeniem funkcjonału f i ma identyczną normę jak f. Ponadto wykazać, że F znika na podprzestrzeni $M^{\perp}$.

10. Skonstruować funkcjonał liniowy ograniczony na podprzestrzeni liniowej przestrzeni $L_1[0, 1]$, który ma dwa różne rozszerzenia liniowe na całą przestrzeń zachowujące jego normę.

11. Wykazać, że jeżeli X i Y są niezerowymi przestrzeniami unormowanymi, to przestrzeń $\mathcal{B}(X, Y)$ jest również niezerowa.

12. Wykazać, że jeżeli przestrzeń unormowana X jest nieskończenie wymiarowa, to przestrzeń do niej sprzężona X' jest również nieskończenie wymiarowa.

13. Udowodnić, że ciąg wektorów (x_n) z przestrzeni unormowanej X jest ograniczony wtedy i tylko wtedy, gdy dla każdego funkcjonału $f \in X'$ ciąg $(f(x_n))$ jest ograniczony.

Rozdział 4

Operatory ograniczone na przestrzeni Hilberta

W tym rozdziale będziemy się zajmowali operatorami liniowymi i ograniczonymi określonymi na (zespolonej) przestrzeni Hilberta H. Przestrzeń takich operatorów oznaczamy symbolem $\mathcal{B}(H)$. Wiemy, że jest ona przestrzenią (nawet algebrą) Banacha z normą operatorową

$$\|A\| = \sup_{\|x\| \leqslant 1} \|Ax\|.$$

4.1. Wstępne informacje

W dalszym ciągu wszystkie rozpatrywane przestrzenie, podprzestrzenie i operatory będą zawsze liniowe i dlatego na ogół słowo *liniowe* będziemy opuszczać.

Często jest użyteczne następujące twierdzenie:

Twierdzenie 4.1. *Niech $A \in \mathcal{B}(H)$. Jeżeli $\langle Ax, x\rangle = 0$ dla dowolnego $x \in H$, to $A = 0$.*

D o w ó d. Dla dowolnych wektorów $x, y \in H$ mamy $\langle A(x+y), x+y\rangle = 0$. Stąd

$$\langle Ax, y\rangle + \langle Ay, x\rangle = 0. \tag{4.1}$$

Zastępując y przez iy w równości (4.1), otrzymujemy

$$-i\langle Ax, y\rangle + i\langle Ay, x\rangle = 0. \tag{4.2}$$

Mnożąc (4.2) przez i, a następnie dodając otrzymaną równość do (4.1), dostajemy

$$\langle Ax, y\rangle = 0.$$

Ponieważ ostatnia równość zachodzi dla dowolnych $x, y \in H$, więc wstawiając do niej $y = Ax$, mamy $\|Ax\|^2 = 0$. Z dowolności x wynika, że $A = 0$.

■

Wniosek 4.2. *Jeżeli dla operatorów $A, B \in \mathcal{B}(H)$ i dowolnego $x \in H$ zachodzi równość $\langle Ax, x\rangle = \langle Bx, x\rangle$, to $A = B$.*

U w a g a. Twierdzenie 4.1 nie jest prawdziwe w przestrzeni Hilberta nad ciałem liczb rzeczywistych. Wystarczy wziąć operator $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ na przestrzeni $\mathbb{R}^2$. Wówczas $\langle Ax, x\rangle = 0$ dla każdego $x \in \mathbb{R}^2$.

4.2. Operator sprzężony

Niech H będzie przestrzenią Hilberta i niech $A \in \mathcal{B}(H)$. Dla dowolnego ustalonego wektora $y \in H$ funkcjonał f_y określony wzorem

$$f_y(x) = \langle Ax, y\rangle$$

jest funkcjonałem liniowym i ograniczonym. Na mocy twierdzenia Riesza (tw. 2.10) istnieje dokładnie jeden wektor $y^* \in H$ taki, że

$$f_y(x) = \langle x, y^*\rangle \qquad \text{dla każdego} \quad x \in H.$$

Wzór $A^*y = y^*$ definiuje odwzorowanie $A^* \colon H \to H$ spełniające związek

$$\langle Ax, y\rangle = \langle x, A^*y\rangle \qquad \text{dla dowolnych} \quad x, y \in H.$$

Odwzorowanie A^* nazywamy *operatorem sprzężonym* z operatorem A.

Twierdzenie 4.3. *Operator A^* sprzężony z operatorem liniowym ograniczonym A na przestrzeni Hilberta H jest liniowy i ograniczony.*

D o w ó d. Najpierw udowodnimy liniowość operatora A^*. Niech $y_1, y_2 \in H$. Wtedy dla dowolnego $x \in H$ mamy

$$\begin{aligned}\langle x, A^*(y_1 + y_2)\rangle &= \langle Ax, y_1 + y_2\rangle = \langle Ax, y_1\rangle + \langle Ax, y_2\rangle = \\ &= \langle x, A^*y_1\rangle + \langle x, A^*y_2\rangle = \langle x, A^*y_1 + A^*y_2\rangle.\end{aligned}$$

Stąd

$$A^*(y_1 + y_2) = A^*y_1 + A^*y_2.$$

Podobnie, jeżeli $y \in H$, a α jest skalarem, to dla dowolnego $x \in H$

$$\langle x, A^*(\alpha y)\rangle = \langle Ax, \alpha y\rangle = \overline{\alpha}\langle Ax, y\rangle = \overline{\alpha}\langle x, A^*y\rangle = \langle x, \alpha A^*y\rangle$$

i w konsekwencji

$$A^*(\alpha y) = \alpha A^*y.$$

Teraz pokażemy, że A^* jest operatorem ograniczonym. Mamy

$$\begin{aligned}\|A^*y\|^2 &= \langle A^*y, A^*y\rangle = \langle AA^*y, y\rangle \leqslant \\ &\leqslant \|AA^*y\|\|y\| \leqslant \|A\|\|A^*y\|\|y\|.\end{aligned}$$

Stąd wynika, że

$$\|A^*y\| \leqslant \|A\|\|y\| \qquad \text{dla} \quad y \in H.$$

Oznacza to, że operator A^* jest ograniczony oraz

$$\|A^*\| \leqslant \|A\|. \tag{4.3}$$ ■

Twierdzenie 4.4. *Niech H będzie przestrzenią Hilberta i niech $A, A_1, A_2 \in \in \mathcal{B}(H)$ oraz $\alpha \in \mathbb{C}$. Wówczas*

(a) $(A^*)^* = A$;

(b) $(A_1 + A_2)^* = A_1^* + A_2^*$;

(c) $(\alpha A)^* = \overline{\alpha} A^*$;

(d) $(A_1 A_2)^* = A_2^* A_1^*$.

D o w ó d. (a) Mamy

$$\langle A^*x, y\rangle = \langle x, (A^*)^*y\rangle.$$

Ale również

$$\langle A^*x, y\rangle = \overline{\langle y, A^*x\rangle} = \overline{\langle Ay, x\rangle} = \langle x, Ay\rangle.$$

Zatem

$$\langle x, Ay\rangle = \langle x, (A^*)^*y\rangle,$$

skąd wynika (a). W dalszym ciągu będziemy pisać A^{**} zamiast $(A^*)^*$.

Dowody (b) i (c) pozostawiamy jako (łatwe) ćwiczenie dla Czytelnika. Część (d) tezy wynika z następujących równości:

$$\begin{aligned}\langle x, (A_1A_2)^*y\rangle &= \langle (A_1A_2)x, y\rangle = \langle A_1(A_2x), y\rangle = \\ &= \langle A_2x, A_1^*y\rangle = \langle x, A_2^*(A_1^*y)\rangle = \langle x, (A_2^*A_1^*)y\rangle.\end{aligned}$$ ■

Wniosek 4.5. *Dla operatora liniowego ograniczonego A na przestrzeni Hilberta zachodzą równości*

$$\|A\| = \|A^*\| = \|A^*A\|^{1/2}.$$

D o w ó d. Ponieważ $A = A^{**}$, więc z (4.3) otrzymujemy

$$\|A\| = \|A^{**}\| \leqslant \|A^*\|.$$

Zatem $\|A^*\| = \|A\|$.

Dalej mamy

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle \leqslant \|x\|\|A^*Ax\| \leqslant \|x\|^2\|A^*A\|.$$

Stąd otrzymujemy nierówność

$$\|A\|^2 \leqslant \|A^*A\|.$$

Z submultyplikatywności normy operatorowej otrzymujemy

$$\|A^*A\| \leqslant \|A^*\|\|A\| = \|A\|^2. \qquad \blacksquare$$

U w a g a. Niech $A \in \mathcal{B}(H)$. Wówczas możemy zdefiniować operator sprzężony z operatorem A w sensie teorii przestrzeni Banacha. Jest to operator $A'\colon H' \to H'$ (H' jest przestrzenią sprzężoną z przestrzenią H) zdefiniowany za pomocą równości: $A'f = f \circ A$. Z twierdzenia Riesza (tw. 2.10) wynika, że istnieje bijekcja $J\colon H' \to H$, która spełnia warunek

$$f(x) = \langle x, J(f)\rangle$$

dla dowolnych $x \in H$ i $f \in H'$. Zatem dla takich x i f mamy

$$\begin{aligned}\langle x, (J \circ A')(f)\rangle &= \langle x, J(A'f)\rangle = (A'f)x = f(Ax) = \langle Ax, J(f)\rangle = \\ &= \langle x, A^*(J(f))\rangle = \langle x, (A^* \circ J)(f)\rangle.\end{aligned}$$

Stąd $J \circ A' = A^* \circ J$ i w konsekwencji $A^* = J \circ A' \circ J^{-1}$. Z tej równości możemy otrzymać wszystkie własności operatora A^* sprzężonego z operatorem A w sensie teorii przestrzeni Hilberta z odpowiednich własności operatora A' sprzężonego z operatorem A w sensie teorii przestrzeni Banacha. Jednakże prościej jest badać wprost operator A^*, nie odwołując się do własności operatora A'.

Przykład 4.1. Wyznaczymy operator sprzężony z operatorem S przesunięcia jednostronnego na przestrzeni ℓ_2. Operator ten określony jest wzorem:

$$S(\xi_1, \xi_2, \xi_3 \dots) = (0, \xi_1, \xi_2, \xi_3, \dots) \quad \text{dla} \quad x = (\xi_j)_{j=1}^{\infty} \in \ell_2.$$

Niech wektory e_n, $n = 1, 2, \dots$ stanowią bazę standardową przestrzeni ℓ_2, tzn. $e_n = (\delta_{nj})_{j=1}^{\infty}$ (δ_{nj} oznacza *deltę Kroneckera*, tzn. $\delta_{nj} = 1$ dla $n = j$ oraz $\delta_{nj} = 0$ w pozostałych przypadkach). Zatem mamy

$$Se_n = e_{n+1} \quad \text{dla} \quad n = 1, 2, \dots$$

Aby wyznaczyć operator S^*, wystarczy znaleźć wartości, jakie przyjmuje on na wektorach e_n. Dla dowolnych wskaźników n, $k \geqslant 1$ mamy

$$\langle S^* e_n, e_k \rangle = \langle e_n, Se_k \rangle = \langle e_n, e_{k+1} \rangle = \delta_{n,k+1}.$$

Stąd dla $n > 1$

$$S^* e_n = \sum_{k=1}^{\infty} \langle S^* e_n, e_k \rangle e_k = \sum_{k=1}^{\infty} \delta_{n,k+1} e_k = e_{n-1}.$$

Natomiast $\langle S^* e_1, e_k \rangle = 0$ dla wszystkich k, stąd $S^* e_1 = 0$. Zatem ostatecznie operator S^* jest wyznaczony wzorem

$$S^*(\xi_1, \xi_2, \xi_3, \dots) = (\xi_2, \xi_3, \xi_4, \dots) \quad \text{dla} \quad x = (\xi_j)_{j=1}^{\infty} \in \ell_2.$$

Definicja 4.1. Niech H będzie przestrzenią Hilberta i niech $A \in \mathcal{B}(H)$. Operator A jest

(i) *samosprzężony* lub *hermitowski*, jeżeli $A^* = A$;

(ii) *normalny*, jeżeli $A^*A = AA^*$;

(iii) *unitarny*, jeżeli $A^*A = AA^* = I$.

U w a g i. 1. Z definicji tej wynika, że operator samosprzężony jest normalny, jak również operator unitarny jest operatorem normalnym.

2. Jeżeli potraktujemy operację brania operatora sprzężonego jako odpowiednik brania liczby sprzężonej do liczby zespolonej, to operatory samosprzężone odpowiadają liczbom rzeczywistym, a unitarne — liczbom zespolonym o module równym jeden.

3. Zauważmy, że operator A jest samosprzężony wtedy i tylko wtedy, gdy

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad \text{dla dowolnych} \quad x,\ y \in H. \tag{4.4}$$

Twierdzenie 4.6. *Operator A jest samosprzężony wtedy i tylko wtedy, gdy iloczyn skalarny $\langle Ax, x \rangle$ jest liczbą rzeczywistą dla dowolnego $x \in H$.*

D o w ó d. Jeżeli A jest operatorem samosprzężonym, to z (4.4) otrzymujemy równość

$$\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}.$$

Zatem liczba $\langle Ax, x\rangle$ jest rzeczywista.

Na odwrót, załóżmy, że liczba $\langle Ax, x\rangle$ jest rzeczywista dla dowolnego $x \in H$. Zatem dla dowolnych $x, y \in H$ i $\alpha \in \mathbb{C}$ liczba $\langle A(y + \alpha x), y + \alpha x\rangle$ jest równa liczbie do niej sprzężonej. Mamy

$$\langle A(y + \alpha x), y + \alpha x\rangle = \langle Ay, y\rangle + \overline{\alpha}\langle Ay, x\rangle + \alpha\langle Ax, y\rangle + |\alpha|^2\langle Ax, x\rangle.$$

Zatem

$$\alpha\langle Ax, y\rangle + \overline{\alpha}\langle Ay, x\rangle = \overline{\alpha}\langle y, Ax\rangle + \alpha\langle x, Ay\rangle = \overline{\alpha}\langle A^*y, x\rangle + \alpha\langle A^*x, y\rangle.$$

Wstawiając do powyższej równości $\alpha = 1$, a następnie $\alpha = i$, otrzymujemy

$$\begin{aligned} \langle Ax, y\rangle + \langle Ay, x\rangle &= \langle A^*y, x\rangle + \langle A^*x, y\rangle, \\ i\langle Ax, y\rangle - i\langle Ay, x\rangle &= -i\langle A^*y, x\rangle + i\langle A^*x, y\rangle. \end{aligned}$$

Mnożąc drugie z tych równań przez $-i$ i dodając stronami, otrzymujemy $\langle Ax, y\rangle = \langle A^*x, y\rangle$. Stąd $A = A^*$. ■

Twierdzenie 4.7. *Norma operatora samosprzężonego A wyraża się wzorem*

$$\|A\| = \sup_{\|x\|\leqslant 1} |\langle Ax, x\rangle|.$$

D o w ó d. Niech $\mu = \sup\{|\langle Ax, x\rangle|\colon \|x\| \leqslant 1\}$. Wykażemy, że $\mu = \|A\|$. Z nierówności Schwarza mamy

$$|\langle Ax, x\rangle| \leqslant \|Ax\|\|x\| \leqslant \|A\|\|x\|^2.$$

Stąd

$$\mu \leqslant \|A\|. \tag{4.5}$$

Dla dowolnych $x, y \in H$ mamy

$$\begin{aligned} \langle A(x + y), x + y\rangle &= \langle Ax, x\rangle + \langle Ax, y\rangle + \langle Ay, x\rangle + \langle Ay, y\rangle, \\ \langle A(x - y), x - y\rangle &= \langle Ax, x\rangle - \langle Ax, y\rangle - \langle Ay, x\rangle + \langle Ay, y\rangle. \end{aligned}$$

Odejmując te nierówności stronami i uwzględniając (4.4), otrzymujemy

$$\begin{aligned} \langle A(x + y), x + y\rangle - \langle A(x - y), x - y\rangle &= 2\left(\langle Ax, y\rangle + \langle Ay, x\rangle\right) = \\ &= 2\left(\langle Ax, y\rangle + \overline{\langle Ax, y\rangle}\right) = 4\mathrm{Re}\,\langle Ax, y\rangle. \end{aligned} \tag{4.6}$$

Ponieważ

$$|\langle Az, z\rangle| = \left|\left\langle A(\frac{1}{\|z\|}z, \frac{1}{\|z\|}z\right\rangle\right| \|z\|^2 \leqslant \mu\|z\|^2,$$

więc z (4.6) mamy

$$|\mathrm{Re}\langle Ax, y\rangle| \leqslant \frac{1}{4}\mu\left(\|x+y\|^2 + \|x-y\|^2\right).$$

Uwzględniając tożsamość równoległoboku, otrzymujemy

$$|\mathrm{Re}\langle Ax, y\rangle| \leqslant \frac{1}{2}\mu\left(\|x\|^2 + \|y\|^2\right). \tag{4.7}$$

Niech $x \in H$ będzie dowolnym wektorem takim, że $\|x\| \leqslant 1$. Jeżeli $Ax \neq 0$, to przyjmując $y = \frac{1}{\|Ax\|}Ax$, mamy $\langle Ax, y\rangle = \|Ax\|$ i $\|y\| = 1$. Z (4.7) otrzymujemy

$$\|Ax\| \leqslant \mu.$$

Zauważmy, że nierówność ta jest również prawdziwa, gdy $Ax = 0$. Z dowolności wektora x otrzymujemy

$$\|A\| = \sup\{\|Ax\|\colon\ \|x\| \leqslant 1\} \leqslant \mu,$$

co wobec (4.5) daje $\|A\| = \mu$. ■

Przykład 4.2. Niech A będzie operatorem liniowym na przestrzeni $\mathbb{C}^n$. Operatorowi A odpowiada jednoznacznie wyznaczona macierz (α^j_k) (względem bazy standardowej w przestrzeni $\mathbb{C}^n$ lub jakiejkolwiek bazy ortonormalnej). Oznacza to, że jeżeli $x = (\xi^1, \ldots, \xi^n) \in \mathbb{C}^n$, to

$$Ax = \left(\sum_{k=1}^n \alpha^1_k \xi^k, \ldots, \sum_{k=1}^n \alpha^n_k \xi^k\right).$$

Wyznaczymy macierz operatora A^* (względem tej samej bazy przestrzeni $\mathbb{C}^n$). Jeżeli $y = (\eta^1, \ldots, \eta^n)$, to

$$\begin{aligned}\langle Ax, y\rangle &= \sum_{j=1}^n \left(\sum_{k=1}^n \alpha^j_k \xi^k\right)\overline{\eta^j} = \sum_{k=1}^n \xi^k \overline{\left(\sum_{j=1}^n \overline{\alpha^j_k}\eta^j\right)} = \\ &= \sum_{j=1}^n \xi^j \overline{\left(\sum_{k=1}^n \overline{\alpha^k_j}\eta^k\right)} = \langle x, A^* y\rangle.\end{aligned}$$

Stąd

$$A^*y = \Big(\sum_{k=1}^{n} \overline{\alpha_1^k}\eta^k, \ldots, \sum_{k=1}^{n} \overline{\alpha_n^k}\eta^k\Big).$$

Udowodniliśmy więc następujący fakt:

Twierdzenie 4.8. *Jeżeli operator liniowy A na przestrzeni $\mathbb{C}^n$ jest określony poprzez macierz (α_k^j) (względem bazy ortonormalnej tej przestrzeni), to operator do niego sprzężony A^* jest określony przez macierz $\left(\overline{\alpha_j^k}\right)$ (względem tej samej bazy). Operator A jest samosprzężony wtedy i tylko wtedy, gdy*

$$\alpha_k^j = \overline{\alpha_j^k} \qquad (j,\, k = 1, 2, \ldots, n).$$

Operator A jest unitarny wtedy i tylko wtedy, gdy

$$\sum_{r=1}^{n} \alpha_r^j \overline{\alpha_r^k} = \delta_{jk}.$$

Macierze spełniające powyższy warunek noszą nazwę macierzy unitarnych.

Jeżeli $A \in \mathcal{B}(H)$, to $B = \dfrac{A + A^*}{2}$ i $C = \dfrac{A - A^*}{2i}$ są operatorami samosprzężonymi oraz $A = B + iC$. Nazywane są one odpowiednio *częścią rzeczywistą* i *częścią urojoną* operatora A.

Twierdzenie 4.9. *Niech $A \in \mathcal{B}(H)$. Następujące warunki są równoważne*:

(a) *A jest operatorem normalnym*;

(b) $\|Ax\| = \|A^*x\|$ *dla każdego $x \in H$*;

(c) *części rzeczywista i urojona operatora A są przemienne.*

D o w ó d. Z równości

$$\|Ax\|^2 - \|A^*x\|^2 = \langle A^*A - AA^*x, x\rangle$$

i twierdzenia 4.1 wynika, że warunki (a) i (b) są równoważne.

Natomiast równoważność (a) i (c) wynika z równości:

$$\begin{aligned} A^*A &= B^2 - iCB + iBC + C^2, \\ AA^* &= B^2 + iCB - iBC + C^2. \qquad \blacksquare \end{aligned}$$

Twierdzenie 4.10. *Operator $A \in \mathcal{B}(H)$ jest unitarny wtedy i tylko wtedy, gdy jest izomorfizmem przestrzeni H na siebie.*

D o w ó d. Niech A będzie operatorem unitarnym. Wówczas z definicji operatora odwrotnego mamy $A^* = A^{-1}$. Ponadto

$$\langle Ax, Ay\rangle = \langle x, A^*Ay\rangle = \langle x, y\rangle.$$

Na odwrót, przypuśćmy, że A jest izomorfizmem przestrzeni Hilberta. Wówczas

$$\langle Ax, Ay\rangle = \langle x, y\rangle \quad \text{dla dowolnych} \quad x, y \in H.$$

To oznacza, że

$$\langle x, A^*Ay\rangle = \langle x, y\rangle,$$

czyli

$$A^*A = I.$$

Ponieważ operator A^{-1} jest również izomorfizmem przestrzeni Hilberta, więc

$$(A^{-1})^*A^{-1} = I.$$

Jeżeli A jest operatorem odwracalnym, to $(A^*)^{-1} = (A^{-1})^*$. Zatem

$$(A^*)^{-1}A^{-1} = I.$$

Biorąc odwrotności obu stron, otrzymujemy ostatecznie

$$AA^* = I. \qquad \blacksquare$$

Twierdzenie 4.11. *Dla operatora $A \in \mathcal{B}(H)$ zachodzi równość*

$$\ker A = (\operatorname{im} A^*)^{\perp}. \tag{4.8}$$

D o w ó d. Jeżeli $x \in \ker A$ i $y \in H$, to mamy

$$\langle x, A^*y\rangle = \langle Ax, y\rangle = 0.$$

Stąd wynika, że $\ker A \subset (\operatorname{im} A^*)^{\perp}$.

Natomiast, jeżeli $x \perp \operatorname{im} A^*$ i $y \in H$, to

$$\langle Ax, y\rangle = \langle x, A^*y\rangle = 0,$$

co oznacza, że $(\operatorname{im} A^*)^{\perp} \subset \ker A$. $\blacksquare$

U w a g i. 1. Ponieważ $A^{**} = A$, więc $\ker A^* = (\operatorname{im} A)^{\perp}$.

2. Natomiast $(\ker A)^{\perp}$ nie musi być równe $\operatorname{im} A^*$, ale mamy

$$(\ker A)^{\perp} = (\operatorname{im} A^*)^{\perp\perp} = \overline{\operatorname{im} A^*}$$

i analogicznie $(\ker A^*)^{\perp} = \overline{\operatorname{im} A}$.

Na zakończenie udowodnimy twierdzenie charakteryzujące operatory rzutowania ortogonalnego.

Twierdzenie 4.12. *Niech operator $P \in \mathcal{B}(H)$ spełnia warunek $P^2 = P$. Następujące warunki są równoważne:*

(a) *P jest operatorem samosprzężonym;*

(b) *P jest operatorem normalnym;*

(c) $\operatorname{im} P = (\ker P)^{\perp}$;

(d) $\langle Px, x\rangle = \|Px\|^2$ *dla każdego $x \in H$.*

D o w ó d. Jest oczywiste, że warunek (a) implikuje (b). Teraz pokażemy, że z (b) wynika (c). Jeżeli P jest operatorem normalnym, to

$$\ker P = \ker P^* = (\operatorname{im} P)^{\perp}.$$

Ponieważ $\operatorname{im} P$ jest zbiorem domkniętym, więc mamy

$$(\ker P)^{\perp} = (\operatorname{im} P)^{\perp\perp} = \overline{\operatorname{im} P} = \operatorname{im} P.$$

Aby wykazać, że (d) jest konsekwencją (c), weźmy dowolne $x \in H$. Wówczas $x = y + z$, gdzie $y \perp z$, $Py = 0$ i $Pz = z$. Ponieważ $Px = z$, więc mamy

$$\langle Px, x\rangle = \langle z, z\rangle = \|z\|^2 = \|Px\|^2.$$

Z twierdzenia 4.6 wynika, że warunek (d) implikuje (a). ■

4.3. Podprzestrzenie niezmiennicze i redukujące

Definicja 4.2. Niech A będzie operatorem liniowym ograniczonym na przestrzeni Hilberta H i niech $M \subset H$ będzie jej domkniętą podprzestrzenią. M jest *podprzestrzenią niezmienniczą* operatora A, jeżeli $A(M) \subset M$. Natomiast M jest *podprzestrzenią redukującą* operator A, jeżeli $A(M) \subset M$ i $A(M^{\perp}) \subset M^{\perp}$.

Niech M będzie domkniętą podprzestrzenią przestrzeni Hilberta H oraz niech $P = P_M$ będzie rzutem ortogonalnym na M. Wówczas $I - P$ jest rzutem ortogonalnym na $M^{\perp}$.

Twierdzenie 4.13. *Niech $A \in \mathcal{B}(H)$. Wówczas M jest podprzestrzenią niezmienniczą operatora A wtedy i tylko wtedy, gdy $PAP = AP$.*

D o w ó d. Załóżmy, że M jest podprzestrzenią niezmienniczą operatora A i niech $x \in H$. Wówczas $Px \in M$ i $APx \in M$. Stąd $P(APx) = APx$, a to oznacza, że $PAP = AP$.

Na odwrót, jeżeli $PAP = AP$ i $x \in M$, to $Ax = APx = PAPx \in M$. ■

Twierdzenie 4.14. *Niech $A \in \mathcal{B}(H)$. Następujące warunki są równoważne:*

(a) *podprzestrzeń M redukuje operator A;*

(b) $PA = AP$;

(c) *M jest podprzestrzenią niezmienniczą dla A i A^*.*

Dowód. Najpierw wykażemy, że (a) implikuje (b). Z poprzedniego twierdzenia wynika, że $AP = PAP$ oraz $A(I-P) = (I-P)A(I-P)$. Z ostatniej równości wynika, że $PA = PAP$ i w konsekwencji $AP = PA$.

Teraz pokazujemy, że (c) jest konsekwencją (b). Jest oczywiste, że z równości $PA = AP$ wynika $PAP = AP$, a więc M jest podprzestrzenią niezmienniczą operatora A. Biorąc sprzężenia po obu stronach równości $PA = AP$, otrzymujemy $A^*P = PA^*$, co oznacza, że M jest podprzestrzenią niezmienniczą operatora A^*.

W końcu udowodnimy, że z (c) wynika (a). Musimy wykazać, że $M^\perp$ jest podprzestrzenią niezmienniczą operatora A. Niech $x \perp M$ i $y \in M$. Wówczas $A^*y \in M$ oraz $0 = \langle A^*y, x\rangle = \langle y, Ax\rangle$. Stąd $Ax \perp M$. ∎

Uwaga. Jeżeli M jest podprzestrzenią niezmienniczą operatora $A \in \mathcal{B}(H)$, to $A\big|_M$ jest operatorem ograniczonym przekształcającym przestrzeń M w siebie, czyli $A\big|_M \in \mathcal{B}(M)$. Mamy ponadto $\big\|A\big|_M\big\| \leqslant \|A\|$. Jeżeli podprzestrzeń M redukuje operator A, to również $A\big|_{M^\perp} \in \mathcal{B}(M^\perp)$. Oznaczmy $A_1 = A\big|_M$ i $A_2 = A\big|_{M^\perp}$. Dla dowolnego $x \in H$ z rozkładu $x = x_1 + x_2$, gdzie $x_1 \in M$ i $x_2 \in M^\perp$, otrzymujemy równość $Ax = A_1x_1 + A_2x_2$. Zatem badanie operatora A jest zredukowane do badania „mniejszych” operatorów A_1 i A_2, które mogą mieć prostszą budowę niż operator A. Stąd też wzięła się nazwa tego rodzaju podprzestrzeni.

Wniosek 4.15. *Jeżeli $A \in \mathcal{B}(H)$ i M jest podprzestrzenią redukującą operator A, to $\left(A\big|_M\right)^* = A^*\big|_M$. Jeżeli ponadto operator A jest normalny, to $A\big|_M$ jest również normalny.*

Dowód. Wiemy, że $A\big|_M$, $A^*\big|_M \in \mathcal{B}(M)$. Ponadto dla dowolnych $x, y \in$ $\in M$ mamy

$$\langle Ax, y\rangle = \langle x, A^*y\rangle.$$

Stąd wynika pierwsza część tezy. Drugą część otrzymujemy z równości

$$\|Ax\| = \|A^*x\|$$

prawdziwej dla dowolnego $x \in M$. ∎

To, czy każdy operator liniowy ograniczony na zespolonej przestrzeni Hilberta H ma nietrywialną (tzn. różną od $\{0\}$ i H) (domkniętą) podprzestrzeń niezmienniczą, jest ciągle problemem otwartym. Stosunkowo łatwo jest podać przykład operatora, który nie ma nietrywialnej podprzestrzeni redukującej.

Przykład 4.3. Operator S jednostronnego przesunięcia na przestrzeni ℓ_2 ma dużo nietrywialnych podprzestrzeni niezmienniczych. Na przykład, każda z podprzestrzeni $M_n = \{x \in \ell_2\colon\ x = (\xi_k),\ \xi_1 = \ldots = \xi_n = 0\}$ dla $(n = 1, 2, \ldots)$ jest niezmiennicza dla S. Natomiast nie ma on nietrywialnych podprzestrzeni redukujących. Istotnie, jeżeli M byłaby taką podprzestrzenią, to zawierałaby ona niezerowy wektor x. Niech $x = \sum\limits_{n=1}^{\infty} \xi_n e_n$ będzie jego rozwinięciem w szereg Fouriera względem standardowej bazy ortonormalnej (e_n) przestrzeni ℓ_2. Załóżmy, że ξ_m jest pierwszym niezerowym współczynnikiem w tym rozwinięciu. Wówczas również $x = \sum\limits_{n=1}^{\infty} \xi_{n+m-1} e_{n+m-1}$. Ponieważ podprzestrzeń M redukuje operator S, więc jest ona niezmiennicza dla operatora S^*. Zatem $y = S^m (S^*)^m x = \sum\limits_{n=1}^{\infty} \xi_{n+m} e_{n+m} \in M$. Stąd wynika, że $\xi_m e_m = x - y \in M$, czyli $e_m \in M$. Ponieważ $e_n = S^{n-1}(S^*)^{m-1} e_m$, więc $e_n \in M$ dla dowolnego n. To oznacza, że $M = H$.

Twierdzenie 4.16. *Jeżeli A jest operatorem normalnym, a λ dowolną liczbą zespoloną, to* $\ker(A - \lambda I) = \ker(A - \lambda I)^*$ *i* $\ker(A - \lambda I)$ *jest podprzestrzenią redukującą operator A.*

D o w ó d. Ponieważ operator A jest normalny, więc normalny jest również $A - \lambda I$. Zatem $\|(A - \lambda I)x\| = \|(A - \lambda I)^* x\|$ i w konsekwencji $\ker(A - \lambda I) = \ker(A - \lambda I)^*$.

Jeżeli $x \in \ker(A - \lambda I)$, to $Ax = \lambda x \in \ker(A - \lambda I)$ oraz $A^* x = \overline{\lambda} x \in \ker(A - \lambda I)$. Zatem podprzestrzeń $\ker(A - \lambda I)$ redukuje operator A. ■

Ćwiczenia

1. Niech X i Y będą przestrzeniami unitarnymi. Wykazać, że każde odwzorowanie $U\colon X \xrightarrow{\text{na}} Y$ spełniające warunek

$$(*) \qquad \langle Ux_1, Ux_2 \rangle = \langle x_1, x_2 \rangle \quad \textit{dla} \quad x_1,\, x_2 \in X$$

jest operatorem liniowym.

2. Niech X i Y będą przestrzeniami unitarnymi. Załóżmy, że operator liniowy $U\colon X \to Y$ spełnia dla każdego $x \in X$ równość $\|Ux\| = \|x\|$. Dowieść, że spełnia on także warunek $(*)$.

3. Wykazać, że jeżeli operator A na przestrzeni unitarnej X ma dokładnie jedną prawą odwrotność, tzn. istnieje dokładnie jeden operator B, dla którego $AB = I_X$, to A jest odwracalny.

4. Niech operator $V\colon L_2(0,\infty) \to L_2(0,\infty)$ będzie określony wzorem

$$Vx(t) = \begin{cases} x(t-1) & \text{dla } t > 1, \\ 0 & \text{dla } t \leqslant 1. \end{cases}$$

Pokazać, że V jest izometrią, która nie jest suriekcją.

5. Niech operator $U\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$ będzie określony wzorem $Ux(t) = x(t-1)$. Pokazać, że U jest operatorem unitarnym (izomorfizmem przestrzeni Hilberta).

6. Załóżmy, że operatory liniowe A i B w przestrzeni Hilberta H spełniają równość $\langle Ax, y\rangle = \langle x, By\rangle$ dla x, $y \in H$. Udowodnić, że operatory A i B są ograniczone (i w konsekwencji $B = A^*$).

7. Znaleźć operatory sprzężone do następujących operatorów działających na przestrzeni ℓ_2:

(a) $A_1(\xi_1, \xi_2, \dots) = (0, 0, \xi_1, \xi_2, \dots)$;

(b) $A_2(\xi_1, \xi_2, \xi_3, \dots) = (\xi_3, \xi_4, \dots)$;

(c) $A_3(\xi_1, \xi_2, \dots) = (\alpha_1\xi_1, \alpha_2\xi_2, \dots)$, gdzie $(\alpha_n) \in \ell_\infty$;

(d) $A_4(\xi_1, \xi_2, \dots) = (0, 0, \alpha_1\xi_1, \alpha_2\xi_2, \dots)$, gdzie $(\alpha_n) \in \ell_\infty$;

(e) $A_5(\xi_1, \xi_2, \dots) = (\xi_1, \xi_2, \dots, \xi_n, 0, 0, \dots)$;

(f) $A_6(\xi_1, \xi_2, \dots) = (\underbrace{0, 0, \dots, 0}_{n-1}, \xi_1, 0, 0, \dots)$;

(g) $A_7(\xi_1, \xi_2, \dots) = (\alpha_n\xi_n, \alpha_{n+1}\xi_{n+1}, \dots)$, gdzie $(\alpha_n) \in \ell_\infty$.

Które z powyższych operatorów są samosprzężone?

8. Znaleźć operatory sprzężone do następujących operatorów działających w przestrzeni $L_2(\mathbb{R})$:

(a) $B_1x(t) = x(t+h)$;

(b) $B_2x(t) = a(t)x(t+h)$, gdzie $a \in L_\infty(\mathbb{R})$;

(c) $B_3x(t) = \frac{1}{2}\left(x(t) + x(-t)\right)$.

Które z powyższych operatorów są samosprzężone?

9. Dla operatora przesunięcia jednostronnego S, $(Se_n = e_{n+1})$, obliczyć SS^*, S^*S oraz S^nS^{*n} i $S^{*n}S^n$.

10. Obliczyć operator sprzężony do operatora Volterry $V \in \mathcal{B}(L_2(0,1))$, $Vx(s) = \int\limits_0^s x(t)\,dt$. Znaleźć $\operatorname{im}(V+V^*)$.

U w a g a. W dalszym ciągu wszystkie operatory będą operatorami na zespolonej przestrzeni Hilberta.

11. Udowodnić, że jeżeli operatory A i B są samosprzężone, to AB jest operatorem samosprzężonym wtedy i tylko wtedy, gdy $AB = BA$.

12. Udowodnić, że jeżeli A jest operatorem normalnym, to $\|A^n\| = \|A\|^n$.

13. Udowodnić, że jeżeli $|\alpha| = |\beta| = 1$, to operator $\alpha A + \beta A^*$ jest operatorem normalnym.

14. Udowodnić, że jeżeli $A^2 = A$ i A jest operatorem normalnym, to A jest samosprzężony.

15. Udowodnić, że jeżeli $A^2 = 0$ i A jest operatorem normalnym, to $A = 0$.

16. Niech operatory A i B będą normalne. Czy operatory AB i $A+B$ muszą być normalne? Przy jakich założeniach będą one normalne?

17. Czy z tego, że dla operatorów A i B zachodzi równość $AB = 0$ wynika, że również $BA = 0$? Czy w przypadku, gdy operatory te są normalne, odpowiedź na to pytanie jest taka sama?

18. Wykazać, że jeżeli operator $A - iI$ jest odwracalny, a $(A+iI)(A-iI)^{-1}$ jest unitarny, to A jest operatorem samosprzężonym.

19. Udowodnić, że jeżeli operator U jest unitarny, a operator $U - I$ — odwracalny, to operator $A = i(U+I)(U-I)^{-1}$ jest samosprzężony.

Rozdział 5

Elementy teorii spektralnej na przestrzeni Banacha

W tym rozdziale wprowadzimy podstawowe pojęcia teorii spektralnej, tj. widmo operatora i jego części oraz promień spektralny, a następnie zbadamy ich własności.

5.1. Widmo operatora na przestrzeni Banacha

W dalszym ciągu będziemy rozważać operatory określone na zespolonych przestrzeniach Banacha. W związku z tym termin *przestrzeń Banacha* będzie zawsze oznaczał zespoloną przestrzeń Banacha. Niech X będzie przestrzenią Banacha i niech $A \in \mathcal{B}(X)$. Zajmiemy się badaniem równania

$$Ax - \lambda x = y, \tag{5.1}$$

gdzie $x, y \in X$ i $\lambda \in \mathbb{C}$. Równanie to możemy zapisać w postaci

$$(A - \lambda I)x = y.$$

Jeżeli operator $A - \lambda I$ jest bijekcją, to równanie (5.1) ma dla każdego $y \in X$ dokładnie jedno rozwiązanie $x = (A - \lambda I)^{-1}y$. Zauważmy przy tym, że z twierdzenia Banacha o operatorze odwrotnym wynika, że wówczas $(A-\lambda I)^{-1}$ jest również operatorem ograniczonym (określonym na całej przestrzeni X).

Definicja 5.1. Zbiór tych liczb zespolonych λ, dla których $(A - \lambda I)^{-1} \notin \notin \mathcal{B}(X)$, nazywamy *widmem* operatora A i oznaczamy symbolem $\sigma(A)$. Zatem

$$\begin{aligned} \sigma(A) &= \{\lambda \in \mathbb{C}\colon\ A - \lambda I \ \text{ nie jest bijekcją}\,\} = \\ &= \{\lambda \in \mathbb{C}\colon\ A - \lambda I \ \text{ nie jest odwracalne w } \ \mathcal{B}(X)\}. \end{aligned}$$

Zbiór tych λ, dla których $A - \lambda I$ nie jest iniekcją, nazywamy *widmem punktowym* operatora A i oznaczamy symbolem $\sigma_{\mathrm{p}}(A)$. Elementy zbioru $\sigma_{\mathrm{p}}(A)$ nazywamy *wartościami własnymi* operatora A. Z definicji wynika, że

$$\sigma_{\mathrm{p}}(A) \subset \sigma(A).$$

Zauważmy, że w przypadku, gdy $X = \mathbb{C}^n$, operator A jest wyznaczony przez macierz (α_k^j) (przyjmujemy w $\mathbb{C}^n$ bazę standardową) i równanie (5.1) ma postać niejednorodnego układu równań liniowych

$$\sum_{k=1}^{n} \alpha_k^j x^k - \lambda x^j = y^j \qquad (j = 1, 2, \ldots, n).$$

Z algebry liniowej wiemy, że układ ten ma rozwiązanie przy każdym y wtedy i tylko wtedy, gdy

$$\det(\alpha_k^j - \lambda\delta_k^j) \neq 0.$$

(δ_k^j oznacza deltę Kroneckera). Zatem operator $(A - \lambda I)^{-1}$ istnieje wtedy i tylko wtedy, gdy $\det(\alpha_k^j - \lambda\delta_k^j) \neq 0$. Ponieważ w przestrzeni skończenie wymiarowej operator jest iniekcją wtedy i tylko wtedy, gdy jest suriekcją (por. [26], tw. 5.7), więc $\sigma_{\mathrm{p}}(A) = \sigma(A)$. Zatem w przypadku operatora A na przestrzeni skończenie wymiarowej mamy

$$\begin{aligned}\sigma(A) &= \{\lambda \in \mathbb{C}\colon\ \det(\alpha_k^j - \lambda\delta_k^j) = 0\} = \\ &= \{\lambda \in \mathbb{C}\colon\ \det(A - \lambda I) = 0\} = \\ &= \{\text{zbiór wartości własnych operatora } A\}.\end{aligned}$$

W przypadku przestrzeni nieskończenie wymiarowych taka sytuacja nie musi mieć miejsca. Może się bowiem zdarzyć, że operator $A - \lambda I$ jest iniekcją, ale jego obraz nie wypełnia całej przestrzeni. Zatem $(A - \lambda I)^{-1}$ nie jest określone na przestrzeni X, ale na pewnej jej podprzestrzeni. Powstaje zatem potrzeba charakteryzacji takich liczb λ. Niech więc A będzie operatorem liniowym ograniczonym na przestrzeni Banacha X i załóżmy, że $(A - \lambda I)^{-1}$ istnieje. Zbiór tych liczb λ, dla których $\operatorname{im}(A - \lambda I)$ nie jest gęsty w X, nazywamy *widmem residualnym* operatora A i oznaczamy symbolem $\sigma_{\mathrm{r}}(A)$. Jeżeli $\operatorname{im}(A - \lambda I)$ jest gęste w X, to $(A - \lambda I)^{-1}$ jest operatorem ograniczonym lub nieograniczonym. Zbiór tych λ, dla których $\overline{\operatorname{im}(A - \lambda I)} = X$ i $(A - \lambda I)^{-1}$ jest ograniczone, nazywamy *zbiorem rezolwenty* operatora A i oznaczamy przez $\varrho(A)$. Zbiór tych λ, dla których $\overline{\operatorname{im}(A - \lambda I)} = X$ i $(A - \lambda I)^{-1}$ jest nieograniczone, nazywamy *widmem ciągłym* operatora A i oznaczamy przez $\sigma_{\mathrm{c}}(A)$. Zatem mamy następującą sytuację:

— $\lambda \in \sigma_{\mathrm{p}}(A)$ wtedy i tylko wtedy, gdy $(A - \lambda I)^{-1}$ nie istnieje;
— $\lambda \in \sigma_{\mathrm{r}}(A)$ wtedy i tylko wtedy, gdy $(A - \lambda I)^{-1}$ istnieje i $\overline{\operatorname{im}(A - \lambda I)} \neq X$;
— $\lambda \in \sigma_{\mathrm{c}}(A)$ wtedy i tylko wtedy, gdy $(A - \lambda I)^{-1}$ istnieje i $\overline{\operatorname{im}(A - \lambda I)} = X$ oraz $(A - \lambda I)^{-1}$ jest nieograniczone;

— $\lambda \in \varrho(A)$ wtedy i tylko wtedy, gdy $(A-\lambda I)^{-1}$ istnieje i $\overline{\mathrm{im}(A-\lambda I)} = X$ oraz $(A-\lambda I)^{-1}$ jest ograniczone.

Wszystkie te fakty możemy zebrać w następującym twierdzeniu:

Twierdzenie 5.1. *Niech X będzie przestrzenią Banacha i niech $A \in \mathcal{B}(X)$. Wówczas $\lambda \in \varrho(A)$ wtedy i tylko wtedy, gdy $A-\lambda I$ jest bijekcją. Widma punktowe, residualne i ciągłe są parami rozłączne i ich suma mnogościowa jest równa widmu operatora A. Jeżeli przestrzeń X jest skończenie wymiarowa, to $\sigma(A) = \sigma_{\mathrm{p}}(A)$.*

D o w ó d. Musimy jedynie udowodnić pierwszą część tezy. Jeżeli $A-\lambda I$ jest bijekcją, to z twierdzenia Banacha o operatorze odwrotnym wynika, że $\lambda \in \varrho(A)$. Aby zakończyć dowód, musimy pokazać, że jest na odwrót, tzn. jeżeli $\lambda \in \varrho(A)$, to $A-\lambda I$ jest bijekcją, czyli $\mathrm{im}(A-\lambda I) = X$. Niech $y \in X$ będzie dowolne. Istnieje taki ciąg (x_n), $x_n \in X$, że $(A-\lambda I)x_n \to y$. Wówczas

$$\begin{aligned}\|x_n - x_m\| &= \|(A-\lambda I)^{-1}(A-\lambda I)x_n - (A-\lambda I)^{-1}(A-\lambda I)x_m\| \leqslant \\ &\leqslant \|(A-\lambda I)^{-1}\|\|(A-\lambda I)x_n - (A-\lambda I)x_m\|.\end{aligned}$$

Ciąg $((A-\lambda I)x_n)$ jest zbieżny, a więc spełnia warunek Cauchy'ego. Zatem i ciąg (x_n) jest ciągiem Cauchy'ego. Ponieważ przestrzeń X jest zupełna, więc $x_n \to x$ dla pewnego $x \in X$. Wówczas $(A-\lambda I)x_n \to (A-\lambda I)x$. Ale ponieważ $(A-\lambda I)x_n \to y$, więc $y = (A-\lambda I)x$. ■

Przykłady 5.1. (a) Najprostszym operatorem liniowym ograniczonym na przestrzeni Banacha X jest operator $M(\alpha)$ mnożenia przez skalar α, tzn. $M(\alpha)x = \alpha x$ dla $x \in X$. Wówczas $M(\alpha) - \lambda I = M(\alpha - \lambda)$ oraz dla $\lambda \neq \alpha$ mamy

$$M(\alpha-\lambda)M\left((\alpha-\lambda)^{-1}\right) = M\left((\alpha-\lambda)^{-1}\right)M(\alpha-\lambda) = I.$$

Natomiast dla $\lambda = \alpha$

$$\left(M(\lambda) - \lambda I\right)x = 0 \quad \text{dla dowolnego} \;\; x \in X.$$

Zatem, jeżeli $X \neq \{0\}$, to

$$\sigma(M(\alpha)) = \sigma_{\mathrm{p}}(M(\alpha)) = \{\alpha\}, \quad \sigma_{\mathrm{r}}(M(\alpha)) = \sigma_{\mathrm{c}}(M(\alpha)) = \emptyset.$$

(b) Niech A będzie operatorem liniowym ograniczonym na przestrzeni $X = \mathcal{C}[a,b]$ określonym wzorem

$$Ax(t) = tx(t) \quad (t \in [a,b]\,).$$

Wyznaczymy widmo operatora A i jego części. Operator odwrotny

$$(A-\lambda I)^{-1}x(t)=\frac{1}{t-\lambda}x(t)$$

istnieje na przestrzeni X i jest ograniczony dla $\lambda\notin[a,b]$, bo funkcja $t\mapsto$ $\mapsto(t-\lambda)^{-1}x(t)$ jest ciągła dla dowolnej funkcji ciągłej x. Natomiast dla $\lambda\in[a,b]$ operator odwrotny $(A-\lambda I)^{-1}$ jest określony na podprzestrzeni Y przestrzeni $\mathcal{C}[a,b]$ tych funkcji y ciągłych na przedziale $[a,b]$, które mają postać $t\mapsto(t-\lambda)x(t)$, gdzie $x\in\mathcal{C}[a,b]$. Jej domknięcie jest właściwą podprzestrzenią przestrzeni X, bo jest ono zawarte w jądrze ciągłego funkcjonału liniowego $x\mapsto x(\lambda)$.

Zauważmy ponadto, że operator $(A-\lambda I)^{-1}\colon Y\to X$ nie jest ograniczony. Istotnie, weźmy ciąg funkcji ciągłych (x_n) określonych wzorem

$$x_n(t)=\begin{cases}0 & \text{dla } a\leqslant t<\lambda-\dfrac{1}{n},\\ -n^2(t-\lambda)-n & \text{dla } \lambda-\dfrac{1}{n}\leqslant t\leqslant\lambda,\\ n^2(t-\lambda)-n & \text{dla } \lambda<t\leqslant\lambda+\dfrac{1}{n},\\ 0 & \text{dla } \lambda+\dfrac{1}{n}<t\leqslant b.\end{cases}$$

Wówczas $\|x_n\|=n$ oraz $\sup\limits_{a\leqslant t\leqslant b}|(t-\lambda)x_n(t)|=\frac{1}{4}$.

Zatem, jeżeli $y_n(t)=(t-\lambda)x_n(t)$, to $y_n\in Y$, $\|y_n\|=\frac{1}{4}$ dla $n=1,2,\ldots$, ale $\|(A-\lambda I)^{-1}y_n\|=\|x_n\|=n\to+\infty$.

Wykazaliśmy, że

$$\sigma(A)=\sigma_{\mathrm{r}}(A)=[a,b],\quad \sigma_{\mathrm{p}}(A)=\sigma_{\mathrm{c}}(A)=\emptyset.$$

Z tego przykładu wynika, że operator liniowy ograniczony na przestrzeni unormowanej nieskończenie wymiarowej może nie mieć wartości własnych.

Twierdzenie 5.2. *Niech A będzie operatorem liniowym ograniczonym na przestrzeni Banacha X. Jeżeli $\|A\|<1$, to operator $I-A$ jest odwracalny w $\mathcal{B}(X)$ i zachodzi wzór*

$$(I-A)^{-1}=\sum_{n=0}^{\infty}A^n.$$

Ponadto

$$\|(I-A)^{-1}\|\leqslant\frac{1}{1-\|A\|}.$$

Uwaga. Aby uprościć zapis, przyjmujemy konwencję, że $A^0 = I$ nawet wtedy, gdy A jest operatorem zerowym.

Dowód. Z założenia $\|A\| < 1$ wynika, że szereg $\sum\limits_{n=0}^{\infty} A^n$ jest zbieżny bezwzględnie. Na mocy twierdzenia 1.6 jest on zbieżny. Ponadto $\lim\limits_{n\to\infty} A^n = 0$. Niech $S_n = I + A + A^2 + \ldots + A^n$ oraz $S = \sum\limits_{n=0}^{\infty} A^n$. Wówczas $S = \lim\limits_{n\to\infty} S_n$. Ponieważ mnożenie w algebrze $\mathcal{B}(X)$ jest ciągłe (por. tw. 1.15), więc mamy

$$S(I - A) = \lim_{n\to\infty} S_n(I - A) = \lim_{n\to\infty} (I - A^{n+1}) = I.$$

Analogicznie dowodzi się, że $(I-A)S = I$. Zatem $S = (I-A)^{-1}$. Nierówność dla normy operatora $(I - A)^{-1}$ wynika z nierówności (1.13) w twierdzeniu 1.6 i ze wzoru na sumę szeregu geometrycznego. ■

Wniosek 5.3. *Niech $A \in \mathcal{B}(X)$ i $\|A\| < 1$. Wówczas równanie*

$$x - Ax = y$$

ma dla każdego $y \in X$ dokładnie jedno rozwiązanie

$$x = y + \sum_{n=1}^{\infty} A^n y.$$

Ponadto

$$\|x\| \leqslant \frac{\|y\|}{1 - \|A\|}.$$

Szereg $\sum\limits_{n=0}^{\infty} A^n y$ nosi nazwę *szeregu Neumanna.*

Wniosek 5.4. *Dla dowolnego $A \in \mathcal{B}(X)$*

$$(I - \lambda A)^{-1} \to I, \quad \textit{gdy} \ \ \lambda \to 0.$$

Dowód. Dla $A = 0$ nie ma czego dowodzić. Niech więc $A \neq 0$ i niech $0 < q < 1$. Jeżeli $|\lambda| < \dfrac{q}{\|A\|}$, to $\|\lambda A\| < q$ i wówczas

$$(I - \lambda A)^{-1} = \sum_{n=0}^{\infty} (\lambda A)^n,$$

$$\|(I - \lambda A)^{-1}\| \leqslant \sum_{n=0}^{\infty} \|\lambda A\|^n \leqslant \sum_{n=0}^{\infty} q^n = \frac{1}{1 - q}.$$

Zatem wyrażenie $\|(I-\lambda A)^{-1}\|$ jest ograniczone na otoczeniu zera. Dalej mamy

$$(I-\lambda A)^{-1} = I + \lambda A(I-\lambda A)^{-1}. \tag{5.2}$$

Ponieważ

$$\|\lambda A(I-\lambda A)^{-1}\| \leqslant |\lambda|\|A\|\|(I-\lambda A)^{-1}\| \leqslant |\lambda|\frac{\|A\|}{1-q},$$

więc

$$\|\lambda A(I-\lambda A)^{-1}\| \to 0, \quad \text{gdy} \;\; \lambda \to 0.$$

Stąd i z (5.2) otrzymujemy

$$(I-\lambda A)^{-1} \to I, \quad \text{gdy} \;\; \lambda \to 0. \qquad \blacksquare$$

Twierdzenie 5.5. *Niech A_0 będzie operatorem odwracalnym w $\mathcal{B}(X)$ i niech $\|A\| < \|A_0^{-1}\|^{-1}$. Wówczas operator $A_1 = A_0 + A$ jest odwracalny w $\mathcal{B}(X)$ i ponadto*

$$A_1^{-1} = (I + A_0^{-1}A)^{-1}A_0^{-1}.$$

D o w ó d. Mamy

$$A_1 = A_0 + A = A_0(I + A_0^{-1}A) = A_0(I-B),$$

gdzie $B = -A_0^{-1}A$. Ponieważ $\|B\| = \|A_0^{-1}A\| < 1$, więc na mocy twierdzenia 5.2 operator $I-B$ jest odwracalny w $\mathcal{B}(X)$. Ponieważ odwrotność iloczynu operatorów odwracalnych jest iloczynem ich odwrotności (mnożonych w odwrotnej kolejności), więc mamy $A_1^{-1} = (I-B)^{-1}A_0^{-1}$. $\blacksquare$

Wniosek 5.6. *Zbiór operatorów odwracalnych w algebrze $\mathcal{B}(X)$ jest otwarty (w normie operatorowej).*

Definicja 5.2. Niech X będzie przestrzenią Banacha i niech $A \in \mathcal{B}(X)$. *Rezolwentą* operatora A nazywamy funkcję $R_\lambda(A)\colon \mathbb{C}\setminus\sigma(A) \to \mathcal{B}(X)$ określoną wzorem

$$R_\lambda(A) = (A-\lambda I)^{-1}.$$

Wniosek 5.7. *Rezolwenta $R_\lambda(A)$ jest funkcją ciągłą zmiennej λ na zbiorze $\varrho(A) = \mathbb{C}\setminus\sigma(A)$.*

D o w ó d. Mamy

$$R_{\lambda_0+\lambda}(A) = (A-\lambda_0 I-\lambda I)^{-1} = (I-\lambda(A-\lambda_0 I)^{-1})^{-1}(A-\lambda_0 I)^{-1}.$$

Ponieważ na mocy wniosku 5.4

$$(I - \lambda(A - \lambda_0 I)^{-1})^{-1} \to I, \quad \text{gdy} \quad \lambda \to 0,$$

więc ostatecznie

$$R_{\lambda_0 + \lambda}(A) \to (A - \lambda_0 I)^{-1} = R_{\lambda_0}(A), \quad \text{gdy} \quad \lambda \to 0. \qquad \blacksquare$$

Twierdzenie 5.8. *Niech* λ, $\mu \in \varrho(A) = \mathbb{C} \setminus \sigma(A)$. *Wówczas*

(a) $R_\lambda(A)R_\mu(A) = R_\mu(A)R_\lambda(A)$;

(b) $R_\lambda(A) - R_\mu(A) = (\lambda - \mu)R_\lambda(A)R_\mu(A)$.

Równość z punktu (b) nosi nazwę *tożsamości Hilberta.*

Dowód. (a)

$$\begin{aligned} R_\lambda(A)R_\mu(A) &= (A - \lambda I)^{-1}(A - \mu I)^{-1} = \\ &= ((A - \mu I)(A - \lambda I))^{-1} = ((A - \lambda I)(A - \mu I))^{-1} = \\ &= (A - \mu I)^{-1}(A - \lambda I)^{-1} = R_\mu(A)R_\lambda(A). \end{aligned}$$

(b) Na mocy (a) i definicji rezolwenty mamy

$$\begin{aligned} R_\lambda(A) &= (A - \mu I)R_\lambda(A)R_\mu(A), \\ R_\mu(A) &= (A - \lambda I)R_\lambda(A)R_\mu(A). \end{aligned}$$

Stąd

$$R_\lambda(A) - R_\mu(A) = (\lambda - \mu)R_\lambda(A)R_\mu(A). \qquad \blacksquare$$

Definicja 5.3. *Promieniem spektralnym* operatora A jest liczba

$$r(A) = \sup\{|\lambda| \colon \lambda \in \sigma(A)\}.$$

Tak więc promień spektralny jest promieniem najmniejszego koła domkniętego o środku w zerze, które zawiera widmo.

Twierdzenie 5.9 (o widmie operatora). *Widmo dowolnego operatora* $A \in \mathcal{B}(X)$ *jest niepustym zwartym podzbiorem płaszczyzny zespolonej. Ponadto zachodzi nierówność* $r(A) \leqslant \|A\|$.

Dowód. Najpierw wykażemy, że widmo dowolnego operatora jest niepuste. Przypuśmy, że dla operatora A mamy $\sigma(A) = \emptyset$. Wówczas $A^{-1} \in \mathcal{B}(X)$. Niech f będzie funkcjonałem liniowym ograniczonym na $\mathcal{B}(X)$, dla którego $f(A^{-1}) \neq 0$. Taki funkcjonał istnieje na mocy wniosku 3.15. Utwórzmy funkcję $F(\lambda) = f(R_\lambda(A))$. Funkcja ta jest określona na całej płaszczyźnie zespolonej. Wykażemy, że jest ona funkcją całkowitą. Weźmy dowolny punkt $\lambda_0 \in \mathbb{C}$. Udowodnimy, że w tym punkcie funkcja F ma pochodną. Mamy

$$\frac{F(\lambda) - F(\lambda_0)}{\lambda - \lambda_0} = \frac{f\left(R_\lambda(A) - R_{\lambda_0}(A)\right)}{\lambda - \lambda_0} = \frac{f\left((\lambda - \lambda_0)R_\lambda(A)R_{\lambda_0}(A)\right)}{\lambda - \lambda_0} =$$
$$= f\left(R_\lambda(A)R_{\lambda_0}(A)\right) \to f\left(R^2_{\lambda_0}(A)\right), \quad \text{gdy} \ \ \lambda \to \lambda_0.$$

Zatem $F'(\lambda_0)$ istnieje i w konsekwencji F jest funkcją holomorficzną na całej płaszczyźnie, a więc funkcją całkowitą. Dla $|\lambda| > \|A\|$ na mocy twierdzenia 5.5 otrzymujemy

$$|F(\lambda)| \leqslant \|f\|\|R_\lambda(A)\| = \|f\|\|(A - \lambda I)^{-1}\| =$$
$$= \frac{\|f\|}{|\lambda|}\left\|\left(I - \frac{1}{\lambda}A\right)^{-1}\right\| \to 0, \quad \text{gdy} \ \ |\lambda| \to +\infty.$$

Zatem na mocy twierdzenia Liouville'a (patrz [27], tw. 16.26 z części III) funkcja $F = \text{const} = 0$. Stąd $0 = F(0) = f(A^{-1})$. Otrzymana sprzeczność dowodzi, że $\sigma(A) \neq \emptyset$ dla dowolnego $A \in \mathcal{B}(X)$.

Teraz udowodnimy zwartość widma. Jeżeli $|\lambda| > \|A\|$, to na mocy twierdzenia 5.5 operator

$$A - \lambda I = -\lambda\left(I - \frac{1}{\lambda}A\right)$$

jest odwracalny w $\mathcal{B}(X)$. Zatem $\sigma(A) \subset \{\lambda \in \mathbb{C}\colon \ |\lambda| \leqslant \|A\|\}$. Stąd wynika, że $\sigma(A)$ jest zbiorem ograniczonym oraz $r(A) \leqslant \|A\|$. Pozostaje udowodnić, że widmo jest domkniętym podzbiorem płaszczyzny. Z wniosku 5.6 wynika, że zbiór operatorów odwracalnych jest otwarty w $\mathcal{B}(X)$. Zbiór rezolwenty $\varrho(A)$ jest przeciwobrazem tego zbioru przez funkcję ciągłą $\lambda \mapsto A - \lambda I$, a zatem $\varrho(A)$ jest zbiorem otwartym (patrz [27], tw. 4.6 z części I). Ponieważ $\sigma(A) = \mathbb{C} \setminus \varrho(A)$, więc widmo jest domknięte jako dopełnienie zbioru otwartego (patrz [27], wn. 2.5 z części I). ■

Twierdzenie 5.10 (o odwzorowaniu spektralnym). *Jeżeli $A \in \mathcal{B}(X)$ i w jest wielomianem o współczynnikach zespolonych, to*

$$w(\sigma(A)) = \sigma(w(A)).$$

D o w ó d. Jeżeli w jest wielomianem stałym, to twierdzenie jest oczywiste. Załóżmy więc, że stopień wielomianu w jest dodatni. Niech $\mu \in \mathbb{C}$ i rozważmy wielomian $w(\lambda) - \mu$. Na mocy podstawowego twierdzenia algebry (por. [26], wn. 3.11) mamy

$$w(\lambda) - \mu = \beta(\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n),$$

gdzie β, λ_1, λ_2, ..., $\lambda_n \in \mathbb{C}$ i $\beta \neq 0$. Wówczas

$$w(A) - \mu I = \beta(A - \lambda_1 I)(A - \lambda_2 I) \dots (A - \lambda_n I).$$

Ponieważ operatory $A - \lambda_j I$ $(j = 1, 2, \dots, n)$ są parami przemienne, więc operator $w(A) - \mu I$ jest odwracalny wtedy i tylko wtedy, gdy każdy z operatorów $A - \lambda_j I$ jest odwracalny. To ma miejsce wtedy i tylko wtedy, gdy $\lambda_j \notin \sigma(A)$ dla $j = 1, 2, \dots, n$. A to z kolei jest równoważne temu, że dla dowolnego $\lambda \in \sigma(A)$ mamy $\beta(\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n) \neq 0$. Dalej to jest równoważne temu, że $w(\lambda) \neq \mu$ dla dowolnego $\lambda \in \sigma(A)$. ■

Twierdzenie 5.11 (wzór Gelfanda-Beurlinga). *Dla dowolnego operatora $A \in \mathcal{B}(X)$ zachodzi równość*

$$r(A) = \inf_n \|A^n\|^{1/n} = \lim_{n\to\infty} \|A^n\|^{1/n}. \tag{5.3}$$

Uwagi. 1. Zauważmy, że istnienie granicy jest częścią tezy.

2. Definicje widma operatora i promienia spektralnego są „czysto" algebraiczne, nie zależą one od normy operatorowej. Natomiast granica we wzorze (5.3) jest granicą ciągu n-tych pierwiastków norm n-tych potęg operatora, a więc zależy ona od normy. Zatem istota wzoru Gelfanda-Beurlinga polega na tym, że te zdefiniowane w różny sposób wielkości są sobie równe.

Dowód. Mamy (por. [27], def. 3.6 i tw. 3.14 z części I)

$$\inf_n \|A^n\|^{1/n} \leqslant \liminf_{n\to\infty} \|A^n\|^{1/n} \leqslant \limsup_{n\to\infty} \|A^n\|^{1/n}.$$

Jeżeli $\lambda \in \sigma(A)$, to z twierdzenia o odwzorowaniu spektralnym wynika, że $\lambda^n \in \sigma(A^n)$ $(n = 1, 2, \dots,)$. Stąd $|\lambda^n| \leqslant \|A^n\|$ i w konsekwencji $|\lambda| \leqslant \|A^n\|^{1/n}$. Z dowolności λ otrzymujemy

$$r(A) \leqslant \inf_n \|A^n\|^{1/n}.$$

Pozostaje wykazać, że

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leqslant r(A).$$

Z twierdzeń 5.5 i 5.2 wynika, że dla $|\lambda| > \|A\|$ mamy

$$R_\lambda(A) = (A - \lambda I)^{-1} = -\frac{1}{\lambda}\left(I - \frac{1}{\lambda}A\right)^{-1} = -\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n. \tag{5.4}$$

Weźmy funkcjonał liniowy ograniczony f na $\mathcal{B}(X)$ i utwórzmy funkcję $F(\lambda) = f(R_\lambda(A))$. Z dowodu twierdzenia 5.9 wynika, że funkcja F jest określona i holomorficzna na zbiorze $\varrho(A)$. W szczególności jest ona holomorficzna na zewnętrzu koła $\{\lambda \in \mathbb{C}\colon |\lambda| > r(A)\}$. Ponadto z (5.4) wynika, że dla $|\lambda| > \|A\|$ zachodzi wzór

$$F(\lambda) = -\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} f(A^n).$$

Jest to rozwinięcie funkcji F w szereg Laurenta w otoczeniu punktu w nieskończoności. Ponieważ funkcja F jest holomorficzna dla $|\lambda| > r(A)$, więc to rozwinięcie w szereg jest również prawdziwe dla takich λ. Zatem

$$\lim_{n\to\infty} f\left(\frac{1}{\lambda^{n+1}} A^n\right) = 0,$$

skąd również

$$\lim_{n\to\infty} f\left(\lambda^{-n} A^n\right) = 0 \qquad (|\lambda| > r(A))\,. \tag{5.5}$$

Ponieważ f było dowolne, możemy zdefiniować funkcjonały liniowe ograniczone φ_n na $\mathcal{B}(X)$ wzorem

$$\varphi_n(f) = f\left(\lambda^{-n} A^n\right) \qquad (n = 1, 2, \dots).$$

Z twierdzenia o wydobywaniu normy za pomocą funkcjonałów wynika, że $\|\varphi_n\| = \|\lambda^{-n} A^n\|$. Równość (5.5) oznacza, że

$$\lim_{n\to\infty} \varphi_n(f) = 0 \quad \text{dla każdego} \quad f \in \mathcal{B}(X)'.$$

Z twierdzenia Banacha-Steinhausa wynika, że dla dowolnego λ takiego, że $|\lambda| > r(A)$, istnieje stała C_λ, dla której $\|\varphi_n\| \leqslant C_\lambda$ $(n = 1, 2, \dots)$. Zatem

$$\left\|\lambda^{-n} A^n\right\| \leqslant C_\lambda \qquad (n = 1, 2, \dots).$$

Stąd otrzymujemy, że jeżeli $|\lambda| > r(A)$, to dla dowolnego naturalnego n

$$\|A^n\|^{1/n} \leqslant |\lambda|\, C_\lambda^{1/n}$$

i w konsekwencji

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leqslant |\lambda|.$$

Wobec dowolności λ, $|\lambda| > r(A)$, mamy

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leqslant r(A). \qquad \blacksquare$$

5.2. Widmo aproksymatywne punktowe

Użyteczne jest rozważanie następującego podzbioru widma operatora.

Definicja 5.4. Niech A będzie operatorem liniowym ograniczonym na przestrzeni Banacha X. *Widmem aproksymatywnym punktowym* operatora A nazywamy zbiór

$$\sigma_{\mathrm{ap}}(A) = \{\lambda \in \mathbb{C}\colon \text{ istnieje ciąg } x_n \in X \text{ taki, że } \|x_n\| = 1 \text{ dla } n = 1, 2, \ldots \text{ i } (A - \lambda I)x_n \to 0, \text{ gdy } n \to \infty\}.$$

Następujące twierdzenie charakteryzuje widmo aproksymatywne punktowe:

Twierdzenie 5.12. *Dla operatora $A \in \mathcal{B}(X)$ i liczby $\lambda \in \mathbb{C}$ następujące warunki są równoważne*:

(a) $\lambda \notin \sigma_{\mathrm{ap}}(A)$;

(b) *istnieje stała $\gamma > 0$ taka, że $\|(A - \lambda I)x\| \geqslant \gamma\|x\|$ dla dowolnego $x \in X$*;

(c) $\ker(A - \lambda I) = \{0\}$, $\overline{\mathrm{im}(A - \lambda I)} = \mathrm{im}(A - \lambda I)$.

D o w ó d. Bez zmniejszania ogólności dowodu możemy założyć, że $\lambda = 0$. Pokażemy, że warunek (b) wynika z warunku (a). Jeżeli (b) nie zachodzi, to dla dowolnego naturalnego n istnieje niezerowy wektor x_n taki, że $\|Ax_n\| < \frac{1}{n}\|x_n\|$. Przyjmując $y_n = \frac{1}{\|x_n\|}x_n$, mamy $\|y_n\| = 1$ i $\|Ay_n\| \to 0$. Zatem $0 \notin \sigma_{\mathrm{ap}}(A)$.

Teraz wykazujemy, że (c) jest konsekwencją (b). Z nierówności $\|Ax\| \geqslant \gamma\|x\|$ wynika, że $\ker A = \{0\}$. Jeżeli $Ax_n \to y$, to mamy

$$\|Ax_n - Ax_m\| \geqslant \gamma\|x_n - x_m\|,$$

co oznacza, że (x_n) jest ciągiem Cauchy'ego. Ponieważ przestrzeń X jest zupełna, więc $x_n \to x$ dla pewnego $x \in X$. Z ciągłości operatora A mamy $Ax_n \to Ax$ i w konsekwencji $y = Ax$. Pokazaliśmy więc, że $\overline{\mathrm{im}\,A} = \mathrm{im}\,A$.

W końcu dowodzimy, że (c) implikuje (a). Z warunku (c) wynika, że $Y = \mathrm{im}\,A$ jest przestrzenią Banacha (jako domknięty podzbiór przestrzeni zupełnej). Zatem $A\colon X \to Y$ jest ciągłą bijekcją. Z twierdzenia Banacha o operatorze odwrotnym $A^{-1}\colon Y \to X$ jest operatorem ograniczonym. Zatem, jeżeli $\|x\| = 1$, to $1 = \|x\| = \|A^{-1}Ax\| \leqslant \|A^{-1}\|\|Ax\|$, czyli $\|Ax\| \geqslant \|A^{-1}\|^{-1} > 0$ dla $\|x\| = 1$. To oznacza, że $0 \notin \sigma_{\mathrm{ap}}(A)$. ■

Twierdzenie 5.13. *Dla dowolnego operatora $A \in \mathcal{B}(X)$ prawdziwe są inkluzje:*

$$\sigma_{\mathrm{p}}(A) \cup \sigma_{\mathrm{c}}(A) \subset \sigma_{\mathrm{ap}} \subset \sigma(A).$$

D o w ó d. Inkluzje $\sigma_{\mathrm{p}}(A) \subset \sigma_{\mathrm{ap}}(A)$ i $\sigma_{\mathrm{ap}} \subset \sigma(A)$ są oczywiste. Pozostaje zatem wykazać, że $\sigma_{\mathrm{c}}(A) \subset \sigma_{\mathrm{ap}}(A)$. Jeżeli $\lambda \in \sigma_{\mathrm{c}}(A)$, to $(A - \lambda I)^{-1}$ jest operatorem nieograniczonym. Dla dowolnej liczby naturalnej n istnieje wektor $x_n \in X$ taki, że $\|x_n\| = 1$ i $\|(A - \lambda I)^{-1} x_n\| \geqslant n$. Przyjmując $y_n = \frac{1}{n}(A - \lambda I)^{-1} x_n$, otrzymujemy $\|y_n\| \geqslant 1$ dla wszystkich n oraz

$$(A - \lambda I) y_n = \frac{1}{n}(A - \lambda I)(A - \lambda I)^{-1} x_n = \frac{1}{n} x_n.$$

Zatem $(A - \lambda I) y_n \to 0$, gdy $n \to \infty$, czyli $\lambda \in \sigma_{\mathrm{ap}}(A)$. ■

Wniosek 5.14. *Jeżeli $\sigma_{\mathrm{r}}(A) = \emptyset$, to $\sigma_{\mathrm{ap}} = \sigma(A)$.*

Topologiczny brzeg podzbioru Z płaszczyzny zespolonej będziemy oznaczać symbolem ∂Z.

Twierdzenie 5.15. *Jeżeli $A \in \mathcal{B}(X)$, to $\partial\sigma(A) \subset \sigma_{\mathrm{ap}}(A)$.*

D o w ó d. Niech $\lambda \in \partial\sigma(A)$. Istnieje ciąg (λ_n) taki, że $\lambda_n \in \mathbb{C} \setminus \sigma(A)$ $(n = 1, 2, \ldots,)$ i $\lambda_n \to \lambda$, gdy $n \to \infty$. Wówczas

$$\|(A - \lambda_n I)^{-1}\| \to +\infty, \quad \text{gdy} \ \ n \to \infty. \tag{5.6}$$

Załóżmy, że to nie jest prawdą. Przechodząc w razie potrzeby do podciągu, znajdziemy stałą $M > 0$ taką, że $\|(A - \lambda_n I)^{-1}\| \leqslant M$ dla wszystkich n. Biorąc takie n, że $|\lambda - \lambda_n| < \frac{1}{M}$, otrzymujemy wówczas

$$\|(A - \lambda I) - (A - \lambda_n I)\| < \frac{1}{\|(A - \lambda_n I)^{-1}\|}.$$

Stąd na mocy twierdzenia 5.5 wynika, że operator

$$A - \lambda I = (A - \lambda I - A + \lambda_n I) + A - \lambda_n I$$

jest odwracalny. Otrzymaliśmy sprzeczność. Zatem relacja (5.6) jest prawdziwa.

Wybierzmy wektor $x_n \in X$, $\|x_n\| = 1$, dla którego

$$\alpha_n = \|(A - \lambda_n I)^{-1} x_n\| > \|(A - \lambda_n I)^{-1}\| - \frac{1}{n} \qquad (n = 1, 2, \ldots).$$

Wówczas $\alpha_n \to +\infty$. Jeżeli oznaczymy $y_n = \dfrac{1}{\alpha_n}(A-\lambda_n I)^{-1}x_n$, to $\|y_n\| = 1$ oraz

$$\begin{aligned}(A-\lambda I)y_n &= (A-\lambda_n I)y_n + (\lambda_n - \lambda)y_n = \\ &= \frac{1}{\alpha_n}x_n + (\lambda_n - \lambda)y_n.\end{aligned}$$

Stąd

$$\|(A-\lambda I)y_n\| \leqslant \frac{1}{\alpha_n} + |\lambda_n - \lambda|,$$

i ostatecznie $\|(A-\lambda I)y_n\| \to 0$, gdy $n \to \infty$, co oznacza, że $\lambda \in \sigma_{\mathrm{ap}}(A)$. ■

Uwaga. Z twierdzenia 5.15 wynika, że widmo aproksymatywne punktowe jest zawsze niepuste (w przeciwieństwie do widma punktowego, residualnego czy też ciągłego). Znaczenie widma aproksymatywnego punktowego będzie widoczne przy rozważaniu operatorów liniowych ograniczonych na przestrzeni Hilberta.

5.3. Widmo operatora na przestrzeni Hilberta

W tym podrozdziale rozważać będziemy operatory (liniowe ograniczone) na zespolonej przestrzeni Hilberta H. Dla podzbioru Z płaszczyzny zespolonej symbolem Z^* oznaczać będziemy zbiór $\{\overline{\lambda}\colon\ \lambda \in Z\}$.

Twierdzenie 5.16. *Jeżeli* $A \in \mathcal{B}(H)$, *to* $\sigma(A^*) = \sigma(A)^*$.

Dowód. Jeżeli $\lambda \notin \sigma(A)$, to istnieje operator $B \in \mathcal{B}(H)$ taki, że zachodzi równość $B(A-\lambda I) = (A-\lambda I)B = I$. Biorąc sprzężenia po obu stronach tej równości, otrzymujemy $(A^* - \overline{\lambda}I)B^* = B^*(A^* - \overline{\lambda}I)$, skąd $\overline{\lambda} \notin \sigma(A^*)$ i w konsekwencji $\lambda \notin \sigma(A^*)^*$. Pokazaliśmy więc, że

$$\sigma(A^*)^* \subset \sigma(A).$$

Wstawiając do tej inkluzji operator A^* w miejsce A, otrzymujemy

$$\sigma(A)^* \subset \sigma(A^*),$$

skąd wynika, że

$$\sigma(A) \subset \sigma(A^*)^*. \qquad ■$$

Przykład 5.2. Znajdziemy widmo i jego części dla operatora jednostronnego przesunięcia S działającego w przestrzeni ℓ_2. Wygodnie jest przy tym rozpatrywać równocześnie operator sprzężony S^*.

Mamy

$$S(\xi_1, \xi_2, \xi_3, \dots) = (0, \xi_1, \xi_2, \xi_3, \dots) \quad \text{dla} \quad x = (\xi_n)_{n=1}^{\infty} \in \ell_2.$$

Ponieważ $\|Sx\| = \|x\|$, więc $\|S\| = 1$. Stąd wynika, że $\sigma(S) \subset \overline{\mathbb{D}} = \{\lambda\colon\ |\lambda| \leqslant \leqslant 1\}$. Ponadto $\sigma(S^*) = \sigma(S)^* \subset \overline{\mathbb{D}}$.

Ponieważ operator S jest izometrią, więc jest on iniekcją, czyli $0 \notin \sigma_{\mathrm{p}}(S)$. Niech $x = (\xi_1, \xi_2, \dots) \in \ell_2$, $\lambda \neq 0$. Gdyby $Sx = \lambda x$, to

$$\begin{aligned} 0 &= \lambda\xi_1, \\ \xi_1 &= \lambda\xi_2, \\ \xi_2 &= \lambda\xi_3, \\ &\vdots \\ \xi_n &= \lambda\xi_{n+1}, \\ &\vdots \end{aligned}$$

Stąd $\xi_1 = \xi_2 = \ldots = \xi_n = \ldots = 0$. Zatem $\lambda \notin \sigma_{\mathrm{p}}(S)$. Wykazaliśmy więc, że

$$\sigma_{\mathrm{p}}(S) = \emptyset.$$

Niech $|\lambda| < 1$. Oznaczmy $x_\lambda = (1, \lambda, \lambda^2, \dots)$. Wówczas $\|x_\lambda\|^2 = \sum\limits_{n=0}^{\infty} |\lambda|^{2n} < < \infty$, czyli $x_\lambda \in \ell_2$. Mamy

$$S^* x_\lambda = (\lambda, \lambda^2, \lambda^3, \dots) = \lambda(1, \lambda, \lambda^2, \dots) = \lambda x_\lambda,$$

co oznacza, że $x_\lambda \in \ker(S^* - \lambda I)$, czyli $\mathbb{D} = \{\lambda\colon\ |\lambda| < 1\} \subset \sigma_{\mathrm{p}}(S^*)$.

Ponadto zauważmy, że $\dim\ker(S^* - \lambda I) = 1$. Jeżeli $x = (\xi_n) \in \ell_2$ spełnia równanie $S^* x = \lambda x$, to

$$(\lambda\xi_1, \lambda\xi_2, \lambda\xi_3, \dots) = (\xi_2, \xi_3, \xi_4, \dots).$$

Tak więc $\xi_{n+1} = \lambda\xi_n$ dla $n = 1, 2, \dots$ Stąd $\xi_{n+1} = \lambda^n \xi_1$, co oznacza, że

$$x = (\xi_1, \lambda\xi_1, \lambda^2\xi_1, \dots) = \xi_1 x_\lambda. \tag{5.7}$$

Ponieważ $x_\lambda \in \ker(S^* - \lambda I)$, więc wektor x_λ tworzy bazę tej podprzestrzeni.

Pokażemy, że jeżeli $|\lambda| = 1$, to $\lambda \notin \sigma_{\rm p}(S^*)$. Jeżeli $\lambda \in \sigma_{\rm p}(S^*)$, to wiemy, że $S^*x = \lambda x$ dla $x \neq 0$ i wówczas zachodzi równość (5.7). Ponieważ $x \neq 0$, więc $\xi_1 \neq 0$. Ponadto

$$\|x\|^2 = |\xi_1| \sum_{n=0}^{\infty} |\lambda|^{2n}.$$

Ten szereg jest zbieżny wtedy i tylko wtedy, gdy $|\lambda| < 1$. Zatem punkt λ, dla którego $|\lambda| = 1$, nie może być wartością własną operatora S^*.

Stąd

$$\sigma_{\rm p}(S^*) = \mathbb{D}.$$

Ponieważ $\mathbb{D} = \sigma_{\rm p}(S^*) \subset \sigma(S^*) \subset \overline{\mathbb{D}}$ oraz $\sigma(S^*)$ jest zbiorem domkniętym, więc $\sigma(S^*) = \overline{\mathbb{D}}$. Stąd wynika, że

$$\sigma(S) = \sigma(S^*) = \overline{\mathbb{D}}.$$

Ponieważ $\ker(S - \lambda I) = \operatorname{im}(S^* - \overline{\lambda} I)^{\perp}$ oraz $\ker(S - \lambda I) = \{0\}$, więc $\overline{\operatorname{im}(S^* - \overline{\lambda} I)} = \ell_2$ dla wszystkich λ. Stąd

$$\sigma_{\rm r}(S^*) = \emptyset, \quad \sigma_{\rm c}(S^*) = \mathbb{T} = \{\lambda\colon\ |\lambda| = 1\}, \quad \sigma_{\rm ap}(S^*) = \overline{\mathbb{D}}.$$

Jeżeli $|\lambda| = 1$, to $\ker(S^* - \overline{\lambda} I) = \{0\}$ i w konsekwencji $\overline{\operatorname{im}(S - \lambda I)} = \ell_2$, czyli $\mathbb{T} \cap \sigma_{\rm r}(S) = \emptyset$.

Natomiast dla $|\lambda| < 1$ mamy $\dim\ker(S^* - \overline{\lambda} I) = 1$. Stąd $\overline{\operatorname{im}(S - \lambda I)} \neq$ $\neq \ell_2$, czyli $\mathbb{D} \subset \sigma_{\rm r}(S)$. Ostatecznie więc

$$\sigma_{\rm r}(S) = \mathbb{D}, \quad \sigma_{\rm c}(S) = \mathbb{T}.$$

Jeżeli $|\lambda| \neq 1$ i $x \in \ell_2$, to

$$\|(S - \lambda I)x\| = \|Sx - \lambda x\| \geqslant \big|\|Sx\| - \|\lambda x\|\big| = \big|1 - |\lambda|\big|\|x\|.$$

Zatem $\lambda \notin \sigma_{\rm ap}(S)$. Ponieważ $\partial\sigma(S) = \mathbb{T} \subset \sigma_{\rm ap}(S)$, więc

$$\sigma_{\rm ap}(S) = \mathbb{T}.$$

Twierdzenie 5.17. *Jeżeli A jest operatorem normalnym, to $r(A) = \|A\|$.*

D o w ó d. Mamy

$$\begin{aligned}\|A\|^4 &= \|A^*A\|^2 = \|(A^*A)^*(A^*A)\| = \|(A^*A)(A^*A)\| = \\ &= \|(A^2)^*A^2\| = \|A^2\|^2.\end{aligned}$$

Zatem $\|A\|^2 = \|A^2\|$. Ponieważ A^n jest również operatorem normalnym, więc przez indukcję otrzymujemy równość $\|A\|^{2^k} = \|A^{2^k}\|$ dla $k = 1, 2, \ldots$ Ostatecznie więc

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \|A^{2^k}\|^{2^{-k}} = \|A\|. \qquad \blacksquare$$

Wniosek 5.18. *Jeżeli A jest operatorem normalnym, to istnieje $\lambda_0 \in \sigma(A)$ takie, że $|\lambda_0| = \|A\|$.*

Twierdzenie 5.19. *Jeżeli A jest operatorem normalnym i λ, μ są jego różnymi wartościami własnymi, to* $\ker(A - \lambda I) \perp \ker(A - \mu I)$.

D o w ó d. Niech $x \in \ker(A - \lambda I)$ i $y \in \ker(A - \mu I)$. Ponieważ $A^* y = \overline{\mu} y$, więc

$$\lambda\langle x, y\rangle = \langle Ax, y\rangle = \langle x, A^* y\rangle = \langle x, \overline{\mu} y\rangle = \mu\langle x, y\rangle.$$

Zatem $(\lambda - \mu)\langle x, y\rangle = 0$. Ponieważ $\lambda \neq \mu$, więc $\langle x, y\rangle = 0$, czyli $x \perp y$. $\blacksquare$

Twierdzenie 5.20. *Jeżeli A jest operatorem normalnym, to $\sigma(A) = \sigma_{\mathrm{ap}}(A)$.*

D o w ó d. Wystarczy pokazać, że jeżeli $\lambda \notin \sigma_{\mathrm{ap}}(A)$, to $\lambda \notin \sigma(A)$. Jeżeli $\lambda \notin \sigma_{\mathrm{ap}}(A)$, to istnieje $\gamma > 0$ takie, że

$$\|(A - \lambda I)x\| \geqslant \gamma\|x\|$$

dla wszystkich $x \in H$. Zatem operator $A - \lambda I$ ma domknięty obraz. Ponieważ $\|(A - \lambda I)^* x\| = \|(A - \lambda I)x\|$, więc również

$$\|(A - \lambda I)^* x\| \geqslant \gamma\|x\|$$

dla każdego $x \in H$. Stąd $\ker(A - \lambda I)^* = \{0\}$. Ponieważ $\ker(A - \lambda I)^* = \operatorname{im}(A - \lambda I)^{\perp}$, więc $\overline{\operatorname{im}(A - \lambda I)} = \operatorname{im}(A - \lambda I) = H$. To oznacza, że operator $A - \lambda I$ jest bijekcją, czyli $\lambda \notin \sigma(A)$. $\blacksquare$

Twierdzenie 5.21. *Jeżeli A jest operatorem samosprzężonym, to $\sigma(A) = \sigma_{\mathrm{ap}}(A) \subset \mathbb{R}$.*

D o w ó d. Niech λ będzie dowolną liczbą zespoloną, dla której $\operatorname{Im}\lambda \neq 0$, i niech $x \neq 0$ będzie dowolnym wektorem z przestrzeni H. Wówczas

$$\begin{aligned} 0 < |\lambda - \overline{\lambda}|\|x\|^2 &= |\langle (A - \lambda I)x, x\rangle - \langle (A - \overline{\lambda})x, x\rangle| = \\ &= \langle (A - \lambda I)x, x\rangle - \langle x, (A - \lambda I)x\rangle| \leqslant 2\|(A - \lambda I)x\|\|x\| \end{aligned}$$

i w konsekwencji

$$\|(A - \lambda I)x\| \geqslant \frac{1}{2}|\lambda - \overline{\lambda}|\|x\|$$

dla dowolnego wektora $x \in H$. Ponieważ $|\lambda - \overline{\lambda}| > 0$, więc $\lambda \notin \sigma_{\mathrm{ap}}(A) = \sigma(A)$. ■

Dla operatora samosprzężonego A definiujemy liczby

$$m = \inf_{\|x\|=1} \langle Ax, x\rangle, \qquad M = \sup_{\|x\|=1} \langle Ax, x\rangle.$$

Twierdzenie 5.22. *Widmo operatora samosprzężonego A jest zawarte w przedziale $[m, M]$. Punkty m i M należą do widma.*

Dowód. Z poprzedniego twierdzenia wiemy, że widmo operatora A leży na osi rzeczywistej. Najpierw pokażemy, że punkty leżące na tej osi poza odcinkiem $[m, M]$ nie należą do widma. Niech $\lambda > M$. Wówczas $\lambda = M + \gamma$, $\gamma > 0$. Dla dowolnego wektora $x \in H$ o normie jeden mamy

$$\langle(\lambda I - A)x, x\rangle = \lambda\langle x, x\rangle - \langle Ax, x\rangle \geqslant \lambda - M = \gamma.$$

Ponieważ ponadto

$$\langle(\lambda I - A)x, x\rangle \leqslant \|(A - \lambda I)x\|,$$

więc $\|(A - \lambda I)x\| \geqslant \gamma$. Stąd wynika, że dla dowolnego wektora $x \in H$ zachodzi nierówność

$$\|A - \lambda I)x\| \geqslant \gamma\|x\|,$$

która oznacza, że $\lambda \notin \sigma(A)$.

W przypadku, gdy $\lambda < m$, dowód przebiega analogicznie.

Teraz pokażemy, że $M \in \sigma(A)$. Zauważmy, że jeżeli zastąpimy operator A przez operator $A - \mu I$, to widmo przesunie się w lewo o μ, a liczby M i m zamienią się na $M - \mu$ i $m - \mu$. Zatem możemy założyć, że $0 \leqslant m \leqslant M$. Wówczas $\langle Ax, x\rangle \geqslant 0$ dla dowolnego wektora $x \in H$ i z twierdzenia 4.7 wynika, że $M = \|A\|$.

Z wniosku 5.18 wynika, że istnieje $\lambda_0 \in \sigma(A)$ takie, że $|\lambda_0| = \|A\| = M$. Ponieważ wszystkie elementy widma operatora A są liczbami nieujemnymi, więc ostatecznie $M = \lambda_0$.

Aby wykazać, że $m \in \sigma(A)$, wystarczy tak przesunąć widmo operatora A, aby $m \leqslant M \leqslant 0$, i zastosować poprzednie rozumowanie do operatora $-A$, bo $\lambda \in \sigma(A)$ wtedy i tylko wtedy, gdy $-\lambda \in \sigma(-A)$. ■

Twierdzenie 5.23. *Jeżeli A jest operatorem unitarnym, to $\sigma(A) = \sigma_{\mathrm{ap}}(A) \subset \mathbb{T}$.*

D o w ó d. Równość $\sigma(A) = \sigma_{\mathrm{ap}}(A)$ wynika z twierdzenia 5.20. Ponieważ operator unitarny jest izometrią, więc mamy $\|A\| = 1$. Stąd $\sigma(A) \subset \overline{\mathbb{D}}$. Jeżeli $|\lambda| < 1$, to $\|\lambda A^*\| < 1$ i z twierdzenia 5.5, wynika, że operator

$$\lambda I - A = -A(I - \lambda A^*)$$

jest odwracalny, czyli $\lambda \notin \sigma(A)$. ■

Ćwiczenia

1. Niech a będzie liczbą taką, że $|a| < 1$. Dowieść, że nieskończony układ równań

$$x_n - ax_{n+1} = y_n \qquad (n = 1, 2, \dots)$$

ma dla każdego ciągu $(y_n) \in \ell_\infty$ dokładnie jedno rozwiązanie $(x_n) \in \ell_\infty$. Znaleźć to rozwiązanie (posługując się szeregiem Neumanna).

2. Dowieść, że jeżeli g jest funkcją ciągłą na przedziale $[0, 1]$, to istnieje dokładnie jedna funkcja f, ciągła na przedziale $[0, 1]$ i taka, że

$$f(x) - \tfrac{1}{2} f(\tfrac{1}{2}x) = g(x) \quad \text{dla} \quad x \in [0, 1].$$

Wyznaczyć tę funkcję (posługując się szeregiem Neumanna).

3. Niech H będzie nieskończenie wymiarową ośrodkową przestrzenią Hilberta. Dowieść, że każdy niepusty zwarty podzbiór płaszczyzny zespolonej jest widmem pewnego operatora $A \in \mathcal{B}(H)$. (*Wskazówka.* Przyjąć $Ax = \sum\limits_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n$, gdzie (e_n) jest bazą ortonormalną przestrzeni H, a (λ_n) jest <u>odpowiednio</u> dobranym ciągiem).

4. Niech H będzie nieskończenie wymiarową przestrzenią Hilberta, a (λ_n) ciągiem liczbowym, $\lambda_n \neq 0$, $\lambda_n \to 0$. Podać przykład operatora $A \in \mathcal{B}(H)$, dla którego $\sigma_{\mathrm{p}}(A) = \{\lambda_n\colon\ n = 1, 2, \dots\}$ (por. ćwiczenie 3).

5. Wykazać, że jeżeli H jest ośrodkową przestrzenią Hilberta, to dla dowolnego operatora samosprzężonego $A \in \mathcal{B}(H)$ jego zbiór wartości własnych jest co najwyżej przeliczalny.

6. Znaleźć wartości własne i odpowiadające im wektory własne operatora rzutowania (ortogonalnego) P na (domkniętą) podprzestrzeń przestrzeni Hilberta H.

7.[*] Wykazać, że jeżeli $A \in \mathcal{B}(H)$ i $|\lambda| = \|A\|$, to $\lambda \in \sigma_{\mathrm{p}}(A)$ wtedy i tylko wtedy, gdy $\overline{\lambda} \in \sigma_{\mathrm{p}}(A^*)$.

8. Pokazać, że jeżeli $A \in \mathcal{B}(H)$ jest operatorem normalnym, to $\lambda \in \sigma_{\mathrm{p}}(A)$ wtedy i tylko wtedy, gdy $\overline{\lambda} \in \sigma_{\mathrm{p}}(A^*)$.

9. Wykazać, że dla operatora $A \in \mathcal{B}(H)$ prawdziwe są następujące własności:

[*] H będzie zawsze oznaczać zespoloną przestrzeń Hilberta.

(a) $\sigma_{\mathrm{r}}(A)^* \subset \sigma_{\mathrm{p}}(A^*) \subset \sigma_{\mathrm{r}}(A)^* \cup \sigma_{\mathrm{p}}(A)^*$;

(b) $\sigma_{\mathrm{ap}}(A^*) \cup \sigma_{\mathrm{ap}}(A)^* = \sigma(A^*)$;

(c) $\sigma_{\mathrm{p}}(w(A)) = w(\sigma_{\mathrm{p}}(A))$, gdzie w jest wielomianem stopnia $\geqslant 1$;

(d) $\sigma_{\mathrm{ap}}(w(A)) = w(\sigma_{\mathrm{ap}}(A))$, gdzie w jest wielomianem;

(e) $\sigma_{\mathrm{r}}(w(A)) \subset w(\sigma_{\mathrm{r}}(A))$, gdzie w jest wielomianem.

10. Operator $A \in \mathcal{B}(L_2[0,1])$ określony jest za pomocą wzoru $Ax(t) = a(t)x(t)$, gdzie $a \in L_\infty[0,1]$. Wykazać, że

(a) liczba λ nie należy do widma operatora A wtedy i tylko wtedy, gdy istnieje $\gamma > 0$ takie, że $|a(t) - \lambda| > \gamma$ prawie wszędzie na przedziale $[0,1]$;

(b) $\lambda \in \sigma_{\mathrm{p}}(A)$ wtedy i tylko wtedy, gdy równość $a(t) = \lambda$ zachodzi na zbiorze miary dodatniej;

(c) jeżeli $\lambda \in \sigma_{\mathrm{p}}(A)$, to przestrzeń $\ker(A - \lambda I)$ jest nieskończenie wymiarowa.

11. Operator $T \in \mathcal{B}(L_2(\mathbb{R}))$ określony jest wzorem $Tx(t) = x(-t)$.

(a) Wyznaczyć wartości własne operatora T i odpowiadające im podprzestrzenie własne;

(b) Czy T jest operatorem normalnym?

(c) Dla jakich liczb zespolonych λ równanie $Tx - \lambda x = f$ ma dokładnie jedno rozwiązanie $x \in L_2(\mathbb{R})$ przy każdej (ustalonej) funkcji $f \in L_2(\mathbb{R})$?

12. Dla operatora $A \in \mathcal{B}(\ell_2)$ określonego wzorem $A(\xi_1, \xi_2, \dots) = (\alpha_1\xi_1, \alpha_2\xi_2, \dots)$, gdzie $(\alpha_n) \in \ell_\infty$, znaleźć widmo i jego części.

13. Dla operatora $B \in \mathcal{B}(L_2[0,1])$ określonego wzorem $Bx(t) = tx(t)$ znaleźć widmo i jego części.

14. Znaleźć widmo operatora Volterry $Vx(s) = \int\limits_0^s x(t)\,dt$ działającego na przestrzeni $L_2[0,1]$.

15. Niech $0 < \alpha < 1$ i niech $M_\alpha = \{x \in L_2[0,1] \colon\ x(t) = 0 \text{ dla p.w. } t \in [0,\alpha]\}$. Wykazać, że M_α jest podprzestrzenią niezmienniczą dla operatora Volterry $Vx(s) = \int\limits_0^s x(t)\,dt$. Czy jest ona również podprzestrzenią redukującą?

16. Dowieść, że jeżeli przestrzeń Hilberta H ma rozkład na (ortogonalną) sumę prostą $H = M \oplus M^\perp$ i M jest podprzestrzenią redukującą dla operatora $A \in \mathcal{B}(H)$, to

$$\sigma(A) = \sigma(A\big|_M) \cup \sigma(A\big|_{M^\perp}).$$

17. Wykorzystać poprzednie ćwiczenie do wyznaczenia widma operatora całkowego A, $Ax(s) = \int\limits_{-1}^1 K(s,t)x(t)\,dt$, na przestrzeni $L_2[-1,1]$ wyznaczonego przez jądro K takie, że $K(s,t) = 1$ dla $st \geqslant 0$ i $K(s,t) = 0$ dla $st < 0$.

Rozdział 6

Operatory zwarte

Badanie własności operatorów liniowych na przestrzeniach nieskończenie wymiarowych jest dużo trudniejsze niż badanie tych operatorów na przestrzeniach skończenie wymiarowych. W całej ogólności praktycznie niewiele możemy o nich udowodnić. Sytuacja jest lepsza, jeżeli ograniczymy się do badania pewnych klas operatorów. Jedną z najważniejszych jest klasa operatorów zwartych. Potrzeba badania tych operatorów wynikła z trzech powodów. Po pierwsze — mają one wiele własności podobnych do operatorów skończenie wymiarowych, po drugie — potrafimy podać wyczerpujący ich opis i po trzecie — odgrywają one bardzo istotną rolę w teorii równań całkowych.

6.1. Ciągowa zwartość i całkowita ograniczoność

Definicja 6.1. Podzbiór E przestrzeni metrycznej X (z metryką d) jest *całkowicie ograniczony*, jeżeli dla dowolnego $\varepsilon > 0$ istnieje skończona liczba punktów x_1, x_2, ..., $x_n \in X$ takich, że

$$E \subset B_\varepsilon(x_1) \cup B_\varepsilon(x_2) \cup \ldots \cup B_\varepsilon(x_n). \tag{6.1}$$

Warunek (6.1) oznacza, że dla dowolnego $x \in E$ istnieje j ($1 \leqslant j \leqslant n$) takie, że $d(x, x_j) < \varepsilon$. Zbiór $\{x_1, x_2, \ldots, x_n\}$ nosi nazwę *ε-sieci* dla zbioru E. Zatem warunek (6.1) oznacza, że zbiór E ma skończoną ε-sieć.

U w a g i. 1. Zauważmy, że możemy zakładać, że punkty x_j należą do zbioru E (o ile zbiór E jest niepusty). Istotnie, niech punkty x_1, x_2, ..., $x_n \in X$ będą takie, że

$$E \subset B_{\varepsilon/2}(x_1) \cup B_{\varepsilon/2}(x_2) \cup \ldots \cup B_{\varepsilon/2}(x_n).$$

Możemy zakładać, że $E \cap B_{\varepsilon/2}(x_j) \neq \emptyset$ dla $j = 1, 2, \ldots, n$. Niech $x'_j \in E \cap B_{\varepsilon/2}(x_j)$. Wówczas, jeżeli $x \in E$, to istnieje j takie, że $d(x, x_j) < \frac{1}{2}\varepsilon$. W konsekwencji, $d(x, x'_j) \leqslant d(x, x_j) + d(x_j, x'_j) < \varepsilon$. Zatem

$$E \subset B_\varepsilon(x'_1) \cup B_\varepsilon(x'_2) \cup \ldots \cup B_\varepsilon(x'_n).$$

2. Każdy zbiór całkowicie ograniczony jest ograniczony. Łatwo można sprawdzić, że w przestrzeni $\mathbb{K}^n$ z metryką euklidesową jest również na odwrót. Natomiast w przypadku przestrzeni nieskończenie wymiarowych nie każdy zbiór ograniczony jest całkowicie ograniczony. Aby się o tym przekonać, wystarczy wziąć (domkniętą) kulę jednostkową B w przestrzeni ℓ_2. Wówczas wektory $e_n = (\delta_{kn})$ $n = 1, 2, \ldots$ są w tej kuli oraz $\|e_n - e_m\|_2 = \sqrt{2}$ dla $n \neq m$. Zatem dla $\varepsilon < \frac{1}{2}\sqrt{2}$ nie istnieje skończone pokrycie zbioru B przez kule o promieniu ε.

Definicja 6.2. Podzbiór K przestrzeni metrycznej X jest *ciągowo zwarty*, jeżeli dowolny ciąg (x_n) elementów tego zbioru zawiera podciąg zbieżny do pewnego elementu zbioru K.

Związki pomiędzy tymi pojęciami a zwartością podane są w następującym twierdzeniu:

Twierdzenie 6.1. *Niech K będzie podzbiorem przestrzeni metrycznej X. Następujące warunki są równoważne:*

(a) *zbiór K jest zwarty;*

(b) *zbior K jest ciągowo zwarty;*

(c) *zbiór K jest zupełny (tzn. dowolny ciąg Cauchy'ego elementów zbioru K jest zbieżny do pewnego elementu tego zbioru) i całkowicie ograniczony.*

Dowód. Zakładamy, że zbiór K jest niepusty, bo w przeciwnym wypadku twierdzenie jest oczywiste. Najpierw pokażemy, że warunek (b) jest konsekwencją (a). Niech $\{x_n \colon n = 1, 2, \ldots\} \subset K$. Przypuśćmy, że żaden podciąg ciągu (x_n) nie jest zbieżny w K. Zatem dla każdego $x \in K$ istnieje $r_x > 0$ takie, że otoczenie $B_{r_x}(x)$ zawiera co najwyżej skończenie wiele wyrazów ciągu (x_n). W przeciwnym wypadku istniałby punkt $x \in K$ taki, że dla dowolnego $\varepsilon > 0$ istniałoby nieskończenie wiele wyrazów ciągu (x_n) takich, że $d(x, x_n) < \varepsilon$. Biorąc $\varepsilon = \dfrac{1}{m}$, skonstruowalibyśmy taki podciąg (x_{n_m}) ciągu (x_n), że $d(x, x_{n_m}) < \dfrac{1}{m}$. A to oznaczałoby, że $x_{n_m} \to x$, wbrew naszemu założeniu. Zatem dowolny punkt $x \in K$ ma otoczenie $B_{r_x}(x)$, które zawiera co najwyżej skończenie wiele wyrazów ciągu (x_n). Ponieważ

$$K \subset \bigcup_{x \in K} B_{r_x}(x),$$

i K jest zbiorem zwartym, więc istnieje skończenie wiele punktów x_1, x_2, $\ldots$, $x_s \in K$ takich, że

$$K \subset \bigcup_{j=1}^{s} B_{r_{x_j}}(x_j).$$

Ponieważ każda z kul $B_{r_{x_j}}(x_j)$ zawiera co najwyżej skończenie wiele wyrazów ciągu (x_n), więc zbiór K zawiera skończenie wiele wyrazów tego ciągu. To oznacza, że zbiór wyrazów ciągu (x_n) jest skończony, a zatem ciąg ten zawiera podciąg (stały) zbieżny. Otrzymaliśmy sprzeczność, zatem zbiór K jest ciągowo zwarty.

Teraz pokażemy, że warunek (b) implikuje (c). Najpierw zauważmy, że K jest zupełne. Wynika to z faktu, że jeżeli ciąg Cauchy'ego (x_n) zawiera podciąg zbieżny do punktu x, to również ciąg (x_n) jest zbieżny do x (por. [27], dowód tw. 3.12 z części I). Załóżmy, że K nie jest całkowicie ograniczone. Istnieje $\varepsilon > 0$ takie, że zbiór nie może być pokryty przez skończenie wiele kul $B_\varepsilon(x)$, $x \in K$. Konstruujemy indukcyjnie ciąg (x_n) punktów zbioru K w następujący sposób:

Punkt $x_1 \in K$ wybieramy w sposób dowolny. Jeżeli punkty x_1, x_2, $\ldots$, $x_{n-1} \in K$ zostały wybrane tak, aby $d(x_i, x_j) \geqslant \varepsilon$ dla $i \neq j$, $1 \leqslant i \leqslant n-1$, $1 \leqslant j \leqslant n-1$, to K nie może być pokryte przez rodzinę $\{B_\varepsilon(x_j)\colon j = 1, 2, \ldots, n-1\}$. Istnieje więc $x_n \in K \setminus \bigcup_{j=1}^{n-1} B_\varepsilon(x_j)$. Wówczas $d(x_j, x_n) \geqslant \varepsilon$ dla $j = 1, 2, \ldots, n-1$. Ciąg (x_n) nie może zawierać podciągu, który jest zbieżny. Wynika to stąd, że taki podciąg (x_{n_m}) byłby ciągiem Cauchy'ego, więc w szczególności mielibyśmy $d(x_{n_i}, x_{n_j}) < \varepsilon$ dla odpowiednio dużych i oraz j. To jest niemożliwe na mocy konstrukcji ciągu (x_n). Otrzymana sprzeczność dowodzi, że zbiór K jest całkowicie ograniczony.

W końcu udowodnimy, że z (c) wynika (a). Przypuśćmy, że K nie jest zbiorem zwartym. Istnieje więc pokrycie otwarte $\{G_\alpha\}$ zbioru K, które nie zawiera skończonego podpokrycia. Definiujemy indukcyjnie ciąg kul (B_n) w następujący sposób:

Ponieważ K jest całkowicie ograniczone, więc jest pokryte przez skończoną rodzinę kul $B_{1/2}(x_j)$, $x_j \in K$, $j = 1, 2, \ldots, s$. Przynajmniej jeden ze zbiorów $B_{1/2}(x_j) \cap K$ nie może być pokryty przez skończoną liczbę zbiorów z rodziny $\{G_\alpha\}$. Oznaczmy kulę o tej własności przez B_1, a jej środek przez ξ_1. Niech $B_{n-1} = B_{1/2^{n-1}}(\xi_{n-1})$, gdzie $\xi_{n-1} \in K$ oraz $B_{n-1} \cap K$ nie może być pokryte przez żadną skończoną liczbę zbiorów z rodziny $\{G_\alpha\}$. Ponieważ K jest całkowicie ograniczone, więc istnieje skończone pokrycie przez rodzinę kul $B_{1/2^n}(y_i)$, $y_i \in K$, $i = 1, 2, \ldots, m$. Wśród kul, dla których $B_{1/2^n}(y_i) \cap B_{n-1} \neq \emptyset$, istnieje co najmniej jedna o tej własności, że zbiór

$B_{1/2^n}(y_i) \cap K$ nie może być pokryty przez żadną skończoną podrodzinę rodziny $\{G_\alpha\}$, gdyż w przeciwnym wypadku zbiór $B_{n-1} \cap K$ byłby pokryty przez skończenie wiele zbiorów z rodziny $\{G_\alpha\}$. Wybieramy więc B_n jako jedną z tych kul $B_{1/2^n}(y_i)$, a jej środek oznaczamy symbolem ξ_n. Ponieważ $B_{n-1} \cap B_n \neq \emptyset$, więc mamy

$$d(\xi_{n-1}, \xi_n) < \frac{1}{2^{n-1}} + \frac{1}{2^n} < \frac{1}{2^{n-2}}.$$

Stąd, jeżeli $n > m \geqslant n_0$, to

$$d(\xi_m, \xi_n) \leqslant d(\xi_m, \xi_{m+1}) + \ldots + d(\xi_{n-1}, \xi_n) < \frac{1}{2^{m-1}} + \ldots + \frac{1}{2^{n-2}} < \frac{1}{2^{n_0-2}}.$$

To oznacza, że (ξ_n) jest ciągiem Cauchy'ego. Na mocy założenia $\xi_n \to x$ dla pewnego $x \in K$. Ponieważ rodzina $\{G_\alpha\}$ pokrywa zbiór K, więc istnieje α_0 takie, że $x \in G_{\alpha_0}$. Zbiór G_{α_0} jest otwarty, zatem istnieje $\varepsilon > 0$ takie, że $B_\varepsilon(x) \subset G_{\alpha_0}$. Ze zbieżności ciągu (ξ_n) otrzymujemy wskaźnik m, dla którego $d(x, \xi_m) < \dfrac{\varepsilon}{2}$ oraz $\dfrac{1}{2^m} < \dfrac{\varepsilon}{2}$. Wówczas, jeżeli $d(y, \xi_m) < \dfrac{1}{2^m}$, to $d(y, x) \leqslant d(y, \xi_m) + d(\xi_m, x) < \varepsilon$, czyli

$$B_m = B_{1/2^m}(\xi_m) \subset B_\varepsilon(x) \subset G_{\alpha_0}.$$

Otrzymaliśmy sprzeczność, bo z konstrukcji ciągu (B_n) wynika, że zbiór B_m nie może być pokryty przez skończoną liczbę zbiorów z rodziny $\{G_\alpha\}$. Zatem zbiór K jest zwarty. ■

Definicja 6.3. Podzbiór K przestrzeni metrycznej X jest *warunkowo zwarty*, jeżeli jego domknięcie $\overline{K}$ jest zbiorem zwartym.

Wniosek 6.2. *Zbiór K jest warunkowo zwarty wtedy i tylko wtedy, gdy z dowolnego ciągu punktów tego zbioru można wybrać podciąg zbieżny.*

Twierdzenie 6.3 (Hausdorffa). *Niech X będzie przestrzenią metryczną zupełną. Podzbiór K przestrzeni X jest warunkowo zwarty wtedy i tylko wtedy, gdy jest całkowicie ograniczony.*

D o w ó d. Jeżeli zbiór $\overline{K}$ jest zwarty, to na mocy twierdzenia 6.1 jest on całkowicie ograniczony. W szczególności K jest całkowicie ograniczone.

Na odwrót, jeżeli zbiór K jest całkowicie ograniczony, to dla dowolnego $\varepsilon > 0$ istnieją x_1, x_2, ..., $x_n \in K$ takie, że

$$K \subset B_{\varepsilon/2}(x_1) \cup B_{\varepsilon/2}(x_2) \cup \ldots \cup B_{\varepsilon/2}(x_n). \tag{6.2}$$

Jeżeli x jest dowolnym punktem z domknięcia zbioru K, to istnieje $y \in \in K$ takie, że $d(x, y) < \frac{1}{2}\varepsilon$. Ponadto z (6.2) wynika, że istnieje j takie, że $d(y, x_j) < \frac{1}{2}\varepsilon$. Z nierówności trójkąta otrzymujemy $d(x, x_j) < \varepsilon$. Stąd

$$\overline{K} \subset B_\varepsilon(x_1) \cup B_\varepsilon(x_2) \cup \ldots \cup B_\varepsilon(x_n).$$

Pokazaliśmy więc, że domknięcie zbioru całkowicie ograniczonego jest zbiorem całkowicie ograniczonym. Ponieważ domknięty podzbiór przestrzeni zupełnej jest zupełny, więc $\overline{K}$ jest zupełne. Na mocy twierdzenia 6.1 zbiór $\overline{K}$ jest zwarty. ■

Na zakończenie tego podrozdziału udowodnimy klasyczne kryterium warunkowej zwartości w przestrzeni funkcji ciągłych na przestrzeni zwartej. W dalszym ciągu słowo *funkcja* będzie oznaczało funkcję o wartościach zespolonych (w szczególności również rzeczywistych).

Definicja 6.4. Niech E będzie dowolnym zbiorem i niech Φ będzie rodziną funkcji określonych na zbiorze E. Funkcje z rodziny Φ są *wspólnie ograniczone*, jeżeli istnieje stała $M > 0$ taka, że $|f(x)| \leqslant M$ dla dowolnej funkcji $f \in \Phi$ i dowolnego $x \in E$.

Definicja 6.5. Niech X będzie przestrzenią metryczną (z metryką d) i niech Φ będzie rodziną funkcji określonych na przestrzeni X. Funkcje z rodziny Φ są *jednakowo ciągłe*, jeżeli dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że $|f(x_1) - f(x_2)| < \varepsilon$ dla dowolnej funkcji $f \in \Phi$ i dla dowolnych punktów x_1, $x_2 \in X$ takich, że $d(x_1, x_2) < \delta$.

Zauważmy, że wówczas każda z funkcji f jest jednostajnie ciągła na X. Istotą tej definicji jest to, że liczba δ jest taka sama dla każdej funkcji f należącej do rodziny Φ.

Twierdzenie 6.4 (Ascoliego-Arzeli). *Niech K będzie przestrzenią metryczną zwartą. Na to, aby rodzina Φ funkcji ciągłych na zbiorze K była warunkowo zwartym podzbiorem przestrzeni $\mathcal{C}(K)$, potrzeba i wystarcza, aby funkcje z tej rodziny były wspólnie ograniczone i jednakowo ciągłe.*

D o w ó d. Załóżmy, że rodzina funkcji Φ jest warunkowo zwartym podzbiorem przestrzeni $\mathcal{C}(K)$. Wówczas na mocy twierdzenia Hausdorffa zbiór Φ jest całkowicie ograniczony, a więc w szczególności ograniczony. Zatem istnieje $M > 0$ takie, że $\|f\|_\infty = \sup\{|f(x)|\colon\ x \in K\} \leqslant M$ dla dowolnego $f \in \Phi$. To oznacza, że funkcje z rodziny Φ są wspólnie ograniczone. Aby pokazać, że są one jednakowo ciągłe, weźmy dowolne $\varepsilon > 0$ i wybierzmy skończoną $\frac{1}{3}\varepsilon$-sieć $\{f_1, f_2, \ldots, f_n\}$ dla Φ. Każda z funkcji f_j $(j = 1, 2, \ldots, n)$ jest jednostajnie ciągła ([27], tw. 4.13 z części I). Zatem dla każdego j istnieje liczba $\delta_j > 0$

taka, że $|f_j(x_1) - f_j(x_2)| < \frac{1}{3}\varepsilon$ dla dowolnych x_1, $x_2 \in K$ spełniających nierówność $d(x_1, x_2) < \delta_j$. Niech $\delta = \min\{\delta_j\colon\ j = 1, 2, \ldots, n\}$. Dla dowolnej funkcji $f \in \Phi$ istnieje j takie, że $\|f - f_j\|_\infty < \frac{1}{3}\varepsilon$. Wówczas dla x_1, $x_2 \in K$ takich, że $d(x_1, x_2) < \delta$, mamy

$$\begin{aligned} |f(x_1) - f(x_2)| &\leqslant |f(x_1) - f_j(x_1)| + |f_j(x_1) - f_j(x_2)| + |f_j(x_2) - f(x_2)| \leqslant \\ &\leqslant 2\|f - f_j\|_\infty + |f_j(x_1) - f_j(x_2)| < \varepsilon. \end{aligned}$$

Zatem rodzina Φ jest jednakowo ciągła.

Na odwrót, załóżmy, że Φ jest rodziną funkcji ciągłych na przestrzeni K wspólnie ograniczonych i jednakowo ciągłych. Aby udowodnić warunkową zwartość Φ, wystarczy na mocy twierdzenia Hausdorffa wykazać, że dla dowolnego $\varepsilon > 0$ istnieje dla Φ skończona ε-sieć. Niech więc będzie dane $\varepsilon > 0$. Na mocy jednakowej ciągłości Φ dla dowolnego $x \in K$ istnieje $\delta_x > 0$ takie, że

$$|f(y) - f(x)| < \frac{1}{4}\varepsilon \tag{6.3}$$

dla $y \in K$ takich, że $d(y, x) < \delta_x$, oraz dla wszystkich $f \in \Phi$. Rodzina kul $B_{\delta_x}(x)$, $x \in K$ stanowi otwarte pokrycie zbioru K. Ze zwartości K otrzymujemy

$$K \subset B_{\delta_{x_1}}(x_1) \cup B_{\delta_{x_2}}(x_2) \cup \ldots \cup B_{\delta_{x_n}}(x_n). \tag{6.4}$$

Niech $\Phi(x_j) = \{f(x_j)\colon\ f \in \Phi\}$ $(j = 1, 2, \ldots, n)$. Ponieważ funkcje z rodziny Φ są wspólnie ograniczone, więc $\Phi(x_j)$ jest ograniczonym podzbiorem zbioru $\mathbb{C}$ (lub $\mathbb{R}$, jeżeli rozważamy funkcje o wartościach rzeczywistych). Zatem $\Phi(x_j)$ jest zbiorem całkowicie ograniczonym. Niech ξ_{j1}, ξ_{j2}, ..., ξ_{jk_j} będzie $\frac{1}{4}\varepsilon$-siecią dla zbioru $\Phi(x_j)$. Dla dowolnego układu $(m_1, m_2, \ldots, m_n)$ wskaźników, gdzie $1 \leqslant m_j \leqslant k_j$ oraz $j = 1, 2, \ldots, n$ oznaczmy

$$\Phi_{m_1 m_2 \ldots m_n} = \left\{ f \in \Phi\colon\ |f(x_j) - \xi_{jm_j}| < \frac{1}{4}\varepsilon \ \text{ dla } \ j = 1, 2, \ldots, n \right\}. \tag{6.5}$$

Jeżeli zbiór $\Phi_{m_1 m_2 \ldots m_n}$ jest niepusty, to wybieramy funkcję $f_{m_1 m_2 \ldots m_n}$ z tego zbioru. W przeciwnym wypadku bierzemy za $f_{m_1 m_2 \ldots m_n}$ dowolną funkcję z rodziny Φ. Wykażemy, że zbiór

$$\{f_{m_1 m_2 \ldots m_n}\colon\ m_j = 1, 2, \ldots, k_j;\ j = 1, 2, \ldots, n\}$$

jest ε-siecią dla Φ.

Niech $f \in \Phi$. Ponieważ zbiór $\{\xi_{j1}, \xi_{j2}, \ldots, \xi_{jk_j}\}$ jest $\frac{1}{4}\varepsilon$-siecią dla $\Phi(x_j)$ $(j = 1, 2, \ldots, n)$, więc istnieje taki wskaźnik m_j, $1 \leqslant m_j \leqslant k_j$, że

$$|f(x_j) - \xi_{jm_j}| < \frac{1}{4}\varepsilon. \tag{6.6}$$

Dla zwięzłości oznaczeń przyjmijmy $f_* = f_{m_1 m_2 \ldots m_n}$. Niech $x \in K$ będzie dowolne. Na mocy (6.4) istnieje taki wskaźnik i, $1 \leqslant i \leqslant n$, że $d(x, x_i) < \delta_{x_i}$. Z nierówności (6.3), (6.5) i (6.6) mamy

$$\begin{aligned} |f(x) - f_*(x)| \leqslant |f(x) - f(x_i)| + |f(x_i) - \xi_{im_i}| + \\ + |\xi_{im_i} - f_*(x_i)| + |f_*(x_i) - f_*(x)| < \varepsilon. \end{aligned}$$

Zatem $\|f - f_*\|_\infty = \sup\{|f(x) - f_*(x)|\colon\ x \in K\} < \varepsilon$, co oznacza, że zbiór Φ jest całkowicie ograniczony. ■

6.2. Zwartość w przestrzeniach skończenie wymiarowych

Udowodnimy, że charakteryzacja zbiorów zwartych wyrażona we wniosku 1.20 jest charakterystyczna dla przestrzeni skończenie wymiarowych.

Lemat 6.5 (Riesza). *Niech X_0 będzie domkniętą podprzestrzenią liniową przestrzeni unormowanej X różną od X. Wówczas dla dowolnego $0 < \varepsilon < 1$ istnieje $y \in X$ takie, że $\|y\| = 1$ i $\|y - x\| \geqslant 1 - \varepsilon$ dla wszystkich $x \in X_0$.*

D o w ó d. Niech $y_0 \in X \setminus X_0$ i niech $\varrho = \inf\{\|y_0 - x\|\colon\ x \in X_0\}$. Wówczas $\varrho > 0$, bo X_0 jest domknięte. Wybierzmy takie $\eta > 0$, aby $\dfrac{\eta}{\varrho + \eta} \leqslant \varepsilon$. Następnie bierzemy takie $x_0 \in X_0$, że $\varrho \leqslant \|y_0 - x_0\| \leqslant \varrho + \eta$. Wektor $y = \dfrac{1}{\|y_0 - x_0\|}(y_0 - x_0)$ ma żądane własności. Mamy bowiem $\|y\| = 1$ oraz dla dowolnego $x \in X_0$

$$\|y - x\| = \frac{1}{\|y_0 - x_0\|}\|y_0 - (x_0 + \|y_0 - x_0\|x)\|.$$

Ponieważ $x_0 + \|y_0 - x_0\|x \in X_0$, więc stąd wynika, że

$$\|y - x\| \geqslant \frac{1}{\|y_0 - x_0\|}\varrho \geqslant \frac{\varrho}{\varrho + \eta} = 1 - \frac{\eta}{\varrho + \eta} \geqslant 1 - \varepsilon. \qquad ■$$

Twierdzenie 6.6. *Domknięta kula jednostkowa w przestrzeni unormowanej X jest zwarta wtedy i tylko wtedy, gdy przestrzeń X jest skończenie wymiarowa.*

D o w ó d. Jeżeli przestrzeń X jest skończenie wymiarowa, to z wniosku 1.20 wynika, że domknięta kula jednostkowa $B = \overline{B_1(0)} = \{x \in X\colon \|x\| \leqslant 1\}$ jest zbiorem zwartym.

Na odwrót, przypuśćmy, że kula B jest zwarta. Wówczas jest ona całkowicie ograniczona, więc ma $\frac{1}{2}$-sieć $\{x_1, x_2, \ldots, x_n\}$. Niech X_0 będzie podprzestrzenią liniową przestrzeni X rozpiętą na wektorach x_1, x_2, $\ldots$, x_n. Jeżeli $X_0 \neq X$, to — ponieważ przestrzeń X_0 jest skończenie wymiarowa — jest ona na mocy wniosku 1.19 domknięta. Z lematu Riesza wynika, że istnieje wektor $y \in B$ taki, że $\|y - x\| \geqslant \frac{1}{2}$ dla dowolnego $x \in X_0$. W szczególności $\|y - x_j\| \geqslant \frac{1}{2}$ dla $j = 1, 2, \ldots, n$. Otrzymaliśmy sprzeczność, bo $\{x_1, x_2, \ldots, x_n\}$ jest $\frac{1}{2}$-siecią dla zbioru B. Zatem $X_0 = X$, czyli przestrzeń X jest skończenie wymiarowa. ■

Wniosek 6.7. *Jeżeli dowolny domknięty i ograniczony podzbiór przestrzeni unormowanej X jest zwarty, to X jest przestrzenią skończenie wymiarową.*

6.3. Operatory zwarte

Definicja 6.6. Operator liniowy A odwzorowujący przestrzeń unormowaną X w przestrzeń unormowaną Y jest *operatorem zwartym*, jeżeli przeprowadza on dowolny zbiór ograniczony w zbiór warunkowo zwarty.

Twierdzenie 6.8. *Niech X i Y będą przestrzeniami unormowanymi oraz niech $A\colon X \to Y$ będzie operatorem liniowym. Następujące warunki są równoważne:*

(a) *operator A jest zwarty;*

(b) *obraz kuli jednostkowej poprzez operator A jest warunkowo zwarty;*

(c) *dla dowolnego ciągu ograniczonego (x_n) w przestrzeni X istnieje podciąg (x_{n_k}) taki, że ciąg (Ax_{n_k}) jest zbieżny.*

D o w ó d. Jest oczywiste, że warunek (b) jest konsekwencją (a). Teraz pokażemy, że z (b) wynika (c). Niech (x_n) będzie ciągiem ograniczonym w przestrzeni X. Zatem $\|x_n\| < M$ dla wszystkich n i pewnej stałej $M > 0$. Jeżeli weźmiemy ciąg $(M^{-1}x_n)$, to jego wyrazy znajdują się w kuli jednostkowej $B_1(0)$ przestrzeni X. Ponieważ zbiór $A(B_1(0))$ jest warunkowo zwarty, więc na mocy wniosku 6.2 z ciągu $(A(M^{-1}x_n))$ można wybrać podciąg zbieżny $(M^{-1}Ax_{n_k})$. Wówczas również i ciąg (Ax_{n_k}) jest zbieżny. W końcu udowodnimy, że (c) implikuje (a). Niech E będzie ograniczonym podzbiorem przestrzeni X i niech (y_n) będzie dowolnym ciągiem elementów zbioru $A(E)$.

Wówczas $y_n = Ax_n$ dla pewnego $x_n \in E$ $(n = 1, 2, \ldots)$. Ponieważ zbiór E jest ograniczony, więc i ciąg (x_n) jest ograniczony. Zatem istnieje podciąg (x_{n_k}) taki, że ciąg (Ax_{n_k}) jest zbieżny. Na mocy wniosku 6.2 zbiór $A(E)$ jest warunkowo zwarty. ■

Uwagi. 1. Ponieważ wszystkie operatory zwarte, jakie będziemy tu rozważać, są liniowe, nazwy *operator zwarty* używamy na określenie operatora liniowego zwartego.

2. W przestrzeni skończenie wymiarowej każdy operator liniowy ograniczony jest zwarty, ponieważ przeprowadza on dowolny zbiór ograniczony w zbiór ograniczony, a w przestrzeni skończenie wymiarowej każdy zbiór ograniczony jest warunkowo zwarty (wynika to z wniosku 6.2 i twierdzenia Bolzano-Weierstrassa).

3. W przestrzeni nieskończenie wymiarowej już tak nie jest. Przykładem operatora, który nie jest zwarty, jest operator identycznościowy I, bo z twierdzenia 6.6 wynika, że domknięta kula jednostkowa jest zbiorem ograniczonym i domkniętym, który nie jest zwarty.

Twierdzenie 6.9. *Każdy operator zwarty jest ograniczony.*

Dowód. Przypuśćmy, że operator A jest zwarty, ale nie jest ograniczony. Istnieje wtedy taki ciąg wektorów (x_n) w przestrzeni X, że $\|x_n\| \leqslant 1$ dla $n = 1, 2, \ldots$ i $\|Ax_n\| \to +\infty$. Ciąg (Ax_n) nie zawiera podciągów zbieżnych, mimo że ciąg (x_n) jest ograniczony. Przeczy to zwartości operatora A. ■

Definicja 6.7. Mówimy, że operator liniowy $A\colon X \to Y$ jest *skończenie wymiarowy*, jeżeli jego obraz $\operatorname{im} A$ jest podprzestrzenią skończenie wymiarową przestrzeni Y.

Twierdzenie 6.10. *Jeżeli X i Y są przestrzeniami unormowanymi, to każdy operator skończenie wymiarowy $A \in \mathcal{B}(X, Y)$ jest zwarty.*

Dowód. Niech E będzie ograniczonym podzbiorem przestrzeni X. Z nierówności $\|Ax\| \leqslant \|A\|\|x\|$ wynika, że operator A przeprowadza zbiór E na zbiór ograniczony $A(E) \subset Y$. $A(E)$ jako podzbiór ograniczony przestrzeni skończenie wymiarowej $\operatorname{im} A$ jest warunkowo zwarty w $\operatorname{im} A$, a więc tym bardziej i w przestrzeni Y. ■

Twierdzenie 6.11. *Jeżeli X i Y są przestrzeniami unormowanymi i operatory A_1, A_2 są zwarte oraz α i β są dowolnymi skalarami, to operator $\alpha A_1 + \beta A_2$ jest również zwarty.*

Łatwy dowód tego twierdzenia pozostawiamy jako ćwiczenie dla Czytelnika.

Z twierdzenia 6.11 wynika, że operatory zwarte tworzą podprzestrzeń liniową przestrzeni $\mathcal{B}(X,Y)$. Udowodnimy, że przy założeniu zupełności przestrzeni Y jest to podprzestrzeń domknięta.

Twierdzenie 6.12. *Niech X będzie przestrzenią unormowaną, a Y — przestrzenią Banacha. Jeżeli $A_n \in \mathcal{B}(X,Y)$ $(n=1,2,\dots)$ są operatorami zwartymi i $A_n \to A$, to operator A też jest zwarty.*

Dowód. Niech $\varepsilon > 0$ będzie dowolne. Aby udowodnić, że operator A jest zwarty, wystarczy, na mocy twierdzenia Hausdorffa, wykazać, że obraz kuli jednostkowej poprzez operator A ma skończoną ε-sieć. Niech n będzie tak duże, że $\|A_n - A\| < \frac{1}{2}\varepsilon$. Ponieważ A_n jest operatorem zwartym, więc istnieje skończona $\frac{1}{2}\varepsilon$-sieć dla zbioru $A_n(B_1(0))$. Załóżmy, że jest to $\{y_1,\dots,y_k\}$. Wówczas zbiór ten jest ε-siecią dla zbioru $A(B_1(0))$, gdyż dla dowolnego $x \in B_1(0)$ istnieje wskaźnik j taki, że $\|A_n x - y_j\| < \frac{1}{2}\varepsilon$, i wówczas

$$\begin{aligned}\|Ax - y_j\| &\leqslant \|Ax - A_n x\| + \|A_n x - y_j\| < \\ &< \|A - A_n\| + \frac{1}{2}\varepsilon < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon. \qquad \blacksquare\end{aligned}$$

W przypadku, gdy $X = Y$, możemy udowodnić, że operatory zwarte stanowią ideał (dwustronny) algebry $\mathcal{B}(X)$. Mianowicie prawdziwe jest następujące twierdzenie:

Twierdzenie 6.13. *Jeżeli $A \in \mathcal{B}(X)$ jest operatorem zwartym, a B operatorem liniowym ograniczonym na X, to operatory AB i BA są zwarte.*

Dowód. Jeżeli zbiór $E \subset X$ jest ograniczony, to również ograniczony jest zbiór $B(E)$. Stąd $AB(E) = A(B(E))$ jest warunkowo zwarte, co oznacza, że operator AB jest zwarty.

Jeżeli E jest zbiorem ograniczonym, to zbiór $A(E)$ jest warunkowo zwarty, a z ciągłości odwzorowania B również zbiór $BA(E) = B(A(E))$ jest warunkowo zwarty, a więc operator BA jest zwarty. $\blacksquare$

Wniosek 6.14. *W nieskończenie wymiarowej przestrzeni unormowanej X operator zwarty nie może mieć ograniczonego operatora odwrotnego.*

Dowód. Istotnie, w przeciwnym wypadku operator $I = A^{-1}A$ byłby zwarty na X, a to jest niemożliwe. $\blacksquare$

W przestrzeni funkcji ciągłych $\mathcal{C}[a,b]$ ważną klasę operatorów zwartych tworzą operatory, które mają postać

$$Ax(s) = \int_a^b K(s,t)x(t)\,dt. \tag{6.7}$$

Udowodnimy następujące twierdzenie:

Twierdzenie 6.15. *Jeżeli funkcja* $(s,t) \mapsto K(s,t)$ *jest ograniczona na kwadracie* $[a,b] \times [a,b]$ *i wszystkie jej punkty nieciągłości leżą na skończonej liczbie krzywych o równaniach* $t = \varphi_k(s)$ $(k = 1, 2, \ldots, n)$, *gdzie* φ_k *są funkcjami ciągłymi, to wzór* (6.7) *określa w przestrzeni* $\mathcal{C}[a,b]$ *operator zwarty.*

Dowód. Zauważmy najpierw, że przy wymienionych założeniach całka we wzorze (6.7) istnieje dla dowolnego s z przedziału $[a,b]$. A więc $y(s) = Ax(s)$ jest funkcją określoną na tym przedziale. Weźmy dowolne $\varepsilon > 0$. Niech $M = \sup\{|K(s,t)|\colon\ a \leqslant s \leqslant b,\ a \leqslant t \leqslant b\}$ i niech G będzie zbiorem tych punktów (s,t), dla których chociaż dla jednego $k = 1, 2, \ldots, n$ spełniona jest nierówność

$$|t - \varphi_k(s)| < \frac{\varepsilon}{12Mn}.$$

Przekrojem $G(s)$ zbioru G z każdą prostą $s = \mathrm{const}$ jest suma przedziałów

$$G(s) = \bigcup_{k=1}^{n} \left\{ t\colon\ |t - \varphi_k(s)| < \frac{\varepsilon}{12Mn} \right\}.$$

Niech F będzie dopełnieniem zbioru G do kwadratu $[a,b] \times [a,b]$. Ponieważ F jest zbiorem zwartym, a funkcja $(s,t) \mapsto K(s,t)$ jest ciągła na F, więc jest jednostajnie ciągła (por. [27], tw. 4.13 z części I). Zatem istnieje $\delta > 0$ takie, że

$$|K(s_1,t_1) - K(s_2,t_2)| < \frac{\varepsilon}{3(b-a)}$$

dla dowolnych punktów (s_1,t_1), $(s_2,t_2) \in F$ spełniających warunek $(s_1 - s_2)^2 + (t_1 - t_2)^2 < \delta^2$. Oszacujemy różnicę $y(s_1) - y(s_2)$ przy założeniu, że $|s_1 - s_2| < \delta$. Mamy

$$|y(s_1) - y(s_2)| \leqslant \int_a^b |K(s_1,t_1) - K(s_2,t_2)| |x(t)|\, dt.$$

Aby oszacować tę całkę, rozbijamy przedział $[a,b]$ na sumę przedziałów $G(s_1) \cup G(s_2)$, którą oznaczymy przez P, i pozostałą część przedziału $[a,b]$, którą oznaczymy przez Q. Zauważmy, że P jest sumą przedziałów, których długość jest mniejsza niż $\dfrac{\varepsilon}{3M}$. Otrzymujemy więc

$$\int_P |K(s_1,t_1) - K(s_2,t_2)| |x(t)|\, dt < \frac{2}{3}\varepsilon \|x\|_\infty.$$

Całkę po Q szacujemy w następujący sposób:

$$\int_Q |K(s_1,t_1) - K(s_2,t_2)||x(t)|\,dt < \frac{1}{3}\varepsilon\|x\|_\infty.$$

Zatem

$$|y(s_1) - y(s_2)| < \varepsilon\|x\|_\infty. \tag{6.8}$$

Konsekwencją tej nierówności jest to, że funkcja y jest ciągła, tzn. wzór (6.7) definiuje operator przeprowadzający przestrzeń $\mathcal{C}[a,b]$ w siebie. Dalej z nierówności (6.8) wynika, że jeżeli funkcje x są zawarte w zbiorze ograniczonym w przestrzeni $\mathcal{C}[a,b]$, to odpowiadający im zbiór funkcji y jest jednakowo ciągły. Ponadto, jeżeli $\|x\|_\infty \leqslant C$, to

$$\begin{aligned}\|y\|_\infty = \sup_{a\leqslant s\leqslant b}|y(s)| \leqslant \sup_{a\leqslant s\leqslant b}\int_a^b |K(s,t)||x(t)|\,dt \leqslant\\ \leqslant M(b-a)\|x\|_\infty \leqslant MC(b-a).\end{aligned}$$

Zatem operator A przeprowadza każdy zbiór ograniczony w przestrzeni $\mathcal{C}[a,b]$ w zbiór funkcji wspólnie ograniczonych i jednakowo ciągłych. Na mocy twierdzenia Ascoliego-Arzeli zbiór ten jest warunkowo zwarty. Zatem operator A jest zwarty. ■

Uwagi. 1. To, że punkty nieciągłości funkcji $(s,t) \mapsto K(s,t)$ leżą na skończonej liczbie krzywych, przecinających proste $s = \text{const}$ tylko w jednym punkcie, jest istotne. Niech np.

$$K(s,t) = \begin{cases} 1 & \text{dla } s < \frac{1}{2},\\ 0 & \text{dla } s \geqslant \frac{1}{2}.\end{cases}$$

Operator (6.7) z tak określonym jądrem na kwadracie $[0,1]\times[0,1]$ i mającym punkty nieciągłości na odcinku $s = \frac{1}{2}$, $0 \leqslant t \leqslant 1$, przeprowadza funkcję $x(t) = 1$ dla $0 \leqslant t \leqslant 1$ w funkcję nieciągłą.

2. Jeżeli $K(s,t) = 0$ dla $t > s$, to operator (6.7) będzie miał postać

$$y(s) = \int_a^s K(s,t)x(t)\,dt.$$

Jeżeli funkcja $K(s,t)$ jest ciągła dla $t < s$, to z tego, co udowodniliśmy, wynika, że tak określony operator jest zwarty na przestrzeni $\mathcal{C}[a,b]$. Operator ten nazywa się *operatorem typu Volterry*.

Aby zakończyć opis ogólnych własności operatorów zwartych, udowodnimy jeszcze następujące twierdzenie:

Twierdzenie 6.16 (Schaudera). *Jeżeli A jest operatorem zwartym na przestrzeni Banacha X, to operator sprzężony A' jest również zwarty.*

D o w ó d. Niech (f_n) będzie ciągiem funkcjonałów liniowych ograniczonych z kuli jednostkowej w przestrzeni sprzężonej X' i niech B będzie kulą jednostkową w przestrzeni X. Ponieważ A jest operatorem zwartym, więc zbiór $K = \overline{A(B)}$ jest zwarty.

Udowodnimy, że zbiór $\Phi = \left\{f_n\big|_K :\ n = 1, 2, \dots\right\}$ jest warunkowo zwarty w przestrzeni $\mathcal{C}(K)$. Ponieważ zbiór K jest ograniczony (por. [27], tw. 2.8 z części I), więc istnieje stała $M > 0$ taka, że $\|x\| \leqslant M$ dla każdego $x \in K$. Stąd dla dowolnego $x \in K$ i $n = 1, 2, \dots$ mamy

$$|f_n(x)| \leqslant \|f_n\|\|x\| \leqslant M.$$

Zatem zbiór Φ jest ograniczony. Ponadto dla x_1, $x_2 \in K$ mamy

$$|f_n(x_1) - f_n(x_2)| \leqslant \|f_n\|\|x_1 - x_2\| \leqslant \|x_1 - x_2\|.$$

Oznacza to, że zbiór funkcji Φ jest jednakowo ciągły. Z twierdzenia Ascoliego-Arzeli wynika, że zbiór Φ jest warunkowo zwarty. Zatem z ciągu $\left(f_n\big|_K\right)$ można wybrać podciąg $\left(f_{n_k}\big|_K\right)$ zbieżny w przestrzeni $\mathcal{C}(K)$. Aby zakończyć dowód, wystarczy pokazać, że ciąg $(A' f_{n_k})$ jest zbieżny.

Niech $\varepsilon > 0$ będzie dowolne. Istnieje k_0 takie, że $|f_{n_k}(y) - f_{n_j}(y)| < \varepsilon$ dla każdego $y \in K$ i dowolnych $k, j \geqslant k_0$. Jeżeli $x \in B$, to $Ax \in K$, a zatem dla $k, j \geqslant k_0$ mamy

$$|(A' f_{n_k})x - (A' f_{n_j})x| = |f_{n_k}(Ax) - f_{n_j}(Ax)| < \varepsilon.$$

Ponieważ $x \in B$ było dowolne, więc stąd otrzymujemy

$$\|A' f_{n_k} - A' f_{n_j}\| \leqslant \varepsilon \qquad (k, j \geqslant k_0).$$

Zatem ciąg $(A' f_{n_k})$ jest ciągiem Cauchy'ego, a ponieważ przestrzeń X' jest zupełna, więc jest on zbieżny. ■

U w a g a. W przypadku, gdy rozważamy operatory na przestrzeni Hilberta, dowód twierdzenia Schaudera jest prostszy.

D o w ó d twierdzenia 6.16 przy założeniu, że X jest przestrzenią Hilberta.

Niech (x_n) będzie dowolnym ciągiem ograniczonym punktów przestrzeni X i niech $M = \sup\limits_n \|x_n\|$. Udowodnimy, że z ciągu $(A^* x_n)$ można wybrać podciąg zbieżny.

Ponieważ ciąg (A^*x_n) jest ograniczony, więc ze zwartości operatora A wynika, że ciąg (AA^*x_n) zawiera podciąg zbieżny $(AA^*x_{n_k})$. Zatem

$$\begin{aligned}\|A^*x_{n_k} - A^*x_{n_j}\|^2 &= \langle A^*(x_{n_k} - x_{n_j}), A^*(x_{n_k} - x_{n_j})\rangle = \\ &= \langle AA^*(x_{n_k} - x_{n_j}), x_{n_k} - x_{n_j}\rangle \leqslant \\ &\leqslant \|AA^*(x_{n_k} - x_{n_j})\|\|x_{n_k} - x_{n_j}\|\end{aligned}$$

dla dowolnych wskaźników k i j. Stąd mamy

$$\|A^*x_{n_k} - A^*x_{n_j}\|^2 \leqslant 2M\|AA^*(x_{n_k} - x_{n_j})\|.$$

Zatem ciąg $(A^*x_{n_k})$ jest ciągiem zbieżnym, bo spełnia warunek Cauchy'ego. ■

6.4. Widmo operatora zwartego

Lemat 6.17. *Niech A będzie operatorem zwartym na przestrzeni Banacha X. Jeżeli $\lambda \neq 0$, to przestrzeń* $\ker(A - \lambda I)$ *jest skończenie wymiarowa.*

D o w ó d. Ze zwartości operatora A wynika, że każdy ciąg ograniczony (x_n), $x_n \in \ker(A - \lambda I)$, zawiera podciąg (x_{n_k}) taki, że (Ax_{n_k}) jest ciągiem zbieżnym. Mamy $Ax_{n_k} = \lambda x_{n_k}$, co oznacza, że ciąg (x_{n_k}) jest zbieżny. Stąd wynika, że każdy ograniczony i domknięty podzbiór przestrzeni $\ker(A - \lambda I)$ jest zwarty. Z wniosku 6.7 otrzymujemy, że przestrzeń $\ker(A - \lambda I)$ jest skończenie wymiarowa. ■

Lemat 6.18. *Niech A będzie operatorem zwartym na przestrzeni Banacha X. Jeżeli $\lambda \neq 0$, to przestrzeń* $\operatorname{im}(A - \lambda I)$ *jest domknięta.*

D o w ó d. Należy udowodnić, że jeżeli (y_n) jest dowolnym ciągiem w obrazie $\operatorname{im}(A - \lambda I)$ i $y_n \to y$, to również $y \in \operatorname{im}(A - \lambda I)$.

Najpierw wykażemy, że istnieje ciąg ograniczony (x_n) taki, że $y_n = (A - \lambda I)x_n$. Ponieważ $y_n \in \operatorname{im}(A - \lambda I)$, więc istnieje ciąg (z_n) (niekoniecznie ograniczony), dla którego $y_n = (A - \lambda I)z_n$. Jeżeli (w_n) jest dowolnym ciągiem w $\ker(A - \lambda I)$ i $x_n = z_n - w_n$, to $y_n = (A - \lambda I)x_n$. Zatem, jeżeli uda się wybrać ciąg (w_n) wektorów z $\ker(A - \lambda I)$ w taki sposób, że ciąg $(z_n - w_n)$ będzie ograniczony, to zakończy dowód stwierdzenia. W tym celu wykażemy, że wyrażenie $d(z_n) = \operatorname{dist}(z_n, \ker(A - \lambda I))$ jest ograniczone jako funkcja zmiennej n. Przypuśćmy, że tak nie jest, tzn. ciąg $(d(z_n))$ jest nieograniczony. Przechodząc w razie potrzeby do podciągu, możemy założyć, że $d(z_n) \to +\infty$, gdy $n \to \infty$. Ponadto możemy założyć, że $d(z_n) \geqslant 1$ dla wszystkich n. Oznaczmy $z_n' = \dfrac{1}{d(z_n)} z_n$. Wykażemy, że wówczas $d(z_n') = 1$. Dla dowolnego $\varepsilon > 0$ istnieje wektor $v_n \in \ker(A - \lambda I)$ taki, że $\|z_n - v_n\| < d(z_n) + \varepsilon$. Stąd

$$\Big\| z_n' - \frac{1}{d(z_n)} v_n \Big\| < 1 + \frac{1}{d(z_n)} \varepsilon \leqslant 1 + \varepsilon.$$

To oznacza, że $d(z_n') = \operatorname{dist}(z_n', \ker(A - \lambda I)) \leqslant 1$. Gdyby $d(z_n') < 1$, to $\|z_n' - u_n\| < 1$ dla pewnego wektora $u_n \in \ker(A - \lambda I)$. Wówczas byłoby $\|z_n - d(z_n) u_n\| < d(z_n)$, co nie jest możliwe, bo $d(z_n) u_n \in \ker(A - \lambda I)$. Zatem $d(z_n') = \operatorname{dist}(z_n', \ker(A - \lambda I)) = 1$. Stąd wynika, że istnieje ciąg wektorów (t_n) taki, że $t_n \in \ker(A - \lambda I)$ i $\|z_n' - t_n\| \leqslant 2$ dla wszystkich n. Zatem ciąg $(w_n) = (z_n' - t_n)$ jest ograniczony. Ponieważ operator A jest zwarty, więc przechodząc w razie potrzeby do podciągu, możemy założyć, że (Aw_n) jest ciągiem zbieżnym. Ponadto

$$(A - \lambda I) w_n = (A - \lambda I) z_n' = \frac{1}{d(z_n)} (A - \lambda I) z_n = \frac{1}{d(z_n)} y_n \to 0,$$

bo (y_n) jest ciągiem zbieżnym, a więc ograniczonym. Zatem ciąg (w_n) jest zbieżny. Niech $w_n \to w$, gdy $n \to \infty$. Wówczas $d(w_n) \to d(w)$ oraz

$$Aw - \lambda w = \lim_{n \to \infty} (Aw_n - \lambda w_n) = 0.$$

Zatem $w \in \ker(A - \lambda I)$, czyli

$$0 = d(w) = \lim_{n \to \infty} d(w_n) = \lim_{n \to \infty} d(z_n') = 1.$$

Otrzymana sprzeczność pokazuje, że istotnie ciąg $(d(z_n))$ jest ograniczony.

Tym samym udowodniliśmy istnienie ograniczonego ciągu (x_n), dla którego

$$y_n = (A - \lambda I) x_n.$$

Na mocy zwartości operatora A ciąg (x_n) zawiera podciąg (x_{n_k}) taki, że (Ax_{n_k}) jest ciągiem zbieżnym. Ponieważ $y_{n_k} = (A - \lambda I) x_{n_k}$ i $y_{n_k} \to y$, więc również ciąg (x_{n_k}) jest zbieżny. Niech $x_{n_k} \to x$. Wówczas

$$y = \lim_{k \to \infty} y_{n_k} = \lim_{k \to \infty} (Ax_{n_k} - \lambda x_{n_k}) = Ax - \lambda x,$$

czyli $y \in \operatorname{im}(A - \lambda I)$. ■

Lemat 6.19. *Jeżeli A jest operatorem zwartym na przestrzeni Banacha X i $\delta > 0$, to istnieje co najwyżej skończenie wiele wartości własnych operatora A o modułach większych bądź równych δ.*

D o w ó d. Przypuśćmy, że teza nie jest prawdziwa. Istnieją więc $\delta > 0$ i ciąg (λ_n) takie, że $\ker(A - \lambda_n I) \neq \{0\}$ i $|\lambda_n| \geqslant \delta$ dla $n = 1, 2, \ldots$ Możemy założyć, że $\lambda_n \neq \lambda_m$ dla $n \neq m$. Niech x_n oznacza wektor własny odpowiadający wartości własnej λ_n ($n = 1, 2, \ldots$). Wektory x_1, x_2, $\ldots$, x_n są liniowo niezależne dla każdego n. W przeciwnym wypadku istniałby wskaźnik n taki, że wektory x_1, x_2, $\ldots$, x_{n-1} byłyby liniowo niezależne, a $x_n = \sum_{j=1}^{n-1} \alpha_j x_j$ dla pewnych α_1, $\ldots$, α_{n-1}. Wówczas

$$0 = Ax_n - \lambda_n x_n = \sum_{j=1}^{n-1} \alpha_j (Ax_j - \lambda_n x_j) = \sum_{j=1}^{n-1} \alpha_j (\lambda_j - \lambda_n) x_j.$$

Ponieważ $\lambda_j \neq \lambda_n$ dla $j < n$, więc $\alpha_1 = \alpha_2 = \ldots = \alpha_{n-1} = 0$ i w konsekwencji $x_n = 0$, co nie jest możliwe.

Niech $X_n = \operatorname{span}\{x_1, x_2, \ldots, x_n\}$. Wówczas

1° X_n jest domkniętą podprzestrzenią przestrzeni X jako podprzestrzeń skończenie wymiarowa;

2° $X_n \subset X_{n+1}$ w sposób właściwy ($n = 1, 2, \ldots$).

Zatem na mocy lematu Riesza istnieje ciąg wektorów (z_n) taki, że

$$z_n \in X_n, \quad \|z_n\| = 1, \quad \|z_n - x\| \geqslant \frac{1}{2} \quad \text{dla} \quad x \in X_{n-1} \quad (n = 2, 3, \ldots).$$

Mamy

$$A\left(\frac{1}{\lambda_n} z_n\right) - A\left(\frac{1}{\lambda_m} z_m\right) = z_n - x,$$

gdzie $x = \dfrac{1}{\lambda_n}(\lambda_n z_n - Az_n) + \dfrac{1}{\lambda_m} Az_m$. Ponadto $\lambda_n z_n - Az_n \in X_{n-1}$. Niech $n > m$. Wtedy $Az_m \in X_{n-1}$ oraz

$$\left\| A\left(\frac{1}{\lambda_n} z_n\right) - A\left(\frac{1}{\lambda_m} z_m\right) \right\| \geqslant \frac{1}{2} \quad \text{dla} \quad n > m.$$

Zatem ciąg $\left(A\left(\dfrac{1}{\lambda_n} z_n\right)\right)$ nie zawiera podciągów zbieżnych, a ponieważ

$$\left\| \frac{1}{\lambda_n} z_n \right\| = \frac{1}{|\lambda_n|} \leqslant \frac{1}{\delta} \qquad (n = 2, 3, \ldots),$$

więc przeczy to zwartości operatora A. ■

Twierdzenie 6.20 (Riesza o widmie operatora zwartego). *Niech A będzie operatorem zwartym na zespolonej przestrzeni Banacha X. Wówczas:*

(a) *Każda liczba $\lambda \neq 0$ należąca do widma operatora A jest jego wartością własną.*

(b) *Dla każdego $\lambda \neq 0$ podprzestrzeń $\ker(A-\lambda I)$ jest skończenie wymiarowa.*

(c) *Zbiór wartości własnych operatora A jest co najwyżej przeliczalny. Jego punktem skupienia może być jedynie punkt $\lambda = 0$.*

D o w ó d. Część (b) jest treścią lematu 6.17. Natomiast część (c) wynika z lematu 6.19. Istotnie dla dowolnego naturalnego n zbiór wartości własnych spełniających nierówność $|\lambda| \geqslant \frac{1}{n}$ jest skończony. Zatem zbiór wartości własnych operatora A jest co najwyżej przeliczalny. Z lematu 6.19 wynika również, że żadna liczba $\lambda \neq 0$ nie może być punktem skupienia zbioru wartości własnych operatora A.

Pozostaje udowodnić, że prawdziwa jest część (a) tezy. Przypuśćmy, że tak nie jest. Istnieje zatem liczba $\lambda \neq 0$, która należy do widma operatora A, ale nie jest jego wartością własną. Oznacza to, że $\ker(A - \lambda I) = \{0\}$ i $\operatorname{im}(A - \lambda I) \neq X$. Zatem istnieje $y_0 \in X$ takie, że $y_0 \notin \operatorname{im}(A - \lambda I)$. Inaczej mówiąc, równanie $(A - \lambda I)x = y_0$ nie ma rozwiązania.

Niech

$$X_0 = X,\ X_1 = \operatorname{im}(A - \lambda I),\ X_2 = \operatorname{im}(A - \lambda I)^2, \dots, X_n = \operatorname{im}(A - \lambda I)^n, \dots$$

Każdy ze zbiorów X_n jest podprzestrzenią liniową przestrzeni X. Ponadto każda z podprzestrzeni X_n jest domknięta na mocy lematu 6.18, bo jest ona obrazem przestrzeni X poprzez operator postaci *operator zwarty minus* $\lambda^n I$. Jest oczywiste, że

$$X_{n+1} \subset X_n \qquad (n = 0, 1, 2, \dots).$$

Pokażemy, że X_{n+1} jest właściwą podprzestrzenią przestrzeni X_n. Aby to udowodnić, zdefiniujmy

$$y_n = (A - \lambda I)^n y_0 \quad \text{dla} \quad n = 1, 2, \dots$$

Wówczas $y_n \in X_n \setminus X_{n+1}$.

Dowód poprowadzimy przez indukcję ze względu na n. Wiemy, że jest to prawdą dla $n = 0$. Załóżmy, że teza zachodzi dla liczby n. Pokażemy, że jest również prawdziwa dla $n + 1$. Gdyby $y_{n+1} \in X_{n+2}$, to istniałby wektor x_0 taki, że $y_{n+1} = (A - \lambda I)^{n+2} x_0$. Zatem

$$(A - \lambda I)^{n+1} y_0 = (A - \lambda I)^{n+2} x_0,$$

czyli

$$(A - \lambda I)y_n = (A - \lambda I)^{n+2}x_0.$$

Stąd wynika, że $y_n - (A - \lambda I)^{n+1}x_0 \in \ker(A - \lambda I) = \{0\}$ i w konsekwencji $y_n = (A - \lambda I)^{n+1}x_0$, czyli $y_n \in X_{n+1}$ wbrew założeniu indukcyjnemu. Otrzymana sprzeczność kończy dowód.

Udowodniliśmy więc, że każda z podprzestrzeni X_n jest domknięta i X_{n+1} jest właściwym podzbiorem przestrzeni X_n. Na mocy lematu Riesza istnieje ciąg wektorów (z_n) o następujących własnościach:

$$z_n \in X_n, \quad \|z_n\| = 1, \quad \|z_n - x\| \geqslant \frac{1}{2} \quad \text{dla} \quad x \in X_{n+1}. \tag{6.9}$$

Mamy

$$Az_n - Az_m = \lambda(z_n - x),$$

gdzie $x = z_m - \frac{1}{\lambda}\big((A-\lambda I)z_n - (A-\lambda I)z_m\big)$. Jeżeli $m > n$, to $z_m \in X_{n+1} \subset \subset X_n$ i z (6.9) wynika, że

$$\|Az_n - Az_m\| \geqslant \frac{1}{2}|\lambda| \quad \text{dla} \quad m > n.$$

Wobec tego ciąg (Az_n) nie zawiera podciągów zbieżnych. Ponieważ $\|z_n\| = 1$ $(n = 1, 2, \dots)$, więc otrzymaliśmy sprzeczność ze zwartością operatora A. ■

Niech X będzie przestrzenią liniową skończenie wymiarową i $A\colon\ X \to X$ operatorem liniowym. Wówczas zachodzi wzór (patrz [26], tw. 5.2)

$$\dim X = \dim \operatorname{im} A + \dim \ker A.$$

Z tego wzoru wynika następujący fakt:

Albo równanie niejednorodne $Ax = y$ ma dla dowolnego wektora $y \in X$ dokładnie jedno rozwiązanie, albo równanie jednorodne $Ax = 0$ ma rozwiązania niezerowe.

Twierdzenie to nosi nazwę *alternatywy Fredholma*. Zauważmy, że alternatywa Fredholma nie zachodzi dla operatorów (liniowych i ograniczonych) na przestrzeniach nieskończenie wymiarowych.

Przykład 6.1. Niech S będzie operatorem jednostronnego przesunięcia na przestrzeni ℓ_2, tzn. $S(\xi_1, \xi_2, \dots) = (0, \xi_1, \xi_2, \dots)$. Wówczas $\ker S = \{0\}$, ale $\operatorname{im} S \neq \ell_2$, bo $\operatorname{im} S$ nie zawiera wektorów postaci $(\xi_1, 0, 0, \dots)$.

Sytuacja jest inna w przypadku operatorów zwartych. Mówi o tym następujące twierdzenie:

Twierdzenie 6.21 (alternatywa Fredholma). *Niech A będzie operatorem zwartym na przestrzeni Banacha X. Wówczas dla dowolnego $\lambda \neq 0$ albo równanie $Ax - \lambda x = y$ ma dokładnie jedno rozwiązanie dla każdego wektora $y \in X$, albo równanie $Ax - \lambda x = 0$ ma rozwiązania niezerowe.*

Dowód. Twierdzenie to jest bezpośrednią konsekwencją twierdzenia Riesza i oczywistego faktu, że dowolna liczba $\lambda \neq 0$ jest albo elementem widma operatora A, albo do niego nie należy. ■

6.5. Operatory całkowe. Twierdzenia Fredholma

W tym podrozdziale zajmiemy się pewną klasą operatorów zwartych na przestrzeni Hilberta, niezwykle ważnych z punktu widzenia zastosowań w teorii równań całkowych. Są to operatory całkowe z jądrem całkowalnym z kwadratem.

Niech $\Omega \subset \mathbb{R}^k$ będzie zbiorem mierzalnym (w sensie Lebesgue'a) o mierze dodatniej. Niech $K \in L_2(\Omega \times \Omega)$. Określamy funkcję

$$(Ax)(s) = \int_\Omega K(s,t)x(t)\,dt. \tag{6.10}$$

Wykażemy, że jeżeli $x \in L_2(\Omega)$, to całka $\int_\Omega K(s,t)x(t)\,dt$ istnieje i jest skończona dla prawie każdego $s \in \Omega$ oraz $Ax \in L_2(\Omega)$. Ponieważ

$$\iint_{\Omega\times\Omega} |K(s,t)|^2\,dsdt < \infty,$$

więc na mocy twierdzenia Fubiniego

$$\int_\Omega |K(s,t)|^2\,dt < \infty,$$

dla prawie każdego $s \in \Omega$ i całka ta jest funkcją mierzalną zmiennej s oraz

$$\int_\Omega ds \int_\Omega |K(s,t)|^2\,dt = \iint_{\Omega\times\Omega} |K(s,t)|^2\,dsdt. \tag{6.11}$$

Z nierówności Schwarza otrzymujemy

$$|(Ax)(s)|^2 = \left|\int_\Omega K(s,t)x(t)\,dt\right|^2 \leqslant \int_\Omega |K(s,t)|^2\,dt \cdot \int_\Omega |x(t)|^2\,dt.$$

Stąd

$$\int\limits_{\Omega} |(Ax)(s)|^2\,ds \leqslant \int\limits_{\Omega}\left(\int\limits_{\Omega} |K(s,t)|^2\,dt\right) ds \cdot \int\limits_{\Omega} |x(t)|^2\,dt, \tag{6.12}$$

czyli $Ax \in L_2(\Omega)$. Z własności całki wynika, że A jest operatorem liniowym. Ponadto uwzględniając (6.11), możemy (6.12) przepisać w postaci

$$\|Ax\|_2 \leqslant M\|x\|_2, \quad \text{gdzie} \quad M = \left(\iint\limits_{\Omega\times\Omega} |K(s,t)|^2\,dsdt\right)^{1/2},$$

co oznacza, że A jest operatorem liniowym ograniczonym.

Definicja 6.8. Funkcję $K(s,t)$ nazywamy *jądrem operatora całkowego* A zdefiniowanego wzorem (6.10).

Zauważmy, że jądro jest wyznaczone jednoznacznie (z dokładnością do zbioru miary zero) przez operator A.

Lemat 6.22. *Jeżeli dwa jądra K_1 i K_2 określają ten sam operator całkowy A, to $K_1(s,t) = K_2(s,t)$ prawie wszędzie na $\Omega \times \Omega$.*

Dowód. Oznaczmy $K(s,t) = K_1(s,t) - K_2(s,t)$. Pokażemy, że jeżeli dla każdej funkcji $x \in L_2(\Omega)$

$$\int\limits_{\Omega} K(s,t)x(t)\,dt = 0 \quad \text{prawie wszędzie na} \quad \Omega, \tag{6.13}$$

to $K(s,t) = 0$ prawie wszędzie na $\Omega \times \Omega$. Weźmy dowolny ciąg funkcji (φ_n), stanowiący bazę ortonormalną przestrzeni $L_2(\Omega)$. Z warunku (6.13) mamy

$$\int\limits_{\Omega} K(s,t)\overline{\varphi_n(t)}\,dt = 0 \quad \text{dla} \quad n = 1,2,\ldots$$

Mnożąc tę równość przez $\overline{\varphi_m(s)}$ $(m = 1,2,\ldots)$ i całkując, zgodnie z twierdzeniem Fubiniego otrzymujemy

$$\iint\limits_{\Omega\times\Omega} K(s,t)\overline{\varphi_m(s)\varphi_n(t)}\,dsdt = 0 \quad \text{dla} \quad m,n = 1,2,\ldots \tag{6.14}$$

Ponieważ na mocy twierdzenia 2.27 zbiór iloczynów $\varphi_m(s)\varphi_n(t)$ jest bazą ortonormalną przestrzeni $L_2(\Omega \times \Omega)$, więc z (6.14) wynika, że $K(s,t) = 0$ prawie wszędzie na $\Omega \times \Omega$. ■

Niech $K \in L_2(\Omega \times \Omega)$ i niech A będzie operatorem na przestrzeni $L_2(\Omega)$ danym wzorem

$$Ax(s) = \int_\Omega K(s,t)x(t)\,dt \qquad (s \in \Omega).$$

Na mocy twierdzenia Fubiniego mamy

$$\begin{aligned}\langle Ax, y\rangle &= \int_\Omega \left(\int_\Omega K(s,t)x(t)\,dt\right)\overline{y(s)}\,ds = \\ &= \iint_{\Omega\times\Omega} K(s,t)\overline{y(s)}x(t)\,dsdt = \int_\Omega \left(x(t)\overline{\int_\Omega \overline{K(s,t)}y(s)\,ds}\right)dt = \\ &= \int_\Omega \left(x(s)\overline{\int_\Omega \overline{K(t,s)}y(t)\,dt}\right)ds = \langle x, A^*y\rangle.\end{aligned}$$

Stąd

$$A^*y(s) = \int_\Omega \overline{K(t,s)}y(t)\,dt \qquad (s \in \Omega).$$

Wobec tego prawdziwe jest następujące twierdzenie:

Twierdzenie 6.23. *Jeżeli A jest operatorem całkowym na przestrzeni $L_2(\Omega)$, określonym przez jądro $(s,t) \mapsto K(s,t)$, $K \in L_2(\Omega \times \Omega)$, to operator z nim sprzężony A^* jest operatorem całkowym określonym przez jądro $(s,t) \mapsto \overline{K(t,s)}$. Operator A jest samosprzężony wtedy i tylko wtedy, gdy*

$$K(s,t) = \overline{K(t,s)} \quad \textit{prawie wszędzie na} \quad \Omega \times \Omega. \tag{6.15}$$

Jądro spełniające warunek (6.15) *nazywamy* jądrem symetrycznym.

Twierdzenie 6.24. *Operator $A \in \mathcal{B}(L_2(\Omega))$ z jądrem $K \in L_2(\Omega \times \Omega)$ jest operatorem zwartym.*

D o w ó d. Najpierw zauważmy, że jeżeli K jest *jądrem zdegenerowanym*, tzn.

$$K(s,t) = \sum_{j=1}^{n} \gamma_j a_j(s) b_j(t),$$

gdzie γ_j są liczbami i a_j, $b_j \in L_2(\Omega)$ $(j = 1, 2, \ldots, n)$, to operator A określony wzorem (6.10) jest skończenie wymiarowy. Istotnie mamy

$$Ax(s) = \sum_{j=1}^{n} \gamma_j a_j(s) \int_{\Omega} b_j(t) x_j(t)\, dt,$$

co oznacza, że wektor Ax należy do podprzestrzeni skończenie wymiarowej rozpiętej na wektorach $a_1, \ldots, a_n$. Ponieważ ograniczone operatory skończenie wymiarowe są zwarte (patrz tw. 6.10), więc stąd wynika, że każdy operator całkowy z jądrem zdegenerowanym jest zwarty.

Niech będzie dane dowolne jądro $K \in L_2(\Omega \times \Omega)$. Weźmy jakąkolwiek bazę ortonormalną (φ_n) w przestrzeni $L_2(\Omega)$. Na mocy twierdzenia 2.27 funkcje

$$(s, t) \mapsto \varphi_n(s)\varphi_m(t) \qquad (n,\, m = 1, 2, \ldots) \tag{6.16}$$

tworzą bazę ortonormalną w przestrzeni $L_2(\Omega \times \Omega)$. Stąd wynika, że zbiór kombinacji liniowych funkcji (6.16) jest gęsty w przestrzeni $L_2(\Omega \times \Omega)$. Każda z tych kombinacji liniowych jest jądrem zdegenerowanym. Zatem istnieje ciąg (K_n) jąder zdegenerowanych zbieżny do jądra K w przestrzeni $L_2(\Omega \times \Omega)$, tzn.

$$\iint_{\Omega \times \Omega} |K_n(s,t) - K(s,t)|^2\, dsdt \to 0, \quad \text{gdy} \quad n \to \infty. \tag{6.17}$$

Niech

$$A_n x(s) = \int_{\Omega} K_n(s,t) x(t)\, dt \quad (n = 1, 2, \ldots).$$

Wtedy $A_n - A$ jest operatorem o jądrze $K_n - K$. Z (6.17) i nierówności

$$\|A_n - A\|^2 \leqslant \iint_{\Omega \times \Omega} |K_n(s,t) - K(s,t)|^2\, dsdt$$

wynika, że $A_n \to A$. Ponieważ każdy z operatorów A_n jest zwarty, więc na mocy twierdzenia 6.12 operator A jest zwarty. ■

Rozważmy równanie całkowe (*Fredholma drugiego rodzaju*)

$$\varphi(s) = \int_{\Omega} K(s,t)\varphi(t)\, dt + f(s), \tag{6.18}$$

gdzie $K \in L_2(\Omega \times \Omega)$ i $\varphi, f \in L_2(\Omega)$. Chcemy znaleźć warunki na to, aby równanie to miało rozwiązanie, oraz zbadamy własności jego rozwiązań. W naszych rozważaniach istotne będzie to, że operator całkowy jest zwarty, a nie sam fakt, że równanie to jest równaniem całkowym. Zatem rozważymy abstrakcyjny wariant tego równania:

$$\varphi = A\varphi + f, \tag{6.19}$$

gdzie A jest operatorem zwartym na przestrzeni Hilberta H oraz φ, $f \in H$. Jeżeli oznaczymy $T = I - A$, to równanie (6.19) przybierze postać

$$T\varphi = f. \tag{6.20}$$

Będziemy również rozpatrywać równanie jednorodne

$$T\varphi_0 = 0 \tag{6.21}$$

i *równania sprzężone*

$$T^*\psi = g, \tag{6.22}$$

$$T^*\psi_0 = 0. \tag{6.23}$$

Związki pomiędzy własnościami rozwiązań tych czterech równań są wyrażone w następujących twierdzeniach:

Twierdzenia 6.25 (Fredholma). (I) *Równanie niejednorodne $T\varphi = f$ ma rozwiązanie wtedy i tylko wtedy, gdy f jest ortogonalne do dowolnego rozwiązania sprzężonego równania jednorodnego $T^*\psi_0 = 0$.*

(II) *Albo równanie $T\varphi = f$ ma dla dowolnego $f \in H$ dokładnie jedno rozwiązanie, albo równanie jednorodne $T\varphi_0 = 0$ ma rozwiązanie niezerowe.*

(III) *Równania jednorodne $T\varphi_0 = 0$, $T^*\psi_0 = 0$ mają skończony zbiór liniowo niezależnych rozwiązań, przy tym wymiary przestrzeni rozwiązań obu tych równań są identyczne.*

Część (II) tego twierdzenia nosi nazwę *alternatywy Fredholma* (por. twierdzenie 6.21).

D o w ó d. (I) Na mocy lematu 6.18 podprzestrzeń $\operatorname{im} T$ jest domknięta. Z twierdzenia 4.11 otrzymujemy więc, że

$$\ker T \oplus \operatorname{im} T^* = H$$

oraz

$$\ker T^* \oplus \operatorname{im} T = H. \tag{6.24}$$

Stąd już wynika teza, bo $f \perp \ker T^*$ wtedy i tylko wtedy, gdy $f \in \operatorname{im} T$, tzn. istnieje φ takie, że $T\varphi = f$.

(II) Dowód tej części wynika z tego, że albo liczba 1 należy do widma operatora zwartego A i wówczas jest ona jego wartością własną, czyli równanie $T\varphi_0 = 0$ ma niezerowe rozwiązania, albo liczba 1 nie należy do widma operatora A i wówczas przekształcenie T jest bijekcją, czyli równanie $T\varphi = f$ ma rozwiązanie (i to dokładnie jedno) dla dowolnego $f \in H$.

(III) Z twierdzeń 6.20 (Riesza) i 6.16 (Schaudera) wynika, że podprzestrzenie $\ker T$ i $\ker T^*$ są skończenie wymiarowe. Pozostaje wykazać, że mają one ten sam wymiar. Niech $\dim \ker T = n$ i $\dim \ker T^* = m$. Przypuśćmy, że $n < m$. Niech $\varphi_1, \ldots, \varphi_n$ będzie bazą ortonormalną dla $\ker T$, a $\psi_1, \ldots, \psi_m$ bazą ortonormalną dla $\ker T^*$. Oznaczmy

$$Sx = Tx + \sum_{j=1}^{n} \langle x, \varphi_j \rangle \psi_j.$$

Operator S powstaje poprzez dodanie do operatora T operatora skończenie wymiarowego. Ma więc on postać $I-$*operator zwarty*, zatem możemy do niego stosować części (I) i (II).

Pokażemy, że równanie $Sx = 0$ ma tylko rozwiązanie zerowe. Zatem niech

$$Tx + \sum_{j=1}^{n} \langle x, \varphi_j \rangle \psi_j = 0. \tag{6.25}$$

Ponieważ na mocy (6.24) wektory ψ_j są ortogonalne do wszystkich wektorów postaci Tx, więc z (6.25) wynika, że

$$Tx = 0$$

i w konsekwencji

$$\langle x, \varphi_j \rangle = 0 \quad \text{dla} \quad j = 1, 2, \ldots, n.$$

Tak więc, z jednej strony wektor x powinien być kombinacją liniową wektorów φ_j, a z drugiej strony jest do nich ortogonalny. Zatem $x = 0$, czyli równanie $Sx = 0$ ma tylko zerowe rozwiązanie. Z alternatywy Fredholma, zastosowanej do operatora S, wynika, że istnieje wektor $y \in H$ taki, że

$$Ty + \sum_{j=1}^{n} \langle y, \varphi_j \rangle \psi_j = \psi_{n+1}.$$

Jeżeli pomnożymy to równanie skalarnie przez ψ_{n+1}, to z prawej strony otrzymamy 1, a z lewej 0, bo $Ty \in \operatorname{im} T$, a $\operatorname{im} T \perp \ker T^*$. Otrzymaliśmy sprzeczność. Nasze przypuszczenie, że $n < m$, okazało się fałszywe. Zatem $n \geqslant m$. Zamieniając operator T na T^*, otrzymujemy $n \leqslant m$ i w konsekwencji $n = m$. ∎

U w a g a. Udowodniliśmy twierdzenia Fredholma dla równań postaci $\varphi = A\varphi + f$, gdzie A jest operatorem zwartym na przestrzeni Hilberta. Twierdzenia te pozostają prawdziwe w przypadku operatorów na przestrzeni Banacha X. Tylko wówczas równanie sprzężone $\psi = A'\psi + g$ jest równaniem na przestrzeni sprzężonej X', a warunek ortogonalności $\langle f, \psi_0 \rangle = 0$ należy rozumieć jako zerowanie się na wektorze $f \in X$ każdego funkcjonału liniowego ograniczonego, należącego do podprzestrzeni $\ker T' \subset X'$ rozwiązań równania $T'\psi_0 = 0$.

Ćwiczenia

1. Niech $(s,t) \mapsto K(s,t)$ będzie funkcją ciągłą na kwadracie $[0,1] \times [0,1]$. Udowodnić, że operator całkowy $Ax(s) = \int\limits_0^1 K(s,t)x(t)\,dt$ jest operatorem zwartym na przestrzeni $\mathcal{C}[0,1]$.

2. Niech $\left(\alpha_k^j\right)_{j,k=1}^{\infty}$ będzie macierzą nieskończoną, spełniającą warunek

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} |\alpha_k^j|^2 < \infty.$$

Operator $A \in \mathcal{B}(\ell_2)$ jest zdefiniowany wzorem $Ax = \left(\sum\limits_{k=1}^{\infty} \alpha_k^j \xi_k\right)_{j=1}^{\infty}$, gdzie $x = (\xi_k)$. Wykazać, że A jest operatorem zwartym.

3. Udowodnić, że operator Volterry $Vx(s) = \int\limits_0^s x(t)\,dt$ jest operatorem zwartym na przestrzeni $L_2[0,1]$.

4. Niech a będzie funkcją mierzalną i ograniczoną na zbiorze mierzalnym $\Omega \subset \mathbb{R}^n$, nierówną zeru prawie wszędzie. Wykazać, że operator A na przestrzeni $L_2(\Omega)$ określony wzorem

$$(Ax)(t) = a(t)x(t) \qquad (t \in \Omega)$$

nie jest operatorem zwartym.

5. Niech X oznacza przestrzeń Banacha, którą tworzą wszystkie funkcje ciągłe i ograniczone na prostej $\mathbb{R}$ z normą $\|x\|_\infty = \sup\{|x(t)|\colon\ t \in (-\infty, +\infty)\}$.

Sprawdzić, że operator całkowy A określony wzorem

$$(Ax)(s) = \int_{-\infty}^{+\infty} e^{-|s-t|} x(t)\, dt$$

przekształca przestrzeń X w siebie oraz jest liniowy i ograniczony. Ponadto wykazać, że nie jest on operatorem zwartym. (*Wskazówka*. Rozważyć ciąg funkcji (x_n) takich, że $x_n(t) = 1$ dla $t \leqslant n$, $x_n(t) = 0$ dla t$\geqslant n+1$ i $x_n(t)$ jest liniowe na przedziale $[n, n+1]$).

6. Czy operator P, spełniający równość $P^2 = P$, może być zwarty?

7. Niech X będzie przestrzenią unormowaną, a Y — przestrzenią Banacha. Wykazać, że jeżeli ciąg operatorów zwartych $A_n \in \mathcal{B}(X, Y)$ $(n = 1, 2, \dots)$ jest zbieżny (w przestrzeni $\mathcal{B}(X, Y)$) do operatora A, to dla każdego zbioru ograniczonego $Z \subset X$ zbiór $\bigcup\limits_{n=1}^{\infty} A_n(Z)$ jest warunkowo zwarty.

8. Niech X będzie nieskończenie wymiarową przestrzenią Banacha i $A \in \mathcal{B}(X)$ — operatorem zwartym. Niech $\lambda \notin \sigma(A)$. Udowodnić, że

$$(A - \lambda I_X)^{-1} = K_\lambda - \frac{1}{\lambda} I_X,$$

gdzie K_λ jest operatorem zwartym.

9. Niech A będzie takim operatorem zwartym na przestrzeni Banacha X, że $A^n x \to$ $\to 0$ dla każdego $x \in X$. Udowodnić, że $1 \notin \sigma(A)$. Wykazać na przykładzie, że założenie zwartości operatora A jest istotne.

10. Niech (e_n) będzie bazą ortonormalną nieskończenie wymiarowej przestrzeni Hilberta H. Sprawdzić, że operator $A \in \mathcal{B}(H)$, określony wzorem

$$Ax = \sum_{n=1}^{\infty} \frac{1}{n} \langle x, e_n \rangle e_{n+1},$$

jest zwarty i nie ma wartości własnych.

11. Niech (e_n) będzie układem ortonormalnym w przestrzeni Hilberta H. Wykazać, że operator A w przestrzeni H, określony wzorem

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n,$$

jest zwarty wtedy i tylko wtedy, gdy $\lambda_n \to 0$.

Rozdział 7

Twierdzenie spektralne

Głównym celem tego rozdziału jest przedstawienie dowodu twierdzenia spektralnego dla operatora samosprzężonego na przestrzeni Hilberta. Twierdzenie to jest bardzo ważnym „narzędziem” do badania własności ograniczonych operatorów liniowych na przestrzeni Hilberta. Jego istota wynika z faktu, że sprowadza ono badanie wielu własności operatorów samosprzężonych (jak również normalnych i unitarnych) do badania własności najprostszych operatorów samosprzężonych, a mianowicie operatorów rzutowania ortogonalnego.

7.1. Operatory rzutowania

Niech M będzie domkniętą podprzestrzenią przestrzeni Hilberta H. Wówczas z twierdzenia o rzucie ortogonalnym (twierdzenie 2.5) wynika, że przestrzeń H ma rozkład na ortogonalną sumę prostą $H = M \oplus M^{\perp}$, a więc istnieje odwzorowanie $P\colon\ H \to H$, które ma następujące własności:

(i) $P(H) = M$ (wn. 2.9 (i));

(ii) P jest przekształceniem liniowym (wn. 2.9 (h));

(iii) $\|P\| = 1$ o ile $M \neq \{0\}$ (wn. 2.9 (d), (f));

(iv) $P = P^*$ (wn. 2.9 (a));

(v) $P^2 = P$ (wn. 2.9 (b));

(vi) $\ker P \perp \operatorname{im} P$ (wn. 2.9 (g), (i)).

Odwzorowanie P jest wyznaczone jednoznacznie i jest nazywane *rzutowaniem ortogonalnym* lub *rzutem ortogonalnym* przestrzeni H na podprzestrzeń M. Ponieważ wszystkie rozważane przez nas rzuty będą rzutowaniami ortogonalnymi, w dalszym ciągu słowo „ortogonalne” będziemy opuszczać.

Zauważmy, że jeżeli P jest rzutem na podprzestrzeń M, to operator $I-P$ jest rzutem na $M^{\perp}$.

Istotnie, $I - P$ jest operatorem samosprzężonym i $(I-P)^2 = I - P$ oraz mamy $x - Px \perp Px$ dla każdego wektora $x \in H$, a to oznacza, że $\mathrm{im}(I-P) = M^{\perp}$ (por. tw. 4.12).

Dwa operatory rzutowania P_1 i P_2 są *ortogonalne*, jeżeli $P_1P_2 = 0$. Warunek ten jest równoważny temu, że $P_2P_1 = 0$. Mamy bowiem $(P_1P_2)^* = P_2P_1 = 0$.

Lemat 7.1. *Na to, aby rzuty P_1 i P_2 były ortogonalne, potrzeba i wystarcza, aby ortogonalne były podprzestrzenie $M_1 = \mathrm{im}\, P_1$ i $M_2 = \mathrm{im}\, P_2$.*

D o w ó d. Jeżeli $P_1P_2 = 0$, to dla $x_1 \in M_1$ i $x_2 \in M_2$ mamy

$$\langle x_1, x_2\rangle = \langle P_1x_1, P_2x_2\rangle = \langle x_1, P_1P_2x_2\rangle = \langle x_1, 0\rangle = 0.$$

Na odwrót, jeżeli $M_1 \perp M_2$, to ponieważ $P_2x \in M_2$ dla dowolnego $x \in H$, więc $P_1P_2x = 0$, czyli $P_1P_2 = 0$. ■

Twierdzenie 7.2. *Na to, aby suma rzutów P_1 i P_2 była rzutem, potrzeba i wystarcza, aby te operatory były ortogonalne. Wówczas* $\mathrm{im}(P_1 + P_2) = \mathrm{im}\, P_1 \oplus \mathrm{im}\, P_2$.

D o w ó d. Niech $P = P_1 + P_2$ będzie rzutem. Wówczas z równości $(P_1 + P_2)^2 = P_1 + P_2$ otrzymujemy $P_1P_2 + P_2P_1 = 0$. Mnożąc tę równość z lewej strony przez P_1, otrzymujemy $P_1P_2 + P_1P_2P_1 = 0$. Z kolei mnożąc otrzymaną równość z prawej strony przez P_1, mamy $P_1P_2P_1 = 0$. Stąd i z poprzedniej równości wynika, że $P_1P_2 = 0$.

Niech $P_1P_2 = P_2P_1 = 0$. Wówczas operator $P_1 + P_2$ jest samosprzężony i zachodzi równość $(P_1+P_2)^2 = P_1+P_2$, a więc $P = P_1+P_2$ jest rzutem (por. tw. 4.12). Z lematu 7.1 wynika, że podprzestrzenie $M_1 = \mathrm{im}\, P_1$ i $M_2 = \mathrm{im}\, P_2$ są ortogonalne.

Dla dowolnego wektora $x \in H$ mamy

$$Px = P_1x + P_2x = x_1 + x_2 \in M_1 \oplus M_2. \tag{7.1}$$

Na odwrót, jeżeli $x = x_1 + x_2 \in M_1 \oplus M_2$, to z równości $P_1x_2 = 0$ i $P_2x_1 = 0$ otrzymujemy

$$\begin{aligned} x &= x_1 + x_2 = P_1x_1 + P_2x_2 = \\ &= P_1(x_1 + x_2) + P_2(x_1 + x_2) = (P_1 + P_2)x = Px. \end{aligned} \tag{7.2}$$

Z (7.1) i (7.2) wynika, że $\mathrm{im}\, P = M_1 \oplus M_2$. ■

Twierdzenie 7.3. *Na to, aby iloczyn dwóch rzutów P_1 i P_2 był rzutem, potrzeba i wystarcza, aby $P_1P_2 = P_2P_1$. Wówczas* $\mathrm{im}(P_1P_2) = \mathrm{im}\, P_1 \cap \mathrm{im}\, P_2$.

D o w ó d. Niech $P = P_1P_2$ będzie rzutem. Wówczas jest on operatorem samosprzężonym, a więc mamy

$$P_1P_2 = (P_1P_2)^* = P_2^*P_1^* = P_2P_1.$$

Na odwrót, jeżeli $P_1P_2 = P_2P_1$, to P jest operatorem samosprzężonym, ponieważ

$$P^* = (P_1P_2)^* = (P_2P_1)^* = P_1P_2 = P.$$

Ponadto

$$P^2 = (P_1P_2)^2 = P_1P_2P_1P_2 = P_1^2P_2^2 = P_1P_2 = P.$$

Zatem P jest rzutem (por. tw. 4.12).

Niech $x \in H$ będzie dowolne. Wówczas $Px = P_1P_2x = P_2P_1x \in \operatorname{im} P_1 \cap \cap \operatorname{im} P_2$. Z kolei, jeżeli $y \in \operatorname{im} P_1 \cap \operatorname{im} P_2$, to $Py = P_1(P_2y) = P_1y = y$. Zatem $\operatorname{im} P = \operatorname{im} P_1 \cap \operatorname{im} P_2$. ■

Rzut P_2 jest *częścią rzutu* P_1, jeżeli $P_1P_2 = P_2$. Przechodząc do operatorów sprzężonych, można sprawdzić, że równość ta jest równoważna równości $P_2P_1 = P_2$.

Lemat 7.4. *Niech P_1 i P_2 będą rzutami na przestrzeni Hilberta H. Następujące warunki są równoważne*:

(a) *rzut P_2 jest częścią rzutu P_1*;

(b) $\|P_2x\| \leqslant \|P_1x\|$ *dla każdego $x \in H$*;

(c) $\operatorname{im} P_2 \subset \operatorname{im} P_1$.

D o w ó d. Udowodnimy najpierw, że warunek (a) implikuje (b). Ponieważ $P_2P_1 = P_2$, więc mamy $P_2P_1x = P_2x$ dla dowolnego $x \in H$. Stąd

$$\|P_2x\| = \|P_2P_1x\| \leqslant \|P_2\|\|P_1x\| \leqslant \|P_1x\|.$$

Teraz pokazujemy, że (c) jest konsekwencją (b). Jeżeli $x \in \operatorname{im} P_2$, to

$$\|x\| = \|P_2x\| \leqslant \|P_1x\| \leqslant \|x\|.$$

Stąd $\|P_1x\| = \|x\|$. Dalej mamy

$$\begin{aligned}\|(I - P_1)x\|^2 &= \langle x - P_1x, x - P_1x\rangle = \|x\|^2 - 2\langle P_1x, x\rangle + \|P_1x\|^2 = \\ &= \|x\|^2 - 2\|P_1x\|^2 + \|P_1x\|^2 = \|x\|^2 - \|P_1x\|^2 = 0.\end{aligned}$$

Zatem $x \in \ker(I - P_1) = \operatorname{im} P_1$.

W końcu pokażemy, że z warunku (c) wynika (a). Jeżeli $\operatorname{im} P_2 \subset \operatorname{im} P_1$, to dla dowolnego $x \in H$ mamy $P_2x \in \operatorname{im} P_1$. Zatem $P_1(P_2x) = P_2x$, skąd $P_1P_2 = P_2$. ■

Twierdzenie 7.5. *Różnica* $P_1 - P_2$ *dwóch rzutów jest rzutem wtedy i tylko wtedy, gdy rzut* P_2 *jest częścią* P_1*. Jeżeli ten warunek jest spełniony, to* $\operatorname{im}(P_1 - P_2)$ *jest dopełnieniem ortogonalnym* $\operatorname{im} P_2$ *w podprzestrzeni* $\operatorname{im} P_1$*.*

Dowód. Jeżeli operator $P_1 - P_2$ jest rzutem, to $I-(P_1-P_2) = (I-P_1)+P_2$ jest także rzutem. Na mocy twierdzenia 7.2 mamy $(I - P_1)P_2 = 0$, czyli $P_1P_2 = P_2$, a to oznacza, że P_2 jest częścią rzutu P_1.

Na odwrót, niech rzut P_2 będzie częścią rzutu P_1. Wówczas rzuty $I - P_1$ i P_2 są ortogonalne i na mocy twierdzenia 7.2 operator $(I - P_1) + P_2$ jest rzutem, a zatem i operator $P_1 - P_2$ jest rzutem.

Z warunku $P_1P_2 = P_2$ wynika, że rzuty $P_1 - P_2$ i P_2 są ortogonalne. Wówczas na mocy twierdzenia 7.2, $\operatorname{im} P_1 = \operatorname{im}(P_1 - P_2) \oplus \operatorname{im} P_2$. ∎

7.2. Twierdzenie spektralne w przestrzeni skończenie wymiarowej

Twierdzenie 7.6 (spektralne w przestrzeni skończenie wymiarowej). *Niech* H *będzie skończenie wymiarową przestrzenią Hilberta i niech* A *będzie operatorem normalnym na* H*. Istnieją wówczas liczby zespolone* $\alpha_1, \ldots, \alpha_r$ *i rzuty ortogonalne* $P_1, \ldots, P_r$ *(gdzie* $r \leqslant \dim H = n$*) takie, że*

(i) *liczby* $\alpha_1, \ldots, \alpha_r$ *są parami różne*;

(ii) *rzuty* P_j *są parami ortogonalne i różne od zera*;

(iii) $\sum_{j=1}^{r} P_j = I$;

(iv) $A = \sum_{j=1}^{r} \alpha_j P_j$.

Dowód. Zakładamy, że $H \neq \{0\}$, gdyż dla przestrzeni zerowej nie ma czego udowadniać. Operator A ma wartości własne, bo $\sigma(A) = \sigma_{\mathrm{p}}(A) \neq \emptyset$. Niech $\alpha_1, \ldots, \alpha_r$ będą wszystkimi parami różnymi wartościami własnymi operatora A. Niech $M_j = \ker(A - \alpha_j I)$, $j = 1, 2, \ldots, r$. Wówczas $M_j \perp M_k$ dla $j \neq k$, bo wektory własne odpowiadające różnym wartościom własnym są ortogonalne (por. tw. 5.19). Ponadto, gdyby $I - \sum_{j=1}^{r} P_j \neq 0$, to $\{0\} \neq$ $\neq H_0 = \operatorname{im}\Big(I - \sum_{j=1}^{r} P_j\Big) \subset H$. Wykażemy, że $A(H_0) \subset H_0$. Podprzestrzenie

$M_j = \ker(A - \alpha_j I)$ redukują operator A. Zatem $P_j A = AP_j$ i jeżeli $y \in \in \operatorname{im}\left(I - \sum_{j=1}^{r} P_j\right)$, to $y = x - \sum_{j=1}^{r} P_j x$ dla pewnego x oraz mamy

$$Ay = Ax - \sum_{j=1}^{r} AP_j x = Ax - \sum_{j=1}^{r} P_j Ax \in \operatorname{im}\left(I - \sum_{j=1}^{r} P_j\right).$$

Zatem $A\big|_{H_0}$ jest operatorem na H_0. Operator $A\big|_{H_0}$ nie miałby wartości własnych. Zauważmy bowiem, że $Ax = \lambda x$ dla $x \neq 0$ implikuje, że $\lambda = \alpha_j$ dla pewnego j. Wówczas $Ax = \alpha_j x$ i $x \in M_j$, czyli $x = P_j x$. Gdyby $x \in \operatorname{im}\left(I - \sum_{j=1}^{r} P_j\right)$, to istniałby wektor y taki, że $y - \sum_{j=1}^{r} P_j y = x$. Stąd

$$0 = P_j y - \sum_{k=1}^{r} P_j P_k y = x,$$

sprzeczność. Zatem

$$I = \sum_{j=1}^{r} P_j.$$

Mamy więc

$$x = \sum_{j=1}^{r} P_j x = \sum_{j=1}^{r} x_j.$$

Stąd

$$Ax = A\left(\sum_{j=1}^{r} x_j\right) = \sum_{j=1}^{r} Ax_j = \sum_{j=1}^{r} \alpha_j P_j x,$$

co kończy dowód. ■

U w a g i. 1. Jeżeli operator A jest samosprzężony, to liczby $\alpha_1, \ldots, \alpha_r$ są rzeczywiste na mocy twierdzenia 5.21.

2. Jeżeli w każdej przestrzeni $M_j = \ker(A - \alpha_j I)$ wybierzemy bazę ortonormalną, to zbiór tych wszystkich wektorów tworzy bazę przestrzeni H, w której macierz operatora A ma postać diagonalną

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ & \ddots & & & & & & \\ 0 & 0 & \alpha_1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \alpha_2 & 0 & 0 & \ldots & 0 \\ & & & & \ddots & & & \\ 0 & 0 & 0 & 0 & 0 & \alpha_2 & \ldots & 0 \\ & & & & & & \ddots & \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \alpha_r \end{pmatrix}.$$

Na odwrót, operator, który ma w pewnej bazie ortonormalnej przestrzeni H macierz tej postaci, jest operatorem normalnym.

7.3. Twierdzenie spektralne dla zwartych operatorów normalnych

Twierdzenie 7.7 (spektralne dla zwartych operatorów normalnych). *Niech A będzie zwartym operatorem normalnym na nieskończenie wymiarowej przestrzeni Hilberta H. Niech $\{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ będzie zbiorem wszystkich niezerowych wartości własnych operatora A oraz niech P_n będzie rzutem ortogonalnym na podprzestrzeń $\ker(A - \lambda_n I)$. Wówczas*:

(i) *Jeżeli zbiór $\{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ jest skończony, to*

$$A = \sum_n \lambda_n P_n \tag{7.3}$$

i A jest operatorem skończenie wymiarowym.

(ii) *W przeciwnym wypadku*

$$A = \sum_{n=1}^{\infty} \lambda_n P_n, \tag{7.4}$$

gdzie szereg ten jest zbieżny w normie operatorowej na przestrzeni $\mathcal{B}(H)$, a ciąg (λ_n) jest uporządkowany w następujący sposób:

$$|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3| \geqslant \ldots \qquad (\lambda_j \neq \lambda_k, \quad \textit{gdy} \quad j \neq k).$$

D o w ó d. W przypadku, gdy A jest operatorem zerowym, zbiór jego niezerowych wartości własnych jest pusty i twierdzenie jest oczywiste. Załóżmy, że $A \neq 0$. Z twierdzeń 5.19 i 4.16 wynika, że $\ker(A - \lambda_n I) \perp \ker(A - \lambda_m I)$ dla $n \neq m$ oraz $\ker(A - \lambda_n I)$ jest podprzestrzenią redukującą operator A. Na mocy wniosku 5.18 i twierdzenia Riesza o widmie operatora zwartego istnieje $\lambda_1 \in \sigma_{\mathrm{p}}(A)$ takie, że $|\lambda_1| = \|A\|$. Niech $M_1 = \ker(A - \lambda_1 I)$, $P_1 \colon H \to M_1$ niech będzie rzutem ortogonalnym na M_1 i niech $H_2 = M_1^{\perp}$. Ponieważ M_1 redukuje operator A, więc również H_2 ma tę samą własność.

Niech $A_2 = A\,\big|_{H_2}$. Jeżeli $A_2 = 0$, to A jest operatorem skończenie wymiarowym i $A = \lambda_1 P_1$. Wówczas dowód jest zakończony. Niech więc $A_2 \neq 0$. Na mocy wniosku 4.15 A_2 jest zwartym operatorem normalnym na przestrzeni H_2. Z wniosku 5.18 wynika istnienie wartości własnej λ_2 dla operatora A_2 takiej, że $|\lambda_2| = \|A_2\|$. Zauważmy, że λ_2 jest również wartością własną operatora A. Ponieważ $\|A_2\| \leqslant \|A\|$, więc $|\lambda_2| \leqslant |\lambda_1|$. Ponadto $\lambda_1 \neq \lambda_2$, bo w przeciwnym wypadku mielibyśmy

$$\{0\} \neq \ker(A_2 - \lambda_2 I) \subset \ker(A - \lambda_2 I) = \ker(A - \lambda_1 I) = M_1,$$

co nie jest możliwe, bo $\ker(A_2 - \lambda_2 I) \subset H_2 \perp M_1$. Niech $M_2 = \ker(A - \lambda_2 I)$. Ponieważ $M_2 \perp M_1$, więc $M_2 \subset H_2$ i w konsekwencji $M_2 = \ker(A_2 - \lambda_2 I)$.

Niech $P_2 \colon H \to M_2$ będzie rzutem ortogonalnym i niech $H_3 = (M_1 \oplus \oplus M_2)^{\perp}$ oraz niech $A_3 = A\,\big|_{H_3}$. Jeżeli $A_3 = 0$, to A jest operatorem skończenie wymiarowym i $A = \lambda_1 P_1 + \lambda_2 P_2$. Jeżeli natomiast $A_3 \neq 0$, to możemy kontynuować dowód w taki sam sposób jak w poprzednim kroku.

Za pomocą indukcji otrzymujemy taki ciąg (λ_n) wartości własnych operatora A, że:

(a) $|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n| \geqslant \ldots$;

(b) jeżeli $M_n = \ker(A - \lambda_n I)$, to $|\lambda_{n+1}| = \big\|A\,\big|_{(M_1 \oplus \ldots \oplus M_n)^{\perp}}\big\|$.

Jeżeli ten proces w którymś kroku się skończy, to operator A jest skończenie wymiarowy i suma w (7.3) jest skończona. Załóżmy więc, że otrzymaliśmy nieskończony ciąg (λ_n) wartości własnych operatora A i ciąg podprzestrzeni M_n spełniających warunki (a) oraz (b). Z warunku (a) wynika, że istnieje nieujemna liczba α taka, że $|\lambda_n| \to \alpha$. Twierdzimy, że $\alpha = 0$, czyli $\lambda_n \to 0$. Istotnie, jeżeli $e_n \in M_n$, $\|e_n\| = 1$, to ze zwartości operatora A wynika, że istnieją wektor $h \in H$ i podciąg (e_{n_k}) takie, że $\|Ae_{n_k} - h\| \to 0$. Ponieważ $e_n \perp e_m$ dla $n \neq m$ i $Ae_{n_k} = \lambda_{n_k} e_{n_k}$, więc

$$\|Ae_{n_k} - Ae_{n_j}\|^2 = |\lambda_{n_k}|^2 + |\lambda_{n_j}|^2 \geqslant 2\alpha^2.$$

Ponieważ ciąg (Ae_{n_k}) jest ciągiem Cauchy'ego, więc $\alpha = 0$.

Niech $P_n\colon H \to M_n$ będzie rzutem ortogonalnym dla $n = 1, 2, \dots$ Weźmy operator $A - \sum\limits_{j=1}^{n} \lambda_j P_j$. Jeżeli $x \in M_k$, $1 \leqslant k \leqslant n$, to

$$\Big(A - \sum_{j=1}^{n} \lambda_j P_j\Big)x = Ax - \lambda_k x = 0.$$

Zatem

$$M_1 \oplus \ldots \oplus M_n \subset \ker\Big(A - \sum_{j=1}^{n} \lambda_j P_j\Big).$$

Jeżeli $x \in (M_1 \oplus \ldots \oplus M_n)^{\perp}$, to $P_j x = 0$ dla $j = 1, 2, \dots, n$. Stąd $\Big(A - \sum\limits_{j=1}^{n} \lambda_j P_j\Big)x = Ax$. Ponieważ przestrzeń $(M_1 \oplus \ldots \oplus M_n)^{\perp}$ redukuje operator A, więc

$$\Big\|A - \sum_{j=1}^{n} \lambda_j P_j\Big\| = \big\|A\big|_{(M_1\oplus\ldots\oplus M_n)^{\perp}}\big\| = |\lambda_{n+1}| \to 0.$$

Zatem szereg $\sum\limits_{n=1}^{\infty} \lambda_n P_n$ jest zbieżny do A w normie operatorowej. ■

Wniosek 7.8. *Jeżeli A jest zwartym operatorem normalnym na przestrzeni Hilberta H, to istnieją ciąg (μ_n) liczb zespolonych i baza ortonormalna dla podprzestrzeni $(\ker A)^{\perp}$ takie, że dla dowolnego $x \in H$ zachodzi równość*

$$Ax = \sum_{n=1}^{\infty} \mu_n \langle x, e_n \rangle e_n.$$

D o w ó d. Jeżeli A jest operatorem zerowym, to twierdzenie jest oczywiste. Zakładamy więc, że $A \neq 0$. Niech (λ_n) będzie ciągiem wszystkich niezerowych wartości własnych operatora A i niech P_n będzie rzutem ortogonalnym na podprzestrzeń $M_n = \ker(A - \lambda_n I)$. Zauważmy, że

$$\ker A = \Big(\overline{\operatorname{span}\{M_n\colon\ n = 1, 2 \dots\}}\Big)^{\perp} = (\operatorname{im} A)^{\perp}.$$

Istotnie, mamy $\ker A = \ker A^* = (\operatorname{im} A)^{\perp}$. Ponieważ $M_n \perp M_k$ dla $n \neq k$, więc dla dowolnego $x \in H$ zachodzi równość

$$\|Ax\|^2 = \sum_{n=1}^{\infty} \|\lambda_n P_n x\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \|P_n x\|^2.$$

Stąd $Ax = 0$ wtedy i tylko wtedy, gdy $P_n x = 0$ dla każdego n, czyli $x \in \ker A$ wtedy i tylko wtedy, gdy $x \perp M_n$ dla każdego n. Zatem $\ker A = \left(\overline{\operatorname{span}\{M_n\colon\, n = 1, 2 \dots\}}\right)^{\perp}$. Niech $L = \overline{\operatorname{im} A}$. Wówczas $L = \overline{\operatorname{span}\{M_n\colon\, n = 1, 2 \dots\}}$.

Przestrzenie L i $\ker A$ są niezmiennicze dla operatora A, a ponieważ $\ker A = L^{\perp}$, więc podprzestrzeń L redukuje ten operator. Możemy rozpatrzyć obcięcie operatora A do podprzestrzeni L, $A\big|_L$. Niech $\{e_j^{(n)}\colon\, j = 1, 2, \dots, N_n\}$ będzie bazą ortonormalną dla M_n. Zatem $Ae_j^{(n)} = \lambda_n e_j^{(n)}$ dla $j = 1, 2, \dots, N_n$. Zbiór $\{e_j^{n)}\colon\, j = 1, 2, \dots, N_n, n = 1, 2, \dots\}$ jest bazą ortonormalną dla przestrzeni L (por. tw. 2.20) i dla $x \in L$ mamy

$$x = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \langle x, e_j^{(n)} \rangle e_j^{(n)}.$$

Stąd

$$Ax = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \langle x, e_j^{(n)} \rangle A e_j^{(n)} = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \lambda_n \langle x, e_j^{(n)} \rangle e_j^{(n)}.$$

Na zakończenie zauważmy, że jeżeli ciąg (λ_n) jest skończony, to operator A jest skończenie wymiarowy i wszystkie powyższe sumy są skończone. ■

Wniosek 7.9. *Jeżeli na przestrzeni Hilberta H istnieje zwarty operator normalny A taki, że* $\ker A = \{0\}$, *to H jest przestrzenią ośrodkową.*

7.4. Operatory dodatnie

Definicja 7.1. Niech H będzie zespoloną przestrzenią Hilberta. Operator $A \in \mathcal{B}(H)$ nazywamy *dodatnim*, jeżeli $\langle Ax, x \rangle \geqslant 0$ dla każdego $x \in H$ (dokładniej należałoby mówić o *operatorze nieujemnym*).

Uwagi. 1. Zauważmy, że na to, aby $\langle Ax, x \rangle \geqslant 0$ dla każdego $x \in H$ potrzeba, aby iloczyn skalarny $\langle Ax, x \rangle$ był dla każdego $x \in H$ liczbą rzeczywistą, a to ma miejsce wtedy i tylko wtedy, gdy operator A jest samosprzężony. Zatem operatory dodatnie są zawsze samosprzężone.

2. Operator identycznościowy I jest dodatni.

3. Dla dowolnego operatora $A \in \mathcal{B}(H)$ operatory A^*A i AA^* są dodatnie.

4. Jeżeli A jest operatorem samosprzężonym, to A^2 jest operatorem dodatnim. Stąd wynika, że dowolny rzut jest dodatni.

Będziemy dalej pisać $A \geqslant 0$ w przypadku, gdy A jest operatorem dodatnim, oraz $A \geqslant B$, jeżeli A i B są operatorami samosprzężonymi oraz $A - B \geqslant 0$. Relacja $A \geqslant B$ ma następujące własności:

(i) $A \geqslant A$;

(ii) jeżeli $A \geqslant B$ i $B \geqslant A$, to $A = B$;

(iii) jeżeli $A \geqslant B$ i $B \geqslant C$, to $A \geqslant C$;

(iv) jeżeli $A \geqslant B$ i $C \geqslant D$, to $A + C \geqslant B + D$;

(v) jeżeli $A \geqslant 0$ i $\lambda \geqslant 0$, to $\lambda A \geqslant 0$;

(vi) jeżeli $A \geqslant 0$ i A^{-1} istnieje (oraz jest operatorem ograniczonym), to $A^{-1} \geqslant 0$.

Dowody tych własności są natychmiastowe. Pozostawiamy je jako ćwiczenie dla Czytelnika.

U w a g i. 1. Operatory samosprzężone tworzą przestrzeń liniową nad ciałem liczb rzeczywistych. Będziemy ją oznaczać symbolem $\mathcal{B}_{\mathrm{sa}}(H)$. Jest to podprzestrzeń przestrzeni wszystkich operatorów liniowych ograniczonych na H z mnożeniem przez skalar zawężonym do ciała liczb rzeczywistych. Z własności (i)–(iii) wynika, że relacja $A \geqslant B$ lub, co jest równoważne, $B \leqslant A$ jest relacją częściowo porządkującą w zbiorze operatorów samosprzężonych na przestrzeni Hilberta H. Jeżeli $\dim H > 1$, to nie jest ona relacją liniowo porządkującą, bo zawsze znajdą się elementy nieporównywalne. Na przykład dla $H = \mathbb{C}^2$ operatory odpowiadające macierzom

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

są nieporównywalne.

2. Jeżeli A jest operatorem samosprzężonym, to ze wzoru (por. tw. 4.7)

$$\|A\| = \sup\{|\langle Ax, x\rangle| \colon \ \|x\| \leqslant 1\} \tag{7.5}$$

wynika, że

$$-\|A\|I \leqslant A \leqslant \|A\|I. \tag{7.6}$$

Dla $x \neq 0$ mamy bowiem

$$\begin{aligned}\langle(\|A\|I - A)x, x\rangle &= \|A\|\langle x, x\rangle - \langle Ax, x\rangle = \\ &= \|A\|\|x\|^2 - \|x\|^2 \left\langle A\left(\frac{1}{\|x\|}x\right), A\left(\frac{1}{\|x\|}x\right)\right\rangle \geqslant \\ &\geqslant \|A\|\|x\|^2 - \|x\|^2\|A\| = 0.\end{aligned}$$

Dla $x = 0$ nierówność $\langle (\|A\|I - A)x, x\rangle \geqslant 0$ jest oczywista. Zatem $\|A\|I - A \geqslant 0$ i w konsekwencji $A \leqslant \|A\|I$. Drugą z nierówności we wzorze (7.6) otrzymujemy w podobny sposób.

3. Jeżeli $m = \inf\{\langle Ax, x\rangle\colon \|x\| = 1\}$ i $M = \sup\{\langle Ax, x\rangle\colon \|x\| = 1\}$, to dla dowolnego $x \in H$ mamy

$$m\langle x, x\rangle \leqslant \langle Ax, x\rangle \leqslant M\langle x, x\rangle$$

i w konsekwencji

$$mI \leqslant A \leqslant MI.$$

Ze wzoru (7.5) otrzymujemy

$$\|A\| = \max\{|m|, |M|\}.$$

Ponadto jeżeli A jest operatorem dodatnim, to $0 \leqslant m \leqslant M$ i $\|A\| = M$.

Twierdzenie 7.10. *Operator samosprzężony A jest dodatni wtedy i tylko wtedy, gdy $\sigma(A) \subset [0, +\infty)$.*

D o w ó d. Istotnie, jeżeli A jest operatorem dodatnim, to $m \geqslant 0$, i z twierdzenia 5.22 wynika, że $\sigma(A) \subset [m, M] \subset [0, +\infty)$.

Na odwrót, jeżeli $\sigma(A) \subset [0, +\infty)$, to ponieważ $m \in \sigma(A)$, na mocy tego samego twierdzenia co poprzednio, więc $m \geqslant 0$ i operator A jest dodatni. ■

Lemat 7.11 (uogólniona nierówność Schwarza). *Dla dowolnego operatora dodatniego A na przestrzeni Hilberta H i dowolnych wektorów $x, y \in H$ zachodzi nierówność*

$$|\langle Ax, y\rangle|^2 \leqslant \langle Ax, x\rangle\langle Ay, y\rangle. \tag{7.7}$$

Nierówność (7.7) nosi nazwę *uogólnionej nierówności Schwarza*.

D o w ó d. Niech $z_\lambda = x + \lambda\langle Ax, y\rangle y$, gdzie λ jest dowolną liczbą rzeczywistą. Wówczas

$$0 \leqslant \langle Az_\lambda, z_\lambda\rangle = \langle Ax, x\rangle + 2\lambda|\langle Ax, y\rangle|^2 + \lambda^2|\langle Ax, y\rangle|^2\langle Ay, y\rangle.$$

Stąd

$$4|\langle Ax, y\rangle|^4 - 4|\langle Ax, y\rangle|^2\langle Ax, x\rangle\langle Ay, y\rangle \leqslant 0$$

i w konsekwencji otrzymujemy nierówność (7.7). ■

Teraz wykażemy twierdzenie, które jest odpowiednikiem twierdzenia o zbieżności ciągów monotonicznych z analizy matematycznej.

Twierdzenie 7.12. (a) *Jeżeli* (A_n) *jest niemalejącym ciągiem operatorów samosprzężonych ograniczonym z góry, to dla dowolnego wektora* $x \in H$ *istnieje* $\lim_{n\to\infty} A_n x = Ax$, A *jest operatorem samosprzężonym oraz* A *jest przemienne z każdym operatorem* B, *który jest przemienny z wszystkimi operatorami* A_n.

(b) *Analogiczne do* (a) *tylko dla nierosnącego ciągu operatorów* (A_n).

D o w ó d. Udowodnimy część (a). Dowód części (b) jest analogiczny. Wystarczy rozważyć przypadek, gdy

$$0 \leqslant A_1 \leqslant A_2 \leqslant \ldots \leqslant I.$$

Niech $m < n$. Wówczas $A_n - A_m \geqslant 0$ i na mocy uogólnionej nierówności Schwarza mamy

$$\begin{aligned}\|(A_n - A_m)x\|^4 &\leqslant \langle (A_n - A_m)x, (A_n - A_m)x\rangle^2 \leqslant \\ &\leqslant \langle (A_n - A_m)x, x\rangle \langle (A_n - A_m)^2 x, (A_n - A_m)x\rangle.\end{aligned}$$

Ponieważ $0 \leqslant A_n - A_m \leqslant I$, więc $\|A_n - A_m\| \leqslant 1$ i mamy

$$\|(A_n - A_m)x\|^4 \leqslant (\langle A_n x, x\rangle - \langle A_m x, x\rangle)\|x\|^2.$$

Ponieważ ciąg liczbowy $(\langle A_n x, x\rangle)$ jest niemalejący i ograniczony, więc jest on zbieżny. Z otrzymanej nierówności wynika, że ciąg wektorów $(A_n x)$ jest również zbieżny. Na mocy twierdzenia 3.3 operator A zdefiniowany równością $Ax = \lim_{n\to\infty} A_n x$ jest liniowy i ograniczony. Z ciągłości iloczynu skalarnego wynika, że A jest operatorem samosprzężonym. Natomiast jeżeli $A_n B = BA_n$ dla operatora B i wszystkich n, to

$$ABx = \lim_{n\to\infty} A_n Bx = \lim_{n\to\infty} BA_n x = BAx. \qquad \blacksquare$$

7.5. Rachunek funkcyjny

W dalszym ciągu A będzie ustalonym samosprzężonym operatorem na przestrzeni Hilberta H. Niech $m = \inf\{\langle Ax, x\rangle \colon \|x\| = 1\}$ oraz $M = \sup\{\langle Ax, x\rangle \colon \|x\| = 1\}$. Wówczas $mI \leqslant A \leqslant MI$. Zauważmy, że gdy $m = M$, to operator A jest równy mI. Aby nie rozważać tego trywialnego przypadku, będziemy zakładać, że $m < M$. Rozważymy przestrzeń liniową $\mathcal{P}$ wszystkich funkcji rzeczywistych określonych na przedziale $[m, M]$ postaci

$$p(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_n\lambda^n, \tag{7.8}$$

gdzie n jest nieujemną liczbą całkowitą, a α_0, α_1, $\ldots$, α_n są liczbami rzeczywistymi. Tak więc funkcje z przestrzeni $\mathcal{P}$ są wielomianami o współczynnikach rzeczywistych. Ponieważ przedział $[m, M]$ jest właściwy, każda funkcja $p \in \mathcal{P}$ jest wyznaczona jednoznacznie poprzez reprezentację (7.8). Dla $p \in \mathcal{P}$ definiujemy operator $p(A)$ za pomocą wzoru

$$p(A) = \alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \ldots + \alpha_n A^n. \tag{7.9}$$

Oczywiście operator $p(A)$ jest samosprzężony. Zatem przyporządkowanie $p \mapsto p(A)$ odwzorowuje przestrzeń $\mathcal{P}$ w przestrzeń $\mathcal{B}_{\mathrm{sa}}(H)$. Jest oczywiste, że ponadto

$$(pq)(A) = p(A)q(A)$$

dla dowolnych $p, q \in \mathcal{P}$. Odwzorowanie to zachowuje porządek, a mianowicie prawdziwy jest następujący lemat.

Lemat 7.13. *Jeżeli $p, q \in \mathcal{P}$ są takie, że $p(\lambda) \geqslant q(\lambda)$ dla $\lambda \in [m, M]$, to $p(A) \geqslant q(A)$.*

Dowód. Wystarczy wykazać, że jeżeli $p(\lambda) \geqslant 0$ dla $\lambda \in [m, M]$, to $p(A) \geqslant 0$. Na mocy twierdzenia 7.10 wystarczy udowodnić, że $\sigma(p(A)) \subset [0, +\infty)$. Z twierdzenia o odwzorowaniu spektralnym (tw. 5.10) wiemy, że $\sigma(p(A)) = p(\sigma(A))$. Ponieważ $\sigma(A) \subset [m, M]$ (patrz tw. 5.22), więc $p(\sigma(A)) \subset [0, +\infty)$. ■

Zbiór wszystkich funkcji dodatnich z przestrzeni $\mathcal{P}$, czyli takich p, że $p(\lambda) \geqslant 0$ dla wszystkich $\lambda \in [m, M]$, oznaczymy symbolem $\mathcal{P}^+$. Zatem lemat 7.13 mówi, że jeżeli $p \in \mathcal{P}^+$, to $p(A)$ jest operatorem dodatnim.

Definicja 7.2. Oznaczmy przez $\mathcal{C}_1$ zbiór wszystkich funkcji rzeczywistych f określonych na przedziale $[m, M]$ takich, że istnieje ciąg (p_n) funkcji z przestrzeni $\mathcal{P}^+$ mający następujące własności:

(a) $0 \leqslant p_{n+1}(\lambda) \leqslant p_n(\lambda)$ dla $n = 1, 2, \ldots$ i $\lambda \in [m, M]$;

(b) $\lim\limits_{n\to\infty} p_n(\lambda) = f(\lambda)$ dla $\lambda \in [m, M]$.

Jest oczywiste, że jeżeli $f, g \in \mathcal{C}_1$ i $\alpha \geqslant 0$, to również $f + g \in \mathcal{C}_1$ oraz $\alpha f \in \mathcal{C}_1$. Ponadto jest jasne, że funkcje należące do zbioru $\mathcal{C}_1$ są ograniczone na przedziale $[m, M]$. Zbiór wszystkich rzeczywistych funkcji ograniczonych na przedziale $[m, M]$, które dadzą się przedstawić w postaci różnicy $f - g$ funkcji f i g ze zbioru $\mathcal{C}_1$, oznaczymy symbolem $\mathcal{C}_2$. Jest to podprzestrzeń liniowa przestrzeni wszystkich funkcji rzeczywistych, które są ograniczone na przedziale $[m, M]$. Zauważmy, że $\mathcal{P} \subset \mathcal{C}_2$, bo dowolny wielomian $p \in \mathcal{P}$

można zapisać w postaci $p = (p+\alpha 1)-\alpha 1$, gdzie α jest taką liczbą dodatnią, że $p + \alpha 1 \geqslant 0$.

Przestrzeń $\mathcal{C}_2$ jest dużo większa, o czym mówi następujące twierdzenie:

Twierdzenie 7.14. *Przestrzeń $\mathcal{C}_{\mathbb{R}}[m, M]$ wszystkich rzeczywistych funkcji ciągłych na przedziale $[a, b]$ jest podprzestrzenią przestrzeni $\mathcal{C}_2$.*

Dowód. Niech $f \in \mathcal{C}_{\mathbb{R}}[m, M]$. Wówczas $f = f^+ - f^-$, gdzie $f^+ = \max\{f, 0\}$ i $f^- = -\min\{f, 0\}$. Ponieważ $f^+ \geqslant 0$ i $f^- \geqslant 0$ oraz funkcje te są również ciągłe na przedziale $[m, M]$, więc możemy założyć, że $f \geqslant 0$. Na mocy twierdzenia aproksymacyjnego Weierstrassa (patrz [27], wn. 7.29 z części II) istnieje ciąg wielomianów (p_n) taki, że

$$\left| p_n(\lambda) - \left(f(\lambda) + \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right)\right)\right| < \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1}\right)$$

dla wszystkich $\lambda \in [m, M]$ i $n = 1, 2, \ldots$ Stąd

$$f(\lambda) + \frac{1}{n+1} < p_n(\lambda) < f(\lambda) + \frac{1}{n}$$

dla wszystkich $\lambda \in [m, M]$. Jest oczywiste, że $p_n \in \mathcal{P}^+$, $p_{n+1} \leqslant p_n$ dla $n = 1, 2, \ldots$ i $\lim\limits_{n\to\infty} p_n(\lambda) = f(\lambda)$ dla $\lambda \in [m, M]$. Zatem $f \in \mathcal{C}_1$. ∎

Naszym celem jest rozszerzenie odwzorowania $p \mapsto p(A)$ z przestrzeni $\mathcal{P}$ na przestrzeń $\mathcal{C}_2$, czyli chcemy zdefiniować operator samosprzężony $f(A)$ odpowiadający funkcji $f \in \mathcal{C}_2$. Ponadto chcemy to zrobić tak, aby zachować algebraiczne i porządkowe własności tego przekształcenia.

Lemat 7.15. *Jeżeli (p_n) jest ciągiem wielomianów z $\mathcal{P}^+$ takim, że $p_{n+1} \leqslant p_n$ dla $n = 1, 2, \ldots$, to ciąg operatorów $p_n(A)$ dąży punktowo do operatora dodatniego.*

Dowód. Z lematu 7.13 wynika, że $0 \leqslant p_{n+1}(A) \leqslant p_n(A)$ dla $n = 1, 2, \ldots$ Zatem na mocy twierdzenia 7.12 (b) ciąg $(p_n(A))$ jest zbieżny punktowo do operatora samosprzężonego B. Zauważmy, że operator B jest dodatni, bo mamy

$$\langle Bx, x\rangle = \left\langle \lim_{n\to\infty} p_n(A)x, x\right\rangle = \lim_{n\to\infty} \langle p_n(A)x, x\rangle \geqslant 0. \qquad ∎$$

Lemat 7.16. *Niech (p_n) i (q_n) będą ciągami w $\mathcal{P}^+$ takimi, że $p_{n+1} \leqslant p_n$ i $q_{n+1} \leqslant q_n$ dla $n = 1, 2, \ldots$ i niech operatory B i C będą granicami punktowymi ciągów $(p_n(A))$ i odpowiednio $(q_n(A))$. Wówczas, jeżeli $\lim\limits_{n\to\infty} p_n(\lambda) \leqslant \lim\limits_{n\to\infty} q_n(\lambda)$ dla $\lambda \in [m, M]$, to $B \leqslant C$.*

D o w ó d. Ponieważ ciągi liczbowe $(p_n(\lambda))$, $(q_n(\lambda))$ są nierosnące i ograniczone z dołu, więc są one zbieżne dla $\lambda \in [m, M]$. Niech k będzie ustaloną liczbą naturalną. Wówczas dla $\lambda \in [m, M]$ mamy

$$\lim_{n\to\infty}(p_n(\lambda) - q_k(\lambda)) \leqslant \lim_{n\to\infty} p_n(\lambda) - \lim_{n\to\infty} q_n(\lambda) \leqslant 0. \tag{7.10}$$

Niech $r_n = \max\{p_n - q_k, 0\}$ dla $n = 1, 2, \ldots$ Z nierówności (7.10) wynika, że $\lim\limits_{n\to\infty} r_n(\lambda) = 0$ dla każdego $\lambda \in [m, M]$. Ponadto z monotoniczności ciągu (p_n) wynika, że $r_{n+1} \leqslant r_n$ dla $n = 1, 2, \ldots$ Z twierdzenia Diniego ([22], tw. 7.13) (które mówi, że *ciąg monotoniczny funkcji ciągłych zbieżny punktowo na przedziale domkniętym do funkcji ciągłej jest na tym przedziale jednostajnie zbieżny*), otrzymujemy jednostajną zbieżność do zera ciągu (r_n) na przedziale $[m, M]$. Zatem dla dowolnego $\varepsilon > 0$ istnieje wskaźnik n_0 taki, że $0 \leqslant r_n(\lambda) < \varepsilon$ dla każdego $\lambda \in [m, M]$ i wszystkich $n \geqslant n_0$. Stąd otrzymujemy $p_n(\lambda) - q_k(\lambda) < \varepsilon$ dla wszystkich $\lambda \in [m, M]$ i $n \geqslant n_0$. Na mocy lematu 7.13 mamy $p_n(A) \leqslant q_k(A) + \varepsilon I$ dla wszystkich $n \geqslant n_0$. Stąd

$$\langle p_n(A)x, x\rangle \leqslant \langle q_k(A)x, x\rangle + \varepsilon\|x\|^2. \tag{7.11}$$

Ponieważ

$$\langle Bx, x\rangle = \Big\langle \lim_{n\to\infty} p_n(A)x, x\Big\rangle = \lim_{n\to\infty}\langle p_n(A)x, x\rangle,$$

więc przechodząc w nierówności (7.11) do granicy, $n \to \infty$, otrzymujemy $B \leqslant q_k(A) + \varepsilon I$. Ponieważ k było dowolną liczbą naturalną, więc przechodząc ponownie do granicy, $k \to \infty$, otrzymujemy $B \leqslant C + \varepsilon I$. Z dowolności $\varepsilon > 0$ wynika, że $B \leqslant C$. ■

Definicja 7.3. Niech $f \in \mathcal{C}_1$ i niech (p_n) będzie dowolnym ciągiem wielomianów z $\mathcal{P}^+$, który ma następujące własności: $p_{n+1} \leqslant p_n$ dla $n = 1, 2, \ldots$ i $\lim\limits_{n\to\infty} p_n(\lambda) = f(\lambda)$ dla $\lambda \in [m, M]$. Na mocy lematu 7.15 ciąg $(p_n(A))$ jest punktowo zbieżny do operatora dodatniego, który oznaczymy przez $f(A)$.

Z lematu 7.16 wynika, że ta definicja jest poprawna, a więc tak zdefiniowany operator $f(A)$ nie zależy od wyboru ciągu wielomianów (p_n). Istotnie, niech (q_n) będzie innym ciągiem wielomianów dodatnich takim, że $q_{n+1} \leqslant q_n$ dla $n = 1, 2 \ldots$ i $\lim\limits_{n\to\infty} q_n(\lambda) = f(\lambda)$ dla $\lambda \in [m, M]$. Niech B będzie punktową granicą ciągu operatorów $(q_n(A))$. Z lematu 7.16 wynika, że $B \leqslant f(A)$ i $f(A) \leqslant B$. Stąd $B = f(A)$.

Ponadto zauważmy, że jeżeli $p \in \mathcal{P}^+$, to operator $p(A)$ określony powyżej jest identyczny z operatorem zdefiniowanym za pomocą wzoru (7.9). Aby się o tym przekonać, wystarczy w definicji 7.3 wziąć ciąg $p_n = p$ dla $n = 1, 2, \ldots$

Lemat 7.17. *Odwzorowanie* $f \mapsto f(A)$ *ze zbioru* $\mathcal{C}_1$ *w zbiór operatorów dodatnich ma następujące własności*:

(a) $(f+g)(A) = f(A) + g(A)$ *dla* $f, g \in \mathcal{C}_1$;

(b) $(\alpha f)(A) = \alpha f(A)$ *dla* $f \in \mathcal{C}_1$, $\alpha \geqslant 0$;

(c) $(fg)(A) = f(A)g(A)$ *dla* $f, g \in \mathcal{C}_1$;

(d) *jeżeli* $f, g \in \mathcal{C}_1$ *i* $f \leqslant g$, *to* $f(A) \leqslant g(A)$.

Dowód. Własności (a) i (b) są oczywiste, a własność (d) wynika z lematu 7.16. Aby udowodnić (c), wybierzmy ciągi (p_n) i (q_n) ze zbioru $\mathcal{P}^+$ spełniające warunki: $p_{n+1} \leqslant p_n$ i $q_{n+1} \leqslant q_n$ dla $n = 1, 2, \dots$ oraz $\lim_{n\to\infty} p_n(\lambda) = f(\lambda)$ i $\lim_{n\to\infty} q_n(\lambda) = g(\lambda)$ dla $\lambda \in [m, M]$. Mamy $p_{m+1}q_{n+1} \leqslant p_n q_n$ oraz $\lim_{n\to\infty}(p_n q_n)(\lambda) = (fg)(\lambda)$ dla $\lambda \in [m, M]$. Zatem operator $(fg)(A)$ jest granicą punktową ciągu operatorów $((p_n q_n)(A))$. Ponieważ operatory $f(A)$ i $p_n(A)$ są samosprzężone, więc dla dowolnego $x \in H$ mamy

$$\begin{aligned}
\langle (f(A)g(A))x, x\rangle &= \langle f(A)(g(A)x), x\rangle = \langle g(A)x, f(A)x\rangle = \\
&= \lim_{n\to\infty} \langle q_n(A)x, p_n(A)x\rangle = \lim_{n\to\infty} \langle p_n(A)(q_n(A)x), x\rangle = \\
&= \lim_{n\to\infty} \langle (p_n(A)q_n(A))x, x\rangle = \lim_{n\to\infty} \langle (p_n q_n)(A)x, x\rangle = \\
&= \langle (fg)(A)x, x\rangle.
\end{aligned}$$

Z wniosku 4.2 otrzymujemy równość $f(A)g(A) = (fg)(A)$. ■

Teraz rozszerzymy odwzorowanie $f \mapsto f(A)$ na przestrzeń $\mathcal{C}_2$.

Definicja 7.4. Niech $f \in \mathcal{C}_2$. Bierzemy $g, h \in \mathcal{C}_1$ takie, że $f = g - h$ i definiujemy operator $f(A)$ za pomocą wzoru $f(A) = g(A) - h(A)$. Oczywiście $f(A)$ jest operatorem samosprzężonym. Zauważmy, że jego definicja nie zależy od wyboru funkcji g i h. Istotnie, jeżeli $f = g_1 - h_1$, gdzie $f_1, g_1 \in \mathcal{C}_1$, to $g + h_1 = h + g_1$. Wówczas z lematu 7.17 mamy

$$g(A) + h_1(A) = (g + h_1)(A) = (h + g_1)(A) = h(A) + g_1(A)$$

i w konsekwencji $g(A) - h(A) = g_1(A) - h_1(A)$. Jest również oczywiste, że jeżeli $f \in \mathcal{C}_1$, to definicje 7.3 i 7.4 operatora $f(A)$ się pokrywają.

Twierdzenie 7.18. *Odwzorowanie* $f \mapsto f(A)$ *z przestrzeni* $\mathcal{C}_2$ *w przestrzeń* $\mathcal{B}_{\mathrm{sa}}(H)$ *określone w definicji 7.4 ma następujące własności*:

(a) *jest liniowe*;

(b) $(fg)(A) = f(A)g(A)$ *dla* $f, g \in \mathcal{C}_2$;

(c) *jeżeli* $f, g \in \mathcal{C}_2$ *i* $f \leqslant g$, *to* $f(A) \leqslant g(A)$.

Dowód. Liniowość przekształcenia $f \mapsto f(A)$ wynika z części (a) i (b) lematu 7.17. Część (c) wynika z własności relacji porządku w zbiorze operatorów dodatnich. Udowodnimy część (b) lematu. Niech $f = s - t$ i $g = u - v$, gdzie $s, t, u, v \in \mathcal{C}_1$. Wówczas $su + tv, sv + tu \in \mathcal{C}_1$ i $fg = su + tv - (sv + tu)$. Zatem $(fg)(A) = (su + tv)(A) - (sv + tu)(A)$ i mamy

$$\begin{aligned} f(A)g(A) &= (s(A) - t(A))(u(A) - v(A)) = \\ &= s(A)u(A) + t(A)v(A) - (s(A)v(A) + t(A)u(A)) = \\ &= (su + tv)(A) - (sv + tu)(A) = (fg)(A). \qquad \blacksquare \end{aligned}$$

Uwagi. 1. Odwzorowanie $f \mapsto f(A)$ jest nazywane *rachunkiem funkcyjnym*. Pozwala ono rozważać funkcje od operatora samosprzężonego A odpowiadające ciągłym funkcjom rzeczywistym określonym na przedziale $[m, M]$.

Zauważmy, że można je rozszerzyć na funkcje ciągłe przyjmujące wartości zespolone. Istotnie, jeżeli $f = u + iv$ jest rozkładem funkcji f na części rzeczywistą i urojoną, to ponieważ są one również funkcjami ciągłymi, istnieją operatory $u(A)$ i $v(A)$. Możemy zatem zdefiniować operator $f(A)$ za pomocą wzoru $f(A) = u(A) + iv(A)$. Oczywiście ten operator nie jest już samosprzężony, ale ponieważ operatory $u(A)$ i $v(A)$ są przemienne, więc $f(A)$ jest operatorem normalnym (por. tw. 4.9).

2. Definicja operatora $f(A)$ jest niekonstruktywna, dlatego trudno ją wykorzystać do praktycznego wyznaczenia tego operatora. W następnym podrozdziale przedstawimy reprezentację całkową operatora $f(A)$, która jest pod tym względem dużo wygodniejsza.

Definicja 7.5. Operator samosprzężony B jest *pierwiastkiem kwadratowym* z operatora dodatniego A, jeżeli $A = B^2$.

Ponieważ funkcja $\lambda \mapsto \sqrt{\lambda}$ jest ciągła na przedziale $[0, +\infty)$, więc z twierdzenia 7.14 i definicji 7.4 wynika, że dla każdego operatora dodatniego A istnieje dokładnie jeden dodatni pierwiastek kwadratowy. Będziemy go oznaczać symbolem $\sqrt{A}$. Na zakończenie tego podrozdziału przedstawimy bezpośredni dowód prawdziwości tego stwierdzenia.

Twierdzenie 7.19. *Istnieje dokładnie jeden dodatni pierwiastek kwadratowy B z operatora dodatniego A. Ponadto jest on przemienny z każdym operatorem przemiennym z operatorem A.*

Dowód. Najpierw udowodnimy istnienie operatora o żądanych własnościach. Z nierówności (7.6) wynika, że bez zmniejszenia ogólności rozważań możemy ograniczyć się do przypadku, gdy $A \leqslant I$. Tworzymy następujący ciąg operatorów:

$$B_0 = 0, \qquad B_{n+1} = B_n + \frac{1}{2}\left(A - B_n^2\right) \quad \text{dla} \;\; n = 0, 1, 2, \ldots \tag{7.12}$$

Na mocy łatwej indukcji otrzymujemy, że operatory B_n są samosprzężone dla dowolnego n i przemienne z dowolnym operatorem, który jest przemienny z operatorem A. W szczególności $B_nB_m = B_mB_n$. Prawdziwe są następujące równości:

$$I - B_{n+1} = \frac{1}{2}\left(I - B_n\right)^2 + \frac{1}{2}\left(I - A\right), \tag{7.13}$$

$$B_{n+1} - B_n = \frac{1}{2}\left(\left(I - B_{n-1}\right) + \left(I - B_n\right)\right)\left(B_n - B_{n-1}\right). \tag{7.14}$$

Z (7.13) wynika, że $B_n \leqslant I$ dla dowolnego n. Natomiast z (7.12) i (7.14), poprzez indukcję, wynika, że $B_{n+1} - B_n \geqslant 0$. Zatem $B_n \leqslant B_{n+1}$ dla $n = 0, 1, 2, \ldots$ Stąd i z (7.12) wynika, że operatory B_n są dodatnie. Ciąg (B_n) jest niemalejącym ciągiem operatorów dodatnich, który jest ograniczony z góry. Na mocy twierdzenia 7.12 (a) jest on zbieżny do pewnego operatora samosprzężonego B. Operator B jest oczywiście dodatni. Po przejściu do granicy w równości (7.12) otrzymujemy

$$B = B + \frac{1}{2}\left(A - B^2\right),$$

skąd wynika, że $B^2 = A$. Ponieważ każdy z operatorów B_n jest przemienny z dowolnym operatorem przemiennym z A, więc taką samą własność ma operator B.

Teraz wykażemy, że dodatni pierwiastek kwadratowy z operatora dodatniego jest wyznaczony jednoznacznie. Niech C będzie operatorem dodatnim, który spełnia równość $C^2 = A$. Wówczas $CA = CC^2 = C^2C = AC$, a zatem $BC = CB$. Wybierzmy dowolny wektor $x \in H$ i niech $y = (B - C)x$. Wówczas

$$\begin{aligned}\langle By, y\rangle + \langle Cy, y\rangle &= \langle (B + C)y, y\rangle = \langle (B + C)(B - C)x, y\rangle = \\ &= \langle (B^2 - C^2)x, y\rangle = 0.\end{aligned}$$

Ponieważ B i C są operatorami dodatnimi, więc stąd wynika, że $\langle By, y\rangle = \langle Cy, y\rangle = 0$. Z pierwszej części dowodu wynika, że $B = D^2$ dla pewnego operatora dodatniego D. Zatem

$$\|Dy\|^2 = \langle Dy, Dy\rangle = \langle D^2y, y\rangle = \langle By, y\rangle = 0.$$

Stąd $Dy = 0$ i w konsekwencji $By = D(Dy) = 0$. Analogicznie $Cy = 0$. Mamy więc

$$\|Bx - Cx\|^2 = \langle (B-C)^2 x, x\rangle = \langle (B-C)y, x\rangle = 0,$$

co oznacza, że $Bx = Cx$ dla dowolnego $x \in H$, czyli $B = C$. ■

Wniosek 7.20. *Iloczyn dwóch przemiennych operatorów dodatnich jest również operatorem dodatnim.*

D o w ó d. Niech A i B będą takimi operatorami dodatnimi, że $AB = BA$. Wówczas operator A jest również przemienny z operatorem $\sqrt{B}$. Zatem dla dowolnego wektora $x \in H$ mamy

$$\langle ABx, x\rangle = \langle A\sqrt{B}\sqrt{B}x, \rangle = \langle \sqrt{B}A\sqrt{B}x, x\rangle = \langle A(\sqrt{B}x), \sqrt{B}x\rangle \geqslant 0. \qquad ■$$

Wniosek 7.21. *Jeżeli A i B są operatorami samosprzężonymi przemiennymi z operatorem dodatnim C i $A \geqslant B$, to również $AC \geqslant BC$.*

7.6. Twierdzenie spektralne dla operatora samosprzężonego

Naszym celem w tym podrozdziale jest wykazanie twierdzenia spektralnego dla operatorów samosprzężonych na nieskończenie wymiarowej przestrzeni Hilberta. Dowód będzie polegał na zbudowaniu pewnej funkcji zwanej *rozkładem identyczności* generowanym przez operator samosprzężony, a następnie za jej pomocą — na znalezieniu reprezentacji całkowej dla danego operatora.

Niech, jak poprzednio, A będzie ustalonym samosprzężonym operatorem na przestrzeni Hilberta H. Niech $m = \inf\{\langle Ax, x\rangle \colon \|x\| = 1\}$ oraz $M = \sup\{\langle Ax, x\rangle \colon \|x\| = 1\}$. Wówczas $mI \leqslant A \leqslant MI$. Aby nie rozważać trywialnego przypadku, będziemy zakładać, że $m < M$.

Najpierw zbudujemy pewną rodzinę rzutów odpowiadającą operatorowi A. Dowolnej liczbie rzeczywistej μ przyporządkujemy funkcję rzeczywistą e_μ określoną w następujący sposób: jeżeli $m \leqslant \mu < M$, to definiujemy

$$e_\mu(\lambda) = \begin{cases} 1, & \text{gdy } \ m \leqslant \lambda \leqslant \mu, \\ 0, & \text{gdy } \ \mu < \lambda \leqslant M; \end{cases}$$

jeżeli natomiast $\mu < m$, to $e_\mu = 0$, a dla $\mu \geqslant M$ definiujemy $e_\mu = 1$.

Lemat 7.22. *Funkcja e_μ należy do zbioru $\mathcal{C}_1$ dla dowolnej liczby rzeczywistej μ.*

D o w ó d. Niech μ będzie liczbą rzeczywistą. Jeżeli $\mu < m$ lub $\mu \geqslant M$, to funcja e_μ jest stała, a więc należy do $\mathcal{C}_1$.

Niech $m \leqslant \mu < M$ i niech N będzie najmniejszą liczbą naturalną, dla której $\mu + \frac{1}{N} \leqslant M$. Dla $n \geqslant N$ definiujemy funkcję

$$f_n(\lambda) = \begin{cases} 1, & \text{gdy } \; m \leqslant \lambda \leqslant \mu, \\ -n\lambda + n\mu + 1, & \text{gdy } \; \mu < \lambda < \mu + \frac{1}{n}, \\ 0, & \text{gdy } \; s + \frac{1}{n} \leqslant \lambda \leqslant M. \end{cases}$$

Wprost z tej definicji wynika, że f_n jest funkcją ciągłą na przedziale $[m, M]$ oraz że $\lim\limits_{n\to\infty} f_n(\lambda) = e_\mu(\lambda)$ dla $\lambda \in [m, M]$. Ponadto $0 \leqslant f_{n+1} \leqslant f_n$ dla $n = N, N+1, \ldots$ Z twierdzenia aproksymacyjnego Weierstrassa otrzymujemy ciąg wielomianów (p_n) taki, że

$$f_n(\lambda) + \frac{1}{2^{n+1}} < p_n(\lambda) < f_n(\lambda) + \frac{1}{2^n}$$

dla każdego $\lambda \in [m, M]$ i wszystkich $n = N, N+1, \ldots$ Jest oczywiste, że $p_n \in \mathcal{P}^+$, $p_{n+1} \leqslant p_n$ dla $n = N, N+1, \ldots$ i $\lim\limits_{n\to\infty} p_n(\lambda) = e_\mu(\lambda)$ dla $\lambda \in [m, M]$. To oznacza, że $e_\mu \in \mathcal{C}_1$. ∎

Na mocy definicji 7.3 każdej funkcji e_μ odpowiada jednoznacznie wyznaczony operator dodatni $e_\mu(A) = E_\mu$, który jest punktową granicą ciągu wielomianów od operatora A. Ponadto, ponieważ $e_\mu^2(\lambda) = e_\mu(\lambda)$ dla $\lambda \in [m, M]$, więc z twierdzenia 7.18 wynika, że $E_\mu^2 = E_\mu$. Ponieważ operator E_μ jest samosprzężony, więc jest on rzutem ortogonalnym (patrz tw. 4.12). Własności operatorów rzutowania E_μ są zawarte w następującym twierdzeniu:

Twierdzenie 7.23. *Dla dowolnego operatora samosprzężonego A istnieje rodzina rzutów ortogonalnych (E_μ) zależna od parametru $\mu \in \mathbb{R}$, spełniająca następujące warunki*:

(a) *dowolny operator C przemienny z operatorem A jest również przemienny z każdym z operatorów E_μ, $\mu \in \mathbb{R}$;*

(b) *jeżeli $\mu \leqslant \nu$, to $E_\mu \leqslant E_\nu$ lub, co jest równoważne, $E_\mu E_\nu = E_\mu$;*

(c) *funkcja $\mu \mapsto E_\mu$ jest prawostronnie ciągła w sensie punktowej zbieżności operatorów, tzn. $\lim\limits_{\varepsilon\to 0+} E_{\mu+\varepsilon}x = E_\mu x$ dla dowolnego $x \in H$;*

(d) *$E_\mu = 0$ dla $\mu < m$ i $E_\mu = I$ dla $\mu \geqslant M$, gdzie $m = \inf\{\langle Ax, x\rangle\colon \|x\| = 1\}$ oraz $M = \sup\{\langle Ax, x\rangle\colon \|x\| = 1\}$.*

Definicja 7.6. Rodzina rzutów (E_μ) nazywa się *rozkładem identyczności* generowanym przez operator A lub *rozkładem spektralnym* operatora A.

D o w ó d   t w i e r d z e n i a   7.23. Część (a) wynika z tego, że operator E_μ jest punktową granicą ciągu wielomianów od operatora A.

Aby wykazać część (b), zauważmy, że jeżeli $\mu \leqslant \nu$, to $e_\mu(\lambda) \leqslant e_\nu(\lambda)$ dla dowolnego $\lambda \in [m, M]$ i lematu 7.17 (d) otrzymujemy, że $E_\mu \leqslant E_\nu$. Ponadto z nierówności $\mu \leqslant \nu$ wynika, że $e_\mu(\lambda)e_\nu(\lambda) = e_\mu(\lambda)$, więc $E_\mu E_\nu = E_\mu$, czyli E_μ jest częścią rzutu E_ν (por. lemat 7.4).

Ponieważ $e_\mu = 0$ dla $\mu < m$, więc $E_\mu = 0$ dla takich μ na mocy lematu 7.17 (a). Analogicznie, z tego, że $e_\mu = 1$ dla $\mu \geqslant M$, z lematu 7.17 (c) otrzymujemy równość $E_\mu = I$ dla $\mu \geqslant M$.

Pozostaje do wykazania część (c), czyli prawostronna ciągłość funkcji $\lambda \mapsto E_\lambda$. Ustalmy punkt μ i skonstruujmy taki ciąg wielomianów (p_n) ze zbioru $\mathcal{P}^+$, że $p_{n+1}(\lambda) \leqslant p_n(\lambda)$, $e_{\mu+1/n}(\lambda) \leqslant p_n(\lambda)$ dla $\lambda \in [m, M]$ i dużych n oraz $\lim\limits_{n\to\infty} p_n(\lambda) = e_\mu(\lambda)$. Taki ciąg (p_n) można skonstruować, odpowiednio modyfikując rozumowanie przeprowadzone w lemacie 7.22. Wówczas $p_n(A) \geqslant E_{\mu+1/n} \geqslant E_\mu$. Ponieważ $p_n(A)x \to E_\mu x$, stąd $E_{\mu+1/n}x \to E_\mu x$. Z monotoniczności funkcji $\lambda \mapsto E_\lambda$ wynika, że $\lim\limits_{\varepsilon\to 0+} E_{\mu+\varepsilon}x = E_\mu x$, co kończy dowód własności (c). ■

U w a g a. Rozważania dotyczące konstrukcji rozkładu identyczności (E_λ) operatora samosprzężonego A były prowadzone przy założeniu, że $m < M$. Zauważmy, że jeżeli $m = M$, to $A = mI$ i rozkład spektralny tego operatora jest trywialny, tzn. $E_\lambda = 0$ dla $\lambda < m$ oraz $E_\lambda = I$ dla $\lambda \geqslant m$.

Aby objaśnić, w jaki sposób operator samosprzężony można aproksymować za pomocą kombinacji liniowych operatorów rzutowania E_λ, potrzebujemy pojęcia *całki Riemanna-Stieltjesa* z funkcji skalarnej względem funkcji przyjmującej wartości wektorowe. Definicja takiej całki jest analogiczna do definicji całki Riemanna-Stieltjesa rozważanej w analizie matematycznej.

Definicja 7.7. Niech $[a, b]$ będzie przedziałem na prostej i niech $\mathcal{X}$ będzie przestrzenią Banacha. Niech $f\colon [a, b] \to \mathbb{R}$ i $\varPhi\colon [a, b] \to \mathcal{X}$. Mówimy, że *funkcja f jest całkowalna względem funkcji $\varPhi$* na przedziale $[a, b]$, jeżeli istnieje wektor $S \in \mathcal{X}$ taki, że dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że dla dowolnego podziału $\varPi\colon\ a = t_0 < t_1 < \ldots < t_n = b$ przedziału $[a, b]$, dla którego $\max_i(t_i - t_{i-1}) < \delta$ i dowolnego wyboru punktów pośrednich s_i, $t_{i-1} \leqslant s_i \leqslant t_i$ dla $i = 1, 2, \ldots, n$, zachodzi nierówność

$$\left\| S - \sum_{i=1}^{n} f(s_i)\,(\varPhi(t_i) - \varPhi(t_{i-1})) \right\| < \varepsilon.$$

Wówczas wektor S o powyższej własności jest wyznaczony jednoznacznie i nosi nazwę *całki Riemanna-Stieltjesa* z funkcji f względem funkcji Φ. Będziemy go oznaczać symbolem

$$S = \int_a^b f(t)\, d\Phi(t).$$

Uwagi. 1. W definicji całki Riemanna-Stieltjesa funkcja f może przyjmować wartości zespolone, jak również przestrzeń Banacha $\mathcal{X}$ może być przestrzenią nad ciałem liczb zespolonych.

2. Jeżeli funkcje f i Φ są określone na całej prostej i funkcja f jest całkowalna względem funkcji Φ na dowolnym skończonym przedziale, to *niewłaściwą całkę Riemanna-Stieltjesa* definiujemy za pomocą równości

$$\int_{-\infty}^{+\infty} f(t)\, d\Phi(t) = \lim_{\substack{a \to -\infty \\ b \to +\infty}} \int_a^b f(t)\, d\Phi(t),$$

pod warunkiem, że ta granica istnieje.

Twierdzenie 7.24 (twierdzenie spektralne dla operatora samosprzężonego). *Niech A będzie operatorem samosprzężonym na przestrzeni Hilberta H i niech (E_λ) będzie jego rozkładem spektralnym. Niech $\eta > 0$ będzie dowolne. Wówczas funkcja $\lambda \mapsto \lambda$ jest całkowalna względem funkcji $\lambda \mapsto E_\lambda$ na przedziale $[m - \eta, M]$ i zachodzi równość*

$$A = \int_{m-\eta}^{M} \lambda\, dE_\lambda, \tag{7.15}$$

gdzie $m = \inf\{\langle Ax, x\rangle\colon\ \|x\| = 1\}$ oraz $M = \sup\{\langle Ax, x\rangle\colon\ \|x\| = 1\}$.

Dowód. Zauważmy, że dla dowolnych liczb rzeczywistych $\mu < \nu$ mamy

$$e_\nu(\lambda) - e_\mu(\lambda) = \begin{cases} 1 & \text{dla } \lambda \in (\mu, \nu] \cap [m, M], \\ 0 & \text{dla } \lambda \in [m, M] \setminus (\mu, \nu]. \end{cases}$$

Zatem dla dowolnego $\lambda \in [m, M]$ mamy

$$\mu(e_\nu(\lambda) - e_\mu(\lambda)) \leqslant \lambda(e_\nu(\lambda) - e_\mu(\lambda)) \leqslant \nu(e_\nu(\lambda) - e_\mu(\lambda)).$$

Stąd otrzymujemy

$$\mu(E_\nu - E_\mu) \leqslant A(E_\nu - E_\mu) \leqslant \nu(E_\nu - E_\mu). \tag{7.16}$$

Wybierzmy dowolne $\varepsilon > 0$. Niech $\Pi\colon\ m-\eta = \lambda_0 < \lambda_1 < \ldots < \lambda_n = M$ będzie dowolnym podziałem przedziału $[m-\eta, M]$. Wówczas na mocy (7.16) mamy

$$\lambda_{i-1}(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant A(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant \lambda_i(E_{\lambda_i} - E_{\lambda_{i-1}})$$

dla $i = 1, 2, \ldots, n$. Sumując te nierówności po wszystkich i oraz biorąc pod uwagę, że $\sum\limits_{i=1}^{n}(E_{\lambda_i} - E_{\lambda_{i-1}}) = I$, otrzymujemy

$$\sum_{i=1}^{n}\lambda_{i-1}(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant A \leqslant \sum_{i=1}^{n}\lambda_i(E_{\lambda_i} - E_{\lambda_{i-1}}).$$

Stąd dla dowolnych liczb $\lambda_{i-1} \leqslant \nu_i \leqslant \lambda_i$, $i = 1, 2, \ldots, n$, otrzymujemy

$$\sum_{i=1}^{n}(\lambda_{i-1} - \nu_i)(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant A - \sum_{i=1}^{n}\nu_i(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant$$
$$\leqslant \sum_{i=1}^{n}(\lambda_i - \nu_i)(E_{\lambda_i} - E_{\lambda_{i-1}}).$$

Niech $\max_i(\lambda_i - \lambda_{i-1}) = \delta$. Wówczas z powyższych nierówności otrzymujemy

$$-\delta I \leqslant A - \sum_{i=1}^{n}\nu_i(E_{\lambda_i} - E_{\lambda_{i-1}}) \leqslant \delta I,$$

co oznacza, że

$$-\delta\langle x, x\rangle \leqslant \left\langle\Big(A - \sum_{i=1}^{n}\nu_i(E_{\lambda_i} - E_{\lambda_{i-1}})\Big)x, x\right\rangle \leqslant \delta\langle x, x\rangle.$$

Stąd wynika, że dla $\delta < \varepsilon$ mamy

$$\left\|A - \sum_{i=1}^{n}\nu_i(E_{\lambda_i} - E_{\lambda_{i-1}})\right\| < \varepsilon, \tag{7.17}$$

czyli

$$A = \int_{m-\eta}^{M}\lambda\, dE_\lambda. \qquad \blacksquare$$

U w a g i. 1. Ponieważ $E_\lambda = 0$ dla $\lambda < m$ i $E_\lambda = I$ dla $\lambda \geqslant M$, więc

$$\int\limits_{m-\eta}^{M} \lambda \, dE_\lambda = \int\limits_{-\infty}^{+\infty} \lambda \, dE_\lambda$$

i wzór (7.15) przyjmuje postać

$$A = \int\limits_{-\infty}^{+\infty} \lambda \, dE_\lambda .$$

2. Ponieważ zbieżność ciągu operatorów (A_n) według normy operatorowej do operatora A implikuje zbieżność punktową, a także zbieżność $\langle A_n x, y\rangle \to \langle Ax, y\rangle$, więc z twierdzenia 7.24 wynika, że dla dowolnego operatora samosprzężonego A i dla dowolnych wektorów $x, y \in H$ zachodzą równości

$$Ax = \int\limits_{m-\eta}^{M} \lambda \, dE_\lambda x, \tag{7.18}$$

$$\langle Ax, y\rangle = \int\limits_{m-\eta}^{M} \lambda \, d\langle E_\lambda x, y\rangle ,$$

gdzie po prawych stronach występują odpowiednie całki Riemanna-Stieltjesa.

3. Reprezentacja wektora Ax wyrażona we wzorze (7.18) jest uogólnieniem przedstawienia

$$Ax = \sum_{n=1}^{\infty} \mu_n \langle x, e_n\rangle e_n$$

uzyskanego we wniosku 7.8 dla zwartego operatora normalnego, a więc w szczególności dla zwartego operatora samosprzężonego. Istotnie, szereg ten może być przedstawiony w postaci całki Riemanna-Stieltjesa typu (7.18), jeżeli zdefiniujemy E_λ za pomocą wzoru

$$E_\lambda x = \begin{cases} \sum\limits_{\mu_k \leqslant \lambda} \langle x, e_k\rangle e_k & \text{dla } \ \lambda < 0, \\ x - \sum\limits_{\mu_k > \lambda} \langle x, e_k\rangle e_k & \text{dla } \ \lambda \geqslant 0. \end{cases}$$

Wówczas E_λ, jako funkcja zmiennej λ, jest stała pomiędzy kolejnymi wartościami własnymi operatora A, jest równa zeru, gdy λ jest mniejsze od wszystkich wartości własnych i jest równa I, gdy λ jest większe od każdej wartości własnej tego operatora. Skok funkcji E_λ, gdy λ przechodzi przez wartość własną μ, tzn. $E(\mu) = E_\mu - E_{\mu-}$, jest rzutem na podprzestrzeń własną odpowiadającą μ:

$$E(\mu)x = \sum_{\mu_k=\mu} \langle x, e_k \rangle e_k \quad \text{dla} \quad \mu \neq 0,$$

$$E(0)x = x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k.$$

4. Twierdzenie spektralne jest również prawdziwe dla operatorów normalnych oraz dla operatorów unitarnych. Zainteresowany Czytelnik może znaleźć te twierdzenia w wielu książkach poświęconych teorii spektralnej, np. w [1], [3], [8] Part II, [15], [19], [21], [24].

Naszym następnym celem jest zbudowanie reprezentacji całkowej dla rachunku funkcyjnego.

W dalszym ciągu zakładamy, że A jest operatorem samosprzężonym, E_λ jest jego rozkładem spektralnym oraz $m = \inf\{\langle Ax, x\rangle\colon \|x\| = 1\}$ i $M = \sup\{\langle Ax, x\rangle\colon \|x\| = 1\}$.

Lemat 7.25. *Dla dowolnej liczby całkowitej nieujemnej k funkcja $\lambda \mapsto \lambda^k$ jest całkowalna względem funkcji $\lambda \mapsto E_\lambda$ na przedziale $[m-\eta, M]$ i zachodzi równość*

$$A^k = \int_{m-\eta}^{M} \lambda^k \, dE_\lambda. \tag{7.19}$$

D o w ó d. Dla $k = 0$ równość (7.19) jest oczywista, a dla $k = 1$ pokrywa się ona z równością (7.15). Załóżmy więc, że $k > 1$. Niech $\Pi\colon m - \eta = \lambda_0 < \lambda_1 < \ldots < \lambda_n = M$ będzie dowolnym podziałem przedziału $[m - \eta, M]$. Ponadto niech liczby $\lambda_{i-1} \leqslant \nu_i \leqslant \lambda_i$, $i = 1, 2, \ldots, n$, będą dowolne. Z twierdzenia 7.23 (b) wynika, że

$$\left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\left(E_{\lambda_j} - E_{\lambda_{j-1}}\right) = 0 \quad \text{dla} \quad i \neq j.$$

Stąd przez łatwą indukcję otrzymujemy równość

$$\left(\sum_{i=1}^{n} \nu_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right)^k = \sum_{i=1}^{n} \nu_i^k \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right).$$

Zatem

$$\begin{aligned}
&\left\|A^k - \sum_{i=1}^n \nu_i^k \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| = \left\|A^k - \left(\sum_{i=1}^n \nu_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right)^k\right\| \leqslant \\
&\leqslant \left\|A - \sum_{i=1}^n \nu_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| \left\|\sum_{j=1}^k A^{j-1}\left(\sum_{i=1}^n \nu_i^{k-j} \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right)\right\| \leqslant \\
&\leqslant \left\|A - \sum_{i=1}^n \nu_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| \sum_{j=1}^k \|A\|^{j-1} \left\|\sum_{i=1}^n \nu_i^{k-j} \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\|.
\end{aligned}$$

Niech $N = \max\{|m-\eta|, |M|\}$. Ponieważ

$$\begin{aligned}
&\left\langle \sum_{i=1}^n \nu_i^{k-j} \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right) x, x \right\rangle = \sum_{i=1}^n \nu_i^{k-j} \left\langle \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right) x, x \right\rangle \leqslant \\
&\leqslant N^{k-j} \sum_{i=1}^n \left\langle \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right) x, x \right\rangle = N^{k-j} \left\langle \sum_{i=1}^n \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right) x, x \right\rangle = \\
&= N^{k-j} \langle x, x \rangle,
\end{aligned}$$

więc mamy

$$\left\|\sum_{i=1}^n \nu_i^{k-j} \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| \leqslant N^{k-j}.$$

Stąd i z poprzednich oszacowań otrzymujemy

$$\left\|A^k - \sum_{i=1}^n \nu_i^k \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| \leqslant C \left\|A - \sum_{i=1}^n \nu_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\|, \tag{7.20}$$

gdzie C jest stałą dodatnią, która zależy tylko od operatora A i liczby k. Biorąc dowolne $\varepsilon > 0$ i podział Π taki, że $\delta = \max_i(\lambda_i - \lambda_{i-1}) < C^{-1}\varepsilon$, z (7.17) i (7.20) otrzymujemy

$$\left\|A^k - \sum_{i=1}^n \nu_i^k \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right)\right\| < \varepsilon,$$

co oznacza, że funkcja $\lambda \mapsto \lambda^k$ jest całkowalna i zachodzi wzór (7.19). ■

Z liniowości całki Riemanna-Stieltjesa ze względu na funkcję, którą całkujemy, otrzymujemy następujący wniosek:

Wniosek 7.26. *Jeżeli* $p(\lambda) = \alpha_n\lambda^n + \alpha_{n-1}\lambda^{n-1} + \ldots + \alpha_1\lambda + \alpha_0$ *jest dowolnym wielomianem o współczynnikach rzeczywistych, to funkcja* $\lambda \mapsto p(\lambda)$ *jest całkowalna względem funkcji* $\lambda \mapsto E_\lambda$ *rozkładu spektralnego operatora* A *i zachodzi równość*

$$p(A) = \int_{m-\eta}^{M} p(\lambda)\, dE_\lambda. \tag{7.21}$$

Twierdzenie 7.27 (całkowa reprezentacja rachunku funkcyjnego). *Niech* f *będzie ciągłą funkcją rzeczywistą na przedziale* $[m, M]$ *i niech* $\eta > 0$ *będzie dowolne. Rozszerzamy funkcję* f *na przedział* $[m-\eta, M]$*, przyjmując, że* $f(\lambda) = f(m)$ *dla* $\lambda \in [m-\eta, m)$*. Wówczas funkcja* f *jest całkowalna na przedziale* $[m-\eta, M]$ *względem funkcji* $\lambda \mapsto E_\lambda$ *rozkładu spektralnego operatora* A *i zachodzi równość*

$$f(A) = \int_{m-\eta}^{M} f(\lambda)\, dE_\lambda. \tag{7.22}$$

D o w ó d. Bierzemy dowolne $\varepsilon > 0$. Na mocy twierdzenia aproksymacyjnego Weierstrassa istnieje wielomian p taki, że dla dowolnego $\lambda \in [m-\eta, M]$ mamy

$$-\frac{1}{3}\varepsilon \leqslant f(\lambda) - p(\lambda) \leqslant \frac{1}{3}\varepsilon. \tag{7.23}$$

W szczególności nierówności te zachodzą na przedziale $[m, M]$. Stąd na mocy twierdzenia 7.18 (c) otrzymujemy

$$-\frac{1}{3}\varepsilon I \leqslant f(A) - p(A) \leqslant \frac{1}{3}\varepsilon I,$$

czyli

$$\|f(A) - p(A)\| \leqslant \frac{1}{3}\varepsilon.$$

Niech $\varPi\colon m-\eta = \lambda_0 < \lambda_1 < \ldots < \lambda_n = M$ będzie dowolnym podziałem przedziału $[m-\eta, M]$. Ponadto niech liczby $\lambda_{i-1} \leqslant \nu_i \leqslant \lambda_i$, $i = 1, 2, \ldots, n$, będą dowolne. Bierzemy sumy $S_f = \sum\limits_{i=1}^{n} f(\nu_i)(E_{\lambda_i} - E_{\lambda_{i-1}})$ i $S_p = \sum\limits_{i=1}^{n} p(\nu_i)(E_{\lambda_i} - E_{\lambda_{i-1}})$. Wówczas

$$S_f - S_p = \sum_{i=1}^{n} (f(\nu_i) - p(\nu_i))(E_{\lambda_i} - E_{\lambda_{i-1}}).$$

Z nierówności (7.23) otrzymujemy

$$-\frac{1}{3}\varepsilon I \leqslant S_f - S_p \leqslant \frac{1}{3}\varepsilon I$$

i w konsekwencji

$$\|S_f - S_p\| \leqslant \frac{1}{3}\varepsilon.$$

Z wniosku 7.26 wynika, że możemy wybrać tak drobny podział Π przedziału $[m-\eta, M]$, aby

$$\|p(A) - S_p\| \leqslant \frac{1}{3}\varepsilon.$$

Ostatecznie więc mamy

$$\|f(A) - S_f\| \leqslant \|f(A) - p(A)\| + \|p(A) - S_p\| + \|S_p - S_f\| \leqslant \varepsilon.$$

Zatem funkcja f jest całkowalna względem rozkładu spektralnego operatora A i zachodzi równość

$$f(A) = \int_{m-\eta}^{m} f(\lambda)\, dE_\lambda. \qquad \blacksquare$$

U w a g i. 1. Ponieważ $E_\lambda = 0$ dla $\lambda < m$ i $E_\lambda = I$ dla $\lambda \geqslant M$, więc wzór (7.22) można zapisać w postaci

$$f(A) = \int_{-\infty}^{+\infty} f(\lambda)\, dE_\lambda.$$

2. Nie jest istotny sposób, w jaki rozszerzamy funkcję f do funkcji ciągłej na przedziale $[m-\eta, M]$. Wartość całki we wzorze (7.22) nie ulegnie zmianie przy dowolnym innym takim rozszerzeniu. Zamiast rozszerzania funkcji f na przedział $[m-\eta, M]$ można by postąpić inaczej. Mianowicie gdybyśmy przyjęli, że $E_m = 0$, to wtedy we wzorze (7.22) wystarczyłaby całka w granicach od m do M. Jednakże tak zmodyfikowany rozkład spektralny operatora A nie byłby prawostronnie ciągły w punkcie m.

3. Ponieważ zbieżność ciągu operatorów (A_n) według normy operatorowej do operatora A implikuje zbieżność punktową, a także zbieżność $\langle A_n x, y\rangle \to \langle Ax, y\rangle$, więc z twierdzenia 7.27 wynika, że dla dowolnego

operatora samosprzężonego A, dowolnej funkcji ciągłej f na widmie tego operatora i dla dowolnych wektorów $x, y \in H$ zachodzą równości

$$f(A)x = \int_{m-\eta}^{M} f(\lambda)\, dE_\lambda x,$$

$$\langle f(A)x, y\rangle = \int_{m-\eta}^{M} f(\lambda)\, d\langle E_\lambda x, y\rangle, \tag{7.24}$$

gdzie po prawych stronach występują odpowiednie całki Riemanna-Stieltjesa.

4. Rozważając rozkład funkcji zespolonej na części rzeczywistą i urojoną, można łatwo sprawdzić, że wniosek 7.26 i twierdzenie 7.27 są prawdziwe również dla wielomianów p o współczynnikach zespolonych i ciągłych funkcji zespolonych f. Wówczas operatory $p(A)$ i $f(A)$ nie są na ogół samosprzężone. Są one operatorami normalnymi.

Na zakończenie wyjaśnimy problem jednoznaczności rozkładu spektralnego operatora samosprzężonego. Mianowicie prawdziwe jest następujące twierdzenie:

Twierdzenie 7.28. *Rozkład spektralny* (E_λ) *operatora samosprzężonego* A *jest wyznaczony jednoznacznie.*

D o w ó d. Jeżeli rodzina (E_λ) jest rozkładem spektralnym operatora A, to dla niej prawdziwe jest twierdzenie 7.27 i w konsekwencji dla dowolnej funkcji ciągłej f na przedziale $[m, M]$ zachodzi wzór (7.24). W szczególności dla dowolnego wektora $x \in H$ zachodzi równość

$$\langle f(A)x, x\rangle = \int_{m-\eta}^{M} f(\lambda)\, d\langle E_\lambda x, x\rangle. \tag{7.25}$$

Lewa strona tego wzoru jest zdefiniowana niezależnie od funkcji rozkładu E_λ i dla dowolnego (ustalonego) wektora $x \in H$ jest ona funkcjonałem liniowym i ograniczonym na przestrzeni $\mathcal{C}_\mathbb{R}[m, M]$. Wzór (7.25) jest więc całkową reprezentacją tego funkcjonału (patrz twierdzenie D.6 w dodatku). Z twierdzenia 7.23 (b), (c) i (d) wynika, że funkcja $\lambda \mapsto \langle E_\lambda x, x\rangle$ jest niemalejąca, prawostronnie ciągła i dla $\lambda < m$ jest równa 0. Stąd na podstawie twierdzenia D.9 jest ona wyznaczona jednoznacznie. ■

Ćwiczenia

1. Niech A będzie operatorem samosprzężonym na przestrzeni Hilberta H. Udowodnić, że jeżeli $\langle Ax, x\rangle \neq 0$ dla $x \neq 0$ i $\langle Ax, x\rangle > 0$ przynajmniej dla jednego $x \in H$, to operator A jest dodatni.

2. Wykazać, że jeżeli $A \neq 0$ jest operatorem zwartym samosprzężonym, to A jest dodatni wtedy i tylko wtedy, gdy wszystkie różne od zera wartości własne operatora A są dodatnie.

3. Wykazać, że jeżeli A jest operatorem zwartym samosprzężonym dodatnim, to $\|A\| = \langle Ae, e\rangle = \lambda$, gdzie λ jest największą wartością własną operatora A, a e odpowiadającym jej wektorem własnym o normie 1.

4. Niech $m \leqslant \mu < M$ i niech $p(\lambda) = 1 + (M-m)^{-1}(\mu - \lambda)$ dla $\lambda \in [m, M]$. Niech (p_n) będzie ciągiem wielomianów zdefiniowanym indukcyjnie w następujący sposób: $p_1 = p$ i $p_{n+1} = p_n(1 - \frac{1}{4}(p_n - 1)^2)$ dla $n = 1, 2, \ldots$ Udowodnić, że $0 \leqslant p_{n+1} \leqslant p_n$ dla $n = 1, 2, \ldots$ i $\lim\limits_{n\to\infty} p_n(\lambda) = e_\mu(\lambda)$, gdzie e_μ jest funkcją z lematu 7.22.

U w a g a. W kolejnych ćwiczeniach zakładamy, że A jest operatorem samosprzężonym, E_λ jest jego rozkładem spektralnym i $m = \inf\{\langle Ax, x\rangle \colon \|x\| = 1\}$ oraz $M = \sup\{\langle Ax, x\rangle \colon \|x\| = 1\}$.

5. Niech A będzie operatorem samosprzężonym i niech $f \in \mathcal{C}_{\mathbb{R}}[m, M])$. Wykazać, że jeżeli $f = \lim\limits_{n\to\infty} p_n$, $p_n \in \mathcal{P}$ dla $n = 1, 2, \ldots$, to $f(A) = \lim\limits_{n\to\infty} p_n(A)$ oraz dowolny operator B przemienny z operatorem A jest również przemienny z operatorem $f(A)$.

6. Udowodnić, że dla operatora samosprzężonego A i funkcji $f \in \mathcal{C}_{\mathbb{R}}[m, M])$ zachodzi równość

$$\sigma(f(A)) = f(\sigma(A)).$$

7. Niech λ będzie liczbą rzeczywistą i załóżmy, że istnieje $\delta > 0$ takie, że $E\mu = E_\nu$ dla $\lambda - \delta \leqslant \mu \leqslant \nu \leqslant \lambda + \delta$. Udowodnić, że $\lambda \in \varrho(A)$. Ponadto wykazać, że dla dowolnej funkcji f takiej, że $f(\mu) = (\mu - \lambda)^{-1}$ dla $\mu \notin [m - \eta, M] \cap [\lambda - \delta, \lambda + \delta]$, zachodzi równość

$$(A - \lambda I)^{-1} = \int\limits_{m-\eta}^{M} f(\mu)\, dE_\mu.$$

8. Wykazać, że jeżeli liczba rzeczywista λ należy do zbioru rezolwenty $\varrho(A)$ operatora A, to istnieje $\delta > 0$ takie, że $E\mu = E_\nu$ dla $\lambda - \delta \leqslant \mu \leqslant \nu \leqslant \lambda + \delta$.

9. Wykazać, że liczba rzeczywista λ jest wartością własną operatora A wtedy i tylko wtedy, gdy $E_\lambda \neq E_{\lambda-}$, i wówczas $\ker(A - \lambda I) = \operatorname{im}(E_\lambda - E_{\lambda-})$.

Dodatek

Przestrzeń sprzężona z przestrzenią $\mathcal{C}[a, b]$

W tym dodatkowym rozdziale naszym celem jest opisanie przestrzeni sprzężonej z przestrzenią $\mathcal{C}[a, b]$ funkcji ciągłych określonych na przedziale $[a, b]$.

Definicja D.1. Niech φ będzie funkcją o wartościach rzeczywistych lub zespolonych określoną na przedziale $[a, b]$. Dla podziału $\Pi\colon\ a = x_0 < x_1 < < x_2 < \ldots < x_n = b$ przedziału $[a, b]$ tworzymy sumę

$$\sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_{j-1})|. \tag{D.1}$$

Jeżeli zbiór tych sum jest ograniczony z góry przy Π przebiegającym rodzinę wszystkich podziałów przedziału $[a, b]$, to mówimy, że φ jest funkcją o *wahaniu ograniczonym* (lub *wariacji ograniczonej*) na tym przedziale, a liczbę

$$\sup_{\Pi} \sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_{j-1})|$$

nazywamy *wahaniem funkcji* (lub *wariacją funkcji*) φ i oznaczamy symbolem $\mathrm{var}(\varphi; a, b)$ lub krótko $\mathrm{var}(\varphi)$, jeżeli wiadomo, o jaki przedział chodzi. Zauważmy, że przy zagęszczaniu podziałów sumy (D.1) nie maleją.

Następujące fakty zawierają elementarne informacje o funkcjach o wahaniu ograniczonym. Są one oczywiste lub łatwe do udowodnienia, dlatego ich dowody pozostawiamy jako ćwiczenia dla Czytelnika.

— Funkcja $\varphi\colon\ [a, b] \to \mathbb{C}$ jest funkcją o wahaniu ograniczonym wtedy i tylko wtedy, gdy funkcje $\mathrm{Re}\,\varphi$ i $\mathrm{Im}\,\varphi$ mają tę własność.
— Funkcja monotoniczna ma wahanie ograniczone.
— Funkcja spełniająca warunek Lipschitza jest funkcją o wahaniu ograniczonym.
— Funkcja o wahaniu ograniczonym jest funkcją ograniczoną.

— Funkcja

$$\varphi(x) = \begin{cases} x \sin \dfrac{\pi}{x}, & \text{gdy } 0 < x \leqslant 2, \\ 0, & \text{gdy } x = 0 \end{cases}$$

jest funkcją ciągłą, której wahanie jest nieograniczone.

— Suma i iloczyn funkcji o wahaniu ograniczonym są funkcjami o tej samej własności. W szczególności zbiór funkcji o wahaniu ograniczonym (na przedziale $[a, b]$) jest przestrzenią liniową. Przestrzeń tę będziemy oznaczać symbolem $BV[a, b]$.

Z przytoczonych własności funkcji o wahaniu ograniczonym wynika, że różnica funkcji monotonicznych ma wahanie ograniczone. Prawdziwe jest również twierdzenie odwrotne, tzn. funkcja rzeczywista o wahaniu ograniczonym jest różnicą dwóch funkcji rosnących (patrz [16]). Stąd wynika, że ma ona granice jednostronne w każdym punkcie przedziału, na którym jest określona (patrz [27], tw. 4.17 z części I). Wykażemy ten fakt bezpośrednio.

Twierdzenie D.1. *Funkcja o wahaniu ograniczonym na przedziale $[a, b]$ ma granicę lewostronną dla każdego $a < x \leqslant b$ i granicę prawostronną dla dowolnego $a \leqslant x < b$. Innymi słowy, funkcja o wahaniu ograniczonym może mieć tylko nieciągłości pierwszego rodzaju.*

D o w ó d. Niech φ będzie funkcją na przedziale $[a, b]$, która nie ma granicy lewostronnej w punkcie $x \in (a, b]$. Pokażemy, że φ nie ma ograniczonego wahania na przedziale $[a, b]$.

Ponieważ φ jest nieciągła lewostronnie w punkcie x, więc istnieje $\varepsilon > 0$ takie, że dla dowolnego $\delta > 0$ istnieją punkty t i s w przedziale $[a, b]$ takie, że $x - \delta < t < s < x$ oraz $|\varphi(s) - \varphi(t)| \geqslant \varepsilon$.

W przeciwnym wypadku dla dowolnego ciągu (t_n) punktów z przedziału $[a, x)$ zbieżnego do x ciąg $(\varphi(t_n))$ byłby ciągiem Cauchy'ego, a więc byłby zbieżny i jego granica byłaby taka sama dla wszystkich ciągów o tej własności. Zatem istniałaby granica lewostronna funkcji φ w punkcie x.

Wybieramy indukcyjnie ciągi (t_n) i (s_n) w taki sposób, że $a < t_1 < s_1 < t_2 < s_2 < \ldots < t_n < s_n < \ldots < x$ oraz $|\varphi(s_n) - \varphi(t_n)| \geqslant \varepsilon$ dla dowolnego n. Definiujemy następujący podział przedziału $[a, b]$:

$$x_0 = a, \quad x_{2j-1} = t_j, \quad x_{2j} = s_j \quad \text{dla} \quad j = 1, 2, \ldots, n, \quad x_{2n+1} = b.$$

Wówczas

$$\sum_{k=1}^{2n+1} |\varphi(x_k) - \varphi(x_{k-1})| \geqslant \sum_{j=1}^{n} |\varphi(s_j) - \varphi(t_j)| \geqslant n\varepsilon.$$

Stąd wynika, że wahanie funkcji φ jest nieograniczone. Dowód w przypadku granicy prawostronnej jest analogiczny. ■

Wniosek D.2. *Funkcja o wahaniu ograniczonym ma co najwyżej przeliczalnie wiele punktów nieciągłości.*

D o w ó d. Na mocy poprzedniego twierdzenia funkcja φ jest nieciągła w punkcie $x \in [a,b]$ wtedy i tylko wtedy, gdy $\varphi(x) \neq \varphi(x-)$ lub $\varphi(x) \neq \varphi(x+)$. Weźmy dowolne punkty $x_1, x_2, \ldots, x_n$ z przedziału (a,b). Niech $a < x_1 < < x_2 < \ldots < x_n < b$. Dla dowolnego $\varepsilon > 0$ tak wybieramy punkty t_j i s_j $(j = 1,2,\ldots,n)$, aby $a < t_1 < x_1 < s_1 < t_2 < x_2 < s_2 < \ldots < t_n < < x_n < s_n < b$ oraz aby $|\varphi(t_j) - \varphi(x_j-)| < \frac{1}{2n}\varepsilon$, $|\varphi(s_j) - \varphi(x_j+)| < \frac{1}{2n}\varepsilon$ dla $j = 1,2,\ldots,n$. Wówczas

$$\sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_j-)| + \sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_j+)| \leqslant$$
$$\leqslant \sum_{j=1}^{n} |\varphi(x_j) - \varphi(t_j)| + \sum_{j=1}^{n} |\varphi(t_j) - \varphi(x_j-)| + \sum_{j=1}^{n} |\varphi(s_j) - \varphi(x_j)| +$$
$$+ \sum_{j=1}^{n} |\varphi(s_j) - \varphi(x_j+)| \leqslant \operatorname{var}(\varphi; a, b) + \varepsilon.$$

Wobec dowolności $\varepsilon > 0$ otrzymujemy

$$\sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_j-)| + \sum_{j=1}^{n} |\varphi(x_j) - \varphi(x_j+)| \leqslant \operatorname{var}(\varphi; a, b).$$

Z tej nierówności wynika, że dla dowolnego $\varepsilon > 0$ istnieje co najwyżej skończenie wiele punktów $x \in [a,b]$, dla których

$$|\varphi(x) - \varphi(x-)| + |\varphi(x) - \varphi(x+)| \geqslant \varepsilon.$$

Stąd z kolei wnioskujemy, że zbiór punktów nieciągłości funkcji φ jest co najwyżej przeliczalny. ■

Łatwo zauważyć, że wahanie funkcji ma na przestrzeni $BV[a,b]$ następujące własności:

(a) $\operatorname{var}(\varphi; a, b) = 0$ wtedy i tylko wtedy, gdy $\varphi = \mathrm{const}$;

(b) $\operatorname{var}(\lambda\varphi; a, b) = |\lambda| \operatorname{var}(\varphi; a, b) \quad (\lambda \in \mathbb{C})$;

(c) $\operatorname{var}(\varphi + \psi; a, b) \leqslant \operatorname{var}(\varphi; a, b) + \operatorname{var}(\psi; a, b)$.

Zatem funkcjonał $\|\cdot\|_v$ zdefiniowany wzorem

$$\|\varphi\|_v = |\varphi(a)| + \operatorname{var}(\varphi; a, b) \tag{D.2}$$

jest normą na przestrzeni $BV[a, b]$.

Twierdzenie D.3. *Przestrzeń funkcji o wahaniu ograniczonym $BV[a, b]$ z normą daną wzorem* (D.2) *jest przestrzenią Banacha.*

D o w ó d. Niech (φ_n) będzie ciągiem Cauchy'ego w przestrzeni $BV[a, b]$. Ponieważ $|\varphi(x) - \varphi(a)| \leqslant \operatorname{var}(\varphi; a, b)$ dla każdego $x \in [a, b]$, więc

$$|\varphi(x)| \leqslant |\varphi(a)| + \operatorname{var}(\varphi; a, b).$$

Zatem

$$\|\varphi\|_\infty = \sup_{a \leqslant x \leqslant b} |\varphi(x)| \leqslant |\varphi(a)| + \operatorname{var}(\varphi; a, b) = \|\varphi\|_v.$$

Stąd wynika, że ciąg (φ_n) jest zbieżny jednostajnie na przedziale $[a, b]$ do pewnej funkcji φ (patrz [27], tw. 7.1 z części II). W istocie w dalszym ciągu wykorzystamy jedynie zbieżność punktową ciągu (φ_n).

Aby zakończyć dowód, wystarczy pokazać, że $\operatorname{var}(\varphi; a, b)$ jest skończone i $\operatorname{var}(\varphi_n - \varphi; a, b) \to 0$, gdy $n \to \infty$.

Niech będzie dane $\varepsilon > 0$. Wybierzmy n_0 tak, aby $\operatorname{var}(\varphi_n - \varphi_m; a, b) < \varepsilon$ dla n, $m \geqslant n_0$. Weźmy dowolny podział $\Pi\colon\ a = x_0 < x_1 < \ldots < x_k = b$. Wówczas

$$\sum_{j=1}^{k} |\varphi_n(x_j) - \varphi_m(x_j) - \varphi_n(x_{j-1}) + \varphi_m(x_{j-1})| < \varepsilon.$$

Przechodząc w tej nierówności do granicy, $m \to \infty$, otrzymujemy

$$\sum_{j=1}^{k} |\varphi_n(x_j) - \varphi(x_j) - \varphi_n(x_{j-1}) + \varphi(x_{j-1})| \leqslant \varepsilon$$

dla dowolnego $n \geqslant n_0$. Ponieważ podział Π był dowolny, więc stąd wynika, że

$$\operatorname{var}(\varphi_n - \varphi; a, b) \leqslant \varepsilon \quad \text{dla} \quad n \geqslant n_0.$$

Zatem $\operatorname{var}(\varphi_n - \varphi; a, b) \to 0$. Ponieważ $\varphi = \varphi_n - (\varphi_n - \varphi)$, więc wahanie funkcji φ jest ograniczone. ■

Definicja D.2. Niech f i α będą dwiema funkcjami rzeczywistymi lub zespolonymi określonymi na przedziale $[a,b]$. Niech $\Pi\colon\ a = x_0 < x_1 < \ldots < x_k = b$ będzie podziałem przedziału $[a,b]$. *Średnicą podziału* Π nazywamy liczbę $\mu(\Pi) = \max\{x_j - x_{j-1}\colon\ j = 1,2,\ldots,k\}$. Wybieramy liczby ξ_j tak, aby $x_{j-1} \leqslant \xi_j \leqslant x_j$ dla $j = 1,2,\ldots,k$ i tworzymy sumę

(D.3)
$$S(f,\alpha;\Pi) = \sum_{j=1}^{k} f(\xi_j)(\alpha(x_j) - \alpha(x_{j-1})).$$

Suma (D.3) nazywa się *sumą Stieltjesa*. Mówimy, że sumy $S(f,\alpha;\Pi)$ dążą do liczby A, gdy $\mu(\Pi) \to 0$, jeżeli dla dowolnego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że dla dowolnego podziału Π przedziału $[a,b]$, dla którego $\mu(\Pi) < \delta$ i dowolnego wyboru punktów ξ_j, $x_{j-1} \leqslant \xi_j \leqslant x_j$, zachodzi nierówność

(D.4)
$$|S(f,\alpha;\Pi) - A| < \varepsilon.$$

Piszemy wówczas

(D.5)
$$\lim_{\mu(\Pi)\to 0} S(f,\alpha;\Pi) = A.$$

Żeby uprościć oznaczenie, nie zaznaczyliśmy, że suma Stieltjesa (D.3) zależy również od *punktów pośrednich* ξ_j. Nie będzie to prowadzić do nieporozumień, jeżeli będziemy pamiętać, że równość (D.5) oznacza, że nierówność (D.4) zachodzi dla dowolnego podziału Π i dla dowolnego wyboru punktów pośrednich ξ_j, o ile tylko $\mu(\Pi) < \delta$.

Definicja D.3. Mówimy, że funkcja f jest *całkowalna* (*w sensie Stieltjesa*) *względem funkcji* α, jeżeli istnieje granica sum Stieltjesa w sensie definicji D.2. Wówczas tę granicę oznaczamy symbolem

$$\int_a^b f d\alpha \quad \text{lub} \quad \int_a^b f(x) d\alpha(x)$$

i nazywamy *całką Riemanna-Stieltjesa* funkcji f względem funkcji α.

U w a g a. W przypadku, gdy α jest funkcją monotoniczną, można również zdefiniować całkę Riemanna-Stieltjesa w analogiczny sposób jak definiowaliśmy całkę Riemanna w [27], część I, tzn. za pomocą całek dolnej i górnej Darboux. W książce Rudina (patrz [22], tw. 6.14) jest udowodnione twierdzenie, które wyjaśnia związek pomiędzy tymi dwiema definicjami całki.

Jeżeli $\alpha(x) = x$, to całka Riemanna-Stieltjesa jest równa całce Riemanna. Zatem całka Riemanna-Stieltjesa jest uogólnieniem całki Riemanna.

Dla całki Riemanna-Stieltjesa zachodzą następujące równości:

$$\int\limits_a^b (\lambda_1 f_1 + \lambda_2 f_2)\, d\alpha = \lambda_1 \int\limits_a^b f_1\, d\alpha + \lambda_2 \int\limits_a^b f_2\, d\alpha,$$

$$\int\limits_a^b f\, d(\lambda_1 \alpha_1 + \lambda_2 \alpha_2) = \lambda_1 \int\limits_a^b f\, d\alpha_1 + \lambda_2 \int\limits_a^b f\, d\alpha_2,$$

gdzie λ_1, λ_2 są dowolnymi liczbami. Przy czym, jeżeli istnieją całki po stronie prawej tych wzorów, to istnieją również całki po stronie lewej (i zachodzi równość). Ponadto prawdą jest również, że

$$\int\limits_a^b f\, d\alpha = \int\limits_a^c f\, d\alpha + \int\limits_c^b f\, d\alpha \quad \text{dla} \quad a < c < b,$$

gdzie istnienie całki po stronie lewej zapewnia istnienie całek po stronie prawej (ale nie na odwrót!).

Dowody tych równości pozostawiamy jako ćwiczenie dla Czytelnika.

Wygodnie jest posługiwać się ciągową wersją definicji granicy sum Stieltjesa. Aby ją sformułować, wprowadzamy następujące pojęcie:

Ciąg podziałów $(\varPi_n)$ przedziału $[a, b]$ nazywamy *normalnym ciągiem podziałów*, jeżeli $\mu(\varPi_n) \to 0$, gdy $n \to \infty$.

Lemat D.4. *Niech f i α będą dwiema funkcjami rzeczywistymi lub zespolonymi określonymi na przedziale $[a, b]$. Wówczas funkcja f jest całkowalna względem funkcji α wtedy i tylko wtedy, gdy dla dowolnego normalnego ciągu podziałów $(\varPi_n)$ przedziału $[a, b]$ i dowolnego wyboru punktów pośrednich ciąg $(S(f, \alpha; \varPi_n))$ jest zbieżny. Ponadto, jeżeli to ma miejsce, to zachodzi równość*

$$\lim_{n\to\infty} S(f, \alpha; \varPi_n) = \int\limits_a^b f\, d\alpha.$$

Dowód lematu jest analogiczny do dowodu tego, że definicje granicy funkcji w punkcie w sensie Cauchy'ego i w sensie Heinego są równoważne. Dlatego pozostawimy go jako łatwe ćwiczenie dla Czytelnika.

Uwaga. Jeżeli dla dowolnego normalnego ciągu podziałów $(\varPi_n)$ i dowolnego wyboru punktów pośrednich $\xi_j^{(n)}$ ciąg sum Stieltjesa $(S(f, \alpha; \varPi_n))$ jest zbieżny, to granica tego ciągu nie zależy od wyboru ciągu normalnego podziałów ani od wyboru punktów $\xi_j^{(n)}$. Jeżeli bowiem $\lim_{n\to\infty} S(f,\alpha;\Pi_n) = A$ (przy ustalonych punktach pośrednich $\xi_j^{(n)}$) i $\lim_{n\to\infty} S(f,\alpha;\Pi_n') = A'$ (przy ustalonych punktach pośrednich $\xi_j^{\prime(n)}$), to ciąg Π_1, Π_1', Π_2, Π_2', $\ldots$, Π_n, Π_n', $\ldots$ jest również ciągiem normalnym podziałów, a więc ciąg

$$S(f,\alpha;\Pi_1),\ \ S(f,\alpha,\Pi_1'),\ \ S(f,\alpha;\Pi_2),\ \ S(f,\alpha,\Pi_2'),\ldots,$$
$$\ldots,S(f,\alpha;\Pi_n),\ \ S(f,\alpha,\Pi_n'),\ldots$$

jest zbieżny. Oznaczmy jego granicę przez A''. Ponieważ ciągi $(S(f,\alpha;\Pi_n))$ i $(S(f,\alpha,\Pi_n'))$ są podciągami tego ciągu, więc mamy $A = A' = A''$.

Podstawą naszych dalszych rozważań jest następujące twierdzenie:

Twierdzenie D.5. *Jeżeli f jest funkcją ciągłą i α funkcją o ograniczonym wahaniu na przedziale $[a,b]$, to całka Riemanna-Stieltjesa funkcji f względem funkcji α istnieje i ponadto*

$$\left|\int_a^b f\,d\alpha\right| \leqslant \max_{a\leqslant x\leqslant b}|f(x)|\ \mathrm{var}(\alpha;a,b).$$

D o w ó d. Możemy założyć, że funkcja f przyjmuje wartości rzeczywiste. Niech (Π_n) będzie dowolnym ciągiem normalnym podziałów przedziału $[a,b]$. Załóżmy, że jest on postaci

$$\Pi_n\colon\ a = x_0^{(n)} < x_1^{(n)} < \ldots < x_{k(n)}^{(n)} = b.$$

Obierzmy dowolne punkty $\xi_j^{(n)}$ tak, aby $x_{j-1}^{(n)} \leqslant \xi_j^{(n)} \leqslant x_j^{(n)}$. Niech

$$S_n(f) = S(f,\alpha;\Pi_n) = \sum_{j=1}^{k(n)} f(\xi_j^{(n)})(\alpha(x_j^{(n)}) - \alpha(x_{j-1}^{(n)})).$$

Wówczas S_n jest funkcjonałem liniowym na przestrzeni $\mathcal{C}[a,b]$ oraz

$$|S_n(f)| \leqslant \max_{a\leqslant x\leqslant b}|f(x)| \sum_{j=1}^{k(n)} |\alpha(x_j^{(n)}) - \alpha(x_{j-1}^{(n)})| \leqslant \max_{a\leqslant x\leqslant b}|f(x)|\,\mathrm{var}\,(\alpha;a,b).$$

Stąd

$$\|S_n\| \leqslant \mathrm{var}(\alpha;a,b),$$

co oznacza, że funkcjonały S_n są ciągłe oraz ich normy są wspólnie ograniczone. Aby udowodnić, że ciąg $(S_n(f))$ jest zbieżny dla dowolnej funkcji ciągłej f, wystarczy na mocy twierdzenia 3.4 wykazać, że jest on zbieżny dla każdej funkcji należącej do zbioru gęstego w $\mathcal{C}[a,b]$. Ponieważ funkcja ciągła na przedziale $[a,b]$ jest na nim jednostajnie ciągła, więc stąd wynika, że każdą funkcję ciągłą możemy przybliżać w normie przestrzeni $\mathcal{C}[a,b]$ funkcjami łamanymi (tzn. wykresy tych funkcji są łamanymi). Zatem wystarczy pokazać, że ciąg $(S_n(f))$ jest zbieżny dla dowolnej funkcji łamanej f. Ponieważ dla takiej funkcji przedział $[a,b]$ można podzielić na skończoną liczbę podprzedziałów, na których funkcja ta jest liniowa, więc aby zakończyć dowód, wystarczy pokazać, że dla dowolnych punktów c, d takich, że $a \leqslant c < d \leqslant b$, istnieją całki Riemanna-Stieltjesa

$$\int_c^d d\alpha(x) \quad \text{i} \quad \int_c^d x\, d\alpha(x).$$

Pierwsza z tych całek wprost z definicji jest równa $\alpha(d) - \alpha(c)$. Aby wykazać istnienie drugiej, weźmy dowolny podział $\Pi_1\colon\ c = x_0 < x_1 < x_2 < \ldots < x_n = d$ i punkty pośrednie $x_{j-1} \leqslant \xi_j \leqslant x_j$, a następnie utwórzmy sumę Stieltjesa

$$S_1 = \sum_{j=1}^{n} \xi_j(\alpha(x_j) - \alpha(x_{j-1})).$$

Wykażemy, że dla dwóch sum Stieltjesa S_1 i S_2 powyższej postaci mamy

$$|S_1 - S_2| \leqslant (\mu(\Pi_1) + \mu(\Pi_2))\operatorname{var}(\alpha; c, d).$$

Niech Π będzie wspólnym rozdrobnieniem podziałów Π_1 i Π_2 (tzn. Π zawiera wszystkie punkty podziałów Π_1 i Π_2) oraz niech $\Pi\colon\ c = t_0 < t_1 < t_2 \ldots < t_m = d$. Utwórzmy sumę

$$S = \sum_{i=1}^{m} t_i(\alpha(t_i) - \alpha(t_{i-1})).$$

Oznaczając przez $s_1, \ldots, s_k$ te punkty podziału Π, które wpadają do przedziału $[x_{j-1}, x_j]$, mamy

$$|\xi_j(\alpha(x_j) - \alpha(x_{j-1}))| - \sum_{i=1}^{k} s_i(\alpha(s_i) - \alpha(s_{i-1}))| =$$
$$= \left|\sum_{i=1}^{k}(\xi_j - s_i)(\alpha(s_i) - \alpha(s_{i-1}))\right| \leqslant \mu(\Pi_1)\sum_{i=1}^{k}|\alpha(s_i) - \alpha(s_{i-1})|.$$

Dodając stronami nierówności utworzone dla przedziałów $[x_{j-1}, x_j]$ $j = 1, 2, \ldots, n$, otrzymujemy

$$|S_1 - S| \leqslant \mu(\Pi_1)\operatorname{var}(\alpha; c, d).$$

Analogicznie

$$|S_2 - S| \leqslant \mu(\Pi_2)\operatorname{var}(\alpha; c, d).$$

Zatem

$$|S_1 - S_2| \leqslant |S_1 - S| + |S - S_2| \leqslant (\mu(\Pi_1) + \mu(\Pi_2))\operatorname{var}(\alpha; c, d).$$

Z otrzymanej nierówności wynika istnienie całki $\int\limits_c^d x\, d\alpha(x)$, bo dowolny ciąg normalny podziałów (Π_n) daje ciąg sum Stieltjesa (S_n) spełniający warunek Cauchy'ego, a więc zbieżny. ■

Z udowodnionego twierdzenia i z własności całki Riemanna-Stieltjesa wynika, że dla funkcji α o wahaniu ograniczonym wzór

$$F(f) = \int_a^b f\, d\alpha \tag{D.6}$$

określa funkcjonał liniowy i ograniczony na przestrzeni $\mathcal{C}[a,b]$. Pokażemy, że dowolny funkcjonał liniowy i ograniczony na tej przestrzeni jest tej postaci.

Twierdzenie D.6 (Riesza). *Dla dowolnego liniowego i ograniczonego funkcjonału F na przestrzeni $\mathcal{C}[a,b]$ istnieje funkcja α o wahaniu ograniczonym na przedziale $[a,b]$, dla której zachodzi równość* (D.6) *dla dowolnej funkcji ciągłej f oraz $\|F\| = \operatorname{var}(\alpha; a, b)$.*

D o w ó d. Ponieważ przestrzeń $\mathcal{C}[a,b]$ jest podprzestrzenią przestrzeni $B[a,b]$ funkcji ograniczonych na przedziale $[a,b]$, więc na mocy twierdzenia Hahna-Banacha istnieje rozszerzenie $\widetilde{F}$ funkcjonału F zachowujące jego normę. Niech

$$\varphi_x(t) = \begin{cases} 1 & \text{dla } a \leqslant t < x, \\ 0 & \text{dla } x \leqslant t \leqslant b \end{cases}$$

oraz $\varphi_a(t) = 0$ dla $a \leqslant t \leqslant b$. Niech $\alpha(x) = \widetilde{F}(\varphi_x)$ dla $a \leqslant x \leqslant b$. Wykażemy, że funkcja α ma wahanie ograniczone.

Weźmy w tym celu dowolny podział $\Pi\colon\ a = x_0 < x_1 < \ldots < x_n = b$. W dalszym ciągu symbol $\operatorname{sgn} z$ będzie oznaczał liczbę $\dfrac{z}{|z|}$, gdy $z \neq 0$ i 0, gdy $z = 0$. Zauważmy, że w przypadku, gdy z jest liczbą rzeczywistą, $\operatorname{sgn} z$ jest tym samym co znak liczby z.

Mamy

$$\begin{aligned}\sum_{j=1}^{n} |\alpha(x_j) - \alpha(x_{j-1})| &= \sum_{j=1}^{n} (\alpha(x_j) - \alpha(x_{j-1}) \operatorname{sgn}(\alpha(x_j) - \alpha(x_{j-1})) = \\ &= \sum_{j=1}^{n} (\widetilde{F}(\varphi_{x_j}) - \widetilde{F}(\varphi_{x_{j-1}}) \operatorname{sgn}(\alpha(x_j) - \alpha(x_{j-1})) = \\ &= \widetilde{F}\left(\sum_{j=1}^{n} (\varphi_{x_j} - \varphi_{x_{j-1}}) \operatorname{sgn}(\alpha(x_j) - \alpha(x_{j-1}))\right) \leqslant \\ &\leqslant \|F\| \left\| \sum_{j=1}^{n} (\varphi_{x_j} - \varphi_{x_{j-1}}) \operatorname{sgn}(\alpha(x_j) - \alpha(x_{j-1})) \right\|_{\infty} = \|F\|,\end{aligned}$$

bo funkcja $\sum\limits_{j=1}^{n} (\varphi_{x_j} - \varphi_{x_{j-1}}) \operatorname{sgn}(\alpha(x_j) - \alpha(x_{j-1}))$ przyjmuje wartości o module równym 1 lub wartość zero. Zatem funkcja α ma wahanie ograniczone oraz

$$\operatorname{var}(\alpha; a, b) \leqslant \|F\|.$$

Ponieważ $\alpha(a) = \widetilde{F}(\varphi_a) = F(0) = 0$, więc

$$\|\alpha\|_v = |\alpha(a)| + \operatorname{var}(\alpha; a, b) = \operatorname{var}(\alpha; a, b) \leqslant \|F\|.$$

Z poprzedniego twierdzenia otrzymamy nierówność przeciwną, jeżeli pokażemy, że funkcjonał F ma reprezentację całkową. Aby to zrobić, weźmy dowolne $f \in \mathcal{C}[a, b]$. Zdefiniujemy taki ciąg funkcji $f_n \in B[a, b]$, że $f_n \to f$ oraz $\widetilde{F}(f_n)$ będą sumami Stieltjesa dla całki $\int\limits_a^b f\, d\alpha$. Niech $s = b - a$. Definiujemy

$$f_n(t) = \sum_{j=1}^{n} f\left(a + \frac{js}{n}\right) \left(\varphi_{a+js/n}(t) - \varphi_{a+(j-1)s/n}(t)\right).$$

Zauważmy, że f_n jest funkcją „schodkową”, przyjmującą wartość $f\left(a + \frac{js}{n}\right)$ na przedziale $a + \frac{(j-1)s}{n} < t \leqslant a + \frac{js}{n}$, $j = 1, 2, \ldots, n$ i $f_n(a) = f(a)$. Z jednostajnej ciągłości funkcji f wynika, że ciąg (f_n) dąży do niej jednostajnie na przedziale $[a,b]$. Ponadto mamy

$$\begin{aligned}\widetilde{F}(f_n) &= \sum_{j=1}^{n} f\left(a+\frac{js}{n}\right)\left(\widetilde{F}\left(\varphi_{a+js/n}\right)-\widetilde{F}\left(\varphi_{a+(j-1)s/n}\right)\right)= \\ &= \sum_{j=1}^{n} f\left(a+\frac{js}{n}\right)\left(\alpha\left(a+\frac{js}{n}\right)-\alpha\left(a+\frac{(j-1)s}{n}\right)\right).\end{aligned}$$

Zatem $(\widetilde{F}(f_n))$ jest ciągiem sum Stieltjesa dla funkcji f i α. Na mocy twierdzenia D.5 mamy $\widetilde{F}(f_n) \to \int_a^b f\,d\alpha$. Z ciągłości funkcjonału $\widetilde{F}$ wynika, że $\widetilde{F}(f_n) \to \widetilde{F}(f) = F(f)$. Z jednoznaczności granicy otrzymujemy równość

$$F(f) = \int_a^b f\,d\alpha.$$

Z twierdzenia D.5 wynika również, że

$$\|F\| \leqslant \operatorname{var}(\alpha; a, b) = \|\alpha\|_v. \qquad \blacksquare$$

Zauważmy, że funkcja o ograniczonym wahaniu α, odpowiadająca funkcjonałowi liniowemu F we wzorze (D.6), nie jest wyznaczona jednoznacznie. Po pierwsze, jest oczywiste, że jeżeli dodamy do funkcji α dowolną stałą, to wzór (D.6) pozostanie nadal prawdziwy. Po drugie, jeżeli x_0 jest punktem wewnętrznym przedziału $[a,b]$ i zdefiniujemy funkcję β na przedziale $[0,1]$ za pomocą wzoru: $\beta(x) = \alpha(x)$ dla $x \neq x_0$ oraz $\beta(x_0) = \alpha(x_0+)$, to otrzymamy funkcję o ograniczonym wahaniu, dla której

$$\text{(D.7)} \qquad \int_a^b f\,d\alpha = \int_a^b f\,d\beta \quad \text{dla} \quad f \in \mathcal{C}[a,b].$$

Zatem, jeżeli chcielibyśmy wyznaczyć przestrzeń sprzężoną z przestrzenią $\mathcal{C}[a,b]$, to należałoby wprowadzić w przestrzeni $BV[a,b]$ relację równoważności w następujący sposób: $\alpha \sim \beta$ wtedy i tylko wtedy, gdy zachodzi wzór (D.7), a następnie rozważać przestrzeń ilorazową. Wygodniej jest wybrać z każdej klasy abstrakcji znormalizowanego reprezentanta i rozważać podprzestrzeń utworzoną z takich funkcji. Najpierw udowodnimy następujące twierdzenie:

Twierdzenie D.7. *Jeżeli α jest funkcją o wahaniu ograniczonym na przedziale $[a, b]$, a β jest zdefiniowane wzorami: $\beta(x) = \alpha(x+)$ dla $x \in (a, b)$, $\beta(a) = \alpha(a)$ oraz $\beta(b) = \alpha(b)$, to β jest również funkcją o wahaniu ograniczonym,* $\mathrm{var}(\beta; a, b) \leqslant \mathrm{var}(\alpha; a, b)$ *oraz dla dowolnej funkcji $f \in \mathcal{C}[a, b]$ zachodzi wzór* (D.7).

D o w ó d. Niech Π: $a = x_0 < x_1 < \ldots < x_n = b$ będzie dowolnym podziałem przedziału $[a, b]$. Aby wykazać, że funkcja β ma ograniczone wahanie i $\mathrm{var}(\beta; a, b) \leqslant \mathrm{var}(\alpha; a, b)$, wystarczy udowodnić, że

$$\sum_{j=1}^{n} |\beta(x_j) - \beta(x_{j-1})| \leqslant \mathrm{var}(\alpha; a, b).$$

Niech dane będzie $\varepsilon > 0$. Wybieramy tak punkty t_j, $j = 1, 2, \ldots, n-1$, aby

$$a = x_0 < x_1 < t_1 < x_2 < t_2 < \ldots < x_{n-1} < t_{n-1} < x_n = b$$

oraz

$$|\alpha(t_j) - \alpha(x_j+)| < \frac{\varepsilon}{2n} \quad \text{dla} \quad j = 1, 2, \ldots, n-1.$$

Wówczas

$$\begin{aligned}
\sum_{j=1}^{n} |\beta(x_j) - \beta(x_{j-1})| &= |\alpha(x_1+) - \alpha(a)| + \\
&\quad + \sum_{j=2}^{n-1} |\alpha(x_j+) - \alpha(x_{j-1}+)| + |\alpha(b) - \alpha(x_{n-1}+)| \leqslant \\
&\leqslant |\alpha(x_1+) - \alpha(t_1)| + |\alpha(t_1) - \alpha(a)| + \sum_{j=2}^{n-1} |\alpha(x_j+) - \alpha(t_j)| + \\
&\quad + \sum_{j=2}^{n-1} |\alpha(t_j) - \alpha(t_{j-1})| + \sum_{j=2}^{n-1} |\alpha(t_{j-1}) - \alpha(x_{j-1}+)| + \\
&\quad + |\alpha(b) - \alpha(t_{n-1})| + |\alpha(t_{n-1}) - \alpha(x_{n-1}+)| = \\
&= |\alpha(t_1) - \alpha(a)| + \sum_{j=2}^{n-1} |\alpha(t_j) - \alpha(t_{j-1})| + |\alpha(b) - \alpha(t_{n-1})| + \\
&\quad + \sum_{j=1}^{n-1} |\alpha(x_j+) - \alpha(t_j)| + \sum_{j=2}^{n} |\alpha(t_{j-1}) - \alpha(x_{j-1}+)| < \\
&< \mathrm{var}(\alpha; a, b) + \varepsilon.
\end{aligned}$$

Z dowolności $\varepsilon > 0$ wynika, że β jest funkcją o żądanych własnościach. Ponieważ całka $\int_a^b f\,d\beta$ istnieje, więc jest ona granicą ciągu sum Stieltjesa, odpowiadającego dowolnemu ciągowi normalnemu podziałów przedziału $[a,b]$. Z wniosku D.2 wynika, że zbiór punktów prawostronnej nieciągłości funkcji α w przedziale (a,b) jest co najwyżej przeliczalny. Możemy zatem brać ciągi normalne podziałów niezawierające tych punktów. Dla takiego podziału Π_n sumy Stieltjesa $S(f,\alpha;\Pi_n)$ i $S(f,\beta;\Pi_n)$ są równe. Stąd wynika, że i całki $\int_a^b f\,d\alpha$, $\int_a^b f\,d\beta$ są równe. ■

Definicja D.4. Symbolem $NBV[a,b]$ oznaczymy zbiór wszystkich funkcji φ o wahaniu ograniczonym na przedziale $[a,b]$ prawostronnie ciągłych w jego wnętrzu i takich, że $\varphi(a) = 0$. Funkcje należące do tej przestrzeni będziemy nazywali funkcjami *znormalizowanymi o wahaniu ograniczonym.*

Wniosek D.8. *Przestrzeń $NBV[a,b]$ jest przestrzenią Banacha z normą $\|\varphi\|_v = \mathrm{var}(\varphi;a,b)$.*

D o w ó d. Teza wynika z oczywistego faktu, że $NBV[a,b]$ jest domkniętą podprzestrzenią przestrzeni $BV[a,b]$. ■

Teraz możemy zidentyfikować przestrzeń sprzężoną z przestrzenią $\mathcal{C}[a,b]$.

Twierdzenie D.9. *Przekształcenie A, przyporządkowujące każdej znormalizowanej funkcji o wahaniu ograniczonym α funkcjonał liniowy ograniczony F zdefiniowanym wzorem* (D.6)*, jest liniową izometrią przestrzeni $NBV[a,b]$ na przestrzeń sprzężoną z przestrzenią $\mathcal{C}[a,b]$.*

D o w ó d. Liniowość przekształcenia A wynika z własności całki Riemanna-Stieltjesa. Z twierdzenia Riesza wiemy, że dla danego funkcjonału F istnieje taka funkcja $\alpha \in BV[a,b]$, że $F(f) = \int_a^b f\,d\alpha$, $\alpha(a) = 0$ i $\|F\| = \mathrm{var}(\alpha;a,b)$. Z twierdzenia D.7 wynika, że istnieje taka funkcja $\beta \in NBV[a,b]$, że $F(f) = \int_a^b f\,d\beta$ oraz $\mathrm{var}(\beta;a,b) \leqslant \mathrm{var}(\alpha;a,b)$. Na mocy twierdzenia D.6 wiemy, że wówczas $\|F\| \leqslant \mathrm{var}(\beta;a,b)$. Ostatecznie więc mamy

$$\|F\| \leqslant \mathrm{var}(\beta;a,b) \leqslant \mathrm{var}(\alpha;a,b) = \|F\|,$$

co oznacza, że przekształcenie A jest izometrią na przestrzeń sprzężoną z przestrzenią $\mathcal{C}[a,b]$. ■

U w a g a. Nie jest istotne w twierdzeniu D.7, że znormalizowaliśmy funkcje z przestrzeni $BV[a,b]$, żądając, aby były one prawostronnie ciągłe w przedziale (a,b). Równie dobrze moglibyśmy wybierać funkcje lewostronnie ciągłe w tym przedziale lub brać funkcje z klasy abstrakcji funkcji α równe $\frac{1}{2}(\alpha(x-) + \alpha(x+))$ w każdym punkcie $x \in (a,b)$ i twierdzenie to nadal pozostałoby prawdziwe.

Literatura

Podręczniki

[1] N. I. Akhiezer, I. M. Glazman, *Theory of Linear Operators in Hilbert Space*, vol. 1, Frederick Ungar, New York 1961.

[2] A. Alexiewicz, *Analiza funkcjonalna*, PWN, Warszawa 1969.

[3] G. Bachman, L. Narici, *Functional Analysis*, Academic Press, New York 1966.

[4] C. Bosch, C. Swartz, *Functional Calculi*, World Scientific, New Jersey 2013.

[5] A. L. Brown, A. Page, *Elements of Functional Analysis*, Van Nostrand Reinhold Company, London 1970.

[6] J. B. Conway, *A Course in Functional Analysis*, Graduate Texts in Mathematics 96, Springer-Verlag, New York 1997.

[7] R. G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press, New York 1972.

[8] N. Dunford, J. T. Schwartz, *Linear Operators*, Part I i Part II, Interscience Publishers, New York 1958 (Part I) i 1963 (Part II).

[9] C. Goffman, G. Pedrick, *First Course in Functional Analysis*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1965.

[10] P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea Publishing Company, New York 1951.

[11] G. Helmberg, *Introduction to spectral theory in Hilbert space*, North-Holland Publishing Co., Amsterdam 1969.

[12] V. C. L. Hutson, J. S. Pym, *Applications of Functional Analysis and Operator Theory*, Academic Press, London 1980.

[13] A. N. Kolmogorov, S. V. Fomin, *Elementy teorii funkcij i funkcional'nogo analiza*, Nauka, Moskva 1976 (w języku rosyjskim).

[14] W. Kołodziej, *Wybrane rozdziały analizy matematycznej*, PWN, Warszawa 1982.

[15] C. S. Kubrusly, *Spectral Theory of Operators on Hilbert Spaces*, Springer, New York 2012.

[16] S. Łojasiewicz, *Wstęp do teorii funkcji rzeczywistych*, PWN, Warszawa 1973.

[17] L. A. Lusternik, W. I. Sobolew, *Elementy analizy funkcjonalnej*, PWN, Warszawa 1959.

[18] K. Maurin, *Metody przestrzeni Hilberta*, PWN, Warszawa 1959.

[19] W. Mlak, *Wstęp do teorii przestrzeni Hilberta*, PWN, Warszawa 1987.
[20] J. Musielak, *Wstęp do analizy funkcjonalnej*, PWN, Warszawa 1989.
[21] F. Riesz, B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar, New York 1955.
[22] W. Rudin, *Podstawy analizy matematycznej*, wyd. 1, PWN, Warszawa 1969.
[23] W. Rudin, *Analiza rzeczywista i zespolona*, Wydawnictwo Naukowe PWN, Warszawa 1998.
[24] W. Rudin, *Analiza funkcjonalna*, Wydawnictwo Naukowe PWN, Warszawa 2001.
[25] G. F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw Hill, New York 1963.
[26] A. Sołtysiak, *Algebra liniowa*, Wydawnictwo Naukowe UAM, Poznań 2003.
[27] A. Sołtysiak, *Analiza matematyczna*, części I, II i III, Wydawnictwo Naukowe UAM, Poznań 2009, 2004 i 2000.
[28] G. E. Šilov, *Matematičeskij analiz. Funkcji odnogo peremennogo*, č. 3, Nauka, Moskva 1970 (w języku rosyjskim).
[29] A. E. Taylor, *Introduction to Functional Analysis*, John Wiley & Sons, Inc., New York 1967.

Zbiory zadań

[30] A. B. Antonevič, P. N. Kniazev, Ja. V. Radyno, *Zadači i uprazhnenija po funkcjonal'nomu analizu*, Vyšejšaja Škola, Minsk 1978 (w języku rosyjskim).
[31] J. Górniak, T. Pytlik, *Analiza funkcjonalna w zadaniach*, Wydawnictwo Politechniki Wrocławskiej, Wrocław 1976.
[32] P.R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York 1982.
[33] A. A. Kirillov, A. D. Gvišiani, *Teoremy i zadači funkcjonal'nogo analiza*, Nauka, Moskva 1988 (w języku rosyjskim).
[34] S. Prus, A. Stachura, *Analiza funkcjonalna w zadaniach*, Wydawnictwo Naukowe PWN, Warszawa 2007.
[35] K. Rudol, *Zbiór zadań z analizy funkcjonalnej. Część I: Przestrzenie*, AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków 2008.

Skorowidz symboli

Skorowidz nazw

K

L

M

N

O

P

R

Z

ISBN 978-83-232-2970-4