

WYKŁADY Z ANALIZY FUNKCJONALNEJ

Bartosz Kosma Kwaśniewski

3 marca 2025

Spis treści

1	Przestrzenie Banacha i operatory ograniczone	2
1.1	Przestrzenie Banacha	2
1.2	Klasyczne przestrzenie Banacha.	7
1.2.1	Przestrzenie funkcji ciągłych	7
1.2.2	Przestrzenie L^p dla $p \geq 1$	11
1.2.3	Przestrzenie L^∞	15
1.3	Operatory ograniczone	17
1.3.1	Przykłady operatorów ograniczonych	20
1.3.2	Równoważność norm i przestrzenie skończone wymiarowe	22
1.3.3	Przestrzenie operatorów ograniczonych	24
2	Przestrzenie Hilberta	26
2.1	Iloczyn skalarny i norma	26
2.2	Rzut ortogonalny	31
2.3	Układy i bazy ortonormalne	37
2.4	Operatory sprzężone	45
2.5	Zadania	51
3	Trzy zasady analizy funkcjonalnej	52
3.1	Twierdzenie Hahna-Banacha	52
3.2	Przykłady przestrzeni sprzężonych. Przestrzenie refleksywne	56
3.3	Twierdzenie Banacha-Steinhaus	60
3.4	Topologia słaba i *-słaba	63
3.5	Twierdzenie Banacha o odwzorowaniu otwartym	68
4	Teoria spektralna	73
4.1	Algebry Banacha i elementy odwracalne	73
4.2	Widmo i promień spektralny	78
4.3	Analityczny rachunek funkcyjny	84
4.4	Twierdzenie spektralne	84

Rozdział 1

Przestrzenie Banacha i operatory ograniczone

1.1 Przestrzenie Banacha

Przypomnijmy, iż elementy przestrzeni liniowej (wektorowej) X nazywamy wektorami. Wektory możemy dodawać i mnożyć przez skalar. Przez skalary będziemy tu zawsze rozumieć liczby rzeczywiste $\mathbb{R}$ lub liczby zespolone $\mathbb{C}$. Przestrzeń X wraz z dodawaniem wektorów tworzy grupę przemianą, a jej element neutralny nazywamy wektorem zerowym i oznaczmy $0 \in X$.

Definicja 1.1. *Przestrzenią unormowaną* nazywamy przestrzeń liniową X nad ciałem $\mathbb{F} := \mathbb{R}, \mathbb{C}$ wyposażoną w *normę*, tzn. funkcję $\|\cdot\| : X \rightarrow [0, \infty)$ spełniającą następujące warunki

$$(N1) \quad \|x\| = 0 \iff x = 0, \quad (\text{niezdegenerowanie})$$

$$(N2) \quad \|\lambda x\| = |\lambda| \cdot \|x\|, \quad (\text{dodatnia jednorodność})$$

$$(N3) \quad \|x + y\| \leq \|x\| + \|y\|, \quad (\text{nierówność trójkąta})$$

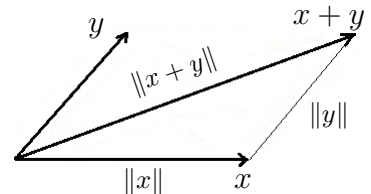
gdzie $x, y \in X$ i $\lambda \in \mathbb{F}$. Liczbę $\|x\|$ nazywamy *normą (długością) wektora* $x \in X$.

Uwaga 1.2. Norma zadaje wzorem $d(x, y) := \|x - y\|$, $x, y \in X$, metrykę (🏠 sprawdzić). Zatem przestrzeń liniowa jest zarówno przestrzenią liniową jak i przestrzenią topologiczną. Zbiory otwarte w X są sumami kul otwartych, gdzie kulą otwartą o środku w $x \in X$ i promieniu $r > 0$ nazywamy zbiór

$$K(x, r) := \{y \in X : \|x - y\| < r\}.$$

Ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ jest zbieżny do $x \in X$ wtedy i tylko wtedy, gdy $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$. Piszemy wtedy $x_n \rightarrow x$, $x_n \xrightarrow{X} x$ lub $x_n \xrightarrow{\|\cdot\|} x$.

Warunek trójkąta (N3) ma tu następującą interpretację geometryczną: suma długości dwóch boków trójkąta jest nie mniejsza, niż długość boku pozostałego. (patrz rysunek po prawej, dla trójkąta o wierzchołkach 0 , x i $x + y$). Implikuje on, że różnica długości dwóch boków trójkąta jest nie większa, niż długość boku pozostałego, co można zapisać następująco:



Lemat 1.3. *Dla dowolnych wektorów przestrzeni unormowanej zachodzi*

$$||\|x\| - \|y\|| \leq \|x - y\|. \quad (\text{odwrotna nierówność trójkąta})$$

Dowód. Korzystając z nierówności trójkąta mamy $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|$, skąd $\|x\| - \|y\| \leq \|x - y\|$. Analogicznie $\|y\| = \|(y - x) + x\| \leq \|y - x\| + \|x\|$, skąd $\|y\| - \|x\| \leq \|y - x\|$. $\square$

Struktura liniowa w przestrzeni unormowanej jest zgodna z jej strukturą topologiczną, w tym sensie, że operacje liniowe są ciągłe:

Stwierdzenie 1.4. *W przestrzeni unormowanej mnożenie przez skalar, dodawanie wektorów i norma są funkcjami ciągłymi.*

Dowód. Interesują nas funkcje „mnożenie przez skalar” $\cdot : \mathbb{F} \times X \rightarrow X$, „dodawanie wektorów” $+: X \times X \rightarrow X$ oraz „norma” $\|\cdot\| : X \rightarrow [0, +\infty)$. Niech $\lambda_n \rightarrow \lambda$, $x_n \rightarrow x$ i $y_n \rightarrow y$, tzn. $|\lambda_n - \lambda| \rightarrow 0$, $\|x_n - x\| \rightarrow 0$ i $\|y_n - y\| \rightarrow 0$. Wtedy

$$\begin{aligned} \|\lambda_n x_m - \lambda x\| &\stackrel{(N3)}{\leq} \|\lambda_n x_m - \lambda_n x\| + \|\lambda_n x - \lambda x\| \\ &\stackrel{(N2)}{=} |\lambda_n| \cdot \|x_m - x\| + |\lambda_n - \lambda| \cdot \|x\| \longrightarrow 0, \quad \text{przy } n, m \rightarrow \infty. \end{aligned}$$

Zatem $\lambda_n x_m \rightarrow \lambda x$, co dowodzi ciągłości mnożenia przez skalar. Podobnie

$$\|(x_n + y_m) - (x + y)\| \stackrel{(N3)}{\leq} \|x_n - x\| + \|y_m - y\| \longrightarrow 0, \quad \text{przy } n, m \rightarrow \infty,$$

dowodzi ciągłości dodawania wektorów. Natomiast ciągłość normy wynika łatwo z „odwrotnej nierówności trójkąta”. Rzeczywiście $\|\|x_n\| - \|x\|\| \stackrel{\text{Lemat 1.3}}{\leq} \|x_n - x\| \rightarrow 0$, skąd $\|x_n\| \rightarrow \|x\|$. $\square$

Przypomnijmy, że ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ w przestrzeni metrycznej X nazywamy *ciągą Cauchy* jeżeli $\forall \varepsilon \exists N \forall n, m \geq N d(x_n, x_m) < \varepsilon$, czyli gdy $d(x_n, x_m) \rightarrow 0$ przy $n, m \rightarrow \infty$. Natomiast przestrzeń metryczną nazywamy *zupełną*, jeżeli każdy jej ciąg Cauchy jest zbieżny (posiada granicę). Teorię przestrzeni unormowanych zupełnych rozwinął Stefan Banach i stąd nazywamy je dziś przestrzeniami Banacha:

Definicja 1.5. *Przestrzenią Banacha* nazywamy przestrzeń unormowaną zupełną, tzn. przestrzeń $(X, \|\cdot\|)$, w której zachodzi

$$\forall \{x_n\}_{n=1}^\infty \subseteq X \quad \underbrace{\lim_{n, m \rightarrow \infty} \|x_n - x_m\| = 0}_{\{x_n\}_{n=1}^\infty \text{ ciąg Cauchy}} \implies \underbrace{\exists x_0 \in X \lim_{n \rightarrow \infty} \|x_n - x_0\| = 0}_{\{x_n\}_{n=1}^\infty \text{ ciąg zbieżny}}.$$

Uwaga 1.6. Każdy ciąg zbieżny jest ciągiem Cauchy (🏠 pokazać). Zatem w przestrzeniach zupełnych ciąg jest zbieżny wtedy i tylko wtedy, gdy jest ciągiem Cauchy.

Przykład 1.7. Zbiór liczb rzeczywistych $\mathbb{R}$ jest niemal z definicji przestrzenią zupełną - jest to rozszerzenie (uzupełnienie) zbioru liczb wymiernych $\mathbb{Q}$, jako przestrzeni metrycznej, do przestrzeni zupełnej, por. Twierdzenie 1.11 poniżej. Zatem $\mathbb{R}$ jest jednowymiarową przestrzenią Banacha z modułem $|t|$, $t \in \mathbb{R}$, jako normą.

Przykład 1.8 (Przestrzeń euklidesowa). Każda skończenie wymiarowa przestrzeń liniowa na ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$ jest izomorficzna z $X = \mathbb{F}^n$, gdzie dodawania wektorów i mnożenie przez skalar jest zdefiniowane po współrzędnych:

$$x + y := (x(1) + y(1), x(2) + y(2), \dots, x(n) + y(n)), \quad \lambda x := (\lambda x(1), \lambda x(2), \dots, \lambda x(n)),$$

gdzie $x = (x(1), \dots, x(n))$, $y = (y(1), \dots, y(n)) \in X$ i $\lambda \in \mathbb{F}$. Jedną z najbardziej standardowych norm na X jest *norma euklidesowa*:

$$\|x\|_2 := \sqrt{\sum_{k=1}^n |x(k)|^2}$$

Zauważmy, że przestrzeń $(\mathbb{R}^n, \|\cdot\|_2)$ to dobrze znana z analizy *przestrzeń euklidesowa*. W szczególności, z analizy wiemy, że jest to przestrzeń zupełna (przestrzeń Banacha). Jej zespolony odpowiednik $(\mathbb{C}^n, \|\cdot\|_2)$ można również traktować jako rzeczywistą przestrzeń euklidesową. Mianowicie przyporządkowanie wektorowi zespolonemu $(x(1), \dots, x(n)) \in \mathbb{C}^n$ wektora rzeczywistego $(a(1), b(1), \dots, a(n), b(n)) \in \mathbb{R}^{2n}$, gdzie $x(k) = a(k) + ib(k)$, zadaje $\mathbb{R}$ -liniowy izomorfizm $\mathbb{C}^n \cong \mathbb{R}^{2n}$, który zachowuje normę 🏠. W szczególności, $(\mathbb{C}^n, \|\cdot\|_2)$ jest również przestrzenią Banacha.

Uwaga 1.9. Jak pokażemy poniżej każda unormowana przestrzeń skończenie wymiarowa jest przestrzenią Banacha, patrz Wniosek 1.69. Zatem jedynie przestrzenie nieskończenie wymiarowe mogą nie być zupełne.

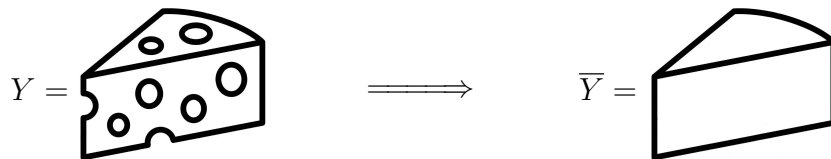
Poniższe stwierdzenie pozwala łatwo konstruować przykłady unormowanych przestrzeni niezupełnych, a ponadto pozwala interpretować „zupełność” jako „domkniętość w nadprzestrzeni”. Dokładniej, jeśli $Y \subseteq X$ jest podprzestrzenią liniową przestrzeni unormowanej X , to będziemy traktować Y jako przestrzeń unormowaną z normą z przestrzeni X (obciętą do podprzestrzeni Y). Tak rozumiana *podprzestrzeń unormowana* $Y \subseteq X$ musi być domkniętym podzbiorem X , jeśli jest zupełna. Implikacja odwrotna zachodzi jeśli X jest zupełna:

Stwierdzenie 1.10. *Podprzestrzeń liniowa Y przestrzeni Banacha X jest przestrzenią Banacha $\iff$ przestrzeń Y jest domknięta w X (implikacja „ $\implies$ ” zachodzi bez założenia o zupełności X).*

Dowód. „ $\implies$ ”. Przypomnijmy, że zbiór Y jest domknięty wtedy i tylko wtedy, gdy zawiera swoje punkty skupienia.¹ Niech $y \in X$ będzie punktem skupienia zbioru Y . Wtedy istnieje ciąg $\{y_n\}_{n=1}^\infty \subseteq Y$ zbieżny do y . Skoro $\{y_n\}_{n=1}^\infty$ zbieżny, to $\{y_n\}_{n=1}^\infty$ jest ciągiem Cauchy w Y . Zatem z zupełności Y , ciąg $\{y_n\}_{n=1}^\infty$ posiada granicę w Y . Czyli $y \in Y$ (z jednoznaczności granicy).

„ $\impliedby$ ”. Każdy ciąg Cauchy $\{y_n\}_{n=1}^\infty \subseteq Y$ w Y jest ciągiem Cauchy w X , a zatem jest zbieżny w X . To znaczy, istnieje $y \in X$ taki, że $y_n \xrightarrow{\|\cdot\|} y$. Ale skoro $\{y_n\}_{n=1}^\infty \subseteq Y$ i $Y = \overline{Y}$, to $y \in Y$. Zatem Y jest zupełna. $\square$

Pokażemy teraz, że każdą przestrzeń unormowaną można zanurzyć w pewną przestrzeń Banacha. Zatem powyższe stwierdzenie rzeczywiście pozwala scharakteryzować zupełność za pomocą domkniętości. Co więcej każdą przestrzeń Y można domknąć wybranej nadprzestrzeni Banacha X , otrzymując w ten sposób przestrzeń Banacha $\overline{Y} \subseteq X$, a tak otrzymane uzupełnienie przestrzeni Y z dokładnością do izomorfizmu nie zależy od wyboru nadprzestrzeni X . Tę procedurę uzupełniania czasami porównuje się do zaklejania dziur w serze:



Jest to jeden ze sposobów konstrukcji przestrzeni Banacha, który formalnie opisuje następujące twierdzenie:

¹Na ogół $\overline{Y} = Y \cup Y^d$, gdzie Y^d to zbiór punktów skupienia zbioru Y

Twierdzenie 1.11 (Uzupełnienie przestrzeni unormowanych). *Każdą przestrzeń unormowaną można „uzupełnić do przestrzeni Banacha”. To znaczy jeśli $(Y, \|\cdot\|)$ jest przestrzenią unormowaną, to istnieją*

- (1) *przestrzeń Banacha $(X, \|\cdot\|)$ - którą będziemy nazywać uzupełnieniem przestrzeni $(Y, \|\cdot\|)$;*
- (2) *liniowa izometria $\Psi : (Y, \|\cdot\|) \rightarrow (X, \|\cdot\|)$ taka, że $\overline{\Psi(Y)} = X$.*

Czyli Y zanurza się w X jako gęsta podprzestrzeń. Ponadto, uzupełnienie $(X, \|\cdot\|)$ jest wyznaczone jednoznacznie z dokładnością do kanonicznego izometrycznego izomorfizmu.

Dowód. Elementy przestrzeni X zdefiniujemy jako klasy abstrakcji ciągów Cauchy $\{x_n\}_{n=1}^\infty$, $\{y_n\}_{n=1}^\infty \subseteq Y$ względem następującej relacji równoważności

$$\{x_n\} \sim \{y_n\} \stackrel{\text{def}}{\iff} \lim_{n \rightarrow \infty} \|x_n - y_n\| = 0.$$

Odpowiednią klasę równoważności oznaczamy $[\{x_n\}]$. Na X wprowadzamy strukturę przestrzeni unormowanej wzorami

$$[\{x_n\}] + [\{y_n\}] := [\{x_n + y_n\}] \quad \lambda[\{x_n\}] := [\{\lambda x_n\}] \quad \|[\{x_n\}]\| := \lim_{n \rightarrow \infty} \|x_n\|.$$

Dodawanie i mnożenie przez skalar są poprawnie określone, ponieważ jeśli $\{x_n\} \sim \{x'_n\}$ i $\{y_n\} \sim \{y'_n\}$, to $\|x_n + y_n - (x'_n + y'_n)\| \leq \|x_n + y_n\| + \|x'_n + y'_n\| \rightarrow 0$, czyli $\{x_n + y_n\} \sim \{x'_n + y'_n\}$, oraz $\|\lambda x_n - \lambda x'_n\| \leq |\lambda| \|x_n - x'_n\| \rightarrow 0$, skąd $\{\lambda x_n\} \sim \{\lambda x'_n\}$. Żeby wykazać poprawność definicji normy zauważmy, że dla $[\{x_n\}] \in X$ z odwrotnej nierówności trójkąta mamy

$$|\|x_n\| - \|x_m\|| \leq \|x_n - x_m\| \rightarrow 0, \quad \text{przy } n, m \rightarrow \infty.$$

Czyli ciąg liczbowy $\{\|x_n\|\}_{n=1}^\infty$ jest Cauchy w $\mathbb{R}$. Zatem jego granica istnieje z zupełności $\mathbb{R}$. Ta granica nie zależy od wyboru reprezentanta, bo jeśli $\{x_n\} \sim \{x'_n\}$, to $|\|x_n\| - \|x'_n\|| \leq \|x_n - x'_n\| \rightarrow 0$, czyli $\lim_{n \rightarrow \infty} \|x_n\| = \lim_{n \rightarrow \infty} \|x'_n\|$.

Aksjomaty przestrzeni wektorowej i normy są łatwe do sprawdzenia 🏠. W szczególności

$$\|[\{x_n\}]\| = 0 \iff \lim_{n \rightarrow \infty} \|x_n\| = 0 \iff x_n \xrightarrow{\|\cdot\|} 0 \iff \{x_n\} \sim \{0\} \iff [\{x_n\}] = [(0, 0, 0, \dots, 0)].$$

Elementy Y zmożemy utożsamić z klasami abstrakcji ciągów stałych. Formalnie, mamy naturalną liniową izometrię

$$Y \ni x_0 \xrightarrow{\Psi} [(x_0, x_0, \dots, x_0)] \in X.$$

Obraz $\Psi(Y)$ jest gęsty w X ponieważ punkt $[\{x_n\}] \in X$ jest granicą ciągu $\{\Psi(x_m)\}_{m=1}^\infty$:

$$\lim_{m \rightarrow \infty} \|[\{x_n\}] - \Psi(x_m)\| = \lim_{m \rightarrow \infty} \|[\{x_n - x_m\}]\| = \lim_{m, n \rightarrow \infty} \|x_n - x_m\| = 0.$$

Żeby pokazać zupełność X rozważmy ciąg Cauchy $\{\tilde{x}_n\}_{n=1}^\infty \subseteq X$ w X . Dla każdego $n \in \mathbb{N}$ możemy znaleźć $x_n \in Y$ takie, że $\|\tilde{x}_n - \Psi(x_n)\| < \frac{1}{n}$. Wtedy ciąg $\{x_n\}_{n=1}^\infty$ jest Cauchy w Y ponieważ $\|x_n - x_m\| = \|\Psi(x_n) - \Psi(x_m)\| \leq \|\Psi(x_n) - \tilde{x}_n\| + \|\tilde{x}_n - \tilde{x}_m\| + \|\tilde{x}_m - \Psi(x_m)\| \rightarrow 0$, przy $n, m \rightarrow \infty$. Ponadto

$$\|\tilde{x}_n - [\{x_n\}]\| \leq \|\tilde{x}_n - \Psi(x_n)\| + \|\Psi(x_n) - [\{x_n\}]\| \rightarrow 0,$$

czyli $\{\tilde{x}_n\}_{n=1}^\infty$ jest zbieżny w X do $[\{x_n\}] \in X$.

Jeśli $\Phi : Y \rightarrow X'$ jest dowolną liniową izometrią w przestrzeń Banacha X' taką, że $\overline{\Phi(Y)} = X'$, to wzór $T(\Psi(x)) := \Phi(x)$ definiuje liniową izometrię z gęstej podprzestrzeni $\Psi(Y) \subseteq X$ na gęstą podprzestrzeń $\Phi(Y) \subseteq X'$. Izometria T przedłuża się do izometrycznego izomorfizmu $T : X \rightarrow X'$. Rzeczywiście, dla każdego $x \in X$ istnieje $\{x_n\} \subseteq \Psi(Y)$ ciąg zbieżny do x i wtedy ciąg Tx_n jest ciągiem Cauchy w X' . Zatem istnieje granica $Tx := \lim_{n \rightarrow \infty} Tx_n$. Ponadto, na mocy ciągłości normy mamy

$$\|Tx\| = \lim_{n \rightarrow \infty} \|Tx_n\| = \lim_{n \rightarrow \infty} \|x_n\| = \|x\|.$$

Zatem w ten sposób otrzymujemy poprawnie określoną liniową izometrię $T : X \rightarrow X'$ (liniowość wynika natychmiast z liniowości granicy). Ponadto, skoro X jest przestrzenią Banacha, to jej izometrycznie izomorficzny obraz $T(X)$ też jest przestrzenią Banacha, a więc jest domknięty w X' . Stąd $T(X) = X'$. Czyli $T : X \rightarrow X'$ jest odwracalną izometrią $\square$

Wniosek 1.12. *Przestrzeń unormowana Y jest zupełna wtedy i tylko wtedy, gdy jej obraz $\Psi(Y)$ przy dowolnej liniowej izometrii $\Psi : Y \rightarrow X$ w przestrzeń unormowaną X jest domknięty w X .*

Dowód. Zauważmy, że dla dowolnej liniowej izometrii $\Psi : Y \rightarrow X$ przestrzeń Y jest zupełna wtedy i tylko wtedy, gdy przestrzeń $\Psi(Y)$ jest zupełna. Jeśli to zachodzi, to $\Psi(Y)$ jest domkniętą podprzestrzenią X na mocy Stwierdzenia 1.10. Na odwrót na mocy Twierdzenia 1.11 istnieje liniowa izometria $\Psi : Y \rightarrow X$, gdzie X jest przestrzenią zupełną. Zatem jeżeli $\Psi(Y)$ jest domknięty w X , to $\Psi(Y) \cong Y$ jest przestrzenią zupełną, znowu na mocy Stwierdzenia 1.10. $\square$

Na koniec przytoczymy jeszcze jedną charakteryzację zupełności wiążącą ją z absolutną zbieżnością szeregów w przestrzeniach unormowanych.

Definicja 1.13. Dla ciągu $\{x_n\}_{n=1}^{\infty} \subseteq X$ w przestrzeni unormowanej X powiemy, że *szereg* $\sum_{n=1}^{\infty} x_n$ *jest zbieżny*, jeżeli ciąg sum częściowych $\{\sum_{n=1}^N x_n\}_{N=1}^{\infty}$ jest zbieżny w X i wtedy granicę $\sum_{n=1}^{\infty} x_n := \lim_{N \rightarrow \infty} \sum_{n=1}^N x_n$ nazywamy sumą szeregu. Powiemy, że *szereg* $\sum_{n=1}^{\infty} x_n$ *jest absolutnie zbieżny* jeżeli szereg liczbowy $\sum_{n=1}^{\infty} \|x_n\| < \infty$ jest zbieżny.

Lemat 1.14. *Niech $\{x_n\}_{n=1}^{\infty} \subseteq X$ będzie ciągiem Cauchy.*

- (1) *Jeżeli pewien podciąg $\{x_{n_k}\}_{k=1}^{\infty}$ jest zbieżny do x , to ciąg $\{x_n\}_{n=1}^{\infty}$ jest zbieżny do x .*
- (2) *Dla każdego $q \in (0, 1)$ istnieje podciąg $\{x_{n_k}\}_{k=1}^{\infty}$ taki, że $\|x_{n_k} - x_{n_l}\| \leq q^k$ dla każdego $l \geq k$.*

Dowód. (1). Załóżmy, że $\{x_{n_k}\}_{k=1}^{\infty}$ jest zbieżny do x . Korzystając z ciągłości normy i warunku Cauchy mamy $\lim_{n \rightarrow \infty} \|x_n - x\| = \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \|x_n - x_{n_k}\| = 0$.

(2). Konstruujemy żądany podciąg indukcyjnie korzystając z warunku Cauchy. Najpierw wybierzmy $n_1 \in \mathbb{N}$ takie, że $\|x_n - x_m\| \leq q$ dla $n, m \geq n_1$. Następnie bierzemy $n_2 \geq n_1$ takie, że $\|x_n - x_m\| \leq q^2$ dla $n, m \geq n_2$. W k -tym kroku indukcyjnym wybieramy $n_k \geq n_{k-1}$ takie, że $\|x_n - x_m\| \leq q^k$ dla $n, m \geq n_k$. $\square$

Stwierdzenie 1.15. *Przestrzeń unormowana X jest zupełna wtedy i tylko wtedy, gdy każdy szereg absolutnie zbieżny jest zbieżny w X , tzn. gdy $\sum_{n=1}^{\infty} \|x_n\| < \infty$ pociąga, że szereg $\sum_{n=1}^{\infty} x_n$ jest zbieżny, dla każdego ciągu $\{x_n\}_{n=1}^{\infty} \subseteq X$.*

Dowód. Załóżmy, że $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Wtedy dla $M \geq N \geq 1$ mamy

$$\left\| \sum_{n=1}^M x_n - \sum_{n=1}^N x_n \right\| = \left\| \sum_{n=N+1}^M x_n \right\| \leq \sum_{n=N+1}^M \|x_n\| \longrightarrow 0, \quad \text{przy } N, M \rightarrow \infty,$$

bo ogon szeregu zbieżnego jest zbieżny do zera. Zatem ciąg sum częściowych $\{\sum_{n=1}^N x_n\}_{N=1}^\infty$ jest ciągiem Cauchy, więc musi być zbieżny jeżeli przestrzeń X jest zupełna.

Założmy teraz że każdy szereg w X , który jest absolutnie zbieżny jest zbieżny. Niech $\{x_n\}_{n=1}^\infty \subseteq X$ będzie dowolnym ciągiem Cauchy. Na mocy Lematu 1.14(2) istnieje podciąg $\{x_{n_k}\}_{k=1}^\infty$ taki, że $\|x_{n_k} - x_{n_l}\| \leq (1/2)^k$ dla każdego $l \geq k$. Rozważmy ciąg $\{y_k\}_{k=1}^\infty \subseteq X$, gdzie $y_1 := x_{n_1}$ oraz $y_{k+1} := x_{n_{k+1}} - x_{n_k}$ dla $k \geq 1$. Odpowiadający mu szereg jest bezwzględnie zbieżny:

$$\sum_{k=1}^{\infty} \|y_k\| = \|x_{n_1}\| + \sum_{k=1}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| \leq \|x_{n_1}\| + \sum_{k=1}^{\infty} (1/2)^k = \|x_{n_1}\| + 1 < \infty.$$

Zatem na mocy założenia szereg $\sum_{k=1}^{\infty} y_k$ jest zbieżny. Ale $\sum_{k=1}^N y_k = x_{n_1} + \sum_{k=1}^N x_{n_{k+1}} - x_{n_k} = x_{n_{N+1}}$, skąd wynika, że podciąg $\{x_{n_k}\}_{k=1}^\infty$ jest zbieżny. Zatem ciąg $\{x_n\}_{n=1}^\infty$ jest zbieżny, patrz Lemat 1.14(1). $\square$

1.2 Klasyczne przestrzenie Banacha.

1.2.1 Przestrzenie funkcji ciągłych

Niech M będzie ustaloną przestrzenią topologiczną. Jako że kombinacja liniowa funkcji ciągłych jest funkcją ciągłą, to *przestrzeń funkcji ciągłych* $C(M) := \{x : M \rightarrow \mathbb{F} \text{ funkcja ciągła}\}$ o wartościach w ciele $\mathbb{F} = \mathbb{R}, \mathbb{C}$ wraz z działaniami określonymi punktowo:

$$(x + y)(t) := x(t) + y(t), \quad (\lambda x)(t) := \lambda x(t),$$

dla $x, y \in C(M)$, $\lambda \in \mathbb{F}$, jest przestrzenią liniową. *Przestrzeń funkcji ciągłych i ograniczonych*

$$C_b(M) := \{x \in C(M) : \sup_{t \in M} |x(t)| < \infty\}$$

jest podprzestrzenią liniową przestrzeni $C(M)$, bo kombinacja liniowa funkcji ograniczonych jest funkcją ograniczoną. Na przestrzeni $C_b(M)$ w naturalny sposób określona jest *norma supremum*

$$\|x\|_\infty := \sup_{t \in M} |x(t)|$$

(🏠 sprawdzić, że $\|\cdot\|_\infty$ jest normą na $C_b(M)$).

Lemat 1.16. *Zbieżność w normie $\|\cdot\|_\infty$ jest równoważna zbieżności jednostajnej.*

Dowód. $x_n \xrightarrow{\|\cdot\|_\infty} x \iff \lim_{n \rightarrow \infty} \sup_{t \in M} |x_n(t) - x(t)| = 0 \iff \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N \sup_{t \in M} |x_n(t) - x(t)| < \varepsilon \iff \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N \forall t \in M |x_n(t) - x(t)| < \varepsilon \stackrel{\text{def}}{\iff} x_n \rightrightarrows x. \quad \square$

Zbieżność jednostajną, tak jak w powyższym dowodzie, będziemy oznaczać symbolem $\rightrightarrows$, który podkreśla, że granica jednostajna ciągu funkcji jest znacznie silniejsza niż zbieżność punktowa. W szczególności,

jednostajnie zbieżny ciąg funkcji ciągłych musi być zbieżny do funkcji ciągłej

(znany fakt z analizy matematycznej/topologii 🏠). Zbieżność punktowa takiej własności nie ma. Np. ciąg $x_n(t) = t^n$ jest zbieżny punktowo na $[0, 1]$ do funkcji nieciągłej $x(t) = \begin{cases} 0, & t < 1 \\ 1, & t = 1 \end{cases}$.

W szczególności, ciąg $\{x_n\}_{n=1}^\infty$ nie jest zbieżny w normie $\|\cdot\|_\infty$.

Stwierdzenie 1.17. Przestrzeń $C_b(M)$ wraz z normą $\|x\|_\infty = \sup_{t \in M} |x(t)|$ jest przestrzenią Banacha.

Dowód. Niech $\{x_n\}_{n=1}^\infty \subseteq C_b(M)$ ciąg Cauchy. Wtedy dla każdego $t \in M$ mamy

$$|x_n(t) - x_m(t)| \leq \sup_{s \in M} |x_n(s) - x_m(s)| = \|x_n - x_m\|_\infty \rightarrow 0, \quad \text{przy } n, m \rightarrow \infty.$$

Czyli ciąg liczbowy $\{x_n(t)\}_{n=1}^\infty$ jest ciągiem Cauchy w ciele $\mathbb{F}$, a zatem jest zbieżny (bo ciało $\mathbb{F}$ jest zupełne). Czyli istnieje $x(t) \in \mathbb{F}$ takie, że $x_n(t) \rightarrow x(t)$ w $\mathbb{F}$. W ten sposób otrzymujemy funkcję liczbową $M \ni t \mapsto x(t) \in \mathbb{F}$, która jest „kandydatem na granicę” ciągu $\{x_n\}_{n=1}^\infty$. Żeby pokazać, że $\{x_n\}_{n=1}^\infty$ zbiega do x w normie, zauważmy, że dla każdego $t \in M$ mamy

$$|x_n(t) - x(t)| = \lim_{m \rightarrow \infty} |x_n(t) - x_m(t)| \leq \limsup_{m \rightarrow \infty} \|x_n - x_m\|_\infty.$$

Stąd oraz stąd, że ciąg $\{x_n\}_{n=1}^\infty$ jest Cauchy dostajemy

$$\lim_{n \rightarrow \infty} \|x_n - x\|_\infty = \lim_{n \rightarrow \infty} \sup_{t \in M} |x_n(t) - x(t)| \leq \lim_{n \rightarrow \infty} \limsup_{m \rightarrow \infty} \|x_n - x_m\|_\infty = 0.$$

Zatem $x_n \xrightarrow{\|\cdot\|_\infty} x$. Granica jednostajna ciągu funkcji ciągłych jest ciągła. Stąd $x \in C(M)$. Co więcej, funkcja x jest ograniczona, bo $\|x\|_\infty \leq \|x - x_n\|_\infty + \|x_n\|_\infty < \infty$. Czyli $x \in C_b(M)$ i ciąg $\{x_n\}_{n=1}^\infty$ jest zbieżny w przestrzeni $C_b(M)$. $\square$

Uwaga 1.18. Rzeczywista funkcja ciągła na zbiorze zwartym osiąga swoje kresy i w szczególności jest ograniczona. Zatem jeśli M jest przestrzenią zwartą, to dla każdego $x \in C(M)$ mamy $\|x\|_\infty = \max_{t \in M} |x(t)| < \infty$ (funkcja $|x(t)|$ osiąga swoje maksimum) oraz $C(M) = C_b(M)$.

Wniosek 1.19. Jeśli M jest przestrzenią zwartą, to przestrzeń $C(M)$ jest przestrzenią Banacha z normą $\|x\|_\infty = \max_{t \in M} |x(t)|$.

Dowód. Stwierdzenie 1.17 + Uwaga 1.18. $\square$

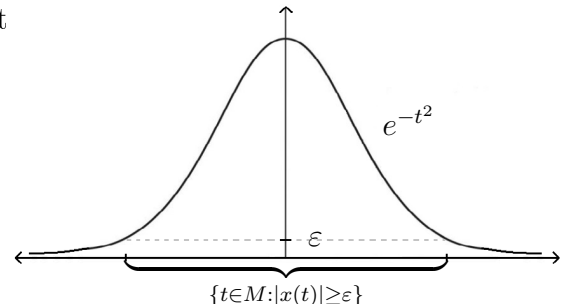
Definicja 1.20. Powiemy, że funkcja $x \in C(M)$ *znika w nieskończoności* jeśli dla każdego $\varepsilon > 0$ jej wartość bezwzględna poza zbiorem zwartym jest mniejsza niż ε . W szczególności, zbiór

$$C_0(M) := \{x \in C(M) : \forall_{\varepsilon > 0} \{t \in M : |x(t)| \geq \varepsilon\} \text{ jest zbiorem zwartym}\}$$

nazywamy *przestrzenią funkcji ciągłych znikających w nieskończoności*.

Uwaga 1.21. Jeśli M jest zwarta, to każda funkcja ciągła znika w nieskończoności, bo każdy domknięty podzbiór przestrzeni zwartej jest zbiorem zwartym. Zatem $C_0(M) = C(M) = C_b(M)$.

Przykład 1.22. Niech $M = \mathbb{R}$. Funkcja $x(t) = e^{-t^2}$ jest ściśle dodatnia, więc „nie znika w żadnym punkcie”. Ale znika w nieskończoności, bo dla $0 < \varepsilon < 1$ zbiór $\{t \in M : |x(t)| \geq \varepsilon\} = [-\sqrt{\ln \varepsilon^{-1}}, \sqrt{\ln \varepsilon^{-1}}]$ jest zwarty. Natomiast na przykład funkcja $x(t) \equiv 1$ nie znika w nieskończoności, bo dla $0 < \varepsilon < 1$ mamy wtedy $\{t \in M : |x(t)| \geq \varepsilon\} = \mathbb{R}$.



Stwierdzenie 1.23. $C_0(M)$ jest domkniętą podprzestrzenią liniową przestrzeni Banacha $(C_b(M), \|\cdot\|_\infty)$. Zatem $(C_0(M), \|\cdot\|_\infty)$ jest przestrzenią Banacha.

Dowód. Niech $x, y \in C_0(M)$ i $\varepsilon > 0$. Zauważmy, że

$$\underbrace{\{t : |x(t) + y(t)| \geq \varepsilon\}}_{\text{domknięty}} \subseteq \underbrace{\{t : |x(t)| \geq \varepsilon/2\} \cup \{t : |y(t)| \geq \varepsilon/2\}}_{\text{zbiór zwarty, jako suma dwóch zwartych}}.$$

Zatem zbiór $\{t : |x(t) + y(t)| \geq \varepsilon\}$ jest zwarty, jako domknięty podzbiór zbioru zwartego. Czyli $x + y \in C_0(M)$. Dla $\lambda \in \mathbb{F}$ zbiór $\{t : |\lambda x(t)| \geq \varepsilon\}$ jest zwarty, bo albo jest pusty (jeśli $\lambda = 0$) albo $\{t : |\lambda x(t)| \geq \varepsilon\} = \{t : |x(t)| \geq \frac{\varepsilon}{|\lambda|}\}$ jest zwarty, bo x znika w nieskończoności. Zatem $\lambda x \in C_0(M)$. Ponadto, skoro zbiór $\{t : |x(t)| \geq \varepsilon\}$ jest zwarty dla każdego $\varepsilon > 0$, to $\|x\|_\infty = \max_{t \in M} |x(t)| < \infty$, bo funkcja ciągła na zbiorze zwartym osiąga swoje kresy. Zatem $C_0(M) \subseteq C_b(M)$ jest podprzestrzenią liniową.

Aby wykazać, że podprzestrzeń $C_0(M)$ jest domknięta w $C_b(M)$ rozważmy dowolny ciąg $\{x_n\}_{n=1}^\infty \subseteq C_0(M)$ zbieżny do $x \in C_b(M)$. Niech $\varepsilon > 0$. Dla dostatecznie dużego $n \in \mathbb{N}$ mamy $\|x_n - x\|_\infty < \frac{\varepsilon}{2}$ i wtedy

$$\{t : |x(t)| \geq \varepsilon\} \subseteq \{t : |x_n(t)| \geq \varepsilon/2\}.$$

Skąd $\{t : |x(t)| \geq \varepsilon\}$ jest zwarty, jako domknięty podzbiór zbioru zwartego. Zatem $x \in C_0(M)$. $\square$

Przykład 1.24. Kładąc $M = \mathbb{N}$ funkcje możemy utożsamić z ciągami. Traktując $\mathbb{N}$ jako przestrzeń dyskretną wszystkie funkcje na $\mathbb{N}$ są ciągłe. Zatem w tym przypadku przestrzeń $C_b(M)$ można utożsamić z *przestrzenią ciągów ograniczonych*. Czyli z przestrzenią Banacha

$$\ell^\infty := \{x = (x(1), x(2), \dots) : \sup_{k \in \mathbb{N}} |x(k)| < \infty\}$$

wyposażoną w normę $\|x\|_\infty := \sup_{k \in \mathbb{N}} |x(k)|$. Natomiast przestrzeń $C_0(M)$ można utożsamić z *przestrzenią ciągów zbieżnych do zera*:

$$c_0 := \{x = (x(1), x(2), \dots) : \lim_{k \rightarrow \infty} x(k) = 0\}.$$

Rzeczywiście, podzbiór przestrzeni dyskretnej jest zwarty wtedy i tylko wtedy, gdy jest skończony. Zatem znikanie w nieskończoności funkcji $x : \mathbb{N} \rightarrow \mathbb{F}$ jest równoważne temu, że ciąg $\{x(k)\}_{k=1}^\infty$ zbiega do zera:

$$x \in C_0(\mathbb{N}) \iff \forall_{\varepsilon > 0} \{k \in \mathbb{N} : |x(k)| \geq \varepsilon\} \text{ skończony} \iff \forall_{\varepsilon > 0} \exists_{N \in \mathbb{N}} \forall_{k > N} |x(k)| < \varepsilon \iff x \in c_0.$$

W kontekście przestrzeni ciągów, często rozważa się również *przestrzeń ciągów zbieżnych*:

$$c := \{x = (x(1), x(2), \dots) : \exists_{x(\infty) \in \mathbb{F}} \lim_{k \rightarrow \infty} x(k) = x(\infty)\}.$$

Jako że ciągi zbieżne są ograniczone oraz granica zachowuje kombinacje liniowe, c jest podprzestrzenią liniową ℓ^∞ . Ponadto c jest domknięta w ℓ^∞ (aby to wykazać, trzeba pokazać, że „ciąg zbieżny ciągów zbieżnych jest zbieżny do ciągu zbieżnego” $\hat{\text{H}}$). Reasumując mamy wstępujący ciąg przestrzeni Banacha

$$c_0 \subseteq c \subseteq \ell^\infty$$

wyposażonych w normę $\|x\|_\infty := \sup_{k \in \mathbb{N}} |x(k)|$.

Na koniec pokażemy, że przestrzeń $C_0(M)$ jest domknięciem przestrzeni funkcji ciągłych o zwartym nośniku. Czyli, że funkcje znikające w nieskończoności to są dokładnie te funkcje, które można aproksymować jednostajnie funkcjami o zwartych nośnikach.

Definicja 1.25. *Nośnikiem funkcji $x : M \rightarrow \mathbb{F}$ nazywamy domknięcie zbioru punktów, w których x się nie zeruje, czyli zbiór $\text{supp}(x) := \overline{\{t \in M : x(t) \neq 0\}}$. Natomiast zbiór*

$$C_c(M) := \{x \in C(M) : \text{supp}(x) \text{ jest zbiorem zwartym}\}$$

nazywamy przestrzenią funkcji ciągłych o zwartych nośnikach.

Lemat 1.26. *Przestrzeń $C_c(M)$ jest podprzestrzenią liniową przestrzeni Banacha $C_0(M)$.*

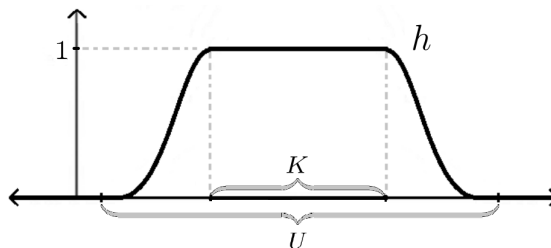
Dowód. Zauważmy, że $\text{supp}(\lambda x) = \text{supp}(x)$ dla $\lambda \neq 0$ oraz

$$\text{supp}(x + y) \subseteq \overline{\{t \in M : x(t) \neq 0\} \cup \{t \in M : y(t) \neq 0\}} = \text{supp}(x) \cup \text{supp}(y),$$

bo domknięcie sumy zbiorów jest sumą ich domknięć. Stąd wynika, że $C_c(M)$ jest przestrzenią liniową. Ponadto $\{t \in M : |x(t)| \geq \varepsilon\} \subseteq \text{supp}(x)$, dla $\varepsilon > 0$. Zatem $C_c(M) \subseteq C_0(M)$. $\square$

W kontekście przestrzeni $C_0(M)$ zazwyczaj zakłada się, że M jest przestrzenią *lokalnie zwartą*, to znaczy, że każdy punkt M posiada otwarte otoczenie, którego domknięcie jest zbiorem zwartym, oraz że M jest *przestrzenią Hausdorffa*, czyli że każde dwa różne punkty M posiadają rozłączne otoczenia otwarte. Wynika to stąd, że jeżeli punkt t nie posiada otoczenia, które jest zwarte, to każda funkcja z $C_0(M)$ musi się na nim zerować. Natomiast, jeżeli dwa punkty t_1, t_2 nie dadzą się rozdzielić zbiorami otwartymi, to $f(t_1) = f(t_2)$ dla każdej funkcji ciągłej f , czyli z punktu widzenia funkcji ciągłych można punkty t_1 i t_2 utożsamić. Zauważmy, że każdy domknięty lub otwarty podzbiór $\mathbb{R}^n$ (lub ogólniej każda różnorodność topologiczna) jest lokalnie zwartą przestrzenią Hausdorffa. Ponadto klasa takich przestrzeni jest zamknięta np. na (nieskończone) iloczyny kartezjańskie i sumy proste, więc jest bardzo obszerna. Na takich przestrzeniach zachodzi Lemat Uryshona, który ma zniezluczone zastosowania:

Twierdzenie 1.27 (Lemat Uryshona, patrz np. [4]). *Załóżmy, że M jest lokalnie zwartą przestrzenią Hausdorffa. Dla dowolnego zbioru zwartego K i otwartego U takich, że $K \subseteq U \subseteq M$, istnieje funkcja ciągła $h : M \rightarrow [0, 1]$ o zwartym nośniku zawartym w U i przyjmująca wartość 1 na zbiorze K , patrz rysunek.*



Wniosek 1.28. *Niech M będzie lokalnie zwartą przestrzenią Hausdorffa. Przestrzeń funkcji ciągłych o zwartych nośnikach $C_c(M)$ jest gęstą podprzestrzenią liniową przestrzeni Banacha $C_0(M)$. Czyli $C_0(M)$ jest uzupełnieniem $C_c(M)$ w normie supremum.*

Dowód. Niech $x \in C_0(M)$. Dla każdego n połóżmy $K_n := \{t : |x(t)| \geq 1/n\}$ oraz $U_n := \{t : |x_n(t)| > 1/(n+1)\}$. Wtedy K_n jest zwartym podzbiorem zbioru otwartego U_n . Zatem na mocy Lematu Uryshona istnieje $h_n \in C_c(M)$ taka, że $0 \leq h_n \leq 1$, $h_n|_{K_n} = 1$ oraz $\text{supp}(h_n) \subseteq U_n$. Kładąc $x_n := x \cdot h_n$ mamy $x_n \in C_c(M)$ oraz

$$\|x_n - x\|_\infty = \sup_{t \in M} |x(t)h_n(t) - x(t)| = \sup_{t \in M \setminus K_n} |x(t)h_n(t) - x(t)| \leq 2/n.$$

Zatem ciąg $\{x_n\}_{n=1}^\infty \subseteq C_c(M)$ zbiega do $x \in C_0(M)$. Stąd $\overline{C_c(M)} = C_0(M)$. $\square$

Przykład 1.29 (Przestrzeń ciągów skończonych). Dla przestrzeni dyskretnej $M = \mathbb{N}$ przestrzeń funkcji $C_c(M)$ możemy utożsamić z *przestrzenią ciągów skończonych*:

$$c_{00} = \{x = (x(1), x(2), \dots, x(N), 0, 0, \dots) : N \in \mathbb{N}, x(k) \in \mathbb{F}\}.$$

Formalnie jest to przestrzeń ciągów ze skończoną ilością wyrazów niezerowych. W szczególności c_{00} jest gęstą podprzestrzenią przestrzeni c_0 ciągów zbieżnych do zera.

1.2.2 Przestrzenie L^p dla $p \geq 1$

Niech (Ω, Σ, μ) będzie ustaloną przestrzenią z miarą. Czyli Σ jest σ -algebrą podzbiorów zbioru Ω , a $\mu : \Sigma \rightarrow [0, +\infty]$ jest miarą (czyli σ -addytywną funkcją zbioru taką, że $\mu(\emptyset) = 0$). Niech $p \geq 1$. Powiemy, że funkcja $x : \Omega \rightarrow \mathbb{F} = \mathbb{R}, \mathbb{C}$ jest *całkowalna w p -tej potęgde*, gdy funkcja $|x|^p$ jest całkowalna, czyli istnieje liczba

$$\|x\|_p := \left(\int_{\Omega} |x|^p d\mu \right)^{\frac{1}{p}}$$

Zbiór funkcji całkowalnych w p -tej potęgde oznaczamy przez $\mathcal{L}^p(\mu)$.

Twierdzenie 1.30 (Nierówność Höldera). *Dla dowolnych $1 < p, q < \infty$ takich, że $\frac{1}{p} + \frac{1}{q} = 1$ oraz dowolnych funkcji $x \in \mathcal{L}^p(\mu)$ i $y \in \mathcal{L}^q(\mu)$ zachodzi*

$$\|x \cdot y\|_1 \leq \|x\|_p \cdot \|y\|_q, \quad (1.1)$$

czyli

$$\int_{\Omega} |xy| d\mu \leq \left(\int_{\Omega} |x|^p d\mu \right)^{\frac{1}{p}} \left(\int_{\Omega} |y|^q d\mu \right)^{\frac{1}{q}}.$$

W szczególności $x \cdot y \in \mathcal{L}^1(\mu)$. Ponadto następujące warunki są równoważne

- (1) w nierówności (1.1) zachodzi równość;
- (2) $|x|^p$ i $|y|^q$ są liniowo zależne poza zbiorem miary zero, tzn. istnieją liczby α, β nie będące jednocześnie zerami takie, że $\alpha|x|^p = \beta|y|^q$ μ -prawie wszędzie;
- (3) $(\int_{\Omega} |y|^q d\mu) \cdot |x|^p = (\int_{\Omega} |x|^p d\mu) \cdot |y|^q$ μ -prawie wszędzie.

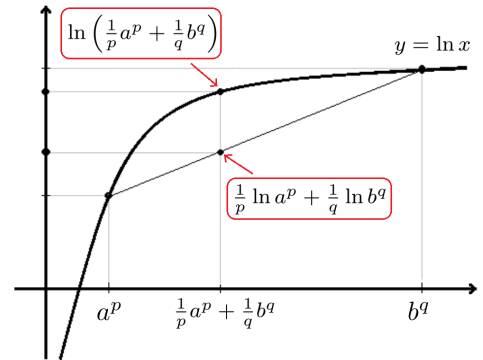
Dowód. Jeśli $\|x\|_p = 0$, to $x = 0$ μ -prawie wszędzie, skąd $x \cdot y = 0$ μ -prawie wszędzie i wtedy warunki (1)-(3) zachodzą trywialnie. Podobnie, gdy $\|y\|_q = 0$. Załóżmy zatem, że $\|x\|_p \neq 0$ i $\|y\|_q \neq 0$. Zastosujemy *nierówność Younga*, która mówi, że dla dowolnych liczb $a, b > 0$ mamy

$$a \cdot b \leq \frac{1}{p} a^p + \frac{1}{q} b^q,$$

a która wynika np. z własności logarytmu. Mianowicie,

$$\ln(ab) = \ln a + \ln b = \frac{1}{p} \ln a^p + \frac{1}{q} \ln b^q \leq \ln \left(\frac{1}{p} a^p + \frac{1}{q} b^q \right),$$

gdzie nierówność wynika z wklęsłości logarytmu, patrz rysunek po prawej. Zdejmując logarytm otrzymujemy nierówność Younga. Równość zachodzi $\iff a^p = b^q$.



Formalnie, funkcja f jest *wklęsła*, jeśli dla $\lambda \in [0, 1]$ oraz t_1, t_2 z dziedziny f zachodzi $f(\lambda t_1 + (1 - \lambda)t_2) \leq \lambda f(t_1) + (1 - \lambda)f(t_2)$. W dowodzie nierówności Younga kładziemy $f(t) = \ln(t)$, $\lambda = \frac{1}{p}$, $1 - \lambda = \frac{1}{q}$, $t_1 = a^p$ oraz $t_2 = b^q$.

Stosując nierówność Younga do $a = \frac{|x(t)|}{\|x\|_p}$ i $b = \frac{|y(t)|}{\|y\|_q}$ otrzymujemy

$$\frac{|x(t)y(t)|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} \cdot \frac{|x(t)|^p}{\|x\|_p^p} + \frac{1}{q} \cdot \frac{|y(t)|^q}{\|y\|_q^q}, \quad \text{dla każdego } t \in \Omega.$$

Zauważmy, że $\|x\|_p^p = \int_{\Omega} |x(t)|^p d\mu$. Zatem całkując powyższą nierówność obustronnie mamy

$$\frac{\int_{\Omega} |x(t)y(t)| d\mu}{\|x\|_p \|y\|_q} \leq \frac{1}{p} \cdot \frac{\int_{\Omega} |x(t)|^p d\mu}{\|x\|_p^p} + \frac{1}{q} \cdot \frac{\int_{\Omega} |y(t)|^q d\mu}{\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Dzieląc powyższą nierówność obustronnie przez $\|x\|_p \|y\|_q$ dostajemy $\|x \cdot y\|_1 \leq \|x\|_p \cdot \|y\|_q$. Przy czym, równość $\|x \cdot y\|_1 = \|x\|_p \cdot \|y\|_q$ zachodzi wtedy i tylko wtedy, gdy $\frac{|x(t)|^p}{\|x\|_p^p} = \frac{|y(t)|^q}{\|y\|_q^q}$ dla μ -prawie każdego $t \in \Omega$, czyli gdy $\|y\|_q^q \cdot |x|^p = \|x\|_p^p \cdot |y|^q$ μ -prawie wszędzie. To dowodzi równoważności (1) $\Leftrightarrow$ (3). Warunek (3) implikuje (2) kładąc $\alpha = \|y\|_q^q$ i $\beta = \|x\|_p^p$. Na odwrót, jeśli $\alpha|x|^p = \beta|y|^q$ μ -prawie wszędzie, dla pewnych $\alpha, \beta \in \mathbb{F}$ to całkując mamy $\alpha\|x\|_p^p = \beta\|y\|_q^q$ lub równoważnie $\alpha = \frac{\|y\|_q^q}{\|x\|_p^p} \beta$. Podstawiając do $\alpha|x|^p = \beta|y|^q$ otrzymujemy, że $\frac{|x|^p}{\|x\|_p^p} = \frac{|y|^q}{\|y\|_q^q}$ μ -prawie wszędzie. Zatem (2) implikuje (3). $\square$

Uwaga 1.31. Liczby $1 < p, q < \infty$ spełniające $\frac{1}{p} + \frac{1}{q} = 1$ nazywamy *sprzężonymi hölderowsko*. Oczywiście liczby te wyznaczają się nawzajem: mamy $q = p/(p-1)$ oraz $p = q/(q-1)$.

Wniosek 1.32 (Nierówność Minkowskiego). *Dla dowolnego $p \geq 1$ oraz $x, y \in \mathcal{L}^p(\mu)$ mamy*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

W szczególności $x + y \in \mathcal{L}^p(\mu)$. Ponadto, równość $\|x + y\|_p = \|x\|_p + \|y\|_p$ zachodzi wtedy i tylko wtedy, gdy

- (1) *jeśli $p = 1$, to “ x i y mają ten sam kierunek”, czyli $x \cdot \bar{y} \geq 0$ μ -prawie wszędzie;*
- (2) *jeśli $p > 1$, to “ x i y są dodatnio proporcjonalne”, tj. istnieją $\alpha, \beta \geq 0$, nie będące jednocześnie zerami, takie że $\alpha x = \beta y$ μ -prawie wszędzie.*

Dowód. Dla $p = 1$ dowód jest łatwy. Rzeczywiście, korzystając z nierówności trójkąta i monotoniczności całki mamy

$$\|x + y\|_1 = \int_{\Omega} |x + y| d\mu \leq \int_{\Omega} |x| + |y| d\mu = \int_{\Omega} |x| d\mu + \int_{\Omega} |y| d\mu = \|x\|_1 + \|y\|_1.$$

Ponadto, $|x + y| = |x|e^{i\arg x} + |y|e^{i\arg y}$ równa się $|x| + |y|$ wtedy i tylko wtedy, gdy $\arg x = \arg y$, co zachodzi wtedy i tylko wtedy, gdy $x \cdot \bar{y} \geq 0$. Zatem $\|x + y\|_1 = \|x\|_1 + \|y\|_1$ wtedy i tylko wtedy, gdy $x \cdot \bar{y} \geq 0$ μ -prawie wszędzie.

Założmy zatem, że $p > 1$. Niech $q := p/(p-1)$. Wtedy $1/p + 1/q = 1$ oraz $q(p-1) = p$. Stąd

$$\begin{aligned} \|x + y\|_p^p &= \int_{\Omega} |x + y| \cdot |x + y|^{p-1} d\mu \stackrel{\text{N. Trójkąta}}{\leq} \int_{\Omega} |x| \cdot |x + y|^{p-1} d\mu + \int_{\Omega} |y| \cdot |x + y|^{p-1} d\mu \\ &\stackrel{\text{N. Höldera } \times 2}{\leq} \|x\|_p \cdot \left(\int_{\Omega} |x + y|^{q(p-1)} d\mu \right)^{\frac{1}{q}} + \|y\|_p \cdot \left(\int_{\Omega} |x + y|^{q(p-1)} d\mu \right)^{\frac{1}{q}} \\ &= \|x\|_p \cdot \|x + y\|_p^{p/q} + \|y\|_p \cdot \|x + y\|_p^{p/q} = (\|x\|_p + \|y\|_p) \cdot \|x + y\|_p^{p/q}. \end{aligned}$$

Dzieląc obie strony przez $\|x + y\|_p^{p/q}$ i korzystając stąd, że $p - p/q = 1$, otrzymujemy żadaną nierówność. Równość $\|x + y\|_p = \|x\|_p + \|y\|_p$ zachodzi wtedy i tylko wtedy, gdy $|x + y| = |x| + |y|$ zachodzi μ -prawie wszędzie oraz gdy zachodzą równości w nierównościach Höldera, czyli pary $|x|^p$ i $|x + y|^p$ oraz $|y|^p$ i $|x + y|^p$ są liniowo zależne, poza zbiorem miary zero. Te warunki są równoważne założeniu, że $x\bar{y} \geq 0$ μ -prawie wszędzie oraz $|x|$ i $|y|$ są liniowo zależne poza zbiorem miary zero. To z kolei jest równoważne dodatniej proporcjonalności opisanej w (2). $\square$

Uwaga 1.33. Warunek (2) w powyższym wniosku jest znacznie silniejszy niż warunek (1). Warunek (1) oznacza, że istnieją funkcje nieujemne α, β takie, że $\alpha x = \beta y$ (można tu wziąć na przykład $\alpha = |y|$ i $\beta = |x|$). Natomiast warunek (2) mówi, że zamiast funkcji można wziąć stałe (w tym warunku można zawsze wziąć $\alpha = \int_{\Omega} |y| d\mu$ i $\beta = \int_{\Omega} |x| d\mu$).

Z nierówności Minkowskiego wynika, że zbiór $\mathcal{L}^p(\mu)$ wraz z działaniami określonymi punktowo jest przestrzenią liniową, a funkcja $\|\cdot\|_p : \mathcal{L}^p(\mu) \rightarrow [0, \infty)$ określa na nim *półnormę*, tj. funkcję spełniającą warunki (N2), (N3) w Definicji 1.1. Co do warunku (N1), to zachodzi on jedynie “ μ -prawie wszędzie”:

$$\|x\|_p = 0 \iff \int_{\Omega} |x|^p d\mu = 0 \iff |x|^p = 0 \text{ } \mu\text{-p.w.} \iff x = 0 \text{ } \mu\text{-p.w.}$$

Zatem aby formalnie móc mówić o przestrzeni unormowanej, w dalszym ciągu będziemy utożsamiać funkcje równe sobie prawie wszędzie. Dokładniej, równość μ -prawie wszędzie:

$$x \stackrel{\mu\text{-pw}}{=} y \stackrel{\text{def}}{\iff} \mu(\{t \in \Omega : x(t) \neq y(t)\}) = 0.$$

jest relacją równoważności. Klasą abstrakcji $[0]$ funkcji zerowej jest przestrzeń liniowa $\mathcal{L}_0(\mu) = \{x : \Omega \rightarrow \mathbb{F} : x \stackrel{\mu\text{-pw}}{=} 0\}$ funkcji mierzalnych równych zero μ -prawie wszędzie. Dla dowolnego $x \in \mathcal{L}^p(\mu)$ klasa abstrakcji $[x]$ pokrywa się z warstwą $x + \mathcal{L}_0(\mu)$ w przestrzeni ilorazowej $\mathcal{L}^p(\mu)/\mathcal{L}_0(\mu)$.

Definicja 1.34. *Przestrzenią funkcji całkowalnych w p -tej potęgzie nazywamy przestrzeń unormowaną, która jako przestrzeń liniowa jest przestrzenią ilorazową*

$$L^p(\mu) := \mathcal{L}^p(\mu)/\stackrel{\mu\text{-pw}}{=} = \mathcal{L}^p(\mu)/\mathcal{L}_0(\mu),$$

a norma na $L^p(\mu)$ jest indukowana przez półnormę $\|\cdot\|_p$ na $\mathcal{L}^p(\mu)$. Czyli formalnie $[x] + [y] := [x + y]$, $\lambda[x] := [\lambda x]$ oraz $\|[x]\|_p := \|x\|_p$ dla $x, y \in \mathcal{L}^p(\mu)$, $\lambda \in \mathbb{F}$.

Konwencja 1.35. Mimo, iż formalnie elementami przestrzeni $L^p(\mu)$ są klasy abstrakcji $[x]$, a nie funkcje x , to dla uproszczenia notacji i języka, w dalszym ciągu

elementy $L^p(\mu)$ będziemy nazywać funkcjami i zapisywać jak funkcje!

Jest to standardowa konwencja, która nie prowadzi do nieporozumień, tak długo jak pamiętamy, że w $L^p(\mu)$ utożsamiamy funkcje równe sobie μ -prawie wszędzie.

Twierdzenie 1.36. *Dla każdego $p \geq 1$ przestrzeń $L^p(\mu)$ jest przestrzenią Banacha.*

Dowód. Niech $\{x_n\}_{n=1}^{\infty} \subseteq L^p(\mu)$ będzie ciągiem Cauchy. Jeżeli ciąg Cauchy posiada podciąg zbieżny, to jest zbieżny. Zatem przechodząc ewentualnie do podciągu możemy założyć, że $\|x_n - x_m\|_p \leq (1/4)^n$ dla $m \geq n$, patrz Lemat 1.14. Połóżmy

$$A_n := \{t \in \Omega : |x_n(t) - x_{n+1}(t)| \geq (1/2)^n\} \quad \text{oraz} \quad A := \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n.$$

Wtedy

$$(1/2)^{np} \mu(A_n) \leq \int_{A_n} \underbrace{|x_n - x_{n+1}|^p}_{\geq (1/2)^{np}} d\mu \leq \|x_n - x_{n+1}\|_p^p \leq (1/4)^{np},$$

skąd $\mu(A_n) \leq (1/2)^{np}$. Zatem

$$0 \leq \mu(A) \leq \mu\left(\bigcup_{n=N}^{\infty} A_n\right) \leq \sum_{N=n}^{\infty} \mu(A_n) \leq \underbrace{\sum_{n=N}^{\infty} (1/2)^{np}}_{\text{ogon szeregu zbieżnego}} \longrightarrow 0, \quad \text{przy } N \rightarrow \infty.$$

Czyli $\mu(A) = 0$. Zauważmy dalej, że

$$\begin{aligned} t \in \Omega \setminus A &\iff t \notin \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n \iff \exists_N \forall_{n \geq N} |x_n(t) - x_{n+1}(t)| < (1/2)^n \\ &\iff \exists_N \forall_{m \geq n \geq N} |x_n(t) - x_m(t)| < \sum_{k=n}^m (1/2)^k \left(\longrightarrow 0, \quad \text{przy } n, m \rightarrow \infty \right). \end{aligned}$$

Zatem dla $t \in \Omega \setminus A$ ciąg liczbowy $\{x_n(t)\}_{n=1}^{\infty}$ jest Cauchy, a więc jest zbieżny. Połóżmy $x(t) := \lim_{n \rightarrow \infty} x_n(t)$, gdy $t \in \Omega \setminus A$, oraz $x(t) = 0$, gdy $t \in A$. Korzystając z Lematu Fatou mamy

$$\begin{aligned} \|x - x_n\|_p^p &= \int_{\Omega} |x(t) - x_n(t)|^p d\mu = \int_{\Omega \setminus A} \lim_{m \rightarrow \infty} |x_m(t) - x_n(t)|^p d\mu \\ &\stackrel{\text{Fatou}}{\leq} \liminf_{m \rightarrow \infty} \int_{\Omega} |x_m - x_n|^p d\mu \leq \sup_{m \geq n} \|x_n - x_m\|_p^p \leq (1/4)^{pn} \longrightarrow 0. \end{aligned}$$

Zatem $x_n \xrightarrow{\|\cdot\|_p} x$. W szczególności $\|x\|_p \leq \|x - x_n\|_p + \|x_n\|_p < \infty$, czyli $x \in L^p(\mu)$. $\square$

Uwaga 1.37. Z powyższego dowodu wynika, jeżeli ciąg $\{x_n\}_{n=1}^{\infty} \subseteq L^p(\mu)$ jest zbieżny w $L^p(\mu)$ do x , to istnieje podciąg $\{x_{n_k}\}_{k=1}^{\infty}$ zbieżny μ -prawie wszędzie do x :

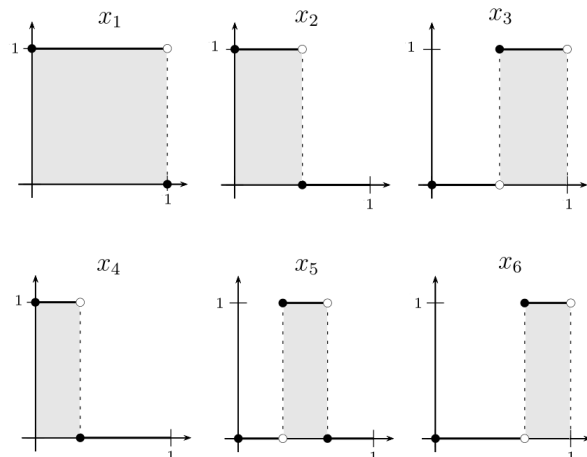
$$x_n \xrightarrow{\|\cdot\|_p} x \implies \exists_{\{x_{n_k}\}_{k=1}^{\infty}} x_{n_k} \xrightarrow{\mu\text{-pw}} x.$$

Na ogół zbieżność w p -tej normie nie pociąga zbieżności prawie wszędzie całego ciągu $\{x_n\}_{n=1}^{\infty}$; przejście do podciągu jest konieczne.

Przykład 1.38 (Wędrujący garb). Rozważmy przestrzeń $L^p(\mu)$, gdzie μ jest miarą Lebesgue'a "długością" na odcinku $[0, 1)$. Dla każdego $k \in \mathbb{N}$ utwórzmy k -elementowy ciąg funkcji charakterystycznych $x_i^{(k)} = \mathbb{1}_{[\frac{i-1}{k}, \frac{i}{k})}$, $i = 1, \dots, k$, odpowiadających podziałowi odcinka $[0, 1)$ na k równych części. Następnie ustawmy te skończone ciągi kolejno, jeden po drugim, tworząc jeden ciąg nieskończony:

$$x_1 = x_1^{(1)}, \quad \underbrace{x_2 = x_1^{(2)}, x_3 = x_2^{(2)}}_{\text{drugi ciąg}}, \quad \underbrace{x_4 = x_1^{(3)}, x_5 = x_2^{(3)}, x_6 = x_3^{(3)}}_{\text{trzeci ciąg}}, \quad \dots,$$

Wykresy funkcji z tego ciągu można wyobrazić sobie jako garb, który wędruje z lewa na prawo. Za każdym razem, gdy przekroczy "prawą stronę ekranu", garb chudnie i pojawia się z "lewej strony ekranu". To że za każdym razem garb chudnie gwarantuje, że ciąg $\{x_n\}_{n=1}^{\infty}$ zbiega do zera w $L^p(\mu)$ (mamy $\|x_i^{(k)}\|_p = (1/k)^{1/p} \rightarrow 0$ przy $k \rightarrow \infty$). Natomiast wędrowanie garbu implikuje, że dla każdego $t \in [0, 1)$ ciąg liczbowy $\{x_n(t)\}_{n=1}^{\infty}$ jest rozbieżny (posiada dwa punkty skupienia 0 i 1).



Konwencja 1.39. Piszemy $L^p[a, b] := L^p(\mu)$ jeśli μ to miara Lebesgue'a na odcinku $[a, b]$.

Przykład 1.40 (Przestrzeń ℓ^p). Jeśli $\Omega = \mathbb{N}$ i μ jest miarą liczącą (określoną na wszystkich podzbiorach Ω), to traktując funkcje na $\mathbb{N}$ jako ciągi, całka względem μ jest sumą szeregu. Stąd dla $p \geq 1$, $L^p(\mu) = \ell^p$ jest przestrzenią ciągów sumowalnych w p -tej potęgę. Innymi słowy, przestrzeń

$$\ell^p := \left\{ x = (x(1), \dots, x(n), \dots) \in \mathbb{F}^{\mathbb{N}} : \sum_{k=1}^{\infty} |x(k)|^p < \infty \right\}$$

wraz z działaniami określonymi po współrzędnych i normą $\|x\|_p := (\sum_{k=1}^{\infty} |x(k)|^p)^{\frac{1}{p}} < \infty$ jest przestrzenią Banacha.

Przykład 1.41 (Przestrzeń $\mathbb{F}^n$). Jeśli $\Omega = \{1, \dots, n\}$ jest zbiorem skończonym i μ miarą liczącą, to $L^p(\mu) \cong \mathbb{F}^n$ jest n -wymiarową przestrzenią wektorową z p -tą normą $\|x\|_p := (\sum_{k=1}^n |x(k)|^p)^{\frac{1}{p}}$. W szczególności, $\|x\|_2 := (\sum_{k=1}^n |x(k)|^2)^{\frac{1}{2}}$ jest normą euklidesową.

Zauważmy, że funkcje charakterystyczne zbiorów o mierze skończonej rozpinają liniowo przestrzeń całkowalnych funkcji prostych

$$\begin{aligned} \mathcal{E}(\mu) &:= \text{lin}\{\mathbf{1}_{A_k} : A_k \in \Sigma, \mu(A_k) < \infty\} / \mu\text{-}\underline{\text{p}}\text{w} \\ &= \left\{ \sum_{k=1}^n a_k \mathbf{1}_{A_k} : a_k \in \mathbb{F}, \mu(A_k) < \infty, k = 1, \dots, n \in \mathbb{N} \right\} / \mu\text{-}\underline{\text{p}}\text{w}. \end{aligned}$$

Ponadto $\mathcal{E}(\mu)$ zawiera się w każdej z przestrzeni $L^p(\mu)$ jako gęsta podprzestrzeń liniowa, co w świetle Twierdzenia 1.11 oznacza, że przestrzeń $L^p(\mu)$ można zdefiniować abstrakcyjnie jako uzupełnienie przestrzeni $\mathcal{E}(\mu)$ w normie $\|\cdot\|_p$:

Stwierdzenie 1.42. Dla każdego $p \in [1, +\infty)$, $\mathcal{E}(\mu) \subseteq L^p(\mu)$ jest gęstą podprzestrzenią liniową. Czyli $L^p(\mu) = \overline{\mathcal{E}(\mu)}^{\|\cdot\|_p}$ jest uzupełnieniem przestrzeni $\mathcal{E}(\mu)$ w normie $\|\cdot\|_p$.

Dowód. □

1.2.3 Przestrzeń L^∞

Powyżej rozpatrywaliśmy przestrzeń L^p dla $1 \leq p < \infty$. Do kompletu zdefiniujemy jeszcze przestrzeń L^∞ .

Definicja 1.43. Funkcja $x : \Omega \rightarrow \mathbb{F}$ jest istotnie ograniczona względem μ , jeśli jest ograniczona poza zbiorem miary zero, tzn. gdy istnieje $A \in \Sigma$ taki, że $\mu(A) = 0$ oraz $\sup_{t \in \Omega \setminus A} |x(t)| < \infty$. Supremum istotnym funkcji x nazywamy wartość

$$\|x\|_\infty := \inf_{\mu(A)=0} \sup_{t \in \Omega \setminus A} |x(t)|. \quad (1.2)$$

Zbiór mierzalnych funkcji istotnie ograniczonych oznaczamy przez $\mathcal{L}^\infty(\mu)$.

Lemat 1.44. Dla każdego $x \in \mathcal{L}^\infty(\mu)$ infimum w (1.2) jest de facto minimum, tzn. istnieje $A \in \Sigma$ taki, że $\mu(A) = 0$ oraz $\|x\|_\infty = \sup_{t \in \Omega \setminus A} |x(t)|$.

Dowód. Z definicji (1.2), dla każdego $n \in \mathbb{N}$ istnieje zbiór $A_n \in \Sigma$ taki, że $\mu(A_n) = 0$ oraz $\sup_{t \in \Omega \setminus A_n} |x(t)| \leq \|x\|_\infty + 1/n$. Wtedy dla $A = \bigcup_{n=1}^{\infty} A_n$ mamy $\mu(A) = 0$ oraz

$$\|x\|_\infty \leq \sup_{t \in \Omega \setminus A} |x(t)| = \sup_{t \in \bigcap_{n=1}^{\infty} \Omega \setminus A_n} |x(t)| \leq \|x\|_\infty + 1/n,$$

dla każdego $n \in \mathbb{N}$. Czyli $\|x\|_\infty = \sup_{t \in \Omega \setminus A} |x(t)|$. □

Zbiór $\mathcal{L}^\infty(\mu)$ mierzalnych funkcji istotnie ograniczonych jest przestrzenią liniową, na której $\|x\|_\infty$ jest półnormą (sprawdzić 🏠). W świetle Lematu 1.44, $\|x\|_\infty = 0$ wtedy i tylko wtedy, gdy $x \stackrel{\mu\text{-pw}}{=} 0$. Zatem możemy w naturalny sposób rozszerzyć Definicję 1.34 o przypadek $p = \infty$:

Definicja 1.45. *Przestrzenią funkcji istotnie ograniczonych nazywamy przestrzeń ilorazową*

$$L^\infty(\mu) := \mathcal{L}^\infty(\mu) / \mu\text{-pw} = \mathcal{L}^\infty(\mu) / \mathcal{L}_0(\mu),$$

wyposażoną w normę $\|\cdot\|_\infty$ zadaną przez supremum istotne. Konwencję 1.35 rozszerzamy na przypadek, gdy $p = \infty$.

Twierdzenie 1.46. *Przestrzeń $L^\infty(\mu)$ z normą $\|x\|_\infty$ jest przestrzenią Banacha.*

Dowód. (Warto porównać z dowodem Twierdzenia 1.36). Niech $\{x_n\}_{n=1}^\infty \subseteq L^\infty(\mu)$ ciąg Cauchy. Przechodząc ewentualnie do podciągu możemy założyć, że $\|x_n - x_m\|_\infty \leq (1/2)^n$ dla $m \geq n$. Wtedy zbiory $A_n := \{t \in \Omega : |x_n(t) - x_{n+1}(t)| > (1/2)^n\}$, $n \in \mathbb{N}$, mają miarę zero. Zatem zbiór

$$A := \bigcup_{n=1}^{\infty} A_n = \{t \in \Omega : \exists_{n \in \mathbb{N}} |x_n(t) - x_{n+1}(t)| > (1/2)^n\}$$

też ma miarę zero (bo $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n) = 0$). Zauważmy dalej, że

$$t \in \Omega \setminus A \iff \forall_{n \in \mathbb{N}} |x_n(t) - x_{n+1}(t)| \leq (1/2)^n \implies \forall_{n \leq m} |x_n(t) - x_m(t)| \leq \sum_{k=n}^m (1/2)^k.$$

Zatem dla $t \in \Omega \setminus A$ ciąg liczbowy $\{x_n(t)\}_{n=1}^\infty$ jest Cauchy, czyli istnieje granica $x(t) := \lim_{n \rightarrow \infty} x_n(t)$. Ponadto

$$|x(t) - x_n(t)| = \lim_{m \rightarrow \infty} |x_m(t) - x_n(t)| \leq \lim_{m \rightarrow \infty} \sum_{k=n}^m (1/2)^k = \sum_{k=n}^{\infty} (1/2)^k = (1/2)^{n-1}.$$

Zatem, kładąc dla $t \in A$ np. $x(t) := 0$, dostajemy funkcję, $x : \Omega \rightarrow \mathbb{F}$ dla której

$$\|x - x_n\|_\infty \leq \sup_{t \in \Omega \setminus A} |x(t) - x_n(t)| \leq (1/2)^{n-1} \longrightarrow 0, \quad \text{przy } n \rightarrow \infty.$$

Czyli $x_n \xrightarrow{\|\cdot\|_\infty} x$ oraz w szczególności $\|x\|_\infty \leq \|x - x_n\|_\infty + \|x_n\|_\infty < \infty$, skąd $x \in L^\infty(\mu)$. $\square$

Uwaga 1.47. Nierówność Höldera (Twierdzenie 1.30) zachodzi dla $p = 1$ i $q = \infty$, tzn.

$$\|xy\|_1 \leq \|x\|_1 \|y\|_\infty.$$

Rzeczywiście, jeśli $A \in \Sigma$ taki, że $\mu(A) = 0$ oraz $\|x\|_\infty = \sup_{t \in \Omega \setminus A} |y(t)|$, to

$$\|xy\|_1 = \int_{\Omega} |xy| d\mu = \int_{\Omega \setminus A} |xy| d\mu \leq \int_{\Omega \setminus A} |x| \cdot \sup_{t \in \Omega \setminus A} |y(t)| d\mu \leq \int_{\Omega} |x| \cdot \|y\|_\infty d\mu = \|x\|_1 \cdot \|y\|_\infty.$$

Przykład 1.48 (Przestrzeń funkcji ograniczonych). Jeśli μ jest miarą liczącą, to jedynym zbiorem miary zero jest zbiór pusty. Zatem w tym przypadku $L^\infty(\mu)$ pokrywa się z *przestrzenią funkcji ograniczonych*

$$B(\Omega) := \{x : \Omega \rightarrow \mathbb{F} : \sup_{t \in \Omega} |x(t)| < \infty\}$$

z normą $\|x\|_\infty := \sup_{t \in \Omega} |x(t)|$. Oczywiście, jeśli $\Omega = \mathbb{N}$, to $L^\infty(\mu) = \ell^\infty$ jest przestrzenią ciągów ograniczonych z normą $\|x\|_\infty = \sup_{k \in \mathbb{N}} |x(k)|$, a jeśli $\Omega = \{1, \dots, n\}$, to $L^\infty(\mu) = \mathbb{F}^n$ z normą $\|x\|_\infty = \max_{k=1, \dots, n} |x(k)|$. Jeśli Ω jest przestrzenią topologiczną, to przestrzeń funkcji ciągłych i ograniczonych $C_b(\Omega) \subseteq B(\Omega)$ jest domkniętą podprzestrzenią liniową przestrzeni $B(\Omega)$.

Zbieżność w przestrzeniach L^p dla $1 \leq p < \infty$ jest trudna do scharakteryzowania w terminach innych znanych zbieżności funkcji, por. Uwaga 1.37. Inaczej jest ze zbieżnością w przestrzeni L^∞ . Przypomnijmy, że zbieżność w normie supremum jest równoważna zbieżności jednostajnej $\Rightarrow$, patrz strona 7. Natomiast zbieżność w L^∞ jest równoważna zbieżności jednostajnej poza zbiorem miary zero:

Lemat 1.49. $x_n \xrightarrow{L^\infty} x \iff x_n|_{\Omega \setminus A} \Rightarrow x|_{\Omega \setminus A}$ dla pewnego $A \in \Sigma$, gdzie $\mu(A) = 0$.

Dowód. Implikacja ‘ $\Leftarrow$ ’ jest jasna. Załóżmy więc, że $x_n \xrightarrow{L^\infty} x$. Na mocy Lematu 1.44, dla każdego $n \in \mathbb{N}$ istnieje zbiór $A_n \in \Sigma$ miary zero taki, że $\|x_n - x\|_\infty = \sup_{t \in \Omega \setminus A_n} |x_n(t) - x(t)|$. Wtedy $A := \bigcup_{n=1}^\infty A_n$ jest zbiorem miary zero takim, że $\|x_n - x\|_\infty = \sup_{t \in \Omega \setminus A} |x_n(t) - x(t)|$ dla każdego $n \in \mathbb{N}$. Czyli $x_n|_{\Omega \setminus A} \Rightarrow x|_{\Omega \setminus A}$. $\square$

1.3 Operatory ograniczone

Ustalmy przestrzeń unormowaną X i Y nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$. Normy na X i Y (mimo, iż formalnie są to dwie różne funkcje) będziemy oznaczać tym samym symbolem $\|\cdot\|$. Z kontekstu będzie wynikać, w której przestrzeni dana norma jest liczona.

Operatorem liniowym z X do Y nazywamy odwzorowanie liniowe $T : X \rightarrow Y$, tj. funkcję spełniającą

$$T(x + y) = T(x) + T(y) \quad \text{oraz} \quad T(\lambda x) = \lambda T(x),$$

dla wszystkich $x, y \in X$ i $\lambda \in \mathbb{F}$. Przy czym, jeśli nie prowadzi to do nieporozumień to, zwyczajowo piszemy Tx w miejsce $T(x)$. Wtedy powyższe wzory przyjmują postać:

$$T(x + y) = Tx + Ty \quad \text{oraz} \quad T\lambda x = \lambda Tx.$$

Nazwa *operator* wzięła się od przykładów takich odwzorowań pochodzących od różnych operacji wykonywanych na funkcjach (np. operacja całkowania lub różniczkowania). Dodatkowo nazwa ta jest poręczna językowo, gdyż elementy rozpatrywanych przez nas przestrzeni są zazwyczaj przestrzeniami funkcji, a wyrażenie “operator w działaniu na funkcji” brzmi nieco lepiej niż “funkcja liniowa w działaniu na funkcji”. Zatem trzeba zapamiętać, że:

$$\text{operator liniowy} \quad \equiv \quad \text{odwzorowanie liniowe}$$

Co więcej często będziemy opuszczać przymiotnik “liniowy”, jako że innych operatorów tu rozpatrywać nie będziemy!

Dodatkowo, będziemy zainteresowani badaniem wyłącznie operatorów ograniczonych, tj. spełniających równoważne warunki w poniższym stwierdzeniu:

Stwierdzenie 1.50. *Operator liniowy $T : X \rightarrow Y$ przekształca zbiory ograniczone na zbiory ograniczone $\iff$ istnieje stała $C \geq 0$, dla której zachodzi “nierówność ograniczoności”:*

$$\forall x \in X \quad \|Tx\| \leq C\|x\|.^2 \tag{1.3}$$

Zachodzi również następująca równość, której wspólną wartość oznaczamy przez $\|T\|$:

$$\|T\| := \inf\{C \geq 0 : \forall x \in X \quad \|Tx\| \leq C\|x\|\} = \sup_{\|x\|=1} \|Tx\|. \tag{1.4}$$

²Zauważmy, że norma $\|Tx\|$ jest liczona w przestrzeni Y , a norma $\|x\|$ w przestrzeni X . W szczególności, formalnie warunek (1.3) powinno się zapisać jako

$$\forall x \in X \quad \|Tx\|_Y \leq C\|x\|_X,$$

gdzie $\|\cdot\|_X$ oznacza normę na X , a $\|\cdot\|_Y$ normę na Y .

Dowód. “ $\implies$ ”: Niech $S := \{x \in X : \|x\| = 1\}$ będzie sferą jednostkową w X . Zbiór S jest ograniczony w X , więc z założenia $T(S) = \{Tx : \|x\| = 1\}$ jest zbiorem ograniczonym w Y . Zatem $C := \sup_{\|x\|=1} \|Tx\| = \sup_{y \in T(S)} \|y\| < \infty$. Pokażemy, że C spełnia (1.3). Niech $x \in X$ dowolny. Jeśli $x = 0$, to $\|Tx\| = \|0\| = 0 = C \cdot 0 = C \cdot \|x\|$, czyli (1.3) zachodzi trywialnie. Niech więc $x \neq 0$. Wtedy $\frac{x}{\|x\|} \in S$ i stąd

$$\left\| T \frac{x}{\|x\|} \right\| \leq C \iff \frac{\|T(x)\|}{\|x\|} \leq C \iff \|Tx\| \leq C\|x\|.$$

To dowodzi (1.3) oraz że $\inf\{C \geq 0 : \forall x \in X \quad \|Tx\| \leq C\|x\|\} \leq \sup_{\|x\|=1} \|Tx\|$.

“ $\impliedby$ ”: Niech C spełnia (1.3) i niech $K \subseteq X$ zbiór ograniczony. Wtedy istnieje $C' > 0$ takie, że dla $x \in K$ mamy $\|x\| \leq C'$. Stąd dla $x \in K$ mamy $\|Tx\| \leq C\|x\| \leq C \cdot C' < \infty$. Zatem $TK = \{Tx : x \in K\}$ jest zbiorem ograniczonym. Co więcej,

$$\forall x \in X \quad \|Tx\| \leq C\|x\| \implies \forall x \in X \neq 0 \quad \|T(x/\|x\|)\| \leq C \implies \sup_{\|x\|=1} \|Tx\| \leq C.$$

Skąd $\sup_{\|x\|=1} \|Tx\| \leq \inf\{C > 0 : \forall x \in X \quad \|Tx\| \leq C\|x\|\}$. Przeciwną nierówność wykazaliśmy w dowodzie implikacji “ $\implies$ ”. Stąd równość w (1.4). $\square$

Definicja 1.51. Operator liniowy T , dla którego istnieje stała C spełniająca nierówność ograniczoności (1.3) nazywamy *operatorem ograniczonym*. Natomiast liczbę $\|T\|$ daną wzorem (1.4) nazywamy *normą operatora* T .

Uwaga 1.52. Operator T jest ograniczony wtedy i tylko wtedy, gdy jego norma (1.4) jest skończona, tj. gdy $\sup_{\|x\|=1} \|Tx\| < \infty$. Norma operatora T jest najmniejszą stałą, dla której zachodzi nierówność ograniczoności. W szczególności

$$\forall x \in X \quad \|Tx\| \leq \|T\| \cdot \|x\|.$$

Uwaga 1.53. Przy obliczaniu normy operatora oba wzory występujące w (1.4) są przydatne. Równość $\|T\| = \inf\{C \geq 0 : \forall x \in X \quad \|Tx\| \leq C\|x\|\}$ pozwala znaleźć *ograniczenie górne*: Jeżeli C spełnia nierówność ograniczoności, to $\|T\| \leq C$. Równość $\|T\| = \sup_{\|x\|=1} \|Tx\|$ daje *ograniczenie dolne*: Jeżeli $\|x\| = 1$ to $\|Tx\| \leq \|T\|$. Schematycznie obliczanie $\|T\|$ składa się zazwyczaj z dwóch kroków:

$$\|T\| = C \iff \begin{cases} 1) & \forall x \in X \quad \|Tx\| \leq C\|x\| \\ 2) & \exists \{x_n\}_{n=1}^\infty \subseteq X, \quad \|x_n\|=1, n \in \mathbb{N} \quad \|Tx_n\| \rightarrow C \end{cases}$$

tzn. trzeba znaleźć jak najmniejsze C takie, żeby zachodziło 1), a następnie ciąg unormowanych wektorów w X takich by zachodziło 2). Wtedy $\|T\| = C$.

Jednym z podstawowych faktów (bez znajomości którego żaden student nie może liczyć na pozytywną ocenę z egzaminu z analizy funkcjonalnej!!!) jest następująca równoważność dla operatorów liniowych:

$$\text{operator ograniczony} \quad \equiv \quad \text{operator ciągły}.$$

Mówiąc bardziej “uczenie” powyższa równoważność oznacza, że w kategorii przestrzeni unormowanych operatory ograniczone są naturalnymi morfizmami, bo zachowują zarówno strukturę liniową jak i topologiczną. Mianowicie, zachodzi następujące:

Twierdzenie 1.54. Niech $T : X \rightarrow Y$ operator liniowy między dwiema przestrzeniami unormowanymi. Następujące warunki są równoważne:

- (1) T jest ograniczony;
- (2) T jest odwzorowaniem ciągłym;
- (3) T jest ciągłe w pewnym punkcie $x_0 \in X$;
- (4) T jest ciągłe w zerze.

Dowód. (1) $\Rightarrow$ (2). Wynika to z własności Lipschitza dla operatorów ograniczonych. Rzeczywiście, dla dowolnych $x, y \in X$ mamy $\|Tx - Ty\| = \|T(x - y)\| \leq \|T\|\|x - y\|$. Zatem jeśli $x \rightarrow y$, to $Tx \rightarrow Ty$.

(2) $\Rightarrow$ (3). Z definicji, funkcja ciągła jest ciągła w każdym punkcie.

(3) $\Rightarrow$ (4). Jeśli $x_n \rightarrow 0$, to $x_n + x_0 \rightarrow x_0$, a zatem z ciągłości T w x_0 mamy $Tx_n = T(x_n + x_0) - Tx_0 \rightarrow Tx_0 - Tx_0 = 0 = T(0)$. Czyli operator T jest ciągły w zerze.

(4) $\Rightarrow$ (1). Załóżmy nie wprost, że T nieograniczony. Wówczas istnieje ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ taki, że $\|x_n\| = 1$ oraz $\|Tx_n\| \geq n$ dla $n \in \mathbb{N}$. Wtedy $\|\frac{x_n}{\sqrt{n}}\| = \frac{1}{\sqrt{n}} \rightarrow 0$, czyli $\frac{x_n}{\sqrt{n}} \rightarrow 0$. Natomiast, $\|T\frac{x_n}{\sqrt{n}}\| \geq n \cdot \frac{1}{\sqrt{n}} = \sqrt{n} \rightarrow \infty$. Zatem T nie jest ciągły w zerze. $\square$

Jedną z ważniejszych metod konstrukcji operatorów ograniczonych jest zdefiniowanie najpierw operatora na gęstej podprzestrzeni, a następnie rozszerzenie go na całą podprzestrzeń za pomocą następującej zasady ciągłego przedłużenia operatorów liniowych.

Twierdzenie 1.55 (Zasada ciągłego przedłużania operatorów). *Każdy ograniczony operator liniowy $T : M \rightarrow Y$ zdefiniowany na podprzestrzeni M przestrzeni unormowanej X o wartościach w przestrzeni Banacha Y przedłuża się jednoznacznie do ograniczonego operatora liniowego $\bar{T} : \bar{M} \rightarrow Y$ na domknięciu przestrzeni M . Ponadto*

$$\|T\| = \|\bar{T}\|.$$

W szczególności, jeśli M jest gęstą podprzestrzenią X , to T jednoznacznie przedłuża się do operatora ograniczonego na X .

Dowód. Jeśli $x_0 \in \bar{M}$, to istnieje ciąg $\{x_n\}_{n=1}^\infty \subseteq M$ zbieżny do x_0 . Wtedy ciąg $\{Tx_n\}_{n=1}^\infty \subseteq Y$ jest ciągiem Cauchy w Y , gdyż

$$\|Tx_n - Tx_m\| \leq \|T\| \cdot \|x_n - x_m\| \rightarrow 0, \quad \text{przy } n, m \rightarrow \infty.$$

Zatem na mocy zupełności Y ciąg $\{Tx_n\}_{n=1}^\infty$ jest zbieżny do pewnego $y_0 \in Y$. Zauważmy, że punkt y_0 zależy tylko od punktu x_0 , tzn. nie zależy od wyboru ciągu $\{x_n\}_{n=1}^\infty$. Rzeczywiście, jeżeli $\{x'_n\}_{n=1}^\infty \subseteq M$ inny ciąg zbieżny do x_0 , to ciąg $\{Tx'_n\}_{n=1}^\infty$ jest zbieżny do y_0 , gdyż

$$\lim_{n \rightarrow \infty} \|y_0 - Tx'_n\| = \lim_{n, m \rightarrow \infty} \|Tx_m - Tx'_n\| \leq \lim_{n, m \rightarrow \infty} \|T\| \cdot \|x_m - x'_n\| = \|T\| \cdot \|x_0 - x_0\| = 0.$$

Zatem kładąc $\bar{T}x_0 := y_0$ otrzymujemy poprawnie zdefiniowaną funkcję $\bar{T} : \bar{M} \rightarrow Y$. Jako że granica jest operacją liniową, z powyższej konstrukcji natychmiast wynika, że $\bar{T}$ jest operatorem liniowym. Jasne jest, że $\bar{T}x_0 = Tx_0$ dla $x_0 \in M$ (bo jeśli $x_0 \in M$, to za ciąg $\{x_n\}_{n=1}^\infty$ zbieżny do x_0 możemy wziąć ciąg stały $\{x_0\}_{n=1}^\infty$). Czyli $\bar{T}$ jest przedłużeniem T . Ponadto, dla każdego $x_0 \in \bar{M}$ oraz ciągu $\{x_n\}_{n=1}^\infty \subseteq M$ zbieżnego do x_0 mamy

$$\|\bar{T}x_0\| = \lim_{n \rightarrow \infty} \|Tx_n\| \leq \lim_{n \rightarrow \infty} \|T\| \cdot \|x_n\| = \|T\| \cdot \|x_0\|.$$

Czyli operator $\bar{T}$ jest liniowy oraz $\|\bar{T}\| \leq \|T\|$. De facto $\|\bar{T}\| = \|T\|$ ponieważ $\|\bar{T}\| \geq \|T\|$ wynika wprost stąd, że $\bar{T}$ jest przedłużeniem T . $\square$

Uwaga 1.56. W końcowej części dowodu Twierdzenia 1.11 mieliśmy do czynienia ze szczególnym przypadkiem Twierdzenia 1.55, gdzie wyjściowy operator $T : M \rightarrow Y$ jest izometrią. Wtedy jego przedłużenie $\bar{T} : \bar{M} \rightarrow Y$ też jest izometrią.

1.3.1 Przykłady operatorów ograniczonych

Uwaga 1.57. Najważniejsze operacje w Analizie są liniowe! Co więcej, operatory liniowe określone na przestrzeni Banacha (unormowanej, zupełnej) “z natury” są ograniczone.³

Przykład 1.58 (Całkowanie). Niech $X := L^1(\mu)$ będzie rzeczywistą przestrzenią funkcji całkowlanych względem miary μ oraz niech $Y := \mathbb{R}$ (jednowymiarowa przestrzeń Banacha nad $\mathbb{R}$). Wtedy całka

$$Tx := \int_{\Omega} x(t) d\mu(t), \quad x \in L^1(\mu),$$

jest operatorem liniowym $T : L^1(\mu) \rightarrow \mathbb{R}$. Jest to operator ograniczony o normie 1. Rzeczywiście,

$$|Tx| = \left| \int_{\Omega} x(t) d\mu(t) \right| \leq \int_{\Omega} |x(t)| d\mu(t) = 1 \cdot \|x\|_1 \implies \|T\| \leq 1.$$

Z drugiej strony biorąc dowolny zbiór mierzalny $A \subseteq \Omega$ taki, że $0 < \mu(A) < \infty$ ⁴ i kładąc $x := \frac{1}{\mu(A)} \mathbb{1}_A$ mamy $\|x\|_1 = 1$ oraz $|Tx| = 1$, skąd $1 \leq \|T\|$.

Przykład 1.59 (Różniczkowanie). Niech $X := C^{(1)}([0, 1])$ będzie przestrzenią funkcji różniczkowalnych w sposób ciągły i niech $Y = C([0, 1])$ przestrzeń funkcji ciągłych. Obie przestrzenie rozpatrujemy z normą supremum (zauważmy, że $C^{(1)}([0, 1])$ jest podprzestrzenią liniową $C([0, 1])$). Operator różniczkowania $T = \frac{d}{dt} : X \rightarrow Y$, tzn.

$$(Tx)(t) := x'(t) \quad x \in C^{(1)}([0, 1]), t \in [0, 1],$$

jest poprawnie określony i liniowy. W rozpatrywanych normach operator T jest **nieograniczony**! Rzeczywiście, kładąc $x_n(t) = t^n$ mamy

$$\|x_n\|_{\infty} = 1 \quad \text{oraz} \quad \|Tx_n\| = \|x'_n\|_{\infty} = \sup_{t \in [0, 1]} |nt^{n-1}| = n \rightarrow \infty.$$

Nie jest to sprzeczne z Uwagą 1.57, gdyż przestrzeń $C^{(1)}([0, 1])$ z normą supremum nie jest zupełna (🏠 wykazać to). Standardowa norma na przestrzeni $C^{(1)}([0, 1])$ jest dana wzorem

$$\|x\|_1 = \|x\|_{\infty} + \|x'\|_{\infty} = \max_{t \in [0, 1]} |x(t)| + \max_{t \in [0, 1]} |x'(t)|.$$

Przestrzeń $X = C^{(1)}([0, 1])$ z normą $\|\cdot\|_1$ jest przestrzenią Banacha (🏠 wykazać to). Z tą normą na X operator T jest ograniczony i jego norma wynosi 1. Rzeczywiście, $\|Tx\|_{\infty} = \|x'\|_{\infty} \leq 1 \cdot \|x\|_1$, czyli $\|T\| \leq 1$. Zauważmy, że aby udowodnić nierówność przeciwną nie uda się nam znaleźć funkcji $x \in X$ takiej, że $\|x\|_1 = 1$ oraz $\|Tx\|_{\infty} = 1$, bo wtedy mielibyśmy $\|x\|_{\infty} + \|x'\|_{\infty} = 1 = \|x'\|_{\infty}$, co implikuje, że $\|x\|_{\infty} = 0$ czyli $x \equiv 0$ oraz $x' \neq 0$ (bo $\|x'\|_{\infty} = 1$) ⚡. Natomiast, możemy znaleźć funkcje, dla których $\|x\|_1 = 1$ oraz $\|Tx\|_{\infty}$ jest dowolnie bliskie 1 (równoważenie $\|x\|_{\infty}$ jest dowolnie bliskie zera). Na przykład, kładąc $x_n(t) = \frac{t^n}{n+1}$ mamy

$$\|x_n\|_1 = \max_{t \in [0, 1]} \left| \frac{t^n}{n+1} \right| + \max_{t \in [0, 1]} \left| \frac{nt^n}{n+1} \right| = \frac{1}{n+1} + \frac{n}{n+1} = 1$$

oraz

$$\|Tx_n\|_{\infty} = \max_{t \in [0, 1]} \left| \frac{nt^n}{n+1} \right| = \frac{n}{n+1} \rightarrow 1.$$

Zatem $\|T\| = 1$, por. Uwaga 1.53.

³Jeśli na teście będzie pytanie, czy dany (zdefiniowany w sposób jawny) operator liniowy na przestrzeni Banacha jest ograniczony, to śmiało można odpowiedzieć, że tak.

⁴Jeśli taki zbiór nie istnieje, to $X = \{0\}$ jest przestrzenią zerową, a tak nierozsądnie zdegenerowaną sytuację będziemy ignorować ☺

Uwaga 1.60. Na Przykładzie 1.59 widzimy, że mimo, iż nie uwzględniamy tego w notacji, to *norma operatora* $T : X \rightarrow Y$ *zależy od norm na przestrzeniach* X *i* Y ! Ponadto, na ogół, supremum we wzorze $\|T\| = \sup_{\|x\|=1} \|Tx\| < \infty$ nie można zastąpić przez maksimum. Dlatego w kroku 2) w Uwadze 1.53, na ogół, potrzebujemy ciągu wektorów $\{x_n\}_{n=1}^\infty$, a nie pojedynczego wektora (jak miało to miejsce w Przykładach 1.61, 1.58).

Przykład 1.61 (Zamiana zmiennych). Niech $X := C_b(N)$ i $Y := C_b(M)$ będą przestrzeniami funkcji ciągłych i ograniczonych na przestrzeniach topologicznych M i N (z normą supremum). Niech $\varphi : M \rightarrow N$ będzie odwzorowaniem ciągłym. *Operator kompozycji* $T : X \rightarrow Y$ dany wzorem

$$(Tx)(t) := x(\varphi(t)), \quad x \in C_b(N), t \in M,$$

jest poprawnie określony (bo $x \circ \varphi : M \xrightarrow{\varphi} N \xrightarrow{x} \mathbb{F}$ jest funkcją ciągłą i ograniczoną) oraz liniowy. Operator ten „zamienia zmienne” za pomocą odwzorowania φ . Zauważmy, że

$$\|Tx\|_\infty = \sup_{t \in M} |x(\varphi(t))| = \sup_{s \in \varphi(M) \subseteq N} |x(s)| \leq \sup_{s \in N} |x(s)| = \|x\|_\infty.$$

Zatem T jest operatorem ograniczonym oraz $\|T\| \leq 1$. W rzeczywistości $\|T\| = 1$, gdyż dla funkcji $x \equiv 1$ (tożsamościowo równej 1) mamy $\|x\|_\infty = 1$ oraz $\|Tx\|_\infty = \|x\|_\infty = 1$, skąd $1 \leq \|T\|$, por. Uwaga 1.53.

To że normy operatorów w powyższych przykładach wynoszą 1 jest przypadkiem. Aby czytelnik nie odniósł mylnego wrażenia, że tak musi być zawsze, rozważmy jeszcze jedną ważną klasę operatorów. Poniższe obliczenia pokazują, że wyznaczenie normy operatora jest często zadaniem nietrywialnym.

Przykład 1.62 (Mnożenie przez funkcje). Niech $X = Y = L^p([0, 1])$, $1 \leq p < \infty$, przestrzeń funkcji całkowalnych w p -tej potęgze względem miary Lebesgue’a na $[0, 1]$. Dla dowolnej funkcji ciągłej $a \in C([0, 1])$ *operator mnożenia*

$$(T_a x)(t) := a(t)x(t), \quad x \in X, t \in [0, 1],$$

jest porownie określonym operatorem liniowym $T_a : X \rightarrow X$. Twierdzimy, że jest to operator ograniczony oraz $\|T_a\| = \|a\|_\infty = \max_{t \in [0, 1]} |a(t)|$. Rzeczywiście,

$$\|T_a x\|_p = \left(\int_\Omega |a(t)x(t)|^p d\mu(t) \right)^{\frac{1}{p}} \leq \left(\int_\Omega \|a\|_\infty \cdot |x(t)|^p d\mu(t) \right)^{\frac{1}{p}} = \|a\|_\infty \cdot \|x\|_p.$$

Skąd $\|T_a\| \leq \|a\|_\infty$. Aby pokazać nierówność przeciwną, dla każdego $n \in \mathbb{N}$ połóżmy $A_n := \{t \in [0, 1] : |a(t)|^p - \|a\|_\infty^p \leq 1/n\}$ oraz $x_n := \frac{1}{\mu(A_n)^{1/p}} \mathbf{1}_{A_n}$. Wtedy $\|x_n\|_p = 1$ oraz

$$\begin{aligned} \|T_a x_n\|_p &= \frac{1}{\mu(A_n)^{1/p}} \cdot \left(\int_{A_n} |a(t)|^p d\mu(t) \right)^{\frac{1}{p}} \geq \frac{1}{\mu(A_n)^{1/p}} \cdot \left(\int_{A_n} \|a\|_\infty^p - \frac{1}{n} d\mu(t) \right)^{\frac{1}{p}} \\ &= \left(\|a\|_\infty^p - \frac{1}{n} \right)^{\frac{1}{p}} \rightarrow \|a\|_\infty. \end{aligned}$$

Zatem $\|T_a x_n\|_p \rightarrow \|a\|_\infty$ i stąd $\|T_a\| \geq \|a\|_\infty$.

1.3.2 Równoważność norm i przestrzenie skończone wymiarowe

Rozważmy dwie normy $\|\cdot\|_1$ i $\|\cdot\|_2$ na przestrzeni liniowej X .

Definicja 1.63. Powiemy, że norma $\|\cdot\|_1$ jest *mocniejsza* od normy $\|\cdot\|_2$ (lub że $\|\cdot\|_2$ jest *słabsza* od $\|\cdot\|_1$) jeżeli

$$\exists_{c>0} \forall_{x \in X} \quad \|x\|_2 \leq c\|x\|_1.$$

Normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są *równoważne* jeżeli norma $\|\cdot\|_1$ jest jednocześnie mocniejsza i słabsza od $\|\cdot\|_2$, czyli gdy

$$\exists_{c_1, c_2 > 0} \forall_{x \in X} \quad \|x\|_2 \leq c_1\|x\|_1 \quad \text{oraz} \quad \|x\|_1 \leq c_2\|x\|_2.$$

Stwierdzenie 1.64. *Następujące warunki są równoważne:*

- (1) norma $\|\cdot\|_1$ jest mocniejsza, niż $\|\cdot\|_2$;
- (2) zbieżność ciągu w normie $\|\cdot\|_1$ pociąga zbieżność w normie $\|\cdot\|_2$;
- (3) topologia zadana przez $\|\cdot\|_1$ jest mocniejsza (większa) od topologii zadanej przez $\|\cdot\|_2$.

Jeśli $\|\cdot\|_1$ jest mocniejsza od $\|\cdot\|_2$ i przestrzeń $(X, \|\cdot\|_2)$ jest zupełna, to $(X, \|\cdot\|_1)$ jest zupełna.

Dowód. Zauważmy, że norma $\|\cdot\|_1$ jest mocniejsza od normy $\|\cdot\|_2$ wtedy i tylko wtedy, gdy operator identycznościowy $id : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ jest ograniczonym. Na mocy Twierdzenia 1.54 jest to równoważne temu, że odwzorowanie $id : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ jest ciągłe. Korzystając z ciągowej charakteryzacji ciągłości funkcji (kryterium Heinego) otrzymujemy, że $id : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ jest ciągłe wtedy i tylko wtedy, gdy zbieżność w $\|\cdot\|_1$ pociąga zbieżność $\|\cdot\|_2$. Zatem (1) i (2) są równoważne. Korzystając z topologicznej definicji ciągłości otrzymujemy, że $id : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ jest ciągłe wtedy i tylko wtedy topologia zadana przez $\|\cdot\|_2$ zawiera się w topologii przez $\|\cdot\|_1$, czyli (1) i (3) są równoważne. $\square$

Wniosek 1.65. *Następujące warunki są równoważne:*

- (1) normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są równoważne;
- (2) zbieżność ciągu w normie $\|\cdot\|_1$ jest równoważna zbieżności ciągu w normie $\|\cdot\|_2$;
- (3) topologie zadane przez $\|\cdot\|_1$ i $\|\cdot\|_2$ są takie same.

Ponadto jeśli normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są równoważne, to

$$(X, \|\cdot\|_1) \text{ jest przestrzenią Banacha} \iff (X, \|\cdot\|_2) \text{ jest przestrzenią Banacha}.$$

Dowód. Równoważność warunków (1)-(3) wynika natychmiast z dwukrotnego zastosowania Stwierdzenia 1.64. Ponadto, jeśli normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są równoważne, to ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ jest Cauchy w normie $\|\cdot\|_1$ wtedy i tylko wtedy, gdy $\{x_n\}_{n=1}^\infty$ jest ciągiem Cauchy w normie $\|\cdot\|_2$, bo

$$\|x_n - x_m\|_2 \leq c_1\|x_n - x_m\|_1 \quad \text{oraz} \quad \|x_n - x_m\|_1 \leq c_2\|x_n - x_m\|_2.$$

Ta uwaga wraz z równoważnością zbieżności w normach $\|\cdot\|_1$ i $\|\cdot\|_2$ daje równoważność zupełności przestrzeni $(X, \|\cdot\|_1)$ i $(X, \|\cdot\|_2)$. $\square$

Przykład 1.66. Na przestrzeni $X = C[0, 1]$ funkcji ciągłych na przedziale $[0, 1]$ normy

$$\|x\|_\infty := \max_{t \in [0, 1]} |x(t)|, \quad \|x\|_p := \left(\int_0^1 |x(t)|^p dt \right)^{1/p}, \quad p \in [1, \infty)$$

nie są równoważne. Rzeczywiście, niech $x_n(t) = t^n$. Wtedy $\|x_n\|_1 = \int_0^1 t^n dt = \frac{1}{n+1} \rightarrow 0$, czyli $\{x_n\}_{n=1}^\infty$ zbiega do zera w normie $\|\cdot\|_1$. Natomiast $\|x_n\|_\infty = \max_{t \in [0, 1]} |t^n| = 1 \not\rightarrow 0$. Zatem $\{x_n\}_{n=1}^\infty$ nie może zbiegać do zera w normie $\|\cdot\|_\infty$ (de facto w tej normie jest rozbieżny).

Przykład 1.67. W przestrzeni $X = \mathbb{F}^n$ normy

$$\|x\|_\infty := \max_{k=1, \dots, n} |x(k)|, \quad \|x\|_p := \left(\sum_{k=1}^n |x(k)|^p \right)^{1/p}, \quad p \in [1, \infty)$$

są równoważne, ponieważ $\|x\|_\infty \leq \|x\|_p$ oraz $\|x\|_p \leq n\|x\|_\infty$ dla każdego $x \in X$ i $p \in [1, \infty)$.

Równoważność norm w Przykładzie 1.67, to nie przypadek, gdyż zachodzi następujące

Twierdzenie 1.68. W przestrzeni skończenie wymiarowej każde dwie normy są równoważne.

Dowód. Niech X będzie przestrzenią skończenie wymiarową. Czyli istnieje skończona baza *Hamela*, tzn. liniowo niezależne elementy $e_1, \dots, e_n \in X$ takie, że $\forall x \in X \exists x(1), \dots, x(n) \in \mathbb{F} x = \sum_{k=1}^n x(k)e_k$. Współczynniki $x(1), \dots, x(n) \in \mathbb{F}$, w powyższej sumie, są jednoznacznie wyznaczone przez x . Zdefiniujmy na X normę wzorem

$$\|x\|_\infty := \max_{k=1, \dots, n} |x(k)|, \quad \text{gdzie } x = \sum_{k=1}^n x(k)e_k.$$

Wystarczy pokazać, że dowolna norma $\|\cdot\|$ na X jest równoważna z $\|\cdot\|_\infty$ (bo równoważność norm jest relacją równoważności, a w szczególności jest relacją przechodnią). Zauważmy, że

$$\|x\| = \left\| \sum_{k=1}^n x(k)e_k \right\| \stackrel{(N3)}{\leq} \sum_{k=1}^n \|x(k)e_k\| \stackrel{(N2)}{=} \sum_{k=1}^n |x(k)| \|e_k\| \leq \|x\|_\infty \sum_{k=1}^n \|e_k\|.$$

Zatem kładąc $c := \sum_{k=1}^n \|e_k\|$ mamy

$$\|x\| \leq c\|x\|_\infty \quad \text{dla każdego } x \in X. \quad (1.5)$$

Założmy nie wprost, że nie istnieje analogiczna stała $c > 0$ taka, że $\|x\|_\infty \leq c\|x\|$ dla każdego $x \in X$. Wtedy dla każdego $n \in \mathbb{N}$ znajdziemy $x_n \in X$ takie, że $\|x_n\|_\infty > n\|x_n\|$. Kładąc $y_n := \frac{x_n}{\|x_n\|_\infty}$ mamy

$$\|y_n\| = \frac{\|x_n\|}{\|x_n\|_\infty} < \frac{1}{n} \rightarrow 0.$$

Zatem z jednej strony $y_n \xrightarrow{\|\cdot\|} 0$. Z drugiej strony, jednak strony $\|y_n\|_\infty = 1$. Czyli ciąg $\{y_n\}_{n=1}^\infty$ jest zawarty w sferze jednostkowej w normie $\|\cdot\|_\infty$. Jako że norma $\|\cdot\|_\infty$ jest równoważna normie euklidesowej $\|\cdot\|_2$ (patrz Przykład 1.67), to ma zastosowanie Twierdzenie Bolzano-Weierstrassa.⁵ Zatem istnieje podciąg $\{y_{n_k}\}_{n_k=1}^\infty$ zbieżny do pewnego $y \in X$ w normie $\|\cdot\|_\infty$. Wtedy $\|y\|_\infty = \lim_{k \rightarrow \infty} \|y_{n_k}\|_\infty = 1$ i w szczególności $y \neq 0$. Ale z (1.5) wynika, że

$$\|y_{n_k} - y\| \leq c\|y_{n_k} - y\|_\infty \rightarrow 0 \text{ przy } k \rightarrow \infty.$$

Czyli $y_{n_k} \xrightarrow{\|\cdot\|} y \neq 0$, co jest sprzeczne z tym, że $y_n \xrightarrow{\|\cdot\|} 0$. □

⁵Twierdzenie Bolzano-Weierstrassa mówi, że w przestrzeni euklidesowej każdy ciąg ograniczony zawiera podciąg zbieżny

Wniosek 1.69. *Każda unormowana przestrzeń skończenie wymiarowa jest zupełna (jest przestrzenią Banacha).*

Dowód. Na mocy Twierdzenia 1.68 i ostatniej części Wniosku 1.65 wystarczy pokazać, że przestrzeń $X \cong \mathbb{F}^n$ jest zupełna w jakiejkolwiek normie. Z analizy wiemy, że np. przestrzeń euklidesowa, czyli $\mathbb{F}^n$ z normą $\|\cdot\|_2$ jest zupełna (można to łatwo pokazać 🏠 korzystając z zupełności $\mathbb{R}$). $\square$

Wniosek 1.70. *Każda skończenie wymiarowa podprzestrzeń Y przestrzeni unormowanej X jest domknięta.*

Dowód. W świetle Wniosku 1.69, Y jest przestrzenią Banacha. Zatem na mocy Stwierdzenia 1.10, Y jest domkniętą podprzestrzenią w X . $\square$

Wniosek 1.71. *Każdy operator liniowy określony na przestrzeni skończenie wymiarowej jest ograniczony (ciągły).*

Dowód. Niech $T : X \rightarrow Y$ liniowy i $\dim(X) = n < \infty$. Wtedy istnieje baza $e_1, \dots, e_n$ przestrzeni X . Skoro na X wszystkie normy są równoważne (Twierdzenie 1.68) możemy założyć, że dla $x = \sum_{i=1}^n \alpha_i e_i$ mamy $\|x\| = \|x\|_1 = \sum_{i=1}^n |\alpha_i|$. Wtedy

$$\|Tx\| = \|T \sum_{i=1}^n \alpha_i e_i\| = \|\sum_{i=1}^n \alpha_i T e_i\| \leq \sum_{i=1}^n |\alpha_i| \|T e_i\| \leq C \cdot \|x\|,$$

gdzie $C := \max_{i=1, \dots, n} \|T e_i\|$. Czyli operator T jest ograniczony (równoważnie ciągły). $\square$

1.3.3 Przestrzenie operatorów ograniczonych

Zbiór operatorów ograniczonych między przestrzeniami unormowanymi tworzy w naturalny sposób przestrzeń normowaną. W szczególności nazwa “norma operatora” nie jest przypadkowa, gdyż wzór (1.4) rzeczywiście określa normę na przestrzeni wszystkich operatorów ograniczonych:

Definicja 1.72. Niech X i Y będą przestrzeniami unormowanymi. Zbiór

$$B(X, Y) := \{T : X \rightarrow Y \text{ operator liniowy i ograniczony}\}$$

nazywamy przestrzenią operatorów ograniczonych z X w Y . Przestrzeń operatorów ograniczonych z X w X oznaczamy skrótowo $B(X) := B(X, X)$.

Przestrzeń $B(X, Y)$ wraz z działaniami określonymi punktowo, tzn.

$$(T + S)x := Tx + Sx, \quad (\lambda T)x := \lambda Tx$$

dla $T, S \in B(X, Y)$, $x \in X$, $\lambda \in \mathbb{F} = \mathbb{R}, \mathbb{C}$, oraz normą operatorową $\|T\| = \sup_{\|x\|=1} \|Tx\|$ jest przestrzenią unormowaną (🏠 sprawdzić to). Jej zupełność jest równoważna zupełności Y :

Twierdzenie 1.73. *Przestrzeń operatorów ograniczonych $B(X, Y)$ jest przestrzenią Banacha wtedy i tylko wtedy, gdy Y jest przestrzenią Banacha.*

Dowód. “ $\Rightarrow$ ” Niech $\{y_n\}_{n=1}^\infty \subseteq$ ciąg Cauchy. Aby pokazać, że ciąg ten jest zbieżny zwiążemy z nim pewien ciąg operatorów $\{T_n\}_{n=1}^\infty \subseteq B(X, Y)$ “jednowymiarowych”. Mianowicie, weźmy dowolny wektor unormowany $x_0 \in X$, $\|x_0\| = 1$. Połóżmy

$$T_n x = \begin{cases} \lambda y_n, & \text{jeśli } x = \lambda x_0 \\ 0, & \text{w przeciwnym razie,} \end{cases}$$

tzn. $T x_0 = y_n$ oraz $T_n \equiv 0$ poza jednowymiarową podprzestrzenią $\{\lambda x_0 : \lambda \in \mathbb{F}\} \subseteq X$ rozpiętą przez x_0 . Wtedy

$$\|T_n\| = \sup_{\|x\|=1} \|T_n x\| = \sup_{|\lambda|=1} \|T_n \lambda x_0\| = \|T_n x_0\| = \|y_n\|.$$

Zatem operatory T_n są ograniczone i podobne rozumowanie pokazuje, że

$$\|T_n - T_m\| = \sup_{\|x\|=1} \|(T_n - T_m)x\| = \|T_n x_0 - T_m x_0\| = \|y_n - y_m\|.$$

Czyli $\{T_n\}_{n=1}^\infty \subseteq B(X, Y)$ ciąg Cauchy. Z założenia więc, ciąg $\{T_n\}_{n=1}^\infty$ jest zbieżny w $B(X, Y)$ do pewnego operatora $T \in B(X, Y)$. Wtedy $T_n x \rightarrow T x$ dla każdego $x \in X$ (bo $\|T_n x - T x\| \leq \|T_n - T\| \|x\| \rightarrow 0$). W szczególności, $y_n = T_n x_0 \rightarrow T x_0$. Czyli ciąg $\{y_n\}_{n=1}^\infty \subseteq Y$ jest zbieżny.

“ $\Leftarrow$ ” Niech $\{T_n\}_{n=1}^\infty \subseteq B(X, Y)$ ciąg Cauchy. Dla każdego $x \in X$ mamy $\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\| \rightarrow 0$. Zatem ciąg $\{T_n x\}_{n=1}^\infty \subseteq Y$ jest ciągiem Cauchy w Y . Z założenia więc jest on zbieżny. Jego granicę oznaczmy przez $T x \in Y$. W sposób otrzymujemy operator liniowy $T : X \rightarrow Y$ (operacja brania granicy jest liniowa!). Ponadto, dla każdego $x \in X$,

$$\|(T - T_n)x\| = \lim_{m \rightarrow \infty} \|(T_m - T_n)x\| \leq \limsup_{m \rightarrow \infty} \|T_m - T_n\| \|x\| \leq \sup_{m \geq n} \|T_m - T_n\| \|x\|.$$

Stąd

$$\|T - T_n\| \leq \sup_{m \geq n} \|T_m - T_n\| \rightarrow 0, \text{ gdy } n \rightarrow \infty.$$

Czyli $\{T_n\}_{n=1}^\infty$ zbiega w normie do T . W szczególności, skoro $T_n \in B(X, Y)$ oraz $T - T_n \in B(X, Y)$ (dla dostatecznie dużych n), to $T = T - T_n + T_n \in B(X, Y)$ jest operatorem ograniczonym ($B(X, Y)$ jest przestrzenią liniową). Czyli $\{T_n\}_{n=1}^\infty$ jest zbieżny w przestrzeni $B(X, Y)$. $\square$

Zauważmy, że ciało skalarów $Y := \mathbb{F}$ jest (jednowymiarową) przestrzenią Banacha. Zatem na mocy Twierdzenia 1.73, przestrzeń operatorów liniowych $B(X, \mathbb{F})$ o wartościach skalarnych jest przestrzenią Banacha. Jej elementy nazywamy *funkcjonałami*. Pełnią one kluczową rolę w analizie funkcjonalnej. W szczególności, wśród znawców tematu przeważa opinia, że nazwa “Analiza Funkcjonalna” wzięła się właśnie od funkcjonałów.⁶ Podstawowym przykładem funkcjonału jest całka, patrz Przykład 1.58.

Definicja 1.74. Przestrzeń Banacha $B(X, \mathbb{F})$ nazywamy *przestrzenią sprzężoną* lub też *przestrzenią dualną* do przestrzeni X i oznaczamy ją przez X^* . Elementy przestrzeni $X^* = B(X, \mathbb{F})$ nazywamy *funkcjonałami ograniczonymi* na X .

Uwaga 1.75. Zasadniczą motywacją do rozważania przestrzeni dualnej X^* jest fakt, że pozwala ona zdefiniować następującą formę dwuliniową (zwanej czasem parowaniem)

$$X^* \times X \ni (f, x) \mapsto \langle f, x \rangle := f(x) \in \mathbb{F},$$

która imituje iloczyn skalarny na X . Zanim jednak zagłębimy się w tę nieco abstrakcyjną teorię przestrzeni dualnych (podrozdział 3.2), omówimy wpierw dokładnie przestrzenie z “prawdziwym” iloczynem skalarnym (patrz Rozdział 2).

⁶Konkurencyjna opinia głosi, że przymiotnik “funkcjonalna” odnosi się do funkcji, jako że na analizie funkcjonalnej zazwyczaj badamy przestrzenie funkcji

Rozdział 2

Przestrzenie Hilberta

Przestrzenie Hilberta to specjalny rodzaj przestrzeni Banacha, w których norma pochodzi od iloczynu skalarnego. Iloczyn skalarny zadaje na danej przestrzeni liniowej strukturę metryczną (normę) oraz geometryczną (pojęcie kąta, ortogonalność).

2.1 Iloczyn skalarny i norma

Zacniemy od omówienia samego iloczynu skalaranego oraz normy, którą definiuje.

Definicja 2.1. *Przestrzenią unitarną (lub też prehilbertowską) nazywamy przestrzeń liniową H nad $\mathbb{F} = \mathbb{R}, \mathbb{C}$ wraz z iloczynem skalarnym, tzn. funkcją dwóch zmiennych $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{F}$ spełniającą*

- (1) $\langle x, x \rangle \geq 0$ oraz $\langle x, x \rangle > 0$ dla $x \neq 0$ (*dodatnio-określoność*)
- (2) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$, (*liniowość w pierwszym argumencie*)
- (3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, (*antysymetria*)

gdzie $x, y, z \in H$ i $\alpha, \beta \in \mathbb{F}$. Liczbę $\langle x, y \rangle$ nazywamy *iloczynem skalarnym* wektorów x, y .

Uwaga 2.2. W warunku (3), symbol $\overline{\langle y, z \rangle}$ oznacza sprzężenie zespolone liczby $\langle y, z \rangle$. W szczególności, w przypadku rzeczywistym, tzn. gdy $\mathbb{F} = \mathbb{R}$, to sprzężenie zespolone możemy pominąć i warunek (3) przyjmuje postać $\langle x, y \rangle = \langle y, z \rangle$. *Rzeczywisty iloczyn skalarny jest funkcją dwuliniową symetryczną!* Należy tu podkreślić, że w literaturze naukowej zazwyczaj przestrzenie Hilberta rozpatrywane są nad ciałem liczb zespolonych $\mathbb{C}$.¹ Natomiast w przypadku rzeczywistym mówi się o rzeczywistych przestrzeniach Hilberta lub o (nieskończenie wymiarowych) przestrzeniach euklidesowych.

Uwaga 2.3. Warunki (2) i (3), implikują

$$(2') \quad \langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle, \quad (\text{antyliniowość w drugim argumencie})$$

gdzie $x, y, z \in H$ i $\alpha, \beta \in \mathbb{F}$. W szczególności, iloczyn skalarny jest *formą półtoraliniową*, a w przypadku rzeczywistym *iloczyn skalarny jest formą dwuliniową!*

Przykład 2.4. Wzór $\langle x, y \rangle = \sum_{k=1}^n x(k) \overline{y(k)}$ definiuje iloczyn skalarny na przestrzeni skończenie wymiarowej $H = \mathbb{F}^n$.

¹Ponadto zazwyczaj zakłada się liniowość ze względu na drugi argument (ale to już inna bajka)

Przykład 2.5. Wzór $\langle x, y \rangle = \sum_{k=1}^{\infty} x(k) \overline{y(k)}$ definiuje iloczyn skalarny na przestrzeni ciągów sumowalnych w kwadracie $H = \ell^2$.

Przykład 2.6. Jeśli (Ω, Σ, μ) przestrzeń z miarą, to wzór $\langle x, y \rangle = \int_{\Omega} x(t) \overline{y(t)} d\mu(t)$ definiuje iloczyn skalarny na przestrzeni $L^2(\mu)$ funkcji całkowalnych w kwadracie względem miary μ .

Uwaga 2.7. Przykład 2.6 obejmuje Przykłady 2.4, 2.5 jako szczególne przypadki (wystarczy wziąć miarę liczącą odpowiednio na zbiorze skończonym i zbiorze nieskończonym przeliczalnym). Zauważmy, że iloczyn skalarny w Przykładzie 2.6 jest poprawnie określony, tzn. funkcja $x(t) \overline{y(t)}$ jest całkowalna. Wynika to z nierówności Höldera dla $p = q = 2$ (patrz wykład 3):

$$\int_{\Omega} |x(t)y(t)| d\mu(t) \leq \|x\|_2 \cdot \|y\|_2.$$

Powyższa nierówność zwana jest również *nierównością Schwartza*. Ponadto norma w przestrzeni $L^2(\mu)$ zadana jest przez iloczyn skalarny wzorem

$$\|x\|_2 = \left(\int_{\Omega} |x(t)|^2 d\mu(t) \right)^{\frac{1}{2}} \stackrel{|z|^2 = z\bar{z}}{=} \left(\int_{\Omega} x(t) \overline{x(t)} d\mu(t) \right)^{\frac{1}{2}} = \sqrt{\langle x, x \rangle}.$$

Okazuje się, że powyższa nierówność jak i zależność między normą i iloczynem skalarnym zachodzi ogólnie dla wszystkich (abstrakcyjnych) przestrzeni unitarnych:

Definicja 2.8. Normą elementu $x \in H$ w przestrzeni unitarnej H nazywamy liczbę

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Pierwiastek jest poprawnie określony, gdyż $\langle x, x \rangle \geq 0$ na mocy Definicji 2.1(1).

Następujący wzór można traktować jako *uogólnienie wzoru skróconego mnożenia*: $(a + b)^2 = a^2 + 2ab + b^2$, gdzie a, b są liczbami rzeczywistymi. Będziemy go używać wielokrotnie.

Lemat 2.9 (“Uogólniony wzór skróconego mnożenia”). Dla dowolnych wektorów x, y w przestrzeni unitarnej zachodzi $\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$.

Dowód. Zastosować półtoraliniowość iloczynu skalarnego oraz wzór $\operatorname{Re} z = \frac{z + \bar{z}}{2}$, $z \in \mathbb{C}$. 🏠 □

Twierdzenie 2.10 (Nierówność Schwartza). Dla dowolnych wektorów x, y przestrzeni unitarnej zachodzi nierówność

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Dowód. (Inny i chyba łatwiejszy do zapamiętania, dowód przedstawimy w Uwadze 2.30.) Weźmy $\lambda \in \mathbb{F}$ takie, że $|\lambda| = 1$ oraz $|\langle x, y \rangle| = \lambda \langle x, y \rangle$ (mianowicie $\lambda = e^{-i \arg \langle x, y \rangle}$, gdzie $\arg z$ oznacza argument liczby zespolonej z). Dla każdego $t \in \mathbb{R}$ mamy

$$0 \stackrel{2.1(1)}{\leq} \langle t\lambda x + y, t\lambda x + y \rangle \stackrel{2.9}{=} \|t\lambda x\|^2 + 2 \operatorname{Re} \langle t\lambda x, y \rangle + \|y\|^2 = t^2 \|x\|^2 + 2t |\langle x, y \rangle| + \|y\|^2.$$

Zatem funkcja kwadratowa $f(t) := t^2 \|x\|^2 + 2t |\langle x, y \rangle| + \|y\|^2 \geq 0$ jest nieujemna. Czyli $f(t)$ ma co najwyżej jedno miejsce zerowe, a stąd jej wyróżnik (czyli tzw. delta) jest niedodatni:

$$\Delta = 4 |\langle x, y \rangle|^2 - 4 \|x\|^2 \|y\|^2 \leq 0.$$

Po przekształceniach daje to $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$. □

Wniosek 2.11. Funkcja $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ jest normą na przestrzeni unitarnej.

Dowód. Z Definicji 2.1(1), $\|\cdot\|$ jest poprawnie określoną funkcją nieujemną taką, że $\|x\| = 0$ wtedy i tylko wtedy, gdy $x = 0$. Dla $\lambda \in \mathbb{F}$ mamy

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda \bar{\lambda} \langle x, x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \cdot \|x\|.$$

Natomiast korzystając z Lematu 2.9, nierówności $\operatorname{Re} z \leq |z|^2$, $z \in \mathbb{C}$, oraz nierówności Schwartza, dla $x, y \in H$ mamy

$$\begin{aligned} \|x + y\|^2 &\stackrel{2.9}{=} \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \stackrel{\operatorname{Re} z \leq |z|^2}{\leq} \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\stackrel{2.10}{\leq} \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Pierwiastkując tę nierówność otrzymujemy nierówność trójkąta. □

Definicja 2.12. *Przestrzeni Hilberta* nazywamy przestrzeń unitarną, która jest zupełna w normie $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ zadanej przez iloczyn skalarny.

Uwaga 2.13. Nierówność Schwartza implikuje, że iloczyn skalarny $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{F}$ jest funkcją ciągłą w normie, którą zadaje, tzn. jeśli $x_n \rightarrow x_0$ i $y_m \rightarrow y_0$, to $\langle x_n, y_m \rangle \rightarrow \langle x_0, y_0 \rangle$. Rzeczywiście,

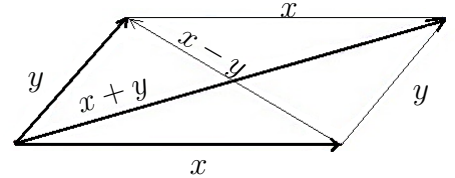
$$\begin{aligned} |\langle x_n, y_m \rangle - \langle x_0, y_0 \rangle| &= |\langle x_n, y_m \rangle - \langle x_0, y_m \rangle + \langle x_0, y_m \rangle - \langle x_0, y_0 \rangle| \\ &\leq |\langle x_n - x_0, y_m \rangle| + |\langle x_0, y_m - y_0 \rangle| \\ &\stackrel{2.10}{\leq} \|x_n - x_0\| \cdot \|y_m\| + \|x_0\| \cdot \|y_m - y_0\| \xrightarrow{n, m \rightarrow \infty} 0. \end{aligned}$$

Uwaga 2.14. Uzupełnienie (patrz Twierdzenie 1.11) przestrzeni unitarnej jest przestrzenią Hilberta, gdyż na mocy zasady ciągłego przedłużenia operatorów liniowych (Twierdzenie 1.55), iloczyn skalarny przedłuża się jednoznacznie do uzupełnienia. Ponadto każda domknięta podprzestrzeń liniowa przestrzeni Hilberta jest przestrzenią Hilberta, por. Stwierdzenie 1.10.

Zatem przestrzeń Hilberta to takie przestrzenie Banacha, w których norma zadana jest przez iloczyn skalarny. Jak rozróżnić mając daną przestrzeń Banacha, czy jest to przestrzeń Hilberta? Innym słowy, jak mając daną normę rozpoznać, czy odpowiedni iloczyn skalarny istnieje? Okazuje się, że eleganckim geometrycznym kryterium charakteryzującym przestrzeń Hilberta jest tożsamość równoległoboku:

Uwaga 2.15. Tożsamość równoległoboku mówi, że:

suma kwadratów długości boków w równoległoboku jest równa sumie kwadratów długości przekątnych, patrz rysunek obok:



Twierdzenie 2.16. W dowolnej przestrzeni unitarnej H zachodzi “tożsamość równoległoboku”

$$\forall_{x, y \in H} \quad \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (2.1)$$

oraz iloczyn skalarny jest jednoznacznie wyznaczony przez normę “wzorem polaryzacyjnym”

$$\langle x, y \rangle = \begin{cases} \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2), & \text{gdy } \mathbb{F} = \mathbb{R}, \\ \frac{1}{4} \sum_{k=0}^3 i^k \|x + i^k y\|^2, & \text{gdy } \mathbb{F} = \mathbb{C}. \end{cases} \quad (2.2)$$

Na odwrót, jeśli $(X, \|\cdot\|)$ jest przestrzenią unormowaną, w której spełniona jest tożsamość równoległoboku (2.1), to wzór polaryzacyjny (2.2) zadaje iloczyn skalarny taki, że $\|x\| = \sqrt{\langle x, x \rangle}$.

Dowód. Pierwszą część tezy dowodzi się poprzez proste rachunki z wykorzystaniem Lematu 2.9. Dlatego dowód tej części pozostawiamy jako pracę domową (🏠 policzyć). Dowód drugiej części twierdzenia jest znacznie trudniejszy i dlatego też pozostawiamy go jako pracę domową (🏠 dla chętnych) - najlepiej poszukać w literaturze. $\square$

Wniosek 2.17. *Przestrzeń Banacha $(X, \|\cdot\|)$ jest przestrzenią Hilberta wtedy i tylko wtedy, gdy spełniona jest tożsamość równoległoboku (2.1).*

Przykład 2.18. Przestrzeń $L^p(\mu)$, $p \geq 1$, dla przestrzeni z miarą (Ω, Σ, μ) , jest przestrzenią Hilberta wtedy i tylko wtedy, gdy $p = 2$ (zakładamy tu, że $L^p(\mu)$ jest co najmniej dwuwymiarowa). Rzeczywiście, to że $L^2(\mu)$ jest przestrzenią Hilberta z iloczynem skalarnym $\langle x, y \rangle = \int_{\Omega} x(t)y(t) d\mu(t)$ już wiemy, patrz Przykład 2.6. W drugą stronę założmy, że w $L^p(\mu)$ jest spełniona tożsamość równoległoboku. Weźmy dwa zbiory rozłączne $A, B \in \Sigma$ takie, że $0 < \mu(A), \mu(B) < \infty$.² Podstawiając $x := \mathbf{1}_A$ i $y = \mathbf{1}_B$ do obu stron tożsamości (2.1) dostajemy (🏠 policzyć)

$$\begin{aligned} \text{Lewa strona} &= \|x + y\|_p^2 + \|x - y\|_p^2 = \mu(A \cup B)^{\frac{2}{p}} + \mu(A \cup B)^{\frac{2}{p}} = 2\mu(A \cup B)^{\frac{2}{p}}, \\ \text{Prawa strona} &= 2(\|x\|_p^2 + \|y\|_p^2) = 2\left(\mu(A)^{\frac{2}{p}} + \mu(B)^{\frac{2}{p}}\right). \end{aligned}$$

Dzieląc prawą stronę przez lewą otrzymujemy, $1 = \left(\frac{\mu(A)}{\mu(A \cup B)}\right)^{\frac{2}{p}} + \left(\frac{\mu(B)}{\mu(A \cup B)}\right)^{\frac{2}{p}}$. Czyli kładąc $\lambda_1 := \frac{\mu(A)}{\mu(A \cup B)}$ oraz $\lambda_2 := \frac{\mu(B)}{\mu(A \cup B)}$ dostajemy dwie liczby takie, że

$$0 < \lambda_1, \lambda_2 < 1, \quad \lambda_1 + \lambda_2 = 1 = (\lambda_1)^{\frac{2}{p}} + (\lambda_2)^{\frac{2}{p}}.$$

Powyższe relacje mogą zachodzić tylko wtedy, gdy $p = 2$, gdyż jeśli $p > 2$ to $\lambda_i > (\lambda_i)^{\frac{2}{p}}$, a jeśli $p < 2$ to $\lambda_i < (\lambda_i)^{\frac{2}{p}}$ dla $i = 1, 2$.

Iloczyn skalarny pozwala zdefiniować cosinus kąta α pomiędzy dwoma wektorami $x, y \in H$, stosując znany ze szkoły podstawowej wzór:

$$\cos \alpha := \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

W przypadku rzeczywistym, biorąc arcus tangens możemy za pomocą tego wzoru zdefiniować sam kąt α . Zauważmy, że wektory x, y są do siebie prostopadłe, gdy $\cos \alpha = \cos 90^\circ = 0$, czyli gdy $\langle x, y \rangle = 0$. Uwaga to prowadzi do następującej definicji pojęcia ortogonalności w ogólnych (abstrakcyjnych) przestrzeniach unitarnych:

Definicja 2.19. Powiemy, że wektory $x, y \in H$ w przestrzeni unitarnej H są *ortogonalne* (lub też *prostopadłe*), co zapisujemy $x \perp y$, jeżeli ich iloczyn skalarny wynosi zero. To znaczy

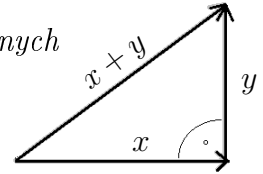
$$x \perp y \iff \langle x, y \rangle = 0.$$

Uwaga 2.20. Z antysymetrii i dodatniej-określoności iloczynu skalarnego, patrz Definicja 2.1, wynika, że relacja ortogonalności jest symetryczna, tzn. $x \perp y \Leftrightarrow y \perp x$, oraz jedynym wektorem prostopadłym do siebie jest $x = 0$.

²Takie zbiory istnieją, ponieważ założyliśmy, że $\dim(L^p(\mu)) \geq 2$

Pojęcie iloczynu skalarnego pozwala na najkrótszy dowód Twierdzenia Pitagorasa jaki do tej pory widzieliście.³ Mianowicie, zauważmy, że twierdzenie:

“W trójkącie prostokątnym suma kwadratów długości przyprostokątnych jest równa kwadratowi długości przeciwprostokątnej”,



w języku wektorów przyjmuje następującą postać:

Lemat 2.21 (Twierdzenie Pitagorasa). *Jeśli $x \perp y$, to $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

Dowód. $\|x + y\|^2 \stackrel{2.9}{=} \|x\|^2 + 2\operatorname{Re}\langle x, y \rangle + \|y\|^2 \stackrel{x \perp y}{=} \|x\|^2 + \|y\|^2.$ □

Uwaga 2.22. Z dowodu widać, że w przypadku rzeczywistym powyższą implikację można odwrócić, tzn. jeśli H to rzeczywista przestrzeń unitarna, to $x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Będziemy też rozważać wektory prostopadłe do podzbiorów:

Definicja 2.23. *Dopełnieniem ortogonalnym podzbioru $M \subseteq H$ przestrzeni unitarnej H nazywamy zbiór*

$$M^\perp := \{x \in H : x \perp y \text{ dla każdego } y \in M\}.$$

Jeśli $x \in M^\perp$ to piszemy też $x \perp M$ i mówimy, że x jest *ortogonalny do zbioru M* . Dla dwóch zbiorów $N, M \subseteq H$ piszemy $N \perp M$ jeżeli $n \perp m$ dla każdego $n \in N$ i $m \in M$.

Stwierdzenie 2.24. *Dopełnienie ortogonalne $M^\perp$ jest domkniętą przestrzenią liniową.*

Dowód. Można to stwierdzenie dowieść wprost korzystając z definicji. My przedstawimy tu dowód, który uwidoczni rolę funkcjonałów liniowych. Przypomnijmy, że iloczyn skalarny jest liniowy ze względu na pierwszą współrzędną. Zatem dla każdego $y \in H$ wzór

$$f_y(x) := \langle x, y \rangle, \quad x \in H,$$

definiuje funkcjonał liniowy $f_y : H \rightarrow \mathbb{F} = \mathbb{R}, \mathbb{C}$. Z nierówności Schwartza wynika, że

$$|f_y(x)| = |\langle x, y \rangle| \leq \|x\| \cdot \|y\|,$$

tzn. funkcjonał f_y jest ograniczony oraz $\|f_y\| \leq \|y\|$ (w rzeczywistości, $\|f_y\| = \|y\|$, bo jeśli $y \neq 0$, to kładąc $x = \frac{y}{\|y\|}$ mamy $\|x\| = 1$ oraz $|f_y(x)| = |\langle \frac{y}{\|y\|}, y \rangle| = \frac{\|y\|^2}{\|y\|} = \|y\|$). Zatem f_y jest ciągły i jego jądro $\ker f_y = f_y^{-1}(0) \subseteq H$ jest domkniętą podprzestrzenią liniową. Teraz wystarczy zauważyć, że

$$M^\perp := \{x \in H : \langle x, y \rangle = 0 \text{ dla każdego } y \in M\} = \bigcap_{y \in M} \ker f_y.$$

Skąd $M^\perp$ jest domkniętą podprzestrzenią liniową (bo przekrój przestrzeni liniowych jest przestrzenią liniową, a przekrój zbiorów domkniętych jest zbiorem domkniętym). □

Wprost z definicji wynika, że $M \subseteq M^{\perp\perp} := (M^\perp)^\perp$ oraz na mocy Stwierdzenia 2.24, żeby w tej inkluzji zachodziła równość M musi być domkniętą podprzestrzenią liniową. W rzeczywistości oczekivalibyśmy, żeby dla domkniętej podprzestrzeni $M \subseteq H$ zachodziły relacje

$$M = M^{\perp\perp} \quad \text{oraz} \quad H = M \oplus M^\perp.$$

Relacje te rzeczywiście mają miejsce w każdej przestrzeni Hilberta (wykażemy to w kolejnym podrozdziale). Zupełność przestrzeni pełni tu kluczową rolę i w przestrzeniach unitarnych niezupełnych dopełnienie ortogonalne na ogół nie ma dobrych własności:

³W książce E. S. Loomis, The Pythagorean Proposition, NCTM, 1968, spisanych jest 367 różnych dowodów tego twierdzenia

Przykład 2.25 (Zdegenerowane dopełnienie ortogonalne w niezupełnej przestrzeni unitarnej). Niech $H = c_{00} = \{x = (x(1), x(2), \dots, x(N), 0, 0, \dots) : N \in \mathbb{N}, x(k) \in \mathbb{F}\}$ będzie przestrzenią ciągów ze skończoną ilością wyrazów niezerowych. Jest to przestrzeń unitarna wraz z iloczynem skalarnym $\langle x, y \rangle = \sum_{k=1}^{\infty} x(k)\overline{y(k)}$, dla $x, y \in H$. Wtedy H jest gęstą podprzestrzenią liniową w przestrzeni Hilberta ℓ^2 , tzn. $\overline{H} = \ell^2$, por. Przykład 2.5. W szczególności, H *nie jest przestrzenią zupełną!* Rozważmy domkniętą podprzestrzeń liniową H daną wzorem

$$M := \left\{ x \in H : \sum_{k=1}^{\infty} x(k) = 0 \right\}$$

(domkniętość wynika z faktu, że zbieżność w H pociąga zbieżność po współrzędnych, a suma $\sum_{k=1}^{\infty} x(k)$ jest de facto skończona, bo $x \in H = c_{00}$). Oczywiście $M \neq H$. Mimo, to twierdzimy, że

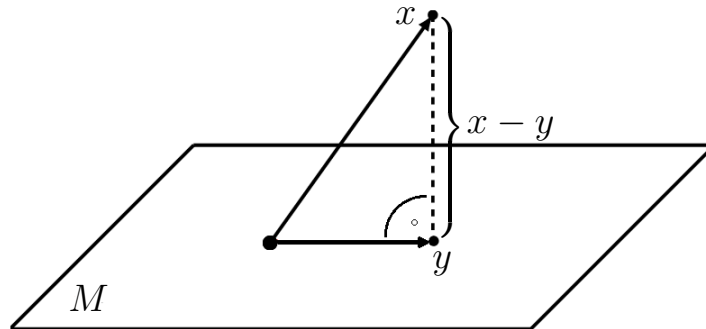
$$M^{\perp} = \{0\}.$$

Rzeczywiście, niech $y \in M^{\perp}$. Połóżmy $x_n := (\underbrace{0, \dots, 0}_{n-1}, 1, -1, 0, 0, \dots)$ dla $n > 0$. Wtedy $x_n \in M$ oraz $\langle x_n, y \rangle = 0$ oznacza, że $y(n) = y(n+1)$, dla każdego $n > 0$. Stąd $y = (y(1), y(1), y(1), \dots)$ jest ciągiem stałym, który od pewnego miejsca się zeruje, bo $y \in H$. Zatem $y = 0$. W szczególności

$$M + M^{\perp} \neq H, \quad M \neq M^{\perp\perp} = \{0\}^{\perp} = H.$$

2.2 Rzut ortogonalny

Niech H będzie przestrzenią Hilberta. Jedną z kluczowych operacji, którą można wykonywać w przestrzeniach Hilberta jest rzut ortogonalny. Intuicyjnie rzecz biorąc (patrz rysunek poniżej), rzut ortogonalny wektora $x \in H$ na podprzestrzeń $M \subseteq H$ to jest taki wektor y z przestrzeni M , że wektor $x - y$ jest ortogonalny do przestrzeni M .



Definicja 2.26. *Rzutem ortogonalnym* wektora $x \in H$ na podprzestrzeń $M \subseteq H$ nazywamy wektor $y \in M$ taki, że $x - y \perp M$. Piszemy wtedy $P_M x = y$. Czyli

$$P_M x = y \quad \stackrel{\text{def}}{\iff} \quad \forall_{z \in M} \langle x - y, z \rangle = 0.$$

Uwaga 2.27. Rzut ortogonalny $P_M x$ jest wyznaczony jednoznacznie! Rzeczywiście, jeśli $y, y' \in M$ takie, że $\langle x - y, z \rangle = 0$ oraz $\langle x - y', z \rangle = 0$ dla każdego $z \in M$, to biorąc za $z = y - y' \in M$ oraz odejmując stronami poprzednie równości otrzymujemy $\langle y - y', y - y' \rangle = 0$, skąd $y = y'$.

W przestrzeniach Hilberta rzut ortogonalny na domkniętą podprzestrzeń zawsze istnieje. Ale dowód jest nietrywialny i zajmie nam kolejnych kilka stron. Co więcej w przestrzeniach unitarnych niezupełnych (tzn. gdy H nie jest przestrzenią Hilberta) rzut ortogonalny może nie istnieć:

Przykład 2.28 (Nieistnienie rzutu ortogonalnego w przestrzeniach niezupełnych). Rozważmy przestrzeń unitarną $H = c_{00}$ i jej domkniętą podprzestrzeń M z Przykładu 2.25. Dla wektora $x := (1, 1, 0, 0, \dots) \in H$ rzut na M nie istnieje. Rzeczywiście, gdyby istniał $y \in M$ taki, że $x - y \perp M$, czyli $x - y \in M^\perp$, to stąd że $M^\perp = \{0\}$ otrzymalibyśmy, że $x = y$, co prowadzi do sprzeczności, bo $x \notin M$ i $y \in M$. W tym przykładzie kluczową rolę odgrywa de facto nie brak zupełności przestrzeni H , lecz brak zupełności podprzestrzeni M , patrz Uwaga 2.35 poniżej.

Rzut ortogonalny na podprzestrzeń jednowymiarową istnieje zawsze, nawet w niezupełnych przestrzeniach unitarnych. Co więcej, możemy na taki rzut wypisać konkretny wzór:

Przykład 2.29 (Rzut ortogonalny na podprzestrzeń jednowymiarową). Niech $M = \{\lambda y : \lambda \in \mathbb{F}\}$ będzie przestrzenią rozpiętą przez niezerowy wektor $y \in H$. Niech $x \in H$. Aby wyznaczyć $P_M x$ trzeba znaleźć skalar $\lambda_0 \in \mathbb{F}$ taki, że $x - \lambda_0 y \perp M$, ale skoro M jest rozpięta przez y , to

$$\forall_{z \in M} \langle x - \lambda_0 y, z \rangle = 0 \xLeftrightarrow{z=y} \langle x - \lambda_0 y, y \rangle = 0 \iff \langle x, y \rangle - \lambda_0 \langle y, y \rangle = 0 \iff \lambda_0 = \frac{\langle x, y \rangle}{\langle y, y \rangle}.$$

Zatem

$$P_M x = \lambda_0 y = \frac{\langle x, y \rangle}{\langle y, y \rangle} y = \frac{\langle x, y \rangle}{\|y\|^2} y.$$

Rzut ten nazywa się też *rzutem ortogonalnym* wektora x na wektor y i oznacza przez $P_y x$. Czyli

$$P_y x := \frac{\langle x, y \rangle}{\|y\|^2} y.$$

Uwaga 2.30 (Alternatywny dowód nierówności Schwartza). Odległość x od jego rzutu $P_y x$, czyli wartość $\|x - P_y x\|$, można, a nawet należy traktować jako odległość wektora x od przestrzeni rozpiętej przez y . Obliczając kwadrat tej odległości otrzymujemy dowód nierówności Schwartza. Rzeczywiście, założmy, że $y \neq 0$. Wtedy $P_y x = \frac{\langle x, y \rangle}{\|y\|^2} y$ i stąd

$$\begin{aligned} \|x - P_y x\|^2 &\stackrel{2.9}{=} \|x\|^2 - 2 \operatorname{Re} \left\langle x, \frac{\langle x, y \rangle}{\|y\|^2} y \right\rangle + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \|y\|^2 = \|x\|^2 - 2 \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}. \end{aligned}$$

Czyli $0 \leq \|x - P_y x\|^2 = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}$. Nierówność ta jest równoważna nierówności Schwartza $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$. Co więcej, widzimy stąd, że równość $|\langle x, y \rangle| = \|x\| \cdot \|y\|$ zachodzi wtedy i tylko wtedy, gdy $x = \frac{\langle x, y \rangle}{\|y\|^2} y$ lub $y = 0$. To z kolei jest równoważne temu, że x i y są liniowo zależne (bo jeśli $x = \lambda y$ i $y \neq 0$, to $\frac{\langle x, y \rangle}{\|y\|^2} = \lambda$).

Wniosek 2.31. Równość $|\langle x, y \rangle| = \|x\| \cdot \|y\|$ w nierówności Schwartza zachodzi wtedy i tylko wtedy, gdy wektory x i y są liniowo zależne.

Rysunek na stronie 31 sugeruje, że jeżeli $y = P_M x$ jest rzutem ortogonalnym wektora $x \in H$ na podprzestrzeń $M \subseteq H$, to długość $\|y - x\|$ powinna być odległością wektora x od przestrzeni M . To znaczy $y = P_M x \in M$ powinien być wektorem, w którym realizuje się poniższe infimum

$$\operatorname{dist}(x, M) = \inf_{z \in M} \|x - z\|.$$

Pokażemy, że w istocie tak jest (o ile M jest domkniętą podprzestrzenią). Interpretacja ta jest kluczowa dla dowodu istnienia rzutu ortogonalnego w przestrzeni Hilberta! Mianowicie, pokażemy,

że dla każdego domkniętego podzbioru wypukłego $M \subseteq H$ przestrzeni Hilberta H oraz wektora $x \in H$ istnieje dokładnie jeden $y \in M$ taki, że

$$\|x - y\| = \inf_{z \in M} \|x - z\|.$$

Przypomnijmy, że zbiór $M \subseteq H$ jest *wypukły*, jeżeli $\forall x, y \in M \forall \lambda \in [0, 1] \lambda x + (1 - \lambda)y \in M$, tzn. gdy każdy odcinek o końcach w M cały zawiera się w M , patrz rysunek:



Oczywiście każda podprzestrzeń liniowa jest zbiorem wypukłym.

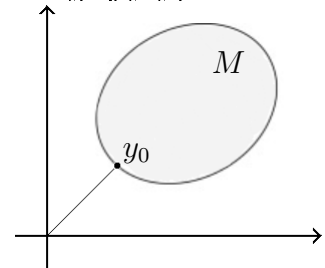
Twierdzenie 2.32 (O odległości od zbioru wypukłego). *Niech H przestrzeń Hilberta. Dla dowolnego domkniętego zbioru wypukłego $M \subseteq H$ oraz punktu $x \in H$ istnieje dokładnie jeden $y \in M$ taki, że $\|x - y\| = \text{dist}(x, M)$.*

Dowód. Zauważmy najpierw, że “przesuwając” zbiór M możemy założyć, że $x = 0$ (nie jest to konieczne, lecz ułatwia rachunki). Rzeczywiście, jeśli M jest domknięty i wypukły, to zbiór $\widetilde{M} = x - M := \{x - z : z \in M\}$ też ma te własności, bo przesunięcie o wektor oraz odbicie względem zera są homeomorfizmami zachowującymi kombinacje wypukłe⁴. Ponadto, $y \in M$ oraz $\|x - y\| = \text{dist}(x, M)$ wtedy i tylko wtedy, gdy $x - y \in \widetilde{M}$ oraz $\|x - y\| = \text{dist}(0, \widetilde{M}) = \inf_{\tilde{y} \in \widetilde{M}} \|\tilde{y}\|$.

Założmy więc, że $x = 0$. Wtedy dowodzone twierdzenie przyjmuje następującą postać:

$$M \subseteq H \text{ domknięty i wypukły} \implies \exists!_{y_0 \in M} \|y_0\| = \inf_{y \in M} \|y\|.$$

Innymi słowy, potrzebujemy pokazać, że w domkniętym i wypukłym zbiorze M istnieje dokładnie jeden element o minimalnej normie.



“Istnienie”. Połóżmy $d := \inf_{y \in M} \|y\|$ i niech $\{y_n\}_{n=1}^\infty \subseteq M$ taki, że $\|y_n\| \rightarrow d$. Wystarczy pokazać, że $\{y_n\}_{n=1}^\infty$ jest ciągiem Cauchy, bo wtedy na mocy zupełności H i domkniętości M , ciąg ten jest zbieżny do pewnego $y_0 \in M$, a z ciągłości normy $\|y_0\| = \lim_{n \rightarrow \infty} \|y_n\| = d$. Warunek Cauchy wynika z tożsamości równoległoboku i wypukłości M :

$$\begin{aligned} \|y_n - y_m\|^2 &\stackrel{(2.1)}{=} 2\|y_n\|^2 + 2\|y_m\|^2 - \|y_n + y_m\|^2 = 2\|y_n\|^2 + 2\|y_m\|^2 - 4\left\|\frac{y_n + y_m}{2}\right\|^2 \\ &\stackrel{\frac{y_n + y_m}{2} \in M}{\leq} 2\|y_n\|^2 + 2\|y_m\|^2 - 4d^2 \xrightarrow{n, m \rightarrow \infty} 2d^2 + 2d^2 - 4d^2 = 0. \end{aligned}$$

“Jednoznaczność”. Jeżeli $y_1, y_2 \in M$ takie, że $\|y_1\| = \|y_2\| = d$, to rachunek niemal identyczny do tego powyżej pokazuje, że $y_1 = y_2$:

$$\begin{aligned} \|y_1 - y_2\|^2 &\stackrel{(2.1)}{=} 2\|y_1\|^2 + 2\|y_2\|^2 - \|y_1 + y_2\|^2 = 2\|y_1\|^2 + 2\|y_2\|^2 - 4\left\|\frac{y_1 + y_2}{2}\right\|^2 \\ &\stackrel{\frac{y_1 + y_2}{2} \in M}{\leq} 2\|y_1\|^2 + 2\|y_2\|^2 - 4d^2 = 2d^2 + 2d^2 - 4d^2 = 0. \end{aligned}$$

□

⁴Odwzorowania zachowujące kombinacje wypukłe nazywamy odwzorowaniami afinicznymi

Uwaga 2.33. W zbiorze niewypukłym M mogą istnieć dwa różne punkty $y_1, y_2 \in H$ takie, że $\|x - y_1\| = \|x - y_2\| = \text{dist}(x, M)$. Takich punktów może być nawet nieskończenie wiele. Weźmy np. $H = \mathbb{C}$, $x = 0$ i $M = \{z \in \mathbb{C} : |z| \geq 1\}$ dopełnienie koła jednostkowego. Wtedy $\|x - y\| = \text{dist}(x, M)$ dla każdego y leżącego na okręgu $S^1 = \{z \in \mathbb{C} : |z| = 1\}$.

Twierdzenie 2.34 (O istnieniu rzutu ortogonalnego). *Dla dowolnej domkniętej podprzestrzeni $M \subseteq H$ przestrzeni Hilberta H oraz punktu $x \in H$ istnieje rzut ortogonalny $y = P_M x$. Co więcej,*

$$\|x - y\| = \text{dist}(x, M) \quad (2.3)$$

i rzut ortogonalny y wektora x na M jest przez tę równość wyznaczony jednoznacznie.

Dowód. Skoro M jest zbiorem wypukłym, to na mocy Twierdzenia 2.32 istnieje $y \in M$ spełniający (2.3). Musimy pokazać, że $x - y \perp M$. Niech $z \in M$. Dla każdego $t \in \mathbb{F}$, $y + tz \in M$, bo M przestrzeń liniowa. Zatem

$$\|x - y\|^2 \stackrel{(2.3)}{\leq} \|x - (y + tz)\|^2 = \|(x - y) + tz\|^2 \stackrel{2.9}{=} \|x - y\|^2 - 2 \operatorname{Re} \langle x - y, tz \rangle + |t|^2 \|z\|^2.$$

Stąd, po skróceniu przez $\|x - y\|^2$, dostajemy, że $0 \leq |t|^2 \|z\|^2 - 2 \operatorname{Re} \langle x - y, tz \rangle$ dla każdego $t \in \mathbb{F}$. Kładąc $t = se^{i\varphi}$, gdzie $s \in \mathbb{R}$ i $\varphi := \arg \langle x - y, z \rangle$ (jeśli $\mathbb{F} = \mathbb{R}$, to $e^{i\varphi} = \operatorname{sign} \langle x - y, z \rangle \in \{-1, 0, 1\}$ jest liczbą rzeczywistą) nierówność ta przyjmuje postać

$$0 \leq s^2 |e^{i\varphi}|^2 \|z\|^2 - 2s \operatorname{Re}(e^{-i\varphi} \langle x - y, z \rangle) = s^2 \|z\|^2 - 2s |\langle x - y, z \rangle|.$$

Czyli rzeczywista funkcja kwadratowa $f(s) = s^2 \|z\|^2 - 2s |\langle x - y, z \rangle|$ jest nieujemna i posiada miejsce zerowe dla $s = 0$. Stąd jej wyróżnik $\Delta = 4 |\langle x - y, z \rangle|^2$ musi być równy zero, czyli $\langle x - y, z \rangle = 0$, tzn. $x - y \perp z$. $\square$

Uwaga 2.35. Z powyższego Twierdzenia wynika twierdzenie pozornie ogólniejsze mówiące, że „rzut ortogonalny $P_M x$ istnieje dla dowolnej zupełnej podprzestrzeni $M \subseteq H$ przestrzeni unitarnej H oraz punktu $x \in H$ ”. Rzeczywiście, jeśli M jest zupełna, to jest domkniętą podprzestrzenią uzupełnienia przestrzeni H do przestrzeni Hilberta $\overline{H}$, por. Uwaga 2.14. Zatem Twierdzenie 2.34 zastosowane do inkluzji $M \subseteq \overline{H}$ implikuje tę ogólniejszą tezę.

Wniosek 2.36. *Dla dowolnej domkniętej podprzestrzeni M przestrzeni Hilberta H mamy*

$$H = M \oplus M^\perp,$$

czyli $\forall x \in H \exists! y \in M \exists! z \in M^\perp x = y + z$.

Dowód. Niech $x \in H$. Połóżmy $y := P_M x$ i $z := x - y$. Wtedy $x = y + z$ oraz z definicji rzutu ortogonalnego $x - y \perp M$, tzn. $z \in M^\perp$. Aby pokazać jednoznaczność tego rozkładu załóżmy, że $x = y' + z'$ dla pewnych $y' \in M$ i $z' \in M^\perp$. Wtedy $y - y' \in M$, $z - z' \in M^\perp$ oraz $y - y' = z - z'$, ale $M \cap M^\perp = \{0\}$ (bo jedynym wektorem ortogonalnym do samego siebie jest wektor zerowy). Zatem $y = y'$ oraz $z = z'$. $\square$

Uwaga 2.37. Tezę powyższego wniosku można też zapisać jako $1 = P_M + P_{M^\perp}$, gdzie 1 jest operatorem identycznościowym na H . W szczególności, jeśli P_M jest rzutem ortogonalnym na domkniętą podprzestrzeń M , to $1 - P_M$ jest rzutem na jej dopełnienie ortogonalne $M^\perp$:

$$P_{M^\perp} = 1 - P_M.$$

Wniosek 2.38. *Jeśli M domknięta podprzestrzeń przestrzeni Hilberta, to $(M^\perp)^\perp = M$.*

Dowód. Inkluzja $M \subseteq (M^\perp)^\perp$ jest jasna, bo $x \in M \implies x \perp M^\perp \implies x \in (M^\perp)^\perp$. Żeby wykazać, inkluzję odwrotną weźmy element $x \in (M^\perp)^\perp$. Zapiszmy go w postaci $x = y + z$ gdzie $y \in M$, $z \in M^\perp$, jak we Wniosku 2.36. Chcemy pokazać, że $z = 0$, bo wtedy $x = y \in M$. W tym celu zauważmy, że zarówno $x \in (M^\perp)^\perp$ jak i $y \in M$ są prostopadłe do $z \in M^\perp$, a stąd $0 = \langle x, z \rangle = \langle y + z, z \rangle = \langle y, z \rangle + \langle z, z \rangle = \langle z, z \rangle = \|z\|^2$. Czyli $z = 0$. $\square$

Na koniec omówimy operację brania rzutu ortogonalnego jako operator liniowy.

Lemat 2.39. *Rzut ortogonalny $H \ni x \rightarrow P_M x \in M \subseteq H$ na domkniętą podprzestrzeń przestrzeni Hilberta jest ograniczonym operatorem liniowym. Ponadto $\|P_M\| = 1$, chyba, że $P_M \equiv 0$ (równoważnie $M = \{0\}$) i wtedy $\|P_M\| = 0$.*

Dowód. Niech $x, y \in H$ oraz $\alpha, \beta \in \mathbb{F}$. Chcemy wykazać, że $P_M(\alpha x + \beta y) = \alpha P_M x + \beta P_M y$. Z definicji rzutu ortogonalnego $P_M(\alpha x + \beta y)$ to jedyny element w M taki, że $(\alpha x + \beta y) - P_M(\alpha x + \beta y) \perp M$. Wystarczy zatem pokazać, że wektor $\alpha P_M x + \beta P_M y$ ma te same własności. Jasne, że $\alpha P_M x + \beta P_M y \in M$ bo M przestrzeń liniowa. Korzystając z (półtora)liniowości iloczynu skalarnego, dla $z \in M$ mamy

$$\langle (\alpha x + \beta y) - (\alpha P_M x + \beta P_M y), z \rangle = \alpha \langle x - P_M x, z \rangle + \beta \langle y - P_M y, z \rangle = 0,$$

bo $x - P_M x$ oraz $y - P_M y$ są ortogonalne do M z definicji rzutu. Zatem $(\alpha x + \beta y) - (\alpha P_M x + \beta P_M y) \perp M$ i stąd $P_M(\alpha x + \beta y) = \alpha P_M x + \beta P_M y$.

Aby wykazać ograniczoność operatora P_M weźmy dowolny $x \in H$. Korzystając z twierdzenia Pitagorasa oraz równości $1 = P_M + P_{M^\perp}$, patrz Uwaga 2.37, otrzymujemy

$$\|P_M x\|^2 \leq \|P_M x\|^2 + \|P_{M^\perp} x\|^2 \stackrel{2.21}{=} \|P_M x + P_{M^\perp} x\|^2 = \|x\|^2.$$

Stąd $\|P_M\| \leq 1$. Jeżeli $P_M \neq 0$ to $M \neq \{0\}$ i istnieje $x \in M$ o normie 1. Skoro $P_M x = x$ to $\|P_M x\| = \|x\| = 1$, czyli $\|P_M\| \geq 1$. Zatem $\|P_M\| = 1$. $\square$

Operator rzutu ortogonalnego można scharakteryzować na wiele różnych sposobów. Jedną z podstawowych własności operatora P_M jest to, że jego obraz jest domkniętą podprzestrzenią M oraz jego obcięcie do tej podprzestrzeni jest identycznością, tzn. $P_M|_M = id_M$. Obie te własności można w języku operatorów zapisać krótko poprzez równość $P_M^2 = P_M$, gdzie $P_M^2 = P_M \circ P_M$ jest składaniem odwzorowań liniowych. Własność tę nazywamy *idempotentnością* (z łac. idem, „taki sam, równy” i potens, „mający moc, siłę”). *Operator idempotentny* to taki, którego działanie nie zmienia wyniku przy kolejnych powtórzeniach.

Definicja 2.40. Operator liniowy $P : H \rightarrow H$ na przestrzeni liniowej H nazywamy *idempotentem* (lub *rzutem*), gdy $P^2 = P$.

Każdy rzut ortogonalny jest idempotentem (rzutem), ale nie na odwrót. Jeśli $P^2 = P$ jest idempotentem, to P na swoim obrazie $\text{Im } P := PH$ jest identycznością. Zatem jeśli tylko $P \neq 0$, to $\|P\| \geq 1$. Ponadto idempotent P zadaje rozkład przestrzeni na sumę prostą

$$H = \text{Im } P \oplus \ker P,$$

gdzie $\ker P = \{x \in H : Px = 0\}$ to jądro operatora P . Rozkład ten wynika z równości $1 = P + (1 - P)$ oraz $\ker P = \text{Im}(1 - P)$ (🏠 sprawdzić szczegóły). Rozkład ten wyznacza rzut P jednoznacznie, gdyż jeśli $x = y + z$, gdzie $y \in \text{Im } P$ i $z \in \ker P$, to $Px = y$. W szczególności, mówi się wtedy, że P jest *rzutem na podprzestrzeń $\text{Im } P$ względem podprzestrzeni $\ker P$* . Rzut P jest rzutem ortogonalnym wtedy i tylko wtedy, gdy $\ker P \perp \text{Im } P$.

Twierdzenie 2.41 (Charakteryzacje operatora rzutu ortogonalnego). *Niech $P : H \rightarrow H$ liniowy idempotent, tzn. $P^2 = P$, na przestrzeni Hilberta H . Następujące warunki są równoważne:*

- (1) P jest rzutem ortogonalnym,
- (2) $\ker P \perp \operatorname{Im} P$,
- (3) $P^* = P$ jest samosprężony, tzn. $\forall_{x,y \in H} \langle Px, y \rangle = \langle x, Py \rangle$,
- (4) $\|P\| \leq 1$ (dokładniej $\|P\| = 1$ lub $P = 0$).

Jeśli powyższe równoważne warunki zachodzą, to $P = P_M$ jest rzutem ortogonalnym na swój obraz $M = \operatorname{Im} P = PH$ oraz $\ker P = M^\perp$.

Dowód. (1) $\Rightarrow$ (2). Ta implikacja wynika z powyższej dyskusji, patrz też Uwaga 2.37.

(2) $\Rightarrow$ (3). Przypomnijmy, iż $\ker P = \operatorname{Im}(1 - P)$. Zatem z założenia $PH \perp (1 - P)H$. Stąd dla $x, y \in H$ mamy

$$\begin{aligned} \langle Px, y \rangle &\stackrel{y=Py+(1-P)y}{=} \langle Px, Py \rangle + \langle Px, (1-P)y \rangle \stackrel{PH \perp (1-P)H}{=} \langle Px, Py \rangle \\ &\stackrel{PH \perp (1-P)H}{=} \langle Px, Py \rangle + \langle (1-P)x, Py \rangle \stackrel{x=Px+(1-P)x}{=} \langle x, Py \rangle. \end{aligned}$$

(3) $\Rightarrow$ (4). Dla dowolnego $x \in H$ mamy

$$\|Px\|^2 = \langle Px, Px \rangle \stackrel{(3)}{=} \langle P(Px), x \rangle \stackrel{P^2=P}{=} \langle Px, x \rangle \stackrel{N.Schwartza}{\leq} \|Px\| \cdot \|x\|.$$

Po skróceniu przez $\|Px\|$ dostajemy stąd, że $\|Px\| \leq \|x\|$. Zatem $\|P\| \leq 1$. Nierówność $\|P\| \geq 1$ zachodzi dla każdego niezerowego idempotentu. Czyli $\|P\| = 1$ lub $P = 0$.

(4) $\Rightarrow$ (1). Pokażemy, że P jest rzutem ortogonalnym na podprzestrzeń $M := \operatorname{Im} P = PH$.⁵ Z definicji trzeba pokazać, że $x - Px \perp M$ dla każdego $x \in H$. Jest to równoważne inkluzji $\ker P \subseteq M^\perp$, bo $\ker P = (1 - P)H$. Weźmy zatem dowolne $x \in \ker P$ oraz $y \in M = \operatorname{Im} P$. Wtedy $Px = 0$ oraz $Py = y$. Stąd dla każdego $t \in \mathbb{F}$ dostajemy

$$\|y\|^2 = \|Py + tPx\|^2 = \|P(y + tx)\|^2 \stackrel{\|P\| \leq 1}{\leq} \|y + tx\|^2 \stackrel{2.9}{=} \|y\|^2 + 2 \operatorname{Re} t \langle x, y \rangle + |t|^2 \|x\|^2.$$

Podobnie jak w dowodzie Twierdzenia 2.34 implikuje to, że $\langle x, y \rangle = 0$. Rzeczywiście, kładąc $t = se^{-i\varphi}$, gdzie $\varphi := \arg \langle x, y \rangle$ i $s \in \mathbb{R}$, powyższa nierówność oznacza, że rzeczywista funkcja kwadratowa $f(s) = 2s|\langle x, y \rangle| + s^2\|x\|^2 \geq 0$ jest nieujemna. Stąd $\langle x, y \rangle = 0$, co należało dowieść. $\square$

Przykład 2.42 (Rzuty na płaszczyźnie). Niech $H = \mathbb{R}^2$ będzie płaszczyzną euklidesową. Wtedy wszystkie rzuty na oś x -sów $M = \{(x, 0) : x \in \mathbb{R}\}$ są dane przez macierze postaci

$$P_a = \begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Dla $a = 0$, P_0 jest rzutem na prostą $y = 0$ względem prostej $x = 0$ (tzn. rzutem ortogonalnym na M). Jeśli $a \neq 0$, to P_a jest rzutem na prostą $y = 0$ względem prostej $y = -x/a$ (tzn. rzutem na M względem podprzestrzeni $N = \{(x, -x/a) : x \in \mathbb{R}\}$). Żeby policzyć normę operatora P_a trzeba

zmaksymalizować funkcję $f(x, y) = x + ay$ pod warunkiem $x^2 + y^2 = 1$.

⁵ M jest domknięta, bo jeśli $\{y_n\}_{n=1}^\infty \subseteq M$ i $y_n \rightarrow y_0$, to $y_0 = \lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} Py_n = Py_0 \in M$

Stosując metodę mnożników Lagrange'a ( przeliczyć to) otrzymujemy, że funkcja f osiąga maksimum w punkcie $(x, y) = (\frac{1}{\sqrt{1+a^2}}, \frac{a}{\sqrt{1+a^2}})$ i jej maksimum wynosi $\sqrt{1+a^2}$. Stąd

$$\|P_a\| = \sqrt{1+a^2}, \quad a \in \mathbb{R}.$$

W szczególności $\|P_a\| \rightarrow \infty$, gdy $a \rightarrow \infty$. Czyli norma rzutu (idempotenta) może być dowolnie duża. Oczywiście macierz P_a jest samosprężona wtedy i tylko wtedy, gdy $a = 0$ wtedy i tylko wtedy, gdy $\|P_a\| = 1$, co jest zgodne z Twierdzeniem 2.41.

Przykład 2.43 (Rzuty ortogonalne na $L^2(\mu)$ zadane przez zbiory). Niech $H = L^2(\mu)$ będzie przestrzenią Hilberta związaną z pewną przestrzenią z miarą (Ω, Σ, μ) . Dla każdego zbioru mierzalnego $A \in \Sigma$, zbiór funkcji “żyjących na tym zbiorze” $\{f \in L^2(\mu) : f \text{ zeruje się poza } A\}$ jest domkniętą podprzestrzenią H , izometrycznie izomorficzną z przestrzenią $L^2(\mu|_{\Sigma_A})$ zadaną przez przestrzeń z miarą $(A, \Sigma_A, \mu|_{\Sigma_A})$, gdzie $\Sigma_A = \{B \in \Sigma : B \subseteq A\}$. Dalej będziemy te izomorficzne przestrzenie utożsamiać. Rzut ortogonalny z $L^2(\mu)$ na $L^2(\mu|_{\Sigma_A})$ jest operatorem mnożenia przez funkcję charakterystyczną $\mathbb{1}_A$ zbioru A , por. Przykład 1.62. W szczególności $\mathbb{1}_A^2 = \mathbb{1}_A$ oraz $\|\mathbb{1}_A\|_\infty = 1$ (o ile $\mu(A) \neq 0$).

Przykład 2.44 (Warunkowa wartość oczekiwana jako rzut ortogonalny). Niech $H = L^2(\mu)$ będzie jak w przykładzie powyżej. Jeśli $\mathcal{F} \subseteq \Sigma$ jest σ -podalgebrą zbiorów, to $\{f \in L^2(\mu) : f \text{ jest } \mathcal{F}\text{-mierzalna}\}$ jest domkniętą podprzestrzenią H izometrycznie izomorficzną z $L^2(\mu|_{\mathcal{F}})$. Zatem istnieje rzut ortogonalny P z $L^2(\mu)$ na $L^2(\mu|_{\mathcal{F}})$ (traktowaną jako podprzestrzeń tej pierwszej). Rzut ten w rachunku prawdopodobieństwa jest nazywany *warunkową wartością oczekiwaną pod warunkiem σ -ciała $\mathcal{F}$* . Rzeczywiście, jeśli miara μ jest skończona, to P spełnia następujące własności, które definiują warunkową wartość oczekiwaną:⁶

- (1) Pf jest $\mathcal{F}$ -mierzalna dla każdej funkcji $f \in H$,
- (2) Dla każdego zbioru $A \in \mathcal{F}$ i funkcji $f \in H$ mamy $\int_A Pf d\mu = \int_A f d\mu$.

Pierwsza własność zachodzi z definicji, natomiast druga wynika z następującego rachunku

$$\int_A Pf d\mu = \int_\Omega (Pf)\mathbb{1}_A d\mu = \langle Pf, \mathbb{1}_A \rangle \stackrel{P=P^*}{=} \langle f, P\mathbb{1}_A \rangle \stackrel{A \in \mathcal{F}}{=} \langle f, \mathbb{1}_A \rangle = \int_\Omega f\mathbb{1}_A d\mu = \int_A f d\mu.$$

2.3 Układy i bazy ortonormalne

Na kursie algebry (na pierwszym roku studiów) uczymy się, że każda przestrzeń liniowa skończenie wymiarowa posiada bazę, tzn. maksymalny układ wektorów liniowo-niezależnych, i wtedy każdy wektor z przestrzeni ma jednoznaczne przedstawienie jako kombinacja liniowa wektorów bazowych. Taka baza, nazywana również *bazą algebraiczną* lub *bazą Hamela*, istnieje też w nieskończenie wymiarowych przestrzeniach liniowych, lecz nie jest już wtedy tak praktyczna.

W przestrzeniach Banacha oprócz struktury liniowej mamy też topologię, a więc naturalnym jest, aby w pojęciu bazy uwzględnić również pojęcie zbieżności. Teorię takich baz opracował Juliusz Schauder (jeden z przedstawicieli lwowskiej szkoły matematycznej). Mianowicie *bazę Schaudera* w przestrzeni Banacha X nazywamy ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ o tej własności, że dla każdego wektora $x \in X$ istnieje dokładnie jeden ciąg skalarów $\{a_n\}_{n=1}^\infty \subseteq \mathbb{F}$ taki, że

$$x = \sum_{n=1}^{\infty} a_n x_n,$$

⁶Chcąc być uczciwym, trzeba powiedzieć, że na rachunku prawdopodobieństwa wartość oczekiwaną rozpatruje się zazwyczaj na przestrzeni $L^1(\mu)$, która w przypadku miary skończonej, jest na ogół “nieco” większa niż $L^2(\mu)$

gdzie szereg jest zbieżny w sensie normy przestrzeni X . Niestety na obecnym (podstawowym) kursie nie poświęcimy tej bazie więcej miejsca. Warto jednak podkreślić, że baza Schaudera nie zawsze istnieje nawet, gdy X jest *przestrzenią ośrodkową*, czyli gdy X zawiera przeliczalny zbiór gęsty.^a Zagadnienie to było jednym z problemów Księgi Szkockiej, które rozwiązał Per Enflo, za co otrzymał od Stanisława Mazura żywą gęś, patrz rysunek obok. (🏠 poczytać o tym w internetach)



^aJeśli przestrzeń X posiada bazę Schaudera, to musi być ośrodkowa (🏠 dlaczego?)

W przestrzeni Hilberta oprócz zbieżności mamy pojęcie ortogonalności. Prowadzi to do kolejnej (trzeciej) definicji bazy, w której zakładamy, że wektory są parami do siebie prostopadłe - tzw. *baza ortogonalna*. Dodatkowo możemy wektory bazy unormować, tak aby każdy z nich miał długość 1. Wtedy mówimy o *bazie ortonormalnej*. Jak pokażemy poniżej bazy ortonormalne istnieją w każdej przestrzeni Hilberta i są kluczowym narzędziem w analizie tych przestrzeni.

Definicja 2.45. Rodzinę wektorów $\{e_i\}_{i \in I} \subseteq H$ w przestrzeni Hilberta H nazywamy

- a) *układem ortogonalnym*, jeśli $e_i \perp e_j$ dla $i \neq j$ (tzn. $\langle e_i, e_j \rangle = 0$ dla $i \neq j$).
- b) *układem ortonormalnym*, jeśli $e_i \perp e_j$, gdy $i \neq j$ oraz $\|e_i\| = 1$, tzn. gdy

$$\langle e_i, e_j \rangle = \delta_{i,j} := \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i, j \in I.$$

($\delta_{i,j}$ - delta Kroneckera).

Uwaga 2.46. Każdy układ niezerowych wektorów $\{u_i\}_{i \in I}$, który jest ortogonalny można unormować do układu normalnego $\{e_i\}_{i \in I}$. Rzeczywiście kładąc $e_i := \frac{u_i}{\|u_i\|}$, $i \in I$ mamy

$$\langle e_i, e_j \rangle = \left\langle \frac{u_i}{\|u_i\|}, \frac{u_j}{\|u_j\|} \right\rangle = \frac{\langle u_i, u_j \rangle}{\|u_i\| \|u_j\|} = \begin{cases} \frac{\langle u_i, u_i \rangle}{\|u_i\| \|u_i\|}, & i = j, \\ 0, & i \neq j, \end{cases} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i, j \in I.$$

Jeśli zbiór indeksów I jest przeliczalny, to można go ponumerować liczbami naturalnymi i w powyższej definicji można zastąpić rodzinę $\{e_i\}_{i \in I}$ przez ciąg $\{e_i\}_{i=1}^\infty$ albo $\{e_i\}_{i=1}^n$, w zależności od tego, czy zbiór indeksów I jest nieskończony, czy skończony.

W skończenie wymiarowych przestrzeniach Hilberta bazę ortogonalną, odpowiednio ortonormalną, można zdefiniować jako bazę algebraiczną, która ma dodatkowo własność, że wektory ją tworzące są ortogonalne, odpowiednio ortonormalne. Takie bazy łatwo skonstruować:

Przykład 2.47 (Ortogonalizacja Grama-Schmidta). Jeśli $M \subseteq H$ podprzestrzeń skończenie wymiarowa, to istnieje układ ortonormalny $\{e_i\}_{i=1}^n$ taki, że $M = \text{lin}\{e_1, \dots, e_n\}$, tzn. elementy M są kombinacjami liniowymi wektorów $\{e_i\}_{i=1}^n$. Rzeczywiście, niech $\{x_i\}_{i=1}^n$ będzie bazą algebraiczną przestrzeni M , tzn. $\{x_i\}_{i=1}^n$ są liniowo-niezależne oraz $M = \text{lin}\{x_1, \dots, x_n\}$. Proces ortogonalizacji Grama-Schmidta są to następujące rekurencyjne wzory

$$u_1 := x_1, \quad u_2 := x_2 - P_{u_1}x_2, \quad u_3 := x_3 - P_{u_1}x_3 - P_{u_2}x_3, \quad \dots, \quad u_n := x_n - \sum_{i=1}^{n-1} P_{u_i}x_n,$$

które definiują układ ortogonalny $\{u_i\}_{i=1}^n$, który rozpina M , tzn. $M = \text{lin}\{u_1, \dots, u_n\}$ (🏠 sprawdzić to). Kładąc $e_i := \frac{u_i}{\|u_i\|}$, $i = 1, \dots, n$, otrzymujemy układ ortonormalny rozpinający M .

Ortogonalizacja Grama-Schmidta wraz z poniższym stwierdzeniem dają prosty przepis na to jak obliczyć rzut ortogonalny wektora na podprzestrzeń skończenie wymiarową.

Stwierdzenie 2.48. Niech M będzie skończenie wymiarową podprzestrzenią przestrzeni Hilberta H . Dla dowolnego układu ortonormalnego $\{e_i\}_{i=1}^n$ takiego, że $M = \text{lin}\{e_1, \dots, e_n\}$ mamy

$$P_M x = \sum_{i=1}^n \langle x, e_i \rangle e_i, \quad \text{dla każdego } x \in H.$$

Ponadto, $\|P_M x\|^2 = \sum_{i=1}^n |\langle x, e_i \rangle|^2$ dla $x \in M$.

Dowód. Rzut ortogonalny na przestrzeń M istnieje, bo M jest skończenie wymiarowa, a więc domknięta (Wniosek 1.70). Niech $x \in H$. Wtedy $P_M x = \sum_{i=1}^n \lambda_i e_i$, dla pewnych $\{\lambda_i\}_{i=1}^n \subseteq \mathbb{F}$. Korzystając z samosprężoności rzutu P_M , patrz Twierdzenie 2.41, mamy

$$\langle x, e_i \rangle \stackrel{e_i \in M}{=} \langle x, P_M e_i \rangle \stackrel{P_M = P_M^*}{=} \langle P_M x, e_i \rangle = \left\langle \sum_{j=1}^n \lambda_j e_j, e_i \right\rangle = \sum_{j=1}^n \lambda_j \langle e_j, e_i \rangle \stackrel{\langle e_j, e_i \rangle = \delta_{i,j}}{=} \lambda_i.$$

Stąd $P_M x = \sum_{i=1}^n \langle x, e_i \rangle e_i$ oraz

$$\|P_M x\|^2 = \left\langle \sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^n \lambda_j e_j \right\rangle = \sum_{i,j=1}^n \lambda_i \bar{\lambda}_j \langle e_i, e_j \rangle = \sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n |\langle x, e_i \rangle|^2.$$

□

Wniosek 2.49 (Nierówność Bessela). Dla dowolnego układu ortonormalnego $\{e_i\}_{i \in I} \subseteq H$ i wektora $x \in H$ mamy

$$\sum_{i \in I} |\langle x, e_i \rangle|^2 \leq \|x\|^2.$$

Dowód. Niech $A \subseteq I$ będzie skończonym podzbiorem i niech M będzie podprzestrzenią liniową rozpiętą przez wektory $\{e_i\}_{i \in A}$. Na mocy Stwierdzenia 2.48 i Twierdzenia 2.41 mamy

$$\sum_{i \in A} |\langle x, e_i \rangle|^2 = \|P_M x\|^2 \stackrel{\|P_M\| \leq 1}{\leq} \|x\|^2.$$

Stąd $\sum_{i \in I} |\langle x, e_i \rangle|^2 = \sup_{\substack{A \subseteq I \\ \text{skończony}}} \sum_{i \in A} |\langle x, e_i \rangle|^2 \leq \|x\|^2.$

□

W nierówności Bessela napisaliśmy sumę szeregu po dowolnym zbiorze indeksów I . Jest to szereg liczb nieujemnych, więc może być traktowany jako supremum po sumach skończonych. Ogólnie zbieżność szeregów po zbiorach niekoniecznie przeliczalnych definiujemy jak następuje. Aby skrócić formuły przyjmujemy następujące oznaczenie na skończone podbiory:

$$K \Subset I \stackrel{\text{def}}{\iff} K \subseteq I \text{ oraz } K \text{ jest zbiorem skończonym.}$$

Definicja 2.50. Niech $\{x_i\}_{i \in I}$ będzie rodziną wektorów w przestrzeni unormowanej X . Powiemy, że szereg $\sum_{i \in I} x_i$ jest (bezwzględnie) zbieżny i jego suma wynosi $x \in X$, jeżeli dla dowolnego $\varepsilon > 0$ istnieje $K \Subset I$ taki, że $K \Subset J \Subset I$ implikuje $\|\sum_{i \in J} x_i - x\| < \varepsilon$.⁷ Piszemy wtedy $x = \sum_{i \in I} x_i$.

⁷Innymi słowy szereg $\sum_{i \in I} x_i$ jest zbieżny jeżeli sieć sum częściowych $\{\sum_{i \in J} x_i\}_{J \Subset I}$ (indeksowana skończonymi podzbiorymi I skierowanymi przez relację inkluzji) jest zbieżna i wtedy granica tej sieci jest sumą szeregu.

Lemat 2.51. *Jeżeli szereg $\sum_{i \in I} x_i$ jest zbieżny w przestrzeni unormowanej X , to dla każdego $\varepsilon > 0$ istnieje $K \in I$ taki, że dla każdego $J \in I$ rozłącznego z K mamy $\|\sum_{i \in J} x_i\| < \varepsilon$. Implikacja odwrotna zachodzi, gdy X jest przestrzenią zupełną.*

Co więcej, jeżeli $\sum_{i \in I} x_i$ zbieżny, to poza zbiorem przeliczalnym wektory $\{x_i\}_{i \in I}$ się zerują, to znaczy istnieje zbiór przeliczalny $N \subseteq I$ taki, że $i \notin N \implies x_i = 0$ i wtedy $\sum_{i \in I} x_i = \sum_{n \in \mathbb{N}} x_n$.

Dowód. Załóżmy, że szereg $\sum_{i \in I} x_i$ jest zbieżny. Niech $\varepsilon > 0$ i niech $K \in I$ taki, że $K \in J \in I$ implikuje $\|\sum_{i \in J} x_i - x\| < \varepsilon/2$. Weźmy teraz dowolny $J \in I$ taki, że $J \cap K = \emptyset$. Wtedy

$$\left\| \sum_{i \in J} x_i \right\| = \left\| \sum_{i \in K \cup J} x_i - \sum_{i \in K} x_i \right\| \leq \left\| \sum_{i \in K \cup J} x_i - x \right\| + \left\| x - \sum_{i \in K} x_i \right\| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Co dowodzi żądanej implikacji.

Założmy teraz, że dla każdego $\varepsilon > 0$ istnieje $K \in I$ taki, że $\|\sum_{i \in J} x_i\| < \varepsilon$ dla każdego dla każdego $J \in I$ rozłącznego z K . Dla każdego $n \in \mathbb{N}$ wybierzmy $K_n \in I$ taki, że $\|\sum_{i \in J \setminus K_n} x_i\| < 1/n$ dla $J \in I$. Przechodząc ewentualnie do zbiorów $K'_n := K_1 \cup \dots \cup K_n$, możemy założyć, ciąg $\{K_n\}_{n=1}^\infty$ jest wstępujący, tzn. $K_1 \subseteq K_2 \subseteq K_3 \subseteq \dots$. Zauważmy, że zbiór $N := \bigcup_{n \in \mathbb{N}} K_n$ jest przeliczalny oraz jeśli $i \notin N$ to $\|x_i\| < 1/n$ dla każdego $n \in \mathbb{N}$, tzn. $x_i = 0$. Dowodzi to drugiej części tezy. Ponadto ciąg sum częściowych $S_n := \sum_{i \in K_n} x_i$ jest ciągiem Cauchy, gdyż dla $m > n$ mamy $\|S_m - S_n\| = \|\sum_{i \in K_m \setminus K_n} x_i\| \leq 1/n$. Zatem jeśli przestrzeń X jest zupełna, to istnieje granica $x := \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{i \in K_n} x_i \in X$. Szereg $\sum_{i \in I} x_i$ jest zbieżny do x , bo dla każdego $n \in \mathbb{N}$ oraz $J \in I$ zawierającego K_n mamy

$$\left\| \sum_{i \in J} x_i - x \right\| = \lim_{m \rightarrow \infty} \left\| \sum_{i \in J} x_i - \sum_{i \in K_m} x_i \right\| \leq \left\| \sum_{i \in J \setminus K_n} x_i \right\| + \lim_{m \rightarrow \infty} \left\| \sum_{i \in K_m \setminus K_n} x_i \right\| < 2/n.$$

□

Bazę algebraiczną definiuje się jako maksymalny układ wektorów liniowo niezależnych. Bazę ortonormalną zdefiniujemy analogicznie, jako maksymalny układ ortonormalny:

Definicja 2.52. *Bazę ortonormalną przestrzeni Hilberta H nazywamy układ ortonormalny $\{e_i\}_{i \in I}$, który jest maksymalny, tzn. nie istnieje $e \in H$ taki, że układ $\{e_i\}_{i \in I} \cup \{e\}$ jest ortonormalny.*

Uwaga 2.53. Analogicznie bazę ortogonalną definiuje się jako maksymalny układ parami ortogonalnych wektorów niezerowych. Jednakże układ taki możemy zawsze unormować by otrzymać bazę ortonormalną. Dlatego skupimy się tu na bazach ortonormalnych.

Baza ortonormalna zawsze istnieje (jeśli wierzymy w aksjomat wyboru):

Stwierdzenie 2.54. *Każdy układ ortonormalny można rozszerzyć do bazy ortonormalnej. W szczególności każda przestrzeń Hilberta posiada bazę ortonormalną.*

Dowód. Teza wynika z Lematu Kuratowskiego-Zorna. Rzeczywiście niech $\{e_i\}_{i \in I} \subseteq H$ będzie ustalonym układem ortonormalnym. Niech $\mathcal{P}$ będzie rodziną wszystkich układów ortonormalnych $\{e'_i\}_{i \in I'} \subseteq H$ rozszerzających $\{e_i\}_{i \in I}$, czyli $\{e_i : i \in I\} \subseteq \{e'_i : i \in I'\}$. Jest to zbiór częściowo uporządkowany ze względu na relację inkluzji. Ponadto każda rodzina układów ortonormalnych $\mathcal{C} \subseteq \mathcal{P}$, która jest liniowo uporządkowana (tzn. jeśli $u, u' \in \mathcal{C}$, to albo $u \subseteq u'$ albo $u' \subseteq u$) ma ograniczenie górne, a mianowicie $\bigcup_{u \in \mathcal{C}} u$. Zatem na mocy Lematu Kuratowskiego-Zorna istnieje element maksymalny w $\mathcal{P}$. □

Twierdzenie 2.55. *Niech $\{e_i\}_{i \in I} \subseteq H$ będzie układem ortonormalnym w przestrzeni Hilberta H . Następujące warunki są równoważne:*

- (1) $\{e_i\}_{i \in I}$ jest bazą ortonormalną, tzn. $\{e_i\}_{i \in I}$ jest maksymalnym układem ortonormalnym.
- (2) układ $\{e_i\}_{i \in I}$ jest liniowo gęsty w H , tzn. przestrzeń liniowa $\text{lin}\{e_i : i \in I\}$ jest gęsta w H .
- (3) Dla każdego $x \in H$, $\|x\|^2 = \sum_{i \in I} |\langle x, e_i \rangle|^2$, tzn. w nierówności Bessela zachodzi równość. Równość tę nazywa się “tożsamością Parsevala”.
- (4) Każdy $x \in H$ jest postaci $x = \sum_{i \in I} \lambda_i e_i$, gdzie $\lambda_i \in \mathbb{F}$, $i \in I$, a szereg jest zbieżny w sensie Definicji 2.50.

Współczynniki $\{\lambda_i\}_{i \in I} \in \mathbb{F}$ w (4) są wyznaczone jednoznacznie przez x , a mianowicie $\lambda_i = \langle x, e_i \rangle$ dla $i \in I$. Liczby te nazywane są “współczynnikami Fouriera” elementu x przy bazie $\{e_i\}_{i \in I}$.

Dowód. Zauważmy, że jeżeli $x = \sum_{i \in I} \lambda_i e_i$ dla pewnych $\lambda_i \in \mathbb{F}$, to ortonormalność układu $\{e_i\}_{i \in I}$ oraz ciągłość iloczynu skalarnego implikują, że

$$\langle x, e_i \rangle = \left\langle \sum_{j \in I} \lambda_j e_j, e_i \right\rangle = \sum_{j \in I} \lambda_j \langle e_j, e_i \rangle = \sum_{j \in I} \lambda_j \delta_{i,j} = \lambda_i, \quad \text{dla } i \in I.$$

Dowodzi to ostatniej części twierdzenia.

(1) $\Rightarrow$ (2). Niech $M := \overline{\text{lin}\{e_i : i \in I\}}$ będzie domkniętą podprzestrzenią H rozpiętą przez $\{e_i\}_{i \in I}$. Jeśli $M \neq H$, to $M^\perp \neq \{0\}$ na mocy Wniosku 2.36. Istnieje wtedy $e \in M^\perp$ taki, że $\|e\| = 1$. Stąd $\{e_i\}_{i \in I} \cup \{e\}$ jest układem ortonormalny, co przeczy maksymalności $\{e_i\}_{i \in I}$. Zatem $M = H$.

(2) $\Rightarrow$ (3). Niech $x \in H$. Ustalmy $\varepsilon > 0$. Z założenia istnieje zbiór skończony $A \subseteq I$ oraz $v \in M := \text{lin}\{e_i : i \in A\}$ taki, że $\|x - v\| < \varepsilon$. Zauważmy, że przestrzeń M przestrzeń skończenie wymiarowa, więc możemy do niej zastosować Stwierdzenie 2.48. Ponadto stosując Twierdzenie Pitagorasa oraz interpretację rzutu ortogonalnego z Twierdzenia 2.34 dostajemy

$$\begin{aligned} \|x\|^2 &= \|P_M x + (x - P_M x)\|^2 \stackrel{2.21}{=} \|P_M x\|^2 + \|x - P_M x\|^2 \leq \left\{ \begin{array}{l} \|x - P_M x\| \stackrel{2.34}{=} \\ = \inf_{y \in M} \|x - y\| \end{array} \right\} \\ &\leq \|P_M x\|^2 + \|x - v\|^2 < \|P_M x\|^2 + \varepsilon^2 \stackrel{2.48}{=} \sum_{i \in A} |\langle x, e_i \rangle|^2 + \varepsilon^2 \\ &\leq \sum_{i \in I} |\langle x, e_i \rangle|^2 + \varepsilon^2. \end{aligned}$$

Przechodząc z ε do zera otrzymujemy, że $\|x\|^2 \leq \sum_{i \in I} |\langle x, e_i \rangle|^2$. Jest to nierówność przeciwna do nierówności Bessela (Wniosek 2.49). Stąd równość.

(3) $\Rightarrow$ (4). Niech $x \in H$ oraz $\varepsilon > 0$. Szereg $\|x\|^2 = \sum_{i \in I} |\langle x, e_i \rangle|^2$ jest zbieżny (na mocy tożsamości Parsevala). Zatem istnieje $K \Subset I$ taki, że dla każdego $J \Subset I$ rozłącznego z K mamy $\sum_{i \in J} |\langle x, e_i \rangle|^2 < \varepsilon$. Stąd $\sum_{i \in I \setminus K} |\langle x, e_i \rangle|^2 \leq \varepsilon$. Niech teraz $J \Subset I$ zawiera K . Stosując tożsamość Parsevala do wektora $x - \sum_{j \in J} \langle x, e_j \rangle e_j$ dostajemy

$$\|x - \sum_{j \in J} \langle x, e_j \rangle e_j\|^2 = \sum_{i \in I \setminus J} |\langle x, e_i \rangle|^2 \leq \varepsilon,$$

bo $\langle x - \sum_{j \in J} \langle x, e_j \rangle e_j, e_i \rangle = \langle x, e_i \rangle - \sum_{j \in J} \langle x, e_i \rangle \delta_{i,j} = \langle x, e_i \rangle (1 - \delta_{i,j})$. Dowodzi to równości $x = \sum_{i \in I} \langle x, e_i \rangle e_i$.

(4) $\Rightarrow$ (1). Załóżmy nie wprost, że istnieje $e \in H$ taki, że układ $\{e_i\}_{i \in I} \cup \{e\}$ jest ortonormalny. Na mocy założenia $e = \sum_{i \in I} \lambda_i e_i$ dla $\lambda_i \in \mathbb{F}$, $i \in I$. Ale $\lambda_i = \langle e, e_i \rangle = 0$ dla każdego $i \in I$. Zatem $e = 0$, co prowadzi do sprzeczności z warunkiem $\|e\| = 1$. $\square$

Uwaga 2.56. Czasem przez tożsamość Parsevala rozumie się nieco bardziej ogólną formułę: $\langle x, y \rangle = \sum_{i \in I} \langle x, e_i \rangle \overline{\langle y, e_i \rangle}$, gdzie $x, y \in H$ i $\{e_i\}_{i \in I}$ baza ortonormalna w przestrzeni Hilberta H (🏠 udowodnić ten wzór).

Wniosek 2.57. Jeśli $\{e_i\}_{i \in I}$ jest bazą ortonormalną przestrzeni Hilberta H , to każdy wektor $x \in H$ jest jednoznacznie wyznaczony przez swoje współczynniki Fouriera $\{\langle x, e_i \rangle\}_{i \in I}$ i wyraża się za pomocą szeregu, nazywanego “szeregiem Fouriera” elementu x względem bazy $\{e_i\}_{i \in I}$:

$$x = \sum_{i \in I} \langle x, e_i \rangle e_i.$$

Tożsamość Parsevala pozwala obliczać normę x za pomocą wzoru $\|x\| = \sqrt{\sum_{i \in I} |\langle x, e_i \rangle|^2}$.

Nazwa szeregi Fouriera i współczynniki Fouriera wzięła się od wzorów wyprowadzonych w 1807 roku przez Josepha Fouriera, które są szczególnym przypadkiem wzorów, które omówiliśmy powyżej. Fourier rozważał bazę ortonormalną zadaną przez funkcje trygonometryczne:

Przykład 2.58 (Baza funkcji trygonometrycznych). W przestrzeni Hilberta $H = L^2([-\pi, \pi])$ mamy bazę ortonormalną $\{e_n\}_{n \in \mathbb{Z}}$, gdzie $e_0 := \frac{1}{\sqrt{2\pi}}$ oraz

$$e_n(t) := \frac{1}{\sqrt{\pi}} \cos(nt), \quad e_{-n}(t) := \frac{1}{\sqrt{\pi}} \sin(nt) \quad \text{dla } n > 0.$$

Rzeczywiście, korzystając ze wzorów trygonometrycznych, dla $n, m \in \mathbb{Z}$,

$$\cos(nt) \cos(mt) = \frac{\cos[(n-m)t] + \cos[(n+m)t]}{2},$$

$$\sin(nt) \sin(mt) = \frac{\cos[(n-m)t] - \cos[(n+m)t]}{2},$$

oraz stąd, że $\int_{-\pi}^{\pi} \cos(nt) dt = \begin{cases} 0, & n \neq 0 \\ 2\pi, & n = 0 \end{cases}$ i $\int_{-\pi}^{\pi} \cos(nt) \sin(mt) dt = 0$ (bo całka z funkcji nieparzystej) łatwo sprawdzić, że $\{e_n\}_{n \in \mathbb{Z}}$ układ ortonormalny (🏠 przeliczyć). Nieco trudniej jest wykazać, że układ ten jest maksymalny, tzn. że jeżeli $\int_{-\pi}^{\pi} f(t) \cos(nt) dt = \int_{-\pi}^{\pi} f(t) \sin(nt) dt = 0$, dla każdego $n \geq 0$, to funkcja f jest równa zero prawie wszędzie. Ale można w to uwierzyć jeśli się zda sobie sprawę, że funkcje $\cos(nt)$, $\sin(nt)$ mają okres $\frac{2\pi}{n}$, który zbiega do zera. Szereg z Wniosku 2.57 prowadzi do klasycznego rozwinięcia w szereg Fouriera. Mianowicie, dla dowolnej funkcji $f \in L^2([-\pi, \pi])$ mamy

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt), \quad (2.4)$$

przy czym równość należy rozumieć tu w sensie zbieżności w $L^2([-\pi, \pi])$, a współczynniki Fouriera dane są wzorami

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt, \quad \text{dla } n \geq 0 \quad \text{i} \quad b_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) dt, \quad \text{dla } n > 0.$$

Rzeczywiście $f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n = \langle f, e_0 \rangle e_0 + \sum_{n=1}^{\infty} \langle f, e_n \rangle e_n + \sum_{n=1}^{\infty} \langle f, e_{-n} \rangle e_{-n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) dt + \sum_{n=1}^{\infty} \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt \cos(nt) + \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) dt \sin(nt)$.

Postać bazy w poprzednim przykładzie ma swoje źródła w zastosowaniach. Na przykład w akustyce, czy w teorii przetwarzania sygnałów, funkcje $\cos(nt)$ i $\sin(nt)$ odpowiadają n -tym składowym harmonicznym, a wzór (2.4) interpretuje się jako rozbitcie dźwięku opisanego funkcją f na składowe harmoniczne. Jednakże tak dobrane funkcje bazowe są bardzo nieporęczne w rachunkach. Fourier wymyślił sposób jak patrzeć na to zagadnienie w inny sposób, tak aby baza była bardziej naturalna. Prowadzi to do pojęcia transformaty Fouriera, która jest niczym innym jak specjalnym rodzajem operatora unitarnego - izomorfizmu przestrzeni Hilberta. Pokażemy, że każdą przestrzeń Hilberta z wyróżnioną bazą ortonormalną można utożsamiać z następującą przestrzenią i jej standardową bazą “zero-jedynkową”.

Przykład 2.59 (Standardowa baza w $\ell^2(I)$). Niech I dowolny zbiór. Rozważmy przestrzeń Hilberta $\ell^2(I) := \{x : I \rightarrow \mathbb{F} : \sum_{i \in I} |x(i)|^2 < \infty\}$. Zbieżność szeregu można tu traktować w sensie Definicji 2.50. Równoważnie $\ell^2(I)$ jest przestrzenią $L^2(I, 2^I, \mu)$ gdzie μ -miara licząca iloczyn skalarany w $\ell^2(I)$ jest dany wzorem $\langle x, y \rangle = \sum_{i \in I} x(i)y(i)$ i standardową bazę ortonormalną stanowią wektory $\{e_i\}_{i \in I}$, gdzie $e_i(j) = \delta_{i,j}$, dla $j \in I$. W szczególności, jeśli $I = \mathbb{N}$, to $\ell^2(\mathbb{N}) = \ell^2$ oraz $e_n = (\underbrace{0, \dots, 0}_{n-1}, 1, 0, \dots)$, dla $n \in \mathbb{N}$.

Definicja 2.60. Operatorem unitarnym nazywamy odwracalną izometrię liniową $U : H \rightarrow K$ między przestrzeniami unitarnymi H i K . Jeśli taki operator istnieje, to mówimy, że H i K są izomorficzne i piszemy $H \cong K$.

Pojęcie operatora unitarnego zdefiniowane powyżej jest naturalnym izomorfizmem w kategorii przestrzeni unitarnych (pozwala utożsamiać odpowiednie przestrzenie unitarne wraz z całą ich strukturą), ponieważ każda liniowa izometria automatycznie zachowuje iloczyn skalarny:

Lemat 2.61. Niech H i K przestrzenie unitarne. Odwzorowanie liniowe $U : H \rightarrow K$ jest izometrią wtedy i tylko wtedy, gdy zachowuje iloczyn skalarny, to znaczy $\langle Ux, Uy \rangle = \langle x, y \rangle$ dla $x, y \in H$.

Dowód. Jeśli U zachowuje iloczyn skalarny, to U jest izometrią, bo $\|Ux\| = \sqrt{\langle Ux, Ux \rangle} = \sqrt{\langle x, x \rangle} = \|x\|$ dla $x \in H$. On odwrót, jeśli U jest izometrią, to na mocy wzorów polaryzacyjnych (2.2), U zachowuje iloczyn skalarny. Rzeczywiście, na przykład w przypadku zespolonym, mamy

$$\langle Ux, Uy \rangle = \frac{1}{4} \sum_{k=0}^3 i^k \|Ux + i^k Uy\|^2 = \frac{1}{4} \sum_{k=0}^3 i^k \|U(x + i^k y)\|^2 = \frac{1}{4} \sum_{k=0}^3 i^k \|x + i^k y\|^2 = \langle x, y \rangle,$$

dla dowolnych $x, y \in H$. □

Twierdzenie 2.62. Jeśli H jest przestrzenią Hilberta z bazą ortonormalną $\{e_i\}_{i \in I}$, to $H \cong \ell^2(I)$, gdzie izomorfizm jest zadany przez operator unitarny $U : H \rightarrow \ell^2(I)$ dany wzorem $(Ux)(i) := \langle x, e_i \rangle$ dla $x \in H$, $i \in I$, czyli U przyporządkowuje wektorowi $x \in H$ jego współczynniki Fouriera.

Dowód. Z tożsamości Parsevala wynika, że $U : H \rightarrow \ell^2(I)$ jest poprawnie określoną izometrią (por. Wniosek 2.57). Liniowość iloczynu skalarnego ze względu na pierwszą współrzędną implikuje liniowość U : dla $x, y \in H$ oraz $\alpha, \beta \in \mathbb{F}$ mamy

$$[U(\alpha x + \beta y)](i) = \langle \alpha x + \beta y, e_i \rangle = \alpha \langle x, e_i \rangle + \beta \langle y, e_i \rangle = \alpha (Ux)(i) + \beta (Uy)(i).$$

Zatem wystarczy pokazać, że U jest surjekcją. Ustalmy więc dowolny element $a \in \ell^2(I)$. W świetle Lematu 2.51 istnieje zbiór przeliczalny $N \subseteq I$ taki, że $a(i) = 0$ jeśli $i \notin N$. Ponumerujemy

elementy tego zbioru: $N = \{i_1, i_2, \dots\}$. Wtedy $\sum_{k=1}^{\infty} |a(i_k)|^2 = \sum_{i \in I} |a(i)|^2 < \infty$. Stąd wynika, że szereg $\sum_{k=1}^{\infty} a(i_k)e_{i_k}$ jest zbieżny w H . Rzeczywiście, ciąg sum częściowych

$$x_n = \sum_{k=1}^n a(i_k)e_{i_k}, \quad n \in \mathbb{N}$$

jest ciągiem Cauchy, gdyż dla $m > n$ mamy

$$\|x_m - x_n\|^2 = \left\| \sum_{k=n+1}^m a(i_k)e_{i_k} \right\|^2 \stackrel{\text{Parseval}}{=} \sum_{k=n+1}^m |a(i_k)|^2 \xrightarrow{n,m \rightarrow \infty} 0.$$

Zatem z zupełności H otrzymujemy, że ciąg $\{x_n\}_{n=1}^{\infty} \subseteq H$ jest zbieżny do pewnego $x \in H$. Korzystając z ciągłości iloczynu skalarnego, dla każdego $i \in I$ mamy

$$\langle x, e_i \rangle = \lim_{n \rightarrow \infty} \langle x_n, e_i \rangle = a(i).$$

Zatem $Ux = a$. □

Stwierdzenie 2.63. *Przestrzeń Hilberta jest ośrodkowa wtedy i tylko wtedy, gdy H posiada przeliczalną bazę ortonormalną.*

Dowód. " $\Leftarrow$ " Jeśli $\{e_i\}_{i \in I}$ jest przeliczalną bazą ortonormalną H , to biorąc dowolny przeliczalny zbiór gęsty $A \subseteq \mathbb{F}$ (np. $A = \mathbb{Q}$ jeśli $\mathbb{F} = \mathbb{R}$ lub $A = \mathbb{Q} + i\mathbb{Q}$ jeśli $\mathbb{F} = \mathbb{C}$) i kładąc

$$C := \{\lambda_1 e_{i_1} + \dots + \lambda_n e_{i_n} : \lambda_i \in A, \{i_1, \dots, i_n\} \subseteq I\}$$

otrzymujemy przeliczalny zbiór gęsty w H .

" $\Rightarrow$ " Załóżmy, że $C \subseteq H$ jest przeliczalnym zbiorem gęstym. Niech $\{e_i\}_{i \in I}$ będzie bazą ortonormalną H . Zauważmy, że dla $i \neq j$ mamy $\|e_i - e_j\| = \sqrt{\|e_i\|^2 + \|e_j\|^2} = \sqrt{2}$, czyli kule otwarte o środkach w e_i i e_j oraz promieniu $\sqrt{2}/2$ są rozłączne: $K(e_i, \sqrt{2}/2) \cap K(e_j, \sqrt{2}/2) = \emptyset$. Z gęstości zbioru C wynika, że każda kula otwarta $K(e_i, \sqrt{2}/2)$ zawiera pewien element C , nazwijmy go c_i . Z poprzedniej uwagi wynika, że $c_i \neq c_j$ dla $i \neq j$. Innymi słowy, skonstruowaliśmy injekcję $I \ni i \mapsto c_i \in C$ ze zbioru indeksów I w przeliczalny zbiór C . Stąd I musi być również przeliczalny. □

Wniosek 2.64. *Każda nieskończenie wymiarowa, ośrodkowa przestrzeń Hilberta jest izomorficzna z ℓ^2 . Każda przestrzeń Hilberta wymiaru $n < \infty$ jest izomorficzna z $\mathbb{F}^n$, gdzie $\langle x, y \rangle = \sum_{k=1}^n x(k)\overline{y(k)}$.*

Pojęcie wymiaru dla przestrzeni skończenie wymiarowych można uogólnić na dowolne (nieskończenie wymiarowe) przestrzenie Hilberta jak następuje.

Definicja 2.65. *Wymiarem przestrzeni Hilberta H nazywamy moc bazy ortonormalnej w tej przestrzeni i oznaczamy $\dim(H)$.*

Lemat 2.66. *Wymiar przestrzeni Hilberta jest dobrze zdefiniowany, tzn. każde dwie bazy tej samej przestrzeni mają tę samą moc.*

Dowód. Załóżmy, że $\{e_i\}_{i \in I}$ i $\{e'_j\}_{j \in I'}$ to dwie bazy ortonormalne przestrzeni Hilberta H . Jeśli któraś z tych baz jest skończona, to nietrudno pokazać, że druga też jest skończona i wtedy ich moce muszą być równe na mocy faktów z algebry liniowej (🏠 uzupełnić szczegóły). Załóżmy, więc że obie bazy są nieskończone. Dla każdego $i \in I$ mamy $e_i = \sum_{j \in I'} \langle e_i, e'_j \rangle e'_j$, patrz Wniosek 2.57. Zatem na mocy Lematu 2.51 zbiór $N_i := \{j \in I' : \langle e_i, e'_j \rangle \neq 0\}$ jest przeliczalny. Jako że $\{e_i\}_{i \in I}$ maksymalny układ ortonormalny, to $I' = \bigcup_{i \in I} N_i$. Stąd $|I| \leq \aleph_0 |I'| = |I'|$. W podobny sposób, poprzez symetrię otrzymujemy, że $|I'| \leq |I|$. Zatem $|I'| = |I|$ na mocy Twierdzenia Cantora–Bernsteina (🏠 przypomnieć sobie to twierdzenie). □

Z Twierdzenia 2.62 wynika, że wymiar klasyfikuje przestrzenie Hilberta z dokładnością do izomorfizmu:

Wniosek 2.67. *Dwie przestrzenie Hilberta H i K są izomorficzne wtedy i tylko wtedy, gdy $\dim(H) = \dim(K)$.*

2.4 Operatory sprzężone

W tym podrozdziale omówimy operację sprzężenia operatorowego, która w teorii operatorów pełni rolę tak samo ważną (o ile nie ważniejszą) jak sprzężenie zespolone w analizie zespolonej. De facto sprzężenie operatorów jest operacją uogólniającą sprzężenia liczb zespolonych, a także sprzężenie macierzy:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \implies A^* = \begin{pmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{mn}} \end{pmatrix}.$$

Dla ogólnych operatorów na (nieskończenie wymiarowych) przestrzeniach Hilberta operację sprzężenia definiuje się za pomocą iloczynu skalarnego. Mówi się, że operatory T i T^* są ze sobą sprzężone jeżeli dla wszystkich wektorów x i y zachodzi równość

$$\langle Tx, y \rangle = \langle x, T^*y \rangle.$$

Istnienie operatora sprzężonego T^* do dowolnego operatora ograniczonego T wynika z ważnego twierdzenia o postaci funkcjonałów ograniczonych na przestrzeni Hilberta, od którego zaczniemy.

Przypomnijmy, że przestrzeń sprzężona (dualna) do przestrzeni unormowanej H jest to przestrzeń Banacha H^* składająca się z funkcjonałów liniowych i ograniczonych $\varphi : H \rightarrow \mathbb{F} = \mathbb{C}, \mathbb{R}$ wyposażona w normę operatorową (Definicja 1.74). Jedną z charakterystycznych własności wyróżniających przestrzenie Hilberta pośród przestrzeni Banacha jest to, że

przestrzeń sprzężona do przestrzeni Hilberta jest izometrycznie izomorficzna z wyjściową przestrzenią (w przypadku zespolonym antyliniowo izomorficzna).

Mianowicie, w 1907 roku Frigyes Riesz i Maurice Fréchet niezależnie wykazali, że każdy ograniczony funkcjonał liniowy na przestrzeni Hilberta jest zadany przez iloczyn skalarny z pewnym ustalonym wektorem. Fakt ten jest czasem nazywany twierdzeniem o reprezentacji funkcjonału liniowego na przestrzeni Hilberta:

Twierdzenie 2.68 (Riesz-Fréchet). *Niech H przestrzeń Hilberta oraz $f : H \rightarrow \mathbb{F}$ odwzorowanie.*

$$f \in H^* \iff \exists!_{y \in H} f(x) = \langle x, y \rangle \text{ dla każdego } x \in H.$$

Ponadto, wtedy $\|f\| = \|y\|$. W szczególności, przyporządkowanie $H \ni y \mapsto \langle \cdot, y \rangle \in H^$ jest izometrycznym izomorfizmem antyliniowym: $H \cong_{\text{anty}} H^*$.*

Dowód. Jeśli $f(x) = \langle x, y \rangle$, $x \in H$, dla pewnego $y \in H$, to jak wykazaliśmy w dowodzie Stwierdzenia 2.24, $f \in H^*$ oraz $\|f\| = \|y\|$. Zatem odwzorowanie $H \ni y \mapsto \langle \cdot, y \rangle \in H^*$ jest poprawnie określoną izometrią. Antyliniowość tego odwzorowania wynika natychmiast z antyliniowości iloczynu skalarnego ze względu na drugą współrzędną, patrz Uwaga 2.3. Zatem aby wykazać, że

jest to izomorfizm wystarczy pokazać jego surjektywność, tzn. że każdy element H^* jest zadany przez iloczyn skalarny.

Założmy, że $f \in H^*$ dowolny. Wtedy $M := \ker f = f^{-1}(0)$ jest domkniętą podprzestrzenią liniową H . Jeśli $M = H$ to $f \equiv 0$ i wtedy $f(x) = \langle x, y \rangle$, $x \in H$, dla $y = 0$. Założmy zatem, że $M \neq H$. Wtedy $\{0\} \neq M^\perp \subseteq H$ jest niezerową podprzestrzenią liniową (Stwierdzenie 2.24 i Wniosek 2.36). Co więcej twierdzimy, że $\dim(M^\perp) = 1$. Rzeczywiście, weźmy dwa niezerowe wektory $y_1, y_2 \in M^\perp$. Wtedy $f(y_1), f(y_2) \neq 0$, więc możemy zdefiniować np. $\lambda := \frac{f(y_2)}{f(y_1)} \in \mathbb{F}$. Stąd

$$f(\lambda y_1 - y_2) = \lambda f(y_1) - f(y_2) = f(y_2) - f(y_2) = 0.$$

Czyli $\lambda y_1 - y_2 \in \ker f = M$. Ale z drugiej strony $\lambda y_1 - y_2 \in M^\perp$, bo $M^\perp$ jest przestrzenią liniową i $y_1, y_2 \in M^\perp$. Zatem $y_2 = \lambda y_1$, bo $M \cap M^\perp = \{0\}$. Stąd $\dim(M^\perp) = 1$.

Weźmy więc dowolny $y_0 \in M^\perp$ taki, że $\|y_0\| = 1$. Wtedy $M^\perp = \{\lambda y_0 : \lambda \in \mathbb{F}\}$. Zatem rzut $x \in H$ na $M^\perp$ jest dany wzorem $P_{M^\perp}x = \langle x, y_0 \rangle y_0$, patrz Przykład 2.29. Stąd dla $x \in H$ mamy

$$f(x) = f(P_M x + P_{M^\perp} x) = f(P_M x) + f(P_{M^\perp} x) = 0 + f(\langle x, y_0 \rangle y_0) = \langle x, y_0 \rangle f(y_0) = \langle x, \overline{f(y_0)} y_0 \rangle.$$

Czyli kładąc $y := \overline{f(y_0)} y_0$ mamy $f(x) = \langle x, y \rangle$, $x \in H$. □

Uwaga 2.69. Z Twierdzenia 2.68 wynika, że przestrzeń dualna H^* do przestrzeni Hilberta H jest przestrzenią Hilberta z iloczynem skalarnym danym wzorem $\langle f_1, f_2 \rangle := \langle y_2, y_1 \rangle$ gdzie $f_i(x) = \langle x, y_i \rangle$ dla $x \in H$, $i = 1, 2$.

Wniosek 2.70. Niech (Ω, Σ, μ) przestrzeń z miarą. Każdy ograniczony funkcjonal liniowy $f : L^2(\mu) \rightarrow \mathbb{F}$ jest postaci

$$f(x) = \int_{\Omega} x(t) \overline{y(t)} d\mu, \quad x \in L^2(\mu),$$

gdzie $y \in L^2(\mu)$. Ponadto wtedy $\|f\| = \left(\int_{\Omega} |y(t)|^2 d\mu \right)^{\frac{1}{2}} = \|y\|_2$.

Uwaga 2.71. Powyższe twierdzenie i wniosek, oprócz znaczenia teoretycznego, są bardzo praktyczne przy obliczaniu normy funkcjonałów liniowych na przestrzeniach Hilberta. Na przykład, jeśli $f(x) = \int_a^b x(t) \sin t dt + i \int_a^b x(t) \cos t dt$ jest funkcjonałem określonym na przestrzeni $L^2[a, b]$, to jego norma wynosi $\|f\| = \left(\int_a^b |\sin t + i \cos t|^2 d\mu \right)^{\frac{1}{2}} = \left(\int_a^b 1 d\mu \right)^{\frac{1}{2}} = \sqrt{b-a}$.

Twierdzenie 2.68 implikuje istnienie i jednoznaczność operatora sprzężonego do dowolnego operatora ograniczonego:

Twierdzenie 2.72. Jeśli $T : H \rightarrow K$ jest ograniczonym operatorem liniowym między dwoma przestrzeniami Hilberta H i K , to istnieje dokładnie jedna funkcja $T^* : K \rightarrow H$ taka, że

$$\langle Tx, y \rangle = \langle x, T^*y \rangle, \quad \text{dla } x \in H, y \in K. \quad (2.5)$$

Ponadto, T^* jest ograniczonym operatorem liniowym oraz $\|T^*\| = \|T\|$ i $(T^*)^* = T$.

Dowód. Dla ustalonego $y \in K$ funkcja $f(x) := \langle Tx, y \rangle$, $x \in H$, jest ograniczonym funkcjonałem liniowym na H (bo jest złożeniem ograniczonego operatora liniowego T oraz ograniczonego funkcjonału liniowego $\langle \cdot, y \rangle$). W szczególności, dla $x \in H$,

$$|f(x)| = |\langle Tx, y \rangle| \leq \|Tx\| \cdot \|y\| \leq \|T\| \cdot \|y\| \cdot \|x\|,$$

skąd $\|f\| \leq \|T\| \cdot \|y\|$. Zatem na mocy Twierdzenia 2.68 istnieje dokładnie jeden wektor w H , oznaczmy go przez T^*y , taki, że $f(x) = \langle x, T^*y \rangle$, $x \in H$, i wtedy $\|T^*y\| \leq \|T\| \cdot \|y\|$.

Czyli $\langle Tx, y \rangle = \langle x, T^*y \rangle$, $x \in H$. To dowodzi istnienia i jednoznaczności funkcji $T^* : K \rightarrow H$ spełniającej (2.5). Jest to funkcja liniowa, bo dla $y_1, y_2 \in K$, $\lambda \in \mathbb{F}$ oraz dowolnego $x \in H$ mamy

$$\begin{aligned}\langle x, T^*(\lambda y_1 + y_2) \rangle &= \langle Tx, \lambda y_1 + y_2 \rangle = \overline{\lambda} \langle Tx, y_1 \rangle + \langle Tx, y_2 \rangle = \overline{\lambda} \langle x, T^*y_1 \rangle + \langle x, T^*y_2 \rangle \\ &= \langle x, \lambda T^*y_1 \rangle + \langle x, T^*y_2 \rangle = \langle x, \lambda T^*y_1 + T^*y_2 \rangle,\end{aligned}$$

skąd $T^*(\lambda y_1 + y_2) = \lambda T^*y_1 + T^*y_2$. Z otrzymanej wcześniej nierówności $\|T^*y\| \leq \|T\| \cdot \|y\|$ wynika, że $\|T^*\| \leq \|T\|$. Żeby wykazać nierówność przeciwną zauważmy, że sytuacja jest symetryczna i możemy zamienić T i T^* rolami. To znaczy,

$$\langle T^*x, y \rangle = \overline{\langle y, T^*x \rangle} = \overline{\langle Ty, x \rangle} = \langle x, Ty \rangle, \quad x \in K, y \in H,$$

skąd wynika, że $(T^*)^* = T$ i w szczególności $\|T\| = \|(T^*)^*\| \leq \|T^*\|$. $\square$

Definicja 2.73. Niech H, K przestrzenie Hilberta oraz $T : H \rightarrow K$ operatora ograniczonego. Operator $T^* : K \rightarrow H$ spełniający (2.5) nazywamy *operatorem sprzężonym* do T .

Przykład 2.74 (Macierze). Jeśli $H = \mathbb{F}^n$ i $K = \mathbb{F}^m$ skończenie wymiarowe przestrzenie Hilberta, to każdy operator liniowy z H do K można utożsamić z macierzą $A = [a_{i,j}]_{i=1,j=1}^{m,n}$. Wtedy

$$\langle Ax, y \rangle = \sum_{i=1}^m (Ax)(i) \overline{y(i)} = \sum_{i=1}^m \left(\sum_{j=1}^n a_{i,j} x(j) \right) \overline{y(i)} = \sum_{j=1}^n x(j) \sum_{i=1}^m \overline{a_{i,j} y(i)} = \left\langle x, [\overline{a_{j,i}}]_{i=1,j=1}^{m,n} y \right\rangle$$

dla dowolnych $x \in H$ i $y \in K$. Stąd $A^* = [\overline{a_{j,i}}]_{i=1,j=1}^{m,n}$, tzn. sprzężenie operatorowe możemy w tym wypadku utożsamić ze sprzężeniem macierzowym.

Powyższy przykład można uogólnić w kontekście operatorów z jądrem całkowym:

Przykład 2.75 (Operatory z jądrem całkowym). Niech $H = L^2(\mu)$, gdzie (Ω, Σ, μ) dowolna przestrzeń z miarą. Dla uproszczenia rozważmy operator $T : H \rightarrow H$ działający na tej samej przestrzeni Hilberta. Powiemy, że operator T posiada *jądro całkowe* jeżeli istnieje funkcja mierzalna dwóch zmiennych $K : \Omega \times \Omega \rightarrow \mathbb{F}$ taka, że zachodzi wzór

$$(Tx)(s) = \int_{\Omega} K(s, t) x(t) d\mu(t), \quad x \in L^2(\Omega, \Sigma, \mu), s \in \Omega.$$

Zauważmy, że każda funkcja całkowalna w kwadracie $K \in L^2(\Omega \times \Omega, \Sigma \otimes \Sigma, \mu \otimes \mu)$ jest jądrem całkowym operatora ograniczonego, gdyż

$$\begin{aligned}\|Tx\|_2^2 &= \int_{\Omega} \left| \int_{\Omega} K(s, t) x(t) d\mu \right|^2 d\mu \stackrel{\text{N. Schwartza}}{\leq} \int_{\Omega} \left| \left(\int_{\Omega} |K(s, t)|^2 d\mu \right)^{\frac{1}{2}} \left(\int_{\Omega} |x(t)|^2 d\mu \right)^{\frac{1}{2}} \right|^2 d\mu \\ &= \int_{\Omega} \int_{\Omega} |K(s, t)|^2 d\mu \cdot \|x\|_2^2 d\mu = \|K\|_2 \cdot \|x\|_2,\end{aligned}$$

skąd $\|T\| \leq \|K\|_2$. Jeżeli K jest jądrem całkowym operatora T , to stosując twierdzenie Fubiniego (zmianę kolejności całkowania) otrzymujemy

$$\begin{aligned}\langle Tx, y \rangle &= \int_{\Omega} \left(\int_{\Omega} K(s, t) x(t) d\mu(t) \right) \overline{y(s)} d\mu(s) = \int_{\Omega} \int_{\Omega} K(s, t) x(t) \cdot \overline{y(s)} d\mu(s) d\mu(t) \\ &= \int_{\Omega} x(t) \left(\int_{\Omega} K(s, t) \overline{y(s)} d\mu(s) \right) d\mu(t) = \int_{\Omega} x(t) \overline{\int_{\Omega} \overline{K(s, t)} y(s) d\mu(s)} d\mu(t) \\ &= \langle x, T^*y \rangle.\end{aligned}$$

Stąd operator sprzężony jest dany wzorem

$$(T^*x)(s) = \int_{\Omega} \overline{K(t,s)}x(t) d\mu, \quad x \in L^2(\Omega, \Sigma, \mu), s \in \Omega.$$

Czyli T posiada jądro całkowe $K(s, t) \iff T^*$ posiada jądro całkowe $\overline{K(t, s)}$. Oczywiście, jeśli $\Omega = \{1, \dots, n\}$ zbiór skończony (i μ to miara licząca), to jądro całkowe $K : \Omega \times \Omega \rightarrow \mathbb{F}$ możemy utożsamić z macierzą $[K(i, j)]_{i,j=1}^n$.

Przykład 2.76 (Operatory mnożenia). Niech $H = L^2(\mu)$, gdzie (Ω, Σ, μ) dowolna przestrzeń z miarą. Dla dowolnej funkcji istotnie ograniczonej $a \in L^\infty(\mu)$. Wzór

$$(M_ax)(t) = a(t)x(t), \quad x \in L^2(\mu), t \in \Omega,$$

definiuje operator mnożenia przez funkcję a , por. Przykład 1.62. Prosty rachunek pokazuje, że $(M_a)^* = M_{\bar{a}}$ tzn. operator sprzężony od operatora mnożenia przez funkcję $a(t)$ jest operatorem mnożenia przez funkcję $\overline{a(t)}$ (🏠 przeliczyć).

Z ostatniej linijki Twierdzenia 2.72 wynika, że sprzężenie operatorowe $*$: $B(H, K) \rightarrow B(K, H)$ jest odwracalną izometrią. Izometria ta jest antyliniowa i antymultiplikatywna:

Stwierdzenie 2.77. *Operacja sprzężenia operatorowego posiada następujące własności:*

- a) *jest inwolucją: $(T^*)^* = T$ dla dowolnego $B(H, K)$;*
- b) *jest antyliniowa: $(\alpha T + \beta S)^* = \bar{\alpha}T^* + \bar{\beta}S^*$, gdzie $T, S \in B(H, K)$, $\alpha, \beta \in \mathbb{F}$;*
- c) *jest antymultiplikatywna: $(ST)^* = T^*S^*$. gdzie $T \in B(H, K)$ i $S \in B(K, L)$;*
- d) *spełnia tzw. C^* -równość: $\|T\|^2 = \|T^*T\|$ dla dowolnego $T \in B(H, K)$.*

Dowód. Własność a) wykazaliśmy w Twierdzeniu 2.72. Dowód własności b) i c) jest bardzo łatwym i dobrym ćwiczeniem (🏠 sprawdzić). By wykazać d) zauważmy, że dla $h \in H$, $\|h\| \leq 1$,

$$\|Th\|^2 = \langle Th, Th \rangle = \langle h, T^*Th \rangle \stackrel{\text{N. Schwartza}}{\leq} \|h\| \|T^*Th\| \leq \|T^*T\|$$

Stąd $\|T\|^2 \leq \|T^*T\|$. Nierówność w drugą stronę wynika stąd, że $*$ jest izometrią, patrz Twierdzeniu 2.72, gdyż stąd $\|T^*T\| \leq \|T^*\| \cdot \|T\| = \|T\|^2$. $\square$

Powyższe stwierdzenie w zastosowaniu do $B(H) = B(H, H)$ mówi, że przestrzeń wszystkich operatorów ograniczonych na przestrzeni Hilberta wraz z mnożeniem jako składaniem operatorów i operacją sprzężenia operatorowego jest C^* -algebrą (żeby zdać egzamin nie trzeba wiedzieć, co to jest C^* -algebra, ale dobrze wiedzieć, że teoria takich algebr jest dziedziną, w której specjalizuje się egzaminator). W szczególności, za pomocą operacji sprzężenia można zdefiniować wiele ważnych klas operatorów w $B(H)$. Poniższa tabela zawiera przykłady takich klas:

Relacje	Nazwa operatora
$T = T^*$	samosprężony
$TT^* = T^*T$	normalny
$P^2 = P, P = P^*$	rzut ortogonalny
$U^*U = 1$	izometria
$U^*U = UU^* = 1$	unitarny

Dwa pierwsze rzędy można traktować jako definicje. Charakteryzację rzutów ortogonalnych wykazaliśmy w Twierdzeniu 2.41. W dwóch ostatnich charakteryzacjach $1 \in B(H)$ jest operatorem identycznościowym na H . Wynikają one z Lematu 2.61:

Stwierdzenie 2.78. *Operator $U \in B(H)$ jest izometrią wtedy i tylko wtedy, gdy $U^*U = 1$. W szczególności, U jest unitarny wtedy i tylko wtedy, gdy $UU^* = U^*U = 1$.*

Dowód. Korzystając Lematu 2.61 mamy

$$U \text{ izometria} \xLeftrightarrow{2.61} \forall_{x,y \in H} \langle Ux, Uy \rangle = \langle x, y \rangle \iff \forall_{x,y \in H} \langle x, U^*Uy \rangle = \langle x, y \rangle \iff U^*U = 1.$$

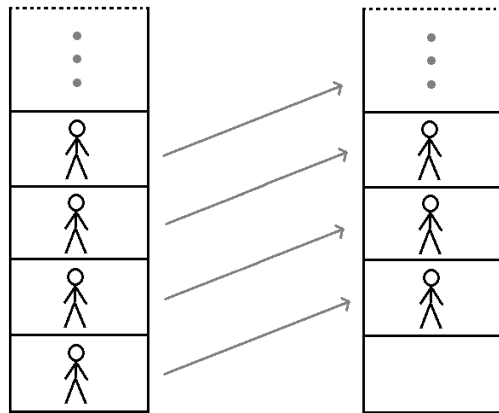
Zatem jeśli $U^*U = UU^* = 1$, to U odwracalna izometria, gdzie $U^* = U^{-1}$, a więc U unitarny. Na odwrót, jeśli U unitarny, to $U^*U = 1$, bo izometria, i stąd $U^{-1} = 1U^{-1} = U^*UU^{-1} = U^*$, czyli $U^*U = UU^* = 1$. $\square$

Jeśli $\dim(H) = n < \infty$, to każdy operator $U \in B(H)$ injektywny jest automatycznie odwracalny, bo $\dim(\ker U) + \dim(\operatorname{Im} U) = \dim(H)$. W szczególności, każda izometria na przestrzeni skończonej wymiarowej jest automatycznie operatorem unitarnym. W przestrzeniach nieskończonej wymiarowej nieodwracalne izometrie istnieją.

Jest to możliwe na mocy *paradoksu Hotelu Hilberta*:

“Jeśli w hotelu jest nieskończona liczba pokoi, to nawet gdy wszystkie pokoje są zajęte, a przychodzi do nas kolejny klient, to możemy go zakwaterować, przesuwając pozostałych, każdego o jeden pokój do góry.”

Na przestrzeni Hilberta ℓ^2 , odpowiednikiem powyższej procedury przekwaterowywania gości, jest tzw. jednostronny operator przesunięcia:



Przykład 2.79 (Operator przesunięcia). *Klasycznym operatorem jednostronnego przesunięcia nazywamy operator $U : \ell^2 \rightarrow \ell^2$ dany wzorem*

$$U(x(1), x(2), x(3), \dots) := (0, x(1), x(2), x(3), \dots).$$

Oczywistym jest, że U jest nieodwracalną izometrią, bo $\|Ux\|^2 = |x(1)|^2 + |x(2)|^2 + \dots = \|x\|^2$ i $\operatorname{Im} U = \{x \in \ell^2 : x(1) = 0\} \neq H$. Operator do niego sprzężony jest postaci (🏠 sprawdzić)

$$U^*(x(1), x(2), x(3), \dots) = (x(2), x(3), \dots),$$

a więc ma nietrywialne jądro $\ker U^* = \{x \in \ell^2 : x = (x(1), 0, 0, \dots)\}$. W szczególności, U nie jest operatorem unitarnym i mamy

$$U^*U = 1, \quad UU^* = P_{\operatorname{Im} U} = 1 - P_{\ker U^*} \neq 1.$$

Uwaga 2.80. Powyższy przykład jest szczególnym przypadkiem następującej sytuacji. Niech (Ω, Σ, μ) będzie przestrzenią z miarą i niech $\varphi : \Delta \rightarrow \Omega$ będzie odwzorowaniem mierzalnym zdefiniowanym na mierzalnym podzbiorze $\Delta \subseteq \Omega$. Można pokazać, że operator kompozycji z takim częściowym odwzorowaniem:

$$(T_\varphi x)(t) = \begin{cases} x(\varphi(t)), & t \in \Delta, \\ 0, & t \notin \Delta, \end{cases} \quad t \in \Omega,$$

jest poprawnie określoną izometrią na $H := L^2(\mu)$ wtedy i tylko wtedy, gdy φ zachowuje miarę μ , tzn. $\mu(\varphi^{-1}(A)) = \mu(A)$ dla każdego $A \in \Sigma$. W szczególności, jeśli $\Omega = \mathbb{N}$, μ to miara licząca, $\Delta = \{2, 3, \dots\}$ oraz $\varphi(n) = n - 1$ dla $n \geq 2$, to $T_\varphi = U$ jest jednostronnym przesunięciem. Ponadto, operator mnożenia M_a przez funkcję $a \in L^\infty(\mu)$, patrz Przykład 2.76, jest izometrią wtedy i tylko wtedy, gdy jest operatorem unitarnym (bo M_a jest operatorem normalnym) wtedy i tylko wtedy, gdy $|a(t)| = 1$ μ -prawie wszędzie (🏠 wyjaśnić szczegóły). Jako że złożenie izometrii jest izometrią, to otrzymujemy, że ważony operator kompozycji

$$(Ux)(t) := [(M_a T_\varphi)x](t) = \begin{cases} a(t)x(\varphi(t)), & t \in \Delta, \\ 0, & t \notin \Delta, \end{cases} \quad t \in \Omega,$$

jest izometrią na $L^2(\Omega, \Sigma, \mu)$ dla dowolnego częściowego odwzorowania $\varphi : \Delta \rightarrow \Omega$ zachowującego miarę μ oraz mierzalnej funkcji (wagi) $a : \Delta \rightarrow \{z \in \mathbb{F} : |z| = 1\}$. Nietrudno zauważyć, że powyższy wzór na U definiuje izometrię na wszystkich przestrzeniach $L^p(\mu)$, $p \geq 1$. Twierdzenie Banacha-Lamperti mówi, że każda izometria na przestrzeni $L^p(\mu)$ dla $p \neq 2$ jest tej postaci.

Na powyższym przykładzie można zaobserwować dwa ogólne fakty. Po pierwsze dla każdej izometrii $U \in B(H)$ operator UU^* jest rzutem ortogonalnym na obraz U (🏠 udowodnić). Po drugie jądro operatora sprzężonego jest dopełnieniem ortogonalnym wyjściowego operatora.

Lemat 2.81. *Jeśli $U \in B(H)$ izometria, to $UU^* = P_{\text{Im } U}$, tzn. UU^* jest rzutem ortogonalnym na obraz operatora U .*

Dowód. Operator UU^* jest samosprężony, bo $(UU^*)^* = (U^*)^*U^* = U^*U$, i jest idempotentem, bo $(UU^*)^2 = U(U^*U)U^* = UU^*$. Zatem, na mocy Twierdzenia 2.41, UU^* jest rzutem ortogonalnym na swój obraz. Oczywiście jest, że $\text{Im } UU^* \subseteq \text{Im } U$. Inkluzja w drugą stronę wynika z relacji $U = U1 = U(U^*U) = (UU^*)U$, która oznacza, że rzut UU^* jest identycznością na $\text{Im } U$. $\square$

Stwierdzenie 2.82. *Dla dowolnego operatora $T \in B(H)$ mamy $(\text{Im } T)^\perp = \ker T^*$.*

Dowód. Jeśli $x \in \ker T^*$, to dla dowolnego $y \in H$ mamy $\langle x, Ty \rangle = \langle T^*x, y \rangle = \langle 0, y \rangle = 0$. Stąd $\ker T^* \subseteq (\text{Im } T)^\perp$. Z drugiej strony jeśli $x \perp \text{Im } T$, to $\|T^*x\|^2 = \langle T^*x, T^*x \rangle = \langle x, T(T^*x) \rangle = 0$, czyli $x \in \ker T^*$. Stąd $(\text{Im } T)^\perp \subseteq \ker T^*$. $\square$

Na koniec powiedzmy parę słów o operatorach normalnych. Pełnią one ważną rolę mimo, iż zdecydowana większość operatorów jest “nienormalna”. Operatory normalne posiadają bardzo dobre własności spektralne (o czym jeszcze mam nadzieję pomówić pod koniec niniejszego wykładu). Na przykład, z algebry liniowej powinniśmy pamiętać, że macierz zespolona A jest normalna wtedy i tylko wtedy, gdy jest diagonalizowalna, tzn. jest postaci $A = UDU^*$, gdzie U to operator unitarny, a D to macierz diagonalna. Oczywiście, każdy operator samosprężony jest normalny, a izometria jest operatorem unitarnym wtedy i tylko wtedy, gdy jest normalna. W szczególności jednostronne przesunięcie (Przykład 2.79) nie jest operatorem normalnym.

Przykład 2.83 (Operatory mnożenia jako operatory normalne). Modelowymi operatorami normalnymi są operatory mnożenia M_a , $a \in L^\infty(\mu)$, z Przykładu 2.76 (twierdzenie spektralne mówi, że z dokładnością do sprzężenia operatorem unitarnym każdy operator normalny jest tej postaci). Rzeczywiście, wprost z definicji operatorów mnożenia wynika, że zachodzi wzór $M_a M_b = M_{ab}$, dla $a, b \in L^\infty(\mu)$, więc skoro $M_a^* = M_{\bar{a}}$, to

$$M_a^* M_a = M_{\bar{a}a} = M_{a\bar{a}} = M_a M_a^*,$$

bo mnożenie funkcji jest przemienne. Podobnie otrzymujemy, że

$$M_a \text{ samosprężony} \iff a \stackrel{\mu\text{-p.w.}}{=} \bar{a} \iff a(t) \in \mathbb{R} \text{ } \mu\text{-prawie wszędzie,}$$

$$M_a \text{ unitarny} \iff |a|^2 \stackrel{\mu\text{-p.w.}}{=} 1 \iff a(t) \in S^1 \text{ } \mu\text{-prawie wszędzie,}$$

$$M_a \text{ rzut ortogonalny} \iff a^2 \stackrel{\mu\text{-p.w.}}{=} a \stackrel{\mu\text{-p.w.}}{=} \bar{a} \iff a(t) \in \{0, 1\} \text{ } \mu\text{-prawie wszędzie.}$$

W szczególności ostatnia równoważność mówi, że operator mnożenia M_a jest rzutem wtedy i tylko wtedy, gdy $a = \mathbb{1}_A$ jest funkcją charakterystyczną pewnego zbioru mierzalnego $A \in \Sigma$.

2.5 Zadania

Zad 2.1. Korzystając ze wzoru z Stwierdzenia 2.48 pokazać, że rzut ortogonalny wektora na podprzestrzeń skończenie wymiarową istnieje w dowolnej przestrzeni unitarnej (niekoniecznie zupełnej).

Zad 2.2. Pokazać, że w Twierdzeniu 2.55 zupełność przestrzeni Hilberta jest kluczowa. To znaczy pokazać, że w dowolnej przestrzeni unitarnej (niekoniecznie zupełnej) można zdefiniować bazy ortonormalne jak w Definicji 2.52 i wykazać ich istnienie jak w Stwierdzeniu 2.54, ale na ogół może wtedy istnieć wektor, którego nie da się przedstawić w postaci szeregu elementów bazowych (Hint: Rozważyć przestrzeń z przykładu 2.25 i bazę ortonormalną podprzestrzeni M).

Zad 2.3. Niech H zespolona przestrzeń Hilberta. Operator $T \in B(H)$ jest samosprężony wtedy i tylko wtedy, gdy $\langle Th, h \rangle \in \mathbb{R}$ dla każdego $h \in H$.

Zad 2.4. Pokazać, że jeśli $T \in B(H)$ jest samosprężony ($T^* = T$), to $\|T\| = \sup_{\|h\|=1} |\langle Th, h \rangle|$.

Zad 2.5. Wykazać, że operator $T \in B(H)$ jest normalny wtedy i tylko wtedy, gdy $\|Th\| = \|T^*h\|$ dla każdego $h \in H$.

Rozdział 3

Trzy zasady analizy funkcjonalnej

W tym rozdziale omówimy trzy podstawowe (najważniejsze) twierdzenia analizy funkcjonalnej:

- Twierdzenie Hahna-Banacha (o przedłużaniu funkcjonałów),
- Twierdzenie Banacha-Steinhaus (o jednostajnej ograniczoności),
- Twierdzenie Banacha o przekształceniu otwartym,

oraz pewne ich zastosowania. W szczególności omówimy tu przykłady przestrzeni dualnych.

3.1 Twierdzenie Hahna-Banacha

Jak już zostało wspomniane, kluczowym narzędziem i obiektem badań analizy funkcjonalnej (od którego wzięła się nazwa tej dziedziny) są funkcjonały liniowe. W latach 20. XX wieku Hans Hahn i Stefan Banach udowodnili twierdzenie, które mówi, że ograniczone funkcjonały liniowe określone na podprzestrzeni zawsze można przedłużyć na całą przestrzeń bez zwiększania normy. Twierdzenie to pozwala konstruować ograniczone funkcjonały liniowe o różnych własnościach, a w szczególności pokazuje, że takich funkcjonałów na każdej przestrzeni unormowanej jest “bardzo dużo”.

Zacznijmy od prostego lematu, za pomocą którego analizę funkcjonałów zespolonych będziemy mogli sprowadzić do analizy funkcjonałów rzeczywistych.

Lemat 3.1. *Niech X będzie zespoloną przestrzenią unormowaną. Wtedy*

$$f : X \rightarrow \mathbb{C} \text{ funkcjonal } \mathbb{C}\text{-liniowy} \iff \exists_{u: X \rightarrow \mathbb{R} \text{ funkcjonal } \mathbb{R}\text{-liniowy}} \forall_{x \in X} f(x) = u(x) + iu(-ix).$$

Ponadto, f jest ograniczony wtedy i tylko wtedy, gdy $u = \operatorname{Re} f$ jest ograniczony i wtedy $\|f\| = \|u\|$.

Dowód. “ $\implies$ ”: Jeśli funkcjonal $f : X \rightarrow \mathbb{C}$ jest $\mathbb{C}$ -liniowy, to kładąc $u := \operatorname{Re} f$ otrzymujemy funkcjonal $\mathbb{R}$ -liniowy $u : X \rightarrow \mathbb{R}$ oraz

$$\operatorname{Im} f(x) = \operatorname{Im} i(-i)f(x) = \operatorname{Im} if(-ix) = \operatorname{Re} f(-ix) = u(-ix).$$

Zatem $f(x) = \operatorname{Re} f(x) + i \operatorname{Im} f(x) = u(x) + iu(-ix)$. Ponadto,

$$\|u\| = \sup_{\|x\|=1} |u(x)| = \sup_{\|x\|=1} |\operatorname{Re} f(x)| \leq \sup_{\|x\|=1} |f(x)| = \|f\|.$$

Czyli jeśli f jest ograniczony, to u jest ograniczony oraz $\|u\| \leq \|f\|$.

“ $\Leftarrow$ ”: Niech $u : X \rightarrow \mathbb{R}$ funkcjonal $\mathbb{R}$ -liniowy. Jasne jest, że wzór $f(x) := u(x) + iu(-ix)$, $x \in X$, definiuje funkcjonal $f : X \rightarrow \mathbb{C}$, który jest $\mathbb{R}$ -liniowy. Co więcej

$$f(ix) = u(ix) + iu(-i^2x) = u(ix) + iu(x) = i(-iu(ix) + u(x)) = i(u(x) + iu(-ix)) = if(x).$$

Stąd f jest $\mathbb{C}$ -liniowy, bo jeśli $z = a + ib$, gdzie $a, b \in \mathbb{R}$, to

$$f(zx) = f(ax + ibx) = f(ax) + f(ibx) = af(x) + ibf(x) = zf(x).$$

Aby oszacować $\|f\|$ weźmy dowolny $x \in X$, $\|x\| = 1$ oraz niech $\lambda \in \mathbb{C}$ będzie taka, że $\lambda f(x) = |f(x)|$. Wtedy albo $f(x) = 0$ albo $|\lambda| = 1$ i stąd

$$|f(x)| = \lambda f(x) = f(\lambda x) = \operatorname{Re} f(\lambda x) = u(\lambda x) \leq \|u\|.$$

Czyli jeśli u jest ograniczony, to f jest ograniczony i $\|f\| \leq \|u\|$. □

Twierdzenie Hahna-Banacha udowodnimy za pomocą ogólniejszego lematu Banacha, który mówi o przedłużaniu funkcjonałów $\mathbb{R}$ -liniowych dominowanych przez tzw. funkcjonal Banacha:

Definicja 3.2. Niech X przestrzeń liniowa nad $\mathbb{R}$. *Funkcjonałem Banacha* nazywamy funkcję $p : X \rightarrow \mathbb{R}$ taką, że

$$(1) \quad \forall_{x,y \in X} p(x+y) \leq p(x) + p(y), \quad (\text{nierówność trójkąta})$$

$$(2) \quad \forall_{x \in X} \forall_{t>0} p(tx) = tp(x). \quad (\text{dodatnia jednorodność})$$

Prototypowymi przykładami funkcjonałów Banacha, są funkcjonały liniowe oraz (pół)normy.

Twierdzenie 3.3 (Lemat Banacha). *Niech $X_0 \subseteq X$ podprzestrzeń liniowa przestrzeni liniowej X i $p : X \rightarrow \mathbb{R}$ funkcjonal Banacha. Każdy funkcjonal liniowy $f_0 : X_0 \rightarrow \mathbb{R}$ majoryzowany przez p przedłuża się do funkcjonału liniowego $f : X \rightarrow \mathbb{R}$ majoryzowanego przez p . Tzn.*

$$\left(\begin{array}{l} X_0 \subseteq X \text{ podprzestrzeń liniowa} \\ f_0 : X_0 \rightarrow \mathbb{R} \text{ funkcjonal liniowy} \\ p : X \rightarrow \mathbb{R} \text{ funkcjonal Banacha} \\ \forall_{x \in X_0} f_0(x) \leq p(x) \end{array} \right) \implies \left(\begin{array}{l} \text{istnieje funkcjonal liniowy } f : X \rightarrow \mathbb{R} \\ f|_{X_0} = f_0 \text{ oraz } \forall_{x \in X} f(x) \leq p(x) \end{array} \right)$$

Dowód. Dowód składa się z dwóch kroków. Najpierw zastosujemy Lemat Kuratowskiego-Zorna, żeby wykazać, że istnieje maksymalne rozszerzenie f majoryzowane przez p . Następnie pokażemy, że to maksymalne rozszerzenie jest de facto określone na całej przestrzeni X .

1) Niech Φ będzie zbiorem wszystkich liniowych przedłużeń dominowanych przez p , tzn.

$$\Phi := \{(\varphi, X_\varphi) : X_\varphi \text{ podprzestrzeń liniowa zawierająca } X_0, \varphi : X_\varphi \rightarrow \mathbb{R} \text{ funkcjonal liniowy taki, że } \varphi|_{X_0} = f_0 \text{ oraz } \varphi \leq p|_{X_\varphi}\}.$$

Na Φ wprowadźmy relację częściowego porządku, gdzie funkcjonal φ_2 jest “większy” niż funkcjonal φ_1 jeżeli jest jego rozszerzeniem. Tzn.

$$(\varphi_1, X_{\varphi_1}) \prec (\varphi_2, X_{\varphi_2}) \stackrel{\text{def}}{\iff} X_{\varphi_1} \subseteq X_{\varphi_2} \text{ oraz } \varphi_2|_{X_{\varphi_1}} = \varphi_1.$$

Nietrudno się przekonać, że każdy łańcuch $\{(\varphi_i, X_i)\}_{i \in I}$ (tzn. zbiór liniowo uporządkowany) ma ograniczenie górne (φ, X_φ) , gdzie $X_\varphi := \bigcup_{i \in I} X_{\varphi_i}$ oraz $\varphi(x) := \varphi_i(x)$, gdy $x \in X_{\varphi_i}$. Zatem na mocy Lematu Kuratowskiego-Zorna istnieje element maksymalny $(\varphi_m, X_{\varphi_m})$ w Φ .

2) Teraz wykażemy, że $X_{\varphi_m} = X$ i tym samym zakończymy dowód kładąc $f = \varphi_m$. Założmy nie wprost, że istnieje $x_0 \in X \setminus X_{\varphi_m}$. Połóżmy

$$X_1 := \text{lin}\{X_{\varphi_m}, x_0\} = \{y + \lambda x_0 : y \in X_{\varphi_m}, \lambda \in \mathbb{R}\}.$$

Wtedy dla dowolnego $u \in \mathbb{R}$ wzór

$$\varphi_1(y + \lambda x_0) := \varphi_m(y) + \lambda u, \quad y \in X_{\varphi_m}, \lambda \in \mathbb{R},$$

definiuje funkcjonal liniowy $\varphi_1 : X_1 \rightarrow \mathbb{R}$, który przedłuża φ_m . Aby uzyskać sprzeczność z maksymalnością φ_m musimy dobrać $u \in \mathbb{R}$ tak, aby $\varphi_1(x) \leq p(x)$ dla $x \in X_1$. W tym celu zauważmy, że

$$\begin{aligned} \varphi_1(x) \leq p(x) &\iff \forall_{\substack{y \in X_{\varphi_m} \\ \lambda \in \mathbb{R}}} \varphi_m(y) + \lambda u \leq p(y + \lambda x_0) \iff \forall_{\substack{y \in X_{\varphi_m} \\ \lambda > 0}} \begin{cases} u \leq \frac{1}{\lambda} (p(y + \lambda x_0) - \varphi_m(y)) \\ u \geq -\frac{1}{\lambda} (p(y - \lambda x_0) - \varphi_m(y)) \end{cases} \\ &\stackrel{3.2(2)}{\iff} \forall_{\substack{y \in X_{\varphi_m} \\ \lambda > 0}} \begin{cases} u \leq p\left(\frac{y}{\lambda} + x_0\right) - \varphi_m\left(\frac{y}{\lambda}\right) \\ u \geq -p\left(\frac{y}{\lambda} - x_0\right) + \varphi_m\left(\frac{y}{\lambda}\right) \end{cases} \\ &\iff \begin{cases} u \leq b := \inf\{p(y + x_0) - \varphi_m(y) : y \in X_{\varphi_m}\} \\ u \geq a := \sup\{-p(y - x_0) + \varphi_m(y) : y \in X_{\varphi_m}\}. \end{cases} \end{aligned}$$

Zatem trzeba tylko pokazać, że dla wartości zdefiniowanych powyżej mamy $a \leq b$, gdyż wtedy biorąc dowolne $u \in [a, b]$ otrzymamy element $(\varphi_1, X_1) \in \Phi$, który jest rozszerzeniem elementu maksymalnego $(\varphi_m, X_m) \in \Phi$, czyli otrzymujemy sprzeczność $\nexists$. Ale dla dowolnych $y_1, y_2 \in X_{\varphi_m}$ mamy

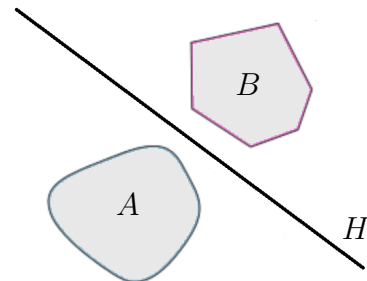
$$\begin{aligned} p(y_1 + x_0) - \varphi_m(y_1) - (-p(y_2 - x_0) + \varphi_m(y_2)) &= p(y_1 + x_0) + p(y_2 - x_0) - \varphi_m(y_1 + y_2) \\ &\stackrel{3.2(1)}{\geq} p(y_1 + y_2) - \varphi_m(y_1 + y_2) \\ &\geq 0 \quad (\text{bo } \varphi_m \leq p). \end{aligned}$$

Stąd $b \geq a$. □

Uwaga 3.4. W literaturze pod nazwą Twierdzenia Hahna-Banacha można spotkać wiele różnych twierdzeń. Wszystkie one wynikają z powyższego Lematu Banacha. My zastosujemy go tylko do przedłużania ograniczonych funkcjonałów liniowych. Innymi ważnymi zastosowaniami, na których omówienie tu niestety nie mamy miejsca, są twierdzenia o rozdzielaniu zbiorów wypukłych. Mianowicie korzystając z Twierdzenia 3.3 można pokazać ($\spadesuit$ dla chętnych), że

Każde dwa niepuste rozłączne wypukłe podzbiory A, B przestrzeni unormowanej X , takie że jeden z nich zawiera punkt wewnętrzny, można rozdzielić hiperpłaszczyzną.

Hiperpłaszczyzną w $\mathbb{R}$ -linowej przestrzeni X nazywamy każdy zbiór postaci $H := \{x \in X : f(x) = \alpha\}$, gdzie $f : X \rightarrow \mathbb{R}$ to funkcjonal liniowy i $\alpha \in \mathbb{R}$. Wtedy zbiory $\{x \in X : f(x) \leq \alpha\}$ i $\{x \in X : f(x) \geq \alpha\}$ nazywamy stronami tej hiperpłaszczyzny. Natomiast zbiory $A, B \subseteq X$ są *rozdzielone* tą hiperpłaszczyzną, jeśli leżą po jej różnych stronach, to znaczy, gdy $f(A) \leq \alpha \leq f(B)$ lub $f(B) \leq \alpha \leq f(A)$.



Twierdzenie 3.5 (Hahn-Banach). *Niech X przestrzeń unormowana (nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$). Każdy ograniczony funkcjonal liniowy $f_0 : X_0 \rightarrow \mathbb{F}$ określony na podprzestrzeni liniowej $X_0 \subseteq X$ przedłuża się (niejednoznacznie) do ograniczonego funkcjonala liniowego $f : X \rightarrow \mathbb{F}$ takiego, że $\|f\| = \|f_0\|$. Tzn.*

$$\left(\begin{array}{l} X \text{ prz. unormowana, nad } \mathbb{F} = \mathbb{R}, \mathbb{C} \\ X_0 \subseteq X \text{ podprzestrzeń liniowa} \\ f_0 : X_0 \rightarrow \mathbb{F} \text{ liniowy i ograniczony} \end{array} \right) \implies \left(\begin{array}{l} \text{istnieje funkcjonal liniowy } f : X \rightarrow \mathbb{F} \\ f|_{X_0} = f_0 \text{ oraz } \|f\| = \|f_0\|. \end{array} \right)$$

Dowód. Rozpatrzmy dwa przypadki, w zależności od tego czy $\mathbb{F} = \mathbb{R}$, czy $\mathbb{F} = \mathbb{C}$.

1) Załóżmy, że $\mathbb{F} = \mathbb{R}$. Połóżmy $p(x) := \|x\| \cdot \|f_0\|$ dla $x \in X$. Wtedy p jest w oczywisty sposób funkcja Banacha takim, że $f_0 \leq p$ na X_0 ($f_0(x) \leq |f_0(x)| \leq \|f_0\| \cdot \|x\| = p(x)$). Zatem na mocy Lematu Banacha (Twierdzenie 3.3) istnieje funkcjonal liniowy $f : X \rightarrow \mathbb{R}$ taki, że $f|_{X_0} = f_0$ oraz $f \leq p$ na X . Stąd

$$\|f\| = \sup_{\|x\|=1} |f(x)| \leq \sup_{\|x\|=1} p(x) = \sup_{\|x\|=1} \|x\| \cdot \|f_0\| = \|f_0\|.$$

Czyli f jest ograniczony oraz $\|f\| \leq \|f_0\|$. Nierówność w drugą stronę jest oczywista (bo $\|f_0\| = \sup_{\|x\|=1, x \in X_0} |f_0(x)| = \sup_{\|x\|=1, x \in X_0} |f(x)| \leq \sup_{\|x\|=1, x \in X} |f(x)| = \|f\|$).

2) Załóżmy, że $\mathbb{F} = \mathbb{C}$. Na mocy Lematu 3.1 mamy $f_0(x) = u_0(x) + iu_0(-ix)$, gdzie $u_0 : X_0 \rightarrow \mathbb{R}$ jest funkcja liniowa oraz $\|u_0\| = \|f_0\|$. Z kroku 1) wiemy, że u_0 przedłuża się do funkcjonala liniowego $u : X \rightarrow \mathbb{R}$ takiego, że $\|u\| = \|u_0\| = \|f_0\|$. Zatem korzystając jeszcze raz z Lematu 3.1 otrzymujemy, że wzór

$$f(x) := u(x) + iu(-ix), \quad x \in X$$

definiuje $\mathbb{C}$ -linowe przedłużenie f_0 oraz $\|f\| = \|u\| = \|f_0\|$. □

Wniosek 3.6. *Dla każdego niezerowego wektora x przestrzeni unormowanej X istnieje funkcjonal $f \in X^*$ taki, że $\|f\| = 1$ oraz $f(x) = \|x\|$. W szczególności funkcjonaly ograniczone rozdziwiają punkty przestrzeni X , tzn.*

$$\forall_{x,y \in X} \quad x \neq y \implies \exists_{f \in X^*} \quad f(x) \neq f(y).$$

Dowód. Niech $x \in X \setminus \{0\}$. Połóżmy $X_0 := \text{lin}\{x\}$ i zdefiniujmy funkcjonal liniowy wzorem $f_0 : X_0 \rightarrow \mathbb{F}$ wzorem $f_0(\lambda x) = \lambda \|x\|$, $\lambda \in \mathbb{F}$. Wtedy $f_0 \in X_0^*$ oraz $\|f_0\| = 1$ (bo $|f_0(\lambda x)| = |\lambda| \|x\| = |\lambda| \cdot \|x\| = \|\lambda x\|$). Zatem f_0 przedłuża się do żadanego funkcjonala f na mocy Twierdzenia 3.5. W szczególności, jeśli $x \neq y$, to $x - y \neq 0$, a więc istnieje $f \in X^*$ taki, że $f(x - y) = \|x - y\| \neq 0$, skąd $f(x) \neq f(y)$. □

Wniosek 3.7. *Każda przestrzeń unormowana X zanurza się w przestrzeń podwójnie sprzężoną $X^{**} := (X^*)^*$. Dokładniej, odwzorowanie $i : X \rightarrow X^{**}$ dane wzorem*

$$i(x)(f) = f(x), \quad x \in X, f \in X^*,$$

jest liniową izometrią.

Dowód. Dla każdego $x \in X$ funkcjonal $i(x) : X^* \rightarrow \mathbb{F}$ jest liniowy, bo

$$i(x)(\lambda f_1 + f_2) = (\lambda f_1 + f_2)(x) = \lambda f_1(x) + f_2(x) = \lambda i(x)f_1 + i(x)f_2.$$

Co więcej,

$$\|i(x)\|_{X^{**}} = \sup_{\|f\|_{X^*}=1} |i(x)(f)| = \sup_{\|f\|_{X^*}=1} |f(x)| \leq \|x\|_X.$$

Czyli $i(x) \in X^{**}$ jest funkcjonalem ograniczonym na X^* oraz $\|i(x)\|_{X^{**}} \leq \|x\|_X$. Aby wykazać inkluzję przeciwną możemy założyć, że $x \neq 0$. Wtedy na mocy Wniosku 3.6 istnieje $f \in X^*$ taki, że $\|f\|_{X^*} = 1$ oraz $f(x) = \|x\|_X$. Zatem $\|i(x)\|_{X^{**}} = \|x\|_X$, czyli odwzorowanie $X \ni x \rightarrow i(x) \in X^{**}$ jest poprawnie określoną izometrią. Ta izometria jest liniowa, bo

$$i(\lambda x_1 + x_2)(f) = f(\lambda x_1 + x_2) = \lambda f(x_1) + f(x_2) = \lambda i(x_1)(f) + i(x_2)(f) = (\lambda i(x_1) + i(x_2))(f).$$

□

3.2 Przykłady przestrzeni sprzężonych. Przestrzenie refleksywne

Każdą przestrzeń unormowaną X możemy zanurzyć w kanoniczny sposób w jej podwójne sprzężenie X^{**} za pomocą izometrii i z Wniosku 3.7. To znaczy utożsamiając X z $i(X)$ możemy napisać $X \subseteq X^{**}$. Jeśli $X = X^{**}$, czyli gdy izometria i jest suriekcją (izomorfizmem), to X nazywamy przestrzenią refleksywną:

Definicja 3.8. Mówimy, że X jest *przestrzenią refleksywną* jeśli $i : X \rightarrow X^{**}$ jest izomorfizmem, czyli gdy każdy funkcjonal na X^* jest postaci $X^* \ni f \mapsto f(x) \in \mathbb{F}$ dla pewnego $x \in X$.

Uwaga 3.9. W powyższej definicji istotne jest by X i X^{**} były izomorficzne kanoniczną izometrią $i(x)(f) = f(x)$, $x \in X$, $f \in X^*$. W szczególności istnieją przestrzenie Banacha X takie, że $X \cong X^{**}$ mimo, iż $i(X) \neq X^{**}$ (🏠 dla chętnych, poczytać w internecie o *przestrzeniach James'a*). Innymi słowy, istnieją nierefleksywne przestrzenie Banacha, dla których podwójne sprzężenie jest izomorficzne z wyjściową przestrzenią (ale nie w sposób kanoniczny).

Żadna przestrzeń niezupełna nie może być refleksywana, gdyż przestrzeń dualna jest zawsze zupełna (na mocy Twierdzenia 1.73). Refleksywne przestrzenie Banacha pod pewnymi względami są bliskie przestrzeniom Hilberta. Poniższa tabela przedstawia, najważniejsze przykłady refleksywnych i nierefleksywnych przestrzeni Banacha:

Refleksywne	Nierefleksywne
przestrzenie skończenie wymiarowe, przestrzenie Hilberta, przestrzenie L^p dla $1 < p < \infty$	c_0 , $C[a, b]$, ℓ^1 , ℓ^∞ , $L^1([a, b])$, $L^\infty([a, b])$

Przestrzenie skończenie wymiarowe i przestrzenie Hilberta są “bardzo refleksywne”, gdyż dla tych przestrzeni nie tylko podwójne sprzężenie, ale już sama przestrzeń sprzężona jest izomorficzna z wyjściową przestrzenią. Dla przestrzeni skończenie wymiarowych jest to dobrze znany fakt z algebry liniowej. Dla przestrzeni Hilberta H jest to Twierdzenie Riesz-Frecheta (Twierdzenie 2.68).

Co prawda izomorfizm $H \xrightarrow{\text{anty}} H^*$ jest antyliniowy, ale złożenie dwóch odwzorowań antyliniowych jest liniowe. W szczególności $H \xrightarrow{\text{anty}} H^* \xrightarrow{\text{anty}} H^{**}$ daje $H \cong H^{**}$ i jest to kanoniczny izomorfizm.

Twierdzenie 3.10 (Przestrzeń dualna do L^p). *Niech $1 \leq p < \infty$ i niech (Ω, Σ, μ) będzie przestrzenią z miarą. Jeśli $p = 1$ zakładamy, że μ jest σ -skończona. Wtedy*

$$L^p(\mu)^* \cong L^q(\mu)$$

gdzie $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ gdy $p = 1$). Dokładniej, $f \in L^p(\mu)^* \iff$ istnieje $y \in L^q(\mu)$ taki, że

$$f(x) = \int_{\Omega} x \cdot y \, d\mu, \quad x \in L^p(\mu).$$

Ponadto wtedy y jest jednoznacznie wyznaczony przez f oraz $\|f\| = \|y\|_q$. To znaczy $\|f\| = \left(\int_{\Omega} |y|^q \, d\mu \right)^{\frac{1}{q}}$ jeśli $p > 1$ oraz $\|f\| = \text{ess sup } |y|$ jeśli $p = 1$.

Dowód. Pokażemy tylko, że $L^q(\mu)$ zanurza się izometrycznie w $L^p(\mu)^*$. Niech $y \in L^q(\mu)$ i połóżmy $f(x) := \int_{\Omega} x \cdot y \, d\mu$, $x \in L^p(\mu)$. Jest to poprawnie określony ograniczony funkcjonal liniowy $f : L^p(\mu) \rightarrow \mathbb{F}$, gdyż z nierówności Höldera dostajemy, że

$$|f(x)| = \left| \int_{\Omega} x \cdot y \, d\mu \right| \leq \int_{\Omega} |x \cdot y| \, d\mu \leq \|x\|_p \cdot \|y\|_q.$$

W szczególności, $\|f\| \leq \|y\|_q$. Jeśli $y = 0$, to $f = 0$. Załóżmy zatem, że $y \neq 0$.

Aby pokazać, że $\|f\| = \|y\|_q$ załóżmy najpierw, że $p > 1$. Wystarczy znaleźć niezerowy $x \in L^p(\mu)$ taki, aby powyższe nierówności były równościami. Przypomnijmy, że w nierówności Höldera równość zachodzi wtedy i tylko wtedy, gdy $|x|^p$ i $|y|^q$ są liniowo zależne, patrz Twierdzenie 1.30. Natomiast, żeby w nierówności $\left| \int_{\Omega} x \cdot y \, d\mu \right| \leq \int_{\Omega} |x \cdot y| \, d\mu$ zachodziła równość wystarczy wziąć x taki, aby $x \cdot y \geq 0$. Biorąc to wszystko pod uwagę połóżmy

$$x := \bar{y}|y|^{q/p-1}.$$

Wtedy $|x|^p = |y|^q$, zatem $\|x\|_p = \|y\|_q^{q/p}$ i $x \in L^p(\mu)$. Ponadto $x \cdot y = |y|^{q/p+1} = |y|^q$. Stąd

$$|f(x)| = \left| \int_{\Omega} x \cdot y \, d\mu \right| = \int_{\Omega} |y|^q \, d\mu = \|y\|_q^q = \|x\|_p \cdot \|y\|_q.$$

Zatem $\|f\| = \|y\|_q$. Załóżmy teraz, że $p = 1$. Jako że miara μ jest σ -skończona, to dla każdego $n \in \mathbb{N}$ znajdziemy zbiór mierzalny $A_n \subseteq \Omega$ taki, że $0 < \mu(A_n) < \infty$ oraz $A_n \subseteq \{t \in \Omega : |y(t)| > \|y\|_{\infty} - 1/n\}$. Skoro $\|y\|_{\infty} > 0$, to dla dostatecznie dużych $n \in \mathbb{N}$ mamy $|y| > 0$ na A_n . Zatem funkcje

$$x_n := \bar{y}|y|^{-1}\mu(A_n)^{-1}\mathbf{1}_{A_n}$$

są poprawnie określone. Ponadto, $\|x_n\|_1 = 1$ oraz

$$|f(x_n)| = \left| \int_{\Omega} x_n \cdot y \, d\mu \right| = \left| \mu(A_n)^{-1} \int_{A_n} |y| \, d\mu \right| \geq \|y\|_{\infty} - 1/n \longrightarrow \|y\|_{\infty}.$$

Stąd $\|f\| = \|y\|_{\infty}$.

Aby wykazać, że każdy funkcjonal z $L^p(\mu)^*$ jest zadany przez całkę z pewną funkcją z $L^q(\mu)$ należy zastosować Twierdzenie Radona-Nikodyma, albo dla ładunków, albo dla miar zespolonych, patrz np. [4, Twierdzenie 6.16]. $\square$

Wniosek 3.11. Dla każdego $1 \leq p < \infty$ mamy $\ell^{p*} \cong \ell^q$ gdzie $\frac{1}{p} + \frac{1}{q} = 1$. Dokładniej $f \in \ell^{p*} \Leftrightarrow \exists_{y \in \ell^q} \forall_{x \in \ell^p} f(x) = \sum_{k=1}^{\infty} x(k)y(k)$ oraz wtedy $\|f\| = \|y\|_q$, to znaczy $\|f\| = (\sum_{k=1}^{\infty} |y(k)|^q)^{\frac{1}{q}}$ jeśli $p > 1$ oraz $\|f\| = \sup_{k \in \mathbb{N}} |y(k)|$ jeśli $p = 1$.

Dowód. Zastosować powyższe twierdzenie do miary liczącej na $\mathbb{N}$. $\square$

Wniosek 3.12. Dla każdego $1 < p < \infty$ oraz każdej przestrzeni (Ω, Σ, μ) z miarą przestrzeni $L^p(\mu)$ jest refleksywna.

Dowód. Skoro $p > 1$, to istnieje skończone $q > 1$ takie, że $\frac{1}{p} + \frac{1}{q} = 1$ (mianowicie $q = \frac{1}{p-1}$). Stosując dwukrotnie Twierdzenie 3.10 mamy

$$L^p(\mu)^{**} = (L^p(\mu)^*)^* \cong L^q(\mu)^* \cong L^p(\mu).$$

Złożenie tych izomorfizmów daje kanoniczny izomorfizm. Zatem $L^p(\mu)$ jest refleksywna. $\square$

Omówimy teraz przestrzeń sprzężoną do przestrzeni funkcji ciągłych. Kluczowym jest tu twierdzenie Riesz o reprezentacji funkcjonałów liniowych dodatnich na przestrzeni funkcji ciągłych za pomocą miar. W szczególności, w przypadku rzeczywistym przestrzenią sprzężoną jest przestrzeń ładunków (miar znakowych). W przypadku zespolonym będzie to przestrzeń miar zespolonych:

Definicja 3.13. Niech Σ będzie σ -algebrą podzbiorów zbioru Ω . *Miarą zespoloną* na przestrzeni mierzalnej (Ω, Σ) nazywamy funkcję $\mu : \Sigma \rightarrow \mathbb{C}$, która jest σ -addytywna, tzn. dla dowolnego ciągu $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma$ parami rozłącznych zbiorów mierzalnych mamy

$$\mu\left(\bigsqcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Normą (lub też *wariacją*) *miary zespolonej* μ nazywamy liczbę

$$\|\mu\| := \sup \left\{ \sum_{n=1}^{\infty} |\mu(A_n)| : \Omega = \bigsqcup_{n=1}^{\infty} A_n, \{A_n\}_{n=1}^{\infty} \subseteq \Sigma \right\}, \quad (3.1)$$

tzn. kres górny bierzemy po wszystkich przeliczalnych rozbiciach mierzalnych przestrzeni Ω .

Uwaga 3.14. Jeśli μ jest miarą zespoloną to $\mu(\emptyset) = 0$ (🏠 pokazać). Jeśli $\mu \geq 0$, tzn. μ jest miarą zespoloną przyjmującą tylko wartości nieujemne, to μ jest “zwykłą” miarą. Aby wyróżnić ten przypadek, mówimy też, że μ jest miarą dodatnią. Wtedy $\|\mu\| = \mu(\Omega)$. Miary probabilistyczne to miary dodatnie, dla których $\|\mu\| = 1$. Jeśli μ przyjmuje wartości rzeczywiste, $\mu(\Omega) \in \mathbb{R}$, to μ jest *ładunkiem* (miarą znakową). Na mocy rozkładu Hahna-Jordana mamy wtedy, że $\mu = \mu_+ - \mu_-$, gdzie $\mu_{\pm}$ wzajemnie osobiwe miary dodatnie, oraz $\|\mu\| = \mu_+(\Omega) + \mu_-(\Omega) = \|\mu_+\| + \|\mu_-\|$.

Definicja 3.15. Niech μ miara zespolona na przestrzeni mierzalnej (Ω, Σ) . Wtedy $\mu_1 = \operatorname{Re} \mu$ i $\mu_2 = \operatorname{Im} \mu$ są ładunkami (miarami znakowymi 🏠 łatwe). Niech $\mu_1 = \mu_{1,+} - \mu_{1,-}$ i $\mu_2 = \mu_{2,+} - \mu_{2,-}$ rozkłady Hahna-Jordana tych ładunków. Mówimy, że funkcja mierzalna $f : \Omega \rightarrow \mathbb{C}$ jest μ -całkowalna jeżeli $\int_{\Omega} |f| d\mu_{i,\pm} < \infty$, $i = 1, 2$, i wtedy kładziemy

$$\int_{\Omega} f d\mu := \int_{\Omega} f d\mu_{1,+} - \int_{\Omega} f d\mu_{1,-} + i \left(\int_{\Omega} f d\mu_{2,+} - \int_{\Omega} f d\mu_{2,-} \right).$$

Przypomnijmy, że całkę z zespolonej funkcji f względem miary dodatniej μ definiujemy wzorem $\int_{\Omega} f d\mu = \int_{\Omega} \operatorname{Re} f d\mu + i \int_{\Omega} \operatorname{Im} f d\mu$.

Uwaga 3.16. Uogólnieniem rozkładu Hahna-Jordana dla miary zespolonej μ jest *rozkład biegunowy*. Mianowicie, patrz [4, Twierdzenie 6.2], dla każdej miary zespolonej μ wzór $|\mu|(E) := \sup \left\{ \sum_{n=1}^{\infty} |\mu(A_n)| : E = \bigsqcup_{n=1}^{\infty} A_n, \{A_n\}_{n=1}^{\infty} \subseteq \Sigma \right\}$, $E \in \Sigma$, definiuje miarę dodatnią na Σ nazywaną *wariacją całkowitą miary* μ . Ponadto, patrz [4, Twierdzenie 6.12], istnieje funkcja mierzalna $h : \Sigma \rightarrow S^1 := \{z \in \mathbb{C} : |z| = 1\}$ taka, że $d\mu = h d|\mu|$, tzn. $\int_{\Omega} x d\mu = \int_{\Omega} x h d|\mu|$, dla $x \in L^1_{\mu}(\Omega)$.

Miarę zespoloną μ określoną na σ -ciele podzbiorów borelowskich przestrzeni topologicznej M nazywamy *miarą borelowską* na M . Mówimy, że przestrzeń topologiczna M jest *lokalnie zwarta*, jeżeli każdy punkt przestrzeni M posiada otoczenie otwarte, którego domknięcie jest zwarte. Każdy podzbiór otwarty (albo domknięty) przestrzeni euklidesowej $\mathbb{R}^k$ jest lokalnie zwarty. Oczywiście, każda przestrzeń zwarta jest także lokalnie zwarta.

Twierdzenie 3.17 (Twierdzenie Riesz o reprezentacji funkcjonałów liniowych na $C_0(M)$). *Niech M będzie lokalnie zwartą przestrzenią metryczną i niech $C_0(M)$ będzie przestrzenią znikających w nieskończoności ciągłych funkcji zespolonych. Dla każdego ograniczonego funkcjonału liniowego $f : C_0(M) \rightarrow \mathbb{C}$ istnieje dokładnie jedna zespolona miara borelowska μ na M taka, że*

$$f(x) = \int_M x d\mu, \quad x \in C_0(M). \quad (3.2)$$

Ponadto, wtedy $\|f\| = \|\mu\|$ oraz

- (1) μ jest miarą dodatnią $\iff f$ jest dodatni, tzn. $f(x) \geq 0$ jeśli $x \geq 0$;
- (2) μ jest ładunkiem $\iff f(x) \in \mathbb{R}$ dla wszystkich rzeczywistych funkcji $x \in C_0(M)$.

Dowód. Teza zachodzi przy nieco bardziej ogólnym założeniu, że M jest lokalnie zwartą przestrzenią Hausdorffa, patrz [4, Twierdzenie 6.19]. Przy czym, wtedy aby uzyskać jednoznaczność miary μ spełniającej (3.2) trzeba założyć, że jest regularna. Dodatnia miara (skończona) jest regularna jeżeli dla dowolnego zbioru borelowskiego E mamy

$$\mu(E) = \inf\{\mu(V) : E \subseteq V \text{ i } V \text{ jest otwarty}\} = \sup\{\mu(K) : K \subseteq E \text{ i } K \text{ jest zwarty}\}.$$

Natomiast zespolona miara borelowska jest regularna, jeśli miara dodatnia $|\mu|$ jest regularna. Dla lokalnie zwartych przestrzeni metrycznych każda miara borelowska jest regularna, patrz [4, Twierdzenie 2.18]. To daje pierwszą część twierdzenia.

Implikacje “ $\implies$ ” w punktach (1), (2) są jasne. Implikacja “ $\impliedby$ ” w punkcie (1) jest treścią klasycznego twierdzenia Riesz, patrz [4, Twierdzenie 2.14]. By wykazać implikację “ $\impliedby$ ” w punkcie (2) założymy nie wprost, że μ nie jest ładunkiem. Wtedy $\mu_2 := \text{Im } \mu$ jest niezerową miarą. Zatem odpowiadający jej funkcjonał $f_2(x) := \int_\Omega x d\mu_2$ (na mocy jednoznaczności przyporządkowania funkcjonałowi miary) też jest niezerowy. Czyli istnieje funkcja $x \in C_0(M)$ taka, że $\int_\Omega x d\mu_2 \neq 0$. Zastępując x przez $\text{Re } x$ lub $\text{Im } x$, możemy przy tym założyć, że x jest funkcją rzeczywistą. Wtedy $f(x) = \int_M x d\text{Re } \mu + i \int_M x d\text{Im } \mu \notin \mathbb{R}$, co przeczy warunkowi w (2). $\square$

Wniosek 3.18 (Przestrzeń dualna do $C_0(M)$). *Niech M będzie lokalnie zwartą przestrzenią metryczną i niech $C_0(M, \mathbb{F})$ będzie przestrzenią znikających w nieskończoności funkcji ciągłych o wartościach w ciele $\mathbb{F} = \mathbb{R}, \mathbb{C}$. Utożsamienie miary z funkcjonałem całkowania $\mu \mapsto \int_M \cdot d\mu$ zadaje izometryczny izomorfizm*

$$C_0(M, \mathbb{F})^* \cong \mathcal{M}_{\mathbb{F}}(M),$$

gdzie $\mathcal{M}_{\mathbb{C}}(M)$ oznacza przestrzeń zespolonych miar borelowskich na M , a $\mathcal{M}_{\mathbb{R}}(M)$ zbiór ładunków borelowskich na M , wraz z działaniami określonymi punktowo $(\lambda\mu)(E) := \lambda\mu(E)$, $(\mu_1 + \mu_2)(E) := \mu_1(E) + \mu_2(E)$, $E \in \Sigma$, $\lambda \in \mathbb{F}$, oraz normą zadaną wzorem (3.1).

Dowód. Dla każdej miary $\mu \in \mathcal{M}_{\mathbb{F}}(M)$ wzór $f(x) := \int_M x d\mu$ definiuje funkcjonał liniowy na $C_0(M, \mathbb{F})$. Jest on poprawnie określony i ograniczony ponieważ korzystając z rozkładu biegunowego $d\mu = h d|\mu|$ miary μ , patrz Uwaga 3.16, dla dowolnego $x \in C_0(M, \mathbb{F})$ mamy

$$|f(x)| = \left| \int_M x(t) h(t) d|\mu|(t) \right| \leq \int_M |x(t)| d|\mu|(t) \leq \|x\|_\infty |\mu|(M) = \|x\|_\infty \|\mu\|.$$

Jasne jest, że przyporządkowanie $\mathcal{M}_{\mathbb{F}}(M) \ni \mu \mapsto \int_M \cdot d\mu \in C_0(M, \mathbb{F})^*$ jest liniowe. Na mocy Twierdzenia 3.17 jest ono izometryczną surjekcją, czyli jest izometrycznym izomorfizmem. W szczególności przestrzeń $\mathcal{M}_{\mathbb{F}}(M)$ jest przestrzenią Banacha. $\square$

Wniosek 3.19. Jeśli $f \in c_0^*$ to istnieje dokładnie jeden $y \in \ell^1$ taki, że

$$f(x) = \sum_{k=1}^{\infty} x(k)y(k), \quad x \in c_0,$$

oraz wtedy $\|f\| = \|y\|_1 = \sum_{k=1}^{\infty} |y(k)|$. W szczególności $c_0^* \cong \ell^1$.

Dowód. Jeśli $M = \mathbb{N}$ jest przestrzenią dyskretną, to c_0 można utożsamić z $C_0(M)$, a odwzorowanie $\mu \mapsto (\mu(\{1\}), \mu(\{2\}), \dots)$ jest izometrycznym izomorfizmem z $\mathcal{M}_{\mathbb{F}}(M)$ na przestrzeń ℓ^1 nad $\mathbb{F}$. Zatem teza jest szczególnym przypadkiem Wniosku 3.18. (ćw. 🏠 przeprowadzić dowód bezpośrednio, bez korzystania z Wniosku 3.18). $\square$

Z Wniosków 3.19, 3.11 otrzymujemy, że przestrzeń podwójnie sprzężona do c_0 to ℓ^∞ :

$$c_0^{**} = (c_0^*)^* \cong (\ell^1)^* \cong \ell^\infty.$$

Przy izomorfizmie $c_0^{**} \cong \ell^\infty$ naturalne zanurzenie $c_0 \subseteq \ell^\infty$ odpowiada naturalnemu zanurzeniu $i(c_0) \subseteq c_0^{**}$. W szczególności, przestrzeń c_0 nie jest refleksywna. Ponadto, przestrzeń c_0 nie może być izomorficzna z przestrzenią ℓ^∞ nawet w sposób niekanoniczny, gdyż c_0 jest ośrodkowa, a ℓ^∞ nie (🏠 pokazać to).

Ogólnym użytecznym kryterium pozwalającym wykazać, że jakaś przestrzeń nie jest refleksywna, jest następujący lemat, który jest kolejnym zastosowaniem Twierdzenia Hahna-Banacha.

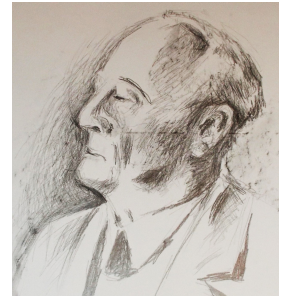
Lemat 3.20. Jeśli przestrzeń Banacha X jest refleksywna, to dla każdego funkcjonału $f \in X^*$ supremum we wzorze $\|f\| = \sup_{\|x\|=1} |f(x)|$ realizuje się dla pewnego $x \in X$, tzn. istnieje $x \in X$ taki, że $\|x\| = 1$ oraz $f(x) = \|f\|$.

Dowód. Niech $f \in X^*$. Na mocy Wniosku 3.6 istnieje funkcjonał $\varphi \in X^{**}$ taki, że $\varphi(f) = \|f\|$ i $\|\varphi\| = 1$. Jeśli $x \in X$ jest elementem odpowiadającym φ przy izomorfizmie $X \cong X^{**}$, (tzn. $i(x) = \varphi$), to $\|x\| = 1$ oraz $f(x) = i(x)(f) = \varphi(f) = \|f\|$. $\square$

Uwaga 3.21. Stosując Lemat 3.20, tzn. pokazując, że jego teza nie zachodzi można wykazać, że przestrzenie ℓ^1 , $C[a, b]$, $L^1[a, b]$ są nie refleksywne (🏠 wykazać, patrz też [3, 6.B.24]). Z nierefleksywności $C[a, b]$ wynika nierefleksywność $L^\infty[a, b]$, gdyż $C[a, b]$ jest domkniętą podprzestrzenią $L^\infty[a, b]$ (🏠 każda podprzestrzeń domknięta przestrzeni refleksywnej jest refleksywna).

3.3 Twierdzenie Banacha-Steinhausa

Legenda głosi, że Hugo Steinhaus w 1916 r. zainteresował się przypadkowo spotkanym Banachem, gdy ten na ławce na Plantach Krakowskich dyskutował żarliwie z Otto Nikodymem o całce Lebesgue'a. Dzięki wstawiennictwu Steinhausa, Banach został zatrudniony na Politechnice Lwowskiej. Współpracowali ze sobą przez wiele lat. Mimo ogromnego dorobku naukowego Steinhausa, zwykł on mawiać, że jego największym odkryciem matematycznym jest Stefan Banach.



W 1927 r. Banach i Steinhaus udowodnili wspólnie twierdzenie zwane *zasadą jednostajnej ograniczoności*. We współczesnym sformułowaniu mówi ono, że jeżeli mamy rodzinę operatorów ograniczonych $\{T_i\}_{i \in I} \subseteq B(X, Y)$ określonych na przestrzeni Banacha X i ograniczonych punktowo, tzn.

$$\forall x \in X \quad \exists M > 0 \quad \forall i \in I \quad \|T_i x\| < M \quad (\text{równoważnie} \quad \forall x \in X \quad \sup_{i \in I} \|T_i x\| < \infty)$$

to rodzina $\{T_i\}_{i \in I}$ jest ograniczona w normie operatorowej (jednostajnie), tzn. zachodzi

$$\exists_{M>0} \forall_{x \in X} \forall_{i \in I} \|T_i x\| < M\|x\| \quad (\text{równoważnie} \quad \sup_{i \in I} \|T_i\| < \infty).$$

Zauważmy, że implikacja odwrotna zawsze zachodzi w sposób trywialny. Natomiast implikacja z Twierdzenia Banacha-Steinhaus'a jest nietrywialna i kluczową rolę pełni tu zupełność przestrzeni X , bez której twierdzenie przestaje być prawdziwe:

Przykład 3.22. Niech $X = c_{00}$ będzie przestrzenią ciągów skończonych z normą supremum i niech $Y = c_0$. Rozważmy ciąg operatorów $\{T_n\}_{n=1}^\infty \subseteq B(X, Y)$ danych wzorem

$$T_n(x(1), x(2), x(3), \dots) = (1 \cdot x(1), 2 \cdot x(2), 3 \cdot x(3), \dots, n \cdot x(n), 0, 0, \dots)$$

Dla każdego $x \in c_{00}$ istnieje $N > 0$ takie, że $x(k) = 0$ dla $k > N$ i stąd ciąg wektorów $\{T_n x\}_{n=1}^\infty$ jest ograniczony. Mianowicie, $\sup_{n \in \mathbb{N}} \|T_n x\| = \max_{n=1, \dots, N} \|T_n x\| \leq N\|x\| < \infty$. Ale ciąg operatorów $\{T_n\}_{n=1}^\infty$ jest nieograniczony, bo

$$\sup_{n \in \mathbb{N}} \|T_n\| = \sup_{n \in \mathbb{N}} n = \infty.$$

Oczywiście nie przeczy to Twierdzeniu Banacha-Steinhaus'a, bo c_{00} nie jest przestrzenią zupełną.

Dowód Twierdzenia Banacha-Steinhaus'a, tak jak również dowód Twierdzenia Banacha o przekształceniu otwartym, które omówimy w kolejnym podrozdziale, opiera się na Twierdzeniu Baire'a. Aby zrozumieć lepiej to ostatnie twierdzenie rozważmy przykład.

Przykład 3.23. Niech $X = [0, 1]$ będzie wyposażony w topologię euklidesową (za X można wziąć jakąkolwiek przestrzeń metryczną bez punktów izolowanych). Wtedy istnieje rodzina $\{A_i\}_{i \in I}$ zbiorów domkniętych i brzegowych (tzn. o pustym wnętrzu) taka, że $X = \bigcup_{i \in I} A_i$, ponieważ

$$X = \bigcup_{x \in X} \{x\}.$$

Zatem przedział $X = [0, 1]$ jest sumą zbiorów nigdziegęstych (*zbiory nigdziegęste* to takie, których domknięcie ma puste wnętrze). Ale suma ta jest nieprzeliczalna. Czy da się wypełnić X przeliczalną ilością zbiorów nigdziegęstych?

Twierdzenie Baire'a na pytanie postawione w Przykładzie 3.23 odpowiada negatywnie:

Twierdzenie 3.24 (Twierdzenie Baire'a). *Przeliczalna suma domkniętych zbiorów brzegowych w przestrzeni metrycznej zupełnej jest zbiorem brzegowym.*

$$\left(\begin{array}{l} X \text{ przestrzeń metryczna zupełna,} \\ A_n \subseteq X \text{ dla } n \in \mathbb{N} \text{ takie, że} \\ \overline{A_n} = A_n \text{ oraz } \text{Int}(A_n) = \emptyset \end{array} \right) \implies \text{Int} \left(\bigcup_{n=1}^\infty A_n \right) = \emptyset.$$

Uwaga 3.25. Poprzez dualność między zbiorami otwartymi i domkniętymi Twierdzenie Baire'a można równoważnie sformułować jak następuje ( sprawdzić). *Przeliczalny przekrój otwartych zbiorów gęstych w przestrzeni metrycznej zupełnej jest zbiorem gęstym, tzn.*

$$\left(\begin{array}{l} X \text{ przestrzeń metryczna zupełna,} \\ U_n \subseteq X \text{ dla } n \in \mathbb{N} \text{ takie, że} \\ \text{Int } U_n = U_n \text{ oraz } \overline{U_n} = X \end{array} \right) \implies \overline{\bigcap_{n=1}^\infty U_n} = X.$$

Dowód. Wykażemy implikację z Uwagi 3.25. Niech $\{U_n\}_{n=1}^\infty$ ciąg otwartych zbiorów gęstych w X . Weźmy dowolny niepusty zbiór otwarty $V \subseteq X$. Potrzebujemy pokazać, że $W \cap \bigcap_{n=1}^\infty U_n \neq \emptyset$. W tym celu skonstruujemy ciąg kul domkniętych $\{K(x_n, r_n)\}_{n=1}^\infty$ zawartych w W takich, że

$$\overline{K}(x_n, r_n) \subseteq U_n \cap K(x_{n-1}, r_{n-1}) \quad \text{oraz} \quad r_n < 1/n, \text{ dla } n > 1. \quad (3.3)$$

Tutaj $\overline{K}(x, r) := \{y \in X : d(x, y) \leq r\}$ jest kulą domkniętą, a $K(x, r) := \{y \in X : d(x, y) < r\}$ kulą otwartą w przestrzeni metrycznej (X, d) . Wtedy z (3.3) wynika, że ciąg kul jest zstępujący:

$$W \supseteq \overline{K}(x_1, r_1) \supseteq \overline{K}(x_2, r_2) \supseteq \overline{K}(x_3, r_3) \supseteq \dots,$$

i ich promienie zbiegają do zera. Stąd $\{x_n\}_{n=1}^\infty$ jest ciągiem Cauchy, a więc na mocy zupełności jest to ciąg zbieżny. Niech $x := \lim_{n \rightarrow \infty} x_n \in X$. Jako że ciąg $\{x_n\}_{n=1}^\infty$ odpewnego miejsca wpada cały do $\overline{K}(x_N, r_N)$ dla każdego $N > 0$, a $\overline{K}(x_N, r_N)$ jest zbiorem domkniętym, to $x = \lim_{n \rightarrow \infty} x_n \in \overline{K}(x_N, r_N)$ dla każdego $N \in \mathbb{N}$. Stąd $x \in \bigcap_{n=1}^\infty \overline{K}(x_n, r_n) \subseteq W \cap \bigcap_{n=1}^\infty U_n$.

Ciąg spełniający (3.3) konstruujemy indukcyjnie. Skoro U_1 jest otwarty i gęsty w X , to $W \cap U_1$ jest niepusty i otwarty. Zatem istnieją $x_1 \in X$ i $r_1 > 0$ takie, że $\overline{K}(x_1, r_1) \subseteq W \cap U_1$, oraz $r_1 < 1$. Załóżmy teraz, że skonstruowaliśmy już x_n i r_n spełniające (3.3). Wtedy przekrój $U_{n+1} \cap K(x_n, r_n)$ jest niepusty i otwarty, bo U_{n+1} jest otwarty i gęsty. Zatem możemy wybrać $x_{n+1} \in X$ oraz r_{n+1} takie, że $\overline{K}(x_{n+1}, r_{n+1}) \subseteq \cap K(x_n, r_n)$, oraz $r_{n+1} < 1$. $\square$

Zupełność w Twierdzeniu Baire'a jest kluczowa:

Przykład 3.26. Niech $X = \mathbb{Q}$ z metryką $d(x, y) = |x - y|$. Ustawmy liczby wymierne w ciąg $q_1, q_2, \dots$ i połóżmy $A_n := \{q_n\}$ dla $n \in \mathbb{N}$. Wtedy $\overline{A_n} = A_n$ oraz $\text{Int}(A_n) = \emptyset$, ale $\bigcup_{n=1}^\infty A_n = \mathbb{Q}$ jest całą przestrzenią, więc wewnątrz ma niepuste (jest zbiorem otwartym). Zatem Twierdzenie Baire'a tutaj nie zachodzi, ale $\mathbb{Q}$ nie jest przestrzenią zupełną.

Przykład 3.27. W przestrzeni $X = c_{00}$ z normą supremum, por. Przykład 3.22, zbiory

$$A_n := \{(x(1), x(2), x(3), \dots) \in c_{00} : (x(1), x(2), x(3), \dots, x(n), 0, 0, \dots)\}, \quad n \in \mathbb{N},$$

są skończenie wymiarowymi podprzestrzeniami liniowymi. Zatem zbiory A_n są domknięte i mają puste wewnątrz w X . Natomiast $X = \bigcup_{n=1}^\infty A_n$ jest całą przestrzenią (ma niepuste wewnątrz).

Twierdzenie 3.28 (Banach-Steinhaus). *Niech $\{T_i\}_{i \in I} \subseteq B(X, Y)$ będzie rodziną operatorów ograniczonych, gdzie X przestrzeń Banacha, a Y przestrzeń unormowana. Wtedy*

$$\forall x \in X \quad \sup_{i \in I} \|T_i x\| < \infty \implies \sup_{i \in I} \|T_i\| < \infty.$$

Czyli rodzina operatorów $\{T_i\}_{i \in I}$ jest ograniczona w $B(X, Y)$ wtedy i tylko wtedy, gdy dla każdego $x \in X$ rodzina wektorów $\{T_i x\}_{i \in I}$ jest ograniczona w Y .

Dowód. Zauważmy, że zbiory $A_n := \{x \in X : \sup_{i \in I} \|T_i x\| \leq n\}$, $n \in \mathbb{N}$, są domknięte, bo T_i są operatorami ciągłymi. Ponadto założenie, że $\sup_{i \in I} \|T_i x\| < \infty$ dla każdego $x \in X$, oznacza, że $X = \bigcup_{n=1}^\infty A_n$. Zatem na mocy Twierdzenia Baire'a jeden ze zbiorów A_n ma niepuste wewnątrz. Czyli istnieją $n \in \mathbb{N}$, $x_0 \in X$ oraz $\varepsilon > 0$ takie, że $K(x_0, \varepsilon) \subseteq A_n$. Niech teraz $x \in X$ taki, że $\|x\| = 1$. Wtedy dla każdego $i \in I$

$$\begin{aligned} \|T_i x\| &= \frac{2}{\varepsilon} \left\| T_i \left(\frac{\varepsilon}{2} x \right) \right\| = \frac{2}{\varepsilon} \left\| T_i \left((x_0 + \frac{\varepsilon}{2} x) - x_0 \right) \right\| \\ &\leq \frac{2}{\varepsilon} \left\| T_i \left(x_0 + \frac{\varepsilon}{2} x \right) \right\| + \frac{2}{\varepsilon} \|T_i(x_0)\| \quad \left\{ \begin{array}{l} x_0 + \frac{\varepsilon}{2} x \in K(x_0, \varepsilon) \subseteq A_n \\ x_0 \in K(x_0, \varepsilon) \subseteq A_n \end{array} \right\} \\ &\leq \frac{2}{\varepsilon} n + \frac{2}{\varepsilon} n = \frac{4}{\varepsilon} n. \end{aligned}$$

Stąd $\sup_{i \in I} \|T_i\| \leq \frac{4}{\varepsilon} n < \infty$. $\square$

Wniosek 3.29. *Operator będący granicą punktową ciągu operatorów ograniczonych określonych na przestrzeni Banacha jest ograniczony. Tzn.*

$$\left(\begin{array}{l} \{T_n\}_{n=1}^\infty \subseteq B(X, Y) \\ X \text{ przestrzeń Banacha} \\ \forall_{x \in X} \{T_n x\}_{n=1}^\infty \text{ zbieżny w } Y \end{array} \right) \implies \left(\begin{array}{l} \text{kładąc } Tx := \lim_{n \rightarrow \infty} T_n x \text{ dla } x \in X \\ \text{mamy } T \in B(X, Y) \end{array} \right)$$

Dowód. Jeśli $\{T_n x\}_{n=1}^\infty$ zbieżny dla każdego $x \in X$, to kładąc $Tx := \lim_{n \rightarrow \infty} T_n x$ otrzymujemy operator liniowy, bo granica jest operacją liniową. Ponadto zbieżność ciągu $\{T_n x\}_{n=1}^\infty$ implikuje jego ograniczoność, dla każdego $x \in X$. Zatem na mocy Twierdzenia Banacha-Steinhausa mamy $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$. Stąd operator T jest ograniczony:

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1} \lim_{n \rightarrow \infty} \|T_n x\| \leq \sup_{\|x\|=1} \sup_{n \in \mathbb{N}} \|T_n\| \|x\| = \sup_{n \in \mathbb{N}} \|T_n\| < \infty.$$

□

3.4 Topologia słaba i *-słaba

Na danym zbiorze można rozpatrywać różne topologie. Im słabsza topologia (zawiera mniej zbiorów otwartych), tym łatwiej danemu ciągowi być zbieżnym, a zbiorowi zwartym. W nieskończonej wymiarowej przestrzeni unormowanej X topologia jest dość mocna. Na przykład, kula domknięta $\{x \in X : \|x\| \leq r\}$ nie jest zwarta w topologii takiej przestrzeni (🏠 pokazać). Dlatego rozważa się na X też słabsze topologie, które pozwalają na szersze zastosowanie właśnie takich pojęć jak zbieżność czy zwartość. Standardowo do tego celu wykorzystuje się topologię zadaną przez ograniczone funkcjonały liniowe. Ogólna procedura definiowania topologii za pomocą odwzorowań jest opisana w poniższym stwierdzeniu:

Lemat 3.30. *Funkcja $f : X \rightarrow \mathbb{F}$ na przestrzeni topologicznej X jest ciągła wtedy i tylko wtedy, gdy dla każdego $x \in X$ i $\varepsilon > 0$ zbiór $U_{f,\varepsilon}(x) := \{y \in X : |f(y) - f(x)| < \varepsilon\}$ jest otwarty w X .*

Dowód. f jest ciągła $\iff$ dla każdego $y \in f(X)$ i $\varepsilon > 0$ przeciwobraz $f^{-1}(K(y, \varepsilon))$ jest otwarty w $X \iff$ dla każdego $x \in X$ i $\varepsilon > 0$ przeciwobraz $f^{-1}(K(f(x), \varepsilon))$ jest otwarty w X . Ale $f^{-1}(K(f(x), \varepsilon)) = \{y \in X : |f(y) - f(x)| < \varepsilon\} = U_{f,\varepsilon}(x)$. □

Stwierdzenie 3.31. *Niech F będzie pewną rodziną funkcji na zbiorze X o wartościach w ciele $\mathbb{F}$ i niech τ_F będzie najslabszą topologią na X , przy której wszystkie funkcje $f : X \rightarrow \mathbb{F}$ należące do F są ciągłe. Topologia taka zawsze istnieje oraz*

- (1) *zbiory postaci $U_{f_1, \dots, f_n, \varepsilon}(x) := \{y \in X : |f_i(y) - f_i(x)| < \varepsilon, 1 \leq i \leq n\}$ gdzie $x \in X$, $f_1, \dots, f_n \in F$, $\varepsilon > 0$, tworzą bazę topologii τ_F .*
- (2) *Topologia τ_F spełnia warunek Hausdorffa wtedy i tylko wtedy, gdy funkcje z F rozdziwiają punkty X , tzn. dla $x \neq y$ istnieje $f \in F$ taki, że $f(x) \neq f(y)$.*
- (3) *Ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ jest zbieżny do $x_0 \in X$ w topologii τ_F wtedy i tylko wtedy, gdy $f(x_n) \rightarrow f(x_0)$ dla każdego $f \in F$.*

Dowód. Na mocy Lematu 3.30 topologia generowana przez zbiory postaci $U_{f,\varepsilon}(x)$, $x \in X$, $f \in F$, generuje najslabszą topologię τ_F , przy której każda funkcja $f \in F$ jest ciągła. Wynika stąd, że

skończone przekroje $\bigcap_{i=1}^n U_{f_i, \varepsilon_i}(x_i)$, gdzie $x_1, \dots, x_n \in X$, $f_1, \dots, f_n \in F$, $\varepsilon_1, \dots, \varepsilon_n > 0$, tworzą bazę topologii τ_F (każdy zbiór z τ_F jest sumą takich skończonych przekrojów). W szczególności,

$$U_{f_1, \dots, f_n, \varepsilon}(x) = \bigcap_{i=1}^n U_{f_i, \varepsilon_i}(x) \in \tau_F.$$

Ponadto, dla dowolnego $x \in \bigcap_{i=1}^n U_{f_i, \varepsilon_i}(x_i)$ kładąc $\varepsilon := \min_{i=1, \dots, n} \varepsilon_i > 0$ mamy,

$$x \in U_{f_1, \dots, f_n, \varepsilon}(x) \subseteq \bigcap_{i=1}^n U_{f_i, \varepsilon_i}(x_i).$$

Stąd zbiory postaci $U_{f_1, \dots, f_n, \varepsilon}(x)$ tworzą bazę topologii τ_F , tzn. dowodzi to punktu (1).

(2). Ten punkt jest jasny (🏠 sprawdzić).

(3). Jeśli $\{x_n\}_{n=1}^\infty \subseteq X$ zbiega do $x_0 \in X$ w τ_F , to dla dowolnego f i $\varepsilon > 0$ ciąg $\{x_n\}_{n=1}^\infty$ od pewnego miejsca wpada do $U_{f, \varepsilon}(x_0)$, tzn. istnieje $N > 0$ taki, że dla każdego $n > N$ mamy $|f(x_n) - f(x_0)| < \varepsilon$. Stąd $f(x_n) \rightarrow f(x_0)$.

Na odwrót, załóżmy, że $f(x_n) \rightarrow f(x_0)$ dla każdego $f \in F$. Ustalmy dowolne bazowe otoczenie punktu x_0 , tzn. $U_{f_1, \dots, f_k, \varepsilon}(x_0)$, gdzie $f_1, \dots, f_k \in F$ oraz $\varepsilon > 0$. Dla każdego $i = 1, \dots, k$ istnieje N_i taki, że dla $n > N_i$, $|f_i(x_n) - f_i(x_0)| < \varepsilon$, tzn. $x_n \in U_{f_i, \varepsilon}(x_0)$. Zatem dla $n > N := \max_{i=1, \dots, k} N_i$ mamy $x_n \in U_{f_1, \dots, f_k, \varepsilon}(x_0) = \bigcap_{i=1}^k U_{f_i, \varepsilon}(x_0)$. Czyli x_n zbiega do x_0 w τ_F $\square$

Uwaga 3.32. Przypomnijmy, że zbieżność ciągu determinuje topologię w przestrzeni metrycznej: podzbiór A przestrzeni metrycznej jest domknięty wtedy i tylko wtedy, gdy A zawiera granicę każdego ciągu zbieżnego $\{x_n\}_{n=1}^\infty \subseteq A$. Twierdzenie to przenosi się na ogólne przestrzenie topologiczne pod warunkiem, że ciągi zastąpimy sieciami. Przypomnijmy, że *sieć* nazywamy rodzinę $\{x_i\}_{i \in I} \subseteq X$ indeksowaną zbiorem skierowanym przez pewną relację częściowego porządku $\leq$. Mówimy, że taka *sieć jest zbieżna* do punktu $x_0 \in X$, w przestrzeni topologicznej X , jeżeli dla dowolnego otwartego otoczenia U punktu x_0 istnieje taki wskaźnik $i_0 \in I$, że dla każdego $i \geq i_0$ mamy $x_i \in U$. Piszemy wtedy $x_i \rightarrow x_0$. Można pokazać (🏠 dla chętnych), że

podzbiór A przestrzeni topologicznej X jest domknięty wtedy i tylko wtedy, gdy dla każdej sieci $\{x_i\}_{i \in I} \subseteq A$ jeśli $x_i \rightarrow x_0$ w X , to $x_0 \in A$.

Sieć $\{x_i\}_{i \in I} \subseteq X$ jest zbieżna do $x_0 \in X$ w topologii τ_F ze Stwierdzenia 3.31 wtedy i tylko wtedy, gdy $f(x_i) \rightarrow f(x_0)$ dla każdego $f \in F$. Zatem opis ten całkowicie determinuje topologię τ_F .

Definicja 3.33. Niech X przestrzeń unormowana. *Słabą topologią* na X nazywamy najslabszą topologię, przy której wszystkie funkcjonały z X^* są ciągłe. Mówimy, że ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ jest *słabo zbieżny* do $x_0 \in X$, jeżeli jest zbieżny do x_0 w słabej topologii. Piszemy wtedy, $x_n \xrightarrow{w} x_0$.

Uwaga 3.34. Na mocy Stwierdzenia 3.31 bazą słabej topologii są zbiory postaci $U_{f_1, \dots, f_n, \varepsilon}(x) := \{y \in X : |f_i(y) - f_i(x)| < \varepsilon, 1 \leq i \leq n\}$ gdzie $x \in X$, $f_1, \dots, f_n \in X^*$, $\varepsilon > 0$. Natomiast ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ jest słabo zbieżny do $x_0 \in X$, wtedy i tylko wtedy, gdy ciąg liczbowy $\{f(x_n)\}_{n=1}^\infty$ jest zbieżny do $f(x_0)$, dla każdego funkcjonału liniowego i ograniczonego $f \in X^*$, tzn.

$$x_n \xrightarrow{w} x_0 \iff \forall f \in X^* f(x_n) \rightarrow f(x_0).$$

Granica słabo zbieżnego ciągu jest wyznaczona jednoznacznie, bo w świetle Stwierdzenia 3.31(2) oraz Wniosku 3.6 (z Twierdzenia Hahna-Banacha), topologia słaba jest topologią Hausdorffa. Powyższe stwierdzenia uogólniają się z ciągów na sieci.

Uwaga 3.35. Jako że ograniczone funkcjonały liniowe są ciągłe na X w normie, to topologia na X zadana przez normę jest mocniejsza od topologii słabej (stąd nazwa tej drugiej). W szczególności, zbieżność w normie implikuje zbieżność słabą:

$$x_n \xrightarrow{\|\cdot\|} x_0 \implies x_n \xrightarrow{w} x_0.$$

Implikacja odwrotna na ogół nie zachodzi. De facto można pokazać, że topologia normowa pokrywa się z topologią słabą wtedy i tylko wtedy, gdy $\dim(X) < \infty$ (🏠 dla chętnych, patrz [3, 6.C.20]).

Stwierdzenie 3.36. *Każdy ciąg słabo zbieżny jest ograniczony, tzn*

$$x_n \xrightarrow{w} x_0 \implies \{x_n\}_{n=1}^{\infty} \text{ ograniczony w normie.}$$

Ponadto, wtedy $\|x_0\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$.

Dowód. Przypomnijmy, iż każdy element $x \in X$ możemy traktować jako funkcjonal $i(x)$ na przestrzeni X^* , gdzie $i(x)(f) = f(x)$, oraz $\|i(x)\| = \|x\|$, patrz Wniosek 3.7. Ponadto,

$$x_n \xrightarrow{w} x_0 \iff \forall f \in X^* \ i(x_n)(f) \longrightarrow i(x_0)(f).$$

Innymi słowy słaba zbieżność ciągu $\{x_n\}_{n=1}^{\infty} \subseteq X$ jest równoważna zbieżności punktowej ciągu funkcjonałów liniowych $\{i(x_n)\}_{n=1}^{\infty} \subseteq X^{**} = B(X^*, \mathbb{F})$. Zatem słaba zbieżność pociąga ograniczoność na mocy Twierdzenia Banacha-Steinhaus'a, por. Wniosek 3.29.

Aby wykazać, oszacowanie na normę x_0 możemy założyć, że $x_0 \neq 0$. Wtedy na mocy Twierdzenia Hahna-Banacha, patrz Wniosek 3.6, istnieje funkcjonal $f \in X^*$ taki, że $\|f\| = 1$ oraz $f(x_0) = \|x_0\|$. Stąd i ze słabej zbieżności otrzymujemy

$$\|x_0\| = f(x_0) = \lim_{n \rightarrow \infty} f(x_n) = \liminf_{n \rightarrow \infty} f(x_n) \leq \liminf_{n \rightarrow \infty} \|f\| \|x_n\| = \liminf_{n \rightarrow \infty} \|x_n\|.$$

□

Przykład 3.37 (Słaba zbieżność w przestrzeniach Hilberta). Jeśli $X = H$ to przestrzeń Hilberta, to na mocy Twierdzenia Riesz-Frechet'a (Twierdzenie 2.68) każdy funkcjonal $f \in H^*$ jest postaci $f(x) = \langle x, y \rangle$, dla pewnego $y \in H$. Zatem dla każdego ciągu $\{x_n\}_{n=1}^{\infty}$ zachodzi

$$x_n \xrightarrow{w} x_0 \iff \forall y \in H \ \langle x_n, y \rangle \longrightarrow \langle x_0, y \rangle.$$

Dla przykładu rozważmy przeliczalny układ ortonormalny $\{e_n\}_{n=1}^{\infty}$ w H , taki układ zawsze istnieje, gdy $\dim(H) = \infty$. Wtedy $\|e_n - e_m\| = \sqrt{2}$ dla $n \neq m$. Zatem ciąg $\{e_n\}_{n=1}^{\infty}$ nie jest zbieżny w normie. Twierdzimy jednak, że jest słabo zbieżny do zera:

$$e_n \xrightarrow{w} 0.$$

Rzeczywiście, niech $y \in H$. Wtedy $\sum_{n=1}^{\infty} |\langle e_i, y \rangle|^2 \leq \|y\|^2$ na mocy nierówności Bessela, patrz Wniosek 2.49. Skoro szereg $\sum_{n=1}^{\infty} |\langle e_i, y \rangle|^2$ jest zbieżny, to $\langle e_i, y \rangle \rightarrow 0 = \langle 0, y \rangle$. Czyli $e_n \xrightarrow{w} 0$. W szczególności, oszacowanie na normę w Stwierdzeniu 3.36 może być ostrą nierównością: $\|0\| = 0 < 1 = \liminf_{n \rightarrow \infty} \|e_n\|$.

Przykład 3.38 (Słaba zbieżność w przestrzeniach L^p). Niech (Ω, Σ, μ) przestrzeń z miarą σ -skończoną oraz niech $1 \leq p < \infty$. Niech $q > 1$ takie, że $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ gdy $p = 1$). Na mocy Twierdzenia 3.10 mamy, że dla każdego ciągu $\{x_n\}_{n=1}^{\infty} \subseteq L^p_{\mu}(\Omega)$ zachodzi

$$x_n \xrightarrow{w} x_0 \text{ w } L^p_{\mu}(\Omega) \iff \forall y \in L^q_{\mu}(\Omega) \ \int_{\Omega} x_n(t)y(t) d\mu(t) \longrightarrow \int_{\Omega} x_0(t)y(t) d\mu(t).$$

W przypadku, gdy μ jest miarą liczącą na $\Omega = \mathbb{N}$ dostajemy

$$x_n \xrightarrow{w} x_0 \text{ w } \ell^p \iff \forall_{y \in \ell^q} \sum_{k=1}^{\infty} x_n(k)y(k) \longrightarrow \sum_{k=1}^{\infty} x_0(k)y(k).$$

Dla $p > 1$, podobnie jak w poprzednim przykładzie można pokazać, że kładąc $e_n = (\underbrace{0, \dots, 0}_{n-1 \text{ razy}}, 1, 0, \dots)$

otrzymujemy ciąg $\{e_n\}_{n=1}^{\infty}$, który nie jest zbieżny w normie ℓ^p , ale jest słabo zbieżny do zera (🏠 sprawdzić). Dla $p = 1$ sytuacja jest wyjątkowa. Okazuje się, że mimo, iż ℓ^1 jest nieskończenie wymiarowa, to słaba zbieżność ciągów w ℓ^1 jest równoważna zbieżności w normie (jest to tzw. Twierdzenie Schura, 🏠 dla chętnych, patrz [3, 6.C.25]). Nie oznacza to jednak, że słaba topologia na ℓ^1 pokrywa się z topologią zadaną przez normę. Wręcz przeciwnie topologie te są różne i w szczególności słaba zbieżność sieci nie jest równoważna zbieżności sieci w normie, por. Uwaga 3.32.

Przykład 3.39 (Słaba zbieżność w przestrzeni $C_0(M)$). Niech M będzie lokalnie zwartą przestrzenią metryczną. Zbieżność w przestrzeni $C_0(M)$, wyposażonej w normę supremum, jest zbieżnością jednostajną. Twierdzimy, że zbieżność słaba w tej przestrzeni jest równoważna zbieżności punktowej modulo ograniczoność ciągu. To znaczy twierdzimy, że

$$x_n \xrightarrow{w} x_0 \text{ w } C_0(M) \iff \forall_{t \in M} x_n(t) \longrightarrow x_0(t) \text{ oraz } \sup_{n \in \mathbb{N}} \|x_n\| < \infty.$$

Przypomnijmy, że na mocy Twierdzenia Riesz, patrz Wniosek 3.18, przestrzeń dualna do $C_0(M)$ to przestrzeń miar $\mathcal{M}_{\mathbb{F}}(M)$. Zatem

$$x_n \xrightarrow{w} x_0 \text{ w } C_0(M) \iff \forall_{\mu \in \mathcal{M}_{\mathbb{F}}(M)} \int_M x_n d\mu \longrightarrow \int_M x_0 d\mu.$$

Zatem jeśli $x_n \xrightarrow{w} x_0$, to oznaczając przez δ_t miarę probabilistyczną skupioną w punkcie $t \in M$, dostajemy $x_n(t) = \int_M x_n d\delta_t \longrightarrow \int_M x_0 d\delta_t = x_0(t)$. Natomiast ograniczoność ciągu w normie: $\sup_{n \in \mathbb{N}} \|x_n\| < \infty$ wynika ze Stwierdzenia 3.36. Załóżmy teraz na odwrót, że ciąg $\{x_n\}_{n=1}^{\infty} \subseteq C_0(M)$ jest punktowo zbieżny oraz ograniczony w normie. Wtedy dla każdej miary dodatniej miary borelowskiej μ mamy $\int_M x_n d\mu \longrightarrow \int_M x_0 d\mu$ na mocy Twierdzenia Lebesgue'a o zbieżności zmajorizowanej. Relacja ta przedłuża się na ładunki poprzez rozkład Hahna-Jordana i dalej na miary zespolone poprzez rozkład na część rzeczywistą i część urojoną.

Możemy też w pewnym sensie odwrócić sytuację i zadać topologię na X^* za pomocą elementów X . W szczególności, tak jak w dowodzie Stwierdzenia 3.36, możemy traktować elementy przestrzeni X jako funkcjonały na przestrzeni dualnej X^* . Prowadzi to do następującej definicji.

Definicja 3.40. Niech X przestrzeń unormowana. *Topologią *-słabą* na przestrzeni sprzężonej X^* nazywamy najśłabszą topologię, przy której wszystkie funkcjonały $X^* \ni f \mapsto f(x) \in \mathbb{F}$ dla $x \in X$ są ciągłe na X^* (są to dokładnie funkcjonały z $i(X) \subseteq X^{**}$). Powiemy, że ciąg $\{f_n\}_{n=1}^{\infty} \subseteq X^*$ jest **-słabo zbieżny* do $f_0 \in X^*$, jeżeli jest zbieżny do f_0 w *-słabej topologii. Piszemy wtedy, $f_n \xrightarrow{w^*} f_0$.

Uwaga 3.41. Na mocy Stwierdzenia 3.31 bazą słabej topologii są zbiory postaci $U_{x_1, \dots, x_n, \varepsilon}(f) := \{g \in X^* : |g(x_i) - f(x_i)| < \varepsilon, 1 \leq i \leq n\}$ gdzie $f \in X^*$, $x_1, \dots, x_n \in X$, $\varepsilon > 0$. Natomiast ciąg funkcjonałów $\{f_n\}_{n=1}^{\infty} \subseteq X^*$ jest *-słabo zbieżny do $f_0 \in X^*$, wtedy i tylko wtedy, gdy jest zbieżny do f_0 punktowo, tzn.

$$f_n \xrightarrow{w^*} f_0 \iff \forall_{x \in X} f_n(x) \longrightarrow f_0(x).$$

Zatem oczywiste jest, że granica *-słabo zbieżnego ciągu jest wyznaczona jednoznacznie, oraz że zbieżność funkcjonałów w normie pociąga zbieżność *-słabą.

Jeśli przestrzeń X jest refleksywna, to możemy ją traktować jako przestrzeń sprzężoną do przestrzeni sprzężonej X^* . Jasne jest, że wtedy topologia $*$ -słaba na $(X^*)^* = X^{**} = X$ zadana przez X^* pokrywa się z topologią słabą na X adaną przez X^* . Czyli dla przestrzeni refleksywnych w zasadzie pojęcia słabej topologii i $*$ -słabej topologii są tożsame. W szczególności, w przestrzeniach Hilberta oraz przestrzeniach L^p dla $p > 1$ nie ma sensu rozważać topologii $*$ -słabych, por. Przykłady 3.37, 3.38. Natomiast na przestrzeni miar już tak:

Przykład 3.42 ($*$ -Słaba zbieżność w przestrzeni miar). Rozpatrzmy przestrzeń miar borelowskich $\mathcal{M}_{\mathbb{F}}(M)$ o wartościach w ciele $\mathbb{F}$, gdzie M jest lokalnie zwartą przestrzenią metryczną. Przestrzeń tę możemy traktować jako przestrzeń sprzężoną do przestrzeni $C_0(M)$. Wtedy dla ciągu miar $\{\mu_n\}_{n=1}^{\infty} \subseteq \mathcal{M}_{\mathbb{F}}(M)$ mamy

$$\mu_n \xrightarrow{w^*} \mu_0 \iff \forall_{x \in C_0(M)} \int_M x d\mu_n \longrightarrow \int_M x d\mu_0.$$

Założmy, że $M = \mathbb{R}$. Wtedy każda miara znakowa (ładunek) $\mu \in \mathcal{M}_{\mathbb{F}}(\mathbb{R})$ jest jednoznacznie wyznaczona przez swoją dystrybuantę $F_{\mu}(t) := \mu((-\infty, t])$, $t \in \mathbb{R}$. Zbieżność $*$ -słaba jest dobrze znaną z Rachunku Prawdopodobieństwa zbieżnością według rozkładu. Mianowicie, można pokazać (🏠 dla chętnych, patrz np. [1, Twierdzenie 8.3.1]), że

$$\mu_n \xrightarrow{w^*} \mu_0 \iff \forall_{t \in \mathbb{R}} \text{ jeśli } F_0 \text{ ciągła w } t, \text{ to } F_{\mu_n}(t) \rightarrow F_{\mu_0}(t).$$

Ogólnie topologia $*$ -słaba ma przewagę nad topologią słabą ze względu na następujące:

Twierdzenie 3.43 (Banach-Alaoglu). *Kula domknięta $\{f \in X^* : \|f\| \leq 1\}$ jest zwarta w $*$ -słabej topologii dla dowolnej przestrzeni unormowanej X .*

Dowód. Dla każdego $x \in X$ połóżmy $D_x := \{z \in \mathbb{F} : |z| \leq \|x\|\}$. Jest to podzbiór zwarty $\mathbb{F}$. Zatem na mocy Twierdzenia Tichonowa (nieskończony iloczyn kartezyjański

$$D := \prod_{x \in X} D_x = \{f : X \rightarrow \mathbb{F} : f(x) \in D_x\} = \{f : X \rightarrow \mathbb{F} : |f(x)| \leq \|x\|\}$$

jest przestrzenią zwartą. Przestrzeń ta zawiera kulę $B := \{f \in X^* : \|f\| \leq 1\}$, gdyż dla $f \in B$ i $x \in X$ mamy $|f(x)| \leq \|x\|$. Aby wykazać tezę wystarczy zauważyć, że: (a) topologia na B indukowana z topologii produktowej D pokrywa się ze $*$ -słabą topologią dziedziczną z X^* ; (b) B jest domkniętym podzbiorem D . Wtedy B jako domknięty podzbiór zbioru zwartego sam jest zwarty.

(a) wynika stąd, że topologia produktowa na D jest najslabszą topologią, przy której rzuty $\pi_x : D \rightarrow D_x$, gdzie $\pi_x(f) = f(x)$, dla $x \in X$ są ciągłe.

(b) wynika z Uwagi 3.32 oraz stąd, że jeżeli $\{f_i\}_{i \in I} \subseteq B$ jest siecią zbieżną do f_0 w D , to dla każdego $x_1, x_2 \in X$ i $\lambda_1, \lambda_2 \in \mathbb{F}$ mamy

$$f_0(\lambda_1 x_1 + \lambda_2 x_2) = \lim_{i \in I} f_i(\lambda_1 x_1 + \lambda_2 x_2) = \lim_{i \in I} \lambda_1 f_i(x_1) + \lambda_2 f_i(x_2) = \lambda_1 f_0(x_1) + \lambda_2 f_0(x_2).$$

Czyli $f_0 \in B$. □

Wniosek 3.44. *W przestrzeni refleksywnej X kula domknięta $\{x \in X : \|x\| \leq 1\}$ jest zbiorem słabo zwartym.*

Dowód. Słabą topologię na $X = (X^*)^*$ możemy traktować jako $*$ -słabą topologię zadaną przez X^* . Zatem teza wynika z Twierdzenia 3.43. □

Powyższe twierdzenie i wniosek mają szereg zastosowań. Wynika z nich na przykład, że każdy ciąg zawarty w domkniętym zbiorze ograniczonym posiada w tym zbiorze *-słabą albo słabą granicę. Umożliwia to konstruowanie funkcjonałów i elementów o żądanych własnościach. Ponadto, Wniosek 3.44 da się odwrócić, co daje elegancką charakteryzację przestrzeni refleksywnych, patrz [2, Twierdzenie 23.9]

Twierdzenie 3.45. *Przestrzeń Banacha X jest refleksywna wtedy i tylko wtedy, gdy kula domknięta $\{x \in X : \|x\| \leq 1\}$ jest zbiorem słabo zwartym.*

3.5 Twierdzenie Banacha o odwzorowaniu otwartym

Przypomnijmy, że odwzorowanie $f : X \rightarrow Y$ między przestrzeniami topologicznymi X i Y jest ciągłe, jeżeli dla każdego zbioru otwartego V w Y przeciwobraz $f^{-1}(V)$ jest zbiorem otwartym w X . Definicja odwzorowania otwartego jest podobna, ale “idzie w przeciwnym kierunku”.

Definicja 3.46. Odwzorowanie $f : X \rightarrow Y$ między przestrzeniami topologicznymi X i Y nazywamy *odwzorowaniem otwartym*, jeżeli obraz $f(U)$ każdego zbioru otwartego U w X jest zbiorem otwartym w Y . Czyli

$$f : X \rightarrow Y \text{ odwzorowanie otwarte} \stackrel{\text{def}}{\iff} \forall U \subseteq X \text{ otwarty } f(U) \text{ jest otwarty w } Y.$$

Uwaga 3.47. Dla dowolnych przestrzeni topologicznych X i Y mamy

- (a) Odwracalne odwzorowanie $f : X \rightarrow Y$ jest otwarte wtedy i tylko wtedy, gdy odwzorowanie odwrotne $f^{-1} : Y \rightarrow X$ jest ciągłe.
- (b) Ciągła bijekcja $f : X \rightarrow Y$ jest homeomorfizmem wtedy i tylko wtedy, gdy jest odwzorowaniem otwartym. W szczególności, istnieją odwracalne odwzorowania ciągłe niebędące otwartymi, np. $f : [0, 1) \rightarrow S^1 = \{z \in \mathbb{C} : |z| = 1\}$, gdzie $f(t) = e^{2\pi it}$, $t \in \mathbb{R}$.

Przykład 3.48. Funkcja kwadratowa $f(x) = x^2$, gdzie $f : \mathbb{R} \rightarrow \mathbb{R}$, nie jest otwarta, bo $f(\mathbb{R}) = [0, +\infty)$ nie jest zbiorem otwartym w $\mathbb{R}$. Staje się ona odwzorowaniem otwartym jeśli jej przeciwdziedzinę ograniczymy do obrazu, tzn. jeśli traktujemy f jako funkcję $f : \mathbb{R} \rightarrow [0, +\infty)$.

Dla odwzorowań liniowych na przestrzeniach unormowanych otwartość implikuje surjektywność. Co więcej, żeby stwierdzić czy odwzorowanie liniowe jest otwarte wystarczy sprawdzić, czy obraz kuli o środku w zerze zawiera kulę o środku w zerze.

Stwierdzenie 3.49. *Oznaczmy przez $K_X := \{x \in X : \|x\| < 1\}$ i $K_Y := \{y \in Y : \|y\| < 1\}$ otwarte kule jednostkowe w przestrzeniach unormowanych X i Y .*

$$\text{Odwzorowanie liniowe } T : X \rightarrow Y \text{ jest otwarte} \iff \exists_{r>0} rK_Y \subseteq T(K_X).$$

Ponadto, otwarte odwzorowanie liniowe $T : X \rightarrow Y$ musi być surjekcją.

Dowód. Zauważmy, że kulę o środku w $x_0 \in X$ (odpowiednio $y_0 \in Y$) i promieniu $r > 0$ można zapisać jako $x_0 + rK_X$ (odpowiednio $y_0 + rK_Y$).

“ $\implies$ ” Jeśli T jest odwzorowaniem otwartym, to $T(K_X)$ jest zbiorem otwartym. Zatem skoro $0 \in T(K_X)$, to istnieje $r > 0$ takie, że $rK_Y \subseteq T(K_X)$. Ponadto wtedy

$$Y = \bigcup_{n=1}^{\infty} nrK_Y \subseteq \bigcup_{n=1}^{\infty} nT(K_X) = T\left(\bigcup_{n=1}^{\infty} nK_X\right) = T(X).$$

Czyli $T(X) = Y$.

“ $\Leftarrow$ ” Załóżmy, że $rK_Y \subseteq T(K_X)$ dla pewnego $r > 0$. Niech U będzie dowolnym zbiorem otwartym w X . Potrzebujemy pokazać, że $T(U)$ jest zbiorem otwartym. Niech $y \in T(U)$ i niech $x \in U$ taki, że $Tx = y$. Skoro U otwarty, to istnieje kula o środku w x zawarta w U , czyli istnieje $\delta > 0$ taka, że $x + \delta K_X \subseteq U$. Obraz tej kuli

$$T(x + \delta K_X) = Tx + \delta T(K_X) = y + \delta T(K_X)$$

zawiera kulę $y + \delta r K_Y$ i zawiera się w $T(U)$. Czyli $T(U)$ wraz z każdym punktem zawiera pewne jego otoczenie. Zatem $T(U)$ jest zbiorem otwartym. $\square$

Okazuje się, że implikację z ostatniej części Stwierdzenia 3.49 da się odwrócić przy założeniu, że $T : X \rightarrow Y$ jest ograniczonym operatorem liniowym oraz X i Y są przestrzeniami Banacha. Jest to kolejny ważny fakt, obok Twierdzenia Banacha-Steinhaus (Twierdzenie 3.28), którego dowód opiera się na Twierdzeniu Baire’a (Twierdzenie 3.24):

Twierdzenie 3.50 (Twierdzenie Banacha o odwzorowaniu otwartym). *Niech $T \in B(X, Y)$ będzie ograniczonym operatorem liniowym, gdzie X i Y przestrzenie Banacha. Wtedy*

$$T \text{ jest surjekcją} \iff T \text{ jest odwzorowaniem otwartym.}$$

Dowód. Implikacja “ $\Leftarrow$ ” wynika ze Stwierdzenia 3.49. Żeby dowieść implikację “ $\Rightarrow$ ” przyjmujemy oznaczenia ze Stwierdzenia 3.49 dotyczące kul. Załóżmy, że T jest surjekcją. Wtedy

$$Y = T(X) = T\left(\bigcup_{n=1}^{\infty} nK_X\right) = \bigcup_{n=1}^{\infty} T(nK_X).$$

Zatem z Twierdzenia Baire’a (wykorzystujemy zupełność Y) któryś ze zbiorów $T(nK_X)$ nie może być nigdziegęsty (wnętrze jego domknięcia jest niepuste). Innymi słowy, istnieje $n \in \mathbb{N}$ takie, że

$$\text{Int}(\overline{T(nK_X)}) \neq \emptyset.$$

Czyli istnieje $y_0 \in Y$ oraz $\varepsilon > 0$ takie, że $y_0 + \varepsilon K_Y \subseteq \overline{T(nK_X)}$. Skoro $T(X) = Y$ to istnieje $x_0 \in X$ taki, że $Tx_0 = y_0$. Zatem

$$\begin{aligned} \varepsilon K_Y &\subseteq \overline{T(nK_X)} - y_0 = \overline{T(nK_X)} - T(x_0) = \overline{T(nK_X - x_0)} \\ &\subseteq \overline{T((n + \|x_0\|)K_X)} = (n + \|x_0\|)\overline{T(K_X)}, \end{aligned}$$

gdzie w drugiej i trzeciej równości wykorzystaliśmy liniowość T i ciągłość operacji liniowych. Dzieląc powyższą inkluzję stronami przez $n + \|x_0\|$ i kładąc $r := \frac{\varepsilon}{n + \|x_0\|}$ otrzymujemy

$$rK_Y \subseteq \overline{T(K_X)}. \quad (3.4)$$

Z dokładnością do domknięcia jest to warunek ze Stwierdzenia 3.49. W celu “pozbycia się domknięcia” po prawej stronie inkluzji (3.4) pokażemy, że

$$\overline{T(K_X)} \subseteq T(2K_X). \quad (3.5)$$

Niech $y \in \overline{T(K_X)}$. Wtedy istnieje $x_1 \in K_X$ taki, że $\|y - Tx_1\| < \frac{r}{2}$. Stąd

$$y - Tx_1 \in \frac{r}{2}K_Y \stackrel{(3.4)}{\subseteq} \frac{1}{2}\overline{T(K_X)} = \overline{T\left(\frac{1}{2}K_X\right)}.$$

Stosując to samo rozumowanie do $y - Tx_1 \in \overline{T\left(\frac{1}{2}K_X\right)}$ możemy znaleźć $x_2 \in \frac{1}{2}K_X$ taki, że $\|(y - Tx_1) - Tx_2\| < \frac{r}{4}$ i stąd

$$y - T(x_1 + x_2) = (y - Tx_1) - Tx_2 \in \frac{r}{4}K_Y \subseteq \overline{T\left(\frac{1}{4}K_X\right)}.$$

Kontynuując w ten sposób otrzymujemy ciąg $\{x_n\}_{n=1}^\infty \subseteq X$ taki, że dla każdego $n \in \mathbb{N}$

$$x_n \in \frac{1}{2^{n-1}}K_X \quad \text{oraz} \quad y - T(x_1 + \dots + x_n) \in \frac{r}{2^n}K_Y.$$

Z tej drugiej relacji wynika, że $T(x_1 + \dots + x_n) \rightarrow y$ w Y . Natomiast pierwsza relacja (oraz zupełność X) gwarantuje, że szereg $\sum_{n=1}^\infty x_n$ jest zbieżny w X . Rzeczywiście, ciąg sum częściowych $S_n := x_1 + x_2 + \dots + x_n$ jest ciągiem Cauchy w X , bo dla $m > n$ mamy

$$\|S_m - S_n\| = \|x_{n+1} + \dots + x_m\| < \sum_{k=n+1}^\infty \|x_k\| < \sum_{k=n+1}^\infty \frac{1}{2^{k-1}} = \frac{1}{2^n} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{1}{2^{n-1}} \xrightarrow{n \rightarrow \infty} 0.$$

Zatem z zupełności X wynika, że granica $\lim_{n \rightarrow \infty} S_n \in X$ istnieje. Z definicji taką granicę nazywa się sumą szeregu i oznacza $\sum_{n=1}^\infty x_n$. Zauważmy, że

$$\left\| \sum_{n=1}^\infty x_n \right\| \leq \sum_{n=1}^\infty \|x_n\| < \sum_{n=1}^\infty \frac{1}{2^{n-1}} = 2.$$

Czyli $\sum_{n=1}^\infty x_n \in 2K_X$. Wykorzystując ciągłość (ograniczoność) operatora T otrzymujemy

$$T\left(\sum_{n=1}^\infty x_n\right) = T\left(\lim_{n \rightarrow \infty} x_1 + \dots + x_n\right) = \lim_{n \rightarrow \infty} T(x_1 + \dots + x_n) = y,$$

Czyli $y \in T(2K_X)$. To kończy dowód inkluzji (3.5).

W świetle inkluzji (3.4) i (3.5) dostajemy, że $rK_Y \subseteq T(2K_X)$ lub równoważnie $r/2K_Y \subseteq T(K_X)$. Zatem T jest odwzorowaniem otwartym na mocy Stwierdzenia 3.49. $\square$

Wniosek 3.51. *Jeśli T jest różnowartościowym operatorem ograniczonym z przestrzeni Banacha X na przestrzeń Banacha Y , to odwzorowanie $T^{-1} : Y \rightarrow X$ jest operatorem ograniczonym. Tzn.*

$$\left(\begin{array}{l} T \in B(X, Y) \text{ oraz } T \text{ bijekcja} \\ X, Y \text{ przestrzenie Banacha} \end{array} \right) \implies T^{-1} \in B(Y, X).$$

Dowód. Odwzorowanie $T^{-1} : Y \rightarrow X$ jest liniowe jako odwzorowanie odwrotne do odwzorowania liniowego $T : X \rightarrow Y$ jest liniowe. Skoro T jest surjekcją, to na mocy Twierdzenia 3.50 jest odwzorowaniem otwartym. Czyli dla każdego otwartego zbioru U w X zbiór $(T^{-1})^{-1}(U) = T(U)$ jest otwarty w Y . Zatem operator T^{-1} jest ciągły, czyli ograniczony. $\square$

Wniosek 3.52. *Każde dwie porównywalne normy zupełne na przestrzeni liniowej X są równoważne. Tzn.*

$$\left(\begin{array}{l} (X, \|\cdot\|_1), (X, \|\cdot\|_2) \text{ przestrzenie Banacha} \\ \text{norma } \|\cdot\|_1 \text{ jest słabsza, niż } \|\cdot\|_2 \end{array} \right) \implies \|\cdot\|_1 \text{ i } \|\cdot\|_2 \text{ są równoważne.}$$

Dowód. Przypomnijmy, że norma $\|\cdot\|_1$ jest słabsza, niż $\|\cdot\|_2$ jeżeli

$$\exists_{c_1 > 0} \forall_{x \in X} \|x\|_1 \leq c_1 \|x\|_2.$$

Jest to to samo co powiedzieć, że operator identycznościowy $id : (X, \|\cdot\|_2) \rightarrow (X, \|\cdot\|_1)$ jest ograniczony. Oczywiście id jest bijekcją, zatem na mocy Wniosku 3.51 operator odwrotny $id : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$ jest ograniczony. Czyli

$$\exists_{c_2 > 0} \forall_{x \in X} \|x\|_2 \leq c_2 \|x\|_1,$$

co oznacza, że norma $\|\cdot\|_2$ jest słabsza, niż $\|\cdot\|_1$. Zatem normy te są równoważne. $\square$

Założenie o zupełności X i Y w powyższych Twierdzeniach jest istotne!

Przykład 3.53. Na przestrzeni $X = L^2[0, 1]$, korzystając z nierówności Schwartz'a, mamy

$$\|x\|_1 = \int_0^1 1 \cdot |x(t)| dt \leq \left(\int_0^1 1 dt \right)^{\frac{1}{2}} \left(\int_0^1 |x(t)|^2 dt \right)^{\frac{1}{2}} = \|x\|_2.$$

Czyli norma $\|\cdot\|_1$ jest słabsza, niż norma $\|\cdot\|_2$. Normy $\|\cdot\|_1$ i $\|\cdot\|_2$ są porównywalne na X , ale nie są równoważne, bo X w normie $\|\cdot\|_2$ jest przestrzenią zupełną, a w normie $\|\cdot\|_1$ nie jest (domknięcie $L^2[0, 1] \subseteq L^1[0, 1]$ w normie $\|\cdot\|_1$ daje całą przestrzeń $L^1[0, 1]$). W szczególności, operator identycznościowy $id : (X, \|\cdot\|_2) \rightarrow (X, \|\cdot\|_1)$ jest odwracalnym operatorem ograniczonym, ale nie jest otwarty i operator do niego odwrotny nie jest ograniczony.

Jako kolejny ważny wniosek z Twierdzenia Banacha o odwzorowaniu otwartym otrzymujemy Twierdzenie Banacha o wykresie domkniętym.

Definicja 3.54. Wykresem funkcji $f : X \rightarrow Y$ nazywamy zbiór

$$\Gamma(f) := \{(x, f(x)) : x \in X\} \subseteq X \times Y.$$

Lemat 3.55. Wykres funkcji ciągłej między przestrzeniami metrycznymi jest domknięty:

$$\left(\begin{array}{l} f : X \rightarrow Y \text{ funkcja ciągła} \\ X, Y \text{ przestrzenie metryczne} \end{array} \right) \implies \Gamma(f) \text{ domknięty w } X \times Y.$$

Dowód. Jeśli (x_0, y_0) jest punktem skupienia wykresu $\Gamma(f)$, to istnieje ciąg $(x_n, y_n) \in \Gamma(f)$ taki, że $(x_n, y_n) \rightarrow (x_0, y_0)$. Ale z definicji wykresu $y_n = f(x_n)$. Natomiast z ciągłości funkcji

$$f(x_0) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} y_n = y_0.$$

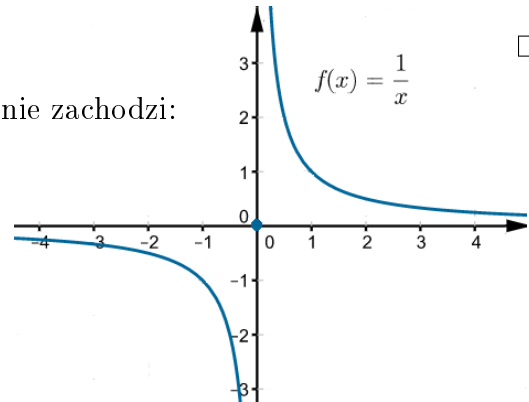
Czyli $(x_0, y_0) \in \Gamma(f)$. $\square$

Na ogół implikacja odwrotna do tej w Lemacie 3.55 nie zachodzi:

$$\text{Niech } X = Y = \mathbb{R} \text{ oraz } f(x) = \begin{cases} \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Funkcja f nie jest ciągła, a jej

wykres $\Gamma(f)$ jest domknięty w $\mathbb{R}^2$.



Banach wykazał, że dla operatorów liniowych między przestrzeniami Banacha, implikację z Lematu 3.55 można odwrócić.

Twierdzenie 3.56 (Twierdzenie Banacha o wykresie domkniętym). *Operator liniowy $T : X \rightarrow Y$ między dwiema przestrzeniami Banacha X i Y jest ograniczony (ciągły) wtedy i tylko wtedy, gdy ma domknięty wykres.*

Dowód. Na mocy Lematu 3.55 trzeba jedynie pokazać, że operator liniowy z domkniętym wykresem jest ograniczony. Załóżmy więc, że wykres $\Gamma(T)$ jest domknięty w $X \times Y$. Zauważmy, że

- 1) $X \times Y$ jest przestrzenią Banacha z normą $\|(x, y)\|_{X \times Y} := \|x\|_X + \|y\|_Y$ (🏠 sprawdzić);
- 2) $\Gamma(T)$ jest domkniętą podprzestrzenią liniową w $X \times Y$ (🏠 sprawdzić).

Zatem $\Gamma(T)$ jest przestrzenią Banacha z normą $\|\cdot\|_{X \times Y}$. Rzuty

$$P_1 : \Gamma(T) \rightarrow X, \quad \text{gdzie} \quad P_1(x, Tx) = x,$$

$$P_2 : \Gamma(T) \rightarrow Y, \quad \text{gdzie} \quad P_2(x, Tx) = Tx,$$

są liniowe i ograniczone, bo

$$\|P_1(x, Tx)\|_X = \|x\|_X \leq \|x\|_X + \|Tx\|_Y = \|(x, Tx)\|_{X \times Y} \implies \|P_1\| \leq 1,$$

$$\|P_2(x, Tx)\|_Y = \|Tx\|_Y \leq \|x\|_X + \|Tx\|_Y = \|(x, Tx)\|_{X \times Y} \implies \|P_2\| \leq 1.$$

Zauważmy, że operator P_1 jest odwracalny. Zatem na mocy Twierdzenia Banacha o odwzorowaniu otwartym (patrz Wniosek 3.51) operator odwrotny

$$P_1^{-1} : X \rightarrow \Gamma(T), \quad \text{gdzie} \quad P_1^{-1}(x) = (x, Tx)$$

jest operatorem ograniczonym. Stąd operator

$$T = P_2 \circ P_1^{-1}$$

jest ograniczony jako złożenie operatorów ograniczonych. □

Uwaga 3.57. Ograniczoność (ciągłość) operatora liniowego $T : X \rightarrow Y$ oznacza, że $x_n \rightarrow x$ implikuje $Tx_n \rightarrow Tx$. Natomiast domkniętość wykresu oznacza, że $x_n \rightarrow x$ oraz $Tx_n \rightarrow y$ implikuje $Tx = y$. Znaczenie twierdzenia o wykresie domkniętym polega na tym, że sprawdzając, czy $x_n \rightarrow x$ implikuje $Tx_n \rightarrow Tx$ możemy założyć, że Tx_n zbiega do czegoś i trzeba tylko pokazać, że granica jest odpowiednia.

Rozdział 4

Teoria spektralna

W tym rozdziale omówimy pojęcie widma (spektrum) operatora ograniczonego. W przypadku przestrzeni nad ciałem liczb rzeczywistych teoria spektralna nie ma dobrych własności (np. widmo operatora może być puste). Ponadto współcześnie teorię spektralną rozwija się głównie w języku algebr Banacha, gdzie na operatory ograniczone na konkretnej przestrzeni Banacha patrzy się jak na elementy algebry, którą te operatory tworzą (zapominając o przestrzeni na której działają). Dlatego w tym rozdziale będziemy głównie rozpatrywać przestrzenie zespolone i zespolone algebry Banacha.

4.1 Algebry Banacha i elementy odwracalne

Niech X będzie przestrzenią Banacha nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$. Jak wiemy (patrz Twierdzenie 1.73), przestrzeń ograniczonych operatorów liniowych $B(X)$ wraz z działaniami określonymi punktowo i normą operatorową jest przestrzenią Banacha. Zauważmy, że $B(X)$ jest nie tylko przestrzenią liniową, ale też algebrą z mnożeniem zdefiniowanym jako złożenie operatorów. Rzeczywiście, złożenie operatorów ograniczonych jest operatorem ograniczonym i norma złożenia jest nie większa, niż iloczyn składników:

$$T, S \in B(X) \implies T \circ S \in B(X) \text{ oraz } \|T \circ S\| \leq \|T\| \cdot \|S\|$$

bo $\|T \circ S\| = \sup_{\|x\|=1} \|T(Sx)\| \leq \sup_{\|x\|=1} \|T\| \|Sx\| = \|T\| \cdot \|S\|$. Co więcej operator identycznościowy $1 \in B(X)$ jest jedyneką w tej algebrze: $1 \circ T = T \circ 1 = T$, oraz $\|1\| = 1$. Własności te są podstawowymi elementami definicji algebr Banacha:

Definicja 4.1. *Algebrą* nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$ nazywamy przestrzeń liniową A nad $\mathbb{F}$ wyposażoną dodatkowo w operację mnożenia $\cdot : A \times A \rightarrow A$, która jest łączna, rozdzielna względem dodawania i łączna ze względu na mnożenie skalarne, tzn.:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c), \quad (a + b) \cdot c = a \cdot c + b \cdot c, \quad a \cdot (b + c) = a \cdot b + a \cdot c, \quad (\lambda a) \cdot b = a \cdot (\lambda b) = \lambda(a \cdot b),$$

dla dowolnych $a, b, c \in A$, $\lambda \in \mathbb{F}$. Innymi słowy, algebra to jednocześnie przestrzeń liniowa oraz pierścień łączny, gdzie struktury te są ze sobą zgodne:

$$\text{algebra} \equiv \left(\begin{array}{l} \text{przestrzeń liniowa} + \\ \text{pierścień} + \text{łączność} \end{array} \right).$$

Uwaga 4.2. Zwyczajowo znak mnożenia w algebrze będziemy opuszczać.

Definicja 4.3. Algebrą Banacha nad ciałem $\mathbb{F} = \mathbb{R}, \mathbb{C}$ nazywamy przestrzeń Banacha A , która jest jednocześnie algebrą nad $\mathbb{F}$ oraz norma jest submultiplikatywna, tzn.

$$\|a \cdot b\| \leq \|a\| \cdot \|b\|, \quad a, b \in A.$$

Jeśli dodatkowo mnożenie w algebrze A posiada element neutralny, to element ten nazywamy jedyneką algebry i oznaczamy przez 1. Zwyczajowo zakładamy wtedy też, że $\|1\| = 1$.

Uwaga 4.4. Submultiplikatywność normy implikuje, że mnożenie $\cdot : A \times A \rightarrow A$ w algebrze Banacha jest ciągłe.

Przykład 4.5 (Algebry operatorów ograniczonych). Przestrzeń liniowych operatorów ograniczonych $B(X)$, na dowolnej przestrzeni Banacha X , jest algebrą Banacha z mnożeniem zdefiniowanym jako złożenie operatorów. Algebra $B(X)$ jest nieprzemieniana, o ile tylko $\dim(X) > 1$. W szczególności, algebrą Banacha jest każda domknięta podalgebra $A \subseteq B(X)$. Jak wykażemy poniżej (Stwierdzenie 4.12) każda algebra Banacha, z dokładnością do izomorfizmu, jest tej postaci.

Przykład 4.6 (Algebry funkcji ciągłych). Przestrzeń funkcji ciągłych i znikających w nieskończoności $C_0(M)$ na przestrzeni topologicznej M , jest algebrą Banacha z mnożeniem zdefiniowanym punktowo

$$(a \cdot b)(t) := a(t)b(t), \quad a, b \in C_0(M).$$

Zauważmy, że tak zdefiniowane mnożenie jest przemienne, tzn. $a \cdot b = b \cdot a$ dla dowolnych $a, b \in C_0(M)$. Innymi słowy, algebra $C_0(M)$ jest *przemieniana*. Zauważmy, że algebra $C_0(M)$ ma jedynekę wtedy i tylko wtedy, gdy przestrzeń M jest zwarta i wtedy jedyneką jest funkcja tożsamościowa równa jeden (🏠 pokazać).

Przykład 4.7 (Algebra splotowa $L^1(\mathbb{R})$). Przestrzeń $L^1(\mathbb{R})$ funkcji całkowalnych względem miary Lebesgue'a na prostej $\mathbb{R}$ jest algebrą Banacha wraz z *mnożeniem splotowym*:

$$(f \star g)(t) := \int_{-\infty}^{\infty} f(t-s)g(s) ds, \quad t \in \mathbb{R}, f, g \in L^1(\mathbb{R}).$$

Na przykład, dla $f, g \in L^1(\mathbb{R})$ mamy

$$\begin{aligned} \|f \star g\|_1 &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} f(t-s)g(s) ds \right| dt \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(t-s)g(s)| ds dt \left\{ \begin{array}{l} \text{zmiana kolejności} \\ \text{całkowania} \end{array} \right\} \\ &= \int_{-\infty}^{\infty} |g(s)| \int_{-\infty}^{\infty} |f(t-s)| dt ds = \int_{-\infty}^{\infty} |g(s)| \|f\|_1 ds \\ &= \|g\|_1 \|f\|_1. \end{aligned}$$

Zatem iloczyn jest poprawnie określony i norma w $L^1(\mathbb{R})$ jest dla mnożenia splotowego submultiplikatywna. Algebra $L^1(\mathbb{R})$ z mnożeniem splotowym jest przemieniana i nie posiada jedynki.

Uwaga 4.8. Przypomnijmy z rachunku prawdopodobieństwa, że jeżeli f i g są gęstościami rozkładów niezależnych zmiennych losowych X i Y , to splot $f \star g$ jest gęstością rozkładu sumy $X + Y$ tych zmiennych.

Przykład 4.9 (Algebry grupowe). Niech G będzie grupą dyskretną (bez topologii). Przestrzeń $\ell^1(G)$ funkcji sumowalnych na G jest algebrą Banacha wraz z *mnożeniem splotowym*:

$$(a \star b)(g) := \sum_{h \in G} a(gh^{-1})b(h), \quad g \in G, a, b \in \ell^1(G).$$

Analogicznie jak w Przykładzie 4.7 można pokazać, że splot jest poprawnie określony i norma w $\ell^1(G)$ jest ze względu nań submultiplikatywna (🏠 pokazać). Algebra ta ma jedynekę $1 : G \rightarrow \mathbb{F}$, która przyjmuje wartość 1 w elemencie neutralnym $e \in G$ grupy G , a w pozostałych punktach się zeruje. Algebra $\ell^1(G)$ jest przemienna wtedy i tylko wtedy, gdy grupa G jest przemienna.

Przykład 4.10 (Algebra z mnożeniem zerowym). Na dowolnej przestrzeni Banacha X można zdefiniować mnożenie wzorem $x \cdot y := 0$ dla wszystkich $x, y \in X$. Wtedy X jest algebrą Banacha. Jest to algebra przemienna i jest to skrajny przypadek algebry bez jedynki.

Do każdej algebry Banacha można w kanoniczny sposób dodać jedynekę - proces taki nazywa się czasem “ujedynkowieniem”. Pozwala to sprowadzić większość rozważań do algebr z jedyneką. *Homomorfizm algebr* to odwzorowanie liniowe i multiplikatywne, tzn. zachowujące mnożenie elementów.

Lemat 4.11 (Ujedynkowanie algebry Banacha). *Niech A będzie algebrą Banacha (z jedyneką lub bez) nad ciałem $\mathbb{F}$. Wtedy suma prosta $\tilde{A} := A \oplus \mathbb{F}$ wraz z mnożeniem i normą*

$$(a, \lambda) \cdot (b, \mu) := (ab + \lambda b + \mu a, \lambda\mu), \quad \|(a, \lambda)\| := \|a\| + |\lambda|,$$

$a, b \in A$, $\lambda, \mu \in \mathbb{F}$, jest algebrą Banacha z jedyneką daną przez element $(0, 1)$ oraz taką, że przyporządkowanie $A \ni a \mapsto (a, 0) \in \tilde{A}$ jest izometrycznym homomorfizmem algebr.

Dowód. Dowód składa się z prostych rachunków 🏠 . □

Każdą algebrę Banacha A można traktować jako algebrę operatorów na pewnej przestrzeni Banacha. Dlatego też elementy algebry Banacha czasem nazywa się operatorami.

Stwierdzenie 4.12. *Dla dowolnej algebry Banacha A istnieje izometryczny homomorfizm $\pi : A \rightarrow B(X)$, gdzie X jest pewną przestrzenią Banacha. Czyli A zanurza się jako domknięta podalgebra w algebrę operatorów ograniczonych $B(X)$ na przestrzeni Banacha X .*

Dowód. Przechodząc ewentualnie do ujedynkowania A , patrz Lemat 4.11, możemy założyć, że A jest algebrą z jedyneką $1 \in A$. Rozważmy $X := A$ jako przestrzeń Banacha. Korzystając z mnożenia w A możemy zdefiniować π wzorem

$$\pi(a)x := a \cdot x, \quad a \in A, x \in X = A.$$

Z rozdzielnosci mnożenia względem dodawania i łączności wynika od razu, że $\pi(a)$ jest operatorem liniowym na X . Ponadto korzystając z submultiplikatywności normy w A mamy

$$\|\pi(a)\| = \sup_{\|x\|=1} \|\pi(a)x\| = \sup_{\|x\|=1} \|a \cdot x\| \leq \sup_{\|x\|=1} \|a\|\|x\| = \|a\|.$$

Czyli $\|\pi(a)\| \leq \|a\|$. Z drugiej strony skoro $\|1\| = 1$, to $\|\pi(a)\| \geq \|\pi(a)1\| = \|a \cdot 1\| = \|a\|$. Zatem $\|\pi(a)\| = \|a\|$. Stąd $\pi : A \rightarrow B(X)$ jest poprawnie zdefiniowaną izometrią. Z rozdzielnosci mnożenia względem dodawania i łączności wynika wprost, że π jest homomorfizmem algebr. □

Od tej pory będziemy zakładać, że

rozważane przez nas algebry Banacha mają jedynekę,

gdyż będą nas interesować elementy odwracalne.

Definicja 4.13. Niech A będzie algebrą Banacha z jedyneką $1 \in A$. Zbiór elementów odwracalnych w A oznaczamy przez

$$\text{Inv}(A) := \{a \in A : \exists_{b \in A} ab = ba = 1\}$$

Jeżeli $a \in \text{Inv}(A)$ i $b \in A$ spełniają $ab = ba = 1$ to piszemy $b = a^{-1}$ i nazywamy go *elementem odwrotnym* do a (element odwrotny jest jednoznacznie wyznaczony przez a).

Uwaga 4.14. $\text{Inv}(A)$ tworzy grupę, której elementem neutralnym jest $1 \in A$. W szczególności

$$a, b \in \text{Inv}(A) \implies \begin{cases} a^{-1} \in \text{Inv}(A) \text{ oraz } (a^{-1})^{-1} = a, \\ ab \in \text{Inv}(A) \text{ oraz } (ab)^{-1} = b^{-1}a^{-1}. \end{cases}$$

Przykład 4.15. Jeśli $A = B(X)$, gdzie X to przestrzeń Banacha, to

$$\text{Inv}(B(X)) = \{T \in B(X) : T : X \rightarrow X \text{ bijekcja}\}.$$

Rzeczywiście, każdy element T odwracalny w $B(X)$ jest funkcją odwracalną, bo złożenia $TT^{-1} = T^{-1}T = 1$ dają operator identycznościowy. Na odwrót jeżeli operator ograniczony $T \in B(X)$ jest odwzorowaniem odwracalnym, to odwzorowanie odwrotne T^{-1} jest operatorem ograniczonym na mocy Twierdzenia Banacha o odwzorowaniu otwartym, patrz Wniosek 3.51.

Przykład 4.16. Jeśli $A = C(M)$, gdzie M przestrzeń zwarta. Wtedy

$$\text{Inv}(C(M)) = \{a \in C(M) : a(t) \neq 0 \text{ dla każdego } t \in M\}.$$

Rzeczywiście, jedyneką w $C(M)$ jest funkcja stale równa jeden, $1(t) = 1$ dla $t \in M$. Zatem jeśli element $a \in C(M)$ jest odwracalny to $a(t)a^{-1}(t) = 1$ dla każdego $t \in M$, skąd $a(t) \neq 0$ dla $t \in M$. Na odwrót, jeśli $a(t) \neq 0$ dla $t \in M$, to kładąc $a^{-1}(t) := a(t)^{-1}$ dla $t \in M$ otrzymujemy funkcję ciągłą $a^{-1} \in C(M)$ oraz $aa^{-1} = 1$, więc $a \in \text{Inv}(C(M))$.

Dobrze znany ze szkoły wzór na sumę szeregu geometrycznego mówi, że dla każdej liczby $t \in \mathbb{F}$ takiej, że $|t| < 1$ zachodzi

$$\sum_{k=0}^{\infty} t^k = \frac{1}{1-t} = (1-t)^{-1}.$$

Uogólnienie tego wzoru na elementy algebr Banacha jest podstawowym faktem, za pomocą którego można wywnioskować najważniejsze własności zbioru elementów odwracalnych i operacji odwracania. W literaturze fakt ten nazywa się *Lematem Neumanna*.¹

Potęgi elementu a algebry A definiujemy w oczywisty sposób wykorzystując mnożenie z A , tzn. $a^n := \underbrace{a \cdot \dots \cdot a}_{n \text{ razy}}$ dla $n > 1$ oraz $a^0 := 1$.

Lemat 4.17 (Lemat Neumanna). Niech A będzie algebrą Banacha z jedyneką $1 \in A$.

$$\forall_{a \in A} \|a\| < 1 \implies 1 - a \in \text{Inv}(A) \text{ oraz } (1 - a)^{-1} = \sum_{k=0}^{\infty} a^k.$$

¹Chodzi tu o Carla Neumanna (1832–1925), a nie Johna von Neumanna (1903–1957)

Dowód. Niech $\|a\| < 1$ i połóżmy $S_n = \sum_{k=0}^n a^k$. Dla $m > n$ mamy

$$\|S_m - S_n\| = \left\| \sum_{k=n+1}^m a^k \right\| \leq \sum_{k=n+1}^m \|a^k\| \leq \sum_{k=n+1}^{\infty} \|a\|^k = \frac{\|a\|^{n+1}}{1 - \|a\|} \xrightarrow{n \rightarrow \infty} 0.$$

Czyli ciąg sum częściowych $\{S_n\}_{n=1}^{\infty}$ jest ciągiem Cauchy w A , a zatem jest on zbieżny. Innymi słowy, szereg $\sum_{k=0}^{\infty} a^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n a^k \in A$ jest zbieżny. Zauważmy dalej, że skoro $\|a^n\| \leq \|a\|^n \rightarrow 0$, to $a^n \rightarrow 0$ i stąd

$$(1 - a) \sum_{k=0}^{\infty} a^k = \lim_{n \rightarrow \infty} (1 - a) \sum_{k=0}^n a^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n a^k - \sum_{k=0}^n a^{k+1} = \lim_{n \rightarrow \infty} 1 - a^{n+1} = 1.$$

Analogicznie otrzymujemy $\sum_{k=0}^{\infty} a^k (1 - a) = 1$. Stąd $(1 - a)^{-1} = \sum_{k=0}^{\infty} a^k$. $\square$

Wniosek 4.18. *Elementy oddalone od $1 \in A$ o mniej niż jeden są odwracalne:*

$$\underbrace{\{a \in A : \|1 - a\| < 1\}}_{\substack{\text{kula jednostkowa o} \\ \text{środku w } 1 \in A}} \subseteq \text{Inv}(A).$$

Dowód. Jeśli $\|1 - a\| < 1$, to $a = 1 - (1 - a) \in \text{Inv}(A)$ na mocy Lematu Neumanna. $\square$

Twierdzenie 4.19. *Dla każdej algebry Banacha A z jedynek $1 \in A$ zbiór elementów odwracalnych $\text{Inv}(A)$ jest otwarty w A . Z Lematy Ponadto, odwzorowanie $\text{Inv}(A) \ni a \mapsto a^{-1} \in \text{Inv}(A)$ jest różniczkowalne (a więc jest dyfeomorfizmem, bo jest samo do siebie odwrotne).*

Dowód. Aby wykazać, że $\text{Inv}(A)$ jest zbiorem otwartym trzeba pokazać, że dla każdego $a \in \text{Inv}(A)$ istnieje $\varepsilon > 0$ taki, że

$$K(a, \varepsilon) = \{b \in A : \|a - b\| < \varepsilon\} \subseteq \text{Inv}(A). \quad (4.1)$$

Z Lematu Neumanna wiemy, że $K(1, 1) \subseteq \text{Inv}(A)$ i skoro iloczyn elementów odwracalnych jest odwracalny to $a \cdot K(1, 1) \subseteq \text{Inv}(A)$. Ponadto twierdzimy, że

$$K(a, \|a^{-1}\|^{-1}) \subseteq a \cdot K(1, 1)$$

co dowodzi (4.1) z $\varepsilon := \|a^{-1}\|^{-1}$. Rzeczywiście, jeśli $b \in K(a, \|a^{-1}\|^{-1})$, to

$$b = a - (a - b) = a(1 - a^{-1}(a - b))$$

oraz

$$\|a^{-1}(a - b)\| \leq \|a^{-1}\| \cdot \|a - b\| < \|a^{-1}\| \cdot \|a^{-1}\|^{-1} = 1.$$

Czyli $b = a(1 - a^{-1}(a - b)) \in aK(1, 1) \subseteq \text{Inv}(A)$. Dokładniej, z Lematu 4.17 wynika, że

$$[1 - a^{-1}(a - b)]^{-1} = \sum_{k=0}^{\infty} [a^{-1}(a - b)]^k \quad \text{oraz} \quad b^{-1} = \sum_{k=0}^{\infty} [a^{-1}(a - b)]^k a^{-1}.$$

Zatem $\text{Inv}(A)$ jest zbiorem otwartym.

Rozważmy teraz odwzorowanie $F : \text{Inv}(A) \rightarrow \text{Inv}(A)$, gdzie $F(a) := a^{-1}$. Zauważmy, że jest to odwzorowanie odwracalne i $F^{-1} = F$, bo $F^2(a) = (a^{-1})^{-1} = a$ (innymi słowy F jest inwolucją).

Aby wykazać, że F jest różniczkowalne w punkcie $a \in \text{Inv}(A)$ trzeba znaleźć operator liniowy $F'(a) : A \rightarrow A$ taki, że

$$\lim_{b \rightarrow a} \frac{\|F(a) - F(b) - F'(a)(a - b)\|}{\|a - b\|} = 0.$$

Położmy, $F'(a)c := -a^{-1}ca^{-1}$, dla $c \in A$ (podpopowiedzią, może tu być przypadek jednowymiarowy, gdy $f(t) = t^{-1}$ wtedy $f'(t) = -t^{-2}$, czyli $f'(t)h = -t^{-2}h$). Wtedy $F'(a) : A \rightarrow A$ liniowy i ograniczony, $\|F'(a)\| \leq \|a^{-1}\|^2$. Dla $b \in K(a, \|a^{-1}\|^{-1})$, jak wyżej, mamy

$$\begin{aligned} \|F(a) - F(b) - F'(a)(a - b)\| &= \|a^{-1} - \sum_{k=0}^{\infty} [a^{-1}(a - b)]^k a^{-1} + a^{-1}(a - b)a^{-1}\| \\ &= \left\| \sum_{k=2}^{\infty} [a^{-1}(a - b)]^k a^{-1} \right\| \leq \sum_{k=2}^{\infty} \| [a^{-1}(a - b)]^k a^{-1} \| \\ &\leq \sum_{k=2}^{\infty} \|a^{-1}\|^{k+1} \|a - b\|^k \stackrel{\frac{\|a-b\|}{\|a^{-1}\|^{-1}} < 1}{=} \frac{\|a^{-1}\|^3 \|a - b\|^2}{1 - \|a^{-1}\| \|a - b\|}. \end{aligned}$$

Stąd

$$\lim_{b \rightarrow a} \frac{\|F(a) - F(b) - F'(a)(a - b)\|}{\|a - b\|} \leq \lim_{b \rightarrow a} \frac{\|a^{-1}\|^3 \|a - b\|}{1 - \|a^{-1}\| \|a - b\|} = 0.$$

□

4.2 Widmo i promień spektralny

W całym podrozdziale zakładamy, że A jest algebrą Banacha z jedyneką 1.

Definicja 4.20. *Widmem (spektrum) elementu $a \in A$ algebry Banacha z jedyneką $1 \in A$ nad ciałem $\mathbb{F}$ nazywamy następujący zbiór liczb*

$$\sigma(a) := \{\lambda \in \mathbb{F} : a - \lambda 1 \notin \text{Inv}(A)\}.$$

Przykład 4.21 (Widmo operatora działającego na przestrzeni Banacha). Jeśli $A = B(X)$, gdzie X to przestrzeń Banacha, to $\text{Inv}(B(X)) = \{T \in B(X) : T : X \rightarrow X \text{ bijekcja}\}$, patrz Przykład 4.15. Zatem *widmo operatora* $T \in B(X)$ wyraża się formułą

$$\sigma(T) := \{\lambda \in \mathbb{F} : T - \lambda 1 : X \rightarrow X \text{ odwzorowanie nieodwracalne}\}.$$

Czyli $\lambda \in \sigma(T)$ wtedy i tylko wtedy, gdy zachodzi któryś z warunków:

- a) $T - \lambda 1$ nie jest injekcją. To zachodzi wtedy i tylko wtedy, gdy $\ker(T - \lambda 1) \neq \{0\}$, tzn. gdy istnieje $x \in X$ różny od zera taki, że $Tx = \lambda x$, czyli λ jest *wartością własną* dla T .
- b) $T - \lambda 1$ nie jest surjekcją.

W przypadku, gdy przestrzeń X jest nieskończenie wymiarowa łatwo znaleźć operatory, które są różnowartościowe, ale nie surjektywne (np. operator przesunięcia, Przykład 2.79). Jeśli jednak $\dim(X) < \infty$, to z racji na równość $\dim(X) = \dim(T(X)) + \dim(\ker T)$, która zachodzi dla dowolnego $T \in B(X)$, powyższe warunki a) i b) są równoważne. Zatem

$$\dim(X) < \infty \implies \sigma(T) = \{\lambda \in \mathbb{F} : \ker(T - \lambda 1) \neq \{0\}\} - \text{zbiór wartości własnych}.$$

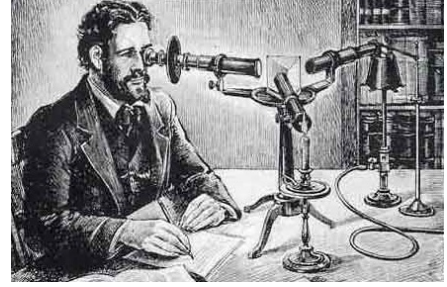
Przypomnijmy też, że jeśli $\dim(X) = n < \infty$, to ustalając bazę w X otrzymujemy izomorfizm algebr $B(X) \cong M_n(\mathbb{F})$ i wtedy odwracalność $T \in M_n(\mathbb{F})$ jest równoważna niezerowaniu się wyznacznika $\det(T)$ macierzy T . Zatem

$$B(X) \cong M_n(\mathbb{F}) \implies \sigma(T) = \{\lambda \in \mathbb{F} : \det(T - \lambda 1) = 0\}.$$

Przykład 4.22 (Widmo elementu algebry funkcji). Niech $A = C(M)$, gdzie M przestrzeń zwarta. Wtedy $\text{Inv}(C(M)) = \{a \in C(M) : a(t) \neq 0 \text{ dla każdego } t \in M\}$, patrz Przykład 4.16. Zatem widmo $a \in C(M)$ jest obrazem tej funkcji

$$\sigma(a) = \{\lambda \in \mathbb{F} : \exists_{t \in M} (a - \lambda 1)(t) = 0\} = \{\lambda \in \mathbb{F} : \exists_{t \in M} a(t) = \lambda\} = a(M).$$

Powyższy przykład pozwala interpretować widmo elementu algebry Banacha jako uogólnienie obrazu funkcji. Jest to zgodne z interpretacją fizyczną widma. Mianowicie, w mechanice klasycznej wielkości fizyczne są opisywane przez funkcje. Natomiast w mechanice kwantowej wielkości fizyczne opisywane są przez operatory, a ich widmo odpowiada temu co badaczka (badacz) widzi na spektroskopie badając daną wielkość.



Myślenie o widmie operatora jako o obrazie “funkcji liczbowej” (którą ten operator zastępuje) ma również swoje umocowanie matematyczne. Mianowicie, jak zobaczymy poniżej operatory można w naturalny sposób składać z funkcjami liczbowymi (analitycznymi w przypadku ogólnym i ciągłymi w przypadku operatorów normalnych) określonymi na widmie, co prowadzi do tak zwanego rachunku funkcyjnego dla operatorów.

Obraz funkcji ciągłej na przestrzeni zwartej jest zbiorem zwartym, i tak też jest w przypadku widma. Ponadto widmo jest zawsze zbiorem niepustym jeśli rozpatrujemy przestrzenie nad ciałem liczb zespolonych $\mathbb{C}$.

Twierdzenie 4.23. *Widmo $\sigma(a)$ dowolnego elementu $a \in A$ algebry Banacha A jest zbiorem zwartym (domkniętym i ograniczonym), zawartym w $\{\lambda \in \mathbb{F} : |\lambda| \leq \|a\|\}$:*

$$\sigma(a) \subseteq \{\lambda \in \mathbb{F} : |\lambda| \leq \|a\|\},$$

oraz niepustym jeśli $\mathbb{F} = \mathbb{C}$.

Dowód. Zauważmy, że dla funkcji $F_a : \mathbb{F} \rightarrow A$ danej wzorem $F_a(\lambda) := a - \lambda 1$, $\lambda \in \mathbb{F}$, zachodzi

$$\mathbb{F} \setminus \sigma(a) = \{\lambda \in \mathbb{F} : a - \lambda 1 \in \text{Inv}(A)\} = F_a^{-1}(\text{Inv}(A)).$$

Jasne jest, że funkcja F_a jest ciągła. Natomiast zbiór $\text{Inv}(A)$ jest otwarty w A na mocy Twierdzenia 4.19. Zatem przeciwobraz $F_a^{-1}(\text{Inv}(A))$ jest zbiorem otwartym w $\mathbb{F}$, skąd jego dopełnienie, czyli widmo $\sigma(a)$ jest zbiorem domkniętym.

Założmy teraz, że $|\lambda| > \|a\| \geq 0$. Wtedy $\lambda \neq 0$ oraz $\|\lambda^{-1}a\| < 1$. Stąd

$$a - \lambda 1 = \underbrace{-\lambda}_{\neq 0} \underbrace{(1 - \lambda^{-1}a)}_{\in \text{Inv}(A)} \in \text{Inv}(A)$$

na mocy Lematu Neumanna (Lemat 4.17). Czyli $\lambda \notin \sigma(a)$. Wykazana implikacja jest równoważna inkluzji $\sigma(a) \subseteq \{\lambda \in \mathbb{F} : |\lambda| \leq \|a\|\}$. W szczególności, widmo jest zbiorem domkniętym i ograniczonym, a więc zwartym.

Niech teraz $\mathbb{F} = \mathbb{C}$ i założmy nie wprost, że $\sigma(a) = \emptyset$. Wtedy wzór $R_a(\lambda) = (a - \lambda 1)^{-1}$ definiuje funkcję (nazywaną *rezolwentą elementu a*) na całej płaszczyźnie zespolonej:

$$R_a : \mathbb{C} \rightarrow A.$$

Ponadto funkcja ta ma następujące własności:

- 1) R_a jest różniczkowalna w każdym punkcie. Rzeczywiście R_a jest złożeniem dwóch funkcji różniczkowalnych $\mathbb{C} \ni \lambda \mapsto F_a(\lambda) = a - \lambda 1 \in A$ oraz $\text{Inv}(A) \ni b \mapsto b^{-1} \in \text{Inv}(A)$. Różniczkowalność F_a jest jasna. Natomiast różniczkowalność operacji odwracania została wykazana w Twierdzeniu 4.19.
- 2) R_a zbiega do zera w nieskończoności. Rzeczywiście, dla $|\lambda| > \|a\|$, mamy

$$\begin{aligned} \|R_a(\lambda)\| &= \|(a - \lambda 1)^{-1}\| = \frac{1}{|\lambda|} \|(1 - \lambda^{-1}a)^{-1}\| \stackrel{\text{Lemat Neumanna}}{=} \frac{1}{|\lambda|} \left\| \sum_{n=0}^{\infty} (\lambda^{-1}a)^n \right\| \\ &\leq \frac{1}{|\lambda|} \sum_{n=0}^{\infty} \|\lambda^{-1}a\|^n = \frac{1}{|\lambda|} \cdot \frac{1}{1 - |\lambda|^{-1}\|a\|} = \frac{1}{|\lambda| - \|a\|} \rightarrow 0, \text{ gdy } |\lambda| \rightarrow \infty. \end{aligned}$$

Twierdzenie Liouville'a głosi, że każda *funkcja całkowita*, to jest funkcja zmiennej zespolonej, która jest analityczna na całej płaszczyźnie zespolonej, która jest ograniczona, jest stała. Warunki 1) i 2) implikują, że funkcja R_a jest całkowita i ograniczona, a co więcej w nieskończoności zbiega do zera. Zatem na mocy Twierdzenia Liouville'a R_a jest funkcją stałą tożsamościowo równą zero:

$$R_a \equiv 0.$$

Prowadzi to do sprzeczności z faktem, że $R_a(0) = a^{-1}$, bo operator odwracalny nie może być zerem. ⚡

Powyższe zastosowanie Twierdzenia Liouville'a wymaga pewnego uzasadnienia, gdyż jego klasyczne sformułowanie dotyczy funkcji o wartościach zespolonych. Natomiast rezolwenta R_a jest funkcją zmiennej zespolonej o wartościach operatorowych (w algebrze Banacha A). Zauważmy jednak, że dla dowolnego funkcjonału $f \in A^*$, z przestrzeni dualnej do przestrzeni Banacha A , na mocy warunków 1), 2) funkcja

$$\mathbb{C} \ni \lambda \mapsto f(R_a(\lambda)) \in \mathbb{C}$$

jest różniczkowalna w sensie zespolonym (a więc analityczna) i znika w nieskończoności. Zatem na mocy klasycznego Twierdzenia Liouville'a $f(R_a(\lambda)) = 0$ dla każdego $\lambda \in \mathbb{C}$ i $f \in A^*$. Jako że funkcjonały z A^* rozdziwiają elementy A , patrz Wniosek 3.6, otrzymujemy, że $R_a(\lambda) = 0$ dla każdego $\lambda \in \mathbb{C}$. $\square$

Uwaga 4.24. W dowodzie niepustości widma zastosowaliśmy Twierdzenia Liouville'a, które nie zachodzi w przypadku rzeczywistym. W przypadku operatorów na przestrzeni skończonej wymiarowej niepustość widma można wykazać korzystając z Zasadniczego Twierdzenia Algebry, które również nie zachodzi w przypadku rzeczywistym.

Przykład 4.25. Niech $A \cong M_n(\mathbb{F})$, por. Przykład 4.21. *Wielomian charakterystyczny* macierzy $a \in A$ jest dany wzorem $p_a(\lambda) = \det(a - \lambda 1)$. Zatem jest to wielomian co najwyżej stopnia n o wyrazach z ciała $\mathbb{F}$: $p_a(\lambda) = \alpha_n \lambda^n + \dots + \alpha_1 \lambda + \alpha_0$, $\alpha_i \in \mathbb{F}$, $i = 1, \dots, n$. Natomiast widmo elementu a składa się z pierwiastków tego wielomianu:

$$\sigma(a) = \{\lambda \in \mathbb{F} : \alpha_n \lambda^n + \dots + \alpha_1 \lambda + \alpha_0 = 0\}.$$

W przypadku zespolonym powyższe równanie charakterystyczne ma zawsze rozwiązanie (Zasadnicze Twierdzenie Algebry). W przypadku rzeczywistym zdarzyć się, że widmo $\sigma(a)$ będzie zbiorem pustym. Na przykład dla $a := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ mamy $p_a(\lambda) = \det(a - \lambda 1) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$, skąd

$$\sigma(a) = \{\lambda \in \mathbb{F} : \lambda^2 + 1 = 0\} = \begin{cases} \emptyset, & \text{jeśli } \mathbb{F} = \mathbb{R}, \\ \{i, -i\}, & \text{jeśli } \mathbb{F} = \mathbb{C}. \end{cases}$$

Powyższe rozważania uzasadniają dlaczego teorię spektralną rozwija się w przypadku zespolonym. Ogólnie teoria ta czerpie garściami z teorii funkcji zespolonych i dlatego

od tej pory będziemy zakładać, że $\mathbb{F} = \mathbb{C}$!

To znaczy będziemy zakładać, że wszystkie przestrzenie i algebry Banacha są rozpatrywane nad ciałem liczb zespolonych.

Rozważmy wielomian $p \in \mathbb{C}[z]$, to znaczy niech $p(z) = \alpha_n z^n + \dots + \alpha_1 z + \alpha_0$, gdzie $\alpha_i \in \mathbb{C}$, $i = 1, \dots, n$. Dla dowolnego elementu $a \in A$ zespolonej algebry Banacha A z jedyneką, definiujemy element $p(a) \in A$ wzorem

$$p(a) := \alpha_n a^n + \dots + \alpha_1 a + \alpha_0 1.$$

Poniższe stwierdzenie jest jedną z najprostszych wersji twierdzenia o odwzorowaniu spektralnym, które mówi, że widmo funkcji od operatora jest obrazem widma przy tej funkcji. Czyli mówiąc krótko i obrazowo: “gdzie idzie operator, tam idzie jego widmo”.

Stwierdzenie 4.26 (Twierdzenie o odwzorowanie spektralnym). *Dla dowolnego wielomianu $p \in \mathbb{C}[z]$ i elementu $a \in A$ zespolonej algebry Banacha z jedyneką $1 \in A$, mamy*

$$\sigma(p(a)) = p(\sigma(a)).$$

Dowód. Niech $\lambda \in \mathbb{C}$. Zapiszmy wielomian $p(z) - \lambda$ w postaci iloczynowej (na mocy Zasadniczego Twierdzenia Algebry każdy wielomian o współczynnikach zespolonych ma postać iloczynową). To znaczy, niech $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{C}$, $n \geq 0$, takie, że

$$p(z) - \lambda = \lambda_0 (z - \lambda_1) \cdot \dots \cdot (z - \lambda_n).$$

Liczby $\lambda_1, \dots, \lambda_n$ są rozwiązaniami równania $p(z) = \lambda$. Nietrudno spostrzec, że powyższe równanie ma swój analogon operatorowy

$$p(a) - \lambda 1 = \lambda_0 (a - \lambda_1 1) \cdot \dots \cdot (a - \lambda_n 1).$$

Jeśli $\lambda_0 = 0$, to $p \equiv \lambda$ i stąd $\sigma(p(a)) = \sigma(\lambda 1) = \{\lambda\} = p(\sigma(a))$. Załóżmy, zatem że $\lambda_0 \neq 0$. Wtedy odwracalność $p(a) - \lambda 1$ jest równoważna odwracalności iloczynu $(a - \lambda_1 1) \cdot \dots \cdot (a - \lambda_n 1)$, która jest z kolei równoważna odwracalności każdego ze składników, gdyż są one przemienne (🏠 sprawdzić). Czyli

$$p(a) - \lambda 1 \in \text{Inv}(A) \iff \forall_{i=1, \dots, n} (a - \lambda_i 1) \in \text{Inv}(A).$$

Stąd

$$\begin{aligned} \lambda \in \sigma(p(a)) &\iff \exists_{i=1, \dots, n} \lambda_i \in \sigma(a) \iff \sigma(a) \cap \{\lambda_1, \dots, \lambda_n\} \neq \emptyset \\ &\iff \sigma(a) \cap \{z \in \mathbb{C} : p(z) = \lambda\} \neq \emptyset \iff \exists_{z \in \sigma(a)} p(z) = \lambda \\ &\iff \lambda \in p(\sigma(a)). \end{aligned}$$

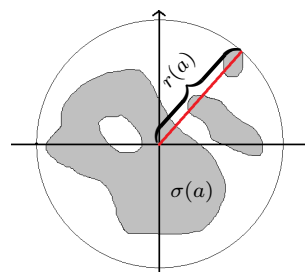
Zatem $\sigma(p(a)) = p(\sigma(a))$. □

Jako że widmo jest niepuste i zwarte, to zawsze istnieje w nim liczba o największym module.

Definicja 4.27. *Promieniem spektralnym elementu $a \in A$, zespolonej algebry Banacha A z jedyneką, nazywamy liczbę*

$$r(a) = \max_{\lambda \in \sigma(a)} |\lambda|.$$

Jest to najmniejszy promień koła o środku w zerze, w którym zawarte jest całe widmo.



Z Twierdzenia 4.23 wynika, że promień spektralny jest zawsze nie większy niż norma danego operatora: $r(a) \leq \|a\|$. Poniższy, bardzo ważny, dokładny wzór wyrażający $r(a)$ w języku norm potęg operatora a jest de facto operatorową wersją wzoru Cauchy-Hadamarda ma promień zbieżności szeregu potęgowego.

Twierdzenie 4.28 (Wzór Beurlinga-Gelfanda-Naimarka). *Promień spektralny elementu $a \in A$ wyraża się wzorem*

$$r(a) = \inf_{n \in \mathbb{N}} \|a^n\|^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}}.$$

Dowód. Niech $\lambda \in \sigma(a)$. Wtedy $\lambda^n \in \sigma(a^n)$ na mocy Stwierdzenia 4.26 i stąd $|\lambda|^n \leq \|a^n\|$ na mocy Twierdzenia 4.23. Czyli $|\lambda| \leq \|a^n\|^{\frac{1}{n}}$ dla dowolnego $\lambda \in \sigma(a)$ i $n \in \mathbb{N}$. Stąd

$$r(a) \leq \inf_{n \in \mathbb{N}} \|a^n\|^{\frac{1}{n}} \leq \liminf_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}}.$$

Zatem wystarczy pokazać, że

$$\limsup_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}} \leq r(a).$$

Przypomnijmy, że rezolwenta $\mathbb{C} \setminus \sigma(a) \ni \lambda \mapsto R_a(\lambda) = (a - \lambda 1)^{-1} \in \text{Inv}(A)$ jest różniczkowalna. Zatem kładąc $\Delta := \{\lambda \in \mathbb{C} : |\lambda| > r(a)\}$ dla dowolnego $f \in A^*$ funkcja

$$\mathbb{C} \setminus \Delta \ni \lambda \mapsto f(R_a(\lambda)) \in \mathbb{C}$$

jest holomorficzna, czyli posiada rozwinięcie postaci

$$f(R_a(\lambda)) = \sum_{k=1}^{\infty} \lambda_n \lambda^{-n}, \quad \lambda \in \mathbb{C} \setminus \Delta,$$

gdzie $\lambda_n \in \mathbb{C}$, $n \in \mathbb{N}$, są pewnymi ustalonymi współczynnikami. Korzystając z Lematu Neumanna, jeśli $|\lambda| > \|a\|$, to

$$R_a(\lambda) = (a - 1\lambda)^{-1} = -\lambda^{-1} (1 - \lambda^{-1}a)^{-1} = -\lambda^{-1} \sum_{n=0}^{\infty} (\lambda^{-1}a)^n = \sum_{n=1}^{\infty} -a^{n-1} \lambda^{-n}.$$

Stąd $f(R_a(\lambda)) = \sum_{n=1}^{\infty} f(-a^{n-1}) \lambda^{-n}$, czyli

$$\lambda_n = f(-a^{n-1}), \quad n \in \mathbb{N}.$$

Zatem dla każdego $|\lambda| > r(a)$ szereg

$$f(R_a(\lambda)) = \sum_{k=1}^{\infty} \lambda_n \lambda^{-n} = \sum_{n=1}^{\infty} f(-a^{n-1}) \lambda^{-n} = \sum_{n=1}^{\infty} f(-a^{n-1} \lambda^{-n})$$

jest zbieżny. W szczególności ciąg $\{f(a^{n-1} \lambda^{-n})\}_{n=1}^{\infty}$ jest zbieżny do zera. Skoro zachodzi to dla dowolnego funkcjonału $f \in A^*$, to ciąg $\{a^{n-1} \lambda^{-n}\}_{n=1}^{\infty}$ jest słabo zbieżny do zera, a więc ograniczony na mocy Stwierdzenia 3.36 (które wynika z Twierdzenia Banacha-Steinhaus). Stąd ciąg $\{a^n \lambda^{-n}\}_{n=1}^{\infty}$ jest również ograniczony. Zatem

$$r(a) < |\lambda| \implies \exists_{M>0} \forall_{n \in \mathbb{N}} \|a^n \lambda^{-n}\| \leq M \implies \forall_{n \in \mathbb{N}} \|a^n\|^{\frac{1}{n}} \leq M^{\frac{1}{n}} |\lambda| \implies \limsup_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}} \leq |\lambda|.$$

Stąd $\limsup_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}} \leq r(a)$, co jak wyjaśniliśmy powyżej kończy dowód. $\square$

Uwaga 4.29. Zwróćmy uwagę, że widmo elementu $a \in A$ jest zdefiniowane w terminach algebraicznych - wykorzystuje tylko pojęcie odwracalności w A i w sposób jawny nie zależy od normy. W konsekwencji to samo dotyczy definicji promienia spektralnego $r(a) = \max_{\lambda \in \sigma(a)} |\lambda|$. Natomiast wzór Beurlinga-Gelfanda-Naimarka na odwrót, nie zależy od struktury algebry A i wyraża $r(a)$ jedynie w terminach norm potęg elementu $a \in A$.

Poniższe przykłady pokazują m.in., że na ogół promień spektralny jest ostro mniejszy od normy operatora.

Przykład 4.30. Jeśli element $a \in A$ jest nilpotentny, to znaczy $a^n = 0$ dla pewnego $n \in \mathbb{N}$, to $r(a) = 0$. Na przykład, jeśli $A = M_2(\mathbb{C})$ oraz $a := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, to $\|a\| = 1$, ale $r(a) = 0$, bo $a^2 = 0$.

Przykład 4.31. Przestrzeń $A = C^{(1)}[0, 1]$ funkcji różniczkowalnych w sposób ciągły jest algebrą Banacha z operacjami zdefiniowanymi punktowo i normą $\|a\| = \max_{t \in [0, 1]} |a(t)| + \max_{t \in [0, 1]} |a'(t)|$. Rozważmy element dany wzorem $a(t) := t$, dla $t \in [0, 1]$. Wtedy $a^n(t) = t^n$ i stąd $\|a^n\| = 1 + n$ dla $n \in \mathbb{N}$. Zatem

$$r(a) = \lim_{n \rightarrow \infty} \|a^n\|^{\frac{1}{n}} = \lim_{n \rightarrow \infty} (1 + n)^{\frac{1}{n}} = 1 < 2 = \|a\|.$$

Dla operatorów normalnych promień spektralny i norma zawsze się sobie równają.

Stwierdzenie 4.32. Niech $A = B(H)$, gdzie H przestrzeń Hilberta. Jeśli $T \in B(H)$ jest operatorem normalnym (tzn. $T^*T = TT^*$), to

$$r(T) = \|T\|.$$

Dowód. Załóżmy najpierw, że $T = T^*$ jest samosprężony. Wtedy $T^n = T^{*n}$ dla każdego $n \in \mathbb{N}$. Korzystając z C^* -równości, patrz Stwierdzenie 2.77, mamy

$$\|T^n\|^2 = \|T^{*n}T^n\| = \|T^{2n}\| \quad \text{dla } n \in \mathbb{N}.$$

Stąd dla dowolnego $n > 0$ otrzymujemy

$$\|T\|^{2^n} = \|T^2\|^{2^{n-1}} = \|T^4\|^{2^{n-2}} = \dots = \|T^{2^n}\|.$$

Zatem na mocy wzoru Beurlinga-Gelfanda-Naimarka (Twierdzenie 4.28) mamy

$$r(T) = \lim_{n \rightarrow \infty} \|T^n\|^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \|T^{2^n}\|^{\frac{1}{2^n}} = \lim_{n \rightarrow \infty} \|T\| = \|T\|.$$

Niech teraz T będzie dowolnym operatorem normalnym, tzn. zakładamy jedynie, że T komutuje ze swoim sprzężeniem T^* . Korzystając z tego, że teza zachodzi dla operatorów samosprężonych otrzymujemy

$$\begin{aligned} r(T)^2 &= \lim_{n \rightarrow \infty} (\|T^n\|^2)^{\frac{1}{n}} \stackrel{C^*\text{-równość}}{=} \lim_{n \rightarrow \infty} \|T^{*n}T^n\|^{\frac{1}{n}} \stackrel{T \text{ normalny}}{=} \lim_{n \rightarrow \infty} \|(T^*T)^n\|^{\frac{1}{n}} \\ &= r(T^*T) \stackrel{T^*T \text{ samosprężony}}{=} \|T^*T\| \stackrel{C^*\text{-równość}}{=} \|T\|^2. \end{aligned}$$

Stąd $r(T) = \|T\|$. □

Wniosek 4.33. Dla dowolnego operatora $T \in B(H, K)$ między dwiema przestrzeniami Hilberta H i K mamy

$$\|T\| = \sqrt{r(T^*T)},$$

gdzie $r(T^*T)$ jest promieniem spektralnym operatora $T^*T \in B(H)$.

Dowód. Z C^* -równości, patrz Stwierdzenie 2.77, mamy $\|T\|^2 = \|T^*T\|$. Natomiast skoro T^*T samosprężony (a więc i normalny), to na mocy Stwierdzenia 4.32 otrzymujemy tezę. $\square$

Przykład 4.34. Rozważmy operator $T : \mathbb{C}^2 \rightarrow \mathbb{C}^3$, gdzie $\mathbb{C}^2$ i $\mathbb{C}^3$ są przestrzeniami Hilberta ze standardowymi normami, natomiast $T(z_1, z_2) = (z_1 + z_2, 2iz_1, 2z_2)$. Wtedy

$$T = \begin{pmatrix} 1 & 1 \\ 2i & 0 \\ 0 & 2 \end{pmatrix} \quad \text{ i } \quad T^*T = \begin{pmatrix} 1 & -2i & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2i & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}.$$

Wielomian charakterystyczny tej ostatniej macierzy to $p(\lambda) = (5 - \lambda)^2 - 1 = \lambda^2 - 10\lambda + 24$. Pierwiastki tego wielomianu, to $\lambda_1 = 4$ i $\lambda_2 = 6$. Zatem $\sigma(T^*T) = \{4, 6\}$ i $r(T^*T) = 6$. Stąd

$$\|T\| = \sqrt{r(T^*T)} = \sqrt{6}.$$

4.3 Analityczny rachunek funkcyjny

4.4 Twierdzenie spektralne

Bibliografia

- [1] J. Jakubowski, R. Sztencel, Wstęp do teorii prawdopodobieństwa, Script, Warszawa, 2001
- [2] J. Musielak, Wstęp do analizy funkcjonalnej, PWN, 1976
- [3] S. Prus, A. Stachura, Analiza funkcjonalna w zadaniach, PWN, 2007.
- [4] W. Rudin, Analiza rzeczywista i zespolona, PWN, 1998.