INSTYTUT MATEMATYCZNY UNIWERSYTETU WROCŁAWSKIEGO

Tadeusz Pytlik

Wrocław 2000

Wstęp

Analiza funkcjonalna, to dziedzina matematyki, która już od początku lat 30-tych, gdy powstawała, była entuzjastycznie przyjmowana przez matematyków. Wielbicielom dawała nowy język matematyczny i nowe, atrakcyjne narzędzie pracy. Dawała im też nową perspektywę patrzenia na matematykę.

Dziś stale jeszcze jest atrakcyjna, choć została już nieco nadwątlona przez upływający czas. Przyjęła też tytuł dziedziny „klasycznej". Osobiście uważam, że po drobnych korektach urody i przy dobrym makijażu może być jeszcze bardzo pociągająca — szczególnie dla studentów, którzy jej wcześniej nie znali.

Skrypt niniejszy został opracowany na podstawie materiałów i notatek z wykładów analizy funkcjonalnej, które prowadziłem przez wiele lat na Uniwersytecie Wrocławskim. Został on uzupełniony o zadania i ćwiczenia do samodzielnego rozwiązywania. Wskazówki do zadań i ich rozwiązania można znaleźć na końcu skryptu. Niestety w trakcie opracowywania przypadkowo uległy przetasowaniu i mimo prób nie udało mi się ustawić ich w kolejności numerów, za co bardzo serdecznie przepraszam.

Zakres materiału i poruszanych zagadnień w zasadzie odpowiada programowi typowego wykładu analizy funkcjonalnej. Potocznie mówiąc jest to „Elementarz" i „pół podręcznika do drugiej klasy". Tych, którym nie wystarczy ta dodatkowa „połówka" odsyłam od razu do innych, lepszych podręczników analizy funkcjonalnej. Spis takich podręczników zamieściłem na końcu skryptu.

Ponieważ do śledzenia rozumowań i rozwiązywania zadań potrzebna jest pewna wiedza z innych dziedzin matematyki, analizy matematycznej, algebry liniowej, teorii funkcji zmiennej zespolonej, topologii oraz teorii miary i funkcji rzeczywistych, dodatek na końcu skryptu zawiera zestaw koniecznych faktów i twierdzeń, podanych z dowodami lub odsyłaczami do literatury.

Tadeusz Pytlik

Wrocław, grudzień 2000

ROZDZIAŁ I

PRZESTRZENIE BANACHA

Przestrzenie unormowane

Podstawowe własności

Funkcję rzeczywistą $\| \ \|$ na przestrzeni liniowej X nazywamy **normą**, jeżeli

1. $\|0\| = 0$ i $\|x\| > 0$ gdy $x \neq 0$,

2. $\|x + y\| \leqslant \|x\| + \|y\|$ (podaddytywność),

3. $\|\lambda x\| = |\lambda| \, \|x\|$ (jednorodność).

Przestrzeń liniową, w której określona jest norma nazywamy **przestrzenią liniową unormowaną** (lub krótko **przestrzenią unormowaną**). Często dla podkreślenia, co jest normą w przestrzeni liniowej X, będziemy pisać $(X, \| \ \|)$.

Norma w przestrzeni liniowej X wyznacza metrykę wzorem

$$\rho(x, y) = \|x - y\|.$$

W ten sposób przestrzeń unormowana staje się przestrzenią topologiczną metryczną. Ilekroć będzie mowa o własnościach topologicznych przestrzeni unormowanej $(X, \| \ \|)$, chodzi o topologię zadaną metryką ρ.

Przestrzeń unormowana zupełna nosi nazwę **przestrzeni Banacha**.

Z określenia normy wynika, że dodawanie wektorów w przestrzeni unormowanej i mnożenie ich przez skalary są operacjami ciągłymi. Ponadto:

1.1. FAKT. *Domknięcie podprzestrzeni liniowej przestrzeni unormowanej jest podprzestrzenią liniową a domknięta podprzestrzeń liniowa przestrzeni Banacha jest przestrzenią Banacha.*

Głównym obiektem naszych zainteresowań będą przestrzenie Banacha. Zupełność przestrzeni pozwala przy badaniu zbieżności ciągu sprawdzać tylko warunek Cauchy'ego.

1.2. Twierdzenie. *Każdą przestrzeń unormowaną można uzupełnić do przestrzeni Banacha.*

Dowód: Niech X będzie przestrzenią unormowaną a $\widetilde{X}$ jej uzupełnieniem metrycznym (patrz **??**), tzn. zbiorem klas równoważności ciągów Cauchy'ego. W $\widetilde{X}$ można określić strukturę liniową kładąc

$$\big[\{x_n\}\big] + \big[\{y_n\}\big] = \big[\{x_n + y_n\}\big], \qquad \lambda\big[\{x_n\}\big] = \big[\{\lambda x_n\}\big]$$

i normę

$$\big\|\big[\{x_n\}\big]\big\| = \lim_{n\to\infty} \|x_n\|. \quad \square$$

Uwaga. Udowodnimy później (wniosek **??**), że w przestrzeni liniowej $\boldsymbol{s}_0$ ciągów, dla których tylko skończenie wiele wyrazów jest różnych od zera, nie da się wprowadzić normy tak, by była w niej przestrzenią zupełmą.

Szereg $\sum_{n=1}^{\infty} x_n$ wektorów przestrzeni unormowanej nazywa się **zbieżny**, jeżeli zbieżny jest ciąg jego sum częściowych $s_n = x_1 + x_2 + \ldots + x_n$ i nazywa się **bezwzględnie zbieżny**, jeżeli $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

1.3. Zadanie. Na to, by przestrzeń unormowana X była zupełna potrzeba i wystarcza, by każdy szereg bezwzględnie zbieżny był zbieżny.

1.4. Ćwiczenia.

1. W każdej przestrzeni liniowej można określić normę.

2. Każdy ciąg Cauchy'ego elementów przestrzeni unormowanej jest ograniczony.

3. W każdej niezerowej przestrzeni unormowanej istnieje szereg zbieżny, który nie jest bezwzględnie zbieżny.

Przykłady przestrzeni Banacha

Podamy tu kilka przykładów typowych przestrzeni Banacha.

1.5. PRZYKŁAD. Zbiór $\boldsymbol{b}$ wszystkich ograniczonych ciągów liczb zespolonych $\{x_n\}_{n=1}^{\infty}$ ze zwykłym dodawaniem ciągów i mnożeniem ich przez skalary oraz normą

$$\big\|\{x_n\}\big\|_\infty = \sup_n |x_n|$$

tworzy przestrzeń Banacha.

Jest oczywiste, że $\boldsymbol{b}$ jest przestrzenią liniową oraz, że $\| \ \|_\infty$ jest normą. Uzasadnienia wymaga jedynie zupełność przestrzeni $\boldsymbol{b}$. Jeżeli $\{x_n^1\}$, $\{x_n^2\}$, $\ldots$ jest ciągiem Cauchy'ego w $\boldsymbol{b}$, tzn. dla każdego $\varepsilon > 0$ istnieje takie k_0, że

$$\sup_n |x_n^k - x_n^m| < \varepsilon \tag{1.1}$$

dla wszystkich $k, m \geqslant k_0$, to dla każdego $n = 1, 2, 3, \ldots$ ciąg $x_n^1, x_n^2, x_n^3, \ldots$ jest liczbowym ciągiem Cauchy'ego, jest zatem zbieżny do pewnej liczby zespolonej x_n. Przechodząc w nierówności (1.1) do granicy przy $m \to \infty$ otrzymujemy

$$\sup_n |x_n^k - x_n| \leqslant \varepsilon. \tag{1.2}$$

Wnioskujemy stąd, że $\{x_n\}$ jest ciągiem ograniczonym i że jest granicą ciągu $\{x_n^1\}$, $\{x_n^2\}$, $\ldots$ w normie przestrzeni $\boldsymbol{b}$.

1.6. ZADANIE. Pokazać, że zbiory $\boldsymbol{c}$ wszystkich ciągów zbieżnych oraz $\boldsymbol{c}_0$ wszystkich ciągów zbieżnych do zera są domkniętymi podprzestrzeniami liniowymi przestrzeni $\boldsymbol{b}$, są zatem przestrzeniami Banacha w normie odziedziczonej z $\boldsymbol{b}$.

1.7. PRZYKŁAD. Zbiór ℓ^1 wszystkich absolutnie sumowalnych ciągów liczb zespolonych z normą

$$\big\|\{x_n\}\big\|_1 = \sum_{n=1}^{\infty} |x_n|$$

także tworzy przestrzeń Banacha.

Zupełność przestrzeni ℓ^1 można dowieść podobnie jak zupełność przestrzeni $\boldsymbol{b}$, należy tylko symbol $\sup_n$ zastąpić przez $\sum_{n=1}^{\infty}$.

1.8. Przykład. Niech $1 < p < \infty$. Oznaczmy przez ℓ^p zbiór tych ciągów zespolonych $\{x_n\}$, dla których $\sum |x_n|^p < \infty$ i oznaczmy

$$\|\{x_n\}\|_p = \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{1/p}.$$

Pokażemy, że ℓ^p jest przestrzenią liniową a $\| \ \|_p$ normą zupełną. Udowodnimy w tym celu dwie ważne nierówności:

1.9. Nierówność Höldera. *Jeśli $\{x_n\} \in \ell^p$, $\{y_n\} \in \ell^q$, gdzie $p, q > 1$ oraz $1/p + 1/q = 1$, to $\{x_n y_n\}$ należy do ℓ^1 i zachodzi nierówność*

$$\sum_{n=1}^{\infty} |x_n y_n| \leqslant \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{1/p} \Big(\sum_{n=1}^{\infty} |y_n|^q\Big)^{1/q}.$$

Dla dowodu zauważmy najpierw, że

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q \quad \text{dla dowolnych} \quad a, b \geqslant 0.$$

Można to łatwo odczytać z rysunku obok, bo

$$\frac{1}{p}a^p = \int_0^a s^{p-1} ds \quad \text{oraz} \quad \frac{1}{q}b^q = \int_0^b t^{q-1} dt,$$

a funkcje $t = s^{p-1}$ oraz $s = t^{q-1}$ są wzajemnie odwrotne.

Jeżeli jedna z sum po prawej stronie nierówności Höldera jest równa zeru, to nierówność oczywiście zachodzi. Jeśli zaś obie sumy, które oznaczymy odpowiednio x^p i y^p, są różne od zera, to nierówność też jest spełniona, bo mamy

$$\frac{1}{xy}\sum_{n=1}^{\infty} |x_n y_n| = \sum_{n=1}^{\infty} \frac{|x_n|}{x}\frac{|y_n|}{y} \leqslant \frac{1}{p}\sum_{n=1}^{\infty} \frac{|x_n|^p}{x^p} + \frac{1}{q}\sum_{n=1}^{\infty} \frac{|y_n|^q}{y^q} = 1.$$

1.10. Nierówność Minkowskiego. *Niech $p > 1$. Jeżeli $\{x_n\}, \{y_n\} \in \ell^p$, to*

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

Istotnie, stosując nierówność Höldera otrzymujemy

$$\begin{aligned}\sum_{n=1}^{\infty}|x_n+y_n|^p &\leqslant \sum_{n=1}^{\infty}|x_n+y_n|^{p-1}|x_n| + \sum_{n=1}^{\infty}|x_n+y_n|^{p-1}|y_n| \leqslant \\ &\leqslant \Big(\sum_{n=1}^{\infty}|x_n|^p\Big)^{1/p}\Big(\sum_{n=1}^{\infty}|x_n+y_n|^{(p-1)q}\Big)^{1/q} + \\ &\quad + \Big(\sum_{n=1}^{\infty}|y_n|^p\Big)^{1/p}\Big(\sum_{n=1}^{\infty}|x_n+y_n|^{(p-1)q}\Big)^{1/q} = \\ &= \Big[\Big(\sum_{n=1}^{\infty}|x_n|^p\Big)^{1/p} + \Big(\sum_{n=1}^{\infty}|y_n|^p\Big)^{1/p}\Big]\Big(\sum_{n=1}^{\infty}|x_n+y_n|^p\Big)^{1/q},\end{aligned}$$

gdyż $(p-1)q = p$. To daje nierówność Minkowskiego.

Nierówności Minkowskiego oznacza podaddytywność funkcji $\| \, \|_p$, a to jest jedyna nieoczywista własność normy. Wynika z niej też, że suma ciągów z ℓ^p także leży w ℓ^p, a więc, że ℓ^p jest przestrzenią liniową. Dowód zupełności jest taki jak dla przestrzeni ℓ^1.

1.11. FAKT. *Jeżeli ciąg $x = \{x_n\}$ należy do ℓ^{p_0}, $p_0 \geqslant 1$, to należy także do ℓ^p dla $p \geqslant p_0$ oraz*

$$\lim_{p\to\infty}\|x\|_p = \|x\|_\infty.$$

Istotnie, dzięki jednorodności normy możemy przyjąć $\max_n |x_n| = 1$. Wybierając wskaźnik n_0 tak, aby $\sum_{n>n_0}|x_n|^{p_0} \leqslant 1$ otrzymamy $\sum_{n>n_0}|x_n|^p \leqslant 1$. Zatem $x \in \ell^p$ oraz

$$1 \leqslant \Big(\sum_{n=1}^{\infty}|x_n|^p\Big)^{1/p} \leqslant (n_0+1)^{1/p} \leqslant 1 + \frac{n_0}{p}. \quad \square$$

Fakt ten wyjaśnia, dlaczego dla normy „supremum" użyliśmy oznaczenia $\| \, \|_\infty$. Dodajmy, że z tego też powodu przestrzeń $\boldsymbol{b}$ często oznaczana jest symbolem ℓ^∞.

1.12. ZADANIE. Pokazać, że gdy $0 < p < 1$, to na zbiorze ℓ^p ciągów sumowalnych z p-tą potęgą funkcja

$$\big\|\{x_n\}\big\|_p = \Big(\sum_{n=1}^{\infty}|x_n|^p\Big)^{1/p}$$

nie jest normą. Czy ℓ^p jest przestrzenią liniową?

1.13. Przykład. Niech Ω będzie przestrzenią topologiczną z nieujemną miarą borelowską μ oraz $S(\Omega,\mu)$ zbiorem wszystkich funkcji mierzalnych $x:\Omega\to\mathbb{C}$, przy czym utożsamimy ze sobą funkcje równe μ-prawie wszędzie. Zbiór ten z naturalnymi działaniami na funkcjach tworzy przestrzenią liniową.

Przez $L^p(\Omega,\mu)$, $1\leqslant p<\infty$, oznaczymy podprzestrzeń przestrzeni $S(\Omega,\mu)$ złożoną z tych funkcji x, dla których całka $\int_\Omega |x(t)|^p\,dt$ jest skończona. Podobnie jak w przykładzie 1.8 dowodzimy, że funkcja

$$\|x\|_p=\left(\int_\Omega |x(t)|^p\,dt\right)^{1/p}$$

jest normą w $L^p(\Omega,\mu)$. Dowód zupełności przestrzeni $L^1(\Omega,\mu)$ został przedstawiony na stronie **??**, zupełność pozostałych przestrzeni $L^p(\Omega,\mu)$, $1<p<\infty$, dowodzi się tak samo.

Oznaczmy ponadto przez $L^\infty(\Omega,\mu)$ podprzestrzeń przestrzeni $S(\Omega,\mu)$, złożoną z funkcji **istotnie ograniczonych**, tj. funkcji x, dla których wielkość

$$\inf_{\substack{A\subset\Omega\\ \mu(A)=0}}\ \sup_{t\in\Omega\setminus A}\big|x(t)\big|$$

jest skończona. Wielkość tą oznaczany $\operatorname{ess\,sup}_{t\in\Omega}|x(t)|$ lub $\|x\|_\infty$ i nazywamy **supremum istotnym** funkcji x. Funkcja $\|\ \|_\infty$ jest normą a $L^\infty(\Omega,\mu)$ w tej normie przestrzenią zupełną.

Utworzyliśmy w ten sposób całą nową klasę przykładów przestrzeni Banacha, uogólniających przestrzenie ℓ^p, $1\leqslant p\leqslant\infty$.

Jako modelu przestrzeni (Ω,μ) będziemy używać na ogól prostej rzeczywistej $\mathbb{R}$ lub przedziałów (a,b) z miarą Lebesgue'a, płaszczyzny zespolonej $\mathbb{C}$ z miarą płaską Lebesgue'a bądź zbioru liczb naturalnych $\mathbb{N}$ z miarą „liczącą”, dla której miara zbioru jest jego licznością. Pisanie symbolu (Ω,μ) ma ustrzec przed uciążliwym rozpatrywaniem przypadków oraz wykorzystywaniem dodatkowych własności konkretnych przestrzeni, np. zwartości przestrzeni czy skończoności miary.

1.14. Przykład. Niech T będzie dowolnym zbiorem niepustym. Zbiór $B(T)$ wszystkich funkcji ograniczonych $x:T\to\mathbb{C}$ jest przestrzenią Banacha w normie $\|x\|_\infty=\sup_{t\in T}|x(t)|$.

Gdy za T przyjmiemy zbiór liczb naturalnych $\mathbb{N}$, to przestrzeń $B(T)$ stanie się identyczna z wcześniej wprowadzoną przestrzenią $\boldsymbol{b}$ ciągów ograniczonych.

1.15. Przykład. Niech S będzie przestrzenią topologiczną. Oznaczmy przez $C(S)$ podprzestrzeń przestrzeni $B(S)$ złożoną z wszystkich funkcji ciągłych. Jest to przestrzeń zupełna, bo granica jednostajnie zbieżnego ciągu funkcji ciągłych jest funkcją ciągłą.

Dla uniknięcia patologicznych przykładów będziemy zawsze zakładali, że S jest przestrzenią topologiczną całkowicie regularną. Typowymi przykładami, do których praktycznie wystarczy się ograniczyć, są przestrzenie $C([a,b])$ oraz $C(\mathbb{R})$.

Załóżmy, że przestrzeń topologiczna S jest lokalnie zwarta, tzn. każdy punkt w S posiada bazę otoczeń złożoną ze zbiorów zwartych. Oznaczmy przez $C_0(S)$ podprzestrzeń $C(S)$ złożoną z tych funkcji x, że dla każdego $\varepsilon > 0$ istnieje zbiór zwarty $K \subset S$ o własności $|x(s)| < \varepsilon$ dla $s \notin K$. Jest to domknięta podprzestrzeń liniowa przestrzeni $C(S)$, jest zatem przestrzenią Banacha. O funkcjach $x \in C_0(S)$ mówi się czasem, że „znikają w nieskończoności".

Przestrzenie Banacha $\boldsymbol{c}$ oraz $\boldsymbol{c}_0$ opisane w zadaniu 1.6 są szczególnymi przykładami odpowiednio przestrzeni $C(S)$ oraz $C_0(S)$. Można je otrzymać np. biorąc za S zbiór $\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ z topologią odziedziczoną z prostej $\mathbb{R}$ (jest to przestrzeń topologiczna normalna, lokalnie zwarta) i utożsamiając funkcje x na S z ciągami $\{x(\frac{1}{n})\}$.

Izomorfizm. Równoważność norm

Powiemy, że przestrzenie unormowane X i Y są **izomorficzne topologicznie**, jeśli istnieje izomorfizm algebraiczny $T : X \to Y$, ciągły jako funkcja z przestrzeni topologicznej X do przestrzeni topologicznej Y i taki, że $T^{-1} : Y \to X$ jest także funkcją ciągłą. Izomorfizm nazwiemy **izometrycznym** (lub krótko **izometrią**), jeżeli $\|Tx\| = \|x\|$ dla wszystkich $x \in X$.

1.16. PRZYKŁAD. Rozpatrzmy w przestrzeni $\mathbb{R}^2$ dwie normy

$$\big\|(x_1, x_2)\big\|_1 = |x_1| + |x_2|, \qquad \big\|(x_1, x_2)\big\|_\infty = \max\big\{|x_1|, |x_2|\big\}.$$

Twierdzimy, choć na pierwszy rzut oka może się to wydać nieprawdopodobne, że przestrzenie $(\mathbb{R}^2, \|\ \|_1)$ i $(\mathbb{R}^2, \|\ \|_\infty)$ są izometrycznie izomorficzne. Odwzorowanie identycznościowe izometrią oczywiście nie jest, jest nią za to odwzorowanie $(x_1, x_2) \to (x_1 + x_2, x_1 - x_2)$. Wynika to z równości

$$\max\big\{|x_1 + x_2|, |x_1 - x_2|\big\} = |x_1| + |x_2|,$$

którą łatwo sprawdzamy rozpatrując przypadki, gdy liczby x_1 i x_2 są tego samego i przeciwnego znaku.

W przestrzeni $\mathbb{R}^3$ sytuacja jest odmienna, kulą jednostkową w normie $\|\ \|_1$ jest ośmiościan, a w normie $\|\ \|_\infty$ sześcian, nie ma zatem odwzorowania liniowego przeprowadzającego jedną z figur na drugą. W przestrzeni $\mathbb{R}^n$, $n > 3$, jest jeszcze gorzej, pierwszą z kul jednostkowych jest wielościan o $2n$ wierzchołkach, a drugą wielościan o 2^n wierzchołkach.

1.17. PRZYKŁAD. Pokażemy, że przestrzenie $\boldsymbol{c}_0$ i $\boldsymbol{c}$ są izomorficznie topologicznie. Pokażemy też, że izometrycznie izomorficzne nie są.

Jest rzeczą jasną, że odwzorowanie $T : \boldsymbol{c}_0 \to \boldsymbol{c}$, określone wzorem

$$T(x_1, x_2, x_3, \ldots) = (x_1 + x_2, x_1 + x_3, x_1 + x_4, \ldots)$$

jest algebraicznym izomorfizmem tych przestrzeni a jego odwzorowanie odwrotne $T^{-1} : \boldsymbol{c} \to \boldsymbol{c}_0$ ma postać

$$T^{-1}(x_1, x_2, x_3, \ldots) = (x_0, x_1 - x_0, x_2 - x_0, x_3 - x_0, \ldots),$$

gdzie $x_0 = \lim_{n\to\infty} x_n$. Ponieważ $\|Tx\| \leqslant 2\,\|x\|$ dla $x \in \boldsymbol{c}_0$ i $\|T^{-1}x\| \leqslant 2\,\|x\|$ dla $x \in \boldsymbol{c}$, więc oba odwzorowania są ciągłe.

Dowód drugiego ze stwierdzeń jest znacznie trudniejszy. Należy wykazać, że żaden z algebraicznych izomorfizmów tych przestrzeni izometrią nie jest. Dla każdego konkretnego izomorfizmu T można zapewne nietrudno wskazać taki element x, że $\|Tx\| \neq \|x\|$, ale dla wszystkich izomorfizmów takiego uniwersalnego elementu nie ma. Uciekniemy się wobec tego do podobnego rozumowania jak poprzednim przykładzie, pokażemy mianowicie, że domknięta kula jednostkowa K przestrzeni $\boldsymbol{c}$ ma bardzo wiele wierzchołków, a domknięta kula jednostkowa K_0 przestrzeni $\boldsymbol{c}_0$ nie ma ich wcale, nie można zatem w sposób izometryczny przeprowadzić jednej na drugą.

Hasło „wierzchołek kuli K" zastąpimy jednak bardziej precyzyjnym „punkt ekstremalny kuli K". Z definicji jest to każdy punkt $x \in K$, którego nie można przedstawić w postaci $x = \lambda y + (1-\lambda)z$ dla pewnych $y, z \in K$, $y \neq z$ i $0 < \lambda < 1$, tj. punkt, który nie leży wewnątrz żadnego odcinka

$$\big\{\lambda y + (1-\lambda)z : 0 \leqslant \lambda \leqslant 1\big\}$$

łączącego dwa różne punkty y i z kuli K. Jest jasne, że każda izometria przeprowadza punkty ekstremalne jednej kuli na punkty ekstremalne drugiej kuli.

W domkniętej kuli jednostkowej K przestrzeni $\boldsymbol{c}$ punktami ekstremalnymi są wszystkie te punkty $x = (x_1, x_2, x_3, \ldots)$, dla których $|x_n| = 1$, $n = 1, 2, 3, \ldots$.

1.18. FAKT. *Domknięta kula jednostkowa przestrzeni* $\boldsymbol{c}_0$ *nie ma punktów ekstremalnych.*

Dowód: Pokażemy, że każdy element x kuli $K_0 = \{x \in \boldsymbol{c}_0 : \|x\| \leqslant 1\}$ jest środkiem pewnego odcinka leżącego całkowicie w K_0, tj., że $x = \frac{1}{2}y + \frac{1}{2}z$ dla pewnych dwóch różnych elementów y i z tej kuli. Jeżeli x ma postać $x = (x_1, x_2, x_3, \ldots)$, to $|x_n| \leqslant 1$ dla $n = 1, 2, 3 \ldots$ oraz $x_n \to 0$ przy $n \to \infty$. Wybierzmy wskaźnik n_0 tak, by $|x_{n_0}| < 1$ i określmy elementy $y = (y_1, y_2, y_3, \ldots)$ i $z = (z_1, z_2, z_3, \ldots)$ kładąc $y_n = z_n = x_n$ dla $n \neq n_0$ i przyjmując za y_{n_0} i z_{n_0} takie liczby, by były różne, by $|y_{n_0}| \leqslant 1$, $|z_{n_0}| \leqslant 1$ oraz, by $x_{n_0} = \frac{1}{2} y_{n_0} + \frac{1}{2} z_{n_0}$. Wtedy $y, z \in K_0$, $y \neq z$ oraz $x = \frac{1}{2} y + \frac{1}{2} z$. □

DEFINICJA. Załóżmy, że w przestrzeni liniowej X dane są dwie normy $\| \, \|_1$ i $\| \, \|_2$. Powiemy, że normy te są **równoważne**, gdy istnieją takie liczby C_1 i C_2, że

$$\|x\|_1 \leqslant C_1 \|x\|_2 \quad \text{oraz} \quad \|x\|_2 \leqslant C_2 \|x\|_2$$

dla wszystkich $x \in X$.

1.19. TWIERDZENIE. *W przestrzeni liniowej X normy $\| \, \|_1$ i $\| \, \|_2$ są równoważne wtedy i tylko wtedy, gdy zbieżność ciągów w normie $\| \, \|_1$ jest równoważne ich zbieżności w normie $\| \, \|_2$, czyli gdy odwzorowanie tożsamościowe $Tx = x$ jest izomorfizmem przestrzeni $(X, \| \, \|_1)$ na przestrzeń $(X, \| \, \|_2)$.*

Dowód: Jest jasne, że warunek

$$\|x\|_1 \leqslant C \, \|x\|_2 \quad \text{dla} \quad x \in X$$

wystarcza na to, by każdy ciąg zbieżny w normie $\| \, \|_2$ był zbieżny w normie $\| \, \|_1$. Pokażemy, że warunek ten jest także konieczny.

Jeżeli nierówność

$$\|x\|_1 \leqslant C \, \|x\|_2 \quad \text{dla} \quad x \in X$$

nie zachodzi przy żadnej stałej C, to w X istnieje taki ciąg $\{x_n\}$, że

$$\|x\|_1 > n \, \|x_n\|_2, \qquad n = 1, 2, 3, \ldots \, .$$

Wtedy dla ciągu

$$y_n = \frac{1}{\sqrt{n} \|x_n\|_2} x_n, \qquad n = 1, 2, 3, \ldots$$

zachodzą nierówności

$$\|y_n\|_1 = \frac{\|x_n\|_1}{\sqrt{n} \|x_n\|_2} > \sqrt{n},$$
$$\|y_n\|_2 \leqslant \frac{1}{\sqrt{n}}.$$

Wynika z nich, że ciąg $\{y_n\}$ jest zbieżny (do zera) w normie $\| \, \|_2$, a nie jest zbieżny w normie $\| \, \|_1$ (nie jest nawet ograniczony). □

Z powyższego twierdzenia wynika, że algebraiczny izomorfizm T przestrzeni unormowanej $\big(X, \|\ \|_X\big)$ na przestrzeń unormowaną $\big(Y, \|\ \|_Y\big)$ jest izomorfizmem topologicznym wtedy i tylko wtedy, gdy w przestrzeni X norma $\|\ \|_1 = \|\ \|_X$ i norma $\|\ \|_2$ określona wzorem $\|x\|_2 = \|Tx\|_Y$ są równoważne.

1.20. Przykład. W przestrzeni $C^1[0,1]$ funkcji mających ciągłą pochodną na przedziale $[0,1]$ normy

$$\|x\|_1 = \max_{t\in[0,1]} \big|x(t)\big| + \max_{t\in[0,1]} \big|x'(t)\big|$$
$$\|x\|_2 = \big|x(0)\big| + \max_{t\in[0,1]} \big|x'(t)\big|$$

są równoważne. Nierówność $\|x\|_2 \leqslant \|x\|_1$ jest oczywista. Z drugiej strony $x(t) = x(0) + \int_0^t x'(s)\,ds$, więc

$$\max_{t\in[0,1]} \big|x(t)\big| \leqslant \big|x(0)\big| + \max_{s\in[0,1]} \big|x'(s)\big| = \|x\|_2,$$

stąd $\|x\|_1 \leqslant 2\,\|x\|_2$.

1.21. Zadanie. Pokazać, że w przestrzeni $C[0,1]$ normy

$$\|x\|_1 = \int_0^1 \big|x(t)\big|\,dt, \qquad \|x\|_2 = \max_{t\in[0,1]} \big|x(t)\big|$$

nie są równoważne.

1.22. Twierdzenie. *W przestrzeni skończenie wymiarowej każde dwie normy są równoważne. Unormowana przestrzeń skończenie wymiarowa jest zupełna.*

Dowód: Określimy w przestrzeni skończenie wymiarowej X pewną normę $\|\ \|_1$ i pokażemy, że każda inna norma $\|\ \|$ jest z nią równoważna.

Ustalmy w tym celu bazę Hamela $e_1, e_2, \ldots, e_m$ przestrzeni X i dla elementu $x \in X$ postaci

$$x = \sum_{k=1}^{m} x_k\, e_k$$

połóżmy

$$\|x\|_1 = \max_{1\leqslant k\leqslant m} |x_k|.$$

Jasne jest, że funkcja $\|\ \|_1$ jest normą na X oraz, że dla dowolnej normy $\|\ \|$ na X zachodzi nierówność

$$\|x\| \leqslant C\,\|x\|_1, \quad x \in X.$$

Wynika to z nierówności

$$\|x\| = \left\| \sum_{k=1}^{m} x_k\, e_k \right\| \leqslant \sum_{k=1}^{m} |x_k| \|e_k\| \leqslant \max_{1 \leqslant k \leqslant m} |x_k| \cdot \sum_{k=1}^{m} \|e_k\|.$$

Załóżmy, nie wprost, że normy $\| \ \|_1$ i $\| \ \|$ nie są równoważne. Możemy więc dla każdej liczby naturalnej n znaleźć taki element $x_n \in X$, że

$$\|x_n\|_1 > n\, \|x_n\|.$$

Stąd dla $y_n = \frac{1}{\|x_n\|}\, x_n$ otrzymamy $\|y_n\|_1 = 1$ oraz $\|y_n\| \to 0$ przy $n \to \infty$. Ciąg $\{y_n\}$ jest zatem zbieżny do zera w normie $\| \ \|$. Ciąg ten nie musi być zbieżny w normie $\| \ \|_1$, ale jako ograniczony, zawiera (na mocy twierdzenia Bolzano-Weierstrassa) pewien podciąg zbieżny $\{y_{n_k}\}$. Dla jego granicy y z jednej strony mamy $y \neq 0$, bo $\|y\|_1 = \lim_{k\to\infty} \|y_{n_k}\| = 1$, a z drugiej $y = 0$, bo zbieżność w normie $\| \ \|_1$ pociąga zbieżność w normie $\| \ \|$. Tu sprzeczność.

Zupełność przestrzeni X w normie $\| \ \|$ wynika z twierdzenia 1.19 i oczywistej zupełności X w normie $\| \ \|_1$. □

Ośrodkowość. Bazy topologiczne

Jeżeli przestrzeń unormowana ma podzbiór przeliczalny gęsty, to mówimy, że jest **ośrodkowa**, a sam podzbiór nazywamy **ośrodkiem**.

1.23. FAKT. *Przestrzeń unormowana X jest ośrodkowa wtedy i tylko wtedy, gdy zawiera przeliczalny podzbiór liniowo gęsty P, tzn. $\overline{\operatorname{lin} P} = X$.*

Istotnie, jeśli zbiór $P = \{x_1, x_2, x_3, \ldots\}$ jest liniowo gęsty w X, to kombinacje liniowe postaci $\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$, $\lambda_k \in \mathbb{W} + i\,\mathbb{W}$, tworzą zbiór przeliczalny gęsty w X.

Każda z przestrzeni ℓ^p, $1 \leqslant p < \infty$, jest ośrodkowa, podzbiór przeliczalny gęsty tworzą tu ciągi z jedynką w dokładnie jednym miejscu. Przestrzeń ℓ^∞ nie jest ośrodkowa, każde dwa różne ciągi zero-jedynkowe są odległe od siebie o 1, a jest ich nieprzeliczalnie wiele. Funkcje charakterystyczne zbiorów borelowskich ograniczonych na prostej $\mathbb{R}$ tworzą przeliczalny zbiór liniowo gęsty w każdej z przestrzeni $L^p(\mathbb{R})$, $1 \leqslant p < \infty$. Także, jak wiemy z twierdzenia Weierstrassa, funkcje $1, t, t^2, t^3, \ldots$ tworzą zbiór liniowo gęsty w przestrzeni $C[a, b]$. Wszystkie te przestrzenie są zatem ośrodkowe. Przestrzenie $C(\mathbb{R})$ i $L^\infty(\mathbb{R})$ nie są ośrodkowe z tego samego powodu, co przestrzeń ℓ^∞

DEFINICJA. Ciąg $e_1, e_2, e_3, \ldots$ elementów przestrzeni Banacha X nazywamy **bazą topologiczną** tej przestrzeni, jeżeli każdy element $x \in X$ ma jednoznaczne przedstawienie w postaci zbieżnego szeregu

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

Jest oczywiste, że baza topologiczna jest zbiorem liniowo niezależnym i liniowo gęstym w przestrzeni. Bazę topologiczną mogą mieć więc tylko ośrodkowe przestrzenie Banacha.

Zbiór ciągów e_1, e_2, $e_3, \ldots$ postaci $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$ z jedynką na n-tym miejscu tworzy bazę topologiczną w każdej z przestrzeni $\mathbf{c}_0$, ℓ^p, $1 \leqslant p < \infty$. Ciekawszy przykład bazy pochodzi od Schaudera.

1.24. PRZYKŁAD. **Baza Schaudera**. Niech $\{t_0, t_1, t_2, \ldots\}$ będzie zbiorem gęstym w przedziale $[a, b]$, przy czym $t_0 = a$, $t_1 = b$, i niech $x_0(t) = 1$, $x_1(t) = \frac{t-a}{t-b}$ dla $t \in [a, b]$. Funkcję x_n dla $n \geqslant 2$ określamy w sposób następujący. Punkty $\{t_0, t_1, t_2, \ldots, t_n\}$ dzielą przedział $[a, b]$ na podprzedziały, do jednego z nich wpada punkt t_n, oznaczmy ten przedział $[\alpha_n, \beta_n]$. Jako x_n określamy funkcję równą 0 dla $t \in [a, \alpha_n] \cup [\beta_b, b]$, równą 1 w punkcie $t = t_n$ i liniową w każdym z przedziałów $[\alpha_n, t_n]$ i $[t_n, \beta_n]$.

Pokażemy, że funkcje $x_0, x_1, x_2, \ldots$ stanowią bazę przestrzeni $C[a, b]$. Zauważmy w tym celu, że $x_n(t_i) = 0$ dla $i = 0, 1, 2, \ldots, n-1$ oraz $x_n(t_n) = 1$. Wynika z tego, że jeśli funkcja y przedstawia się w postaci jednostajnie zbieżnego szeregu

$$y(t) = \sum_{k=0}^{\infty} \lambda_k \, x_k(t), \tag{1.3}$$

to $y(t_0) = \lambda_0$, $y(t_1) = \lambda_0 \, x_0(t_1) + \lambda_1$, $y(t_2) = \lambda_0 \, x_0(t_2) + \lambda_1 \, x(t_2) + \lambda_3$, itd. Pozwala to jednoznacznie wyznaczyć współczynniki λ_k. Udowodniliśmy w ten sposób, że jeżeli przedstawienie (1.3) istnieje, to jest jedyne.

Pozostaje pokazać, że jeśli współczynniki λ_k określimy w wyżej opisany sposób, to ciąg $\{y_n\}$ n-tych sum częściowych szeregu (1.3) jest jednostajnie zbieżny do y. Z konstrukcji wynika, że wykresem funkcji y_n jest łamana o wierzchołkach w punktach o odciętych $t_0, t_1, t_2, \ldots, t_n$ i pokrywająca się z wykresem funkcji y w tych punktach. Zatem $\|y_n - y\|_\infty \to 0$, bo funkcja y jest jednostajnie ciągła.

1.25. ZADANIE. Pokazać, że układ Schaudera nie tworzy bazy topologicznej w żadnej z przestrzeni $L^p(a, b)$, $1 \leqslant p < \infty$.

Uwaga. W roku 1932 Stanisław Mazur postawił problem, czy każda ośrodkowa przestrzeń Banacha ma bazę. Za jego rozwiązanie wyznaczył nagrodę w postaci żywej gęsi. Gęś otrzymał matematyk szwedzki Per Enflo za podanie w roku 1973 kontrprzykładu a zdjęcia z tej uroczystości zamieściła prasa.

Przestrzenie ilorazowe

Niech Y będzie podprzestrzenią liniową przestrzeni unormowanej X. Funkcja

$$\operatorname{dist}(x,Y)=\inf_{y\in Y}\|x-y\|,\quad x\in X,$$

ma następujące własności:

1. $\{x\in X:\operatorname{dist}(x,Y)=0\}=\overline{Y}$,
2. $\operatorname{dist}(x+y,Y)\leqslant\operatorname{dist}(x,Y)+\operatorname{dist}(y,Y)$,
3. $\operatorname{dist}(\lambda\,x,Y)=|\lambda|\,\operatorname{dist}(x,Y)$.

Wynika stąd, że jeżeli Y jest domkniętą podprzestrzenią X oraz $[x]$ jest warstwą przestrzeni ilorazowej X/Y, to dla każdych $x_1,x_2\in[x]$ mamy $\operatorname{dist}(x_1,Y)=\operatorname{dist}(x_2,Y)$, a więc funkcjonał

$$\|\,[x]\,\|=\operatorname{dist}(x,Y)=\inf_{y\in Y}\|x-y\|$$

jest normą w przestrzeni X/Y.

1.26. Zadanie. Niech $x=(x_1,x_2,x_3,\ldots)\in\ell^\infty$. Pokazać, że w przestrzeni ilorazowej $\ell^\infty/\boldsymbol{c}_0$

$$\|\,[x]\,\|=\limsup_{n\to\infty}|x_n|.$$

1.27. Przykład. Niech Y oznacza domkniętą podprzestrzeń przestrzeni $\boldsymbol{c}$, złożoną z ciągów stałych. Pokażemy, że przestrzeń ilorazowa $\boldsymbol{c}/Y$ jest izomorficzna z przestrzenią $\boldsymbol{c}_0$.

Dla ciągu $x=(x_1,x_2,x_3,\ldots)\in\boldsymbol{c}$ w warstwie $[x]$ istnieje dokładnie jeden ciąg z podprzestrzeni $\boldsymbol{c}_0$, mianowicie ciąg $x'=(x_1-x_0,x_2-x_0,x_3-x_0,\ldots)$, gdzie $x_0=\lim_{n\to\infty}x_n$. Pozwala to utożsamić przestrzeń ilorazową $\boldsymbol{c}/Y$ z przestrzenią $\boldsymbol{c}_0$. Ponieważ $\|x'\|\leqslant 2\|x\|$ oraz $x'\in[x]$, więc

$$\|\,[x]\,\|\leqslant\|x'\|\leqslant 2\,\|\,[x]\,\|.$$

Zauważmy jeszcze, że norma w $\boldsymbol{c}/Y$ nie pokrywa się z normą w $\boldsymbol{c}_0$, a w układzie powyższych nierówności każda z równości jest możliwa.

1.28. TWIERDZENIE. *Jeżeli X jest przestrzenią Banacha a Y jej domkniętą podprzestrzenią, to X/Y jest też przestrzenią Banacha.*

Dowód: Wystarczy pokazać (por. zadanie 1.3), że w X/Y każdy szereg bezwzględnie zbieżny jest zbieżny. Niech $x_1, x_2, x_3, \ldots$ będzie ciągiem w X, dla którego

$$\sum_{k=1}^{\infty} \big\| [x_k] \big\| < \infty.$$

Z każdej z warstw $[x_k]$ wybierzmy element x_k' tak, by $\|x_k'\| \leqslant \big\| [x_k] \big\| + 1/2^k$. Wtedy $\sum_{k=1}^{\infty} \|x_k'\| < \infty$, a ponieważ przestrzeń X jest zupełna, więc szereg $\sum_{k=1}^{\infty} x_k'$ jest zbieżny do pewnego elementu $x' \in X$. Z nierówności

$$\left\| [x'] - \sum_{k=1}^{n} [x_k] \right\| \leqslant \left\| x' - \sum_{k=1}^{n} x_k' \right\|, \quad n = 1, 2, 3, \ldots,$$

wynika, że $[x']$ jest granicą sum częściowych szeregu $\sum_{k=1}^{\infty} [x_k]$. □

1.29. PRZYKŁAD. Niech S_0 będzie dowolnym niepustym podzbiorem domkniętym przestrzeni normalnej S. Oznaczmy przez X domkniętą podprzestrzeń przestrzeni $C(S)$ złożoną z funkcji zerujących się na S_0. Pokażemy, że przestrzeń $C(S_0)$ jest izometrycznie izomorficzna z przestrzenią ilorazową $C(S)/X$.

Odwzorowanie przyporządkowujące funkcji $x \in C(S)$ jej obcięcie $x|_{S_0}$ do zbioru S_0 jest kontrakcją z przestrzeni $C(S)$ do $C(S_0)$. Jego jądrem jest zbiór X. Wzór $T[x] = x|_{S_0}$ określa zatem odwzorowanie liniowe T z $C(S)/X$ w $C(S_0)$, także będące kontrakcją. Z drugiej strony z twierdzenia Tietzego-Urysohna (patrz **??**) wynika, że każdą ograniczoną funkcję ciągłą x_0 na S_0 można przedłużyć do ograniczonej funkcji ciągłej x na S i to tak, by $\|x\|_\infty = \|x_0\|_\infty$. Oznacza to, że T odwzorowuje „na" i nie zmniejsza normy. Wobec tego jest izometrią.

Udowodnimy jeszcze twierdzenie o uniwersalności przestrzeni ℓ^1 dla klasy wszystkich ośrodkowych przestrzeni Banacha.

1.30. TWIERDZENIE. *Każda ośrodkowa przestrzeń Banacha jest izometrycznie izomorficzna z przestrzenią ilorazową ℓ^1/Y, gdzie Y jest pewną domkniętą podprzestrzenią liniową przestrzeni ℓ^1.*

Dowód: Niech $x_1, x_2, x_3, \ldots$ będzie dowolnym podzbiorem gęstym sfery jednostkowej $\big\{x \in X : \|x\| = 1\big\}$ przestrzeni X. Określmy odwzorowanie $T : \ell^1 \to X$ wzorem

$$T(\lambda_1, \lambda_2, \lambda_3, \ldots) = \sum_{n=1}^{\infty} \lambda_n x_n.$$

Jest oczywiste, że T jest kontrakcją liniową. Pokażemy, że obrazem T jest cała przestrzeń X. Zauważmy w tym celu, że dla każdego $x \in X$, każdej liczby naturalnej k i każdego $\varepsilon > 0$ istnieje taki wskaźnik $n > k$, że

$$\big\| x - \|x\|\, x_n \big\| < \varepsilon.$$

Ustalmy dowolnie liczbę $\varepsilon > 0$ i wybierzmy ciąg $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ liczb dodatnich tak, by $\sum_{k=1}^{\infty} \varepsilon_k < \varepsilon$. Niech x będzie dowolnym elementem przestrzeni X. Wybierzmy wskaźnik n_1 tak, by $\|x - \lambda_{n_1} x_{n_1}\| < \varepsilon_1$, gdzie przyjęliśmy oznaczenie $\lambda_{n_1} = \|x\|$, następnie wskaźnik $n_2 > n_1$ tak, by $\big\|(x - \lambda_{n_1} x_{n_1}) - \lambda_{n_2} x_{n_2}\big\| < \varepsilon_2$, z oznaczeniem $\lambda_{n_2} = \|x - \lambda_{n_1} x_{n_1}\|$, itd. W rezultacie otrzymamy ciąg $\lambda_{n_1}, \lambda_{n_2}, \ldots$ o własności

$$\lambda_{n_{k+1}} = \big\| x - (\lambda_{n_1} x_{n_1} + \lambda_{n_2} x_{n_2} + \ldots + \lambda_{n_k} x_{n_k}) \big\| < \varepsilon_k.$$

Uzupełnijmy go do ciągu $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ przez przyjęcie zer za brakujące składniki. Wtedy $T(\lambda) = x$ oraz

$$\|\lambda\|_1 = \sum_{k=1}^{\infty} \lambda_{n_k} \leqslant \|x\| + \sum_{k=1}^{\infty} \varepsilon_k < \|x\| + \varepsilon.$$

Wnosimy z tej konstrukcji, że T odwzorowuje ℓ^1 na całą przestrzeń X.

Określmy podprzestrzeń $Y \subset \ell^1$ jako jądro

$$Y = \big\{ \lambda \in \ell^1 : T(\lambda) = 0 \big\}$$

odwzorowania T. Wtedy T staje się algebraicznym izomorfizmem ℓ^1/Y na X. Ciągłość odwzorowania T gwarantuje, że Y jest domkniętą podprzestrzenią liniową przestrzeni ℓ^1. Z powyższej konstrukcji wynika, że jeżeli $T(\lambda) = x$, to dla dowolnego $\varepsilon > 0$ zachodzi nierówność

$$\big\| [\lambda] \big\| \leqslant \|x\| + \varepsilon,$$

zaś z faktu, że T jest kontrakcją, wynika nierówność przeciwna

$$\|x\| \leqslant \big\| [\lambda] \big\|,$$

musi zatem być $\big\| [\lambda] \big\| = \|x\|$. To dowodzi, że T jest izometrią ℓ^1/Y na X. $\square$

Produkt przestrzeni

Niech $(X_k, \| \, \|_k)$, $k = 1, 2, \ldots, n$, będą przestrzeniami unormowanymi nad tym samym ciałem ($\mathbb{R}$ lub $\mathbb{C}$). **produktem** (topologicznym) $\prod_{k=1}^n X_k$ nazywamy produkt kartezjański zbiorów X_k z działaniami i normą określonymi następująco:

$$
\begin{aligned}
(x_1, x_2, \ldots, x_k) + (x_1', x_2', \ldots, x_k') &= (x_1 + x_1', x_2 + x_2', \ldots, x_k +'_k), \\
\lambda\,(x_1, x_2, \ldots, x_k) &= (\lambda x_1, \lambda x_2, \ldots, x_\lambda k), \\
\big\|(x_1, x_2, \ldots, x_k)\big\| &= \sum_{k=1}^n \|x_k\|_k .
\end{aligned}
$$

1.31. FAKT. *Produkt $\prod_{k=1}^n X_k$ przestrzeni unormowanych jest przestrzenią Banacha wtedy i tylko wtedy, gdy każda z przestrzeni X_k jest zupełna.*

1.32. PRZYKŁAD. Przestrzeń $C^n[a,b]$ funkcji mających ciągłe pochodne do rzędu n włącznie jest izomorficzna z produktem $\mathbb{C}^n \times C[a,b]$. Istotnie, funkcja f jest wyznaczona jednoznacznie przez wektor $\big(f(a), f'(a), \ldots, f^{(n-1)}(a)\big)$ i funkcję $f^{(n)}$. Można w tym celu użyć wzoru Taylora

$$
f(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}\,(t-a)^k + \int_a^t \frac{(t-a)^{n-1}}{(n-1)!}\, f^{(n)}(s)\, ds.
$$

DEFINICJA. Jeżeli X jest przestrzenią liniową oraz $X_1, X_2, \ldots, X_n$ takimi jej podprzestrzeniami, że

$$
\operatorname{lin} \bigcup_{k=1}^n X_k = X \quad \text{oraz} \quad X_k \cap \sum_{\substack{j=1 \\ j\neq k}}^n X_j = \{0\}, \qquad k = 1, 2, \ldots, n,
$$

to X nazywamy **algebraiczną sumą prostą** podprzestrzeni X_k. Każdy wektor x ma wtedy jednoznaczne przedstawienie $x = \sum_{k=1}^n x_k$, gdzie $x_k \in X_k$, a odwzorowanie $\varphi : (x_1, x_2, \ldots, x_n) \to \sum_{k=1}^n x_k$ jest algebraicznym izomorfizmem $\prod_{k=1}^n X_k$ na X. Jeżeli założymy dodatkowo, że X jest przestrzenią unormowaną, to odwzorowanie φ jest ciągłe. Gdy jest ono izomorfizmem (topologicznym), to X nazywamy **sumą prostą** (topologiczną) podprzestrzeni X_k i piszemy $X = X_1 \oplus X_2 \oplus \ldots \oplus X_n$.

1.33. ZADANIE. Wykazać, że przestrzeń $C[-1,1]$ jest sumą prostą dwóch swoich domkniętych podprzestrzeni, złożonych odpowiednio z funkcji parzystych i funkcji nieparzystych.

Zadania uzupełniające

1.34. Dowieść, że przestrzeń ℓ^1 ma nieprzeliczalną bazę Hamela.

1.35. W przestrzeni $C^k(\mathbb{R})$ złożonej z funkcji ograniczonych mających ciągłe i ograniczone pochodne do rzędu k włącznie wprowadzić tak normę, by stała się przestrzenią Banacha.

1.36. Dowieść, że przestrzenie $L^p(\mathbb{R})$ i $L^p(0,1)$, $1 \leqslant p < \infty$, są izometrycznie izomorficzne.

1.37. Czy przestrzenie $L^1(0,1)$ oraz $L^1(0,1) \times L^1(0,1)$ są izomorficzne?

1.38. Załóżmy, że X i Y są podprzestrzeniami pewnej przestrzeni liniowej. Dowieść, że przestrzenie $(X+Y)/Y$ oraz $X/(X \cap Y)$ są algebraicznie izomorficzne.

1.39. Jeżeli Y jest domkniętą podprzestrzenią przestrzeni Banacha X, to X/Y jest przestrzenią Banacha. Odwzorowanie kanoniczne $\tau : X \to X/Y$ określone wzorem $\tau(x) = [x]$ jest ciągłe i otwarte (obraz zbioru otwartego jest otwarty).

ROZDZIAŁ II

PRZESTRZENIE FUNKCJI CIĄGŁYCH

Przestrzenie funkcji ciągłych, obok przestrzeni Hilberta i przestrzeni typu L^p, to podstawowe klasy przestrzeni Banacha. Szczególne miejsce zajmują w teorii aproksymacji i teorii szeregów Fouriera.

Dwa twierdzenia Weierstrassa

Oba twierdzenia mówią o możliwości jednostajnej aproksymacji na przedziale $[a, b]$ funkcji ciągłych wielomianami, pierwsze wielomianami zwykłymi, a drugie wielomianami trygonometrycznymi.

2.1. Twierdzenie Weierstrassa. *Każdą funkcję ciągłą na przedziale* $[a,\, b]$ *można jednostajnie aproksymować wielomianami.*

Dowód: Transformacja liniowa

$$t = \frac{s-a}{b-a}, \quad s \in [a, b],$$

sprowadza zagadnienie do przedziału $[0, 1]$. Tu posłużymy się dowodem pochodzącym od Bernsteina.

Dla funkcji ciągłej x na przedziale $[0,\, 1]$ niech x_n oznacza jej n–ty **Wielomian Bernsteina**

$$x_n(t) = \sum_{k=0}^{n} \binom{n}{k} x\left(\tfrac{k}{n}\right) t^k (1-t)^{n-k}.$$

Pokażemy, że ciąg x_n zbiega do x jednostajnie na całym przedziale $[0,\, 1]$.

Ustalmy $\varepsilon > 0$. Ponieważ funkcja x jest ciągła jednostajnie, więc dla pewnego $\delta > 0$ nierówność $|t-s| < \delta$ pociąga $|x(t)-x(s)| < \varepsilon$. Korzystając z tożsamości $\sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} = 1$ łatwo otrzymujemy

$$\big|x(t) - x_n(t)\big| \leqslant \sum_{k=0}^{n} \binom{n}{k} \Big|x(t) - x\big(\tfrac{k}{n}\big)\Big| \, t^k (1-t)^{n-k}.$$

Jeżeli teraz osobno zsumujemy po zbiorze tych wskaźników k dla których $\big|t - \frac{k}{n}\big| < \delta$ i osobno po zbiorze pozostałych to

$$\begin{aligned}\big|x(t) - x_n(t)\big| &< \varepsilon \sum_{|t-\frac{k}{n}|<\delta} \binom{n}{k} t^k (1-t)^{n-k} + 2M \sum_{|t-\frac{k}{n}|\geqslant\delta} \binom{n}{k} t^k (1-t)^{n-k} \\ &< \varepsilon + \frac{2M}{\delta^2} \sum_{k=0}^{n} \binom{n}{k} \Big(t - \frac{k}{n}\Big)^2 t^k (1-t)^{n-k},\end{aligned}$$

gdzie $M = \max\limits_{0\leqslant t\leqslant 1} |x(t)|$. Z równości

$$\sum_{k=0}^{n} \binom{n}{k} \big(t - \tfrac{k}{n}\big)^2 t^k (1-t)^{n-k} = \frac{t\,(1-t)}{n} \tag{2.4}$$

daje to oszacowanie $\varepsilon + \dfrac{2M}{\delta^2}\dfrac{1}{n}$. A więc $\big|x(t) - x_n(t)\big| < 2\varepsilon$ jeśli tylko n jest dostatecznie duże.

Pozostaje weryfikacja równości (2.4). Najłatwiej można ją otrzymać ze wzorów

$$\sum_{k=0}^{n} \binom{n}{k} t^k s^{n-k} = (t+s)^n,$$

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} t^k s^{n-k} = t\,(t+s)^{n-1},$$

$$\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} t^k s^{n-k} = \frac{1}{n} t (t+s)^{n-1} + \frac{n-1}{n} t^2 (t+s)^{n-2},$$

po podstawieniu $s = 1-t$, pomnożeniu kolejno stronami przez t^2, $-2t$ oraz 1 i dodaniu do siebie. Pierwszy ze wzorów, to oczywiście wzór Newtona, drugi powstaje z pierwszego przez zróżniczkowanie po zmiennej t i pomnożeniu stronami przez $\frac{1}{n}\,t$, a trzeci w ten sam sposób z drugiego. $\square$

Drugie z twierdzeń Weierstrassa jest odpowiednikiem pierwszego dla wielomianów trygonometrycznych.

2.2. Drugie twierdzenie Weierstrassa. *Wielomiany trygonometryczne*

$$a_0 + \sum_{k=1}^{n} \big(a_k \cos kt + b_k \sin kt\big)$$

leżą gęsto w zbiorze wszystkich funkcji ciągłych okresowych o okresie 2π

Dowód: Załóżmy, że x jest funkcją ciągłą okresową o okresie 2π. Jeżeli x jest parzysta, to możemy traktować ją jako funkcję na przedziale $[0,\pi]$ i zamienić na funkcję $y \in C[-1,1]$ wzorem

$$y(\cos t) = x(t), \quad t \in [0,\pi],$$

a tą, na mocy twierdzenia Weierstrassa, aproksymować jednostajnie na wielomianami. Wielomian od $\cos t$ można sprowadzić do postaci wielomianu trygonometrycznego parzystego. To rozumowanie dowodzi tezy dla funkcji parzystych. Jeśli x jest nieparzysta, to funkcja

$$x_1(t) = x(t)\sin t$$

jest parzysta, więc na mocy udowodnionej już części tezy, można ją jednostajnie aproksymować wielomianami trygonometrycznymi. Wynika stąd, że jeśli x jest funkcją dowolną, to możliwość jednostajnej aproksymacji wielomianami trygonometrycznymi zachodzi w każdym razie dla funkcji $x(t)\sin t$, a co tym idzie, także dla funkcji

$$x(t)\sin^2 t.$$

Przyjęcie tutaj funkcji $x\big(\frac{\pi}{2} - t\big)$ zamiast $x(t)$ i zamiana zmiennej t na $\frac{\pi}{2} - t$ (taka zamiana zachowuje zbiór wielomianów trygonometrycznych), daje możliwość aproksymacji wielomianami trygonometrycznymi funkcji

$$x(t)\cos^2 t,$$

zatem także funkcji $x(t) = x(t)\sin^2 t + x(t)\cos^2 t$. $\square$

Twierdzenie Stone'a

Oba przedstawione wyżej twierdzenia Weierstrassa są szczególnymi przypadkami znacznie ogólniejszego twierdzenia Stone'a, które udowodnimy.

Niech S będzie zwartą przestrzenia Hausdorffa. Jak wiemy Zbiór $C(S)$ jest przestrzenią Banacha w normie

$$\|x\| = \max_{t\in S} \bigl|x(t)\bigr|.$$

Zauważmy, że w istocie zbiór $C(S)$ ma strukturę algebry, wraz z funkcjami x, y zawiera ich iloczyn xy, nadto $\|xy\| \leqslant \|x\|\,\|y\|$. W przestrzeni $C_{\mathbb{R}}(S)$ wszystkich funkcji ciągłych na S o wartościach rzeczywistych można dodatkowo rozpatrywać działania $\vee$ i $\wedge$ określone wzorami

$$(x \vee y)(t) = \max\{x(t), y(t)\}, \qquad (x \wedge y)(t) = \min\{x(t), y(t)\}.$$

2.3. LEMAT. *Niech $\mathcal{A}$ będzie podzbiorem przestrzeni $C_{\mathbb{R}}(S)$ zamkniętym na działania $\vee$ i $\wedge$. Jeżeli dane są $x \in \mathcal{A}$ oraz $\varepsilon > 0$, a dla dowolnych $s, t \in S$ istnieje taka funkcja $y_{su} \in \mathcal{A}$, że*

$$\bigl|x(s) - y_{su}(s)\bigr| < \varepsilon, \qquad \bigl|x(u) - y_{su}(u)\bigr| < \varepsilon,$$

to w $\mathcal{A}$ istnieje też taka funkcja y, że

$$\bigl|x(t) - y(t)\bigr| < \varepsilon$$

dla wszystkich $t \in S$.

Dowód: Oznaczmy przez U_{su} i V_{su} zbiory tych punktów t, dla których odpowiednio $y_{su}(t) < x(t) + \varepsilon$ i $y_{su}(t) > x(t) - \varepsilon$. Są to zbiory otwarte, zawierające oba punkty s, u. Przy ustalonym u zbiory U_{su} tworzą pokrycie przestrzeni zwartej S, wybierając z tego pokrycia skończone podpokrycie $\{U_{s_1u}, U_{s_2u}, \ldots, U_{s_nu}\}$ i kładąc $y_u = y_{s_1u} \wedge y_{s_2u} \cdots \wedge y_{s_nu}$, otrzymamy funkcję $y_u \in \mathcal{A}$, spełniającą warunki

$$y_u(t) < x(t) + \varepsilon \quad \text{dla wszystkich} \quad t \in S,$$
$$y_u(t) > x(t) - \varepsilon \quad \text{dla} \quad t \in V_u = \bigcap_{k=1}^{n} V_{s_ku}.$$

Wybierzmy teraz z pokrycia przestrzeni S zbiorami V_u skończone podpokrycie $\{V_{u_1}, V_{u_2}, \ldots, V_{u_m}\}$ i połóżmy $y = y_{u_1} \vee y_{u_2} \cdots \vee y_{u_m}$. Otrzymamy w ten sposób funkcję $y \in \mathcal{A}$ spełniającą obie nierówności $x(t) - \varepsilon < y(t) < x(t) + \varepsilon$, funkcję której oczekiwaliśmy. $\square$

2.4. LEMAT. *Jeżeli domknięta podalgebra $\mathcal{A}$ algebry $C_{\mathbb{R}}(S)$ zawiera funkcję stałą 1, to jest zamknięta na działania $\vee$ i $\wedge$.*

Dowód: Ponieważ

$$x \vee y = \tfrac{1}{2}\big(x + y + |x - y|\big), \qquad x \wedge y = \tfrac{1}{2}\big(x + y - |x - y|\big),$$

więc wystarczy pokazać, że wraz z funkcją x zbiór $\mathcal{A}$ zawiera także funkcję $|x|$. Możemy przy tym dodatkowo założyć, że $\|x\| \leqslant 1$. Wtedy

$$\big|x(t)\big| = \sqrt{1 - \big(1 - x^2(t)\big)} = 1 - \sum_{n=1}^{\infty} \frac{(2n-3)!!}{(2n)!!}\big(1 - x^2(t)\big)^n,$$

a szereg jest zbieżny na S jednostajnie, zatem $|x| \in \mathcal{A}$. $\square$

Możemy teraz przystąpić do dowodu zapowiadanego twierdzenia Stone'a. Ustalmy przedtem, że o zbiorze $\mathcal{A}$ funkcji na S powiemy, iż **rozdziela punkty** przestrzeni S, gdy dla dowolnych dwóch różnych punktów $s, t \in S$ istnieje pewna funkcja $x \in \mathcal{A}$, spełniająca $x(s) \neq x(t)$.

2.5. TWIERDZENIE STONE'A. *Niech S będzie zwartą przestrzenią Hausdorffa oraz $\mathcal{A}$ domkniętą podalgebrą algebry $C(S)$. Jeżeli $\mathcal{A}$ rozdziela punkty S, zawiera funkcję stałą 1 i jest zamknięta na operację sprzężenia $x \to \overline{x}$, to $\mathcal{A} = C(S)$.*

Dowód: Oznaczmy przez $\mathcal{A}_{\mathbb{R}}$ zbiór wszystkich funkcji rzeczywistych $x \in \mathcal{A}$. Wtedy $\mathcal{A}_{\mathbb{R}}$ jest domkniętą podalgebrą algebry $C_{\mathbb{R}}(S)$, zawierającą funkcję 1. Z lematu 2.4 wynika, że zbiór $\mathcal{A}_{\mathbb{R}}$ jest zamknięty na działania $\vee$ i $\wedge$. Jeżeli $s, u \in S$, $s \neq u$, to $y(s) \neq y(u)$ dla pewnej funkcji $y \in \mathcal{A}$. Ponieważ obie funkcje

$$\operatorname{Re} y = \tfrac{1}{2}(y + \overline{y}), \qquad \operatorname{Im} y = \tfrac{1}{2i}(y - \overline{y})$$

należą do $\mathcal{A}_{\mathbb{R}}$, więc $\operatorname{Re} y(s) \neq \operatorname{Re} y(u)$ lub $\operatorname{Im} y(s) \neq \operatorname{Im} y(u)$, zatem $\mathcal{A}_{\mathbb{R}}$ rozdziela punkty przestrzeni S. Dla każdej funkcji $x \in C_{\mathbb{R}}(S)$ i każdej pary punktów $s, u \in S$ istnieje taka funkcja $y_{su} \in \mathcal{A}_{\mathbb{R}}$, że $y_{su}(s) = x(s)$, $y_{su}(u) = x(u)$. Wystarczy w tym celu w $\mathcal{A}_{\mathbb{R}}$ obrać taką funkcję y, że $y(s) \neq y(u)$ i położyć

$$y_{su}(t) = \frac{x(s) - x(u)}{y(s) - y(u)}\, y(t) - \frac{x(s)y(u) - x(u)y(s)}{y(s) - y(u)}\,.$$

Z lematu 2.3 wnioskujemy, że $\mathcal{A}_{\mathbb{R}} = C_{\mathbb{R}}(S)$, a dalej że $\mathcal{A} = C(S)$. Dla $x \in C(S)$ mamy bowiem $\operatorname{Re} x, \operatorname{Im} x \in C_{\mathbb{R}}(S) = \mathcal{A}_{\mathbb{R}}$, zatem $x = \operatorname{Re} x + i \operatorname{Im} x \in \mathcal{A}$. $\square$

UWAGA. Założenie w twierdzeniu Stone'a, że podalgebra $\mathcal{A}$ wraz z funkcją x zawiera także funkcję sprzężoną $\overline{x}$ jest istotne. Widać to z poniższego przykładu.

2.6. PRZYKŁAD. Okrąg jednostkowy

$$\mathbb{T} = \{z \in \mathbb{C}; |z| = 1\}$$

z topologią odziedziczoną z $\mathbb{C}$ jest przestrzenią zwartą. Najmniejsza domknięta podalgebra $\mathcal{A} \subset C(\mathbb{T})$ zawierająca funkcje $x(z) = 1$ oraz $x(z) = z$ rozdziela punkty $\mathbb{T}$. Pokażemy, że dla funkcji $x_0(z) = \overline{z}$ zachodzi równość

$$\operatorname{dist}(x_0, \mathcal{A}) = 1.$$

Wielomiany $w(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$ leżą gęsto w $\mathcal{A}$ oraz

$$\|x_0 - w\|^2 \geqslant \frac{1}{2\pi} \int_0^{2\pi} \left|x_0(e^{it}) - w(e^{it})\right|^2 dt = 1 + \sum_{k=0}^{n} |a_k|^2 \geqslant 1.$$

Istotnie

$$\left|x_0(e^{it}) - w(e^{it})\right|^2 = 1 - \sum_{k=0}^{n} a_k e^{i(k+1)t} - \sum_{k=0}^{n} \overline{a_k} e^{-i(k+1)t} + \sum_{k=0}^{n} \sum_{r=0}^{n} a_k \overline{a_r} e^{i(k-r)t},$$

a całka $\int_0^{2\pi} e^{imt} dt$ jest równa 0, gdy $m \neq 0$ oraz równa 2π, gdy $m = 0$. Z tego wnosimy, że $\operatorname{dist}(x_0, \mathcal{A}) \geqslant 1$. Nierówność przeciwna jest oczywista.

Lemat Urysohna i twierdzenie Tietzego-Urysohna

Pomocnym narzędziem przy konstrukcji funkcji ciągłych na ogólnych przestrzeniach topologicznych są dwa poniższe twierdzenia. Wynika z nich w szczególności, że gdy S jest przestrzenią topologiczną normalną, to przestrzeń $C(S)$ jest niezdegenerowana.

Pierwsze twierdzenie tradycyjnie nosi nazwę lematu Urysohna i jego dowód zamieszczony jest w rozdziale Uzupełnienia **??**.

2.7. LEMAT URYSOHNA. *Jeżeli S jest przestrzenią topologiczną normalną, to dla każdej pary K_0, K_1 rozłącznych podzbiorów domkniętych S istnieje taka funkcja ciągła $x : S \to [0,1]$, że $x|_{K_0} = 0$ oraz $x|_{K_1} = 1$.*

2.8. Twierdzenie Tietzego-Urysohna. *Niech S_0 będzie podzbiorem domkniętym przestrzeni normalnej S. Dla każdej funkcji ciągłej $x_0 : S_0 \to [-1,1]$ istnieje przedłużenie ciągłe $x : S \to [-1,1]$.*

Dowód: Załóżmy, że $|x_0(s)| \leqslant c \leqslant 1$ dla $s \in S_0$. Podamy konstrukcję takiego ciągu $x_1, x_2, \ldots$ funkcji ciągłych na S, że

$$|x_n(s)| \leqslant \tfrac{1}{3}\big(\tfrac{2}{3}\big)^{n-1} \quad \text{dla} \quad s \in S$$

oraz

$$\Big|x_0(s) - \sum_{k=1}^{n} x_k(s)\Big| \leqslant \big(\tfrac{2}{3}\big)^{n} \quad \text{dla} \quad x \in S_0.$$

Ponieważ zbiory $K_0 = x_0^{-1}\big([-c, -\frac{1}{3}c]\big)$ i $K_1 = x_0^{-1}\big([\frac{1}{3}c, c]\big)$ są rozłączne i domknięte w S_0, a więc domknięte w S, to z lematu Urysohna **??** wynika istnienie takiej funkcji $k : S \to [-1,1]$, że $k|_{K_0} = 0$ i $k|_{K_1} = 1$. Łatwo sprawdzić, że funkcja $x_1(s) = \frac{2}{3}c\,\big(k(s) - \frac{1}{2}\big)$ spełnia żądane warunki. Funkcję x_2 konstruujemy w podobny sposób, przyjmując $x_0 - x_1$ w miejsce x_0 i $\frac{2}{3}c$ w miejsce c, itd.

Szereg $\sum_{k=1}^{\infty} x_k$ jest jednostajnie zbieżny. Jego suma x jest więc funkcją ciągłą na S, spełnia warunek $|x(t)| \leqslant 1$ dla $s \in S$ oraz $x(s) = x_0(s)$ dla $s \in S_0$, jest zatem poszukiwanym przedłużeniem funkcji x_0. $\square$

2.9. Wniosek. *Niech S_0 będzie podzbiorem domkniętym przestrzeni normalnej S. Każdą ograniczoną funkcję ciągłą $x_0 : S_0 \to \mathbb{C}$ można przedłużyć do ograniczonej funkcji ciągłej $x : S \to \mathbb{C}$ i to tak, by $\|x\|_\infty = \|x_0\|_\infty$.*

Dowód: Możemy założyć, że $\|x_0\|_\infty = 1$. Z twierdzenia Tietzego-Urysohna wynika, że funkcje $\operatorname{Re} x_0$ i $\operatorname{Im} x_0$ można przedłużyć do ograniczonych funkcji rzeczywistych na S. Nazwijmy je y_1 oraz y_2 i połóżmy $y = \max\{1, |y_1 + i\,y_2|\}$. Funkcja $x = (y_1 + i\,y_2)/y$ ma żądaną własność, gdyż $y|_{S_0} = 1$. $\square$

Twierdzenia o najlepszej aproksymacji

2.10. TWIERDZENIE KOŁMOGOROWA. *Niech S będzie zwartą przestrzenią Hausdorffa oraz W domkniętą podprzestrzenią liniową przestrzeni $C(S)$. Ustalmy funkcje $x_0 \in C(S)$ oraz $w_0 \in W$ i oznaczmy*

$$S_0 = \{t \in S : \big|x_0(t) - w_0(t)\big| = \|x_0 - w_0\|_\infty\}.$$

Funkcja w_0 najlepiej w W aproksymuje funkcję x_0 w normie jednostajnej wtedy i tylko wtedy, gdy dla każdej funkcji $w \in W$ zachodzi nierówność

$$\min_{t\in S_0} \operatorname{Re} \overline{(x_0(t) - w_0(t))}\, w(t) \leqslant 0. \tag{2.5}$$

Dowód: Niech na razie S_0 będzie dowolnym niepustym podzbiorem domkniętym S. Pokażemy, że własność (2. 5) pociąga

$$\operatorname{dist}(x_0, W) \geqslant \min_{t\in S_0} \big|x_0(t) - w_0(t)\big|.$$

Przyjmując za S_0 zbiór określony w twierdzeniu, otrzymamy wtedy $\operatorname{dist}(x_0, W) \geqslant \|x_0 - w_0\|_\infty$, a więc, że w_0 najlepiej w W aproksymuje x_0.

Załóżmy nie wprost, że $\operatorname{dist}(x_0, W) < \min_{t\in S_0} \big|x_0(t) - w_0(t)\big|$ a więc, że

$$\operatorname{dist}(x_0, W) < \|x_0 - w_1\|_\infty < \min_{t\in S_0} \big|x_0(t) - w_0(t)\big|$$

dla pewnej funkcji $w_1 \in W$, wtedy

$$\big|x_0(t) - w_1(t)\big| < \big|x_0(t) - w_0(t)\big|$$

dla wszystkich $t \in S_0$, a przyjmując w (2. 5) za w funkcję $w_1 - w_0$ otrzymamy

$$\begin{aligned}\operatorname{Re} \overline{(x_0 - w_0)}\,(w_1 - w_0) &= |x_0 - w_0|^2 - \operatorname{Re} \overline{(x_0 - w_0)}\,(x_0 - w_1)\\ &\geqslant |x_0 - w_0|^2 - |x_0 - w_0|\,|x_0 - w_1|\\ &= |x_0 - w_0|\Big(|x_0 - w_0| - |x_0 - w_1|\Big) > 0\end{aligned}$$

na S_0, co stoi w sprzeczności z (2. 5).

Dowód w drugą stronę także przeprowadzimy nie wprost. Załóżmy wobec tego, że

$$\min_{t\in S_0} \operatorname{Re} \overline{(x_0(t) - w_0(t))}\, w_1(t) = a > 0$$

dla pewnej funkcji $w_1 \in W$. Wybierzmy w S zbiór otwarty $U \supset S_0$ tak, aby

$$\operatorname{Re} \overline{(x_0(t) - w_0(t))}\, w_1(t) > \frac{a}{2}$$

dla $t \in U$ i oznaczmy

$$\|x_0 - w_0\|_\infty - \max_{t \in S \setminus U} \big|x_0(t) - w_0(t)\big| = b > 0.$$

Jeśli przyjmiemy

$$\delta = \min\Big\{\frac{a}{2c^2}, \frac{b}{2c}\Big\},$$

gdzie $c = \|w_1\|_\infty$, oraz $w_2 = w_0 + \delta\, w_1$, to dla $t \in U$ spełniona jest nierówność

$$\begin{aligned}\big|x_0(t) - w_2(t)\big|^2 &\leqslant \|x_0 - w_0\|_\infty^2 - 2\delta \operatorname{Re} \overline{(x_0(t) - w_0(t))}\, w_1(t) + \delta^2 \|w_1\|_\infty^2 \\ &\leqslant \|x_0 - w_0\|_\infty^2 - \frac{a\delta}{2},\end{aligned}$$

a dla $t \in S \setminus U$ nierówność

$$\big|x_0(t) - w_2(t)\big| \leqslant \big|x_0(t) - w_0(t)\big| + \delta\, \|w_1\|_\infty \leqslant \|x_0 - w_0\|_\infty - \frac{b}{2}.$$

To oznacza, że

$$\|x_0 - w_2\|_\infty < \|x_0 - w_0\|_\infty$$

tj., że funkcja w_2 daje lepszą aproksymację x_0 niż funkcja w_0. $\square$

Ze względu na liczne zastosowania praktyczne, szczególnie ważnym przypadkiem twierdzeń o najlepszej aproksymacji jest ten, w którym S jest domkniętym przedziałem na prostej, zaś W zbiorem wielomianów nie przekraczających ustalonego stopnia. Udowodnimy:

2.11. TWIERDZENIE. *Dla każdej funkcji $x_0 \in C[a, b]$ pośród wielomianów stopnia niższego niż n dokładnie jeden wielomian w_0 leży najbliżej x_0. W tym przypadku zbiór*

$$S_0 = \big\{t \in [a, b] : \big|x_0(t) - w_0(t)\big| = \|x_0 - w_0\|_\infty\big\}$$

ma przynajmniej $n + 1$ elementów.

Dowód: Przestrzeń W wielomianów stopnia niższego niż n ma wymiar n, a wielomianu w_0, najbliższego x_0, wystarczy szukać w zbiorze

$$\{w \in W : \|w\|_\infty \leqslant 2\, \|x_0\|_\infty\},$$

który jest zwarty. Funkcja $\|x_0 - w\|_\infty$ jest ciągła, osiąga więc na nim swój kres dolny. To dowodzi istnienia w_0. Także zbiór S_0 zawiera przynajmniej $n+1$ elementów, w przeciwnym przypadku w zbiorze W moglibyśmy wskazać wielomian w_1, dla którego

$$w_1(t) = x_0(t) - w_0(t), \quad t \in S_0$$

i po wstawieniu w_1 do (2.5) w miejsce w otrzymalibyśmy

$$\min_{t \in S_0} \operatorname{Re} \overline{(x_0(t) - w_0(t))}\, w_1(t) = \min_{t \in S_0} \left| x_0(t) - w_0(t) \right|^2 > 0.$$

Dodajmy, że jeśli $S_0 = \{t_1, t_2, \ldots, t_{m+1}\}$ oraz $m < n$, to za w_1 można wybrać m-ty wielomian interpolacyjny Lagrange'a dla funkcji $x_0 - w_0$

$$w_1(t) = \sum_{k=1}^{m+1} \frac{\varphi(t)\big(x_0(t_k) - w_0(t_k)\big)}{(t - t_k)\,\varphi'(t_k)}, \quad \text{gdzie} \quad \varphi(t) = \prod_{k=1}^{m+1} (t - t_k).$$

Do dowodu pozostaje jedyność wielomianu w_0. Gdy wielomiany w_0' i w_0'' ze zbioru W dają najlepszą aproksymację funkcji x_0, to daje ją też wielomian

$$w_0 = \frac{1}{2} w_0' + \frac{1}{2} w_0''. \tag{2.6}$$

W zbiorze S_0 punktów ekstremalnych funkcji $x_0 - w_0$ wybierzmy $n+1$ punktów $t_1, t_2, \ldots, t_{n+1}$. Mamy wtedy $x_0(t_k) - w_0(t_k) = d\, e^{i s_k}$ dla $d = \operatorname{dist}(x_0, W)$ i pewnych liczb rzeczywistych $s_1, s_2, \ldots, s_{n+1}$. Ponieważ

$$\left| x_0(t_k) - w_0'(t_k) \right| \leqslant d, \quad \left| x_0(t_k) - w_0''(t_k) \right| \leqslant d$$

dla $k = 1, 2, \ldots, n+1$, więc z określenia (2.6) wynika, że

$$x_0(t_k) - w_0'(t_k) = x_0(t_k) - w_0''(t_k) = d\, e^{i s_k},$$

tj. $w_0'(t_k) - w_0''(t_k) = 0$ dla $k = 1, 2, \ldots, n+1$, a ponieważ niezerowy wielomian w W nie może mieć więcej niż n zer, więc musi być $w_0' = w_0''$. □

Mimo, że w klasie wielomianów stopnia niższego niż n, jak widać z twierdzenia 2.11, wielomian w_0 najlepiej aproksymujący ustaloną funkcję x_0 na przedziale $[a, b]$ jest jedyny, nie jest znany żaden algorytm pozwalający wyznaczyć w_0. Jest bardzo niewiele przykładów, w których w_0 można wskazać konkretnie. Oto jeden z nich.

2.12. PRZYKŁAD. Pokażemy, że na przedziale $[-1, 1]$ pośród wielomianów stopnia niższego niż n najlepiej aproksymującym jednomian t^n jest wielomian (stopnia $n-2$)

$$w_0(t) = t^n - T_n(t), \tag{2.7}$$

gdzie T_n jest n-tym **wielomianem Czebyszewa**

$$T_0(t) = 1, \quad T_n(t) = \frac{1}{2^{n-1}} \cos(n \arccos t), \quad n = 1, 2, 3, \dots .$$

Wykresy funkcji $\cos(n \arccos t)$ dla $n = 0, 1, 2, 3, 4, 5$.

Kolejne wielomiany Czebyszewa mają postać $T_0(t) = 1$, $T_1(t) = t$, $T_2(t) = t^2 - \frac{1}{2}$, następne można łatwo wyznaczyć ze wzoru rekurencyjnego

$$T_n(t) = t\,T_{n-1}(t) - \tfrac{1}{4}\,T_{n-2}(t) \quad \text{dla} \quad n = 3, 4, 5, \dots,$$

zatem $T_3(t) = t^3 - \frac{3}{4}t$, $T_4(t) = t^4 - \frac{5}{4}t^2 + \frac{1}{4}$, itd. Widać, że w istocie T_n jest wielomianem stopnia n ze współczynnikiem przy najwyższej potędze równym 1, zatem funkcja w_0, określona wzorem (2.7), jest wielomianem stopnia $n-2$.

Zbiór S_0 punktów ekstremalnych funkcji $x_0 - w_0 = T_n$ składa się z liczb

$$t_k = \arccos\left(\frac{k\pi}{n}\right), \quad k = 0, 1, 2, \dots, n.$$

Gdyby w_0 nie był wielomianem optymalnym, to na mocy twierdzenia Kołmogorowa mielibyśmy

$$\min_{0 \leqslant k \leqslant n} (-1)^k \operatorname{Re} w(t_k) > 0$$

dla pewnego wielomianu w stopnia niższego niż n. Wielomian $\operatorname{Re} w$ musiałby wtedy zmieniać znak w każdym z przedziałów (t_k, t_{k+1}), a taki wielomian musi mieć stopień przynajmniej n.

2.13. Zadanie. Dowieść, że pośród wielomianów stopnia niższego niż n wielomian w_0 najlepiej przybliża funkcję rzeczywistą $x_0 \in C[a,b]$ wtedy i tylko wtedy, gdy sam jest rzeczywisty a w przedziale $[a,b]$ można tak wybrać podzbiór $t_1 < t_2 < \cdots < t_{n+1}$, zwany **alternansem**, że funkcja błędu $x_0 - w_0$ przyjmuje na nim na przemian wartości $\|x_0 - w_0\|_\infty$ oraz $-\|x_0 - w_0\|_\infty$ (niekoniecznie rozpoczynając od wartości dodatniej).

Przedstawimy jeszcze jeden przykład najlepszej aproksymacji funkcji przez wielomiany. Przykład ten pochodzi od Bernsteina, można go znaleźć w wielu podręcznikach i monografiach z teorii aproksymacji, wymaga jednak znajomości dość zaawansowanych technik teorii funkcji zmiennej zespolonej. Poniższa prezentacja jest dłuższa, za to elementarna.

2.14. Przykład. Znajdziemy wielomiany najlepiej aproksymujące jednostajnie na przedziale $[-1,1]$ funkcję $x_0(t) = \dfrac{1}{a-t}$, gdzie a jest ustaloną liczbą rzeczywistą, większą niż 1.

Ustalmy $n = 0, 1, 2, \ldots$ i rozważmy funkcję $y_n : [-1,1] \to \mathbb{C}$ określoną wzorem

$$y_n(t) = z^n \frac{z-\alpha}{1-\alpha z},$$

gdzie $z = t + i\sqrt{1-t^2}$ oraz $\alpha = a - \sqrt{a^2-1}$. Ponieważ $|z| = 1$, więc $\left|\dfrac{z-\alpha}{1-\alpha z}\right| = 1$ oraz

$$\operatorname{Re} \frac{z-\alpha}{1-\alpha z} = \frac{(1+\alpha^2)t - 2\alpha}{1 - 2\alpha t + \alpha^2} = \frac{at-1}{a-t} = (a^2-1)\, x_0(t) - a.$$

Wynika stąd, że $\big|y_n(t)\big| \equiv 1$ oraz

$$\begin{aligned} x_0(t) - \frac{\alpha^n}{a^2-1} \operatorname{Re} y_n(t) &= \frac{a}{a^2-1} + \frac{1}{a^2-1} \operatorname{Re}\big(1 - \alpha^n z^n\big) \frac{z-\alpha}{1-\alpha z} \\ &= \frac{a}{a^2-1} + \frac{1}{a^2-1} \operatorname{Re}(z-\alpha)\big(1 + \alpha z + \ldots + \alpha^{n-1} z^{n-1}\big). \end{aligned} \tag{2.8}$$

Ponieważ $\operatorname{Re} z^k = \cos(k \arccos t) = 2^{k-1} T_k(t)$ jest wielomianem Czebyszewa stopnia k, więc $x_0 - \dfrac{\alpha^n}{a^2-1} \operatorname{Re} y_n$ jest wielomianem stopnia n, który oznaczymy przez w_n. Ze wzoru (2.8) otrzymujemy $w_0(t) = \dfrac{a}{a^2-1}$ oraz

$$w_n(t) = \frac{(2\alpha)^{n-1}}{a^2-1} T_n(t) + \frac{1}{\sqrt{a^2-1}} \sum_{k=0}^{n-1} (2\alpha)^k T_k(t) \tag{2.9}$$

dla $n = 1, 2, 3, \ldots$.

Twierdzimy, że spośród wszystkich wielomianów stopnia $\leqslant n$ to właśnie wielomian w_n daje najlepszą aproksymację funkcji x_0. By tego dowieść, wystarczy dla funkcji błędu $x_0 - w_n = \frac{\alpha^n}{a^2-1}\operatorname{Re} y_n$ wskazać alternans długości $n+2$.

Dla $u = \arg z = \arccos t$ oznaczmy

$$\varphi(u) = \arg \frac{z-\alpha}{1-z\alpha} = \arg(e^{iu} - \alpha) - \arg(\tfrac{1}{\alpha} - e^{iu}).$$

Widać (patrz rysunek), że $\varphi(u)$ rośnie od 0 do π na przedziale $[0, \pi]$. Dlatego $\arg y_n(\cos u) = nu + \varphi(u)$ rośnie od 0 do $(n+1)\pi$ na tym przedziale i można wskazać takie punkty

$$0 = u_0 < u_1 < u_2 < \cdots < u_{n+1} = \pi,$$

dla których $\arg y_n(\cos u_k) = k\pi$, $k = 0, 1, 2, \ldots, n+1$, tj. $y_n(\cos u_k) = (-1)^k$. Stwierdzamy zatem, że układ punktów $\cos u_{n+1}, \cos u_n, \ldots, \cos u_0$ jest szukanym alternansem funkcji $x_0 - w_n$.

Kolejne funkcje błędu dla $a = \frac{5}{4}$ i $n = 0, 1, 2, 3, 4, 5$.

Zadania uzupełniające

2.15. Dla $n = 0, 1, 2, 3, 4$ obliczyć n-ty wielomian Bernsteina funkcji $x(t) = \cos \pi t$.

2.16. Czy potrafisz wskazać konkretny ciąg wielomianów zbieżny jednostajnie do funkcji $x(t) = \cos \pi t$ na przedziale $[0, 1]$?

2.17. Obliczyć n-ty wielomian Bernsteina funkcji: $x(t) = t^2$, $x(t) = t^3$.

2.18. Uogólnić wzór na n-ty wielomian Bernsteina dla funkcji określonych na dowolnym przedziale $[a, b]$.

2.19. Dowieść, że jeżeli przestrzenie topologiczne normalne S_1 i S_2 są homeomorficzne, to $C(S_1)$ i $C(S_2)$ są izometrycznie izomorficzne. Wykazać na przykładzie, że twierdzenie odwrotne nie zawsze jest prawdziwe.

2.20. Niech X będzie najmniejszą domkniętą podprzestrzenią przestrzeni $C(\mathbb{R})$ zawierającą wszystkie ograniczone funkcje jednostajnie ciągłe. Czy $X = C(\mathbb{R})$?

ROZDZIAŁ III

PRZESTRZENIE HILBERTA

Przestrzenie unitarne

Przestrzenią unitarną nazywamy przestrzeń liniową $\mathcal{H}$ nad ciałem $\mathbb{C}$ liczb zespolonych wraz z zespoloną funkcją $\langle\ ,\ \rangle$ określoną na $\mathcal{H}\times\mathcal{H}$ i mającą następujące własności:

(1) $\langle x, x\rangle \geqslant 0$ dla $x \in \mathcal{H}$ oraz $\langle x, x\rangle = 0$ wtedy i tylko wtedy, gdy $x = 0$.

(2) $\langle x, x\rangle \geqslant 0$ dla $x \in \mathcal{H}$.

(3) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ dla $x, y, z \in \mathcal{H}$ (addytywność).

(4) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ dla $\lambda \in \mathbb{C}$, $x, y \in \mathcal{H}$ (jednorodność).

(5) $\langle y, x\rangle = \overline{\langle x, y\rangle}$ dla $x, y \in \mathcal{H}$ (hermitowskość).

Funkcja $\langle\ ,\ \rangle$ nosi nazwę **iloczynu skalarnego** w $\mathcal{H}$. Normę w przestrzeni unitarnej $\mathcal{H}$ określamy wzorem

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

Przestrzeń zupełną w normie $\|\ \|$ nazywamy **przestrzenią Hilberta**.

3.1. Przykład. Na przestrzeni $\mathbb{C}^n$ funkcja $\langle\ ,\ \rangle$ określona wzorem

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k\,\overline{y_k},$$

gdzie $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$, jest iloczynem skalarnym. Przestrzeń $\mathbb{C}^n$ ma skończony wymiar, jest zatem przestrzenią Hilberta. Jest to dobrze nam znana przestrzeń euklidesowa.

3.2. Przykład. Także przestrzeń ℓ^2 z iloczynem skalarnym

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k\,\overline{y_k}$$

jest przestrzenią Hilberta. Zauważmy, że $|x_k\,\overline{y_k}| \leqslant \frac{1}{2}\,|x_k|^2 + \frac{1}{2}\,|y_k|^2$, a więc szereg powyższy jest bezwzględnie zbieżny.

3.3. PRZYKŁAD. Przestrzeń $C\big([0,1]\big)$ z iloczynem skalarnym

$$\langle x,\, y\rangle = \int_0^1 x(t)\,\overline{y(t)}\,dt$$

jest przestrzenią unitarną. Nie jest jednak przestrzenią zupełną w normie

$$\|x\|_2 = \Big(\int_0^1 |x(t)|^2\,dt\Big)^{1/2}.$$

Po uzupełnieniu otrzymamy przestrzeń Hilberta $L^2\big([0,1]\big)$ funkcji całkowalnych z kwadratem w sensie Lebesgue'a na przedziale $[0,1]$.

3.4. PRZYKŁAD. Zamiast przestrzeń $L^2\big([0,1]\big)$ z poprzedniego przykładu można badać bardzo ogólną klasę przestrzeni $L^2(\Omega,\mathcal{B},\mu)$, gdzie $(\Omega,\mathcal{B},\mu)$ jest dowolną przestrzenią miarową σ-skończoną. Są to wszystko przestrzenie Hilberta a iloczyn skalarny ma postać

$$\langle x,\, y\rangle = \int_\Omega x(t)\,\overline{y(t)}\,d\mu(t).$$

Każdą z wcześniejszych przestrzeni można zrealizować jako szczególny przypadek przestrzeni $L^2(\Omega,\mathcal{B},\mu)$. Przestrzeń $\mathbb{C}^n$ otrzymamy biorąc $\Omega=\{1,2,\ldots,n\}$, za $\mathcal{B}$ σ-ciało wszystkich podzbiorów Ω a za μ miarę liczącą ilość elementów zbioru; przestrzeń ℓ^2 biorąc $\Omega=\mathbb{N}$ a $\mathcal{B}$ i μ jak wyżej, zaś $L^2\big([0,1]\big)$ biorąc $\Omega=[0,1]$, za $\mathcal{B}$ σ-ciało wszystkich podzbiorów borelowskich $[0,1]$ a za μ miarę Lebesgue'a na $[0,1]$.

3.5. PRZYKŁAD. Podamy jeszcze jeden ważny przykład przestrzeni Hilberta. Oznaczmy przez H^2 zbiór wszystkich tych funkcji holomorficznych f na dysku jednostkowym $D=\{z\in\mathbb{C}:|z|<1\}$, dla których

$$\|f\| = \Bigg(\sup_{0\leqslant r<1}\frac{1}{2\pi}\int_0^{2\pi}\big|f(r\,e^{it})\big|^2\,dt\Bigg)^{1/2} < \infty.$$

Pokażemy, że przestrzeń H^2 jest zupełna i można w niej tak wprowadzić iloczyn skalarny $\langle\ ,\ \rangle$, by $\|f\|=\sqrt{\langle f,\, f\rangle}$.

Jak wiadomo, każdą funkcję f holomorficzną na D można przedstawić w postaci zbieżnego szeregu potęgowego

$$f(z) = \sum_{n=0}^{\infty} a_n\, z^n.$$

Mamy wtedy

$$\frac{1}{2\pi}\int_0^{2\pi}\left|f(r\,e^{it})\right|^2 dt = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} r^{n+k}\,a_n\,\overline{a_k}\,\frac{1}{2\pi}\int_0^{2\pi} e^{i(n-k)t}\,dt = \sum_{n=0}^{\infty} r^{2n}\,|a_n|^2.$$

Wyrażenie $\sup_{0\leqslant r<1}$ w definicji normy $\|\ \|$ można więc zastąpić przez $\lim_{r\to 1-}$ otrzymując

$$\|f\| = \Big(\lim_{r\to 1-}\sum_{n=0}^{\infty} r^{2n}\,|a_n|^2\Big)^{1/2} = \Big(\sum_{n=0}^{\infty}|a_n|^2\Big)^{1/2},$$

Oznacza to, że przestrzeń H^2 jest izometrycznie izomorficzna z przestrzenią ℓ^2, jest zatem przestrzenią Hilberta a iloczyn skalarny do niej można przenieść z przestrzeni ℓ^2. Otrzymujemy

$$\langle f,\,g\rangle = \lim_{r\to 1-}\frac{1}{2\pi}\int_0^{2\pi} f(r\,e^{it})\,\overline{g(r\,e^{it})}\,dt.$$

Podamy teraz kilka prostych twierdzeń o przestrzeniach unitarnych.

3.6. Twierdzenie Pitagorasa. *Jeśli* $\langle x,\,y\rangle = 0$, *to* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Dowód: Korzystając z własności (2) – (4) iloczynu skalarnego otrzymujemy

$$\|x+y\|^2 = \langle x+y,\,x+y\rangle = \langle x,\,x\rangle + \langle x,\,y\rangle + \overline{\langle x,\,y\rangle} + \langle y,\,y\rangle = \\ = \|x\|^2 + \|y\|^2. \quad \square$$

W interpretacji geometrycznej kąt α, utworzony przez wektory x i y, to kąt przy wierzchołku 0 trójkąta o bokach długości $\|x\|$, $\|y\|$ i $\|x-y\|$, zatem na mocy twierdzenia cosinusów

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\,\|x\|\,\|y\|\,\cos\alpha.$$

Z drugiej strony, z definicji iloczynu skalarnego, otrzymujemy

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\,\mathrm{Re}\langle x,\,y\rangle.$$

Porównanie tych wielkości daje następującą interpretację kąta α:

$$\cos\alpha = \frac{\mathrm{Re}\langle x,\,y\rangle}{\|x\|\,\|y\|}.$$

W interpretacji tej trójkąt, rozpatrywany w twierdzeniu Pitagorasa 3.6, to trójkąt prostokątny.

3.7. NIERÓWNOŚĆ SCHWARZA. *Dla dowolnych* $x, y \in \mathcal{H}$ *zachodzi nierówność*

$$\big|\langle x,\, y\rangle\big| \leqslant \|x\|\,\|y\|.$$

Dowód: Dla $y = 0$ nierówność jest oczywista, dla $y \neq 0$ zaś wynika bezpośrednio z nierówności

$$0 \leqslant \langle x - \lambda y,\, x - \lambda y\rangle = \|x\|^2 - \overline{\lambda}\langle x,\, y\rangle - \lambda\,\overline{\langle x,\, y\rangle} + |\lambda|^2\|y\|^2,$$

jeśli podstawić $\lambda = \dfrac{\langle x,\, y\rangle}{\|y\|^2}$. $\square$

Z nierówności Schwarza łatwo wynika, że funkcjonał $\|x\| = \sqrt{\langle x,\, x\rangle}$ istotnie jest normą w $\mathcal{H}$. Mamy bowiem

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x,\, y\rangle \leqslant \\ &\leqslant \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| = \big(\|x\| + \|y\|\big)^2,\end{aligned}$$

co dowodzi podaddytywności normy. Pozostałe postulaty normy są spełnione w sposób oczywisty.

Z nierówności Schwarza wynika też, że iloczyn skalarny $\langle\ ,\ \rangle$ jest ciągły jako funkcja dwóch zmiennych. Rzeczywiście, jeżeli $x_n \to x$ i $y_n \to y$, to

$$\big|\langle x_n,\, y_n\rangle - \langle x,\, y\rangle\big| \leqslant \|x_n - x\|\,\|y\| + \|x\|\,\|y_n - y\| + \|x_n - x\|\,\|y_n - y\| \to 0.$$

3.8. WZÓR POLARYZACYJNY. *Dla dowolnych* $x, y \in \mathcal{H}$ *zachodzi równość*

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

W rzeczywistej przestrzeni Hilberta wzór polaryzacyjny przyjmuje postać

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

3.9. RÓWNOŚĆ RÓWNOLEGŁOBOKU. *Dla dowolnych* $x, y \in \mathcal{H}$ *zachodzi równość*

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2. \quad \square$$

Nazwa „równość równoległoboku” pochodzi od znanego twierdzenia geometrii mówiącego, że suma kwadratów długości przekątnych równoległoboku jest równa sumie kwadratów długości jego boków.

Równość równoległoboku charakteryzuje przestrzenie unitarne, mianowicie:

3.10. ZADANIE. Jeżeli w przestrzeni unormowanej X równość równoległoboku zachodzi dla każdej pary wektorów, to wzór polaryzacyjny 3.8 określa w niej iloczyn skalarny zgodny z normą, tzn. $\|x\| = \sqrt{\langle x, x\rangle}$, zatem X jest przestrzenią unitarną.

Twierdzenia o najlepszej aproksymacji

Twierdzenia, które teraz udowodnimy, dotyczą centralnych zagadnień teorii optymalizacji w przestrzeniach Hilberta.

Przedtem przypomnijmy, że podzbiór W przestrzeni liniowej jest nazywany wypukłym, gdy wraz z każdymi dwoma punktami x i y zawiera cały odcinek

$$\big\{(1-\lambda)x + \lambda y : 0 \leqslant \lambda \leqslant 1\big\}$$

łączący te punkty.

3.11. TWIERDZENIE (O NAJLEPSZEJ APROKSYMACJI). *Niech $\mathcal{H}$ będzie przestrzenią Hilberta oraz W jej domkniętym podzbiorem wypukłym. Dla każdego $x_0 \in \mathcal{H}$ istnieje dokładnie jeden taki element $w_0 \in W$, że*

$$\|x_0 - w_0\| = \operatorname{dist}(x_0, W) = \inf_{w\in W} \|x_0 - w\|. \tag{3.10}$$

Dowód: Oznaczmy $d = \operatorname{dist}(x_0, W)$ i wybierzmy w zbiorze W taki ciąg $\{w_n\}$, że $\|x_0 - w_n\| \to d$. Pokażemy, że jest to ciąg zbieżny do pewnego elementu $w_0 \in W$. Z równości równoległoboku mamy

$$\|w_n - w_m\|^2 = 2\,\|x_0 - w_n\|^2 + 2\,\|x_0 - w_m\|^2 - 4\,\big\|(\tfrac{1}{2}w_n + \tfrac{1}{2}w_m) - x_0\big\|^2.$$

Ponieważ $\frac{1}{2}w_n + \frac{1}{2}w_m \in W$, więc $\|(\frac{1}{2}w_n + \frac{1}{2}w_m) - x_0\| \geqslant d$, stąd

$$\|w_n - w_m\|^2 \leqslant 2\,\|x_0 - w_n\|^2 + 2\,\|x_0 - w_m\|^2 - 4d^2.$$

Prawa strona tej nierówności dąży do zera, gdy $n, m \to \infty$. Oznacza to, że $\{w_n\}$ jest ciągiem Cauchy'ego w $\mathcal{H}$, jest zatem zbieżny do pewnego elementu w_0. Z domkniętości zbioru W wnosimy, że $w_0 \in W$.

Jest oczywiste, że $\|x_0 - w_0\| = d$. Jeśli równość taka jest spełniona dla dwóch elementów $w_0', w_0'' \in W$, to przyjmując jako ciąg $\{w_n\}$ ciąg $\{w_0', w_0'', w_0', w_0'', \ldots\}$, który musi być zbieżny, otrzymujemy $w_0' = w_0''$. □

Twierdzenie o najlepszej aproksymacji istotnie wyróżnia przestrzenie Hilberta spośród przestrzeni Banacha. Ilustruje to następujący przykład:

3.12. PRZYKŁAD. Ustalmy niezerowy wektor $y = (y_1, y_2, \ldots) \in \ell^\infty$ i określmy funkcję $\varphi : \ell^1 \to \mathbb{C}$ wzorem

$$\varphi(x) = \sum_{n=1}^{\infty} x_n\, y_n, \quad \text{gdy} \quad x = (x_1, x_2, \ldots).$$

Jest to funkcja liniowa i ciągła, liniowość jest oczywista, a ciągłość wynika z nierówności $|\varphi(u) - \varphi(v)| \leqslant \|u - v\|_1\, \|y\|_\infty$; zatem zbiór

$$W = \Big\{x \in \ell^1 : \sum_{n=1}^{\infty} x_n\, y_n = 1\Big\}$$

jest wypukły i domknięty w ℓ^1. Zbadamy czy w W istnieją wektory o najmniejszej normie.

Dla wektora $y = (1, 1, \ldots)$ otrzymujemy

$$W = \Big\{x \in \ell^1 : \sum_{n=1}^{\infty} x_n = 1\Big\}$$

i jeśli $x \in W$, to $\|x\|_1 = \sum_{n=1}^{\infty} |x_n| \geqslant \sum_{n=1}^{\infty} x_n = 1$; więc wektorem o najmniejszej normie jest każdy wektor $x \in W$, dla którego $x_n \geqslant 0$, $n = 1, 2, \ldots$, np. każdy wektor postaci $x = (\lambda, 1 - \lambda, 0, 0 \ldots)$, $0 \leqslant \lambda \leqslant 1$. W zbiorze W istnieje zatem nieprzeliczalnie wiele wektorów o najmniejszej normie.

Przyjmijmy teraz $y = (\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots)$, wtedy

$$W = \Big\{x \in \ell^1 : \sum_{n=1}^{\infty} \frac{n}{n+1}\, x_n = 1\Big\}.$$

Zauważmy, że każdy z wektorów $\frac{n+1}{n}\, e_n$, $n = 1, 2, \ldots$, gdzie e_n jest wektorem $e_n = (0, \ldots, 0, 1, 0, \ldots)$ z jedynką na n-tym miejscu, leży w W oraz $\|\frac{n+1}{n}\, e_n\|_1 = \frac{n+1}{n}$, więc $\operatorname{dist}(0, W) \leqslant 1$. Jednak w zbiorze W nie ma wektorów o normie 1, gdyż

$$\|x\|_1 = \sum_{n=1}^{\infty} |x_n| > \sum_{n=1}^{\infty} \frac{n}{n+1} |x_n| \geqslant \sum_{n=1}^{\infty} \frac{n}{n+1} x_n = 1.$$

Ogólnie $\operatorname{dist}(0, W) = 1/\|y\|_\infty$, a odpowiedź na pytanie, czy w W istnieją wektory o najmniejszej normie, zależy od liczności zbioru $\{n \in \mathbb{N} : |y_n| = \|y\|_\infty\}$. Jeżeli jest to zbiór pusty — jak w drugim przypadku — to w zbiorze W nie ma wektorów o najmniejszej normie; jeżeli jest jednoelementowy, powiedzmy postaci $\{n_0\}$, to w W istnieje dokładnie jeden wektor o najmniejszej normie, jest nim wektor e_{n_0}/y_{n_0}, a jeżeli posiada więcej niż jeden element — jak w przypadku pierwszym — to w W jest nieprzeliczalnie wiele wektorów o najmniejszej normie.

3.13. TWIERDZENIE. *Wektor* $w_0 \in W$ *ma własność* (3.10) z twierdzenia o najlepszej aproksymacji wtedy i tylko wtedy, gdy

$$\mathrm{Re}\langle x_0 - w_0,\, w_0 - w\rangle \geqslant 0$$

dla wszystkich $w \in W$.

Dowód: Załóżmy, że $\|x_0 - w_0\| = \inf_{w\in W}\|x_0 - w\|$ i niech w będzie dowolnym elementem zbioru W. Ponieważ $(1-\lambda)w_0 + \lambda w \in W$ dla $\lambda \in [0,1]$, więc funkcja

$$\begin{aligned} g(\lambda) &= \big\|x_0 - (1-\lambda)w_0 - \lambda\, w\big\|^2 \\ &= \|x_0 - w_0\|^2 + 2\lambda\,\mathrm{Re}\langle x_0 - w_0,\, w_0 - w\rangle + \lambda^2\|w_0 - w\|^2 \end{aligned}$$

przyjmuje na przedziale $[0,1]$ najmniejszą wartość w punkcie $\lambda = 0$, musi zatem być $g'(0) \geqslant 0$. To dowodzi postulowanej nierówności $\mathrm{Re}\langle x_0 - w_0,\, w_0 - w\rangle \geqslant 0$.

Z drugiej strony nierówność powyższa pociąga $g(0) \leqslant g(1)$, czyli $\|x_0 - w_0\| \leqslant \|x_0 - w\|$. $\square$

Twierdzenie to ma następującą interpretację geometryczną: zbiór tych wektorów x w przestrzeni $\mathcal{H}$, dla których

$$\mathrm{Re}\langle x_0 - w_0,\, x\rangle = \mathrm{Re}\langle x_0 - w_0,\, w_0\rangle = c$$

tworzy hiperpłaszczyznę przechodzącą przez punkt w_0 (i normalną do wektora $x_0 - w_0$). Hiperpłaszczyznę taką nazywamy **podpierającą** zbiór W w punkcie w_0, bowiem $\mathrm{Re}\langle x_0 - w_0,\, w_0\rangle = c$ oraz $\mathrm{Re}\langle x_0 - w_0,\, w\rangle \leqslant c$ dla wszystkich $w \in W$ (zbiór W styka się z hiperpłaszczyzną, ale leży tylko po jej jednej stronie).

3.14. WNIOSEK. *Jeżeli* W *jest domkniętą podprzestrzenią liniową przestrzeni Hilberta* $\mathcal{H}$, *to wektor* w_0 *najlepiej spośród wszystkich wektorów* W *aproksymuje wektor* $x_0 \in \mathcal{H}$ *wtedy i tylko wtedy, gdy* $\langle x_0 - w_0,\, w\rangle = 0$ *dla wszystkich* $w \in W$.

Dowód: Ponieważ wektor $w_0 + \langle x_0 - w_0,\, w\rangle w$ także należy do W, więc z twierdzenia 3.13 otrzymujemy $-\big|\langle x_0 - w_0,\, w\rangle\big|^2 \geqslant 0$, czyli $\langle x_0 - w_0,\, w\rangle = 0$. $\square$

Aby obliczyć odległość wektora x od domkniętej podprzestrzeni $\mathcal{H}_0$ przestrzeni Hilberta nie zawsze potrzebna jest znajomość wektora najlepiej w $\mathcal{H}_0$ aproksymującego wektor x. Pokażemy jak można ją obliczyć posługując się pojęciem wyznacznika Gramma.

DEFINICJA. **Wyznacznikiem Gramma** $G(x_1, x_2, \ldots, x_n)$ układu wektorów $x_1, x_2, \ldots, x_n$ przestrzeni unitarnej $\mathcal{H}$ nazywamy liczbę

$$G(x_1, x_2, \ldots, x_n) = \det\big\{\langle x_i,\, x_j\rangle\big\}_{1\leqslant i,j\leqslant n}.$$

3.15. LEMAT. *Niech $x_1, x_2, \ldots, x_n$ będzie układem wektorów przestrzeni unitarnej $\mathcal{H}$ i niech $\mathcal{H}_{n-1}$ oznacza podprzestrzeń liniową w $\mathcal{H}$ rozpiętą na wektorach $x_1, x_2, \ldots, x_{n-1}$. Wtedy*

$$G(x_1, x_2, \ldots, x_n) = G(x_1, x_2, \ldots, x_{n-1}) \cdot \mathrm{dist}(x_n, \mathcal{H}_{n-1})^2.$$

Dowód: Zauważmy na wstępie, że założenie zupełności przestrzeni $\mathcal{H}$ nie jest nam potrzebne. Cała „akcja" odbywa się w podprzestrzeni $\mathcal{H}_n = \mathrm{lin}\{x_1, x_2, \ldots, x_n\}$, a ta, jako skończenie wymiarowa, jest zupełna.

Załóżmy, że wektor w, najlepiej w przestrzeni $\mathcal{H}_{n-1}$ aproksymujący wektor x_n, ma postać $w = \lambda_1 x_1 + \lambda_2 x_2 + \ldots \lambda_{n-1} x_{n-1}$. Odejmując od ostatniego wiersza macierzy

$$\begin{matrix} \langle x_1,\, x_1\rangle & \ldots & \langle x_1,\, x_{n-1}\rangle & \langle x_1,\, x_n\rangle \\ \vdots & & \vdots & \vdots \\ \langle x_{n-1},\, x_1\rangle & \ldots & \langle x_{n-1},\, x_{n-1}\rangle & \langle x_{n-1},\, x_n\rangle \\ \langle x_n,\, x_1\rangle & \ldots & \langle x_n,\, x_{n-1}\rangle & \langle x_n,\, x_n\rangle \end{matrix}$$

(której wyznacznikiem jest $G(x_1, x_2, \ldots, x_n)$) kombinację liniową pozostałych wierszy, ze współczynnikami wynoszącymi odpowiednio $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$, otrzymamy macierz, której wyznacznik się nie zmieni, a ostatni wiersz przyjmie postać

$$\langle x_n - w,\, x_1\rangle \quad \ldots \quad \langle x_n - w,\, x_{n-1}\rangle \quad \langle x_n - w,\, x_n\rangle.$$

Z wniosku 3.14 wynika, że $\langle x_n - w,\, x\rangle = 0$ dla każdego $x \in \mathcal{H}_{n-1}$, zatem w istocie wiersz ten przyjmie postać

$$0 \quad \ldots \quad 0 \quad \langle x_n - w,\, x_n\rangle.$$

Postępując podobnie z ostatnią kolumną, tym razem odejmując od niej kombinację liniową pozostałych kolumn ze współczynnikami $\overline{\lambda_1}, \overline{\lambda_2}, \ldots, \overline{\lambda_{n-1}}$ otrzymamy macierz

$$\begin{matrix} \langle x_1,\, x_1\rangle & \ldots & \langle x_1,\, x_{n-1}\rangle & 0 \\ \vdots & & \vdots & \vdots \\ \langle x_{n-1},\, x_1\rangle & \ldots & \langle x_{n-1},\, x_{n-1}\rangle & 0 \\ 0 & \ldots & 0 & \langle x_n - w,\, x_n - w\rangle \end{matrix}$$

o wyznaczniku wynoszącym $G(x_1, x_2, \ldots, x_{n-1}) \cdot \|x_n - w\|^2$. $\square$

3.16. FAKT. *Dla dowolnych wektorów $x_1, x_2, \ldots, x_n$ przestrzeni unitarnej zachodzi nierówność $G(x_1, x_2, \ldots, x_n) \geqslant 0$, przy czym $G(x_1, x_2, \ldots, x_n) = 0$ wtedy i tylko wtedy, gdy wektory te są liniowo zależne.*

Dowód: Przyjmując $\mathcal{H}_0 = \{0\}$, z lematu 3.15 otrzymujemy

$$G(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} \operatorname{dist}(x_k, \mathcal{H}_{k-1})^2,$$

więc $G(x_1, x_2, \ldots, x_n) \geqslant 0$. Wektory $x_1, x_2, \ldots, x_n$ są liniowo niezależne wtedy i tylko wtedy, gdy każdy składnik powyższego iloczynu jest różny od zera. □

Liczbę $\sqrt{G(x_1, x_2, \ldots, x_n)}$ możemy interpretować jako (n-wymiarową) objętość równoległościanu rozpiętego na wektorach $x_1, x_2, \ldots, x_n$. Istotnie, objętość ta $V(x_1, x_2, \ldots, x_n)$ jest iloczynem pola podstawy równoległościanu i jego wysokości, podstawą jest równoległościan rozpięty na wektorach $x_1, x_2, \ldots, x_{n-1}$ a wysokością odległość wierzchołka x_n od hiperpłaszczyzny podstawy, zatem

$$V(x_1, x_2, \ldots, x_n) = V(x_1, x_2, \ldots, x_{n-1}) \cdot \operatorname{dist}(x_n, \mathcal{H}_{n-1}),$$

a w konsekwencji $V(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} \operatorname{dist}(x_k, \mathcal{H}_{k-1})$.

Natychmiastową konsekwencją lematu 3.15 jest następujące twierdzenie:

3.17. TWIERDZENIE. *Niech $x_1, x_2, \ldots, x_n$ będzie liniowe niezależnym układem wektorów przestrzeni unitarnej $\mathcal{H}$ i niech $\mathcal{H}_n$ oznacza podprzestrzeń liniową w $\mathcal{H}$ rozpiętą na tych wektorach. Dla dowolnego wektora $x \in \mathcal{H}$ zachodzi równość*

$$\operatorname{dist}(x, \mathcal{H}_n) = \sqrt{\frac{G(x_1, x_2, \ldots, x_n, x)}{G(x_1, x_2, \ldots, x_n)}}. \quad \square$$

3.18. PRZYKŁAD. Obliczymy odległość w $L^2(0,1)$ jednomianu t^n od podprzestrzeni $\mathcal{H}_{n-1}$ wszystkich wielomianów stopnia niższego niż n, tj. podprzestrzeni rozpiętej na jednomianach $1, t, \ldots, t^{n-1}$. Łatwo sprawdzamy, że

$$G(1, t, \ldots, t^n) = \det\left\{\frac{1}{i+j+1}\right\}_{0 \leqslant i,j \leqslant n},$$

zatem ze wzoru wyznacznikowego Cauchy'ego (patrz **??**)

$$G(1, t, \ldots, t^n) = \frac{\prod_{i<j}(i-j)^2}{\prod_{i,j}(i+j+1)} = \frac{[1!\,2!\,\cdots\,n!\,]^3}{(n+1)!\,(n+2)!\,\cdots\,(2n+1)!},$$

stąd na mocy twierdzenia 3.17

$$\operatorname{dist}(t^n, \mathcal{H}_{n-1}) = \sqrt{\frac{[n!]^4}{(2n)!\,(2n+1)!}} = \frac{1}{\sqrt{2n+1}\binom{2n}{n}}.$$

3.19. ZADANIE. Niech $x_1, x_2, \ldots, x_n$ będzie dowolnym układem funkcji z przestrzeni $L^2(\mathbb{R})$. Dowieść, że wyznacznik Gramma

$$G(x_1, x_2, \ldots, x_n) = \det\left\{\langle x_i,\, x_j\rangle\right\}_{1\leqslant i,j\leqslant n}$$

tego układu można wyrazić w postaci

$$G(x_1, x_2, \ldots, x_n) = \frac{1}{n!}\int_{\mathbb{R}^n}\left|\det\left\{x_i(t_j)\right\}_{1\leqslant i,j\leqslant n}\right|^2 dt_1\, dt_2 \ldots dt_n.$$

Ortogonalność

Dwa wektory x, y przestrzeni unitarnej $\mathcal{H}$ nazywamy **ortogonalnymi** jeżeli $\langle x,\, y\rangle = 0$. Gdy każdy element zbioru $A \subset \mathcal{H}$ jest ortogonalny do każdego elementu zbioru $B \subset \mathcal{H}$, to powiemy, że zbiory te są **ortogonalne** do siebie. Ortogonalność wektorów będziemy oznaczać $x \perp y$ a zbiorów $A \perp B$. Jeżeli $\{x\} \perp A$, to piszemy $x \perp A$.

Z liniowości i ciągłości iloczynu skaranego łatwo wyprowadzić, że wektory ortogonalne do zbioru A tworzą domkniętą podprzestrzeń liniową w $\mathcal{H}$. Będziemy ją oznaczać $A^{\perp}$ i nazywać **dopełnieniem ortogonalnym** zbioru A.

3.20. TWIERDZENIE (O ROZKŁADZIE ORTOGONALNYM). *Jeżeli $\mathcal{H}_0$ jest domkniętą podprzestrzenią liniową przestrzeni Hilberta $\mathcal{H}$, to każdy wektor $x \in \mathcal{H}$ można przedstawić, i to dokładnie na jeden sposób, w postaci*

$$x = x_1 + x_2, \quad \textit{gdzie} \quad x_1 \in \mathcal{H}_0,\ x_2 \in \mathcal{H}_0^{\perp}.$$

Dowód: Jeżeli $x \in \mathcal{H}$, to w podprzestrzeni $\mathcal{H}_0$ — na mocy twierdzenia o najlepszej aproksymacji 3.11 — jest taki wektor x_0, że

$$\|x - x_0\| = \operatorname{dist}(x, \mathcal{H}_0).$$

Wykażemy, że wektor $x_1 = x - x_0$ leży w $\mathcal{H}_0^{\perp}$, tzn., że $\langle x_1,\, y\rangle = 0$ dla wszystkich $y \in \mathcal{H}_0$.

Obierzmy $\lambda = \langle x_1, y\rangle/\|y\|^2$. Ponieważ $x_0 + \lambda y \in \mathcal{H}_0$, więc $\|x - (x_0 + \lambda y)\| \geqslant \operatorname{dist}(x, \mathcal{H}_0)$, a zatem

$$\begin{aligned} 0 \leqslant \|x - (x_0 + \lambda y)\|^2 - \|x - x_0\| = \|x_1 - \lambda y\|^2 - \|x_1\|^2 = \\ = -\overline{\lambda}\,\langle x_1, y\rangle - \lambda\,\overline{\langle x_1, y\rangle} + |\lambda|^2\|y\|^2 \end{aligned}$$

Podstawiając przyjętą wartość λ dostajemy $-|\langle x_1, y\rangle|^2/\|y\|^2 \geqslant 0$, a w konsekwencji $\langle x_1, y\rangle = 0$.

Do udowodnienia pozostaje jednoznaczność rozkładu. Przypuśćmy, że mamy jeszcze rozkład $x = x_0' + x_1'$, $x_0' \in \mathcal{H}_0$, $x_1' \in \mathcal{H}_0^{\perp}$. Wówczas $x_0 + x_1 = x_0' + x_1'$, czyli $x_0 - x_0' = x_1 - x_1'$. Wektor z lewej strony równości leży w podprzestrzeni $\mathcal{H}_0$, a wektor z prawej w $\mathcal{H}_0^{\perp}$; jednak $\mathcal{H}_0 \cap \mathcal{H}_0^{\perp} = \{0\}$, co oznacza $x_0 - x_0' = x_1 - x_1' = 0$, lub inaczej $x_0 = x_0'$ i $x_1 = x_1'$. $\square$

3.21. Wniosek. *Dla podzbioru $A \subset \mathcal{H}$ zbiór $(A^{\perp})^{\perp}$ jest najmniejszą domkniętą podprzestrzenią liniową w $\mathcal{H}$ zawierającą A.*

Dowód: Należy wykazać, że jeżeli $\mathcal{H}_0$ jest domkniętą podprzestrzenią liniową w $\mathcal{H}$ oraz $\mathcal{H}_0 \supset A$, to także $\mathcal{H}_0 \supset (A^{\perp})^{\perp}$. Jest tak w istocie, bo inkluzja $\mathcal{H}_0 \supset A$ pociąga inkluzję $\mathcal{H}_0^{\perp} \subset A^{\perp}$, a ta z kolei inkluzję $(\mathcal{H}_0^{\perp})^{\perp} \supset (A^{\perp})^{\perp}$, zaś z twierdzenia o rozkładzie ortogonalnym otrzymujemy równość $(\mathcal{H}_0^{\perp})^{\perp} = \mathcal{H}_0$. $\square$

3.22. Zadanie. Niech $\mathcal{H}$ będzie przestrzenią Hilberta oraz $\mathcal{H}_0$ jej domkniętą podprzestrzenią. Dowieść, że $\mathcal{H}/\mathcal{H}_0$ jest przestrzenią Hilberta, izometrycznie izomorficzną z $\mathcal{H}_0^{\perp}$. Opisać ten izomorfizm.

Układy ortogonalne

Zbiór $A \subset \mathcal{H}$ nazywamy **ortogonalnym** jeżeli każde dwa wektory A są ortogonalne. Jeżeli ponadto norma każdego wektora z A wynosi 1, to mówimy, że jest to zbiór **ortonormalny**. Zbiór ortonormalny nazywamy **zupełnym**, jeżeli nie ma w $\mathcal{H}$ niezerowego wektora ortogonalnego do A.

3.23. Ćwiczenia.

1. Zbiór ortonormalny jest liniowo niezależny.

2. W ośrodkowej przestrzeni unitarnej zbiór ortonormalny musi być skończony lub przeliczalny.

3.24. Przykład. **Wielomiany Legendre'a** L_n, $n = 0, 1, 2, 3, \ldots$, gdzie

$$L_n(t) = \frac{1}{n!2^n} \frac{d^n}{dt^n}(t^2 - 1)^n$$

tworzą układ ortogonalny w przestrzeni Hilberta $L^2(-1, 1)$.

Istotnie, L_n jest wielomianem stopnia n a dla dowolnego wielomianu P, całkując n-krotnie przez części, dostajemy

$$\langle P,\, L_n \rangle = \frac{(-1)^n}{n!2^n} \int_{-1}^{1} \frac{d^n}{dt^n} P(t)\, (t^2 - 1)^n dt,$$

więc dostajemy $\langle P,\, L_n \rangle = 0$ dla każdego wielomianu P stopienia niższego niż n, stąd $< L_m, L_n >= 0$ dla $m \neq n$.

Sprawdźmy jeszcze, czy jest to układ ortonormalny. Ponieważ

$$\frac{d^n}{dt^n} L_n(t) = \frac{1}{(2n)!!} \frac{d^{2n}}{dt^{2n}}(t^2 - 1)^n = \frac{(2n)!}{(2n)!!} = (2n-1)!!,$$

więc

$$\langle L_n,\, L_n \rangle = \frac{(2n-1)!!}{(2n)!!} \int_{-1}^{1} (1 - t^2)^n dt = \frac{2(2n-1)!!}{(2n)!!} \int_0^{\pi/2} \cos^{2n+1} u\, du = \frac{2}{2n+1}.$$

Zatem układ ortonormalny tworzą funkcje

$$\frac{\sqrt{2n+1}}{\sqrt{2}} L_n, \quad n = 0, 1, 2, \ldots\ .$$

Przez analogię w przestrzeni $L^2(a, b)$ możemy określić układ wielomianów ortogonalnych

$$P_n(t) = \frac{d^n}{dt^n}(t - a)^n (t - b)^n.$$

3.25. Przykład. W przykładzie 3.18 obliczyliśmy odległość w $L^2(0, 1)$ jednomianu t^n od podprzestrzeni $\mathcal{H}_{n-1}$ wszystkich wielomianów stopnia niższego niż n, Jak wiemy odległość ta jest równa normie wielomianu $t^n - W(t)$, gdzie W jest wielomianem stopnia niższego niż n, najlepiej aproksymującym jednomian t^n w $L^2(0, 1)$. Jest on jednoznacznie wyznaczony przez własność

$$\int_0^1 \big[t^n - W(t)\big] P(t)\, dt = 0$$

dla wszystkich $P \in \mathcal{H}_{n-1}$. W poprzednim przykładzie poznaliśmy już wielomian stopnia n o takiej własności, mianowicie wielomian

$$P_n(t) = \frac{d^n}{dt^n} t^n (t-1)^n.$$

Jego współczynnikiem przy t^n jest $\frac{(2n)!}{n!}$, zatem musi być

$$W(t) = t^n - \frac{n!}{(2n)!} \frac{d^n}{dt^n} t^n (t-1)^n.$$

Pozwala to obliczyć

$$\operatorname{dist}(t^n, \mathcal{H}_{n-1})^2 = \left\| \frac{n!}{(2n)!} P_n \right\|_2^2 = \frac{[n!]^2}{(2n)!} \int_0^1 t^n (1-t)^n dt = \frac{[n!]^4}{(2n)!\,(2n+1)!}.$$

3.26. ZADANIE. Pokazać, że **funkcje Rademachera**

$$r_n(t) = \operatorname{sgn} \sin 2^{n-1} \pi t, \quad n = 1, 2, 3, \ldots$$

tworzą układ ortonormalny w $L^2(0,1)$. Czy jest to układ zupełny?

Podamy teraz opis procesu ortonormalizacyjnego, pozwalającego ze zbioru liniowo niezależnego zbudować zbiór ortonormalny.

3.27. TWIERDZENIE ORTONORMALIZACYJNE SCHMIDTA. *Załóżmy, że* $\{x_1, x_2, \ldots, x_n\}$ *jest zbiorem liniowo niezależnym w przestrzeni unitarnej* $\mathcal{H}$. *Wtedy w* $\mathcal{H}$ *istnieje dokładnie jeden taki zbiór ortonormalny* $\{y_1, y_2, \ldots, y_n\}$, *że*

1. $\operatorname{lin}\{x_1, x_2, \ldots, x_k\} = \operatorname{lin}\{y_1, y_2, \ldots, y_k\}$,

2. $\langle x_k, y_k \rangle > 0$,

dla $k = 1, 2, \ldots, n$.

Dowód: Przeprowadzimy go przez indukcję względem k. Dla $k = 1$ wystarczy przyjąć $y_1 = \frac{1}{\|x_1\|} x_1$, a jeżeli elementy $y_1, y_2, \ldots, y_k$ są już określone, to kładziemy

$$y_{k+1} = \frac{1}{\|z_{k+1}\|} z_{k+1}, \quad \text{gdzie} \quad z_{k+1} = x_{k+1} - \sum_{j=1}^{k} \langle x_{k+1}, y_j \rangle \, y_j.$$

Zauważmy, że $z_{k+1} \neq 0$, w przeciwnym przypadku elementy $x_1, x_2, \ldots, x_{k+1}$ byłyby liniowo zależne. Z konstrukcji jest oczywiste, że $y_{k+1} \in \operatorname{lin}\{x_1, x_2, \ldots, x_{k+1}\}$, że $\|y_{k+1}\| = 1$ oraz, że $\langle y_{k+1}, y_j \rangle = 0$ dla $j = 1, 2, \ldots, k$. Stąd

$$\langle y_{k+1}, x_{k+1} \rangle = \langle y_{k+1}, z_{k+1} \rangle = \|z_{k+1}\| > 0.$$

Pokażemy jeszcze, że element y_{k+1} jest jedyny. Jeżeli pewien element y' ma te same własności, co y_{k+1}, to jest postaci

$$y' = \sum_{j=1}^{k} \lambda_j y_j.$$

Jednak y' musi być ortogonalny do każdego z elementów $y_1, y_2, \ldots, y_k$, więc w istocie $y' = \lambda_{k+1} y_{k+1}$. Wiemy też, że $\langle y', x_{k+1}\rangle = \lambda_{k+1}\langle y_{k+1}, x_{k+1}\rangle > 0$ oraz, że $\|y'\| = 1$. Z warunków tych wynika, że λ_{k+1} jest liczbą dodatnią i $|\lambda_{k+1}| = 1$. Daje to $y' = y_{k+1}$. $\square$

Dla wyznaczenia elementów $y_1, y_2, \ldots, y_n$ można posłużyć się wyznacznikami Gramma. Wzór, który przedstawimy może być szczególnie przydatny do obliczeń numerycznych.

3.28. TWIERDZENIE GRAMMA. *Elementy zbioru $\{y_1, y_2, \ldots, y_n\}$ w twierdzeniu ortonormalizacyjnym można otrzymać z formalnego wyznacznika*

$$y_k = c_k \cdot \det \begin{pmatrix} \langle x_1, x_1\rangle & \langle x_1, x_2\rangle & \ldots & \langle x_1, x_{k-1}\rangle & x_1 \\ \langle x_2, x_1\rangle & \langle x_2, x_2\rangle & \ldots & \langle x_2, x_{k-1}\rangle & x_2 \\ \vdots & \vdots & & \vdots & \vdots \\ \langle x_k, x_1\rangle & \langle x_k, x_2\rangle & \ldots & \langle x_k, x_{k-1}\rangle & x_k \end{pmatrix},$$

gdzie

$$c_k = \frac{1}{\sqrt{G(x_1, x_2, \ldots, x_{k-1})\, G(x_1, x_2, \ldots, x_k)}},$$

rozwijając go względem ostatniej kolumny.

Dowód: Jest oczywiste, że $y_k \in \operatorname{lin}\{x_1, x_2, \ldots, x_k\}$. Obliczając iloczyn skalarny $\langle y_k, x_j\rangle$ przy $j < k$ otrzymamy wyznacznik, w którym ostatnia kolumna pokrywa się z kolumną j-tą. Taki wyznacznik jest równy zeru, zatem $\langle y_k, x_j\rangle = 0$ dla $j < k$, w szczególności $y_k \perp y_j$ dla takich j. Dla $j = k$ otrzymujemy

$$\langle y_k, x_k\rangle = c_k\, G(x_1, x_2, \ldots, x_k) > 0.$$

Pozostaje sprawdzić, czy $\|y_k\| = 1$. Obliczając $\|y_k\|^2 = \langle y_k, y_k\rangle$ otrzymamy wyznacznik, w którym ostatnia kolumna przyjmuje postać

$$0 \quad 0 \quad \ldots \quad 0 \quad \langle x_k, y_k\rangle,$$

więc jego wartość jest równa $G(x_1, x_2, \ldots, x_{k-1})\,\langle x_k, y_k\rangle$. Stąd

$$\|y_k\|^2 = c_k^2\, G(x_1, x_2, \ldots, x_{k-1})\, G(x_1, x_2, \ldots, x_k) = 1. \quad \square$$

3.29. PRZYKŁAD. W $L^2(-1,1)$ proces ortonormalizacji Schmidta układu jednomianów 1, t, t^2, $t^3, \ldots$ prowadzi do układu funkcji

$$\frac{\sqrt{2n+1}}{\sqrt{2}} L_n. \quad n = 0, 1, 2, 3, \ldots,$$

gdzie

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

jest n-tym wielomianem Legendre'a.

Stwierdzenie to nie jest widoczną konsekwencją twierdzenia ortonormalizacyjnego a wynika z własności wielomianów Legendre'a, które pokazaliśmy w przykładzie 3.24: L_n jest wielomianem stopnia n, $\int_{-1}^{1} P(t)L_n(t)\,dt = 0$ dla każdego wielomianu P stopnia niższego niż n oraz

$$\int_{-1}^{1} t^n L_n(t)\,dt = \frac{1}{2^n} \int_{-1}^{1} (1-t^2)^n dt > 0.$$

Inne klasyczne wielomiany ortogonalne można otrzymać zastępując miarę Lebesgue'a na $(-1,1)$ innymi miarami. Dla przykładu, **wielomiany Czebyszewa**

$$T_0(t) = 1, \quad T_n(t) = \frac{1}{2^{n-1}} \cos(n \arccos t), \quad n = 1, 2, 3, \ldots$$

(a dokładniej wielomiany $2^{n-1/2}\pi^{-1/2}\, T_n(t)$) otrzymamy dla miary μ określonej wzorem

$$\mu(E) = \int_E \frac{dt}{\sqrt{1-t^2}}.$$

Jest tak istotnie, bo podstawiając $u = \arccos t$ otrzymujemy

$$\int_{-1}^{1} \cos(m \arccos t)\, \cos(n \arccos t)\, \frac{dt}{\sqrt{1-t^2}}$$
$$= \int_0^{\pi} \cos mu\, \cos nu\, du.$$

Kolejne wielomiany Czebyszewa mają postać $T_0(t) = 1$, $T_1(t) = t$, $T_2(t) = t^2 - \frac{1}{2}$. Następne można łatwo obliczyć ze wzoru rekurencyjnego

$$T_n(t) = t\, T_{n-1}(t) - \frac{1}{4} T_{n-2}(t) \quad \text{dla} \quad n = 3, 4, 5, \ldots .$$

3.30. ZADANIE. Pokazać, że w przestrzeni $L^2\big((-1,1), \sqrt{1-t^2}\,dt\big)$ proces ortonormalizacji Schmidta układu jednomianów 1, t, t^2, $t^3, \ldots$ prowadzi do funkcji $2^{n+1/2}\pi^{-1/2}\, U_n(t)$, gdzie

$$U_n(t) = \frac{1}{n+1} \frac{d}{dt} T_{n+1}(t) \quad n = 0, 1, 2, \ldots$$

jest n-tym **wielomianem Czebyszewa 2-giego rodzaju**.

Szeregi ortogonalne

Z twierdzenia Pitagorasa 3.6 wynika, że jeżeli $x_1, x_2, \ldots, x_n$ jest zbiorem ortogonalnym, to

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

Dla nieskończonych zbiorów ortonormalnych można to uogólnić w następujący sposób:

3.31. TWIERDZENIE. *Jeżeli $x_1, x_2, \ldots$ jest ciągiem ortogonalnym w przestrzeni Hilberta $\mathcal{H}$, to warunkiem koniecznym i wystarczającym na to, by szereg $\sum_{n=1}^{\infty} x_n$ był zbieżny, jest aby $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$. Wtedy*

$$\left\| \sum_{n=1}^{\infty} x_n \right\|^2 = \sum_{n=1}^{\infty} \|x_n\|^2.$$

Dowód: Jeśli $\{s_n\}$ jest ciągiem sum częściowych szeregu $\sum_{k=1}^{\infty} x_k$, zaś $\{S_n\}$ ciągiem sum częściowych szeregu $\sum_{k=1}^{\infty} \|x_k\|^2$, to dla $n > m$, z twierdzenia Pitagorasa, mamy

$$\begin{aligned} \|s_n - s_m\|^2 &= \|x_{m+1} + x_{m+2} + \ldots + x_n\|^2 = \\ &= \|x_{m+1}\|^2 + \|x_{m+2}\|^2 + \ldots + \|x_n\|^2 = S_n - S_m. \end{aligned}$$

Zatem $\{s_n\}$ jest ciągiem Cauchy'ego w $\mathcal{H}$ wtedy i tylko wtedy, gdy $\{S_n\}$ jest liczbowym ciągiem Cauchy'ego. $\square$

3.32. WNIOSEK. *Zbieżny szereg ortogonalny jest zbieżny bezwarunkowo, tj. zbieżny przy każdej permutacji swoich wyrazów.*

Dowód: Niech σ będzie dowolną permutacją wskaźników. Ustalmy dowolnie $\varepsilon > 0$ i wybierzmy wskaźnik n_0 tak, aby $\sum_{n>n_0} \|x_n\|^2 < \varepsilon$ a wskaźnik n_1 tak, aby zbiór $\{\sigma(1), \sigma(3), \ldots, \sigma(n_1)\}$ zawierał wszystkie liczby $1, 2, \ldots, n_0$. Jeżeli $n > n_1$ to

$$\left\| \sum_{k=1}^{n} x_k - \sum_{k=1}^{n} x_{\sigma(k)} \right\|^2 \leqslant 2 \sum_{k>n_0} \|x_k\|^2 \leqslant 2\,\varepsilon^2,$$

więc

$$\lim_{n\to\infty} \sum_{k=1}^{n} x_k = \lim_{n\to\infty} \sum_{k=1}^{n} x_{\sigma(k)}. \quad \square$$

Uwaga. Przypomnijmy, że w zbiorze $\mathbb{C}$ liczb zespolonych, a więc także w każdej skończenie wymiarowej przestrzeni unormowanej zbieżność bezwarunkowa szeregów pokrywa się z ich zbieżnością bezwzględną. W nieskończenie wymiarowych przestrzeniach Hilberta jest inaczej, istnieją szeregi bezwarunkowo zbieżne, które nie są zbieżne bezwzględnie. Dla przykładu w przestrzeni ℓ^2 szereg ortogonalny $\sum_{n=1}^{\infty} x_n$, gdzie x_n jest wektorem $(0,\ldots,0,\frac{1}{n},0,\ldots)$ z liczbą $\frac{1}{n}$ na n-tym miejscu, jest zbieżny bezwarunkowo (do wektora $(1,\frac{1}{2}\frac{1}{3},\ldots)$) lecz nie jest zbieżny bezwzględnie.

Klasyczne szeregi Fouriera

Dla funkcji $x \in L^2(0,1)$ szereg

$$a_0 + \sum_{k=1}^{\infty}\big(a_k \cos 2k\pi it + b_k\, sin 2k\pi it\big), \tag{3.11}$$

gdzie $a_0 = \displaystyle\int_0^1 x(t)\,dt$ oraz

$$a_k = 2\int_0^1 x(t)\cos 2k\pi it\,dt, \quad b_k = 2\int_0^1 x(t)\sin 2k\pi it\,dt,$$

nosi nazwę **klasycznego szeregu Fouriera** funkcji x. Szereg ten jest zbieżny bezwarunkowo, bo funkcje 1, $\sqrt{2}\cos 2\pi it$, $\sqrt{2}\sin 2\pi it$, $\sqrt{2}\cos 4\pi it$, $\sqrt{2}\sin 4\pi it$, $\sqrt{2}\cos 6\pi it$, itd, tworzą układ ortonormalny w $L^2(0,1)$. Naszym celem będzie pokazanie, że jest to układ zupełny. Układ ten można przenieść na dowolny skończony przedział (a,b) wstawiając $\frac{t}{\sqrt{b-a}}$ w miejsce zmiennej t i dopisując dodatkowy współczynnik $\frac{1}{\sqrt{b-a}}$

Zauważmy, że suma (3.11) jest równa sumie szeregu

$$\sum_{k=-\infty}^{\infty} c_k\, e^{2k\pi it},$$

gdy $c_0 = a_0$ oraz $c_k = \frac{1}{2}(a_k - i\,b_k)$, $c_{-k} = \frac{1}{2}(a_k - i\,b_k)$ dla $k = 1,2,3,\ldots$ Funkcje $y_k(t) = e^{2k\pi it}$, $k = 0,\pm1,\pm2,\ldots$ także tworzą układ ortonormalny w $L^2(0,1)$. Układ ten jest wygodniejszy od układu opisanego na początku, dlatego tylko nim będziemy się posługiwali. Dodatkowo, współczynniki c_k będziemy oznaczali $\widehat{x}(k)$, tj.

$$\widehat{x}(k) = \langle x,\, y_k\rangle = \int_0^1 x(t)\,e^{-2k\pi it}dt.$$

Zaczniemy od udowodnienia dwóch pomocniczych lematów.

3.33. LEMAT. *Jeżeli funkcja* $x : [0,1] \to \mathbb{C}$ *jest klasy* C^1 *oraz* $x(0) = x(1)$, *to*

$$\widehat{x'}(k) = 2k\pi i\, \widehat{x}(k) \quad \textit{dla} \quad k = 0, \pm 1, \pm 2, \dots .$$

Dowód: Całkując przez części otrzymujemy

$$\widehat{x'}(k) = x(t)\, e^{-2k\pi i t}\Big|_0^1 - \int_0^1 x(t)(-2k\pi i) e^{-2k\pi i t} dt = 2k\pi i\, \widehat{x}(k). \quad \square$$

3.34. LEMAT (RIEMANNA-LEBESGUE'A). *Jeżeli* $x \in L^1(a,b)$, *to*

$$\lim_{\lambda \to \infty} \int_a^b x(t) \sin \lambda t \, dt = 0.$$

Dowód: Dla funkcji y klasy C^1 lemat wynika z równości

$$\int_a^b y(t) \sin \lambda t \, dt = \frac{1}{\lambda}\Big(- y(t) \cos \lambda t \Big|_a^b + \int_a^b y'(t) \cos \lambda t \, dt \Big).$$

Ustalmy teraz dowolnie $\varepsilon > 0$ i dobierzmy funkcję y klasy C^1 tak, by $\|x - y\|_1 < \frac{\varepsilon}{2}$ a następnie wskaźnik λ_0, by $\left| \int_a^b y(t) \sin \lambda t \, dt \right| < \frac{\varepsilon}{2}$ dla $\lambda \geqslant \lambda_0$. Wtedy

$$\left| \int_a^b x(t) \sin \lambda t \, dt \right| \leqslant \big\| x - y \big\|_1 + \left| \int_a^b y(t) \sin \lambda t \, dt \right| < \varepsilon$$

dla $\lambda \geqslant \lambda_0$, o co chodziło. $\square$

3.35. TWIERDZENIE. *Jeżeli* $x : \mathbb{R} \to \mathbb{C}$ *jest funkcją klasy* C^1, *okresową o okresie* 1, *to ciąg*

$$s_n(t) = \sum_{|k| \leqslant n} \widehat{x}(k)\, e^{2k\pi i t}, \quad n = 0, 1, 2, \dots ,$$

sum częściowych szeregu Fouriera funkcji x, *jest do niej zbieżny jednostajnie. Co więcej*

$$\big\| x - s_n \big\|_\infty \leqslant \frac{C}{\sqrt{n}}$$

dla pewnej stałej C.

Dowód: Gdy $m < n$, to korzystając kolejno z lematu 3.33, nierówności Schwarza i nierówności Bessela otrzymujemy

$$
\begin{aligned}
\left|s_n(t) - s_m(t)\right| &\leqslant \sum_{m<|k|\leqslant n} \left|\widehat{x}(k)\right| = \sum_{m<|k|\leqslant n} \left|\widehat{x'}(k)\right| \frac{1}{2k\pi|k|} \\
&\leqslant \Big(\sum_{m<|k|\leqslant n} \left|\widehat{x'}(k)\right|^2\Big)^{1/2} \Big(\sum_{m<|k|\leqslant n} \frac{1}{4\pi^2k^2}\Big)^{1/2} \qquad (3.12)\\
&\leqslant \frac{1}{2\pi} \left\|x'\right\|_2 m^{-1/2},
\end{aligned}
$$

gdyż

$$\sum_{k>m} \frac{1}{k^2} \leqslant \sum_{k>m} \frac{1}{(k-1)k} = \frac{1}{m}.$$

Wynika stąd, że ciąg $\{s_n\}$ jest zbieżny jednostajnie. Pokażemy, że jego granicą jest funkcja x.

Funkcję s_n można zapisać w postaci

$$s_n(t) = \sum_{|k|\leqslant n} \int_0^1 x(u)\, e^{-2k\pi iu} du\, e^{2k\pi it} = \int_0^1 x(u)\, D_n(t-u)\, du, \qquad (3.13)$$

gdzie

$$D_n(t) = \sum_{|k|\leqslant n} e^{2k\pi it} = \frac{\sin(2n+1)\pi t}{\sin \pi t}$$

jest tzw. **jądrem Dirichleta**. Zauważmy, że D_n jest funkcją parzystą, okresową o okresie 1 oraz

$$\int_0^1 D_n(t)\, dt = \sum_{|k|\leqslant n} \int_0^1 y_k(t)\, dt = \sum_{|k|\leqslant n} \langle y_k, y_0\rangle = 1.$$

Ponieważ funkcja x jest także okresowa o okresie 1, więc przedział całkowania w drugiej z całek (3.13) możemy zmienić na dowolny przedział długości 1. Jeżeli wybierzemy przedział $(-t, 1-t)$ a następnie w całce zamienimy zmienną u na $t+u$, to otrzymamy

$$s_n(t) = \int_0^1 x(t+u)\, D_n(u)\, du.$$

Uwzględniając teraz fakt, że $\int_0^1 D_n(u)\, du = 1$, różnicę $s_n(t) - x(t)$ możemy przedstawić w postaci

$$s_n(t) - x(t) = \int_0^1 \big[x(t+u) - x(t)\big]\, D_n(u)\, du = \int_0^1 \varphi_t(u)\, \sin(2n+1)\pi u\, du,$$

gdzie

$$\varphi_t(u) = \frac{x(t+u) - x(t)}{\sin \pi u}.$$

Funkcja φ_t jest ciągła i ograniczona, bo $\lim_{u\to 0} \varphi_t(u) = \frac{x'(t)}{\pi}$ oraz $\lim_{u\to 1} \varphi_t(u) = -\frac{x'(t)}{\pi}$, zatem na mocy lematu Riemanna-Lebesgue'a $s_n(t) \to x(t)$ dla każdego ustalonego $t \in [0,1]$. Przechodząc w nierówności 3.35 do granicy przy $n \to \infty$ otrzymamy

$$\left\|x - s_m\right\|_\infty \leqslant \frac{1}{2\pi} \left\|x'\right\|_2 m^{-1/2}. \quad \Box$$

3.36. WNIOSEK. *W przestrzeni $L^2(0,1)$ trygonometryczny układ ortonormalny $e^{2k\pi it}$, $k = 0, \pm 1, \pm 2, \ldots$ jest zupełny.*

Dowód: W przestrzeni $L^2(0,1)$ funkcje, które dają się przedłużyć do funkcji klasy C^1, okresowych o okresie 1 (tj. takie funkcje x klasy C^1 na $(0,1)$, dla których $\lim_{t\to 0} x(t) = \lim_{t\to 1} x(t)$ oraz $\lim_{t\to 0} x'(t) = \lim_{t\to 1} x'(t)$) tworzą zbiór gęsty. Z twierdzenia 3.35 wiemy, że każdą taką funkcję w normie $\|\ \|_\infty$, a wiec tym bardziej w normie $\|\ \|_2$, można dowolnie przybliżać skończonymi kombinacjami liniowymi funkcji układu. Zatem funkcje układu tworzą zbiór liniowo gęsty w $L^2(0,1)$. $\Box$

3.37. PRZYKŁAD. Rozwiniemy w szereg Fouriera funkcję $x(t) = t - \frac{1}{2}$. Łatwy rachunek daje $\widehat{x}(0) = 0$ oraz

$$\widehat{x}(k) = \int_0^1 \left(t - \tfrac{1}{2}\right) e^{-2\pi ikt} dt = \frac{1}{-2\pi ik} \quad \text{dla} \quad k = \pm 1, \pm 2, \ldots,$$

zatem

$$t - \frac{1}{2} = \sum_{k\neq 0} \frac{1}{-2k\pi i} e^{2k\pi it} = \sum_{k=1}^{\infty} \frac{\sin 2k\pi t}{-k\pi}. \tag{3.14}$$

Ponieważ

$$\sum_{k=-\infty}^{\infty} \left|\widehat{x}(k)\right|^2 = \frac{1}{2\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} \quad \text{oraz} \quad \left\|x\right\|_2^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \frac{1}{12},$$

więc z równości Parsevala otrzymujemy

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Rozwinięcie (3.14) można łatwo wyprowadzić bez użycia techniki szeregów Fouriera, opierając się na wzorze

$$\sum_{k=1}^{n} \cos 2k\pi t = \frac{\sin(2n+1)\pi t}{2\sin \pi t} - \frac{1}{2}.$$

Mianowicie

$$\begin{aligned}\sum_{k=1}^{n} \frac{\sin 2k\pi t}{-k\pi} &= -2\int_0^t \sum_{k=1}^{n} \cos 2k\pi s\, ds = t - \int_0^t \frac{\sin(2n+1)\pi s}{\sin \pi s}\, ds \\ &= t - \int_0^t \left[\frac{1}{\sin \pi s} - \frac{1}{\pi s}\right] \sin(2n+1)\pi s\, ds + \frac{1}{\pi}\int_0^t \frac{\sin(2n+1)\pi s}{s}\, ds.\end{aligned}$$

Ponieważ funkcja $\frac{1}{\sin \pi s} - \frac{1}{\pi s}$ jest ograniczona na przedziale $(0,t)$, więc na mocy lematu Riemanna-Lebesgue'a pierwsza z całek po prawej stronie ostatniej równości dąży do 0 przy $n \to \infty$. Druga całka przez postawienie $u = (2n+1)\pi s$ sprowadza się do całki

$$\frac{1}{\pi}\int_0^{(2n+1)\pi t} \frac{\sin u}{u}\, du,$$

a więc dąży do $\frac{1}{\pi}\int_0^\infty \frac{\sin u}{u}\, du = \frac{1}{2}$. W konsekwencji

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{\sin 2k\pi t}{-k\pi} = t - \frac{1}{2}.$$

3.38. Przykład. Lemat 3.33 nasuwa pomysł uogólnienia rezultatów przykładu 3.37 tak, by móc obliczyć także sumy szeregów

$$\sum_{k=1}^{\infty} \frac{1}{k^{2n}}, \quad n = 2, 3, \dots .$$

Wybierzmy w przestrzeni $L^2(0,1)$ tak ciąg funkcji $x_0, x_1, \dots$, by $x_0 = 1$ oraz

$$x_n' = x_{n-1} \quad \text{i} \quad \int_0^1 x_n(t)\, dt = 0 \quad \text{dla} \quad n = 1, 2, 3, \dots .$$

Istnieje tylko jeden taki ciąg funkcji. Mamy mianowicie

$$\begin{aligned}&x_1(t) = t - \frac{1}{2}, \quad x_2(t) = \frac{1}{2}t^2 - \frac{1}{2}t + \frac{1}{12}, \quad x_3(t) = \frac{1}{6}t^3 - \frac{1}{4}t^2 + \frac{1}{12}t, \\ &x_4(t) = \frac{1}{24}t^4 - \frac{1}{12}t^3 + \frac{1}{24}t^2 - \frac{1}{720}, \quad \text{itd.},\end{aligned}$$

ogólnie

$$x_n(t) = \int_0^t x_{n-1}(s)\,ds + a_n,$$

gdzie

$$a_n = -\int_0^1 \int_0^t x_{n-1}(s)\,ds\,dt = \int_0^1 (s-1)\,x_{n-1}(s)\,ds.$$

Mamy zatem

$$x_n(t) = \sum_{k=0}^{n} \frac{a_k}{(n-k)!}\,t^{n-k},$$

gdzie $a_0 = 1$ i dla każdego $n = 1, 2, 3, \ldots$ zachodzi równość

$$\sum_{k=0}^{n} \frac{a_k}{(n+1-k)!} = 0.$$

Mnożąc obie strony tej równości przez $(n+1)!$ i przyjmując $k!\,a_k = \beta_k$ otrzymamy

$$\sum_{k=0}^{n} \binom{n+1}{k} \beta_k = 0, \qquad n = 1, 2, \ldots\ ,$$

zatem (patrz [3], 449]) $\beta_0, \beta_1, \beta_2, \ldots$ jest ciągiem współczynników rozwinięcia w szereg Taylora funkcji $t/(e^t - 1)$, tzn.

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} \frac{\beta_n}{n!}\,t^n.$$

W szczególności $\beta_{2n} = (-1)^{n-1} B_n$, gdzie $B_1, B_2, \ldots$ jest ciągiem **liczb Bernoulli'ego**

$$B_1 = \frac{1}{6}, \quad B_2 = \frac{1}{30}, \quad B_3 = \frac{1}{42}, \quad B_4 = \frac{1}{30}, \quad B_5 = \frac{5}{66}, \quad B_6 = \frac{691}{2730}, \quad \ldots$$

Wiemy, że funkcje $e_k(t) = \exp 2\pi i k t$, $k = 0, \pm 1, \pm 2, \ldots$ tworzą układ ortonormalny zupełny w $L^2(0,1)$ a dla funkcji $x \in L^2(0,1)$ zachodzi **równość Parsevala**

$$\int_0^1 \left|x(t)\right|^2 dt = \sum_{-\infty}^{+\infty} \left|\widehat{x}(k)\right|^2, \quad \text{gdzie} \quad \widehat{x}(k) = \int_0^1 x(t)\,\overline{e_k(t)}\,dt$$

jest ciągiem współczynników Fouriera funkcji x.

Łatwo obliczamy

$$\widehat{x_1}(0) = 0, \qquad \widehat{x_1}(k) = \int_0^1 \left(t - \tfrac{1}{2}\right) e^{-2\pi i k t} dt = \frac{1}{-2\pi i k} \quad \text{dla} \quad k = \pm 1, \pm 2, \ldots,$$

a dla $n > 1$ całkując przez części dostajemy

$$\widehat{x_n}(k) = \frac{1}{-2\pi ik}\,\widehat{x_{n-1}}(k) = \ldots = \frac{1}{(-2\pi ik)^n}, \quad \text{dla} \quad k \neq 0 \quad \text{oraz} \quad \widehat{x_n}(0) = 0.$$

Z drugiej strony, także całkując przez części, dla $n, m = 1, 2, \ldots$ otrzymujemy

$$\int_0^1 x_n(t)\overline{x_m(t)}\,dt = x_{n+1}(t)\overline{x_m(t)}\Big|_0^1 - \int_0^1 x_{n+1}(t)\overline{x_{m-1}(t)}\,dt.$$

Gdy $m = 1$, to lewa strona przyjmuje wartość

$$\frac{1}{2}\big(x_{n+1}(1) + x_{n+1}(0)\big) = x_{n+1}(0) = a_{n+1},$$

a gdy $m > 1$, to przyjmie wartość $-\int_0^1 x_{n+1}(t)\overline{x_{m-1}(t)}\,dt$ i można znów całkować przez części otrzymując w końcu

$$\int_0^1 x_n(t)\overline{x_m(t)}\,dt = (-1)^{m-1} x_{n+m}(0) = (-1)^{m-1}\,a_{n+m}.$$

W szczególności

$$\int_0^1 \big|x_n(t)\big|^2\,dt = (-1)^{n-1} a_{2n} = \frac{B_n}{(2n)!}.$$

Uwzględniając obliczone wartości współczynników Fouriera funkcji x_n, z równości Parsevala, otrzymujemy

$$\sum_{k=1}^{\infty} \frac{1}{k^{2n}} = \frac{(2\pi)^{2n}}{2\,(2n)!} \cdot B_n.$$

Funkcje Hermite'a

Pokażemy, że **funkcje Hermite'a**

(3.15) $$h_n(t) = \frac{(-1)^n}{n!}\, e^{\pi t^2}\, \frac{d^n}{dt^n}\, e^{-2\pi t^2}, \quad n = 0, 1, 2, \ldots$$

są funkcjami własnymi transformaty Fouriera, odpowiadającymi wartościom własnym $(-i)^n$, tj.

$$\widehat{h_n} = (-i)^n\, h_n.$$

Zamiast funkcji Hermite'a wygodniej jest posługiwać się układem funkcji przeskalowanych

$$e_n = \frac{h_n}{\|h_n\|_2} = \sqrt{\frac{\sqrt{2}\,n!}{(4\pi)^n}}\, h_n,$$

które tworzą układ ortonormalny zupełny w $L^2(\mathbb{R})$.

3.39. TWIERDZENIE. *Funkcje*

$$e_n = \frac{h_n}{\|h_n\|_2} = \sqrt{\frac{\sqrt{2}\,n!}{(4\pi)^n}}\,h_n, \quad n = 0, 1, 2, \ldots,$$

gdzie

$$h_n(t) = \frac{(-1)^n}{n!}\, e^{\pi t^2}\, \frac{d^n}{dt^n}\, e^{-2\pi t^2}$$

jest n-tą funkcją Hermite'a, tworzą układ ortonormalny zupełny w $L^2(\mathbb{R})$. Ponadto każda z funkcji e_n jest funkcją własną transformaty Fouriera odpowiadającą wartości własnej $(-i)^n$, tj.

$$\widehat{e_n} = (-i)^n\, e_n.$$

Dowód: Zauważmy, że n-ta funkcja Hermite'a jest postaci

$$h_n(t) = H_n(t)\, e^{-\pi t^2},$$

gdzie H_n jest pewnym wielomianem stopnia n, tzw. n-tym **wielomianem Hermite'a**. Wynika stąd, że $h_n \in \mathcal{S}$

Różniczkując (3.15) otrzymamy

$$h_n'(t) - 2\pi t\, h_n(t) = -(n+1)\, h_{n+1}(t), \tag{3.16}$$

a ponieważ ze wzoru Leibniza dla funkcji gładkiej x zachodzi równość

$$\frac{d^n}{dt^n}\big(4\pi t\, x(t)\big) = 4\pi t \frac{d^n}{dt^n} x(t) + 4\pi n \frac{d^{n-1}}{dt^{n-1}} x(t),$$

więc przyjmując $x(t) = e^{-2\pi t^2}$ i mnożąc obie strony przez $e^{\pi t^2}$ otrzymamy

$$(n+1)\, h_{n+1} = 4\pi t\, h_n(t) - 4\pi\, h_{n-1},$$

co po wstawieniu w (3.16) daje (przyjmujemy $h_{-1}(t) \equiv 0$)

$$h_n'(t) + 2\pi t\, h_n(t) = 4\pi\, h_{n-1}(t). \tag{3.17}$$

Jeśli teraz wyrugujemy z równań (3.16) i (3.17) pochodną $h_n'(t)$, to otrzymamy następujący wzór rekurencyjny dla funkcji Hermite'a:

$$4\pi t\, h_n(t) = (n+1)\, h_{n+1}(t) + 4\pi\, h_{n-1}(t), \quad n = 0, 1, 2, \ldots \tag{3.18}$$

Posługując się wzorem (3.18) łatwo możemy obliczyć wszystkie iloczyny skalarne $\langle h_n, h_m \rangle$. Mianowicie, gdy $m \geqslant 1$, to całkując dla przez części funkcję

$$\left[e^{2\pi t^2} \frac{d^n}{dt^n} e^{-2\pi t^2}\right] \frac{d^m}{dt^m} e^{-2\pi t^2}$$

i uwzględniając (3.18) otrzymujemy

$$\langle h_n, h_m\rangle = \frac{1}{m}\int_{-\infty}^{\infty}\big[4\pi t\, h_n(t) - (n+1)\, h_{n+1}(t)\big]\, h_{m-1}(t)\, dt = \frac{4\pi}{m}\langle h_{n-1}, h_{m-1}\rangle,$$

a ponieważ $\langle h_0, h_0\rangle = \int_{-\infty}^{\infty} e^{-2\pi t^2} dt = \frac{1}{\sqrt{2}}$, więc $\langle h_n, h_m\rangle = 0$, gdy $n \neq m$ oraz

$$\|h_n\|_2^2 = \langle h_n, h_n\rangle = \frac{(4\pi)^n}{\sqrt{2}\, n!}.$$

Udowodniliśmy w ten sposób, że funkcje $e_0, e_1, e_2, \ldots$ tworzą układ ortonormalny. Pokażemy teraz, że jest to układ zupełny. Niech $x \in L^2(\mathbb{R})$. Wtedy transformata Fouriera funkcji $y(t) = x(t)\, e^{-\pi t^2}$ ma postać

$$\widehat{y}(t) = \int_{-\infty}^{\infty} y(s) \sum_{n=0}^{\infty} \frac{(-2\pi t s)^n}{n!}\, ds = \sum_{n=0}^{\infty} \frac{(-2\pi t)^n}{n!} \int_{-\infty}^{\infty} x(s)\, s^n e^{-\pi s^2}\, ds.$$

Jeżeli funkcja x jest ortogonalna do wszystkich funkcji e_n, to każda z całek po prawej stronie jest równa zeru, więc $\widehat{y} = 0$, stąd $y = 0$, a w konsekwencji także $x = 0$.

Pozostaje dowieść, że $\widehat{h_n} = (-i)^n h_n$ dla wszystkich n. Stosując transformatę Fouriera równości (3.16) i (3.17) oraz uwzględniając **??** dostaniemy

$$\begin{aligned} 2\pi i t\, \widehat{h_n}(t) + i\, \widehat{h_n}'(t) &= -(n+1)\, \widehat{h_{n+1}}(t), \\ 2\pi i t\, \widehat{h_n}(t) - i\, \widehat{h_n}'(t) &= 4\pi\, \widehat{h_{n+1}}(t), \end{aligned}$$

skąd

$$4\pi i t\, \widehat{h_n}(t) = -(n+1)\, \widehat{h_{n+1}}(t) + 4\pi\, \widehat{h_{n-1}}(t).$$

Porównując tą równość z (3.18) widzimy, że funkcje $\widehat{h_n}$ oraz $(-i)^n h_n$ spełniają ten sam wzór rekurencyjny, aby pokazać, że są równe wystarczy sprawdzić, że $\widehat{h_{-1}} = i\, h_{-1}$ oraz $\widehat{h_0} = h_0$. Pierwsza z równości jest oczywista, a drugą udowodniliśmy już w **??**. $\square$

ROZDZIAŁ IV

CIĄGŁE FUNKCJONAŁY LINIOWE

Własności ogólne

4.1. TWIERDZENIE. *Niech X będzie przestrzenią unormowaną. Dla funkcjonału liniowego x^* na X następujące warunki są równoważne:*

(i). *x^* jest ciągły;*

(ii). *x^* jest ciągły w jednym punkcie;*

(iii). *${x^*}^{-1}(\{0\})$ jest zbiorem domkniętym;*

(iv). *istnieje taki zbiór otwarty $U \subset X$ oraz liczba $\lambda \in \mathbb{C}$, że $\lambda \notin x^*(U)$;*

(v). *istnieje taka stała M, że*

$$\left|x^*(x)\right| \leqslant M\,\|x\|, \quad x \in X.$$

Dowód:

(i)$\Rightarrow$(ii). Oczywiste.

(ii)$\Rightarrow$(iii). Załóżmy, że $x_n \to x_0$ i $x^*(x_n) = 0$ dla $n = 1, 2, \ldots$. Pokażemy, że także $x^*(x_0) = 0$. Jeżeli funkcjonał x^* jest ciągły w punkcie y_0, to $x_n - x_0 + y_0 \to y_0$, zatem $x^*(x_n - x_0 + y_0) \to x^*(y_0)$, a ponieważ $x^*(x_n - x_0 + y_0) = x^*(x_n) - x^*(x_0) + x^*(y_0)$, więc $x^*(x_0) = 0$.

(iii)$\Rightarrow$(iv). Jeśli $x^* = 0$, to $1 \notin x^*(X) = \{0\}$ a jeśli $x^* \neq 0$, to $0 \notin x^*(U)$, gdzie $U = X - {x^*}^{-1}\{0\})$.

(iv)$\Rightarrow$(v). Ustalmy element x_0 należący do U. Zbiór U jest otwarty, zatem dla pewnego $\varepsilon > 0$ do U należą wszystkie elementy postaci $x_0 + \varepsilon x$, $\|x\| \leqslant 1$. Jeżeli $\lambda \notin x^*(U)$, to kładąc $\mu = \frac{1}{\varepsilon}[\lambda - x^*(x_0)]$ otrzymujemy łatwo, że $\mu \notin x^*(K)$, gdzie K jest kulą jednostkową w X. Wynika stąd, że $\left|x^*(x)\right| \leqslant |\mu|$ dla $x \in K$. Rzeczywiście, gdyby dla pewnego $x \in K$ zachodziła nierówność $\left|x^*(x)\right| > |\mu|$, to dla elementu $y = \frac{\mu}{x^*(x)}\, x$ leżącego w K mielibyśmy $x^*(y) = \mu$. Jeżeli teraz x jest dowolnym elementem w X i $x \neq 0$, to $\frac{1}{\|x\|}\, x \in K$. Stąd

$$|x^*(x)| = \left| x^*\Big(\frac{1}{\|x\|}\, x\Big)\right| \|x\| \leqslant |\mu|\, \|x\|. \tag{4.19}$$

(v)⇒(i). Oczywiste. □

Kres dolny liczb M spełniających nierówność (v) oznaczamy przez $\|x^*\|$ i nazywamy **normą funkcjonału** x^*.

4.2. TWIERDZENIE. *Dla każdego ciągłego funkcjonału liniowego x^* na X zachodzi nierówność*

$$|x^*(x)| \leqslant \|x^*\|\, \|x\|, \qquad x \in X,$$

a także

$$\|x^*\| = \sup_{\|x\| \leqslant 1} |x^*(x)| = \sup_{\|x\|=1} |x^*(x)|.$$

Dowód: Niech $M = \sup\{|x^*(x)| : \|x\| = 1\}$. Wtedy rozumując jak w (4.19) otrzymujemy, że $|x^*(x)| \leqslant M\, \|x\|$ dla wszystkich $x \in X$, a więc $M \geqslant \|x^*\|$. Z drugiej strony, gdy $\|x\| = 1$, to $|x^*(x)| \leqslant \|x^*\|$, zatem $M \leqslant \|x^*\|$. □

4.3. PRZYKŁAD. Niech x_0 będzie dowolną funkcją całkowalną w sensie Lebesgue'a na przedziale $[0,1]$. Na przestrzeni $C[0,1]$ określmy funkcjonał liniowy x^* wzorem

$$x^*(x) = \int_0^1 x(t)\, x_0(t)\, dt.$$

Wtedy

$$|x^*(x)| = \left| \int_0^1 x(t)\, x_0(t)\, dt \right| \leqslant \int_0^1 |x(t)|\, |x_0(t)|\, dt \leqslant$$
$$\leqslant \|x\|_\infty \int_0^1 |x_0(t)|\, dt = \|x\|_\infty\, \|x_0\|_1.$$

Zatem x^* jest ciągły i $\|x^*\| \leqslant \|x_0\|_1$.

Pokażemy, że zachodzi równość obu norm. Z regularności miary Lebesgue'a (patrz **??**) wynika, że dla dowolnego $\varepsilon > 0$ istnieje taka funkcja $x_1 \in C[0,1]$, że $\|x_0 - x_1\|_1 < \varepsilon$. Możemy przy tym dodatkowo zażądać, by $x_1(t) \neq 0$ dla $t \in [0,1]$. Połóżmy $x = \operatorname{sgn} \overline{x_1}$, tzn.

$$x(t) = \frac{\overline{x_1(t)}}{|x_1(t)|} \quad \text{dla} \quad t \in [0,1].$$

Oczywiście $x \in C[0,1]$, $\|x\|_\infty = 1$ oraz

$$\begin{aligned}\left|x^*(x)\right| &\geqslant \left|\int_0^1 x(t)\,x_1(t)\,dt\right| - \int_0^1 \left|x(t)\right| \left|x_1(t)-x_0(t)\right| dt \geqslant \\ &\geqslant \int_0^1 \left|x_1(t)\right| dt - \|x_1 - x_0\|_1 \geqslant \|x_0\|_1 - 2\varepsilon,\end{aligned}$$

a to daje nierówność przeciwną $\|x^*\| \geqslant \|x_0\|_1$.

4.4. Przykład. Niech x^* będzie funkcjonałem z poprzedniego przykładu, lecz tym razem określonym na przestrzeni $C^1[0,1]$ z normą $\|x\|_{C^1} = |x(0)| + \|x'\|_\infty$. Jest to oczywiście funkcjonał ciągły, gdyż $\|\ \|_\infty \leqslant \|\ \|_{C^1}$. Obliczymy $\|x^*\|$.

Jeżeli funkcję $x \in C^1[0,1]$ przedstawimy w postaci $x(t) = x(0) + \int_0^t x'(s)\,ds$, to

$$\begin{aligned}x^*(x) &= x(0)\int_0^1 x_0(t)\,dt + \int_0^1\int_0^t x'(s)\,x_0(t)\,ds\,dt = \\ &= x(0)\,\widetilde{x}_0(0) + \int_0^1 x'(s)\,\widetilde{x}_0(s)\,ds,\end{aligned}$$

gdzie $\widetilde{x}_0(s) = \int_s^1 x_0(t)\,dt$. Ponieważ x' może być dowolną funkcją ciągłą na $[0,1]$, więc rozumując jak przy obliczaniu normy funkcjonału w poprzednim przykładzie i opierając się na oczywistej równości

$$\sup_{|\alpha|+|\beta|\leqslant 1} |\alpha a + \beta b| = \max\{|a|,|b|\}$$

otrzymamy

$$\|x^*\| = \max\left\{\left|\widetilde{x}_0(0)\right|,\ \|\widetilde{x}_0\|_1\right\}.$$

Ta wielkość jest na ogół istotnie mniejsza niż $\|x_0\|_1$. Niech na przykład x_ε^*, dla $0 < \varepsilon \leqslant 1$, oznacza funkcjonał, w którym w miejsce x_0 wstawiono funkcję x_ε określoną wzorem

$$x_\varepsilon(t) = \begin{cases} \dfrac{1}{2t-1} & \text{gdy } \varepsilon < |2t-1| \leqslant 1; \\ 0 & \text{poza tym.}\end{cases}$$

Wtedy

$$\|x_\varepsilon^* - x_\delta^*\| = \|\widetilde{x}_\varepsilon - \widetilde{x}_\delta\|_1 = \frac{1}{4}(\delta - \varepsilon) \quad \text{dla} \quad 0 < \varepsilon < \delta \leqslant 1. \tag{4.20}$$

W szczególności $\|x_\varepsilon^*\| = \frac{1}{4}(1-\varepsilon)$, natomiast $\|x_\varepsilon\|_1 = \ln\frac{1}{\varepsilon}$.

Z (4.20) wynika ponadto, że dla każdej funkcji $x \in C^1[0,1]$ istnieje granica $\lim_{\varepsilon\to 0} x_\varepsilon^*(x)$, Pozwala to określić na przestrzeni $C^1[0,1]$ ciągły funkcjonał liniowy w postaci wartości głównej całki niewłaściwej

$$\lim_{\varepsilon\to 0} \int\limits_{\varepsilon<|2t-1|\leqslant 1} \frac{x(t)}{2t-1}\, dt.$$

Norma tego funkcjonału wynosi $\frac{1}{4}$.

Dla funkcji x z przestrzeni $C[0,1]$ całka powyższa zwykle jest rozbieżna.

Przestrzeń sprzężona

4.5. Twierdzenie. *Niech X będzie przestrzenią unormowaną. Zbiór X^* wszystkich ciągłych funkcjonałów liniowych na X, po wprowadzeniu działań dodawania funkcjonałów i mnożenia funkcjonału przez liczbę w sposób następujący*

$$\big(x_1^* + x_2^*\big)(x) = x_1^*(x) + x_2^*(x),$$
$$\big(\lambda x^*\big)(x) = \lambda\, x^*(x)$$

oraz normy jako normy funkcjonału, staje się przestrzenią Banacha, nazywaną przestrzenią sprzężoną do przestrzeni X.

Dowód: Sprawdzenie, że X^* jest przestrzenią liniową unormowaną jest nietrudne. Pokażemy więc tylko jej zupełność.

Niech x_n^* będzie ciągiem Cauchy'ego w X^*. Ponieważ

$$\big|x_n^*(x) - x_m^*(x)\big| \leqslant \|x_n^* - x_m^*\|\,\|x\|,$$

więc $x_n^*(x)$ jest dla każdego $x \in X$ ciągiem Cauchy'ego w $\mathbb{C}$. Zatem istnieje granica $\lim_{n\to\infty} x_n^*(x)$, którą oznaczymy $x^*(x)$. Pokażemy kolejno, że x^* jest liniowy, ciągły oraz, że $\|x_n^* - x^*\| \to 0$.

Liniowość funkcjonału x^* otrzymujemy łatwo przechodząc do granicy w równościach

$$x_n^*(x_1 + x_2) = x_n^*(x_1) + x_n^*(x_2),$$
$$x_n^*(\lambda x) = \lambda\, x_n^*(x).$$

Pokażemy, że funkcjonał x^* jest ciągły. Zauważmy w tym celu, że $\|x_n^*\|$ jest liczbowym ciągiem Cauchy'ego, a więc dla pewnej stałej M zachodzi $\|x_n^*\| \leqslant M$, $n = 1, 2, \ldots$. Stąd dla każdego $x \in X$ mamy $\big|x_n^*(x)\big| \leqslant M\,\|x\|$, a to z kolei pociąga

$$\big|x^*(x)\big| \leqslant M\,\|x\|.$$

Pozostaje więc tylko dowieść, że x_n^* dąży do x^* w X^*. Dla dowolnego $\varepsilon > 0$ mamy

$$\left|x_n^*(x) - x_m^*(x)\right| < \|x_n^* - x_m^*\| \, \|x\| < \varepsilon \, \|x\|$$

przy dostatecznie dużych m i n, Ustalając n i przechodząc do granicy przy $m \to \infty$ otrzymamy

$$\left|x_n^*(x) - x^*(x)\right| \leqslant \varepsilon \, \|x\|,$$

co z definicji normy daje $\|x_n^* - x^*\| \leqslant \varepsilon$. $\square$

4.6. TWIERDZENIE. *Na przestrzeni skończenie wymiarowej każdy funkcjonał liniowy jest ciągły, oraz przestrzeń sprzężona jest izomorficzna z wyjściową.*

Dowód: Ustalmy bazę Hamela $e_1, e_2, \ldots, e_n$ przestrzeni X i wprowadźmy normę $\| \ \|$ kładąc

$$\|x\| = \max_{1 \leqslant i \leqslant n} |x_i|, \quad \text{dla} \quad x = \sum_{i=1}^{n} x_i e_i.$$

Każdy funkcjonał liniowy x^* na X jest jednoznacznie wyznaczony przez swoje wartości na wektorach bazy, gdyż

$$x^*(x) = x^*\Big(\sum_{i=1}^{n} x_1 e_i\Big) = \sum_{i=1}^{n} x_1 x^*(e_i).$$

Ponadto

$$\left|x^*(x)\right| = \left|\sum_{i=1}^{n} x_i x^*(e_i)\right| \leqslant \max_{1 \leqslant i \leqslant n} |x_i| \sum_{i=1}^{n} \left|x^*(e_1)\right| \leqslant M \, \|x\|,$$

gdzie $M = \sum_{i=1}^{n} |x^*(e_i)|$. Wynika stąd, że x^* jest ciągły w normie $\| \ \|$, a więc jest ciągły także w każdej innej normie przestrzeni X, bo w X każde dwie normy są równoważne.

Z drugiej strony dowolny układ liczb $\alpha_1, \alpha_2 \ldots, \alpha_n$ wzorem

$$x^*(x) = \sum_{i=1}^{n} \alpha_i x_i$$

określa (ciągły) funkcjonał liniowy na X. Ponieważ układ taki określa także (jednoznacznie) element $y = \sum_{i=1}^{n} \alpha_i e_i$ przestrzeni X, zatem przyporządkowanie $y \to x_y^*$

$$x_y^*(x) = \sum_{i=1}^{n} \alpha_i x_i, \quad x \in X,$$

elementom przestrzeni X funkcjonałów liniowych z X^* jest wzajemnie jednoznaczne. Jest ono liniowe — jest więc algebraicznym izomorfizmem X na X^*, a ponieważ w przestrzeni X^* każde dwie normy są także równoważne, jest homeomorfizmem. $\square$

4.7. Przykład. Podobne rozumowanie jak w powyższym dowodzie można przeprowadzić przy opisie przestrzeni ${\ell^1}^*$ wszystkich ciągłych funkcjonałów liniowych na przestrzeni ℓ^1.

Jak wiemy (patrz str. 16), każdy wektor $x = (x_1, x_2, \ldots)$ przestrzeni ℓ^1 można jednoznacznie przestawić w postaci (bezwzględnie) zbieżnego szeregu

$$x = \sum_{n=1}^{\infty} x_n e_n, \tag{4.21}$$

gdzie e_n, $n = 1, 2, \ldots$, jest wektorem $e_n = (0, \ldots, 0, 1, 0, \ldots)$ z jedynką na n-tym miejscu. Jeżeli y^* jest ciągłym funkcjonałem liniowym na przestrzeni ℓ^1, to

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

Połóżmy $y_n = y^*(e_n)$, Ciąg $\{y_n\}$ jest oczywiście ograniczony, definiuje więc wektor $y = (y_1, y_2, \ldots)$ z przestrzeni ℓ^∞, o normie $\|y\|_\infty \leqslant \|y^*\|$. Odwrotnie, każdy wektor $y = (y_1, y_2, \ldots) \in \ell^\infty$ wzorem

$$y^*(x) = \sum_{n=1}^{\infty} x_n y_n, \quad \text{gdzie} \quad x = (x_1, x_2, \ldots) \tag{4.22}$$

określa funkcjonał liniowy na przestrzeni ℓ^1. Mamy

$$\big|y^*(x)\big| \leqslant \sum_{n=1}^{\infty} |x_n|\,|y_n| \leqslant \|x\|_1 \|y\|_\infty,$$

co pociąga, że y^* jest ciągły i $\|y^*\| \leqslant \|y\|_\infty$. Wraz z wcześniej udowodnioną nierównością $\|y^*\| \geqslant \|y\|_\infty$ oznacza to, że przyporządkowanie $y \to y^*$ określone wzorem (4.22) jest izometrycznym izomorfizmem ℓ^∞ na ${\ell^1}^*$.

Wzór (4.22) może też posłużyć do określenia ciągłych funkcjonałów liniowych na przestrzeni $\boldsymbol{c_0}$. Jeżeli $y = (y_1, y_2 \ldots)$ leży w ℓ^1, to szereg (4.22) jest zbieżny oraz $|y^*(x)| \leqslant \|x\|_\infty \|y\|_1$, zatem y^* jest ciągły na $\boldsymbol{c_0}$ i $\|y^*\| \leqslant \|y\|_1$. Aby dowieść, że przyporządkowanie $y \to y^*$ jest izometrycznym izomorfizmem ℓ^1 na $\boldsymbol{c_0}$ wystarczy pokazać, że każdy ciągły funkcjonał liniowy y^* na $\boldsymbol{c_0}$ ma postać (4.22) oraz, że $\|y^*\| \geqslant \|y\|_1$. Pierwszy fakt wynika z tego, że ciąg $\{e_n\}$ jest bazą przestrzeni $\boldsymbol{c_0}$, a więc każdy wektor ma przedstawienie (4. 21). Dla dowodu drugiego niech x będzie wektorem $x = (x_1, \ldots, x_n, 0, 0, \ldots)$, gdzie $x_k = \operatorname{sgn} \overline{y_k}$ (przypomnijmy, że dla liczby zespolonej z, $\operatorname{sgn} z = z/|z|$, gdy $z \neq 0$ oraz $\operatorname{sgn} z = 0$ w przeciwnym przypadku). Wtedy

$$y^*(x) = \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} |y_k|.$$

Ponieważ $\|x\|_\infty \leqslant 1$ a n jest dowolne, więc $\sum_{k=1}^{\infty} |y_k| \leqslant \|y^*\|$.

4.8. Zadanie. Czy przestrzeń $\ell^{\infty *}$ jest izometrycznie izomorficzna z przestrzenią ℓ^1 ?

4.9. Zadanie. Udowodnić, że jeżeli $1 < p < \infty$, to przyporządkowanie $y \to y^*$ określone wzorem (4. 22) jest izometrycznym izomorfizmem przestrzeni ℓ^q na przestrzeń ℓ^{P*}, gdzie $q = \frac{p}{p-1}$.

4.10. Twierdzenie (Riesz). *Niech $\mathcal{H}$ będzie przestrzenią Hilberta. Dla każdego $y^* \in \mathcal{H}^*$ istnieje, i to tylko jeden, taki wektor $y \in \mathcal{H}$, że*

$$y^*(x) = \langle x,\, y\rangle, \quad x \in \mathcal{H}.$$

Przyporządkowanie $\sigma : y^* \to y$ jest wzajemnie jednoznacznym odwzorowaniem izometrycznym $\mathcal{H}^*$ na $\mathcal{H}$, przy czym $\sigma(y^* + z^*) = \sigma(y^*) + \sigma(z^*)$ i $\sigma(\lambda y^*) = \overline{\lambda}\, \sigma(y^*)$.

Dowód: Niech $\mathcal{H}_0 = \{x \in \mathcal{H} : y^*(x) = 0\}$. Ponieważ y^* jest ciągłym funkcjonałem liniowym, $\mathcal{H}_0$ jest domkniętą podprzestrzenią liniową w $\mathcal{H}$.

Jeżeli $\mathcal{H}_0 = \mathcal{H}$, to $y^*(x) = 0$ dla każdego $x \in \mathcal{H}$, wystarczy wtedy przyjąć $y = 0$. Załóżmy wobec tego, że $\mathcal{H}_0 \neq \mathcal{H}$. W przestrzeni $\mathcal{H}_0^\perp$ istnieje wtedy taki niezerowy element y, że $y^*(y) = \|y\|^2$. Wystarczy dla dowolnego $0 \neq x \in \mathcal{H}_0^\perp$ położyć

$$y = \frac{y^*(x)}{\|x\|^2}\, x.$$

Pokażemy, że element y ma żądane własności. Jeżeli $x \in \mathcal{H}$, to

$$y^*\Big(x - \frac{y^*(x)}{\|y\|^2}\, y\Big) = y^*(x) - y^*(x) = 0,$$

stąd $x - \dfrac{y^*(x)}{\|y\|^2}\, y$ należą do $\mathcal{H}_0$, a więc

$$0 = \Big\langle x - \frac{y^*(x)}{\|y\|^2}\, y\,,\, y\Big\rangle = \langle x,\, y\rangle - y^*(x).$$

Element y jest wyznaczony jednoznacznie przez funkcjonał y^*. Rzeczywiście, jeżeli $y', y'' \in \mathcal{H}$ oraz

$$y^*(x) = \langle x,\, y'\rangle = \langle x,\, y''\rangle,$$

to $\langle x,\, y'-y''\rangle = 0$ dla każdego $x \in \mathcal{H}$, W szczególności dla $x = y'-y''$ otrzymamy $\|y' - y''\|^2 = 0$, a stąd $y' = y''$.

Weźmy teraz dowolny element $y \in \mathcal{H}$ i określmy funkcjonał y^* wzorem

$$y^*(x) = \langle x,\, y\rangle.$$

Z liniowości iloczynu skalarnego względem pierwszej zmiennej wynika, że y^* jest liniowy, a z nierówności Schwarza

$$\big|y^*(x)\big| = \big|\langle x,\, y\rangle\big| \leqslant \|x\|\,\|y\|,$$

że y^* jest ciągły. Wobec tego $y^* \in \mathcal{H}^*$ i $\|y^*\| \leqslant \|y\|$. Z drugiej strony $|\langle x,\, y\rangle| = |y^*(x)| \leqslant \|y^*\|\,\|x\|$ dla $x \in \mathcal{H}$. Podstawiając tu $x = y$ otrzymujemy $\|y\|^2 \leqslant \|y^*\|\,\|y\|$, czyli $\|y\| \leqslant \|y^*\|$. W konsekwencji $\|y^*\| = \|y\|$. □

4.11. TWIERDZENIE HAHNA-BANACHA. *Niech X_0 będzie podprzestrzenią (niekoniecznie domkniętą) przestrzeni liniowej unormowanej X. Każdy ciągły funkcjonał liniowy x_0^* na X_0 można przedłużyć do ciągłego funkcjonału liniowego x^* na całej przestrzeni X bez podnoszenia jego normy.*

Dowód: W X określmy półnormę p wzorem

$$p(x) = \|x\|\,\|x^*\|_{X_0^*}$$

i zastosujmy wersję twierdzenia Hahna-Banacha dla zespolonych przestrzeni liniowych (patrz **??**). Dla przedłużenia x^* otrzymamy wtedy

$$\big|x^*(x)\big| \leqslant p(x) = \|x\|\,\|x_0^*\|_{X_0^*},$$

czyli

$$\|x^*\|_{X^*} \leqslant \|x_0^*\|_{X_0^*}. \quad \square$$

UWAGA. Dla przestrzeni Hilberta przedłużenie, o którym mowa w twierdzeniu Hahna-Banacha jest jednoznaczne. Rzeczywiście, przedłużenie funkcjonału ciągłego z podprzestrzeni do jej domknięcia jest zawsze jednoznaczne. Załóżmy wobec tego, że y_0^* jest ciągłym funkcjonałem liniowym na właściwej domkniętej podprzestrzeni $\mathcal{H}_0$ przestrzeni Hilberta $\mathcal{H}$. Z twierdzenia Riesza 4.10 wynika, że

$$y_0^*(x) = \langle x,\, y_0\rangle, \quad x \in \mathcal{H}_0,$$

dla pewnego wektora $y_0 \in \mathcal{H}_0$, przy czym $\|y_0^*\| = \|y_0\|$. Funkcjonał y^* jest przedłużeniem y_0^* na $\mathcal{H}$ wtedy i tylko wtedy, gdy

$$y^*(x) = \langle x,\, y\rangle, \quad x \in \mathcal{H},$$

oraz $\langle x,\, y - y_0\rangle = 0$ dla każdego $x \in \mathcal{H}_0$ czyli, gdy $y - y_0 \perp \mathcal{H}_0$. W konsekwencji

$$\|y^*\| = \|y\| = \sqrt{\|y_0\|^2 + \|y - y_0\|^2} \geqslant \|y_0\| = \|y_0^*\|,$$

a równość zachodzi tylko wtedy, gdy $y = y_0$.

4.12. Twierdzenie (Hahn). *Niech X_0 będzie podprzestrzenią przestrzeni unormowanej X oraz $x_0 \in X$, przy czym*

$$\operatorname{dist}(x_0, X_0) = \inf_{x \in X_0} \|x_0 - x\| = d.$$

Wtedy istnieje taki funkcjonał liniowy $x^* \in X^*$, że

$$x^*(x) = 0 \quad \textit{dla} \quad x \in X_0, \qquad x^*(x_0) = d, \qquad \|x^*\| \leqslant 1.$$

Dowód: Możemy oczywiście założyć, że $d > 0$. Wtedy każdy element y przestrzeni $X_1 = \operatorname{lin}(X_0 \cup \{x\})$ ma jednoznaczne przedstawienie w postaci $= x + \lambda x_0$, gdzie $x \in X_0$ i $\lambda \in \mathbb{C}$. Na przestrzeni X_1 określmy funkcjonał liniowy x_1^* wzorem

$$x_1^*(y) = \lambda\, d.$$

Mamy wówczas $x_1^*(x) = 0$ dla $x \in X_0$ i $x_1^*(x_0) = d$. Funkcjonał x_1^* jest ciągły na X_1 i $\|x_1^*\| \leqslant 1$. Rzeczywiście

$$\|x + \lambda x_0\| = |\lambda| \left\| \frac{1}{\lambda} x + x_0 \right\| \geqslant |\lambda|\, d = \left| x_1^*(x + \lambda x_0)\right|$$

dla $\lambda \neq 0$, a więc $\|x_1^*\| \leqslant 1$. Dowolne przedłużenie x_1^* na całą przestrzeń, zachowujące jego normę, spełnia tezę. □

4.13. Wniosek. *Jeżeli $x^*(x) = 0$ dla wszystkich $x^* \in X^*$, to $x = 0$. Co więcej*

$$\|x\| = \max_{\|x^*\| \leqslant 1} \left| x^*(x)\right|. \quad \square$$

Zauważmy, że dla każdego $x \in X$ wzór

$$x^{**}(x^*) = x^*(x), \quad x^* \in X^*, \tag{4.23}$$

określa ciągły funkcjonał liniowy na X^*. Z Wniosku 4.13 wynika, że $\|x^{**}\| = \|x\|$. Udowodniliśmy w ten sposób:

4.14. TWIERDZENIE. *Przyporządkowanie* $x \to x^{**}$ *określone wzorem* (4.23) jest izometrycznym izomorfizmem przestrzeni unormowanej X na podprzestrzeń przestrzeni X^{**}. $\square$

Refleksywność

Przestrzeń Banacha X nazywamy **refleksywną**, jeżeli naturalne włożenie $\tau_X : X \to X^{**}$, określone wzorem

$$\tau_X x(x^*) = x^*(x), \quad x^* \in X^*,$$

jest (izometrycznym izomorfizmem) przestrzeni X na przestrzeń X^{**} tzn., gdy dla każdego $x^{**} \in X^{**}$ istnieje taki wektor $x \in X$, że $x^{**}(x^*) = x^*(x)$ dla każdego $x^* \in X^*$.

4.15. FAKT. Z wniosku 4.13 wynika, że jeżeli przestrzeń Banacha jest refleksywna, to każdy ciągły funkcjonał liniowy $x^* \in X^*$ osiąga swoją normę na kuli jednostkowej przestrzeni X.
Oznacza to, że ani przestrzeń $\boldsymbol{c_0}$, ani ℓ^1 nie są refleksywne. Na przestrzeni $\boldsymbol{c_0}$ normy nie osiąga żaden funkcjonał wyznaczony przez ciąg $y = (y_1, y_2, \dots) \in \ell^1$, dla którego $y_n \neq 0$ dla nieskończenie wielu n, zaś na przestrzeni ℓ^1 żaden funkcjonał wyznaczony przez ciąg $y = (y_1, y_2, \dots) \in \ell^\infty$, dla którego $\|y\|_\infty \neq |y_n|$ dla wszystkich n, np. przez ciąg $y_n = n/(n+1)$.

Prawdziwe jest twierdzenie odwrotne do powyższego, ale dowód jest znaczne trudniejszy.

4.16. LEMAT. *Niech* $\varphi : X \to Y$ *będzie ciągłym odwzorowaniem liniowym przestrzeni Banacha* X *w przestrzeń Banacha* Y. *Wówczas odwzorowanie sprzężone* $\varphi_* : Y^* \to X^*$, *określone wzorem*

$$\varphi_* y^*(x) = y^*(\varphi x), \quad x \in X,$$

też jest ciągłe oraz $\|\varphi_*\| \leqslant \|\varphi\|$.

Jeśli odwzorowanie φ *jest izomorfizmem, to izomorfizmem jest również odwzorowanie* φ_*.

Niech $\varphi_{**} : X^{**} \to Y^{**}$ *oznacza odwzorowanie* $\varphi_{**} = (\varphi_*)_*$ *sprzężone do* φ_*, *zaś* τ_X *i* τ_Y *naturalne włożenia odpowiednio przestrzeni* X *w* X^{**} *oraz* Y *w* Y^{**}. *Wtedy*

$$\varphi_{**}\tau_X = \tau_Y \varphi,$$

tzn. następujący diagram jest przemienny

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ \downarrow{\scriptstyle \tau_X} & & \downarrow{\scriptstyle \tau_Y} \\ X^{**} & \xrightarrow{\varphi_{**}} & Y^{**} \end{array}$$

Dowód: Pierwsza część lematu jest oczywista.

Jeżeli φ jest izomorfizmem X na Y, to prosty rachunek daje $(\varphi^{-1})_*\varphi_* = \mathrm{id}_{Y^*}$ oraz $\varphi_*(\varphi^{-1})_* = \mathrm{id}_{X^*}$, tzn. $(\varphi_*)^{-1} = (\varphi^{-1})_*$. Z pierwszej części lematu wnosimy, że φ_* i $(\varphi_*)^{-1}$ są ciągłe, czyli że φ_* jest homeomorfizmem.

Dla dowodu ostatniej części lematu wybierzmy dowolnie $x \in X$ oraz $y^* \in Y^*$. Wtedy

$$\varphi_{**}\tau_X x(y^*) = \tau_X x(\varphi_* y^*) = \varphi_* y^*(x)$$

oraz

$$\tau_Y \varphi x(y^*) = y^*(\varphi x) = \varphi_* y^*(x),$$

a więc $\varphi_{**}\tau_X = \tau_Y\varphi$. $\square$

4.17. Twierdzenie. *Przestrzeń Banacha izomorficzna z przestrzenią refleksywną jest także refleksywna.*

Dowód: Załóżmy, że $\varphi : X \to Y$ jest izomorfizmem refleksywnej przestrzeni Banacha X na przestrzeń Banacha Y. Wtedy odwzorowania $\varphi_{**} : X^{**} \to Y^{**}$ oraz włożenie $\tau_X : X \to X^{**}$ są także izomorfizmami. Z lematu 4.16 wynika, że włożenie $\tau_Y : Y \to Y^{**}$ ma postać $\tau_Y = \varphi_{**}\tau_X\varphi^{-1}$, jest więc izomorfizmem. $\square$

4.18. Lemat. *Domknięta podprzestrzeń przestrzeni refleksywnej jest refleksywna.*

Dowód: Niech X_0 będzie domkniętą podprzestrzenią liniową przestrzeni refleksywnej X i niech $\kappa : X^* \to X_0^*$ oznacza operację zwężania funkcjonału. Jest to operacja sprzężona do identycznościowego włożenia X_0 w X. Z twierdzenia Hahna-Banacha wynika, że κ jest kontrakcją oraz $\kappa(X^*) = X_0^*$. Dla $x_0^{**} \in X_0^{**}$ określmy wektor $x_0 \in X$ wzorem $x_0 = \tau_X^{-1}\kappa_* x_0^{**}$. Dla $x^* \in X^*$ mamy wtedy

$$x^*(x_0) = \kappa_* x_0^{**}(x^*) = x_0^{**}(\kappa x^*).$$

Pokażemy, że w istocie $x_0 \in X_0$. Gdyby tak nie było, to na mocy twierdzenia Hahna w X^* istniałby taki funkcjonał x^*, że $x^*(x_0) = 1$, lecz $x^*(x) = 0$ dla wszystkich $x \in X_0$, tzn. $\kappa x^* = 0$, to nie jest możliwe, bo wtedy mielibyśmy

$$1 = x^*(x_0) = x_0^{**}(\kappa x^*) = x_0^{**}(0) = 0.$$

Skoro $x_0 \in X_0$, to dla $x^* \in X^*$ mamy

$$x^*(x_0) = \kappa x^*(x_0) = \tau_{X_0} x_0(\kappa x^*),$$

a ponieważ $\kappa(X^*) = X_0^*$, więc $x_0^{**} = \tau_{X_0} x_0$, w konsekwencji $\tau_{X_0}(X_0) = X_0^{**}$. □

4.19. PRZYKŁAD. Wiemy, że przestrzeń $\boldsymbol{c_0}$ nie jest refleksywna. Z lematu 4.18 wynika, że refleksywne nie są wobec tego także przestrzenie $\boldsymbol{c}$ i $\boldsymbol{m}$ $(= \ell^\infty)$.

4.20. TWIERDZENIE. *Przestrzeń Banacha jest refleksywna wtedy i tylko wtedy, gdy przestrzeń z nią sprzężona jest refleksywna.*

Dowód: Jeżeli X jest przestrzenią refleksywną, to odwzorowanie $\tau_X^{-1} : X^{**} \to X$ jest izomorfizmem. Z lematu 4.16 wynika, że wtedy odwzorowanie $\tau_{X^*} = (\tau_X^{-1})_*$ jest izomorfizmem X^* na X^{***}, to zaś oznacza refleksywność przestrzeni sprzężonej X^*.

Załóżmy teraz, że X^* jest przestrzenią refleksywną. Z udowodnionej już części twierdzenia wynika, że X^{**} jest refleksywna. Przestrzeń X jest izomorficzna z domkniętą podprzestrzenią liniową $\tau_X(X)$ przestrzeni X^{**}, jest więc przestrzenią refleksywną na mocy lematów 4.16 i 4.18. □

4.21. WNIOSEK. *Jeżeli X jest nierefleksywną przestrzenią Banacha, to wszystkie kanoniczne inkluzje $X \subset X^{**} \subset X^{(iv)} \subset \dots$ oraz $X^* \subset X^{***} \subset X^{(v)} \subset \dots$ są właściwe.*

4.22. WNIOSEK. *Jeżeli X_0 jest domkniętą podprzestrzenią liniową przestrzeni refleksywnej X, to X_0 i X/X_0 są refleksywne.*

Dowód: Refleksywność podprzestrzeni X_0 udowodniliśmy już w lemacie 4.18. Przestrzeń $(X/X_0)^*$ jest izometrycznie izomorficzna z anihilatorem

$$X_0^\perp = \{x^* \in X^* : x^*|_{X_0} \equiv 0\}$$

podprzestrzeni X_0. Ponieważ $X_0^\perp$ jest domkniętą podprzestrzenią liniową w X^*, jest refleksywna, dlatego przestrzeń $(X/X_0)^*$, a w konsekwencji także przestrzeń X/X_0, jest refleksywna. □

4.23. PRZYKŁAD. Niech S będzie przestrzenią topologiczną normalną. Pokażemy, że przestrzeń $C(S)$ jest refleksywna wtedy i tylko wtedy, gdy S jest zbiorem skończonym.

Jeżeli zbiór S jest skończony, to przestrzeń $C(S)$ jest refleksywna, bo ma skończony wymiar. Jeżeli zaś S jest zbiorem nieskończonym, to zawiera podzbiór domknięty S_0 będący ciągiem zbieżnym lub ciągiem izolowanym. Przestrzeń $C(S_0)$

jest wtedy izometrycznie izomorficzna odpowiednio z przestrzenią $\boldsymbol{c}$ lub $\boldsymbol{m}$, w każdym razie z przestrzenią nierefleksywną. Z drugiej strony, jak wiemy z przykładu 1.29, przestrzeń $C(S_0)$ jest izometrycznie izomorficzna z przestrzenią ilorazową $C(S)/X$, gdzie X jest domkniętą podprzestrzenią liniową w $C(S)$ złożoną z funkcji zerujących się na zbiorze S_0. Z wniosku 4.22 wynika, że $C(S)$ nie jest przestrzenią refleksywną.

Przestrzeń sprzężona z przestrzenią C(S)

Opis wszystkich ciągłych funkcjonałów liniowych na przestrzeni $C(S)$ jest trudny i wymaga znajomości ogólnej teorii miary, dlatego poprzedzimy go opisem ciągłych funkcjonałów liniowych na przestrzeni $C[a,b]$. Opis taki można otrzymać w sposób zdecydowanie bardziej elementarny.

W przykładzie 4.3 zobaczyliśmy, że każda funkcja x_0, całkowalna w sensie Lebesgue'a na przedziale $[0,1]$, określa ciągły funkcjonał liniowy na przestrzeni $C[0,1]$ wzorem

$$x^*(x) = \int_0^1 x(t)\, x_0(t)\, dt. \tag{4.24}$$

Z drugiej strony przestrzeń $C[0,1]$ posiada ciągłe funkcjonały liniowe zupełnie innej natury, np. takie, jak

$$x^*(x) = \tfrac{1}{2}\, x(0) + \tfrac{1}{2}\, x(1). \tag{4.25}$$

Co łączy te przykłady? Otóż oba można wyrazić w postaci całki z funkcji x, trzeba jednak posłużyć się ogólniejszym pojęciem całki niż znana nam całka Riemanna (a w pewnym sensie ogólniejszym nawet niż całka Lebesgue'a), mianowicie całką Sieltjesa.

Niech x i f będą funkcjami na przedziale $[a,b]$, z których pierwsza jest ciągła a druga ograniczona. Dla dowolnego rozbicia P przedziału $[a,b]$ punktami

$$a = t_0 < t_1 < t_2 < \ldots < t_n = b$$

oznaczmy przez $S(P,x,f)$ sumę

$$S(P,x,f) = \sum_{k=1}^{n} x(t_k)\big[f(t_k) - f(t_{k-1})\big].$$

Liczbę $\Delta P = \max\{|t_k - t_{k-1}| : k = 1, 2, \dots, n\}$ będziemy nazywać **średnicą rozbicia** P. Powiemy, że funkcja x jest **całkowalna względem funkcji** f w sensie Stieltjesa, jeżeli istnieje taka liczba A, że dla każdej liczby $\varepsilon > 0$ można tak dobrać liczbę $\delta > 0$, aby dla rozbicia P warunek $\Delta P < \delta$ pociągał

$$\big|A - S(P, x, f)\big| < \varepsilon.$$

Liczbę A oznaczymy wtedy

$$\int_0^1 x(t)\, df(t)$$

i nazwiemy całką z funkcji x względem funkcji f.

Zbadamy jaki warunek winna spełniać funkcja f, aby całka Stieltjesa względem f określała ciągły funkcjonał liniowy na przestrzeni $C[a, b]$, tj. aby każda funkcja ciągła x była całkowalna względem f i aby

$$\Big|\int_0^1 x(t)\, df(t)\Big| \leqslant C\, \|x\|_\infty.$$

Jeżeli dla rozbicia $P = \{t_0, t_1, t_2, \dots, t_n\}$ wybierzemy funkcję x tak, żeby $x(t_k) = \operatorname{sgn}\big[f(t_k) - f(t_{k-1})\big]$ dla $k = 1, 2, \dots, n$ oraz $\|x\|_\infty \leqslant 1$, to suma $S(P, x, f)$ przyjmie wartość

$$\sum_{k=1}^{n} \big|f(t_k) - f(t_{k-1})\big|. \tag{4.26}$$

Należy więc oczekiwać, że dla funkcji f wszystkie takie sumy są wspólnie ograniczone.

Definicja. Jeżeli zbiór sum (4.26) jest ograniczony, to mówimy, że funkcja f ma w przedziale $[a, b]$ **wahanie ograniczone**, a kres górny tych sum nazywamy **wahaniem funkcji** f w przedziale $[a, b]$ i oznaczamy $\operatorname*{Var}_{[a,b]} f$.

Przyjrzyjmy się bliżej funkcjom o wahaniu ograniczonym. Jest oczywiste, że gdy funkcja f ma w przedziale $[a, b]$ ograniczoną pochodną, to ma też w tym przedziale ograniczone wahanie. Ciągłość na to jednak nie wystarcza, gdyż np. funkcja $f : [0, 1] \to \mathbb{R}$ postaci $f(0) = 0$, $f(t) = t \cos\frac{\pi}{t}$ dla $t \neq 0$, jest ciągła, a dla rozbicia $P = \big\{0, \frac{1}{n}, \frac{1}{n-1}, \dots, \frac{1}{2}, 1\big\}$ otrzymujemy

$$\operatorname*{Var}_{[0,1]} f \geqslant \frac{1}{n} + \sum_{k=2}^{n} \Big(\frac{1}{k} + \frac{1}{k-1}\Big) \geqslant \log n.$$

4.24. TWIERDZENIE. *Funkcja zespolona f ma w przedziale $[a,b]$ wahanie ograniczone wtedy i tylko wtedy, gdy obie funkcje rzeczywiste $\operatorname{Re} f$ oraz $\operatorname{Im} f$ mają tą własność. Funkcja rzeczywista ma w przedziale $[a,b]$ wahanie ograniczone wtedy i tylko wtedy, gdy jest sumą dwóch funkcji monotonicznych.*

Dowód: Pierwsza część twierdzenia jest oczywista. Dla dowodu części drugiej zauważmy, że gdy funkcja f jest monotoniczna w przedziale $[a,b]$, to ma w nim ograniczone wahanie i $\operatorname{Var}_{[a,b]} f = \big|f(b) - f(a)\big|$. Zatem suma skończonej liczby takich funkcji też ma wahanie skończone w tym przedziale.

Załóżmy teraz, że $\operatorname{Var}_{[a,b]} f < \infty$. Funkcja $f_1(t) = \operatorname{Var}_{[a,t]} f$ jest rosnąca w przedziale $[a,b]$. Wystarczy zatem pokazać, że funkcja $f_2 = f_1 - f$ też jest rosnąca. Otóż dla $s < t$ z przedziału $[a,b]$ mamy

$$f_2(t) - f_2(s) = \operatorname*{Var}_{[s,t]} f - \big(f(t) - f(s)\big) \geqslant \big|f(t) - f(s)\big| - \big(f(t) - f(s)\big) \geqslant 0$$

na mocy definicji wahania funkcji. □

4.25. ZADANIE. Pokazać, że gdy f jest funkcją klasy C^1 na przedziale $[a,b]$, to

$$\operatorname*{Var}_{[a,b]} f = \int_a^b \big|f'(t)\big|\, dt.$$

Jesteśmy gotowi teraz do przedstawienia charakteryzacji wszystkich ciągłych funkcjonałów liniowych na przestrzeni $C[a,b]$.

4.26. TWIERDZENIE. *Jeżeli funkcja $f : [a,b] \to \mathbb{C}$ ma wahanie ograniczone, to każda funkcja $x \in C[a,b]$ jest całkowalna względem f oraz*

$$\Big|\int_a^b x(t)\, df(t)\Big| \leqslant \|x\|_\infty \cdot \operatorname*{Var}_{[a,b]} f.$$

Zatem funkcja f wzorem

$$x^*(x) = \int_a^b x(t)\, df(t) \tag{4.27}$$

określa ciągły funkcjonał liniowy na przestrzeni $C[a,b]$.

Każdy funkcjonał $x^* \in C[a,b]^*$ jest takiej postaci.

Dowód: Ustalmy dowolnie funkcję $x \in C[a,b]$. Ponieważ jest to funkcja jednostajnie ciągła, więc dla każdego $\varepsilon > 0$ istnieje taka liczba $\delta > 0$, że $|s-t| < \delta$

pociąga $\left|x(s)-x(t)\right| < \varepsilon$. Wybierzmy teraz dowolne rozbicie P przedziału $[a,b]$ o średnicy $\Delta P < \delta$ i niech P' będzie dowolnym jego rozdrobnieniem. Twierdzimy, że

$$\left|S(P,x,f) - S(P',x,f)\right| < \varepsilon \operatorname*{Var}_{[a,b]} f. \tag{4.28}$$

Jest tak w istocie, bo jeśli rozbicie P' ma postać $a = t_0 < t_1 < t_2 < \ldots < t_n = b$, zaś rozbicie P postać $a = t_0 = t_{k_0} < t_{k_1} < \ldots < t_{k_r} = t_n = b$, to

$$S(P,x,f) - S(P',x,f) = \sum_{k=1}^{n} \left[x(s_k) - x(t_k)\right]\left[f(t_k) - f(t_{k-1})\right],$$

gdzie $s_k = t_{k_i}$, gdy $k_{i-1} < k \leqslant k_i$. Wystarczy więc zauważyć, że $|s_k - t_k| < \delta$ dla wszystkich k.

Z nierówności (4.28) wynika, że gdy P i Q są rozbiciami przedziału $[a,b]$, dla których $\Delta P < \delta$, $\Delta Q < \delta$, to

$$\left|S(P,x,f) - S(Q,x,f)\right| < 2\,\varepsilon \operatorname*{Var}_{[a,b]} f, \tag{4.29}$$

wystarczy bowiem w nierówności (4.28) wybrać za P' wspólne rozdrobnienie rozbić P i Q.

Wybierzmy teraz dowolny ciąg normalny $\{Q_n\}$ rozbić przedziału $[a,b]$, tj. ciąg rozbić, dla którego ΔQ_n dąży do zera przy $n \to \infty$. Z nierówności (4.29) wynika, że sumy $S(Q_n,x,f)$ tworzą ciąg fundamentalny liczb zespolonych, zatem ciąg zbieżny do pewnej liczby A. Podstawiając w (4.29) Q_n w miejsce Q i przechodząc do granicy przy $n \to \infty$ otrzymamy

$$\left|S(P,x,f) - A\right| \leqslant 2\,\varepsilon \operatorname*{Var}_{[a,b]} f,$$

gdy tylko $\Delta P < \delta$. Oznacza to z definicji całki Stieltjesa, że funkcja x jest całkowalna względem funkcji f. W ten sposób zakończyliśmy dowód pierwszej części twierdzenia.

Pozostało nam do pokazania, że każdy funkcjonał $x^* \in C[a,b]^*$ jest postaci (4.27) dla pewnej funkcji f o ograniczonym wahaniu na przedziale $[a,b]$. Ponieważ $C[a,b]$ jest podprzestrzenią liniową przestrzeni $B[a,b]$ wszystkich funkcji ograniczonych na $[a,b]$, funkcjonał x^* można na mocy twierdzenia Hahna-Banacha przedłużyć do ciągłego funkcjonału liniowego na całą przestrzeń $B[a,b]$, i to bez podwyższania normy. Weźmy taki przedłużony funkcjonał i oznaczy go tym samym symbolem x^* a dla $t \in [a,b]$ połóżmy

$$f(t) = x^*\big(\chi_{[a,t]}\big).$$

Twierdzimy, że to f jest poszukiwaną przez nas funkcją.

Najpierw pokażemy, że f ma ograniczone wahanie na $[a,b]$. Wybierzmy w tym celu dowolne rozbicie $P=\{t_0,t_1,t_2,\dots,t_n\}$ tego przedziału i oznaczmy przez x_P funkcję z $B[a,b]$ postaci

$$x_P = a_1\chi_{[t_0,t_1]} + \sum_{k=2}^{n} a_k\chi_{(t_{k-1},t_k]}.$$

Jeśli wybierzemy $a_k = \operatorname{sgn}\big[f(t_k)-f(t_{k-1})\big]$ i uwzględnimy, że $f(t_0)=0$, to otrzymamy

$$x^*(x_P) = \sum_{k=1}^{n} a_k\big[f(t_k)-f(t_{k-1})\big] = \sum_{k=1}^{n}\big|f(t_k)-f(t_{k-1})\big|,$$

a ponieważ $x^*(x_P) \leqslant \|x^*\|\,\|x_P\|_\infty \leqslant \|x^*\|$, więc

$$\operatorname*{Var}_{[a,b]} f = \sup_P x^*(x_P) \leqslant \|x^*\|.$$

Aby zakończyć dowód obierzmy w definicji funkcji x_P liczby $a_k = x(t_k)$. Dostaniemy wtedy

$$x^*(x_P) = S(P,x,f).$$

Jeśli $\{P_n\}$ jest ciągiem normalnym rozbić przedziału $[a,b]$, to ciąg funkcji $\{x_{P_n}\}$ zbiega jednostajnie do funkcji x, więc na mocy ciągłości funkcjonału x^*

$$x^*(x) = \lim_{n\to\infty} x^*\big(x_{P_n}\big) = \lim_{n\to\infty} S(P_n,x,f) = \int_a^b x(t)\,df(t). \quad \square$$

Uwaga. Funkcja f z twierdzenia 4.26 nie jest określona jednoznacznie. Dodanie do niej funkcji stałej nie zmienia ani wartości całki ani wahania funkcji. Dlatego od funkcji f wymaga się często, by $f(a)=0$. Jednak i ten warunek nie gwarantuje jednoznaczności, nawet przy dodatkowym żądaniu, by wahanie funkcji na $[a,b]$ i norma odpowiadającego jej funkcjonału były równe. Widać to na przykładzie funkcji f_a, określonej na przedziale $[0,1]$ wzorem: $f_a(t)=0$, gdy $t<\frac12$, $f_a(\frac12)=a$ oraz $f_a(t)=1$ gdy $t>\frac12$. Wtedy

$$\int_0^1 x(t)\,df_a(t) = x\big(\tfrac12\big),$$

niezależnie od przyjętej wartości parametru a, zatem wszystkie funkcje f_a wyznaczają ten sam ciągły funkcjonał liniowy na przestrzeni $C[0,1]$, a jego norma wynosi 1. Jeżeli $0\leqslant a\leqslant 1$, to także $\operatorname{Var}_{[0,1]} f_a = 1$.

Powróćmy jeszcze na chwilę do przykładów funkcjonałów z początku tego podrozdziału. Oba funkcjonały (4. 24) i (4. 25) mają postać (4. 27), pierwszy dla funkcji

$$f(t) = \int_0^t x_0(s)\,ds, \quad 0 \leqslant t \leqslant 1,$$

a drugi dla funkcji

$$f(0) = 0,\ f(t) = \tfrac{1}{2} \text{ dla } 0 < t < 1 \text{ oraz } f(1) = 1.$$

Przypadek ogólny

Niech S będzie przestrzenią topologiczną lokalnie zwartą. Zespoloną funkcję μ, określoną na σ-ciele $\mathcal{B}$ wszystkich podzbiorów borelowskich S, nazywamy **miarą Radona** jeżeli μ jest:

1. **przeliczalnie addytywna**, tj.

$$\mu\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \sum_{n=1}^{\infty} \mu(B_n),$$

gdy zbiory $B_1, B_2, B_3, \ldots$ są parami rozłączne.

2. **regularna**, tj. dla każdego $\varepsilon > 0$ istnieją takie zbiory K i U, z których pierwszy jest zwarty a drugi otwarty, że $K \subset B \subset U$ oraz $\big|\mu(A)\big| < \varepsilon$ dla każdego zbioru borelowskiego A zawartego w $U \setminus K$.

Zbiór wszystkich miar Radona na S tworzy przestrzeń liniową, oznaczaną przez $M(S)$.

Jeżeli μ jest miarą Radona, to $\operatorname{Re}\mu$ oraz $\operatorname{Im}\mu$ są także miarami Radona, a jeżeli μ jest miarą nieujemną, to warunek regularności przyjmuje prostszą formę: dla każdego $\varepsilon > 0$ istnieją takie zbiory K i U, z których pierwszy jest zwarty a drugi otwarty, że $K \subset B \subset U$ oraz $\mu(U \setminus K) < \varepsilon$.

4.27. LEMAT. *Miara Radona jest funkcją ograniczoną.*

Dowód: Pokażemy, że w przeciwnym przypadku można wskazać taki ciąg zbiorów $B_1, B_2, B_3, \ldots$ parami rozłącznych, że szereg

$$\sum_{n=1}^{\infty} \mu(B_n)$$

jest rozbieżny. Niech $A_1, A_2, A_3, \ldots$ będzie takim ciągiem zbiorów borelowskich, że $\big|\mu(A_n)\big| \geqslant n + \sum_{k=1}^{n-1} \big|\mu(A_k)\big|$. Wtedy

$$\Big|\mu\Big(\bigcup_{k=1}^{n} A_k\Big)\Big| \geqslant \big|\mu(A_n)\big| - \sum_{k=1}^{n-1} \big|\mu(A_k)\big| \geqslant n.$$

Położmy $B_1 = A_1$ oraz $B_{n+1} = A_{n+1} \setminus \bigcup_{k=1}^n A_k$. Wtedy zbiory $B_1, B_2, B_3, \dots$ są parami rozłączne i

$$\Big|\sum_{k=1}^n \mu(B_k)\Big| = \Big|\mu\Big(\bigcup_{k=1}^n A_k\Big)\Big| \to \infty. \quad \Box$$

Wahaniem miary Radona μ na zbiorze borelowskim B nazywamy kres górny sum $\sum_{k=1}^n \big|\mu(B_k)\big|$, gdzie $B_1, B_2, \dots, B_n$ jest dowolnym rozbiciem zbioru B na skończoną liczbę zbiorów borelowskich i oznaczamy ją $|\mu|(B)$. Wahanie miary to odpowiednik wahania funkcji, o którym była mowa poprzednio.

4.28. LEMAT. *Funkcja $|\mu|$ jest miarą Radona na S a dla każdego zbioru borelowskiego B zachodzą nierówności*

$$\sup_{A\subset B} \big|\mu(A)\big| \leqslant |\mu|(B) \leqslant 4 \sup_{A\subset B} \big|\mu(A)\big|.$$

Dowód: Pierwszą część lematu dowodzi się w sposób standartowy, dlatego dowód ten pominiemy. Pozostaje więc tylko dowód drugiej z nierówności. Załóżmy, że zbiory $B_1, B_2, \dots, B_n$ dają rozbicie zbioru B. Zbiór wskaźników $\{1, 2, \dots, n\}$ rozdzielmy na podzbiory I_1 i I_2, zaliczając do pierwszego z nich wszystkie te wskaźniki k, dla których $\operatorname{Re}\mu(B_k) \geqslant 0$ a do drugiego pozostałe wskaźniki. Wtedy

$$\begin{aligned}\sum_{k=1}^n \big|\operatorname{Re}\mu(B_k)\big| &= \sum_{k\in I_1} \operatorname{Re}\mu(B_k) - \sum_{k\in I_2} \operatorname{Re}\mu(B_k)\\ &= \operatorname{Re}\Big(\bigcup_{k\in I_1} B_k\Big) - \operatorname{Re}\Big(\bigcup_{k\in I_2} B_k\Big) \leqslant 2 \sup_{A\subset B} \big|\mu(A)\big|.\end{aligned}$$

Podobnego oszacowania dowodzimy dla sumy $\sum_{k=1}^n \big|\operatorname{Im}\mu(B_k)\big|$. W konsekwencji otrzymamy

$$\sum_{k=1}^n \big|\mu(B_k)\big| \leqslant 4 \sup_{A\subset B} \big|\mu(A)\big|,$$

a stąd postulowaną nierówność. $\Box$

Możemy teraz przystąpić do sformułowania zapowiadanego twierdzenia opisującego ciągłe funkcjonały liniowe na przestrzeni $C(S)$ w przypadku ogólnym.

4.29. TWIERDZENIE (O POSTACI FUNKCJONAŁÓW LINIOWYCH NA PRZESTRZENI $C(S)$). *Niech S będzie zwartą przestrzenią Hausdorffa. Dla każdego ciągłego funkcjonału liniowego x^* na przestrzeni $C(S)$ istnieje dokładnie jedna miara Radona μ na S o tej własności, że*

$$x^*(x) = \int_S x(s)\, d\mu(s), \qquad x \in C(S).$$

Każdy funkcjonał powyższej postaci jest liniowy i ciągły na przestrzeni $C(S)$, a jego norma wynosi

$$\|x^*\| = |\mu|(S).$$

Dowód istnienia miary μ jest trudny. Przypomina konstrukcję miary Lebesgue'a, a dodatkowym utrudnieniem jest fakt, że w rezultacie otrzymujemy funkcję zbioru o wartościach zespolonych a nie miarę nieujemną. Także dowód jedyności miary i dowód drugiej części twierdzenia używa wyłącznie metod ogólnej teorii miary, wymaga więc znajomości tych metod. Z tego powodu dowód pomijamy, odsyłając np. do podręcznika [1], Twierdzenie VIII.2.2.

Na pierwszy rzut oka może się wydawać, że mimo zapowiedzianej ogólności twierdzenie 4.29 dotyczy tylko bardzo szczególnej klasy przestrzeni topologicznych. Nie daje wobec tego opisu ciągłych funkcjonałów liniowych nawet dla takiej przestrzeni, jak $C(\mathbb{R})$. Spodziewamy się tu, że przestrzenią sprzężoną $C(\mathbb{R})^*$ jest przestrzeń miar Radona $M(\mathbb{R})$. Okazuje się, że opis taki z twierdzenia 4.29 można otrzymać oraz, że jest o wiele bardzie skomplikowany niż możemy przypuszczać.

Wiadomo (patrz **??**), że jeżeli S jest przestrzenią topologiczną całkowicie regularną, to każdą funkcję $x \in C(S)$ można jednoznaczne przedłużyć do funkcji ciągłej $\widetilde{x}$ na uzwarcenia Čecha-Stone'a βS przestrzeni S. Ponieważ S leży gęsto w βS, więc przestrzenie $C(S)$ i $C(\beta S)$ można utożsamić. Twierdzenie 4.29 daje więc opis ciągłych funkcjonałów liniowych na przestrzeni $C(S)$ jako miar Radona na uzwarceniu βS. Dokładniej:

4.30. TWIERDZENIE. *Niech S będzie przestrzenią topologiczną całkowicie regularną i dla funkcji $x \in C(S)$ niech $\widetilde{x}$ oznacza jej jednoznaczne przedłużenie do funkcji ciągłej na uzwarcenia Čecha-Stone'a βS przestrzeni S. Dla każdego funkcjonału $x^* \in C(S)^*$ istnieje dokładnie jedna taka miara Radona μ na βS, że*

$$x^*(x) = \int_{\beta S} \widetilde{x}(s)\, d\mu(s). \quad \square$$

Widać z tego twierdzenia, że przestrzeń $C(\mathbb{R})$ ma znacznie więcej ciągłych funkcjonałów liniowych niż te wyznaczone przez miary Radona na prostej $\mathbb{R}$. W szczególności, jeżeli za miarę μ wybierzemy miarę o masie 1 skupioną w punkcie s_0 narostu $\beta\mathbb{R} \setminus \mathbb{R}$, to otrzymamy funkcjonał

$$x^*(x) = \widetilde{x}(s_0),$$

który nie może być reprezentowany jako całka z funkcji x względem żadnej miary Radona na $\mathbb{R}$.

Z twierdzenia 4.30 można też otrzymać opis przestrzeni $(\ell^\infty)^*$. Przestrzeń ℓ^∞ możemy utożsamić z przestrzenią $C(\mathbb{N})$, gdzie $\mathbb{N}$ oznacza zbiór liczb naturalnych z topologią dyskretną. Zatem ciągłe funkcjonały liniowe na przestrzeni ℓ^∞ mają postać miar Radona na uzwarceniu Čecha-Stone'a $\beta\mathbb{N}$ zbioru $\mathbb{N}$.

Przedstawimy jeszcze jedno zastosowanie twierdzenia 4.29. Niech S będzie przestrzenią topologiczną lokalnie zwartą Hausdorffa. Oznaczmy przez $C_0(S)$ zbiór wszystkich funkcji ciągłych „znikających w nieskończoności", tj. takich funkcji ciągłych $x : S \to \mathbb{C}$, że dla każdego $\varepsilon > 0$ zbiór $\{s \in S : |x(s)| \geqslant \varepsilon\}$ jest zwarty. Każda taka funkcja ma jednoznaczne przedłużenie do funkcji ciągłej $\widetilde{x}$ na uzwarceniu Aleksandrowa $S_\infty = S \cup \{\infty\}$ przestrzeni S (patrz **??**), należy bowiem położyć $\widetilde{x}(\infty) = 0$. Zatem przestrzeń $C_0(S)$ można traktować jako domkniętą podprzestrzeń przestrzeni $C(S_\infty)$. Z twierdzenia Hahna-Banacha wiemy, że każdy ciągły funkcjonał liniowy na przestrzeni $C_0(S)$ możemy przedłużyć do ciągłego funkcjonału liniowego na przestrzeni $C(S_\infty)$. Na mocy twierdzenia 4.29 musi on być reprezentowany przez pewną miarę Radona na przestrzeni $S \cup \{\infty\}$. W szczególności funkcjonał $x^* \in C(S_0)^*$ ma postać

$$x^*(x) = \int_S x(s)\, d\mu(s) + \lambda\widetilde{x}(\infty),$$

jeśli miara μ ma atom w punkcie ∞ o masie λ. Ponieważ $\widetilde{x}(\infty) = 0$, więc w istocie x^* jest reprezentowany przez miarę Radona $\mu \in M(S)$. Udowodniliśmy w ten sposób:

4.31. TWIERDZENIE. *Niech S będzie lokalnie zwartą przestrzenią topologiczną Hausdorffa i niech $C_0(S)$ oznacza przestrzeń Banacha wszystkich zespolonych funkcji ciągłych na S „znikających w nieskończoności". Ciągłym funkcjonałom liniowym odpowiadają w sposób wzajemnie jednoznaczny miary Radona na S, mianowicie*

$$x^*(x) = \int_S x(s)\, d\mu(s), \qquad x \in C_0(S). \quad \square$$

Słaba i *-słaba zbieżność

Ciąg $\{x_n\}$ elementów przestrzeni unormowanej X nazywamy **słabo zbieżnym** do elementu $x_0 \in X$, jeżeli dla każdego ciągłego funkcjonału liniowego $x^* \in X^*$ istnieje granica $\lim\limits_{n\to\infty} x^*(x_n)$ i zachodzi równość

$$\lim_{n\to\infty} x^*(x_n) = x^*(x_0).$$

Element x_0 nazywamy wtedy **słabą granicą** ciągu $\{x_n\}$.

Ciąg $\{x_n\}$ nazywamy **słabo fundamentalnym**, jeżeli dla każdego $x^* \in X^*$ ciąg liczbowy $\{x^*(x_n)\}$ jest fundamentalny (Cauchy'ego) w ciele liczb zespolonych.

4.32. Ćwiczenia.

1. Ciąg słabo zbieżny może mieć tylko jedną granicę.

2. W przestrzeni unormowanej X każdy ciąg zbieżny (w normie) jest słabo zbieżny, i to do tej samej granicy.

3. W przestrzeni skończenie wymiarowej słaba zbieżność jest równoważna zbieżności w normie.

Uwaga. Zbieżność w normie, dla odróżnienia od słabej zbieżności nazywamy czasem **mocną zbieżnością**.

4.33. Twierdzenie. *Słabo zbieżny ciąg elementów przestrzeni unormowanej jest ograniczony.*

Dowód: Każdy element x_n ciągu słabo zbieżnego w przestrzeni unormowanej X wyznacza ciągły funkcjonał liniowy x_n^{**} na przestrzeni X^*, określony wzorem $x_n^{**}(x^*) = x^*(x_n)$, przy czym (jak wiemy z wniosku 4.13) $\|x_n^{**}\| = \|x_n\|$. Dla każdego $x^* \in X^*$ ciąg $\{x_n^{**}(x^*)\}$ jest ograniczony (zbieżny), więc z twierdzenia Banacha-Steinhausa wnosimy, że wszystkie funkcjonały x_n^{**} mają wspólnie ograniczone normy. $\square$

4.34. Przykład. Ciąg $\{x_n\} \subset \ell^p$, $1 < p < \infty$, jest słabo zbieżny do 0 wtedy i tylko wtedy, gdy $\sup_n \|x_n\|_p < \infty$ oraz $\lim\limits_{n\to\infty} x_n(k) = 0$ dla każdego $k = 1, 2, 3, \ldots$. W szczególności ciąg bazowy $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$, z jedynką na n-tym miejscu, jest słabo zbieżny do zera.

Istotnie, ponieważ ciąg $\{x_n\}$ jest ograniczony, więc zbieżność $x^*(x_n) \to 0$ wystarczy sprawdzić tylko dla funkcjonałów x^* z dowolnego podzbioru gęstego w przestrzeni $(\ell^p)^* = \ell^q$, $q = p/(p-1)$, a w konsekwencji z dowolnego podzbioru liniowe gęstego. Takim podzbiorem jest baza $\{e_1, e_2, e_3, \ldots\}$ przestrzeni ℓ^q.

Podobna charakteryzacja słabej zbieżności zachodzi dla przestrzeni $\mathbf{c}_0$ i $\mathbf{c}$, ale nie dla ℓ^1 i ℓ^∞.

4.35. Zadanie. Dowieść, że ciąg funkcji $\{x_n\}$ w przestrzeni $L^p(\mathbb{R})$, $1 < p < \infty$, jest słabo zbieżny do zera wtedy i tylko wtedy, gdy jest ograniczony oraz

$$\int_a^b x_n(t)\, dt \to 0.$$

dla dowolnego przedziału $[a, b]$

4.36. Przykład. Ciąg $\{x_n\}$ funkcji w przestrzeni $C[a, b]$ jest słabo zbieżny do funkcji $x_0 \in C[a, b]$ wtedy i tylko wtedy, gdy $\sup_n \|x_n\|_\infty < \infty$ oraz $\lim\limits_{n\to\infty} x_n(t) = x_0(t)$ dla każdego $t \in [a, b]$. Wynika to z twierdzenia Riesza 4.26 o postaci funkcjonału i twierdzenia Lebesgue'a o zbieżności ograniczonej. Zauważmy, że w przestrzeni $C[a, b]$ pojęcia „słabo zbieżny" i „słabo fundamentalny" są różne. Wystarczy wskazać ograniczony ciąg funkcji ciągłych, zbieżny do funkcji nieciągłej.

Jeszcze lepiej tą różnicę widać w przestrzeni $\mathbf{c}_0$. Ciąg $x_n = (1, 1, \ldots, 1, 0, 0, \ldots)$ z jedynkami do miejsca n-tego jest słabo fundamentalny a nie jest słabo zbieżny.

4.37. Zadanie. Dowieść, że każdy słabo zbieżny ciąg w $C[a, b]$ jest mocno zbieżny w $L^p(a, b)$ dla $1 \leqslant p < \infty$.

4.38. Twierdzenie (Schur). *W przestrzeni ℓ^1 ciąg $\{x_n\}$ jest słabo zbieżny wtedy i tylko wtedy, gdy jest zbieżny w normie, tj. słaba i mocna zbieżność pokrywają się.*

Dowód: Załóżmy nie wprost, że istnieje ciąg $\{x_n\}$, $\|x_n\| \geqslant 1$, słabo zbieżny do zera. Zauważmy, że dla każdego ustalonego p wyrażenie

$$\sum_{k=1}^{p} |x_n(k)|$$

dąży do zera przy $n \to \infty$. Przez indukcję wybieramy takie dwa ciągi rosnące $\{n_i\}$ i $\{p_i\}$, że

$$\sum_{k=1}^{p_i} |x_{n_i}(k)| < \frac{1}{4}, \qquad \sum_{k=p_i+1}^{p_{i+1}} |x_{n_i}(k)| > \frac{3}{4}, \qquad \sum_{k=p_{i+1}}^{\infty} |x_{n_i}(k)| < \frac{1}{4}.$$

Określmy funkcjonał $x^* \in (\ell^1)^* = \ell^\infty$ wzorem

$$x^* = (\mu_1, \mu_2, \mu_3, \ldots),$$

gdzie $\mu_k = \overline{\operatorname{sgn} x_{n_i}(k)}$, gdy $p_i < k \leqslant p_{i+1}$. Wtedy

$$\left|x^*(x_{n_i})\right| \geqslant \sum_{k=p_i+1}^{p_{i+1}} \left|x_{n_i}(k)\right| - \sum_{k=1}^{p_i} \left|x_{n_i}(k)\right| - \sum_{k=p_{i+1}}^{\infty} \left|x_{n_i}(k)\right| > \frac{1}{4}$$

nie dąży do zera. □

4.39. Twierdzenie Mazura. *Jeżeli podzbiór A przestrzeni Banacha jest wypukły i domknięty, to wraz z każdym ciągiem słabo zbieżnym zawiera jego granicę.*

Dowód: Zauważmy, że zbiór A można tak przesunąć (ciąg też), by $0 \in A$. Załóżmy teraz nie wprost, że ciąg $\{x_n\}$ leży w A i jest słabo zbieżny do pewnego $x_0 \notin A$. Niech $0 < \varepsilon < \operatorname{dist}(x_0, A)$ i niech A_ε oznacza zbiór $a_\varepsilon = A + \varepsilon K$, tj.

$$A_\varepsilon = \{x + y : x \in A,\ \|y\| \leqslant \varepsilon\}.$$

Oczywiście A_ε jest również zbiorem wypukłym domkniętym i $x_0 \notin A_\varepsilon$. Ponadto zbiór A_ε jest pochłaniający, bo zawiera pewne otoczenie zera. Niech p oznacza funkcjonał Minkowskiego zbioru A_ε. Wtedy $p(x_0) = \delta > 1$, zatem element $x_1 = \delta^{-1}x_0$ leży w A_ε. Określmy na przestrzeni jednowymiarowej $\mathbb{C}x_1$ rzeczywisty funkcjonał liniowy x_0^* kładąc

$$x_0^*(\lambda x_1) = \lambda.$$

Mamy wtedy $x_0^*(\lambda x_1) = \lambda \leqslant p(\lambda x_1)$ (w istocie dla $\lambda \geqslant 0$ zachodzi równość, a dla $\lambda < 0$ oczywista nierówność). Na mocy twierdzenia Hahna-Banacha możemy x_0^* przedłużyć do funkcjonału liniowego x^* na całe X tak, aby $x^*(x) \leqslant p(x)$. Ponieważ $p(x) \leqslant \varepsilon^{-1}\|x\|$, więc x^* jest ciągłym rzeczywistym funkcjonałem liniowym na przestrzeni X. Ponadto $x^*(x_0) = \delta > 1$ oraz $x^*(x) \leqslant 1$ dla $x \in A$. Zatem x^* oddziela x_0 od zbioru A, co przeczy słabej zbieżności ciągu $\{x_n\}$ do x_0.

W zespolonej przestrzeni Banacha bierzemy funkcjonał liniowy y^* postaci

$$y^*(x) = x^*(x) - i\,x^*(i\,x). \quad \square$$

4.40. Wniosek. *Jeżeli ciąg $\{x_n\}$ jest słabo zbieżny do x_0, to istnieje ciąg (skończonych) kombinacji wypukłych tego ciągu mocno zbieżny do x_0.*

4.41. Przykład. Niech $x_0 \in C(\mathbb{R})$ będzie funkcją, dla której

$$\lim_{t\to-\infty} x_0(t) = 0, \quad \lim_{t\to+\infty} x_0(t) = 1.$$

Zbadamy, czy ciąg $\{x_n\}$, określony równością $x_n(t) = (t+n)$, jest słabo zbieżny w przestrzeni $C(\mathbb{R})$.

Przypuśćmy, że jest to ciąg słabo zbieżny do pewnej funkcji y, wtedy

$$y(t) = \lim_{n\to\infty} x_n(t) = \lim_{n\to\infty} x(t+n) = 1$$

dla każdego $t \in \mathbb{R}$, zatem musi być $y = 1$. Przeczy to jednak twierdzeniu Mazura, gdyż zbiór

$$W = \big\{x \in C(\mathbb{R}) : \lim_{t\to-\infty} x(t) = 0, \lim_{t\to+\infty} x(t) = 1\big\}$$

jest wypukły i domknięty w $C(\mathbb{R})$, zawiera wszystkie funkcje x_n, a nie zawiera funkcji y.

4.42. Zadanie. Wskazać funkcjonał $x^* \in C(\mathbb{R})^*$, dla którego

$$\lim_{n\to\infty} x^*(x_n) \neq x^*(1).$$

W przestrzeni X^*, sprzężonej do przestrzeni unormowanej X poza słabą zbieżnością rozważa się jeszcze inny rodzaj zbieżności.

Definicja. Ciąg $\{x_n^*\}$ elementów przestrzeni X^*, sprzężonej do przestrzeni unormowanej X, nazywamy ***-słabo zbieżnym** do $x_0^* \in X^*$, jeżeli dla każdego $x \in X$ istnieje granica $\lim\limits_{n\to\infty} x_n^*(x)$ i zachodzi równość

$$\lim_{n\to\infty} x_n^*(x) = x_0^*(x).$$

Wtedy x_0^* nazywamy ***-słabą granicą** ciągu $\{x_n^*\}$.

Ciąg $\{x_n^*\}$ jest ***-słabo fundamentalny**, jeżeli dla każdego $x \in X$ ciąg liczbowy $\{x_n^*(x)\}$ jest fundamentalny w $\mathbb{C}$.

4.43. Twierdzenie. *Każdy ciąg $\{x_n^*\}$ słabo zbieżny jest *-słabo zbieżny w tej przestrzeni. Każdy ciąg *-słabo fundamentalny jest *-słabo zbieżny, tzn. każda przestrzeń X^* jest ciągowo *-słabo zupełna.*

Dowód: Jeśli ciąg $\{x_n^*\}$ jest *-słabo fundamentalny, to jest ograniczony, a więc równość

$$x_0^*(x) = \lim_{n\to\infty} x_n^*(x)$$

określa ciągły funkcjonał liniowy na X. $\square$

Uwaga. W przestrzeni refleksywnej pojęcia słabej i *-słabej zbieżności pokrywają się.

4.44. Twierdzenie. *Niech X będzie ośrodkową przestrzenią unormowaną. Z każdego ciągu ograniczonego $\{x_n^*\}$ w X^* można wybrać podciąg *-słabo zbieżny w X^**

Dowód: Niech zbiór $\{e_1, e_2, e_3, \ldots\}$ będzie ośrodkiem w X. Metodą „przekątniową" z ciągu $\{x_n^*\}$ możemy wybrać taki podciąg $\{x_{n_k}\}$, aby ciąg $\{x_{n_k}^*(e_i)\}$ był zbieżny dla każdego $i = 1, 2, 3, \ldots$. Podciąg $\{x_{n_k}\}$ jest wtedy *-słabo fundamentalny w X^*, a na mocy twierdzenia 4.43 jest *-słabo zbieżny. $\square$

Zadania uzupełniające

4.45. W nieskończenie wymiarowej przestrzeni Hilberta każdy ciąg ortonormalny $e_1, e_2, e_3, \ldots$ jest słabo zbieżny do zera. Wskazać konkretny ciąg kombinacji wypukłych tego ciągu mocno zbieżny do zera.

4.46. Wykazać, że jeżeli ciąg $\{x_n\}$ elementów przestrzeni Hilberta jest słabo zbieżny do x_0 oraz $\|x_n\| \to \|x_0\|$, to x_n dąży do x_0 w normie.

4.47. W przestrzeni $\left(\ell^1\right)^* = \ell^\infty$ podać przykład ciągu *-słabo zbieżnego, który nie jest słabo zbieżny.

4.48. Pokazać, że dla dowolnej funkcji $x \in C_0(\mathbb{R})$ ciąg $x_n(t) = x(t+n)$ jest słabo zbieżny w $C_0(\mathbb{R})$.

ROZDZIAŁ V

ZASTOSOWANIA TWIERDZENIA BAIRE'A

Przypomnijmy, że zbiór E przestrzeni metrycznej X nazywamy **brzegowym**, gdy ma puste wnętrze a **zbiorem pierwszej kategorii** w X, jeżeli zawarty jest w skończonej lub przeliczalnej sumie zbiorów domkniętych brzegowych. Twierdzenie Baire'a (patrz **??**) mówi, że: *przestrzeń metryczna zupełna nie jest zbiorem pierwszej kategorii.*

5.1. Twierdzenie Banacha-Steinhausa. *Niech $T_1, T_2, T_3, \ldots$ będzie ciągiem odwzorowań liniowych ciągłych z przestrzeni Banacha X do przestrzeni unormowanej Y. Jeżeli dla każdego $x \in X$ ciąg $\{T_n x\}$ jest ograniczony w Y, to wszystkie operatory T_n mają wspólnie ograniczone normy.*

Dowód: Zauważmy, że dla ustalonego n funkcja $x \to \|T_n x\|$ jest ciągła, zatem każdy ze zbiorów $\{x \in X : \|T_n x\| \leqslant k\}$, $k = 1, 2, 3, \ldots$ jest domknięty w X jako przeciwobraz przedziału domkniętego $[-k, k]$. Wynika stąd, że zbiory

$$E_k = \{x \in X : \sup_n \|T_n x\| \leqslant k\}$$

także są domknięte w X. Z założenia twierdzenia wiemy, że $X = \bigcup_{k=1}^{\infty} E_k$ i że X jest przestrzenią metryczną zupełną. Możemy więc stosować do niej twierdzenie Baire'a **??**. Wnosimy z niego, że przynajmniej jeden ze zbiorów E_k, powiedzmy zbiór E_{k_0}, ma niepuste wnętrze. Możemy więc tak wybrać element $x_0 \in X$ i liczbę $\varepsilon > 0$, by E_{k_0} zawierał wszystkie elementy postaci $x_0 + \varepsilon z$, $z \in X$, $\|z\| \leqslant 1$; innymi słowy, by nierówność $\|T_n(x_0 + \varepsilon z)\| \leqslant k_0$ zachodziła dla wszystkich wskaźników n i wszystkich elementów z z kuli jednostkowej przestrzeni X. Wtedy $\|T_n z\| \leqslant \frac{1}{\varepsilon}\big(k_0 + \|T_n x_0\|\big) \leqslant \frac{2k_0}{\varepsilon}$, co daje oszacowanie

$$\|T_n\| = \sup_{\|z\| \leqslant 1} \|T_n z\| \leqslant \frac{2k_0}{\varepsilon},$$

jednostajne względem n. $\square$

5.2. Wniosek. *Załóżmy, że $T_1, T_2, T_3, \ldots$ są takimi ciągłymi odwzorowaniami liniowymi z przestrzeni Banacha X do przestrzeni unormowanej Y, że dla każdego $x \in X$ ciąg $\{T_n x\}$ jest zbieżny w Y. Wtedy wzór*

$$Tx = \lim_{n\to\infty} T_n x$$

określa ciągłe odwzorowanie liniowe z X do Y i

$$\|T\| \leqslant \liminf_{n\to\infty} \|T_n\|.$$

Dowód: Jest oczywiste, że T jest dobrze określonym odwzorowaniem liniowym z X do Y i że zachodzi postulowane nierówność dla normy T. Ciągłość odwzorowania T jest konsekwencją wspólnej ograniczoności norm $\|T_n\|$, ta zaś wynika bezpośrednio z twierdzenia Banacha-Steinhausa. $\square$

5.3. Zadanie. Pokazać na przykładzie, że jeśli przestrzeń X nie jest zupełna, to T może nie być odwzorowaniem ciągłym.

5.4. Przykład. Pokażemy jak, korzystając z twierdzenia Banacha-Steinhausa, można wykazać istnienie funkcji ciągłej okresowej o okresie 1, której szereg Fouriera nie jest zbieżny w wybranym punkcie.

Na podprzestrzeni

$$X = \big\{x \in C[0,1] : x(0) = x(1)\big\}$$

przestrzeni $C[0,1]$, reprezentującej wszystkie funkcje ciągłe okresowe o okresie 1, określmy ciąg $\{x_n^*\}$ ciągłych funkcjonałów liniowych, przyporządkowujących funkcji x wartość n-tej sumy częściowej jej szeregu Fouriera w punkcie $t = 0$, tj.

$$x_n^*(x) = \sum_{|k|\leqslant n} \widehat{x}(k).$$

Jak pamiętamy z dowodu twierdzenia 3.35, funkcjonał x_n^* ma postać

$$x_n^*(x) = \int_0^1 x(u)\,\frac{\sin(2n+1)\pi u}{\sin \pi u}\,du.$$

Ponieważ

$$\|x_n^*\| = 2\int_0^{1/2} \left|\frac{\sin(2n+1)\pi u}{\sin \pi u}\right| du \geqslant 2\sum_{k=0}^{n-1} \frac{1}{\sin \pi\frac{k+1}{2n+1}} \int_{\frac{k}{2n+1}}^{\frac{k+1}{2n+1}} \big|\sin(2n+1)\pi u\big|\,du$$
$$\geqslant \frac{4}{\pi^2}\sum_{k=0}^{n-1}\frac{1}{k+1} \geqslant \frac{4}{\pi^2}\log n,$$

więc z twierdzenia Banacha-Steinhausa wnosimy, że dla pewnej funkcji $x_0 \in X$ ciąg $\{x_n^*(x_0)\}$ musi być nieograniczony. Szereg Fouriera funkcji x_0 nie jest wobec tego zbieżny w punkcie $t = 0$.

Brak zbieżności w innych punktach można uzyskać dokonując przesunięcia w argumencie funkcji x_0.

Definicja. Powiemy, że ciąg $\{x_n\}$ elementów przestrzeni liniowej unormowanej jest **słabo zbieżny** do elementu x_0, jeżeli dla każdego $x^* \in X^*$ ciąg $\{x^*(x_n)\}$ jest zbieżny do $x^*(x)$.

5.5. Wniosek. *W przestrzeni Banacha każdy ciąg słabo zbieżny jest ograniczony oraz* $\|x_0\| \leqslant \liminf_{n\to\infty} \|x_n\|$. $\square$

5.6. Przykład. W przestrzeni Hilberta każdy ciąg ortonormalny $e_1, e_2, e_3, \ldots$ jest słabo zbieżny do zera. Istotnie, dla każdego x ciąg $\{\langle e_n, x\rangle\}$ dąży do zera, bo z nierówności Bessela

$$\sum_{n=1}^{\infty} \left|\langle e_n, x\rangle\right|^2 \leqslant \|x\|^2 < \infty.$$

5.7. Twierdzenie Banacha o odwzorowaniu otwartym. *Każde ciągłe odwzorowanie liniowe T przestrzeni Banacha X na przestrzeń Banacha Y jest otwarte, tzn. obraz $T(A)$ dowolnego zbioru otwartego A w X jest zbiorem otwartym w Y.*

Dowód: Pokażemy najpierw, że domknięcie $\overline{T(U)}$ obrazu dowolnego otoczenia zera U przestrzeni X zawiera pewne otoczenie zera przestrzeni Y.

Zauważmy, że w przestrzeniach X i Y wystarczy rozpatrywać otoczenia zera postaci εU, εV, gdzie $U = \{x \in X : \|x\| < 1\}$, $V = \{y \in Y : \|y\| < 1\}$ są otwartymi kulami jednostkowymi odpowiednio w X i Y a ε jest liczbą dodatnią.

Ponieważ zbiór wartości operatora T wypełnia przestrzeń Y, więc dla każdego $y \in Y$ istnieje taka liczba naturalna n, że $y \in T(n\,U)$, dostajemy więc rozkład przestrzeni Y

$$Y = \bigcup_{n=1}^{\infty} T(n\,U).$$

Skorzystamy w tym miejscu z twierdzenia Baire'a. Musimy jednak w miejsce zbiorów $T(n\,U)$ wziąć ich domknięcia $\overline{T(n\,U)}$, gdyż same zbiory $T(n\,U)$ nie są na ogół domknięte w Y. Piszemy zatem

$$Y = \bigcup_{n=1}^{\infty} \overline{T(n\,U)}.$$

Wobec założonej zupełności przestrzeni Y, z twierdzenia Baire'a wnioskujemy, że przynajmniej jeden ze zbiorów $\overline{T(n\,U)}$ ma niepuste wnętrze; ale $\overline{T(n\,U)} = n\,\overline{T(U)}$, zatem $\overline{T(U)}$ także ma niepuste wnętrze. Oznaczmy $A = \operatorname{int} \overline{T(U)}$. Ponieważ zbiór $\frac{1}{2}(A - A)$ jest otwartym otoczenie zera w Y i

$$\tfrac{1}{2}(A - A) \subset \tfrac{1}{2}\overline{\big(T(U) - T(U)\big)} = \overline{T(\tfrac{1}{2}U - \tfrac{1}{2}U)} = \overline{T(U)},$$

to dla pewnego $\varepsilon > 0$ zachodzi inkluzja $\varepsilon\, V \subset \frac{1}{2}(A - A) \subset \overline{T(U)}$.

Pokażemy teraz, że inkluzja $\varepsilon\, V \subset \overline{T(U)}$ pociąga inkluzję $\varepsilon\, V \subset T(2U)$, czyli że dla dowolnego $y \in \varepsilon\, V$ możemy tak wybrać element $x \in X$, żeby $\|x\| < 2$ i $Tx = y$. Istotnie, skoro $\varepsilon\, V \subset \overline{T(U)}$, to w U możemy tak wybrać x_0, żeby $\|y - Tx_0\| < \frac{1}{2}\varepsilon$. Element $y - Tx_0$ należy do zbioru $\frac{1}{2}\varepsilon\, V \subset \overline{T(\frac{1}{2}\,U)}$, zatem w kuli $\frac{1}{2}\,U$ możemy tak wybrać element x_1, żeby $\|(y - Tx_0) - Tx_1\| < \frac{1}{4}\varepsilon$. Powtarzając to postępowanie wielokrotnie zbudujemy ciąg $x_0, x_1, x_2, \ldots$ o tej własności, że $\|x_n\| < \frac{1}{2^n}$ oraz

$$\|y - Tz_n\| < \frac{1}{2^{n+1}}\varepsilon. \qquad n = 0, 1, 2, \ldots,$$

gdzie $z_n = x_0 + x_1 + \ldots + x_n$. Ponieważ szereg $\sum_{k=0}^{\infty} \|x_k\|$ jest zbieżny, więc ciąg $\{z_n\}$ jest fundamentalny w X a ponieważ X jest przestrzenią zupełną, jest zbieżny do pewnego elementu x. Dla tego elementu x mamy

$$\|x\| \leqslant \sum_{k=0}^{\infty} \|x_k\| < \sum_{k=0}^{\infty} \frac{1}{2^k} = 2,$$

oraz z ciągłości odwzorowania T

$$\|y - Tx\| = \lim_{n\to\infty} \|y - Tz_n\| = 0,$$

tj. $y = Tx$.

Łatwo już teraz uzyskać tezę twierdzenia. Jeżeli A jest zbiorem otwartym w X oraz $x \in A$, to także $x + \delta\, U \subset A$ dla pewnego $\delta > 0$. Dlatego

$$T(A) \supset T(x + \delta\, U) = Tx + \delta\, T(U) \supset Tx + \tfrac{1}{2}\varepsilon\delta\, V,$$

co oznacza, że element Tx leży w $T(A)$ wraz z pewnym swoim otoczeniem. □

Wszystkie dalsze twierdzenia tego paragrafu są wnioskami z twierdzenia Banacha o odwzorowaniu otwartym.

5.8. Twierdzenie Banacha o odwzorowaniu odwrotnym. *Jeżeli T jest wzajemnie jednoznacznym i ciągłym odwzorowaniem liniowym z przestrzeni Banacha X na przestrzeń Banacha Y, to odwzorowanie odwrotne $T^{-1} : Y \to X$ jest także ciągłe, tzn. przestrzenie X i Y są izomorficzne.*

Dowód: Ciągłość odwzorowania T^{-1} to otwartość odwzorowania T. □

5.9. PRZYKŁAD. Wzajemnie jednoznaczne odwzorowanie liniowe przestrzeni Banacha na przestrzeń Banacha może nie być izomorfizmem topologicznym, choć oczywiście jest izomorfizmem algebraicznym. Niech X będzie przestrzenią Banacha z bazą topologiczną $\{e_k\}_{k=1}^{\infty}$, np. jedną z przestrzeni $\boldsymbol{c}_0$, $\boldsymbol{c}$ lub ℓ^p, $1 \leqslant p < \infty$. Bazę tą uzupełnijmy dowolnie do bazy Hamela $\{e_\alpha\}_{\alpha \in A}$ przestrzeni X i wybierzmy dwa różne wektory e_{α_1} i e_{α_2} tego uzupełnienia. Odwzorowanie liniowe $T : X \to X$ zamieniające miejscami wektory e_{α_1} i e_{α_2} jest wzajemnie jednoznaczne i odwzorowuje X na X. Nie jest odwzorowaniem ciągłym, bo ciągłe odwzorowanie liniowe przestrzeni X jest jednoznacznie wyznaczone przez swoje wartości na wektorach bazy topologicznej, a na nich T jest identycznością.

5.10. TWIERDZENIE. *Niech T będzie ciągłym odwzorowaniem liniowym przestrzeni Banacha X na przestrzeń Banacha Y i niech $X_0 = \{x \in X : Tx = 0\}$ będzie jego jądrem. Wtedy przestrzenie X/X_0 i Y są izomorficzne.* $\square$

5.11. TWIERDZENIE O DWU NORMACH. *Załóżmy, że w przestrzeni liniowej X dane są dwie normy zupełne $\| \ \|_1$ i $\| \ \|_2$ czyli, że przestrzenie $(X, \| \ \|_1)$ i $(X, \| \ \|_2)$ są zupełne. Jeżeli istnieje taka stała C_1, że*

$$\|x\|_1 \leqslant C_1 \|x\|_2, \qquad x \in X,$$

to także

$$|x\|_2 \leqslant C_2 \|x\|_1, \qquad x \in X,$$

dla pewnej stałej C_2

Dowód: Identyczność I odwzorowuje wzajemnie jednoznacznie przestrzeń Banacha $(X, \| \ \|_1)$ na przestrzeń Banacha $(X, \| \ \|_2)$. Ponieważ $\|x\|_1 \leqslant C_1 \|x\|_2$ dla wszystkich $x \in X$, to I jest odwzorowaniem ciągłym. Z twierdzenia Banacha o odwzorowaniu odwrotnym wynika, że także odwzorowanie odwrotne I^{-1} jest ciągłe. Zatem

$$\|x\|_2 = \|I^{-1}x\|_2 \leqslant \|I^{-1}\| \|x\|_1$$

dla wszystkich $x \in X$. $\square$

UWAGA. W przestrzeni X z przykładu 5.9 normy $\|x\|_1 = \|x\|$ i $\|x\|_2 = \|Tx\|$ są zupełne, ale nie są równoważne.

DEFINICJA. Mówimy, że dwie normy $\| \ \|_1$ i $\| \ \|_2$ w przestrzeni liniowej X są **zgodne**, jeżeli dla każdego ciągu $\{x_n\}$ w X warunki $\|x_n - x'\|_1 \to 0$ i $\|x_n - x''\|_2 \to 0$ pociągają $x' = x''$.

Jeżeli w przestrzeni liniowej X dane są trzy normy $\| \|_1$, $\| \|_2$ i $\| \|_3$ oraz

$$\|x\|_3 \leqslant C_1\|x\|_1, \quad \|x\|_3 \leqslant C_2\|x\|_2, \quad x \in X,$$

dla pewnych stałych C_1 i C_2, to wszystkie te normy są zgodne.

5.12. TWIERDZENIE. *Jeżeli normy zupełne $\| \|_1$ i $\| \|_2$ w przestrzeni liniowej są zgodne, to są równoważne.*

Dowód: Określmy jeszcze dodatkową, trzecią normę, kładąc $\|x\|_3 = \|x\|_1 + \|x\|_2$. Dzięki zgodności norm $\| \|_1$ i $\| \|_2$ norma ta jest także zupełna a ponieważ $\|x\|_1 \leqslant \|x\|_3$, więc z twierdzenia o dwu normach wnioskujemy, że $\|x\|_3 \leqslant C_1\|x\|_1$ dla pewnej stałej C_1, w szczególności $\|x\|_2 \leqslant C_1\|x\|_1$. Podobnie pokazujemy, że $\|x\|_1 \leqslant C_2\|x\|_2$. $\square$

DEFINICJA. **Wykresem** odwzorowania liniowego $T : X \to Y$ nazywamy podzbiór W_T produktu $X \times Y$

$$W_T = \{(x, Tx) : x \in X\}.$$

Odwzorowanie T nazywamy **domkniętym**, jeżeli jego wykres jest podzbiorem domkniętym przestrzeni $X \times Y$ z topologią produktową.

Zauważmy, że T jest odwzorowaniem domkniętym wtedy i tylko wtedy, gdy ze zbieżności ciągów $\{x_n\}$ w X i $\{Tx_n\}$ w Y wynika równość

$$T\big(\lim_{n\to\infty} x_n\big) = \lim_{n\to\infty} Tx_n.$$

Każde odwzorowanie ciągłe jest oczywiście domknięte.

5.13. TWIERDZENIE O WYKRESIE DOMKNIĘTYM. *Domknięte odwzorowanie liniowe z przestrzeni Banacha X do przestrzeni Banacha Y jest ciągłe.*

Dowód: Przestrzeń $X\times Y$ z normą $\|(x,y)\| = \|x\|+\|y\|$ jest przestrzenią Banacha a wykres W_T odwzorowania T jej domkniętą podprzestrzenią liniową, W_T jest zatem przestrzenią Banacha.

Odwzorowanie $S : (x, Tx) \to x$ przestrzeni W_T na X jest wzajemnie jednoznaczne, liniowe i ciągłe. Z twierdzenia Banacha o odwzorowaniu odwrotnym wynika istnienie takiej stałej C, że

$$\|(x, Tx)\| \leqslant C\|x\|, \qquad x \in X,$$

co daje

$$\|Tx\| \leqslant C\|x\|, \qquad x \in X. \quad \square$$

5.14. PRZYKŁAD. Na przestrzeni $\boldsymbol{s}_0$ tych ciągów $(x_1, x_2, x_3, \ldots)$ liczb zespolonych, dla których tylko skończenie wiele wyrazów może być różnych od zera, wprowadźmy normę $\|(x_1, x_2, x_3, \ldots)\| = \max_n |x_n|$. Wtedy odwzorowanie $T : \boldsymbol{s}_0 \to \boldsymbol{s}_0$ określone wzorem

$$T(x_1, x_2, x_3, \ldots) = (x_1, 2x_2, 3x_3, \ldots)$$

jest domknięte a nie jest ciągłe. Nie przeczy to twierdzeniu o wykresie domkniętym, bo przestrzeń $\boldsymbol{s}_0$ nie jest zupełna.

Zadania uzupełniające

5.15. Niech $\{a_n\}$ będzie ustalonym ciągiem liczbowym. Jeżeli szereg $\sum\limits_{n=1}^{\infty} a_n b_n$ jest zbieżny dla każdego ciągu ograniczonego $\{b_n\}$, to $\sum\limits_{n=1}^{\infty} |a_n| < \infty$.

5.16. Niech będą przestrzeniami Banacha oraz T ciągłym odwzorowaniem liniowym z X na Y. Dowieść, że dla pewnej stałej C i każdego $y \in Y$ można tak wybrać element $x \in X$, by $Tx = y$ i $\|x\| \leqslant C\,\|y\|$.

5.17. Niech $1 \leqslant p < q \leqslant \infty$. Czy na przestrzeni $L^p(\mathbb{R}) \cap L^q(\mathbb{R})$ normy $\| \ \|_p$ i $\| \ \|_q$ są zgodne?

5.18. Niech $\| \ \|_1$ i $\| \ \|_2$ oznaczają normy standartowe w przestrzeniach ℓ^1 i ℓ^2. Jeżeli T jest odwzorowaniem liniowym z przestrzeni Banacha X do ℓ^1 oraz $\|Tx\|_2 \leqslant \|x\|$ dla każdego $x \in X$, to T jest odwzorowaniem ciągłym, tj. $\|Tx\|_1 \leqslant C\,\|x\|$ dla $x \in X$.

5.19. Jeżeli X jest domkniętą podprzestrzenią liniową przestrzeni $C[0,1]$ złożoną z funkcji mających ciągłą pochodną, to $\dim X < \infty$.

ROZDZIAŁ VI

PRZESTRZENIE LINIOWE TOPOLOGICZNE

Topologie liniowe

Przestrzeń liniowa X z topologią Hausdorffa $\mathcal{T}$ nazywamy **przestrzenią liniową topologiczną** jeżeli działania dodawania wektorów i mnożenia wektora przez liczbę są ciągłe. Chodzi tu o ciągłość odwzorowań $(x,y)\to x+y$ z $X\times X$ do X oraz $(\lambda,x)\to\lambda x$ z $\mathbb{C}\times X$ do X, przy czym $X\times X$ i $\mathbb{C}\times X$ traktujemy jako przestrzenie z topologią produktową.

Niech x będzie punktem przestrzeni liniowej topologicznej X. Zbiór $U\subset X$ nazwiemy **otoczeniem punktu** x, gdy zawiera on x w swoim wnętrzu, a rodzinę $\mathcal{U}_x$ otoczeń punktu x nazwiemy **bazą otoczeń punktu** x, jeżeli w każdym otoczeniu punktu x leży pewien zbiór rodziny $\mathcal{U}_x$. Zauważmy, że odwzorowanie $y\to x+y$ jest homeomorfizmem X na X. Wynika z tego, że jeżeli $\mathcal{U}$ jest bazą otoczeń zera, to rodzina $\mathcal{U}+x=\{U+x : u\in\mathcal{U}\}$ jest bazą otoczeń punktu x.

6.1. ZADANIE. Rodzina $\mathcal{U}$ podzbiorów przestrzeni liniowej X jest bazą otoczeń zera topologii liniowej wtedy i tylko wtedy, gdy

a. $\bigcap_{U\in\mathcal{U}} U=\{0\}$,
a. $\forall_{U_1,U_2\in\mathcal{U}}\ \exists_{V\in\mathcal{U}}\ : V\subset U_1\cap U_2$,
c. $\forall_{U\in\mathcal{U}}\ \exists_{V\in\mathcal{U}}\ : V+V\subset U$,
d. $\forall_{U\in\mathcal{U}}\ \exists_{V\in\mathcal{U}}\ \forall_{|\lambda|\leqslant 1}\ : \lambda V\subset U$,
e. $\forall_{U\in\mathcal{U}}\ \forall_{x\in X}\ \exists_{\lambda>0}\ : x\in\lambda U$, tzn. że U jest zbiorem pochłaniającym.

6.2. PRZYKŁAD. Rozważmy przestrzeń liniową $\mathcal{L}(0,1)$ wszystkich zespolonych, mierzalnych w sensie Lebesgue'a funkcji na przedziale $(0,1)$. Dla $x\in\mathcal{L}(0,1)$ połóżmy

$$\|x\|=\int_0^1\frac{|x(t)|}{1+|x(t)|}\,dt.$$

Wtedy funkcja $d(x,y) = \|x-y\|$ jest metryką i wyznacza w $\mathcal{L}(0,1)$ topologię liniową. Bazą otoczeń zera tej topologii tworzą np. zbiory $\{U_\varepsilon\}_{\varepsilon>0}$, gdzie

$$U_\varepsilon = \{x \in \mathcal{L}(0,1) : \|x\| < \varepsilon\}.$$

Ciągłość dodawania wektorów wynika natychmiast z nierówności $\|x+y\| \leqslant \|x\| + \|y\|$. Aby dowieść ciągłości mnożenia przez liczbę zobaczmy, że zbieżność w metryce d jest równoważna zbieżności według miary. Rzeczywiście, jeżeli $\|x_n - x\| \to 0$ przy $n \to \infty$ i $E_\varepsilon(n)$, dla $\varepsilon > 0$ i $n = 1,2,\ldots$, oznacza zbiór

$$E_\varepsilon(n) = \{t \in (0,1) : |x_n(t) - x(t)| \geqslant \varepsilon\},$$

to

$$\|x_n - x\| \geqslant \int_{E_\varepsilon(n)} \frac{|x_n(t)-x(t)|}{1+|x_n(t)-x(t)|}\, dt \geqslant \int_{E_\varepsilon(n)} \frac{\varepsilon}{1+\varepsilon}\, dt = |E_\varepsilon(n)| \frac{\varepsilon}{1+\varepsilon},$$

więc $|E_\varepsilon(n)| \to 0$ przy $n \to \infty$.

Z drugiej strony, gdy $|E_\varepsilon(n)| < \varepsilon$ dla $n \geqslant n_0$, to

$$\|x_n - x\| = \left(\int_{E_\varepsilon(n)} + \int_{E'_\varepsilon(n)}\right) \frac{|x(t)|}{1+|x(t)|}\, dt \leqslant |E_\varepsilon(n)| + \varepsilon|E'_\varepsilon(n)| < 2\varepsilon.$$

Łatwo jest już teraz sprawdzić, że mnożenie przez liczbę jest ciągłe, tzn. że gdy $\lambda_n \to \lambda$ a ciąg x_n zbiega do x według miary, to $\lambda_n x_n$ zbiega według miary do λx.

Przestrzeń $\mathcal{L}(0,1)$ z topologią zbieżności według miary jest przestrzenią liniową topologiczną.

Przestrzeń $\mathcal{L}(0,1)$ jest ważnym obiektem w teorii funkcji rzeczywistych. Z punktu widzenia analizy funkcjonalnej niej jest jednak aż tak atrakcyjna. Pokażemy, że

Na $\mathcal{L}(0,1)$ nie ma ciągłych niezerowych funkcjonałów liniowych.

Przypuśćmy, że taki funkcjonał x^* istnieje. Istnieje zatem także taka funkcja $x \in \mathcal{L}(0,1)$, że $x^*(x) = 1$. Podzielmy przedział $[0,1]$ na n równych przedziałów i niech χ_k oznacza funkcję charakterystyczną k-tego z nich. Zauważmy, że co najmniej dla jednej z funkcji $nx\chi_k$ zachodzi nierówność $x^*(nx\chi_k) \geqslant 1$. W przeciwnym razie z równości $x = x\chi_1 + x\chi_2 + \ldots + x\chi_n$ wynikałoby, że $|x^*(x)| < 1$. Wybierzmy jedną z nich i oznaczmy ją przez x_n. Otrzymujemy w ten sposób ciąg funkcji w $\mathcal{L}(0,1)$ o tej własności, że $x^*(x_n) \geqslant 1$ oraz $\big|\{t \in [0,1] : x_n(t) \neq 0\}\big| \leqslant 1/n$. Ta ostatnia nierówność oznacza, że x_n dąży do zera według miary. Powinno stąd wynikać, że $x^*(x_n) \to 0$, a tak nie jest.

6.3. ZADANIE. Sprawdzić, że na zbiorze $L^p(0,1)$, $0 < p < 1$, funkcji całkowalnych w sensie Lebesgue'a z p-tą potęgą na $[0,1]$, funkcja $d(x,y) = \|x-y\|$, gdzie

$$\|x\| = \int_0^1 |x(t)|^p dt,$$

jest metryką. W metryce tej $L^p(0,1)$ staje się przestrzenią liniową topologiczną zupełną. Dowieść, że jedynym zbiorem wypukłym w $L^p(0,1)$ o niepustym wnętrzu jest cała przestrzeń.

Aby nie mieć tego typu patologicznych przykładów najczęściej żąda się by topologia liniowa w przestrzeni X była **lokalnie wypukła** tj., by istniała baza otoczeń zera złożona ze zbiorów wypukłych.

6.4. LEMAT. *Przestrzeń liniowa lokalnie wypukła posiada bazę otoczeń zera złożoną ze zbiorów absolutnie wypukłych i pochłaniających.*

Dowód: Pokażemy, że jeżeli U jest otoczeniem zera, to zbiór

$$\bigcap_{|\lambda|\geqslant 1} \lambda U$$

jest także otoczeniem zera. Ponieważ operacja mnożenia przez skalary

$$\mathbb{C} \times X \ni (\lambda, x) \longrightarrow \lambda x \in X$$

jest ciągła, więc istnieje takie otoczenie zera V i takie $\varepsilon > 0$, że $\mu V \subset U$, gdy $|\mu| \leqslant \varepsilon$. Wobec tego $\varepsilon V \subset \lambda U$ dla każdego $\lambda \in \mathbb{C}$, $|\lambda| \geqslant 1$. Jeżeli U jest wypukłym otoczeniem zera, to zbiór $\bigcap_{|\lambda|\geqslant 1} \lambda U$ jest absolutnie wypukłym otoczeniem zera. Jest też zbiorem pochłaniającym, bo takim jest każde otoczenie zera. $\square$

Wiemy, że funkcjonał Minkowskiego p_W zbioru absolutnie wypukłego i pochłaniającego W jest półnormą, stąd otrzymujemy:

6.5. TWIERDZENIE. *Przestrzeń liniowa topologiczna jest lokalnie wypukła wtedy i tylko wtedy, gdy jej topologia jest wyznaczona przez rodzinę półnorm $\{p_\alpha\}_{\alpha\in A}$ o tej własności, że $p_\alpha(x) = 0$ dla wszystkich $\alpha \in A$ pociąga $x = 0$.* $\square$

Ważną klasę przestrzeni liniowych topologicznych to przestrzenie lokalnie wypukłe metryzowalne. Przestrzeń taka ma przeliczalną bazę otoczeń zera, a więc jej topologia jest wyznaczona przez przeliczalną rodzinę półnorm $\{p_n\}_{n=1}^{\infty}$. Możemy przy tym założyć, że półnormy spełniają nierówności $p_1 \leqslant p_2 \leqslant p_3 \leqslant \dots$ przyjmując np.

$$q_n = \sum_{k=1}^{n} p_k$$

zamiast p_n. Za metrykę zaś można przyjąć

$$d(x,y) = \sum_{n=0}^{\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)}. \tag{6.30}$$

Jeśli przestrzeń X jest zupełna w tej metryce, to nosi nazwę **przestrzeni typu $\boldsymbol{B}_0$**

6.6. TWIERDZENIE. *Załóżmy, że X jest przestrzenią lokalnie wypukłą z topologią wyznaczoną przez rodziną półnorm $p_1 \leqslant p_2 \leqslant p_3 \leqslant \dots$. Jeżeli $x^* \in X^*$, to istnieje taki wskaźnik n i stała C, że*

$$\big|x^*(x)\big| \leqslant C\, p_n(x), \quad x \in X. \tag{6.31}$$

Wynika stąd, że x^* jest ciągłym funkcjonałem liniowym na przestrzeni unormowanej $\big(X/\ker p_n,\, p_n\big)$. Każdy funkcjonał liniowy tej postaci jest ciągły na X.

Dowód: Jeśli warunek (6.31) nie jest spełniony, to w X istnieje taki ciąg $\{x_n\}$, że

$$1 = \big|x^*(x_n)\big| \geqslant n\, p_n(x_n).$$

Ponieważ $p_k(x_n) \leqslant p_n(x_n) = 1/n$ dla $n \geqslant k$, więc $x_n \to 0$ w X, zaś $x^*(x_n)$ nie dąży do zera, więc sprzeczność. Reszta jest oczywista. □

6.7. WNIOSEK. *Dla każdego wektora $x_0 \neq 0$ przestrzeni liniowej metrycznej lokalnie wypukłej istnieje taki ciągły funkcjonał liniowy x^*, że $x^*(x_0) \neq 0$.* □

6.8. PRZYKŁAD. W przestrzeni $C^{\infty}[0,1]$ funkcji zespolonych dowolną liczbę razy różniczkowalnych na $[0,1]$ można określić topologię lokalnie wypukłą wybierając rodzinę półnorm

$$p_n(x) = \max_{t\in[0,1]} \big|x^{(n)}(t)\big|, \quad n = 0,1,2,\dots\ .$$

Widać, że w metryce (6.30) przestrzeń $C^{\infty}[0,1]$ jest zupełna, jest zatem przestrzenią typu B_0. Zauważmy też, że wielomiany tworzą w niej podzbiór gęsty.

Posłużymy się twierdzeniem 6.6 dla wyznaczenia wszystkich ciągłych funkcjonałów liniowych $x^* \in \big(C^\infty[0,1]\big)^*$. Zauważmy w tym celu, że topologia przestrzeni $C^\infty[0,1]$ nie zmieni się, gdy definiujące ją półnormy p_n zastąpimy przez

$$p_n(x) = \sum_{k=0}^{n} \max_{t\in[0,1]} \big|x^{(k)}(t)\big|.$$

Wtedy $p_0 \leqslant p_1 \leqslant p_2 \leqslant \ldots$ a każda z nowych półnorm p_n staje się normą. Uzupełnienie przestrzeni $C^\infty[0,1]$ w normie p_n daje przestrzeń $C^n[0,1]$, funkcji mających ciągłe pochodne do rzędu n włącznie. Z twierdzenia 6.6 wynika zatem, że x^* jest ciągłym funkcjonałem liniowym na przestrzeni $C^\infty[0,1]$ wtedy i tylko wtedy, gdy dla pewnego n jest ciągłym funkcjonałem liniowym na $C^n[0,1]$. Ponieważ przestrzeń $C^n[0,1]$ jest izomorficzna z produktem $\mathbb{C}^n \times C[0,1]$ (patrz przykład 1.32), więc x^* musi być postaci

$$x^*(x) = \sum_{k=0}^{n-1} \lambda_k x^{(k)}(0) + \int_0^1 x^{(n)}(t)\, d\mu(t),$$

dla pewnego wektora $(\lambda_0, \lambda_1, \ldots, \lambda_{n-1}) \in \mathbb{C}^n$ oraz miary zespolonej $\mu \in M[0,1]$.

6.9. Przykład. Przestrzeń Schwartza $\mathcal{S}(\mathbb{R})$ składa się z funkcji zespolonych na prostej $\mathbb{R}$ dowolnie wiele razy różniczkowalnych i szybko dążących do zera w nieskończoności, tj. takich funkcji x, że

$$\sup_{t\in\mathbb{R}} \big|t^k x^{(n)}(t)\big| < \infty$$

dla wszystkich $k, n = 0, 1, 2, \ldots$. Jest to przestrzeń typu $\boldsymbol{B}_0$.

Funkcjonały $x^* \in \mathcal{S}^*(\mathbb{R})$ noszą nazwę **dystrybucji temperowanych**. Przestrzeń Schwartza zajmuje ważne miejsce w analizie z tego względu, że transformacja Fouriera przeprowadza ją z powrotem na siebie (patrz **??**). Pozwala to określić transformatę Fouriera dystrybucji temperowanej wzorem

$$\widehat{x^*}(x) = x^*(\widehat{x}), \quad x \in \mathcal{S}(\mathbb{R}).$$

Dla przykładu, dla dystrybucji $x^*(x) = x''(0)$ otrzymujemy

$$\widehat{x^*}(x) = \widehat{x}''(0) = \int_{\mathbb{R}} 4\pi^2 t^2\, x(t)dt.$$

6.10. Zadanie. Niech X oznacza przestrzeń liniową wszystkich funkcji ciągłych na przedziale $[0,1]$ z topologią lokalnie wypukłą wyznaczoną przez rodzinę półnorm $\{p_t\}_{t\in[0,1]}$ postaci

$$p_t(x) = |x(t)|.$$

Czy topologia ta jest metryzowalna? Opisać przestrzeń X^*.

Słabe topologie w przestrzeniach Banacha

Niech X będzie przestrzenią Banacha oraz X^* jej przestrzenią sprzężoną. Symbolem $\mathcal{T}_\omega$ oznaczymy najsłabszą topologię w X, w której każdy funkcjonał $x^* \in X^*$ jest ciągły. Topologię $\mathcal{T}_\omega$ nazywamy **słabą topologią** w X. Jest to topologia wyznaczona przez rodzinę półnorm p_{x^*}, $x^* \in X^*$, postaci

$$p_{x^*}(x) = \left|x^*(x)\right|.$$

Przestrzeń $(X, \mathcal{T}_\omega)$ jest zatem lokalnie wypukła.

6.11. TWIERDZENIE. *Topologia słaba i normowa w przestrzeni Banacha pokrywają się wtedy i tylko wtedy, gdy jest to przestrzeń skończenie wymiarowa.*

Dowód: Załóżmy, że przestrzeń X ma wymiar skończony i niech $x_1^*, x_2^*, \ldots, x_n^*$ będzie bazą Hamela przestrzeni X^*. Wtedy funkcjonał

$$\|x\|_1 = \max_{1 \leqslant k \leqslant n} \left|x_k^*(x)\right|$$

jest normą w X wyznaczającą słabą topologię. Jest to norma równoważna z normą wyjściową.

Jeżeli X jest przestrzenią nieskończenie wymiarową, to każde jej otoczenie zera U topologii słabej jest zbiorem nieograniczonym. Rzeczywiście, możemy założyć, że U jest zbiorem bazowym tzn., że dla pewnej liczby $\varepsilon > 0$ i zbioru skończonego $x_1^*, x_2^*, \ldots, x_n^*$ w X^*

$$U = \left\{x \in X : \left|x_k^*(x)\right| < \varepsilon \quad \text{dla} \quad k = 1, 2, \ldots, n\right\}.$$

Zauważmy, że U zawiera niezerową podprzestrzeń liniową

$$\bigcap_{k=1}^{n} \ker x_k^*,$$

nie może być zatem zbiorem ograniczonym. W przestrzeni unormowanej istnieją ograniczone otoczenia zera, np. kule. $\square$

6.12. TWIERDZENIE. *Słaba topologia w nieskończenie wymiarowej przestrzeni Banacha nie jest metryzowalna.*

Dowód: Załóżmy, że taka metryka istnieje. Rozumując jak w dowodzie poprzedniego twierdzenia stwierdzamy, że każda z kul o środku w zerze jest zbiorem nieograniczonym. Możemy zatem znaleźć ciąg zbieżny do zera w metryce, ale nieograniczony w normie. Jak wiemy z twierdzenia Banacha-Steinhausa (patrz także 5.1), taki ciąg nie może być słabo zbieżny. Tu sprzeczność. $\square$

6.13. Przykład. Niech $e_1, e_2, e_3, \dots$ będzie układem ortonormalnym w nieskończenie wymiarowej przestrzeni Hilberta $\mathcal{H}$ oraz A podzbiorem $\mathcal{H}$, złożonym z elementów postaci $\sqrt{n}\, e_n$. Pokażemy, że zero leży w słabym domknięciu zbioru A ale, jak widać, nie jest granicą żadnego słabo zbieżnego ciągu elementów zbioru A.

Niech
$$U = \{x \in \mathcal{H} : \big|\langle x,\, y_i\rangle\big| < \varepsilon,\ i = 1, 2, \dots, k\}$$
będzie bazowym otoczeniem zera w słabej topologii $\mathcal{H}$. Ponieważ
$$\sum_{n=1}^{\infty} \big|\langle e_n,\, y_i\rangle\big|^2 < \infty$$
dla każdego i, więc
$$\sum_{n=1}^{\infty} \Big(\sum_{i=1}^{k} \big|\langle e_n,\, y_i\rangle\big|\Big)^2 < \infty.$$
Wynika stąd, że dla pewnego wskaźnika n_0 zachodzi
$$\sum_{i=1}^{k} \big|\langle e_{n_0},\, y_i\rangle\big| < \frac{\varepsilon}{\sqrt{n_0}},$$
więc $\sqrt{n_0}\, e_{n_0} \in U$.

6.14. Zadanie. Dowieść, że jeżeli punkt x_0 leży w słabym domknięciu zbioru ograniczonego A w przestrzeni ℓ^2, to jest on słabą granicą pewnego ciągu elementów zbioru A.

Z twierdzenia **??** o oddzielaniu zbiorów wypukłych wynika, że każdy domknięty podzbiór wypukły przestrzeni Banacha jest słabo domknięty. Takie stwierdzenie zdecydowanie nie jest prawdziwe dla dowolnych zbiorów domkniętych.

6.15. Zadanie. Pokazać, że w nieskończenie wymiarowej przestrzeni Banacha X słabe domknięcie sfery $\{x \in X : \|x\|_2 = 1\}$ jest kulą $\{x \in X : \|x\|_2 \leqslant 1\}$.

W przestrzeni X^*, sprzężonej do pewnej przestrzeni Banacha X, poza topologią normową i topologią słabą, ważną rolę odgrywa **topologia *-słaba**, oznaczana $\mathcal{T}_{\omega^*}$ lub $\sigma(X^*, X)$. Jest to z definicji najsłabsza topologia, w której wszystkie funkcjonały $\tau(x) \in X^{**}$, $x \in X$, postaci
$$\tau(x)(x^*) = x^*(x)$$
są ciągłe.

Topologia $\mathcal{T}_{\omega^*}$ jest lokalnie wypukła i może być wyznaczona przez rodzinę półnorm $\{p_x\}_{x\in X}$ postaci
$$p_x(x^*) = \big|x^*(x)\big|.$$

Uwaga. Przestrzenie $\mathbf{c}^*$ i $\mathbf{c}_0^*$ są identyczne z przestrzenią ℓ^1, a więc ℓ^1 ma dwie słabe topologie $\mathcal{T}$ i $\mathcal{T}_0$. Topologie te są nieporównywalne, tzn. nie zachodzi żadna z inkluzji $\mathcal{T} \subset \mathcal{T}_0$, $\mathcal{T}_0 \subset \mathcal{T}$. Istotnie, niech φ i φ_0 oznaczają odpowiednio izometryczny izomorfizm przestrzeni ℓ^1 na $\mathbf{c}^*$ i $\mathbf{c}_0^*$

$$\varphi(y)(x) = y_1 \lim_{k\to\infty} x_k + \sum_{k=1}^{\infty} x_k\, y_{k+1}, \qquad x \in \mathbf{c},$$
$$\varphi_0(y)(x) = \sum_{k=1}^{\infty} x_k\, y_k, \qquad x \in \mathbf{c}_0.$$

Zbiór $U_1 = \big\{y \in \ell^1 : \big|\sum y_k\big| < 1\big\}$ jest otoczeniem zera w topologii $\mathcal{T}$, a nie jest otoczeniem zera w topologii $\mathcal{T}_0$, zaś dla zbioru $U_2 = \big\{y \in \ell^2 : |y_1| < 1\big\}$ jest odwrotnie.

6.16. Twierdzenie Banacha-Alaoglu. *Niech X będzie przestrzenią Banacha. Zbiór $A \subset X^*$ jest *-słabo zwarty wtedy i tylko wtedy, gdy jest ograniczony (w normie) i *-słabo domknięty.*

Dowód: Załóżmy, że zbiór A jest *-słabo zwarty w X^*. Ponieważ $\mathcal{T}_{\omega^*}$ jest topologią Hausdorffa w X^*, otrzymujemy natychmiast, że A jest *-słabo domknięty.

Nim przystąpimy do dalszej części dowodu, przedstawimy inny, wygodniejszy model przestrzeni $\big(X^*, \mathcal{T}_{\omega^*}\big)$.

Dla każdego $x \in X$ połóżmy $\mathbb{C}_x = \mathbb{C}$ (ciało liczb zespolonych) i określmy odwzorowanie $T : x^* \to \mathbb{C}^X = \prod_{x\in X} \mathbb{C}_x$ wzorem

$$Tx^* = \big\{x^*(x)\big\}_{x\in X},$$

tj. Tx^* jest elementem produktu $\mathbb{C}^X$, dla którego x-tą współrzędną jest $x^*(x)$. Jeżeli $\mathbb{C}^X$ traktować jako przestrzeń topologiczną z topologią produktową, a zbiór $T(X^*)$ jako podprzestrzeń z topologią indukowaną z $\mathbb{C}^X$, to porównując otwarte zbiory bazowe w $\big(X^*, \mathcal{T}_{\omega^*}\big)$ oraz w $T(X^*)$ dochodzimy do wniosku, że przestrzenie te są homeomorficzne.

Dla każdego $y \in X$ niech $P_y : \mathbb{C}^X \to \mathbb{C}_y$ oznacz rzut z produktu na oś $\mathbb{C}_y$. Ponieważ każde odwzorowanie P_y jest ciągłe, a T jest homeomorfizmem, to odwzorowanie $P_yT : X^* \to \mathbb{C}_y$ jest ciągłe. W szczególności $P_yT(A) = \{x^*(y) : x^* \in A\}$ jest zbiorem zwartym w $\mathbb{C}_y = \mathbb{C}$ dla każdego $y \in X$. Istnieją więc takie stałe $M_y > 0$, że

$$\sup_{x^*\in A} \big|x^*(y)\big| \leqslant M_y.$$

Korzystając teraz z twierdzenia Banacha-Steinhausa dochodzimy do wniosku, że

$$\sup_{x^*\in A} \|x^*\| < \infty.$$

W dowodzie w drugą stronę możemy założyć, bez ograniczania ogólności, że A jest domkniętą kulą jednostkową $K^* = \{x^* \in X^* : \|x^*\| \leqslant 1\}$ przestrzeni X^*. Oczywiście K^* jest zbiorem ograniczonym w X^*, łatwo też można pokazać, że K^* jest zbiorem *-słabo domkniętym. Jeżeli wiemy już, że K^* jest zbiorem *-słabo zwartym, to ponieważ A jest zbiorem ograniczonym, to istnieje taka stała $c > 0$, że $A \subset cK^*$. Zbiór cK^* jest *-słabo zwarty, bo mnożenie przez skalary jest homeomorfizmem w $(X^*, \mathcal{T}_{\omega^*})$, a więc A jest *-słabo zwarty, jako *-słabo domknięty podzbiór w cK^*.

Niech więc $A = K^*$. Dla każdego $x \in X$ niech

$$K_x = \{z \in \mathbb{C} : |z| \leqslant \|x\|\}.$$

Oczywiście K_x jest zwartym podzbiorem w $\mathbb{C}$, ponadto

$$T(K^*) \subset \prod_{x\in X} K_x.$$

Istotnie, jeżeli $x^* \in X^*$, to $T(x^*) = \{x^*(x)\}_{x\in X}$ oraz $|x^*(x)| \leqslant \|x^*\|\,\|x\|$. Twierdzimy, że $T(K^*)$ jest podzbiorem zwartym w $T(X^*)$. Wystarczy w tym celu pokazać, że jest domkniętym podzbiorem $\prod_{x\in X} K_x$, gdyż ten ostatni zbiór z twierdzenia Tichonowa **??** jest zwarty w $\mathbb{C}^X$. Załóżmy, że punkt $\{s_x\}_{x\in X}$ należy do domknięcia zbioru $T(K^*)$ w $\mathbb{C}^X$. Ustalmy dowolne $x_1, x_2 \in X$ i $\varepsilon > 0$. Niech

$$U_\varepsilon = \left\{\{u_x\}_{x\in X} \in \mathbb{C}^X : \left|u_{x_1} - s_{x_1}\right| < \frac{\varepsilon}{3},\ \left|u_{x_2} - s_{x_2}\right| < \frac{\varepsilon}{3},\ \left|u_{x_1+x_2} - s_{x_1+x_2}\right| < \frac{\varepsilon}{3}\right\}$$

będzie bazowym otoczeniem punktu $\{s_x\}_{x\in X}$ w $\mathbb{C}^X$. Z założenia istnieje taki element $x^* \in X^*$, że $Tx^* \in U_\varepsilon$. Stąd

$$\left|s_{x_1+x_2} - s_{x_1} - s_{x_2}\right| \leqslant \left|s_{x_1+x_2} - x^*(x_1+x_2)\right| + \left|s_{x_1} - x^*(x_1)\right| + \left|s_{x_2} - x^*(x_2)\right| < \varepsilon,$$

a ponieważ $\varepsilon > 0$ jest dowolne, więc w rezultacie

$$s_{x_1+x_2} = s_{x_1} + s_{x_2}.$$

Podobnie pokazujemy, że

$$s_{\lambda x} = \lambda\, s_x, \quad \text{dla} \quad x \in X,\ \lambda \in \mathbb{C}.$$

Wynika stąd, że przyporządkowanie $x \to s_x$ jest funkcjonałem liniowym na X, a ponieważ $|s_x| \leqslant \|x\|$, więc $\{s_x\}_{x \in X} \in T(K^*)$.

Reasumując, $T(K^*)$ jest domkniętym podzbiorem w $\prod_{x \in X} K_x$, ten zaś z kolei zwartym podzbiorem w $\mathbb{C}^X$; odwzorowanie T jest homeomorfizmem, zatem A jest $*$-słabo zwarty w X^*. $\square$

Jako wniosek z Twierdzenia Banacha-Alaoglu otrzymujemy poniższe twierdzenie o uniwersalności przestrzeni $C(S)$ dla wszystkich przestrzeni Banacha.

6.17. TWIERDZENIE. *Dla każdej przestrzeni Banacha X istnieje zwarta przestrzeń Hausdorffa S o tej własności, że X jest izometrycznie izomorficzna z domkniętą podprzestrzenią liniową przestrzeni $C(S)$.*

Dowód: Za przestrzeń S wybierzmy kulę jednostkową K^* przestrzeni sprzężonej X^* i zaopatrzmy ją w topologię $*$-słabą. Z twierdzenia Banacha-Alaoglu wynika, że S jest zwartą przestrzenią Hausdorffa. Jeśli teraz elementowi $x \in X$ przyporządkujemy funkcję $f_x \in C(S)$ wzorem

$$f_x(x^*) = x^*(x),$$

to $\|x\| = \|f_x\|_\infty$ na mocy wniosku 4.13). $\square$

Zadania uzupełniające

6.18. ZADANIE. W przestrzeni ℓ^p, $0 < p < 1$, ciągów sumowalnych z p-tą potęgą, analogicznie jak w $L^p(0,1)$, topologia liniowa jest wyznaczona przez metrykę $d(x,y) = \|x - y\|$, gdzie

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

Opisać przestrzeń $(\ell^p)^*$.

6.19. ZADANIE. Niech X będzie przestrzenią liniową wszystkich funkcji ciągłych na przedziale $(0,1)$. Dla $x \in X$ i $\varepsilon > 0$ oznaczmy przez $U_{x,\varepsilon}$ zbiór tych $y \in X$, że $|x(t) - y(t)| < \varepsilon$ dla wszystkich $t \in (0,1)$. Niech $\mathcal{T}$ będzie topologią w X, wyznaczoną przez wszystkie takie zbiory $U_{x,\varepsilon}$. Pokazać, że w topologii $\mathcal{T}$ dodawanie funkcji jest operacją ciągłą, a mnożenie funkcji przez liczby nie.

ROZDZIAŁ VII

OPERATORY LINIOWE

Przestrzeń operatorów linowych

Norma operatora

Operator liniowy T z przestrzeni liniowej unormowanej X do przestrzeni liniowej unormowaną Y nazywamy **ograniczonym**, jeżeli istnieje taka stała M, że

$$\|Tx\| \leqslant M\,\|x\|$$

dla wszystkich $x \in X$. Kres dolny takich stałych oznaczamy $\|T\|$ i nazywamy **normą operatora** T.

7.1. Twierdzenie. *Dla każdego operatora ograniczonego $T : X \to Y$ zachodzi nierówność $\|Tx\| \leqslant \|T\|\,\|x\|$, $x \in X$, a także*

$$\|T\| = \sup_{\|x\| \leqslant 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|.$$

Dowód: Oznaczmy $M = \sup_{\|x\|=1} \|Tx\|$ i dla $x \in X$, $x \neq 0$, połóżmy $y = \frac{1}{\|x\|}x$. Wtedy $\|y\| = 1$, więc

$$\|Tx\| = \|Ty\|\,\|x\| \leqslant M\,\|x\|$$

co daje $\|T\| \leqslant M$. Z drugiej strony, gdy $\|x\| \leqslant 1$, to $\|Tx\| \leqslant \|T\|$, więc

$$\sup_{\|x\| \leqslant 1} \|Tx\| \leqslant \|T\|. \quad \square$$

7.2. Zadanie. Operator liniowy jest ograniczony wtedy i tylko wtedy, gdy jest ciągły. Dla ograniczoności operatora wystarcza jego ciągłość w zerze.

Uwaga. Niech $X = Y = \mathbb{C}^n$. Każdemu operatorowi $A : X \to Y$ odpowiada w sposób wzajemnie jednoznaczny macierz $\{a_{i,j}\}_{i,j=1}^n$ wzorem

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Bezpośrednie obliczanie normy $\|A\|$ zwykle jest bardzo skomplikowane. Pokażemy później, że można to zrobić w następujący sposób:

1. znaleźć macierz A^*A, gdzie $A^* = \overline{A}^\top$,

2. znaleźć wszystkie pierwiastki wielomianu charakterystycznego

$$\det\big(zI - A^*A\big)$$

(okazuje się, że są to liczby rzeczywiste nieujemne).

3. obliczyć

$$\|A\| = \max_{1\leqslant k\leqslant n} |z_k|^{1/2}.$$

7.3. Przykład. Porównajmy bezpośrednią i opisaną wyżej metodę obliczania normy $\|A\|$ na macierzy $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$. Mamy tu

$$\|A\| = \sup_{|x|^2+|y|^2=1} \big(|x-y|^2 + |x|^2\big)^{1/2} = \sup_{|x|^2+|y|^2=1} \big(2|x|^2 + |y|^2 - 2\,\mathrm{Re}\, x\overline{y}\big)^{1/2}.$$

Ponieważ $|x|^2 + |y|^2 = 1$, więc $|x| = \cos t$, $|y| = \sin t$ dla pewnego $t \in [0, \frac{\pi}{2})$. Oznaczając $s = \arg x$, $u = \arg y$, otrzymamy

$$\begin{aligned} \|A\| &= \sup_{\substack{t\in[0,\frac{\pi}{2})\\ s,u\in\mathbb{R}}} \big(2\cos^2 t + \sin^2 t - 2\sin t\cos t\cos(s-u)\big)^{1/2} \\ &= \sup_{t\in[0,\frac{\pi}{2})} (2\cos^2 t + \sin^2 t + 2\sin t\cos t)^{1/2} \\ &= \sup_{t\in[0,\frac{\pi}{2})} \big(\tfrac{3}{2} + \tfrac{1}{2}\cos 2t + \sin 2t\big)^{1/2}. \end{aligned}$$

Supremum to jest osiągnięte dla $t = \frac{1}{2}\operatorname{arc\,tg} 2$ i wynosi $\sqrt{\frac{3+\sqrt{5}}{2}}$, zatem

$$\|A\| = \sqrt{\frac{3+\sqrt{5}}{2}}.$$

Z drugiej strony $A^* = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $A^*A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$, a ponieważ wielomian charakterystyczny $\det(zI - A^*A) = z^2 - 3z + 1$ ma pierwiastki $z_1 = \frac{3+\sqrt{5}}{2}$, $z_2 = \frac{3-\sqrt{5}}{2}$, więc

$$\|A\| = \max\left\{\sqrt{\tfrac{3+\sqrt{5}}{2}}, \sqrt{\tfrac{3-\sqrt{5}}{2}}\right\} = \sqrt{\tfrac{3+\sqrt{5}}{2}}.$$

Zwróćmy jeszcze uwagę, że do obliczenia $\|A\|$ nie wystarczy znajomość pierwiastków wielomianu charakterystycznego samej macierzy A. Dla rozpatrywanej macierzy wielomian charakterystyczny

$$\det(zI - A) = z^2 - z + 1$$

ma pierwiastki $\frac{1+i\sqrt{3}}{2}$ i $\frac{1-i\sqrt{3}}{2}$, oba o module 1.

7.4. Zadanie. Przeprowadzić analogiczne porównanie obu metod obliczania normy $\|A\|$ na macierzy

$$A = \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ i & 0 & 1 \end{pmatrix}.$$

W praktyce, przy badaniu własności operatorów, znajomość ich normy nie jest aż tak istotna, zadowala nas znajomość oszacowań od góry. Zdarza się jednak, że rozstrzygnięcie, czy badany operator jest ograniczony, a więc znalezienie jakiegokolwiek oszacowania od góry, może być trudne. Zilustrujemy to na przykładzie transformaty Laplace'a.

7.5. Przykład. **Transformacja Laplace'a** to operacja przyporządkowująca funkcji x na półprostej $(0,\infty)$ funkcję Lx postaci

$$Lx(t) = \int_0^\infty e^{-st}\, x(s)\, ds, \qquad t \in (0,\infty). \tag{7.32}$$

Jeżeli $x \in L^2(0,\infty)$, to funkcja Lx jest dobrze określona. Wynika to z nierówności Schwarza

$$\begin{aligned} \big|Lx(t)\big| &\leqslant \int_0^\infty e^{-st}\, |x(s)|\, ds \leqslant \left(\int_0^\infty e^{-2st}\, ds\right)^{1/2} \|x\|_2 \\ &= \frac{1}{\sqrt{2t}}\, \|x\|_2. \end{aligned} \tag{7.33}$$

Powstaje jednak pytanie, czy $Lx \in L^2(0,\infty)$? W każdym razie z oszacowania (7.33) tego nie widać, bo funkcja $\dfrac{1}{\sqrt{2t}}$ nie leży $L^2(0,\infty)$, nie leży nawet w żadnej z przestrzeni $L^p(0,\infty)$, $1 \leqslant p \leqslant \infty$. Sprawdźmy wobec tego na przykładach czy można się spodziewać odpowiedzi pozytywnej na postawione pytanie.

Dla $x = \chi_{(0,1)}$ otrzymujemy

$$Lx(t) = \frac{1-e^{-t}}{t}.$$

Jest to funkcja z $L^2(0,\infty)$, bo $Lx(t) \leqslant 1$ dla $t \in (0,1]$ i $Lx(t) \leqslant t^{-1}$ dla $t \in [1,\infty)$.

Niech teraz $x(t) = \dfrac{1}{1+t}$. Oczywiście $x \in L^2(0,\infty)$ oraz $\|x\|_2 = 1$, zaś

$$Lx(t) = \int_0^\infty e^{-st}\,\frac{ds}{1+s}.$$

Jeśli powyższą całkę rozdzielimy na dwie $\int_0^{1/t} + \int_{1/t}^\infty$ i w pierwszej z nich składnik e^{-st} zastąpimy przez 1 a w drugiej składnik $\frac{1}{1+s}$ przez $\frac{1}{1+1/t}$, to po obliczeniu otrzymanych całek dostaniemy

$$Lx(t) \leqslant \ln\Big(1+\frac{1}{t}\Big) + \frac{1}{1+t}.$$

Prawa strona tej nierówności przedstawia funkcję z $L^2(0,\infty)$, więc tym bardziej $Lx \in L^2(0,\infty)$.

Proponuję trochę inną metodę dowodu tego faktu. Wybierzmy dowolnie liczbę $a \in (-1,1)$ i po przedstawieniu funkcji $e^{-st}\frac{1}{1+s}$ w postaci iloczynu $\big(s^{-a/2}e^{-st}\big)\cdot\big(s^{a/2}\frac{1}{1+s}\big)$ zastosujmy nierówność Schwarza. Wtedy otrzymamy

$$\big|Lx(t)\big|^2 \leqslant C_a\, t^{a-1},$$

gdzie

$$\begin{aligned} C_a &= 2^{a-1}\int_0^\infty u^{-a}e^{-u}\,du \int_0^\infty \frac{s^a}{(1+s)^2}\,ds \\ &= 2^{a-1}\Gamma(1-a)^2\Gamma(1+a). \end{aligned}$$

Wybór $a < 0$ pociąga całkowalność funkcji $|Lx|^2$ na $(1,\infty)$ a wybór $a > 0$ całkowalność na $(0,1)$.

Jeszcze lepszy efekt otrzymamy wybierając przedstawienie postaci $e^{-st}\frac{1}{1+s} = \left(s^{-a/2}e^{-st/2}\right)\left(s^{a/2}e^{-st/2}\,\frac{1}{1+s}\right)$.

$$\begin{aligned}\|Lx\|_2^2 &\leqslant \int_0^\infty \left(\int_0^\infty s^{-a}e^{-st}\,ds\right)\left(\int_0^\infty s^a e^{-st}\frac{1}{(1+s)^2}\,ds\right)dt\\ &= \int_0^\infty \int_0^\infty \Gamma(1-a)\,t^{a-1}s^a e^{-st}\frac{1}{(1+s)^2}\,ds\,dt\\ &= \Gamma(1-a)\int_0^\infty s^a\frac{1}{(1+s)^2}\int_0^\infty t^{a-1}e^{-st}\,dt\,ds\\ &= \Gamma(a)\Gamma(1-a)\int_0^\infty \frac{1}{(1+s)^2}\,ds = \frac{\pi}{\sin a\pi}\,\|x\|_2^2.\end{aligned}$$

Trzeba jednak założyć, że $0 < a < 1$, aby wszystkie całki były zbieżne.

Zwróćmy uwagę, że w powyższym rozumowaniu nigdzie nie korzystaliśmy z postaci funkcji x. Można zatem je powtórzyć dla dowolnej funkcji z $L^2(0,\infty)$. To dowodzi ograniczoności operatora L. Zwróćmy też uwagę, że spośród różnych stałych a najlepszą jest $a = \frac{1}{2}$, dla niej $\|Lx\|_2 \leqslant \sqrt{\pi}\,\|x\|_2$.

7.6. Twierdzenie. *Transformacja Laplace'a*

$$Lx(t) = \int_0^\infty e^{-st}\,x(s)\,ds$$

jest ograniczoną operacją liniową z $L^2(0,\infty)$ do $L^2(0,\infty)$ i ma normę $\sqrt{\pi}$.

Dowód: Niech x będzie dowolną funkcją ciągłą o nośniku zwartym na półprostej $(0,\infty)$. Wtedy

$$\begin{aligned}\|Lx\|_2^2 &= \int_0^\infty \left|\int_0^\infty \left(e^{-st/2}s^{-1/4}\right)\cdot\left(e^{-ts/2}s^{1/4}x(s)\right)ds\right|^2 dt\\ &\int_0^\infty \int_0^\infty e^{-st}s^{-1/2}ds\cdot\int_0^\infty e^{-ts}s^{1/2}|x(s)|^2 ds\,dt\\ &= \int_0^\infty \int_0^\infty \Gamma(1/2)t^{-1/2}e^{-st}s^{1/2}|x(s)|^2 ds\,dt\\ &= \Gamma(1/2)\int_0^\infty s^{1/2}|x(s)|^2\int_0^\infty t^{-1/2}e^{-st}dt\,ds\\ &= \Gamma(1/2)^2\int_0^\infty s^{1/2}|x(s)|^2 ds = \Gamma(1/2)^2\|x\|_2^2,\end{aligned}$$

zatem $\|L\| \leqslant \Gamma(1/2) = \sqrt{\pi}$.

Z drugiej strony dla $\varepsilon > 0$ i funkcji x_ε określonej wzorem $x_\varepsilon(t) = t^{-1/2+\varepsilon} e^{-t}$ mamy

$$\|x_\varepsilon\|_2^2 = \int_0^\infty t^{2\varepsilon-1} e^{-2t} dt = 2^{-2\varepsilon}\Gamma(2\varepsilon)$$

oraz

$$Lx_\varepsilon(t) = \int_0^\infty s^{-1/2+\varepsilon} e^{-(t+1)s} ds = \Gamma(1/2+\varepsilon)\,(t+1)^{-1/2-\varepsilon},$$

więc

$$\|Lx_\varepsilon\|_2^2 = \Gamma(1/2+\varepsilon)^2 \int_0^\infty \frac{dt}{(t+1)^{1+2\varepsilon}} = \frac{1}{2\varepsilon}\Gamma(1/2+\varepsilon)^2.$$

Ponieważ $\Gamma(2\varepsilon) = \dfrac{\Gamma(1+2\varepsilon)}{2\varepsilon}$, więc

$$\frac{\|Lx_\varepsilon\|_2^2}{\|x_\varepsilon\|_2^2} = \frac{\Gamma(1/2+\varepsilon)^2}{2^{-2\varepsilon}\Gamma(1+2\varepsilon)} \longrightarrow \pi. \quad \square$$

Metodę „preparowania” funkcji podcałkowej przed stosowaniem nierówności Schwarza, którą z sukcesem stosowaliśmy dla dowodu ograniczoności transformaty Laplace'a, można uogólnić w następujący sposób:

7.7. Kryterium Schura. *Niech (Ω, μ) będzie dowolną przestrzenią miarową oraz $\mathcal{K} : \Omega \times \Omega \to \mathbb{C}$ funkcją $\mu \times \mu$ mierzalną. Załóżmy, że istnieje taka funkcja mierzalna dodatnia φ na Ω i takie dwie stałe M, N, że*

$$\begin{aligned} &\int_\Omega \big|\mathcal{K}(s,t)\big|\,\varphi(s)\,ds \leqslant M\,\varphi(t) \quad \textit{dla} \quad t \in \Omega, \\ &\int_\Omega \big|\mathcal{K}(s,t)\big|\,\varphi(t)\,dt \leqslant N\,\varphi(s) \quad \textit{dla} \quad s \in \Omega. \end{aligned} \tag{7.34}$$

Wtedy operator całkowy K, określony na przestrzeni $L^2(\Omega, \mu)$ wzorem

$$Kx(t) = \int_\Omega \mathcal{K}(s,t)\,x(s)\,ds,$$

jest ograniczony i zachodzi nierówność $\|K\| \leqslant \sqrt{MN}$.

Dowód: Dla $x \in L^2(\Omega, \mu)$, po zastosowaniu nierówności Schwarza a następnie kolejno nierówności (7.34), dostajemy

$$
\begin{aligned}
\|Kx\|_2^2 &\leqslant \int_\Omega \left(\int_\Omega \sqrt{|\mathcal{K}(s,t)|\,\varphi(s)} \sqrt{|\mathcal{K}(s,t)|\,|x(s)|^2 \varphi(s)^{-1}}\, ds \right) dt \\
&\leqslant \int_\Omega \int_\Omega |\mathcal{K}(s,t)|\,\varphi(s)\, ds \int_\Omega |\mathcal{K}(s,t)|\,|x(s)|^2 \varphi(s)^{-1}\, ds\, dt \\
&\leqslant M \int_\Omega \int_\Omega |\mathcal{K}(s,t)|\,|x(s)|^2\, \varphi(s)^{-1} \varphi(t)\, ds\, dt \\
&= M \int_\Omega |x(s)|^2\, \varphi(s)^{-1} \int_\Omega |\mathcal{K}(s,t)|\,\varphi(t)\, dt\, ds \\
&\leqslant MN \int_\Omega |x(s)|^2\, ds = MN\, \|x\|_2^2. \quad \square
\end{aligned}
$$

Posługując się kryterium Schura można łatwo dowodzić ograniczoności wielu operatorów. Kłopot polega na odgadnięciu funkcji φ. Proponuję spróbować swych sił i pomysłowości przy rozwiązywaniu następującego zadania:

7.8. Zadanie. Pokazać, że na przestrzeni ℓ^2 operator wyznaczony przez **macierz Hilberta**

$$\left(\frac{1}{j+k-1} \right)_{j.k=1}^{\infty}$$

przyporządkowujący wektorowi $x = (x_1, x_2, x_3, \ldots)$ wektor $y = (y_1, y_2, y_3, \ldots)$ określony wzorem

$$y_k = \sum_{j=1}^{\infty} \frac{1}{j+k-1}\, x_j, \qquad k = 1, 2, 3, \ldots$$

jest ograniczony i jego norma nie przekracza π.

Podamy na zakończenie jeszcze jeden przykład. W nim posłużymy się nierównością Schwarza dla obliczenia normy operatora Volterry na $L^2(0,1)$. Kryterium Schura nie ma tu zastosowania.

7.9. Przykład. Na przestrzeni $L^2[0,1]$ **operator Volterry** V

$$Vx(t) = \int_0^t x(s)\, ds$$

ma normę $\|V\| = \dfrac{2}{\pi}$.

Dla funkcji $x \in L^2[0,1]$ z nierówności Schwarza otrzymujemy

$$\begin{aligned}\|Vx\|_2^2 &= \int_0^1 \left| \int_0^t x(s)\, ds \right|^2 dt = \int_0^1 \left| \int_0^t \sqrt{\cos \frac{\pi}{2}s} \cdot \frac{x(s)}{\sqrt{\cos \frac{\pi}{2}s}}\, ds \right|^2 dt \\ &\leqslant \int_0^1 \int_0^t \cos \frac{\pi}{2}s\, ds \int_0^t \frac{|x(s)|^2}{\cos \frac{\pi}{2}s}\, ds\, dt = \frac{2}{\pi} \int_0^1 \int_0^t \sin \frac{\pi}{2}t\, \frac{|x(s)|^2}{\cos \frac{\pi}{2}s}\, ds\, dt \\ &= \frac{2}{\pi} \int_0^1 \int_s^1 \sin \frac{\pi}{2}t\, dt\, \frac{|x(s)|^2}{\cos \frac{\pi}{2}s}\, ds = \frac{4}{\pi^2} \int_0^1 |x(s)|^2 ds = \frac{4}{\pi^2}\|x\|_2^2,\end{aligned}$$

co dowodzi, że $\|V\| \leqslant \frac{2}{\pi}$. Równość jest wymuszona przez funkcję $x(t) = \cos \frac{\pi}{2}t$, bo $Vx(t) = \frac{2}{\pi} \sin \frac{\pi}{2}t$, a funkcje $\sin \frac{\pi}{2}t$ i $\cos \frac{\pi}{2}t$ mają te same normy. Można też zauważyć, że dla funkcji $x(t) = \cos \frac{\pi}{2}t$ powyższa nierówność Schwarza przechodzi w równość.

Przestrzeń $\mathcal{L}(X, Y)$

7.10. TWIERDZENIE. *Niech X i Y będą przestrzeniami unormowanymi. Zbiór $\mathcal{L}(X, Y)$ wszystkich ograniczonych operatorów liniowych z X w Y, po wprowadzeniu działań dodawania operatorów i mnożenia operatorów przez liczby w sposób następujący*

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

oraz normy jako normy operatorowej, staje się przestrzenia liniową unormowaną. Jeżeli przestrzeń Y jest zupełna, to $\mathcal{L}(X, Y)$ także jest przestrzenią zupełną.

Dowód: Sprawdzenie, że $\mathcal{L}(X, Y)$ jest przestrzenią liniową unormowaną jest nietrudne. Udowodnimy więc tylko drugą część twierdzenia. Załóżmy, że Y jest przestrzenią zupełną i niech $\{T_n\}$ będzie ciągiem Cauchy'ego w $\mathcal{L}(X, Y)$. Ponieważ dla każdego $x \in X$

$$\|T_n x - T_m x\| \leqslant \|T_n - T_m\|\, \|x\|,$$

więc $\{T_n x\}$ jest ciągiem Cauchy'ego w przestrzeni Y. Z zupełności Y wynika istnienie granicy $\lim_{n\to\infty} T_n x$, którą oznaczymy Tx. Pokażemy kolejno, że przyporządkowanie $x \to Tx$ jest liniowe, ciągłe oraz, że $\|T_n - T\| \to 0$

Liniowość odwzorowania T otrzymujemy łatwo przechodząc do granicy w równościach

$$T_n(x_1 + x_2) = T_n x_1 + T_n x_2,$$
$$T_n(\lambda x) = \lambda T_n x.$$

Pokażemy, że operator T jest ciągły. Zauważmy w tym celu, że $\{\|T_n\|\}$ jest liczbowym ciągiem Cauchy'ego, a więc dla pewnej stałej M zachodzi $\|T_n\| \leqslant M$, $n = 1, 2, 3, \ldots$, stąd dla każdego $x \in X$ otrzymujemy $\|T_n x\| \leqslant M\,\|x\|$, a to z kolei pociąga

$$\|Tx\| \leqslant M\,\|x\|.$$

Pozostaje więc tylko dowieść, że T_n dąży do T w $\mathcal{L}(X, Y)$. Dla dowolnego $\varepsilon > 0$ mamy

$$\|T_n x - T_m x\| \leqslant \|T_n - T_m\|\,\|x\| < \varepsilon\,\|x\|$$

przy dostatecznie dużych m i n. Ustalając n i przechodząc do granicy przy $m \to \infty$ otrzymamy

$$\|T_n x - Tx\| \leqslant \varepsilon\,\|x\|,$$

a to z definicji normy operatorowej daje $\|T_n - T\| \leqslant \varepsilon$. $\square$

7.11. Zadanie. Gdy Y jest przestrzenią niezupełną, to $\mathcal{L}(X, Y)$ także jest niezupełna.

Operatory odwracalne

Definicja. Niech $T \in \mathcal{L}(X, Y)$. Jeżeli istnieje taki operator $S \in \mathcal{L}(Y, X)$, że ST i TS są operatorami identycznościowymi odpowiednio na przestrzeniach X i Y, to mówimy, że S jest operatorem **odwrotnym** do T i piszemy $S = T^{-1}$. Operator odwrotny jest jedyny, bo jeżeli S_1, S_2 są odwrotnymi do T, to

$$S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2. \tag{7.35}$$

7.12. Uwagi.

1. Może się zdarzyć, że zachodzi tylko jedna z równości, np. $ST = I$, zaś $TS \neq I$. Mówimy wtedy, że S jest **lewym odwrotnym** dla T, a T **prawym odwrotnym** dla S. Rozpatrzmy na przestrzeni ℓ^2 operatory

$$S(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots),$$
$$T(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots).$$

Wtedy $ST = I$, zaś $TS(x_1, x_2, x_3, \ldots) = (0, x_2, x_3, \ldots) \neq (x_1, x_2, x_3, \ldots)$, więc $TS \neq I$.

2. Ze wzoru (7.35) wynika, że jeżeli operator S_1 jest lewym odwrotnym dla T, a S_2 prawym odwrotnym dla T, to $S_1 = S_2$, tzn. operator T jest odwracalny.
3. Jeżeli operator $T \in \mathcal{L}(X, Y)$ jest odwracalny, to przestrzenie X i Y są izomorficzne. Pozwala to sprowadzić badania do przypadku $X = Y$.

7.13. Twierdzenie. *Niech X będzie przestrzenią Banacha. Zbiór elementów odwracalnych w $\mathcal{L}(X)$ tworzy grupę względem mnożenia i jest podzbiorem otwartym w $\mathcal{L}(X)$. Na nim odwzorowanie $T \to T^{-1}$ jest ciągłe.*

Dowód: Jeżeli operatory T i S są odwracalne, to odwracalny jest też operator TS i $(TS)^{-1} = S^{-1}T^{-1}$, bo

$$(TS)(S^{-1}T^{-1}) = T(SS^{-1})T^{-1} = TT^{-1} = I$$

oraz $(S^{-1}T^{-1})(ST) = I$. Oznacza to, że operatory odwracalne w $\mathcal{L}(X)$ z operacją składania tworzą grupę.

Pokażemy teraz, że jeżeli T jest operatorem odwracalnym oraz $\|S - T\| < 1/\|T^{-1}\|$, to S jest także operatorem odwracalnym. Wystarczy w tym celu pokazać tylko, że $A = T^{-1}S$ jest operatorem odwracalnym. Z założenia mamy

$$\|I - A\| = \|T^{-1}(T - S)\| \leqslant \|T^{-1}\|\,\|T - S\| < 1,$$

więc szereg

$$B = \sum_{n=0}^{\infty}(I - A)^n$$

(przyjmujemy tu $(I - A)^0 = I$) jest bezwzględnie zbieżny w $\mathcal{L}(X)$. Ponieważ

$$AB = B - (I - A)B = \sum_{n=0}^{\infty}(I - A)^n - \sum_{n=1}^{\infty}(I - A)^n = I$$

oraz $BA = B - B(I - A) = I$, więc $B = A^{-1}$.

Operator S^{-1} można wprost napisać wzorem

$$S^{-1} = \sum_{n=0}^{\infty}\left[T^{-1}(T - S)\right]^n T^{-1}.$$

Jeżeli $\|S - T\| \leqslant \varepsilon\,\|T^{-1}\|^{-1}$, gdzie $0 < \varepsilon < 1$, to $\|S^{-1}\| \leqslant \frac{1}{1-\varepsilon}\,\|T^{-1}\|$, więc

$$\|S^{-1} - T^{-1}\| = \|S^{-1}(T - S)T^{-1}\| \leqslant \|S^{-1}\|\,\|T - S\|\,\|T^{-1}\| \leqslant \frac{\varepsilon}{1-\varepsilon}\,\|T^{-1}\|. \quad \square$$

7.14. Twierdzenie (o promieniu spektralnym). *Niech X będzie przestrzenią unormowaną. Dla każdego operatora $T \in \mathcal{L}(X, Y)$ istnieje granica*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Liczba $r(T)$ nosi nazwę **promień spektralny** *operatora T.*

Dowód: Ustalmy liczbę naturalną m i przedstawmy n w postaci $n = km + r$, gdzie $0 \leqslant r < m$. Niech $M = \max\{1, \|T\|, \|T^2\|, \dots, \|T^{m-1}\|\}$, wtedy

$$\|T^n\| \leqslant \|T^m\|^k \|T\|^r \leqslant M\, \|T^m\|^k,$$

skąd

$$\limsup_{n\to\infty} \|T^n\| \leqslant \lim_{n\to\infty} M^{1/n}\, \|T^m\|^{k/n}.$$

Ponieważ $\dfrac{k}{n} = \dfrac{1}{m} - \dfrac{r}{m} \to \dfrac{1}{m}$ przy $n \to \infty$, otrzymujemy

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \|T^m\|^{1/m},$$

co daje

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leqslant \liminf_{m\to\infty} \|T^m\|^{1/m}. \quad \square$$

7.15. Przykład. Podamy przykład operatora **operatora quasi-nilpotentnego**, tzn. operatora, którego promień spektralny wynosi zero. Jest nim operator V określony na $C[0,1]$ wzorem

$$Vx(t) = \int_0^t x(s)\, ds,$$

zwany **operatorem Volterry**.

Łatwo sprawdzamy przez indukcję, że

$$V^n x(t) = \int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\, x(s)\, ds, \qquad n = 1, 2, 3, \dots \tag{7.36}$$

Pokażemy, że $\|V^n\| = \dfrac{1}{n!}$. Jeżeli $x \in C[0,1]$ i $\|x\|_\infty \leqslant 1$, to

$$\begin{aligned}
\sup_{t\in[0,1]} |V^n x(t)| &\leqslant \sup_{t\in[0,1]} \int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\, |x(s)|\, ds \\
&\leqslant \sup_{t\in[0,1]} \int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\, ds = \int_0^1 \frac{(t-s)^{n-1}}{(n-1)!}\, ds = \frac{1}{n!}.
\end{aligned}$$

Ponieważ nierówności te przechodzą w równości dla $x \equiv 1$, więc $\|V^n\| = \dfrac{1}{n!}$. Wynika stąd, że $r(T) = \lim_{n\to\infty} \|V^n\|^{1/n} = 0$.

7.16. TWIERDZENIE. *Niech T będzie ograniczonym operatorem liniowym na przestrzeni Banacha X i niech $r(T)$ oznacza jego promień spektralny. Jeżeli liczba zespolona λ spełnia nierówność $|\lambda| > r(T)$, to operator $(\lambda I - T)^{-1}$ istnieje i można go przedstawić w postaci bezwzględnie zbieżnego szeregu*

$$(\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}. \tag{7.37}$$

Dowód: Zauważmy, że

$$\lim_{n\to\infty} \sqrt[n]{\left\|\frac{T^n}{\lambda^{n+1}}\right\|} = \frac{1}{|\lambda|} \lim_{n\to\infty} \sqrt[n]{\|T^n\|} = \frac{r(T)}{|\lambda|} < 1,$$

a więc szereg (7.37) jest bezwzględnie zbieżny w $\mathcal{L}(X)$, a skoro $\mathcal{L}(X)$ jest przestrzenią zupełną, to szereg ten jest także zbieżny. Niech S oznacza jego sumę. Sprawdzamy łatwo, że $S(\lambda I - T) = (\lambda I - T)S = I$, a więc, że $S = (\lambda I - T)^{-1}$. $\square$

7.17. PRZYKŁAD. Wiemy już, że dla operatora Volterry V zachodzi równość $r(V) = 0$. Wobec tego dla dowolnej liczby zespolonej $\lambda \neq 0$ operator $(\lambda I - V)$ jest odwracalny. Wynika stąd, że równanie całkowe

$$y(t) = \lambda\, x(t) - \int_0^t x(s)\, ds,$$

tzn. równanie $y = (\lambda I - V)x$ ma rozwiązanie

$$x = \sum_{n=0}^{\infty} \frac{V^n y}{\lambda^{n+1}}.$$

Można przy tym korzystać ze wzoru (7.36).

7.18. PRZYKŁAD. Załóżmy, że chcemy zbudować kolumnę mostu tak, by naprężenie w każdym punkcie kolumny pod naciskiem siły $\vec{Q}$ było stałe i wynosiło c. Jeżeli $x(t)$ oznacza pole przekroju kolumny na wysokości t mierzonej od szczytu a ρ ciężar właściwy materiału z którego wykonamy kolumnę, to otrzymamy równanie

$$Q + \rho \int_0^t x(s)\, ds = c\, x(t),$$

czyli równanie typu

$$y(t) = \lambda x(t) - \int_0^t x(s)\, ds,$$

gdzie $y(t) = \frac{Q}{\rho}\mathbf{1}(t)$ i $\lambda = \frac{c}{\rho}$. Ponieważ $V\mathbf{1}(t) = \int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\,ds = \frac{t^n}{n!}$, więc

$$x(t) = \sum_{n=0}^{\infty} \frac{V^n y(t)}{\lambda^{n+1}} = \frac{Q}{c}\sum_{n=0}^{\infty} \frac{t^n}{n!\,\lambda^n} = \frac{Q}{c}\,e^{t/\lambda} = \frac{Q}{c}\,\exp(\rho\, t/c).$$

7.19. PRZYKŁAD. Równanie różniczkowe 1-szego rzędu

$$x' = \alpha x + x_0, \qquad \alpha \in \mathbb{C}, \quad x_0 \in C[0,1]$$

ma rozwiązanie

$$x(t) = \int_0^t e^{\alpha(t-s)}\, x_0(s)\, ds.$$

Istotnie, połóżmy $\lambda = \frac{1}{\alpha}$ i $y = \lambda V x_0$. Wtedy $y' = \lambda x' - x$, lub inaczej $y = \lambda x - Vx$. Zatem

$$x = \sum_{n=0}^{\infty} \frac{V^n y}{\lambda^{n+1}} = \sum_{n=0}^{\infty} \frac{V^{n+1} x_0}{\lambda^n} = \sum_{n=0}^{\infty} \alpha^n\, V^{n+1} x_0.$$

Uwzględniając wzór (7.36) otrzymamy

$$x(t) = \int_0^t \sum_{n=0}^{\infty} \alpha^n\, \frac{(t-s)^n}{n!}\, x_0(s)\, ds = \int_0^t e^{\alpha(t-s)}\, x_0(s)\, ds.$$

Operator sprzężony

7.20. TWIERDZENIE (O OPERATORZE SPRZĘŻONYM). *Niech $\mathcal{H}$ będzie przestrzenią Hilberta. Dla każdego operatora $T \in \mathcal{L}(\mathcal{H})$ istnieje dokładnie jeden operator $T^* \in \mathcal{L}(\mathcal{H})$ o tej własności, że*

$$\langle Tx,\, y\rangle = \langle x,\, T^*y\rangle \tag{7.38}$$

dla każdej pary $x, y \in \mathcal{H}$. Ponadto zachodzi równość $\|T^*\| = \|T\|$.

Dowód: Ustalmy $y \in \mathcal{H}$. Odwzorowanie $x \to \langle Tx,\, y\rangle$ jest liniowe, a ponieważ $|\langle Tx,\, y\rangle| \leqslant \|T\|\,\|x\|\,\|y\|$, jest też ciągłe. Określa zatem ciągły funkcjonał liniowy na $\mathcal{H}$. Z twierdzenia Riesza o postaci funkcjonału liniowego 4.10 wynika, że $\langle Tx,\, y\rangle = \langle x,\, z\rangle$ dla pewnego wektora $z \in \mathcal{H}$. Wektor z jest jedyny, ale zależy od wyboru y. Oznaczmy go zatem $z = T^*y$. Z własności iloczynu skalarnego wnioskujemy, że przyporządkowanie $y \to T^*y$ jest liniowe, a ponieważ

$$\|T^*y\| = \sup_{\|x\|\leqslant 1} |\langle x,\, T^*y\rangle| = \sup_{\|x\|\leqslant 1} |\langle Tx,\, y\rangle| \leqslant \|T\|\,\|y\|,$$

to określa ograniczony operator liniowy T^* na $\mathcal{H}$ o normie $\|T^*\| \leqslant \|T\|$. W istocie zachodzi równość $\|T^*\| = \|T\|$, gdyż z nierówności $\|T^*\| \leqslant \|T\|$ wynika nierówność $\|T^{**}\| \leqslant \|T^*\|$, a jak widać z (7.38), $T^{**} = T$. $\square$

DEFINICJA. Operator T^* nazywamy **operatorem sprzężonym** do T. Jeżeli zachodzi równość $T^* = T$, to mówimy, że T jest operatorem **samosprzężonym** lub **hermitowskim**. Operator T, dla którego $T^*T = TT^*$, tzn. operator komutujący ze swoim sprzężonym, nosi nazwę operatora **normalnego**.

7.21. PRZYKŁADY.

1. Operator S^* sprzężony do operatora przesunięcia S na ℓ^2

$$S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$$

ma postać

$$S^*(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots),$$

bo

$$\langle Sx,\, y\rangle = \sum_{n=1}^{\infty} x_n\, \overline{y_{n+1}} = \langle x,\, S^*y\rangle.$$

2. Dla operatora Volterry $Vx(t) = \int_0^t x(s)\,ds$ na przestrzeni $L^2(0,1)$ otrzymujemy

$$V^*x(t) = \int_t^1 x(s)\,ds,$$

bo

$$\langle Vx,\, y\rangle = \int_0^1 \int_0^t x(s)\,\overline{y(t)}\,ds\,dt = \int_0^1 \int_s^1 x(s)\,\overline{y(t)}\,dt\,ds = \langle x,\, V^*y\rangle.$$

3. Jeżeli A jest operatorem liniowym na $\mathbb{C}^n$, o macierzy $\{a_{i,j}\}_{i,j=1}^n$, to operator sprzężony A^* ma macierz $\{\overline{a_{j,i}}\}_{i,j=1}^n$, transponowaną sprzężoną.

7.22. TWIERDZENIE. *Operacja $T \to T^*$ ma następujące własności:*

1. $(T+S)^* = T^* + S^*$,

2. $(\lambda T)^* = \overline{\lambda} T^*$,

3. $T^{**} = T$,

4. $(ST)^* = T^*S^*$,

5. $\|T^*\| = \|T\|$,

6. $\|T^*T\| = \|T\|^2$,

7. *jeśli T^{-1} istnieje, to istnieje też $(T^*)^{-1}$ i $(T^*)^{-1} = (T^{-1})^*$.*

7.23. Twierdzenie. *Norma operatora hermitowskiego wyraża są wzorem*

$$\|T\| = \sup_{\|x\|\leqslant 1} |\langle Tx,\, x\rangle|.$$

Dowód: Wystarczy ograniczyć się do przypadku, gdy $\|T\| = 1$. Istotnie, dla $T = 0$ twierdzenie jest oczywiście prawdziwe, a gdy $T \neq 0$, to kładąc $S = \frac{1}{\|T\|}T$ otrzymamy

$$1 = \sup_{\|x\|\leqslant 1} \big|\langle Sx,\, x\rangle\big| = \frac{1}{\|T\|} \sup_{\|x\|\leqslant 1} \big|\langle Tx,\, x\rangle\big|,$$

a stąd

$$\|T\| = \sup_{\|x\|\leqslant 1} \big|\langle Tx,\, x\rangle\big|.$$

Załóżmy więc, że $\|T\| = 1$ i oznaczmy $\sup_{\|x\|\leqslant 1} \big|\langle Tx,\, x\rangle\big| = a$. Mamy pokazać, że $a = 1$. Z nierówności Schwarza wynika, że

$$\big|\langle Tx,\, x\rangle\big| \leqslant \|Tx\|\,\|x\| \leqslant \|x\|^2,$$

zatem $a \leqslant 1$. Ustalmy $x \in \mathcal{H}$ i oznaczmy $y = x + Tx$, $z = x - Tx$. Wtedy

$$\begin{aligned}\langle Ty,\, y\rangle &= \langle Tx,\, x\rangle + \langle T^2x,\, x\rangle + \langle Tx,\, Tx\rangle + \langle T^2x,\, Tx\rangle,\\ \langle Tz,\, z\rangle &= -\langle Tx,\, x\rangle + \langle T^2x,\, x\rangle + \langle Tx,\, Tx\rangle - \langle T^2x,\, Tx\rangle.\end{aligned}$$

Odejmując te nierówności stronami i uwzględniając, że $\langle T^2x,\, x\rangle = \langle Tx,\, Tx\rangle = \|Tx\|^2$, otrzymujemy

$$\langle Ty,\, y\rangle - \langle Tz,\, z\rangle = 4\,\|Tx\|^2.$$

Z drugiej strony $\big|\langle Tx,\, y\rangle\big| \leqslant a\,\|y\|^2$ oraz $\big|\langle Tz,\, z\rangle\big| \leqslant a\,\|z\|^2$, więc

$$4\,\|Tx\|^2 \leqslant \big|\langle Ty,\, y\rangle\big| + \big|\langle Tz,\, z\rangle\big| \leqslant a\big(\|y\|^2 + \|z\|^2\big) = a\big(\|x + Tx\|^2 + \|x - Tx\|^2\big).$$

Korzystając z równości równoległoboku i tego, że $a \leqslant 1$ możemy napisać

$$4\|Tx\|^2 \leqslant 2a\big(\|x\|^2 + \|Tx\|^2\big) \leqslant 2a\|x\|^2 + 2\|Tx\|^2,$$

czyli

$$\|Tx\|^2 \leqslant a\,\|x\|^2,$$

a stąd już łatwo otrzymujemy $1 = \|T\| = \sup_{\|x\|\leqslant 1} \|Tx\| \leqslant a$. $\square$

7.24. TWIERDZENIE.

1. *Każdy operator* $T \in \mathcal{L}(\mathcal{H})$ *ma jednoznaczne przedstawienie w postaci*

$$T = T_1 + iT_2,$$

gdzie T_1 *i* T_2 *są operatorami hermitowskimi.*

2. *Operator* T *jest normalny wtedy i tylko wtedy, gdy* T_1 *i* T_2 *komutują.*

3. *Dla operatora normalnego zachodzi równość*

$$\|T\| = \sqrt{\|T_1^2 + T_2^2\|}.$$

Dowód: Jeżeli równość $T = T_1 + iT_2$ zachodzi, to $T^* = T_1 - iT_2$, więc $T_1 = \frac{1}{2}(T + T^*)$ i $T_2 = \frac{1}{2i}(T - T^*)$, można zatem te dwie ostatnie równości przyjąć jako definicje operatorów T_1 i T_2. Punkt 2 wynika z równości $T^*T - TT^* = 2i(T_1T_2 - T_2T_1)$ a punkt 3 z tego, że $\|T\|^2 = \|T^*T\|$ oraz $T^*T = T_1^2 + T_2^2$. □

Operatory normalne

Operatory normalne, w szczególności macierze normalne $n \times n$, tworzą ważną klasę operatorów. Zachodzi dla nich twierdzenie spektralne (patrz **??**), które dla macierzy oznacza możliwość sprowadzenia do postaci diagonalnej, tzn. pozwala przedstawić macierz A w postaci $A = U^{-1}BU$, gdzie U jest macierzą unitarną a B macierzą diagonalną.

7.25. LEMAT. *W* $\mathcal{L}(\mathcal{H})$ *promień spektralny ma własności:*

1. $r(T) \leqslant \|T\|$,

2. $r(T^k) = r(T)^k$,

3. $r(T^*) = r(T)$.

Dowód: Ponieważ $\|T^{kn}\|^{1/n} \leqslant \|T^n\|^{k/n}$, więc przy $n \to \infty$ otrzymujemy $r(T^k) \leqslant r(T)^k$. Z drugiej strony, gdy $n = km + r$, $0 \leqslant r < k$, to $\|T^n\|^{k/n} \leqslant M\,\|T^{km}\|$, gdzie $M = \max_{0 \leqslant r < k} \|T^r\|$, więc $\|T^k\|^{k/n} \leqslant \|T^{km}\|^{1/m} M^{k/n}$, dlatego $r(T^k) \leqslant r(T^k)$. □

7.26. TWIERDZENIE. *Dla operatora normalnego* $T \in \mathcal{L}(\mathcal{H})$ *zachodzi równość*

$$\|T^n\| = \|T\|^n, \qquad n = 1, 2, 3, \ldots,$$

w szczególności $r(T) = \|T\|$.

Dowód: Mamy

$$\|T^2\|^2 = \|(T^*)^2T^2\| = \|(T^*T)(T^*T)\| = \|T^*T\|^2 = \|T\|^4.$$

a dalej przez indukcję dostajemy $\|T^{2^n}\| = \|T\|^{2^n}$. Wynika stąd, że $r(T) = \|T\|$. Ponieważ T^n jest także operatorem normalnym, więc $\|T^n\| = r(T^n)$, ale z Lematu wynika, że $t(T^n) = r(T)^n$ więc $\|T^n\| = \|T\|^n$, co kończy dowód. □

7.27. Twierdzenie (o charakteryzacji operatorów normalnych). *Operator $T \in \mathcal{L}(\mathcal{H})$ jest normalny wtedy i tylko wtedy, gdy $\|Tx\| = \|T^*x\|$ dla każdego $x \in \mathcal{H}$.*

Dowód: Niech $T \in \mathcal{L}(\mathcal{H})$ i niech $S = TT^* - T^*T$. Wtedy dla dowolnego $x \in \mathcal{H}$

$$\begin{aligned}\langle Sx,\, x\rangle &= \langle TT^*x - T^*Tx,\, x\rangle = \langle TT^*x,\, x\rangle - \langle T^*Tx,\, x\rangle \\ &= \langle T^*x,\, T^*x\rangle - \langle Tx,\, Tx\rangle = \|T^*x\|^2 - \|Tx\|^2.\end{aligned}$$

Jeżeli T jest operatorem normalnym, to $S = 0$, a zatem równość $\|T^*x\| = \|Tx\|$ zachodzi dla każdego $x \in \mathcal{H}$.

Z drugiej strony wiemy, że ta równość pociąga $\langle Sx,\, x\rangle = 0$ dla wszystkich $x \in \mathcal{H}$. Ponieważ S jest operatorem hermitowskim, to

$$\|S\| = \sup_{\|x\|\leqslant 1} \left|\langle Sz,\, x\rangle\right| = 0,$$

czyli $T^*T = TT^*$. □

7.28. Twierdzenie. *Niech T będzie operatorem normalnym w $\mathcal{L}(\mathcal{H})$. Operator odwrotny T^{-1} istnieje wtedy i tylko wtedy, gdy*

$$\inf_{\|x\|=1} \|Tx\| > 0.$$

Dowód: Oznaczmy

$$\inf_{\|x\|=1} \|Tx\| = c.$$

Jeżeli operator odwrotny T^{-1} istnieje, to dla dowolnego $x \in \mathcal{H}$ mamy

$$\|x\| = \|T^{-1}(Tx)\| \leqslant \|T^{-1}\|\,\|Tx\|,$$

a stąd

$$\|Tx\| \geqslant \|x\|\,\|T^{-1}\|^{-1}.$$

Załóżmy teraz, że $c > 0$. Łatwo pokazuje się, że dla dowolnego x zachodzi nierówność
$$\|x\| \leqslant \frac{1}{c}\,\|Tx\|.$$
wynika stąd, że T jest operatorem różnowartościowym, gdyż równość $Tx_1 = Tx_2$ pociąga $\|x_1 - x_2\| \leqslant \frac{1}{c}\,\|Tx_1 - Tx_2\| = 0$, czyli $x_1 = x_2$.

Aby dowieść, że T^{-1} istnieje, należy jeszcze pokazać, że $T(\mathcal{H}) = \mathcal{H}$. Spełnione wtedy będą założenia twierdzenia o odwzorowaniu odwrotnym.

Pokażemy najpierw, że $T(\mathcal{H})$ jest domkniętą podprzestrzenią liniową w $\mathcal{H}$. Niech $y \in \overline{T(\mathcal{H})}$. Istnieje taki ciąg $\{x_n\}$, że $Tx_n \to y$. ciąg $\{Tx_n\}$ jest ciągiem Cauchy'ego w $\mathcal{H}$, a z tego, że $\|x_n - x_m\| \leqslant \frac{1}{c}\,\|Tx_n - Tx_m\|$, $n, m = 1, 2, 3, \ldots$ wynika, że także $\{x_n\}$ jest ciągiem Cauchy'ego w $\mathcal{H}$. Oznaczmy $\lim_{n\to\infty} x_n = x$. Wtedy
$$\|Tx - y\| = \lim_{n\to\infty} \|Tx_n - y\| = 0,$$
w rezultacie $y \in T(\mathcal{H})$.

Teraz dowód sprowadza się do pokazania, że $T(\mathcal{H})^{\perp} = \{0\}$. Weźmy $y \in T(\mathcal{H})^{\perp}$. Wtedy
$$\langle x,\, T^*y\rangle = \langle Tx,\, y\rangle = 0$$
dla każdego $x \in \mathcal{H}$. Jest to możliwe tylko wtedy, gdy $T^*y = 0$. Z założenia operator T jest normalny, zatem $\|T^*y\| = \|Ty\|$, co daje
$$0 = \|T^*y\| = \|Ty\| \geqslant \frac{1}{c}\,\|y\|,$$
czyli $y = 0$. $\square$

Forma kwadratowa operatora

Definicja. Niech $T \in \mathcal{L}(\mathcal{H})$. Funkcję $f_T : \mathcal{H} \to \mathbb{C}$ postaci
$$f_T(x) = \langle Tx,\, x\rangle$$
nazywamy **formą kwadratową** operatora T.

7.29. Twierdzenie. *Niech T będzie operatorem w zespolonej przestrzeni Hilberta $\mathcal{H}$ a f_T jego formą kwadratową. Wtedy:*

- **a.** *f_T jednoznacznie wyznacza T.*
- **b.** *f_T jest funkcją rzeczywistą wtedy i tylko wtedy, gdy T jest operatorem hermitowskim.*
- **c.** *Jeżeli oznaczymy $\|f_T\| = \sup_{\|x\|\leqslant 1} \left|f_T(x)\right|$, to*
$$\|f_T\| \leqslant \|T\| \leqslant 2\,\|f_t\|.$$

Dowód: Z równości równoległoboku mamy

$$(7.39)\qquad \langle Tx,\, y\rangle = \frac{1}{4}\big[f_T(x+y) - f_T(x-y) + i\, f_T(x+iy) - i\, f_T(x-iy)\big].$$

wynika stąd, że jeżeli $f_{T_1} = f_{T_2}$, to $\langle T_1 x,\, y\rangle = \langle T_2 x,\, y\rangle$ dla wszystkich $x, y \in \mathcal{H}$, a więc $T_1 = T_2$. Zauważmy, że T jest operatorem hermitowskim wtedy i tylko wtedy, gdy f_T jest funkcją rzeczywistą. Ponieważ $f_T(x) \leqslant \|f_T\|\,\|x\|^2$ dla dowolnego wektora $x \in \mathcal{H}$, więc z (7.39) wynika, że

$$\begin{aligned}\big|\langle Tx,\, y\rangle\big| &\leqslant \frac{1}{4}\,\|f_t\|\,\big(\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2\big)\\ &\leqslant \|f_T\|\,\big(\|x\|^2 + \|y\|^2\big),\end{aligned}$$

więc

$$\|T\| = \sup_{\substack{\|x\|\leqslant 1\\ \|y\|\leqslant 1}} \big|\langle Tx,\, y\rangle\big| \leqslant 2\,\|f_T\|.$$

Nierówność $\|f_T\| \leqslant \|T\|$ jest oczywista. $\square$

Uwaga. W rzeczywistej przestrzeni Hilberta żadna z własności a, b, c nie zachodzi. Kontrprzykładem na każdą z nich jest operator $T = \begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}$ w $\mathbb{R}^2$. Mamy tu $f_T \equiv 0$.

Uwaga. Dla dowolnego operatora T na przestrzeni Hilberta zbiór

$$\{\langle Tx,\, x\rangle : \|x\| \leqslant 1\}$$

jest domknięty i wypukły (Twierdzenie Toeplitza-Hausdorffa, patrz [H])

Operatory dodatnie

Operator $T \in \mathcal{L}(\mathcal{H})$ nazywamy **dodatnim** i piszemy $T \geqslant 0$, jeżeli jest hermitowski oraz $\langle Tx, x\rangle \geqslant 0$ dla wszystkich $x \in \mathcal{H}$

7.30. TWIERDZENIE. *Niech $\mathcal{H}$ będzie przestrzenią Hilberta. Wtedy*

1. *T^*T jest operatorem dodatnim dla dowolnego $T \in \mathcal{L}(\mathcal{H})$.*

2. *W zespolonej przestrzeni Hilberta warunek $\langle Tx, x\rangle \geqslant 0$ dla wszystkich $x \in \mathcal{H}$ pociąga hermitowskość operatora T.*

3. *Jeżeli $T \geqslant 0$, to $(I+T)^{-1}$ istnieje i $\left\|(I+T)^{-1}\right\| \leqslant 1$.*

Dowód: ad 3. Mamy

$$\|(I+T)x\|^2 = \|x\|^2 + \|Tx\|^2 + 2\langle Tx, x\rangle \geqslant \|x\|^2,$$

więc spełnione są założenia twierdzenia o istnieniu operatora odwrotnego dla operatora normalnego 7.28. Otrzymana nierówność pociąga także nierówność $\left\|(I+T)^{-1}\right\| \leqslant 1$. $\square$

DEFINICJA. Dla operatorów $S, T \in \mathcal{L}(\mathcal{H})$ zapis $S \leqslant T$, lub $T \geqslant S$ oznacza, że $T - S$ jest operatorem dodatnim.

7.31. ZADANIE. Niech A, B, C będą operatorami na przestrzeni Hilberta $\mathcal{H}$. Pokazać, że:

1. $A \leqslant B \wedge B \leqslant A \Rightarrow A = B$.

2. $A \leqslant B \wedge B \leqslant C \Rightarrow A \leqslant C$.

3. $A_1 \leqslant A_2 \wedge B_1 \leqslant B_2 \Rightarrow A_1 + B_1 \leqslant A_2 + B_2$.

4. $A \leqslant B \wedge \lambda \geqslant 0 \Rightarrow \lambda A \leqslant \lambda B$.

5. $A \leqslant B \Rightarrow -A \geqslant -B$.

7.32. PRZYKŁAD. Niech T będzie operatorem hermitowskim. Wtedy

$$m\,I \leqslant T \leqslant M\,I,$$

gdzie $m = \inf\limits_{\|x\|=1} \langle Tx, x\rangle$, $M = \sup\limits_{\|x\|=1} \langle Tx, x\rangle$. Ponadto

$$\|T\| = \max\left\{|m|, |M|\right\}.$$

DEFINICJA. Liczby m, M nazywamy odpowiednio **kresem dolnym** oraz **kresem górnym** operatora T.

7.33. TWIERDZENIE. *Jeżeli $m > 0$ oraz $m\,I \leqslant T \leqslant M\,I$, to operator odwrotny T^{-1} istnieje i spełnia nierówności*

$$\frac{1}{M}\,I \leqslant T^{-1} \leqslant \frac{1}{m}\,I.$$

Dowód: Z założenia $T = m\,I + A$, gdzie $A \geqslant 0$, zatem

$$\|Tx\|^2 \geqslant m^2\|x\|^2 + \|Ax\|^2 + 2m\langle Ax,\, x\rangle \geqslant m^2\|x\|^2.$$

Z twierdzenia o operatorze odwrotnym do operatora normalnego 7.28 wnosimy, że operator T^{-1} istnieje i $\|T^{-1}\| \leqslant \frac{1}{m}$. Więc $\langle T^{-1}x,\, x\rangle \leqslant \|T^{-1}\|\,\|x\|^2$, a to oznacza, że

$$T^{-1} \leqslant \frac{1}{m}\,I.$$

Jeśli oznaczymy $m' = \inf_{\|x\|=1}\langle T^{-1}x,\, x\rangle$, to $\langle T^{-1}x,\, x\rangle = \langle y,\, Ty\rangle \geqslant 0$ dla $x = Ty$, więc rozumując jak poprzednio dla T otrzymamy $M = \|T\| \leqslant \frac{1}{m'}$, czyli $m' \geqslant \frac{1}{M}$, a stąd

$$\frac{1}{M}\,I \leqslant T^{-1}. \quad \square$$

7.34. TWIERDZENIE (UOGÓLNIONA NIERÓWNOŚĆ SCHWARZA). *Dla operatora dodatniego T w $\mathcal{L}(\mathcal{H})$ i dowolnych $x, y \in \mathcal{H}$ zachodzi nierówność*

$$\big|\langle Tx,\, y\rangle\big|^2 \leqslant \langle Tx,\, x\rangle\langle Ty,\, y\rangle.$$

Dowód: Funkcja $\langle\langle x,\, y\rangle\rangle = \langle tx,\, y\rangle$ ma wszystkie własności iloczynu skalarnego, z wyjątkiem być może $\langle\langle x,\, y\rangle\rangle = 0$ wtedy i tylko wtedy, gdy $x = 0$. Dla tej funkcji można powtórzyć standartowy dowód nierówności Schwarza. $\square$

7.35. WNIOSKI.

1. *Jeżeli $S \geqslant 0$, to dla każdego $x \in \mathcal{H}$*

$$\|Sx\|^2 \leqslant \|S\|\,\langle Sx,\, x\rangle.$$

2. *Jeżeli $S \geqslant 0$, to równość $\langle Sx_0,\, x_0\rangle = 0$ pociąga $Sx_0 = 0$.*

Dowód: Z uogólnionej nierówności Schwarza dostajemy

$$\|Sx\|^4 = \big|\langle Sx,\, Sx\rangle\big|^2 \leqslant \langle Sx,\, x\rangle\,\langle S(Sx),\, Sx\rangle \leqslant \langle Sx,\, x\rangle\,\|S\|\,\|Sx\|^2$$

a stąd żądaną nierówność. $\square$

Ciągi monotoniczne operatorów

Definicja. Ciąg $\{T_n\}$ operatorów hermitowskich na przestrzeni Hilberta nazywamy **monotonicznym**, jeżeli $T_1 \leqslant T_2 \leqslant T_3 \leqslant \ldots$ lub $T_1 \geqslant T_2 \geqslant T_3 \geqslant \ldots$.

7.36. Twierdzenie (o ciągach monotonicznych). *Każdy monotoniczny i ograniczony ciąg operatorów hermitowskich jest mocno zbieżny, tzn. istnieje taki operator $T \in \mathcal{L}(\mathcal{H})$, że*

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0$$

dla każdego $x \in \mathcal{H}$.

Dowód: Wystarczy pokazać, że dla każdego $x \in \mathcal{H}$ istnieje granica $\lim_{n\to\infty} T_n x$, gdyż równość $Tx = \lim_{n\to\infty} T_n x$, $x \in \mathcal{H}$ określa na przestrzeni $\mathcal{H}$ operator liniowy i ograniczony.

Dla ustalonego $x \in \mathcal{H}$ ciąg $\{\langle t_n x, x\rangle\}$ jest monotonicznym i ograniczonym ciągiem liczb rzeczywistych, jest więc zbieżny. Pokażemy, że pociąga to zbieżność ciągu $\{T_n x\}$. Dla dowolnych wskaźników m, n mamy

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| = \|(T_m - T_n)x\|.$$

Ponieważ przynajmniej jeden z operatorów $T_n - T_m$, $T_m - T_n$ jest dodatni, powiedzmy $T_n - T_m$, zatem z wniosku 7.35.1 otrzymujemy

$$\begin{aligned}\|T_n x - T_m x\|^2 = \|(T_n - T_m)x\|^2 &\leqslant \|T_n - T_m\| \, \langle (T_n - T_m)x, x\rangle \\ &\leqslant 2M\Big(\langle T_n x, x\rangle - \langle T_m x, x\rangle\Big),\end{aligned}$$

gdzie $M = \sup_{k\geqslant 1} \|T_k\|$.

Wiemy już, że ciąg $\{\langle T_n x, x\rangle\}$ jest zbieżny, jest więc ciągiem Cauchy'ego a to, na mocy dopiero co udowodnionej nierówności oznacza, że $\{T_n x\}$ jest ciągiem Cauchy'ego w $\mathcal{H}$, jest zatem zbieżny. □

7.37. Twierdzenie (o pierwiastku kwadratowym). *Dla dowolnego operatora dodatniego T istnieje, i to dokładnie jeden, taki operator dodatni S, że $S^2 = T$. Operator ten komutuje z każdym operatorem komutującym z T.*

Dowód: Idea dowodu jest ukryta w następującym prostym twierdzeniu z analizy: dla dowolnej liczby $a \in [0, 1]$ ciąg rekurencyjny

$$a_0 = 0, \qquad a_{n+1} = \frac{1}{2}(a_n^2 + 1 - a), \quad n = 0, 1, 2, \ldots$$

jest monotonicznie rosnący i zbieżny do $1-\sqrt{a}$. Zastosujemy to twierdzenie do operatora T, o którym bez ograniczenia ogólności możemy założyć, że $0 \leqslant T \leqslant I$. Utwórzmy rekurencyjny ciąg operatorów

$$S_0 = 0, \qquad S_{n+1} = \frac{1}{2}(S_n^2 + I - T), \quad n = 0, 1, 2, \ldots .$$

Wtedy $0 \leqslant S_1 \leqslant S_2 \leqslant \ldots \leqslant I$, a więc z twierdzenia o ciągach monotonicznych 7.36 ciąg ten jest mocno zbieżny w $\mathcal{L}(\mathcal{H})$ do pewnego operatora dodatniego B. Ponieważ $0 \leqslant B \leqslant I$, więc operator $S = I - B$ jest dodatni. Przechodząc do granicy w równości

$$S_{n+1}x = \frac{1}{2}(S_n^2 + I - T)x,$$

prawdziwej dla każdego $x \in \mathcal{H}$, otrzymamy

$$Bx = \frac{1}{2}(B^2 + I - T)x,$$

czyli

$$Tx = (I - B)^2 x = S^2 x,$$

a stąd $T = S^2$.

Dowód będzie zakończony, jeżeli pokażemy, że S jest jedynym operatorem dodatnim o własności $S^2 = T$. Jeżeli dany jest operator $S' \geqslant 0$, $S'^2 = T$, to $S'T = S'^3 = TS'$, tzn. S' jest przemienny z T. Łatwo pokazujemy, że wtedy $S'S_n = S_nS'$ dla $n = 1, 2, 3, \ldots$ i w konsekwencji, że $S'B = BS'$. Mamy więc $S'S = SS'$ a równość $S' = S$ wynika z poniższego lematu. $\square$

7.38. LEMAT. *Jeżeli operatory dodatnie S, T komutują, to równość $S^2 = T^2$ pociąga równość $S = T$.*

Dowód: Niech $x \in \mathcal{H}$ i $y = (S - T)x$. Wtedy

$$0 = \langle (S^2 - T^2)x, y\rangle = \langle (S + T)y, y\rangle = \langle Sy, y\rangle + \langle Ty, y\rangle.$$

Ponieważ S i T są operatorami dodatnimi, więc z wniosku 7.35.2 otrzymujemy $Sy = Ty = 0$, a w konsekwencji

$$\|Sx - Tx\| = \langle Sy - Ty, x\rangle = 0. \quad \square$$

DEFINICJA. Operator S z twierdzenia 7.36 nazywamy **pierwiastkiem kwadratowym** z operatora T i oznaczamy $T^{1/2}$.

7.39. WNIOSEK. *Operator T jest dodatni wtedy i tylko wtedy, gdy jest postaci*

$$T = S^*S. \quad \square$$

Definicja. Operator $U \in \mathcal{L}(\mathcal{H})$ nazywamy **izometrią**, jeżeli $\|Ux\| = \|x\|$ dla każdego $x \in \mathcal{H}$. Jest to równoważne równości $U^*U = I$. Operator U jest **częściową izometrią**, jeżeli $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$, oraz $U|_{\mathcal{H}_0} \equiv 0$ i $U|_{\mathcal{H}_1}$ jest izometrią.

7.40. Twierdzenie o rozkładzie biegunowym. *Każdy operator $T \in \mathcal{L}(\mathcal{H})$ można jednoznacznie przedstawić w postaci*

$$T = US,$$

gdzie S jest operatorem dodatnim a U częściową izometrią na $\mathcal{H}$.

Dowód: Dla operatora $S = (T^*T)^{1/2}$ i $x \in \mathcal{H}$ mamy

$$\|Sx\|^2 = \langle S^2x,\, x\rangle = \langle T^*Tx,\, x\rangle = \langle Tx,\, Tx\rangle = \|Tx\|^2. \tag{7.40}$$

Oznaczmy $\mathcal{H}_1 = S(\mathcal{H})$; jest to podprzestrzeń liniowa w $\mathcal{H}$, choć może nie być domknięta. Dla $y \in \mathcal{H}_1$ istnieje taki wektor $x \in \mathcal{H}$, że $y = Sx$. Połóżmy $Uy = Tx$. Wektor Uy jest dobrze określony, bo jeśli $Sx_1 = Sx_2$, tzn. $S(x_1 - x_2) = 0$, to z (7.40) także $Tx_1 = Tx_2$. Z (7.40) wynika także, że U jest izometrią liniową na $\mathcal{H}_1$, zatem przedłuża się jednoznacznie do izometrii na $\overline{\mathcal{H}_1}$. Kładąc $Ux = 0$ dla $x \in \mathcal{H}_0 = \mathcal{H}_1^{\perp}$ otrzymamy częściową izometrię spełniającą równość $T = US$. □

Widmo operatora, wartości własne

Liczbę zespoloną λ nazywamy **wartością własną** operatora $T \in \mathcal{L}(X)$ jeżeli dla pewnego $x_0 \in X$, $x_0 \neq 0$, zachodzi

$$Tx_0 = \lambda\, x_0.$$

Zbiór $\sigma_p(T)$ wszystkich wartości własnych operatora T nazywamy **widmem punktowym** lub **spektrum punktowym** operatora T.

Niech $\lambda \in \sigma_p(T)$. Wektor $x \in X$ spełniający równość $Tx = \lambda\, x$ nazywamy **wektorem własnym** odpowiadającym wartości własnej λ. Zbiór wszystkich takich wektorów tworzy domkniętą podprzestrzeń liniową w X, zwaną **podprzestrzenią własną** operatora T.

7.41. Przykład. Na przestrzeni ℓ^2 rozpatrzmy operator przesunięcia S postaci

$$S(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots).$$

Jeśli liczba $\lambda \in \mathbb{C}$ jest wartością własną operatora S, a $x = (x_1, x_2, x_3, \ldots)$ odpowiadającym jej, niezerowym wektorem własnym, to

$$x_{k+1} = \lambda\, x_k, \quad k = 1, 2, 3, \ldots, \quad \text{oraz} \quad x_1 \neq 0. \tag{7.41}$$

Gdy $|\lambda| < 1$, to dla $x = (1, \lambda, \lambda^2, \ldots)$ mamy $Sx = \lambda x$, więc $\lambda \in \sigma_p(S)$. Jeżeli $|\lambda| \geqslant 1$, to warunek (7.41) nie może być spełniony dla żadnego wektora $x \in \ell^2$. Dlatego $\sigma_p(S) = \{z \in \mathbb{C} : |z| < 1\}$. Dla operatora sprzężonego S^*

$$S^*(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$$

otrzymujemy $\sigma_p(S^*) = \emptyset$.

7.42. TWIERDZENIE. *Niech T będzie operatorem hermitowskim na przestrzeni Hilberta $\mathcal{H}$. Wtedy:*

- **a.** *Wartości własne operatora T są liczbami rzeczywistymi.*
- **b.** *Podprzestrzenie własne odpowiadające różnym wartościom własnym są ortogonalne.*
- **c.** *W ośrodkowej przestrzeni Hilberta zbiór $\sigma_p(T)$ jest co najwyżej przeliczalny.*

UWAGA. Z przykładu 7.41 widać, że żadna z własności a, b, c nie przysługuje wszystkim ograniczonym operatorom na przestrzeni Hilberta.

DEFINICJA. Niech T będzie ograniczonym operatorem na przestrzeni Banacha X. **Widmem** lub **spektrum** $\sigma(T)$ operatora T nazywamy zbiór

$$\sigma(T) = \{\lambda \in \mathbb{C} : (\lambda I - T)^{-1} \text{ nie istnieje}\}.$$

Zauważmy, że $\sigma_p(T) \subset \sigma(T)$.

7.43. TWIERDZENIE. *Niech $T \in \mathcal{L}(\mathcal{H})$. Wtedy:*

- **1.** *$\sigma(T)$ jest podzbiorem domkniętym płaszczyzny.*
- **b.** *$\sigma(\lambda T) = \lambda\, \sigma(T)$.*
- **c.** *$\sigma(T^*) = \overline{\sigma(T)}$.*
- **d.** *$\sup\{|\lambda| : \lambda \in \sigma(T)\} \leqslant r(T)$.*

Dowód: Domkniętość $\sigma(T)$ wynika wprost z otwartości zbioru elementów odwracalnych przestrzeni $\mathcal{L}(\mathcal{H})$ a równość zbiorów $\sigma(T^*)$ i $\overline{\sigma(T)}$ z tożsamości

$$\left[(\lambda I - T)^{-1}\right]^* = (\overline{\lambda} I - T^*)^{-1}.$$

Jeżeli $\lambda > r(T)$, to szereg

$$\sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$$

jest bezwzględnie zbieżny w $\mathcal{L}(\mathcal{H})$, jego suma to operator odwrotny do operatora $\lambda I - T$, zatem $\lambda \notin \sigma(T)$. $\square$

UWAGA. Pokażemy potem **??**, że $\sigma(T)$ jest zbiorem niepustym oraz, że w **d** zachodzi równość. To tłumaczy nazwę „promień spektralny” dla liczby $r(T)$.

7.44. PRZYKŁADY.

1. Dla operatora S z poprzedniego przykładu mamy

$$\sigma(S) = \sigma(S^*) = \{z \in \mathbb{C} : |z| \leqslant 1\}.$$

Istotnie, operator S jest kontrakcją, więc $r(S) \leqslant \|S\| \leqslant 1$, co oznacza, że $\sigma(S) \subset \{z \in \mathbb{C} :: |z| \leqslant 1\}$. Zawieranie $\supset$ wynika z $\sigma(S) \supset \sigma_p(S) = \{z \in \mathbb{C} : |z| < 1\}$ i domkniętości $\sigma(S)$.

2. Niech operator $T \in \mathcal{L}\big(L^2(0,1)\big)$ będzie określony wzorem

$$Tx(t) = f(t)\,x(t), \quad t \in (0,1),$$

gdzie f jest ograniczoną funkcją mierzalną na $(0,1)$. Liczba λ jest wartością własną operatora T wtedy i tylko wtedy, gdy równość $f(t) = \lambda$ zachodzi na zbiorze miary dodatniej. Liczba λ leży w spektrum operatora T wtedy i tylko wtedy, gdy dla każdego $\varepsilon > 0$ zbiór $\{t \in (0,1) : \big|f(t) - \lambda\big| < \varepsilon\}$ jest miary dodatniej. W szczególności dla funkcji f ciągłej otrzymujemy $\sigma(T) = \overline{f\big((0,1)\big)}$.

7.45. TWIERDZENIE. *Jeżeli T jest operatorem normalnym w $\mathcal{L}(\mathcal{H})$, to*

$$\sigma(T) \subset \overline{\{\langle Tx,\, x\rangle : \|x\| = 1\}}.$$

Wynika stąd, że widmo operatora hermitowskiego jest podzbiorem prostej rzeczywistej a widmo operatora dodatniego podzbiorem półprostej $[0,\infty)$.

Dowód: Jeżeli $\lambda \notin \overline{\{\langle Tx,\, x\rangle : \|x\| = 1\}}$, to

$$0 < \inf_{\|x\|=1} \big|\langle(\lambda I - T)x,\, x\rangle\big| \leqslant \inf_{\|x\|=1} \big\|(\lambda I - T)\big\|,$$

a więc operator $\lambda I - T$ jest odwracalny na mocy twierdzenia 7.28 o istnieniu operatora odwrotnego dla operatora normalnego. $\square$

Rezolwenta operatora

Dla operatora ograniczonego T na przestrzeni Banacha X oznaczmy przez $\rho(T)$ dopełnienie widma $\sigma(T)$ operatora T. Jest to otwarty podzbiór płaszczyzny zespolonej a dla każdego $\lambda \in \rho(T)$ istnieje operator

$$R(\lambda, T) = (\lambda I - T)^{-1},$$

zwany **rezolwentą** operatora T.

7.46. LEMAT (RÓWNANIE HILBERTA). *Niech $T \in \mathcal{L}(X)$. Jeśli $\lambda, \mu \in \rho(T)$, to*

$$R(\lambda, T) - R(\mu, T) = (\mu - \lambda)\, R(\lambda, T)\, R(\mu, T).$$

Dowód: Operatory $R(\lambda, T)$ i $R(\mu, T)$ komutują, więc

$$\begin{aligned}(\lambda\, I - T)\, R(\lambda, T)\, R(\mu, T) &= R(\mu, T),\\ (\mu\, I - T)\, R(\lambda, T)\, R(\mu, T) &= R(\lambda, T).\end{aligned}$$

Po odjęciu tych równości stronami otrzymamy postulowane równanie. $\square$

7.47. LEMAT. *Niech $T \in \mathcal{L}(X)$. Dla dowolnych $x \in X$, $x^* \in X^*$ funkcja*

$$\lambda \longrightarrow x^*\big(R(\lambda, T)x\big)$$

jest holomorficzna na $\rho(T)$.

Dowód: Pokazaliśmy jeż wcześniej (patrz **??**), że funkcja $\lambda \to R(\lambda, T)$ jest ciągła. Z lematu 7.46 mamy więc

$$\begin{aligned}\lim_{\mu\to\lambda} \frac{x^*\big(R(\lambda, T)x\big) - x^*\big(R(\mu, T)x\big)}{\lambda - \mu} &= \lim_{\mu\to\lambda} x^*\big(- R(\lambda, T)R(\mu, T)x\big)\\ &= -x^*\big(R(\lambda, T)^2 x\big).\end{aligned}$$

Oznacza to, że funkcja $\lambda \to x^*\big(R(\lambda, T)\big)$ ma pochodną w każdym punkcie $\lambda \in \rho(T)$. $\square$

7.48. TWIERDZENIE. *W zespolonej przestrzeni Banacha X widmo dowolnego operatora ograniczonego jest zbiorem niepustym.*

Dowód: Jeżeli $|\lambda| > \|T\|$, to $R(\lambda, T) = \sum\limits_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$, więc

$$\big\|R(\lambda, T)\big\| \leqslant \sum_{n=0}^{\infty} \frac{\|T^n\|}{|\lambda|^{n+1}} = \frac{1}{|\lambda| - \|T\|}\,.$$

Wynika stąd, że $\lim_{|\lambda|\to\infty} \big\|R(\lambda, T)\big\| = 0$. Załóżmy niewprost, że $\sigma(T)$ jest zbiorem pustym. Wtedy dla każdych $x \in X$ i $x^* \in X^*$ funkcja $\lambda \longrightarrow x^*\big(R(\lambda, T)x\big)$ jest całkowita (holomorficzna na całej płaszczyźnie zespolonej) i dąży do zera przy $|\lambda| \to \infty$. Jest więc ograniczona. Z twierdzenia Liouville'a wynika, że jest funkcją stałą, musi więc być funkcją tożsamościowo równą zeru. Ponieważ dla każdych $x^* \in X^*$ i $x \in X$ mamy $x^*\big(R(\lambda, T)x\big) = 0$, więc $R(\lambda, T) = 0$, co przeczy odwracalności operatora $R(\lambda, T)$. $\square$

7.49. TWIERDZENIE (O PROMIENIU SPEKTRALNYM). *Niech T będzie ograniczonym operatorem na zespolonej przestrzeni Banacha X i niech*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$$

oznacza jego promień spektralny, wtedy

$$r(T) = \max_{\lambda\in\sigma(T)} |\lambda|,$$

tzn. $r(T)$ jest promieniem najmniejszego koła domkniętego w $\mathbb{C}$ zawierającego zbiór $\sigma(T)$.

Dowód: Niech A oznacza zbiór

$$A = \{z \in \mathbb{C} : \tfrac{1}{z} \in \sigma(T)\}$$

i niech $r_0 = \operatorname{dist}(0, A)$. Ustalmy dowolnie $\varepsilon > 0$ i wybierzmy $0 < r_1 < r_0 < r_2$ tak, aby $\frac{1}{r_1} - \frac{1}{r_2} < \varepsilon$. Z określenia r_0 wynika istnienie takiej liczby $\lambda_0 \in \sigma(T)$, że $\frac{1}{|\lambda_0|} \leqslant r_2$. Dla $x \in X$ i $x^* \in X^*$ niech f będzie funkcją określoną na $\mathbb{C} \setminus A$ wzorem

$$f(z) = x^*\big(R(\tfrac{1}{z}, T)\, x\big), \qquad z \neq 0,$$

oraz $f(0) = 0$. Ponieważ $\lim_{z\to 0} f(z) = 0 = f(0)$, więc funkcja f jest holomorficzna w $\mathbb{C}\setminus A$. Oznacza to, że można ją przedstawić w postaci szeregu potęgowego

$$f(z) = \sum_{n=1}^{\infty} a_n\, z^n,$$

którego promień zbieżności wynosi r_0. Wynika stąd w szczególności, że gdy $|z| = r_1$, to szereg ten jest zbieżny, więc

$$\sup_{n\geqslant 1} |a_n|\, r_1^n < \infty.$$

Zauważmy, że gdy $|z| < \frac{1}{\|T\|}$, to $R(\frac{1}{z}, T) = \sum\limits_{n=0}^{\infty} T^n\, z^{n+1}$, więc

$$a_n = x^*(T^n x).$$

Mamy zatem

$$\sup_n \big|x^*(r_1^n T^n x)\big| < \infty \quad \text{dla każdych} \quad x \in X,\ x^* \in X^*.$$

Z twierdzenia Banacha-Steinchausa wynika, że

$$\sup_n \|r_1^n T^n x\| < \infty,$$

a dalej, że

$$\sup_n \|r_1^n T^n\| < \infty$$

co oznacza, że $r_1\, r(T) \leqslant 1$ Możemy zatem napisać

$$r(T) \leqslant \frac{1}{r_1} < \frac{1}{r_2} + \varepsilon \leqslant |\lambda_0| + \varepsilon.$$

Widać stąd, że poza kołem o promieniu $r(T) - \varepsilon$ można wskazać punkt z widma $\sigma(T)$. To dowodzi tezy twierdzenia. $\square$

Operatory unitarne

Niech X i Y będą przestrzeniami unormowanymi. Operator $T \in \mathcal{L}(X,Y)$ nazywamy **izometrią**, gdy

$$\|Tx\| = \|x\|$$

dla wszystkich $x \in X$.

7.50. TWIERDZENIE. *Niech $\mathcal{H}$ będzie przestrzenią Hilberta. Dla operatora $T \in \mathcal{L}(\mathcal{H})$ następujące warunki są równoważne:*

i. *T jest izometrią,*

ii. $\langle Tx, Ty\rangle = \langle x, y\rangle$, $x, y \in \mathcal{H}$,

iii. $T^*T = I$.

Jeżeli T jest izometrią, to $T(\mathcal{H})$ jest domkniętą podprzestrzenią w $\mathcal{H}$. $\square$

7.51. PRZYKŁADY.

1. Na przestrzeni ℓ^2 operator przesunięcia

$$S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$$

jest izometrią. Operator sprzężony

$$S^*(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)$$

już izometrią nie jest.

2. Jeżeli f jest funkcją ciągłą, odwzorowującą przedział $[0,1]$ na $[0,1]$, to operator T, określony na przestrzeni $C[0,1]$ wzorem

$$Tx(t) = x\big(f(t)\big), \tag{7.42}$$

jest izometrią.

7.52. Zadanie. Czy każda izometria przestrzeni $C[0,1]$ ma postać (7.42)?

Definicja. Operator $U \in \mathcal{L}(\mathcal{H})$ nazywamy operatorem **unitarnym**, gdy jest izometrycznym izomorfizmem $\mathcal{H}$ na $\mathcal{H}$.

7.53. Twierdzenie. *Dla operatora $U \in \mathcal{L}(\mathcal{H})$ następujące warunki są równoważne:*

i. *U jest unitarny,*

ii. *U i U^* są izometriami,*

iii. *$U^{-1} = U^*$, tzn. $U^*U = UU^* = I$.* $\square$

Uwaga. Operator unitarny na przestrzeni Hilberta przeprowadza układy ortonormalne wektorów na układy ortonormalne. Dla dowolnych układów ortonormalnych zupełnych $\{e_\alpha\}_{\alpha\in A}$ i $\{f_\alpha\}_{\alpha\in A}$ istnieje dokładnie jeden operator unitarny U o tej własności, że $Ue_a = f_a$, $\alpha \in A$.

7.54. Przykład. Niech $f \in C(\mathbb{R})$. Operator $T \in \mathcal{L}\big(L^2(\mathbb{R})\big)$ określony wzorem

$$Tx(t) = f(t)\,x(t)$$

jest unitarny wtedy i tylko wtedy, gdy $|f(t)| \equiv 1$.

7.55. Twierdzenie (von Neumann). *Niech T będzie operatorem hermitowskim w $\mathcal{L}(\mathcal{H})$ i niech*

$$U_T = (T - iI)(T + iI)^{-1}.$$

Wtedy:

a. *U_T jest operatorem unitarnym,*

b. *operator $(I - U_T)^{-1}$ istnieje,*

c. *zachodzi wzór*

$$T = i\,(I + U_T)(I - U_T)^{-1}.$$

Operator U_T nazywamy **transformatą Kelley'a** *operatora T.*

Dowód: Zauważmy, że $U_T^* = (T + iI)(T - iI)^{-1}$. Ponieważ operatory $T + iI$ i $T - iI$ komutują, więc

$$U_T^* U_T = U_T U_T^* = I,$$

co dowodzi **a**. Punkt **b** wynika z równości

$$I - U_T = \big[(T + iI) - (T - iI)\big](T + iI)^{-1} = 2i\,(T + iI)^{-1},$$

zaś **c** z równości

$$i\,(I+U_T)(I-U_T)^{-1} = i\,\big[2T(T+iI)^{-1}\big]\big[-\tfrac{i}{2}\,(T+iI)\big] = T. \quad \square$$

7.56. Zadanie. Czy każdy operator unitarny jest transformatą Kelley'a pewnego operatora hermitowskiego?

Projektory

Niech X będzie przestrzenią Banacha. Operator $P \in \mathcal{L}(X)$ nazywamy **projektorem** lub **rzutem**, gdy $P^2 = P$.

7.57. Twierdzenie. *Jeżeli P jest projektorem w $\mathcal{L}(X)$, to istnieje jednoznaczny rozkład przestrzeni X*

$$X = X_1 \oplus X_2$$

na topologiczną sumę prostą dwóch podprzestrzeni domkniętych X_1 i X_2, przy czym $Px = x$ dla $x \in X_1$ oraz $Px = 0$ dla $x \in X_2$.

Dowód: Oznaczmy

$$X_1 = \{x \in X : Px = x\}, \qquad X_2 = \{x \in X : Px = 0\}.$$

Ponieważ $X_1 = \ker(I-P)$, $X_2 = \ker P$, oba zbiory $X_!$, X_2 są domkniętymi podprzestrzeniami liniowymi w X. Oczywiście $X_1 \cap X_2 = \{0\}$.

Dla dowolnego wektora $x \in X$ połóżmy $x_1 = Px$, $x_2 = x - Px$. Wtedy $x_1 \in X_1$, $x_2 \in X_2$. Istotnie,

$$\begin{gathered} Px_1 = P^2x = Px = x_1, \\ Px_2 = P(x-Px) = Px - P^2x = Px - Px = 0. \end{gathered}$$

Ponadto $\|x_1\| \leqslant \|P\|\,\|x\|$, $\|x_2\| \leqslant \|I-P\|\,\|x\|$, a ponieważ $x = x_1 + x_2$ i przedstawienie to jest jednoznaczne (bo $X_1 \cap X_2 = \{0\}$), mamy $X = X_1 \oplus X_2$. $\square$

Uwaga. O projektorze P mówimy, że jest rzutem na podprzestrzeń X_1 wzdłuż podprzestrzeni X_2.

7.58. Wniosek. *Na to, by przestrzeń Banacha była topologiczną sumą prostą podprzestrzeni X_i, $i = 1,2,3,\ldots,n$, potrzeba i wystarcza, aby w przestrzeni X istniały takie projektory P_i, że $P_iP_j = 0$, gdy $i \neq j$, $I = P_1 + P_2 + P_3 + \ldots + P_n$ oraz $P_i(X) = X_i$ dla $i = 1,2,3,\ldots,n$.* $\square$

DEFINICJA. Projektor P na przestrzeni Hilberta nazywamy **projektorem ortogonalnym** lub **rzutem ortogonalnym**, jeżeli odpowiadające mu podprzestrzenie X_1, X_2 z poprzedniego twierdzenia są ortogonalne.

7.59. TWIERDZENIE. *Dla rzutu P na przestrzeni Hilberta następujące warunki są równoważne:*

- **i.** *P jest projektorem ortogonalnym,*
- **ii.** $P^* = P$,
- **iii.** $\|P\| \leqslant 1$, *tzn. P jest kontrakcją.*

Dowód: i$\Rightarrow$ii. Załóżmy, że P jest projektorem ortogonalnym w $\mathcal{H}$. Dla $x, y \in \mathcal{H}$ oznaczmy $x_1 = Px$, $x_2 = x - Px$, $y_1 = Py$, $y_2 = y - Py$. Z założenia $x_1 \perp y_2$ oraz $x_2 \perp y_1$, a więc

$$\begin{aligned}\langle x_1, y\rangle &= \langle x_1, y_1\rangle + \langle x_1, y_2\rangle = \langle x_1, y_1\rangle,\\ \langle x, y_1\rangle &= \langle x_1, y_1\rangle + \langle x_2, y_1\rangle = \langle x_1, y_1\rangle.\end{aligned}$$

Widzimy zatem, że

$$\langle Px, y\rangle = \langle Px, Py\rangle = \langle x, Py\rangle$$

dla dowolnych $x, y \in \mathcal{H}$, czyli $P^* = P$.

ii$\Rightarrow$iii. Dla dowolnego $x \in \mathcal{H}$ mamy

$$\begin{aligned}\|Px\|^2 &= \langle Px, Px\rangle = \langle P^*Px, x\rangle = \langle P^2x, x\rangle\\ &= \langle Px, x\rangle,\end{aligned}$$

co z nierówności Schwarza daje $\|Px\|^2 \leqslant \|Px\|\,\|x\|$, i w rezultacie

$$\|Px\| \leqslant \|x\|.$$

iii$\Rightarrow$i. Niech $\mathcal{H}_1 = \{x \in \mathcal{H} : Px = x\}$, $\mathcal{H}_2 = \{x \in \mathcal{H} : Px = 0\}$. Mamy pokazać, że $\mathcal{H}_1 \perp \mathcal{H}_2$ tzn., że dla dowolnych $x_1 \in \mathcal{H}_1$, $x_2 \in \mathcal{H}_2$ zachodzi $\langle x_1, x_2\rangle = 0$. Jest tak, gdy $x_2 = 0$. Jeżeli $x_2 \neq 0$, to dla wektora $x = x_1 - \lambda\, x_2$ otrzymamy

$$Px = Px_1 - \lambda\, Px_2 = x_1,$$

zatem

$$\|x_1\| = \|Px\| \leqslant \|P\|\,\|x\| \leqslant \|x\|,$$

co oznacza, że nierówność

$$0 \leqslant \|x\|^2 - \|x_1\|^2 = -\overline{\lambda}\langle x_1, x_2\rangle - \lambda\overline{\langle x_1, x_2\rangle} + |\lambda|^2\|x_2\|^2$$

jest spełniona dla dowolnej liczby λ. Po podstawieniu $\lambda = \frac{\langle x_1, x_2\rangle}{\|x_2\|^2}$ otrzymamy

$$\frac{-\left|\langle x_1, x_2\rangle\right|^2}{\|x_2\|^2} \geqslant 0,$$

a stąd $\langle x_1, x_2\rangle = 0$. $\square$

7.60. Przykład. W przestrzeni $\mathbb{R}^2$ rozpatrzmy rzut P na prostą X_1 wzdłuż prostej X_2, tworzących ze sobą kąt α. Jeżeli wybierzemy wektor x tak jak na rysunku, to

$$\|Px\| = \frac{1}{\sin\alpha}\,\|x\| \geqslant \|x\|.$$

Równość zachodzi dokładnie, gdy $\alpha = \frac{\pi}{2}$.

7.61. Twierdzenie. *Dla każdej domkniętej podprzestrzeni liniowej $\mathcal{H}_1 \subset \mathcal{H}$ istnieje dokładnie jeden rzut ortogonalny P na $\mathcal{H}_1$. Jeżeli $x \in \mathcal{H}$, to Px realizuje minimum odległości wektorów z podprzestrzeni $\mathcal{H}_1$ od x.*

Dowód: Niech x_1 będzie dowolnym wektorem w $\mathcal{H}_1$. Wtedy

$$\|x - x_1\|^2 = \|x - Px + Px - x_1\|^2 = \|x - Px\|^2 + \|Px - x_1\|^2,$$

bo wektory $x - Px$ i $Px - x_1$ są do siebie ortogonalne. Wyrażenie po prawej stronie równości przyjmuje najmniejszą wartość, gdy $\|Px - x_1\| = 0$, tzn. gdy $x_1 = Px$. $\square$

Operatory zwarte

Niech X i Y będą przestrzeniami Banacha oraz $K = \{x \in X : \|x\| \leqslant 1\}$ kulą jednostkową w przestrzeni X. Operator $T \in \mathcal{L}(X, Y)$ nazywamy **zwartym**, jeżeli obraz $T(K)$ kuli K jest zbiorem warunkowo zwartym w Y. Zauważmy, że operator T jest zwarty wtedy i tylko wtedy, gdy dla każdego ograniczonego ciągu $\{x_n\}$ w X, ciąg $\{Tx_n\}$ zawiera podciąg zbieżny.

7.62. Przykład. Z twierdzenia Arzeli **??** wynika, że operator identycznościowy z przestrzeni $C^1[a, b]$ w przestrzeń $C[a, b]$ jest zwarty.

7.63. TWIERDZENIE. *Operator zwarty $T \in \mathcal{L}(X,Y)$ przeprowadza ciągi słabo zbieżne w ciągi mocno zbieżne. Jeżeli X jest przestrzenią refleksywną, to prawdziwe jest też twierdzenie odwrotne.*

Dowód: Załóżmy, że $\{x_n\}$ jest ciągiem w X, słabo zbieżnym do x. Z ciągłości operatora T wynika, że ciąg $\{Tx_n\}$ jest wtedy słabo zbieżny w Y do Tx. Pokażemy, że dla każdego $\varepsilon > 0$ zbiór

$$\{n \in \mathbb{N} : \|Tx_n - Tx\| \geqslant \varepsilon\}$$

jest skończony. Gdyby zbiór ten był nieskończony, to na mocy ograniczoności ciągu $\{x_n\}$ i zwartości operatora T, moglibyśmy z ciągu $\{Tx_n\}$ wybrać taki podciąg $\{Tx_{n_k}\}$, który byłby mocno zbieżny do pewnego wektora y i jednocześnie spełniał nierówność $\|Tx_{n_k} - Tx\| \geqslant \varepsilon$. Ponieważ zbieżność mocna pociąga zbieżność słabą, otrzymalibyśmy $y = Tx$, a to przecież jest niemożliwe.

W przestrzeni refleksywnej z każdego ciągu ograniczonego można wybrać podciąg słabo zbieżny (patrz 4.44). $\square$

7.64. WNIOSEK. *Każdy ograniczony operator liniowy z przestrzeni ℓ^p, $1 < p < \infty$, do przestrzeni ℓ^1 jest zwarty.*

Dowód: Twierdzenie Schura 4.38 mówi, że w przestrzeni ℓ^1 pojęcia słabej zbieżności ciągów i mocnej zbieżności pokrywają się. Teza wynika więc z drugiej części poprzedniego twierdzenia. $\square$

UWAGA. W skończenie wymiarowej przestrzeni unormowanej każdy zbiór ograniczony jest warunkowo zwarty. Wynika stąd, że jeżeli $T \in \mathcal{L}(X,Y)$ jest operatorem skończenie wymiarowym, tzn. $\dim T(X) < \infty$, to jest zwarty.

7.65. ZADANIE. Wykazać, że operator $T \in \mathcal{L}(X,Y)$ jest skończenie wymiarowy wtedy i tylko wtedy, gdy dla pewnego n istnieją takie elementy $x_1^*, x_2^*, \ldots, x_n^*$ w X^* oraz $y_1, y_2, \ldots, y_n$ w Y, że

$$Tx = \sum_{k=1}^{n} x_k^*(x)\, y_k.$$

7.66. PRZYKŁAD. Niech $\lambda_1, \lambda_2, \lambda_3, \ldots$ będzie ograniczonym ciągiem liczb zespolonych. Na przestrzeni ℓ^2 określmy operator T wzorem

$$T(x_1, x_2, x_3, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \lambda_3 x_3, \ldots).$$

Pokażemy, że T jest operatorem zwartym wtedy i tylko wtedy, gdy $\lambda_n \to 0$.

Załóżmy, że $\lambda_n \to 0$. Mamy pokazać, że dla każdego $\varepsilon > 0$ dla obrazu $T(K)$ kuli jednostkowej K istnieje skończona ε-sieć. Najpierw dobierzmy tak wskaźnik m, aby $|\lambda_k| < \frac{\varepsilon}{2}$, gdy $k > m$ a następnie w zbiorze $\{(t_1, t_2, \ldots, t_m, 0, 0, \ldots) : \sum_{k=1}^{m} |t_k|^2 \leqslant 1\}$ wybierzmy skończoną $\frac{\varepsilon}{2\|T\|}$-sieć $\{x_1, x_2, \ldots, x_n\}$. Wtedy zbiór $\{Tx_1, Tx_2, \ldots, Tx_n\}$ jest ε-siecią dla $T(K)$.

Warunek $\lambda_n \to 0$ jest konieczny. Istotnie, wiemy (patrz 4.34), że ciąg $\{e_n\}$ wektorów bazy standartowej przestrzeni ℓ^2, $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$ z 1 na miejscu n-tym, jest słabo zbieżny do zera a ciąg $Te_n = \lambda_n x_n$ jest mocno zbieżny (do zera) tylko wtedy, gdy $\lambda_n \to 0$.

Uwaga. Można sprawdzić, że wszystkie operatory z powyższego przykładu są normalne. Pokażemy potem **??**, że na przestrzeni Hilberta nie ma innych operatorów zwartych normalnych niż te, opisane w Przykładzie 7.66.

7.67. Twierdzenie. *Granica zbieżnego ciągu operatorów zwartych jest także operatorem zwartym. Operatory zwarte tworzą zatem w $\mathcal{L}(X, Y)$ domkniętą podprzestrzeń liniową.*

Dowód: Załóżmy, że $T_1, T_2, T_3, \ldots$ są operatorami zwartymi i że $T_n \to T$ w $\mathcal{L}(X, Y)$. Niech K oznacza kulę jednostkową w X. Jeżeli $\|T - T_n\| \leqslant \frac{\varepsilon}{2}$ oraz $\{y_1, y_2, \ldots, y_m\}$ jest $\frac{\varepsilon}{2}$-siecią dla zbioru $T_n(K)$, to jest też ε-siecią dla zbioru $T(K)$. $\square$

7.68. Twierdzenie. *Niech $T \in \mathcal{L}(X, Y)$, $S \in \mathcal{L}(Y, Z)$. Jeżeli przynajmniej jeden z operatorów T, S jest zwarty, to zwarty jest także operator ST.*

7.69. Wniosek. *Operator $T \in \mathcal{L}(\mathcal{H})$ jest zwarty wtedy i tylko wtedy, gdy operator sprzężony T^* jest zwarty.*

Dowód: Zauważmy najpierw, że $\{T^* x_n\}$ jest ciągiem Cauchy'ego wtedy i tylko wtedy, gdy $\{TT^* x_n\}$ jest ciągiem Cauchy'ego. Implikacja w jedną stronę jest oczywista a w drugą wynika z nierówności

$$\|T^* x_n - T^* x_m\|^2 = \langle TT^* x_n - TT^* x_m,\, x_n - x_m \rangle \leqslant 2M\, \|TT^* x_n - TT^* x_m\|.$$

Załóżmy teraz, że $\{x_n\}$ jest dowolnym ciągiem ograniczonym w $\mathcal{H}$. Jeżeli T jest operatorem zwartym, to ciąg $\{TT^* x_n\}$ zawiera podciąg zbieżny, więc ciąg $\{T^* x_n\}$ również zawiera podciąg zbieżny. $\square$

Zobaczymy za chwilę, że prawdziwe jest ogólniejsze twierdzenie.

Definicja. Dla operatora $T \in \mathcal{L}(X, Y)$ określmy operator $T^* : Y^* \to X^*$ kładąc

$$(T^*y^*)x = y^*(Tx).$$

Jest jasne, że $T^* \in \mathcal{L}(Y^*, X^*)$ i $\|T^*\| = \|T\|$. Operator ten nosi nazwę operatora **sprzężonego** do T.

Następuje tu pewna kolizja oznaczeń, operator sprzężony był już definiowany wcześniej dla operatora na przestrzeni Hilberta i był znów operatorem na tej samej przestrzeni. Z twierdzenia 4.10 o postaci ciągłych funkcjonałów liniowych na przestrzeni Hilberta wynika, że możemy utożsamić $\mathcal{H}$ i $\mathcal{H}^*$, mianowicie wektorowi y przyporządkować funkcjonał $x \to < x, y >$. Należy pamiętać jednak, że nie jest to operacja liniowa, mnożeniu wektorów przez skalary odpowiada mnożenie funkcjonałów przez skalary sprzężone. Należy więc stale deklarować w jakim sensie T^* jest operatorem sprzężonym to T. W istocie różnica obu pojęć nie jest aż tak istotna. Dla przykładu, jeśli $T \in \mathcal{L}(\ell^2)$ jest operatorem mnożenia przez ciąg ograniczony $\{\lambda_n\}$, to operator sprzężony T^* w sensie powyższej definicji pokrywa się z T, a w zwykłym sensie jest operatorem mnożenia przez ciąg $\{\overline{\lambda_n}\}$.

7.70. Twierdzenie (Schauder). *Operator $T \in \mathcal{L}(X, Y)$ jest zwarty wtedy i tylko wtedy, gdy zwarty jest operator T^*.*

Dowód: Załóżmy najpierw, że T jest operatorem zwartym. Niech K oznacza kulę jednostkową w przestrzeni X a S domknięcie w Y zbioru $T(K)$. Jeżeli $\{y_n^*\}$ jest ciągiem ograniczonym w Y^*, to jego obcięcia do S tworzą jednakowo ciągłą i ograniczoną rodzinę funkcji na zbiorze zwartym S. Z twierdzenia Arzeli wynika więc istnienie podciągu $\{y_{n_k}^*\}$ jednostajnie zbieżnego na S, czyli zbieżność podciągu $\{T^*y_{n_k}^*\}$ w X^*

Jeżeli T^* jest operatorem zwartym, to zwarty jest także operator $T^{**} : X^{**} \to Y^{**}$. Operator T jest izometryczny z obcięciem T^{**} do domkniętej podprzestrzeni $\tau(X) \subset X^{**}$ (patrz (4.23)), jest więc też zwarty. $\square$

7.71. Przykład. Operator Volterry $V \in \mathcal{L}\big(C[0,1]\big)$

$$Vx(t) = \int_0^t x(u)\,du$$

jest zwarty. Wynika to natychmiast z twierdzenia Arzeli. Zobaczmy, czy jest on zwarty także jako operator z $L^p(0,1)$ do $L^p(0,1)$, $1 \leqslant p \leqslant \infty$.

Jeżeli $1 < p \leqslant \infty$, to z nierówności Höldera

$$\big|Vx(t_1) - Vx(t_2)\big| \leqslant \int_{t_1}^{t_2} \big|x(s)\big|\,ds \leqslant |t_1 - t_2|^{\frac{p-1}{p}} \|x\|_p.$$

Wynika stąd, że jeżeli $\{x_n\}$ jest ciągiem ograniczonym w $L^p(0,1)$, to $\{Vx_n\}$ jest ograniczonym ciągiem funkcji jednakowo ciągłych. Istnieje więc podciąg $\{Vx_{n_k}\}$ zbieżny jednostajnie na $[0,1]$, a więc zbieżny także w $L^p(0,1)$.

Dla $p=1$ mamy

$$V^*x(t) = \int_t^1 x(s)\,ds,$$

więc można powołać się na twierdzenie Schaudera i dokładnie powtórzyć poprzednie rozumowanie dla operatora V^* na $L^\infty(0,1)$.

Definicja. Ciąg $T_n \in \mathcal{L}(X,Y)$ nazywamy **mocno zbieżnym**, gdy dla każdego $x \in X$ ciąg $\{T_n x\}$ jest zbieżny w Y. Z twierdzenia Banacha-Steinhausa wynika, że $\sup_n \|T_n\| < \infty$. Zatem operator $T : X \to Y$, określony wzorem

$$Tx = \lim_{n\to\infty} T_n x, \qquad x \in X,$$

leży w $\mathcal{L}(X,Y)$. Ponadto dla każdego ciągu zbieżnego $x_n \to x$ w X ciąg $\{T_n x_n\}$ jest zbieżny do Tx. Mamy bowiem

$$\|T_n x_n - Tx\| \leqslant M\,\|x_n - x\| + \|T_n x - Tx\|,$$

gdzie $M = \sup_n \|T_n\|$, a oba składniki po prawej stronie dążą do zera gdy $n \to \infty$.

7.72. Twierdzenie. *Niech X, Y, Z będą przestrzeniami Banacha. Jeżeli operator $S \in \mathcal{L}(X,Y)$ jest zwarty a ciąg operatorów $T_n \in \mathcal{L}(Y,Z)$ jest mocno zbieżny, to ciąg $\{T_n S\}$ jest zbieżny w normie.*

Dowód: Załóżmy niewprost, że $\|T_n S - TS\| \geqslant \varepsilon > 0$ dla $n = 1,2,3,\ldots$. Niech $\{x_n\}$ będzie takim ciągiem wektorów z kuli jednostkowej przestrzeni X, że

$$\|T_n S x_n - TSx_n\| \geqslant \frac{\varepsilon}{2}. \tag{7.43}$$

Ponieważ istnieje podciąg zbieżny $\{Sx_{n_k}\}$ ciągu $\{Sx_n\}$, więc $\|T_{n_k} S x_{n_k} - TSx_{n_k}\|$ dąży do 0, a to przeczy nierówności (7.43). $\square$

Uwaga. Odpowiednik tego twierdzenia nie zachodzi dla ciągu $\{ST_n\}$. Jeżeli dla przykładu przyjmiemy $T_n = T^n$, gdzie T jest operatorem przesunięcia w ℓ^2

$$T(t_1, t_2, t_3, \ldots) = (t_2, t_3, t_4, \ldots)$$

a S rzutem na pierwszą oś

$$S(t_1, t_2, t_3, \ldots) = (t_1, 0, 0, \ldots),$$

to S jest zwarty a $T_n \to 0$ mocno. Ponieważ $\|ST_n\| = 1$, więc $\{ST_n\}$ nie dąży do zera w normie. Nawet więcej $\|ST_n - ST_m\| = \sqrt{2}$, gdy $m \neq n$.

7.73. WNIOSEK. *Jeżeli przestrzeń Y ma bazę topologiczną, to każdy operator zwarty $T \in \mathcal{L}(X, Y)$ jest granicą (w normie) ciągu operatorów skończenie wymiarowych.*

Dowód: Jeżeli $y_1, y_2, y_3, \ldots$ jest bazą topologiczną w Y, oraz

$$P_n\Big(\sum_{k=1}^{\infty} \lambda_k\, y_k\Big) = \sum_{k=1}^{n} \lambda_k\, y_k, \qquad n = 1, 2, 3, \ldots,$$

to $\{P_n\}$ jest ciągiem projektorów skończenie wymiarowych, zbieżnym mocno do operatora identycznościowego na Y, więc z twierdzenia 7.72 ciąg $\{P_n T\}$ jest zbieżny w normie do T. $\square$

UWAGA. Przypomnijmy, że nie każda ośrodkowa przestrzeń Banacha ma bazę topologiczną. Zauważmy jednak, że na to, by teza Wniosku 7.73 była prawdziwa wystarczy mniej, mianowicie, by istniał ciąg operatorów skończenie wymiarowych zbieżny mocno do identyczności. Przestrzeń o tej własności nosi nazwę przestrzeni Banacha z **z własnością aproksymacji**. Musi ona być ośrodkowa, ale nie każda przestrzeń ośrodkowa jest taka.

7.74. WNIOSEK. *Operator T w ośrodkowej przestrzeni Hilberta jest zwarty wtedy i tylko wtedy, gdy jest granicą zbieżnego w normie, ciągu operatorów skończenie wymiarowych.*

7.75. PRZYKŁAD. Niech $e_1, e_2, e_3, \ldots$ będzie bazą ortonormalną w przestrzeni Hilberta $\mathcal{H}$. Jeżeli operator $T \in \mathcal{L}(\mathcal{H})$ spełnia warunek

$$\sum_{k=1}^{\infty} \|Te_k\|^2 < \infty,$$

to jest zwarty. Rzeczywiście, połóżmy

$$T_n = \sum_{k=1}^{n} \langle x,\, e_k\rangle\, Te_k, \qquad x \in \mathcal{H}, \quad n = 1, 2, 3, \ldots,$$

wtedy z nierówności Schwarza dostaniemy

$$\|T_n x - Tx\| \leqslant \sum_{k=n+1}^{\infty} \big|\langle x,\, e_k\rangle\big|\, \|Te_k\| \leqslant \|x\| \Big(\sum_{k=n+1}^{\infty} \|Te_k\|^2\Big)^{1/2}$$

a stąd $\|T_n - T\| \to 0$. $\square$

Rozkład spektralny operatora zwartego

Przejdziemy teraz do badania własności spektralnych operatorów zwartych.

7.76. TWIERDZENIE. *Niech X, Y będą przestrzeniami Banacha i niech $S, T \in \mathcal{L}(X, Y)$. Jeżeli T jest operatorem zwartym a S operatorem odwracalnym, to obraz*

$$\operatorname{Im}(T+S) = (T+S)(X)$$

jest domkniętą podprzestrzenią liniową w Y.

Dla dowodu potrzebne nam będą następujące lematy:

7.77. LEMAT. *Jeżeli $\{x_n\}$ jest ciągiem ograniczonym w X a ciąg $\{(T+S)x_n\}$ jest zbieżny, to $\{x_n\}$ zawiera podciąg zbieżny.*

Dowód: Niech $y = \lim_{n\to\infty}(T+S)x_n$. Ponieważ T jest operatorem zwartym a $\{x_n\}$ ciągiem ograniczonym, więc z ciągu $\{Tx_n\}$ można wybrać podciąg zbieżny $\{Tx_{n_k}\}$ do pewnego elementu $y_0 \in Y$. Wtedy podciąg $\{Sx_{n_k}\}$ dąży do $y - y_0$ a ponieważ S jest izomorfizmem X na Y, więc sam ciąg $\{x_{n_k}\}$ jest zbieżny. $\square$

7.78. LEMAT. *Jeżeli $y \in \overline{(T+S)(X)}$, to istnieje taki ograniczony ciąg $\{x_n\}$ w X, że $(T+S)x_n \to y$.*

Dowód: Dla pewnego ciągu $\{x_n\}$ w przestrzeni X zachodzi $(T+S)x_n \to y$. Oznaczmy $X_0 = \ker(T+S)$ oraz $\alpha_n = \operatorname{dist}(x_n, X_0)$. W podprzestrzeni X_0 wybierzmy taki ciąg $\{x'_n\}$, że $\|x_n - x'_n\| \leqslant 2\alpha_n$. Jeżeli $\{\alpha_n\}$ jest ciągiem ograniczonym, to ciąg $\{x_n - x'_n\}$ ma żądaną własność

$$(T+S)(x_n - x'_n) = (T+S)x_n \to y.$$

Pozostaje zatem pokazać, że ciąg $\{\alpha_n\}$ jest ograniczony. Gdyby $\alpha_n \to \infty$, to kładąc $z_n = \frac{x_n - x'_n}{\|x_n - x'_n\|}$ otrzymalibyśmy $\|z_n\| = 1$, $\operatorname{dist}(z_n, X_0) = \frac{\alpha_n}{\|x_n - x'_n\|} \geqslant \frac{1}{2}$ oraz $(T+S)z_n \to 0$. Z lematu 7.77 wynika, że ciąg $\{z_n\}$ miałby podciąg $\{z_{n_k}\}$ zbieżny do pewnego z_0. Ponieważ $(T+S)z_0 = \lim_{k\to\infty}(T+S)x_{n_k} = 0$, więc otrzymalibyśmy $z_0 \in X_0$, to jednak nie jest możliwe, gdyż $\operatorname{dist}(z_0, X_0) = \lim_{k\to\infty}\operatorname{dist}(z_{n_k}, X_0) \geqslant \frac{1}{2}$. $\square$

Dowód Twierdzenia: Jeżeli $y \in \overline{(T+S)(X)}$, to z lematu 7.78 wynika istnienie takiego ciągu ograniczonego $\{x_n\}$ w X, że $(T+S)x_n \to y$. Dzięki lematowi 7.77 możemy założyć, że ciąg $\{x_n\}$ jest zbieżny do pewnego $x \in X$. Wtedy $y = (T+S)x$. $\square$

7.79. WNIOSEK. *Niech T będzie operatorem zwartym w przestrzeni Banacha X i niech $\lambda \in \sigma(T)$, $\lambda \neq 0$. Wtedy λ jest wartością własną operatora T lub operatora sprzężonego T^*.*

Dowód: Jeżeli λ nie jest wartością własną operatora T, to $X_0 = (\lambda I - T)(X)$ jest właściwą podprzestrzenią w X. Z twierdzenia 7.76 wynika, że X_0 jest domknięta w X. Jeżeli $0 \neq x^* \in X^*$ i $x^*(X_0) = \{0\}$, to $(\lambda I - T)^*(X^*) = \{0\}$. $\square$

UWAGA. Pokażemy potem, że we Wniosku 7.79 słowo „lub" w tezie można zastąpić słowem „i".

7.80. WNIOSEK. *Jeżeli T jest operatorem zwartym i normalnym w przestrzeni Hilberta, to*

$$\sigma(T) \subset \sigma_p(T) \cup \{0\}.$$

Dowód: Pamiętamy, że dla operatora normalnego $\lambda \in \sigma_p(T)$ wtedy i tylko wtedy, gdy $\overline{\lambda} \in \sigma_p(T^*)$. Teza wynika więc z Wniosku 7.79. $\square$

7.81. TWIERDZENIE. *Niech T będzie operatorem zwartym na przestrzeni Hilberta $\mathcal{H}$ i niech $\lambda \neq 0$. Operator $\lambda I - T$ jest „na" wtedy i tylko wtedy, gdy jest różnowartościowy.*

Dowód: Załóżmy niewprost, że operator $S = \lambda I - T$ odwzorowuje $\mathcal{H}$ na $\mathcal{H}$, a mimo to $\ker S \neq \{0\}$. Utwórzmy w $\mathcal{H}$ ciąg domkniętych podprzestrzeni liniowych $\mathcal{H}_0 = \{0\}$, $\mathcal{H}_n = \ker S^n$, $n = 1, 2, 3, \ldots$. Wtedy $\mathcal{H}_0 \subset \mathcal{H}_1 \subset \mathcal{H}_2 \subset \ldots$ i każda z inkluzji jest właściwa. Istotnie, niech x_1 będzie niezerowym wektorem w przestrzeni $\mathcal{H}_1 = \ker S$. Ponieważ $S(X) = X$, więc istnieją takie wektory $x_2, x_3, x_4, \ldots$, że $Sx_n = x_{n-1}$. Mamy wtedy $S^n x_n = 0$ oraz $S^{n-1}x_n = x_1 \neq 0$ a w konsekwencji $\mathcal{H}_n \neq \mathcal{H}_{n-1}$. W każdej z przestrzeni $\mathcal{H}_n$ wybierzmy tak wektor z_n, by $\|z_n\| = 1$ oraz $z_n \perp \mathcal{H}_{n-1}$. Pokażemy, że ciąg $\{Tz_n\}$ nie ma podciągów zbieżnych, choć T jest operatorem zwartym. Da nam to sprzeczność z założeniem $\ker S \neq \{0\}$.

Dla $n > m$ mamy

$$\|Tz_n - Tz_m\| = \|\lambda z_n - z\|,$$

gdzie $z = \lambda z_m - Sz_m + Sz_n$. Ponieważ $S^{n-1}z = \lambda S^{n-1}z_m - S^n z_m + S^n z_n = 0$, więc $z \in X^{n-1}$. W rezultacie

$$\|Tz_n - Tz_m\| \geqslant |\lambda|.$$

Dla dowodu w drugą stronę załóżmy, że $\ker S = \{0\}$. Wtedy $S^*(\mathcal{H})^\perp = \ker S = \{0\}$, czyli $S^*(\mathcal{H})$ jest gęstą podprzestrzenią liniową w $\mathcal{H}$. Z twierdzenia 7.76 wiadomo, że jest podprzestrzenią domkniętą. Musi więc być $S^*(\mathcal{H}) = \mathcal{H}$. Powtarzając rozumowanie z pierwszej części dowodu dla operatora S^*, otrzymamy $\ker S^* = \{0\}$, a stąd, tak jak przed chwilą $S(\mathcal{H}) = \mathcal{H}$. $\square$

Uwaga. Twierdzenie to jest prawdziwe dla operatorów zwartych na dowolnych przestrzeniach Banacha. Dowód jest jednak bardziej skomplikowany.

7.82. Wniosek. *Jeżeli T jest operatorem zwartym (na przestrzeni Hilberta), to liczba $\lambda \neq 0$ jest jego wartością własną operatora T wtedy i tylko wtedy, gdy $\overline{\lambda}$ jest wartością własną operatora T^*.*

7.83. Wniosek. *Dla operatora zwartego T (na przestrzeni Hilberta) mamy*

$$\sigma(T) \subset \sigma_p(T) \cup \{0\}$$

a jedynym możliwym punktem skupienia zbioru $\sigma(T)$ jest 0. Zatem zbiór $\sigma(T)$ jest co najwyżej przeliczalny.

Dowód: Pierwsza część wniosku jest oczywista. Gdyby istniał taki ciąg liczb $\{\lambda_n\}$, że $\lambda_n \neq \lambda_m$, gdy $n \neq m$, $\lambda_n \to \lambda_0 \neq 0$ oraz $\lambda_n \in \sigma(T)$, to moglibyśmy dobrać taki ciąg niezerowych wektorów x_n, że $Tx_n = \lambda_n x_n$, $n = 1, 2, 3, \ldots$. Wektory takie muszą być liniowo niezależne. Wobec tego dla przestrzeni $\mathcal{H}_n = \operatorname{lin}\{x_1, x_2, \ldots, x_n\}$ wszystkie inkluzje $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{H}_3 \subset \ldots$ są właściwe. Jeśli teraz, podobnie jak poprzednio, w każdej z przestrzeni $\mathcal{H}_n$ wybierzemy taki wektor z_n, by $\|z_n\| = 1$ oraz $z_n \perp \mathcal{H}_{n-1}$, to ciąg $\{Tz_n\}$ nie będzie zawierał żadnego podciągu zbieżnego, a to przeczy założonej zwartości operatora T. Istotnie, ponieważ dla pewnej liczby α_n wektor $z_n - \alpha_n x_n$ leży w $\mathcal{H}_{n-1}$, więc $Tz_n - \lambda_n \alpha_n x_n \in \mathcal{H}_{n-1}$, a stąd $Tz_n - \lambda_n z_n \in \mathcal{H}_{n-1}$. Jeżeli $n > m$, to także $Tz_m \in \mathcal{H}_{n-1}$, więc $Tz_n - Tz_m = \lambda_n z_n + y$ dla pewnego $y \in \mathcal{H}_{n-1}$. Stąd natychmiast wynika, że $\|Tz_n - Tz_m\| \geqslant |\lambda_n|$, a więc ciąg $\{Tz_n\}$ nie ma podciągów zbieżnych. □

Twierdzenie spektralne dla operatorów zwartych

7.84. Lemat. *Wektory własne, odpowiadające różnym wartościom własnym operatora normalnego są do siebie ortogonalne.*

Dowód: Jeśli $Tx_1 = \lambda_1 x_1$, $Tx_2 = \lambda_2 x_2$ i $\lambda_1 \neq \lambda_2$, to

$$\begin{aligned}\lambda_1 \langle x_1, x_2 \rangle &= \langle \lambda_1 x_1, x_2 \rangle = \langle Tx_1, x_2 \rangle = \langle x_1, T^* x_2 \rangle \\ &= \langle x_1, \overline{\lambda_2} x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle,\end{aligned}$$

więc $(\lambda_1 - \lambda_2)\langle x_1, x_2 \rangle = 0$, a stąd $\langle x_1, x_2 \rangle = 0$. Wykorzystaliśmy tu wcześniej udowodniony fakt (patrz **??**), że $Tx = \lambda x$ wtedy i tylko wtedy, gdy $T^* x = \overline{\lambda} x$. □

Definicja. Jeżeli $\lambda \in \sigma_p(T)$, to zbiór $X_\lambda = \{x \in X : Tx = \lambda x\}$ jest domkniętą podprzestrzenią liniową w X, zwaną **podprzestrzenią własną** operatora T, odpowiadającą wartości własnej λ.

7.85. LEMAT. *Podprzestrzenie własne operatora zwartego są skończenie wymiarowe.* $\square$

7.86. TWIERDZENIE (O ROZKŁADZIE SPEKTRALNYM OPERATORÓW ZWARTYCH). *Niech T będzie operatorem zwartym normalnym na przestrzeni Hilberta $\mathcal{H}$. Wtedy w $\mathcal{H}$ istnieje taki układ ortonormalny $e_1, e_2, e_3, \ldots$ (być może skończony) i taki układ liczb zespolonych $\lambda_1, \lambda_2, \lambda_3, \ldots$, że $|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3| \geqslant \ldots$ oraz*

1. $Tx = \sum\limits_{n=1}^{\infty} \lambda_k \langle x, e_k \rangle e_k$ *dla* $x \in \mathcal{H}$,
2. $\sigma(T) = \{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ *lub* $\sigma(T) = \{\lambda_1, \lambda_2, \lambda_3, \ldots\} \cup \{0\}$, *gdy* 0 *nie jest wartością własną a przestrzeń $\mathcal{H}$ jest nieskończenie wymiarowa.*

Dowód: Ustawmy wszystkie różne od zera wartości własne $\mu_1, \mu_2, \mu_3, \ldots$ operatora T w taki ciąg, aby $|\mu_1| \geqslant |\mu_2| \geqslant |\mu_3| \geqslant \ldots$. Jest to możliwe, bo $\sigma(T)$ jest zbiorem najwyżej przeliczalnym, a jedynym możliwym jego punktem skupienia jest zero. W każdej (skończenie wymiarowej z lematu 7.85) podprzestrzeni własnej $\mathcal{H}_{\mu_n}$ operatora T, odpowiadającej wartości własnej μ_n, wybierzmy w dowolny sposób bazę ortonormalną B_{μ_n} i niech $B = \bigcup_{n=1}^{\infty} B_{\mu_n}$. Z lematu 7.84 wynika, że B jest zbiorem przeliczalnym ortonormalnym. Ustawmy elementy zbioru B w ciąg $e_1, e_2, e_3, \ldots$. Jest jasne, że są to wektory własne operatora T, a więc dla każdego wektora x z podprzestrzeni $\mathcal{H}_0 = \overline{\operatorname{lin} B}$ mamy

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

przy czym każda z liczb λ_n jest jedną z liczb $\mu_1, \mu_2, \mu_3, \ldots$, jest więc wartością własną operatora T i w powyższym rozkładzie występuje tyle razy, ile wynosi wymiar podprzestrzeni $\mathcal{H}_{\lambda_n}$. Także $|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3| \geqslant \ldots$.

Pokażemy, że dla $x \in X^{\perp}$ zachodzi $Tx = 0$. Ponieważ $T(X) \subset X$ oraz $T^*(X) \subset X$, bo

$$T^*x = \sum_{n=1}^{\infty} \overline{\lambda_n} \langle x, e_n \rangle e_n,$$

więc oba operatory T, T^* zachowują podprzestrzeń $X^{\perp}$. Niech T_0 oznacza obcięcie $T_0 = T|_{X^{\perp}}$ operatora T do podprzestrzeni $x^{\perp}$. Operator T_0 jest oczywiście zwarty i normalny. Jeżeli $T_0 \neq 0$, to $r(T_0) = \|T_0\| \neq 0$, więc $\sigma(T_0) \neq \{0\}$ (patrz twierdzenie 7.49). Wobec tego T_0 ma niezerowy wektor własny odpowiadający wartości własnej różnej od zera. Jest to także wektor własny operatora T. Tu sprzeczność, bo wszystkie taki wektory leżą już w X.

Część druga twierdzenia była dowiedziona już wcześniej. $\square$

Uwaga. Operatory zwarte „nienormalne” mogą w ogóle nie mieć wartości własnych. Jako przykład niech służy operator Volterry $Vx(t) = \int_0^t x(s)\,ds$ na $L^2(0,1)$. Wiemy, że $\sigma(V) = \{0\}$. Liczba 0 nie jest jednak wartością własną operatora V. Istotnie, gdy $Vx = 0$, to $\int_a^b x(s)\,ds = Vx(b) - Vx(a) = 0$ dla dowolnego przedziału $[a,b] \subset [0,1]$. Jest to możliwe tylko wtedy, gdy $x(s) = 0$ prawie wszędzie.

7.87. Wniosek. *Dla dowolnej macierzy normalnej A $n\times n$ istnieje taka macierz unitarna U i macierz diagonalna B, że*

$$A = U^{-1}BU. \tag{7.44}$$

Uwaga. Własność (7.44) charakteryzuje macierze normalne. Podobnie z twierdzeniem spektralnym, które właśnie udowodniliśmy. Jeśli zachodzi dla jakiegoś operatora, to jest to operator normalny.

Operatory Hilberta-Schmidta

Niech $e_1, e_2, e_3, \ldots$ będzie bazą ortonormalną w ośrodkowej przestrzeni Hilberta $\mathcal{H}$. Operator $T \in \mathcal{L}(\mathcal{H})$ nazywamy **operatorem Hilberta-Schmidta** i piszemy $T \in$ H-S, jeżeli

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

Liczbę $|||T||| = \sqrt{\sum_{n=1}^{\infty} \|Te_n\|^2}$ nazywamy **normą Hilberta-Schmidta** operatora T

Wiemy już (patrz **??**), że operatory Hilberta-Schmidta są zwarte.

7.88. Lemat. *Norma Hilberta-Schmidta nie zależy od wyboru bazy ortonornalnej w $\mathcal{H}$.*

Dowód: Niech $e_1', e_2', e_3', \ldots$ oraz $e_1'', e_2'', e_3'', \ldots$ będą bazami ortonornalnymi w $\mathcal{H}$ i niech $|||T|||'$ oraz $|||T|||''$ oznaczają normy Hilberta-Schmidta operatora T, określone przy pomocy tych baz. Równość Parsevala mówi, że dla każdego wektora $x \in \mathcal{H}$ zachodzi równość

$$\|x\|^2 = \sum_{n=1}^{\infty} \left|\langle x,\, e_n'\rangle\right|^2 = \sum_{n=1}^{\infty} \left|\langle x,\, e_n''\rangle\right|^2.$$

Wobec tego

$$\begin{aligned}|||T|||'^2 &= \sum_{n=1}^{\infty} \|Te_n'\|^2 = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \left|\langle Te_n',\, e_m''\rangle\right|^2 \\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \left|\langle e_n',\, T^*e_m''\rangle\right|^2 = \sum_{m=1}^{\infty} \|T^*e_m''\|^2 \\ &= |||T^*|||''^2.\end{aligned}$$

Jeżeli w roli w roli pierwszej bazy weźmiemy drugą, to otrzymamy $|||T|||'' = |||T^*|||''$ a stąd $|||T|||' = |||T^*|||'' = |||T|||''$. □

7.89. Wniosek. *Jeżeli U jest operatorem unitarnym w $\mathcal{H}$ a T operatorem Hilberta-Schmidta, to $U^{-1}TU \in$ H-S oraz $|||U^{-1}TU||| = |||T|||$.*

Dowód: Ponieważ ciąg $Ue_1, Ue_2, Ue_3, \ldots$ jest bazą ortonormalną w $\mathcal{H}$, więc

$$|||U^{-1}TU|||^2 = \sum_{n=1}^{\infty} \|U^{-1}TUe_n\|^2 = \sum_{n=1}^{\infty} \|TUe_n\|^2 = |||T|||^2. \quad \square$$

7.90. Wniosek. *Jeżeli $\in$ H-S i $e_1, e_2, e_3, \ldots$ jest bazą ortonormalną, to*

$$|||T||| = \left(\sum_{n,m=1}^{\infty} \left|\langle Te_n,\, e_m\rangle\right|^2\right)^{1/2}.$$

Dowód: Wynika to natychmiast z faktu, że $\|Te_n\|^2 = \sum_{m=1}^{\infty} \left|\langle Te_n,\, e_m\rangle\right|^2$. □

7.91. Lemat. *Jeżeli $T \in$ H-S, to $\|T\| \leqslant |||T|||$.*

Dowód: Dla każdego $\varepsilon > 0$ znajdziemy taki wektor x_0, że $\|x_0\| = 1$ oraz $\|T\|^2 < \|Tx_0\|^2 + \varepsilon$. Ponieważ istnieje baza ortonormalna w $\mathcal{H}$, zawierająca wektor x_0, więc $\|T\|^2 < |||T|||^2 + \varepsilon$, a stąd $\|T\| \leqslant |||T|||$. □

7.92. Twierdzenie. *Zbiór wszystkich operatorów Hilberta-Schmidta z normą Hilberta-Schmidta jest przestrzenią Banacha. Ponadto przestrzeń H-S jest algebrą, przy czym $|||TS||| \leqslant |||T|||\,|||S|||$ dla dowolnych $T, S \in$ H-S.*

Dowód: Jest jasne, że jeżeli $T \in$ H-S oraz λ jest skalarem, to $|||\lambda T||| = |\lambda|\,|||T|||$. Niech $T, S \in$ H-S i niech $e_1, e_2, e_3, \ldots$ będzie bazą ortonormalną w $\mathcal{H}$. Z wniosku

7.90 i nierówności Minkowskiego wynika, że

$$\begin{aligned}
|||T+S||| &= \Big(\sum_{n,m=1}^{\infty}\big|\langle (T+S)e_n,\, e_m\rangle\big|^2\Big)^{1/2} \\
&= \Big(\sum_{n,m=1}^{\infty}\big|\langle Te_n,\, e_m\rangle\big|^2\Big)^{1/2} + \Big(\sum_{n,m=1}^{\infty}\big|\langle Se_n,\, e_m\rangle\big|^2\Big)^{1/2} \\
&= |||T||| + |||S|||,
\end{aligned}$$

a więc $T+S \in$ H-S. Aby dowieść zupełności przestrzeni H-S rozpatrzmy ciąg $T_1, T_2, T_3, \ldots$ operatorów z H-S, dla którego $|||T_n - T_m||| \to 0$ przy $m, n \to \infty$. Z Lematu 7.91 wynika, że $\|T_n - T_m\| \to 0$, dlatego istnieje taki ograniczony operator liniowy T na $\mathcal{H}$, że $\|T - T_n\| \to 0$. Aby dowieść, że T należy do H-S, oznaczmy przez M kres górny ciągu $\{|||T_n|||\}$. Dla dowolnej liczby naturalnej r mamy

$$\sum_{k=1}^{r} \|Te_k\|^2 = \lim_{n\to\infty} \sum_{k=1}^{r} \|T_n e_k\|^2 \leqslant M^2,$$

i dlatego

$$|||T|||^2 = \sum_{k=1}^{\infty} \|Te_k\|^2 \leqslant M^2.$$

Wynika stąd, że T leży w H-S. Dla $\varepsilon > 0$ dobierzmy tak wskaźnik $m(\varepsilon)$, aby $|||T_n - T_m||| < \varepsilon$ przy $m, n \geqslant m(\varepsilon)$. Wtedy przy $m \geqslant m(\varepsilon)$

$$\sum_{k=1}^{r} \big\|(T - T_m)e_k\big\|^2 = \lim_{n\to\infty} \sum_{k=1}^{r} \big\|(T_n - T_m)e_k\big\|^2 \leqslant \limsup_{n\to\infty} |||T_n - T_m||| \leqslant \varepsilon^2,$$

a stąd $|||T - T_m||| \leqslant \varepsilon$ przy $m \geqslant m(\varepsilon)$. Tak więc przestrzeń H-S z normą Hilberta-Schmidta jest przestrzenią Banacha.

Niech teraz $T \in$ H-S i $B \in \mathcal{L}(\mathcal{H})$. Wtedy

$$|||BT|||^2 = \sum_{n=1}^{\infty} \|BTe_n\|^2 \leqslant \|B\|^2 \sum_{n=1}^{\infty} \|Te_n\|^2 = \|B\|\; |||T|||,$$

więc

$$|||TB||| = |||(TB)^*||| = |||B^*T^*||| \leqslant \|B\|\; |||T|||.$$

wynika stąd w szczególności, że jeżeli $S \in$ H-S, to

$$|||ST||| \leqslant \|S\|\; |||T||| \leqslant |||S|||\; |||T|||,$$

bo $\|S\| \leqslant |||S|||$. $\square$

7.93. WNIOSEK. *Zbiór operatorów Hilberta-Schmidta jest ideałem dwustronnym w algebrze Banacha* $\mathcal{L}(\mathcal{H})$, *wszystkich ograniczonych operatorów liniowych na przestrzeni Hilberta* $\mathcal{H}$. *Ponadto, jeśli* $T \in$ *H-S oraz* $B \in \mathcal{L}(\mathcal{H})$, *to* $|||BT||| \leqslant \|B\|\,|||T|||$ *oraz* $|||TB||| \leqslant |||T|||\,\|B\|$.

7.94. PRZYKŁAD. Niech T będzie zwartym operatorem normalnym w przestrzeni Hilberta $\mathcal{H}$ i niech $\lambda_1, \lambda_2, \lambda_3, \ldots$ będzie ciągiem jego wartości własnych, przy czym każda spośród liczb λ_k występuje tyle razy, ile wynosi jej krotność. Operator $T \in$ H-S wtedy i tylko wtedy, gdy

$$\sum_{k=1}^{\infty} |\lambda_k|^2 < \infty.$$

Rzeczywiście, jeżeli w $\mathcal{H}$ wybrać bazę złożoną z wektorów własnych operatora T, to

$$\sum_{k=1}^{\infty} |\lambda_k| = \sum_{k=1}^{\infty} \|Te_k\|^2 = |||T|||.$$

Widać stąd, że istnieją operatory zwarte, nie będące operatorami Hilberta--Schmidta, np. operatory postaci

$$T(x_1, x_2, x_3, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \lambda_3 x_3, \ldots),$$

gdy ciąg $\lambda_1, \lambda_2, \lambda_3, \ldots$ zbiega do zera, ale $\sum_{k=1}^{\infty} |\lambda_k| = \infty$.

7.95. TWIERDZENIE. *Niech* $\mathcal{K} \in \mathcal{L}^2\big((0,1) \times (0,1)\big)$. *Operator całkowy* K, *określony na przestrzeni* $L^2(0,1)$ *wzorem*

$$Kx(t) = \int_0^1 \mathcal{K}(s,t)\, x(s)\, ds$$

jest operatorem Hilberta-Schmidta, oraz

$$|||K|||^2 = \int_0^1 \int_0^1 \big|\mathcal{K}(s,t)\big|^2 ds\, dt.$$

Każdy operator Hilberta-Schmidta na $L^2(0,1)$ *jest takiej postaci.*

Dowód: Z nierówności Schwarza wynika całkowalność funkcji $\mathcal{K}(s,t)x(s)$ na zbiorze $(0,1) \times (0,1)$. Twierdzenie Fubiniego orzeka zatem, że dla prawie każdego $t \in (0,1)$ funkcja $s \longrightarrow \mathcal{K}(s,t)\, x(s)$ jest całkowalna oraz, że $\int_0^1 \mathcal{K}(s,t)\, x(s)\, ds$ jest mierzalną (nawet całkowalną) funkcją zmiennej t. Oznacza to, że operator K jest dobrze określony. Nie wiemy jednak, czy funkcja Kx leży w $L^2(0,1)$, czy K jest ciągły no i czy leży w H-S?

Niech $\varphi_1(s), \varphi_2(s), \varphi_3(s), \dots$ będzie bazą ortonormalną w $L^2(0,1)$. Wtedy rodzina funkcji $\overline{\varphi_m(s)}\varphi_n(t)$, $m,n = 1,2,3,\dots$ tworzy bazę ortonormalną w przestrzeni $\mathcal{L}^2\big((0,1)\times(0,1)\big)$. Z równości Parsevala mamy więc

$$\sum_{m,n=1}^{\infty}\left|\int_0^1\int_0^1 \mathcal{K}(s,t)\,\varphi_m(s)\overline{\varphi_n(t)}\,ds\,dt\right|^2 = \int_0^1\int_0^1 |\mathcal{K}(s,t)|^2 ds\,dt < \infty.$$

Z drugiej strony, stosując Wniosek 7.90 do operatora K, otrzymamy

$$|||K|||^2 = \sum_{m,n=1}^{\infty}\big|\langle K\varphi_m,\,\varphi_m\rangle\big|^2 = \sum_{m,n=1}^{\infty}\left|\int_0^1\int_0^1 \mathcal{K}(s,t)\,\varphi_m(s)\,\overline{\varphi_n(t)}\,ds\,dt\right|^2,$$

tzn.

$$|||K||| = \sqrt{\int_0^1\int_0^1 |\mathcal{K}(s,t)|^2 ds\,dt} < \infty.$$

Załóżmy teraz, że K jest operatorem Hilberta-Schmidta i określmy na zbiorze $(0,1)\times(0,1)$ funkcję $\mathcal{K}$ wzorem

$$\mathcal{K}(s,t) = \sum_{m,n=1}^{\infty} a_{m,n}\,\overline{\varphi_m(s)}\varphi_n(t),$$

gdzie $a_{m,n} = \langle K\varphi_m,\,\varphi_m\rangle$, $m,n = 1,2,3,\dots$. Wtedy, jak to pokazaliśmy przed chwilą,

$$\int_0^1\int_0^1 |\mathcal{K}(s,t)|^2 ds\,dt = \sum_{m,n=1}^{\infty} |a_{m,n}|^2 = |||K|||,$$

a dla $x\in\mathcal{L}^2(0,1)$, postaci $x(s) = \sum_{m=1}^{\infty} b_m\,\varphi_m(s)$ dostaniemy

$$\begin{aligned} Kx(t) &= \sum_{n=1}^{\infty}\langle Kx,\,\varphi_n\rangle\,\varphi_n(t) = \sum_{m,n=1}^{\infty} a_{m,n}\,b_m\,\varphi_n(t) \\ &= \int_0^1 \mathcal{K}(s,t)\,x(s)\,ds. \quad \square \end{aligned}$$

7.96. Przykład. Na przestrzeni $L^2(0,1)$ operator Volterry

$$Vx(t) = \int_0^t x(s)\,ds$$

ma jądro całkowe

$$\mathcal{K}(s,t) = \begin{cases} 1 & \text{gdy } 0 \leqslant s < t \leqslant 1, \\ 0 & \text{gdy } 0 \leqslant t \leqslant s \leqslant 1, \end{cases}$$

leżące w $L^2\big((0,1)\times(0,1)\big)$. Z Twierdzenia 7.95 otrzymujemy więc

$$|||V||| = \frac{1}{\sqrt{2}}.$$

7.97. TWIERDZENIE. *Każdy operator Hilberta-Schmidta można przedstawić jako granicę zbieżnego w normie Hilberta-Schmidta ciągu operatorów skończenie wymiarowych.*

Dowód: Niech $T \in$ H-S i niech $e_1, e_2, e_3, \ldots$ będzie bazą ortonormalną w $\mathcal{H}$. Ponieważ $|||T|||^2 = \sum_{n=1}^{\infty} \|Te_n\|^2 < \infty$, więc istnieje taki ciąg wskaźników $n_1, n_2, n_3, \ldots$, że

$$\sum_{n>n_k} \|Te_n\|^2 < \frac{1}{k^2}.$$

Niech P_k, $k = 1,2,3,\ldots$, oznacza rzut ortogonalny na podprzestrzeń liniową w $\mathcal{H}$, rozpiętą na wektorach $e_1, e_2, \ldots, e_{n_k}$. Wtedy każdy z operatorów $T_k = TP_m$ jest skończenie wymiarowy i

$$|||T - T_k||| = \sqrt{\sum_{n>n_k} \|Te_n\|^2} < \frac{1}{k}. \quad \square$$

Jak wiemy, zbiór operatorów Hilberta-Schmidta tworzy algebrę Banacha (bez jedności, jeśli przestrzeń $\mathcal{H}$ jest nieskończenie wymiarowa). W zbiorze tym można określić iloczyn skalarny.

7.98. TWIERDZENIE. *Niech S i T będą operatorami Hilberta-Schmidta w przestrzeni Hilberta $\mathcal{H}$. Jeżeli $e_1, e_2, e_3, \ldots$ jest bazą ortonormalną w $\mathcal{H}$, to szereg*

$$\langle\langle S,\, T\rangle\rangle = \sum_{n=1}^{\infty} \langle Se_n,\, Te_n\rangle$$

jest bezwzględnie zbieżny i jego granica nie zależy od wyboru bazy. Funkcja $\langle\langle\ ,\ \rangle\rangle$ ma wszystkie własności iloczynu skalarnego oraz

$$\langle\langle T,\, T\rangle\rangle = |||T|||^2$$

dla $T \in$ H-S. Ponadto, dla operatorów $S, T, R \in$ H-S zachodzi równość

$$\langle\langle ST,\, R\rangle\rangle = \langle\langle T,\, S^*R\rangle\rangle.$$

Dowód: Bezwzględna zbieżność szeregu wynika natychmiast z nierówności Schwarza a niezależność od wyboru bazy z nierówności

$$\langle Se_n, Te_n \rangle = \sum_{n=1}^{\infty} \langle Se_n, f_n \rangle \langle f_m, Te_n \rangle, \quad n = 1, 2, 3, \ldots,$$

gdzie $f_1, f_2, f_3, \ldots$ jest dowolną bazą ortonormalną w $\mathcal{H}$. Pozostałe stwierdzenia są oczywiste. $\square$

Oznacza to, że H-S$(\mathcal{H})$ jest jednocześnie algebrą Banacha i przestrzenią Hilberta, a inwolucja $S \longrightarrow S^*$ spełnia tożsamość

$$\langle\langle ST, R \rangle\rangle = \langle\langle T, S^* R \rangle\rangle.$$

Takie algebry noszą nazwę H^***-alebr**. Wiadomo, że każda H^*-algebra jest izometrycznie *-izomorficzna algebrze operatorów Hilberta-Schmidta na pewnej przestrzeni Hilberta.

ROZDZIAŁ VIII

ALGEBRY BANACHA

Własności ogólne

Spotkaliśmy się już wielokrotnie z przestrzeniami Banacha, w których określone było dodatkowe działanie, tworzące strukturę algebry. Na przykład w przestrzeni $C(\mathbb{R})$ tym działaniem było zwykłe mnożenie funkcji, w przestrzeni $L^1(\mathbb{R})$ splot

$$x * y(t) = \int_{-\infty}^{+\infty} x(t-s)\, y(s)\, ds,$$

a w przestrzeni $\mathcal{L}(X)$ ograniczonych operatorów liniowych na przestrzeni Banacha, składanie operatorów. Obiekty tego typu noszą ogólną nazwę algebr Banacha.

Definicja. Algebrę $\mathcal{A}$ z jednością e nad ciałem liczb zespolonych nazywamy **algebrą Banacha** jeżeli wyposażona jest w normę $\| \ \|$ spełniającą warunki

$$\|e\| = 1, \qquad \|x+y\| \leqslant \|x\| + \|y\|, \quad x, y \in \mathcal{A},$$

i jest w tej normie przestrzenią Banacha. Algebra Banacha jest **przemienną**, gdy $xy = yx$ dla wszystkich $x, y \in \mathcal{A}$. **Inwolucją** w algebrze Banacha $\mathcal{A}$ nazywamy takie odwzorowanie $x \to x^*$ algebry $\mathcal{A}$ w siebie, że

$$\begin{aligned} (x+y)^* &= x^* + y^*, \qquad & (xy)^* &= y^* x^*, \\ (\lambda x)^* &= \overline{\lambda} x^*, & (x^*)^* &= x. \end{aligned}$$

W myśl tej definicji $L^1(\mathbb{R})$ ze splotem nie jest algebrą Banacha, bo nie posiada jedności. Nie jest to jednak wielka przeszkoda. Pokażemy, że można tej algebrze dołączyć jedność tak, by stała się algebrą Banacha.

Załóżmy, że w algebrze $\mathcal{A}$ nie ma jedności, ale że spełnione są wszystkie pozostałe aksjomaty algebry Banacha. Połóżmy $\widetilde{\mathcal{A}} = \mathbb{C} \times \mathcal{A}$, gdzie $\mathbb{C}$ jest ciałem liczb zespolonych. Zbiór $\widetilde{\mathcal{A}}$ składa się z wszystkich takich par $[\alpha, x]$, że $\alpha \in \mathbb{C}$, $x \in \mathcal{A}$. Działania i normę w $\widetilde{\mathcal{A}}$ określamy następująco

$$\begin{aligned} [\alpha, x] + [\beta, y] &= [\alpha + \beta, x + y], \\ [\alpha, x][\beta, y] &= [\alpha\beta, \alpha y + \beta x + xy], \\ \lambda\,[\alpha, x] &= [\lambda\alpha, \lambda x], \\ \big\|[\alpha, x]\big\| &= |\alpha| + \|x\|. \end{aligned}$$

Jest oczywiste, że $\widetilde{\mathcal{A}}$ jest algebrą Banacha z jednością $e = [1, 0]$, przy czym $\|e\| = 1$ i odwzorowanie $x \to [0, x]$ jest izometrycznym włożeniem $\mathcal{A}$ w $\widetilde{\mathcal{A}}$.

Opiszemy jeszcze jedną metodę „poprawiania" algebry tak, by stała się algebrą Banacha. Załóżmy, że przestrzeń Banacha $\mathcal{A}$ jest jednocześnie algebrą nad ciałem liczb zespolonych z jednością e ale, że zamiast własności $\|e\| = 1$ i $\|xy\| \leqslant \|x\|\|y\|$ wiemy tylko, że $\|e\| \neq 1$ i że iloczyn xy jest ciągły po każdej ze zmiennych przy ustalonej drugiej. Określmy odwzorowanie $\tau : x \to T_x$ algebry $\mathcal{A}$ w algebrę $\mathcal{L}(\mathcal{A})$ wszystkich ograniczonych operatorów liniowych na $\mathcal{A}$ kładąc $T_x y = xy$. Jest jasne, że τ jest algebraicznym izomorfizmem algebry $\mathcal{A}$ w podalgebrę $\tau(\mathcal{A})$ algebry $\mathcal{L}(\mathcal{A})$. Pokażemy, że τ jest także homeomorfizmem. Będzie to oznaczało, że algebra $\mathcal{A}$ jest topologicznie i algebraicznie równoważna algebrze Banacha $\tau(\mathcal{A})$. Ponieważ $\|x\| = \|xe\| = \|T_x e\| \leqslant \|e\|\|T_x\|$, więc odwzorowanie $\tau^{-1} : Tx \to x$ jest ciągłe. Aby dowieść ciągłości τ pokażemy najpierw, że operator $T \in \mathcal{L}(\mathcal{A})$ leży w $\tau(\mathcal{A})$ wtedy i tylko wtedy, gdy $(Ty)z = T(yz)$ dla wszystkich $y, z \in \mathcal{A}$. Rzeczywiście, jeśli warunek ten jest spełniony, to kładąc $x = Te$, otrzymamy $Ty = T(ey) = (Te)y = xy$, tzn. $T = T_x \in \tau(\mathcal{A})$. To, że każdy operator w $\tau(\mathcal{A})$ ma wspomnianą własność wynika natychmiast z łączności mnożenia. Jeżeli teraz $\{T_n\}$ jest ciągiem w $\tau(\mathcal{A})$ i $T_n \to T$, to z równości

$$(Ty)z = \lim_{n\to\infty} (T_n y)z = \lim_{n\to\infty} T_n(yz) = T(yz)$$

wynika, że $T \in \tau(\mathcal{A})$. Dlatego $\tau(\mathcal{A})$ jest domkniętą podprzestrzenią w $\mathcal{L}(\mathcal{A})$. Oznacza to, że τ jest wzajemnie jednoznacznym odwzorowaniem liniowym przestrzeni Banacha $\mathcal{A}$ na przestrzeń Banacha $\tau(\mathcal{A})$, a ponieważ odwzorowanie odwrotne τ^{-1} jest ciągłe, więc z twierdzenia Banacha o odwzorowaniu odwrotnym wynika, że τ jest homeomorfizmem.

8.1. Wniosek. *Każda algebra Banacha $\mathcal{A}$ jest izometrycznie izomorficzna z domkniętą podalgebrą algebry Banacha $\mathcal{L}(X)$ wszystkich ciągłych operatorów liniowych na pewnej przestrzeni Banacha X.*

8.2. Przykład. Algebra $C_0(\mathbb{R})$ funkcji ciągłych na prostej $\mathbb{R}$ dążących do zera w nieskończoności, z normą $\|x\| = \max_{t\in\mathbb{R}} \big|x(t)\big|$, ma wszystkie własności algebry

Banacha, z wyjątkiem posiadania jedności. Mamy dwie możliwości jej dołączenia. Pierwszą „standartową”, gdzie $\widetilde{\mathcal{A}} = \mathbb{C}\times C_0(\mathbb{R})$, a drugą przez wybranie w algebrze $C(\mathbb{R})$ najmniejszej podalgebry (domkniętej) zawierającej $C_0(\mathbb{R})$. Jest nią algebra wszystkich funkcji mających granicę w nieskończoności, tzn. funkcji postaci $\lambda\boldsymbol{1} + x$, gdzie $\lambda \in \mathbb{C}$, $x \in C_0(\mathbb{R})$. Jest to więc algebra izomorficzna z $\widetilde{\mathcal{A}}$. Nie jest jednak izometrycznie izomorficzna. Bo np. dla funkcji $x(t) = t^2/(1+t^2) = 1 - 1/(1+t^2)$ mamy

$$\|x\|_{\widetilde{\mathcal{A}}} = 1 + \max_{t\in\mathbb{R}} \left|\frac{1}{1+t^2}\right| = 2, \qquad \|x\|_\infty = 1.$$

Wydaje się, że drugi sposób „dołączania jedności” jest lepszy.

8.3. Przykład. Podobnie jest dla algebry $L^1(\mathbb{R})$. Drugi sposób dołączania jedności może tu polegać na wybraniu w algebrze $M(\mathbb{R})$ wszystkich miar borelowskich skończonych z normą $\|\mu\| = \mathrm{Var}(\mu)$, splotem

$$\mu * \nu(A) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \chi_A(t+s)\,d\mu(t)\,d\nu(s)$$

i jednością δ_0, najmniejszej domkniętej podalgebry zawierającej $L^1(\mathbb{R})$ (jako miary absolutnie ciągłe względem miary Lebesgue'a). Jest nią zbiór wszystkich miar postaci

$$d\mu(t) = \lambda\,\delta_0 + x(t)\,dt,$$

gdzie $\lambda \in \mathbb{C}$, $x \in L^1(\mathbb{R})$. Mamy tu $\|\mu\| = |\lambda| + \|x\|_1$, a więc algebra ta jest izometrycznie izomorficzna z $\widetilde{\mathcal{A}}$.

Definicja. Element x algebry Banacha $\mathcal{A}$ nazywamy **regularnym**, jeśli posiada element odwrotny x^{-1} a **singularnym** lub **osobliwym** w przeciwnym przypadku. **Widmo** $\sigma(x)$ elementu x składa się z tych liczb zespolonych λ, dla których element $\lambda e - x$ jest singularny. Liczba $|\sigma(x)| = \sup_{\lambda\in\sigma(x)}|\lambda|$ nosi nazwę **promienia spektralnego** elementu x. **Zbiorem rezolwenty** $\rho(x)$ jest dopełnienie $\mathbb{C}\setminus\sigma(x)$ spektrum $\sigma(x)$. Funkcja $x(\lambda) = (\lambda e - x)^{-1}$ określona dla $\lambda \in \rho(x)$ nosi nazwę **rezolwenty** elementu x. Element $x \in \mathcal{A}$ nazywamy **prawym** (odp. **lewym**) **topologicznym dzielnikiem zera**, jeżeli w $\mathcal{A}$ istnieje taki ciąg $\{x_n\}$, że $\|x_n\| = 1$, $n = 1,2,3,\ldots$, oraz $x_n x \to 0$ (odp. $x x_n \to 0$). Jeżeli element x jest jednocześnie prawym i lewym topologicznym dzielnikiem zera, to będziemy go nazywać **dwustronnym topologicznym dzielnikiem zera**.

Uwaga. Jak widzieliśmy we wniosku 8.1, każdą algebrę Banacha $\mathcal{A}$ można traktować jako domkniętą podalgebrą algebry $\mathcal{L}(X)$ dla pewnej przestrzeni Banacha X (jest nią $\mathcal{A}$). Powstaje pytanie, czy pojęcia — które przed chwilą wprowadziliśmy — pokrywają się ze znanymi nam już odpowiednimi pojęciami z teorii operatorów. Odpowiedź brzmi TAK.

Dla $x \in \mathcal{A}$ niech T_x oznacza operator $T_x y = xy$, $y \in \mathcal{A}$. Jeżeli element x^{-1} istnieje, to

$$T_{x^{-1}}(T_x y) = y = T_x(T_{x^{-1}} y),$$

więc $T_{x^{-1}} = T_x^{-1}$. Z drugiej strony, gdy T_x^{-1} istnieje, to

$$T_x\big[(T_x^{-1}y)z\big] = yz \quad \text{tzn.} \quad (T_x^{-1}y)z = T_x^{-1}(yz)$$

i jeśli $a = T_x^{-1}e$, to $az = T_x^{-1}z$ dla wszystkich $z \in \mathcal{A}$. Oprócz tego

$$xa = T_x a = e = T_x^{-1}(T_x e) = T_x^{-1}(ex) = (T_x^{-1}e)x = ax,$$

a więc $x^{-1} = a$ istnieje i $T_x^{-1}z = x^{-1}z$.

Z rozumowania tego wynika, że $\sigma(x) = \sigma(T_x)$. Pozwala nam to zaadoptować niektóre spośród twierdzeń teorii spektralnej operatorów do teorii algebr Banacha i odwrotnie. Na przykład:

8.4. LEMAT. *Multiplikatywna grupa G elementów regularnych algebry Banacha $\mathcal{A}$ jest zbiorem otwartym w $\mathcal{A}$, a odwzorowanie $x \to x^{-1}$ jest homeomorfizmem G na siebie.*

Dowód: Pokażemy najpierw, że G zawiera kulę $\{x \in \mathcal{A} : \|e - x\| < 1\}$. Rzeczywiście, jeżeli $\|e - x\| < 1$, to szereg $y = \sum_{n=0}^{\infty}(e-x)^n$ jest zbieżny, dlatego

$$yx = xy = y - (e-x)y = \sum_{n=0}^{\infty}(e-x)^n - \sum_{n=1}^{\infty}(e-x)^n = e.$$

W konsekwencji $y = x^{-1}$ i

$$\|x^{-1} - e\| = \left\| \sum_{n=1}^{\infty}(e-x)^n \right\| \leqslant \frac{\|e-x\|}{1 - \|e-x\|}.$$

Pokazaliśmy, że G zawiera otoczenie jedności i x^{-1} jest ciągłą funkcją x w punkcie e. Niech teraz $x \in G$ i $\|y - x\| \leqslant \frac{1}{\|x^{-1}\|}$. Wtedy

$$\|x^{-1}y - e\| = \|x^{-1}(y-x)\| \leqslant \|x^{-1}\|\|y-x\| < 1,$$

a więc z tego, co udowodniliśmy poprzednio, $x^{-1}y \in G$, skąd $y \in G$. Jeżeli $y_n \to y$, to $y_n y^{-1} \to e$ i dlatego $yy_n^{-1} = (y_n y^{-1})^{-1} \to e$. W konsekwencji $y_n^{-1} \to y^{-1}$. $\square$

8.5. LEMAT. *Każdy punkt brzegowy grupy elementów regularnych algebry Banacha jest dwustronnym topologicznym dzielnikiem zera.*

Dowód: Niech $x \notin G$, $x_n \in G$, $x_n \to x$. Jeżeli ciąg $\{\|x_n^{-1}\|\}$ jest ograniczony, to $x_n x - e = x_n^{-1}(x - x_n) \to 0$, co z lematu 8.4 oznacza, że $x_n^{-1}x \in G$ dla dużych n. Przeczy to założeniu $x \notin G$. Tak więc ciąg $\{\|x_n^{-1}\|\}$ jest nieograniczony i możemy założyć, że $\|x_n^{-1}\| \to \infty$. Połóżmy $y_n = x_n^{-1}/\|x_n^{-1}\|$. Ponieważ $\|y_n\| = 1$ oraz

$$y_n x = y_n(x - x_n) + \frac{e}{\|x_n^{-1}\|} \to 0,$$
$$x y_n = (x - x_n) y_n + \frac{e}{\|x_n^{-1}\|} \to 0,$$

więc x jest dwustronnym topologicznym dzielnikiem zera. □

8.6. LEMAT. *Widmo $\sigma(x)$ elementu x algebry Banacha jest zbiorem niepustym, ograniczonym i domkniętym. Rezolwenta $x(\lambda) = (\lambda e - x)^{-1}$ elementu x jest analityczna na zbiorze $\rho(x)$, dąży do zera przy $|\lambda| \to \infty$ i spełnia równanie*

$$x(\lambda) - x(\mu) = (\mu - \lambda)\, x(\lambda)\, x(\mu), \qquad \lambda, \mu \in \rho(x).$$

Dowód: To, że zbiór $\rho(x)$ jest otwarty, a w konsekwencji $\sigma(x)$ domknięty, wynika z lematu 8.4. Lemat ten pokazuje też, że $\lambda e - x = \lambda(e - x/\lambda) \in G$ dla dużych λ (bo element $e - x/\lambda$ jest bliski e). Stąd wynika ograniczoność zbioru $\sigma(x)$. Ponieważ $e - x/\lambda \to e$ przy $|\lambda| \to \infty$, to z lematu 8.4 $x(\lambda) = \lambda^{-1}(e - x/\lambda)^{-1} \to 0$ przy $|\lambda| \to \infty$. Dla dowolnych $\lambda, \mu \in \rho(x)$ elementy $x(\mu)$ i $x(\lambda)$ komutują i

$$(\lambda e - x)\, x(\lambda)\, x(\mu) = x(\mu),$$
$$(\mu e - x)\, x(\lambda)\, x(\mu) = x(\lambda),$$
$$x(\lambda) - x(\mu) = (\mu - \lambda)\, x(\lambda)\, x(\mu),$$

skąd

$$\frac{x(\lambda) - x(\mu)}{\lambda - \mu} = -x(\lambda)\, x(\mu).$$

Ponieważ funkcja $x(\lambda)$ jest ciągła dla $\lambda \in \rho(x)$, ostatnia równość dowodzi, że $x'(\lambda) = -[x(\lambda)]^2$, a w konsekwencji dowodzi analityczności rezolwenty $x(\lambda)$ na zbiorze $\rho(x)$.

Na zakończenie zauważmy jeszcze, że jeżeli widmo $\sigma(x)$ jest zbiorem pustym, to dla każdego funkcjonału liniowego $x^* \in \mathcal{A}^*$ funkcja $\lambda \to x^*(x(\lambda))$ jest analityczna na całej płaszczyźnie $\mathbb{C}$ i dąży do zera w nieskończoności, dlatego równa jest zeru wszędzie. Ponieważ za x^* można wybrać dowolny element przestrzeni sprzężonej $\mathcal{A}^*$, to $0 = x(\lambda) = x(\lambda)(\lambda e - x) = e$, co przeczy własności $\|e\| = 1$. □

8.7. TWIERDZENIE. *Jeżeli algebra Banacha nie zawiera różnych od zera dwustronnych topologicznych dzielników zera, to jest izometrycznie izomorficzna z ciałem liczb zespolonych.*

Dowód: Niech $x \in \mathcal{A}$, Zgodnie z lematem 8.6, widmo $\sigma(x)$ jest zbiorem niepustym i ograniczonym, istnieje więc punkt λ należący do jego brzegu. Wtedy, jak to wynika z lematu 8.5, $\lambda e - x$ jest dwustronnym topologicznym dzielnikiem zera, zatem $x = \lambda e$. Ponieważ $\|e\| = 1$, więc $\|x\| = |\lambda|$. $\square$

8.8. WNIOSEK (TWIERDZENIE MAZURA-GELFANDA). *Jeżeli algebra Banacha jest ciałem, to jest izometrycznie izomorficzna z ciałem liczb zespolonych.*

UWAGA. Mazur udowodnił ogólniejsze twierdzenie: Jeżeli algebra Banacha nad ciałem liczb rzeczywistych jest ciałem, to jest izomorficzna albo z ciałem liczb rzeczywistych, albo z ciałem liczb zespolonych, albo z ciałem kwaternionów (patrz np. [10]).

8.9. LEMAT. *Promień spektralny $|\sigma(x)|$ elementu x algebry Banacha spełnia równość*

$$\big|\sigma(x)\big| = \lim_{n\to\infty} \|x^n\|^{1/n} \leqslant \|x\|.$$

Dowód: Jeżeli $|\lambda| > \|x\|$, to szereg $\sum\limits_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}}$ jest zbieżny, a ponieważ

$$(\lambda e - x)\sum_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}} = \sum_{n=0}^{\infty}\Big(\frac{x^n}{\lambda^n} - \frac{x^{n+1}}{\lambda^{n+1}}\Big) = e,$$

więc λ należy do zbioru rezolwenty $\rho(x)$. Dlatego $|\sigma(x)| \leqslant \|x\|$. Z lematu 8.6 wynika, że rezolwenta $x(\lambda)$ jest funkcją analityczną na zbiorze $\rho(x)$ i dlatego dla każdego funkcjonału $x^* \in \mathcal{A}^*$ wszystkie punkty osobliwe skalarnej funkcji analitycznej $x^*\big(x(\lambda)\big)$ zawarte są w kole $|\lambda| \leqslant |\sigma(x)|$. Wynika stąd, że szereg $x^*\big(x(\lambda)\big) = \sum_{n=0}^{\infty} x^*(x^n)/\lambda^{n+1}$ jest zbieżny przy $|\lambda| > |\sigma(x)|$, a w konsekwencji, że dla takich λ

$$\sup_n \left|\frac{x^*(x^n)}{\lambda^{n+1}}\right| < \infty.$$

Ponieważ x^* jest dowolnym elementem $\mathcal{A}^*$, z twierdzenia Banacha-Steinhausa wnosimy, że

$$\left\|\frac{x^n}{\lambda^{n+1}}\right\| \leqslant M_\lambda < \infty, \qquad n = 1, 2, 3, \ldots,$$

a stąd

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leqslant |\sigma(x)|.$$

Dla zakończenia dowodu zauważmy, że jeżeli element $\lambda e - x$ jest singularny, to singularny jest również element $\lambda^n e - x^n$, bo jest wielokrotnością $\lambda e - x$. Dlatego $\lambda \in \sigma(x)$ pociąga $\lambda^n \in \sigma(x^n)$, skąd $|\lambda^n| \leqslant \|x^n\|$, a dalej $|\lambda| < \liminf_{n\to\infty} \|x^n\|^{1/n}$ czyli, że $|\sigma(x)| \leqslant \liminf_{n\to\infty} \|x^n\|^{1/n}$. $\square$

Jeżeli $\mathcal{A}_0$ jest podalgebrą Banacha (tzn. domkniętą podalgebrą z jednością) algebry $\mathcal{A}$ i $x \in \mathcal{A}_0$, to wraz z widmem $\sigma(x)$ elementu x względem algebry $\mathcal{A}$ można rozpatrywać jego widmo $\sigma_0(x)$ względem podalgebry $\mathcal{A}_0$. Zbiory $\sigma(x)$ i $\sigma_0(x)$, ogólnie rzecz biorąc, są różne, jednak — jak to widać z powyższego lematu — ich promienie spektralne $|\sigma(x)|$ i $|\sigma_0(x)|$ są sobie równe. Jeżeli e_0 jest idempotentem w $\mathcal{A}$ (tzn. $e_0^2 = e_0$) różnym od e i 0, a $\mathcal{A}_0 = e_0\mathcal{A}e_0$, to jest jasne, że $\sigma_0(x) \subset \sigma(x)$ dla każdego $x \in \mathcal{A}_0$ Poniższy lemat pokazuje, że jeżeli jedynka podalgebry $\mathcal{A}_0$ pokrywa się z e, to zachodzi zawieranie odwrotne.

8.10. Lemat. *Niech $\mathcal{A}_0$ będzie podalgebrą Banacha algebry Banacha $\mathcal{A}$, mającą wspólną z $\mathcal{A}$ jedynkę e. Wtedy $\sigma(x) \subset \sigma_0(x)$ dla każdego $x \in \mathcal{A}_0$, a dla brzegów tych zbiorów zachodzi zawieranie przeciwne $\partial\big(\sigma_0(x)\big) \subset \partial\big(\sigma(x)\big)$.*

Dowód: Ponieważ jedynka e algebry $\mathcal{A}$ leży w $\mathcal{A}_0$, to każdy element regularny w $\mathcal{A}_0$ jest regularny w $\mathcal{A}$. Dlatego $\rho_0(x) \subset \rho(x)$, czyli $\sigma(x) \subset \sigma_0(x)$. Jeżeli $\lambda \in \partial\big(\sigma_0(x)\big)$, to $\lambda e - x$ jest punktem brzegu grupy elementów odwracalnych algebry $\mathcal{A}_0$. Z lematu 8.5 wynika, że $\lambda e - x$ jest dwustronnym topologicznym dzielnikiem zera w $\mathcal{A}_0$, a więc i w $\mathcal{A}$. Wynika stąd, że $\partial\big(\sigma_0(x)\big) \subset \sigma(x)$, co w połączeniu z zawieraniem $\rho_0(x) \subset \rho(x)$ daje

$$\partial\big(\sigma_0(x)\big) = \overline{\rho_0(x)} \cap \partial\big(\sigma_0(x)\big) \subset \overline{\rho(x)} \cap \sigma(x) = \partial\big(\sigma(x)\big). \quad \square$$

8.11. Przykład. Przy oznaczeniach $D = \{z \in \mathbb{C} : |z| < 1\}$ oraz $T = \partial D = \{z \in \mathbb{C} : |z| = 1\}$ rozpatrzmy algebrę Banacha $\mathcal{A} = C(T)$ i jej podzbiór $\mathcal{A}_0$ złożony z tych funkcji holomorficznych w D, które posiadają przedłużenie do funkcji ciągłej na $\overline{D}$. Z twierdzenia Cauchy'ego wynika, że $\mathcal{A}_0$ jest domkniętą podalgebrą Banacha algebry $\mathcal{A}$. Dla funkcji $x(z) = z$ mamy $\sigma(x) = T$, zaś $\sigma_0(x) = \overline{D}$. Tak więc $\sigma(x) \neq \sigma_0(x)$, $\partial\big(\sigma(x)\big) = \partial\big(\sigma_0(x)\big)$. Łatwo sprawdzić, że dla dowolnej funkcji $x \in \mathcal{A}_0$ mamy $\sigma(x) = x(T)$, zaś $\sigma_0(x) = x(\overline{D})$, czyli $\sigma(x) \subset \sigma_0(x)$. Z zasady maksimum dla funkcji holomorficznych wynika, że $\partial\big(x(T)\big) = \partial\big(x(\overline{D})\big)$, tzn. $\partial\big(\sigma_0(x)\big) = \partial\big(\sigma(x)\big)$.

8.12. Wniosek. *Jeżeli w lemacie 8.10 założymy dodatkowo, że $\sigma_0(x)$ jest zbiorem brzegowym lub, że $\rho(x)$ jest zbiorem spójnym, to $\sigma_0(x) = \sigma(x)$.*

Dowód: Jeżeli $\sigma_0(x)$ jest zbiorem brzegowym, to

$$\sigma_0(x) = \partial\big(\sigma_0(x)\big) \subset \partial\big(\sigma(x)\big) \subset \sigma(x) \subset \sigma_0(x).$$

Jeżeli zbiór $\rho(x)$ jest spójny i istnieje punkt $\lambda \in \sigma_0(x) \cap \rho(x)$, to można go połączyć z nieskończonością krzywą ciągłą całkowicie leżącą w zbiorze $\rho(x)$. Ale wtedy znajdziemy punkt brzegowy zbioru $\sigma_0(x)$ leżący w $\rho(x)$, a to przeczy lematowi 8.10. Tak więc zbiór $\sigma_0(x) \cap \rho(x)$ jest pusty, co oznacza, że $\sigma_0(x) \subset \sigma(x) \subset \sigma_0(x)$. $\square$

8.13. WNIOSEK. *Jeżeli w lemacie 8.10 założyć dodatkowo, że widmo $\sigma_0(x)$ jest rzeczywiste, to pokrywa się ono z widmem $\sigma_1(x)$ elementu x względem dowolnej podalgebry Banacha $\mathcal{A}_1 \subset \mathcal{A}$, zawierającej x i jedynkę algebry $\mathcal{A}$.*

Dowód: Ponieważ $\sigma(x) \subset \sigma_0(x)$, więc $\sigma(x)$ leży na prostej rzeczywistej. Ponieważ $\sigma(x)$ jest zbiorem ograniczonym, więc $\rho(x)$ jest zbiorem spójnym. Poprzedni wniosek pokazuje, że $\sigma_0(x) = \sigma(x) = \sigma_1(x)$. $\square$

DEFINICJA. **Prawym** (odp. **lewym**) **ideałem** w algebrze Banacha $\mathcal{A}$ nazywamy taką niezerową właściwą podprzestrzeń liniową J w $\mathcal{A}$, że $J\mathcal{A} = J$ (odp. $\mathcal{A}J = J$). Jeżeli jednocześnie $J\mathcal{A} = J$ i $\mathcal{A}J = J$, to ideał J nazywamy **dwustronnym**. Zbiory $\mathcal{A}$ i $\{0\}$ traktujemy jako ideały trywialne.

Ponieważ ideał jest właściwym podzbiorem $\mathcal{A}$, nie może zawierać żadnego elementy regularnego i dlatego jest zawarty w dopełnieniu G^c grupy elementów regularnych. Ponieważ zbiór G jest otwarty (lemat 8.4), to $\overline{J} \subset G^c$, a więc $\overline{J} \neq \mathcal{A}$. Stąd i z ciągłości operacji algebraicznych w $\mathcal{A}$ wynika, że $\overline{J}$ jest także ideałem. Tak więc domknięcie prawego, lewego lub dwustronnego ideału jest znów ideałem prawym, lewym lub dwustronnym. Wynika stąd, że ideał maksymalny (tzn. nie zawierający się w żadnym innym ideale tego samego typu) jest zawsze domknięty.

Załóżmy, że J jest prawym ideałem. Uporządkujmy przez zawieranie zbiór wszystkich prawych ideałów zawierających J. Stosując lemat Zorna otrzymamy, że zbiór ten zawiera element maksymalny. Tak więc każdy prawy (analogicznie lewy i dwustronny) ideał jest zawarty w pewnym maksymalnym prawym (odp. lewym i dwustronnym) ideale. Wynika stąd w szczególności, że jeżeli element x jest singularny, to przynajmniej jeden ze zbiorów $x\mathcal{A}$ lub $\mathcal{A}x$ jest ideałem, jest więc zawarty w pewnym ideale maksymalnym.

Własności ideałów zebrane są w następującym lemacie:

8.14. LEMAT. *Ideały prawostronne, lewostronne i dwustronne mają następujące własności:*

a. *Ideał nie zawiera elementów regularnych.*

b. *Domknięcie ideału jest znów ideałem.*

c. *Ideał maksymalny jest domknięty.*

d. *Każdy ideał jest zawarty w pewnym ideale maksymalnym.*

e. *Element x leży w przynajmniej jednym ideale maksymalnym prawostronnym (odp. lewostronnym) wtedy i tylko wtedy, gdy nie ma prawego (odp. lewego) odwrotnego.* $\square$

Zwróćmy uwagę, że klasy równoważności $x+J$, $x \in \mathcal{A}$, względem ideału dwustronnego J tworzą algebrę względem następujących działań

$$(x+J)+(y+J)=(x+y)+J,$$
$$\alpha(x+J)=(\alpha x)+J, \qquad (x+J)(y+J)=(xy)+J.$$

Algebra ta nosi nazwę **algebry ilorazowej** algebry $\mathcal{A}$ względem ideału J i oznaczana jest $\mathcal{A}/J$. Normę w algebrze ilorazowej określamy wzorem

$$\|x+J\| = \inf_{y \in J} \|x+y\|.$$

8.15. LEMAT. *Jeżeli J jest domkniętym ideałem dwustronnym w algebra Banacha $\mathcal{A}$, to $\mathcal{A}/J$ jest algebrą Banacha.*

Dowód: Dla uproszczenia oznaczmy przez $[x]$ klasę równoważności elementu x w $\mathcal{A}/J$. Jest oczywiste, że $\mathcal{A}/J$ jest algebrą z jednością $[e]$. Pozostaje tylko dowieść, że spełnione są aksjomaty normy. Jeżeli $\|[x]\|=0$, to znajdziemy taki ciąg $\{x_n\}$ w J, że $\|x+x_n\| \to 0$, a ponieważ J jest zamknięty, więc $x \in J$. W konsekwencji $[x]=[0]$ tzn., że $[x]$ jest zerem w $\mathcal{A}/J$. Załóżmy teraz, że z,u,v przebiegają niezależnie zbiór J. Wtedy

$$\big\|[x][y]\big\| = \big\|[xy]\big\| = \inf_z \|xy+z\| \leqslant \inf_{u,v} \big\|(x+u)(y+v)\big\| \leqslant \big\|[x]\big\|\big\|[y]\big\|.$$

Ponieważ $\|[e]\| = \|[e^2]\| \leqslant \|[e]\|^2$ i $\|[e]\| \neq 0$, więc $\|[e]\| \geqslant 1$. Z drugiej strony $\|[e]\| = \inf_{z\in J}\|e+z\| \leqslant \|e\| = 1$. W konsekwencji $\|[e]\|=1$. Podaddytywność i jednorodność normy ilorazowej są oczywiste. Także zupełność przestrzeni $\mathcal{A}/J$. $\square$

Przemienne algebry Banacha

W przemiennej algebrze Banacha $\mathcal{A}$ każdy ideał J jest ideałem dwustronnym, a algebra ilorazowa $\mathcal{A}/J$ jest przemienna. Jest ona algebrą Banacha, gdy J jest ideałem domkniętym. Łatwo zauważyć, że każdy ideał J' algebry $\mathcal{A}$ zawierający J jako właściwy podzbiór określa w algebrze $\mathcal{A}/J$ ideał $\widetilde{J'}$, złożony z wszystkich tych warstw $\widetilde{x}=x+J$, że $x \in J$. Odwrotnie, każdy ideał algebry ilorazowej $\mathcal{A}/J$ można otrzymać w ten sposób.

8.16. TWIERDZENIE. *Niech J będzie ideałem domkniętym w przemiennej algebrze Banacha $\mathcal{A}$. Algebra ilorazowa $\mathcal{A}/J$ jest izometrycznie izomorficzna ciału liczb zespolonych wtedy i tylko wtedy, gdy J jest ideałem maksymalnym.*

Dowód: Jeśli ideał J nie jest maksymalny, to jest właściwym podzbiorem pewnego ideału. Wtedy algebra ilorazowa $\mathcal{A}/J$ posiada nietrywialny ideał i nie może być ciałem. Z drugiej strony, gdy ideał J jest maksymalny, to $\mathcal{A}/J$ nie zawiera nietrywialnych ideałów, a więc jest ciałem. Teza wynika z Twierdzenia Mazura-Gelfanda. $\square$

Niech $\mathfrak{M}$ oznacza zbiór wszystkich ideałów maksymalnych przemiennej algebry Banacha $\mathcal{A}$. Z twierdzenia 8.16 wynika, że dla każdego $M \in \mathfrak{M}$ i każdego $x \in \mathcal{A}$ istnieje taka liczba zespolona $x(M)$, że $x + M = x(M)\,e + M$. Jest jasne, że odwzorowania $x \to x(M)$ jest homomorfizmem algebry $\mathcal{A}$ w ciało liczb zespolonych $\mathbb{C}$. Homomorfizm ten jest ciągły, bo $|x(M)| \leqslant \|x + M\| \leqslant \|x\|$.

8.17. LEMAT. *Niech μ będzie niezerowym homomorfizmem przemiennej algebry Banacha $\mathcal{A}$ w ciało liczb zespolonych i niech*

$$M_\mu = \{x \in \mathcal{A} : \mu(x) = 0\}$$

oznacza jego jądro. Wtedy M_μ jest ideałem maksymalnym algebry $\mathcal{A}$ i $x(M_\mu) = \mu(x)$ dla $x \in \mathcal{A}$. Odwzorowanie $\mu \to M_\mu$ między zbiorem wszystkich niezerowych homomorfizmów i zbiorem $\mathfrak{M}$ wszystkich ideałów maksymalnych jest wzajemnie jednoznaczne.

Dowód: Ponieważ element $\mu(x)\,e - x$ leży w ideale M_μ, to jest singularny i dlatego liczba $\mu(x)$ leży w widmie $\sigma(x)$. Z lematu 8.9 wynika, że $|\mu(x)| \leqslant \|x\|$. Oznacza to, że homomorfizm μ jest ciągły, ideał M_μ domknięty a $\mathcal{A}/M_\mu$ jest algebrą Banacha. Homomorfizm μ jest stały na każdej warstwie $x + M_\mu$ i dlatego wyznacza pewien homomorfizm $\widetilde{\mu}$ algebry ilorazowej $\mathcal{A}/M_\mu$ w ciało liczb zespolonych $\mathbb{C}$. Ponieważ $\mu(\alpha e) = \alpha$, to $\widetilde{\mu}$ odwzorowuje $\mathcal{A}/M_\mu$ na całe $\mathbb{C}$. Wynika stąd, że $\mathcal{A}/M_\mu$ jest ciałem, a w konsekwencji, że ideał M_μ jest maksymalny. Równość $\mu(x) = x(M_\mu)$ wynika bezpośrednio z określenia $x(M_\mu)$. Odwzorowanie $\mu \to M_\mu$ jest wzajemnie jednoznaczne, bo jeśli $M_{\mu_1} = M_{\mu_2}$, to $\mu_1(x) = x(M_{\mu_1}) = x(M_{\mu_2}) = \mu_2(x)$ dla wszystkich $x \in \mathcal{A}$. $\square$

8.18. WNIOSEK. *Każdy homomorfizm przemiennej algebry Banacha w ciało liczb zespolonych jest ciągły.*

8.19. LEMAT. *Niech $\mathfrak{M}$ oznacza zbiór wszystkich ideałów maksymalnych przemiennej algebry Banacha $\mathcal{A}$. Wtedy $x(\mathfrak{M}) = \sigma(x)$ oraz*

$$\sup_{M\in\mathfrak{M}} \left|x(M)\right| = \lim_{n\to\infty} \|x^n\|^{1/n}. \tag{8.45}$$

Dowód: Element $x(M)e - x$ leży w ideale maksymalnym M, dlatego jest singularny, więc $x(M) \in \sigma(x)$. Odwrotnie, jeżeli $\lambda \in \sigma(X)$, to element singularny $\lambda e - x$ leży w pewnym ideale maksymalnym M. Wtedy $\lambda = x(M)$. Równość (8.45) wynika z lematu 8.9. $\square$

DEFINICJA. Element x algebry Banacha $\mathcal{A}$ nazywamy **topologicznym nilpotentem**, jeśli $\|x^n\|^{1/n} \to 0$. **Radykał** $\mathcal{R}$ to zbiór wszystkich topologicznych nilpotentów algebry $\mathcal{A}$. Jeżeli $\mathcal{R} = \{0\}$, to algebra nosi nazwę **półprostej**.

8.20. LEMAT. *Radykał przemiennej algebry Banacha pokrywa się z przekrojem wszystkich jej ideałów maksymalnych.*

Dowód: Ponieważ $x(M) = 0$ wtedy i tylko wtedy, gdy $x \in M$, więc teza wynika z lematu 8.19. $\square$

DEFINICJA. **Przestrzenią strukturalną** przemiennej algebry Banacha $\mathcal{A}$ nazywamy zbiór $\mathfrak{M}$ wszystkich jej ideałów maksymalnych wyposażony w topologię, w której bazą otoczeń punktu $M_0 \in \mathfrak{M}$ tworzą zbiory

$$U(M_0, \varepsilon, F) = \left\{M \in \mathfrak{M} : \left|x(M) - x(M_0)\right| < \varepsilon \quad \text{dla} \quad x \in F\right\},$$

gdzie $\varepsilon > 0$ a F jest dowolnym skończonym podzbiorem $\mathcal{A}$. Zauważmy, że topologia ta, to najsłabsza topologia w $\mathfrak{M}$, w której wszystkie funkcje $M \to x(M)$ są ciągłe.

8.21. LEMAT. *Przestrzeń strukturalna $\mathfrak{M}$ przemiennej algebry Banacha $\mathcal{A}$ jest zwartą przestrzenią Hausdorffa i dla każdego $x \in \mathcal{A}$ funkcja $x(m)$ jest ciągła na $\mathfrak{M}$.*

Dowód: Oznaczmy przez Q_x domknięty dysk $\{\lambda \in \mathbb{C} : |\lambda| \leqslant \|x\|\}$ płaszczyzny zespolonej i niech $Q = \prod_{x\in\mathcal{A}} Q_x$ będzie produktem topologicznym wszystkich tych dysków. Ponieważ $|x(M)| \leqslant \|x\|$, to $x(M) \in Q_x$ dla każdego $x \in \mathcal{A}$ i dlatego każdemu $M \in \mathfrak{M}$ odpowiada pewien punkt $q \in Q$, mianowicie punkt o współrzędnych $q(x) = x(M)$. Różnym ideałom maksymalnym M_1 i M_2 odpowiadają w Q różne punkty, bo jeżeli $x \in M_1$, $x \notin M_2$, to $x(M_1) = 0 \neq x(M_2)$. Z twierdzenia Tichonowa (patrz **??**) wynika, że Q jest zwartą przestrzenią Hausdorffa, a

$\mathfrak{M}$ jest topologicznie równoważna pewnemu podzbiorowi przestrzeni Q. Dla zakończenia dowodu wystarczy pokazać, że $\mathfrak{M}$ jest domkniętym podzbiorem w Q. Niech $\lambda \in \overline{\mathfrak{M}}$, $\varepsilon > 0$, $F = \{x, y, x+y\}$, gdzie x, y są dowolnymi elementami $\mathcal{A}$. Otoczenie $U(\lambda, \varepsilon, F)$ punktu λ w Q ma niepusty przekrój z $\mathfrak{M}$, dlatego istnieje taki ideał maksymalny M, że

$$\big|\lambda(x) - x(M)\big| < \varepsilon, \qquad \big|\lambda(y) - y(M)\big| < \varepsilon,$$
$$\big|\lambda(x+y) - (x+y)(M)\big| < \varepsilon.$$

Ponieważ $(x+y)(M) = x(M) + y(M)$, a $\varepsilon > 0$ jest dowolne, to $\lambda(x+y) = \lambda(x) + \lambda(y)$. W ten sam sposób można dowieść, że $\lambda(e) = 1$, $\lambda(\alpha x) = \alpha\lambda(x)$ i $\lambda(xy) = \lambda(x)\lambda(y)$. Tak więc punkt $\lambda \in \overline{\mathfrak{M}}$ określa niezerowy homomorfizm algebry $\mathcal{A}$ w ciało liczb zespolonych. Z lematu 8.17 wynika istnienie takiego ideału maksymalnego M_μ w $\mathfrak{M}$, że $x(M_\mu) = \lambda(x)$, $x \in \mathcal{A}$. Tak więc $\lambda \in \mathfrak{M}$ tzn., że $\mathfrak{M}$ jest domkniętym zbiorem w Q. $\square$

8.22. TWIERDZENIE. *Niech $\mathfrak{M}$ będzie przestrzenią strukturalną przemiennej algebry Banacha $\mathcal{A}$, zaś $C(\mathfrak{M})$ algebrą Banacha wszystkich zespolonych funkcji ciągłych na $\mathfrak{M}$. Wtedy odwzorowanie $x \to x(\cdot)$ jest ciągłym homomorfizmem algebry $\mathcal{A}$ w $C(\mathfrak{M})$, przy czym $\sup_{M\in\mathfrak{M}} \big|x(M)\big| \leqslant \|x\|$. Odwzorowanie to jest wzajemnie jednoznaczne wtedy i tylko wtedy, gdy algebra $\mathcal{A}$ jest półprosta.*

Dowód: To, że odwzorowanie $x \to x(\cdot)$ jest homomorfizmem wynika z określenia $x(M)$. Z lematu 8.21 wynika, że $x(\cdot) \in C(\mathfrak{M})$. Nierówność $|x(M)| \leqslant \|x\|$ pokazana została w lemacie 8.19. Gdy $\mathcal{A}$ jest algebrą półprostą, to z lematu 8.20 wnosimy, że gdy $x(M) = 0$ dla wszystkich $M \in \mathfrak{M}$, to $x = 0$. Wtedy odwzorowanie $x \to x(\cdot)$ jest wzajemnie jednoznaczne. Odwrotnie, jeżeli to odwzorowanie jest wzajemnie jednoznaczne, to z lematu 8.20 wynika, że $\mathcal{A}$ jest półprosta. $\square$

8.23. PRZYKŁAD. Niech K będzie zwartą przestrzenią Hausdorffa. Pokażemy, że w algebrze $C(K)$ każdy funkcjonał liniowy i multiplikatywny μ dany jest wzorem

$$\mu(x) = x(t_0), \quad \text{gdzie} \quad t_0 \in K.$$

Pokażemy najpierw, że istnieje taki punkt $t_0 \in K$, że jeżeli $\mu(x) = 0$, to $x(t_0) = 0$. Istotnie, gdyby takiego punktu nie było, to dla każdego $t \in K$ istniałby taki element $x_t \in C(K)$, że $\mu(x_t) = 0$ oraz $x_t(t) \neq 0$. Kładąc w razie potrzeby $x_t\overline{x_t}$ w miejsce x_t, możemy założyć, że $x_t(s) \geqslant 0$ i $x_t(t) > 0$. Zbiory otwarte $U_t = \{s \in K : x_t(s) > 0\}$ pokrywają cała przestrzeń K, ze zwartości K wynika możliwość wybrania takich punktów $t_1, t_2, \ldots, t_n \in K$, że $K = \bigcup_{k=1}^n U_{t_k}$. Wynika stąd, że $y(s) = x_{t_1}(s) + x_{t_2}(s) + \ldots + x_{t_n}(s) > 0$ dla wszystkich $s \in K$, czyli, że y jest elementem odwracalnym w $C(K)$. Z drugiej strony $\mu(y) = \mu(x_{t_1}) + \mu(x_{t_2}) +$

$\ldots + \mu(x_{t_n}) = 0$, co prowadzi do sprzeczności. Istnieje więc żądany punkt t_0. Niech x będzie dowolnym elementem algebry $C(K)$. Jeżeli $\mu(x) = \alpha$, to $\mu(x - \alpha e) = 0$, skąd $(x - \alpha e)(t_0) = 0$, czyli $x(t_0) = \alpha = \mu(x)$. Widzimy więc, że $\mathfrak{M} = K$ i że odwzorowanie $x \to x(\cdot)$ jest po prostu tożsamością.

8.24. Przykład. Niech S będzie przestrzenią topologiczną lokalnie zwartą i niech $C_0(S)$ oznacza algebrę funkcji ciągłych na S, „znikających w nieskończoności”, tzn. takich funkcji $x : S \to \mathbb{C}$, że dla każdego $\varepsilon > 0$ istnieje taki zbiór zwarty $K \subset S$, że $|x(t)| < \varepsilon$ dla $t \notin K$. Rozpatrzmy algebrę Banacha $\mathcal{A}$ wszystkich funkcji „mających granicę w nieskończoności”, tzn. postaci $x(t) = \lambda + x_0(t)$, gdzie $\lambda \in \mathbb{C}$, $x_0 \in C_0(S)$, z normą $\|x\| = \sup_{t \in S} |x(t)|$. Dla takiej funkcji napiszemy $\lim_{t\to\infty} x(t) = \lambda$. Rozumując podobnie jak w poprzednim przykładzie pokazujemy, że jedynymi funkcjonałami liniowymi multyplikatywnymi algebry $\mathcal{A}$ są funkcjonały postaci

$$\mu(x) = x(t_0), \quad t_0 \in S, \quad \text{lub} \quad \mu(x) = \lim_{t\to\infty} x(t).$$

Oznacza to, że przestrzeń strukturalna $\mathfrak{M}$ algebry $\mathcal{A}$ pokrywa się ze zbiorem $S \cup \{\infty\}$. Jak poprzednio, łatwo pokazujemy, że topologia na przestrzeni S odziedziczona z $\mathfrak{M}$ pokrywa się z wyjściową topologią S. Oznacza to, że $\mathfrak{M}$ jest uzwarceniem jednopunktowym, tzw. **uzwarceniem Aleksandrowa**, przestrzeni S.

8.25. Przykład. Niech S będzie przestrzenią topologiczną całkowicie regularną. Pokażemy, że przestrzeń strukturalna $\mathfrak{M}$ algebry $C(S)$ jest homeomorficzna z uzwarceniem Čecha-Stone'a βS przestrzeni S.

Rzeczywiście, ponieważ S jest przestrzenią całkowicie regularną, to przyporządkowanie $t \to M_t = \{x \in C(S) : x(t) = 0\} \in \mathfrak{M}$ jest wzajemnie jednoznaczne. Niech $\mathfrak{M}_S$ oznacza obraz zbioru S przez to odwzorowanie. Pokażemy, że $\mathfrak{M}_S$ leży gęsto w $\mathfrak{M}$. Gdyby tak nie było, to w $\mathfrak{M}$ istniałoby otoczenie postaci

$$\big\{M \in \mathfrak{M} : \big|x_k(M) - x_k(M_0)\big| < \varepsilon,\ k = 1, 2, \ldots, n\big\}$$

rozłączne z $\mathfrak{M}_S$. Połóżmy $y_k = x_k - x_k(M_0)\, e$. Wtedy dla każdego $t \in S$ istnieje taki wskaźnik k, że $|y_k(t)| \geqslant \varepsilon$. Niech

$$y(t) = \sum_{k=1}^{n} y_k(t)\, \overline{y_k(t)}.$$

Wtedy $y(t) \geqslant \varepsilon^2$ dla wszystkich $t \in S$, więc dla y w $C(S)$ istnieje funkcja odwrotna y^{-1}. Wtedy y nie leży w żadnym ideale maksymalnym, choć $y(M_0) = 0$. Tu sprzeczność. Rozumowanie to pokazuje, że $\mathfrak{M}_S$ leży gęsto w $\mathfrak{M}$. Ponadto dla każdego $x \in C(S)$ funkcja $x(M)$ jest ciągła na $\mathfrak{M}$ i jest przedłużeniem na $\mathfrak{M}$ funkcji $x(M_t) = x(t)$. Dla zakończenia dowodu wystarczy więc pokazać, że odwzorowanie $t \to M_t$ jest homeomorfizmem S na $\mathfrak{M}_S$. Przeciwobrazem otoczenia

$$(a) \qquad \left\{ M_t \in \mathfrak{M}_S : \left| x_k(M_t) - x_k(M_{t_0}) \right| < \varepsilon,\ k = 1, 2, \ldots, n \right\}$$

przez to odwzorowanie jest zbiór

$$(b) \qquad \left\{ t \in S : \left| x_k(t) - x_k(t_0) \right| < \varepsilon,\ k = 1, 2, \ldots, n \right\},$$

który z powodu ciągłości funkcji x_k, $k = 1, 2, \ldots, n$ jest otwarty. Oznacza to, że rozpatrywane odwzorowanie jest ciągłe i wystarczy tylko sprawdzić, że jest otwarte. Zauważmy w tym celu, że każdy zbiór otwarty w S jest sumą zbiorów typu (b). Istotnie, jeżeli U jest otoczeniem punktu t_0 a $x \in C(S)$ jest funkcją oddzielającą punkt t_0 od zbioru $S \setminus U$, (tzn. $x(t_0) = 1$, $x(t) = 0$ dla $t \notin U$) to zbiór

$$\left\{ t \in S : \left| x(t) - x(t_0) \right| < \tfrac{1}{2} \right\}$$

jest otoczeniem punktu t_0 typu (b), zawartym u otoczeniu U.

8.26. Przykład. Wykażemy, że w algebrze $L^1(\mathbb{R})$ każdy funkcjonał multiplikatywny μ dany jest wzorem

$$\mu(x) = \int_{-\infty}^{+\infty} x(t)\, e^{its} dt, \quad \text{gdzie} \quad s \in \mathbb{R}.$$

Zauważmy najpierw, że $L^1(\mathbb{R})$ nie jest algebrą Banacha, bo nie posiada jedności, winniśmy więc badać algebrę $\mathcal{A} = \mathbb{C} \oplus L^1(\mathbb{R})$ powstałą przez dołączenie jedności. Każdy niezerowy funkcjonał multiplikatywny na $\mathcal{A}$ jest postaci

$$\mu(\lambda e + x) = \lambda + \mu(x).$$

Jednym spośród dopuszczalnych funkcjonałów jest $\mu_0(\lambda e + x) = \lambda$. Pozostałe są niezerowymi funkcjonałami multyplikatywnymi algebry $L^1(\mathbb{R})$. Oznacza to, że przestrzeń strukturalna $\mathfrak{M}$ algebry $\mathcal{A}$ można utożsamić ze zbiorem $\mathbb{R} \cup \{\mu_0\}$. Pozostawiamy jako ćwiczenie pokazanie, że $\mathfrak{M}$ jest uzwarceniem Aleksandrowa przestrzeni $\mathbb{R}$.

Funkcjonał μ, jako ciągły funkcjonał liniowy na przestrzeni $L^1(\mathbb{R})$ musi mieć postać

$$\mu(x) = \int_{-\infty}^{+\infty} x(t)\, \xi(t)\, dt,$$

gdzie $\xi \in L^{\infty}(\mathbb{R})$. Z własności multyplikatywności μ otrzymujemy warunek

$$\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x(t-s)\, y(s)\, ds \right) \xi(t)\, dt = \int_{-\infty}^{+\infty} x(t)\, \xi(t)\, dt \int_{-\infty}^{+\infty} y(s)\, \xi(s)\, ds.$$

Równość ta zachodzi dla wszystkich $x, y \in L^1(\mathbb{R})$. Zatem $\xi(t+s) = \xi(t)\,\xi(s)$ prawie wszędzie. Wobec mierzalności funkcji ξ, musi ona być postaci $\xi(t) = e^{zt}$ dla pewnego $z \in \mathbb{C}$. Ponieważ funkcja ξ jest istotnie ograniczona, więc z musi być czysto urojona, tzn. $z = is$, $s \in \mathbb{R}$. Oczywiście dla różnych s funkcjonały, które otrzymamy są różne. Ponieważ funkcja

$$\widehat{x}(s) = \int_{-\infty}^{+\infty} x(t)\, e^{its} dt$$

jest ciągłą funkcją zmiennej s, dążącą do zera w nieskończoności, więc istotnie przestrzeń strukturalną $\mathfrak{M}$ algebry $\mathcal{A}$ możemy utożsamić z uzwarceniem Aleksandrowa $\mathbb{R} \cup \{\infty\}$ prostej $\mathbb{R}$. Punkt ∞ jest tu reprezentowany przez funkcjonał μ_0.

8.27. PRZYKŁAD. Stosując podobną technikę można pokazać, że każdy funkcjonał multyplikatywny μ na algebrze $\ell^1(\mathbb{Z})$ wszystkich ciągów sumowalnych $x = (\ldots, x_{-1}, x_0, x_1, \ldots)$ ze splotem, jest postaci

$$\mu(x) = \sum_{-\infty}^{+\infty} x_n\, e^{int},$$

gdzie t jest liczbą rzeczywistą (spełniającą warunek $0 \leqslant t < 2\pi$, by odpowiedniość była wzajemnie jednoznaczna).

DEFINICJA. Powiemy, że algebra Banacha $\mathcal{A}$ jest **generowana** przez zbiór $A \subset \mathcal{A}$, jeżeli najmniejszą domkniętą podalgebrą z jedynką e i zawierającą A jest $\mathcal{A}$. Zbiór A nazywamy wtedy **zbiorem generatorów** algebry $\mathcal{A}$.

8.28. PRZYKŁAD. Algebra dyskowa $\mathcal{A}(D)$ jest generowana przez zbiór jednoelementowy, złożony z funkcji identycznościowej $f_0(z) = z$. Algebra $C(T)$ jest generowana przez zbiór $\{f_0, \overline{f_0}\}$. Wynika to z twierdzenia Stone'a-Weierstrassa. Żaden zbiór jednoelementowy nie generuje $C(T)$.

8.29. TWIERDZENIE. *Przestrzeń strukturalna przemiennej algebry Banacha $\mathcal{A}$ generowanej przez zbiór $A \subset \mathcal{A}$ jest homeomorficzna z domkniętym podzbiorem produktu zbiorów $\prod \sigma(a)$, gdzie a przebiega zbiór A.*

Dowód: Zgodnie z lematem 8.19, $x(\mathfrak{M}) = \sigma(x)$ i dlatego przyporządkowanie $M \to x(M)$ określa odwzorowanie przestrzeni $\mathfrak{M}$ w $\prod_{x\in A}\sigma(x)$. Odwzorowanie to jest ciągłe, bo ciągłe są wszystkie funkcje $x(\cdot)$. Załóżmy, że $x(M_1) = x(M_2)$ dla wszystkich $x \in A$. Ponieważ zbiór A generuje $\mathcal{A}$, to $y(M_1) = y(M_2)$ dla wszystkich $y \in \mathcal{A}$, stąd $M_1 = M_2$, bo jeżeli $y \in M_1$ oraz $y \notin M_2$, to $y(M_1) = 0 \neq y(M_2)$. Rozpatrywane odwzorowanie jest więc ciągłe i wzajemnie jednoznaczne z przestrzeni zwartej Hausdorffa $\mathfrak{M}$ w (zwartą) przestrzeń Hausdorffa $\prod_{x\in A}\sigma(x)$, jest wobec tego homeomorfizmem a obraz $\mathfrak{M}$ jest domknięty w $\prod_{x\in A}\sigma(x)$.

8.30. WNIOSEK. *Przestrzeń strukturalna przemiennej algebry Banacha z jednym generatorem jest homeomorficzna ze spektrum tego generatora.*

Przemienne C*-algebry

C*-**algebrą** nazywamy algebrę Banacha z inwolucją *, spełniającą dodatkowo warunek $\|x^*x\| = \|x\|^2$.

Jest jasne, że algebra $C(K)$, gdzie K jest przestrzenią zwartą, jest C*-algebrą przemienną, jeżeli za inwolucję $*$ wybrać zwykłe sprzężenie zespolone $x^*(t) = \overline{x(t)}$. Mamy wtedy $\|x^*x\|_\infty = \|\,|x|^2\|_\infty = \|x\|_\infty^2$. Innym dobrze znanym nam przykładem C*-algebry, tym razem nieprzemiennej, jest algebra $\mathcal{L}(\mathcal{H})$ wszystkich ograniczonych operatorów na przestrzeni Hilberta $\mathcal{H}$. Inwolucja, to sprzęganie operatorów $\langle T^*x, y\rangle = \langle x, Ty\rangle$, $x, y \in \mathcal{H}$. Wiemy, że $\|T^*T\| = \|T\|^2$ dla każdego $T \in \mathcal{L}(\mathcal{H})$.

Niech T będzie operatorem normalnym na przestrzeni Hilberta $\mathcal{H}$, zaś $\mathcal{A}$ najmniejszą domkniętą podalgebrą w $\mathcal{L}(\mathcal{H})$ zawierającą T, T^* oraz I. Wtedy $\mathcal{A}$ jest przemienną C*-algebrą.

W algebrze dyskowej $\mathcal{A}(D)$ można wprowadzić inwolucję kładąc $x^*(z) = \overline{x(\overline{z})}$. Jednak $\mathcal{A}(D)$ nie jest C*-algebrą. Mamy bowiem dla funkcji $x(z) = z+i$; $\|x^*x\| = \sup_{z\in D}|z^2+1| = 2$, natomiast $\|x\| = \sup_{z\in D}|z+i| = 2$.

8.31. LEMAT. *Jeżeli $\mathcal{A}$ jest przemienną algebrą Banacha, to $\|x^2\| = \|x\|^2$ i $\|x^*\| = \|x\|$ dla każdego $x \in \mathcal{A}$, a także $e^* = e$.*

Dowód: Mamy

$$\|x^2\|^2 = \big\|(x^2)^*x^2\big\| = \big\|(x^*)^2x^2\big\| = \big\|(x^*x)(x^*x)^*\big\| = \|x^*x\| = \|x\|^4.$$

Tak więc $\|x^2\| = \|x\|^2$. Dalej $\|x\|^2 = \|x^*x\| = \|xx^*\| = \|x^{**}x^*\| = \|x^*\|^2$, skąd $\|x^*\| = \|x\|$. W końcu $e^* = ee^* = e^{**}e^* = (e^*e)^* = e^{**} = e$. □

Definicja. Homomorfizm h z C*-algebry $\mathcal{A}$ w C*-algebrę $\mathcal{B}$ nazywamy *-**homomorfizmem**, jeśli zachowuje inwolucję tzn. $\big[h(x)\big]^* = h(x^*)$. Jeżeli *-homomorfizm jest izomorfizmem, to nazywamy go *-**izomorfizmem**.

8.32. Lemat (Arens). *Jeżeli $\mathfrak{M}$ oznacza przestrzeń strukturalną przemiennej C*-algebry $\mathcal{A}$, to odwzorowanie $x \to x(\cdot)$ algebry $\mathcal{A}$ w algebrę $C(\mathfrak{M})$ jest *-homomorfizmem.*

Dowód: Z twierdzenia 8.22 poprzedniego paragrafu wiemy, że odwzorowanie $x \to x(\cdot)$ jest homomorfizmem. Pozostaje więc wykazać, że $x^*(\lambda) = \overline{x(\lambda)}$ dla wszystkich $\lambda \in \mathfrak{M}$ oraz $x \in \mathcal{A}$. Niech $x(\lambda) = \alpha + i\beta$, zaś $x^*(\lambda) = \gamma + i\delta$, gdzie liczby $\alpha, \beta, \gamma\, d$ są rzeczywiste. Pokażemy, że założenie $\beta + \delta \neq 0$ prowadzi do sprzeczności. Połóżmy

$$y = \frac{1}{\beta+\delta}\Big(x + x^* - (\alpha+\gamma)e\Big).$$

Wtedy $y^* = y$ oraz $y(\lambda) = i$ a dla dowolnej liczby rzeczywistej m mamy

$$(y + i\,m\,e)(\lambda) = y(\lambda) + i\,m = i(1+m).$$

Wobec tego $|1+m| \leqslant \|y + i\,m\,e\|$, czyli

$$\begin{aligned}(1+m)^2 &\leqslant \|y + i\,m\,e\|^2 = \|(y + i\,m\,e)^*(y + i\,m\,e)\| = \|(y - i\,m\,e)(y + i\,m\,e)\ | \\ &= \|y^2 + m^2 e\| \leqslant \|y\|^2 + m^2.\end{aligned}$$

Nierówność ta nie może być prawdziwa dla wszystkich liczb rzeczywistych, np. dla $m = \|y\|^2$, tu sprzeczność. Dowiedliśmy w ten sposób, że $\beta + \delta = 0$, tzn. $x(\lambda) = \alpha + i\beta$, $x^*(\lambda) = \gamma - i\beta$. Ale wtedy $(ix)(\lambda) = i\,x(\lambda) = -\beta + i\alpha$, zaś $(ix)^*(\lambda) = -i\,x^*(\lambda) = -\beta - i\gamma$. Ma mocy udowodnionej już części twierdzenia wnosimy, że $\alpha - \gamma = 0$, tzn. $x^*(\lambda) = \overline{x(\lambda)}$. $\square$

8.33. Twierdzenie Gelfanda-Naimarka. *Przemienna C*-algebra jest izometrycznie *-izomorficzna algebrze $C(\mathfrak{M})$ wszystkich funkcji ciągłych na swojej przestrzeni strukturalnej.*

Dowód: Niech $\mathfrak{M}$ oznacza przestrzeń strukturalną algebry $\mathcal{A}$. Z lematu 8.31 wynika, że $\|x^m\| = \|x\|^m$, gdy m jest potęgą 2. Z lematu 8.19 poprzedniego paragrafu wynika, że

$$\sup_{\lambda\in\mathfrak{M}} \big|x(\lambda)\big| = \lim_{n\to\infty} \sqrt[n]{\|x^n\|} = \|x\|.$$

Wobec tego odwzorowanie $x \to x(\cdot)$ jest izometrią $\mathcal{A}$ w $C(\mathfrak{M})$ i $\mathcal{A}$ nie posiada radykału. Z lematu 8.32 wiemy, że odwzorowanie to jest *-homomorfizmem. Pozostaje dowieść, że jest ono „na” tzn., że każdej funkcji ciągłej na $\mathfrak{M}$ odpowiada pewien element $x \in \mathcal{A}$. Skorzystamy w tym celu z twierdzenia Stone'a. Oznaczmy przez $\mathcal{B}$ podalgebrę $C(\mathfrak{M})$ rozpiętą na wszystkich funkcjach $x(\cdot)$, $x \in \mathcal{A}$. Ponieważ $\|x\| = \sup_{\lambda \in \mathfrak{M}} \left|x(\lambda)\right|$, oraz przestrzeń $\mathcal{A}$ jest zupełna, więc $\mathcal{B}$ jest domkniętą podalgebrą w $C(\mathfrak{M})$. Niech λ_1, λ_2 będą dwoma różnymi ideałami maksymalnymi w $\mathfrak{M}$ i niech $y \in \lambda_1$, $y \notin \lambda_2$. Wtedy $y(\lambda_1) \neq y(\lambda_2)$, tzn. $\mathcal{B}$ rozdziela punkty $\mathfrak{M}$. Lemat 8.32 pokazuje, że spełnione są założenia twierdzenia Stone'a. Dlatego $\mathcal{B} = C(\mathfrak{M})$. □

8.34. Wniosek. *Każda przemienna C^*-algebra operatorów na przestrzeni Hilberta jest izometrycznie $*$-izomorficzna C^*-algebrze wszystkich funkcji ciągłych na swojej przestrzeni strukturalnej.*

8.35. Wniosek. *Jeżeli $\mathcal{B}$ jest domkniętą C^*-podalgebrą przemiennej C^*-algebry $\mathcal{A}$ i $\mathcal{B}$ i $\mathcal{A}$ mają wspólną jedynkę e, to element $y \in \mathcal{B}$ jest odwracalny w $\mathcal{A}$ wtedy i tylko wtedy, gdy jest odwracalny w $\mathcal{B}$. Wynika stąd, że widmo $\sigma_{\mathcal{B}}(x)$ elementu x względem algebry $\mathcal{B}$ pokrywa się z widmem $\sigma_{\mathcal{A}}(x)$ względem całej algebry $\mathcal{A}$.* □

Dowód: Jeżeli y jest odwracalny w $\mathcal{B}$, to tym bardziej jest odwracalny w $\mathcal{A}$, bo obie algebry mają wspólną jedynkę. Załóżmy teraz, że y jest odwracalny w $\mathcal{A}$. Ponieważ $(y^{-1})^* y^* = (yy^{-1})^* = e^* = e$, więc także y^* jest odwracalny w $\mathcal{A}$ a w konsekwencji odwracalny jest także element yy^*, Z twierdzenia Gelfanda-Naimarka wynika, że widmo elementu yy^* jest rzeczywiste, dlatego zbiór $\rho(yy^*)$ jest spójny. Z wniosku 8.12 wynika, że yy^* jest odwracalny w $\mathcal{B}$. □

8.36. Wniosek. *Niech x będzie elementem przemiennej C^*-algebry $\mathcal{A}$ i niech $C^*(x)$ oznacza najmniejszą domkniętą C^*-podalgebrą zawierającą x i jedynkę e algebry $\mathcal{A}$. Wtedy algebra $C^*(x)$ jest izometrycznie $*$-izomorficzna z algebrą $C(\sigma(x))$.*

Dowód: Z poprzedniego wniosku wynika, że widmo $\sigma(x)$ elementu x względem algebr $C^*(x)$ i $\mathcal{A}$ pokrywają się. Dlatego możemy założyć, że wielomiany od x, x^* i e leżą gęsto w $\mathcal{A}$, tzn. $\mathcal{A} = C^*(x)$. Niech $\mathfrak{M}$ oznacza przestrzeń strukturalną algebry $\mathcal{A}$. Mamy wtedy $\sigma(x) = x(\mathfrak{M})$. Rozważmy ciągłe odwzorowanie $\lambda \to x(\lambda)$ przestrzeni $\mathfrak{M}$ na $\sigma(x)$. Pokażemy, że jest ono wzajemnie jednoznaczne. Jeżeli $x(\lambda) = x(\lambda')$, to $x^*(\lambda) = \overline{x(\lambda)} = \overline{x(\lambda')} = x^*(\lambda')$, a więc $y(\lambda) = y(\lambda')$ dla każdego elementu y będącego wielomianem od x, x^* i e. Ponieważ elementy takiej postaci leżą gęsto w $\mathcal{A}$, otrzymujemy $y(\lambda) = y(\lambda')$ dla wszystkich $y \in \mathcal{A}$. Z twierdzenia Gelfanda-Naimarka wynika, że każda funkcja ciągła na $\mathfrak{M}$ przyjmuje jednakowe wartości w punktach λ i λ' i dlatego musi być $\lambda = \lambda'$. Reasumując, funkcja $\lambda \to x(\lambda)$ odwzorowuje w sposób ciągły i wzajemnie jednoznaczny przestrzeń zwartą $\mathfrak{M}$ na przestrzeń zwartą $x(\mathfrak{M}) = \sigma(x)$. Zbiory te są więc homeomorficzne. □

ROZDZIAŁ IX

TWIERDZENIE SPEKTRALNE

Twierdzenie spektralne, które udowodnimy, jest uogólnieniem znanego nam twierdzenia z algebry o sprowadzaniu macierzy normalnej $n \times n$ do postaci diagonalnej. Obiektem naszych rozważań będzie dowolny ograniczony operator normalny na przestrzeni Hilberta $\mathcal{H}$, tzn. operator spełniający równość $TT^* = T^*T$, gdzie T^* oznacza operator sprzężony, związany z T równością $\langle Tx,\, y\rangle = \langle x,\, T^*y\rangle$ dla wszystkich $x, y \in \mathcal{H}$. Zauważmy, że najmniejsza domknięta podalgebra z jednością w $\mathcal{L}(\mathcal{H})$ zawierająca T i T^* jest przemienną C*-algebrą. Z twierdzenia Gelfanda-Naimarka wynika, że jest izometrycznie *-izomorficzna z C*-algebrą $C\big(\sigma(T)\big)$ (patrz wniosek 8.36 z poprzedniego rozdziału). Pozwala to każdej funkcji ciągłej na $\sigma(T)$ przyporządkować operator działający na przestrzeni $\mathcal{H}$, przy czym funkcji $f(\lambda) = \lambda$ odpowiada operator T. Naszym celem będzie rozszerzenie tego odwzorowania na szerszą klasę funkcji — wszystkich funkcji borelowskich ograniczonych na $\sigma(T)$. Funkcjom charakterystycznym podzbiorów będą odpowiadały projektory ortogonalne. Rodzina tych projektorów nosi nazwę **miary spektralnej**. Jej związek z operatorem T odpowiada związkowi funkcji $f(\lambda) = \lambda$ z funkcjami charakterystycznymi podzbiorów widma $\sigma(T)$ i będzie opisany w twierdzeniu spektralnym.

Miara spektralna

Załóżmy, że dane jest ciało Σ podzbiorów ustalonego zbioru S oraz funkcja E odwzorowująca Σ w algebrą $\mathcal{L}(X)$ ciągłych operatorów liniowych na przestrzeni Banacha X. Załóżmy, że funkcja E jest **addytywna** i **ograniczona** tzn., że

$$E(\delta \cup \sigma) = E(\delta) + E(\sigma) \quad \text{oraz} \quad \|E(\sigma)\| \leqslant K \tag{9.46}$$

dla każdej pary rozłącznych zbiorów δ, σ w Σ i pewnej stałej K. Potem, w twierdzeniu spektralnym będziemy zakładali, że E jest miarą spektralną tzn., że X jest przestrzenią Hilberta a operatory $E(\delta)$ projektorami ortogonalnymi. Na razie jednak takie założenie nie jest potrzebne. Wymagamy jednak by funkcje, które będziemy całkować były ograniczone i Σ-mierzalne.

Zespoloną funkcję f na S nazywamy **Σ-mierzalną**, jeżeli $f^{-1}(A) \in \Sigma$ dla każdego borelowskiego podzbioru $A \subset \mathbb{C}$. Jeżeli Σ jest rodziną wszystkich podzbiorów borelowskich przestrzeni topologicznej S, to funkcje Σ-mierzalne noszą nazwę **borelowskich**. Przestrzeń $B(S, \Sigma)$ wszystkich ograniczonych i Σ-mierzalnych funkcji zespolonych na S jest domkniętą podprzestrzenią liniową w przestrzeni wszystkich zespolonych funkcji ograniczonych na S z normą $\| \|_\infty$. Zwróćmy uwagę, że funkcje charakterystyczne zbiorów z Σ tworzą zbiór liniowo gęsty w $B(S, \Sigma)$

Funkcję $f : S \to \mathbb{C}$ nazywamy **Σ-prostą**, gdy jest postaci

$$f = \sum_{k=1}^{n} a_k \chi_{\sigma_k}, \tag{9.47}$$

gdzie χ_{σ_k} jest funkcją charakterystyczną zbioru $\sigma_k \in \Sigma$. Z addytywności funkcji E łatwo wyprowadzić, że jeżeli $\sum_{i=1}^{n} \alpha_i \chi_{\sigma_i} = \sum_{j=1}^{m} \beta_j \chi_{\delta_j}$, to $\sum_{i=1}^{n} \alpha_i E(\sigma_i) = \sum_{j=1}^{m} \beta_j E(\delta_j)$. Dlatego **całkę** z funkcji Σ-prostej postaci (9.47) można określić równością

$$\int_S f(s)\, E(ds) = \sum_{i=1}^{n} \alpha_i\, E(\sigma_i).$$

Ponieważ pełne wahanie każdej skalarnej funkcji addytywnej zbiorów μ ma Σ nie przekracza $4 \sup_{\sigma \in \Sigma} \big|\mu(\sigma)\big|$, to dla każdej funkcji Σ-prostej f otrzymamy

$$\begin{aligned} \left| x^* \Big(\int_S f(s)\, E(ds) \Big) x \right| &= \left| \int_S f(s)\, x^* \big(E(ds0) x \right| \\ &\leqslant 4 \sup_{si \in S} \big| f(s) \big| \sup_{\sigma \in \Sigma} \big\| E(\sigma) \big\| \|x\| \|x^*\| \end{aligned}$$

dla dowolnych $x \in X$ oraz $x^* \in X^*$. W takim razie

$$\left\| \int_S f(s)\, E(ds) \right\| \leqslant 4K \sup_{s \in S} \big| f(s) \big|,$$

gdzie K jest stałą z nierówności (9.46). Nierówność ta dowodzi, że jeżeli ciąg $\{f_n\}$ funkcji Σ-prostych jest zbieżny w $B(S, \Sigma)$ do funkcji f, to ciąg całek $\big\{ \int_S f_n(s)\, E(ds) \big\}$ jest zbieżny w $\mathcal{L}(X)$ i jego granica zależy tylko od f, a nie od konkretnego wyboru ciągu $\{f_n\}$. Możemy zatem określić całkę z funkcji f wzorem

$$\int_S f(s)\, E(ds) = \lim_{n \to \infty} \int_S f_n(s)\, E(ds).$$

Dla zbioru $\sigma \in \Sigma$ całkę $\int_\sigma f(s)\,E(ds)$ będziemy rozumieć jako $\int_S f(s)\,\chi_\sigma(s)\,E(ds)$. Jest jasne, że odwzorowanie $f \to \int_S f(s)\,E(ds)$ jest ciągłym odwzorowaniem liniowym przestrzeni $\sigma B(S,\Sigma)$ w algebrę $\mathcal{L}(X)$, ograniczonych operatorów liniowych na X.

Jeżeli funkcja E jest miarą spektralną, to odwzorowanie $f \to \int_S f(s)\,E(ds)$ okazuje się być homomorfizmem algebr. Rzeczywiście, niech $f, g \in B(S,\Sigma)$. Zauważmy, że oba operatory $\int_S f(s)\,g(s)\,E(ds)$ i $\left(\int_S f(s)\,E(ds)\right) \times \left(\int_S g(s)\,E(ds)\right)$ w sposób liniowy i ciągły zależą od g i że jeżeli g jest funkcją charakterystyczną zbioru δ w Σ, to

$$\begin{aligned}\int_S f(s)g(s)\,E(ds) &= \int_S f(s)\,E(ds \cap \delta) = \int_S f(s)\,E(ds)E(\delta)\\ &= \left[\int_S f(s)\,E(ds)\right]\left[\int_S g(s)\,E(ds)\right].\end{aligned}$$

Wynika stąd, że jeżeli $f \in B(S,\Sigma)$, to równość

$$\int_S f(s)g(s)\,E(ds) = \left[\int_s f(s)\,E(ds)\right]\left[\int_S g(s)\,E(ds)\right]$$

zachodzi dla wszystkich funkcji Σ-prostych g, a w konsekwencji, z ciągłości obu stron względem g, dla dowolnych funkcji $f, g \in B(S,\Sigma)$, Jak widać, niezbędne jest, by funkcja E spełniała warunek

$$E(\sigma \cap \delta) = E(\sigma)\,E(\delta), \quad \text{dla} \quad \sigma, \delta \in \Sigma.$$

W szczególności wszystkie operatory $E(\sigma)$ muszą być projektorami na X. Jeżeli ponadto X jest przestrzenią Hilberta, a E funkcją samosprzężoną tzn. $E(\sigma)^* = E(\sigma)$ dla wszystkich $\sigma \in \Sigma$, to odwzorowanie $f \to \int_S f(s)\,E(ds)$ jest *-homomorfizmem C*-algebry $B(S,\Sigma)$ w C*-algebrę wszystkich ciągłych operatorów liniowych na przestrzeni Hilberta. Aby tego dowieść rozpatrzmy funkcję Σ-prostą f. Mamy tu

$$\left[\int_S f(s)\,E(ds)\right]^* = \sum_{i=1}^{n} \overline{\alpha_i}\,E(\delta_i) = \int_S \overline{f(s)}\,E(ds),$$

a ponieważ funkcje proste leżą gęsto w $B(S,\Sigma)$, otrzymujemy $\left[\int_S f(s)\,E(ds)\right]^* = \int_S \overline{f(s)}\,E(ds)$ dla wszystkich $f \in B(S,\Sigma)$.

Definicja. Niech Σ będzie ciałem podzbiorów ustalonego zbioru S a $\mathcal{H}$ przestrzenią Hilberta. Funkcję $E : \Sigma \to \mathcal{L}(\mathcal{H})$ nazywamy **miarą spektralną**, jeżeli

a. $E(\delta \cup \sigma) = E(\delta) + E(\sigma)$ dla rozłącznych δ, σ w Σ.

b. $F(\delta \cap \sigma) = E(\delta)\, E(\sigma)$ dla wszystkich δ, σ w Σ.
c. $E(\delta) = E(\delta)^*$ dla wszystkich $\delta \in \Sigma$.

Mamy wtedy

9.1. Twierdzenie. *Niech E będzie miarą spektralną w przestrzeni Hilberta $\mathcal{H}$, określoną na ciele Σ podzbiorów zbioru S. Wtedy odwzorowanie $f \to T(f)$, określone wzorem*

$$T(f) = \int_S f(s)\, E(ds), \qquad f \in B(S, \Sigma),$$

*jest ciągłym *-homomorfizmem C^*-algebry $B(S,\Sigma)$ ograniczonych, Σ-mierzalnych funkcji zespolonych na S w C^*-algebrę ograniczonych operatorów na przestrzeni Hilberta $\mathcal{H}$.* $\square$

Twierdzenie spektralne

Udowodnimy najpierw bardzo ogólne twierdzenie (ogólne twierdzenie spektralne). Zapowiadane wcześniej twierdzenie spektralne dla ograniczonych operatorów normalnych będzie jego szczególnym przypadkiem.

9.2. Ogólne twierdzenie spektralne. *Każda przemienna C^*-algebra $\mathcal{A}$ operatorów na przestrzeni Hilberta $\mathcal{H}$ jest izometrycznie *-izomorficzna z C^*-algebrą $C(\mathfrak{M})$ wszystkich funkcji ciągłych na swojej przestrzeni strukturalnej. Ponadto każdy izometryczny *-izomorfizm $f \to T(f)$ między tymi algebrami jednoznacznie określa miarę spektralną E na ciele $\mathcal{B}$ wszystkich podzbiorów borelowskich przestrzeni strukturalnej $\mathfrak{M}$ i posiadającą następujące własności:*

I. *dla dowolnych $x, y \in \mathcal{H}$ funkcja zbiorów $\langle E(\sigma)x,\, y\rangle$, $\sigma \in \mathcal{B}$ jest regularna i przeliczalnie addytywna;*

II. $E(\delta)T = TE(\delta)$, $\delta \in \mathcal{B}$, $T \in \mathcal{A}$;
III. $T(f) = \int_{\mathfrak{M}} f(\lambda)\, E(d\lambda)$, $f \in C(\mathfrak{M})$.

Dowód: Pierwsza część twierdzenia, to przytoczone twierdzenie Gelfanda-Naimarka. Dla dowodu drugiej potrzebny będzie lemat:

9.3. Lemat. *Każda ograniczona forma dwuliniowa $[x, y]$ na przestrzeni Hilberta $\mathcal{H}$, spełniająca warunek $[y, x] = \overline{[x, y]}$ jednoznacznie wyznacza taki ograniczony operator samosprzężony A, że $[x, y] = \langle Ax,\, y\rangle$.*

Dowód: Dla ustalonego $y \in \mathcal{H}$, wartość $[x, y]$ w sposób liniowy i ciągły zależy od x, a więc w $\mathcal{H}$ istnieje taki element Ay, że $[x, y] = \langle x,\, Ay\rangle$. Ponieważ forma $[x, y]$ jest ograniczona i dwuliniowa, więc operator A jest liniowy i ograniczony a równość $[x, y] = \overline{[y, x]}$ pokazuje, że jest on samosprzężony. $\square$

Kontynuując dowód twierdzenia zauważmy, że dla każdej pary wektorów x, y w $\mathcal{H}$ liczba $< T(f)\, x, y >$ zależy w sposób liniowy od f oraz $\left|\langle T(f)\, x,\, y\rangle\right| \leqslant \|f\|\|x\|\|y\|$. Wobec tego z twierdzenia Riesza o postaci funkcjonału na $C(\mathfrak{M})$ istnieje dokładnie jedna taka miara regularna $\mu_{x,y}$, że

$$\langle T(f)\, x,\, y\rangle = \int_{\mathfrak{M}} f(\lambda)\, \mu_{x,y}(d\lambda), \qquad f \in C(\mathfrak{M}), \tag{9.48}$$

$$\left|\mu_{x,y}(\sigma)\right| \leqslant |\mu_{x,y}|(\sigma) \leqslant \|x\|\|y\|, \qquad \sigma \in \mathcal{B}. \tag{9.49}$$

Ponieważ $\langle T(f)(\alpha x),\, y\rangle = \alpha\,\langle T(f)\, x,\, y\rangle$ to z (9.48) otrzymujemy

$$\int_{\mathfrak{M}} f(\lambda)\, \mu_{\alpha x,y}(d\lambda) = \int_{\mathfrak{M}} f(\lambda)\, \alpha\, \mu_{x,y}(d\lambda), \qquad f \in C(\mathfrak{M}),$$

a ponieważ miara jest jednoznacznie wyznaczona przez funkcjonał, który reprezentuje, więc $\mu_{\alpha x,y} = \alpha\, \mu_{x,y}$. Analogicznie pokazujemy, że $\mu_{x,y}$ w sposób dwuliniowy zależy od x i y. Jeżeli funkcja f jest rzeczywista, to $T(f) = T(\overline{f}) = T(f)^*$, a stąd $\langle T(f)\, x,\, y\rangle = \overline{\langle T(f)\, y,\, x\rangle}$, co z (9.48) daje

$$\int_{\mathfrak{M}} f(\lambda)\, \mu_{x,y}(d\lambda) = \int_{\mathfrak{M}} f(\lambda)\, \overline{\mu_{y,x}(d\lambda)}, \qquad f \in C(\mathfrak{M}),$$

a z jednoznaczności przedstawienia $\mu_{x,y} = \overline{\mu_{y,x}}$. Lemat 9.2 i nierówność (9.49) pokazują, że dla każdego $\delta \in \mathcal{B}$ istnieje dokładnie jeden taki operator samosprzężony $E(\delta)$ w $\mathcal{L}(\mathcal{H})$, że $\mu_{x,y}(\delta) = \langle E(\delta)\, x,\, y\rangle$. Jest jasne, że funkcja $\delta \to E(\delta)$ jest addytywna na $\mathcal{B}$, a z (9.49) wynika, że $\|E(\delta)\| \leqslant 1$. Dlatego punkt III twierdzenia wynika z (9.48). Aby dowieść, że $E(\delta \cap \sigma) = E(\delta)\, E(\sigma)$ zauważmy, że dla każdej pary f, g w $C(\mathfrak{M})$ mamy

$$\begin{aligned}
\int_{\mathfrak{M}} f(\lambda) \int_{\mathfrak{M}} g(\mu)\, E(d\lambda \cap d\mu) &= \int_{\mathfrak{M}} f(\lambda) \int_{d\lambda} g(\mu)\, E(d\mu) \\
&= \int_{\mathfrak{M}} f(\lambda) g(\lambda)\, E(d\lambda) = T(fg) = T(f)T(g) \\
&= \int_{\mathfrak{M}} f(\lambda)\, T(g)\, E(d\lambda) = \int_{\mathfrak{M}} f(\lambda) \int_{\mathfrak{M}} g(\mu)\, E(d\mu) E(d\lambda).
\end{aligned}$$

Wobec czego

$$\int_{\mathfrak{M}} g(\mu)\, E(d\mu \cap \delta) = \int_{\mathfrak{M}} g(\mu)\, E(d\lambda) E(\delta), \qquad \delta \in \mathcal{B},\ g \in C(\mathfrak{M}),$$

a dalej $E(\sigma \cap \delta) = E(\sigma)E(\delta)$ dla dowolnych $\delta, \sigma \in \mathcal{B}$. Wobec tego E jest miarą spektralną i w szczególności wszystkie projektory komutują. Z własności 3 wynika wtedy, że projektory ortogonalne $E(\delta)$ komutują także z $T(f)$, a to kończy dowód. $\square$

9.4. WNIOSEK. *Miara spektralna jest przeliczalnie addytywna w mocnej topologii operatorowej.*

Dowód: Jeżeli ciąg $\{\delta_n\}$ zbiorów borelowskich zstępuje do zbioru pustego, to na mocy własności I

$$\big\|E(\delta_n)\,x\big\|^2 = \langle E(\delta_n)\,x,\, E(\delta_n)x\rangle = \langle E(\delta_n)\,x,\, x\rangle \to 0. \quad \square$$

Przeprowadzone tu rozumowanie pokazuje także, że regularność miary skalarnej $\langle E(\delta)\,x,\, y\rangle$ dla dowolnych $x, y \in \mathcal{H}$ pociąga regularność miary wektorowej $E(\delta)\,x$ względem mocnej topologii w $\mathcal{H}$, tzn. dla dowolnych $\delta \in \mathcal{B}$ i $\varepsilon > 0$ istnieje taki zbiór domknięty $F \subset \delta$ i taki zbiór otwarty $G \supset \delta$, że $\big\|E(\delta)\,x\big\| < \varepsilon$ dla każdego $\sigma \in \mathcal{B}$ zawartego w $G \backslash F$. Wynika to z nierówności $\|E(\sigma)\,x\|^2 = \langle E(\sigma)\,x,\, x\rangle$. Dlatego dla miary spektralnej pojęcie przeliczalnej addytywności i regularności nie zależy od tego, czy rozpatrujemy je w słabej, czy w mocnej topologii.

9.5. WNIOSEK. *Każdy ograniczony operator normalny jednoznacznie wyznacza na podzbiorach borelowskich płaszczyzny zespolonej $\mathbb{C}$ przeliczalnie addytywną regularną miarę spektralną E, zerującą się na $\rho(T)$ i mającą własność*

$$f(T) = \int_{\sigma(T)} f(\lambda)\,E(d\lambda), \qquad f \in C(\mathfrak{M}). \quad \square$$

DEFINICJA. Miara spektralna, o której mowa we wniosku 9.5 nosi nazwę **rozkładu jedności** dla operatora normalnego T.

9.6. WNIOSEK. *Regularna, przeliczalnie addytywna miara spektralna E, określona na podzbiorach borelowskich płaszczyzny zespolonej jest rozkładem spektralnym jedności dla operatora normalnego T wtedy i tylko wtedy, gdy*

$$T = \int_{\sigma(T)} \lambda\,E(d\lambda).$$

Dowód: Jeżeli $T = \int_{\sigma(T)} \lambda\,E(d\lambda)$, to $T^* = \int_{\sigma(T)} \overline{\lambda}\,E(d\lambda)$, więc dla każdego wielomianu $p(\lambda, \overline{\lambda})$ zmiennych λ i $\overline{\lambda}$ mamy $p(T, T^*) = \int_{\sigma(T)} p(\lambda, \overline{\lambda})\,E(d\lambda)$. Z twierdzenia Weierstrassa wynika wtedy, że $f(T) = \int_{\sigma(T)} f(\lambda)\,E(d\lambda)$ dla każdej funkcji $f \in C\big(\sigma(T)\big)$. Wobec tego z wniosku 9.5 wynika, że E jest rozkładem jedności dla T. Przeciwna implikacja została dowiedziona we wniosku 9.5. $\square$

9.7. Wniosek. *Niech E będzie rozkładem jedności dla ograniczonego operatora normalnego T. Dla dowolnej ograniczonej, zespolonej funkcji borelowskiej f na widmie $\sigma(T)$ połóżmy*

$$f(T) = \int_{\sigma(T)} f(\lambda)\, E(d\lambda).$$

Wtedy odwzorowanie $f \to f(T)$ jest ciągłym $*$-izomorfizmem C^*-algebry ograniczonych funkcji borelowskich na $\sigma(T)$ w C^*-algebrę $\mathcal{L}(\mathcal{H})$, przy czym funkcji $f(\lambda) = \lambda$ i $f(\lambda) = 1$ odpowiadają operatory T oraz I. Homomorfizm ten ma jeszcze własności:

(i). $\|f(T)\, x\|^2 = \int_{\sigma(T)} |f(\lambda)|^2 \langle E(d\lambda)\, x,\, x\rangle,\ x \in \mathcal{H}$*;*

(ii). *jeżeli ograniczony ciąg funkcji borelowskich jest jednocześnie zbieżny w każdym punkcie widma do funkcji f, to $f_n(T) \to f(T)$ w silnej topologii operatorowej.* $\square$

ROZDZIAŁ X

DODATEK

Przestrzenie liniowe

10.1. OKREŚLENIE PRZESTRZENI LINIOWEJ. Zbiór X nazywamy **przestrzenią liniową**, jeśli:

1. dla dowolnych dwóch elementów x, $y \in X$ jest określona ich suma $x + y$, także będąca elementem zbioru X;

2. dla dowolnej liczby zespolonej α i dowolnego elementu $x \in X$ określony jest ich iloczyn αx, także będący elementem zbioru X.

Zakładamy ponadto, że operacje dodawania elementów i mnożenia elementu przez liczbę spełniają następujące warunki:

- $x + y = y + x$;
- $(x + y) + z = x + (y + z)$;
- w X istnieje element 0 o tej własności, że $x + 0 = x$ dla każdego elementu $x \in X$;
- dla każdego elementu $x \in X$ istnieje element $-x$, o własności $x + (-x) = 0$;
- $1\,x = x$;
- $\alpha(\beta x) = (\alpha\beta)x$;
- $(\alpha + \beta)x = \alpha x + \beta x$;
- $\alpha(x + y) = \alpha x + \alpha y$.

Elementy przestrzeni liniowej nazywamy na ogół **wektorami**. Jeśli w zbiorze X określone jest mnożenie tylko przez liczby rzeczywiste, to X nazywamy **rzeczywistą** przestrzenia liniową. Oczywiście każdą (zespoloną) przestrzeń liniową można także traktować jako rzeczywista przestrzeń liniową.

PRZYKŁADY.

1. Oznaczmy przez $\mathbb{C}^n$ zbiór wszystkich układów $x = (x_1, x_2, \ldots, x_n)$ liczb zespolonych. Jeśli w zbiorze $\mathbb{C}^n$ określimy operacje dodawania i mnożenia przez liczbę zespoloną wzorem

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$
$$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n),$$

to staje się on przestrzenią liniową.

Zbiór tych wektorów $x = (x_1, x_2, \ldots, x_n)$, dla których wszystkie liczby x_k są rzeczywiste tworzy rzeczywistą przestrzeń liniową, oznaczaną $\mathbb{R}^n$.

Przestrzeń $\mathbb{C}^1$ jest, po prostu, zbiorem wszystkich liczb zespolonych, a $\mathbb{R}^1$ zbiorem wszystkich liczb rzeczywistych.

2. Zbiór $\boldsymbol{c}$ wszystkich zbieżnych ciągów liczb zespolonych $x = (x_1, x_2, x_3, \ldots)$ z podobnie określonymi operacjami dodawania ciągów i mnożenia ciągu przez liczbę zespoloną tworzy przestrzeń liniową.

3. Zbiór $C[a, b]$ wszystkich zespolonych funkcji ciągłych $x : [a, b] \to \mathbb{C}$ na ustalonym przedziale $[a, b]$, jest przestrzenią liniową ze zwykłym dodawaniem funkcji i mnożeniem funkcji przez liczbę. Zbiór funkcji ciągłych rzeczywistych na przedziale $[a, b]$ jest rzeczywistą przestrzenią liniową, oznaczaną $\mathbb{C}_{\mathbb{R}}[a, b]$.

10.2. Liniowa niezależność wektorów, bazy Hamela. **Kombinacją liniową** wektorów $x_1, x_2, \ldots, x_k$ przestrzeni liniowej X nazywamy każdą sumę postaci $\alpha_1 x_1 + \alpha_2 x_2 + \ldots \alpha_k x_k$. Wektory $x_1, x_2, \ldots, x_k$ nazywamy **liniowo niezależnymi** jeśli ich kombinacja liniowa $\alpha_1 x_1 + \alpha_2 x_2 + \ldots \alpha_k x_k$ staje się wektorem zerowym 0 tylko wtedy, gdy $\alpha_1 = \alpha_2 = \ldots = a_k = 0$. Jeśli wektory $x_1, x_2, \ldots, x_k$ nie są liniowo niezależnymi, to nazywamy je **liniowo zależnymi**. Zbiór $P \subset X$ nazywamy **liniowo niezależnym**, jeżeli elementy tworzące dowolny skończony podzbiór tego zbioru są liniowo niezależne. Zbiór liniowo niezależny $B \subset X$ nazywamy **bazą Hamela** przestrzeni X, jeżeli każdy wektor $x \in X$ można przedstawić w postaci skończonej kombinacji liniowej

$$x = \alpha_1 b_1 + \alpha_2 b_2 + \ldots \alpha_k b_k$$

wektorów zbioru B. Z liniowej niezależności zbioru B wynika, że przedstawienie takie jest jednoznaczne.

Każda przestrzeń liniowa ma bazę Hamela, a nawet więcej (patrz np. [1], Rozdział II, Twierdzenia 4.5, 4.3):

> *Każdy zbiór liniowo niezależny w przestrzeni liniowej można uzupełnić do bazy Hamela tej przestrzeni. Każde dwie bazy Hamela ustalonej przestrzeni liniowej są równej mocy.*

Moc bazy Hamela przestrzeni liniowej X oznaczamy $\dim X$ i nazywamy jej **wymiarem**.

Przestrzeń $\mathbb{C}^n$ ma wymiar n, a jej bazę Hamela tworzą na przykład wektory $e_1 = (1,0,0,\ldots,0)$, $e_2 = (0,1,0,\ldots,0)$,..., $e_n = (0,0,0,\ldots,1)$. Obie przestrzenie $\mathbf{c}$ i $C[a,b]$ mają wymiar nieskończony; w każdej z nich można wskazać nieskończone zbiory liniowo niezależne, w pierwszej zbiór ciągów $e_n = (0,0,\ldots,0,1,0,\ldots)$ z jedynką na n-tym miejscu, $n = 1,2,3,\ldots$, a w drugiej zbiór wszystkich jednomianów 1, t, t^2, $t^3,\ldots$.

W istocie bazy Hamela obu przestrzeni $\mathbf{c}$ oraz $C[a,b]$ muszą być nieprzeliczalne. W przestrzeni $\mathbf{c}$ ciągi postaci $\{n^{-s}\}_{n=1}^{\infty}$, $s \in \mathbb{R}^+$, tworzą nieprzeliczalny zbiór liniowo niezależny, bo gdy przy $0 \leqslant s_1 < s_2 < \ldots < s_k$ równość

$$\alpha_1 n^{-s_1} + \alpha_2 n^{-s_2} + \ldots + \alpha_k n^{-s_k} = 0$$

zachodzi dla wszystkich n, to mnożąc obie strony przez n^{s_1} i przechodząc do granicy przy $n \to \infty$ otrzymujemy $\alpha_1 = 0$. Usuwając z tej równości zerowy składnik $\alpha_1 n^{-s_1}$ i powtarzając rozumowanie dla wykładnika s_2 otrzymujemy $\alpha_2 = 0$, itd. W końcu po k-krokach otrzymamy $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. W przestrzeni $C[a,b]$ nieprzeliczalny zbiór liniowo niezależny tworzą funkcje postaci $x_s(t) = (t-a)^s$, $s \in \mathbb{R}^+$.

10.3. Podprzestrzenie liniowe, przestrzenie ilorazowe. Podzbiór X_0 przestrzeni liniowej X nazywamy **podprzestrzenią liniową**, gdy wraz z każdymi dwoma elementami x, y zawiera ich sumę $x+y$ i dla każdej liczby zespolonej α wraz z elementem x zawiera także element αx; a więc, gdy X_0 jest przestrzenią liniową z tymi samymi działaniami co w X. Dla dowolnego zbioru $A \subset X$ symbol $\operatorname{lin} A$ oznacza zbiór wszystkich skończonych kombinacji liniowych elementów zbioru A. Zatem $\operatorname{lin} A$ jest najmniejszą podprzestrzenią liniową w przestrzeni X, zawierającą zbiór A.

W przestrzeni $\mathbb{C}^n$ zbiór tych wektorów $x = (x_1, x_2, \ldots, x_n)$, dla których $x_1 + x_2 + \ldots + x_n = 0$ jest podprzestrzenią liniową. Jej wymiar wynosi $n-1$. W przestrzeni $\mathbf{c}$ podprzestrzeń liniową tworzą wszystkie ciągi zbieżne do zera a w przestrzeni $C[a,b]$ wszystkie wielomiany.

Niech X_0 będzie podprzestrzenią liniową przestrzeni X. Określmy w zbiorze X relację $\sim$ pisząc

$$x \sim y \quad \text{gdy} \quad x - y \in X_0.$$

Jest to relacja równoważności, zatem dzieli ona przestrzeń X na klasy abstrakcji, które będziemy nazywali **warstwami** i oznaczali

$$[x] = \{y \in X : y \sim x\} = \{x + z : z \in X_0\}.$$

Przestrzeń ilorazowa $X/\sim$, oznaczana dalej symbolem X/X_0, wszystkich warstw po wprowadzeniu operacji dodawania warstw i mnożenia warstwy przez skalar w sposób następujący

$$[x_1] + [x_2] = [x_1 + x_2], \qquad \alpha\,[x] = [\alpha\, x]$$

staje się przestrzenią liniową, noszącą nazwę **przestrzeni ilorazowej** X/X_0.

10.4. ODWZOROWANIA LINIOWE, IZOMORFIZM. Niech X, Y będą dwiema przestrzeniami liniowymi. Funkcję $T : X \to Y$ nazywamy **funkcją liniową**, **odwzorowaniem liniowym** lub **operatorem liniowym**, jeżeli dla dowolnych wektorów $x_1, x_2 \in X$

$$T(x_1 + x_2) = Tx_1 + Tx_2$$

oraz dla dowolnej liczby α

$$T(\alpha x) = \alpha\, T x.$$

Odwzorowanie liniowe z przestrzeni X w zbiór $\mathbb{C}$ liczb zespolonych (lub w zbiór $\mathbb{R}$ liczb rzeczywistych, gdy X jest rzeczywistą przestrzenią liniową) nazywamy **funkcjonałem liniowym**.

Dla $X = \mathbb{C}^n$, $Y = \mathbb{C}^m$ odwzorowanie $A : X \to Y$

$$A(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m),$$

gdzie

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, 2, \ldots, m,$$

zaś a_{ij} jest układem liczb zespolonych, jest liniowe. Każde odwzorowanie liniowe $A : \mathbb{C}^n \to \mathbb{C}^m$ jest takiej postaci. Macierz

$$a = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

nosi nazwę **macierzą odwzorowania** A.

Funkcja $x^* : \boldsymbol{c} \to \mathbb{C}$, określona dla $x = (x_1, x_2, x_3, \ldots) \in \boldsymbol{c}$ wzorem

$$x^*(x) = \lim_{n\to\infty} x_n$$

jest funkcjonałem liniowym na przestrzeni $\boldsymbol{c}$.

Odwzorowanie $V : C[a, b] \to C[a, b]$

$$Vx(t) = \int_a^t x(u)\, du, \qquad t \in [a, b],$$

jest operatorem liniowym na przestrzeni $C[a,b]$, zwanym **operatorem Volterry**.

Odwzorowanie liniowe T z przestrzeni liniowej X do przestrzeni liniowej Y jest jednoznacznie wyznaczone przez swoje wartości na wektorach dowolnej bazy Hamela przestrzeni X, mogą nimi być dowolne wektory przestrzeni Y. Zatem

Niech $\{b_t\}_{t\in T}$ będzie bazą Hamela przestrzeni liniowej X. Na to by odwzorowanie T z X do przestrzeni liniowej Y było liniowe potrzeba i wystarcza, aby istniała taka funkcja $\varphi : T \to Y$, że

$$Tx = \sum_{k=1}^{n} \alpha_k \, \varphi(t_k), \quad \textit{gdy} \quad x = \sum_{k=1}^{n} \alpha_k \, b_{t_k}.$$

Dwie przestrzenie liniowe X i Y nazywają się **algebraicznie izomorficzne**, gdy istnieje wzajemnie jednoznaczne odwzorowanie liniowe T przestrzeni X na przestrzeń Y. Odwzorowanie takie nosi nazwę **algebraicznego izomorfizmu**. Odwzorowanie odwrotne $T^{-1} : Y \to X$ jest także algebraicznym izomorfizmem.

Dla przestrzeni $X = \mathbb{C}^n$ i jej podprzestrzeni $X_0 = \{z \in \mathbb{C}^n : z_1 + z_2 + \ldots + z_n\}$ warstwa $[x]$ elementu $x \in \mathbb{C}^n$ składa się z tych wektorów $y \in \mathbb{C}^n$, dla których $y_1 + y_2 + \ldots + y_n = x_1 + x_2 + \ldots + x_n$. Zatem każdej warstwie $[x]$ można przyporządkować wzajemnie jednoznacznie liczbę zespoloną $x_1 + x_2 + \ldots + x_n$. Przyporządkowanie to jest algebraicznym izomorfizmem przestrzeni X/X_0 na przestrzeń $\mathbb{C}^1$.

10.5. Zbiory wypukłe. Odcinkiem łączącym punkty x_1, x_2 przestrzeni liniowej X nazywamy zbiór wszystkich punktów postaci $\lambda x_1 + (1-\lambda)x_2$, gdzie $0 \leqslant \lambda \leqslant 1$.

Zbiór $W \subset X$ jest **wypukły**, jeżeli wraz z każdymi dwoma punktami zawiera również odcinek łączący te punkty.

Przekrój dowolnej rodziny zbiorów wypukłych też jest zbiorem wypukłym. Wynika stąd, że dla dowolnego zbioru $P \subset X$ istnieje najmniejszy zbiór wypukły zawierający zbiór P. Jest on przekrojem rodziny wszystkich zbiorów wypukłych zawierających P. Zbiór ten oznaczamy $\operatorname{conv} P$ i nazywamy **uwypukleniem** zbioru P.

Jeżeli $P = \{x_1, x_2, \ldots, x_n\}$, to $\operatorname{conv} P$ jest zbiorem elementów postaci

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n, \tag{10.50}$$

gdzie wszystkie liczby λ_k są nieujemne oraz $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$. Istotnie, zbiór wszystkich elementów postaci (10.50) jest wypukły, wystarczy zatem pokazać, że zbiór $\operatorname{conv} P$ musi wszystkie takie elementy posiadać. Gdy $n = 2$, wynika to bezpośrednio z definicji zbioru wypukłego. Dla większych n dowód można przeprowadzić przez indukcję. Na przykład, gdy $\lambda_1+\lambda_2>0$ oraz $\lambda_1+\lambda_2+\lambda_3=1$, to $\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\in\operatorname{conv}P$, dlatego

$$\lambda_1x_1+\lambda_2x_2+\lambda_3x_3=(\lambda_1+\lambda_2)\Big(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1+\frac{\lambda_2}{\lambda_1+\lambda_2}x_2\Big)+\lambda_3x_3$$

także leży w $\operatorname{conv}P$.

Zbiór wypukły $W\subset X$ nazywamy **pochłaniającym**, jeżeli dla każdego elementu $x\in X$ istnieje taka liczba dodatnia λ, że $x\in\lambda W=\{\lambda w: w\in W\}$, tj. że $\frac{1}{\lambda}x\in W$. Funkcję $p:X\to\mathbb{R}^+$

$$p(x)=\inf\{\lambda>0: x\in\lambda W\}$$

nazywamy **funkcjonałem Minkowskiego** zbioru W.

Jest oczywiste, że gdy $x\in W$, to $p(x)\leqslant 1$, a gdy $p(x)<1$, to $x\in W$.

Funkcjonał Minkowskiego ma następujące własności:

1. $p(x)\geqslant 0$*;*

2. $p(x+y)\leqslant p(x)+p(y)$*;*

3. $p(\lambda x)=\lambda p(x)$ *dla* $\lambda\geqslant 0$.

Każda funkcja p na X spełniający powyższe warunki jest funkcjonałem Minkowskiego pewnego zbioru wypukłego pochłaniającego.

Własność pierwsza jest oczywista. Dla dowodu własności 2 zauważmy, że z definicji funkcjonału Minkowskiego wynika, iż dla dowolnego $\varepsilon>0$

$$\frac{1}{p(x)+\varepsilon}\,x\in W,\qquad \frac{1}{p(y)+\varepsilon}\,y\in W.$$

Połóżmy

$$\lambda=\frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon},$$

wtedy

$$1-\lambda=\frac{p(y)+\varepsilon}{p(x)+p(y)+2\varepsilon},$$

więc na mocy wypukłości zbioru W, do W należy element

$$\lambda\frac{1}{p(x)+\varepsilon}\,x+(1-\lambda)\frac{1}{p(y)+\varepsilon}\,y,$$

tj. element

$$\frac{1}{p(x)+p(y)+2\varepsilon}\,(x+y).$$

Stąd $p(x+y) \leqslant p(x)+p(y)+2\varepsilon$, a ponieważ ε mogło być dowolną liczbą dodatnią, więc $p(x+y) \leqslant p(x)+p(y)$. Równość 3 dla $\lambda = 0$ jest oczywista, a jeżeli $\lambda > 0$, to

$$\begin{aligned} p(\lambda x) &= \inf\{\alpha > 0 : \lambda x \in \alpha W\} = \inf\{\alpha > 0 : x \in \tfrac{\alpha}{\lambda} W\} \\ &= \inf\{\lambda\beta > 0 : x \in \beta W\} = \lambda\, p(x). \end{aligned}$$

Załóżmy teraz, że funkcjonał p na przestrzeni X spełnia warunki 1, 2 i 3. Określmy $W = \{x \in X : p(x) \leqslant 1\}$. Zbiór W jest wypukły, bo jeśli $x_1, x_2 \in W$ oraz $0 \leqslant \lambda \leqslant 1$, to

$$p\big(\lambda x_1 + (1-\lambda)x_2\big) \leqslant \lambda p(x) + (1-\lambda)p(y) \leqslant \lambda + (1-\lambda) = 1,$$

więc $\lambda x_1 + (1-\lambda)x_2 \in W$. Zbiór W jest też pochłaniający, bo $\dfrac{1}{p(x)+1}\, x \in W$ dla każdego $x \in X$. Wreszcie

$$p(x) = \inf\{\lambda > 0 : p(x) \leqslant \lambda\} = \inf\{\lambda > 0 : x \in \lambda W\},$$

czyli p jest funkcjonałem Minkowskiego zbioru W.

10.6. Twierdzenie Hahna-Banacha. To jedno z najważniejszych twierdzeń analizy funkcjonalnej. Podana poniżej postać tego twierdzenia jest bardzo ogólna, choć dotyczy tylko rzeczywistych przestrzeni liniowych. W zastosowaniach do przestrzeni unormowanych (rzeczywistych lub zespolonych) korzystać będziemy na ogół z innej, nieco słabszej, jego wersji

> Twierdzenie Hahna-Banacha. *Niech p będzie funkcjonałem Minkowskiego na rzeczywistej przestrzeni liniowej X oraz x_0^* funkcjonałem liniowym, określonym na pewnej podprzestrzeni liniowej $X_0 \subset X$ i spełniającym nierówność*
>
> $$x_0^*(x) \leqslant p(x) \quad \textit{dla} \quad x \in X_0.$$
>
> *Wtedy x_0^* można przedłużyć od funkcjonału liniowego x^*, określonego na całej przestrzeni X i spełniającego warunek*
>
> $$x^*(x) \leqslant p(x) \quad \textit{dla} \quad x \in X.$$

Dowód: Możemy oczywiście założyć, że $X_0 \neq X$. Weźmy dowolny element $x_0 \notin X_0$ i przyjmijmy

$$X_1 = \operatorname{lin}\big(X_0 \cup \{x_0\}\big).$$

X_1 jest więc zbiorem wszystkich elementów y postaci

$$y = x + \lambda\, x_0, \qquad x \in X_0,\ \lambda \in \mathbb{R}.$$

Zauważmy, że przedstawienie takie jest jednoznaczne. Rzeczywiście, jeżeli

$$y = x' + \lambda' x_0 = x'' + \lambda'' x_0 \quad \text{i} \quad \lambda' \neq \lambda'', \quad \text{to} \quad x_0 = \frac{1}{\lambda'' - \lambda'}(x' - x''),$$

a stąd $x_0 \in X_0$ wbrew założeniu. Wobec tego musi być $\lambda' = \lambda''$ i w konsekwencji $x' = x''$.

Określmy na przestrzeni X_1 funkcjonał x_1^* przyjmując

$$x_1^*(y) = x_1^*(x + \lambda x_0) = x_0^*(x) + \lambda c,$$

gdzie c jest dowolną stałą rzeczywistą. Funkcjonał x_1^* jest oczywiście liniowy i jest przedłużeniem x_0^*. Należy tylko dobrać c tak, aby

$$x_1^*(y) \leqslant p(y), \quad \text{tzn. aby} \quad x_0^*(x) + \lambda c \leqslant p(x + \lambda x_0).$$

Ostatnia z tych nierówności równoważna jest dwu następującym:

$$\begin{gathered} x_0^*\bigl(\tfrac{1}{\lambda} x\bigr) + c \leqslant p\bigl(x_0 + \tfrac{1}{\lambda} x\bigr) \quad \text{dla} \quad \lambda > 0, \\ x_0^*\bigl(\tfrac{1}{-\lambda} x\bigr) - c \leqslant p\bigl(-x_0 + \tfrac{1}{-\lambda} x\bigr) \quad \text{dla} \quad \lambda < 0. \end{gathered}$$

Aby spełnić te nierówności, dobieramy c tak, by

$$x_0^*(x') - p(x' - x_0) \leqslant c \leqslant p(x'' + x_0) - x_0^*(x'')$$

dla wszystkich x', $x'' \in X_0$. Wybór taki jest możliwy. Rzeczywiście,

$$\begin{aligned} x_0^*(x') + x_0^*(x'') &= x_0^*(x' + x'') \leqslant p(x' + x'') \\ &= p(x' - x_0 + x'' + x_0) \leqslant p(x' - x_0) + p(x'' + x_0), \end{aligned}$$

a stąd

$$\sup_{x' \in X_0} \bigl(x_0^*(x') - p(x' - x_0)\bigr) \leqslant \inf_{x'' \in X_0} \bigl(p(x'' + x_0) - x_0^*(x'')\bigr).$$

Rozpatrzmy zbiór F wszystkich funkcjonałów liniowych y^*, będących przedłużeniami x_0^* i takich, że $y^*(x) \leqslant p(x)$. Zbiór F można częściowo uporządkować przyjmując $y_1^* \prec y_2^*$, jeśli funkcjonał y_2^* jest przedłużeniem y_1^*. Z lematu Zorna wynika istnienie maksymalnego przedłużenia $x^* \in F$ funkcjonału x_0^*. Jest nim funkcjonał określony na całej przestrzeni X, w przeciwnym przypadku — na mocy tego, co udowodniliśmy wyżej — miałby on dalsze przedłużenie, wbrew definicji elementu maksymalnego. $\square$

Funkcjonał Minkowskiego p na (rzeczywistej lub zespolonej) przestrzeni liniowej X nosi nazwę półnormy, gdy $p(\lambda x) = |\lambda|\, p(x)$ dla wszystkich liczb (odpowiednio rzeczywistych lub zespolonych) λ i wszystkich $x \in X$.

Twierdzenia Hahna-Banacha. **(wersja zespolona)** *Niech p będzie półnormą na zespolonej przestrzeni liniowej X oraz x_0^* funkcjonałem liniowym, określonym na pewnej podprzestrzeni liniowej $X_0 \subset X$ i spełniającym nierówność*

$$\left|x_0^*(x)\right| \leqslant p(x) \quad dla \quad x \in X_0.$$

Wtedy x_0^* można przedłużyć od funkcjonału liniowego x^*, określonego na całej przestrzeni X i spełniającego warunek

$$\left|x^*(x)\right| \leqslant p(x) \quad dla \quad x \in X.$$

Dowód: Będziemy traktować X jako rzeczywistą przestrzeń liniową. Oznaczmy

$$x_0^*(x) = y_0^*(x) + iz_0^*(x),$$

gdzie y_0^* i z_0^* są odpowiednio częścią rzeczywistą i częścią urojoną funkcjonału x_0^*. Oczywiście y_0^* i z_0^* są rzeczywistymi funkcjonałami liniowymi na podprzestrzeni X_0, a ponieważ

$$i\big(y_0^*(x) + iz_0^*(x)\big) = ix_0^*(x) = x_0^*(ix) = y_0^*(ix) + i\,z_0^*(ix),$$

więc

$$z_0^*(x) = -y_0^*(ix) \quad \text{dla} \quad x \in X_0.$$

Z założenia wiemy, że $y_0^*(x) \leqslant \left|x_0^*(x)\right| \leqslant p(x)$ dla $x \in X_0$, zatem na mocy udowodnionej już wersji twierdzenia Hahna-Banacha funkcjonał y_0^* można przedłużyć do rzeczywistego funkcjonału liniowego y^* na całej przestrzeni X, z zachowaniem warunku $y^*(x) \leqslant p(x)$ dla $x \in X$.

Określmy teraz funkcjonał x^* wzorem

$$x^*(x) = y^*(x) - iy^*(ix).$$

Widać, że x^* jest przedłużeniem funkcjonału x_0^*. Widać również, że x^* jest funkcjonałem addytywnym i jednorodnym przy mnożeniu przez liczby rzeczywiste. Aby udowodnić jego jednorodność względem mnożenia przez liczby zespolone wystarczy zauważyć, że

$$x^*(ix) = y^*(ix) - iy^*(-x) = i\big(y^*(x) - iy^*(ix)\big) = ix^*(x).$$

Pozostaje dowieść, że $\left|x^*(x)\right| \leqslant p(x)$. Przedstawmy w tym celu liczbę $x^*(x)$ w postaci $x^*(x) = r\,e^{-i\theta}$. Wtedy $\left|x^*(x)\right| = e^{i\theta}x^*(x) = x^*\big(e^{i\theta}x\big)$, a ponieważ $x^*\big(e^{i\theta}x\big)$ jest liczbą rzeczywistą nieujemną, więc

$$\left|x^*(x)\right| = x^*\big(e^{i\theta}x\big) = y^*\big(e^{i\theta}x\big) \leqslant p\big(e^{i\theta}x\big) = p(x). \quad \square$$

Literatura.

[1] Andrzej Alexiewicz, *Analiza funkcjonalna*, PWN, Warszawa 1969.

Przestrzenie topologiczne

10.7. Pojęcie przestrzeni topologicznej. **Przestrzenią topologiczną** nazywamy dowolny zbiór X wraz z rodziną $\mathcal{T}$ jego podzbiorów o tej własności, że

T1. zbiór pusty $\emptyset$ oraz X należą do rodziny $\mathcal{T}$,

T2. jeżeli $U_1 \in \mathcal{T}$ i $U_2 \in \mathcal{T}$, to $U_1 \cap U_2 \in \mathcal{T}$,

T3. jeżeli $U_s \in \mathcal{T}$ dla każdego $s \in S$, to $\bigcup_{s \in S} U_s \in \mathcal{T}$.

Samą rodzinę $\mathcal{T}$ nazywamy **topologią**, a jej elementy **zbiorami otwartymi**. Dopełnienia zbiorów otwartych noszą nazwę **zbiorów domkniętych**. Dla jednego zbioru X można wybrać na ogół wiele rodzin $\mathcal{T}$ spełniających T1–T3. Przez wprowadzenie topologii rozumiemy wybranie jednej z nich. Jeżeli zachodzi inkluzja $\mathcal{T}_1 \subset \mathcal{T}_2$, to powiemy, że topologia $\mathcal{T}_1$ jest **słabsza** od topologii $\mathcal{T}_2$ lub, że topologia $\mathcal{T}_2$ jest **silniejsza** od topologii $\mathcal{T}_1$. Najsłabszą topologią w X jest rodzina $\{\emptyset, X\}$, zwana **topologią trywialną** a najsilniejszą rodzina wszystkich podzbiorów X, zwana **topologią dyskretną**.

Dowolny podzbiór X_0 przestrzeni topologicznej X można przekształcić w przestrzeń topologiczną przyjmując za zbiory otwarte w X_0 przekroje X_0 ze zbiorami otwartymi w X. Przestrzeń taką nazywamy **podprzestrzenią** przestrzeni topologicznej X, a jej topologię **topologią odziedziczoną**.

10.8. Wnętrze zbioru, otoczenia. Z każdym zbiorem A przestrzeni topologicznej związane są dwa zbiory: $\operatorname{int} A$, zwany **wnętrzem zbioru** A i $\overline{A}$ zwany **domknięciem zbioru** A. Pierwszy jest sumą wszystkich zbiorów otwartych zawartych w A, a więc największym zbiorem otwartym zawartym w A, a drugi przekrojem wszystkich zbiorów domkniętych zawierających A.

Zbiór A nosi nazwę **otoczenia punktu** x, jeżeli $x \in \operatorname{int} A$. O rodzinie $\mathcal{U}_x$ złożonej z otoczeń punktu x powiemy, że jest **bazą otoczeń punktu** x, jeżeli w każdym otoczeniu punktu x leży pewien zbiór z $\mathcal{U}_x$. Zauważmy, że punkt x leży w domknięciu $\overline{A}$ zbioru $A \subset X$ wtedy i tylko wtedy, gdy $U \cap A \neq \emptyset$ dla każdego otoczenia $U \in \mathcal{U}_x$. Jeżeli $\overline{A} = X$, to powiemy, że zbiór A jest **gęsty** w X. Jeżeli w X istnieje przeliczalny podzbiór gęsty to X nosi nazwę **przestrzeni ośrodkowej**, a sam zbiór nazwę **ośrodka**.

10.9. Przekształcenia przestrzeni topologicznych. Przekształcenie $f : X \to Y$ przestrzeni topologicznej X w przestrzeń topologiczną Y nazywamy **otwartym**, jeżeli obraz $f(U)$ każdego zbioru otwartego $U \subset X$ jest otwarty w Y. Jeżeli zaś dla każdego zbioru otwartego V w Y jego przeciwobraz $f^{-1}(V)$ jest otwarty w X, to powiemy, że f jest przekształceniem **ciągłym**. Ciągłość przekształcenia $f : X \to Y$ w punkcie x oznacza, że każde otoczenie V punktu $y = f(x)$ w Y zawiera obraz $f(U)$ pewnego otoczenia U punktu x. Widać stąd, że odwzorowanie $f : X \to Y$ jest ciągłe wtedy i tylko wtedy, gdy jest ciągłe w każdym punkcie $x \in X$. Wzajemnie jednoznaczne przekształcenie ciągłe i otwarte f przestrzeni X na przestrzeń Y nosi nazwę **homeomorfizmu**. Otwartość f to ciągłość przekształcenia odwrotnego f^{-1}.

10.10. Przestrzenie Hausdorffa. Przestrzeń topologiczną X nazywamy **przestrzenią Hausdorffa**, jeżeli każde dwa różne jej punkty posiadają otoczenia będące zbiorami rozłącznymi. Z określenia tego wynika, że w przestrzeni Hausdorffa każdy zbiór jednopunktowy jest domknięty. Przestrzeń Hausdorffa X, w której dla każdych dwóch rozłącznych zbiorów domkniętych A, $B \subset X$ istnieją rozłączne zbiory otwarte U, V, zawierające odpowiednio A i B, nosi nazwę **przestrzeni normalnej**. Jeżeli w przestrzeni Hausdorffa X dla każdego punktu x i każdego zbioru domkniętego A nie zawierającego punktu x istnieje taka funkcja ciągła $f : X \to [0,1]$, że $f(x) = 1$ oraz $f(A) = \{0\}$, to X nazywa się **przestrzenią całkowicie regularną**. Z lematu Urysohna, który za chwilę udowodnimy wynika, że każda przestrzeń normalna jest całkowicie regularna.

10.11. Lemat Urysohna. *Dla każdych dwóch rozłącznych zbiorów domkniętych A, B przestrzeni normalnej X istnieje taka funkcja ciągła $f : X \to [0,1]$, że $f(A) = \{0\}$ oraz $f(B) = \{1\}$.*

Dowód: Niech U, V będą rozłącznymi zbiorami otwartymi zawierającymi odpowiednio A i B. Oznaczmy zbiór U przez U_1 i rozpatrzmy parę A, $X - U_1$ zbiorów domkniętych rozłącznych. Z normalności przestrzeni X wynika istnienie takich rozłącznych zbiorów otwartych U_0 i V_0, że $A \subset U_0$ oraz $X - U_1 \subset V_0$. W szczególności mamy wtedy $\overline{U_0} \subset U_1$. W analogiczny sposób dla pary $\overline{U_0}$ i $X - U_1$ dowodzimy istnienia takiego zbioru otwartego $U_{1/2}$, że $\overline{U_0} \subset U_{1/2}$, $\overline{U_{1/2}} \subset U_1$. Później dla par $\overline{U_0}$, $X - U_{1/2}$ oraz $\overline{U_{1/2}}$, $X - U_1$ odpowiednio zbiorów $U_{1/4}$ i $U_{3/4}$. Powtarzając to rozumowanie wielokrotnie, dla każdej liczby r postaci $\frac{m}{2^n}$, $0 \leqslant m \leqslant 2^n$, określimy taki zbiór otwarty U_r, że $\overline{U_{r_1}} \subset U_{r_2}$ przy $r_1 < r_2$.

Przyjmijmy teraz $U_t = \bigcup_{r \leqslant t} U_r$ dla $t \in [0,1]$ oraz $U_t = \emptyset$ dla $t < 0$ i $U_t = X$ dla $t > 1$. Otrzymana rodzina $\{U_t\}_{t \in R}$ zbiorów otwartych ma tą własność, że $A \subset U_0$, $\overline{U_1} \subset X - B$ oraz

$$\overline{U_{t_1}} \subset U_{t_2}, \quad \text{gdy} \quad t_1 < t_2. \tag{10.51}$$

Istotnie, jeśli $0 \leqslant t_1 < t_2 \leqslant 1$, to wybierając liczby wymierne r_1 i r_2 postaci $\frac{m}{2^n}$ tak, aby $t_1 < r_1 < r_2 < t_2$ otrzymamy $\overline{U_{t_1}} \subset \overline{U_{r_1}} \subset U_{r_2} \subset U_{t_2}$., Jeśli zaś $t_1 < 0$ lub $t_2 > 1$ to inkluzja (10.51) jest spełniona, bo $U_{t_1} = \emptyset$ lub $U_{t_2} = X$.

Pokażemy, że funkcja

$$f(x) = \inf\{t : x \in U_t\}$$

spełnia wszystkie postulaty tezy. Ponieważ $U_t = X$, gdy $t > 1$ oraz $U_t = \emptyset$ gdy $t < 0$, więc $f(x)$ przyjmuje tylko wartości z przedziału $[0, 1]$. Łatwo sprawdzamy, że dla $x \in A$ wartość ta stale wynosi 0 a dla $x \in B$ stale 1. Do dowiedzenia pozostaje więc tylko, że f jest funkcją ciągłą. Obierzmy $x_0 \in X$, $\varepsilon > 0$ i oznaczmy $t_0 = f(x_0)$. Z (10.51) wynika, że $x_0 \in U_t$, gdy $t > t_0$ i $x_0 \notin \overline{U_t}$, gdy $t < t_0$. W szczególności x_0 leży w zbiorze otwartym $U = U_{t_0+\varepsilon} - \overline{U_{t_0-\varepsilon}}$. Dla każdego punktu $x \in U$ mamy $x \in U_{t_0+\varepsilon}$, skąd $f(x) \leqslant t_0 + \varepsilon$ oraz $x \notin \overline{U_{t_0-\varepsilon}}$, skąd $f(x) \geqslant t_0 - \varepsilon$. W rezultacie

$$f(x_0) - \varepsilon \leqslant f(x) \leqslant f(x_0) + \varepsilon. \quad \square$$

10.12. Przestrzenie zwarte. Rodzina zbiorów $\{U_\alpha\}_{\alpha\in A}$ nosi nazwę **pokrycia** zbioru X, jeżeli $X \subset \bigcup_{\alpha\in A} U_\alpha$. Przestrzeń topologiczną X nazywamy **zwartą**, jeżeli każde jej pokrycie zbiorami otwartymi zawiera pokrycie skończone tzn., gdy w zbiorze wskaźników A można wybrać taki skończony podzbiór $\alpha_1, a_2, \ldots, a_n$, że $X = \bigcup_{k=1}^{n} U_{\alpha_k}$.

Niepustą rodzinę $\{F_\alpha\}_{\alpha\in A}$ podzbiorów zbioru X nazywamy **scentrowaną**, jeżeli $\bigcap_{k=1}^{n} F_{\alpha_k} \neq \emptyset$ dla każdego skończonego zbioru wskaźników $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Przestrzeń topologiczna jest zwarta wtedy i tylko wtedy, gdy każda rodzina scentrowana jej podzbiorów ma wspólny punkt skupienia.

Jeśli rodzina $\{F_\alpha\}_{\alpha\in A}$ podzbiorów przestrzeni zwartej X nie ma wspólnego punktu skupienia, to $\bigcap_{\alpha\in A} \overline{F_\alpha} = \emptyset$. Dopełnienia $U_\alpha = X \setminus \overline{F_\alpha}$ zbiorów $\overline{F_\alpha}$ są otwarte i tworzą pokrycie przestrzeni X, zatem można wybrać taki skończony układ wskaźników $\alpha_1, a_2, \ldots, a_n$, że $X = \bigcup_{k=1}^{n} U_{\alpha_k}$. Wtedy jednak $\bigcap_{k=1}^{n} \overline{F_{\alpha_k}} = \emptyset$, a w konsekwencji także $\bigcap_{k=1}^{n} F_{\alpha_k} \neq \emptyset$, więc rodzina $\{F_\alpha\}_{\alpha\in A}$ nie jest scentrowana. Jeżeli przestrzeń X nie jest zwarta, to posiada rodzinę $\{U_\alpha\}_{\alpha\in A}$ podzbiorów otwartych pokrywających X, lecz takich, że X nie da się pokryć żadną skończoną ich ilością. Ich dopełnienia $F_\alpha = X \setminus U_\alpha$ tworzą scentrowaną rodzinę podzbiorów domkniętych X o pustym przekroju, a więc bez wspólnego punktu skupienia.

Zbiór $K \subset X$ nazywamy **zwartym**, gdy rozpatrywany jako podprzestrzeń w X, jest zwarty.

Domknięty podzbiór przestrzeni zwartej jest zbiorem zwartym.

Jeżeli rodzina zbiorów otwartych $\{U_\alpha\}_{\alpha\in A}$ pokrywa domknięty podzbiór K przestrzeni zwartej X, to uzupełniona o zbiór $X - K$ jest otwartym pokryciem całej przestrzeni X. Ze zwartości X wynika istnienie skończonej podrodziny $\{U_{\alpha_k}\}_{k=1}^{n}$, która wraz z $X - K$ pokrywa X. Podrodzina ta pokrywa wtedy K.

Obraz zbioru zwartego przez przekształcenie ciągłe jest zbiorem zwartym.

Załóżmy, że f jest ciągłym przekształceniem przestrzeni topologicznej X na przestrzeń topologiczną Y oraz, że K jest zbiorem zwartym w X. Jeżeli $\{V_\alpha\}_{\alpha\in A}$ jest otwartym pokryciem zbioru $f(K)$, to $\{f^{-1}(V_\alpha)\}_{\alpha\in A}$ jest otwartym pokryciem K. Zbiór K jest zwarty, więc pokrywa go też pewna skończona podrodzina $\{f^{-1}(V_{\alpha_k})\}_{k=1}^n$. Wobec tego rodzina $\{V_{\alpha_k}\}_{k=1}^n$ pokrywa zbiór $f(K)$.

Zbiór zwarty w przestrzeni Hausdorffa jest domknięty.

Chcemy pokazać, że jeżeli punkt y nie leży w zbiorze zwartym K, to posiada otoczenie V rozłączne z K. Dla każdego punktu $x \neq y$ wybierzmy parę rozłącznych zbiorów otwartych U_x i V_x zawierających punkty x i y. Ponieważ zbiory $\{U_x\}_{x\neq y}$ pokrywają nasz zbiór zwarty K, więc pokrywa K też pewna podrodzina $\{U_{x_k}\}_{k=1}^n$. Zbiór $V = \bigcap_{k=1}^n V_{x_k}$ jest otwartym otoczeniem punktu y, rozłącznym ze zbiorem $U = \bigcup_{k=1}^n U_{x_k}$, a tym bardziej rozłącznym z K.

Zwarta przestrzeń Hausdorffa jest przestrzenią normalną.

Niech K_1, K_2 będą rozłącznymi zbiorami domkniętymi w przestrzeni Hausdorffa X. Są wtedy zbiorami zwartymi. Jak pokazaliśmy przed chwilą, dla każdego punktu y leżącego poza zbiorem K_1 istnieją rozłączne zbiory otwarte U i V zawierające odpowiednio y i K_1. Dla zaznaczenia zależności od y oznaczmy je U_y i V_y. Dalej rozumujemy dokładnie tak jak poprzednio. Ponieważ zbiory $\{U_y\}_{y\notin K_1}$ pokrywają zbiór K_2, więc pokrywa go też pewna podrodzina $\{U_{y_k}\}_{k=1}^n$. Zbiory $\bigcup_{k=1}^n U_{y_k}$ oraz $\bigcap_{k=1}^n V_{y_k}$ są otwarte, rozłączne i zawierają odpowiednio K_1, K_2. Wnosimy stąd, że X jest przestrzenią normalną.

10.13. PRZESTRZENIE LOKALNIE ZWARTE. Przestrzeń topologiczną X nazywamy **lokalnie zwartą**, jeżeli każdy punkt $x \in X$ ma otoczenie, którego domknięcie jest zbiorem zwartym. Przykładami przestrzeni lokalnie zwartych, poza przestrzeniami zwartymi, są prosta $\mathbb{R}$ i każda nieskończona przestrzeń dyskretna.

Przestrzeń lokalnie zwartą, która nie jest zwarta, można tak uzupełnić jednym punktem ∞, by stała się przestrzenią zwartą.

Oznaczmy $X_\infty = X \cup \{\infty\}$ i uznajmy za zbiory otwarte w X_∞ wszystkie zbiory otwarte w X, a także wszystkie zbiory postaci $U \cup \{\infty\}$, gdzie U jest zbiorem otwartym w X, którego dopełnienie jest zbiorem zwartym.

Widać, że X_∞ jest przestrzenią zwartą. Widać też, że gdy X jest przestrzenią Hausdorffa, to X_∞ jest także przestrzenią Hausdorffa. Istotnie, wystarczy zobaczyć, że dowolny punkt $x \in X$ oraz punkt ∞ posiadają otoczenia rozłączne. Jeżeli

U jest otoczeniem punktu x, którego domknięcie jest zbiorem zwartym, tu zbiór $(X \setminus \overline{U}) \cup \{\infty\}$ jest otoczeniem punktu ∞, rozłącznym z U.

Przestrzeń X_∞ nazywamy **uzwarceniem jednopunktowym**, lub **uzwarceniem Aleksandrowa** przestrzeni X.

Z powyższego rozumowania i lematu Urysohna wynika, że

> *W przestrzeni lokalnie zwartej Hausdorffa X dla dowolnego zbioru otwartego U i dowolnego zbioru zwartego $K \subset U$ istnieje taka funkcja ciągła $f : X \to [0,1]$, że $f(x) = 1$ na zbiorze K oraz $f(x) = 0$ poza zbiorem U.*

Wynika stąd w szczególności, że każda przestrzeń lokalnie zwarta Hausdorffa jest całkowicie regularna.

Niech X będzie dowolną przestrzenią lokalnie zwartą, lecz niezwartą. Dla funkcji ciągłej $f : X \to \mathbb{C}$ zapis

$$\lim_{x\to\infty} f(x) = A$$

oznacza, że dla każdej liczby $\varepsilon > 0$ zbiór $\{x \in X : \left|f(x) - A\right| \geqslant \varepsilon\}$ jest zwarty. Liczba A o powyższej własności (jeśli istnieje) jest wyznaczona jednoznacznie. O funkcji f mówimy wtedy, że ma **granicę w nieskończoności**. Tylko takie funkcje dają się przedłużyć do funkcji ciągłych na uzwarceniu Aleksandrowa X_∞. Jeśli $\lim_{x\to\infty} f(x) = 0$, to mówimy, że funkcja f **znika w nieskończoności** a zbiór wszystkich zespolonych funkcji ciągłych na X, znikających w nieskończoności oznaczamy $C_0(X)$.

10.14. Uzwarcenie Čecha-Stone'a. Jeśli X jest przestrzenią topologiczną całkowicie regularną, to daje się zanurzyć homeomorficznie w przestrzeń zwartą (patrz [1], 3.5.2). Prawdą jest też, że jeśli takie zanurzenie jest możliwe, to X jest przestrzenią całkowicie regularna. Takich zanurzeń przestrzeni X jest bardzo wiele. Interesują nas te, dla których obraz zbioru X leży gęsto. Takie zanurzenie nazywamy **uzwarceniem**. W poprzednim punkcie pokazaliśmy, że każda niezwarta przestrzeni lokalnie zwarta posiada uzwarcenie jednopunktowe. Pośród wszystkich uzwarceń ustalonej przestrzeni całkowicie regularnej istnieje dokładnie jedno uzwarcenie maksymalne $\beta : X \to \beta X$ (patrz [1], 3.5.10), zwane **uzwarceniem Čecha-Stone'a**. Maksymalność należy rozumieć w tym sensie, że jeżeli f jest ciągłym odwzorowaniem z przestrzeni X do przestrzeni zwartej Y, to istnieje takie odwzorowanie ciągłe F z przestrzeni βX do Y, że $f = F \circ \beta$.

Przestrzeń całkowicie regularną X wygodnie jest traktować jako gęstą podprzestrzeń swojego uzwarcenia βX.

Z maksymalności uzwarcenia Čecha-Stone'a wynika bezpośrednio (patrz [1], 3.6.1), że

Jeżeli X jest przestrzenią topologiczną całkowicie regularną oraz βX jest jej uzwarceniem Čecha-Stone'a, to każda ograniczona funkcja ciągła $f : X \to \mathbb{C}$ ma jednoznaczne przedłużenie do funkcji ciągłej $\widetilde{f} : \beta X \to \mathbb{C}$.

10.15. Słaba topologia, wyznaczona przez rodzinę funkcji. Niech X będzie dowolnym zbiorem oraz $\{f_\alpha\}_{\alpha\in A}$ rodziną funkcji ze zbioru X do przestrzeni topologicznych Y_α. Jako bazę zbiorów otwartych w X weźmy rodzinę wszystkich skończonych przekrojów zbiorów postaci $f_\alpha^{-1}(U_\alpha)$, gdzie U_α jest zbiorem otwartym w przestrzeni Y_α. Topologię w X z tak wybraną bazą nazywamy **słabą topologią w X wyznaczoną przez rodzinę funkcji** $\{f_\alpha\}_{\alpha\in A}$.

Słaba topologia wyznaczona przez rodzinę funkcji $\{f_\alpha\}_{\alpha\in A}$ to najsłabsza topologia, w której wszystkie funkcje f_α są ciągłe.

10.16. Topologiczny produkt przestrzeni. Załóżmy, że dana jest rodzina $\{X_\alpha\}_{\alpha\in A}$ przestrzeni topologicznych. Oznaczmy przez $X = \prod_{\alpha\in A} X_\alpha$ zbiór wszystkich takich funkcji $x = \{x_\alpha\}_{\alpha\in A}$ na zbiorze A, że $x_\alpha \in X_\alpha$. Element x_α nazywamy α-tą **współrzędną punktu** x. Odwzorowanie $\pi_\alpha : X \to X_\alpha$, przyporządkowujące punktowi x jego α-tą współrzędną x_α, nazywamy **rzutem na oś** X_α. Słabą topologię w przestrzeni X wyznaczoną przez rodzinę funkcji $\{\pi_\alpha\}_{\alpha\in A}$ nazywamy **topologią produktową** w przestrzeni X. Jest to więc najsłabsza topologia w X, dla której wszystkie rzuty π_α są ciągłe. Bazę topologii produktowej tworzą wszystkie zbiory postaci $\prod_{\alpha\in A} U_\alpha$, gdzie U_α jest zbiorem otwartym w przestrzeni X_α, lecz $U_\alpha \neq X_\alpha$ może zachodzić tylko dla skończenie wielu wskaźników α.

Produkt topologiczny przestrzeni Hausdorffa jest przestrzenią Hausdorffa.

Jeżeli $x, y \in \prod_{\alpha\in A} X_\alpha$ oraz $x \neq y$, to $x_\alpha \neq y_\alpha$ dla pewnego $\alpha \in A$. Ponieważ X_α jest przestrzenią Hausdorffa, więc istnieją w niej rozłączne otoczenia U_α i V_α

punktów x_α i y_α. Zbiory $\pi_\alpha^{-1}(U_\alpha)$ i $\pi_\alpha^{-1}(V_\alpha)$ są wtedy rozłącznymi otoczeniami punktów x i y.

TWIERDZENIE TICHONOWA. *Produkt topologiczny przestrzeni zwartych jest przestrzenią zwartą.*

Dowód: Niech $X = \prod_{\alpha\in A} X_\alpha$ będzie produktem topologicznym przestrzeni zwartych X_α. Niech także $\mathcal{F}_0$ będzie dowolną scentrowaną rodziną podzbiorów X. Z lematu Kuratowskiego-Zorna wynika, że rodzina $\mathcal{F}_0$ jest zawarta w pewnej maksymalnej rodzinie scentrowanej $\mathcal{F}$ podzbiorów X. Dla dowodu twierdzenia wystarczy oczywiście pokazać, że wszystkie zbiory rodziny $\mathcal{F}$ mają wspólny punkt skupienia.

Z maksymalności rodziny $\mathcal{F}$ wynika, że:

1. jeżeli $F_0 \subset X$ oraz $F_0 \cap F \neq \emptyset$ dla każdego $F \in \mathcal{F}$, to $F_0 \in \mathcal{F}$.

2. jeżeli $F_1, F_2, \ldots, F_n \in \mathcal{F}$, to $F_1 \cap F_2 \cap \ldots \cap F_n \in \mathcal{F}$.

Ponieważ $\mathcal{F}$ jest rodziną scentrowaną, więc dla każdego $\alpha \in A$ rodzina $\mathcal{F}_\alpha = \{\pi_\alpha(F) : F \in \mathcal{F}\}$ jest scentrowana w odpowiedniej przestrzeni X_α. Ze zwartości przestrzeni X_α wynika, że wszystkie zbiory rodziny $\mathcal{F}_\alpha$ mają wspólny punkt skupienia x_α, tj.

$$x_\alpha \in \bigcap_{F\in\mathcal{F}} \overline{\pi_\alpha(F)}, \quad \alpha \in A.$$

Pokażemy, że punkt $x = \{x_\alpha\}$ ma żądaną własność. Dla każdego otoczenia U_α punktu x_α w X_α i każdego $F \in \mathcal{F}$ mamy $U_\alpha \cap \pi_\alpha(F) \neq \emptyset$, stąd $\pi_\alpha^{-1}(U_\alpha) \cap F \neq \emptyset$ dla każdego $F \in \mathcal{F}$. Z własności 1 rodziny $\mathcal{F}$ wynika, że $\pi_\alpha^{-1}(U_\alpha) \in \mathcal{F}$, a z własności 2, że do $\mathcal{F}$ należą też wszystkie skończone przecięcia $\bigcap_{k=1}^n \pi_{\alpha_k}^{-1}(U_{\alpha_k})$. Zatem $\mathcal{F}$ zawiera wszystkie otoczenia punktu x z bazy kanonicznej produktu topologicznego $X = \prod_{\alpha\in A} X_\alpha$. Z tego, że $\mathcal{F}$ jest rodziną scentrowaną wynika z kolei, że $x \in \overline{F}$ dla wszystkich $F \in \mathcal{F}$. $\square$

Przestrzenie metryczne

10.17. POJĘCIE PRZESTRZENI METRYCZNEJ. Zbiór X nazywamy **przestrzenią metryczną**, jeśli określona jest funkcja $d : X \times X \to [0, \infty)$, o własnościach:

1. $d(x_1, x_2) = 0$ wtedy i tylko wtedy, gdy $x_1 = x_2$,

2. $d(x_1, x_2) = d(x_2, x_1)$ (symetryczność),

3. $d(x_1, x_3) \leqslant d(x_1, x_2) + d(x_2, x_3)$ (nierówność trójkąta).

Funkcję d nazywamy **metryką** lub **odległością** w X.

Dla $x_0 \in X$ i $r > 0$ zbiór

$$K_r(x_0) = \{x \in X : d(x, x_0) < r\}$$

nazywamy **kulą** (otwartą) o środku x_0 i promieniu r. W przestrzeni metrycznej wprowadzamy topologię, przyjmując za rodzinę $\mathcal{T}$ zbiorów otwartych wszystkie te zbiory $A \subset X$, które wraz z każdym punktem x_0 zawierają także pewną kulę o środku x_0. W ten sposób przestrzeń metryczna staje się przestrzenia topologiczną Hausdorffa. Topologię $\mathcal{T}$ nazywamy **topologią wyznaczoną przez metrykę** d. W tej topologii każda z kul $K_r(x_0)$ jest zbiorem otwartymi a każda z kul domkniętych $\overline{K}_r(x_0) = \{x \in X : d(x, x_0) \leqslant r\}$ zbiorem domkniętym. Przestrzeń topologiczną nazywamy **metryzowalną** jeżeli można w niej wprowadzić metrykę wyznaczającą topologię zgodną z topologią wyjściową.

Odwzorowanie f przestrzeni metrycznej X w przestrzeń metryczną Y nazywamy **izometrią** lub **izometrycznym włożeniem**, gdy zachowuje odległość punktów, tzn. gdy $d\big(f(x_1), f(x_2)\big) = d(x_1, x_2)$ dla wszystkich $x_1, x_2 \in X$. Jest oczywiste, że izometria jest homomorfizmem a izometria „na" homeomorfizmem.

10.18. Przestrzenie metryczne zupełne. Mówimy, że ciąg $\{x_n\}$ elementów przestrzeni metrycznej jest **zbieżny** do x_0, gdy $d(x_n, x_0) \to 0$ przy $n \to \infty$. Jeżeli dla każdej liczby dodatniej ε istnieje taki wskaźnik N, że $d(x_n, x_m) < \varepsilon$, gdy $m, n \geqslant N$, to ciąg $\{x_n\}$ nazywamy **fundamentalnym**.

Z nierówności $d(x_n, x_m) < d(x_n, x_0) + d(x_m, x_0)$ wynika, że

Każdy ciąg zbieżny jest fundamentalny.

Przestrzeń metryczną, w której każdy ciąg fundamentalny jest zbieżny, nazywamy **przestrzenią zupełną**.

Twierdzenie o uzupełnianiu. *Niezupełną przestrzeń metryczną można włożyć izometrycznie w przestrzeń metryczną zupełną.*

Dowód: Podamy konstrukcję przestrzeni metrycznej $\widetilde{X}$ i izometrii $i : X \to \widetilde{X}$, dla której $i(X)$ będzie podzbiorem gęstym w $\widetilde{X}$. Oznaczmy w tym celu przez F_X zbiór wszystkich ciągów fundamentalnych przestrzeni X.

Dla ciągów $(x_1, x_2, x_3 \ldots)$, $(y_1, y_2, y_3 \ldots) \in F_X$ istnieje granica

$$\lim_{n\to\infty} d(x_n, y_n).$$

Istotnie, z nierówności

$$\big|d(x_m, y_m) - d(x_n, y_n)\big| \leqslant d(x_m, x_n) + d(y_m, y_n)$$

wynika, że ciąg $\{d(x_n, y_n)\}$ jest fundamentalny w zbiorze $\mathbb{R}$ liczb rzeczywistych, jest zatem ciągiem zbieżnym.

Równość $\lim_{n\to\infty} d(x_n, y_n) = 0$ określa relację równoważności $\sim$ w F_X. Przestrzeń $\widetilde{X}$ definiujemy jako zbiór klas abstrakcji relacji $\sim$ a metrykę d w $\widetilde{X}$ wzorem

$$d(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} d(x_n, y_n),$$

przy czym $\widetilde{x}$ i $\widetilde{y}$ oznaczają tu odpowiednio klasy abstrakcji ciągów $(x_1, x_2, x_3 \ldots)$ i $(y_1, y_2, y_3 \ldots)$. Jest oczywiste, że funkcja i, przyporządkowująca elementowi $x \in X$ klasę abstrakcji ciągu $(x, x, x, \ldots)$, jest izometrycznym włożeniem X w $\widetilde{X}$. Obraz $i(X)$ zbioru X leży gęsto w $\widetilde{X}$, bo dla klasy abstrakcji $\widetilde{x}$ dowolnego ciągu $(x_1, x_2, x_3, \ldots) \in F_X$ zachodzi równość

$$\lim_{m\to\infty} d(\widetilde{x}, i(x_m)) = \lim_{m,n\to\infty} d(x_n, x_m) = 0.$$

Do udowodnienia pozostaje zupełność przestrzeni $\widetilde{X}$. Niech $\{\widetilde{x}_n\}$ będzie ciągiem fundamentalnym w $\widetilde{X}$. Wybierzmy w X tak ciąg $(x_1, x_2, x_3, \ldots)$, aby

$$d(\widetilde{x}_n, i(x_n)) < \frac{1}{n}, \quad n = 1, 2, 3, \ldots$$

Z nierówności

$$d(x_n, x_m) < d(\widetilde{x}_n, \widetilde{x}_m) + \frac{1}{n} + \frac{1}{m}$$

wynika, że $(x_1, x_2, x_3, \ldots) \in F_X$. Jeśli $\widetilde{x}$ oznacza klasę abstrakcji tego ciągu, to

$$\lim_{n\to\infty} d(\widetilde{x}_n, \widetilde{x}) = \lim_{n\to\infty} d(\widetilde{x}_n, i(x_n)) + \lim_{n\to\infty} d(i(x_n), \widetilde{x}) = 0,$$

więc $\lim_{n\to\infty} \widetilde{x}_n = \widetilde{x}$. $\square$

10.19. Twierdzenie Baire'a. *Jeśli $A_1, A_2, A_3, \ldots$ jest ciągiem domkniętych podzbiorów przestrzeni metrycznej zupełnej X oraz*

$$X = \bigcup_{n=1}^{\infty} A_n,$$

to przynajmniej jeden ze zbiorów A_n ma niepuste wnętrze, tzn. zawiera pewną kulę otwartą.

Dowód: Przypuśćmy, że teza twierdzenia nie zachodzi tj., że żaden ze zbiorów A_n nie zawiera żadnej kuli. Wybierzmy dowolnie kulę $K_{r_0}(x_0)$. Ponieważ zbiór A_1 nie zawiera tej kuli i jest domknięty, to zbiór $K_{r_0}(x_0) \setminus A_1$ jest niepusty i otwarty. Zawiera zatem pewną kulę $K_{2r_1}(x_1)$. Następnie zbiór $K_{r_1}(x_1) \setminus A_2$ zawiera pewną kulę $K_{2r_2}(x_2)$, zbiór $K_{r_2}(x_2) \setminus A_3$ kulę $K_{2r_2}(x_3)$, itd. W ten sposób możemy indukcyjnie określić ciąg kul $K_{r_0}(x_0), K_{r_1}(x_1), K_{r_2}(x_2), \ldots$ o tej własności, że

$$K_{2r_n}(x_n) \subset K_{r_{n-1}}(x_{n-1}) \setminus A_n \quad \text{dla} \quad n = 1, 2, 3, \ldots$$

Z inkluzji powyższych wynika w szczególności, że $\lim_{n\to\infty} r_n = 0$, bo $r_n \leqslant \frac{1}{2} r_{n-1} \leqslant \cdots \leqslant \frac{1}{2^n} r_0$. Wynika też, że każdy zbiór A_n jest rozłączny z kulą domkniętą $\overline{K}_{r_n}(x_n)$, gdyż $\overline{K}_{r_n}(x_n) \subset K_{2r_n}(x_n)$.

Zauważmy, że ciąg $x_1, x_2, x_3, \ldots$ środków kul spełnia warunek $d(x_n, x_m) < r_n$, dla $m \geqslant n$, jest więc ciągiem fundamentalnym w X a ponieważ X jest przestrzenią zupełną, jest zbieżny do pewnego $x \in X$. Dla dowolnego wskaźnika n zachodzi wobec tego nierówność

$$d(x_n, x) = \lim_{m\to\infty} d(x_n, x_m) \leqslant r_n.$$

To z kolei implikuje, że $x \in \overline{K}_{r_n}(x_n)$, a dalej, że $x \notin A_n$. W konsekwencji dostajemy $x \notin \bigcup_{n=1}^{\infty} A_n$, co przeczy założeniu $\bigcup_{n=1}^{\infty} A_n$=X. $\square$

10.20. Zbiory zwarte w przestrzeniach metrycznych. Podzbiór E przestrzeni metrycznej nazywamy **całkowicie ograniczonym** jeżeli dla każdej liczby $\varepsilon > 0$ można go pokryć skończoną ilością kul o promieniu ε tj., gdy istnieje taki zbiór skończony F_ε, że $E \subset \bigcup_{x\in F_\varepsilon} K_\varepsilon(x)$.

Z każdego ciągu elementów zbioru całkowicie ograniczonego można wybrać podciąg fundamentalny

Załóżmy, że E jest zbiorem całkowicie ograniczonym a $\{x_n\}$ ciągiem jego elementów. Wybierzmy dowolnie wskaźnik n_1. Wyraz x_{n_1} będzie pierwszym wyrazem podciągu. Niech F_1 będzie skończoną 1-siecią dla E. Ponieważ $E \subset \bigcup_{x\in F_1} K_1(x)$, więc jedna z kul $K_1(x)$ zawiera nieskończenie wiele wyrazów ciągu $\{x_n\}$. Z ciągu usuńmy wszystkie te wyrazy, które nie leżą w wybranej kuli. Z pozostałych wyrazów ciągu wybierzmy wyraz x_{n_2} dowolnie, ale tak, aby $n_2 > n_1$. Podobnie postępując dla sieci $F_{1/2}$, $F_{1/3}$ itd. zbudujemy podciąg $\{x_{n_k}\}$ ciągu $\{x_n\}$ o tej własności, że $d(x_{n_i}, x_{n_j}) < 2/k$, gdy $i, j > k$. Ciąg taki jest fundamentalny.

Zbiory całkowicie ograniczone często nazywa się zbiorami **warunkowo zwartymi**. Wyjaśnienie znajduje się w następującym twierdzeniu:

Podzbiór E przestrzeni metrycznej zupełnej X jest zwarty wtedy i tylko wtedy, gdy jest domknięty i całkowicie ograniczony.

Dowód: Załóżmy, że zbiór E jest zwarty. Ponieważ każda przestrzeń metryczna jest przestrzenią Hausdorffa, więc E jest zbiorem domkniętym. Pokażemy, że jest także zbiorem całkowicie ograniczonym. Niech ε będzie dowolną liczbą dodatnią. Rodzina kul $\{K_\varepsilon(x)\}_{x\in E}$ tworzy pokrycie otwarte zbioru E, a ponieważ E jest zbiorem zwartym, więc pewna skończona podrodzina $K_\varepsilon(x_1), K_\varepsilon(x_2), \ldots, K_\varepsilon(x_n)$ także pokrywa E. Oznacza to, że zbiór $F_\varepsilon = \{x_1, x_2, \ldots, x_n\}$ jest ε-siecią dla E.

Załóżmy teraz, że zbiór E jest domknięty i całkowicie ograniczony. Niech $\{U_\alpha\}_{\alpha\in A}$ będzie dowolnym pokryciem E zbiorami otwartymi U_α. Pokażemy najpierw, że z pokrycia tego można wybrać podpokrycie przeliczalne E, a potem, że z pokrycia przeliczalnego można wybrać podpokrycie skończone.

Niech F_1 będzie skończoną 1-siecią dla E. Ze zbioru kul $\{K_1(x)\}_{x\in F_1}$ wybierzmy te, które zawarte są całkowicie przynajmniej w jednym ze zbiorów U_α pokrycia. Podobnie postępując dla sieci $F_{1/2}$, $F_{1/3}$ itd. otrzymamy przeliczalną rodzinę kul, tworzących pokrycie zbioru E. Jeżeli z rodziny A usuniemy wszystkie te zbiory U_α, które nie zawierają żadnej z wybranej kul, to pozostanie nam przeliczalna podrodzina $\{U_{\alpha_k}\}_{k=1}^{n}$ rodziny $\{U_\alpha\}_{\alpha\in A}$, dalej dająca pokrycie E.

Twierdzimy, że jeden ze zbiorów $U_n = \bigcup_{k=1}^{n} U_{\alpha_k}$, $n = 1, 2, 3, \ldots$ zawiera całkowicie zbiór E. Jeśli tak nie jest, to wybierając z każdego ze zbiorów $E \setminus U_n$ po elemencie x_n, otworzymy ciąg $\{x_n\}$, który — na mocy całkowitej ograniczoności zbioru E — musi zawierać podciąg fundamentalny. Granica x_0 tego podciągu oczywiście należy do E, bo E jest zbiorem domkniętym. Jednak nie może należeć do żadnego ze zbiorów U_n, bo do takiego zbioru musiałyby należeć prawie wszystkie wyrazy podciągu. Tu sprzeczność z założeniem, że $E \subset \bigcup_{k=1}^{n} U_k$. □

10.21. Produkt topologiczny przestrzeni metrycznych. Niech $X = \prod_{k=1}^{\infty} X_k$ będzie produktem topologicznym przeliczalnej rodziny przestrzeni metrycznych X_k i niech d_k oznacza metrykę w przestrzeni X_k. Dla elementów $x = \{x_k\}$ i $y = \{y_k\}$ produktu X połóżmy

$$d(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{d_k(x_k, y_k)}{1 + d_k(x_k, y_k)}.$$

Nietrudno sprawdzamy, że d jest metryką w X.

Pokażemy, że topologia produktowa w X i topologia wyznaczona przez metrykę d pokrywają się. Wystarczy w tym celu porównać bazowe otoczenia punktu $y = \{y_k\} \in X$ w topologii produktowej:

$$U_{n,\varepsilon} = \{x \in X : d_k(x_k, y_k) < \varepsilon \quad \text{dla} \quad k = 1, 2, \ldots, n\}, \qquad n = 1, 2, 3, \ldots, \ \varepsilon > 0,$$

z otoczeniami w metryce d:

$$V_\delta = \{x \in X : d(x,y) < \delta\}, \qquad \delta > 0.$$

Gdy

$$\frac{1}{2^n} + \frac{\varepsilon}{1+\varepsilon} < \delta,$$

to zachodzi zawieranie $U_{n,\varepsilon} \subset V_\delta$, a gdy

$$\delta < \frac{1}{2^n}\frac{\varepsilon}{1+\varepsilon},$$

to zawieranie przeciwne $V_\delta \subset U_{n,\varepsilon}$. Udowodniliśmy tym samym, że:

> *Produkt topologiczny przeliczalnej rodziny przestrzeni metrycznych jest przestrzenią metryzowalną.*

Literatura.

[1] Ryszard Engelking, *Topologia ogólna*, PWN, Warszawa 1989.

WSKAZÓWKI

1.25 Pokazać, że funkcję x_0 można przedstawić w postaci zbieżnego w $L^p(a,b)$ szeregu $\sum_{k=1}^{\infty} \lambda_k\, x_k$.

1.34 Wskazać w ℓ^1 nieprzeliczalny zbiór liniowo niezależny.

1.36 Funkcja $x = \operatorname{ctg} \pi t$ jest homeomorfizmem $(0,1)$ na $\mathbb{R}$.

1.39 W dowodzie zupełności przestrzeni X/Y skorzystać z zadania 1.3.

2.18 Znaleźć funkcje liniową przeprowadzającą przedział $[0,1]$ na $[a,b]$.

2.20 Nie.

3.10 Równość równoległoboku potrzebna jest tylko przy dowodzie addytywności iloczynu skalarnego.

5.15 Skorzystać z twierdzenia Banacha-Steinhausa.

5.19 Wykazać, że różniczkowanie jest ciągłą operacją z X w $C[0,1]$. Wywnioskować stąd, że kula jednostkowa w X jest zbiorem jednakowo ciągłym. Następnie skorzystać z twierdzenia Arzeli i twierdzenia Riesza.

6.3 Dowieść, że dla każdej funkcji $x \in L^p(0,1)$ i dowolnej liczby $r > 0$ istnieje taki podział $0 = t_0 < t_1 < \cdots < t_n = 1$ odcinka $[0,1]$, że funkcje

$$x_k(t) = \begin{cases} nx(t) & \text{gdy } t_{k-1} \leqslant t \leqslant t_k, \\ 0 & \text{poza tym,} \end{cases} \qquad k = 1, 2, \ldots, n$$

spełniają nierówność $\|x_k\| < r$ oraz $x = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$.

6.10 Topologia ta nie jest metryzowalna a każdy ciągły funkcjonał liniowy x^* na X ma postać $x^*(x) = \sum_{k=1}^{n} \alpha_k x(t_k)$ przy pewnym wyborze punktów $t_1, t_2, \ldots, t_n$ w $[0,1]$ i liczb $\alpha_1, \alpha_2, \ldots, \alpha_n$.

7.9 Posłużyć się kryterium Schura z funkcją $\varphi(k) = \frac{1}{\sqrt{k-1/2}}$.

ROZWIĄZANIA ZADAŃ

4.43 Na podprzestrzeni $X_0 \subset C(\mathbb{R})$, złożonej z tych funkcji x, dla których granica $\lim\limits_{t\to-\infty} x(t)$ istnieje, określmy funkcjonał liniowy x_0^* wzorem

$$x_0^*(x) = \lim_{t\to-\infty} x(t),$$

a następnie przedłużmy go do ciągłego funkcjonału liniowego x^* na całą przestrzeń $C(\mathbb{R})$. Wtedy $x^*(1) = 1$, zaś $x^*(x_n) = 0$ dla wszystkich n.

2.17 Można się posłużyć wzorami z końca dowodu twierdzenia Weierstrassa 2.1 otrzymując dla funkcji $x(t) = t^2$

$$x_n(t) = \tfrac{n-1}{n}\, t^2 + \tfrac{1}{n}\, t,$$

a dla funkcji $x(t) = t^3$

$$x_n(t) = \tfrac{(n-1)(n-2)}{n^2}\, t^3 + 3\tfrac{n-1}{n^2}\, t^2 + \tfrac{1}{n^2}\, t.$$

1.33 Każda funkcja $f \in C[-1,1]$ ma jednoznaczne przedstawienie w postaci sumy $f = f_1 + f_2$ funkcji parzystej i nieparzystej, mianowicie

$$f_1(t) = \frac{f(t) + f(-t)}{2}, \quad f_2(t) = \frac{f(t) - f(-t)}{2},$$

zachodzą przy tym nierówności $\|f\|_\infty \leqslant \|f_1\|_\infty + \|f_2\|_\infty \leqslant 2\,\|f\|_\infty$.

6.19 Ciągłość dodawania funkcji jest oczywista. Mnożenie przez liczby nie jest ciągłe, bo na przykład dla funkcji $x(t) = 1/t$ nierówność $|\lambda x(t) - x(t)| < 1$ nie zachodzi dla żadnej liczby $\lambda \neq 1$.

1.12 Funkcja $\|\ \|_p$ nie jest podaddytywna. Dla ciągów $x = (1, t^{1/p}, 0, \ldots)$, gdzie $t > 0$, oraz $y = (1, 0, 0, \ldots)$ otrzymujemy

$$\|x + y\|_p - \|x\|_p - \|y\|_p = (2^p + t)^{1/p} - (1 + t)^{1/p} - 1 > 0,$$

a nierówność łatwo jest sprawdzić różniczkując po t.

Zbiór ℓ^p jest przestrzenią liniową. Z nierówności $|x_n + y_n| \leqslant 2\max\{|x_n|, |y_n|\}$ wynika

$$|x_n + y_n|^p \leqslant 2^p \max\{|x_n|^p, |y_n|^p\} \leqslant 2^p\big(|x_n|^p + |y_n|^p\big).$$

Dlatego $\sum |x_n + y_n|^p < \infty$, gdy oba ciągi $\{x_n\}$, $\{y_n\}$ leżą w ℓ^p.

4.38 Wystarczy tezy dowieść dla ciągu $\{x_n\}$ słabo zbieżnego do zera. Wtedy ciąg $\{x_n\}$ jest jednostajnie ograniczony oraz $x_n(t) \to 0$ dla wszystkich $t \in [a, b]$, więc

$$\lim_{n\to\infty} \int_a^b \big|x_n(t)\big|^p dt = 0$$

na mocy twierdzenia Lebesgue'a o zbieżności ograniczonej. $\square$

2.18 $f_n(x) = \dfrac{1}{(b-a)^n} \displaystyle\sum_{k=0}^{n} f\big(\tfrac{n-k}{n}a + \tfrac{k}{n}b\big)\,(x-a)^k(b-x)^{n-k}$

1.39 Odwzorowanie τ jest kontrakcją liniową, jest wobec tego funkcją ciągłą. Zauważmy, że obrazem kuli otwartej $\{x \in X : \|x - x_0\| < r\}$ o środku x_0 i promieniu $r > 0$ przez odwzorowanie τ jest kula $\big\{[x] \in X/Y : \big\|[x] - [x_0]\big\| < r\big\}$. Istotnie, zawieranie $\subset$ jest oczywiste, a jeżeli element $[x]$ leży w drugiej kuli, to z definicji normy ilorazowej wynika istnienie takiego elementu $x' \in [x]$, że $\|x' - x_0\| < r$, więc x' leży w pierwszej kuli oraz $\tau(x') = [x]$. Zachodzi zatem także zawieranie w drugą stronę. Ponieważ każdy zbiór otwarty U w przestrzeni X jest sumą kul otwartych, więc jego obraz $\tau(U)$ musi być zbiorem otwartym w przestrzeni X/Y.

Aby dowieść zupełności przestrzeni X/Y wystarczy pokazać, że każdy szereg bezwzględnie zbieżny jest zbieżny (patrz zad. 1.3). Załóżmy, że dla ciągu $\big\{[x_n]\big\}$ w przestrzeni X/Y szereg norm $\sum \big\|[x_n]\big\|$ jest zbieżny. Z każdej z warstw $[x_n]$ wybierzmy element x'_n tak, aby $\|x'_n\| \leqslant \big\|[x_n]\big\| + 1/2^n$. Wtedy $\sum \|x'_n\| < \infty$, więc, na mocy założonej zupełności przestrzeni X, szereg $\sum x'_n$ jest zbieżny w X. Z ciągłości i liniowości odwzorowania τ wynika, że szereg $\sum [x_n]$, będąc jego obrazem, musi być zbieżny w X/Y.

3.23 Dla wektora $x \in \mathcal{H}$ warstwa $[x] = x + \mathcal{H}_0$ przestrzeni $\mathcal{H}/\mathcal{H}_0$ zawiera dokładnie jeden element przestrzeni $\mathcal{H}_0^{\perp}$, jest nim wektor x_1 z rozkładu ortogonalnego $x = x_0 + x_1$.

7.12 Załóżmy, że $\{y_n\}$ jest ciągiem Cauchy'ego w Y, który nie jest zbieżny. Ustalmy niezerowy funkcjonał $x^* \in X^*$ i określmy ciąg operatorów $\{T_n\}$ w

$\mathcal{L}(X,Y)$ wzorem
$$T_n x = x^*(x)\,y_n, \quad n = 1,2,3,\ldots$$
Łatwo zobaczyć, że gdy $x^*(x) \neq 0$, to ciąg $\{T_n x\}$ nie jest zbieżny, nie może więc być zbieżny też ciąg $\{T_n\}$.

3.27 Funkcja r_n jest stała na każdym z przedziałów $I_{k,n} = \left(\frac{k}{2^{n-1}}, \frac{k+1}{2^{n-1}}\right)$ i na przemian przyjmuje wartości $+1$ i -1, zatem gdy $m < n$, to całka z niej po każdym z przedziałów $I_{j,m}$ jest zerem. Dlatego
$$\langle r_m,\, r_n\rangle = \sum_{j=0}^{2^m-1} (-1)^j \int_{I_{j,m}} r_n(t)\,dt = 0.$$
Łatwo sprawdzamy, że funkcja $r_2\,r_3$ jest ortogonalna do wszystkich funkcji r_n, z tego powodu układ Rademachera nie jest zupełny.

Stosując podobne rozumowanie jak wyżej, można pokazać, że układ ortonormalny tworzą wraz z funkcją r_1 wszystkie skończone iloczyny $r_{n_1} r_{n_2} \cdots r_{n_k}$, $1 < n_1 < n_2 < \ldots < n_k$, $k = 1,2,3,\ldots$, funkcji Rademachera.

2.20 Funkcje jednostajnie ciągłe tworzą domkniętą podprzestrzeń liniową w przestrzeni $C(\mathbb{R})$. Istotnie, jeżeli funkcja x jest granicą jednostajnie zbieżnego ciągu $\{x_n\}$ funkcji jednostajnie ciągłych z $C(\mathbb{R})$, to dla dowolnej liczby dodatniej ε można wybrać taki wskaźnik n, by $\|x - x_n\| < \varepsilon$; następnie dla funkcji x_n taką liczbę dodatnią δ, by dla dowolnej pary $s,t \in \mathbb{R}$ nierówność $|s-t| < \delta$ pociągała $\big|x_n(s) - x_n(t)\big| < \varepsilon$. Wtedy
$$\begin{aligned}\big|x(s) - x(t)\big| &\leqslant \big|x(s) - x_n(s)\big| + \big|x_n(s) - x_n(t)\big| + \big|x_n(t) - x(t)\big| \\ &\leqslant 2\|x - x_n\| + \big|x_n(s) - x_n(t)\big| < 3\varepsilon\end{aligned}$$
dla każdej pary $s,t \in \mathbb{R}$ spełniającej nierówność $|s-t| < \delta$.

Przestrzeń X jest właściwym podzbiorem $C(\mathbb{R})$, gdyż istnieją ograniczone funkcje ciągłe, które nie są jednostajnie ciągłe, np. funkcja $x(t) = \sin t^2$.

4.8 Przyporządkowanie $y \to y^*$ określone wzorem (4.22) jest izometrycznym izomorfizmem ℓ^1 na podprzestrzeń przestrzeni $\ell^{\infty *}$, jednak nie każdy funkcjonał y^* ma postać (4.22), np. funkcjonał LIM z przykładu **??**. Rzeczywiście, mielibyśmy $y_k = \operatorname{LIM} e_k = 0$ dla wszystkich k.

4.26 Z twierdzenia o wartości średniej wynika, że jeżeli $P = \{t_0, t_1, \ldots, t_n\}$ jest rozbiciem przedziału $[a,b]$, to $\big|f(t_k) - f(t_{k-1})\big| = \big|f'(s_k)\big|(t_k - t_{k-1})$ dla pewnego punktu $s_k \in (t_{k-1}, t_k)$, więc $\sum_{k=1}^n \big|f(t_k) - f(t_{k-1})\big|$ jest sumą Riemanna całki $\int_a^b \big|f'(t)\big|\,dt$.

4.46 Taką własność ma ciąg $x_n = \frac{1}{n}(e_1 + e_2 + \cdots + e_n)$, gdyż $\|x_n\| = \frac{1}{\sqrt{n}}$.

1.4.3 Niech x będzie dowolnym wektorem niezerowym. Szereg $\sum_{n=1}^\infty \frac{(-1)^n}{n} x$ jest zbieżny, bo zbieżny jest szereg $-1 + \frac{1}{2} - \frac{1}{3} + \ldots$, a nie jest zbieżny bezwzględnie.

3.24.2 Jeśli zbiór $\{x_t\}_{t\in T}$ jest ortonormalny, to kule $\{x \in \mathcal{H} : \|x - x_t\| < 1\}$, $t \in T$ są parami rozłączne, a każda zawiera przynajmniej jeden element ośrodka. Zbiór T musi być więc najwyżej przeliczalny.

2.15 $x_0(t) = 1$, $x_1(t) = x_2(t) = 1 - 2t$, $x_3(t) = 1 - \frac{3}{2}t - \frac{3}{2}t^2 + x^3$, $x_4(t) = (1-t)^4 + 2\sqrt{2}\,t(1-t)^3 - 2\sqrt{2}\,t^3(1-x) - t^4$.

4.9 Rozumujemy identycznie jak dla przestrzeni ℓ^1, korzystając przy oszacowaniu $|y^*(x)| \leqslant \|x\|_p \|y\|_q$ z nierówności Höldera. Dla dowodu nierówności $\|y^*\| \geqslant \|y\|_q$ wybieramy wektor $x = (x_1, x_2, \ldots)$, postaci $x_k = \operatorname{sgn} \overline{y_k}\, |y_k|^{q-1}$. Mamy wtedy $\|x\|_p = \|y\|_q^{q/p}$ oraz $y^*(x) = \|y\|_q^q$, zatem $\|y^*\| \geqslant \|y\|_q$.

1.4.2 Niech $\{x_n\}$ będzie ciągiem Cauchy'ego. Jeżeli m jest takim wskaźnikiem, że $\|x_n - x_m\| \leqslant 1$ dla $n > m$, oraz $M = \max\{\|x_1\|, \ldots, \|x_{m-1}\|, \|x_m\| + 1\}$, to $\|x_n\| \leqslant M$ dla wszystkich $n = 1, 2, \ldots$.

5.15 Na przestrzeni ℓ^∞ funkcjonały liniowe

$$x_n^*(b) = \sum_{k=1}^n a_k b_k, \quad \text{gdzie} \quad b = (b_1, b_2, b_3, \ldots) \in \ell^\infty,\ n = 1, 2, 3, \ldots$$

są ciągłe i $\big\|x_n^*\big\| = \sum\limits_{k=1}^n |a_k|$. Także dla każdego $b \in \ell^\infty$ ciąg $\{x_n^*(b)\}$ jest ograniczony. Z twierdzenia Banacha-Steinhausa wynika, że normy $\|x_n^*\|$ muszą być wspólnie ograniczone.

1.21 Nierówność $\|x\|_2 \leqslant C_2\,\|x\|_1$, $x \in C[a,b]$, nie zachodzi przy żadnej stałej C_2, bo dla funkcji $x_n(t) = t^n$, $n = 1, 2, 3, \ldots$ mamy $\|x\|_1 = \frac{1}{n+1}$, $\|x_n\|_2 = 1$.

1.3 Załóżmy, że X jest przestrzenią zupełną oraz, że szereg $\sum_{k=1}^\infty \|x_k\|$ jest zbieżny. Aby dowieść zbieżności szeregu $\sum_{k=1}^\infty x_k$ w X, wystarczy pokazać, że ciąg $s_n = \sum_{k=1}^{n} x_k$ jego sum częściowych jest ciągiem Cauchy'ego w X. Zauważmy w tym celu, że dla $m < n$ mamy

$$\begin{aligned}\|s_n - s_m\| &= \|x_{m+1} + x_{m+2} + \ldots + x_n\| \\ &\leqslant \|x_{m+1}\| + \|x_{m+2}\| + \ldots + \|x_n\| \\ &= \sigma_n - \sigma_m,\end{aligned}$$

gdzie σ_n i σ_m są odpowiednio n-tą i m-tą sumą częściową szeregu $\sum_{k=1}^{\infty} \|x_k\|$. Stąd, dla dowolnych już m, n zachodzi

$$\|s_n - s_m\| \leqslant |\sigma_n - \sigma_m|,$$

a ponieważ $\{\sigma_n\}$ jest liczbowym ciągiem Cauchy'ego, więc

$$\lim_{m,n\to\infty} \|s_n - s_m\| = 0.$$

Przejdźmy do dowodu w drugą stronę. Niech $\{x_n\}$ będzie dowolnym ciągiem Cauchy'ego w X. Pokażemy, że jest on zbieżny. Wybierzmy z ciągu $\{x_n\}$ podciąg $\{x_{n_k}\}$ tak, aby $\|x_{n_k} - x_m\| \leqslant 1/2^k$ dla $k = 1, 2, \ldots$ oraz wszystkich $m > n_k$ i oznaczmy

$$y_1 = x_{n_1}, \qquad y_k = x_{n_k} - x_{n_{k-1}}.$$

Ponieważ

$$\sum_{k=1}^{\infty} \|y_k\| \leqslant \|x_{n_1}\| + \sum_{k=2}^{\infty} \frac{1}{2^k} < \infty,$$

więc z założenia wynika, że szereg $\sum_{k=1}^{\infty} y_k$ jest zbieżny w X. Ciąg $\{x_{n_k}\}$ jest ciągiem sum częściowych tego szeregu, jest więc także zbieżny. Niech $x = \lim_{k\to\infty} x_{n_k}$. Sprawdźmy, że x jest także granicą samego ciągu $\{x_n\}$. Rzeczywiście,

$$\|x - x_m\| = \|x - x_{n_k}\| + \|x_{n_k} - x_m\| < \varepsilon + \frac{1}{2^k}$$

dla dostatecznie dużych k i $m > n_k$.

7.66 Oczywiście nie. Jeżeli $g : [0,1] \to \mathbb{C}$ jest funkcją ciągłą i $|g(t)| \equiv 1$, to operator $Tx(t) = x(t)\,g(t)$ jest izometrią na $C[0,1]$, ale nie ma postaci (7.42).

5.3 Na przestrzeni $\boldsymbol{s}_0$ tych ciągów liczb zespolonych, dla których tylko skończenie wiele wyrazów może być różnych od zera, z normą „sup” określmy operatory $T_1, T_2, T_3, \ldots$ wzorem

$$T_n(x_1, x_2, x_3, \ldots) = (x_1, 2x_2, \ldots, nx_n, 0, 0, \ldots).$$

Założenia wniosku 5.2 są wtedy spełnione, bo $\|T_n\| = n$ a dla $x \in \boldsymbol{s}_0$ ciąg $\{T_n x\}$ jest od pewnego miejsca stały. Operator T ma postać

$$T(x_1, x_2, x_3, \ldots) = (x_1, 2x_2, 3x_3, \ldots),$$

jest więc operatorem nieciągłym.

3.31 Ponieważ $U_n(\cos u) = \dfrac{\sin(n+1)u}{2^n \sin u}$, więc ortonormalność układu jest konsekwencją równości

$$\int_{-1}^{1} U_m(t)\, U_n(t)\, \sqrt{1-t^2}\, dt = \frac{1}{2^{m+n}} \int_0^{\pi} \sin(m+1)u\, \sin(n+1)u\, du.$$

6.10 Bazę otoczeń zera dla omawianej topologii tworzą zbiory $U_{\varepsilon,F}$ postaci

$$U_{\varepsilon,F} = \{x \in X : |x(t)| < \varepsilon \text{ dla } t \in F\},$$

gdzie ε jest dowolną liczbą dodatnią, a F skończonym podzbiorem $[0,1]$. Gdyby topologia ta była metryzowalna, to spośród zbiorów $U_{\varepsilon,F}$ można by wybrać przeliczalną rodzinę, tworzącą bazę otoczeń zera. Powiedzmy, że tworzyłyby ją zbiory U_{ε_n,F_n}, $n = 1, 2, \ldots$. Ponieważ suma $\bigcup_{n=1}^{\infty} F_n$ jest zbiorem najwyżej przeliczalnym, więc w $[0,1]$ istnieje punkt t_0 leżący poza tą sumą. Mamy tu sprzeczność, bo zbiór $U_{1,\{t_0\}}$ jest otwarty, a nie zawiera żadnego ze zbiorów U_{ε_n,F_n}.

Niech x^* będzie ciągłym funkcjonałem liniowy na X. Z ciągłości w zerze wynika istnienie takiego otoczenia $U_{\varepsilon,F}$, że $|x^*(x)| \leqslant 1$ dla wszystkich $x \in U_{\varepsilon,F}$. Zauważmy, że jeżeli funkcja x zeruje się na zbiorze F, to $x^*(x) = 0$. Rzeczywiście, ponieważ $nx \in U_{\varepsilon,F}$ dla $n = 1, 2, \ldots$, więc $|x^*(x)| \leqslant 1/n$, a w konsekwencji $x^*(x) = 0$. Oznaczmy punkty zbioru F przez $t_1, t_2, \ldots, t_n$ i wybierzmy w X taki układ funkcji $e_1, e_2, \ldots, e_n$, że $e_m(t_k) = \delta_{mk}$ dla $m, k = 1, 2, \ldots, n$. Oznaczmy także $x^*(e_k) = \alpha_k$. Jeżeli $x \in X$, to funkcja $x - \sum_{k=1}^{n} x(t_k) e_k$ znika na zbiorze F, więc $x^*(x - \sum_{k=1}^{n} x(t_k) e_k) = 0$, a stąd

$$x^*(x) = \sum_{k=1}^{n} \alpha_k x(t_k).$$

Z drugiej strony wzór powyższy określa ciągły funkcjonał liniowy x^* na X przy dowolnym wyborze zbioru skończonego $F = \{t_1, t_2, \ldots, t_n\}$ w $[0,1]$ i układu liczb zespolonych $\alpha_1, \alpha_2, \ldots, \alpha_n$.

1.6 Aby dowieść zupełności obu przestrzeni można powtórzyć rozumowanie z przykładu 1.5 przy dodatkowym założeniu, że każdy z ciągów $\{x_n^k\}$, $k = 1, 2, \ldots,$

jest zbieżny, powiedzmy do liczby g_k. Przechodząc w (1.1) do granicy przy $n \to \infty$ otrzymujemy $|g_k - g_m| \leqslant \varepsilon$, a w konsekwencji zbieżność ciągu $g_1, g_2 \ldots$ Oznaczmy $\lim g_k = g$. Wybierzmy i ustalmy wskaźnik $k \geqslant k_0$ tak, by $|g_k - g| < \varepsilon$ a następnie n_0 tak, by $|x_n^k - g_k| < \varepsilon$ dla $n \geqslant n_0$. Wtedy z (1.2)

$$|x_n - g| \leqslant |x_n - x_n^k| + |x_n^k - g_k| + |g_k - g| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon$$

dla wszystkich $n \geqslant n_0$. Zatem $\lim x_n = g$.

1.34 Wystarczy w przestrzeni ℓ^1 wskazać nieprzeliczalny zbiór liniowo niezależny. Takim zbiorem jest na przykład zbiór wszystkich ciągów postaci $\{n^{-s}\}_{n=1}^{\infty}$, $s > 1$, bo gdy przy $1 < s_1 < s_2 < \ldots < s_k$ równość

$$\alpha_1 n^{-s_1} + \alpha_2 n^{-s_2} + \ldots + \alpha_k n^{-s_k} = 0$$

zachodzi dla wszystkich n, to mnożąc obie strony przez n^{s_1} i przechodząc do granicy przy $n \to \infty$ otrzymujemy $\alpha_1 = 0$. Usuwając z tej równości zerowy składnik $\alpha_1 n^{-s_1}$ i powtarzając rozumowanie dla wykładnika s_2 otrzymujemy $\alpha_2 = 0$, itd. W końcu po k-krokach otrzymamy $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

4.33.1 Jeśli $x^*(x_n) \to x^*(x_0')$ oraz $x^*(x_n) \to x^*(x_0'')$, to $x^*(x_0' - x_0'') = 0$, więc $x_0' = x_0''$.

7.72 Jest tak tylko wtedy, gdy $1 \notin \sigma(U)$. Warunek ten nie jest spełniony np. dla operatora $U = I$.

5.16 Niech $X_0 = \{x \in X : Tx = 0\}$ oznacza jądro odwzorowania T. Z twierdzenia 5.10 wiemy, że T jest izomorfizmem z przestrzeni X/X_0 na Y. Gdy $y = 0$, to przyjmijmy $x = 0$, a gdy $y \neq 0$ to w warstwie $T^{-1}y$ przestrzeni X/X_0 wybierzmy x tak, aby $\|x\| \leqslant 2\,\|T^{-1}y\|$. Wtedy $Tx = y$ oraz $\|x\| \leqslant 2\,\|T^{-1}\|\,\|y\|$.

1.36 Ponieważ funkcja $x = \operatorname{ctg} \pi t$ jest homeomorfizmem $(0,1)$ na $\mathbb{R}$, więc możemy określić odwzorowanie $T : L^p(\mathbb{R}) \to L^p(0,1)$ wzorem

$$Tf(t) = \pi^{-p} \sin^{2/p} \pi t\, f(\operatorname{ctg} \pi t).$$

Wtedy

$$\|Tf\|_p^p = \int_0^1 \left|f(\operatorname{ctg} \pi t)\right|^p \frac{\sin^2 t}{\pi}\, dt = \int_{-\infty}^{+\infty} \left|f(x)\right|^p dx = \|f\|_p^p.$$

7.5 Mamy obliczyć bezpośrednio

$$\begin{aligned}\|A\| &= \sup_{|x|^2+|y|^2+|z|^2=1} \left(|x+iy|^2+|y+iz|^2+|z+ix|^2\right)^{1/2}\\ &= \sup_{|x|^2+|y|^2+|z|^2=1} \left(2+2\,\mathrm{Im}(x\overline{y}+y\overline{z}+z\overline{x})\right)^{1/2}.\end{aligned}$$

Jeśli przyjmiemy $x = x_1 + ix_2$, $y = y_1 + iy_2$, $z = z_1 + iz_2$, to zadanie sprowadzi się do obliczenia maksymalnej wartości wyrażenia

$$\begin{aligned}V &= \mathrm{Im}(x\overline{y}+y\overline{z}+z\overline{x})\\ &= x_2y_1 - x_1y_2 + y_2z_1 - y_1z_2 + z_2x_1 - z_1x_2,\end{aligned}$$

pod warunkiem $x_1^2 + x_2^2 + y_1^2 + y_2^2 + z_1^2 + z_2^2 = 1$. Do obliczenia tego maksimum warunkowego zastosujemy metodę czynników nieoznaczonych Lagrange'a (patrz [3] I. 212). Przyrównanie do zera pochodnych cząstkowych funkcji

$$x_2y_1 - x_1y_2 + y_2z_1 - y_1z_2 + z_2x_1 - z_1x_2 - \lambda(x_1^2 + x_2^2 + y_1^2 + y_2^2 + z_1^2 + z_2^2)$$

prowadzi do układu równań

$$\begin{aligned} z_2 - y_2 - 2\lambda x_1 &= 0, \qquad & y_1 - z_1 - 2\lambda x_2 &= 0,\\ x_2 - z_2 - 2\lambda y_1 &= 0, & z_1 - x_1 - 2\lambda y_2 &= 0,\\ y_2 - x_2 - 2\lambda z_1 &= 0, & x_1 - y_1 - 2\lambda z_2 &= 0.\end{aligned}$$

Z niego

$$\begin{aligned}V_{\max} &= (z_2-y_2)x_1 + (x_2-z_2)y_1 + (y_2-x_2)z_1 = 2\lambda(x_1^2+y_1^2+z_1^2)\\ &= (y_1-z_1)x_2 + (z_1-x_1)y_2 + (x_1-y_1)z_2 = 2\lambda(x_2^2+y_2^2+z_2^2),\end{aligned}$$

co po uwzględnieniu równości $x_1^2 + x_2^2 + y_1^2 + y_2^2 + z_1^2 + z_2^2 = 1$ daje $V_{\max} = \lambda$. Wartość $\lambda = \frac{\sqrt{3}}{2}$ łatwo wyznaczamy z układu równań, zatem

$$\|A\| = \sqrt{2+\sqrt{3}}.$$

Obliczanie współrzędnych wektora optymalnego (x, y, z) można kontynuować, otrzymując w rezultacie

$$x = \frac{e^{it}}{\sqrt{3}}, \quad y = \frac{e^{i(t+2\pi/3)}}{\sqrt{3}}, \quad z = \frac{e^{i(t+4\pi/3)}}{\sqrt{3}}, \qquad t \in \mathbb{R}.$$

Tworzą one wierzchołki trójkąta foremnego wpisanego w okrąg o środku 0 i promieniu $\frac{1}{\sqrt{3}}$.

Druga metoda obliczania $\|A\|$ znacznie szybciej prowadzi do celu. Ponieważ

$$A^* = \begin{pmatrix} 1 & 0 & -i \\ -i & 1 & 0 \\ 0 & -i & 1 \end{pmatrix}, \qquad A^*A = \begin{pmatrix} 2 & i & -i \\ -i & 2 & i \\ i & -i & 2 \end{pmatrix},$$

więc wielomian $\det(zI - A^*A)$ przyjmuje postać

$$\det(zI - A^*A) = z^3 - 6z^2 + 9z - 2 = (z-2)(z-2-\sqrt{3})(z-2+\sqrt{3}),$$

skąd

$$\|A\| = \max\left\{\sqrt{2}, \sqrt{2+\sqrt{3}}, \sqrt{2-\sqrt{3}}\right\} = \sqrt{2+\sqrt{3}}.$$

2.16 Ciąg s_n sum częściowych szeregu Taylora funkcji $x(t) = \cos \pi t$

$$s_n(t) = 1 - \frac{\pi^2}{2!}t^2 + \frac{\pi^4}{4!}t^4 - \frac{\pi^6}{6!}t^6 + \ldots + (-1)^n \frac{\pi^{2n}}{(2n)!}t^{2n}$$

jest zbieżny jednostajnie do funkcji x na każdym przedziale skończonym.

4.33.3 Przestrzeń skończenie wymiarowa jest izomorficzna z $\mathbb{C}^k$ (względnie $\mathbb{R}^k$, gdy jest to przestrzeń rzeczywista). W przestrzeni $\mathbb{C}^k$ słaba zbieżność ciągu $x_n = (x_{n1}, x_{n2}, \ldots, x_{nk})$ pociąga zbieżność ciągów $\{x_{ni}\}_{n=1}^{\infty}$, $i = 1, 2, \ldots, k$ jego współrzędnych a to, jak wiadomo, gwarantuje zbieżność samego ciągu $\{x_n\}$.

2.19 Jeżeli odwzorowanie $\varphi : S_1 \to S_2$ jest homeomorfizmem, to odwzorowanie liniowe $T : C(S_2) \to C(S_1)$ postaci

$$Tx(s) = x\big(\varphi(s)\big)$$

jest izometrią.

Niech $\beta\mathbb{R}$ oznacza uzwarcenie Čecha-Stone'a prostej $\mathbb{R}$. Z określenia uzwarcenia Čecha-Stone'a wynika (patrz 10.14), że przestrzenie $C(\mathbb{R})$ i $C(\beta\mathbb{R})$ są izometrycznie izomorficzne, choć jak widać przestrzenie $\mathbb{R}$ i $\beta\mathbb{R}$ homeomorficzne nie są (druga jest zwarta a pierwsza nie).

1.35 Normę zupełną w przestrzeni $C^k(\mathbb{R})$ można wprowadzić na wiele sposobów. Jednym z nich jest

$$\|x\| = \sup_{t\in\mathbb{R}} \big|x(t)\big| + \sup_{t\in\mathbb{R}} \big|x'(t)\big|.$$

Jeżeli $\{x_n\}$ jest ciągiem fundamentalnym w tej normie, to jest jednostajnie zbieżny do pewnej funkcji x_0, a także ciąg pochodnych $\{x'_n\}$ jest zbieżny jednostajnie do pewnej funkcji y_0. Ponieważ

$$x_n(t) = x_n(0) + \int_0^t x'_n(s)\,ds,$$

więc $x_0(t) = x_0(0) + \int_0^t y_0(s)\,ds$ dla każdego $t \in \mathbb{R}$. Wynika stąd, że x_0 jest funkcją różniczkowalną oraz $x'_0 = y_0$, zatem y_0 jest granicą ciągu $\{x_n\}$ w normie $\|\ \|$.

4.36 Zbiór złożony z punkcji charakterystycznych wszystkich przedziałów $[a, b]$ jest liniowo gęsty w przestrzeni $L^p(\mathbb{R})$, można zatem przeprowadzić analogiczne rozumowanie jak w przykładzie 4.34.

3.20 Mamy

$$\left| \det \left\{x_i(t_j)\right\}_{1\leqslant i,j\leqslant n} \right|^2 = \sum_{\tau,\sigma} \operatorname{sgn}(\tau\sigma) \prod_{i=1}^{n} x_{\tau(i)}(t_i)\,\overline{x_{\sigma(i)}(t_i)},$$

gdzie τ i σ przebiegają grupę permutacji liczb $\{1, 2, \ldots, n\}$ a sgn oznacza znak permutacji. Przyjmując

$$a_{i,j} = \langle x_i,\, x_j\rangle = \int_{-\infty}^{\infty} x_i(t)\,\overline{x_j(t)}\,dt,$$

prawa strona dowodzonej równości przyjmie postać

$$\begin{aligned}
\frac{1}{n!}\sum_{\tau,\sigma} \operatorname{sgn}(\tau\sigma) \prod_{i=1}^{n} a_{\tau(i),\sigma(i)} &= \frac{1}{n!}\sum_{\tau,\sigma} \operatorname{sgn}(\tau\sigma^2) \prod_{i=1}^{n} a_{\tau\sigma(i),\sigma(i)} = \\
&= \frac{1}{n!}\sum_{\sigma}\sum_{\tau} \operatorname{sgn}(\tau) \prod_{i=1}^{n} a_{\tau(i),i} = \\
&= \det \left\{a_{i,j}\right\}_{1\leqslant i,j\leqslant n}
\end{aligned}$$

6.14 Niech $\mathcal{O}$ będzie ośrodkiem w ℓ^2. Jeżeli x_0 należy do słabego domknięcia zbioru A, to w A metodą „przekątniową" można wybrać taki ciąg $\{x_n\}$, że

$$\lim_{n\to\infty} \langle x_n - x_0,\, y\rangle = 0$$

dla wszystkich $y \in \mathcal{O}$. Ponieważ ciąg $\{x_n - x_0\}$ jest ograniczony, więc równość

$$\lim_{n\to\infty} \langle x_n - x_0,\, y\rangle = 0$$

zachodzi dla wszystkich $x \in \ell^2$.

1.25 Pokażemy, że układ Schura nie daje jednoznaczności rozkładu. Przedstawimy w tym celu funkcję x_0 w postaci zbieżnego szeregu

$$x_0 = \sum_{k=1}^{\infty} \lambda_k \, x_k.$$

Szereg ten będzie zbieżny na całym półotwartym przedziale $(a, b]$, w szczególności zbieżny w punktach $t_1, t_2, t_3, \ldots$. Z własności $x_k(t_i) = 0$ dla $i = 1, 2, \ldots, k-1$ i $x_k(t_k) = 1$, podobnie jak w przykładzie 1.24, otrzymamy układ równań $1 = x_0(t_1) = \lambda_1$, $1 = x_0(t_2) = \lambda_1 \, x_1(t_2) + \lambda_2$, itd., który pozwala wyznaczyć współczynniki $\lambda_1, \lambda_2, \lambda_3, \ldots$ jednoznacznie. Jeśli y_n jest n-tą sumą częściową tego szeregu oraz s_n oznacza najmniejszą spośród liczb $t_1, t_2, \ldots, t_n$, to wykresem funkcji y_n jest łamana złożona z dwóch odcinków, łączących punkt $(a, 0)$ z punktem $(s_n, 1)$ i punkt $(s_n, 1)$ z punktem $(b, 1)$. Dlatego

$$\|x_0 - y_n\|_p^p = \int_a^{s_n} \left(\frac{t-a}{s_n - a}\right)^p dt = \frac{s_n - a}{p+1} \to 0.$$

7.2 Jeśli T jest operatorem ograniczonym, to $\|Tx - Ty\| \leqslant \|T\| \, \|x - y\|$ dla wszystkich $x, y \in X$, co oznacza, że T jest funkcją jednostajnie ciągłą.

Załóżmy teraz, że T jest ciągły w pewnym punkcie $x_0 \in X$, tzn.

$$\|Tx - Tx_0\| \leqslant \varepsilon, \quad \text{gdy} \quad \|x - x_0\| \leqslant \delta.$$

Niech $y \in X$, $\|y\| \leqslant 1$. Wtedy dla wektora $x = x_0 + \delta y$ zachodzi $\|x - x_0\| \leqslant \delta$, zatem $\|\delta T y\| = \|Tx - Tx_0\| \leqslant \varepsilon$. Stąd $\|Ty\| \leqslant \frac{\varepsilon}{\delta}$, a w konsekwencji $\|T\| \leqslant \frac{\varepsilon}{\delta}$.

5.17 Niech $x_n = \chi_{[0,1/n]}$, $y_n = \chi_{[0,n]}$. Wtedy

$$\frac{\|x_n\|_p}{\|x_n\|_q} = n^{1/q - 1/p}, \qquad \frac{\|y_n\|_p}{\|y_n\|_q} = n^{1/p - 1/q}.$$

Gdy $n \to \infty$, to pierwszy z ilorazów dąży do zera, a drugi do nieskończoności. Zatem normy te nie są zgodne.

4.47 Wynika to natychmiast z równości

$$\|x_n - x_0\|^2 = \|x_n\|^2 - 2\,\mathrm{Re}\langle x_n, \, x_0\rangle + \|x_0\|^2,$$

gdyż prawa strona dąży do zera przy $n \to \infty$.

4.49 Z twierdzenia 4.31 wiadomo, że każdy ciągły funkcjonał liniowy x^* na przestrzeni $C_0(\mathbb{R})$ ma postać

$$x^*(x) = \int_0^\infty x(t)\, d\mu(t),$$

gdzie μ jest miarą Radona na prostej $\mathbb{R}$. Słaba zbieżność do zera ciągu $\{x_n\}$ wynika więc natychmiast z twierdzenia Lebesgue'a o zbieżności ograniczonej.

Przedstawimy dowód nie korzystający z opisu przestrzeni $\big(C_0(\mathbb{R})\big)^*$. Zauważmy przedtem, że dla każdego $\varepsilon > 0$ można dobrać taką liczbę $d = d(\varepsilon)$, aby $\big|x(t)\big| < \varepsilon$ dla t poza przedziałem $[-d, d]$, a jeśli $n_1, n_2, n_3, \ldots$ jest ciągiem wskaźników spełniającym warunek $n_{k+1} \geqslant n_k + 2d$, to

$$\big\|\tfrac{1}{m}x_{n_1} + \tfrac{1}{m}x_{n_2} + \cdots + \tfrac{1}{m}x_{n_m}\big\|_\infty \leqslant \frac{1}{m}\|x\|_\infty + \varepsilon.$$

Załóżmy, że ciąg $\{x_n\}$ nie jest zbieżny do zera. Wtedy istnieje taki ciągły funkcjonał liniowego x^* na przestrzeni $C_0(\mathbb{R})$ oraz podciąg $\{x_{n_k}\}$ ciągu $\{x_n\}$, że $\operatorname{Re} x^*(x_{n_k}) \geqslant 1$. Od ciągu wskaźników $\{n_k\}$ możemy przy tym wymagać dodatkowo, by $n_{k+1} > n_k + 2d(\varepsilon)$. Wtedy

$$! \leqslant \big|x^*\big(\tfrac{1}{m}x_{n_1} + \tfrac{1}{m}x_{n_2} + \cdots + \tfrac{1}{m}x_{n_m}\big)\big| \leqslant \big(\tfrac{1}{m}\|x\|_\infty + \varepsilon\big)\|x^*\|.$$

To jednak jest niemożliwe, gdy liczba ε jest dostatecznie mała a m dostatecznie duże.

6.15 Niech x_0 będzie dowolnym elementem kuli oraz

$$U = \big\{x \in X : \big|x_k^*(x - x_0)\big| < \varepsilon \quad \text{dla} \quad k = 1, 2, \ldots, n\big\}$$

dowolnym jego bazowym otoczeniem w słabej topologii. Pokażemy, że U zawiera pewien element sfery. Jeżeli $x_1 \in \bigcap_{k=1}^n \ker x_k^*$, $x_1 \neq 0$, to $x_0 + \lambda x_1 \in U$ dla każdego λ. Funkcja $\|x_0 + \lambda x_1\|$ musi na $[0, \infty)$ przyjmować wszystkie wartości z przedziału $\big[\|x_0\|, \infty\big)$, w szczególności wartość 1.

Słabe domknięcie sfery żadnych innych elementów zawierać już nie może, bo kula jest zbiorem domkniętym wypukłym.

6.18 Pokażemy, że przestrzeń $(\ell^p)^*$ można utożsamić z ℓ^∞. Z każdym ciągiem ograniczonym $y = (y_1, y_2, \ldots) \in \ell^\infty$ zwiążemy funkcjonał liniowy y^* na ℓ^p wzorem

$$y^*(x) = \sum_{n=1}^\infty x_n y_n \quad \text{gdy} \quad x = (x_1, x_2, \ldots) \in \ell^p.$$

Szereg ten jest bezwzględnie zbieżny, a gdy $\|x\| \leqslant 1$, to $|x_n| \leqslant |x_n|^p$ dla wszystkich n, więc

$$|y^*(x) \leqslant \sum_{n=1}^{\infty} |x_n|^p |y_n| \leqslant \|y\|_\infty.$$

Z nierówności tej i jednorodności $\|\lambda x\| = |\lambda|^p \|x\|$ wynika, że gdy $\|x\| < \varepsilon$ to $|y^*(x)| < \varepsilon^{1/p} \|y\|_\infty$. To zaś oznacza ciągłość funkcjonału y^*.

Z drugiej strony, jeżeli $y^* \in (\ell^p)^*$, to dla $n = 1, 2, , \ldots$ połóżmy $y_n = y^*(e_n)$, gdzie e_n jest ciągiem z jedynką na n-tym miejscu. Twierdzimy, że ciąg $\{y_n\}$ jest ograniczony. Gdyby dla pewnego podciągu było $\lim_{k\to\infty} |y_{n_k}| = \infty$, to ciąg wektorów $x_k = y_{n_k}^{-1} e_{n_k}$ byłby zbieżny do zera w ℓ^p, zaś $y^*(x_n) \equiv 1$, co przeczy ciągłości y^*. Skoro ciąg $\{y_n\}$ jest ograniczony, to dla $x \in \ell^p$ mamy $y^*(x) = \sum_{n=1}^\infty x_n y_n$, bo szereg po prawej stronie jest zbieżny.

7.9 Oczywiście zastosujemy tu kryterium Schura. Jako funkcję φ wybierzemy $\varphi(k) = \frac{1}{\sqrt{k-1/2}}$. Ponieważ macierz Hilberta jest symetryczna, więc wystarczy dowodzić tylko jednej z nierówności (7.34).

$$\begin{aligned}\sum_{j=1}^{\infty} \frac{1}{j+k-1}\,\varphi(j) &= \sum_{j=1}^{\infty} \frac{1}{j-1/2+k-1/2} \frac{1}{\sqrt{j-1/2}} \\ &\leqslant \int_0^\infty \frac{dt}{(t+k-1/2)\sqrt{t}} = 2\int_0^\infty \frac{ds}{s^2+k-1/2} \\ &= \frac{2}{\sqrt{k-1/2}} \int_0^\infty \frac{du}{u^2+1} = \frac{\pi}{\sqrt{k-1/2}}.\end{aligned}$$

1.26 Oznaczmy $g = \limsup |x_n|$. Jeżeli $\{y_n\} \in \mathbf{c}_0$, to $\limsup |x_n - y_n| \geqslant g - \limsup |y_n| = g$, więc $\|[x]\| \geqslant g$. Równość otrzymamy dla ciągu $\{y_n\}$ określonego wzorem $y_n = x_n(1 - g/|x_n|)$, gdy $|x_n| > g$, oraz $y_n = 0$ dla pozostałych n.

7.84 Niech $\{y_1, y_2, \ldots, y_n\}$ będzie bazą przestrzeni $T(X)$ oraz $\{y_1^*, y_2^*, \ldots, y_n^*\}$ takim układem funkcjonałów w Y^*, że $y_i^*(y_k) = \delta_{i,k}$ (patrz **??**). Określmy funkcjonał $x_k^* \in X^*$, $k = 1, 2, \ldots, n$, wzorem $x_k^*(x) = y_k^*(Tx)$. Jeżeli wektor Tx ma postać $Tx = \lambda_1 y_1 + \lambda_2 y_2 + \ldots + \lambda_n y_n$, to $\lambda_k = y_k^*(Tx) = x_k^*(x)$.

4.33.2 Wynika to natychmiast z nierówności $\big|x^*(x_n) - x^*(x_0)\big| \leqslant \|x^*\|\, \|x_n - x_0\|$.

5.18 Wystarczy pokazać, że odwzorowanie T ma domknięty wykres. Załóżmy, że $x_n \to x_0$ w X oraz $Tx_n \to y$ w ℓ^1. Ponieważ $\|\ \|_2 \leqslant \|\ \|_1$, więc także $Tx_n \to y$

w ℓ^2. Z założenia wiemy, że T jako odwzorowanie z X do ℓ^2 jest ciągłe. Zatem musi być $y = Tx_0$.

1.38 Dla $x, x' \in X$ i $y, y' \in Y$ równość warstw $[x+y] = [x'+y']$ w przestrzeni $(X+Y)/Y$ zachodzi, gdy $x+y-x'-y' \in Y$, tj. gdy $x - x' \in X \cap Y$, to z zaś oznacza równość warstw $[x] = [x']$ w $X/(X \cap Y)$.

5.19 Operacja różniczkowania z przestrzeni X do $C[0,1]$ ma wykres domknięty. Wynika to ze znanego twierdzenia analizy: jeżeli ciąg x_n zbiega do x_0 oraz x_n' zbiega do y_0 jednostajnie na przedziale $[0,1]$, to x_0 jest funkcją różniczkowalną i $x_0' = y_0$. Z twierdzenia o wykresie domkniętym wnosimy, że mamy do czynienia z operacją ciągłą. Zatem $\|x'\|_\infty \leqslant C\,\|x\|_\infty$ dla $x \in X$. Wynika stąd, że kula jednostkowa przestrzeni X jest zbiorem jednakowo ciągłym. Twierdzenie Arzeli orzeka, że jest zbiorem zwartym a twierdzenie Riesza z kolei, że X musi być przestrzenią skończenie wymiarową.

6.1.e Warunki 1–5 są konieczne. Pierwsze dwa, bo chodzi o topologię Hausdorffa. Trzeci, bo dodawanie wektorów jest operacją ciągłą w zerze. Dla każdego $U \in \mathcal{U}$ istnieją wtedy takie otoczenia zera V_1 i V_2, że $V_1 + V_2 \subset U$. Wystarczy zatem obrać $V \in \mathcal{U}$, by $V \subset V_1 \cap V_2$. Warunek czwarty wynika z ciągłości w zerze operacji mnożenia przez liczby. Dla $U \in \mathcal{U}$ istnieje bowiem taka liczba $\delta > 0$ i otoczenie zera W, że $\lambda W \subset U$ jeśli tylko $|\lambda| \leqslant \delta$. Ponieważ mnożenie przez liczbę różną od zera jest homeomorfizmem, więc zbiór δW jest także otoczeniem zera. Aby dowieść (v), zauważmy, że dla dowolnego $x \in X$ funkcja $\varphi(\mu) = \mu x$ jest ciągła z $\mathbb{C}$ w X i $\varphi(0) = 0$. Zatem dla dowolnego $U \in \mathcal{U}$ istnieje taka liczba $\delta > 0$, że $|\mu| \leqslant \delta$ pociąga $\mu x \in U$. Wystarczy zatem wziąć $\lambda = \delta^{-1}$.

Z warunków (i) oraz (ii) wynika, że rodzina zbiorów $\{x + U : x \in X,\ U \in \mathcal{U}\}$ jest bazą otoczeń pewnej topologii $\mathcal{T}$. Pokażemy, że jest to topologia Hausdorffa i że operacje liniowe są ciągłe w $\mathcal{T}$. Niech $x, y \in X$, $x \neq y$. Zbiór $U \in \mathcal{U}$ wybierzmy tak, by $y - x \notin \mathcal{U}$. Na mocy (iii) i (iv) w $\mathcal{U}$ istnieje taki zbiór V, że $V - V \subset U$. Wtedy $x+V$ i $y+V$ są rozłącznymi otoczeniami punktów x i y. Oznacza to, że $\mathcal{T}$ jest topologią Hausdorffa. Ciągłość operacji dodawania wektorów $(x,y) \to x+y$ w dowolnym punkcie (x_0, y_0) wynika natychmiast z (iii). Jeżeli bowiem $x_0 + y_0 + U$ jest otoczeniem punktu $x_0 + y_0$ oraz $V \in \mathcal{U}$ jest takie, by $V + V \subset U$, to $(x_0 + V) + (y_0 + V) \subset x_0 + y_0 + U$. Pozostaje wykazać, że operacja mnożenia przez liczby jest ciągła. Niech $\lambda_0 x_0 + U$ będzie dowolnym bazowym otoczeniem punktu $\lambda_0 x_0$. Dobierzemy tak liczbę $\varepsilon > 0$ i zbiór $V \in \mathcal{U}$, by obraz otoczenia $W = \{(\lambda, x) : |\lambda - \lambda_0| < \varepsilon,\ x - x_0 \in V\}$ punktu (λ_0, x_0) w $\mathbb{C} \times X$ przez odwzorowanie $(\lambda, x) \to \lambda x$ wpadał w zbiór $\lambda_0 x_0 + U$. Wybierzmy w $\mathcal{U}$ kolejno $V_1 \subset V_2 \subset V_3$, by $V_1 + V_1 + V_1 \subset U$, by $\mu V_2 \subset V_1$ gdy tylko $|\mu| \leqslant 1$ i by $\lambda_0 V_3 \subset V_1$. Niech także δ

będzie taką liczbą dodatnią, że $x_0 \in \delta V_2$. Połóżmy $\varepsilon = \min\{1, 1/\delta\}$ oraz $V = V_3$. Jeżeli $(\lambda, x) \in W$, to $\lambda x = \lambda_0 x_0 + (\lambda - \lambda_0)x_0 + \lambda_0(x - x_0) + (\lambda - \lambda_0)(x - x_0)$. A ponieważ $(\lambda - \lambda_0)x_0 \in (\lambda - \lambda_0)\delta V_2 \subset V_1$, $\lambda_0(x - x_0) \in \lambda_0 V_3 \subset V_1$ oraz $(\lambda - \lambda_0)(x - x_0) \in (\lambda - \lambda_0)V_2 \subset V_1$, więc $\lambda x \in \lambda_0 x_0 + U$.

6.3 Sprawdzamy łatwo, że $\|x + y\| \leqslant \|x\| + \|y\|$ oraz $\|\lambda x\| = |\lambda|^p \|x\|$ dla dowolnych $x, y \in L^p(0,1)$ i $\lambda \in \mathbb{C}$. Wynika stąd natychmiast, że $L^p(0,1)$ jest przestrzenią liniową, a operacje dodawania wektorów i mnożenia wektorów przez liczby są ciągłe. Dowód zupełności jest identyczny jak dla $p \geqslant 1$ (patrz **??**).

Załóżmy, że W jest zbiorem wypukłym $L^p(0,1)$ z niepustym wnętrzem. Nie ograniczając ogólności, możemy założyć, że dla pewnego $r > 0$ zbiór W zawiera kulę $\{x \in L^p(0,1) : \|x\| < r\}$. Ustalmy dowolną funkcję $x \in L^p(0,1)$. Ponieważ całka $\int_0^t |x(s)|^p ds$ jest ciągłą funkcją parametru t (patrz **??**), więc dla każdego n istnieje taki podział $0 = t_0 < t_1 < \cdots < t_n = 1$ odcinka $[0,1]$, że $\int_{t_{k-1}}^{t_k} |x(s)|^p ds = \|x\|/n$ dla wszystkich $k = 1, 2, \ldots, n$. Połóżmy $x_k(s) = nx(s)$ gdy $s \in [t_{k-1}, t_k]$ oraz $x_k(s) = 0$ poza tym. Wtedy $\|x_k\| = n^{p-1}\|x\|$ oraz $x = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$. Ponieważ n^{p-1} dąży do zera przy $n \to \infty$, więc dla dostatecznie dużego n mamy $\|x_k\| < r$, $k = 1, 2, \ldots, n$. Oznacza to, że każda z funkcji x_k, a w konsekwencji także funkcja x leży w W.

3.24.1 Jeżeli wektory $x_1, x_2, \ldots, x_n$ tworzą zbiór ortonormalny oraz $\sum\limits_{k=1}^{n} \lambda_k x_k = 0$, to

$$0 = \Big\langle \sum_{k=1}^{n} \lambda_k x_k, x_m \Big\rangle = \sum_{k=1}^{n} \lambda_k \langle x_k,\, x_m \rangle = \lambda_m$$

dla $m = 1, 2, \ldots, n$.

1.37 TAK, izomorfizmem jest odwzorowanie $\varphi : L^1(0,1) \to L^1(0,1) \times L^1(0,1)$ określone wzorem $\varphi(x)(t) = \big(x(\frac{t}{2}), x(\frac{1+t}{2})\big)$.

3.10 Mamy $\langle x,\, x\rangle = \|x\|^2$, zatem jest spełniona własność (1) iloczynu skalarnego, także $\|x\| = \sqrt{\langle x,\, x\rangle}$. Hermitowskość funkcji $\langle\ ,\ \rangle$ wyprowadzamy łatwo wprost z definicji, podobnie identyczność

$$\langle i\, x,\, y\rangle = i\,\langle x,\, y\rangle.$$

Pozostaje do udowodnienia addytywność funkcji $\langle\ ,\ \rangle$ i jej jednorodność przy mnożeniu przez skalary rzeczywiste.

Korzystając z równości równoległoboku łatwo dowodzimy, że

$$\langle u+v,\, z\rangle + \langle u-v,\, z\rangle = 2\,\langle u,\, x\rangle,$$

a przyjmując tu $v = u$ otrzymujemy

$$\langle 2u,\, z\rangle = 2\langle u,\, z\rangle.$$

Z porównania lewych stron obu tożsamości, po wstawieniu $u = \frac{1}{2}(x+y)$ oraz $v = \frac{1}{2}(x-y)$, dostajemy postulowaną addytywność funkcji $\langle\ ,\ \rangle$.

Z addytywności i ciągłości badanej funkcji wynika już jej jednorodność przy mnożeniu przez liczby rzeczywiste. Istotnie, z addytywności przez indukcję po n dostajemy równość $\langle n\,x,\, y\rangle = n\,\langle x,\, y\rangle$, a po wstawieniu $\frac{1}{n}x$ w miejsce x równość $\langle \frac{1}{n}\,x,\, y\rangle = \frac{1}{n}\langle x,\, y\rangle$; łącznie zatem $\langle \frac{m}{n}\,x,\, y\rangle = \frac{m}{n}\,\langle x,\, y\rangle$. Ponieważ zachodzi także równość $\langle -x,\, y\rangle = -\langle x,\, y\rangle$, funkcja $\lambda \to \langle \lambda\, x,\, y\rangle - \lambda, \langle x,\, y\rangle$ zeruje się na zbiorze wszystkich liczb wymiernych. Skoro jest ciągła a liczby wymierne leżą gęsto w $\mathbb{R}$ — musi być funkcją zerową.

1.4.1 Niech $\{b_t\}_{t\in T}$ będzie bazą Hamela tej przestrzeni. Dla $x = \sum_{t\in T}\lambda_t b_t$ położmy $\|x\| = \max_{t\in T}|\lambda_t|$ ($\lambda_t \neq 0$ tylko dla skończenie wielu $t \in T$). Funkcja $\|\ \|$ jest normą.

4.48 Ciąg $x_n = (0, 0, \ldots, 0, 1, 1, \ldots)$ z zerami do miejsca n-tego jest *-słabo zbieżny do zera. Nie jest jednak słabo zbieżny do zera, gdyż dla funkcjonału LIM z przykładu **??** zachodzi $\mathrm{LIM}\, x_n = 1$.

2.13 Ponieważ wielomian optymalny jest jedyny, więc z nierówności

$$\|x_0 - w_0\|_\infty \geqslant \|x_0 - \mathrm{Re}\, w_0\|_\infty$$

wynika, że musi być rzeczywisty. Gdyby ze zbioru S_0 punktów ekstremalnych funkcji $x_0 - w_0$ nie można było wybrać żadnego alternansu długości $n+1$, to dla pewnej liczby $m < n$ przedział $[a, b]$ da się tak podzielić na $m+1$ przedziałów domkniętych punktami $a < s_1 < s_2 < \cdots < s_m < b$, że w punktach zbioru S_0 wybranych z jednego przedziału funkcja $x_0 - w_0$ przyjmuje wartości tego samego znaku a z przedziałów sąsiednich znaków przeciwnych (oczywiście żaden z punktów s_k nie może należeć do S_0). Wielomian stopnia m

$$w(t) = \prod_{k=1}^{m}(t - s_k)$$

zmienia znak w każdym z punktów s_k, a więc funkcja $(x_0 - w_0)w$ przyjmuje na zbiorze S_0 wartości wyłącznie jednego znaku. Z twierdzenia 2.11 wynika, że w w_0 nie może być w takim razie wielomianem optymalnym.

Dalej rozumujemy jak w przykładzie 2.12. Gdyby w_0 nie był wielomianem optymalnym mimo istnienia alternansu $t_1, t_2, \ldots, t_{n+1}$, to na mocy twierdzenia Kołmogorowa mielibyśmy

$$\min_{1 \leqslant k \leqslant n+1} (-1)^k \operatorname{Re} w(t_k) > 0$$

dla pewnego wielomianu w stopnia niższego niż n. Wielomian $\operatorname{Re} w$ musiałby wtedy zmieniać znak w każdym z przedziałów (t_k, t_{k+1}), a taki wielomian musi mieć stopień przynajmniej n.

LITERATURA UZUPEŁNIAJĄCA

[1] A. Alexiewicz, *Analiza funkcjonalna*, PWN, Warszawa 1969.

[2] R. Engelking, *Topologia ogólna*, PWN, Warszawa 1989.

[3] G. M. Fichtenholz, *Rachunek różniczkowy i całkowy*, PWN 1995.

[4] J. Górniak, T. Pytlik, *Analiza funkcjonalna w zadaniach*, Skrypt Politechniki Wrocławskiej 1976.

[5] R. Larsen, *Functional analysis, an introduction*, Marcel Dekker inc., New York 1973.

[6] G. Meinardus, *Aproksymacja funkcji i jej metody numeryczne*, PWN, Warszawa 1968.

[7] J. Musielak, *Wstęp do analizy funkcjonalnej*, PWN, Warszawa 1989.

[8] W. Rudin, *Analiza funkcjonalna*, PWN, Warszawa 2001.

[9] K. Yosida, *Functional analysis*, Springer-Verlag , Berlin Heildelberg New York 1980.

[10] W. Żelazko, *Algebry Banacha*, PWN, Warszawa 1968.

SKOROWIDZ NAZW

SPIS RZECZY