POLITECHNIKA LUBELSKA

PODRĘCZNIKI

Adam Bobrowski

Analiza funkcjonalna jeden i pół

Szkic o zupełności

Lublin 2015

Tytanicznym wysiłkiem woli, przełamawszy wrodzone opory, zabrał się do myślenia.

Andrzej Pilipiuk „Homo Bimbrownikus"

Żeby przynajmniej przy wszystkim było napisane, czy to na poważnie, czy na jaja.

Andrzej Mleczko

(...) nagle wykrzyknął, bez związku z jakąkolwiek poprzednią rozmową: Jestem oczarowany tymi wyspami. Wyrwał się ni stąd ni zowąd, a propos de bottes, *jak mówią Francuzi (...). Może doprawdy ktoś urok na niego rzucił? Więcej jest czarów na świecie, niż śniło się przeciętnym czarownikom.*

J. Conrad „Zwycięstwo"

Podręczniki – Politechnika Lubelska

Politechnika Lubelska
Wydział Elektrotechniki i Informatyki
ul. Nadbystrzycka 38A
20-618 Lublin

Politechnika Lubelska
Wydział Podstaw Techniki
ul. Nadbystrzycka 38
20-618 Lublin

Adam Bobrowski

Analiza funkcjonalna jeden i pół

Szkic o zupełności

Politechnika Lubelska
Lublin 2015

Recenzent:
Prof. dr hab. inż. Jacek Banasiak

Publikacja wydana za zgodą Rektora Politechniki Lubelskiej

ISBN: 978-83-7947-107-2

Wydawca: Politechnika Lubelska
ul. Nadbystrzycka 38D, 20-618 Lublin
Realizacja: Biblioteka Politechniki Lubelskiej
Ośrodek ds. Wydawnictw i Biblioteki Cyfrowej
ul. Nadbystrzycka 36A, 20-618 Lublin
tel. (81) 538-46-59, email: wydawca@pollub.pl
www.biblioteka.pollub.pl
Druk: TOP Agencja Reklamowa Agnieszka Łuczak
www.agencjatop.pl

Elektroniczna wersja książki dostępna w Bibliotece Cyfrowej PL www.bc.pollub.pl
Nakład: 100 egz.

Spis treści

Wstęp

W roku 1923 H. Steinhaus wydał książeczkę „Czem jest a czem nie jest matematyka” [31]. Niniejszy wstęp wyjaśnić powinien zagadnienie znacznie prostsze „Czym jest analiza funkcjonalna”, ale nawet to zadanie znacznie przekracza możliwości niżej podpisanego. Bo analiza funkcjonalna, która początkowo, choć z trudem, mogłaby być sprowadzona do analizy funkcjonałów, rozwinęła się w dziedzinę matematyki tak obszerną i tak wieloma wypustkami połączoną z innymi, że coraz bardziej wymyka się porządnej definicji, a granice wyznacza sobie tylko sama. Nie jest ona bowiem tylko teorią przestrzeni Banacha [5, 20, 21, 35], ani tym bardziej przestrzeni Hilberta (choć nawet o tej ostatniej można pisać grube, piękne tomy [6, 24]) bo jej częścią jest też teoria lokalnie wypukłych przestrzeni topologicznych [8]. Nie jest też tylko teorią operatorów, choć to być może jeden z najżywiej rozwijających się jej działów. Nie jest też tym bardziej tylko językiem dla procesów stochastycznych czy równań różniczkowych cząstkowych, choć początkowo wielu ją (z pożytkiem) tak widziało. Faktem jest natomiast, że nowoczesne książki ze „zwykłej” analizy coraz częściej traktują analizę funkcjonalną jako swą integralną część, a przestrzeń Banacha jako naturalny punkt wyjścia do rozważań [15, 27]: nauczyliśmy się różniczkować i całkować funkcje o wartościach w przestrzeniach Banacha niemal tak samo jak funkcje rzeczywiste, a wiele twierdzeń „klasycznej” analizy (takie jak twierdzenie Taylora, twierdzenie o funkcji uwikłanej, czy twierdzenie o zbieżności zmajoryzowanej) znalazło uogólnienia na przypadek banachowski. Także rachunek prawdopodobieństwa uprawia się dziś w przestrzeniach Banacha.

I nie ma w tym nic dziwnego. Bo przestrzeń Banacha jest idealnym połączeniem pojęć algebry liniowej, geometrii i topologii. Tworem doskonale dopasowanym do potrzeb współczesnej analizy, która zmagać się musi z zastosowaniami w informatyce, biologii, fizyce, ekonomii, bioinformatyce, statystyce i wielu innych -ikach. Strzałem w dziesiątkę – bo z jednej strony uogólnia znane nam z geometrii prostą, płaszczyznę i przestrzeń trójwymiarową, a z drugiej jest na tyle „konkretna”, że można o niej udowodnić całą gamę pięknych, użytecznych, zaskakujących, intrygujących twierdzeń. Bo jest też punktem wyjścia do dalszych uogólnień. Bo pozwala na rozwijanie potężnych narzędzi. Bo „żyją” w niej rozkłady procesów stochastycznych i rozwiązania równań różniczkowych (już nie mówiąc o żyjących „z niej” matematykach).

Wagę jej odkrycia porównać można chyba tylko do odkrycia koła. Także dlatego, że koła używamy dziś mimochodem wszędzie, nawet nie zastanawiając się nad innymi rozwiązaniami – choć bez niego cywilizacja po prostu nie mogłaby istnieć.

A jednak nawet w Polsce, kolebce analizy funkcjonalnej mało jest studentów, którzy rozumieją „czem jest, a czem nie jest przestrzeń Banacha". Choć to tu (w Krakowie) wspomniany H. Steinhaus dokonał swego największego odkrycia, którym jego zdaniem był Stefan Banach, a ten ostatni wydał (we Lwowie) „sprzedającą się jak świeże bułeczki" monografię „Teoria operatorów liniowych", pierwszy, znakomity podręcznik analizy funkcjonalnej. Brak też jest chyba refleksji nad tym jak użyteczne, potężne i, że tak powiem „centralne", jest w.w. pojęcie. Analiza funkcjonalna wielu wydaje się sucha, abstrakcyjna, zbędna.

Trudno walczyć z uprzedzeniami, ale gdy otrzymałem od p. dr A. Kuczmaszewskiej propozycję poprowadzenia wykładów z analizy funkcjonalnej dla studentów studiów magisterskich (którzy jej elementy mieli już na studiach licencjackich, a część z nich zdążyła ją szczerze znienawidzić) pomyślałem sobie, że to dobry moment na zastanowienie się nad sednem i skutkami zupełności przestrzeni Banacha; że zupełność może być osią wykładu, kluczem do zrozumienia całości. Powstała w ten sposób „Analiza funkcjonalna jeden i pół": jeszcze nie „Analiza funkcjonalna dwa", która by chociaż zarysowała najważniejsze działy tej dziedziny, ale też już nie „Analiza funkcjonalna jeden" – bo patrzymy na znane częściowo wyniki z nieco innej perspektywy, trochę się ucząc, a trochę delektując i zachwycając istotą rozumowania. Przeglądamy też trochę zastosowań w innych działach matematyki. Taki skrypt z natury rzeczy musiał traktować materiał bardzo wybiórczo, a kryterium wyboru pozostać musiało tylko moje „widzimisię". Jest to więc tekst trochę nietypowy – nie jest naukową rozprawą, lecz raczej szkicem, esejem.

Zupełność jest dla niego pojęciem centralnym. To ona pozwala Achillesowi dogonić żółwia, powoduje istnienie $\sqrt{2}$ oraz rozwiązań równań różniczkowych (twierdzenie Picarda), ona też pozwala rozwiązać równanie odnowienia związane z modelem McKendricka (opisującym rozwój populacji) i aproksymować funkcje ciągłe wielomianami. Dzięki niej znajdujemy punkt minimalizujący odległość między zbiorem domkniętym i wypukłym w przestrzeni Hilberta a punktem leżącym poza nim. Ona też jest prawdziwym powodem, dla którego pozornie różne przestrzenie $L^2[0,1]$ i l^2 okazują się takie same – a to pozwala nam rozwiązać równanie ciepła. Bez zupełności nie jest prawdziwe zaskakujące twierdzenie Banacha–Steinhausa (zasada wspólnej ograniczoności); zupełność wreszcie pozwala dobrze zdefiniować eksponentę operatora, a potem wykorzystywać ją do opisu procesów Markowa.

Nie łudzę się, że dzięki temu tekstowi wszyscy studenci Politechniki Lubelskiej pokochają analizę funkcjonalną. Sądząc po ilości ocen niedostatecznych i terminów poprawkowych egzaminu, w których brałem udział, wielu z nich przeklina ją teraz bardziej, niż zanim podjąłem się tych wykładów. Mam tylko nadzieję, że ci z nich, którzy najpierw zrozumieją, a potem wybaczą mój upór w twierdzeniu, że w dzisiejszych czasach nie można być matematykiem, jeśli nie pojmuje się pojęcia zupełności – staną się matematykami lepszymi.

Skrypt ten zawiera nieco za dużo materiału na jeden semestr, to jest 15 dwugodzinnych wykładów – z mojego doświadczenia wynika, że jeśli chce się wszystko studentom dokładnie wytłumaczyć, czasu wystarcza na omówienie większości, ale nie całości. Cechę tę miały już moje odręcznie przygotowywane do dwóch edycji kursu notatki – raz musiałem rezygnować z tej części materiału na rzecz innej, a innym razem odwrotnie. A gdy doprowadzałem ten skrypt do stanu, w którym można go było pokazać potencjalnemu czytelnikowi – lekkomyślnie dodałem jeszcze więcej treści. Mam nadzieję, że z korzyścią dla stanu ducha studentów.

Jeszcze słowo o zadaniach. W każdym rozdziale starałem się umieścić ich kilka i to takich, które dobrze ilustrują materiał. Znakomita większość z nich jest stosunkowo prosta; zadania trudniejsze wyróżniłem znaczkiem ☞. W rozdziale 14 zebrane zostały też ćwiczenia, które w trakcie semestru i potem w sesji egzaminacyjnej przygotowywaliśmy z p. mgr. Adamem Gregosiewiczem, który prowadził część zajęć; wszystkie rozwiązania są jego autorstwa.[1] Jemu także należą się gorące podziękowania za pomoc LaTeXową, oraz uwagi merytoryczne – to on w szczególności zwrócił mi uwagę na nowy, elementarny dowód twierdzenia Banacha–Steinhausa, który podał A. D. Sokal (zob. rozdział 13).

Specjalnie podziękowania należą się też p. Agacie Klementewicz, której wierne notatki z wykładów pozwoliły mi odtworzyć to, co mówiłem, a czego w ferworze walki nie zapisałem.

Liczne brzydkie wypryski na twarzy skryptu udało się usunąć dzięki benedyktyńskiej cierpliwości p. mgr Renaty Rososzczuk, która – z rzadka tylko tarzając się ze śmiechu – całość przeczytała, wyłapała błędy i zasugerowała dermatologiczną kurację. Z kolei recenzent, prof. dr hab. Jacek Banasiak, udzielił kilku ojcowskich rad, których skutek można by porównać do dobroczynnego efektu oddziaływania odpowiednio grubego papieru ściernego lub wręcz tarnika na łeb drewnianej rzeźby.[2] Mimo ich wysiłków efekt końcowy nie pozostawia wielu złudzeń: rozbieżny zez, haczykowaty nos i uśmiech budzący z grobu Bolesława Krzywoustego w pełnej krasie, raczej wykluczają więzy krwi z Julią Roberts. Pozostaje więc nadzieja, że opisywana skryptu twarz mieści się gdzieś w rozsądnym miejscu między obliczem a gębą. I że, choć piękna nie jest, ma w sobie to coś.

Adam Bobrowski
Lublin, 5 października 2015

[1] Znacznie więcej zadań różnych typów znaleźć można w dostępnych na rynku zbiorach S. Prusa i A. Stachury [26] oraz J. Rusinka [28].

[2] Kilka jego światłych uwag pozwoliłem sobie zachować w skrypcie w formie przypisów, by uświadomić Czytelnikowi, z kim musiałem się zmagać.

Rozdział 1

Przestrzenie metryczne zupełne

1.1 Mapy

Wyobraźmy sobie, że mapę województwa lubelskiego położono gdzieś na jego terenie. Można dowieść, a i intuicja podpowiada nam, że to prawda, iż jest na tej mapie punkt, który leży dokładnie w miejscu, które opisuje. Powyższe twierdzenie pozostaje w mocy niezależnie od tego w jakiej skali sporządzona jest mapa. Nie ma też oczywiście znaczenia, czy mówimy o województwie lubelskim, czy wrocławskim. Ważne jest tylko, by mapa obejmowała całe województwo i by w tymże województwie leżała. To samo możemy powiedzieć o mapach miast: Lublina, Warszawy a nawet Tokio, małych miasteczek, powiatów i regionów.

W powyższym twierdzeniu nie ma też nic „dwuwymiarowego". Gdybyśmy sporządzili jednowymiarową „mapę" drogi z Houston do San Antonio (w Teksasie) i położyli ją na tejże drodze, to jeden z punktów na mapie leżałby dokładnie w miejscu, które opisuje. Innymi słowy, gdybyśmy „ścisnęli" odcinek i położyli go na odcinku wyjściowym, to jeden punkt w ogóle by się nie ruszył. Tak samo jest w wymiarach trzech: na „trójwymiarowej mapie" sali wykładowej byłby punkt, który opisywałby to miejsce, w którym leży (o ile tylko mapa ta zostałaby wniesiona do sali).

Znaleźliśmy więc wspólny mianownik, wspólną cechę wielu „przestrzeni". Mają one tę własność, że na każdej ich „mapie" leży przynajmniej jeden punkt, który opisuje sam siebie.

Oczywiście nie wszystkie przestrzenie własność tę posiadają. Na przykład koło z wyrzuconym środkiem, nazwijmy go O. Jedną z jego „map" jest dowolne mniejsze koło z (wyrzuconym) środkiem w tym samym miejscu: obrazem punktu P tego koła jest na mapie punkt P' leżący na promieniu łączącym O i P, którego odległość od O jest k-krotnie mniejsza od odległości między O i P, gdzie k jest daną stałą > 1. Jak widać, żaden punkt mapy nie opisuje samego siebie. Powodem jest oczywiście wyrzucony środek – gdyby był częścią przestrzeni, on właśnie sam siebie by opisywał.

Odróżniliśmy więc dwie klasy przestrzeni: dziurawe i „niedziurawe”. Te ostatnie fachowo nazywa się zupełnymi, te pierwsze – niezupełnymi (dokładniejszą definicję podamy potem).

1.2 Pierwiastki

Z przestrzeniami zupełnymi spotykaliśmy się już w matematyce wielokrotnie, choć być może o tym nie wiedzieliśmy (tak jak molierowski pan Jourdain, który wykrzyknął „U licha! Już przeszło 40 lat mówię prozą, nic o tym nie wiedząc”). Nim przedstawimy jeden z typowych przykładów, zacznijmy od łatwej do dowiedzenia indukcyjnie nierówności Bernouliego:

$$(x+1)^n \geq 1+nx, \text{ gdzie } x \geq -1, n \geq 1. \tag{1.1}$$

Pokażemy za Lechem Maligrandą (zob. [22] i prace cytowane tamże; wcześniej prawie identyczny dowód podał Bengt Åkerberg [2]), iż (1.1) implikuje ważną nierówność między średnią arytmetyczną a średnią geometryczną:[3]

$$\sqrt[n]{x_1 \cdot x_2 \cdots x_n} \leq \frac{x_1+x_2+\cdots+x_n}{n}, \quad x_1, x_2, ..., x_n > 0, n \geq 1. \tag{1.2}$$

Niech $A_n = \frac{x_1+x_2+...+x_n}{n}$. Mamy $\frac{A_n}{A_{n-1}} > 0$, więc przyjmując w (1.1) $x = \frac{A_n}{A_{n-1}} - 1 > -1$, otrzymujemy

$$\left(\frac{A_n}{A_{n-1}}\right)^n \geq 1+n\left(\frac{A_n}{A_{n-1}}-1\right) = \frac{nA_n-(n-1)A_{n-1}}{A_{n-1}} = \frac{x_n}{A_{n-1}}.$$

Stąd $A_n^n \geq x_n A_{n-1}^{n-1}$. To pozwala już udowodnić (1.2) indukcyjnie: dla $n=1$ nierówność jest oczywista, a zakładając, że nierówność ta zachodzi dla $n-1$ otrzymujemy $A_n^n \geq x_n A_{n-1}^{n-1} \geq x_n G_{n-1}^{n-1} = G_n^n$, to znaczy $A_n \geq G_n$, gdzie $G_n = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$.

Korzystając z tej nierówności pokażemy,[4] że dla każdej liczby dodatniej a i każdej liczby naturalnej n istnieje taka liczba b_n, oznaczana $\sqrt[n]{a}$ i zwana pierwiastkiem n-tego stopnia z a, że $b_n^n = a$ (jak się przekonać, że liczba taka jest tylko jedna?). W tym celu rozważmy ciąg $(x_n)_{n\geq 1}$ dany regułą rekurencyjną

$$x_1 = a, \qquad x_{k+1} = \frac{1}{n}\left((n-1)x_k + \frac{a}{x_k^{n-1}}\right). \tag{1.3}$$

Wobec nierówności (1.2) mamy:

$$x_{k+1}^n = \left(\frac{\overbrace{x_k+...+x_k}^{(n-1)\text{czynników}} + \frac{a}{x_k^{n-1}}}{n}\right)^n \geq a\ ,$$

[3] Inny bardzo sprytny dowód (1.2) sugeruje Kazimierz Kuratowski na stronie 18 klasycznego podręcznika [17].

[4] Za Danielem Danersem, Ulmer Seminare 2013, Heft 18, Three Line Proofs.

więc ciąg ten jest ograniczony z dołu przez 0. Korzystając z tego oszacowania, otrzymujemy także:

$$nx_{k+1} = \left((n-1) + \frac{a}{x_k^n}\right) x_k \leq nx_k,$$

co dowodzi, że ciąg ten jest nierosnący. Stąd ma granicę $b_n = \lim_{k\to\infty} x_k$. Przechodząc z k do nieskończoności w (1.3), otrzymujemy $b_n = \frac{1}{n}\left((n-1)b_n + \frac{a}{b_n^{n-1}}\right)$. Proste przekształcenia prowadzą do $b_n^n = a$, co kończy dowód.

Przyjrzyjmy się bliżej temu rozumowaniu. Poza dość oczywistymi rachunkami i oszacowaniami użyliśmy w nim ważnego argumentu:

> każdy ograniczony z dołu nierosnący ciąg liczb rzeczywistych ma granicę.

Jak zobaczymy nieco później (zadanie 1.3) stwierdzenie powyższe jest zakamuflowaną informacją o tym, że zbiór (przestrzeń) liczb rzeczywistych „nie ma dziur", jest zupełny, kompletny.

Czy zupełność ta jest zupełnie oczywista? Tak się wydaje: od dzieciństwa przyzwyczajano nas (ale do czasów Kartezjusza oczywiste to wcale nie było), że zbiór liczb rzeczywistych można przedstawić jako prostą, która przecież dziur nie ma. Tłumaczono nam również, że liczby rzeczywiste można otrzymać jako granice ciągów liczb wymiernych – tak jak to widzieliśmy przed chwilą w przypadku pierwiastka (zob. też niżej); podobnie też jest z liczbą π. Liczby rzeczywiste więc widzimy jako wypełnienie, uzupełnienie zbioru liczb wymiernych, który raczej jest dziurawy.

No właśnie, skąd wiemy, że ten ostatni zbiór nie jest zupełny, co skądinąd stanowi główny powód naszego zainteresowania liczbami rzeczywistymi? Dotychczas zapewne tłumaczono nam, że powodem, dla którego musimy rozszerzyć zbiór liczb wymiernych jest to, że nie można w nim wykonać pewnych operacji algebraicznych. Rzeczywiście, na przykład w świecie liczb wymiernych symbol $\sqrt{2}$ nie ma sensu – inaczej mówiąc, $\sqrt{2}$ nie jest liczbą wymierną. Istotnie, gdyby

$$\sqrt{2} = \frac{l}{m} = \frac{2^{l_2}3^{l_3}\cdots p^{l_p}}{2^{m_2}3^{m_3}\cdots q^{m_q}}$$

gdzie l_2 jest ilością dwójek w rozkładzie l na czynniki pierwsze itd., to podnosząc do kwadratu i mnożąc obie strony przez m^2, otrzymalibyśmy

$$2^{2m_2+1}3^{2m_3}\cdots q^{2m_q} = 2^{2l_2}3^{2l_3}\cdots p^{2l_p}.$$

Wobec faktu, że po lewej stronie mamy nieparzystą ilość potęg dwójki, a po prawej ich ilość parzystą, przeczyłoby to jednoznaczności rozkładu na czynniki proste.

Popatrzmy na to jednak z drugiej strony: rozważmy raz jeszcze rozumowanie prowadzące do istnienia $\sqrt[n]{a}$. Jeśli a jest liczbą wymierną, to wszystkie wyrazy ciągu (1.3) są wymierne. Algebra pozostaje taka sama i dowodzi jego monotoniczności i ograniczoności z dołu. Granica, jak wiemy, okazuje się jednak nie być liczbą wymierną gdy $a = n = 2$ (i w wielu innych przypadkach). Znaleźliśmy zatem ciąg liczb wymiernych, który dąży do liczby niewymiernej. Gdybyśmy nie znali pojęcia liczby niewymiernej (bo przecież niektórym coś, co nie daje się przedstawić w postaci $\frac{m}{n}$, gdzie m i n są całkowite może się wydawać tworem dziwnym, niewiele liczbę przypominającym) musielibyśmy stwierdzenie to ująć tak:

nie każdy ograniczony z dołu nierosnący ciąg liczb wymiernych ma granicę.

To zaś znaczy ni mniej ni więcej tyle, że liczby wymierne nie są zbiorem zupełnym; zbiór liczb wymiernych ma „dziury”. W istocie jest on niezwykle dziurawy: między każdymi dwiema różnymi liczbami wymiernymi leży nieskończenie wiele liczb niewymiernych; w pewnym sensie dziur tych jest znacznie więcej niż „nie-dziur” (zbiór liczb wymiernych jest przeliczalny, a niewymiernych – nieprzeliczalny). To podstawowa różnica między zbiorem liczb wymiernych a zbiorem liczb rzeczywistych – ten ostatni jest zupełny.

1.3 Achilles

Niektórym Czytelnikom przykład z poprzedniego rozdziału może nie przypaść do gustu. Nie wszyscy chcą umieć pierwiastkować. Na taki argument oczywiście trudno odpowiedzieć, bo przecież można też żyć nie znając liter (a nawet być analfabetą-milionerem). Ale mówiąc o zupełności, dotykamy czegoś fundamentalnego, czegoś co od wieków fascynowało filozofów.

Zacznijmy od przypomnienia postaci Zenona z Elei, który „był najwybitniejszym i najsamodzielniejszym z następców Parmenidesa. O życiu jego i charakterze nic prawie nie wiadomo. Żył prawdopodobnie między r. 490 a 430;” [32], str. 21. Zenon znany jest przede wszystkim z przypisywanych mu paradoksów. Reprezentatywnym ich przykładem jest „Achilles”. Wszyscy wiemy, że nikomu nie uda się uciec przed chyżym Achillesem – a już na pewno nie żółwiowi. Pomyślmy jednak, czy rzeczywiście ten ostatni jest bez szans. Załóżmy, że początkowo zwierzątko oddalone jest od swego prześladowcy o odległość d, no i że Achilles biega od niego $\frac{1}{k}$ razy szybciej (gdzie $k < 1$). Zaczynamy nierówną rywalizację – żółw ucieka, a Achilles go goni. Nim ten ostatni doścignie swą ofiarę, żółw przesunie się nieco (a dokładniej na odległość kd). Achilles więc znów będzie miał przed sobą podobne zadanie – dogonić żółwia, który znów się nieco przesunie, a dzielny wojak, by go dogonić będzie musiał najpierw dotrzeć do miejsca, w którym tamten był poprzednio. I tak w kółko, bez końca. A zatem nie dogoni go nigdy. Żółwie górą! Ot i paradoks.

Niektórym rozumowanie takie wyda się zwykłym sofizmatem. Wielu też – tych, którzy dojrzą zasadnicze trudności w obaleniu przedstawionego argumentu – może zwątpić w rzeczywistość tego, co widzi na co dzień. Przypadki takie znane są w historii – na przykład jeden z filozoficznych następców Zenona, „Georgiasz z Leontinoi na Sycylii” przekazał nam „trzy tezy nihilistyczne i sceptyczne, uzasadniane

Rys. 1.1: Żółwie górą!

w sposób wyraźnie zależny od eleatów, a zwłaszcza od Zenona. Tezy te brzmią: 1) nie ma nic, 2) gdyby nawet było coś, to nie byłoby poznawalne, 3) gdyby nawet było poznawalne, to nie mogłoby być przedmiotem porozumienia między ludźmi." [32], s. 23. Co prowadzi nas do śmiałej tezy:

Nieznajomość matematyki szkodzi.

Nie idźmy ani jedną ani drugą drogą, bo obie są zgubne. Paradoksów Zenona nie można ważyć sobie lekce, skoro jak pisze W. Tatarkiewicz ([32] str. 25): „Argumenty Zenona aż do czasów najnowszych zachowały swą pobudzającą moc i były dyskutowane przez najwybitniejszych filozofów, jak Bayle, Descartes, Leibniz, Kant, Hegel, Herbart, Hamilton, Mili, Renouvier, Bergson, Russell.". Nie byle kto sobie nad nimi głowę łamał, warto więc docenić rozwiązanie, które przyszło dopiero z rozwojem analizy matematycznej, a w szczególności teorii zbieżności szeregów liczbowych, w wieku XIX (po ponad dwóch tysiącleciach!).

Pierwszym krokiem do wyjaśnienia paradoksu jest zauważenie luki w końcowej fazie rozumowania Zenona: to, że coś wydarzy się nieskończenie wiele razy nie oznacza, że będzie trwało w nieskończoność. Nieco dokładniej: suma nieskończenie wielu składników wcale nie musi być nieskończona. Przyjrzyjmy się omawianej sytuacji bliżej. Niech t_0 oznacza czas potrzebny Achillesowi do tego, by dotrzeć do miejsca, gdzie na początku był żółw. Jak zauważyliśmy wcześniej, przez ten czas ścigany uciekł na odległość dk. Ten drugi odcinek Achilles przebiegł więc w czasie kt_0. Żółw przesunął się teraz o odcinek k^2d, który Achilles przebiegł w czasie k^2t_0, itd. Teraz już słówko „itd.” nie straszy, bo

$$t_\infty := \sum_{n=0}^{\infty} k^n t_0 < \infty.$$

Właśnie w czasie t_∞ Achilles dogoni żółwia. Dokładniej rzecz biorąc, dla dowolnej liczby naturalnej N mamy

$$\sum_{n=0}^{N} k^n t_0 = \frac{1-k^{N+1}}{1-k} t_0$$

(wystarczy obie strony równości pomnożyć przez $1-k$ a potem zredukować powtarzające się wyrazy), a ostatnie wyrażenie dąży do $\frac{t_0}{1-k}$ (gdy $N \to \infty$). Dąży do tej wartości, dodajmy, bo $k < 1$, co znaczy, że Achilles biegnie szybciej niż żółw; gdyby $k > 1$, suma po prawej stronie dążyłaby do nieskończoności – Achilles okazałby się niezbyt chyży; dla $k = 1$ zresztą też, ale powyższy wzór na sumę wyglądałby inaczej (jak?).

Być może pouczające jest przestudiowanie szczególnego przypadku, gdy $k = \frac{1}{2}$ (Achilles biegnie 2 razy szybciej niż żółw) i $t_0 = 1$. Mamy teraz do czynienia z sumą nieskończonego szeregu

$$1 + \frac{1}{2} + \frac{1}{4} + \dots .$$

Ktoś mógłby powiedzieć, że suma ta nie może być równa dwa, bo przecież nigdy do dwóch „nie dochodzi”. Nawet takiego uparciucha da się jednak przekonać, bo nawet on nie ma wątpliwości, że żadna z sum częściowych

$$1 + \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^N}, \qquad N \geq 1$$

nie przekracza 2 i że wraz ze wzrostem N sumy te są coraz większe. Paradoks Zenona da się więc wyjaśnić, jeśli zgodzimy się, że niemalejący ciąg ograniczony z góry ma granicę – nieważne, czy w danym wypadku będzie to dwa, czy też co innego – odpowiednia suma da nam czas, w którym Achilles dogoni żółwia. Innymi słowy, jeśli zgodzimy się (zob. zadanie 1.4), że

czas jest przestrzenią zupełną, bez dziur.

Jeśli natomiast postulatu tego nie przyjmiemy, będziemy musieli uznać rzeczywistość za sprzeczną z rozumem. Jak widać, zupełność wkracza w nasze życie bez pardonu, z butami. Nie daje się bez niej żyć, a nawet – zob. zadanie 1.8 – zjeść urodzinowego tortu.

1.4 Przestrzenie metryczne, ciągi Cauchy'ego, zupełność

Powróćmy do rozważanego w poprzednim rozdziale przykładu z mapą. Co tak naprawdę sprawiało, że na każdej mapie woj. lubelskiego umieszczonej w tym województwie znajdował się punkt, który opisywał miejsce na którym leżał? Wygląda na to, że kluczem jest tu pojęcie odległości: istnieje taka stała $q < 1$ (zwana skalą mapy), iż odległość między dowolnymi punktami na mapie równa jest q razy odległość między odpowiadającymi im punktami w terenie. Musimy więc zacząć od pojęcia odległości, czyli metryki.

Przypomnijmy, że przestrzenią metryczną nazywamy zbiór X wyposażony w funkcję d (zwaną metryką), przekształcającą $X \times X$ w $\mathbb{R}^+$, która spełnia następujące trzy warunki:

(a) dla wszystkich $x, y \in X$ równość $d(x, y) = 0$ zachodzi wtedy i tylko wtedy, gdy $x = y$,

(b) dla wszystkich $x, y \in X$ mamy $d(x, y) = d(y, x)$,

(c) dla wszystkich $x, y, z \in X$ mamy $d(x, z) \leq d(x, y) + d(y, z)$.

Oczywiście liczbę $d(x, y)$ interpretujemy jako odległość między punktami x i y. Podane wyżej warunki stają się wtedy intuicyjnie oczywiste: odległość między dwoma punktami równa jest zero, wtedy i tylko wtedy, gdy te punkty się pokrywają, odległość mierzona z punktu x do punktu y jest taka sama jak mierzona z punktu y do punktu x i droga wiodąca z punktu x do punktu z poprzez punkt pośredni y nie może być krótsza niż ta wiodąca z x do z bezpośrednio. Warunek (c) jest w związku z tym nazywany nierównością trójkąta, bo wyobrażamy sobie, że punkty x, y i z są jego (tzn. trójkąta) wierzchołkami.

Jeśli pomyślimy sobie teraz o województwie lubelskim jako o przestrzeni metrycznej X (z odległością między punktami mierzoną zwykłą miarką – czasem dość długą), to położenie mapy w jego obrębie zdefiniuje nam przekształcenie X w siebie. Punktowi $x \in X$ przyporządkujemy punkt x', nad którym leży jego odpowiednik na mapie. Otrzymamy wtedy

$$d(x', y') = \kappa d(x, y), \tag{1.4}$$

gdzie jak poprzednio, $\kappa < 1$ jest skalą mapy.

Zastanówmy się teraz, czy przy pomocy metryki można precyzyjnie zdefiniować pojęcie zupełności. Wróćmy do przykładu z poprzedniego rozdziału. Zobaczyliśmy tam, że zbiór liczb wymiernych nie jest zupełny, a dowodem na to było istnienie nierosnącego ciągu liczb wymiernych, który był ograniczony z dołu, ale nie miał granicy (będącej liczbą wymierną).

Pójdziemy właśnie tą drogą, drogą wytyczoną przez Augustina Louisa Cauchy: użyjemy ciągów (nie będziemy natomiast używać pojęcia monotoniczności, bo w przestrzeni metrycznej nie ma ono sensu). Przypomnijmy, że ciąg $(x_n)_{n\geq 1}$ elementów przestrzeni metrycznej X nazywamy zbieżnym, jeśli istnieje taki $x \in X$, że $\lim_{n\to\infty} d(x_n, x) = 0$; innymi słowy, dla dowolnego $\epsilon > 0$ istnieje takie n_0, że $d(x_n, x) < \epsilon$ dla $n \geq n_0$.

Ponadto

ciąg $(x_n)_{n\geq 1}$ nazywamy ciągiem Cauchy'ego (lub ciągiem podstawowym), jeśli dla dowolnego $\epsilon > 0$ istnieje takie $n_0 \geq 0$, że $d(x_n, x_m) < \epsilon$, o ile tylko $n, m \geq n_0$.

Łatwo się sprawdza, że każdy ciąg zbieżny jest ciągiem Cauchy'ego. Jeśli ciąg jest zbieżny do x, to ustalając $\epsilon > 0$, możemy znaleźć takie n_0, że $d(x_n, x) < \frac{\epsilon}{2}$. Z nierówności trójkąta wynika wtedy jednak, że $d(x_n, x_m) < \epsilon$ dla $n, m \geq n_0$, co kończy rozumowanie. Twierdzenie odwrotne nie jest jednak prawdziwe: nie każdy ciąg podstawowy jest zbieżny.

By zrozumieć powód tego stanu rzeczy i związek ciągów Cauchy'ego z istnieniem dziur w przestrzeni, wyobraźmy sobie osobnika, którego przestrzenią życiową jest rozważane w poprzednim podrozdziale koło z usuniętym środkiem, a znana jest mu nasza, euklidesowa metryka. Nie może on popatrzeć z góry na swój ogródek i stwierdzić „O, tu jest dziura. Żyję w przestrzeni, która nie jest zupełna". Jemu po prostu punkt, który ma obie współrzędne równe zero może się wydawać dziwny, może wynaturzony.[5] Narzucamy to ograniczenie, bo chcemy zdefiniować zupełność „bez pomocy z zewnątrz", używając jedynie tego, co widoczne jest „wewnątrz przestrzeni". Nasz osobnik może więc stwierdzić jedynie coś w stylu: „rozważmy ciąg punktów $(x_n)_{n\geq 1}$, w którym $x_n = (\frac{1}{n}, \frac{1}{n}), n \geq 1$." (Studenci wymyślą z łatwością inne ciągi, które osobnikowi temu mogą przyjść do głowy). „Dla dowolnego $\epsilon > 0$, prawie wszystkie odległości

$$d(x_n, x_m) = \sqrt{2}\left|\frac{1}{n} - \frac{1}{m}\right| < \frac{\sqrt{2}}{\min(m, n)}$$

są mniejsze od ϵ: wystarczy, by $n, m > \sqrt{2}\epsilon^{-1}$. Ciąg ten zatem zachowuje się tak jak by był zbieżny." „A przecież" pomyśli dalej „żaden punkt mojej przestrzeni nie może być granicą tego ciągu". „Bo każdy punkt w mojej przestrzeni, powiedzmy x, ma dwie współrzędne, z których co najmniej jedna, nazwijmy ją a, jest dodatnia. A zatem odległość $d(x_n, x)$ będzie większa niż $|a - \frac{1}{n}|$, co dowodzi, że nie dąży do zera (bo $\lim_{n\to\infty} |a - \frac{1}{n}| = a \neq 0$). Znalazłem ciąg Cauchy'ego, który nie jest zbieżny. Przestrzeń, w której żyję jest co najmniej dziwaczna."

Jak widać, nasze hipotetyczne inteligentne żyjątko potrafi przy pomocy ciągów Cauchy'ego zauważyć dziurę w przestrzeni, w której żyje. W ten sposób dochodzimy do definicji zupełności.

[5] Do dziś wielu nie może zaakceptować istnienia zera. Szczególnie moralnego. (Uwaga recenzenta: A takie istnieje? S.J. Lec napisał: „Kiedy znalazłem się na dnie, usłyszałem pukanie od spodu.").

Mówimy, że przestrzeń metryczna jest *zupełna* jeśli każdy ciąg podstawowy ma granicę.

Jak już wiemy, podstawowym przykładem przestrzeni zupełnej jest zbiór liczb rzeczywistych z metryką $d(x,y) = |x-y|$; z kolei najprostszą przestrzenią niezupełną, w istocie „pełną dziur" jest przestrzeń liczb wymiernych (z tą samą metryką).

Zupełności przestrzeni liczb rzeczywistych dowodzić nie będziemy – studenci powinni znać ten podstawowy fakt z kursu analizy matematycznej. Pokażemy natomiast, że przestrzenią zupełną jest każda z przestrzeni $\mathbb{R}^k, k \in \mathbb{N}$, z metryką euklidesową; więcej przestrzeni zupełnych poznamy w dalszym trakcie wykładu.

Dowód jest prosty. Załóżmy, że $(x_n)_{n\geq 1}$, gdzie

$$x_n = (\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,k}) \in \mathbb{R}^k,$$

jest ciągiem Cauchy'ego w $\mathbb{R}^k$. Z łatwej do sprawdzenia nierówności

$$|\xi_{n,i} - \xi_{m,i}| \leq d(x_n, x_m), \tag{1.5}$$

(która zachodzi dla dowolnego $i = 1, \ldots, k$) wynika, że ciąg liczbowy $(\xi_{n,i})_{n\geq 1}$ także jest ciągiem podstawowym. Faktycznie, jeśli odległość $d(x_n, x_m)$ jest dla $n, m \geq n_0$ mniejsza niż ϵ to tym bardziej mniejsze od ϵ jest $|\xi_{n,i} - \xi_{m,i}|$. Z zupełności $\mathbb{R}$ wynika zatem istnienie granicy $\xi_i = \lim_{n\to\infty} \xi_{n,i}$. Wektor $x = (\xi_1, \xi_2, \ldots, \xi_k)$ jest elementem $\mathbb{R}^k$. Wystarczy pokazać, że jest on granicą $(x_n)_{n\geq 1}$. Z założenia wiemy, że dla dowolnego $\epsilon > 0$ istnieje takie $n_0(\epsilon)$, że

$$d(x_n, x_m) = \sqrt{\sum_{i=1}^{k} (\xi_{n,i} - \xi_{m,i})^2} < \epsilon$$

o ile tylko $n, m \geq n_0(\epsilon)$. Przechodząc z m do ∞, z nierówności tej otrzymujemy

$$d(x_n, x) = \sqrt{\sum_{i=1}^{k} (\xi_{n,i} - \xi_i)^2} \leq \epsilon,$$

dla $n \geq n_0(\epsilon)$. A zatem dla $n \geq n_1(\epsilon) := n_0(\epsilon/2)$, mamy $d(x_n, x) < \epsilon$, a to kończy dowód.

1.5 Zadania

Zadanie 1.1. Udowodnij nierówność Bernoulliego.

Zadanie 1.2. Sprawdź, że $\sqrt{3}$ i $\sqrt{5}$ nie są liczbami wymiernymi.

Zadanie 1.3. ☞ Udowodnij, że z zupełności przestrzeni liczb rzeczywistych wynika, iż każdy nierosnący ciąg ograniczony z dołu ma granicę. Udowodnij też, że to, iż każdy nierosnący ciąg ograniczony z dołu ma granicę pociąga za sobą zupełność przestrzeni liczb rzeczywistych.

Zadanie 1.4. Udowodnij, że jeśli każdy ciąg nierosnący ograniczony z dołu ma granicę, to ma ją też każdy ciąg niemalejący ograniczony z góry i vice versa.

Zadanie 1.5. Wyposażmy $X = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\} \subset \mathbb{R}$ w standardową metrykę $d(x,y) = |x-y|$. Czy jest to przestrzeń zupełna? A $X \cup \{0\}$?

Zadanie 1.6. Udowodnij (1.5). Wskazówka: suma nieujemnych składników jest większa lub równa każdemu z nich.

Zadanie 1.7. Jeśli już damy się przekonać, że Achilles dogoni żółwia, to tego, iż dogoni go po czasie $t_\infty = \frac{t_0}{1-k}$ można też dowieść nie sumując szeregu, lecz wyliczając t_∞ z zależności:

$$k(d+x) = x, \qquad t_\infty = \frac{x}{v_{\text{żółwia}}}, \qquad t_0 = \frac{d}{v_{\text{Achillesa}}} = \frac{dk}{v_{\text{żółwia}}}.$$

Podaj szczegóły odpowiedniego rozumowania.

Zadanie 1.8. Jubilat obżera się po wyjściu gości urodzinowym ciastem, którego ze skąpstwa nie postawił na stole. Chcąc, by chwila przyjemności trwała wiecznie je najpierw połowę tortu wielkości opony samochodowej, potem połowę połowy, potem połowę tego, co zostało i tak ad infinitum. Studenci, którzy nie chcą mieć nudności lub pragną uratować tracących na wieki pracę cukierników, proszeni są o wyjaśnienie tego „paradoksu”.

Rozdział 2

Zasada Banacha

2.1 Zasada Banacha

Tematem tego rozdziału jest pierwsze, proste, ale bardzo eleganckie i użyteczne twierdzenie, w którym zupełność odgrywa kluczową rolę. Wykorzystuje ono „geometryczne” intuicje, przedstawione w poprzednim rozdziale, w przykładzie dotyczącym map, ale przestawia je w terminach abstrakcyjnych. Jak zobaczymy później, takie uogólnienie pozwala na cały szereg zastosowań.

Zacznijmy jednak od prostego wniosku z definicji przestrzeni zupełnej.

Twierdzenie 2.1. *Domknięty (niepusty) podzbiór przestrzeni metrycznej zupełnej jest przestrzenią zupełną (gdy wyposażymy go w metrykę odziedziczoną z przestrzeni, której jest podzbiorem).*

Dowód. Przypomnijmy, że podzbiór $Y \subset X$ przestrzeni metrycznej nazywamy domkniętym, jeśli granica dowolnego ciągu $(x_n)_{n\geq 1}$ elementów Y także należy do Y. Zwróćmy uwagę na to, że dla dowolnych dwóch elementów $y_1, y_2 \in Y \subset X$ zdefiniowana jest ich odległość w przestrzeni X; Y wyposażona w tę odległość jest także przestrzenią metryczną. To właśnie mamy na myśli pisząc, że Y dziedziczy metrykę z X.

Musimy dowieść, że każdy ciąg Cauchy'ego $(x_n)_{n\geq 1}$ elementów Y jest zbieżny (do elementu z Y). Ciąg $(x_n)_{n\geq 1}$ jest jednakże równocześnie ciągiem elementów X. Jest więc zbieżny do pewnego $x \in X$. Z domkniętości Y wnioskujemy, że $x \in Y$, a to kończy dowód. $\square$

Uwaga. Znaczek $\square$, który widzą tu Państwo być może po raz pierwszy, a który towarzyszyć nam będzie przez cały skrypt, jest standardowym graficznym oznaczeniem końca dowodu.

Twierdzenie 2.2 (Zasada Banacha)**.** *Załóżmy, że funkcja T odwzorowuje zupełną przestrzeń metryczną w siebie w taki sposób, że istnieje taka stała $q \in (0, 1)$, iż dla dowolnych $x, y \in X$ zachodzi nierówność*

$$d(Tx, Ty) \leq qd(x, y). \tag{2.1}$$

Istnieje wtedy dokładnie jeden punkt stały $\tilde{x} \in X$ odwzorowania T (co znaczy, że dla tego i tylko dla tego punktu mamy $T\tilde{x} = \tilde{x}$); co więcej $\tilde{x}$ jest granicą $\tilde{x} = \lim_{n\to\infty} T^n x$ dla dowolnego $x \in X$.

$T^n x$ definiuje się tu indukcyjnie: $T^{n+1}x = T(T^n x)$ i $T^1 x = x$.

Dowód. Wybierzmy dowolne $x \in X$. Pokażemy najpierw, że ciąg $(T^n x)_{n\geq 1}$ jest zbieżny. Skoro X jest przestrzenią Banacha, wystarczy dowieść, że $(T^n x)_{n\geq 1}$ jest ciągiem podstawowym. W tym celu oszacujemy odległość między $T^n x$ a $T^m x$ gdzie $m > n$. Z wielokrotnie stosowanej nierówności trójkąta wynika, że

$$d(T^n x, T^m x) \leq d(T^n x, T^{n+1}x) + d(T^{n+1}x, T^{n+2}x) + \cdots + d(T^{m-1}x, T^m x).$$

Pierwszy ze składników prawej strony możemy, używając n-krotnie założenia (2.1), oszacować przez $q^n d(x, Tx)$. Podobnie drugi szacujemy przez $q^{n+1}d(x, Tx)$, itd. Stąd

$$\begin{aligned} d(T^n x, T^m x) &\leq d(x, T^1 x)[q^n + q^{n+1} + \cdots + q^{m-1}] \\ &= q^n d(x, Tx)[1 + q + q^2 + \cdots + q^{m-n-1}] \\ &\leq q^n d(x, Tx)[1 + q + q^2 + \cdots + q^{m-n-1} + \cdots] \\ &\leq \frac{q^n d(x, Tx)}{1-q}. \end{aligned}$$

Zwróćmy uwagę, że prawa strona nie zależy tu już w żaden sposób od m. To dowodzi, że $(T^n x)_{n\geq 1}$ jest ciągiem Cauchy'ego, bo w.w. prawa strona dąży do zera, gdy $n \to \infty$. Jeśli chcemy, by odległość między $T^n x$ a $T^m x$ była mała, wystarczy wybrać odpowiednio duże n.

Niech teraz $\tilde{x}$ będzie granicą $(T^n x)_{n\geq 1}$. Założenie (2.1) pokazuje, że T jest odwzorowaniem ciągłym. Mamy zatem $\lim_{n\to\infty} T(T^n x) = T(\tilde{x})$. Z drugiej strony $T(T^n x) = T^{n+1}x$, a ciąg $\left(T^{n+1}x\right)_{n\geq 1}$ ma oczywiście tę samą granicę co $(T^n x)_{n\geq 1}$. To dowodzi, że $\tilde{x} = \lim_{n\to\infty} T^{n+1}x = \lim_{n\to\infty} T(T^n x) = T(\tilde{x})$, a to znaczy, że $\tilde{x}$ jest punktem stałym przekształcenia T. Fakt, że punkt taki jest tylko jeden wynika znów z (2.1): gdyby były dwa punkty stałe, powiedzmy $\tilde{x}$ i $\hat{x}$, to mielibyśmy

$$d(\tilde{x}, \hat{x}) = d(T\tilde{x}, T\hat{x}) \leq q d(\tilde{x}, \hat{x}).$$

Wobec tego, że $q \in (0, 1)$, nierówność ta dowodzi, iż $d(\tilde{x}, \hat{x}) = 0$, co znaczy, że $\tilde{x} = \hat{x}$. □

Wnioskiem z zasady Banacha jest chociażby twierdzenie o mapach z poprzedniego rozdziału. Jak widzieliśmy bowiem, z każdą mapą położoną na odwzorowanym na niej terenie można związać przekształcenie tego terenu, spełniające zależność (1.4). Oczywiście jest to zależność znacznie mocniejsza niż (2.1) (z $q = \kappa$). Tym samym założenia zasady Banacha są spełnione ... o ile wspomniany teren można uważać za przestrzeń zupełną. Skoro jest on jednak podzbiorem $\mathbb{R}^2$,[6] to wobec

[6] Uwaga recenzenta: przy założeniu, że Ziemia jest płaska, co jest dyskusyjne.

twierdzenia 2.1 wystarczy założyć, że województwo, czy miasto, które mamy na myśli zawiera swoje granice, że – innymi słowy – tworzy ono zbiór domknięty.

Przy tych naturalnych założeniach wnioskujemy, że wspomniane wyżej odwzorowanie ma dokładnie jeden punkt stały. To jednak oznacza, że dokładnie jeden punkt na mapie leży w miejscu, które opisuje.

2.2 Zadania

Zadanie 2.1. Niech $y \in \mathbb{R}^k$ będzie danym wektorem. Rozważmy odwzorowanie $x \mapsto x + y$. O ile $y \neq 0$, nie ma ono punktów stałych. Na podstawie tego przykładu odpowiedz na pytanie, czy w zasadzie Banacha można założenie $q \in (0, 1)$ osłabić do $q \in (0, 1]$.

Zadanie 2.2. Pokaż, że odwzorowanie $T : [1, \infty) \to [1, \infty)$ dane wzorem $f(x) = \frac{3}{4}(x + \frac{1}{x})$ jest kontrakcją (przy standardowej metryce). Co jest jego punktem stałym?

Zadanie 2.3. (Por. [12] str. 201). Niech $\alpha \geq 1$ będzie ustaloną liczbą. Udowodnij, że równanie

$$x^\alpha - (1 + \alpha)(1 - x) = 0$$

ma dokładnie jedno rozwiązanie w przedziale $(0, 1)$. Wskazówka: rozważ odwzorowanie $T : [0, 1] \to [0, 1]$ dane wzorem $T(x) = 1 - \frac{x^\alpha}{1+\alpha}$.

Rozdział 3

Twierdzenie Picarda

Pierwszym zastosowaniem zasady Banacha, a tym samym zastosowaniem pojęcia zupełności, które tu chcemy omówić, jest twierdzenie Picarda o istnieniu i jedyności rozwiązań równań różniczkowych. Rozumowanie, które tu przeprowadzimy jest typowe dla analizy funkcjonalnej, która patrzy na funkcje (a więc w szczególności i rozwiązania równań różniczkowych) jako na punkty pewnej przestrzeni – i to z własności tej przestrzeni (takich, jak zupełność) wnioskuje o istnieniu (lub nie istnieniu) szukanego obiektu. Jej przedmiotem nie są więc – tak jak a analizie „klasycznej" własności konkretnej funkcji (takie jak monotoniczność, ciągłość, całkowalność), lecz własności całej przestrzeni.

Oczywiście głównym celem tego rozdziału jest uwydatnienie roli zupełności przestrzeni funkcji ciągłych, ale jest też i drugi powód: w Lublinie, mieście prof. prof. Adama Bieleckiego i Kazimierza Goebla, którzy twórczo rozwinęli teorię punktów stałych, nie wypada przemilczeć eleganckiego dowodu twierdzenia Picarda wykorzystującego zasadę Banacha i sprytnie dobraną normę, zwaną normą Bieleckiego [3, 8].

3.1 Istnienie i jedyność rozwiązań równań różniczkowych

Rozważmy równanie różniczkowe zwyczajne (rozwiązane względem pochodnej) postaci:

$$u'(t) = f(t, u(t)), \qquad t \geq 0, \tag{3.1}$$

w którym u jest funkcją szukaną, a f daną. Jeśli f jest funkcją o wartościach rzeczywistych, to szukamy u o wartościach rzeczywistych. Dla tak zwanego równania Malthusa $f(t, u) = au$, a dla równania logistycznego $f(t, u) = a\left(1 - \frac{u}{K}\right)u$. ($a$ i K są tu danymi stałymi). Jeśli $f : \mathbb{R}^+ \times \mathbb{R}^k \to \mathbb{R}^k$, gdzie k jest jakąś liczbą naturalną, to szukamy u, którego wartości są wektorami k wymiarowymi. Inaczej mówiąc, szukamy k funkcji rzeczywistych $u_1, u_2, ..., u_k$, które są współrzędnymi u:

$$u(t) = (u_1(t), u_2(t), \ldots, u_k(t)).$$

Na przykład w przypadku tak zwanego równania Lotki–Volterry szukamy funkcji $u = (u_1, u_2) = (x, y)$ o wartościach w przestrzeni $\mathbb{R}^2$; a funkcja f dana jest wzorem

$$f(t, x, y) = (ax - bxy, -cy + dxy) \tag{3.2}$$

i przekształca $\mathbb{R}^+ \times \mathbb{R}^2$ w $\mathbb{R}^2$; trójce liczb przyporządkowuje parę; a, b, c i d są tu stałymi. (We wszystkich dotychczasowych przykładach, f nie zależy w żaden sposób od swojego pierwszego argumentu, interpretowanego jako czas; to znaczy, że mamy tu do czynienia z tak zwanymi równaniami autonomicznymi. Już za chwilę zobaczymy przykład równania nieautonomicznego.)

Podstawowe twierdzenie teorii równań różniczkowych – twierdzenie Peano – mówi, że jeśli funkcja f jest ciągła w otoczeniu punktu $(0, \alpha) \in R^{k+1}$, to w pewnym małym otoczeniu $t = 0$ istnieje rozwiązanie równania (3.1) (zwane również jego *całką*) z warunkiem początkowym $u(0) = \alpha$. Podkreślmy, że w.w. otoczenie może być małe i w ogólności jego wielkość zależy od f i od punktu. Na przykład rozwiązaniem równania

$$u'(t) = [u(t)]^2 t^3 \tag{3.3}$$

z warunkiem początkowym $u(0) = \alpha$ jest

$$u(t) = \frac{4\alpha}{4 - \alpha t^4}.$$

Dla ujemnych α, dziedziną u jest cała oś rzeczywista, ale dla dodatnich mianownik przyjmuje wartość zero dla $t = \sqrt[4]{\frac{4}{\alpha}}$ i definicja traci sens. Funkcja u jest zatem określona tylko w przedziale $[0, \sqrt[4]{\frac{4}{\alpha}})$ (zob. rys. 3.1).

Podkreślmy, że twierdzenie Peano stwierdza jedynie, iż rozwiązanie równania różniczkowego istnieje lokalnie, ale nic nie mówi o jedyności. Może się więc zdarzyć, że z jednego punktu wychodzić będzie wiele rozwiązań. Jest tak na przykład dla funkcji $f(t, u) = 2\sqrt{u}$. Związane z nią zagadnienie początkowe:

$$u'(t) = 2\sqrt{|u|}, \qquad u(0) = 0, \tag{3.4}$$

ma dwa różne rozwiązania: $u(t) = 0$ i $u(t) = t^2$; w istocie całek jest nieskończenie wiele: dla każdego $a > 0$ definiując

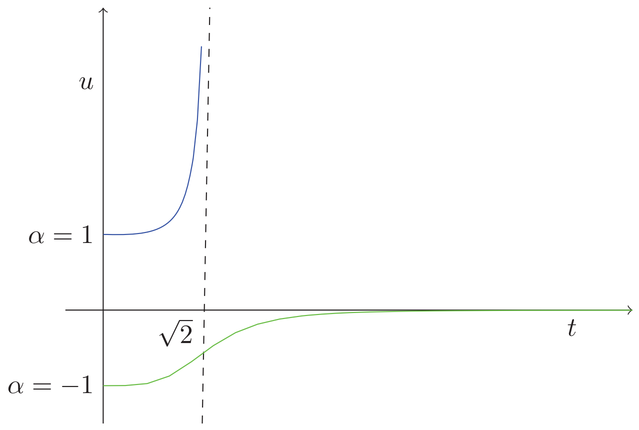
$$u(t) = \begin{cases} 0, & \text{dla } t \in [0, a], \\ (t - a)^2, & \text{dla } t > a, \end{cases}$$

otrzymujemy inną. Podobnie jest z problemem:

$$u'(t) = \sqrt[3]{u(t)}, \qquad u\ (0) = 0,$$

który ma trzy rozwiązania: $u(t) = 0, u(t) = -\left(\frac{2}{3}t\right)^{\frac{3}{2}}$ i $u(t) = \left(\frac{2}{3}t\right)^{\frac{3}{2}}$, z których można sklejać kolejne.

Z matematycznego punktu widzenia warunkiem, który zapewnia istnienie i jedyność rozwiązań jest warunek Lipschitza: słynne twierdzenie Picarda mówi, iż



Rys. 3.1: Lokalność rozwiązań równań różniczkowych: wykres niebieski to rozwiązanie równania (3.3) z $\alpha = 1$, a zielony z $\alpha = -1$. To, że ten ostatni wydaje się od pewnego miejsca pokrywać z osią $u = 0$ nie oznacza, iż rozwiązania rzeczywiście równe są zero, lecz to, że program wykorzystany do rysowania otrzymywał wartości tak małe, że do zera je zaokrąglał.

jeśli f jest ciągła i istnieje taka stała L (od nazwiska Lipschitza), że dla dowolnych u i $\tilde{u}$ mamy

$$\|f(t,u) - f(t,\tilde{u})\| \leq L\|u - \tilde{u}\|, \tag{3.5}$$

to dla każdego warunku początkowego rozwiązania istnieją, są zdefiniowane na całej osi i do tego wyznaczone jednoznacznie. W warunku powyższym $\|\cdot\|$ oznacza normę euklidesową:

$$\|(\xi_1, ..., \xi_k)\| = \sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_k^2},$$

lub jakąkolwiek inną, równoważną, na przykład normę maksimum:

$$\|(\xi_1, ..., \xi_k)\|_{\max} = \max\{|\xi_1|, |\xi_2|, \cdots, |\xi_k|\}.$$

(Nie jest to pierwszy kurs analizy funkcjonalnej, więc student wybaczy mi, że posługuję się tu pojęciem normy, które formalnie zdefiniowane zostanie nieco później).

W przypadku równania (3.4) otrzymaliśmy zamiast jednego całą rodzinę rozwiązań, gdyż funkcja $f(t,u) = \sqrt{u}$ nie jest lipschitzowska. Nawet jeśli ograniczymy się do $u \in [0,1]$, nie znajdziemy takiej stałej L, by

$$\sqrt{u} = \|f(t,u) - f(t,0)\| \leq L\|u - 0\| = Lu\ . \tag{3.6}$$

Czy równanie Lotki–Volterry jest związane z funkcją lipschitzowską? Tak i nie. By to wyjaśnić, będziemy bardziej precyzyjnie mówić, że funkcja spełniająca warunek (3.5) ma *globalną stałą Lipschitza* (ze względu na u). Otóż funkcja f z równania

(3.2) stałej globalnej nie posiada, ale posiada stałe lokalne, co znaczy, że dla dowolnego $r > 0$ istnieje takie $L(r)$, iż

$$\|f(t,u) - f(t,\tilde{u})\| \leq L(r)\|u - \tilde{u}\|, \tag{3.7}$$

o ile tylko normy u i $\tilde{u}$ nie przekraczają r. Rzeczywiście, warunek (3.7) zachodzi w tym przypadku z $L(r) = \max\{a+br, c+dr\}$ (zob. zad 3.5). Zauważmy, że mamy tu typową sytuację, w której stałe $L(r)$ rosną wraz z r.

Niestety, lokalna lipschitzowskość, to znaczy istnienie lokalnych stałych Lipschitza, choć zapewnia jednoznaczność rozwiązań, nie gwarantuje, że są one zdefiniowane dla wszystkich $t \geq 0$. Widać to chociażby z przykładu,

$$u'(t) = 1 + [u(t)]^2.$$

Jak łatwo sprawdzić (zadanie 3.4), funkcja $f(u) = 1 + u^2$ jest lokalnie, ale nie globalnie lipschitzowska. Jej wszystkie rozwiązania są postaci $u(t) = \operatorname{tg}(t + C)$, gdzie C jest dowolną stałą, z odpowiednią dziedziną. W szczególności, rozwiązanie wychodzące z $\alpha = 0$ ma postać $u(t) = \operatorname{tg} t$ i jak widać jest zdefiniowane tylko dla $t \in [0, \frac{\pi}{2})$.

3.2 Przestrzeń funkcji ciągłych

Kluczem do dowodu twierdzenia Picarda jest zupełność przestrzeni funkcji ciągłych na odcinku $[a,b]$, gdzie $a < b$ to dowolne liczby rzeczywistych. Dokładniej, niech $C[a,b]$ oznacza zbiór funkcji ciągłych określonych na przedziale $[a,b]$ i przyjmujących wartości w $\mathbb{R}^k$. (Gdybyśmy chcieli być bardziej formalni, powinniśmy pisać $C([a,b],\mathbb{R}^k)$, ale ze względu na prostotę oznaczeń piszemy tak jak to zaznaczyliśmy wyżej). Wybierzmy z tego zbioru dwie funkcje, x i y. Funkcja $\phi(t) = \|x(t) - y(t)\|$, gdzie $\|\cdot\|$ oznacza normę w przestrzeni $\mathbb{R}^k$ jest wtedy ciągła i ma wartości rzeczywiste. Zgodnie ze znanym twierdzeniem z analizy matematycznej, zwartość przedziału $[a,b]$ powoduje, iż istnieje punkt $t_0 \in [a,b]$, w którym ϕ przyjmuje wartość największą (zob. rysunek 3.2). Możemy zatem zdefiniować odległość między x i y wzorem:

$$d(x,y) = \max_{t\in[a,b]} \|x(t) - y(t)\| = \|x(t_0) - y(t_0)\|. \tag{3.8}$$

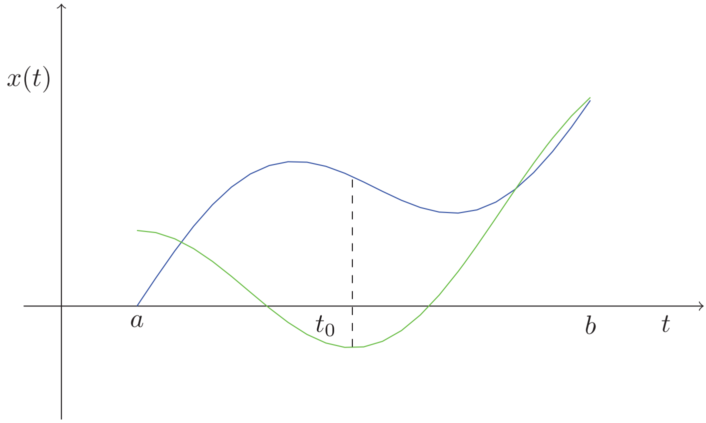
Podkreślmy, że występujące tu t_0 zmienia się wraz ze zmianą x i y.

Łatwo sprawdza się (zadanie 3.7), że tak zdefiniowana funkcja jest metryką. Pokażemy, że $C[a,b]$ z tą metryką jest zupełna. W tym celu pomyślmy o ciągu $(x_n)_{n\geq 1}$ elementów $C[a,b]$, który jest ciągiem Cauchy'ego: musimy pokazać, że jest on zbieżny. Najpierw pokażemy, że dla każdego $s \in [a,b]$, ciąg $(x_n(s))_{n\geq 0}$ jest zbieżny. Z założenia wiemy, że dla dowolnego $\epsilon > 0$ istnieje takie n_0, że dla $n,m \geq n_0$ mamy

$$\max_{t\in[a,b]} \|x_n(t) - x_m(t)\| < \epsilon. \tag{3.9}$$

Skoro

$$\|x_n(s) - x_m(s)\| \leq \max_{t\in[a,b]} \|x_n(t) - x_m(t)\|,$$



Rys. 3.2: Metryka supremum: odległość między dwiema funkcjami to maksymalna długość odcinka prostopadłego do osi $0t$ łączącego ich wykresy; na rysunku jest to długość odcinka narysowanego linią przerywaną.

to wynika stąd, że ciąg $(x_n(s))_{n\geq 0}$ jest ciągiem Cauchy'ego. Jest to jednak ciąg, którego wyrazy są elementami przestrzeni $\mathbb{R}^k$, która jest zupełna. Stąd istnieje $x(s) \in \mathbb{R}^k$, do którego ciąg nasz zbiega.

Funkcja x, przyporządkowująca $s \in [a, b]$ granicę $x(s) = \lim_{n\to\infty} x_n(s)$ ciągu $(x_n(s))_{n\geq 0}$ jest naturalnym kandydatem na granicę ciągu $(x_n)_{n\geq 0}$. Musimy jednak dowieść, że jest to element przestrzeni $C[a, b]$ (a zatem, że x jest funkcją ciągłą) i że $(x_n)_{n\geq 1}$ dąży do x w sensie metryki d. Z warunku (3.9) wiemy, że jeśli $n, m \geq n_0$ to, niezależnie od tego, jakie jest t, zachodzi nierówność

$$\|x_n(t) - x_m(t)\| < \epsilon. \tag{3.10}$$

Przechodząc do granicy z m, przekonujemy się, że niezależnie od tego, jakie jest $t \in [a, b]$, mamy

$$\|x_n(t) - x(t)\| \leq \epsilon, \tag{3.11}$$

o ile tylko $n \geq n_0$. To jednak znaczy, że $d(x_n, x)$ jest dobrze zdefiniowane i że $\lim_{n\to\infty} d(x_n, x) = 0$.

Pozostał nam dowód, że x jest funkcją ciągłą. W tym celu ustalmy $t \in [a, b]$ i $\epsilon > 0$. Wybierzmy tak n_0, by dla $n \geq n_0$ zachodziła nierówność (3.11) z ϵ zamienionym na $\frac{\epsilon}{3}$. Funkcja x_{n_0} jest ciągła, więc możemy dobrać tak $\delta > 0$, by z tego, że $t \in [a, b]$ i $|t - t_0| < \delta$ wynikało, iż $\|x_{n_0}(t) - x_{n_0}(t_0)\| < \frac{\epsilon}{3}$. Dla takich t mamy

$$\begin{aligned}\|x(t) - x(t_0)\| &\leq \|x(t) - x_{n_0}(t)\| + \|x_{n_0}(t) - x_{n_0}(t_0)\| + \|x_{n_0}(t_0) - x(t_0)\| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

Wobec tego, że ϵ i t były dowolne, dowodzi to, że x jest ciągła.

Przypomnijmy, że wyżej opisana funkcja d w $C[a,b]$ nazywana jest metryką supremum, a zbieżność w sensie d – zbieżnością jednostajną na odcinku $[a,b]$.

3.3 Metryki równoważne

Przypomnijmy, że dwie metryki d_1 i d_2 zdefiniowane w tej samej przestrzeni X nazywamy równoważnymi (dokładniej: równoważnymi lipschitzowsko) jeśli istnieją takie stałe dodatnie m i M, że dla każdych $x, y \in X$ zachodzi nierówność

$$md_1(x,y) \leq d_2(x,y) \leq Md_1(x,y).$$

Twierdzenie 3.1. *Jeśli przestrzeń X wyposażona w metrykę d_1 jest przestrzenią zupełną, a d_2 jest metryką równoważną d_1, to X z metryką d_2 też jest przestrzenią zupełną.*

Dowód. Rozważmy ciąg $(x_n)_{n\geq 1}$, który spełnia warunek Cauchy'ego w metryce d_2. Lewa nierówność definicji metryk równoważnych dowodzi, że ciąg ten spełnia też warunek Cauchy'ego w metryce d_1 (bo $d_1(x_n, x_m) \leq \frac{1}{m} d_2(x_n, x_m)$). Z założenia istnieje więc taki $x \in X$, że $\lim_{n\to\infty} d_1(x_n, x) = 0$. Stosując prawą nierówność definicji oraz twierdzenie o trzech ciągach wnioskujemy, że

$$\lim_{n\to\infty} d_2(x_n, x) = 0,$$

co dowodzi tezy. □

Oto przykład metryk równoważnych w przestrzeni $C[a,b]$. Niech ψ będzie dowolną funkcją ciągłą na $[a,b]$, która przyjmuje tylko wartości dodatnie. Dla $x, y \in C[a,b]$ rozważamy

$$d_\psi(x,y) = \sup_{t\in[a,b]} \psi(t)\|x(t) - y(t)\|. \tag{3.12}$$

Supremum to istnieje i jest skończone z tych samych powodów, dla których skończone jest (3.8). Łatwo też sprawdzić, że d_ψ jest metryką (zadanie 3.8). Twierdzimy, że metryka ta jest równoważna metryce supremum. Faktycznie, z dodatniości i ciągłości ψ wynika istnienie takich liczb dodatnich, że

$$m \leq \psi(t) \leq M$$

dla wszystkich $t \in [a,b]$. Stąd dla dowolnych $x, y \in C[a,b]$ mamy

$$\begin{aligned} md(x,y) &= \sup_{t\in[a,b]} m\|x(t) - y(t)\| \\ &\leq \sup_{t\in[a,b]} \psi(t)\|x(t) - y(t)\| \quad (= d_\psi(x,y)) \\ &\leq \sup_{t\in[a,b]} M\|x(t) - y(t)\| = Md(x,y), \end{aligned}$$

co dowodzi naszej tezy. W szczególności $C[a,b]$ z metryką d_ψ jest przestrzenią zupełną.

3.4 Lokalne twierdzenie Picarda

Jesteśmy już gotowi na dowód twierdzenia Picarda; użyjemy zasady Banacha. Zacznijmy od jego wersji lokalnej.

Twierdzenie 3.2 (Lokalne twierdzenie Picarda I)**.** *Niech będą dane* $a, b > 0$, $t_0 \in \mathbb{R}$ *i* $\alpha \in \mathbb{R}^k$ *oraz funkcja* f *o wartościach w* $\mathbb{R}^k$ *zdefiniowana na „prostokącie"*

$$P = \{(t,x) | \, |t - t_0| \leq a \ , \|x - \alpha\| \leq b\}.$$

Załóżmy, że funkcja f *jest ciągła na* P *i lipschitzowska względem drugiego argumentu ze stałą* L*: dla każdych* $(t,x), (t,y) \in P$,

$$\|f(t,x) - f(t,y)\| \leq L\|x - y\|.$$

Niech $M = \sup_{(t,x)\in P} \|f(t,x)\|$ *(*M *jest skończone, bo* f *jest funkcją ciągłą na zbiorze zwartym) i*

$$\kappa = \min\left(a, \frac{b}{M}, \frac{1}{2L}\right). \tag{3.13}$$

Istnieje wtedy dokładnie jedno rozwiązanie zagadnienia

$$u'(t) = f(t, u(t)), \qquad u(t_0) = \alpha, \tag{3.14}$$

okreslone w przedziale $|t - t_0| \leq \kappa$.

Dowód. Dowód przebiega w trzech krokach. Najpierw opiszemy przestrzeń, w której znajduje się rozwiązanie, a potem zlokalizujemy je samo jako punkt stały pewnego przekształcenia tej przestrzeni w siebie. Krokiem drugim, pośrednim, jest powiązanie równania z w.w. przekształceniem.

1. Jak wiemy z poprzedniego podrozdziału, przestrzeń $C[t_0 - \kappa, t_0 + \kappa]$ jest przestrzenią zupełną. Jej podzbiór X złożony z funkcji, które mają wartości w przedziale $[\alpha - b, \alpha + b]$ jest zbiorem domkniętym. Jest to więc także przestrzeń metryczna zupełna (zob twierdzenie 2.1).

2. Równanie (3.14) napiszmy w równoważnej *postaci całkowej*:

$$u(t) = \alpha + \int_{t_0}^{t} f(s, u(s))\, \mathrm{d}s. \tag{3.15}$$

Całka po prawej stronie jest całką z funkcji wektorowej i powinna być rozumiana jako wektor, którego składowymi są całki z jej współrzędnych. Zwróćmy uwagę na fakt, że wypisana wyżej postać całkowa w zwartej formie zawiera informację o warunku początkowym $u(t_0) = \alpha$. Jej główną zaletą jest jednak to, że pozwala sprowadzić zagadnienie istnienia rozwiązania równania różniczkowego do zagadnienia istnienia punktu stałego pewnego odwzorowania.

3. Rozważamy odwzorowanie $T : X \to X$ dane wzorem $x \mapsto Tx$,

$$(Tx)(t) = \alpha + \int_{t_0}^{t} f(s, x(s))\, \mathrm{d}s, \qquad t \in [t_0 - \kappa, t_0 + \kappa]. \tag{3.16}$$

Całka jest tu dobrze zdefiniowana, bo z założenia, że $x \in X$ wynika, iż $(s, x(s))$ należy do P. Zwróćmy uwagę, że dla $t \leq t_0$ całkę definiuje się tu jako $-\int_t^{t_0}$. T rzeczywiście przekształca X w siebie: po pierwsze Tx jest funkcją ciągłą, po drugie

$$\begin{aligned}\|(Tx)(t) - \alpha\| = \| \int_{t_0}^t f(s, x(s))\,\mathrm{d}s\| &\leq \left| \int_{t_0}^t \|f(s, x(s))\|\,\mathrm{d}s \right| \\ &\leq M|t - t_0| \leq M\kappa \leq b;\end{aligned}$$

w drugim kroku wykorzystaliśmy tu zadanie 3.11. Ponadto

$$\begin{aligned}d(Tx, Ty) &= \sup_{t\in[t_0-\kappa, t_0+\kappa]} \left\| \int_{t_0}^t [f(s, x(s)) - f(s, y(s)))\,\mathrm{d}s \right\| \\ &\leq \sup_{t\in[t_0-\kappa, t_0+\kappa]} \left| \int_{t_0}^t \|f(s, x(s)) - f(s, y(s)))\|\,\mathrm{d}s \right| \\ &\leq L \sup_{t\in[t_0-\kappa, t_0+\kappa]} \left| \int_{t_0}^t \|x(s) - y(s)\|\,\mathrm{d}s \right| \\ &\leq L \sup_{t\in[t_0-\kappa, t_0+\kappa]} \left| \int_{t_0}^t d(x, y)\,\mathrm{d}s \right| \\ &\leq Ld(x, y) \sup_{t\in[t_0-\kappa, t_0+\kappa]} |t - t_0| \\ &= L\kappa\, d(x, y) \\ &\leq \frac{1}{2} d(x, y).\end{aligned}$$

Rachunek ten dowodzi, że T spełnia założenia zasady Banacha. Istnieje więc dokładnie jedna funkcja u, która jest punktem stałym odwzorowania T, to znaczy dokładnie jedna taka funkcja, że równość (3.15) zachodzi dla wszystkich t z przedziału $[t_0 - \kappa, t_0 + \kappa]$. To oznacza jednak, że u jest rozwiązaniem zagadnienia (3.14). Z drugiej strony każde rozwiązanie zagadnienia (3.14) spełnia też równość (3.15). □

Dokładna analiza przedstawionego wyżej dowodu pokazuje, że założenia naszego twierdzenia można osłabić – w definicji κ zamiast $\frac{1}{2L}$ można napisać wziąć qL^{-1}, gdzie $q \in (0, 1)$ i rozumowanie będzie tak samo poprawne. W istocie w „prawdziwym" twierdzeniu Picarda stała Lipschitza w ogóle nie ma wpływu na κ, mamy bowiem

$$\kappa = \min\left(a, \frac{b}{M}\right). \tag{3.17}$$

Tego wyniku nasze rachunki jednak nie dowiodą. Do dowodu pełnego twierdzenia Picarda potrzebna nam jest inna, równoważna metryka w $C[a, b]$ (por. [12], str. 203, tw. Lipschitza).

Twierdzenie 3.3 (Lokalne twierdzenie Picarda II)**.** *Przy założeniach twierdzenia 3.2, istnieje dokładnie jedno rozwiązanie zagadnienia* (3.14)*, okreslone w przedziale* $|t - t_0| \leq \kappa$*, gdzie* κ *dane jest wzorem* (3.17)*.*

Dowód. Ustalmy $\lambda > 0$ i niech $\psi(t) = \mathrm{e}^{-\lambda|t-t_0|}, t \in [t_0 - \kappa, t_0 + \kappa]$. ψ jest funkcją ciągłą, przyjmującą wartości dodatnie. Jak wiemy z podrozdziału 3.3, metryka

$$d_\psi(x, y) = \sup_{t\in[t_0-\kappa,t_0+\kappa]} \mathrm{e}^{-\lambda|t-t_0|}\|x(t) - y(t)\|$$

jest zatem równoważna metryce supremum. W szczególności $C[a, b]$, a także jej podzbiór X opisany w dowodzie twierdzenia 3.2, z metryką d_ψ są przestrzeniami zupełnymi. Metrykę d_ψ nazywamy metryką Bieleckiego od nazwiska jej pomysłodawcy, Adama Bieleckiego, jednego z założycieli UMCS.

Podobnie jak w pierwszym dowodzie definiujemy przestrzeń X jako zbiór tych funkcji, które mają wartości w przedziale $[\alpha-b, \alpha+b]$, a odwzorowanie T – wzorem (3.16); wiemy już, że T przekształca X w X. Zauważmy ponadto, że

$$\begin{aligned}
d_\psi(Tx, Ty) &= \sup_{t\in[t_0-\kappa,t_0+\kappa]} \mathrm{e}^{-\lambda|t-t_0|} \left\| \int_{t_0}^{t} [f(s, x(s)) - f(s, y(s))] \,\mathrm{d}s \right\| \\
&\le \sup_{t\in[t_0-\kappa,t_0+\kappa]} \mathrm{e}^{-\lambda|t-t_0|} \left| \int_{t_0}^{t} \|f(s, x(s)) - f(s, y(s))\| \,\mathrm{d}s \right| \\
&\le L \sup_{t\in[t_0-\kappa,t_0+\kappa]} \left| \int_{t_0}^{t} \mathrm{e}^{-\lambda[|t-t_0|-|t_0-s|]} \{\mathrm{e}^{-\lambda|t_0-s|}\|x(s) - y(s)\|\} \,\mathrm{d}s \right|.
\end{aligned}$$

Zwróćmy uwagę na to, że – niezależnie od tego, czy t jest mniejsze, czy też większe niż t_0 – punkt s leży zawsze między nimi. Stąd $|t - t_0| - |t_0 - s| = |t - s|$. Ponadto wyrażenie w nawiasie klamrowym nie przekracza $d_\psi(x, y)$. Tak więc

$$\begin{aligned}
d_\psi(Tx, Ty) &\le L \sup_{t\in[t_0-\kappa,t_0+\kappa]} \left| \int_{t_0}^{t} \mathrm{e}^{-\lambda|t-s|} d_\psi(x, y) \,\mathrm{d}s \right| \\
&= L d_\psi(x, y) \sup_{t\in[t_0-\kappa,t_0+\kappa]} \int_0^{|t-t_0|} \mathrm{e}^{-\lambda u} \,\mathrm{d}u \\
&= L \int_0^{\kappa} \mathrm{e}^{-\lambda u} \,\mathrm{d}u \;\; d_\psi(x, y) < \frac{L}{\lambda} d_\psi(x, y).
\end{aligned}$$

Jeśli teraz wybierzemy $\lambda > L$, to przekonamy się, że T jest przy metryce d_ψ kontrakcją w X. Innymi słowy zasada Banacha jest w mocy i reszta rozumowania przebiega jak poprzednio. □

3.5 Globalne twierdzenie Picarda

Na zakończenie udowodnimy globalną wersję twierdzenia Picarda. Twierdzenie lokalne dotyczy rozwiązań w małym otoczeniu punktu, w którym zadano warunek początkowy. Twierdzenie globalne, przy nieco mocniejszych założeniach, zapewnia istnienie rozwiązań na całej półosi.

Twierdzenie 3.4 (Globalne twierdzenie Picarda)**.** *Załóżmy, że funkcja f jest ciągła i spełnia globalny warunek Lipschitza* (3.5) *oraz że istnieją takie stałe M_0 i ω_0, iż*

$$\left\| \int_0^t f(s,0)\,\mathrm{d}s \right\| \le M_0 \mathrm{e}^{\omega_0 t}. \tag{3.18}$$

Dla dowolnego $\alpha \in \mathbb{R}^k$, równanie (3.1) *ma wtedy dokładnie jedno rozwiązanie spełniające warunek $u(0) = \alpha$; rozwiązanie to zdefiniowane jest dla wszystkich $t \ge 0$ i rośnie najwyżej wykładniczo.*

Uwagi. Podobnie jak poprzednio, całka po lewej stronie nierówności (3.18) jest całką z funkcji o wartościach wektorowych i należy ją rozumieć jako wektor całek z jej składowych. Warunek typu (3.18) jest konieczny jeśli chcemy uzyskać wykładnicze ograniczenie na rozwiązania. Widać to na przykładzie funkcji $f(t,u) = 2t\mathrm{e}^{t^2}$, która (względem u) ma globalną stałą Lipschitza równą zero, a jednak rozwiązania związanego z nią równania różniczkowego (3.1) mają postać $u(t) = \alpha \mathrm{e}^{t^2}$ i w szczególności rosną szybciej niż wykładniczo. Warunek (3.18) jest automatycznie spełniony gdy funkcja f nie zależy od t (to znaczy dla równań autonomicznych).

Dowód globalnego twierdzenia Picarda. Dowód jest analogiczny do dowodu twierdzenia lokalnego. Nieco bardziej jest tylko skomplikowana przestrzeń, w której szukać będziemy rozwiązania.

(a) Dla danej liczby $\omega > 0$ niech $C_\omega(\mathbb{R}^+)$ oznacza przestrzeń takich funkcji ciągłych $x : [0,\infty) \to \mathbb{R}^k$, że istnieje $M = M(x)$ spełniające warunek

$$\sup_{t \ge 0} \mathrm{e}^{-\omega t} \|x(t)\| \le M.$$

Innymi słowy, elementy $C_\omega(\mathbb{R}^+)$ to funkcje ciągłe o co najwyżej wykładniczym wzroście (z wykładnikiem ω). Pokażemy teraz, że przestrzeń ta z odległością Bieleckiego

$$d_\omega(x,y) = \sup_{t \ge 0} \mathrm{e}^{-\omega t} \|x(t) - y(t)\|$$

jest zupełną przestrzenią metryczną. Sprawdzenie, że d_ω jest dobrze zdefiniowana i spełnia warunki metryki pozostawiamy jako ćwiczenie (zadanie 3.9), ograniczając się do trudniejszego dowodu, że mamy tu do czynienia z przestrzenią zupełną.

Niech $(x_n)_{n \ge 1}$ będzie ciągiem Cauchy'ego w $C_\omega(\mathbb{R}^+)$. Z nierówności

$$\sup_{t \in [0,T]} \|x_n(t) - x_m(t)\| \le \mathrm{e}^{\omega T} d_\omega(x_n, x_m), \tag{3.19}$$

która zachodzi dla dowolnego $T > 0$, wynika, iż dla każdego $t \ge 0$ istnieje granica $x(t) := \lim_{n\to\infty} x_n(t)$. Co więcej, funkcje x_n dążą do granicznej funkcji x jednostajnie na każdym przedziale $[0,T]$. Faktycznie, nierówność ta pokazuje, że funkcje $x_n, n \ge 0$ obcięte do przedziału $[0,T]$ tworzą w przestrzeni $C[0,T]$ ciąg podstawowy. W szczególności x jest ciągła na każdym przedziale postaci $[0,T]$ (zob. dowód zupełności przestrzeni $C[a,b]$), a więc ciągła na całym $\mathbb{R}^+$.

Przeprowadźmy to rozumowanie dokładniej: dla dowolnego $\epsilon > 0$ i przy ustalonym $T > 0$ możemy tak dobrać n_0, że dla $n, m \ge n_0$ zachodzi nierówność,

$d_\omega(x_n, x_m) < \epsilon e^{-\omega T}$. Z (3.19) wynika, że dla takich n i m mamy $\sup_{t\in[0,T]} \|x_n(t) - x_m(t)\| < \epsilon$. Skoro ϵ jest tu dowolne, dowodzi to, że ciąg obcięć funkcji x_n do przedziału $[0, T]$ jest ciągiem Cauchy'ego.

Musimy jeszcze pokazać, że x należy do $C_\omega(\mathbb{R}^+)$ i że $(x_n)_{n\geq 1}$ do niej dąży w sensie metryki d_ω. To, że $(x_n)_{n\geq 1}$ jest ciągiem podstawowym znaczy, iż dla każdego $\epsilon > 0$ istnieje takie n_0, że

$$\bigwedge_{t\geq 0} e^{-\omega t}\|x_n(t) - x_m(t)\| < \epsilon.$$

Przechodząc z m do nieskończoności, otrzymujemy stąd

$$\bigwedge_{t\geq 0} e^{-\omega t}\|x_n(t) - x(t)\| \leq \epsilon, \tag{3.20}$$

o ile tylko $n \geq n_0$. W szczególności, wybierając M tak, by $\|x_{n_0}(t)\| \leq M e^{\omega t}$ otrzymujemy oszacowanie $\|x(t)\| \leq \epsilon e^{\omega t} + M e^{\omega t} \leq (M + \epsilon)e^{\omega t}$, co znaczy, że x nie rośnie szybciej niż wykładniczo z wykładnikiem ω; innymi słowy, $x \in C_\omega(\mathbb{R}^+)$.

Nierówność (3.20) implikuje też, że $d_\omega(x_n, x) \leq \epsilon$ o ile $n \geq n_0$, a to (ze względu na dowolność ϵ) dowodzi zbieżności $(x_n)_{n\geq 1}$ do x.

(b) Niech ω będzie teraz większe niż stała Lipschitza z warunku (3.5) i większe niż ω_0 z warunku (3.18). W przestrzeni $C_\omega(\mathbb{R}^+)$ definiujemy odwzorowanie T, które funkcji x przyporządkowuje funkcję Tx daną wzorem:

$$(Tx)(t) = \alpha + \int_0^t f(s, x(s))\, ds. \tag{3.21}$$

Funkcja Tx jest oczywiście ciągła, ale na razie nie jest jasne, czy należy do $C_\omega(\mathbb{R}^+)$.

Rozważmy jednak dwie funkcje x i y z przestrzeni $C_\omega(\mathbb{R}^+)$ i dla ustalonego $t \geq 0$ oszacujmy wyrażenie

$$\begin{aligned} e^{-\omega t}\|(Tx)(t) - (Ty)(t)\| &= e^{-\omega t}\|\int_0^t f(s, x(s))\, ds - \int_0^t f(s, y(s))\, ds\| \\ &\leq e^{-\omega t}\int_0^t \|f(s, x(s)) - f(s, y(s))\|\, ds \\ &\leq e^{-\omega t}\int_0^t L\|x(s) - y(s)\|\, ds; \end{aligned}$$

w drugim kroku wykorzystaliśmy tu zadanie 3.11. Dalej wielkość tę szacować można tak:

$$\begin{aligned} e^{-\omega t}\|(Tx)(t) - (Ty)(t)\| &\leq \int_0^t L e^{-\omega(t-s)} e^{-\omega s}\|x(s) - y(s)\|\, ds \\ &\leq \int_0^t L e^{-\omega s}\, ds\, d_\omega(x, y) \\ &< \frac{L}{\omega} d_\omega(x, y). \end{aligned} \tag{3.22}$$

Jeśli w szczególności $y = 0$, to z założenia (3.18) wynika, iż Ty należy do $C_\omega(\mathbb{R}^+)$, a stąd i z powyższej nierówności można już wywnioskować, że do $C_\omega(\mathbb{R}^+)$ należy również Tx. Nierówność ta dowodzi także, że odwzorowanie T spełnia założenia zasady Banacha ze stałą $q = \frac{L}{\omega} < 1$. A zatem istnieje dokładnie jeden element przestrzeni $C_\omega(\mathbb{R}^+)$, nazwijmy go u, który jest punktem stałym odwzorowania T: on właśnie jest szukanym rozwiązaniem równania różniczkowego. Rzeczywiście, u spełnia warunek (3.15). I odwrotnie, rozwiązanie równania (3.1) też musi spełniać warunek (3.15), a to znaczy, że musi być punktem stałym odwzorowania T. □

Pozostało nam jeszcze tylko rozwiać jedną wątpliwość: czy mogą istnieć rozwiązania równania (3.1), które lokalnie będą się różnić od znalezionego wyżej? Odpowiedź negatywna wynika z lokalnego twierdzenia Picarda.

3.6 Zadania

Zadanie 3.1. Jak nazywał się słynny matematyk, który dostał od swej wybranki kosza, bo w peanach o jej urodzie nie chciał użyć słowa „jedyna”?

Zadanie 3.2. Sprawdź, że podzbiór przestrzeni $C[t_0 - \alpha, t_0 + \alpha]$, złożony z funkcji, które mają wartości w przedziale $[\alpha - b, \alpha + b]$ jest zbiorem domkniętym.

Zadanie 3.3. Dowiedź, że nie istnieje stała L spełniająca warunek (3.6).

Zadanie 3.4. Pokaż, że funkcja $f(u) = 1 + u^2$ jest lokalnie, ale nie globalnie lipschitzowska.

Zadanie 3.5. Pokaż, że jeśli normy wektorów (x, y) i $(\tilde{x}, \tilde{y})$ nie przekraczają r, to mamy

$$\begin{aligned}\|(ax - bxy, -cy + dxy) - (a\tilde{x} - b\tilde{x}\tilde{y}, -c\tilde{y} + d\tilde{x}\tilde{y})\|_{\max} \\ \leq \max\{a + br, c + dr\}\|(x, y) - (\tilde{x}, \tilde{y})\|_{\max}.\end{aligned}$$

Zadanie 3.6. Znajdź takie stałe c i C, by dla każdych $x, y \in \mathbb{R}$ zachodziła nierówność:

$$c \max\{|x|, |y|\} \leq \sqrt{x^2 + y^2} \leq C \max\{|x|, |y|\}.$$

To znaczy, że normy euklidesowa i maksimum w $\mathbb{R}^2$ są równoważne.

Zadanie 3.7. Sprawdź, że wzór (3.8) definiuje metrykę w przestrzeni funkcji ciągłych.

Zadanie 3.8. Dowiedź, że supremum w (3.12) jest skończone i że d_ψ jest metryką w $C[a, b]$.

Zadanie 3.9. Sprawdź, że d_ω jest dobrze zdefiniowana (w szczególności mamy $d_\omega(x, y) < \infty$ dla wszystkich $x, y \in C_\omega(\mathbb{R}^+)$) w przestrzeni $C_\omega(\mathbb{R}^+)$ i jest metryką.

Zadanie 3.10. Udowodnij nierówność (3.19).

Zadanie 3.11. Niech $x : [a,b] \to \mathbb{R}^k$ będzie funkcją ciągłą. Udowodnij, że

$$\left\| \int_a^b x(t)\,\mathrm{d}t \right\| \leq \int_a^b \|x(t)\|\,\mathrm{d}t.$$

Zadanie 3.12. Niech g i ϕ będą funkcjami ciągłymi na $[0,1]$. Udowodnij, że istnieje dokładnie jedna funkcja $f \in C[0,1]$ spełniająca równanie:

$$f(x) - \int_0^x \phi(x-y)f(y)\,\mathrm{d}y = g(x), \qquad x \in [0,1].$$

Wskazówka: rozważ $C[0,1]$ z odległością Bieleckiego

$$d_\lambda(f_1, f_2) = \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x}|f_1(x) - f_2(x)|$$

oraz odwzorowanie T w tej przestrzeni, dane wzorem

$$(Tf)(x) = g(x) + \int_0^x \phi(x-y)f(y)\,\mathrm{d}y, \qquad x \in [0,1].$$

Rozdział 4

Przestrzenie Banacha

4.1 Definicja przestrzeni Banacha

Przestrzeń Banacha, centralny obiekt całej analizy funkcjonalnej, to z definicji przestrzeń liniowa, unormowana i zupełna. W rozdziale niniejszym będziemy stopniowo wprowadzać i systematycznie tłumaczyć występujące w podanej wyżej definicji pojęcia.

4.1.1 Przestrzeń liniowa

Niech $\mathbb{X}$ będzie zbiorem; jego elementy oznaczać będziemy x, y, z, itd. i nazywać wektorami. Trójkę $(\mathbb{X}, +, \cdot)$, w której $+$ jest odzorowaniem $+ : \mathbb{X} \times \mathbb{X} \to \mathbb{X}, (x, y) \mapsto x + y$ a $\cdot$ odwzorowaniem $\cdot : \mathbb{R} \times \mathbb{X} \to \mathbb{X}, (\alpha, x) \mapsto \alpha x$, nazywamy (rzeczywistą) przestrzenią liniową (lub – zamiennie – przestrzenią wektorową) jeśli spełnione są następujące warunki

(a1) dla dowolnych $x, y, z \in \mathbb{X}$ zachodzi równość $(x + y) + z = x + (y + z)$,

(a2) istnieje takie $\Theta \in \mathbb{X}$, że $x + \Theta = x$ dla każdego $x \in \mathbb{X}$,

(a3) dla każdego $x \in \mathbb{X}$ istnieje takie $x' \in \mathbb{X}$, że $x + x' = \Theta$,

(a4) $x + y = y + x$, dla wszystkich $x, y \in \mathbb{X}$,

(m1) $\alpha(\beta x) = (\alpha\beta)x$, dla wszystkich $\alpha, \beta \in \mathbb{R}, x \in \mathbb{X}$,

(m2) $1x = x$, dla wszystkich $x \in \mathbb{X}$,

(d) $\alpha(x + y) = \alpha x + \alpha y$ oraz$(\alpha + \beta)x = \alpha x + \beta x$ dla każdych liczb $\alpha, \beta \in \mathbb{R}$ i wektorów $x, y \in \mathbb{X}$.

Warunki (a1)–(a4) znaczą, że $(\mathbb{X}, +)$ jest grupą przemienną. Choć z definicji przestrzeń liniowa to trójka uporządkowana, bardzo często, jeśli nie będzie to prowadzić do nieporozumień, mówić będziemy, że samo $\mathbb{X}$ jest przestrzenią liniową.

Wyróżniony element $\Theta \in \mathbb{X}$ wyznaczony jest jednoznacznie. Nazywa się go wektorem zerowym, lub po prostu zerem i oznacza 0 (choć trzeba go oczywiście odróżnić od liczby zero) – okazuje się, że dla każdego x, wektor $0x$ jest wektorem zerowym. Wektor x' nazywamy wektorem przeciwnym do x; okazuje się, że jest on równy $(-1)x$. Stąd przyjęło się pisać $x' = -x$.

Podstawowym przykładem przestrzeni liniowej jest przestrzeń wszystkich funkcji zdefiniowanych na danym zbiorze S. Dodawanie funkcji i mnożenie przez skalar wprowadza się tu standardowymi wzorami:

$$(x+y)(p) = x(p) + y(p), \qquad (\alpha x)(p) = \alpha x(p),$$

w których x i y są liczbami, a $p \in S$. Szczególnym przypadkiem takiej przestrzeni liniowej jest przestrzeń macierzy $n \times m$, które utożsamiać można z funkcjami na iloczynie karezjańskim $\{1, \ldots, n\} \times \{1, \ldots, m\}$. Podobnie przestrzeń ciągów to po prostu przestrzeń funkcji zdefiniowanym na zbiorze liczb naturalnych. Zerem w tej przestrzeni jest funkcja tożsamościowo równa zero, a wektorem przeciwnym do funkcji x jest x' dana wzorem $x'(p) = -x(p), p \in S$.

Podprzestrzeń (w sensie algebraicznym) przestrzeni liniowej $\mathbb{X}$ to jej podzbiór $\mathbb{Y}$ o tej własności, że warunki $x, y \in \mathbb{Y}, \alpha \in \mathbb{R}$ pociągają za sobą $x + y \in \mathbb{Y}$ i $\alpha x \in \mathbb{Y}$. Przykładem podprzestrzeni funkcji zdefiniowanych na zbiorze S jest podprzestrzeń tych funkcji, które są ograniczone: dla każdej z nich (oddzielnie) istnieje stała, która jest większa od wszystkich jej wartości. Stwierdzenie powyższe jest tylko innym sformułowaniem faktów, że suma dwóch funkcji ograniczonych jest ograniczona i że iloczyn liczby i funkcji ograniczonej jest ograniczony. Jeśli S jest przestrzenią metryczną (lub ogólniej: topologiczną), to podprzestrzenią wszystkich funkcji na S jest też przestrzeń funkcji ciągłych – suma dwóch funkcji ciągłych, oraz iloczyn funkcji ciągłej przez liczbę nadal jest funkcją ciągłą.

Podprzestrzeń liniowa z działaniami odziedziczonymi ze swej „nadprzestrzeni" sama w sobie jest przestrzenią liniową.

4.1.2 Przestrzeń unormowana

Przestrzeń liniową $\mathbb{X}$ nazywamy unormowaną jeśli zdefiniowana jest w niej funkcja

$$\| \cdot \| : \mathbb{X} \to \mathbb{R},$$

interpretowana jako długość wektora, która spełnia następujące warunki. Dla dowolnych $x, y \in \mathbb{X}$ oraz $\alpha \in \mathbb{R}$,

(n1) $\|x\| \geq 0$,

(n2) $\|x\| = 0$ wtedy i tylko wtedy, gdy $x = \Theta$,

(n3) $\|\alpha x\| = |\alpha| \, \|x\|$,

(n4) $\|x + y\| \leq \|x\| + \|y\|$.

Podobnie jak to było w przypadku definicji metryki, czwarty warunek nazywany jest warunkiem trójkąta. Można pokazać (zadanie 4.1), że z (n3-n4) wynika, iż dla dowolnych $x, y \in \mathbb{X}$, zachodzi nierówność

$$\big|\|x\| - \|y\|\big| \le \|x \pm y\|. \tag{4.1}$$

Zwróćmy uwagę, że jeśli $\|\cdot\|$ jest normą, to $d(x, y) = \|x - y\|$ jest metryką (zob. zadanie 4.2). Innymi słowy, każda przestrzeń unormowana jest przestrzenią metryczną. Możemy więc w niej używać pojęć topologicznych takich, jak zbieżność, zbiór otwarty, domknięty itd. Struktura przestrzeni liniowej jest jednak oczywiście znacznie bogatsza niż przestrzeni topologicznej (ze względu na związki ze strukturą liniową i szczególną postać metryki – by wymienić powody najbardziej oczywiste). To właśnie bogactwo tej struktury jest powodem, dla którego tak wiele ciekawych rzeczy „dzieje się" w przestrzeniach unormowanych, a w szczególności w przestrzeniach Banacha.

4.2 Przykłady przestrzeni unormowanych

Przyjrzyjmy się z bliska klasycznym przykładom przestrzeni unormowanych.

4.2.1 Przestrzeń c ciągów zbieżnych

Jest to przestrzeń ciągów liczbowych $x = (\xi_i)_{i\ge 1}$, które mają granicę w nieskończoności: istnieje $\lim_{i\to\infty} \xi_i$. Dość jasne jest, że jest to przestrzeń liniowa. Jeśli ciągi $x = (\xi_i)_{i\ge 1}$ i $y = (\eta_i)_{i\ge 1}$ mają granice, to dla dowolnych α i β ma ją także ciąg $\alpha x + \beta y$. Każdy element $x \in c$ jest też ciągiem ograniczonym: jeśli g jest jego granicą, to prawie wszystkie wyrazy ciągu leżą w przedziale $(g - 1, g + 1)$. Stąd

$$\sup_{i\in N} |\xi_i| \le \max\{M, |g| + 1\} < \infty,$$

gdzie M jest wartością bezwzględną największego z wyrazów, które leżą poza przedziałem $(g - 1, g + 1)$. Każdemu $x \in c$ możemy więc przyporządkować

$$\|x\| = \sup_{i\in N} |\xi_i| < \infty.$$

Jak łatwo sprawdzić, funkcja $\|\cdot\|$ jest normą.

4.2.2 Przestrzeń c_0 ciągów zbieżnych do zera

Jeśli dwa ciągi, powiedzmy $x = (\xi_i)_{i\ge 1}$ oraz $y = (\eta_i)_{i\ge 1}$ dążą do zera, to do zera dąży też, dla dowolnych α i β, ciąg $\alpha x+\beta y$. Dowodzi to, że – w sensie algebraicznym – zbiór ciągów zbieżnych jest podprzestrzenią c. Podprzestrzeń tę oznacza się c_0. Z normą oddziedziczoną z c, c_0 jest przestrzenią liniową unormowaną.

c_0 jest też podzbiorem domkniętym zbioru c. Załóżmy bowiem, że $x_n \in c_0$ dąży do $x = (\xi_i)_{i\ge 1} \in c$ i wybierzmy dowolne $\epsilon > 0$. Niech N będzie tak dużą liczbą naturalną, że dla $n \geq N$, zachodzi nierówność $\|x_n - x\| < \frac{\epsilon}{2}$. Ciąg $x_N = (\xi_{N,i})_{i\geq 1}$ jest (z założenia) ciągiem zbieżnym do zera, możemy więc tak dobrać m, by dla każdego $i \geq m$, $|\xi_{N,i}| < \frac{\epsilon}{2}$. Z drugiej strony dla każdego i zachodzi nierówność $|\xi_{N,i} - \xi_i| < \frac{\epsilon}{2}$. Dowodzi to, że dla $i \geq m$ mamy $|\xi_i| < \epsilon$. Innymi słowy, prawie wszystkie wyrazy ciągu $(\xi_i)_{i\geq 1}$ są mniejsze niż ϵ: z dowolności ϵ wnioskujemy, że x zbiega do 0. Tak więc c_0 jest podprzestrzenią c w innym, mocniejszym sensie: jest „zamknięta" nie tylko algebraicznie, ale i topologicznie. Nie tylko dodawanie i mnożenie przez skalar nie wyprowadza z c_0 – nie wyjdziemy z niej także, przechodząc do granicy. Informacja ta będzie bardzo ważna później.

4.2.3 Przestrzeń $C[a,b]$ funkcji ciągłych na odcinku $[a,b]$

Jak to wspominaliśmy wyżej, zbiór wszystkich rzeczywistych funkcji ciągłych na odcinku $[a,b]$ (gdzie oczywiście $a < b$) jest podprzestrzenią liniową wszystkich funkcji tamże zdefiniowanych. Oznaczamy ją $C[a,b]$ (to nie powinno prowadzić do nieporozumień – przestrzeń rozważana tu jest szczególnym przypadkiem przestrzeni z podrozdziału 3.2). Z podrozdziału 3.2 wiemy też, że jest to przestrzeń metryczna. Okazuje się, że metryka supremum pochodzi od normy, także zwanej normą supremum. By wyjaśnić to dokładniej przypomnijmy znane twierdzenie analizy rzeczywistej mówiące, iż każda funkcja ciągła przyjmuje na zbiorze zwartym (a takim jest odcinek) swoje – skończone – kresy. Jeśli więc $x \in C[a,b]$, to istnieje takie $t_0 \in [a,b]$, że

$$\max_{t\in[a,b]} |x(t)| = |x(t_0)|.$$

Jak łatwo sprawdzić funkcja

$$x \mapsto \|x\| := \max_{t\in[a,b]} |x(t)|,$$

jest normą. Co więcej dla metryki supremum d, zachodzi wzór

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

Przestrzeń $C[a,b]$ jest więc przestrzenią liniową unormowaną. Z podrozdziału 3.2 wiemy ponadto, że jest to przestrzeń zupełna. Poznaliśmy więc tym samym pierwszy przykład przestrzeni Banacha – zob. podrozdział 4.3.

4.2.4 Przestrzeń l^1 szeregów absolutnie sumowalnych

Mówimy, że szereg $(\xi_i)_{i\geq 1}$ jest absolutnie sumowalny, jeśli suma

$$\|(\xi_i)_{i\geq 1}\| = \sum_{i=1}^{\infty} |\xi_i|$$

jest skończona. Oznaczenie $\|\cdot\|$ wprowadzone wyżej nie jest przypadkowe – $\|\cdot\|$ rzeczywiście jest normą w przestrzeni l^1, ciągów absolutnie zbieżnych. Oczywiście najpierw wypadałoby uzasadnić, że l^1 jest przestrzenią liniową – to jednak nie ulega wątpliwości, bo kombinacja liniowa dwóch szeregów absolutnie sumowalnych jest szeregiem absolutnie sumowalnym.

4.2.5 Przestrzeń $L^1(\mathbb{R})$ funkcji całkowalnych

Ciągłym odpowiednikiem przestrzeni l^1 jest przestrzeń $L^1(\mathbb{R})$ funkcji absolutnie całkowalnych w sensie Lebesgue'a na osi $\mathbb{R}$; normą jest tu

$$\|f\| = \int_{\mathbb{R}} |f|.$$

Ściśle rzecz biorąc, $L^1(\mathbb{R})$ nie składa się z funkcji, lecz z ich klas równoważności. Chodzi o to, że z równości $\|f\| = 0$ *nie wynika* $f \equiv 0$, lecz jedynie to, że f równa jest 0 prawie wszędzie (to znaczy poza zbiorem miary zero). Jeśli umówimy się, że dwie funkcje absolutnie całkowalne różniące się na zbiorze miary zero są w relacji, to otrzymamy relację równoważności. (Dla przykładu, jeśli $f(x)$ równa się 0 dla x niewymiernych, a 1 na zbiorze liczb wymiernych, to f jest w relacji z funkcją tożsamościowo równą zero). Co więcej, przypisując każdej klasie równoważności całkę z bezwzględnej wartości jej reprezentanta, przekonamy się, że nie zależy ona od w.w. reprezentanta i że definiuje normę na liniowej przestrzeni klas (sumą dwóch klas jest klasa odpowiadająca sumie dowolnych ich reprezentantów, podobnie z mnożeniem przez skalar). Wektorem zerowym w tej przestrzeni jest klasa równoważności funkcji będących w relacji z funkcją tożsamościowo równej zero – to znaczy funkcji równych zero niemal wszędzie.

4.3 Przykłady przestrzeni Banacha

4.3.1 Przestrzeń c ciągów zbieżnych

Przestrzeń c jest przestrzenią zupełną. Załóżmy bowiem, że $(x_n)_{n\geq 1}$ jest ciągiem Cauchy'ego w c i niech $x_n = (\xi_{n,i})_{i\geq 1}$. Dla dowolnego $\epsilon > 0$ istnieje zatem takie $N \in \mathbb{N}$, że dla $n, m \geq N$ mamy

$$\sup_{i\geq 1} |\xi_{n,i} - \xi_{m,i}| = \|x_n - x_m\| < \epsilon.$$

Ustalmy na chwilę $i \in \mathbb{N}$. Oczywista nierówność

$$|\xi_{n,i} - \xi_{m,i}| \leq \|x_n - x_m\|,$$

przekonuje nas, w połączeniu z tym co napisaliśmy wyżej, że ciąg liczbowy $(\xi_{n,i})_{n\geq 1}$ spełnia warunek Cauchy'ego, a zatem ma granicę (bo $\mathbb{R}$ jest przestrzenią zupełną). Niech granicą tą będzie ξ_i. Ciąg $x = (\xi_i)_{i\geq 1}$ złożony z tak otrzymanych granic jest naturalnym kandydatem na granicę ciągu $(x_n)_{n\geq 1}$.

Pozostaje nam zatem pokazać, że $x \in c$ i że $\lim_{n\to\infty} x_n = x$. Zwróćmy uwagę na to, że dla dowolnego $\epsilon > 0$ możemy tak dobrać N, iż

$$\text{dla każdego } i \in \mathbb{N} \qquad |\xi_{n,i} - \xi_{m,i}| < \frac{\epsilon}{2}, \tag{4.2}$$

o ile tylko $n, m \geq N$. Przechodząc z i do nieskończoności, przekonujemy się, że dla $n, m \geq N$,

$$|\xi_n - \xi_m| \leq \frac{\epsilon}{2} < \epsilon,$$

co dowodzi, iż $(\xi_i)_{i\geq 1}$ spełnia warunek Cauchy'ego – a zatem jest zbieżny, należy do c. Przechodząc zaś w (4.2) z m do nieskończoności (dla każdego i oddzielnie), otrzymujemy

$$\text{dla każdego } i \in \mathbb{N} \qquad |\xi_{n,i} - \xi_i| \leq \frac{\epsilon}{2},$$

(proszę zauważyć „osłabienie" nierówności w granicy), co pociąga za sobą

$$\|x_n - x\| \leq \frac{\epsilon}{2} < \epsilon.$$

To jednak oznacza, że granicą $(x_n)_{n\geq 1}$ jest x.

4.3.2 Przestrzeń c_0 ciągów zbieżnych do zera

Jak już widzieliśmy, c_0 jest podprzestrzenią c w sensie algebraicznym i dziedziczy po niej normę. Udowodniliśmy też poprzednio, że jest jej podzbiorem domkniętym. Twierdzenie 2.1 przekonuje nas więc, że c_0 jest sama w sobie zupełna. A to dowodzi, że jest przestrzenią Banacha.

4.3.3 Przestrzeń $C[a, b]$ funkcji ciągłych na odcinku $[a, b]$

Z podrozdziału 3.2 wiemy, że $C[a, b]$ z metryką supremum jest przestrzenią metryczną zupełną, a w podrozdziale 4.2.3 przekonaliśmy się, ze jest też przestrzenią liniową, unormowaną. Dodatkowo metryka supremum pochodzi od naturalnej normy w $C[a, b]$. To wszystko razem dowodzi, że $C[a, b]$ jest przestrzenią Banacha.

4.3.4 Przestrzeń l^1 szeregów absolutnie sumowalnych

To, że przestrzeń l^1 jest liniowa i unormowana zostało ustalone w podrozdziale 4.2.4. By dowieść jej zupełności, załóżmy teraz, że $(x_n)_{n\geq 1}$ jest ciągiem podstawowym w c i niech $x_n = (\xi_{n,i})_{i\geq 1}$. Dla każdego ϵ możemy zatem znaleźć takie N, że

$$\sum_{i=1}^{\infty} |\xi_{n,i} - \xi_{m,i}| < \epsilon,$$

o ile $n, m \geq N$. Podobnie jak poprzednio posługujemy się teraz nierównością

$$|\xi_{n,i} - \xi_{m,i}| \leq \|x_n - x_m\|,$$

by przekonać się, że dla każdego naturalnego i ciąg $(\xi_{n,i})_{n\geq 1}$ jest zbieżny, bo jest ciągiem Cauchy'ego w przestrzeni zupełnej.[7]

---

[7] Choć nierówność powyższa jest pozornie powtórzeniem nierówności z podrozdziału 4.3.1, mamy tu do czynienia z zupełnie innymi zależnościami. Inaczej mówiąc, po prawej stronie tych nierówności występują inne normy. Inne też więc są argumenty, których trzeba użyć w dowodzie.

I znów $x = (\xi_i)_{i\geq 1}$, gdzie $\xi_i = \lim_{n\to\infty} \xi_{i,n}$ jest naturalnym kandydatem na granicę ciągu $(x_n)_{n\geq 1}$. Z założenia wiemy, że dla dowolnego ϵ można tak dobrać naturalne $N = N(\epsilon)$, iż dla $n, m \geq N$ zajdzie nierówność

$$\sum_{i=1}^{\infty} |\xi_{n,i} - \xi_{m,i}| < \frac{\epsilon}{2}.$$

Dowodzi to, że dla każdego naturalnego k mamy $\sum_{i=1}^{k} |\xi_{n,i} - \xi_{m,i}| < \frac{\epsilon}{2}$. Przechodząc z m do nieskończoności, wnioskujemy, że dla każdego k zachodzi nierówność $\sum_{i=1}^{k} |\xi_{n,i} - \xi_i| \leq \frac{\epsilon}{2}$, a zatem

$$\sum_{i=1}^{\infty} |\xi_{n,i} - \xi_i| \leq \frac{\epsilon}{2} < \epsilon. \tag{4.3}$$

Przyjmując $\epsilon = 1$ i wybierając $N = N(1)$, przekonujemy się, że

$$\sum_{i=1}^{\infty} |\xi_i| \leq \sum_{i=1}^{\infty} (|\xi_i - \xi_{N,i}| + |\xi_{N,i}|) < 1 + \|x_N\| < \infty.$$

To znaczy jednak, że $x \in l^1$. Nierówność (4.3) świadczy teraz o tym, że $\|x_n - x\| < \epsilon$ dla odpowiednio dużych n. A to znaczy po prostu, że x jest granicą $(x_n)_{n\geq 1}$.

4.3.5 Przestrzeń $L^1(\mathbb{R})$ funkcji całkowalnych

Dowód zupełności przestrzeni (klas) funkcji całkowalnych jest analogiczny do dowodu zupełności przestrzeni l^1, ale wymaga znajomości teorii miary. Pozwolimy więc sobie go opuścić – można go znaleźć w wielu podręcznikach, na przykład w [4].

4.4 Przestrzenie unormowane, które nie są przestrzeniami Banacha

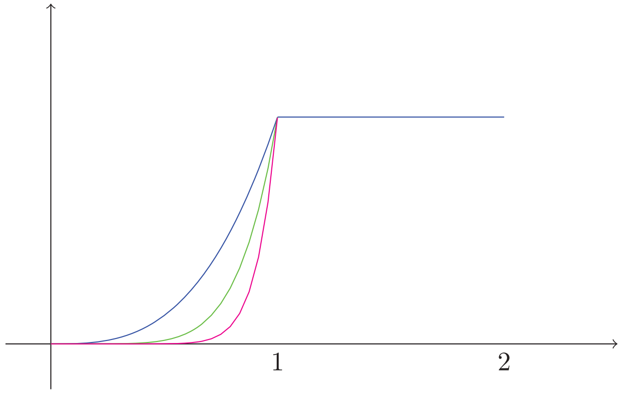
4.4.1 $C[0,2]$ z „niewłaściwą" normą

Przykładem przestrzeni unormowanej, która nie jest zupełna jest $C[0,2]$ wyposażona w normę

$$\|x\| = \int_0^2 |x(s)|\,\mathrm{d}s.$$

To, że $C[0,2]$ jest przestrzenią liniową już wiemy; wiemy też, że $\|\cdot\|$ jest normą. Pokażemy, że $C[0,2]$ z taką normą nie jest przestrzenią Banacha. Oczywiście wystarczy znaleźć ciąg elementów tej przestrzeni, który nie jest zbieżny, a jednak spełnia warunek Cauchy'ego. W tym celu rozważamy funkcje $x_n(s) = \min\{s^n, 1\}, s \in [0,2]$, przedstawione na rysunku 4.1. Skoro

$$\|x_n - x_m\| = \int_0^1 (s^n - s^m)\,\mathrm{d}s = \frac{1}{n+1} - \frac{1}{m+1},$$



Rys. 4.1: Wykresy funkcji x_n dla $n = 3$ (kolor niebieski), $n = 6$ (kolor zielony) o $n = 11$ (czerwony); w przedziale $[1, 2]$ wykresy te są takie same.

o ile tylko $m \geq n$, jasne jest, że $(x_n)_{n\geq 1}$ jest ciągiem podstawowym. Nie jest to jednak ciąg zbieżny. Wyobraźmy sobie bowiem jak wyglądać musiałaby jego granica x. Dla $s > 1$, x musiałoby być równe 1: w przeciwnym wypadku mielibyśmy $\|x_n - x\| \geq \int_1^2 |1 - x(s)| \,\mathrm{d}s > 0$ dla każdego $n \geq 1$. Podobnie musimy mieć $x(s) = 0$ dla $s < 1$, bo w przeciwnym razie otrzymalibyśmy $\lim_{n\to\infty} \|x_n - x\| \geq \lim_{n\to\infty} \int_0^1 |x(s) - s^n| \,\mathrm{d}s = \int_0^1 |x(s)| \,\mathrm{d}s > 0$. Nie ma jednak funkcji ciągłej x, która spełniałaby równocześnie warunki: $x(s) = 0, s < 1$ i $x(s) = 1, s > 1$.

4.4.2 Przestrzeń wielomianów

Niech $\mathbb{X}$ będzie przestrzenią wielomianów wszystkich stopni zdefiniowanych na przedziale $[0, 1]$, to znaczy funkcji postaci

$$x(t) = a_0 + a_1 t + \cdots + a_n t^n,$$

gdzie n jest liczbą naturalną (zależną od x), a $a_i, i = 0, \ldots, n$ są liczbami rzeczywistymi. Łatwo sprawdza się, że $\mathbb{X}$ jest przestrzenią liniową. Funkcja

$$x \mapsto \|x\| = \sup_{t\in[0,1]} |x(s)|$$

jest dla niej normą. Okazuje się jednak, że przestrzeń ta nie jest przestrzenią zupełną.

By się o tym przekonać, rozważmy ciąg $(x_n)_{n\geq 1}$ dany wzorem

$$x_n(t) = 1 + t + \frac{t^2}{2} + \cdots + \frac{t^n}{n!}, \qquad t \in [0, 1], n \geq 1.$$

Oczywiście wszystkie x_n są wielomianami, elementami $\mathbb{X}$. Wobec wzoru (4.4) mamy też dla dowolnych m, n całkowitych

$$\begin{aligned}\|x_m - x_n\| &\le \sup_{t\in[0,1]} \int_0^t \frac{(t-s)^n}{n!} \mathrm{e}^s \,\mathrm{d}s + \sup_{t\in[0,1]} \int_0^t \frac{(t-s)^m}{m!} \mathrm{e}^s \,\mathrm{d}s \\ &\le \mathrm{e}^1 \int_0^1 \frac{s^n}{n!} \,\mathrm{d}s + \mathrm{e}^1 \int_0^1 \frac{s^m}{m!} \,\mathrm{d}s \le \mathrm{e}^1 \left(\frac{1}{(n+1)!} + \frac{1}{(m+1)!} \right).\end{aligned}$$

Stąd łatwo już wnioskujemy, że ciąg $(x_n)_{n\ge1}$ spełnia warunek Cauchy'ego.

Jak przekonać się, że nie jest on jednak zbieżny? Powodem, oczywistym dla tych, którzy pamiętają definicję funkcji wykładniczej, jest to, że jego granicą punktową (a nawet jednostajną – zob. zadanie 4.6) jest $x(t) = \mathrm{e}^t$. Gdyby zatem istniała granica jednostajna, to musiałaby być równa opisanej wyżej eksponencie. Wystarczy zatem dowieść, że x nie należy do naszej przestrzeni, to znaczy, że nie jest wielomianem. W tym celu zauważmy, że pochodną funkcji x jest ona sama, a w konsekwencji wszystkie jej pochodne są równe jej samej; w szczególności są różne od zera. Gdyby natomiast x była wielomianem, to jej pochodna odpowiednio dużego stopnia byłaby równa zero.

4.5 Zadania

Zadanie 4.1. Dowiedź (4.1).

Zadanie 4.2. Pokaż, że każda przestrzeń unormowana jest przestrzenią metryczną.

Zadanie 4.3. Sprawdź, że wszystkie normy wprowadzone w podrozdziale 4.2 są nimi rzeczywiście.

Zadanie 4.4. Niech będzie dane $r > 0$ i niech l^1_r oznacza zbiór ciągów $(\xi_i)_{i\ge1}$, które mają tę własność, że

$$\|x\|_r = \sum_{i\ge1} |\xi_i| r^i < \infty.$$

Udowodnij, że l^1_r jest przestrzenią liniową, że $\|\cdot\|_r$ jest normą i że z tak dobraną normą przestrzeń ta jest przestrzenią Banacha.

Zadanie 4.5. Udowodnij, że $C[0,1]$ z normą $\|x(\cdot)\| = \int_0^1 |x(s)| \,\mathrm{d}s$ nie jest przestrzenią Banacha, rozważając ciąg $(x_n)_{n\ge1}$ dany wzorem

$$x_n(s) = \begin{cases} 0, & 0 \le s \le \frac{1}{2} - \frac{1}{n}, \\ \frac{1}{2} + \frac{n}{2}(s - \frac{1}{2}), & \frac{1}{2} \le s \le \frac{1}{2} + \frac{1}{n}, \\ 1, & \frac{1}{2} + \frac{1}{n} \le s \le 1. \end{cases}$$

Zadanie 4.6. Udowodnij za pomocą indukcji, że dla każdego n zachodzi równość

$$\mathrm{e}^t = \sum_{k=0}^{n} \frac{t^k}{k!} + \int_0^t \frac{(t-s)^n}{n!} \mathrm{e}^s \,\mathrm{d}s. \tag{4.4}$$

Wywnioskuj stąd, że ciąg $(x_n)_{n\geq 1}$ z podrozdziału 4.4.2 dąży do funkcji wykładniczej w przestrzeni $C[0,1]$ (to znaczy w normie supremum).

Zadanie 4.7. Sprawdź, że zbiór c_{00} nieskończonych ciągów $x = (\xi_i)_{i\geq 1}$, których prawie wszystkie wyrazy są zerami jest przestrzenią liniową unormowaną z normą

$$\|x\| = \sup_{i\geq 1} |\xi_i|.$$

Dowiedź, że nie jest to jednak przestrzeń zupełna. (Nie zapomnij dowieść, że $\|\cdot\|$ przyjmuje wartości skończone!)

Zadanie 4.8. Powtórz zadanie poprzednie, zmieniając poprzednio rozważaną normę na

$$\|x\| = \sum_{i}^{\infty} |\xi_i|.$$

Rozdział 5

Równanie odnowienia w modelu McKendricka–von Foerstera

Naszym celem w tym rozdziale jest rozwiązanie równania odnowienia – klucza do modelu rozwoju populacji, który pochodzi od McKendricka i von Foerstera. A głównym i jedynym powodem istnienia i jedyności jego rozwiązań, poza samą jego postacią, jest zupełność przestrzeni $L^1(\mathbb{R}^+)$. Dowód będzie prostym zastosowaniem zasady Banacha, ale musimy zacząć „od pieca", to znaczy przedstawić przestrzeń, w której będziemy operować – tak jak dobry, choć może trochę staroświecki pisarz, czy wręcz gawędziarz (by wspomnieć choćby J.F. Coopera, czy też Tytusa Karpowicza), nim zacznie swą opowieść, opisuje teren, na którym rozgrywać się będzie akcja.

5.1 Algebry Banacha i algebra splotowa $L^1(\mathbb{R}^+)$

Przestrzeń $L^1(\mathbb{R}^+)$ (klas) funkcji absolutnie całkowalnych na $\mathbb{R}^+$ jest, podobnie jak $L^1(\mathbb{R})$, przestrzenią Banacha, z normą

$$\|\phi\| = \int_0^\infty |\phi(t)|\,\mathrm{d}t.$$

W istocie $L^1(\mathbb{R}^+)$ można traktować jako podprzestrzeń $L^1(\mathbb{R})$ o ile utożsamimy funkcję całkowalną na prawej półosi z jej rozszerzeniem na całą oś, które na półosi lewej równe jest stale zero. Obie przestrzenie są też przykładami tak zwanych algebr splotowych Banacha, to znaczy (przemiennych) algebr Banacha, w których „mnożenie" jest splotem. Dokładniej rzecz biorąc, przestrzeń Banacha $\mathbb{X}$ nazywana jest algebrą Banacha, jeśli zdefiniowano w niej dodatkowe działanie łączne $\mathbb{X} \times \mathbb{X} \to \mathbb{X}$, przyporządkowujące parze $x, y \in \mathbb{X}$ jej iloczyn $xy \in \mathbb{X}$, o następujących własnościach, wiążących je ze strukturą przestrzeni Banacha:

(a) $\|xy\| \le \|x\|\|y\|$,

(b) $(\alpha x)y = \alpha(xy) = x(\alpha y), \alpha \in \mathbb{R}$,

(c) $x(y_1 + y_2) = xy_1 + xy_2$,

(c′) $(y_1 + y_2)x = y_1 x + y_2 x$.

W powyższych warunkach x, y, y_1 i y_2 to elementy $\mathbb{X}$, a α jest skalarem. Jeśli zachodzi ponadto warunek

(d) $xy = yx$,

to mówimy, że algebra jest przemienna – w tym wypadku (c′) jest zbędne, gdyż wynika z (c) i (d).

„Mnożeniem” w algebrze $L^1(\mathbb{R}^+)$ jest splot. To znaczy, że dwóm funkcjom $\phi_1, \phi_2 \in L^1(\mathbb{R}^+)$ przyporządkowujemy $\varphi = \psi_1 * \psi_2$ dane wzorem

$$\varphi(t) = \int_0^t \psi_1(t-s)\psi_2(s)\,\mathrm{d}s, \qquad t \ge 0. \tag{5.1}$$

Zdefiniowane tak φ jest rzeczywiście elementem $L^1(\mathbb{R}^+)$. Dowodzi tego następujący rachunek:

$$\begin{aligned} \|\psi_1 * \psi_2\|_{L^1(\mathbb{R}^+)} &= \int_0^\infty \left| \int_0^t \psi_1(t-s)\psi_2(s)\,\mathrm{d}s \right| \mathrm{d}t \\ &\le \int_0^\infty \int_s^\infty |\psi_1(t-s)\psi_2(s)|\,\mathrm{d}t\,\mathrm{d}s \\ &= \int_0^\infty \int_0^\infty |\psi_1(t)|\,|\psi_2(s)|\,\mathrm{d}t\,\mathrm{d}s \\ &= \|\psi_1\|_{L^1(\mathbb{R}^+)}\,\|\psi_2\|_{L^1(\mathbb{R}^+)}. \end{aligned} \tag{5.2}$$

Innymi słowy, jeśli ψ_1 i ψ_2 należą do $L^1(\mathbb{R}^+)$, to ich splot również i zachodzi nierówność

$$\|\psi_1 * \psi_2\|_{L^1(\mathbb{R}^+)} \le \|\psi_1\|_{L^1(\mathbb{R}^+)}\,\|\psi_2\|_{L^1(\mathbb{R}^+)}.$$

Pokazuje to, że warunek (a) definicji algebry Banacha jest spełniony. Pozostałe warunki sprawdza się bardzo łatwo.

5.2 Model McKendricka–von Foerstera

Po tym wstępie przejdźmy do opisu modelu McKendricka. Nieujemną funkcję ϕ zdefiniowaną na $[0, \infty)$ nazywamy gęstością danej populacji, jeśli dla dowolnych nieujemnych liczb $c < d$ całka

$$\int_c^d \phi(a)\,\mathrm{d}a$$

równa jest ilości osobników tejże populacji w wieku między c i d (zgodnie z ustaloną konwencją, zmienną niezależną oznaczamy tu a bo opisuje ona wiek, a po angielsku wiek to „age"). Pojęcie to należy odróżniać od gęstości prawdopodobieństwa: dla gęstości populacji całka $\int_0^\infty \phi(a)\,\mathrm{d}a$ jest całkowitą wielkością populacji, zwykle zmienia się w czasie i raczej nie jest równa 1.

W dalszej analizie będziemy zdecydowanie nieściśle, ale za to przemawiając do intuicji, mówić, iż osobników w wieku a jest $\phi(a)$. Dokładniejsze sformułowanie brzmiałoby tak: osobników w wieku mieszczącym się w przedziale $[a, a+\delta a]$ jest $\int_a^{a+\delta a} \phi(c)\,\mathrm{d}c$, ale w praktyce operowanie nim jest niewygodne.

No właśnie: ϕ powinno zmieniać się w czasie. Czy znając początkową gęstość populacji, możemy opisać ewolucję jej gęstości w przyszłości? Oczywiście zależy to od tego, czy znamy reguły nią rządzące. W tym celu rozważamy dwie nieujemne funkcje: intensywność zgonów

$$\mu = \mu(a)$$

i tempo narodzin

$$b = b(a),$$

zależne od wieku osobnika. Funkcja $\mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}$ opisuje więc prawdopodobieństwo, że osobnik dożyje do wieku $a > 0$. Zakładamy, że μ jest funkcją całkowalną, a b jest ograniczone – co z biologicznego punktu widzenia wydaje się być założeniem sensownym.

Gdyby $b \equiv 0$, to dynamika byłaby dość prosta: osobniki, które w czasie $t > 0$ są w wieku $a > t$, w czasie 0 miały lat $a - t$. Z tych zaś, które w czasie 0 były w wieku $a - t$ przeżyła tylko frakcja równa $\mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c}$. Bierze się to stąd, że

$$\mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c} = \frac{\mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}}{\mathrm{e}^{-\int_0^{a-t} \mu(c)\,\mathrm{d}c}}$$

jest prawdopodobieństwem warunkowym przeżycia osobnika do wieku a pod warunkiem, że wiemy, iż dożył on wieku $a - t$. Innymi słowy, jeśli ϕ_0 jest gęstością populacji w czasie 0, to w czasie $t > 0$ dla $a > t$ gęstość równa będzie

$$\mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c}\phi_0(a-t).$$

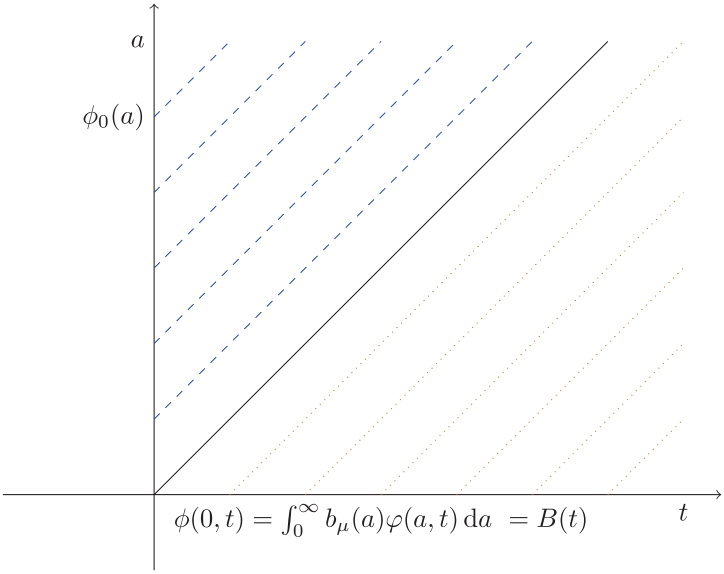
Dla $a < t$, wobec braku narodzin gęstość równa będzie 0. Pisząc $\phi(t,\cdot)$ na oznaczenie gęstości w czasie $t > 0$, mamy więc

$$\phi(t,a) = \begin{cases} \mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c}\phi_0(a-t), & a \ge t, \\ 0, & a < t. \end{cases} \tag{5.3}$$

By opisać ciekawszą sytuację, gdy $b \not\equiv 0$, rozważmy funkcję

$$B(t) = \int_0^\infty b(a)\phi(t,a)\,\mathrm{d}a, \qquad t \ge 0.$$

$B(t)$ jest ilością osobników, którzy rodzą się w całej populacji w czasie t. Gdybyśmy znali B, to moglibyśmy uzupełnić wzór (5.3), wpisując w jego dolną część (zamiast



Rys. 5.1: Dynamika populacji w modelu McKendricka: ϕ „propaguje się" wzdłuż prostych, na których $a - t$ jest stałe. W górnej części pierwszej ćwiartki ϕ zależy od ϕ_0, w dolnej od B, które trzeba wyliczyć z równania odnowienia.

zera), wielkość $\mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}B(t-a)$:

$$\phi(t, a\ ) = \begin{cases} \mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c}\phi_0(a-t), & a\ \geq t, \\ \mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}B(t-a), & a\ < t, \end{cases} \tag{5.4}$$

(zob. rysunek 5.1). Rzeczywiście, każdy osobnik, który w czasie t ma $a < t$ lat musiał narodzić się a czasu temu z kogoś, kto żył w czasie $t - a$, a $\mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}$ jest prawdopodobieństwem, że osobnik wtedy narodzony dożył do czasu t.

Musimy jeszcze znaleźć B, bo na razie wzór (5.4) wyraża $\phi(t,\cdot)$ w terminach samej siebie. Wstawiając (5.4) do definicji B, otrzymujemy jednak:

$$\begin{aligned} B(t) &= \int_0^t b(a)\mathrm{e}^{-\int_0^a \mu(c)\,\mathrm{d}c}B(t-a)\,\mathrm{d}a\ + \int_t^\infty b(a)\mathrm{e}^{-\int_{a-t}^a \mu(c)\,\mathrm{d}c}\phi_0(a-t)\,\mathrm{d}a \quad (5.5) \\ &= \int_0^t \mathrm{e}^{-\int_0^{t-a} \mu(c)\,\mathrm{d}c}b(t-a)B(a)\,\mathrm{d}a\ + \int_0^\infty \mathrm{e}^{-\int_a^{a+t} \mu(c)\,\mathrm{d}c}\phi_0(a)b(a+t)\,\mathrm{d}a. \end{aligned}$$

Wprowadzając oznaczenie

$$C(t) = \int_0^\infty \mathrm{e}^{-\int_a^{a+t} \mu(c)\,\mathrm{d}c}\phi_0(a)b(a+t)\,\mathrm{d}a,$$

wzór ten możemy napisać prościej:

$$B = b_\mu * B + C, \tag{5.6}$$

gdzie $b_\mu(t) = \mathrm{e}^{-\int_0^t \mu(c)\,\mathrm{d}c} b(t)$. Zwróćmy uwagę na fakt, że funkcja C zależy już tylko od początkowego rozkładu oraz od μ i b, możemy więc traktować ją jako daną, a szukamy oczywiście B. Równanie (5.6) jest szczególnym przypadkiem równania Volterry drugiego rodzaju, a także równaniem splotowym. W kontekście modelu McKendricka jest to tak zwane równanie odnowienia.

5.3 Równanie odnowienia i jego rozwiązanie

Jego interpretacja jest następująca. Osobników, którzy rodzą się w czasie t można podzielić na dwie klasy. Do pierwszej z nich zaliczymy dzieci tych, którzy żyli już w czasie 0. Na początku osobników w wieku a było $\phi_0(a)$, do czasu $t > 0$ przeżyło ich $\mathrm{e}^{-\int_a^{a+t} \mu(c)\,\mathrm{d}c} \phi(a)$ i teraz są w wieku $a + t$. Dzieci z nich zrodzonych w czasie t jest więc $C(t)$. Druga klasa to dzieci osobników, którzy urodzili się w międzyczasie (czy to z pokolenia wyjściowego, czy też z kolejnych). Rodzic osobnika z tej klasy musiał urodzić się przed czasem t. Rodziców narodzonych w $a \in [0, t]$ jest $B(a)$; dożywają oni do czasu t z prawdopodobieństwem $\mathrm{e}^{-\int_0^{t-a} \mu(c)\,\mathrm{d}c}$ i są wtedy w wieku $t - a$, stąd całkowita ilość narodzin osobników tej klasy równa jest pierwszej całce w (5.5).

Jeśli zgodzimy się, że splot jest bardzo podobny do mnożenia (a tak jest – zob. zadanie 5.2), to równanie (5.6) okaże się bardzo podobne do równania liniowego. Jego rozwiązaniem powinien więc być szereg

$$B = \sum_{n=0}^{\infty} b_\mu^{n*} * C, \tag{5.7}$$

gdzie b_μ^{n*} jest n-tą potęgą splotową funkcji b_μ (zdefiniowaną oczywiście indukcyjnie). W rozdziale 12 wyjaśni się w jakim sensie powyższy szereg jest zbieżny (w zadaniu 5.4 podajemy też konkretny przykład zastosowania w.w. wzoru do rozwiązania równania (5.6)). Na razie zadowolimy się stwierdzeniem, że istnieje dokładnie jedno rozwiązanie równania odnowienia (które tym samym w pełni determinuje $\phi(t, \cdot)$ ze wzoru (5.5)).

Nim dowiedziemy tego ostatniego twierdzenia, zwróćmy uwagę na to, że niewielkie są szanse na to, iż B należeć będzie do $L^1(\mathbb{R}^+)$: już samo b jest tylko ograniczone i (na przykład jeśli $\mu = 0$) nie musi być całkowalne. Podobnie trudno sprawdzić w ogólności, czy $C \in L^1(\mathbb{R}^+)$. To w sumie logiczne, bo jeśli nie ma śmierci a intensywność narodzin będzie duża, to wielkość populacji będzie szybko rosła a wraz z nią lawinowo rosnąć będzie B. Niezerowe μ wzrost ten będzie hamować, ale dopóki nie znamy dokładniejszych zależności między μ i b trudno stwierdzić jaka będzie dynamika ilości narodzin w czasie.

Musimy więc poszukać B w większej przestrzeni. W tym celu dla $\omega > 0$, zdefiniujmy $L^1_\omega(\mathbb{R}^+)$ jako przestrzeń funkcji mierzalnych ϕ określonych na $\mathbb{R}^+$ dla których całka $\int_0^\infty \mathrm{e}^{-\omega a} |\phi(a)|\,\mathrm{d}a$ jest skończona. Z normą równą powyższej całce, jest to przestrzeń Banacha (zaraz to uzasadnimy, choć można by oczywiście zrobić dowód „na piechotę”).

Choć może się to wydawać dziwne, $L^1_\omega(\mathbb{R}^+)$ mimo iż wyraźnie znacznie większa niż $L^1(\mathbb{R}^+)$ jest z pewnego punktu widzenia ... taka sama jak ta ostatnia. To troszkę tak jak z liczbami naturalnymi i liczbami parzystymi. Choć tych drugich jest wyraźnie mniej niż pierwszych, jest ich w istocie tyle samo. Tu jest podobnie: istnieje wzajemnie jednoznaczne i do tego liniowe przekształcenie I jednego zbioru na drugi. Mówiąc precyzyjniej, każdemu $\phi \in L^1_\omega(\mathbb{R}^+)$ przyporządkować możemy $I\phi \in L^1(\mathbb{R}^+)$ (zgodnie z konwencją przyjętą dla odwzorowań liniowych piszemy $I\phi$ zamiast $I(\phi)$, patrz rozdział 12), daną wzorem

$$(I\phi)(t) = \mathrm{e}^{-\omega t}\phi(t), \qquad t \geq 0, \tag{5.8}$$

a liniowość przekształcenia I to dokładnie tyle, że dla dowolnych $\phi, \varphi \in L^1_\omega(\mathbb{R}^+)$ oraz dowolnych skalarów α i β zachodzi zależność:

$$I(\alpha\phi + \beta\varphi) = \alpha I\phi + \beta I\varphi.$$

Dość jasne jest, że I jest różnowartościowe i „na”: dla każdego $\psi \in L^1(\mathbb{R}^+)$ funkcja ϕ dana wzorem $\phi(t) = \mathrm{e}^{\omega t}\psi(t)$ jest elementem $L^1_\omega(\mathbb{R}^+)$ i do tego $I\phi = \psi$. Z samej definicji normy w obu przestrzeniach mamy też

$$\|I\phi\|_{L^1(\mathbb{R}^+)} = \|\phi\|_{L^1_\omega(\mathbb{R}^+)}. \tag{5.9}$$

I już, bo wobec tej równości zupełność $L^1_\omega(\mathbb{R}^+)$ jest wnioskiem z zupełności $L^1(\mathbb{R}^+)$.

Rzeczywiście, jeśli ciąg $(\phi_n)_{n\geq 1}$ spełnia warunek Cauchy'ego w $L^1_\omega(\mathbb{R}^+)$ to jego obraz $(I\phi_n)_{n\geq 1}$ spełnia tenże warunek w $L^1(\mathbb{R}^+)$. Z zupełności $L^1(\mathbb{R}^+)$ wnioskujemy, że istnieje $\psi \in L^1(\mathbb{R}^+)$, które jest granicą ciągu $(I\phi_n)_{n\geq 1}$. To ψ jest jednak obrazem pewnego $\phi \in L^1_\omega(\mathbb{R}^+)$: $\psi = I\phi$. Z równości norm (i liniowości I) wnioskujemy zatem, że

$$\|\phi_n - \phi\|_{L^1_\omega(\mathbb{R}^+)} = \|I(\phi_n - \phi)\|_{L^1(\mathbb{R}^+)} = \|I\phi_n - I\phi\|_{L^1_\omega(\mathbb{R}^+)},$$

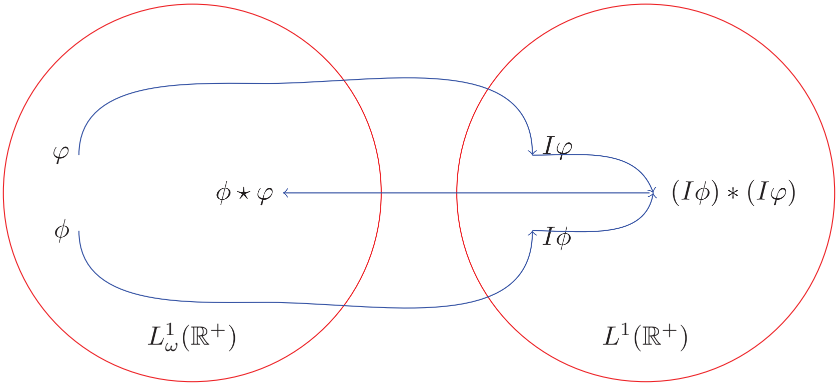
co jak wiemy dąży do zera, gdy $n \to \infty$. To świadczy jednak o tym, że $\phi = \lim_{n\to\infty} \phi_n$ i tym samym kończy dowód.

Tożsamość (5.9) dowodzi nieco więcej: $L^1_\omega(\mathbb{R}^+)$ jest algebrą Banacha. W tym celu wystarczy wprowadzić mnożenie w $L^1_\omega(\mathbb{R}^+)$, chwilowo oznaczane symbolem $\star$, wzorem (zob. rysunek 5.2)

$$\phi \star \varphi = I^{-1}[(I\phi) * (I\varphi)].$$

Sprawdzenie, że $\star$ spełnia warunki podane w definicji algebry Banacha jest bardzo proste: na przykład własność (a) dowodzi się następująco:

$$\begin{aligned}\|\phi \star \varphi\|_{L^1_\omega(\mathbb{R}^+)} &= \|(I\phi) * (I\varphi)\|_{L^1(\mathbb{R}^+)} \leq \|I\phi\|_{L^1(\mathbb{R}^+)}\|I\varphi\|_{L^1(\mathbb{R}^+)} \\ &= \|\phi\|_{L^1_\omega(\mathbb{R}^+)}\|\varphi\|_{L^1_\omega(\mathbb{R}^+)};\end{aligned}$$



Rys. 5.2: Definicja mnożenia w $L^1_\omega(\mathbb{R}^+)$: najpierw znajdujemy obrazy ϕ i ψ w $L^1(\mathbb{R}^+)$, potem te obrazy splatamy, a na końcu wracamy do przestrzeni $L^1_\omega(\mathbb{R}^+)$. Przedstawiany tu schemat rozumowania opiera się na śmiałej hipotezie Szara Szarikowa, iż przestrzenie Banacha są kuliste. Ze względu na to, że inni naukowcy sugerują raczej kształt nerki (grupa badawcza z Nierestädt), lub wręcz jaja (Eugene Egg), formalny dowód hipotezy tej nie używa.

w rachunku tym dwukrotnie użyliśmy (5.9) i raz faktu, że $L^1(\mathbb{R}^+)$ jest algebrą. Jak się okazuje, można podać jawny wzór na działanie $\star$. W tym celu zauważmy, że

$$\begin{aligned}[(I\phi) * (I\varphi)](t) &= \int_0^t \mathrm{e}^{-\omega(t-s)}\phi(t-s)\mathrm{e}^{-\omega s}\varphi(s)\,\mathrm{d}s \\ &= \mathrm{e}^{-\omega t}\int_0^t \phi(t-s)\varphi(s)\,\mathrm{d}s \\ &= [I(\phi * \varphi)](t),\end{aligned}$$

a to dowodzi, że

$$(\phi \star \varphi)(t) = (\phi * \varphi)(t).$$

Innymi słowy, $\star$ okazuje się zwykłym splotem.

Po tych wszystkich przygotowaniach przechodzimy do dowodu, że (5.6) ma dokładnie jedno rozwiązanie – należące do wszystkich algebr $L^1_\omega(\mathbb{R}^+)$ z odpowiednio dużym ω. Zacznijmy od tego, że C należy do $L^1_\omega(\mathbb{R}^+)$. Rzeczywiście:

$$\begin{aligned}\|C\|_{L^1_\omega(\mathbb{R}^+)} &\le \int_0^\infty \mathrm{e}^{-\omega t}\int_0^\infty |\phi_0(a)|b(a+t)\,\mathrm{d}a\ \mathrm{d}t \le \frac{\|b\|_\infty}{\omega}\int_0^\infty |\phi_0(a)|\,\mathrm{d}a \\ &= \frac{\|b\|_\infty}{\omega}\|\phi_0\|_{L^1(\mathbb{R}^+)} < \infty.\end{aligned}$$

Recenzent o podpisie do rysunku 5.2: Czy studenci zrozumieją tego suchara? Oby ta uwaga nie stała się podwaliną nowej szkoły.

Podobnie jest z b_μ: funkcja ta jest ograniczona, więc $t \mapsto \mathrm{e}^{-\omega t} b_\mu(t)$ jest całkowalna i

$$\|b_\mu\|_{L^1_\omega(\mathbb{R}^+)} \leq \frac{\|b_\mu\|_\infty}{\omega}.$$

Wynika stąd, że odwzorowanie

$$T(\psi) = C + b_\mu * \psi, \tag{5.10}$$

przekształca $L^1_\omega(\mathbb{R}^+)$ w siebie. Mamy też

$$\|T(\psi_1) - T(\psi_2)\|_{L^1_\omega(\mathbb{R}^+)} = \|b_\mu * (\psi_1 - \psi_2)\|_{L^1_\omega(\mathbb{R}^+)} \leq \frac{\|b\|_\infty}{\omega} \|\psi_1 - \psi_2\|_{L^1_\omega(\mathbb{R}^+)}.$$

Dla $\omega > \|b\|_\infty$, odwzorowanie T jest zatem kontrakcją i istnienie (jedynego) B jest zagwarantowane.

5.4 Zadania

Zadanie 5.1. Niech $e_\lambda(t) = \mathrm{e}^{-\lambda t}, t \geq 0$, $e_\lambda \in L^1(\mathbb{R}^+), \lambda > 0$. Udowodnij, że zachodzi tożsamość Hilberta:

$$e_\lambda - e_\mu = (\mu - \lambda) e_\lambda * e_\mu.$$

Zadanie 5.2. Pokaż, że dla dowolnych funkcji f, g, h i stałych a i b zachodzą wzory $f * g = g * f, (af + bg) * h = af * h + bg * h$. Pokaż też, że dla funkcji $i(x) = 1, x \geq 0$ zachodzi wzór $i * f(x) = \int_0^x f(y)\,\mathrm{d}y$.

Zadanie 5.3. Podaj indukcyjną definicję n-tej potęgi splotowej danej funkcji. Sprawdź ponadto, że dla dowolnego dodatniego λ, n-ta potęga splotowa funkcji $\phi(t) = \lambda \mathrm{e}^{-\lambda t}, t \geq 0$, dana jest wzorem

$$\phi^{*n}(t) = \mathrm{e}^{-\lambda t} \frac{\lambda^n t^n}{(n-1)!}, \qquad t \geq 0.$$

Zadanie 5.4. Rozważmy przypadek, gdy intensywności śmierci i narodzin w modelu McKendricka nie zależą od wieku i obie są równe stałej dodatniej λ. Sprawdź, że mamy tu

$$C(t) = c\lambda \mathrm{e}^{-\lambda t}, \qquad b_\mu(t) = \lambda \mathrm{e}^{-\lambda t}, \qquad t \geq 0,$$

dla $c = \|\phi_0\| = \int_0^\infty \phi_0(a)\,\mathrm{d}a$. Korzystając z zadania poprzedniego pokaż także, iż formalny szereg $\sum_{n=1}^\infty b_\mu^{*n}$ sprowadza się do funkcji stale równej λ. Wywnioskuj z (5.7), że rozwiązanie równania (5.6) powinno być dane wzorem $B(t) = c\lambda$, a następnie sprawdź, że funkcja ta rzeczywiście spełnia to równanie.

Zadanie 5.5. Udowodnij, że przestrzeń l^1 absolutnie sumowalnych ciągów $(\xi_i)_{i \geq 0}$ jest algebrą splotową jeśli splot zdefiniujemy wzorem

$$(\xi_i)_{i\geq 0} * (\eta_i)_{i \geq 0} = (\zeta_i)_{i \geq 0},$$

w którym $\zeta_i = \sum_{k=0}^i \xi_{i-k} \eta_k$.

Zadanie 5.6. Niech $\alpha \in (-1, 1)$ będzie daną liczbą. Zdefiniujmy $a \in l^1$ wzorem

$$a \ = (1, \alpha, \alpha^2, \cdots).$$

Udowodnij, że dla każdego $y \in l^1$ istnieje dokładnie jeden $x \in l^1$ spełniający równanie

$$x - \alpha(1 - \alpha)a * x = y.$$

Zadanie 5.7. Wprowadźmy mnożenie funkcji z przestrzeni $C[a, b]$ wzorem:

$$(xy)(t) = x(t)y(t), \qquad t \in [a, b].$$

Udowodnij, że $C[a, b]$ z tym działaniem jest przemienną algebrą Banacha.

Zadanie 5.8. Udowodnij, że przestrzeń Banacha $L^1(\mathbb{R})$ jest (splotową) algebrą Banacha; splot jest tu zdefiniowany tak:

$$(\phi * \varphi)(t) = \int_{-\infty}^{\infty} \phi(t - s)\varphi(s)\,\mathrm{d}s, \qquad t \geq 0.$$

Sprawdź, że jeśli ϕ i φ są równe zero na lewej półosi, splot ten pokrywa się z (5.1).

Zadanie 5.9. W przestrzeni Banacha $L^1(\mathbb{R}^+)$ wprowadzamy nowe mnożenie

$$(\phi \diamond \varphi)(t) = \frac{1}{2} \int_{-\infty}^{\infty} \phi(|t - s|)\varphi(|s|)\,\mathrm{d}s, \qquad t \geq 0.$$

Udowodnij, że z tym działaniem $L^1(\mathbb{R}^+)$ jest algebrą Banacha (raczej nie możemy jej w tym kontekście nazywać algebrą splotową, choć wyrażenie po prawej stronie jest splotem parzystych rozszerzeń ϕ i ψ – wymnożonym przez $\frac{1}{2}$).

Rozdział 6

Całka Riemanna w przestrzeni Banacha

6.1 Całka Riemanna w przestrzeni unormowanej

Jednym z fundamentalnych wniosków z zupełności jest to, że klasyczny rachunek różniczkowy można rozszerzyć na funkcje, które przyjmują wartości w przestrzeni Banacha. W szczególności można mówić o całkowalności funkcji w sensie Riemanna (całka typu Lebesgue'a, zwana całką Bochnera, też jest znana, ale jej wprowadzenie jest trudniejsze – zob. np. [13]). Oto szczegóły. Niech $a < b$ będą liczbami rzeczywistymi i załóżmy, że $x : [a, b] \to \mathbb{X}$, $t \mapsto x(t)$ jest funkcją na $[a, b]$ o wartościach w przestrzeni unormowanej $\mathbb{X}$. Rozważmy dwa ciągi $\mathcal{T} = (t_i)_{i=0,..,k}$ i $\mathcal{P} = (p_i)_{i=0,...,k-1}$ punktów przedziału $[a, b]$ (k jest tu liczbą naturalną) tak dobrane, że

$$a \ = t_0 < t_1 < \cdots < t_k = b, \qquad t_0 \le p_0 \le t_1 \le \cdots \le t_{k-1} \le p_{k-1} \le t_k. \tag{6.1}$$

Zdefiniujmy związaną z nimi liczbę $\Delta(\mathcal{T}) = \sup_{0 \le i \le k}(t_i - t_{i-1})$ oraz element $\mathbb{X}$ dany wzorem

$$S(\mathcal{T}, \mathcal{P}, x) = \sum_{i=0}^{k} x(p_i)(t_{i+1} - t_i). \tag{6.2}$$

(Wzór ten ma sens, bo w przestrzeni liniowej możliwe jest dodawanie wektorów i mnożenie ich przez skalary). Jeśli dla dowolnego ciągu par $(\mathcal{T}_n, \mathcal{P}_n)$ spełniającego warunek $\lim_{n\to\infty} \Delta(\mathcal{T}_n) = 0$, istnieje granica

$$\lim_{n\to\infty} S(\mathcal{T}_n, \mathcal{P}_n, x),$$

i nie zależy od wyboru $(\mathcal{T}_n, \mathcal{P}_n)$, to funkcję x nazywamy całkowalną w sensie Riemanna. Granica wyżej wspomniana nazywana jest całką (Riemanna); oznacza się ją symbolem $\int_a^b x(t)\,\mathrm{d}t$ – oczywiście o jej istnieniu nie można byłoby mówić, gdyby w przestrzeni $\mathbb{X}$ nie została wprowadzona norma, a przynajmniej metryka.

W następnym podrozdziale dowiedziemy, że dla funkcji o wartościach w przestrzeni Banacha prawdziwe jest następujące twierdzenie znane z kursu funkcji rzeczywistych; Czytelnik zwróci uwagę na fakt, że założenie zupełności odgrywa w nim kluczową rolę:

funkcje ciągłe są całkowalne.

6.2 Całkowalność funkcji ciągłych

Niech będzie dane $\epsilon > 0$. Skoro $[a, b]$ jest zbiorem zwartym, ciągłość x implikuje jej ciągłość jednostajną. Możemy więc tak wybrać $\delta > 0$, że warunki $|s - t| < \delta$ i $s, t \in [a, b]$ implikować będą $\|x(s) - x(t)\| < \epsilon$. Niech ciągi $\mathcal{T} = (t_i)_{i=0,...,k}$ i $\mathcal{T}' = (t'_i)_{i=0,...,k'}$ spełniają warunki $\Delta(\mathcal{T}) < \delta$ i $\Delta(\mathcal{T}') < \delta$. Niech też $\mathcal{P}$ i $\mathcal{P}'$ będą związanymi z nimi ciągami punktów pośrednich (tak więc zachodzi (6.1)).

Niech $\mathcal{T}'' = (t''_i)_{i=1,...,k''}$ będzie ciągiem zawierającym wszystkie wyrazy $\mathcal{T}$ i $\mathcal{T}'$. Przyjmijmy też $\mathcal{P}'' = (p''_i)_{i=0,...,k''-1} \equiv (t''_i)_{i=0,...,k''-1}$ (tak więc punkty pośrednie leżą na końcach przedziału). Kluczowym krokiem dowodu jest przekonanie się, że

$$\|S(\mathcal{T}, \mathcal{P}, x) - S(\mathcal{T}'', \mathcal{P}'', x)\| \leq \epsilon(b - a). \tag{6.3}$$

Zauważmy w tym celu, że $\Delta(\mathcal{T}'') < \delta$ i $k'' \leq k + k' - 2$, bo poza $t_0 = t'_0 = a$ i $t_k = t'_{k'} = b$ mogą być jeszcze inne $t_i = t'_j$, $i = 1, ..., k-1, j = 1, ..., k'-1$. Odcinek $[t_i, t_{i+1}]$, $i = 0, ..., k$ jest albo tożsamy z pewnym $[t''_j, t''_{j+1}], j \in \{0, ..., k'' - 1\}$ albo jest takich odcinków sumą: powiedzmy, że $[t_i, t_{i+1}] = [t''_j, t''_{j+1}] \cup ... \cup [t''_{j+l}, t''_{j+l+1}]$ dla pewnego l. A zatem

$$\begin{aligned} &\left\| x(p_i)(t_{i+1} - t_i) - \sum_{m=0}^{l} x(t''_{j+m})(t''_{j+m+1} - t''_{j+m}) \right\| \\ &= \left\| \sum_{m=0}^{l} [x(p_i) - x(t''_{j+m})](t''_{j+m+1} - t''_{j+m}) \right\| \\ &\leq \epsilon \sum_{m=0}^{l} (t''_{j+m+1} - t''_{j+m}) = \epsilon(t_{i+1} - t_i), \end{aligned}$$

ponieważ zarówno p_i jak i t''_{j+m} należą do $[t_i, t_{i+1}]$, co powoduje, że $|p_i - t''_{j+m}| < \delta$. Sumując po i, otrzymujemy więc (6.3).

Rozumowanie to działa zarówno wobec $\mathcal{T}'$ jak i wobec $\mathcal{T}$. Dowodzi to, że

$$\|S(\mathcal{T}, \mathcal{P}, x) - S(\mathcal{T}', \mathcal{P}', x)\| \leq 2\epsilon(b - a), \tag{6.4}$$

dla dowolnych ciągów $\mathcal{T}$ i $\mathcal{T}'$ spełniających warunki $\Delta(\mathcal{T}) < \delta$ i $\Delta(\mathcal{T}') < \delta$, oraz dowolnych ciągów $\mathcal{P}$ i $\mathcal{P}'$ punktów pośrednich. To przekonuje nas, że dla dowolnego ciągu $(\mathcal{T}_n, \mathcal{P}_n)$ spełniającego warunek $\lim_{n\to\infty} \Delta(\mathcal{T}_n) = 0$, $S(\mathcal{T}_n, \mathcal{P}_n, x)$ jest ciągiem Cauchy'ego, a zatem ma granicę. Wobec (6.4) granica ta nie zależy od wyboru $(\mathcal{T}_n, \mathcal{P}_n)$. To kończy dowód.

6.3 Zadania

Zadanie 6.1. Niech $\mathbb{X}$ będzie przestrzenią Banacha i załóżmy, że $t \mapsto x(t) \in \mathbb{X}$ jest ciągła na przedziale $[a, b]$. Funkcja rzeczywista $t \to \|x(t)\|$ jest wtedy również ciągła (a zatem i całkowalna). Udowodnij, że (por. zadanie 3.11)

$$\left\| \int_a^b x(t)\,\mathrm{d}t \right\| \le \int_a^b \|x(t)\|\,\mathrm{d}t. \tag{6.5}$$

Zadanie 6.2. Niech $(a_i)_{i\ge1}$ będzie ciągiem liczb dodatnich, dążącym do $a > 0$ i niech funkcja $x : [0, 1] \to c$ będzie dana wzorem $x(t) = (\mathrm{e}^{-a_i t})_{i\ge1}$, $t \in [0, 1]$. Udowodnij jej ciągłość, a następnie oblicz $\int_0^1 x(t)\,\mathrm{d}t \in c$. Napisz nierówność (6.5) dla tej konkretnej całki.

Zadanie 6.3. Niech $t \mapsto u(t)$ będzie skalarną funkcją ciągłą na przedziale $[a, b]$, a x – elementem przestrzeni unormowanej. Przekonaj się, że funkcja o wartościach wektorowych $t \mapsto u(t)x$ jest ciągła i całkowalna oraz zachodzi wzór

$$\int_a^b u(t)x\,\mathrm{d}t = \left(\int_a^b u(t)\,\mathrm{d}t \right) x.$$

(Nie zakładamy tu, że przestrzeń, do której należy x jest zupełna).

Rozdział 7

Twierdzenie Stone'a–Weierstrassa

7.1 Główne twierdzenie

Naszym głównym celem w tym rozdziale jest następujące twierdzenie Weierstrassa, do którego będziemy jeszcze wracać w tej książeczce parokrotnie.

Twierdzenie 7.1 (Weierstrassa)**.** *Każdą funkcję ciągłą na przedziale* $[a, b]$ $(a < b)$ *można aproksymować jednostajnie wielomianami. Innymi słowy, dla każdej funkcji* $f \in C[a, b]$ *i* $\epsilon > 0$ *można znaleźć taki wielomian* w*, że*

$$\|f - w\| = \sup_{x \in [a,b]} \|f(x) - w(x)\| < \epsilon.$$

W istocie będziemy w stanie udowodnić twierdzenie znacznie mocniejsze, zwane twierdzeniem Stone'a–Weierstrassa. Dotyczy ono sytuacji, gdy – zamiast funkcji na $[a, b]$ – mamy do czynienia z funkcjami na abstrakcyjnej zwartej przestrzeni topologicznej S (Hausdorffa). Pojęcie wielomianu oczywiście nie ma wtedy sensu, właściwe uogólnienie polega więc na podaniu *własności przestrzeni wielomianów*, które są rzeczywistym powodem, dla którego przestrzeń ta jest zbiorem gęstym w $C(S)$. Proszę zwrócić uwagę na fakt, że dokonujemy tu manewru typowego dla analizy funkcjonalnej: nie interesują nas wielomiany jako takie ani ich własności – interesują nas własności *przestrzeni* wielomianów.

Nim pójdziemy dalej, chciałbym zatrzymać się na chwilę na $C(S)$, przypominając kilka pojęć z topologii. Przypomnijmy, że przestrzeń topologiczną S nazywamy zwartą, jeśli z każdego jej pokrycia zbiorami otwartymi można wybrać podpokrycie skończone. Rzeczywista funkcja ciągła na S to taka, dla której przeciwobrazy zbiorów otwartych w $\mathbb{R}$ są zbiorami otwartymi w S. Okazuje się, że – podobnie jak w przypadku zwartych podzbiorów $\mathbb{R}$ – funkcje ciągłe na przestrzeniach zwartych są ograniczone i osiągają swoje suprema. W szczególności dla $f \in C(S)$ możemy

myśleć o
$$\|f\| = \sup_{s\in S} |f(s)|.$$
Zupełnie analogicznie, jak w znanym nam już przypadku dowodzi się, że jest to norma, a $C(S)$ w nią wyposażona staje się przestrzenią Banacha.

Wracamy do wspomnianych wyżej kluczowych własności przestrzeni wielomianów $W \subset C[a, b]$. Po pierwsze

(a) jeśli $s \neq t$ należą do $[a, b]$, to istnieje taki $w \in W$, że $w(s) \neq w(t)$.

Własność ta nazywana jest „rozróżnianiem punktów”[9]. Wielomianów spełniających (a) jest w istocie wiele: jednym z nich jest na przykład funkcja liniowa z odpowiednio dobranymi współczynnikami
$$w(x) = ax + b = \frac{1}{t-s}x - \frac{s}{t-s}.$$

Własność druga jest jeszcze prostsza:

(b) dla każdych w i v z W wielomianem jest także ich iloczyn wv, a także dowolna kombinacja liniowa $\alpha w \ + \beta v$.

Sprawdzić ją łatwo (choćby indukcyjnie ze względu na stopień wielomianu), ale jej znaczenie jest nie do przecenienia. Ta własność algebraiczna (nieprawdaż?) okazuje się kluczem do interesującego nas wyniku (patrz niżej).

Punktem wyjścia w twierdzeniu Stone'a – Weierstrassa jest podzbiór $W \subset C(S)$. Nie jest on zbiorem wielomianów, bo jak już to zauważyliśmy wcześniej, o wielomianach nie ma tu sensu mówić. Jest on jednak do ich zbioru podobny w tym sensie, że jest podalgebrą $C(S)$ (to odpowiednik własności (b)): wraz z w i v z W należą do W także ich iloczyn i każda kombinacja liniowa. Nie zwlekając dłużej przedstawmy wreszcie ten kluczowy wynik.

Twierdzenie 7.2 (Stone'a–Weierstrassa). *Każda rozróżniająca punkty S podalgebra $W \subset C(S)$, która zawiera funkcje stałe, jest zbiorem gęstym w $C(S)$.*

Rozróżnianie punktów jest tu oczywiście odpowiednikiem własności (a): dla każdych $s \neq t, s, t \in S$ istnieje takie $w \in W$, że $w(s) \neq w(t)$. Zapewniliśmy też sobie przynależność do W funkcji stałych – tak jak to jest w przypadku wielomianów. Skomentujmy jednak sprawę mniej oczywistą: dlaczego sformułowane wyżej uogólnienie twierdzenie Weiestrassa jest istotne? Początkujący Czytelnik mógłby bowiem zastanawiać się, po co w ogóle zajmujemy się podalgebrami $C(S)$. Poza niepohamowaną u matematyków żądzą poznania motywuje nas fakt, że – jak to wynika z teorii Gelfanda (zob. np [4, 7, 36]) – niemal każda algebra Banacha jest „taka sama”, nieodróżnialna od pewnej podalgebry algebry $C(S)$, z odpowiednio dobranym S. Na przykład algebra splotowa l^1 (zob. zadanie 5.5) okazuje się być podalgebrą $C[-1, 1]$, algebry funkcji ciągłych na $[-1, 1]$, a $L^1(\mathbb{R}^+)$ – podalgebrą

[9] W literaturze przedmiotu mówi się zwykle o „oddzielaniu”, ale wydaje się, że to tradycyjne pojęcie jest nieco mniej adekwatne niż „rozróżnianie”.

algebry funkcji ciągłych na $[0, \infty]$ (a więc funkcji ciągłych na prawej półosi i z granicami w nieskończoności).

Skoro domknięcie podalgebry rozdzielającej punkty jest podalgebrą rozdzielającą punkty, twierdzenie Stone'a–Weierstrassa można równoważnie sformułować następująco.

Twierdzenie 7.3 (Stone'a–Weierstrassa)**.** *Każda rozróżniająca punkty S* domknięta *podalgebra $W \subset C(S)$, która zawiera funkcje stałe, jest równa całemu $C(S)$.*

7.2 Dowód

Przechodzimy do dowodu twierdzenia Stone'a–Weierstrassa, którego pomysłodawcą jest J. Zemánek (matematyk czeski, który lwią część życia spędził w Polsce, jako pracownik Instytutu Matematycznego PAN w Warszawie). Oczywiście dowodów jest cała masa – ten, który tu przedstawiamy, oparty na [37], szczególnie pasuje do niniejszej monografii, gdyż podkreśla rolę zupełności przestrzeni $C(S)$.

Lemat 7.1. *Załóżmy, że w przemiennej algebrze Banacha $\mathbb{X}$ istnieje element u, zwany jedynką, spełniający warunek*

$$ux = xu = x, \qquad \textit{dla wszystkich } x \in \mathbb{X}.$$

Niech też $\epsilon \in (0, 1)$ będzie daną liczbą, a x takim wektorem, że $\|x - u\| \le \epsilon$. Istnieje wtedy $y \in \mathbb{X}$ o tej własności, że

$$\|y\| \le \epsilon \qquad \textit{oraz} \qquad x = (u - y)^2.$$

Dowód. Domknięta kula $K \subset \mathbb{X}$ o promieniu ϵ, to znaczy zbiór tych wektorów z $\mathbb{X}$, których norma nie przekracza ϵ, jest zbiorem domkniętym, a zatem – jako podzbiór przestrzeni Banacha – przestrzenią metryczną zupełną. Rozważmy odwzorowanie $T : K \to K$ dane wzorem

$$Tz = \frac{1}{2}\left((u - x) + z^2\right);$$

(mamy $\|Tz\| \le \frac{1}{2}(\epsilon + \epsilon^2) \le \epsilon$). Poniższy rachunek pokazuje, że T jest kontrakcją (ze stałą $\epsilon < 1$):

$$\|Tz_1 - Tz_2\| = \left\|\frac{z_1^2 - z_2^2}{2}\right\| \le \left\|\frac{z_1 + z_2}{2}\right\| \|z_1 - z_2\| \le \epsilon \|z_1 - z_2\|.$$

(W drugim kroku korzystamy z zależności $z_1^2 - z_2^2 = (z_1 - z_2)(z_1 + z_2)$, a więc de facto z przemienności algebry $\mathbb{X}$). Z zasady Banacha wynika teraz istnienie punktu stałego odwzorowania T, to znaczy istnienie $y \in K$ o tej własności, że

$$2y = u - x + y^2.$$

Innymi słowy $x = u - 2y + y^2 = (u - y)^2$. To kończy dowód. $\square$

Lemat 7.2 (Twierdzenie Kakutaniego–Kreina)**.** *Niech $V \subset C(S)$ będzie odróżniającą punkty podprzestrzenią $C(S)$ (w sensie algebraicznym), zawierającą funkcję stale równą 1 (oznaczaną u). Jeśli V jest kratą, to jest gęsta w $C(S)$.*

Dowód. Przypomnijmy, że kratą nazywamy taki podzbiór $V \subset C(S)$, iż razem z $f, g \in V$ należą do niego funkcje $f \vee g$ i $f \wedge g$:

$$(f \vee g)(s) = \max(f(s), g(s)), \qquad (f \wedge g) = \min(f(s), g(s)).$$

Ustalmy $f \in C(S)$. Mając dane $s, t \in S$, tak wybierzmy $g \in V$ by $g(s) \neq g(t)$. Istnieje wtedy dokładnie jedna taka para liczb α, β, że $\alpha + \beta g(s) = f(s)$ i $\alpha + \beta g(t) = f(t)$ (bowiem wyznacznik powyższego układu równań jest różny od zera). Otrzymujemy zatem należącą do V funkcję $g_{s,t} = \alpha u + \beta g$, która w punktach s i t przyjmuje dokładnie te same wartości co f.

Wybierzmy teraz $\epsilon > 0$ oraz $s \in S$. Skoro funkcje $g_{s,t}$ i f są ciągłe, to możemy znaleźć takie otoczenie otwarte $U_{\epsilon,s}$ punktu s, że $g_{s,t}(r) < f(r) + \epsilon$ dla $r \in U_{\epsilon,s}$. Otoczenia te tworzą pokrycie S. Zwartość S powoduje, że możemy z nich wybrać podpokrycie skończone. Niech $t_1, ..., t_n$ będą punktami definiującymi to podpokrycie. Funkcja $g_s = g_{s,t_1} \wedge g_{s,t_2} \wedge ... \wedge g_{s,t_n}$ należy na podstawie założenia do V i spełnia warunek $g_s < f + \epsilon u$. Rzeczywiście, każde $r \in S$ jest elementem (przynajmniej) jednego z otoczeń $U_{t_i,\epsilon}$: dla odpowiedniego i mamy więc $g_s(r) \leq g_{s,t_i}(r) < f(r) + \epsilon$. Ciągle też $g_s(s) = f(s)$.

Niech teraz V_s będzie otwartym otoczeniem s o tej własności, że $g_s(r) > f(r) - \epsilon$ dla $r \in V_s$. Otoczenia $V_s, s \in S$ tworzą otwarte pokrycie S. Niech $s_1, ..., s_k$ będą punktami definiującymi jego skończone podpokrycie. Funkcja $h = g_{s_1} \vee g_{s_2} \vee ... \vee g_{s_n}$ należy do V i rozumowanie podobne do przedstawionego wyżej dowodzi, iż $h > f - \epsilon u$. Skoro oczywiście $h < f + \epsilon u$, wnioskujemy, że $\|f - h\| < \epsilon$. Ze względu na dowolność ϵ kończy to dowód. □

Dowód twierdzenia 7.3 *(Stone'a–Weierstrassa).* Niech V oznacza domknięcie podalgebry W. Naszym pierwszym celem jest uzasadnienie, że wraz z f z algebry V należy do niej także $|f|$. W tym celu rozważmy

$$f_n = \frac{1}{\|f^2\| + \frac{1}{n}} \left(f^2 + \frac{1}{n} \right).$$

Są to funkcje dodatnie o normach równych jeden i tej własności, że

$$f_n \geq \frac{\frac{1}{n}}{\|f^2\| + \frac{1}{n}} := \delta_n \in (0, 1).$$

Stąd ich odległość $\|u - f_n\|$ od u (funkcji stale równej 1) nie przekracza $\epsilon_n := 1 - \delta_n \in (0, 1)$. Skoro f_n są dodatnie, możemy rozważać $\sqrt{f_n}$, a lemat 7.1 przekonuje nas, że nowopowstałe funkcje są elementami V (korzystamy tu z faktu, że pierwiastkowanie w dziedzinie rzeczywistej jest działaniem jednoznacznym). Z drugiej strony, dla każdej liczby rzeczywistej x zachodzą zależności

$$\sqrt{x^2 + \frac{1}{n}} - |x| = \frac{\frac{1}{n}}{\sqrt{x^2 + \frac{1}{n}} + |x|} \leq \frac{\frac{1}{n}}{\sqrt{\frac{1}{n}}} = \frac{1}{\sqrt{n}}.$$

To dowodzi, że

$$\left\| \sqrt{f^2 + \frac{1}{n}} - |f| \right\| = \sup_{s \in S} \left| \sqrt{f^2(s) + \frac{1}{n}} - |f(s)| \right| \leq \frac{1}{\sqrt{n}},$$

a zatem $\left\| \sqrt{f_n} - \frac{|f|}{\sqrt{\|f^2\| + \frac{1}{n}}} \right\|$ dąży do zera, gdy $n \to \infty$. Domkniętość V zapewnia więc przynależność $\frac{1}{\sqrt{\|f^2\|}}|f|$, a tym samym i samego $|f|$ do tej algebry.

Zwróćmy teraz uwagę na łatwe do sprawdzenia wzory:

$$f \vee g = \frac{1}{2}(f + g + |f - g|) \qquad \text{oraz} \qquad f \wedge g = \frac{1}{2}(f + g - |f - g|), \tag{7.1}$$

które zachodzą dla dowolnych funkcji f i g. Dowodzą one (wobec tego, co udowodniliśmy powyżej i faktu, że V jest podprzestrzenią liniową $C(S)$), że wraz z $f, g \in V$ należą do V także $f \vee g$ i $f \wedge g$. A to oznacza, że V jest kratą. Z twierdzenia Kakutaniego–Kreina wiemy zatem, że V jest zbiorem gęstym w $C(S)$. Skoro jest też zbiorem domkniętym mamy $V = C(S)$. □

7.3 Zadania

Zadanie 7.1. Sprawdź, że domknięcie podalgebry jest podalgebrą.

Zadanie 7.2. Udowodnij wzory (7.1).

Zadanie 7.3. Niech $C[0, \infty]$ będzie przestrzenią funkcji ciągłych na prawej półosi rzeczywistej, które mają granice w nieskończoności. Dla $\lambda \geq 0$ niech też $e_\lambda(x) = \mathrm{e}^{-\lambda x}$, $x \geq 0$. Udowodnij, że kombinacje liniowe funkcji e_λ tworzą zbiór gęsty w $C[0, \infty]$.

Rozdział 8

Norma normie nie równa

8.1 Uzupełnianie przestrzeni

Co możemy począć, jeśli okaże się, że pracujemy w przestrzeni unormowanej, która jest „dziurawa", a chcielibyśmy skorzystać z dobrodziejstw przestrzeni zupełnej? To samo co starożytni, którzy liczby wymierne uzupełnili liczbami niewymiernymi: uzupełnić ją do przestrzeni Banacha.

Okazuje się bowiem, na nasze szczęście, że każdą przestrzeń unormowaną można uważać za podzbiór przestrzeni Banacha. Dokładniej, dla każdej przestrzeni unormowanej $\mathbb{X}$ istnieje przestrzeń Banacha $\mathbb{Y}$ oraz zawarta w niej podprzestrzeń (w sensie algebraicznym) $\mathbb{X}'$, które spełniają dwa warunki.

1. Po pierwsze, $\mathbb{X}$ jest „takie samo" jak $\mathbb{X}'$, co znaczy, że istnieje wzajemnie jednoznaczne przekształcenie liniowe L przestrzeni $\mathbb{X}$ na $\mathbb{X}'$ zachowujące normę:
$$\|Lx\|_{\mathbb{Y}} = \|x\|_{\mathbb{X}}; \tag{8.1}$$
liniowość znaczy tu tyle, że $L(\alpha x + \beta y) = \alpha(Lx) + \beta(Ly)$ dla dowolnych $x, y \in \mathbb{X}$ oraz $\alpha, \beta \in \mathbb{R}$ (o odwzorowaniach takich będziemy jeszcze wiele mówić w rozdziale 12). Zwróćmy uwagę, że znak $+$ po lewej stronie oznacza dodawanie w przestrzeni $\mathbb{X}$, a tenże znak po prawej – w przestrzeni $\mathbb{Y}$. Podobnie jest z normami w równości (8.1), co podkreślają indeksy.

2. Po drugie, $\mathbb{X}'$ jest zbiorem gęstym w $\mathbb{Y}$. To znaczy, że dla każdego $y \in \mathbb{Y}$ istnieje ciąg $(y_n)_{n\geq 1}$ elementów $\mathbb{X}'$, którego granicą jest y.

Co ciekawe, przestrzeń $\mathbb{Y}$ jest jednoznacznie wyznaczona (z dokładnością do izomorfizmu). To znaczy, że każde dwie przestrzenie o wymienionych wyżej własnościach są w pewnym sensie „takie same", nieodróżnialne.

Powyższego twierdzenia dowodzi się w sposób dość abstrakcyjny, choć naturalny, zob. np. [4]. Najpierw „zanurza się" przestrzeń $\mathbb{X}$ w przestrzeń b_c ciągów $(x_n)_{n\geq 1}$, w których $x_n \in \mathbb{X}$, spełniających warunek Cauchy'ego (zauważmy, że każdemu $x \in \mathbb{X}$ odpowiada pewien element b_c, który jest ciągiem stałym o wszystkich wyrazach równych x), a następnie dzieli się b_c przez relację równoważności, w której dwa ciągi, które są nieskończenie blisko siebie w nieskończoności należą do tej samej klasy. Nie będziemy powtarzać tu tego rozumowania, bo w praktyce, a przynajmniej w zadaniach, które napotyka początkujący adept analizy funkcjonalnej, $\mathbb{Y}$ zwykle okazuje się naturalną nadprzestrzenią $\mathbb{X}$. Stwierdzenie powyższe ilustrują przykłady podane niżej.

8.2 Dwa uzupełnienia przestrzeni c_{00}

W przestrzeni liniowej ciągów, które od pewnego (charakterystycznego dla danego ciągu, nie dla przestrzeni) miejsca równe są zero, można wprowadzić różne normy. Na przykład dla $x = (\xi_i)_{i\geq 1}$ możemy myśleć o

$$\|x\|_\infty = \sup_{i\geq 1} |\xi_i| \qquad \text{albo o} \qquad \|x\|_1 = \sum_{i=1}^{\infty} |\xi_i|,$$

(ostatnia suma składa się ze skończonej ilości składników). Jak wiemy z poprzedniego podrozdziału, przestrzeń c_{00}, która zupełna nie jest, można uzupełnić do przestrzeni Banacha. Chcemy podkreślić, że różne normy prowadzą jednak do różnych uzupełnień – wydaje się to jasne, gdy patrzymy na pierwszy warunek uzupełnienia, to znaczy zależność (8.1).

Twierdzimy, że uzupełnieniem c_{00} z pierwszą normą jest c_0. Odwzorowanie L, o którym mowa w definicji w tym przypadku przyporządkowuje ciągowi $x \in c_{00}$ ten sam ciąg, ale widziany jako element c_0. Odwzorowanie to jest oczywiście liniowe, a warunek (8.1) jest spełniony, bo normy w obu przestrzeniach są takie same. Pozostaje nam udowodnić, że dla każdego $x = (\xi_i)_{i\geq 1} \in c_0$ znaleźć można ciąg $(x_n)_{n\geq 1}$ elementów c_{00}, zbieżny do x. W tym celu rozważamy

$$x_n = (\xi_1, \xi_2, \cdots, \xi_n, 0, 0, 0, \dots). \tag{8.2}$$

Ciąg x_n należy do c_{00}, bo wszystkie jego elementy począwszy od $(n+1)$-szego miejsca są równe zero. Z drugiej strony $x - x_n$ to ciąg, którego pierwszych n współrzędnych równych jest zero, a pozostałe są takie same jak w x. Dowodzi to, że

$$\|x - x_n\|_\infty = \sup_{i\geq n+1} |\xi_i|.$$

Skoro ostatnie supremum dąży do zera (bo $x \in c_0$ jest ciągiem zbieżnym do zera), udowodniliśmy tezę.

Twierdzimy teraz, że uzupełnieniem c_{00} w normie $\|\cdot\|_1$ jest l^1. Dowód tego stwierdzenia przebiega analogicznie jak wyżej. Jeśli zbiór c_{00} wyposażymy w w.w. normę, to możemy go uznać za podprzestrzeń przestrzeni l^1, więc rolę L odgrywać będzie odwzorowanie identycznościowe. Jeśli natomiast $x \in l^1$, to każdy z ciągów (8.2) jest elementem c_{00} i

$$\|x - x_n\| = \sum_{i=n+1}^{\infty} |\xi_i|.$$

Występująca tu suma dąży do 0 gdy $n \to \infty$, bo l^1 składa się z szeregów absolutnie sumowalnych.

8.3 Dwa uzupełnienia przestrzeni wielomianów

Oto nieco bardziej zaawansowany, a równocześnie bardziej interesujący przykład. Jak przekonaliśmy się w rozdziale 7 (zob. też podrozdział 13.1), każdą funkcję ciągłą aproksymować można wielomianami z dowolną zadaną dokładnością (w normie supremum). Innymi słowy uzupełnieniem przestrzeni wielomianów, wyposażonej w normę supremum (zob. podrozdział 4.4.2) jest $C[0,1]$. Czytelnik zwróci zapewne uwagę na fakt, że powyższe zdanie pozostanie w mocy i w sumie nie zmieni znaczenia, jeśli słowo „uzupełnieniem" zamienimy na „domknięciem". To pokazuje, jak naturalna jest procedura, którą tu omawiamy.

Jeśli jednak wyposażymy przestrzeń wielomianów w normę

$$\|x\| = \int_0^1 |x(s)| \,\mathrm{d}s,$$

to jej uzupełnieniem nie będzie $C[0,1]$, lecz przestrzeń funkcji bezwzględnie całkowalnych w sensie Lebesgue'a na odcinku $[0,1]$.

8.4 Iloczyny i normy tensorowe

Iloczyn tensorowy

Dla $x = (\xi_i)_{i\geq 1} \in c_0$ i $y = (\eta_j)_{j\geq 1} \in l^1$ definiujemy iloczyn tensorowy

$$x \otimes y = (\xi_i \eta_j)_{i,j\geq 1}\,. \tag{8.3}$$

Iloczyn tensorowy jest więc macierzą nieskończoną. Gdyby dla przykładu oba wektory x, y były elementami l^1 i to takimi, że wszystkie ich współrzędne są nieujemne, a ich suma w obu wypadkach równa jest jeden, to można by je uważać za rozkłady pewnych zmiennych losowych:

$$\mathbb{P}(X = i) = \xi_i, \qquad \mathbb{P}(Y = j) = \eta_j,$$

a macierz $x \otimes y$ byłaby wtedy rozkładem łącznym pary (X, Y) pod warunkiem, że X i Y byłyby zmiennymi losowymi niezależnymi:

$$\xi_i \eta_j = \mathbb{P}(X = i, Y = j).$$

Macierze postaci (8.3) nazywamy tensorami prostymi, a ich kombinacje liniowe, tzn. macierze postaci

$$m = \sum_{k=1}^{l} x_k \otimes y_k \tag{8.4}$$

gdzie $x_k \in c_0$ i $y_k \in l_1$ są dowolnymi wektorami, nazywamy tensorami; wskaźnik sumowania l jest dowolną liczbą naturalną i zmienia się od tensora do tensora. Przestrzeń wszystkich tensorów nazywamy **iloczynem tensorowym** (przestrzeni c_0 i l^1) i oznaczamy

$$c_0 \otimes l^1.$$

Dla ustalenia terminologii przyjmijmy, że wszystkie elementy macierzy $(m_{i,j})_{i,j\geq 1}$ o tym samym indeksie i będziemy nazywać jej i-tym wierszem, a te o ustalonym indeksie j – jej j-tą kolumną.

Szczególne znaczenie mają tensory otrzymane z wektorów $e_k \in c_0 \cap l^1$, które na k-tym miejscu mają 1 a zera poza nim. Tymi szczególnymi tensorami są

$$e_{i,j} = e_i \otimes e_j, \qquad i,j \geq 1$$

(jak zinterpretować je probabilistycznie?). Zwróćmy uwagę na fakt, że przedstawienie (8.4) nie jest jednoznaczne. Na przykład

$$e_{2,2} = e_2 \otimes e_2, \text{ ale też } e_{2,2} = (e_1 + e_2) \otimes (e_1 + e_2) - e_1 \otimes e_2 - e_2 \otimes e_1 - e_1 \otimes e_1.$$

Norma tensorowa największa

Jaka jest „naturalna” norma w $c_0 \otimes l^1$? To zależy co to znaczy „naturalna”. Będziemy rozważać tylko normy spełniające warunek:

$$\|x \otimes y\| = \|x\|_{c_0} \|y\|_{l^1}. \tag{8.5}$$

Ze względu na nierówność trójkąta, przy każdym przedstawieniu (8.4) musimy mieć

$$\|m\| \leq \sum_{k=1}^{l} \|x_i\|_{c_0} \|y_i\|_{l^1}$$

więc „największym” naturalnym kandydatem na normę jest:

$$\|m\|_\pi = \inf \sum_{k=1}^{l} \|x_i\|_{c_0} \|y_i\|_{l^1}; \tag{8.6}$$

we wzorze tym infimum liczone jest po wszystkich reprezentacjach (8.4). Bezpośredni, choć długi rachunek dowodzi, iż $\|\cdot\|_\pi$ można zapisać też wzorem:

$$\|m\|_\pi = \sum_{k=1}^{\infty} \sup_{i\geq 1} |m_{i,k}|. \tag{8.7}$$

Uwaga terminologiczna. *Norma $\|\cdot\|_\pi$ wprowadzona wyżej nazywana jest w języku angielskim, w którym publikowana jest praktycznie cała istotna literatura matematyczna, albo „the projective norm” albo, rzadziej, „the largest norm”. O ile w języku tym pierwszy z tych terminów daje się sensownie uzasadnić w jednym zdaniu (ale po dowodach góry twierdzeń), o tyle w naszym ojczystym wymaga to* *cyrkowych zaiste akrobacji (ze względu na rozbieżności pojęciowe pozornie równoważnych słów). Można by go od biedy przetłumaczyć „norma rzutowa", choć mówi to przeciętnemu adeptowi matematyki z Polski niewiele. Napotykamy wtedy jednak na poważny problem z normą, którą omawiać będziemy za chwilę, nazywaną „the injective norm" albo, rzadziej, „the smallest norm". Otóż tu nie można pisać o „wstrzykniętej" bo nie jest to „injected" tylko jakiejś „wstrzyknięciowej", co po polsku nie znaczy już zupełnie nic. By uniknąć dziwactw spowodowanych bezrefleksyjnymi kalkami językowymi postanowiłem użyć terminu drugiego, rzadszego, ale łatwiejszego do przetłumaczenia na język polski: norma tensorowa największa i najmniejsza. Ta terminologia jest odbiciem faktu, że $\|\cdot\|_\pi$ jest największą z pewnej klasy norm, nazywanych tensorowymi (które między innymi spełniają warunek* (8.5)*), a ta druga, którą poznamy potem – najmniejszą. Jak widać, nazwy te można uzasadnić jednym zdaniem.*

Będzie to miało jednak niewygodną konsekwencję: trudno nam będzie teraz znaleźć tłumaczenie „the projective tensor product" (określenia na przestrzeń $c_0 \hat{\otimes}_\pi l^1$, zdefiniowaną nieco niżej) – nie może to być „największy iloczyn tensorowy" bo, jak zobaczymy, przestrzeń ta będzie właśnie najmniejsza z możliwych.

Od tego momentu traktować będziemy (8.7) jako definicję. Łatwo sprawdzić, że $\|\cdot\|_\pi$ jest normą. Okazuje się jednak, że $c_0 \otimes l^1$ z tak dobraną normą nie jest zupełna. Jej uzupełnieniem okazuje się być przestrzeń macierzy $m = (m_{i,j})_{i,j\geq 1}$, których każda kolumna jest elementem c_0, a dodatkowo

$$\sum_{j=1}^{\infty} \sup_{i\geq 1} |m_{i,j}| < \infty. \tag{8.8}$$

Tę ostatnią przestrzeń będziemy oznaczać

$$c_0 \hat{\otimes}_\pi l^1;$$

normą jest w niej oczywiście wyrażenie występujące po prawej stronie (8.8). Zwróćmy uwagę, że sumę występującą w (8.8) można napisać następująco:

$$\sum_{j=1}^{\infty} \|m_j\|_{c_0}, \tag{8.9}$$

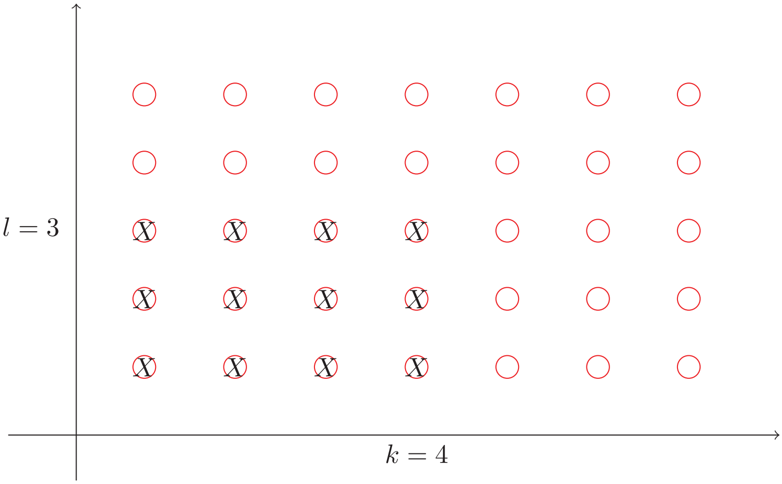
gdzie m_j oznacza j-tą kolumnę macierzy m.

By udowodnić, że $c_0 \hat{\otimes}_\pi l^1$ jest rzeczywiście uzupełnieniem $c_0 \otimes l^1$, ustalmy $m \in c_0 \hat{\otimes}_\pi l^1$. Dla dowolnego $\epsilon > 0$ możemy tak dobrać $k \in N$, by

$$\sum_{j=k+1}^{\infty} \sup_{i\geq 1} |m_{i,j}| < \frac{\epsilon}{2}.$$

Możemy ponadto tak dobrać l, by dla każdego $j \in \{1, ..., k\}$ zachodziła nierówność $\sup_{i>l} |m_{i,j}| \leq \frac{\epsilon}{2k}$. Zdefiniujmy $m_\epsilon = (m_{\epsilon,i,j})_{i,j\geq 1}$ następująco:

$$m_{\epsilon,i,j} = \begin{cases} m_{i,j}, & j \leq k \text{ oraz } i \leq l, \\ 0, & \text{w pozostałych przypadkach.} \end{cases} \tag{8.10}$$



Rys. 8.1: W miejscach oznaczonych krzyżykami macierz m_ϵ jest taka sama jak m, a w pozostałych równa jest zero. Suma supremów bezwzględnych wartości wszystkich wyrazów wchodzących w skład kolumn od $(k+1)$-szej włącznie (tu $k = 4$) jest mała. W każdej z k pierwszych kolumn małe są też suprema bezwzględnych wartości wyrazów macierzy m w miejscach nieoznaczonych krzyżykami.

Inaczej mówiąc, $m_\epsilon = \sum_{i=1}^{l} \sum_{j=1}^{k} m_{i,j} e_i \otimes e_j \in c_0 \otimes l^1$, zob. rysunek 8.1. Mamy zatem

$$\begin{aligned}\|m - m_\epsilon\|_\pi &= \sum_{j=1}^{\infty} \sup_{i\geq 1} |m_{i,j} - m_{\epsilon,i,j}| \\ &= \sum_{j=1}^{k} \sup_{i>l} |m_{i,j}| + \sum_{j=k+1}^{\infty} \sup_{i\geq 1} |m_{i,j}| \leq \sum_{j=1}^{k} \frac{\epsilon}{2k} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

To dowodzi naszej tezy.

Norma tensorowa najmniejsza

Czy zamieniając w (8.7) kolejność znaków $\sum$ i sup, otrzymamy ciekawą normę? Czy

$$\|m\|_\varepsilon = \sup_{i\geq 1} \sum_{j=1}^{\infty} |m_{i,j}| \tag{8.11}$$

jest w ogóle normą w $c_0 \otimes l^1$? Na szczęście, tak. A co będzie uzupełnieniem $c_0 \otimes l^1$ wyposażonej w tę normę?

Pokażemy, że jest nim przestrzeń tych macierzy $m = (m_{i,j})_{i,j \geq 1}$, dla których wektor $Im = \left(\sum_{j=1}^{\infty} |m_{i,j}|\right)_{i \geq 1}$ należy do c_0. Przestrzeń tę będziemy oznaczać

$$c_0 \hat{\otimes}_\varepsilon l^1.$$

Przede wszystkim, zob. zadanie 8.3, $c_0 \hat{\otimes}_\varepsilon l^1$ jest przestrzenią Banacha. Dowiedziemy teraz, że tensory tworzą w niej zbiór gęsty. Rozumować będziemy podobnie jak w podrozdziale poprzednim. Mając dane $m \in c_0 \hat{\otimes}_\varepsilon l^1$ i $\epsilon > 0$, wybieramy najpierw tak l, że $\sup_{i \geq l+1} \sum_{j=1}^{\infty} |m_{i,j}| < \epsilon$. Następnie tak dobieramy k, by $\sup_{1 \leq i \ \leq l} \sum_{j=k+1}^{\infty} |m_{i,j}| < \epsilon$. Wtedy dla m_ϵ zdefiniowanego wzorem (8.10) mamy

$$\begin{aligned}
\|m - m_\epsilon\|_\varepsilon &= \sup_{i \geq 1} \sum_{j=1}^{\infty} |m_{i,j} - m_{\epsilon,i,j}| \\
&= \max\left(\sup_{1 \leq i \ \leq l} \sum_{j=1}^{\infty} |m_{i,j} - m_{\epsilon,i,j}|, \sup_{i \geq l+1} \sum_{j=1}^{\infty} |m_{i,j} - m_{\epsilon,i,j}| \right) \\
&= \max\left(\sup_{1 \leq i \ \leq l} \sum_{j=k+1}^{\infty} |m_{i,j}|, \sup_{i \geq l+1} \sum_{j=1}^{\infty} |m_{i,j}| \right) < \epsilon.
\end{aligned}$$

To kończy dowód wspomnianej własności.

Podsumujmy nasze rozważania. W przestrzeni tensorów można wprowadzić (co najmniej) dwie naturalne normy: „największą” i „najmniejszą”. Domykając tę przestrzeń względem nich, otrzymamy dwie różne przestrzenie Banacha: $c_0 \hat{\otimes}_\pi l^1$ oraz $c_0 \hat{\otimes}_\varepsilon l^1$.

Jaka jest zależność między tymi przestrzeniami? Zwróćmy przede wszystkim uwagę na nierówność zachodzącą między w.w. normami:

$$\|m\|_\varepsilon \leq \|m\|_\pi. \tag{8.12}$$

Dla jej dowodu zauważmy, że dla każdego i $\sum_{j=1}^{\infty} |m_{i,j}| \leq \sum_{j=1}^{\infty} \sup_{i \geq 1} |m_{i,j}|$, a licząc supremum lewej strony, otrzymujemy tezę. Z nierówności (8.12) wynika, że każdy ciąg Cauchy'ego względem normy największej jest też ciągiem Cauchy'ego względem normy najmniejszej. Innymi słowy domknięcie przestrzeni tensorów w normie najmniejszej zawiera domknięcie przestrzeni tensorów w normie największej

$$c_0 \hat{\otimes}_\varepsilon l^1 \subset c_0 \hat{\otimes}_\pi l^1.$$

Czy zachodzi też inkluzja odwrotna? Nie. Na przykład macierz m, która ma na głównej przekątnej ciąg $(1, \frac{1}{2}, \frac{1}{3}, ...)$, a poza tym złożona jest z samych zer należy do $c_0 \otimes_\varepsilon l^1$, bo ciąg Im równy jest właśnie $(1, \frac{1}{2}, \frac{1}{3}, ...)$, ale nie należy do $c_0 \otimes_\pi l^1$, gdyż $\sum_{j=1}^{\infty} |m_{i,j}| = \sum_{j=1}^{\infty} \frac{1}{j} = \infty$. Obraz jest tu więc podobny do tego z rozdziałów 8.2 i 4.4.2: im mocniejsza norma, tym mniejsze uzupełnienie.

Główny morał z tej opowieści jest jednak taki, że

> domknięcie przestrzeni liniowej w sposób kluczowy zależy do tego, jak dobrana została w niej norma.

Systematyczny wykład iloczynów i norm tensorowych można znaleźć w książce [29] – to teoria pochodząca w lwiej części od A. Grothendiecka, urodzonego w Niemczech matematyka francuskiego, jednego z naukowych gigantów XX wieku.

8.5 Zadania

Zadanie 8.1. Sprawdź, że $\|\cdot\|_\pi$ zdefiniowana wzorem (8.7) jest normą i spełnia warunek (8.5).

Zadanie 8.2. Pokaż, że przestrzeń macierzy spełniających warunek (8.8) jest przestrzenią Banacha.

Zadanie 8.3. Sprawdź, że $\|\cdot\|_\varepsilon$ jest normą w $c_0 \otimes l^1$ i spełnia warunek (8.5).

Zadanie 8.4. Pokaż, że przestrzeń $c_0 \hat{\otimes}_\varepsilon l^1$ jest przestrzenią Banacha.

Zadanie 8.5. Wyposażmy c_{00} w normę

$$\|x\| = \|(\xi_i)_{i\geq 1}\| = \sum_{i=1}^{\infty} 2^n |\xi_i|,$$

(powyższa suma jest skończona, bo ma w istocie skończoną ilość składników). Scharakteryzuj uzupełnienie przestrzeni c_{00} z taką normą.

Zadanie 8.6. Powtórz poprzednie zadanie dla normy

$$\|x\| = \|(\xi_i)_{i\geq 1}\| = \sum_{i=1}^{\infty} |\xi_{2i}| + \sup_{i\geq 0} |\xi_{2i+1}|.$$

Rozdział 9

Przestrzenie Hilberta

Przestrzeń Hilberta to zupełna przestrzeń liniowa z normą pochodząca od iloczynu skalarnego. Jest to więc bardzo szczególny przykład przestrzeni Banacha, gdyż iloczyn skalarny narzuca normę wyjątkowo przypominającą tę z przestrzeni euklidesowej – geometria takich przestrzeni jest nam wyjątkowo bliska.

9.1 Przestrzeń unitarna

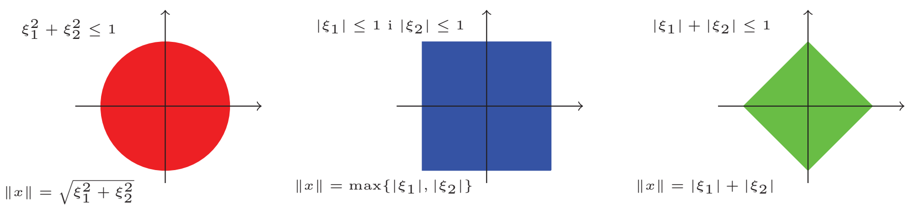
9.1.1 Definicja i podstawowe własności

Przestrzeń unitarna to przestrzeń liniowa $\mathbb{X}$, w której wprowadzono iloczyn skalarny, to znaczy funkcję $\mathbb{X} \times \mathbb{X} \to \mathbb{R}$ (będziemy tu myśleli tylko o przestrzeniach liniowych nad ciałem rzeczywistym, więc i iloczyn skalarny będzie miał wartości rzeczywiste) o następujących własnościach:

(s1) $(x + y, z) = (x, z) + (y, z)$,

(s2) $(\alpha x, y) = \alpha(x, y)$,

(s3) $(x, x) \geq 0$,



Rys. 9.1: Kule jednostkowe w przestrzeni $\mathbb{R}^2$ wyposażonej w różne normy: (a) euklidesową, tzn. pochodzącą od iloczynu skalarnego, (b) normę supremum i (c) normę typu l^1.

(s4) $(x, x) = 0$ wtedy i tylko wtedy, gdy $x = 0$,

(s5) $(x, y) = (y, x)$;

x, y i z są tu dowolnymi wektorami z $\mathbb{X}$, α dowolną liczbą rzeczywistą, (x, y) zaś jest wartością iloczynu skalarnego na parze złożonej z x i y.

Przykładem przestrzeni unitarnej jest przestrzeń $\mathbb{R}^k$ z iloczynem skalarnym danym wzorem

$$(x, y) = \sum_{i=1}^{k} \xi_i \eta_i; \tag{9.1}$$

sprawdzenie warunków (s1) – (s5) jest tu wyjątkowo proste.

Lemat 9.1 (Cauchy'ego-Schwarza)**.** *Dla dowolnych x i y z przestrzeni unitarnej zachodzi nierówność*

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$

Dowód. Ustalmy x i y i rozważmy funkcję

$$\mathbb{R} \ni t \mapsto (tx + y, tx + y) \in \mathbb{R}.$$

Wobec własności (s3) przyjmuje ona jedynie wartości nieujemne. Z drugiej strony rachunki oparte na (s1), (s2) i (s5) pozwalają nam stwierdzić, że

$$(tx + y, tx + y) = t^2(x, x) + 2t(x, y) + (y, y).$$

Jeśli $(x, x) = 0$ to wykresem tej funkcji jest prosta, ale ten przypadek może zajść tylko, gdy $x = 0$; wtedy zaś nierówność Cauchy'ego–Schwarza jest oczywista. Jeśli zaś $(x, x) > 0$ (a wobec (s3), nie jest możliwe, by $(x, x) < 0$), to wykresem jej jest parabola ramionami skierowana w górę. Warunkiem koniecznym i dostatecznym na to, by leżała ona nad osią iksów jest $\Delta \leq 0$, to znaczy

$$[(x, y)]^2 - (x, x)(y, y) \leq 0,$$

co jest równoważne naszej tezie. □

Na przykład, stosując tę nierówność w przestrzeni $\mathbb{R}^k$, otrzymujemy

$$\left|\sum_{i=1}^{k} \xi_i \eta_i\right| \leq \sqrt{\sum_{i=1}^{k} \xi_i^2}\sqrt{\sum_{i=1}^{k} \eta_i^2}. \tag{9.2}$$

Dzięki (9.2) możemy podać inny, ciekawszy przykład przestrzeni unitarnej. Jest nią przestrzeń l^2 ciągów sumowalnych z kwadratem, to znaczy takich $x = (\xi_i)_{i \geq 1}$, że

$$\sum_{i=1}^{\infty} \xi_i^2 < \infty.$$

Iloczynem skalarnym w tej przestrzeni jest

$$(x, y) = \sum_{i=1}^{\infty} \xi_i \eta_i. \tag{9.3}$$

Skąd wiemy, że występujący po prawej stronie równości szereg jest zbieżny? Z zależności (9.2). Dla dowolnych $x, y \in l^2$ mamy bowiem

$$\sum_{i=1}^{k} |\xi_i|\,|\eta_i| \leq \sqrt{\sum_{i=1}^{k} \xi_i^2}\,\sqrt{\sum_{i=1}^{k} \eta_i^2} \leq \sqrt{\sum_{i=1}^{\infty} \xi_i^2}\,\sqrt{\sum_{i=1}^{\infty} \eta_i^2} < \infty.$$

Z dowolności k wnioskujemy, że

$$\sum_{i=1}^{\infty} |\xi_i|\,|\eta_i| \leq \sqrt{\sum_{i=1}^{\infty} \xi_i^2}\,\sqrt{\sum_{i=1}^{\infty} \eta_i^2} < \infty,$$

więc szereg jest nawet zbieżny bezwzględnie. Z tą nierównością w chlebaku zostajemy łatwo generałami i dowodzimy, że funkcja zdefiniowana wzorem (9.3) rzeczywiście jest iloczynem skalarnym.

Bezpośrednim wnioskiem z nierówności Cauchy'ego–Schwarza jest to, że każda przestrzeń unitarna jest przestrzenią unormowaną. Faktycznie, definiując

$$\|x\| = \sqrt{(x, x)} \tag{9.4}$$

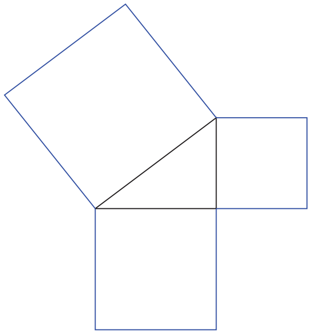
otrzymujemy funkcję o wszystkich własnościach normy. My sprawdzimy tu tylko nierówność trójkąta, pozostawiając Czytelnikowi przeliczenie własności pozostałych. Mamy zatem

$$\begin{aligned}
\|x + y\|^2 &= (x + y, x + y) = (x, x) + 2(x, y) + (y, y) \\
&\leq (x, x) + 2\|x\|\,\|y\| + (y, y) \\
&= \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\
&= \left(\|x\| + \|y\|\right)^2,
\end{aligned}$$

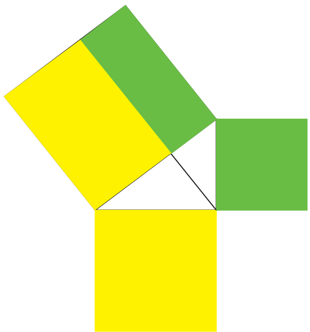
a to jest równoważne temu, że $\|x + y\| \leq \|x\| + \|y\|$.

9.1.2 Zastosowania

Niektórzy studenci widzieli definicję iloczynu skalarnego tyle razy, że dziś wydaje im się oczywista. Nie zapominajmy jednak, że czekaliśmy na jej sformułowanie dobrych kilka tysięcy lat od momentu, gdy ludzie zaczęli się interesować geometrią. Tak, rzeczywiście mam na myśli geometrię. Analiza funkcjonalna często traktowana jest jako uogólnienie tej starożytnej dziedziny nauki. Myślę, że wartość iloczynu skalarnego choć trochę wzrośnie w naszych oczach, gdy zobaczymy nowoczesny, „funkcjonalno-analityczny" dowód twierdzenia nazywanego imieniem Pitagorasa:



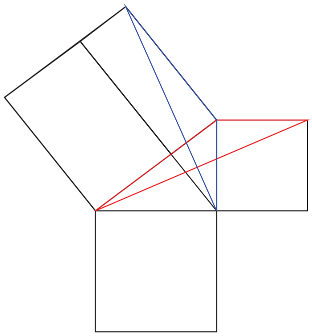
Rys. 9.2: Suma pól mniejszych kwadratów równa jest polu większego bo ...



Rys. 9.3: ... pola prostokątów zaznaczonych tym samym kolorem są równe; na przykład pola zielone są równe, bo ...

kwadrat długości przeciwprostokątnej w trójkącie prostokątnym równy jest sumie kwadratów przyprostokątnych. Dowodów tego wyniku jest wiele; „klasyczny" pochodzący od Euklidesa – przedstawiają rysunki 9.2-9.4. Są też rozumowania jeszcze starsze, zob. [16] str. 81, a wszystkie wymagają wielkiej spostrzegawczości i inteligencji. Tymczasem twierdzenia tego i to w znacznie ogólniejszej wersji może dowieść każdy, kto zna pojęcie iloczynu skalarnego – niekoniecznie inteligentny czy spostrzegawczy. W przestrzeni Hilberta (której przykładem jest oczywiście $\mathbb{R}^2$ z iloczynem (9.1)) sprowadza się ono bowiem do stwierdzenia, że jeśli dwa wektory, x i y, są do siebie prostopadłe, to znaczy jeśli $(x, y) = 0$, to

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2;$$



Rys. 9.4: ... zaznaczone trójkąty są przystające (obracając zgodnie z ruchem wskazówek zegara trójkąt czerwony wokół wspólnego wierzchołka z trójkątem niebieskim, nałożymy go na ten drugi), a ich pola równe połowom pól prostokątów.

x i y są tu oczywiście wektorami „przyprostokątnymi" a $x+y$ – „przeciwprostokątnym". Dowód jest rzeczywiście banalny:

$$x \perp y \implies \|x+y\|^2 = (x+y, x+y) = (x,x) + \underbrace{2(x,y)}_{=0} + (y,y)$$
$$= \|x\|^2 + \|y\|^2.$$

Zachęceni jego prostotą dowiedźmy też tak zwanej tożsamości równoległoboku, która stwierdza, że suma kwadratów długości przekątnych w równoległoboku równa jest podwojonej sumie kwadratów boków. Jej standardowy dowód polega na zastosowaniu twierdzenia Pitagorasa lub twierdzenia kosinusów. My dowiedziemy jej bezpośrednio i bezboleśnie (to znaczy: niewiele myśląc). Twierdzimy, że w dowolnej przestrzeni unitarnej $\mathbb{H}$ zachodzi równość

$$\|x+y\|^2 + \|x-y\|^2 = 2\big[\|x\|^2 + \|y\|^2\big], \tag{9.5}$$

w której $x, y \in \mathbb{H}$, a $\|\cdot\|$ jest normą unitarną (pochodzącą od iloczynu skalarnego). Dowód jest równie banalny jak poprzednio. Mamy:

$$\|x+y\|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y),$$
$$\|x-y\|^2 = (x-y, x-y) = (x,x) - 2(x,y) + (y,y),$$

a dodając stronami, otrzymujemy tezę.

Powie ktoś, że stare dowody są ładniejsze, bardziej godne uwagi. Tak, to prawda. Wejście na Rysy od strony polskiej to przeżycie estetyczne, fizyczne i duchowe oraz pewne osiągnięcie – wejście od strony słowackiej nie może się z nim pod tym względem równać. Ale nie zapominajmy o jednym: skoro maszyneria iloczynu skalarnego pozwala nam tak łatwo dowieść twierdzeń dobrze znanych, to jakichże twierdzeń możemy przy jej pomocy dowieść z pewnym trudem! Na piechotę możemy dojść na Rysy, ale samochód pomoże nam znacznie łatwiej wejść na nie od strony słowackiej; bez niego do Madrytu dotrzemy albo i nie.

9.2 Element minimalizujący odległość

Jak się Państwo domyślają, przestrzenie unitarne, które nie są zupełne, nie mogą być w tej książce doceniane. Przeciwnie, skupiać się będziemy na tych „bez dziur".

Oto kluczowa definicja:

Niech $\mathbb{H}$ będzie przestrzenią unitarną i niech $\|\cdot\|$ będzie unitarną normą, to znaczy normą daną wzorem (9.4). Jeśli para $(\mathbb{H}, \|\cdot\|)$ jest przestrzenią Banacha, to nazywamy ją przestrzenią Hilberta.

Bardzo często mówimy też, że samo $\mathbb{H}$ jest przestrzenią Hilberta – możemy tak jednak zrobić tylko wtedy, gdy z kontekstu jasne jest jaki iloczyn skalarny mamy na myśli.

Podstawowym przykładem przestrzeni Hilberta jest przestrzeń l^2 z podrozdziału 9.1.1 – opuszczamy dowód jej zupełności, gdyż jest on analogiczny do dowodu tej własności dla l^1. Przykładem zdecydowanie najważniejszym jest przestrzeń $L^2(S, \mathcal{F}, \mu)$ funkcji całkowalnych z kwadratem na przestrzeni z miarą $(S, \mathcal{F}, \mu)$ – nie będziemy go jednak omawiać dokładniej, bo znamy bolączki dzisiejszych studentów matematyki, którym oszczędza się nawet teorii całki Lebesgue'a. Stwierdzimy tylko, że przestrzeń l^2 jest szczególnym przykładem $L^2(S, \mathcal{F}, \mu)$ – to samo można powiedzieć o przestrzeni l^2_p z zadania 9.5.

Jak widzieliśmy na przykładzie twierdzenia Pitagorasa i tożsamości równoległoboku, geometria przestrzeni Hilberta jest nam bliska. Twierdzenie, które chcemy teraz omówić jest na to kolejnym dowodem.

Twierdzenie 9.1. *Niech C będzie niepustym, domkniętym i wypukłym podzbiorem przestrzeni Hilberta $\mathbb{H}$ i niech $x \notin C$. Istnieje wtedy dokładnie jeden element $y \in C$ o tej własności, że*

$$\|x - y\| = d := \inf_{z \in C} \|x - z\|.$$

Przypomnijmy, że zbiór C nazywamy wypukłym jeśli, wraz z dowolnymi jego elementami x i y, leży w nim cały odcinek o końcach x i y, to znaczy wszystkie punkty postaci $z = \alpha x + (1 - \alpha)y, \alpha \in [0, 1]$.

Dowód. Wobec tożsamości równoległoboku dla dowolnych $z, z' \in C$ mamy

$$\begin{aligned} \|z - z'\|^2 &= \|(z - x) + (x - z')\|^2 \\ &= 2\{\|z - x\|^2 + \|z' - x\|^2\} - \|z + z' - 2x\|^2 \\ &= 2\{\|z - x\|^2 + \|z' - x\|^2\} - 4\Big\|\frac{z + z'}{2} - x\Big\|^2 \\ &\le 2\{\|z - x\|^2 + \|z' - x\|^2\} - 4d^2, \end{aligned} \tag{9.6}$$

gdyż $\frac{z+z'}{2}$ należy do C. Z definicji d wiemy, że istnieje ciąg $(z_n)_{n \ge 1}$ elementów C o tej własności, że $\lim_{n\to\infty} \|z_n - x\| = d$. Pokażemy, że spełnia on warunek Cauchy'ego. W tym celu dla danego $\epsilon > 0$ tak wybieramy n_0, by dla $n \ge n_0$

wyrażenie $\|z_n - x\|^2$ było mniejsze niż $d^2 + \frac{\epsilon}{4}$. Używając (9.6), z $z = z_n$ i $z' = z_m$ przekonujemy się, że dla $n, m \geq n_0$, mamy $\|z_n - z_m\|^2 \leq \epsilon$ – to właśnie było naszym celem. Skoro $\mathbb{H}$ jest przestrzenią Hilberta, ciąg $(z_n)_{n\geq 1}$ ma granicę. Ale dla $y = \lim_{n\to\infty} z_n$, mamy $\|x - y\| = \lim_{n\to\infty} \|x - z_n\| = d$. Dowodzi to, że y o wskazanych warunkach istnieje: jest on granicą $(z_n)_{n\geq 1}$.

Musimy jeszcze dowieść, że takie y jest tylko jedno. Jeśli jednak $\|y' - x\| = d$, to (9.6) z $z = y$ i $z' = y'$ implikuje $\|y - y'\| = 0$, to znaczy $y' = y$. □

Uwagi, przykład i kontrprzykłady

Analogiczne twierdzenie w przestrzeni Banacha nie jest prawdziwe. Na przykład zbiór C wszystkich nieujemnych ciągów, sumujących się do 1 jest w l^1 zbiorem wypukłym i domkniętym. Ale dla dowolnego ujemnego ciągu x jego odległość od C równa jest $\|x\| + 1$ i realizowana jest w każdym elemencie C. Ważnym elementem założeń jest więc geometria ukryta w normie unitarnej – w innych geometriach twierdzenie zwykle prawdziwe nie jest.

Jak widzieliśmy w pierwszej części dowodu drugim, równie ważnym kluczem jest tu zupełność – bez niej nie bylibyśmy w stanie stwierdzić, że ciąg $(z_n)_{n\geq 1}$ jest zbieżny, a zatem i znaleźć y.

Popatrzmy na kolejny przykład – niech $\mathbb{H}$ będzie przestrzenią $\mathbb{R}^2$ a C niech będzie kołem jednostkowym $\{(\xi_1, \xi_2), \xi_1^2 + \xi_2^2 \leq 1\}$ z wyrzuconym punktem $(0, 1)$. Zbiór ten jest wypukły, ale nie domknięty i nie ma w nim punktu, który minimalizowałby odległość od $(0, 2)$. Tam, gdzie punkt ten powinien leżeć jest „dziura”.

Rozważmy też w tej samej przestrzeni zbiór

$$C = \{(\xi_1, \xi_2) | \xi_1 \in [-1, 1], \xi_2 = \pm\xi_1\},$$

który tworzy literę X umiejscowioną w środku układu współrzędnych. Jest on domknięty, ale nie wypukły. Dlatego też w zależności od tego jaki wybierzemy $x \notin C$, punkt minimalizujący odległość od C może być jeden albo dwa, może też ich być nieskończenie wiele.

Na koniec przyjrzyjmy się prostemu przykładowi „pozytywnemu”, w którym jawnie da się znaleźć element minimalizujący odległość. Niech $\mathbb{H}$ będzie dowolną przestrzenią Hilberta (a nawet tylko unitarną) i niech C będzie (domkniętą) kulą jednostkową, to znaczy zbiorem tych $y \in \mathbb{H}$, dla których $\|y\| \leq 1$. Niech też $x \notin C$, a tym samym $\|x\| > 1$. Który element $y \in C$ leży w najmniejszej odległości od x? Intuicja podpowiada nam, że jest nim $y = \frac{x}{\|x\|}$. Sprawdźmy. Dla dowolnego $y \in C$ mamy

$$\begin{aligned}\|x - y\|^2 &= \|x\|^2 - 2(x, y) + \|y\|^2 \geq \|x\|^2 - 2\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| - \|y\|)^2 \geq (\|x\| - 1)^2.\end{aligned}$$

Z drugiej strony, wstawiając $y = \frac{x}{\|x\|}$, zamiast wszystkich nierówności, otrzymujemy równość. Wektor ten leży więc rzeczywiście najbliżej x.

9.3 Rzut na podprzestrzeń

Szczególnie interesującym przypadkiem twierdzenia z poprzedniego rozdziału jest sytuacja, gdy C jest (domkniętą) podprzestrzenią liniową; będziemy ją odtąd oznaczać $\mathbb{H}_1$.

Twierdzenie 9.2. *Dla każdego x w $\mathbb{H}$ istnieje, i jest wyznaczony jednoznacznie, taki wektor $Px \in \mathbb{H}_1$, że dla dowolnego $z \in \mathbb{H}_1$ mamy $(x - Px, z) = 0$. Px minimalizuje odległość x od $\mathbb{H}_1$, a nazywany jest rzutem x na $\mathbb{H}_1$.*

Dowód. Jeśli x należy do $\mathbb{H}_1$, to przyjmujemy $Px = x$; bardzo prosto sprawdza się, że to jedyny możliwy wybór (gdyby x nie było równe Px to wybierając $z = x - Px$ otrzymalibyśmy sprzeczność z definicją Px).

Załóżmy więc, że $x \notin \mathbb{H}_1$ i niech $Px = y$ będzie elementem minimalizującym odległość od $\mathbb{H}_1$. Weźmy z z $\mathbb{H}_1$. Funkcja

$$f(t) = \|x - y + tz\|^2$$

zmiennej rzeczywistej t osiąga swoje minimum w punkcie $t_{\min} = 0$, bo $y - tz$ należy do $\mathbb{H}_1$, a y ma odległość x od $\mathbb{H}_1$ minimalizować. Z drugiej strony wykonywane już dwukrotnie rachunki prowadzą do wzoru

$$f(t) = \|x - y\|^2 + 2t(z, x - y) + t^2\|z\|,$$

który dowodzi, że $f(t)$ jest funkcją kwadratową i że $t_{\min} = -\frac{2(z,x-y)}{\|x-y\|^2}$. A zatem $(x - y, z) = 0$. Innymi słowy, tak wybrane Px spełnia warunki podane w twierdzeniu.

Musimy jeszcze pokazać, że Px nie można wybrać inaczej. Załóżmy zatem, że dla pewnego $y' \in \mathbb{H}_1$ mamy $(x - y', z) = 0$ o ile tylko $z \in \mathbb{H}_1$. Gdyby $y \neq y'$, to zależność

$$\|x - y\|^2 = \|x - y'\|^2 + 2(x - y', y' - y) + \|y' - y\|^2 = \|x - y'\|^2 + \|y' - y\|^2,$$

wynikająca z tego, że $y' - y$ należy do $\mathbb{H}_1$, dowodziłaby, iż $\|x - y\| > \|x - y'\|$. A to oczywiście przeczyłoby definicji y. Nie ma zatem innej opcji: Px musi minimalizować odległość. □

Dla przykładu rozważmy podprzestrzeń $\mathbb{H}_1$ przestrzeni l^2, do której należą te ciągi $(\eta_i)_{i\geq 1}$, które spełniają warunek

$$\eta_{2i} = \eta_{2i-1}, \qquad i \geq 1. \tag{9.7}$$

(Pozostawiamy Czytelnikowi dowód, że jest to w istocie (domknięta) podprzestrzeń $\mathbb{H}$). Mając dane $x = (\xi_i)_{i\geq 1} \notin \mathbb{H}_1$, szukamy w tej podprzestrzeni Px. Możemy to zrobić na (co najmniej) dwa sposoby. Pomyślmy najpierw o kwadracie odległości między x a y należącym do $\mathbb{H}_1$:

$$\|x - y\|^2 = \sum_{j=1}^{\infty}(\xi_j - \eta_j)^2 = \sum_{i=1}^{\infty} \left[(\xi_{2i} - \eta_{2i})^2 + (\xi_{2i-1} - \eta_{2i})^2\right]^2;$$

w drugim kroku sąsiednie współrzędne połączyliśmy w pary i skorzystaliśmy z warunku definiującego $\mathbb{H}_1$. Jak dobrać y, by wyrażenie to było najmniejsze? Oczywiście wystarczy sprawić, by każdy ze składników był możliwie najmniejszy (zwróćmy uwagę, że elementy η_{2i} możemy wybierać dowolnie i niezależnie od siebie!). W każdym składniku liczby ξ_{2i} i ξ_{2i-1} są dane. Rozwiązujemy więc problem typu: mając dane stałe a i b, znajdź u, dla którego wyrażenie

$$(a-u)^2 + (b-u)^2$$

jest najmniejsze. Tego uczono nas w liceum: rozwiązaniem jest $u = \frac{a+b}{2}$. Wracając zatem do naszego głównego zadania, przekonujemy się, że rzutem x na $\mathbb{H}_1$ jest

$$Px = \frac{1}{2}(\xi_1+\xi_2, \xi_1+\xi_2, \xi_3+\xi_4, \xi_3+\xi_4, \xi_5+\xi_6, \dots).$$

Zróbmy teraz to zadanie inaczej, korzystając z twierdzenia 9.2. Zacznijmy od tego, że wektor

$$z_1 = (1,1,0,0,0,\dots)$$

(który począwszy od trzeciego miejsca ma same zera) jest oczywiście elementem $\mathbb{H}_1$. Dla szukanego $Px = (\eta_i)_{i\geq 1}$ mamy więc

$$0 = (x - Px, z_1) = (\xi_1 - \eta_1)\cdot 1 + (\xi_2 - \eta_2)\cdot 1.$$

Skoro $\eta_2 = \eta_1$ (bo $Px \in \mathbb{H}_1$), dowodzi to, że

$$\eta_1 = \eta_2 = \frac{1}{2}(\xi_1 + \xi_2).$$

To samo rozumowanie dla wektora $z_i, i\ \geq 2$ który poza współrzędnymi $2i-1$ i $2i$ równymi 1 wszystkie pozostałe ma równe zero dowodzi, że

$$\eta_{2i-1} = \eta_{2i} = \frac{1}{2}(\xi_{2i} + \xi_{2i-1}).$$

A to oczywiście jest ten sam wynik co poprzednio.

Oto przykład znacznie ciekawszy. Niech X i Y będą zmiennymi losowymi o wartościach naturalnych, zdefiniowanymi na tej samej przestrzeni prawdopodobieństwa. Załóżmy ponadto (dla ustalenia uwagi), że

$$\Pr(X=i, Y=j) = p_{i,j} > 0, \qquad i,j \geq 1.$$

Załóżmy wreszcie, że istnieją drugie momenty zmiennych X i Y, tzn

$$\sum_{i=1}^{\infty} i^2 \Pr(X=i) < \infty \qquad \text{oraz} \qquad \sum_{i=1}^{\infty} i^2 \Pr(Y=i) < \infty. \tag{9.8}$$

Przestrzeń $\mathbb{H}$ macierzy $x = (\xi_{i,j})_{i,j\geq 1}$, dla których suma

$$\sum_{i,j=1}^{\infty} \xi_{i,j}^2 p_{i,j} \tag{9.9}$$

jest skończona, jest przestrzenią Hilberta z iloczynem skalarnym

$$(x, y) = \sum_{i,j=1}^{\infty} \xi_{i,j} \eta_{i,j} p_{i,j}$$

(por. zadania 9.5 – 9.7), a suma występująca w (9.9) jest unitarną normą macierzy x.

Zmienne losowe X i Y możemy traktować jako macierze: pierwsza z nich jest macierzą $x = (\xi_{i,j})_{i,j\geq 1}$ daną wzorem $\xi_{i,j} = i$, a druga – macierzą $y = (\eta_{i,j})_{i,j\geq 1}$, dla której $\eta_{i,j} = j$. Założenie (9.8) oznacza teraz, że

$$\sum_{i,j=1}^{\infty} \xi_{i,j}^2 p_{i,j} = \sum_{i=1}^{\infty} i^2 \sum_{j=1}^{\infty} p_{i,j} = \sum_{i=1}^{\infty} i^2 \Pr(X = i) < \infty,$$

co świadczy, że x jest elementem $\mathbb{H}$. Podobny rachunek dowodzi, iż także $y \in \mathbb{H}$.

Po tych przygotowaniach rozważmy podprzestrzeń $\mathbb{H}_1$ złożoną z tych macierzy $z = (\zeta_{i,j})_{i,j\geq 1}$, których wartości nie zależą od i. Mówiąc inaczej, wartości tych macierzy w każdej kolumnie są stałe (w konwencji, którą przyjęliśmy w rozdziale 8.4). Łatwo sprawdza się, że jest to rzeczywiście podprzestrzeń domknięta $\mathbb{H}$, oczywiście też $y \in \mathbb{H}_1$. Naszym zadaniem jest znalezienie rzutu x na $\mathbb{H}_1$. Macierz x oczywiście do tej przestrzeni nie należy; w jej przypadku – wręcz przeciwnie – $\xi_{i,j}$ zależy tylko i wyłącznie od i.

W tym celu rozważmy macierz z_k, w której k-ta kolumna złożona jest z samych jedynek, a pozostałe z samych zer. Jest to element $\mathbb{H}$, bo $\sum_{i=1}^{\infty} p_{i,k} = P(Y = k) < \infty$; jest to też oczywiście element $\mathbb{H}_1$. Jeśli więc $Px = (\delta_{i,j})_{i,j\geq 1}$, to musimy mieć

$$0 = (x - Px, z_k) = \sum_{i=1}^{\infty} (\xi_{i,k} - \delta_{i,k}) p_{i,k} = \sum_{i=1}^{\infty} (\xi_{i,k} - \delta_{i,k}) p_{i,k}.$$

Uwzględniając fakt, że $\delta_{i,k}$ nie zależy od i, a $\xi_{i,j} = i$ otrzymujemy stąd

$$\delta_{i,k} = \delta_k = \frac{\sum_{j=1}^{\infty} j p_{j,k}}{\sum_{j=1}^{\infty} p_{j,k}}.$$

Wyrażenie po prawej stronie jest nam znane z rachunku prawdopodobieństwa: pokazaliśmy, że δ_k jest wartością oczekiwaną zmiennej losowej X pod warunkiem $Y = k$.

Czyż to nie piękne, że abstrakcyjna wydawałoby się procedura prowadzi do dającego się naturalnie zinterpretować wyniku? Przedstawiona tu w szczególnym przypadku idea ma daleko idące konsekwencje w teorii rachunku prawdopodobieństwa; pozwala między innymi na zdefiniowanie tak zwanej warunkowej wartości oczekiwanej zmiennej losowej względem innej zmiennej losowej (która niekoniecznie musi przyjmować wartości naturalne) a także względem σ-ciała (zob. np. [4, 14]; pierwszym, który tę piękną ideę przedstawił był, jak się wydaje, A.N. Kołmogorow).

9.4 Zadania

Zadanie 9.1. Udowodnij, że wzór (9.4) definiuje normę.

Zadanie 9.2. Udowodnij, że w każdej przestrzeni unitarnej zachodzi następująca równość, nazywana czasem tożsamością polaryzacyjną:

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

Zadanie 9.3. Sprawdź, że l^2 jest przestrzenią unitarną, to znaczy, że wzór (9.3) definiuje iloczyn skalarny.

Zadanie 9.4. Ustalmy k naturalne i niech $p_1, \ldots, p_k$ będą liczbami dodatnimi. Sprawdź, że wzór

$$\sum_{i=1}^{k} \xi_i \eta_i p_i$$

definiuje iloczyn skalarny w $\mathbb{R}^k$. Jaką figurą geometryczną będzie kula jednostkowa odpowiadająca temu iloczynowi skalarnemu w przypadkach $k = 1, 2, 3$?

Zadanie 9.5. Niech $p = (p_i)_{i \geq 1}$ będzie ciągiem liczb dodatnich. Rozważmy przestrzeń l^2_p tych ciągów $(\xi_i)_{i \geq 1}$, dla których suma

$$\sum_{i=1}^{\infty} \xi_i^2 p_i$$

jest skończona. Korzystając z zadania poprzedniego, udowodnij, że wzór

$$(x, y) = \sum_{i=1}^{\infty} \xi_i \eta_i p_i$$

definiuje iloczyn skalarny w l^2_p, a powyższa suma jest normą wektora x.

Zadanie 9.6. Załóżmy, że w warunkach zadania poprzedniego spełniony jest dodatkowo warunek $\sum_{i=1}^{\infty} p_i = 1$. Dwa elementy l^2_p, powiedzmy x i y, możemy wtedy utożsamiać ze zmiennymi losowymi, powiedzmy X i Y, zdefiniowanymi na przestrzeni probabilistycznej $\Omega = \mathbb{N}$ z funkcją prawdopodobieństwa zadaną wzorem $\Pr(\{i\}) = p_i$. Uzasadnij, że te zmienne losowe posiadają drugi moment. Udowodnij też, że

$$(x, y) = E\,XY.$$

Zadanie 9.7. Udowodnij, że l^2 i l^2_p (zob. zadanie 9.5) są przestrzeniami Hilberta.

Zadanie 9.8. Niech $\mathbb{H}$ będzie przestrzenią Hilberta, a $\mathbb{H}_1$ – jej podprzestrzenią domkniętą.

(a) Udowodnij, że zbiór $\mathbb{H}_1^{\perp}$ złożony z wektorów prostopadłych do wszystkich wektorów należących do $\mathbb{H}_1$ jest domkniętą podprzestrzenią $\mathbb{H}$.

(b) Co będzie przestrzenią $\mathbb{H}_1^{\perp}$ dla $\mathbb{H}_1 \subset l^2$ zdefiniowanego wzorem (9.7)?

(c) Niech x należy do $\mathbb{H}$. Udowodnij, Px jest rzutem na $\mathbb{H}_1$ wtedy i tylko wtedy, gdy $x - Px$ jest rzutem na $\mathbb{H}_1^{\perp}$. Wykorzystaj ten wynik, by znaleźć rzut na podprzestrzeń $\mathbb{H}_1^{\perp}$ z punktu (b).

Zadanie 9.9. Niech $\mathbb{H}_1$ będzie podprzestrzenią l^2 złożoną z tych $(\eta_i)_{i\geq 1}$, dla których

$$\eta_1 = 2\eta_2 \qquad \text{oraz} \qquad \eta_3 = 0.$$

Udowodnij, że jest to podprzestrzeń domknięta. Sprawdź też, że rzutem $(\xi_i)_{i\geq 1}$ na $\mathbb{H}_1$ jest $y = (\eta_i)_{i\geq 1}$ zdefiniowany wzorem

$$\eta_i = \begin{cases} \frac{4\xi_1 + 2\xi_2}{5}, & i = 1, \\ \frac{4\xi_1 + 2\xi_2}{10}, & i = 2, \\ 0, & i = 3, \\ \xi_i, & i \geq 4. \end{cases}$$

Zadanie 9.10. Niech $\mathbb{H}_1 \subset L^2[0,1]$ składa się z tych funkcji całkowalnych z kwadratem na $[0,1]$, które na przedziale $[0, \frac{1}{2}]$ są stałe. Udowodnij, że jest to podprzestrzeń domknięta i znajdź wzór na rzut prostopadły na nią.

Rozdział 10

Układy ortonormalne zupełne

Jednym z najważniejszych przykładów rzutów prostopadłych jest rzut na podprzestrzeń rozpinaną przez układ ortonormalny. Ten temat z kolei wiedzie w naturalny sposób do twierdzenia mówiącego, że pozornie różne przestrzenie $L^2[a,b]$ i l^2 są w istocie nieodróżnialne. Z takim ekwipunkiem zaś w kolejnym rozdziale bez większego trudu rozwiążemy równanie ciepła.

10.1 Układy ortonormalne

Definicja 10.1. *Mówimy, że ciąg $(e_n)_{n\geq 1}$ elementów przestrzeni Hilberta $\mathbb{H}$ jest ortonormalny, gdy*

$$(e_n, e_m) = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Zamiennie mówimy też o układzie ortonormalnym $(e_n)_{n\geq 1}$.

Dla przykładu, niech

$$e_n = (0, 0, \dots, 0, \underbrace{1}_{n\text{-te miejsce}}, 0, \dots) \in l^2, \qquad n \geq 1. \tag{10.1}$$

Łatwo sprawdzamy, że jest to układ ortonormalny. Ciekawszy jest ciąg $(e_n)_{n\geq 1}$ wektorów przestrzeni $L^2[-\pi, \pi]$, zdefiniowany wzorem

$$e_n(x) = \frac{1}{\sqrt{\pi}} \sin nx, \qquad x \in [-\pi, \pi], n \geq 1. \tag{10.2}$$

Skoro w przestrzeni tej $(f, g) = \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x$, to

$$(e_n, e_m) = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin nx \sin mx \,\mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\cos(n-m)x - \cos(n+m)x] \,\mathrm{d}x.$$

Ta ostatnia całka równa jest zero jeśli $n \neq m$. Jeśli zaś $m = n$, równa jest ona 2π, a całe wyrażenie równe jest 1. To dowodzi ortonormalności ciągu $(e_n)_{n\geq 1}$.

Niech $(e_n)_{n\geq 1}$ będzie ustalonym układem ortonormalnym. Ustalmy $m \geq 1$ i rozważmy przestrzeń $\mathbb{H}_1$ wektorów postaci

$$y = \sum_{n=1}^{m} \alpha_n e_n,$$

to znaczy kombinacji liniowych $e_1, \dots, e_m$ – mówimy często, że jest to przestrzeń rozpinana przez $e_1, \dots, e_m$. Jak będzie wyglądał rzut Px wektora $x \in \mathbb{H}$ na $\mathbb{H}_1$? Skoro $x - Px$ ma być prostopadłe to wszystkich wektorów z $\mathbb{H}_1$, to w szczególności musi być prostopadłe do każdego $e_n, n = 1, \dots, m$. Jeśli więc $Px = \sum_{i=1}^m \alpha_i e_i$, to musimy mieć

$$0 = (x - Px, e_n) = (x, e_n) - \alpha_n;$$

w ostatnim kroku wykorzystaliśmy orgonormalność $(e_n)_{n\geq 1}$. Innymi słowy

$$Px = \sum_{n=1}^{m} (x, e_n) e_n;$$

wzór elegancki i godny zapamiętania.

Zwróćmy uwagę, że dla rzutu Px wektora x na dowolną podprzestrzeń, $x - Px$ jest prostopadły do Px. Stąd, wobec twierdzenia Pitagorasa,

$$\|x\|^2 = \|x - Px\|^2 + \|Px\|^2;$$

to zaś pociąga za sobą

$$\|Px\|^2 \leq \|x\|^2.$$

To samo twierdzenie zastosowane $(m-1)$-krotnie (lub bezpośredni rachunek) dowodzi, że

$$\left\| \sum_{n=1}^{m} (x, e_m) e_n \right\|^2 = \sum_{n=1}^{m} [(x, e_n)]^2.$$

Łącząc te fakty, otrzymujemy poniższy, ważny lemat.

Lemat 10.1 (Nierówność Parsevala[10])**.** *Jeśli $(e_n)_{n\geq 1}$ jest układem ortonormalnym, to dla każdego $x \in \mathbb{H}$*

$$\sum_{n=1}^{\infty} [(x, e_n)]^2 \leq \|x\|^2.$$

Dowód. Dowód w zasadzie został już zrobiony; pokazaliśmy bowiem, że dla każdego $m \geq 1$ zachodzi nierówność

$$\sum_{n=1}^{m} [(x, e_n)]^2 \leq \|x\|^2.$$

To zaś jest tylko innym sformułowaniem nierówności Parsevala. □

[10] Marc-Antoine Parseval des Chênes (1755–1836), nie mylić z Parsifalem, synem Pellinore'a, rycerzem Okrągłego Stołu. Nierówność tę nazywa się też często nierównością Bessela.

Nasuwa się naturalne pytanie, czy nierówność Paresevala można zastąpić równością. Odpowiedź brzmi: czasem tak, czasem nie. W tym pierwszym przypadku można myśleć, że wektorów $e_n, n \geq 1$ jest wystarczająco dużo, że innymi słowy tworzą one układ pełny, kompletny, po prostu – zupełny. To słowo już znamy w innym kontekście, kontekście dla tej książeczki podstawowym. Choć pojęcia zupełności przestrzeni i zupełności układu ortonormalnego są formalnie różne, opisują podobną sytuację – tę, w której przestrzeni i układowi „nic nie brakuje”.

Definicja 10.2. *Układ ortonormalny w przestrzeni $\mathbb{H}$ nazywamy zupełnym, jeśli dla każdego $x \in \mathbb{H}$ nierówność Parsevala zastąpić można równością. Wzór*

$$\|x\|^2 = \sum_{n=1}^{\infty} [(x, e_n)]^2.$$

nazywamy wtedy tożsamością Parsevala.

Przejdźmy do przykładów pokazujących, że nie każdy układ – nawet nieskończony – jest zupełny. Pomyślmy o przestrzeni l^2 i o układzie $(e'_n)_{n\geq 1}$ gdzie $e'_n = e_{3n}$ a $(e_n)_{n\geq 1}$ jest zdefiniowany wzorem (10.1). Jest to oczywiście układ ortonormalny, bo jest złożony z elementów większego układu ortonormalnego. Dla tego układu i dowolnego $x = (\xi_i)_{i\geq 1}$ nierówność Parsevala przyjmuje postać:

$$\sum_{n=1}^{\infty} \xi_{3n}^2 \leq \sum_{n=1}^{\infty} \xi_n^2$$

i oczywiste jest, że znaku nierówności na znak równości zamienić w ogólności nie można. Równość zajść może bowiem tylko dla tych x, które mają wszystkie wyrazy, poza być może tymi, które mają indeksy podzielne przez 3, równe zero. Bardzo podobne rozumowanie pokazuje, że $(e_n)_{n\geq 1}$ jest zupełny – ciąg $(e'_n)_{n\geq 1}$ zupełny nie jest, bo powstał przez odrzucenie części wyrazów $(e_n)_{n\geq 1}$.

Oto przykład mniej oczywisty. Rozważmy przestrzeń $L^2[-\pi, \pi]$ oraz układ ortonormalny ze wzoru (10.2). Czy jest to układ zupełny? Jeśli $f(x) = 1, x \in [-\pi, \pi]$, to dla każdego n

$$(f, e_n) = 0,$$

bo sin jest funkcją nieparzystą. Lewa strona nierówności Parsevala równa jest więc zero, a prawa wynosi $\int_{-\pi}^{\pi} 1 \, dx = 2\pi$. Równość zatem nie zachodzi. To już dowodzi, że układ zupełny nie jest. Swoją drogą, powyższe rozumowanie pozostaje w mocy jeśli za f wybierzemy dowolną, niezerową funkcję parzystą. W.w. układ można uzupełnić, dopisując do niego funkcje kosinus:

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \cos 2x, \ldots;$$

(inna stała w pierwszej funkcji nie jest pomyłką). Fundamentem całej teorii szeregów Fouriera jest to, że tak uzupełniony układ funkcji jest zupełny. W szczególności dla każdej funkcji f całkowalnej z kwadratem zachodzi tożsamość Parsevala:

$$\left(\int_{-\pi}^{\pi}[f(x)]^2\,\mathrm{d}x\right)^2 = \frac{1}{\pi}\sum_{n=1}^{\infty}\left(\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x\right)^2 + \frac{1}{2\pi}\left(\int_{-\pi}^{\pi}f(x)\,\mathrm{d}x\right)^2 + \frac{1}{\pi}\sum_{n=1}^{\infty}\left(\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x\right)^2.$$

Po lewej stronie mamy w niej kwadrat normy funkcji f, po prawej – sumę kwadratów jej współczynników Fouriera, będących iloczynami skalarnymi (f, e_n).

Zupełności w.w. układu, a zatem i wspomnianej tożsamości dowodzić nie będziemy. Odpowiednie rozumowanie znaleźć można w wielu książkach, na przykład w [6].

Następnym krokiem jest przeformułowanie definicji układu zupełnego.

Twierdzenie 10.1. *Niech $(e_n)_{n\geq 1}$ będzie układem ortonormalnym w $\mathbb{H}$. Następujące warunki są równoważne.*

(a) $(e_n)_{n\geq 1}$ *jest układem zupełnym.*

(b) *Dla każdego* $x \in \mathbb{H}$

$$x = \sum_{n=1}^{\infty}(x, e_n)e_n \left(:= \lim_{m\to\infty}\sum_{n=1}^{m}(x, e_n)e_n \right).$$

(c) *Jeśli dla ustalonego* x *i każdego* n *mamy* $(x, e_n) = 0$, *to* $x = 0$.

Dowód. Pokażmy najpierw, że z (a) wynika (c). W tym celu rozważmy x, który wymnożony skalarnie z dowolnym elementem układu zupełnego daje zero. Wobec tożsamości Parsevala normą x jest zero, a to znaczy, że sam x równy jest zero.

Dla dowodu, że (c) pociąga (b), zauważmy najpierw, że niezależnie od tego, czy układ jest zupełny czy nie, szereg

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

jest zbieżny. Istotnie, jego sumy częściowe $s_m = \sum_{n=1}^{m}(x, e_n)e_n, m \geq 1$ spełniają warunek

$$\|s_m - s_k\|^2 = \|\sum_{n=k+1}^{m}(x, e_n)\|^2 = \sum_{n=k+1}^{m}[(x, e_n)]^2, \qquad k < m.$$

Skoro, wobec nierówności Parsevala, szereg $\sum_{n=1}^{\infty}[(x, e_n)]^2$ jest zbieżny, to wspomniane sumy częściowe tworzą ciąg Cauchy'ego. Stąd teza.

Rozważmy teraz

$$y := x - \sum_{n=1}^{\infty}(x, e_n)e_n.$$

Dla dowolnego $n \geq 1$ mamy

$$(y, e_n) = \lim_{m\to\infty}(x - s_m, e_n) = (x, e_n) - \lim_{m\to\infty}(s_m, e_n) = (x, e_n) - (x, e_n) = 0.$$

Z założenia (b) mamy zatem $y = 0$, to znaczy $x = \sum_{n=1}^{\infty}(x, e_n)e_n$.

Pozostał nam dowód implikacji (b) $\Rightarrow$ (a). W tym celu przypomnijmy, że dla dowolnego ciagu $(x_n)_{n\geq 1}$ z równości $x = \lim_{n\to\infty} x_n$ wynika (wobec nierówności trójkąta) $\|x\| = \lim_{n\to\infty}\|x_n\|$, a więc także $\|x\|^2 = \lim_{n\to\infty}\|x_n\|^2$. Przy założeniu (b), warunek (a) jest bezpośrednim skutkiem powyższego faktu. □

10.2 Izomorfizm l^2 i L^2

Zatrzymajmy się. Wbiegliśmy na ten szczyt tak szybko, że nie zauważyliśmy widoku, który się przed nami odsłonił. A udowodniliśmy coś godnego najwyższej uwagi. Wynik zupełnie nieoczekiwany. Jeszcze Państwo nie widzą?

> Jeśli w przestrzeni Hilberta istnieje ortonormalny ciąg zupełny, to przestrzeń ta jest nieodróżnialna od l^2.

Weźmy do ręki lornetkę. Jeśli $(e_n)_{n\geq 1}$ jest ciągiem zupełnym w $\mathbb{H}$, to każdemu $x \in \mathbb{H}$ możemy przyporządkować ciąg złożony z liczb $\xi_i = (x, e_i)$, a nierówność Parsevala dowodzi, że $(\xi_i)_{i\geq 1}$ jest wtedy elementem l^2. Przyporządkowanie

$$\mathbb{H} \ni x \mapsto Ix = (\xi_i)_{i\geq 1} \in l^2$$

jest liniowe (tzn $I(\alpha x + \beta y) = \alpha Ix + \beta Iy$), więc struktura liniowa w l^2 jest odpowiednikiem struktury liniowej w $\mathbb{H}$. Jest ono też „na”, bo mając dany ciąg $(\xi_i)_{i\geq 1}$ w l^2, możemy zdefiniować x w l^2 wzorem (zob. dowód twierdzenia 10.1)

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

a wtedy $Ix = (\xi_i)_{i\geq 1}$. Jeśli dodatkowo $(e_n)_{n\geq 1}$ jest zupełny, to nierówność Parsevala zamienia się w jego tożsamość, którą możemy napisać teraz tak

$$\|x\|^2_{\mathbb{H}} = \|Ix\|^2_{l^2},$$

a to sprawia, że normy w obu przestrzeniach są „nieodróżnialne”. W szczególności odwzorowanie I jest różnowartościowe: $Ix = 0$ implikuje $\|Ix\|_{l^2} = 0$, co daje $\|x\|_{\mathbb{H}} = 0$ i w konsekwencji $x = 0$.

Odwzorowania, które mają takie własności jak opisane wyżej I, nazywamy izometrycznymi izomorfizmami przestrzeni Hilberta (Banacha). Formalnie więc rzecz biorąc, twierdzenie w ramce brzmi następująco:[11]

[11]Założenie istnienia układu ortonormalnego można w nim jeszcze osłabić i zastąpić ośrodkowością. Zob. np. [4].

> Jeśli w przestrzeni Hilberta istnieje ortonormalny układ zupełny, to przestrzeń ta jest izometrycznie izomorficzna z l^2.

Twierdzenie to jeszcze nie zrobiło na Państwu wrażenia? To napiszmy bezpośrednio z niego wynikający wniosek (wobec faktu, że ciąg cosinusów i sinusów jest zupełny w $L^2[-\pi, \pi]$).

> Przestrzeń $L^2[-\pi, \pi]$ jest izometrycznie izomorficzna z l^2.

Co? Naprawdę?

10.3 Zadania

Zadanie 10.1. Funkcję $f(x) = x, x \in [-\pi, \pi]$ rozwiń w szereg Fouriera. Sprawdź, że wstawiając $x = \pi$ w tym rozwinięciu, nie otrzymasz $f(\pi)$ lecz 0. (Przekonujemy się tym samym, że zbieżność w $L^2[-\pi, \pi]$ nie oznacza zbieżności punktowej). Napisz tożsamość Parsevala dla tej f.

Zadanie 10.2. Ważny przypadek układów ortonormalnych stanowią te związane z wielomianami ortogonalnymi [18], obejmującymi między innymi wielomiany Legendre'a, Hermite'a i Laguerre'a. W tym zadaniu omawiamy (po krótce) tylko te drugie. Definiuje się je wzorem:

$$H_n(x) = (-1)^n e^{x^2} \left(e^{-x^2}\right)^{(n)}, \qquad n \geq 0,$$

w którym (n) oznacza oczywiście n-tą pochodną po x.

1. Sprawdź bezpośrednim rachunkiem, że
$$H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2 \quad \text{oraz} \quad H_3(x) = 8x^3 - 12x.$$

2. Udowodnij przy pomocy indukcji, że
$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}, \qquad n \geq 1 \tag{10.3}$$
i użyj tej rekurencji do znalezienia jawnego wzoru na kilka kolejnych wielomianów ($n \geq 4$).

3. Przekonaj się teraz, że
$$H_n'(x) = 2nH_{n-1},$$
oraz (wobec (10.3))
$$H_{n+1}(x) = 2xH_n(x) - H_n'. \qquad n \geq 1 \tag{10.4}$$

4. Wywnioskuj, że H_n jest rozwiązaniem (a ściślej: jednym z rozwiązań) równania
$$u'' - 2xu' + 2nu = 0. \tag{10.5}$$

5. Niech $L^2(\mathrm{e}^{-x^2/2})$ oznacza przestrzeń funkcji, które pomnożone przez wagę $x \mapsto w(x) = \mathrm{e}^{-x^2/2}$ stają się całkowalne z kwadratem. Przekonaj się (podobnie jak zrobiliśmy to w rozdziale 5 z algebrą $L^1_\omega(\mathbb{R}^+)$), że jest to przestrzeń Hilberta nieodróżnialna od $L^2(\mathbb{R})$. Jaki jest wzór na iloczyn skalarny w tej przestrzeni?

6. Sprawdź, że wielomiany Hermite'a (jako wielomiany właśnie) są elementami $L^2(\mathrm{e}^{-x^2/2})$ i tworzą w niej ciąg ortogonalny. Wskazówka. Wykorzystaj punkt czwarty dla dowodu, że jeśli $G_n = w\, H_n$, to dla każdych m i n
$$G_n'' + (2n + 1 - 2x)G_n = 0 \qquad \text{i} \qquad G_m'' + (2m + 1 - 2x)G_m = 0.$$
Mnożąc pierwsze równanie przez G_m, a drugie przez G_n, przekonaj się, że
$$(G_n' G_m - G_m' G_n)' + 2(n - m)G_n G_m = 0,$$
a potem obie strony tej równości scałkuj po $\mathbb{R}$.

Rozdział 11

Równanie ciepła

11.1 Wygodny układ zupełny w $L^2[0,\pi]$

To, że w zaskakującym twierdzeniu z poprzedniego rozdziału przedział ma długość 2π, nie jest istotne. Wszystkie przestrzenie $L^2[a,b]$ $(a < b)$ są izometrycznie izomorficzne, a zatem izometrycznie izomorficzne z l^2 (zob. zadanie 11.2). W szczególności dotyczy to także przestrzeni $L^2[0,\pi]$, choć wydaje się ona być tylko połową $L^2[-\pi,\pi]$. Wykorzystamy ten fakt do rozwiązania równania ciepła, ale najpierw znajdziemy w niej wygodny układ zupełny (zupełnych układów jest tam wiele).

Pokażemy mianowicie, że odpowiednie własności ma

$$e_n(x) = \sqrt{\frac{2}{\pi}} \sin nx, \qquad x \in [0,\pi], n \geq 1. \tag{11.1}$$

(Sprawdzenie, że jest to układ ortonormalny pozostawiamy Czytelnikowi). Zwróćmy uwagę na to, że wcześniej pokazaliśmy, iż podobny (inaczej tylko skalowany) układ nie jest zupełny w $L^2[-\pi,\pi]$. A zaraz pokażemy, że w $L^2[0,\pi]$ jest. (A co, nie mówiłem, że $L^2[0,\pi]$ jest tylko połową $L^2[-\pi,\pi]$? W tej ostatniej przestrzeni oprócz sinusów trzeba mieć jeszcze kosinusy.)

Użyjemy w tym celu twierdzenia 10.1. Załóżmy, że dla $f \in L^2[0,\pi]$ wszystkie iloczyny skalarne (f, e_n) są równe zero. Niech $\tilde{f}$ będzie nieparzystym rozszerzeniem f na $[-\pi,\pi]$, to znaczy przyjmijmy, że $\tilde{f}(-x) = -f(x), x \in (0,\pi]$. Wobec parzystości funkcji kosinus mamy wtedy

$$\int_{-\pi}^{\pi} \tilde{f}(x) \cos nx \, \mathrm{d}x = 0, \qquad n \geq 0.$$

Nieparzystość sin pociąga za sobą z kolei parzystość funkcji $x \mapsto \tilde{f}(x) \sin nx, n \geq 1$ i równość

$$\int_{-\pi}^{\pi} \tilde{f}(x) \sin nx \, \mathrm{d}x = 2 \int_{0}^{\pi} f(x) \sin nx \, \mathrm{d}x = 0$$

(to ostatnie to z założenia). A zatem zupełność w $L^2[-\pi,\pi]$ układu złożonego z kosinusów i sinusów dowodzi, że $\tilde{f}=0$, a w szczególności $f=0$ (prawie wszędzie). To z kolei przekonuje nas, że

układ (11.1) jest zupełny.

W następnych podrozdziałach wykorzystamy tę informację do rozwiązania równania ciepła.

11.2 Postawienie problemu

Wyobraźmy sobie stosunkowo krótki i bardzo, bardzo cienki drucik – nie myślimy o kawałku drewna, czy plastiku, bo chcemy by nasz obiekt dobrze przewodził ciepło. Wyobraźmy sobie również, że w chwili $t=0$ znamy rozkład temperatury wzdłuż drutu, a chcielibyśmy przewidzieć, jak będzie on wyglądał w przyszłości. Zakładamy dodatkowo, że temperatura na końcach jest stała i dla ustalenia uwagi równa zero. Jak się okazuje (zob. np. piękny artykuł [11]) z matematycznego punktu widzenia prowadzi to do następującego zagadnienia: mając daną funkcję $f_0:[0,\pi]\to\mathbb{R}$, modelującą rozkład początkowy (jak widać przyjęliśmy, że standardowy drucik ma długość π – w istocie jednak założenie to nie ma większego znaczenia, przyjęcie innej długości prowadziłoby jedynie do nieco bardziej skomplikowanych wzorów), znajdź taką funkcję dwóch zmiennych $u:\mathbb{R}^+\times[0,\pi]$, rozkład w czasie i przestrzeni, że

(a) zachodzi następujaca równość, zwana równaniem ciepła[12]

$$\frac{\partial u(t,x)}{\partial t}=\frac{\partial^2 u(t,x)}{\partial x^2},\qquad x\in[0,\pi], t\geq 0, \tag{11.2}$$

(na końcach przedziału pochodne po x liczymy jednostronnie),

(b) spełnione są warunki brzegowe (Dirichleta):

$$u(t,0)=u(t,\pi)=0,\qquad t\geq 0,$$

(c) oraz warunek początkowy:

$$u\ (0,x)=f_0(x),\qquad x\in[0,\pi].$$

Zagadnienie to można rozwiązać na wiele sposobów – całe działy matematyki rozwinęły się w dużej mierze dlatego, że chciano lepiej zrozumieć równanie ciepła – jedno z najważniejszych równań fizyki matematycznej [9, 25, 33, 34]. My za chwilę przeformułujemy je tak, by wykorzystać izomorfizm przestrzeni l^2 i $L^2[0,\pi]$, ale na razie musimy przejść krótki kurs różniczkowania elementów przestrzeni $L^2[0,\pi]$.

[12] Trochę oszukujemy. Po prawej stronie tej równości powinna występować ważna fizycznie stała (przenikliwość cieplna drutu). Ale jesteśmy matematykami – dla nas stała to coś, co możemy usunąć z równania, nieco zmieniając zmienne.

11.3 Uogólnione różniczkowanie, przestrzenie Sobolewa

Podstawowe twierdzenie rachunku różniczkowego i całkowego mówi, że jeśli funkcja f zdefiniowana na odcinku otwartym (a, b) jest w nim różniczkowalna w sposób ciągły, to można ją „odzyskać” z pochodnej, o tak:

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(y)\,\mathrm{d}y, \tag{11.3}$$

dla dowolnych x_0 i x z w.w. przedziału. Funkcja $f(x) = |x|$, rozważana powiedzmy w przedziale $(-1, 1)$ nie jest różniczkowalna w $x = 0$, więc powyższe twierdzenie jej nie obejmuje. A jednak mamy:

$$f(x) = f(x_0) + \int_{x_0}^{x} g(y)\,\mathrm{d}y \tag{11.4}$$

dla funkcji

$$g(y) = \operatorname{sgn} y = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases}$$

Oczywiście nie możemy twierdzić, że g jest „prawdziwą” pochodną. Widzimy jednak, że odgrywa rolę analogiczną do niej (tej, której nie ma).

Definicja 11.1. *Jeśli istnieje x_0 i taka funkcja całkowalna g, że*

$$f(x) = f(x_0) + \int_{x_0}^{x} g(y)\,\mathrm{d}y, \tag{11.5}$$

dla x z pewnego przedziału (a, b), to mówimy, że f jest absolutnie ciągła w tym przedziale. Jeśli wzór powyższy można rozciągnąć także na a i b, to mówimy, że f jest absolutnie ciągła w $[a, b]$.

Dla g spełniających warunek (11.5) będziemy pisali

$$g = f', \tag{11.6}$$

pamiętając o tym, że jest to pochodna w uogólnionym sensie. Zwróćmy uwagę na to, że zmiana wartości g w jednym czy kilku miejscach (czy też wręcz na zbiorze miary zero) nie zmienia jej kluczowej własności (11.5). Innymi słowy, warunek (11.6) wyznacza g tylko prawie wszędzie.

Co ciekawe, z różniczkowalnych w uogólnionym sensie funkcji można budować przestrzenie Hilberta (i Banacha); są one objęte nazwą „przestrzenie Sobolewa” – okazują się one być fundamentalne dla równań różniczkowych cząstkowych (zob. np. [23]). Na przykład przestrzenią Hilberta jest $H_0^1(a, b]$ złożona z zerujących się w $x = a$ funkcji absolutnie ciągłych z pochodną w $L^2[a, b]$. Dość łatwo się sprawdza (zob. zadanie 11.4), że wzór

$$(f_1, f_2) = \int_a^b f_1' f_2' \tag{11.7}$$

definiuje iloczyn skalarny w $H_0^1(a, b]$. Całka ta ma sens, bo zarówno f_1' jak i f_2' należą do $L^2[a, b]$. W istocie wzór ten moglibyśmy napisać tak (odróżniając iloczyn skalarny w $H_0^1(a, b]$ od tegoż w $L^2[a, b]$ przez indeksy):

$$(f_1, f_2)_{H_0^1(a,b]} = (f_1', f_2')_{L^2[a,b]}.$$

W szczególności zachodzi równość:

$$\|f\|_{H_0^1(a,b]} = \|f'\|_{L^2[a,b]}. \tag{11.8}$$

W sumie też łatwy jest dowód zupełności $H_0^1(a, b]$. Załóżmy, że ciąg $(f_n)_{n\geq 1}$ spełnia warunek Cauchy'ego w tej przestrzeni. Z definicji normy w $H_0^1(a, b]$ widzimy, że $(f_n')_{n\geq 1}$ jest ciągiem podstawowym w $L^2[a, b]$ – ma więc granicę, powiedzmy g. Funkcja h dana wzorem $h(x) = \int_a^x g(y)\,\mathrm{d}y, x \in [a, b]$, jest elementem $H_0^1(a, b]$ i naturalnym kandydatem na granicę $(f_n)_{n\geq 1}$. By dowieść, że jest on rzeczywiście tą granicą sprawdzamy, że

$$\|f_n - h\|_{H_0^1(a,b]} = \|f_n' - h'\|_{L^2[a,b]} = \|f_n' - g\|_{L^2[a,b]},$$

co jak wiemy dąży do zera.

Zakończmy uwagą, że wobec tego, iż dla każdej f z tej przestrzeni mamy $f(x) = \int_a^x f'(y)\,\mathrm{d}y, x \in [a, b]$, prawdziwe są oszacowania

$$\begin{aligned} |f(x)| &\leq \int_a^x |f'(y)|\,\mathrm{d}y \leq \int_a^b |f'(y)| \cdot 1\,\mathrm{d}y \leq \sqrt{\int_a^b |f'(y)|^2\,\mathrm{d}y}\sqrt{\int_a^b 1^2\,\mathrm{d}x} \\ &= \sqrt{b-a}\|f\|_{H_0^1(a,b]} \end{aligned}$$

(trzeci krok powyżej to nierówność Cauchy'ego–Schwarza w $L^2[a, b]$). To pociąga następującą nierówność typu Sobolewa [1, 23]:

$$\|f\|_{C[a,b]} = \sup_{x\in[a,b]} |f(x)| \leq \sqrt{b-a}\|f\|_{H_0^1(a,b]}.$$

Bezpośrednim z niej wnioskiem jest to, że ciągi zbieżne w przestrzeni $H_0^1(a, b]$ są też zbieżne jednostajnie – w przestrzeniach typu $L^2[a, b]$ tak nie jest, zbieżność w nich nie pociąga nawet zbieżności punktowej.

11.4 Definicja rozwiązania

Sprecyzujmy teraz, w jakim sensie chcemy rozwiązać równanie ciepła. Przede wszystkim zamiast myśleć o nim jako o równaniu cząstkowym, potraktujemy je jak równanie różniczkowe zwyczajne w przestrzeni Banacha $L^2[0,\pi]$. Będziemy więc rozważali

$$u'(t) = Au(t), \qquad u(0) = f_0, \tag{11.9}$$

w którym f_0 jest daną funkcją z $L^2[0,\pi]$, a A jest operatorem dwukrotnego różniczkowania po x. Dokładniej: A jest operatorem (przekształceniem przyporządkowującym funkcjom funkcje), którego dziedzina $D(A)$ składa się z różniczkowalnych (w zwykłym sensie) funkcji f, których pierwsze pochodne są absolutnie ciągłe, a (uogólnione) pochodne f'' z tych pochodnych należą do $L^2[0,\pi]$; o funkcjach tych zakładamy dodatkowo, że $f(0) = f(\pi) = 0$. Jak widać warunek brzegowy dla równania ciepła ukryliśmy w dziedzinie operatora A. Oczywiście też, z definicji, $Af = f''$.

Nie szukamy zatem funkcji rzeczywistej dwóch zmiennych $t \geq$ i $x \in [0,\pi]$, lecz raczej funkcji jednej zmiennej $t \geq 0$, która ma wartości w przestrzeni $L^2[0,\pi]$. Mówiąc precyzyjniej, szukamy $t \mapsto u(t) \in L^2[0,\pi]$ o następujących własnościach.

(a) Dla każdego $t > 0$ $u(t)$ należy do $D(A)$ (inaczej prawa strona równania (11.9) nie miałaby sensu).

(b) Funkcja u jest w sposób ciągły różniczkowalna w każdym punkcie $t > 0$ w sensie normy z przestrzeni $L^2[0,\pi]$: istnieje taki element $u'(t) \in L^2[0,\pi]$, że

$$\lim_{h\to 0} \left\| \frac{u(t+h) - u(t)}{h} - u'(t) \right\| = 0, \tag{11.10}$$

a funkcja $t \mapsto u'(t)$ jest ciągła (znów w sensie normy w $L^2[0,\pi]$).

(c) Zachodzi równość $\lim_{t\to 0+} u(t) = f_0$ – to jest oczywiście odpowiednik warunku początkowego.

(d) Dla każdego $t > 0$, zachodzi równość $u'(t) = Au(t)$.

Uważnemu Czytelnikowi rzuci się zapewne w oczy fakt, że mimo formalnego podobieństwa równań (11.9) i (11.2) są one w istocie różne. Przypomnijmy w szczególności, że zbieżność w przestrzeni $L^2[0,\pi]$ ciągu funkcji $(f_n)_{n\geq 1}$ nie pociąga za sobą jego zbieżności punktowej i odwrotnie, zbieżność punktowa nie pociąga zbieżności w sensie $L^2[0,\pi]$ (zob. zadania 10.1 i 11.5). Pochodnej, o której mowa w punkcie (b) nie można więc a priori utożsamiać z pochodną cząstkową $\frac{\partial u}{\partial t}$. To jeden z matematycznych cudów, że – jak zobaczymy później – rozwiązując (11.9), otrzymamy rozwiązanie (11.2).

Uwagi powyższe dowodzą też, że przestrzeń $L^2[0,\pi]$ jest dość niekomfortowa. Z lenistwa, a może dla wygody, przejedziemy więc do przestrzeni l^2, która dla naszych celów będzie znacznie milsza.

11.5 Rozwiązanie: jak to zrobić łatwiej

Przypomnijmy, że – jak to już pokazaliśmy na początku tego rozdziału – $L^2[0,\pi]$ jest nieodróżnialna (izometrycznie izomorficzna) od l^2. Jeśli funkcji f z $L^2[0,\pi]$

przyporządkujemy ciąg $(a_n)_{n\geq 1}$ jej współczynników Fouriera:

$$a_n = (f, e_n) = \sqrt{\frac{2}{\pi}} \int_0^\pi f(x) \sin nx \,\mathrm{d}x,$$

to otrzymamy element l^2. I odwrotnie, mając dany ciąg $(a_n)_{n\geq 1}$ z l^2, możemy mu przypisać

$$f = \sum_{n=1}^{\infty} a_n e_n$$

gdzie $e_n(x) = \sqrt{\frac{2}{\pi}} \sin nx$. Co więcej, norma f w $L^2[0, \pi]$ jest równa normie $(a_n)_{n\geq 1}$ w l^2. Skoro l^2 jest nam bardziej przyjazna, dobrze byłoby przetransponować (11.9) na równanie w tej przestrzeni.

W tym celu wprowadzamy w l^2 operator B dany wzorem

$$B(a_n)_{n\geq 1} = -\left(n^2 a_n\right)_{n\geq 1},$$

którego dziedziną są wszystkie te ciągi $(a_n)_{n\geq 1}$, dla których $\left(n^2 a_n\right)_{n\geq 1}$ jest elementem l^2 (dla przykładu, dowolny ciąg, którego prawie wszystkie wyrazy są zerami należy do $D(B)$, podobnie jest z $\left(\frac{1}{n^4}\right)_{n\geq 1}$, ale już $\left(\frac{1}{n^2}\right)_{n\geq 1}$ elementem dziedziny nie jest). Ten operator wygląda milej, znacznie łatwiej też się go definiuje niż A, a jednak okazuje się, że są one w sprecyzowanym niżej sensie identyczne.

Lemat 11.1. *B jest odpowiednikiem A w l^2. Innymi słowy, oznaczając przez I odwzorowanie przyporządkowujące funkcji f ciąg jej współczynników Fouriera, mamy*

$$f \in D(A) \textit{ wtedy i tylko wtedy, gdy } If \in D(B)$$

i zachodzi równość (zob. rysunek 11.1)

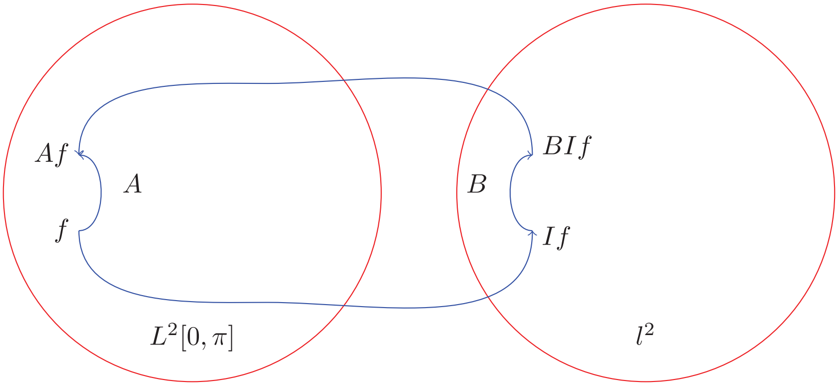
$$Af = I^{-1} BI\, f.$$

Dowód. Każda funkcja $f \in D(A)$ jest postaci

$$f(x) = \int_0^x (c + \int_0^y g(z) \,\mathrm{d}z\ ) \,\mathrm{d}y;$$

wykorzystaliśmy tu fakt, że $f(0) = 0$. Funkcja g jest tu drugą pochodną f i z założenia należy do $L^2[0, \pi]$, c zaś pewną stałą ($= f'(0)$). Oznaczając przez a_n i b_n współczynniki Fouriera funkcji f i g i dwukrotnie całkując przez części (wykorzystując $f(0) = f(\pi) = 0$), otrzymujemy:

$$\begin{aligned}
a_n \sqrt{\frac{\pi}{2}} &= \int_0^\pi f(x) \sin nx \,\mathrm{d}x = \int_0^\pi \int_0^x (c + \int_0^y g(z) \,\mathrm{d}z\ ) \,\mathrm{d}y \sin nx \,\mathrm{d}x \\
&= \int_0^\pi (c + \int_0^x g(z) \,\mathrm{d}z) \frac{\cos nx}{n} \,\mathrm{d}x = -\int_0^\pi g(x) \frac{\sin nx}{n^2} \,\mathrm{d}x \\
&= -\frac{1}{n^2} b_n \sqrt{\frac{\pi}{2}}.
\end{aligned}$$



Rys. 11.1: W l^2 obrazem działającego w $L^2[0,\pi]$ operatora A jest B.

To oznacza, że $a_n = -\frac{1}{n^2}b_n$, a zatem $\left(n^2 a_n\right)_{n\geq 1} = -\left(b_n\right)_{n\geq 1}$ należy do l^2, bo z założenia własność tę ma $\left(b_n\right)_{n\geq 1}$ (jako ciąg współczynników Fouriera g). Mamy też $If = (a_n)_{n\geq 1}$, $BIf = -\left(n^2 a_n\right)_{n\geq 1} = (b_n)_{n\geq 1}$ oraz $I^{-1}BIf = I^{-1}\left(b_n\right)_{n\geq 1} = g$, co dowodzi drugiej zależności przedstawionej w lemacie.

Musimy jeszcze tylko udowodnić, że jeśli $If \in D(B)$, to $f \in D(A)$ (na razie udowodniliśmy implikację odwrotną). Jeśli zatem $If = (a_n)_{n\geq 1}$ ma tę własność, że $\left(n^2 a_n\right)_{n\geq 1}$ należy do l^2, to możemy rozważyć $g = -I^{-1}\left(n^2 a_n\right)_{n\geq 1}$ oraz $\tilde{f} \in L^2[0,\pi]$ dane wzorem

$$\tilde{f}(x) = \int_0^x (c + \int_0^y g(z)\,\mathrm{d}z)\,\mathrm{d}y;$$

w którym $c = -\frac{1}{\pi}\int_0^\pi \int_0^y g(z)\,\mathrm{d}z\ \mathrm{d}y$, co powoduje, że $\tilde{f}(\pi) = 0$. Łatwo się przekonać, że $\tilde{f}$ jest elementem $D(A)$. Rachunki przedstawione wyżej dowodzą ponadto, że ciąg $(\tilde{a}_n)_{n\geq 1}$ współczynników Fouriera funkcji $\tilde{f}$ dany jest wzorem

$$(\tilde{a}_n)_{n\geq 1} = -\left(\frac{1}{n^2}a_n\right)_{n\geq 1} = (a_n)_{n\geq 1}.$$

To oznacza jednak, że zarówno f jak i $\tilde{f}$ są sumami tego samego szeregu $\sum_{n=1}^\infty a_n e_n$. To kończy dowód. $\square$

Jeśli więc sobie wyobrazić I jako działanie zwierciadła, to wszystko to, co dzieje się $L^2[0,\pi]$ ma swoje odbicie w l^2. A cała zaleta lustra w tym, że widziany w nim obraz (operator B) jest znacznie jaśniejszy niż rzeczywistość (operator A). To odwrotnie niż u świętego Pawła, który pisał „Teraz widzimy jakby w zwierciadle, niejasno; wtedy zaś zobaczymy twarzą w twarz” (1 Kor 13,12, Biblia Tysiąclecia, wydanie drugie, Warszawa 1971). Pomijając fakt, że pisał on o rzeczach dalekich od matematyki (jeśli już, to związanych z ekonomią, Bożą ekonomią [19]), za jego czasów lustra nie były tak doskonałe jak dzisiejsze (robione były chyba z polerowanego brązu), a już na pewno rzeczywistości nie udoskonalały.

Wracając do meritum: co nam daje powyższy lemat? Możemy zamiast równania różniczkowego w $L^2[0,\pi]$ rozwiązywać równanie różniczkowe w l^2. Jeśli u jest rozwiązaniem (11.9), to funkcja v o wartościach w l^2 dana wzorem $v(t) = Iu(t)$ rozwiązuje równanie

$$v'(t) = Bv(t), \tag{11.11}$$

z warunkiem początkowym $v(0) = (a_n)_{n\geq 1} = If_0$. I odwrotnie. Znajdziemy więc najpierw v, a potem policzymy u.

Szukamy v o następujących własnościach (analogicznych do własności u).

(a) Dla każdego $t > 0$ $v(t)$ należy do $D(B)$.

(b) Funkcja u jest różniczkowalna w każdym punkcie $t > 0$ w sensie normy z przestrzeni l^2: istnieje taki element $v'(t) \in l^2$, że

$$\lim_{h\to 0} \left\| \frac{v(t+h) - v(t)}{h} - v'(t) \right\| = 0, \tag{11.12}$$

a funkcja $t \mapsto v'(t)$ jest ciągła w sensie normy w l^2.

(c) Zachodzi równość $\lim_{t\to 0+} v(t) = (a_n)_{n\geq 1}$.

(d) Dla każdego $t > 0$, zachodzi równość $v'(t) = Bv(t)$.

Jedną z zalet pracy w l^2 jest to, że dla każdego $i \geq 1$

$$|a_i| \leq \|(a_n)_{n\geq 1}\|_{l^2} = \sqrt{\sum_{n=1}^{\infty} a_n^2},$$

co powoduje, że zbieżność w normie jest tu – w przeciwieństwie do zbieżności w $L^2[0,\pi]$ – mocniejsza niż zbieżność po współrzędnych. Jeśli więc szukane v dane jest wzorem

$$v(t) = (a_n(t))_{n\geq 1}\,, \tag{11.13}$$

to z różniczkowalności v i równości $v'(t) = (b_n(t))_{n\geq 1}$ wynika ciągła różniczkowalność wszystkich funkcji rzeczywistych $t \mapsto a_n(t)$ oraz równość $a_n'(t) = b_n(t), t > 0$. Ze względu na to, że $Bv(t) = B(a_n)_{n\geq 1} = \left(-n^2 a_n(t)\right)_{n\geq 1}$, funkcje $t \mapsto a_n(t)$ okazują się być rozwiązaniami zwyczajnych równań różniczkowych

$$a_n'(t) = -n^2 a_n(t),$$

z warunkami początkowymi $a_n(0) = a_n$. Stąd wynika już bezpośrednio, że

$$a_n(t) = \mathrm{e}^{-n^2 t} a_n. \tag{11.14}$$

11.6 Rozwiązanie: sprawdzenie

Czy to już koniec zadania? Niestety nie. Pokazaliśmy bowiem tylko, używając faktu, że zbieżność w normie w l^2 pociąga zbieżność po współrzędnych, iż *jeśli* istnieje rozwiązanie równania (11.11), to musi być dane wzorami (11.13) i (11.14). Jego istnienia nie możemy jednak a priori zakładać. Próba odwrócenia rozumowania przedstawionego wyżej też nie może się powieść: załamie się ono w momencie, gdy ze zbieżności po współrzędnych będziemy chcieli wywnioskować zbieżność w normie. Nie ma innego wyjścia, trzeba dowieść bezpośrednio, że v zdefiniowane w.w. wzorami rzeczywiście rozwiązuje (11.11).

Czy spełniony jest warunek (a)? Czy, innymi słowy, $\left(\mathrm{e}^{-n^2 t} a_n\right)_{n\geq 1}$ jest elementem $D(A)$? Tak, bo ze zbieżności szeregu $\sum_{n=1}^{\infty} a_n^2$ wynika, że

$$\limsup_{n\to\infty} \sqrt[n]{a_n^2} \leq 1,$$

a stąd wobec

$$\lim_{n\to\infty} \sqrt[n]{n^4 \mathrm{e}^{-2n^2 t}} = \lim_{n\to\infty} \sqrt[n]{n^4} \lim_{n\to\infty} \mathrm{e}^{-2nt} = 0,$$

kryterium Cauchy'ego przesądza o zbieżności szeregu $\sum_{n=1}^{\infty} n^4 \mathrm{e}^{-2n^2 t} a_n^2$. (Gdybyśmy zakładali, że $f_0 \in D(B)$ – a tego nie robimy – to dowód byłby jeszcze prostszy, bo majorantą szeregu $\sum_{n=1}^{\infty} n^4 \mathrm{e}^{-2n^2 t} a_n^2$ jest szereg $\sum_{n=1}^{\infty} n^4 a_n^2$, który byłby wtedy zbieżny). Warunki (b) i (c) zostaną dowiedzione, jeśli pokażemy, że

$$\lim_{h\to 0} \left\| \frac{v(t+h)-v(t)}{h} - Bv(t) \right\| = 0.$$

W tym celu wprowadzamy oznaczenie

$$\frac{v(t+h)-v(t)}{h} - Bv(t) = \left(c_{n,h}(t)\right)_{n\geq 1}.$$

Mamy oczywiście

$$c_{n,h}(t) = \mathrm{e}^{-n^2 t} a_n \left(\frac{\mathrm{e}^{-hn^2} - 1}{h} + n^2 \right) = \mathrm{e}^{-n^2 t} n^2 a_n \left(1 - h^{-1} \int_0^h \mathrm{e}^{-n^2 y}\, \mathrm{d}y \right).$$

Wynika stąd, że dla każdych n i $t > 0$ granica $\lim_{t\to 0} c_{n,h}(t) = 0$, a $|c_{n,h}(t)| \leq \mathrm{e}^{-n^2 t} n^2 a_n$. Mając dane ϵ, tak wybieramy k, by $\sum_{n=k}^{\infty} \left(\mathrm{e}^{-n^2 t} n^2 a_n\right)^2 \leq \frac{\epsilon}{2}$. Jeśli teraz h_0 dobierzemy tak, by $\sum_{n=1}^{k-1} \left(c_{n,h}(t)\right)^2 \leq \epsilon$, to norma $\left(c_{n,h}(t)\right)_{n\geq 1}$ nie będzie przekraczać

$$\sum_{n=1}^{k-1} \left(c_{n,h}(t)\right)^2 + \sum_{n=k}^{\infty} \left(c_{n,h}(t)\right)^2 \leq \frac{\epsilon}{2} + \sum_{n=k}^{\infty} \left(\mathrm{e}^{-n^2 t} n^2 a_n\right)^2 \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Dowód warunku (d) jest podobny i pozostawiamy go Czytelnikom.

Uwaga 11.1. *Zwróćmy uwagę na fakt, że o* $If_0 = (a_n)_{n\geq 1}$ *nie zakładamy, iż jest elementem* $D(B)$*. A mimo to, dla każdego, nawet bardzo małego* $t > 0$*,* $v(t)$ *do* $D(B)$ *już należy. To sytuacja dość typowa dla równań typu parabolicznego, których przedstawicielem jest równanie ciepła.*

Udowodniwszy, że v rzeczywiście jest rozwiązaniem (11.11) wracamy do równania (11.9). Z naszej analizy wynika, że jego rozwiązaniem jest

$$u(t) = I^{-1}v(t) = I^{-1}\left(\mathrm{e}^{-n^2t}a_n\right)_{n\geq 1} = \sum_{n=1}^{\infty} \mathrm{e}^{-n^2t}a_n e_n.$$

Podkreślmy raz jeszcze, że szereg występujący tu po prawej stronie jest zbieżny w sensie normy w $L^2[0,\pi]$. Dość zaskakujące jest więc odkrycie, że także odpowiadający mu szereg funkcyjny

$$u(t,x) = \sqrt{\frac{2}{\pi}}\sum_{n=1}^{\infty} \mathrm{e}^{-n^2t}a_n \sin nx$$

jest zbieżny i to jednostajnie względem $x \in [0,\pi]$ (przy ustalonym $t > 0$). Można tego dowieść, używając kryterium Weierstrassa, mamy tu bowiem jednostajną majorantę

$$\sqrt{\frac{2}{\pi}}\sum_{n=1}^{\infty} \mathrm{e}^{-n^2t}a_n.$$

Podobne rozumowanie dowodzi też, że tak zdefiniowane u jest rozwiązaniem (11.9) w „zwykłym sensie”. Nie będziemy tu tego robić, bo nie temu tematowi poświęcony jest ten szkic. Napiszemy tylko, że prawdziwym powodem tego niewątpliwego cudu są niezwykłe „wygładzające” własności równań parabolicznych.

11.7 Zadania

Zadanie 11.1. Oblicz (f, e_n) dla funkcji $f(x) = \mathrm{e}^x, x \in [0,\pi]$. Odpowiedni szereg nazywany jest rozwinięciem funkcji f w szereg sinusów. Powtórz ćwiczenie dla $f(x) = x$.

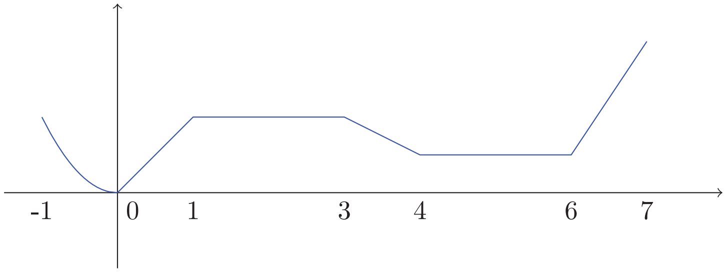
Zadanie 11.2. Udowodnij, że wszystkie przestrzenie $L^2[a,b]$ są izometrycznie izomorficzne z $L^2[0,1]$, a tym samym między sobą. (Zakładamy oczywiście, że $a < b$). Wskazówka: rozważ odwzorowanie $I : L^2[a,b] \to L^2[0,1]$ dane wzorem

$$If(x) = \sqrt{b-a}\, f(k(x)),$$

dla funkcji $k(x) = (b-a)x + a$.

Zadanie 11.3. Sprawdź, że funkcja, której wykres podano na rysunku 11.2 jest absolutnie ciągła, choć nie jest różniczkowalna w kilku punktach.

Zadanie 11.4. Sprawdź, że wzór (11.7) definiuje iloczyn skalarny w $H_0^1(a,b]$.



Rys. 11.2: Nieróżniczkowalna w zwykłym sensie funkcja absolutnie ciągła (zob. zadanie 11.3).

Zadanie 11.5.

(a) W przestrzeni $L^2(0,\infty)$ funkcji całkowalnych z kwadratem na prawej półosi definiujemy ciąg $f_n = 1_{(n,n+1)}$ (funkcja równa 1 na przedziale $(0,1)$ i 0 poza nim). Udowodnij, że dla każdego x granica $\lim_{n\to\infty} f_n(x) = 0$ a jednak $\|f_n\|_{L^2(0,\infty)} = 1$.

(b) W przestrzeni $L^2[0,1]$ następująco definiujemy ciąg funkcji. Pierwsza z nich równa jest stale 1. Funkcja druga równa jest 1 na przedziale $[0,\frac{1}{2}]$ a trzecia na przedziale $[\frac{1}{2},1]$ – poza nimi równe są zero. Następne cztery funkcje przyjmują wartość 1 kolejno na pierwszym, drugim, trzecim i czwartym z odcinków o długości $\frac{1}{4}$, na które składa się $[0,1]$ – a poza nim są równe zero. Kolejnych osiem funkcji to funkcje charakterystyczne ośmiu odcinków o długości $\frac{1}{8}$, i tak dalej. Sprawdź, że granicą tych funkcji w przestrzeni $L^2[0,1]$ jest funkcja zerowa, a jednak ich granica punktowa nie istnieje.

Zadanie 11.6. Udowodnij, że v dane wzorami (11.13) i (11.14) spełnia warunek (d) definicji rozwiązania. Wskazówka: albo rozumuj tak, jak w dowodzie własności (b) i (c), albo skorzystaj z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej.

Rozdział 12

Zupełność algebry $\mathcal{L}(\mathbb{X})$

W trakcie tego wykładu parokrotnie już napotkaliśmy operatory, to znaczy przekształcenia przyporządkowujące funkcjom funkcje, lub ogólniej wektorom wektory – najczęściej wektory te są elementami przestrzeni unormowanej lub Banacha. Operatorami są przekształcenia zdefiniowane w (3.16) i (3.21), które rozważaliśmy, dowodząc twierdzenia Picarda. Jest nim też przekształcenie ze wzoru (5.8) z rozdziału o równaniu McKendricka–Von Foerstera pokazujące, że przestrzenie $L^1(\mathbb{R}^+)$ i $L^1_\omega(\mathbb{R}^+)$ są nieodróżnialne (izometrycznie izomorficzne). Kolejne przykłady to przekształcenia opisane w (5.10) i (6.2) – w tym ostatnim przypadku funkcji o wartościach w przestrzeni Banacha przyporządkowujemy element tej przestrzeni. Ważną klasę operatorów tworzą też operatory rzutowania na podprzestrzeń, które omawialiśmy – nie używając tej nazwy – w rozdziale 9. W rozdziałach 10 i 11 napotkaliśmy natomiast kolejny izometryczny izomorfizm pokazujący, że przestrzenie $L^2[0,\pi]$ i l^2 są nieodróżnialne i dwie kopie tego samego operatora w tychże przestrzeniach (operatory A i B).

Najważniejsze twierdzenie w tym rozdziale mówi o tym, że przestrzeń operatorów ograniczonych nad przestrzenią Banacha sama jest przestrzenią (a w istocie algebrą) Banacha. Ten elementarny, ale piękny wynik pozwala traktować operatory tak, jak elementy przestrzeni (oczywiście innej niż wyjściowa, ale o podobnej strukturze) – a to znakomicie ułatwia sprawę. W istocie znamy to podejście z doświadczenia: macierze $m \times n$ czasem traktujemy jako operatory przekształcające przestrzeń $\mathbb{R}^n$ w $\mathbb{R}^m$, a czasem jako elementy przestrzeni macierzy.

Ale zacznijmy od początku.

12.1 Operatory liniowe

Niech $\mathbb{X}$ i $\mathbb{Y}$ będą przestrzeniami liniowymi. Operator A przekształcający swoją dziedzinę $D(A) \subset \mathbb{X}$ w $\mathbb{Y}$ nazywamy liniowym, gdy spełnione są następujące dwa warunki. Po pierwsze

$$x, y \in D(A), \alpha, \beta \in \mathbb{R} \Rightarrow \alpha x + \beta y \in D(A);$$

po drugie
$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay.$$
Pierwszy z nich stwierdza po prostu, że $D(A)$ jest liniową podprzestrzenią $\mathbb{X}$; kluczem jest drugi. Zwróćmy uwagę, że zgodnie z ogólnie przyjętą konwencją, dla przekształceń liniowych zwykle opuszcza się nawiasy: zamiast pisać $A(x)$ piszemy zwykle Ax (o ile nie prowadzi to do nieporozumień).

Popatrzmy na typowe przykłady

Przykład 12.1. Mnożenie przez macierz. Niech $X = \mathbb{R}^n$ a $Y = \mathbb{R}^m$ i niech
$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$
będzie macierzą o m wierszach i n kolumnach. Przekształcenie
$$Ax = A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{R}^m$$
jest, jak łatwo sprawdzić, liniowe. Dziedziną jest tu cała przestrzeń X.

Przykład 12.2. Niech X będzie przestrzenią wielomianów stopnia najwyżej n, a Y – przestrzenią wielomianów stopnia co najwyżej $n + 1$. Wielomianowi
$$x(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots a_1 t + a_0$$
przyporządkujmy wielomian
$$(Ax)(t) = a_n \frac{t^{n+1}}{n+1} + a_{n-1} \frac{t^n}{n} + \cdots a_1 \frac{t^2}{2} + a_0 t + c.$$
Odwzorowanie to jest liniowe wtedy i tylko wtedy, gdy $c = 0$. Zwróćmy też uwagę, że jeśli utożsamimy X z $\mathbb{R}^{n+1}$ (utożsamiając wielomian z jego współczynnikami) i podobnie postąpimy z Y, to okaże się, że przykład ten jest szczególnym przypadkiem punktu a).

Przykład 12.3. Niech obie przestrzenie (X i Y) będą równe przestrzeni $C_0(\mathbb{R})$ funkcji ciągłych na $\mathbb{R}$, które mają granice zarówno w $+\infty$ jak i w $-\infty$ równe 0. Operator A przekształcający funkcję f w funkcję g zdefiniowaną wzorem
$$g(t) = f(t + 1) + f(t - 1), \qquad t \in \mathbb{R}$$
jest liniowy. Czytelnik powinien sprawdzić również, że A przekształca $C_0(\mathbb{R})$ w siebie. (W szczególności trzeba sprawdzić istnienie granic)!

Przykład 12.4. Niech $P = (p_{i,j})_{i,j \geq 1}$ będzie (nieskończoną) macierzą stochastyczną, to znaczy taką, że $p_{ij} \geq 0$ oraz, dla każdego $i \geq 1$, $\sum_{j=1}^{\infty} p_{ij} = 1$. Niech ponadto $X = Y = l^1$ oznacza przestrzeń bezwzględnie sumowalnych ciągów $(\xi)_{i\geq 1}$. Jeśli wektorowi $(\xi)_{i\geq 1} \in l^1$ przyporządkujemy wektor

$$y = (\xi_1, \xi_2, \dots) \begin{pmatrix} p_{11} & p_{11} & \cdots \\ p_{21} & p_{22} & \cdots \\ \dots & \dots & \dots \end{pmatrix},$$

to otrzymamy odwzorowanie liniowe. Jedynym faktem, który wymaga tu naszej uwagi jest to, czy wartości tego przekształcenia rzeczywiście mieszczą się w l^1. Wynika to jednak z następujących oszacowań. Jeśli $\eta_i = \sum_{j=1}^{\infty} \xi_j p_{ji}$, to

$$\begin{aligned} \sum_{i=1}^{\infty} |\eta_i| &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |p_{ji}\xi_j| = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ji}|\xi_j| \\ &\leq \sum_{j=1}^{\infty} |\xi_j| \sum_{i=1}^{\infty} p_{ji} = \sum_{j=1}^{\infty} |\xi_j| < \infty. \end{aligned} \tag{12.1}$$

Przykład 12.5. Niech c_{00} będzie przestrzenią ciągów $x = (\xi_i)_{i\geq 1}$ nieskończonych, które od pewnego miejsca (zależnego od ciągu) równe są zero. Operator zdefiniowany wzorem

$$Ax = \sum_{i=1}^{\infty} i\xi_i$$

jest liniowy i przekształca c_{00} w $\mathbb{R}$. (Zwróćmy uwagę, że suma powyższa jest w istocie skończona – prawie wszystkie jej składniki równe są zero).

Przykład 12.6. Niech $\mathbb{X} = C^1[0,1]$ będzie przestrzenią funkcji określonych na $[0,1]$, które są różniczkowalne w sposób ciągły, a $\mathbb{Y}$ – przestrzenią funkcji ciągłych na tym samym przedziale. Operator przyporządkowujący funkcji $f \in C^1[0,1]$ jej pochodną (element $C[0,1]$) jest liniowy.

Przykład 12.7. W przestrzeni c_0 (ciągów $x = (\xi_i)_{i\geq 0}$, zbieżnych do zera) definiujemy dwa operatory:

$$Ax = (\xi_2, \xi_3, \cdots)$$

oraz

$$Bx = (0, \xi_1, \xi_2, \cdots).$$

Przekształcenia te, zwane operatorami przesunięć w lewo i w prawo, są liniowe.

Przykład 12.8. Niech $X = Y = C_0[0,\infty)$ będzie przestrzenią funkcji ciągłych na $\mathbb{R}^+$, które w nieskończoności mają granicę równą 0. Niech X będzie zmienną losową o rozkładzie wykładniczym z parametrem λ, to znaczy zmienną losową o wartościach dodatnich, której gęstość jest dana wzorem $x \mapsto \lambda \mathrm{e}^{-\lambda x}$, $x \geq 0$. Niech też

$$Af(t) = E\, f(t+X) = \lambda \int_0^{\infty} \mathrm{e}^{-\lambda s} f(t+s)\,\mathrm{d}s, \qquad t \geq 0,$$

gdzie E oznacza wartość oczekiwaną. Operator A jest liniowy.

Przykład 12.9. Jeśli operatory A i B o tej samej dziedzinie (i tej samej przestrzeni wartości Y) są liniowe, to liniowy jest też operator $\alpha A\ + \beta B$ zdefiniowany wzorem $(\alpha A\ + \beta B)x = \alpha A\ x + \beta Bx$.

Przykład 12.10. Niech $X = Y$. Operator identycznościowy przekształcający każdy wektor w siebie jest liniowy. Oznaczamy go I_X, albo po prostu I.

Czytelnik powinien samodzielnie sprawdzić teraz, które z operatorów omawianych we wstępie do tego wykładu są liniowe (zadanie 12.1).

12.2 Operatory ograniczone i ich normy

Przekształcenie liniowe $A : \mathbb{X} \to \mathbb{Y}$ z przestrzeni unormowanej $\mathbb{X}$ do takiejże przestrzeni $\mathbb{Y}$ będziemy nazywać ograniczonym, jeśli istnieje $M \geq 0$ o tej własności, że dla każdego $x \in \mathbb{X}$ zachodzi nierówność

$$\|Ax\|_{\mathbb{Y}} \leq M\|x\|_{\mathbb{X}};$$

w szczególności zakładamy, że dziedziną A jest całe $\mathbb{X}$.

Rozważmy kilka operatorów z poprzedniego podrozdziału. W przykładzie 12.1 załóżmy dla ustalenia uwagi, że $\|x\|_{\mathbb{X}} = \max_{i=1,\dots,n} |\xi_i|$, a $\|y\|_{\mathbb{Y}} = \sqrt{\sum_{i=1}^m \eta_i^2}$. Oznaczając $M_0 = \max_{i,j} |a_{i,j}|$, otrzymujemy:

$$\begin{aligned}\|Ax\| &= \sqrt{\sum_{i=1}^m \eta_i^2} = \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^n a_{i,j}\xi_j\right)^2} \leq \sqrt{\sum_{i=1}^m \left(M_0^2 \sum_{j=1}^n |\xi_j|\right)^2} \\ &\leq \sqrt{mM_0^2 n^2 \|x\|^2} = \sqrt{m} M_0 n \|x\|.\end{aligned}$$

Dowodzi to, że A jest ograniczony (stałą M z definicji ograniczoności może być na przykład $\sqrt{m}M_0 n$).

Zwróćmy uwagę na to, że w przykładzie 12.2 operator A można zapisać nieco inaczej:

$$Ax(t) = \int_0^t x(s)\,\mathrm{d}s.$$

Jeśli zarówno $\mathbb{X}$ jak i $\mathbb{Y}$ rozważać będziemy z normą supremum, to otrzymamy oszacowanie:

$$\begin{aligned}\|Ax\| &= \sup_{t\in[0,1]} |Ax(t)| = \sup_{t\in[0,1]} |\int_0^t x(s)\,\mathrm{d}s| \leq \sup_{t\in[0,1]} \int_0^t |x(s)|\,\mathrm{d}s \\ &\leq \sup_{t\in[0,1]} \int_0^t \|x\|\,\mathrm{d}s = \|x\|,\end{aligned}$$

dowodzące, że A jest ograniczone; stałą M jest tu na przykład 1.

Przykład 12.3 na razie opuszczamy, omówimy go dokładnie później. Natomiast oszacowanie (12.1) dowodzi, że operator opisany w przykładzie 12.4 jest ograniczony (znów $M = 1$).

Operator z przykładu 12.5 zaś ograniczony nie jest. By tego dowieść rozważmy ciąg $e_n \in c_{00}$, którego n-ty wyraz równy jest 1, a pozostałe zero. Mamy $\|e_n\| = 1$ i $Ae_n = n$ więc gdyby istniała stała M z definicji ograniczoności, to musielibyśmy mieć $M \geq n$. Dowolność n oczywiście na to nie pozwala.

W przykładzie 12.6 odpowiedź na pytanie, czy jest to operator ograniczony, nie jest prosta. Zależy ona bowiem od tego, jaką normę wybierzemy w rozważanych przestrzeniach. Oczywiście ostatnie zdanie dotyczy każdej sytuacji, bo ograniczoność zawsze zależy od wyboru normy: przy jednej normie operator ograniczony będzie, a przy innej nie. Tutaj kwestia polega na tym, co jest normą „naturalną". Jeśli obie przestrzenie, to znaczy $C^1[0,1]$ i $C[0,1]$ wyposażymy w normę supremum, to operator różniczkowania nie będzie ograniczony. Jeśli natomiast w tej pierwszej przestrzeni wprowadzimy normę

$$\|f\|_{C^1[0,1]} = \|f\|_{C[0,1]} + \|f'\|_{C[0,1]},$$

to ograniczony będzie. To drugie stwierdzenie jest oczywiste, bo zachodzi nierówność $\|f'\|_{C[0,1]} \leq \|f\|_{C^1[0,1]}$. To pierwsze wymaga wyjaśnienia: przyjmując $f_n(t) = t^n$, widzimy, że $\|f\|_{C[0,1]} = 1$, podczas gdy $f_n'(t) = nt^{n-1}$, co implikuje $\|f_n'\| = n$. Tak samo jak w przykładzie e) pokazujemy zatem, że M spełniające warunki definicji nie może istnieć.

Odkładając na razie analizę pozostałych przykładów, przejdźmy do opisu przestrzeni operatorów. Operatory liniowe i ograniczone, przekształcające przestrzeń $\mathbb{X}$ w $\mathbb{Y}$, same tworzą przestrzeń liniową, oznaczaną symbolem $\mathcal{L}(\mathbb{X}, \mathbb{Y})$: jeśli A i B do niej należą, a α i β są skalarami, to możemy zdefiniować nowy operator $\alpha A + \beta B$ wzorem

$$(\alpha A + \beta B)(x) = \alpha Ax + \beta Bx, \qquad x \in \mathbb{X},$$

i okazuje się, że taka definicja czyni zadość wszystkim warunkom przestrzeni liniowej. Zwróćmy uwagę na to, że jeśli $\|Ax\| \leq L\|x\|$ oraz $\|Bx\| \leq M\|x\|$, to

$$\|(\alpha A + \beta B)(x)\| \leq (|\alpha|L + |\beta|M)\|x\|,$$

co dowodzi, że operator $\alpha A + \beta B$ jest ograniczony (załatwiliśmy tym samym sprawę ograniczoności operatora z przykładu 12.9). W przypadku, gdy $\mathbb{X} = \mathbb{Y}$, zamiast $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ używamy symbolu $\mathcal{L}(\mathbb{X})$ i mówimy o przestrzeni (ograniczonych) operatorów na $\mathbb{X}$. Jeśli $\mathbb{Y} = \mathbb{R}$, piszemy po prostu $\mathbb{X}^*$ – ta przestrzeń nazywana jest przestrzenią (ograniczonych) funkcjonałów na $\mathbb{X}$.

Następujące twierdzenie wyjaśnia dlaczego w przypadku operatorów liniowych przymiotników „ograniczony" i „ciągły" używa się zamiennie.

Twierdzenie 12.1. *Niech $\mathbb{X}$ i $\mathbb{Y}$ będą przestrzeniami unormowanymi. Dla operatora liniowego $A : \mathbb{X} \to \mathbb{Y}$ następujące warunki są równoważne:*

(a) A jest ciągły,

(b) A jest ciągły w pewnym punkcie $x \in \mathbb{X}$,

(c) A jest ciągły w zerze,

(d) $\sup_{\|x\|_{\mathbb{X}}=1} \|Ax\|_{\mathbb{Y}}$ *jest skończone,*

(e) A *jest ograniczony.*

Jeśli A jest ciągły, to zachodzi równość

$$\sup_{\|x\|_{\mathbb{X}}=1} \|Ax\|_{\mathbb{Y}} = \min\{M \in \mathcal{S}\} \tag{12.2}$$

w której $\mathcal{S}$ jest zbiorem stałych spełniających warunek $\|Ax\|_{\mathbb{Y}} \le M\|x\|_{\mathbb{X}}$ dla każdego $x \in \mathbb{X}$.

Dowód. Implikacja $(a) \Rightarrow (b)$ jest oczywista, bo pierwszy z tych warunków wymaga, by ciąg $(Ax_n)_{n\ge1}$ zbiegał do Ax dla każdego x i dążącego do niego ciągu $(x_n)_{n\ge1}$, podczas gdy drugi – tego, by było tak dla pewnego x i odpowiednich ciągów. Idąc dalej, jeśli $\lim_{n\to\infty} x_n = 0$ to $\lim_{n\to\infty}(x_n + x) = x$. Jeśli więc zachodzi (b), to $A(x_n+x)$, które jest równe Ax_n+Ax, dąży do Ax; innymi słowy granicą Ax_n jest 0, a to dowodzi (c). Dla dowodu, że (c) implikuje (d), załóżmy że ten ostatni warunek nie zachodzi, a więc, że istnieje taki ciąg $(x_n)_{n\ge1}$ elementów $\mathbb{X}$, iż $\|x_n\|_{\mathbb{X}} = 1$ oraz $\|Ax_n\| > n$. W takiej sytuacji granicą wektorów $y_n = \frac{1}{\sqrt{n}}x_n$ jest zero, podczas gdy $\|Ay_n\|_{\mathbb{Y}} > \sqrt{n}$ do zera zbiegać nie może, a zatem nie zachodzi (c). O tym, że (d) implikuje (e) można się przekonać, definiując $M = \sup_{\|x\|_{\mathbb{X}}=1} \|Ax\|_{\mathbb{Y}}$. Rzeczywiście, nierówność $\|Ax\|_{\mathbb{Y}} \le M\|x\|_{\mathbb{X}}$ jest oczywista dla $x = 0$, a dla niezerowego wektora x norma $\frac{1}{\|x\|_{\mathbb{X}}}x$ równa jest 1, więc $\|A\frac{1}{\|x\|_{\mathbb{X}}}x\|_{\mathbb{Y}} \le M$. Mnożąc obie strony przez $\|x\|_{\mathbb{X}}$, otrzymujemy stąd (e). Na zakończenie zauważmy, że (a) wynika z (e), bo $\|Ax_n - Ax\|_{\mathbb{Y}} \le \|A(x_n - x)\| \le M\|x_n - x\|_{\mathbb{X}}$.

By dowieść drugiej części twierdzenia zauważmy, że w dowodzie implikacji (d)$\Rightarrow$ (e) pokazaliśmy, iż $M_1 := \sup_{\|x\|_{\mathbb{X}}=1} \|Ax\|_{\mathbb{Y}}$ jest elementem $\mathcal{S}$. Z drugiej strony zaś, jeśli nierówność $\|Ax\|_{\mathbb{Y}} \le M\|x\|_{\mathbb{X}}$ zachodzi dla wszystkich $x \in X$, to ograniczając się do x o tej własności, że $\|x\|_{\mathbb{X}} = 1$ przekonujemy się, iż $M_1 \le M$, a to znaczy, że M_1 jest minimum zbioru $\mathcal{S}$. □

Przyjrzyjmy się drugiej części twierdzenia chwilę dłużej. Jeśli operator A jest ograniczony, to zbiór $\mathcal{S}$ jest niepusty. Stałych M, które do niego należą jest nieskończenie wiele – jeśli należy do niego jakaś liczba, to należą też do niego wszystkie liczby od niej większe. Wprowadzona w (12.2) wielkość

$$\|A\| := \sup_{\|x\|_{\mathbb{X}}=1} \|Ax\|_{\mathbb{Y}} = \min\{M \in \mathcal{S}\} \tag{12.3}$$

jest najmniejszą z takich stałych i to policzoną efektywnie. Nazywamy ją normą operatora ograniczonego. W ramach ćwiczenia Czytelnik powinien sprawdzić, że odwzorowanie $A \mapsto \|A\|$ jest rzeczywiście normą w sensie definicji z podrozdziału 4.1.2.

Zacznijmy od przykładu 12.3. Występujący tam operator jest ograniczony, bo

$$\sup_{t\in\mathbb{R}} |g(t)| \le \sup_{t\in\mathbb{R}} |f(t+1)| + \sup_{t\in\mathbb{R}} |f(t-1)| = 2\sup_{t\in\mathbb{R}} |f(t)| = 2\|f\|.$$

Przy okazji pokazaliśmy, że $\|A\| \leq 2$. By przekonać się, że $\|A\| = 2$ rozważamy na przykład funkcję, która w przedziale $[-1, 1]$ równa jest stale jeden, poza przedziałem $[-2, 2]$ równa jest zero, a jej wykres w przedziałach $[-2, -1]$ i $[1, 2]$ jest odcinkiem łączącym odpowiednie punkty. Oczywiście $\|f\| = 1$ a $Af(0) = f(1) + f(-1) = 2$. Stąd $\|Af\| \geq |Af(0)| \geq 2$, więc i $\|A\| \geq 2$, co kończy rozumowanie.

Pokażemy też, że wszystkie operatory związane z macierzami stochastycznymi mają normy równe 1 (patrz przykład 12.4). Oszacowania (12.1) dowodzą jedynie, że 1 jest elementem zbioru $\mathcal{S}$. Na razie wiemy więc tylko, że $\|A\| \leq 1$. Gdybyśmy jednak rozważyli wektor x o wszystkich składowych nieujemnych, to nierówności w (12.1) zamieniłyby się w równości. W szczególności dla x, którego wyrazy sumują się do 1, lewa strona okazałaby się równa 1. Dowodzi to, że supremum występujące w (12.3) równe jest jeden, a to znaczy, że $\|A\| = 1$. Dla wygody zanotujmy, że operatory związane z macierzami stochastycznymi nazywane są operatorami Markowa.

Ogólniej, macierz $A = (a_{i,j})_{i,j \in \mathbb{N}}$ definiuje operator ograniczony na l^1, dany wzorem

$$A(\xi_i)_{i \geq 1} = \left(\sum_{j=1}^{\infty} \xi_j a_{j,i} \right)_{i \geq 1},$$

o ile wielkość $\sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} |a_{i,j}|$ jest skończona, a wtedy

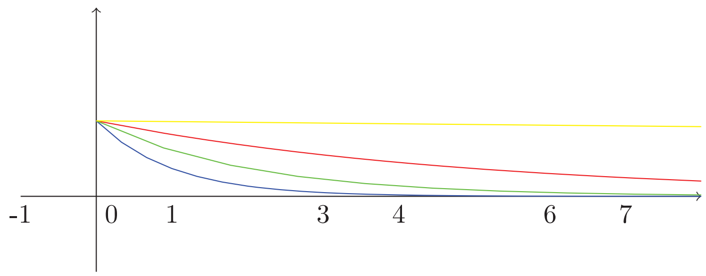
$$\|A\| = \sup_{j \in \mathbb{N}} \sum_{i=1}^{\infty} |a_{j,i}|. \tag{12.4}$$

(Operator A sprowadza się do mnożenia „poziomego” wektora $(\xi_i)_{i \geq 1}$ przez ustawioną po jego prawej stronie macierz A – operator i macierz utożsamiamy ze sobą, więc używamy do ich oznaczenia tej samej litery). Dowód, że $\|A\|$ nie przekracza wyrażenia po prawej stronie, nazwijmy je M_0, jest niemal natychmiastowy, ponieważ

$$\left\| \left(\sum_{j=1}^{\infty} \xi_j a_{j,i} \right)_{i \geq 1} \right\| = \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} \xi_j a_{j,i} \right| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |\xi_j| \, |a_{j,i}| = \sum_{j=1}^{\infty} |\xi_j| \sum_{i=1}^{\infty} |a_{j,i}|$$

$$\leq M_0 \sum_{j=1}^{\infty} |\xi_j| = M_0 \left\| (\xi_i)_{i \geq 1} \right\|.$$

By udowodnić nierówność przeciwną, rozważmy wektor $e_j \in l^1$, który na j-tej współrzędnej ma 1, a poza tym zera. Oczywiście $\|e_j\| = 1$. Mamy też $Ae_j = (a_{j,i})_{i \geq 1}$ (prawa strona to j-ty wiersz macierzy A), więc $\|Ae_j\| = \sum_{i=1}^{\infty} |a_{j,i}|$. Skoro e_j ma opisane wyżej własności dla każdego j, to warunkiem koniecznym na to, by operator A był ograniczony jest, by M_0, które jest równe $\sup_{j \in \mathbb{N}} \|Ae_j\|$ było skończone. Otrzymujemy też $\|A\| \geq M_0$. To kończy dowód. Pozostawiam Czytelnikowi sprawdzenie, że z właśnie udowodnionego rezultatu wynika bezpośrednio, iż operatory Markowa mają normę równą 1.



Rys. 12.1: Funkcje wykładnicze $\mathrm{e}_\mu(x) = \mathrm{e}^{-\mu x}$ dla małych μ (kolor niebieski $\mu = 1$, kolor zielony $\mu = 0,5$ kolor czerwony $\mu = 0,2$ i kolor żółty $\mu = 0,01$). Dla $\mu \to 0$ funkcje te przybliżają w pewnym sensie funkcję stale równą jeden.

Oba operatory w przykładzie 12.7 są ograniczone i mają normy równe 1. Z jednej bowiem strony

$$\|Ax\| = \sum_{i=2}^{\infty} |\xi_i| \le \sum_{i=1}^{\infty} |\xi_i| = \|x\|,$$

a z drugiej, jeśli $\xi_1 = 0$, to powyższa nierówność zamienia się w równość. Podobnie postępujemy z B.

Najbardziej pouczający jest chyba przykład 12.8. Tu też wyliczymy normę, choć nie umiemy znaleźć elementu $x \in \mathbb{X}$ z normą równą jeden, dla którego $\|Ax\| = \|A\|$. Oszacowanie $\|A\| \le 1$ otrzymujemy łatwo, korzystając z faktu, że $\lambda \int_0^\infty \mathrm{e}^{-\lambda x}\, \mathrm{d}x = 1$:

$$\begin{aligned} \|Af\| &= \sup_{t\ge 0} |\lambda \int_0^\infty \mathrm{e}^{-\lambda s} f(t+s)\, \mathrm{d}s| \le \sup_{t\ge 0} \lambda \int_0^\infty \mathrm{e}^{-\lambda s} |f(t+s)|\, \mathrm{d}s \\ &\le \sup_{t\ge 0} \lambda \int_0^\infty \mathrm{e}^{-\lambda s} \|f\|\, \mathrm{d}s = \|f\|. \end{aligned}$$

Gdyby funkcja stale równa jeden, nazwijmy ją e_0, należała do przestrzeni $C_0[0,\infty)$, w której operujemy, dowód nierówności $\|A\| \ge 1$ byłby natychmiastowy, bo $A\mathrm{e}_0$ równałoby się e_0. Skoro jednak $\mathrm{e}_0 \notin C_0[0,\infty)$ (bo $\lim_{t\to\infty} \mathrm{e}_0(t) = 1 \ne 0$) musimy postępować nieco ostrożniej (zob. rysunek 12.1). W tym celu definiujemy $\mathrm{e}_\mu \in C_0[0,\infty)$ dane wzorem $\mathrm{e}_\mu(t) = \mathrm{e}^{-\mu t}, t \ge 0, \mu > 0$. Oczywiście $\|\mathrm{e}_\mu\| = 1$, podczas gdy

$$A\mathrm{e}_\mu(t) = \lambda \int_0^\infty \mathrm{e}^{-\lambda s} \mathrm{e}^{-\mu(t+s)}\, \mathrm{d}s = \lambda \mathrm{e}^{-\mu t} \int_0^\infty \mathrm{e}^{-(\lambda+\mu)s}\, \mathrm{d}s = \frac{\lambda}{\lambda+\mu} \mathrm{e}^{-\mu t}.$$

W konsekwencji $\|Ae_\mu\| = \frac{\lambda}{\lambda+\mu}$, a stąd

$$\|A\| \ge \sup_{\mu>0} \|A\mathrm{e}_\mu\| \ge \lim_{\mu\to 0} \|A\mathrm{e}_\mu\| = 1.$$

12.3 Zupełność przestrzeni operatorów

Skoro przestrzeń operatorów $\mathcal{L}(\mathbb{X},\mathbb{Y})$ jest unormowana, możemy mówić o ich zbieżności: ciąg $(A_n)_{n\geq 1}$ dąży do A jeśli

$$\lim_{n\to\infty} \|A_n - A\| = 0.$$

Dokładniej rzecz biorąc, w takim przypadku stwierdzamy, że $(A_n)_{n\geq 1}$ zbiega do A w normie operatorowej (lub po prostu: w normie).

Na przykład, wobec nierówności (12.17), jeśli ciąg $(A_n)_{n\geq 1}$ operatorów z $\mathcal{L}(\mathbb{X})$ zbiega w normie operatorowej do A, to dla dowolnego operatora $B \in \mathcal{L}(\mathbb{X})$, ciągi $(BA_n)_{n\geq 1}$ i $(A_nB)_{n\geq 1}$ zbiegają odpowiednio do BA i AB. Ogólniej, jeśli $(A_n)_{n\geq 1}$ i $(B_n)_{n\geq 1}$ zbiegają odpowiednio do A i B, to $(A_nB_n)_{n\geq 1}$ zbiega do AB. By się o tym przekonać szacujemy następująco:

$$\|A_nB_n - AB\| \leq \|A_nB_n - A_nB\| + \|A_nB - AB\| \leq \|A_n\|\,\|B_n - B\| + \|B\|\,\|A_n - A\|$$

i korzystamy z faktu, że ciąg liczbowy $(\|A_n\|)_{n\geq 1}$ jest ograniczony (ten ostatni fakt wynika ze zbieżności ciągu $(A_n)_{n\geq 1}$ i nierówności trójkąta dla normy w $\mathcal{L}(\mathbb{X})$, która powoduje, że $|\,\|A_n\| - \|A_m\|\,| \leq \|A_n - A_m\|$).

Naturalne pytanie brzmi, czy $\mathcal{L}(\mathbb{X},\mathbb{Y})$ jest przestrzenią zupełną. Okazuje się, że odpowiedź brzmi „tak” o ile tylko zupełna jest $\mathbb{Y}$.

Twierdzenie 12.2. *Jeśli $\mathbb{Y}$ jest przestrzenią Banacha, to jest nią też $\mathcal{L}(\mathbb{X},\mathbb{Y})$.*

Dowód. Jeśli $(A_n)_{n\geq 1}$ jest ciągiem podstawowym w $\mathcal{L}(\mathbb{X},\mathbb{Y})$, to dla każdego $x \in \mathbb{X}$ i $\epsilon > 0$ zachodzą nierówności

$$\|A_nx - A_mx\| \leq \|A_n - A_m\|\,\|x\| \leq \epsilon\|x\|,$$

o ile tylko n i m są odpowiednio duże. Dowodzi to, że $(A_nx)_{n\geq 1}$ o wartościach w $\mathbb{Y}$ jest ciągiem Cauchy'ego. Z zupełności $\mathbb{Y}$ wnioskujemy, że $(A_nx)_{n\geq 1}$ ma granicę. Zależy ona w sposób oczywisty od x; otrzymujemy więc odwzorowanie $x \mapsto Ax = \lim_{n\to\infty} A_nx$. Bez wątpienia $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$, co dowodzi, że A jest operatorem liniowym. Przechodząc do granicy $m \to \infty$ w powyższym oszacowaniu, przekonujemy się, że dla odpowiednio dużych n (i wszystkich x) zachodzi nierówność

$$\|A_nx - Ax\| \leq \epsilon\|x\|.$$

Dowodzi to zarówno, że A jest ograniczony (bo $\|Ax\| \leq \|Ax - A_nx\| + \|A_nx\| \leq (\epsilon + \|A_n\|)\|x\|$), a tym samym, że jest elementem $\mathcal{L}(\mathbb{X},\mathbb{Y})$, jak i tego, że $(A_n)_{n\geq 1}$ zbiega do A (bo nierówność ta implikuje $\|A_n - A\| \leq \epsilon$ dla odpowiednio dużych n). □

Jednym z najważniejszych przypadków jest ten, gdy $\mathbb{X} = \mathbb{Y}$ i ta ostatnia przestrzeń jest zupełna. Jak to wynika z nierówności (12.17), $\mathcal{L}(\mathbb{X}) := \mathcal{L}(\mathbb{X},\mathbb{X})$ jest wtedy algebrą Banacha. Posiada ona jedynkę (jest nią identyczność), ale jest nieprzemienna. Chociażby operatory A i B rozważane w przykładzie 12.7, które są elementami $\mathcal{L}(c_0)$, nie komutują. Jeśli bowiem ciąg przesunę najpierw w prawo, a potem w lewo to wynik nie będzie taki sam jak gdybym przesunął go najpierw w lewo, a potem w prawo. W przypadku pierwszym otrzymam ciąg wyjściowy, w tym drugim – „zgubię” pierwszą współrzędną, która stanie się zerem.

Zastosowania twierdzenia 12.2, które omówimy poniżej dotyczą właśnie $\mathcal{L}(\mathbb{X})$, gdy $\mathbb{X}$ jest przestrzenią Banacha.

12.4 Zastosowanie: rozwiązywanie równań

Fakt, że przestrzeń ograniczonych operatorów liniowych o wartościach w przestrzeni Banacha sama jest przestrzenią zupełną, ma głębokie znaczenie (pomijając piękno i swoistą symetrię powyższego stwierdzenia). Jednym z najprostszych wynikających z niego wniosków (o wielorakich zastosowaniach) jest to, że dla każdego operatora $A \in \mathcal{L}(\mathbb{X})$ działającego w przestrzeni Banacha $\mathbb{X}$ warunek

$$\|A - I\| < 1 \tag{12.5}$$

gwarantuje (przypomnijmy, że operator identycznościowy I zdefiniowany jest wzorem $Ix = x$) istnienie operatora doń odwrotnego. Znaczy to, że istnieje takie $B \in \mathcal{L}(\mathbb{X})$, oznaczane dalej A^{-1}, że

$$BA = AB = I.$$

Dla dowodu przyjmijmy $q = \|A - I\| < 1$ i rozważmy operatory:

$$B_n = I + (I - A) + (I - A)^2 + ... + (I - A)^n,$$

gdzie $(I - A)^i$ jest itą potęgą operatora $I - A$, to znaczy i-krotnym złożeniem $I - A$ ze sobą. Wobec (12.17),

$$\begin{aligned}\|B_n - B_m\| &= \| \sum_{k=m+1}^{n} (I - A)^k \| \le \sum_{k=m+1}^{n} \|(I - A)^k\| \le \sum_{k=m+1}^{n} q^k \\ &= q^{m+1} \frac{1 - q^{n-m}}{1 - q} \le \frac{q^{m+1}}{1 - q},\end{aligned}$$

(wykorzystaliśmy tu nierówność trójkąta oraz (12.17)) o ile $n > m$. Wynika stąd, że $(B_n)_{n \ge 1}$ jest ciągiem Cauchy'ego w $\mathcal{L}(\mathbb{X})$. Niech B będzie jego granicą. Skoro, jak tego dowodzą proste rachunki

$$(I - A)B_n = B_n(I - A) = B_{n+1} - I,$$

to przechodząc z $n \to \infty$, otrzymamy $B - AB = B - BA = B - I$ (wykorzystujemy tu pierwszy przykład zbieżności operatorów opisany na początku podrozdziału 12.3), a to już dowodzi tezy.

Zanotujmy wynikający z powyższych rozważań wzór

$$A^{-1} = \sum_{k=0}^{\infty} (I - A)^k, \tag{12.6}$$

który jest operatorową kalką znanego rozwinięcia

$$\frac{1}{x} = \frac{1}{1-(1-x)} = \sum_{k=0}^{\infty}(1-x)^k$$

prawdziwego dla x spełniających warunek $|1-x| < 1$. Szereg występujący po prawej stronie równania (12.6) nazywany jest szeregiem Neumanna[13].

Rozwiązywanie równań algebraicznych w przestrzeniach Banacha

Niech $\mathbb{X}$ będzie przestrzenią Banacha, a A – operatorem, elementem $\mathcal{L}(\mathbb{X})$. Mówimy, że A ma prawy odwrotny, jeśli istnieje taki operator $B \in \mathcal{L}(\mathbb{X})$, że dla każdego $y \in \mathbb{X}$ zachodzi równość:

$$ABy = y.$$

Istnienie prawego odwrotnego dowodzi, że operator A jest „na": dla każdego $y \in \mathbb{X}$ istnieje taki x (równy By), że $Ax = y$. Jeśli więc taki operator istnieje, to równanie

$$Ax = y, \tag{12.7}$$

w którym y jest dane, a x szukane, ma (przynajmniej jedno) rozwiązanie.

Operator A z przykładu 12.7 ma prawy odwrotny, a jest nim B z tegoż samego przykładu – jeśli ciąg przesuniemy najpierw w prawo, a potem w lewo, to w efekcie otrzymamy ciąg wyjściowy.

Warto podkreślić, że prawych odwrotnych do danego operatora A może być wiele. Na przykład zarówno B_1 jak i B_2 zdefiniowane niżej pełnią taką rolę wobec omawianego tu A:

$$B_1x = (\xi_1, \xi_1, \xi_2, \cdots), \qquad B_2x = \left(\sum_{i=1}^{\infty}\frac{1}{2^i}\xi_i, \xi_1, \xi_2, \cdots\right), \tag{12.8}$$

a zdolny student wytrzepie z rękawa tego typu operatorów całą masę. Ma to związek z niejedynością rozwiązań równania (12.7). Powtórzmy zatem, że

> istnienie operatora odwrotnego prawego gwarantuje *istnienie* rozwiązań, ale nie ich *jedyność.*

Tę ostatnią cechę opisuje operator odwrotny lewy. Z definicji jest to taki operator B, że dla każdego $y \in \mathbb{X}$ zachodzi równość

$$BAy = y.$$

[13]Karla Gottfrieda Neumanna, matematyka niemieckiego nie należy mylić z Johnem von Neumannem, matematykiem austriacko-węgierskim, a później amerykańskim.

Gwarantuje on, że rozwiązanie równania (12.7), jeśli istnieje, to jest jedyne. Rzeczywiście, gdyby istniały dwa, powiedzmy x_1 i x_2, to mielibyśmy

$$x_1 - x_2 = BAx_1 - BAx_2 = y - y = 0,$$

a zatem $x_1 = x_2$. W takim wypadku operator A jest różnowartościowy: jeśli $Ax = 0$, to $x = BAx = B0 = 0$, co dowodzi, że jądro A jest trywialne.

Lewym odwrotnym dla operatora B z przykładu 12.7 jest opisany tam operator A. B jest więc (jak to też widać gołym okiem) różnowartościowy. Nie jest natomiast „na", bo w zbiorze jego wartości nie leży na przykład $(1, 0, 0, \dots)$. Jak widać na tym przykładzie, posiadanie lewego odwrotnego nie pociąga za sobą posiadania odwrotnego prawego. Faktem jest natomiast, że lewy odwrotny, jeśli istnieje, jest tylko jeden.

Przeprowadzone wyżej rozważania rzucają więcej światła na warunek (12.5). Skoro gwarantuje on, że istnieje zarówno prawy jak i lewy odwrotny do A, to gwarantuje także, że dla każdego $y \in \mathbb{X}$ równanie (12.7) ma dokładnie jedno rozwiązanie.

Jeśli zachodzi warunek (12.5), to równanie (12.7) ma dokładnie jedno rozwiązanie.

Oto prosty przykład (por. zadanie 3.12). Pokażemy, że niezależnie od tego jak dobierzemy funkcję g z $C[0, 1]$ i stałą $k \in \mathbb{R}$, istnieć będzie w tej przestrzeni dokładnie jedno f spełniające warunek:

$$f(t) - k \int_0^t f(s)\,\mathrm{d}s = g(s), \qquad t \in [0, 1]. \tag{12.9}$$

Jeśli zdefiniujemy operator A wzorem $Af(t) = f(t) - k \int_0^t f(s)\,\mathrm{d}s, t \in [0, 1]$, to (12.9) będzie można napisać tak:

$$Af = g,$$

a wobec twierdzenia w ramce nasze zadanie sprowadzi się do sprawdzenia, iż zachodzi (12.5). Jak się okazuje, w standardowej normie supremum (12.5) nie jest spełniony (chyba, że $|k|$ jest mała) – nie przeszkadza nam to jednak, bo możemy (podobnie jak w drugim, lepszym dowodzie lokalnego twierdzenia Picarda) użyć normy równoważnej, typu Bieleckiego[14]:

$$\|f\| = \max_{t \in [0,1]} \mathrm{e}^{-\lambda t} |f(t)|,$$

[14]Zdolnemu do refleksji Czytelnikowi, któremu nie może się pomieścić w głowie, że dla jednej normy (na przykład tej standardowej) twierdzenie z poprzedniego rozdziału nie działa, a dla innej i owszem, śpieszę wyjaśnić, że (12.5) jest warunkiem dostatecznym, a nie koniecznym istnienia i jedyności rozwiązań.

gdzie λ jest liczbą dodatnią, której wartość ustalimy później. Skoro $(A-I)f(t) = -k\int_0^t f(s)\,\mathrm{d}s$, to

$$\begin{aligned}\|(A-I)f\| &= \sup_{t\in[0,1]} |k|\mathrm{e}^{-\lambda t}\left|\int_0^t f(s)\,\mathrm{d}s\right| \le \sup_{t\in[0,1]} |k|\mathrm{e}^{-\lambda t}\int_0^t |f(s)|\,\mathrm{d}s \\ &= \sup_{t\in[0,1]} \int_0^t |k|\mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{-\lambda s}|f(s)|\,\mathrm{d}s \le \sup_{t\in[0,1]} \int_0^t |k|\mathrm{e}^{-\lambda(t-s)}\|f\|\,\mathrm{d}s \\ &= \sup_{t\in[0,1]} \int_0^t |k|\mathrm{e}^{-\lambda s}\,\mathrm{d}s\|f\| \le \int_0^\infty |k|\mathrm{e}^{-\lambda s}\,\mathrm{d}s\|f\| \\ &= \frac{|k|}{\lambda}\|f\|.\end{aligned}$$

A zatem jeśli $\lambda > |k|$, to warunek (12.5) jest spełniony i nasz cel został osiągnięty.

Zanotujmy, że wzór (12.6) pozwala znaleźć jawny wzór na f będące rozwiązaniem (12.9). Można też je dość łatwo zgadnąć: jest nim

$$f(t) = g(t) + k\int_0^t \mathrm{e}^{k(t-s)}g(s)\,\mathrm{d}s.$$

W ogólności jednak rozwinięcie (12.6) nie prowadzi do wzorów jawnych, więc rola warunku (12.5) rzeczywiście sprowadza się do zapewnienia istnienia i jedyności rozwiązań (12.7).

12.5 Zastosowanie: eksponenta operatora

Zupełność algebry $\mathcal{L}(\mathbb{X})$ zapewnia też istnienie operatorowej funkcji wykładniczej, zgodnie przez analityków uważanej za jedną z najważniejszych funkcji dzisiejszej analizy, a być może całej matematyki [27]. Pamiętając, że dla rzeczywistego (a nawet zespolonego) a i dla $t \in \mathbb{R}$ funkcję $\mathbb{R} \ni t \mapsto \mathrm{e}^{ta}$ definiuje się najwygodniej jako

$$\mathrm{e}^{ta} = \sum_{k=0}^{\infty} \frac{t^k a^k}{k!},$$

mając dany operator $A \in \mathcal{L}(\mathbb{X})$ oraz $t \in \mathbb{R}$, rozważamy

$$S_n(t) = \sum_{k=0}^{n} \frac{t^k A^k}{k!}.$$

Chcielibyśmy oczywiście udowodnić, że ciąg $(S_n(t))_{n\ge1}$ jest zbieżny, a e^{tA} zdefiniować jako jego granicę.

Gdybyśmy nie wiedzieli, że $\mathcal{L}(\mathbb{X})$ jest przestrzenią Banacha, bylibyśmy w kropce. Bo jak udowodnić zbieżność powyższego ciągu? Jeśli chcemy pokazać, że coś jest jego granicą, to sprawdzamy, czy odległości elementów ciągu od tego „cosia” dążą do zera. Ale tu nie mamy „cosia”, od którego mielibyśmy liczyć odległości, i koło się zamyka. Zupełność natomiast pozwala stwierdzić istnienie granicy na podstawie analizy odległości elementów ciągu między sobą. Biorąc więc $n > m$ szacujemy odległość

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^{n} \frac{t^k A^k}{k!} \right\| \le \sum_{k=m+1}^{n} \left\| \frac{t^k A^k}{k!} \right\| \le \sum_{k=m+1}^{n} \frac{|t|^k \|A^k\|}{k!}$$
$$\le \sum_{k=m+1}^{n} \frac{|t|^k \|A\|^k}{k!};$$

wykorzystaliśmy tu nierówność trójkąta oraz (12.17). Ta ostatnia suma dąży do zera, gdy $m, n \to \infty$, bo jest to różnica między n-tą i m-tą sumą częściową szeregu e^b dla liczby rzeczywistej $b = |t|\,\|A\|$. I koniec: zupełność dowodzi istnienia granicy ciągu $(S_n(t))_{n\ge 1}$, a to pozwala nam zdefiniować

$$\mathrm{e}^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

Przykład 12.11. Niech $\mathbb{X} = \mathbb{R}^2$ i niech naszym operatorem będzie macierz $A = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$, w której liczby $a \ge 0$ i $b \ge 0$ spełniające warunek $a + b = 1$ są dane. Pokażemy, że

$$\mathrm{e}^{tA} = \begin{pmatrix} b + a\,\mathrm{e}^{-t} & a - a\,\mathrm{e}^{-t} \\ b - b\mathrm{e}^{-t} & b\mathrm{e}^{-t} + a \end{pmatrix}, t \in \mathbb{R}. \tag{12.10}$$

W tym celu sprawdzamy, że wielomian charakterystyczny A równy jest $\lambda(\lambda + a + b) = \lambda(\lambda + 1)$, więc jej wartościami charakterystycznymi są 0 i -1. Obliczając następnie wektory własne, przekonujemy się, że A można sprowadzić do następującej postaci diagonalnej:

$$A = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix}.$$

To dowodzi, że

$$A^i = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & (-1)^i \end{pmatrix} \begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix}, \qquad i \ge 1,$$

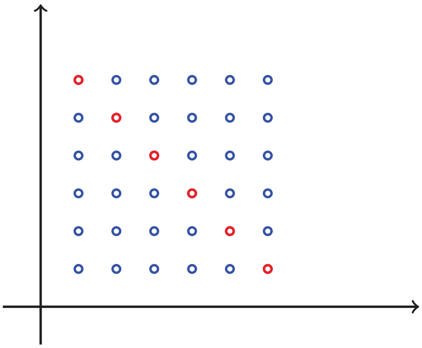
oraz

$$\sum_{i=0}^{n} \frac{t^i}{i!} A^i = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sum_{i=0}^{n} \frac{t^i(-1)^i}{i!} \end{pmatrix} \begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix}, \qquad n \ge 1,$$

skąd już tylko krok do

$$\mathrm{e}^{tA} = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-t} \end{pmatrix} \begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix}.$$

Wymnażając powyższe macierze, otrzymujemy (12.10).



Rys. 12.2: Sumowanie w (12.11) odbywa się po całym „kwadracie” a w (12.12) – tylko po jego części, po „trójkącie” leżącym poniżej czerwonej diagonali.

Rozwiązywanie równań różniczkowych w przestrzeni Banacha

Jednym z zasadniczych powodów, dla których funkcja wykładnicza jest tak ważna jest to, że opisuje rozwiązania równań różniczkowych. Rozwiązaniem najprostszego równania liniowego:

$$u'(t) = au(t), \qquad u(0) = x,$$

w którym a jest ustaloną liczbą rzeczywistą, x (które też jest liczbą) – warunkiem początkowym, a u – szukaną funkcją, jest

$$u(t) = \mathrm{e}^{ta}x.$$

Okazuje się, że podobne własności ma operatorowa funkcja wykładnicza, a to pozwala w prosty sposób rozwiązywać równania, które wcale proste nie są.

Dowiedziemy tego wyniku, przechodząc przez serię lematów. Aż do końca tego podrozdziału $\mathbb{X}$ jest przestrzenią Banacha, a A elementem $\mathcal{L}(\mathbb{X})$.

Lemat 12.1. *Dla dowolnych $s, t \in \mathbb{R}$ zachodzi zależność*

$$\mathrm{e}^{tA}\mathrm{e}^{sA} = \mathrm{e}^{(s+t)A}.$$

Dowód. Ustalmy n i rozważmy iloczyn sum częściowych

$$S_n(t)S_n(s) = \sum_{k=0}^{n} \frac{t^k A^k}{k!} \sum_{k=0}^{n} \frac{s^k A^k}{k!} = \sum_{0\le k,l\le n} \frac{t^k s^l A^{k+l}}{k!\,l!}. \tag{12.11}$$

Gdy $n \to \infty$, dąży on oczywiście do $\mathrm{e}^{tA}\mathrm{e}^{sA}$. Rozważmy więc podobną sumę (zob. rysunek 12.2):

$$R_n = \sum_{\substack{0\le k,l\le n\\ 0\le k+l\le n}} \frac{t^k s^l A^{k+l}}{k!\,l!}, \tag{12.12}$$

która, skoro dla liczb dodatnich warunek $k+l \le n$ implikuje $k, l \le n$, jest równa

$$\sum_{k,l\ge 0, k+l\le n} \frac{t^k s^l A^{k+l}}{k!\, l!} = \sum_{m=0}^{n} \sum_{k=0}^{m} \frac{t^k s^{m-k} A^m}{k!(m-k)!}$$

(myślimy o m jako o sumie $k+l$ i sumujemy po diagonalach równoległych do tej czerwonej na rysunku 12.2). Korzystając ze wzoru Newtona:

$$\sum_{k=0}^{m} \frac{t^k s^{m-k}}{k!(m-k)!} = \frac{(t+s)^m}{m!},$$

widzimy, że

$$R_n = \sum_{m=0}^{n} \frac{(t+s)^m A^m}{m!},$$

a zatem $R_n = S_n(t+s)$ i wobec tego $\lim_{n\to\infty} R_n = \mathrm{e}^{(t+s)A}$.

Pozostaje nam udowodnić, że różnica między $S_n(t)S_n(s)$ i R_n dąży do zera, gdy $n \to \infty$. Norma tej różnicy nie przekracza

$$\begin{aligned}
&\sum_{\substack{0\le k,l\le n\\ k+l\ge n+1}} \left\| \frac{t^k s^l A^{k+l}}{k!\, l!} \right\| && \text{(z nierówności trójkąta)}\\
\le &\sum_{k+l\ge n+1} \left\| \frac{t^k s^l A^{k+l}}{k!\, l!} \right\| && \text{(ta suma ma więcej składników)}\\
\le &\sum_{k+l=n+1}^{\infty} \frac{|t|^k |s|^l \|A\|^{k+l}}{k!\, l!}\\
\le &\sum_{m=n+1}^{\infty} \sum_{k=0}^{n+1} \frac{|t|^k |s|^{n+1-k} \|A\|^m}{k!(n+1-k)!} && \text{(ta sztuczka już była)}\\
\le &\sum_{m=n+1}^{\infty} \frac{(|t|+|s|)^m \|A\|^m}{m!}.
\end{aligned}$$

Ostatnie wyrażenie dąży do 0, gdy $n \to \infty$, bo jest to różnica między n-tą sumą częściową szeregu $\mathrm{e}^{(|t|+|s|)\|A\|}$, a jego pełną sumą (mamy tu do czynienia z szeregami liczbowymi). To kończy dowód. □

Lemat 12.2. *Funkcja $t \mapsto \mathrm{e}^{tA}$ jest funkcją ciągłą. Ciągła jest też funkcja $t \mapsto A\mathrm{e}^{tA} = \mathrm{e}^{tA}A$.*

Dowód. Zaczynamy oczywiście od części pierwszej.

1. Dla każdego n, $S_n(t)$ jest funkcją ciągłą (jako wielomian argumentu t).

2. Ustalmy $T > 0$. Mamy

$$\sup_{t\in[-T,T]} \|S_n(t) - \mathrm{e}^{tA}\| = \sup_{t\in[-T,T]} \left\| \sum_{k=n+1}^{\infty} \frac{t^k A^k}{k!} \right\|$$
$$\leq \sup_{t\in[-T,T]} \sum_{k=n+1}^{\infty} \frac{|t|^k \|A\|^k}{k!} \leq \sum_{k=n+1}^{\infty} \frac{T^k \|A\|^k}{k!}.$$

   Argumentując tak, jak pod koniec poprzedniego lematu, przekonujemy się, że ostatnie wyrażenie dąży do zera, gdy $n \to \infty$. Dowodzi to, że $S_n(t)$ dąży do e^{tA} jednostajnie ze względu na t w przedziale $[-T, T]$.

3. Granica jednostajna funkcji ciągłych jest funkcją ciągłą. Wobec 1. i 2., $t \mapsto \mathrm{e}^{tA}$ jest ciągła na przedziale $[-T, T]$. Dowolność T kończy rozumowanie.

Druga część wynika z oszacowań $\|A\mathrm{e}^{tA} - A\mathrm{e}^{sA}\| \leq \|A\|\|\mathrm{e}^{tA} - \mathrm{e}^{sA}\|$ i $\|\mathrm{e}^{tA}A - \mathrm{e}^{sA}A\| \leq \|A\|\|\mathrm{e}^{tA} - \mathrm{e}^{sA}\|$. □

Lemat 12.3. *Dla każdego t zachodzi tożsamość*

$$I + \int_0^t A\mathrm{e}^{sA}\,\mathrm{d}s = \mathrm{e}^{tA}.$$

Co za tym idzie funkcja $t \mapsto \mathrm{e}^{tA}$ jest różniczkowalna i mamy

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\ \mathrm{e}^{tA} = \mathrm{e}^{tA}A.$$

Dowód. Zacznijmy od przeprowadzenia następujących rachunków (wykorzystujących zadanie 6.3):

$$I + \int_0^t AS_n(s)\,\mathrm{d}s = I + \int_0^t \sum_{k=0}^{n} \frac{s^k A^{k+1}}{k!}\,\mathrm{d}s = I + \sum_{k=0}^{n} \int_0^t \frac{s^k A^{k+1}}{k!}\,\mathrm{d}s$$
$$= I + \sum_{k=0}^{n} \frac{t^{k+1} A^{k+1}}{(k+1)!} = S_{n+1}(t).$$

By otrzymać pierwszą tożsamość, wystarczy zatem udowodnić, że $\int_0^t AS_n(s)\,\mathrm{d}s$ dąży do $\int_0^t A\mathrm{e}^{sA}\,\mathrm{d}s$ gdy $n \to \infty$. Oszacowania są podobne jak poprzednio: norma różnicy między tymi całkami równa jest

$$\left\| \int_0^t A\,[S_n(s) - \mathrm{e}^{sA}]\,\mathrm{d}s \right\| \leq \|A\| \left| \int_0^t \|S_n(s) - \mathrm{e}^{sA}\|\,\mathrm{d}s \right|$$
$$\leq \|A\|\,|t| \sup_{s\in[0,|t|]} \sum_{k=n+1}^{\infty} \frac{|s|^k \|A\|^k}{k!}$$
$$\leq \|A\|\,|t| \sum_{k=n+1}^{\infty} \frac{|t|^k \|A\|^k}{k!},$$

(pierwsza nierówność wynika z (6.5) i (12.17)) i ostatnie wyrażenie dąży do zera, gdy $n \to \infty$.

W części drugiej dla prostoty ograniczymy się do dowodu prawostronnej różniczkowalności (znane twierdzenie mówi, że jeśli prawostronna pochodna jest funkcją ciągłą, to jest też pochodną „pełną", a ciągłość $t \mapsto Ae^{tA}$ gwarantuje nam poprzedni lemat). Z pierwszej części wynika, że dla dowolnego t i $h > 0$ mamy

$$\frac{e^{(t+h)A} - e^{tA}}{h} = \frac{1}{h}\int_t^{t+h} Ae^{sA}\, ds,$$

a zatem

$$\frac{e^{(t+h)A} - e^{tA}}{h} - Ae^{tA} = \frac{1}{h}\int_t^{t+h} [Ae^{sA} - Ae^{tA}]\, ds.$$

Jako że funkcja $t \mapsto Ae^{sA}$ jest ciągła (zob. lemat 12.2), mając dane $\epsilon > 0$, możemy tak dobrać h, by norma funkcji podcałkowej po prawej stronie nie przekraczała ϵ dla $s \in [t, t+h]$. Dla takiego h norma lewej strony nie przekracza

$$\frac{1}{h} \times (\text{max. normy funkcji podcałkowej}) \times (\text{długość przedziału}) \le \frac{1}{h}\epsilon h = \epsilon.$$

To kończy dowód (fakt, że $Ae^{tA} = e^{tA}A$ pozostawiamy do uzasadnienia Czytelnikowi). □

Dotarliśmy wreszcie do miejsca, w którym możemy poznać związek funkcji wykładniczych z rozwiązaniami liniowych i autonomicznych równań różniczkowych w przestrzeniach Banacha.

Twierdzenie 12.3. *Niech x będzie elementem przestrzeni Banacha $\mathbb{X}$, a A liniowym, ograniczonym operatorem w tej przestrzeni. Dla dowolnego $T > 0$ równanie różniczkowe*

$$u'(t) = A\, u(t), \qquad u(0) = x, \quad t \in [0, T]$$

ma dokładnie jedno rozwiązanie dane wzorem $u(t) = e^{tA}x$.

Dowód. Fakt, że funkcja $u(t) = e^{tA}x$ jest rozwiązaniem jest bezpośrednim wnioskiem z lematu 12.3. Pozostaje do udowodnienia jednoznaczność. Załóżmy więc, że $t \mapsto u(t)$ jest rozwiązaniem naszego równania, ustalmy $s \in (0, T]$ i rozważmy funkcję $t \mapsto v(t) = e^{(s-t)A}u(t)$, zdefiniowaną na $[0, s]$. Jest ona różniczkowalna, a jej pochodna wynosi

$$\begin{aligned}\frac{d}{dt}\left[e^{(s-t)A}u(t)\right] &= -Ae^{(s-t)A}u(t) + e^{(s-t)A}u'(t)\\ &= -Ae^{(s-t)A}u(t) + e^{(s-t)A}Au(t).\end{aligned}$$

Wyrażenie to jest równe zero, gdyż $e^{(s-t)A}A = Ae^{(s-t)A}$. A zatem funkcja u jest stała. Stąd $u(s) = v(s) = v(0) = e^{sA}u(0) = e^{sA}x$. Wobec dowolności s twierdzenie jest udowodnione. (To, że $u(0) = e^{0A}x = x$ wynika z założenia, iż u jest rozwiązaniem). □

Uwaga 12.1. Zauważmy, że powyższe twierdzenie poświęcone jest rozwiązaniom istniejącym tylko dla $t \geq 0$. To sytuacja typowa w zastosowaniach (biologii, fizyce itd) i teorii równań różniczkowych. Rozwiązania rozwijane wstecz często nie mają dobrego sensu fizycznego czy interpretacji biologicznej. Teoria matematyczna jest tu zgodna z życiem, gdyż operatory występujące po prawej stronie równania najczęściej są nieograniczone i nasze twierdzenie się nie stosuje (*vide* równanie ciepła (11.9)). Trzymamy się tej konwencji (to znaczy rozważamy tylko $t \geq 0$), by przyzwyczajać Czytelnika do tej ogólniejszej sytuacji. Z drugiej strony wypada podkreślić, że gdy – tak jak u nas – operator jest ograniczony, rozwiązania istnieją też na półosi lewej (zob. też przykłady podane niżej).

Przykład 12.12. Łącząc wzór (12.11) z zadaniem 12.7 widzimy, że eksponenta macierzy

$$A = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix},$$

w której $a, b > 0$, dana jest wzorem

$$\mathrm{e}^{tA} = \frac{1}{a+b}\begin{pmatrix} b + a\ \mathrm{e}^{-(a+b)t} & a - a\ \mathrm{e}^{-(a+b)t} \\ b - b\mathrm{e}^{-(a+b)t} & b\mathrm{e}^{-(a+b)t} + a \end{pmatrix}, t \in \mathbb{R}.$$

Jeśli $\mathbb{R}^2$ wyposażymy w normę $\|(\xi_1, \xi_2)\| = |\xi_1| + |\xi_2|$ i utożsamimy macierz A z operatorem danym wzorem

$$A\ (\xi_1, \xi_2) = (\xi_1, \xi_2) \cdot A\ ,$$

w którym $\cdot$ oznacza mnożenie macierzowe, to okaże się, że rozwiązaniem równania

$$u'(t) = Au(t), \qquad u(0) = (\xi_1, \xi_2),$$

jest

$$(\xi_1, \xi_2) \cdot \mathrm{e}^{tA}. \tag{12.13}$$

Wynik ten ma ładną interpretację probabilistyczną. Otóż jeśli ξ_1, ξ_2 są liczbami nieujemnymi sumującymi się do jedynki, to wektor (ξ_1, ξ_2) można interpretować jako rozkład prawdopodobieństwa pewnego łańcucha Markowa w czasie zero, a wektor po prawej stronie omawianego równania – jako rozkład prawdopodobieństwa w czasie zero $t \geq 0$. Wspomniany łańcuch ma dwa stany, powiedzmy 1 i 2 (ξ_1 jest zatem prawdopodobieństwem, że łańcuch Markowa startuje ze stanu 1, a ξ_2 – że startuje ze stanu 2). Jeśli proces znajdzie się w pierwszym z nich, to pozostaje tam przez wykładniczy czas z parametrem a (tak więc prawdopodobieństwo, że czas spędzony w 1 będzie większy niż t wynosi e^{-at}), a następnie skacze do 2. Tam pozostaje przez czas wykładniczy z parametrem b, by skoczyć z powrotem do 1, itd. Jak napisałem wyżej, wyrażenie

$$\xi_1 \left(\frac{b}{a+b} + \frac{a}{a+b}\mathrm{e}^{-(a+b)t} \right) + \xi_2 \left(\frac{b}{a+b} - \frac{b}{a+b}\mathrm{e}^{-(a+b)t} \right),$$

które jest pierwszą współrzędną iloczynu (12.13), jest prawdopodobieństwem, że taki łańcuch w czasie t znajdzie się w stanie 1.

Formalnie rzecz biorąc, powyższy wzór podaje też rozwiązania dla $t < 0$, ale w takim przypadku podana wyżej interpretacja traci sens.

Przykład 12.13. Wykorzystamy otrzymane wyżej wyniki do rozwiązania równania różniczkowego

$$u''(t) = -\alpha^2 u, \qquad t \geq 0 \tag{12.14}$$

z warunkami początkowymi $u(0) = \xi_1, u'(0) = \xi_2$. Jest to równanie drugiego rzędu i w tej postaci nie mieści się w ramach twierdzenia 12.3. Zamienimy je więc na równoważne mu równanie rzędu pierwszego w $\mathbb{R}^2$. Jeśli wprowadzimy oznaczenie $v = u'$, to (12.14) przyjmie postać

$$\begin{cases} u' = v \\ v' = u'' = -\alpha u \end{cases} \qquad \text{albo równoważnie} \quad \begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = A \begin{pmatrix} u \\ v \end{pmatrix}. \tag{12.15}$$

Zgodnie z twierdzeniem 12.3 (jednoznacznie wyznaczone) rozwiązania mają postać

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \mathrm{e}^{tA} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \qquad t \geq 0.$$

By znaleźć jawną postać e^{tA} zauważamy, że

$$A^2 = \begin{pmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{pmatrix} = -\alpha^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -\alpha^2 I,$$

co z kolei implikuje poprzez indukcję

$$A^{2n} = (-\alpha^2)^n I \qquad A^{2n+1} = (-\alpha^2)^n A, \qquad n \geq 1.$$

Stąd już wnioskujemy, że

$$\begin{aligned} \mathrm{e}^{tA} &= \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} A^{2n} + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} A^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{t^{2n}(-\alpha^2)^n}{(2n)!} I + \sum_{n=0}^{\infty} \frac{t^{2n+1}(-\alpha^2)^n}{(2n+1)!} A = \cos \alpha t I + \frac{\sin \alpha t}{\alpha} A. \end{aligned}$$

Reasumując,

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} \cos \alpha t & \frac{\sin \alpha t}{\alpha} \\ -\alpha \sin \alpha t & \cos \alpha t \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

W szczególności więc rozwiązaniem równania (12.14) jest

$$u(t) = \xi_1 \cos \alpha t + \xi_2 \frac{\sin \alpha t}{\alpha}$$

(wzór na v nie jest nam już potrzebny, bo v jest po prostu pochodną z u).

Przykład 12.14. Rozważmy przestrzeń $C[0,\infty]$ funkcji ciągłych, które mają granice w nieskończoności oraz operatory A i B dane wzorami:

$$A = B - \lambda I, \quad Bf(x) = \lambda f(x+1), \qquad x \in [0,\infty],$$

w których λ jest daną liczbą dodatnią. Z drugiej części zadania 12.7 wynika, że

$$\mathrm{e}^{tA} = \mathrm{e}^{-\lambda t}\mathrm{e}^{tB}.$$

Jako że $B^n f(x) = \lambda^n f(x+n)$, mamy

$$\mathrm{e}^{tB} f(x) = \sum_{n=0}^{\infty} \frac{\lambda^n t^n}{n!} f(x+n) \qquad \text{oraz} \qquad \mathrm{e}^{tA} f(x) = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t}\frac{\lambda^n t^n}{n!} f(x+n), \tag{12.16}$$

dla $x \geq 0, t \geq 0$. Przypomnijmy (zob. [10]), że $\mathrm{e}^{-\lambda t}\frac{\lambda^n t^n}{n!}$ jest prawdopodobieństwem, iż proces Poissona $\{N(t), t \geq 0\}$ o intensywności λ przyjmuje w czasie t wartość n. Tak więc ostatecznie rozwiązaniem równania różniczkowego

$$u'(t) = Au(t), \qquad u(0) = f$$

w $C[0,\infty]$ jest

$$u(t)(x) = \mathrm{e}^{tA} f(x) = \sum_{n=0}^{\infty} \Pr(N(t) = n) f(x+n) = E\, f(x + N(t)), \qquad x \geq 0, t \geq 0$$

(litera E jest tu symbolem wartości oczekiwanej). Zwróćmy uwagę na to, że ten ostatni wzór ma sens tylko dla $t \geq 0$, podczas gdy (12.16) tego ograniczenia nie posiada.

12.6 Zadania

Zadanie 12.1. Sprawdź, które z operatorów omawianych we wstępie do tego wykładu są liniowe. Które z nich są ograniczone? Policz ich normy.

Zadanie 12.2. Policz normy operatorów ograniczonych rozważanych w podrozdziale 12.1 (część z tych norm została już znaleziona w tekście).

Zadanie 12.3. Niech $\mathbb{X}, \mathbb{Y}$ oraz $\mathbb{Z}$ będą przestrzeniami unormowanymi, a $A \in \mathcal{L}(\mathbb{X},\mathbb{Y})$ i $B \in \mathcal{L}(\mathbb{Y},\mathbb{Z})$ – liniowymi operatorami ograniczonymi. Operator C, zdefiniowany wzorem $Cx = B(Ax)$ i oznaczany BA, nazywamy złożeniem operatorów A i B. Pokaż, że BA jest również ograniczony, że jest elementem $\mathcal{L}(\mathbb{X},\mathbb{Z})$ oraz

$$\|BA\| \leq \|B\|\,\|A\|. \tag{12.17}$$

Zadanie 12.4. Dowiedź, że jeśli $x : [a,b] \mapsto \mathbb{X}$ jest ciągłą funkcją o wartościach w przestrzeni Banacha, a $A \in L(\mathbb{X},\mathbb{Y})$ jest ograniczonym operatorem liniowym odwzorowującym $\mathbb{X}$ w inną przestrzeń Banacha, to

$$A\int_a^b x(t)\,\mathrm{d}t = \int_a^b Ax(t)\,\mathrm{d}t. \tag{12.18}$$

Wskazówka: rozważ postać A na sumach Riemanna całki po lewej stronie.

Zadanie 12.5. Dowiedź, że jeśli $\mathbb{X}$ jest algebrą Banacha z jedynką u, to warunki $a \in \mathbb{X}$ i $\|a - u\| < 1$ implikują istnienie a^{-1} (to znaczy takiego elementu, że $a^{-1}a = aa^{-1} = u$) i oszacowanie $\|a^{-1} - u\| \leq \frac{\|a-u\|}{1-\|a-u\|}$. Ogólniej, jeśli a ma odwrotny a^{-1}, a b leży na tyle blisko a, że $\|a - b\| < \frac{1}{\|a^{-1}\|}$, to b też ma odwrotny i zachodzi nierówność

$$\|a^{-1} - b^{-1}\| \leq \frac{\|a^{-1}\|^2 \|a - b\|}{1 - \|a^{-1}\| \|a - b\|}.$$

Zadanie 12.6. Udowodnij ograniczoność operatorów B_1 i B_2 zdefiniowanych poprzez (12.8) i oblicz ich normy. Znajdź jeszcze przynajmniej jeden inny prawy odwrotny do A z omawianego tam przykładu.

Zadanie 12.7. Uzasadnij, że dla każdej liczby rzeczywistej a oraz operatora A zachodzą następujące związki między funkcjami wykładniczymi operatorów A, $A - aI$ oraz aA:

$$\mathrm{e}^{t(aA)} = \mathrm{e}^{atA} \qquad \text{i} \qquad \mathrm{e}^{t(A-aI)} = \mathrm{e}^{-at}\mathrm{e}^{tA}.$$

Wskazówka: przy dowodzie drugiego wzoru skorzystaj z twierdzenia 12.3.

Rozdział 13

Twierdzenie Banacha–Steinhausa

13.1 Zbieżność mocna a zbieżność w normie operatorowej

Opisana w poprzednim rozdziale zbieżność operatorów w normie, choć ze wszech miar użyteczna i elegancka, nie jest w stanie opisać nieco delikatniejszych twierdzeń granicznych. Na przykład, niech dla danego $t \geq 0$ operator $A(t)$ działający w l^1 będzie dany wzorem

$$A(t)\,(\xi_i)_{i\geq 1} = \left(\mathrm{e}^{-it}\xi_i\right)_{i\geq 1}. \tag{13.1}$$

Łatwo sprawdzić, że każdy z $A(t), t \geq 0$ ma normę nie przekraczającą 1. Można też pokazać, że dla każdego $x \in l^1$,

$$\lim_{t\to 0} \|A(t)x - x\| = \lim_{t\to 0} \sum_{i=1}^{\infty} |(\mathrm{e}^{-it} - 1)\xi_i| = 0.$$

Jest to bezpośredni wniosek z twierdzenia o zbieżności zmajoryzowanej, ale można też podać dowód elementarny: otóż powyższa suma równa jest

$$\sum_{i=1}^{\infty} (1 - \mathrm{e}^{-it})|\xi_i|,$$

a szereg $\sum_{i=1}^{\infty} |\xi_i|$ jest zbieżny, więc mając dane $\epsilon > 0$, możemy tak dobrać n, by

$$\sum_{i=n+1}^{\infty} (1 - \mathrm{e}^{-it})|\xi_i| \leq \sum_{i=n+1}^{\infty} |\xi_i| < \frac{\epsilon}{2}.$$

Teraz zaś możemy tak dobrać t, by $\sum_{i=1}^{n}(1 - \mathrm{e}^{-it})|\xi_i| < \frac{\epsilon}{2}$ (bo dla każdego $i = 1, ..., n$ mamy $\lim_{t\to 0} \mathrm{e}^{-it} = 1$). To kończy dowód.

Jak widzimy, operatory $A(t)$ w pewnym sensie, dość naturalnym zresztą, zbiegają do I. A jednak, dla każdego $t > 0$,

$$\|A(t) - I\| = \sup_{x \in l^1, \|x\|=1} \|A(t)x - x\| \geq \sup_{n \geq 1} \|A(t)e_n - e_n\| = \sup_{n \geq 1} (1 - \mathrm{e}^{-nt}) = 1.$$

Dowodzi to, że odległość $A(t)$ od I w sensie normy operatorowej jest stale równa 1, co pokazuje, że $A(t)$ w sensie tej normy do I zdążać nie może. (Ten sam wniosek można wysnuć z (12.4)). Mamy tu więc do czynienia z nieco innym, bardziej subtelnym, rodzajem zbieżności.

Definicja 13.1. *Niech $\mathbb{X}$ i $\mathbb{Y}$ będą przestrzeniami unormowanymi. Mówimy, że ciąg $(A_n)_{n \geq 1}$ elementów $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ zbiega do operatora A w mocnej topologii, jeśli dla każdego $x \in \mathbb{X}$ mamy $\lim_{n\to\infty} A_n x = Ax$; innymi słowy*

$$\lim_{n\to\infty} \|A_n x - Ax\|_{\mathbb{Y}} = 0.$$

Zbieżność w mocnej topologii jest słabsza niż zbieżność w normie: jeśli ciąg $(A_n)_{n \geq 1}$ zbiega w normie do A, to dla każego $x \in \mathbb{X}$ mamy

$$\|A_n x - Ax\| \leq \|A_n - A\| \, \|x\|_{\substack{\\ n\to\infty}} \longrightarrow 0,$$

co implikuje zbieżność w mocnej topologii. Jak widzieliśmy też w poprzednim przykładzie, może się zdarzyć, że ciąg operatorów $(A_n)_{n \geq 1}$ w przestrzeni Banacha $\mathbb{X}$ nie będzie zbiegał do operatora A w normie operatorowej:

$$\|A_n - A\| \underset{n\to\infty}{\not\to} 0,$$

a jednak dla każdego $x \in \mathbb{X}$,

$$\lim_{n\to\infty} \|A_n x - Ax\| = 0. \tag{13.2}$$

W sumie to nie dziwi, bo zbieżność w normie operatorowej to zbieżność jednostajna na każdej kuli – dlaczego wszystkie ciągi operatorów miałyby być zbieżne jednostajnie? Co ważniejsze, zbieżność w normie operatorowej (choć bardzo użyteczna i elegancka) okazuje się zdarzać stosunkowo rzadko – ciekawe twierdzenia analizy matematycznej, czy rachunku prawdopodobieństwa zwykle są zbyt subtelne, by ująć je w jej ramach. Jak zobaczymy podrozdziale 13.3, znane nam już twierdzenie Weiestrassa można sformułować jako twierdzenie o zbieżności w mocnej topologii, która nie jest zbieżnością w normie operatorowej.

13.2 Twierdzenie Banacha–Steinhausa

Wprowadzone wyżej pojęcie zbieżności mocnej wydaje się jak najbardziej naturalne. Ale przyglądając się definicji 13.1, natychmiast zaczynamy się zastanawiać, czy aby na pewno jest użyteczne. Powodem jest oczywiście to, że zbieżność punktowa ciągu funkcji ciągłych nie gwarantuje ciągłości funkcji granicznej: dowodzi tego znany powszechnie przykład funkcji $f_n(x) = x^n, x \in [0,1]$, które dążą do funkcji nieciągłej (jakiej?). Z góry więc nie możemy zakładać, że dla danego ciągu operatorów $(A_n)_{n\geq 1}$ istnienie, dla każdego $x \in \mathbb{X}$, granicy

$$\lim_{n\to\infty} A_n x =: Ax \tag{13.3}$$

definiującej nam operator A, gwarantuje tegoż operatora ciągłość. To dość kłopotliwe, prawda?

W kategoriach lekkiego cudu należy zatem myśleć o twierdzeniu mówiącym, że

w przestrzeniach Banacha zbieżność mocna gwarantuje ciągłość granicznego operatora

i podchodzić do niego z należnym szacunkiem. Dowodzi ono, że zbieżność mocna ma „przyjazne" własności. Raz jeszcze okazuje się, że zupełność jest kluczem do elegancji i użyteczności teorii.

By zrozumieć powód, dla którego twierdzenie powyższe jest prawdziwe, przyjrzyjmy się jego założeniom. Po pierwsze przyjmujemy milcząco, że mamy tu do czynienia z operatorami liniowymi. Podany wyżej przykład zbieżności punktowej, w której granica nie jest funkcją ciągłą oczywiście się tu nie stosuje, bo funkcje w nim występujące liniowe nie są. Co więcej, próby zbudowania przykładu jednowymiarowego spełzają na niczym: jak łatwo się przekonać, punktowa zbieżność ciągu funkcji liniowych:[15]

$$f_n(x) = a_n x + b_n$$

pociąga za sobą zbieżność ciągów liczbowych $(a_n)_{n\geq 1}$ i $(b_n)_{n\geq 1}$, a zatem zbieżność jednostajną na dowolnym przedziale domkniętym $x \in [\alpha, \beta] \subset \mathbb{R}$. Liniowość (lub bliskość liniowości) jest więc tu założeniem kluczowym. Poniższy przykład dowodzi jednak, że niewystarczającym.

Rozważmy mianowicie przestrzeń c_{00} ciągów $x = (\xi_i)_{i\geq 1}$, które od pewnego (zależnego od ciągu) miejsca równe są zero, oraz ciąg operatorów (a dokładniej: funkcjonałów, to znaczy operatorów, których wartościami są liczby) działających w tej przestrzeni, dany wzorem

$$A_n\,(\xi_i)_{i\geq 1} = \sum_{i=1}^{n} \xi_i \qquad n \geq 1.$$

Ciąg ten, jak łatwo sprawdzić, zdąża w mocnej topologii do granicy

$$A\ \ (\xi_i)_{i\geq 1} = \sum_{i=1}^{\infty} \xi_i.$$

[15]Zdefiniowane tu funkcje f_n nie są operatorami liniowymi nad przestrzenią Banacha $\mathbb{R}$, chyba że $b_n = 0$. Są im jednak dostatecznie bliskie, by nie mogły służyć za kontrprzykład.

W istocie dla dowolnego $x \in c_{00}$ i dla dowolnego $n \geq n_0(x)$, gdzie $n_0(x)$ jest miejscem, od którego ciąg x składa się z samych zer, mamy $A_n x = Ax$. Operator graniczny A nie jest jednak ograniczony (a zatem nie jest ciągły): wektor x_j, który na pierwszych j współrzędnych ma jedynki, a poza tym składa się z samych zer, ma normę równą jeden, a $Ax_j = j$, więc $\sup_{\|x\|=1} |Ax| = \infty$.

Tak więc w twierdzeniu w ramce konieczne jest zarówno założenie liniowości operatorów jak i zupełności przestrzeni, w której działają (rozważana wyżej przestrzeń c_{00} nie jest przestrzenią Banacha).[16] Jak zobaczymy, jest ono bezpośrednim wnioskiem z następującego twierdzenia Banacha–Steinhausa (zwanego również zasadą wspólnej ograniczoności).

Twierdzenie 13.1 (Banacha–Steinhausa)**.** *Załóżmy, że operatory A_n działające w przestrzeni Banacha $\mathbb{X}$, o wartościach w przestrzeni unormowanej $\mathbb{Y}$, są liniowe i ograniczone. Jeśli dla każdego $x \in \mathbb{X}$,*

$$\sup_{n\geq 1} \|A_n x\| < \infty \tag{13.4}$$

to

$$\sup_{n\geq 1} \|A_n\| < \infty. \tag{13.5}$$

Podkreślmy, że norma w warunku (13.4) pochodzi z przestrzeni $\mathbb{Y}$, a ta w (13.5) – z przestrzeni $\mathcal{L}(\mathbb{X}, \mathbb{Y})$.

Podawane w większości podręczników dowody powyższego twierdzenia są zwykle dość skomplikowane – opierają się albo na zaawansowanym twierdzeniu Baira o kategoriach, albo na „metodzie przesuwającego się garbu" [4, 5, 27]. Elementarne rozumowanie, które przedstawiamy niżej pochodzi z niedawnego artykułu amerykańskiego matematyka A. D. Sokala [30].

Lemat 13.1. *Jeśli A jest operatorem ograniczonym, to dla dowolnego $x \in \mathbb{X}$ i $r > 0$ zachodzi warunek*

$$r\|A\| \leq \sup_{y\in K(x,r)} \|Ay\|,$$

gdzie $K(x,r)$ oznacza kulę o środku x i promieniu r (zbiór tych $y \in \mathbb{X}$, dla których $\|x-y\| \leq r$).

Dowód. Jeśli $\|z\| \leq 1$, to $\|rz\| \leq r$ i mamy

$$2rz \ = rz - x + x + rz,$$

co sprawia, że

$$r\|Az\| \leq \frac{1}{2}\| - A(x-rz)\| + \frac{1}{2}\|A(x+rz)\|.$$

Zarówno $x - rz$ jak i $x + rz$ należą do $K(x,r)$, prawa strona zatem nie przekracza $\sup_{y\in K(x,r)} \|Ay\|$. Licząc supremum po $\|z\| = 1$, otrzymujemy tezę. □

[16] Za radą recenzenta bądźmy bardziej precyzyjni: pokazaliśmy, że bez *jakichś* dodatkowych założeń twierdzenie w ramce nie jest prawdziwe. Twierdzenie Banacha–Steinhausa dowodzi, że liniowość i zupełność z powodzeniem odgrywają tę rolę.

Wniosek 13.1. *Niezależnie od tego jak wybiorę x i $r > 0$, w kuli o promieniu r i środku w x znajdę takie y, że $\|Ay\| \geq \frac{2}{3}r\|A\|$.*

Dowód twierdzenia Banacha–Steinhausa. Załóżmy, że teza nie jest prawdziwa. Dla każdego n znajdziemy zatem takie $m = m(n)$, że $\|A_m\| \geq 4^n$. By nie wprowadzać dodatkowego zamieszania związanego ze skomplikowanymi oznaczeniami, założymy bez straty ogólności, iż $\|A_n\| \geq 4^n$ – jeśli warunek ten nie jest spełniony, wystarczy rozważyć ciąg $B_n := A_{m(n)}$.

Kluczem do dowodu jest konstrukcja następującego ciągu $(x_n)_{n\geq 0}$: przyjmujemy, że $x_0 = 0$, a mając już dane x_{n-1}, definiujemy x_n jako taki element kuli $K(x_{n-1}, \frac{1}{3^n})$, że

$$\|A_n x_n\| \geq \frac{2}{3}\frac{1}{3^n}\|A_n\|;$$

wektor taki można znaleźć na podstawie wniosku podanego wyżej.

Twierdzimy, że ciąg ten jest ciągiem Cauchy'ego. Rzeczywiście dla $m > n$,

$$\begin{aligned}\|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + ... + \|x_{n+1} - x_n\| \\ &\leq \frac{1}{3^m} + ... + \frac{1}{3^{n+1}} < \frac{\frac{1}{3^{n+1}}}{1-\frac{1}{3}} = \frac{1}{2}\frac{1}{3^n}.\end{aligned}$$

Skoro $\mathbb{X}$ jest przestrzenią Banacha, istnieje granica $x = \lim_{n\to\infty} x_n$. Z powyższego oszacowania wynika też, że $\|x - x_n\| \leq \frac{1}{2}\frac{1}{3^n}$. Stąd,

$$\begin{aligned}\|A_n x\| &\geq \|A_n x_n\| - \|A_n(x - x_n)\| \\ &\geq \frac{2}{3}\frac{1}{3^n}\|A_n\| - \frac{1}{2}\frac{1}{3^n}\|A_n\| = \frac{1}{6}\frac{1}{3^n}\|A_n\|,\end{aligned}$$

bo przecież $\|A_n(x - x_n)\| \leq \frac{1}{2}\frac{1}{3^n}\|A_n\|$. Ale to dowodzi, że

$$\|A_n x\| \geq \frac{1}{6}\left(\frac{4}{3}\right)^n,$$

co stoi w jawnej sprzeczności z (13.4). Ta sprzeczność dowodzi tezy. □

Dowód „twierdzenia w ramce" jest już teraz bardzo prosty. Chcemy udowodnić, że operator

$$Ax := \lim_{n\to\infty} A_n \tag{13.6}$$

jest liniowy i ograniczony. Pierwsza cecha wynika wprost z liniowości operacji liczenia granicy i liniowości operatorów aproksymujących A_n:

$$\begin{aligned}A(\alpha x + \beta y) &= \lim_{n\to\infty} A_n(\alpha x + \beta y) = \lim_{n\to\infty} (\alpha A_n x + \beta A_n y) \\ &= \alpha \lim_{n\to\infty} A_n x + \beta \lim_{n\to\infty} A_n y = \alpha Ax + \beta Ay.\end{aligned}$$

Dla dowodu ograniczoności A zaś rozumujemy tak: nierówność

$$|\|A_n x\| - \|Ax\|| \leq \|A_n x - Ax\|$$

powoduje, że zbieżność ciągu wektorów $(A_n x)_{n\geq 1}$ pociąga za sobą zbieżność ciągu ich norm $(\|A_n x\|)_{n\geq 1}$ (do $\|Ax\|$). Każdy ciąg zbieżny jest jednak ograniczony. Tak więc istnienie granicy (13.6) implikuje założenie (13.4) twierdzenia Banacha–Steinhausa, z którego wnioskujemy, że $M = \sup_{n\geq 1} \|A_n\|$ jest mniejsze od nieskończoności. Dla tego M i dowolnego $x \in \mathbb{X}$ mamy też

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| \leq \limsup_{n\to\infty} \|A_n\| \, \|x\| \leq M\|x\|,$$

co kończy dowód.

13.3 Wielomiany Bernsteina

Opisany w rozdziale 7 dowód twierdzenia Weierstrassa jest niekonstruktywny w tym sensie, że nie podaje receptury na znalezienie konkretnego ciągu wielomianów aproksymującego daną funkcję ciągłą. Przepis taki podał Siergiej Bernstein w roku 1912. Pokazał on mianowicie, że dla każdej funkcji $f \in C[0,1]$ skonstruowane na jej podstawie wielomiany

$$w_{n,f}(x) = \sum_{k=1}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k},$$

zdążają, gdy $n \to \infty$, do wyjściowego $f \in C[0,1]$ – nazywamy je wielomianami Bernsteina.

Popatrzymy na to twierdzenie z punktu widzenia zbieżności operatorów. Niech mianowicie A_n będzie operatorem przyporządkowującym funkcji odpowiadający jej n-ty wielomian Bernsteina:

$$A_n f(x) = w_{n,f}(x) = \sum_{k=1}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Łatwo stwierdzamy, że operatory $A_n, n \geq 1$, są ograniczone:

$$\begin{aligned} \|A_n f\| &= \sup_{x\in[0,1]} |A_n f(x)| \leq \sup_{x\in[0,1]} \sum_{k=1}^{n} \left| f\left(\frac{k}{n}\right)\right| \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \|f\| \sup_{x\in[0,1]} \sum_{k=1}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \|f\| \quad \text{(skutek wzoru Newtona).} \end{aligned}$$

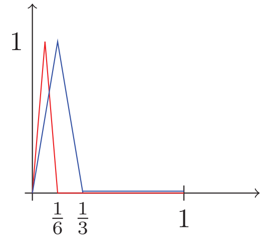
Co więcej, jeśli f_0 jest stale równa jeden, to $Af_0 = f_0$ (i oczywiście $\|f_0\| = 1$) a to dowodzi, że nie tylko $\|A_n\| \leq 1$ ale wręcz $\|A_n\| = 1$.

Pokażemy, że dla każdego f

$$\lim_{n\to\infty} \|A_n f - f\| = 0, \tag{13.7}$$

(co jest po prostu innym sformułowaniem twierdzenia Weiestrassa w wersji Bernsteina), podczas gdy dla wszystkich n

$$\|A_n - I\| \geq 1. \tag{13.8}$$



Rys. 13.1: Funkcja f_3 (kolor niebieski) i f_6 (kolor czerwony); na przedziale $[\frac{1}{3}, 1]$ funkcje te się pokrywają.

Ta ostatnia nierówność wyklucza zbieżność ciągu $(A_n)_{n\geq 1}$ w normie operatorowej. Wyjaśnijmy to dokładniej: zależność (13.7), gdy już ją udowodnimy, przekona nas, iż jedyną możliwą granicą $(A_n)_{n\geq 1}$ w normie operatorowej jest I, bo jak pamiętamy, zbieżność w normie pociąga za sobą zbieżność mocną do tego samego operatora. Z drugiej strony I granicą tą być nie może ze względu na (13.8).

Jak widzimy, klasyczne twierdzenie analizy matematycznej, kamień węgielny wielu rozumowań nie daje się opisać przy pomocy zbieżności w normie – występuje tu typ zbieżności bardziej subtelny: zbieżność mocna.

Uważny Czytelnik zauważył zapewne, że wielomiany Bernsteina są ściśle powiązane z rozkładem dwumianowym. Przypomnijmy, że mówimy, iż zmienna losowa X ma rozkład dwumianowy z parametrami n i x jeśli

$$\Pr(X = k) = \binom{n}{k} x^k (1-x)^{n-k}, \qquad \text{dla} \qquad k = 0, \ldots, n;$$

opisuje ona ilość sukcesów w n niezależnych próbach Bernoulliego, gdy prawdopodobieństwo sukcesu w jednej próbie wynosi $x \in [0, 1]$. W tych terminach $A_n f$ przyjmuje postać

$$A_n f(x) = E\, f(n^{-1} X_n),$$

w której X_n jest zmienną losową dwumianową o parametrach n i x.

Przypomnijmy też, że wartością oczekiwaną dwumianowej zmiennej losowej o parametrach n i p jest np, a jej wariancją $np(1-p)$ (zob. np [10]). Do dowodu (13.7) potrzebna nam będzie też nierówność Czebyszewa: jeśli X jest zmienną losową o wartości oczekiwanej μ i wariancji σ^2, to dla każdej liczby dodatniej δ

$$\Pr(|X - \mu| > \delta) \leq \frac{\sigma^2}{\delta^2}.$$

Zacznijmy od dowodu (13.8). Niech $f_n(x) = 0 \vee (1 - |2nx - 1|)$ (zob. rysunek 13.1). Jak widać $\|f_n\| = 1$, ale $f_n(j/n) = 0$, dla $j = 0, 1, ..., n$, co skutkuje tym, że $A_n f_n = 0$. Tak więc $\|A_n - I\| \geq \|A_n f_n - f_n\| = \|f_n\| = 1$. To dowodzi (13.8).

Przejdźmy do dowodu (13.7). Funkcja f, jako funkcja ciągła na przedziale domkniętym, jest jednostajnie ciągła. Dla danego $\epsilon > 0$ możemy więc tak dobrać δ, by

$|f(x)-f(y)|$ było mniejsze od $\frac{\epsilon}{2}$ dla wszystkich x, y z przedziału $[0,1]$ spełniających warunek $|x-y|<\delta$. Ze względu na fakt, iż

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}=1,$$

mamy też

$$f(x)=\sum_{k=0}^{n}f(x)\binom{n}{k}x^k(1-x)^{n-k}.$$

A zatem wartość bezwzględna różnicy między $f(x)$ i $A_nf(x)$ nie przekracza

$$\sum_{k=1}^{n}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k}.$$

Wyrażenie to można rozbić na dwie sumy: na sumę Σ_1 po wskaźnikach spełniających nierówność $|\frac{k}{n}-x|\geq\delta$ i sumę Σ_2 po wskaźnikach pozostałych. Ze względu na dobór δ, Σ_2 nie będzie przekraczać $\frac{\epsilon}{2}$ pomnożonego przez sumę prawdopodobieństw odpowiednich wskaźników – skoro suma tych prawdopodobieństw nie przekracza 1, Σ_2 nie przekracza $\frac{\epsilon}{2}$. W Σ_1 z kolei występują pewne liczby mnożone przez prawdopodobieństwa. Łatwo zauważamy, że liczby te nie przekraczają $2\|f\|$, a suma prawdopodobieństw to prawdopodobieństwo $\Pr(|n^{-1}X_n-x|\geq\delta)$, które wobec nierówności Czebyszewa nie przekracza $\frac{x(1-x)}{n\delta^2}$ A zatem Σ_1 nie przekracza $2\|f\|\frac{x(1-x)}{n\delta^2}$. Ze względu na fakt, iż maksimum funkcji $x\mapsto x(1-x)$ w przedziale $[0,1]$ osiągane jest w $x=\frac{1}{2}$ i równe jest $\frac{1}{4}$, Σ_2 nie przekracza $\frac{\|f\|}{2n\delta^2}$. To dowodzi, że dla $n\geq\frac{\|f\|}{\epsilon\delta^2}$ norma f_n-f nie przekracza $\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon$, c.b.d.o.

13.4 Aproksymacja poissonowska

W poprzednich podrozdziałach dość buńczucznie przekonywałem Czytelnika, że zbieżność operatorów w sensie normy jest narzędziem zbyt brutalnym, by dobrze opisywać delikatną naturę niektórych twierdzeń. I miałem rację. Ale nie chciałbym pozostawić wrażenia, że przy jego pomocy nie da się zrobić nic ciekawego, bo to po prostu nieprawda. Dla przykładu pokażemy, że klasyczne twierdzenie rachunku prawdopodobieństwa, nazywane aproksymacją poissonowską, jest związane ze zbieżnością według normy właśnie.

Twierdzenie to mówi, że

jeśli X_n jest zmienną losową o rozkładzie dwumianowym z parametrami n i p_n, a parametry te dobrane są tak, że $\lim_{n\to\infty}np_n=\lambda>0$, to

$$\lim_{n\to\infty}\Pr(X_n=k)=\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}=:\lambda_k. \tag{13.9}$$

(Po prawej stronie równości występują tu prawdopodobieństwa rozkładu Poissona z parametrem λ).

By przepisać je w języku operatorów wyobraźmy sobie, że początkowo mamy pewną losową liczbę złotówek, a kolejne zdobywamy w niezależnych próbach Bernoulliego z prawdopodobieństwem sukcesu w jednej próbie równym p – sukces powoduje, że otrzymujemy dodatkową złotówkę, porażka – że nasz stan posiadania się nie zmienia. Jeśli $(\xi_i)_{i\geq 0}$ jest rozkładem początkowego kapitału, a więc

$$\xi_i = \Pr(\text{początkowy kapitał} = i),$$

to po pierwszej próbie

$$\xi_i' = \Pr(\text{nowy kapitał} = i) = p\xi_{i-1} + (1-p)\xi_i \qquad (\text{przyjmujemy } \ \xi_{-1} = 0).$$

Zapominając o tym, że $(\xi)_{i\geq 0}$ jest rozkładem prawdopodobieństwa i rozważając w zamian dowolny element l^1, otrzymujemy operator

$$A\,(\xi_i)_{i\geq 0} = (p\xi_{i-1} + (1-p)\xi_i)_{i\geq 0} \tag{13.10}$$

działający w tej przestrzeni. Jest to jak łatwo sprawdzić operator ograniczony o normie równej 1.

Dość oczywiste jest, że jeśli operatorem takim podziałamy dwukrotnie na rozkład początkowy $(\xi_i)_{i\geq 0}$, to znaczy policzymy $A^2\,(\xi_i)_{i\geq 0}$ to otrzymamy rozkład naszego kapitału po dwóch (niezależnych) próbach Bernoulliego. Ogólniej, $A^n\,(\xi_i)_{i\geq 0}$ jest rozkładem po n próbach. Innymi słowy, A^n jest operatorem związanym ze zmienną losową o rozkładzie dwumianowym i parametrach n i p: do początkowego kapitału dodajemy tyle złotówek, ile osiągnęliśmy sukcesów w n niezależnych próbach Bernoulliego.

Jak widać, naszym zadaniem jest zbadanie granicy operatorów $A_n^n, n\geq 1$, gdy A_n jest zdefiniowany wzorem (13.10) z p zamienionym na p_n. Chcielibyśmy mianowicie pokazać, że

$$\lim_{n\to\infty} A_n^n = P_\lambda, \tag{13.11}$$

gdzie P_λ jest operatorem związanym z rozkładem Poissona z parametrem λ. P_λ opisuje więc zmianę rozkładu naszego kapitału w doświadczeniu, w którym zyskujemy losową ilość złotówek o rozkładzie poissonowskim. Po takim doświadczeniu mamy i złotych, o ile wcześniej mieliśmy tyle, a zmienna losowa poissonowska równa jest zero, albo wcześniej mieliśmy $i-1$ a zmienna równa jest 1, itd., albo nie mieliśmy nic, a zmienna losowa wynosi i. To znaczy (zob (13.9))

$$\xi_i' = \xi_i\lambda_0 + \xi_{i-1}\lambda_1 + \cdots \xi_0\lambda_i,$$

co sprawia, że

$$P_\lambda\,(\xi_i)_{i\geq 0} = \left(\sum_{j=0}^{i} \xi_{i-j}\lambda_i\right)_{i\geq 0}.$$

Dowód zależności (13.11). Najważniejszym krokiem dowodu jest przedstawienie P_λ jako eksponenty pewnego operatora. Pokażemy mianowicie, że zachodzi zależność

$$P_\lambda = \mathrm{e}^{\lambda(P-I)}, \tag{13.12}$$

w której P jest operatorem przesunięcia w prawo: $P\,(\xi_i)_{i\geq 0} = (\xi_{i-1})_{i\geq 0}$, $(\xi_{-1} = 0)$. Opisuje on, swoją drogą, deterministyczny wzrost losowego kapitału o złotówkę.[17] Tożsamy z (13.12) wzór (użyj zadania 12.7)

$$P_\lambda = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} P^k$$

można zinterpretować następująco: operator po prawej stronie z prawdopodobieństwem $\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$ przesuwa ciągi z l^1 o k pozycji w prawo. Jest to jak widać równoważna definicja operatora związanego z rozkładem Poissona, a to dowodzi (13.12).

„Uzmienniając" λ, możemy popatrzeć na tę zależność nieco inaczej. Stwierdza ona mianowicie, że $\lambda \to P_\lambda$ jest „prawą częścią" eksponenty operatorowej $\lambda \mapsto \mathrm{e}^{\lambda(R-I)}$, to znaczy obcięciem tej ostatniej do prawej półosi. Gdy sprawdzimy (dość łatwo), że P_λ ma normę 1 dla każdego $\lambda > 0$, przekonamy się tym samym, że kontrakcjami są wszystkie operatory funkcji wykładniczej $\mathrm{e}^{\lambda(P-I)}$ dla $\lambda > 0$.

Skorzystamy także z następującej nierówności: jeśli $B_1, \cdots, B_n$ i $C_1, \cdots, C_n$ są kontrakcjami w przestrzeni Banacha $\mathbb{X}$, to (zob. zadanie 13.1)

$$\|B_n B_{n-1} \cdots B_1 - C_n C_{n-1} \cdots C_1\| \leq \sum_{i=1}^{n} \|B_i - C_i\|, \tag{13.13}$$

co w szczególności powoduje, że dla dowolnych kontrakcji B i C mamy

$$\|B^n - C^n\| \leq n\|B - C\|.$$

Teraz dowód jest już łatwy. Skoro

$$\begin{aligned} \|A_n^n - P_\lambda\| &= \|A_n^n - P_{\frac{\lambda}{n}}^n\| \qquad \text{(bo } \lambda \to P_\lambda \text{ jest częścią funkcji wykładniczej)} \\ &\leq n\|A_n - P_{\frac{\lambda}{n}}\| \\ &\leq \|n(A_n - I) - n(P_{\frac{\lambda}{n}} - I)\|, \end{aligned}$$

to wystarczy uzasadnić, że granicą $n(A_n - I)$ i $n(P_{\frac{\lambda}{n}} - I)$ (w sensie normy) jest ten sam operator. Z drugiej części lematu 12.3 wiemy jednak, że

$$\lim_{n\to\infty} n(P_{\frac{\lambda}{n}} - I) = \lambda \lim_{n\to\infty} \frac{n}{\lambda}(\mathrm{e}^{\frac{\lambda}{n}(R-I)} - I) = \lambda \lim_{t\to 0} \frac{1}{t}(\mathrm{e}^{t(R-I)} - I) = \lambda(R - I),$$

natomiast z definicji A_n mamy $A_n = p_n R + (1 - p_n)I$, więc

$$n(A_n - I) = np_n(R - I) \longrightarrow \lambda(R - I),$$

a to kończy dowód. □

[17] Uśmiech na naszych twarzach zniknie, gdy zdamy sobie sprawę z tego, że operator ten opisuje także dynamikę rozkładu naszego wieku, jakichkolwiek byśmy nie przyjęli jednostek.

By zobaczyć zależność między (13.11) a (13.9) rozważmy szczególny wektor, nazwijmy go e_0, którego wszystkie współrzędne są równe zero, poza pierwszą, która równa jest 1. Odpowiada on sytuacji, gdy nasz początkowy kapitał wynosi (deterministycznie, niestety) zero. $A_n^n e_0$ jest zatem rozkładem dwumianowym: po doświadczeniu mamy tyle złotówek ile było sukcesów w n próbach Bernoulliego. Z (13.11) wnioskujemy, że $A_n^n e_0$ dąży do $P_\lambda e_0$ to znaczy do rozkładu Poissona. Skoro zbieżność w l^1 pociąga za sobą zbieżność po współrzędnych, otrzymujemy stąd (13.9). Ale nasz wynik jest w istocie mocniejszy: pokazaliśmy mianowicie, że rozkład dwumianowy dąży do rozkładu Poissona w normie przestrzeni l^1 (podczas gdy (13.9) dotyczy zbieżności po współrzędnych). Ten sam wniosek można wysnuć z tak zwanego twierdzenia Scheffé (zob. [4]). Nasz wynik jest jednak jeszcze mocniejszy: dowiedliśmy, że operatory A_n^n dążą do P_λ jednostajnie na całej kuli jednostkowej w przestrzeni l^1. A tego łatwo z (13.9) nie widać.

Zadanie 13.1. Stosując indukcję, udowodnij (13.13).

Rozdział 14

Zbiór zadań

14.1 Zadania

14.1.1 Przestrzenie Banacha

Zadanie 14.1. Niech $C[0,2]$ oznacza przestrzeń funkcji ciągłych na przedziale $[0,2]$. Sprawdź, że jest to przestrzeń liniowa i unormowana z normą

$$\|f\| = \int_0^2 |f(x)|\,\mathrm{d}x.$$

Rozważając ciąg $f_n(x) = \min\{x^n, 1\}, x \in [0,2], n \geq 1$, sprawdź, że przestrzeń ta nie jest przestrzenią Banacha.

Zadanie 14.2. Sprawdź, że przestrzeń absolutnie sumowalnych ciągów $x = (\xi_i)_{i\geq 1}$ jest przestrzenią Banacha, jeśli wyposażymy ją w normę

$$\|x\| = \sup_{i=1,\ldots,1000} \frac{|\xi_i|}{i} + \sum_{i=1001}^{\infty} |\xi_i|.$$

Zadanie 14.3. Sprawdź, że przestrzeń c_{00} ciągów, których wszystkie współrzędne od pewnego miejsca równe są zero, nie stanie się przestrzenią Banacha, jeśli wprowadzimy w niej normę

$$\|(\xi_i)_{i\geq 0}\| = \sup_{i\geq 0}(i+1)^2|\xi_i|.$$

Zadanie 14.4. Sprawdź, że przestrzeń absolutnie sumowalnych ciągów $x = (\xi_i)_{i\geq 1}$ jest przestrzenią Banacha, jeśli wyposażymy ją w normę

$$\|x\| = \sum_{i=1}^{\infty} \frac{i}{i+10^{100}}|\xi_{2i}| + \sum_{i=0}^{\infty} |\xi_{2i+1}|.$$

Zadanie 14.5. Niech $C[0,2]$ oznacza przestrzeń funkcji ciągłych na przedziale $[0,2]$. Sprawdź, że jest to przestrzeń liniowa i unormowana z normą

$$\|f\| = \sup_{x\in[0,1]} |f(x)| + \int_1^2 |f(x)|\,\mathrm{d}x,$$

ale nie jest to przestrzeń Banacha.

Zadanie 14.6. Niech $\mathbb{X}$ oznacza przestrzeń funkcji ograniczonych, zdefiniowanych na zbiorze $S = [0,2] \cup \{5\}$, które mają tę własność, że $f(x) = 0$ dla $x \in [1,2]$. Czy z normą $\|f\| = \sup_{x\in[0,2]}|f(x)| + |f(5)|$ jest to przestrzeń Banacha?

Zadanie 14.7. Niech $\mathbb{X}$ będzie, wyposażoną w normę supremum, przestrzenią funkcji ograniczonych na przedziale $[0,3]$, które na przedziale $[0,1]$ są ciągłe, a dodatkowo spełniają warunek

$$f(x) = \frac{1}{2}(f(x-1) + f(x+1)), \qquad x \in [1,2].$$

Narysuj wykres typowego przedstawiciela $\mathbb{X}$. Udowodnij, że $\mathbb{X}$ jest przestrzenią Banacha.

Zadanie 14.8.

(a) Wykaż, że przestrzeń

$$\mathbb{X} = \left\{ (\xi_i)_{i\in\mathbb{N}} \in l^\infty \colon \sum_{i=1}^{\infty}|\xi_i| < \infty,\ \sum_{i=1}^{\infty}\xi_i = 0 \right\}$$

z normą taką jak w przestrzeni l^1, jest przestrzenią Banacha.

(b) Uzasadnij, że jeżeli w definicji przestrzeni $\mathbb{X}$ z punktu (a) zastąpimy warunek $\sum_{i=1}^{\infty}\xi_i = 0$, warunkiem

istnieje takie $i \geq 1$, *że* $\xi_i = 0$,

to $\mathbb{X}$ nie będzie przestrzenią Banacha.

Zadanie 14.9. Niech $\mathbb{X}$ oznacza przestrzeń funkcji ciągłych na przedziale $[0,1]$, których wykresy są symetryczne względem $\frac{1}{2}$ (tzn. $f(x) = f(1-x)$ dla $x \in [0,1]$) i które mają tę własność, że $f(0) = f(\frac{1}{2})$. Czy z normą supremum $\|f\| = \sup_{x\in[0,1]}|f(x)|$ jest to przestrzeń Banacha?

Zadanie 14.10. W przestrzeni $\mathbb{X}$ absolutnie sumowalnych ciągów $x = (\xi_i)_{i\geq 1}$ zdefiniujmy funkcję $\|\cdot\|$ wzorem

$$\|x\| = \sum_{i=1}^{\infty} \mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_i|.$$

Sprawdź, że jest to norma oraz wykaż, że przestrzeń $\mathbb{X}$ wyposażona w tę normę jest przestrzenią Banacha.

Zadanie 14.11. Udowodnij, że z zupełności przestrzeni liczb rzeczywistych wynika, iż każdy nierosnący ciąg ograniczony z dołu ma granicę. Udowodnij też, że to, iż każdy nierosnący ciąg ograniczony z dołu ma granicę pociąga za sobą zupełność przestrzeni liczb rzeczywistych.

Zadanie 14.12. Sprawdź, że istnieje dokładnie jedna taka funkcja ciągła f, zdefiniowana na przedziale $[0,1]$, iż

$$\bigwedge_{x\in[0,1]} \qquad f(x)+\int_0^x [\sin(xy)+\cos^4(x+y)]f(y)\,\mathrm{d}y=\frac{1}{1+x^2}.$$

Wskazówka: Udowodnij, że przestrzeń funkcji ciągłych na $[0,1]$ z normą typu Bieleckiego $\|f\|_\lambda=\sup_{x\in[0,1]}\mathrm{e}^{-\lambda x}|f(x)|$, jest dla dowolnego λ przestrzenią Banacha.

Zadanie 14.13. Udowodnij, że istnieje dokładnie jeden ciąg $(\xi_n)_{n\in\mathbb{Z}}$ spełniający warunki:

$$\sum_{n\in\mathbb{Z}}|\xi_n|<\infty, \quad (n^2+1)\xi_n+\tfrac{n^2\cos n}{2}\xi_{n-2}=\tfrac{(n^2+1)\sin n^2}{5}\xi_{n-1}+\tfrac{1}{8}\xi_{n+1}+1, \quad n\in\mathbb{Z}.$$

Zadanie 14.14. Udowodnij, że istnieje dokładnie jedna funkcja ciągła f na przedziale $[0,1]$, spełniająca zależność

$$f(x)+\sin x^2=\frac{1}{2}f\left(\frac{1}{3}x\right)+\frac{1}{3}f\left(\frac{1}{2}x\right)+\mathrm{e}^{\pi x}, \qquad x\in[0,1].$$

Zadanie 14.15. Sprawdź, że istnieje dokładnie jedna taka funkcja ciągła f, zdefiniowana na przedziale $[0,1]$, iż

$$\forall_{x\in[0,1)} \qquad f(x)-\int_0^x[\sin(x+y)+\cos^4(xy)]f(y)\,\mathrm{d}y=\frac{x^2-1}{x^2-3x+2}.$$

Zadanie rozwiąż dwoma sposobami.

Wskazówka: Skorzystaj z faktu, że przestrzeń funkcji ciągłych na $[0,1]$ z normą typu Bieleckiego $\|f\|_\lambda=\sup_{x\in[0,1]}\mathrm{e}^{-\lambda x}|f(x)|$, jest dla dowolnego λ przestrzenią Banacha.

Zadanie 14.16. Niech będą dane ciągi

$$a_n=\frac{1}{n^4+n^2-1} \quad \text{i} \quad b_n=\mathrm{e}^{-4}\frac{2^n}{n!}(-1)^{n+3}, \qquad n\geq 0.$$

Sprawdź, że w przestrzeni l^1 ciągów bezwzględnie sumowalnych istnieje dokładnie jeden ciąg $(\xi_n)_{n\geq 0}$ spełniający równanie

$$\xi_n+\sum_{k=0}^n\xi_k b_{n-k}=a_n, \qquad n\geq 0.$$

Zadanie rozwiąż dwiema metodami.

Zadanie 14.17. Udowodnij, że istnieje dokładnie jeden ciąg $(\xi_n)_{n\geq 1}$ spełniający warunki:

$$\sum_{n=1}^{\infty} |\xi_n| < \infty, \qquad n\xi_{n+1} = 4n^2\xi_n + 2\xi_{n+2} + \sin n, \quad n \geq 1.$$

14.1.2 Przestrzenie Hilberta

Zadanie 14.18. W przestrzeni l^2 wyznacz jawny wzór rzutu ortogonalnego na podprzestrzeń liniową

$$A \;= \big\{(\xi_n)_{n\geq 0} \in l^2 \colon \xi_{2n} = 2\xi_{2n+1},\ n \geq 0\big\}.$$

Zadanie 14.19. Niech $\mathbb{X}$ oznacza przestrzeń ciągów $x = (\xi_i)_{i\geq 1}$, dla których szereg $\sum_{i=1}^{\infty} i\ \xi_i^2$ jest zbieżny.

(A) Wprowadź tak iloczyn skalarny, by $\mathbb{X}$ stała się przestrzenią Hilberta z normą $\|x\| = \sqrt{\sum_{i=1}^{\infty} i\ \xi_i^2}$.

(B) Który z następujących zbiorów jest podprzestrzenią $\mathbb{X}$? (Odpowiedź dokładnie uzasadnij).

   (a) $\{x \in \mathbb{X} \colon \|x\| = 1\}$,

   (b) $\{x = (\xi_i)_{i\geq 1} \in \mathbb{X} \colon \xi_4 + \xi_9 = 0\}$.

(C) Znajdź wzór na rzut wektora $(\eta_i)_{i\geq 1}$ na podprzestrzeń z poprzedniego punktu.

Zadanie 14.20. Niech $a_n, n \geq 1$, będzie ciągiem liczb dodatnich. Załóżmy, że istnieją takie stałe $0 < c < C < \infty$, iż $c < a_n < C$ dla każdego $n \geq 1$. Niech $\mathbb{H}$ będzie przestrzenią Hilberta ciągów liczbowych $x = (\xi_i)_{i\geq 1}$, dla których

$$\|x\|^2 := \sum_{i\geq 1} a_i\xi_i^2 < \infty.$$

(a) Podaj wzór na iloczyn skalarny w tej przestrzeni.

(b) Sprawdź, że zbiór $\mathbb{H}_0$ złożony z tych elementów $\mathbb{H}$, dla których zachodzą zależności:

$$\xi_8 = \xi_5, \qquad \xi_{121} + \xi_{500} = 0,$$

jest podprzestrzenią $\mathbb{H}$.

(c) Znajdź wzór na rzut ortogonalny na podprzestrzeń $\mathbb{H}_0$.

14.1.3 Operatory liniowe

Zadanie 14.21. Operator $A\colon \mathbb{R}^3 \to \mathbb{R}^2$ dany jest wzorem

$$A(x, y, z) = (x - z, x + y).$$

Wyznacz normę operatora A , jeżeli w $\mathbb{R}^3$ określimy normę $\|(x, y, z)\| = |x| + |y| + |z\ |$, a w przestrzeni $\mathbb{R}^2$ normę $\|(x, y)\| = \max(|x|, |y|)$.

Zadanie 14.22. Niech c_{00} oznacza przestrzeń ciągów $x = (\xi_n)_{n\geq 1}$, które od pewnego miejsca (zwykle różnego dla różnych ciągów) są równe zero, z normą $\|x\| = \sup_{n\geq 1} |\xi_n|$. Które z następujących operatorów liniowych przekształcających tę przestrzeń w siebie są ograniczone? Oblicz ich normy.

$$\begin{aligned} A\ (\xi_n)_{n\geq 1} &= (0, 2\xi_2, 0, 4\xi_4, 0, 6\xi_6, ...),\\ B(\xi_n)_{n\geq 1} &= (\xi_2, \xi_1, \xi_4, \xi_3, \xi_6, \xi_5, ...),\\ C(\xi_n)_{n\geq 1} &= (\sum_{n=1}^{\infty} \xi_{2n}, 0, \sum_{n=1}^{\infty} \xi_{2n+1}, 0, 0, 0, 0, ...). \end{aligned}$$

Zadanie 14.23. Który z następujących operatorów nie jest liniowy? Który jest liniowy, ale nieograniczony? Odpowiedzi dokładnie uzasadnij. Znajdź normę operatora, który jest liniowy i ograniczony. We wszystkich przypadkach rozważamy operatory w przestrzeni c_{00} ciągów, które od pewnego miejsca równe są stale zero (z normą supremum).

(a) $A(\xi_i)_{i\geq 1} = (i\xi_i)_{i\geq 1}$.

(b) $B(\xi_i)_{i\geq 1} = (\xi_i \sin \frac{i\pi}{12})_{i\geq 1}$.

(c) $C(\xi_i)_{i\geq 1} = (1, \xi_1, \xi_2, ...)$.

Zadanie 14.24. Niech $(a_i)_{n\geq 1}$ będzie takim ciągiem liczb rzeczywistych z przedziału $(0, 1)$, że $\lim_{i\to\infty} a_i = 1$. Zdefiniujmy operator $A\colon l^1 \to \mathbb{R}$ wzorem

$$Ax = \sum_{i=1}^{\infty} a_i \xi_i, \qquad \text{gdzie } x = (\xi_i)_{i\geq 1} \in l^1.$$

Sprawdź, że $\|A\ \| = 1$.

Zadanie 14.25. Niech operator $A\colon C[0, 1] \to C[0, 1]$ będzie dany wzorem

$$(Af)(x) = \int_0^x \frac{t}{1+t} f(t^2) dt, \qquad f \in C[0, 1].$$

Oblicz normę operatora A .

Zadanie 14.26. W przestrzeni $C[-\infty, \infty]$ funkcji ciągłych o granicach w plus i minus nieskończoności (z normą supremum) rozważamy operator

$$A\ f(x) = \int_{-\infty}^{\infty} \frac{1}{1+y^2} f(x+y)\, \mathrm{d}y, \qquad x \in \mathbb{R}.$$

Udowodnij, że jest to operator ograniczony i znajdź jego normę.

Zadanie 14.27. Wykaż, że operator $A\colon C^2[0,1] \to C[0,1]$ dany wzorem $Af = f''$, $f \in C^2[0,1]$, nie jest ograniczony, jeżeli w przestrzeni $C^2[0,1]$ definiujemy normę $\|f\|_{C^2[0,1]} = \|f\|_{C[0,1]} + \|f'\|_{C[0,1]}$, a w $C[0,1]$ rozważamy normę standardową.

Zadanie 14.28. W przestrzeni funkcji ciągłych na przedziale $[0,1]$ rozważamy operator dany wzorem

$$Af = (\int_0^1 f(y)\,\mathrm{d}y) f_*,$$

gdzie $f_*(x) = x^2$. Znajdź wzór na $\mathrm{e}^{tA}, t \geq 0$.

Zadanie 14.29. W przestrzeni $C[0,\infty]$ funkcji ciągłych na $[0,\infty)$ i posiadających granice w nieskończoności (z normą supremum) rozważamy operator

$$Af = f(0) f_*,$$

gdzie $f_*(x) = 2\mathrm{e}^{-x}, x \geq 0$. Sprawdź, że jest to operator ograniczony i udowodnij, że

$$\mathrm{e}^{tA} f = f + \frac{f(0)}{2}(\mathrm{e}^{2t} - 1) f_*, \qquad t \in \mathbb{R}.$$

Zadanie 14.30. Niech $L^1[0,1]$ będzie przestrzenią funkcji bezwzględnie całkowalnych na przedziale jednostkowym, wyposażoną w normę $\|f\| = \int_0^1 |f(x)|\,\mathrm{d}x$. Sprawdź, że operator dany wzorem

$$Af(x) = f(x) + f(1-x) + f\left(\frac{1}{2}x\right), \quad x \in [0,1],$$

jest ograniczony i oblicz jego normę.

Zadanie 14.31. Niech $L^1[0,1]$ będzie przestrzenią funkcji bezwzględnie całkowalnych na przedziale jednostkowym, wyposażoną w normę $\|f\| = \int_0^1 |f(x)|\,\mathrm{d}x$. Sprawdź, że operator dany wzorem

$$Af(x) = f(x) + f(1-x) + f\left(\frac{1}{3}x + \frac{2}{3}\right), \quad x \in [0,1],$$

jest ograniczony i oblicz jego normę.

Zadanie 14.32. Niech $C_0(0,1]$ oznacza, wyposażoną w normę supremum, przestrzeń funkcji ciągłych na przedziale $[0,1]$, które przyjmują wartość zero w $x = 0$. Każdą z funkcji $f \in C_0(0,1]$ rozszerzamy do funkcji $\tilde{f}$ zdefiniowanej na całej osi liczbowej za pomocą następującego wzoru:

$$\tilde{f}(x) = \begin{cases} f(x), & x \in [0,1], \\ f(1), & x > 1, \\ -f(-x), & x < 0. \end{cases}$$

Udowodnij, że operatory A_n, zdefiniowane wzorem

$$A_n f(x) = \frac{1}{2}\Big[\tilde{f}(x + n^{-1}) + \tilde{f}(x - n^{-1})\Big],$$

zbiegają w mocnej topologii do operatora identycznościowego (tzn. że $\lim_{n\to\infty} \|A_n f - f\| = 0$ dla każdego $f \in C_0(0,1]$), ale nie jest prawdą, że $\lim_{n\to\infty}\|A_n - I\,\| = 0$.

Zadanie 14.33. Niech $a_n, n \geq 1$, będzie dążącym do 1 ciągiem liczb spełniających warunek $a_n \in (0,1)$. W przestrzeni $C[0,1]$ ze standardową normą supremum definiujemy operatory

$$A_n f(x) = f(1 - a_n(1-x)), \qquad n \geq 1.$$

Udowodnij, że operatory A_n dążą do I w mocnej topologii, ale nie w normie operatorowej.

Zadanie 14.34. Niech $C[0,1]$ będzie przestrzenią funkcji ciągłych na odcinku $[0,1]$. Zdefiniujmy operator $T_n\colon C[0,1] \to C[0,1]$, $n = 1, 2, \ldots$, wzorem

$$T_n f(x) = \begin{cases} f\big(\sqrt[n]{3}x\big), & x \in [0, 1/\sqrt[n]{3}], \\ f(1), & x \in (1/\sqrt[n]{3}, 1]. \end{cases}$$

(a) Wykaż, że dla dowolnego n operator T_n jest ograniczony oraz oblicz jego normę.

(b) Sprawdź, że dla dowolnej funkcji $f \in C[0,1]$ zachodzi

$$\lim_{n\to\infty} T_n f = f,$$

jednak nie jest prawdą, że

$$\lim_{n\to\infty} \|T_n - I\,\| = 0,$$

gdzie $I = I_{C[0,1]}$ jest operatorem identycznościowym w $C[0,1]$.

Zadanie 14.35. Niech $C[0,1]$ będzie przestrzenią Banacha funkcji ciągłych na przedziale $[0,1]$, które mają tę własność, że $f(0) = f(1)$. Operator A działający w tej przestrzeni dany jest wzorem

$$Af(x) = -2f(x \oplus 0.25),$$

w którym $\oplus$ oznacza dodawanie modulo 1.

(a) Opisz działanie operatora A.

(b) Sprawdź jego ograniczoność i oblicz normę.

(c) Oblicz e^A.

Zadanie 14.36. W przestrzeni l^1 rozważamy operatory $T_n, n \geq 1$, dane wzorem

$$T_n(\xi_i)_{i\geq 1} = \left(\sum_{k=1}^{n} \xi_i, \xi_1 + \xi_2, \xi_1 - \xi_2, \xi_3 + \xi_4, \xi_3 - \xi_4, \xi_5 + \xi_6, \xi_5 - \xi_6, \ldots\right),$$

w którym $(\xi_i)_{i\geq 1} \in l^1$, a $n \geq 1$. Oblicz ich normy i znajdź ich granicę w mocnej topologii. Udowodnij, że nie jest ona granicą w normie operatorowej.

14.2 Rozwiązania

Zadanie 14.1. Sprawdzenie warunków przestrzeni liniowej nie sprawia trudności.

Pokażemy, że funkcja $\|\cdot\|\colon C[0,2] \to \mathbb{R}$ jest normą. Załóżmy, że $\|f\| = 0$. Wtedy $\int_0^2 |f(x)|\,\mathrm{d}x = 0$. Gdyby dla pewnego punktu $x_0 \in [0,2]$ było $|f(x_0)| > 0$, to z ciągłości funkcji f istniałby taki przedział $(a,b) \ni x_0$, że $|f(x)| > 0$ dla $x \in (a,b)$. Stąd $\|f\| \geq \int_a^b |f(x)|\,\mathrm{d}x > 0$. Otrzymana sprzeczność dowodzi, że $|f(x)| = 0$ dla $x \in [0,2]$, skąd oczywiście wynika, że $f(x) = 0$ dla $x \in [0,2]$.

Nierówność trójkąta oraz jednorodność wynikają wprost z odpowiednich własności całki.

Wykażemy teraz, że przestrzeń $C[0,2]$ z tak zdefiniowaną normą nie jest przestrzenią Banacha. Sprawdźmy, że ciąg $(f_n)_{n\geq 1}$ spełnia warunek Cauchy'ego. Załóżmy, bez zmniejszania ogólności rozważań, że $n \geq m$. Wtedy

$$\begin{aligned}\|f_n - f_m\| &= \int_0^2 |\min(x^n,1) - \min(x^m,1)|\,\mathrm{d}x = \\ &= \int_0^1 (x^m - x^n)\,\mathrm{d}x = \frac{1}{m+1} - \frac{1}{n+1} \leq \frac{1}{m+1}.\end{aligned}$$

Dowodzi to, że ciąg $(f_n)_{n\geq 1}$ jest ciągiem Cauchy'ego, gdyż dla dowolnego $\epsilon > 0$ oraz liczby naturalnej $N > \frac{1}{\epsilon} - 1$ mamy $\frac{1}{m+1} < \epsilon$, o ile $m > N$.

Pozostaje nam uzasadnić, że ciąg $(f_n)_{n\geq 1}$ nie jest zbieżny w tej przestrzeni. Przypuśćmy, że istnieje takie $f \in C[0,2]$, że $\|f_n - f\| \to 0$, gdy $n \to \infty$. Zauważmy, że dla dowolnego punktu $x_0 \in (1,2)$ musi być $f(x_0) = 1$. W przeciwnym wypadku, na mocy ciągłości funkcji f, byłoby $f(x) < 1$ (bądź, analogicznie, > 1) na pewnym przedziale $[a,b] \subset (1,2)$ zawierającym x_0. Wtedy

$$\|f_n - f\| \geq \int_a^b (1 - f(x))\,\mathrm{d}x \geq (b-a) \inf_{x\in[a,b]} (1 - f(x)).$$

Liczba $\inf_{x\in[a,b]}(1 - f(x))$ jest większa od 0 na mocy założenia oraz niezależna od n. Otrzymujemy zatem sprzeczność z $\|f_n - f\| \to 0$, gdy $n \to \infty$, co pokazuje, że $f(x) = 1$ dla $x \in (1,2)$.

Podobnie, załóżmy, że dla pewnego punktu $x_0 \in (0,1)$ zachodzi $f(x_0) > 0$. Istnieje wtedy taki przedział $[a,b] \subset (0,1)$, że $f(x) > 0$ dla $x \in [a,b]$. Ponieważ ciąg $(f_n)_{n\geq 1}$ zbiega jednostajnie na $[a,b]$ do funkcji stale równej 0, to dla dowolnego $0 < \epsilon < \inf_{x\in[a,b]} f(x)$ oraz odpowiednio dużych n mamy

$$\begin{aligned}\|f_n - f\| &\geq \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x \geq (b-a) \inf_{x\in[a,b]} |f_n(x) - f(x)| = \\ &= (b-a) \inf_{x\in[a,b]} (f(x) - f_n(x)) > (b-a)\Big(\inf_{x\in[a,b]} f(x) - \epsilon\Big).\end{aligned}$$

Otrzymana sprzeczność dowodzi, że $f(x) = 0$ dla $x \in (0,1)$. Ostatecznie otrzymujemy zatem przeczący ciągłości f wzór:

$$f(x) = \begin{cases} 0, & x \in (0,1), \\ 1, & x \in (1,2). \end{cases}$$

Zadanie 14.2. Oznaczmy badaną przestrzeń przez $\mathbb{X}$. Z definicji przestrzeni Banacha musimy wykazać, że dowolny ciąg $(x_n)_{n\geq 1} \subset \mathbb{X}$ spełniający warunek Cauchy'ego

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|x_n - x_m\|_{\mathbb{X}} < \epsilon, \tag{14.1}$$

jest zbieżny do pewnego elementu $x \in \mathbb{X}$. Załóżmy zatem, że ciąg $(x_n)_{n\geq 1}$ w przestrzeni $\mathbb{X}$ spełnia warunek Cauchy'ego. Oznaczmy $x_n = (\xi_{n,i})_{i\geq 1}$ dla $n \geq 1$ i przepiszmy warunek (14.1) następująco

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_{m,i}|}{i} + \sum_{i=1001}^{\infty} |\xi_{n,i} - \xi_{m,i}| < \epsilon. \tag{14.2}$$

Krok 1 — zbieżność „po współrzędnych" ciągu $(x_n)_{n\geq 1}$.

Niech $i \in \{1, 2, \ldots, 1000\}$ będzie ustalone. Wtedy

$$\frac{|\xi_{n,i} - \xi_{m,i}|}{i} \leq \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_{m,i}|}{i} < \epsilon, \qquad n \geq 1,$$

a zatem $|\xi_{n,i} - \xi_{m,i}| \leq i\,\epsilon$. Z dowolności ϵ wynika, że ciąg liczbowy $(\xi_{n,i})_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$. Ma on zatem granicę, którą oznaczmy $\xi_i = \lim_{n\to\infty} \xi_{n,i}$, $i = 1, 2, \ldots, 1000$. Podobnie, jeżeli $i \geq 1001$, to

$$|\xi_{n,i} - \xi_{m,i}| \leq \sum_{i=1001}^{\infty} |\xi_{n,i} - \xi_{m,i}| < \epsilon.$$

Stąd również ciąg $(\xi_{n,i})_{n\geq 1}$ (dla $i \geq 1001$) spełnia warunek Cauchy'ego w $\mathbb{R}$ i jego granicę oznaczamy jak wcześniej, $\xi_i = \lim_{n\to\infty} \xi_{n,i}$, $i \geq 1001$.

Dowiedliśmy zatem, że ciąg $(x_n)_{n\geq 1}$ jest zbieżny „po współrzędnych" do $x = (\xi_i)_{i\geq 1}$.

Krok 2 — ciąg $x = (\xi_i)_{i\geq 1}$ należy do przestrzeni $\mathbb{X}$.

Aby pokazać, że $x \in \mathbb{X}$ trzeba sprawdzić absolutną sumowalność ciągu x, to znaczy, że $\sum_{i=1}^{\infty} |\xi_i| < \infty$. Ponieważ na zbieżność szeregu wpływa tylko jego reszta, to $\sum_{i=1}^{\infty} |\xi_i| < \infty$ wtedy i tylko wtedy, gdy $\sum_{i=1001}^{\infty} |\xi_i| < \infty$.

Zauważmy, że warunek (14.2) implikuje

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{k\geq 1001} \sum_{i=1001}^{k} |\xi_{n,i} - \xi_{m,i}| < \epsilon. \tag{14.3}$$

Ponieważ N nie zależy od k, więc

$$\lim_{m\to\infty} \sum_{i=1001}^{k} |\xi_{n,i} - \xi_{m,i}| = \sum_{i=1001}^{k} |\xi_{n,i} - \xi_i| \leq \epsilon, \tag{14.4}$$

dla dowolnego $k \geq 1001$. Z nierówności trójkąta dla wartości bezwzględnej wynika zatem, że

$$\sum_{i=1001}^{k} |\xi_i| - \sum_{i=1001}^{k} |\xi_{n,i}| \leq \sum_{i=1001}^{k} |\xi_{n,i} - \xi_i| \leq \epsilon,$$

skąd

$$\sum_{i=1001}^{k} |\xi_i| \leq \sum_{i=1001}^{k} |\xi_{n,i}| + \epsilon.$$

Z dowolności $k \geq 1001$ widzimy, że szereg $\sum_{i=1001}^{k} |\xi_i|$ jest zbieżny, ponieważ $x_n \in \mathbb{X}$ (ciąg x_n jest absolutnie sumowalny). Zatem $x \in \mathbb{X}$, czego należało dowieść.

Krok 3 — zbieżność ciągu $(x_n)_{n\geq 1}$ do x w $\mathbb{X}$.

Pozostaje sprawdzić, że $\lim_{n\to\infty} \|x_n - x\|_{\mathbb{X}} = 0$, co tak naprawdę już (prawie) zrobiliśmy.

Podobnie jak w kroku drugim, warunek (14.2) implikuje

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{k\geq 1001} \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_{m,i}|}{i} < \epsilon.$$

Ponieważ supremum jest liczone po zbiorze skończonym, to również

$$\lim_{m\to\infty} \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_{m,i}|}{i} = \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_i|}{i} \leq \epsilon. \tag{14.5}$$

Sumując nierówności (14.3) i (14.5), otrzymujemy

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \sup_{i=1,\ldots,1000} \frac{|\xi_{n,i} - \xi_i|}{i} + \sum_{i=1001}^{k} |\xi_{n,i} - \xi_{m,i}| \leq 2\epsilon.$$

Lewa strona powyższej nierówności jest normą $\|x_n - x\|_{\mathbb{X}}$, a zatem z dowolności ϵ warunek ten implikuje zbieżność ciągu $(x_n)_{n\geq 1}$ do x w przestrzeni $\mathbb{X}$.

Zadanie 14.3. Dla danego $n \geq 1$ niech $x_n = (\xi_{n,i})_{i\geq 0}$ będzie dany wzorem

$$\xi_{n,i} = \begin{cases} \frac{1}{(i+1)^3}, & i \leq n, \\ 0, & i > n, \end{cases}$$

czyli

$$x_n = (1, \tfrac{1}{2^3}, \tfrac{1}{3^3}, \ldots, \tfrac{1}{n^3}, \ldots).$$

Wtedy oczywiście $x_n \in c_{00}$ dla $n \geq 1$. Ustalmy teraz dowolne $\epsilon > 0$. Dla $m, n \geq 1$, $m < n$, mamy

$$\begin{aligned} \|x_n - x_m\| &= \left\| \left(0, \ldots, 0, \frac{1}{(m+1)^3}, \ldots, \frac{1}{n^3}, 0, \ldots \right) \right\| = \\ &= (m+1)^2 \frac{1}{(m+1)^3} = \frac{1}{m+1}. \end{aligned}$$

Oznacza to, że ciąg $(x_n)_{n\geq 1}$ spełnia w przestrzeni c_{00} warunek Cauchy'ego. Przypuśćmy, że ma on granicę $x = (\xi_i)_{i\geq 0} \in c_{00}$. Zauważmy, że ciąg $(x_n)_{n\geq 1}$ jest zbieżny „po współrzędnych”. Istotnie, dla ustalonego $i \geq 0$ prawdziwa jest nierówność

$$|\xi_{n,i} - \xi_i| \leq \frac{1}{(i+1)^2} \|x_n - x\|.$$

Jeżeli zatem $x_n \to x$ dla $n \to \infty$ w c_{00}, to $\lim_{n\to\infty} \xi_{n,i} = \xi_i$ dla każdego $i \geq 0$. Jednak ciąg liczbowy $(\xi_{n,i})_{n\geq 1}$ jest ciągiem stałym od pewnego miejsca i $\lim_{n\to\infty} \xi_n^i = \frac{1}{(i+1)^3}$, skąd

$$x = \left(1, \frac{1}{2^3}, \frac{1}{3^3}, \dots\right).$$

Taki ciąg nie jest elementem przestrzeni c_{00}, co prowadzi do sprzeczności.

Zadanie 14.4. Oznaczmy badaną przestrzeń przez $(\mathbb{X}, \|\cdot\|)$. Z definicji przestrzeni Banacha, musimy wykazać, że dowolny ciąg $(x_n)_{n\geq 1} \subset \mathbb{X}$, spełniający warunek Cauchy'ego

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|x_n - x_m\| < \epsilon, \tag{14.6}$$

jest zbieżny do pewnego elementu $x \in \mathbb{X}$. Załóżmy zatem, że ciąg $(x_n)_{n\geq 1}$ spełnia w przestrzeni $\mathbb{X}$ warunek Cauchy'ego. Oznaczmy ponadto $x_n = (\xi_{n,i})_{i\geq 1}$ dla $n \geq 1$ i przepiszmy warunek (14.6) w postaci

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sum_{k=1}^{\infty} \frac{k}{k+10^{100}} |\xi_{n,2k} - \xi_{m,2k}| + \sum_{k=0}^{\infty} |\xi_{n,2k+1} - \xi_{m,2k+1}| < \epsilon. \tag{14.7}$$

Krok 1 — zbieżność „po współrzędnych" ciągu $(x_n)_{n\geq 1}$.

Niech $\epsilon' > 0$ oraz $i \geq 1$ będą dowolnie wybrane. Udowodnimy najpierw, że ciągi współrzędnych o numerach parzystych ciągu $(x_n)_{n\geq 1}$ są zbieżne w $\mathbb{R}$. Na mocy warunku (14.7) mamy

$$\frac{i}{i+10^{100}} |\xi_{n,2i} - \xi_{m,2i}| \leq \sum_{k=1}^{\infty} \frac{k}{k+10^{100}} |\xi_{n,2k} - \xi_{m,2k}| < \epsilon, \qquad m, n > N.$$

Stąd

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} |\xi_{n,2i} - \xi_{m,2i}| < \frac{i+10^{100}}{i} \epsilon.$$

Kładąc $\epsilon := \frac{i}{i+10^{100}} \epsilon'$, otrzymujemy, iż

$$\bigvee_{N\geq 1} \bigwedge_{m,n>N} |\xi_{n,2i} - \xi_{m,2i}| < \epsilon'.$$

Z dowolności ϵ' wynika, że ciąg liczbowy $(\xi_{n,2i})_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$. Ma on zatem granicę, którą oznaczmy $\xi_{2i} := \lim_{n\to\infty} \xi_{n,2i}$, $i \geq 1$.

Z drugiej strony, przechodząc do współrzędnych o numerach nieparzystych, dla $i \geq 0$ mamy

$$|\xi_{n,2i+1} - \xi_{m,2i+1}| \leq \sum_{k=0}^{\infty} |\xi_{n,2k+1} - \xi_{m,2k+1}| < \epsilon, \qquad m, n > N.$$

Widzimy zatem, że ciąg $(\xi_{n,2i+1})_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$ i jego granicę oznaczamy jak wcześniej, $\xi_{2i+1} := \lim_{n\to\infty} \xi_{n,2i+1}$, $i \geq 0$.

Dowiedliśmy zatem, że ciąg $(x_n)_{n\geq 1}$ jest zbieżny „po współrzędnych" do ciągu liczbowego $x = (\xi_i)_{i\geq 1} \subset \mathbb{R}$.

Krok 2 — ciąg $x = (\xi_i)_{i\geq 1}$ należy do przestrzeni $\mathbb{X}$.

Aby pokazać, że $x \in \mathbb{X}$, trzeba sprawdzić absolutną sumowalność ciągu x, to znaczy, że $\sum_{i=1}^{\infty}|\xi_i| < \infty$.

Zauważmy, że warunek (14.7) implikuje

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{k\geq 1} \sum_{i=1}^{k} \frac{i}{i+10^{100}}|\xi_{n,2i} - \xi_{m,2i}| + \sum_{i=0}^{k}|\xi_{n,2i+1} - \xi_{m,2i+1}| < \epsilon. \quad (14.8)$$

Ponadto, ponieważ

$$1 > \frac{i}{i+10^{100}} = 1 - \frac{10^{100}}{i+10^{100}} \geq 1 - \frac{10^{100}}{1+10^{100}} = \frac{1}{1+10^{100}},$$

to dla $k \geq 1$ mamy

$$\sum_{i=1}^{k} \frac{i}{i+10^{100}}|\xi_{n,2i} - \xi_{m,2i}| + \sum_{i=0}^{k}|\xi_{n,2i+1} - \xi_{m,2i+1}| \geq \frac{1}{1+10^{100}} \sum_{i=1}^{2k+1}|\xi_{n,i} - \xi_{m,i}|. \quad (14.9)$$

Ponieważ (suma jest skończona!)

$$\lim_{m\to\infty} \sum_{i=1}^{2k+1}|\xi_{n,i} - \xi_{m,i}| = \sum_{i=1}^{2k+1}|\xi_{n,i} - \xi_i|, \qquad k, n \geq 1,$$

to przechodząc w nierówności (14.9) do granicy przy $m \to \infty$, oraz wykorzystując (14.8), otrzymujemy

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \bigwedge_{k\geq 1} \frac{1}{1+10^{100}} \sum_{i=1}^{2k+1}|\xi_{n,i} - \xi_i| \leq \epsilon. \quad (14.10)$$

Wybierając $\epsilon = 1$, z nierówności trójkąta dla wartości bezwzględnej wiemy zatem, że

$$\sum_{i=1}^{2k+1}|\xi_i| - \sum_{i=1}^{2k+1}|\xi_{n,i}| \leq \sum_{i=1}^{2k+1}|\xi_{n,i} - \xi_i| \leq 1 + 10^{100}, \qquad k \geq 1,\ n > N.$$

Ponieważ $x_n \in \mathbb{X} \subset l^1$, to

$$\sum_{i=1}^{2k+1}|\xi_i| \leq \sum_{i=1}^{2k+1}|\xi_{n,i}| + 1 + 10^{100} \leq \|x_n\|_{l^1} + 1 + 10^{100}, \qquad k \geq 1,\ n > N.$$

Liczba po prawej stronie nierówności nie zależy od k, więc z dowolności $k \geq 1$ widzimy, że szereg $\sum_{i=1}^{\infty}|\xi_i|$ jest zbieżny. Stąd $x \in \mathbb{X}$, czego należało dowieść.

Krok 3 — zbieżność ciągu $(x_n)_{n\geq 1}$ do x w $\mathbb{X}$.

Pozostaje sprawdzić, że $\lim_{n\to\infty}\|x_n - x\| = 0$, co tak naprawdę już (prawie) zrobiliśmy.

Podobnie jak w kroku drugim, przechodząc w warunku (14.7) do granicy przy $m \to \infty$, otrzymujemy

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \bigwedge_{k\geq 1} \sum_{i=1}^{k} \frac{i}{i+10^{100}} |\xi_{n,2i} - \xi_{2i}| + \sum_{i=0}^{k} |\xi_{n,2i+1} - \xi_{2i+1}| \leq \epsilon.$$

Stąd

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \sum_{i=1}^{\infty} \frac{i}{i+10^{100}} |\xi_{n,2i} - \xi_{2i}| + \sum_{i=0}^{\infty} |\xi_{n,2i+1} - \xi_{2i+1}| \leq \epsilon.$$

Ostatecznie widzimy zatem, iż

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \|x_n - x\| \leq \epsilon,$$

co z definicji oznacza, że

$$\lim_{n\to\infty} x_n = x$$

w przestrzeni $\mathbb{X}$.

Uwaga. Wykorzystując jedną z nierówności z kroku drugiego, łatwo sprawdzić, że dla $x \in \mathbb{X}$ mamy

$$\frac{1}{1+10^{100}} \|x\|_{l^1} \leq \|x\| \leq \|x\|_{l^1}.$$

Pokazuje to, że normy $\|\cdot\|$ i $\|\cdot\|_{l^1}$ są równoważne, zatem dowód zupełności przestrzeni $\mathbb{X}$ można sprowadzić do dowodu zupełności przestrzeni l^1.

Zadanie 14.5. W badanej przestrzeni wprowadzamy naturalne działania

$$(f+g)(x) := f(x) + g(x), \qquad f, g \in C[0,2],\ x \in [0,2],$$

oraz

$$(\alpha f)(x) := \alpha f(x), \qquad f \in C[0,2], \alpha \in \mathbb{R},\ x \in [0,2].$$

Z tak wprowadzoną sumą wektorów oraz mnożeniem przez skalar, przestrzeń $C[0,2]$ jest przestrzenią liniową. Rzeczywiście, łączność, przemienność oraz prawa rozdzielności wynikają wprost z tych własności dla liczb rzeczywistych. Ponadto, funkcja stale równa zero jest elementem neutralnym, a funkcja $-f$, gdzie $(-f)(x) := -f(x)$, $x \in [0,2]$, elementem przeciwnym do f.

Sprawdzimy teraz, że funkcja $\|\cdot\|$ jest rzeczywiście normą.

- Niech $\|f\| = 0$ dla $f \in C[0,2]$. Wtedy

$$\sup_{x\in[0,1]} |f(x)| + \int_1^2 |f(x)|\, \mathrm{d}x = 0,$$

  skąd $\sup_{x\in[0,1]}|f(x)| = 0$ oraz $\int_1^2 |f(x)|\, \mathrm{d}x = 0$. Zatem, z pierwszego warunku, $f(x) = 0$ dla $x \in [0,1]$. Gdyby natomiast istniał taki punkt $x_0 \in (1,2]$, że

$f(x_0) \neq 0$, to z ciągłości funkcji f możemy znaleźć przedział $[a,b] \subset (1,2]$, dla którego $\inf_{x\in[a,b]}|f(x)| > 0$. Wtedy

$$\int_1^2 |f(x)|\,\mathrm{d}x \geq \int_a^b |f(x)|\,\mathrm{d}x \geq \int_a^b \inf_{x\in[a,b]}|f(x)|\,\mathrm{d}x = \\ = (b-a)\inf_{x\in[a,b]}|f(x)| > 0.$$

Powyższa sprzeczność dowodzi, że $f(x) = 0$ dla $x \in (1,2]$, czyli ostatecznie $f(x) = 0$ dla $x \in [0,2]$.

- Niech $f \in C[0,2]$, $\alpha \in \mathbb{R}$. Wtedy

$$\begin{aligned}\|\alpha f\| &= \sup_{x\in[0,1]}|(\alpha f)(x)| + \int_1^2 |(\alpha f)(x)|\,\mathrm{d}x = \\ &= \sup_{x\in[0,1]}|\alpha||f(x)| + \int_1^2 |\alpha||f(x)|\,\mathrm{d}x = \\ &= |\alpha|\sup_{x\in[0,1]}|f(x)| + |\alpha|\int_1^2 |f(x)|\,\mathrm{d}x = \alpha\|f\|.\end{aligned}$$

- Niech $f, g \in C[0,2]$. Wykorzystując nierówność trójkąta dla wartości bezwzględnej i supremum oraz monotoniczność całki Riemanna, mamy

$$\begin{aligned}\|f+g\| &= \sup_{x\in[0,1]}|f(x)+g(x)| + \int_1^2 |f(x)+g(x)|\,\mathrm{d}x \leq \\ &\leq \sup_{x\in[0,1]}\big(|f(x)|+|g(x)|\big) + \int_1^2 \big(|f(x)|+|g(x)|\big)\,\mathrm{d}x \leq \\ &\leq \sup_{x\in[0,1]}|f(x)| + \sup_{x\in[0,1]}|g(x)| + \int_1^2 |f(x)|\,\mathrm{d}x + \int_1^2 |g(x)|\,\mathrm{d}x = \\ &= \|f\| + \|g\|.\end{aligned}$$

Żeby dowieść, że przestrzeń $C[0,2]$ z tak wprowadzoną normą nie jest przestrzenią zupełną, niech $f_n \in C[0,2]$, $n \geq 1$, gdzie

$$f_n(x) = \begin{cases} 1, & x \in [0,1], \\ (2-x)^n, & x \in (1,2]. \end{cases}$$

Wtedy, dla $n > m$ oraz dowolnego $\epsilon > 0$ mamy

$$\begin{aligned}\|f_n - f_m\| &= \int_1^2 \big[(2-x)^m - (2-x)^n\big]\,\mathrm{d}x = \\ &= \int_0^1 \big(t^m - t^n\big)\,\mathrm{d}t = \frac{1}{m} - \frac{1}{n} < \frac{1}{m} < \epsilon,\end{aligned}$$

o ile $m > N := \left\lfloor \frac{1}{\epsilon} \right\rfloor$. Pokazuje to, że ciąg $(f_n)_{n\geq 1}$ spełnia warunek Cauchy'ego w $C[0,2]$. Załóżmy, że jest to ciąg zbieżny w zadanej normie do funkcji $f \in C[0,2]$. Ponieważ

$$\|f_n - f\| \geq \sup_{x\in[0,1]} |f_n(x) - f(x)|,$$

to ciąg $(f_n)_{n\geq 1}$ jest zbieżny jednostajnie do f na przedziale $[0,1]$. Zatem dla $x \in [0,1]$, $f(x)$ jest granicą punktową ciągu $(f_n(x))_{n\geq 1}$, więc

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} 1 = 1, \qquad x \in [0,1].$$

Z drugiej strony, na przedziale $(1,2]$ funkcja f musi być równa zero. Gdyby było inaczej, to na mocy ciągłości funkcji f istniałby przedział $[a,b] \subset (1,2]$, o tej własności, że $\inf_{x\in[a,b]}|f(x)| > 0$. Wtedy

$$\begin{aligned}
\lim_{n\to\infty}\|f_n - f\| &\geq \lim_{n\to\infty}\int_1^2 |f_n(x) - f(x)|\,\mathrm{d}x \geq \\
&\geq \lim_{n\to\infty}\int_a^b |f_n(x) - f(x)|\,\mathrm{d}x \geq \\
&\geq \lim_{n\to\infty}\int_a^b |f(x)|\,\mathrm{d}x - \lim_{n\to\infty}\int_a^b |f_n(x)|\,\mathrm{d}x \geq \\
&\geq (b-a)\inf_{x\in[a,b]}|f(x)| - \lim_{n\to\infty}\int_a^b (2-x)^n\,\mathrm{d}x = \\
&= (b-a)\inf_{x\in[a,b]}|f(x)| - \lim_{n\to\infty}\frac{1}{n+1}\left[(2-a)^{n+1} - (2-b)^{n+1}\right] = \\
&= (b-a)\inf_{x\in[a,b]}|f(x)| > 0,
\end{aligned}$$

co daje sprzeczność. Stąd $f(x) = 0$ dla $x \in (1,2]$, więc ostatecznie

$$f(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & x \in (1,2], \end{cases}$$

jednak ta funkcja nie jest funkcją ciągłą, to znaczy $f \notin C[0,2]$. Ciąg $(f_n)_{n\geq 1}$ nie jest zatem zbieżny w przestrzeni $C[0,2]$, co pokazuje, że nie jest ona przestrzenią Banacha.

Zadanie 14.6. Niech $(f_n)_{n\geq 1} \subset \mathbb{X}$ będzie ciągiem spełniającym warunek Cauchy'ego, to znaczy

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n,m>N} \|f_n - f_m\| < \epsilon.$$

Wtedy

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n,m>N} \sup_{x\in[0,2]} |f_n(x) - f_m(x)| + |f_n(5) - f_m(5)| < \epsilon, \tag{14.11}$$

skąd

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n,m>N}\bigwedge_{x\in S}|f_n(x)-f_m(x)|<\epsilon. \tag{14.12}$$

Oznacza to, że dla każdego ustalonego $x\in S$, ciąg liczbowy $(f_n(x))_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$, a zatem jest zbieżny. Niech

$$f(x):=\lim_{n\to\infty}f_n(x),\qquad x\in S.$$

Należy teraz wykazać, że $f\in\mathbb{X}$ oraz $\lim_{n\to\infty}\|f_n-f\|=0$.

Na mocy (14.12) istnieje liczba naturalna N, dla której

$$|f_n(x)-f_N(x)|<1,\qquad n>N,\ x\in S.$$

Z nierówności trójkąta otrzymujemy zatem

$$|f_n(x)|<1+|f_N(x)|,\qquad n>N,\ x\in S.$$

Przechodząc z n to nieskończoności, powyższa nierówność przyjmuje postać

$$|f(x)|\leq 1+|f_N(x)|,\qquad x\in S.$$

Ponieważ f_N jest funkcją ograniczoną na S, to również f jest funkcją ograniczoną na S. Ponadto, dla $x\in[1,2]$ mamy

$$f(x)=\lim_{n\to\infty}f_n(x)=\lim_{n\to\infty}0=0,$$

co dowodzi, że $f\in\mathbb{X}$. Zauważmy ponadto, że warunek (14.11) implikuje

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\bigwedge_{x\in[0,2]}|f_n(x)-\lim_{m\to\infty}f_m(x)|+|f_n(5)-\lim_{m\to\infty}f_m(5)|\leq\epsilon,$$

czyli

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\|f_n-f\|\leq\epsilon.$$

Dowodzi to, że $\mathbb{X}$ jest przestrzenią Banacha.

Zadanie 14.7. Załóżmy, że ciąg $(f_n)_{n\geq 1}\subset\mathbb{X}$ spełnia warunek Cauchy'ego i niech $\epsilon>0$ będzie dowolnie wybrane. Istnieje zatem taka liczba naturalna N, że dla dowolnych liczb naturalnych $m,n\geq N$ zachodzi nierówność

$$\|f_n-f_m\|<\epsilon,$$

czyli

$$\sup_{x\in[0,3]}|f_n(x)-f_m(x)|<\epsilon. \tag{14.13}$$

Krok 1 — zbieżność punktowa ciągu $(f_n)_{n\geq 1}$.

Niech $x \in [0,3]$. Wtedy z warunku (14.13) wynika, że

$$|f_n(x) - f_m(x)| < \epsilon \tag{14.14}$$

o ile $m, n \geq N$. Oznacza to, że ciąg liczbowy $(f_n(x))_{n\geq 1}$ spełnia w $\mathbb{R}$ warunek Cauchy'ego, czyli jest zbieżny. Zdefiniujmy funkcję f wzorem

$$f(x) := \lim_{n\to\infty} f_n(x), \qquad x \in [0,3].$$

Krok 2 — funkcja f jest elementem przestrzeni $\mathbb{X}$.

Musimy wykazać, że

(a) f jest funkcją ograniczoną na przedziale $[0,3]$,

(b) f jest funkcją ciągłą na przedziale $[0,1]$,

(c) spełniona jest równość

$$f(x) = \frac{1}{2}(f(x-1) + f(x+1)), \qquad x \in [1,2].$$

(a) Z nierówność (14.13) otrzymujemy

$$|f_n(x) - f_N(x)| < \epsilon, \qquad x \in [0,3],\ n \geq N,$$

co implikuje

$$|f_n(x)| < \epsilon + |f_N(x)|, \qquad x \in [0,3].$$

Przechodząc z n do nieskończoności, otrzymujemy więc

$$|f(x)| \leq \epsilon + |f_N(x)|, \qquad x \in [0,3].$$

Ponieważ f_N jest funkcją ograniczoną, to również f jest funkcją ograniczoną.

(b) Aby funkcja f była ciągła w ustalonym punkcie $x_0 \in [0,1]$, musimy pokazać, że dla dowolnego $\epsilon > 0$ istnieje taka $\delta > 0$, że

$$|x_0 - x| < \delta \Longrightarrow |f(x_0) - f(x)| < \epsilon.$$

Na mocy nierówności trójkąta

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x) - f(x)|, \quad n \geq 1.$$

Oszacujemy każdą z powyższych wartości bezwzględnych oddzielnie. Niech $\epsilon > 0$. Przechodząc w warunku (14.13) z m do $+\infty$ otrzymujemy, że istnieje taka liczba naturalna N, że dla $n \geq N$ oraz dowolnego $x \in [0,1]$ zachodzi nierówność

$$|f_n(x) - f(x)| < \epsilon. \tag{14.15}$$

Ponieważ funkcja f_N jest ciągła na $[0,1]$, to istnieje $\delta_N > 0$, iż

$$|f_N(x_0) - f_N(x)| < \epsilon \tag{14.16}$$

o ile $|x_0 - x| < \delta_N$. Łącząc warunki (14.15) i (14.16), otrzymujemy

$$\begin{aligned}|f(x_0) - f(x)| &\le |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| \le \\ &\le \epsilon + \epsilon + \epsilon = 3\epsilon,\end{aligned}$$

jeżeli tylko $|x_0 - x| < \delta_N$. Dowolność ϵ dowodzi zatem ciągłości funkcji f w każdym punkcie $x_0 \in [0, 1]$.

(c) Ponieważ $f_n \in \mathbb{X}$, to

$$\begin{aligned}f(x) &= \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{1}{2}(f_n(x-1) + f_n(x+1)) = \\ &= \frac{1}{2}(f(x-1) + f(x+1)).\end{aligned}$$

Krok 3 — zbieżność ciągu $(f_n)_{n\ge 1}$ do f w normie $\|\cdot\|$.

Zauważmy, że dla dowolnego $x \in [0, 3]$ z warunku (14.14) oraz zbieżności punktowej ciągu $(f_n)_{n\ge 1}$ do f wynika, że

$$\lim_{m\to\infty} |f_n(x) - f_m(x)| = |f_n(x) - f(x)| \le \epsilon,$$

dla $n \ge N$, co możemy przepisać w postaci

$$\sup_{x\in[0,3]} |f_n(x) - f(x)| \le \epsilon,$$

czyli

$$\bigwedge_{\epsilon>0} \bigvee_{N\ge 1} \bigwedge_{m,n>N} \|f_n - f\| \le \epsilon.$$

Z dowolności ϵ, dowodzi to zbieżności ciągu $(f_n)_{n\ge 1}$ w normie supremum.

Zadanie 14.8. (a) Z definicji przestrzeni Banacha, musimy wykazać, że dowolny ciąg $(x_n)_{n\ge 1} \subset \mathbb{X}$ spełniający warunek Cauchy'ego

$$\bigwedge_{\epsilon>0} \bigvee_{N\ge 1} \bigwedge_{m,n>N} \|x_n - x_m\|_{\mathbb{X}} < \epsilon, \tag{14.17}$$

jest zbieżny do pewnego elementu $x \in \mathbb{X}$. Załóżmy zatem, że ciąg $(x_n)_{n\ge 1}$ w przestrzeni $\mathbb{X}$ spełnia warunek Cauchy'ego. Oznaczmy $x_n = (\xi_{n,i})_{i\ge 1}$ dla $n \ge 1$ i przepiszmy warunek (14.17) następująco, pamiętając że $\|\cdot\|_{\mathbb{X}} = \|\cdot\|_{l^1}$,

$$\bigwedge_{\epsilon>0} \bigvee_{N\ge 1} \bigwedge_{m,n>N} \sum_{i=1}^{\infty} |\xi_{n,i} - \xi_{m,i}| < \epsilon. \tag{14.18}$$

Ponieważ dla dowolnie wybranego $i \ge 1$ mamy $\sum_{j=1}^{\infty} |\xi_{n,j} - \xi_{m,j}| \ge |\xi_{n,i} - \xi_{m,i}|$, więc, w szczególności, warunek (14.18) implikuje

$$\bigwedge_{\epsilon>0} \bigvee_{N\ge 1} \bigwedge_{m,n>N} \bigwedge_{i\ge 1} |\xi_{n,i} - \xi_{m,i}| < \epsilon.$$

Oznacza to, że ciąg i -tych współrzędnych $(\xi_{n,i})_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$. Ponieważ zbiór $\mathbb{R}$ z naturalną normą jest przestrzenią Banacha, więc dla ustalonego $i \geq 1$ ciąg $(\xi_{n,i})_{n\geq 1}$ jest zbieżny, powiedzmy do $\xi_i \in \mathbb{R}$.

Niech teraz $x = (\xi_i)_{i\geq 1}$. Pokażemy, że $x_n \to x$ w $\mathbb{X}$. W tym celu musimy sprawdzić dwie rzeczy. Po pierwsze, że $x \in \mathbb{X}$, a po drugie, że $\lim_{n\to\infty} \|x_n - x\|_{\mathbb{X}} = 0$. Zauważmy, że warunek (14.18) można zapisać w postaci

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{k\geq 1} \sum_{i=1}^{k} |\xi_{n,i} - \xi_{m,i}| < \epsilon. \tag{14.19}$$

Ponieważ N nie zależy od k, więc

$$\lim_{m\to\infty} \sum_{i=1}^{k} |\xi_{n,i} - \xi_{m,i}| = \sum_{i=1}^{k} |\xi_{n,i} - \xi_i| \leq \epsilon, \tag{14.20}$$

dla dowolnego $k \geq 1$. Z nierówności trójkąta dla wartości bezwzględnej wiemy, że

$$\sum_{i=1}^{k} |\xi_i| - \sum_{i=1}^{k} |\xi_{n,i}| \leq \epsilon,$$

więc

$$\sum_{i=1}^{k} |\xi_i| \leq \sum_{i=1}^{k} |\xi_{n,i}| + \epsilon.$$

Z dowolności k widzimy, że szereg $\sum_{i=1}^{k} |\xi_i|$ jest zbieżny, ponieważ $x_n \in \mathbb{X}$. Zatem $x \in l^1$ i, żeby $x \in \mathbb{X}$, wystarczy pokazać, że $\sum_{i=1}^{\infty} \xi_i = 0$. Dla $\epsilon > 0$ mamy

$$\begin{aligned} \left|\sum_{i=1}^{\infty} \xi_i\right| &= \left|\sum_{i=1}^{\infty} (\xi_i - \xi_{n,i} + \xi_{n,i})\right| \leq \left|\sum_{i=1}^{\infty} (\xi_i - \xi_{n,i})\right| + \left|\sum_{i=1}^{\infty} \xi_{n,i}\right| = \\ &= \left|\sum_{i=1}^{\infty} (\xi_i - \xi_{n,i})\right| \leq \sum_{i=1}^{\infty} |\xi_i - \xi_{n,i}| \leq \epsilon, \end{aligned}$$

o ile $n > N$, na mocy (14.19) (zwróćmy uwagę, że korzystaliśmy tutaj również ze zbieżności bezwzględnej szeregów liczbowych). Z dowolności ϵ otrzymujemy $\sum_{i=1}^{\infty} \xi_i = 0$, więc $x \in \mathbb{X}$.

Pozostaje sprawdzić, że $\lim_{n\to\infty} \|x_n - x\|_{\mathbb{X}} = 0$, co tak naprawdę już zrobiliśmy. Warunek (14.19) i nierówność (14.20) pokazują, że

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \bigwedge_{k\geq 1} \sum_{i=1}^{k} |\xi_{n,i} - \xi_i| \leq \epsilon,$$

co można zapisać jako

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \sum_{i=1}^{\infty} |\xi_{n,i} - \xi_i| \leq \epsilon,$$

czyli

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{n>N} \|x_n - x\| \leq \epsilon.$$

Ostatni warunek jest definicją zbieżności ciągu $(x_n)_{n\geq 1}$ do x w przestrzeni $\mathbb{X}$.

(b) Niech $x_n = (1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^n}, 0\ldots)$, $n \geq 1$. Ciąg ten jest elementem (zmienionej) przestrzeni $\mathbb{X}$ i spełnia warunek Cauchy'ego, gdyż dla $n > m$ mamy

$$\|x_n - x_m\| = \sum_{i=m+1}^{n} \frac{1}{2^i} = \frac{1}{2^{m+1}} \frac{1 - \frac{1}{2^{n-m}}}{1 - \frac{1}{2}} = \frac{1}{2^m}\left(1 - \frac{1}{2^{n-m}}\right) < \frac{1}{2^m}.$$

Z drugiej strony, gdyby $(x_n)_{n\geq 1}$ był zbieżny w $\mathbb{X}$, to byłby również zbieżny po współrzędnych (czego dowiedliśmy w punkcie (a)), a wtedy jego granicą byłby ciąg $(1, \frac{1}{2}, \frac{1}{4}, \ldots)$, który nie należy do $\mathbb{X}$, gdyż wszystkie jego współrzędne są różne od 0.

Uwaga: Punktu (a) można dowieść dużo prościej, wykorzystując następujące twierdzenie: Jeżeli $\mathbb{X}$ jest przestrzenią Banacha, a $\mathbb{Y} \subset \mathbb{X}$ jej podprzestrzenią domkniętą, to $\mathbb{Y}$ jest również przestrzenią Banacha.

Zadanie 14.9. Załóżmy, że ciąg $(f_n)_{n\geq 1}$ spełnia w przestrzeni $\mathbb{X}$ warunek Cauchy'ego, to znaczy

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|f_n - f_m\| < \epsilon. \tag{14.21}$$

Krok 1 — zbieżność punktowa ciągu $(f_n)_{n\geq 1}$.

Z warunku (14.21) wiemy, że

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sup_{x\in[0,1]} |f_n(x) - f_m(x)| < \epsilon.$$

Stąd

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{x\in[0,1]} |f_n(x) - f_m(x)| < \epsilon. \tag{14.22}$$

Zatem dla każdego $x \in [0,1]$ ciąg liczbowy $(f_n(x))_{n\geq 1}$ spełnia (w $\mathbb{R}$) warunek Cauchy'ego, czyli jest zbieżny. Zdefiniujmy funkcję f wzorem

$$f(x) := \lim_{n\to\infty} f_n(x), \qquad x \in [0,1].$$

Krok 2 — funkcja f jest elementem przestrzeni $\mathbb{X}$.

Aby funkcja f była ciągła w ustalonym punkcie $x_0 \in [0,1]$, musimy pokazać, że

$$\bigwedge_{\epsilon>0} \bigvee_{\delta>0} |x_0 - x| < \delta \Longrightarrow |f(x_0) - f(x)| < \epsilon.$$

Na mocy nierówności trójkąta

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x) - f(x)|, \quad n \geq 1.$$

Oszacujemy każdą z powyższych wartości bezwzględnych oddzielnie. Niech $\epsilon > 0$. Przechodząc w warunku (14.22) z m do $+\infty$ otrzymujemy, że istnieje taka liczba naturalna N, że dla $n \geq N$ oraz dowolnego $x \in [0,1]$ zachodzi nierówność

$$|f_n(x) - f(x)| < \epsilon. \tag{14.23}$$

Ponieważ funkcja f_N jest ciągła, to istnieje $\delta_N > 0$, iż

$$|f_N(x_0) - f_N(x)| < \epsilon \tag{14.24}$$

o ile $|x_0 - x| < \delta_N$. Łącząc warunki (14.23) i (14.24), otrzymujemy

$$\begin{aligned}|f(x_0) - f(x)| &\leq |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| \leq \\ &\leq \epsilon + \epsilon + \epsilon = 3\epsilon,\end{aligned}$$

jeżeli tylko $|x_0 - x| < \delta_N$. Dowolność ϵ dowodzi zatem ciągłości funkcji f w każdym punkcie $x_0 \in [0,1]$.

Ponieważ $f_n \in \mathbb{X}$, to

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(1-x) = f(1-x),$$

oraz

$$f(0) = \lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} f_n(1/2) = f(1/2).$$

Oznacza to, że $f \in \mathbb{X}$.

Krok 3 — zbieżność ciągu $(f_n)_{n\geq 1}$ do f w normie $\|\cdot\|$.

Przechodząc w warunku (14.22) do granicy przy $m \to \infty$, otrzymujemy (na mocy zbieżności *punktowej* ciągu $(f_n)_{n\geq 1}$ do f), że

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{x\in[0,1]} \lim_{m\to\infty} |f_n(x) - f_m(x)| = |f_n(x) - f(x)| \leq \epsilon,$$

co możemy przepisać w postaci

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sup_{x\in[0,1]} |f_n(x) - f(x)| \leq \epsilon,$$

czyli

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|f_n - f\| \leq \epsilon.$$

Ponownie, z dowolności ϵ, dowodzi to zbieżności ciągu $(f_n)_{n\geq 1}$ w normie supremum.

Zadanie 14.10. Niech $x = (\xi_n)_{n\geq 1}$ i załóżmy, że $\|x\| = 0$. Wtedy

$$\sum_{n=1}^{\infty} \mathrm{e}^{\sin \frac{n\pi}{7}} |\xi_n| = 0,$$

skąd dla dowolnego $n \geq 1$ mamy $\mathrm{e}^{\sin\frac{n\pi}{7}}|\xi_n| = 0$, czyli $\xi_n = 0$, gdyż $\mathrm{e}^{\sin\frac{n\pi}{7}} \neq 0$. Dla $a \in \mathbb{R}$ widzimy, że

$$\|ax\| = \sum_{n=1}^{\infty} \mathrm{e}^{\sin\frac{n\pi}{7}}|a\xi_n| = |a| \sum_{n=1}^{\infty} \mathrm{e}^{\sin\frac{n\pi}{7}}|\xi_n| = |a| \cdot \|x\|.$$

Ponadto, dla $y = (\zeta_n)_{n\geq 1} \in l^1$,

$$\|x+y\| \leq \sum_{n=1}^{\infty} \mathrm{e}^{\sin\frac{n\pi}{7}}|\xi_n + \zeta_n| \leq \sum_{n=1}^{\infty} \mathrm{e}^{\sin\frac{n\pi}{7}}|\xi_n| + \sum_{n=1}^{\infty} \mathrm{e}^{\sin\frac{n\pi}{7}}|\zeta_n| = \|x\| + \|y\|.$$

Funkcja $\|\cdot\|$ jest zatem normą.

Z definicji przestrzeni Banacha, musimy wykazać, że dowolny ciąg $(x_n)_{n\geq 1}$ spełniający w przestrzeni $\mathbb{X}$ warunek Cauchy'ego

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|x_n - x_m\| < \epsilon, \tag{14.25}$$

jest zbieżny do pewnego elementu $x \in \mathbb{X}$. Załóżmy zatem, że ciąg $(x_n)_{n\geq 1}$ spełnia w przestrzeni $\mathbb{X}$ warunek Cauchy'ego. Oznaczmy ponadto $x_n = (\xi_{n,i})_{i\geq 1}$ dla $n \geq 1$ i przepiszmy warunek (14.25) w postaci

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sum_{k=1}^{\infty} \mathrm{e}^{\sin\frac{k\pi}{7}}|\xi_{n,k} - \xi_{m,k}| < \epsilon. \tag{14.26}$$

Krok 1 — zbieżność „po współrzędnych" ciągu $(x_n)_{n\geq 1}$

Niech $\epsilon' > 0$ oraz $i \geq 1$ będą dowolnie wybrane. Na mocy warunku (14.26) mamy

$$\mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_{n,i} - \xi_{m,i}| \leq \sum_{k=1}^{\infty} \mathrm{e}^{\sin\frac{k\pi}{7}}|\xi_{n,k} - \xi_{m,k}| < \epsilon, \qquad m,n > N.$$

Stąd

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} |\xi_{n,i} - \xi_{m,i}| < \mathrm{e}^{-\sin\frac{i\pi}{7}}\epsilon.$$

Kładąc $\epsilon := \mathrm{e}^{-\sin\frac{i\pi}{7}}\epsilon'$, otrzymujemy, iż

$$\bigvee_{N\geq 1} \bigwedge_{m,n>N} |\xi_{n,i} - \xi_{m,i}| < \epsilon'.$$

Z dowolności ϵ' wynika, że ciąg liczbowy $(\xi_{n,i})_{n\geq 1}$ spełnia warunek Cauchy'ego w $\mathbb{R}$. Ma on zatem granicę, którą oznaczmy $\xi_i := \lim_{n\to\infty} \xi_{n,i}$, $i \geq 1$.

Dowiedliśmy zatem, że ciąg $(x_n)_{n\geq 1}$ jest zbieżny „po współrzędnych" do ciągu liczbowego $x = (\xi_i)_{i\geq 1} \subset \mathbb{R}$.

Krok 2 — ciąg $x = (\xi_i)_{i\geq 1}$ należy do przestrzeni $\mathbb{X}$.

Aby pokazać, że $x \in \mathbb{X}$, trzeba sprawdzić absolutną sumowalność ciągu x, to znaczy, że $\sum_{i=1}^{\infty}|\xi_i| < \infty$.

Zauważmy, że warunek (14.26) implikuje

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{m,n>N}\bigwedge_{k\geq 1}\sum_{i=1}^{k}\mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_{n,i}-\xi_{m,i}| < \epsilon. \tag{14.27}$$

Ponadto, ponieważ $\mathrm{e}^{\sin\frac{i\pi}{7}} \geq \frac{1}{\mathrm{e}}$, $i \geq 1$, to dla $k \geq 1$ mamy

$$\sum_{i=1}^{k}\mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_{n,i}-\xi_{m,i}| \geq \frac{1}{\mathrm{e}}\sum_{i=1}^{k}|\xi_{n,i}-\xi_{m,i}|. \tag{14.28}$$

Ponieważ (suma jest skończona!)

$$\lim_{m\to\infty}\sum_{i=1}^{k}|\xi_{n,i}-\xi_{m,i}| = \sum_{i=1}^{k}|\xi_{n,i}-\xi_i|, \qquad k,n \geq 1,$$

to przechodząc w nierówności (14.28) do granicy przy $m \to \infty$, oraz wykorzystując (14.27), otrzymujemy

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\bigwedge_{k\geq 1}\frac{1}{\mathrm{e}}\sum_{i=1}^{k}|\xi_{n,i}-\xi_i| \leq \epsilon. \tag{14.29}$$

Wybierając $\epsilon = 1$, z („odwrotnej”) nierówności trójkąta dla wartości bezwzględnej wiemy zatem, że

$$\sum_{i=1}^{k}|\xi_i| - \sum_{i=1}^{k}|\xi_{n,i}| \leq \sum_{i=1}^{k}|\xi_{n,i}-\xi_i| \leq \mathrm{e}, \qquad k \geq 1,\ n > N.$$

Ponieważ $x_n \in \mathbb{X}$, to

$$\sum_{i=1}^{k}|\xi_i| \leq \sum_{i=1}^{k}|\xi_{n,i}| + \mathrm{e} \leq \|x_n\|_{l^1} + \mathrm{e}, \qquad k \geq 1,\ n > N.$$

Liczba po prawej stronie nierówności nie zależy od k, więc z dowolności $k \geq 1$ widzimy, że szereg $\sum_{i=1}^{\infty}|\xi_i|$ jest zbieżny. Stąd $x \in \mathbb{X}$, czego należało dowieść.

Krok 3 — zbieżność ciągu $(x_n)_{n\geq 1}$ do x w $\mathbb{X}$.

Pozostaje sprawdzić, że $\lim_{n\to\infty}\|x_n - x\| = 0$, co tak naprawdę już (prawie) zrobiliśmy.

Podobnie jak w kroku drugim, przechodząc w warunku (14.26) do granicy (sumy są skończone!) przy $m \to \infty$, otrzymujemy

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\bigwedge_{k\geq 1}\sum_{i=1}^{k}\mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_{n,i}-\xi_i| \leq \epsilon.$$

Stąd

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\sum_{i=1}^{\infty}\mathrm{e}^{\sin\frac{i\pi}{7}}|\xi_{n,i}-\xi_i|\leq\epsilon.$$

Ostatecznie widzimy, iż

$$\bigwedge_{\epsilon>0}\bigvee_{N\geq 1}\bigwedge_{n>N}\|x_n-x\|\leq\epsilon,$$

co z definicji oznacza, że

$$\lim_{n\to\infty}x_n=x$$

w przestrzeni $\mathbb{X}$.

Zadanie 14.11. Załóżmy najpierw, że każdy ciąg monotoniczny i ograniczony jest zbieżny. Niech $(\xi_n)_{n\geq 1}$ będzie ciągiem podstawowym. Zauważmy, że musi on być ograniczony. Istotnie, na mocy warunku Cauchy'ego, istnieje taka liczba naturalna N, że dla $m,n\geq N$ mamy $|\xi_n-\xi_m|<1$, skąd

$$|\xi_n|<|\xi_N|+1,$$

o ile tylko $n\geq N$. Otrzymujemy zatem, że

$$|\xi_n|\leq\max\big\{|\xi_N|+1,|\xi_1|,|\xi_2|,\ldots,|\xi_N|\big\},\qquad n\geq 1.$$

Liczba po prawej stronie jest oczywiście skończona, więc ciąg $(\xi_n)_{n\geq 1}$ jest ograniczony.

Udowodnimy teraz, że z każdego ciągu podstawowego można wybrać podciąg monotoniczny. Powiemy, że wyraz ξ_k ciągu $(\xi_n)_{n\geq 1}$ jest *dominujący w prawo*, gdy

$$\xi_k\geq\xi_n,\qquad n>k.$$

Przypuśćmy, że w ciągu $(\xi_n)_{n\geq 1}$ jest nieskończenie wiele wyrazów dominujących w prawo. Oznaczmy je przez $\xi_{k_1},\xi_{k_2},\ldots$, gdzie $k_1<k_2<\ldots$. Ponieważ każdy wyraz ξ_{k_i} jest dominujący w prawo, to podciąg $(\xi_{k_i})_{i\geq 1}$ jest malejący.

Z drugiej strony, załóżmy, że w ciągu $(\xi_n)_{n\geq 1}$ jest skończenie wiele wyrazów dominujących w prawo. Oznaczmy przez N najmniejszą liczbę naturalną, dla której żaden wyraz ξ_n dla $n\geq N$ nie jest dominujący w prawo. Ponieważ ξ_N nie jest dominujący w prawo, to można znaleźć wyraz $\xi_{k_1}>\xi_N$, gdzie $k_1>N$. Podobnie, ponieważ ξ_{k_1} nie jest dominujący w prawo, to istnieje taki indeks k_2, że $\xi_{k_2}>\xi_{k_1}$ i $k_2>k_1$. W ten sposób konstruujemy indukcyjnie rosnący podciąg $(\xi_{k_i})_{i\geq 1}$.

Wykazaliśmy do tej pory, że ciąg podstawowy $(\xi_n)_{n\geq 1}$ jest ograniczony i można z niego wybrać podciąg monotoniczny $(\xi_{k_i})_{i\geq 1}$. Na mocy założenia, podciąg ten ma granicę, powiedzmy ξ. Niech $\epsilon>0$. Istnieje taka liczba naturalna K, że

$$|\xi_{k_i}-\xi|<\epsilon,\qquad i>K.$$

Ponadto, ponieważ $(\xi_n)_{n\geq 1}$ jest ciągiem podstawowym, to można znaleźć $N\geq 1$, że $|\xi_n-\xi_{k_i}|<\epsilon$ dla $n,i>N$. Ostatecznie

$$|\xi_n-\xi|\leq|\xi_n-\xi_{k_i}|+|\xi_{k_i}-\xi|<2\epsilon$$

dla $n, i > \max(N, K)$, co dowodzi zbieżności ciągu $(\xi_n)_{n\geq 1}$.

Załóżmy teraz, że każdy ciąg podstawowy jest zbieżny i przypuśćmy, że istnieje ograniczony i monotoniczny ciąg $(\xi_n)_{n\geq 1}$, który nie ma granicy. Bez straty ogólności rozważań, możemy założyć, że jest on rosnący. Gdyby ciąg $(\xi_n)_{n\geq 1}$ nie spełniał warunku Cauchy'ego, to istniałaby taka liczba $\epsilon > 0$, że dla dowolnej liczby naturalnej k można znaleźć wskaźniki n_k i m_k, dla których $n_k > m_k$ oraz

$$|\xi_{n_k} - \xi_{m_k}| = \xi_{n_k} - \xi_{m_k} \geq \epsilon.$$

Ponadto, liczbę m_k można wybrać tak, aby $m_k > n_{k-1}$ dla $k \geq 2$. Stąd

$$\begin{aligned}\xi_{n_k} &= \xi_{n_k} - \xi_{m_k} + \xi_{m_k} \geq \epsilon + \xi_{n_{k-1}} = \\ &= \epsilon + \xi_{n_{k-1}} - \xi_{m_{k-1}} + \xi_{m_{k-1}} \geq 2\epsilon + \xi_{n_{k-2}} \geq \\ &\geq \cdots \geq (k-1)\epsilon + \xi_{n_1},\end{aligned}$$

gdzie $k \geq 1$. Na mocy powyższej nierówności otrzymujemy $\lim_{k\to\infty} \xi_{n_k} = +\infty$, co przeczy ograniczoności ciągu $(\xi_n)_{n\geq 1}$. Zatem musi on spełniać warunek Cauchy'ego. Na mocy założenia ma on więc granicę, co prowadzi do sprzeczności. Pokazuje to, że każdy ciąg ograniczony i monotoniczny jest zbieżny, co kończy dowód.

Zadanie 14.12. Udowodnimy najpierw, że przestrzeń $\mathbb{X}$ funkcji ciągłych na $[0, 1]$ z normą Bieleckiego jest przestrzenią Banacha. Musimy zatem wykazać, że dowolny ciąg $(f_n)_{n\geq 1}$ spełniający w przestrzeni $\mathbb{X}$ warunek Cauchy'ego

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|f_n - f_m\|_\lambda < \epsilon, \tag{14.30}$$

jest zbieżny do pewnego elementu $f \in \mathbb{X}$. Załóżmy zatem, że ciąg $(f_n)_{n\geq 1}$ w przestrzeni $\mathbb{X}$ spełnia warunek Cauchy'ego.

Krok 1 — zbieżność punktowa ciągu $(f_n)_{n\geq 1}$.

Z warunku (14.30) wiemy, że

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x}|f_n(x) - f_m(x)| < \epsilon.$$

Stąd

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{x\in[0,1]} \mathrm{e}^{-\lambda x}|f_n(x) - f_m(x)| < \epsilon, \tag{14.31}$$

czyli

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{x\in[0,1]} |f_n(x) - f_m(x)| < \mathrm{e}^{\lambda x}\epsilon \leq \max(1, \mathrm{e}^{\lambda}) \cdot \epsilon. \tag{14.32}$$

Z dowolności ϵ, dla każdego $x \in [0, 1]$ ciąg liczbowy $(f_n(x))_{n\geq 1}$ spełnia (w $\mathbb{R}$) warunek Cauchy'ego, a zatem jest zbieżny. Oznaczmy więc

$$f(x) := \lim_{n\to\infty} f_n(x), \qquad x \in [0, 1].$$

Krok 2 — ciągłość funkcji f.
Aby funkcja f była ciągła w ustalonym punkcie $x_0 \in [0,1]$, musimy pokazać, że

$$\bigwedge_{\epsilon>0} \bigvee_{\delta>0} \bigwedge_{x\in[0,1]} |x_0 - x| < \delta \Longrightarrow |f(x_0) - f(x)| < \epsilon.$$

Na mocy nierówności trójkąta

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x) - f(x)|, \quad n \geq 1.$$

Oszacujemy każdą z powyższych wartości bezwzględnych oddzielnie. Niech $\epsilon > 0$. Przechodząc w warunku (14.32) z m do $+\infty$ otrzymujemy, że istnieje taka liczba naturalna N, że dla $n \geq N$ oraz dowolnego $x \in [0,1]$ zachodzi nierówność

$$|f_n(x) - f(x)| < \epsilon. \tag{14.33}$$

Ponadto, ponieważ funkcja f_N jest ciągła, to istnieje $\delta_N > 0$, iż

$$|f_N(x_0) - f_N(x)| < \epsilon \tag{14.34}$$

o ile $|x_0 - x| < \delta_N$. Łącząc (14.33) i (14.34), otrzymujemy

$$\begin{aligned} |f(x_0) - f(x)| &\leq |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| \leq \\ &\leq \epsilon + \epsilon + \epsilon = 3\epsilon, \end{aligned}$$

jeżeli tylko $|x_0 - x| < \delta_N$. Dowolność ϵ dowodzi zatem ciągłości funkcji f w każdym punkcie $x_0 \in [0,1]$.

Krok 3 — zbieżność ciągu $(f_n)_{n\geq 1}$ do f w normie $\|\cdot\|_\lambda$.
Przechodząc w warunku (14.31) do granicy przy $m \to \infty$, otrzymujemy (na mocy zbieżności **punktowej** ciągu $(f_n)_{n\geq 1}$ do f), że

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \bigwedge_{x\in[0,1]} \lim_{m\to\infty} \mathrm{e}^{-\lambda x}|f_n(x) - f_m(x)| = \mathrm{e}^{-\lambda x}|f_n(x) - f(x)| \leq \epsilon,$$

co możemy przepisać w postaci

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x}|f_n(x) - f(x)| \leq \epsilon,$$

czyli

$$\bigwedge_{\epsilon>0} \bigvee_{N\geq 1} \bigwedge_{m,n>N} \|f_n - f\|_\lambda \leq \epsilon.$$

Ponownie, z dowolności ϵ, dowodzi to zbieżności ciągu $(f_n)_{n\geq 1}$ w normie $\|\cdot\|_\lambda$.

Aby wykazać, że istnieje dokładnie jedna funkcja ciągła na przedziale $[0,1]$ dla której

$$\bigwedge_{x\in[0,1]} f(x) + \int_0^x [\sin(xy) + \cos^4(x+y)]f(y)\,\mathrm{d}y = \frac{1}{1+x^2},$$

zauważmy na początku, że powyższą równość można przepisać w postaci

$$(I - A)f = g,$$

gdzie $g \in C[0,1]$, $g(x) = \frac{1}{1+x^2}$ oraz A jest operatorem liniowym zdefiniowanym następująco:

$$A\colon C[0,1] \to C[0,1], \quad (Af)(x) = -\int_0^x [\sin(xy) + \cos^4(x+y)]f(y)\,\mathrm{d}y, \; x \in [0,1].$$

Ponadto, dla $\lambda > 0$, mamy

$$\begin{aligned}
\|Af\|_\lambda &= \sup_{x\in[0,1]} \left| \mathrm{e}^{-\lambda x} \int_0^x [\sin(xy) + \cos^4(x+y)]f(y)\,\mathrm{d}y \right| \leq \\
&\leq \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x |\sin(xy) + \cos^4(x+y)| \cdot |f(y)|\,\mathrm{d}y = \\
&= \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x \mathrm{e}^{\lambda y}|\sin(xy) + \cos^4(x+y)| \cdot \mathrm{e}^{-\lambda y}|f(y)|\,\mathrm{d}y \leq \\
&\leq \|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x \mathrm{e}^{\lambda y}|\sin(xy) + \cos^4(x+y)|\,\mathrm{d}y \leq \\
&\leq \|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x 2\mathrm{e}^{\lambda y}\,\mathrm{d}y = 2\frac{1}{\lambda}\|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x}(\mathrm{e}^{\lambda x} - 1) = \\
&= \frac{2}{\lambda}\|f\|_\lambda \sup_{x\in[0,1]} (1 - \mathrm{e}^{-\lambda x}) = \frac{2}{\lambda}(1 - \mathrm{e}^{-\lambda})\|f\|_\lambda \leq \frac{2}{\lambda}\|f\|_\lambda.
\end{aligned}$$

Jeżeli zatem $\lambda > 2$, to $\|A\| < 1$. Ponieważ udowodniliśmy, że przestrzeń funkcji ciągłych z normą Bieleckiego jest przestrzenią Banacha, to na mocy twierdzenia z podrozdziału 12.4, operator $I - A$ ma ograniczony operator odwrotny $(I - A)^{-1}\colon C[0,1] \to C[0,1]$. Zatem $f := (I - A)^{-1}g$ jest (jedyną) funkcją spełniającą żądaną równość.

Zadanie 14.13. Przepiszmy równość z treści zadania w postaci

$$\xi_n + \frac{n^2\cos n}{2(n^2+1)}\xi_{n-2} - \frac{\sin n^2}{5}\xi_{n-1} - \frac{1}{8}\frac{1}{1+n^2}\xi_{n+1} = \frac{1}{n^2+1}, \qquad n \in \mathbb{Z}. \quad (14.35)$$

Jeżeli zatem na przestrzeni ciągów obustronnie nieskończonych i absolutnie sumowalnych określimy operator A wzorem

$$A(\xi_n)_{n\in\mathbb{Z}} = \left(-\frac{n^2\cos n}{2(n^2+1)}\xi_{n-2} + \frac{\sin n^2}{5}\xi_{n-1} + \frac{1}{8}\frac{1}{1+n^2}\xi_{n+1}\right)_{n\in\mathbb{Z}},$$

to równość (14.35) przyjmie postać operatorową

$$(I - A)\big((\xi_n)_{n\in\mathbb{Z}}\big) = (\tfrac{1}{n^2+1})_{n\in\mathbb{Z}}.$$

Badana przestrzeń jest przestrzenią Banacha, jeżeli wprowadzimy w niej naturalną normę $\|(\xi_n)_{n\in\mathbb{Z}}\| = \sum_{n\in\mathbb{Z}}|\xi_n|$. Ponadto, mamy

$$\begin{aligned}\|A(\xi_n)_{n\in\mathbb{Z}}\| &\leq \sum_{n\in\mathbb{Z}}\left|-\frac{n^2\cos n}{2(n^2+1)}\xi_{n-2}+\frac{\sin n^2}{5}\xi_{n-1}+\frac{1}{8}\frac{1}{1+n^2}\xi_{n+1}\right| \leq \\ &\leq \sum_{n\in\mathbb{Z}}\frac{|\cos n|}{2}|\xi_{n-2}|+\sum_{n\in\mathbb{Z}}\frac{|\sin n^2|}{5}|\xi_{n-1}|+\sum_{n\in\mathbb{Z}}\frac{1}{8(1+n^2)}|\xi_{n+1}| \leq \\ &\leq \frac{1}{2}\sum_{n\in\mathbb{Z}}|\xi_{n-2}|+\frac{1}{5}\sum_{n\in\mathbb{Z}}|\xi_{n-1}|+\frac{1}{8}\sum_{n\in\mathbb{Z}}|\xi_{n+1}| = \\ &= \left(\frac{1}{2}+\frac{1}{5}+\frac{1}{8}\right)\sum_{n\in\mathbb{Z}}|\xi_n| = \frac{33}{40}\|(\xi_n)_{n\in\mathbb{Z}}\|.\end{aligned}$$

Stąd $\|A\| \leq \frac{33}{40} < 1$, a zatem na mocy twierdzenia z podrozdziału 12.4 istnieje ograniczony operator odwrotny do $I - A$. Ponieważ ciąg $(\frac{1}{n^2+1})_{n\in\mathbb{Z}}$ jest elementem rozważanej przestrzeni, to

$$(\xi_n)_{n\in\mathbb{Z}} := (I-A)^{-1}(\tfrac{1}{n^2+1})_{n\in\mathbb{Z}}$$

jest jedynym rozwiązaniem wyjściowego równania.

Zadanie 14.14. Niech $C[0,1]$ będzie przestrzenią Banacha funkcji ciągłych na odcinku $[0,1]$ ze standardową normą supremum. Przepiszmy równość z treści zadania w postaci

$$f(x)-\frac{1}{2}f\Big(\frac{1}{3}x\Big)-\frac{1}{3}f\Big(\frac{1}{2}x\Big) = \mathrm{e}^{\pi x}-\sin(x^2), \qquad x\in[0,1], \tag{14.36}$$

oraz zdefiniujmy operator liniowy $A\colon C[0,1]\to C[0,1]$ wzorem

$$(Af)(x) = \frac{1}{2}f\Big(\frac{1}{3}x\Big)+\frac{1}{3}f\Big(\frac{1}{2}x\Big), \qquad f\in C[0,1],\ x\in[0,1].$$

Ponadto, niech $g\in C[0,1]$ będzie funkcją ciągłą, gdzie $g(x) = \mathrm{e}^{\pi x}-\sin(x^2)$, $x\in[0,1]$. Równość (14.36) przyjmuje wtedy postać

$$(I-A)f = g. \tag{14.37}$$

Zatem, aby wykazać, że istnieje dokładnie jedna funkcja $f\in C[0,1]$ spełniająca (14.37), wystarczy sprawdzić, że $\|A\| < 1$. W tym celu, niech $f\in C[0,1]$. Mamy

$$\begin{aligned}\|Af\|_{C[0,1]} &= \sup_{x\in[0,1]}\left|\frac{1}{2}f\Big(\frac{1}{3}x\Big)+\frac{1}{3}f\Big(\frac{1}{2}x\Big)\right| \leq \\ &\leq \sup_{x\in[0,1]}\left|\frac{1}{2}f\Big(\frac{1}{3}x\Big)\right|+\sup_{x\in[0,1]}\left|\frac{1}{3}f\Big(\frac{1}{2}x\Big)\right| = \\ &= \frac{1}{2}\sup_{x\in[0,1]}\left|f\Big(\frac{1}{3}x\Big)\right|+\frac{1}{3}\sup_{x\in[0,1]}\left|f\Big(\frac{1}{2}x\Big)\right|.\end{aligned}$$

Ponieważ $0 \leq \frac{1}{2}x, \frac{1}{3}x \leq 1$ dla dowolnego $x \in [0,1]$, to

$$\sup_{x\in[0,1]}\left|f\left(\frac{1}{2}x\right)\right| \leq \sup_{t\in[0,1]}|f(t)| = \|f\|_{C[0,1]},$$

oraz

$$\sup_{x\in[0,1]}\left|f\left(\frac{1}{3}x\right)\right| \leq \sup_{t\in[0,1]}|f(t)| = \|f\|_{C[0,1]}.$$

Stąd

$$\|Af\|_{C[0,1]} \leq \frac{1}{2}\|f\|_{C[0,1]} + \frac{1}{3}\|f\|_{C[0,1]} = \frac{5}{6}\|f\|_{C[0,1]},$$

a zatem ostatecznie $\|A\| \leq \frac{5}{6}$. Na mocy twierdzenia z podrozdziału 12.4, operator $I - A$ jest bijekcją oraz istnieje do niego operator odwrotny $(I-A)^{-1}\colon C[0,1] \to C[0,1]$. Równość (14.37) jest zatem równoważna równości

$$f = (I-A)^{-1}g,$$

i tak zdefiniowana funkcja $f \in C[0,1]$ jest (jedynym) szukanym rozwiązaniem.

Zadanie 14.15. Zauważmy najpierw, że

$$\lim_{x\to 1^-}\frac{x^2-1}{x^2-3x+2} = \lim_{x\to 1^-}\frac{x+1}{x-2} = -2.$$

Oznacza to, że funkcja g określona wzorem

$$g(x) = \begin{cases} \frac{x^2-1}{x^2-3x+2}, & x \in [0,1), \\ -2, & x = 1, \end{cases}$$

jest funkcją ciągłą na przedziale $[0,1]$. Zatem *jeżeli* istnieje funkcja f ciągła na przedziale $[0,1)$, spełniająca równość z treści zadania, to można ją w jednoznaczny sposób rozszerzyć do funkcji ciągłej na przedziale $[0,1]$, kładąc

$$f(1) = -2 + \int_0^1 [\sin(1+y) + \cos^4(y)]f(y)\,\mathrm{d}y.$$

Wystarczy zatem wykazać, że istnieje dokładnie jedna funkcja $f \in C[0,1]$, spełniająca równość

$$f(x) - \int_0^x [\sin(x+y) + \cos^4(xy)]f(y)\,\mathrm{d}y = g(x), \qquad x \in [0,1]. \tag{14.38}$$

Metoda I. Równość (14.38) można zapisać w postaci

$$(I-A)f = g,$$

gdzie $A\colon C[0,1] \to C[0,1]$ jest operatorem liniowym zdefiniowanym następująco:

$$(Af)(x) = \int_0^x [\sin(x+y) + \cos^4(xy)]f(y)\,\mathrm{d}y, \qquad x \in [0,1].$$

Ponadto, dla $\lambda > 0$, mamy

$$\begin{aligned}
\|Af\|_\lambda &= \sup_{x\in[0,1]} \left| \mathrm{e}^{-\lambda x} \int_0^x [\sin(x+y) + \cos^4(xy)] f(y)\,\mathrm{d}y \right| \leq \\
&\leq \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x |\sin(x+y) + \cos^4(xy)| \cdot |f(y)|\,\mathrm{d}y = \\
&= \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x \mathrm{e}^{\lambda y} |\sin(x+y) + \cos^4(xy)| \cdot \mathrm{e}^{-\lambda y} |f(y)|\,\mathrm{d}y \leq \\
&\leq \|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x \mathrm{e}^{\lambda y} |\sin(x+y) + \cos^4(xy)|\,\mathrm{d}y \leq \\
&\leq \|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x} \int_0^x 2\mathrm{e}^{\lambda y}\,\mathrm{d}y = 2\frac{1}{\lambda}\|f\|_\lambda \sup_{x\in[0,1]} \mathrm{e}^{-\lambda x}(\mathrm{e}^{\lambda x} - 1) = \\
&= \frac{2}{\lambda}\|f\|_\lambda \sup_{x\in[0,1]} (1 - \mathrm{e}^{-\lambda x}) = \frac{2}{\lambda}(1 - \mathrm{e}^{-\lambda})\|f\|_\lambda \leq \frac{2}{\lambda}\|f\|_\lambda.
\end{aligned}$$

Jeżeli zatem $\lambda > 2$, to $\|A\| < 1$. Ponieważ przestrzeń funkcji ciągłych z normą Bieleckiego jest przestrzenią Banacha, to operator $I - A$ ma ograniczony operator odwrotny $(I - A)^{-1}\colon C[0,1] \to C[0,1]$. Stąd $f := (I - A)^{-1}g$ jest jedyną funkcją ciągłą spełniającą równość

$$f(x) - \int_0^x [\sin(x+y) + \cos^4(xy)] f(y)\,\mathrm{d}y = g(x), \qquad x \in [0,1].$$

Metoda II. Niech $S\colon C[0,1] \to C[0,1]$ będzie operatorem (nieliniowym!) zdefiniowanym wzorem

$$(Sf)(x) = \int_0^x [\sin(x+y) + \cos^4(xy)] f(y)\,\mathrm{d}y + g(x), \qquad x \in [0,1].$$

Zauważmy, że wykorzystując te same oszacowania jak w metodzie I, mamy

$$\begin{aligned}
\|Sf_1 - Sf_2\|_\lambda &= \sup_{x\in[0,1]} \left| \mathrm{e}^{-\lambda x} \int_0^x [\sin(x+y) + \cos^4(xy)] \big[f_1(y) - f_2(y)\big]\,\mathrm{d}y \right| \leq \\
&\leq \frac{2}{\lambda}\|f_1 - f_2\|_\lambda.
\end{aligned}$$

Zatem dla $\lambda > 2$ na mocy zasady Banacha istnieje dokładnie jedna funkcja ciągła $f \in C[0,1]$, będąca punktem stałym odwzorowania S, to znaczy $Sf = f$, czego należało dowieść.

Zadanie 14.16. Metoda I. Niech operator $A\colon l^1 \to l^1$ będzie dany wzorem

$$A(\xi_n)_{n\geq 0} = (a_n)_{n\geq 0} - \left(\sum_{k=0}^n \xi_k b_{n-k} \right)_{n\geq 0}, \qquad (\xi_n)_{n\geq 0} \in l^1.$$

Zwróćmy uwagę, że definicja operatora A jest poprawna. Rzeczywiście, $(a_n)_{n\geq 0} \in l^1$ $(a_n \sim \frac{1}{n^4})$ oraz $(b_n)_{n\geq 0} \in l^1$, gdyż

$$\sum_{n=0}^{\infty} |b_n| = \mathrm{e}^{-4} \sum_{n=0}^{\infty} \frac{2^n}{n!} = \mathrm{e}^{-4}\mathrm{e}^2 = \mathrm{e}^{-2},$$

co powoduje, że $\left(\sum_{k=0}^{n} \xi_k b_{n-k}\right)_{n\geq 0} \in l^1$, ze względu na

$$\begin{aligned}\sum_{n=0}^{\infty} \left|\sum_{k=0}^{n} \xi_k b_{n-k}\right| \leq \sum_{n=0}^{\infty}\sum_{k=0}^{n} |\xi_k| \cdot |b_{n-k}| = \sum_{k=0}^{\infty} |\xi_k| \sum_{n=k}^{\infty} |b_{n-k}| \\ = \|(\xi_n)_{n\geq 0}\|_{l^1} \cdot \|(b_n)_{n\geq 0}\|_{l^1}.\end{aligned}$$

Ponadto, dla $(\xi_n)_{n\geq 0}, (\zeta_n)_{n\geq 0} \in l^1$, mamy

$$\begin{aligned}\|A(\xi_n)_{n\geq 0} - A(\zeta_n)_{n\geq 0}\|_{l^1} &= \left\| \left(\sum_{k=0}^{n} \zeta_k b_{n-k}\right)_{n\geq 0} - \left(\sum_{k=0}^{n} \xi_k b_{n-k}\right)_{n\geq 0} \right\|_{l^1} = \\ &= \sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} \zeta_k b_{n-k} - \sum_{k=0}^{n} \xi_k b_{n-k} \right| \leq \\ &\leq \sum_{n=0}^{\infty}\sum_{k=0}^{n} |\zeta_k - \xi_k||b_{n-k}| = \\ &= \mathrm{e}^{-4} \sum_{n=0}^{\infty}\sum_{k=0}^{n} |\zeta_k - \xi_k| \frac{2^{n-k}}{(n-k)!} = \\ &= \mathrm{e}^{-4} \sum_{k=0}^{\infty} |\zeta_k - \xi_k| \sum_{n=k}^{\infty} \frac{2^{n-k}}{(n-k)!} = \\ &= \mathrm{e}^{-4} \|(\zeta_k)_{k\geq 0} - (\xi_k)_{k\geq 0}\|_{l^1} \sum_{n=0}^{\infty} \frac{2^n}{n!} = \\ &= \mathrm{e}^{-4}\mathrm{e}^2 \|(\zeta_k)_{k\geq 0} - (\xi_k)_{k\geq 0}\|_{l^1} = \\ &= \mathrm{e}^{-2} \|(\zeta_k)_{k\geq 0} - (\xi_k)_{k\geq 0}\|_{l^1}.\end{aligned}$$

Ponieważ $\mathrm{e}^{-2} < 1$, zatem na mocy zasady Banacha istnieje dokładnie jedno rozwiązanie $(\xi_n)_{n\geq 0} \in l^1$ równania

$$(\xi_n)_{n\geq 0} = A(\xi_n)_{n\geq 0},$$

co można zapisać jako

$$\xi_n = a_n - \sum_{k=0}^{n} \xi_k b_{n-k}, \qquad n \geq 0.$$

Metoda II. Zdefiniujmy operator (liniowy) $A\colon l^1 \to l^1$ wzorem

$$A(\xi_n)_{n\geq 0} = \left(-\sum_{k=0}^{n} \xi_k b_{n-k} \right)_{n\geq 0}, \qquad (\xi_n)_{n\geq 0} \in l^1.$$

Wtedy równanie z treści zadania przyjmuje postać

$$(I-A)(\xi_n)_{n\geq 0}=(a_n)_{n\geq 0}.$$

Ponieważ $(a_n)_{n\geq 0}\in l^1$, wystarczy zatem dowieść, że A jest operatorem ograniczonym i $\|A\|<1$. Podobnie jak wcześniej, łatwo sprawdzić, że dla $(\xi_n)_{n\geq 0}\in l^1$ mamy

$$\|A(\xi_n)_{n\geq 0}\|_{l^1}\leq \sum_{n=0}^{\infty}\sum_{k=0}^{n}|\xi_k|\frac{2^{n-k}}{(n-k)!}=\mathrm{e}^{-2}\|(\xi_n)_{n\geq 0}\|_{l^1}.$$

Stąd ostatecznie $\|A\|\leq \mathrm{e}^{-2}<1$.

Zadanie 14.17. Przepiszmy równanie z treści zadania w postaci

$$\xi_n=\frac{1}{4n}\xi_{n+1}-\frac{1}{2n^2}\xi_{n+2}-\frac{\sin n}{4n^2},\qquad n\geq 1,$$

oraz zdefiniujmy operator $A\colon l^1\to l^1$ wzorem

$$A(\xi_n)_{n\geq 1}=\left(\frac{1}{4n}\xi_{n+1}-\frac{1}{2n^2}\xi_{n+2}-\frac{\sin n}{4n^2}\right)_{n\geq 1},\qquad (\xi_n)_{n\geq 1}\in l^1.$$

Definicja operatora A jest poprawna, gdyż $\frac{1}{4n},\frac{1}{2n^2}<1$, $(\xi_n)_{n\geq 1}\in l^1$ oraz $\frac{|\sin n|}{n^2}\leq\frac{1}{n^2}$.

Niech $(\xi_n)_{n\geq 1},(\zeta_n)_{n\geq 1}\in l^1$. Wtedy

$$\begin{aligned}\|A(\xi_n)_{n\geq 1}-A(\zeta_n)_{n\geq 1}\|_{l^1}&=\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n}|\xi_{n+1}-\zeta_{n+1}|+\frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{n^2}|\xi_{n+2}-\zeta_{n+2}|\leq\\&\leq\frac{1}{4}\sum_{n=1}^{\infty}|\xi_{n+1}-\zeta_{n+1}|+\frac{1}{2}\sum_{n=1}^{\infty}|\xi_{n+2}-\zeta_{n+2}|\leq\\&\leq\frac{1}{4}\sum_{n=1}^{\infty}|\xi_n-\zeta_n|+\frac{1}{2}\sum_{n=1}^{\infty}|\xi_n-\zeta_n|\leq\\&\leq\left(\frac{1}{4}+\frac{1}{2}\right)\|(\xi_n)_{n\geq 1}-(\zeta_n)_{n\geq 1}\|_{l^1}=\\&=\frac{3}{4}\|(\xi_n)_{n\geq 1}-(\zeta_n)_{n\geq 1}\|_{l^1}.\end{aligned}$$

Zatem na mocy zasady Banacha istnieje dokładnie jeden ciąg $(\xi_n)_{n\geq 1}\in l^1$, dla którego

$$(\xi_n)_{n\geq 1}=A(\xi_n)_{n\geq 1},$$

czego należało dowieść.

Zadanie 14.18. Dla dowolnego $x=(\xi_i)_{i\geq 0}\in l^2$, element $y=(\eta_i)_{i\geq 0}\in A$ jest rzutem x na podprzestrzeń A, jeżeli dla dowolnego $z\in A$ mamy

$$(x-y,z)=0,$$

gdzie $(\cdot,\cdot)$ jest standardowym iloczynem skalarnym w l^2. Dla $n \geq 0$ niech

$$z^n = (\zeta_i^n)_{i\geq 0} = (0,\ldots,0,1,2,0,\ldots),$$

przy czym „1” występuje na miejscu o indeksie $2n$. Wtedy oczywiście $z^n \in A$ oraz

$$(x-y,z^n) = \sum_{i=0}^{\infty}(\xi_i - \eta_i)z_i^n = (\xi_{2n} - \eta_{2n}) + 2(\xi_{2n+1} - \eta_{2n+1}), \qquad n \geq 0.$$

Ponieważ chcemy, aby $y \in A$, więc $\eta_{2i} = 2\eta_{2i+1}$, $i \geq 0$. Zatem $(x-y,z^n) = 0$ wtedy i tylko wtedy, gdy

$$\eta_{2n+1} = \tfrac{1}{3}(\xi_{2n} + 2\xi_{2n+1}).$$

Z dowolności n oraz faktu, że $y \in A$, otrzymujemy ostatecznie

$$\eta_i = \begin{cases} \frac{2}{3}(\xi_{2k} + 2\xi_{2k+1}), & i = 2k, \\ \frac{1}{3}(\xi_{2k} + 2\xi_{2k+1}), & i = 2k+1, \end{cases}$$

dla $i,k \geq 0$.

Zadanie 14.19. (A) Łatwo zauważyć, że iloczyn skalarny powinien być zdefiniowany worem

$$\big((\xi_i)_{i\geq 1},(\eta_i)_{i\geq 1}\big) = \sum_{i=1}^{\infty} i\ \xi_i\eta_i.$$

Sprawdzenie, że jest to rzeczywiście iloczyn skalarny, który generuje normę $\|\cdot\|$, jest prostym ćwiczeniem.

Uwaga. Można również wykorzystać jawny wzór na iloczyn skalarny w przestrzeni Hilberta

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

(B) Zbiór z podpunktu (a) nie jest oczywiście podprzestrzenią liniową, gdyż nie zawiera on wektora zerowego.

Aby uzasadnić, że zbiór z podpunktu (b) jest podprzestrzenią liniową, zauważmy że

$$(\zeta_i)_{i\geq 1} := a(\xi_i)_{i\geq 1} + b(\eta_i)_{i\geq 1} = (a\xi_i + b\eta_i)_{i\geq 1}, \qquad a,b \in \mathbb{R},$$

a zatem jeżeli $\xi_4 + \xi_9 = \eta_4 + \eta_9 = 0$, to

$$\zeta_4 + \zeta_9 = a\xi_4 + b\eta_4 + a\xi_9 + b\eta_9 = a(\xi_4 + \xi_9) + b(\eta_4 + \eta_9) = 0.$$

Niech teraz $(x_n)_{n\geq 1} = \big((\xi_{n,i})_{i\geq 1}\big)_{n\geq 1}$ będzie ciągiem elementów z rozważanej podprzestrzeni, to znaczy

$$\xi_{n,4} + \xi_{n,9} = 0, \qquad n \geq 1.$$

Jeżeli dla $x = (\xi_i)_{i\geq 1}$ mamy $\|x_n - x\| \to 0$, $n \to \infty$, to

$$2|\xi_4 - \xi_{n,4}|^2 + 3|\xi_9 - \xi_{n,9}|^2 \leq \sum_{i=1}^{\infty} i\ (\xi_i - \xi_{n,i})^2 \to 0,$$

dla $n \to \infty$, skąd $\xi_{n,4} \to \xi_4$ i $\xi_{n,9} \to \xi_9$. Ostatecznie

$$0 = \lim_{n\to\infty} (\xi_{n,4} + \xi_{n,9}) = \xi_4 + \xi_9,$$

co dowodzi domkniętości zbioru z podpunktu (b).

(C) Niech

$$z = (\zeta_i)_{i\geq 1} = (0, 0, 0, 1, 0, 0, 0, 0, -1, 0, \ldots).$$

Oczywiście $\zeta_4 + \zeta_9 = 0$. Jeżeli $x = (\xi_i)_{i\geq 1}$ ma być rzutem wektora $y = (\eta_i)_{i\geq 1}$, to również

$$0 = (x - y, z) = 4(\xi_4 - \eta_4) + 9(\xi_9 - \eta_9).$$

Ponieważ chcemy, aby $\xi_4 + \xi_9 = 0$, to $\xi_4 = -\frac{1}{5}(4\eta_4 + 9\eta_9)$ i $\xi_9 = \frac{1}{5}(4\eta_4 + 9\eta_9)$.

Aby wyznaczyć pozostałe współrzędne, niech $i \geq 1$, $i \neq 4$, $i \neq 9$, oraz $z = (\zeta_j)_{j\geq 1} = (\delta_{ji})_{j\geq 1}$. Wtedy

$$0 = (x - y, z) = i(\xi_i - \eta_i),$$

skąd $\xi_i = \eta_i$. Ostatecznie otrzymujemy

$$x = (\xi_i)_{i\geq 1} = \begin{cases} \eta_i, & i \neq 4,\ i \neq 9, \\ -\frac{1}{5}(4\eta_4 + 9\eta_9), & i = 4, \\ \frac{1}{5}(4\eta_4 + 9\eta_9), & i = 9. \end{cases}$$

Zadanie 14.20. (a) Ze wzoru na iloczyn skalarny w rzeczywistej przestrzeni Hilberta otrzymujemy

$$\begin{aligned}(x, y) &= \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right) = \\ &= \frac{1}{4}\left(\sum_{i\geq 1} a_i(\xi_i + \eta_i)^2 - \sum_{i\geq 1} a_i(\xi_i - \eta_i)^2\right) = \\ &= \frac{1}{4}\sum_{i\geq 1} a_i\left[(\xi_i + \eta_i)^2 - (\xi_i - \eta_i)^2\right] = \\ &= \sum_{i\geq 1} a_i\xi_i\eta_i,\end{aligned}$$

gdzie $x = (\xi_i)_{i\geq 0} \in \mathbb{H}$ i $y = (\eta_i)_{i\geq 0} \in \mathbb{H}$.

(b) Niech $a, b \in \mathbb{R}$, $x, y \in \mathbb{H}_0$, gdzie $x = (\xi_i)_{i\geq 1}$ i $y = (\eta_i)_{i\geq 1}$. Wtedy

$$z = (\zeta_i)_{i\geq 0} := ax + by = (a\xi_i + b\eta_i)_{i\geq 0}.$$

Stąd

$$\zeta_8 = a\xi_8 + b\eta_8 = a\xi_5 + b\eta_5$$

oraz

$$\zeta_{121} + \zeta_{500} = a\xi_{121} + b\eta_{121} + a\xi_{500} + b\eta_{500} = 0,$$

ponieważ $x, y \in \mathbb{H}_0$.

Niech teraz $(x_n)_{n\geq 1} = \big((\xi_{n,i})_{i\geq 1}\big)_{n\geq 1}$ będzie ciągiem elementów z rozważanej podprzestrzeni, to znaczy

$$\xi_{n,8} = \xi_{n,5}, \qquad \xi_{n,121} + \xi_{n,500} = 0, \quad n \geq 1.$$

Jeżeli dla $x = (\xi_i)_{i\geq 1}$ mamy $\|x_n - x\| \to 0$, $n \to \infty$, to z nierówności

$$(\xi_{n,i} - \xi_i)^2 \leq \frac{1}{a_i}\|x_n - x\|$$

oraz faktu, że $a_n > c > 0$, wynika zbieżność „po współrzędnych", to znaczy

$$\lim_{n\to\infty} \xi_{n,i} = \xi_i.$$

Stąd

$$\xi_8 = \lim_{n\to\infty} \xi_{n,8} = \lim_{n\to\infty} \xi_{n,5} = \xi_5$$

oraz

$$\xi_{121} + \xi_{500} = \lim_{n\to\infty} (\xi_{n,121} + \xi_{n,500}) = \lim_{n\to\infty} 0 = 0,$$

co dowodzi domkniętości zbioru $\mathbb{H}_0$.

(c) Niech $x = (\xi_i)_{i\geq 1} \in \mathbb{H}$, a $y = (\eta_i)_{i\geq 1} \in \mathbb{H}_0$ będzie jego rzutem ortogonalnym na $\mathbb{H}_0$. Z definicji, $y \in \mathbb{H}_0$ jest wektorem, dla którego najmniejsza jest norma

$$\|x - y\| = \sqrt{\sum_{i\geq 1} a_i(\xi_i - \eta_i)^2}.$$

Ponieważ na współrzędne o indeksach różnych od 5, 8, 121 i 500 nie jest nałożony żaden warunek, możemy przyjąć $\eta_i := \xi_i$, $i \neq 5, 8, 121, 500$. Rozważając współrzędne o numerach 5 i 8, dla ustalonych ξ_5, ξ_8 chcemy minimalizować wyrażenie

$$a_5(\xi_5 - \eta_5)^2 + a_8(\xi_8 - \eta_8)^2,$$

co, w połączeniu z $y \in \mathbb{H}_0$, prowadzi do

$$a_5(\xi_5 - \eta_5)^2 + a_8(\xi_8 - \eta_5)^2 = \eta_5^2(a_5 + a_8) - 2\eta_5(a_5\xi_5 + a_8\xi_8) + a_5\xi_5^2 + a_8\xi_8^2.$$

Funkcja kwadratowa $\eta_5 \mapsto \eta_5^2(a_5 + a_8) - 2\eta_5(a_5\xi_5 + a_8\xi_8) + a_5\xi_5^2 + a_8\xi_8^2$ osiąga minimum dla

$$\eta_5 = \frac{a_5\xi_5 + a_8\xi_8}{a_5 + a_8}.$$

Podobnie, rozważając współrzędne o numerach 121 i 500, szukamy liczby η_{121} minimalizującej wyrażenie

$$a_{121}(\xi_{121}-\eta_{121})^2 + a_{500}(\xi_{500}+\eta_{500})^2 = \eta_{121}^2(a_{121}+a_{500}) + \\ -2\eta_{121}(a_{121}\xi_{121}-a_{500}\xi_{500}) + a_{121}\xi_{121}^2 + a_{500}\xi_{500}^2.$$

Otrzymujemy zatem

$$\eta_{121} = \frac{a_{121}\xi_{121}+a_{500}\xi_{500}}{a_{121}+a_{500}}.$$

Ostatecznie

$$y=(\eta_i)_{i\geq 0} = \begin{cases} \xi_i, & i \neq 5,8,121,500, \\ \frac{a_5\xi_5+a_8\xi_8}{a_5+a_8}, & i = 5,8, \\ \frac{a_{121}\xi_{121}+a_{500}\xi_{500}}{a_{121}+a_{500}}, & i = 121, \\ -\frac{a_{121}\xi_{121}+a_{500}\xi_{500}}{a_{121}+a_{500}}, & i = 500. \end{cases}$$

Zadanie 14.21. Mamy

$$\|A(x,y,z)\|_{\mathbb{R}^2} = \|(x-z,x+y)\|_{\mathbb{R}^2} = \max(|x-z|,|x+y|) \leq \\ \leq \max(|x|+|z\ |,|x|+|y|) = |x| + \max(|z\ |,|y|) \leq \\ \leq |x|+|z\ |+|y| = \|(x,y,z)\|_{\mathbb{R}^3}.$$

Widzimy zatem, że operator A jest ograniczony i $\|A\ \| \leq 1$. Aby uzasadnić, że $\|A\ \| = 1$, niech $(x,y,z)=(1,0,0)\in\mathbb{R}^3$. Oczywiście $\|(1,0,0)\|_{\mathbb{R}^3}=1$, a ponadto

$$\|A\ \| = \sup_{\substack{(x,y,z)\in\mathbb{R}^3 \\ \|(x,y,z)\|_{\mathbb{R}^3}=1}} \|A(x,y,z)\|_{\mathbb{R}^2} \geq \|A(1,0,0)\|_{\mathbb{R}^2} = \max(1,1) = 1.$$

Stąd ostatecznie $\|A\ \| = 1$.

Zadanie 14.22. Operator A nie jest ograniczony. Niech $M>0$ oraz

$$x=(0,\ldots,0,1,0,\ldots),$$

przy czym 1 występuje na miejscu o numerze (parzystym) $2\lceil M\rceil$, gdzie $\lceil M\rceil$ jest najmniejszą liczbą całkowitą większą od M. Oczywiście $x\in c_{00}$ oraz

$$\|Ax\| = \sup(0,\ldots,0,2\lceil M\rceil\cdot 1,0,\ldots) = 2\lceil M\rceil > M = M\|x\|,$$

co dowodzi nieograniczoności operatora A .

Operator B jest ograniczony, ponieważ

$$\|B(\xi_n)_{n\geq 1}\| = \sup_{n\geq 1}(|\xi_2|,|\xi_1|,|\xi_4|,|\xi_3|,|\xi_6|,|\xi_5|,\dots) = \\ = \sup_{n\geq 1}(|\xi_1|,|\xi_2|,|\xi_3|,|\xi_4|,|\xi_5|,|\xi_6|,\dots) = \|(\xi_n)_{n\geq 1}\|,$$

co dowodzi również, że $\|B\| = 1$.

Operator C nie jest ograniczony. Niech $M > 0$ oraz

$$x = (0, 1, 0, 1, 0, 1, \ldots, 0, 1, 0, 0, \ldots),$$

przy czym ostatnia „1” występuje na miejscu o numerze (parzystym) $2\lceil M\rceil$. Wtedy $x \in c_{00}$ oraz

$$\|Cx\| = \sup(2\lceil M\rceil, 0, 0, 0, \ldots) = 2\lceil M\rceil > M = M\|x\|.$$

Zadanie 14.23. Operator A jest liniowy. Istotnie, niech $x, y \in c_{00}$, gdzie $x = (\xi_i)_{i\geq 0}$, $y = (\eta_i)_{i\geq 0}$ oraz $a, b \in \mathbb{R}$. Wtedy

$$A(ax + by) = \big(i(a\xi_i + b\eta_i)\big)_{i\geq 0} = a(i\xi_i)_{i\geq 0} + b(i\eta_i)_{i\geq 0} = aAx + bAy.$$

Zauważmy jednak, że operator A nie jest ograniczony. Niech

$$x_n = (1, 1, \ldots, 1, 0, \ldots),$$

przy czym ostatnia jedynka występuje na miejscu o numerze n. Wtedy $x_n \in c_{00}$ oraz

$$\|Ax_n\|_{c_{00}} = \|(1, 2, \ldots, n, 0, \ldots)\|_{c_{00}} = n.$$

Nie istnieje zatem taka stała M, że $\|Ax\|_{c_{00}} \leq M\|x\|_{c_{00}}$ dla $x \in c_{00}$.

Również operator B jest liniowy. Dowód tego faktu przebiega analogicznie jak dla operatora A . Wykażemy teraz, że operator B jest ograniczony. Niech $x = (\xi_i)_{i\geq 0} \in c_{00}$. Wtedy

$$\|Bx\|_{c_{00}} = \sup_{i\geq 0}|\xi_i \sin\tfrac{i\pi}{12}| \leq \sup_{i\geq 0}|\xi_i| = \|x\|_{c_{00}},$$

skąd $\|B\| \leq 1$. Ponadto dla $x = (0, \ldots, 0, 1, 0, \ldots)$, gdzie jedynka występuje na miejscu o indeksie 6, mamy

$$\|Bx\|_{c_{00}} = \|x\|,$$

czyli $\|B\| = 1$.

Operator C nie jest natomiast liniowy, gdyż $C0 = (1, 0, 0, \ldots) \neq 0$.

Zadanie 14.24. Niech $x = (\xi_i)_{n\geq 1} \subset l^1$. Ponieważ $0 < a_i < 1$, to

$$\|Ax\|_{\mathbb{R}} = |Ax| = \left|\sum_{i=1}^{\infty} a_i\xi_i\right| \leq \sum_{i=1}^{\infty}|a_i\xi_i| = \sum_{i=1}^{\infty} a_i|\xi_i| \leq \sum_{i=1}^{\infty} 1\cdot|\xi_i| = \|x\|_{l^1}.$$

Oznacza to, że operator A jest ograniczony oraz $\|A\| \leq 1$.

Wykażemy teraz, że $\|A\| = 1$. Niech

$$x_n = (0, 0, \ldots, 0, 1, 0, \ldots) \in l^1, \qquad i \geq 1,$$

przy czym „1” występuję na miejscu o numerze n. Przypuśćmy, że $\|A\| = M < 1$. Wtedy

$$\|Ax_n\|_{\mathbb{R}} = |Ax_n| \leq M\|x_n\|_{l^1} = M, \qquad i \geq 1.$$

Z drugiej strony

$$|Ax_n| = a_n \xrightarrow{n\to\infty} 1 > M,$$

co daje sprzeczność.

Zadanie 14.25. Zauważmy, że dla $t \in [0,1]$ mamy $t^2 \in [0,1]$, skąd

$$|f(t^2)| \leq \sup_{u\in[0,1]} |f(u)| = \|f\|_{C[0,1]}.$$

Z nierówności trójkąta dla całki Riemanna otrzymujemy zatem

$$|(Af)(x)| = \left|\int_0^x \frac{t}{1+t} f(t^2)\,\mathrm{d}t\right| \leq \int_0^x \frac{t}{1+t}|f(t^2)|\,\mathrm{d}t \leq \|f\|_{C[0,1]} \cdot \int_0^x \frac{t}{1+t}\,\mathrm{d}t =$$
$$= \|f\|_{C[0,1]} \cdot \int_0^x \Big(1 - \frac{1}{1+t}\Big)\,\mathrm{d}t = \big(x - \ln(1+x)\big)\|f\|_{C[0,1]}.$$

Stąd

$$\|Af\|_{C[0,1]} = \sup_{x\in[0,1]} |(Af)(x)| \leq \sup_{x\in[0,1]} \Big|\big(x - \ln(1+x)\big)\|f\|_{C[0,1]}\Big| =$$
$$= \|f\|_{C[0,1]} \sup_{x\in[0,1]} |x - \ln(1+x)| = (1 - \ln 2) \cdot \|f\|_{C[0,1]},$$

ponieważ

$$\sup_{x\in[0,1]} |x - \ln(1+x)| = 1 - \ln 2.$$

Operator A jest zatem ograniczony i $\|A\| \leq 1 - \ln 2$. Ponadto dla funkcji $f \equiv 1$ mamy $\|Af\|_{C[0,1]} = 1 - \ln 2$, co dowodzi, że $\|A\| = 1 - \ln 2$.

Zadanie 14.26. Ustalmy $f \in C[-\infty, \infty]$. Wtedy

$$\begin{aligned}\left|\int_{-\infty}^{\infty} \frac{1}{1+y^2} f(x+y)\,\mathrm{d}y\right| &\leq \int_{-\infty}^{\infty} \frac{1}{1+y^2}|f(x+y)|\,\mathrm{d}y \leq \\ &\leq \int_{-\infty}^{\infty} \frac{1}{1+y^2}\|f\|_{C[-\infty,\infty]}\,\mathrm{d}y = \\ &= \|f\|_{C[-\infty,\infty]} \int_{-\infty}^{\infty} \frac{1}{1+y^2}\,\mathrm{d}y = \\ &= \pi\|f\|_{C[-\infty,\infty]}.\end{aligned}$$

Powyższa nierówność pokazuje, że operator A jest ograniczony i $\|A\| \leq \pi$. Z drugiej strony, dla $f(x) = 1$, $x \in \mathbb{R}$ mamy $f \in C[-\infty, \infty]$ oraz $Af = \pi$, co dowodzi, że $\|A\| = \pi$.

Zadanie 14.27. Aby pokazać, że operator A nie jest ograniczony, wystarczy sprawdzić, że

$$\bigwedge_{M>0} \bigvee_{f\in C^2[0,1]} \|Af\|_{C[0,1]} > M\|f\|_{C^2[0,1]}. \tag{14.39}$$

Niech $M > 0$ będzie dowolnie ustalone. Zdefiniujmy funkcję $f \in C^2[0,1]$ wzorem

$$f(x) = x^{M+3}, \qquad x \in [0,1].$$

Wtedy

$$\|Af\|_{C[0,1]} = \|f''\|_{C[0,1]} = \sup_{x\in[0,1]} |(M+3)(M+2)x^{M+1}| = (M+3)(M+2).$$

Z drugiej strony $\|f\|_{C^2[0,1]} = 1 + (M+3) = M+4$. Ostatecznie, mamy

$$\|Af\|_{C[0,1]} = M^2 + 5M + 6 > M^2 + 4M = M(M+4) = M\|f\|_{C^2[0,1]},$$

co dowodzi (14.39).

Zadanie 14.28. Zauważmy, że

$$A^2 f = A\left(\int_0^1 f(y)\,\mathrm{d}y \cdot f_*\right) = \int_0^1 f(y)\,\mathrm{d}y \cdot A\ f_* =$$
$$= \int_0^1 f(y)\,\mathrm{d}y \cdot \int_0^1 f_*(y)\,\mathrm{d}y \cdot f_* = \frac{1}{3}\int_0^1 f(y)\,\mathrm{d}y \cdot f_* = \frac{1}{3}Af.$$

Zatem, na mocy zasady indukcji matematycznej, mamy

$$A^n f = \frac{1}{3^{n-1}} Af, \qquad n \geq 1,$$

skąd

$$\mathrm{e}^{tA} f = \sum_{n=0}^{\infty} \frac{(tA)^n f}{n!} = A^0 f + \sum_{n=1}^{\infty} \frac{t^n A^n f}{n!} = f + \sum_{n=1}^{\infty} \frac{t^n Af}{3^{n-1}} \frac{1}{n!} =$$
$$= f + 3Af \cdot \sum_{n=1}^{\infty} \left(\frac{t}{3}\right)^n \frac{1}{n!} = f + 3(\mathrm{e}^{t/3} - 1)Af.$$

Ostatecznie otrzymujemy zatem

$$\mathrm{e}^{tA} = I\ + 3(\mathrm{e}^{t/3} - 1)A.$$

Zadanie 14.29. Niech $f \in C[0,\infty]$. Mamy

$$\|Af\| = \|f(0)f_*\| = |f(0)|\|f_*\| = 2|f(0)| \leq 2\|f\|.$$

Zatem operator A jest ograniczony i $\|A\ \| \leq 2$.

Zauważmy, że dla dowolnego $f \in C[0,\infty]$ zachodzi $A^2 f = f(0)Af_* = 2f(0)f_*$. Analogicznie $A^3 f = 4f(0)f_*$. Wykorzystując zasadę indukcji matematycznej, łatwo pokazać, że

$$A^n f = 2^{n-1} f(0) f_*, \qquad n \geq 1.$$

Stąd

$$\mathrm{e}^{tA} f = \sum_{n=0}^{\infty} \frac{t^n A^n f}{n!} = f + \sum_{n=1}^{\infty} \frac{t^n}{n!} A^n f =$$
$$= f + \sum_{n=1}^{\infty} \frac{t^n}{n!} 2^{n-1} f(0) f_* = f + \frac{1}{2} f(0) \sum_{n=1}^{\infty} \frac{(2t)^n}{n!} f_* =$$
$$= f + \frac{1}{2}(\mathrm{e}^{2t} - 1) f_*$$

dla $t \in \mathbb{R}$.

Zadanie 14.30. Dla dowolnej funkcji $f \in L^1(0,1)$ mamy

$$\begin{aligned}
\|Af\| &= \int_0^1 \Big| f(x) + f(1-x) + f\big(\tfrac{1}{2}x\big) \Big| \,\mathrm{d}x \leq \\
&\leq \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(1-x)| \,\mathrm{d}x + \int_0^1 |f\big(\tfrac{1}{2}x\big)| \,\mathrm{d}x = \\
&= \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(t)| \,\mathrm{d}t + 2\int_0^{1/2} |f(u)| \,\mathrm{d}u \ \leq \\
&\leq \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(t)| \,\mathrm{d}t + 2\int_0^{1} |f(u)| \,\mathrm{d}u \ = \\
&= 4\|f\|,
\end{aligned}$$

zatem $\|A\| \leq 4$. Z drugiej strony, jeżeli funkcja $g \in L^1(0,1)$ dana jest wzorem

$$g(x) = \begin{cases} 2, & x \in [0, 1/2], \\ 0, & x \in (1/2, 1], \end{cases}$$

to oczywiście $\|g\| = 1$ oraz

$$\|A\| = \sup_{\substack{f \in L^1(0,1) \\ \|f\|=1}} \|Af\| \geq \|Ag\| = 4.$$

Zadanie 14.31. Dla dowolnej funkcji $f \in L^1(0,1)$ mamy

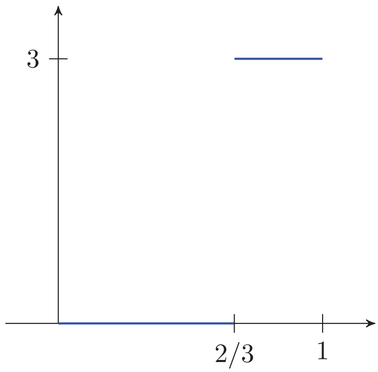
$$\begin{aligned}
\|Af\| &= \int_0^1 \Big| f(x) + f(1-x) + f\big(\tfrac{1}{3}x + \tfrac{2}{3}\big) \Big| \,\mathrm{d}x \leq \\
&\leq \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(1-x)| \,\mathrm{d}x + \int_0^1 \Big|f\big(\tfrac{1}{3}x + \tfrac{2}{3}\big)\Big| \,\mathrm{d}x = \\
&= \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(t)| \,\mathrm{d}t + 3\int_{2/3}^{1} |f(u)| \,\mathrm{d}u \ \leq \\
&\leq \int_0^1 |f(x)| \,\mathrm{d}x + \int_0^1 |f(t)| \,\mathrm{d}t + 3\int_0^{1} |f(u)| \,\mathrm{d}u \ = \\
&= 5\|f\|,
\end{aligned}$$

zatem operator A jest ograniczony i $\|A\| \leq 5$. Z drugiej strony, jeżeli funkcja $g \in L^1(0,1)$ dana jest wzorem

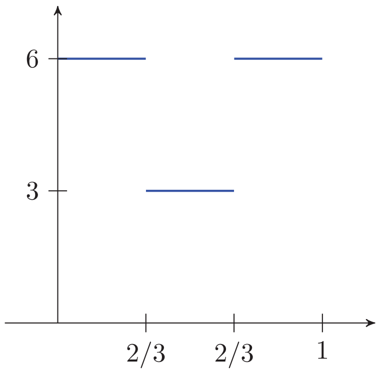
$$g(x) = \begin{cases} 0, & x \in [0, 2/3], \\ 3, & x \in (2/3, 1], \end{cases}$$

to oczywiście $\|g\| = 1$ oraz

$$\|A\| = \sup_{\substack{f \in L^1(0,1) \\ \|f\|=1}} \|Af\| \geq \|Ag\| = 5.$$



Rys. 14.1: Wykres funkcji g.



Rys. 14.2: Wykres funkcji Ag.

Zadanie 14.32. Zauważmy, że dla $f \in C_0(0,1]$ i dowolnego $x \in [0,1]$ mamy

$$A_n f(x) - f(x) = \frac{1}{2}\Big[\tilde{f}(x+n^{-1}) - f(x)\Big] + \frac{1}{2}\Big[\tilde{f}(x-n^{-1}) - f(x)\Big] = \\ = \frac{1}{2}\Big[\tilde{f}(x+n^{-1}) - \tilde{f}(x)\Big] + \frac{1}{2}\Big[\tilde{f}(x-n^{-1}) - \tilde{f}(x)\Big].$$

Z definicji rozszerzenia wynika w szczególności, że $\tilde{f}$ jest funkcją ciągłą na $[-1,2]$, a zatem również jednostajnie ciągłą na tym przedziale. Oznacza to, że

$$\bigwedge_{\epsilon>0} \bigvee_{\delta>0} \bigwedge_{x,y\in[-1,2]} |x-y| < \delta \Rightarrow |\tilde{f}(x) - \tilde{f}(y)| < \epsilon.$$

Ustalmy zatem $\epsilon > 0$ i niech n będzie tak duże, że $n^{-1} < \delta$. Wtedy $|x \pm n^{-1} - x| < \delta$, skąd $\left|\tilde{f}(x+n^{-1}) - \tilde{f}(x)\right| < \epsilon$ i $\left|\tilde{f}(x-n^{-1}) - \tilde{f}(x)\right| < \epsilon$. Ostatecznie

$$|A_n f(x) - f(x)| \le \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon,$$

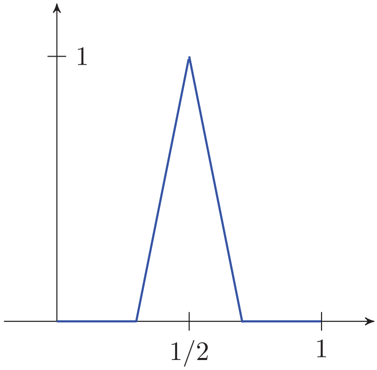
co dowodzi, że $\lim_{n\to\infty} A_n f = f$ dla dowolnego $f \in C_0(0,1]$.

Aby wykazać, że ciąg operatorów A_n nie jest zbieżny do I w normie operatorowej, zdefiniujmy ciąg funkcji $f_n \in C_0(0,1]$, $n \ge 2$, wzorem

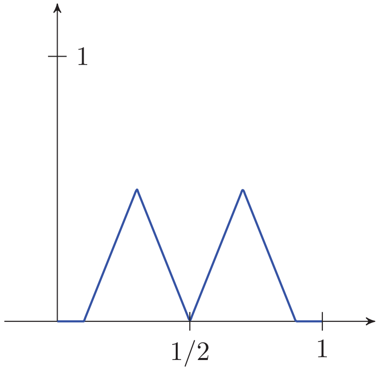
$$f_n(x) = \begin{cases} 0, & x \in [0, 1/2 - n^{-1}], \\ n(x - 1/2 + n^{-1}), & x \in (1/2 - n^{-1}, 1/2], \\ n(-x + 1/2 + n^{-1}), & x \in (1/2, 1/2 + n^{-1}], \\ 0, & x \in (1/2 + n^{-1}, 1]. \end{cases}$$

Wtedy $\|f_n\| = 1$ i

$$\sup_{f\in C_0(0,1], \|f\|=1} \|A_n f - f\| \ge |A_n f_n(1/2) - f_n(1/2)| = 1,$$



Rys. 14.3: Wykres funkcji f_5.



Rys. 14.4: Wykres funkcji $A_5 f_5$.

co pokazuje, że $\|A_n - I\| \geq 1$ dla $n \geq 2$, a więc tym bardziej nie jest prawdą, że $\lim_{n\to\infty}\|A_n - I\| = 0$.

Zadanie 14.33. Niech $\epsilon > 0$ i $f \in C[0,1]$. Ponieważ f jest funkcją ciągłą na przedziale domkniętym, to jest ona również jednostajnie ciągła, to znaczy istnieje liczba $\delta > 0$, dla której

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon, \qquad x, y \in [0,1].$$

Ponieważ ciąg $(a_n)_{n\geq 1}$ jest zbieżny do 1, to istnieje taka liczba naturalna N, że

$$|x - [1 - a_n(1-x)]| < \delta, \qquad n > N.$$

Zatem

$$\|A_n f - f\|_{C[0,1]} = \sup_{x\in[0,1]} |f(1 - a_n(1-x)) - f(x)| \leq \epsilon$$

dla $n > N$. Dowodzi to zbieżności operatorów A_n do I w mocnej topologii.

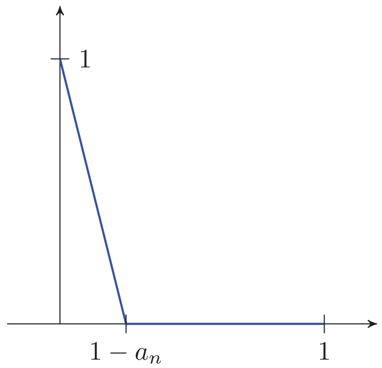
Wykażemy teraz, że operatory A_n nie są zbieżne do I w normie operatorowej. Dla ustalonego $n \geq 1$ niech funkcja $f_n \in C[0,1]$ będzie dana wzorem

$$f_n(x) = \begin{cases} \frac{1}{1-a_n}((1-a_n) - x), & x \in [0, 1-a_n], \\ 0, & x \in (1-a_n, 1]. \end{cases}$$

Wtedy $f_n(a_n(1-x)) = 0$ dla $x \in [0,1]$, skąd $A_n f_n = 0$ oraz

$$\|A_n - I\| = \sup_{\substack{f\in C[0,1] \\ \|f\|_{C[0,1]}=1}} \|A_n f - f\|_{C[0,1]} \geq \|A_n f_n - f_n\|_{C[0,1]} = 1.$$

Pokazuje to, że operatory A_n nie mogą być zbieżne w normie operatorowej do operatora I.



Rys. 14.5: Wykres funkcji f_n.

Zadanie 14.34. (a) Niech $f \in C[0,1]$. Wtedy

$$|T_n f(x)| \leq \sup_{x\in[0,1]} |f(x)| = \|f\|.$$

Oznacza to, że operator T_n jest ograniczony i $\|T_n\| \leq 1$. Z drugiej strony, niech $F \in C[0,1]$, $F \equiv 1$. Wtedy

$$\|T_n\| \geq \|T_n F\| = 1,$$

skąd $\|T_n\| = 1$.

(b) Niech $f \in C[0,1]$. Mamy

$$T_n f(x) - f(x) = \begin{cases} f\big(\sqrt[n]{3}x\big) - f(x), & x \in [0, 1/\sqrt[n]{3}], \\ f(1) - f(x), & x \in (1/\sqrt[n]{3}, 1]. \end{cases}$$

Ponieważ f jest funkcją ciągłą na $[0,1]$, a zatem również jednostajnie ciągłą, to

$$\bigwedge_{\epsilon>0} \bigvee_{\delta>0} \bigwedge_{x,y\in[0,1]} |x-y| \leq \delta \Rightarrow |f(x)-f(y)| < \epsilon.$$

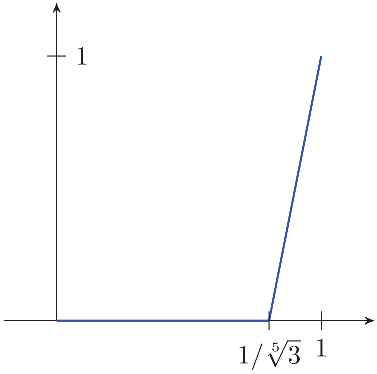
Ze względu na to, iż $\lim_{n\to\infty} \sqrt[n]{3} = 1$, możemy dobrać takie $N \geq 1$, że $0 < \sqrt[n]{3} - 1 < \epsilon$ dla $n > N$. Wtedy dla dowolnego $x \in [0, 1/\sqrt[n]{3}]$ otrzymujemy $0 \leq \sqrt[n]{3}x - x \leq \epsilon$ oraz dla $x \in [1/\sqrt[n]{3}, 1]$, $0 \leq 1 - x \leq \epsilon$, o ile tylko $n > N$. Ostatecznie

$$|T_n f(x) - f(x)| \leq \epsilon, \qquad n > N,\ x \in [0,1].$$

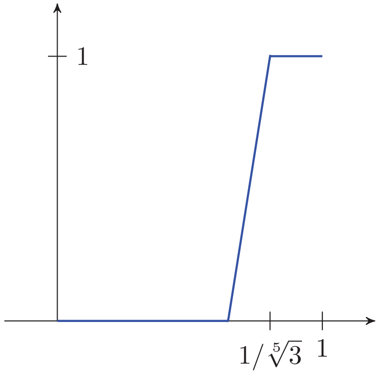
Dowodzi to, że ciąg operatorów T_n jest zbieżny w mocnej topologii do operatora identycznościowego I .

Niech teraz $f_n \in C[0,1]$, $n \geq 2$, gdzie

$$f_n(x) = \begin{cases} 0, & x \in [0, 1/\sqrt[n]{3}], \\ \frac{\sqrt[n]{3}}{\sqrt[n]{3}-1}(x - 1/\sqrt[n]{3}), & x \in (1/\sqrt[n]{3}, 1]. \end{cases}$$



Rys. 14.6: Wykres funkcji f_5.



Rys. 14.7: Wykres funkcji $T_5 f_5$.

Wtedy $\|f_n\| = 1$, $n \geq 2$, oraz

$$\|T_n f_n - f_n\| \geq T_n f_n(1/\sqrt[n]{3}) - f_n(1/\sqrt[n]{3}) = 1 - 0 = 1,$$

co pokazuje, że $\|T_n - I\| \geq 1$.

Zadanie 14.35. (a) Operator A można zapisać jako złożenie operatorów B, C, D, to znaczy

$$A = BCD,$$

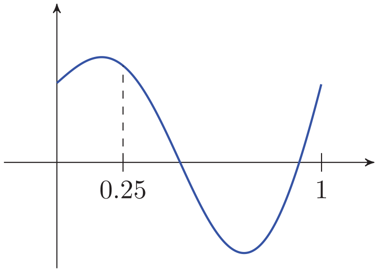
gdzie dla dowolnego $f \in C[0,1]$ mamy

$$Bf = -f, \qquad Cf = 2f,$$

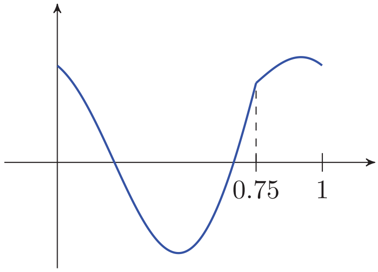
oraz

$$Df(x) = f(x \oplus 0.25), \qquad x \in [0,1].$$

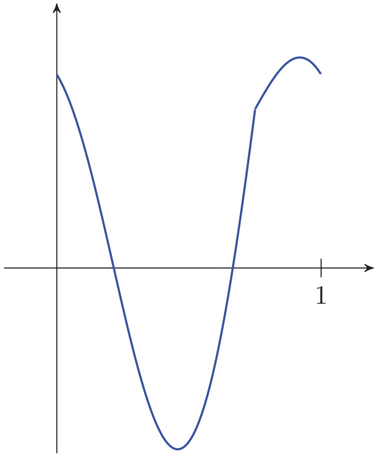
Przy tych oznaczeniach, operator D przesuwa wykres funkcji w lewo o 0.25, a fragment wykresu, który leży teraz w półpłaszczyźnie ujemnej dokleja z prawej strony w punkcie 0.75. Następnie, operator C rozszerza wykres w skali 2 względem osi Ox,



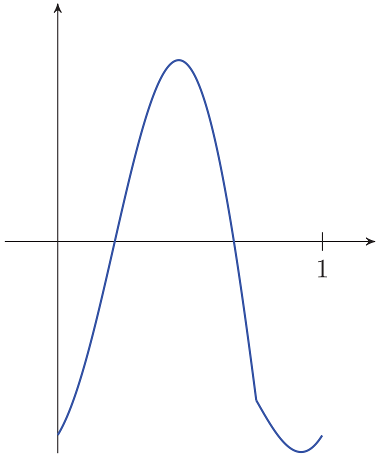
Rys. 14.8: Wykres funkcji f.



Rys. 14.9: Wykres funkcji Df.



Rys. 14.10: Wykres funkcji CDf.



Rys. 14.11: Wykres funkcji Af.

a operator B odbija względem osi Ox.

(b) Niech $f \in C[0,1]$. Wtedy

$$\|Af\|_{C[0,1]} = \sup_{x\in[0,1]} |-2f(x \oplus 0.25)| = 2 \sup_{x\in[0,1]} |f(x)| = 2\|f\|_{C[0,1]}.$$

Oznacza to, że operator A jest ograniczony i $\|A\| \leq 2$. Z drugiej strony, niech $g \in C[0,1]$, $g(x) = 1$ dla $x \in [0,1]$. Mamy $\|g\|_{C[0,1]} = 1$ oraz $Ag(x) = 2$, $x \in [0,1]$, zatem

$$\|A\| = \sup_{\substack{f\in C[0,1]\\ \|f\|\leq 1}} \|Af\|_{C[0,1]} \geq \|Ag\|_{C[0,1]} = 2,$$

co dowodzi, że $\|A\| = 2$.

(c) Zauważmy, że $A^2f(x) = 4f(x \oplus 0.5)$, $A^3f(x) = -8f(x \oplus 0.75)$ i $A^4f(x) = 16f(x \oplus 1) = 16f(x)$ dla $x \in [0,1]$. Stąd $A^{4n+k} = 2^{4n}A^k$ dla $n \geq 0$ oraz $k = 0,1,2,3$, co prowadzi do

$$\begin{aligned}
\mathrm{e}^A f(x) &= \sum_{n=0}^{\infty} \frac{A^n f(x)}{n!} = \\
&= \sum_{n=0}^{\infty} 2^{4n}\left(\frac{1}{(4n!)} I + \frac{1}{(4n+1)!} A + \frac{1}{(4n+2)!} A^2 + \frac{1}{(4n+3)!} A^3\right)(f)(x) = \\
&= \sum_{n=0}^{\infty} \left[\frac{2^{4n}}{(4n)!} f(x) - \frac{2^{4n+1}}{(4n+1)!} f(x \oplus 0.25) + \right. \\
&\qquad \left. + \frac{2^{4n+2}}{(4n+2)!} f(x \oplus 0.5) - \frac{2^{4n+3}}{(4n+3)!} f(x \oplus 0.75)\right],
\end{aligned}$$

dla dowolnego $x \in [0,1]$.

Zadanie 14.36. Niech $x = (\xi_i)_{i\geq 1} \in l^1$. Zauważmy, że

$$\begin{aligned}\|T_n x\|_{l^1} &= \left|\sum_{i=1}^{n} \xi_i\right| + \sum_{i=1}^{\infty} |\xi_{2i-1} + \xi_{2i}| + \sum_{i=1}^{\infty} |\xi_{2i-1} - \xi_{2i}| \leq \\ &\leq \sum_{i=1}^{n} |\xi_i| + \sum_{i=1}^{\infty} \big(|\xi_{2i-1}| + |\xi_{2i}|\big) + \sum_{i=1}^{\infty} \big(|\xi_{2i-1}| + |\xi_{2i}|\big) \leq \\ &\leq 3\|x\|_{l^1}.\end{aligned}$$

Z drugiej strony, niech

$$x = (1, 0, 0, \ldots).$$

Wtedy oczywiście $x \in l^1$ i $\|x\|_{l^1} = 1$. Ponadto

$$\|T_n x\|_{l^1} = \|(1, 1, 1, 0, 0, \ldots)\|_{l^1} = 3 = 3\|x\|_{l^1},$$

co dowodzi, że $\|T_n\| = 3$ dla $n \geq 1$.

Ponieważ $\lim_{n\to\infty} \sum_{i=1}^{n} \xi_i = \sum_{i=1}^{\infty} \xi_i$, to naturalnym kandydatem na granicę ciągu $(T_n)_{n\geq 1}$ wydaje się być operator $T\colon l^1 \to l^1$ dany wzorem

$$T(\xi_i)_{i\geq 1} = \left(\sum_{i=1}^{\infty} \xi_i, \xi_1 + \xi_2, \xi_1 - \xi_2, \xi_3 + \xi_4, \xi_3 - \xi_4, \ldots\right).$$

Ustalmy $x = (\xi_i)_{i\geq 1} \in l^1$ i zauważmy, że

$$\|T_n x - Tx\|_{l^1} = \left\|\left(\sum_{i=n+1}^{\infty} \xi_i\right), 0, 0, \ldots\right\|_{l^1} = \left|\sum_{i=n+1}^{\infty} \xi_i\right|.$$

Ponieważ $x \in l^1$, to $\sum_{i=n+1}^{\infty} \xi_i$ jest resztą szeregu zbieżnego, więc

$$\lim_{n\to\infty} \sum_{i=n+1}^{\infty} \xi_i = 0,$$

a zatem rzeczywiście $\lim_{n\to\infty} \|T_n x - Tx\|_{l^1} = 0$. Oznacza to, że operatory T_n zbiegają w mocnej topologii do operatora T.

Dla $n \geq 1$ połóżmy

$$x_n = (0, \ldots, 0, 1, 0, 0, \ldots),$$

gdzie jedynka występuje na miejscu o numerze $n + 1$. Wtedy

$$T_n x_n - T x_n = (1, 0, 0, \ldots),$$

skąd

$$\|T_n - T\| = \sup_{\substack{x\in l^1 \\ \|x\|_{l^1}=1}} \|T_n x - Tx\|_{l^1} \geq 1.$$

Pokazuje to, że operatory T_n nie mogą być zbieżne do T w normie operatorowej.

Skorowidz

Bibliografia

[1] R. A. Adams, J. J. F. Fournier. *Sobolev spaces*, tom 140 serii *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, wydanie drugie, 2003. (Pozycja cytowana na stronie 96.)

[2] B. Åkerberg. Classroom Notes: A Proof of the Arithmetic-Geometric Mean Inequality. *Amer. Math. Monthly*, 70(9):997–998, 1963. (Pozycja cytowana na stronie 14.)

[3] A. Bielecki. Une remarque sur la méthode de Banach–Cacciopoli–Tikhonov. *Bull. Polish Acad. Sci.*, 4:261–268, 1956. (Pozycja cytowana na stronie 26.)

[4] A. Bobrowski. *Functional Analysis for Probability and Stochastic Processes.* Cambridge University Press, Cambridge, 2005. (Pozycja cytowana na stronach 45, 62, 66, 83, 90, 129 i 136.)

[5] N. L. Carothers. *A Short Course on Banach Space Theory*, tom 64 serii *London Mathematical Society Student Texts.* Cambridge University Press, 2004. (Pozycja cytowana na stronach 9 i 129.)

[6] L. Debnath, P. Mikusiński. *Introduction to Hilbert spaces with applications.* Elsevier, Amsterdam, wydanie trzecie, 2005. (Pozycja cytowana na stronach 9 i 89.)

[7] R. G. Douglas. *Banach algebra techniques in operator theory*, tom 179 serii *Graduate Texts in Mathematics.* Springer-Verlag, New York, wydanie drugie, 1998. (Pozycja cytowana na stronie 62.)

[8] R. E. Edwards. *Functional Analysis. Theory and Applications.* Dover Publications, 1995. (Pozycja cytowana na stronach 9 i 26.)

[9] L. C. Evans. *Partial differential equations*, tom 19 serii *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, wydanie drugie, 2010. Tłumaczenie polskie „Równania różniczkowe cząstkowe”, Wydawnictwo Naukowe PWN, Warszawa 2004. (Pozycja cytowana na stronie 94.)

[10] W. Feller. *An Introduction to Probability Theory and Its Applications. I.* Wiley, New York, wydanie trzecie, 1970. Tłumaczenie polskie „Wstęp do rachunku

prawdopodobieństwa, tom 1", Państwowe Wydawnictwo Naukowe, Warszawa 1977. (Pozycja cytowana na stronach 124 i 132.)

[11] G. R. Goldstein. Derivation and physical interpretation of general boundary conditions. *Adv. Differential Equations*, 11(4):457–480, 2006. (Pozycja cytowana na stronie 94.)

[12] J. Górnicki. *Okruchy matematyki.* Wydawnictwo Naukowe PWN, Warszawa, 1995. (Pozycja cytowana na stronach 25 i 33.)

[13] E. Hille, R. S. Phillips. *Functional Analysis and Semi-Groups.* Amer. Math. Soc. Colloq. Publ. 31. Amer. Math. Soc., Providence, R. I., 1957. (Pozycja cytowana na stronie 58.)

[14] O. Kallenberg. *Foundations of Modern Probability.* Springer, 1997. (Pozycja cytowana na stronie 83.)

[15] S. Kantorovitz. *Introduction to modern analysis*, tom 8 serii *Oxford Graduate Texts in Mathematics.* Oxford University Press, Oxford, 2003. (Pozycja cytowana na stronie 9.)

[16] M. Kordos. *Wykłady z historii matematyki.* Script, Warszawa, 2006. (Pozycja cytowana na stronie 77.)

[17] K. Kuratowski. *Rachunek różniczkowy i całkowy*, tom 22 serii *Biblioteka Matematyczna.* Państwowe Wydawnictwo Naukowe, 1964 i 1967. (Pozycja cytowana na stronie 14.)

[18] N. N. Lebedev. *Special functions and their applications.* Dover Publications, Inc., New York, 1972. (Pozycja cytowana na stronie 91.)

[19] W. Lee. *The Economy of God.* Living Stream Ministry, Los Angeles, 1968. Tłumaczenie polskie „Boża ekonomia", Wydawnictwo Strumień Życia, Warszawa, 1994. (Pozycja cytowana na stronie 99.)

[20] J. Lindenstrauss, L. Tzafriri. *Classical Banach spaces. I*, tom 92 serii *Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas].* Springer-Verlag, Berlin-New York, 1977. (Pozycja cytowana na stronie 9.)

[21] J. Lindenstrauss, L. Tzafriri. *Classical Banach spaces. II*, tom 97 serii *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas].* Springer-Verlag, Berlin-New York, 1979. Function spaces. (Pozycja cytowana na stronie 9.)

[22] L. Maligranda. Why Hölder's inequality should be called Rogers' inequality. *Math. Inequal. Appl.*, 1(1):69–83, 1998. (Pozycja cytowana na stronie 14.)

[23] H. Marcinkowska. *Dystrybucje, przestrzenie Sobolewa, równania różniczkowe*, tom 75 serii *Biblioteka Matematyczna.* Wydawnictwo Naukowe PWN, 1993. (Pozycja cytowana na stronach 95 i 96.)

[24] K. Maurin. *Metody przestrzeni Hilberta*, tom 45 serii *Monografie Matematyczne*, Państwowe Wydawnictwo Naukowe, Warszawa, 1967. (Pozycja cytowana na stronie 9.)

[25] M. A. Pinsky. *Partial differential equations and boundary-value problems with applications*, tom 15 serii *Pure and Applied Undergraduate Texts*. American Mathematical Society, Providence, RI, 2011. (Pozycja cytowana na stronie 94.)

[26] S. Prus, A. Stachura. *Analiza funkcjonalna w zadaniach*. Wydawnictwo Naukowe PWN, Warszawa, 2007. (Pozycja cytowana na stronie 11.)

[27] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, wydanie drugie, 1976. Tłumaczenie polskie „Analiza rzeczywista i zespolona", Państwowe Wydawnictwo Naukowe, Warszawa 1986. (Pozycja cytowana na stronach 9, 116 i 129.)

[28] J. Rusinek. *Zadania z analizy funkcjonalnej z rozwiązaniami*. Wydawnictwo Uniwersytetu Kardynała Stefana Wyszyńskiego, Warszawa, wydanie drugie, 2006. (Pozycja cytowana na stronie 11.)

[29] R. A. Ryan. *Introduction to Tensor Products of Banach Spaces*. Springer, 2002. (Pozycja cytowana na stronie 73.)

[30] A. D. Sokal. A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly*, 118(5):450–452, 2011. (Pozycja cytowana na stronie 129.)

[31] H. Steinhaus. *Czem jest a czem nie jest matematyka*. Księgarnia Nakładowa H. Altenberga, 1923. (Pozycja cytowana na stronie 9.)

[32] W. Tatarkiewicz. *Historia filozofii*. Państwowe Wydawnictwo Naukowe, Warszawa, 1958. (Pozycja cytowana na stronach 16 i 17.)

[33] M. E. Taylor. *Partial differential equations I. Basic theory*, tom 115 serii *Applied Mathematical Sciences*. Springer, New York, wydanie drugie, 2011. (Pozycja cytowana na stronie 94.)

[34] H. F. Weinberger. *A first course in partial differential equations with complex variables and transform methods*. Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1965. (Pozycja cytowana na stronie 94.)

[35] P. Wojtaszczyk. *Banach spaces for analysts*, tom 25 serii *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1991. (Pozycja cytowana na stronie 9.)

[36] K. Yosida. *Functional Analysis*. Springer, 1965. (Pozycja cytowana na stronie 62.)

[37] J. Zemánek. A simple proof of the Weierstrass–Stone theorem. *Comment. Math. Prace Mat.*, 20(2):495–497 (loose errata), 1977/78. (Pozycja cytowana na stronie 63.)