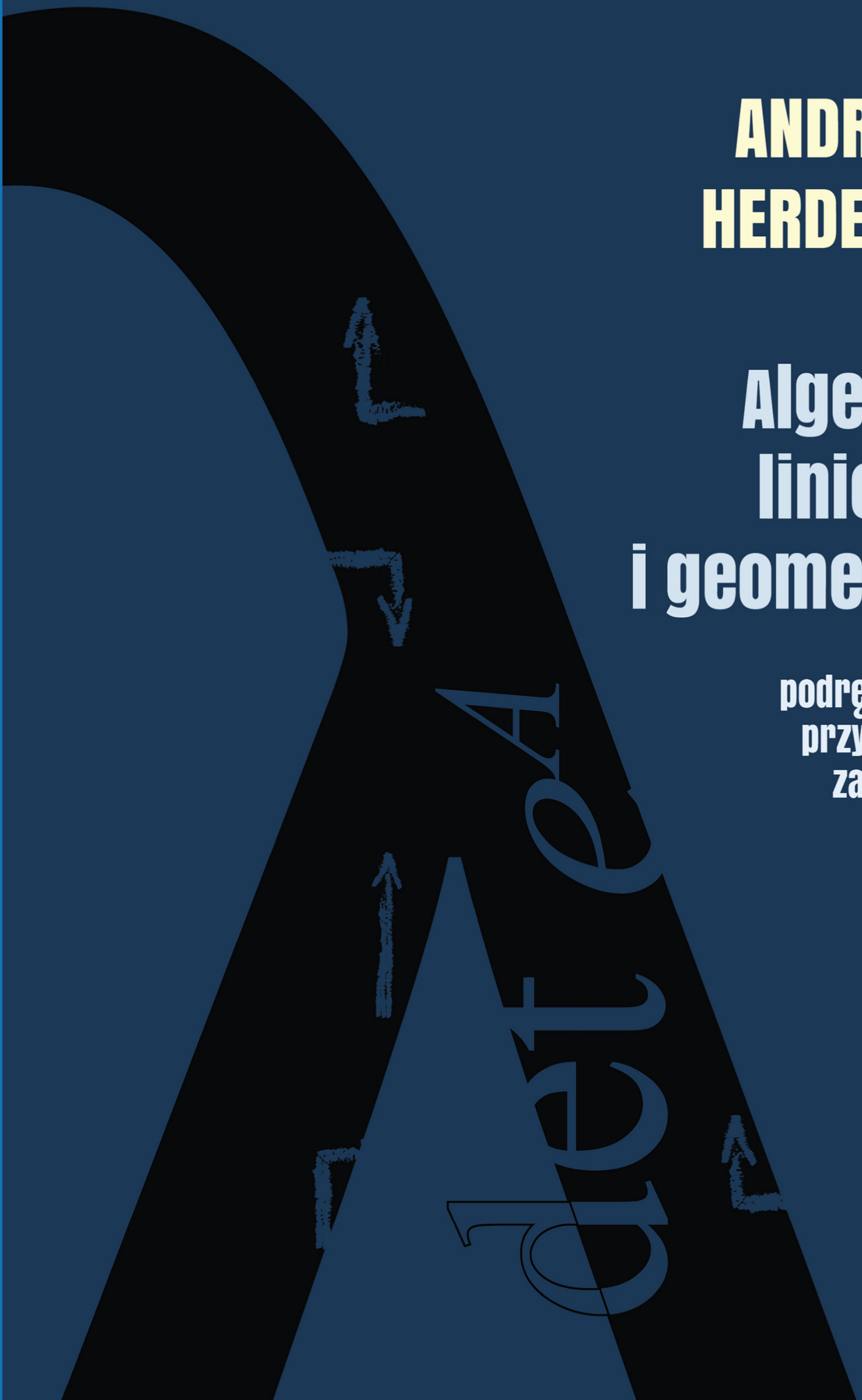
ANDRZEJ HERDEGEN

Algebra liniowa i geometria

podręcznik •
przykłady •
zadania •



Podręcznik niniejszy zawiera rozszerzony materiał kursu algebry liniowej i geometrii prowadzonego na Uniwersytecie Jagiellońskim dla studentów pierwszego roku fizyki, wraz ze zbiorem zadań ze wskazówkami i odpowiedziami. Poza ogólnym obyciem z logiką i charakterem języka matematyki na poziomie licealnym wykład nie stawia wymagań co do wiadomości wstępnych czytelnika.

Wybór zakresu materiału omówionego w książce jest dostosowany do potrzeb fizyków. Równocześnie zakres ten jest wystarczający dla większości studentów innych kierunków, szczególnie przyrodniczych i technicznych, gdzie język matematyczny jest wykorzystywany jako narzędzie.

Kolejność i sposób prezentacji materiału motywowane są przeznaczeniem książki. Podręcznik stara się pogodzić przekazanie głębszego zrozumienia struktur algebraicznych z nauką praktycznego stosowania opartych na nich metod, ilustrowaną licznymi przykładami.

Obecne wydanie elektroniczne jest poprawioną wersją drugiego drukowanego wydania książki.

Andrzej Herdegen

ALGEBRA LINIOWA I GEOMETRIA

◂ EIGENSPACE.PL ▸
KRZYSZKOWICE 2018

Andrzej Herdegen
Instytut Fizyki im. M. Smoluchowskiego
Uniwersytet Jagielloński
ul. prof. Stanisława Łojasiewicza 11
30-348 Kraków

andrzej.herdegen@uj.edu.pl

redakcja: autor

korekta: Magdalena Herdegen

skład i łamanie: autor
przy użyciu programu $\mathcal{AMS}$-LaTeX

projekt okładki: Sebastian Książkiewicz

wydanie III poprawione

ISBN 978-83-921987-2-7

eigenspace.pl

Krzyszkowice 2018

Pamięci Tymka

Autor oddaje tę książkę do użytku wszystkim zainteresowanym.

Równocześnie autor prosi wszystkich jej użytkowników
o wsparcie dzieła pomocy dzieciom,
którym możliwość studiowania nie będzie nigdy dana.

Szczegóły na stronie

eigenspace.pl

SPIS TREŚCI

WSTĘP

Kurs Podręcznik niniejszy zawiera rozszerzony materiał kursu algebry liniowej i geometrii prowadzonego na Uniwersytecie Jagiellońskim dla studentów pierwszego roku fizyki, wraz ze zbiorem zadań ze wskazówkami i odpowiedziami. Poza ogólnym obyciem z logiką i charakterem języka matematyki na poziomie licealnym wykład nie stawia wymagań co do wiadomości wstępnych czytelnika. Omówieniu matematycznego zaplecza właściwego przedmiotu wykładu poświęcony jest pierwszy rozdział książki. Dyskusja przestrzeni wektorowych i struktur w nich zadawanych stanowi właściwy przedmiot algebry liniowej, podjęty w drugim rozdziale. Dalsze rozdziały, nie wychodząc poza język algebry liniowej, omawiają zagadnienia określane na ogół jako geometryczne – w ramach geometrii *płaskich*. Z drugiej strony dyskusję przestrzeni unitarnych zamieszczoną w trzecim rozdziale (i uzupełnioną o iloczyn tensorowy w rozdziale czwartym) można uważać za przygotowanie do wykładu teorii przestrzeni Hilberta. W szóstym rozdziale zamieszczono omówienie dodatkowych zagadnień, na które zabrakło miejsca w głównym tekście.

Integralną częścią podręcznika są zgrupowane w osobnych punktach przykłady, które ilustrują poprzedzające je partie materiału, oraz zgrupowane w oddzielnym rozdziale zadania. Dyskusja zamieszczona w przykładach ma częstokroć charakter szkicowy, szczegółowe jej uzupełnienie czytelnik powinien traktować jako obowiązkowe ćwiczenie. To samo dotyczy niektórych prostszych dowodów pozostawionych do uzupełnienia przez czytelnika. Rozwiązywanie zadań powinno na bieżąco towarzyszyć studiowaniu kolejnych partii materiału. Zadania są zróżnicowane pod względem trudności; dla wielu z nich zamieszczono w kolejnym rozdziale wskazówki, a dla większości odpowiedzi.

Układ i zakres Zamierzeniem podręcznika jest przygotowanie podstaw do nowoczesnego ujęcia teorii fizycznych, a także do dalszych studiów matematycznych. Takie określenie przeznaczenia książki nie pozostawia wiele miejsca na dowolność wyboru zakresu kursu. Kanon materiału, jaki można pomieścić w rocznym kursie, jest niemal jednoznacznie zakreślony potrzebami języka matematycznego podstawowych teorii fizycznych, jak mechanika klasyczna i kwantowa, teorie pola. Równocześnie, w przekonaniu autora, materiał podręcznika, po rozszerzeniu o rozdział VI, pokrywa wszystkie podstawowe zagadnienia z zakresu algebry liniowej i geometrii wykładane dla początkowych lat studiów czy to czysto matematycznych, czy też przyrodniczo-technicznych.

Podręcznik kładzie nacisk na oba aspekty teorii matematycznej: teoretyczny i rachunkowy. Przedstawiamy nowoczesne ujęcie struktur algebraicznych, ale pokazujemy również jak teoretyczne wyniki stosować do praktycznych obliczeń. Dbamy o matematyczną elegancję wywodów, ale staramy się też wyrabiać intuicję. Z jednej strony stosujemy na ogół zasadę ekonomii: omawiać problemy tak wcześnie, jak to możliwe z punktu widzenia minimalnej struktury do tego niezbędnej. Z drugiej strony, nie pomijamy wskazania ich miejsca w ramach struktur i pojęć ogólniejszych, wykładanych na dalszych etapach. Na przykład, do omówienia podstawowych własności macierzy wystarczy znajomość pojęcia struktury liniowej na zbiorze funkcji o wartościach w zadanym ciele, abstrakcyjną definicję przestrzeni wektorowej można odłożyć na później. Ta kolejność wykładania ma również tę zaletę, że dostarcza gotowych przykładów struktur ogólniejszych. Inny przykład: do nowoczesnego omówienia układów równań liniowych niezbędna jest znajomość powłok liniowych, ale nie są konieczne inne struktury, w kontekście których takie układy mogą się pojawiać, jak odwzorowania liniowe i inne. Wskazanie roli układów liniowych w tych innych kontekstach można odłożyć na późniejszy etap. Dowody wszystkich twierdzeń są dostosowane do tego układu podręcznika.

Organizacja tekstu Podstawowymi jednostkami tekstu tej książki są paragrafy, numerowane kolejno na przestrzeni całego podręcznika. Paragrafy podzielono na punkty, numerowane osobno wewnątrz każdego paragrafu. Numeracja twierdzeń przebiega osobno wewnątrz każdego paragrafu, niezależnie od numeracji punktów (i obejmuje łącznie *twierdzenia*, *lematy* i *wnioski*). Odsyłacze do punktów lub twierdzeń mają przykładowo postać: p. 3, §7 lub tw. 6, §9 odpowiednio. Przy odwołaniach do punktów lub twierdzeń wewnątrz bieżącego paragrafu jego numer pomijamy. Zgrupowanie paragrafów w rozdziały ma służyć lepszej orientacji w tekście, ich numeracja nie występuje w odsyłaczach.

Nazwy pojęć definiowanych w tekście zapisano drukiem tłustym, pochyłym. Niedefiniowane pojęcia pierwotne, omawiane w §1, zapisano dla odróżnienia drukiem tłustym, prostym.

Podstawowe oznaczenia Zbiory liczbowe oznaczane będą następującymi symbolami:

liczby naturalne (nie obejmujące zera): $\mathbb{N}$;
liczby całkowite: $\mathbb{Z}$;
liczby wymierne: $\mathbb{Q}$;
liczby rzeczywiste: $\mathbb{R}$;
liczby zespolone (zdefiniowane w §4): $\mathbb{C}$.

Posługujemy się znakiem $:=$ dla oznaczenia równości, której lewa strona jest symbolem definiowanym przez tę równość. Przykładowo, jeśli przez $\mathbb{K}$ oznaczyć dowolny z powyższych zbiorów liczbowych, z wyjątkiem zbioru liczb naturalnych, to wprowadzamy oznaczenie:

$$\mathbb{K}^* := \mathbb{K} \setminus \{0\}\,.$$

Jeśli $\mathbb{K}$ symbolizuje $\mathbb{Q}$ lub $\mathbb{R}$, to oznaczamy ponadto

$$\mathbb{K}^+ := \{x \in \mathbb{K} \,|\, x > 0\} \quad \text{i} \quad \mathbb{K}^- := \{x \in \mathbb{K} \,|\, x < 0\}\,.$$

Będziemy niekiedy używać znaku $\equiv$ na oznaczenie równości w przypadkach, gdy równość przypomina jedynie przyjęte oznaczenia lub znaczenie symboli. Przykładowo, możemy napisać $\mathbb{N} \equiv \{1, 2, \ldots\}$.

Konwencje i interpretacje Istnieje kilka tradycyjnych punktów w algebrze liniowej, w których przyjęte konwencje, a nawet interpretacja pewnych pojęć, są na ogół nieco inne w znacznej części literatury fizycznej, niż w nowoczesnych podręcznikach „czystej" matematyki. Kończymy ten wstęp zestawieniem dokonanych w tej książce wyborów.

notacja wektorów i odwzorowań liniowych
Stosujemy konsekwentnie do oznaczania tych obiektów zwykłą czcionkę matematyczną (bez graficznych dodatków). Pogrubioną czcionką oznaczamy macierzowe reprezentacje tych obiektów w zadanych bazach. Strzałkę nad symbolem rezerwujemy na oznaczenie wektora łączącego dwa punkty w przestrzeni afinicznej.

współczynniki rozkładu wektora w bazie
W przeważającej części podręczników matematycznych te współczynniki określane są jako *współrzędne* (ang. *coordinates*), podczas gdy większa część literatury fizycznej nazywa je *składowymi* (ang. *components*). Wprowadzamy oba określenia, ale częściej stosujemy pierwsze z nich.

metryki półtoraliniowe
Według konwencji stosowanej w fizyce metryki te są liniowe w prawym, a antyliniowe w lewym argumencie, i taką definicję stosujemy w tej książce (użył

jej P. Dirac w swoim, powszechnie dziś przyjętym, sformułowaniu mechaniki kwantowej). Konwencja przeciwna – będąca standardem w literaturze matematycznej – jest oczywiście merytorycznie równoważna, choć mniej wygodna przy notowaniu transformacji macierzowych reprezentacji metryki i operatorów.

termin „afiniczny”

Określenie to można spotkać w dwóch znaczeniach: jako nazwę jednorodnej przestrzeni punktowej zbudowanej nad przestrzenią wektorową; oraz: na oznaczenie tych własności przestrzeni wektorowych i punktowych, które nie zależą od dodatkowych struktur takich jak metryka. Tutaj używamy tego terminu wyłącznie w pierwszym znaczeniu.

pseudotensory

Interpretacja tych obiektów jest zagadnieniem najcięższym gatunkowo spośród tutaj wymienionych, wykraczającym poza prosty wybór konwencji. Pojęcie pseudotensora jest nieobecne w nowoczesnych matematycznych podręcznikach algebry liniowej. Z drugiej strony, każdy student fizyki dowiaduje się na wykładach i z podręczników fizyki, że istnieją *wektory* („zwykłe”, lub: *biegunowe*) i *pseudowektory* (wektory *osiowe*), i że przykładami fizycznych realizacji tego drugiego typu są: moment pędu i pole magnetyczne. Ta niezgodność wynika stąd, że w literaturze fizycznej ciągle jeszcze tensory są często rozumiane jako tablice liczb transformujących się w określony sposób przy zmianie bazy, a dodanie odpowiedniego czynnika w prawie transformacyjnym prowadzi w tym ujęciu do pseudotensora. Na takie czynniki nie ma jednak miejsca w nowoczesnym rozumieniu tensorów. Stanowisko przyjęte w niniejszym podręczniku wyjaśnia ten problem, pokazując, że odpowiednia interpretacja usuwa trudność. Rozważamy jedynie „prawdziwe” tensory i pola tensorowe: każdy tensorowy model może być skonstruowany przy ich pomocy. Problem rozróżnienia pojawia się przy analizie *symetrii* modeli tensorowych. Przy afinicznej czynnej transformacji przestrzeni afinicznej istnieje naturalny sposób czynnego transformowania pól tensorowych w inne pola tensorowe. Dla uzyskania symetrii modelu tensorowego okazuje się jednak niekiedy koniecznym zmodyfikowanie tej czynnej transformacji dla niektórych spośród pól tensorowych występujących w modelu. Pola transformujące się w taki zmodyfikowany sposób to właśnie *pseudotensory*. Ta interpretacja pozostaje w zgodzie z faktyczną procedurą, w wyniku której przekonujemy się, że pole magnetyczne musi być reprezentowane pseudowektorem (jeśli ładunek jest skalarnym parametrem): należy sytuację doświadczalną poddać odbiciu w lustrze, a więc transformacji *czynnej*.

Książka jest poprawioną wersją drugiego wydania. Dziękuję wszystkim moim Kolegom i Studentom, którzy przyczynili się do usunięcia usterek poprzednich wersji książki.

I

WIADOMOŚCI WSTĘPNE

§1 Zbiory i zdania

1 Pojęcia pierwotne i podstawowe zasady

Zanim zaczniemy mówić o matematyce, zastanówmy się nad problemem z życia codziennego: chcemy wytłumaczyć sens słowa *szubienica* osobie słabo znającej język polski. Moglibyśmy zacząć tak: „szubienica to rusztowanie służące do ...". Jeśli jednak okazałoby się, że osoba ta nie zna również słowa *rusztowanie*, to musielibyśmy kontynuować: „rusztowanie to konstrukcja z belek ...", i tak dalej aż do wyjaśnienia wszystkich użytych przez nas niezrozumiałych wyrazów. Jednak pomyślne zakończenie takich wysiłków zakłada, że w końcu dojdziemy do pojęć danej osobie znanych. W przeciwnym wypadku zaczniemy po wielu krokach używać do określania nowych pojęć terminów, które chcieliśmy wytłumaczyć. W pewnym momencie tłumaczenia przestają mieć sens, i trzeba raczej wskazać przedmioty, i powiedzieć: „to jest belka, to sznur, a to hak". Terminy tego typu utworzą zasób pojęć danego języka, które możemy nazwać pojęciami pierwotnymi: pojęć tych nie definiujemy zdaniami języka – zakładamy ich znajomość „z życia", ale wszystkie inne pojęcia można przy ich pomocy określić. Staramy się, oczywiście, zminimalizować zasób tak rozumianych pojęć pierwotnych.

Ustalenie pojęć – jednych jako pierwotne, innych jako zdefiniowane za pomocą pierwotnych – to jednak dopiero wstęp do sensownej rozmowy. Po wytłumaczeniu słowa *szubienica* moglibyśmy powiedzieć do naszego rozmówcy: „morderców należy wieszać", i dopiero w tym momencie rozmowa zaczyna być ciekawa – w zdaniu tym przedstawiamy pewną teorię w zalążku, wypowiadając je mamy na myśli, że według nas jest ono prawdziwe. Nasz rozmówca mógłby jednak się z nami nie zgodzić. Wtedy, pragnąc uzasadnić nasz pogląd, moglibyśmy powiedzieć: „wynika to z zasady oko za oko, ząb za ząb". A na to nasz rozmówca:

„pozostaję przy swoim zdaniu, gdyż wynika ono z konwencji uzgodnionej przez państwa europejskie". W rzeczywistości prawdopodobnie każda ze stron użyje subtelniejszych argumentów, ale jedyne co każda z nich może zrobić, to powiązać swoje zdanie z grupą innych zasad, które uważa za bardziej oczywiste i podstawowe. Jak długo nie zostanie osiągnięta zgoda co do tego, który z tych konkurencyjnych systemów „oczywistych" zasad należy przyjąć, nie unikniemy rozbieżności ich szczegółowych zastosowań. Ciekawe, jak często nie dostrzegają tego adwersarze w podobnego typu dyskusjach.

2 Zbiory i zdania

Matematyka też posługuje się pewnym językiem – choć znacznie bardziej sformalizowanym od codziennego – i tam też zachodzi konieczność przyjęcia pojęć pierwotnych oraz podstawowych, niedowiedlnych, zasad.

Pojęciami pierwotnymi, na których buduje się matematykę w dzisiejszym ujęciu, są zbiory i zdania. Podstawowe zasady to przede wszystkim zasady logiki orzekające o prawdziwości zdań i poprawności wnioskowań oraz aksjomaty rachunku zbiorów. Jak już rozumiemy, pojęć pierwotnych nie możemy zdefiniować, ani zasad pierwotnych udowodnić, ale możemy, i powinniśmy je omówić, wiążąc z nimi intuicyjne wyobrażenia pomagające w posługiwaniu się nimi.

Pojęcie zbioru występuje nieodłącznie z pojęciem jego elementów i jest uściśleniem jego potocznego rozumienia. Pomyślmy o całym świecie „przedmiotów" – czy to materialnych, czy pomyślanych – i wybierzmy niektóre z nich odgradzając je w myśli od całej reszty świata. Otrzymany twór to właśnie **zbiór**, a wybrane przedmioty to **elementy** tego zbioru. Na oznaczenie zbiorów używamy na ogół dużych liter łacińskich, a na oznaczenie elementów – liter małych. Stwierdzenia 'przedmiot a jest elementem zbioru A' oraz 'przedmiot b nie jest elementem zbioru A' zapisujemy symbolicznie odpowiednio:

$$a \in A\,, \qquad b \notin A\,.$$

Zbiór złożony z elementów a, b, c, $\ldots$ zapisujemy też jako $\{a, b, c, \ldots\}$. Zbiór $\{a\}$ złożony z jednego elementu należy odróżniać od samego przedmiotu a – są to twory innej kategorii: a jest elementem zbioru $\{a\}$, ale $\{a\}$ nie jest elementem samego siebie, choć może być elementem innego zbioru, np. zbioru $\{\{a\}, \{b\}\}$. Możliwość tworzenia takich piętrowych konstrukcji stanowi pewne zagrożenie dla pojęcia zbioru. Rozważmy słynny przykład (pochodzący od B. Russella) próby zdefiniowania pewnego zbioru R jako 'zbioru wszystkich takich zbiorów S, że $S \notin S$'. Pytamy, czy $R \in R$? Jeśli tak, to R nie spełnia warunku, więc $R \notin R$. Jeśli nie, to R spełnia warunek, więc $R \in R$. Fiasko tej definicji polega na niemożności rozstrzygnięcia, czy dany przedmiot jest jego elementem czy też nie. Na ogół niebezpieczeństwo to eliminujemy zakładając pewne uniwersum wszystkich przedmiotów, które są dostępne do rozważań, czyli pewien niesprzecznie zadany maksymalny zbiór, który nazywamy często przestrzenią. Inne zbiory tworzymy wybierając z tej przestrzeni pewne elementy, a więc jednoznacznie wskazując, które z nich wchodzą do danego zbioru. Zbiory tak utworzone nazywamy **podzbiorami** tej przestrzeni. Zbiór, który nie zawiera żadnego elementu, nazywamy **zbiorem pustym** i oznaczamy $\varnothing$. Zbiór wszystkich podzbiorów przestrzeni X (do którego należy też podzbiór pusty) nazywamy **zbiorem potęgowym** tej przestrzeni i oznaczamy $P(X)$. Aksjomatyczna teoria mnogości (w jednej ze swych wersji) wyklucza ogólnie istnienie zbiorów, które byłyby własnymi elementami.

Pojęcie zdania odpowiada potocznemu pojęciu zdania sensownego, od którego nieodłączne są pojęcia prawdy i fałszu (nieprawdy). **Zdanie** to każda wypowiedź o przedmiotach świata – materialnego czy też pomyślanego – której można jednoznacznie przypisać jedną z dwóch wykluczających się cech: **prawdy** lub **fałszu**, zwanych **wartościami logicznymi**. Wartość ta może być dla danego zdania nieznana, ale wiadomo z natury stwierdzenia, że jedna, i tylko jedna z nich przysługuje temu zdaniu. Jeśli mamy przy tym powody, by sądzić, iż wartością tą jest *prawda*, to mówimy niekiedy o takim zdaniu, że jest hipotezą. Na przykład, zdanie „każda liczba parzysta większa od 2 jest sumą dwóch liczb pierwszych" jest hipotezą (Goldbacha): jej wartość logiczna, według najlepszej wiedzy autora, nie jest w chwili pisania tych słów komukolwiek znana, choć wiele wskazuje na to, że hipoteza ta jest prawdziwa. Zastrzeżenie o sensowności wypowiedzi jest istotne: np. wypowiedzi „R należy do R", gdzie R próbujemy określić jak wyżej w paradoksie Russella, nie można przypisać konsystentnie wartości logicznej, więc wypowiedź ta nie jest zdaniem w sensie logiki.

Podstawowe zasady matematyki to zasady logiki i podstawowe aksjomaty teorii zbiorów (teorii mnogości). Ponieważ nie jest naszym celem systematyzacja matematyki, a tym mniej logiki, nie będziemy się starali sprowadzać logiki czy teorii zbiorów do najprostszych elementów, ograniczymy się do wyliczenia głównych zasad operowania zdaniami i zbiorami.

3 Operacje logiczne

Istnieje pięć podstawowych sposobów tworzenia z danych zdań – zdań nowych. Niech p i q będą dowolnymi zdaniami. Wprowadzamy nowe zdania:

(i) 'nie p', ozn. $\sim p$, — ***negacja***;

(ii) 'p i q', ozn. $p \wedge q$, — ***koniunkcja***;

(iii) 'p lub q', ozn. $p \vee q$, — ***alternatywa***;

(iv) 'jeśli p to q', ozn. $p \Rightarrow q$, — ***implikacja***;

(v) ‘p wtedy, i tylko wtedy, gdy q’ , ozn. $p \Leftrightarrow q$, — ***równoważność***.

Wartości logiczne tych zdań zależą od wartości logicznych zdań tworzących je. Oznaczmy wartość *prawda* przez 1, a wartość *fałsz* przez 0. Wtedy wartości logiczne nowych zdań wyznaczone są tabelkami:

p	$\sim p$
0	1
1	0

,

p	q	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \Leftrightarrow q$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

.

Mówiąc słowami:

(i) $\sim p$ ma wartość przeciwną do p ;

(ii) $p \wedge q$ jest prawdziwe tylko wtedy, gdy oba zdania są prawdziwe;

(iii) $p \vee q$ jest prawdziwe zawsze z wyjątkiem przypadku, gdy oba zdania są fałszywe;

(iv) $p \Rightarrow q$ jest prawdziwe zawsze z wyjątkiem przypadku, gdy pierwsze zdanie jest prawdziwe, a drugie fałszywe; jeśli ta implikacja jest prawdziwa, to mówimy, że p ***pociąga*** q, lub że q ***wynika z*** p;

(v) $p \Leftrightarrow q$ jest prawdziwe, gdy oba zdania mają tę samą wartość logiczną; gdy tak jest, to mówimy, że te zdania są ***równoważne***.

Pojęciem, które wiąże pojęcia zbioru i zdania, jest **forma zdaniowa**. Niech będą dane zbiory $X, Y, \ldots$ i oznaczmy ich dowolne elementy symbolami x, y, ... odpowiednio. Utwórzmy wypowiedź, która oprócz słów desygnujących konkretne przedmioty zawiera symbole x, y, ... Wypowiedź tę nazywamy **formą zdaniową**, gdy staje się ona zdaniem po wstawieniu za te symbole dowolnych konkretnych elementów zbiorów $X, Y, \ldots$. Symbole x, y, ... nazywamy argumentami formy zdaniowej. Niech na przykład słowo *człowiek* oznacza dowolnego przedstawiciela zbioru wszystkich ludzi, a symbol MD jedną z dwuelementowego zbioru ról społecznych {morderca, detektyw}. Wypowiedź „człowiek jest MD” jest formą zdaniową: nie ma ona określonej wartości logicznej, ale jeśli wstawimy w odpowiednie miejsca dowolne konkretne elementy obu zbiorów, dostaniemy zdanie, np. fałszywe zdanie „Sherlock Holmes jest mordercą”. Inny przykład: wypowiedź „x jest podzielna przez 2” jest formą zdaniową, w której x symbolizuje dowolny element zbioru liczb naturalnych. Z form zdaniowych tworzy się nowe formy zdaniowe za pomocą tych samych operacji co dla zdań, i obowiązują dla nich te same reguły prawdziwości, jak te ujęte dla zdań w powyższe tabele.

Oznaczmy przez $\varphi(x)$ formę zdaniową o argumentach ze zbioru X. Wstawienie za x konkretnego elementu z X czyni z niej zdanie. Istnieją dwa dalsze, często używane, sposoby uzyskania zdań z niej zbudowanych. Każda z wypowiedzi

(i) dla każdego elementu x zbioru X zachodzi $\varphi(x)$,

(ii) istnieje element x zbioru X, dla którego zachodzi $\varphi(x)$

jest zdaniem (można sprawdzić, czy jest prawdziwa). Frazy słowne *dla każdego* oraz *istnieje* nazywamy **kwantyfikatorami**, pierwszą – **kwantyfikatorem ogólnym**, drugą – **kwantyfikatorem szczegółowym**. Powyższe zdania notujemy odpowiednio jako:

$$\begin{aligned} &\text{(i)} \qquad \forall x \in X : \varphi(x)\,, \\ &\text{(ii)} \qquad \exists x \in X : \varphi(x)\,. \end{aligned}$$

Ogólniej, jeśli dana jest forma zdaniowa $\varphi(x, y, \ldots)$ o argumentach z X, Y, ..., to wypowiedzi

$$\forall x \in X : \varphi(x, y, \ldots)\,, \quad \exists x \in X : \varphi(x, y, \ldots)$$

są formami zdaniowymi o argumentach z Y, ... Stąd, w szczególności, wypowiedzi

$$\forall x \in X\, \exists y \in Y : \varphi(x, y) \quad \text{oraz} \quad \exists y \in Y\, \forall x \in X : \varphi(x, y)$$

są zdaniami – różnymi! (dwukropki pomiędzy kwantyfikatorami pomijamy).

Zasady zaprzeczania zdań (lub form) zaczynających się kwantyfikatorami ujmują **prawa de Morgana dla kwantyfikatorów**. Mówią one, że prawdziwe są następujące równoważności

$$\begin{aligned} \sim (\forall x \in X : \varphi(x, y, \ldots)) &\Longleftrightarrow \exists x \in X : \sim \varphi(x, y, \ldots)\,, \\ \sim (\exists x \in X : \varphi(x, y, \ldots)) &\Longleftrightarrow \forall x \in X : \sim \varphi(x, y, \ldots)\,, \end{aligned}$$

czyli każda z form po dwóch stronach równoważności ma taką samą wartość logiczną dla wszystkich argumentów $y, \ldots$.

4 Definicje i twierdzenia

Matematyka nie składa się, oczywiście, z samych pojęć pierwotnych i podstawowych zasad – jej zasadniczą treścią są teorie dotyczące nowych pojęć wprowadzanych za pomocą definicji, formułowane w postaci twierdzeń. Choć określenia *definicja* i *twierdzenie* są znane każdemu, kto zetknął się z matematyką, skomentujemy zwięźle ich znaczenie.

Definicje to w istocie nic innego, jak pewne symbole i sformułowania słowne symbolizujące pojęcie złożone, utworzone z pojęć wcześniej określonych. Wprowadzimy, na przykład, nowy symbol $\exists!$ pokrewny kwantyfikatorowi szczegółowemu, używany do zbudowania zdania $\exists! x \in X : \varphi(x)$ czytanego: „***istnieje dokładnie jeden*** x w zbiorze X taki, że $\varphi(x)$” (podobnie ogólniej dla form zależnych od większej liczby argumentów). Wypowiedź ta jest, z definicji, skrótowym oznaczeniem zdania

$$\big(\exists x \in X : \varphi(x)\big) \wedge \big(\forall x_1, x_2 \in X : [\varphi(x_1) \wedge \varphi(x_2)] \Rightarrow x_1 = x_2\big) \,.$$

Twierdzeniem, najogólniej mówiąc, jest dowolne prawdziwe zdanie, lub forma zdaniowa, która jest prawdziwa dla wszystkich możliwych wartości swoich argumentów. Dowód twierdzenia to demonstracja, że istotnie jest ono zawsze prawdziwe, przy milczącym założeniu prawdziwości podstawowych zasad logiki. W rzeczywistości rozróżnienie pomiędzy twierdzeniem w postaci zdania lub w postaci formy zdaniowej jest mało istotne, gdyż każdą zawsze prawdziwą formę zdaniową można sprowadzić do postaci prawdziwego zdania przez poprzedzenie jej ogólnymi kwantyfikatorami. Na przykład, jeśli forma $\varphi(x)$, $x \in X$, jest zawsze prawdziwa, to twierdzenie

$$\varphi(x)$$

jest tym samym, co prawdziwe zdanie

$$\forall x \in X : \;\varphi(x) \,.$$

Układając tabelki wartości logicznych złożonych zdań w zależności od wartości zdań p, q i r wykazujemy prawdziwość następujących twierdzeń logiki:

(i)	$p \vee \sim p \,,$	***prawo wyłączonego środka,***
(ii)	$p \wedge (q \vee r) \iff [(p \wedge q) \vee (p \wedge r)] \,,$	***prawa***
(iii)	$p \vee (q \wedge r) \iff [(p \vee q) \wedge (p \vee r)] \,,$	***rozdzielności,***
(iv)	$\sim (\sim p) \iff p \,,$	***prawo podwójnego przeczenia,***
(v)	$\sim (p \wedge q) \iff [(\sim p) \vee (\sim q)] \,,$	***prawa***
(vi)	$\sim (p \vee q) \iff [(\sim p) \wedge (\sim q)] \,,$	***de Morgana,***
(vii)	$[(p \Rightarrow q) \wedge p] \Longrightarrow q \,,$	***zasada wnioskowania***,
(viii)	$(p \Rightarrow q) \iff \sim [p \wedge (\sim q)] \,,$	***zasada dowodów* nie wprost**,
(ix)	$(p \Rightarrow q) \iff [(\sim q) \Rightarrow (\sim p)] \,,$	***zasada kontrapozycji***,
(x)	$(p \Leftrightarrow q) \iff [(p \Rightarrow q) \wedge (q \Rightarrow p)] \,.$	

Twierdzenia matematyczne mają zwykle postać implikacji, jej poprzednik nazywamy założeniem, a następnik tezą. Użyteczność takiej formy wynika z zasady wnioskowania (punkt (vii) powyżej): jeśli wiemy, że prawdziwe jest twierdzenie oraz jego założenie, to prawdziwa jest teza. Tę samą zasadę stosuje się przy konstrukcji dowodów.

Zasady (viii) i (ix) są pomocne przy dowodach niektórych twierdzeń. Korzystając z pierwszej z nich, dowody *nie wprost* przeprowadza się w następujący sposób: zakładamy, że jest przeciwnie niż to mówi teza, i zakładamy również, że prawdziwe jest założenie. Jeśli dojdziemy do sprzeczności, to ta koniunkcja jest fałszywa, więc – na mocy powyższej zasady – twierdzenie prawdziwe. Zasady kontrapozycji używamy wtedy, gdy od dowodu bezpośredniego prostsze jest wykazanie, że z zaprzeczenia tezy wynika zaprzeczenie założenia. Wreszcie – zasada (x) pokazuje, że twierdzenie sformułowane jako równoważność to nic innego, jak dwie implikacje, w których założenie i teza zamieniają się rolami.

5 Algebra zbiorów

Niech będzie zadana przestrzeń X. Istnieje wzajemnie jednoznaczna odpowiedniość między podzbiorami tej przestrzeni a klasami wzajemnie równoważnych form zdaniowych o jednym argumencie z tej przestrzeni. Jeśli zadana jest forma $\varphi(x)$, to zbiór elementów $x \in X$ takich, że prawdziwe jest zdanie $\varphi(x)$ (tutaj x symbolizuje konkretny element, więc $\varphi(x)$ jest zdaniem), jest podzbiorem przestrzeni X; oznaczamy ten podzbiór $\{x \in X \mid \varphi(x)\}$, lub – jeśli przestrzeń jest oczywista – $\{x \mid \varphi(x)\}$. Zauważmy, że jeśli forma zdaniowa $\psi(x)$ jest równoważna formie $\varphi(x)$ (dla wszystkich $x \in X$), to $\{x \in X \mid \psi(x)\} = \{x \in X \mid \varphi(x)\}$. Odwrotnie, dla danego podzbioru A oznaczmy przez φ_A formę

$$x \in A\,.$$

Wtedy zbiór A może być zapisany jako:

$$A = \{x \in X |\, \varphi_A(x)\}\,.$$

Przy tych oznaczeniach definiujemy ***dopełnienie*** $A^\complement$ ***zbioru*** A oraz ***przecięcie*** $A \cap B$ i ***sumę*** $A \cup B$ zbiorów A i B za pomocą odpowiadających im form:

$$\varphi_{A^\complement} :=\sim \varphi_A\,, \quad \varphi_{A\cap B} := \varphi_A \wedge \varphi_B\,, \quad \varphi_{A\cup B} := \varphi_A \vee \varphi_B\,.$$

Twierdzenia logiki (i) – (vi) z poprzedniego punktu przekładają się teraz natychmiast na własności operacji na zbiorach:

$$
\begin{aligned}
&\text{(i)} && A \cup A^{\complement} = X\,, \\
&\text{(ii)} && A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\,, \\
&\text{(iii)} && A \cup (B \cap C) = (A \cup B) \cap (A \cup C)\,, \\
&\text{(iv)} && (A^{\complement})^{\complement} = A\,, \\
&\text{(v)} && (A \cap B)^{\complement} = A^{\complement} \cup B^{\complement}\,, \\
&\text{(vi)} && (A \cup B)^{\complement} = A^{\complement} \cap B^{\complement}\,.
\end{aligned}
$$

Definiujemy dalej ***różnicę*** $A \setminus B$ zbiorów A i B jako

$$A \setminus B := A \cap B^{\complement}\,.$$

Korzystając z tożsamości (ii), (v) i (vi) pokazuje się, że

$$
\begin{aligned}
C \setminus (A \cap B) &= (C \setminus A) \cup (C \setminus B)\,, \\
C \setminus (A \cup B) &= (C \setminus A) \cap (C \setminus B)\,.
\end{aligned}
$$

Zwróćmy uwagę, że jeśli $C = X$, to $X \setminus A = A^{\complement}$, i wtedy ostatnie dwie tożsamości sprowadzają się do tożsamości (v) i (vi).

Na język zbiorów przekłada się też pojęcie implikacji. Definiujemy zdanie: A ***zawiera się w*** B (lub A ***jest podzbiorem*** B), symbolicznie $A \subseteq B$, za pomocą równoważności

$$A \subseteq B \iff \forall x \in X : \varphi_A(x) \Rightarrow \varphi_B(x)\,.$$

Jeśli $B = X$, to $\varphi_B(x)$ jest prawdziwe dla wszystkich $x \in X$, więc prawa strona równoważności jest prawdziwa (jeśli następnik w implikacji jest prawdziwy, to implikacja jest prawdziwa niezależnie od wartości poprzednika). Zatem w tym przypadku pojęcie podzbioru pokrywa się z wcześniej wprowadzonym pojęciem podzbioru przestrzeni. Za pomocą pierwszej z poniższych równoważności wprowadzamy też oznaczenie zaprzeczenia zdania $A \subseteq B$:

$$A \nsubseteq B \iff \sim (A \subseteq B) \iff A \setminus B \neq \varnothing\,,$$

a druga równoważność jest twierdzeniem, którego dowód zostawiamy jako ćwiczenie dla czytelnika.

W celu wprowadzenia ostatniego z niezbędnych pojęć rachunku zbiorów musimy wyjaśnić pojęcie pary uporządkowanej (x, y), gdzie x jest elementem przestrzeni X, a y – przestrzeni Y. Intuicyjnie to pojęcie jest oczywiste – dwa elementy, jeden z nich wskazujemy jako pierwszy, drugi jako drugi. Intuicję tę można ująć w następujący formalny warunek: pary uporządkowane to obiekty (x, y) o następującej własności:

$$(x, y) = (x', y') \Longleftrightarrow x = x' \wedge y = y' .$$

Warunek ten jest jedyną i wystarczającą do dalszych rozważań cechą par. Warto jednak przekonać się, że para uporządkowana może być skonstruowana jako szczególnego typu zbiór, a więc jej wprowadzenie nie wymaga wyjścia poza ramy teorii zbiorów. Są możliwe różne konstrukcje tego typu, podajemy tzw. definicję Kuratowskiego: ***para (uporządkowana)*** (x, y) to zbiór $\{\{x\}, \{x, y\}\}$. Elementem pierwszym w parze (x, y) jest ten, który jest elementem obu zbiorów, będących elementami podanego zbioru, co jednoznacznie określa kolejność. Fakt ten zasadniczo odróżnia pojęcie pary uporządkowanej (x, y) od zbioru $\{x, y\}$, w którym kolejność zapisania elementów jest nieistotna. Zbiór wszystkich par (x, y), gdzie $x \in X$ i $y \in Y$, nazywamy ***iloczynem kartezjańskim przestrzeni*** X ***i*** Y i oznaczamy $X \times Y$. Jeśli $A \subseteq X$ i $B \subseteq Y$, to określamy też ***iloczyn kartezjański zbiorów*** A ***i*** B jako

$$A \times B = \{(x, y) \in X \times Y | \, x \in A \wedge y \in B\} .$$

W podobny sposób określamy iloczyn kartezjański większej liczby przestrzeni $X_1 \times \ldots \times X_n$ oraz iloczyn kartezjański ich podzbiorów.

6 Przykłady

(i) Implikacja
Zdania:

$$(3 \text{ jest podzielne przez } 2) \Rightarrow \big((-1)^3 = 1\big) ,$$
$$(3 \text{ jest podzielne przez } 2) \Rightarrow \big((-1)^2 = 1\big)$$

są prawdziwe – z fałszywych przesłanek można wywnioskować zarówno zdanie prawdziwe, jak i fałszywe.

(ii) Twierdzenie
Niech rozważaną przestrzenią będzie przestrzeń $\mathbb{R}$ liczb rzeczywistych. Forma zdaniowa o argumencie $x \in \mathbb{R}$:

$$\exists \lim_{n \to \infty} x^n$$

zadaje podzbiór przestrzeni $\mathbb{R}$. Łatwo pokazać, że podzbiór ten to przedział $(-1, 1\rangle$. Stąd forma zdaniowa o argumencie $x \in \mathbb{R}$

$$\big(\exists \lim_{n \to \infty} x^n\big) \Leftrightarrow \big(x \in (-1, 1\rangle\big)$$

jest prawdziwa dla wszystkich argumentów x, a więc jest twierdzeniem. Można je zapisać w postaci prawdziwego zdania

$$\forall x \in \mathbb{R}: \ \Big[\big(\exists \lim_{n\to\infty} x^n\big) \Leftrightarrow \big(x \in (-1, 1\rangle\big)\Big].$$

(iii) Przecięcie i suma dowolnej rodziny podzbiorów
Niech I będzie pewnym zbiorem, nazwiemy go zbiorem indeksów. Rodziną podzbiorów przestrzeni X indeksowaną zbiorem I nazywamy każde przyporządkowanie (czyli 'odwzorowanie' – formalną definicję tego pojęcia znajdzie czytelnik poniżej, w punkcie 6, §2) $I \ni i \mapsto A_i \subseteq X$. Suma i przecięcie zbiorów tej rodziny są podzbiorami przestrzeni X określonymi za pomocą następujących form zdaniowych:

$$x \in \bigcup_{i\in I} A_i \Longleftrightarrow \exists i \in I : x \in A_i\,,$$
$$x \in \bigcap_{i\in I} A_i \Longleftrightarrow \forall i \in I : x \in A_i\,.$$

(iv) Podzbiory i formy zdaniowe
Rozważmy przestrzeń $\mathbb{Z}$ liczb całkowitych. Forma zdaniowa o jednym argumencie x z tej przestrzeni:

$$\text{istnieje liczba } k \in \mathbb{Z} \text{ taka, że } x = 2k$$

definiuje zbiór liczb nazywanych parzystymi. Dopełnieniem tego zbioru w przestrzeni $\mathbb{Z}$ jest zbiór liczb nieparzystych, zadany zaprzeczeniem powyższej formy zdaniowej, które to zaprzeczenie można w tym przypadku równoważnie sformułować jako:

$$\text{istnieje liczba } k \in \mathbb{Z} \text{ taka, że } x = 2k + 1.$$

Rozważmy teraz przestrzeń $\mathbb{R}$ liczb rzeczywistych. Zbiór liczb całkowitych parzystych można w niej zdefiniować tą samą co wyżej formą zdaniową, gdzie teraz argument $x \in \mathbb{R}$. Dopełnienie tego zbioru w przestrzeni $\mathbb{R}$ jest jednak większe niż w poprzednim przypadku, i może być zapisane jako

$$\bigcup_{k\in\mathbb{Z}} (2k, 2k+2)\,.$$

(v) Kontrapozycja
Formy zdaniowe o argumencie $p \in \mathbb{N} \setminus \{1, 2\}$:

$$(p \text{ jest liczbą parzystą}) \Rightarrow (\exists \text{ liczby pierwsze } m, n : p = m + n)\,,$$
$$(\forall \text{ liczb pierwszych } m, n : p \neq m + n) \Rightarrow (p \text{ jest liczbą nieparzystą})$$

mają na mocy zasady kontrapozycji tę samą wartość logiczną. Po podstawieniu za p liczby nieparzystej, a także wielu spośród liczb parzystych otrzymujemy zdania prawdziwe. Jednak dla nieskończenie wielu liczb parzystych wartość logiczna otrzymanych zdań nie jest znana, ale dla danej liczby jednakowa dla obu zdań.

§2 Relacje. Odwzorowania

1 Relacje

Niech dany będzie zbiór X. ***Relacją (binarną) w zbiorze*** X nazywamy każdy podzbiór iloczynu kartezjańskiego $X \times X$. Niech będzie zadana relacja $\mathcal{R} \subseteq X \times X$. O każdej parze $(x, y) \in \mathcal{R}$ mówimy, że x jest w relacji $\mathcal{R}$ z y, i piszemy równoważnie $x \,\mathcal{R}\, y$. Mówimy też, że relacja $\mathcal{R}$ działa w zbiorze X.

Wyróżniamy następujące własności relacji. Mówimy, że relacja $\mathcal{R} \subseteq X \times X$ jest:

(i) ***zwrotna***, gdy $\forall x \in X : x \,\mathcal{R}\, x$;

(ii) ***symetryczna***, gdy $\forall x, y \in X : x \,\mathcal{R}\, y \Rightarrow y \,\mathcal{R}\, x$;

(iii) ***antysymetryczna***, gdy $\forall x, y \in X : x \,\mathcal{R}\, y \wedge y \,\mathcal{R}\, x \Rightarrow x = y$;

(iv) ***przechodnia***, gdy $\forall x, y, z \in X : x \,\mathcal{R}\, y \wedge y \,\mathcal{R}\, z \Rightarrow x \,\mathcal{R}\, z$;

(v) ***spójna***, gdy $\forall x, y \in X : x \,\mathcal{R}\, y \vee y \,\mathcal{R}\, x$.

2 Relacje równoważności

Mówimy, że relacja jest ***relacją równoważności***, gdy jest ona zwrotna, symetryczna i przechodnia. Najprostszym przykładem relacji tego typu jest relacja identyczności zadana zbiorem $\mathcal{R}_{\text{id}} = \{(x, x) \mid x \in X\}$, dla której

$$x \,\mathcal{R}_{\text{id}}\, y \Longleftrightarrow x = y\,.$$

Spełnienie warunków definicyjnych jest oczywiste. Relacja ta jest minimalną relacją równoważności – ze względu na zwrotność każda inna relacja równoważności musi ją zawierać.

Jednym z powodów doniosłości relacji równoważnościowych jest fakt, że dzielą one zbiory, w których działają, na podzbiory w taki sposób, że wewnątrz każdego z nich wszystkie elementy są identycznie względem siebie usytuowane. Dla każdego $x \in X$ zdefiniujmy, mianowicie, następujący podzbiór zbioru X: ***klasą równoważności*** (lub ***klasą abstrakcji***) ***elementu*** x nazywamy zbiór

$$[x]_{\mathcal{R}} := \{y \in X \mid y \,\mathcal{R}\, x\}\,.$$

Będziemy też pisać $[x]_{\mathcal{R}} \equiv [x]$, jeśli relacja jest ustalona. Ponieważ $x \in [x]$, to $X = \bigcup_{x\in X} [x]$, więc klasy pokrywają cały zbiór X. Pokrycie to jest rozłączne, co wynika z następującego twierdzenia.

Twierdzenie 1. *Jeśli $x \,\mathcal{R}\, y$, to $[x] = [y]$. W przeciwnym przypadku $[x] \cap [y] = \varnothing$.*

Dowód. Jeśli $x \,\mathcal{R}\, y$, to z symetryczności również $y \,\mathcal{R}\, x$, więc z przechodniości mamy wtedy $z \,\mathcal{R}\, x \Longleftrightarrow z \,\mathcal{R}\, y$, co dowodzi pierwszego stwierdzenia. Dowód drugiego stwierdzenia z zasady kontrapozycji: jeśli istnieje $z \in [x] \cap [y]$, to $x \,\mathcal{R}\, z$ i $z \,\mathcal{R}\, y$, więc $x \,\mathcal{R}\, y$. □

Otrzymane wyniki można zreasumować tak: przestrzeń X została podzielona na rozłączne zbiory, elementy różnych zbiorów nie wchodzą w relację ze sobą, a wewnątrz każdego ze zbiorów każde dwa elementy są w relacji ze sobą. Zbiór klas równoważności nazywamy ***zbiorem ilorazowym***, i oznaczamy go

$$X/\mathcal{R} = \{[x] \mid x \in X\}\,.$$

Powyższe podsumowanie podziału zbioru na klasy pokazuje, że relacja równoważności zadająca ten podział może być z niego bezpośrednio odczytana. Co więcej, zachodzi następujące twierdzenie.

Twierdzenie 2. *Każde pokrycie przestrzeni zbiorami rozłącznymi zadaje jednoznacznie relację równoważności, której klasami są zbiory tego pokrycia.*

Dowód. Zadajemy relację warunkiem: $x \,\mathcal{R}\, y$ wtedy, i tylko wtedy, gdy x i y są elementami tego samego zbioru pokrycia. Łatwo sprawdzić, że jest to relacja równoważności, a jej klasy to zbiory pokrycia. Jednoznaczność z dyskusji poprzedzającej twierdzenie. □

3 Przykłady

(i) Symetria i przechodniość
W zbiorze ludzi relacja zadana przez: $a \,\mathcal{R}\, b \Longleftrightarrow$ (a jest bratem b) jest przechodnia, ale nie jest symetryczna. W tym samym zbiorze relacja $a \,\mathcal{R}'\, b \Longleftrightarrow$ (a i b są rodzeństwem) jest symetryczna i przechodnia.

(ii) Relacja równoważności
W zbiorze punktów płaszczyzny relacja $(x, y)\,\mathcal{R}\,(x', y') \Leftrightarrow y = y'$ jest relacją równoważności. Klasami równoważności są poziome proste.

(iii) Rozłączne pokrycie
Dla każdej liczby rzeczywistej $r \geq 0$ określamy podzbiór przestrzeni $\mathbb{R} \times \mathbb{R}$: $O_r = \{(x, y) \,|\, x^2 + y^2 = r^2\}$. Rodzina zbiorów $\{O_r \mid r \geq 0\}$ jest rozłącznym pokryciem przestrzeni $\mathbb{R} \times \mathbb{R}$, więc zadaje w niej relację równoważności. Relacja ta ma postać $(x, y)\,\mathcal{R}\,(x', y') \iff x^2 + y^2 = x'^2 + y'^2$.

(iv) Relacja mod m w zbiorze $\mathbb{Z}$
Dla każdego $m \in \mathbb{N} \setminus \{1\}$ określamy w zbiorze liczb całkowitych $\mathbb{Z}$ relację

$$k = l \mod m \iff k - l \text{ dzieli się przez } m$$

(czytamy: k równa się l modulo m). Relacja ta jest relacją równoważności, ma ona m różnych klas równoważności

$$[k] = \{k + lm \mid l \in \mathbb{Z}\}\,, \quad k = 0, 1, \ldots, m - 1\,, \quad \text{więc}$$
$$\mathbb{Z}/\mathrm{mod}\, m = \{[k] \mid k = 0, \ldots, m - 1\}\,.$$

(v) Relacja mod c w zbiorze $\mathbb{R}$
Dla każdego $c \in \mathbb{R}^+$ definiujemy w zbiorze $\mathbb{R}$ relację

$$x = y \mod c \iff x - y = kc\,, \quad k \in \mathbb{Z}\,.$$

Relacja ta jest relacją równoważności, jej klasy równoważności to

$$[x] = \{x + kc \mid k \in \mathbb{Z}\}\,, \quad \text{więc} \quad \mathbb{R}/\mathrm{mod}\, c = \{[x] \mid x \in \langle 0, c)\}\,.$$

4 Relacje porządkujące

Relacja $\mathcal{R}$ działająca w zbiorze X nazywa się ***relacją częściowego porządku***, gdy jest ona zwrotna, antysymetryczna i przechodnia. Mówimy wtedy, że ***zbiór X jest częściowo uporządkowany (relacją $\mathcal{R}$)***. Jeśli relacja częściowego porządku ma ponadto własność spójności, to nazywa się ***relacją porządku***, a zbiór X określamy wtedy jako ***uporządkowany (relacją $\mathcal{R}$)***. Za pierwowzór relacji porządku uważać można znak nierówności słabej wśród liczb naturalnych: $k \leq n$.

5 Przykłady

(i) Relacja zawierania zbiorów
Relacja zawierania w zbiorze podzbiorów danego zbioru X jest częściowym porządkiem, ale nie jest porządkiem: nie dla każdej pary zbiorów jeden z nich zawiera się w drugim.

(ii) Porządek leksykograficzny
W zbiorze par $(i, j) \in \mathbb{N} \times \mathbb{N}$ wprowadzamy relację:

$$(i_1, j_1) \preceq (i_2, j_2) \iff i_1 < i_2 \vee (i_1 = i_2 \wedge j_1 \leq j_2) .$$

Relacja ta jest relacją porządku, nazywamy ją ***porządkiem leksykograficznym***. Jej uogólnienie i odpowiednie zmodyfikowanie określa porządek wyrazów w słowniku.

6 Odwzorowania

Określiliśmy wyżej relację w zbiorze X jako podzbiór iloczynu kartezjańskiego $X \times X$. Definicja ta jest szczególnym przypadkiem ogólnej ***relacji binarnej***, którą określamy jako dowolny podzbiór iloczynu kartezjańskiego $X \times Y$ dwóch zadanych zbiorów X i Y. Notacja uogólnia się w oczywisty sposób do $\mathcal{R} \subseteq X \times Y$, $x \,\mathcal{R}\, y \iff (x, y) \in \mathcal{R}$. Wśród tych ogólniejszych relacji interesować nas będą jedynie relacje innego szczególnego typu, mianowicie takie, dla których każdy element zbioru X jest w relacji z dokładnie jednym elementem zbioru Y. Formalny zapis tego warunku ma następującą postać:

$$\forall x \in X \ \exists! y \in Y : x \,\mathcal{R}\, y .$$

Widocznym jest więc, że zadanie relacji tego typu to tyle samo, co wybranie dla każdego $x \in X$ dokładnie jednego $y \in Y$, oznaczmy go $f(x)$, który ma z nim wchodzić w relację. Relacja ta zadaje więc jednoznacznie przepis f, według którego dokonujemy wyboru. Oznaczamy ten typ relacji przez $\mathcal{R}_f$ i piszemy

$$\mathcal{R}_f = \{(x, f(x)) \mid x \in X\} .$$

Jeśli zadana jest taka relacja, to mówimy, że f jest ***odwzorowaniem zbioru*** X ***w zbiór*** Y i piszemy

$$f : X \mapsto Y , \ x \mapsto f(x) ,$$

lub w uproszczeniu

$$X \ni x \mapsto f(x) \in Y .$$

Zbiór X nazywamy ***dziedziną***, a zbiór Y ***przeciwdziedziną odwzorowania*** f. Relację $\mathcal{R}_f$ nazywamy ***grafikiem odwzorowania*** f. Jeśli para $(x, f(x)) \in \mathcal{R}_f$, to x nazywamy ***argumentem odwzorowania***, a $f(x)$ ***wartością odwzorowania*** dla argumentu x. Dla zadanego odwzorowania f ***obrazem zbioru*** $A \subseteq X$ nazywamy zbiór

$$f(A) := \{f(x) \mid x \in A\} \subseteq Y\,,$$

a ***przeciwobrazem zbioru*** $B \subseteq Y$ nazywamy zbiór

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\} \subseteq X\,.$$

(Część podręczników przeciwdziedziną odwzorowania f nazywa obraz $f(X)$, a więc zbiór wartości osiąganych przez to odwzorowanie; my będziemy konsekwentnie trzymać się konwencji wprowadzonej wyżej.) Niech X' będzie podzbiorem dziedziny, $X' \subseteq X$. Jeśli nie zmieniając przepisu $f(x)$ będziemy rozpatrywać tylko argumenty $x \in X'$, to otrzymamy odwzorowanie

$$f|_{X'} : X' \mapsto Y\,, x \mapsto f(x)$$

nazywane ***zacieśnieniem odwzorowania*** f do zbioru X'. Odwrotnie, niech $X \subseteq X''$. Każde odwzorowanie $F : X'' \mapsto Y$, którego zacieśnienie do zbioru X jest identyczne z f nazywamy ***rozszerzeniem odwzorowania*** f.

Twierdzenie 3. *Niech* $f : X \mapsto Y$, $A, A' \subseteq X$, $B, B' \subseteq Y$. *Wtedy*

$$\begin{aligned} A \subseteq A' &\Rightarrow f(A) \subseteq f(A')\,, & B \subseteq B' &\Rightarrow f^{-1}(B) \subseteq f^{-1}(B')\,,\\ f(A \cup A') &= f(A) \cup f(A')\,, & f^{-1}(B \cup B') &= f^{-1}(B) \cup f^{-1}(B')\,,\\ f(A \cap A') &\subseteq f(A) \cap f(A')\,, & f^{-1}(B \cap B') &= f^{-1}(B) \cap f^{-1}(B')\,. \end{aligned}$$

Dowód. Dowody wszystkich własności przebiegają według podobnych schematów – dla ilustracji metody udowodnimy tylko jedną z nich, pozostawiając pozostałe dowody jako ćwiczenie dla czytelnika. Jeśli $y \in f(A \cap A')$, to istnieje $x \in A \cap A'$ taki, że $f(x) = y$. Ale stąd $x \in A$ i $x \in A'$, więc $y \in f(A)$ i $y \in f(A')$, czyli $y \in f(A) \cap f(A')$. Wykazaliśmy, że $f(A \cap A') \subseteq f(A) \cap f(A')$. □

Zwróćmy uwagę, że dla dowolnego odwzorowania nie wszystkie elementy przeciwdziedziny muszą brać udział w relacji definiującej to odwzorowanie, a także, że inne mogą pojawiać się wielokrotnie. Obserwacja ta nasuwa wprowadzenie określeń odwzorowań dwóch szczególnych typów. Niech będzie dane odwzorowanie $f : X \mapsto Y$. Mówimy, że f jest ***surjekcją*** (lub ***odwzorowaniem surjektywnym***, lub ***odwzorowaniem zbioru*** X **na** ***zbiór*** Y), gdy $f(X) = Y$ (a więc każdy element przeciwdziedziny wchodzi w relację z pewnym elementem dziedziny). Łatwo zauważyć, że każde odwzorowanie można zmodyfikować bez zmiany przepisu funkcyjnego $f(x)$ tak, by otrzymać surjekcję: wystarczy w tym celu zawęzić przeciwdziedzinę do zbioru $f(X)$. Mówimy z kolei, że f jest ***injekcją*** (lub ***odwzorowaniem injektywnym***, lub ***odwzorowaniem różnowartościowym***), gdy dowolny element przeciwdziedziny może być w relacji co najwyżej z jednym elementem dziedziny. Zapisuje się ten warunek niekiedy tak:

$$\forall x_1, x_2 \in X : f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \,.$$

Możemy ten warunek sformułować również tak: przeciwobrazem dowolnego elementu przeciwdziedziny (dokładniej: jednoelementowego podzbioru przeciwdziedziny) jest zbiór co najwyżej jednoelementowy.

7 Bijekcje. Odwzorowania odwrotne. Odwzorowanie identycznościowe

Odwzorowanie, które jest równocześnie surjektywne i injektywne nazywamy ***bijekcją*** (lub ***odwzorowaniem bijektywnym***, lub ***odwzorowaniem wzajemnie jednoznacznym***). Jeśli $\mathcal{R}_f$ jest relacją zadającą odwzorowanie

$$f : X \mapsto Y \,,$$

to warunek bijektywności f powstaje z połączenia wymagań surjektywności i injektywności:

$$\forall y \in Y \ \exists! x \in X : x \,\mathcal{R}_f\, y \,.$$

Definiujemy ogólnie dla dowolnej relacji binarnej $\mathcal{R} \subseteq X \times Y$ relację do niej odwrotną $\mathcal{R}^{-1} \subseteq Y \times X$ za pomocą warunku

$$y \,\mathcal{R}^{-1}\, x \Longleftrightarrow x \,\mathcal{R}\, y \,.$$

Widzimy teraz, że warunek bijektywności odwzorowania f jest identyczny ze stwierdzeniem, iż $(\mathcal{R}_f)^{-1}$ jest relacją zadającą odwzorowanie zbioru Y w zbiór X. Każda bijekcja $f : X \mapsto Y$ określa więc jednoznacznie odwzorowanie

$$f^{-1} : Y \mapsto X \,, \ y \mapsto f^{-1}(y) \,,$$

nazywane ***odwzorowaniem odwrotnym do*** f, takie, że

$$(\mathcal{R}_f)^{-1} = \mathcal{R}_{f^{-1}} = \{(y, f^{-1}(y)) \mid y \in Y\} \,.$$

Z łatwością udowodnimy teraz następujące własności bijekcji.

Twierdzenie 4. *Jeśli $f : X \mapsto Y$ jest bijekcją, to $f^{-1} : Y \mapsto X$ jest też bijekcją, i wtedy $(f^{-1})^{-1} = f$. Dla każdej pary $(x, y) \in X \times Y$ zachodzi w tym przypadku równoważność:*

$$y = f(x) \Longleftrightarrow x = f^{-1}(y) \,.$$

Dowód. Wiemy już, że jeśli f jest bijekcją, to $(\mathcal{R}_f)^{-1}$ zadaje f^{-1}. Łatwo zauważyć, że warunek zadawania przez $\mathcal{R}_f$ odwzorowania jest identyczny z warunkiem bijektywności f^{-1}. Wprost z definicji relacji odwrotnej widać, że dla każdej relacji $\mathcal{R}$ mamy $(\mathcal{R}^{-1})^{-1} = \mathcal{R}$. W szczególności $(\mathcal{R}_{f^{-1}})^{-1} = \mathcal{R}_f$, więc $(f^{-1})^{-1} = f$. Ostatnie stwierdzenie tezy jest wynikiem ciągu równoważności:

$$y = f(x) \iff (x, y) \in \mathcal{R}_f \iff (y, x) \in (\mathcal{R}_f)^{-1} \iff x = f^{-1}(y)\,.$$

□

Omawiając relacje określone w zbiorze X wprowadziliśmy relację identycznościową $\mathcal{R}_{\mathrm{id}}$. Widzimy teraz, że relacja ta zadaje ***odwzorowanie identycznościowe zbioru*** X

$$\mathrm{id}_X : X \mapsto X\,,\ x \mapsto \mathrm{id}_X(x) = x\,.$$

Odwzorowanie to w sposób oczywisty spełnia warunki bijektywności, a odwrotne do niego jest z nim identyczne.

8 Przykłady

(i) Własności odwzorowania i jego zacieśnienia

Relacja

$$\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}\,, \quad x\,\mathcal{R}\,y \ \Leftrightarrow\ \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1\,, \quad a, b > 0\,,$$

nie zadaje odwzorowania. Relacja $\mathcal{R}'$ określona tym samym równaniem, ale rozpatrywana jako podzbiór przestrzeni $\langle -a, +a\rangle \times \langle 0, +\infty)$ zadaje odwzorowanie

$$f : \langle -a, +a\rangle \mapsto \langle 0, +\infty)\,, \quad f(x) = \frac{b}{a}\sqrt{a^2 - x^2}\,.$$

Odwzorowanie f nie jest injekcją, ani surjekcją. Niech

$$-a \le s \le t \le +a \ \text{ i oznaczmy } \ m = \min\{|s|, |t|\}\,, \quad M = \max\{|s|, |t|\}\,.$$

Obraz zbioru $\langle s, t\rangle$ w odwzorowaniu f jest równy

$$f(\langle s, t\rangle) = \begin{cases} \left\langle \frac{b}{a}\sqrt{a^2 - M^2},\, b \right\rangle, & \text{gdy } \ st < 0\,, \\ \left\langle \frac{b}{a}\sqrt{a^2 - M^2},\, \frac{b}{a}\sqrt{a^2 - m^2} \right\rangle, & \text{gdy } \ st \ge 0\,. \end{cases}$$

Stąd dostajemy w szczególności

$$f(\langle -a, +a/2\rangle \cap \langle -a/2, +a\rangle) = f(\langle -a/2, +a/2\rangle) = \left\langle \frac{\sqrt{3}}{2}b, b \right\rangle,$$
$$f(\langle -a, +a/2\rangle) \cap f(\langle -a/2, +a\rangle) = \langle 0, b\rangle\,,$$

co jest ilustracją faktu, że dla dowolnego odwzorowania $f(A \cap A') \subseteq f(A) \cap f(A')$, ale w ogólności $f(A \cap A') \neq f(A) \cap f(A')$ (tw. 3).

Niech dalej $r \geq 0$. Przeciwobraz zbioru $\langle r, +\infty)$ jest równy

$$f^{-1}(\langle r, +\infty)) = \begin{cases} \varnothing\,, & \text{gdy} \quad r > b\,, \\ \Big\langle -\dfrac{a}{b}\sqrt{b^2 - r^2},\, +\dfrac{a}{b}\sqrt{b^2 - r^2} \Big\rangle\,, & \text{gdy} \quad r \leq b\,. \end{cases}$$

Oznaczmy przez $\tilde{f}$ zacieśnienie odwzorowania f do podzbioru dziedziny $\langle 0, a \rangle$. Odwzorowanie $\tilde{f}$ jest injektywne, ale nie jest surjekcją. Obraz całej jego dziedziny jest równy $\tilde{f}(\langle 0, a \rangle) = \langle 0, b \rangle$. Niech teraz $\hat{f}$ różni się od $\tilde{f}$ tylko zawężeniem przeciwdziedziny do zbioru $\langle 0, b \rangle$. Dostajemy odwzorowanie

$$\hat{f} : \langle 0, a \rangle \mapsto \langle 0, b \rangle\,, \quad \hat{f}(x) = \frac{b}{a}\sqrt{a^2 - x^2}\,,$$

które jest bijekcją. Odwzorowanie odwrotne jest równe

$$\hat{f}^{-1} : \langle 0, b \rangle \mapsto \langle 0, a \rangle\,, \quad \hat{f}^{-1}(y) = \frac{a}{b}\sqrt{b^2 - y^2}\,.$$

W szczególnym przypadku gdy $a = b$, odwzorowanie $\hat{f}$ jest bijekcją zbioru $\langle 0, a \rangle$ na samego siebie, o dodatkowej szczególnej własności $\hat{f}^{-1} = \hat{f}$.

(ii) Część całkowita i ułamkowa liczby rzeczywistej
Definiujemy odwzorowania $[.] : \mathbb{R} \mapsto \mathbb{Z}$ i $m : \mathbb{R} \mapsto \langle 0, 1)$ za pomocą warunków

$$x - [x] \in \langle 0, 1)\,, \quad m(x) = x - [x]\,.$$

Oba odwzorowania są surjekcjami, ale żadne nie jest injekcją.

(iii) Rzeczywiste funkcje wielomianowe
Wybierzmy ciąg liczb rzeczywistych $(a_0, a_1, \ldots, a_n)$, $a_n \neq 0$. Odwzorowanie

$$\mathbb{R} \ni x \mapsto P(x) = a_0 + a_1 x + \ldots + a_n x^n \in \mathbb{R}$$

nazywamy funkcją wielomianową stopnia n, a liczby a_i – jej współczynnikami. (Dokładniejszą dyskusję wielomianów i funkcji wielomianowych znajdzie czytelnik w punktach 13 i 14, §3). Wszystkie funkcje wielomianowe stopnia parzystego są niesurjektywne i nieinjektywne. Wszystkie funkcje wielomianowe stopnia nieparzystego są surjektywne, ale tylko niektóre spośród nich są injektywne.

(iv) Bijektywne zacieśnienie
Odwzorowanie $\mathbb{R} \ni x \mapsto \sin x \in \mathbb{R}$ jest niesurjektywne i nieinjektywne, ale odwzorowanie $\langle -(\pi/2), +(\pi/2) \rangle \ni x \mapsto \sin x \in \langle -1, +1 \rangle$ jest bijekcją; odwzorowanie odwrotne do niego oznacza się arcsin.

9 Składanie odwzorowań

Niech dane będą odwzorowania $f : X \mapsto Y$ i $g : Y \mapsto Z$. ***Złożeniem odwzorowań** f **i** g* nazywamy odwzorowanie zadane następującym przepisem:

$$g \circ f : X \mapsto Z\,, \ \ (g \circ f)(x) = g\big(f(x)\big)\,.$$

Z definicji odwzorowania identycznościowego mamy natychmiast

$$f \circ \mathrm{id}_X = f\,, \quad \mathrm{id}_Y \circ f = f\,.$$

Jeśli ponadto $h : Z \mapsto W$, to łatwo sprawdzić, że

$$h \circ (g \circ f) = (h \circ g) \circ f\,,$$

czyli składanie odwzorowań jest łączne.

Jeśli $A \subseteq X$ i $B \subseteq Z$, to zachodzą oczywiste równości

$$(g \circ f)(A) = g\big(f(A)\big)\,, \quad (g \circ f)^{-1}(B) = f^{-1}\big(g^{-1}(B)\big)\,.$$

Z pomocą tych identyczności udowodnimy następujące twierdzenie.

Twierdzenie 5. *Niech $f : X \mapsto Y$ i $g : Y \mapsto Z$. Wtedy*

(i) jeśli f i g są surjekcjami (injekcjami), to $g \circ f$ jest surjekcją (odpowiednio: injekcją); jeśli f i g są bijekcjami, to $g \circ f$ jest bijekcją i

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}\,;$$

(ii) jeśli $g \circ f$ jest surjekcją, to g jest surjekcją;

(iii) jeśli $g \circ f$ jest injekcją, to f jest injekcją.

Dowód. (i) Jeśli $f(X) = Y$ i $g(Y) = Z$, to $(g \circ f)(X) = g\big(f(X)\big) = g(Y) = Z$. Jeśli zarówno dla f jak i g przeciwobrazem dowolnego elementu jest zbiór co najwyżej jednoelementowy, to jest to prawdą również dla złożenia. Stąd łącznie: jeśli oba odwzorowania są bijekcjami, to ich złożenie jest bijekcją. Ponadto mamy ciąg równoważności $z = g\big(f(x)\big) \iff g^{-1}(z) = f(x) \iff f^{-1}\big(g^{-1}(z)\big) = x$, więc $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
(ii) Jeśli $(g \circ f)(X) = Z$, to $Z = g\big(f(X)\big) \subseteq g(Y) \subseteq Z$, więc $g(Y) = Z$.
(iii) Niech $f(x_1) = f(x_2)$. Wtedy $g\big(f(x_1)\big) = g\big(f(x_2)\big)$, więc z injektywności $g \circ f$ mamy $x_1 = x_2$. □

Jeśli $f : X \mapsto Y$ jest bijekcją, to, jak wiemy, $y = f(x) \iff x = f^{-1}(y)$, więc $\forall x \in X : x = f^{-1}\big(f(x)\big)$ oraz $\forall y \in Y : y = f\big(f^{-1}(y)\big)$, co oznacza, że

$$f^{-1} \circ f = \mathrm{id}_X\,, \quad f \circ f^{-1} = \mathrm{id}_Y\,.$$

Zachodzi też wynikanie odwrotne, co formułujemy w postaci twierdzenia.

Twierdzenie 6. *Niech $f : X \mapsto Y$. Wtedy następujące warunki są równoważne:*

(i) f jest bijekcją;

(ii) istnieje $g : Y \mapsto X$ takie, że $g \circ f = \mathrm{id}_X$ i $f \circ g = \mathrm{id}_Y$.

Jeśli spełniony jest warunek (ii), to $g = f^{-1}$.

Dowód. Jeśli f jest bijekcją, to $g = f^{-1}$ realizuje warunek (ii), więc (i) $\Rightarrow$ (ii). Jeśli spełniony jest warunek (ii), to poprzednie twierdzenie mówi nam, że f jest surjekcją i injekcją (bo identyczności są bijekcjami), więc jest bijekcją, czyli (ii) $\Rightarrow$ (i). Składając obie strony pierwszego równania warunku (ii) prawostronnie z f^{-1} dostajemy $g = f^{-1}$. □

Ostatnie twierdzenie daje nam gwarancję bijektywności dwu odwzorowań, które po złożeniu w dowolnej kolejności dają identyczności. W następnym twierdzeniu zakładamy podobną własność, ale tylko przy jednym porządku złożenia.

Twierdzenie 7. *Jeśli $f : X \mapsto Y$ i $g : Y \mapsto X$ są takie, że*

$$g \circ f = \mathrm{id}_X \,,$$

to odwzorowania $f' : X \mapsto f(X)\,,\ x \mapsto f(x)$ oraz $g' : f(X) \mapsto X\,,\ y \mapsto g(y)$ są bijekcjami i $g' = f'^{-1}$.

Dowód. Z założenia twierdzenia i z definicji odwzorowań f' i g' wynika, że $g' \circ f' = \mathrm{id}_X$. Twierdzenie 5 mówi nam więc, że f' jest injekcją, a z określenia f' jest też surjekcją, więc jest bijekcją. Składając ostatnią równość prawostronnie z f'^{-1} dostajemy tezę. □

W przypadku, gdy przeciwdziedzina odwzorowania pokrywa się z dziedziną, $f : X \mapsto X$, możemy składać odwzorowanie f z samym sobą. Dla wielokrotnych takich złożeń wprowadzamy notację

$$f^k := \underbrace{f \circ \ldots \circ f}_{k \text{ razy}} \,.$$

Jeśli ponadto f jest bijekcją, to piszemy też

$$f^{-k} := \underbrace{f^{-1} \circ \ldots \circ f^{-1}}_{k \text{ razy}}, \quad f^0 = \mathrm{id}_X \ .$$

Przy tych oznaczeniach dla dowolnych liczb $k, n \in \mathbb{Z}$ mamy

$$f^k \circ f^n = f^{k+n} \,.$$

10 Przykłady

(i) Składanie odwzorowań
Określmy odwzorowania g i h o dziedzinie i przeciwdziedzinie równej zbiorowi liczb całkowitych $\mathbb{Z}$, zadane wzorami $g(x) = |x|$ i $h(x) = -x$. Złożenia w obu porządkach mają jako dziedzinę i przeciwdziedzinę zbiór $\mathbb{Z}$, ale $h \circ g \neq g \circ h$.

(ii) Odwzorowanie odwrotne i całkowita potęga odwzorowania
Dla każdego $m \in \mathbb{N}$ funkcja potęgowa

$$\langle 0, +\infty) \ni x \mapsto P_m(x) = x^m \in \langle 0, +\infty)$$

jest bijekcją. Odwzorowanie odwrotne określa przepis $P_m^{-1}(x) = x^{\frac{1}{m}}$. Dla dowolnej liczby całkowitej k mamy $P_m^k(x) = x^{m^k}$.

11 Zbiory równoliczne

Mówimy, że dwa ***zbiory są równoliczne***, jeśli istnieje bijekcja jednego z nich na drugi. To określenie zadaje relację wśród zbiorów. Identyczność jest bijekcją, więc ta relacja jest zwrotna, odwzorowanie odwrotne do bijekcji jest bijekcją, więc relacja jest symetryczna, oraz złożenie bijekcji jest bijekcją, więc relacja jest przechodnia. Równoliczność jest stąd relacją równoważności, dzieli więc wszystkie zbiory na rozłączne klasy. Mówimy, że zbiory w jednej klasie mają ***równą moc***. Okazuje się (nie będziemy tego dowodzić), że jeśli zbiory mają różną moc, to zawsze istnieje injekcja jednego z nich w drugi, i fakt ten zadaje relację porządku w zbiorze mocy zbiorów. Zbiory skończone o równej liczbie elementów mają tę samą moc (wzrastającą wraz z tą liczbą). Każdy zbiór w klasie równoważności zbioru liczb naturalnych nazywamy ***przeliczalnym nieskończonym***, a moc zbioru liczb rzeczywistych nazywamy mocą ***continuum***. Przeliczalność nieskończonego zbioru X oznacza więc istnienie bijekcji $x : \mathbb{N} \mapsto X\,,\ i \mapsto x_i$. Każdą taką bijekcję nazwiemy ***ponumerowaniem*** (***indeksacją***) elementów zbioru X, który można wtedy zapisać jako: $X = \{x_1, x_2, \ldots\}$. Podobnie w przypadku zbioru o $n \in \mathbb{N}$ elementach mamy $x : \{1, \ldots, n\} \mapsto X\,,\ i \mapsto x_i$ i $X = \{x_1, \ldots, x_n\}$.

§3 Działania, grupa, ciało

1 Działania, struktury algebraiczne

Pojęcie działania jest podstawowe dla wszystkich struktur pojawiających się w algebrze liniowej. Jest to pojęcie bardzo ogólne, ale największe znaczenie mają tzw. działania dwuargumentowe, i do takich ograniczymy się w poniższej definicji działania. Niech będzie dany zbiór X. ***Działaniem wewnętrznym w zbiorze*** X (lub krótko: ***działaniem w*** X, lub ***na*** X) nazywamy każde odwzorowanie $h : X \times X \mapsto X$. Jeśli e jest elementem zbioru X takim, że dla każdego $x \in X$ zachodzi

$$h(e, x) = h(x, e) = x\,,$$

to e nazywamy ***elementem neutralnym*** (***elementem jednostkowym***, ***jedynką***) ***działania*** h. Ile elementów neutralnych może mieć działanie? W ogólności, może nie mieć żadnego, natomiast założenie, że posiada więcej niż jeden, prowadzi do sprzeczności. Rzeczywiście, niech e i e' będą elementami neutralnymi, wtedy wprost z definicji $e' = h(e', e) = e$. Otrzymujemy proste, ale ważne twierdzenie:

Twierdzenie 1. *Działanie wewnętrzne ma co najwyżej jeden element neutralny.*

Jeżeli działanie h ma element neutralny oraz dla danego $x \in X$ istnieje $x' \in X$ taki, że

$$h(x, x') = h(x', x) = e\,,$$

to mówimy, że x posiada ***element odwrotny (względem działania*** h***)*** x'. Wówczas również x' ma element odwrotny równy x. Element neutralny, jeśli istnieje, zawsze posiada element odwrotny równy jemu samemu. W ogólności nie ma natomiast gwarancji istnienia innych elementów odwrotnych, ani ich jednoznaczności dla danego elementu, gdy istnieją. Aby uzyskać jednoznaczność elementów odwrotnych, jeśli istnieją, wyposażymy działanie w następną własność. Mówimy, że działanie h jest ***łączne***, gdy dla każdych trzech elementów x, y, z zbioru X zachodzi $h(h(x, y), z) = h\big(x, h(y, z)\big)$. Dla działań łącznych (a nawet dla niektórych działań pozbawionych tej własności) stosujemy na ogół notację innego typu:

$$\circ : X \times X \mapsto X\,,\ (x, y) \mapsto x \circ y\,,$$

przy której łączność ma naturalny zapis

$$(x \circ y) \circ z = x \circ (y \circ z)\,.$$

Większość działań, z którymi spotkamy się w niniejszej książce, będzie działaniami łącznymi. Dla takich działań mamy następującą własność jednoznaczności.

Twierdzenie 2. *Jeśli działanie w zbiorze X jest łączne i ma element neutralny, to każdy element tego zbioru ma co najwyżej jeden element odwrotny.*

Dowód. Niech $x', x'' \in X$ będą elementami odwrotnymi do $x \in X$. Wtedy z definicji elementu odwrotnego oraz łączności mamy

$$x'' = x'' \circ e = x'' \circ (x \circ x') = (x'' \circ x) \circ x' = e \circ x' = x' ,$$

co należało pokazać. □

Własnością działań niezależną od łączności jest przemienność. Mówimy, że ***działanie** h **jest przemienne***, jeśli dla każdych dwóch elementów x, y zachodzi $h(x, y) = h(y, x)$.

Oprócz działań wewnętrznych istotną rolę w algebrze grają też działania dwuargumentowe, w których jeden z elementów pochodzi z dodatkowego, *pomocniczego* zbioru. Niech F i X będą dwoma zbiorami. Każde odwzorowanie $g : F \times X \mapsto X$ nazywamy ***działaniem zewnętrznym na** X **zadanym elementami zbioru** F*.

Uzbrojeni w dotychczasowe definicje możemy teraz wskazać twory matematyczne, których szczególne przypadki będą obiektami badań w tym podręczniku. ***Strukturą algebraiczną*** nazywamy zbiór, w którym zadano działania wewnętrzne oraz, ewentualnie, działania zewnętrzne z elementami innych zbiorów.

2 Grupa

Grupa jest szczególnym, i niezwykle ważnym, przypadkiem struktury algebraicznej z jednym działaniem wewnętrznym. Niech G będzie zbiorem, a $\circ$ – zadanym w nim działaniem wewnętrznym. Mówimy, że para $(G, \circ)$ tworzy ***grupę***, gdy spełnione są następujące warunki:

(i) działanie $\circ$ jest łączne,

(ii) istnieje element neutralny,

(iii) dla każdego elementu zbioru G istnieje element odwrotny.

Na mocy ogólnego twierdzenia element neutralny jest jedyny w grupie, oznaczamy go e, a na mocy łączności działania element odwrotny do g jest dla niego jedyny, oznaczamy go g^{-1}. Wykorzystując łączność działania wyliczamy $(h^{-1} \circ g^{-1}) \circ (g \circ h) = e$ i podobnie w przeciwnej kolejności. Stąd mamy:

Twierdzenie 3. *Dla każdych dwóch elementów g, h dowolnej grupy zachodzi:*

$$(g \circ h)^{-1} = h^{-1} \circ g^{-1} .$$

Upraszczając mówimy niekiedy, że G jest grupą (jeśli działanie jest znane). Jeśli działanie grupowe jest przemienne, to G nazywamy ***grupą przemienną*** lub ***abelową*** (Niels Henrik Abel – matematyk norweski).

Jeśli w definicji grupy opuścimy założenie (iii), to otrzymamy definicję struktury ogólniejszej, zwanej ***półgrupą z jedynką*** (***monoidem***). Niech G będzie półgrupą z jedynką, a g i h dowolnymi jej elementami posiadającymi elementy odwrotne. Rachunek prowadzący do ostatniego twierdzenia pokazuje, że ich złożenie $g \circ h$ również posiada element odwrotny, $h^{-1} \circ g^{-1}$. Stąd działanie w G zawężone do podzbioru złożonego z elementów posiadających odwrotne działa wewnątrz tego podzbioru. Dostajemy więc:

Twierdzenie 4. *Niech $(G, \circ)$ będzie półgrupą z jedynką. Podzbiór złożony z wszystkich elementów G mających odwrotne stanowi grupę z tym samym działaniem.*

Na koniec tego punktu uwaga o elementach odwrotnych w grupach. Definicja elementu odwrotnego wymaga, aby jego złożenie z danym elementem w każdej z dwu kolejności dawało element neutralny. Jeśli jednak już wiemy, że element odwrotny istnieje (jak to ma miejsce dla każdego elementu grupy), to warunek $g \circ h = e \quad \textit{lub} \quad h \circ g = e$ pociąga $h = g^{-1}$ – wystarczy pierwszą z równości pomnożyć przez g^{-1} z lewej strony, lub drugą z prawej. Można ten wynik sformułować też tak:

Twierdzenie 5. *Jeśli g i h są elementami grupy spełniającymi $g \circ h = e$, to są one wzajemnie odwrotne.*

3 Przykłady

(i) Działanie wewnętrzne
Odwzorowanie $\mathbb{R} \times \mathbb{R} \ni (x, y) \mapsto x^2 + y^2 \in \mathbb{R}$ jest działaniem wewnętrznym w $\mathbb{R}$. To działanie jest przemienne, ale nie jest łączne i nie posiada elementu neutralnego.

(ii) Działanie łączne i przemienne
W zbiorze podzbiorów zbioru X definiujemy działanie $(A, B) \mapsto A \cup B$. To działanie jest przemienne i łączne, a jego elementem neutralnym jest zbiór pusty. Nie istnieją nietrywialne elementy odwrotne (tj. dla zbiorów różnych od pustego).

(iii) Dodawanie w zbiorach liczbowych
Dodawanie liczb jest łącznym i przemiennym działaniem w zbiorach liczb $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ i $\mathbb{R}$. W $\mathbb{N}$ nie ma elementu neutralnego, w pozostałych zbiorach elementem

neutralnym jest 0, a odwrotnym do liczby x jest liczba $-x$. Odejmowanie może być określone za pomocą dodawania, kładziemy z definicji $x - y := x + (-y)$. Zbiory $\mathbb{Z}$, $\mathbb{Q}$ i $\mathbb{R}$ tworzą z dodawaniem przemienne grupy.

(iv) Mnożenie w zbiorach liczbowych
Mnożenie liczb jest łącznym i przemiennym działaniem w zbiorach liczb $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ i $\mathbb{R}$. W każdym z tych zbiorów elementem neutralnym mnożenia jest 1. W zbiorze $\mathbb{N}$ jedynie 1, a w $\mathbb{Z}$ jedynie $+1$ i -1 mają elementy odwrotne, w zbiorach $\mathbb{Q}$ i $\mathbb{R}$ – wszystkie elementy z wyjątkiem zera. Zbiory $\mathbb{Q}^*$ i $\mathbb{R}^*$ tworzą z mnożeniem przemienne grupy. Dzielenie jest w nich określone jako mnożenie przez element odwrotny.

(v) Różnica symetryczna zbiorów
W zbiorze podzbiorów zbioru X określamy różnicę symetryczną za pomocą przepisu $(A, B) \mapsto A \triangle B := (A \setminus B) \cup (B \setminus A)$. Otrzymujemy w wyniku grupę abelową, z elementem neutralnym równym zbiorowi pustemu, w której elementem odwrotnym do każdego zbioru A jest ten sam zbiór A.

(vi) Rząd grupy
Jeśli grupa ma skończoną liczbę elementów, to tę liczbę nazywamy ***rzędem grupy***. Jedynymi grupami skończonego rzędu utworzonymi z liczb rzeczywistych, z działaniem mnożenia liczbowego, są: $\{1\}$, $\{1, -1\}$.

(vii) Grupy cykliczne
Niech w zbiorze X zadane będzie działanie łączne $\circ$, mające element neutralny e (a więc $(X, \circ)$ jest półgrupą z jedynką). Oznaczmy dla $k \in \mathbb{N}$:

$$g^k := \underbrace{g \circ \ldots \circ g}_{k \text{ razy}} .$$

Jeśli g jest takim elementem zbioru X, że $g^n = e$, i n jest najmniejszą liczbą, dla której zachodzi ta równość, to zbiór

$$\{e, g, g^2, \ldots, g^{n-1}\}$$

jest grupą rzędu n z działaniem $\circ$ (gdyby $g^k = g^m$ dla $k < m < n$, to mnożąc obie strony przez g^{n-k} uzyskalibyśmy $e = g^{m-k}$, co jest sprzeczne z założeniem). Każdą grupę o tej postaci nazywamy ***grupą cykliczną rzędu*** n. Każda grupa cykliczna jest abelowa.

Mówimy, że element g' generuje grupę cykliczną $\{e, g, g^2, \ldots, g^{n-1}\}$, jeśli

$$\{e, g', g'^2, \ldots, g'^{n-1}\} = \{e, g, g^2, \ldots, g^{n-1}\} .$$

Niech m będzie taką liczbą naturalną, że $\dfrac{m}{n} = \dfrac{m_0}{n_0}$ dla pewnych liczb naturalnych $m_0 < m$ i $n_0 < n$. Wtedy $(g^m)^{n_0} = (g^n)^{m_0} = e$, więc g^m nie generuje powyższej grupy cyklicznej. Można pokazać, że zachodzi też implikacja odwrotna, i dostaje się następujące twierdzenie.

> Element g^m, $m < n$, generuje grupę $\{e, g, \ldots, g^{n-1}\}$ wtedy, i tylko wtedy, gdy liczby m i n są względem siebie pierwsze (tj. nie mają wspólnych podzielników poza 1).

W dowodzie brakującej implikacji korzysta się z następującej własności liczb względnie pierwszych m i n: istnieją liczby całkowite p i q takie, że $pm + qn = 1$.

(viii) Grupa trzeciego rzędu
Niech G będzie grupą trzeciego rzędu, o elementach e, a, b. Gdyby $a \circ b = a$ lub b, to $b = e$ lub $a = e$, co jest wykluczone. Stąd $b = a^{-1}$, więc $a^2 \neq e$. Gdyby $a^2 = a$, to $a = e$, wbrew założeniu. Stąd $a^{-1} = a^2$, czyli $a^3 = e$. Każda grupa trzeciego rzędu jest cykliczna.

(ix) Iloczyn prosty
Niech $(G, \circ)$ i $(G', \circ')$ będą dwoma grupami. Na iloczynie kartezjańskim $G \times G'$ wprowadzamy działanie $\times$ za pomocą przepisu

$$(g_1, g_1') \times (g_2, g_2') := (g_1 \circ g_2 \,,\, g_1' \circ' g_2') \,.$$

Z tak określonym działaniem zbiór $G \times G'$ stanowi grupę; nazywamy ją ***iloczynem prostym grup G i G'*** i oznaczamy $G \times G'$.

4 Podgrupy

Niech będzie dana grupa $(G, \circ)$ oraz podzbiór H zbioru G. Mówimy, że H jest ***podgrupą*** grupy G, jeśli $(H, \circ)$ jest grupą, tj. H jest grupą względem działania dziedziczonego od G. Niech w tym przypadku e będzie elementem neutralnym grupy G, a e' elementem neutralnym grupy H. Stąd, w szczególności, $e' \circ e' = e'$. Składając obie strony tego równania z elementem odwrotnym do e' w grupie G dostajemy $e' = e$. Element neutralny podgrupy jest więc dziedziczony od całej grupy. Stąd również element odwrotny do $h \in H$ w grupie H jest identyczny z odwrotnym do h jako elementu grupy G. Łatwo teraz wykazać prawdziwość następującego kryterium.

Twierdzenie 6. *Jeśli G jest grupą, to niepusty zbiór $H \subseteq G$ jest jej podgrupą wtedy, i tylko wtedy, gdy spełnione są łącznie następujące warunki:*

(i) $\forall h_1, h_2 \in H : h_1 \circ h_2 \in H$,

(ii) $\forall h \in H : h^{-1} \in H$.

Dowód. Konieczność kryterium oczywista w świetle definicji podgrupy i dyskusji poprzedzającej twierdzenie. Pokazujemy dostateczność. Punkt (i) mówi, że $\circ$ jest działaniem w zbiorze H. Łączność tego działania wynika z jego łączności na G. Jeśli h jest dowolnym elementem zbioru H (który nie jest pusty), to łącznie z obu punktów wynika, że także $h \circ h^{-1} = e$ jest elementem H. Istnieje więc element neutralny w H, a drugi punkt mówi wtedy o istnieniu odwrotnych w H. □

Użycie tego kryterium pozwala natychmiast przekonać się o słuszności następującego stwierdzenia (szczegóły pozostawiamy czytelnikowi).

Twierdzenie 7. *Jeśli H_1 i H_2 są podgrupami grupy G, to $H_1 \cap H_2$ jest też podgrupą G.*

Każda podgrupa H grupy G zadaje w niej dwie relacje binarne, $\mathcal{R}^l_H$ i $\mathcal{R}^r_H$. Określamy je następująco:

$$g_1 \, \mathcal{R}^l_H \, g_2 \iff {g_1}^{-1} \circ g_2 \in H \,, \quad g_1 \, \mathcal{R}^r_H \, g_2 \iff g_1 \circ {g_2}^{-1} \in H \,.$$

Twierdzenie 8. *$\mathcal{R}^l_H$ i $\mathcal{R}^r_H$ są relacjami równoważności.*

Dowód. Prowadzimy dowód dla pierwszej z relacji, dla drugiej przebiega analogicznie. Mamy: $g^{-1} \circ g = e \in H$, więc relacja jest zwrotna; jeśli ${g_1}^{-1} \circ g_2 \in H$, to ${g_2}^{-1} \circ g_1 = ({g_1}^{-1} \circ g_2)^{-1} \in H$, więc jest symetryczna; jeśli ponadto ${g_2}^{-1} \circ g_3 \in H$, to ${g_1}^{-1} \circ g_3 = ({g_1}^{-1} \circ g_2) \circ ({g_2}^{-1} \circ g_3) \in H$, więc jest przechodnia. □

Warunki definicyjne wprowadzonych relacji można zapisać równoważnie jako

$$g \, \mathcal{R}^l_H \, g' \iff \exists h \in H : g' = g \circ h \,, \quad g \, \mathcal{R}^r_H \, g' \iff \exists h \in H : g' = h \circ g \,.$$

Dla dowolnych elementów $g, g' \in G$ oznaczamy

$$gHg' := \{ g \circ h \circ g' \mid h \in H \} \,.$$

Klasą równoważności dowolnego elementu $g \in G$ względem relacji $\mathcal{R}^l_H$ jest więc zbiór gH; każdy taki zbiór nazywamy ***warstwą lewostronną grupy** G **względem podgrupy** H*, a zbiór wszystkich warstw lewostronnych oznaczamy G/H. Podobnie: klasą równoważności elementu g względem relacji $\mathcal{R}^r_H$ jest zbiór Hg; każdy taki zbiór nazywamy ***warstwą prawostronną grupy** G **względem podgrupy** H*, a zbiór wszystkich warstw prawostronnych oznaczamy $H \backslash G$.

5 Podgrupy niezmiennicze. Grupy ilorazowe

Warstwy lewo- i prawostronne na ogół nie pokrywają się, jednak te szczególne przypadki, gdy są one identyczne, prowadzą do dalszych interesujących struktur. Wprowadzamy dlatego następującą definicję. Mówimy, że podgrupa $N \subseteq G$ jest ***podgrupą niezmienniczą grupy*** G (lub: ***dzielnikiem normalnym grupy*** G), gdy dla każdego $g \in G$ zachodzi $gN = Ng$. Na mocy twierdzenia 2, §2, wynika stąd, że relacje $\mathcal{R}^l_N$ i $\mathcal{R}^r_N$ pokrywają się; oznaczymy wtedy tę relację $\mathcal{R}_N$.

Ogólnie, jeśli H jest dowolną podgrupą grupy G, to łatwo przekonać się, że dla dowolnego elementu $g \in G$ zbiór gHg^{-1} również stanowi podgrupę. Możemy więc podać równoważne sformułowanie warunku definicyjnego podgrupy niezmienniczej N: dla każdego $g \in G$ podgrupa gNg^{-1} jest równa N. Przyjmijmy pozornie słabszy warunek

$$\forall g \in G : gNg^{-1} \subseteq N\,.$$

Zastępując w inkluzji zbiorów g przez g^{-1} dostajemy stąd również $\forall g \in G$: $g^{-1}Ng \subseteq N$. Mnożąc ostatnią inkluzję przez g z lewej i g^{-1} z prawej dostajemy $\forall g \in G : N \subseteq gNg^{-1}$. Łącznie: $\forall g \in G : gNg^{-1} = N$. Wykazaliśmy następujące kryterium.

Twierdzenie 9. *Niech N będzie podgrupą grupy G. N jest podgrupą niezmienniczą wtedy, i tylko wtedy, gdy dla każdego $g \in G$ zachodzi $gNg^{-1} \subseteq N$.*

Niech grupa N będzie podgrupą niezmienniczą grupy G. W zbiorze warstw G/N chcemy wprowadzić działanie określone następującym przepisem

$$g_1N \circ g_2N = (g_1 \circ g_2)N\,,$$

gdzie po prawej stronie znak $\circ$ jest działaniem w grupie G, a po lewej ten sam znak określa wprowadzane działanie w zbiorze G/N. Określenie powyższe ma na razie charakter prowizoryczny – musimy sprawdzić jego poprawność. Przypuśćmy, że w warstwach g_1N i g_2N wybrano inne elementy g_1' i g_2' odpowiednio, i użyto ich zamiast g_1 i g_2 w powyższym przepisie. Należy pokazać, że wtedy również $g_1' \circ g_2' \in (g_1 \circ g_2)N$. Z przyjętych założeń wynika, że $g_i' = g_i \circ n_i$, $n_i \in N$ $(i = 1, 2)$. Stąd

$$g_1' \circ g_2' = g_1 \circ g_2 \circ ({g_2}^{-1} \circ n_1 \circ g_2) \circ n_2\,.$$

Na podstawie warunku definicyjnego dzielnika normalnego mamy

$${g_2}^{-1} \circ n_1 \circ g_2 \in N\,, \quad \text{więc} \quad g_1' \circ g_2' = g_1 \circ g_2 \circ n_{12}\,, \quad n_{12} \in N\,,$$

co dowodzi poprawności definicji. Zwróćmy uwagę, że założenie o niezmienniczości podgrupy N odgrywa w dowodzie poprawności istotną rolę, dla innych podgrup przepis powyższy nie jest konsystentny.

Wprowadzone powyższym przepisem działanie jest łączne (wynika to natychmiast z łączności działania w grupie G), ma element neutralny eN oraz elementy odwrotne $(gN)^{-1} = g^{-1}N$. Para $(G/N, \circ)$ stanowi więc grupę, nazywaną ***grupą ilorazową***.

Jeśli grupa G jest grupą przemienną, to każda jej podgrupa jest dzielnikiem normalnym, więc każda z nich zadaje podgrupę ilorazową, która również jest przemienna.

W dalszym ciągu wykładu stosować będziemy uproszczoną notację działania grupowego, przy której znak $\circ$ po prostu pomijamy: $gg' \equiv g \circ g'$. Dla pewnych grup abelowych stosuje się notację addytywną przejętą od dodawania liczb: $g + g' \equiv g \circ g'$.

6 Przykłady

(i) Moc warstw
Wszystkie warstwy grupy względem jej wybranej podgrupy mają tę samą moc.

(ii) Twierdzenie Lagrange'a
Jeśli G jest skończonego rzędu n, a H jest jej podgrupą rzędu k, to n dzieli się przez k. Jeśli ponadto H jest dzielnikiem normalnym, to rząd grupy G/H jest równy n/k. W dowodzie korzysta się z poprzedniego przykładu oraz własności podziału zbioru na klasy równoważności.

(iii) Podgrupa generowana elementem grupy
Każdy ***element*** g ***grupy*** G ***generuje jej podgrupę***: cykliczną podgrupę

$$\{e, g, g^2, \dots, g^{k-1}\},$$

jeśli istnieje takie (minimalne) $k \in \mathbb{N}$, że $g^k = e$, lub podgrupę

$$\{\dots, g^{-2}, g^{-1}, e, g, g^2, \dots\}$$

(nazywaną cykliczną grupą nieskończonego rzędu), w przeciwnym przypadku.

(iv) Podgrupa generowana podzbiorem grupy
Niech Y będzie dowolnym podzbiorem grupy G. Tworzymy zbiór H złożony ze wszystkich skończonych iloczynów postaci $y_1 y_2 \dots$, gdzie każdy z czynników w iloczynie lub jego odwrotny jest elementem zbioru Y. Łatwo wykazać, że H jest podgrupą grupy G. Mówimy, że ***zbiór*** Y ***generuje*** H. Każda podgrupa grupy G zawierająca zbiór Y zawiera również H. Gdy $Y = \{g\}$, to Y generuje podgrupę cykliczną (poprzedni przykład).

(v) Grupa o rzędzie danym liczbą pierwszą
Jeśli rząd G jest liczbą pierwszą, to G jest grupą cykliczną i każdy element różny od e jest jej generatorem. (Konsekwencja poprzednich przykładów.)

(vi) Podgrupa $\mathbb{R}^+ \subseteq \mathbb{R}^*$
$\mathbb{R}^+$ jest podgrupą multiplikatywnej grupy $\mathbb{R}^*$, a $\mathbb{R}^-$ nie jest grupą (podobnie dla $\mathbb{Q}$).

(vii) Grupa ilorazowa $\mathbb{R}^*/\mathbb{R}^+$
$\mathbb{R}^*/\mathbb{R}^+ = \{\mathbb{R}^+, \mathbb{R}^-\}$ tworzy przemienną grupę ilorazową, z tablicą mnożenia: $\mathbb{R}^+ \cdot \mathbb{R}^+ = \mathbb{R}^- \cdot \mathbb{R}^- = \mathbb{R}^+$, $\mathbb{R}^+ \cdot \mathbb{R}^- = \mathbb{R}^-$.

(viii) Grupy ilorazowe $\mathbb{Z}_m$
Dla każdej liczby naturalnej $m \geq 2$ zbiór liczb

$$m\mathbb{Z} := \{mk \mid k \in \mathbb{Z}\} \subseteq \mathbb{Z}$$

jest podgrupą addytywnej grupy $\mathbb{Z}$. Relacja $\mathcal{R}_{m\mathbb{Z}}$ pokrywa się z relacją równości modulo m. Stąd grupa ilorazowa

$$\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$$

jako zbiór pokrywa się z $\mathbb{Z}/\mathrm{mod}\, m$ (patrz przykład (iv), p. 3, §2), a działanie jest dane przez $[k] + [l] = [k + l]$. Grupa $\mathbb{Z}_m$ jest cykliczną grupą rzędu m.

(ix) Grupy ilorazowe $\mathbb{R}_c$
Dla każdej liczby $c \in \mathbb{R}^+$ zbiór liczb

$$c\mathbb{Z} := \{ck \mid k \in \mathbb{Z}\} \subseteq \mathbb{R}$$

jest podgrupą addytywnej grupy $\mathbb{R}$. Relacja $\mathcal{R}_{c\mathbb{Z}}$ pokrywa się z relacją równości modulo c. Grupa ilorazowa

$$\mathbb{R}_c := \mathbb{R}/c\mathbb{Z}$$

jako zbiór pokrywa się z $\mathbb{R}/\mathrm{mod}\, c$ (patrz przykład (v), p. 3, §2), a działanie jest dane przez $[x] + [y] = [x + y]$.

(x) Równość relacji $\mathcal{R}^l_H$ i $\mathcal{R}^r_H$
Jeśli dla pewnego $g \in G$ istnieje $g' \in G$ taki, że $gH = Hg'$, to $gH = Hg$. Stąd, jeśli $\mathcal{R}^l_H = \mathcal{R}^r_H$, to H jest dzielnikiem normalnym.

7 Homomorfizmy

Niech będą dane dwie grupy G i G'; ich elementy neutralne oznaczmy e i e' odpowiednio. Mówimy, że odwzorowanie $h : G \mapsto G'$ jest ***homomorfizmem (grupy G w grupę G')*** gdy dla każdych dwóch elementów g_1, g_2 grupy G zachodzi

$$h(g_1 g_2) = h(g_1)h(g_2)\,.$$

Zwróćmy uwagę, że mnożenie elementów g_i wykorzystuje strukturę grupy G, a mnożenie elementów $h(g_i)$ – strukturę grupy G'. Fakt istnienia homomorfizmu mówi więc o pokrewieństwie tych struktur w dwóch grupach. Zauważmy ponadto, że z definicji wynika w szczególności, że $h(e)h(e) = h(e)$. Mnożąc to równanie obustronnie przez $\big(h(e)\big)^{-1}$ dostajemy

$$h(e) = e'\,.$$

Wybierając teraz w definicji homomorfizmu elementy g_1 i g_2 jako wzajemnie odwrotne dostajemy również

$$\big(h(g)\big)^{-1} = h(g^{-1})$$

dla każdego $g \in G$.

Niektóre szczególne przypadki homomorfizmów mają tak wielkie znaczenie, że zasługują na specjalne nazwy. Jeśli h jest homomorfizmem grupy w samą siebie (czyli $G' = G$), to nazywamy go ***endomorfizmem***. Bijektywny homomorfizm między dowolnymi grupami nazywamy ***izomorfizmem***. Bijektywny endomorfizm (czyli izomorfizm grupy na samą siebie) nazywamy ***automorfizmem***. Dla każdego homomorfizmu h obraz jego dziedziny nazywamy ***obrazem homomorfizmu*** i oznaczamy

$$\operatorname{Im} h := h(G) \subseteq G'\,.$$

Definiujemy ponadto ***jądro homomorfizmu*** przez:

$$\operatorname{Ker} h := h^{-1}(\{e'\}) \subseteq G\,.$$

Własności homomorfizmów są ściśle związane z własnościami grupowymi przestrzeni, między którymi działają. Ważną ilustracją tego stwierdzenia są dwa następne twierdzenia.

Twierdzenie 10. *Jeśli $h : G \mapsto G'$ jest homomorfizmem, to $\operatorname{Im} h$ jest podgrupą grupy G', a $\operatorname{Ker} h$ jest podgrupą niezmienniczą grupy G.*

Dowód. Kryterium własności bycia podgrupą pokazuje, że $\operatorname{Im} h$ i $\operatorname{Ker} h$ są podgrupami odpowiednich grup; szczegóły pozostawiamy czytelnikowi. Jeśli $g \in G$ i $g_0 \in \operatorname{Ker} h$, to $h(gg_0g^{-1}) = h(g)e'h(g^{-1}) = e'$, więc $g(\operatorname{Ker} h)g^{-1} \subseteq \operatorname{Ker} h$. Kryterium niezmienniczości jest tym samym spełnione dla $\operatorname{Ker} h$. □

Twierdzenie 11. *Niech $h : G \mapsto G'$ będzie homomorfizmem. Odwzorowanie h jest różnowartościowe wtedy, i tylko wtedy, gdy* $\operatorname{Ker} h = \{e\}$.

Dowód. $\Rightarrow$ (nie wprost) Jeśli $g \in \operatorname{Ker} h$ i $g \neq e$, to $h(g) = e' = h(e)$, więc h nie jest różnowartościowe.
$\Leftarrow$ (nie wprost) Jeśli $g_1 \neq g_2$ i $h(g_1) = h(g_2)$, to ${g_1}^{-1}g_2 \neq e$ oraz

$$h({g_1}^{-1}g_2) = (h(g_1))^{-1}h(g_2) = e' \,, \text{ więc } \operatorname{Ker} h \neq \{e\} \,.$$

□

8 Grupy izomorficzne

Izomorfizmy pomiędzy grupami tworzą strukturę określoną następującym twierdzeniem.

Twierdzenie 12. *Jeśli $h : G \mapsto G'$ oraz $h' : G' \mapsto G''$ są izomorfizmami grup, to $h' \circ h$ oraz h^{-1} są też izomorfizmami.*

Dowód. Jeśli h i h' są izomorfizmami, to są bijekcjami, więc $h' \circ h$ i h^{-1} są bijekcjami. Dla dowolnych $g_1, g_2 \in G$ mamy

$$h'\big(h(g_1g_2)\big) = h'\big(h(g_1)h(g_2)\big) = h'\big(h(g_1)\big)h'\big(h(g_2)\big) \,,$$

więc $h' \circ h$ jest homomorfizmem, a zatem izomorfizmem. Kładąc $g_i = h^{-1}(g'_i)$ $(i = 1, 2)$ w warunku definicyjnym homomorfizmu h dostajemy

$$h\big(h^{-1}(g'_1)h^{-1}(g'_2)\big) = g'_1g'_2 \,, \text{ czyli } h^{-1}(g'_1)h^{-1}(g'_2) = h^{-1}(g'_1g'_2) \,,$$

więc h^{-1} również jest izomorfizmem. □

Grupy, pomiędzy którymi istnieje izomorfizm, nazywamy ***izomorficznymi***. Twierdzenie powyższe pokazuje, że relacja izomorficzności jest relacją równoważności. Grupy będące w jednej klasie równoważności są bardzo ściśle ze sobą związane – ich struktury grupowe są identyczne.

Homomorfizmy i izomorfizmy wprowadza się również dla innych struktur algebraicznych w sposób analogiczny do teorii grup. Analogicznie też określa się struktury izomorficzne, które w swoich własnościach algebraicznych są identyczne.

9 Przykłady

(i) Projekcja $G \mapsto G/N$
Niech N będzie podgrupą niezmienniczą grupy G. Określmy odwzorowanie $\pi : G \mapsto G/N$, $\pi(g) = gN$. Sposób określenia działania w grupie ilorazowej G/N pokazuje, że π jest homomorfizmem. Jego obraz jest równy G/N, a jądro pokrywa się z N.

(ii) Przykład projekcji
Kładziemy w poprzednim przykładzie $G = \mathbb{R}$ i $N = c\mathbb{Z}$ (patrz przykł. (ix), p. 6). Wtedy $\operatorname{Im}\pi = \mathbb{R}_c$, $\operatorname{Ker}\pi = c\mathbb{Z}$.

(iii) Rozkład homomorfizmu
Niech $h : G \mapsto H$ będzie homomorfizmem grupowym. Wtedy $\operatorname{Ker} h$ jest niezmienniczą podgrupą grupy G, a $\operatorname{Im} h$ jest podgrupą grupy H. Określamy odwzorowanie

$$\tilde{h} : G/\operatorname{Ker} h \mapsto \operatorname{Im} h\,, \quad \tilde{h}(g \operatorname{Ker} h) = h(g)\,.$$

Mamy $h(g') = h(g)$ wtedy, i tylko wtedy, gdy $g' = gk$, $k \in \operatorname{Ker} h$, co jest równoważne $g' \operatorname{Ker} h = g \operatorname{Ker} h$. To pokazuje, że definicja $\tilde{h}$ jest poprawna, oraz że to odwzorowanie jest injektywne. Jest ono również surjektywne, więc jest bijekcją. Ponadto

$$\begin{aligned}\tilde{h}(g \operatorname{Ker} h \circ g' \operatorname{Ker} h) &= \tilde{h}(gg' \operatorname{Ker} h)\\ &= h(gg') = h(g)h(g') = \tilde{h}(g \operatorname{Ker} h)\tilde{h}(g' \operatorname{Ker} h)\,,\end{aligned}$$

więc $\tilde{h}$ jest izomorfizmem.

Homomorfizm h można teraz rozłożyć w następujący sposób:

$$h = j \circ \tilde{h} \circ \pi\,,$$

gdzie $\pi : G \mapsto G/\operatorname{Ker} h$ jest projekcją omówioną w przykładzie (i) (z podstawieniem $N = \operatorname{Ker} h$), a $j : \operatorname{Im} h \mapsto H$ jest naturalnym włożeniem $j(k) = k$.

(iv) Izomorfizmy grup $\mathbb{R}_c$
Oznaczmy elementy grupy $\mathbb{R}_c$ przez $[x]_c$. Odwzorowanie

$$j : \mathbb{R}_c \mapsto \mathbb{R}_{c'}\,, \quad j([x]_c) = \left[\frac{c'x}{c}\right]_{c'}$$

jest poprawnie określone (nie zależy od wyboru reprezentanta x klasy $[x]_c$), i jest bijektywne. Z definicji dodawania w grupie $\mathbb{R}_{c'}$ mamy

$$\left[\frac{c'(x+y)}{c}\right]_{c'} = \left[\frac{c'x}{c}\right]_{c'} + \left[\frac{c'y}{c}\right]_{c'}\,.$$

Biorąc pod uwagę, że również $[x+y]_c = [x]_c + [y]_c$, dostajemy stąd

$$j([x]_c + [y]_c) = j([x]_c) + j([y]_c)\,,$$

czyli stwierdzenie homomorficzności j (z addytywną notacją działań). Wszystkie grupy $\mathbb{R}_c$ są więc izomorficzne.

(v) Grupy różnych rzędów
Jeśli dwie grupy mają różne rzędy, to nie są izomorficzne. W szczególności, grupy $\mathbb{Z}_m$ i $\mathbb{Z}_n$ nie są izomorficzne dla $m \neq n$.

(vi) Grupy cykliczne
Każda grupa cykliczna rzędu m jest izomorficzna z grupą $\mathbb{Z}_m$.

10 Ciało

Własności następnej struktury algebraicznej modelowane są na własnościach zbioru liczb wymiernych lub zbioru liczb rzeczywistych. Niech będzie dany zbiór $\mathbb{K}$ oraz określone na nim dwa działania:

$$+ : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K}\,,\ (x,y) \mapsto x+y\,, \qquad \cdot : \mathbb{K} \times \mathbb{K} \mapsto \mathbb{K}\,,\ (x,y) \mapsto xy\,,$$

nazwane odpowiednio ***dodawaniem*** oraz ***mnożeniem***. Strukturę $(\mathbb{K}, +, \cdot)$ nazywamy ***ciałem*** gdy spełnione są następujące warunki:

(i) $(\mathbb{K}, +)$ jest grupą przemienną; jej element neutralny oznaczamy 0, a odwrotny do x względem dodawania oznaczamy $-x$;

(ii) mnożenie jest łączne, przemienne, ma element neutralny 1 różny od 0, oraz dla każdego $x \in \mathbb{K}^* \equiv \mathbb{K} \setminus \{0\}$ istnieje element odwrotny x^{-1};

(iii) mnożenie jest rozdzielne względem dodawania, tj. dla każdych $x, y, z \in \mathbb{K}$ zachodzi
$$(x+y)z = xz + yz\,.$$

Z definicji wynika, że $\mathbb{K}^*$ tworzy grupę z mnożeniem, więc element neutralny oraz elementy odwrotne są jednoznacznie zadane. Z własności grupowych mamy też

$$-(x+y) = (-x) + (-y)\,; \qquad \text{dla} \quad x, y \neq 0:\ (xy)^{-1} = y^{-1}x^{-1}\,.$$

Bezpośrednio z definicji otrzymujemy też dalsze proste, lecz podstawowe własności ciała, znane nam ze zbioru liczb rzeczywistych.

Twierdzenie 13. *W każdym ciele prawdziwe są następujące własności:*

(iv) dla każdych x i y zachodzi

$$0x = x0 = 0\,, \quad (-x)y = -(xy) = x(-y)\,;$$

(v) jeśli $xy = 0$, to $x = 0$ lub $y = 0$;

Dowód. (iv) $0 + 0 = 0$, więc z rozdzielności $0x + 0x = 0x$. Dodając obustronnie $-(0x)$ dostajemy $0x = 0$, a z przemienności również $x0 = 0$. Z rozdzielności mamy również $xy + (-x)y = (x + (-x))y = 0y = 0$, więc $-(xy) = (-x)y$. Podobnie ostatnia równość.

(v) Dowód nie wprost. Jeśli $x \neq 0$ i $y \neq 0$, to mnożąc równanie $xy = 0$ przez elementy odwrotne dostajemy $1 = 0$, co jest wykluczone. □

W dalszym toku wykładu stosować będziemy, podobnie jak dla grup, uproszczone oznaczenia ciał, w których działania będą pomijane: $(\mathbb{K}, +, \cdot) \equiv \mathbb{K}$. Wprowadzamy również znaki odejmowania i dzielenia jako uproszczoną notację dodawania i mnożenia przez odpowiednie elementy odwrotne działań:

$$x - y := x + (-y)\,, \quad \frac{x}{y} := xy^{-1}\,.$$

Elementy odwrotne względem dodawania nazywa się ***elementami przeciwnymi***.

Dwa kryteria klasyfikacyjne wszystkich ciał będą miały znaczenie w dalszym toku wykładu. Po pierwsze, ciało może być zbiorem skończonym lub nieskończonym (przykłady obu przypadków w następnym punkcie). Po drugie, rozważmy ciało $\mathbb{K}$ ze względu na jego strukturę grupy addytywnej. Element jednostkowy mnożenia 1 generuje ciąg elementów

$$1,\ 1+1,\ 1+1+1,\ \ldots\,.$$

Zastosowanie wyniku przykładu (iii), p. 6, pokazuje, że zachodzą dwie możliwości: albo (i) ciąg ten jest nieskończony i żaden element w ciągu nie powtarza się, albo (ii) istnieje taka minimalna liczba naturalna k, że

$$\underbrace{1 + \ldots + 1}_{k \text{ razy}} = 0\,.$$

W pierwszym przypadku mówimy, że ciało ma ***nieskończoną charakterystykę*** i piszemy $\operatorname{char}\mathbb{K} = \infty$, a w drugim – ***charakterystykę równą*** k, i piszemy

$\operatorname{char}\mathbb{K} = k$. Charakterystyka nieskończona nazywana też bywa ***charakterystyką zerową***. Każde ciało skończone ma skończoną (niezerową) charakterystykę. Z prawa rozdzielności mamy $x + \ldots + x = x(1 + \ldots + 1)$, więc w ciele o charakterystyce równej k nie tylko 1, ale również każdy element x spełnia

$$\underbrace{x + \ldots + x}_{k \text{ razy}} = 0\,.$$

W dalszych rozważaniach będziemy wielokrotnie spotykać się z warunkiem

$$x = -x\,.$$

Jeśli ciało ma charakterystykę równą 2, to każdy jego element spełnia to równanie. Jeśli charakterystyka ciała jest różna od dwóch, to jedynym rozwiązaniem jest $x = 0$ (gdyby $x \neq 0$ spełniało to równanie, to dzieląc przez x dostalibyśmy $1 = -1$, a więc ciało miałoby charakterystykę równą 2). Ta podstawowa różnica będzie źródłem pojawiającego się często w dalszym ciągu zastrzeżenia $\operatorname{char}\mathbb{K} \neq 2$.

Niezależnie od wartości charakterystyki ciała stosujemy oznaczenia

$$\underbrace{1 + \ldots + 1}_{k \text{ razy}} = k1 \equiv k\,, \quad \underbrace{(-1) + \ldots + (-1)}_{k \text{ razy}} = -k1 \equiv -k\,,$$

gdzie skrajne prawe strony oznaczają elementy ciała $\mathbb{K}$ (kontekst zwykle wskazuje, o którym znaczeniu symbolu k w danym przypadku jest mowa). Stąd również dla dowolnego elementu $x \in \mathbb{K}$ mamy

$$\underbrace{(1 + \ldots + 1)}_{k \text{ razy}}\, x = kx\,.$$

Jeśli więc ciało ma nieskończoną charakterystykę, to zawiera jako addytywną podgrupę (a także jako przemienny pierścień z jedynką – patrz przykład (iv) poniżej) zbiór liczb całkowitych. Jeśli natomiast charakterystyka ciała jest równa $k \in \mathbb{N}$, to przy powyższym oznaczeniu $k = 0$. Łatwo zauważyć, że ten przypadek może zajść tylko gdy k jest liczbą pierwszą. Istotnie, gdyby $k = mn = 0$, $m, n \in \mathbb{N} \setminus \{1\}$, to z własności ciała musiałoby zachodzić $m = 0$ lub $n = 0$, więc charakterystyka ciała byłaby mniejsza od k, wbrew założeniu.

W niniejszym podręczniku jesteśmy zainteresowani głównie ciałami liczb rzeczywistych i zespolonych. Dla poprawności wywodów zamieszczamy tam, gdzie to jest konieczne, uwagi i zastrzeżenia dotyczące wyboru ciała, ale jeśli czytelnik chce ograniczyć uwagę do przypadków $\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$, to te dodatkowe ograniczenia może zignorować.

11 Przykłady

(i) Liczby wymierne i rzeczywiste
Zbiory liczb wymiernych $\mathbb{Q}$ oraz rzeczywistych $\mathbb{R}$ są ciałami o nieskończonej charakterystyce. Zbiór liczb całkowitych nie stanowi ciała.

(ii) Ciało dwuelementowe
Zbiór $\{0,1\}$, z działaniami zadanymi w ten sposób, że 0 jest elementem neutralnym dodawania, 1 – elementem neutralnym mnożenia, oraz $0 \cdot 0 = 0$ i $1+1=0$, jest ciałem. Każde ciało dwuelementowe jest izomorficzne z tym ciałem, i ma charakterystykę równą 2.

(iii) Ciało o nieskończonej charakterystyce
Niech d będzie taką liczbą wymierną, że $\sqrt{d}$ jest liczbą niewymierną. Zbiór liczb rzeczywistych postaci $a + b\sqrt{d}$, $a, b \in \mathbb{Q}$, jest ciałem o nieskończonej charakterystyce.

(iv) Pierścienie i ciała $(\mathbb{Z}/\mathrm{mod}\, m, +, \cdot)$
Zbiór $\mathbb{Z}/\mathrm{mod}\, m$ z działaniem dodawania tworzy grupę, którą oznaczyliśmy $\mathbb{Z}_m$ (przykład (viii), p. 6). W tym samym zbiorze wprowadzamy teraz mnożenie za pomocą przepisu

$$[k]\,[l] = [kl]\,;$$

jeśli $k' = k + rm$, $l' = l + sm$, to $[k'l'] = [kl + (rsm + ks + lr)m] = [kl]$, więc działanie to jest poprawnie określone. Łatwo sprawdzić, że otrzymana struktura $(\mathbb{Z}/\mathrm{mod}\, m, +, \cdot)$ ma następujące własności (pierwszy punkt powtarza fakt grupowości $\mathbb{Z}_m$):

(a) zbiór $\mathbb{Z}/\mathrm{mod}\, m$ tworzy grupę z dodawaniem;

(b) zbiór $\mathbb{Z}/\mathrm{mod}\, m$ tworzy półgrupę z mnożeniem, z elementem jednostkowym $[1]$, przy czym $[0] \neq [1]$;

(c) obowiązuje prawo rozdzielności

$$(\,[j] + [k]\,)\,[l] = [j]\,[l] + [k]\,[l]\,.$$

Każdą strukturę spełniającą warunki (a) – (c) (z przemiennymi działaniami) nazywamy ***przemiennym pierścieniem z jedynką*** (określenie *przemienny* odnosi się tutaj do mnożenia, które dla ogólnych pierścieni może nie posiadać tej własności). Pierścień taki staje się ciałem, jeśli dodatkowo każdy element różny od zerowego ma element odwrotny względem mnożenia.

Pierścienie $(\mathbb{Z}/\mathrm{mod}\, m, +, \cdot)$ na ogół nie są ciałami: jeśli m ma rozkład $m = pq$, to $[p][q] = [m] = [0]$, więc wbrew twierdzeniu prawdziwemu dla ciał istnieje nietrywialny multiplikatywny rozkład zera. Niech jednak m będzie liczbą pierwszą i $1 < k < m$. Ponieważ zbiór $\mathbb{Z}/\mathrm{mod}\, m$ jest skończony, to w ciągu $[k], [k^2], \ldots$ elementy muszą się powtarzać, i niech np.

$$[k^s] = [k^r]\,, \ s > r\,, \quad \text{a więc} \ \ k^s = k^r + lm\,, \quad \text{czyli} \ \ k^r(k^{s-r} - 1) = lm\,.$$

Z jednoznaczności rozkładu liczby naturalnej na czynniki pierwsze wynika, że $k^{s-r} - 1$ musi dzielić się przez m (bo $k < m$, więc nie dzieli się przez m), czyli $k^{s-r} = jm + 1$. Dostajemy $[k][k^{s-r-1}] = [1]$, a więc każdy element $[k]$ zbioru $\mathbb{Z}/\mathrm{mod}\, m$ różny od $[0]$ ma element odwrotny względem mnożenia. Reasumujemy:

> Struktura $(\mathbb{Z}/\mathrm{mod}\, m, +, \cdot)$ jest ciałem wtedy, i tylko wtedy, gdy m jest liczbą pierwszą. Charakterystyka tego ciała jest w tym przypadku równa m.

Ciało określone w przykładzie (ii) jest szczególnym przypadkiem tej struktury.

12 Funkcje o wartościach w ciele

Termin ***funkcja*** jest w zasadzie inną nazwą odwzorowania, jednak w praktyce używa się go na ogół dla odwzorowań $f : X \mapsto \mathbb{K}$, gdzie X jest dowolnym zbiorem, a $\mathbb{K}$ ciałem. W szczególności, jeśli $\mathbb{K} = \mathbb{R}$, to f nazywamy ***funkcją rzeczywistą na*** X, a jeśli $\mathbb{K}$ jest ciałem liczb zespolonych (omawianych w następnym paragrafie), to f nazywa się ***funkcją zespoloną na*** X.

Na zbiorze funkcji z X w ciało $\mathbb{K}$ określa się trzy działania. Niech f i g będą funkcjami w tym zbiorze, a λ niech będzie elementem $\mathbb{K}$. Pierwsze dwa spośród trzech wspomnianych działań to:

(i) ***dodawanie funkcji*** (działanie wewnętrzne): ***suma*** funkcji f i g jest zadana przepisem

$$f + g : X \mapsto \mathbb{K}\,, \quad (f+g)(x) = f(x) + g(x)\,;$$

(ii) ***mnożenie funkcji przez elementy ciała*** (działanie zewnętrzne): ***iloczyn funkcji*** f ***przez liczbę*** λ jest zadany przez

$$\lambda f : X \mapsto \mathbb{K}\,, \quad (\lambda f)(x) = \lambda f(x)\,.$$

Zwróćmy uwagę, że po prawej stronie definicji znaki dodawania i mnożenia dotyczą elementów ciała $\mathbb{K}$, a po lewej stronie te same znaki są użyte w nowym znaczeniu, definiowanym właśnie przez te równości.

Dodawanie funkcji jest łączne (na mocy łączności dodawania w ciele), ma element neutralny – jest to funkcja zerowa (jej wartość jest równa zeru dla każdego argumentu), a każda funkcja ma element odwrotny względem dodawania, nazywamy ją funkcją przeciwną: $(-f)(x) := -f(x)$. Zbiór funkcji z X w $\mathbb{K}$ stanowi więc z działaniem dodawania grupę abelową. Łatwo sprawdzić następujące własności mnożenia funkcji przez liczbę:

$$1f = f\,, \quad \alpha(\beta f) = (\alpha\beta)f\,,$$

oraz prawa rozdzielności:

$$(\alpha + \beta)f = \alpha f + \beta f\,, \quad \alpha(f + g) = \alpha f + \alpha g\,.$$

Tak wprowadzone działania, wraz z ich własnościami, nazywamy ***strukturą liniową*** na zbiorze funkcji. Funkcję $\alpha f + \beta g$ nazywamy ***kombinacją liniową funkcji*** f ***i*** g, ze współczynnikami α i β.

Trzecim działaniem na zbiorze funkcji $X \mapsto \mathbb{K}$ jest:

(iii) ***mnożenie funkcji*** (działanie wewnętrzne): ***iloczyn funkcji*** f ***i*** g jest zadany przez
$$fg : X \mapsto \mathbb{K}\,, \quad (fg)(x) = f(x)g(x)\,.$$

Działanie to jest łączne i przemienne, a jego elementem neutralnym jest funkcja jednostkowa – jej wartość jest równa 1 dla każdego argumentu. Zbiór funkcji z tak określonym mnożeniem jest więc półgrupą z jedynką.

Działania mnożenia i dodawania funkcji spełniają prawo rozdzielności:

$$(f + g)h = fh + gh\,.$$

Zbiór funkcji $X \mapsto \mathbb{K}$ z działaniem dodawania (i) i mnożenia funkcji (iii) jest więc przemiennym pierścieniem z jedynką (patrz przykład (iv), p. 11).

13 Wielomiany

Rozważmy szczególny przypadek struktury omówionej w poprzednim punkcie, funkcje $\mathbb{K} \mapsto \mathbb{K}$. Wśród tych funkcji możemy wyróżnić klasę ***funkcji wielomianowych***, które można zadać przepisem

$$P : \mathbb{K} \mapsto \mathbb{K}\,, \quad P(x) = a_0 + a_1 x + \ldots + a_n x^n\,,$$

gdzie wszystkie liczby a_i są elementami ciała $\mathbb{K}$. Łatwo przekonać się, że ta klasa tworzy pierścień. Mając intuicje wyrobione na ciele liczb rzeczywistych skłonni bylibyśmy przyjąć, że przedstawienie takiej funkcji za pomocą współczynników a_i jest jednoznaczne. Tak jednak w ogólności nie jest, np. łatwo sprawdzić, że jeśli x jest dowolnym elementem ciała $\mathbb{Z}/\mathrm{mod}\,3$ (przykład (iv), p. 11), to $x^3 = x$, więc przepisy $\mathbb{Z}/\mathrm{mod}\,3 \ni x \mapsto x \in \mathbb{Z}/\mathrm{mod}\,3$ i $\mathbb{Z}/\mathrm{mod}\,3 \ni x \mapsto x^3 \in \mathbb{Z}/\mathrm{mod}\,3$ zadają tę samą funkcję. Ten właśnie fakt jest przyczyną użycia w powyższej definicji terminu *funkcja* wielomianowa, a zarezerwowania terminu ***wielomian*** na konkretny przepis funkcyjny w postaci kombinacji liniowej ***jednomianów*** $x \mapsto x^i$, przy czym wszystkie jednomiany traktujemy jako niezależne. Mówimy, że wielomian P jest ***stopnia*** $n > 0$, gdy x^n jest jednomianem najwyższej potęgi wchodzącym do tej kombinacji z niezerowym współczynnikiem, czyli P ma postać

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n\,, \quad a_n \neq 0\,, \quad n > 0\,,$$

oraz że P jest ***stopnia zerowego***, gdy

$$P(x) = a_0\,.$$

W szczególności ***wielomian zerowy*** $P = 0$ jest w ten sposób dołączony do zbioru wielomianów stopnia zerowego. Wielomiany pozostają więc we wzajemnie jednoznacznej odpowiedniości z ciągami współczynników $(a_0, a_1, \ldots)$, w których tylko skończona liczba wyrazów jest różna od zera; notację upraszczamy urywając ciąg na ostatnim wyrazie różnym od zera.

We wspomnianym powyżej przykładzie ciała $\mathbb{Z}/\mathrm{mod}\,3$ wielomian pierwszego stopnia $P_1(x) = x$ jest zadany ciągiem współczynników $(0, 1)$, wielomian trzeciego stopnia $P_3(x) = x^3$ – ciągiem $(0, 0, 0, 1)$, ale oba wielomiany określają tę samą funkcję wielomianową.

Dodawanie i mnożenie wielomianów jest określone tak samo jak dla funkcji, przy uwzględnieniu równości $x^i x^j = x^{i+j}$, ale z utrzymaniem założenia o niezależności jednomianów. Tak więc, jeśli $P(x)$ jest określony powyższym przepisem, a $Q(x) = b_0 + \ldots b_m x^m$, $b_m \neq 0$, to współczynniki wielomianów będących wynikiem działań są dane odpowiednio przez:

$$\begin{aligned} \alpha P + \beta Q: &\qquad c_i = \alpha a_i + \beta b_i\,,\\ PQ: &\qquad d_i = \sum_{j+l=i} a_j b_l\,. \end{aligned}$$

Iloczyn niezerowych wielomianów jest więc wielomianem niezerowym, którego stopień jest równy sumie stopni wielomianów w iloczynie. Prawo rozdzielności, jak łatwo sprawdzić, przenosi się z ciała na wielomiany. Wielomiany o współczynnikach w ciele tworzą więc przemienny pierścień z jedynką. Pierścień ten nie jest ciałem, ale nie ma w nim nietrywialnych podzielników zera.

Wartościami wielomianu będziemy nazywać wartości określanej przez niego funkcji.

Twierdzenie 14 (Wielomian interpolacyjny Lagrange'a)**.**
Niech będą dane dwa ciągi liczb ciała $\mathbb{K}$: $(x_1, \ldots, x_{n+1})$ i $(p_1, \ldots, p_{n+1})$, przy czym pierwszy z tych ciągów jest różnowartościowy, $x_i \neq x_j$ dla $i \neq j$. Wtedy wśród wielomianów stopnia mniejszego lub równego n istnieje dokładnie jeden wielomian P taki, że $P(x_i) = p_i$, $i = 1, \ldots, n+1$. Wielomian ten jest równy

$$P(x) = \sum_{i=1}^{n+1} p_i \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)} .$$

Dowód. Spełnienie warunków przez wielomian P jest widoczne. Gdyby istniał inny wielomian P' spełniający te same warunki, to wartości wielomianu $P' - P$ byłyby równe zeru w punktach $x_1, \ldots, x_{n+1}$, co jest sprzeczne z wynikiem twierdzenia z następnego punktu (wielomian $P' - P$ jest stopnia co najwyżej n). Inny dowód twierdzenia podamy w dalszej części podręcznika (przykład (v), p. 8, §7). □

Z twierdzenia otrzymujemy wniosek pełniej charakteryzujący relację pomiędzy funkcjami a wielomianami.

Wniosek 15.
(i) Jeśli ciało $\mathbb{K}$ ma nieskończenie wiele elementów, to każda funkcja wielomianowa $\mathbb{K} \mapsto \mathbb{K}$ jest zadana dokładnie jednym wielomianem, więc te dwa pojęcia można utożsamić.
(ii) Jeśli ciało ma m elementów, to każda funkcja $\mathbb{K} \mapsto \mathbb{K}$ jest funkcją wielomianową, i wśród wielomianów stopnia mniejszego lub równego $m - 1$ istnieje dokładnie jeden, który ją zadaje.

Dowód. (i) Załóżmy, że wielomiany: P stopnia n i Q stopnia n' zadają tę samą funkcję wielomianową i $n \geq n'$. Wybierzmy $n + 1$ różnych elementów ciała $x_1, \ldots, x_{n+1}$. Istnieje dokładnie jeden wielomian stopnia mniejszego lub równego n przyjmujący dla tych argumentów wartości $P(x_i) = Q(x_i)$, więc $Q = P$.

(ii) Niech $\mathbb{K} = \{x_1, \ldots, x_m\}$. Dowolna funkcja $f : \mathbb{K} \mapsto \mathbb{K}$ jest jednoznacznie zadana swoimi wartościami $f(x_i)$. Z poprzedniego twierdzenia wnioskujemy o istnieniu dokładnie jednego wielomianu stopnia mniejszego lub równego $m-1$, przyjmującego wartości $f(x_i)$. □

Założenie punktu (i) powyższego wniosku jest spełnione, w szczególności, dla liczb rzeczywistych i dla liczb zespolonych, które to ciała będą miały dla nas największe znaczenie. Zwróćmy uwagę, że w przypadku (ii) jednoznaczność przypisania funkcji wielomianu obowiązuje tylko w określonej klasie wielomianów.

14 Pierwiastki wielomianu

Pierwiastkami wielomianu nazywamy miejsca zerowe określanej przez niego funkcji, czyli argumenty, dla których wartość wielomianu jest równa zeru. Znajdziemy w tym punkcie ich związek z rozkładem wielomianu na prostsze czynniki.

Zauważamy najpierw, że po wybraniu dowolnego elementu $y \in \mathbb{K}$ każdy wielomian stopnia $n > 0$

$$P(x) = a_0 + \ldots + a_n x^n\,, \quad a_n \neq 0\,,$$

można przedstawić w postaci

$$P(x) = b_0 + b_1(x-y) + \ldots + b_n(x-y)^n\,, \quad b_n = a_n \neq 0\,.$$

Dla dowodu wystarczy zapisać $x = (x-y) + y$ i wyliczyć potęgi x^i jako kombinacje liniowe wyrażeń $y^k(x-y)^l$. Zauważmy, że $b_0 = P(y)$. Niech k będzie najmniejszą liczbą naturalną, dla której $b_k \neq 0$ (zawsze taka liczba istnieje, bo $b_n \neq 0$). Wtedy

$$P(x) = (x-y)^k Q(x) + P(y)\,,$$

gdzie Q jest wielomianem stopnia $n-k$, niezerującym się w $x = y$ (zależnym od wyboru y). Jeśli wielomian P ma pierwiastki $x_1, \ldots, x_s$, to wybierzmy w tym schemacie $y = x_1$. Wtedy dla pewnej liczby naturalnej n_1

$$P(x) = (x-x_1)^{n_1} P_1(x)\,,$$

gdzie P_1 jest wielomianem stopnia $n - n_1$, o pierwiastkach równych $x_2, \ldots, x_s$. Stosujemy teraz tę samą procedurę do wielomianu P_1 i pierwiastka x_2, itd. aż do wyczerpania wszystkich pierwiastków. Po ich wyczerpaniu dostajemy następujący wynik.

Twierdzenie 16. *Liczby $x_1, \ldots, x_s$ są wszystkimi pierwiastkami niezerowego wielomianu P stopnia n wtedy, i tylko wtedy, gdy istnieje przedstawienie*

$$P(x) = (x-x_1)^{n_1} \ldots (x-x_s)^{n_s}\, Q(x)\,,$$

gdzie Q jest wielomianem stopnia $n - n_1 - \ldots - n_s$, nie posiadającym pierwiastków. W szczególności: liczba pierwiastków niezerowego wielomianu jest mniejsza lub równa jego stopniowi.

Liczbę n_i będącą wykładnikiem w czynniku $(x - x_i)^{n_i}$ powyższego rozkładu nazywamy ***krotnością pierwiastka*** x_i.

15 Przykłady

(i) Dwumian Newtona
Dla rzeczywistych liczb x, y dowodzi się przez indukcję wzoru Newtona

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \,.$$

Czy formuła pozostaje w mocy dla innych ciał? Można przekonać się, że dowód indukcyjny pozostaje w mocy, należy tylko zauważyć, że symbol $\binom{n}{k}$ jest w przypadku liczb rzeczywistych liczbą naturalną, i rozszerzyć jego znaczenie dla innych ciał za pomocą reguły podanej przy końcu punktu 10. Wzorem Newtona można więc posłużyć się w ostatnim punkcie dla wyliczenia potęg wyrażenia $(x-y)+y$.

(ii) Trójmian kwadratowy
Wielomian drugiego stopnia, $P(x) = ax^2 + bx + c$, nazywa się często trójmianem kwadratowym. Jeśli charakterystyka ciała $\mathbb{K}$ jest różna od 2, to ma ono co najmniej trzy elementy $0, \pm 1$. W takim ciele funkcja wielomianowa drugiego stopnia zadaje więc jednoznacznie określający ją trójmian (i podobnie dla stopnia równego 1).

(iii) Dzielenie wielomianów
Niech P i Q będą wielomianami stopnia n i m odpowiednio, przy czym $n \geq m \geq 1$, i niech a_n i b_m będą ich najwyższymi różnymi od zera współczynnikami. Wielomian

$$P_1(x) = P(x) - \frac{a_n}{b_m} x^{n-m} Q(x)$$

jest stopnia niższego niż n. Jeśli ten stopień jest większy lub równy m, to powtarzamy powyższy schemat biorąc za punkt wyjścia wielomiany P_1 i Q i otrzymując wielomian P_2. Kontynuujemy, aż uzyskamy wielomian $P_k \equiv R$ stopnia niższego niż Q. Składając rezultaty kolejnych kroków dostajemy:

$$P = WQ + R \,,$$

gdzie W jest wielomianem stopnia $n-m$, a stopień wielomianu R jest mniejszy od m. Jeśli w otrzymanym wyniku $R = 0$, to mówimy, że ***wielomian*** P ***dzieli się przez wielomian*** Q, lub że ***wielomian*** Q ***dzieli*** P. (Jeśli Q jest niezerowym wielomianem stopnia $m = 0$, to każdy wielomian P dzieli się przez Q bez reszty.)

Jeśli x_1 jest znanym nam pierwiastkiem wielomianu P, to zagadnienie znalezienia pozostałych pierwiastków upraszcza się przez podzielenie $P(x)$ przez

$x - x_1$ (ewentualnie wielokrotne – aż do otrzymania wielomianu niezerującego się w x_1).

(iv) Wielomiany pierwsze, wielomiany względnie pierwsze
Wielomian stopnia niezerowego nazywamy ***wielomianem pierwszym*** (lub ***nieredukowalnym***), jeśli jedynymi jego podzielnikami są: 1 i on sam (i wielomiany do nich proporcjonalne). W każdym ciele dwumiany $P(x) = x - x_0$ są wielomianami pierwszymi. W ciele liczb rzeczywistych każdy trójmian kwadratowy o wyróżniku mniejszym od zera również jest pierwszy. Pokażemy później, że wyczerpuje to zbiór rzeczywistych wielomianów pierwszych.

Niezerowe wielomiany $P_1, \ldots, P_s$, $s \geq 2$, są ***względnie pierwsze***, gdy jedynym ich wspólnym podzielnikiem jest 1. Jeśli ten warunek jest spełniony, to istnieją wielomiany $Q_1, \ldots, Q_s$ takie, że

$$\sum_{i=1}^{s} P_i Q_i = 1\,.$$

Dowód tego faktu przeprowadza się metodą indukcji względem najmniejszego ze stopni wielomianów P_i. Załóżmy, bez ograniczania ogólności, że stopień wielomianu P_s jest najmniejszy (w sensie nierówności słabej). Jeśli ten stopień jest równy 0, to teza oczywista (kładziemy $Q_s = 1/P_s$ i $Q_i = 0$, $i = 1, \ldots, s-1$). W przeciwnym razie dzielimy: $P_i = W_i P_s + P'_i$, $i = 1, \ldots, s-1$, i przenosimy $P_s = P'_s$. Stopnie wielomianów P'_i, $i = 1, \ldots, s-1$, są mniejsze (nierówność silna) od stopnia P_s i co najmniej jeden z tych wielomianów jest niezerowy. Wielomiany niezerowe spośród wielomianów $P'_1, \ldots, P'_s$ są względnie pierwsze (przeciwne przypuszczenie prowadzi do sprzeczności z założeniem). Zakładamy więc, że dla tego układu wielomianów teza jest prawdziwa: dla pewnych wielomianów Q'_i jest $\sum_{i=1}^{s} P'_i Q'_i = 1$ (dla zerowych P'_i kładziemy $Q'_i = 0$). Stąd dla P_i teza jest spełniona przy podstawieniu $Q_i = Q'_i$, $i = 1, \ldots, s-1$, $Q_s = Q'_s - \sum_{i=1}^{s-1} W_i Q'_i$.

Dowód powyższy daje też metodę konstrukcji szukanych wielomianów Q_i: redukując minimalny stopień należy go sprowadzić do stopnia zerowego. Niech na przykład dane będą rzeczywiste wielomiany $P_1(x) = x^2 - 1$, $P_2(x) = x - 2$. Przy oznaczeniach jak wyżej kładziemy $P'_2 = P_2$ i znajdujemy $W_1(x) = x + 2$, $P'_1(x) = 3$. Dla wielomianów P'_i teza jest spełniona z wielomianami $Q'_1(x) = \frac{1}{3}$, $Q'_2(x) = 0$. Stąd $Q_1(x) = \frac{1}{3}$, $Q_2(x) = -\frac{1}{3}(x+2)$.

(v) Rozkład wielomianu na wielomiany pierwsze
Jeśli wielomian P nie jest pierwszy, to ma nietrywialne podzielniki, więc daje

się przedstawić jako iloczyn wielomianów. Każdy z tych wielomianów albo jest pierwszy, albo znów może być rozseparowany. Kontynuując dostajemy rozkład wielomianu na wielomiany pierwsze:

$$P = U_1 \dots U_s \,.$$

Rozkład ten jest jednoznaczny z dokładnością do kolejności czynników i pomnożenia ich przez niezerowe liczby (których łączny iloczyn jest równy 1). Dowód tego faktu opiera się na następującym pomocniczym rezultacie. Jeśli iloczyn wielomianów $W_1 W_2$ dzieli się przez wielomian pierwszy U, to jeden z wielomianów W_i dzieli się przez U. Istotnie, załóżmy, że W_1 nie dzieli się przez U (w przeciwnym przypadku teza jest spełniona). Wtedy W_1 i U są względnie pierwsze, więc istnieją wielomiany R, S takie, że $W_1 R + US = 1$. Stąd $W_2 = W_2 1 = (W_2 W_1)R + W_2 US$, więc W_2 dzieli się przez U, co zamyka dowód pomocniczego rezultatu. Dokończenie dowodu jednoznaczności pozostawiamy czytelnikowi jako ćwiczenie.

(vi) Redukcja pierwiastków
Wielomian rzeczywisty $P(x) = x^4 - x^3 + 2x^2 - 5x + 3$ zeruje się dla $x = 1$. Dzielimy go przez $(x - 1)$:

$$P(x) = x^3(x-1) + 2x^2 - 5x + 3 = (x^3 + 2x)(x-1) - 3x + 3 = (x^3 + 2x - 3)(x-1)\,.$$

Wielomian $P_1(x) = x^3 + 2x - 3$ nadal zeruje się w $x = 1$, więc znów dzielimy go przez $(x - 1)$:

$$P_1(x) = x^2(x-1) + x^2 + 2x - 3 = (x^2 + x)(x-1) + 3x - 3 = (x^2 + x + 3)(x-1)\,.$$

Wyróżnik trójmianu kwadratowego $Q(x) = x^2 + x + 3$ jest ujemny, więc nie ma on pierwiastków. Dostajemy rozkład $P(x) = (x - 1)^2(x^2 + x + 3)$, o którym mówi twierdzenie 16.

(vii) Wyrażenia wymierne
Konstrukcja wyrażeń wymiernych z wielomianów przebiega podobnie jak konstrukcja liczb wymiernych z liczb naturalnych. W zbiorze par wielomianów w ciele $\mathbb{K}$: (P, Q), gdzie $Q \neq 0$, wprowadza się relację:

$$(P, Q) \sim (P', Q') \iff PQ' = P'Q\,.$$

Ta relacja jest relacją równoważności; jej klasy równoważności $[(P, Q)]$ oznaczamy $\dfrac{P}{Q}$ i nazywamy ***wyrażeniami wymiernymi***. Do tych wyrażeń stosują się zwykłe reguły działań na ułamkach. Każde wyrażenie wymierne można sprowadzić do postaci, w której P i Q są względnie pierwsze; wtedy przepis $x \mapsto \dfrac{P(x)}{Q(x)}$ określa ***funkcję wymierną*** na zbiorze $\mathbb{K} \setminus Z_Q$, gdzie Z_Q jest zbiorem pierwiastków wielomianu Q.

(viii) Wielomiany wielu zmiennych
Wielomiany nad ciałem $\mathbb{K}$ w zmiennych $x_1, \ldots, x_k$ konstruuje się przez uogólnienie przypadku jednej zmiennej:

$$P(x_1, \ldots, x_k) = \sum_{l_1, \ldots, l_k} a_{l_1 \ldots l_k} x_1^{l_1} \ldots x_k^{l_k} \,,$$

gdzie współczynniki $a_{l_1 \ldots l_k}$ są elementami ciała $\mathbb{K}$ i tylko skończona ich liczba jest różna od zera. Stopniem wielomianu w zmiennej x_i jest najwyższa potęga tej zmiennej występująca w niezerowym wyrazie wielomianu, lub liczba zero, gdy wielomian jest zerowy. Wprowadźmy następujące oznaczenia zmiennych

$$X' = (x_1, \ldots, x_{k-1}) \,, \quad x = x_k \,, \quad X = (x_1, \ldots, x_{k-1}, x) \,.$$

Każdy wielomian wielu zmiennych można wtedy przedstawić w postaci

$$P(X) = p_0(X') + p_1(X')x + \ldots + p_n(X')x^n \,,$$

gdzie $p_i(X')$ są wielomianami, $p_n(X') \neq 0$ i n jest wtedy stopniem wielomianu P w zmiennej x. Jeśli $n = 0$, to wielomian P redukuje się do wielomianu w zmiennych X'. Podobny do powyższego zapis można zastosować przy zastąpieniu zmiennej x dowolną inną spośród zmiennych wielomianu.

Stosując powyższe przedstawienie wielomianów łatwo pokazać przez indukcję względem liczby zmiennych, że iloczyn wielomianów niezerowych jest niezerowy.

(ix) Dzielenie wielomianów wielu zmiennych
Przejmujemy oznaczenia zmiennych z poprzedniego przykładu. Jeśli wielomian $P(X)$ ma rozkład $P(X) = W(X)Q(X)$, to mówimy, że P dzieli się przez Q. Jeśli $P(X)$ nie ma podzielników stopnia niezerowego, to powiemy, że jest wielomianem pierwszym. Ogólnie, jeśli m jest stopniem wielomianu $Q(X)$ w zmiennej x i $n \geq m \geq 1$, to istnieje rozkład

$$s(X')P(X) = W(X)Q(X) + R(X) \,,$$

przy czym stopień wielomianu $R(X)$ w zmiennej x jest mniejszy od stopnia wielomianu $Q(X)$ w tej zmiennej, a wielomian $s(X')$ jest niezerowy.

Dowód tego faktu przebiega podobnie, jak w przykładzie (iii), ale z pomnożeniem pierwszego równania w tym przykładzie przez b_m, który to symbol oznacza teraz wielomian $b_m(X')$. Analogicznie zostają zmodyfikowane następne kroki. Jeśli $R = 0$, a wielomiany współczynnikowe przy potęgach zmiennej x w wielomianie $W(X) = w_0(X') + w_1(X')x + \ldots + w_k(X')x^k$ dzielą się przez $s(X')$, to P dzieli się przez Q.

(x) Rozkład wielomianu wielu zmiennych na wielomiany pierwsze
Podobnie jak w przypadku wielomianów jednej zmiennej, rozkład wielomianu na wielomiany pierwsze

$$P = U_1 \ldots U_s$$

jest jednoznaczny z dokładnością do kolejności czynników i pomnożenia ich przez niezerowe stałe, których łączny iloczyn jest równy 1.

Zastosujemy zasadę indukcji względem liczby zmiennych wielomianu.

1° Jednoznaczność rozkładu obowiązuje dla wielomianów jednej zmiennej.

2° Zakładamy więc na potrzeby pozostałej części rozumowania, że jednoznaczność rozkładu obowiązuje dla wielomianów zmiennych X' (przy oznaczeniach poprzednich przykładów). Mamy wykazać, że pozostaje słuszna dla zmiennych X.

Ten dowód przeprowadzamy znów metodą indukcji, teraz względem stopnia rozkładanego wielomianu w zmiennej x.

1°′ Jeśli wielomian jest stopnia zerowego w x, to jest wielomianem w zmiennych X', więc na mocy założenia indukcyjnego 2° jego rozkład jest jednoznaczny.

2°′ Załóżmy więc, że jednoznaczność obowiązuje dla wielomianów stopnia mniejszego od n w zmiennej x i niech wielomian $P(X)$ będzie stopnia n w tej zmiennej. Mamy wykazać, że jego rozkład jest jednoznaczny.

Zauważmy najpierw, że w rozkładzie P czynniki pierwsze stopnia zerowego w zmiennej x są określone jednoznacznie na mocy założenia 2°. Istotnie, jeśli bowiem $P(X)$ dzieli się przez wielomian $q(X')$, to wszystkie wielomiany współczynnikowe $p_i(X')$ (zob. przykład (viii)), których rozkład na czynniki jest jednoznaczny, dzielą się przez $q(X')$. Tym samym również czynnik pierwszy stopnia n w x, jeśli występuje w rozkładzie P, jest jednoznaczny. Wystarczy zatem rozważyć dalej czynniki pierwsze stopni $1, \ldots, n-1$.

Załóżmy, że $P(X)$ dzieli się przez dwa istotnie różne (tj. nie różniące się jedynie stałym czynnikiem) wielomiany pierwsze $U_1(X)$ i $U_2(X)$ takich stopni. Istnieją zatem wielomiany $Q_1(X)$ i $Q_2(X)$ takie, że $P = Q_1U_1 = Q_2U_2$. Na mocy założenia indukcyjnego 2°′ rozkłady wielomianów Q_i na wielomiany pierwsze są jednoznaczne, więc każdy z wielomianów U_i określa jednoznacznie rozkład wielomianu P, w którym występuje. Wystarczy więc pokazać, że oba wielomiany

U_i występują w jednym rozkładzie. Niech stopień wielomianu $U_2(X)$ w zmiennej x będzie większy lub równy stopniowi wielomianu $U_1(X)$ w tej zmiennej i podzielmy: $s(X')U_2(X) = W(X)U_1(X) + R(X)$ (zgodnie z przykładem (ix)). Obie strony tej równości są wielomianami stopnia mniejszego od n w x, więc ich rozkłady są jednoznaczne na mocy założenia $2^{\circ\prime}$, a zatem $R \neq 0$. Ze związków między wielomianami wyliczamy $Q_2R = (sQ_1 - Q_2W)U_1$, więc wielomian Q_2R dzieli się przez U_1. Na mocy założenia indukcyjnego $2^{\circ\prime}$ rozkład wielomianu Q_2R na wielomiany pierwsze jest jednoznaczny i równocześnie U_1 nie może wystąpić w rozkładzie R (stopień R w x jest mniejszy od stopnia U_1). Stąd U_1 wystąpi w rozkładzie Q_2, czyli $Q_2 = TU_1$ dla pewnego wielomianu T. Ostatecznie $P = TU_1U_2$, co należało dowieść.

§4 Liczby zespolone

1 Motywacja

Niedogodnością liczb rzeczywistych jest fakt, że nie każdy wielomian zmiennej rzeczywistej ma pierwiastki. Weźmy pod uwagę najprostszy wielomian tego typu, $P(x) = x^2 + 1$. Chcemy rozszerzyć ciało liczb rzeczywistych w taki sposób, aby równanie $x^2 + 1 = 0$ miało rozwiązanie. Załóżmy, że jest to możliwe, i oznaczmy rozwiązanie tego równania przez $\sqrt{-1}$. Postępując tak, jakby element ten mógł być dołączony do liczb rzeczywistych, otrzymamy zbiór *liczb* postaci $x + \sqrt{-1}\,y$, $x, y \in \mathbb{R}$. Żądając, aby prawa łączności i rozdzielności pozostały w mocy, dostaniemy następujące reguły działań:

$$(x_1 + \sqrt{-1}\,y_1) + (x_2 + \sqrt{-1}\,y_2) = (x_1 + x_2) + \sqrt{-1}\,(y_1 + y_2)\,,$$
$$(x_1 + \sqrt{-1}\,y_1)(x_2 + \sqrt{-1}\,y_2) = (x_1x_2 - y_1y_2) + \sqrt{-1}\,(x_1y_2 + x_2y_1)\,.$$

2 Ciało liczb zespolonych

Okazuje się, że powyższe heurystyczne rozważania mogą być wykorzystane do poprawnej konstrukcji nowego ciała spełniającego nasze żądania. Wystarczy w tym celu zastąpić nieformalne wyrażenie $x + \sqrt{-1}\,y$ parą liczb rzeczywistych (x, y).

Twierdzenie 1. *Zbiór $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ z działaniami zadanymi przez*

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2,\ y_1 + y_2)\,,$$
$$(x_1, y_1)(x_2, y_2) := (x_1x_2 - y_1y_2,\ x_1y_2 + x_2y_1)$$

stanowi ciało.

Ciało wprowadzone tym twierdzeniem nazywamy ***ciałem liczb zespolonych***. Zwróćmy uwagę, że mnożenie i dodawanie występujące po prawej stronie tych definicji są znanymi działaniami w zbiorze liczb rzeczywistych; te same znaki po lewej stronie wprowadzają działania w nowym ciele.

Elementem neutralnym dodawania jest w ciele liczb zespolonych $(0,0)$, a elementem neutralnym mnożenia jest $(1,0)$. Elementy przeciwne oraz odwrotne dane są przez

$$-(x,y) = (-x,-y)\,, \qquad (x,y)^{-1} = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right).$$

Przy uwzględnieniu tych identyfikacji dowód twierdzenia sprowadza się do rachunkowego sprawdzenia warunków definicyjnych ciała, pozostawiamy go czytelnikowi jako ćwiczenie.

Podzbiór ciała $\mathbb{C}$ złożony z liczb postaci $(x,0)$, $x \in \mathbb{R}$, również jest ciałem. Odwzorowanie zbioru liczb rzeczywistych w ten zbiór zadane przez $x \mapsto (x,0)$ jest bijektywne oraz homomorficzne, co w przypadku ciała oznacza zachowywanie obu działań:

$$(x,0)+(y,0) = (x+y,0)\,, \qquad (x,0)(y,0) = (xy,0)\,.$$

Odwzorowanie to zadaje więc izomorfizm ciał, co pozwala na wprowadzenie utożsamienia $(x,0) \equiv x$. Wprowadźmy oznaczenie $i := (0,1)$; nazywamy ten element ***jednostką urojoną***. Używając reguły mnożenia liczb zespolonych pokazujemy łatwo, że

$$i^2 = (-1,0) \equiv -1\,.$$

Każdą liczbę zespoloną można zapisać jako

$$(x,y) = (x,0) + (0,1)(y,0)\,,$$

co przy powyższych utożsamieniach upraszcza się do zapisu zbliżonego do początkowej, nieformalnej notacji:

$$(x,y) = x + iy\,.$$

Jeśli oznaczyć tę liczbę zespoloną przez

$$z = x+iy\,, \quad \text{to piszemy} \quad x = \mathfrak{Re}\, z \ \text{ i } \ y = \mathfrak{Im}\, z\,,$$

i nazywamy te liczby rzeczywiste odpowiednio ***częścią rzeczywistą liczby*** z i ***częścią urojoną liczby*** z. Z praw dodawania liczb zespolonych wynika, że

$$\mathfrak{Re}(z+z') = \mathfrak{Re}\, z + \mathfrak{Re}\, z'\,, \quad \mathfrak{Im}(z+z') = \mathfrak{Im}\, z + \mathfrak{Im}\, z'\,.$$

3 Pierwiastki wielomianu

Widzimy teraz, że równanie $z^2 + 1 = 0$ ma w dziedzinie liczb zespolonych dokładnie dwa rozwiązania, gdyż jest równoważne równaniu $(z - i)(z + i) = 0$, a stąd i z własności ciała mamy $z = i$ lub $z = -i$. Okazuje się, że wraz z rozszerzeniem ciała liczb rzeczywistych do ciała liczb zespolonych otrzymujemy, niejako *gratis*, znacznie więcej: rozkład każdego wielomianu na iloczyn dwumiennych czynników prostych. Przytaczamy najpierw następujące twierdzenie.

Twierdzenie 2 („Zasadnicze twierdzenie algebry”)**.** *Każdy wielomian zmiennej zespolonej, o zespolonych współczynnikach, stopnia większego lub równego jeden, ma co najmniej jeden pierwiastek.*

Twierdzenia tego dowodzi się stosując metody wykraczające poza algebrę. Zamieszczamy taki dowód poniżej wśród przykładów (punkt 9).

Posługując się tym twierdzeniem łatwo wykazać zapowiedziany rezultat. Niech P będzie wielomianem stopnia $n \geq 1$. Posiada on rozkład dany twierdzeniem 16, §3, obowiązującym w każdym ciele. Wielomian Q występujący w tym rozkładzie nie ma pierwiastków, co w ciele liczb zespolonych jest możliwe tylko wtedy, gdy jest on niezerową stałą. Dostajemy więc następujący rezultat.

Twierdzenie 3. *Jeśli P jest wielomianem zespolonym stopnia $n \geq 1$, to istnieje rozkład*

$$P(z) = a\,(z - w_1)^{n_1} \ldots (z - w_s)^{n_s}\,, \quad n_1 + \ldots + n_s = n\,.$$

Uwzględniając krotność pierwiastków możemy ująć krótko to twierdzenie mówiąc, że każdy wielomian zespolony stopnia $n \geq 1$ ma dokładnie n pierwiastków. Zapisujemy ten rezultat niekiedy tak:

$$P(z) = a(z - z_1) \ldots (z - z_n)\,,$$

gdzie w ciągu $(z_1, \ldots, z_n)$ każda z liczb w_i występuje n_i razy $(i = 1, \ldots, s)$.

4 Płaszczyzna Gaussa

Liczby zespolone posiadają naturalną interpretację geometryczną. Wybierzmy na płaszczyźnie euklidesowej prostokątny układ współrzędnych. Liczby zespolone są parami liczb rzeczywistych, więc można im przypisać punkty na tej płaszczyźnie: liczbie $z = x + iy$ odpowiada punkt o odciętej x i rzędnej y. Płaszczyznę, na której w ten sposób reprezentujemy liczby zespolone, nazywamy ***płaszczyzną Gaussa***, punkty na niej oznaczamy odpowiadającymi liczbami zespolonymi. Przy tej reprezentacji dodawanie liczb zespolonych ma prostą interpretację geometryczną: punkt $z + z'$ jest końcem skierowanego odcinka o początku w punkcie 0, otrzymanego regułą równoległoboku ze złożenia skierowanych odcinków $\overrightarrow{0z}$ i $\overrightarrow{0z'}$.

5 Moduł liczby zespolonej. Liczba sprzężona

Modułem (wartością bezwzględną) liczby zespolonej $z = x + iy$ nazywamy rzeczywistą liczbę

$$|z| := \sqrt{x^2 + y^2}\,.$$

Dla wszystkich liczb zespolonych mamy

$$|z| \geq 0\,, \quad \text{oraz} \quad |z| = 0 \Longleftrightarrow z = 0\,.$$

Wartość modułu $|z|$ jest równa długości odcinka $0z$ na płaszczyźnie Gaussa. Jeśli z jest liczbą rzeczywistą, to jej moduł pokrywa się z określeniem wartości bezwzględnej dla liczb rzeczywistych. Prawdziwe są oczywiste nierówności:

$$|\,\mathfrak{Re}\, z| \leq |z|\,, \quad |\,\mathfrak{Im}\, z| \leq |z|\,.$$

Liczbą sprzężoną do liczby $z = x + iy$ nazywamy liczbę

$$\overline{z} := x + i(-y) \equiv x - iy$$

(ostatnia równość wprowadza wygodną konwencję notacyjną). Na płaszczyźnie Gaussa liczba sprzężona jest odbiciem symetrycznym danej liczby względem osi odciętych. Pozostawiamy czytelnikowi jako ćwiczenie sprawdzenie następujących prostych własności operacji sprzężenia:

$$\overline{w + z} = \overline{w} + \overline{z}\,, \quad \overline{\overline{z}} = z\,, \quad \overline{wz} = \overline{w}\,\overline{z}\,, \quad \overline{z^{-1}} = \overline{z}^{\,-1}\,, \quad |\overline{z}| = |z|\,, \quad \overline{z}z = |z|^2\,.$$

Posługując się liczbą sprzężoną uzyskujemy nowy zapis definicji modułu, części rzeczywistej i urojonej:

$$|z| = \sqrt{\overline{z}z}\,, \quad \mathfrak{Re}\, z = \frac{1}{2}(z + \overline{z})\,, \quad \mathfrak{Im}\, z = \frac{1}{2i}(z - \overline{z})\,.$$

Użycie pojęcia liczby sprzężonej ułatwia przeprowadzanie wielu rachunków, np. wyliczenie elementu odwrotnego można teraz zapisać jako:

$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{\overline{z}z} = \frac{\overline{z}}{|z|^2} = \frac{\mathfrak{Re}\, z}{|z|^2} - i\frac{\mathfrak{Im}\, z}{|z|^2}\,.$$

Również dowód następujących dalszych własności modułu daje się prosto zapisać przy użyciu tego pojęcia.

Twierdzenie 4. *Dla dowolnej pary liczb zespolonych w i z zachodzi:*

(*i*) $|wz| = |w||z|\,,$ *w szczególności:* $|-z| = |z|\,;$

(*ii*) $|w + z| \leq |w| + |z|\,,$ *nierówność trójkąta*;

(*iii*) $|w - z| \geq \big||w| - |z|\big|\,.$

Dowód. (i) Tożsamość uzyskujemy z ciągu równości

$$|wz| = \sqrt{\overline{wz}\,wz} = \sqrt{\overline{w}w\,\overline{z}z} = |w||z|\,.$$

(ii) Z własności sprzężenia i operacji brania części rzeczywistej liczby zespolonej mamy ciąg (nie)równości

$$\begin{aligned}|w+z|^2 &= (\overline{w}+\overline{z})(w+z) = |w|^2+|z|^2+\overline{w}z+\overline{z}w = \\ &= |w|^2+|z|^2+2\,\mathfrak{Re}[\overline{w}z] \leq |w|^2+|z|^2+2|w||z| = (|w|+|z|)^2\,.\end{aligned}$$

(iii) Zastępując we własności (ii) liczbę w liczbą $w-z$ dostajemy

$$|w-z| \geq |w|-|z|\,.$$

Zamieniając tutaj miejscami w i z wyczerpujemy treść punktu (iii). □

Przy interpretacji gaussowskiej liczb zespolonych moduły $|w|$, $|z|$ i $|w+z|$ są długościami boków trójkąta, stąd nazwa własności (ii).

6 Argument liczby zespolonej. Reprezentacja trygonometryczna i wykładnicza

Argumentem liczby zespolonej $z = x + iy$ różnej od zera nazywamy każdą taką liczbę rzeczywistą φ, że

$$\cos\varphi = \frac{x}{|z|}\,, \quad \sin\varphi = \frac{y}{|z|}$$

(pamiętamy, że $x^2+y^2 = |z|^2$, więc taka liczba zawsze istnieje). Jeśli φ jest dowolnym argumentem z, to zbiór wszystkich argumentów tej liczby składa się z liczb $\varphi + 2k\pi$, $k \in \mathbb{Z}$. Przy użyciu argumentu liczbie zespolonej można nadać ***postać trygonometryczną***:

$$z = r(\cos\varphi + i\sin\varphi)\,,$$

gdzie $r = |z|$. Argument liczby z jest kątem, jaki tworzy na płaszczyźnie Gaussa odcinek $0z$ z półosią dodatnich odciętych, przy czym jeśli liczymy ten kąt od półosi do odcinka idąc przeciwnie do wskazówek zegara, to kąt ten ma znak dodatni; dla kierunku zgodnego ze wskazówkami kąt ma znak ujemny.

Stosując znane wzory trygonometryczne otrzymujemy tożsamość:

$$(\cos\varphi + i\sin\varphi)(\cos\varphi' + i\sin\varphi') = \cos(\varphi+\varphi') + i\sin(\varphi+\varphi')\,.$$

Tożsamość ta przypomina znaną własność rzeczywistej funkcji eksponencjalnej: $\exp(x)\exp(y) = \exp(x+y)$, lub w zapisie wykładniczym: $e^x e^y = e^{x+y}$. Okazuje się, że podobieństwo to nie jest przypadkowe. Metodami analizy matematycznej pokazuje się, że definicję funkcji eksponencjalnej można rozszerzyć na dziedzinę liczb zespolonych za pomocą rozwinięcia potęgowego

$$\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!} .$$

Oznacza to, że dla każdej liczby zespolonej z istnieje liczba oznaczana

$$\exp(z) \equiv e^z \quad \text{taka, że} \quad \lim_{N\to\infty} \left| e^z - \sum_{k=0}^{N} \frac{z^k}{k!} \right| = 0 .$$

Funkcja eksponencjalna ma następujące podstawowe własności:

$$e^z e^{z'} = e^{z+z'} , \quad \overline{e^z} = e^{\overline{z}} , \quad |e^z| = e^{\mathfrak{Re}\, z} ;$$

pierwszą z nich otrzymuje się przez przegrupowanie wyrazów w iloczynie dwóch szeregów, pozostałe są konsekwencją własności sprzężenia. W szczególności jeśli wybrać za argument eksponenty liczbę $i\varphi$, gdzie φ jest liczbą rzeczywistą, to mamy stąd

$$\left| e^{i\varphi} \right| = 1 , \quad e^{i\varphi} e^{i\varphi'} = e^{i(\varphi+\varphi')} .$$

Zespolona funkcja $e^{i\varphi}$ zmiennej rzeczywistej φ ma więc te same własności co funkcja $\cos\varphi + i \sin\varphi$. Okazuje się, że w istocie funkcje te są równe,

$$e^{i\varphi} = \cos\varphi + i \sin\varphi ,$$

co można pokazać na różne sposoby, np. przez porównanie rozwinięć potęgowych obu stron. Przy użyciu tej tożsamości możemy nadać każdej liczbie zespolonej ***postać wykładniczą***, którą zapisujemy poniżej wraz z jej podstawowymi własnościami:

$$z = re^{i\varphi} , \quad z' = r' e^{i\varphi'} , \quad zz' = rr' e^{i(\varphi+\varphi')} , \quad \forall n \in \mathbb{Z} : z^n = r^n e^{in\varphi} ;$$

ostatnia równość zapisana w postaci trygonometrycznej zwana jest ***wzorem de Moivre'a***. Warte specjalnego wyróżnienia są uderzające tożsamości otrzymane przez położenie w postaci wykładniczej i trygonometrycznej $r = 1$ oraz $\varphi = \pm 2\pi , \pm\pi , \pm\frac{\pi}{2}$:

$$e^{\pm i2\pi} = 1 , \quad e^{\pm i\pi} = -1, \quad e^{\pm i(\pi/2)} = \pm i .$$

7 Pierwiastkowanie liczb zespolonych

Z wcześniejszej dyskusji wiemy, że każdy wielomian zespolony stopnia $n \in \mathbb{N}$ ma n pierwiastków. Praktyczne znalezienie ich jednak jest proste tylko w przypadku $n = 1, 2$ (do przypadku $n = 2$ wrócimy poniżej), dla przypadków $n = 3, 4$ wymaga złożonej procedury, a dla $n \geq 5$ nie jest w ogólności możliwe algebraicznie. Jednak dla szczególnej klasy wielomianów dowolnego stopnia o postaci $P_{w,n}(z) = z^n - w$, $w \in \mathbb{C}$, problem ma pełne rozwiązanie, które jest wynikiem następującego twierdzenia. (Dla $w = 0$ problem jest trywialny – wszystkie pierwiastki są równe 0.)

Twierdzenie 5. *Niech będzie dana liczba zespolona $w \neq 0$ w postaci wykładniczej $w = re^{i\varphi}$. Wtedy dla $n \in \mathbb{N}$ wszystkie rozwiązania równania*

$$z^n = w \quad \textit{tworzą zbiór} \quad \left\{ r^{\frac{1}{n}} e^{i\frac{\varphi+2k\pi}{n}} \right\}_{k=0}^{n-1} .$$

Dowód. Zauważamy po pierwsze, że w zapisie zbioru liczb z tezy nie ma powtórzeń, tj. wszystkie liczby są różne między sobą, ich liczba wynosi więc n. Korzystając ze wzoru de Moivre'a łatwo sprawdzamy, że spełniają one równanie, są więc pierwiastkami wielomianu $z^n - w$, który nie może mieć ich więcej. □

Każdą liczbę z spełniającą warunki tezy nazywamy ***n-tym pierwiastkiem liczby*** w.

8 Wielomian kwadratowy

Twierdzenie z poprzedniego punktu daje kompletne rozwiązanie problemu pierwiastków liczb zespolonych, ale wymaga znajomości argumentu liczby pierwiastkowanej. W tym punkcie uzyskamy najpierw zapis pierwiastków drugiego stopnia nie odwołujący się do pojęcia argumentu. Przy oznaczeniach ostatniego twierdzenia pierwiastki te mają postać $\sqrt{r}e^{i(\varphi/2)}$ i $\sqrt{r}e^{i[(\varphi/2)+\pi]}$, czyli $\pm\sqrt{|w|}(\alpha + i\beta)$, gdzie $\alpha = \cos(\varphi/2)$, $\beta = \sin(\varphi/2)$. Wyciągając na zewnątrz nawiasu znak liczby α, te same dwa pierwiastki (choć może w innej kolejności) można zapisać jako $\pm\sqrt{|w|}\,(|\alpha| + i\,\mathrm{sgn}(\alpha\beta)|\beta|)$. Posługując się wzorami trygonometrycznymi

$$\big[\cos(\varphi/2)\big]^2 = \frac{1}{2}\,(1 + \cos\varphi)\ , \quad \big[\sin(\varphi/2)\big]^2 = \frac{1}{2}\,(1 - \cos\varphi)\ ,$$
$$2\sin(\varphi/2)\cos(\varphi/2) = \sin\varphi$$

wyliczamy: $|\alpha| = \sqrt{\frac{1}{2}\left(1 + \frac{\mathfrak{Re}\,w}{|w|}\right)}$, $|\beta| = \sqrt{\frac{1}{2}\left(1 - \frac{\mathfrak{Re}\,w}{|w|}\right)}$ i $\mathrm{sgn}[\alpha\beta] = \mathrm{sgn}[\mathfrak{Im}\,w]$.

Otrzymujemy zatem pierwiastki kwadratowe liczby w w następującej postaci:

$$\pm\left\{\sqrt{\frac{|w|+\mathfrak{Re}\,w}{2}}+i\,\mathrm{sgn}(\mathfrak{Im}\,w)\sqrt{\frac{|w|-\mathfrak{Re}\,w}{2}}\right\}.$$

Innym sposobem uzyskania tego samego rezultatu jest położenie $w=x+iy$, $z=a+ib$ i rozpisanie równania $z^2=w$ na część rzeczywistą i urojoną. Otrzymuje się układ dwóch równań kwadratowych na rzeczywiste liczby a i b. Pozostawiamy czytelnikowi przeprowadzenie rachunków tą metodą jako ćwiczenie.

Rozważmy następnie dowolny zespolony trójmian kwadratowy az^2+bz+c, $a,b,c\in\mathbb{C}$, $a\neq 0$. Proste przekształcenie sprowadza go do postaci rozważonej powyżej:

$$az^2+bz+c=a\left[\left(z+\frac{b}{2a}\right)^2-\frac{b^2-4ac}{4a^2}\right],$$

więc pierwiastki mają postać

$$z=\frac{-b\pm\sqrt{b^2-4ac}}{2a},$$

gdzie $\sqrt{b^2-4ac}$ jest jednym z pierwiastków wyliczonych w sposób opisany w poprzednim akapicie.

9 Przykłady

(i) Moduł i liczba odwrotna

$$|5-7i|=\sqrt{74}\,,\quad (5-7i)^{-1}=\frac{5+7i}{|5-7i|^2}=\frac{5}{74}+\frac{7}{74}i\,.$$

(ii) Twory geometryczne na płaszczyźnie Gaussa
Równanie $|z-z_0|=r$ zadaje na płaszczyźnie Gaussa okrąg o środku w z_0 i promieniu r.

Jeśli $z=re^{i\varphi}$, to równanie $r=a(1+\kappa\sin^2\varphi)^{-1/2}$ $(a>0,\ \kappa>-1)$ zadaje na płaszczyźnie Gaussa elipsę.

(iii) Interpretacja geometryczna iloczynu $\overline{z}z'$
Pomnożenie wszystkich liczb zespolonych przez liczbę $e^{i\psi}$, gdzie $\psi\in\mathbb{R}$, powoduje obrócenie płaszczyzny Gaussa o kąt ψ. Dla dowolnych liczb z, z' iloczyn $\overline{z}z'$ jest więc niezmiennikiem obrotów. Jeśli $z=x+iy=re^{i\varphi}$, $z'=x'+iy'=r'e^{i\varphi'}$, to $\mathfrak{Re}(\overline{z}z')=xx'+yy'=rr'\cos(\varphi'-\varphi)$ jest symetrycznym iloczynem (skalarnym) skierowanych odcinków $\overrightarrow{0z}$, $\overrightarrow{0z'}$, a $\mathfrak{Im}(\overline{z}z')=xy'-yx'=rr'\sin(\varphi'-\varphi)$ jest ich antysymetrycznym iloczynem, przy czym liczba $|\,\mathfrak{Im}(\overline{z}z')|$ jest polem powierzchni opartego na nich równoległoboku.

(iv) Funkcje hiperboliczne i trygonometryczne zmiennej zespolonej
Przyjmuje się definicje:

$$\cosh z := \frac{1}{2}(e^z + e^{-z}), \quad \sinh z := \frac{1}{2}(e^z - e^{-z}),$$
$$\cos z := \cosh iz, \quad \sin z := -i \sinh iz.$$

Funkcje te spełniają następujące tożsamości:

$$\begin{gathered}(\cosh z)^2 - (\sinh z)^2 = 1, \quad (\cos z)^2 + (\sin z)^2 = 1, \\ \cosh(z + z') = \cosh z \cosh z' + \sinh z \sinh z', \\ \sinh(z + z') = \sinh z \cosh z' + \cosh z \sinh z', \\ \cos(z + z') = \cos z \cos z' - \sin z \sin z', \\ \sin(z + z') = \sin z \cos z' + \cos z \sin z'.\end{gathered}$$

(v) Grupy pierwiastków z 1
Pierwiastki n-tego stopnia z 1 tworzą zbiór

$$\left\{ e^{ik\frac{2\pi}{n}} \right\}_{k=0}^{n-1} = \{1, \omega_n, (\omega_n)^2, \ldots, (\omega_n)^{n-1}\}, \quad \text{gdzie} \quad \omega_n = e^{i\frac{2\pi}{n}},$$

więc stanowią z mnożeniem grupę cykliczną rzędu n. ***Pierwiastkiem pierwotnym z*** 1 nazywamy każdą liczbę z tego zbioru, która generuje tę grupę, a więc każdą liczbę postaci $(\omega_n)^k$, gdzie k i n są względnie pierwsze (przykład (vii) w punkcie 3, §3).

(vi) Pierwiastki drugiego i trzeciego stopnia
Pierwiastki kwadratowe liczby $e^{i\frac{\pi}{6}} = \frac{\sqrt{3}}{2} + \frac{i}{2}$ wyliczone metodą punktu 8 to

$$\pm e^{i\frac{\pi}{12}} = \pm\frac{1}{2}\left[\sqrt{2+\sqrt{3}} + i\sqrt{2-\sqrt{3}}\right] = \pm\frac{1}{2\sqrt{2}}\Big[(\sqrt{3}+1) + i(\sqrt{3}-1)\Big].$$

Pierwiastki trzeciego stopnia z liczby $2 + 2i$ to liczby

$$\begin{gathered}\sqrt{2}e^{i\frac{\pi}{12}} = \frac{1}{2}\Big[(\sqrt{3}+1) + i(\sqrt{3}-1)\Big], \\ \sqrt{2}e^{i\frac{9\pi}{12}} = -\sqrt{2}e^{-i\frac{\pi}{4}} = -1 + i, \\ \sqrt{2}e^{i\frac{17\pi}{12}} = -\sqrt{2}ie^{-i\frac{\pi}{12}} = -\frac{1}{2}\Big[(\sqrt{3}-1) + i(\sqrt{3}+1)\Big].\end{gathered}$$

(vii) Równanie kwadratowe
Wyróżnik równania kwadratowego

$$z^2 + (-1+i)z + 4 + 7i = 0$$

jest równy $\Delta = -16 - 30i$. Szukamy jego pierwiastków w postaci $a + ib$, więc żądamy $a^2 - b^2 = -16$, $ab = -15$. Z drugiego równania widzimy, że $a \neq 0$. Mnożąc pierwsze równanie przez a^2 i eliminując $(ab)^2$ za pomocą drugiego dostajemy $a^4 + 16a^2 - 225 = 0$. Stąd $a^2 = 9$ (drugi pierwiastek jest ujemny), więc $a = \pm 3$. Teraz z drugiego równania wyliczamy b i dostajemy $a + ib = \pm(3 - 5i)$. Rozwiązaniami równania są liczby $2 - 3i$, $-1 + 2i$.

(viii) Dowód „zasadniczego twierdzenia algebry"
Najbardziej elegancki i zwięzły dowód wykorzystuje teorię funkcji analitycznych. Czytelnikom, którzy zapoznają się z tą teorią, poddajemy pod rozwagę ten dowód. Można zamknąć go w jednym zdaniu: jeśli wielomian $P(z)$ nie ma miejsc zerowych, to funkcja $1/P(z)$ jest analityczna na całej płaszczyźnie $\mathbb{C}$ i ograniczona, więc na mocy twierdzenia Liouville'a musi być stałą, co jest możliwe tylko gdy wielomian P jest stopnia zerowego.

Podamy teraz inny dowód, który oprócz pojęcia ciągłości wymaga zastosowania tylko jednego analitycznego, intuicyjnie oczywistego wyniku: jeśli funkcja ciągła $\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ dąży do $+\infty$ w nieskończoności (tj. dla $x^2 + y^2 \to \infty$), to istnieje punkt w $\mathbb{R}^2$, w którym ma ona globalne minimum.

Rozważamy wielomian $P(z)$ stopnia ≥ 1 jako funkcję x, y, gdzie $z = x + iy$. Można sprawdzić, że funkcja $|P(z)|^2$ ma własności założone w przytoczonym wyżej stwierdzeniu, więc musi posiadać globalne minimum (wynik ten milcząco wykorzystaliśmy też w pierwszej wersji dowodu). Chcemy pokazać, że wartość tego minimum wynosi zero. Wykażemy w tym celu, że jeśli $P(z_0) \neq 0$, to w sąsiedztwie punktu z_0 istnieją punkty, w których wartość $|P(z)|^2$ jest mniejsza niż $|P(z_0)|^2$. Bez ograniczenia ogólności wystarczy wykazać tę własność dla $z_0 = 0$, zakładając przy tym, że $P(0) = 1$. Przy tych założeniach wielomian P można zapisać w postaci $P(z) = 1 + z^k Q(z)$, gdzie k jest pewną liczbą naturalną $\leq n$, a Q jest wielomianem stopnia $n - k$ takim, że $Q(0) \neq 0$. Niech w będzie taką liczbą, że $w^k = -\overline{Q(0)}$ (zauważmy, że dla wskazania pierwiastków liczby zespolonej nie korzystaliśmy z zasadniczego twierdzenia) i połóżmy $z = w\lambda$, gdzie $\lambda \geq 0$. Wtedy $P(w\lambda) = 1 - \lambda^k \overline{Q(0)} Q(w\lambda)$, więc

$$|P(w\lambda)|^2 = 1 - \lambda^k W_w(\lambda)\,, \quad W_w(\lambda) \equiv 2\,\mathfrak{Re}\left[\overline{Q(0)}Q(w\lambda)\right] - \lambda^k \left|\overline{Q(0)}Q(w\lambda)\right|^2.$$

Funkcja $\mathbb{R} \ni \lambda \mapsto W_w(\lambda) \in \mathbb{R}$ jest ciągła oraz $W_w(0) = 2|Q(0)|^2 > 0$, więc również w pewnym otoczeniu zera $W_w(\lambda) > 0$. Stąd $z = 0$ nie jest miejscem globalnego minimum, co należało wykazać.

(ix) Wielomian o współczynnikach rzeczywistych
Niech $P(z)$ będzie wielomianem zmiennej zespolonej o rzeczywistych współczynnikach. Ta szczególna własność wielomianu jest równoważna, jak łatwo sprawdzić, spełnieniu tożsamości $\overline{P(\bar{z})} = P(z)$. Podstawienie do tej tożsamości rozkładu

$$P(z) = a(z - z_1)\dots(z - z_n)$$

daje

$$\bar{a}(z - \overline{z_1})\dots(z - \overline{z_n}) = a(z - z_1)\dots(z - z_n)\,.$$

Stąd a jest liczbą rzeczywistą, a ciągi $(z_1, \dots, z_n)$ i $(\overline{z_1}, \dots, \overline{z_n})$ mogą się różnić tylko porządkiem wyrazów. Wnioskujemy stąd, że pierwiastki można uszeregować w taki ciąg, aby rozkład przyjął postać

$$P(z) = a(z - z_1)(z - \overline{z_1})\dots(z - z_m)(z - \overline{z_m})(z - z_{2m+1})\dots(z - z_n)\,,$$
$$\text{gdzie}\quad z_1, \dots, z_m \in \mathbb{C} \setminus \mathbb{R}\,,\quad z_{2m+1}, \dots, z_n \in \mathbb{R}\,.$$

Niech teraz $P(x)$ będzie wielomianem zmiennej rzeczywistej, o rzeczywistych współczynnikach. Rozszerza się on jednoznacznie do wielomianu zmiennej zespolonej o tych samych współczynnikach, więc jest zacieśnieniem do zmiennej rzeczywistej takiego wielomianu, jaki rozważaliśmy wyżej. Stąd, biorąc pod uwagę rzeczywistość zmiennej x,

$$P(x) = a|x - z_1|^2 \dots |x - z_m|^2 (x - z_{2m+1})\dots(x - z_n)\,.$$

Czynniki $|x - z_i|^2 = (x - \mathfrak{Re}\, z_i)^2 + (\mathfrak{Im}\, z_i)^2$, $\mathfrak{Im}\, z_i \neq 0$, są rzeczywistymi trójmianami kwadratowymi, nie mającymi pierwiastków rzeczywistych (o wyróżniku ujemnym). Dostaliśmy ogólną postać rozkładu wielomianu rzeczywistego na czynniki pierwsze.

(x) Uzwarcona płaszczyzna zespolona
W przestrzeni kartezjańskiej $\mathbb{R}^3$, której punkty są trójkami liczb rzeczywistych (x_1, x_2, x_3), rozważymy następującą konstrukcję geometryczną. Przez punkt $N = (0, 0, 1)$ poprowadzimy wszystkie możliwe proste nierównoległe do płaszczyzny $x_3 = 0$. Każda taka prosta przebija dokładnie jeden raz płaszczyznę $x_3 = 0$ oraz sferę jednostkową

$$S = \{(x_1, x_2, x_3) \mid (x_1)^2 + (x_2)^2 + (x_3)^2 = 1\}$$

– dla tej ostatniej nie liczymy punktu N. Ustala to, jak łatwo się przekonać, wzajemnie jednoznaczną odpowiedniość pomiędzy punktami płaszczyzny, a punktami na sferze S bez jednego punktu N. Płaszczyznę $x_3 = 0$ możemy zinterpretować jak płaszczyznę Gaussa liczb zespolonych $z = x_1 + ix_2$, i dostajemy w ten sposób wzajemnie jednoznaczną odpowiedniość

$$\mathbb{C} \longleftrightarrow S \setminus \{N\}\,.$$

Jest widoczne, że jeśli moduł $|z|$ liczby zespolonej dąży do nieskończoności, to odpowiadający liczbie z punkt na sferze zbliża się do punktu N. W teorii funkcji zmiennej zespolonej okazuje się naturalnym rozszerzenie zbioru liczb zespolonych $\mathbb{C}$ o jeden element, oznaczany ∞, i rozszerzenie powyższej odpowiedniości do

$$\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\} \longleftrightarrow S\,, \quad \text{gdzie} \quad \infty \longleftrightarrow N\,.$$

Zbiór $\overline{\mathbb{C}}$ nazywamy ***uzwarconą płaszczyzną zespoloną***. Rozważmy wyniki działań $z_1 + z_2$, $z_1 z_2$, $\dfrac{z_1}{z_2}$. Przez odpowiednie przejście do granicy z modułem jednej lub dwóch liczb do zera lub nieskończoności przyjmuje się naturalne rozszerzenia działań:

$$\begin{aligned}
\text{dla } z \in \mathbb{C}: &\quad z + \infty = \infty + z = \infty\,, \quad \frac{z}{\infty} = 0\,;\\
\text{dla } z \in \overline{\mathbb{C}} \setminus \{0\}: &\quad \frac{z}{0} = \infty\,, \quad z \cdot \infty = \infty \cdot z = \infty\,;\\
\text{nieokreślone symbole}: &\quad \frac{0}{0}\,, \quad \frac{\infty}{\infty}\,, \quad 0 \cdot \infty\,, \quad \infty + \infty\,.
\end{aligned}$$

Cztery nieokreślone symbole pozostają takimi, gdyż w tych przypadkach wynik podwójnego przejścia granicznego zależy od sposobu jego wykonania.

Odpowiedniość pomiędzy liczbami zespolonymi $z \in \mathbb{C}$ a punktami na sferze $P \in S \setminus \{N\}$ można prosto wyrazić analitycznie. Wprowadźmy w przestrzeni $\mathbb{R}^3$ współrzędne sferyczne (r, ϑ, φ) za pomocą przepisu

$$\begin{gathered}
x_1 = r \sin\vartheta \cos\varphi\,, \quad x_2 = r \sin\vartheta \sin\varphi\,, \quad x_3 = r\cos\vartheta\,,\\
(r, \vartheta, \varphi) \in \langle 0, +\infty) \times \langle 0, \pi\rangle \times \langle 0, 2\pi)\,.
\end{gathered}$$

Elementarne rozważania geometryczne pokazują wtedy, że określona powyżej odpowiedniość może być zapisana analitycznie jako

$$\begin{gathered}
P = (\sin\vartheta\, \cos\varphi, \sin\vartheta\, \sin\varphi, \cos\vartheta) \longleftrightarrow z = \operatorname{ctg}(\vartheta/2)\, e^{i\varphi}\,,\\
(\vartheta, \varphi) \in (0, \pi\rangle \times \langle 0, 2\pi)\,.
\end{gathered}$$

Dla $\vartheta = 0$ dostajemy na sferze punkt N, który przyporządkowaliśmy elementowi ∞.

§5 Grupy odwzorowań. Permutacje

1 Grupy symetryczne

Niech X będzie dowolnym zbiorem. Rozważmy zbiór wszystkich bijekcji zbioru X na samego siebie. Składanie odwzorowań określa działanie wewnętrzne w tym zbiorze: parze bijekcji f i g zbioru X na X przypisujemy ich złożenie,

$$(g, f) \mapsto g \circ f\,,$$

które też jest bijekcją X na X. Z ogólnych własności składania odwzorowań działanie to jest łączne, posiada element neutralny równy odwzorowaniu identycznościowemu, a każdy element zbioru bijekcji posiada element odwrotny względem tego działania równy odwzorowaniu odwrotnemu, jak je określiliśmy wyżej. Stwierdzenia te składają się na treść następującego twierdzenia.

Twierdzenie 1. *Zbiór bijekcji dowolnego zbioru X na siebie tworzy grupę z działaniem składania odwzorowań.*

Grupę, o której mówi twierdzenie, nazywamy ***grupą symetryczną zbioru*** X i oznaczamy S_X.

Struktury grupowe grup symetrycznych zbiorów o tej samej mocy można utożsamić, mówi o tym następne twierdzenie.

Twierdzenie 2. *Jeśli moc zbioru X jest równa mocy zbioru Y, to grupy S_X i S_Y są izomorficzne. Każda bijekcja $j : X \mapsto Y$ zadaje izomorfizm*

$$h : S_X \mapsto S_Y\,, \;\; f \mapsto h[f] = j \circ f \circ j^{-1}\,.$$

Dowód. Zgodnie z definicją równej mocy istnieje bijekcja j, o której mowa w tezie, oraz odwzorowania: h określone w tezie oraz

$$h' : S_Y \mapsto S_X\,, \;\; g \mapsto h'[g] = j^{-1} \circ g \circ j\,.$$

Widzimy, że dla każdych: $g \in S_Y$ i $f \in S_X$ jest

$$h[h'[g]] = g \;\text{ i }\; h'[h[f]] = f\,, \quad \text{więc} \quad h \circ h' = \mathrm{id}_{S_Y} \;\text{ i }\; h' \circ h = \mathrm{id}_{S_X}\,.$$

Twierdzenie 6, §2, mówi teraz, że h jest bijekcją. Mamy również

$$h[f_2] \circ h[f_1] = j \circ f_1 \circ f_2 \circ j^{-1} = h[f_2 \circ f_1]\,,$$

więc h jest homomorfizmem. Otrzymaliśmy izomorfizm S_X na S_Y. □

Istnienie izomorfizmu, o którym mówi powyższe twierdzenie, nie oznacza bynajmniej, że byłoby wskazanym ograniczenie się tylko do jednego zbioru danej mocy. Rzecz w tym, że dodatkowe struktury istniejące w zbiorze X mogą w naturalny sposób wyznaczać podgrupy grupy S_X. Następny punkt zawiera ilustrację tego stwierdzenia.

2 Grupy automorfizmów grupowych

Rozważmy teraz bardziej szczególną sytuację, gdy zbiór X ma strukturę grupy, oznaczamy go wtedy G. Jak w ogólnym przypadku, tak i teraz mamy grupę symetryczną, S_G. Wiemy jednak, że w przypadku grupy pewne jej odwzorowania na samą siebie są szczególnie ciekawe, mianowicie określone wcześniej automorfizmy grupy G. Zbiór wszystkich automorfizmów grupy G oznaczamy przez $\operatorname{Aut} G \subseteq S_G$. Kładąc w twierdzeniu 12, §3, $G'' = G' = G$ oraz stosując kryterium 6, §3, otrzymujemy natychmiast następujące twierdzenie.

Twierdzenie 3. *Zbiór automorfizmów* $\operatorname{Aut} G$ *tworzy podgrupę grupy symetrycznej* S_G.

Wśród wszystkich automorfizmów wyróżnić możemy ważną podklasę. Dla każdego $g \in G$ określamy odwzorowanie

$$h_g : G \mapsto G\,, \ g' \mapsto h_g(g') = gg'g^{-1}\,.$$

Mamy $h_g(g'g'') = gg'g^{-1}gg''g^{-1} = h_g(g')h_g(g'')$, więc wszystkie h_g są homomorfizmami. Ponadto łatwo sprawdzić, że

$$h_g \circ h_{g^{-1}} = h_{g^{-1}} \circ h_g = \mathrm{id} \quad \text{oraz} \quad h_{g_1} \circ h_{g_2} = h_{g_1 g_2}\,.$$

Stąd h_g są bijekcjami, więc automorfizmami, a ich złożenia i odwrotne są w tej samej klasie. Automorfizmy tego typu nazywamy ***automorfizmami wewnętrznymi***. Zbiór automorfizmów wewnętrznych grupy G oznaczamy $\operatorname{Inn} G$. Na podstawie powyższej dyskusji dostajemy w oparciu o kryterium 6, §3, następujący wynik.

Twierdzenie 4. *Automorfizmy wewnętrzne* $\operatorname{Inn} G$ *tworzą podgrupę grupy* $\operatorname{Aut} G$.

3 Przykłady

(i) Homografie zespolone

Dla każdej czwórki liczb $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ takich, że $\alpha\delta - \beta\gamma \neq 0$, zadajemy odwzorowanie $h : \overline{\mathbb{C}} \mapsto \overline{\mathbb{C}}$, zwane ***homografią***, (definicja uzwarconej płaszczyzny zespolonej $\overline{\mathbb{C}}$ w przykładzie (x), p. 9, §4) w następujący sposób:

$$\text{dla } z \in \mathbb{C}: \quad h(z) = \frac{\alpha z + \beta}{\gamma z + \delta}\,, \qquad h(\infty) = \frac{\alpha}{\gamma}\,.$$

Pierwsza z tych formuł obejmuje również przypadek zerowania się mianownika: licznik jest wtedy różny od zera (w przeciwnym razie byłoby $\alpha\delta - \beta\gamma = 0$), więc zgodnie z regułami działań w $\overline{\mathbb{C}}$ wartość homografii jest równa ∞. Homografia h wyznacza swoje współczynniki $\alpha, \beta, \gamma, \delta$ z dokładnością do wspólnego czynnika.

Wykazuje się bezpośrednim rachunkiem, że jeśli homografia h jest zadana współczynnikami $\alpha, \beta, \gamma, \delta$, a homografia h' – współczynnikami $\alpha', \beta', \gamma', \delta'$, to złożenie $h' \circ h$ jest również homografią, zadaną współczynnikami

$$\alpha'\alpha + \beta'\gamma\,, \quad \alpha'\beta + \beta'\delta\,, \quad \gamma'\alpha + \delta'\gamma\,, \quad \gamma'\beta + \delta'\delta\,.$$

Odwzorowanie identycznościowe jest homografią o współczynnikach $1, 0, 0, 1$, a każda homografia ma odwzorowanie odwrotne, będące homografią o współczynnikach

$$\frac{\delta}{\alpha\delta - \beta\gamma}\,, \quad \frac{-\beta}{\alpha\delta - \beta\gamma}\,, \quad \frac{-\gamma}{\alpha\delta - \beta\gamma}\,, \quad \frac{\alpha}{\alpha\delta - \beta\gamma}\,.$$

Homografie tworzą więc podgrupę grupy symetrycznej $S_{\overline{\mathbb{C}}}$.

(ii) Automorfizmy grupy cyklicznej
Jeśli grupa G jest abelowa, to grupa jej automorfizmów wewnętrznych jest trywialna. Otrzymamy pełną charakterystykę grupy automorfizmów grupy cyklicznej G. Niech

$$G = \{e, g, \ldots, g^{n-1}\} \quad \text{oraz} \quad h(g) = g^k \equiv g'\,, \quad 1 \leq k \leq n-1\,.$$

Jeśli h ma być endomorfizmem, to musimy mieć $h(g^l) = g'^l$. Z dyskusji przykładu (vii), punkt 3, §3, wynika więc, że h jest automorfizmem wtedy, i tylko wtedy, gdy liczby k i n są względnie pierwsze. Grupa $\operatorname{Aut} G$ ma tyle elementów, ile jest liczb $k \in \{1, \ldots, n-1\}$ pierwszych względem liczby n; liczbę tę oznacza się $\varphi(n)$, a funkcję $\mathbb{N} \ni n \mapsto \varphi(n) \in \mathbb{N}$ nazywa się ***funkcją*** φ ***Eulera*** – nie mylić z funkcją Γ Eulera.

(iii) Automorfizm iloczynu prostego grup
Niech G będzie dowolną grupą. Utwórzmy iloczyn prosty $G \times G$ (patrz przykład (ix), p. 3, §3). Odwzorowanie

$$G \times G \ni (g, h) \mapsto (h, g) \in G \times G$$

jest automorfizmem, ale nie jest automorfizmem wewnętrznym. Grupa automorfizmów wewnętrznych $\mathrm{Inn}_{G\times G}$ jest izomorficzna z grupą $\mathrm{Inn}_G \times \mathrm{Inn}_G$.

4 Grupy permutacji

Jeśli zbiór X jest skończony, to jego grupę symetryczną S_X nazywamy ***grupą permutacji*** jego elementów. Jeśli liczba elementów zbioru X jest równa n, to grupa S_X jest izomorficzna z grupą symetryczną zbioru liczb $\{1, \ldots, n\}$, którą oznaczamy S_n. Bijekcje zbioru $\{1, \ldots, n\}$ na X notujemy w tym przypadku jako $\{1, \ldots, n\} \ni i \mapsto a_i \in X$ i nazywać będziemy ***numeracjami elementów zbioru*** X. Jeśli $\pi \in S_n$ to izomorfizm zadany numeracją a daje permutację

$$p = a \circ \pi \circ a^{-1}\,, \quad \text{czyli} \quad p(a_i) = a_{\pi(i)}\,.$$

Permutacje $\pi \in S_n$ notujemy często za pomocą symbolu $\left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{smallmatrix}\right)$. Jest widoczne, że permutacje te pozostają we wzajemnie jednoznacznej odpowiedniości z ciągami liczb w drugim wierszu tego symbolu. Stąd:

Twierdzenie 5. *Liczba elementów zbioru S_n jest równa $n!$.*

Jako wniosek otrzymujemy w tym przypadku wynikanie odwrotne do zawartego w tezie twierdzenia 2.

Wniosek 6. *Niech X i Y będą skończonymi zbiorami. Wtedy S_X jest izomorficzna z S_Y $\Longleftrightarrow$ zbiory X i Y są równoliczne.*

Niech X' będzie podzbiorem zbioru X i oznaczmy $X'' = X \setminus X'$. Łatwo przekonać się, że te permutacje zbioru X, których zacieśnienie do zbioru X'' jest identycznością, tworzą podgrupę grupy S_X, oznaczmy ją $\tilde{S}_{X'}$. Zacieśnienie każdej z tych permutacji do zbioru X' daje permutację zbioru X' i ustala izomorfizm grup $\tilde{S}_{X'}$ i $S_{X'}$. Składanie permutacji jest na ogół nieprzemienne, ale jeśli $p' \in \tilde{S}_{X'}$, a $p'' \in \tilde{S}_{X''}$, to $p' \circ p'' = p'' \circ p'$. Takie dwie permutacje nazywamy ***niezależnymi***.

Jeśli elementy zbioru X ponumerujemy, $X = \{a_1 \ldots, a_n\}$, to uogólniając notację wprowadzoną dla grupy S_n, każdą permutację $p \in S_X$ możemy zapisać w formie $\left(\begin{smallmatrix} a_1 & \cdots & a_n \\ p(a_1) & \cdots & p(a_n) \end{smallmatrix}\right)$. Całe kolumny w tym symbolu można przestawiać bez wpływu na odwzorowanie.

5 Permutacje cykliczne. Transpozycje

Niech X będzie zbiorem n-elementowym, a p jego permutacją różną od identyczności. Wybierzmy element $a \in X$ taki, że $p(a) \neq a$. Utwórzmy następnie ciąg

$$(a, p(a), p^2(a), \ldots, p^{m-1}(a))$$

taki, że wszystkie jego elementy są różne, ale $p^m(a)$ pokrywa się z jednym z nich. Ciąg taki zawsze istnieje, bo X jest skończony oraz z założenia ma co najmniej dwa wyrazy. Ponadto z konstrukcji ciągu wynika, że $p^m(a) = a$. Rzeczywiście, gdyby $p^m(a) = p^l(a)$ dla pewnego $l \in \{1, \ldots, m-1\}$, to mielibyśmy stąd $p^{m-l}(a) = a$, co przeczyłoby założeniom konstrukcji ciągu. Oznaczmy

$$X' = \{a, p(a), \ldots, p^{m-1}(a)\} \quad \text{i} \quad X'' = X \setminus X'\,.$$

Z konstrukcji wynika, że $p(X') = X'$, więc z bijektywności p mamy również $p(X'') = X''$. Rozpatrzymy teraz dwie możliwości.

(i) Jeśli p zacieśniona do X'' jest identycznością, to permutację p nazywamy ***permutacją cykliczną (cyklem)***. Jest widoczne, że jeśli b jest dowolnym wyrazem ciągu $(a, p(a), p^2(a), \dots, p^{m-1}(a))$, to ciąg $(b, p(b), p^2(b), \dots, p^{m-1}(b))$ jest cyklicznym przestawieniem ciągu $(a, p(a), p^2(a), \dots, p^{m-1}(a))$ i $p^m(b) = b$. Permutację cykliczną symbolicznie oznaczamy dowolnym z tych ciągów.

(ii) Jeśli $p|_{X''}$ nie jest identycznością, to niech c_1 będzie takim rozszerzeniem $p|_{X'}$, które jest identycznością na X'', a p_1 - takim rozszerzeniem $p|_{X''}$, które jest identycznością na X'. Widać z konstrukcji, że permutacje c_1 i p_1 są niezależne, c_1 jest cyklem, $c_1 = (a, p(a), \dots, p^{m-1}(a))$ oraz

$$p = c_1 \circ p_1 = p_1 \circ c_1 \,.$$

Do permutacji p_1 możemy teraz znów zastosować całą procedurę opisaną w tym punkcie, otrzymując w wyniku rozkład na niezależne permutacje $p = c_1 \circ c_2 \circ p_2$. Po skończonej liczbie kroków albo wszystkie elementy zbioru X znajdą się w cyklach $c_1, \dots, c_k$, albo osiągniemy sytuację, w której $p_k = \mathrm{id}_X$. W obu przypadkach uzyskujemy rozkład, o którym mówi następujące twierdzenie.

Twierdzenie 7. *Każda permutacja skończonego zbioru X ma jednoznaczny rozkład na niezależne cykle,*

$$p = c_1 \circ c_2 \circ \dots \circ c_k \,.$$

Niech m_i oznacza długość cyklu c_i oraz niech $M =$ najmniejsza wspólna wielokrotność liczb *$\{m_1, \dots, m_k\}$. Wtedy M jest najmniejszą liczbą, dla której $p^M = \mathrm{id}$, i permutacja p generuje cykliczną podgrupę*

$$\{\mathrm{id}, p, p^2, \dots, p^{M-1}\} \subseteq S_X \,.$$

Dowód. Pewien rozkład skonstruowaliśmy w dyskusji poprzedzającej twierdzenie; należy teraz pokazać, że jest on jedyny. Niech $p = c_1 \circ \dots \circ c_k$ będzie dowolnym rozkładem. Elementy zbioru X wchodzące w obręb cykli są zadane jednoznacznie jako te, które transformują się nietrywialnie pod działaniem permutacji p. Jak widzieliśmy wyżej, każdy taki element zadaje jednoznacznie cykl, w skład którego wchodzi, skąd jednoznaczność rozkładu.

Jeśli c jest cyklem o długości m, to na mocy konstrukcji cyklu $c^m = \mathrm{id}$ i m jest najmniejszą liczbą, dla której to zachodzi. W przypadku ogólnym z przemienności cykli w rozkładzie mamy $p^l = c_1^l \circ \dots \circ c_k^l$. Z niezależności permutacji c_i^l otrzymujemy równoważność

$$p^l = \mathrm{id} \quad \Longleftrightarrow \quad \forall i \in \{1, \dots, k\} : l \text{ dzieli się przez } m_i \,,$$

więc M określona w tezie jest najmniejszą z takich liczb l. Podgrupa cykliczna jest generowana tak, jak to opisaliśmy w §3 (przykłady (vii), p. 3 i (iii), p. 6). □

Dla każdego cyklu $c \in S_X$ można wybrać taką numerację zbioru X, przy której cykl ten ma szczególnie prosty zapis. Niech

$$c = (a, c(a), \ldots, c^{m-1}(a)) .$$

Wybierzmy numerację zadaną przez $a_i = c^{i-1}(a)$ dla $i = 1, \ldots, m$, a pozostałe elementy zbioru X ponumerujmy w dowolny sposób jako $a_{m+1}, \ldots, a_n$. Wtedy

$$c = (a_1, \ldots, a_m) \equiv \begin{pmatrix} a_1 & a_2 & \ldots & a_{m-1} & a_m & a_{m+1} & \ldots & a_n \\ a_2 & a_3 & \ldots & a_m & a_1 & a_{m+1} & \ldots & a_n \end{pmatrix} .$$

Łatwo sprawdzić, że ten sam fakt można zapisać jako

$$c = a \circ (1, \ldots, m) \circ a^{-1} ,$$

czyli w izomorfizmie S_n i S_X zadanym przez numerację a cykl c jest obrazem cyklu $(1, \ldots, m)$. Prowadząc analogiczne rozumowanie dla dowolnej numeracji pokazuje się, że dla każdej z nich cykl c jest obrazem m-wyrazowego cyklu z grupy S_n.

Każdy cykl o dwóch wyrazach, $(a_1, a_2) \equiv \begin{pmatrix} a_1 & a_2 & a_3 & \ldots & a_n \\ a_2 & a_1 & a_3 & \ldots & a_n \end{pmatrix}$, nazywamy ***transpozycją***. Każdy cykl można rozłożyć zgodnie z następującym wzorem

$$(a_1, \ldots, a_m) = (a_1, a_m) \circ (a_1, \ldots, a_{m-1}) ,$$

co łatwo sprawdzić wyliczając złożenie permutacji po prawej stronie wzoru. Stąd metodą indukcji względem długości cyklu dostaje się rozkład

$$(a_1, \ldots, a_m) = (a_1, a_m) \circ (a_1, a_{m-1}) \circ \ldots \circ (a_1, a_2) .$$

Czynniki tego rozkładu nie są niezależne; łatwo też sprawdzić, że *żaden* rozkład cyklu nie może rozpadać się na permutacje niezależne. Rozkład cyklu na transpozycje w połączeniu z poprzednim twierdzeniem daje następujący wynik.

Twierdzenie 8. *Każdą permutację można rozłożyć na transpozycje.*

Rozkład, o którym mówi twierdzenie nie jest jednoznaczny, np. $\mathrm{id} = (a, b) \circ (a, b)$; z ostatniej równości dla każdej transpozycji (a, b) mamy $(a, b)^{-1} = (a, b)$.

6 Inwersje ciągów liczbowych

Niech $(a_1, \ldots, a_n)$ będzie dowolnym różnowartościowym ciągiem liczb naturalnych (ogólniej: można dopuścić liczby rzeczywiste). Mówimy, że para liczb w tym ciągu $\{a_j, a_k\}$ tworzy ***inwersję***, gdy $j < k$ i $a_j > a_k$. ***Liczbą inwersji*** w ciągu $(a_1, \ldots, a_n)$ nazywamy liczbę par elementów tego ciągu, które tworzą inwersję. Następująca własność inwersji okaże się przydatna w następnym punkcie:

Lemat 9. *Zamiana miejscami dwóch wyrazów różnowartościowego ciągu liczb naturalnych (lub rzeczywistych) prowadzi do zmiany liczby inwersji o liczbę nieparzystą.*

Dowód. Operacja, o której mowa w twierdzeniu, może być zapisana jako

$$(s_1, a, s_2, b, s_3) \to (s_1, b, s_2, a, s_3)\,,$$

gdzie a, b są dwoma wyrazami ciągu, a s_1, s_2 i s_3 są jego podciągami. Jest widoczne, że porządek w ciągu zmieniają pary następujących, i tylko tych, typów: $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, gdzie w drugim i trzecim przypadku c przebiega wszystkie wyrazy podciągu s_2. Jeśli więc liczba wyrazów podciągu s_2 wynosi m, to liczba par zmieniających porządek jest równa $2m + 1$. Każda z tych zmian porządku prowadzi do zmiany liczby inwersji o $+1$ lub -1, skąd teza. □

7 Znak i parzystość permutacji

Jak widzieliśmy, rozkład permutacji na transpozycje nie jest jednoznaczny, a nawet liczba transpozycji w różnych rozkładach może być różna. Istnieje jednak cecha wspólna rozkładów danej permutacji.

Twierdzenie 10. *Niech X będzie zbiorem skończonym, p – dowolną permutacją z S_X, i niech*

$$p = t_s \circ t_{s-1} \circ \ldots \circ t_1$$

będzie jej rozkładem na transpozycje. Wtedy liczba transpozycji w dowolnym innym rozkładzie p różni się od s o liczbę parzystą, więc liczba

$$\operatorname{sgn} p := (-1)^s$$

nie zależy od wyboru rozkładu.

Dowód. Wybierając dowolną numerację a zbioru X znajdujemy permutację $\pi = a^{-1} \circ p \circ a$ i transpozycje $\sigma_i = a^{-1} \circ t_i \circ a$. Z założenia twierdzenia mamy $\pi = \sigma_s \circ \ldots \circ \sigma_1$. Twierdzenie wystarczy więc wykazać dla grupy S_n. Permutacja π jest jednoznacznie zadana ciągiem $(\pi(1), \ldots, \pi(n))$. Oznaczmy liczbę inwersji w tym ciągu przez I_π. Ciąg ten powstaje z ciągu $(1,\ldots,n)$ przez poddanie go kolejno działaniu transpozycji $\sigma_1,\ldots,\sigma_s$. Po każdej takiej operacji dwa wyrazy w ciągu zamieniają się miejscami, a pozostałe zachowują swoje pozycje. Każda z tych operacji prowadzi więc do nieparzystej zmiany liczby inwersji w przekształcanym ciągu. Dodając te liczby dla operacji od 1 do s dostajemy liczbę postaci: liczba parzysta $+\ s\cdot 1$. Liczba inwersji w początkowym ciągu $(1,\ldots,n)$ wynosi zero, więc I_π ma taką parzystość jak liczba s (tj. jest parzysta wtedy, i tylko wtedy, gdy s jest parzysta). Liczba I_π nie zależy od rozkładu na transpozycje, skąd teza. $\square$

Funkcję $S_X \ni p \mapsto \operatorname{sgn} p \in \{1,-1\}$ określoną ostatnim twierdzeniem nazywa się ***znakiem permutacji***. Ma ona interesującą własność wynikającą natychmiast z pomnożenia rozkładów na transpozycje dwóch permutacji.

Twierdzenie 11. *Znak permutacji jest homomorfizmem grupy S_X na multiplikatywną grupę $\{1,-1\}$, tj.*

$$\operatorname{sgn}(p_1\circ p_2) = \operatorname{sgn} p_1\ \operatorname{sgn} p_2\,.$$

Jądro tego homomorfizmu (czyli zbiór permutacji o dodatnim znaku), które na mocy twierdzenia 10, §3, jest podgrupą niezmienniczą grupy S_X, nazywamy ***grupą alternującą zbioru*** X i oznaczamy A_X.

8 Przykłady

(i) Składanie i odwracanie permutacji

$$\begin{pmatrix}1&2&3&4&5\\3&1&5&4&2\end{pmatrix}\circ\begin{pmatrix}1&2&3&4&5\\2&4&1&5&3\end{pmatrix}=\begin{pmatrix}1&2&3&4&5\\1&4&3&2&5\end{pmatrix}=(2,4)\,,$$

$$\begin{pmatrix}1&2&3&4&5&6\\3&5&2&6&1&4\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3&4&5&6\\5&3&1&6&2&4\end{pmatrix}.$$

(ii) Liczba inwersji a znak permutacji

Dowód twierdzenia 10 pokazuje, że znak permutacji liczbowej π można ustalić znając parzystość liczby inwersji w ciągu $(\pi(1),\ldots,\pi(n))$. Mając daną permutację

$$\pi=\begin{pmatrix}1&2&3&4&5&6&7&8&9\\9&5&8&3&7&2&4&1&6\end{pmatrix}$$

liczymy więc liczbę inwersji I w ciągu $(9,5,8,3,7,2,4,1,6)$. Wystarczy policzyć dla każdej kolejnej liczby w tym ciągu, ile liczb mniejszych od niej następuje po niej. Dostajemy $I = 8+4+6+2+4+1+1+0 =$ liczba parzysta. Stąd $\operatorname{sgn}\pi = (-1)^I = 1$.

(iii) Rozkład permutacji na cykle rozłączne
Permutacja $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 8 & 1 & 7 & 3 & 2 & 5 & 6 \end{pmatrix}$ ma rozkłady na cykle rozłączne i na transpozycje (znak złożenia między permutacjami często pomijamy):

$$\pi = (1,4,7,5,3)(2,8,6) = (1,3)(1,5)(1,7)(1,4)(2,6)(2,8)\,.$$

Rozkład na transpozycje ma 6 czynników, więc znak tej permutacji jest równy $\operatorname{sgn}\pi = (-1)^6 = 1$. Mamy $(1,4,7,5,3)^5 = \mathrm{id}$ i $(2,8,6)^3 = \mathrm{id}$, więc $\pi^{15} = \mathrm{id}$, i jest to najniższa potęga, dla której ta równość jest prawdziwa. Zbiór

$$\{\mathrm{id}, \pi, \ldots, \pi^{14}\}$$

jest cykliczną podgrupą grupy S_8. Dla dowolnej potęgi mamy

$$\pi^k = (1,4,7,5,3)^{k \bmod 5}\,(2,8,6)^{k \bmod 3}\,,$$
$$\text{na przykład} \qquad \pi^{71} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 1 & 7 & 3 & 8 & 5 & 2 \end{pmatrix}.$$

(iv) Grupa S_3
Grupa ta ma 6 elementów:

$$\mathrm{id}\,,\ (1,2,3)\,,\ (1,3,2)\,,\ (1,2)\,,\ (1,3)\,,\ (2,3)\,.$$

Pierwsze trzy elementy mają dodatni znak i składają się na podgrupę A_3.

(v) Wewnętrzne automorfizmy grupy S_3
Odwzorowanie

$$S_3 \ni \pi \mapsto h_\sigma(\pi) = \sigma^{-1} \circ \pi \circ \sigma \in S_3$$

jest automorfizmem wewnętrznym. Na przykład dla $\sigma = (1,3,2)$:

$$\mathrm{id} \mapsto \mathrm{id}\,,\ (1,2,3) \mapsto (1,2,3)\,,\ (1,3,2) \mapsto (1,3,2)\,,$$
$$(1,2) \mapsto (2,3)\,,\ (1,3) \mapsto (1,2)\,,\ (2,3) \mapsto (1,3)\,.$$

§6 Macierze

1 Definicja macierzy

Niech $\mathbb{K}$ będzie ciałem. ***Macierzą o*** m ***wierszach i*** n ***kolumnach i wartościach w*** $\mathbb{K}$ (krótko: ***macierzą*** $m \times n$) nazywamy każde odwzorowanie $A : \{1, \dots, m\} \times \{1, \dots, n\} \mapsto \mathbb{K}\,,\ (i, j) \mapsto A_{ij}$; notujemy je w formie tablicy

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} .$$

Wyrazy A_{ij} nazywamy ***elementami macierzy*** A. Stosuje się też niekiedy notację, w której elementy oznacza się innym symbolem niż symbol macierzy, na ogół małą czcionką tej samej litery: $A_{ij} = a_{ij}$. Piszemy też $A = (A_{ij}) = (a_{ij})$. Równość dwóch macierzy takiego samego wymiaru (tj. o wspólnej charakterystyce $m \times n$), $A = B$, oznacza równość ich wszystkich odpowiednich elementów, $A_{ij} = B_{ij}$ $(i = 1, \dots, m,\ j = 1, \dots, n)$.

Określiliśmy wyżej macierz jako odwzorowanie. Istnieje też alternatywny, użyteczny sposób ujmowania macierzy. Ciąg

$$A_{i*} = \begin{pmatrix} A_{i1} & A_{i2} & \dots & A_{in} \end{pmatrix} , \quad i \in \{1, \dots, m\} ,$$

nazwiemy ***i-tym wierszem macierzy*** A, a ciąg

$$A_{*j} = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{pmatrix} , \quad j \in \{1, \dots, n\} ,$$

nazwiemy ***j-tą kolumną macierzy*** A. Macierz A możemy teraz potraktować jak:

(i) kolumnę wierszy A_{i*}, a więc element zbioru $\underbrace{\mathbb{K}^n \times \dots \times \mathbb{K}^n}_{m \text{ razy}}$ (elementy $\mathbb{K}^n$ to wiersze):

$$A = \begin{pmatrix} A_{1*} \\ A_{2*} \\ \vdots \\ A_{m*} \end{pmatrix} ;$$

(ii) wiersz kolumn A_{*j}, a więc element zbioru $\underbrace{\mathbb{K}^m \times \dots \times \mathbb{K}^m}_{n \text{ razy}}$ (elementy $\mathbb{K}^m$ to kolumny):

$$A = \begin{pmatrix} A_{*1} & A_{*2} & \dots & A_{*n} \end{pmatrix} .$$

Zbiór macierzy $m \times n$ o elementach w ciele $\mathbb{K}$ oznaczać będziemy $\mathbb{K}^{m\times n}$. Macierze $n \times n$ nazywamy ***kwadratowymi***. ***Diagonalą macierzy kwadratowej*** A nazywamy ciąg elementów $(A_{11}, A_{22}, \ldots, A_{nn})$, a elementy tego ciągu nazywamy ***elementami diagonalnymi*** macierzy A. Macierz kwadratową, której wszystkie pozadiagonalne elementy są równe zeru, nazywamy ***macierzą diagonalną***, a taką, której wszystkie elementy poniżej diagonali są równe zeru – ***macierzą trójkątną***.

2 Transpozycja i sprzężenie

Określamy trzy operacje na macierzach. ***Transpozycją*** macierzy nazywamy odwzorowanie zadane przez

$$\mathbb{K}^{m\times n} \ni A \mapsto A^{\mathrm{T}} \in \mathbb{K}^{n\times m}\,, \quad (A^{\mathrm{T}})_{ij} = A_{ji}\,.$$

Macierz A^{T} nazywamy ***macierzą transponowaną (do macierzy A)***. Definicję transpozycji można ująć mówiąc, że macierz A^{T} powstaje przez zamianę wierszy macierzy A na kolumny. Druga operacja dotyczy macierzy o elementach w ciele liczb zespolonych $\mathbb{C}$. ***Zespolonym sprzężeniem*** macierzy $A \in \mathbb{C}^{m\times n}$ nazywamy macierz $\overline{A} \in \mathbb{C}^{m\times n}$ o elementach będących liczbami sprzężonymi do odpowiednich elementów macierzy A: $(\overline{A})_{ij} = \overline{A_{ij}}$. Dla macierzy zespolonych złożenie operacji transpozycji i sprzężenia zespolonego (w dowolnej kolejności) prowadzi do operacji ***sprzężenia hermitowskiego***: dla danej macierzy A ***macierz sprzężoną do niej po hermitowsku*** $A^\dagger$ zadajemy przez

$$A^\dagger = \overline{A}^{\mathrm{T}} = \overline{A^{\mathrm{T}}}\,,$$

czyli

$$\mathbb{C}^{m\times n} \ni A \mapsto A^\dagger \in \mathbb{C}^{n\times m}\,, \quad (A^\dagger)_{ij} = \overline{A_{ji}}\,.$$

Łatwo sprawdzić (dowód pozostawiamy czytelnikowi), że podwójne złożenie każdej z trzech tak określonych operacji jest operacją identycznościową:

$$(A^{\mathrm{T}})^{\mathrm{T}} = A\,, \quad \overline{\overline{A}} = A\,, \quad \left(A^\dagger\right)^\dagger = A\,.$$

3 Działania na macierzach

Zbiór macierzy ustalonego wymiaru $m \times n$, o elementach w ciele $\mathbb{K}$, jest szczególnym przypadkiem zbioru funkcji na zadanym zbiorze. Z ogólnych rozważań punktu 12, §3, wynika więc, że na $\mathbb{K}^{m\times n}$ istnieje w naturalny sposób określona ***struktura liniowa***: dla każdych $\lambda, \mu \in \mathbb{K}$ i $A, B \in \mathbb{K}^{m\times n}$ zadajemy macierz $\lambda A + \mu B$

$$(\lambda A + \mu B)_{ij} := \lambda A_{ij} + \mu B_{ij}\,.$$

Ogólne własności takich struktur, omówione w punkcie 12, §3, mają zastosowanie w tym szczególnym przypadku. Dla każdego wymiaru macierzowego $m \times n$ otrzymujemy odrębne struktury.

Rozważmy teraz zbiór wszystkich macierzy. Wprowadzamy w nim nowe działanie wewnętrzne – mnożenie – o tyle ogólniejsze od działań dotychczas przez nas rozważanych, że nie każde dwie macierze można przez siebie pomnożyć: pierwsza macierz musi mieć tyle samo kolumn co druga wierszy. Dla każdych liczb naturalnych m, n, p definiujemy ***mnożenie macierzy*** w następujący sposób:

$$\mathbb{K}^{m\times n} \times \mathbb{K}^{n\times p} \mapsto \mathbb{K}^{m\times p}\,, \quad (A, B) \mapsto AB\,,$$
$$(AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}\,.$$

Mówimy, że element (i, j)-ty macierzy AB jest iloczynem i-tego wiersza macierzy A i j-tej kolumny macierzy B, przez co rozumiemy sumę iloczynów kolejnych elementów wiersza i kolumny.

Twierdzenie 1 (Łączność mnożenia macierzy)**.** *Jeśli* $A \in \mathbb{K}^{m\times n}$*,* $B \in \mathbb{K}^{n\times p}$*,* $C \in \mathbb{K}^{p\times r}$*, to*
$$(AB)C = A(BC)\,.$$

Dowód. Wykorzystując łączność mnożenia, przemienność i łączność dodawania oraz prawo rozdzielności w ciele $\mathbb{K}$ dostajemy ciąg równości

$$[(AB)C]_{ik} = \sum_{l=1}^{p} (AB)_{il} C_{lk} = \sum_{j=1}^{n} \sum_{l=1}^{p} A_{ij} B_{jl} C_{lk} =$$
$$= \sum_{j=1}^{n} A_{ij} (BC)_{jk} = [A(BC)]_{ik}\,,$$

co kończy dowód łączności. □

Jak zauważyliśmy wyżej, mnożenie macierzy nie jest działaniem typu dyskutowanego w §3. Własności tego mnożenia przypominają raczej strukturę składania odwzorowań: jeśli mamy odwzorowanie prowadzące z X do Y oraz inne odwzorowanie prowadzące z Y do Z, to przez ich złożenie otrzymujemy odwzorowanie prowadzące z X do Z. Analogicznie, patrząc na iloczyn macierzy od prawej strony: pomnożenie macierzy o p kolumnach i n wierszach przez macierz o n kolumnach i m wierszach daje macierz o p kolumnach i m wierszach. Jak zobaczymy później, to podobieństwo nie jest przypadkowe. Analogia rozciąga się dalej na istnienie elementów neutralnych. Jak wiemy, złożenie odwzorowania prowadzącego z X do Y z identycznością w X od prawej, lub z identycznością w Y od lewej strony, nie zmienia tego odwzorowania. Macierzami analogicznymi do identycznościowych odwzorowań są macierze jednostkowe. Wprowadźmy najpierw dla każdego naturalnego n odwzorowanie:

$$\{1,\ldots,n\}\times\{1,\ldots,n\}\ni(i,j)\mapsto\delta_{ij}=\begin{cases}1 & \text{gdy } i=j\,,\\ 0 & \text{gdy } i\neq j\,.\end{cases}$$

Nazywamy to odwzorowanie ***symbolem Kroneckera***. Jeśli dany jest dowolny ciąg liczb $(a_1,\ldots,a_n)$ w ciele $\mathbb{K}$, to dla każdego $k\in\{1,\ldots,n\}$

$$\sum_{i=1}^{n}a_i\delta_{ik}=\sum_{i=1}^{n}\delta_{ki}a_i=a_k\,.$$

Dla każdego naturalnego n symbol δ określa ***macierz jednostkową***

$$\mathbf{1}_n\in\mathbb{K}^{n\times n}\quad\text{o elementach}\quad(\mathbf{1}_n)_{ij}=\delta_{ij}\,.$$

Mówiąc obrazowo, macierz ta ma jedynki na diagonali i zera poza nią. Jeśli $A\in\mathbb{K}^{m\times n}$, to widać natychmiast z własności symbolu Kroneckera, że

$$A\mathbf{1}_n=\mathbf{1}_mA=A\,.$$

Jeśli wymiar macierzy jednostkowej $\mathbf{1}_n$ jest oczywisty z kontekstu, to często pomijamy indeks w jej symbolu i piszemy $\mathbf{1}$. Macierz jednostkowa ma następujące oczywiste własności względem transpozycji i – gdy $\mathbb{K}=\mathbb{C}$ – względem sprzężeń:

$$\mathbf{1}^{\mathrm{T}}=\mathbf{1}\,,\quad\overline{\mathbf{1}}=\mathbf{1}\,,\quad\mathbf{1}^{\dagger}=\mathbf{1}\,.$$

W następującym twierdzeniu zbieramy własności łączące różne poznane operacje i działania macierzowe.

Twierdzenie 2. *Jeśli* $\lambda\in\mathbb{K}$, $A,B\in\mathbb{K}^{m\times n}$, $C,D\in\mathbb{K}^{n\times p}$, *to*

$$\lambda(AC)=(\lambda A)C=A(\lambda C)\,,$$
$$A(C+D)=AC+AD\,,\qquad(A+B)C=AC+BC\,,$$
$$(\lambda A)^{\mathrm{T}}=\lambda A^{\mathrm{T}}\,,\qquad(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}\,,\qquad(AC)^{\mathrm{T}}=C^{\mathrm{T}}A^{\mathrm{T}}\,.$$

Jeśli przy tym $\mathbb{K}=\mathbb{C}$, *to*

$$\overline{\lambda A}=\overline{\lambda}\,\overline{A}\,,\qquad\overline{A+B}=\overline{A}+\overline{B}\,,\qquad\overline{AC}=\overline{A}\,\overline{C}\,,$$
$$(\lambda A)^{\dagger}=\overline{\lambda}A^{\dagger}\,,\qquad(A+B)^{\dagger}=A^{\dagger}+B^{\dagger}\,,\qquad(AC)^{\dagger}=C^{\dagger}A^{\dagger}\,.$$

Dowód. Dowody wszystkich własności sprowadzają się do jawnego wypisania elementów macierzowych lewych i prawych stron równości. Pokażemy tylko przykładowo, że $(AC)^{\mathrm{T}} = C^{\mathrm{T}} A^{\mathrm{T}}$. Tożsamość ta jest wynikiem ciągu równości

$$(C^{\mathrm{T}} A^{\mathrm{T}})_{ij} = \sum_{k=1}^{n} (C^{\mathrm{T}})_{ik} (A^{\mathrm{T}})_{kj} = \sum_{k=1}^{n} A_{jk} C_{ki} = (AC)_{ji} = [(AC)^{\mathrm{T}}]_{ij} .$$

□

Jeśli A jest macierzą $n \times n$, to również A^{T} i $A^{\dagger}$ (w przypadku macierzy zespolonej) są macierzami kwadratowymi. Mówimy, że macierz A jest

– ***symetryczna***, gdy $A^{\mathrm{T}} = A$ (czyli $A_{ij} = A_{ji}$),

– ***antysymetryczna***, gdy $A^{\mathrm{T}} = -A$ (czyli $A_{ij} = -A_{ji}$).

Mówimy, że macierz zespolona jest

– ***hermitowska***, gdy $A^{\dagger} = A$ (czyli $\overline{A_{ij}} = A_{ji}$),

– ***antyhermitowska***, gdy $A^{\dagger} = -A$ (czyli $\overline{A_{ij}} = -A_{ji}$).

Zauważmy, że jeśli charakterystyka ciała $\mathbb{K}$ jest różna od 2, to diagonalne elementy macierzy antysymetrycznej są równe zeru. Elementy diagonalne macierzy hermitowskiej są rzeczywiste, a macierzy antyhermitowskiej - urojone. Pojęcie macierzy antyhermitowskiej ma mniejsze znaczenie od pozostałych, gdyż każda macierz antyhermitowska A może być przedstawiona w postaci $A = iB$, gdzie B jest macierzą hermitowską (wystarczy zauważyć, że jeśli A jest antyhermitowska, to $-iA$ jest hermitowska; tę ostatnią macierz oznaczamy B i dostajemy żądany wynik).

4 Ślad macierzy

Dla każdej liczby naturalnej n ***śladem macierzy kwadratowej*** $n \times n$ nazywamy wyrażenie wielomianowe względem elementów macierzy:

$$\operatorname{Tr} A = \sum_{i=1}^{n} A_{ii}$$

– używamy tego samego symbolu dla wszystkich n. Dla konkretnej macierzy jej śladem nazywamy też w skrócie wartość tego wyrażenia. Dla ustalonego n otrzymujemy wprost z definicji następujące własności śladu

$$\operatorname{Tr} A^{\mathrm{T}} = \operatorname{Tr} A , \quad \operatorname{Tr}(\lambda A + \mu B) = \lambda \operatorname{Tr} A + \mu \operatorname{Tr} B .$$

Dla $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{n\times m}$ obie macierze AB i BA są kwadratowe (choć mogą być różnych wymiarów). Obliczając ślad każdej z nich i korzystając z przemienności mnożenia i sumowania liczb pokazuje się, że

$$\operatorname{Tr} AB = \operatorname{Tr} BA\,.$$

5 Ogólny układ równań liniowych

Niech będą zadane liczby $a_{ij} \in \mathbb{K}$ ($i = 1, \ldots, m$, $j = 1, \ldots, n$) oraz $b_i \in \mathbb{K}$ ($i = 1, \ldots, m$). ***Ogólnym układem równań liniowych*** nazywamy koniunkcję warunków na ciąg ***niewiadomych*** $x_j \in \mathbb{K}$ ($j = 1, \ldots, n$), zadanych równaniami

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1\,,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2\,,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\,,\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m\,.
\end{aligned}$$

Każdy ciąg $(x_1, x_2, \ldots, x_n)$ spełniający ten układ nazywamy ***rozwiązaniem układu***, natomiast zbiór wszystkich rozwiązań nazywamy ***ogólnym rozwiązaniem układu***. Jeśli układ nie ma rozwiązania, mówimy, że jest ***sprzeczny***.

Ze współczynników układu oraz niewiadomych utwórzmy macierze:

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Macierz A nazywamy ***macierzą główną układu***, macierz B – ***kolumną wyrazów wolnych***, a macierz X – ***kolumną niewiadomych***. Ogólny układ równań liniowych można teraz zapisać zwięźle przy wykorzystaniu mnożenia macierzowego jako równanie macierzowe

$$AX = B\,.$$

Ogólną metodę rozwiązywania tego typu równań poznamy w dalszej części podręcznika.

6 Blokowa postać macierzy

Niech j i k będą zadanymi liczbami naturalnymi i załóżmy, że dla każdego $r \in \{1, \ldots, j\}$ i $s \in \{1, \ldots, k\}$ dana jest macierz $A_{rs} \in \mathbb{K}^{m(r)\times n(s)}$, gdzie $m(1), \ldots, m(j)$ i $n(1), \ldots, n(k)$ są zadanymi ciągami wymiarów tych macierzy.

Tworzymy nową macierz ustawiając wiersze i kolumny tych macierzy obok siebie w następujący sposób:

$$A = \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots} \\ A_{j1} & \dots & A_{jk} \end{pmatrix} .$$

Zauważmy, że wymiary bloków A_{rs} są zgodne z tym zapisem – przy ustalonym r bloki mają tyle samo wierszy, a przy ustalonym s – tyle samo kolumn. Mówimy, że macierz A jest zapisana w ***postaci blokowej***. W definicji macierzy za zbiór indeksujący wiersze oraz kolumny przyjęliśmy zbiór kolejnych liczb naturalnych. Zauważmy jednak, że można wiersze i kolumny ponumerować też w inny sposób, byleby zbiór numerujący był uporządkowany, tj. zapewniał określenie kolejności wierszy i kolumn. Wybierzemy numerację zgodną z powyższym zapisem blokowym. Wiersze i kolumny numerować będziemy parami liczb naturalnych, rx i sy odpowiednio, i położymy:

$$A_{rx,sy} = (A_{rs})_{xy} ,$$

gdzie $r \in \{1, \dots, j\}$, $s \in \{1, \dots, k\}$, $x \in \{1, \dots, m(r)\}$, $y \in \{1, \dots, n(s)\}$. Zapis ten oznacza, że elementy macierzy A identyfikujemy jako elementy bloków składających się na nią. Porządek leksykograficzny par rx jest zgodny z porządkiem wierszy macierzy A, (i podobnie dla par sy i kolumn).

Uzasadnieniem dla wprowadzenia postaci blokowej jest następujący wynik.

Twierdzenie 3. *Niech będą dane dwie macierze w postaci blokowej*

$$A = \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots} \\ A_{j1} & \dots & A_{jk} \end{pmatrix} , \quad B = \begin{pmatrix} B_{11} & \dots & B_{1l} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots} \\ B_{k1} & \dots & B_{kl} \end{pmatrix} ,$$

gdzie bloki A_{rs} mają wymiary $m(r) \times n(s)$, a bloki B_{st} wymiary $n(s) \times p(t)$. Wtedy

$$AB = \begin{pmatrix} C_{11} & \dots & C_{1l} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots} \\ C_{j1} & \dots & C_{jl} \end{pmatrix} , \quad \textit{gdzie} \quad C_{rt} = \sum_{s=1}^{k} A_{rs} B_{st} .$$

Dowód. Przy zastosowaniu numeracji elementów macierzy A i B wprowadzonej w dyskusji poprzedzającej twierdzenie, macierz AB dziedziczy numerację wierszy od numeracji wierszy macierzy A, a numerację kolumn – od numeracji kolumn macierzy B. Stąd dla $r \in \{1, \dots, j\}$, $x \in \{1, \dots, m(r)\}$, $t \in \{1, \dots, l\}$,

$z \in \{1, \dots, p(t)\}$ mamy

$$(AB)_{rx,tz} = \sum_{s=1}^{k} \sum_{y=1}^{n(s)} A_{rx,sy} B_{sy,tz} = \sum_{s=1}^{k} \sum_{y=1}^{n(s)} (A_{rs})_{xy} (B_{st})_{yz} =$$
$$= \sum_{s=1}^{k} (A_{rs} B_{st})_{xz} = \left(\sum_{s=1}^{k} A_{rs} B_{st} \right)_{xz} ,$$

co dowodzi prawdziwości postaci blokowej z tezy. □

Wzór na mnożenie macierzy blokowych wygląda więc identycznie jak definicyjny wzór mnożenia (po zamianie elementów na bloki); pamiętać należy jedynie o koniecznej zgodności wymiarów bloków oraz o kolejności bloków w iloczynach bloków (nieprzemienność mnożenia macierzy).

7 Półgrupy macierzy kwadratowych. Ogólne grupy liniowe

Rozważmy dla zadanej liczby $n \in \mathbb{N}$ zbiór macierzy kwadratowych $\mathbb{K}^{n \times n}$. W każdym z tych zbiorów mnożenie macierzowe jest łącznym działaniem wewnętrznym, z elementem neutralnym równym macierzy jednostkowej $\mathbf{1}_n$. Dla każdego n zbiór $\mathbb{K}^{n \times n}$ tworzy więc półgrupę (z jedynką) z działaniem mnożenia. Jeśli macierz A ma element odwrotny – ***macierz odwrotną*** A^{-1} spełniającą warunki

$$AA^{-1} = A^{-1}A = \mathbf{1}$$

– to nazywamy ją ***macierzą nieosobliwą***, w przeciwnym przypadku mówimy, że jest ona ***osobliwa***. Jak stwierdzić, czy macierz jest nieosobliwa i jak znaleźć jej odwrotną, jeśli istnieje, dowiemy się w następnym paragrafie. Z ogólnych rozważań dotyczących półgrup wiemy, że podzbiór półgrupy multiplikatywnej $\mathbb{K}^{n \times n}$ złożony z macierzy nieosobliwych stanowi grupę. Nazywamy ją ***ogólną grupą liniową (wymiaru*** n***, nad ciałem*** $\mathbb{K}$***)*** i oznaczamy $GL(n, \mathbb{K})$. Operacje transpozycji i sprzężeń działają wewnątrz grup liniowych:

Twierdzenie 4. *Jeśli* $A \in GL(n, \mathbb{K})$*, to także* $A^{\mathrm{T}} \in GL(n, \mathbb{K})$*, i wtedy*

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}} .$$

Jeśli $A \in GL(n, \mathbb{C})$*, to także* $\overline{A} \in GL(n, \mathbb{C})$ *i* $A^{\dagger} \in GL(n, \mathbb{C})$*, i wtedy*

$$(\overline{A})^{-1} = \overline{A^{-1}} , \quad (A^{\dagger})^{-1} = (A^{-1})^{\dagger} .$$

Dowód. Korzystając z własności transpozycji i mnożenia macierzowego mamy $(A^{-1})^{\mathrm{T}} A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = \mathbf{1}^{\mathrm{T}} = \mathbf{1}$ i podobnie w odwrotnym porządku macierzy, co dowodzi pierwszego stwierdzenia. Dowód dla sprzężenia hermitowskiego jest analogiczny. Z własności sprzężenia zespolonego $\overline{A^{-1}}\,\overline{A} = \overline{A^{-1}A} = \mathbf{1}$ i podobnie dla odwrotnego porządku macierzy, co kończy dowód. □

8 Przykłady

(i) Kombinacja liniowa macierzy

$$5\begin{pmatrix}1 & 3 & 5\\ 8 & -1 & 3\end{pmatrix} - 2\begin{pmatrix}-2 & 1 & 3\\ -3 & 0 & 2\end{pmatrix} = \begin{pmatrix}9 & 13 & 19\\ 46 & -5 & 11\end{pmatrix}$$

(ii) Transpozycja i sprzężenie hermitowskie

$$\begin{pmatrix}1 & -7 & 9 & -14\\ 3 & -2 & 5 & 8\\ 0 & 10 & 1 & -5\end{pmatrix}^{\mathrm{T}} = \begin{pmatrix}1 & 3 & 0\\ -7 & -2 & 10\\ 9 & 5 & 1\\ -14 & 8 & -5\end{pmatrix},$$

$$\begin{pmatrix}1+i & -3i & 2\\ 4-2i & 7+i & 1\end{pmatrix}^{\dagger} = \begin{pmatrix}1-i & 4+2i\\ +3i & 7-i\\ 2 & 1\end{pmatrix}.$$

(iii) Własności symetrii

$$S = \begin{pmatrix}1 & 4 & 3\\ 4 & 7 & 8\\ 3 & 8 & 2\end{pmatrix}, \quad A = \begin{pmatrix}0 & 6 & 3\\ -6 & 0 & -9\\ -3 & 9 & 0\end{pmatrix}, \quad H = \begin{pmatrix}6 & 3-4i & 1+2i\\ 3+4i & 1 & 5+8i\\ 1-2i & 5-8i & 4\end{pmatrix}.$$

Macierz S jest symetryczna, macierz A jest antysymetryczna, macierz H jest hermitowska.

(iv) Mnożenie macierzy

$$\begin{pmatrix}1 & 2 & 4\\ -2 & 1 & 3\end{pmatrix}\begin{pmatrix}3 & -2 & 1 & 0\\ -1 & 2 & 1 & -2\\ 5 & 2 & 1 & 3\end{pmatrix} = \begin{pmatrix}21 & 10 & 7 & 8\\ 8 & 12 & 2 & 7\end{pmatrix};$$

$$\text{niech} \qquad X = \begin{pmatrix}1\\ 3\\ -2\end{pmatrix}, \quad Y = \begin{pmatrix}2\\ -1\\ 4\end{pmatrix},$$

$$\text{wtedy} \qquad X^{\mathrm{T}}Y = -9, \quad YX^{\mathrm{T}} = \begin{pmatrix}2 & 6 & -4\\ -1 & -3 & 2\\ 4 & 12 & -8\end{pmatrix}.$$

(v) Ślad macierzy

$$\operatorname{Tr}\begin{pmatrix}1 & 3 & -7\\ 0 & 2 & 4\\ 6 & -5 & 5\end{pmatrix} = 8\,.$$

(vi) Mnożenie blokowe

$$\text{Niech}\quad A=\begin{pmatrix}K & L\\ M & N\end{pmatrix}\,,\quad B=\begin{pmatrix}P\\ Q\end{pmatrix}\,,\quad \text{gdzie}\quad K=\begin{pmatrix}0 & 1\\ -1 & 2\end{pmatrix}\,,\quad L=\begin{pmatrix}2\\ 3\end{pmatrix}\,,$$

$$M=\begin{pmatrix}1 & -1\\ 3 & 1\\ 2 & 0\end{pmatrix}\,,\quad N=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}\,,\quad P=\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}\,,\quad Q=\begin{pmatrix}1 & 1\end{pmatrix}\,.$$

$$\text{Wtedy}\qquad AB=\begin{pmatrix}C\\ D\end{pmatrix}\,,\quad \text{gdzie}$$

$$C=KP+LQ=\begin{pmatrix}2 & 4\\ 2 & 8\end{pmatrix}\,,\quad D=MP+NQ=\begin{pmatrix}2 & -2\\ 3 & -1\\ 3 & -1\end{pmatrix}\,.$$

§7 Wyznaczniki. Macierz odwrotna

1 Macierz 2×2

Chcemy w tym paragrafie znaleźć odpowiedź na pytanie: kiedy macierz ma odwrotną i jak ją znaleźć? Na wstępie rozpatrzmy najprostszy przypadek, macierze 2×2. Niech $A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$ i załóżmy, że istnieje macierz odwrotna w postaci $A^{-1}=\begin{pmatrix}w & x\\ y & z\end{pmatrix}$. Wypisując warunek $AA^{-1}=\mathbf{1}$ jako układ równań na w, x, y i z dostajemy:

$$a_{11}w+a_{12}y=1\,,\ a_{21}w+a_{22}y=0\,,\ a_{11}x+a_{12}z=0\,,\ a_{21}x+a_{22}z=1\,.$$

Z pierwszego lub czwartego z tych równań widać, że A nie może być macierzą zerową. Mnożąc pierwsze równanie przez a_{22}, a drugie przez a_{12}, i odejmując stronami dostajemy równanie $(a_{11}a_{22}-a_{12}a_{21})w=a_{22}$. Analogicznie otrzymujemy $(a_{11}a_{22}-a_{12}a_{21})y=-a_{21}$, a z pozostałych dwóch równań $(a_{11}a_{22}-a_{12}a_{21})x=-a_{12}$ i $(a_{11}a_{22}-a_{12}a_{21})z=a_{11}$. Ponieważ $A\neq 0$, to nie mogą jednocześnie znikać strony wszystkich czterech ostatnich równań, a stąd $a_{11}a_{22}-a_{12}a_{21}\neq 0$. Oznaczmy $\det A=a_{11}a_{22}-a_{12}a_{21}$ i nazwijmy tę liczbę wyznacznikiem macierzy A. Wyliczając w, x, y i z dostajemy

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}\,.$$

Odwrotnie, załóżmy że $\det A \neq 0$. Wtedy łatwo sprawdzić bezpośrednim rachunkiem, że powyższy przepis zadaje macierz spełniającą $AA^{-1} = A^{-1}A = \mathbf{1}$, więc A ma odwrotną A^{-1}.

W przypadku macierzy 2×2 otrzymaliśmy więc rezultat: macierz ma odwrotną wtedy, i tylko wtedy, gdy jej wyznacznik jest różny od zera. Okaże się, jak zobaczymy, że po odpowiednim uogólnieniu pojęcia wyznacznika podobne twierdzenie jest prawdziwe dla dowolnych macierzy kwadratowych.

2 Wyznaczniki

Wyznacznikiem macierzy ze zbioru $\mathbb{K}^{n\times n}$ nazywamy następujący wielomian względem jej elementów

$$\det A = \sum_{\pi\in S_n} \operatorname{sgn}\pi\, A_{1\pi(1)}A_{2\pi(2)}\dots A_{n\pi(n)}\,.$$

Sumowanie w tym wzorze definicyjnym przebiega po wszystkich permutacjach z grupy S_n. Mnożenie przez $\operatorname{sgn}\pi$ jest rozumiane tak, jak to wyjaśniliśmy z końcem punktu 10, §3. Dla konkretnej macierzy jej wyznacznikiem nazywamy też w skrócie wartość tego wielomianu. Jeśli $A = (A_{ij})$, to notujemy też inaczej:

$$\det A \equiv \begin{vmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{vmatrix}\,.$$

W przypadku $n = 2$ grupa S_2 zawiera dwa elementy: identyczność, o znaku dodatnim, i transpozycję (12), o znaku ujemnym. Powyższa definicja daje wtedy przepis identyczny z określonym w poprzednim punkcie.

Dla otrzymania własności wyznaczników pomocne będzie następujące pojęcie. Załóżmy najpierw, że $\operatorname{char}\mathbb{K} \neq 2$. Nazywamy ***symbolem całkowicie antysymetrycznym*** każde odwzorowanie

$$\underbrace{\{1,\dots,n\}\times\dots\times\{1,\dots,n\}}_{n \text{ razy}} \ni (i_1,\dots,i_n) \mapsto a_{i_1\dots i_n} \in \mathbb{K}$$

o tej własności, że dowolne przestawienie pary liczb w argumencie $(i_1,\dots,i_n)$ prowadzi do zmiany znaku wartości $a_{i_1\dots i_n}$. Stąd, jeśli dowolna liczba ze zbioru $\{1,\dots,n\}$ występuje dwukrotnie w ciągu $(i_1,\dots,i_n)$, to $a_{i_1\dots i_n} = -a_{i_1\dots i_n}$ (bo przestawienie miejscami w ciągu tych identycznych liczb równocześnie nie zmienia symbolu $a_{i_1\dots i_n}$ i zmienia jego znak). Dla ciała o charakterystyce różnej od 2 wynika stąd $a_{i_1\dots i_n} = 0$. Jeśli wszystkie liczby w ciągu $(i_1,\dots,i_n)$ są różne, to ciąg ten otrzymuje się z ciągu $(1,\dots,n)$ przez permutację $\left(\begin{smallmatrix} 1 & \dots & n \\ i_1 & \dots & i_n \end{smallmatrix}\right)$, czyli przez kolejne transpozycje, na które rozkłada się ta permutacja. Pod działaniem każdej z nich początkowa wartość $a_{1\dots n}$ zmienia znak, więc po osiągnięciu $a_{i_1\dots i_n}$ multiplikatywna zmiana znaku jest dana przez $\operatorname{sgn}\left(\begin{smallmatrix}1 & \cdots & n\\ i_1 & \dots & i_n\end{smallmatrix}\right)$. Reasumujemy: dla symbolu całkowicie antysymetrycznego jest

$$a_{i_1\dots i_n} = \begin{cases} 0\,, & \text{jeśli w ciągu } (i_1,\dots,i_n) \text{ dowolna liczba} \\ & \text{powtarza się,} \\ \operatorname{sgn}\left(\begin{smallmatrix}1 & \cdots & n\\ i_1 & \dots & i_n\end{smallmatrix}\right) a_{1\dots n}\,, & \text{w przypadku przeciwnym.}\end{cases}$$

Widać stąd, że dla danego n symbol całkowicie antysymetryczny jest w pełni określony wartością $a_{1\dots n}$ i każde dwa takie symbole różnią się co najwyżej o czynnik.

Symbol całkowicie antysymetryczny, dla którego $a_{1\dots n} = 1$, nazywamy ***symbolem Levi-Civity*** i oznaczamy $\epsilon_{i_1\dots i_n}$. Dowolny inny symbol całkowicie antysymetryczny $a_{i_1\dots i_n}$ może być zapisany jako $a_{i_1\dots i_n} = a_{1\dots n}\,\epsilon_{i_1\dots i_n}$. Zanotujemy ponadto tożsamość wiążącą symbole Levi-Civity o różnej liczbie wskaźników:

$$\epsilon_{i_1\dots i_{n-1}n} = \epsilon_{i_1\dots i_{n-1}} \quad \text{dla} \quad i_1,\dots,i_{n-1} \in \{1,\dots,n-1\}\,.$$

Jeśli $\operatorname{char}\mathbb{K} = 2$, to równość $a = -a$ jest spełniona dla każdego elementu ciała, więc w tym przypadku definicja symbolu całkowicie antysymetrycznego nic nie mówi o wartościach symbolu dla powtarzających się wskaźników. Wszystkie powyższe rezultaty pozostaną jednak w mocy, jeśli *założymy* dodatkowo znikanie symbolu dla powtarzających się wskaźników.

Twierdzenie 1. *Niech $A \in \mathbb{K}^{n\times n}$. Wtedy następujące stwierdzenia są równoważne:*

(i) $\det A = \sum_{\pi\in S_n} \operatorname{sgn}\pi\, A_{1\pi(1)}\dots A_{n\pi(n)}$;

(ii) $\det A = \sum_{\pi\in S_n} \operatorname{sgn}\pi\, A_{\pi(1)1}\dots A_{\pi(n)n}$;

(iii) $\det A\,\epsilon_{i_1\dots i_n} = \sum_{j_1,\dots,j_n} \epsilon_{j_1\dots j_n} A_{i_1j_1}\dots A_{i_nj_n}$;

(iv) $\det A\,\epsilon_{i_1\dots i_n} = \sum_{j_1,\dots,j_n} \epsilon_{j_1\dots j_n} A_{j_1i_1}\dots A_{j_ni_n}$.

Dowód. Pierwsza formuła jest definicją wyznacznika. Pokażemy równoważność z nią pozostałych.

(ii) Każda liczba $j \in \{1,\dots,n\}$ znajduje się dokładnie jeden raz w ciągu

$(\pi(1),\ldots,\pi(n))$; jeśli przy tym $j=\pi(l)$, to $A_{\pi(l)l}=A_{j\pi^{-1}(j)}$. Z przemienności mnożenia liczb mamy więc $A_{\pi(1)1}\ldots A_{\pi(n)n} = A_{1\pi^{-1}(1)}\ldots A_{n\pi^{-1}(n)}$. Z własności znaku permutacji, tw. 11, §5, mamy $\operatorname{sgn}\pi^{-1} = \operatorname{sgn}\pi$, więc prawą stronę formuły (ii) można zapisać w równoważnej postaci

$$\sum_{\pi\in S_n}\operatorname{sgn}\pi^{-1}\, A_{1\pi^{-1}(1)}\ldots A_{n\pi^{-1}(n)}\,.$$

Gdy π przebiega zbiór wszystkich permutacji z S_n, to również π^{-1} przebiega ten sam zbiór, więc sumowanie po π w ostatnim wyrażeniu można zastąpić sumowaniem po π^{-1}. Pisząc σ zamiast π^{-1} otrzymujemy wyrażenie identyczne z prawą stroną formuły (i). Prawe strony formuł (i) i (ii) są więc równe.

(iii) Prawą stronę formuły (iii) możemy zapisać w postaci

$$\sum_{j_3,\ldots,j_n}\Bigg(\sum_{j_1,j_2}\epsilon_{j_1\ldots j_n}A_{i_1j_1}A_{i_2j_2}\Bigg)A_{i_3j_3}\ldots A_{i_nj_n}\,.$$

Korzystając z własności symbolu Levi-Civity przepisujemy wyrażenie w nawiasach w postaci

$$\begin{aligned}\sum_{j_1<j_2}\epsilon_{j_1j_2\ldots j_n}A_{i_1j_1}A_{i_2j_2} - \sum_{j_1>j_2}\epsilon_{j_2j_1\ldots j_n}A_{i_1j_1}A_{i_2j_2} &= \\ = \sum_{j_1<j_2}\epsilon_{j_1\ldots j_n}(A_{i_1j_1}A_{i_2j_2}-A_{i_2j_1}A_{i_1j_2})\,,\end{aligned}$$

gdzie w drugim kroku wskaźniki j_1, j_2 zamieniliśmy nazwami w drugiej sumie po lewej stronie. Stąd jest widoczne, że prawa strona formuły (iii) zmienia znak przy przestawieniu wskaźników i_1, i_2, i jest równa zero dla $i_1 = i_2$. Podobnie dla każdej pary wskaźników. Z wyników poprzedzających twierdzenie wnioskujemy więc, że całe wyrażenie jest proporcjonalne do symbolu Levi-Civity. Formułę wystarczy więc sprawdzić dla ciągu wskaźników $(i_1,\ldots,i_n)=(1,\ldots,n)$. Dla tego przypadku lewa strona formuły jest równa $\det A$. Po prawej stronie przyczynek do sumy dają tylko wyrazy, dla których $j_1,\ldots,j_n$ są wszystkie różne, sumowanie przebiega więc efektywnie po wszystkich permutacjach $\pi := \left(\begin{smallmatrix}1 & \ldots & n\\ j_1 & \ldots & j_n\end{smallmatrix}\right)$. Przy tych oznaczeniach $\epsilon_{j_1\ldots j_n}=\operatorname{sgn}\pi$ i $A_{lj_l}=A_{l\pi(l)}$, więc dostajemy formułę identyczną z formułą (i).

(iv) Dowód jak dla (iii), przez sprowadzenie do formuły (ii). □

3 Własności wyznaczników

W następnym twierdzeniu podajemy podstawowe własności wyznaczników. Niektóre z nich są parafrazami formuł z poprzedniego twierdzenia.

Twierdzenie 2. *Dla dowolnych macierzy kwadratowych wyznacznik ma następujące własności:*

(i) $\det A^{\mathrm{T}} = \det A$.

(ii) *dla* $\mathbb{K} = \mathbb{C}$: $\quad \det \overline{A} = \overline{\det A}\,, \ \det A^{\dagger} = \overline{\det A}$.

(iii) $\det AB = \det A \det B$ *(twierdzenie Cauchy'ego).*

(iv) *Jeśli macierz* A' *powstaje z macierzy* A *przez przestawienie wierszy lub kolumn zgodnie z permutacją* π, *to* $\det A' = \operatorname{sgn}\pi \det A$. *W szczególności: zamiana miejscami dwóch wierszy lub kolumn macierzy prowadzi do zmiany znaku jej wyznacznika.*

(v) *Jeśli wiersz lub kolumna macierzy jest równa zeru, to wyznacznik tej macierzy jest równy zeru.*

(vi) *Jeśli dwa wiersze macierzy lub dwie kolumny macierzy są identyczne, to wyznacznik tej macierzy jest równy zero.*

(vii) *Jeśli macierze* A', A'' *i* A *różnią się tylko wybranym wierszem (lub wybraną kolumną), który w macierzy* A' *jest równy* X', *w macierzy* A'' *jest równy* X'', *a w macierzy* A *jest kombinacją liniową* $X = \alpha' X' + \alpha'' X''$ *(lub podobnie dla kolumn), to*

$$\det A = \alpha' \det A' + \alpha'' \det A''\,.$$

W szczególności, jeśli przy tych samych oznaczeniach dla A', A, X' *i* X *jest* $X = \alpha' X'$, *to* $\det A = \alpha' \det A'$.

(viii) *Wyznacznik macierzy nie zmienia się, jeśli do dowolnego wiersza (kolumny) tej macierzy dodać dowolny inny wiersz (odpowiednio: kolumnę) pomnożony przez dowolną liczbę.*

(ix) *Jeśli wymiar macierzy* A *jest równy* n, *to* $\det(\alpha A) = \alpha^n \det A$.

Dowód. (i) Stwierdzenie wynika bezpośrednio z definicji macierzy transponowanej i z równoważności formuł (i) i (ii) z poprzedniego twierdzenia.

(ii) Stwierdzenie wynika z własności sprzężenia i z punktu (i).

(iii) Własność wynika z ciągu równości:

$$\begin{aligned}\det AB = &\sum_{j_1,\ldots,j_n} \epsilon_{j_1\ldots j_n}(AB)_{j_11}\ldots(AB)_{j_nn} = \\ &= \sum_{k_1,\ldots,k_n}\sum_{j_1,\ldots,j_n} \epsilon_{j_1\ldots j_n}A_{j_1k_1}\ldots A_{j_nk_n}B_{k_11}\ldots B_{k_nn} = \\ &\qquad= \sum_{k_1,\ldots,k_n} \det A\,\epsilon_{k_1\ldots k_n}B_{k_11}\ldots B_{k_nn} = \det A\det B\,.\end{aligned}$$

Pierwsza równość jest konsekwencją zastosowania formuły (iv) poprzedniego twierdzenia do macierzy AB i indeksów $(i_1,\ldots,i_n)=(1,\ldots,n)$, druga wynika z definicji mnożenia macierzy, trzecia i czwarta znów z zastosowania formuły (iv) poprzedniego twierdzenia (do A i B odpowiednio).

(iv) Niech $\pi = \left(\begin{smallmatrix}1 & \cdots & n\\ i_1 & \cdots & i_n\end{smallmatrix}\right)$. Jeśli macierz A' powstaje z macierzy A przez przestawienie wierszy zgodnie z permutacją π, to $A'_{lj_l} = A_{i_lj_l}$. Stosując dwukrotnie formułę (iii) poprzedniego twierdzenia dostajemy:

$$\det A' = \sum_{j_1,\ldots,j_n} \epsilon_{j_1\ldots j_n}A'_{1j_1}\ldots A'_{nj_n} = \sum_{j_1,\ldots,j_n} \epsilon_{j_1\ldots j_n}A_{i_1j_1}\ldots A_{i_nj_n} = \det A\,\epsilon_{i_1\ldots i_n}\,,$$

co kończy dowód dla permutacji wierszy. Dla kolumn podobny dowód przy użyciu formuły (iv) poprzedniego twierdzenia.

(v) Znikanie k-tego wiersza oznacza, że dla wszystkich $i\in\{1,\ldots,n\}$ jest $A_{ki}=0$. Formuła (i) poprzedniego twierdzenia daje teraz żądany wynik. Podobnie dla kolumn z formuły (ii) tego twierdzenia.

(vi) Niech w macierzy A dwa wiersze będą identyczne. Przestawiając je na pierwsze dwa miejsca w ciągu wierszy zmieniamy tylko ewentualnie znak wyznacznika (na podstawie punktu (iv)). Jeśli pierwsze dwa wiersze są identyczne, $A_{1j}=A_{2j}$, to formuła (iii) poprzedniego twierdzenia daje

$$\det A\,\epsilon_{123\ldots n} = \det A\,\epsilon_{113\ldots n}\,,\quad \text{czyli}\quad \det A = 0\,.$$

Dla identycznych kolumn podobny dowód z formuły (iv) poprzedniego twierdzenia.

(vii) Niech założenia twierdzenia będą spełnione dla k-tego wiersza, czyli $A_{ki}=\alpha'A'_{ki}+\alpha''A''_{ki}$ dla wszystkich $i\in\{1,\ldots,n\}$, oraz $A_{ji}=A'_{ji}=A''_{ji}$ dla $j\neq k$ i wszystkich i. Używając tych równości w formule (i) poprzedniego twierdzenia dostajemy tezę. Podobnie dla kolumn z formuły (ii).

(viii) Oznaczmy początkową macierz przez A', a końcową przez A. Opisana procedura jest szczególnym przypadkiem związku z poprzedniego punktu twierdzenia, gdzie $\alpha'=1$, a A'' ma dwa identyczne wiersze lub kolumny. Teza wynika więc z punktów (vi) i (vii).

(ix) Stwierdzenie to wynika natychmiast z dowolnej z formuł twierdzenia 1. $\square$

4 Rozwinięcie Laplace'a

Niech A będzie dowolną macierzą prostokątną. ***Minorem macierzy*** A nazywamy wyznacznik dowolnej macierzy kwadratowej powstającej z macierzy A przez wykreślenie z niej pewnej liczby wierszy i/lub kolumn. W szczególności, jeśli A jest macierzą kwadratową, to przez $\mathcal{M}_{ij}(A)$ oznaczymy minor zadany przez wykreślenie z macierzy A i-tego wiersza i j-tej kolumny.

Lemat 3. *Niech A będzie macierzą kwadratową. Jeśli $A_{ij} = 1$ dla zadanych i, j, a pozostałe elementy i-tego wiersza (lub pozostałe elementy j-tej kolumny) są równe zero, to* $\det A = (-1)^{i+j}\mathcal{M}_{ij}(A)$.

Dowód. Jeśli oznaczymy przez $\hat{A}$ macierz, która powstaje z macierzy A przez skreślenie i-tego wiersza i j-tej kolumny, to $\mathcal{M}_{ij}(A) = \det \hat{A}$. Załóżmy, że $A_{ij} = 1$, a pozostałe wyrazy i-tego wiersza są równe zero. Przestawienie i-tego wiersza oraz j-tej kolumny na ostatnie pozycje prowadzi do powstania macierzy blokowej $A' = \left(\begin{smallmatrix} \hat{A} & * \\ 0 & 1 \end{smallmatrix}\right)$. Przestawienie i-tego wiersza na ostatnią pozycję jest permutacją wierszy o znaku $(-1)^{n-i}$ (można ją otrzymać jako wynik zamiany miejscami tego wiersza z kolejnymi wierszami, począwszy od $i+1$-szego do n-tego), przestawienie j-tej kolumny na ostatnią pozycję ma znak $(-1)^{n-j}$. Z własności wyznaczników mamy więc $\det A = (-1)^{2n-i-j} \det A' = (-1)^{i+j} \det A'$. Ponieważ $A'_{ni} = \delta_{ni}$, to

$$\begin{aligned}
\det A' = \sum_{i_1,\ldots,i_n \in \{1,\ldots,n\}} \epsilon_{i_1\ldots i_n} A'_{1i_1} \ldots A'_{n-1\,i_{n-1}} \delta_{ni_n} &= \\
= \sum_{i_1,\ldots,i_{n-1} \in \{1,\ldots,n-1\}} \epsilon_{i_1\ldots i_{n-1}n} A'_{1i_1} \ldots A'_{n-1\,i_{n-1}} &= \\
= \sum_{i_1,\ldots,i_{n-1} \in \{1,\ldots,n-1\}} \epsilon_{i_1\ldots i_{n-1}} \hat{A}_{1i_1} \ldots \hat{A}_{n-1\,i_{n-1}} &= \det \hat{A}\,.
\end{aligned}$$

Pierwsza równość wynika z formuły (iii) twierdzenia 1, druga z własności symbolu Kroneckera, trzecia z własności symboli Levi-Civity i postaci macierzy A', czwarta znów z formuły (iii) twierdzenia 1. Zestawiając teraz otrzymane równości dostajemy tezę. $\square$

Dla dowolnej macierzy kwadratowej A wyrażenie

$$\mathcal{A}_{ij}(A) := (-1)^{i+j}\mathcal{M}_{ij}(A)$$

nazywamy ***dopełnieniem algebraicznym elementu*** A_{ij}. Korzystając z udowodnionego lematu, z łatwością otrzymamy teraz podstawowe narzędzie do obliczania wyznaczników.

Twierdzenie 4 (Rozwinięcie Laplace'a)**.** *Niech A będzie dowolną macierzą kwadratową $n \times n$. Wtedy dla każdego $i \in \{1, \dots, n\}$:*

$$\det A = \sum_{k=1}^{n} A_{ik}\, \mathcal{A}_{ik}(A) = \sum_{k=1}^{n} A_{ki}\, \mathcal{A}_{ki}(A)\,.$$

Dowód. Aby udowodnić drugą z formuł przedstawiamy i-tą kolumnę macierzy A w postaci kombinacji liniowej

$$A_{1i} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + A_{2i} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + A_{ni} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}\,.$$

Z własności (vii), twierdzenie 2, mamy $\det A = \sum_k A_{ki} \det C_{ki}$, gdzie C_{ki} jest macierzą powstającą z macierzy A przez zastąpienie w niej i-tej kolumny przez kolumnę złożoną z samych zer, z wyjątkiem pozycji k-tej, na której stoi 1. Ponieważ i-ta kolumna ulega wykreśleniu przy liczeniu minora $\mathcal{M}_{ki}$, to przy uwzględnieniu wyniku lematu dla macierzy C_{ki} dostajemy

$$\det C_{ki} = (-1)^{k+i} \mathcal{M}_{ki}(C_{ki}) = (-1)^{k+i} \mathcal{M}_{ki}(A)\,,$$

co kończy dowód drugiej formuły. Dowód pierwszej z formuł twierdzenia dostajemy w analogiczny sposób przez rozpisanie i-tego wiersza w postaci kombinacji liniowej. □

Pierwsza formuła twierdzenia nazywa się ***rozwinięciem wyznacznika względem i-tego wiersza***, a druga – ***rozwinięciem wyznacznika względem i-tej kolumny***.

Rozwinięcie Laplace'a jest szczególnie efektywne, gdy stosuje się je w połączeniu z innymi własnościami wyznaczników, tw. 2. Dodając odpowiednie wielokrotności wybranej kolumny do pozostałych kolumn, można sprowadzić do zera wszystkie wyrazy wybranego wiersza, z wyjątkiem wyrazu na przecięciu tej kolumny i tego wiersza. Rozwinięcie Laplace'a względem wybranego wiersza ma wtedy tylko jeden wyraz (patrz przykład (iii) poniżej). Podobnie dla kolumn i wierszy zamienionych rolami.

5 Przykłady

(i) Wyznacznik trzeciego stopnia
Wprost z definicji wyznacznika dla macierzy 3×3 dostajemy

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

$$- a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}\,.$$

Następujące schematyczne przedstawienie tego wyrażenia może ułatwić jego zapamiętanie:

$$+\begin{vmatrix} * & & \\ & * & \\ & & * \end{vmatrix} + \begin{vmatrix} & * & \\ & & * \\ * & & \end{vmatrix} + \begin{vmatrix} & & * \\ * & & \\ & * & \end{vmatrix} - \begin{vmatrix} & & * \\ & * & \\ * & & \end{vmatrix} - \begin{vmatrix} & * & \\ * & & \\ & & * \end{vmatrix} - \begin{vmatrix} * & & \\ & & * \\ & * & \end{vmatrix}\,.$$

(ii) Rozwinięcie Laplace'a
Rozwijając dany niżej wyznacznik względem trzeciego wiersza dostajemy

$$\begin{vmatrix} 1 & 2 & -1 & 2 \\ 3 & 1 & 4 & -3 \\ 5 & 2 & -2 & 3 \\ 1 & 2 & 4 & -1 \end{vmatrix} = 5\,(-1)^4 \begin{vmatrix} 2 & -1 & 2 \\ 1 & 4 & -3 \\ 2 & 4 & -1 \end{vmatrix} + 2\,(-1)^5 \begin{vmatrix} 1 & -1 & 2 \\ 3 & 4 & -3 \\ 1 & 4 & -1 \end{vmatrix} +$$

$$+ (-2)\,(-1)^6 \begin{vmatrix} 1 & 2 & 2 \\ 3 & 1 & -3 \\ 1 & 2 & -1 \end{vmatrix} + 3\,(-1)^7 \begin{vmatrix} 1 & 2 & -1 \\ 3 & 1 & 4 \\ 1 & 2 & 4 \end{vmatrix} =$$

$$= 5 \cdot 13 - 2 \cdot 24 - 2 \cdot 15 - 3 \cdot (-25) = 62\,.$$

(iii) Upraszczanie wyznaczników
Wyliczenie tego samego wyznacznika, co w poprzednim przykładzie, przy wykorzystaniu zarówno rozwinięcia Laplace'a, jak i własności wyznaczników, tw. 2, może przebiegać np. następująco:

$$\begin{vmatrix} 1 & 2 & -1 & 2 \\ 3 & 1 & 4 & -3 \\ 5 & 2 & -2 & 3 \\ 1 & 2 & 4 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 3 & -5 & 7 & -9 \\ 5 & -8 & 3 & -7 \\ 1 & 0 & 5 & -3 \end{vmatrix} = \begin{vmatrix} -5 & 7 & -9 \\ -8 & 3 & -7 \\ 0 & 5 & -3 \end{vmatrix} =$$

$$= \frac{1}{5} \begin{vmatrix} -5 & 7 & -45 \\ -8 & 3 & -35 \\ 0 & 5 & -15 \end{vmatrix} = \frac{1}{5} \begin{vmatrix} -5 & 7 & -24 \\ -8 & 3 & -26 \\ 0 & 5 & 0 \end{vmatrix} = - \begin{vmatrix} -5 & -24 \\ -8 & -26 \end{vmatrix} = 62\,.$$

Pierwszą równość uzyskaliśmy dodając odpowiednie wielokrotności pierwszej kolumny do pozostałych, a zastosowanie do wyniku rozwinięcia Laplace'a względem pierwszego wiersza (które to rozwinięcie ma tylko jeden składnik w sumie) daje drugą równość. Dla uzyskania trzeciej równości pomnożyliśmy trzecią kolumnę i podzielili cały wyznacznik przez tę samą liczbę (co nie zmienia wartości wyznacznika), tak aby ułatwić otrzymanie w następnym kroku drugiego zera w trzecim wierszu (przez dodanie wielokrotności drugiej kolumny do trzeciej). Przedostatni krok to rozwinięcie Laplace'a względem trzeciego wiersza (dające znów tylko jeden składnik). Ostatni krok oczywisty.

(iv) Wyznacznik macierzy trójkątnej

Rozwińmy wyznacznik macierzy trójkątnej $\begin{pmatrix} a_1 & * & * & \dots & * \\ 0 & a_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix}$ względem ostatniego wiersza. Dostaniemy wtedy

$$\begin{vmatrix} a_1 & * & * & \dots & * \\ 0 & a_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n \end{vmatrix} = a_n \begin{vmatrix} a_1 & * & * & \dots & * \\ 0 & a_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} \end{vmatrix} .$$

Przez indukcję otrzymujemy więc

$$\begin{vmatrix} a_1 & * & * & \dots & * \\ 0 & a_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n \end{vmatrix} = a_1 \dots a_n \,;$$

podobnie dla macierzy o formie transponowanej względem powyższej (zera nad diagonalą). W szczególności:

$$\det \mathbf{1} = 1 \,.$$

(v) Wyznacznik macierzy blokowo trójkątnej szczególnej postaci
Niech A będzie blokiem kwadratowym. Wtedy

$$\det \begin{pmatrix} A & B \\ 0 & \mathbf{1} \end{pmatrix} = \det A \,.$$

Metoda dowodu jest taka, jak w poprzednim przykładzie. Podobnie dla macierzy, w której blok A stoi w prawym dolnym rogu macierzy, a blok jednostkowy – w lewym górnym (rozwijanie względem kolumn).

(vi) Wyznacznik macierzy blokowo trójkątnej o czterech blokach
Macierz blokową $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ można przedstawić jako iloczyn

$$\begin{pmatrix} \mathbf{1} & B \\ 0 & C \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & \mathbf{1} \end{pmatrix} .$$

Z twierdzenia Cauchy'ego i poprzedniego przykładu dostajemy

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C \, .$$

Podobny wynik jest słuszny dla macierzy z blokiem zerowym powyżej diagonali.

(vii) Wyznacznik macierzy blokowo trójkątnej – ogólny przypadek
Przez indukcję z poprzedniego przykładu dostajemy dla macierzy blokowo trójkątnej:

$$\det \begin{pmatrix} A_1 & * & * & \dots & * \\ 0 & A_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_n \end{pmatrix} = \prod_{i=1}^{n} \det A_i \, .$$

(viii) Wyznacznik Vandermonde'a
Niech $z_1, \dots, z_n \in \mathbb{K}$. Wtedy

$$V_n := \begin{vmatrix} 1 & z_1 & z_1^2 & \dots & z_1^{n-1} \\ 1 & z_2 & z_2^2 & \dots & z_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & z_n & z_n^2 & \dots & z_n^{n-1} \end{vmatrix} = \prod_{1 \le j < i \le n} (z_i - z_j) \, .$$

Aby otrzymać ten wynik, w wyznaczniku V_n odejmujemy od n-tej kolumny kolumnę $(n-1)$-szą pomnożoną przez z_n, następnie od $(n-1)$-szej kolumny odejmujemy kolumnę $(n-2)$-gą pomnożoną przez z_n itd., aż od drugiej kolumny odejmiemy kolumnę pierwszą pomnożoną przez z_n. Rozwijając teraz otrzymany wyznacznik (równy wyjściowemu) względem ostatniego wiersza dostajemy

$$V_n = (-1)^{n+1} \begin{vmatrix} (z_1 - z_n) & (z_1 - z_n) z_1 & \dots & (z_1 - z_n) z_1^{n-2} \\ (z_2 - z_n) & (z_2 - z_n) z_2 & \dots & (z_2 - z_n) z_2^{n-2} \\ \dots & \dots & \dots & \dots \\ (z_{n-1} - z_n) & (z_{n-1} - z_n) z_{n-1} & \dots & (z_{n-1} - z_n) z_{n-1}^{n-2} \end{vmatrix} .$$

Wyciągając wspólne czynniki w wierszach dostajemy wzór rekurencyjny

$$V_n = (z_n - z_{n-1}) \dots (z_n - z_2)(z_n - z_1) V_{n-1} \, ,$$

z którego przez indukcję dostajemy żądany wynik.

(ix) Wyznacznik wielomianowy
Niech $a_0, \ldots, a_k, x \in \mathbb{K}$. Wtedy przy oznaczeniu $P_k(x) = a_0 + a_1 x + \ldots + a_k x^k$ mamy

$$\begin{vmatrix} a_0 & -x & 0 & \ldots & 0 \\ a_1 & 1 & -x & \ldots & 0 \\ a_2 & 0 & 1 & \ldots & 0 \\ \multicolumn{5}{c}{\dotfill} \\ a_{k-1} & 0 & 0 & \ldots & -x \\ a_k & 0 & 0 & \ldots & 1 \end{vmatrix} = \begin{vmatrix} x & 0 & \ldots & 0 & a_0 \\ -1 & x & \ldots & 0 & a_1 \\ 0 & -1 & \ldots & 0 & a_2 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \ldots & x & a_{k-1} \\ 0 & 0 & \ldots & -1 & a_k \end{vmatrix} = P_k(x)\,.$$

Aby otrzymać ten wynik rozwijamy drugą postać wyznacznika względem ostatniego wiersza, co daje wzór rekurencyjny

$$P_k(x) = a_k x^k + P_{k-1}(x)\,,$$

skąd przez indukcję otrzymujemy żądany rezultat.

6 Macierz odwrotna

Posługując się rozwinięciem Laplace'a udowodnimy następujący wynik.

Lemat 5. *Niech A będzie dowolną macierzą $n \times n$. Wtedy dla każdych liczb $i, j \in \{1, \ldots, n\}$:*

$$\sum_{k=1}^{n} A_{ik}\, \mathcal{A}_{jk}(A) = \sum_{k=1}^{n} A_{ki}\, \mathcal{A}_{kj}(A) = \det A\, \delta_{ij}\,.$$

Dowód. Dla $i = j$ twierdzenie sprowadza się do udowodnionego rozwinięcia Laplace'a. Jeśli $j \neq i$, to utwórzmy macierz A' przez zastąpienie w macierzy A jej j-tej kolumny przez kolumnę i-tą. Wyznacznik macierzy A' jest równy zero (ma ona dwie identyczne kolumny). Równocześnie rozwinięcie tego wyznacznika względem j-tej kolumny daje drugą z sum tezy. To kończy dowód drugiej równości tezy. Równość skrajnych stron tezy uzyskuje się podobną metodą zastosowaną do wierszy. □

Twierdzenie 6 (O macierzy odwrotnej)**.** *Macierz odwrotna do macierzy kwadratowej A istnieje wtedy, i tylko wtedy, gdy* $\det A \neq 0$*. Jeśli ten warunek jest spełniony, to*

$$\left(A^{-1}\right)_{ij} = \frac{\mathcal{A}_{ji}(A)}{\det A}\,.$$

Dowód. Jeśli istnieje A^{-1}, to z twierdzenia Cauchy'ego $\det A \det A^{-1} = 1$, więc $\det A \neq 0$. Odwrotnie, jeśli $\det A \neq 0$, to ostatni lemat mówi, że macierz A^{-1} zdefiniowana formułą obecnej tezy spełnia $AA^{-1} = A^{-1}A = \mathbf{1}$, więc istotnie jest macierzą odwrotną. □

Zgodnie z ogólną definicją elementów odwrotnych macierze A i B są wzajemnie odwrotne, gdy $AB = \mathbf{1}$ i $BA = \mathbf{1}$. Jednak twierdzenie Cauchy'ego w połączeniu z twierdzeniem o macierzy odwrotnej daje następujące uproszczenie.

Wniosek 7. *Jeśli macierze kwadratowe A, B spełniają $AB = \mathbf{1}$, to obie są nieosobliwe i wzajemnie odwrotne.*

Dowód. Jeśli $AB = \mathbf{1}$, to z twierdzenia Cauchy'ego obie macierze są nieosobliwe. Mnożąc obie strony przez A^{-1} z lewej lub B^{-1} z prawej dostajemy tezę. □

Przy zastosowaniu ostatniego wniosku i blokowej postaci macierzy udowodnimy pomocnicze twierdzenie, które okaże się przydatne w dalszym ciągu wykładu.

Lemat 8. *Jeśli macierze $A \in \mathbb{K}^{m\times n}$ i $B \in \mathbb{K}^{n\times m}$ spełniają warunek $AB = \mathbf{1}_m$, to $m \leq n$. Stąd, jeśli również $BA = \mathbf{1}_n$, to $m = n$ i macierze A i B są wzajemnie odwrotne.*

Dowód. Dowód nie wprost. Załóżmy, że $m > n$. Wtedy danym macierzom można nadać następującą postać blokową:

$$A = \begin{pmatrix} A' \\ A'' \end{pmatrix}, \qquad B = \begin{pmatrix} B' & B'' \end{pmatrix},$$

gdzie A' i B' są macierzami $n \times n$. Warunek $AB = \mathbf{1}_m$ zapisuje się przy tych oznaczeniach w postaci blokowej:

$$\begin{pmatrix} A'B' & A'B'' \\ A''B' & A''B'' \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n & 0 \\ 0 & \mathbf{1}_{m-n} \end{pmatrix},$$

która jest równoważna czterem warunkom macierzowym:

$$A'B' = \mathbf{1}_n\,, \quad A'B'' = 0\,, \quad A''B' = 0\,, \quad A''B'' = \mathbf{1}_{m-n}\,.$$

Z pierwszego równania, na podstawie ostatniego wniosku wynika, że macierze kwadratowe A' i B' są nieosobliwe. Mnożąc drugie równanie przez A'^{-1} z lewej strony, a trzecie przez B'^{-1} z prawej strony, dostajemy $B'' = 0$ i $A'' = 0$, co jest sprzeczne z ostatnim warunkiem. Dowód pierwszego stwierdzenia lematu jest tym samym zakończony. Jeśli spełniony jest ponadto drugi warunek założenia, to również $n \leq m$, więc $m = n$. □

7 Wzory Cramera

Poznaliśmy wcześniej pojęcie ogólnego układu równań liniowych i jego zapis macierzowy $AX = B$. W ogólności liczba równań w układzie m może być różna od liczby niewiadomych n. Załóżmy, że $m = n$ – wtedy macierz główna układu A jest macierzą kwadratową. Jeśli przy tym A jest nieosobliwa, to układ taki nazywamy ***układem Cramera***.

Twierdzenie 9. *Układ Cramera $AX = B$ ma dokładnie jedno rozwiązanie dane wzorami*

$$x_i = \frac{\det A_i}{\det A}\,,$$

gdzie macierz A_i powstaje z macierzy A przez zastąpienie jej i-tej kolumny kolumną B.

Dowód. Jeśli $\det A \neq 0$, to mnożąc równanie $AX = B$ przez A^{-1} z lewej strony dostajemy $X = A^{-1}B$. Odwrotnie, mnożąc podobnie ostatnie równanie przez A dostajemy równanie pierwsze – równania są równoważne. Rozwiązanie istnieje więc i jest jednoznaczne. Zapiszmy jawnie ostatnią równość:

$$x_i = \sum_k (A^{-1})_{ik} B_k = (\det A)^{-1} \sum_k \mathcal{A}_{ki}(A) B_k\,.$$

Występująca tu po prawej stronie suma jest rozwinięciem względem i-tej kolumny wyznacznika macierzy A_i określonej w twierdzeniu, skąd teza. □

8 Przykłady

(i) Macierz odwrotna

Wyznacznik macierzy $\begin{pmatrix} 0 & 1 & -1 \\ 3 & 2 & -1 \\ 5 & 1 & 4 \end{pmatrix}$ jest równy -10, więc jest to macierz nieosobliwa. Macierz do niej odwrotna jest równa

$$-\frac{1}{10}\begin{pmatrix} +\begin{vmatrix} 2 & -1 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} & +\begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} \\ -\begin{vmatrix} 3 & -1 \\ 5 & 4 \end{vmatrix} & +\begin{vmatrix} 0 & -1 \\ 5 & 4 \end{vmatrix} & -\begin{vmatrix} 0 & -1 \\ 3 & -1 \end{vmatrix} \\ +\begin{vmatrix} 3 & 2 \\ 5 & 1 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 5 & 1 \end{vmatrix} & +\begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix} \end{pmatrix} = \frac{1}{10}\begin{pmatrix} -9 & 5 & -1 \\ 17 & -5 & 3 \\ 7 & -5 & 3 \end{pmatrix}.$$

(ii) Macierz odwrotna do macierzy blokowo trójkątnej
Macierz odwrotna do macierzy blokowo trójkątnej ma następującą postać:

$$A = \begin{pmatrix} A_1 & * & * & \dots & * \\ 0 & A_2 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_n \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} A_1^{-1} & * & * & \dots & * \\ 0 & A_2^{-1} & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_n^{-1} \end{pmatrix}$$

– najprościej dowieść tego przez indukcję względem liczby bloków, wprost z definicji (a nie z twierdzenia o macierzy odwrotnej). W szczególności, macierz odwrotna do trójkątnej jest trójkątna.

(iii) Przykład macierzy niekwadratowych, spełniających warunki lematu 8
Niech

$$A = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} -2\beta & 1-2\delta \\ \alpha & \gamma \\ \beta & \delta \\ 1-2\alpha & -2\gamma \end{pmatrix}.$$

Każda macierz B o powyższej ogólnej postaci spełnia warunek $AB = \mathbf{1}_2$.

(iv) Układ Cramera
Wyznacznik macierzy głównej układu

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

jest równy 20, więc jest to układ Cramera. Stosując wzory Cramera dostajemy:

$$x = \frac{1}{20} \begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 1 \end{vmatrix} = \frac{5}{20}, \quad y = \frac{1}{20} \begin{vmatrix} 1 & 1 & 2 \\ 3 & 0 & 1 \\ -1 & -1 & 1 \end{vmatrix} = \frac{-9}{20},$$

$$z = \frac{1}{20} \begin{vmatrix} 1 & -1 & 1 \\ 3 & 2 & 0 \\ -1 & 2 & -1 \end{vmatrix} = \frac{3}{20}.$$

(v) Dowód twierdzenia interpolacyjnego Lagrange'a
Załóżmy, jak w założeniach twierdzenia 14, §3, że dane są dwa ciągi elementów ciała $\mathbb{K}$: $(x_1, \dots, x_{n+1})$ i $(p_1, \dots, p_{n+1})$, przy czym pierwszy z tych ciągów jest różnowartościowy, $x_i \neq x_j$ dla $i \neq j$. Szukamy wielomianu

$$P(x) = a_0 + a_1 x + \dots + a_n x^n \quad \text{takiego, że} \quad P(x_i) = p_i\,, \quad i = 1, \dots, n\,.$$

Jawne wypisanie tego układu równań pokazuje, że jest to układ Cramera ze względu na współczynniki $a_0, \ldots, a_n$ (wyznacznik macierzy głównej jest różnym od zera wyznacznikiem Vandermonde'a), więc ma jednoznaczne rozwiązanie.

9 Pfaffian

W tym punkcie skupimy uwagę na macierzach kwadratowych antysymetrycznych, o elementach w ciele $\mathbb{K}$ o charakterystyce różnej od 2. Jeśli A jest taką macierzą o wymiarze n, to z własności wyznacznika dostajemy

$$\det A = \det A^{\mathrm{T}} = \det(-A) = (-1)^n \det A\,.$$

Stąd, jeśli n jest nieparzyste i $\operatorname{char}\mathbb{K} \neq 2$, to $\det A = 0$.

Dla macierzy antysymetrycznych parzystego wymiaru określa się wielkość spokrewnioną z wyznacznikiem. Niech A będzie macierzą antysymetryczną wymiaru $2p$, o elementach w ciele $\mathbb{K}$, $\operatorname{char}\mathbb{K} \neq 2$. ***Pfaffianem*** macierzy A nazywamy wielomian względem jej elementów

$$\operatorname{Pf} A = {\sum}' \epsilon_{i_1\ldots i_p\, j_1\ldots j_p} A_{i_1 j_1} \ldots A_{i_p j_p}\,,$$

gdzie wszystkie wskaźniki sumacyjne należą do zbioru $\{1, \ldots, 2p\}$ i zakres sumowania jest dodatkowo ograniczony warunkami:

$$i_1 < \ldots < i_p\,, \qquad i_1 < j_1\,, \quad \ldots, \quad i_p < j_p\,,$$

co zasygnalizowaliśmy primem. Dla konkretnej macierzy jej pfaffianem nazywamy też w skrócie wartość tego wielomianu. (Spotyka się również w literaturze definicję odpowiadającą zamianie we wzorze definicyjnym symbolu $\epsilon_{i_1\ldots i_p\, j_1\ldots j_p}$ na $\epsilon_{i_1\, j_1\ldots i_p\, j_p}$, co skutkuje zmianą pfaffianu o czynnik $(-1)^{p(p-1)/2}$.)

Jeśli charakterystyka ciała $\mathbb{K}$ jest większa od p, a więc $p! \neq 0$, to pfaffian można wyrazić równoważnie za pomocą formuły

$$\operatorname{Pf} A = \frac{1}{2^p p!} \sum_{i_1, j_1, \ldots, i_p, j_p} \epsilon_{i_1\ldots i_p\, j_1\ldots j_p} A_{i_1 j_1} \ldots A_{i_p j_p}\,,$$

gdzie teraz wszystkie wskaźniki sumacyjne przebiegają wartości od 1 do $2p$. Aby się o tym przekonać zauważamy najpierw, że przestawiając wskaźniki w parach i_k, j_k ($k = 1, \ldots, p$) i zamieniając je nazwami możemy każdy z niezerowych wyrazów ostatniej sumy sprowadzić do postaci, w której $i_k < j_k$ (korzystamy z antysymetrii macierzy A_{ij} oraz symbolu Levi-Civity). To pokazuje, że mamy po 2^p wyrazów identycznych w sumie, i pozwala ograniczyć zakres wskaźników warunkami $i_k < j_k$, $k = 1, \ldots, p$, przy równoczesnym pominięciu czynnika $1/2^p$.

Istnieje dalsza symetria sumowanych wyrazów: ich wartość nie zmienia się przy poddaniu ciągów $(i_1, \ldots, i_p)$ oraz $(j_1, \ldots, j_p)$ takiej samej permutacji. Wybieramy więc jeden wyraz z każdej grupy, ten dla którego $i_1 < \ldots < i_p$, i pomijamy czynnik $1/p!$, co sprowadza sumę do postaci przyjętej przez nas w definicji.

Na potrzeby następnego twierdzenia zauważmy, że jeśli A jest macierzą antysymetryczną, to (i) jej i-ty wiersz jest równy zeru wtedy, i tylko wtedy, gdy zeruje się również jej i-ta kolumna; (ii) dla dowolnej macierzy kwadratowej γ macierz $\gamma^{\mathrm{T}} A\gamma$ również jest antysymetryczna.

Twierdzenie 10. *Pfaffian macierzy antysymetrycznej (o elementach w* $\mathbb{K}$*,* $\operatorname{char}\mathbb{K} \neq 2$*) ma następujące własności (*A *jest macierzą antysymetryczną parzystego wymiaru, a* γ *– dowolną macierzą kwadratową tego samego wymiaru):*

(i) *jeśli wiersz i odpowiednia kolumna macierzy* A *są równe zeru, to*
$$\operatorname{Pf} A = 0\,;$$

(ii)
$$\operatorname{Pf}\begin{pmatrix} 0 & \mathbf{1}_p \\ -\mathbf{1}_p & 0 \end{pmatrix} = 1\,;$$

(iii)
$$\operatorname{Pf}(\gamma^{\mathrm{T}} A\gamma) = \operatorname{Pf} A \, \det\gamma\,.$$

Dowód. (i) W tym przypadku wszystkie składniki sumy określającej pfaffian są równe zeru, bo jeśli $\epsilon_{i_1\ldots i_p\, j_1\ldots j_p} \neq 0$, to któryś z czynników $A_{i_1 j_1}, \ldots, A_{i_p j_p}$ nosi numer zerującego się wiersza lub zerującej się kolumny.

(ii) Oznaczmy przez E macierz będącą argumentem pfaffianu w drugim punkcie. Jej elementy są równe $E_{ij} = \delta_{(i+p)\,j} - \delta_{i\,(j+p)}$. Zauważamy, że drugi wyraz w tej definicji nie znika tylko dla $i = j + p > j$, więc nie daje przyczynku po wstawieniu do definicji pfaffianu. Wykonanie sumowania po j-tach daje więc
$$\operatorname{Pf} E = \sum_{1\leq i_1<\ldots<i_p\leq p} \epsilon_{i_1\ldots i_p\,(i_1+p)\ldots(i_p+p)} = \epsilon_{1\ldots 2p} = 1\,.$$

(iii) Ograniczamy się do przypadku $\operatorname{char}\mathbb{K} > p$ i używamy drugiego z podanych przedstawień pfaffianu (jako ćwiczenie polecamy dostosowanie dowodu do ogólnego przypadku). Wstawiając w tym wyrażeniu
$$(\gamma^{\mathrm{T}} A\gamma)_{i_s j_s} = \sum_{k_s, l_s} \gamma_{k_s i_s}\gamma_{l_s j_s} A_{k_s l_s}\,,$$
dostajemy
$$\operatorname{Pf}(\gamma^{\mathrm{T}} A\gamma) = \frac{1}{2^p p!}\sum \epsilon_{i_1\ldots i_p\, j_1\ldots j_p}\gamma_{k_1 i_1}\ldots\gamma_{k_p i_p}\gamma_{l_1 j_1}\ldots\gamma_{l_p j_p} A_{k_1 l_1}\ldots A_{k_p l_p} =$$
$$= \det\gamma \frac{1}{2^p p!}\sum \epsilon_{k_1\ldots k_p\, l_1\ldots l_p} A_{k_1 l_1}\ldots A_{k_p l_p} = \det\gamma \operatorname{Pf} A\,,$$

gdzie sumowanie w krokach pośrednich przebiega po wszystkich wskaźnikach. □

Niech E nadal oznacza macierz będącą argumentem pfaffianu w punkcie (ii) powyższego twierdzenia. Macierz ta sprowadza się do macierzy $\begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_p \end{pmatrix}$, której wyznacznik jest równy $(-1)^p$, przez przestawienie kolejno kolumn $(p+1)$-szej, …, $2p$-tej przed pierwszych p kolumn, a więc złożenie p^2 transpozycji. Stąd $\det E = (-1)^{p(p+1)} = 1$.

Niech teraz G będzie macierzą blokową

$$G = \begin{pmatrix} E & 0 \\ 0 & 0_{2r} \end{pmatrix} .$$

Z ostatniego twierdzenia oraz z twierdzenia Cauchy'ego dla wyznaczników dostajemy związki

$$\operatorname{Pf}(\gamma^{\mathrm{T}} G \gamma) = \operatorname{Pf} G \, \det \gamma \,, \quad \det(\gamma^{\mathrm{T}} G \gamma) = \det G \, (\det \gamma)^2 \,.$$

Jeśli diagonalny blok zerowy w macierzy G znika, to $\operatorname{Pf} G = \det G = 1$, a jeśli $r > 0$, to $\operatorname{Pf} G = \det G = 0$. W każdym przypadku dla macierzy antysymetrycznej o postaci $A = \gamma^{\mathrm{T}} G \gamma$ otrzymujemy więc zależność

$$\det A = (\operatorname{Pf} A)^2 \,.$$

Zobaczymy później (w punkcie 6, §14), że każda macierz antysymetryczna parzystego wymiaru może być przedstawiona w założonej postaci, więc ten związek między wyznacznikiem a pfaffianem jest prawdziwy dla wszystkich macierzy antysymetrycznych parzystego wymiaru.

10 Podgrupy ogólnej grupy liniowej

Określimy w tym punkcie pewną klasę grup macierzowych; wszystkie one są podgrupami grup $GL(n, \mathbb{K})$.

Twierdzenie 11. *Niech będzie dana nieosobliwa macierz $G \in \mathbb{K}^{n \times n}$. Zbiór macierzy $A \in \mathbb{K}^{n \times n}$ spełniających warunek*

$$A^{\mathrm{T}} G A = G$$

tworzy podgrupę grupy $GL(n, \mathbb{K})$. Jeśli $\mathbb{K} = \mathbb{C}$, to macierze spełniające warunek

$$A^{\dagger} G A = G$$

również tworzą podgrupę grupy $GL(n, \mathbb{C})$.

Dowód. Z warunków definicyjnych powyższych zbiorów macierzy i z twierdzenia Cauchy'ego widzimy, że zadane macierze są nieosobliwe, więc należą do $GL(n,\mathbb{K})$. Zgodnie z kryterium 6, §3, wystarczy więc pokazać, że jeśli A i B spełniają warunek definicyjny, to również A^{-1} i AB spełniają ten sam warunek. Przeprowadzimy dowód dla przypadku pierwszego warunku, dowód dla przypadku warunku drugiego przebiega analogicznie. Mnożąc warunek definicyjny z lewej przez $(A^{-1})^{\mathrm{T}}$, a z prawej przez A^{-1}, i korzystając z własności $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$ dostajemy $(A^{-1})^{\mathrm{T}}GA^{-1} = G$, więc A^{-1} spełnia warunek definicyjny. Dla iloczynu macierzy A i B spełniających warunek mamy – na mocy własności transpozycji oraz warunków definicyjnych kolejno dla A i B – ciąg równości

$$(AB)^{\mathrm{T}}G(AB) = B^{\mathrm{T}}A^{\mathrm{T}}GAB = B^{\mathrm{T}}GB = G\,,$$

więc AB również spełnia warunek definicyjny. □

Dla ogólnej macierzy nieosobliwej G warunki definicyjne w powyższym twierdzeniu są zbyt ograniczające. Ponadto istnieje wiele izomorfizmów pomiędzy grupami zadanymi tym twierdzeniem. Podstawowe znaczenie mają następujące grupy tego typu (przekonamy się o ich algebraicznym znaczeniu w §16).

(i) Grupy $O(p,q,\mathbb{K})$, $p,q \in \mathbb{N}\cup\{0\}$, $n = p+q \geq 1$, z definicji złożone są z macierzy $n\times n$ spełniających warunek

$$A^{\mathrm{T}}\begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} A = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix}.$$

Z twierdzenia Cauchy'ego widzimy, że $(\det A)^2 = 1$, więc $\det A = \pm 1$. Oba przypadki znaku wyznacznika są realizowane, gdyż – jak łatwo sprawdzić – każda macierz diagonalna o dowolnie na diagonali rozłożonych liczbach $+1$ i -1 spełnia warunek definicyjny tej grupy. Gdy $\mathbb{K} = \mathbb{R}$, piszemy $O(p,q)$.

(ii) Na wyróżnienie zasługuje szczególny przypadek grup tego rodzaju otrzymany przez położenie $q = 0$. Grupy te, nazywane ***grupami ortogonalnymi*** i oznaczane $O(n,\mathbb{K})$, składają się z macierzy spełniających warunek

$$A^{\mathrm{T}}A = \mathbf{1}\,.$$

Macierze takie nazywamy ***macierzami ortogonalnymi***. Gdy $\mathbb{K} = \mathbb{R}$, to mówimy o rzeczywistej grupie ortogonalnej i piszemy $O(n)$.

(iii) Grupy $U(p,q)$, $p,q \in \mathbb{N}\cup\{0\}$, $n = p+q \geq 1$, z definicji złożone są z macierzy zespolonych $n\times n$ spełniających warunek

$$A^{\dagger}\begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} A = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix}.$$

Z twierdzenia Cauchy'ego widzimy, że $|\det A| = 1$, i można pokazać (zostawiamy to zadanie jako ćwiczenie), że zbiór wartości $\det A$ pokrywa wszystkie liczby o module jednostkowym.

(iv) ***Grupy unitarne*** $U(n)$ składają się z macierzy zespolonych $n \times n$, spełniających warunek

$$A^\dagger A = \mathbf{1}\,,$$

co jest szczególnym przypadkiem punktu poprzedniego ($q = 0$). Macierze takie nazywamy ***macierzami unitarnymi***.

(v) ***Grupy symplektyczne*** $Sp(2p, \mathbb{K})$, gdzie $\operatorname{char}\mathbb{K} \neq 2$, złożone są z macierzy $2p \times 2p$, spełniających warunek

$$A^{\mathrm{T}} \begin{pmatrix} 0 & \mathbf{1}_p \\ -\mathbf{1}_p & 0 \end{pmatrix} A = \begin{pmatrix} 0 & \mathbf{1}_p \\ -\mathbf{1}_p & 0 \end{pmatrix}.$$

Wyliczając pfaffian obu stron tego warunku i korzystając z twierdzenia 10, dostajemy $\det A = 1$. Grupę rzeczywistych macierzy symplektycznych notujemy $Sp(2p)$.

Z definicji ciała wiemy, że $\mathbb{K}^*$ stanowi grupę multiplikatywną. Twierdzenie Cauchy'ego mówi, że odwzorowanie $\det : GL(n, \mathbb{K}) \mapsto \mathbb{K}^*$ jest homomorfizmem grup. Jądro tego homomorfizmu, czyli zbiór macierzy o wyznaczniku równym 1, stanowi podgrupę grupy $GL(n, \mathbb{K})$; nazywamy ją ***specjalną grupą liniową (nad ciałem*** $\mathbb{K}$***)*** i oznaczamy $SL(n, \mathbb{K})$. Grupa symplektyczna $Sp(2p, \mathbb{K})$ jest podgrupą specjalnej grupy liniowej $SL(2p, \mathbb{K})$, a w szczególnym przypadku $p = 1$ jest z nią identyczna — $Sp(2, \mathbb{K}) = SL(2, \mathbb{K})$ (patrz następny punkt).

Przecięcia specjalnych grup liniowych z grupami określonymi w (i) – (iv) dają dalsze grupy macierzowe:

(i)′ $SO(p, q, \mathbb{K}) := O(p, q, \mathbb{K}) \cap SL(n, \mathbb{K})$;

(ii)′ $SO(n, \mathbb{K}) := O(n, \mathbb{K}) \cap SL(n, \mathbb{K})$ – ***specjalna grupa ortogonalna***;
$SO(n) := O(n) \cap SL(n, \mathbb{R})$;

(iii)′ $SU(p, q) := U(p, q) \cap SL(n, \mathbb{C})$;

(iv)′ $SU(n) := U(n) \cap SL(n, \mathbb{C})$ – ***specjalna grupa unitarna***.

W przypadku rzeczywistej grupy liniowej wyznaczniki macierzy są liczbami rzeczywistymi różnymi od zera, a więc albo dodatnimi, albo ujemnymi. Twierdzenie Cauchy'ego pokazuje, że podzbiór grupy $GL(n, \mathbb{R})$ złożony z macierzy o wyznaczniku dodatnim, $GL(n, \mathbb{R})_+$, jest podgrupą. Podzbiór złożony z macierzy o wyznaczniku ujemnym, $GL(n, \mathbb{R})_-$, nie stanowi grupy.

11 Przykłady

(i) Pfaffiany niskich stopni
Wprost z definicji wyliczamy pfaffiany macierzy antysymetrycznych A najniższych wymiarów:

$$2\times 2: \ \mathrm{Pf}\, A = A_{12}\,, \quad 4\times 4: \ \mathrm{Pf}\, A = A_{12}A_{43} + A_{13}A_{24} + A_{14}A_{32}\,.$$

(ii) Związek pfaffianu z wyznacznikiem
Otrzymany na końcu punktu 9 związek $\det A = (\mathrm{Pf}\, A)^2$ wykazaliśmy jako równość funkcyjną (przedstawienie $A = \gamma^{\mathrm{T}} G\gamma$, do którego się tam odwołujemy, ma charakter równości *wartości* obu stron). Ale obie strony tej równości są funkcjami wielomianowymi drugiego stopnia względem każdego z niezależnych elementów antysymetrycznej macierzy A, więc są równe także jako wielomiany względem elementów macierzy (należy wykorzystać rozszerzenie wyniku przykładu (ii), p. 15, §3, na przypadek wielu zmiennych – przez indukcję względem liczby zmiennych).

(iii) Macierze ortogonalne i unitarne
Zapiszmy macierz kwadratową A w postaci blokowej kolumnowej:

$$A = (X_1 X_2 \ldots X_n)\,.$$

Macierz ta jest ortogonalna (lub unitarna), gdy kolumny spełniają warunki: $X_i^{\mathrm{T}} X_j = \delta_{ij}$ (lub, odpowiednio, $X_i^{\dagger} X_j = \delta_{ij}$).

(iv) Grupa $O(1)$
Ta grupa jest multiplikatywną grupą liczb $\{1, -1\}$. Grupa $SO(1)$ składa się z samego elementu jednostkowego 1.

(v) Grupa $SO(2)$
Na tę grupę składają się rzeczywiste macierze 2×2 o wyznaczniku równym 1, spełniające warunek $A^{-1} = A^{\mathrm{T}}$. Ogólna postać macierzy spełniających te warunki to $R = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, gdzie liczby a i b spełniają warunek $a^2 + b^2 = 1$. Ostatni warunek ma ogólne rozwiązanie: $a = \cos\varphi$, $b = \sin\varphi$, gdzie $\varphi \in \langle 0, 2\pi)$. Każdy element grupy $SO(2)$ ma więc postać

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\,.$$

Mnożenie macierzy tego typu spełnia równanie $R(\varphi)R(\varphi') = R(\varphi + \varphi')$. Stąd pokazuje się, że odwzorowanie

$$\mathbb{R}_{2\pi} \ni [\varphi] \mapsto R(\varphi) \in SO(2)$$

jest izomorfizmem. Grupa $SO(2)$ jest izomorficzna z grupą $\mathbb{R}_{2\pi}$.

(vi) Zbiór $O(2) \setminus SO(2)$
Elementy grupy $O(2)$ o wyznaczniku ujemnym (więc równym -1) mają postać

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}, \quad \varphi \in \langle 0, 2\pi) .$$

(vii) Grupa $U(1)$
Multiplikatywna grupa liczb zespolonych o module równym 1. Grupa ta jest izomorficzna z grupą $\mathbb{R}_{2\pi}$, dzięki istnieniu izomorfizmu

$$\mathbb{R}_{2\pi} \ni [\varphi] \mapsto e^{i\varphi} \in U(1) .$$

(viii) Grupa $SU(2)$
Grupa składa się z macierzy zespolonych o postaci $\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$, gdzie liczby a i b spełniają warunek $|a|^2 + |b|^2 = 1$. Każda macierz grupy $U(2)$ może być zapisana w postaci αU, gdzie $U \in SU(2)$, a α jest liczbą zespoloną o module równym 1.

(ix) Grupa $Sp(2, \mathbb{K})$
Niech A będzie macierzą 2×2. Bezpośrednim rachunkiem łatwo sprawdzić, że

$$A^{\mathrm{T}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A = \det A \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

Stąd $Sp(2, \mathbb{K}) = SL(2, \mathbb{K})$.

(x) Grupa $SL(2, \mathbb{C})$ i grupa homografii
Niech $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ będzie macierzą z grupy $SL(2, \mathbb{C})$, a więc jej elementy spełniają równanie $ad - bc = 1$. Z warunku tego wynika, że liczby a, b, c, d można przyjąć za współczynniki pewnej homografii zmiennej zespolonej (patrz przykład (i), punkt 3, §5), co zadaje odwzorowanie

$$SL(2, \mathbb{C}) \mapsto \text{ grupa homografii} .$$

Przez bezpośrednie porównanie prawa składania w każdej z tych grup sprawdza się, że odwzorowanie to jest homomorfizmem. Homomorfizm ten jest surjektywny, gdyż dowolna homografia o współczynnikach $\alpha, \beta, \gamma, \delta$ jest obrazem macierzy $\frac{1}{\sqrt{\alpha\delta - \beta\gamma}} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2, \mathbb{C})$. Jądrem tego homomorfizmu jest niezmiennicza podgrupa $\{\mathbf{1}, -\mathbf{1}\} \subseteq SL(2, \mathbb{C})$. Homomorfizm ten zadaje więc izomorfizm grup

$$SL(2, \mathbb{C})/\{\mathbf{1}, -\mathbf{1}\} \mapsto \text{ grupa homografii}\,.$$

II

PRZESTRZENIE WEKTOROWE

§8 Podstawowe pojęcia

1 Przestrzeń wektorowa

Przystępujemy do określenia podstawowego obiektu algebry liniowej – algebraicznej struktury zadającej jej zakres.

Niech będzie dane ciało $\mathbb{K}$ oraz zbiór V. Załóżmy, że na zbiorze V określono dwa działania:

(i) ***dodawanie*** (działanie wewnętrzne)

$$+ : V \times V \ni (x, y) \mapsto x + y \in V \,;$$

(ii) ***mnożenie przez liczbę*** (działanie zewnętrzne)

$$\cdot : \mathbb{K} \times V \ni (\lambda, x) \mapsto \lambda x \in V \,.$$

Utworzona została w ten sposób struktura algebraiczna $(V, \mathbb{K}, +, \cdot)$. Mówimy, że V jest ***przestrzenią wektorową*** (lub ***przestrzenią liniową***) ***nad ciałem*** $\mathbb{K}$, gdy ta struktura ma następujące własności:

(i) $(V, +)$ jest grupą abelową;

(ii) $\forall x \in V \; \forall \alpha, \beta \in \mathbb{K}$:

$$1x = x \,, \quad \alpha(\beta x) = (\alpha\beta)x \quad (\textbf{\textit{prawo łączności}});$$

(iii) $\forall x, y \in V \; \forall \alpha, \beta \in \mathbb{K}$:

$$(\alpha + \beta)x = \alpha x + \beta x \,, \quad \alpha(x + y) = \alpha x + \alpha y \quad (\textbf{\textit{prawa rozdzielności}}).$$

Elementy przestrzeni wektorowej nazywamy ***wektorami***. Element neutralny dodawania wektorów (czyli element neutralny grupy abelowej $(V, +)$) nazywamy ***wektorem zerowym*** i oznaczamy $\vec{0}$, a element odwrotny w tej grupie do wektora x nazywamy ***wektorem przeciwnym do*** x i oznaczamy $-x$. Ponieważ każda grupa ma element neutralny, również przestrzeń wektorowa ma co najmniej jeden element – wektor zerowy. Bezpośrednio z definicji przestrzeni wektorowej otrzymujemy jej dalsze podstawowe własności.

Twierdzenie 1. *W każdej przestrzeni wektorowej V prawdziwe są następujące własności:*

(iv) dla każdego $x \in V$:

$$(-1)x = -x\,;$$

(v) dla każdych $\alpha \in \mathbb{K}$ i $x \in V$:

$$\alpha x = \vec{0} \iff \alpha = 0 \lor x = \vec{0}\,.$$

Dowód. Udowodnimy najpierw punkt drugi.

(v) $\Leftarrow$ Dla $\alpha = 0$ mamy $0x = (0 + 0)x = 0x + 0x$ (z rozdzielności). Dodając do obu stron wektor przeciwny do wektora $0x$ dostajemy $\vec{0} = 0x$. Podobnie dla przypadku $x = \vec{0}$: $\alpha\vec{0} = \alpha(\vec{0} + \vec{0}) = \alpha\vec{0} + \alpha\vec{0}$, więc $\alpha\vec{0} = \vec{0}$.
$\Rightarrow$ Niech $\alpha x = \vec{0}$. Jeśli $\alpha = 0$, to następnik prawdziwy. Jeśli $\alpha \neq 0$, to mnożymy założone równanie przez α^{-1} i dostajemy $x = \alpha^{-1}\vec{0} = \vec{0}$ na mocy łączności oraz udowodnionego wyżej wynikania w lewą stronę. Następnik i w tym przypadku jest prawdziwy.

(iv) Dla każdego x mamy ciąg równości

$$x + (-1)x = 1x + (-1)x = (1 - 1)x = 0x = \vec{0}\,, \quad \text{a stąd} \quad -x = (-1)x\,.$$

□

Zanotujmy, że własność (iv) w połączeniu z łącznością daje

$$(-\alpha)x = \alpha(-x) = -\alpha x\,.$$

Zwykle zamiast $\vec{0}$ piszemy dla prostoty 0 – z kontekstu jest na ogół oczywiste, czy w danym przypadku chodzi o element zerowy w grupie, czy w ciele. Z drugiej strony, niekiedy stosuje się notację, w której wszystkie wektory notuje się ze strzałką: $\vec{x}$; w tym podręczniku nie używamy tej notacji.

Przestrzeń wektorową nad ciałem liczb rzeczywistych nazywamy ***rzeczywistą przestrzenią wektorową***, a przestrzeń utworzoną nad ciałem liczb zespolonych – ***zespoloną przestrzenią wektorową***.

2 Kombinacja liniowa

Składając działania mnożenia wektorów przez liczbę i dodawania wektorów otrzymujemy pojęcie kombinacji liniowej. Wybierzmy dowolny skończony ciąg $x_1, \ldots, x_k$ wektorów przestrzeni V oraz dowolny ciąg $\lambda_1, \ldots, \lambda_k$ liczb ciała $\mathbb{K}$ i utwórzmy wektor

$$x = \sum_{i=1}^{k} \lambda_i \, x_i \, .$$

Wektor tak określony nazywamy ***kombinacją liniową wektorów*** $x_1, \ldots, x_k$. Liczby $\lambda_1, \ldots, \lambda_k$ nazywamy ***współczynnikami kombinacji liniowej***. Ze względu na przemienność dodawania jest widoczne, że wartość kombinacji liniowej nie zależy od wyboru wspólnej numeracji wektorów i liczb, a tylko od przypisania współczynników wektorom. Jeśli ciąg $x_1, \ldots, x_k$ jest różnowartościowy, to mówimy też, że x jest ***kombinacją liniową wektorów rodziny*** $\{x_1, \ldots, x_k\} \subseteq V$. Jeśli F jest dowolną rodziną wektorów w V, niekoniecznie skończoną, i $\{x_1, \ldots, x_k\} \subseteq F$, to x nazywamy ***kombinacją liniową wektorów rodziny*** F. Mówimy też wtedy inaczej, że powyższa równość daje ***rozkład wektora*** x ***na wektory rodziny*** F. Podkreślmy, że w ogólności w ciągu $x_1, \ldots, x_k$ wektory mogą się powtarzać, jednak gdy mówimy o kombinacji liniowej wektorów *rodziny*, to mamy zawsze na myśli ciąg różnowartościowy.

Iloczyn λx wektora x przez liczbę λ wygodnie będzie niekiedy notować w odwrotnym porządku, przyjmiemy więc $x\lambda \equiv \lambda x$. Jeśli wiadomo, który symbol opisuje wektor, a który liczbę, to ta dodatkowa swoboda nie prowadzi do niejednoznaczności. Kombinację liniową możemy więc zgodnie z tą konwencją zapisać również jako

$$x = \sum_{i=1}^{k} x_i \, \lambda_i \, .$$

Pojęcie przestrzeni wektorowej ujmuje w precyzyjny język matematyczny ideę zasady superpozycji (nazywanej też zasadą liniowości). Zasada ta stwierdza, że obiekty w rozpatrywanej klasie można mnożyć przez liczbę i dodawać, otrzymując znów obiekty w tej samej klasie (odpowiada to tworzeniu kombinacji liniowej w przestrzeni wektorowej). Jest uderzające, jak powszechne zastosowanie ma ta zasada (a więc również jej precyzyjne ujęcie w postaci przestrzeni wektorowej) do opisu zjawisk świata materialnego. Z jednej strony językiem wielu teorii fizycznych, modeli fizycznych, chemicznych, ekonomicznych itd., jest formalizm przestrzeni wektorowych. Z drugiej strony, nawet w przypadkach, gdy język ten okazuje się zbyt prosty dla dokładnego opisu rzeczywistości (zjawiska są, jak mówimy, nieliniowe), liniowe przybliżenia tego opisu odgrywają ogromną rolę w zrozumieniu teorii, a także są punktem wyjścia dla konstrukcji dokładniejszego opisu.

3 Modele przestrzeni wektorowych

Mówiąc o przykładach przestrzeni wektorowych, należy zrobić wcześniej pewną ogólną uwagę (która dotyczy zresztą również innych struktur algebraicznych). Przestrzeń wektorową określiliśmy aksjomatycznie, jako strukturę spełniającą pewne postulaty, ale poza tym zupełnie dowolną. Przez przykłady przestrzeni wektorowych rozumiemy istniejące obiekty *matematyczne*, dla których można prawdziwość tych postulatów *wykazać na drodze dedukcji*. Kilka takich przykładów podamy poniżej. Ich istnienie gwarantuje, że postulaty nie są sprzeczne. Badanie konsekwencji postulatów prowadzimy jednak na poziomie ogólnym, dodając w miarę potrzeby następne własności, czyniące strukturę bardziej szczegółową.

Jeśli myślimy o matematyce nie tylko jako o przedmiocie dla samego siebie, ale również jako o źródle modeli stosowanych do opisu rzeczywistości materialnej, to często za model określonego fragmentu rzeczywistości obieramy pewną strukturę matematyczną, np. przestrzeń wektorową, w jej abstrakcyjnej postaci. Taki model matematyczny może dobrze spełniać swoją rolę, pozwalając dokonywać w oparciu o niego trafne przewidywania. Błędny byłby jednak wniosek, że opisany w ten sposób fragment rzeczywistości jest przykładem modelującej go struktury matematycznej w sensie opisanym powyżej. Trafności postulatów dla tego wycinka rzeczywistości nie da się dowieść drogą samego rozumowania. Mierzy się ją zgodnością faktów doświadczalnych z konsekwencjami wydedukowanymi z postulatów.

Wyjaśnijmy to bliżej na przykładzie newtonowskiej koncepcji przestrzeni fizycznej. W mechanice klasycznej dobrym modelem przestrzeni fizycznej jest pewna przestrzeń wektorowa (dokładniej: przestrzeń afiniczna związana z trójwymiarową przestrzenią euklidesową – pojęcia te omówimy w dalszej części podręcznika). Punktom przestrzeni fizycznej odpowiadają elementy tej przestrzeni. Związki i relacje zachodzące między punktami przestrzeni fizycznej można więc przewidywać badając jej model matematyczny, jednak trafność tych przewidywań może być sprawdzona jedynie na drodze doświadczalnej.

4 Przykłady

(i) Przestrzeń jednoelementowa

Najprostszym, choć trywialnym, przykładem przestrzeni wektorowej nad ciałem $\mathbb{K}$ jest zbiór jednoelementowy zawierający tylko wektor zerowy.

(ii) Ciało jako przestrzeń wektorowa

Ciało $\mathbb{K}$ spełnia postulaty przestrzeni wektorowej nad tym samym ciałem – aksjomaty przestrzeni wektorowej wynikają w tym przypadku z aksjomatów ciała. Ogólniej: jeśli $\mathbb{K}$ jest podciałem ciała $\mathbb{K}'$, to $\mathbb{K}'$ jest przestrzenią wektorową nad ciałem $\mathbb{K}$, np. zbiór liczb zespolonych $\mathbb{C}$ jest przestrzenią wektorową nad ciałem liczb rzeczywistych $\mathbb{R}$.

(iii) Przestrzeń funkcyjna
W punkcie 12, §3, omawialiśmy zbiór funkcji na zadanym zbiorze X o wartościach w ciele $\mathbb{K}$, wraz z wprowadzoną na tym zbiorze strukturą liniową. Własności tej struktury otrzymane w tym punkcie pokrywają się z aksjomatami przestrzeni wektorowej, więc otrzymujemy przykład przestrzeni wektorowej.

(iv) Przestrzeń macierzowa
Zbiór macierzy $\mathbb{K}^{m\times n}$ jest, jak wiemy, szczególnym przypadkiem zbioru funkcji na zadanym zbiorze, stanowi więc szczególny przypadek przestrzeni wektorowej z poprzedniego przykładu.

(v) Przestrzenie kolumn i wierszy
Zbiór $\mathbb{K}^{m\times 1}$ można utożsamić z potęgą kartezjańską $\mathbb{K}^m$, o elementach notowanych jako kolumny, a zbiór $\mathbb{K}^{1\times n}$ z potęgą $\mathbb{K}^n$, o elementach notowanych jako wiersze. Daje to dwa szczególne przypadki przestrzeni wektorowych z przykładu (iv).

5 Podprzestrzenie. Powłoki liniowe

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$, a W podzbiorem zbioru V. Mówimy, że W jest ***podprzestrzenią (wektorową) przestrzeni*** V, gdy W jest przestrzenią wektorową nad $\mathbb{K}$ z działaniami odziedziczonymi od przestrzeni V. Każda przestrzeń wektorowa V ma co najmniej dwie podprzestrzenie: zbiór złożony tylko z wektora zerowego $\{0\}$ i całą przestrzeń V. Nazywamy te podprzestrzenie trywialnymi. Ogólnie, kryterium dla podprzestrzeni jest analogiczne do kryterium dla podgrup, tw. 6, §3.

Twierdzenie 2. *Niepusty podzbiór W przestrzeni wektorowej V jest jej podprzestrzenią wtedy, i tylko wtedy, gdy spełniony jest warunek:*

$$\forall \alpha, \beta \in \mathbb{K}\ \forall x, y \in W :\ \alpha x + \beta y \in W\,.$$

Dowód. Konieczność warunku jest oczywista z definicji przestrzeni wektorowej. Jeśli warunek jest spełniony, to na W jest określone dodawanie wektorów oraz mnożenie wektorów przez liczby (żadne z działań nie wyprowadza poza W). Wtedy również dla każdych $x, y \in W$ mamy $0 = 0x \in W$, $-x = (-1)x \in W$ oraz $x + y \in W$, więc $(W, +)$ jest abelową grupą. Pozostałe własności działań wynikają z ich prawdziwości na całej przestrzeni V. Stąd teza. □

Niech F będzie dowolną rodziną wektorów przestrzeni V. Zbiór wszystkich kombinacji liniowych wektorów tej rodziny nazywamy ***powłoką liniową zbioru*** F i oznaczamy $L(F)$. Jeśli zbiór F jest zbiorem wartości ciągu wektorów $(x_1, \ldots, x_n)$, niekoniecznie różnowartościowego, to $L(F)$ jest zbiorem wszystkich kombinacji liniowych wektorów tego ciągu; piszemy wtedy też

$$L(F) = L(x_1, \ldots, x_n) .$$

Z praw łączności i rozdzielności widzimy natychmiast, że iloczyn kombinacji liniowej wektorów rodziny F przez liczbę oraz suma kombinacji liniowych wektorów rodziny F są znów kombinacjami liniowymi wektorów rodziny F, więc na podstawie ostatniego kryterium mamy następujący wniosek.

Twierdzenie 3. *Powłoka liniowa każdego zbioru $F \subseteq V$ jest podprzestrzenią przestrzeni V.*

Własności podprzestrzeni liniowych powodują, że każdy podzbiór podprzestrzeni zawiera się w niej wraz ze swoją powłoką liniową. Mówi o tym następne twierdzenie.

Twierdzenie 4.
(i) Dla każdej rodziny wektorów F i podprzestrzeni W zachodzi równoważność:

$$L(F) \subseteq W \iff F \subseteq W .$$

(ii) Dla każdych dwóch rodzin wektorów F_1, F_2 zachodzi równoważność:

$$L(F_1) = L(F_2) \iff F_1 \subseteq L(F_2) \wedge F_2 \subseteq L(F_1) .$$

Dowód. (i) Wynikanie $\Rightarrow$ jest oczywiste. Jeśli $F \subseteq W$, to na podstawie kryterium podprzestrzeni również dowolna kombinacja liniowa wektorów rodziny F należy do W, co kończy dowód pierwszego punktu. Punkt (ii) jest bezpośrednim wnioskiem z punktu (i). □

Jeśli $W = L(F)$, to mówimy, że ***rodzina*** F ***generuje podprzestrzeń*** W. Podobnie, jeśli $W = L(x_1, \ldots, x_n)$, to mówimy, że ***ciąg wektorów*** $(x_1, \ldots, x_n)$ ***generuje podprzestrzeń*** W.

6 Liniowa niezależność wektorów

Wybierzmy skończony ciąg $(x_1, \ldots, x_k)$ wektorów przestrzeni V. Mówimy, że wektory tego ciągu są ***liniowo niezależne***, jeśli dla każdej liniowej kombinacji

$\sum_{i=1}^{k} \lambda_i x_i$ tych wektorów zachodzi wynikanie:

$$\sum_{i=1}^{k} \lambda_i x_i = 0 \implies \forall i \in \{1, \ldots, k\} : \lambda_i = 0 \, .$$

W przypadku przeciwnym, tj. gdy istnieje kombinacja liniowa $\sum_{i=1}^{k} \lambda_i x_i = 0$ o współczynnikach nie wszystkich równych zero, mówimy, że wektory te są ***liniowo zależne***. W szczególności, jeśli ciąg nie jest różnowartościowy, to wektory są liniowo zależne (jeśli $x_l = x_m$, to $1x_l + (-1)x_m = 0$).

Niech teraz F będzie podzbiorem przestrzeni wektorowej V. Mówimy, że F jest ***rodziną wektorów liniowo niezależnych***, jeśli wektory każdego skończonego, *różnowartościowego* ciągu utworzonego z wektorów tej rodziny są liniowo niezależne. Inaczej mówiąc, nie istnieje nietrywialny rozkład wektora zerowego na wektory rodziny F. W przeciwnym przypadku, tj. gdy istnieje kombinacja liniowa $\sum_{i=1}^{k} \lambda_i x_i = 0$ o współczynnikach nie wszystkich równych zero, mówimy, że ***wektory rodziny F są liniowo zależne***.

Jest oczywistym, że jeśli F jest ciągiem lub rodziną liniowo niezależnych wektorów, a F' jest jego podciągiem lub podrodziną, to również F' jest odpowiednio takim ciągiem lub rodziną.

Treść pojęcia liniowej niezależności wyjaśnia bliżej następujące twierdzenie.

Twierdzenie 5.
(i) Wektory rodziny F są liniowo niezależne wtedy, i tylko wtedy, gdy żaden z nich nie może być przedstawiony jako liniowa kombinacja innych wektorów tej rodziny.
(ii) Jeśli wektory rodziny F są liniowo niezależne, a wektory rodziny $F \cup \{x\}$ liniowo zależne, to x ma jednoznaczne przedstawienie w postaci kombinacji liniowej wektorów rodziny F.

Dowód. (i) Skorzystamy z zasady kontrapozycji. Liniowa zależność wektorów rodziny F oznacza, że istnieje nietrywialny rozkład zera

$$\sum_{i=1}^{k} \lambda_i x_i = 0 \, , \quad \{x_1, \ldots, x_k\} \subseteq F \, .$$

Niech $\lambda_s \neq 0$. Wtedy ostatnia równość jest równoważna równości

$$x_s = \sum_{i \neq s} (-\lambda_i / \lambda_s) x_i \, ,$$

co kończy dowód tego punktu.

(ii) Wektory rodziny F są liniowo niezależne, a wektory rodziny $F \cup \{x\}$ – liniowo zależne, więc $x \notin F$ i istnieje nietrywialny rozkład zera

$$\alpha x + \sum_{i=1}^{k} \lambda_i x_i = 0\,, \quad \{x_1, \ldots, x_k\} \subseteq F\,.$$

Gdyby $\alpha = 0$, to wektory rodziny F byłyby liniowo zależne, co jest sprzeczne z założeniem. Stąd $\alpha \neq 0$; dzieląc równanie przez tę liczbę i przenosząc kombinację wektorów x_i na drugą stronę dostajemy rozkład x na wektory rodziny F. Załóżmy, wbrew tezie, że istnieją dwa różne rozkłady x na wektory z F. Łącznie wykorzystują one skończony zbiór wektorów $\{x_1, \ldots, x_k\}$, dają się więc zapisać jako $x = \sum_{i=1}^{k} \mu_i x_i$ i $x = \sum_{i=1}^{k} \mu'_i x_i$. Rozkłady są różne, więc nie wszystkie liczby $\mu_i - \mu'_i$ są równe zero. Odejmując rozkłady stronami dostajemy więc nietrywialny rozkład zera, co przeczy założeniu. □

7 Bazy. Wymiar przestrzeni

Bazą przestrzeni wektorowej V nazywamy każdą rodzinę $E \subseteq V$ liniowo niezależnych wektorów, która generuje V (tj. $L(E) = V$). Użyteczność pojęcia bazy opiera się na następującym, równoważnym jego sformułowaniu.

Twierdzenie 6. *Rodzina $E \subseteq V$ jest bazą wtedy, i tylko wtedy, gdy każdy wektor w przestrzeni V ma jednoznaczny rozkład na wektory rodziny E.*

Dowód. Jeśli E jest bazą, to $L(E) = V$, więc każdy wektor x ma rozkład na wektory rodziny E. Wektory rodziny E są liniowo niezależne, a rodziny $E \cup \{x\}$ – liniowo zależne, więc na podstawie twierdzenia 5 rozkład ten jest jednoznaczny. Odwrotnie, jeśli każdy wektor ma rozkład na wektory rodziny E, to $L(E) = V$. Ponadto, jeśli rozkłady te są jednoznaczne, to w szczególności nie istnieje nietrywialny rozkład zera, co jest stwierdzeniem liniowej niezależności wektorów rodziny E. □

Istotne pytania, na które musimy teraz odpowiedzieć, to: czy każda przestrzeń ma bazę? a jeśli tak, to ile jest wektorów w bazie? Częściowej odpowiedzi na te pytania dostarcza następne twierdzenie.

Twierdzenie 7.
(i) Jeśli przestrzeń wektorowa ma skończoną bazę, to każda baza tej przestrzeni ma tyle samo elementów. Jeśli $\{e_1, \ldots, e_n\}$ jest bazą, to rodzina wektorów $\{e'_1 \ldots, e'_n\}$ również jest bazą wtedy, i tylko wtedy, gdy macierz β ($\equiv (\beta_{ij})$)

współczynników rozkładów

$$e'_j = \sum_{i=1}^{n} e_i \, \beta_{ij} \,, \quad j = 1, \ldots, n$$

jest nieosobliwa. Wówczas

$$e_l = \sum_{k=1}^{n} e'_k \, (\beta^{-1})_{kl} \,, \quad l = 1, \ldots, n \,.$$

(ii) Jeśli przestrzeń wektorowa nie ma skończonej bazy, to każdą skończoną rodzinę liniowo niezależnych wektorów można powiększyć z zachowaniem liniowej niezależności.

Dowód. (i) Niech E i E' będą dwiema bazami, przy czym E jest skończona i $E = \{e_1, \ldots, e_n\}$. Każdy wektor bazy E ma rozkład na wektory bazy E' i odwrotnie, każdy wektor bazy E' rozkłada się na wektory bazy E. Ponieważ liczba elementów E jest skończona, a rozkład każdego z nich używa skończonej liczby elementów E', to łącznie wszystkie rozkłady wektorów E używają skończonego zbioru wektorów $\{e'_1, \ldots, e'_m\} \subseteq E'$. Twierdzimy, że zbiór ten wyczerpuje cały zbiór E'. Rzeczywiście, gdyby w E' istniał dodatkowy wektor f', to rozkładałby się on na wektory bazy E, zaś wstawienie rozkładów tych wektorów bazowych na wektory $\{e'_1, \ldots, e'_m\}$ dałoby rozkład wektora f' na wektory $\{e'_1, \ldots, e'_m\}$, co przeczy liniowej niezależności wektorów E'. Stąd $E' = \{e'_1, \ldots, e'_m\}$. Wzajemne rozkłady baz możemy teraz zapisać jako

$$e'_j = \sum_{i=1}^{n} e_i \beta_{ij} \quad (j = 1, \ldots, m) \,, \quad e_l = \sum_{k=1}^{m} e'_k \gamma_{kl} \quad (l = 1, \ldots, n) \,,$$

gdzie β_{ij} i γ_{kl} są współczynnikami kombinacji. Wstawmy rozkłady wektorów e_i po prawej stronie pierwszej równości, a rozkłady wektorów e'_k po prawej stronie drugiej równości. Przegrupowując sumy dostajemy

$$e'_j = \sum_{k=1}^{m} e'_k \sum_{i=1}^{n} \gamma_{ki} \beta_{ij} \,, \qquad e_l = \sum_{i=1}^{n} e_i \sum_{k=1}^{m} \beta_{ik} \gamma_{kl} \,.$$

Stąd ze względu na jednoznaczność rozkładów wektorów w bazach:

$$\sum_{i=1}^{n} \gamma_{ki} \beta_{ij} = \delta_{kj} \,, \qquad \sum_{k=1}^{m} \beta_{ik} \gamma_{kl} = \delta_{il} \,.$$

Wprowadzając macierze $\beta = (\beta_{ij}) \in \mathbb{K}^{n\times m}$ i $\gamma = (\gamma_{kl}) \in \mathbb{K}^{m\times n}$ zapisujemy te równania w równoważnej macierzowej postaci

$$\gamma\beta = \mathbf{1}_m\,, \qquad \beta\gamma = \mathbf{1}_n\,.$$

Twierdzenie 8, §7, pozwala teraz stwierdzić, że $m = n$ i macierze β i γ są nieosobliwe, wzajemnie odwrotne.

Odwrotnie, niech wektory n-elementowej rodziny E' mają takie rozkłady na wektory bazy E jak w założeniu i niech macierz β będzie nieosobliwa. Wtedy tworząc odpowiednie kombinacje liniowe wektorów rodziny E' dostajemy

$$\sum_{k=1}^{n} e'_k(\beta^{-1})_{kl} = \sum_{j=1}^{n} e_j\Big(\sum_{k=1}^{n}\beta_{jk}(\beta^{-1})_{kl}\Big) = e_l\,,$$

więc każdy wektor bazy E rozkłada się na wektory rodziny E', czyli E' generuje V. Dla dokończenia dowodu pozostaje pokazać, że E' jest rodziną wektorów liniowo niezależnych. Załóżmy, że $\sum_{j=1}^{n}\lambda_j\, e'_j = 0$. Wstawiając tu rozkłady wektorów e'_j i przegrupowując sumy dostajemy $\sum_{i=1}^{n} e_i\Big(\sum_{j=1}^{n}\beta_{ij}\lambda_j\Big) = 0$. Ze względu na liniową niezależność wektorów bazy E wynikają stąd równania $\sum_{j=1}^{n}\beta_{ij}\lambda_j = 0$ dla każdego $i = 1, \ldots, n$. Otrzymaliśmy układ Cramera na niewiadome λ_i (macierz β jest kwadratowa i nieosobliwa). Jego jedynym rozwiązaniem jest $\lambda_1 = \ldots = \lambda_n = 0$, więc wektory rodziny E' są liniowo niezależne.

(ii) Niech F będzie skończoną rodziną liniowo niezależnych wektorów. Zaprzeczając tezie załóżmy, że nie da się jej powiększyć z zachowaniem liniowej niezależności. Oznacza to, że dla każdego $x \in V \setminus F$ wektory rodziny $F \cup \{x\}$ są liniowo zależne. Stosując twierdzenie 5 wnioskujemy więc, że każdy wektor $x \in V \setminus F$ ma jednoznaczny rozkład na wektory rodziny F. Ale z liniowej niezależności również żaden wektor $f \in F$ nie daje się inaczej rozłożyć na wektory rodziny F. Reasumując: każdy wektor przestrzeni V ma jednoznaczny rozkład na wektory F, więc F jest bazą. Przeczy to założeniu, bo F jest skończona. □

Jeśli przestrzeń wektorowa ma skończoną bazę, to nazywamy ją ***skończenie wymiarową***, a jej ***wymiar*** definiujemy jako liczbę

$$\dim V := \text{liczba wektorów w dowolnej bazie.}$$

W przeciwnym przypadku mówimy, że przestrzeń jest ***nieskończenie wymiarowa***, a jej ***wymiar jest nieskończony***, i piszemy $\dim V = \infty$. Relację silnej nierówności określonej w $\mathbb{N}$ rozszerza się na ten zbiór powiększony o jeden element – ∞ – przez położenie $k < \infty$ dla każdego $k \in \mathbb{N}$.

Postawiliśmy wyżej pytanie o istnienie i liczebność baz. Ostatnie twierdzenie pozostawia te problemy otwartymi w przypadku przestrzeni nieskończenie wymiarowych. Okazuje się, że każda przestrzeń wektorowa ma bazę – jest to konsekwencja twierdzenia zamieszczonego w następnym punkcie. W przypadku przestrzeni nieskończenie wymiarowych pojęcie to jednak traci na praktycznym znaczeniu w rachunkach, choć nadal może być przydatne w niektórych dowodach. Dla wielu przestrzeni wektorowych jest nawet trudno efektywnie skonstruować jakąkolwiek bazę. W pewnych przestrzeniach nieskończenie wymiarowych używa się natomiast bardziej „ekonomicznych" baz innego typu. Na przykład, w przestrzeniach Hilberta (ośrodkowych) bazy takie są przeliczalne i każdy wektor ma jednoznaczny rozkład na „przeliczalną kombinację liniową" wektorów bazowych. Zbieżność szeregu określającego taką kombinację wymaga jednak użycia pewnej topologii (uogólnienie pojęcia odległości). Problemy takie wychodzą poza zakres algebry.

Naszym zasadniczym przedmiotem jest badanie przestrzeni skończenie wymiarowych, jednak tam, gdzie dowody nie wymagają użycia skończonej bazy, twierdzenia będą na ogół słuszne niezależnie od założeń o wymiarze.

8 Uzupełnianie do bazy. Monotoniczność wymiaru

Każdy zbiór liniowo niezależnych wektorów można uzupełnić do bazy (i to przy dodatkowych założeniach o bazie). Mówi o tym następujące twierdzenie, słuszne dla wszystkich przestrzeni wektorowych, którego jednak dowiedziemy tutaj tylko dla przestrzeni skończenie wymiarowych (dowód w ogólnym przypadku znajdzie czytelnik w Uzupełnieniach: p. 1, §24).

Twierdzenie 8. *Niech $A \subseteq V$ generuje V i niech $F \subseteq A$ będzie rodziną liniowo niezależnych wektorów (lub zbiorem pustym). Wtedy istnieje baza E taka, że $F \subseteq E \subseteq A$.*

Dowód. Zakładamy, że $n \equiv \dim V < \infty$. Twierdzenie implikuje w tym przypadku, że liczebność dowolnej rodziny liniowo niezależnych wektorów jest mniejsza bądź równa wymiarowi przestrzeni. Zaczniemy od udowodnienia tego rezultatu.

Niech E' będzie dowolną bazą. Wybierajmy z E' i dołączajmy do F kolejne wektory $e'_1, e'_2, \ldots$ w taki sposób, aby wektory kolejnych rodzin $F \cup \{e'_1, \ldots, e'_i\}$ $(i = 1, 2\ldots)$ były liniowo niezależne. Ta procedura musi się urwać, bo E' jest skończonym zbiorem. Niech dla $0 \leq k \leq n$ rodzina $F \cup \{e'_1, \ldots, e'_k\}$ nie da się już powiększyć. Oznacza to, że pozostałe wektory bazy E' (jeśli takie jeszcze zostały) dają się rozłożyć na wektory tej rodziny (na mocy twierdzenia 5). Stąd

$E' \subseteq L(F \cup \{e'_1, \ldots, e'_k\})$, więc $V = L(F \cup \{e'_1, \ldots, e'_k\})$ (bo każdy wektor rozkłada się na wektory E'). To pokazuje, że rodzina $F \cup \{e'_1, \ldots, e'_k\}$ jest bazą. Stąd liczba wektorów w F jest mniejsza bądź równa n.

Teraz przeprowadźmy podobne rozumowanie dołączając do F w podobny sposób kolejne wektory $a_1, a_2, \ldots$ zbioru A. Teraz procedura też musi się urwać, ale z innego powodu: wiemy już, że liczba elementów w rodzinie liniowo niezależnych wektorów nie może przekroczyć n. Poza tym rozumowanie przebiega analogicznie, aż do uzyskania $A \subseteq L(F \cup \{a_1, \ldots, a_k\})$. Stąd każda kombinacja liniowa wektorów A może być przedstawiona jako kombinacja liniowa wektorów $F \cup \{a_1, \ldots, a_k\}$, które są liniowo niezależne. Ale $V = L(A)$, więc $F \cup \{a_1, \ldots, a_k\}$ jest bazą spełniającą warunki tezy. $\square$

Dowód przeprowadziliśmy tylko dla przestrzeni skończenie wymiarowych, ale, jak zaznaczyliśmy, twierdzenie jest prawdziwe w ogólnym przypadku. Wynika z niego wspomniane wcześniej istnienie bazy w każdej przestrzeni wektorowej – wystarczy położyć $F = \varnothing$ i $A = V$.

Wnioskiem z ostatniego twierdzenia jest również następujące przydatne stwierdzenie.

Wniosek 9. *Z każdego ciągu wektorów $x_1, \ldots, x_n$ można wybrać wektory tworzące bazę powłoki $L(x_1, \ldots, x_n)$.*

Dowód. Wybierzmy zbiór A jako zbiór wartości ciągu $(x_1, \ldots, x_n)$. Wtedy $L(x_1, \ldots, x_n) = L(A)$. Założenia twierdzenia są spełnione z wyborem $F = \varnothing$ i $V = L(A)$, więc pewien podzbiór zbioru A jest bazą $L(A)$, co należało wykazać. $\square$

Każdą bazę, o której mówi wniosek, nazywamy ***maksymalną rodziną liniowo niezależnych wektorów ciągu*** $x_1, \ldots, x_n$. Dowód twierdzenia 8 pokazuje też metodę znalezienia tej bazy: dobieramy kolejne wektory $x_{i_1}, x_{i_2}, \ldots$ tego ciągu tak, aby za każdym krokiem otrzymać zbiór wektorów liniowo niezależnych. Zbiór $\{x_{i_1}, \ldots, x_{i_k}\}$, którego nie da się już powiększyć w ten sposób z zachowaniem liniowej niezależności, jest jedną z szukanych baz.

Twierdzenie o rozszerzaniu do bazy jest szczególnie przydatne w następującym kontekście. Niech $W \subseteq V$ będzie podprzestrzenią. Można tak wybrać bazę przestrzeni V, aby podzbiór tej bazy był bazą podprzestrzeni W. Wystarczy w tym celu wybrać dowolną bazę F przestrzeni W i uzupełnić ją do bazy E przestrzeni V (kładziemy w twierdzeniu $A = V$). Jeśli przy tym $\dim W = \dim V < \infty$, to $E = F$, więc $W = V$. W ten sposób udowodniliśmy następujący wniosek z twierdzenia.

Wniosek 10 (Monotoniczność wymiaru)**.** *Niech W będzie podprzestrzenią przestrzeni V. Wtedy*

(i) $\dim W \leq \dim V$*;*

(ii) $\dim W = \dim V < \infty \Rightarrow W = V$*.*

9 Bazy uporządkowane, współrzędne (składowe) wektora w bazie, zmiana bazy

Jeśli zadana jest baza E przestrzeni wektorowej, to każdy wektor x tej przestrzeni rozkłada się jednoznacznie na wektory bazowe, jest więc zadany jednoznacznie swoimi współczynnikami rozkładu w tej bazie. Jeśli przestrzeń jest skończenie wymiarowa (co dalej zakładamy w tym punkcie), wymiaru n, to zadaje go układ n liczb, musimy jednak kontrolować przypisanie liczb wektorom bazowym. Najłatwiej zrobić to wprowadzając następujące dodatkowe pojęcie. Różnowartościowy ciąg wektorów $(e_1, \ldots, e_n)$ nazywamy ***bazą uporządkowaną*** (***reperem***), jeśli zbiór wartości tego ciągu, $\{e_1, \ldots, e_n\}$, tworzy bazę. Inaczej mówiąc: bazę uporządkowaną otrzymujemy z bazy poprzez ustawienie wektorów w dowolny ciąg (a więc istnieje $n!$ baz uporządkowanych utworzonych z wektorów danej bazy). Dalej terminu *baza* często używać będziemy na oznaczanie każdego z tych pojęć, rozróżnianych kształtem użytych nawiasów.

Przy zadanej bazie uporządkowanej $(e_1, \ldots, e_n)$ współczynniki rozkładu $x = \sum_{i=1}^{n} \alpha_i e_i$ nazywamy ***współrzędnymi*** (lub ***składowymi*** – w literaturze fizycznej ta nazwa jest bardziej rozpowszechniona) ***wektora*** x ***w tej bazie***. Rozkład ten daje więc wzajemnie jednoznaczną odpowiedniość pomiędzy wektorami a elementami przestrzeni $\mathbb{K}^n$ – ciągami współrzędnych $(\alpha_1, \ldots, \alpha_n)$. Odpowiedniość ta stoi u podstaw użyteczności pojęcia bazy uporządkowanej i będzie używana w wielu twierdzeniach i dowodach. Zaznaczmy jednak, że wektor oraz jego współrzędne są pojęciowo różnymi obiektami. Co więcej, ich wzajemne przyporządkowanie zależy od wyboru bazy.

Znajdziemy związek współrzędnych wektora przestrzeni skończenie wymiarowej w dwóch różnych bazach. Niech $(e_1, \ldots, e_n)$ i $(e'_1, \ldots, e'_n)$ będą dwiema bazami uporządkowanymi, związanymi transformacją

$$e'_j = \sum_{i=1}^{n} e_i \beta_{ij} \quad (j = 1, \ldots, n)\,, \quad e_l = \sum_{k=1}^{n} e'_k (\beta^{-1})_{kl} \quad (l = 1, \ldots, n)\,.$$

Pierwszą z tych baz nazywać będziemy nieprimowaną, a drugą – primowaną. Macierz $\beta = (\beta_{ij})$ nazywać będziemy ***macierzą przejścia z bazy*** $(e_1, \ldots, e_n)$

do bazy $(e'_1, \dots, e'_n)$. Rozkładamy wektor x w każdej z tych baz:

$$x = \sum_{j=1}^{n} \alpha_j e_j = \sum_{j=1}^{n} \alpha'_j e'_j \,.$$

Jeśli do pierwszego z tych rozkładów wstawić rozkłady wektorów bazy nieprimowanej na wektory bazy primowanej, to po przegrupowaniu sum dostajemy

$$x = \sum_{i=1}^{n} e'_i \sum_{j=1}^{n} (\beta^{-1})_{ij} \alpha_j \,.$$

Z jednoznaczności rozkładu na wektory bazy primowanej mamy więc

$$\alpha'_i = \sum_{j=1}^{n} (\beta^{-1})_{ij} \alpha_j \quad (i = 1, \dots, n) \,.$$

Ustawmy wektory baz w symboliczne wiersze, np. $\begin{pmatrix} e_1 & \dots & e_n \end{pmatrix}$, i stosujmy do mnożenia tych wierszy przez macierze liczbowe zwykły przepis mnożenia macierzowego. Wtedy otrzymane rezultaty zapiszemy w następującej zwięzłej postaci.

Twierdzenie 11. *Jeśli bazy* $(e_1, \dots, e_n)$ *i* $(e'_1, \dots, e'_n)$ *przestrzeni skończenie wymiarowej wiążą się transformacją*

$$\begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \beta \,,$$

to współrzędne (składowe) każdego wektora w tych bazach są związane relacją

$$\begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix} = \beta^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \,.$$

Zwróćmy na koniec uwagę na niezbędność użycia w powyższym twierdzeniu baz wraz z określonym ich porządkiem. Jeśli $\{e_1, \dots, e_n\}$ i $\{e'_1, \dots, e'_n\}$ są równe jako zbiory, ale ciągi $(e_1, \dots, e_n)$ i $(e'_1, \dots, e'_n)$ są różne, to macierz przejścia między nimi jest różna od jednostkowej, co pociąga również zmianę kolumny współrzędnych.

10 Przykłady

(i) Przestrzeń trywialna
W przestrzeni wektorowej nad ciałem $\mathbb{K}$ złożonej tylko z wektora zerowego nie ma żadnej rodziny liniowo niezależnych wektorów. Istotnie, jeśli $\lambda\vec{0} = \vec{0}$, to λ może być dowolną liczbą. Przestrzeń ta jest więc 0-wymiarowa.

(ii) Bazy ciała jako przestrzeni wektorowej
W ciele $\mathbb{K}$ traktowanym jako przestrzeń wektorowa nad nim samym (patrz przykład (ii), p. 4) każdy zbiór jednoelementowy $\{\alpha\}$, gdzie $\alpha \neq 0$, jest bazą. Rzeczywiście, każda liczba $\beta \in \mathbb{K}$ ma jednoznaczne przedstawienie $\beta = (\beta/\alpha)\,\alpha$. Przestrzeń ta jest więc jednowymiarowa.

(iii) Baza przestrzeni wielomianowej
Niech $\mathbb{K} = \mathbb{C}$ lub $\mathbb{R}$. Zbiór wszystkich funkcji o dziedzinie i przeciwdziedzinie równej $\mathbb{K}$ stanowi szczególny przypadek przestrzeni z przykładu (iii), p. 4. Zbiór wszystkich wielomianów, które można w przypadku tych ciał utożsamić z funkcjami wielomianowymi, jest jej podprzestrzenią (fakt widoczny natychmiast w oparciu o kryterium podprzestrzeni).

Jednomiany $e_i(z) = z^i$, $i \in \{0\} \cup \mathbb{N}$, stanowią bazę podprzestrzeni wielomianów. Istotnie, współczynniki wielomianu P są przez niego jednoznacznie określone (patrz przykład (v), p. 8, §7 oraz punkt (i) we wniosku 15, §3), $P(z) = a_0 + a_1 z + \ldots + a_n z^n$, więc ma on jednoznaczny rozkład na wektory e_i: $P = \sum_{i=0}^{n} a_i e_i$. Przestrzeń wielomianów jest więc nieskończenie wymiarowa, a z monotoniczności wymiaru również przestrzeń wszystkich funkcji jest nieskończenie wymiarowa. Wiemy na podstawie twierdzenia, że musi istnieć jej baza zawierająca bazę przestrzeni wielomianów. O ile jednak przestrzeń wielomianów ma naturalną, przeliczalną bazę, o tyle jest niezmiernie trudno podać konstrukcyjnie jakąkolwiek bazę całej przestrzeni.

Dla każdego naturalnego n zbiór wielomianów stopnia mniejszego bądź równego n jest podprzestrzenią przestrzeni wszystkich wielomianów (znów w oparciu o kryterium tw. 2). Bazą tej podprzestrzeni jest skończony podzbiór $\{e_0, \ldots, e_n\}$ bazy wszystkich wielomianów, więc przestrzeń ta jest $(n+1)$-wymiarowa. Dla $m < n$ przestrzeń wielomianów stopnia $\leq m$ jest podprzestrzenią przestrzeni wielomianów stopnia $\leq n$.

(iv) Bazy przestrzeni macierzowych
Przestrzeń wektorowa macierzy $\mathbb{K}^{m\times n}$ ma naturalną bazę, zwaną ***bazą kanoniczną***. Wektory tej bazy wygodnie jest ponumerować parą wskaźników $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, i zdefiniować wektory bazowe E_{ij} o elementach macierzowych

$$(E_{ij})_{kl} = \delta_{ik}\delta_{jl} = \begin{cases} 1 & \text{gdy } k = i \text{ i } l = j\,, \\ 0 & \text{w pozostałych przypadkach}\,; \end{cases}$$

inaczej mówiąc: jedyny niezerowy element macierzy E_{ij} to jedynka na przecięciu i-tego wiersza i j-tej kolumny. Macierz $A = (A_{ij}) \in \mathbb{K}^{m\times n}$ ma jednoznaczne przedstawienie

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} \, E_{ij} \,,$$

więc zbiór $\{E_{ij}\}$ istotnie jest bazą, a przestrzeń jest mn-wymiarowa.

Zbiór macierzy symetrycznych $n \times n$ jest podprzestrzenią przestrzeni $\mathbb{K}^{n \times n}$. Bazą tej podprzestrzeni jest zbiór wektorów

$$\{E_{ij} + E_{ji}\}_{i<j} \cup \{E_{ii}\} \,,$$

gdzie E_{ij} są wektorami kanonicznej bazy przestrzeni $\mathbb{K}^{n \times n}$. W pierwszym podzbiorze jest $(n^2 - n)/2$ macierzy, a w drugim – n macierzy. Podprzestrzeń jest więc $n(n+1)/2$-wymiarowa.

Podobnie: jeśli $\operatorname{char} \mathbb{K} \neq 2$, to zbiór macierzy antysymetrycznych $n \times n$ jest podprzestrzenią $n(n-1)/2$-wymiarową o bazie

$$\{E_{ij} - E_{ji}\}_{i<j} \,.$$

Inna sytuacja zachodzi dla zbioru macierzy hermitowskich w przestrzeni zespolonej $\mathbb{C}^{n \times n}$. Zbiór ten nie jest podprzestrzenią – jeśli A jest macierzą hermitowską, to iA jest antyhermitowska. Rozpatrzmy jednak ten zbiór sam w sobie i zacieśnijmy ciało liczb, przez które można mnożyć macierze, do liczb rzeczywistych. Teraz łatwo sprawdzić, że otrzymujemy przestrzeń wektorową nad ciałem liczb rzeczywistych. Bazą tej przestrzeni jest zbiór wektorów

$$\{E_{ii}\} \cup \{E_{ij} + E_{ji}\}_{i<j} \cup \{iE_{kl} - iE_{lk}\}_{k<l} \,,$$

więc przestrzeń jest n^2-wymiarowa (wymiar przestrzeni $\mathbb{C}^{n \times n}$ jest też n^2, ale nie ma tu sprzeczności – zauważmy, że te wymiary są określone nad różnymi ciałami).

(v) Bazy przestrzeni kolumn i przestrzeni wierszy
Przestrzenie $\mathbb{K}^{n \times 1}$ i $\mathbb{K}^{1 \times n}$ są szczególnymi przypadkami przestrzeni macierzowych. Obie można uważać za realizacje przestrzeni $\mathbb{K}^n$: pierwszą w postaci wierszy, drugą w postaci kolumn. Wymiar każdej z nich jest równy n.

Baza kanoniczna przestrzeni $\mathbb{K}^{n \times 1}$ jest szczególnym przypadkiem bazy kanonicznej wprowadzonej w poprzednim przykładzie i ma postać:

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \ \ldots , \ E_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} .$$

Podobnie dla przestrzeni wierszy $\mathbb{K}^{1 \times n}$.

(vi) Transformacja współrzędnych wektorów przy zmianie bazy
W przestrzeni rzeczywistej trójwymiarowej dana jest baza uporządkowana (e_1, e_2, e_3). Określamy nową bazę

$$e'_1 = e_1 + 2e_2 - 3e_3\,, \quad e'_2 = 2e_1 - e_2 + 4e_3\,, \quad e'_3 = e_1 - e_2 + 2e_3\,.$$

Macierz przejścia z bazy (e_i) do bazy (e'_i) otrzymujemy ustawiając współrzędne wektorów nowej bazy w starej bazie jako jej kolejne *kolumny*:

$$\beta = \begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & -1 \\ -3 & 4 & 2 \end{pmatrix}\,.$$

Związek baz można teraz zapisać jako $\begin{pmatrix} e'_1 & e'_2 & e'_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \beta$. Dla otrzymania związku współrzędnych dowolnego wektora w tych dwu bazach wyliczamy macierz odwrotną do macierzy przejścia:

$$\beta^{-1} = \frac{1}{5} \begin{pmatrix} 2 & 0 & -1 \\ -1 & 5 & 3 \\ 5 & -10 & -5 \end{pmatrix}\,,$$

ustawiamy współczynniki rozkładów

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = \alpha'_1 e'_1 + \alpha'_2 e'_2 + \alpha'_3 e'_3$$

w kolumny, które podstawiamy do ogólnego przepisu

$$\begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \alpha'_3 \end{pmatrix} = \beta^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}\,.$$

Przepis ten mówi, że kolejne *wiersze* macierzy β^{-1} określają kolejne nowe współrzędne:

$$\alpha'_1 = \frac{2}{5}\alpha_1 - \frac{1}{5}\alpha_3\,, \quad \alpha'_2 = -\frac{1}{5}\alpha_1 + \alpha_2 + \frac{3}{5}\alpha_3\,, \quad \alpha'_3 = \alpha_1 - 2\alpha_2 - \alpha_3\,.$$

Na przykład, wektor $x = 10e_1 + 5e_2 - 15e_3$ ma w nowej bazie rozkład
$x = 7e'_1 - 6e'_2 + 15e'_3$.

§9 Układy równań liniowych

1 Istnienie rozwiązań układu równań liniowych

W tym paragrafie przerwiemy chwilowo ogólną analizę struktur pojawiających się w przestrzeniach wektorowych, aby zastosować aparat pojęciowy dotychczas poznany do przeanalizowania problemu rozwiązywalności, i znalezienia wszystkich rozwiązań, ogólnego układu równań liniowych (patrz p. 5, §6)

$$AX = B\,,$$

gdzie $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{m\times 1}$, a $X \in \mathbb{K}^{n\times 1}$ jest kolumną niewiadomych. Przypomnijmy, że macierz A nazwaliśmy macierzą główną układu. Tworzymy ponadto macierz blokową $\begin{pmatrix} A & B\end{pmatrix}$, którą oznaczać będziemy też inaczej $(A|B)$, i nazywamy ją ***macierzą dołączoną układu***.

Istnieje kilka sposobów interpretowania tego układu z punktu widzenia struktury przestrzeni wektorowych. Każdy z nich, jak zobaczymy później, wiąże się z innym, bardziej ogólnym, problemem w teorii skończenie wymiarowych przestrzeni wektorowych.

Dla rozważenia rozwiązywalności układu potraktujemy kolumnę niewiadomych jako ciąg współczynników liczbowych. Zapiszmy macierz A w postaci bloku kolumn $A = \begin{pmatrix} A_{*1} & \dots & A_{*n}\end{pmatrix}$. Mnożymy blokowo tę macierz przez kolumnę $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ i zapisujemy układ w postaci

$$x_1 A_{*1} + \ldots + x_n A_{*n} = B\,.$$

Tak zapisany, układ ma naturalną interpretację jako problem w przestrzeni $\mathbb{K}^m = \mathbb{K}^{m\times 1}$: znaleźć wszystkie sposoby, na jakie wektor B można rozłożyć na wektory A_{*i}. Problem ma więc rozwiązanie wtedy, i tylko wtedy, gdy

$$B \in L(A_{*1}, \ldots, A_{*n})\,.$$

Na podstawie punktu (ii) twierdzenia 4, §8, można ten warunek równoważnie sformułować jako

$$L(A_{*1}, \ldots, A_{*n}) = L(A_{*1}, \ldots, A_{*n}, B)\,.$$

Wniosek 9, §8, i uwagi następujące po nim pokazują, że ostatni zapis warunku jest równoważny stwierdzeniu: liczba elementów w każdej maksymalnej rodzinie liniowo niezależnych wektorów ciągu $(A_{*1}, \ldots, A_{*n}, B)$ jest równa liczbie elementów w każdej maksymalnej rodzinie liniowo niezależnych wektorów ciągu $(A_{*1}, \ldots, A_{*n})$.

Zanim przystąpimy do ostatecznego sformułowania otrzymanego wyniku, wprowadzimy definicję ułatwiającą zredagowanie wielu stwierdzeń. Niech A będzie dowolną macierzą o kolumnach $A_{*1}, \ldots, A_{*n}$. Liczbę elementów w każdej maksymalnej rodzinie liniowo niezależnych spośród kolumn tego ciągu nazywamy ***rzędem macierzy*** A i oznaczamy $\operatorname{rk} A$, a zatem

$$\operatorname{rk} A = \dim L(A_{*1}, \ldots, A_{*n}) .$$

Udowodniony wyżej wynik ma teraz zwięzłe sformułowanie.

Twierdzenie 1 (Kroneckera-Capellego)**.** *Układ $AX = B$ ma rozwiązanie wtedy, i tylko wtedy, gdy*

$$\operatorname{rk} A = \operatorname{rk}(A|B) .$$

2 Struktura ogólnego rozwiązania. Układ jednorodny

Konstrukcyjnym znalezieniem rozwiązań, jeśli istnieją, zajmiemy się w dalszych punktach. W tym punkcie zastanówmy się nad strukturą zbioru wszystkich rozwiązań. Naturalny będzie teraz inny punkt widzenia układu liniowego: traktujemy równanie $AX = B$ jako warunek zadający zbiór wektorów X przestrzeni $\mathbb{K}^n = \mathbb{K}^{n\times 1}$. Załóżmy, że nasz układ posiada rozwiązania, i niech wektor X_0 będzie jednym z nich, $AX_0 = B$ (a więc rzędy macierzy głównej i dołączonej są zgodne). Wprowadźmy wektor X' kładąc $X = X' + X_0$. Jest widoczne, że układ $AX = B$ jest równoważny układowi

$$AX' = 0 .$$

Układ o tej postaci (z zerową kolumną wyrazów wolnych) nazywamy ***układem jednorodnym***. Układ taki posiada zawsze co najmniej jedno rozwiązanie: wektor zerowy. Jeśli X_1 i X_2 są jego rozwiązaniami, to korzystając z własności mnożenia macierzowego widzimy, że dowolna ich kombinacja liniowa jest też rozwiązaniem tego samego układu. A więc: zbiór rozwiązań układu jednorodnego jest podprzestrzenią przestrzeni $\mathbb{K}^n$. Niech $\{X_1, \ldots, X_k\}$ będzie bazą tej podprzestrzeni. Wtedy ogólne rozwiązanie układu $AX = B$ jest k-parametrową rodziną wektorów

$$X = X_0 + \sum_{i=1}^{k} t_i X_i , \quad t_i \in \mathbb{K} .$$

Musimy teraz wypełnić konkretną treścią ten ogólny schemat.

3 Kryterium liniowej niezależności w przestrzeni $\mathbb{K}^m$

Znalezienie rozwiązania układu liniowego wymaga wcześniejszego zbadania własności rzędu macierzy. Udowodnimy najpierw twierdzenie pomocnicze.

Lemat 2. *Następujące warunki są równoważne:*

(i) wektory $X_1, \ldots, X_n \in \mathbb{K}^m$ są liniowo niezależne;

(ii) macierz utworzona z wektorów $X_1, \ldots, X_n$ zapisanych jako jej kolejne kolumny ma nieznikający minor $n \times n$;

(iii) macierz utworzona z wektorów $X_1, \ldots, X_n$ zapisanych jako jej kolejne wiersze ma nieznikający minor $n \times n$.

Dowód. (i)⇒(ii) Zapisujemy wektory $X_1, \ldots, X_n$ jako kolumny, czyli elementy przestrzeni $\mathbb{K}^{m\times 1}$. Jeśli są one liniowo niezależne, to można je uzupełnić do bazy $X_1, \ldots, X_m$ (stąd w szczególności $n \leq m$), która wiąże się nieosobliwą (twierdzenie 7, §8) macierzą β z bazą kanoniczną, $X_j = \sum_{i=1}^{m} E_i\,\beta_{ij}$. Posługując się jawną postacią bazy kanonicznej (patrz przykład (v) w punkcie 10, §8) mamy $X_j = \begin{pmatrix} \beta_{1j} \\ \vdots \\ \beta_{mj} \end{pmatrix} = \beta_{*j}$. Rozwińmy wyznacznik macierzy β względem m-tej kolumny, tj. ostatniej spośród dopisanych kolumn. Następnie minory wchodzące w skład tego rozwinięcia rozwińmy względem $(m-1)$-szych kolumn. Kontynuujmy aż do kolumn $(n+1)$-szych, czyli do wyczerpania kolumn dopisanych. Otrzymamy w ten sposób rozwinięcie wyznacznika macierzy β na minory wymiaru n utworzone z kolumn $X_1, \ldots, X_n$. Ponieważ wyznacznik jest różny od zera, to co najmniej jeden z tych minorów musi być różny od zera.

(ii)⇒(i) Zakładamy, że macierz $A = \begin{pmatrix} X_1 & \ldots & X_n \end{pmatrix}$ ma nieznikający minor $n \times n$, czyli po wykreśleniu pewnych $m-n$ spośród jej wierszy wyznacznik powstałej macierzy jest różny od zera. Istnieje więc taki podciąg $(i_1, \ldots, i_n)$ ciągu liczb $(1, \ldots, m)$, że macierz $A' \in \mathbb{K}^{n\times n}$ o elementach $A'_{kl} = A_{i_k l}$ jest nieosobliwa. Mamy pokazać, że kolumny $X_j = A_{*j}$ są liniowo niezależne. Załóżmy, że $\sum_{j=1}^{n} \lambda_j\, A_{*j} = 0$, czyli $\sum_{j=1}^{n} A_{ij}\,\lambda_j = 0$ dla $i = 1, \ldots, m$. Stąd w szczególności mamy $\sum_{j=1}^{n} A'_{kj}\,\lambda_j = 0$ dla $k = 1, \ldots, n$. Otrzymaliśmy układ Cramera, którego jedynym rozwiązaniem w tym przypadku jest $\lambda_1 = \ldots = \lambda_n = 0$.

(ii)⇔(iii) Macierz utworzona w punkcie (iii) jest macierzą transponowaną do macierzy utworzonej w punkcie (ii). Wartości odpowiadających sobie minorów tych macierzy są takie same, bo wyznacznik każdej macierzy jest równy wyznacznikowi macierzy do niej transponowanej. Stąd (ii) i (iii) są równoważne. □

4 Własności rzędu macierzy

Wnioskiem z lematu udowodnionego w poprzednim punkcie jest podstawowe twierdzenie o rzędzie.

Twierdzenie 3. *Następujące określenia rzędu macierzy są równoważne:*

(i) liczba elementów w każdej maksymalnej rodzinie liniowo niezależnych kolumn macierzy;

(ii) liczba elementów w każdej maksymalnej rodzinie liniowo niezależnych wierszy macierzy;

(iii) wymiar każdego maksymalnego (pod względem liczby kolumn i wierszy) nieznikającego minora macierzy.

Minor macierzy spełnia warunek zawarty w punkcie (iii) wtedy, i tylko wtedy, gdy powstaje na przecięciu rodzin wierszy i kolumn określonych warunkami (i) i (ii). Rząd macierzy jest równy rzędowi macierzy do niej transponowanej.

Dowód. Z równoważności punktów (i) i (ii) lematu 2 wynika, że wybrane kolumny macierzy tworzą rodzinę określoną punktem (i) twierdzenia wtedy, i tylko wtedy, gdy istnieje minor określony punktem (iii) i obejmujący te, i tylko te, kolumny. Podobne stwierdzenie dla wierszy wynika z równoważności punktów (i) i (iii) lematu, co kończy dowód równoważności punktów (i) – (iii) twierdzenia, a także pokazuje, że zarówno kolumny, jak i wiersze wchodzące w skład każdego maksymalnego nieznikającego minora tworzą maksymalny układ liniowo niezależny. Odwrotnie, rozważmy minor powstający na przecięciu dowolnego maksymalnego układu liniowo niezależnych kolumn i maksymalnego układu liniowo niezależnych wierszy; mamy wykazać, że jest on różny od zera. Bez ograniczenia ogólności możemy założyć, że jest to minor stojący w lewym górnym rogu macierzy (sprowadzamy go tam przestawiając wiersze i kolumny, co prowadzi co najwyżej do zmiany jego znaku). Zapiszmy macierz w postaci blokowej $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, gdzie A jest macierzą $k \times k$ i $\det A$ jest rozważanym minorem. Z naszych założeń wynika, że wszystkie kolumny macierzy $\begin{pmatrix} B \\ D \end{pmatrix}$ są kombinacjami liniowymi kolumn $\begin{pmatrix} A \\ C \end{pmatrix}$. Stąd, w szczególności, wszystkie kolumny macierzy B

są kombinacjami liniowymi kolumn macierzy A. Maksymalny układ liniowo niezależnych kolumn macierzy $\begin{pmatrix} A & B \end{pmatrix}$ można więc wybrać spośród kolumn macierzy A. Ponieważ równocześnie $\operatorname{rk}\begin{pmatrix} A & B \end{pmatrix} = k$ (bo wszystkie wiersze tej macierzy są liniowo niezależne), to $\det A \neq 0$, co kończy dowód. □

Powyższe twierdzenie ma zasadnicze znaczenie teoretyczne. Do obliczania rzędu stosujemy natomiast na ogół metody poznane już przy liczeniu wyznaczników:

Twierdzenie 4. *Rząd macierzy nie ulega zmianie przy następujących operacjach na macierzy:*

(i) przestawieniu kolumn lub wierszy;

(ii) pomnożeniu dowolnej kolumny lub wiersza przez liczbę różną od zera;

(iii) dodaniu do dowolnej kolumny lub wiersza macierzy dowolnej kombinacji liniowej pozostałych kolumn lub wierszy odpowiednio.

Dowód. Niech $X_1, \ldots, X_n$ będą wszystkimi kolejnymi kolumnami macierzy, a kolumny macierzy powstającej w wyniku zastosowania dowolnej z operacji (i) – (iii) oznaczmy $X'_1, \ldots, X'_n$. Tak otrzymane kolumny są kombinacjami liniowymi kolumn początkowych, więc $L(X'_1, \ldots, X'_n) \subseteq L(X_1, \ldots, X_n)$. Ale każda z operacji (i) – (iii) jest odwracalna, tj. $X_1, \ldots, X_n$ można przedstawić jako kombinacje liniowe kolumn $X'_1, \ldots, X'_n$, więc $L(X_1, \ldots, X_n) \subseteq L(X'_1, \ldots, X'_n)$. Stąd $L(X'_1, \ldots, X'_n) = L(X_1, \ldots, X_n)$, co dowodzi równości rzędów. Podobnie dla wierszy. □

5 Rozwiązania układu równań liniowych

Załóżmy, że układ równań $AX = B$ ma rozwiązania, a więc

$$\operatorname{rk} A = \operatorname{rk}(A|B)\,,$$

i oznaczmy $k = \operatorname{rk} A$. Przystępując do znalezienia ogólnego rozwiązania układu wróćmy chwilowo do jego jawnego, nieco zmodyfikowanego, zapisu jako m równań:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n - b_1 &= 0\,, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n - b_2 &= 0\,, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\,, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n - b_m &= 0\,. \end{aligned}$$

Dla uproszczenia dalszej dyskusji i zapisu rozwiązań założymy, że minor wymiaru k stojący w górnym lewym rogu macierzy A jest różny od zera. Nie ogranicza to ogólności. Po pierwsze: kolejność równań nie ma znaczenia, a przez ich zmianę można zawsze maksymalny układ liniowo niezależnych wierszy macierzy A sprowadzić na pierwszych k pozycji. Po drugie: przestawmy kolumny macierzy A tak, aby maksymalny układ liniowo niezależnych kolumn również zajął pierwszych k pozycji, dokonując równocześnie takiego samego przestawienia kolejności niewiadomych $x_1, \dots, x_n$ w kolumnie X; postać układu użyta w punkcie 1 pokazuje, że otrzymujemy układ równoważny. Przez zastosowanie tych operacji sprowadziliśmy układ do żądanej postaci.

Współczynniki każdego z równań zadane są jednym wierszem macierzy dołączonej $(A|B)$. Rozważmy te wiersze jako wektory w przestrzeni $\mathbb{K}^{1\times(n+1)}$. Jest widoczne, że jeśli pewien wiersz jest kombinacją liniową innych wierszy, to równanie przez niego określone wynika z równań zadanych wierszami tworzącymi tę kombinację liniową. Odrzucenie tego równania prowadzi do powstania układu równoważnego wyjściowemu. Zgodnie więc z naszymi założeniami otrzymujemy układ równoważny zachowując tylko k pierwszych równań. Zapisujemy go w postaci

$$\begin{pmatrix} A' & A'' \end{pmatrix} \begin{pmatrix} X' \\ X'' \end{pmatrix} = B' \,,$$

gdzie A' jest nieosobliwą macierzą kwadratową $k \times k$ (jest to blok, który zajmował lewy górny róg macierzy A), a $X = \begin{pmatrix} X' \\ X'' \end{pmatrix}$ (X' jest kolumną k-wyrazową, a X'' – $(n-k)$-wyrazową). Mnożąc macierze blokowo i przenosząc część wyniku na drugą stronę dostajemy równoważną postać równania

$$A'X' = B' - A''X'' \,.$$

Jesteśmy teraz przygotowani do sformułowania ogólnego rozwiązania układu równań liniowych.

Twierdzenie 5. *Niech rzędy macierzy głównej i dołączonej układu m równań liniowych o n niewiadomych $AX = B$ będą zgodne, i oznaczmy*

$$k = \operatorname{rk} A = \operatorname{rk}(A|B) \,.$$

Załóżmy, że minor wymiaru k stojący w lewym górnym rogu macierzy A jest różny od zera. Oznaczmy przez $\begin{pmatrix} A' & A'' \end{pmatrix}$ i B' macierze powstające przez skreślenie ostatnich $(n-k)$ wierszy z macierzy A i B odpowiednio, gdzie A' jest macierzą kwadratową $k \times k$. Wtedy:

(i) *ogólne rozwiązanie układu ma postać*

$$X = \begin{pmatrix} A'^{-1}(B' - A''X'') \\ X'' \end{pmatrix},$$

gdzie X'' *jest dowolnym wektorem w* $\mathbb{K}^{(n-k)\times 1}$*;*

(ii) *zbiór rozwiązań układu jednorodnego* $AX = 0$ *jest* $(n-k)$*-wymiarową podprzestrzenią przestrzeni* $\mathbb{K}^{n\times 1}$ *o bazie*

$$X_i = \begin{pmatrix} -A'^{-1}A''E_i \\ E_i \end{pmatrix}, \quad i = 1, \ldots, n-k,$$

gdzie $\{E_1, \ldots, E_{n-k}\}$ *jest bazą kanoniczną przestrzeni* $\mathbb{K}^{(n-k)\times 1}$*;*

(iii) *ogólne rozwiązanie układu* $AX = B$ *można przedstawić jako* $(n-k)$*-parametrową rodzinę*

$$X = X_0 + \sum_{i=1}^{n-k} t_i\, X_i\,,$$

gdzie t_i *są dowolnymi parametrami z ciała* $\mathbb{K}$*,*

$$X_0 = \begin{pmatrix} A'^{-1}B' \\ 0 \end{pmatrix},$$

a $X_1, \ldots, X_{n-k}$ *jest znalezioną w punkcie (ii) bazą podprzestrzeni rozwiązań układu jednorodnego.*

Dowód. W dyskusji poprzedzającej twierdzenie pokazaliśmy, że dany układ jest równoważny układowi $A'X' = B' - A''X''$. Przy dowolnie zadanym wektorze X'' układ ten jest układem Cramera ze względu na wektor X', skąd wynika teza (i). Jeśli zapisać wektor X'' jawnie jako $X'' = \begin{pmatrix} t_1 \\ \vdots \\ t_{n-k} \end{pmatrix}$, to jego rozkład w bazie kanonicznej ma postać $X'' = \sum_{i=1}^{n-k} t_i\, E_i$. Wstawiając ten rozkład do rozwiązania danego punktem (i) dostajemy

$$X = \begin{pmatrix} A'^{-1}B' \\ 0 \end{pmatrix} + \sum_{i=1}^{n-k} t_i \begin{pmatrix} -A'^{-1}A''E_i \\ E_i \end{pmatrix},$$

czyli rozkład taki, jak w punkcie (iii). Dla układu jednorodnego $X_0 = 0$. Dla dokończenia dowodu (ii) pozostaje pokazać, że wektory $X_1, \ldots, X_{n-k}$ są liniowo niezależne. Ale to jest natychmiast widoczne, bo jeśli $\sum_{i=1}^{n-k} t_i X_i = 0$, to w szczególności $\sum_{i=1}^{n-k} t_i E_i = 0$, więc $t_1 = \ldots = t_{n-k} = 0$. □

6 Metoda eliminacji zmiennych (Gaussa)

Opiszemy w tym punkcie metodę rozwiązywania układu równań liniowych, która nie daje tak wyczerpującego zrozumienia struktur obecnych w tym zagadnieniu, jakie uzyskaliśmy na dotychczasowej drodze, ale jest efektywna rachunkowo.

Przedstawmy symbolicznie układ $AX = B$ za pomocą jego macierzy dołączonej $(A|B)$. Niech macierz $(\tilde{A}|\tilde{B})$ powstaje z macierzy $(A|B)$ przez złożenie następujących operacji:

(i) pomnożenie wybranego wiersza przez dowolną liczbę różną od zera,

(ii) dodanie do wybranego wiersza innego wiersza pomnożonego przez dowolną liczbę,

(iii) zmiana kolejności wierszy.

Macierz $(\tilde{A}|\tilde{B})$ przedstawia wtedy równoważny układ równań $\tilde{A}X = \tilde{B}$. Operacja (iii) jest jedynie zamianą kolejności równań, więc równoważność jest oczywista. Dla operacji (i) i (ii) jest widoczne, że jeśli X jest rozwiązaniem początkowego układu, to spełnia również przekształcony układ równań. Ale obie operacje są odwracalne, więc zachodzi też wynikanie odwrotne, więc równoważność. Wprowadzamy też pomocniczo, dla ułatwienia zapisu dalszego rozumowania, równoczesną transformację (przy zachowaniu kolumny B):

(iv) $A \to \tilde{A}$, $X \to \tilde{X}$, gdzie $\tilde{A}$ powstaje z A przez zmianę kolejności kolumn, a $\tilde{X}$ powstaje z X przez taką samą zmianę kolejności wyrazów w kolumnie.

Wtedy układ $\tilde{A}\tilde{X} = B$ jest równoważny układowi $AX = B$, co uzasadniliśmy już na początku poprzedniego punktu.

Załóżmy, że macierz dołączona układu ma blokową postać $\begin{pmatrix} A' & A'' & B' \\ 0 & C & B'' \end{pmatrix}$, gdzie A' jest kwadratową macierzą trójkątną o wyrazach diagonalnych różnych od zera, a $\begin{pmatrix} B' \\ B'' \end{pmatrix}$ jest kolumną wyrazów wolnych. Ogólna postać dowolnego układu jest szczególnym przypadkiem tej postaci – z blokiem A' zerowego wymiaru. Niech w bloku C istnieje wyraz różny od zera. Do wierszy i kolumn wchodzących w skład tego bloku oraz do kolumny niewiadomych zastosujemy operacje typu (ii) – (iv), co nie wpłynie na postać bloków $\begin{pmatrix} A' \\ 0 \end{pmatrix}$. Posługując się najpierw operacjami (iii) i (iv) umieszczamy niezerowy wyraz bloku C w jego lewym górnym rogu. Następnie używając transformacji (ii) sprowadzamy do zera wszystkie wyrazy leżące pod tym rogiem. Rezultatem tych operacji jest znów postać założonego typu, ale z każdym wymiarem bloku typu A' powiększonym o jeden, a każdym wymiarem bloku typu C pomniejszonym o jeden. Powtarzamy tę procedurę aż do wyczerpania niezerowych elementów w bloku typu C (lub do zniknięcia tego bloku – jeśli któryś z jego wymiarów spadnie do zera). Otrzymujemy układ równoważny wyjściowemu, o macierzy dołączonej typu:

$$\begin{pmatrix} A' & A'' & B' \\ 0 & 0 & B'' \end{pmatrix},$$

z trójkątną nieosobliwą macierzą A'. Stąd rząd macierzy głównej otrzymanego układu jest równy wymiarowi macierzy A', oznaczmy go k. Część z równań tego układu ma postać $0 = B''$, więc spełnienie tego warunku jest konieczne do istnienia rozwiązań. Jest ono również wystarczające, za chwilę wyliczymy rozwiązanie w tym przypadku. Zauważmy również, że warunek $B'' = 0$ pokrywa się z kryterium Kroneckera-Capellego. Istotnie, po pierwsze: ponieważ w stosowanych operacjach nie mieszaliśmy kolumn macierzy głównej z kolumną wyrazów wolnych, to rzędy macierzy głównej i dołączonej nie uległy zmianie; po drugie: jest widoczne, że znikanie B'' jest warunkiem koniecznym i wystarczającym, aby rząd macierzy dołączonej w powyższej postaci był równy k.

Załóżmy więc, że $B'' = 0$. Zerowe wiersze macierzy dołączonej symbolizują tożsamościowo spełnione równania, więc równoważna postać układu jest wtedy zadana macierzą dołączoną

$$\begin{pmatrix} A' & A'' & B' \end{pmatrix}.$$

Stosując operacje typu (ii) sprowadzamy do zera wszystkie wyrazy stojące nad elementem A'_{kk}, następnie nad elementem $A'_{(k-1)(k-1)}$, itd. Otrzymujemy postać jak wyżej, ale z macierzą A' diagonalną. Mnożąc wiersze przez odwrotności diagonalnych elementów macierzy A' (punkt (i)) otrzymujemy ostatecznie układ o macierzy dołączonej i kolumnie niewiadomych typu $\begin{pmatrix} \mathbf{1} & A'' & B' \end{pmatrix}$ i $\begin{pmatrix} X' \\ X'' \end{pmatrix}$ odpowiednio. Stąd

$$X' = B' - A''X'',$$

gdzie X'' jest dowolną kolumną $(n-k)$-wyrazową. Interpretacja rozwiązania została przedyskutowana we wcześniejszych punktach. Dodajmy na koniec, że mnożenie wierszy przez liczby różne od zera możemy stosować również w krokach pośrednich, dla ułatwienia rachunków.

7 Metoda Gaussa – wariant z zachowaniem kolejności kolumn

Metodę eliminacji zmiennych można stosować w kilku równoważnych wariantach. W poprzednim punkcie opisaliśmy wariant najprostszy do uzasadnienia metody przy pomocy symbolicznego zapisu. W pewnych sytuacjach, które opiszemy poniżej, wygodniejsze w zastosowaniu są warianty nie posługujące się operacją (iv) z poprzedniego punktu, a więc nie zmieniające kolejności kolumn macierzy głównej układu, ani kolejności niewiadomych.

Stosujemy postępowanie rekurencyjne podobnie jak w poprzednim punkcie, ale z następującymi modyfikacjami:

– nie zakładamy, że blok A' jest kwadratowy (a więc założenie o niezerowości elementów diagonalnych też traci moc);

– elementu niezerowego szukamy nie w całym bloku C, lecz w jego pierwszej kolumnie; jeśli taki istnieje, to podnosimy go do pierwszego wiersza bloku C, przeprowadzamy odjęcia i zmniejszenie bloku C jak w poprzedniej wersji;

– jeśli ta pierwsza kolumna jest zerowa, to dołączamy ją do bloku zerowego odpowiednio również przesuwając granicę między blokami A' i A'', po czym kontynuujemy procedurę rekurencyjną.

Po zakończeniu procedury rekurencyjnej blok $\begin{pmatrix} A'' \\ C \end{pmatrix}$ zostaje całkowicie wyeliminowany. Podobnie jak poprzednio układ ma rozwiązanie wtedy, i tylko wtedy, gdy $B'' = 0$, i sprowadza się wówczas do układu określonego macierzą dołączoną $\begin{pmatrix} A' & B' \end{pmatrix}$, gdzie ciąg początkowych zer w każdym wierszu bloku A' jest dłuższy niż w poprzednim, ale nie wypełnia żadnego wiersza całkowicie. Skalujemy wiodące niezerowe wyrazy w każdym wierszu do jedynki, po czym przez odjęcia wierszy sprowadzamy wszystkie leżące ponad nimi elementy do zera. Wtedy macierz A' uzyskuje postać

$$A' = \begin{pmatrix}
0\cdots 0 & 1 & *\cdots * & 0 & *\cdots * & 0 & *\cdots & \cdots * & 0 & *\cdots * \\
0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * & 0 & *\cdots & \cdots * & 0 & *\cdots * \\
0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots & \cdots * & 0 & *\cdots * \\
0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots & \cdots * & 0 & *\cdots * \\
\vdots\ \ \vdots & \vdots & \vdots\ \ \vdots & \vdots & \vdots\ \ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\ \ \vdots \\
0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots & \cdots * & 0 & *\cdots * \\
0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots & \cdots 0 & 1 & *\cdots *
\end{pmatrix}.$$

$$\begin{array}{cccccccccc}
 & \uparrow & & \uparrow & & \uparrow & & \cdots & \uparrow &
\end{array}$$

Niewiadome przypadające na kolumny wskazane strzałkami wylicza się natychmiast w funkcji pozostałych zmiennych, które są dowolnymi parametrami.

8 Metoda Gaussa – liniowe zależności między wektorami w $\mathbb{K}^m$

Wrócimy teraz jeszcze raz do problemu liniowych zależności między wektorami w przestrzeni $\mathbb{K}^m$. Niech $C_1, \ldots, C_n$ będą wektorami w $\mathbb{K}^m$. Chcemy znaleźć maksymalną rodzinę liniowo niezależnych spośród nich i zapisać pozostałe jako ich kombinacje liniowe. Przekonamy się, że metoda Gaussa opisana w ostatnim punkcie da nam prosty sposób rozwiązania tego problemu.

Notujemy wektory C_i w postaci kolumn (bez zmiany oznaczeń) i rozpatrujemy jednorodne równanie wektorowe

$$x_1 C_1 + \ldots x_n C_n = 0$$

ze względu na niewiadome współczynniki $x_1, \ldots, x_n$. Jeśli utworzyć macierz blokową $A = \begin{pmatrix} C_1 & \ldots & C_n \end{pmatrix}$, to równanie to jest równoważne jednorodnemu równaniu liniowemu $AX = 0$, gdzie kolumna X jest utworzona z niewiadomych $x_1, \ldots, x_n$. Stosując metodę Gaussa w wersji opisanej w poprzednim punkcie sprowadzamy równanie do równoważnej postaci

$$x_1 A'_{*1} + \ldots + x_n A'_{*n} = 0\,,$$

gdzie A' ma taką postać, jaką uzyskaliśmy w poprzednim punkcie. Stąd liczby $x_1, \ldots, x_n$ są współczynnikami wszystkich możliwych liniowych zależności pomiędzy kolumnami macierzy A', które to zależności są zatem identyczne, jak dla kolumn $C_1, \ldots, C_n$. Ale liniowe zależności pomiędzy kolumnami macierzy A' odczytuje się natychmiast z jej postaci: kolumny wskazane strzałkami są liniowo niezależne, a pozostałe są ich kombinacjami liniowymi o współczynnikach odczytywanych bezpośrednio z ich postaci. Identyczne związki zachodzą więc dla kolumn $C_1, \ldots, C_n$.

9 Metoda Gaussa – macierz odwrotna

Niech macierz główna układu $AX = B$ będzie kwadratowa. Zastosujmy do tego układu metodę Gaussa w wersji z punktu 7. Jeśli układ okaże się być rozwiązywalny, to dostaniemy w wyniku układ o macierzy dołączonej $\begin{pmatrix} A' & B' \\ 0 & B'' \end{pmatrix}$, z macierzą A' o postaci uzyskanej w punkcie 7. Jest natychmiast widoczne, że macierz A jest nieosobliwa wtedy, i tylko wtedy, gdy bloki $\begin{pmatrix} 0 & B'' \end{pmatrix}$ są nieobecne, a macierz A' ma tylko kolumny wskazane strzałką, czyli $A' = \mathbf{1}$. Macierz dołączona została sprowadzona do postaci $\begin{pmatrix} \mathbf{1} & B' \end{pmatrix}$, więc uzyskujemy jednoznaczne rozwiązanie $X = B'$.

Niech teraz A będzie macierzą kwadratową $n \times n$, dla której chcemy wyliczyć, jeśli istnieje, macierz odwrotną. Szukamy więc macierzy kwadratowej Y spełniającej równanie $AY = \mathbf{1}$. Zapisując to równanie jako równość poszczególnych kolumn uzyskujemy układ

$$AY_{*1} = E_1 \,, \quad \ldots, \quad AY_{*n} = E_n \,,$$

gdzie $E_1, \ldots, E_n$ jest bazą kanoniczną w $\mathbb{K}^n$. Do każdego z równań możemy zastosować identyczną procedurę według schematu użytego w poprzednim akapicie, i jeśli macierz A okaże się nieosobliwa, to macierze dołączone otrzymane w wyniku będą miały postać $\begin{pmatrix}\mathbf{1} & B'_1\end{pmatrix}, \ldots, \begin{pmatrix}\mathbf{1} & B'_n\end{pmatrix}$, skąd odczytujemy $Y_{*i} = B'_i$.

Otrzymany rezultat możemy teraz opisać w następujący sposób. Chcąc znaleźć macierz odwrotną do macierzy A tworzymy macierz blokową $\begin{pmatrix}A & \mathbf{1}\end{pmatrix}$ i stosujemy do jej wierszy operacje opisane w punkcie 7 (a więc bez zmiany kolejności kolumn!), dążąc do sprowadzenia bloku A do bloku jednostkowego. Jeśli to jest możliwe, to uzyskamy w wyniku macierz $\begin{pmatrix}\mathbf{1} & A^{-1}\end{pmatrix}$.

10 Przykłady

(i) Układ sprzeczny

Sprawdzamy, czy następujący układ równań liniowych ma rozwiązania:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - x_4 &= 3 \,, \\ 4x_1 - x_2 + 2x_3 + 2x_4 &= 5 \,, \\ 2x_1 - 5x_2 - 4x_3 + 4x_4 &= 0 \,. \end{aligned}$$

Liczymy rząd macierzy głównej układu:

$$\operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 \\ 4 & -1 & 2 & 2 \\ 2 & -5 & -4 & 4\end{pmatrix} = \operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 \\ 0 & -9 & -10 & 6 \\ 0 & -9 & -10 & 6\end{pmatrix} = \\ = \operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 \\ 0 & -9 & -10 & 6\end{pmatrix} = 2 \,.$$

Pierwszą równość otrzymaliśmy przez dodanie wielokrotności pierwszego wiersza do pozostałych wierszy (przez wielokrotność wiersza lub kolumny będziemy rozumieć ich iloczyn przez dowolną liczbę rzeczywistą), drugą – przez pominięcie wiersza identycznego z drugim, trzecią – z niezależności otrzymanych wierszy.

Liczymy rząd macierzy dołączonej

$$\operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 & 3 \\ 4 & -1 & 2 & 2 & 5 \\ 2 & -5 & -4 & 4 & 0\end{pmatrix} = \operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 & 3 \\ 0 & -9 & -10 & 6 & -7 \\ 0 & -9 & -10 & 6 & -6\end{pmatrix} = \\ = \operatorname{rk}\begin{pmatrix}1 & 2 & 3 & -1 & 3 \\ 0 & -9 & -10 & 6 & -7 \\ 0 & 0 & 0 & 0 & 1\end{pmatrix} = 3 \,.$$

Pierwszą równość otrzymaliśmy podobnie jak w poprzednim rachunku, drugą – przez odjęcie drugiego wiersza od trzeciego, trzecią – z nieznikania minora utworzonego z pierwszej, drugiej i piątej kolumny.

Rzędy nie są zgodne, więc układ jest sprzeczny.

(ii) Ogólne rozwiązanie niesprzecznego układu
Szukamy ogólnego rozwiązania układu równań:

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & -1 & -1 & 4 \\ 3 & 1 & 2 & -2 \\ 6 & 2 & 4 & 2 \\ 2 & -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \\ 6 \\ 0 \end{pmatrix} .$$

Oznaczmy macierz główną tego układu przez A, a kolumnę wyrazów wolnych przez B. Wtedy:

$$\operatorname{rk} A = \operatorname{rk} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -5 & -7 & 4 \\ 3 & -5 & -7 & -2 \\ 6 & -10 & -14 & 2 \\ 2 & -5 & -7 & 2 \end{pmatrix} = 1 + \operatorname{rk} \begin{pmatrix} -5 & -7 & 4 \\ -5 & -7 & -2 \\ -10 & -14 & 2 \\ -5 & -7 & 2 \end{pmatrix} =$$

$$= 1 + \operatorname{rk} \begin{pmatrix} 1 & 4 \\ 1 & -2 \\ 2 & 2 \\ 1 & 2 \end{pmatrix} = 1 + \operatorname{rk} \begin{pmatrix} 1 & 3 \\ 1 & -3 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} = 3 .$$

Pierwszą równość otrzymaliśmy przez dodanie wielokrotności pierwszej kolumny do pozostałych kolumn, drugą – z liniowej niezależności pierwszej kolumny od pozostałych, trzecią – przez skreślenie drugiej kolumny (proporcjonalnej do pierwszej) i pomnożenie pierwszej kolumny przez liczbę, czwartą – przez odjęcie pierwszej kolumny od drugiej, piątą – z liniowej niezależności otrzymanych kolumn. Podobnie wyliczamy $\operatorname{rk}(A|B) = 3$, więc układ ma rozwiązania. Minor macierzy A powstający na przecięciu wierszy o numerach 1, 2 i 3 z kolumnami o numerach 2, 3 i 4 jest różny od zera, więc równanie czwarte i piąte można odrzucić, a wyrazy proporcjonalne do x_1 przenieść na prawą stronę. Dany układ jest więc równoważny układowi

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & -1 & 4 \\ 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix} - x_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 - x_1 \\ 2 - 2x_1 \\ -3x_1 \end{pmatrix} .$$

Stosując metodę Cramera wyliczamy

$$x_2 = 2 + 7x_1\,, \quad x_3 = -5x_1\,, \quad x_4 = 1\,,$$

gdzie x_1 jest parametrem. Kładąc $x_1 = t$ i ustawiając znalezione zmienne w kolumnę dostajemy

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 7 \\ -5 \\ 0 \end{pmatrix}\,, \ t \in \mathbb{R}\,.$$

Pierwsza kolumna po prawej stronie jest szczególnym rozwiązaniem danego układu równań, a druga kolumna tworzy bazę jednowymiarowej przestrzeni rozwiązań odpowiedniego układu jednorodnego.

(iii) Metoda Gaussa – układ sprzeczny
Sprawdzamy, czy następujący układ równań ma rozwiązania:

$$\begin{pmatrix} 2 & 1 & 3 & 5 \\ 0 & 3 & -2 & 1 \\ -1 & 4 & 1 & 0 \\ 1 & 11 & 0 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \end{pmatrix}\,.$$

Zamieniamy miejscami pierwsze i trzecie równanie (aby ułatwić dalsze rachunki), i stosując zasady metody otrzymujemy ciąg układów równoważnych danemu, reprezentowanych kolejnymi macierzami dołączonymi (znakiem $\sim$ zaznaczamy równoważność odpowiednich układów):

$$\begin{pmatrix} -1 & 4 & 1 & 0 & -1 \\ 0 & 3 & -2 & 1 & 2 \\ 2 & 1 & 3 & 5 & 1 \\ 1 & 11 & 0 & 7 & 3 \end{pmatrix} \sim \begin{pmatrix} -1 & 4 & 1 & 0 & -1 \\ 0 & 3 & -2 & 1 & 2 \\ 0 & 9 & 5 & 5 & -1 \\ 0 & 15 & 1 & 7 & 2 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} -1 & 4 & 1 & 0 & -1 \\ 0 & 3 & -2 & 1 & 2 \\ 0 & 0 & 11 & 2 & -7 \\ 0 & 0 & 11 & 2 & -8 \end{pmatrix} \sim \begin{pmatrix} -1 & 4 & 1 & 0 & -1 \\ 0 & 3 & -2 & 1 & 2 \\ 0 & 0 & 11 & 2 & -7 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}\,.$$

Pierwszą równoważność otrzymaliśmy przez dodanie wielokrotności pierwszego wiersza macierzy dołączonej do pozostałych, drugą – przez dodanie wielokrotności drugiego wiersza do znajdujących się poniżej, trzecią – przez odjęcie trzeciego wiersza od ostatniego. Bloki niezerowe macierzy głównej otrzymanego układu kończą się na trzecim wierszu, a kolumna wyrazów wolnych ma wyraz niezerowy poniżej trzeciego wiersza, więc układ jest sprzeczny.

(iv) Metoda Gaussa – ogólne rozwiązanie
Stosujemy metodę Gaussa do znalezienia ogólnego rozwiązania układu równań

$$\begin{pmatrix} 1 & 2 & -1 & 3 \\ 3 & -2 & 1 & 1 \\ -1 & 6 & -3 & 5 \\ 3 & -10 & 5 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 6 \\ -7 \end{pmatrix} .$$

Dodając odpowiednie wielokrotności pierwszego wiersza macierzy dołączonej do pozostałych wierszy, dostajemy równoważny układ o macierzy dołączonej

$$\begin{pmatrix} 1 & 2 & -1 & 3 & 5 \\ 0 & -8 & 4 & -8 & -11 \\ 0 & 8 & -4 & 8 & 11 \\ 0 & -16 & 8 & -16 & -22 \end{pmatrix} .$$

Dodajemy wielokrotności drugiego wiersza macierzy dołączonej do pozostałych jej wierszy, skreślamy wiersze zerowe, mnożymy drugi wiersz przez liczbę i dodajemy jego wielokrotność do pierwszego wiersza, dostając w wyniku równoważny układ o macierzy dołączonej

$$\begin{pmatrix} 1 & 0 & 0 & 1 & \frac{9}{4} \\ 0 & 1 & -\frac{1}{2} & 1 & \frac{11}{8} \end{pmatrix} .$$

Stąd $x_1 = \frac{9}{4} - x_4$, $x_2 = \frac{11}{8} + \frac{1}{2}x_3 - x_4$. Kładąc $x_3 = 2s$, $x_4 = t$ dostajemy

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 18 \\ 11 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} , \quad s, t \in \mathbb{R} .$$

Pierwsza kolumna po prawej stronie jest szczególnym rozwiązaniem układu, a dwie następne tworzą bazę dwuwymiarowej przestrzeni rozwiązań odpowiedniego układu jednorodnego.

(v) Metoda Gaussa – liniowe zależności wektorów w $\mathbb{K}^n$
W przestrzeni $\mathbb{R}^5$ dane są wektory

$$X_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \\ 1 \end{pmatrix} , \ X_2 = \begin{pmatrix} 2 \\ 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} , \ X_3 = \begin{pmatrix} 0 \\ 3 \\ 8 \\ -3 \\ 2 \end{pmatrix} , \ X_4 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \\ 2 \end{pmatrix} , \ X_5 = \begin{pmatrix} 3 \\ -2 \\ -4 \\ 2 \\ 1 \end{pmatrix} .$$

Szukamy maksymalnej liniowo niezależnej rodziny wśród tych wektorów i przedstawienia pozostałych jako kombinacji liniowych tej rodziny.

Postępujemy zgodnie z metodą opisaną w punkcie 8:

$$\begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 2 & 1 & 3 & -1 & -2 \\ 3 & -2 & 8 & 1 & -4 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & 0 & 2 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 0 & -3 & 3 & -5 & -8 \\ 0 & -8 & 8 & -5 & -13 \\ 0 & 3 & -3 & 2 & 5 \\ 0 & -2 & 2 & 0 & -2 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Stąd odczytujemy: wektory X_1, X_2 i X_4 są liniowo niezależne, a pozostałe wyrażają się jako: $X_3 = 2X_1 - X_2$, $X_5 = -X_1 + X_2 + X_4$.

(vi) Metoda Gaussa – macierz odwrotna
Szukamy macierzy odwrotnej do macierzy

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 5 & 4 & 1 \\ 2 & 1 & -2 \end{pmatrix}.$$

Stosując metodę Gaussa dostajemy:

$$\begin{pmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 5 & 4 & 1 & 0 & 1 & 0 \\ 2 & 1 & -2 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & -6 & 16 & -5 & 1 & 0 \\ 0 & -3 & 4 & -2 & 0 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 3 & -4 & 2 & 0 & -1 \\ 0 & 0 & 8 & -1 & 1 & -2 \end{pmatrix} \sim \begin{pmatrix} 24 & 0 & 0 & -9 & 1 & 14 \\ 0 & 6 & 0 & 3 & 1 & -4 \\ 0 & 0 & 8 & -1 & 1 & -2 \end{pmatrix}.$$

Stąd

$$A^{-1} = \frac{1}{24} \begin{pmatrix} -9 & 1 & 14 \\ 12 & 4 & -16 \\ -3 & 3 & -6 \end{pmatrix}.$$

§10 Odwzorowania liniowe

1 Definicje i konwencje

Niech V i W będą dwiema przestrzeniami wektorowymi nad tym samym ciałem $\mathbb{K}$. Podobnie jak dla grup, wśród wszystkich odwzorowań zbioru V w zbiór W wyróżniamy te, które są zgodne ze strukturami algebraicznymi zadanymi w tych zbiorach – w przypadku struktur liniowych oznacza to zgodność z dodawaniem wektorów i mnożeniem ich przez liczbę. Mówimy, że odwzorowanie $A : V \mapsto W$ jest ***odwzorowaniem liniowym*** (lub ***homomorfizmem liniowym***; będziemy z reguły oznaczać takie odwzorowania dużymi literami łacińskimi), gdy dla każdych dwóch wektorów x i y oraz każdej liczby $\lambda \in \mathbb{K}$ jest

$$\text{(i)} \qquad A(x + y) = A(x) + A(y)\,,$$
$$\text{(ii)} \qquad A(\lambda x) = \lambda A(x)\,.$$

Zwróćmy uwagę, że – podobnie jak dla homomorfizmów grup – działania po lewej stronie odnoszą się do przestrzeni V, a po prawej – do przestrzeni W. Z drugiego warunku definicyjnego mamy $A(\vec{0}) = A(0\vec{0}) = 0A(\vec{0})$, więc

$$A(\vec{0}) = \vec{0}\,.$$

Często stosowaną konwencją, którą przyjmiemy też w tym podręczniku, jest notowanie wartości odwzorowania liniowego dla zadanego argumentu bez nawiasów funkcyjnych:

$$A : V \mapsto W\,, \quad x \mapsto Ax\,.$$

Warunki definicyjne odwzorowania liniowego można zapisać łącznie w postaci jednego równania

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay\,.$$

Analogicznie do pojęć teorii grup wprowadzamy dalsze terminy: bijektywne odwzorowanie liniowe nazywamy ***izomorfizmem liniowym***, dowolne odwzorowanie liniowe przestrzeni w samą siebie ($W = V$) nazywamy ***operatorem (liniowym) w*** V (lub ***endomorfizmem liniowym***), a bijektywny endomorfizm (czyli izomorfizm przestrzeni na samą siebie) nazywamy ***automorfizmem liniowym***.

Odwzorowania liniowe zaliczyć trzeba do najważniejszych obiektów algebry liniowej, znaczna część dalszego wykładu im właśnie będzie poświęcona. Aby lepiej przyswoić sens ich definicji, rozważymy kilka przykładów.

2 Przykłady

(i) Pochodna i całka
Oznaczmy przez $\mathcal{C}^k$, $k = 0, 1 \ldots$, przestrzeń wektorową funkcji zespolonych na zbiorze liczb rzeczywistych, k-krotnie różniczkowalnych w sposób ciągły ($\mathcal{C}^0$ jest przestrzenią funkcji ciągłych). Odwzorowania

$$\mathcal{C}^{k+1} \ni f \mapsto Df \in \mathcal{C}^k\,, \quad (Df)(x) = \frac{df(x)}{dx}\,,$$
$$\mathcal{C}^k \ni f \mapsto If \in \mathcal{C}^{k+1}\,, \quad (If)(x) = \int_0^x f(y)\,dy$$

są odwzorowaniami liniowymi. Podobnie określamy przestrzenie $\mathcal{C}^k_{\mathbb{R}}$ złożone z funkcji rzeczywistych oraz działające między nimi odwzorowania D i I.

(ii) Pochodna, całka i mnożenie przez argument jako operatory
Oznaczmy przez $\mathcal{C}^\infty$ przestrzeń wektorową funkcji zespolonych na zbiorze liczb rzeczywistych, które mają pochodne wszystkich rzędów. Przy zacieśnieniu dziedzin i przeciwdziedzin odwzorowań określonych w poprzednim przykładzie do tej przestrzeni, dostajemy operatory liniowe (oznaczone, nie całkiem właściwie, ale zwięźle, tymi samymi symbolami):

$$D : \mathcal{C}^\infty \mapsto \mathcal{C}^\infty\,, \quad I : \mathcal{C}^\infty \mapsto \mathcal{C}^\infty\,.$$

Operatorem liniowym jest również odwzorowanie mnożenia funkcji przez jej argument:
$$\mathcal{C}^\infty \ni f \mapsto Xf \in \mathcal{C}^\infty\,, \quad (Xf)(x) = xf(x)\,.$$

Podobnie dla rzeczywistych przestrzeni $\mathcal{C}^\infty_{\mathbb{R}}$.

(iii) Pochodna, całka i mnożenie przez argument na przestrzeniach wielomianowych
Niech $\mathcal{P}_n$, $n = 0, 1 \ldots$, oznacza przestrzeń zespolonych wielomianów zmiennej rzeczywistej, stopnia $\leq n$. Odwzorowania D_n, I_n i X_n o takich przepisach działania, jak D, I i X odpowiednio, ale działające między przestrzeniami:

$$D_n : \mathcal{P}_n \mapsto \mathcal{P}_{n-1}\,, \quad I_n : \mathcal{P}_n \mapsto \mathcal{P}_{n+1}\,, \quad X_n : \mathcal{P}_n \mapsto \mathcal{P}_{n+1}\,,$$

są odwzorowaniami liniowymi między skończenie wymiarowymi przestrzeniami. ($\mathcal{P}_{-1}$ jest przestrzenią zero-wymiarową, złożoną z samej funkcji zerowej.)

(iv) Całka oznaczona
Zbiór zespolonych funkcji ciągłych na przedziale $\langle 0,1\rangle$ tworzy przestrzeń wektorową, oznaczmy ją $\mathcal{C}(\langle 0,1\rangle)$. Odwzorowanie

$$\mathcal{C}(\langle 0,1\rangle) \ni f \mapsto \int_0^1 f\,dx \in \mathbb{R}$$

jest odwzorowaniem liniowym.

(v) Odwzorowanie macierzowe
Niech A będzie macierzą z przestrzeni $\mathbb{K}^{m\times n}$. Odwzorowanie

$$\mathbb{K}^{n\times p} \ni X \mapsto AX \in \mathbb{K}^{m\times p}$$

jest odwzorowaniem liniowym. Jeśli $B \in \mathbb{K}^{p\times r}$, to również odwzorowanie

$$\mathbb{K}^{n\times p} \ni X \mapsto AXB \in \mathbb{K}^{m\times r}$$

jest liniowe.

3 Konstrukcja odwzorowania liniowego

Wskażemy jeden ze sposobów zadawania odwzorowań spełniających warunki liniowości. Ma on zarówno znaczenie praktyczne, jak też jest narzędziem wielu dowodów w teorii odwzorowań liniowych. Zauważmy najpierw, że jeśli $A : V \mapsto W$ jest dowolnym odwzorowaniem liniowym, a $E \subseteq V$ dowolną bazą, to z liniowości

$$A\sum \lambda_i e_i = \sum \lambda_i A e_i$$

dla dowolnej kombinacji liniowej wektorów bazowych (sumowanie po dowolnym skończonym podzbiorze bazy). Ponieważ każdy wektor ma jednoznaczny rozkład na wektory bazowe, to równość powyższa odtwarza wartość odwzorowania A dla dowolnego argumentu za pomocą jego wartości na wektorach bazowych. Okazuje się, że fakt ten może również posłużyć do konstrukcji odwzorowania liniowego.

Twierdzenie 1. *Niech V i W będą przestrzeniami wektorowymi, a E dowolną bazą przestrzeni V. Niech ponadto $A|_E : E \mapsto W$ będzie dowolnym odwzorowaniem. Wtedy istnieje dokładnie jedno odwzorowanie liniowe $A : V \mapsto W$ będące rozszerzeniem odwzorowania $A|_E$.*

Dowód. Z założenia odwzorowanie A jest zadane na wektorach bazowych. Jeśli A ma być liniowe, to jego wartość na kombinacji liniowej wektorów bazowych musi być zadana równością poprzedzającą twierdzenie. Każdy wektor ma jednoznaczny rozkład na wektory bazowe, więc równość ta zadaje jednoznacznie A

na całej przestrzeni. Pozostaje jedynie sprawdzić, że otrzymane odwzorowanie jest liniowe. Niech wektory x i y mają rozkłady

$$x = \sum_{i=1}^{n} \alpha_i e_i \quad \text{i} \quad y = \sum_{i=1}^{n} \beta_i e_i$$

(x i y wykorzystują łącznie skończony podzbiór bazy, więc ich rozkłady zawsze można tak zapisać, ewentualnie z niektórymi spośród współczynników α_i, β_i równymi zero). Wtedy stosując definicję odwzorowania A do wektorów $\lambda x + \mu y$ oraz x i y dostajemy

$$A(\lambda x + \mu y) = A \sum_{i=1}^{n} (\lambda \alpha_i + \mu \beta_i) e_i = \sum_{i=1}^{n} (\lambda \alpha_i + \mu \beta_i) A e_i =$$
$$= \lambda \sum_{i=1}^{n} \alpha_i A e_i + \mu \sum_{i=1}^{n} \beta_i A e_i = \lambda A x + \mu A y \,,$$

co kończy dowód. □

4 Obraz i przeciwobraz podprzestrzeni

Z zachowywania przez odwzorowania liniowe struktury liniowej wynika zachowywanie własności bycia podprzestrzenią.

Twierdzenie 2. *Jeśli $A : V \mapsto W$ jest odwzorowaniem liniowym, V' jest podprzestrzenią przestrzeni V, a W' podprzestrzenią przestrzeni W, to $A(V') \subseteq W$ i $A^{-1}(W') \subseteq V$ są podprzestrzeniami.*

Dowód. Zauważmy na początku, że zarówno $A(V')$ jak i $A^{-1}(W')$ są niepuste, bo należą do nich wektory zerowe. Jeśli $w, v \in A(V')$, to istnieją $x, y \in V'$ takie, że $w = Ax$, $v = Ay$. Wtedy z liniowości A mamy

$$\alpha w + \beta v = A(\alpha x + \beta y) \in A(V') \,;$$

z kryterium tw. 2, §8, wnioskujemy, że $A(V')$ jest podprzestrzenią. Podobnie, niech $r, s \in A^{-1}(W')$. Oznacza to, że $Ar, As \in W'$, więc również

$$A(\alpha r + \beta s) = \alpha A r + \beta A s \in W' \,, \quad \text{czyli} \;\; \alpha r + \beta s \in A^{-1}(W') \,.$$

□

Dwa szczególne przypadki podprzestrzeni, o których mówi twierdzenie, zasługują na wyróżnienie – otrzymujemy je kładąc $V' = V$ lub $W' = \{0\}$ odpowiednio. ***Obrazem odwzorowania liniowego*** A nazywamy podprzestrzeń

$$\operatorname{Im} A := A(V) \equiv \{Ax \in W \,|\, x \in V\} \subseteq W \,,$$

a ***jądrem odwzorowania liniowego*** A – podprzestrzeń

$$\operatorname{Ker} A := A^{-1}(\{0\}) \equiv \{x \in V |\, Ax = 0\} \subseteq V\,.$$

Z liniowości A widać, że jeśli E jest dowolną bazą przestrzeni V, to

$$\operatorname{Im} A = L\big(A(E)\big)\,.$$

Wymiary jądra oraz obrazu odwzorowania liniowego spełniają zależność daną następującym twierdzeniem.

Twierdzenie 3. *Jeśli $A : V \mapsto W$ jest odwzorowaniem liniowym, to*

$$\dim \operatorname{Ker} A + \dim \operatorname{Im} A = \dim V\,.$$

Dowód. Przeprowadzimy dowód dla przypadku, gdy $\dim V < \infty$, ale dość oczywiste modyfikacje prowadzą do uogólnienia tego dowodu również na przypadek $\dim V = \infty$ (w tej sytuacji teza sprowadza się do stwierdzenia: $\dim \operatorname{Ker} A = \infty$ lub $\dim \operatorname{Im} A = \infty$). Niech $E = \{e_1, \ldots, e_n\}$ będzie taką bazą przestrzeni V, że $E' = \{e_1, \ldots, e_m\}$ $(m \leq n)$ jest bazą $\operatorname{Ker} A$. Obraz $\operatorname{Im} A$ jest powłoką liniową zbioru $A(E)$, ale ponieważ $A(E') = \{0\}$, to również

$$\operatorname{Im} A = L(Ae_{m+1}, \ldots, Ae_n)\,.$$

Wektory $Ae_{m+1}, \ldots, Ae_n$ są liniowo niezależne: jeśli $\displaystyle\sum_{i=m+1}^{n} \lambda_i A e_i = 0$, to z liniowości A mamy $\displaystyle\sum_{i=m+1}^{n} \lambda_i e_i \in \operatorname{Ker} A$, a więc $\displaystyle\sum_{i=m+1}^{n} \lambda_i e_i = \sum_{i=1}^{m} \mu_i e_i$. To jest możliwe, tylko gdy $\lambda_i = 0$, $i = m+1, \ldots, n$ (oraz $\mu_i = 0$, $i = 1, \ldots, m$). Tak więc wektory $Ae_{m+1}, \ldots, Ae_n$ stanowią bazę $\operatorname{Im} A$, co kończy dowód twierdzenia. $\square$

5 Składanie odwzorowań liniowych

Odwzorowania liniowe, podobnie jak każde inne, można składać zgodnie z ogólnymi regułami. W przyjętej przez nas notacji złożenie odwzorowań liniowych notuje się bez znaku $\circ$ między symbolami odwzorowań: BA. Należy przy tym pamiętać, że w zapisie BAx argumentem odwzorowania A jest wektor x, a odwzorowania B – wektor Ax. Co istotne, liniowość jest zachowywana przez operację złożenia, zgodnie z następującym twierdzeniem.

Twierdzenie 4. *Jeśli $A : V \mapsto W$ i $B : W \mapsto Y$ są odwzorowaniami liniowymi, to BA jest też odwzorowaniem liniowym. Jeśli ponadto A i B są izomorfizmami liniowymi, to A^{-1} i BA są też izomorfizmami liniowymi.*

Dowód. Jeśli A i B są liniowe, to dla każdej kombinacji liniowej wektorów $x, y \in X$ mamy

$$BA(\alpha x + \beta y) = B(\alpha Ax + \beta Ay) = \alpha\, BAx + \beta\, BAy\,,$$

więc BA jest odwzorowaniem liniowym. Jeśli ponadto A jest bijekcją, to A^{-1} jest też bijekcją oraz dla każdej kombinacji liniowej wektorów $u, v \in W$ mamy

$$\alpha u + \beta v = A(\alpha A^{-1}u + \beta A^{-1}v) \qquad (\text{z liniowości } A)\,.$$

Wyliczając wartość odwzorowania A^{-1} dla obu stron dostajemy

$$A^{-1}(\alpha u + \beta v) = \alpha A^{-1}u + \beta A^{-1}v\,.$$

Stąd A^{-1} jest bijektywnym odwzorowaniem liniowym, czyli izomorfizmem liniowym. Jeśli zarówno A jak i B są izomorfizmami liniowymi, to jak już wiemy, BA jest homomorfizmem liniowym, ale z bijektywności A i B jest też bijekcją, więc izomorfizmem liniowym. □

6 Przykłady

(i) Jądro i obraz odwzorowania liniowego
Dla odwzorowań liniowych D i I, określonych tak, jak w przykładzie (i), p. 2, mamy

$$\operatorname{Ker} D = \mathcal{P}_0\,,\ \operatorname{Im} D = \mathcal{C}^k\,,\ \operatorname{Ker} I = \{0\}\,,\ \operatorname{Im} I = \{f \in \mathcal{C}^{k+1} \mid f(0) = 0\}\,.$$

Żadne z tych odwzorowań nie jest bijektywne, ale zachodzi

$$DI = \mathrm{id}_{\mathcal{C}^k}\ .$$

Twierdzenie 7, §2, mówi więc, że jeśli wprowadzić I' i D' przez zacieśnienie przeciwdziedziny I i dziedziny D do $\operatorname{Im} I$, to I' i D' są wzajemnie odwrotnymi bijekcjami, więc izomorfizmami liniowymi.

(ii) Nieprzemienność składania
Rozważmy odwzorowania liniowe D i X jako operatory w przestrzeni $\mathcal{C}^\infty$ (przykład (ii), p. 2). Złożenia tych operatorów są znów operatorami w tej przestrzeni, o działaniu:

$$(DXf)(x) = \frac{d}{dx}(Xf)(x) = \frac{d}{dx}\big(xf(x)\big) = f(x) + x\frac{df(x)}{dx}\,,$$
$$(XDf)(x) = x(Df)(x) = x\frac{df(x)}{dx}\,.$$

Stąd

$$DX - XD = \mathrm{id}_{\mathcal{C}^\infty}\ .$$

7 Izomorfizmy. Przestrzenie izomorficzne

Podobnie jak dla homomorfizmów grupowych, tak i dla homomorfizmów liniowych istnieje związek między różnowartościowością homomorfizmu a trywialnością jego jądra.

Twierdzenie 5. *Odwzorowanie liniowe A jest różnowartościowe wtedy, i tylko wtedy, gdy jego jądro jest trywialne (tj.* $\operatorname{Ker} A = \{0\}$*). Stąd A jest izomorfizmem wtedy, i tylko wtedy, gdy* $\operatorname{Ker} A = \{0\}$ *i* $\operatorname{Im} A = W$.

Dowód. Jeśli A jest różnowartościowe, to w szczególności $A^{-1}(\{0\})$ może mieć co najwyżej jeden element, więc $\operatorname{Ker} A = \{0\}$. Odwrotnie, niech $\operatorname{Ker} A = \{0\}$ i niech $Ax_1 = Ax_2$. Wtedy z liniowości

$$A(x_1 - x_2) = 0\,, \quad \text{więc} \ \ x_1 - x_2 \in \operatorname{Ker} A = \{0\}\,.$$

Stąd $x_1 = x_2$, więc A jest różnowartościowe. Drugie zdanie twierdzenia jest teraz oczywiste. □

Odwzorowanie identycznościowe id_V jest w oczywisty sposób operatorem liniowym. Jest ono bijekcją, więc automorfizmem liniowym, i w zgodzie z twierdzeniem $\operatorname{Ker} \mathrm{id}_V = \{0\}$, $\operatorname{Im} \mathrm{id}_V = V$.

Również analogicznie jak w teorii grup określamy klasy przestrzeni, które są podobne w swojej strukturze liniowej. Mówimy, że przestrzenie wektorowe V i W są ***izomorficzne***, gdy istnieje izomorfizm liniowy przestrzeni V na W. Warunek ten zadaje relację równoważności w zbiorze przestrzeni wektorowych: identyczność jest izomorfizmem liniowym przestrzeni na samą siebie, więc relacja jest zwrotna, a z twierdzenia 4 wynika też, że jest symetryczna i przechodnia. Relacja ta dzieli więc zbiór wszystkich przestrzeni wektorowych na klasy przestrzeni izomorficznych.

8 Izomorfizmy przestrzeni skończenie wymiarowych

Istnieje inne, bardzo użyteczne, kryterium izomorficzności dla przypadku przestrzeni skończenie wymiarowych.

Twierdzenie 6. *Niech $A : V \mapsto W$ będzie odwzorowaniem liniowym i niech* $\dim V < \infty$. *Wtedy następujące warunki są równoważne:*

(i) *A jest izomorfizmem;*

(ii) $\dim W = \dim V$ *i* $\operatorname{Ker} A = \{0\}$*;*

(iii) $\dim W = \dim V$ *i* $\operatorname{Im} A = W$.

Dowód. Jeśli A jest izomorfizmem, to $\operatorname{Ker} A = \{0\}$ i $\operatorname{Im} A = W$, więc

$$\dim \operatorname{Ker} A = 0 \quad \text{i} \quad \dim \operatorname{Im} A = \dim W\,.$$

Z twierdzenia 3 dostajemy wtedy $\dim W = \dim V$, więc warunek (i) pociąga (ii) i (iii). Jeśli $\dim W = \dim V < \infty$, to na mocy tego samego twierdzenia mamy

$$\dim \operatorname{Ker} A + \dim \operatorname{Im} A = \dim W < \infty\,.$$

Stąd widać natychmiast, że warunki (ii) i (iii) są równoważne. Łącznie warunki (ii) i (iii) pociągają (i) na mocy poprzedniego twierdzenia, co kończy dowód. □

Wniosek 7. *Jeśli przestrzeń wektorowa V jest skończenie wymiarowa, to przestrzeń W jest izomorficzna z V wtedy, i tylko wtedy, gdy ma taki sam wymiar jak V.*

Dowód. Jeśli istnieje izomorfizm przestrzeni V w W, to przestrzenie te mają równy wymiar na mocy poprzedniego twierdzenia. Odwrotnie, jeśli wymiary przestrzeni V i W są równe $n \in \mathbb{N}$, to wybierzmy dowolną bazę $(e_1, \ldots, e_n)$ przestrzeni V oraz dowolną bazę $(f_1, \ldots, f_n)$ przestrzeni W, i zadajmy odwzorowanie liniowe $F : V \mapsto W$ kładąc $Fe_i = f_i$. Wtedy $\operatorname{Im} F = W$, więc F jest izomorfizmem (znów na mocy poprzedniego twierdzenia). □

9 Izomorfizm V z $\mathbb{K}^{\dim V}$. Notacja

Niech V będzie n-wymiarową przestrzenią wektorową, a $(e_1, \ldots, e_n)$ jej dowolną bazą uporządkowaną. Wprowadzamy następujące konwencje, których odtąd stale będziemy używać. Współrzędne wektora x w tej bazie oznaczać będziemy x^i, a więc tym samym symbolem rdzeniowym co symbol wektora, z dodanym górnym indeksem. W zapisie rozkładu wektora na wektory bazowe pomijamy znak sumy, traktując go jako domyślny:

$$x = \sum_{i=1}^{n} x^i e_i = x^i e_i\,.$$

Podobnie w innych wyrażeniach, jeśli ten sam wskaźnik występuje jednokrotnie jako indeks górny i jednokrotnie jako indeks dolny, to traktujemy go jak indeks sumacyjny i sumujemy po całym jego zakresie. Umowa ta nosi nazwę ***konwencji sumacyjnej Einsteina***. Kolumnę współrzędnych wektora x będziemy oznaczać

$$\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}\,.$$

Baza $(e_1, \ldots, e_n)$ przestrzeni V zadaje naturalny izomorfizm J tej przestrzeni z przestrzenią $\mathbb{K}^{n\times 1}$, otrzymujemy go kładąc

$$Je_i = \mathbf{e}_i\,, \quad i = 1, \ldots, n\,,$$

gdzie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest bazą kanoniczną przestrzeni $\mathbb{K}^{n\times 1}$. Dla dowolnego wektora $x \in V$ dostajemy stąd

$$Jx = \mathbf{x}\,.$$

10 Przestrzeń wektorowa odwzorowań liniowych

Niech będzie dany dowolny zbiór X oraz przestrzeń wektorowa W nad ciałem $\mathbb{K}$. Rozważmy zbiór wszystkich odwzorowań $X \mapsto W$. Na tym zbiorze można wprowadzić strukturę liniową w sposób analogiczny do tego, jakiego użyliśmy w prostszym przypadku, gdy $W = \mathbb{K}$ (w punkcie 12, §3; wprowadzona tam struktura pierścienia nie uogólnia się na obecny przypadek). Jeśli więc A i B są odwzorowaniami z tej rodziny, to definiujemy

$$(\lambda A + \mu B)(x) = \lambda A(x) + \mu B(x)\,.$$

Zwracamy uwagę, jak czyniliśmy to już wcześniej w analogicznych przypadkach, że mnożenie przez liczbę i dodawanie po prawej stronie powyższej równości są znanymi operacjami (w przestrzeni W), a te same znaki po lewej stronie równości są nowymi, definiowanymi przez tę równość, działaniami. Analogicznie jak w prostszym przypadku $W = \mathbb{K}$ łatwo pokazać, że zbiór odwzorowań z tą strukturą liniową tworzy przestrzeń wektorową. Elementem neutralnym dodawania odwzorowań jest odwzorowanie zerowe. Sprawdzenie warunków definicyjnych przestrzeni wektorowej jest prostym ćwiczeniem, które pozostawiamy czytelnikowi do uzupełnienia.

Niech teraz zbiór X będzie przestrzenią wektorową V. Wśród wszystkich odwzorowań $V \mapsto W$ wyróżniliśmy odwzorowania liniowe. Niech A i B będą odwzorowaniami liniowymi. Wtedy ciąg równości:

$$\begin{aligned}
(\lambda A{+}\mu B)(\alpha x + \beta y) &= \lambda A(\alpha x + \beta y) + \mu B(\alpha x + \beta y) = \\
&= \lambda(\alpha Ax + \beta Ay) + \mu(\alpha Bx + \beta By) = \\
&= \alpha(\lambda Ax + \mu Bx) + \beta(\lambda Ay + \mu By) = \alpha(\lambda A + \mu B)x + \beta(\lambda A + \mu B)y
\end{aligned}$$

pokazuje, że odwzorowanie $\lambda A{+}\mu B$ jest również liniowe (skorzystaliśmy z liniowości A i B oraz dwukrotnie z definicji $\lambda A{+}\mu B$). Z podstawowego kryterium dla podprzestrzeni, tw. 2, §8, mamy więc pierwszą część następującego twierdzenia.

Twierdzenie 8. *Niech V i W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$. Wtedy zbiór odwzorowań liniowych z V w W również tworzy przestrzeń wektorową nad ciałem $\mathbb{K}$; oznaczamy ją $\mathcal{L}(V, W)$. Wymiar tej przestrzeni jest równy*

$$\dim \mathcal{L}(V, W) = \dim V \dim W \,.$$

Stwierdzenie dotyczące wymiaru przestrzeni wynika z dyskusji, którą przeprowadzimy w następnym punkcie.

Składanie odwzorowań liniowych, podobnie jak mnożenie macierzy, posiada własność liniowości względem każdego ze składanych odwzorowań.

Twierdzenie 9. *Niech $A, B \in \mathcal{L}(V, W)$ oraz $C, D \in \mathcal{L}(W, U)$. Wtedy dla każdych $\alpha, \beta \in \mathbb{K}$ jest*

$$C(\alpha A + \beta B) = \alpha CA + \beta CB \,, \quad (\alpha C + \beta D)A = \alpha CA + \beta DA \,.$$

Dowód. Dla każdego wektora $x \in V$ mamy ciąg równości

$$\begin{aligned}[C(\alpha A + \beta B)]x = C[(\alpha A + \beta B)x] &= C(\alpha Ax + \beta Bx) = \\ &= \alpha CAx + \beta CBx = [\alpha CA + \beta CB]x \,,\end{aligned}$$

co dowodzi pierwszej równości. Analogicznie wykazuje się drugą równość. □

11 Izomorfizm $\mathcal{L}(V, W)$ z $\mathbb{K}^{\dim W \times \dim V}$

Wybierzmy w n-wymiarowej przestrzeni V bazę $(e_1, \ldots, e_n)$ oraz w m-wymiarowej przestrzeni W bazę $(f_1, \ldots, f_m)$. Jeśli A jest odwzorowaniem liniowym V w W, to rozkładając wektory $(Ae_1, \ldots, Ae_n)$ w bazie $(f_1, \ldots, f_m)$ dostajemy

$$Ae_i = f_j A^j{}_i \,,$$

gdzie liczby $A^j{}_i$ są współczynnikami rozkładów (sumowanie po j zgodnie z konwencją Einsteina). Macierz $\mathbf{A} := (A^j{}_i)$ (górny wskaźnik numeruje wiersze, a dolny – kolumny) nazywamy ***macierzą odwzorowania liniowego*** A ***w bazach*** $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$. Sposób oznaczania tej macierzy i jej elementów jest następną ogólną konwencją, którą będziemy stosować w dalszym ciągu wykładu. Powyższą definicję macierzy $\mathbf{A}$ można teraz zapisać symbolicznie jako

$$\begin{pmatrix} Ae_1 & \ldots & Ae_n \end{pmatrix} = \begin{pmatrix} f_1 & \ldots & f_m \end{pmatrix} \mathbf{A} \,.$$

Odwrotnie, jeśli zadana jest macierz $\mathbf{A}$, to powyższe związki jednoznacznie zadają odwzorowanie liniowe A, dostajemy zatem wzajemnie jednoznaczne odwzorowanie (związane z wyborem baz $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$)

$$T : \mathcal{L}(V, W) \mapsto \mathbb{K}^{m \times n} \,, \quad A \mapsto T(A) = \mathbf{A} \,.$$

Ponadto mamy

$$(\lambda A + \mu B)e_i = \lambda A e_i + \mu B e_i = \lambda f_j A^j{}_i + \mu f_j B^j{}_i = f_j(\lambda A^j{}_i + \mu B^j{}_i)\,,$$

więc

$$T(\lambda A + \mu B) = \lambda \mathbf{A} + \mu \mathbf{B} = \lambda T(A) + \mu T(B)\,.$$

Odwzorowanie T jest więc izomorfizmem liniowym (mimo liniowości zachowujemy w tym przypadku nawiasy funkcyjne w symbolu $T(A)$, aby odróżnić go od symbolu złożenia odwzorowań liniowych). Wynika stąd ostatnie stwierdzenie w twierdzeniu 8 z poprzedniego punktu.

Wyjaśnimy bliżej znaczenie elementów macierzowych $A^j{}_i$. Wprowadźmy układ odwzorowań liniowych

$$E^i{}_j \in \mathcal{L}(V,W)\,, \quad E^i{}_j\, x = x^i\, f_j\,, \quad i = 1,\ldots,n\,,\ j = 1,\ldots,m\,,$$

gdzie x^i są współrzędnymi wektora x w bazie $(e_1,\ldots,e_n)$. Z przepisu definicyjnego dostajemy w szczególności

$$E^i{}_j e_k = \delta^i_k\, f_j = f_l\, \delta^l_j\, \delta^i_k\,, \quad \text{skąd} \quad (E^i{}_j)^l{}_k = \delta^l_j\, \delta^i_k$$

(δ^j_i ma takie samo znaczenie jak δ_{ji}, sposób notowania wskaźników podyktowany jest obraną konwencją). Oznacza to, że macierze

$$T(E^i{}_j) = \mathbf{E}^i{}_j$$

są wektorami bazy kanonicznej przestrzeni $\mathbb{K}^{m\times n}$, macierz $\mathbf{E}^i{}_j$ ma jedynkę na skrzyżowaniu j-tego wiersza i i-tej kolumny. Ponieważ T jest izomorfizmem, wynika stąd, że odwzorowania $(E^i{}_j)$, $i \in \{1,\ldots,n\}$, $j \in \{1,\ldots,m\}$, stanowią bazę przestrzeni $\mathcal{L}(V,W)$. Przy jej użyciu możemy zapisać

$$Ax = x^i A e_i = x^i f_j\, A^j{}_i = A^j{}_i E^i{}_j x$$

dla każdego wektora $x \in V$, więc

$$A = A^j{}_i E^i{}_j$$

– elementy macierzowe $A^j{}_i$ są współrzędnymi odwzorowania A, jako wektora przestrzeni $\mathcal{L}(V,W)$, w bazie $(E^i{}_j)$.

Jeśli A jest operatorem liniowym, tj. $A \in \mathcal{L}(V,V)$, to zawsze wybieramy w V jako dziedzinie i jako przeciwdziedzinie tę samą bazę, więc w tym przypadku macierz operatora w bazie $(e_1,\ldots,e_n)$ określona jest przez

$$Ae_i = e_j A^j{}_i\,.$$

W szczególności, dla odwzorowania identycznościowego dostajemy $\mathrm{id}\, e_i = e_j\, \delta^j_i$, więc obrazem identyczności w izomorfizmie zadanym bazą jest macierz jednostkowa,

$$T(\mathrm{id}_V) = \mathbf{1}\,.$$

12 Pokrewieństwo struktur funkcyjnych przestrzeni $\mathcal{L}(V,W)$ i $\mathbb{K}^{\dim W\times\dim V}$

W poprzednich punktach mówiliśmy o podobieństwie struktur liniowych przestrzeni V i $\mathbb{K}^{\dim V}$, przestrzeni $\mathcal{L}(V,W)$ i $\mathbb{K}^{\dim W\times\dim V}$, oraz o konkretnych realizacjach tych pokrewieństw za pomocą izomorfizmów zadanych wyborem baz w V i W. Pokażemy teraz, że te same izomorfizmy wyznaczają też związek własności funkcyjnych przestrzeni $\mathcal{L}(V,W)$ i przestrzeni $\mathbb{K}^{\dim W\times\dim V}$.

Zauważmy na początku, że w przypadku przestrzeni $\mathbb{K}^p$ istnieje naturalny wybór bazy, jako bazy kanonicznej. Oznaczmy przez $(\mathbf{e}_i)$ i $(\mathbf{f}_j)$ bazy kanoniczne w $\mathbb{K}^n$ i $\mathbb{K}^m$ odpowiednio. Wtedy łatwo sprawdzić, że dla każdej macierzy $\mathbf{A}=(A^i{}_j)\in\mathbb{K}^{m\times n}$ mamy

$$\mathbf{A}\mathbf{e}_i=\mathbf{f}_jA^j{}_i\,.$$

Prawa strona powyższej równości jest identyczna z określeniem macierzy pewnego odwzorowania liniowego z $\mathcal{L}(\mathbb{K}^n,\mathbb{K}^m)$. Odwzorowanie to, jak pokazuje równość, jest zadane macierzą $\mathbf{A}$ jako

$$\mathbb{K}^n\ni\mathbf{x}\mapsto\mathbf{A}\mathbf{x}\in\mathbb{K}^m$$

– jego liniowość jest natychmiast widoczna na podstawie własności mnożenia macierzowego. Widzimy więc, że przy takim określeniu działania macierzy z $\mathbb{K}^{m\times n}$ na wektory z $\mathbb{K}^n$, izomorfizm przestrzeni $\mathcal{L}(\mathbb{K}^n,\mathbb{K}^m)$ z przestrzenią $\mathbb{K}^{m\times n}$ można uważać za utożsamienie. Mówiąc dalej o macierzach z $\mathbb{K}^{m\times n}$ jako odwzorowaniach liniowych, będziemy mieli na myśli to właśnie działanie.

Możemy teraz sformułować główny rezultat tego punktu.

Twierdzenie 10. *Niech będą zadane bazy uporządkowane w przestrzeniach V i W oraz określone przez nie izomorfizmy przestrzeni V, W i $\mathcal{L}(V,W)$ odpowiednio z $\mathbb{K}^n$, $\mathbb{K}^m$ i $\mathbb{K}^{m\times n}$. Wtedy dla $x\in V$, $y\in W$ i $A\in\mathcal{L}(V,W)$*

$$y=Ax\Longleftrightarrow\mathbf{y}=\mathbf{A}\mathbf{x}\,,$$

czyli działaniu odwzorowań $\mathcal{L}(V,W)$ odpowiada działanie odwzorowań $\mathbb{K}^{m\times n}$. Jeśli ponadto wybrano bazę w przestrzeni Y oraz odpowiednie izomorfizmy przestrzeni Y i $\mathcal{L}(W,Y)$ z $\mathbb{K}^k$ i $\mathbb{K}^{k\times m}$, to dla $B\in\mathcal{L}(W,Y)$ i $C\in\mathcal{L}(V,Y)$ mamy

$$C=BA\Longleftrightarrow\mathbf{C}=\mathbf{B}\mathbf{A}\,,$$

czyli składaniu odwzorowań $\mathcal{L}(V,W)$ z odwzorowaniami $\mathcal{L}(W,Y)$ odpowiada mnożenie macierzy $\mathbb{K}^{m\times n}$ przez macierze $\mathbb{K}^{k\times m}$.

Dowód. Oznaczmy bazy wybrane w V i W odpowiednio przez (e_i) i (f_j). Korzystając z rozkładów $y = y^j f_j$, $x = x^i e_i$ oraz z definicji macierzy odwzorowania liniowego, zapisujemy równanie $y = Ax$ w równoważnej postaci

$$y^j f_j = x^i A e_i = f_j A^j{}_i x^i\,, \quad \text{czyli} \quad y^j = A^j{}_i x^i\,,$$

lub w postaci macierzowej $\mathbf{y} = \mathbf{A}\mathbf{x}$, co dowodzi pierwszego stwierdzenia. Stąd dla odwzorowania B jak w założeniu również mamy $z = By$ wtedy, i tylko wtedy, gdy $\mathbf{z} = \mathbf{B}\mathbf{y}$. Łącząc obie równoważności dostajemy: $z = BAx$ wtedy, i tylko wtedy, gdy $\mathbf{z} = \mathbf{B}\mathbf{A}\mathbf{x}$, skąd wynika drugie stwierdzenie. □

Odpowiedniość struktur liniowych oraz struktur funkcyjnych pozwala zamienić każdy problem dotyczący dowolnej przestrzeni skończenie wymiarowej oraz odwzorowań takich przestrzeni na problem w przestrzeniach $\mathbb{K}^p$, czyli macierzowy. W szczególności, wektory $(x_1, \ldots, x_r)$ są liniowo niezależne wtedy, i tylko wtedy, gdy wektory $(\mathbf{x}_1, \ldots, \mathbf{x}_r)$ są liniowo niezależne.

Rozważmy problem znalezienia wymiaru obrazu odwzorowania liniowego. Wiemy, że $\operatorname{Im} A = L(Ae_1, \ldots, Ae_n)$, gdzie $(e_1, \ldots, e_n)$ jest dowolną bazą dziedziny odwzorowania A. Zamieniając ten problem na problem macierzowy mamy więc $\dim \operatorname{Im} A = \dim L(\mathbf{A}\mathbf{e}_1, \ldots, \mathbf{A}\mathbf{e}_n)$. Ale $\mathbf{A}\mathbf{e}_i$ jest i-tą kolumną macierzy $\mathbf{A}$, więc wykazaliśmy następujące twierdzenie.

Twierdzenie 11. *Dla każdego odwzorowania liniowego wymiar jego obrazu jest równy rzędowi jego macierzy w dowolnych bazach.*

Twierdzenie to sugeruje następującą definicję: ***rzędem odwzorowania liniowego*** nazywać będziemy wymiar jego obrazu,

$$\operatorname{rk} A := \dim \operatorname{Im} A = \operatorname{rk} \mathbf{A}\,,$$

ostatnia równość na podstawie twierdzenia.

Rozważmy z kolei, przy zadanym odwzorowaniu A i wektorze b w jego przeciwdziedzinie, problem znalezienia wszystkich takich wektorów x z jego dziedziny, że $Ax = b$. Po zamianie na problem macierzowy zagadnienie sprowadza się do rozwiązania układu równań liniowych $\mathbf{A}\mathbf{x} = \mathbf{b}$. Przechodząc z powrotem od kolumn $\mathbf{x}$ do wektorów x, dostajemy ogólne rozwiązanie postawionego problemu. W szczególności, jądro $\operatorname{Ker} A$ otrzymamy rozwiązując ten problem z $b = 0$. Widzimy teraz, że w przypadku przestrzeni skończenie wymiarowych twierdzenie 3 jest już zawarte w postaci ogólnego rozwiązania znalezionego w §9.

13 Zmiana baz

Jeśli problemy przestrzeni skończenie wymiarowych można zamienić na problemy przestrzeni $\mathbb{K}^p$, to powstaje pytanie, czy nie bardziej ekonomicznie byłoby tylko do tych ostatnich się ograniczyć. Chociaż rzeczywiście można byłoby tak zrobić, to stracilibyśmy w ten sposób pewien niezmiernie istotny aspekt teorii przestrzeni wektorowych: jej niezależność od wyboru baz. Zachowując więc ogólne ramy przestrzeni wektorowych, wyjaśnijmy teraz zależność opisanych izomorfizmów od wyboru baz. Nowe bazy w przestrzeniach V i W oznaczymy $(e'_1, \ldots, e'_n)$ i $(f'_1, \ldots, f'_m)$ odpowiednio. Współrzędne wektorów $x \in V$ i $y \in W$ oznaczymy podobnymi symbolami jak poprzednio, ale z dodanymi „primami": x'^i i y'^j odpowiednio. Podobnie, przy tym nowym wyborze baz, macierz odwzorowania A oznaczymy $\mathbf{A}' = (A'^j{}_i)$.

Twierdzenie 12. *Niech $x \in V$ i $A \in \mathcal{L}(V, W)$, i niech będą dane bazy*

$$\begin{aligned} &(e_1, \ldots, e_n) \quad \textit{i} \quad (e'_1, \ldots, e'_n) \quad \textit{w przestrzeni } V, \\ &(f_1, \ldots, f_m) \quad \textit{i} \quad (f'_1, \ldots, f'_m) \quad \textit{w przestrzeni } W, \end{aligned}$$

wraz ze związkami

$$\begin{aligned} e'_i &= e_j \beta^j{}_i\,, \quad \textit{czyli} \quad \begin{pmatrix} e'_1 & \ldots & e'_n \end{pmatrix} = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} \boldsymbol{\beta}\,, \\ f'_k &= f_l \gamma^l{}_k\,, \quad \textit{czyli} \quad \begin{pmatrix} f'_1 & \ldots & f'_m \end{pmatrix} = \begin{pmatrix} f_1 & \ldots & f_m \end{pmatrix} \boldsymbol{\gamma}\,. \end{aligned}$$

Wtedy

$$\begin{aligned} x'^i &= \beta^{-1i}{}_j x^j\,, & &\textit{czyli} \quad \mathbf{x}' = \boldsymbol{\beta}^{-1}\mathbf{x}\,, \\ A'^i{}_k &= \gamma^{-1i}{}_j\, A^j{}_l\, \beta^l{}_k\,, & &\textit{czyli} \quad \mathbf{A}' = \boldsymbol{\gamma}^{-1}\mathbf{A}\,\boldsymbol{\beta}\,. \end{aligned}$$

W szczególności, jeśli $A \in \mathcal{L}(V, V)$, to

$$A'^i{}_k = \beta^{-1i}{}_j\, A^j{}_l\, \beta^l{}_k\,, \quad \textit{czyli} \quad \mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\,\boldsymbol{\beta}\,.$$

Dowód. Transformacja współrzędnych wektora przy zmianie bazy otrzymana została już wcześniej w p. 9, §8, tutaj przepisaliśmy ją tylko w nowych oznaczeniach. Mamy ciąg równoważności $\mathbf{y}' = \mathbf{A}'\mathbf{x}' \iff y = Ax \iff \mathbf{y} = \mathbf{A}\mathbf{x}$. Wstawiając $\mathbf{x} = \boldsymbol{\beta}\mathbf{x}'$ i analogicznie $\mathbf{y} = \boldsymbol{\gamma}\mathbf{y}'$ dostajemy $\mathbf{y}' = \mathbf{A}'\mathbf{x}' \iff \boldsymbol{\gamma}\mathbf{y}' = \mathbf{A}\boldsymbol{\beta}\mathbf{x}'$, czyli $\boldsymbol{\gamma}\mathbf{A}'\mathbf{x}' = \mathbf{A}\boldsymbol{\beta}\mathbf{x}'$ dla każdego $\mathbf{x}'$, skąd wynikają związki macierzy $\mathbf{A}$ i $\mathbf{A}'$. □

14 Wyznacznik i ślad operatora

Transformację macierzy kwadratowej zadaną ostatnią relacją twierdzenia 12 z poprzedniego punktu nazywamy ***transformacją podobieństwa za pomocą***

macierzy $\boldsymbol{\beta}$. Z twierdzenia Cauchy'ego dla wyznaczników oraz z zachowania śladu iloczynu macierzy przy przestawieniu macierzy (tutaj: macierzy $\boldsymbol{\beta}^{-1}$ i $\mathbf{A}\boldsymbol{\beta}$) wnioskujemy, że ta transformacja zachowuje wyznacznik oraz ślad macierzy:

$$\det \mathbf{A}' = \det \mathbf{A}\,, \quad \operatorname{Tr} \mathbf{A}' = \operatorname{Tr} \mathbf{A}\,,$$

gdzie równość rozumiemy w sensie równości wartości wielomianów po obu stronach równości. Widać stąd, że wielkości te, będąc niezależnymi od bazy, są charakterystykami samego operatora. Definiujemy więc ***wyznacznik operatora liniowego,*** $\det A$, oraz ***ślad operatora liniowego,*** $\operatorname{Tr} A$, jako odpowiednio wartość wyznacznika i wartość śladu jego macierzy w dowolnej bazie. Ze względu na izomorfizm liniowy odwzorowań liniowych i ich macierzy dostajemy liniowość śladu dla operatorów:

$$\operatorname{Tr}(\lambda A + \mu B) = \lambda \operatorname{Tr} A + \mu \operatorname{Tr} B\,,$$

a ze względu na pokrewieństwo ich struktur funkcyjnych – własność śladu:

$$\operatorname{Tr} AB = \operatorname{Tr} BA\,,$$

oraz twierdzenie Cauchy'ego dla operatorów:

$$\det AB = \det A \det B\,.$$

15 Wielomianowe funkcje operatorów

Niech P będzie wielomianem w ciele $\mathbb{K}$, o współczynnikach $(a_0, \ldots, a_n)$, a $A : V \mapsto V$ – dowolnym operatorem w przestrzeni V nad ciałem $\mathbb{K}$. Przez wielokrotne złożenia można nadać sens potęgom A^k, co daje nam możliwość wprowadzenia ***funkcji wielomianowej operatorów***

$$\mathcal{L}(V, V) \ni A \mapsto P(A) = a_0 \operatorname{id} + a_1 A + \ldots a_n A^n \in \mathcal{L}(V, V)$$

dla każdego wielomianu P. Podobnie jak dla funkcji wielomianowych w ciele, porządek składników nie ma znaczenia, a także przy rozkładzie wielomianu na czynniki kolejność tych czynników jest dowolna. Jeśli ciało jest nieskończone, to funkcje wielomianowe w ciele pozostają w bijektywnym związku z wielomianami, więc zadają jednoznacznie odpowiednie funkcje operatorowe. Jeśli jednak ciało jest skończone, to dwa różne wielomiany mogą zadawać tę samą funkcję wielomianową w ciele, ale dwie różne funkcje operatorowe (przykład (viii), p. 17 poniżej).

16 Postać kanoniczna odwzorowania liniowego między różnymi przestrzeniami

Niech $A : V \mapsto W$ będzie odwzorowaniem liniowym, przy czym $V \neq W$. Wybierzmy bazę $(e_1, \ldots, e_n)$ przestrzeni V tak, aby ciąg $(e_{k+1}, \ldots, e_n)$, $k \leq n$, tworzył bazę podprzestrzeni $\operatorname{Ker} A \subseteq V$. Z dowodu twierdzenia 3 wynika, że liniowo niezależne wektory $(Ae_1, \ldots, Ae_k)$ tworzą wtedy bazę podprzestrzeni $\operatorname{Im} A \subseteq W$. Oznaczmy $f_i = Ae_i$ dla $i = 1, \ldots, k$ i ciąg wektorów $(f_1, \ldots, f_k)$ uzupełnijmy w dowolny sposób do bazy $(f_1, \ldots, f_m)$ przestrzeni W.

Zapisujemy teraz działanie odwzorowania A w bazach $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$. Zbierając relacje i definicje tworzące powyższą konstrukcję dostajemy:

$$Ae_i = f_i \ \text{ dla } \ i = 1, \ldots, k\,, \qquad Ae_j = 0 \ \text{ dla } \ j = k+1, \ldots, n\,.$$

Otrzymaliśmy metodę konstrukcji następującej postaci kanonicznej odwzorowania liniowego.

Twierdzenie 13. *Jeśli $A : V \mapsto W$ jest odwzorowaniem liniowym, przy czym $V \neq W$, oraz $\dim V = n$, $\dim W = m$, $\operatorname{rk} A = k$, to istnieją bazy przestrzeni V i W, w których macierz odwzorowania A ma postać*

$$\mathbf{A} = \begin{pmatrix} \mathbf{1}_k & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{K}^{m \times n}\,.$$

Wymiary bloków zerowych są tą relacją jednoznacznie określone i bloki te (lub część z nich) mogą w szczególności znikać.

Konstrukcja powyższa opiera się w istotny sposób na założeniu $V \neq W$ – dzięki niemu mogliśmy wybrać bazę w każdej z tych przestrzeni niezależnie. Trudniejszym zadaniem jest uproszczenie postaci operatora liniowego w przestrzeni V, gdzie mamy możliwość wybrania tylko jednej bazy. Zajmiemy się tym zagadnieniem w §13 oraz w Uzupełnieniach, §25.

17 Przykłady

(i) Uzupełnianie do bazy

Niech $(e_1, \ldots, e_5)$ będzie bazą przestrzeni rzeczywistej V. Dane są trzy liniowo niezależne wektory tej przestrzeni $f_1 = e_1 + e_2 + 2e_3 - e_4$, $f_2 = e_1 + e_3 + 2e_4 + e_5$, $f_3 = -e_1 + 2e_2 + e_3 - 3e_4 - e_5$. Chcemy tę liniową niezależność potwierdzić i uzupełnić układ do bazy przestrzeni V.

Korzystając z izomorfizmu przestrzeni V na $\mathbb{R}^5$ zadanego bazą (e_i), wystarczy rozwiązać odpowiedni problem w $\mathbb{R}^5$. Z kolumn współrzędnych kolejnych wektorów f_1, f_2, f_3 tworzymy macierz

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \\ -1 & 2 & -3 \\ 0 & 1 & -1 \end{pmatrix} .$$

Licząc rząd tej macierzy operujmy jedynie kolumnami – wówczas nie zmienimy liniowych zależności występujących między wierszami (analogicznie, jak przy omawianiu metody Gaussa operując wierszami nie zmienialiśmy liniowych zależności pomiędzy kolumnami). Stwierdzamy w ten sposób, że wiersze drugi, trzeci i czwarty są liniowo niezależne, skąd wnioskujemy, że wszystkie kolumny są liniowo niezależne. Wybrana metoda ma jednak dodatkową zaletę: pozwala nam wnioskować również, że wszystkie wiersze macierzy poszerzonej w następujący sposób:

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 1 & -1 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ -1 & 2 & -3 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 \end{pmatrix} .$$

są liniowo niezależne. Zatem również wszystkie kolumny są liniowo niezależne, co daje nam rozwiązanie postawionego problemu. Kolumny macierzy $\boldsymbol{\beta}$ tworzą bazę przestrzeni $\mathbb{R}^5$, więc wektory $f_4 = e_1$, $f_5 = e_5$ uzupełniają wektory f_1, f_2, f_3 do bazy przestrzeni V. Macierz $\boldsymbol{\beta}$ jest macierzą przejścia z bazy (e_i) do bazy (f_i).

(ii) Macierz odwzorowania liniowego w bazie
Wybierzmy w przestrzeniach $\mathcal{P}_n$ bazy $e_k(x) = x^k$, $k = 0, \ldots, n$. W tych bazach odwzorowania liniowe D_n, I_n, X_n, jak w przykładzie (iii), p. 2, mają działanie

$$D_n e_k = k e_{k-1} , \quad I_n e_k = \frac{1}{k+1} e_{k+1} , \quad X_n e_k = e_{k+1} ,$$

więc ich macierze są postaci

$$\mathbf{D}_n = \left((D_n)^j{}_k\right) = \left(k\delta^j_{k-1}\right) = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & n \end{pmatrix} \in \mathbb{R}^{n \times (n+1)} ,$$

$$\mathbf{I}_n = \big((I_n)^j{}_k\big) = \left(\frac{1}{k+1}\delta^j_{k+1}\right) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \frac{1}{1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{2} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{n+1} \end{pmatrix} \in \mathbb{R}^{(n+2)\times(n+1)},$$

$$\mathbf{X}_n = \big((X_n)^j{}_k\big) = \left(\delta^j_{k+1}\right) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{(n+2)\times(n+1)}.$$

Złożenia $D_{n+1}I_n$, $D_{n+1}X_n$ i $X_{n-1}D_n$ są operatorami w przestrzeni $\mathcal{P}_n$, spełniającymi relacje

$$D_{n+1}I_n = \mathrm{id}_{\mathcal{P}_n}, \quad D_{n+1}X_n - X_{n-1}D_n = \mathrm{id}_{\mathcal{P}_n},$$

co można sprawdzić zarówno wprost z definicji odwzorowań D_l, I_l i X_l, jak też mnożąc ich macierze.

(iii) Transformacja macierzy operatora przy zmianie bazy

Niech $A = \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$, gdzie b i c są różnymi od zera liczbami zespolonymi. Odwzorowanie

$$\mathbb{C}^{2\times 2} \ni X \mapsto S(X) = AXA^{-1} \in \mathbb{C}^{2\times 2}$$

jest operatorem liniowym. Utwórzmy bazę uporządkowaną kładąc $E_1 = E_{12}$, $E_2 = E_{21}$, $E_3 = E_{11}$, $E_4 = E_{22}$, gdzie E_{ij} jest macierzową bazą kanoniczną. Macierz operatora S w tej bazie jest

$$\mathbf{S} = \begin{pmatrix} 0 & e^{\lambda} & 0 & 0 \\ e^{-\lambda} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

gdzie $e^{\lambda} = b/c$. Wprowadzamy nową bazę:

$$E_1' = \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; E_2' = \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; E_3' = \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \; E_4' = \mathbf{1}$$

– macierze σ_i nazywane są ***macierzami Pauliego***. Macierz przejścia z bazy nieprimowanej do primowanej jest równa

$$\beta = \begin{pmatrix} 1 & -i & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

więc macierz operatora S w nowej bazie jest

$$\mathbf{S}' = \boldsymbol{\beta}^{-1}\mathbf{S}\boldsymbol{\beta} = \begin{pmatrix} \cosh\lambda & i\sinh\lambda & 0 & 0 \\ i\sinh\lambda & -\cosh\lambda & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(iv) Jądro i obraz odwzorowania liniowego. Postać kanoniczna
Szukamy jądra i obrazu odwzorowania liniowego $A : V \mapsto W$. Niech $(e_1, \ldots, e_5)$ będzie bazą rzeczywistej przestrzeni V, a $(f_1, \ldots, f_4)$ – bazą rzeczywistej przestrzeni W, i niech macierzą odwzorowania A w tych bazach będzie

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 & 1 & 3 \\ -2 & 1 & 4 & -1 & 1 \\ 1 & 7 & 1 & 2 & 10 \\ -8 & 9 & 18 & -3 & 11 \end{pmatrix}.$$

Posłużymy się metodą omówioną w punkcie 8, §9, i zastosowaną już wcześniej w przykładzie (v), p. 10, §9. Wykonując odpowiednie odjęcia wierszy stwierdzamy, że układ jednorodny z macierzą główną $\mathbf{A}$ jest równoważny układowi z macierzą

$$\begin{pmatrix} 5 & 0 & -9 & 3 & 1 \\ 0 & 5 & 2 & 1 & 7 \end{pmatrix}.$$

Odczytujemy więc, że pierwsze dwie kolumny macierzy $\mathbf{A}$ są liniowo niezależne, a pozostałe są ich kombinacjami liniowymi. Stąd wektory

$$f_1' = Ae_1 = f_1 - 2f_2 + f_3 - 8f_4\,, \quad f_2' = Ae_2 = 2f_1 + f_2 + 7f_3 + 9f_4$$

tworzą bazę przestrzeni $\operatorname{Im} A$.

Jądro odwzorowania A jest utworzone ze wszystkich wektorów $x \in V$ spełniających równanie $Ax = 0$, czyli $\mathbf{Ax} = 0$ w zadanych bazach. Rozwiązanie odczytujemy natychmiast z przekształconej postaci macierzy $\mathbf{A}$:

$$x^1 = \frac{1}{5}(9x^3 - 3x^4 - x^5)\,, \quad x^2 = \frac{1}{5}(-2x^3 - x^4 - 7x^5)\,.$$

Kładąc $x^3 = 5r$, $x^4 = 5s$, $x^5 = 5t$ dostajemy

$$\begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ x^4 \\ x^5 \end{pmatrix} = r \begin{pmatrix} 9 \\ -2 \\ 5 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -3 \\ -1 \\ 0 \\ 5 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -7 \\ 0 \\ 0 \\ 5 \end{pmatrix} .$$

Bazę przestrzeni $\operatorname{Ker} A$ tworzą wektory

$$e'_3 = 9e_1 - 2e_2 + 5e_3 , \quad e'_4 = -3e_1 - e_2 + 5e_4 , \quad e'_5 = -e_1 - 7e_2 + 5e_5 .$$

Oznaczmy ponadto $e'_1 = e_1$, $e'_2 = e_2$. Układ wektorów $(e'_1, \dots, e'_5)$ tworzy bazę przestrzeni V: jeśli $y'^i e'_i = 0$, to działając na tę równość odwzorowaniem A dostajemy $y'^1 Ae_1 + y'^2 Ae_2 = 0$, więc $y'^1 = y'^2 = 0$ (bo wektory Ae_1, Ae_2 są bazą obrazu $\operatorname{Im} A$). Stąd również $y'^3 = y'^4 = y'^5 = 0$, bo wektory e'_3, e'_4, e'_5 są liniowo niezależne. Nową bazę przestrzeni W otrzymamy uzupełniając w dowolny sposób do bazy wektory f'_1, f'_2, np. $f'_3 = f_3$, $f'_4 = f_4$ (ten sposób daje rzeczywiście bazę, gdyż wyznacznik utworzony z pierwszych dwóch wierszy kolumn współrzędnych wektorów f'_1, f'_2 w bazie (f_j) jest różny od zera). W nowych bazach działanie odwzorowania A ma prostą formę:

$$Ae'_1 = f'_1 , \quad Ae'_2 = f'_2 , \quad Ae'_3 = Ae'_4 = Ae'_5 = 0 , \quad \text{więc}$$
$$Ax'^i e'_i = x'^1 f'_1 + x'^2 f'_2 .$$

Macierz odwzorowania A w nowych bazach ma postać

$$\mathbf{A}' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Ze związków wektorów bazowych odczytujemy macierze przejścia

$$\begin{pmatrix} e'_1 & \dots & e'_5 \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_5 \end{pmatrix} \boldsymbol{\beta} , \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & 9 & -3 & -1 \\ 0 & 1 & -2 & -1 & -7 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} ,$$

$$\begin{pmatrix} f'_1 & \dots & f'_4 \end{pmatrix} = \begin{pmatrix} f_1 & \dots & f_4 \end{pmatrix} \boldsymbol{\gamma} , \qquad \boldsymbol{\gamma} = \begin{pmatrix} 1 & 2 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 7 & 1 & 0 \\ -8 & 9 & 0 & 1 \end{pmatrix} .$$

Nowa macierz odwzorowania łączy się ze starą transformacją $\mathbf{A}' = \boldsymbol{\gamma}^{-1} \mathbf{A} \boldsymbol{\beta}$.

(v) Rząd złożenia odwzorowań liniowych
Niech $A : V \mapsto W$ i $B : W \mapsto Y$ będą odwzorowaniami liniowymi skończenie wymiarowych przestrzeni, więc $BA : V \mapsto Y$ jest również takim odwzorowaniem. Zachodzą oczywiste inkluzje

$$\operatorname{Im} BA \subseteq \operatorname{Im} B\,, \qquad \operatorname{Ker} BA \supseteq \operatorname{Ker} A\,.$$

Stąd

$$\operatorname{rk} BA \leq \operatorname{rk} B\,, \qquad \operatorname{rk} BA \leq \operatorname{rk} A\,,$$

gdzie w drugim przypadku wykorzystaliśmy równości

$$\operatorname{rk} BA = \dim V - \dim \operatorname{Ker} BA\,, \quad \operatorname{rk} A = \dim V - \dim \operatorname{Ker} A\,.$$

Powyższe nierówności (szczególnie zapisane dla macierzy) noszą nazwę twierdzenia Sylwestra.

(vi) Niezależność rzędu od baz
Rząd odwzorowania liniowego, który jako wymiar jego obrazu nie zależy od wyboru baz, można wyliczyć jako rząd jego macierzy w dowolnych bazach. Stąd dla każdej macierzy $\mathbf{A} \in \mathbb{K}^{m\times n}$ i nieosobliwych macierzy $\boldsymbol{\beta} \in \mathbb{K}^{n\times n}$, $\boldsymbol{\gamma} \in \mathbb{K}^{m\times m}$ mamy

$$\operatorname{rk}(\boldsymbol{\gamma}^{-1}\mathbf{A}\boldsymbol{\beta}) = \operatorname{rk} \mathbf{A}\,.$$

(vii) Ślad operatora
Niech V będzie przestrzenią skończenie wymiarową. Nie istnieją działające w niej operatory liniowe $\hat{D}$ i $\hat{X}$ takie, które spełniałyby

$$\hat{D}\hat{X} - \hat{X}\hat{D} = \mathrm{id}_V\,,$$

co można pokazać licząc ślad obu stron równania. (Porównaj z przykładem (ii) oraz z przykładem (ii) w punkcie 6.)

(viii) Wielomianowe funkcje operatorów
Rozważmy przestrzeń macierzy 2×2 o elementach w ciele $\mathbb{K} = \mathbb{Z}/\mathrm{mod}\,2$ jako przestrzeń operatorów w przestrzeni $\mathbb{K}^2$. Wielomiany $P_1(x) = x$ oraz $P_2(x) = x^2$ zadają tę samą funkcję wielomianową w ciele, ale różne funkcje operatorowe. Na przykład: dla macierzy $\mathbf{A} = \left(\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right)$ jest $\mathbf{A}^2 = \left(\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right)$.

18 Odwzorowania antyliniowe

Kończymy ten paragraf krótkim omówieniem pewnej klasy odwzorowań spotykanych dla przestrzeni zespolonych. Niech V i W będą przestrzeniami nad ciałem liczb zespolonych. Mówimy, że odwzorowanie $A : V \mapsto W$ jest ***odwzorowaniem antyliniowym***, gdy dla każdych wektorów $x, y \in V$ i liczb $\alpha, \beta \in \mathbb{C}$ zachodzi

$$A(\alpha x + \beta y) = \overline{\alpha} Ax + \overline{\beta} Ay \,.$$

Analogicznie jak dla odwzorowań liniowych wprowadzamy pojęcia: ***izomorfizm antyliniowy*** – bijektywne odwzorowanie antyliniowe, ***operator antyliniowy*** – antyliniowe odwzorowanie przestrzeni w samą siebie, ***automorfizm antyliniowy*** – bijektywny operator antyliniowy.

Dla odwzorowań antyliniowych przenoszą się definicje jądra i obrazu odwzorowania i pozostają w mocy, po oczywistych modyfikacjach sformułowań, twierdzenia 1, 2, 3, 5, 6, 7. W twierdzeniu 4 wynik modyfikuje się:
– jeśli jedno z odwzorowań A i B jest antyliniowe, a drugie liniowe, to BA jest antyliniowe;
– jeśli oba są antyliniowe, to BA jest liniowe;
– jeśli A jest izomorfizmem antyliniowym, to A^{-1} jest też izomorfizmem antyliniowym;
– jeśli A i B są izomorfizmami, to BA jest też izomorfizmem, liniowym lub antyliniowym – zgodnie z wyżej podaną regułą składania.

Definicja macierzy odwzorowania liniowego w zadanych bazach przenosi się bez zmian na przypadek odwzorowania antyliniowego. Twierdzenie 10 ulega jednak modyfikacji:
– jeśli A jest odwzorowaniem antyliniowym, to

$$y = Ax \iff \mathbf{y} = \mathbf{A}\overline{\mathbf{x}} \,;$$

– jeśli B jest liniowe, a A jest liniowe lub antyliniowe, to

$$C = BA \iff \mathbf{C} = \mathbf{B}\mathbf{A} \,;$$

– jeśli B jest antyliniowe, a A jest liniowe lub antyliniowe, to

$$C = BA \iff \mathbf{C} = \mathbf{B}\overline{\mathbf{A}} \,.$$

Reguła transformacji macierzy odwzorowania antyliniowego ma postać

$$\mathbf{A}' = \boldsymbol{\gamma}^{-1} \mathbf{A} \, \overline{\boldsymbol{\beta}} \,.$$

W szczególności, dla operatora antyliniowego dostajemy

$$\mathbf{A}' = \boldsymbol{\beta}^{-1} \mathbf{A} \, \overline{\boldsymbol{\beta}} \,,$$

więc pojęcia wyznacznika i śladu nie mają w tym przypadku inwariantnego, niezależnego od bazy znaczenia.

Operator antyliniowy K nazywamy ***operatorem sprzężenia*** (lub krótko ***sprzężeniem***), gdy spełnia on warunek

$$K^2 = \mathrm{id} \,.$$

Sprzężenia są szczególnie prostymi operatorami antyliniowymi, ich postać zanalizujemy poniżej w przykładach. Mając dane dowolne sprzężenie K w przestrzeni V, z dowolnego odwzorowania antyliniowego $A : V \mapsto W$ można utworzyć odwzorowanie liniowe $B = AK$, działające między tymi samymi przestrzeniami. Wtedy $A = BK$, więc dostajemy rozkład dowolnego odwzorowania antyliniowego na sprzężenie i odwzorowanie liniowe. Należy jednak zwrócić uwagę na to, że rozkład ten nie jest jednoznaczny – każdy wybór sprzężenia daje inne odwzorowanie liniowe w tym rozkładzie.

19 Przykłady

(i) Operator sprzężenia
Niech K będzie operatorem sprzężenia na przestrzeni V i oznaczmy przez R zbiór wektorów spełniających równanie $Ky = y$. Wprost z definicji zbiór ten jest rzeczywistą (ale nie zespoloną!) przestrzenią wektorową. Każdy wektor $x \in V$ ma jednoznaczne przedstawienie $x = x_+ + ix_-$, $x_\pm \in R$. Aby to wykazać potraktujmy tę równość jako równanie na $x_\pm$, zadziałajmy na nią sprzężeniem: $Kx = x_+ - ix_-$, i wyliczmy z otrzymanego układu:

$$x_+ = \tfrac{1}{2}(K + \mathrm{id})x \,, \quad x_- = \tfrac{i}{2}(K - \mathrm{id})x \,.$$

Łatwo sprawdzić używając definicyjnego warunku dla sprzężenia, że dane tymi wzorami wektory istotnie należą do R, oraz spełniają $x = x_+ + ix_-$, co kończy dowód.

Odwrotnie, niech $R \subseteq V$ będzie rzeczywistą przestrzenią wektorową taką, że każdy wektor $x \in V$ ma jednoznaczne przedstawienie $x = x_+ + ix_-$, $x_\pm \in R$. Wtedy operator zadany równaniem $Kx = x_+ - ix_-$ jest sprzężeniem. Spełnienie warunku $K^2 = \mathrm{id}$ i liniowości względem mnożenia przez liczby rzeczywiste jest oczywiste, a antyliniowość dla mnożenia przez jednostkę urojoną wynika z ciągu równości

$$K(ix) = K(-x_- + ix_+) = -x_- - ix_+ = -i(x_+ - ix_-) = -iKx \,.$$

Zanotujmy związek przestrzeni R spełniającej powyższe warunki z bazami przestrzeni V.

(i) Każda baza E przestrzeni R jako przestrzeni rzeczywistej jest bazą przestrzeni V, jako przestrzeni zespolonej.
(ii) Rzeczywista liniowa powłoka każdej bazy E przestrzeni V jest rzeczywistą przestrzenią R w omawianej klasie.

(ii) Sprzężenie funkcji zespolonych
Niech $\mathcal{F}(X)$ będzie przestrzenią wektorową zespolonych funkcji na dowolnym zbiorze X. Odwzorowanie

$$\mathcal{F}(X) \ni f \mapsto \overline{f} \in \mathcal{F}(X)\,, \quad \overline{f}(x) := \overline{f(x)}$$

jest operatorem sprzężenia. Przestrzenią rzeczywistą R wektorów spełniających $Kf = f$ jest w tym przypadku zbiór funkcji rzeczywistych. W szczególności, operacja sprzężenia zespolonego macierzy $m \times n$ jest operatorem sprzężenia, z przestrzenią R złożoną z macierzy rzeczywistych.

(iii) Sprzężenie hermitowskie
Dla macierzy $m \times n$ operacja sprzężenia hermitowskiego jest odwzorowaniem antyliniowym, rozkładającym się na operator sprzężenia zespolonego i transpozycję, która jest odwzorowaniem liniowym.

Dla $m = n$ operacja sprzężenia hermitowskiego jest operatorem sprzężenia, z przestrzenią R złożoną z macierzy hermitowskich.

§11 Grupy operatorowe. Orientacja

1 Grupy operatorowe i ich izomorfizm z grupami macierzowymi

Operatory liniowe na przestrzeni V tworzą półgrupę (z jedynką) z działaniem składania – składanie operatorów jest łącznym działaniem w ich zbiorze, a operator identycznościowy jest elementem neutralnym składania. Zgodnie z ogólnym wynikiem z teorii grup zbiór operatorów mających odwrotne jest więc grupą, oznaczamy ją $GL(V)$. Ograniczamy się w tym paragrafie do przestrzeni skończenie wymiarowych.

Twierdzenie 1.
Niech V będzie przestrzenią o skończonym wymiarze.
(i) Operator liniowy w przestrzeni V jest automorfizmem wtedy, i tylko wtedy, gdy jego wyznacznik jest różny od zera.
(ii) Grupa automorfizmów liniowych $GL(V)$ przestrzeni wektorowej V jest izomorficzna z ogólną grupą liniową $GL(\dim V, \mathbb{K})$.

Dowód. (i) Operator A jest bijektywny wtedy, i tylko wtedy, gdy dla pewnego operatora B jest $AB = BA = \mathrm{id}$, co na podstawie drugiej części twierdzenia 10, §10, jest równoważne istnieniu macierzy $\mathbf{B}$ takiej, że $\mathbf{AB} = \mathbf{BA} = \mathbf{1}$. Wiemy, że dla macierzy ostatni warunek jest równoważny warunkowi $\det \mathbf{A} \neq 0$, co jest tym samym co $\det A \neq 0$.

(ii) Wszystkie rozważane tu odwzorowania liniowe są operatorami w jednej przestrzeni V, więc określone w punkcie 11, §10, odwzorowanie $A \mapsto T(A) = \mathbf{A}$ jest zadane dla nich jednym przepisem (związanym z wyborem jednej bazy w V). Na podstawie punktu (i) widzimy, że T zacieśnione do automorfizmów jest bijekcją $GL(V)$ na $GL(\dim V; \mathbb{K})$, a z twierdzenia 10, §10, mamy tożsamość $T(AB) = T(A)T(B)$. Odwzorowanie $T : GL(V) \mapsto GL(\dim V, \mathbb{K})$ jest więc izomorfizmem grupowym. □

Podobnie jak dla macierzy, zbiór $SL(V)$ automorfizmów o wyznaczniku równym jeden jest podgrupą wszystkich automorfizmów liniowych, a jego obraz w izomorfizmie grup $GL(V)$ i $GL(\dim V, \mathbb{K})$ jest grupą $SL(\dim V, \mathbb{K})$. Dla przestrzeni rzeczywistej zbiory automorfizmów: o wyznaczniku dodatnim, $GL(V)_+$, oraz o wyznaczniku ujemnym, $GL(V)_-$, przechodzą pod działaniem izomorfizmu w $GL(\dim V, \mathbb{R})_+$ i $GL(\dim V, \mathbb{R})_-$ odpowiednio, przy czym $GL(V)_+$ jest grupą.

2 Związek $GL(V)$ ze zbiorem baz uporządkowanych przestrzeni V

Wybierzmy w przestrzeni V bazę $(e_1, \ldots, e_n)$. Dla dowolnego automorfizmu $\beta \in GL(V)$ zbiór wektorów $(\beta e_1, \ldots, \beta e_n)$ również tworzy bazę, której wektory mają w zadanej bazie rozkłady $\beta e_i = e_j \beta^j{}_i$, zgodnie z ogólną definicją macierzy operatora liniowego w zadanej bazie. Ponieważ prawa strona tego związku pokrywa się z przepisem na dowolną zmianę bazy, możemy wyciągnąć następujący wniosek.

Twierdzenie 2. *Każda baza $(e_1, \ldots, e_n)$ przestrzeni V zadaje wzajemnie jednoznaczny związek między automorfizmami z $GL(V)$ i zbiorem baz uporządkowanych przestrzeni V, dany przyporządkowaniem $\beta \mapsto (\beta e_1, \ldots, \beta e_n)$.*

Korzystając z tego rezultatu, własności związków zachodzących między bazami można przełożyć na własności grupy $GL(V)$, a dzięki izomorfizmowi omówionemu w poprzednim punkcie – na własności grupy $GL(\dim V, \mathbb{K})$.

3 Ciągłość i spójność

Od tego miejsca ograniczamy się w tym paragrafie do przestrzeni rzeczywistych lub zespolonych. Zbadamy teraz własności ciągłości i spójności w grupach automorfizmów przestrzeni wektorowych takich typów. Korzystając następnie z wyniku poprzedniego punktu przełożymy je na własności zbioru baz uporządkowanych. Najpierw jednak musimy wyjaśnić, jak należy rozumieć ciągłość i spójność w przestrzeniach wektorowych.

Niech V będzie rzeczywistą lub zespoloną przestrzenią wektorową i niech będzie dane odwzorowanie $\langle a, b\rangle \ni t \mapsto x(t) \in V$, gdzie $\langle a, b\rangle$ jest dowolnym przedziałem domkniętym w zbiorze liczb rzeczywistych $\mathbb{R}$. Wybieramy dowolną bazę $(e_1, \ldots, e_n)$ przestrzeni V i rozkładamy $x(t) = x^i(t)e_i$. Mówimy, że dane odwzorowanie jest ***krzywą ciągłą w*** V, gdy ciągłe są wszystkie funkcje $x^i(t)$ – w przypadku przestrzeni zespolonej oznacza to ciągłość wszystkich funkcji $\mathfrak{Re}\, x^i(t)$ i $\mathfrak{Im}\, x^i(t)$. Zmienną rzeczywistą t nazywamy parametrem krzywej. Z postaci transformacji współrzędnych wektora przy zmianie bazy widać natychmiast, że własność ciągłości współrzędnych nie zależy od wyboru bazy, więc definicja jest poprawna i charakteryzuje samo odwzorowanie.

Jeśli $\langle a', b'\rangle \subseteq \mathbb{R}$ i funkcja $\langle a', b'\rangle \ni s \mapsto f(s) \in \langle a, b\rangle$ jest ciągłą bijekcją, to przy powyższych oznaczeniach odwzorowanie $\langle a', b'\rangle \ni s \mapsto y(s) := x(f(s)) \in V$ również jest krzywą ciągłą. Korzystając z podstawowych wyników analizy matematycznej stwierdzamy, że w tym przypadku odwzorowanie odwrotne f^{-1} jest również ciągłą funkcją bijektywną; podobnie złożenia takich bijekcji są ciągłe. Na tej podstawie łatwo przekonać się, że określona powyżej relacja pomiędzy krzywymi $x(.)$ i $y(.)$ jest relacją równoważności; krzywe w jednej klasie równoważności różnią się jedynie wyborem parametru. Używając określenia *krzywa ciągła* ma się niekiedy na myśli całą klasę, a elementy klasy nazywa się wtedy dokładniej *sparametryzowanymi krzywymi ciągłymi*. Z kontekstu na ogół jest jasne, o którym pojęciu mowa.

Jeśli dwie krzywe ciągłe mają wspólny punkt będący jednym z końców każdej z nich, to można z tych krzywych utworzyć jedną krzywą ciągłą przez odpowiedni dobór parametru – pozostawiamy wykazanie tego faktu jako ćwiczenie.

Elementy macierzy odwzorowania liniowego są, jak wiemy, współrzędnymi (w pewnej bazie) odwzorowania jako wektora w przestrzeni wektorowej odwzorowań liniowych między zadanymi przestrzeniami. Powyższa definicja ciągłości stosuje się więc w szczególności do odwzorowań liniowych: $A(t)$ jest ciągłą krzywą, gdy wszystkie funkcje (lub ich części rzeczywiste i urojone) ${A^i}_j(t)$ są ciągłe. Jest oczywiste, że jeśli $A(t) \in \mathcal{L}(V, W)$ oraz $x(t) \in V$ są krzywymi ciągłymi o wspólnym zakresie parametru, to $A(t)x(t) \in W$ jest też krzywą ciągłą o tym samym zakresie parametru.

Niech będzie dany podzbiór X przestrzeni wektorowej V. Mówimy, że X jest ***spójny łukowo***, gdy dla każdych dwóch wektorów $x, y \in X$ istnieje krzywa ciągła $\langle a, b\rangle \ni t \mapsto x(t) \in X$ taka, że $x(a) = x$ i $x(b) = y$. Mówiąc bardziej obrazowo: każde dwa wektory zbioru X dają się połączyć krzywą ciągłą leżącą wewnątrz tego zbioru.

4 Własności spójności grup $GL(V)$

Zbadamy, czy grupy automorfizmów rzeczywistej oraz zespolonej przestrzeni wektorowej są zbiorami spójnymi łukowo. Zgodnie z definicjami spójności łukowej i ciągłości nasze zadanie sprowadza się do zbadania spójności łukowej grup $GL(n, \mathbb{C})$ oraz $GL(n, \mathbb{R})$. Zaczniemy od udowodnienia pomocniczego wyniku.

Lemat 3. *Niech* $\mathbf{A} \in GL(n, \mathbb{K})$*, gdzie* $\mathbb{K} = \mathbb{R}$ *lub* $\mathbb{C}$*. Istnieje ciągła krzywa* $\langle a, b\rangle \ni t \mapsto \mathbf{A}(t) \in GL(n, \mathbb{K})$ *taka, że* $\det \mathbf{A}(t) = \det \mathbf{A}$*,* $\mathbf{A}(a) = \mathbf{A}$ *oraz* $\mathbf{A}(b)$ *jest macierzą diagonalną.*

Dowód. Dowód przeprowadzimy w trzech rekurencyjnych krokach.
(i) Jeśli $A_{nn} \neq 0$, to przechodzimy do kroku (ii). Jeśli $A_{nn} = 0$, to z nieosobliwości macierzy istnieje taka liczba $k \in \{1, \dots, n-1\}$, że $A_{nk} \neq 0$. Dla $t \in \langle 0, 1\rangle$ tworzymy macierze $\mathbf{A}(t)$ zadając ich kolumny przez:

$$\mathbf{A}_{*i}(t) = \mathbf{A}_{*i} \quad \text{dla} \quad i = 1, \dots, n-1 \quad \text{oraz} \quad \mathbf{A}_{*n}(t) = \mathbf{A}_{*n} + t\mathbf{A}_{*k}\,.$$

Mamy $\mathbf{A}(0) = \mathbf{A}$, $\det \mathbf{A}(t) = \det \mathbf{A}$ oraz $A_{nn}(1) \neq 0$.
(ii) Przedłużamy w sposób ciągły krzywą na $t \in \langle 1, 2\rangle$ odejmując wielokrotności ostatniej kolumny macierzy $\mathbf{A}(1)$ od pozostałych kolumn:

$$\mathbf{A}_{*i}(t) = \mathbf{A}_{*i}(1) - (t-1)\frac{A_{ni}(1)}{A_{nn}(1)}\mathbf{A}_{*n}(1) \quad \text{dla} \quad i \in \{1, \dots, n-1\}\,.$$

Mamy $\det \mathbf{A}(t) = \det \mathbf{A}(1)$ oraz $A_{ni}(2) = 0$ dla $i \in \{1, \dots, n-1\}$.
(iii) Przedłużamy w sposób ciągły krzywą na $t \in \langle 2, 3\rangle$ odejmując wielokrotności ostatniego wiersza macierzy $\mathbf{A}(2)$ od pozostałych wierszy:

$$\mathbf{A}_{i*}(t) = \mathbf{A}_{i*}(2) - (t-2)\frac{A_{in}(2)}{A_{nn}(2)}\mathbf{A}_{n*}(2) \quad \text{dla} \quad i \in \{1, \dots, n-1\}\,.$$

Mamy $\det \mathbf{A}(t) = \det \mathbf{A}(2)$, a $\mathbf{A}(3)$ jest macierzą, której wszystkie wyrazy w ostatniej kolumnie i ostatnim wierszu znikają, z wyjątkiem wyrazu na diagonali.
Powtarzamy teraz te same kroki dla lewego górnego bloku $(n-1)\times(n-1)$ macierzy $\mathbf{A}(3)$ i rekurencyjnie aż do otrzymania macierzy diagonalnej. Otrzymaliśmy ciągłą krzywą o żądanych własnościach. □

Stosując ten lemat łatwo udowodnimy teraz następujące wyniki.

Twierdzenie 4.
(i) Dla zespolonej przestrzeni wektorowej V *grupa* $GL(V)$ *jest spójna łukowo.*
(ii) Dla rzeczywistej przestrzeni wektorowej grupa $GL(V)$ *nie jest spójna łukowo. Podgrupa* $GL(V)_+$ *oraz podzbiór* $GL(V)_-$ *są zbiorami spójnymi łukowo.*

Dowód. Wystarczy udowodnić odpowiednie własności dla grup $GL(n,\mathbb{C})$ oraz $GL(n,\mathbb{R})$.

(i) Wystarczy udowodnić, że każdą nieosobliwą macierz diagonalną można połączyć w sposób ciągły wewnątrz $GL(n,\mathbb{C})$ z macierzą jednostkową – wtedy w świetle lematu każde dwie macierze z tej grupy można połączyć w sposób ciągły poprzez macierz jednostkową. Niech elementy diagonalne macierzy diagonalnej $\mathbf{A}$ będą równe $A_{ii} = r_i \exp(i\varphi_i)$. Macierze diagonalne $\mathbf{A}(t)$ o elementach diagonalnych

$$A_{ii}(t) = (r_i(1-t)+t)\exp(i(1-t)\varphi_i)$$

tworzą szukaną krzywą ciągłą, taką że $\mathbf{A}(0) = \mathbf{A}$ i $\mathbf{A}(1) = \mathbf{1}$.

(ii) Wyznacznik macierzy jest ciągłą funkcją jej elementów, więc dla dowolnej ciągłej krzywej $\mathbf{A}(t)$ funkcja $\det \mathbf{A}(t)$ jest ciągła. Ale dla wszystkich macierzy z $GL(n,\mathbb{R})$ wyznacznik jest różny od zera, więc macierzy o dodatnim wyznaczniku nie można w sposób ciągły połączyć z macierzą o wyznaczniku ujemnym. Stąd $GL(n,\mathbb{R})$ nie jest spójna łukowo.

Załóżmy teraz, że $\mathbf{A} \in GL(n,\mathbb{R})_+$. Korzystając z lematu wystarczy przyjąć, że $\mathbf{A}$ jest diagonalna. Zapiszmy wyrazy diagonalne jako $A_{ii} = \epsilon_i r_i$, gdzie $r_i > 0$, $\epsilon_i = \pm 1$. Za pomocą ciągłej krzywej macierzy diagonalnych $\langle 0,1\rangle \ni t \mapsto \mathbf{A}(t)$ o elementach $A_{ii} = (r_i(1-t)+t)\epsilon_i$ sprowadzamy macierz $\mathbf{A}$ do macierzy $\mathbf{A}(1)$, dla której $A_{ii}(1) = \epsilon_i$. Wyznacznik jest dodatni, więc liczba ujemnych liczb wśród ϵ_i jest parzysta, niech będą to liczby $\epsilon_{i_1} = \epsilon_{j_1} = \ldots = \epsilon_{i_k} = \epsilon_{j_k} = -1$. Przedłużamy krzywą w sposób ciągły na $t \in \langle 1,2\rangle$ kładąc dla $l \in \{1,\ldots,k\}$:

$$\begin{aligned} A_{i_l i_l}(t) &= \cos(2-t)\pi\,, \qquad & A_{i_l j_l}(t) &= -\sin(2-t)\pi\,,\\ A_{j_l i_l}(t) &= \sin(2-t)\pi\,, & A_{j_l j_l}(t) &= \cos(2-t)\pi \end{aligned}$$

oraz $A_{ii}(t) = 1$ dla pozostałych elementów diagonalnych, a $A_{ij}(t) = 0$ dla pozostałych elementów pozadiagonalnych. Krzywa $A(t)$ jest ciągła oraz $\mathbf{A}(2) = \mathbf{1}$. Pokażemy, że $\det A(t) = 1$. W tym celu przesuńmy w macierzy $A(t)$ wiersze o numerach $i_1, j_1, \ldots, i_k, j_k$ na pierwszych $2k$ pozycji, bez zmiany kolejności pozostałych wierszy. Następnie dokonajmy takiej samej permutacji kolumn. Otrzymana macierz $A'(t)$, która ma wyznacznik równy wyznacznikowi macierzy $A(t)$ (własność (iv), tw. 2, §7), ma postać

$$A'(t) = \begin{pmatrix} R(t) & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & R(t) & 0 \\ 0 & 0 & 0 & \mathbf{1}_{n-2k} \end{pmatrix}, \quad R(t) = \begin{pmatrix} \cos(2-t)\pi & -\sin(2-t)\pi \\ \sin(2-t)\pi & \cos(2-t)\pi \end{pmatrix}.$$

Stąd $\det A'(t) = \big(\det R(t)\big)^k = 1$, co kończy dowód spójności $GL(n,\mathbb{R})_+$.

Jeśli $\mathbf{A} \in GL(n, \mathbb{R})_-$, to można ją przedstawić w formie $\mathbf{A} = \mathbf{R}_1\mathbf{B}$, gdzie $\mathbf{R}_1 \in GL(n, \mathbb{R})_-$ jest macierzą diagonalną o wyrazach na diagonali równych $(-1, 1, \dots, 1)$. Wtedy $\mathbf{B} \in GL(n, \mathbb{R})_+$, więc istnieje w $GL(n, \mathbb{R})_+$ krzywa ciągła $\mathbf{B}(t)$ taka, że $\mathbf{B}(0) = \mathbf{B}$ i $\mathbf{B}(1) = \mathbf{1}$. Stąd każdą macierz z $GL(n, \mathbb{R})_-$ można połączyć krzywą ciągłą z macierzą $\mathbf{R}_1$, więc zbiór ten jest spójny. □

5 Orientacja

Mówimy, że dwie bazy uporządkowane $(e_1, \dots, e_n)$ i $(e'_1, \dots, e'_n)$ przestrzeni V są ***jednakowo zorientowane***, gdy dla wszystkich $t \in \langle a, b\rangle$ istnieją bazy $\big(e_1(t), \dots, e_n(t)\big)$ takie, że $e_i(a) = e_i$, $e_i(b) = e'_i$, oraz wszystkie krzywe $e_i(t)$ są ciągłe. W skrócie: jedną bazę można w sposób ciągły przenieść na drugą (zachowując stale liniową niezależność). Jest widoczne, że własność jednakowego zorientowania pary baz jest relacją równoważnościową w zbiorze baz uporządkowanych, dzieli więc zbiór wszystkich takich baz na klasy równoważności baz jednakowo zorientowanych.

Warunek jednakowego zorientowania przekłada się na ciągłość automorfizmów łączących bazy:

Twierdzenie 5. *Niech będą dane: baza $(e_1, \dots, e_n)$ przestrzeni V oraz dla każdego $t \in \langle a, b\rangle \subseteq \mathbb{R}$ automorfizm liniowy $\beta(t)$. Wtedy ciągłość wszystkich krzywych $\beta(t)e_i$, $i = 1, \dots, n$, jest równoważna ciągłości krzywej $\beta(t)$.*

Dowód. Jeśli $\beta(t)$ jest ciągła, to w oczywisty sposób krzywe wektorów bazowych są ciągłe. Odwrotnie, ciągłość krzywych bazowych oznacza, że ich współrzędne w bazie $(e_1, \dots, e_n)$ są ciągłe. Współrzędne te są elementami macierzy operatorów $\beta(t)$ w tej bazie, skąd teza. □

Uwzględniając wyniki poprzedniego punktu otrzymujemy główne twierdzenie o orientacji.

Twierdzenie 6.
(i) Każde dwie bazy przestrzeni zespolonej są jednakowo zorientowane.
(ii) W przestrzeni rzeczywistej dwie bazy są jednakowo zorientowane wtedy, i tylko wtedy, gdy wyznacznik łączącego je automorfizmu jest dodatni. Relacja jednakowej orientacji dzieli zbiór baz uporządkowanych na dwie klasy równoważności.

Dowód. Niech będą dane dwie bazy $(e_1, \dots, e_n)$ i $(\beta e_1, \dots, \beta e_n)$. Bazy te są jednakowo zorientowane wtedy, i tylko wtedy, gdy automorfizm β można w sposób ciągły połączyć z identycznością. Dla przestrzeni zespolonej automorfizmy stanowią grupę spójną łukowo, co dowodzi punktu (i). W przestrzeni rzeczywistej $GL(V)_+$ jest spójna, a $GL(V)$ nie jest spójna, więc β można w sposób ciągły połączyć z identycznością wtedy, i tylko wtedy, gdy $\det\beta > 0$, co dowodzi pierwszego stwierdzenia punktu (ii). Niech wyznacznik automorfizmu łączącego bazy $(e_1, \dots, e_n)$ i $(e'_1, \dots, e'_n)$ będzie ujemny, przez co bazy te należą do dwu różnych klas równoważności. Dla każdej innej bazy wyznacznik automorfizmu łączącego ją z jedną, i tylko jedną, z tych baz jest dodatni, co dowodzi drugiego stwierdzenia w punkcie (ii). □

Każdą z dwu klas równoważności baz przestrzeni rzeczywistej nazywamy jej ***orientacją*** – jedną z nich, dowolnie wybraną, nazywamy orientacją dodatnią, a drugą – ujemną.

§12 Sumy proste przestrzeni i operatorów. Operatory rzutowe

1 Suma i przecięcie podprzestrzeni

Niech V będzie przestrzenią wektorową nad dowolnym ciałem $\mathbb{K}$, a V_1 i V_2 jej podprzestrzeniami. Z podprzestrzeni tych konstruujemy za pomocą działań na zbiorach dwa dalsze podzbiory przestrzeni V: $V_1 \cap V_2$ i $V_1 \cup V_2$. Podstawowe kryterium podprzestrzeni, tw. 2, §8, pokazuje natychmiast, że pierwszy z tych podzbiorów jest podprzestrzenią przestrzeni V, ale drugi nią nie jest. Najmniejsza podprzestrzeń zawierająca dany zbiór wektorów to jego powłoka liniowa. Łatwo przekonać się, że dowolna kombinacja liniowa wektorów ze zbioru $V_1 \cup V_2$ ma postać $x_1 + x_2$, gdzie $x_1 \in V_1$ i $x_2 \in V_2$. Odwrotnie, zbiór wektorów tej postaci zawiera się w powłoce liniowej zbioru $V_1 \cup V_2$. Stąd

$$V_1 + V_2 := \{x_1 + x_2 \mid x_1 \in V_1\,,\ x_2 \in V_2\} = L(V_1 \cup V_2)\,.$$

Ogólniej, jeśli $V_1, \dots, V_s$ są podprzestrzeniami przestrzeni V, to podprzestrzeń $\bigcap_{i=1}^{s} V_i$ nazywamy ***przecięciem podprzestrzeni***, a podprzestrzeń

$$\sum_{i=1}^{s} V_i := \Big\{ \sum_{i=1}^{s} x_i \;\Big|\; x_i \in V_i \Big\} = L\Big(\bigcup_{i=1}^{s} V_i \Big)$$

– ich ***sumą (wektorową)***. Będziemy mówić, że podprzestrzenie $V_1, \dots, V_s$ ***generują*** przestrzeń V, gdy $\sum_{i=1}^{s} V_i = V$.

Podprzestrzenie $V_1, \dots, V_s$ będziemy nazywać ***liniowo niezależnymi***, gdy spełniony będzie następujący warunek:

$$\sum_{i=1}^{s} x_i = 0\,,\ x_i \in V_i \implies x_i = 0\,, i = 1, \dots, s\,.$$

W przypadku dwóch podprzestrzeni V_1 i V_2 warunek ich liniowej niezależności sprowadza się do żądania $V_1 \cap V_2 = \{0\}$. Jednak dla większej liczby przestrzeni zastosowanie tego warunku do wszystkich par nie jest wystarczające.

Twierdzenie 1. *Liniowa niezależność podprzestrzeni $V_1, \ldots, V_s \subseteq V$ jest równoważna koniunkcji warunków:*

$$V_i \cap \Big(\sum_{j \neq i} V_j \Big) = \{0\}\,, \ i = 1, \ldots, s$$

Dowód. Przy ustalonym $i = 1, \ldots, s$ przecięcie po lewej stronie warunku składa się z wektorów $x_i \in V_i$, które dają się przedstawić w postaci $x_i = \sum_{j \neq i} y_j$, $y_j \in V_j$. Ale wtedy przy oznaczeniu $y_i = -x_i$ mamy $\sum_{k=1}^{s} y_k = 0$, i jeśli podprzestrzenie są liniowo niezależne, to stąd $x_i = 0$. Odwrotnie, niech y_k sumują się do zera, więc dla każdego i jest $-y_i = \sum_{j \neq i} y_j$. Jeśli warunki z tezy są spełnione, to stąd $y_i = 0$, więc podprzestrzeni są liniowo niezależne. □

Mówimy, że podprzestrzenie dowolnej rodziny podprzestrzeni (niekoniecznie skończonej) są liniowo niezależne, gdy podprzestrzenie każdego skończonego, różnowartościowego ciągu z tej rodziny są liniowo niezależne.

W szczególnym przypadku, gdy $e_1, \ldots, e_s$ jest ciągiem niezerowych wektorów, a $V_i = L(e_i)$, to warunek liniowej niezależności podprzestrzeni V_i jest identyczny z warunkiem liniowej niezależności wektorów e_i.

2 Suma prosta podprzestrzeni

Mówimy, że przestrzeń V jest ***sumą prostą podprzestrzeni*** $V_1, \ldots, V_s$ i piszemy $V = \bigoplus_{i=1}^{s} V_i$, gdy podprzestrzenie $V_1, \ldots, V_s$ są liniowo niezależne i generują V.

Twierdzenie 2. *Niech $V_1, \ldots, V_s$ będą podprzestrzeniami przestrzeni V. Następujące warunki są równoważne:*

(i) V jest sumą prostą podprzestrzeni V_i;

(ii) każdy wektor przestrzeni V ma jednoznaczny rozkład na sumę wektorów podprzestrzeni V_i;

(iii) dla każdych baz E_i podprzestrzeni V_i, $i = 1, \ldots, s$, zachodzi $E_i \cap E_j = \varnothing$ dla $i \neq j$ i rodzina wektorów $E_1 \cup \ldots \cup E_s$ jest bazą przestrzeni V;

(iv) istnieją bazy E_i podprzestrzeni V_i, $i = 1, \ldots, s$ takie, że $E_i \cap E_j = \varnothing$ dla $i \neq j$ i rodzina wektorów $E_1 \cup \ldots \cup E_s$ jest bazą przestrzeni V.

Jeśli wymiar przestrzeni V jest skończony, to również następujący warunek jest równoważny wcześniejszym:

(v) podprzestrzenie V_i generują V i $\sum_{i=1}^{s} \dim V_i = \dim V$.

Dowód. Wykażemy ciąg implikacji (i)⇒(ii)⇒(iii)⇒(iv)⇒(i), a przy dodatkowym założeniu: (v)⇔(iv). Stąd wynikają wszystkie równoważności.

(i)⇒(ii) Podprzestrzenie generują V, więc każdy wektor rozkłada się na sumę wektorów podprzestrzeni. Jeśli $x = \sum_i x_i = \sum_i x'_i$, to $\sum_i (x_i - x'_i) = 0$, więc z liniowej niezależności podprzestrzeni $x_i = x'_i$ – rozkłady są jednoznaczne.

(ii)⇒(iii) Gdyby istniał wektor $e \in E_i \cap E_j$ dla $i \neq j$, to rozkład tego wektora na wektory podprzestrzeni nie byłby jednoznaczny, więc bazy E_i są rozłączne. Każdy wektor rozkłada się na wektory rodziny $E_1 \cup \ldots \cup E_s$. Z jednoznaczności rozkładu wektorów na wektory podprzestrzeni oraz z jednoznaczności rozkładów na wektory bazowe w podprzestrzeniach rozkład ten jest jednoznaczny, więc ta rodzina jest bazą przestrzeni V, skąd teza.

(iii)⇒(iv) – oczywiste.

(iv)⇒(i) Jest oczywiste, że przy danych założeniach podprzestrzenie V_i generują V. Należy wykazać liniową niezależność tych podprzestrzeni. Niech $\sum_i x_i = 0$, $x_i \in V_i$. Rozkładając każdy wektor x_i w bazie E_i dostajemy rozkład wektora x w bazie $E_1 \cup \ldots \cup E_s$. Wszystkie współczynniki rozkładu muszą więc znikać, a stąd $x_i = 0$.

(iv)⇒(v) – oczywiste.

(v)⇒(iv) Przy oznaczeniach punktu (iii): każdy wektor rozkłada się na wektory rodziny $E_1 \cup \ldots \cup E_s$. Gdyby istniały między wektorami rodzin E_i liniowe zależności, to wymiar przestrzeni byłby mniejszy, wbrew założeniu, od sumy wymiarów podprzestrzeni (tutaj istotne jest założenie o skończoności wymiaru). Stąd $E_i \cap E_j = \varnothing$ dla $i \neq j$, a rodzina $E_1 \cup \ldots \cup E_s$ jest bazą V. □

Twierdzenie 3. *Jeśli W jest podprzestrzenią przestrzeni V, to istnieje podprzestrzeń W' taka, że $V = W \oplus W'$. Jeśli W'' jest inną podprzestrzenią, dla której również $V = W \oplus W''$, to W'' i W' są izomorficzne.*

Dowód. Niech F będzie bazą podprzestrzeni W. Uzupełniamy ją do bazy E przestrzeni V i kładziemy $W' = L(E \setminus F)$. Z równoważności punktu (i) i (iv) poprzedniego twierdzenia dostajemy $V = W \oplus W'$. Załóżmy, że również $V = W \oplus W''$. Jeśli wymiar przestrzeni V jest skończony, to drugie zdanie twierdzenia jest natychmiastowym wnioskiem z implikacji (i)⇒(v) w poprzednim twierdzeniu. W przypadku ogólnym zauważamy, że każdy wektor $x'' \in W'' \subseteq V$ ma jednoznaczny rozkład $x'' = x + x'$, $x \in W$, $x' \in W'$. Odwzorowanie $x'' \mapsto x'$ jest odwzorowaniem liniowym, ma trywialne jądro (jeśli $x' = 0$, to $x'' \in W \cap W'' = \{0\}$) i jest surjektywne (każdy wektor $x' \in W' \subseteq V$ ma jednoznaczny rozkład $x' = y + y''$, $y \in W$, $y'' \in W''$, czyli $y'' = -y + x'$; stąd $y'' \mapsto x'$), więc jest izomorfizmem. □

3 Przykłady

(i) Rozkład sumy podprzestrzeni na sumę prostą
Niech V_1, V_2 będą podprzestrzeniami przestrzeni V i rozłóżmy zgodnie z twierdzeniem 3:

$$V_1 = (V_1 \cap V_2) \oplus W_1\,, \qquad V_2 = (V_1 \cap V_2) \oplus W_2\,.$$

Wtedy

$$V_1 + V_2 = (V_1 \cap V_2) \oplus W_1 \oplus W_2\,,$$

więc

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)\,.$$

Istotnie, jeśli $x_i \in W_i$, $y \in V_1 \cap V_2$ i $y + x_1 + x_2 = 0$, to $x_1 \in W_1 \cap V_2 = \{0\}$ i $x_2 \in W_2 \cap V_1 = \{0\}$, więc też $y = 0$.

Stąd jeśli $h_1, \dots, h_k$ jest bazą $V_1 \cap V_2$, wektory $f_1, \dots, f_m$ uzupełniają ją do bazy V_1, a wektory $g_1, \dots, g_n$ – do bazy V_2, to łącznie te trzy układy wektorów tworzą bazę $V_1 + V_2$.

(ii) Wyliczenie baz podprzestrzeni
W rzeczywistej przestrzeni pięciowymiarowej V dana jest baza $(e_1, \dots, e_5)$ oraz wektory

$$a_1 = e_1 + 3e_2 + 2e_3 - e_4 + 5e_5\,, \quad a_2 = 2e_1 + e_2 + 3e_3 + 2e_5\,,$$
$$a_3 = -e_1 + 2e_2 - e_3 + e_4 - 3e_5\,,$$
$$b_1 = e_1 + 5e_2 + 4e_3 + e_4 + e_5\,, \quad b_2 = 4e_1 + 8e_2 + 8e_3 + 7e_5\,,$$
$$b_3 = 3e_1 + 3e_2 + 4e_3 - e_4 + 6e_5\,, \quad b_4 = e_1 + 4e_2 + 3e_3 + e_4\,.$$

Szukamy baz podprzestrzeni $V_1 = L(a_1, a_2, a_3)$, $V_2 = L(b_1, b_2, b_3, b_4)$, $V_1 \cap V_2$, $V_1 + V_2$.

Badając liniową niezależność kolumn współrzędnych wektorów stwierdzamy, że wektory (a_1, a_2, a_3) tworzą bazę przestrzeni V_1, a wektory (b_1, b_2, b_4) – bazę przestrzeni V_2. Podprzestrzeń $V_1 \cap V_2$ jest złożona z wszystkich wektorów mających przedstawienia

$$x = \alpha^1 a_1 + \alpha^2 a_2 + \alpha^3 a_3 = \beta^1 b_1 + \beta^2 b_2 + \beta^4 b_4\,.$$

Znalezienie tych wektorów sprowadza się do rozwiązania jednorodnego układu równań o macierzy głównej utworzonej z kolumn współrzędnych wektorów $(a_1, a_2, a_3, b_1, b_2, b_4)$:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 & 1 & 4 & 1 \\ 3 & 1 & 2 & 5 & 8 & 4 \\ 2 & 3 & -1 & 4 & 8 & 3 \\ -1 & 0 & 1 & 1 & 0 & 1 \\ 5 & 2 & -3 & 1 & 7 & 0 \end{pmatrix},$$

z niewiadomymi $(\alpha^1, \alpha^2, \alpha^3, -\beta^1, -\beta^2, -\beta^4)$. Stosując metodę Gaussa (z zachowaniem porządku kolumn) uzyskujemy równoważny układ o macierzy $\mathbf{A}'$ i jego rozwiązanie:

$$\mathbf{A}' = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \alpha^3 \\ -\beta^1 \\ -\beta^2 \\ -\beta^4 \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \\ 1 \\ 1 \\ -2 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ -2 \end{pmatrix}.$$

Bazę $V_1 \cap V_2$ tworzą więc wektory

$$\begin{aligned} h_1 &= 2a_1 + 3a_2 + a_3 = -b_1 + 2b_2 = 7e_1 + 11e_2 + 12e_3 - e_4 + 13e_5 \,, \\ h_2 &= a_2 + a_3 = -b_1 + 2b_4 = e_1 + 3e_2 + 2e_3 + e_4 - e_5 \,. \end{aligned}$$

Z macierzy $\mathbf{A}'$ możemy też odczytać, że bazę przestrzeni $V_1 + V_2$ tworzą wektory (a_1, a_2, a_3, b_1), jednak chcemy też znaleźć bazę zgodną z jednym z rozkładów, które omówiliśmy w poprzednim przykładzie. Musimy zatem znaleźć wektory f_1 i g_1 takie, że transformacje

$$(a_1, a_2, a_3) \;\to\; (h_1, h_2, f_1)\,, \qquad (b_1, b_2, b_4) \;\to\; (h_1, h_2, g_1)$$

są nieosobliwe. Zapewniają to w szczególności wybory:

$$f_1 = a_3 \,, \qquad g_1 = b_1 \,.$$

(iii) Przecięcie i suma nieskończonej rodziny podprzestrzeni
Niech V_i, $i \in I$, będzie rodziną podprzestrzeni przestrzeni V indeksowaną dowolnym zbiorem I, tj. dane jest odwzorowanie

$$I \ni i \mapsto V_i \in (\text{zbiór podprzestrzeni przestrzeni } V)\,.$$

Definicje przecięcia oraz sumy wektorowej podprzestrzeni uogólniają się bezpośrednio na ten przypadek:

$$\bigcap_{i\in I} V_i\,, \qquad \sum_{i\in I} V_i = L\Big(\bigcup_{i\in I} V_i\Big)\,,$$

(por. przykład (iii) w punkcie 6, §1). Suma wektorowa jest zbiorem wszystkich wektorów postaci

$$x = \sum_{i\in I} x_i\,, \quad x_i \in V_i\,,$$

przy czym tylko skończona liczba wektorów x_i jest różna od zera (więc powyższa suma jest sensownie określona). Podprzestrzenie V_i są liniowo niezależne, jeśli z założenia $x = 0$ wynika $x_i = 0$ dla wszystkich $i \in I$. Definicja sumy prostej przenosi się teraz bez zmian na ten ogólny przypadek, obowiązuje także twierdzenie 2.

4 Podprzestrzenie niezmiennicze operatorów

Niech V będzie dowolną przestrzenią wektorową, $W \subseteq V$ jej podprzestrzenią, a A operatorem liniowym na V. Mówimy, że W jest ***podprzestrzenią niezmienniczą*** (lub ***inwariantną***) ***(względem) operatora*** A, gdy $A\,W \subseteq W$, tj. dla każdego $x \in W$ również $Ax \in W$. W tym przypadku zacieśniając zarówno dziedzinę, jak i przeciwdziedzinę, do podprzestrzeni W dostajemy operator liniowy A_W na przestrzeni W, o przepisie

$$A_W : W \mapsto W\,, \quad A_W\, x := Ax\,.$$

Będziemy mówili, że A_W jest ***zacieśnieniem operatora*** A ***do operatora na przestrzeni*** W.

Niech przestrzeń V będzie skończenie wymiarowa, a $(e_1, \ldots, e_n)$ niech będzie taką jej bazą, że $(e_1, \ldots, e_m)$, $m \leq n$, jest bazą podprzestrzeni W. Oznaczmy elementy macierzowe operatora A w bazie (e_i) przez $A^j{}_i$. Jeśli W jest niezmiennicza względem A, to wektory $Ae_1, \ldots, Ae_m$ rozkładają się na wektory $e_1, \ldots, e_m$, więc dla $i = 1, \ldots, m$ i $j = m+1, \ldots, n$ mamy $A^j{}_i = 0$. Odwrotnie, jeśli spełnione są te równości, to dla każdego wektora $x \in W$ jest $Ax \in W$. Podprzestrzeń W jest więc inwariantna względem operatora A wtedy, i tylko wtedy, gdy jego macierz w bazie $(e_1, \ldots, e_n)$ ma postać blokową

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_W & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix}\,,$$

gdzie A_W jest macierzą kwadratową $m \times m$. W tym przypadku macierz $\mathbf{A}_W$ jest macierzą operatora A_W w bazie $(e_1, \ldots, e_m)$.

Zauważmy, że dla każdego operatora A podprzestrzenie $\operatorname{Ker} A$ oraz $\operatorname{Im} A$ są niezmienniczymi podprzestrzeniami.

5 Sumy proste operatorów

Rozważymy teraz sytuację bardziej szczególną: niech $V = W \oplus W'$ i niech obie podprzestrzenie W i W' będą niezmiennicze względem operatora A. Operator A wyznacza wtedy dwa operatory A_W i $A_{W'}$, będące jego zacieśnieniami do operatorów na W i W' odpowiednio. Każdy wektor ma jednoznaczny rozkład $x = y + y'$, gdzie $y \in W$ i $y' \in W'$. Korzystając z definicji operatorów A_W i $A_{W'}$ mamy więc

$$Ax = Ay + Ay' = A_W y + A_{W'} y' .$$

Odwrotnie, niech A_W i $A_{W'}$ będą dowolnymi operatorami na W i W' odpowiednio. Wtedy ostatni przepis zadaje jednoznacznie operator A, względem którego podprzestrzenie W i W' są inwariantne, i którego zacieśnienia do operatorów na W i W' są równe A_W i $A_{W'}$ odpowiednio. Mówimy, że operator A jest ***sumą prostą operatorów*** A_W ***i*** $A_{W'}$, i piszemy

$$A = A_W \oplus A_{W'} .$$

Jeśli V jest skończenie wymiarowa i $(e_1, \ldots, e_n)$ jest taką jej bazą, że $(e_1, \ldots, e_m)$ jest bazą W, a $(e_{m+1}, \ldots, e_n)$ – bazą W', to obie podprzestrzenie są inwariantne względem operatora A wtedy, i tylko wtedy, gdy jego macierz w tej bazie ma postać

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_W & 0 \\ 0 & \mathbf{A}_{W'} \end{pmatrix} ,$$

gdzie A_W jest macierzą $m \times m$, a $A_{W'}$ – macierzą $(n-m)\times(n-m)$. W tym przypadku macierz $\mathbf{A}_W$ jest macierzą operatora A_W w bazie $(e_1, \ldots, e_m)$, a macierz $\mathbf{A}_{W'}$ – macierzą operatora $A_{W'}$ w bazie $(e_{m+1}, \ldots, e_n)$.

Ogólniej, jeśli $V = \bigoplus_{i=1}^{s} V_i$ oraz dane są operatory $A_i : V_i \mapsto V_i$, to zadają one sumę prostą:

$$A = \bigoplus_{i=1}^{s} A_i .$$

Podprzestrzenie V_i są niezmienniczymi podprzestrzeniami operatora A, a operatory A_i – jego zacieśnieniami do operatorów na V_i. Niech V będzie skończenie wymiarowa, a $(e_1, \ldots, e_n)$ niech będzie taką jej bazą, że kolejne podciągi w tym ciągu wektorów są bazami kolejnych podprzestrzeni $V_1, \ldots, V_s$. Wtedy w tej bazie

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{A}_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mathbf{A}_s \end{pmatrix} .$$

Postać ta pokazuje, że

$$\det A = \prod_{i=1}^{s} \det A_i \, .$$

Jeśli B jest innym operatorem o podobnym rozkładzie, $B = \bigoplus_{i=1}^{s} B_i$, z tymi samymi podprzestrzeniami V_i, to

$$\lambda A + \mu B = \bigoplus_{i=1}^{s} (\lambda A_i + \mu B_i) \, , \quad AB = \bigoplus_{i=1}^{s} A_i B_i \, .$$

6 Operatory rzutowe

Niech V będzie dowolną przestrzenią wektorową. Wprowadzimy następujące definicje. Każdy operator P na V spełniający $P^2 = P$ nazywamy ***operatorem rzutowym***. Rodzinę operatorów $P_1, \ldots, P_s$ o własnościach:

$$\sum_{j=1}^{s} P_j = \mathrm{id} \, , \quad P_i P_j = 0 \quad \text{dla} \quad i \neq j \, ,$$

nazywamy ***rozkładem jedności*** (***identyczności***). Zauważmy, że każdy z operatorów P_i jest operatorem rzutowym. Istotnie, składając pierwszy warunek z operatorem P_i i korzystając z drugiego dostajemy $P_i^2 = P_i$ dla każdego i. Stąd równoważnym określeniem rozkładu jedności jest

$$\sum_{j=1}^{s} P_j = \mathrm{id} \, , \quad P_i P_j = \delta_{ij} P_i \, .$$

Motywacją do wprowadzenia powyższych definicji jest następująca ich realizacja. Niech będzie dany rozkład przestrzeni V na sumę prostą, $V = \bigoplus_{i=1}^{s} V_i$. Zadajemy na V operatory liniowe P_i przepisem

$$P_i x = x_i \, , \quad \text{gdzie} \quad x = \sum_{i=1}^{s} x_i \, , \; x_i \in V_i \, .$$

Łatwo sprawdzić wprost z tego określenia, że operatory $P_1, \ldots, P_s$ tworzą rozkład jedności. Mówimy w tym przypadku, że operatory P_i ***rzutują wektory na podprzestrzenie*** V_i (są ***operatorami rzutowania na podprzestrzenie*** V_i). Wszystkie podprzestrzenie $V_1, \ldots, V_s$ są inwariantnymi podprzestrzeniami każdego z operatorów $P_1, \ldots, P_s$. Operator P_i jest sumą prostą identyczności na podprzestrzeni V_i i operatora zerowego na podprzestrzeni $\bigoplus_{j \neq i} V_j$.

Okazuje się, że każda realizacja rozkładu jedności daje się przedstawić w powyższej formie. Mówi o tym następujące twierdzenie.

Twierdzenie 4. *Niech operatory $P_1, \dots, P_s$ tworzą rozkład identyczności w przestrzeni V i oznaczmy $V_i = \operatorname{Im} P_i \equiv P_i V$. Wtedy*

$$V = \bigoplus_{i=1}^{s} V_i$$

i operatory P_i są operatorami rzutowania na podprzestrzenie V_i.

Dowód. Ponieważ $\sum_i P_i = \mathrm{id}$, to dla każdego wektora $x \in V$ mamy $x = \sum_i P_i x$, więc $V = \sum_i V_i$. Jeśli $x_i \in V_i$, to $x_i = P_i\, y_i$ dla pewnego $y_i \in V$. Stąd

$$P_j x_i = P_j P_i y_i = \delta_{ji} P_i y_i = \delta_{ji} x_i\,, \quad \text{więc} \quad P_j \sum_{i=1}^{s} x_i = x_j\,.$$

Stąd też jeśli $\sum_i x_i = 0$, to $x_j = 0$ dla każdego j. Przestrzenie $V_1, \dots, V_s$ są więc liniowo niezależne, co kończy dowód. □

Pojęcie rzutowania na podprzestrzeń zależy nie tylko od wyboru podprzestrzeni, na którą rzutujemy, ale także od sposobu, w jaki uzupełniamy ją do całej przestrzeni. Aby to zrozumieć lepiej, wybierzmy podprzestrzeń W na którą chcemy rzutować. Szukamy więc operatora rzutowego P, dla którego $\operatorname{Im} P = W$. Warunek ten ma wiele rozwiązań, i wybór jednego z nich jest równoważny wyborowi podprzestrzeni W' takiej, że $V = W \oplus W'$ i położeniu $PW' = 0$. Istotnie, wybierzmy jedną z takich podprzestrzeni. Wtedy każdy wektor ma jednoznaczny rozkład $x = y + y'$, gdzie $y \in W$, $y' \in W'$. Z warunków nałożonych na P mamy $y = Pz$ (dla pewnego wektora z), $Py = P^2 z = Pz = y$, $Py' = 0$. Stąd $Px = y$, co określa ten operator jednoznacznie, przy czym $W' = \operatorname{Ker} P$. Odwrotnie, Niech będzie dany operator rzutowy, dla którego $W = \operatorname{Im} P$. Operator ten zadaje jednoznacznie operator $P' := \mathrm{id} - P$, który również jest operatorem rzutowym i wspólnie z operatorem P tworzy rozkład jedności (bo $PP' = P'P = P - P^2 = 0$, $P'P' = P' - PP' = P'$). Mamy w tym przypadku $\operatorname{Im} P' = \operatorname{Ker} P$, więc $V = \operatorname{Im} P \oplus \operatorname{Ker} P$, co kończy dowód. Mówiąc obrazowo, aby określić rzutowanie, oprócz docelowej przestrzeni rzutowania musimy też określić kierunki leżące wzdłuż rzutowania, czyli $\operatorname{Ker} P$.

7 Przykłady

(i) Suma prosta podprzestrzeni, operatory rzutowania na podprzestrzenie
W rzeczywistej czterowymiarowej przestrzeni V dana jest baza (e_1, e_2, e_3, e_4) oraz dwie podprzestrzenie V_1, V_2 zadane swoimi bazami:

$$\begin{aligned} &\text{baza } V_1: \quad e'_1 = e_1 + e_3\,, \quad e'_2 = 2e_1 + e_2\,, \\ &\text{baza } V_2: \quad e'_3 = e_3 + 2e_4\,, \quad e'_4 = e_1 + e_2 - e_3 + e_4\,. \end{aligned}$$

Wyliczając rząd macierzy utworzonej z kolumn współrzędnych wektorów $(e'_1, \ldots, e'_4)$ stwierdzamy, że te wektory również tworzą bazę, więc $V = V_1 \oplus V_2$. Każdy wektor ma jednoznaczny rozkład $x = x_1 + x_2$, gdzie $x_i \in V_i$, co wyznacza odpowiedni rozkład identyczności: $P_i x := x_i$. Działanie operatorów P_i w bazie primowanej (zgodnej z rozkładem) jest

$$\begin{aligned} &P_1 e'_1 = e'_1\,, \quad P_1 e'_2 = e'_2\,, \quad P_1 e'_3 = 0\,, \quad P_1 e'_4 = 0\,, \\ &P_2 e'_1 = 0\,, \quad P_2 e'_2 = 0\,, \quad P_2 e'_3 = e'_3,\,, \quad P_2 e'_4 = e'_4\,, \end{aligned}$$

więc macierze tych operatorów w bazie primowanej są równe

$$\mathbf{P}'_1 = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{P}'_2 = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{1}_2 \end{pmatrix}.$$

Aby otrzymać działanie operatorów rzutowych w pierwotnej bazie, wystarczy odczytać macierz przejścia i znaleźć jej odwrotną: $e'_i = e_j \beta^j{}_i$,

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & 0 & 2 & 1 \end{pmatrix}, \quad \boldsymbol{\beta}^{-1} = \begin{pmatrix} 3 & -6 & -2 & 1 \\ -2 & 5 & 2 & -1 \\ -1 & 2 & 1 & 0 \\ 2 & -4 & -2 & 1 \end{pmatrix},$$

oraz wyliczyć macierze operatorów ze standardowych związków $\mathbf{P}_i = \boldsymbol{\beta}\,\mathbf{P}'_i\,\boldsymbol{\beta}^{-1}$:

$$\mathbf{P}_1 = \begin{pmatrix} -1 & 4 & 2 & -1 \\ -2 & 5 & 2 & -1 \\ 3 & -6 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 2 & -4 & -2 & 1 \\ 2 & -4 & -2 & 1 \\ -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(ii) Podprzestrzeń niezmiennicza z wektorem cyklicznym
Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni V nad dowolnym ciałem $\mathbb{K}$, a x dowolnym wektorem tej przestrzeni. Tworzymy ciąg liniowo niezależnych wektorów $x, Ax, A^2x, \ldots, A^kx$ o tej własności, że ciąg

$x, Ax, \dots, A^{k+1}x$ nie jest liniowo niezależny. Ciąg taki zawsze istnieje: przedłużamy go dotąd, póki można to zrobić z zachowaniem liniowej niezależności, co musi stać się niemożliwe w pewnym kolejnym kroku, gdyż przestrzeń V ma skończony wymiar. Wynika stąd, że wektor $A^{k+1}x$ jest kombinacją liniową wektorów wcześniejszych:

$$A^{k+1}x = \sum_{i=0}^{k} a_i A^i x\,.$$

Oznaczmy $e_i = A^i x$, $i = 0, \dots, k$. Wektory $(e_0, \dots, e_k)$ tworzą bazę podprzestrzeni $W \subseteq V$, która jest inwariantną podprzestrzenią operatora A.

Rozważmy zacieśnienie A_W operatora A do operatora w tej inwariantnej podprzestrzeni. Każdy wektor przestrzeni W można otrzymać przez działanie funkcją wielomianową operatora A_W na wektor x:

$$y = y^i e_i = \sum_{i=0}^{k} y^i A^i x = \left(\sum_{i=0}^{k} y^i (A_W)^i\right) x\,.$$

Każdy wektor x o tej własności nazywamy ***wektorem cyklicznym*** operatora A_W. Operator A_W działa na wektory bazowe według reguły

$$A_W e_i = e_{i+1} \ \text{ dla } \ i = 0, \dots, k-1\,, \quad A_W e_k = \sum_{i=0}^{k} a_i e_i\,,$$

więc macierz tego operatora w tej bazie ma postać

$$\mathbf{A}_W = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & a_k \end{pmatrix}.$$

Jeśli uzupełnić bazę $(e_0, \dots, e_k)$ podprzestrzeni W do bazy $(e_0, \dots, e_n)$ całej przestrzeni V, to macierz operatora A w tej bazie ma postać

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_W & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix}.$$

(iii) Rozkład operatora na składniki

Niech A będzie operatorem liniowym w przestrzeni $V = W \oplus W'$ i niech P i P' będą operatorami rzutowania odpowiednio na W i W'. Wtedy

$$A = PAP + P'AP' + PAP' + P'AP\,.$$

Podprzestrzeń W jest inwariantna względem A wtedy, i tylko wtedy, gdy

$$AP = PAP\,, \qquad \text{czyli} \quad A = PAP + P'AP' + PAP'\,,$$

natomiast warunkiem koniecznym i wystarczającym, aby *obie* podprzestrzenie W i W' były inwariantne jest, by

$$AP = PA\,, \qquad \text{czyli} \quad A = PAP + P'AP'\,.$$

(iv) Przestrzeń funkcji Schwartza
Mówimy, że funkcja $f \in \mathcal{C}^\infty$ jest ***klasy Schwartza***, gdy wszystkie jej pochodne (łącznie z zerową) dążą do zera dla $|x| \to \infty$ szybciej niż dowolna potęga $(1/x)$, tj. spełniają ograniczenia

$$|x^k f^{(l)}(x)| \leq c_{kl} \quad k, l = 0, 1, \ldots\,,$$

gdzie c_{kl} są zależnymi od funkcji stałymi. Zbiór tych funkcji tworzy podprzestrzeń przestrzeni $\mathcal{C}^\infty$, oznaczaną $\mathcal{S}$. Przestrzeń Schwartza $\mathcal{S}$ jest niezmiennicza względem operatorów D i X określonych w przykładzie (ii), p. 2, §10. Każda funkcja o postaci $P(x)\exp[-c\,x^{2n}]$, gdzie $c > 0$, $n \in \mathbb{N}$, a P jest funkcją wielomianową, jest przykładem funkcji klasy Schwartza.

(v) Rozkład operatora na sumę prostą
Rozważamy operator $H = -D^2 + X^2$ na przestrzeni rzeczywistych funkcji nieskończenie różniczkowalnych $\mathcal{C}^\infty$, gdzie D i X są operatorami różniczkowania i mnożenia przez argument, określonymi w przykładzie (ii), p. 2, §10. Funkcję f nazywamy ***parzystą***, gdy $f(-x) = f(x)$ dla każdego argumentu x, a ***nieparzystą***, gdy $f(-x) = -f(x)$ dla każdego argumentu. Podzbiór $\mathcal{C}^\infty_P$ przestrzeni $\mathcal{C}^\infty$ złożony z funkcji parzystych oraz podzbiór $\mathcal{C}^\infty_N$ złożony z funkcji nieparzystych, są podprzestrzeniami wektorowymi. Każdą funkcję $f \in \mathcal{C}^\infty$ można rozłożyć na sumę funkcji parzystej i nieparzystej:

$$f = f_P + f_N\,, \quad \text{gdzie} \quad f_P(x) = \frac{1}{2}\big(f(x) + f(-x)\big)\,, \quad f_N(x) = \frac{1}{2}\big(f(x) - f(-x)\big)\,.$$

Rozkład ten jest jednoznaczny: jeśli $f = f'_P + f'_N$, to podstawiając tu za argument x oraz $-x$ łatwo sprawdzić, że f'_P i f'_N muszą być identyczne z f_P i f_N odpowiednio. Dostajemy więc rozkład na sumę prostą

$$\mathcal{C}^\infty = \mathcal{C}^\infty_P \oplus \mathcal{C}^\infty_N\,.$$

Operator H zachowuje parzystość funkcji: jeśli f jest parzysta (nieparzysta), to Hf jest też parzysta (odpowiednio: nieparzysta). Oznacza to, że $\mathcal{C}_P^\infty$ i $\mathcal{C}_N^\infty$ są podprzestrzeniami niezmienniczymi operatora H. Dostajemy rozkład

$$H = H_P \oplus H_N\,,$$

gdzie H_P (H_N) jest zacieśnieniem operatora H do operatora na przestrzeni $\mathcal{C}_P^\infty$ (odpowiednio: $\mathcal{C}_N^\infty$).

(vi) Rozkład na sumę prostą podprzestrzeni z rozkładu identyczności
W rzeczywistej przestrzeni trójwymiarowej V dana jest baza (e_1, e_2, e_3) oraz operator P_1, którego macierz w tej bazie jest równa

$$\mathbf{P}_1 = \begin{pmatrix} 3 & 2 & -2 \\ -4 & -3 & 4 \\ -1 & -1 & 2 \end{pmatrix}.$$

Sprawdzamy, że $\mathbf{P}_1^2 = \mathbf{P}_1$, co jest równoważne tożsamości $P_1^2 = P_1$, więc P_1 jest operatorem rzutowym i wspólnie z operatorem $P_2 := \mathrm{id} - P_1$ tworzy rozkład identyczności. Macierz operatora P_2 w bazie (e_i) jest równa

$$\mathbf{P}_2 = \mathbf{1} - \mathbf{P}_1 = \begin{pmatrix} -2 & -2 & 2 \\ 4 & 4 & -4 \\ 1 & 1 & -1 \end{pmatrix}.$$

Z rozkładem identyczności związany jest jednoznacznie rozkład V na sumę prostą: $V = V_1 \oplus V_2$, gdzie $V_i = \operatorname{Im} P_i$. Wyliczamy: $\operatorname{rk} P_1 = \operatorname{rk} \mathbf{P}_1 = 2$, $\operatorname{rk} P_2 = \operatorname{rk} \mathbf{P}_2 = 1$, więc $\dim V_1 = 2$, $\dim V_2 = 1$. Stwierdzamy, że np. pierwsze dwie kolumny macierzy $\mathbf{P}_1$ są liniowo niezależne, oraz że każda z kolumn macierzy $\mathbf{P}_2$ jest różna od zera, co daje nam bazy przestrzeni V_i:

$$\begin{aligned} &\text{baza } V_1: \quad e_1' := P_1 e_1 = 3e_1 - 4e_2 - e_3\,, \quad e_2' := P_1 e_2 = 2e_1 - 3e_2 - e_3\,, \\ &\text{baza } V_2: \quad e_3' := P_2 e_1 = -2e_1 + 4e_2 + e_3\,. \end{aligned}$$

Wspólnie wektory (e_1', e_2', e_3') tworzą nową bazę przestrzeni V:

$$\begin{pmatrix} e_1' & e_2' & e_3' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \begin{pmatrix} 3 & 2 & -2 \\ -4 & -3 & 4 \\ -1 & -1 & 1 \end{pmatrix}.$$

§13 Zagadnienie własne operatora liniowego

1 Wartości, podprzestrzenie i wektory własne

Niech A będzie operatorem liniowym na dowolnej przestrzeni wektorowej V nad ciałem $\mathbb{K}$. Mówimy, że liczba $\lambda \in \mathbb{K}$ jest ***wartością własną operatora*** A gdy $\operatorname{Ker}(A-\lambda\,\mathrm{id}) \neq \{0\}$. Podprzestrzeń $\operatorname{Ker}(A-\lambda\,\mathrm{id})$ nazywamy wtedy ***podprzestrzenią własną operatora*** A ***do wartości własnej*** λ. Każdy niezerowy wektor tej podprzestrzeni, czyli każdy niezerowy wektor $x \in V$ spełniający ***równanie własne dla wartości własnej*** λ

$$Ax = \lambda x\,,$$

nazywamy ***wektorem własnym operatora*** A ***do wartości własnej*** λ. Jeśli przestrzeń V ma skończony wymiar, to operator w tej przestrzeni jest surjektywny wtedy, i tylko wtedy, gdy jest injektywny (zob. tw. 6, §10). Zatem operator $A-\lambda\,\mathrm{id}$ jest bijekcją wtedy, i tylko wtedy, gdy λ nie jest jego wartością własną. Ogólnie, ***widmem operatora*** A (lub jego ***spektrum***) nazywamy zbiór $\sigma(A)$ liczb takich, że $A-\lambda\,\mathrm{id}$ nie jest bijektywny dla $\lambda \in \sigma(A)$. W przestrzeni o skończonym wymiarze widmo jest zbiorem wartości własnych operatora, ale w przypadku nieskończenie wymiarowym widmo może być większe, gdyż operator $A-\lambda\,\mathrm{id}$ może nie być surjekcją, pomimo injektywności.

Niech $V_1, \ldots, V_s$ będą podprzestrzeniami własnymi operatora A do wartości własnych $\lambda_1, \ldots, \lambda_s$ odpowiednio oraz niech $x = \sum_{i=1}^{s} x_i$, $x_i \in V_i$. Wtedy $(A-\lambda_j\,\mathrm{id})x_i = (\lambda_i-\lambda_j)x_i$, więc stosując kolejne operatory w poniższym iloczynie (przy zadanym k) do wszystkich wektorów składowych dostajemy:

$$\left(\prod_{\substack{j=1\\ j\neq k}}^{s}(A-\lambda_j\,\mathrm{id})\right)x = \left(\prod_{\substack{j=1\\ j\neq k}}^{s}(\lambda_k-\lambda_j)\right)x_k$$

– pozostałe składniki zerują się, gdyż dla każdego $l \neq k$ w iloczynie wystąpi operator $A-\lambda_l\,\mathrm{id}$, który działając na x_l daje zero. Otrzymujemy stąd następujący wynik.

Twierdzenie 1. *Podprzestrzenie własne każdego operatora liniowego są liniowo niezależnymi podprzestrzeniami niezmienniczymi tego operatora. Jeśli V_λ jest podprzestrzenią własną operatora do wartości własnej λ, to jego zacieśnienie do operatora na V_λ jest równe $\lambda\,\mathrm{id}_{V_\lambda}$.*

Dowód. Z formuły poprzedzającej twierdzenie wynika, że jeśli $x = 0$, to $x_k = 0$ dla wszystkich $k = 1, \ldots, s$, więc podprzestrzenie $V_1, \ldots, V_s$ są liniowo niezależne. Niezmienniczość podprzestrzeni V_λ i postać zawężenia operatora do tej podprzestrzeni wynikają natychmiast z równania własnego. □

2 Zagadnienie własne w przestrzeni skończenie wymiarowej

Niech V będzie przestrzenią o skończonym wymiarze równym n. W takiej przestrzeni maksymalna liczebność rodziny liniowo niezależnych podprzestrzeni wynosi n: gdyby dla pewnej rodziny była ona większa, to wybierając z każdej podprzestrzeni po jednym wektorze dostalibyśmy rodzinę liniowo niezależnych wektorów o liczebności większej niż n, co jest sprzeczne z założeniem.

Niech A będzie operatorem liniowym w V, a $V_1, \ldots, V_s$, $s \leq n$, niech będą jego wszystkimi podprzestrzeniami własnymi, do wartości własnych $\lambda_1, \ldots, \lambda_s$ odpowiednio. Z twierdzenia 1 wiemy, że podprzestrzenie są liniowo niezależne. Załóżmy, że $V = \bigoplus_{i=1}^{s} V_i$. Wtedy, zgodnie z tym samym twierdzeniem, mamy $A = \bigoplus_{i=1}^{s} \lambda_i \,\mathrm{id}_{V_i}$. Operator taki nazywamy ***diagonalizowalnym***. Jako prosty wniosek z dyskusji przeprowadzonej w poprzednim paragrafie dostajemy następujący opis takich operatorów.

Twierdzenie 2. *Operator A w przestrzeni wektorowej V o skończonym wymiarze jest diagonalizowalny wtedy, i tylko wtedy, gdy istnieje rozkład identyczności $\{P_1, \ldots, P_s\}$ i liczby $\lambda_1 \ldots, \lambda_s$ takie, że*

$$A = \sum_{i=1}^{s} \lambda_i P_i \,.$$

*Jeśli ten warunek jest spełniony, to liczby $\lambda_1, \ldots, \lambda_s$ są wartościami własnymi operatora A, a operatory $P_1, \ldots, P_s$ są operatorami rzutowania na odpowiednie podprzestrzenie własne. Powyższe przedstawienie diagonalizowalnego operatora A nazywamy jego **rozkładem spektralnym**.*

Na późniejszy użytek zanotujmy ponadto następujące przedstawienie operatorów rozkładu spektralnego.

Lemat 3. *Przy spełnionych warunkach ostatniego twierdzenia zachodzi*

$$P_k = \left(\prod_{\substack{j=1 \\ j \neq k}}^{s} (\lambda_k - \lambda_j) \right)^{-1} \prod_{\substack{j=1 \\ j \neq k}}^{s} (A - \lambda_j \,\mathrm{id}).$$

Dowód. Teza jest bezpośrednim wynikiem zastosowania, przy obecnych założeniach, formuły poprzedzającej twierdzenie 1. □

Jeśli operator A jest diagonalizowalny, i jeśli przy powyższych oznaczeniach wybrać bazę $(e_1, \dots, e_n)$ przestrzeni V taką, że kolejne podciągi tej bazy są bazami kolejnych podprzestrzeni własnych, to w tej bazie:

$$\mathbf{A} = \begin{pmatrix} \lambda_1 \mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s \mathbf{1}_{n_s} \end{pmatrix} .$$

W tej samej bazie operatory rzutowe jego rozkładu spektralnego mają macierze:

$$\mathbf{P}_1 = \begin{pmatrix} \mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix} , \ \dots, \ \mathbf{P}_s = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{1}_{n_s} \end{pmatrix} .$$

W ogólnym przypadku suma prosta $\bigoplus_{i=1}^{s} V_i$ jest mniejsza od V. Konstruujemy bazę V dołączając, podobnie jak uprzednio, bazy kolejnych podprzestrzeni własnych, ale na końcu uzupełniając otrzymany układ wektorów w dowolny sposób do bazy całej przestrzeni. W takiej bazie

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_\oplus & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix} , \quad \mathbf{A}_\oplus = \begin{pmatrix} \lambda_1 \mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s \mathbf{1}_{n_s} \end{pmatrix} .$$

Macierz $\mathbf{A}_\oplus$ jest macierzą operatora $A_\oplus$, będącego zacieśnieniem operatora A do operatora na przestrzeni $\bigoplus_{i=1}^{s} V_i$.

3 Wielomian charakterystyczny

Opisaliśmy strukturę operatora liniowego, przy założeniu, że jego wartości i podprzestrzenie własne są znane. Zagadnienie ich wyliczenia sprowadza się do znalezienia wszystkich rozwiązań równania $Ax = \lambda x$, gdzie zarówno λ, jak i x są niewiadomymi. W ogólności jest to więc problem na tym etapie nieliniowy. W przypadku, gdy wymiar przestrzeni jest skończony, zadanie można uprościć rozbijając je na dwa kroki: w pierwszym znajdujemy wartości własne, a w drugim, znając je już, rozwiązujemy liniowe już teraz równania na wektory własne. W tym punkcie rozwiązujemy pierwszy z tych problemów.

Niech A będzie operatorem liniowym na skończenie wymiarowej przestrzeni V, a $\mathbf{A}$ – jego macierzą w dowolnie wybranej bazie. Wyznacznik $\det \mathbf{A}$ wyraża się wielomianowo przez elementy macierzy $\mathbf{A}$, jeśli więc wstawić w miejsce $\mathbf{A}$ macierz $\mathbf{A} - \lambda\mathbf{1}$, $\lambda \in \mathbb{K}$, to wyrażenie

$$P_A(\lambda) = \det(A - \lambda\,\mathrm{id}) = \det(\mathbf{A} - \lambda\mathbf{1})$$

określa wielomian zmiennej λ – nazywamy go ***wielomianem charakterystycznym operatora*** A. W wyrażeniu określającym wyznacznik macierzy $\mathbf{A} - \lambda\mathbf{1}$ składnikiem, który zawiera najwyższą potęgę λ, jest iloczyn wyrazów diagonalnych

$$(A^1{}_1 - \lambda)\dots(A^n{}_n - \lambda)\,,$$

gdzie n jest wymiarem przestrzeni V. Stąd wielomian P_A jest stopnia n i jego składnik najwyższego stopnia w λ ma postać $(-\lambda)^n$. Jeśli bazę wybrać tak, jak w poprzednim punkcie, to przy tych samych oznaczeniach

$$P_A(\lambda) = \det(\mathbf{A}_\oplus - \lambda\mathbf{1})\det(\mathbf{C} - \lambda\mathbf{1}) = \det(\mathbf{C} - \lambda\mathbf{1})\prod_{i=1}^{s}(\lambda_i - \lambda)^{n_i}\,.$$

Twierdzenie 4. *Niech A będzie operatorem liniowym w przestrzeni V o skończonym wymiarze n. Wtedy:*

(i) widmo operatora A jest zbiorem co najwyżej n-elementowym, równym zbiorowi pierwiastków jego wielomianu charakterystycznego;

(ii) wymiar podprzestrzeni własnej do wartości własnej λ jest mniejszy bądź równy krotności λ jako pierwiastka wielomianu charakterystycznego;

(iii) jeśli wielomian charakterystyczny operatora A ma n różnych pierwiastków, to operator A jest diagonalizowalny.

Dowód. (i) W przestrzeni skończenie wymiarowej jądro operatora jest trywialne wtedy, i tylko wtedy, gdy jego wyznacznik jest różny od zera. Przez kontrapozycję wynika stąd punkt pierwszy twierdzenia.

(ii) Z formuły poprzedzającej twierdzenie wynika, że krotność pierwiastka λ_i jest większa lub równa n_i, co jest treścią drugiego punktu twierdzenia.

(iii) Jeśli operator A ma n różnych wartości własnych, to w przestrzeni V o wymiarze n jest n liniowo niezależnych podprzestrzeni własnych. To jest możliwe tylko wtedy, gdy wymiar każdej z nich jest równy jeden, a ich suma daje całą przestrzeń. □

Jeśli V jest przestrzenią zespoloną o wymiarze n, to wielomian charakterystyczny P_A ma następujący rozkład na czynniki proste:

$$P_A(\lambda) = (\lambda_1 - \lambda)^{m_1}\dots(\lambda_s - \lambda)^{m_s}\,,$$

gdzie $s \leq n$ i $\sum_{i=1}^{s} m_i = n$ (współczynnik przed całością wyznaczony jest wcześniej otrzymanym warunkiem, że człon o najwyższej potędze λ jest równy $(-\lambda)^n$, co jest zgodne z powyższą postacią). Suma prosta $\bigoplus_{i=1}^{s} V_i$ podprzestrzeni własnych operatora A, o wymiarach $n_1, \ldots, n_s$, ma wymiar $\sum_{i=1}^{s} n_i$. Na podstawie ostatniego twierdzenia $n_i \leq m_i$, $i = 1, \ldots, s$, więc wymiar ten jest równy n wtedy, i tylko wtedy, gdy $n_i = m_i$, $i = 1, \ldots, s$. Otrzymujemy następujące wnioski.

Twierdzenie 5.
(i) Operator liniowy w przestrzeni zespolonej ma co najmniej jedną wartość własną i odpowiednią podprzestrzeń własną.
(ii) Operator w przestrzeni zespolonej jest diagonalizowalny wtedy, i tylko wtedy, gdy wymiary wszystkich jego podprzestrzeni własnych są równe krotnościom odpowiednich pierwiastków wielomianu charakterystycznego.

Na zakończenie tego punktu wracamy do ogólniejszej sytuacji przestrzeni wektorowej V nad dowolnym ciałem $\mathbb{K}$. Niech A będzie operatorem w przestrzeni V i niech W będzie jego inwariantną podprzestrzenią. W bazie wybranej tak, jak w punkcie 4, §12, i przy przejętym z tego punktu oznaczeniu macierzy operatora A w tej bazie, mamy

$$
\begin{aligned}
P_A(\lambda) = \det(\mathbf{A} - \lambda \mathbf{1}_n) = \det(\mathbf{A}_W - \lambda \mathbf{1}_m) \det(\mathbf{C} - \lambda \mathbf{1}_{n-m}) \\
= P_{A_W}(\lambda) \det(\mathbf{C} - \lambda \mathbf{1}_{n-m}) \,.
\end{aligned}
$$

Wykazaliśmy następujące stwierdzenie.

Twierdzenie 6. *Niech A będzie operatorem liniowym w przestrzeni V nad dowolnym ciałem, a W – jego inwariantną podprzestrzenią. Wtedy wielomian charakterystyczny $P_A(\lambda)$ operatora A dzieli się przez wielomian charakterystyczny $P_{A_W}(\lambda)$ jego zacieśnienia A_W do operatora w przestrzeni W.*

4 Wyliczenie wektorów własnych

Niech A będzie operatorem w przestrzeni o skończonym wymiarze n. Zgodnie z poprzednim punktem jego wartości własne znajdujemy jako pierwiastki jego wielomianu charakterystycznego. Wyliczenie wszystkich wektorów własnych do wartości własnej λ sprowadza się teraz do rozwiązania liniowego równania $(A - \lambda\,\mathrm{id})x = 0$. Wybierając dowolną bazę $(e_1, \ldots, e_n)$ w przestrzeni V i znajdując w tej bazie macierz $\mathbf{A}$ operatora A zamieniamy ten problem, zgodnie z tw. 10, §10, na równanie macierzowe

$$(\mathbf{A} - \lambda\mathbf{1})\mathbf{x} = 0\,.$$

Ogólne rozwiązanie tego układu jednorodnego jest podprzestrzenią przestrzeni $\mathbb{K}^n$, wyliczaną tak, jak to opisaliśmy w §9. Podprzestrzeń własna w V do wartości własnej λ jest więc złożona z wektorów $x = x^i e_i$, gdzie $\mathbf{x} \equiv (x^i)$ są wszystkimi rozwiązaniami tego układu.

Dla kolejnych wartości własnych $\lambda_1, \dots, \lambda_s$ wyliczmy bazy podprzestrzeni rozwiązań układu $(\mathbf{A} - \lambda_i\mathbf{1})\mathbf{x} = 0$; liczebność bazy dla wartości własnej λ_i oznaczmy n_i. Tworzymy ciąg $\mathbf{e}'_1, \dots, \mathbf{e}'_m$ liniowo niezależnych wektorów w $\mathbb{K}^n$ ustawiając te bazy jako kolejne jego podciągi, $m = \sum_{i=1}^s n_i$. Jeśli $m = n$, to otrzymany ciąg tworzy bazę $\mathbb{K}^n$, a jeśli $m < n$, to uzupełniamy go w dowolny sposób do bazy $\mathbf{e}'_1, \dots, \mathbf{e}'_n$. Wektory $\mathbf{e}'_i \in \mathbb{K}^n$ są kolumnami współrzędnych wektorów $e'_i \in V$, więc otrzymujemy nową bazę w V:

$$\begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \begin{pmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{pmatrix}\,.$$

Z konstrukcji wektory $e'_1, \dots, e'_m$ są wektorami własnymi operatora A. Jego macierz w bazie primowanej jest więc

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} \mathbf{A}'_\oplus & \mathbf{B}' \\ 0 & \mathbf{C}' \end{pmatrix}\,, \quad \mathbf{A}'_\oplus = \begin{pmatrix} \lambda_1\mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2\mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s\mathbf{1}_{n_s} \end{pmatrix}\,.$$

Operator A jest diagonalizowalny wtedy, i tylko wtedy, gdy $\displaystyle\sum_{i=1}^s n_i = n$, i wówczas

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} \lambda_1\mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2\mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s\mathbf{1}_{n_s} \end{pmatrix}\,.$$

Otrzymany algorytm sprowadzania operatora do diagonalnej, lub częściowo diagonalnej, postaci można też zinterpretować w języku macierzowym. Dla danej macierzy $\mathbf{A}$ znaleźliśmy macierz $\boldsymbol{\beta}$, która przez transformację podobieństwa sprowadza macierz $\mathbf{A}$ do postaci diagonalnej, lub, jeśli całkowita diagonalizacja nie jest możliwa, do postaci częściowo diagonalnej.

5 Przykłady

(i) Operator inwolutywny
Mówimy, że operator Q w dowolnej przestrzeni wektorowej jest ***inwolutywny***, gdy spełnia warunek
$$Q^2 = \mathrm{id}\ .$$
Z pomocą takiego operatora określmy dwa nowe operatory $P_\pm := \frac{1}{2}(\mathrm{id} \pm Q)$. Pokazuje się z łatwością, że operatory te spełniają relacje
$$P_\pm^2 = P_\pm\,, \quad P_+P_- = P_-P_+ = 0\,, \quad P_+ + P_- = \mathrm{id}\,,$$
a więc tworzą rozkład identyczności. Ponadto
$$Q = P_+ - P_-\ ,$$
więc operator Q jest diagonalizowalny, jego widmem jest zbiór $\{+1, -1\}$ (z wyjątkiem szczególnych przypadków, gdy $Q = \pm\,\mathrm{id}$, więc jeden z operatorów rzutowych zeruje się), a powyższa relacja jest jego rozkładem spektralnym.

(ii) Operator diagonalizowalny
W rzeczywistej przestrzeni trójwymiarowej operator A jest zadany swoją macierzą w bazie (e_1, e_2, e_3):
$$\mathbf{A} = \begin{pmatrix} -3 & 3 & 15 \\ 4 & 1 & -10 \\ -2 & 1 & 8 \end{pmatrix} .$$
Wyliczamy wielomian charakterystyczny
$$P_A(\lambda) = \begin{vmatrix} -3-\lambda & 3 & 15 \\ 4 & 1-\lambda & -10 \\ -2 & 1 & 8-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 9\lambda = -\lambda(\lambda-3)^2\,.$$
Wartości własne są równe $0, 3$. Dla $\lambda = 0$ wyliczamy $\mathrm{rk}\,\mathbf{A} = 2$, a dla $\lambda = 3$: $\mathrm{rk}(\mathbf{A} - 3\,\mathbf{1}) = 1$. Stąd $\dim \mathrm{Ker}\, A = 1$, $\dim \mathrm{Ker}(A - 3\,\mathrm{id}) = 2$, więc operator A jest diagonalizowalny:
$$V_1 = \mathrm{Ker}\, A\,,\ V_2 = \mathrm{Ker}(A - 3\ \mathrm{id})\,,\ V = V_1 \oplus V_2\,,\ A = 0\,\mathrm{id}_{V_1} \oplus 3\,\mathrm{id}_{V_2}\ .$$
Rozwiązując układy równań otrzymujemy wektory w $\mathbb{R}^3$: $\mathbf{e}_1'$ napinający przestrzeń rozwiązań układu $\mathbf{A}\mathbf{x} = 0$ oraz $\mathbf{e}_2', \mathbf{e}_3'$ – bazę przestrzeni rozwiązań układu $(\mathbf{A} - 3\,\mathbf{1})\mathbf{x} = 0$:
$$\mathbf{e}_1' = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{e}_2' = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3' = \begin{pmatrix} 0 \\ -5 \\ 1 \end{pmatrix} .$$

Stąd

$$\begin{pmatrix} e_1' & e_2' & e_3' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}\,, \quad \text{gdzie} \quad \boldsymbol{\beta} = \begin{pmatrix} 3 & 1 & 0 \\ -2 & 2 & -5 \\ 1 & 0 & 1 \end{pmatrix},$$

jest bazą własną operatora A, w której jego macierz jest równa

$$\mathbf{A}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta}\,.$$

Ponieważ $\det \boldsymbol{\beta} > 0$, to bazy (e_i) i (e_i') mają wspólną orientację.

(iii) Operator niediagonalizowalny
W bazie (e_1, e_2) w przestrzeni zespolonej operator A ma macierz

$$\mathbf{A} = \begin{pmatrix} i & 1 \\ -2i & -2-i \end{pmatrix}.$$

Wielomian charakterystyczny jest równy $P_A(\lambda) = (\lambda + 1)^2$, więc jedyną wartością własną jest -1. Przestrzeń rozwiązań równania $(\mathbf{A} + \mathbf{1})\mathbf{x} = 0$ jest jednowymiarowa, napięta wektorem $\mathbf{e}_1' = \begin{pmatrix} 1+i \\ -2i \end{pmatrix}$. Operator nie jest diagonalizowalny. Wybieramy w dowolny sposób drugi wektor w $\mathbb{C}^2$ liniowo niezależny od pierwszego, np.: $\mathbf{e}_2' = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$, $\alpha \neq 0$, i określamy nową bazę w V:

$$\begin{pmatrix} e_1' & e_2' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \begin{pmatrix} 1+i & \alpha \\ -2i & 0 \end{pmatrix}.$$

W primowanej bazie macierz operatora A jest

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} -1 & \alpha \\ 0 & -1 \end{pmatrix}.$$

(iv) Operator z jednokrotnymi wartościami własnymi
Macierz operatora A w bazie (e_1, e_2, e_3) trójwymiarowej przestrzeni rzeczywistej jest

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & -2 \\ 4 & -2 & -3 \\ 2 & -2 & -1 \end{pmatrix}.$$

Wielomian charakterystyczny ma trzy pierwiastki: $1, -1, 0$, więc operator jest diagonalizowalny. Znajdujemy:

$$\begin{pmatrix} e_1' & e_2' & e_3' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix},$$

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(v) Operator w przestrzeni rzeczywistej
Niech A będzie operatorem w n-wymiarowej przestrzeni rzeczywistej V. Jego wielomian charakterystyczny ma postać (zob. przykład (ix), p. 9, §4)

$$P(\lambda) = (\lambda_1 - \lambda)^{n_1} \ldots (\lambda_s - \lambda)^{n_s} \big[(\mu_1 - \lambda)^2 + \nu_1^2\big]^{m_1} \ldots \big[(\mu_r - \lambda)^2 + \nu_r^2\big]^{m_r}$$

gdzie $\nu_i \neq 0$. Do każdej z wartości λ_i istnieje podprzestrzeń własna V_i. Czynniki kwadratowe nie mają miejsc zerowych. Niech jednak $\mathbf{A}$ będzie macierzą operatora w dowolnej bazie, i rozpatrzmy zagadnienie własne dla operatora w $\mathbb{C}^n$ działającego jako mnożenie przez tę macierz. Wtedy każdy z czynników kwadratowych daje dwie dodatkowe wartości własne: $\lambda_j = \mu_j + i\nu_j$ oraz $\overline{\lambda_j} = \mu_j - i\nu_j$. Jeśli $\mathbf{A}\mathbf{z} = \lambda_j\mathbf{z}$, to na mocy rzeczywistości $\mathbf{A}$ mamy $\mathbf{A}\overline{\mathbf{z}} = \overline{\lambda_j}\overline{\mathbf{z}}$. Wektory $\mathbf{z}$ i $\overline{\mathbf{z}}$ są liniowo niezależne. Oznaczmy $\mathbf{z} = \mathbf{x} - i\mathbf{y}$. Wówczas $\mathbf{x}$ i $\mathbf{y}$ są liniowo niezależne, a rozkładając równanie własne na część rzeczywistą i urojoną dostajemy $\mathbf{A}\mathbf{x} = \mu_i\mathbf{x} + \nu_i\mathbf{y}$, $\mathbf{A}\mathbf{y} = -\nu_i\mathbf{x} + \mu_i\mathbf{y}$. Teraz możemy wrócić do wyjściowej przestrzeni V, wprowadzając wektory x i y o kolumnach współrzędnych $\mathbf{x}$ i $\mathbf{y}$ odpowiednio. Wektory te rozpinają podprzestrzeń inwariantną operatora A, który działa na nie według przepisu

$$Ax = \mu_j x + \nu_j y\,, \quad Ay = -\nu_j x + \mu_j x\,.$$

Pozostawiamy czytelnikowi uzupełnienie dyskusji prowadzącej do następującego wniosku.

> Operator rzeczywisty, którego wielomian charakterystyczny ma wyżej podaną postać, ma do każdego czynnika $\big[(\mu_j - \lambda)^2 + \nu_j^2\big]^{m_j}$ maksymalną inwariantną podprzestrzeń W_j parzystego wymiaru taką, że w pewnej bazie tej podprzestrzeni macierz zacieśnienia operatora do W_j ma postać blokową
>
> $$\mathbf{A}_j' = \begin{pmatrix} \Lambda_j & 0 & \ldots & 0 \\ 0 & \Lambda_j & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \Lambda_j \end{pmatrix}, \quad \text{gdzie} \quad \Lambda_j = \begin{pmatrix} \mu_j & -\nu_j \\ \nu_j & \mu_j \end{pmatrix},$$
>
> przy czym $\dim W_j \leq 2m_j$. Podprzestrzenie $V_1, \ldots, V_s, W_1, \ldots, W_r$ są liniowo niezależne, ale nie muszą wyczerpywać całej przestrzeni.

(vi) Twierdzenie Cayley-Hamiltona
Niech A będzie operatorem w skończenie wymiarowej przestrzeni V nad dowolnym ciałem. Jeśli A jest diagonalizowalny, to licząc w bazie własnej łatwo przekonać się, że jest spełnione równanie operatorowe $P_A(A) = 0$, gdzie P_A jest wielomianem charakterystycznym operatora A. Okazuje się, że wynik ten jest prawdziwy dla dowolnego operatora (założenie diagonalizowalności jest zbędne).

Niech $x \in V$ będzie dowolnym wektorem. Wektor x generuje podprzestrzeń niezmienniczą W operatora A tak, jak to opisaliśmy w przykładzie (ii), p. 7, §12, którego oznaczenia tu przejmujemy. Z twierdzenia 6 mamy

$$P_A(\lambda) = P_{A_W}(\lambda)\,Q(\lambda)\,, \quad \text{gdzie} \quad Q(\lambda) \text{ jest pewnym wielomianem}\,.$$

Wyliczamy wielomian

$$P_{A_W}(\lambda) = \begin{vmatrix} -\lambda & \dots & 0 & 0 & a_0 \\ 1 & \dots & 0 & 0 & a_1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -\lambda & 0 & a_{k-2} \\ 0 & \dots & 1 & -\lambda & a_{k-1} \\ 0 & \dots & 0 & 1 & a_k - \lambda \end{vmatrix} = (-1)^k \left(\sum_{i=0}^{k} a_i \lambda^i - \lambda^{k+1} \right)$$

(przy wykorzystaniu wyniku przykładu (ix), p. 5, §7). Porównując z wynikami przykładu (ii), p. 7, §12, mamy

$$P_{A_W}(A)x = 0\,, \quad \text{więc} \quad P_A(A)x = Q(A)P_{A_W}(A)x = 0\,.$$

Ponieważ wektor x był wybrany dowolnie, dostajemy twierdzenie (Cayley-Hamiltona):

> Jeśli P_A jest wielomianem charakterystycznym operatora A w przestrzeni wektorowej nad dowolnym ciałem, to
>
> $$P_A(A) = 0\,.$$

(vii) Rezolwenta operatora
Rezolwentą operatora A w przestrzeni V nad dowolnym ciałem nazywa się funkcję

$$\mathbb{K} \setminus \sigma(A) \ni \lambda \mapsto (\lambda\,\mathrm{id} - A)^{-1} \in \mathcal{L}(V, V)\,,$$

gdzie $\sigma(A)$ jest widmem operatora A.

Niech wymiar V będzie skończony. Wielomiany dwóch zmiennych $\lambda^l - \mu^l$ dzielą się przez $\lambda - \mu$ w każdym ciele:

$$\lambda^l - \mu^l = (\lambda - \mu)\sum_{i=0}^{l-1} \lambda^i \mu^{l-1-i}\,,$$

więc dla wielomianu charakterystycznego P operatora A dostajemy

$$P(\lambda) - P(\mu) = (\lambda - \mu) R(\lambda, \mu)\,,$$

gdzie $R(\lambda, \mu)$ jest wielomianem w obu zmiennych. Wstawiając za zmienną μ operator A, a za λ operator $\lambda\,\mathrm{id}$, na mocy twierdzenia Cayley-Hamiltona dostajemy

$$P(\lambda)\,\mathrm{id} = (\lambda\,\mathrm{id} - A) R(\lambda\,\mathrm{id}, A)\,, \quad \text{więc} \quad (\lambda\,\mathrm{id} - A)^{-1} = [P(\lambda)]^{-1} R(\lambda\,\mathrm{id}, A)\,.$$

Stąd w szczególności mamy wielomianowe względem operatora (jeśli nie liczyć wyznacznika w mianowniku) wyrażenie na operator odwrotny (jeśli istnieje): $A^{-1} = -[\det A]^{-1} R(0, A)$.

6 Komutator

Niech V będzie przestrzenią nad ciałem $\mathbb{K}$, a A i B dowolnymi operatorami działającymi w V. ***Komutatorem operatorów*** A i B nazywamy operator

$$[A, B] := AB - BA\,.$$

Niech C będzie jeszcze jednym operatorem w tej samej przestrzeni. Komutator posiada następujące własności.

(i) $[A, \mu B + \nu C] = \mu[A, B] + \nu[A, C]\,, \quad [\mu A + \nu B, C] = \mu[A, C] + \nu[B, C]\,,$ — ***liniowość***,

(ii) $[A, B] = -[B, A]\,,$ — ***antysymetria***,

(iii) $[AB, C] = A[B, C] + [A, C]B\,, \quad [A, BC] = B[A, C] + [A, B]C\,,$

(iv) $\big[[A, B], C\big] + \big[[B, C], A\big] + \big[[C, A], B\big] = 0\,,$ — ***tożsamość Jacobiego***.

Sprawdzenie tych własności pozostawiamy czytelnikowi jako ćwiczenie. Mówimy, że ***operatory*** A ***i*** B ***komutują***, gdy ich komutator znika, $[A, B] = 0$.

Twierdzenie 7. *Jeśli operatory A i B komutują, to dla dowolnych wielomianów P i Q operatory $P(A)$ i $Q(B)$ też komutują.*

Dowód. Pierwsza równość punktu (iii) daje $[A^n, B] = A[A^{n-1}, B] + [A, B]A^{n-1}$, więc jeśli A komutuje z B, to przez indukcję względem n pokazuje się, że również A^n komutuje z B. Wykorzystując teraz w podobny sposób drugą tożsamość punktu (iii) pokazuje się, że w tym przypadku również $[A^n, B^m] = 0$ dla dowolnych $n, m = 0, 1, \ldots$. Z punktu (i) wynika teraz teza. □

7 Komutujące operatory diagonalizowalne

Rozpatrywaliśmy dotąd zagadnienie własne i warunki diagonalizowalności pojedynczego operatora. Istotnym zagadnieniem algebraicznym jest wyjaśnienie, kiedy rodzina diagonalizowalnych operatorów ma równoczesną (to jest we wspólnej bazie) diagonalną postać. Odpowiedź sformułujemy dla przypadku dwóch operatorów, ale twierdzenie ma oczywiste uogólnienie na przypadek dowolnej komutującej rodziny (komplikacja leży wyłącznie w notacji). Pozostawiamy jako ćwiczenie dla czytelnika prześledzenie odpowiednich modyfikacji.

Twierdzenie 8. *Niech A i B będą diagonalizowalnymi operatorami w przestrzeni V, o rozkładach spektralnych*

$$A = \sum_{i=1}^{r} \lambda_i P_i \,, \qquad B = \sum_{j=1}^{s} \mu_j Q_j \,.$$

Wtedy następujące warunki są równoważne:

(i) operatory A i B komutują:

$$[A, B] = 0\,;$$

(ii) operatory rzutowe ich rozkładów spektralnych komutują:

$$[P_i, Q_j] = 0\,, \quad i = 1, \ldots, r\,, \quad j = 1, \ldots, s\,;$$

(iii) operatory A i B mają wspólną bazę wektorów własnych.

Jeśli te warunki są spełnione, to układ operatorów

$$R_{ij} := P_i Q_j\,, \quad i = 1, \ldots, r\,, \quad j = 1, \ldots, s\,,$$

jest rozkładem jedności, przy czym operator R_{ij} rzutuje na podprzestrzeń złożoną ze wspólnych wektorów własnych operatorów A i B do wartości własnych odpowiednio λ_i, μ_j. Każda baza, o której mówi punkt (iii), jest sumą mnogościową pewnych baz wszystkich podprzestrzeni $R_{ij}V$.

Dowód. Wykażemy ciąg implikacji (i)⇒(ii)⇒(iii)⇒(i), skąd wynikają równoważności punktów (i) – (iii).

(i)⇒(ii) Lemat 3 daje reprezentację operatorów rozkładu spektralnego operatora diagonalizowalnego w postaci funkcji wielomianowych tego operatora. Z twierdzenia 7 dostajemy więc natychmiast żądaną implikację.

(ii)⇒(iii) Wykażemy najpierw wynikanie z warunku (ii) pierwszego zdania następującego po punkcie (iii). Mamy

$$\sum_{ij} R_{ij} = \sum_i P_i \sum_j Q_j = \mathrm{id}\ .$$

Ponadto, przy spełnionym warunku (ii), zachodzi

$$R_{i_1j_1}R_{i_2j_2} = P_{i_1}P_{i_2}Q_{j_1}Q_{j_2} = \delta_{i_1i_2}\delta_{j_1j_2}R_{i_1j_1}\ ,$$

więc rodzina operatorów $\{R_{ij}\}$ tworzy rozkład jedności. Przy nadal założonym warunku (ii) mamy też:

$$AR_{ij} = \sum_k \lambda_k P_k P_i Q_j = \lambda_i R_{ij}\ , \quad BR_{ij} = \sum_l \mu_l P_i Q_l Q_j = \mu_j R_{ij}\ ,$$

co kończy dowód zdania następującego po warunku (iii). Wybierzmy teraz bazę każdej z podprzestrzeni $R_{ij}V$. Łącznie bazy te dają żądaną wspólną bazę własną operatorów A i B.

(iii)⇒(i) Działanie operatorów A i B na wektory wspólnej bazy własnej sprowadza się do mnożenia przez liczby (odpowiednie wartości własne). Z przemienności mnożenia liczb dostajemy więc natychmiast implikację.

Pozostaje wykazać ostatnie stwierdzenie twierdzenia. Jeśli dana jest dowolna wspólna baza własna operatorów A i B, to jej podrodzina złożona z wektorów będących jednocześnie wektorami własnymi tych operatorów do wartości własnych λ_i, μ_j jest bazą podprzestrzeni $R_{ij}V$, co kończy dowód. □

8 Funkcje operatorów

Używając elementarnych operacji składania operatorów i tworzenia ich kombinacji liniowych, mogliśmy przyporządkować każdemu wielomianowi Q w ciele $\mathbb{K}$ funkcję operatorową $\mathcal{L}(V,V) \ni A \mapsto Q(A) \in \mathcal{L}(V,V)$, gdzie V jest dowolną, ustaloną przestrzenią nad ciałem $\mathbb{K}$ (punkt 15, §10). Przyporządkowanie to zachowuje strukturę liniową i multiplikatywną w następującym sensie. Niech f i g będą dowolnymi wielomianami, a przez $f(..)$ i $g(..)$ oznaczmy definiowane za ich pomocą funkcje operatorowe. Przyporządkowanie funkcji $f(..)$ wielomianowi f zapiszmy jako $f \mapsto f(..)$. Wtedy dla dowolnych $\alpha, \beta \in \mathbb{K}$ jest

$$\alpha f + \beta g \mapsto \alpha f(..) + \beta g(..)\,, \qquad fg \mapsto f(..)g(..)\,.$$

Istnieją różne rozszerzenia tak określonej struktury, ale w ogólnym przypadku powstają one za cenę ograniczenia dziedziny operatorowej. Poniżej, w przykładzie (vi), p. 9, wskazujemy na rozszerzenie określone przez wyrażenia wymierne. Tutaj omówimy najpierw funkcje operatorów diagonalizowalnych, a potem zajmiemy się przypadkiem operatorów w rzeczywistych lub zespolonych przestrzeniach.

Niech x będzie wektorem własnym operatora A, $Ax = \lambda x$. Wtedy, używając elementarnych własności operatorów, dostajemy $Q(A)x = Q(\lambda)x$ dla dowolnego wielomianu Q. Stąd, jeśli operator A jest diagonalizowalny, to operator $Q(A)$ jest również diagonalizowalny, wektory własne operatora A są wektorami własnymi operatora $Q(A)$, z wartościami własnymi równymi wartościom wielomianu dla odpowiednich wartości własnych operatora A. Sformułujemy to w języku rozkładu spektralnego: jeśli $A = \sum_{i=1}^{s} \lambda_i P_i$, to $Q(A) = \sum_{i=1}^{s} Q(\lambda_i)P_i$. Kierując się tą własnością wielomianów wprowadzamy następującą definicję:

$$\text{jeśli} \quad A = \sum_{i=1}^{s} \lambda_i P_i \quad \text{i} \quad f : \mathbb{K} \mapsto \mathbb{K}\,, \quad \text{to} \quad f(A) := \sum_{i=1}^{s} f(\lambda_i) P_i\,.$$

Analogicznie jak na początku tego punktu, otrzymujemy przyporządkowanie $f \mapsto f(..)$, ale teraz dziedziną tego przyporządkowania jest zbiór wszystkich *funkcji* w ciele $\mathbb{K}$, a funkcje operatorowe $f(..)$ są ograniczone do operatorów diagonalizowalnych. Przy tych znaczeniach symboli zachowywanie struktury liniowej i struktury multiplikatywnej obowiązuje także w tym przypadku.

W przypadku przestrzeni rzeczywistej lub zespolonej, przy czym ograniczamy się tutaj do przestrzeni o skończonym wymiarze, istnieje sposób uogólnienia wielomianowych (a także tych określonych przez wyrażenia wymierne) funkcji operatorowych przy użyciu przejścia granicznego. W takich przestrzeniach istnieje w naturalny sposób określone pojęcie zbieżności: ciąg wektorów x_l zmierza do granicy x,

$$\lim_{l\to\infty} x_l = x\,, \quad \text{gdy w pewnej bazie} \quad \lim_{l\to\infty} x_l^i = x^i\,, \quad i = 1, \ldots, n\,.$$

Jeśli zachodzi to w jednej bazie, to zachodzi również w każdej innej, więc pojęcie zbieżności nie zależy od wyboru bazy. To pojęcie obejmuje również przypadek ciągu operatorów A_l zbieżnego do operatora A:

$$\lim_{l\to\infty} A_l = A\,, \quad \text{gdy} \quad \lim_{l\to\infty} (A_l)^i{}_j = A^i{}_j\,.$$

Ograniczymy się do przejścia granicznego polegającego na sumowaniu nieskończonego szeregu potęgowego zmiennej operatorowej. Przypomnijmy ponadto, że w ciałach $\mathbb{R}$ i $\mathbb{C}$, jak w każdym ciele nieskończonym, wielomiany można utożsamić z funkcjami wielomianowymi.

Niech funkcja f określona będzie dla $|z| < R$ za pomocą szeregu potęgowego $f(z) = \sum_{l=0}^{\infty} c_l \, z^l$ o promieniu zbieżności R. Jeśli dla wszystkich $l \in \mathbb{N}$ w pewnej bazie wszystkie elementy macierzy operatorów A^l spełniają ograniczenia $|(A^l)^i{}_j| \leq cr^l$, gdzie c i r są stałymi, $r < R$, to szeregi liczbowe $\sum_{l=0}^{\infty} c_l (A^l)^i{}_j$ są bezwzględnie zbieżne. Są one wtedy zbieżne w każdej bazie. Szereg operatorowy w następującej definicji operatora $f(A)$ jest więc zbieżny z definicji zbieżności w przestrzeni nad $\mathbb{R}$ lub $\mathbb{C}$:

$$f(A) := \sum_{l=0}^{\infty} c_l \, A^l \, .$$

Jeśli A jest operatorem diagonalizowalnym, to rozpatrując szereg określający funkcję $f(A)$ w jego bazie własnej łatwo wykazać, że jest on bezwzględnie zbieżny wtedy, i tylko wtedy, gdy $\max\{|\lambda_1|, \ldots, |\lambda_s|\} < R$, oraz że obecna definicja funkcji operatora pokrywa się wtedy z poprzednią. Do dyskusji zakresu stosowalności ogólnych potęgowych szeregów operatorów zespolonych lub rzeczywistych wrócimy w Uzupełnieniach (w przykładzie (iii), p. 8, §25), tutaj ograniczymy się do szeregów o nieskończonym promieniu zbieżności.

Jeśli f jest funkcją holomorficzną na całej płaszczyźnie $\mathbb{C}$ (tj. określający ją szereg ma nieskończony promień zbieżności), to operator $f(A)$ jest poprawnie określony za pomocą szeregu dla każdego operatora A. Wynika to z następującej własności macierzy kwadratowych: jeśli $|A^i{}_j| \leq a$ dla wszystkich elementów macierzy, to elementy wszystkich potęg tej macierzy spełniają założone wyżej ograniczenia z podstawieniami $c = \frac{1}{n}$, $r = na$ (n jest wymiarem macierzy). Wykazuje się tę własność przez indukcję, korzystając z nierówności: $|(A^{l+1})^i{}_j| \leq \sum_{k=1}^{n} |(A^l)^i{}_k||A^k{}_j|$; szczegóły zostawiamy jako ćwiczenie dla czytelnika. Dla funkcji holomorficznych przyporządkowanie $f \mapsto f(..)$ definiuje więc funkcje $f(..)$ określone na całej dziedzinie $\mathcal{L}(V, V)$; zachowanie struktury liniowej oraz struktury multiplikatywnej jest wypełnione także dla tej definicji (dla funkcji nieholomorficznych obowiązuje po odpowiednim zacieśnieniu dziedziny zmiennej operatorowej).

Kończymy ten punkt ważnym wzmocnieniem twierdzenia 7.

Twierdzenie 9. *Jeśli operatory A i B komutują, to również $[f(A), g(B)] = 0$, gdzie każda z funkcji operatorowych $f(A)$ i $g(B)$ jest określona na jeden z wyżej omówionych sposobów.*

Dowód. (i) Jeśli jeden z operatorów jest diagonalizowalny, np. $A = \sum_i \lambda_i P_i$, to wtedy $[P_i, B] = 0$ (bo P_i można wyrazić jako wielomian operatora A, patrz lemat 3). Stąd, jeśli

$$x \in P_i V\,, \quad \text{a więc} \quad P_i x = x\,, \quad \text{to} \quad Bx = BP_i x = P_i Bx \in P_i V\,,$$

czyli każda podprzestrzeń własna operatora A jest podprzestrzenią niezmienniczą operatora B, więc także operatora $g(B)$. Ale na każdej podprzestrzeni własnej operator $f(A)$ jest proporcjonalny do operatora identycznościowego, więc komutuje z $g(B)$.

(ii) Jeśli oba operatory $f(A)$ i $g(B)$ są określone za pomocą szeregów, to ich komutator jest iterowaną granicą komutatorów skończonych szeregów, które są równe zeru. □

9 Przykłady

(i) Jednoczesna diagonalizacja dwóch operatorów
W bazie (e_1, e_2, e_3) przestrzeni rzeczywistej V dane są macierze dwóch operatorów A i B:

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 4 \\ 6 & 4 & -12 \\ -3 & 0 & 10 \end{pmatrix}\,, \quad \mathbf{B} = \begin{pmatrix} 22 & 10 & 0 \\ -30 & -18 & 0 \\ 15 & 5 & -8 \end{pmatrix}\,.$$

Pytamy, czy operatory te są jednocześnie diagonalizowalne, i jeśli tak, to jaka jest ich wspólna baza wektorów własnych.

Sprawdzamy, że $\mathbf{AB} = \mathbf{BA}$, więc operatory komutują. Jeśli są diagonalizowalne, to mają wspólną bazę wektorów własnych.

Rozwiązujemy zagadnienie własne dla operatora A. Jego wartości własne to 8 i 4. Wyliczamy bazy podprzestrzeni własnych

$$\begin{aligned} &\text{dla } \lambda = 8: && f_1 = -2e_1 + 6e_2 - 3e_3\,, \\ &\text{dla } \lambda = 4: && f_2 = e_2\,, \quad f_3 = 2e_1 + e_3\,. \end{aligned}$$

Operator A jest diagonalizowalny.

Operator B komutuje z A, więc podprzestrzenie własne operatora A są niezmiennicze względem operatora B. Stąd f_1 jest jego wektorem własnym. Działając macierzą $\mathbf{B}$ na kolumnę współrzędnych wektora f_1 dostajemy

$$Bf_1 = -8f_1\,.$$

Oznaczmy zacieśnienie operatora B do podprzestrzeni rozpiętej wektorami f_2, f_3 przez B_{23}. Wyliczamy:

$$\begin{aligned} B_{23}f_2 &= 10e_1 - 18e_2 + 5e_3 = -18f_2 + 5f_3\,, \\ B_{23}f_3 &= 44e_1 - 60e_2 + 22e_3 = -60f_2 + 22f_3\,. \end{aligned}$$

Macierz operatora B_{23} w bazie (f_2, f_3) jest więc równa

$$\mathbf{B}_{23} = \begin{pmatrix} -18 & -60 \\ 5 & 22 \end{pmatrix}.$$

Rozwiązujemy zagadnienie własne dla operatora B_{23} w bazie (f_2, f_3). Znajdujemy:

$$B_{23}e'_2 = -8\,e'_2\,, \quad B_{23}e'_3 = 12\,e'_3\,,$$

gdzie

$$e'_2 = -6f_2 + f_3 = 2e_1 - 6e_2 + e_3\,, \quad e'_3 = -2f_2 + f_3 = 2e_1 - 2e_2 + e_3\,.$$

Oznaczamy ponadto $e'_1 = f_1$. Dostajemy ostatecznie

$$\begin{pmatrix} e'_1 & e'_2 & e'_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix}\boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \begin{pmatrix} -2 & 2 & 2 \\ 6 & -6 & -2 \\ -3 & 1 & 1 \end{pmatrix},$$

$$\mathbf{A}' = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta}\,, \quad \mathbf{B}' = \begin{pmatrix} -8 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 12 \end{pmatrix} = \boldsymbol{\beta}^{-1}\mathbf{B}\boldsymbol{\beta}\,.$$

(ii) Operator o macierzy rosnącej liniowo z potęgą
Rozważmy operator A na dwuwymiarowej przestrzeni rzeczywistej, którego macierz w pewnej bazie jest równa $\mathbf{A} = \mathbf{1} + \alpha E$, gdzie $E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Ponieważ $E^2 = 0$, to $\mathbf{A}^k = \begin{pmatrix} 1 & k\alpha \\ 0 & 1 \end{pmatrix}$. Jeśli promień zbieżności szeregu $f(z) = \sum_{k=1}^{\infty} c_k z^k$ jest większy od 1, to istnieje operator $f(A)$.

(iii) Operatory rzutowania zgodne z rozkładem spektralnym
Niech $\chi_\Omega : \mathbb{K} \mapsto \mathbb{K}$ będzie funkcją charakterystyczną podzbioru $\Omega \subseteq \mathbb{K}$:

$$\chi_\Omega(\alpha) = 1\,, \text{ gdy } \alpha \in \Omega\,, \quad \text{i} \quad \chi_\Omega(\alpha) = 0\,, \text{ gdy } \alpha \notin \Omega\,.$$

Niech A będzie dowolnym operatorem diagonalizowalnym i $A = \sum_{i=1}^{s} \lambda_i P_i$. Wtedy $\chi_\Omega(A) = \sum_{i:\lambda_i\in\Omega} P_i$. Operator $\chi_\Omega(A)$ jest operatorem rzutowania na sumę podprzestrzeni $\bigoplus_{i:\lambda_i\in\Omega} P_i V$.

(iv) Eksponenta operatora i jej wyznacznik
Szereg określający funkcję eksponencjalną zespolonego lub rzeczywistego argumentu $e^z \equiv \exp(z) = \sum_{k=0}^{\infty} \dfrac{z^k}{k!}$ ma nieskończony promień zbieżności, więc operator $e^A \equiv \exp(A)$ jest określony dla każdego operatora A w przestrzeni rzeczywistej lub zespolonej. Eksponenta operatora posiada następującą własność:

$$\text{jeśli} \quad [A,B]=0\,, \quad \text{to} \quad e^{(A+B)} = e^A e^B\,.$$

Dowód prowadzi się tak samo, jak dla liczb – przez przegrupowanie wyrazów w iloczynie szeregów po prawej stronie.

Rozważymy dokładniej funkcję $\mathbb{R} \ni t \mapsto \exp(tA)$, gdzie A jest ustalonym operatorem. Mamy:

$$e^{(s+t)A} = e^{sA} e^{tA}\,.$$

Niech przestrzeń, w której działa A, będzie n-wymiarowa i niech w pewnej bazie $|A^i{}_j| \le a$. Wtedy dla $|s| \le (na)^{-1}$ mamy

$$|(e^{sA} - \mathrm{id} - sA)^i{}_j| \le \frac{1}{n} \sum_{k=2}^{\infty} \frac{(|s|na)^k}{k!} \le \frac{(sna)^2}{n} \sum_{k=2}^{\infty} \frac{1}{k!} \le cs^2\,,$$

gdzie c jest stałą. Ponieważ $e^{(s+t)A} - e^{tA} = (e^{sA} - \mathrm{id})e^{tA}$, to przy użyciu powyższego ograniczenia można pokazać, że funkcja e^{tA} jest ciągła i różniczkowalna (różniczkowalność jest określona, podobnie jak ciągłość, przez przejście do współrzędnych w dowolnej bazie) oraz

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A\,,$$

więc jej pochodna jest ciągła. Ponadto, wykorzystanie powyższej nierówności w definicji wyznacznika prowadzi do wyniku:

$$\lim_{s\to 0} \frac{1}{s}\big(\det e^{sA} - 1\big) = \lim_{s\to 0} \frac{1}{s}\big(\det(\mathrm{id} + sA) - 1\big) = \mathrm{Tr}\, A\,.$$

Ostatni wynik użyty w równaniu otrzymanym z twierdzenia Cauchy'ego:

$$\det e^{(s+t)A} = \det e^{sA} \det e^{tA}$$

daje

$$\frac{d}{dt}\big[\det e^{tA}\big] = \operatorname{Tr} A\big[\det e^{tA}\big]\,.$$

Rozwiązując to równanie z warunkiem początkowym $\det e^{0A} = 1$ dostajemy $\det e^{tA} = e^{t \operatorname{Tr} A}$. Stąd mamy wynik:

$$\det e^{A} = e^{\operatorname{Tr} A}\,.$$

Inny dowód tej tożsamości otrzymuje się przez użycie postaci kanonicznej operatora zespolonego, omówionej w Uzupełnieniach (i zastosowanej do obliczenia szeregu operatorowego w przykładzie (iii), p. 8, §25).

(v) Jednoparametrowe, ciągłe grupy operatorów

Niech dla każdego $t \in \mathbb{R}$ dany będzie operator $U(t)$ w pewnej skończenie wymiarowej przestrzeni rzeczywistej lub zespolonej V. Mówimy, że operatory $U(t)$ tworzą ***jednoparametrową, ciągłą grupę operatorów***, gdy dla każdych $t, s \in \mathbb{R}$ spełnione są warunki

$$U(s)U(t) = U(s+t)\,, \qquad U(0) = \mathrm{id}\,,$$

oraz odwzorowanie $\mathbb{R} \ni t \mapsto U(t) \in \mathcal{L}(V,V)$ jest ciągłe. Z pierwszego warunku widać w szczególności, że $U(t)U(-t) = \mathrm{id}$, więc $U(t) \in GL(V)$. Odwzorowanie $\mathbb{R} \ni t \mapsto U(t) \in GL(V)$ jest homomorfizmem grupowym.

Jeśli $A \in \mathcal{L}(V,V)$, to operatory $U(t) = \exp(tA)$ tworzą jednoparametrową, ciągłą grupę, co wynika z dyskusji przeprowadzonej w poprzednim przykładzie. Mówimy w tym przypadku, że A jest ***generatorem*** jednoparametrowej grupy $U(t)$. Trudniejszym zadaniem jest wykazanie następującego stwierdzenia odwrotnego.

> Dla każdej jednoparametrowej, ciągłej grupy operatorów $U(t)$ istnieje dokładnie jeden operator A, który ją generuje: $U(t) = \exp(tA)$.

Dla dowodu wykazuje się najpierw, że z definicji jednoparametrowej, ciągłej grupy operatorów wynika różniczkowalność funkcji $t \mapsto U(t)$ – naszkicujemy ten krok poniżej, a zróżniczkowanie warunku homomorficzności po s i położenie $s = 0$ daje równanie z warunkiem początkowym:

$$\frac{d}{dt}U(t) = AU(t)\,, \quad U(0) = \mathrm{id}\,, \quad \text{gdzie} \quad A \equiv \frac{d}{ds}U(s)\Big|_{s=0}\,.$$

Stąd łatwo pokazać, że $\frac{d}{dt}[\exp(-tA)U(t)] = 0$, $[\exp(-tA)U(t)]_{t=0} = \mathrm{id}$, więc $U(t) = \exp(tA)$. Odwrotnie, jeśli $U(t)$ ma taką postać, to generator A wyraża się jako pochodna, więc jest jednoznacznie określony.

Dla wykazania różniczkowalności funkcji $U(t)$ wybiera się funkcję $f : \mathbb{R} \mapsto \mathbb{R}$, o ciągłej pochodnej, taką że $f \geq 0$, $f(s) = 0$ dla wszystkich s na zewnątrz pewnego przedziału $\langle -\delta, \delta\rangle$ oraz $\int_{\mathbb{R}} f(s)ds = 1$. Mnożąc warunek definicyjny jednoparametrowej grupy przez $f(s)$ i całkując po s dostaje się (po zmianie zmiennej całkowania po prawej stronie)

$$\int_{\mathbb{R}} f(s)U(s)ds\ U(t) = \int_{\mathbb{R}} f(r-t)U(r)dr$$

– całkowanie wykonuje się dla elementów macierzy operatorów w dowolnej bazie, które z założenia są ciągłymi funkcjami. Operator po prawej stronie tej równości ma pochodną po t równą $-\int f'(r-t)U(r)dr$ – wynika to z twierdzenia o różniczkowaniu pod znakiem całki. Pierwszy czynnik po lewej stronie można zapisać jako

$$\mathrm{id} + \int_{\mathbb{R}} f(s)[U(s) - \mathrm{id}]ds\,.$$

Jeśli parametr δ jest dostatecznie mały, to wyznacznik tego operatora jest bliski 1 (z ciągłości $U(s)$), więc operator jest odwracalny, co kończy dowód różniczkowalności $U(t)$.

(vi) Funkcje wymierne operatorów
Rozszerzamy określenie funkcji operatorowych (dla dowolnego ciała) z pomocą wyrażeń wymiernych (przykład (vii), p. 15, §3). Jeśli wielomiany P i Q są względnie pierwsze, to wyrażenie wymierne $\dfrac{P}{Q}$ określa funkcję

$$A \mapsto \frac{P(A)}{Q(A)} \equiv P(A)[Q(A)]^{-1} = [Q(A)]^{-1}P(A)$$

na zbiorze operatorów, dla których $\det Q(A) \neq 0$. Ostatnia równość w powyższym przepisie wynika z równości $P(A)Q(A) = Q(A)P(A)$. Przyporządkowanie $\dfrac{P}{Q} \mapsto \dfrac{P(..)}{Q(..)}$ (oznaczenia analogiczne jak w punkcie 8, ale z uwzględnieniem ograniczeń na dziedziny) zachowuje strukturę liniową i strukturę multiplikatywną w następującym sensie. Jeśli

$$\alpha'\frac{P'}{Q'} + \alpha''\frac{P''}{Q''} = \frac{P_1}{Q_1}\,, \qquad \frac{P'}{Q'}\frac{P''}{Q''} = \frac{P_2}{Q_2}\,,$$

gdzie w parach (P_i, Q_i) wielomiany są względnie pierwsze, to

$$\alpha' \frac{P'(A)}{Q'(A)} + \alpha'' \frac{P''(A)}{Q''(A)} = \frac{P_1(A)}{Q_1(A)}\,, \qquad \frac{P'(A)}{Q'(A)} \frac{P''(A)}{Q''(A)} = \frac{P_2(A)}{Q_2(A)}\,,$$

dla tych operatorów A, dla których lewe strony są określone.

III

PRZESTRZENIE WEKTOROWE Z ILOCZYNEM SKALARNYM

§14 Iloczyny skalarne

1 Podstawowe definicje

W tym paragrafie zaczynamy dyskusję zagadnień, które można uważać za *geometryczne*. Pojęcie geometrii jest dziś niezmiernie szerokie i trudno jest podać jej precyzyjną definicję i zakres. Najprościej będzie odwołać się do historycznych jej źródeł, a także raczej do intuicji, niż precyzyjnych pojęć, i powiedzieć, że zajmuje się ona rozważaniem relacji przestrzennych między tworami takimi jak punkty, linie, powierzchnie itd., ujmowaniem tych relacji w takie liczbowe charakterystyki jak odległości, kąty, itp., badaniem transformacji tych obiektów i symetrii otrzymanych struktur. Może być rozwijana aksjomatycznie, na wzór aksjomatyki euklidesowskiej. Dziś przeważa jednak inne podejście, bardziej efektywne, wykorzystujące struktury i twory dostarczane przez różne gałęzie matematyki. Gdy do rozważań wystarcza język algebry liniowej, to mówimy o geometriach *płaskich*, czy *liniowych*.

Wśród geometrii płaskich kluczowe znaczenie dla nauk przyrodniczych mają geometrie uzyskane przez wyposażenie przestrzeni wektorowej w funkcję przypisującą liczby parom wektorów. Na funkcję tę nakładamy warunki zgodne z duchem algebry liniowej i wprowadzamy następującą definicję.

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$. ***Iloczynem skalarnym*** (lub ***metryką***, lub ***formą metryczną***) nazywać będziemy odwzorowanie

$$g : V \times V \mapsto \mathbb{K}\,, \quad (x, y) \mapsto g(x, y)\,,$$

spełniające następujące warunki dla wszystkich wektorów i liczb:

(i) g jest liniowe w prawym argumencie, czyli

$$g(x, \alpha y + \beta z) = \alpha g(x, y) + \beta g(x, z)\,;$$

(ii) g jest liniowe w lewym argumencie, czyli

$$g(\alpha x + \beta y, z) = \alpha g(x, z) + \beta g(y, z)\,,$$

lub (ten warunek zakłada, że $\mathbb{K} = \mathbb{C}$):
g jest antyliniowe w lewym argumencie, czyli

$$g(\alpha x + \beta y, z) = \overline{\alpha} g(x, z) + \overline{\beta} g(y, z)\,;$$

(iii) spełniona jest równoważność

$$g(x, y) = 0 \Longleftrightarrow g(y, x) = 0\,.$$

Każde odwzorowanie spełniające warunki (i) i (ii) nazywamy ***formą biliniową*** w przypadku warunku liniowości w (ii), lub ***formą półtoraliniową*** w przypadku warunku antyliniowości w (ii). Najbardziej ogólna definicja iloczynu skalarnego ogranicza się do tych dwóch punktów, jednak dodanie punktu (iii) pozwala uprościć wiele zapisów, a tak otrzymana definicja jest dostatecznie szeroka dla większości zastosowań i będzie konsekwentnie stosowana w tym podręczniku.

Wprowadzonego tutaj terminu *metryka* nie należy utożsamiać z tak samo brzmiącym terminem wprowadzanym w topologii – związek tych pojęć wyjaśnimy później.

O parze wektorów x i y mówimy, że są ***ortogonalne*** (lub ***prostopadłe***), gdy $g(x, y) = 0$. Trzeci warunek przyjętej przez nas definicji iloczynu skalarnego można więc ująć mówiąc, że relacja ortogonalności jest symetryczna. Wprowadzamy oznaczenie tej relacji:

$$x \perp y \Longleftrightarrow g(x, y) = 0\,.$$

Będziemy również mówić, że ***podzbiory W i W' przestrzeni V są ortogonalne***, i pisać $W \perp W'$, gdy każdy wektor ze zbioru W jest prostopadły do każdego wektora ze zbioru W'. Z liniowości metryki w prawym argumencie wynika, że

$$\text{jeśli} \quad x \perp y \quad \text{i} \quad x \perp z\,, \quad \text{to} \quad x \perp (\alpha y + \beta z)$$

dla każdej pary liczb α, β.

Przestrzeń wektorową V wraz z zadanym na niej iloczynem skalarnym g nazywamy ***przestrzenią z iloczynem skalarnym*** i oznaczamy (V, g).

2 Ortogonalne dopełnienie. Jądro metryki

Niech $W \subseteq V$ będzie dowolnym podzbiorem przestrzeni. ***Ortogonalnym dopełnieniem zbioru*** W nazywamy zbiór utworzony ze wszystkich wektorów ortogonalnych do wszystkich wektorów zbioru W:

$$W^{\perp} := \{x \in V \mid \forall y \in W : y \perp x\} .$$

Z własności zamykającej ostatni punkt jest widoczne, że jeśli $x, z \in W^{\perp}$, to również dowolna ich kombinacja liniowa należy do $W^{\perp}$, więc ortogonalne dopełnienie każdego zbioru jest podprzestrzenią.

Jądrem metryki nazywamy podprzestrzeń $\operatorname{Ker} g := V^{\perp}$, czyli zbiór wektorów prostopadłych do wszystkich wektorów przestrzeni. Iloczyn skalarny nazywamy ***niezdegenerowanym***, gdy $\operatorname{Ker} g = \{0\}$.

Jeśli W jest podprzestrzenią przestrzeni V, to zacieśnieniem g_W metryki g do podprzestrzeni W nazywamy odwzorowanie

$$g_W : W \times W \mapsto \mathbb{K} , \quad g_W(x, y) = g(x, y) .$$

Podprzestrzeń W nazywamy ***niezdegenerowaną (względem metryki*** g***)***, gdy zacieśnienie g_W jest niezdegenerowaną metryką na W.

Wnioskujemy z twierdzenia 3, §12, że dla każdej metryki g istnieje podprzestrzeń W taka, że $V = W \oplus \operatorname{Ker} g$.

Twierdzenie 1. *Jeśli $V = W \oplus \operatorname{Ker} g$, to W jest niezdegenerowana względem metryki g.*

Dowód. Załóżmy wbrew tezie, że istnieje niezerowy wektor $x \in W$ taki, że $g(x, w) = 0$ dla każdego $w \in W$. Wtedy również $g(x, y) = 0$ dla każdego $y \in V$, gdyż $y = u + z$, $u \in W$, $z \in \operatorname{Ker} g$. Stąd $x \in \operatorname{Ker} g$, co przeczy założeniu. □

3 Macierz metryki w bazie

W dwóch kolejnych punktach ograniczmy rozważania do przestrzeni skończenie wymiarowych. ***Macierzą metryki*** g ***w bazie*** $(e_1, \ldots, e_n)$ (lub ***macierzą Grama bazy*** $(e_1, \ldots, e_n)$) nazywamy macierz

$$\mathbf{g} = (g_{ij}) , \quad g_{ij} = g(e_i, e_j) .$$

Twierdzenie 2. *Niech w bazie $(e_1, \ldots, e_n)$ współrzędne wektorów x i y dane będą kolumnami $\mathbf{x}$ i $\mathbf{y}$, a macierz metryki g niech będzie równa $\mathbf{g}$. Wtedy*

$$g(x, y) = x^i \, g_{ij} \, y^j = \mathbf{x}^{\mathrm{T}} \mathbf{g} \, \mathbf{y} \qquad \textit{w przypadku biliniowym,}$$
$$g(x, y) = \overline{x^i} \, g_{ij} \, y^j = \mathbf{x}^{\dagger} \mathbf{g} \, \mathbf{y} \qquad \textit{w przypadku półtoraliniowym.}$$

Jeśli $\mathbf{g}'$ *jest macierzą metryki* g *w bazie* $(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)\boldsymbol{\beta}$*, to*

$$\mathbf{g}' = \boldsymbol{\beta}^{\mathrm{T}} \mathbf{g}\, \boldsymbol{\beta} \qquad \textit{w przypadku biliniowym,}$$
$$\mathbf{g}' = \boldsymbol{\beta}^{\dagger} \mathbf{g}\, \boldsymbol{\beta} \qquad \textit{w przypadku półtoraliniowym.}$$

Dowód. Dla przypadku półtoraliniowego: $g(x, y) = g(x^i e_i, y^j e_j) = \overline{x^i} g(e_i, e_j) y^j$. Podstawiając tu $x = e'_k$, $y = e'_l$ dostajemy $g'_{kl} = \overline{\beta^i{}_k} g_{ij} \beta^j{}_l$. Podobnie dla przypadku biliniowego. □

Forma metryczna, podobnie jak operator liniowy, jest w dowolnej bazie w pełni scharakteryzowana swoją macierzą, jednak sens algebraiczny tej macierzy jest różny w obu przypadkach. Zwróćmy uwagę na różnicę w sposobie transformowania się macierzy metryki i macierzy operatora przy zmianie bazy. W szczególności, dla metryki nie istnieje niezależne od bazy pojęcie wyznacznika. Fakt znikania (bądź nieznikania) wyznacznika macierzy metryki ma jednak sens inwariantny.

Twierdzenie 3. *Wektor* x *należy do jądra metryki* g *wtedy, i tylko wtedy, gdy w dowolnej bazie* $\mathbf{g}\mathbf{x} = 0$*. Metryka* g *jest więc niezdegenerowana wtedy, i tylko wtedy, gdy* $\det \mathbf{g} \neq 0$.

Dowód. Zgodnie z poprzednim twierdzeniem wektor x należy do jądra wtedy, i tylko wtedy, gdy $\mathbf{y}^{\mathrm{T}} \mathbf{g}\, \mathbf{x} = 0$ ($\mathbf{y}^{\dagger} \mathbf{g}\, \mathbf{x} = 0$) dla każdego $\mathbf{y}$, co jest równoważne równaniu $\mathbf{g}\mathbf{x} = 0$. Ten układ jednorodny ma niezerowe rozwiązania wtedy, i tylko wtedy, gdy $\det \mathbf{g} = 0$, skąd wynika teza. □

4 Ortogonalne dopełnienie podprzestrzeni

Twierdzenie 3, §12, pokazało, że każdą podprzestrzeń można uzupełnić liniowo niezależną podprzestrzenią do pełnej przestrzeni. Następujące twierdzenie daje wyróżniony rozkład tego typu w obecności metryki.

Twierdzenie 4. *Niech* V *będzie skończenie wymiarową przestrzenią, a* W *jej niezdegenerowaną względem metryki* g *podprzestrzenią. Wtedy*

$$V = W \oplus W^{\perp}\,.$$

Macierz metryki g *w bazie* $(e_1, \ldots, e_n)$ *takiej, że* $(e_1, \ldots, e_m)$ *jest bazą* W*, a* $(e_{m+1}, \ldots, e_n)$ *jest bazą* $W^{\perp}$*, ma postać*

$$\mathbf{g} = \begin{pmatrix} \mathbf{g}_W & 0 \\ 0 & \mathbf{g}_{W^{\perp}} \end{pmatrix}.$$

Jeśli ponadto metryka g *jest niezdegenerowana na* V*, to podprzestrzeń* $W^{\perp}$ *jest również niezdegenerowana.*

Dowód. Pokażemy, że każdy wektor $x \in V$ można rozłożyć jednoznacznie na sumę wektorów z W i $W^\perp$. Niech $(e_1, \dots, e_m)$ będzie dowolną bazą przestrzeni W. Dla danego wektora $x \in V$ szukamy takiego wektora $x_W \in W$, aby $x - x_W$ był prostopadły do każdego wektora z W. Wystarczy zażądać, aby $g(e_i, x - x_W) = 0$ dla $i = 1, \dots, m$ – z liniowości lub antyliniowości w lewym argumencie mamy wtedy $g(y, x - x_W) = 0$ dla każdego wektora $y \in W$. Wstawiając do warunku rozkład x_W w bazie, $x_W = \sum_{j=1}^{m} \alpha^j e_j$, dostajemy układ równań liniowych na niewiadome α^j:

$$\sum_{j=1}^{m} g(e_i, e_j)\alpha^j = g(e_i, x)\,, \quad i = 1, \dots, m\,.$$

Macierz główna tego układu jest macierzą zacieśnienia g_W metryki g do podprzestrzeni W, w bazie $(e_1, \dots, e_m)$. Ponieważ podprzestrzeń W jest niezdegenerowana, to $\det \mathbf{g}_W \neq 0$. Układ równań, jako układ Cramera, ma więc jednoznaczne rozwiązanie. Szukany rozkład ma postać $x = x_W + (x - x_W)$, i jest jednoznaczny.

Stwierdzenie o postaci macierzy metryki w bazie jest bezpośrednią konsekwencją ortogonalności podprzestrzeni W i $W^\perp$. Jeśli g jest niezdegenerowana na V, to $\det \mathbf{g} \neq 0$. Ponieważ $\det \mathbf{g} = \det \mathbf{g}_W \det \mathbf{g}_{W^\perp}$, to mamy $\det \mathbf{g}_{W^\perp} \neq 0$. □

5 Warunki symetrii

Określimy trzy klasy iloczynów skalarnych. Każda klasa zdefiniowana jest dodatkowym warunkiem symetrii, gwarantującym symetryczność relacji ortogonalności.

(i) ***Metryką symetryczną*** nazywamy formę biliniową, spełniającą warunek

$$g(x, y) = g(y, x)$$

dla każdej pary wektorów. Wstawiając za x i y dowolne wektory bazowe widzimy, że macierz tej metryki w dowolnej bazie jest symetryczna. Przestrzeń rzeczywistą wyposażoną w taką metrykę nazywamy ***przestrzenią ortogonalną***. Szczególne przypadki tych przestrzeni dostarczają w fizyce modeli przestrzeni lub czasoprzestrzeni (przy zaniedbaniu grawitacji).

(ii) ***Metryką hermitowską*** nazywamy półtoraliniową formę na zespolonej przestrzeni wektorowej, spełniającą warunek

$$\overline{g(x, y)} = g(y, x)$$

dla każdej pary wektorów. Jej macierz w dowolnej bazie jest hermitowska. Przestrzeń wyposażoną w taką metrykę nazywamy ***przestrzenią hermitowską***. Szczególny przypadek tej przestrzeni ma zastosowanie dla opisu prostych układów kwantowych.

(iii) ***Metryką symplektyczną*** nazywamy formę biliniową na przestrzeni wektorowej nad ciałem $\mathbb{K}$ o charakterystyce różnej od 2, spełniającą warunek

$$g(x, y) = -g(y, x)$$

dla każdej pary wektorów. Jej macierz w dowolnej bazie jest antysymetryczna. Przestrzeń wyposażoną w taką metrykę nazywamy ***przestrzenią symplektyczną***. Znajduje ona zastosowanie do opisu układów mechanicznych w fizyce klasycznej.

Można pokazać, że te trzy klasy wyczerpują iloczyny skalarne spełniające warunki definicyjne (i) – (iii) z punktu 1 (w przypadku półtoraliniowym – z dokładnością do czynnika o module jeden). Zakładamy odtąd, że metryka spełnia jeden z trzech powyżej określonych warunków symetrii, a wykazanie twierdzenia ujętego w poprzednim zdaniu odkładamy do przykładu (iii), p. 7, §18.

6 Postać kanoniczna formy symplektycznej

Pokazujemy tutaj, jak w przestrzeni symplektycznej o skończonym wymiarze wybrać bazę, aby macierz metryki przyjęła szczególnie prostą postać. Rozwiązanie analogicznego problemu dla przestrzeni ortogonalnych i hermitowskich odkładamy do następnego paragrafu.

Twierdzenie 5. *W przestrzeni V z iloczynem symplektycznym g (nad ciałem $\mathbb{K}$,* char $\mathbb{K} \neq 2$*) istnieje baza uporządkowana, w której macierz formy g ma postać:*

$$\begin{pmatrix} 0 & \mathbf{1}_m & 0 \\ -\mathbf{1}_m & 0 & 0 \\ 0 & 0 & 0_r \end{pmatrix}.$$

Ostatnich r wektorów w bazie stanowi bazę jądra metryki, więc liczby r i m nie zależą od wyboru bazy, w której metryka przyjmuje postać tego typu.

Dowód. Istnienia bazy, o której mówi twierdzenie, dowodzimy przez indukcję względem wymiaru przestrzeni V. Szukamy w przestrzeni V wektorów x i y, dla których $g(x, y) \neq 0$. Jeśli takich wektorów nie ma, to metryka jest zerowa, więc teza jest spełniona (z $m = 0$). Jeśli takie wektory istnieją, to mnożąc jeden z nich przez liczbę uzyskujemy parę wektorów u, v, dla których $g(u, v) = 1$. Powłoka liniowa $L(u, v)$ jest niezdegenerowaną podprzestrzenią – w bazie (u, v) metryka zacieśniona do tej podprzestrzeni ma macierz $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Na podstawie twierdzenia 4 mamy stąd $V = L(u,v) \oplus L(u,v)^{\perp}$. Przestrzeń $L(u,v)^{\perp}$ ma wymiar o dwa mniejszy od wymiaru V, i zakładamy, że istnieje dla niej baza $(e_1, \ldots, e_{2m'+r})$, o której mówi twierdzenie. Dopisujemy w tym ciągu wektor u po wektorze $e_{m'}$, a wektor v – po wektorze $e_{2m'}$. Dostajemy w wyniku bazę przestrzeni V, dla której spełniona jest teza z $m = m' + 1$. W szczególności przypadek $L(u,v)^{\perp} = \{0\}$ daje tezę dla $\dim V = 2$. Tutaj indukcja posuwa się co 2 (od wymiaru V równego n do $n+2$), więc dla jej zapoczątkowania trzeba sprawdzić tezę także dla $\dim V = 1$. Ale każda macierz 1×1 antysymetryczna jest macierzą zerową, więc teza jest spełniona także w tym przypadku.

Jeśli $(e_1, \ldots, e_n)$ jest dowolną bazą, w której metryka ma taką postać jak w tezie, to rozwiązując równanie $\mathbf{g}\mathbf{x} = 0$ widzimy, że ostatnich r wektorów bazowych jest bazą jądra, więc r jest niezmiennikiem metryki (nie zależy od wyboru bazy, w której ma ona taką postać). Stąd również $m = \frac{1}{2}(\dim V - r)$ jest niezmiennikiem. □

Jeśli w danej bazie metryka ma macierz określoną udowodnionym twierdzeniem, to mówimy, że przyjmuje ona w niej ***postać kanoniczną***. Każdą taką bazę nazywamy ***bazą symplektyczną***.

Otrzymany rezultat można ująć w języku macierzowym. W tym celu załóżmy, że $(e_1, \ldots, e_n)$ jest dowolną bazą przestrzeni V i określmy w tej przestrzeni formę symplektyczną g zadając jej macierz w tej bazie jako dowolnie wybraną macierz antysymetryczną $\mathbf{g}$. Niech $(e'_1, \ldots, e'_n)$ będzie bazą symplektyczną tej formy i oznaczmy macierz przejścia z bazy nieprimowanej do primowanej przez $\boldsymbol{\beta}$. Twierdzenie mówi nam, że

$$\boldsymbol{\beta}^{\mathrm{T}} \mathbf{g}\, \boldsymbol{\beta} = \begin{pmatrix} 0 & \mathbf{1}_m & 0 \\ -\mathbf{1}_m & 0 & 0 \\ 0 & 0 & 0_r \end{pmatrix}, \quad \text{więc} \quad \mathbf{g} = (\boldsymbol{\beta}^{-1})^{\mathrm{T}} \begin{pmatrix} 0 & \mathbf{1}_m & 0 \\ -\mathbf{1}_m & 0 & 0 \\ 0 & 0 & 0_r \end{pmatrix} \boldsymbol{\beta}^{-1}.$$

Stąd każda antysymetryczna macierz $\mathbf{g}$ ma przedstawienie dane drugim z tych wzorów, co jest wynikiem zapowiedzianym, i użytym, przy końcu punktu 9, §7.

7 Przykłady

(i) Iloczyny w przestrzeniach $\mathbb{K}^n$

Niech $\mathbf{g}$ będzie macierzą antysymetryczną $n \times n$ o elementach w ciele $\mathbb{K}$ o charakterystyce różnej od 2 lub macierzą symetryczną o elementach w ciele $\mathbb{R}$. Wtedy przepis

$$g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{T}} \mathbf{g}\, \mathbf{y}$$

określa metrykę symplektyczną lub, odpowiednio, symetryczną w przestrzeni $\mathbb{K}^n$. Podobnie, jeśli $\mathbf{g}$ jest macierzą zespoloną hermitowską, to przepis

$$g(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\dagger \mathbf{g}\, \mathbf{y}$$

określa metrykę hermitowską w przestrzeni $\mathbb{C}^n$.

(ii) Iloczyn określony śladem macierzowym
Niech $\mathbf{H}$ będzie macierzą hermitowską. Przepis

$$g(\mathbf{A}, \mathbf{B}) = \mathrm{Tr}(\mathbf{A}^\dagger \mathbf{H}\, \mathbf{B})$$

określa iloczyn hermitowski w przestrzeni macierzy zespolonych $n \times n$.

(iii) Iloczyny hermitowskie w zespolonych przestrzeniach funkcyjnych
Niech $\mu : \mathbb{R} \mapsto \mathbb{R}$ będzie funkcją kawałkami ciągłą (ciągłą poza skończoną liczbą punktów nieciągłości), a $V \subseteq \mathcal{C}^0$ pewną podprzestrzenią przestrzeni zespolonych funkcji ciągłych $\mathcal{C}^0$ o tej własności, że dla każdej funkcji $f \in V$ liczba $\int_{\mathbb{R}} |f(x)|^2 |\mu(x)|\, dx$ jest skończona. Wtedy przepis

$$g(f_1, f_2) = \int_{\mathbb{R}} \overline{f_1(x)} f_2(x) \mu(x)\, dx$$

zadaje hermitowski iloczyn skalarny na V. Założenia ciągłości można osłabić dopuszczając funkcje mierzalne.
Wskazówka: Dla wykazania bezwzględnej zbieżności całki określającej ten iloczyn skorzystać z nierówności $2|f(x)g(x)| \leq |f(x)|^2 + |g(x)|^2$.

(iv) Iloczyny symetryczne w rzeczywistych przestrzeniach funkcyjnych
Podobnie jak w poprzednim przykładzie, dla podprzestrzeni V rzeczywistych funkcji ciągłych $\mathcal{C}^0_{\mathbb{R}}$, spełniających warunek z poprzedniego punktu, przepis

$$g(f_1, f_2) = \int_{\mathbb{R}} f_1(x) f_2(x) \mu(x)\, dx$$

zadaje symetryczny iloczyn skalarny na V.

(v) Funkcje całkowalne w kwadracie
Wybierzmy w przykładzie (iii) funkcję μ oraz niech V będzie zbiorem wszystkich ciągłych (lub mierzalnych) funkcji zespolonych na $\mathbb{R}$ takich, że

$$\int_{\mathbb{R}} |f(x)|^2 |\mu(x)|\, dx < \infty\,.$$

Korzystając z nierówności $2|f(x)g(x)| \leq |f(x)|^2 + |g(x)|^2$ pokazuje się, że jeśli f i g spełniają powyższy warunek, to spełnia go również dowolna kombinacja liniowa tych funkcji. Stąd V jest istotnie przestrzenią wektorową i przykład (iii) zadaje na niej iloczyn hermitowski.

(vi) Podprzestrzenie zdegenerowane
Metryki niezdegenerowane posiadają, w ogólności, zdegenerowane podprzestrzenie – nie należy mylić tych pojęć. Na przykład:
(a) jeśli metryka symplektyczna ma w bazie (e_1, e_2, e_3, e_4) macierz $\begin{pmatrix} 0 & \mathbf{1}_2 \\ -\mathbf{1}_2 & 0 \end{pmatrix}$, to zawężenie metryki do podprzestrzeni $L(e_1, e_2)$ jest zerowe;
(b) jeśli metryka symetryczna ma w bazie (e_0, e_1, e_2, e_3) macierz $\begin{pmatrix} 1 & 0 \\ 0 & -\mathbf{1}_3 \end{pmatrix}$, to macierz jej zawężenia do podprzestrzeni o bazie $(e_0 + e_3, e_1, e_2)$ ma w tej bazie macierz $\begin{pmatrix} 0 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix}$, więc ta podprzestrzeń jest zdegenerowana.

(vii) Algorytm sprowadzania formy symplektycznej do postaci kanonicznej
Niech g będzie metryką symplektyczną w przestrzeni V, z macierzą $\mathbf{g}$ w bazie $(e_1, \ldots, e_n)$. Szukamy macierzy przejścia do bazy, w której metryka przyjmuje postać kanoniczną. Sformułujemy praktyczny algorytm oparty na rekurencyjnej procedurze zastosowanej w dowodzie twierdzenia 5.

Załóżmy, że $g_{12} = 1$; spełnienie tego wstępnego warunku dla niezerowej metryki można zawsze uzyskać przez zmianę kolejności wektorów bazowych i pomnożenie jednego z nich przez niezerową liczbę. Teraz macierz metryki ma postać blokową

$$\mathbf{g} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ -\mathbf{F}^{\mathrm{T}} & \mathbf{G} \end{pmatrix}, \quad \text{gdzie} \quad \mathbf{E} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Szukamy takiej macierzy $\tilde{\boldsymbol{\beta}}$, aby macierz $\tilde{\mathbf{g}} = \tilde{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{g} \tilde{\boldsymbol{\beta}}$ była blokowo diagonalna (miała blok zerowy zamiast $\mathbf{F}$). Postulujemy, by wektory e_1, e_2 nie ulegały zmianie, więc $\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \mathbf{1} & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix}$. Wyliczamy blok pozadiagonalny w macierzy $\tilde{\mathbf{g}}$ i stwierdzamy, że warunek jego znikania najłatwiej osiągnąć kładąc $\mathbf{C} = \mathbf{1}$, $\mathbf{B} = \mathbf{EF}$. Dostajemy

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \mathbf{1} & \mathbf{EF} \\ 0 & \mathbf{1} \end{pmatrix}, \quad \tilde{\mathbf{g}} = \begin{pmatrix} \mathbf{E} & 0 \\ 0 & \tilde{\mathbf{G}} \end{pmatrix}, \quad \text{gdzie} \quad \tilde{\mathbf{G}} = \mathbf{G} - \mathbf{F}^{\mathrm{T}} \mathbf{EF}.$$

Metryka g ma macierz $\tilde{\mathbf{g}}$ w bazie $\begin{pmatrix} \tilde{e}_1 & \ldots & \tilde{e}_n \end{pmatrix} = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} \tilde{\boldsymbol{\beta}}$. Powtarzamy teraz procedurę na podprzestrzeni $L(\tilde{e}_3, \ldots, \tilde{e}_n)$ i na kolejno otrzymanych w ten sposób podprzestrzeniach aż do wyczerpania niezerowych wyrazów macierzy metryki. Uzyskamy w wyniku macierz metryki w bazie $(e'_1, \ldots, e'_n)$ w postaci blokowo diagonalnej, o blokach na diagonali $(\mathbf{E}, \ldots, \mathbf{E}, 0_r)$. Jeśli liczba bloków typu $\mathbf{E}$ wynosi m, to wystarczy teraz dokonać następującej zmiany kolejności pierwszych $2m$ wektorów bazowych:

$$(e'_1, \ldots, e'_{2m}) \longrightarrow (e'_1, e'_3, \ldots, e'_{2m-1}, e'_2, e'_4, \ldots, e'_{2m}),$$

z wektorami jądra następującymi po nich bez zmiany. W nowej bazie metryka ma postać kanoniczną. Składając wszystkie transformacje pośrednie dostajemy macierz przejścia do bazy kanonicznej.

(viii) Sprowadzenie metryki symplektycznej do postaci kanonicznej
Metryka symplektyczna ma w bazie $(e_1, \dots, e_5)$ macierz

$$\mathbf{g} = \begin{pmatrix} 0 & 1 & 0 & 2 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 0 \\ -2 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Postępując jak w ogólnym opisie w poprzednim przykładzie w pierwszym kroku uzyskujemy

$$\mathbf{EF} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 1 \end{pmatrix} , \quad \tilde{\mathbf{G}} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} .$$

Stąd w bazie

$$\tilde{e}_1 = e_1\,,\ \tilde{e}_2 = e_2\,,\ \tilde{e}_3 = e_1 + e_3\,,\ \tilde{e}_4 = -2e_2 + e_4\,,\ \tilde{e}_5 = e_2 + e_5$$
$$\text{mamy} \quad \tilde{g}_{12} = -\tilde{g}_{21} = \tilde{g}_{53} = -\tilde{g}_{35} = 1\,,$$

a pozostałe elementy są równe zeru. W uporządkowanej bazie $(\tilde{e}_1, \tilde{e}_2, \tilde{e}_5, \tilde{e}_3, \tilde{e}_4)$ macierz metryki jest blokowo diagonalna z blokami $(\mathbf{E}, \mathbf{E}, 0)$ na diagonali, a w bazie $(\tilde{e}_1, \tilde{e}_5, \tilde{e}_2, \tilde{e}_3, \tilde{e}_4)$ uzyskuje postać kanoniczną. Oznaczając tę ostatnią bazę przez $\begin{pmatrix} e'_1 & \dots & e'_5 \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_5 \end{pmatrix} \boldsymbol{\beta}$ dostajemy ostatecznie

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} , \quad \mathbf{g}' = \begin{pmatrix} 0 & \mathbf{1}_2 & 0 \\ -\mathbf{1}_2 & 0 & 0 \\ 0 & 0 & 0_1 \end{pmatrix} .$$

§15 Przestrzenie ortogonalne i hermitowskie

1 Formy kwadratowe. Normalizacja wektorów

W tym paragrafie dyskutujemy geometrię przestrzeni ortogonalnych i hermitowskich, które mają wiele punktów wspólnych. Dla metryk określających ich geometrię istotne znaczenie ma następujący wynik.

Twierdzenie 1.
(i) Forma biliniowa g jest symetryczna wtedy, i tylko wtedy, gdy

$$g(x,y) = \frac{1}{4}\big[g(x+y,x+y) - g(x-y,x-y)\big]\,.$$

(ii) Dla każdej formy półtoraliniowej g zachodzi

$$g(x,y) = \frac{1}{4}\big[g(x+y,x+y) - g(x-y,x-y) \\ - ig(x+iy,x+iy) + ig(x-iy,x-iy)\big]\,.$$

(iii) Forma półtoraliniowa g jest hermitowska wtedy, i tylko wtedy, gdy dla każdego wektora x jest

$$\overline{g(x,x)} = g(x,x)\,.$$

Dowód. (i) Jeśli g jest biliniowa, to prawa strona równości jest równa $\frac{1}{2}g(x,y) + \frac{1}{2}g(y,x)$, skąd natychmiast wynika teza.

(ii) Stosując własności półtoraliniowości do wyrażenia po prawej stronie uzyskuje się tożsamość.

(iii) Jeśli forma jest hermitowska, to równość oczywista. Odwrotnie, zakładając warunek i korzystając z punktu (ii) dostaje się własność hermitowskości: korzystamy z faktu, że dla formy półtoraliniowej $g(x\pm y, x\pm y) = g(y\pm x, y\pm x)$, $g(x\pm iy, x\pm iy) = g(y\mp ix, y\mp ix)$. □

Kierując się wynikiem tego twierdzenia wprowadza się następujące pojęcie: ***formą kwadratową metryki symetrycznej lub hermitowskiej*** g nazywamy odwzorowanie

$$V \ni x \mapsto g(x,x) \in \mathbb{R}\,.$$

Twierdzenie mówi, że metryki takie są wyznaczone wartościami swoich form kwadratowych.

Każdy wektor u spełniający warunek $g(u,u) = \pm 1$ nazywamy ***wektorem unormowanym*** (lub ***jednostkowym***). Z własności iloczynów rzeczywistych symetrycznych i zespolonych hermitowskich wynika, że $g(\lambda x, \lambda x) = |\lambda|^2 g(x,x)$. Jeśli $g(x,x) \neq 0$, to dla $x' = |g(x,x)|^{-1/2}x$ mamy

$$g(x',x') = |g(x,x)|^{-1} g(x,x) = \pm 1\,.$$

Mówimy, że każdy wektor x, dla którego $g(x,x) \neq 0$, można ***unormować*** mnożąc go przez liczbę rzeczywistą dodatnią.

2 Postać kanoniczna formy symetrycznej i hermitowskiej

Twierdzenie 2. *W przestrzeni ortogonalnej oraz w przestrzeni hermitowskiej, o wymiarze skończonym, istnieje baza uporządkowana, w której macierz formy ma postać:*

$$\begin{pmatrix} \mathbf{1}_p & 0 & 0 \\ 0 & -\mathbf{1}_q & 0 \\ 0 & 0 & 0_r \end{pmatrix} ;$$

Liczby p, q i r nie zależą od wyboru bazy, w której metryka przyjmuje postać tego typu – są ***niezmiennikami metryki***. *Ostatnich r wektorów w bazie stanowi bazę jądra metryki.*

Dowód. Dowód istnienia bazy, o której mówi twierdzenie, przez indukcję względem wymiaru przestrzeni V. Szukamy wektora x, dla którego $g(x,x) \neq 0$. Jeśli takiego wektora nie ma, to na podstawie twierdzenia 1 metryka jest zerowa, i teza jest spełniona. Jeśli taki wektor istnieje, to normując go dostajemy wektor u, dla którego $g(u,u) = \pm 1$. Powłoka liniowa $L(u)$ wektora u jest więc niezdegenerowaną podprzestrzenią, i na podstawie twierdzenia 4, §14, mamy $V = L(u) \oplus L(u)^\perp$. Przestrzeń $L(u)^\perp$ ma wymiar o jeden mniejszy od wymiaru V, i zakładamy, że istnieje dla niej baza uporządkowana $(e_1, \ldots, e_{p'+q'+r})$, o której mówi twierdzenie. Jeśli $g(u,u) = 1$, to dopisujemy w tym ciągu wektor u przed wektorem e_1, jeśli $g(u,u) = -1$, to dopisujemy go po wektorze $e_{p'}$; dostajemy w wyniku bazę przestrzeni V, dla której spełniona jest teza (z $p = p'+1$, $q = q'$ lub $p = p'$, $q = q'+1$ odpowiednio). W szczególności, przypadek $L(u)^\perp = \{0\}$ daje pierwszy krok indukcji ($\dim V = 1$).

Jeśli metryka ma w danej bazie taką postać jak w tezie, to rozwiązując równanie $\mathbf{g}\mathbf{x} = 0$ widzimy, że ostatnich r wektorów bazowych jest bazą jądra, więc r jest niezmiennikiem metryki (nie zależy od wyboru bazy, w której ma ona taką postać). Pozostaje do wykazania, że również p i q są niezmiennikami. Niech $(e_1, \ldots, e_n)$ i $(e'_1, \ldots, e'_n)$ będą dwiema bazami, w których metryka przyjmuje postać jak w tezie, z liczbami charakterystycznymi (p,q,r) i (p',q',r) odpowiednio. Jeśli $x \in L(e_1, \ldots, e_p)$,

$$\text{czyli} \quad x = \sum_{i=1}^{p} x^i e_i \,, \quad \text{to mamy} \quad g(x,x) = \sum_{i=1}^{p} |x^i|^2 > 0 \,,$$

z wyjątkiem przypadku $x = 0$ (gdy wszystkie liczby x^i znikają). Natomiast dla każdego wektora $y \in L(e'_{p'+1}, \ldots, e'_n)$, mającego więc rozkład

$$y = \sum_{i=p'+1}^{n} y'^i e'_i \,, \quad \text{jest} \quad g(y,y) = -\sum_{i=p'+1}^{p'+q'} |y^i|^2 \leq 0 \,.$$

Stąd przecięcie powłok liniowych $L(e_1, \dots, e_p)$ i $L(e'_{p'+1}, \dots, e'_n)$ zawiera tylko wektor zerowy, więc powłoki te są liniowo niezależne. Suma wymiarów tych podprzestrzeni jest więc mniejsza lub równa n, a stąd $p \leq p'$. Zamieniając w tym rozumowaniu bazy nieprimowaną i primowaną rolami dostajemy $p' \leq p$, więc $p' = p$, co pociąga też $q' = q$. □

Jeśli w danej bazie metryka ma macierz określoną udowodnionym twierdzeniem, to mówimy, że przyjmuje ona w niej ***postać kanoniczną***, a bazę nazywamy ***bazą kanoniczną***. Układ liczb (p, q, r) nazywamy ***sygnaturą*** metryki.

3 Układy ortonormalne

Każdy układ wektorów $(f_1, \dots, f_k)$, dla którego $g(f_i, f_j) = \pm\delta_{ij}$, nazywamy ***układem ortonormalnym***. Jeśli metryka jest niezdegenerowana, to ostatnie twierdzenie pokazuje, że istnieją wtedy układy ortonormalne, które są bazami. Każdą taką bazę nazywamy ***bazą ortonormalną***.

Twierdzenie 3. *Każdy układ ortonormalny jest liniowo niezależny. Jeśli metryka jest niezdegenerowana, to każdy układ ortonormalny może być uzupełniony do bazy ortonormalnej.*

Dowód. Jeśli $(f_1, \dots, f_k)$ jest układem ortonormalnym i $\displaystyle\sum_{i=1}^{k} \alpha^i f_i = 0$, to licząc iloczyn skalarny lewej strony tego równania z wektorem f_j dostajemy $\alpha^j = 0$, $j = 1, \dots, k$, więc wektory są liniowo niezależne. Podprzestrzeń $W = L(f_1, \dots, f_k)$ jest niezdegenerowana, więc $V = W \oplus W^\perp$, i jeśli metryka jest niezdegenerowana, to podprzestrzeń $W^\perp$ jest niezdegenerowana względem metryki. Wystarczy teraz wybrać w $W^\perp$ dowolną bazę ortonormalną. □

4 Sprowadzanie formy kwadratowej do sumy kwadratów

Otrzymamy w tym punkcie prosty algorytmiczny sposób sprowadzenia metryki symetrycznej lub hermitowskiej do postaci kanonicznej. Metoda ta opiera się na jednoznacznej odpowiedniości tych metryk z ich formami kwadratowymi.

Niech $(e_1, \dots, e_n)$ będzie dowolną bazą w przestrzeni z iloczynem symetrycznym lub hermitowskim. Wtedy

$$g(x, x) = \sum_i g_{ii}(x^i)^2 + 2\sum_{i<j} g_{ij} x^i x^j \quad \text{lub} \quad g(x, x) = \sum_{i,j} g_{ij} \overline{x^i} x^j \,,$$

gdzie w pierwszym przypadku wykorzystaliśmy symetrię macierzy $\mathbf{g}$. Mając dane te formy, jednoznacznie odczytuje się z nich macierz metryki $\mathbf{g}$ (znów przy uwzględnieniu symetrii w pierwszym przypadku). W szczególności, jeśli po pewnej nieosobliwej transformacji współrzędnych $\mathbf{x} = \boldsymbol{\beta}\mathbf{x}'$ forma kwadratowa przyjmie postać $g(x,x) = \sum_i g'_{ii}\,|x'^i|^2$, to w bazie $(e'_1,\dots,e'_n) = (e_1,\dots,e_n)\boldsymbol{\beta}$ macierz metryki $\mathbf{g}'$ jest diagonalna, z ciągiem liczb $(g'_{11},\dots,g'_{nn})$ na diagonali. Odpowiednio przeskalowując teraz współrzędne $\mathbf{x}'$ i zmieniając ewentualnie ich kolejność, dostaniemy nieosobliwą transformację

$$\mathbf{x} = \boldsymbol{\gamma}\mathbf{x}'' \;\text{ i bazę }\; (e''_1,\dots,e''_n) = (e_1,\dots,e_n)\boldsymbol{\gamma}\,,$$

w której metryka uzyska postać kanoniczną. Zadanie sprowadziliśmy więc do znalezienia transformacji $\mathbf{x} \to \mathbf{x}'$. Wyliczymy ją przy użyciu dwóch rekurencyjnych kroków. Zapiszemy je najpierw dla metryki rzeczywistej symetrycznej.

(i) Jeśli dla pewnego $i \in \{1,\dots,n\}$ jest $g_{ii} \neq 0$, to ewentualnie zmieniając kolejność zmiennych możemy założyć, że $g_{11} \neq 0$, i przechodzimy do kroku (ii).

Jeśli dla wszystkich $i = 1,\dots,n$ jest $g_{ii} = 0$, ale dla pewnej pary (i,j), $i \neq j$, jest $g_{ij} \neq 0$, to po ewentualnej zmianie kolejności współrzędnych możemy przyjąć, że $g_{12} \neq 0$. Podstawiamy teraz

$$x^1 = y^1 + y^2\,, \quad x^2 = y^1 - y^2\,, \quad x^i = y^i \;\text{ dla }\; i = 3,\dots,n\,,$$

co jest nieosobliwą transformacją współrzędnych. Po tym podstawieniu dostajemy

$$g(x,x) = 2g_{12}(y^1)^2 + \text{wyrazy nie zawierające } (y^1)^2\,,$$

więc w nowych współrzędnych $(y^1,\dots,y^n)$ współczynnik przy kwadracie pierwszej współrzędnej jest różny od zera, i przechodzimy do kroku (ii).

Jeśli wszystkie elementy $g_{ij} = 0$, to forma jest zerowa.

(ii) Jeśli $g_{11} \neq 0$, to wyrazy zawierające x^1 zbieramy do pełnego kwadratu:

$$\begin{aligned} g(x,x) &= g_{11}\left[(x^1)^2 + 2\sum_{i=2}^n \frac{g_{1i}}{g_{11}} x^1 x^i\right] + \sum_{i,j=2}^n g_{ij}x^i x^j \\ &= g_{11}\left[x^1 + \sum_{i=2}^n \frac{g_{1i}}{g_{11}} x^i\right]^2 + \sum_{i,j=2}^n \left[g_{ij} - \frac{g_{1i}g_{1j}}{g_{11}}\right] x^i x^j = g'_{11}(x'^1)^2 + \sum_{i,j=2}^n g'_{ij}x'^i x'^j\,, \end{aligned}$$

gdzie podstawiliśmy

$$x^1 = x'^1 - \sum_{i=2}^n \frac{g_{1i}}{g_{11}} x'^i\,, \quad x^i = x'^i \;\text{ dla }\; i = 2,\dots,n\,,$$

$$g'_{11} = g_{11}\,, \quad g'_{ij} = g_{ij} - \frac{g_{1i}g_{1j}}{g_{11}} \;\text{ dla }\; i,j = 2,\dots,n\,.$$

Transformacja $(x^1,\dots,x^n) \to (x'^1,\dots,x'^n)$ jest nieosobliwa.

Po wykonaniu powyższych kroków uzyskujemy postać formy, w której pierwsza współrzędna występuje już tylko w postaci kwadratu. Odłączamy ten wyraz i stosujemy teraz powyższy schemat do pozostałej formy kwadratowej, nie zawierającej już pierwszej zmiennej. Rekurencyjne zastosowanie tego schematu sprowadza formę kwadratową do żądanej postaci. Wyrażenie współrzędnych $\mathbf{x}$ przez nowe współrzędne uzyskuje się przez podstawianie kolejnych transformacji.

W przypadku metryki hermitowskiej zachodzą tylko nieznaczne modyfikacje opisanej procedury. W kroku (i), po znalezieniu $g_{12} \neq 0$, dla takiej samej transformacji współrzędnych jak poprzednio mamy

$$g(x,x) = 2\,\mathfrak{Re}\,g_{12}\,|y^1|^2 + \text{wyrazy nie zawierające } |y^1|^2\,.$$

Jeśli natomiast $\mathfrak{Re}\,g_{12} = 0$, to wystarczy nieco zmodyfikować transformację współrzędnych:

$$x^1 = y^1 + y^2\,, \quad x^2 = -iy^1 + iy^2\,, \quad x^i = y^i \ \text{ dla } \ i = 3, \ldots, n\,,$$

aby dostać

$$g(x,x) = 2\,\mathfrak{Im}\,g_{12}\,|y^1|^2 + \text{wyrazy nie zawierające } |y^1|^2\,.$$

W kroku (ii) mamy

$$\begin{aligned} g(x,x) &= g_{11}\left[\overline{x^1}\,x^1 + \overline{x^1}\sum_{i=2}^n \frac{g_{1i}}{g_{11}}x^i + x^1\,\overline{\sum_{i=2}^n \frac{g_{1i}}{g_{11}}x^i}\right] + \sum_{i,j=2}^n g_{ij}\overline{x^i}\,x^j \\ &= g_{11}\left|x^1 + \sum_{i=2}^n \frac{g_{1i}}{g_{11}}x^i\right|^2 + \sum_{i,j=2}^n \left[g_{ij} - \frac{\overline{g_{1i}}\,g_{1j}}{g_{11}}\right]\overline{x^i}\,x^j\,. \end{aligned}$$

5 Przykłady

(i) Diagonalizacja metryki symetrycznej metodą form kwadratowych
W bazie (e_1, e_2, e_3, e_4) rzeczywistej przestrzeni V macierz symetrycznej metryki g jest równa

$$\mathbf{g} = \frac{1}{2}\begin{pmatrix} 0 & 0 & -3 & 3 \\ 0 & 2 & 2 & 4 \\ -3 & 2 & 2 & -5 \\ 3 & 4 & -5 & 8 \end{pmatrix}.$$

Szukamy bazy, w której metryka ma postać kanoniczną.

Forma kwadratowa metryki g ma postać

$$g(x,x) = (x^2)^2 + (x^3)^2 + 4(x^4)^2 - 3x^1x^3 + 3x^1x^4 + 2x^2x^3 + 4x^2x^4 - 5x^3x^4\,.$$

Zmieniamy zmienne

$$x^1 = y^2\,, \quad x^2 = y^1\,, \quad x^3 = y^3\,, \quad x^4 = y^4$$

i przekształcamy

$$\begin{aligned} g(x,x) &= (y^1)^2 + (y^3)^2 + 4(y^4)^2 + 2y^1y^3 + 4y^1y^4 - 3y^2y^3 + 3y^2y^4 - 5y^3y^4 \\ &= (y^1 + y^3 + 2y^4)^2 - 3y^2y^3 + 3y^2y^4 - 9y^3y^4 = (z^1)^2 - 3z^2z^3 + 3z^2z^4 - 9z^3z^4\,, \end{aligned}$$

gdzie dokonaliśmy transformacji

$$y^1 = z^1 - z^3 - 2z^4\,, \quad y^2 = z^2\,, \quad y^3 = z^3\,, \quad y^4 = z^4\,.$$

Następna transformacja

$$z^1 = t^1\,, \quad z^2 = t^2 + t^3\,, \quad z^3 = t^2 - t^3\,, \quad z^4 = t^4$$

sprowadza formę do postaci

$$\begin{aligned} g(x,x) &= (t^1)^2 - 3(t^2)^2 + 3(t^3)^2 - 6t^2t^4 + 12t^3t^4 \\ &= (t^1)^2 - 3(t^2 + t^4)^2 + 3(t^3)^2 + 3(t^4)^2 + 12t^3t^4 \\ &= (u^1)^2 - 3(u^2)^2 + 3(u^3)^2 + 3(u^4)^2 + 12u^3u^4 \\ &= (u^1)^2 - 3(u^2)^2 + 3(u^3 + 2u^4)^2 - 9(u^4)^2 \\ &= (v^1)^2 - 3(v^2)^2 + 3(v^3)^2 - 9(v^4)^2\,, \end{aligned}$$

gdzie podstawiliśmy kolejno

$$\begin{aligned} &t^1 = u^1\,, \quad t^2 = u^2 - u^4\,, \quad t^3 = u^3\,, \quad t^4 = u^4\,, \\ &u^1 = v^1\,, \quad u^2 = v^2\,, \quad u^3 = v^3 - 2v^4\,, \quad u^4 = v^4\,. \end{aligned}$$

Kończymy przekształcenia transformacją zmiany kolejności i przeskalowania zmiennych

$$v^1 = x'^1\,, \quad v^2 = \frac{1}{\sqrt{3}}x'^3\,, \quad v^3 = \frac{1}{\sqrt{3}}x'^2\,, \quad v^4 = \frac{1}{3}x'^4\,,$$

która sprowadza formę kwadratową do postaci kanonicznej

$$g(x,x) = (x'^1)^2 + (x'^2)^2 - (x'^3)^2 - (x'^4)^2\,.$$

Składając zmiany zmiennych dostajemy transformację współrzędnych $\mathbf{x} = \boldsymbol{\beta}\mathbf{x}'$, z której odczytujemy macierz $\boldsymbol{\beta}$, a stąd związek baz:

$$\boldsymbol{\beta} = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -1 \\ 1 & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -1 \\ 0 & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}, \quad \begin{pmatrix} e'_1 & e'_2 & e'_3 & e'_4 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 & e_4 \end{pmatrix} \boldsymbol{\beta}\,.$$

W bazie primowanej macierz formy metrycznej ma postać kanoniczną

$$\mathbf{g}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \boldsymbol{\beta}^{\mathrm{T}} \mathbf{g}\, \boldsymbol{\beta}\,.$$

(ii) Diagonalizacja metryki hermitowskiej metodą form kwadratowych
W bazie (e_1, e_2, e_3) zespolonej przestrzeni V macierz hermitowskiej metryki jest równa

$$\mathbf{g} = \begin{pmatrix} 1 & 2-i & i \\ 2+i & 4 & 1+2i \\ -i & 1-2i & -3 \end{pmatrix}\,.$$

Szukamy bazy, w której metryka ma postać kanoniczną.

Forma kwadratowa metryki g ma postać

$$\begin{aligned} g(x,x) &= |x^1|^2 + 4|x^2|^2 - 3|x^3|^2 + (2-i)\overline{x^1}x^2 + (2+i)\overline{x^2}x^1 \\ &\qquad + i\overline{x^1}x^3 - i\overline{x^3}x^1 + (1+2i)\overline{x^2}x^3 + (1-2i)\overline{x^3}x^2 \\ &= |x^1 + (2-i)x^2 + ix^3|^2 - |x^2|^2 - 4|x^3|^2 + 2\overline{x^2}x^3 + 2\overline{x^3}x^2 \\ &= |x^1 + (2-i)x^2 + ix^3|^2 - |x^2 - 2x^3|^2\,. \end{aligned}$$

Transformujemy zmienne $x'^1 = x^1 + (2-i)x^2 + ix^3$, $x'^2 = x^2 - 2x^3$, $x'^3 = x^3$. Stąd $\mathbf{x} = \boldsymbol{\beta}\mathbf{x}'$,

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & -2+i & -4+i \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} e'_1 & e'_2 & e'_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}\,.$$

W bazie primowanej macierz formy metrycznej ma postać kanoniczną

$$\mathbf{g}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{\beta}^{\dagger} \mathbf{g}\, \boldsymbol{\beta}\,.$$

6 Ortogonalizacja Grama-Schmidta

Poznamy teraz inny sposób diagonalizowania symetrycznej lub hermitowskiej formy metrycznej, przeprowadzany bezpośrednio na wektorach bazowych. Zanotujmy najpierw następującą prostą obserwację.

Lemat 4. *Niech* $(f_1, \ldots, f_n)$ *i* $(e_1, \ldots, e_n) = (f_1, \ldots, f_n)\boldsymbol{\beta}$ *będą dwiema bazami przestrzeni* V*. Macierz* $\boldsymbol{\beta}$ *jest macierzą trójkątną wtedy, i tylko wtedy, gdy* $L(f_1, \ldots, f_k) = L(e_1, \ldots, e_k)$ *dla wszystkich* $k = 1, \ldots, n$*.*

Dowód. Jeśli zachodzą równości powłok, to dla każdego i jest $e_i \in L(f_1, \ldots, f_i)$, więc $e_i = \sum_{j=1}^{i} f_j \beta^j{}_i$, czyli macierz $\boldsymbol{\beta}$ jest trójkątna. Odwrotnie, jeśli dla każdego i zachodzi ostatnia równość, to $L(e_1, \ldots, e_k) \subseteq L(f_1, \ldots, f_k)$ dla każdego k. Ale macierz odwrotna do trójkątnej jest też macierzą trójkątną, więc wtedy również $f_i = \sum_{j=1}^{i} e_j (\beta^{-1})^j{}_i$, i w konsekwencji $L(f_1, \ldots, f_k) \subseteq L(e_1, \ldots, e_k)$. Łącznie dostajemy równość powłok. □

Przypomnijmy, że równoważne ujęcie faktu niezdegenerowania podprzestrzeni o bazie $(f_1, \ldots, f_k)$, to $\det\big(g(f_i, f_j)_{i,j=1,\ldots,k}\big) \neq 0$.

Twierdzenie 5. *Niech* $(f_1, \ldots, f_n)$ *będzie taką bazą przestrzeni ortogonalnej lub hermitowskiej, że wszystkie powłoki liniowe* $L(f_1, \ldots, f_k)$*,* $k = 1, \ldots, n$*, są niezdegenerowane, i oznaczmy* $\Delta_k = \det\big(g(f_i, f_j)_{i,j \leq k}\big)$*. Wtedy istnieje dokładnie jedna macierz trójkątna* $\boldsymbol{\beta}$ *o jedynkach na diagonali, dla której baza*

$$(e_1, \ldots, e_n) = (f_1, \ldots, f_n)\boldsymbol{\beta}$$

jest ortogonalna. Baza ta może być otrzymana za pomocą rekurencyjnego przepisu

$$e_1 = f_1\,, \qquad e_{k+1} = f_{k+1} - \sum_{i=1}^{k} \frac{g(e_i, f_{k+1})}{g(e_i, e_i)} e_i\,, \quad k = 1, \ldots, n-1\,.$$

Diagonalne elementy macierzy Grama tej bazy są równe

$$g(e_1, e_1) = \Delta_1\,, \qquad g(e_j, e_j) = \frac{\Delta_j}{\Delta_{j-1}}\,, \quad j = 2, \ldots, n\,.$$

Dowód. Dowód przeprowadzimy przez indukcję względem wymiaru przestrzeni. Dla $n = 1$ prawdziwość tezy jest oczywista. Zakładamy więc, że teza jest prawdziwa dla każdej przestrzeni o wymiarze mniejszym lub równym n i chcemy ją udowodnić dla przestrzeni o wymiarze $n+1$. Niech będzie dana baza takiej przestrzeni $(f_1, \ldots, f_{n+1})$ spełniająca założenia twierdzenia. Szukamy ortogonalnej bazy $(e_1, \ldots, e_{n+1})$ takiej, że

$$e_{k+1} = f_{k+1} + \sum_{i=1}^{k} f_i \, \beta^i{}_{k+1} \,, \quad k = 0, \ldots, n \,.$$

Ale stąd w szczególności warunki te spełniać ma baza $(e_1, \ldots, e_n)$ powłoki $L(f_1, \ldots, f_n)$, a ten problem, z założenia indukcyjnego, ma jednoznaczne rozwiązanie posiadające wszystkie własności podane w tezie. Zadanie polega więc jedynie na znalezieniu wektora e_{n+1} ortogonalnego do wektorów $(e_1, \ldots, e_n)$ i wykazaniu, że stwierdzenia tezy pozostają w mocy. Równoważnym sposobem przedstawienia tego wektora jest teraz

$$e_{n+1} = f_{n+1} + \sum_{i=1}^{n} \gamma^i e_i \,,$$

i jeśli ten problem ma jednoznaczne rozwiązanie, to również rozkład wektora w bazie $(f_1, \ldots, f_{n+1})$ będzie jednoznaczny. Narzucamy więc warunki ortogonalności $g(e_k, e_{n+1}) = 0$, $k = 1, \ldots, n$. Korzystając z ortogonalności układu $(e_1, \ldots, e_n)$ i liniowości metryki w prawym argumencie dostajemy stąd

$$0 = g(e_k, f_{n+1}) + \gamma^k g(e_k, e_k) \,, \quad k = 1, \ldots, n \,.$$

Wyliczenie współczynnika γ^k i wstawienie do rozkładu wektora e_{n+1} daje rozszerzenie wzoru rekurencyjnego podanego w tezie do wartości $k = n$.

Pozostaje do wykazania wzór na element $g(e_{n+1}, e_{n+1})$. Bazy $(e_1, \ldots, e_{n+1})$ i $(f_1, \ldots, f_{n+1})$ wiążą się macierzą trójkątną o jedynkach na diagonali, której wyznacznik jest więc równy jeden. Stąd wyznaczniki macierzy Grama tych dwóch baz są równe, czyli $\Delta_{n+1} = g(e_1, e_1) \ldots g(e_{n+1}, e_{n+1})$. Z założenia indukcyjnego mamy więc $\Delta_{n+1} = \Delta_n \, g(e_{n+1}, e_{n+1})$, co kończy dowód. □

Zwróćmy uwagę, że jeśli V jest rzeczywistą przestrzenią, to ortogonalizacja jej dowolnej bazy daje bazę o tej samej orientacji.

Ortogonalizację Grama-Schmidta można stosować również do układu wektorów, który nie stanowi bazy całej przestrzeni – stosujemy wtedy twierdzenie do $V = L(f_1, \ldots, f_n)$. Wykonalność ortogonalizacji wymaga jednak, aby kolejne podprzestrzenie były niezdegenerowane.

7 Przykłady

(i) Ortogonalizacja Grama-Schmidta w przestrzeni ortogonalnej
W bazie (f_1, f_2, f_3) przestrzeni rzeczywistej symetryczna metryka g ma macierz

$$\mathbf{g}_f = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 5 \\ 3 & 5 & 4 \end{pmatrix} .$$

Wyliczamy

$$\Delta_1 = 1 \,, \quad \Delta_2 = \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = -4 \,, \quad \Delta_3 = \det \mathbf{g}_f = 19 \,,$$

więc schemat Grama-Schmidta może być stosowany i po ortogonalizacji dostaniemy

$$g(e_1, e_1) = 1 \,, \quad g(e_2, e_2) = -4 \,, \quad g(e_3, e_3) = -\frac{19}{4} \,.$$

Stosując wzór rekurencyjny mamy

$$\begin{gathered} e_1 = f_1 \,, \quad g(e_1, f_2) = g(f_1, f_2) = 2 \,, \quad \text{więc} \quad e_2 = f_2 - 2f_1 \,, \\ g(e_1, f_3) = 3 \,, \quad g(e_2, f_3) = g(f_2, f_3) - 2g(f_1, f_3) = -1 \,, \\ \text{więc} \quad e_3 = f_3 - \frac{1}{4} e_2 - 3e_1 = f_3 - \frac{1}{4} f_2 - \frac{5}{2} f_1 \,. \end{gathered}$$

Otrzymujemy ostatecznie

$$\begin{gathered} \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \boldsymbol{\beta} \,, \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 & -2 & -\frac{5}{2} \\ 0 & 1 & -\frac{1}{4} \\ 0 & 0 & 1 \end{pmatrix} \,, \\ \mathbf{g}_e = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -\frac{19}{4} \end{pmatrix} = \boldsymbol{\beta}^{\mathrm{T}} \mathbf{g} \, \boldsymbol{\beta} \,. \end{gathered}$$

(ii) Ortogonalizacja Grama-Schmidta w przestrzeni hermitowskiej
W bazie (f_1, f_2, f_3) przestrzeni zespolonej hermitowska metryka g ma macierz

$$\mathbf{g}_f = \begin{pmatrix} 1 & i & 0 \\ -i & 2 & 1-i \\ 0 & 1+i & 1 \end{pmatrix} .$$

Wyliczamy

$$\Delta_1 = 1\,, \quad \Delta_2 = \begin{vmatrix} 1 & i \\ -i & 2 \end{vmatrix} = 1\,, \quad \Delta_3 = \det \mathbf{g}_f = -1\,,$$

więc schemat Grama-Schmidta może być stosowany i po ortogonalizacji będzie

$$g(e_1, e_1) = 1\,, \quad g(e_2, e_2) = 1\,, \quad g(e_3, e_3) = -1\,.$$

Tutaj metryka jest hermitowska, a nie symetryczna, więc trzeba zwracać uwagę na kolejność wektorów w iloczynie skalarnym. Znów stosujemy procedurę rekurencyjną, ale w nieco inny sposób, przy którym nie jest konieczne pamiętanie współczynników we wzorze rekurencyjnym. Kładziemy

$$e_1 = f_1\,, \quad e_2 = f_2 + re_1\,, \quad e_3 = f_3 + se_2 + te_1\,.$$

Z warunków ortogonalności

$$0 = g(e_1, e_2) = g(e_1, f_2) + rg(e_1, e_1)\,, \quad 0 = g(e_1, e_3) = g(e_1, f_3) + tg(e_1, e_1)$$

dostajemy $r = -i$, $t = 0$. Znając $e_2 = f_2 - if_1$ możemy wyliczyć

$$g(e_2, f_3) = g(f_2 - if_1, f_3) = g(f_2, f_3) + ig(f_1, f_3) = 1 - i\,.$$

Z warunku ortogonalności

$$0 = g(e_2, e_3) = g(e_2, f_3) + sg(e_2, e_2)$$

mamy $s = -1 + i$, więc

$$e_3 = f_3 + (-1 + i)e_2 = f_3 + (-1 + i)f_2 + (1 + i)f_1\,.$$

Otrzymujemy ostatecznie

$$\begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \boldsymbol{\beta}\,, \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 & -i & 1+i \\ 0 & 1 & -1+i \\ 0 & 0 & 1 \end{pmatrix}\,,$$

$$\mathbf{g}_e = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \boldsymbol{\beta}^\dagger \mathbf{g}\, \boldsymbol{\beta}\,.$$

(iii) Transformacja bazy
Istnieje prosty symboliczny przepis na nową bazę według schematu Grama-Schmidta. Przy oznaczeniach twierdzenia 5 i dla (g_{ij}) będącej macierzą Grama bazy (f_i) ma on postać

$$e_k = \frac{1}{\Delta_{k-1}} \begin{vmatrix} g_{11} & \dots & g_{1k} \\ \vdots & & \vdots \\ g_{k-1\,1} & \dots & g_{k-1\,k} \\ f_1 & \dots & f_k \end{vmatrix} = f_k + \frac{1}{\Delta_{k-1}} \begin{vmatrix} g_{11} & \dots & g_{1\,k-1} & g_{1k} \\ \vdots & & \vdots & \vdots \\ g_{k-1\,1} & \dots & g_{k-1\,k-1} & g_{k-1\,k} \\ f_1 & \dots & f_{k-1} & 0 \end{vmatrix}$$

dla $k = 2, \dots, n$. Po zastosowaniu symbolicznego rozwinięcia Laplace'a wyznaczników względem ostatniego wiersza otrzymamy kombinacje liniowe wektorów f_i.

Dla wykazania tych wzorów zauważamy po pierwsze, że macierz przejścia do nowej bazy spełnia warunki twierdzenia, wystarczy więc sprawdzić, czy e_k jest prostopadły do $f_1, \dots, f_{k-1}$. Ale to jest natychmiastową konsekwencją postaci pierwszego przepisu: dla iloczynu $g(f_j, e_k)$, $j < k$, w wyznaczniku w miejsce wiersza wektorów pojawi się wiersz $g_{j1} \dots g_{jk}$ i wyznacznik zniknie (ma wtedy dwa takie same wiersze).

Polecamy czytelnikowi przeliczenie poprzednich przykładów przy użyciu wykazanego wzoru.

8 Przestrzenie euklidesowe i unitarne

Metrykę rzeczywistą symetryczną lub zespoloną hermitowską nazywamy ***dodatnio określoną***, gdy

$$g(x, x) > 0 \quad \text{dla każdego niezerowego wektora } x\,.$$

Twierdzenie 6. *Niech g będzie metryką rzeczywistą symetryczną lub zespoloną hermitowską. Wtedy:*

(i) metryka g jest dodatnio określona wtedy, i tylko wtedy, gdy jej macierz w postaci kanonicznej jest macierzą jednostkową;

(ii) jeśli g jest dodatnio określona, to jest niezdegenerowana na każdej podprzestrzeni;

(iii) **(Kryterium Sylwestra)** *metryka g jest dodatnio określona wtedy, i tylko wtedy, gdy w dowolnie wybranej bazie $(f_1, \dots, f_n)$ jest*

$$\Delta_k \equiv \det\big(g(f_i, f_j)_{i,j\leq k}\big) > 0 \qquad \text{dla wszystkich} \quad k = 1, \dots, n\,.$$

Dowód. Stwierdzenie (i) jest oczywiste. Ponieważ dodatnio określona metryka jest również dodatnio określona po zacieśnieniu do dowolnej podprzestrzeni, to stwierdzenie (ii) wynika natychmiast z pierwszego. Wynika stąd także, że dla dodatnio określonej metryki ortogonalizację Grama-Schmidta można stosować do dowolnej bazy $(f_1, \ldots, f_n)$. Elementy diagonalne macierzy metryki po przeprowadzeniu ortogonalizacji są równe $(\Delta_1, \Delta_2/\Delta_1, \ldots, \Delta_n/\Delta_{n-1})$, więc $\Delta_k > 0$, $k = 1, \ldots, n$. Odwrotnie, jeśli spełnione są te nierówności, to podprzestrzenie $L(f_1, \ldots, f_k)$ są niezdegenerowane, i po przeprowadzeniu ortogonalizacji elementy diagonalne są dodatnie, więc metryka jest dodatnio określona. □

Rzeczywistą, skończenie wymiarową przestrzeń z symetrycznym, dodatnio określonym iloczynem skalarnym nazywamy ***przestrzenią euklidesową***, a skończenie wymiarową przestrzeń zespoloną z hermitowskim, dodatnio określonym iloczynem skalarnym – ***przestrzenią unitarną***. W przestrzeniach tego typu iloczyn skalarny oznacza się często

$$g(x, y) \equiv (x, y)\,.$$

Dla przestrzeni euklidesowych stosowana jest też notacja

$$g(x, y) \equiv x \cdot y\,.$$

Normą wektora x (w przypadku euklidesowym także jego ***długością***) nazywamy liczbę

$$\|x\| := \sqrt{(x, x)}\,.$$

Twierdzenie 7. *Określona powyższym przepisem norma wektora ma następujące własności*

$$\begin{aligned}
&\textit{(i)} && \|0\| = 0\,, \quad \|x\| > 0 \quad \textit{dla} \quad x \neq 0\,; \\
&\textit{(ii)} && \|\alpha x\| = |\alpha|\,\|x\|\,; \\
&\textit{(iii)} && \|x + y\| \leq \|x\| + \|y\| \qquad \text{—} \qquad \textbf{\textit{nierówność trójkąta}}\,; \\
&\textit{(iv)} && \|x - y\| \geq \big|\|x\| - \|y\|\big|\,.
\end{aligned}$$

Dowód. Pierwsze dwie własności są bezpośrednimi konsekwencjami definicji normy i dodatniej określoności metryki. Dowód trzeciej własności będzie dołączony w następnym punkcie do dowodu nierówności Schwarza, jako jej wniosek. Czwarta własność jest konsekwencją trzeciej: podstawiając w nierówności (iii) $x - y$ w miejsce x dostajemy $\|x\| - \|y\| \leq \|x - y\|$. Zamieniając x i y rolami mamy $\|y\| - \|x\| \leq \|y - x\| = \|x - y\|$, co kończy dowód. □

Ogólniej, normą na przestrzeni wektorowej nad ciałem liczb rzeczywistych lub zespolonych nazywamy każdą funkcję

$$V \ni x \mapsto \|x\| \in \mathbb{R}$$

o własnościach (i) – (iii) z ostatniego twierdzenia. Przestrzeń z określoną normą jest przestrzenią metryczną w sensie topologicznym, z metryką (topologiczną) zadaną przepisem $d(x, y) = \|x - y\|$.

9 Nierówność Schwarza. Kąt w przestrzeni euklidesowej

Podstawowe znaczenie dla geometrii przestrzeni euklidesowej i unitarnej ma następująca nierówność.

Twierdzenie 8 (Nierówność Schwarza)**.** *Dla każdych wektorów x, y w przestrzeni euklidesowej lub unitarnej zachodzi*

$$|(x, y)| \leq \|x\| \, \|y\| \,,$$

przy czym równość zachodzi wtedy, i tylko wtedy, gdy wektory x i y są liniowo zależne.

Dowód. Jeśli wektory x i y są liniowo niezależne, to są bazą dwuwymiarowej podprzestrzeni $L(x, y)$, na której metryka jest dodatnio określona (jak na każdej podprzestrzeni). Kryterium Sylwestra mówi więc, że wyznacznik macierzy Grama bazy x, y jest dodatni, czyli

$$\begin{vmatrix} (x, x) & (x, y) \\ (y, x) & (y, y) \end{vmatrix} > 0 \,,$$

co daje przypadek silnej nierówności w nierówności Schwarza.

Niech teraz wektory x i y będą liniowo zależne i niech np. $y = \lambda x$. Wtedy $|(x, y)| = |\lambda| \, \|x\|^2 = \|x\| \, \|y\|$, co kończy dowód twierdzenia.

Trzecia własność normy wektora (tw. 7) jest teraz konsekwencją nierówności Schwarza:

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

□

Z nierówności Schwarza wynika, że dla każdych niezerowych wektorów x i y jest

$$\frac{|(x, y)|}{\|x\| \, \|y\|} \leq 1 \,.$$

Pozwala to na wprowadzenie następującej definicji. ***Kątem miedzy wektorami*** x i y (niezerowymi) w przestrzeni euklidesowej nazywamy liczbę $\measuredangle(x,y) \in \langle 0,\pi\rangle$ określoną równaniem

$$\cos[\measuredangle(x,y)] = \frac{(x,y)}{\|x\|\,\|y\|}\,.$$

Wektory x i y są prostopadłe wtedy, i tylko wtedy, gdy tworzą kąt prosty, $\measuredangle(x,y) = \pi/2$. Dla dowolnych wektorów

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\|\cos\measuredangle(x,y)\,.$$

W szczególności, dla wektorów prostopadłych uzyskujemy twierdzenie Pitagorasa:

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

prawdziwe także w przestrzeni unitarnej.

10 Bazy ortonormalne w przestrzeniach unitarnych i euklidesowych

Jeśli $(e_1,\ldots,e_n)$ jest bazą ortonormalną w przestrzeni unitarnej lub euklidesowej, to macierz Grama tej bazy jest macierzą jednostkową. Stąd iloczyn skalarny wyraża się poprzez współrzędne wektorów w takiej bazie za pomocą przepisów

$$(x,y) = \mathbf{x}^\dagger\mathbf{y}\,,\quad (x,y) = \mathbf{x}^{\mathrm{T}}\mathbf{y}$$

w przestrzeni unitarnej i euklidesowej odpowiednio. Stąd widoczne jest również, że współrzędne wektora w takiej bazie można otrzymać jako

$$x^i = (e_i, x)\,.$$

Niech $(e'_1,\ldots,e'_n) = (e_1,\ldots,e_n)\boldsymbol{\beta}$ będzie inną bazą ortonormalną. Macierz Grama każdej z baz jest więc wtedy macierzą jednostkową. Przypominając sposób transformowania się macierzy Grama przy zmianie bazy (tw. 2, §14) widzimy, że w tym przypadku mamy odpowiednio

$$\boldsymbol{\beta}^\dagger\boldsymbol{\beta} = \mathbf{1}\,,\quad \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta} = \mathbf{1}\,,$$

a więc macierz $\boldsymbol{\beta}$ jest unitarna dla przestrzeni unitarnej, a ortogonalna dla euklidesowej. Transformację macierzy operatora liniowego pomiędzy dwiema bazami *ortonormalnymi* można więc zapisać jako

$$\mathbf{A}' = \boldsymbol{\beta}^\dagger\mathbf{A}\,\boldsymbol{\beta}\,,\quad \mathbf{A}' = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{A}\,\boldsymbol{\beta}$$

odpowiednio w przypadku unitarnym i euklidesowym.

11 Przykłady

(i) Przestrzeń funkcji Schwartza
Funkcje klasy Schwartza (patrz przykład (iv), p. 7, §12) spełniają założenia przykładu (iii), p. 7, §14, z funkcją $\mu = 1$. Iloczyn skalarny

$$(f_1, f_2) = \int_{\mathbb{R}} \overline{f_1(x)} f_2(x)\, dx$$

jest dodatnio określony.

(ii) Przestrzenie wielomianowe z dodatnio określonym iloczynem skalarnym
Niech $\mathcal{P}$ będzie przestrzenią zespolonych wielomianów zmiennej rzeczywistej. Każdy z poniższych przepisów zadaje na niej dodatnio określony iloczyn skalarny z klasy określonej przykładem (iii), p. 7, §14:

$$(P, Q)_L = \int_{-1}^{1} \overline{P(x)} Q(x)\, dx\,, \quad (P, Q)_H = \int_{\mathbb{R}} \overline{P(x)} Q(x) e^{-x^2}\, dx\,.$$

Ortogonalizacja Grama-Schmidta naturalnej bazy $e_k(x) = x^k$, $k = 0, 1, \ldots$, daje w każdym z tych przypadków ciąg wielomianów ortogonalnych względem odpowiedniego iloczynu skalarnego. Oznaczmy przez p_k bazę zortogonalizowaną względem $(.\,,.)_L$, a przez h_k bazę zortogonalizowaną względem $(.\,,.)_H$. Ciągi wielomianów

$$P_k = \frac{(2k)!}{2^k (k!)^2}\, p_k\,, \quad H_k = 2^k\, h_k$$

noszą odpowiednio nazwy wielomianów Legendre'a i wielomianów Hermite'a. Ponieważ $(e_k, e_l)_L = (e_k, e_l)_H = 0$, jeśli $k + l$ jest liczbą nieparzystą, to

$$P_k(-x) = (-1)^k P_k(x)\,, \qquad H_k(-x) = (-1)^k H_k(x)\,.$$

Początkowe wielomiany tych ciągów są równe

$$P_0(x) = 1\,, \quad P_1(x) = x\,, \quad P_2(x) = \frac{1}{2}(3x^2 - 1)\,, \quad P_3(x) = \frac{1}{2}(5x^3 - 3x)\,,$$
$$H_0(x) = 1\,, \quad H_1(x) = 2x\,, \quad H_2(x) = 4x^2 - 1\,, \quad H_3(x) = 8x^3 - 12x\,.$$

Do otrzymania ogólnej postaci i innych własności tych wielomianów stosuje się inne, analityczne metody.

(iii) Ortonormalizacja bazy w przestrzeni euklidesowej – rachunek w bazie ortonormalnej

W ortonormalnej bazie (e_1, e_2, e_3) przestrzeni euklidesowej dana jest inna baza (f_1, f_2, f_3) kolumnami swoich współrzędnych:

$$\mathbf{f}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{f}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{f}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Przeprowadzamy ortogonalizację Grama Schmidta bazy (f_1, f_2, f_3). Kładziemy

$$f_1' = f_1\,, \quad f_2' = f_2 + r f_1'\,, \quad f_3' = f_3 + s f_2' + t f_1'\,,$$

i nakładamy warunki ortogonalności. Dostajemy

$$0 = (f_1', f_2') = \mathbf{f}_1'^{\mathrm{T}} \mathbf{f}_2 + r \mathbf{f}_1'^{\mathrm{T}} \mathbf{f}_1' = 1 + 3r\,, \quad \text{więc} \quad \mathbf{f}_2' = \frac{2}{3} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

$$0 = (f_1', f_3') = \mathbf{f}_1'^{\mathrm{T}} \mathbf{f}_3 + t \mathbf{f}_1'^{\mathrm{T}} \mathbf{f}_1' = 1 + 3t\,,$$

$$0 = (f_2', f_3') = \mathbf{f}_2'^{\mathrm{T}} \mathbf{f}_3 + s \mathbf{f}_2'^{\mathrm{T}} \mathbf{f}_2' = \frac{4}{3}(-1 + 2s)\,, \quad \text{więc} \quad \mathbf{f}_3' = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Wektory (f_1', f_2', f_3') normujemy mnożąc je przez odwrotność norm (czynnik $2/3$ w $\mathbf{f}_2'$ można przy tym pominąć). Dostajemy bazę ortonormalną (f_1'', f_2'', f_3'') o współrzędnych

$$\mathbf{f}_1'' = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{f}_2'' = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{f}_3'' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

czyli

$$\begin{pmatrix} f_1'' & f_2'' & f_3'' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}\,, \qquad \boldsymbol{\beta} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Macierz $\boldsymbol{\beta}$ łączy dwie bazy ortonormalne, więc jest macierzą ortogonalną.

(iv) Ortonormalizacja bazy w przestrzeni unitarnej – rachunek w bazie ortonormalnej

W ortonormalnej bazie (e_1, e_2, e_3) przestrzeni unitarnej dana jest inna baza (f_1, f_2, f_3) kolumnami swoich współrzędnych:

$$\mathbf{f}_1 = \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{f}_2 = \begin{pmatrix} 2-2i \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{f}_3 = \begin{pmatrix} 0 \\ i \\ 2 \end{pmatrix}.$$

Przeprowadzamy ortogonalizację Grama Schmidta bazy (f_1, f_2, f_3), a następnie bazę normujemy. Postępowanie jest podobne jak w poprzednim przykładzie, pamiętać jedynie należy o zamianie operacji transpozycji kolumn współrzędnych na operację sprzężenia hermitowskiego. Dostaje się w ten sposób

$$\mathbf{f}'_1 = \begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{f}'_2 = \begin{pmatrix} 1-i \\ 1 \\ 1+i \end{pmatrix}, \quad \mathbf{f}'_3 = \frac{1}{5}\begin{pmatrix} -1-2i \\ -2+6i \\ 2-i \end{pmatrix}.$$

Po unormowaniu wektorów f'_1, f'_2, f'_3 dostajemy bazę ortonormalną:

$$\begin{pmatrix} f''_1 & f''_2 & f''_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}, \qquad \boldsymbol{\beta} = \begin{pmatrix} \frac{i}{\sqrt{2}} & \frac{1-i}{\sqrt{5}} & \frac{-1-2i}{5\sqrt{2}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{-2+6i}{5\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1+i}{\sqrt{5}} & \frac{2-i}{5\sqrt{2}} \end{pmatrix}.$$

Macierz $\boldsymbol{\beta}$ łączy dwie bazy ortonormalne, więc jest unitarna.

12 Przestrzeń Minkowskiego

Rzeczywistą przestrzeń czterowymiarową V z symetrycznym iloczynem skalarnym g o sygnaturze $(1, 3, 0)$, a więc o postaci kanonicznej

$$\mathbf{g} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

nazywamy ***przestrzenią Minkowskiego***. Iloczyn skalarny w tej przestrzeni nazywa się ***metryką lorentzowską*** i oznacza się często

$$g(x, y) = x \cdot y\,.$$

Każdą bazę kanoniczną tej przestrzeni (w której macierz metryki ma powyższą postać) nazywa się ***bazą Minkowskiego*** i według często używanej konwencji numeruje się (e_0, e_1, e_2, e_3).

Geometria przestrzeni Minkowskiego jest matematycznym modelem, na którym oparta jest szczególna teoria względności Einsteina (dokładniej: tym modelem jest przestrzeń afiniczna nad wektorową przestrzenią Minkowskiego, patrz przykład (v), p. 9, §21 w dalszej części książki). To fizyczne zastosowanie dyktuje terminologię używaną przez fizyków do opisu tej geometrii. Wprowadzamy następującą klasyfikację: wektor x nazywamy ***czasowym***, gdy $x \cdot x > 0$; ***przestrzennym***, gdy $x \cdot x < 0$; ***świetlnym***, gdy $x \cdot x = 0$. Wektory czasowe i wektory świetlne nazywa się łącznie ***kauzalnymi***.

Niech t będzie dowolnym unormowanym wektorem czasowym i dokonajmy rozkładu ortogonalnego $V = L(t) \oplus L(t)^\perp$. Przestrzeń $L(t)^\perp$ jest trójwymiarowa, a metryka $(-g)$ zacieśniona do tej podprzestrzeni jest metryką euklidesową: aby się o tym przekonać wystarczy uzupełnić wektor t do dowolnej bazy ortonormalnej. Dla każdego wektora x mamy $[x - (x \cdot t)\, t\,] \cdot t = 0$, więc jego rozkład ortogonalny ma postać

$$x = (x \cdot t)\, t + x_\perp\,, \quad x_\perp \in L(t)^\perp\,,$$

a stąd

$$x \cdot x = (x \cdot t)^2 - \|x_\perp\|_\perp^2\,,$$

gdzie $\|.\|_\perp$ jest normą w podprzestrzeni euklidesowej $L(t)^\perp$. Wektor x jest więc czasowy, gdy $\|x_\perp\|_\perp < |x \cdot t|$, przestrzenny, gdy $\|x_\perp\|_\perp > |x \cdot t|$, oraz świetlny, gdy $\|x_\perp\|_\perp = |x \cdot t|$. Interpretacja fizyczna tej struktury, w ramach szczególnej teorii względności, jest następująca. Dla każdego wektora czasowego t może istnieć poruszający się inercjalnie punkt materialny, z punktu widzenia którego wektor t wskazuje kierunek czasowy w czasoprzestrzeni; kierunki ściśle przestrzenne względem tego punktu leżą w podprzestrzeni $L(t)^\perp$. Wektory świetlne wskazują możliwe kierunki wysłania sygnałów świetlnych.

Rozważymy ważny przykład dokładniej ilustrujący odmienność geometrii Minkowskiego od geometrii euklidesowej. Niech x i y będą wektorami czasowymi i połóżmy $t = x/\sqrt{x \cdot x}$. Wtedy

$$y_\perp = y - \frac{x \cdot y}{x \cdot x}\, x\,,$$

więc

$$-(y_\perp \cdot y_\perp)(x \cdot x) = (x \cdot y)^2 - (x \cdot x)(y \cdot y)\,.$$

Metryka $-g$ jest dodatnio określona na $L(t)^\perp$, więc $-y_\perp \cdot y_\perp \geq 0$, a równość zachodzi wtedy, i tylko wtedy, gdy $y_\perp = 0$. Stąd mamy:

Twierdzenie 9. *Dla każdej pary wektorów czasowych x i y spełniona jest nierówność*

$$(x\cdot y)^2 \geq (x\cdot x)(y\cdot y)\,.$$

Równość w tej nierówności zachodzi wtedy, i tylko wtedy, gdy wektory x i y są liniowo zależne.

Zwróćmy uwagę, że strony otrzymanej nierówności mają postać taką, jak w nierówności Schwarza w przestrzeni euklidesowej, ale kierunek tej nierówności jest przeciwny. Wynika z niej w szczególności, że dla dowolnych wektorów czasowych x i y jest $x \cdot y \neq 0$.

Lemat 10. *Dla każdych trzech wektorów czasowych x, y, z zachodzi:*

$$(x\cdot y)(y\cdot z)(z\cdot x) > 0\,.$$

Dowód. Połóżmy $t = x/\sqrt{x\cdot x}$. Wtedy podstawiając rozkłady wektorów y i z w lewej stronie nierówności dostajemy

$$(x\cdot x)(y\cdot t)(z\cdot t)\big[y_\perp\cdot z_\perp + (y\cdot t)(z\cdot t)\big] = (x\cdot x)(y\cdot t)^2(z\cdot t)^2\left(\frac{y_\perp}{y\cdot t}\cdot\frac{z_\perp}{z\cdot t}+1\right) > 0\,,$$

gdyż wyrażenie w ostatnim nawiasie jest dodatnie na mocy nierówności Schwarza w przestrzeni $L(t)^\perp$ i czasowości wektorów y i z. □

Natychmiastowym wnioskiem z tego lematu jest następujący wynik.

Twierdzenie 11. *Relacja w zbiorze wektorów czasowych*

$$x \sim y \iff x\cdot y > 0$$

jest relacją równoważności, dzielącą ten zbiór na dwie klasy równoważności.

Dowód. Zwrotność i symetryczność relacji są oczywiste, a przechodniość wynika z ostatniego lematu. Niech x będzie wektorem czasowym. Wtedy dla dowolnego wektora czasowego y mamy $y\cdot x > 0$ lub $y \cdot x < 0$, więc klasy równoważności wektorów x oraz $-x$ wyczerpują zbiór wektorów czasowych. □

Istnienie dwóch rozłącznych klas wektorów czasowych, o których mówi twierdzenie, pozwala nadać przestrzeni Minkowskiego charakterystykę o cechach podobnych do orientacji. Mówimy, że przestrzeń ta ma nadaną ***orientację czasową***, jeśli wyróżniono jedną z dwu klas wektorów czasowych: mówimy o tych wektorach, że są skierowane w przyszłość, a o wektorach przeciwnych – skierowane w przeszłość. Mówimy dalej, że baza Minkowskiego (e_0, e_1, e_2, e_3) ma dodatnią orientację czasową, jeśli e_0 jest skierowany w przyszłość, a ujemną –

gdy jest skierowany w przeszłość. Zauważmy, że własności orientacji czasowej nie można określić w klasie wszystkich baz – baza musi mieć wyróżniony wektor czasowy.

Niech $(e_0, \ldots, e_3)$ będzie bazą ortonormalną. Wtedy wektory świetlne zadane są równaniem

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 0\,,$$

które opisuje stożek w $\mathbb{R}^4$. Stąd zbiór wektorów świetlnych nazywa się ***stożkiem świetlnym***. Wektory czasowe odpowiadają punktom we wnętrzu tego stożka (znak $>$ zamiast równości w powyższej formule). Jeśli przy tym baza ma dodatnią orientację czasową, to dla czasowych wektorów przyszłości $x^0 > 0$, a dla wektorów przeszłości $x^0 < 0$.

Najbardziej fundamentalną cechą odróżniającą przestrzeń Minkowskiego od przestrzeni euklidesowej jest jej ***nieizotropowość***: kierunki w tej przestrzeni nie są równoprawne, obecność stożka świetlnego wprowadza ich inwariantną (niezależną od wyboru bazy) klasyfikację. Przestrzeń euklidesowa przeciwnie, jest ***izotropowa***: wszystkie kierunki są równoprawne, tj. sama metryka nie dostarcza żadnej metody wprowadzenia ich klasyfikacji. Jeśli myśleć o trójwymiarowej geometrii euklidesowej jako modelu przestrzeni fizycznej w mechanice klasycznej, a o geometrii Minkowskiego jako modelu *czasoprzestrzeni* w szczególnej teorii względności, to pojawienie się takiej różnicy jest fizycznie niezbędne. Mechanika newtonowska oparta jest na podstawowym założeniu, że pusta przestrzeń nie ma wyróżnionych kierunków. Z drugiej strony, jeśli przedmiotem opisu jest przestrzeń wraz z czasem, to fizyczna różnica między czasem a przestrzenią musi znaleźć odbicie w formalizmie matematycznym.

§16 Odwzorowania liniowe przestrzeni z iloczynem skalarnym

1 Izometrie. Przestrzenie izometryczne

Niech V i $\tilde{V}$ będą przestrzeniami wektorowymi nad tym samym ciałem, a $A : V \mapsto \tilde{V}$ odwzorowaniem liniowym. Jeśli przestrzeń $\tilde{V}$ jest wyposażona w metrykę $\tilde{g}$, to łatwo sprawdzić, że odwzorowanie

$$V \times V \ni (x, y) \mapsto \tilde{g}(Ax, Ay) \in \mathbb{K}$$

jest iloczynem skalarnym na przestrzeni V. Iloczyn ten ma takie same własności symetrii, jak iloczyn $\tilde{g}$ (wspólnie z nim jest symetryczny, hermitowski, lub antysymetryczny). Niech przestrzeń V będzie wyposażona w swoją własną, niezależną od tej konstrukcji, metrykę g o tych samych własnościach symetrii. Odwzorowanie A nazywamy ***izometrią przestrzeni V w przestrzeń $\tilde{V}$***, gdy metryka generowana w powyższy sposób na V przez $\tilde{g}$ i A pokrywa się z metryką g, czyli gdy dla każdej pary wektorów x i y jest spełniona równość

$$\tilde{g}(Ax, Ay) = g(x, y)\,.$$

Wszystkie przestrzenie wektorowe nad danym ciałem liczbowym sklasyfikowaliśmy wcześniej za pomocą izomorfizmów: przestrzenie izomorficzne są do siebie podobne pod względem struktury liniowej. W przypadku porównywania przestrzeni wyposażonych w iloczyn skalarny możemy przyjąć teraz silniejsze, bardziej wymagające rozumienie podobieństwa, które obejmować będzie nie tylko strukturę liniową, ale też strukturę metryczną. Mówimy, że dwie przestrzenie z iloczynem skalarnym (V, g) i $(\tilde{V}, \tilde{g})$ nad tym samym ciałem są ***izometryczne***, gdy istnieje bijektywna izometria $A : V \mapsto \tilde{V}$ (odwzorowanie takie możemy też nazwać izometrycznym izomorfizmem).

Twierdzenie 1. *Relacja izometryczności przestrzeni wektorowych jest relacją równoważności, dzieli więc wszystkie przestrzenie z iloczynem skalarnym na klasy przestrzeni izometrycznych.*

Należy pokazać, że izomorfizm odwrotny do izomorfizmu izometrycznego jest też izometryczny, oraz że złożenie dwóch izomorfizmów izometrycznych $(V, g) \mapsto (\tilde{V}, \tilde{g}) \mapsto (\hat{V}, \hat{g})$ jest również izometrią. Pozostawiamy to zadanie jako ćwiczenie dla czytelnika.

Twierdzenie 2. *Skończenie wymiarowe przestrzenie z iloczynem skalarnym (V, g) i $(\tilde{V}, \tilde{g})$ są izometryczne wtedy, i tylko wtedy, gdy mają ten sam wymiar i metryki g i $\tilde{g}$ mają te same niezmienniki, tj. taką samą sygnaturę w przypadku przestrzeni ortogonalnych lub hermitowskich, lub taki sam wymiar jądra w przypadku przestrzeni symplektycznych.*

Dowód. Niech przestrzenie (V, g) i $(\tilde{V}, \tilde{g})$ będą izometryczne, więc istnieje izometryczny izomorfizm $A : V \mapsto \tilde{V}$. Niech metryka g przyjmuje postać kanoniczną w bazie $(e_1, \ldots, e_n)$. Wtedy metryka $\tilde{g}$ przyjmuje taką samą postać kanoniczną w bazie $(Ae_1, \ldots, Ae_n)$. Odwrotnie, niech metryki g i $\tilde{g}$ przyjmują takie same postacie kanoniczne w bazach $(e_1, \ldots, e_n)$ i $(\tilde{e}_1, \ldots, \tilde{e}_n)$ odpowiednio. Wtedy odwzorowanie liniowe zadane przez $Ae_i = \tilde{e}_i$, $i = 1, \ldots, n$, jest izometrycznym izomorfizmem. □

2 Operator sprzężony

Ograniczamy się odtąd do niezdegenerowanych form metrycznych na przestrzeniach skończenie wymiarowych i rozważać będziemy operatory liniowe na takich przestrzeniach. Zobaczymy, że obecność metryki zadaje dodatkową strukturę w zbiorze tych operatorów.

Jeśli g jest formą metryczną na V, to dla każdego ustalonego wektora $x \in V$ przepis

$$V \ni y \mapsto g(x, y) \in \mathbb{K}$$

zadaje odwzorowanie liniowe z V w $\mathbb{K}$, a więc element przestrzeni wektorowej $\mathcal{L}(V, \mathbb{K})$. Oznaczymy to odwzorowanie $g(x, .\,)$. Pytamy teraz, jak odwzorowanie to zależy od wyboru wektora x, czyli jakie własności ma przyporządkowanie $x \mapsto g(x, .\,)$.

Twierdzenie 3. *Jeśli g jest niezdegenerowaną formą metryczną na skończenie wymiarowej przestrzeni V, to odwzorowanie*

$$V \ni x \mapsto g(x, .\,) \in \mathcal{L}(V, \mathbb{K})$$

jest izomorfizmem liniowym w przypadku formy biliniowej, lub izomorfizmem antyliniowym w przypadku formy hermitowskiej.

Dowód. Jeśli g jest formą półtoraliniową, to

$$g(\alpha x + \beta y, .\,) = \overline{\alpha} g(x, .\,) + \overline{\beta} g(y, .\,)\,,$$

co jest natychmiast widoczne po wstawieniu dowolnego wektora z jako argumentu tych odwzorowań. Stąd odwzorowanie $x \mapsto g(x, .\,)$ jest antyliniowe. Podobne rozumowanie pokazuje jego liniowość w przypadku biliniowym.

Niech x będzie takim wektorem, że $g(x, .\,) = 0$. Z definicji odwzorowania $g(x, .\,)$ oznacza to, że $g(x, z) = 0$ dla każdego $z \in V$, co jest równoznaczne ze stwierdzeniem $x \in \operatorname{Ker} g$. Metryka jest niezdegenerowana, stąd $x = 0$. Wykazaliśmy, że jądro odwzorowania $x \mapsto g(x, .\,)$ składa się tylko z wektora zerowego, więc to odwzorowanie jest injektywne. Ponieważ wymiary przestrzeni V i $\mathcal{L}(V, \mathbb{K})$ są równe i skończone, to z injektywności wynika surjektywność na mocy twierdzenia 6, §10, i dyskusji w punkcie 18, §10. □

Udowodniony rezultat gra kluczową rolę w dowodzie następującego twierdzenia.

Twierdzenie 4. *Jeśli (V, g) jest skończenie wymiarową przestrzenią z niezdegenerowaną metryką, to dla każdego operatora liniowego A na przestrzeni V istnieje dokładnie jeden operator liniowy A^* na przestrzeni V spełniający dla każdej pary wektorów x i y warunek*

$$g(A^* x, y) = g(x, Ay)\,.$$

Dowód. W całym poniższym rozumowaniu operator A jest ustalony. Dla każdego wektora x przepis $y \mapsto g(x, Ay)$ zadaje odwzorowanie liniowe z V w $\mathbb{K}$, a więc element przestrzeni $\mathcal{L}(V, \mathbb{K})$; oznaczymy go $g(x, A\,.\,)$. Wiemy z poprzedniego twierdzenia, że każdy element tej przestrzeni można jednoznacznie przedstawić jako $g(z, .\,)$, dla pewnego $z \in V$. W naszym przypadku z jest wyznaczony przez wektor x, więc oznaczając tę zależność funkcyjną przez $z = A^*x$ dostajemy $g(x, A\,.\,) = g(A^*x, .\,)$. Równość ta oznacza, że spełniony jest warunek z tezy. Pozostaje do wykazania, że odwzorowanie $x \mapsto A^*x$ jest liniowe. Korzystając z definicji odwzorowania A^* oraz z (anty)liniowości metryki dostajemy ciąg równości

$$g(A^*(\alpha x + \beta y), z) = g(\alpha x + \beta y, Az) = \overline{\alpha} g(x, Az) + \overline{\beta} g(y, Az) = \\ = \overline{\alpha} g(A^*x, z) + \overline{\beta} g(A^*y, z) = g(\alpha A^*x + \beta A^*y, z)\,,$$

zatem wektor $A^*(\alpha x + \beta y) - (\alpha A^*x + \beta A^*y)$ jest prostopadły do wszystkich wektorów, więc znika na mocy niezdegenerowania metryki. To zamyka dowód liniowości A. □

Określony powyższym twierdzeniem operator A^* nazywamy ***operatorem sprzężonym (względem metryki*** g***) do operatora*** A (lub ***sprzężeniem operatora*** A).

3 Własności sprzężenia operatorowego

Operacja sprzężenia operatorowego będzie nas interesować głównie w przestrzeniach unitarnych i euklidesowych, ale jej najprostsze podstawowe własności mają taką samą postać we wszystkich przestrzeniach przez nas tutaj rozważanych. Sformułowanie ich poprzedzimy prostym twierdzeniem pomocniczym.

Lemat 5. *Niech g będzie niezdegenerowaną metryką na przestrzeni V, a A i B operatorami w V. Jeśli dla każdej pary wektorów x i y jest*

$$g(x, Ay) = g(x, By)\,, \quad \text{to} \quad A = B\,.$$

Dowód. Z założenia mamy $g(x, (A - B)y) = 0$ dla każdej pary wektorów, więc $(A - B)y \in \operatorname{Ker} g$ dla każdego y. Metryka jest niezdegenerowana, więc $Ay = By$ dla każdego y, czyli $A = B$. □

Z własności symetrii iloczynu skalarnego widać, że równoważne sformułowanie założenia to $g(Ay, x) = g(By, x)$.

Twierdzenie 6. *Niech A i B będą operatorami liniowymi na skończenie wymiarowej przestrzeni V wyposażonej w niezdegenerowaną metrykę. Operacja sprzężenia operatorowego ma następujące własności:*

(i) $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$ *antyliniowość – w przypadku półtoraliniowym,*

$(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$ *liniowość – w przypadku biliniowym;*

(ii) $(\mathrm{id})^* = \mathrm{id}\,;$

(iii) $(A^*)^* = A\,;$

(iv) $(AB)^* = B^* A^*\,.$

Dowód. (i) Dla metryki półtoraliniowej mamy ciąg równości

$$g\big((\alpha A + \beta B)^* x, y\big) = g\big(x, (\alpha A + \beta B) y\big) = \alpha g(x, Ay) + \beta g(x, By) = \\ = \alpha g(A^* x, y) + \beta g(B^* x, y) = g\big((\overline{\alpha} A^* + \overline{\beta} B^*) x, y\big)\,,$$

skąd przy użyciu lematu wynika punkt (i) (i podobnie dla metryki biliniowej).

(ii) – punkt oczywisty.

(iii) Ze względu na własności symetrii formy metrycznej (symetria, antysymetria lub hermitowskość) warunek definicyjny na operator sprzężony jest równoważny warunkowi $g(Ay, x) = g(y, A^* x)$, a to oznacza, że operatorem sprzężonym do A^* jest operator A.

(iv) Ta własność wynika, w oparciu o lemat, z ciągu równości

$$g\big((AB)^* x, y\big) = g(x, ABy) = g(A^* x, By) = g(B^* A^* x, y)\,.$$

□

Wybierzmy w przestrzeni V dowolną bazę $(e_1, \ldots, e_n)$ i niech $\mathbf{g}$, $\mathbf{A}$ i $\mathbf{A}^*$ będą odpowiednio macierzami metryki g, operatora A i operatora A^* w tej bazie. Podstawiając $x = e_i$ i $y = e_j$ w warunku $g(x, A^* y) = g(Ax, y)$ (równoważnym definicyjnemu) dostajemy $g(e_i, e_k {A^*}^k{}_j) = g(e_k A^k{}_i, e_j)$, skąd – po uwzględnieniu półtora-(bi-)liniowości – mamy

$$\mathbf{g}\,\mathbf{A}^* = \mathbf{A}^\dagger\,\mathbf{g} \quad \text{lub} \quad \mathbf{g}\,\mathbf{A}^* = \mathbf{A}^{\mathrm{T}}\mathbf{g}$$

w przypadku hermitowskim lub biliniowym odpowiednio. Stąd

$$\mathbf{A}^* = \mathbf{g}^{-1}\mathbf{A}^\dagger\,\mathbf{g} \quad \text{lub} \quad \mathbf{A}^* = \mathbf{g}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{g}\,.$$

4 Izometrie wewnętrzne

W pierwszym punkcie tego paragrafu określiliśmy ogólne pojęcie izometrii. Dla danej przestrzeni V rozważymy teraz wszystkie izometrie tej przestrzeni w siebie samą, czyli takie operatory $A : V \mapsto V$, dla których

$$g(Ax, Ay) = g(x, y)$$

dla każdej pary wektorów x i y. Nazywamy je ***izometriami wewnętrznymi***.

Twierdzenie 7. *Niech (V, g) będzie skończenie wymiarową przestrzenią z niezdegenerowaną formą metryczną, a A operatorem liniowym w przestrzeni V. Wtedy następujące warunki są równoważne:*

(i) A jest izometrią wewnętrzną;

(ii) A jest izomorfizmem i spełnia warunek

$$A^* = A^{-1}\,;$$

(iii) w dowolnie wybranej bazie

$$\mathbf{A}^\dagger \mathbf{g}\,\mathbf{A} = \mathbf{g} \quad \textit{lub} \quad \mathbf{A}^{\mathrm{T}} \mathbf{g}\,\mathbf{A} = \mathbf{g}\,,$$

w przypadku hermitowskim lub biliniowym odpowiednio.

Dowód. Równoważnym sformułowaniem definicji izometrii wewnętrznej jest warunek: $g(A^*Ax, y) = g(x, y)$ dla każdej pary wektorów, co na podstawie lematu 5 jest tym samym, co $A^*A = \mathrm{id}$. Dla operatorów w skończenie wymiarowej przestrzeni warunek ten jest równoważny stwierdzeniu (ii). Ponadto, macierzowy zapis tego warunku, $\mathbf{A}^*\mathbf{A} = \mathbf{1}$, jest równoważny warunkom (iii) (należy użyć przepisów na macierze $\mathbf{A}^*$ uzyskanych na końcu poprzedniego punktu). □

Wniosek 8. *Izometrie wewnętrzne skończenie wymiarowej przestrzeni z niezdegenerowaną metryką tworzą grupę izomorficzną z jedną z grup macierzowych: $O(p, q)$ dla metryki rzeczywistej symetrycznej, $U(p, q)$ dla metryki hermitowskiej, $Sp(2m; \mathbb{K})$ dla metryki symplektycznej.*

Dowód. Warunek izometryczności (iii) z ostatniego twierdzenia zapisany w bazie kanonicznej sprowadza się do jednego z warunków definiujących grupy macierzowe omówione w §7, p. 10. Stąd, w połączeniu z rezultatami punktu 1, §11, otrzymujemy tezę. □

Omawialiśmy w §11 związek baz przestrzeni z jej automorfizmami, ujęty w twierdzenie 2. Otrzymujemy teraz jego bardziej szczególną postać.

Twierdzenie 9. *Niech V będzie przestrzenią z niezdegenerowanym iloczynem skalarnym. Każda jej baza kanoniczna (tj. symplektyczna lub ortonormalna w zależności od rodzaju metryki) $(e_1, \ldots, e_n)$ zadaje wzajemnie jednoznaczny związek między izometriami wewnętrznymi przestrzeni V i jej bazami kanonicznymi, dany przyporządkowaniem $\beta \mapsto (\beta e_1, \ldots, \beta e_n)$.*

Dowód tego twierdzenia otrzymuje się natychmiast przez zastosowanie dowodu twierdzenia 2 do przypadku izometrii wewnętrznej.

5 Transformacje Lorentza

Własności grup izometrii wewnętrznych silnie zależą od metryki. Przypadkiem najprostszym – metryki dodatnio określonej – zajmiemy się szczegółowo w następnym paragrafie. Tutaj krótko omówimy grupę $\mathcal{L}$ izometrii przestrzeni Minkowskiego, nazywanych ***transformacjami Lorentza***, i spełniających na mocy definicji

$$Ax \cdot Ay = x \cdot y$$

dla każdej pary wektorów x i y. W szczególności wynika stąd, że transformacja Lorentza przeprowadza wektory czasowe, świetlne i przestrzenne w wektory odpowiednio tych samych typów. Grupę transformacji Lorentza nazywamy ***grupą Lorentza***.

Grupa ta jest izomorficzna z $O(1,3)$, więc w szczególności jeśli $A \in \mathcal{L}$, to $\det A = \pm 1$. Grupa $\mathcal{L}$ rozpada się więc na podgrupę $\mathcal{L}_+$ złożoną z transformacji Lorentza o wyznaczniku $+1$ oraz nie będący grupą zbiór $\mathcal{L}_-$ złożony z transformacji o wyznaczniku -1. W działaniu na dowolną bazę transformacje z $\mathcal{L}_+$, które nazywamy ***właściwymi transformacjami Lorentza***, nie zmieniają orientacji, a te z $\mathcal{L}_-$ – zmieniają ją na przeciwną.

Przestrzeń Minkowskiego posiada, jak wiemy, inną charakterystykę podobną do orientacji – orientację czasową. Zbadamy strukturę grupy $\mathcal{L}$ ze względu na tę orientację. Niech u będzie dowolnym wektorem czasowym. Wtedy na mocy twierdzenia 9, §15, mamy $|u \cdot Au| \geq u \cdot u$; jeśli przy tym $u \cdot Au > 0$, to wektory u i Au mają tę samą orientację czasową, a w przeciwnym wypadku – przeciwną. Niech teraz v będzie innym wektorem czasowym. Wtedy przy użyciu definicji transformacji Lorentza uzyskujemy

$$(u \cdot Av)\,(Av \cdot Au) = (Av \cdot u)\,(u \cdot v)\,.$$

Lewa strona tej równości ma na mocy lematu 10, §15, taki znak jak $u \cdot Au$, a prawa – jak $v \cdot Av$. Stąd, dla danej transformacji Lorentza A, znaki tych wyrażeń zgadzają się dla dowolnych wektorów czasowych. Wykazaliśmy następujące twierdzenie.

Twierdzenie 10. *Transformacja Lorentza albo nie zmienia orientacji czasowej żadnego wektora czasowego, albo zmienia orientację czasową wszystkich takich wektorów na przeciwną.*

Transformacje pierwszego typu nazywamy ***ortochronicznymi***. Transformacje ortochroniczne tworzą podgrupę $\mathcal{L}^{\uparrow}$, a transformacje nie-ortochroniczne – podzbiór $\mathcal{L}^{\downarrow}$. Przecięcia tych podzbiorów z podzbiorami $\mathcal{L}_{\pm}$ dają rozkład grupy Lorentza na cztery rozłączne i nie mające spójnego połączenia podzbiory: $\mathcal{L}^{\uparrow}_{+}$, $\mathcal{L}^{\downarrow}_{+}$, $\mathcal{L}^{\uparrow}_{-}$, $\mathcal{L}^{\downarrow}_{-}$, z których tylko pierwszy stanowi podgrupę, utworzoną z właściwych, ortochronicznych transformacji Lorentza. Można wykazać, że grupa $\mathcal{L}^{\uparrow}_{+}$ jest spójna łukowo (patrz poniżej przykład (viii), punkt 8).

6 Operatory samosprzężone. Operatory normalne

Wracamy do dyskusji operatorów w ogólnych przestrzeniach z niezdegenerowanym iloczynem skalarnym. Inną klasę, obok izometrii wewnętrznych, operatorów o szczególnym znaczeniu tworzą operatory spełniające warunek

$$A^* = A\,.$$

Każdy taki operator nazywamy ***operatorem samosprzężonym***.

Zarówno operatory izometryczne, jak samosprzężone, są szczególnymi przypadkami operatorów, dla których zachodzi

$$[A, A^*] = 0\,.$$

Każdy operator spełniający ten warunek nazywamy ***operatorem normalnym***.

7 Ortogonalne sumy proste i ortogonalne rozkłady jedności

Niech $(V_1, \dots, V_s)$ będą liniowo niezależnymi podprzestrzeniami przestrzeni V, więc ich suma jest sumą prostą,

$$\bigoplus_{i=1}^{s} V_i \subseteq V\,.$$

Jeśli, ponadto, przestrzeń V jest wyposażona w niezdegenerowaną formę metryczną g, względem której przestrzenie $(V_1, \dots, V_s)$ są wzajemnie ortogonalne,

$$V_i \perp V_j\,, \quad i \neq j\,,$$

to sumę tych podprzestrzeni nazywamy ***ortogonalną sumą prostą***. W często spotykanych sytuacjach sprawdzenie liniowej niezależności można pominąć, dzięki następującej implikacji.

Twierdzenie 11. *Jeśli podprzestrzenie* $(V_1, \ldots, V_s)$ *są niezdegenerowane i wzajemnie prostopadłe, to są liniowo niezależne.*

Dowód. Jeśli $\sum_{j=1}^{s} x_j = 0$, $x_j \in V_j$, to dla każdego wektora $y_i \in V_i$ mamy

$$0 = g\Big(y_i, \sum_{j=1}^{s} x_j\Big) = g(y_i, x_i)\,,$$

więc z niezdegenerowania przestrzeni V_i jest $x_i = 0$ dla każdego $i = 1, \ldots, s$. □

Załóżmy, że rozkład

$$V = \bigoplus_{i=1}^{s} V_i$$

jest ortogonalną sumą prostą i niech $P_1, \ldots, P_s$ będzie rozkładem jedności złożonym z operatorów rzutowania na te podprzestrzenie. Niech $x = \sum x_i$ i $y = \sum y_i$ będą rozkładami dowolnych wektorów. Wtedy z ciągu równości

$$g(x, P_i y) = g(x, y_i) = g(x_i, y_i) = g(x_i, y) = g(P_i x, y)$$

wnioskujemy, że wszystkie operatory P_i są samosprzężone. Okazuje się, że zachodzi też wynikanie odwrotne.

Twierdzenie 12. *Niech* $P_1, \ldots, P_s$ *będzie rozkładem jedności na przestrzeni* V *i niech* g *będzie niezdegenerowaną metryką na przestrzeni* V*. Suma*

$$V = \bigoplus_{i=1}^{s} P_i V$$

jest ortogonalną sumą prostą wtedy, i tylko wtedy, gdy wszystkie operatory rozkładu są samosprzężone względem metryki g*:*

$$P_i^* = P_i\,, \quad i = 1, \ldots, s\,.$$

Dowód. Dyskusja poprzedzająca twierdzenie pokazuje, że samosprzężoność jest warunkiem koniecznym ortogonalności; należy pokazać, że jest wystarczającym. Niech $x_i \in P_i V$, $i = 1, \ldots, s$, i niech operatory rzutowania będą samosprzężone. Wtedy dla $i \neq j$ mamy

$$g(x_i, x_j) = g(x_i, P_j x_j) = g(P_j x_i, x_j) = g(0, x_j) = 0\,,$$

co kończy dowód. □

Rozkład jedności, którego wszystkie operatory są samosprzężone, będziemy nazywać ***ortogonalnym rozkładem jedności***.

8 Przykłady

(i) Operatory w przestrzeni funkcji Schwartza
Rozważamy operatory D i X działające w przestrzeni funkcji Schwartza wyposażonej w iloczyn skalarny określony w przykładzie (i), p. 11, §15. Chociaż przestrzeń $\mathcal{S}$ jest nieskończenie wymiarowa, to warunek na operator sprzężony ma w tych przypadkach – jednoznaczne (z niezdegenerowania iloczynu) – rozwiązanie:

$$X^* = X\,, \qquad D^* = -D\,,$$

co jest konsekwencją tożsamości całkowych

$$\int_{\mathbb{R}} \overline{f(x)} x g(x)\, dx = \int_{\mathbb{R}} \overline{x f(x)} g(x)\, dx\,, \qquad \int_{\mathbb{R}} \overline{f(x)} g'(x)\, dx = -\int_{\mathbb{R}} \overline{f'(x)} g(x)\, dx\,.$$

Jeśli wprowadzić operator $P = -iD$, to z własności sprzężenia operatorowego mamy

$$P^* = P\,.$$

Operatory X i P, jako operatory na $\mathcal{S}$, są samosprzężone.

(ii) Operatory w przestrzeni wielomianowej $\mathcal{P}$
Rozważamy operatory D i X działające w przestrzeni $\mathcal{P}$ wyposażonej w iloczyn skalarny $(.\,,.)_H$ (przykład (ii), p. 11, §15). Postępując podobnie jak w poprzednim przykładzie pokazuje się, że w tej przestrzeni

$$X^* = X\,, \qquad D^* = -D + 2X\,.$$

Operator $T = -i(D - X)$ jest samosprzężony.

(iii) Operatory w przestrzeni wielomianowej $\mathcal{P}_n$
Podprzestrzeń $\mathcal{P}_n \subseteq \mathcal{P}$ złożona z wielomianów stopnia $\leq n$ jest inwariantna względem operatora D, więc można rozważać jego zacieśnienie do $\mathcal{P}_n$, ale nie jest inwariantna względem X, więc nie istnieje jego zacieśnienie do tej podprzestrzeni. Przestrzeń $\mathcal{P}_n$ ma skończony wymiar, więc istnieje operator D^* sprzężony do D względem iloczynu skalarnego $(.\,,.)_H$. Całkując przez części jak w poprzednim przykładzie dostajemy

$$(P, (D^* + D)Q)_H = 2 \int_{\mathbb{R}} \overline{P(x)}\, x Q(x)\, e^{-x^2}\, dx\,.$$

Jak pogodzić ten wynik z nieistnieniem operatora X w przestrzeni $\mathcal{P}_n$? Musimy znaleźć taki operator $\tilde{X}$, że dla każdych dwóch wielomianów $P, Q \in \mathcal{P}_n$ będzie

$$\int_{\mathbb{R}} \overline{P(x)}\, x Q(x)\, e^{-x^2}\, dx = \int_{\mathbb{R}} \overline{P(x)}\, (\tilde{X} Q)(x)\, e^{-x^2}\, dx \equiv (P, \tilde{X} Q)_H\,.$$

Łatwo sprawdzić, że operator ten jest dany przepisem

$$\begin{aligned} \tilde{X}Q &= XQ \qquad \text{dla} \quad Q \in \mathcal{P}_{n-1} \subseteq \mathcal{P}_n \,, \\ \tilde{X}e_n &= e_{n+1} - h_{n+1} \,, \end{aligned}$$

gdzie e_k są wektorami naturalnej bazy w $\mathcal{P}$, $e_k(x) = x^k$, a (h_k) jest bazą powstającą z (e_k) przez ortogonalizację Grama-Schmidta (przykład (ii), p. 11, §15). Stąd

$$D^* + D = 2\tilde{X} \,.$$

(iv) Jednoparametrowe, ciągłe grupy izometrii wewnętrznych
Jednoparametrowe, ciągłe grupy operatorów w rzeczywistej lub zespolonej przestrzeni o skończonym wymiarze zdefiniowaliśmy w przykładzie (v), p. 9, §13, gdzie pokazaliśmy też, że każda taka grupa ma postać $U(t) = \exp(tA)$, z generatorem $A \in \mathcal{L}(V,V)$. Jeśli przestrzeń jest wyposażona w niezdegenerowany iloczyn skalarny, to sprzężenie równania różniczkowego spełnionego przez $U(t)$ daje

$$\frac{d}{dt}U^*(t) = U^*(t)A^* \,, \quad \text{a stąd} \quad \frac{d}{dt}\big[U(t)^*U(t)\big] = U(t)^*\big[A^* + A\big]U(t) \,.$$

Na podstawie ostatniej równości łatwo wykazać następujące stwierdzenie.

Jednoparametrowa, ciągła grupa operatorów jest utworzona z operatorów izometrycznych wtedy, i tylko wtedy, gdy jej generator spełnia warunek

$$A^* + A = 0 \,.$$

(v) Operator sprzężony do operatora antyliniowego
Twierdzenie o operatorze sprzężonym zachodzi również, w następującej zmodyfikowanej formie, dla operatorów antyliniowych.

Jeśli V jest skończenie wymiarową zespoloną przestrzenią z niezdegenerowanym iloczynem hermitowskim g, to dla każdego operatora antyliniowego A istnieje dokładnie jeden operator antyliniowy A^* taki, że dla wszystkich wektorów x, y zachodzi

$$g(x, Ay) = g(y, A^*x) \,.$$

Dowód opiera się na obserwacji, że dla każdego takiego operatora A i wektora x funkcja $y \mapsto g(Ay, x)$ jest formą liniową. Dalej dowód przebiega z niewielkimi modyfikacjami jak dla operatora liniowego. Własność sprzężenia $(AB)^* = B^*A^*$ pozostaje w mocy, gdy jeden lub oba operatory są antyliniowe; dla A, B antyliniowych zachodzi $(A+B)^* = A^* + B^*$, natomiast $(\alpha A)^* = \alpha A^*$.

(vi) Szczególne transformacje Lorentza
Niech $(e_0, \ldots, e_3)$ będzie bazą ortonormalną w przestrzeni Minkowskiego. Definiujemy operator S określając jego działanie w tej bazie:

$$\begin{aligned} Se_0 &= \cosh\chi\, e_0 + \sinh\chi\, e_3\,, & Se_1 &= e_1\,, \\ Se_3 &= \sinh\chi\, e_0 + \cosh\chi\, e_3\,, & Se_2 &= e_2\,, \end{aligned}$$

gdzie χ jest dowolnym parametrem rzeczywistym. Operator S zachowuje iloczyny wektorów bazowych, więc jest transformacją Lorentza. Ponadto $\det S = 1$ oraz $e_0 \cdot Se_0 > 0$, więc $S \in \mathcal{L}_+^\uparrow$. Każdy operator, który w pewnej bazie ortonormalnej przyjmuje powyższą postać, nazywamy ***szczególną transformacją Lorentza***.

(vii) Odbicia
Niech $\{P_t, P_s\}$ będzie ortogonalnym rozkładem jedności związanym z rozkładem przestrzeni Minkowskiego $V = L(t) \oplus L(t)^\perp$, gdzie t jest dowolnym wektorem czasowym. Wtedy operatory $A_t = -P_t + P_s$ oraz $A_s = -A_t$ są transformacjami Lorentza. Transformacja A_t nazywa się odbiciem czasowym, transformacja A_s – odbiciem przestrzennym, a $A_tA_s = -\,\mathrm{id}$ – odbiciem całkowitym. Przez złożenie z właściwymi transformacjami Lorentza odbicia dają bijektywne obrazy:

$$A_t\mathcal{L}_+^\uparrow = \mathcal{L}_-^\downarrow\,, \quad A_s\mathcal{L}_+^\uparrow = \mathcal{L}_-^\uparrow\,, \quad A_tA_s\mathcal{L}_+^\uparrow = \mathcal{L}_+^\downarrow\,.$$

(viii) Rozkład właściwej, ortochronicznej transformacji Lorentza i spójność grupy $\mathcal{L}_+^\uparrow$.
Niech A będzie dowolną właściwą, ortochroniczną transformacją Lorentza, a e_0 – dowolnym unormowanym wektorem czasowym. Na mocy ortochroniczności A oraz twierdzenia 9, §15, mamy $e_0 \cdot Ae_0 \geq 1$, możemy więc oznaczyć $e_0 \cdot Ae_0 = \cosh\chi$, $\chi \geq 0$. Jeśli $\chi > 0$, to wektor $e_3 \equiv (\sinh\chi)^{-1}(Ae_0 - \cosh\chi e_0)$ jest przestrzennym, unormowanym wektorem prostopadłym do e_0. Wybieramy wektory e_1, e_2, które uzupełniają e_0, e_3 do bazy ortonormalnej. Oznaczmy przez S szczególną transformację Lorentza działającą w tej bazie tak, jak w przykładzie (vi). Wtedy $Ae_0 = Se_0$, więc jeśli oznaczyć $R = S^{-1}A$, to $Re_0 = e_0$. Jeśli natomiast $e_0 \cdot Ae_0 = 1$ ($\chi = 0$), to mamy $Ae_0 = e_0$ (bo Ae_0 jest wtedy unormowanym wektorem czasowym liniowo zależnym od e_0 i o tej samej orientacji czasowej). W tym przypadku kładziemy $S = \mathrm{id}$ i $A = R$. W obu przypadkach R jest transformacją Lorentza zachowującą wektor e_0. Pokazuje się przy użyciu definicyjnej własności transformacji Lorentza, że stąd również podprzestrzeń $L(e_0)^\perp$ jest inwariantna względem operatora R. Stąd $R = \mathrm{id}_0 \oplus R'$, gdzie id_0 jest identycznością na $L(e_0)$, a R' jest izometrią o dodatnim wyznaczniku na euklidesowej przestrzeni $L(e_1, e_2, e_3)$. Dla dowolnej właściwej transformacji Lorentza

otrzymaliśmy rozkład $A = SR$. Każdą szczególną transformację Lorentza można połączyć krzywą ciągłą takich transformacji z identycznością ($\chi \to 0$), a grupa izometrii o wyznaczniku dodatnim przestrzeni euklidesowej jest łukowo spójna (wykazujemy to w następnym paragrafie), więc grupa $\mathcal{L}_+^\uparrow$ jest łukowo spójna.

§17 Operatory normalne w przestrzeni unitarnej i euklidesowej

1 Izometria przestrzeni unitarnej (euklidesowej) z przestrzenią $\mathbb{C}^n$ ($\mathbb{R}^n$) z naturalnym iloczynem skalarnym

Wzory na macierz operatora sprzężonego otrzymane w punkcie 3, §16, redukują się w bazie ortonormalnej do

$$\mathbf{A}^* = \mathbf{A}^\dagger\,, \quad \mathbf{A}^* = \mathbf{A}^{\mathrm{T}}$$

dla przypadku unitarnego i euklidesowego odpowiednio. W bazie ortonormalnej macierz operatora samosprzężonego jest więc odpowiednio hermitowska lub symetryczna,

$$\mathbf{A}^\dagger = \mathbf{A}\,, \quad \mathbf{A}^{\mathrm{T}} = \mathbf{A}\,,$$

a macierz izometrii wewnętrznej jest unitarna lub ortogonalna,

$$\mathbf{A}^{-1} = \mathbf{A}^\dagger\,, \quad \mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}\,.$$

Nazwy odpowiednich macierzy przenosi się w tych przypadkach na operatory, i operator samosprzężony w przestrzeni unitarnej nazywa się ***operatorem hermitowskim***, w przestrzeni euklidesowej – ***operatorem symetrycznym***, izometrię w przestrzeni unitarnej – ***operatorem unitarnym***, a izometrię w przestrzeni euklidesowej – ***operatorem ortogonalnym***.

Wyniki punktu 10, §15, wraz z omówionymi powyżej, zyskują dodatkowe znaczenie w świetle następującej interpretacji. Przestrzeń wektorowa $\mathbb{C}^n$ ($\mathbb{R}^n$) staje się przestrzenią unitarną (euklidesową) po wyposażeniu jej w ***naturalny iloczyn skalarny*** zadany przepisem $\mathbf{x}^\dagger\mathbf{y}$ ($\mathbf{x}^{\mathrm{T}}\mathbf{y}$). Niech teraz $(e_1, \ldots, e_n)$ będzie dowolną bazą ortonormalną w przestrzeni unitarnej (euklidesowej) V. Wtedy omówiony w punkcie 9, §10, izomorfizm przestrzeni V z przestrzenią $\mathbb{C}^n$ ($\mathbb{R}^n$),

$$Jx = \mathbf{x}\,, \quad \text{gdzie} \quad x = x^i e_i\,,$$

staje się izometrią przestrzeni V z przestrzenią $\mathbb{C}^n$ ($\mathbb{R}^n$). Pod działaniem stowarzyszonego izomorfizmu operatorów liniowych na V z macierzami, p. 11, §10, sprzężenie operatorowe staje się sprzężeniem hermitowskim (transpozycją) macierzy, a operatory samosprzężone i izometryczne przechodzą w macierze o wyżej wyszczególnionych własnościach.

2 Diagonalizowalność operatorów normalnych w przestrzeni unitarnej

Niech A będzie operatorem w przestrzeni unitarnej, który posiada ortonormalną bazę wektorów własnych. Macierz $\mathbf{A}$ takiego operatora w jego bazie własnej jest diagonalna, a ponieważ baza jest ortonormalna, to macierz operatora sprzężonego jest jej hermitowskim sprzężeniem $\mathbf{A}^\dagger$, więc jest również diagonalna. Macierze diagonalne komutują, więc komutują również operatory A i A^*, stąd operator A jest normalny.

Podstawowy rezultat teorii operatorów w przestrzeni unitarnej mówi, że kierunek powyższego wnioskowania można odwrócić.

Lemat 1. *Niech A będzie operatorem normalnym w przestrzeni unitarnej. Wtedy:*

(i) $Ax = \lambda x \iff A^*x = \overline{\lambda}x$*;*

(ii) podprzestrzenie własne operatora A do różnych wartości własnych są ortogonalne.

Dowód. (i) Połóżmy $B = A - \lambda\,\mathrm{id}$. Jeśli A jest normalny, to również $[B, B^*] = 0$. Stąd

$$\|B^*x\|^2 = (B^*x, B^*x) = (x, BB^*x) = (x, B^*Bx) = (Bx, Bx) = \|Bx\|^2 ,$$

więc z własności normy $B^*x = 0$ wtedy, i tylko wtedy, gdy $Bx = 0$, co jest równoważne stwierdzeniu (i).
(ii) Niech $Ax_i = \lambda_i x_i$, $i = 1, 2$, $\lambda_1 \neq \lambda_2$. Wtedy korzystając z wyniku (i) mamy

$$\lambda_2(x_1, x_2) = (x_1, Ax_2) = (A^*x_1, x_2) = (\overline{\lambda_1}x_1, x_2) = \lambda_1(x_1, x_2) ,$$

więc $(x_1, x_2) = 0$. □

Twierdzenie 2. *Niech A będzie operatorem liniowym w przestrzeni unitarnej. Wtedy następujące warunki są równoważne:*

(i) A jest operatorem normalnym;

(ii) istnieje ortogonalny rozkład jedności $P_1, \ldots, P_s$ i liczby $\lambda_1, \ldots, \lambda_s$ takie, że

$$A = \sum_{i=1}^{s} \lambda_i P_i ;$$

(iii) operator A ma ortonormalną bazę wektorów własnych.

Dowód. Implikację (iii) $\Rightarrow$ (i) wykazaliśmy w dyskusji poprzedzającej lemat. Jeśli spełniony jest warunek (ii), to podprzestrzenie P_iV są wzajemnie ortogonalnymi podprzestrzeniami własnymi operatora A. Wybierając ortonormalną bazę każdej z nich uzyskujemy bazę pełnej przestrzeni, o której mówi punkt (iii). Pozostaje do wykazania implikacja (i) $\Rightarrow$ (ii).

Niech $V_1, \ldots, V_s$ będą wszystkimi podprzestrzeniami własnymi operatora A, do wartości własnych $\lambda_1, \ldots, \lambda_s$ odpowiednio. Podprzestrzenie V_i są wzajemnie ortogonalne, więc liniowo niezależne (tw. 11, §16). Pokażemy, że $V = \bigoplus_{i=1}^{s} V_i$, skąd wynika punkt (ii), z podstawieniem za P_i operatorów rzutowania na V_i. Dowód nie wprost: załóżmy, że $W = \bigoplus_{i=1}^{s} V_i$ nie wyczerpuje całej przestrzeni V, więc $V = W \oplus W^\perp$, $W^\perp \neq \{0\}$. Z lematu wiemy, że podprzestrzeń W jest inwariantna zarówno względem operatora A, jak i A^*. Stąd dla dowolnych wektorów $x \in W$ i $y \in W^\perp$ mamy $(x, Ay) = (A^*x, y) = 0$, więc podprzestrzeń $W^\perp$ jest również inwariantną podprzestrzenią operatora A, skąd $A = A_W \oplus A_{W^\perp}$. Operator $A_{W^\perp}$ ma co najmniej jeden wektor własny, więc przestrzeń W nie wyczerpuje wszystkich wektorów własnych operatora A, co jest sprzeczne z założeniem. □

Jeśli A jest operatorem normalnym i

$$A = \sum_{i=1}^{s} \lambda_i P_i\,, \quad \text{to} \quad A^* = \sum_{i=1}^{s} \overline{\lambda_i}\, P_i \quad \text{oraz} \quad A^*A = \sum_{i=1}^{s} |\lambda_i|^2\, P_i\,.$$

Stąd otrzymujemy natychmiast następujący wniosek.

Wniosek 3. *Operator normalny w przestrzeni unitarnej jest hermitowski (odpowiednio: unitarny) wtedy, i tylko wtedy, gdy wszystkie jego wartości własne są rzeczywiste (odpowiednio: mają moduł jednostkowy).*

3 Wyliczenie ortonormalnej bazy własnej operatora normalnego w przestrzeni unitarnej

Pierwszy krok wyliczenia bazy własnej przebiega według ogólnego schematu wyliczania wektorów własnych operatorów, omówionego w punktach 3 i 4, §13. Niech operator normalny A zadany będzie macierzą $\mathbf{A}$ w bazie $(e_1, \ldots, e_n)$. Wspomniany schemat pozwala wyznaczyć kolumny współrzędnych $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ jego bazy własnej $(f_1, \ldots, f_n)$ (której istnienie jest zagwarantowane twierdzeniem) w bazie $(e_1, \ldots, e_n)$. Możemy przyjąć taki porządek bazy $(f_1, \ldots, f_n)$, aby kolejne podciągi tej bazy były bazami podprzestrzeni własnych do kolejnych wartości własnych $\lambda_1, \ldots, \lambda_s$. Twierdzenie udowodnione w poprzednim punkcie gwarantuje, że podprzestrzenie własne do różnych wartości własnych są ortogonalne. Wewnątrz każdej z podprzestrzeni własnych stosujemy teraz do wektorów bazy f_i rozpinających tę podprzestrzeń ortogonalizację Grama-Schmidta. Rachunki ortogonalizacyjne prowadzimy na współrzędnych w bazie $(e_1, \dots, e_n)$, przy użyciu znanego nam już przepisu $(x, y) = \mathbf{x}^\dagger \mathbf{g}\, \mathbf{y}$, gdzie $\mathbf{g}$ jest macierzą Grama bazy $(e_1, \dots, e_n)$. Ostatni krok to unormowanie otrzymanych wektorów. Otrzymujemy ortonormalną bazę własną $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)\boldsymbol{\beta}$, gdzie $\boldsymbol{\beta} = \begin{pmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{pmatrix}$, oraz postać diagonalną macierzy operatora

$$\mathbf{A}' = \boldsymbol{\beta}^{-1} \mathbf{A} \boldsymbol{\beta} = \begin{pmatrix} \lambda_1 \mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s \mathbf{1}_{n_s} \end{pmatrix}.$$

W przestrzeni unitarnej najczęściej ograniczamy się do używania baz ortonormalnych. Jeśli $(e_1, \dots, e_n)$ jest taką bazą, to w tej bazie normalne operatory zadane są macierzami, dla których $[\mathbf{A}, \mathbf{A}^\dagger] = 0$. W szczególności, operatory hermitowskie zadane są macierzami hermitowskimi, a unitarne – unitarnymi. Macierz Grama zadanej bazy jest jednostkowa, więc na etapie ortogonalizacji używamy prostszego przepisu $(x, y) = \mathbf{x}^\dagger \mathbf{y}$. Znaleziona w powyższej procedurze baza własna jest też ortonormalna, więc macierz przejścia $\boldsymbol{\beta}$ jest w tym przypadku unitarna (p. 10, §15), stąd $\mathbf{A}' = \boldsymbol{\beta}^\dagger \mathbf{A}\, \boldsymbol{\beta}$.

Zapiszemy rezultaty dotyczące diagonalizacji w języku macierzowym.

Wniosek 4. *Niech* $\mathbf{A}$ *będzie zespoloną macierzą hermitowską lub unitarną. Wtedy istnieje macierz unitarna* $\boldsymbol{\beta}$ *taka, że macierz*

$$\mathbf{A}' = \boldsymbol{\beta}^\dagger \mathbf{A}\, \boldsymbol{\beta}$$

jest diagonalna. Liczby na diagonali macierzy $\mathbf{A}'$ *są pierwiastkami wielomianu charakterystycznego* $\det(\mathbf{A} - \lambda \mathbf{1})$*, z poprawnym uwzględnieniem krotności; są one rzeczywiste dla macierzy hermitowskiej, a o module jednostkowym – dla macierzy unitarnej.*

4 Diagonalizacja operatorów symetrycznych w przestrzeni euklidesowej

W przestrzeni euklidesowej klasa operatorów diagonalizowalnych w bazach ortonormalnych zawęża się do operatorów samosprzężonych. Wykażemy to w twierdzeniach analogicznych do twierdzeń z punktu 2 o diagonalizacji operatorów normalnych w przestrzeni unitarnej.

Lemat 5. *Niech A będzie operatorem symetrycznym w przestrzeni euklidesowej. Wtedy:*

(i) wielomian charakterystyczny operatora A rozkłada się na dwumiany:

$$\det(A - \lambda\,\mathrm{id}) = (\lambda_1 - \lambda)^{n_1} \dots (\lambda_s - \lambda)^{n_s}\,, \quad \lambda_1, \dots, \lambda_s \in \mathbb{R}\,;$$

(ii) podprzestrzenie własne operatora A do różnych wartości własnych są ortogonalne.

Dowód. (i) W bazie ortonormalnej macierz $\mathbf{A}$ operatora symetrycznego A jest symetryczna. Ponieważ rozpatrywana przestrzeń jest rzeczywista, to macierz ta jest macierzą rzeczywistą, więc jest również hermitowska. Przez zastosowanie wniosku 4 mamy więc $\mathbf{A} - \lambda\mathbf{1} = \boldsymbol{\beta}(\mathbf{A}' - \lambda\mathbf{1})\boldsymbol{\beta}^\dagger$, gdzie $\mathbf{A}'$ jest macierzą rzeczywistą diagonalną. Stąd

$$\det(A - \lambda\,\mathrm{id}) = \det(\mathbf{A} - \lambda\mathbf{1}) = \det(\mathbf{A}' - \lambda\mathbf{1})\,,$$

co daje żądany rozkład.

(ii) Niech $Ax_i = \lambda_i x_i$, $i = 1, 2$, $\lambda_1 \neq \lambda_2$. Wtedy

$$\lambda_2(x_1, x_2) = (x_1, Ax_2) = (Ax_1, x_2) = \lambda_1(x_1, x_2)\,,$$

skąd wynika teza. □

Twierdzenie 6. *Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni euklidesowej. Wtedy następujące warunki są równoważne:*

(i) A jest operatorem symetrycznym;

(ii) istnieje ortogonalny rozkład jedności $P_1, \dots, P_s$ i liczby $\lambda_1, \dots, \lambda_s \in \mathbb{R}$ takie, że

$$A = \sum_{i=1}^{s} \lambda_i P_i\,;$$

(iii) operator A ma ortonormalną bazę wektorów własnych.

Dowód. Implikację (ii) ⇒ (iii) wykazuje się tak samo, jak dla operatorów normalnych w przestrzeni unitarnej. Jeśli spełniony jest warunek (iii), to macierz operatora w bazie własnej jest diagonalna. Macierz operatora sprzężonego jest jej transpozycją, więc jest z nią identyczna. Stąd otrzymujemy $A^* = A$, czyli punkt (i).

Dowód implikacji (i) $\Rightarrow$ (ii) wymaga tylko niewielkich modyfikacji odpowiedniej części dowodu twierdzenia 2. Liczby $\lambda_1, \ldots, \lambda_s$ z lematu 5 są wszystkimi wartościami własnymi operatora A. Podprzestrzenie własne są ortogonalne, więc liniowo niezależne. Wprowadzamy przestrzeń $W = \bigoplus_{i=1}^{s} V_i$ i znów twierdzimy, że $V = W$. Gdyby tak nie było, to na mocy twierdzenia 4, §14, mielibyśmy $V = W \oplus W^\perp$, $W^\perp \neq \{0\}$. Podprzestrzeń W jest inwariantna względem operatora A, więc dla dowolnych wektorów $x \in W$ i $y \in W^\perp$ mamy $(x, Ay) = (Ax, y) = 0$, więc podprzestrzeń $W^\perp$ jest również inwariantna względem A, skąd $A = A_W \oplus A_{W^\perp}$. Dla dowolnych wektorów $y, z \in W^\perp$ mamy $(y, A_{W^\perp} z) = (y, Az) = (Ay, z) = (A_{W^\perp} y, z)$, więc operator $A_{W^\perp}$ jest symetryczny. Jego wielomian charakterystyczny jest rozkładalny, więc ma on co najmniej jeden wektor własny, zatem podprzestrzeń W nie wyczerpuje wszystkich wektorów własnych operatora A, co przeczy założeniu. □

Wyliczenie bazy własnej i postaci diagonalnej macierzy operatora symetrycznego przebiega według takiego schematu, jak dla operatora normalnego w przestrzeni unitarnej (poprzedni punkt), z zastąpieniem sprzężenia hermitowskiego macierzy transpozycją oraz unitarnych macierzy przejścia między bazami ortonormalnymi – macierzami ortogonalnymi. W szczególności otrzymujemy też macierzowe ujęcie diagonalizacji.

Wniosek 7. *Niech* $\mathbf{A}$ *będzie rzeczywistą macierzą symetryczną. Wtedy istnieje macierz ortogonalna* $\boldsymbol{\beta}$ *taka, że macierz*

$$\mathbf{A}' = \boldsymbol{\beta}^{\mathrm{T}} \mathbf{A}\, \boldsymbol{\beta}$$

jest diagonalna. Liczby na diagonali macierzy $\mathbf{A}'$ *są pierwiastkami wielomianu charakterystycznego* $\det(\mathbf{A} - \lambda \mathbf{1})$*, z poprawnym uwzględnieniem krotności.*

5 Przykłady

(i) Diagonalizacja operatora hermitowskiego

Operator A w przestrzeni unitarnej V ma w bazie ortonormalnej (e_1, e_2, e_3) macierz

$$\mathbf{A} = \begin{pmatrix} 3 & i & 1-i \\ -i & 3 & -1-i \\ 1+i & -1+i & 4 \end{pmatrix}.$$

Macierz $\mathbf{A}$ jest hermitowska, więc operator jest samosprzężony (hermitowski). Szukamy jego ortonormalnej bazy własnej.

Wyliczamy wielomian charakterystyczny:

$$P_A(\lambda) = \begin{vmatrix} 3-\lambda & i & 1-i \\ -i & 3-\lambda & -1-i \\ (1+i)(-2+\lambda) & 0 & 2-\lambda \end{vmatrix} = -(\lambda-2)^2(\lambda-6)\,.$$

Z ogólnych rozważań wynika, że podprzestrzeń własna do wartości własnej 2 jest dwuwymiarowa, a podprzestrzeń własna do wartości własnej 6 – jednowymiarowa. Równania: $(\mathbf{A} - 2\,\mathbf{1})\mathbf{x} = 0$ oraz $(\mathbf{A} - 6\,\mathbf{1})\mathbf{x} = 0$ sprowadzają się odpowiednio do jednego oraz do dwóch liniowo niezależnych równań:

$$\lambda = 2: \quad x^1 = -ix^2 + (-1+i)x^3\,, \qquad \lambda = 6: \quad \begin{pmatrix} -3 & i \\ -i & -3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} -1+i \\ 1+i \end{pmatrix} x^3\,.$$

Rozwiązując je otrzymujemy bazę w $\mathbb{C}^3$:

$$\mathbf{f}_1 = \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{f}_2 = \begin{pmatrix} -1+i \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{f}_3 = \begin{pmatrix} 1-i \\ -1-i \\ 2 \end{pmatrix},$$

gdzie $\mathbf{f}_1, \mathbf{f}_2$ rozpinają podprzestrzeń rozwiązań dla $\lambda = 2$, a $\mathbf{f}_3$ – dla $\lambda = 6$. Znów na podstawie ogólnych rozważań pierwsze dwa wektory są prostopadłe do trzeciego, ale wymagają ortogonalizacji względem siebie. Przenosząc więc trzeci wektor bez zmiany, a stosując schemat Grama-Schmidta do pierwszych dwóch, dostajemy:

$$\mathbf{f}_1' = \mathbf{f}_1 = \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{f}_2' = \frac{1}{2}\begin{pmatrix} -1+i \\ 1+i \\ 2 \end{pmatrix}, \quad \mathbf{f}_3' = \mathbf{f}_3 = \begin{pmatrix} 1-i \\ -1-i \\ 2 \end{pmatrix}.$$

Pozostaje już tylko unormować otrzymane wektory. Zauważmy, że w przypadku drugiego wektora czynnik $\frac{1}{2}$ można przy tym pominąć (bo dla $\lambda > 0$ i dowolnego wektora x mamy $\dfrac{\lambda x}{\|\lambda x\|} = \dfrac{x}{\|x\|}$). Dostajemy

$$\mathbf{e}_1' = \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2' = \frac{1}{2\sqrt{2}}\begin{pmatrix} -1+i \\ 1+i \\ 2 \end{pmatrix}, \quad \mathbf{e}_3' = \frac{1}{2\sqrt{2}}\begin{pmatrix} 1-i \\ -1-i \\ 2 \end{pmatrix},$$

a stąd ortonormalna baza własna operatora A w przestrzeni V jest równa

$$(e_1' \quad e_2' \quad e_3') = (e_1 \quad e_2 \quad e_3)\,\boldsymbol{\beta}\,, \qquad \boldsymbol{\beta} = \begin{pmatrix} \dfrac{-i}{\sqrt{2}} & \dfrac{-1+i}{2\sqrt{2}} & \dfrac{1-i}{2\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1+i}{2\sqrt{2}} & \dfrac{-1-i}{2\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}.$$

Zarówno baza „nieprimowana”, jak i „primowana”, są ortonormalne, więc macierz $\boldsymbol{\beta}$ jest unitarna. W nowej bazie macierz operatora A jest równa

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \boldsymbol{\beta}^{\dagger}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

– prawą stronę tego ciągu równości otrzymujemy natychmiast na mocy konstrukcji, dwie pośrednie równości wypisaliśmy jedynie dla przypomnienia związku macierzy $\mathbf{A}$ i $\mathbf{A}'$.

(ii) Diagonalizacja operatora symetrycznego
Operator A w przestrzeni euklidesowej V ma w bazie ortonormalnej (e_1, e_2, e_3) macierz

$$\mathbf{A} = \begin{pmatrix} -4 & 2 & 3 \\ 2 & -1 & 6 \\ 3 & 6 & 4 \end{pmatrix} .$$

Macierz $\mathbf{A}$ jest symetryczna, więc operator jest samosprzężony (symetryczny). Szukamy jego ortonormalnej bazy własnej.

Postępujemy tak, jak w przykładzie (i). Wyliczamy wielomian charakterystyczny $P_A(\lambda) = -(\lambda+5)^2(\lambda-9)$ i bazę przestrzeni $\mathbb{R}^3$

$$\mathbf{f}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} , \quad \mathbf{f}_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} , \quad \mathbf{f}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} ,$$

gdzie $\mathbf{f}_1, \mathbf{f}_2$ rozpinają podprzestrzeń rozwiązań równania $(\mathbf{A} + 5\,\mathbf{1})\mathbf{x} = 0$, a $\mathbf{f}_3$ – równania $(\mathbf{A} - 9\,\mathbf{1})\mathbf{x} = 0$. Po ortogonalizacji pierwszych dwóch wektorów mamy

$$\mathbf{f}'_1 = \mathbf{f}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} , \quad \mathbf{f}'_2 = \frac{1}{5}\begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix} , \quad \mathbf{f}'_3 = \mathbf{f}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} ,$$

a w wyniku normalizacji dostajemy

$$\mathbf{e}'_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} , \quad \mathbf{e}'_2 = \frac{1}{\sqrt{70}}\begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix} , \quad \mathbf{e}'_3 = \frac{1}{\sqrt{14}}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} .$$

Ortonormalna baza własna operatora A w przestrzeni V jest równa

$$\begin{pmatrix} e'_1 & e'_2 & e'_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta} , \qquad \boldsymbol{\beta} = \begin{pmatrix} \frac{-2}{\sqrt{5}} & \frac{-3}{\sqrt{70}} & \frac{1}{\sqrt{14}} \\ \frac{1}{\sqrt{5}} & \frac{-6}{\sqrt{70}} & \frac{2}{\sqrt{14}} \\ 0 & \frac{5}{\sqrt{70}} & \frac{3}{\sqrt{14}} \end{pmatrix} .$$

Obie bazy są ortonormalne, więc macierz $\boldsymbol{\beta}$ jest ortogonalna; jej wyznacznik jest równy $+1$, więc bazy mają wspólną orientację. W nowej bazie macierz operatora A jest równa

$$\mathbf{A}' = \boldsymbol{\beta}^{-1}\mathbf{A}\boldsymbol{\beta} = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{A}\boldsymbol{\beta} = \begin{pmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

– znów na mocy konstrukcji.

6 Komutujące operatory

Rozwijając w §13 teorię diagonalizowalności operatorów w ogólnych przestrzeniach wektorowych, otrzymaliśmy zawarte w twierdzeniu 8, §13, warunki, przy których operatory diagonalizowalne mają wspólną bazę własną. Dostosujemy teraz to twierdzenie do klas operatorów rozważanych w bieżącym paragrafie. Podobnie jak w przypadku ogólnym, twierdzenie ma oczywiste uogólnienie na dowolną liczbę komutujących operatorów.

Twierdzenie 8. *Niech A i B będą operatorami normalnymi w przestrzeni unitarnej lub operatorami symetrycznymi w przestrzeni euklidesowej, o (ortogonalnych) rozkładach spektralnych*

$$A = \sum_{i=1}^{r} \lambda_i P_i \,, \qquad B = \sum_{j=1}^{s} \mu_j Q_j \,.$$

Wtedy następujące warunki są równoważne:

(i) operatory A i B komutują:

$$[A, B] = 0 \,;$$

(ii) operatory rzutowe ich rozkładów spektralnych komutują:

$$[P_i, Q_j] = 0 \,, \quad i = 1, \ldots, r \,, \quad j = 1, \ldots, s \,;$$

(iii) operatory A i B mają wspólną ortonormalną bazę wektorów własnych.

Jeśli te warunki są spełnione, to układ operatorów

$$R_{ij} := P_i Q_j \,, \quad i = 1, \ldots, r \,, \quad j = 1, \ldots, s \,,$$

jest ortogonalnym rozkładem jedności, a operator R_{ij} rzutuje na podprzestrzeń złożoną ze wspólnych wektorów własnych operatorów A i B do wartości własnych odpowiednio λ_i, μ_j. Każda baza, o której mówi punkt (iii), jest sumą mnogościową pewnych baz wszystkich podprzestrzeni $R_{ij}V$.

Dowód. Operatory A i B są diagonalizowalne na mocy udowodnionych twierdzeń, więc stosuje się do nich twierdzenie 8, §13. Do wykazania pozostaje jedynie zastrzeżenie o ortonormalności bazy w warunku (iii) oraz własność ortogonalności rozkładu jedności $\{R_{ij}\}$, czyli samosprzężoności wszystkich operatorów R_{ij}. Ta ostatnia wynika natychmiast z samosprzężoności wszystkich operatorów P_i i Q_j oraz z ich przemienności:

$$R_{ij}^* = Q_j^* P_i^* = P_i Q_j = R_{ij}\,.$$

Podprzestrzenie $R_{ij}V$ są więc ortogonalne i rozpinają całą przestrzeń. Wybierając bazę ortonormalną każdej z nich dostajemy łącznie ortonormalną bazę własną, o której mówi punkt (iii) □

7 Odpowiedniość między operatorami samosprzężonymi i formami metrycznymi

Niech A będzie operatorem samosprzężonym w przestrzeni unitarnej (lub euklidesowej) V. Wtedy łatwo pokazać, że funkcja na $V \times V$ określona przepisem $\tilde{g}(x,y) = (x, Ay)$ jest nową hermitowską (odpowiednio: symetryczną) formą metryczną na V – wykazanie tego pozostawiamy czytelnikowi jako ćwiczenie. Pokażemy, że każdej formie można nadać takie przedstawienie.

Twierdzenie 9. *Niech V będzie przestrzenią unitarną (lub euklidesową) z iloczynem skalarnym oznaczanym przez $(\,.\,,.\,)$. Wtedy istnieje wzajemnie jednoznaczna odpowiedniość pomiędzy operatorami samosprzężonymi (względem iloczynu $(\,.\,,.\,)$) na V, a hermitowskimi (odpowiednio: symetrycznymi) metrykami na V, zadana przepisem*

$$\tilde{g}(x,y) = (x, Ay)\,,$$

gdzie A jest operatorem, a $\tilde{g}$ metryką. Przy tej odpowiedniości zachodzi tożsamość

$$\operatorname{Ker}\tilde{g} = \operatorname{Ker} A\,.$$

Jeśli $\mathbf{g}$, $\tilde{\mathbf{g}}$ i $\mathbf{A}$ są macierzami w dowolnej bazie odpowiednio: iloczynu skalarnego $(\,.\,,.\,)$, metryki $\tilde{g}$ i operatora A, to zachodzi

$$\tilde{\mathbf{g}} = \mathbf{g}\,\mathbf{A}\,; \quad \textit{w szczególności w bazie ortonormalnej:} \quad \tilde{\mathbf{g}} = \mathbf{A}\,.$$

Dowód. Widzieliśmy, że operator samosprzężony wyznacza formę $\tilde{g}$. Odwrotnie, niech $\tilde{g}$ będzie hermitowską (symetryczną) formą metryczną. Przy każdym ustalonym $x \in V$ funkcja $y \mapsto \tilde{g}(x,y)$ jest elementem przestrzeni $\mathcal{L}(V,\mathbb{K})$ ($\mathbb{K} = \mathbb{C}$ lub $\mathbb{R}$), więc zgodnie z twierdzeniem 3, §16 (stosowanym tutaj do metryki $(\,.\,,.\,)$), istnieje wektor x^*, dla którego $\tilde{g}(x,y) = (x^*, y)$ dla wszystkich $y \in V$. Wektor x^*

jest jednoznacznie określony przez wektor x, oznaczamy tę zależność funkcyjną przez $x^* = Ax$. Stąd dla każdej pary wektorów $x, y \in V$ mamy $\tilde{g}(x, y) = (Ax, y)$. Dalszy ciąg dowodu zapisujemy dla przypadku hermitowskiego, w przypadku symetrycznym należy jedynie pominąć sprzężenia. Odwzorowanie A jest liniowe: dla każdego y jest

$$\begin{aligned}(A(\alpha x + \beta z), y) = \tilde{g}(\alpha x + \beta z, y) &= \overline{\alpha}\, \tilde{g}(x, y) + \overline{\beta}\, \tilde{g}(z, y) \\ &= \overline{\alpha}\,(Ax, y) + \overline{\beta}\,(Az, y) = (\alpha Ax + \beta Az, y)\,,\end{aligned}$$

więc $A(\alpha x + \beta z) = \alpha Ax + \beta Az$. Ponadto mamy ciąg równości

$$(Ax, y) = \tilde{g}(x, y) = \overline{\tilde{g}(y, x)} = \overline{(Ay, x)} = (x, Ay)\,,$$

więc operator A jest samosprzężony. Stwierdzenie o jądrach wynika wprost z definicji jąder i niezdegenerowania metryki g. Wzór na związek macierzy metryki i operatora w dowolnej bazie dostaje się przez podstawienie za argumenty dowolnych wektorów bazowych: $\tilde{g}(e_i, e_j) = (e_i, e_k)A^k{}_j$. □

8 Jednoczesna diagonalizacja dwóch form hermitowskich lub symetrycznych

Niech g i $\tilde{g}$ będą dwiema formami hermitowskimi lub symetrycznymi odpowiednio w przestrzeni zespolonej lub rzeczywistej, przy czym niech forma g będzie dodatnio określona, a przestrzeń ma skończony wymiar. Wtedy przestrzeń V można uznać za przestrzeń unitarną (euklidesową) względem iloczynu skalarnego g. Stosując wynik poprzedniego punktu wnioskujemy więc o istnieniu dokładnie jednego operatora A, samosprzężonego względem iloczynu g i takiego, że

$$\tilde{g}(x, y) = g(x, Ay)$$

dla każdej pary wektorów. Stąd dla każdej liczby rzeczywistej λ forma $\tilde{g} - \lambda g$ jest hermitowska (symetryczna) i jej wartości wyrażają się przez

$$(\tilde{g} - \lambda g)(x, y) = g(x, (A - \lambda\, \mathrm{id})y)\,.$$

Na podstawie poprzedniego twierdzenia $V_\lambda \equiv \mathrm{Ker}(\tilde{g} - \lambda g)$ jest nietrywialna wtedy, i tylko wtedy, gdy λ jest wartością własną operatora A, i wówczas ta podprzestrzeń jest podprzestrzenią własną operatora A. Jeśli $y \in V_\lambda$, to dla każdego wektora x jest $\tilde{g}(x, y) = \lambda g(x, y)$. Stąd podprzestrzenie V_λ, które są ortogonalne względem g, są też ortogonalne względem $\tilde{g}$, a dla zacieśnień tych metryk mamy $\tilde{g}_{V_\lambda} = \lambda g_{V_\lambda}$. Wykorzystując teraz diagonalizowalność operatora A dostajemy następujące twierdzenie.

Twierdzenie 10. *Niech g i $\tilde{g}$ będą formami hermitowskimi (lub symetrycznymi) na zespolonej (odpowiednio: rzeczywistej), skończenie wymiarowej przestrzeni V, przy czym g jest dodatnio określona. Niech dalej $\lambda_1, \dots, \lambda_s$ będą wszystkimi liczbami charakterystycznymi, dla których* $\mathrm{Ker}(\tilde{g} - \lambda_i g) \neq \{0\}$. *Wtedy*

$$V = \bigoplus_{i=1}^{s} \mathrm{Ker}(\tilde{g} - \lambda_i g)\,,$$

przy czym ta suma prosta jest ortogonalna względem każdej z metryk g i $\tilde{g}$. W bazie ortonormalnej względem g, zgodnej z tym rozkładem przestrzeni, macierze metryk mają postać

$$\mathbf{g}' = \mathbf{1}\,, \quad \tilde{\mathbf{g}}' = \begin{pmatrix} \lambda_1 \mathbf{1}_{n_1} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbf{1}_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_s \mathbf{1}_{n_s} \end{pmatrix}.$$

Niech metryki g i $\tilde{g}$ będą zadane swoimi macierzami $\mathbf{g}$ i $\tilde{\mathbf{g}}$ w bazie $(e_1, \dots, e_n)$. Dodatnią określoność jednej z nich sprawdzamy np. metodą Sylwestra; niech g będzie dodatnio określona. Wyliczenie omówionej diagonalizacji sprowadza się do rozwiązania równań

$$\det(\tilde{\mathbf{g}} - \lambda \mathbf{g}) = 0\,, \quad (\tilde{\mathbf{g}} - \lambda \mathbf{g})\mathbf{x} = 0\,.$$

Liniowo niezależne wektory do tej samej wartości charakterystycznej ortogonalizujemy względem metryki g metodą Grama-Schmidta, a następnie normujemy, pamiętając przy obu operacjach, że w danej bazie $g(x, y) = \mathbf{x}^\dagger \mathbf{g} \mathbf{y}$ (lub $\mathbf{x}^{\mathrm{T}} \mathbf{g} \mathbf{y}$). W wyniku dostajemy

$$\begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \boldsymbol{\beta}\,, \quad \mathbf{g}' = \boldsymbol{\beta}^\dagger \mathbf{g}\, \boldsymbol{\beta}\,, \quad \tilde{\mathbf{g}}' = \boldsymbol{\beta}^\dagger \tilde{\mathbf{g}}\, \boldsymbol{\beta}\,,$$

dla przypadku hermitowskiego, lub z transpozycją dla przypadku symetrycznego, gdzie $\mathbf{g}'$ i $\tilde{\mathbf{g}}'$ są takie jak w tezie twierdzenia.

9 Przykłady

(i) Jednoczesna diagonalizacja komutujących operatorów samosprzężonych W ortonormalnej bazie (e_1, e_2, e_3) przestrzeni unitarnej dane są macierze samosprzężonych operatorów A i B:

$$\mathbf{A} = \begin{pmatrix} 3 & -i & \sqrt{2}i \\ i & 3 & \sqrt{2} \\ -\sqrt{2}i & \sqrt{2} & 2 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 5 & -3i & 3\sqrt{2}i \\ 3i & -3 & -\sqrt{2} \\ -3\sqrt{2}i & -\sqrt{2} & -2 \end{pmatrix}.$$

Sprawdzamy, że $\mathbf{AB} = \mathbf{BA}$, więc operatory A i B mają wspólną bazę wektorów własnych. Wartości własne operatora A są równe $\lambda = 0, 4$. Wyliczamy bazy podprzestrzeni własnych:

$$\begin{aligned}
&\text{dla } \lambda = 4: \qquad && f_1 = -ie_1 + e_2\,, \quad f_2 = \sqrt{2}ie_1 + e_3\,,\\
&\text{dla } \lambda = 0: \qquad && f_3 = e_1 - ie_2 + \sqrt{2}ie_3\,.
\end{aligned}$$

Ortogonalizujemy wektory f_1, f_2, normujemy całą bazę i dostajemy ortonormalną bazę własną operatora A:

$$\begin{pmatrix} f_1' & f_2' & f_3' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \frac{1}{2}\begin{pmatrix} -\sqrt{2}i & i & 1 \\ \sqrt{2} & 1 & -i \\ 0 & \sqrt{2} & \sqrt{2}i \end{pmatrix},$$

w której macierze operatorów mają postać

$$\mathbf{A}'_f = \boldsymbol{\beta}^\dagger \mathbf{A} \boldsymbol{\beta} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{B}'_f = \boldsymbol{\beta}^\dagger \mathbf{B} \boldsymbol{\beta} = 4\begin{pmatrix} 1 & -\sqrt{2} & 0 \\ -\sqrt{2} & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Podprzestrzeń rozpięta wektorami bazowymi f_1', f_2' jest inwariantna względem operatora B – jak należało oczekiwać na mocy komutowania A i B. Trzeba teraz znaleźć ortonormalną bazę własną zacieśnienia operatora B do operatora na tej podprzestrzeni. Macierz tego zacieśnienia w bazie (f_1', f_2') jest równa

$$\mathbf{B}'_{f12} = 4\begin{pmatrix} 1 & -\sqrt{2} \\ -\sqrt{2} & 0 \end{pmatrix}.$$

Rozwiązujemy zagadnienie własne i dostajemy wartości własne $\lambda = -4, 8$ i unormowane wektory własne

$$\begin{aligned}
&\text{dla } \lambda = 8: \qquad && e_1' = \frac{1}{\sqrt{3}}(-\sqrt{2}f_1' + f_2') = \frac{1}{2\sqrt{3}}(3ie_1 - e_2 + \sqrt{2}e_3)\,,\\
&\text{dla } \lambda = -4: \qquad && e_2' = \frac{1}{\sqrt{3}}(f_1' + \sqrt{2}f_2') = \frac{1}{\sqrt{3}}(\sqrt{2}e_2 + e_3)\,.
\end{aligned}$$

Trzeci wektor bazowy przenosimy bez zmiany

$$e_3' = f_3' = \frac{1}{2}(e_1 - ie_2 + \sqrt{2}ie_3)\,.$$

W nowej bazie operatory mają macierze diagonalne

$$\mathbf{A}' = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{\gamma}^\dagger \mathbf{A} \boldsymbol{\gamma}\,, \quad \mathbf{B}' = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{pmatrix} = \boldsymbol{\gamma}^\dagger \mathbf{B} \boldsymbol{\gamma}\,,$$

gdzie $\boldsymbol{\gamma}$ jest unitarną macierzą przejścia z bazy (e_1, e_2, e_3) do bazy (e_1', e_2', e_3').

(ii) Jednoczesna diagonalizacja dwóch form metrycznych
Niech V będzie trójwymiarową przestrzenią rzeczywistą. Na $V \times V$ zadajemy funkcję rzeczywistą

$$F(p,x) = g(p,p) + \tilde{g}(x,x)\,,$$

gdzie g i $\tilde{g}$ są symetrycznymi formami metrycznymi, przy czym g jest dodatnio określona. Szukamy bazy, w której F wyraża się jedynie przez kwadraty współrzędnych. Zadanie sprowadza się do jednoczesnej diagonalizacji form g i $\tilde{g}$.

Niech w bazie (e_1, e_2, e_3) funkcja F będzie dana przez

$$\begin{aligned} F(p,x) = {} & 2(p^1)^2 + 3(p^2)^2 + (p^3)^2 + 2p^1p^2 - 2p^2p^3 \\ & + (x^1)^2 + 15(x^2)^2 - 4(x^3)^2 + 10x^1x^2 - 18x^1x^3 - 10x^2x^3\,, \end{aligned}$$

więc macierze form są równe

$$\mathbf{g} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad \tilde{\mathbf{g}} = \begin{pmatrix} 1 & 5 & -9 \\ 5 & 15 & -5 \\ -9 & -5 & -4 \end{pmatrix}.$$

Stosując kryterium Sylwestra sprawdzamy, że g istotnie jest dodatnio określona. Wyliczamy $\det(\tilde{\mathbf{g}} - \lambda\mathbf{g}) = -3(\lambda-5)^2(\lambda+10)$ i znajdujemy bazę jądra $\mathrm{Ker}(\tilde{g} - \lambda g)$ dla każdej z wartości λ:

$$\begin{aligned} &\text{dla } \lambda = 5: && f_1 = e_2\,, \quad f_2 = e_1 - e_3\,, \\ &\text{dla } \lambda = -10: && f_3 = e_1 + e_2 + 4e_3\,. \end{aligned}$$

Przeprowadzamy ortogonalizację Grama-Schmidta układu f_1, f_2 względem metryki g. Kładziemy $f_1' = f_1$, $f_2' = f_2 + \alpha f_1$ i z warunku ortogonalności znajdujemy

$$\alpha = -\frac{g(f_1,f_2)}{g(f_1,f_1)} = -\frac{\mathbf{f}_1^{\mathrm{T}}\,\mathbf{g}\,\mathbf{f}_2}{\mathbf{f}_1^{\mathrm{T}}\,\mathbf{g}\,\mathbf{f}_1} = -\frac{2}{3}\,, \quad \text{więc} \quad f_2' = e_1 - \frac{2}{3}e_2 - e_3\,.$$

Przepisujemy $f_3' = f_3$. Normujemy wektory względem metryki g i dostajemy ostatecznie bazę

$$e_1' = \frac{1}{\sqrt{3}}e_2\,, \quad e_2' = \frac{1}{\sqrt{15}}(3e_1 - 2e_2 - 3e_3)\,, \quad e_3' = \frac{1}{\sqrt{15}}(e_1 + e_2 + 4e_3)\,,$$

w której funkcja F ma postać

$$F(p,x) = (p'^1)^2 + (p'^2)^2 + (p'^3)^2 + 5(x'^1)^2 + 5(x'^2)^2 - 10(x'^3)^2\,.$$

(iii) Przestrzenie operatorów samosprzężonych i metryk hermitowskich (lub symetrycznych)

Zbiór metryk hermitowskich (lub symetrycznych) w przestrzeni zespolonej (odpowiednio: rzeczywistej) tworzy rzeczywistą przestrzeń wektorową. Jeśli przestrzeń jest wyposażona w metrykę dodatnio określoną, to również zbiór operatorów samosprzężonych względem tej wybranej metryki tworzy rzeczywistą przestrzeń wektorową. Określone na początku punktu 7 przyporządkowanie $A \mapsto \tilde{g}$ jest różnowartościowym odwzorowaniem liniowym (jego jądro jest trywialne). Obie przestrzenie są izomorficzne z przestrzenią macierzy hermitowskich (odpowiednio: symetrycznych; por. przykład (iv) p. 10, §8), odpowiednie izomorfizmy przypisują metrykom i operatorom ich macierze w dowolnie wybranej bazie ortonormalnej. Stąd również przestrzenie metryk i przestrzenie operatorów samosprzężonych są izomorficzne, więc mają taki sam wymiar. Zatem wspomniane wyżej odwzorowanie jest izomorfizmem, czego konsekwencją jest między innymi treść twierdzenia 9.

(iv) Od metryk symetrycznych do operatorów symetrycznych

W przestrzeni euklidesowej operatory symetryczne oraz metryki symetryczne są diagonalizowalne w pewnej bazie ortonormalnej. Wykazaliśmy ten fakt najpierw dla operatorów korzystając z wyników uzyskanych dla przestrzeni unitarnych (patrz punkt (i) lematu 5), a odpowiednia własność metryk była konsekwencją ich związku z operatorami symetrycznymi. Odwrotnie, jeśli wykazalibyśmy niezależnie diagonalizowalność metryk (tj. twierdzenie 10), to ten sam związek dałby w wyniku diagonalizowalność operatorów symetrycznych. Naszkicujemy w tym przykładzie dowód diagonalizowalności metryk nie odwołujący się do wyników z przestrzeni unitarnych, który wymaga jednak użycia dwóch wyników analitycznych: (i) twierdzenia: rzeczywista funkcja ciągła na zbiorze zwartym osiąga na nim maksimum, oraz (ii) metody czynników Lagrange'a szukania ekstremum warunkowego.

Niech S oznacza zbiór wektorów x, dla których $g(x,x) = 1$ (przejmujemy oznaczenia punktu 8). Zapisany we współrzędnych w dowolnej bazie ortonormalnej zbiór S jest sferą, więc zbiorem zwartym. Stąd funkcja $S \ni x \mapsto \tilde{g}(x,x)$ ma na nim maksimum. Szukamy więc maksimum funkcji $V \ni x \mapsto \tilde{g}(x,x)$ przy warunku $x \in S$. Zgodnie z metodą Lagrange'a szukamy wektorów x i wartości parametru λ, przy których

$$\delta\big(\tilde{g}(x,x) - \lambda g(x,x)\big) = 0\,, \quad g(x,x) = 1\,,$$

gdzie δ oznacza zmianę przy dowolnej wariacji wektora x. Przy wykorzystaniu symetrii metryk mamy $\delta[g(x,x)] = g(x,\delta x) + g(\delta x, x) = 2g(\delta x, x)$ i podobnie dla $\tilde{g}$. Z dowolności δx oraz liniowości metryk w lewym argumencie wektor δx

można zastąpić w warunku dowolnym wektorem y (niekoniecznie *infinitezymalnym*). Dostajemy układ

$$\tilde{g}(.\,,x) - \lambda g(.\,,x) = 0\,, \quad g(x,x) = 1\,,$$

który na mocy istnienia maksimum na pewno ma rozwiązanie. Jeśli istnieje kilka wartości λ, to $\tilde{g}(x,x)$ osiąga maksimum dla największej z nich, oznaczmy ją λ_1. Powłoka liniowa wszystkich wektorów x spełniających układ dla λ_1 jest podprzestrzenią $\operatorname{Ker}(\tilde{g} - \lambda_1 g)$. Podprzestrzeń $\operatorname{Ker}(\tilde{g} - \lambda_1 g)^\perp$ ortogonalna do niej względem g, jest również ortogonalna względem $\tilde{g}$, więc wystarczy teraz zawęzić obie formy do tej podprzestrzeni. Kontynuując w ten sposób dostajemy tezę twierdzenia 10.

10 Postać kanoniczna operatora normalnego w przestrzeni euklidesowej

Operatory normalne w przestrzeni euklidesowej na ogół nie są diagonalizowalne (z wyjątkiem, jak już wiemy, operatorów symetrycznych). Zawsze jednak istnieje baza, w której macierz operatora normalnego przyjmuje szczególnie przejrzystą i łatwą do zinterpretowania postać. Zaczniemy analizę od lematu rozwiązującego szczególny przypadek problemu.

Lemat 11. *Niech C będzie operatorem w przestrzeni euklidesowej V spełniającym warunki*

$$C^{-1} = C^* = -C\,.$$

Wtedy istnieje baza ortonormalna, w której macierz tego operatora ma postać blokowo diagonalną:

$$\mathbf{C} = \begin{pmatrix} \mathbf{E} & 0 & 0 & \dots & 0 \\ 0 & \mathbf{E} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathbf{E} \end{pmatrix}\,, \quad \textit{gdzie} \quad \mathbf{E} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\,.$$

Dowód. Wybierzmy dowolny unormowany wektor e_1 i wprowadźmy $e_2 = Ce_1$. Operator C jest ortogonalny, więc wektor e_2 jest również unormowany. Ponadto, z własności operatora C mamy

$$(e_2, e_1) = (Ce_1, e_1) = -(e_1, Ce_1) = 0 \quad \text{oraz} \quad Ce_2 = -e_1\,,$$

więc uzyskaliśmy ortonormalną bazę dwuwymiarowej, inwariantnej podprzestrzeni operatora C, w której macierz zacieśnienia operatora do tej podprzestrzeni jest równa $\mathbf{E}$. Jeśli x jest dowolnym wektorem prostopadłym do obu wektorów e_i, to $(Cx, e_i) = -(x, Ce_i) = 0$. Uzyskujemy więc rozkład przestrzeni $V = L(e_1, e_2) \oplus L(e_1, e_2)^\perp$ na inwariantne podprzestrzenie operatora C. Na podprzestrzeni $L(e_1, e_2)^\perp$ operator spełnia założenia twierdzenia, więc można rekurencyjnie stosować opisaną procedurę, aż do wyczerpania całej przestrzeni V. W bazie otrzymanej w wyniku tego postępowania macierz operatora ma postać podaną w tezie. $\square$

Niech teraz A będzie dowolnym operatorem normalnym w przestrzeni euklidesowej V. Wtedy symetryczne operatory $A + A^*$ i A^*A komutują, więc istnieje rozkład ortogonalny przestrzeni na sumę prostą wspólnych podprzestrzeni własnych obu operatorów. Niech $V_{\mu\rho}$ będzie podprzestrzenią wektorów własnych tych operatorów do wartości własnych 2μ i ρ odpowiednio:

$$(A + A^*)x = 2\mu x\,, \quad A^*Ax = \rho x\,.$$

Ponieważ operatory A i A^* komutują, to stąd również wektor Ax spełnia te same równania, więc podprzestrzeń $V_{\mu\rho}$ jest inwariantna względem operatora A. Oznaczmy przez B zacieśnienie operatora $A - \mu\,\mathrm{id}$ do operatora na podprzestrzeni $V_{\mu\rho}$. Wtedy mamy

$$B + B^* = 0\,, \quad B^*B = (\rho - \mu^2)\,\mathrm{id}\;.$$

Dla dowolnego unormowanego wektora x w tej podprzestrzeni jest

$$0 \leq \|Bx\|^2 = (x, B^*Bx) = \rho - \mu^2 \equiv \nu^2\,, \quad \nu \geq 0\,.$$

Możliwe są dwa przypadki:
(i) Jeśli $\nu = 0$, to $B = 0$, więc operator A jest na tej podprzestrzeni równy $\mu\,\mathrm{id}$.
(ii) Jeśli $\nu > 0$, to oznaczmy $C = \nu^{-1}B$. Operator C spełnia założenia lematu, więc istnieje baza ortonormalna, w której macierz operatora C ma postać podaną w lemacie. Operator A na rozważanej podprzestrzeni jest równy $\mu\,\mathrm{id} + \nu C$, więc jego macierz w tej bazie ma postać blokowo diagonalną, z klatkami $\mu\mathbf{1} + \nu\mathbf{E}$ na diagonali. Uzyskaliśmy zasadniczą część wyniku następującego twierdzenia.

Twierdzenie 12. *Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni euklidesowej. Wtedy następujące warunki są równoważne:*

(i) A jest operatorem normalnym;

(ii) istnieje baza ortonormalna, w której operator ma postać blokowo diagonalną

$$\mathbf{A} = \begin{pmatrix} \Lambda_1 & & & & & \\ & \ddots & & & & \\ & & \Lambda_m & & & \\ & & & \mu_{m+1} & & \\ & & & & \ddots & \\ & & & & & \mu_{n-m} \end{pmatrix}, \qquad \Lambda_i = \mu_i \mathbf{1} + \nu_i \mathbf{E} = \begin{pmatrix} \mu_i & -\nu_i \\ \nu_i & \mu_i \end{pmatrix},$$

przy czym można dodatkowo żądać, aby $\nu_i > 0$.

W szczególności, operator A jest ortogonalny wtedy, i tylko wtedy, gdy z dokładnością do porządku liczb ± 1 na diagonali ta macierz ma postać

$$\mathbf{A} = \begin{pmatrix} O(\varphi_1) & & & & \\ & \ddots & & & \\ & & O(\varphi_m) & & \\ & & & -\mathbf{1}_q & \\ & & & & \mathbf{1}_p \end{pmatrix}, \quad O(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

gdzie można dodatkowo żądać $\varphi_i \in (0, \pi)$.

Dowód. Implikacja (i) $\Rightarrow$ (ii) wynika z dyskusji poprzedzającej twierdzenie. Odwrotnie, niech macierz operatora w pewnej bazie ortonormalnej ma postać daną w punkcie (ii). W bazie ortonormalnej macierz operatora sprzężonego jest transpozycją jego macierzy, więc macierz operatora A^*A w tej bazie ma podobną postać z blokami $\Lambda_i^{\mathrm{T}}\Lambda_i$ i liczbami μ_j^2, a dla operatora AA^* zmienia się jedynie porządek mnożenia bloków na $\Lambda_i\Lambda_i^{\mathrm{T}}$. Ale te iloczyny bloków dają w każdym z porządków $(\mu_i^2 + \nu_i^2)\mathbf{1}$, więc operator A jest normalny. Łatwo też teraz zauważyć, że operator jest ortogonalny wtedy, i tylko wtedy, gdy $\mu_i^2 + \nu_i^2 = 1$, $i = 1, \ldots, m$, oraz $\mu_j^2 = 1$, $j = m+1, \ldots, n-m$, co dowodzi ostatniego stwierdzenia. □

Operator ortogonalny, którego macierz w postaci kanonicznej jest złożona tylko z bloków $-\mathbf{1}_q$ i $\mathbf{1}_p$, przy czym $q > 0$, będziemy nazywać ***odbiciem*** (w podprzestrzeni rozpiętej wektorami zmieniającymi znak). W szczególności, jeśli $q = 1$, to powiemy, że jest to odbicie jednego kierunku (odbicie jednokierunkowe, lub jednowymiarowe). Każde odbicie jest złożeniem odbić kierunków wzajemnie ortogonalnych.

Operator liniowy w dwuwymiarowej przestrzeni euklidesowej, którego macierz w bazie ortonormalnej jest równa $O(\varphi)$, nazywamy ***obrotem o kąt*** φ w tej bazie.

Wyżej otrzymana postać kanoniczna pokazuje, że operator ortogonalny A jest równoczesnym obrotem w dwuwymiarowych, wzajemnie ortogonalnych podprzestrzeniach rozpiętych parami wektorów $(e'_1, e'_2), \ldots, (e'_{2m-1}, e'_{2m})$ i odbiciem kierunków $e'_{2m+1}, \ldots, e'_{2m+q}$ (ortogonalnych do podprzestrzeni obrotów).

11 Grupy operatorów unitarnych i ortogonalnych. Orientacja

Niech V będzie n-wymiarową przestrzenią unitarną (lub euklidesową). Zastosowanie wniosku 8, §16, pokazuje, że grupa operatorów unitarnych $U(V)$ (odpowiednio: ortogonalnych $O(V)$) w przestrzeni V jest izomorficzna z grupą macierzową $U(n)$ (odpowiednio: $O(n)$). W szczególności, dla każdego operatora ortogonalnego A jest $\det A = \pm 1$, i grupa $O(V)$ rozpada się na dwie części nie mające ciągłego połączenia: podgrupę $SO(V)$ złożoną z operatorów o wyznaczniku $+1$ oraz podzbiór $O(V)_-$ (nie tworzący grupy) operatorów o wyznaczniku -1. Operatory z grupy $SO(V)$ nazywamy ***właściwymi transformacjami ortogonalnymi***. Przy wykorzystaniu postaci kanonicznej operatorów unitarnych i ortogonalnych uzyskamy własności spójności ich grup, w analogii do dyskusji przeprowadzonej w §11 dla ogólnych grup liniowych.

Twierdzenie 13.
(i) Dla przestrzeni unitarnej V grupa izometrii wewnętrznych $U(V)$ jest łukowo spójna.
(ii) Dla przestrzeni euklidesowej V grupa izometrii wewnętrznych $O(V)$ nie jest spójna. Podgrupa $SO(V)$ i podzbiór $O(V)_-$ są łukowo spójne.

Dowód. (i) Jeśli A jest operatorem unitarnym, to w bazie własnej jego macierz jest diagonalna, z liczbami postaci $e^{i\varphi}$ na diagonali. Nie zmieniając bazy sprowadzamy wszystkie liczby φ w sposób ciągły do zera. Dostajemy w wyniku ciągłe połączenie operatora A z operatorem identycznościowym, nie opuszczające po drodze grupy $U(V)$.
(ii) Niespójność całej grupy jest oczywista. Wyznacznik bloku $O(\varphi)$ jest równy 1, więc jeśli A jest operatorem ortogonalnym o wyznaczniku $+1$, to liczba q w jego postaci kanonicznej jest parzysta. Blok $-\mathbf{1}_q$ można wtedy otrzymać poprzez ustawienie $q/2$ bloków diagonalnych $O(\pi)$. Bez ograniczenia ogólności możemy więc przyjąć, że w jego postaci kanonicznej $q = 0$. Łącząc teraz we wszystkich blokach $O(\varphi)$ liczby φ z zerem (przy zachowaniu bazy) dostajemy ciągłe połączenie operatora A z operatorem identycznościowym, pozostając po drodze wewnątrz grupy $SO(V)$. Jeśli A jest operatorem ortogonalnym o wyznaczniku -1, to wybierzmy pomocniczo i ustalmy inny taki operator R. Wtedy $AR \in SO(V)$, więc AR łączy się w sposób ciągły z operatorem identycznościowym. Stąd, każdy operator z podzbioru $O(V)_-$ łączy się w sposób ciągły z R^{-1}. □

Korzystając ze związku baz ortonormalnych z izometriami, tw. 9, §16, oraz z rezultatów punktu 5, §11, otrzymujemy następujący wniosek.

Wniosek 14.
(i) Każde dwie bazy ortonormalne w przestrzeni unitarnej można połączyć krzywą ciągłą baz ortonormalnych.
(ii) Dwie bazy ortonormalne w przestrzeni euklidesowej można połączyć krzywą ciągłą baz ortonormalnych wtedy, i tylko wtedy, gdy są one jednakowo zorientowane.

12 Postać kanoniczna operatora ortogonalnego w przestrzeni dwu- lub trójwymiarowej

Niech V będzie dwuwymiarową przestrzenią euklidesową. Jeśli A jest operatorem ortogonalnym o wyznaczniku $+1$, to jego macierz w dowolnej bazie ortonormalnej (e_1, e_2) jest elementem grupy $SO(2)$, więc – zgodnie z przykładem (v) z punktu 11, §7 – ma postać $O(\varphi)$. Operator z $SO(V)$ ma więc postać kanoniczną w każdej bazie ortonormalnej. Pozostawiamy jako ćwiczenie dla czytelnika wykazanie, że postać ta jest taka sama we wszystkich bazach ortonormalnych o tej samej orientacji, a różni się znakiem argumentu φ dla baz o orientacji przeciwnej. Stąd wynika też, że pary wektorów (e'_1, e'_2), ..., (e'_{2m-1}, e'_{2m}) bazy kanonicznej operatora ortogonalnego w dowolnej przestrzeni euklidesowej można obracać w odpowiednich podprzestrzeniach rozpiętych tymi parami bez zmiany postaci kanonicznej.

Jeśli A jest operatorem ortogonalnym o wyznaczniku -1, to jego macierz w bazie kanonicznej musi mieć postać $\left(\begin{smallmatrix} -1 & 0 \\ 0 & +1 \end{smallmatrix}\right)$. Operator ma więc dwie rzeczywiste wartości własne $+1$ i -1 i baza kanoniczna jest bazą własną.

Niech teraz V będzie trójwymiarową przestrzenią euklidesową. Jeśli A jest operatorem ortogonalnym, to jego macierz w bazie kanonicznej musi mieć postać blokową

$$\begin{pmatrix} O(\varphi) & 0 \\ 0 & \pm 1 \end{pmatrix}$$

(postać kanoniczna diagonalna może być przedstawiona jak wyżej z $\varphi = 0$ lub π), więc A ma wartość własną równą $\det A$. Znajdujemy wektor własny do tej wartości własnej, oznaczmy go e'_3. Z dyskusji w punkcie 10 wynika, że istnieją wektory e'_1 i e'_2 tworzące wspólnie z e'_3 bazę ortonormalną, w której operator ma postać kanoniczną. Wynik dyskusji przypadku dwuwymiarowego mówi nam, że za wektory (e'_1, e'_2) można przyjąć dowolny układ ortonormalny w podprzestrzeni $L(e'_3)^\perp$. Postać kanoniczną operatora A dostajemy teraz wyliczając jego działanie na wektory e'_1, e'_2. W szczególności, jeśli $\det A = +1$, to otrzymaną postać operatora nazywamy ***obrotem wokół osi*** $L(e_3')$. Dostajemy więc następujący ciekawy wynik.

Twierdzenie 15. *Każda właściwa transformacja ortogonalna w trójwymiarowej przestrzeni euklidesowej jest obrotem wokół pewnej osi.*

Omówione wyniki mają też inną ciekawą konsekwencję. Widzieliśmy, że w przestrzeni dwuwymiarowej operator ortogonalny o wyznaczniku -1 jest odbiciem jednego kierunku. Jeśli A jest operatorem ortogonalnym o wyznaczniku $+1$, to jego złożenie $AR = R'$ z dowolnym jednokierunkowym odbiciem R jest operatorem ortogonalnym o wyznaczniku -1, a więc znów odbiciem pewnego kierunku. Ponieważ dla każdego odbicia R jest $R^2 = \mathrm{id}$, to widzimy, że każdy operator ortogonalny o dodatnim wyznaczniku można złożyć z odbić dwóch kierunków, $A = R'R$ (zwracamy uwagę, że odbijane kierunki mogą przyjmować dowolne względne ułożenie). Z postaci kanonicznej operatora ortogonalnego w dowolnej przestrzeni euklidesowej dostajemy teraz następujące twierdzenie.

Twierdzenie 16. *Każdy operator ortogonalny w dowolnej przestrzeni euklidesowej może być przedstawiony jako złożenie odbić. Każdą właściwą transformację ortogonalną w przestrzeni dwu- lub trójwymiarowej można przedstawić jako złożenie odbić dwóch kierunków.*

13 Wyliczenie postaci kanonicznej operatora ortogonalnego

Podamy tutaj inną metodę sprowadzenia operatora ortogonalnego do postaci kanonicznej. (Po oczywistych modyfikacjach stosuje się ona również do operatorów normalnych.) Opiera się ona na zamianie problemu na zagadnienie macierzowe, i skorzystaniu z wyników dotyczących macierzy unitarnych.

Niech A będzie operatorem ortogonalnym w przestrzeni euklidesowej V. Wybieramy w tej przestrzeni dowolną bazę ortonormalną $(e_1, \dots, e_n)$. Macierz $\mathbf{A}$ operatora A w tej bazie jest ortogonalna. Ponieważ jest ona rzeczywista, to jest równocześnie unitarna, a więc można ją interpretować jak unitarny operator w przestrzeni $\mathbb{C}^n$ z naturalnym iloczynem skalarnym $\mathbf{x}^\dagger\mathbf{y}$. Operator ten posiada ortonormalną bazę wektorów własnych. Rzeczywistość macierzy $\mathbf{A}$ pociąga przy tym równoważność

$$\mathbf{A}\mathbf{z} = \lambda\mathbf{z} \iff \mathbf{A}\overline{\mathbf{z}} = \overline{\lambda}\overline{\mathbf{z}}\,.$$

Widmo operatora można więc zapisać w postaci

$$(e^{i\varphi_1}, e^{-i\varphi_1}, \dots, e^{i\varphi_m}, e^{-i\varphi_m}, \underbrace{-1, \dots, -1}_{q \text{ razy}}, \underbrace{1, \dots, 1}_{p \text{ razy}})\,, \quad \varphi_1, \dots, \varphi_m \in (0, \pi)\,,$$

– wyjątkowo w tym punkcie notujemy ciąg wartości własnych z uwzględnieniem krotności każdej z nich. Znajdujemy ortogonalną bazę wektorów własnych, przy czym wektory własne do rzeczywistych wartości własnych możemy wybrać jako rzeczywiste:

$$\mathbf{A}\mathbf{e}'_j = -\mathbf{e}'_j\,,\ j = 2m+1,\dots,2m+q\,, \quad \mathbf{A}\mathbf{e}'_j = +\mathbf{e}'_j\,,\ j = 2m+q+1,\dots,n\,,$$

a wektory własne do wartości $e^{-i\varphi_k}$ – jako sprzężenia wektorów do wartości $e^{+i\varphi_k}$:

$$\mathbf{A}\mathbf{f}_i = e^{+i\varphi_i}\mathbf{f}_i\,, \quad \mathbf{A}\overline{\mathbf{f}_i} = e^{-i\varphi_i}\overline{\mathbf{f}_i}\,, \quad i = 1,\dots,m\,.$$

Przyjmujemy jednostkową normalizację rzeczywistych wektorów oraz żądamy, aby kwadrat normy wektorów zespolonych był równy 2. Warunki te, łącznie z warunkami ortogonalności, mają postać

$$\begin{gathered}\mathbf{f}_i^\dagger\mathbf{f}_j = 2\delta_{ij}\,, \quad \overline{\mathbf{f}_i}^\dagger\mathbf{f}_j = 0\,, \quad i,j = 1,\dots,m\,,\\ \mathbf{f}_i^\dagger\mathbf{e}'_k = \overline{\mathbf{f}_i}^\dagger\mathbf{e}'_k = 0\,, \quad i = 1,\dots,m\,, \quad k = 2m+1,\dots,n\,,\\ {\mathbf{e}'_k}^{\mathrm{T}}\mathbf{e}'_l = \delta_{kl}\,, \quad k,l = 2m+1,\dots,n\,.\end{gathered}$$

Podstawiamy teraz

$$\mathbf{f}_1 = \mathbf{e}'_1 - i\mathbf{e}'_2\,, \quad \dots\,, \quad \mathbf{f}_m = \mathbf{e}'_{2m-1} - i\mathbf{e}'_{2m}\,,$$

gdzie wszystkie wektory $\mathbf{e}'_i$ są rzeczywiste. Prosty rachunek pokazuje, że otrzymujemy w ten sposób rzeczywistą bazę ortonormalną w $\mathbb{R}^n$

$$(\mathbf{e}'_1,\dots,\mathbf{e}'_n)\ : \quad {\mathbf{e}'_i}^{\mathrm{T}}\mathbf{e}'_j = \delta_{ij}\,.$$

Rozpisując równania własne dla wektorów $\mathbf{f}_i$ na część rzeczywistą i urojoną dostajemy działanie operatora $\mathbf{A}$ na pierwszych $2m$ wektorów tej bazy:

$$\mathbf{A}\mathbf{e}'_{2k-1} = \cos\varphi_k\mathbf{e}'_{2k-1} + \sin\varphi_k\mathbf{e}'_{2k}\,, \quad \mathbf{A}\mathbf{e}'_{2k} = -\sin\varphi_k\mathbf{e}'_{2k-1} + \cos\varphi_k\mathbf{e}'_{2k}\,.$$

Wróćmy teraz do przestrzeni euklidesowej V. Układamy ortogonalną macierz przejścia do nowej bazy ortonormalnej

$$\boldsymbol{\beta} = \begin{pmatrix}\mathbf{e}'_1 \dots \mathbf{e}'_n\end{pmatrix}\,, \quad (e'_1,\dots,e'_n) = (e_1,\dots,e_n)\boldsymbol{\beta}\,,$$

w której macierz $\mathbf{A}' = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{A}\boldsymbol{\beta}$ operatora A przyjmuje postać kanoniczną, taką jak podana w twierdzeniu 12.

14 Operatory dodatnie. Rozkład polarny

Niech A będzie operatorem samosprzężonym w przestrzeni unitarnej (lub euklidesowej). Wykazaliśmy w punkcie 7, że określa on jednoznacznie formę hermitowską (odpowiednio: symetryczną) przepisem (x, Ay). Mówimy, że operator A jest ***dodatni***, gdy zadana przez niego forma jest dodatnio określona:

$$(x, Ax) > 0 \quad \text{dla każdego wektora} \quad x \neq 0\,.$$

Twierdzenie 17. *Operator samosprzężony w przestrzeni unitarnej lub euklidesowej jest dodatni wtedy, i tylko wtedy, gdy wszystkie jego wartości własne są większe od zera. Operator dodatni jest więc nieosobliwy.*

Dowód. Niech A ma rozkład spektralny $A = \sum_{i=1}^{s} \lambda_i P_i$. Wtedy

$$(x, Ax) = \sum_{i=1}^{s} \lambda_i (x, P_i x) = \sum_{i=1}^{s} \lambda_i \|P_i x\|^2\,,$$

skąd teza jest natychmiast widoczna. □

Na dodatniej półosi rzeczywistej określona jest funkcja pierwiastka

$$\mathbb{R}^+ \ni x \mapsto x^{1/2} \in \mathbb{R}^+\,.$$

Jeśli więc A jest operatorem dodatnim, to istnieje jednoznacznie określony operator $A^{1/2}$, utworzony zgodnie z ogólną regułą omówioną w punkcie 8, §13:

$$A^{1/2} = \sum_{i=1}^{s} \lambda_i^{1/2} P_i\,.$$

Lemat 18. *Jeśli A i B są dodatnimi operatorami takimi, że*

$$A = B^2\,, \quad \textit{to} \ \ B = A^{1/2}\,.$$

Dowód. Jeśli spełnione jest równanie $A = B^2$ i

$$B = \sum_{i=1}^{s} \mu_i P_i\,, \quad \mu_i > 0\,, \quad \text{to} \ \ A = \sum_{i=1}^{s} \mu_i^2 P_i\,, \quad \text{więc} \ \ A^{1/2} = \sum_{i=1}^{s} \mu_i P_i = B\,.$$

□

Lemat 19. *Jeśli A jest nieosobliwym operatorem w przestrzeni unitarnej lub euklidesowej, to operator A^*A jest dodatni.*

Dowód. Dla dowolnego wektora x mamy

$$(x, A^*Ax) = (Ax, Ax) = \|Ax\|^2 \geq 0\,,$$

przy czym równość zachodzi wtedy, i tylko wtedy, gdy $Ax = 0$. Ponieważ A jest z założenia nieosobliwy, ostatni warunek jest równoważny stwierdzeniu $x = 0$. □

Twierdzenie 20. *Każdy nieosobliwy operator A w przestrzeni unitarnej (lub euklidesowej) ma jednoznaczne przedstawienie w postaci $A = U|A|$, gdzie $|A|$ jest operatorem samosprzężonym dodatnim, a U – operatorem unitarnym (odpowiednio: ortogonalnym). Dodatni operator w tym przedstawieniu jest wyznaczony przez*

$$|A| = (A^*A)^{1/2}\,.$$

Dowód. Szukamy rozkładu $A = U|A|$ o własnościach zadanych w twierdzeniu. Jeśli U ma być operatorem unitarnym, to musimy mieć

$$A^*A = |A|U^*U|A| = |A|^2\,.$$

Rozwiązanie tego równania ze względu na $|A|$ istnieje i jest jednoznacznie dane wzorem z tezy (korzystamy z dwóch ostatnich lematów). Utwórzmy operator $U = A|A|^{-1}$ i wyliczmy:

$$U^*U = |A|^{-1}A^*A|A|^{-1} = |A|^{-1}|A|^2|A|^{-1} = \mathrm{id}\,.$$

Operator U jest więc unitarny (lub ortogonalny). Otrzymaliśmy żądany rozkład wraz z gwarancją jego jednoznaczności. □

Rozkład operatora dany twierdzeniem nazywa się jego ***rozkładem polarnym***.

15 Przykłady

(i) Postać kanoniczna operatora ortogonalnego
Operator A ma w ortonormalnej bazie (e_1, e_2, e_3, e_4) przestrzeni euklidesowej V macierz

$$\mathbf{A} = \frac{1}{2}\begin{pmatrix} \sqrt{2} & 0 & -1 & 1 \\ 0 & \sqrt{2} & 1 & 1 \\ 1 & -1 & \sqrt{2} & 0 \\ -1 & -1 & 0 & \sqrt{2} \end{pmatrix}.$$

Macierz jest ortogonalna, więc operator również. Sprowadzamy go do postaci kanonicznej za pomocą procedury opisanej w punkcie 13. Znajdujemy

$$\det(\mathbf{A} - \lambda\mathbf{1}) = \left(\left(\lambda - \frac{1}{\sqrt{2}}\right)^2 + \frac{1}{2}\right)^2,$$

więc wartości własne operatora unitarnego w $\mathbb{C}^4$ będącego mnożeniem przez macierz $\mathbf{A}$ są równe $\lambda = \lambda_1, \overline{\lambda_1}$, gdzie $\lambda_1 = \exp(i\pi/4)$. Wyliczamy ortogonalną bazę podprzestrzeni własnej do wartości λ_1 i uwzględniamy warunek, aby kwadraty norm wektorów bazowych były równe 2. Otrzymujemy:

$$\mathbf{f}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ -i \\ \sqrt{2} \\ 0 \end{pmatrix}, \qquad \mathbf{f}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ -i \\ 0 \\ \sqrt{2} \end{pmatrix}$$

Rozdzielamy te wektory na części rzeczywiste i urojone i oznaczamy zgodnie z ogólnym przepisem wyłożonym w głównym tekście: $\mathbf{f}_1 = \mathbf{e}'_1 - i\mathbf{e}'_2$, $\mathbf{f}_2 = \mathbf{e}'_3 - i\mathbf{e}'_4$, więc

$$\mathbf{e}'_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}'_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}'_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{e}'_4 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Wracając do operatora A stwierdzamy więc, że w nowej bazie

$$e'_1 = e_3\,, \quad e'_2 = \frac{1}{\sqrt{2}}(-e_1 + e_2)\,, \quad e'_3 = e_4\,, \quad e'_4 = \frac{1}{\sqrt{2}}(e_1 + e_2)$$

jego macierz ma postać kanoniczną

$$\mathbf{A}' = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

(ii) Postać kanoniczna operatora ortogonalnego w przestrzeni trójwymiarowej

W przestrzeni trójwymiarowej zagadnienie można rachunkowo znacznie uprościć. Wyjaśniamy to na przykładzie.

Operator A ma w ortonormalnej bazie (e_1, e_2, e_3) przestrzeni euklidesowej V macierz

$$\mathbf{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}.$$

Macierz jest ortogonalna, więc operator również. Wyliczamy $\det \mathbf{A} = 1$, więc A ma rzeczywistą wartość własną równą 1, z unormowanym wektorem własnym

$$e'_3 = \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3).$$

Operator A jest więc obrotem wokół e'_3. Znajdując dowolne szczególne rozwiązanie równania $(e'_3, x) = 0$, czyli $x^1 + x^2 + x^3 = 0$, wybieramy wektor e'_1 prostopadły do e'_3 i normujemy go. Na przykład niech

$$e'_1 = \frac{1}{\sqrt{2}}(e_1 - e_2).$$

W postaci kanonicznej powinniśmy uzyskać

$$Ae'_1 = \cos\varphi e'_1 + \sin\varphi e'_2, \quad \sin\varphi \geq 0,$$

więc

$$\cos\varphi = (e'_1, Ae'_1), \quad \sin\varphi\, e'_2 = Ae'_1 - \cos\varphi\, e'_1.$$

Posługując się macierzą operatora A w bazie (e_1, e_2, e_3) wyliczamy

$$Ae'_1 = \frac{1}{\sqrt{2}}(e_1 - e_3), \quad \cos\varphi = \frac{1}{2}, \quad e'_2 = \frac{1}{\sqrt{6}}(e_1 + e_2 - 2e_3).$$

Stąd w nowej bazie macierz operatora A ma kanoniczną postać

$$\mathbf{A}' = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Załóżmy dodatkowo, że w przestrzeni V wybrano jedną z orientacji jako dodatnią. Jeśli baza (e'_1, e'_2, e'_3) ma orientację ujemną, to możemy utworzyć bazę o orientacji dodatniej zamieniając $e'_3 \to -e'_3$ i pozostawiając pozostałe wektory bazy bez zmiany. Macierz operatora nie zmienia się przy takiej zmianie bazy.

(iii) Jednoparametrowe, ciągłe grupy operatorów unitarnych i ortogonalnych

Generatory jednoparametrowych, ciągłych grup izometrii są operatorami spełniającymi warunek $A^*+A=0$ (przykład (iv), p. 8, §16). W przestrzeni unitarnej to oznacza $A=iH$, gdzie H jest operatorem samosprzężonym. Jeśli więc H ma rozkład spektralny

$$H=\sum_{i=1}^{s}\lambda_i P_i\,, \quad \text{to} \quad U(t)=\exp(itH)=\sum_{i=1}^{s}\exp(it\lambda_i)P_i\,.$$

W przestrzeni euklidesowej operator A może nie być diagonalizowalny, ale jest operatorem normalnym. W bazie ortonormalnej jego macierz jest antysymetryczna, więc postać kanoniczna określona twierdzeniem 12 ma w tym przypadku szczególną formę

$$\mathbf{A}=\begin{pmatrix}\Lambda_1 &&&&& \\ & \ddots &&&& \\ && \Lambda_m &&& \\ &&& 0 && \\ &&&& \ddots & \\ &&&&& 0\end{pmatrix}, \qquad \Lambda_i=\nu_i\mathbf{E}=\begin{pmatrix}0 & -\nu_i \\ \nu_i & 0\end{pmatrix},$$

Wtedy w tej samej bazie macierze operatorów $U(t)$ mają postać

$$\mathbf{U}(t)=\begin{pmatrix}O(t\nu_1) &&&&& \\ & \ddots &&&& \\ && O(t\nu_m) &&& \\ &&& 1 && \\ &&&& \ddots & \\ &&&&& 1\end{pmatrix}.$$

(iv) Rozkład polarny operatora w przestrzeni unitarnej

Operator A ma w ortonormalnej bazie (e_1, e_2) macierz

$$\mathbf{A}=\begin{pmatrix}8i & 11 \\ 4 & 7i\end{pmatrix}.$$

Szukamy jego rozkładu polarnego $A=U|A|$.

Diagonalizujemy operator A^*A. Jego macierz w danej bazie jest

$$\mathbf{A}^\dagger\mathbf{A}=10\begin{pmatrix}8 & -6i \\ 6i & 17\end{pmatrix}.$$

Znajdujemy ortonormalną bazę wektorów własnych

$$\begin{pmatrix} e_1' & e_2' \end{pmatrix} = \begin{pmatrix} e_1 & e_2 \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2i \\ 2i & 1 \end{pmatrix}\,, \quad \mathbf{A}'^{\dagger}\mathbf{A}' = 50 \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}\,.$$

Stąd

$$|\mathbf{A}'| = 5\sqrt{2} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\,,$$

więc

$$|\mathbf{A}| = \boldsymbol{\beta}\,|\mathbf{A}'|\,\boldsymbol{\beta}^{\dagger} = \sqrt{2} \begin{pmatrix} 6 & -2i \\ 2i & 9 \end{pmatrix}\,, \quad \mathbf{U} = \mathbf{A}\,|\mathbf{A}|^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\,.$$

IV

ALGEBRA TENSOROWA

§18 Iloczyn tensorowy

1 Proste przykłady iloczynów tensorowych

Rozważmy przestrzenie wektorowe $\mathbb{R}^m$ jako przestrzenie m-wyrazowych ciągów liczb rzeczywistych, czyli funkcji $\{1,\dots,m\} \ni i \mapsto x^i \in \mathbb{R}$. Jeśli $\mathbf{x} \equiv (x^i) \in \mathbb{R}^m$ i $\mathbf{y} \equiv (y^j) \in \mathbb{R}^n$, to korzystając z operacji mnożenia w ciele $\mathbb{R}$ możemy utworzyć funkcję

$$\{1,\dots,m\} \times \{1,\dots,n\} \ni (i,j) \mapsto x^i y^j \in \mathbb{R}\,,$$

a więc element przestrzeni wektorowej macierzy rzeczywistych $m \times n$; oznaczymy tę macierz $\mathbf{x} \otimes \mathbf{y}$. Co więcej, każda macierz z tej przestrzeni może być przedstawiona jako kombinacja liniowa takich iloczynów. Istotnie, jeśli $(\mathbf{e}_1,\dots,\mathbf{e}_m)$ jest bazą kanoniczną $\mathbb{R}^m$, a $(\mathbf{f}_1 \dots, \mathbf{f}_n)$ – bazą kanoniczną $\mathbb{R}^n$, to łatwo sprawdzić, że macierze $\mathbf{e}_i \otimes \mathbf{f}_j$, $i = 1,\dots,m$, $j = 1,\dots,n$, tworzą bazę kanoniczną przestrzeni macierzy $m \times n$. Ten sposób utworzenia przestrzeni $\mathbb{R}^{m\times n}$ z przestrzeni $\mathbb{R}^m$ i $\mathbb{R}^n$ nazwiemy iloczynem tensorowym i oznaczymy

$$\mathbb{R}^{m\times n} = \mathbb{R}^m \otimes \mathbb{R}^n\,.$$

Rozważmy drugi przykład. Niech $\mathcal{P}_m$ i $\mathcal{P}_n$ będą przestrzeniami wielomianów jednej zmiennej rzeczywistej, stopnia $\le m$ i $\le n$ odpowiednio. Dla każdej pary $P \in \mathcal{P}_m$, $Q \in \mathcal{P}_n$ tworzymy wielomian dwóch zmiennych przepisem

$$(P \otimes Q)(x,y) = P(x)Q(y)\,.$$

Wszystkie kombinacje liniowe takich wielomianów tworzą przestrzeń

$$\mathcal{P}_m \otimes \mathcal{P}_n = \mathcal{P}_{m,n}$$

wielomianów dwóch zmiennych rzeczywistych, stopnia $\leq m$ w pierwszej zmiennej i stopnia $\leq n$ w drugiej. Oznaczmy naturalne bazy przestrzeni $\mathcal{P}_m$ i $\mathcal{P}_n$ przez $e_{mi}(x) = x^i$ i $e_{nj}(y) = y^j$ odpowiednio. Wektory $e_{mi} \otimes e_{nj}$ stanowią bazę przestrzeni $\mathcal{P}_m \otimes \mathcal{P}_n$.

Naszym zadaniem w tym paragrafie jest uogólnienie tych konstrukcji na inne przestrzenie wektorowe. W kolejnych punktach omawiamy struktury, które posłużą do otrzymania tego uogólnienia. Ograniczamy się na ogół w tym rozdziale do przestrzeni skończenie wymiarowych.

2 Przestrzeń dualna

Wiemy z dyskusji w punkcie 10, §10, że zbiór odwzorowań liniowych $V \mapsto W$ tworzy przestrzeń wektorową $\mathcal{L}(V, W)$ o wymiarze równym iloczynowi wymiarów przestrzeni V i W. W szczególności, jeśli przestrzeń V jest zbudowana nad ciałem $\mathbb{K}$, to przestrzeń $\mathcal{L}(V, \mathbb{K})$ nazywamy ***przestrzenią dualną*** (lub ***sprzężoną***) ***do przestrzeni*** V i oznaczamy

$$V^* := \mathcal{L}(V, \mathbb{K}) .$$

Wymiar przestrzeni dualnej do V jest równy wymiarowi przestrzeni V. Elementy przestrzeni dualnej nazywamy ***formami liniowymi na*** V. Dla $\varphi \in V^*$, $x \in V$ zależność funkcyjną będziemy notować z nawiasami: $\varphi(x)$.

Niech $(e_1, \ldots, e_n)$ będzie dowolną bazą przestrzeni V. Określamy ciąg wektorów $(e^1, \ldots, e^n)$ w przestrzeni dualnej V^*, a więc ciąg form liniowych na V, za pomocą warunków:

$$e^i(e_j) = \delta^i_j , \quad i, j = 1, \ldots, n .$$

Wektory $(e^1, \ldots, e^n)$ są liniowo niezależne. Rzeczywiście, jeśli $0 = \sum_{j=1}^{n} \alpha_j e^j$, to podstawiając za argument tej formy wektor e_i dostajemy $0 = \sum_{j=1}^{n} \alpha_j e^j(e_i) = \alpha_i$, dla wszystkich $i = 1, \ldots, n$. Ponieważ wymiar przestrzeni V^* jest równy n, to wektory $(e^1, \ldots, e^n)$ tworzą bazę przestrzeni V^*, nazywamy ją ***bazą dualną do bazy*** $(e_1, \ldots, e_n)$.

Numery wektorów bazy dualnej notujemy jako wskaźniki górne, więc numery współrzędnych w tej bazie – jako wskaźniki dolne. Schemat ten ma użyteczną konsystencję. Niech dowolne wektory $x \in V$ i $\varphi \in V^*$ mają rozkłady

$$x = x^i e_i , \quad \varphi = \varphi_i e^i .$$

Wtedy z definicji bazy dualnej otrzymuje się natychmiast wyrażenia

$$x^i = e^i(x)\,, \quad \varphi_i = \varphi(e_i)\,.$$

Dla dowolnej formy φ i wektora x mamy teraz

$$\varphi(x) = \varphi_i x^i\,.$$

Każda para baz $(e_1, \ldots, e_n)$ i $(e^1, \ldots, e^n)$ zadaje izomorfizm między przestrzeniami V i V^*, określony przyporządkowaniem $e_i \mapsto e^i$. Izomorfizm ten zależy od wyboru baz, i w ogólnym przypadku nie ma między przestrzeniami V i V^* izomorfizmu wyróżnionego.

3 Przykłady

(i) Jądro formy liniowej

Jądro niezerowej formy na przestrzeni V jest podprzestrzenią o wymiarze równym $\dim V - 1$. Istotnie, jeśli forma jest niezerowa, to jej obraz jest równy $\mathbb{K}$, więc zastosowanie twierdzenia 3, §10, daje żądany wynik. Na przykład: jądro formy e^i jest powłoką liniową wektorów e_j, $j \neq i$.

Znajomość jądra niezerowej formy wyznacza tę formę z dokładnością do niezerowego czynnika. Niech bowiem wektor f nie należy do $\operatorname{Ker}\varphi$. Wtedy $V = L(f) \oplus \operatorname{Ker}\varphi$ i dla $x = \alpha f + y$, $y \in \operatorname{Ker}\varphi$, mamy $\varphi(x) = \alpha\varphi(f)$. Czynnik $\varphi(f)$ zawiera całą pozostałą informację o φ.

(ii) Formy liniowe na $\mathbb{K}^n$

Niech $\varphi_1, \ldots, \varphi_n$ będą liczbami z ciała $\mathbb{K}$. Odwzorowanie

$$\mathbb{K}^n \ni \mathbf{x} \mapsto \boldsymbol{\varphi}(\mathbf{x}) = \sum_{i=1}^{n} \varphi_i x^i \in \mathbb{K}$$

jest formą liniową na $\mathbb{K}^n$. W tym przypadku forma $\boldsymbol{\varphi}$ może być utożsamiona z elementem $\mathbb{K}^n$, a przestrzeń sprzężona do $\mathbb{K}^n$ – z nią samą. Istnienie tego naturalnego utożsamienia wynika tutaj z istnienia wyróżnionej bazy w $\mathbb{K}^n$: bazy kanonicznej.

(iii) Układ równań liniowych

Układ równań liniowych $\mathbf{A}\mathbf{x} = \mathbf{b}$ ma naturalną interpretację w kategoriach odwzorowań liniowych: macierz $\mathbf{A}$ zadaje odwzorowanie liniowe

$$\mathbb{K}^n \ni \mathbf{x} \mapsto \mathbf{A}\mathbf{x} \in \mathbb{K}^m\,,$$

układ jest warunkiem na to, aby $\mathbf{x} \mapsto \mathbf{b}$. Teraz możemy nadać układowi inną interpretację. Niech $\boldsymbol{\varphi}^i$, $i = 1, \ldots, m$, będą formami na $\mathbb{K}^n$:

$$\boldsymbol{\varphi}^i(\mathbf{x}) = \sum_{j=1}^{n} A^i{}_j \, x^j \, .$$

Układ możemy zapisać jako zespół warunków

$$\boldsymbol{\varphi}^i(\mathbf{x}) = b^i \, , \quad i = 1, \ldots, m \, .$$

Zbiór rozwiązań jest więc częścią wspólną zbiorów określonych oddzielnymi warunkami.

(iv) Współrzędne form liniowych
Każda z następujących funkcji rzeczywistych na przestrzeni $\mathcal{P}_n$ wielomianów zmiennej rzeczywistej stopnia $\leq n$ jest formą liniową:

$$\varphi(P) = \int f(x) P(x) \, dx \, , \quad \psi(P) = P(a) \, , \quad \rho(P) = P'(a) \, ,$$

gdzie f jest ustaloną ciągłą funkcją spełniającą ograniczenie

$$|f(x)| \leq \mathrm{const}\, (1 + |x|)^{-n-1-\epsilon} \, , \quad \epsilon > 0 \, ;$$

a jest dowolną liczbą rzeczywistą; P' jest pochodną wielomianu P. Wybierzmy bazę $e_i(x) = x^i$, $i = 0, \ldots, n$ przestrzeni $\mathcal{P}_n$. Współrzędne wyżej określonych form w bazie dualnej są kolejno dane przez

$$\varphi_i = \varphi(e_i) = \int f(x) x^i \, dx \, , \quad \psi_i = \psi(e_i) = a^i \, , \quad \rho_i = \rho(e_i) = i a^{i-1} \, .$$

4 Kanoniczny izomorfizm przestrzeni i jej dwusprzężonej

Operacja sprzężenia (dualności), jak widzieliśmy, produkuje z przestrzeni wektorowej V nową przestrzeń wektorową V^* o tym samym wymiarze. Operację sprzężenia możemy teraz zastosować powtórnie, produkując z przestrzeni V^* przestrzeń $(V^*)^*$, którą oznaczać będziemy dla uproszczenia przez V^{**} i nazwiemy ***przestrzenią dwusprzężoną***. Przestrzeń ta, zgodnie z definicją, jest przestrzenią form liniowych na przestrzeni V^*. Czy otrzymujemy w ten sposób znów istotnie nowy obiekt? Okazuje się, że tak nie jest w przypadku przestrzeni o skończonym wymiarze; pokażemy mianowicie, że przestrzenie V i V^{**} można wtedy w naturalny sposób utożsamić.

Twierdzenie 1. *Niech V będzie skończenie wymiarową przestrzenią nad ciałem $\mathbb{K}$. Dla każdego $x \in V$ określamy odwzorowanie*

$$x^{**} : V^* \mapsto \mathbb{K}\,, \quad x^{**}(\varphi) := \varphi(x)\,.$$

Wtedy

(i) odwzorowanie x^{**} jest liniowe, więc $x^{**} \in V^{**}$;

(ii) przyporządkowanie $V \ni x \mapsto x^{**} \in V^{**}$ jest izomorfizmem liniowym.

Dowód. (i) Ten punkt wynika z ciągu równości

$$x^{**}(\mu\varphi + \mu'\varphi') = (\mu\varphi + \mu'\varphi')(x) = \mu\varphi(x) + \mu'\varphi'(x) = \mu x^{**}(\varphi) + \mu' x^{**}(\varphi')\,.$$

(ii) Łatwo sprawdzamy, że odwzorowanie $x \mapsto x^{**}$ jest liniowe:

$$(\alpha\, x + \beta\, y)^{**} = \alpha\, x^{**} + \beta\, y^{**}$$

– wystarczy zadziałać obiema stronami na dowolną formę $\varphi \in V^*$, skorzystać z definicji x^{**} oraz liniowości φ. Przypomnijmy, że odwzorowanie liniowe pomiędzy dwiema przestrzeniami wektorowymi o tym samym, skończonym wymiarze jest izomorfizmem wtedy, i tylko wtedy, gdy jego jądro jest trywialne (tw. 6, §10). Załóżmy więc, że dla pewnego $x \in V$ jest $x^{**} = 0$. Oznacza to, że dla każdego $\varphi \in V^*$ jest $x^{**}(\varphi) = 0$, czyli $\varphi(x) = 0$. Kładąc $\varphi = e^i$ otrzymujemy $x^i = 0$ dla każdego i, a więc $x = 0$. Pokazaliśmy, że jądro odwzorowania $x \mapsto x^{**}$ składa się tylko z elementu zerowego, więc to odwzorowanie jest izomorfizmem przestrzeni V i V^{**}. □

Podkreślmy, że określenie izomorfizmu $V \ni x \mapsto x^{**} \in V^{**}$ nie wymagało zadania żadnej dodatkowej struktury, ani wybrania baz na przestrzeniach V lub V^{**}, a jego postać zdeterminowana była samą definicją dualności przestrzeni. Taki naturalny izomorfizm pozwala utożsamić obie przestrzenie: nie będziemy odtąd ich rozróżniać, utożsamiając x^{**} z x i zadając działanie elementów przestrzeni V na V^* przepisem

$$x(\varphi) := \varphi(x)\,.$$

Z podobnymi izomorficznymi utożsamieniami dwóch przestrzeni na mocy ich definicji będziemy mieć w tym rozdziale wielokrotnie do czynienia. Będziemy mówić w takich przypadkach, że przestrzenie są ***kanonicznie izomorficzne***, lub ***kanonicznie utożsamione***.

Jednym z elementów konstrukcji z pierwszego punktu tego paragrafu było przedstawienie wektorów jako funkcji liczbowych. Uzyskaliśmy teraz podobne przedstawienie w ogólnym przypadku skończenie wymiarowym: utożsamiliśmy wektor przestrzeni V z pewną funkcją na V^*.

Dokonane utożsamienie pozwala nam teraz spojrzeć w nowy sposób na parę przestrzeni V i V^*. Przy tym utożsamieniu dwukrotne złożenie sprzężenia przestrzeni jest identycznością. Wektory z przestrzeni V^* to formy liniowe na V, a wektory z przestrzeni V można również interpretować jako formy liniowe na V^*. Podkreślimy tę symetrię wprowadzając odwzorowanie, które nazywać będziemy ***funkcją*** (lub ***formą***) ***dualności przestrzeni V i V^****, określone przepisem:

$$V^* \times V \ni (\varphi, x) \mapsto \langle \varphi, x\rangle := \varphi(x) \equiv x(\varphi) \in \mathbb{K}\,.$$

Funkcja dualności jest odwzorowaniem liniowym w każdym z dwóch argumentów.

Dwukrotne złożenie sprzężenia baz jest również identycznością: jeśli bazą dualną do bazy $(e_1, \ldots, e_n)$ jest $(e^1, \ldots, e^n)$, to bazą dualną do bazy $(e^1, \ldots, e^n)$ jest $(e_1, \ldots, e_n)$. Związek baz wzajemnie dualnych oraz współczynniki rozkładów w tych bazach dane są w nowej notacji formułami

$$\langle e^i, e_j\rangle = \delta^i_j\,, \qquad x^i = \langle e^i, x\rangle\,, \qquad \varphi_i = \langle \varphi, e_i\rangle\,.$$

5 Podprzestrzenie ortogonalne względem dualności

Obecność funkcji dualności wiążącej przestrzenie V i V^* pozwala na wprowadzenie pojęcia ortogonalności, przypominającego analogiczne pojęcie w przypadku przestrzeni z iloczynem skalarnym. Powiemy, że podprzestrzenie $W \subseteq V$ oraz $U \subseteq V^*$ są ***ortogonalne względem dualności***, gdy

$$\forall x \in W, \varphi \in U : \langle \varphi, x\rangle = 0\,.$$

Ortogonalnym dopełnieniem (względem dualności) podprzestrzeni W przestrzeni V nazywamy podprzestrzeń

$$W^{\perp^*} := \{\varphi \in V^* \mid \forall x \in W : \langle \varphi, x\rangle = 0\} \subseteq V^*\,.$$

Podobnie określamy ortogonalne dopełnienie $U^{\perp^*} \subseteq V$ (względem dualności) podprzestrzeni $U \subseteq V^*$. Łatwo przekonać się, że ortogonalne dopełnienie podprzestrzeni jest największą podprzestrzenią ortogonalną do tej podprzestrzeni.

Niech podprzestrzeń V ma skończony wymiar i niech $(e_1, \ldots, e_n)$ będzie jej dowolną bazą. Jeśli $m \leq n$, to ortogonalne dopełnienie podprzestrzeni $L(e_1, \ldots, e_m)$ jest złożone z form φ, dla których $\varphi_i \equiv \langle \varphi, e_i\rangle = 0$, $i = 1, \ldots, m$, a więc jest podprzestrzenią $L(e^{m+1}, \ldots, e^n)$. Podobnie rozumując pokazujemy, że ortogonalne dopełnienie podprzestrzeni $L(e^{m+1}, \ldots, e^n)$ jest podprzestrzenią $L(e_1, \ldots, e_m)$. Wybierając odpowiednio bazy uzyskujemy więc następujący wynik.

Twierdzenie 2. *Jeśli $W \subseteq V$ i $U \subseteq V^*$ są podprzestrzeniami przestrzeni o skończonym wymiarze, to*

$$\dim W^{\perp^*} + \dim W = \dim V\,, \quad (W^{\perp^*})^{\perp^*} = W\,, \quad (U^{\perp^*})^{\perp^*} = U\,.$$

6 Odwzorowania wieloliniowe

Niech $V_1, \ldots, V_r$ oraz W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$. Zgodnie z wynikiem dyskusji z początku punktu 10, §10, zbiór wszystkich odwzorowań

$$V_1 \times \ldots \times V_r \mapsto W$$

tworzy przestrzeń wektorową nad $\mathbb{K}$. Odwzorowanie T z tej przestrzeni nazywamy ***odwzorowaniem wieloliniowym***, gdy jest ono liniowe w każdym z argumentów z osobna, tj. dla każdego $\alpha = 1, \ldots, r$ jest

$$T(x_1, \ldots, x_{\alpha-1}, \mu x_\alpha + \mu' x'_\alpha, x_{\alpha+1}, \ldots, x_r) =$$
$$= \mu\, T(x_1, \ldots, x_{\alpha-1}, x_\alpha, x_{\alpha+1}, \ldots, x_r) + \mu'\, T(x_1, \ldots, x_{\alpha-1}, x'_\alpha, x_{\alpha+1}, \ldots, x_r)\,.$$

Twierdzenie 3. *Zbiór odwzorowań wieloliniowych $V_1 \times \ldots \times V_r \mapsto W$ tworzy przestrzeń wektorową; oznaczać ją będziemy $\mathcal{L}(V_1, \ldots, V_r; W)$. Jej wymiar jest równy*

$$\dim \mathcal{L}(V_1, \ldots, V_r; W) = \dim W \prod_{\alpha=1}^{r} \dim V_\alpha\,.$$

Dowód. Łatwo przekonać się, że kombinacja liniowa odwzorowań wieloliniowych jest znów odwzorowaniem wieloliniowym (dowód jest analogiczny jak w przypadku odwzorowań liniowych, p. 10, §10), tworzą więc one podprzestrzeń przestrzeni wszystkich odwzorowań określonych na wstępie tego punktu. Stwierdzenia o wymiarze dowieść można przez indukcję względem r. Dla $r = 1$ sprowadza się ono do znanej nam własności przestrzeni odwzorowań liniowych. Spełnienie ogólnego kroku indukcyjnego wynika natychmiast z poniższego twierdzenia. □

Twierdzenie 4. *Istnieje kanoniczne, izomorficzne utożsamienie*

$$\mathcal{L}(V_1, \ldots, V_r; W) = \mathcal{L}\big(V_1, \mathcal{L}(V_2, \ldots, V_r; W)\big)\,,$$
więc $$\dim \mathcal{L}(V_1, \ldots, V_r; W) = \dim V_1 \dim \mathcal{L}(V_2, \ldots, V_r; W)\,.$$

Dowód. Rozważmy dowolne odwzorowanie $S \in \mathcal{L}\big(V_1, \mathcal{L}(V_2, \ldots, V_r; W)\big)$: przyporządkowuje ono każdemu wektorowi $x_1 \in V_1$ element $S(x_1)$ przestrzeni $\mathcal{L}(V_2, \ldots, V_r; W)$, czyli odwzorowanie wieloliniowe

$$V_2 \times \ldots \times V_r \ni (x_2, \ldots, x_r) \mapsto S(x_1)(x_2, \ldots, x_r) \in W\,.$$

Z określenia S wyrażenie $S(x_1)(x_2,\dots,x_r)$ jest również liniowe w swoim pierwszym argumencie x_1, więc przyporządkowanie

$$V_1\times\ldots\times V_r \ni (x_1,\dots,x_r)\mapsto S(x_1)(x_2,\dots,x_r)\in W$$

jest odwzorowaniem wieloliniowym, tj. elementem przestrzeni $\mathcal{L}(V_1,\dots,V_r;W)$. Odwrotnie, oddzielając pierwszy argument dowolnego odwzorowania wieloliniowego od pozostałych, dostajemy element przestrzeni $\mathcal{L}\big(V_1,\mathcal{L}(V_2,\dots,V_r;W)\big)$. Przestrzenie, o których mówi twierdzenie, można więc uważać za tę samą przestrzeń, odmiennie zinterpretowaną w każdym z dwóch przypadków. Utożsamienie to zachowuje strukturę liniową, więc jest izomorfizmem. □

7 Przykłady

(i) Dualne dopełnienie przecięcia i sumy podprzestrzeni
Dla każdych dwóch podprzestrzeni $W_1,W_2\subseteq V$, $\dim V<\infty$, zachodzi

$$(W_1\cap W_2)^{\perp^*}=W_1^{\perp^*}+W_2^{\perp^*}\,,\quad (W_1+W_2)^{\perp^*}=W_1^{\perp^*}\cap W_2^{\perp^*}\,.$$

Pierwsza tożsamość ma równoważne sformułowanie $W_1\cap W_2=(W_1^{\perp^*}+W_2^{\perp^*})^{\perp^*}$, a ta postać wynika z ciągu równoważności

$$x\in W_1\cap W_2 \Leftrightarrow x\in W_1\wedge x\in W_2 \Leftrightarrow x\perp W_1^{\perp^*}\wedge x\perp W_2^{\perp^*} \Leftrightarrow x\perp (W_1^{\perp^*}+W_2^{\perp^*})\,.$$

Druga tożsamość ma podobny dowód. Tożsamości mają oczywiste uogólnienia na dowolną skończoną rodzinę podprzestrzeni.

(ii) Funkcja dualności
Funkcja dualności przestrzeni V i V^* jest odwzorowaniem biliniowym, co jest szczególnym przypadkiem odwzorowania wieloliniowego.

(iii) Metryka i jej symetrie
Metryki symetryczne oraz symplektyczne są odwzorowaniami biliniowymi.

Wykażemy, że forma biliniowa, dla której relacja ortogonalności jest symetryczna, jest albo symetryczna, albo symplektyczna. Ze względu na twierdzenie 1, §14, wystarczy ograniczyć się do metryk niezdegenerowanych. Wtedy dla każdego wektora x forma $g(x,.\,)$ jest niezerowa.

Symetryczność ortogonalności oznacza, że dla każdego wektora x formy liniowe $g(\,.\,,x)$ i $g(x,.\,)$ mają identyczne jądro. Stąd na podstawie przykładu (i), p. 3, mamy $g(\,.\,,x)=\mu g(x,.\,)$. Niech y będzie liniowo niezależny od x. Podobnie rozumując dostajemy $g(\,.\,,y)=\nu g(y,.\,)$ oraz $g(\,.\,,x+y)=\lambda g(x+y,.\,)$.

Łącząc te trzy równania otrzymujemy związek $g((\lambda-\mu)x+(\lambda-\nu)y,.)=0$. Ale metryka jest niezdegenerowana, a wektory liniowo niezależne, więc $\mu=\nu=\lambda$ – czynnik proporcjonalności jest wspólny dla wszystkich wektorów przestrzeni.

Niech $g(x,z)\neq 0$. Wtedy $g(x,z)=\lambda g(z,x)=\lambda^2 g(x,z)$, więc $\lambda=+1$ lub $\lambda=-1$.

Dla metryki półtoraliniowej pokazuje się analogiczną metodą, że g ma postać "liczba o module jeden $\times$ metryka hermitowska"; zostawiamy ten problem jako zadanie.

(iv) Wyznacznik macierzy

Każdą macierz $m\times n$ o elementach w ciele $\mathbb{K}$ możemy interpretować jako n-wyrazowy ciąg kolumn o m elementach, a więc element n-krotnego iloczynu kartezjańskiego przestrzeni $\mathbb{K}^m$ (patrz p. 1, §6); w szczególności dla macierzy kwadratowych $m=n$. Własność (vii) wyznacznika macierzy, tw. 2, §7, można teraz ująć mówiąc, że odwzorowanie

$$\underbrace{\mathbb{K}^n\times\ldots\times\mathbb{K}^n}_{n\ \text{razy}}\ni(\mathbf{A}_{*1},\ldots,\mathbf{A}_{*n})\mapsto\det\mathbf{A}\in\mathbb{K}$$

jest odwzorowaniem wieloliniowym. Podobnie dla macierzy traktowanej jak ciąg wierszy.

(v) Zbiór wartości odwzorowania wieloliniowego

Podobnie jak odwzorowania liniowe, odwzorowania wieloliniowe mogą być jednoznacznie zadane działaniem na wektory bazowe. Niech na przykład ciąg wektorów $(e_{11},\ldots,e_{1n_1})$ będzie bazą przestrzeni V_1, a ciąg $(e_{21},\ldots,e_{2n_2})$ – bazą przestrzeni V_2. Wtedy, przy dowolnym wyborze wektorów $B(e_{1i},e_{2j})\in W$, przepis

$$B(x_1,x_2)={x_1}^i{x_2}^j B(e_{1i},e_{2j})\,,\quad\text{gdzie}\quad x_\alpha={x_\alpha}^{i_\alpha}e_{\alpha i_\alpha}\,,\ \alpha=1,2\,,$$

jednoznacznie określa odwzorowanie wieloliniowe $B\in\mathcal{L}(V_1,V_2;W)$.

Rozważmy zbiór wartości odwzorowania wieloliniowego. Przeciwnie niż w przypadku liniowym, ten zbiór nie tworzy, na ogół, podprzestrzeni przestrzeni W: powłoka liniowa zbioru wartości odwzorowania B jest utworzona przez wszystkie kombinacje liniowe $\alpha^{ij}B(e_{1i},e_{2j})$, a nie tylko te, dla których współczynniki separują się, jak w powyższym określeniu B. Widać też stąd, że w ogólnym przypadku maksymalny możliwy wymiar tej powłoki dla odwzorowania z przestrzeni $\mathcal{L}(V_1,\ldots,V_r;W)$ to

$$\min\left\{\dim W,\ \prod_{i=1}^r\dim V_i\right\}.$$

Wróćmy do przypadku $B \in \mathcal{L}(V_1, V_2; W)$. Przy powyższych oznaczeniach niech $n_1 = n_2 = 2$ oraz niech (f_1, f_2, f_3, f_4) będzie bazą W. Połóżmy

$$B(e_{11}, e_{21}) = f_2\,, \quad B(e_{11}, e_{22}) = f_1\,, \quad B(e_{12}, e_{21}) = -f_1\,, \quad B(e_{12}, e_{22}) = f_3\,.$$

Powłoka zbioru wartości odwzorowania B jest trójwymiarowa, ze względu na zależność liniową $B(e_{11}, e_{22}) + B(e_{12}, e_{21}) = 0$. Jednak $B(x_1, x_2) = 0$, tylko gdy $x_1 = 0$ lub $x_2 = 0$.

8 Iloczyn tensorowy: model odwzorowań (form) wieloliniowych

Niech $V_1, \ldots, V_r$ będą skończenie wymiarowymi przestrzeniami wektorowymi nad ciałem $\mathbb{K}$, a $V_1^*, \ldots, V_r^*$ – ich przestrzeniami dualnymi. Dzięki wynikom punktu 4 możemy przestrzenie V_α traktować jako przestrzenie funkcji – form liniowych na V_α^*,

$$x_\alpha : V_\alpha^* \ni \varphi^\alpha \mapsto \langle \varphi^\alpha, x_\alpha \rangle_\alpha \in \mathbb{K}\,,$$

gdzie $\langle .\,, . \rangle_\alpha$ jest funkcją dualności pary przestrzeni V_α^*, V_α. Funkcje takie można w naturalny sposób mnożyć, zadajemy ich iloczyn jako odwzorowanie

$$\begin{gathered} x_1 \otimes \ldots \otimes x_r : V_1^* \times \ldots \times V_r^* \mapsto \mathbb{K}\,, \\ [x_1 \otimes \ldots \otimes x_r]\,(\varphi^1, \ldots, \varphi^r) := \langle \varphi^1, x_1 \rangle_1 \ldots \langle \varphi^r, x_r \rangle_r\,. \end{gathered}$$

Z liniowości każdej z funkcji dualności w pierwszym argumencie wynika, że $x_1 \otimes \ldots \otimes x_r$ jest odwzorowaniem wieloliniowym, a więc wektorem przestrzeni $\mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})$. Co więcej wektory tej postaci generują całą przestrzeń,

$$\mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K}) = L\big(\{x_1 \otimes \ldots \otimes x_r \mid x_\alpha \in V_\alpha,\ \alpha = 1, \ldots, r\}\big)$$

– przekonamy się o tym w następnym punkcie, gdzie z wektorów tego typu utworzymy bazę. Ten sposób otrzymania przestrzeni $\mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})$ z przestrzeni $V_1, \ldots, V_r$ nosi więc istotne cechy iloczynu. Powiemy, że przestrzeń

$$V_1 \otimes \ldots \otimes V_r := \mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})\,.$$

jest iloczynem tensorowym przestrzeni $V_1, \ldots, V_r$. Wkrótce poznamy ogólniejszy sposób określenia iloczynu tensorowego, dlatego powyższą konstrukcję nazwiemy dokładniej ***modelem form (lub odwzorowań) wieloliniowych iloczynu tensorowego***. Ponieważ $\dim V_\alpha^* = \dim V_\alpha$, to zgodnie z twierdzeniem 3 wymiar tej przestrzeni jest równy

$$\dim(V_1 \otimes \ldots \otimes V_r) = \prod_{i=1}^{r} \dim V_i\,.$$

Wektor o postaci $x_1 \otimes \ldots \otimes x_r$ nazwiemy ***iloczynem tensorowym wektorów*** $x_1, \ldots, x_r$ lub ***wektorem iloczynowym*** przestrzeni $V_1 \otimes \ldots \otimes V_r$.

9 Bazy iloczynowe w modelu form wieloliniowych

Pozostajemy w tym punkcie w ramach modelu iloczynu tensorowego wprowadzonego w punkcie poprzednim.

Twierdzenie 5. *Niech* $\{e_{\alpha 1}, \ldots, e_{\alpha n_\alpha}\}$ *będzie bazą przestrzeni* V_α *dla każdego* $\alpha = 1, \ldots, r$. *Wtedy wektory*

$$e_{1i_1} \otimes \ldots \otimes e_{ri_r}\,, \qquad i_1 = 1, \ldots, n_1, \quad \ldots, \quad i_r = 1, \ldots, n_r$$

tworzą bazę przestrzeni $V_1 \otimes \ldots \otimes V_r = \mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})$.

Dowód. Liczba wektorów w określonej twierdzeniem rodzinie jest $n_1 \ldots n_r$, więc jest równa wymiarowi przestrzeni $\mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})$. Dla wykazania twierdzenia wystarczy więc przekonać się, że wektory tej rodziny są liniowo niezależne. Załóżmy, że

$$0 = \sum_{i_1, \ldots, i_r} \mu^{i_1 \ldots i_r}\, e_{1i_1} \otimes \ldots \otimes e_{ri_r}\,.$$

Wyliczając wartość odwzorowania po prawej stronie tej równości dla ciągu argumentów $e_1^{j_1}, \ldots, e_r^{j_r}$, gdzie $(e_\alpha^1, \ldots, e_\alpha^{n_\alpha})$ jest bazą dualną do bazy $(e_{\alpha 1}, \ldots, e_{\alpha n_\alpha})$, $\alpha = 1, \ldots, r$, dostajemy

$$0 = \sum_{i_1, \ldots, i_r} \mu^{i_1 \ldots i_r}\, \langle e_1^{j_1}, e_{1i_1} \rangle_1 \ldots \langle e_r^{j_r}, e_{ri_r} \rangle_r = \mu^{j_1 \ldots j_r}$$

dla wszystkich ciągów wskaźników $j_1 \ldots j_r$. Wektory rodziny są więc liniowo niezależne. $\square$

Każdą bazę przestrzeni $V_1 \otimes \ldots \otimes V_r$ utworzoną tak, jak w twierdzeniu, nazywać będziemy ***bazą iloczynową***.

10 Przykłady

(i) Rozkład wektora w przestrzeni iloczynowej w bazie iloczynowej
W rzeczywistej przestrzeni dwuwymiarowej V_1 wybieramy bazę (e_1, e_2), a w rzeczywistej przestrzeni trójwymiarowej V_2 – bazę (f_1, f_2, f_3). Przestrzeń

$$V_1 \otimes V_2 = \mathcal{L}(V_1^*, V_2^*; \mathbb{R})$$

jest sześciowymiarowa, z bazą $e_i \otimes f_j$. Wybieramy w tej przestrzeni wektor Z kładąc

$$Z(\varphi, \psi) = (\varphi_1 + 2\varphi_2)(\psi_1 - \psi_3) + 3(\varphi_1 - 2\varphi_2)(\psi_2 - 2\psi_3)\,,$$

gdzie φ_i i ψ_j są współrzędnymi form φ i ψ w bazach (e^1, e^2) i (f^1, f^2, f^3) odpowiednio. Pamiętamy, że $\varphi_i = \langle \varphi, e_i \rangle_1$, $\psi_j = \langle \psi, f_j \rangle_2$, więc $\varphi_i \psi_j = (e_i \otimes f_j)(\varphi, \psi)$. Stąd

$$Z = e_1 \otimes f_1 + 3e_1 \otimes f_2 - 7e_1 \otimes f_3 + 2e_2 \otimes f_1 - 6e_2 \otimes f_2 + 10e_2 \otimes f_3 .$$

(ii) Funkcja dualności
Funkcja dualności przestrzeni V i V^* jest biliniową funkcją $V^* \times V \mapsto \mathbb{K}$, może więc być interpretowana jako wektor przestrzeni $V \otimes V^*$.

11 Iloczyn tensorowy: ogólna definicja

Omówiony w punktach 8 i 9 model iloczynu tensorowego może być zastosowany do każdego ciągu przestrzeni skończenie wymiarowych. Został on zbudowany na zasadzie analogicznej do prostych modeli omówionych w punkcie 1, które jednak nie są w bezpośredni sposób jego szczególnymi przypadkami. Powstaje pytanie, co łączy te modele i czy można włączyć je w jedną, ogólniejszą definicję iloczynu tensorowego. Okazuje się, że następująca definicja pozwoli zbudować pełną teorię iloczynu tensorowego, również w przypadku przestrzeni o nieskończonym wymiarze.

Niech będą dane przestrzenie $V_1, \ldots, V_r$ nad ciałem $\mathbb{K}$. Przestrzeń V nad tym samym ciałem nazywamy ***iloczynem tensorowym przestrzeni*** $V_1, \ldots, V_r$, i piszemy $V = V_1 \otimes \ldots \otimes V_r$, jeśli

(i) zadane jest odwzorowanie wieloliniowe

$$\otimes : V_1 \times \ldots \times V_r \ni (x_1, \ldots, x_r) \mapsto x_1 \otimes \ldots \otimes x_r \in V ,$$
$$\otimes \in \mathcal{L}(V_1, \ldots, V_r; V) ;$$

(ii) zbiór wektorów $\{e_{1i_1} \otimes \ldots \otimes e_{ri_r} \mid i_\alpha \in I_\alpha,\ \alpha = 1, \ldots, r\}$ jest bazą przestrzeni V dla każdego doboru baz $E_\alpha = \{e_{\alpha i_\alpha} \mid i_\alpha \in I_\alpha\}$ przestrzeni V_α (przestrzenie mogą być nieskończenie wymiarowe, więc odpowiednie zbiory indeksów I_α mogą mieć dowolną moc).

Każdy wektor $x_1 \otimes \ldots \otimes x_r$ nazywać będziemy ***iloczynem tensorowym wektorów*** $x_1, \ldots, x_r$ lub ***wektorem iloczynowym*** w przestrzeni $V_1 \otimes \ldots \otimes V_r$, a bazy $\{e_{1i_1} \otimes \ldots \otimes e_{ri_r}\}$ – ***bazami iloczynowymi*** tej przestrzeni. Bezpośrednio z definicji iloczynu tensorowego wynika, że

$$\dim(V_1 \otimes \ldots \otimes V_r) = \prod_{i=1}^{r} \dim V_i .$$

Z wieloliniowości odwzorowania $\otimes$ każdy wektor iloczynowy ma rozkład postaci

$$x_1 \otimes \ldots \otimes x_r = \sum {x_1}^{i_1} \ldots {x_r}^{i_r}\, e_{1i_1} \otimes \ldots \otimes e_{ri_r}\,, \quad \text{gdzie} \quad x_\alpha = \sum {x_\alpha}^{i_\alpha} e_{\alpha i_\alpha}\,,$$

i wszystkie sumy przebiegają po skończonych zbiorach. Wynika stąd własność iloczynu tensorowego wspólna z iloczynem liczb:

Twierdzenie 6.

$$x_1 \otimes \ldots \otimes x_r = 0 \quad \iff \quad \textit{dla pewnego } \alpha\textit{:} \ \ x_\alpha = 0\,.$$

Dowód. Udowodnimy równoważność zaprzeczenia obu stron. Zaprzeczenie warunku po lewej stronie jest równoważne stwierdzeniu, że ${x_1}^{i_1} \ldots {x_r}^{i_r} \neq 0$ dla pewnego ciągu wskaźników $i_1, \ldots, i_r$, co zachodzi wtedy, i tylko wtedy, gdy wszystkie wektory x_α są różne od zera. □

Model form wieloliniowych

$$V_1 \otimes \ldots \otimes V_r = \mathcal{L}(V_1^*, \ldots, V_r^*; \mathbb{K})$$

spełnia warunki ogólnej definicji: istnieją w nim bazy iloczynowe, a wieloliniowość odwzorowania $(x_1, \ldots, x_r) \mapsto x_1 \otimes \ldots \otimes x_r$ wynika z liniowości w prawych argumentach wszystkich funkcji dualności par przestrzeni V_α^*, V_α:

$$\begin{aligned}
&\big[(\mu x_1 + \mu' x_1') \otimes x_2 \otimes \ldots \otimes x_r\big]\,(\varphi^1, \ldots, \varphi^r) = \\
&\quad = \langle \varphi^1, \mu x_1 + \mu' x_1' \rangle_1 \langle \varphi^2, x_2 \rangle_2 \ldots \langle \varphi^r, x_r \rangle_r = \\
&\qquad = \mu \langle \varphi^1, x_1 \rangle_1 \ldots \langle \varphi^r, x_r \rangle_r + \mu' \langle \varphi^1, x_1' \rangle_1 \ldots \langle \varphi^r, x_r \rangle_r = \\
&\qquad\quad = \mu\,[\,x_1 \otimes \ldots \otimes x_r\,]\,(\varphi^1, \ldots, \varphi^r) + \mu'\,[\,x_1' \otimes \ldots \otimes x_r\,]\,(\varphi^1, \ldots, \varphi^r)\,,
\end{aligned}$$

i podobnie w innych argumentach. Również modele omówione w punkcie 1 spełniają warunki ogólnej definicji (drugi punkt definicyjny sprawdza się prosto dla szczególnych baz mnożonych przestrzeni, skąd ogólne spełnienie warunku wynika na mocy wniosku 8 otrzymanego w następnym punkcie). Przykład szczególnego modelu dla przestrzeni nieskończenie wymiarowych otrzymuje się mnożąc wielomiany dowolnego stopnia dwóch różnych zmiennych. Uniwersalny model dla przestrzeni o nieskończonym wymiarze omawiamy w Uzupełnieniach (p. 10, §24).

Przyjęta ogólna definicja iloczynu tensorowego dopuszcza wiele modeli. Musimy więc jeszcze wyjaśnić, czy stopień pokrewieństwa różnych modeli uzasadnia wspólną nazwę iloczynu tensorowego. Ostatecznym uzasadnieniem przyjętej definicji będzie wykazanie (w następnym punkcie), że wszystkie one są kanonicznie podobne.

12 Iloczyn tensorowy: alternatywna definicja i uniwersalność

Zakładamy w tym punkcie początkowo, że $\otimes : V_1 \times \ldots \times V_r \mapsto V$ jest dowolnym odwzorowaniem wieloliniowym (bez nakładania drugiego warunku definicyjnego iloczynu tensorowego). Rozważmy przestrzeń odwzorowań liniowych z przestrzeni V w dowolną przestrzeń W. Dla każdego $A \in \mathcal{L}(V, W)$ utwórzmy złożenie $A \circ \otimes$, czyli

$$(A \circ \otimes)(x_1, \ldots, x_r) = A(x_1 \otimes \ldots \otimes x_r)\,.$$

Z wieloliniowości odwzorowania $\otimes$ i liniowości A łatwo przekonać się, że złożenie $A \circ \otimes$ jest odwzorowaniem wieloliniowym. Ponadto dla każdego ciągu argumentów (x_i) mamy:

$$\begin{aligned}[(\mu A + \nu B) \circ \otimes](x_1, \ldots, x_r) &= (\mu A + \nu B)(x_1 \otimes \ldots \otimes x_r) = \\ &= \mu A(x_1 \otimes \ldots \otimes x_r) + \nu B(x_1 \otimes \ldots \otimes x_r) = \\ &= \mu(A \circ \otimes)(x_1, \ldots, x_r) + \nu(B \circ \otimes)(x_1, \ldots, x_r)\,,\end{aligned}$$

skąd wynika, że odwzorowanie $\mathcal{L}(V, W) \mapsto \mathcal{L}(V_1, \ldots, V_r; W)$, $A \mapsto A \circ \otimes$ jest liniowe. Poniższe twierdzenie daje teraz równoważne dopełnienie definicji iloczynu tensorowego.

Twierdzenie 7. *Niech $\otimes \in \mathcal{L}(V_1, \ldots, V_r; V)$. Wtedy następujące warunki nałożone na to odwzorowanie są równoważne:*

(i) dla każdej przestrzeni W odwzorowanie

$$\iota : \mathcal{L}(V, W) \mapsto \mathcal{L}(V_1, \ldots, V_r; W)\,, \quad \iota(A) = A \circ \otimes$$

jest izomorfizmem;

(ii) warunek (ii) definicji iloczynu tensorowego.

Dowód. Dopuszczenie przestrzeni nieskończenie wymiarowych jest jedynie niewielką komplikacją, więc przeprowadzimy ogólny dowód. Wybierzmy w dowolny sposób bazy $E_\alpha = \{e_{\alpha i_\alpha}\} \subseteq V_\alpha$ i utwórzmy zbiór $E = \{e_{1i_1} \otimes \ldots \otimes e_{ri_r}\}$. Wybierzmy też dowolną niezerową przestrzeń W nad tym samym ciałem co V. Wykażemy najpierw, że odwzorowanie ι jest injektywne wtedy, i tylko wtedy, gdy E generuje V. Istotnie, jądro odwzorowania ι jest utworzone z odwzorowań liniowych A, dla których $A \circ \otimes = 0$, czyli $A(x_1 \otimes \ldots \otimes x_r) = 0$ dla wszystkich ciągów (x_i). Ale wektory iloczynowe generują podprzestrzeń $L(E)$, więc równoważnym warunkiem na A jest $A\,L(E) = 0$. Zbiór wszystkich odwzorowań A spełniających ten warunek jest równy $\{0\}$ wtedy, i tylko wtedy, gdy $L(E) = V$, co mieliśmy wykazać. Uzupełniamy teraz wynikania w obu kierunkach:

(ii)⇒(i) Jeśli E jest bazą, to wiemy już, że odwzorowanie ι jest injektywne. Wybierzmy dowolnie $\tilde{A} \in \mathcal{L}(V_1, \ldots, V_r; W)$ i określmy A kładąc

$$A(e_{1i_1} \otimes \ldots \otimes e_{ri_r}) = \tilde{A}(e_{1i_1}, \ldots, e_{ri_r})$$

dla wszystkich $e_{\alpha i_\alpha} \in E_\alpha$. Odwzorowanie A rozszerza się do liniowego odwzorowania w $\mathcal{L}(V, W)$, a z wieloliniowości $\tilde{A}$ jest $\tilde{A}(x_1, \ldots, x_r) = A(x_1 \otimes \ldots \otimes x_r)$ dla dowolnych (x_i), więc $\tilde{A} = A \circ \otimes$. Stąd odwzorowanie ι jest surjektywne.
(i)⇒(ii) Jeśli odwzorowanie ι jest izomorfizmem, to wiemy już, że E generuje V. Niech

$$\sum \lambda^{i_1 \ldots i_r} e_{1i_1} \otimes \ldots \otimes e_{ri_r} = 0\,,$$

gdzie sumowanie przebiega po skończonym zbiorze. Działając na to równanie odwzorowaniem $A \in \mathcal{L}(V, W)$ mamy stąd

$$\sum \lambda^{i_1 \ldots i_r} \tilde{A}(e_{1i_1}, \ldots, e_{ri_r}) = 0\,,$$

gdzie $\tilde{A} = A \circ \otimes$ jest dowolnym odwzorowaniem z $\mathcal{L}(V_1, \ldots, V_r; W)$ – z surjektywności ι. W szczególności możemy zażądać, aby jedyną niezerową wartością $\tilde{A}$ na wektorach bazowych był wektor $\tilde{A}(e_{1j_1}, \ldots, e_{rj_r})$. Stąd $\lambda^{j_1 \ldots j_r} = 0$ i podobnie dla pozostałych współczynników. Wektory rodziny E są więc liniowo niezależne. □

Na mocy udowodnionej równoważności widzimy, że warunek (ii) w definicji iloczynu tensorowego w punkcie 11 moglibyśmy zastąpić warunkiem (i) z ostatniego twierdzenia, i w istocie niekiedy tak właśnie definiuje się iloczyn tensorowy.

Z postaci dowodu twierdzenia możemy wyciągnąć dodatkowy wniosek. Zwróćmy uwagę, że w dowodzie wystarczyło w (i) wybrać jedną nietrywialną przestrzeń W aby wykazać (ii) dla dowolnej rodziny E, i odwrotnie, wystarczyło wybrać jedną rodzinę w (ii), aby wykazać (i) dla dowolnej przestrzeni W. Stąd mamy następujące uproszczenie.

Wniosek 8. *Warunek (ii) definicji iloczynu tensorowego wystarczy sprawdzić dla jednego wyboru baz $E_\alpha \subseteq V_\alpha$, $\alpha = 1, \ldots, r$. Równoważnie, wystarczy sprawdzić warunek (i) z ostatniego twierdzenia dla jednej nietrywialnej przestrzeni W.*

W oparciu o udowodnione twierdzenie wykażemy teraz zapowiedziane podobieństwo modeli iloczynu tensorowego.

Twierdzenie 9. *Niech $V_1 \otimes \ldots \otimes V_r$ i $V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r$ będą dwoma modelami iloczynu tensorowego przestrzeni $V_1, \ldots, V_r$, zadanymi odwzorowaniami $\otimes$ i $\tilde{\otimes}$ odpowiednio. Wtedy przepis*

$$V_1 \otimes \ldots \otimes V_r \mapsto V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r\,, \quad x_1 \otimes \ldots \otimes x_r \mapsto x_1 \tilde{\otimes} \ldots \tilde{\otimes} x_r$$

jednoznacznie zadaje izomorfizm tych przestrzeni. Stąd, w szczególności, współrzędne wektorów w bazach iloczynowych zachowują się pod działaniem tego izomorfizmu.

Dowód. Na podstawie twierdzenia 7 (kładziemy $W = V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r$) dla odwzorowania

$$\tilde{\otimes} \in \mathcal{L}(V_1, \ldots, V_r; V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r)$$

istnieje dokładnie jedno odwzorowanie

$$J \in \mathcal{L}(V_1 \otimes \ldots \otimes V_r, V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r) \quad \text{takie, że}$$
$$J \circ \otimes = \tilde{\otimes}\,, \quad \text{czyli} \quad J(x_1 \otimes \ldots \otimes x_r) = x_1 \tilde{\otimes} \ldots \tilde{\otimes} x_r$$

dla wszystkich wektorów iloczynowych. Odwracając rolami $\otimes$ i $\tilde{\otimes}$ znajdujemy odwzorowanie $\tilde{J} \in \mathcal{L}(V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r, V_1 \otimes \ldots \otimes V_r)$, dla którego $\tilde{J} \circ J = \mathrm{id}$ na wektorach iloczynowych w $V_1 \otimes \ldots \otimes V_r$ i $J \circ \tilde{J} = \mathrm{id}$ na wektorach iloczynowych w $V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r$. Wektory iloczynowe generują całe przestrzenie, więc J i $\tilde{J}$ są wzajemnie odwrotne. □

13 Łączność i przemienność iloczynu przestrzeni wektorowych

Rozważmy teraz operację pomnożenia tensorowego dwóch przestrzeni, z których każda jest już iloczynem tensorowym innych przestrzeni. Ściśle rzecz biorąc, przestrzenie $(V_1 \otimes \ldots \otimes V_{r'}) \otimes (V_{r'+1} \otimes \ldots \otimes V_r)$ oraz $V_1 \otimes \ldots \otimes V_r$ są tworami różnych kategorii, jednak pominięcie nawiasów okazuje się naturalnym izomorficznym ich utożsamieniem. Podobnie zmiana kolejności przestrzeni w iloczynie tensorowym nie prowadzi do istotnie nowego obiektu.

Twierdzenie 10. *Istnieją kanoniczne, izomorficzne utożsamienia*

$$(V_1 \otimes \ldots \otimes V_{r'}) \otimes (V_{r'+1} \otimes \ldots \otimes V_r) = V_1 \otimes \ldots \otimes V_r\,,$$
$$V_{\pi(1)} \otimes \ldots \otimes V_{\pi(r)} = V_1 \otimes \ldots \otimes V_r \quad \textit{dla dowolnej permutacji } \pi \in S_r\,,$$

jednoznacznie zadane utożsamieniami wektorów iloczynowych, odpowiednio:

$$(x_1 \otimes \ldots \otimes x_{r'}) \otimes (x_{r'+1} \otimes \ldots \otimes x_r) = x_1 \otimes \ldots \otimes x_r\,,$$
$$x_{\pi(1)} \otimes \ldots \otimes x_{\pi(r)} = x_1 \otimes \ldots \otimes x_r\,.$$

Dowód. Zastąpmy początkowo znak $\otimes$ nowym znakiem $\tilde{\otimes}$ po prawych stronach wszystkich równości pojawiających się w powyższym twierdzeniu. Łatwo sprawdzić, że lewe strony tych równości definiują wtedy poprawnie w każdym z dwóch przypadków nowy model iloczynu tensorowego, zadany odwzorowaniem $\tilde{\otimes}$. Stosując kanoniczne utożsamienie modeli $V_1 \otimes \ldots \otimes V_r$ i $V_1 \tilde{\otimes} \ldots \tilde{\otimes} V_r$ dostajemy tezę. □

Musimy tutaj dodać komentarz do opisanej w tym twierdzeniu przemienności. Zwracamy uwagę, że przy utożsamieniu zadanym permutacją wektory w iloczynach podlegają *jednej i tej samej* zmianie kolejności dla *wszystkich* wektorów iloczynowych. Każda permutacja zadaje inne utożsamienie modeli. Należy tę operację odróżnić od innej, która dyskutowana będzie obszerniej w następnym paragrafie. Mianowicie, jeśli w iloczynie $V_1 \otimes \ldots \otimes V_r$ wszystkie przestrzenie są kopiami tej samej przestrzeni V, to możemy kolejność czynników w wektorach iloczynowych poddawać zmianie w ramach *tego samego modelu*. Wynik takiej operacji jest na ogół wektorem innym od początkowego, więc przemienność w tym sensie nie obowiązuje.

14 Przestrzeń dualna do iloczynu przestrzeni

Ważny szczególny przypadek twierdzenia o uniwersalności otrzymuje się dla przestrzeni skończenie wymiarowych kładąc $W = \mathbb{K}$. Wtedy z definicji przestrzeni dualnej mamy

$$(V_1 \otimes \ldots \otimes V_r)^* = \mathcal{L}(V_1 \otimes \ldots \otimes V_r; \mathbb{K}),$$

a za model iloczynu tensorowego przestrzeni $V_1^*, \ldots, V_r^*$ możemy przyjąć

$$V_1^* \otimes \ldots \otimes V_r^* = \mathcal{L}(V_1, \ldots, V_r; \mathbb{K}),$$

więc przestrzenie $(V_1 \otimes \ldots \otimes V_r)^*$ i $V_1^* \otimes \ldots \otimes V_r^*$ wiążą się kanonicznym izomorfizmem danym twierdzeniem 7.

Twierdzenie 11. *Dla przestrzeni skończenie wymiarowych $V_1, \ldots, V_r$ istnieje kanoniczne izomorficzne utożsamienie*

$$(V_1 \otimes \ldots \otimes V_r)^* = V_1^* \otimes \ldots \otimes V_r^* .$$

Przy tym utożsamieniu funkcja dualności przyjmuje dla wektorów iloczynowych postać

$$\langle \varphi^1 \otimes \ldots \otimes \varphi^r, x_1 \otimes \ldots \otimes x_r \rangle = \langle \varphi^1, x_1 \rangle_1 \ldots \langle \varphi^r, x_r \rangle_r .$$

Dowód. Pozostała do wykazania jedynie postać funkcji dualności. Twierdzenie o uniwersalności wiąże jednoznacznie odwzorowanie

$$\varphi^1 \otimes \ldots \otimes \varphi^r \in \mathcal{L}(V_1, \ldots, V_r; \mathbb{K})$$

z odwzorowaniem

$$\varphi^1 \otimes \ldots \otimes \varphi^r \circ \otimes^{-1} \in \mathcal{L}(V_1 \otimes \ldots \otimes V_r, \mathbb{K}) .$$

Przy utożsamieniu tych przestrzeni mamy więc

$$\langle\varphi^1 \otimes \ldots \otimes \varphi^r, x_1 \otimes \ldots \otimes x_r\rangle = [\,\varphi^1 \otimes \ldots \otimes \varphi^r \circ \otimes^{-1}\,](x_1 \otimes \ldots \otimes x_r) =$$
$$= [\varphi^1 \otimes \ldots \otimes \varphi^r](x_1, \ldots, x_r) = \langle\varphi^1, x_1\rangle_1 \ldots \langle\varphi^r, x_r\rangle_r\,,$$

co kończy dowód. □

15 Iloczyn tensorowy odwzorowań liniowych

Niech przestrzenie V i W będą tensorowymi iloczynami takiej samej liczby przestrzeni wektorowych. Wśród odwzorowań liniowych $V \mapsto W$ istnieje szczególna klasa odwzorowań zgodnych ze strukturami tensorowymi, zadana następującym twierdzeniem.

Twierdzenie 12. *Niech $A_\alpha : V_\alpha \mapsto W_\alpha$, $\alpha = 1, \ldots, r$, będą odwzorowaniami liniowymi przestrzeni wektorowych nad wspólnym ciałem liczbowym. Wtedy istnieje dokładnie jedno odwzorowanie liniowe $A : V_1 \otimes \ldots \otimes V_r \mapsto W_1 \otimes \ldots \otimes W_r$, które na wektory iloczynowe działa według przepisu*

$$A(x_1 \otimes \ldots \otimes x_r) = A_1 x_1 \otimes \ldots \otimes A_r x_r\,.$$

Dowód. Przepis $\tilde{A}(x_1, \ldots, x_r) = A_1x_1 \otimes \ldots \otimes A_r x_r$ zadaje odwzorowanie w przestrzeni odwzorowań wieloliniowych $\mathcal{L}(V_1, \ldots, V_r; W_1 \otimes \ldots \otimes W_r)$. Na mocy tw. 7 istnieje dokładnie jedno odwzorowanie $A \in \mathcal{L}(V_1 \otimes \ldots \otimes V_r, W_1 \otimes \ldots \otimes W_r)$ takie, że $\tilde{A} = A \circ \otimes$, czyli $A(x_1 \otimes \ldots \otimes x_r) = A_1x_1 \otimes \ldots \otimes A_r x_r$. □

Odwzorowanie określone udowodnionym twierdzeniem nazywamy ***iloczynem tensorowym odwzorowań*** $A_1, \ldots, A_r$ i oznaczamy

$$A = A_1 \otimes \ldots \otimes A_r\,.$$

W szczególności, gdy A_α są operatorami liniowymi na V_α, to $A_1 \otimes \ldots \otimes A_r$ jest operatorem działającym w przestrzeni $V_1 \otimes \ldots \otimes V_r$.

Wprost z definicji iloczynu tensorowego odwzorowań wynika, że

$$\begin{array}{ll} \text{jeśli} & A_\alpha : V_\alpha \mapsto W_\alpha\,, \quad B_\alpha : W_\alpha \mapsto Y_\alpha\,, \\ \text{to} & (B_1 \otimes \ldots \otimes B_r)(A_1 \otimes \ldots \otimes A_r) = B_1A_1 \otimes \ldots \otimes B_rA_r\,. \end{array}$$

16 Przykłady

(i) Wektory iloczynowe

Wybierzmy bazę iloczynową $\{e_{1i_1} \otimes \ldots \otimes e_{ri_r}\}$ przestrzeni skończenie wymiarowej $V_1 \otimes \ldots \otimes V_r$. Każdy wektor tej przestrzeni rozkłada się na wektory bazowe,

$$Z = Z^{i_1 \ldots i_r} e_{1i_1} \otimes \ldots \otimes e_{ri_r}\,.$$

Wektor Z jest wektorem iloczynowym wtedy, i tylko wtedy, gdy

$$Z^{i_1\ldots i_r} = x_1{}^{i_1}\ldots x_r{}^{i_r}\,, \quad \text{i wtedy} \quad Z = x_1 \otimes \ldots \otimes x_r\,, \quad \text{gdzie} \quad x_\alpha = x_\alpha{}^{i_\alpha} e_{\alpha i_\alpha}\,.$$

Dla przypadku $r = 2$ można podać proste kryterium: $Z = Z^{i_1 i_2} e_{1i_1} \otimes e_{2i_2}$ jest wektorem iloczynowym wtedy, i tylko wtedy, gdy rząd macierzy współczynników $(Z^{i_1 i_2})$ jest równy 1. Na przykład, niech $X, Y \in V_1 \otimes V_2$,

$$X = e_1 \otimes f_1 + 2e_1 \otimes f_2 - 2e_2 \otimes f_1 - 4e_2 \otimes f_2 + 3e_3 \otimes f_1 + 6e_3 \otimes f_2\,,$$
$$Y = e_1 \otimes f_2 + 2e_2 \otimes f_1 + 4e_3 \otimes f_1 - e_3 \otimes f_2\,,$$

gdzie (e_1, e_2, e_3) jest bazą V_1, a (f_1, f_2) – bazą V_2. Mamy

$$(X^{ij}) = \begin{pmatrix} 1 & 2 \\ -2 & -4 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}\,, \quad \operatorname{rk}(Y^{ij}) = \operatorname{rk} \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 4 & -1 \end{pmatrix} = 2\,,$$

więc $X = (e_1 - 2e_2 + 3e_3) \otimes (f_1 + 2f_2)$, a wektor Y nie jest wektorem iloczynowym.

(ii) Operator odwrotny do iloczynu tensorowego operatorów
Operator identycznościowy na przestrzeni $V_1 \otimes \ldots \otimes V_r$ można zapisać jako $\mathrm{id}_1 \otimes \ldots \otimes \mathrm{id}_r$. Operator odwrotny do $A_1 \otimes \ldots \otimes A_r$ istnieje wtedy, i tylko wtedy, gdy istnieją odwrotne do wszystkich operatorów w iloczynie. Gdy ten warunek jest spełniony, to

$$(A_1 \otimes \ldots \otimes A_r)^{-1} = A_1^{-1} \otimes \ldots \otimes A_r^{-1}\,.$$

(iii) Macierz operatora iloczynowego w bazie iloczynowej
Niech A i B będą operatorami w skończenie wymiarowych przestrzeniach V i W odpowiednio. Wybierzmy bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ tych przestrzeni i niech $Ae_i = e_k A^k{}_i$, $Bf_j = f_l\, B^l{}_j$. Wtedy

$$(A \otimes B)(e_i \otimes f_j) = (e_k \otimes f_l) A^k{}_i B^l{}_j\,.$$

Jeśli bazę $(e_i \otimes f_j)$ ustawić w ciąg, np. zgodnie z porządkiem leksykograficznym par (i, j), to zespół współczynników $(A^k{}_i B^l{}_j)$ tworzy macierz operatora $A \otimes B$ w tej bazie.

(iv) Iloczyn tensorowy operatorów diagonalizowalnych
Przy oznaczeniach poprzedniego przykładu niech A i B będą diagonalizowalne i niech $Ae_i = \mu_i e_i$, $Bf_j = \nu_j f_j$. Wtedy $e_i \otimes f_j$ jest bazą własną iloczynu tensorowego operatorów:

$$(A \otimes B)(e_i \otimes f_j) = \mu_i \nu_j\, (e_i \otimes f_j)\,.$$

Iloczyn tensorowy operatorów diagonalizowalnych jest diagonalizowalny.

(v) Wyznacznik iloczynu tensorowego operatorów
Przy oznaczeniach przykładu (iii) zachodzi tożsamość:

$$\det(A \otimes B) = (\det A)^{\dim W} (\det B)^{\dim V} .$$

Korzystając z rozkładu $A \otimes B = (A \otimes \mathrm{id}_W)(\mathrm{id}_V \otimes B)$ oraz z twierdzenia Cauchy'ego dla wyznaczników, wystarczy pokazać, że $\det(A \otimes \mathrm{id}_W) = (\det A)^{\dim W}$. W tym celu wybierzmy dowolne bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ przestrzeni V i W odpowiednio. Niech $Ae_i = e_j A^j{}_i$. Wtedy $(A \otimes \mathrm{id}_W)(e_i \otimes f_k) = (e_j \otimes f_k) A^j{}_i$. Stąd każda podprzestrzeń $V \otimes L(f_k) = \{x \otimes f_k \mid x \in V\}$ jest inwariantna względem $A \otimes \mathrm{id}$ i na każdej z nich macierz zacieśnienia tego operatora w danej bazie jest równa $\mathbf{A}$, co prowadzi do żądanej tożsamości.

17 Kontrakcja

Rozważamy w tym punkcie przypadek, gdy wśród przestrzeni skończenie wymiarowych $V_1, \ldots, V_r$ znajduje się pewna przestrzeń V oraz dualna do niej V^*. Założymy dla uproszczenia notacji, że zajmują one dwa pierwsze miejsca w iloczynie, a pozostałe przestrzenie można wtedy objąć wspólnym oznaczeniem W. Podobne wyniki obowiązują jednak również dla dowolnego położenia przestrzeni V i V^* w ciągu $V_1, \ldots, V_r$.

Wybierzmy w przestrzeniach V i W bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ odpowiednio. Tworzymy przestrzeń $V \otimes V^* \otimes W$ i wybieramy w niej bazę $e_i \otimes e^j \otimes f_k$. Każdy wektor Z w tej przestrzeni ma więc rozkład $Z = Z^i{}_j{}^k \, e_i \otimes e^j \otimes f_k$, gdzie $Z^i{}_j{}^k$ są jego współrzędnymi w tej bazie.

Twierdzenie 13. *Istnieje dokładnie jedno odwzorowanie liniowe*

$$C : V \otimes V^* \otimes W \mapsto W ,$$

które na wektory iloczynowe działa według przepisu

$$C(x \otimes \varphi \otimes y) = \langle \varphi, x \rangle \, y .$$

Jeśli $Z = Z^i{}_j{}^k \, e_i \otimes e^j \otimes f_k$, to $C(Z) = Z^i{}_i{}^k \, f_k$ (sumowanie po i zgodnie z konwencją Einsteina).

Dowód. Przepis $\tilde{C}(x, \varphi, y) = \langle \varphi, x \rangle \, y$ zadaje odwzorowanie z przestrzeni odwzorowań wieloliniowych $\mathcal{L}(V, V^*, W; W)$. Z uniwersalności iloczynu tensorowego, tw. 7, istnieje dokładnie jedno odwzorowanie liniowe $C \in \mathcal{L}(V \otimes V^* \otimes W; W)$ takie, że $\tilde{C} = C \circ \otimes$, czyli $C(x \otimes \varphi \otimes y) = \langle \varphi, x \rangle \, y$. Dla wektorów bazowych mamy stąd $C(e_i \otimes e^j \otimes f_k) = \delta^j_i f_k$, więc z liniowości $C(Z) = Z^i{}_j{}^k \delta^j_i f_k = Z^i{}_i{}^k \, f_k$. □

Każde odwzorowanie zdefiniowane wykazanym twierdzeniem nazywamy ***kontrakcją*** (lub ***zwężeniem***).

18 Iloczyn tensorowy $W \otimes V^*$ jako przestrzeń odwzorowań liniowych. Odwzorowanie transponowane

Składając operację mnożenia tensorowego i kontrakcji otrzymamy w tym punkcie dwa dalsze kanoniczne izomorfizmy przestrzeni wektorowych.

Niech będą dane dwie przestrzenie skończenie wymiarowe V i W nad ciałem $\mathbb{K}$ i wybierzmy w nich bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ odpowiednio. Tworzymy przestrzeń $W \otimes V^*$ i wybieramy w niej bazę $f_j \otimes e^i$. Każdy wektor S z tej przestrzeni ma więc rozkład $S = S^j{}_i\, f_j \otimes e^i$, gdzie $S^j{}_i$ są jego współrzędnymi w tej bazie.

Twierdzenie 14. *Istnieją kanoniczne, izomorficzne utożsamienia*

$$W \otimes V^* = \mathcal{L}(V, W)\,, \quad W \otimes V^* = \mathcal{L}(W^*, V^*)\,,$$

jednoznacznie zadane działaniem wektorów iloczynowych jako odwzorowań liniowych, odpowiednio:

$$(y \otimes \varphi)\, x := \langle \varphi, x \rangle_V\, y\,, \quad (y \otimes \varphi)^{\mathrm{T}}\, \psi := \langle \psi, y \rangle_W\, \varphi\,.$$

gdzie $\langle .\,, . \rangle_V$ i $\langle .\,, . \rangle_W$ są funkcjami dualności par V, V^* i W, W^* odpowiednio. Działanie dowolnego $S \in W \otimes V^*$, o rozkładzie

$$S = S^j{}_i\, f_j \otimes e^i\,,$$

ma wtedy postać

$$Se_i = f_j\, S^j{}_i\,, \quad S^{\mathrm{T}} f^k = S^k{}_l e^l\,.$$

Dowód. Mnożąc tensorowo $S \in W \otimes V^*$ i $x \in V$ dostajemy $S \otimes x \in W \otimes V^* \otimes V$. Dokonując teraz kontrakcji w parze przestrzeni V i V^* dostajemy wektor z przestrzeni W, przy czym obie składane operacje są liniowe. Dla $S = y \otimes \varphi$ dostajemy w wyniku tego złożenia wektor $\langle \varphi, x \rangle_V y$. Podobnie mnożąc $\psi \in W^*$ przez S dostajemy $\psi \otimes S \in W^* \otimes W \otimes V^*$. Dokonując kontrakcji w parze przestrzeni W i W^* dostajemy wektor z przestrzeni V^*. Dla $S = y \otimes \varphi$ otrzymujemy w wyniku formę $\langle \psi, y \rangle_W \varphi$.

Wykazaliśmy, że zadane w twierdzeniu przepisy określają odwzorowania liniowe $W \otimes V^* \mapsto \mathcal{L}(V, W)$ i $W \otimes V^* \mapsto \mathcal{L}(W^*, V^*)$, a ich jednoznaczność wynika z generowania przez wektory iloczynowe pełnych przestrzeni. Działanie operatorów zadanych wektorami bazowymi na wektory bazowe ma postać

$$(f_j \otimes e^i)\, e_k = \delta^i_k\, f_j\,, \quad (f_j \otimes e^i)^{\mathrm{T}}\, f^l = \delta^l_j\, e^i\,,$$

skąd przez liniowość dostajemy przepisy działania dowolnego $S \in W \otimes V^*$. Ponieważ zespoły liczb $S^j{}_i$ są dowolne, dostajemy w wyniku wszystkie odwzorowania z $\mathcal{L}(V, W)$ i $\mathcal{L}(W^*, V^*)$ odpowiednio. Przestrzenie $W \otimes V^*$, $\mathcal{L}(V, W)$

i $\mathcal{L}(W^*, V^*)$ mają równe wymiary, więc z surjektywności wynika bijektywność, co dowodzi izomorficzności otrzymanych odwzorowań. Izomorfizmy te przyjmujemy za utożsamienie. □

Przy oznaczeniach powyższego twierdzenia wprowadźmy macierz $\mathbf{S} = (S^j{}_i)$ utworzoną ze współrzędnych wektora $S \in W \otimes V^*$. Twierdzenie mówi nam, że macierzą operatora $S : V \mapsto W$ jest macierz $\mathbf{S}$, a macierzą operatora $S^{\mathrm{T}} : W^* \mapsto V^*$ jest macierz transponowana $\mathbf{S}^{\mathrm{T}}$. Zwróćmy uwagę na jeszcze jeden aspekt twierdzenia: pokazuje ono, że istnieje kanoniczny izomorfizm $\mathcal{L}(V, W) \mapsto \mathcal{L}(W^*, V^*)$ przypisujący każdemu odwzorowaniu $S \in \mathcal{L}(V, W)$ odwzorowanie $S^{\mathrm{T}} \in \mathcal{L}(W^*, V^*)$, które nazywa się ***odwzorowaniem transponowanym do*** S, o macierzy transponowanej do macierzy odwzorowania S.

19 Iloczyn tensorowy przestrzeni unitarnych lub euklidesowych

Wykażemy w tym punkcie, że iloczyn tensorowy przestrzeni unitarnych (lub euklidesowych) jest przestrzenią unitarną (euklidesową), z kanonicznie określonym iloczynem skalarnym.

Twierdzenie 15. *Niech V_α, $\alpha = 1, \ldots, r$, będą przestrzeniami unitarnymi lub euklidesowymi, każda wyposażona we własny iloczyn skalarny $(\,.\,,\,.\,)_\alpha$. Wtedy istnieje dokładnie jeden dodatnio określony iloczyn skalarny $(\,.\,,\,.\,)$ na przestrzeni $V_1 \otimes \ldots \otimes V_r$, który dla wektorów iloczynowych przyjmuje wartość*

$$(x_1 \otimes \ldots \otimes x_r\,,\, y_1 \otimes \ldots \otimes y_r) = (x_1, y_1)_1 \ldots (x_r, y_r)_r\,.$$

Dowód. Dowód prowadzimy dla przestrzeni unitarnych, dla euklidesowych należy jedynie pominąć sprzężenie. Przepis

$$B(x_1, \ldots, x_r; y_1, \ldots, y_r) = (x_1, y_1)_1 \ldots (x_r, y_r)_r$$

zadaje przy ustalonych x-ach funkcję wieloliniową w y-kach, a jego sprzężenie zespolone zadaje przy ustalonych y-kach funkcję wieloliniową w x-ach. Z twierdzenia o uniwersalności iloczynu tensorowego, tw. 7, istnieje więc forma półtoraliniowa na $V_1 \otimes \ldots \otimes V_r$ jednoznacznie zadana warunkiem określonym w tezie. Forma ta w oczywisty sposób spełnia warunek hermitowskości dla wektorów iloczynowych, a stąd dla wszystkich wektorów. Aby wykazać dodatnią określoność formy, obliczamy jej wartość dla bazy iloczynowej utworzonej z baz ortonormalnych przestrzeni V_α:

$$(e_{1i_1} \otimes \ldots \otimes e_{ri_r}\,,\, e_{1j_1} \otimes \ldots \otimes e_{rj_r}) = \delta_{i_1 j_1} \ldots \delta_{i_r j_r}\,.$$

Baza $e_{1i_1} \otimes \ldots \otimes e_{ri_r}$ jest więc bazą kanoniczną tego iloczynu, a sam iloczyn jest dodatnio określony. □

Z określenia iloczynu skalarnego danego powyższym twierdzeniem wykazuje się, że dla iloczynu tensorowego operatorów mamy

$$(A_1 \otimes \ldots \otimes A_r)^* = A_1^* \otimes \ldots \otimes A_r^*$$

– wynika to z przekształcenia wyrażenia $(x_1 \otimes \ldots \otimes x_r, (A_1 \otimes \ldots \otimes A_r)\, y_1 \otimes \ldots \otimes y_r)$ zgodnie z definicją iloczynu tensorowego operatorów i zastosowania sprzężenia operatorowego do operatorów A_i, szczegóły pozostawiamy czytelnikowi jako ćwiczenie. Stąd, w szczególności: iloczyn tensorowy operatorów samosprzężonych (izometrycznych, normalnych) jest operatorem samosprzężonym (odpowiednio: izometrycznym, normalnym).

20 Przykłady

(i) Baza przestrzeni odwzorowań liniowych
Niech A będzie odwzorowaniem liniowym z przestrzeni V w przestrzeń W. W punkcie 11, §10, określiliśmy, z pomocą baz $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ przestrzeni V i W odpowiednio, bazowe odwzorowania $E^i{}_j$. Rozkład A ma w tej bazie postać $A = A^j{}_i E^i{}_j$, gdzie $(A^j{}_i)$ jest macierzą odwzorowania A w zadanych bazach. Porównując teraz ten wynik z działaniem elementów przestrzeni $W \otimes V^*$ jako odwzorowań z $\mathcal{L}(V, W)$ (tw. 14) widzimy, że również $A = A^j{}_i f_j \otimes e^i$, więc

$$E^i{}_j = f_j \otimes e^i\,.$$

Istotnie, potwierdza to bezpośredni rachunek:

$$(f_j \otimes e^i)x = \langle e^i, x\rangle f_j = x^i f_j = E^i{}_j x\,.$$

(ii) Podstawowa definicja odwzorowania transponowanego
Oznaczmy funkcje dualności par przestrzeni V i V^* oraz W i W^* przez $\langle\,.\,,\,.\,\rangle_V$ oraz $\langle\,.\,,\,.\,\rangle_W$ odpowiednio. Dla każdego odwzorowania liniowego $A : V \mapsto W$ istnieje dokładnie jedno odwzorowanie $A^{\mathrm{T}} : W^* \mapsto V^*$ zadane warunkiem

$$\langle \psi, Ax\rangle_W = \langle A^{\mathrm{T}}\psi, x\rangle_V$$

i nazywane odwzorowaniem transponowanym. Dowód tego faktu jest podobny do dowodu istnienia operatora sprzężonego do A względem niezdegenerowanego iloczynu skalarnego: odwzorowanie $V \ni x \mapsto \langle \psi, Ax\rangle_W \in \mathbb{K}$ jest formą liniową na V, którą oznaczamy $A^{\mathrm{T}}\psi$, i dowodzimy, że A^{T} jest liniowe. Zauważmy jednak, że dla określenia operatora sprzężonego potrzebowaliśmy istnienia izomorfizmu pomiędzy V i $\mathcal{L}(V, \mathbb{K})$ zadanego przez iloczyn skalarny. Pojęcie to zależy zatem od tego iloczynu, a ponadto dla przestrzeni nieskończenie wymiarowych izomorfizm może nie obowiązywać. Pojęcie operatora transponowanego jest natomiast uniwersalne.

Jeśli przestrzenie V i W mają skończony wymiar, to wybierzmy w nich bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, f_m)$ odpowiednio, oraz dualne do nich bazy w V^* i W^*. Jeśli $\mathbf{A}$ jest macierzą odwzorowania A w bazach (e_i), (f_j), to odwzorowanie transponowane A^{T} ma w bazach (f^j), (e^i) macierz $\mathbf{A}^{\mathrm{T}}$. Dyskusja odwzorowania transponowanego zamieszczona w punkcie 18 jest równoważna tym stwierdzeniom.

(iii) Iloczyn tensorowy operatorów hermitowskich
Niech V i W będą dwuwymiarowymi przestrzeniami unitarnymi, w których wybieramy bazy ortonormalne (e_1, e_2) i (f_1, f_2) odpowiednio. Wektory $e_i \otimes f_j$ tworzą bazę ortonormalną przestrzeni unitarnej $V \otimes W$. Niech A_1, A_2, A_3 będą operatorami w V, a B_1, B_2, B_3 – operatorami w W, takimi, że ich macierze w zadanych bazach są równe

$$\mathbf{A}_i = \mathbf{B}_i = \sigma_i \,,$$

gdzie σ_i są podanymi w przykładzie (iii), p. 17, §10, macierzami Pauliego. Macierze Pauliego są hermitowskie, więc wszystkie zadane operatory są samosprzężone. Tworzymy samosprzężony operator

$$C = \sum_{i=1}^{3} A_i \otimes B_i \,.$$

Chcemy znaleźć ortonormalną bazę własną operatora C.

Znając macierze operatorów A_i, B_j wiemy, jak operatory te działają na wektory bazowe. Posługując się definicją działania iloczynu tensorowego operatorów na iloczyn wektorów wyliczamy działanie operatora C na bazę iloczynową $\{e_i \otimes f_j\}$:

$$C(e_1 \otimes f_1) = e_1 \otimes f_1 \,, \quad C(e_2 \otimes f_2) = e_2 \otimes f_2 \,,$$
$$C(e_1 \otimes f_2) = -e_1 \otimes f_2 + 2e_2 \otimes f_1 \,, \quad C(e_2 \otimes f_1) = 2e_1 \otimes f_2 - e_2 \otimes f_1 \,.$$

Stąd $e_1 \otimes f_1$ i $e_2 \otimes f_2$ są wektorami własnymi, a macierz zacieśnienia C do ortogonalnej do nich podprzestrzeni jest równa $\left(\begin{smallmatrix} -1 & 2 \\ 2 & -1 \end{smallmatrix}\right)$ w bazie uporządkowanej $(e_1 \otimes f_2, e_2 \otimes f_1)$. Rozwiązując w standardowy sposób zagadnienie własne dostajemy pozostałe dwa unormowane wektory własne:

$$C\frac{1}{\sqrt{2}}(e_1 \otimes f_2 + e_2 \otimes f_1) = \frac{1}{\sqrt{2}}(e_1 \otimes f_2 + e_2 \otimes f_1) \,,$$
$$C\frac{1}{\sqrt{2}}(e_1 \otimes f_2 - e_2 \otimes f_1) = -3\frac{1}{\sqrt{2}}(e_1 \otimes f_2 - e_2 \otimes f_1) \,.$$

§19 Tensory

1 Przestrzenie tensorowe

Szczególny, ale niezmiernie ważny przypadek przestrzeni iloczynowej otrzymuje się mnożąc tensorowo kilka kopii jednej przestrzeni wektorowej i jej dualnej. Algebrze takich przestrzeni poświęcone są dwa kolejne paragrafy. Rozpatrujemy w nich wyłącznie przestrzenie o skończonym wymiarze.

Przestrzenią tensorów o walencji $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$ ***nad przestrzenią*** V nazywamy przestrzeń wektorową

$$\mathcal{T}^p_q(V) := \underbrace{V \otimes \ldots \otimes V}_{p \text{ razy}} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{q \text{ razy}} \equiv V^{\otimes p} \otimes V^{*\otimes q}\,, \quad p + q \geq 1\,,$$

gdzie druga równość wprowadza wygodną notację. Wygodnie jest ponadto przyjąć, że

$$\mathcal{T}^0_0(V) = \mathbb{K}\,.$$

Z ogólnych własności iloczynu tensorowego mamy

$$\dim \mathcal{T}^p_q(V) = (\dim V)^{p+q}\,.$$

Każdy element przestrzeni $\mathcal{T}^p_q(V)$ nazywamy ***tensorem o walencji*** $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$. Tensory, które są wektorami iloczynowymi jako wektory przestrzeni $\mathcal{T}^p_q(V)$, nazywamy ***tensorami prostymi***.

Przy konstrukcji baz iloczynowych przestrzeni $\mathcal{T}^p_q(V)$ stosować będziemy zawsze zasadę, że w każdej kopii przestrzeni V wybieramy tę samą bazę, a w każdej kopii przestrzeni V^* – bazę do niej dualną. Mówiąc o bazach iloczynowych ograniczamy się więc do baz postaci

$$e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}\,, \quad \langle e^j, e_i \rangle = \delta^j_i\,,$$

gdzie $(e_1, \ldots, e_n)$ jest dowolną bazą przestrzeni V. ***Współrzędnymi*** (lub ***składowymi***) ***tensora*** t ***w bazie*** (e_i) nazywamy w skrócie współrzędne (składowe) jego rozkładu, jako wektora przestrzeni $\mathcal{T}^p_q(V)$, w bazie iloczynowej zadanej bazą (e_i):

$$t = t^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}\, e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}\,.$$

Stosujemy tu jak zwykle konwencję sumacyjną Einsteina, która wyznacza położenie wskaźników współrzędnych.

Analiza wielu własności tensorów będzie ułatwiona przy zastosowaniu modelu iloczynu tensorowego określonego formami wieloliniowymi. Dla przestrzeni tensorowych oznacza to położenie

$$\mathcal{T}^p_q(V) = \mathcal{L}(\underbrace{V^*, \ldots, V^*}_{p \text{ razy}}, \underbrace{V, \ldots, V}_{q \text{ razy}}; \mathbb{K})\,.$$

Obliczmy wartość tensora $t \in \mathcal{T}^p_q(V)$ w tym modelu dla argumentów bazowych. Z definicji wektorów iloczynowych, p. 8, §18, i relacji dualności baz mamy

$$[e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}](e^{k_1}, \ldots, e^{k_p}; e_{l_1}, \ldots, e_{l_q}) = \delta^{k_1}_{i_1} \ldots \delta^{k_p}_{i_p} \delta^{j_1}_{l_1} \ldots \delta^{j_q}_{l_q} \,,$$

więc współrzędne tensora w bazie iloczynowej mają postać

$$t^{k_1 \ldots k_p}{}_{l_1 \ldots l_q} = t(e^{k_1}, \ldots, e^{k_p};\, e_{l_1}, \ldots, e_{l_q}) \,.$$

Przy ich użyciu wartość odwzorowania t dla dowolnych argumentów wylicza się z wieloliniowości:

$$t(\varphi^1, \ldots, \varphi^p; x_1, \ldots, x_q) = t^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \varphi^1{}_{i_1} \ldots \varphi^p{}_{i_p} x_1{}^{j_1} \ldots x_q{}^{j_q} \,.$$

2 Transformacje współrzędnych (składowych) przy zmianie bazy. Klasyczna definicja tensora

Zbadamy w tym punkcie zmianę współrzędnych tensora przy zmianie bazy. Wartość współrzędnych w bazach iloczynowych nie zależy od modelu iloczynu tensorowego, więc możemy nadal stosować model wieloliniowych odwzorowań.

Niech będą dane dwie bazy $(e_1, \ldots, e_n)$ i $(e'_1, \ldots, e'_n)$ przestrzeni V, związane transformacją $e'_i = e_j \beta^j{}_i$. Wtedy dla baz dualnych $(e^1, \ldots, e^n)$ i $(e'^1, \ldots, e'^n)$ mamy

$$\langle e'^i, e_j \rangle = \langle e'^i, e'_k \, \beta^{-1k}{}_j \rangle = \beta^{-1i}{}_j = \beta^{-1i}{}_l \langle e^l, e_j \rangle = \langle \beta^{-1i}{}_l e^l, e_j \rangle$$

dla każdego wektora e_j, więc $e'^i = \beta^{-1i}{}_l \, e^l$. Podstawiamy teraz transformacje baz w argumentach wyrażenia $t'^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = t(e'^{i_1}, \ldots, e'^{i_p}; e'_{j_1}, \ldots, e'_{j_q})$, korzystamy z wieloliniowości, i otrzymujemy pierwszą część następującego twierdzenia.

Twierdzenie 1. *Transformacja par baz wzajemnie dualnych (e_i) i (e^j) oraz (e'_i) i (e'^j), ma postać*

$$e'_i = e_k \beta^k{}_i \,, \quad e'^j = \beta^{-1j}{}_l e^l \,,$$

czyli symbolicznie

$$\begin{pmatrix} e'_1 \ldots e'_n \end{pmatrix} = \begin{pmatrix} e_1 \ldots e_n \end{pmatrix} \boldsymbol{\beta} \,, \qquad \begin{pmatrix} e'^1 \\ \vdots \\ e'^n \end{pmatrix} = \boldsymbol{\beta}^{-1} \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} \,,$$

i pociąga transformację współrzędnych (składowych) tensora

$$t'^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = \beta^{-1i_1}{}_{k_1} \ldots \beta^{-1i_p}{}_{k_p} t^{k_1 \ldots k_p}{}_{l_1 \ldots l_q} \beta^{l_1}{}_{j_1} \ldots \beta^{l_q}{}_{j_q} \,.$$

Odwrotnie, jeśli z każdą bazą (e_i) związać tablicę liczb $t^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}$ transformujących się przy zmianie bazy w powyższy sposób, to te tablice będą współrzędnymi tensora, jednoznacznie określonego tym warunkiem.

Dowód. Pozostaje do wykazania jedynie ostatnia część twierdzenia. Załóżmy, że z każdą bazą związano tablicę liczb, które transformują się między bazami w zadany sposób. Wybierzmy dowolną bazę (e_i) i określmy tensor t wybierając liczby tablicy $t^{i_1\dots i_p}{}_{j_1\dots j_q}$ związanej z tą bazą jako jego współrzędne w tej bazie. Tensor taki jest określony jednoznacznie tym warunkiem. Na podstawie pierwszej części twierdzenia jego współrzędne w dowolnej innej bazie będą się również pokrywać z tablicą przypisaną tej bazie. □

Określenie tensora jako zespołu tablic liczbowych przypisanych bazom, transformujących się w zadany powyższym twierdzeniem sposób, nazywa się ***klasyczną definicją tensora***. Jest ona równoważna, jak pokazuje ostatnie twierdzenie, bardziej nowoczesnemu ujęciu przedstawionemu wcześniej, ale nie daje zrozumienia inwariantnego znaczenia definiowanego pojęcia. Niemniej jednak, notowanie niektórych własności tensorów jest znacznie ułatwione przy użyciu ich współrzędnych. Górne indeksy współrzędnych tensora nazywa się ***wskaźnikami kontrawariantnymi tensora***, a dolne – ***wskaźnikami kowariantnymi***. Tak więc tensor o walencji $\left[\begin{smallmatrix}p\\q\end{smallmatrix}\right]$ ma p wskaźników kontrawariantnych i q wskaźników kowariantnych.

3 Przykłady

(i) Tensory o walencji $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ oraz tensory o walencji $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$
Tensory pierwszego rodzaju to wektory przestrzeni V, a tensory drugiego rodzaju to wektory przestrzeni V^*. Nazywa się niekiedy te pierwsze ***wektorami kontrawariantnymi***, a drugie – ***wektorami kowariantnymi***. Współrzędne tych tensorów pokrywają się z ich współrzędnymi wektorowymi w przestrzeni V i V^* odpowiednio.

(ii) Tensory o walencji $\left[\begin{smallmatrix}0\\2\end{smallmatrix}\right]$
Tensory te to wektory przestrzeni $\mathcal{T}^0_2(V) = V^* \otimes V^*$. Wybierzmy dla tej przestrzeni model odwzorowań wielo- (w tym przypadku dwu-) liniowych, $\mathcal{T}^0_2(V) = \mathcal{L}(V,V;\mathbb{K})$. Metryki biliniowe g są tensorami z tej przestrzeni. Elementy macierzy metryki, $g_{ij} = g(e_i,e_j)$, pokrywają się ze współrzędnymi metryki jako tensora:

$$g = g_{ij} e^i \otimes e^j \,.$$

(iii) Tensory o walencji $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$
Tensory o tej walencji to wektory przestrzeni $\mathcal{T}^1_1(V) = V \otimes V^*$. Wyniki punktu 18, §18, pokazują, że tensory te można utożsamić z operatorami liniowymi w przestrzeni V, a także z operatorami liniowymi w przestrzeni V^*

(w twierdzeniu 14 z tego punktu należy podstawić $W = V$). Przy tej interpretacji tensory bazowe działają według przepisów

$$(e_i \otimes e^j)\, e_k = \delta^j_k\, e_i\,, \quad (e_i \otimes e^j)^{\mathrm{T}}\, e^l = \delta^l_i\, e^j$$

w przestrzeni V lub V^* odpowiednio. Rozkład dowolnego tensora $t \in \mathcal{T}^1_1(V)$ oraz jego działanie jako operatora w przestrzeniach V i V^* ma postać odpowiednio

$$t = t^i{}_j\, e_i \otimes e^j\,, \quad t\, e_i = e_j\, t^j{}_i\,, \quad t^{\mathrm{T}} e^j = e^k\, t^j{}_k\,.$$

W modelu wieloliniowych odwzorowań $\mathcal{T}^1_1(V) = \mathcal{L}(V^*, V; \mathbb{K})$ określmy szczególny tensor

$$\delta(\varphi, x) = \langle \varphi, x\rangle\,, \quad \text{więc} \quad \delta^i{}_j = \delta(e^i, e_j) = \delta^i_j\,,$$

gdzie po prawej stronie ostatniej równości stoi symbol Kroneckera. Rozkład tego tensora w dowolnej bazie iloczynowej ma więc postać

$$\delta = e_i \otimes e^i\,,$$

a jego działanie jako operatora jest identycznością.

(iv) Transformacja współrzędnych tensora

Zmianę współrzędnych tensora przy zmianie bazy można wyliczyć z ogólnego wzoru zawartego w twierdzeniu 1, ale często wygodniej jest posłużyć się jawnym rozkładem tensora w bazie. Niech t będzie tensorem z przestrzeni $\mathcal{T}^2_1(V)$ nad trójwymiarową przestrzenią rzeczywistą V, danym rozkładem w bazie (e_1, e_2, e_3):

$$t = 2e_1 \otimes e_2 \otimes e^3 - e_2 \otimes e_2 \otimes e^1\,.$$

Można stąd odczytać, że jedynymi nieznikającymi współrzędnymi tensora t w danej bazie są:

$$t^{12}{}_3 = 2\,, \quad t^{22}{}_1 = -1\,.$$

Niech (e'_1, e'_2, e'_3) będzie nową bazą, zadaną tak, jak w twierdzeniu 1 za pomocą macierzy

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{skąd} \quad \boldsymbol{\beta}^{-1} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}.$$

Odwracając relacje wiążące obie bazy wyliczamy

$$\begin{aligned} e_1 &= e'_1\,, & e_2 &= 2e'_1 + e'_2\,, & e_3 &= 4e'_1 + 2e'_2 - e'_3\,, \\ e^1 &= e'^1 - 2e'^2\,, & e^2 &= e'^2 + 2e'^3\,, & e^3 &= -e'^3\,. \end{aligned}$$

Użycie tych związków prowadzi do rozkładu tensora t w nowej bazie:

$$t = -2e'_1 \otimes (2e'_1 + e'_2) \otimes e'^3 - (2e'_1 + e'_2) \otimes (2e'_1 + e'_2) \otimes (e'^1 - 2e'^2)\,.$$

Stąd łatwo odczytać wszystkie współrzędne w nowej bazie, np. $t'^{11}{}_2 = 8$.

4 Kombinacja liniowa tensorów

W trzech kolejnych punktach operacje znane nam już dla dowolnych przestrzeni iloczynowych zapisujemy dla szczególnego przypadku przestrzeni tensorowych.

Tensory o ustalonej walencji $[\substack{p\\q}]$ tworzą przestrzeń wektorową $\mathcal{T}^p_q(V)$, więc jest dla nich określona operacja kombinacji liniowej. W modelu odwzorowań wieloliniowych kombinacja liniowa tensorów $t, s \in \mathcal{T}^p_q(V)$ jest odwzorowaniem $\alpha t + \beta s \in \mathcal{T}^p_q(V)$, o przepisie

$$(\alpha t + \beta s)(\psi^1, \ldots, \psi^p; z_1, \ldots, z_q) = \\ = \alpha\, t(\psi^1, \ldots, \psi^p; z_1, \ldots, z_q) + \beta\, s(\psi^1, \ldots, \psi^p; z_1, \ldots, z_q)\,.$$

Ta sama operacja zapisana we współrzędnych tensorów ma postać

$$(\alpha t + \beta s)^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = \alpha\, t^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} + \beta\, s^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}\,.$$

5 Iloczyn tensorowy tensorów

Rozważmy teraz mnożenie tensorowe przestrzeni tensorowych. Iloczyn przestrzeni $\mathcal{T}^p_q(V) = V^{\otimes p} \otimes V^{*\otimes q}$ i przestrzeni $\mathcal{T}^{p'}_{q'}(V) = V^{\otimes p'} \otimes V^{*\otimes q'}$ jest przestrzenią

$$\mathcal{T}^p_q(V) \otimes \mathcal{T}^{p'}_{q'}(V) = (V^{\otimes p} \otimes V^{*\otimes q}) \otimes (V^{\otimes p'} \otimes V^{*\otimes q'})\,.$$

Do prawej strony możemy zastosować utożsamienia opisane twierdzeniem 10, §18: pominięcie nawiasów oraz zamiana kolejności $V^{*\otimes q} \otimes V^{\otimes p'} = V^{\otimes p'} \otimes V^{*\otimes q}$. Łącznie dostajemy więc utożsamienie

$$\mathcal{T}^p_q(V) \otimes \mathcal{T}^{p'}_{q'}(V) = \mathcal{T}^{p+p'}_{q+q'}(V)\,,$$

zadane jednoznacznie przepisem określonym dla dowolnych tensorów prostych:

$$(x_1 \otimes \ldots \otimes x_p \otimes \varphi^1 \otimes \ldots \otimes \varphi^q) \otimes (y_1 \otimes \ldots \otimes y_{p'} \otimes \psi^1 \otimes \ldots \otimes \psi^{q'}) = \\ = x_1 \otimes \ldots \otimes x_p \otimes y_1 \otimes \ldots \otimes y_{p'} \otimes \varphi^1 \otimes \ldots \otimes \varphi^q \otimes \psi^1 \otimes \ldots \otimes \psi^{q'}\,.$$

Zauważmy, że przy tym utożsamieniu nie zmieniamy następstwa w grupach wektorów: kontrawariantnych $x_1, \ldots, x_p, y_1, \ldots, y_q$, ani kowariantnych $\varphi^1, \ldots, \varphi^q, \psi^1, \ldots, \psi^{q'}$, a jedynie przesuwamy wszystkie wektory kontrawariantne przed wektory kowariantne. Mnożenie tensorowe jest więc na ogół operacją nieprzemienną (por. uwagi na końcu punktu 13, §18).

W modelu form wieloliniowych operacja mnożenia tensorowego zadana była mnożeniem funkcji. W przypadku dowolnych tensorów $t \in \mathcal{T}^p_q(V)$ i $s \in \mathcal{T}^{p'}_{q'}(V)$

dostajemy więc tensor $t \otimes s \in \mathcal{T}^{p+p'}_{q+q'}(V)$ zadany przepisem

$$(t \otimes s)(\psi^1, \dots, \psi^{p+p'}; z_1, \dots, z_{q+q'}) = \\ = t(\psi^1, \dots, \psi^p; z_1, \dots, z_q)\, s(\psi^{p+1}, \dots, \psi^{p+p'}; z_{q+1}, \dots, z_{q+q'})\,.$$

W szczególności, wstawiając za argumenty wektory dowolnej bazy i dualnej do niej dostajemy sformułowanie tej operacji we współrzędnych tensorów

$$(t \otimes s)^{i_1 \dots i_p j_1 \dots j_{p'}}{}_{k_1 \dots k_q l_1 \dots l_{q'}} = t^{i_1 \dots i_p}{}_{k_1 \dots k_q} s^{j_1 \dots j_{p'}}{}_{l_1 \dots l_{q'}}\,.$$

Nieprzemienność iloczynu tensorowego jest widoczna także w tych dwóch ostatnich sformułowaniach. Z drugiej strony widocznym jest również, że

$$\text{jeśli} \qquad t \in \mathcal{T}^p_0(V)\,,\ s \in \mathcal{T}^0_q(V)\,, \qquad \text{to} \qquad t \otimes s = s \otimes t\,.$$

6 Kontrakcja tensorów

Kontrakcja tensorów przebiega według ogólnego schematu określonego twierdzeniem 13, §18. Jak wspomnieliśmy we wstępie do tego twierdzenia, kontrakcję można wykonać w parze przestrzeni V i V^* stojących na dowolnych miejscach w ciągu mnożonych przestrzeni. Kontrakcja tensorów z przestrzeni $\mathcal{T}^p_q(V)$ w parze: r-ta przestrzeń V i s-ta przestrzeń V^* jest więc odwzorowaniem

$$C^r_s : \mathcal{T}^p_q(V) \mapsto \mathcal{T}^{p-1}_{q-1}(V)$$

zadanym przepisem

$$C^r_s(x_1 \otimes \dots \otimes x_p \otimes \varphi^1 \otimes \dots \otimes \varphi^q) = \\ = \langle \varphi^s, x_r \rangle\, x_1 \otimes \dots \otimes x_{r-1} \otimes x_{r+1} \otimes \dots \otimes x_p \otimes \varphi^1 \otimes \dots \otimes \varphi^{s-1} \otimes \varphi^{s+1} \otimes \dots \otimes \varphi^q\,.$$

Dla wektorów bazowych dostajemy, w szczególności,

$$C^r_s(e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}) = \\ = \delta^{j_s}_{i_r}\, e_{i_1} \otimes \dots \otimes e_{i_{r-1}} \otimes e_{i_{r+1}} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_{s-1}} \otimes e^{j_{s+1}} \otimes \dots \otimes e^{j_q}\,.$$

Wynika stąd, że operacja kontrakcji dowolnego tensora $t \in \mathcal{T}^p_q(V)$ zapisana we współrzędnych ma postać

$$(C^r_s t)^{i_1 \dots i_{p-1}}{}_{j_1 \dots j_{q-1}} = t^{i_1 \dots i_{r-1} k\, i_r \dots i_{p-1}}{}_{j_1 \dots j_{s-1} k\, j_s \dots j_{q-1}}\,,$$

gdzie po prawej stronie, zgodnie z konwencją Einsteina, sumujemy po wskaźniku k. Mówimy też, że tensor ten powstaje przez ***zwężenie*** r***-tego górnego***

wskaźnika z s-tym wskaźnikiem dolnym. Sposób zaznaczenia numerów r i s w symbolu operatora kontrakcji C^r_s dostosowaliśmy do notacji wskaźnikowej.

W modelu odwzorowań wieloliniowych kontrakcja przybiera formę:

$$\begin{aligned}(C^r_s t)(\psi^1, \dots, \psi^{p-1}; y_1, \dots, y_{q-1}) &= \\ = t(\psi^1, \dots, \psi^{r-1}, e^k, \psi^r, \dots, \psi^{p-1}; y_1, \dots, y_{s-1}, e_k, y_s, \dots, y_{q-1})\,,\end{aligned}$$

gdzie (e_i) i (e^j) jest dowolną parą wzajemnie dualnych baz.

Złożenie operacji iloczynu tensorowego z kontrakcją daje dalsze możliwości tworzenia nowych tensorów. Notacja wskaźnikowa jest najwygodniejsza do notowania takich bardziej złożonych operacji. Niech, na przykład, $t \in \mathcal{T}^2_2(V)$ i $s \in \mathcal{T}^3_3(V)$. Tworząc ich iloczyn tensorowy $t \otimes s$, a następnie zwężając pierwszy górny wskaźnik z czwartym dolnym, drugi z trzecim i trzeci z drugim, dostajemy tensor $u \in \mathcal{T}^2_2(V)$ o współrzędnych

$$u^{i_1 i_2}{}_{j_1 j_2} = t^{kl}{}_{j_1 m} s^{m i_1 i_2}{}_{l k j_2}\,.$$

7 Permutacja wskaźników tensora

Tensory zbudowane są z wielokrotnych kopii jednej przestrzeni i jej dualnej. Pozwala to na wprowadzenie dodatkowej operacji.

Twierdzenie 2. *Dla każdej permutacji $\pi \in S_p$ ($\sigma \in S_q$) istnieje dokładnie jeden operator liniowy $P^\pi : \mathcal{T}^p_q(V) \mapsto \mathcal{T}^p_q(V)$ (odpowiednio: $P_\sigma : \mathcal{T}^p_q(V) \mapsto \mathcal{T}^p_q(V)$), który na tensory proste działa według przepisu*

$$P^\pi(x_1 \otimes \ldots \otimes x_p \otimes \varphi^1 \otimes \ldots \otimes \varphi^q) = x_{\pi^{-1}(1)} \otimes \ldots \otimes x_{\pi^{-1}(p)} \otimes \varphi^1 \otimes \ldots \otimes \varphi^q$$

lub odpowiednio:

$$P_\sigma(x_1 \otimes \ldots \otimes x_p \otimes \varphi^1 \otimes \ldots \otimes \varphi^q) = x_1 \otimes \ldots \otimes x_p \otimes \varphi^{\sigma^{-1}(1)} \otimes \ldots \otimes \varphi^{\sigma^{-1}(q)}\,.$$

Dowód. Dowód przebiega według wielokrotnie już używanego schematu. Odwzorowanie o przepisie

$$\tilde{P}^\pi(x_1, \dots, x_p; \varphi^1, \dots, \varphi^q) = x_{\pi^{-1}(1)} \otimes \ldots \otimes x_{\pi^{-1}(p)} \otimes \varphi^1 \otimes \ldots \otimes \varphi^q$$

jest wieloliniowe, więc jednoznacznie zadaje operator P^π spełniający warunek twierdzenia. Podobnie dla operatora P_σ. □

Zwróćmy uwagę, że tutaj operatory P^π i P_σ działają wewnątrz *jednej* przestrzeni, a więc dokonujemy permutacji mnożonych wektorów bez zmiany modelu iloczynu tensorowego (por. uwagi na końcu punktu 13, §18).

Operatory P^π i P_σ nazywamy ***operatorami permutacji wskaźników tensora*** – nazwy te mają uzasadnienie w formułach działania tych operatorów w notacji wskaźnikowej. Aby je otrzymać, wystarczy znaleźć przepis działania tych operatorów na tensory z bazy iloczynowej. Podstawmy $x_k = e_{i_k}$ oraz $\varphi^l = e^{j_l}$ w podanej w treści twierdzenia 2 regule działania operatorów permutacji na tensory proste. Zastępując w tych podstawieniach k przez $\pi^{-1}(k)$ oraz l przez $\sigma^{-1}(l)$ dostajemy $x_{\pi^{-1}(k)} = e_{i_{\pi^{-1}(k)}}$ oraz $\varphi^{\sigma^{-1}(l)} = e^{j_{\sigma^{-1}(l)}}$, więc

$$P^\pi(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}) = e_{i_{\pi^{-1}(1)}} \otimes \ldots \otimes e_{i_{\pi^{-1}(p)}} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}\,,$$
$$P_\sigma(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}) = e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_{\sigma^{-1}(1)}} \otimes \ldots \otimes e^{j_{\sigma^{-1}(q)}}\,.$$

Niech teraz t będzie dowolnym tensorem o walencji $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$. Rozkładamy go w bazie iloczynowej i działamy operatorem P^π. Biorąc pod uwagę liniowość operatora dostajemy pierwszą z następujących równości

$$\begin{aligned} P^\pi t = t^{i_1\ldots i_p}{}_{j_1\ldots j_q} e_{i_{\pi^{-1}(1)}} \otimes \ldots \otimes e_{i_{\pi^{-1}(p)}} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q} = \\ = t^{i_{\pi(1)}\ldots i_{\pi(p)}}{}_{j_1\ldots j_q} e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \ldots \otimes e^{j_q}\,. \end{aligned}$$

Druga z powyższych równości powstaje przez zmianę oznaczeń $i_k \to i_{\pi(k)}$ (oznaczenie wskaźników sumacyjnych można zmieniać), przy którym $i_{\pi^{-1}(k)} \to i_k$. Podobne tożsamości uzyskuje się dla operatora P_σ. Z tożsamości tych odczytujemy reguły transformacyjne:

$$\begin{aligned} (P^\pi t)^{i_1\ldots i_p}{}_{j_1\ldots j_q} &= t^{i_{\pi(1)}\ldots i_{\pi(p)}}{}_{j_1\ldots j_q}\,, \\ (P_\sigma t)^{i_1\ldots i_p}{}_{j_1\ldots j_q} &= t^{i_1\ldots i_p}{}_{j_{\sigma(1)}\ldots j_{\sigma(q)}}\,. \end{aligned}$$

Podobnie jak dla operatorów kontrakcji, sposób oznaczenia operatorów P^π i P_σ przystosowaliśmy do notacji wskaźnikowej.

Rozważmy działanie operatorów P^π i P_σ w modelu form wieloliniowych. Jest widoczne, że część kowariantna tensora nie bierze udziału w transformacji zadanej operatorem P^π, więc w następującym przekształceniu pomijamy ją. Z definicji P^π oraz definicji modelu mamy

$$(P^\pi[x_1 \otimes \ldots \otimes x_p])(\psi^1, \ldots, \psi^p) = \langle \psi^1, x_{\pi^{-1}(1)} \rangle \ldots \langle \psi^p, x_{\pi^{-1}(p)} \rangle\,.$$

Jeśli $\pi^{-1}(i) = k$, to $i = \pi(k)$, a każda liczba ze zbioru $\{1, \ldots, p\}$ występuje w ciągu $\pi^{-1}(1), \ldots, \pi^{-1}(p)$ dokładnie jeden raz. Zmieniając porządek czynników liczbowych możemy więc prawą stronę zapisać jako

$$\langle \psi^{\pi(1)}, x_1 \rangle \ldots \langle \psi^{\pi(p)}, x_p \rangle = (x_1 \otimes \ldots \otimes x_p)(\psi^{\pi(1)}, \ldots, \psi^{\pi(p)})\,.$$

Przez liniowość dostajemy działanie operatora P^π i – w analogiczny sposób – operatora P_σ, na dowolny tensor:

$$(P^\pi t)(\psi^1, \dots, \psi^p; y_1, \dots, y_q) = t(\psi^{\pi(1)}, \dots, \psi^{\pi(p)}; y_1, \dots, y_q)\,,$$
$$(P_\sigma t)(\psi^1, \dots, \psi^p; y_1, \dots, y_q) = t(\psi^1, \dots, \psi^p; y_{\sigma(1)}, \dots, y_{\sigma(q)})\,.$$

Podstawiając tutaj $\psi^k = e^{i_k}$ oraz $y_l = e_{j_l}$ moglibyśmy znów otrzymać uzyskane wcześniej przepisy działania operatorów permutacji wyrażone we współrzędnych tensora.

8 Metryka kontrawariantna

W dwóch kolejnych punktach wykazujemy, że obecność w przestrzeni V wyróżnionej metryki prowadzi do dalszych operacji tensorowych.

Przypomnijmy najpierw, że w twierdzeniu 14, §18, otrzymaliśmy naturalne izomorficzne utożsamienia iloczynów tensorowych z przestrzeniami odwzorowań liniowych. Połóżmy w tym twierdzeniu $W = V^*$. Dostajemy wtedy dwa izomorfizmy pomiędzy $V^* \otimes V^* \equiv \mathcal{T}^0_2(V)$ i $\mathcal{L}(V, V^*)$. Podstawiamy w przepisach tych izomorfizmów $y \equiv \psi \in V^*$ i zapisujemy te przepisy w notacji wskaźnikowej:

$$\text{(i)}\quad [(\psi \otimes \varphi)x]_i = \psi_i \varphi_j x^j\,, \qquad \text{(ii)}\quad [(\psi \otimes \varphi)^{\mathrm{T}} x]_i = x^j \psi_j \varphi_i\,.$$

Dla dowolnego tensora $t \in \mathcal{T}^0_2(V)$ formuły te rozszerzają się do odwzorowań liniowych:

$$\text{(i)}\qquad V \ni x \mapsto tx \in V^*\,, \quad (tx)_i = t_{ij} x^j\,,$$
$$\text{(ii)}\qquad V \ni x \mapsto t^{\mathrm{T}} x \in V^*\,, \quad (t^{\mathrm{T}} x)_i = x^j t_{ji}\,.$$

Zastępując w powyższych utożsamieniach $V \to V^*$ dostajemy dwa izomorfizmy przestrzeni tensorowej $\mathcal{T}^2_0(V)$ z przestrzenią odwzorowań $\mathcal{L}(V^*, V)$: dla $s \in \mathcal{T}^2_0(V)$ mamy

$$\text{(i)}\qquad V^* \ni \varphi \mapsto s\varphi \in V\,, \quad (s\varphi)^i = s^{ij} \varphi_j\,,$$
$$\text{(ii)}\qquad V^* \ni \varphi \mapsto s^{\mathrm{T}} \varphi \in V\,, \quad (s^{\mathrm{T}} \varphi)^i = \varphi_j s^{ji}\,.$$

Niech w przestrzeni V zadana będzie niezdegenerowana metryka biliniowa g. Metryka ta, jako tensor o walencji $\left[\begin{smallmatrix}0\\2\end{smallmatrix}\right]$, zadaje odwzorowanie liniowe $V \mapsto V^*$ (wybieramy sposób (ii)):

$$V \ni x \mapsto g^{\mathrm{T}} x \in V^*\,, \quad (g^{\mathrm{T}} x)_i = x^j g_{ji}\,.$$

Odwzorowanie to nazywamy ***opuszczeniem wskaźnika wektora kontrawariantnego***. Ponieważ metryka jest niezdegenerowana, to jądro tego odwzorowania jest trywialne (gdyby dla pewnego $x \neq 0$ było $g^{\mathrm{T}} x = 0$, to mielibyśmy

$g(x, y) = 0$ dla wszystkich $y \in V$). Przestrzenie V i V^* mają równe wymiary, więc odwzorowanie to jest izomorfizmem. Istnieje więc odwzorowanie odwrotne $V^* \mapsto V$, które można utożsamić z pewnym tensorem $\hat{g} \in \mathcal{T}^2_0(V)$, działającym według przepisu (wybieramy sposób (i))

$$V^* \ni \varphi \mapsto \hat{g}\varphi \in V\,, \quad (\hat{g}\varphi)^i = \hat{g}^{ij}\varphi_j\,.$$

To odwzorowanie nazywamy ***podniesieniem wskaźnika wektora kowariantnego***. Odwzorowania g^{T} i $\hat{g}$ są wzajemnie odwrotne, więc w notacji wskaźnikowej mamy

$$\hat{g}^{ik}g_{jk} = \hat{g}^{ki}g_{kj} = \delta^i{}_j\,.$$

Widać stąd, że przy dowolnym wyborze bazy jest

$$\hat{\mathbf{g}} = (\mathbf{g}^{-1})^{\mathrm{T}}\,.$$

Tensor $\hat{g}$ jednoznacznie zadany przez metrykę g nazywamy ***metryką kontrawariantną***. Ze związku ich macierzy w bazie widać natychmiast, że $\hat{g}$ ma taką samą symetrię, jak g. W szczególności, dla metryk symetrycznych odwzorowania typu (i) i (ii) pokrywają się, a dla metryk antysymetrycznych różnią się znakiem.

Izomorfizm $V \ni x \mapsto g^{\mathrm{T}}x \in V^*$ jest tym samym, który – w nieco innym języku – był omówiony w twierdzeniu 3, §16.

9 Podnoszenie i opuszczanie wskaźników tensora

Operacje opuszczenia i podniesienia wskaźnika mają uogólnienia stosujące się do innych tensorów, opisane w następującym twierdzeniu.

Twierdzenie 3. *Niech przestrzeń V będzie wyposażona w niezdegenerowaną metrykę $g \in \mathcal{T}^0_2(V)$ i oznaczmy stowarzyszoną z nią metrykę kontrawariantną przez $\hat{g} \in \mathcal{T}^2_0(V)$. Dla dowolnych liczb $r \in \{1, \ldots, p\}$ i $s \in \{1, \ldots, q+1\}$, oraz $r \in \{1, \ldots, p+1\}$ i $s \in \{1, \ldots, q\}$ istnieją odwzorowania liniowe*

$$\mathcal{T}^p_q(V) \ni t \mapsto g^{\mathrm{T}r}_{\ \ s}t \in \mathcal{T}^{p-1}_{q+1}\ (p \geq 1)\,, \quad \mathcal{T}^p_q(V) \ni t \mapsto \hat{g}^r_s t \in \mathcal{T}^{p+1}_{q-1}\ (q \geq 1)\,,$$

jednoznacznie określone działaniem na tensory proste:

$$\begin{aligned} g^{\mathrm{T}r}_{\ \ s}\big(x_1 \otimes \ldots \otimes x_p \otimes \varphi^1 \otimes \ldots \otimes \varphi^q\big) &= \\ = x_1 \otimes \ldots \otimes x_{r-1} \otimes x_{r+1} \otimes \ldots \otimes x_p \otimes \\ \otimes \varphi^1 \otimes \ldots \otimes \varphi^{s-1} \otimes (g^{\mathrm{T}}x_r) \otimes \varphi^s \otimes \ldots \otimes \varphi^q\,, \\ \hat{g}^r_s\big(x_1 \otimes \ldots \otimes x_p \otimes \varphi^1 \otimes \ldots \otimes \varphi^q\big) &= \\ = x_1 \otimes \ldots \otimes x_{r-1} \otimes (\hat{g}\varphi^s) \otimes x_r \otimes \ldots \otimes x_p \otimes \\ \otimes \varphi^1 \otimes \ldots \otimes \varphi^{s-1} \otimes \varphi^{s+1} \otimes \ldots \otimes \varphi^q\,. \end{aligned}$$

Działanie tych odwzorowań na współrzędne tensorów w dowolnej bazie ma postać:

$$(g^{\mathrm{T}r}_{\ s}t)^{i_1\dots i_{p-1}}{}_{j_1\dots j_{q+1}} = t^{i_1\dots i_{r-1}ki_r\dots i_{p-1}}{}_{j_1\dots j_{s-1}j_{s+1}\dots j_{q+1}}\, g_{kj_s}\,,$$
$$(\hat{g}^r_s t)^{i_1\dots i_{p+1}}{}_{j_1\dots j_{q-1}} = \hat{g}^{i_r k}\, t^{i_1\dots i_{r-1}i_{r+1}\dots i_{p+1}}{}_{j_1\dots j_{s-1}kj_s\dots j_{q-1}}\,.$$

Pierwszą z tych operacji będziemy nazywać ***opuszczeniem*** $\boldsymbol{r}$***-tego wskaźnika na miejsce*** $\boldsymbol{s}$***-te****, a drugą –* ***podniesieniem*** $\boldsymbol{s}$***-tego wskaźnika na miejsce*** $\boldsymbol{r}$***-te****.*

Dowód. Dowód istnienia i jednoznaczności opiera się na znanym nam schemacie: prawe strony pierwszych dwóch równań określają odwzorowania wieloliniowe, które na mocy twierdzenia 7, §18, wyznaczają swoje odpowiedniki liniowe. Dla tensorów prostych postać działania odwzorowań $g^{\mathrm{T}r}_{\ s}$ i $\hat{g}^r_s$ na współrzędne wynika natychmiast z przepisów na opuszczanie i podnoszenie wskaźników dla wektorów otrzymanych w poprzednim punkcie. Na dowolne tensory działanie rozszerza się przez liniowość. □

Odwzorowania $g^{\mathrm{T}r}_{\ s}$ i $\hat{g}^r_s$ można zinterpretować jako złożenia prostszych operacji tensorowych: mnożenia tensorowego przez g lub $\hat{g}$, zwężenia i permutacji. Dokładniej, pomnożenie tensora $t \in \mathcal{T}^p_q(V)$ tensorowo przez g z prawej strony, a następnie zwężenie r-tego wskaźnika górnego z $(q+1)$-szym dolnym daje tensor o współrzędnych $t^{i_1\dots i_{r-1}ki_r\dots i_{p-1}}{}_{j_1\dots j_q} g_{kj_{q+1}}$. Teraz permutacja polegająca na cyklicznym przestawieniu wskaźników $(j_s, \dots, j_q, j_{q+1})$ prowadzi do tensora $g^{\mathrm{T}r}_{\ s}t$. Podobnie dla operacji $\hat{g}^r_s$ (mnożymy przez $\hat{g}$ z lewej strony).

Często wygodne jest przyjęcie zmodyfikowanej konwencji notowania współrzędnych tensora, która istotnie upraszcza notację podnoszenia i opuszczania wskaźników. Dla każdego tensora zadajemy jedną wspólną kolejność notowania wszystkich wskaźników, zarówno dolnych, jak i górnych, np.

$$t^{i_1}{}_{i_2}{}^{i_3}{}_{i_4 i_5}\, e_{i_1} \otimes e_{i_3} \otimes e^{i_2} \otimes e^{i_4} \otimes e^{i_5}\,.$$

Zauważmy, że przyjęcie takiej notacji nie zmienia tensora – kolejność w ramach wskaźników kontrawariantnych i kowariantnych z osobna jest zachowana. Przydatność takiej konwencji ujawnia się dopiero przy podnoszeniu i opuszczaniu wskaźników: przyjmujemy, że ten zadany porządek wskaźników przy tych operacjach nie ulega zmianie, a tensor w którym podniesiono lub opuszczono wskaźnik(i) oznaczamy tym samym symbolem rdzeniowym. Na przykład podniesienie drugiego i piątego, a opuszczenie trzeciego wskaźnika w powyższym przykładzie daje tensor

$$t^{i_1 i_2}{}_{i_3 i_4}{}^{i_5}\, e_{i_1} \otimes e_{i_2} \otimes e_{i_5} \otimes e^{i_3} \otimes e^{i_4}\,,$$

o współrzędnych określonych przepisem

$$t^{i_1 i_2}{}_{i_3 i_4}{}^{i_5} = \hat{g}^{i_2 k}\,\hat{g}^{i_5 l}\,t^{i_1}{}_k{}^m{}_{i_4 l}\,g_{m i_3}\,.$$

Opuszczenie w tym tensorze drugiego i piątego wskaźnika, a podniesienie trzeciego, przywraca tensor wyjściowy.

Ze związku tensorów g i $\hat{g}$ wynika, że

$$\hat{g}^{ij} = \hat{g}^{ik}\hat{g}^{jl}g_{kl}\,,$$

co można zinterpretować tak: tensor $\hat{g}$ powstaje z tensora g przez podniesienie obu wskaźników. Przy powyższej konwencji zachowania symbolu rdzeniowego przy podnoszeniu i opuszczaniu wskaźników można więc pisać $\hat{g}^{ij} \equiv g^{ij}$.

10 Podnoszenie i opuszczanie wskaźników w bazie ortonormalnej

Jeśli g jest niezdegenerowaną metryką symetryczną o sygnaturze $(n-r, r)$ (w przestrzeni rzeczywistej), to w bazie kanonicznej zarówno (g_{ij}), jak i (g^{ij}) są macierzami diagonalnymi, o wyrazach na diagonali równych

$$g_{ii} = g^{ii} = \begin{cases} +1 & \text{dla } i = 1, \ldots, n-r\,, \\ -1 & \text{dla } i = n-r+1, \ldots, n\,. \end{cases}$$

Stąd dla dowolnego wektora:

$$\text{w bazie kanonicznej} \quad x^i = \begin{cases} x_i & \text{dla } i = 1, \ldots, n-r\,, \\ -x_i & \text{dla } i = n-r+1, \ldots, n\,, \end{cases}$$

i podobnie przy podnoszeniu i opuszczaniu każdego ze wskaźników dowolnego tensora. W szczególności: w bazie ortonormalnej w przestrzeni euklidesowej podnoszenie i opuszczanie wskaźników nie zmienia wartości numerycznej współrzędnych.

11 Przykłady

(i) Iloczyn tensorowy
Niech $t = e_1 \otimes e^1 + 2e_2 \otimes e^3 \in \mathcal{T}^1_1(V)$, $s = -e_2 \otimes e_2 + 3e_3 \otimes e_1 \in \mathcal{T}^2_0(V)$. Wtedy

$$t \otimes s = (-e_1 \otimes e_2 \otimes e_2 + 3e_1 \otimes e_3 \otimes e_1) \otimes e^1 + \\ + (-2e_2 \otimes e_2 \otimes e_2 + 6e_2 \otimes e_3 \otimes e_1) \otimes e^3 \in \mathcal{T}^3_1(V)\,.$$

(ii) Kontrakcja tensora
Dany jest tensor

$$t = (3e_2 + 2e_3) \otimes (5e_1 + e_2) \otimes e^1 \otimes e^3 + \\ + (2e_1 - e_2) \otimes (3e_2 - e_3) \otimes e^2 \otimes e^3 \in \mathcal{T}^2_2(V)\,.$$

W wyniku zastosowania do niego kontrakcji C^2_1 dostajemy

$$C^2_1 t = \langle e^1, 5e_1 + e_2\rangle\,(3e_2 + 2e_3) \otimes e^3 + \langle e^2, 3e_2 - e_3\rangle\,(2e_1 - e_2) \otimes e^3 = \\ = (6e_1 + 12e_2 + 10e_3) \otimes e^3 \in \mathcal{T}^1_1(V)\,.$$

Rozważmy tę samą operację w notacji wskaźnikowej, w której operacja C^2_1 jest zwężeniem drugiego górnego wskaźnika z pierwszym dolnym. Jedyne różne od zera współrzędne tensora t dające przyczynek do $C^2_1 t$ to

$$t^{21}{}_{13} = 15\,, \quad t^{31}{}_{13} = 10\,, \quad t^{12}{}_{23} = 6\,, \quad t^{22}{}_{23} = -3\,,$$

a stąd jedyne różne od zera współrzędne tensora $C^2_1 t$ to

$$t^{1i}{}_{i3} = 6\,, \quad t^{2i}{}_{i3} = 12\,, \quad t^{3i}{}_{i3} = 10\,.$$

(iii) Permutacja wskaźników tensora
Niech $\pi = \left(\begin{smallmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{smallmatrix}\right)$ oraz

$$t = (2e_1 - 3e_2) \otimes e_3 \otimes (e_2 + 4e_3) + e_3 \otimes (e_1 - e_2) \otimes e_1 \in \mathcal{T}^3_0(V)\,.$$

Szukamy tensora $P^\pi t$.

W zastosowaniu do tensorów prostych mamy

$$P^\pi(x_1 \otimes x_2 \otimes x_3) = x_2 \otimes x_3 \otimes x_1\,,$$

skąd przez liniowość

$$P^\pi t = e_3 \otimes (e_2 + 4e_3) \otimes (2e_1 - 3e_2) + (e_1 - e_2) \otimes e_1 \otimes e_3\,.$$

Otrzymanie tego samego wyniku przez zastosowanie prawa transformacji współrzędnych tensora pod działaniem operatora permutacji jest zadaniem prostym, lecz rachunkowo bardziej okrężnym. Polecamy je jako ćwiczenie.

(iv) Metryka kontrawariantna, podnoszenie i opuszczanie wskaźników
Trójwymiarowa przestrzeń V wyposażona jest w niezdegenerowaną metrykę

$$g = e^1 \otimes e^1 + e^2 \otimes e^2 + 3e^3 \otimes e^3 + 2e^1 \otimes e^3 + 2e^3 \otimes e^1 .$$

Szukamy tensora $\hat{g}^1_1 t$, dla danego

$$t = e_1 \otimes (e^2 - 2e^3) \otimes (e^1 + e^2) + e_2 \otimes (e^1 + e^2) \otimes (e^2 - e^3) \in \mathcal{T}^1_2(V) .$$

Odwracając macierz metryki g w zadanej bazie dostajemy

$$\hat{g} = -3e_1 \otimes e_1 + e_2 \otimes e_2 - e_3 \otimes e_3 + 2e_1 \otimes e_3 + 2e_3 \otimes e_1$$

i wyliczamy

$$\hat{g}(e^2 - 2e^3) = -4e_1 + e_2 + 2e_3 , \quad \hat{g}(e^1 + e^2) = -3e_1 + e_2 + 2e_3 .$$

Stąd

$$\hat{g}^1_1 t = (-4e_1 + e_2 + 2e_3) \otimes e_1 \otimes (e^1 + e^2) + (-3e_1 + e_2 + 2e_3) \otimes e_2 \otimes (e^2 - e^3) .$$

(v) Inwolutywność ortogonalnego dopełnienia podprzestrzeni
Jeśli g jest metryką w przestrzeni V o skończonym wymiarze, to

$$g(y, x) = \langle g^{\mathrm{T}} y, x\rangle \quad \text{oraz} \quad g(x, y) = \langle g^{\mathrm{T}} x, y\rangle .$$

Jeśli więc dla danego x zażądamy, aby te wielkości znikały dla każdego $y \in W$, gdzie $W \subseteq V$ jest podprzestrzenią, to widzimy, że

$$W^\perp = (g^{\mathrm{T}} W)^{\perp^*} \quad \text{oraz} \quad x \in W^\perp \iff g^{\mathrm{T}} x \in W^{\perp^*} .$$

Jeśli metryka g jest niezdegenerowana, więc istnieje stowarzyszona z nią metryka kontrawariantna, to przez zadziałanie izomorfizmem $\hat{g}$ dostajemy z drugiej z tych formuł

$$W^\perp = \hat{g}(W^{\perp^*}) .$$

Zastępując w ostatniej tożsamości W przez $W^\perp$ i podstawiając po prawej stronie pierwszą reprezentację $W^\perp$ dostajemy przy wykorzystaniu własności operacji $\perp^*$ (p. 5, §18):

$$(W^\perp)^\perp = W .$$

Stąd mamy dalej

$$\operatorname{Ker} g_W = W \cap W^\perp = W^\perp \cap (W^\perp)^\perp = \operatorname{Ker} g_{W^\perp} .$$

§20 Tensory symetryczne i antysymetryczne

1 Własność grupowa operatorów permutacji

Określenia operatorów permutacji wskaźników P^π i P_σ dane twierdzeniem 2, §19, używają permutacji *odwrotnych* do π i σ odpowiednio. Nie jest to przypadkowe, definicja ta pociąga następującą istotną ich własność. Przed jej sformułowaniem przypomnijmy, że przez $GL(W)$ oznaczyliśmy grupę wszystkich nieosobliwych (tj. bijektywnych) operatorów na przestrzeni W.

Twierdzenie 1. *Operatory P^π i P_σ są nieosobliwe, a przyporządkowania*

$$S_p \ni \pi \mapsto P^\pi \in GL\big(\mathcal{T}^p_q(V)\big)\,, \quad S_q \ni \sigma \mapsto P_\sigma \in GL\big(\mathcal{T}^p_q(V)\big)$$

są homomorfizmami grupowymi, tj. zachodzą relacje

$$P^\rho P^\pi = P^{\rho\circ\pi}\,, \quad P_\tau P_\sigma = P_{\tau\circ\sigma}\,.$$

Dowód. Dowód przeprowadzimy dla operatorów P^π na przestrzeni $\mathcal{T}^p_0(V)$, uogólnienia są oczywiste. Wykażemy najpierw tożsamość operatorową z ostatniej linii twierdzenia. Wprowadzamy pomocniczo oznaczenie $y_k \equiv x_{\pi^{-1}(k)}$, które pomoże wykonać pośrednie kroki. Tożsamość wynika teraz z następującego ciągu równości:

$$\begin{aligned}P^\rho P^\pi(x_1\otimes\ldots\otimes x_p) &= P^\rho\big(x_{\pi^{-1}(1)}\otimes\ldots\otimes x_{\pi^{-1}(p)}\big) = P^\rho(y_1\otimes\ldots\otimes y_p) = \\ &= y_{\rho^{-1}(1)}\otimes\ldots\otimes y_{\rho^{-1}(p)} = x_{\pi^{-1}(\rho^{-1}(1))}\otimes\ldots\otimes x_{\pi^{-1}(\rho^{-1}(p))} = \\ &= x_{(\rho\circ\pi)^{-1}(1)}\otimes\ldots\otimes x_{(\rho\circ\pi)^{-1}(p)} = P^{\rho\circ\pi}(x_1\otimes\ldots\otimes x_p)\,.\end{aligned}$$

Ponieważ tensory proste generują całą przestrzeń, to tożsamość operatorowa jest spełniona. Jest oczywiste na mocy definicji, że

$$P^{\mathrm{id}} = \mathrm{id}\,, \quad \text{więc} \quad P^\pi P^{\pi^{-1}} = P^{\pi^{-1}} P^\pi = \mathrm{id}\,.$$

Stąd operatory P^π są nieosobliwe, co kończy dowód. $\square$

2 Symetria i antysymetria tensorów

W całym bieżącym paragrafie zakładamy, że ciało $\mathbb{K}$, nad którym zbudowana jest przestrzeń V, ma charakterystykę nieskończoną (więc zawiera liczby całkowite, p. 10, §3).

Prowadzimy w tym punkcie dyskusję dla permutacji wskaźników kontrawariantnych, ale całkowicie analogiczne rezultaty obowiązują dla wskaźników kowariantnych.

Niech $\tau \in S_p$ będzie dowolną transpozycją. Wtedy z własności grupowej operatorów P^τ mamy

$$P^\tau P^\tau = \mathrm{id}\,,$$

a więc operator ten jest inwolutywny. Przykład (i) z punktu 5, §13, pokazuje, że każdy taki operator jest diagonalizowalny, o wartościach własnych $+1$ i -1. Każdy tensor $t \in \mathcal{T}^p_q(V)$ taki, że $P^\tau t = t$ ($P^\tau t = -t$), nazywamy ***tensorem symetrycznym*** (odpowiednio: ***antysymetrycznym***) ***w parze wskaźników*** zamienianych transpozycją τ.

Jeśli tensor t jest symetryczny (lub antysymetryczny) w każdej parze wskaźników kontrawariantnych, to mówimy, że jest on ***(całkowicie) symetryczny*** (odpowiednio: ***antysymetryczny***) ***we wskaźnikach kontrawariantnych***.

Twierdzenie 2. *Tensor t jest symetryczny (lub: antysymetryczny) we wskaźnikach kontrawariantnych wtedy, i tylko wtedy, gdy dla każdej permutacji $\pi \in S_p$ zachodzi:*

$$\begin{aligned} P^\pi t &= t\,, && \textit{w przypadku symetrycznym}\,,\\ P^\pi t &= \operatorname{sgn}\pi\, t\,, && \textit{w przypadku antysymetrycznym}\,. \end{aligned}$$

Dowód. Niech $\pi = \tau_s \circ \ldots \circ \tau_1$ będzie rozkładem permutacji na transpozycje. Korzystając z homomorfizmu grupowego opisanego twierdzeniem 1 mamy $P^\pi = P^{\tau_s} \ldots P^{\tau_1}$. Jeśli więc t jest symetryczny (antysymetryczny), to $P^\pi t = t$ (odpowiednio: $P^\pi t = (-1)^s t = \operatorname{sgn}\pi\, t$). Wynikanie odwrotne jest oczywiste. $\square$

3 Operatory symetryzacji i antysymetryzacji

Wprowadzamy dwa operatory działające w przestrzeniach $\mathcal{T}^p_q(V)$:

$$\mathcal{S} = \frac{1}{p!}\sum_{\pi\in S_p} P^\pi \qquad \text{– } \textbf{\textit{operator symetryzacji,}}$$

$$\mathcal{A} = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, P^\pi \qquad \text{– } \textbf{\textit{operator antysymetryzacji.}}$$

Ich użyteczność oraz uzasadnienie ich nazw wynikają z następnego twierdzenia, które poprzedzimy prostym lematem.

Lemat 3. *Dla każdej permutacji $\sigma \in S_p$ jest*

$$P^\sigma \mathcal{S} = \mathcal{S}\, P^\sigma = \mathcal{S}\,, \quad P^\sigma \mathcal{A} = \mathcal{A}\, P^\sigma = \operatorname{sgn}\sigma\, \mathcal{A}\,.$$

Dowód. Ostatnia tożsamość wynika z ciągu równości

$$\mathcal{A}\, P^\sigma = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, P^{\pi\circ\sigma} = \operatorname{sgn}\sigma \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}(\pi\circ\sigma)\, P^{\pi\circ\sigma} = \operatorname{sgn}\sigma \mathcal{A}\,.$$

Pierwsza równość w tym ciągu wynika z własności grupowej operatorów permutacji, druga z tożsamości $\operatorname{sgn}\sigma\ \operatorname{sgn}(\pi\circ\sigma) = \operatorname{sgn}\pi$, a ostatnia z następującego faktu: jeśli π przebiega całą grupę S_p, to dla ustalonej permutacji σ również $\pi\circ\sigma$ przebiega całą grupę. W podobny sposób upraszcza się iloczyn $P^\sigma\mathcal{A}$, a dla operatora $\mathcal{S}$ wystarczy pominąć znaki permutacji. □

Twierdzenie 4. *Operatory $\mathcal{S}$ i $\mathcal{A}$ są operatorami rzutowymi oraz*

$$\mathcal{S}\mathcal{A} = \mathcal{A}\mathcal{S} = 0\,.$$

Ponadto,

$$\begin{aligned} t \text{ jest symetryczny} \quad &\Longleftrightarrow \quad \mathcal{S}\,t = t\,,\\ t \text{ jest antysymetryczny} \quad &\Longleftrightarrow \quad \mathcal{A}\,t = t\,, \end{aligned}$$

tj. tensory symetryczne (antysymetryczne) tworzą podprzestrzeń, na którą rzutuje operator $\mathcal{S}$ (odpowiednio: $\mathcal{A}$).

Dowód. Korzystając z lematu mamy

$$\mathcal{A}\mathcal{A} = \frac{1}{p!}\sum_\pi \operatorname{sgn}\pi\, P^\pi\mathcal{A} = \frac{1}{p!}\sum_\pi \mathcal{A} = \mathcal{A}\,,$$

bo w grupie S_p jest $p!$ permutacji. Dla operatora $\mathcal{S}$ wystarczy pominąć znaki permutacji. Dalej,

$$\mathcal{A}\,\mathcal{S} = \frac{1}{p!}\sum_\pi \operatorname{sgn}\pi\, P^\pi\mathcal{S} = \frac{1}{p!}\sum_\pi \operatorname{sgn}\pi\,\mathcal{S} = 0\,,$$

bo w grupie S_p jest tyle samo permutacji parzystych, co nieparzystych. Jeśli t jest tensorem antysymetrycznym, to $\mathcal{A}\,t = \dfrac{1}{p!}\displaystyle\sum_\pi t = t$, i analogicznie dla symetrycznego. Odwrotnie,

$$\text{jeśli}\quad \mathcal{A}\,t = t\,,\qquad \text{to}\quad P^\sigma t = P^\sigma\mathcal{A}\,t = \operatorname{sgn}\sigma\,\mathcal{A}\,t = \operatorname{sgn}\sigma\,t\,,$$

i podobnie dla przypadku symetrycznego. □

4 Dalsze własności symetryzacji i antysymetryzacji

Odwołując się do działania operatorów permutacji na współrzędne tensora (punkt 7, §19), dostajemy działanie operatorów symetryzacji i antysymetryzacji w notacji wskaźnikowej:

$$(\mathcal{S}t)^{i_1\dots i_p}{}_{j_1\dots j_q} = \frac{1}{p!}\sum_{\pi\in S_p} t^{i_{\pi(1)}\dots i_{\pi(p)}}{}_{j_1\dots j_q} \equiv t^{(i_1\dots i_p)}{}_{j_1\dots j_q}\,,$$

$$(\mathcal{A}t)^{i_1\dots i_p}{}_{j_1\dots j_q} = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, t^{i_{\pi(1)}\dots i_{\pi(p)}}{}_{j_1\dots j_q} \equiv t^{[i_1\dots i_p]}{}_{j_1\dots j_q}\,,$$

gdzie skrajne prawe strony wprowadzają wygodną notację.

Rozważania ostatnich dwóch punktów można uogólnić przez zastosowanie ich do pewnej podrodziny wszystkich wskaźników kontrawariantnych (lub kowariantnych). Mówimy wtedy o tensorach ***symetrycznych lub antysymetrycznych w rodzinie wskaźników*** oraz o ***symetryzacji lub antysymetryzacji względem tej rodziny***. Symetryzację lub antysymetryzację najwygodniej zaznaczać wtedy w notacji wskaźnikowej. Jeśli (anty-)symetryzowane wskaźniki są przeplecione wskaźnikami nie biorącymi udziału w (anty-)symetryzacji, to te ostatnie ujmujemy między pionowe kreski. Na przykład antysymetryzacja tensora o współrzędnych $t^{i_1i_2j_1i_3j_2}$ względem wskaźników i_1, i_2, i_3 ma postać

$$\begin{aligned} t^{[i_1i_2|j_1|i_3]j_2} &= \frac{1}{3!}\sum_{\pi\in S_3}\operatorname{sgn}\pi\, t^{i_{\pi(1)}i_{\pi(2)}j_1i_{\pi(3)}j_2} \\ &= \frac{1}{6}(+t^{i_1i_2j_1i_3j_2} + t^{i_2i_3j_1i_1j_2} + t^{i_3i_1j_1i_2j_2} - t^{i_2i_1j_1i_3j_2} - t^{i_1i_3j_1i_2j_2} - t^{i_3i_2j_1i_1j_2})\,. \end{aligned}$$

Posługując się notacją wskaźnikową zapisujemy dalsze użyteczne własności operacji symetryzacji i antysymetryzacji. W punktach (iii) i (iv) stosujemy, jak zwykle, sumacyjną konwencję Einsteina.

Twierdzenie 5. *Jeśli $p \ge p' \ge 2$, to*

$$\begin{array}{lll} (i) & t^{[[i_1\dots i_{p'}]i_{p'+1}\dots i_p]} = t^{[i_1\dots i_p]}\,, & t^{((i_1\dots i_{p'})i_{p'+1}\dots i_p)} = t^{(i_1\dots i_p)}\,; \\ (ii) & t^{[(i_1\dots i_{p'})i_{p'+1}\dots i_p]} = 0\,, & t^{([i_1\dots i_{p'}]i_{p'+1}\dots i_p)} = 0\,; \\ (iii) & t^{[i_1\dots i_p]}{}_{(i_1\dots i_p)} = 0\,, & t^{(i_1\dots i_p)}{}_{[i_1\dots i_p]} = 0\,; \\ (iv) & t^{[i_1\dots i_p]}{}_{[i_1\dots i_p]} = t^{[i_1\dots i_p]}{}_{i_1\dots i_p}\,, & t^{(i_1\dots i_p)}{}_{(i_1\dots i_p)} = t^{(i_1\dots i_p)}{}_{i_1\dots i_p}\,. \end{array}$$

Dowód. (i) Oznaczmy przez $\mathcal{A}$ i $\mathcal{A}'$ operatory antysymetryzacji względem wszystkich i względem pierwszych p' wskaźników odpowiednio. Każda permutacja σ początkowych p' wskaźników jest szczególną permutacją wszystkich wskaźników, więc

$$\mathcal{A}\mathcal{A}' = \frac{1}{p'!}\sum_{\sigma\in S_{p'}}\operatorname{sgn}\sigma\,\mathcal{A}P^\sigma = \frac{1}{p'!}\sum_{\sigma\in S_{p'}}\mathcal{A} = \mathcal{A}\,,$$

co kończy dowód pierwszej równości. Drugiej równości dowodzi się analogicznie.

(ii) Przy podobnych oznaczeniach

$$\mathcal{A}\mathcal{S}' = \frac{1}{p'!}\sum_{\sigma\in S_{p'}} \mathcal{A}P^\sigma = \frac{1}{p'!}\sum_{\sigma\in S_{p'}} \operatorname{sgn}\sigma\,\mathcal{A} = 0$$

(bo permutacji parzystych jest tyle samo co nieparzystych). Podobnie dowodzi się drugiej równości.

(iii) Pierwsza tożsamość wynika z ciągu równości

$$t^{[i_1 i_2\dots i_p]}{}_{(i_1 i_2\dots i_p)} = -t^{[i_2 i_1\dots i_p]}{}_{(i_1 i_2\dots i_p)} = -t^{[i_2 i_1\dots i_p]}{}_{(i_2 i_1\dots i_p)} = 0\,.$$

Analogicznie dla drugiej tożsamości.

(iv) Pierwsza tożsamość wynika z ciągu równości

$$t^{[i_1\dots i_p]}{}_{[i_1\dots i_p]} = \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi\, t^{[i_1\dots i_p]}{}_{i_{\pi(1)}\dots i_{\pi(p)}} =$$

$$= \frac{1}{p!}\sum_{\pi\in S_p} \operatorname{sgn}\pi^{-1}\, t^{[j_{\pi^{-1}(1)}\dots j_{\pi^{-1}(p)}]}{}_{j_1\dots j_p} = t^{[j_1\dots j_p]}{}_{j_1\dots j_p}\,.$$

Druga równość w tym ciągu wynika z podstawienia $j_k = i_{\pi(k)}$ (więc $i_l = j_{\pi^{-1}(l)}$), a trzecia z faktu, że antysymetryzacja tensora antysymetrycznego (tutaj w górnych wskaźnikach) nie zmienia go. Analogicznie dla drugiej tożsamości. $\square$

Twierdzenie w oczywisty sposób uogólnia się na przypadki, gdy tensory mają również wskaźniki nie biorące udziału w operacjach (anty-) symetryzacji.

5 Przykłady

(i) Symetryzacja i antysymetryzacja w przestrzeniach $\mathcal{T}^0_2(V)$ i $\mathcal{T}^0_3(V)$ Zastosowanie definicji operatora antysymetryzacji daje w tych dwóch najprostszych przypadkach, w notacji wskaźnikowej,

$$t_{[ij]} = \frac{1}{2}(t_{ij} - t_{ji})\,,$$

$$s_{[ijk]} = \frac{1}{6}(s_{ijk} + s_{jki} + s_{kij} - s_{ikj} - s_{jik} - s_{kji})\,.$$

Dla operatora symetryzacji należy jedynie zamienić znaki ujemne na dodatnie.

Z drugiego z powyższych wzorów oraz z analogicznego wzoru dla operacji symetryzacji, dostajemy w szczególności:

$$\text{jeśli}\quad s_{ijk} = s_{i[jk]}\,,\quad \text{to}\quad s_{[ijk]} = \frac{1}{3}(s_{ijk} + s_{jki} + s_{kij})\,;$$

$$\text{jeśli}\quad s_{ijk} = s_{i(jk)}\,,\quad \text{to}\quad s_{(ijk)} = \frac{1}{3}(s_{ijk} + s_{jki} + s_{kij})\,.$$

(ii) Część symetryczna i część antysymetryczna tensora
Niech $t \in \mathcal{T}_p^0(V)$. Tensory $\mathcal{S}t$ i $\mathcal{A}t$ nazywa się niekiedy ***częścią całkowicie symetryczną*** (odpowiednio: ***antysymetryczną***) tensora t. Dla $p = 2$ części te wyczerpują cały tensor, gdyż

$$t_{ij} = t_{(ij)} + t_{[ij]}\,, \quad \text{czyli na } \mathcal{T}_2^0(V): \quad \mathcal{S} + \mathcal{A} = \mathrm{id}\,.$$

Własności te nie obowiązują dla wyższych walencji tensorowych, np. tensor

$$t = e^1 \otimes e^2 \otimes e^3 + e^2 \otimes e^1 \otimes e^3 - e^1 \otimes e^3 \otimes e^2 - e^2 \otimes e^3 \otimes e^1 \in \mathcal{T}_3^0(V)$$

ma zerowe części symetryczną i antysymetryczną.

(iii) Rekurencyjna antysymetryzacja
Następująca tożsamość daje prosty sposób rozszerzania grupy antysymetryzowanych wskaźników:

$$\begin{aligned} t_{[i_1\ldots\, i_{p+1}]} &= \frac{1}{p+1}\Big(t_{[i_1\ldots\, i_p]i_{p+1}} - \sum_{s=1}^{p} t_{[i_1\ldots\, i_{s-1}i_{p+1}i_{s+1}\ldots\, i_p]i_s}\Big) = \\ &= \frac{1}{p+1}\sum_{s=1}^{p+1}(-1)^{p+1-s} t_{[i_1\ldots\, i_{s-1}i_{s+1}\ldots\, i_{p+1}]i_s}\,. \end{aligned}$$

Dowód tożsamości opiera się na dwóch, wymagających rozwinięcia, obserwacjach: (i) tensory po obu stronach każdej z dwu równości są całkowicie antysymetryczne; (ii) zastosowanie do każdego z tych tensorów operatora antysymetryzacji ze względu na wszystkie wskaźniki daje tożsamości.

6 p-wektory i p-formy

Przestrzenie tensorów całkowicie antysymetrycznych

$$\bigwedge^p(V) := \mathcal{A}\mathcal{T}_0^p(V) \subseteq \mathcal{T}_0^p(V)\,, \quad \bigwedge^p(V^*) := \mathcal{A}\mathcal{T}_p^0(V) \subseteq \mathcal{T}_p^0(V)\,, \quad p \geq 1\,,$$

nazywa się odpowiednio ***przestrzenią*** p***-wektorów*** i ***przestrzenią*** p***-form***. Podobnie jak dla przestrzeni tensorowych przyjmujemy ponadto

$$\bigwedge^0(V) = \bigwedge^0(V^*) = \mathbb{K}\,.$$

Jeśli $t_i \in \bigwedge^{p_i}(V)$ oraz $s^i \in \bigwedge^{p_i}(V^*)$, $i = 1, \ldots, k$, to oznaczamy

$$t_1 \wedge \ldots \wedge t_k = \mathcal{A}\,(t_1 \otimes \ldots \otimes t_k)\,, \quad s^1 \wedge \ldots \wedge s^k = \mathcal{A}\,(s^1 \otimes \ldots \otimes s^k)\,,$$

i nazywamy te tensory ***iloczynem zewnętrznym*** p***-wektorów (***p***-form)***. Jest widoczne, że iloczyn zewnętrzny jest liniowy we wszystkich argumentach.

Ponadto, z punktu (i) twierdzenia 5 wynika, że operator antysymetryzacji stojący wewnątrz zakresu działania innego operatora antysymetryzacji można pominąć, więc

$$(t_1 \wedge \ldots \wedge t_l) \wedge (t_{l+1} \wedge \ldots \wedge t_k) = t_1 \wedge \ldots \wedge t_k\,,$$
$$(s^1 \wedge \ldots \wedge s^l) \wedge (s^{l+1} \wedge \ldots \wedge s^k) = s^1 \wedge \ldots \wedge s^k\,.$$

Mówimy, że element t (lub s) przestrzeni $\bigwedge^p(V)$ (odpowiednio: $\bigwedge^p(V^*)$) jest ***p*-wektorem prostym** (***p*-formą prostą**), gdy jest iloczynem zewnętrznym wektorów z przestrzeni V (odpowiednio: form z przestrzeni V^*):

$$t = x_1 \wedge \ldots \wedge x_p\,, \quad s = \varphi^1 \wedge \ldots \wedge \varphi^p\,, \quad x_i \in V\,,\ \varphi^i \in V^*\,.$$

W notacji wskaźnikowej warunki te przyjmują postać:

$$t^{i_1 \ldots i_p} = {x_1}^{[i_1} \ldots {x_p}^{i_p]}\,, \quad s_{i_1 \ldots i_p} = {\varphi^1}_{[i_1} \ldots {\varphi^p}_{i_p]}\,.$$

Zwróćmy uwagę, że p-wektor prosty (p-forma prosta) jest tensorem, ale nie jest tensorem prostym, więc nazw tych trzeba używać z odpowiednią starannością.

Jeśli w ciągu wektorów $x_1, \ldots, x_p$ pewien wektor powtarza się, np. $x_i = x_j$ dla pewnej pary $i \neq j$, to tensor $x_1 \otimes \ldots \otimes x_p$ jest symetryczny w odpowiednich wskaźnikach (i-tym i j-tym). Jego antysymetryzacja daje tensor zerowy (tw. 5), więc p-wektor utworzony z tych wektorów jest tensorem zerowym. Następujące twierdzenie znacznie wzmacnia ten wynik.

Twierdzenie 6. *Niech $x_i \in V$, $i = 1, \ldots, p$. Wtedy*

$$\begin{aligned} &(i) \quad \textit{wektory}\ \ x_1, \ldots, x_p\ \ \textit{są liniowo niezależne} \\ &\qquad \iff \quad x_1 \wedge \ldots \wedge x_p \neq 0; \\ &(ii) \quad x_{i_1} \wedge \ldots \wedge x_{i_p} = \epsilon_{i_1 \ldots i_p}\, x_1 \wedge \ldots \wedge x_p\,, \end{aligned}$$

gdzie $\epsilon_{i_1 \ldots i_k}$ jest symbolem Levi-Civity. Podobnie dla p-form.

Dowód. (i) Jeśli wektory $x_1, \ldots, x_p$ są liniowo niezależne, to można je uzupełnić do bazy $x_1, \ldots, x_n$. Wtedy wszystkie składowe tensory $x_{\pi^{-1}(1)} \otimes \ldots \otimes x_{\pi^{-1}(p)}$ występujące w sumie określającej p-wektor $x_1 \wedge \ldots \wedge x_p$ są różnymi wektorami bazy iloczynowej w $\mathcal{T}^p_0(V)$, więc ten p-wektor jest różny od zera.

Jeśli wektory są liniowo zależne, to jeden z nich jest kombinacją liniową pozostałych. Podstawiając do p-wektora jego rozkład dostajemy z liniowości kombinację liniową p-wektorów, w każdym z których jeden z czynników występuje dwukrotnie, a więc tensor zerowy.

(ii) Jeśli $i_k = i_l$ dla pewnych k, l, to p-wektor jest zerowy, więc tożsamość jest spełniona. Jeśli liczby i_k są wszystkie różne i jeśli oznaczyć $\pi^{-1} = \left(\begin{smallmatrix} 1 & \dots & p \\ i_1 & \dots & i_p \end{smallmatrix}\right)$, to $x_{i_1} \otimes \ldots \otimes x_{i_p} = P^\pi(x_1 \otimes \ldots \otimes x_p)$. Działając na tę równość operatorem $\mathcal{A}$ i biorąc pod uwagę, że tutaj $\operatorname{sgn}\pi = \epsilon_{i_1 \ldots i_p}$, dostajemy tezę. □

Otrzymane wyniki pozwalają teraz skonstruować bazę przestrzeni tensorów antysymetrycznych.

Twierdzenie 7. *Niech $(e_1, \ldots, e_n)$ będzie dowolną bazą przestrzeni V. Wtedy dla $p \geq 1$ wektory*

$$e_{i_1} \wedge \ldots \wedge e_{i_p}\,, \quad \textit{(odpowiednio: } e^{i_1} \wedge \ldots \wedge e^{i_p}\textit{)}, \quad 1 \leq i_1 < \ldots < i_p \leq n\,,$$

tworzą bazę przestrzeni $\bigwedge^p(V)$ (odpowiednio: $\bigwedge^p(V^*)$), więc

$$\dim \textstyle\bigwedge^p(V) = \dim \bigwedge^p(V^*) = \displaystyle\binom{n}{p}\,.$$

Dowód. Przestrzeń $\bigwedge^p(V)$ jest generowana wektorami

$$\mathcal{A}(e_{i_1} \otimes \ldots \otimes e_{i_p}) = e_{i_1} \wedge \ldots \wedge e_{i_p}\,, \quad i_k \in \{1, \ldots, n\}\,.$$

Ale p-wektory proste różniące się tylko porządkiem czynników są wzajemnie proporcjonalne, więc można przyjąć $i_1 < \ldots < i_p$ (jeśli którykolwiek wektor powtarza się, to p-wektor znika). Załóżmy, że

$$0 = \sum_{j_1 < \ldots < j_p} \alpha^{j_1 \ldots\, j_p} e_{j_1} \wedge \ldots \wedge e_{j_p}\,.$$

Pomnożymy tę tożsamość zewnętrznie z prawej strony przez $e_{k_1} \wedge \ldots \wedge e_{k_{n-p}}$, gdzie $k_1 < \ldots < k_{n-p}$. Wśród ciągów $(j_1, \ldots, j_p)$, gdzie $j_1 < \ldots < j_p$, jest tylko jeden ciąg $(i_1, \ldots, i_p)$ taki, że ciąg $(i_1, \ldots, i_p, k_1, \ldots, k_{n-p})$ nie zawiera powtórzeń. Z twierdzenia 6 dostajemy więc

$$\begin{aligned} &e_{j_1} \wedge \ldots \wedge e_{j_p} \wedge e_{k_1} \wedge \ldots \wedge e_{k_{n-p}} = 0 \quad \text{dla} \quad (j_1, \ldots, j_p) \neq (i_1, \ldots, i_p)\,, \\ &e_{i_1} \wedge \ldots \wedge e_{i_p} \wedge e_{k_1} \wedge \ldots \wedge e_{k_{n-p}} \neq 0\,. \end{aligned}$$

Stąd $\alpha^{i_1 \ldots\, i_p} = 0$. Dobierając odpowiednio ciągi $(k_1, \ldots, k_{n-p})$ dostajemy znikanie wszystkich współczynników, więc p-wektory podane w tezie są liniowo niezależne, a więc stanowią bazę. Liczba tych wektorów jest równa liczbie kombinacji p-elementowych ze zbioru n-elementowego, co kończy dowód. □

Z twierdzenia wynika w szczególności, że nietrywialne przestrzenie $\bigwedge^p(V)$ i $\bigwedge^p(V^*)$ istnieją tylko dla $p = 0, \ldots, n$.

7 p-wektory proste i podprzestrzenie. Związek z orientacją

Zajmiemy się w tym punkcie dalszymi własnościami p-wektorów prostych.

Twierdzenie 8. *Niech wektory $x_1, \ldots, x_p \in V$ będą liniowo niezależne. Wtedy*

$$L(x_1, \ldots, x_p) = \{y \in V \mid x_1 \wedge \ldots \wedge x_p \wedge y = 0\} .$$

Dowód. Jeśli $x_1, \ldots, x_p$ są liniowo niezależne, to y jest ich kombinacją liniową wtedy, i tylko wtedy, gdy wektory $x_1, \ldots, x_p, y$ są liniowo zależne. Ten ostatni warunek jest równoważny warunkowi $x_1 \wedge \ldots \wedge x_p \wedge y = 0$ (tw. 6), co kończy dowód. □

Jesteśmy teraz przygotowani do wykazania bliskiego związku pomiędzy prostymi p-wektorami a p-wymiarowymi podprzestrzeniami.

Twierdzenie 9. *Niezerowe p-wektory proste $x_1 \wedge \ldots \wedge x_p$ i $y_1 \wedge \ldots \wedge y_p$ są wzajemnie proporcjonalne wtedy, i tylko wtedy, gdy $L(x_1, \ldots, x_p) = L(y_1, \ldots, y_p)$. Równoważność ta ustala wzajemnie jednoznaczną odpowiedniość pomiędzy p-wymiarowymi podprzestrzeniami przestrzeni V a klasami niezerowych, wzajemnie proporcjonalnych p-wektorów prostych. Jeśli te warunki są spełnione*

$$i \quad y_i = \sum_{j=1}^{p} x_j \gamma^j{}_i \,, \quad to \qquad y_1 \wedge \ldots \wedge y_p = \det \boldsymbol{\gamma}\; x_1 \wedge \ldots \wedge x_p \,.$$

Dowód. Jeśli $x_1 \wedge \ldots \wedge x_p = \alpha\, y_1 \wedge \ldots \wedge y_p \neq 0$, to $x_1 \wedge \ldots \wedge x_p \wedge y_i = 0$ dla każdego $i = 1, \ldots, p$, więc z poprzedniego twierdzenia

$$L(y_1, \ldots, y_p) \subseteq L(x_1, \ldots, x_p) .$$

Zamieniając x-y z y-kami rolami dostajemy inkluzję przeciwną, więc równość. Odwrotnie, niech układy wektorów $x_1, \ldots, x_p$ oraz $y_1, \ldots, y_p$ będą liniowo niezależne (co jest równoważne nieznikaniu utworzonych z nich p-wektorów prostych), a ich powłoki liniowe równe. Wtedy $y_i = \sum_{j=1}^{p} x_j \gamma^j{}_i$ dla pewnej nieosobliwej macierzy $\boldsymbol{\gamma}$. Przy użyciu tych rozkładów i wykorzystaniu punktu (ii) twierdzenia 6 dostajemy ciąg równości

$$\begin{aligned} y_1 \wedge \ldots \wedge y_p &= \sum_{i_1, \ldots, i_p} x_{i_1} \wedge \ldots \wedge x_{i_p} \gamma^{i_1}{}_1 \ldots \gamma^{i_p}{}_p = \\ &= \sum_{i_1, \ldots, i_p} \epsilon_{i_1 \ldots i_p}\, \gamma^{i_1}{}_1 \ldots \gamma^{i_p}{}_p\, x_1 \wedge \ldots \wedge x_p = \det \boldsymbol{\gamma}\, x_1 \wedge \ldots \wedge x_p \,, \end{aligned}$$

co kończy dowód równoważności oraz końcowego zdania twierdzenia. Stwierdzenie o odpowiedniości jest podsumowaniem treści równoważności. □

Natychmiastowym wnioskiem z ostatniego twierdzenia jest następujące kryterium.

Wniosek 10. *Niech V będzie przestrzenią rzeczywistą, a W jej dowolną podprzestrzenią p-wymiarową. Bazy uporządkowane $(e_1, \dots, e_p)$ i $(e'_1, \dots, e'_p)$ podprzestrzeni W mają tę samą orientację wtedy, i tylko wtedy, gdy czynnik proporcjonalności między p-wektorami $e_1 \wedge \dots \wedge e_p$ i $e'_1 \wedge \dots \wedge e'_p$ jest dodatni.*

Wszystkie wyniki tego punktu pozostają w mocy po zamianie p-wektorów na p-formy, przy równoczesnym zastąpieniu przestrzeni V przestrzenią V^*.

8 Przykłady

(i) Współrzędne p-wektorów i p-form
Rozważając współrzędne tensorów tego typu należy zwrócić uwagę na pewien subtelny, ale istotny punkt. Każdy tensor t o walencji $\left[\begin{smallmatrix} p \\ 0 \end{smallmatrix}\right]$ rozkłada się w bazie iloczynowej:

$$t = t^{i_1 \dots i_p} e_{i_1} \otimes \dots \otimes e_{i_p} .$$

Jeśli t jest p-wektorem, więc $t = \mathcal{A}t$, to ten sam rozkład można zapisać jako

$$t = t^{i_1 \dots i_p} e_{i_1} \wedge \dots \wedge e_{i_p}$$

Prawa strona jest kombinacją liniową wektorów bazowych przestrzeni $\bigwedge^p(V)$, ale o składnikach powtarzających się, gdyż

$$t^{i_1 \dots i_p} e_{i_1} \wedge \dots \wedge e_{i_p} = t^{j_1 \dots j_p} e_{j_1} \wedge \dots \wedge e_{j_p} \quad \text{— bez sumowania,}$$
$$\text{jeśli ciąg } (j_1, \dots, j_p) \text{ jest wynikiem permutacji ciągu } (i_1, \dots, i_p)$$

– permutacja w $t^{i_1 \dots i_p}$ oraz w $e_{i_1} \wedge \dots \wedge e_{i_p}$ prowadzi jedynie do takiej samej zmiany znaku każdego z wyrażeń. Grupując jednakowe wyrazy dostajemy rozkład p-wektora t w bazie przestrzeni $\bigwedge^p(V)$:

$$t = \sum_{i_1 < \dots < i_p} p!\, t^{i_1 \dots i_p} e_{i_1} \wedge \dots \wedge e_{i_p} .$$

Część literatury matematycznej nazywa współczynniki tego rozkładu współrzędnymi p-wektora t; powyższy wzór daje ich związek ze współrzędnymi *tensorowymi* tensora t. W tej książce mówimy zawsze o współrzędnych tensorowych.

Niech np. w przestrzeni czterowymiarowej t będzie 3-wektorem danym rozkładem

$$t = 7e_1 \wedge e_2 \wedge e_3 - 3e_2 \wedge e_3 \wedge e_4 .$$

Odczytujemy jego współrzędne t^{ijk}: jedyne *niezależne* niezerowe wśród nich to

$$t^{123} = \frac{7}{6}\,, \qquad t^{234} = -\frac{1}{2}\,,$$

a pozostałe otrzymuje się przez odpowiednią permutację wskaźników z ewentualną zmianą znaku.

(ii) Zewnętrzny kwadrat 2-wektora
Jeśli t jest 2-wektorem, to $t \wedge t$ jest 4-wektorem. Przepis na współrzędne upraszcza się w tym przypadku do wzoru

$$t^{[ij}t^{kl]} = \frac{1}{3}\big(t^{ij}t^{kl} + t^{ki}t^{jl} + t^{jk}t^{il}\big)\,.$$

Przez porównanie z przykładem (i), p. 5 (z zamianą p-form na p-wektory), dostajemy więc tożsamość

$$t^{[ij}t^{kl]} = t^{[ij}t^{k]l}\,.$$

(iii) Anihilator i prostota p-wektora
Niech t będzie dowolnym p-wektorem. Jego ***anihilatorem*** nazywamy podprzestrzeń przestrzeni V złożoną z wszystkich wektorów y takich, że $t \wedge y = 0$. Twierdzenie 8 można teraz ująć tak: wektory $x_1, \ldots, x_p$ tworzą bazę anihilatora niezerowego p-wektora prostego $x_1 \wedge \ldots \wedge x_p$. Twierdzenie to posiada następujące uogólnienie obejmujące dowolne niezerowe p-wektory.

> Jeśli liniowo niezależne wektory $x_1, \ldots, x_r$ są elementami anihilatora niezerowego p-wektora t, to istnieje $p-r$-wektor s taki, że
>
> $$t = s \wedge x_1 \wedge \ldots \wedge x_r\,.$$

Idea dowodu polega na uzupełnieniu ciągu wektorów x_i do bazy całej przestrzeni $x_1, \ldots, x_n$ i rozłożeniu tensora t w tej bazie. Szczegóły zostawiamy jako ćwiczenie dla czytelnika. Tensor s nie jest jednoznacznie zadany warunkiem tezy, ale metoda dowodu pokazuje, że można dodatkowo przyjąć, że s jest utworzony z wektorów $x_{r+1}, \ldots, x_n$, i wtedy rozwiązanie jest jednoznaczne. W takim przypadku można podać zwarty przepis na s. Niech $x^1, \ldots, x^n$ będzie bazą dualną. Warunek tezy zapisany we współrzędnych w dowolnej bazie ma postać

$$t^{i_1\ldots i_p} = s^{[i_1\ldots i_{p-r}}{x_1}^{i_{p-r+1}} \ldots {x_r}^{i_p]}\,.$$

Zwężenie tego równania z ${x^1}_{i_{p-r+1}} \ldots {x^r}_{i_p}$ prowadzi do równości

$$s^{i_1\ldots i_{p-r}} = \frac{p!}{(p-r)!}\, t^{i_1\ldots i_p}{x^1}_{i_{p-r+1}} \ldots {x^r}_{i_p}\,,$$

gdyż przy dodatkowym warunku zwężenie s z dowolną z form $x^1, \dots, x^r$ (po dowolnym wskaźniku) daje zero.

Wnioskiem z powyższego twierdzenia jest następujące kryterium.

> Wymiar anihilatora niezerowego p-wektora jest mniejszy lub równy p; równość zachodzi wtedy, i tylko wtedy, gdy p-wektor jest prosty.

(iv) Wyliczenie anihilatora p-wektora prostego
W rzeczywistej przestrzeni czterowymiarowej dany jest 2-wektor

$$t = e_1 \wedge e_2 + 4e_1 \wedge e_3 - 5e_1 \wedge e_4 + 11e_2 \wedge e_3 - 10e_2 \wedge e_4 + 15e_3 \wedge e_4\,.$$

Chcemy stwierdzić, czy jest to 2-wektor prosty, i jeśli tak, to znaleźć jego rozkład.

Szukamy anihilatora. Niech $x = x^i e_i$ będzie dowolnym wektorem. Wtedy

$$\begin{aligned} t \wedge x = {} & (11x^1 - 4x^2 + x^3) e_1 \wedge e_2 \wedge e_3 + (-10x^1 + 5x^2 + x^4) e_1 \wedge e_2 \wedge e_4 + \\ & + (15x^1 + 5x^3 + 4x^4) e_1 \wedge e_3 \wedge e_4 + (15x^2 + 10x^3 + 11x^4) e_2 \wedge e_3 \wedge e_4\,. \end{aligned}$$

Warunek $t \wedge x = 0$ jest równoważny warunkowi znikania współczynników przy wszystkich 3-wektorach bazowych. Otrzymany układ równań liniowych na $\mathbf{x}$ ma dwa liniowo niezależne rozwiązania, dające bazę anihilatora:

$$x_1 = e_1 + 2e_2 - 3e_3\,, \qquad x_2 = e_2 + 4e_3 - 5e_4\,.$$

2-wektor t jest więc prosty, proporcjonalny do $x_1 \wedge x_2$ (wykorzystujemy wynik przykładu (iii)). Sprawdzamy czynnik proporcjonalności i dostajemy $t = x_1 \wedge x_2$.

(v) Bezpośrednie kryterium prostoty p-wektorów
Istnieje inna, bardziej bezpośrednia postać kryterium prostoty p-wektorów: p-wektor t jest prosty wtedy, i tylko wtedy, gdy spełnione jest równanie

$$t^{[i_1 \dots i_p} t^{j_1] j_2 \dots j_p} = 0\,.$$

Naszkicujemy ideę dowodu. Jeśli t jest prosty, to łatwo sprawdzić, że równanie jest spełnione. Odwrotnie, jeśli $t \neq 0$ (przypadek $t = 0$ jest trywialny), to zawsze istnieją 1-formy $y^2, \dots, y^p$ takie, że

$$x^{j_1} \equiv t^{j_1 \dots j_p} y^2{}_{j_2} \dots y^p{}_{j_p} \neq 0\,.$$

Przez zwężenie warunku kryterium z tymi formami dostajemy więc $t \wedge x = 0$, a stąd $t = s \wedge x$. Korzystając z jednego z rezultatów przykładu (iii) możemy przyjąć $s^{i_1 \dots i_{p-1}} = p\, t^{i_1 \dots i_p} \varphi_{i_p}$, gdzie φ jest taką 1-formą, że $\varphi(x) = 1$. Zwężamy teraz warunek kryterium z $\varphi_{i_p}\varphi_{j_p}$. Przy wykorzystaniu tożsamości podanej w przykładzie (iii), p. 5, dostajemy

$$s^{[i_1\dots i_{p-1}}s^{j_1]j_2\dots j_{p-1}} = 0\,.$$

Dowód kończy się teraz przez indukcję względem p.

Porównanie warunku kryterium z wynikiem przykładu (ii) pozwala podać jeszcze inną postać kryterium dla $p = 2$: 2-wektor t jest prosty wtedy, i tylko wtedy, gdy

$$t \wedge t = 0\,.$$

(vi) Operatory $\bigwedge^k(A)$

Niech B będzie operatorem w przestrzeni $\mathcal{T}^k_0(V)$. Jeśli $[B, P^\pi] = 0$ dla każdej permutacji $\pi \in S_k$, to $[B, \mathcal{A}] = 0$, więc podprzestrzeń $\bigwedge^k(V) = \mathcal{A}\mathcal{T}^k_0(V)$ jest inwariantna względem B (przyjęty warunek jest jednak silniejszy od koniecznego). Rozważany warunek jest spełniony dla operatora

$$A^{\otimes k} = \underbrace{A \otimes \ldots \otimes A}_{k \text{ razy}}\,.$$

Zawężenie tego operatora do operatora na $\bigwedge^k(V)$ oznacza się $\bigwedge^k(A)$.

W podobny sposób można wprowadzać operatory na przestrzeniach symetrycznych tensorów $\mathcal{S}\mathcal{T}^k_0(V)$.

9 Przestrzenie $\bigwedge^n(V)$ i $\bigwedge^n(V^*)$

Z twierdzenia 7 wynika w szczególności, że przestrzenie $\bigwedge^n(V)$ i $\bigwedge^n(V^*)$ są jednowymiarowe, o wektorach bazowych odpowiednio $e_1 \wedge \ldots \wedge e_n \neq 0$ oraz $e^1 \wedge \ldots \wedge e^n \neq 0$, gdzie $(e_1, \ldots, e_n)$ jest dowolną bazą V. Stąd każdy n-wektor jest n-wektorem prostym i każda n-forma – n-formą prostą. Niech $\omega = \alpha\, e^1 \wedge \ldots \wedge e^n$. Wprost z definicji antysymetryzacji jest widoczne, że współrzędna $\omega_{1\dots n}$ tensora ω w bazie (e_i) jest równa $\alpha/n!$, więc każdą n-formę można zapisać jako

$$\omega = \omega_{1\dots n}\, n!\, e^1 \wedge \ldots \wedge e^n\,,$$

i podobnie dla n-wektorów

$$u = u^{1\dots n}\, n!\, e_1 \wedge \ldots \wedge e_n\,.$$

Dla innych współrzędnych w tej bazie dostajemy

$$\omega_{i_1\dots i_n} = \omega_{1\dots n}\epsilon_{i_1\dots i_n}\,, \qquad u^{i_1\dots i_n} = u^{1\dots n}\epsilon^{i_1\dots i_n}\,,$$

gdzie $\epsilon_{i_1\dots i_n}$ i $\epsilon^{i_1\dots i_n}$ są symbolami Levi-Civity (wynika to natychmiast stąd, że $\omega_{i_1\dots i_n}$ i $u^{i_1\dots i_n}$ są symbolami całkowicie antysymetrycznymi). Zauważmy, że przez przeskalowanie bazy (e_i) można zawsze uzyskać $\omega_{1\dots n} = 1$ lub $u^{1\dots n} = 1$. Pomnożymy tensorowo tensory u i ω, i otrzymany tensor $u \otimes \omega \in \mathcal{T}^n_n(V)$ zwęzimy po wszystkich wskaźnikach (pierwszy kontrawariantny z pierwszym kowariantnym, itd.). Dostajemy w wyniku liczbę

$$u^{i_1\dots i_n}\omega_{i_1\dots i_n} = n!\, u^{1\dots n}\omega_{1\dots n}\,.$$

Możemy teraz natychmiast wyciągnąć następujący wniosek.

Twierdzenie 11. *Warunek*

$$\omega_{i_1\dots i_n}\hat{\omega}^{i_1\dots i_n} = n!$$

ustala wzajemnie jednoznaczny związek n-form ω i n-wektorów $\hat{\omega}$. Równoważna postać tego warunku to:

$$\omega = n!\, e^1 \wedge \ldots \wedge e^n\,, \qquad \hat{\omega} = n!\, e_1 \wedge \ldots \wedge e_n$$

w pewnej bazie (e_i).

10 Częściowe zwężenie n-formy ω z n-wektorem $\hat{\omega}$

Niech $\omega \in \bigwedge^n(V^*)$ i $\hat{\omega} \in \bigwedge^n(V)$ będą związane takim warunkiem jak w ostatnim twierdzeniu, a więc w dowolnym układzie współrzędnych

$$\omega_{i_1\dots i_n} = \omega_{1\dots n}\epsilon_{i_1\dots i_n} \quad \text{i} \quad \hat{\omega}^{i_1\dots i_n} = \hat{\omega}^{1\dots n}\epsilon^{i_1\dots i_n}, \quad \text{gdzie} \quad \omega_{1\dots n}\,\hat{\omega}^{1\dots n} = 1\,.$$

Dokonując w iloczynie $\hat{\omega} \otimes \omega$ zwężenia ostatnich $(n-p)$ wskaźników kontrawariantnych z kowariantnymi dostajemy tensor t o współrzędnych

$$t^{i_1\dots i_p}{}_{j_1\dots j_p} = \epsilon^{i_1\dots i_p k_1\dots k_{n-p}}\epsilon_{j_1\dots j_p k_1\dots k_{n-p}}$$

(dopuszczamy tu również przypadek $p = n$, a więc bez kontrakcji). Tensor ten jest z konstrukcji całkowicie antysymetryczny w górnych wskaźnikach i całkowicie antysymetryczny we wskaźnikach dolnych, więc dla wyliczenia jego niezależnych składowych wystarczy przyjąć $i_1 < \ldots < i_p$ oraz $j_1 < \ldots < j_p$. Jeśli te ciągi różnią się, to przy każdym układzie wskaźników sumacyjnych $k_1 \ldots, k_{n-p}$ jeden z czynników $\epsilon^{i_1\dots i_p k_1\dots k_{n-p}}$ lub $\epsilon_{j_1\dots j_p k_1\dots k_{n-p}}$ zeruje się, więc taka współrzędna tensora t jest równa zeru. Jeśli $(i_1, \ldots, i_p) = (j_1, \ldots, j_p)$, to symbole Levi-Civity nie zerują się dla takich ciągów wskaźników $k_1, \ldots, k_{n-p}$, że $(i_1, \ldots, i_p, k_1, \ldots, k_{n-p})$ jest efektem permutacji liczb $1, \ldots, n$. Układów takich jest $(n-p)!$, więc

$$t^{i_1\dots i_p}{}_{i_1\dots i_p} = (n-p)!\,, \quad \text{dla } i_1 < \ldots < i_p\,, \text{ bez sumowania.}$$

Utwórzmy teraz inny tensor s przez p-krotne pomnożenie tensorowe tensora δ, a następnie zantysymetryzowanie otrzymanego tensora w górnych wskaźnikach. Współrzędne otrzymanego tensora w dowolnej bazie są równe

$$s^{i_1\dots i_p}{}_{j_1\dots j_p} = \delta^{[i_1}{}_{j_1}\dots\delta^{i_p]}{}_{j_p}\,.$$

Prawą stronę tej równości możemy przekształcić w następujący sposób:

$$\frac{1}{p!}\sum_{\pi\in S_p}\operatorname{sgn}\pi\,\delta^{i_{\pi(1)}}{}_{j_1}\dots\delta^{i_{\pi(p)}}{}_{j_p} = \frac{1}{p!}\sum_{\pi^{-1}\in S_p}\operatorname{sgn}\pi^{-1}\,\delta^{i_1}{}_{j_{\pi^{-1}(1)}}\dots\delta^{i_p}{}_{j_{\pi^{-1}(p)}}\,,$$

gdzie wykorzystaliśmy równość $\delta^{i_{\pi(k)}}{}_{j_k} = \delta^{i_l}{}_{j_{\pi^{-1}(l)}}$ dla $l = \pi(k)$ oraz przemienność czynników w iloczynie. Stąd antysymetryzacja iloczynu tensorów δ w dolnych wskaźnikach daje efekt identyczny jak antysymetryzacja we wskaźnikach górnych:

$$s^{i_1\dots i_p}{}_{j_1\dots j_p} = \delta^{i_1}{}_{[j_1}\dots\delta^{i_p}{}_{j_p]}\,.$$

Tensor s, podobnie jak t, jest więc całkowicie antysymetryczny w górnych wskaźnikach i całkowicie antysymetryczny we wskaźnikach dolnych. Liczymy jego niezależne współrzędne w podobny sposób, jak dla tensora t. Zakładamy, że $i_1 < \dots < i_p$ i $j_1 < \dots < j_p$. Jeśli $(i_1,\dots,i_p) \neq (j_1,\dots,j_p)$, to odpowiednia współrzędna znika (co najmniej jeden z czynników δ musi być równy zeru). Jeśli $(i_1,\dots,i_p) = (j_1,\dots,j_p)$, to niezerowy przyczynek do współrzędnej pochodzi jedynie od permutacji tożsamościowej, więc

$$s^{i_1\dots i_p}{}_{i_1\dots i_p} = \frac{1}{p!}\,,\quad \text{dla } i_1 < \dots < i_p\,, \text{ bez sumowania.}$$

Porównując wyniki dla t i s widzimy, że $t = p!(n-p)!\ s$. Zapisujemy ten wynik jako twierdzenie.

Twierdzenie 12. *Jeśli $\omega \in \bigwedge^n(V^*)$ i $\hat\omega \in \bigwedge^n(V)$ związane są warunkiem*

$$\omega_{i_1\dots i_n}\hat\omega^{i_1\dots i_n} = n!\,,$$

to dla każdego $p = 0,\dots,n$ zachodzą równości tensorowe

$$\begin{aligned}\hat\omega^{i_1\dots i_p k_1\dots k_{n-p}}\,\omega_{j_1\dots j_p k_1\dots k_{n-p}} &= p!(n-p)!\ \delta^{[i_1}{}_{j_1}\dots\delta^{i_p]}{}_{j_p} = \\ &= p!(n-p)!\ \delta^{i_1}{}_{[j_1}\dots\delta^{i_p}{}_{j_p]}\,.\end{aligned}$$

11 Odwzorowania dualności między $\bigwedge^p(V)$ i $\bigwedge^{n-p}(V^*)$

Wybierzmy niezerową n-formę ω. Zadaje ona odwzorowanie liniowe

$$\bigwedge^p(V) \ni t \mapsto {}^*t \in \bigwedge^{n-p}(V^*)$$

w następujący sposób: dla każdego $t \in \bigwedge^p(V)$ tworzymy tensor $\dfrac{1}{(n-p)!}\, t \otimes \omega$ i zwężamy wszystkie wskaźniki tensora t z p początkowymi wskaźnikami tensora ω (z zachowaniem kolejności wskaźników). Otrzymany tensor to z definicji ${}^*t$. Przepis tego odwzorowania najwygodniej jest przedstawić w notacji wskaźnikowej:

$$ {}^*t_{i_1\dots i_{n-p}} = \frac{1}{(n-p)!}\, t^{k_1\dots k_p}\, \omega_{k_1\dots k_p i_1\dots i_{n-p}} \;;$$

nazywamy to przyporządkowanie ***odwzorowaniem dualności*** p***-wektorów na*** $(n-p)$***-formy***. Zauważmy, że

$$\dim \bigwedge^p(V) = \binom{n}{p} = \binom{n}{n-p} = \dim \bigwedge^{n-p}(V^*)\,.$$

Pokażemy, że odwzorowanie dualności jest izomorfizmem. Niech $\hat{\omega}$ będzie n-wektorem związanym z ω warunkiem normalizacyjnym, jak w poprzednim punkcie. Tworzymy tensor $\dfrac{1}{p!}\,\hat{\omega} \otimes {}^*t$ i zwężamy wszystkie wskaźniki tensora ${}^*t$ z $n-p$ końcowymi wskaźnikami tensora $\hat{\omega}$. W notacji wskaźnikowej dostajemy

$$\begin{aligned}\frac{1}{p!}\hat{\omega}^{j_1\dots j_p i_1\dots i_{n-p}}\, {}^*t_{i_1\dots i_{n-p}} &= \frac{1}{p!(n-p)!}\,\hat{\omega}^{j_1\dots j_p i_1\dots i_{n-p}}\, \omega_{k_1\dots k_p i_1\dots i_{n-p}}\, t^{k_1\dots k_p} = \\ &= \delta^{j_1}{}_{[k_1} \dots \delta^{j_p}{}_{k_p]}\, t^{k_1\dots k_p} = t^{j_1\dots j_p}\,,\end{aligned}$$

gdzie w drugiej równości wykorzystaliśmy tożsamość daną twierdzeniem 12, a w trzeciej punkt (iv) twierdzenia 5. Stąd, jeśli ${}^*t = 0$, to $t = 0$, więc odwzorowanie dualności jest izomorfizmem (trywialność jądra implikuje to przy równości wymiarów dziedziny i przeciwdziedziny). Ponadto, uzyskaliśmy przepis odwzorowania odwrotnego

$$\bigwedge^{n-p}(V^*) \ni s \mapsto \hat{{}^*}s \in \bigwedge^p(V)\,,$$

$$\hat{{}^*}s^{i_1\dots i_p} = \frac{1}{p!}\,\hat{\omega}^{i_1\dots i_p k_1\dots k_{n-p}}\, s_{k_1\dots k_{n-p}}\,.$$

Dla każdego p-wektora t i p-formy s mamy więc

$$\hat{{}^*}{}^*t = t\,, \qquad {}^*\hat{{}^*}s = s\,.$$

Przypomnijmy, że odwzorowanie dualności jest określone liniowo przez wybraną formę ω: jeśli ją przeskalować o czynnik, to odwzorowanie dualności skaluje się o ten sam czynnik, a odwrotne odwzorowanie dualności – o czynnik odwrotny.

12 Dualność p-wektorów i $(n-p)$-form prostych

Odwzorowanie dualności, jak pokazuje następne twierdzenie, ustala w szczególności wzajemnie jednoznaczną odpowiedniość p-wektorów prostych z $(n-p)$-formami prostymi. Zanim je sformułujemy, przypomnijmy, że jeśli $x_1, \dots, x_n$ jest bazą przestrzeni V, to podprzestrzeń $W = L(x_1, \dots, x_p)$ jest wyznaczona p-wektorem prostym $x_1 \wedge \dots \wedge x_p$ (zgodnie z twierdzeniem 9). Wtedy też podprzestrzeń $W^{\perp^*} = L(x^{p+1}, \dots, x^n) \subseteq V^*$, gdzie $x^1, \dots, x^n$ jest bazą dualną, jest wyznaczona $(n-p)$-formą prostą $x^{p+1} \wedge \dots \wedge x^n$.

Twierdzenie 13.
(i) Jeśli $(x_1, \dots, x_n)$ jest bazą V, a $(x^1, \dots, x^n)$ bazą do niej dualną, to

$$
\begin{aligned}
{}^*(x_1 \wedge \dots \wedge x_p) &= \omega(x_1, \dots, x_n)\, x^{p+1} \wedge \dots \wedge x^n\,, \\
\hat{{}^*}(x^{p+1} \wedge \dots \wedge x^n) &= \hat\omega(x^1, \dots, x^n)\, x_1 \wedge \dots \wedge x_p\,;
\end{aligned}
$$

n-forma ω i n-wektor $\hat\omega$ traktowane są tutaj jako odwzorowania wieloliniowe.
(ii) p-wektor prosty t wyznacza podprzestrzeń $W \subseteq V$ wtedy, i tylko wtedy, gdy $(n-p)$-forma dualna ${}^*t$ wyznacza podprzestrzeń $W^{\perp^*} \subseteq V^*$.

Dowód. (i) Dowodzimy pierwszej równości. Z jednowymiarowości przestrzeni $\bigwedge^n(V)$ mamy $\omega = \lambda\, n!\, x^1 \wedge \dots \wedge x^n$. Używając tego przedstawienia mamy dalej ciąg równości:

$$
\begin{aligned}
{}^*(x_1 \wedge \dots \wedge x_p)_{i_{p+1} \dots i_n} &= \frac{\lambda\, n!}{(n-p)!}\, {x_1}^{i_1} \dots {x_p}^{i_p}\, {x^1}_{[i_1} \dots {x^n}_{i_n]} = \\
&= \frac{\lambda}{(n-p)!} \sum_{\pi \in S_n} \operatorname{sgn}\pi\ {x_1}^{i_1} \dots {x_p}^{i_p}\, {x^1}_{i_{\pi(1)}} \dots {x^n}_{i_{\pi(n)}} = \\
&= \frac{\lambda}{(n-p)!} \sum_{\pi \in S_{\{p+1,\dots,n\}}} \operatorname{sgn}\pi\ {x^{p+1}}_{i_{\pi(p+1)}} \dots {x^n}_{i_{\pi(n)}} = \\
&\qquad\qquad = \lambda\, (x^{p+1} \wedge \dots \wedge x^n)_{i_{p+1} \dots i_n}\,.
\end{aligned}
$$

Pierwsza równość w tym ciągu wynika z definicji dualności i punktu (iv) twierdzenia 5, druga z definicji antysymetryzacji. Do sumy otrzymanej w wyniku drugiej równości niezerowy przyczynek dadzą tylko te permutacje, które nie zmieniają liczb $1, \dots, p$ (dualność baz), skąd wynika równość trzecia. Ostatnia równość znów z definicji antysymetryzacji. Pozostało jedynie zauważyć, że λ jest współrzędną $\omega_{1\dots n}$ w bazie (x_i), więc $\lambda = \omega(x_1, \dots, x_n)$. Ten sam wynik można też otrzymać kładąc w powyższym rachunku $p = n$:

$$
\lambda = {}^*(x_1 \wedge \dots \wedge x_n) = {x_1}^{i_1} \dots {x_n}^{i_n}\, \omega_{i_1 \dots i_n} = \omega(x_1, \dots, x_n)\,,
$$

co kończy dowód pierwszej równości punktu (i). Druga równość jest odwróceniem pierwszej ($\hat{\omega}(x^1, \ldots, x^n) = [\omega(x_1, \ldots, x_n)]^{-1}$, gdyż w dowolnej bazie te tensory związane są warunkiem $\omega_{1\ldots n}\hat{\omega}^{1\ldots n} = 1$). Punkt (ii) jest natychmiastowym wnioskiem z punktu (i). □

Wniosek 14. *Jeśli* $\dim V = n$, *to każdy* $(n-1)$*-wektor i każda* $(n-1)$*-forma są proste.*

Dowód. Każda $(n-1)$-forma jest dualna do 1-wektora, który zawsze jest prosty, więc teza wynika natychmiast z poprzedniego twierdzenia. Podobnie dla $(n-1)$-wektorów. □

13 Wyróżniona n-forma w przestrzeni rzeczywistej z symetryczną metryką. Związek z orientacją

Niech V będzie przestrzenią rzeczywistą wyposażoną w niezdegenerowaną metrykę symetryczną g. Istnieje wtedy, jak wiadomo, klasa baz ortonormalnych przestrzeni V. Jak pokazaliśmy w punkcie 4, §16, każde dwie bazy ortonormalne są związane izometrią, więc macierz przejścia między nimi jest elementem grupy $O(n-r, r)$, gdzie $(n-r, r)$ jest sygnaturą g. Stąd wyznacznik macierzy przejścia jest równy ± 1; w przypadku, gdy wynosi on $+1$, bazy mają tę samą orientację, a gdy wynosi -1 – mają orientacje przeciwne. Łącząc te fakty z tożsamością z twierdzenia 9 dostajemy

$$e'_1 \wedge \ldots \wedge e'_n = \pm\, e_1 \wedge \ldots \wedge e_n$$

dla dowolnych baz ortonormalnych, przy czym znak minus pojawia się wtedy, i tylko wtedy, gdy bazy mają przeciwną orientację. Łatwo przekonać się, że dla odpowiednich baz dualnych mamy wtedy również

$$e'^1 \wedge \ldots \wedge e'^n = \pm\, e^1 \wedge \ldots \wedge e^n\,,$$

przy czym znaki w ostatnich dwóch relacjach są zgodne.

Wybierzmy w przestrzeni V jedną z orientacji jako dodatnią. Powyższa dyskusja pokazuje, że istnieje wtedy wyróżniona n-forma

$$e := n!\, e^1 \wedge \ldots \wedge e^n\,,$$

gdzie $(e_1, \ldots, e_n)$ jest dowolną bazą ortonormalną dodatnio zorientowaną – definicja nie zależy od wyboru bazy w tej klasie. Definicja ta jest równoważna stwierdzeniu, że w dowolnej bazie ortonormalnej dodatnio zorientowanej

$$e_{1\ldots n} = +1\,.$$

Posługując się regułą podnoszenia wskaźników w bazie ortonormalnej, omówioną w punkcie 10, §19, dla współrzędnej kontrawariantnej tego tensora dostajemy

$$e^{1\dots n} = (-1)^r\,,$$

skąd dla n-wektora $\hat{e}$ związanego z e mamy

$$\hat{e}^{i_1\dots i_n} = (-1)^r e^{i_1\dots i_n}\,.$$

Podstawiając tę równość do tożsamości w twierdzeniu 12 dostajemy

$$\begin{aligned} e^{i_1\dots i_p k_1\dots k_{n-p}}\, e_{j_1\dots j_p k_1\dots k_{n-p}} &= (-1)^r\, p!(n-p)!\ \delta^{[i_1}{}_{j_1}\dots\delta^{i_p]}{}_{j_p} = \\ &= (-1)^r\, p!(n-p)!\ \delta^{i_1}{}_{[j_1}\dots\delta^{i_p}{}_{j_p]}\,. \end{aligned}$$

W szczególności, w przypadku przestrzeni euklidesowej $(-1)^r = 1$, a w przypadku przestrzeni Minkowskiego $(-1)^r = -1$.

Podkreślmy, że dla jednoznacznego ustalenia n-formy e konieczne jest zadanie orientacji. Przy zmianie orientacji forma określona powyższą definicją zmienia znak.

14 Dualność w przypadku przestrzeni z metryką symetryczną

W przypadku przestrzeni rzeczywistej z niezdegenerowaną symetryczną metryką odwzorowanie dualności, dzięki złożeniu z podnoszeniem lub opuszczaniem wskaźników, można interpretować jako odwzorowanie $\bigwedge^p(V) \mapsto \bigwedge^{n-p}(V)$:

$$^*t^{i_1\dots i_{n-p}} = \frac{1}{(n-p)!}\, t^{k_1\dots k_p}\, e_{k_1\dots k_p}{}^{i_1\dots i_{n-p}}\,,$$

lub odwzorowanie $\bigwedge^p(V^*) \mapsto \bigwedge^{n-p}(V^*)$:

$$^*t_{i_1\dots i_{n-p}} = \frac{1}{(n-p)!}\, t_{k_1\dots k_p}\, e^{k_1\dots k_p}{}_{i_1\dots i_{n-p}}\,.$$

Przy tej interpretacji odwzorowanie odwrotne $\hat{*}$ można wyrazić jako

$$\begin{aligned} {}^{\hat{*}}s^{i_1\dots i_p} &= \frac{1}{p!}\, \hat{e}^{i_1\dots i_p}{}_{k_1\dots k_{n-p}}\, s^{k_1\dots k_{n-p}} = \\ &= (-1)^{r+p(n-p)}\, \frac{1}{p!}\, s^{k_1\dots k_{n-p}}\, e_{k_1\dots k_{n-p}}{}^{i_1\dots i_p} = (-1)^{r+p(n-p)}\, {}^*s^{i_1\dots i_p}\,. \end{aligned}$$

Druga równość wynika z wyrażenia $\hat{e}$ przez e oraz z przesunięcia wskaźników $k_1,\dots,k_{n-p}$ w tensorze e przed wskaźniki $i_1,\dots,i_p$. Stąd w działaniu na $t \in \bigwedge^p(V)$ mamy

$$^{**}t = (-1)^{r+(n-1)p}\, t\,,$$

gdzie wzięliśmy pod uwagę, że $(n-1)p - p(n-p) = p(p-1)$ jest liczbą parzystą.

15 Iloczyn wektorowy w trójwymiarowej przestrzeni euklidesowej

W przestrzeni euklidesowej trójwymiarowej dualność wiąże trójwektory (lub trójformy) ze skalarami (0-formami, 0-wektorami) oraz dwuwektory (lub dwuformy) z jednoformami (lub jednowektorami). Dla wektora x, 2-wektora t oraz 3-wektora s dostajemy z ogólnych wzorów

$$ {}^*x^{jk} = \frac{1}{2}\, x^i\, e_i{}^{jk}\,, \quad {}^*t^i = t^{jk}\, e_{jk}{}^i\,, \quad {}^{**}x = x\,, \quad {}^{**}t = t\,, $$
$$ {}^*s = s^{ijk}\, e_{ijk}\,. $$

Ponadto, w trzech wymiarach wszystkie p-wektory i p-formy są proste (3-wektory i 3-formy na mocy punktu 9, a 2-wektory i 2-formy na mocy wniosku 14). Wektor dualny do $x \wedge y$ nazywamy ***iloczynem wektorowym wektorów*** x ***i*** y i oznaczamy:

$$ x \times y := {}^*(x \wedge y)\,, \quad \text{czyli} \quad (x \times y)^i = x^j y^k e_{jk}{}^i = e^i{}_{jk}\, x^j y^k\,. $$

Iloczyn wektorowy ma następujące dwie podstawowe własności:

$$ x \times y = -y \times x\,, $$
$$ \text{wektory } x \text{ i } y \text{ są równoległe} \quad \Longleftrightarrow \quad x \times y = 0\,, $$

– pierwsza wynika wprost z definicji, a druga z punktu (i) twierdzenia 6 (przez równoległość dwóch wektorów rozumiemy ich liniową zależność).

Skalar (czyli 0-wektor) dualny do 3-formy $x \wedge y \wedge z$ ma postać $x^i y^j z^k\, e_{ijk}$. Jeśli iloczyn skalarny wektorów a i b oznaczyć przez $a\cdot b$, to z pierwszej postaci wskaźnikowej wektora $x \times y$ jest natychmiast widoczna następująca tożsamość:

$$ {}^*(x \wedge y \wedge z) = (x \times y) \cdot z = x^i y^j z^k\, e_{ijk}\,; $$

nazywamy tę wielkość ***iloczynem mieszanym wektorów*** x***,*** y ***i*** z. Każda cykliczna permutacja zbioru trójelementowego jest parzysta, skąd wynikają tożsamości

$$ (x \times y) \cdot z = (y \times z) \cdot x = (z \times x) \cdot y\,. $$

Podstawiając tu w miejsce z wektor x lub y dostajemy w szczególności

$$ x \times y \perp x\,, \quad x \times y \perp y\,. $$

Tożsamości dla zwężeń po jednym i po dwóch wskaźnikach dwóch tensorów e przyjmują postać

$$ e^{i_1 i_2 k}\, e_{j_1 j_2 k} = 2\delta^{[i_1}{}_{j_1}\delta^{i_2]}{}_{j_2} = \delta^{i_1}{}_{j_1}\delta^{i_2}{}_{j_2} - \delta^{i_2}{}_{j_1}\delta^{i_1}{}_{j_2} = 2\delta^{i_1}{}_{[j_1}\delta^{i_2}{}_{j_2]}\,, $$
$$ e^{i k_1 k_2}\, e_{j k_1 k_2} = 2\delta^i{}_j\,. $$

Posługując się pierwszą z nich mamy

$$
\begin{aligned}
[\,x\times(y\times z)]^i &= e^i{}_{jk}x^j e^k{}_{lm}\,y^l z^m = e^{ijk}\,e_{lmk}\,x_j\,y^l z^m = \\
&= (\delta^i{}_l\delta^j{}_m - \delta^j{}_l\delta^i{}_m)\,x_j\,y^l z^m = x_j z^j\,y^i - x_j y^j\,z^i = x^k g_{kj} z^j\,y^i - x^k g_{kj} y^j\,z^i = \\
&= (x\cdot z)\,y^i - (x\cdot y)\,z^i\,.
\end{aligned}
$$

Rezultat zapisujemy w postaci pierwszej z poniższych tożsamości wektorowych, a drugą można uzyskać z pierwszej przy użyciu antysymetrii iloczynu wektorowego:

$$x\times(y\times z) = (x\cdot z)\,y - (x\cdot y)\,z\,,\quad (x\times y)\times z = (x\cdot z)\,y - (y\cdot z)\,x\,.$$

Skorzystamy z cykliczności iloczynu mieszanego oraz z drugiej z powyższych tożsamości dla wyliczenia długości wektora $x\times y$:

$$(x\times y)\cdot(x\times y) = [(x\times y)\times x]\cdot y = [(x\cdot x)\,y - (x\cdot y)\,x]\cdot y = (x\cdot x)\,(y\cdot y) - (x\cdot y)^2\,.$$

Podstawiając $x\cdot y = \|x\|\,\|y\|\cos\measuredangle(x,y)$ dostajemy

$$\|x\times y\| = \|x\|\,\|y\|\sin[\measuredangle(x,y)]\,.$$

Niech n będzie wektorem jednostkowym. Zgodnie z twierdzeniem 4, §14, istnieje ortogonalny rozkład

$$V = L(n)\oplus L(n)^{\perp}\,,\quad x = x_{\|} + x_{\perp}\,.$$

Licząc iloczyn skalarny tego rozkładu z wektorem n oraz jego dwukrotny iloczyn wektorowy z wektorem n (szczegóły rachunków pozostawiamy jako ćwiczenie dla czytelnika) dostajemy

$$x_{\|} = (n\cdot x)\,n\,,\quad x_{\perp} = (n\times x)\times n\,.$$

Stąd

$$\|x_{\|}\| = |\,n\cdot x\,|\,,\qquad \|x_{\perp}\| = \|\,n\times x\,\|\,.$$

Zauważmy na koniec, że przyporządkowanie parze wektorów x, y ich iloczynu wektorowego $x\times y$ jest odwzorowaniem zbioru $V\times V$ w zbiór V, a więc działaniem w zbiorze V. To działanie nie jest łączne, nie posiada elementu neutralnego, i jest nieprzemienne.

16 Przykłady

(i) Wyznacznik operatora liniowego
Dla każdego operatora $A \in \mathcal{L}(V, V)$ określiliśmy w przykładzie (vi), p. 8, operatory $\bigwedge^k(A)$. W szczególności, operator $\bigwedge^n(A)$ działa w jednowymiarowej przestrzeni $\bigwedge^n(V)$. Każdy operator w jednowymiarowej przestrzeni jest proporcjonalny do operatora identycznościowego na tej przestrzeni. Korzystając z twierdzenia 9 łatwo pokazać, że czynnikiem proporcjonalności jest w przypadku operatora $\bigwedge^n(A)$ wyznacznik operatora A,

$$\bigwedge\nolimits^n(A) = \det A \ \mathrm{id} \ .$$

W istocie, relacja ta może być wykorzystana do *zdefiniowania* wyznacznika operatora. Notujemy też dwie pokrewne formuły:

$$A^{[i_1}{}_{j_1} \dots A^{i_n]}{}_{j_n} = \det A \, \delta^{[i_1}{}_{j_1} \dots \delta^{i_n]}{}_{j_n}$$

oraz

$$\det A = A^{[i_1}{}_{i_1} \dots A^{i_n]}{}_{i_n} \, .$$

Wstawiając tutaj w miejsce A operator $A - \lambda \, \mathrm{id}$ można też uzyskać związek współczynników wielomianu charakterystycznego z wielkościami $\mathrm{Tr}\big[\bigwedge^k(A)\big]$, $k = 1, \dots, n$.

(ii) Warunek prostoty 2-wektora w przestrzeni czterowymiarowej
W czterowymiarowej przestrzeni wszystkie p-wektory, z wyjątkiem 2-wektorów, są zawsze proste. Wybierzmy 4-formę ω i związany z nią 4-wektor $\hat{\omega}$. Dla każdego 2-wektora F w takiej przestrzeni mamy

$$F \wedge F = \lambda \, \hat{\omega} \, , \quad \text{gdzie} \quad \lambda = \frac{1}{4!} \, \omega_{ijkl} F^{ij} F^{kl} = \frac{1}{12} {}^*\!F_{kl} \, F^{kl} \, .$$

Uwzględniając ostatni wynik przykładu (v), p. 8, dostajemy więc następujące kryterium: 2-wektor F w przestrzeni czterowymiarowej jest prosty wtedy, i tylko wtedy, gdy

$${}^*\!F_{kl} \, F^{kl} = 0 \, .$$

Ponieważ $\omega_{ijkl} = \omega_{1234} \epsilon_{ijkl}$, to widzimy ponadto, że kryterium to jest równoważne warunkowi

$$\mathrm{Pf}(F^{ij}) = 0 \, , \quad \text{a także} \quad \det(F^{ij}) = 0$$

(patrz punkt 9, §7).

(iii) Dualność p-wektorów i q-form
W przestrzeni czterowymiarowej dana jest baza (e_1, e_2, e_3, e_4), związana z nią 4-forma $\omega = 4! e^1 \wedge e^2 \wedge e^3 \wedge e^4$ oraz pewien 3-wektor

$$s = e_1 \wedge e_2 \wedge e_3 + 2e_1 \wedge e_2 \wedge e_4 + 5e_1 \wedge e_3 \wedge e_4 + 3e_2 \wedge e_3 \wedge e_4 \,.$$

Szukamy 1-formy dualnej do s oraz przedstawienia 3-wektora s w formie jawnie prostej.

Mamy $\omega(e_1, e_2, e_3, e_4) = \omega_{1234} = 1$, więc na podstawie twierdzenia 13 dostajemy:

$$\begin{aligned} {}^*(e_1 \wedge e_2 \wedge e_3) &= e^4 \,, & {}^*(e_1 \wedge e_2 \wedge e_4) &= -e^3 \,, \\ {}^*(e_1 \wedge e_3 \wedge e_4) &= e^2 \,, & {}^*(e_2 \wedge e_3 \wedge e_4) &= -e^1 \,, \end{aligned}$$

a stąd przez liniowość

$${}^*s = -3e^1 + 5e^2 - 2e^3 + e^4 \,.$$

Rozkład 3-wektora s moglibyśmy uzyskać znaną nam już metodą wyliczenia anihilatora. Tu postąpimy inaczej. Oznaczmy ${}^*s = f^4$, i uzupełnijmy f^4 do bazy (f^1, f^2, f^3, f^4) przestrzeni dualnej, kładąc np. $f^i = e^i$ dla $i = 1, 2, 3$. Znajdujemy teraz (co można zrobić na różne sposoby):

$$f_1 = e_1 + 3e_4 \,, \quad f_2 = e_2 - 5e_4 \,, \quad f_3 = e_3 + 2e_4 \,, \quad f_4 = e_4 \,.$$

Ponieważ $f^1 \wedge f^2 \wedge f^3 \wedge f^4 = e^1 \wedge e^2 \wedge e^3 \wedge e^4$, to $\omega(f_1, f_2, f_3, f_4) = 1$. Stąd

$${}^*(f_1 \wedge f_2 \wedge f_3) = f^4 \,, \quad \text{więc} \quad s = f_1 \wedge f_2 \wedge f_3 \,.$$

(iv) 2-wektory i 2-formy w przestrzeni Minkowskiego
W przestrzeni Minkowskiego w bazie kanonicznej (e_0, e_1, e_2, e_3) współrzędne 2-formy F można zapisać jako macierz

$$(F_{ij}) = \begin{pmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & -B^3 & B^2 \\ -E^2 & B^3 & 0 & -B^1 \\ -E^3 & -B^2 & B^1 & 0 \end{pmatrix} .$$

Podniesienie wskaźników tej formy prowadzi do 2-wektora o współrzędnych $F^{0\alpha} = -F_{0\alpha}$, $F^{\alpha\beta} = F_{\alpha\beta}$ dla $\alpha, \beta = 1, 2, 3$, czyli

$$(F^{ij}) = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix} .$$

Wyróżniona 4-forma dana jest wzorem $e = 4!\, e_0 \wedge e_1 \wedge e_2 \wedge e_3$, więc jej współrzędne w danej bazie są równe $e_{ijkl} = \epsilon_{ijkl}$. Współrzędne 2-formy dualnej wyliczamy więc z przepisu

$$^*F_{ij} = \frac{1}{2} e_{ijkl} F^{kl} = \sum_{k<l} \epsilon_{ijkl} F^{kl} \,.$$

skąd dostajemy

$$(^*F_{ij}) = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & -E^3 & E^2 \\ B^2 & E^3 & 0 & -E^1 \\ B^3 & -E^2 & E^1 & 0 \end{pmatrix} .$$

Formę $^*F$ otrzymuje się z formy F przez zamianę $E^i \to -B^i$, $B^i \to E^i$.

(v) Iloczyn wektorowy

Wybierzmy w trójwymiarowej przestrzeni euklidesowej jedną z orientacji jako dodatnią. Niech (e_1, e_2, e_3) będzie ortonormalną, dodatnio zorientowaną bazą. Iloczyn wektorowy określa się przy użyciu wyróżnionej 3-formy związanej z wybraną orientacją, skąd dostaje się

$$e_1 \times e_2 = e_3 \,, \quad e_2 \times e_3 = e_1 \,, \quad e_3 \times e_1 = e_2$$

– wystarczy zauważyć, że należy spełnić warunki

$$e_i \times e_j \perp e_i, e_j \,, \; i \neq j \,, \quad \text{oraz} \quad (e_1 \times e_2) \cdot e_3 = e_{123} = 1 \,.$$

Stąd, jeśli $x = x^i e_i$, $y = y^i e_i$, to

$$x \times y = (x^2 y^3 - x^3 y^2) e_1 + (x^3 y^1 - x^1 y^3) e_2 + (x^1 y^2 - x^2 y^1) e_3 \,,$$

co symbolicznie możemy przedstawić tak:

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x^1 & x^2 & x^3 \\ y^1 & y^2 & y^3 \end{vmatrix} .$$

Niech $x = -e_1 + 2e_2 - 3e_3$, $z = 3e_1 - 4e_2 + 5e_3$. Szukamy części x równoległej oraz prostopadłej do z. Wyliczamy

$$z \cdot z = 50 \,, \quad x \cdot z = -26 \,, \quad z \times x = 2e_1 + 4e_2 + 2e_3 \,,$$
$$(z \times x) \times z = 28e_1 - 4e_2 - 20e_3 \,,$$

a stąd

$$x_\parallel = \frac{13}{25}(-3e_1 + 4e_2 - 5e_3) \,, \qquad x_\perp = \frac{1}{25}(14e_1 - 2e_2 - 10e_3) \,.$$

Sprawdzamy, że istotnie $x_\parallel + x_\perp = x$.

V

GEOMETRIA PRZESTRZENI AFINICZNYCH

§21 Przestrzenie afiniczne

1 Jednorodność. Przestrzeń afiniczna

Pojęcie przestrzeni wektorowej wciela w ścisły sposób ideę liniowej superpozycji, leżącą u podstaw wielu modeli w naukach przyrodniczych, i nie tylko. Z punktu widzenia niektórych zastosowań ma ono jednak pewną wadę: obecność wyróżnionego elementu – wektora zerowego. W wielu sytuacjach idea liniowości występuje łącznie z ideą ***jednorodności***, przez którą rozumiemy jednakowe położenie każdego elementu (czyli brak wyróżnienia któregokolwiek z nich przez samą definicję przestrzeni). Łączy te dwie idee pojęcie blisko spokrewnione z przestrzenią wektorową.

Niech $\mathcal{M}$ oznacza pewien zbiór, a V – przestrzeń wektorową. Zbiór $\mathcal{M}$ nazywamy ***przestrzenią afiniczną nad przestrzenią wektorową*** V, jeśli określone jest działanie

$$\mathcal{M} \times V \mapsto \mathcal{M}\,, \quad (Y, x) \mapsto Y + x$$

o własnościach:

$$\begin{array}{ll} \text{(i)} & \forall\, Y, Z \in \mathcal{M} \;\; \exists!\, x \in V : \quad Z = Y + x\,, \\ \text{(ii)} & \forall\, Y \in \mathcal{M}\; \forall\, x, z \in V : \quad Y + (x + z) = (Y + x) + z\,. \end{array}$$

Zauważmy, że znak $+$ występuje w definicji w podwójnej roli: jako znak nowego działania określonego tą definicją oraz jako znak dodawania wektorów (w wyrażeniu $(x+z)$). Własność (ii) gwarantuje, że takie użycie tego znaku nie prowadzi do niejednoznaczności.

Jeśli spełnione są warunki definicyjne, to przestrzeń V nazywamy ***przestrzenią wektorową stowarzyszoną z przestrzenią afiniczną*** $\mathcal{M}$. Mówimy też niekiedy, że *para* $(\mathcal{M}, V)$ jest przestrzenią afiniczną nad ciałem $\mathbb{K}$ (tym samym, nad którym zbudowana jest V). Przestrzeń afiniczna $\mathcal{M}$ może być zbiorem pustym – w odróżnieniu od przestrzeni wektorowej, która musi zawierać wektor zerowy. ***Wymiarem niepustej przestrzeni afinicznej*** nazywamy wymiar stowarzyszonej z nią przestrzeni wektorowej, $\dim \mathcal{M} := \dim V$. W całym rozdziale dotyczącym geometrii afinicznej będziemy na ogół milcząco zakładać, że przestrzeń ma skończony wymiar. Pozostawiamy uwadze czytelnika zauważenie, które z wyników obowiązują także w przypadku nieskończonego wymiaru.

2 Punkty i wektory

Elementy przestrzeni $\mathcal{M}$ nazywamy ***punktami***; notować je będziemy za pomocą dużych liter $X, Y, \ldots$ – dla odróżnienia od wektorów, które nadal oznaczać będziemy literami małymi.

Twierdzenie 1. *Jeśli $Y \in \mathcal{M}$, $x \in V$, to*

$$Y + x = Y \quad \Longleftrightarrow \quad x = 0\,.$$

Dowód. Na podstawie punktu (i) definicji dla każdego punktu Y istnieje dokładnie jeden wektor x spełniający równość po lewej stronie równoważności. Ale na mocy punktu (ii) mamy stąd $Y + 0 = (Y + x) + 0 = Y + x = Y$, więc $x = 0$. □

Jeśli Y i Z są dowolnymi punktami, to jedyny wektor x, który spełnia $Z = Y + x$ oznacza się przez $\overrightarrow{YZ}$, a więc

$$Z = Y + \overrightarrow{YZ}\,.$$

Punkt Y nazywamy ***początkiem*** wektora $\overrightarrow{YZ}$, a punkt Z – jego ***końcem***. Powyższe twierdzenie ma teraz równoważne sformułowanie:

$$Z = Y \quad \Longleftrightarrow \quad \overrightarrow{YZ} = 0\,.$$

Łatwo sprawdzić na podstawie aksjomatów, że dla dowolnych punktów X, Y, Z jest

$$-\overrightarrow{XY} = \overrightarrow{YX} \quad \text{oraz} \quad \overrightarrow{XY} + \overrightarrow{YZ} = \overrightarrow{XZ}\,.$$

Powyższe własności sugerują wprowadzenie wygodnego oznaczenia

$$Y - X := \overrightarrow{XY}\,,$$

przy którym uzyskują one postać

$$Z = Y + (Z - Y)\,, \quad Z = Y \Longleftrightarrow Z - Y = 0\,,$$
$$-(Y - X) = X - Y\,, \quad (Y - X) + (Z - Y) = Z - X\,.$$

Wprowadzamy ponadto notację $X - y := X + (-y)$.

Podkreślmy, że znaki ‘minus’ pomiędzy punktami lub pomiędzy punktem i wektorem nie wprowadzają nowych działań, a jedynie inne zapisy wprowadzonych wcześniej obiektów.

3 Przykłady

(i) Przestrzeń wektorowa jako przestrzeń afiniczna
Każdą przestrzeń wektorową V można uważać za przestrzeń afiniczną nad przestrzenią wektorową V. Punktami tej przestrzeni są wektory, a dodawanie wektora do punktu jest identyczne z dodawaniem dwóch wektorów. W tym modelu przestrzeni afinicznej istnieje wyróżniony punkt (pokrywający się z wektorem zerowym), ale wynika to jedynie ze szczególności modelu, a nie z ogólnej definicji. (W analogiczny sposób istnienie w przestrzeni wektorowej $\mathbb{K}^n$ wyróżnionej bazy wynika ze szczególności modelu.)

(ii) Podzbiór liniowy przestrzeni wektorowej
Niech V będzie przestrzenią wektorową, $x \in V$ – jej dowolnym wektorem, a W – dowolną podprzestrzenią. Zbiór wektorów

$$\mathcal{N} := x + W \equiv \{x + y \mid y \in W\}$$

jest przestrzenią afiniczną nad przestrzenią W, z dodawaniem wektorów przestrzeni W do punktów $\mathcal{N}$ określonym przez dodawanie wektorów w V. Jeśli $x \in W$, to $\mathcal{N} = W$. Jeśli jednak $x \notin W$, to $\mathcal{N}$ nie zawiera wektora zerowego przestrzeni V. W tym przypadku definicja przestrzeni $\mathcal{N}$ nie wyróżnia żadnego punktu. (Jeśli tworzymy obraz graficzny zbioru $\mathcal{N}$, to może nam się wydać, że oddalenie punktów od wektora zerowego zadaje wśród nich niesymetryczną relację. Błąd w tym rozumowaniu polega na milczącym przyjęciu – dla określenia odległości – struktury przestrzeni euklidesowej, nieobecnej w ogólnej przestrzeni wektorowej.)

4 Punkt odniesienia, wektory wodzące

Wybierzmy pewien punkt $O \in \mathcal{M}$. Z definicji dodawania wektorów do punktów oraz z własności (i) tego działania wynika, że każdy punkt $X \in \mathcal{M}$ można jednoznacznie reprezentować wektorem $\overrightarrow{OX}$, czyli że odwzorowanie

$$\mathcal{M} \ni X \mapsto \overrightarrow{OX} \in V$$

jest bijekcją. W przypadku takiej reprezentacji punkt O będziemy nazywać ***punktem odniesienia*** (lub ***punktem centralnym***), a wektor $\overrightarrow{OX}$ – ***wektorem wodzącym punktu*** X (względem punktu odniesienia O). Przy zmianie punktu odniesienia wektory wodzące transformują się według reguły

$$\overrightarrow{O'X} = \overrightarrow{OX} - \overrightarrow{OO'} .$$

5 Podprzestrzenie afiniczne

Podzbiór $\mathcal{N}$ przestrzeni afinicznej $\mathcal{M}$ nad przestrzenią wektorową V nazywamy jej ***podprzestrzenią afiniczną***, jeśli istnieje taka podprzestrzeń $W \subseteq V$, że $\mathcal{N}$ jest przestrzenią afiniczną nad W, z działaniem dodawania wektorów do punktów odziedziczonym po parze $(\mathcal{M}, V)$. Jeśli ten warunek jest spełniony, to W nazywamy ***przestrzenią kierunkową*** podprzestrzeni $\mathcal{N}$.

Twierdzenie 2. *Niech $\mathcal{N}$ będzie niepustym podzbiorem przestrzeni afinicznej $\mathcal{M}$ i niech $X \in \mathcal{N}$. Zbiór $\mathcal{N}$ jest podprzestrzenią afiniczną o przestrzeni kierunkowej W wtedy, i tylko wtedy, gdy dla każdego $Y \in \mathcal{M}$ zachodzi równoważność*

$$Y \in \mathcal{N} \quad \iff \quad Y - X \in W .$$

Dowód. Jeśli $\mathcal{N}$ jest niepustą podprzestrzenią afiniczną z przestrzenią kierunkową W, to spełnienie warunku jest oczywiste. Odwrotnie, niech warunek będzie spełniony. Jeśli $Y \in \mathcal{N}$ i $x \in W$, to $(Y + x) - X = (Y - X) + x \in W$, więc $Y + x \in \mathcal{N}$. Dodawanie wektorów z przestrzeni W do punktów zbioru $\mathcal{N}$ jest więc poprawnie określone w $\mathcal{N}$. Dalej, niech $Y, Z \in \mathcal{N}$. Wtedy

$$Z - Y = (Z - X) - (Y - X) \equiv x \in W , \quad Z = Y + x ,$$

więc warunek (i) definicji przestrzeni afinicznej jest spełniony (jednoznaczność wektora x dla danych punktów Y, Z wynika z jego jednoznaczności w przestrzeni V). Punkt (ii) definicji jest w oczywisty sposób dziedziczony od przestrzeni $(\mathcal{M}, V)$. □

Twierdzenie pokazuje, że każda niepusta podprzestrzeń afiniczna o przestrzeni kierunkowej W jest zbiorem o postaci

$$X + W := \{X + y \mid y \in W\} ,$$

gdzie X jest dowolnym punktem tej podprzestrzeni, i taki zapis będziemy najczęściej stosować.

Używa się kilku dodatkowych nazw na oznaczenie różnych typów podprzestrzeni afinicznych. Jednowymiarowe podprzestrzenie afiniczne nazywa się ***prostymi***, a dwuwymiarowe podprzestrzenie – ***płaszczyznami***. Terminu ***płaszczyzna*** używa się też niekiedy szerzej – na oznaczenie dowolnej podprzestrzeni o wymiarze większym niż jeden, a mniejszym niż wymiar całej przestrzeni; przy tej konwencji dokładniejszym określeniem na oznaczenie podprzestrzeni o wymiarze równym p jest ***p-płaszczyzna***. Podprzestrzenie o wymiarze mniejszym o jeden od wymiaru całej przestrzeni nazywa się ***hiperpłaszczyznami***.

6 Kombinacje afiniczne. Powłoki afiniczne

Niech $X_0, \ldots, X_m$ będzie ciągiem punktów przestrzeni $\mathcal{M}$, a $\lambda_0, \ldots, \lambda_m$ ciągiem liczb takim, że $\sum_{i=0}^{m} \lambda_i = 1$. Wybierzmy dowolny punkt odniesienia $O \in \mathcal{M}$ i utwórzmy punkt

$$Z = O + \sum_{i=0}^{m} \lambda_i \, (X_i - O) \, .$$

Przy zastąpieniu w tym wyrażeniu punktu O innym punktem $O' = O + x$ dostaniemy punkt

$$Z' = O + x + \sum_{i=0}^{m} \lambda_i \, (X_i - O - x) = Z + x - \sum_{i=0}^{m} \lambda_i \, x = Z \, ,$$

więc określenie tego punktu nie zależy od wyboru punktu odniesienia O. Oznaczamy tak określony punkt symbolem

$$\sum_{i=0}^{m} \lambda_i X_i \, ,$$

i nazywamy go ***kombinacją afiniczną*** (lub ***barycentryczną***) ***punktów*** $X_0, \ldots, X_m$, a liczby $\lambda_0, \ldots, \lambda_m$ – ***współczynnikami tej kombinacji***. Zauważmy, że symbol ten ma sens tylko jako całość, oddzielne dodawanie punktów i mnożenie ich przez liczbę nie są określone. Pisząc dalej ten symbol będziemy zawsze milcząco zakładać, że liczby λ_i spełniają warunek sensowności symbolu, tj. $\sum_{i=0}^{m} \lambda_i = 1$. Zbiór wszystkich kombinacji afinicznych punktów $X_0, \ldots, X_m$ nazywamy ich ***powłoką afiniczną*** i oznaczamy $A(X_0, \ldots, X_m)$.

Wybierzmy w definicji kombinacji afinicznej za punkt odniesienia $O = X_0$. Dostajemy wtedy związek

$$\sum_{i=0}^{m} \lambda_i X_i = X_0 + \sum_{i=1}^{m} \lambda_i (X_i - X_0) \, .$$

Po prawej stronie tej równości liczby λ_i są dowolne (bo zawsze można dobrać $\lambda_0 = 1 - \sum_{i=1}^{m} \lambda_i$), więc do punktu X_0 dodaje się dowolny wektor powłoki liniowej $L(\overrightarrow{X_0X_1}, \ldots, \overrightarrow{X_0X_m})$. Porównanie z wynikiem poprzedniego punktu daje natychmiast następujący wynik.

Twierdzenie 3. *Powłoka afiniczna ciągu punktów $X_0, \ldots, X_m$ jest podprzestrzenią afiniczną oraz*

$$A(X_0, \ldots, X_m) = X_0 + L(\overrightarrow{X_0X_1}, \ldots, \overrightarrow{X_0X_m}) .$$

Zauważmy, że powłoka afiniczna nie zależy od porządku punktów, więc w powyższej formule rolę punktu X_0 może przejąć dowolny inny w ciągu.

Pojęcie powłoki afinicznej można określić szerzej, w analogii do powłoki liniowej w przestrzeni wektorowej. Jeśli $C \subseteq \mathcal{M}$ jest dowolnym niepustym podzbiorem przestrzeni afinicznej, to ***powłoką afiniczną zbioru*** C nazywamy zbiór wszystkich kombinacji afinicznych dowolnych ciągów punktów tego podzbioru i oznaczamy go $A(C)$. Pokazuje się, że $A(C)$ jest podprzestrzenią afiniczną, oraz jeśli $O \in C$ jest dowolnym punktem w tym podzbiorze, to

$$A(C) = O + L(\{\overrightarrow{OX} \mid X \in C\}) .$$

Łatwo też przekonać się, że $A(C)$ jest najmniejszą podprzestrzenią afiniczną zawierającą zbiór C. Dowody tych faktów zostawiamy jako ćwiczenie dla czytelnika (skorzystać między innymi z tego, że baza powłoki liniowej zbioru B może być wybrana spośród wektorów tego podzbioru).

7 Afiniczna niezależność punktów

Punkty $X_0, \ldots, X_m$ nazywamy ***afinicznie niezależnymi***, jeśli wartość ich kombinacji afinicznej zadaje jednoznacznie współczynniki tej kombinacji, czyli zachodzi implikacja:

$$\sum_{i=0}^{m} \mu_i X_i = \sum_{i=0}^{m} \lambda_i X_i \quad \Longrightarrow \quad \mu_i = \lambda_i \,, \quad i = 0, \ldots, m .$$

Reprezentując kombinacje afiniczne tak, jak to zrobiliśmy przed twierdzeniem 3, możemy poprzednik tej implikacji zapisać równoważnie jako

$$\sum_{i=1}^{m} \mu_i (X_i - X_0) = \sum_{i=1}^{m} \lambda_i (X_i - X_0) \,, \quad \text{czyli} \quad \sum_{i=1}^{m} (\mu_i - \lambda_i)(X_i - X_0) = 0 .$$

Dostajemy więc następującą równoważność.

Twierdzenie 4. *Punkty $X_0, \ldots, X_m$ są afinicznie niezależne wtedy, i tylko wtedy, gdy wektory $\overrightarrow{X_0X_1}, \ldots, \overrightarrow{X_0X_m}$ są liniowo niezależne.*

Afiniczna niezależność punktów nie zależy od ich porządku, więc tutaj także punkt X_0 można zamienić z dowolnym innym z ciągu.

8 Afiniczne układy odniesienia

Jeśli punkty $O_0, \ldots, O_n$ są afinicznie niezależne, a ich powłoka afiniczna jest równa przestrzeni afinicznej $\mathcal{M}$, to ciąg przez nie utworzony nazywamy ***barycentrycznym układem odniesienia przestrzeni*** $\mathcal{M}$. Zwróćmy uwagę, że podprzestrzeń afiniczna jest równa przestrzeni $\mathcal{M}$ wtedy, i tylko wtedy, gdy jej przestrzeń kierunkowa jest równa V. Na podstawie twierdzeń 3 i 4 dostajemy więc natychmiast następujące kryterium.

Twierdzenie 5. *Ciąg punktów $(O_0, \ldots, O_n)$ jest barycentrycznym układem odniesienia przestrzeni $\mathcal{M}$ nad przestrzenią wektorową V wtedy, i tylko wtedy, gdy ciąg wektorów $(\overrightarrow{O_0O_1}, \ldots, \overrightarrow{O_0O_n})$ jest bazą przestrzeni V.*

Niech punkty $O_0, \ldots, O_n$ tworzą barycentryczny układ odniesienia i oznaczmy $O = O_0$, $e_i = \overrightarrow{O_0O_i}$, $i = 1, \ldots, n$. Wtedy mówimy, że para $(O, (e_i))$, złożona z punktu centralnego O i bazy uporządkowanej $(e_1, \ldots, e_n)$, tworzy ***afiniczny układ odniesienia***. Każdy punkt $X \in \mathcal{M}$ ma teraz jednoznaczne przedstawienie

$$X = \sum_{i=0}^{n} X^i O_i = O + \sum_{i=1}^{n} X^i e_i\,, \quad \text{gdzie} \quad X^0 = 1 - \sum_{i=1}^{n} X^i\,.$$

Ciąg liczb $X^1, \ldots, X^n$ nazywamy ***afinicznymi współrzędnymi punktu*** X ***w układzie*** $(O, (e_i))$. Stosujemy dla nich podobne umowy jak dla współrzędnych w przestrzeni wektorowej:

$$X = O + X^i e_i = O' + X'^i e'_i\,,$$

ale gdy mowa jest o punktach, to *nie używamy dla nich alternatywnej nazwy składowych.* Ta różnica pochodzi stąd, że dla punktów można rozważać ogólniejsze układy współrzędnych, które nie mają bezpośredniego pochodzenia wektorowego, tzw. współrzędne krzywoliniowe, których jednak nie omawiamy w tym podręczniku.

Twierdzenie 6. *Niech $(O, (e_i))$ i $(O', (e'_i))$ będą dwoma afinicznymi układami odniesienia. Jeśli*

$$(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)\boldsymbol{\beta}\,, \quad O' = O + a\,, \quad a = a^i e_i\,,$$

to

$$X'^i = {\beta^{-1}}^i{}_j (X^j - a^j)\,, \quad \text{czyli} \quad \mathbf{X}' = \boldsymbol{\beta}^{-1}(\mathbf{X} - \mathbf{a})\,.$$

Dowód. Wstawiając związek punktów odniesienia do równości poprzedzającej twierdzenie mamy $(X^i - a^i)e_i = X'^i e'_i$. Podstawienie transformacji baz prowadzi teraz bezpośrednio do tezy. □

Jeśli do reprezentacji punktu X użyć bardziej symetrycznego przedstawienia w układzie barycentrycznym, $X = \sum_{i=0}^{n} X^i O_i$, to liczby $X^0, \ldots, X^n$ nazywamy ***barycentrycznymi współrzędnymi punktu*** X ***w układzie*** $(O_0, \ldots, O_n)$.

9 Przykłady

(i) Transformacja współrzędnych afinicznych przy zmianie układu
Rozważamy dwa barycentryczne układy odniesienia: układ $(O_0, \ldots, O_n)$ oraz $(O'_0, \ldots, O'_n)$, przy czym $O'_0 = O_k$, $O'_k = O_0$, oraz $O'_i = O_i$ dla $i \neq 0, k$. Wiążemy z nimi afiniczne układy $(O, (e_j))$, $(O', (e'_j))$, gdzie $O = O_0$, $O' = O'_0$, $e_j = \overrightarrow{O_0 O_j}$, $e'_j = \overrightarrow{O'_0 O'_j}$. Pytamy o transformacje współrzędnych afinicznych punktu X przy zmianie afinicznego układu z nieprimowanego na primowany.

Mamy

$$X = \sum_{i=0}^{n} X^i O_i = \sum_{i=0}^{n} X'^i O'_i = X'^0 O_k + X'^k O_0 + \sum_{j \neq 0,k} X'^j O_j \,,$$

więc (z afinicznej niezależności) $X'^0 = X^k$, $X'^k = X^0$, $X'^j = X^j$ dla $j \neq 0, k$. Dla zmiennych afinicznych otrzymujemy zatem transformację:

$$X'^j = X^j \quad \text{dla} \quad j = 1, \ldots, k-1, k+1, \ldots, n, \quad X'^k = 1 - \sum_{i=1}^{n} X^i \,.$$

Przez porównanie ze standardową postacią transformacji podaną w twierdzeniu 6 odczytujemy

$$\boldsymbol{\beta}^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ -1 & \ldots\ldots & -1 & & \ldots\ldots & & -1 \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

oraz $\mathbf{a} = \mathbf{E}_k$ (wektor bazy kanonicznej). Ten sam wynik można uzyskać odczytując macierz $\boldsymbol{\beta}$ i wektor a bezpośrednio z określenia primowanego układu afinicznego. Zauważmy także, że dwukrotne złożenie opisanej transformacji układu barycentrycznego jest w oczywisty sposób identycznością. Dwukrotne zastosowanie odpowiedniej transformacji współrzędnych $\mathbf{X}$ musi więc też być identycznością, skąd dostaje się $\boldsymbol{\beta}^2 = \mathbf{1}$ i $\boldsymbol{\beta}\mathbf{a} = -\mathbf{a}$, co łatwo też sprawdzić bezpośrednim rachunkiem.

(ii) Przestrzeń $\mathbb{K}^n \times \{1\}$
Rozważmy przestrzeń wektorową $\mathbb{K}^{n+1}$ (zapisaną jako przestrzeń kolumn) z bazą kanoniczną $(\mathbf{e}_1, \ldots, \mathbf{e}_{n+1})$. Zbiór $\mathbf{e}_{n+1} + L(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest szczególnym przypadkiem n-wymiarowej przestrzeni afinicznej skonstruowanej w przykładzie (ii) w punkcie 3. Przestrzeń tę można zapisać jako $\mathbb{K}^n \times \{1\}$, a jej stowarzyszoną przestrzeń wektorową – jako $\mathbb{K}^n \times \{0\}$. Punkty i wektory mają odpowiednio postać $\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$ i $\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix}$, gdzie $\mathbf{x} \in \mathbb{K}^n$. Ta przestrzeń afiniczna posiada wyróżniony układ afiniczny utworzony z punktu centralnego $\mathbf{e}_{n+1}$ i bazy $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$.

(iii) Wektorowa kombinacja punktów
Niech $X_0, \ldots, X_m$ będzie ciągiem punktów, a $\mu_0, \ldots, \mu_m$ ciągiem liczb takim, że $\sum_{i=1}^{m} \mu_i = 0$. Wybieramy dowolny punkt O i definiujemy *wektor*

$$\sum_{i=0}^{m} \mu_i X_i := \sum_{i=0}^{m} \mu_i (X_i - O)\,.$$

Łatwo sprawdzić, że wektor ten nie zależy od wyboru punktu O; moglibyśmy nazwać go wektorową kombinacją punktów. Przy pomocy tego pojęcia afiniczną niezależność punktów można określić równoważnie do wcześniejszej definicji za pomocą warunku:

$$\sum_{i=0}^{m} \mu_i = 0\,,\ \sum_{i=0}^{m} \mu_i X_i = 0 \quad \Longrightarrow \quad \mu_i = 0\,,\ i = 0, \ldots, m\,.$$

Ponadto, jeśli $\sum_{i=0}^{m} \lambda_i X_i$ oraz $\sum_{j=0}^{k} \nu_j Y_j$ są afinicznymi kombinacjami punktów, to różnicę punktów $\sum_{i=0}^{m} \lambda_i X_i - \sum_{j=0}^{k} \nu_j Y_j$ można zinterpretować jako wektorową kombinację (w powyższym sensie) punktów $X_0, \ldots, X_m, Y_0, \ldots, Y_k$ ze współczynnikami $\lambda_1, \ldots, \lambda_m, -\nu_1, \ldots, -\nu_k$.

(iv) Krzywe ciągłe i różniczkowalne
Krzywą ciągłą w rzeczywistej lub zespolonej przestrzeni afinicznej określa się podobnie jak w przestrzeni wektorowej. Zakładamy, że dane jest odwzorowanie $\langle a, b\rangle \ni s \mapsto X(s) \in \mathcal{M}$, takie, że w dowolnie wybranym układzie afinicznym współrzędne są ciągłymi funkcjami parametru s. Pokazuje się, że tę samą własność mają wtedy współrzędne w dowolnym innym układzie afinicznym. Mówimy ponadto, że ta krzywa jest klasy $\mathcal{C}^m$, jeśli współrzędne jako funkcje parametru nie tylko są ciągłe, ale też mają ciągłe pochodne, aż do m-tej włącznie. Jeśli krzywa jest klasy $\mathcal{C}^1$ to określa się ***wektor styczny*** do krzywej:

$$\frac{dX(s)}{ds} := \lim_{\Delta\to 0} \frac{X(s+\Delta) - X(s)}{\Delta}\,.$$

Licznik wyrażenia po prawej stronie jest wektorem, granica jest wykonalna, a współrzędne tak określonego wektora są pochodnymi współrzędnych punktu $X(s)$. Dla krzywej klasy $\mathcal{C}^m$ podobnie określa się wektory $d^kX(s)/ds^k$, $k \leq m$. Jeśli $\langle a', b'\rangle \ni \tau \mapsto f(\tau) \in \langle a, b\rangle$ jest bijektywną funkcją klasy $\mathcal{C}^1$ i $Y(\tau) = X\big(f(\tau)\big)$, to

$$\frac{dY(\tau)}{d\tau} = f'(\tau)\frac{dX(s)}{ds}\bigg|_{s=f(\tau)}\,.$$

Ta ostatnia własność może być użyta w szczególnych przypadkach do wyróżnienia pewnej parametryzacji krzywej. Na przykład, jeśli przestrzeń $\mathcal{M}$ jest euklidesowa (tj. jej stowarzyszona przestrzeń V jest euklidesowa, patrz następny paragraf), a wektor styczny nigdzie na krzywej nie znika, to można zażądać, aby $\|dY(\tau)/d\tau\| = 1$.

(v) Czasoprzestrzeń Minkowskiego i krzywe czasopodobne
Afiniczna przestrzeń Minkowskiego (przestrzeń afiniczna nad wektorową przestrzenią Minkowskiego), jak zapowiadaliśmy w punkcie 12, §15, jest modelem matematycznym czasoprzestrzeni w szczególnej teorii względności (***czasoprzestrzeni Minkowskiego***). Fizycy mówią, że punkty X i Y w takiej przestrzeni są rozdzielone czasowo (świetlnie, przestrzennie), gdy wektor $\overrightarrow{XY}$ jest czasowy (odpowiednio: świetlny, przestrzenny). Prosta przechodząca przez dwa punkty rozdzielone czasowo (świetlnie) może opisywać ruch inercjalnej cząstki materialnej (odpowiednio: promienia świetlnego).

Ogólnie, od trajektorii cząstki materialnej (niekoniecznie inercjalnej) wymagamy, aby dla każdego punktu na tej trajektorii punkty z nim na niej sąsiadujące były czasowo z nim rozdzielone. Dochodzi się w ten sposób do następującego pojęcia. Mówimy, że krzywa klasy $\mathcal{C}^1$ jest ***czasopodobna***, gdy wektor do niej styczny jest w każdym jej punkcie wektorem czasowym. Niech będzie dana krzywa czasopodobna $X(\tau)$, a przestrzeń niech ma wybraną orientację czasową. Mówimy, że τ jest ***czasem własnym*** na tej krzywej, gdy $dX(\tau)/d\tau$ jest w każdym punkcie krzywej unormowanym wektorem czasowym skierowanym w przyszłość.

10 Równania podprzestrzeni afinicznych

Każda niepusta podprzestrzeń afiniczna, jak już wiemy, jest zbiorem postaci $P + W$, gdzie P jest dowolnym punktem tej podprzestrzeni, a W – jej przestrzenią kierunkową. Istnieje kilka sposobów zadania warunków na przynależność ogólnego punktu X do tej podprzestrzeni.

Wybierzmy dowolną bazę $(f_1, \ldots, f_k)$ podprzestrzeni W. Podprzestrzeń $P + W$ może być wtedy zadana warunkiem

$$X = P + \sum_{i=1}^{k} t^i f_i\,, \quad \text{czyli} \quad \overrightarrow{PX} = \sum_{i=1}^{k} t^i f_i\,,$$

gdzie $t^1, \ldots, t^k$ są dowolnymi parametrami. Jeśli ponadto wybrać niezależnie punkt odniesienia O i oznaczyć $\overrightarrow{OP} = a$, to warunek przyjmie postać

$$\overrightarrow{OX} = a + \sum_{i=1}^{k} t^i f_i\,.$$

Równanie to zadaje podprzestrzeń afiniczną wymiaru k przy dowolnym wyborze wektora a i liniowo niezależnych wektorów $f_1, \ldots, f_k$. Nazywamy je ***równaniem parametrycznym*** k***-płaszczyzny***.

Na podstawie punktu 5, §18, wiemy, że przestrzeń W można przedstawić jako $W = (W^{\perp^*})^{\perp^*}$, co oznacza, że jest ona złożona z wszystkich wektorów prostopadłych do $W^{\perp^*}$. Uzupełnijmy bazę $(f_1, \ldots, f_k)$ podprzestrzeni W do bazy $(f_1, \ldots, f_n)$ całej przestrzeni V. Wtedy $W^{\perp^*} = L(f^{k+1}, \ldots, f^n)$, gdzie $(f^1, \ldots, f^n)$ jest bazą dualną do (f_i), więc wektor x należy do podprzestrzeni W wtedy, i tylko wtedy, gdy $f^{k+1}(x) = \ldots = f^n(x) = 0$. Równoważną postacią równania podprzestrzeni afinicznej jest więc:

$$f^i(\overrightarrow{PX}) = 0\,, \quad i = k+1, \ldots, n\,,$$

lub, przy ustalonym punkcie odniesienia O:

$$f^i(\overrightarrow{OX}) = a^i\,, \quad i = k+1, \ldots, n\,, \quad \text{gdzie} \quad a^i = f^i(a)\,.$$

Przy dowolnym wyborze $n-k$ liniowo niezależnych form $f^{k+1}, \ldots, f^n$ oraz liczb a^i, równanie to zadaje k-wymiarową podprzestrzeń afiniczną. Nazywamy je ***ogólnym równaniem*** k***-płaszczyzny***.

Wiemy z twierdzenia 8, §20, że wektor x należy do W wtedy, i tylko wtedy, gdy $f_1 \wedge \ldots \wedge f_k \wedge x = 0$, więc warunek przynależności punktu X do podprzestrzeni można zapisać w trzeciej formie

$$f_1 \wedge \ldots \wedge f_k \wedge \overrightarrow{PX} = 0$$

lub za pomocą wektora wodzącego względem punktu odniesienia O:

$$f_1 \wedge \ldots \wedge f_k \wedge \overrightarrow{OX} = f_1 \wedge \ldots \wedge f_k \wedge a\,.$$

Odwrotnie, dla każdego niezerowego k-wektora prostego t równanie

$$t \wedge \overrightarrow{OX} = t \wedge a$$

zadaje k-płaszczyznę.

11 Względne położenie podprzestrzeni afinicznych

Dwoma aspektami względnego ułożenia podprzestrzeni afinicznych są ich przecinanie się oraz wielkość ich wspólnej powłoki afinicznej. Formułujemy w tym punkcie proste kryteria tych cech.

Twierdzenie 7. *Podprzestrzenie afiniczne $P + W$ i $P' + W'$ przecinają się wtedy, i tylko wtedy, gdy $\overrightarrow{PP'} \in W + W'$. Jeśli ten warunek jest spełniony i $Q \in (P + W) \cap (P' + W')$, to*

$$(P + W) \cap (P' + W') = Q + W \cap W'\,.$$

Dowód. Jeśli podprzestrzenie przecinają się i punkt Q leży w części wspólnej, to $Q - P \in W$ i $Q - P' \in W'$, więc $P' - P = (Q - P) - (Q - P') \in W + W'$. Odwrotnie, jeśli $P' - P = x - x'$, $x \in W$ i $x' \in W'$, to punkt $P' + x' = P + x$ należy do obu podprzestrzeni.

Jeśli ten warunek jest spełniony i punkt Q należy do przecięcia, to $Q + W = P + W$ i $Q + W' = P' + W'$, i teza wynika z ciągu równoważności

$$\begin{aligned} X \in (Q + W) \cap (Q + W') &\Longleftrightarrow \overrightarrow{QX} \in W \text{ i } \overrightarrow{QX} \in W' \Longleftrightarrow \\ &\Longleftrightarrow \overrightarrow{QX} \in W \cap W' \Longleftrightarrow X \in Q + W \cap W'. \qquad \square \end{aligned}$$

Jako natychmiastowy wniosek otrzymujemy stwierdzenia dotyczące dwóch szczególnych przypadków przecięcia.

Wniosek 8.
(i) Niech $W \subseteq W'$. Wtedy: jeśli $\overrightarrow{PP'} \in W'$, to $P + W \subseteq P' + W'$, a jeśli $\overrightarrow{PP'} \notin W'$, to te podprzestrzenie nie przecinają się.
(ii) Jeśli $W + W' = V$, to podprzestrzenie $P + W$ i $P' + W'$ przecinają się (i przecięcie dane jest wzorem z ostatniego twierdzenia).

O parze przestrzeni $P+W$ i $P'+W'$ takich, że $W \subseteq W'$ lub $W' \subseteq W$ (jak w punkcie (i) powyższego wniosku) mówimy, że są ***równoległe***.

Jak wspomnieliśmy z końcem punktu 6, powłoka afiniczna zbioru jest najmniejszą podprzestrzenią zawierającą ten zbiór. Następujące twierdzenie wyznacza powłokę sumy mnogościowej dwóch podprzestrzeni.

Twierdzenie 9.
(i) Powłoką afiniczną podprzestrzeni $P+W$ i $P'+W'$ jest podprzestrzeń

$$P + \big(W + W' + L(\overrightarrow{PP'})\big)\,.$$

(ii) Jeśli podprzestrzenie $P+W$ i $P'+W'$ przecinają się, to ich powłoką afiniczną jest podprzestrzeń

$$P + (W + W')\,.$$

Dowód. (i) Punkty P i P' należą do szukanej podprzestrzeni, więc jeśli jej przestrzeń kierunkową oznaczyć przez U, to ma ona postać $P+U = P'+U$. Wynika stąd, że $\overrightarrow{PP'} \in U$, więc $L(\overrightarrow{PP'}) \subseteq U$. Dalej, ponieważ zarówno $P+W$, jak i $P'+W'$ mają się w niej zawierać, to muszą zachodzić inkluzje $W \subseteq U$ i $W' \subseteq U$. Stąd łącznie $W + W' + L(\overrightarrow{P'P}) \subseteq U$, skąd wynika teza. Stwierdzenie (ii) jest oczywiste przy uwzględnieniu twierdzenia 7. □

12 Element objętości w przestrzeni rzeczywistej

Wszyscy znamy z codziennego doświadczenia oraz elementarnej geometrii ideę objętości. W tym punkcie chcemy nadać jej ścisły sens w kontekście przestrzeni afinicznych.

Przez objętość rozumie się pewną miarę zbiorów, czyli funkcję przypisującą zbiorom nieujemne liczby rzeczywiste. Metodami algebraicznymi można to zrobić tylko dla pewnej wąskiej klasy zbiorów. W rzeczywistej przestrzeni n-wymiarowej niech dany będzie dowolny punkt P oraz n wektorów $x_1, \dots, x_n$ (niekoniecznie liniowo niezależnych). ***Równoległościanem*** opartym na tych wektorach, wyprowadzonych z punktu P, nazywamy zbiór punktów

$$\{P + \sum_{i=1}^{n} \lambda_i x_i \mid \lambda_i \in \langle 0, 1\rangle\,,\ i = 1, \dots, n\}\,.$$

Można się przekonać, że zbiór ten tworzą punkty położone między parami równoległych $n-1$-wymiarowych płaszczyzn. Określimy najpierw skierowaną objętość $\omega(P; x_1, \dots, x_n)$ równoległościanu, która może być dowolną liczbą, a objętość zadamy jako wartość bezwzględną tej liczby. Żądamy, aby skierowana objętość równoległościanu nie zależała od punktu jego zaczepienia, czyli od P: $\omega(x_1, \dots, x_n)$ (inaczej mówiąc: ma pozostawać niezmieniona pod wpływem translacji, które omawiamy w §23 poniżej). Po drugie, zgodnie z duchem algebry liniowej żądamy, aby objętość skierowana zależała liniowo od każdego z wektorów. Oznacza to, że ω jest tensorem o walencji $\left[\begin{smallmatrix}0\\n\end{smallmatrix}\right]$. Po trzecie, żądamy zerowania się objętości skierowanej, jeśli wektory $x_1, \dots, x_n$ nie są liniowo niezależne, a więc gdy równoległościan jest zawarty w niżej wymiarowej podprzestrzeni afinicznej. Z tego ostatniego warunku wynika, w szczególności, że

$$\begin{aligned}0 = \omega(x_1 + x_2, x_1 + x_2, x_3, \dots, x_n) = \omega(x_1, x_1, x_3, \dots, x_n) + \\ + \omega(x_1, x_2, x_3, \dots, x_n) + \omega(x_2, x_1, x_3, \dots, x_n) + \omega(x_2, x_2, x_3, \dots, x_n) = \\ = \omega(x_1, x_2, x_3, \dots, x_n) + \omega(x_2, x_1, x_3, \dots, x_n)\end{aligned}$$

i podobnie dla każdej pary argumentów. Tensor ω jest więc całkowicie antysymetryczny, czyli jest n-formą. Odwrotnie, z antysymetrii wynika zerowanie się objętości skierowanej dla wektorów liniowo zależnych.

Motywowani powyższą dyskusją wprowadzamy następujące określenie. ***Elementem objętości*** nazywamy wybraną n-formę ω. Przy tym wyborze, ***skierowaną objętością*** równoległościanu opartego na wektorach $x_1, \dots, x_n$ nazywamy liczbę $\omega(x_1, \dots, x_n)$, a jego ***objętością*** – jej wartość bezwzględną

$$\mathrm{vol}\,(x_1, \dots, x_n) := |\omega(x_1, \dots, x_n)|\,.$$

Wykazuje się metodami analizy matematycznej, że ta definicja objętości rozszerza się niesprzecznie i jednoznacznie do addytywnej miary zbiorów otrzymanych z klasy równoległościanów w wyniku tworzenia przeliczalnych sum i przecięć mnogościowych oraz dopełnień.

W ogólnych przestrzeniach wektorowych nie ma wyróżnionej n-formy, ale wszystkie różnią się między sobą tylko czynnikiem. Funkcja objętości zgodna z założonymi postulatami jest więc określona z dokładnością do dodatniego czynnika.

Objętość skierowana równoległościanu wchodzi w następującą relację z n-wektorami:

$$x_1 \wedge \dots \wedge x_n = \frac{1}{n!}\,\omega(x_1, \dots, x_n)\,\hat{\omega}\,,$$

gdzie $\hat{\omega}$ jest jedynym n-wektorem związanym z n-formą ω warunkiem $\omega_{i_1 \dots i_n}\hat{\omega}^{i_1 \dots i_n} = n!$ (p. 9, §20). Jest to widoczne natychmiast po wyliczeniu 0-form dualnych do obu stron równości. Ta relacja będzie przydatna w następnym punkcie. Ponadto, jeśli ${x_k}^i$ są współrzędnymi k-tego wektora w bazie, w której $\omega_{1 \dots n} = 1$, to

$$\omega(x_1, \dots, x_n) = \det({x_i}^k)\,.$$

13 p-wymiarowe objętości w podprzestrzeniach p-wymiarowych

Każda podprzestrzeń afiniczna jest sama przestrzenią afiniczną i można w niej określić objętość tak, jak to omówiliśmy w poprzednim punkcie. Wygodnie jest jednak podać charakterystykę takiej p-wymiarowej objętości z punktu widzenia całej przestrzeni.

Wiemy z dyskusji przeprowadzonej w punkcie 7, §20, że każdej p-wymiarowej podprzestrzeni $W \subseteq V$ odpowiada jednoznacznie klasa różniących się co najwyżej czynnikami p-wektorów prostych, utworzonych z wektorów bazowych tej podprzestrzeni. Dla danej W wybierzmy jeden p-wektor $\hat{\omega}_W$ z tej klasy. Niech $x_1, \ldots, x_p \in W$ będzie dowolnym ciągiem wektorów tej podprzestrzeni. Wtedy ***skierowana p-wymiarowa objętość*** $\omega_W(x_1, \ldots, x_p)$ równoległościanu opartego na tych wektorach jest wyznaczona przez relację

$$x_1 \wedge \ldots \wedge x_p = \frac{1}{p!}\,\omega_W(x_1, \ldots, x_p)\,\hat{\omega}_W\,,$$

która uogólnia analogiczny związek dla całej przestrzeni V, otrzymany w poprzednim punkcie. Podobnie, jeśli ${x_k}^i$ są współrzędnymi k-tego wektora w bazie podprzestrzeni W, w której $\omega_{W1\ldots p} = 1$, to

$$\omega_W(x_1, \ldots, x_p) = \det({x_k}^i)\,.$$

Zaznaczmy, że w ogólnych przestrzeniach p-wektor $\hat{\omega}_W$ jest wybierany dla każdej p-wymiarowej podprzestrzeni niezależnie – nie ma metody uzgadniania wyborów dla różnych podprzestrzeni.

14 Przykłady

(i) Od ogólnego do parametrycznego równania podprzestrzeni
W przestrzeni czterowymiarowej dany jest układ afiniczny $(O, (e_i))$ oraz liniowo niezależne formy $\varphi^1 = e^1 + 2e^2 - e^3 + 3e^4$, $\varphi^2 = 2e^1 + 3e^2 - 7e^3$. Szukamy parametrycznego równania 2-płaszczyzny danej ogólnym równaniem

$$\varphi^1(\overrightarrow{OX}) = 5\,, \quad \varphi^2(\overrightarrow{OX}) = -2\,.$$

Zadanie sprowadza się do rozwiązania układu równań liniowych będącego zapisem powyższych równań we współrzędnych afinicznych:

$$X^1 + 2X^2 - X^3 + 3X^4 = 5\,, \quad 2X^1 + 3X^2 - 7X^3 = -2\,.$$

Rozwiązanie ma postać

$$\mathbf{X} = \begin{pmatrix} -19 \\ 12 \\ 0 \\ 0 \end{pmatrix} + X^3 \begin{pmatrix} 11 \\ -5 \\ 1 \\ 0 \end{pmatrix} + X^4 \begin{pmatrix} 9 \\ -6 \\ 0 \\ 1 \end{pmatrix}\,.$$

Jeśli oznaczyć

$$a = -19e_1 + 12e_2\,, \quad f_1 = 11e_1 - 5e_2 + e_3\,, \quad f_2 = 9e_1 - 6e_2 + e_4\,,$$

to równanie parametryczne rozważanej 2-płaszczyzny ma postać

$$\overrightarrow{OX} = a + t^1 f_1 + t^2 f_2\,.$$

(ii) Od parametrycznego do ogólnego równania podprzestrzeni
W przestrzeni czterowymiarowej dane jest równanie parametryczne prostej

$$\overrightarrow{OX} = a + tf\,, \quad \text{gdzie} \quad a = e_1 + 2e_2 + e_4\,, \quad f = 2e_1 - 3e_2 + e_3 + e_4\,.$$

Szukamy ogólnego równania tej prostej.

Zadanie polega na znalezieniu bazy podprzestrzeni $L(f)^{\perp^*}$. Ta podprzestrzeń złożona jest z form φ spełniających w danej bazie warunek:

$$2\varphi_1 - 3\varphi_2 + \varphi_3 + \varphi_4 = 0\,.$$

Znajdujemy trzy liniowo niezależne formy będące jego rozwiązaniem:

$$\varphi^1 = e^1 - 2e^4\,, \quad \varphi^2 = e^2 + 3e^4\,, \quad \varphi^3 = e^3 - e^4\,.$$

Wyliczamy: $\varphi^1(a) = -1$, $\varphi^2(a) = 5$, $\varphi^3(a) = -1$, i dostajemy ogólne równanie danej prostej:

$$\varphi^1(\overrightarrow{OX}) = -1\,, \quad \varphi^2(\overrightarrow{OX}) = 5\,, \quad \varphi^3(\overrightarrow{OX}) = -1\,.$$

Zapisane przy użyciu współrzędnych afinicznych równanie to ma postać

$$X^1 - 2X^4 = -1\,, \quad X^2 + 3X^4 = 5\,, \quad X^3 - X^4 = -1\,.$$

(iii) Przecięcie i powłoka afiniczna podprzestrzeni
W przestrzeni czterowymiarowej dane są równania parametryczne dwóch płaszczyzn:

$$P + W: \qquad \overrightarrow{OX} = a + t^1 f_1 + t^2 f_2\,,$$
$$\text{gdzie} \quad \overrightarrow{OP} = a = e_2 + e_3\,, \quad f_1 = 3e_1 + e_2\,, \quad f_2 = 2e_2 + 5e_4\,,$$

oraz

$$P' + W': \qquad \overrightarrow{OX'} = a' + t'^1 f'_1 + t'^2 f'_2\,,$$
$$\text{gdzie} \quad \overrightarrow{OP'} = a' = -3e_1 + 5e_3 - 4e_4\,, \quad f'_1 = e_3 - e_4\,, \quad f'_2 = -6e_1 + 4e_3 + e_4\,.$$

Szukamy przecięcia i powłoki tych płaszczyzn.

Przecięcie jest zbiorem punktów spełniających równania obu płaszczyzn jednocześnie. Równość prawych stron tych równań jest równaniem liniowym na parametry:

$$t^1 f_1 + t^2 f_2 - t'^1 f'_1 - t'^2 f'_2 = a' - a\,.$$

Badając rzędy macierzy głównej i dołączonej tego układu (zapisanego w bazie (e_i)) stwierdzamy, że wektory f_1, f_2, f'_1 są liniowo niezależne, a wektory f'_2 i $a'-a$ są ich kombinacjami liniowymi. Układ posiada więc jednoparametrowe ogólne rozwiązanie, z wolnym parametrem t'^2, które znajdujemy w standardowy sposób. W istocie wystarczy wyliczyć $t'^1 = -4 - 4t'^2$, aby otrzymać równanie przecięcia

$$\overrightarrow{OX} = a' + (-4 - 4t'^2) f'_1 + t'^2 f'_2 = (a' - 4f'_1) + t'^2 (f'_2 - 4f'_1)\,,$$

czyli równanie prostej

$$\overrightarrow{OX} = b + sc\,, \quad s \in \mathbb{R}\,, \quad \text{gdzie} \quad b = -3e_1 + e_3\,, \quad c = -6e_1 + 5e_4\,.$$

Uwzględniając wynik (ii) twierdzenia 9 otrzymujemy również równanie trójwymiarowej powłoki afinicznej obu płaszczyzn:

$$\overrightarrow{OX} = a + t^1 f_1 + t^2 f_2 + t^3 f'_1\,.$$

(iv) Przecięcie i powłoka afiniczna podprzestrzeni – cd.
Posługując się metodą przykładu (ii) znajdujemy równania ogólne płaszczyzn rozważonych w poprzednim przykładzie:

$$\varphi^1(\overrightarrow{OX}) = 1\,, \quad \varphi^2(\overrightarrow{OX}) = -15\,, \quad \text{gdzie} \quad \varphi^1 = e^3\,, \quad \varphi^2 = 5e^1 - 15e^2 + 6e^4\,,$$

oraz

$$\varphi'^1(\overrightarrow{OX'}) = 0\,, \quad \varphi'^2(\overrightarrow{OX'}) = 21\,, \quad \text{gdzie} \quad \varphi'^1 = e^2\,, \quad \varphi'^2 = 5e^1 + 6e^3 + 6e^4\,.$$

Rozważymy zadanie rozwiązane już w poprzednim przykładzie, ale teraz przy użyciu powyższych równań ogólnych. Znalezienie przecięcia podprzestrzeni sprowadza się do położenia $\overrightarrow{OX'} = \overrightarrow{OX}$ i rozwiązania układu złożonego z wszystkich czterech równań; zostawiamy szczegóły dla czytelnika.

Ponieważ przestrzenie przecinają się, to podprzestrzeń kierunkowa sumy jest sumą ich podprzestrzeni kierunkowych, $W + W'$. Tak więc równanie ogólne sumy zadane jest formami rozpinającymi podprzestrzeń form $(W + W')^{\perp *}$. Zgodnie z tożsamością omówioną w przykładzie (i), p. 7, §18, ta podprzestrzeń jest równa $W^{\perp *} \cap W'^{\perp *} = L(\varphi^1, \varphi^2) \cap L(\varphi'^1, \varphi'^2)$. Rozwiązując równanie

$$s_1 \varphi^1 + s_2 \varphi^2 = s'_1 \varphi'^1 + s'_2 \varphi'^2$$

znajdujemy formę

$$\psi = 6\varphi^1 + \varphi^2 = -15\varphi'^1 + \varphi'^2 = 5e^1 - 15e^2 + 6e^3 + 6e^4 \,,$$

która rozpina to przecięcie. Ponieważ każda z podprzestrzeni afinicznych zawiera się w powłoce, więc wartość $\psi(\overrightarrow{OX})$ można wyliczyć na jednej z nich. Otrzymujemy równanie ogólne powłoki afinicznej:

$$\psi(\overrightarrow{OX}) = -9\,.$$

(v) Objętość jako miara

Wybierzmy n-formę ω zadającą objętość w przestrzeni $\mathcal{M}$. Niech $(e_1, \ldots, e_n)$ będzie dowolną z baz, w których $\omega_{1\ldots n} = 1$. Wtedy objętość vol zadana tą formą jest miarą Lebesgue'a we współrzędnych afinicznych w układzie $(O, (e_i))$, gdzie O jest dowolnym punktem odniesienia. Dokładniej: wybierzmy zbiór $B \subseteq \mathcal{M}$ i niech $\mathbf{B} \subseteq \mathbb{R}^n$ oznacza zakres współrzędnych taki, że $X \in B \iff \mathbf{X} \in \mathbf{B}$. Wtedy

$$\mathrm{vol}\,(B) = \int_{\mathbf{B}} dX^1 \ldots dX^n\,.$$

(vi) Sympleks

Niech w n-wymiarowej przestrzeni $\mathcal{M}$ dane będzie $n+1$ afinicznie niezależnych punktów $X_0, \ldots, X_n$. ***Sympleksem*** o wierzchołkach w tych punktach, oznaczanym $\langle X_0, \ldots, X_n \rangle$, nazywamy zbiór wszystkich kombinacji afinicznych $\sum\limits_{i=0}^{n} \lambda_i X_i$, gdzie liczby λ_i oprócz warunku poprawności symbolu $\sum\limits_{i=0}^{n} \lambda_i = 1$ spełniają warunki $\lambda_i \in \langle 0, 1 \rangle$. Sympleks jednowymiarowy nazywa się ***odcinkiem***, dwuwymiarowy – ***trójkątem***, a trójwymiarowy – ***tetraedrem***. Sympleks jest wielościanem, którego krawędziami są wszystkie odcinki $\langle X_i X_j \rangle$.

Korzystając z poprzedniego punktu dostaje się

$$\mathrm{vol}\,(\langle X_0, \ldots, X_n \rangle) = \big|\omega(\overrightarrow{X_0X_1}, \ldots, \overrightarrow{X_0X_n})\big| \int_{\substack{\lambda_i \geq 0 \\ \sum_{i=1}^n \lambda_i \leq 1}} d\lambda_1 \ldots d\lambda_n =$$
$$= \frac{1}{n!}\big|\omega(\overrightarrow{X_0X_1}, \ldots, \overrightarrow{X_0X_n})\big|\,.$$

Każdy sympleks $\langle X_0, \ldots, X_n \rangle$ może być zanurzony w więcej wymiarowej przestrzeni, i wtedy jego n-objętość jest wyznaczona w podprzestrzeni n-wymiarowej, w której się zawiera (zgodnie z punktem 13).

(vii) Objętość 3-wymiarowa
Wybieramy 3-wektor $\hat{\omega}_W = 3!\, f_1 \wedge f_2 \wedge f_3$ zadający 3-wymiarową objętość w podprzestrzeni W rozpiętej wektorami

$$f_1 = 3e_1 + e_2 + 4e_3 + e_4\,,\ \ f_2 = 4e_1 - 3e_2 + e_3 + e_4\,,\ \ f_3 = e_1 + e_2 + e_3 - e_4\,.$$

Chcemy sprawdzić czy równoległobok oparty na wektorach

$$y_1 = 8e_1 - e_2 + 6e_3 + e_4\,,\ y_2 = -e_1 + e_2 - 2e_3 - 3e_4\,,\ y_3 = -7e_1 + 7e_2 - e_3 - 3e_4$$

leży w podprzestrzeni afinicznej o przestrzeni kierunkowej W, i jeśli tak, to znaleźć jego 3-objętość.

Tworzymy macierz z kolumn współrzędnych wektorów $(f_1, f_2, f_3, y_1, y_2, y_3)$. Stosując metodę Gaussa znajdowania liniowych zależności między kolumnami otrzymujemy

$$\begin{pmatrix} 3 & 4 & 1 & 8 & -1 & -7 \\ 1 & -3 & 1 & -1 & 1 & 7 \\ 4 & 1 & 1 & 6 & -2 & -1 \\ 1 & 1 & -1 & 1 & -3 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 & 2 & 1 \end{pmatrix}\,,$$

skąd odczytujemy

$$\begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \boldsymbol{\beta}\,, \quad \boldsymbol{\beta} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Zatem 3-objętość równoległoboku wynosi $|\det \boldsymbol{\beta}\,| = 7$.

15 Pola tensorowe

Do tej pory rozważaliśmy w tym paragrafie geometrię afiniczną jako strukturę samej przestrzeni afinicznej. Z punktu widzenia potrzeb wielu zastosowań tak rozumiana geometria dostarcza dopiero „areny” dla pojawienia się dalszych obiektów. Wśród takich obiektów zasadnicze znaczenie mają pola tensorowe.

Polem tensorowym (o walencji $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$***)*** na przestrzeni afinicznej $(\mathcal{M}, V)$ nazywamy każde odwzorowanie

$$\mathcal{M} \ni X \mapsto t(X) \in \mathcal{T}^p_q(V)\,.$$

Współrzędne pola tensorowego w dowolnej bazie $(e_1, \ldots, e_n)$ przestrzeni V są funkcjami

$$\mathcal{M} \ni X \mapsto t^{i_1 \ldots\, i_p}{}_{j_1 \ldots\, j_q}(X) \in \mathbb{K}\,.$$

Jeśli wybrać ponadto w przestrzeni $\mathcal{M}$ punkt odniesienia O i oznaczyć kolumnę współrzędnych punktu X w układzie odniesienia $(O,(e_i))$ przez $\mathbf{X}$, to współrzędne pola tensorowego można traktować jako funkcje:

$$\mathbb{K}^n \ni \mathbf{X} \mapsto t^{i_1\dots i_p}{}_{j_1\dots j_q}(\mathbf{X}) \in \mathbb{K}\,.$$

Przyjęliśmy tu nie całkiem poprawną, ale często używaną konwencję, aby współrzędne pola tensorowego jako funkcje punktu przestrzeni $\mathcal{M}$ i te same współrzędne jako funkcje współrzędnych tego punktu oznaczać tym samym symbolem funkcyjnym:

$$t^{i_1\dots i_p}{}_{j_1\dots j_q}(X) = t^{i_1\dots i_p}{}_{j_1\dots j_q}(\mathbf{X})\,,$$

odróżniając je tylko po sposobie notowania argumentu. Umawiamy się przy tym, że dla wybranego układu afinicznego $(O,(e_i))$ zarówno współrzędne punktu X odnosimy do tego układu, jak i współrzędne tensora do bazy (e_i). Dla współrzędnych w nowym układzie mamy więc

$$t'^{i_1\dots i_p}{}_{j_1\dots j_q}(X) = t'^{i_1\dots i_p}{}_{j_1\dots j_q}(\mathbf{X}')\,.$$

Niech układ primowany $(O',(e'_i))$ będzie związany z nieprimowanym takimi relacjami, jak w założeniach twierdzenia 6. Wtedy

$$t'^{i_1\dots i_p}{}_{j_1\dots j_q}(X) = \beta^{-1 i_1}{}_{k_1}\dots\beta^{-1 i_p}{}_{k_p}\beta^{l_1}{}_{j_1}\dots\,\beta^{l_q}{}_{j_q}t^{k_1\dots k_p}{}_{l_1\dots l_q}(X)\,.$$

Wyrażamy teraz współrzędne tensorowe jako funkcje współrzędnych punktu X. Korzystając z przyjętej wyżej konwencji oraz ze związku współrzędnych punktu: $\mathbf{X} = \boldsymbol{\beta}\mathbf{X}' + \mathbf{a}$, dostajemy relację

$$t'^{i_1\dots i_p}{}_{j_1\dots j_q}(\mathbf{X}') = \beta^{-1 i_1}{}_{k_1}\dots\beta^{-1 i_p}{}_{k_p}\beta^{l_1}{}_{j_1}\dots\,\beta^{l_q}{}_{j_q}t^{k_1\dots k_p}{}_{l_1\dots l_q}(\boldsymbol{\beta}\mathbf{X}' + \mathbf{a})\,.$$

16 Pochodna pola tensorowego

Ograniczamy się dalej do przestrzeni rzeczywistych. Mówimy, że pole tensorowe $t(.)$ jest ***klasy*** $\mathcal{C}^m$, $m \geq 0$, jeśli w dowolnym układzie odniesienia istnieją, i są ciągłe, wszystkie cząstkowe pochodne

$$\frac{\partial^m}{\partial X^{k_1}\dots\partial X^{k_m}}\; t^{i_1\dots i_p}{}_{j_1\dots j_q}(\mathbf{X})\,.$$

Warunek istnienia i ciągłości pochodnych wystarczy sprawdzić w jednym, dowolnie wybranym układzie odniesienia: otrzymany w poprzednim punkcie sposób transformowania się współrzędnych pola pociąga wtedy spełnienie warunku w każdym innym układzie. Warunek pociąga też istnienie i ciągłość wszystkich pochodnych niższego stopnia, więc klasa $\mathcal{C}^m$ jest zawarta w klasie $\mathcal{C}^l$ dla $l \leq m$.

Jeśli wartości pola tensorowego klasy $\mathcal{C}^m$ na $\mathcal{M}$ są tensorami o walencji $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$, to oznaczamy dokładniej tę klasę przez $\mathcal{C}^m\big(\mathcal{M}, \mathcal{T}^p_q(V)\big)$.

Przyjmujemy w pozostałej części tego punktu model tensorów jako wieloliniowych funkcji. Wybierzmy dowolny punkt X przestrzeni $\mathcal{M}$ i poprowadźmy z niego prostą $X + L(y)$ w kierunku określonym dowolnym wektorem y. Chcemy zmierzyć prędkość zmiany pola tensorowego przy przemieszczaniu się z punktu X wzdłuż tej prostej. Dla funkcji prędkość zmian mierzy się, jak wiadomo, pochodną, a przemieszczanie się wzdłuż zadanej wyżej prostej w okolicy punktu X oznacza poruszanie się po punktach $X+sy$, dla parametru s w otoczeniu zera. Motywowani powyższą dyskusją wprowadzamy następującą definicję.

Pochodną kierunkową pola tensorowego $t(.) \in \mathcal{C}^{m+1}\big(\mathcal{M}, \mathcal{T}^p_q(V)\big)$ ***w kierunku wektora*** y nazywamy odwzorowanie

$$\underset{y}{\partial} : \mathcal{C}^{m+1}\big(\mathcal{M}, \mathcal{T}^p_q(V)\big) \mapsto \mathcal{C}^m\big(\mathcal{M}, \mathcal{T}^p_q(V)\big)$$

zadane dla każdego $m \geq 0$ przepisem

$$\underset{y}{\partial} t\,(\varphi^1, \ldots, \varphi^p, x_1, \ldots, x_q)(X) = \left\{\frac{d}{ds} t(\varphi^1, \ldots, \varphi^p, x_1, \ldots, x_q)(X + sy)\right\}_{s=0} .$$

Musimy pokazać, że ta definicja jest poprawna, tj. że różniczkowanie po s jest wykonalne i działając na pole klasy $\mathcal{C}^{m+1}$ daje pole klasy $\mathcal{C}^m$. Wykażemy to za chwilę, a na razie załóżmy, że tak jest istotnie, i zbadajmy zależność otrzymanego pola od wektora y. Ze znanych reguł różniczkowania dostajemy

$$\frac{d}{ds} t(\ldots)(X + s\,\alpha y) = \alpha \left\{\frac{d}{dr} t(\ldots)(X + ry)\right\}_{r=\alpha s} ,$$

$$\begin{aligned} &\frac{d}{ds} t(\ldots)(X + s(y+z)) = \\ &\qquad = \left\{\frac{d}{ds} t(\ldots)(X + sy + rz)\right\}_{r=s} + \left\{\frac{d}{ds} t(\ldots)(X + ry + sz)\right\}_{r=s} . \end{aligned}$$

Podstawiając w tych związkach $s = 0$ dostajemy

$$\underset{\alpha y}{\partial} t(\ldots)(X) = \alpha\, \underset{y}{\partial} t(\ldots)(X) , \qquad \underset{y+z}{\partial} t(\ldots)(X) = \underset{y}{\partial} t(\ldots)(X) + \underset{z}{\partial} t(\ldots)(X) ,$$

czyli stwierdzenie liniowości wyrażenia $\underset{y}{\partial} t(\ldots)(X)$ w argumencie y. Dla każdego punktu X i pola tensorowego z klasy $\mathcal{C}^{m+1}\big(\mathcal{M}, \mathcal{T}^p_q(V)\big)$ wyrażenie $\partial t(X)$ jest więc tensorem o walencji $\left[\begin{smallmatrix} p \\ q+1 \end{smallmatrix}\right]$; jego wartość dla argumentów $\varphi^1, \ldots, \varphi^p$, $x_1, \ldots, x_{q+1}$ wynosi

$$\underset{x_1}{\partial} t\,(\varphi^1, \ldots, \varphi^p, x_2, \ldots, x_{q+1})(X) .$$

Otrzymaliśmy więc w wyniku odwzorowanie

$$\partial : \mathcal{C}^{m+1}\big(\mathcal{M}, \mathcal{T}^p_q(V)\big) \mapsto \mathcal{C}^m\big(\mathcal{M}, \mathcal{T}^p_{q+1}(V)\big)\,.$$

Wyliczymy współrzędne pola tensorowego $\partial t(.)$ w dowolnym układzie odniesienia, przy okazji otrzymując brakujące uzasadnienie poprawności definicji pochodnej kierunkowej. W tym celu za argumenty podstawiamy wektory bazowe. Dostajemy ciąg równości

$$\begin{aligned}(\partial t)^{i_1\dots i_p}{}_{j_1\dots j_{q+1}}(X) &= \underset{e_{j_1}}{\partial}\, t(e^{i_1},\dots,e^{i_p},e_{j_2},\dots,e_{j_{q+1}})(X) = \\ &= \left\{\frac{d}{ds}t^{i_1\dots i_p}{}_{j_2\dots j_{q+1}}(X+se_{j_1})\right\}_{s=0} = \frac{\partial}{\partial X^{j_1}}t^{i_1\dots i_p}{}_{j_2\dots j_{q+1}}(X)\,,\end{aligned}$$

gdzie wykorzystaliśmy fakt, że

$$(X+se_{j_1})^k = X^k \quad \text{dla } k\neq j_1, \quad (X+se_{j_1})^{j_1} = X^{j_1}+s\,.$$

Zapisujemy ostateczny wynik:

$$(\partial t)^{i_1\dots i_p}{}_{j_1\dots j_{q+1}}(X) = \frac{\partial}{\partial X^{j_1}}t^{i_1\dots i_p}{}_{j_2\dots j_{q+1}}(X) \equiv \partial_{j_1}t^{i_1\dots i_p}{}_{j_2\dots j_{q+1}}(X)\,,$$

gdzie druga równość wprowadza wygodną notację. Otrzymane pole tensorowe nazywamy ***pochodną pola tensorowego*** $t(.)$. Z wykazanej reguły tworzenia jego współrzędnych wynika poprawność definicji. Licząc pochodną pola w dowolnym układzie współrzędnych i korzystając z reguł różniczkowania dostajemy własności pochodnej tensorowej:

$$\partial(\alpha t+\beta s) = \alpha\partial t+\beta\partial s\,, \qquad \partial(t\otimes u) = \partial t\otimes u + t\otimes\partial u$$

dla dowolnych pól tensorowych (różniczkowalnych) t, s, u i stałych α, β.

Jeśli $t(.)$ jest polem skalarnym (wartości są liczbami), to $\partial t(.)$ jest polem 1-form, nazywanym ***gradientem*** tego pola, o współrzędnych $\partial_i t$. Jeśli $t(.)$ jest polem wektorowym (wartości są wektorami kontrawariantnymi), to $\partial t(.)$ jest polem tensorów o walencji $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, o współrzędnych $\partial_i t^j(X)$. Zwężenie wskaźników daje pole skalarne $\partial_i t^i(X)$, które nazywamy ***dywergencją*** pola wektorowego. Jeśli $t(.)$ jest polem p-form, to pole $(p{+}1)$-form o współrzędnych $\partial_{[i_1}t_{i_2\dots i_{p+1}]}(X)$ nazywamy ***pochodną zewnętrzną*** pola $t(.)$ i oznaczamy standardowo $dt(X)$ (choć symbol $\partial\wedge t(X)$ byłby bliższą analogią do zapisu rotacji - patrz niżej).

Jeśli $\mathcal{M}$ jest przestrzenią euklidesową, to dywergencję pola wektorowego oznacza się często $\partial\cdot t\,(X)$. W szczególnym przypadku trójwymiarowej przestrzeni euklidesowej definiujemy przy użyciu pochodnej zewnętrznej jeszcze jedną operację różniczkową. Mając dane pole wektorowe $t^i(X)$, obniżamy jego wskaźnik i stosujemy pochodną zewnętrzną dostając pole 2-form $\partial_{[i}t_{j]}(X)$, a następnie stosujemy operację dualności dostając pole wektorowe $e^{ijk}\partial_j t_k(X)$. Tak otrzymane pole nazywa się ***rotacją*** pola wektorowego i z oczywistych względów oznacza się je symbolem $\partial \times t\,(X)$.

Stosuje się często inne oznaczenie pochodnej tensorowej: $\nabla t \equiv \partial t$ i odpowiednio też dla pochodnej kierunkowej $\underset{x}{\nabla} t \equiv \underset{x}{\partial} t$.

17 Przykłady

(i) Ruch w polu elektromagnetycznym

Fizyczne pole elektromagnetyczne jest opisywane, w ramach szczególnej teorii względności, antysymetrycznym polem $F(X)$ o walencji $\left[\begin{smallmatrix}0\\2\end{smallmatrix}\right]$ w przestrzeni afinicznej Minkowskiego (przykł. (v), p. 9). Jeśli wybrać bazę ortonormalną (e_0, e_1, e_2, e_3) i bazę do niej dualną, to składowe $F_{ij}(X)$ są postrzegane przez eksperymentatora z punktu widzenia geometrii przestrzeni $L(e_1, e_2, e_3)$ jako składowe E^i i B^i, $i = 1, 2, 3$, pól elektrycznego i magnetycznego, przy czym związek tych współrzędnych jest taki, jak to wynika z notacji użytej w przykładzie (iv), p. 16, §20. Przestrzeń Minkowskiego jest wyposażona w metrykę, więc wskaźniki pola F można podnosić.

Ruch w polu elektromagnetycznym cząstki o masie m i ładunku e określony jest równaniem na jego trajektorię $X(\tau)$:

$$m\frac{d^2X^i(\tau)}{d\tau^2} = eF^i{}_j\big(X(\tau)\big)\frac{dX^j(\tau)}{d\tau}\,,$$

z warunkiem dodatkowym, że τ jest czasem własnym na trajektorii (przykład (v) w punkcie 9). Pokazuje się, że ten warunek jest zgodny z równaniem, tj. jeśli $X(\tau)$ jest rozwiązaniem równania, przy czym $dX(\tau)/d\tau$ jest unormowanym wektorem czasowym skierowanym w przyszłość dla pewnego τ_0, to pozostaje takim dla wszystkich τ. Fakt ten jest konsekwencją równania

$$\frac{d}{d\tau}\left(\frac{dX^i(\tau)}{d\tau}g_{ik}\frac{dX^k(\tau)}{d\tau}\right) = 0\,,$$

które wynika z równania ruchu.

(ii) Różniczkowanie p-form

Jeśli pole tensorowe t jest klasy $\mathcal{C}^m$, to pochodne cząstkowe stopnia $\leq m$ są przemienne, czyli tensory $\partial_{k_1}\ldots\partial_{k_s}t^{i_1\ldots i_p}{}_{j_1\ldots j_q}(X)$ są symetryczne we wskaźnikach $k_1\ldots k_s$ dla $s \leq m$. Zapiszemy ten fakt jako symboliczną tożsamość

$$\partial \wedge \partial = 0\,.$$

W szczególności, jeśli α jest takim polem $(p+1)$-form, klasy $\mathcal{C}^1$, że $\alpha = d\beta$ dla pewnego pola p-form β, to $d\alpha = 0$. Okazuje się, że prawdziwe jest również twierdzenie odwrotne: jeśli α jest klasy $\mathcal{C}^1$ i $d\alpha = 0$, to $\alpha = d\beta$ dla pewnego pola form β klasy $\mathcal{C}^2$. Twierdzenie pozostaje również prawdziwe, gdy obszar określenia pola α nie jest całą przestrzenią, lecz jej podzbiorem z pewnej klasy, która zawiera między innymi zbiory wypukłe (zawierające wraz z każdą parą punktów cały łączący je odcinek). Twierdzenie nosi nazwę lematu Poincaré; jego dokładne sformułowanie i dowód można znaleźć w standardowych podręcznikach analizy.

Pole elektromagnetyczne w czasoprzestrzeni Minkowskiego opisane jest dwuformą F spełniającą równania Maxwella. Jednym z tych równań jest warunek $dF = 0$. Z lematu Poincaré istnieje więc pole jednoformy A, nazywane potencjałem elektromagnetycznym, takie że $F = 2\,dA$, czyli $F_{ij} = \partial_i A_j - \partial_j A_i$.

(iii) Tożsamości różniczkowe w 3-wymiarowej przestrzeni euklidesowej
W tej przestrzeni w operacjach iloczynu wektorowego można użyć operacji różniczkowania ∇, co prowadzi do następujących tożsamości dla pola skalarnego Φ i pól wektorowych B i C:

$$\nabla \times \nabla\Phi = 0\,, \qquad \nabla\cdot(\nabla\times C) = 0\,,$$
$$\nabla\times(\nabla\times C) = \nabla(\nabla\cdot C) - \nabla^2 C\,,$$
$$\nabla\times(B\times C) = (C\cdot\nabla)B - (B\cdot\nabla)C + B\,(\nabla\cdot C) - C\,(\nabla\cdot B)\,,$$

gdzie notację użytą w ostatnim wzorze należy czytać: $\big[(C\cdot\nabla)B\big]^i = C^j\nabla_j B^i$.

§22 Przestrzeń afiniczna euklidesowa

1 Iloczyn skalarny i odległość

Jeśli przestrzeń wektorowa V, nad którą zbudowana jest przestrzeń afiniczna $\mathcal{M}$, jest wyposażona w metrykę, to mówimy również, że na $\mathcal{M}$ jest zadana metryka. W szczególności, jeśli V jest przestrzenią wektorową euklidesową, to $\mathcal{M}$ nazywamy ***afiniczną przestrzenią euklidesową***. Iloczyn skalarny w przestrzeni V będziemy w tym paragrafie oznaczać przez $x\cdot y$.

Wiele szczególnych cech przestrzeni afinicznej euklidesowej $\mathcal{M}$ jest bezpośrednią konsekwencją odpowiednich własności wektorowej przestrzeni euklidesowej V. Na przykład, wśród afinicznych układów odniesienia $(O,(e_i))$ wyróżnione miejsce zajmują układy, w których baza (e_i) jest ortonormalna. Dalej, w przestrzeni $\mathcal{M}$ można wprowadzić funkcję ***odległości pary punktów***:

$$d(P,Q) := \big\|\overrightarrow{PQ}\big\|\,.$$

Przestrzeń $\mathcal{M}$ wyposażona w tę miarę odległości jest przestrzenią metryczną w sensie topologicznym, tj. funkcja odległości ma następujące własności:

(i) $d(X,Y) > 0$ dla $X \neq Y$, $d(X,X) = 0$;

(ii) $d(X,Y) = d(Y,X)$;

(iii) $d(X,Y) + d(Y,Z) \geq d(X,Z)$.

Wszystkie te własności są konsekwencjami odpowiednich własności normy; szczegóły pozostawiamy czytelnikowi jako ćwiczenie.

Definiujemy ponadto ***odległość zbiorów*** B ***i*** C, gdzie $B, C \subseteq \mathcal{M}$, wzorem

$$d(B,C) := \inf\{d(X,Y) \mid X \in B\,,\ Y \in C\}\,.$$

Jeśli B jest jednopunktowym zbiorem, $B = \{P\}$, to mówimy wtedy o ***odległości punktu*** P ***od zbioru*** C i piszemy w uproszczeniu

$$d(P,C) = \inf\{d(P,Y) \mid Y \in C\}\,.$$

Łatwo przekonać się, że odległość zbiorów ma następujące własności:

$$B \cap C \neq \varnothing \Rightarrow d(B,C) = 0\,, \qquad d(B,C) = d(C,B)\,,$$
$$d(B,C) = \inf\{d(X,C) \mid X \in B\}\,.$$

2 Równania afinicznych podprzestrzeni przestrzeni euklidesowej

Ogólne równanie dowolnej k-płaszczyzny $P + W$ jest warunkiem jednoczesnego zerowania się $n - k$ liniowo niezależnych form na wszystkich wektorach $\overrightarrow{PX}$ (p. 10, §21). Jednak w przypadku przestrzeni z niezdegenerowanym iloczynem skalarnym formy liniowe można utożsamić, przez podniesienie wskaźników, z wektorami, i wtedy wartość formy f na wektorze x możemy zapisać jako

$$f(x) = f_j x^j = f^i g_{ij} x^j = f \cdot x\,,$$

gdzie po prawej stronie f oznacza wektor utożsamiony z formą, a kropką oznaczyliśmy iloczyn skalarny. Ograniczmy się dalej do przestrzeni euklidesowej (w przypadku dowolnej niezdegenerowanej metryki stosuje się wynik przykładu (v), p. 11, §19). W tym przypadku każda podprzestrzeń wektorowa W i jej ortogonalne uzupełnienie $W^\perp$ są niezdegenerowane. Niech $f_1, \ldots, f_k$ będzie bazą podprzestrzeni W. Wtedy bazę podprzestrzeni $W^\perp$ uzyskujemy wybierając $n-k$ liniowo niezależnych rozwiązań $f_{k+1}, \ldots, f_n$ układu równań

$$f_i \cdot x = 0\,, \quad i = 1, \ldots, k\,.$$

Układ wektorów $f_1, \ldots, f_n$ tworzy łącznie bazę całej przestrzeni V. Odwrotnie, jeśli dana jest baza $f_{k+1}, \ldots, f_n$ przestrzeni $W^\perp$, to podprzestrzeń W jest rozpięta wektorami bazowymi $f_1, \ldots, f_k$, otrzymanymi jako liniowo niezależne rozwiązania układu równań

$$f_i \cdot x = 0\,, \quad i = k+1, \ldots, n\,.$$

Zastosowanie wyniku powyższej dyskusji do równań podprzestrzeni afinicznych daje następujące ich sformułowanie. Jeśli ciągi wektorów liniowo niezależnych $f_1, \ldots, f_k$ oraz $f_{k+1}, \ldots, f_n$ są związane omówionymi wyżej warunkami, to parametryczne równanie podprzestrzeni afinicznej

$$\overrightarrow{OX} = a + \sum_{i=1}^{k} t^i f_i$$

jest równoważne równaniu ogólnemu

$$f_i \cdot \overrightarrow{OX} = a_i\,, \quad i = k+1, \ldots, n\,, \quad \text{gdzie} \quad a_i = f_i \cdot a\,.$$

3 Podprzestrzenie w trójwymiarowej przestrzeni afinicznej euklidesowej

Nietrywialne podprzestrzenie afiniczne (tj. poza 0-wymiarowymi punktami i całą przestrzenią) przestrzeni trójwymiarowej to proste i płaszczyzny. Na ich równania parametryczne nie ma wpływu obecność struktury euklidesowej, i mają one postać:

$$\begin{aligned} \overrightarrow{OX} &= a + tb_1 + sb_2\,, && \text{— płaszczyzna,} \\ \overrightarrow{OX} &= c + td_3\,, && \text{— prosta,} \end{aligned}$$

gdzie b_1 i b_2 są wektorami liniowo niezależnymi, a t i s – dowolnymi parametrami.

Ogólne równania tych podprzestrzeni piszemy w formie omówionej w poprzednim punkcie. Niech niezerowy wektor b_3 będzie prostopadły do wektorów b_1 i b_2, a liniowo niezależne wektory d_1 i d_2 – prostopadłe do d_3. Oznaczmy ponadto $\mu = b_3 \cdot a$, $\nu_i = d_i \cdot c$, $i = 1, 2$. Równania ogólne podprzestrzeni mają wtedy postać:

$$\begin{aligned} b_3 \cdot \overrightarrow{OX} &= \mu\,, && \text{— płaszczyzna,} \\ d_i \cdot \overrightarrow{OX} &= \nu_i\,,\ i = 1, 2\,, && \text{— prosta.} \end{aligned}$$

W trójwymiarowej przestrzeni euklidesowej wektor prostopadły do pary liniowo niezależnych wektorów najprościej wyznacza się jako ich iloczyn wektorowy. Jeśli więc znane są wektory b_1, b_2, to można przyjąć $b_3 = b_1 \times b_2$. Podobnie, jeśli dane są wektory d_1, d_2, to można przyjąć $d_3 = d_1 \times d_2$.

Zapisujemy teraz trzecią z omówionych w punkcie 10, §21, postaci równań podprzestrzeni:

$$\overrightarrow{OX} \wedge b_1 \wedge b_2 = a \wedge b_1 \wedge b_2\,, \qquad \text{— płaszczyzna,}$$
$$\overrightarrow{OX} \wedge d_3 = c \wedge d_3\,, \qquad \text{— prosta.}$$

Na ogół wygodniejsze są postaci otrzymane stąd przez dualność. Dla płaszczyzny odtwarzamy wtedy ponownie równanie ogólne, z wektorem b_3 w postaci iloczynu wektorowego wektorów b_1 i b_2. Dla prostej natomiast dostajemy:

$$\overrightarrow{OX} \times d_3 = c \times d_3\,, \qquad \text{— prosta.}$$

Przy zmieniającym się wektorze c prawa strona tego równania może stać się dowolnym wektorem prostopadłym do wektora d_3 (pozostawiamy wykazanie tego faktu jako ćwiczenie), więc przy dowolnym $k \perp d_3$ równanie

$$\overrightarrow{OX} \times d_3 = k$$

jest poprawnym równaniem prostej.

4 Odległość podprzestrzeni w przestrzeni euklidesowej

Znajdziemy w tym punkcie przepis na odległość pary dowolnych podprzestrzeni $P + W$ i $P' + W'$. Wnioskujemy z twierdzenia 7, §21, że

$$\text{jeśli } \overrightarrow{PP'} \in W + W'\,, \quad \text{to } d(P + W, P' + W') = 0\,.$$

Zobaczymy, że również w ogólnym przypadku odległość jest określona przez wektor $\overrightarrow{PP'}$ i jego ułożenie względem podprzestrzeni $W + W'$.

Zastosujmy twierdzenie 4, §14, o ortogonalnym dopełnieniu do podprzestrzeni $W + W'$. Dostaniemy wtedy rozkład

$$V = (W + W') \oplus (W + W')^\perp\,,$$

a więc jednoznaczne przedstawienie dowolnego wektora $x \in V$ w postaci sumy

$$x = x_\parallel + x_\perp\,, \quad \text{gdzie } x_\parallel \in W + W'\,, \quad x_\perp \in (W + W')^\perp\,.$$

W szczególności mamy $\overrightarrow{PP'} = \overrightarrow{PP'}_\parallel + \overrightarrow{PP'}_\perp$. Wybierzmy po jednym punkcie z każdej z podprzestrzeni afinicznych:

$$Q = P + x\,, \quad x \in W\,, \quad \text{i} \quad Q' = P' + x'\,, \quad x' \in W'\,.$$

Wtedy ortogonalny rozkład wektora $\overrightarrow{QQ'}$ ma postać

$$\overrightarrow{QQ'} = \overrightarrow{PP'} + x' - x = \big(\overrightarrow{PP'}_{\parallel} + x' - x\big) + \overrightarrow{PP'}_{\perp}\,.$$

Korzystając z twierdzenia Pitagorasa (p. 9, §15) dostajemy

$$d(Q,Q')^2 = \big\|\overrightarrow{PP'}_{\parallel} + x' - x\big\|^2 + \big\|\overrightarrow{PP'}_{\perp}\big\|^2\,, \quad \text{więc} \;\; d(Q,Q') \geq \big\|\overrightarrow{PP'}_{\perp}\big\|\,.$$

Wybierzmy wektory x i x' tak, aby $\overrightarrow{PP'}_{\parallel} = x - x'$ – jest to zawsze możliwe, bo $\overrightarrow{PP'}_{\parallel} \in W + W'$. Przy takim wyborze punktów Q i Q' ich odległość osiąga dolne ograniczenie $\big\|\overrightarrow{PP'}_{\perp}\big\|$. Ponieważ odległość podprzestrzeni $P + W$ i $P' + W'$ jest dolnym kresem zbioru liczb $d(Q,Q')$, to dostajemy następujący wynik.

Twierdzenie 1. *Niech $P + W$ i $P' + W'$ będą podprzestrzeniami afinicznymi przestrzeni $\mathcal{M}$, i oznaczmy*

$$\overrightarrow{PP'} = \overrightarrow{PP'}_{\parallel} + \overrightarrow{PP'}_{\perp}\,, \quad \overrightarrow{PP'}_{\parallel} \in W + W'\,, \quad \overrightarrow{PP'}_{\perp} \in (W + W')^{\perp}\,.$$

Wtedy:

(i) $d(P + W, P' + W') = \big\|\overrightarrow{PP'}_{\perp}\big\|$;

(ii) jeśli $\overrightarrow{PP'}_{\parallel} = x - x'$, $x \in W$, $x' \in W'$, *i oznaczyć* $Q = P + x$, $Q' = P' + x'$, *to*

$$\overrightarrow{QQ'} \perp W + W' \quad i \quad d(Q,Q') = d(P + W, P' + W')\,.$$

Rozkład wektora $\overrightarrow{PP'}_{\parallel}$ na wektory z podprzestrzeni W i W' jest jednoznaczny wtedy, i tylko wtedy, gdy $W \cap W' = \{0\}$, i wtedy też istnieje tylko jedna para punktów Q, Q' takich, jak w punkcie (ii) twierdzenia. W przeciwnym przypadku, jeśli (Q, Q') jest taką parą, to zbiór wszystkich par spełniających warunek tego punktu jest równy $\{(Q+y, Q'+y) \mid y \in W \cap W'\}$. W szczególnym przypadku, gdy $W \subseteq W'$, jest $W \cap W' = W$, więc punkty $Q + y$ w tym zbiorze par przebiegają wtedy całą podprzestrzeń $P + W$. Stąd mamy natychmiastowy wniosek.

Wniosek 2. *Jeśli $W \subseteq W'$, to każdy punkt podprzestrzeni afinicznej $P + W$ leży w takiej samej odległości od podprzestrzeni $P' + W'$.*

W szczególnym przypadku $W = \{0\}$ twierdzenie 1 wyznacza odległość punktu P od podprzestrzeni $P' + W'$. Mamy wtedy $\overrightarrow{PP'}_{\parallel} \in W'$, więc w punkcie (ii) twierdzenia 1 należy podstawić $Q = P$, $Q' = P' - \overrightarrow{PP'}_{\parallel}$. Podstawiając w ostatnim wzorze $\overrightarrow{PP'}_{\parallel} = \overrightarrow{PP'} - \overrightarrow{PP'}_{\perp}$ dostajemy równoważnie $Q' = P + \overrightarrow{PP'}_{\perp}$. Stąd mamy drugi wniosek z twierdzenia.

Wniosek 3. *Punktem podprzestrzeni afinicznej $P' + W'$ położonym najbliżej punktu P jest punkt $P' - \overrightarrow{PP'}_{\|} = P + \overrightarrow{PP'}_{\perp}$ i zachodzi równość*

$$d(P, P' + W') = d(P, P' - \overrightarrow{PP'}_{\|}) = d(P, P + \overrightarrow{PP'}_{\perp}) = \left\|\overrightarrow{PP'}_{\perp}\right\| .$$

5 Objętość w przestrzeni euklidesowej

W zorientowanej euklidesowej przestrzeni wektorowej mamy do dyspozycji wyróżnioną n-formę i związany z nią n-wektor

$$e = n!\, e^1 \wedge \ldots \wedge e^n , \qquad \hat{e} = n!\, e_1 \wedge \ldots \wedge e_n ,$$

gdzie $e_1, \ldots, e_n$ jest dowolną dodatnio zorientowaną bazą ortonormalną (p. 13, §20). Używając tych wielkości dla określenia wyróżnionej objętości (zgodnie z dyskusją w punkcie 12, §21), dla równoległościanu opartego na wektorach $x_1, \ldots, x_n$ dostajemy

$$x_1 \wedge \ldots \wedge x_n = e(x_1, \ldots, x_n)\, e_1 \wedge \ldots \wedge e_n ,$$
$$\mathrm{vol}\,(x_1, \ldots, x_n) = |e(x_1, \ldots, x_n)| .$$

Podobnie jak w ogólnym przypadku, jeśli $x_i = x_i{}^k e_k$, to

$$e(x_1, \ldots, x_n) = \det(x_i{}^k) , \quad \mathrm{vol}\,(x_1, \ldots, x_n) = |\det(x_i{}^k)| .$$

Struktura euklidesowa pozwala jednak uzyskać dodatkowe przedstawienie objętości. Dzięki ortonormalności bazy (e_i) elementy macierzy Grama bazy (x_i) mają postać

$$x_i \cdot x_j = \sum_{k=1}^{n} x_i{}^k x_j{}^k , \quad \text{więc} \quad \det(x_i \cdot x_j) = \big(\det(x_i{}^k)\big)^2 .$$

Porównując z poprzednią reprezentacją objętości dostajemy

$$\mathrm{vol}\,(x_1, \ldots, x_n) = \sqrt{\det(x_i \cdot x_j)} .$$

Jeśli W jest podprzestrzenią przestrzeni V, to wybieramy jej ortonormalną bazę $e_1, \ldots, e_p$. Dla wektorów $x_1, \ldots, x_p$ z tej podprzestrzeni mamy

$$x_1 \wedge \ldots \wedge x_p = e_W(x_1, \ldots, x_p)\, e_1 \wedge \ldots \wedge e_p ,$$
$$\mathrm{vol}_W(x_1, \ldots, x_p) = |e_W(x_1, \ldots, x_p)| .$$

Dalsze wzory pozostają w mocy z zamianą n na p.

W przestrzeni euklidesowej trójwymiarowej możemy zapisać wyniki za pomocą iloczynu wektorowego. Równanie dualne do równania

$$x \wedge y \wedge z = e(x, y, z)\, e_1 \wedge e_2 \wedge e_3$$

ma postać

$$e(x, y, z) = (x \times y) \cdot z\,, \quad \text{więc} \quad \operatorname{vol}(x, y, z) = |(x \times y) \cdot z\,|\,.$$

Jeśli W jest podprzestrzenią dwuwymiarową z bazą ortonormalną e_1, e_2, to dla dowolnych wektorów $x, y \in W$ mamy

$$x \wedge y = e_W(x, y)\, e_1 \wedge e_2\,.$$

Wybierzmy wektor e_3 tak, aby (e_1, e_2, e_3) była dodatnio zorientowaną bazą ortonormalną całej przestrzeni. Biorąc wektory dualne do obu stron ostatniej równości dostajemy

$$x \times y = e_W(x, y)\, e_3\,, \quad \text{więc} \quad \operatorname{vol}_W(x, y) = \|x \times y\|\,.$$

Objętość dwuwymiarową nazywamy ***polem powierzchni***.

6 Przykłady

(i) Odległość punktu od płaszczyzny w trzech wymiarach
W trójwymiarowej przestrzeni euklidesowej dany jest punkt Y i płaszczyzna wyznaczona takimi równaniami, jak w punkcie 3. Płaszczyzna przechodzi przez punkt $X_0 = O + a$, więc punktem płaszczyzny położonym najbliżej punktu Y jest $X_0 - \overrightarrow{YX_0}_{\parallel} = Y + \overrightarrow{YX_0}_{\perp}$, gdzie część równoległa i część prostopadła odnoszą się do podprzestrzeni rozpiętej wektorami b_1 i b_2. Część prostopadła to inaczej: część równoległa do wektora b_3. Na podstawie wyników punktu 15, §20,

$$\overrightarrow{YX_0}_{\perp} = \frac{b_3}{\|b_3\|} \cdot \overrightarrow{YX_0}\, \frac{b_3}{\|b_3\|}\,, \quad \text{więc} \quad \|\overrightarrow{YX_0}_{\perp}\| = \frac{|b_3 \cdot \overrightarrow{YX_0}|}{\|b_3\|}\,.$$

Podstawiając $\overrightarrow{YX_0} = a - \overrightarrow{OY}$ oraz korzystając z równania ogólnego płaszczyzny dostajemy dwie inne postaci wzoru na odległość

$$\|\overrightarrow{YX_0}_{\perp}\| = \frac{|b_3 \cdot (a - \overrightarrow{OY})|}{\|b_3\|} = \frac{|\mu - b_3 \cdot \overrightarrow{OY}|}{\|b_3\|}\,.$$

(ii) Odległość punktu od prostej w trzech wymiarach
W trójwymiarowej przestrzeni euklidesowej dany jest punkt Y i prosta wyznaczona takimi równaniami, jak w punkcie 3. Prosta przechodzi przez punkt $X_0 = O + c$, więc punktem prostej najbliżej położonym punktu Y jest $X_0 - \overrightarrow{YX_0}_{\|} = Y + \overrightarrow{YX_0}_{\perp}$, gdzie teraz część równoległa i część prostopadła odnoszą się do podprzestrzeni rozpiętej wektorem d_3. Na podstawie wyników punktu 15, §20, część prostopadła to

$$\overrightarrow{YX_0}_{\perp} = \left(\frac{d_3}{\|d_3\|} \times \overrightarrow{YX_0}\right) \times \frac{d_3}{\|d_3\|}, \quad \text{więc} \quad \|\overrightarrow{YX_0}_{\perp}\| = \left\|\frac{d_3}{\|d_3\|} \times \overrightarrow{YX_0}\right\|.$$

Podstawiając $\overrightarrow{YX_0} = c - \overrightarrow{OY}$ oraz korzystając z równania prostej w postaci iloczynu wektorowego dostajemy dwie inne postaci wzoru na odległość

$$\|\overrightarrow{YX_0}_{\perp}\| = \frac{\|d_3 \times (c - \overrightarrow{OY})\|}{\|d_3\|} = \frac{\|k - \overrightarrow{OY} \times d_3\|}{\|d_3\|}.$$

(iii) Przybliżone rozwiązania sprzecznego równania liniowego (tzw. metoda najmniejszych kwadratów)
Niech $A \in \mathcal{L}(V, W)$, przy czym W jest przestrzenią euklidesową z metryką g, i $b \in W$. Zbiór $-b + AV$ jest podprzestrzenią afiniczną przestrzeni W traktowanej jak przestrzeń afiniczna. Jeśli ta podprzestrzeń przechodzi przez wektor zerowy, to równanie $Ax = b$ ma rozwiązania ze względu na $x \in V$. W przeciwnym przypadku równanie nie ma rozwiązań, ale można pytać o takie wektory x, które najlepiej aproksymują spełnienie tego równania. Dokładniej, szukamy takich wektorów, dla których euklidesowa odległość Ax od b jest minimalna. Równoważnie możemy powiedzieć, że $-b + Ax$ ma być wektorem podprzestrzeni afinicznej $-b + AV$ leżącym najbliżej wektora zerowego.

Przy użyciu rezultatów punktu 4 łatwo pokazać, że szukane wektory x to te, dla których zachodzi:

$$g(Ax - b, Ay) = 0 \quad \text{dla każdego} \quad y \in V,$$

przy czym rozwiązanie jest jednoznaczne wtedy, i tylko wtedy, gdy $\operatorname{Ker} A = \{0\}$. Przez opuszczenie wskaźnika w wektorze $Ax - b$ dostajemy formę $g(Ax - b)$, więc ten warunek można zapisać jako

$$\langle g(Ax - b), Ay \rangle_W = 0 \quad \text{dla każdego} \quad y \in V.$$

Definicja odwzorowania transponowanego (przykład (ii) w punkcie 20, §18) pozwala zapisać lewą stronę tego równania jako $\langle A^{\mathrm{T}} g(Ax - b), y \rangle_V$. Stąd ostatecznie szukane wektory są rozwiązaniami równania

$$A^{\mathrm{T}} g(Ax - b) = 0.$$

Jeśli wybrać dowolną bazę $(e_1, \ldots, e_n)$ przestrzeni V oraz bazę ortonormalną $(f_1, \ldots, f_m)$ przestrzeni W i odnieść wszystkie obiekty do wybranych baz, to równanie przyjmie postać

$$\mathbf{A}^{\mathrm{T}} \mathbf{A}\, \mathbf{x} = \mathbf{A}^{\mathrm{T}}\, \mathbf{b}\,.$$

(iv) Zmiana p-powierzchni przy nieosobliwej transformacji liniowej
Niech W będzie p-wymiarową podprzestrzenią przestrzeni euklidesowej V, $(f_1, \ldots, f_p)$ – jej dowolną bazą, a $A \in \mathcal{L}(V, V)$ – nieosobliwym operatorem. Równoległościan p-wymiarowy oparty na wektorach (f_i) przechodzi pod działaniem tego operatora w równoległościan oparty na wektorach (Af_i). Kwadrat p-powierzchni początkowego równoległościanu to $\det(f_i, f_j)$, a kwadrat p-powierzchni nowego równoległościanu wyraża się przez

$$\det(Af_i, Af_j) = \det(f_i, A^* A f_j) = \det(f_i, |A|^2 f_j) = \det(f_i, [P|A|^2 P]_W f_j)\,,$$

gdzie $\{P, \mathrm{id} - P\}$ jest ortogonalnym rozkładem identyczności związanym z rozkładem $V = W \oplus W^{\perp}$, a $[P|A|^2 P]_W$ jest zawężeniem operatora $P|A|^2 P$ do operatora na jego inwariantnej podprzestrzeni W. Stosunek p-powierzchni tych równoległościanów nie zależy więc od kształtu początkowego równoległościanu (czyli bazy (f_i) przestrzeni W) i wynosi $\sqrt{\det[P|A|^2 P]_W}$. Przy wykorzystaniu wyników zadania 116 można ten wynik zapisać prościej, zauważając, że $[P|A|^2 P]_W = |A|^2{}_{WW}$ (część operatora $|A|^2$ interpolująca pomiędzy W i W, przy rozkładzie $W = W \oplus W^{\perp}$).

(v) Objętość w przestrzeni z niezdegenerowaną symetryczną metryką
Część wyników dotyczących objętości, omówionych w punkcie 5 (z pominięciem specjalnych rezultatów w trzech wymiarach) przenosi się – z pewnymi zastrzeżeniami – na inne rzeczywiste przestrzenie z niezdegenerowaną symetryczną metryką (np. przestrzeń Minkowskiego). W formule wyrażającej objętość przez $\det(x_i \cdot x_j)$ należy dodać znak bezwzględnej wartości pod pierwiastkiem. Ponadto, pojęcie objętości p-wymiarowych wyznaczonych układami ortonormalnymi ma zastosowanie tylko do niezdegenerowanych podprzestrzeni.

§23 Odwzorowania afiniczne

1 Podstawowe definicje

Spotkaliśmy się już dwukrotnie w tym podręczniku z odwzorowaniami zachowującymi struktury algebraiczne: były to homomorfizmy grupowe – zachowujące strukturę mnożenia grupowego, oraz homomorfizmy liniowe – zachowujące strukturę przestrzeni wektorowej, a więc dodawania wektorów i mnożenia ich przez liczby. Obiektami rozważanymi w tym paragrafie będą odwzorowania pełniące analogiczną rolę w przypadku przestrzeni afinicznych.

Niech będą dane dwie przestrzenie afiniczne $(\mathcal{M}, V)$ i $(\mathcal{N}, W)$ – zaznaczamy jawnie przestrzenie wektorowe, nad którymi te przestrzenie afiniczne są zbudowane. Para odwzorowań (f, A), gdzie $f : \mathcal{M} \mapsto \mathcal{N}$, $A : V \mapsto W$, zadaje odwzorowanie pierwszej struktury w drugą. Niech przestrzenie V i W będą utworzone nad wspólnym ciałem $\mathbb{K}$. Mówimy, że para (f, A) jest ***odwzorowaniem afinicznym*** (lub ***homomorfizmem afinicznym***), jeśli zachowuje ona strukturę przestrzeni wektorowych oraz działanie dodawania wektorów do punktów, tj. spełnione są warunki:

(i) A jest odwzorowaniem liniowym;

(ii) $\forall X \in \mathcal{M},\ y \in V : \quad f(X + y) = f(X) + A\,y.$

Odwzorowanie A nazywamy ***częścią liniową*** tego odwzorowania afinicznego. Widać wprost z drugiego warunku definicyjnego, że

$$Ay = \overrightarrow{f(X)f(X+y)}$$

przy dowolnym wyborze punktu X. Stąd część liniowa A odwzorowania afinicznego (f, A) jest jednoznacznie określona przez odwzorowanie f, mówimy więc często w uproszczeniu, że f jest odwzorowaniem afinicznym.

Wprowadzamy dla oznaczenia szczególnych typów odwzorowań afinicznych nazwy analogiczne do przypadku liniowego. Bijektywne odwzorowanie afiniczne nazywamy ***izomorfizmem afinicznym***, dowolne odwzorowanie afiniczne przestrzeni w samą siebie nazywamy ***endomorfizmem afinicznym***, a izomorfizm przestrzeni na samą siebie – ***automorfizmem afinicznym*** . Wprost z definicji odwzorowania afinicznego wynika, że częścią liniową endomorfizmu afinicznego jest endomorfizm liniowy.

2 Kryteria afiniczności odwzorowań

Następujące twierdzenie podaje dwa rozstrzygające warunki afiniczności.

Twierdzenie 1. *Niech $f : \mathcal{M} \mapsto \mathcal{N}$ będzie odwzorowaniem przestrzeni afinicznych. Następujące warunki są równoważne:*

(i) f jest odwzorowaniem afinicznym;

(ii) odwzorowanie stowarzyszonych z $\mathcal{M}$ i $\mathcal{N}$ przestrzeni

$$A : V \mapsto W\,, \quad Ay := \overrightarrow{f(O)f(O+y)}\,,$$

gdzie O jest ustalonym, dowolnie wybranym punktem przestrzeni $\mathcal{M}$, jest odwzorowaniem liniowym;

(iii) dla dowolnej kombinacji afinicznej punktów przestrzeni $\mathcal{M}$ zachodzi

$$f\Big(\sum_{i=0}^{m}\lambda_i X_i\Big) = \sum_{i=0}^{m}\lambda_i f(X_i)\,.$$

Dowód. Wykażemy ciąg implikacji (i)⇒(iii)⇒(ii)⇒(i).

(i)⇒(iii) Jeśli f jest afiniczne, z częścią liniową A, to mamy ciąg równości

$$\begin{aligned} f\Big(\sum_{i=0}^{m}\lambda_i X_i\Big) &= f\Big(X_0 + \sum_{i=1}^{m}\lambda_i \overrightarrow{X_0X_i}\Big) = f(X_0) + A\Big(\sum_{i=1}^{m}\lambda_i\, \overrightarrow{X_0X_i}\Big) = \\ &= f(X_0) + \sum_{i=1}^{m}\lambda_i\, A\,\overrightarrow{X_0X_i} = f(X_0) + \sum_{i=1}^{m}\lambda_i \overrightarrow{f(X_0)f(X_i)} = \sum_{i=0}^{m}\lambda_i f(X_i)\,, \end{aligned}$$

co kończy dowód tej implikacji.

(iii)⇒(ii) Wybierzmy punkt $O \in \mathcal{M}$. Niech będzie spełniony warunek (iii), więc w szczególności dla dowolnych liczb α, β oraz wektorów $x, y \in V$ mamy

$$\begin{aligned} &f\big((1-\alpha-\beta)O + \alpha(O+x) + \beta(O+y)\big) = \\ &\qquad\qquad = (1-\alpha-\beta)f(O) + \alpha f(O+x) + \beta f(O+y)\,. \end{aligned}$$

Wybierając dla kombinacji afinicznej stojącej wewnątrz nawiasów funkcyjnych po lewej stronie punkt odniesienia O, a dla kombinacji afinicznej stojącej po prawej stronie punkt odniesienia $f(O)$, dostajemy stąd

$$\begin{aligned} f\big(O + (\alpha x + \beta y)\big) &= f(O) + \alpha\overrightarrow{f(O)f(O+x)} + \beta\overrightarrow{f(O)f(O+y)}\,, \\ \text{czyli}\quad \overrightarrow{f(O)f\big(O + (\alpha x + \beta y)\big)} &= \alpha\overrightarrow{f(O)f(O+x)} + \beta\overrightarrow{f(O)f(O+y)}\,, \end{aligned}$$

co jest równoważne stwierdzeniu o liniowości odwzorowania A określonego w punkcie (ii) twierdzenia.

(ii)⇒(i) Niech będzie spełniony warunek (ii) dla pewnego punktu O. Wybierzmy dowolny punkt $X = O + y \in \mathcal{M}$ i dowolny wektor $z \in V$. Wtedy mamy ciąg równości

$$\begin{aligned} &f(X+z) = f\big(O + (y+z)\big) = f(O) + A(y+z) = \\ &\qquad\qquad = f(O) + Ay + Az = f(O+y) + Az = f(X) + Az\,, \end{aligned}$$

co jest stwierdzeniem afiniczności f, z częścią liniową A. □

Wniosek 2. *Niech* $(\mathcal{M}, V)$ *i* $(\mathcal{N}, W)$ *będą dwiema przestrzeniami afinicznymi, a* $O \in \mathcal{M}$ *dowolnie wybranym punktem. Wtedy:*

(i) warunek

$$f(X) = P + A\overrightarrow{OX}$$

zadaje wzajemnie jednoznaczną odpowiedniość pomiędzy odwzorowaniami afinicznymi $f : \mathcal{M} \mapsto \mathcal{N}$*, a parami* (P, A)*, gdzie* $P \in \mathcal{N}$*,* $A \in \mathcal{L}(V, W)$*; odwzorowanie* A *jest częścią liniową* f*; równoważna postać warunku to:*

$$f(O) = P\,, \qquad \textit{liniowa część } f = A\,;$$

(ii) odwzorowanie f *jest injektywne (odpowiednio: surjektywne, bijektywne) wtedy, i tylko wtedy, gdy odwzorowanie* A *jest injektywne (odpowiednio: surjektywne, bijektywne).*

Dowód. Punkt (i) jest prostą konsekwencją równoważności warunków (i) i (ii) w poprzednim twierdzeniu, a bezpośrednio z pierwszej postaci relacji między f i parami (P, A) wynika punkt (ii) wniosku. □

Punkt (ii) twierdzenia dla przypadku bijekcji można ująć tak: odwzorowanie afiniczne jest izomorfizmem wtedy, i tylko wtedy, gdy jego część liniowa jest izomorfizmem liniowym.

3 Obraz i przeciwobraz podprzestrzeni afinicznej

Obraz i przeciwobraz podprzestrzeni afinicznej w odwzorowaniu afinicznym jest podprzestrzenią afiniczną. Dokładniej, zachodzi następujące twierdzenie.

Twierdzenie 3. *Niech* $f : \mathcal{M} \mapsto \mathcal{N}$ *będzie odwzorowaniem afinicznym o części liniowej* A*, a* $O + V' \subseteq \mathcal{M}$ *i* $P + W' \subseteq \mathcal{N}$ *podprzestrzeniami afinicznymi. Wtedy*

$$f(O + V') = f(O) + AV'$$

oraz, jeśli przeciwobraz $f^{-1}(P + W')$ *jest niepusty i* $Q \in f^{-1}(P + W')$*, to*

$$f^{-1}(P + W') = Q + A^{-1}W'\,.$$

Dowód. Pierwsza równość jest natychmiastową konsekwencją punktu (ii) definicji odwzorowania afinicznego. Jeśli przeciwobraz podprzestrzeni $P + W'$ jest niepusty i należy do niego punkt Q, to $f(Q) \in P+W'$, więc $P+W' = f(Q)+W'$. Mamy teraz ciąg równoważności

$$X \in f^{-1}(P + W') \Longleftrightarrow f(X) = f(Q) + A\overrightarrow{QX} \in f(Q) + W' \Longleftrightarrow$$
$$\Longleftrightarrow A\overrightarrow{QX} \in W' \Longleftrightarrow \overrightarrow{QX} \in A^{-1}W'\,,$$

skąd wynika druga równość podana w tezie. □

4 Przykłady

(i) Formy afiniczne

Odwzorowanie afiniczne $f : \mathcal{M} \mapsto \mathbb{K}$ nazywamy ***formą afiniczną***. Wybierzmy punkt odniesienia $O \in \mathcal{M}$. Wtedy zgodnie z wnioskiem 2 zadanie formy afinicznej f to tyle, co wybranie liczby $p \in \mathbb{K}$ i formy φ na V – części liniowej f, i przyjęcie $f(X) = p + \varphi(\overrightarrow{OX})$.

Jeśli forma φ jest zerowa, to forma f jest funkcją stałą. W dalszym ciągu tego przykładu zakładamy, że $\varphi \neq 0$. Łatwo sprawdzić, że wówczas $f(\mathcal{M}) = \mathbb{K}$. Każdy zbiór jednoelementowy $\{r\} \subseteq \mathbb{K}$ stanowi podprzestrzeń afiniczną ciała $\mathbb{K}$ jako przestrzeni afinicznej, z zerową przestrzenią kierunkową. Stąd

$$f^{-1}(\{r\}) = Y + \operatorname{Ker}\varphi\,,$$

gdzie Y jest dowolnym punktem, dla którego $f(Y) = r$. W szczególności, jeśli punkt odniesienia wybrać tak, aby $f(O) = 0$, to

$$f(X) = \varphi(\overrightarrow{OX})\,, \qquad f^{-1}(\{0\}) = O + \operatorname{Ker}\varphi\,.$$

Ogólne równania podprzestrzeni afinicznych można zwięźle zapisać przy użyciu form afinicznych. Niech $f_1, \ldots, f_{n-k}$ ($n = \dim V$) będą formami afinicznymi o liniowo niezależnych częściach liniowych. Wtedy układ równań:

$$f_1(X) = 0\,, \quad \ldots, \quad f_{n-k}(X) = 0$$

zadaje k-wymiarową podprzestrzeń afiniczną, i każda taka podprzestrzeń może być tak określona.

(ii) Równania liniowe

Łącząc wyniki poprzedniego przykładu z dyskusją zamieszczoną w przykładach (ii) i (iii) z punktu 3, §18, przekonujemy się, że każdy układ równań liniowych może być zapisany w postaci

$$\mathbf{f}_i(\mathbf{X}) = 0\,, \quad i = 1, \ldots, m\,,$$

gdzie $\mathbf{f}_i$ są formami afinicznymi na przestrzeni afinicznej $\mathbb{K}^n$. Rozwiązanie ogólne układu jest podprzestrzenią afiniczną w $\mathbb{K}^n$.

(iii) Czasoprzestrzeń Galileusza

Niech $(\mathcal{M}, V)$ będzie czterowymiarową rzeczywistą przestrzenią afiniczną, a $t : \mathcal{M} \mapsto \mathbb{R}$ niech będzie formą afiniczną o niezerowej części liniowej τ. Oznaczmy $S = \operatorname{Ker}\tau$. Na trójwymiarowej podprzestrzeni $S \subseteq V$ zadajemy dodatnio określoną metrykę g. Struktura $(\mathcal{M}, t, g)$ jest w fizyce klasycznej modelem ***czasoprzestrzeni Galileusza***. Wartość $t(X)$ ma interpretację czasu przypisanego zdarzeniu X. Każda z afinicznych podprzestrzeni stałego czasu, $Y + S$, $Y \in \mathcal{M}$, tworzy z metryką g euklidesową przestrzeń afiniczną. Struktura metryczna służy do mierzenia odległości i kątów wśród zdarzeń równoczasowych. Każda prosta $Z + L(u)$, gdzie u jest dowolnym wektorem dla którego $\tau(u) = 1$, opisuje inercjalny ruch punktu materialnego.

5 Składanie odwzorowań afinicznych

Składanie odwzorowań zachowuje afiniczność – w całkowitej analogii do przypadku liniowego.

Twierdzenie 4. *Niech $f : \mathcal{M} \mapsto \mathcal{N}$, $g : \mathcal{N} \mapsto \mathcal{R}$ będą odwzorowaniami afinicznymi o częściach liniowych A i B odpowiednio, a $O \in \mathcal{M}$ – dowolnym punktem. Wtedy*

$$g \circ f(X) = g\big(f(O)\big) + BA\,\overrightarrow{OX}\,,$$

więc $g \circ f$ jest odwzorowaniem afinicznym, o części liniowej BA. Jeśli f jest bijekcją, to f^{-1} jest odwzorowaniem afinicznym o części liniowej A^{-1}.

Dowód. Jeśli f i g są afiniczne, to

$$g\big(f(X)\big) = g(f(O) + A\overrightarrow{OX}) = g\big(f(O)\big) + BA\,\overrightarrow{OX},$$

co dowodzi równości z tezy. Z wniosku 2 wynikają teraz pozostałe stwierdzenia tezy. □

6 Przestrzenie afinicznie izomorficzne

Podobnie jak w przypadku liniowym mówimy, że przestrzenie $(\mathcal{M}, V)$ i $(\mathcal{N}, W)$ są ***afinicznie izomorficzne***, jeśli istnieje izomorfizm jednej przestrzeni na drugą. Jeśli $f : \mathcal{M} \mapsto \mathcal{N}$ jest izomorfizmem, to część liniowa jest też izomorfizmem. Odwrotnie, jeśli $A : V \mapsto W$ jest izomorfizmem liniowym, to wystarczy wybrać parę punktów $O \in \mathcal{M}$, $P \in \mathcal{N}$ i położyć $f(O+x) = P+Ax$, aby otrzymać izomorfizm przestrzeni $\mathcal{M}$ i $\mathcal{N}$. Kryterium izomorficzności przestrzeni przenosi się więc z przypadku wektorowego (tw. 7, §10) na afiniczny.

Twierdzenie 5. *Przestrzenie afiniczne skończonego wymiaru są izomorficzne wtedy, i tylko wtedy, gdy mają taki sam wymiar.*

7 Translacje. Odwzorowania o wspólnej części liniowej

Rozważymy w tym punkcie endomorfizmy afiniczne (a więc odwzorowania przestrzeni w samą siebie) szczególnego typu: takie, których część liniowa jest identycznością. Zgodnie z twierdzeniem 1, odwzorowanie f jest w tej klasie wtedy, i tylko wtedy, gdy $\overrightarrow{f(O)f(X)} = \overrightarrow{OX}$ dla dowolnie wybranego ustalonego punktu O i każdego punktu X. Warunek ten to tyle co $f(X) = f(O) + (X - O)$ lub inaczej $f(X) = X + (f(O) - O)$. Oznaczając $y \equiv f(O) - O \in V$ dostajemy ostatecznie $f(X) = X + y$.

Dla dowolnie wybranego wektora $y \in V$ odwzorowanie

$$t_y : \mathcal{M} \mapsto \mathcal{M}\,, \quad t_y(X) = X + y$$

nazywamy ***translacją o wektor*** y. Pierwszy punkt następującego twierdzenia podsumowuje powyższą dyskusję.

Twierdzenie 6.
(i) Endomorfizm afiniczny jest translacją wtedy, i tylko wtedy, gdy jego część liniowa jest identycznością. Stąd translacje są bijekcjami.
(ii) Zbiór translacji $T(\mathcal{M})$ jest podgrupą grupy symetrycznej $S_{\mathcal{M}}$, izomorficzną z grupą addytywną V, tj. spełnione są równości:

$$t_x \circ t_y = t_{x+y}\,, \qquad t_x = \mathrm{id} \Longleftrightarrow x = 0\,.$$

Dowód. Pozostaje wykazać relacje z punktu (ii). Dla dowolnego punktu Z jest

$$t_x \circ t_y(Z) = t_x(Z + y) = Z + (x + y) = t_{x+y}(Z)\,,$$

więc pierwsza równość jest spełniona. Druga relacja wynika natychmiast z twierdzenia 1, §21. □

Jeśli $f : \mathcal{M} \mapsto \mathcal{N}$ jest odwzorowaniem afinicznym, to dla dowolnej translacji t_x przestrzeni $\mathcal{N}$ odwzorowanie $t_x \circ f : \mathcal{M} \mapsto \mathcal{N}$ ma taką samą część liniową jak f. Odwrotnie, niech f i $g : \mathcal{M} \mapsto \mathcal{N}$ mają taką samą część liniową A. Wybierzmy punkt $O \in \mathcal{M}$. Wtedy zgodnie z punktem (i) wniosku 2 istnieją punkty $P, Q \in \mathcal{N}$ takie, że

$$f(X) = P + A\overrightarrow{OX}\,, \quad g(X) = Q + A\overrightarrow{OX}\,, \quad \text{więc} \quad g(X) = f(X) + \overrightarrow{PQ}\,.$$

Dostajemy więc następujący wynik.

Twierdzenie 7. *Odwzorowania afiniczne $f, g : \mathcal{M} \mapsto \mathcal{N}$ mają taką samą część liniową wtedy, i tylko wtedy, gdy istnieje translacja t_x przestrzeni $\mathcal{N}$ taka, że*

$$g = t_x \circ f\,.$$

8 Punkty stałe endomorfizmu

Punktem stałym odwzorowania f zbioru $\mathcal{M}$ w samego siebie nazywamy taki $X \in \mathcal{M}$, że $f(X) = X$. Zbadamy istnienie punktów stałych dla endomorfizmu afinicznego. W tym przypadku pytanie o punkt stały jest szukaniem punktu X takiego, że wektor $\overrightarrow{Xf(X)}$ jest zerowy. Jest widoczne w szczególności, że translacje nie mają punktów stałych (z wyjątkiem translacji o wektor zerowy, która jest identycznością).

Niech f będzie endomorfizmem afinicznym przestrzeni $\mathcal{M}$, o części liniowej A. Wybierzmy punkt odniesienia O. Wtedy dla dowolnego punktu $X \in \mathcal{M}$ mamy

$$f(X) - X = f(O) + A\overrightarrow{OX} - (O + \overrightarrow{OX}) = (f(O) - O) + (A - \mathrm{id})\overrightarrow{OX}\,.$$

Stąd przy dowolnym wyborze punktu odniesienia O mamy

$$\{\overrightarrow{Xf(X)} \mid X \in \mathcal{M}\} = \overrightarrow{Of(O)} + \mathrm{Im}(A - \mathrm{id})\,.$$

Zbiór ten jest podprzestrzenią afiniczną przestrzeni wektorowej V. Przechodzi ona przez zero wtedy, i tylko wtedy, gdy wektor $\overrightarrow{Of(O)}$ należy do jej przestrzeni kierunkowej (patrz przykład (ii), p. 3, §21). Dostajemy więc kryterium ujęte w następujące twierdzenie.

Twierdzenie 8. *Niech f będzie endomorfizmem afinicznym przestrzeni $\mathcal{M}$ nad przestrzenią wektorową V, a $O \in \mathcal{M}$ dowolnym punktem. Odwzorowanie f ma punkt stały wtedy, i tylko wtedy, gdy $\overrightarrow{Of(O)} \in \mathrm{Im}(A - \mathrm{id})$. Jeśli ten warunek jest spełniony, to punkty stałe X są rozwiązaniami równania*

$$(A - \mathrm{id})\overrightarrow{OX} = -\overrightarrow{Of(O)}$$

i tworzą podprzestrzeń afiniczną $X_0 + \mathrm{Ker}(A - \mathrm{id})$, gdzie X_0 jest dowolnym rozwiązaniem tego równania.

Ostatnie stwierdzenie otrzymuje się natychmiast po wybraniu $O = X_0$.

9 Rozkład endomorfizmu względem punktu odniesienia

Wybierzmy w przestrzeni afinicznej $\mathcal{M}$ punkt odniesienia O. Działanie endomorfizmu tej przestrzeni można zapisać jako

$$f(X) = f(O) + A\,\overrightarrow{OX} = O + \overrightarrow{Of(O)} + A\,\overrightarrow{OX}\,.$$

Oznaczmy $\overrightarrow{Of(O)} = x$. Odpowiedniość, o której mowa w punkcie (i) wniosku 2, można więc w tym przypadku ująć jako wzajemnie jednoznaczną odpowiedniość endomorfizmów i par (x, A), gdzie $x \in V$, $A \in \mathcal{L}(V, V)$. Jeśli przy tym wyborze punktu odniesienia endomorfizmowi f odpowiada para (x, A), to będziemy pisać

$$f \equiv (x, A)_O\,, \quad \text{czyli} \quad (x, A)_O(Z) = O + x + A\,\overrightarrow{OZ}\,.$$

Jest widoczne, że odwzorowania postaci $(x, \mathrm{id})_O$ są translacjami, a dla każdego odwzorowania $(x, A)_O$ punkt odniesienia O jest punktem stałym wtedy, i tylko wtedy, gdy $x = 0$. Odwzorowanie tego ostatniego typu sprowadza się do liniowego odwzorowania wektorów wodzących.

Prawo składania symboli $(x, A)_O$ oraz ich zależność od punktu odniesienia O podaje następujące twierdzenie.

Twierdzenie 9. *Niech punkt O i wektor a będą dowolnie wybrane z przestrzeni afinicznej $\mathcal{M}$ i stowarzyszonej z nią przestrzeni wektorowej V odpowiednio. Wtedy*

$$\begin{array}{ll} (i) & (y, B)_O \circ (x, A)_O = (y + Bx, BA)_O\,; \\ (ii) & (x, A)_O = (x, \mathrm{id})_O \circ (0, A)_O\,; \\ (iii) & (x, A)_O = (x + (A - \mathrm{id})a, A)_{O+a}\,. \end{array}$$

Dowód. Mamy ciąg równości

$$(y, B)_O \circ (x, A)_O\,(Z) = (y, B)_O\,(O + x + A\,\overrightarrow{OZ}) = O + y + B\,(x + A\,\overrightarrow{OZ})\,,$$

skąd wynika punkt (i). Rozkład (ii) jest szczególnym przypadkiem punktu (i). Punkt (iii) wynika z ciągu równości

$$\begin{aligned}(x, A)_O(Z) &= O + x + A\overrightarrow{OZ} = O + x + Aa + A\overrightarrow{(O+a)Z} = \\ &= (O + a) + x + (A - \mathrm{id})a + A\overrightarrow{(O+a)Z} = (x + (A - \mathrm{id})a, A)_{O+a}(Z)\,.\end{aligned}$$

□

Punkt (iii) twierdzenia dostarcza innego dowodu twierdzenia o punkcie stałym: odwzorowanie $(x, A)_O$ ma punkt stały $O + a$ wtedy, i tylko wtedy, gdy równanie $x + (A - \mathrm{id})a = 0$ ma rozwiązanie ze względu na a.

10 Automorfizmy afiniczne

Złożenie izomorfizmów afinicznych jest izomorfizmem afinicznym, a także odwzorowanie odwrotne do izomorfizmu afinicznego jest izomorfizmem afinicznym. Stąd, w szczególności, automorfizmy przestrzeni afinicznej $\mathcal{M}$ tworzą grupę; oznaczamy ją $AF(\mathcal{M})$.

Niech przestrzenią wektorową stowarzyszoną z $\mathcal{M}$ będzie V. Zastosowanie twierdzenia 4 do automorfizmów daje następujący wynik. Odwzorowanie

$$AF(\mathcal{M}) \mapsto GL(V)\,, \qquad f \mapsto \text{ część liniowa } f$$

jest homomorfizmem, przy tym jest to homomorfizm surjektywny. Jądro tego homomorfizmu tworzą odwzorowania o identycznościowej części liniowej, czyli grupa translacji $T(\mathcal{M})$. Stąd na podstawie twierdzenia 10, §3, grupa $T(\mathcal{M})$ jest podgrupą niezmienniczą grupy $AF(\mathcal{M})$. Grupa ilorazowa $AF(\mathcal{M})/T(\mathcal{M})$ składa się z klas równoważności $[f]$, w każdej klasie zgrupowane są odwzorowania różniące się co najwyżej o translację, a więc wszystkie odwzorowania o takiej samej części liniowej. Określony wyżej homomorfizm indukuje teraz homomorfizm

$$AF(\mathcal{M})/T(\mathcal{M}) \mapsto GL(V)\,, \qquad [f] \mapsto \text{ część liniowa } f\,.$$

Jest to homomorfizm surjektywny, a jego jądro jest trywialne, więc

$$\text{grupy } \; AF(\mathcal{M})/T(\mathcal{M}) \; \text{ i } \; GL(V) \; \text{ są izomorficzne.}$$

Mówimy, że $AF(\mathcal{M})$ jest ***afinicznym rozszerzeniem*** grupy $GL(V)$.

Naturę afinicznego rozszerzenia można opisać dokładniej przy użyciu reprezentacji endomorfizmów afinicznych wprowadzonej w poprzednim punkcie. Twierdzenie 9 zastosowane do automorfizmów daje natychmiast następujący wynik.

Twierdzenie 10. *Niech $(\mathcal{M}, V)$ będzie przestrzenią afiniczną, a O jej dowolnym punktem. Automorfizmy afiniczne $AF(\mathcal{M})$ tworzą grupę złożoną z odwzorowań $(x, A)_O$, gdzie $x \in V$, $A \in GL(V)$, o prawie składania*

$$(y, B)_O \circ (x, A)_O = (y + Bx, BA)_O$$

oraz elemencie neutralnym i elementach odwrotnych odpowiednio:

$$\mathrm{id} = (0, \mathrm{id})_O\,, \qquad {(x, A)_O}^{-1} = (-A^{-1}x, A^{-1})_O\,.$$

Z dyskusji poprzedzającej twierdzenie widać, że jeśli $H(V)$ jest dowolną podgrupą grupy $GL(V)$, to zbiór automorfizmów

$$\{f \in AF(\mathcal{M}) \mid \text{część liniowa } f \in H(V)\}$$

jest podgrupą grupy $AF(\mathcal{M})$; nazywamy ją ***afinicznym rozszerzeniem podgrupy*** $H(V)$. Powyższe twierdzenie pozostaje w mocy dla każdego takiego rozszerzenia, z zawężeniem $A \in H(V)$ w założeniu twierdzenia. Rozważmy, w szczególności, rzeczywistą przestrzeń afiniczną. Powiemy, że jej automorfizm ***zachowuje orientację***, jeśli jest elementem afinicznego rozszerzenia grupy $GL_+(V)$.

Zwróćmy jeszcze raz uwagę na niejednakowy status elementów pary $(x, A)_O$ zadającej odwzorowanie afiniczne. Operator A jest częścią liniową odwzorowania afinicznego, jednoznacznie przez nie określoną, podczas gdy wektor x zmienia się wraz ze zmianą punktu odniesienia zgodnie z regułą (iii) z twierdzenia 9. W literaturze fizycznej ta różnica często nie jest dostatecznie uwypuklona.

11 Działanie endomorfizmu w afinicznym układzie odniesienia. Czynne i bierne transformacje przestrzeni afinicznej

Dla dalszego wyliczenia działania endomorfizmu przestrzeni afinicznej odnosimy go do afinicznego układu odniesienia.

Twierdzenie 11. *Niech $\big(O, (e_1, \ldots, e_n)\big)$ będzie dowolnym afinicznym układem odniesienia w przestrzeni $\mathcal{M}$. Wtedy dla każdego endomorfizmu afinicznego $(a, A)_O$ zachodzi równoważność*

$$\tilde{X} = (a, A)_O(X) \iff \tilde{\mathbf{X}} = \mathbf{a} + \mathbf{A}\mathbf{X}\,,$$

gdzie $\mathbf{X}$, $\tilde{\mathbf{X}}$, $\mathbf{a}$ są kolumnami współrzędnych w tym układzie, a $\mathbf{A}$ jest macierzą operatora A w bazie (e_i).

Dowód. Twierdzenie wynika z ciągu równoważności

$$\tilde{X} = (a, A)_O(X) \iff \tilde{X} = O + a + A\overrightarrow{OX} \iff$$
$$\iff \overrightarrow{O\tilde{X}} = a + A\overrightarrow{OX} \iff \tilde{\mathbf{X}} = \mathbf{a} + \mathbf{A}\mathbf{X}\,.$$

□

Niech, w szczególności, rozważany wyżej endomorfizm będzie automorfizmem, co jest równoważne nieosobliwości operatora A, a tym samym nieosobliwości macierzy $\mathbf{A}$. Przy tym założeniu poddajmy transformacji *odwrotnej* $(-A^{-1}a, A^{-1})_O$ układ odniesienia $(O, (e_i))$. Dostajemy

$$O' = (-A^{-1}a, A^{-1})_O(O) = O - A^{-1}a\,, \quad e'_i = A^{-1}e_i = e_j A^{-1j}{}_i\,.$$

Oznaczmy przez $\mathbf{X}$ i $\mathbf{X}'$ współrzędne *tego samego* punktu X w starym i nowym układzie odniesienia. Podstawiając związek między tymi układami do twierdzenia 6, §21, dostajemy

$$\mathbf{X}' = \mathbf{a} + \mathbf{A}\mathbf{X}\,.$$

Numeryczny związek współrzędnych $\mathbf{X}$ i $\tilde{\mathbf{X}}$ w ostatnim twierdzeniu jest taki sam, jak otrzymany teraz związek $\mathbf{X}$ i $\mathbf{X}'$, ale sposób jego otrzymania jest inny.

W pierwszym przypadku transformowaliśmy punkty przestrzeni (zgodnie z odwzorowaniem afinicznym) i badaliśmy związek współrzędnych w jednym ustalonym układzie odniesienia: tę operację nazywa się często ***czynną transformacją*** przestrzeni. W drugim przypadku transformowaliśmy układ odniesienia i porównywaliśmy współrzędne każdego punktu w tych dwóch układach: tę operację nazywa się często ***bierną transformacją*** przestrzeni (nie całkiem poprawnie – sama przestrzeń nie podlega tu transformacji). Uzyskanemu powyżej wynikowi możemy więc teraz nadać następujące sformułowanie.

Wniosek 12. *Zmiana współrzędnych punktów przestrzeni afinicznej przy czynnej transformacji przestrzeni jest taka sama, jak przy odwrotnej do niej transformacji biernej.*

12 Przykłady

(i) Izomorfizm afiniczny $\mathcal{M}$ z $\mathbb{K}^{\dim \mathcal{M}} \times \{1\}$

Jeśli wymiar przestrzeni afinicznej $\mathcal{M}$ jest równy n, to jest ona izomorficzna z przestrzenią $\mathbb{K}^n \times \{1\}$ (omówioną w przykładzie (ii), p. 9, §21). Wybierzmy w $\mathcal{M}$ układ afiniczny $(O, (e_i))$ i oznaczmy przez $\mathbf{X}$ kolumnę współrzędnych afinicznych punktu X w tym układzie, $X = O + X^i e_i$. Wtedy odwzorowanie

$$j : \mathcal{M} \mapsto \mathbb{K}^n \times \{1\}, \quad X \mapsto j(X) = \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix}$$

jest izomorfizmem afinicznym. Rozkład punktu w układzie $(O, (e_i))$ możemy symbolicznie zapisać przy użyciu notacji macierzowej jako

$$X = \begin{pmatrix} e_1 & \dots & e_n & O \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix},$$

gdzie powstające przy mnożeniu wyrażenie $O1 = O$ ma sens jako kombinacja afiniczna. Niech $(O', (e'_i))$ będzie innym układem afinicznym w $\mathcal{M}$, w którym kolumnę współrzędnych punktu X oznaczamy $\mathbf{X}'$, a związany z układem izomorfizm zapisujemy jako j'. Związkom primowanych i nieprimowanych układów i współrzędnych zanotowanym w twierdzeniu 6, §21, można teraz nadać formę:

$$\begin{pmatrix} e'_1 & \dots & e'_n & O' \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n & O \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} & \mathbf{a} \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{X}' \\ 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}^{-1} & -\boldsymbol{\beta}^{-1}\mathbf{a} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix}$$

(mnożenie $O0 = 0$ ma sens jako kombinacja wektorowa punktów, przykład (iii) w punkcie 9, §21).

(ii) Odwzorowania afiniczne hiperpłaszczyzn w przestrzeniach wektorowych
Niech w przestrzeniach V_i $(i = 1, 2)$ nad ciałem $\mathbb{K}$ dane będą wektory $z_i \in V_i$ i podprzestrzenie $W_i \subseteq V_i$ takie, że $V_i = L(z_i) \oplus W_i$. Podzbiory $z_i + W_i \subseteq V_i$ tworzą przestrzenie afiniczne nad przestrzeniami wektorowymi W_i. Istnieje następująca wzajemnie jednoznaczna odpowiedniość.

> Każde odwzorowanie afiniczne $z_1 + W_1 \mapsto z_2 + W_2$ rozszerza się jednoznacznie do liniowego odwzorowania $V_1 \mapsto V_2$. Odwrotnie, jeśli $A \in \mathcal{L}(V_1, V_2)$ i $A(z_1 + W_1) = Az_1 + AW_1 \subseteq z_2 + W_2$, to zawężenie dziedziny i przeciwdziedziny A do tych podzbiorów prowadzi do odwzorowania afinicznego $z_1 + W_1 \mapsto z_2 + W_2$.

(iii) Odwzorowania afiniczne $\mathbb{K}^n \times \{1\} \mapsto \mathbb{K}^m \times \{1\}$
Na podstawie poprzedniego przykładu odwzorowania te to zacieśnienia tych odwzorowań liniowych $\mathbb{K}^{n+1} \mapsto \mathbb{K}^{m+1}$, które wektory postaci $\begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix}$, $\mathbf{X} \in \mathbb{K}^n$, przeprowadzają w wektory $\begin{pmatrix} \mathbf{Y} \\ 1 \end{pmatrix}$, $\mathbf{Y} \in \mathbb{K}^m$. Posługując się utożsamieniem liniowych odwzorowań $\mathcal{L}(\mathbb{K}^r, \mathbb{K}^s)$ z macierzami $\mathbb{K}^{s \times r}$ (p. 12, §10), uzyskuje się działanie omawianych odwzorowań afinicznych jako mnożenie macierzowe:

$$\begin{pmatrix} \mathbf{Y} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix},$$

gdzie $\mathbf{A}$ jest macierzą $m \times n$. W szczególności dla $m = n$ grupa automorfizmów afinicznych składa się z macierzy tego typu, dla których $\det \mathbf{A} \neq 0$, z operacjami grupowymi

$$\begin{pmatrix} \mathbf{B} & \mathbf{b} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{BA} & \mathbf{b} + \mathbf{Ba} \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{a} \\ 0 & 1 \end{pmatrix}.$$

(iv) Macierzowa reprezentacja odwzorowań afinicznych
Niech $\mathcal{M}$ i $\mathcal{N}$ będą przestrzeniami afinicznymi z wybranymi układami afinicznymi $\big(O, (e_1, \ldots, e_n)\big)$ i $\big(P, (f_1, \ldots, f_m)\big)$ odpowiednio, a $f : \mathcal{M} \mapsto \mathcal{N}$ – odwzorowaniem afinicznym z częścią liniową A. Oznaczmy ponadto $f(O) = P + a$. Wybrane układy afiniczne określają izomorfizmy przestrzeni $\mathcal{M}$ i $\mathcal{N}$ z $\mathbb{K}^n \times \{1\}$ i $\mathbb{K}^m \times \{1\}$ odpowiednio, jak to opisaliśmy w przykładzie (i). Teraz łatwo pokazać, że związek $Y = f(X)$ zachodzi wtedy, i tylko wtedy, gdy kolumny współrzędnych punktów X i Y z dodanymi jedynkami wiążą się tak, jak w poprzednim przykładzie, przy czym $\mathbf{A}$ jest macierzą odwzorowania A w bazach (e_i) i (f_j),

natomiast **a** jest kolumną współrzędnych wektora a w bazie (f_j). Jeśli $\mathcal{R}$ jest kolejną przestrzenią afiniczną z wybranym układem afinicznym $\big(Q,(h_1,\ldots,h_k)\big)$, a $g:\mathcal{N}\mapsto\mathcal{R}$ odwzorowaniem afinicznym, to złożeniu odwzorowań $g\circ f$ odpowiada iloczyn macierzy reprezentujących odwzorowania f i g, z zachowaniem kolejności. W szczególności, grupa automorfizmów afinicznych przestrzeni $\mathcal{M}$ jest izomorficzna z grupą macierzową opisaną na końcu poprzedniego przykładu.

13 Izometrie afiniczne przestrzeni z iloczynem skalarnym

Niech $(\mathcal{M},V)$ i $(\tilde{\mathcal{M}},\tilde{V})$ będą przestrzeniami afinicznymi z iloczynami skalarnymi g i $\tilde{g}$ odpowiednio. Odwzorowanie afiniczne $f:\mathcal{M}\mapsto\tilde{\mathcal{M}}$ nazywamy ***izometrią afiniczną***, jeśli jego część liniowa jest izometrią liniową. Podobnie jak w przypadku liniowym powiemy, że przestrzenie są ***afinicznie izometryczne***, jeśli istnieje izometryczny izomorfizm odwzorowujący jedną przestrzeń na drugą.

Afiniczne izometrie wewnętrzne przestrzeni z niezdegenerowanym iloczynem skalarnym to automorfizmy afiniczne, których liniowe części są wewnętrznymi izometriami liniowymi przestrzeni wektorowej stowarzyszonej z przestrzenią afiniczną. Liniowe izometrie tworzą grupę, więc zgodnie z dyskusją z punktu 10 mamy następujący wynik.

Twierdzenie 13. *Afiniczne izometrie wewnętrzne stanowią podgrupę automorfizmów afinicznych, będącą afinicznym rozszerzeniem grupy liniowej izometrii wewnętrznych.*

14 Izometrie wewnętrzne rzeczywistej przestrzeni z niezdegenerowaną metryką symetryczną

Niech $(\mathcal{M},V)$ będzie rzeczywistą przestrzenią z niezdegenerowanym, symetrycznym iloczynem skalarnym g i oznaczmy $q(x)=g(x,x)$ dla każdego $x\in V$. Jeśli f jest afiniczną izometrią przestrzeni $\mathcal{M}$, to dla dowolnych dwóch punktów X i Y mamy $q\big(\overrightarrow{f(X)f(Y)}\big)=q\big(\overrightarrow{XY}\big)$ – mówimy, że f ***zachowuje interwały między punktami***; w szczególnym przypadku przestrzeni euklidesowej oznacza to zachowanie odległości: $d\big(f(X),f(Y)\big)=d(X,Y)$. Okazuje się, że zachodzi również wynikanie odwrotne. Zwracamy uwagę, że w następującym twierdzeniu nie zakłada się z góry afiniczności odwzorowania f.

Twierdzenie 14. *Niech $\mathcal{M}$ będzie rzeczywistą, skończenie wymiarową przestrzenią afiniczną z niezdegenerowanym, symetrycznym iloczynem skalarnym. Odwzorowanie $f:\mathcal{M}\mapsto\mathcal{M}$ jest afiniczną izometrią wewnętrzną wtedy, i tylko wtedy, gdy zachowuje interwały między punktami.*

Dowód. Niech f zachowuje interwały. Wybierzmy dowolny punkt odniesienia O i oznaczmy $Ay=\overrightarrow{f(O)f(O+y)}$. Afiniczność odwzorowania f jest równoważna liniowości odwzorowania A (tw. 1), a wówczas izometryczność f – izometryczności A. Wystarczy więc pokazać, że A jest liniową izometrią wewnętrzną.

Dla każdej pary wektorów $x, y \in V$ mamy

$$Ax - Ay = \overrightarrow{f(O)f(O+x)} - \overrightarrow{f(O)f(O+y)} = \overrightarrow{f(O+y)f(O+x)} .$$

Z zachowywania interwałów przez odwzorowanie f dostajemy

$$\begin{aligned} q(Ay) &= q\big(\overrightarrow{f(O)f(O+y)}\big) = q\big(\overrightarrow{O(O+y)}\big) = q(y) , \\ q(Ax - Ay) &= q\big(\overrightarrow{f(O+y)f(O+x)}\big) = q\big(\overrightarrow{(O+y)(O+x)}\big) = q(x-y) , \end{aligned}$$

a stąd dla każdej pary wektorów

$$\begin{aligned} g(Ax, Ay) = \frac{1}{2}\Big(q(Ax) + q(Ay) - q(Ax - Ay)\Big) = \\ = \frac{1}{2}\Big(q(x) + q(y) - q(x-y)\Big) = g(x, y) . \end{aligned}$$

Pozostaje wykazać liniowość A. Niech (e_i) będzie dowolną bazą ortonormalną. Wtedy, na mocy powyższej równości, (Ae_i) jest również bazą ortonormalną, więc wektory Az, $z \in V$, generują całą przestrzeń V. Dla dowolnej kombinacji liniowej $\alpha x + \beta y$ i dowolnego wektora z mamy z liniowości g w prawym argumencie i z zachowywania iloczynów skalarnych przez A:

$$g\big(Az, A(\alpha x + \beta y) - \alpha Ax - \beta Ay\big) = g(z, \alpha x + \beta y) - \alpha g(z, x) - \beta g(z, y) = 0 .$$

Z niezdegenerowania iloczynu skalarnego dostajemy więc liniowość A. $\square$

15 Kanoniczna postać wewnętrznej izometrii przestrzeni euklidesowej

W tym punkcie ograniczamy rozważania do przestrzeni euklidesowej. Izometria wewnętrzna, tak jak każdy endomorfizm, może być przedstawiona w postaci $(x, A)_O$, omówionej w punkcie 9. Jej część liniowa A jest operatorem ortogonalnym, jednoznacznie przez tę izometrię wyznaczonym, ale wektor translacji x zależy od wyboru punktu odniesienia O (tw. 9). Wykażemy, że przy odpowiednim doborze punktu O postać ta przyjmuje wyróżnioną formę. Poprzedzimy twierdzenie prostym rezultatem pomocniczym.

Lemat 15. *Jeśli B jest operatorem normalnym w przestrzeni euklidesowej, to*

$$\operatorname{Ker} B = (\operatorname{Im} B)^{\perp} .$$

Dowód. Na mocy punktu (i) lematu 1, §17, mamy $\operatorname{Ker} B = \operatorname{Ker} B^*$. Mamy ponadto ciąg równoważności

$$B^*x = 0 \Longleftrightarrow \forall y : (B^*x, y) = 0 \Longleftrightarrow \forall y : (x, By) = 0 \Longleftrightarrow x \perp \operatorname{Im} B\,,$$

czyli $\operatorname{Ker} B^* = (\operatorname{Im} B)^\perp$. Łącząc obie równości dostajemy tezę. □

Twierdzenie 16. *Dla każdej izometrii wewnętrznej f przestrzeni euklidesowej $\mathcal{M}$ nad przestrzenią V istnieje dokładnie jeden wektor $x \in V$ taki, że*

$$Ax = x \qquad \textit{oraz} \qquad f = (x, A)_O$$

dla pewnego punktu odniesienia O. Wszystkie punkty odniesienia, dla których spełniony jest powyższy warunek, tworzą podprzestrzeń afiniczną $O+\operatorname{Ker}(A-\mathrm{id})$.

Dowód. Wybierzmy dowolny punkt odniesienia O' i oznaczmy $x' = f(O') - O'$, więc wtedy $f = (x', A)_{O'}$. Dla dowolnego innego punktu odniesienia $O' + a$ na podstawie punktu (iii) twierdzenia 9 mamy

$$f = (x' + (A - \mathrm{id})a, A)_{O'+a}\,.$$

Aby spełnić warunki tezy szukamy wszystkich takich wektorów a, że

$$x' + (A - \mathrm{id})a \in \operatorname{Ker}(A - \mathrm{id}) = \big(\operatorname{Im}(A - \mathrm{id})\big)^\perp\,,$$

gdzie równość podprzestrzeni obowiązuje na mocy lematu. Dokonajmy ortogonalnego (jednoznacznego) rozkładu wektora x':

$$x' = x + y\,, \quad x \in \big(\operatorname{Im}(A - \mathrm{id})\big)^\perp\,, \quad y \in \operatorname{Im}(A - \mathrm{id})\,.$$

Wektor a spełnia więc powyższy warunek wtedy, i tylko wtedy, gdy jest rozwiązaniem równania

$$(A - \mathrm{id})a = -y\,.$$

Równanie to ma rozwiązania (bo y leży w obrazie operatora $A - \mathrm{id}$) i jeśli a' jest jego szczególnym rozwiązaniem, to wszystkie rozwiązania tworzą zbiór $a' + \operatorname{Ker}(A - \mathrm{id})$. Wszystkie szukane punkty odniesienia tworzą więc zbiór $O' + a' + \operatorname{Ker}(A - \mathrm{id}) \equiv O + \operatorname{Ker}(A - \mathrm{id})$, w zgodności z tezą. Dla każdego z tych punktów $x' + (A - \mathrm{id})a = x$, co dowodzi jednoznaczności wektora x i kończy dowód. □

Zastosowanie twierdzenia do przypadku przestrzeni trójwymiarowej daje następujący ciekawy wynik.

Wniosek 17. *Każda zachowująca orientację izometria wewnętrzna trójwymiarowej afinicznej przestrzeni euklidesowej jest złożeniem obrotu wokół pewnej osi z translacją wzdłuż tej samej osi.*

16 Transformacje pól tensorowych przy czynnej transformacji przestrzeni afinicznej

Jeśli na przestrzeni afinicznej, którą poddajemy wewnętrznej transformacji afinicznej, określone jest pole tensorowe (p. 15, §21), to musimy również określić pewien sposób transformowania się tego pola, wymuszony tym odwzorowaniem. Jest to więc problem wyboru najbardziej naturalnej definicji takiej transformacji.

Niech f będzie automorfizmem afinicznym i dla każdego punktu X oznaczmy $\tilde{X} = f(X)$. Niech będzie dane pole tensorowe $\mathcal{M} \ni X \mapsto t(X) \in \mathcal{T}^p_q(V)$, a przez $\tilde{t}(.)$ oznaczmy szukane pole przetransformowane. Przyjmiemy naturalne założenie, że pole przetransformowane ma tę samą walencję, oraz że jego wartość w punkcie $\tilde{X}$, czyli tensor $\tilde{t}(\tilde{X})$, jest wyznaczona wartością pola pierwotnego w punkcie X, a więc tensorem $t(X)$. Dalej, zgodnie z duchem algebry liniowej przyjmiemy, że odwzorowanie $t(X) \mapsto \tilde{t}(\tilde{X})$ jest liniowym odwzorowaniem niezależnym od punktu. Sprowadziliśmy więc zadanie do określenia tego liniowego operatora w przestrzeni $\mathcal{T}^p_q(V)$.

Jeśli $p = q = 0$, czyli $\mathcal{T}^0_0 = \mathbb{K}$, to przyjmujemy naturalne założenie, że ten operator jest identycznością, a więc

$$\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]: \qquad \tilde{t}(\tilde{X}) = t(X)\,.$$

W przypadku pola wektorowego kierujemy się następującym rozumowaniem: jeśli w punkcie X wartość pola jest równa $t(X) = \overrightarrow{XY}$ dla pewnego punktu Y, to w punkcie $\tilde{X}$ powinniśmy mieć $\tilde{t}(\tilde{X}) = \overrightarrow{\tilde{X}\tilde{Y}}$. Oznacza to, że jeśli część liniowa odwzorowania f jest równa A, to przyjmujemy

$$\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]: \qquad \tilde{t}(\tilde{X}) = A\,t(X)\,.$$

Niech teraz $t(.)$ będzie polem 1-form, a $s(.)$ dowolnym polem wektorowym. Ich zwężenie jest polem skalarnym, więc żądamy

$$\langle \tilde{t}(\tilde{X}), \tilde{s}(\tilde{X})\rangle = \langle t(X), s(X)\rangle\,,$$

gdzie $\langle .\,,.\rangle$ jest funkcją dualności. Kładąc za $s(X)$ dowolny wektor łatwo pokazać, że dostaje się stąd warunek

$$\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]: \qquad \tilde{t}(\tilde{X}) = (A^{-1})^{\mathrm{T}}\,t(X)\,,$$

gdzie $(A^{-1})^{\mathrm{T}}$ jest operatorem transponowanym do operatora A^{-1} (punkt 18 oraz przykład (ii) w punkcie 20, §18). Dowolny tensor w punkcie jest sumą iloczynów tensorowych wektorów i form, przyjmujemy więc naturalną postać transformacji:

$$\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]: \qquad \tilde{t}(\tilde{X}) = \underbrace{A \otimes \ldots \otimes A}_{p \text{ razy}} \otimes \underbrace{(A^{-1})^{\mathrm{T}} \otimes \ldots \otimes (A^{-1})^{\mathrm{T}}}_{q \text{ razy}} \, t(X)\,.$$

Wybierzmy afiniczny układ odniesienia $(O, (e_i))$ i niech $f = (a, A)_O$. Wtedy transformacja $t \mapsto \tilde{t}$ zapisana dla współrzędnych pól ma postać

$$\tilde{t}^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}(\tilde{\mathbf{X}}) = A^{i_1}{}_{k_1} \ldots A^{i_p}{}_{k_p} A^{-1 l_1}{}_{j_1} \ldots A^{-1 l_q}{}_{j_q} t^{k_1 \ldots k_p}{}_{l_1 \ldots l_q}\big(\mathbf{A}^{-1}(\tilde{\mathbf{X}} - \mathbf{a})\big)\,.$$

Porównując ten przepis z transformacją współrzędnych pola tensorowego przy zmianie układu odniesienia (p. 15, §21) widzimy, że stwierdzenie zawarte we wniosku 12 rozszerza się do współrzędnych pól tensorowych:

Wniosek 18. *Zmiana współrzędnych pól tensorowych przy czynnej transformacji przestrzeni afinicznej jest taka sama, jak przy odwrotnej do niej transformacji biernej.*

Otrzymana wyżej transformacja $t \mapsto \tilde{t}$ jest zadana przez automorfizm (f, A), więc oznaczamy ją bardziej jawnie jako $T_f t = \tilde{t}$. Dla pola tensorowego t o walencji $\left[\begin{smallmatrix} p \\ q \end{smallmatrix}\right]$ mamy więc:

$$[T_f t](X) = A^{\otimes p} \otimes [(A^{-1})^{\mathrm{T}}]^{\otimes q}\, t\big(f^{-1}(X)\big)\,.$$

Twierdzenie 19. *Niech (f, A) będzie automorfizmem przestrzeni afinicznej $\mathcal{M}$, a T_f odpowiadającą mu transformacją pól tensorowych. Wtedy:*

(i) transformacja T_f zachowuje operacje tensorowe, tj.

$$T_f(\alpha t + \beta s) = \alpha T_f t + \beta T_f s\,, \quad T_f(t \otimes u) = (T_f t) \otimes (T_f u)\,,$$
$$T_f(Ct) = C(T_f t)\,,$$

gdzie C jest dowolną kontrakcją;

(ii) tensor δ jest inwariantem transformacji T_f,

$$T_f \delta = \delta\,;$$

(iii) transformacja T_f zachowuje pochodną, tj.

$$T_f(\partial t) = \partial(T_f t)\,;$$

(iv) odwzorowanie $f \mapsto T_f$ jest homomorficzne, tj.

$$T_{g\circ f} = T_g T_f \,.$$

Dowód. Wszystkie własności można sprawdzić bezpośrednio z definicji T_f lub skorzystać z poprzedzającej twierdzenie relacji dla współrzędnych oraz z towarzyszącego jej wniosku. Wykażemy tą drugą metodą punkt (iii), pozostawiając pozostałe punkty jako ćwiczenie. Oznaczmy $\mathbf{X} = \mathbf{A}^{-1}(\tilde{\mathbf{X}} - \mathbf{a})$. Wtedy

$$\begin{aligned}
\widetilde{(\partial t)}^{i_1\dots i_p}{}_{j_0\dots j_q}(\tilde{\mathbf{X}}) &= A^{i_1}{}_{k_1}\dots A^{-1l_0}{}_{j_0}\dots(\partial t)^{k_1\dots k_p}{}_{l_0\dots l_q}(\mathbf{X}) = \\
= A^{i_1}{}_{k_1}\dots A^{-1l_0}{}_{j_0}\dots \frac{\partial}{\partial X^{l_0}} t^{k_1\dots k_p}{}_{l_1\dots l_q}(\mathbf{X}) &= A^{-1l_0}{}_{j_0}\frac{\partial}{\partial X^{l_0}}\tilde{t}^{i_1\dots i_p}{}_{j_1\dots j_q}(\tilde{\mathbf{X}}) = \\
&= \frac{\partial}{\partial \tilde{X}^{j_0}}\tilde{t}^{i_1\dots i_p}{}_{j_1\dots j_q}(\tilde{\mathbf{X}}) = (\partial\tilde{t})^{i_1\dots i_p}{}_{j_0\dots j_q}(\tilde{\mathbf{X}})\,,
\end{aligned}$$

co kończy dowód własności (iii). W przedostatnim kroku skorzystaliśmy z zamiany zmiennych w pochodnej cząstkowej:

$$\frac{\partial}{\partial\tilde{X}^{j_0}} = \frac{\partial X^{l_0}}{\partial\tilde{X}^{j_0}}\frac{\partial}{\partial X^{l_0}} = A^{-1l_0}{}_{j_0}\frac{\partial}{\partial X^{l_0}}\,.$$

□

Wniosek 18 wskazuje na pewną symetrię między transformacjami czynnymi i biernymi. Nie należy jednak tej symetrii brać zbyt dosłownie: podkreślmy, że transformacja bierna jednoznacznie wynika z definicji tensora i opisuje jedynie odniesienie *tego samego* pola tensorowego do nowego układu afinicznego, podczas gdy transformacja czynna jest zdefiniowanym w tym punkcie odwzorowaniem prowadzącym do *nowego* pola tensorowego. W istocie, jak przekonamy się wkrótce, naturalnym jest niekiedy nieznaczne zmodyfikowanie definicji tego odwzorowania.

17 Przykłady

(i) Przenoszenie pól przy ogólnych odwzorowaniach afinicznych
Niech $(\mathcal{M}, V)$ i $(\mathcal{N}, W)$ będą przestrzeniami afinicznymi nad tym samym ciałem $\mathbb{K}$, a $f : \mathcal{M} \mapsto \mathcal{N}$ niech będzie odwzorowaniem afinicznym, z częścią liniową A. Pytamy, czy istnieje uogólnienie otrzymanej w ostatnim punkcie transformacji pól tensorowych, pozwalające powiązać pola na tych dwóch różnych, w ogólności, przestrzeniach.

Jeśli f jest odwzorowaniem bijektywnym, to A jest izomorfizmem, i definicja odwzorowania T_f może być przeniesiona na ten ogólniejszy przypadek bez zmian. W przeciwnym przypadku nie jest to możliwe dla wszystkich pól tensorowych. Zapiszmy jednak otrzymany w ostatnim punkcie związek pól t i $\tilde{t}$ w postaci:

$$\underbrace{\mathrm{id}\otimes\ldots\otimes\mathrm{id}}_{p\ \text{razy}}\otimes\underbrace{A^{\mathrm{T}}\otimes\ldots\otimes A^{\mathrm{T}}}_{q\ \text{razy}}\ \tilde{t}(\tilde{X}) = \underbrace{A\otimes\ldots\otimes A}_{p\ \text{razy}}\otimes\underbrace{\mathrm{id}\otimes\ldots\otimes\mathrm{id}}_{q\ \text{razy}}\ t(X)\,.$$

W ogólności równanie to nie zadaje zależności funkcyjnej. Istnieją jednak dwa szczególne przypadki, gdy taka zależność istnieje. Jeśli $\tilde{t}(.)$ jest polem na przestrzeni $\mathcal{N}$ o wartościach tensorowych w $\mathcal{T}^0_q(W)$, to transformacja

$$[S_f\tilde{t}](X) = (A^{\mathrm{T}})^{\otimes q}\,\tilde{t}\big(f(X)\big)$$

określa pole tensorów $\mathcal{T}^0_q(V)$ na przestrzeni $\mathcal{M}$. Tę transformację nazywa się ***cofnięciem*** pola tensorowego (kowariantnego). Jeśli natomiast $t(.)$ jest polem na $\mathcal{M}$ tensorów $\mathcal{T}^p_0(V)$, a transformacja f jest *injektywna*, to transformacja

$$[T_f t]\big(f(X)\big) = A^{\otimes p}\,t(X)$$

określa pole tensorów $\mathcal{T}^p_0(W)$ na podprzestrzeni $f(\mathcal{M}) \subseteq \mathcal{N}$. Tę transformację nazywa się ***pchnięciem*** pola tensorowego (kontrawariantnego). Założenie o injektywności jest konieczne dla konsystencji definicji pola na $f(\mathcal{M})$ – zrezygnować z niego można tylko wtedy, gdy rozważamy rezultat pchnięcia w zależności od początkowego punktu X.

Zwróćmy uwagę, że określone wyżej transformacje działają w różnych kierunkach.

(ii) Pola tensorowe nad podprzestrzenią przestrzeni afinicznej
Niech przestrzeń afiniczna $(\mathcal{M}, V)$ będzie podprzestrzenią przestrzeni $(\mathcal{N}, W)$. Rozważmy odwzorowanie identyczności na $\mathcal{M}$ i rozszerzmy jego przeciwdziedzinę do $\mathcal{N}$. Otrzymujemy kanoniczne odwzorowanie $\iota : \mathcal{M} \mapsto \mathcal{N}$, $X \mapsto X$. Część liniowa tego odwzorowania jest kanonicznym odwzorowaniem $V \mapsto W$, $y \mapsto y$, a jej transpozycja – kanonicznym odwzorowaniem $W^* \mapsto V^*$, $\varphi \mapsto \varphi$. Dla odwzorowań cofnięcia i pchnięcia dostajemy więc odpowiednio utożsamienia:

$$[S_\iota\tilde{t}](X) = \tilde{t}(X)\,,\ X \in \mathcal{M}\,; \qquad [T_\iota t](Y) = t(Y)\,,\ Y \in \mathcal{M} \subseteq \mathcal{N}\,.$$

18 Symetrie modeli afinicznych

Jedną z kategorii charakteryzujących teorie czy też modele matematyczne, istotną zarówno dla ich postaci czystej, jak i dla zastosowań, jest symetria. Nie ma potrzeby precyzyjnego definiowania tego pojęcia na poziomie ogólnym; wystarczy przyjąć, że o symetrii mówimy wtedy, gdy pewne charakterystyki czy struktury matematyczne nie ulegają zmianie po poddaniu obiektów występujących w opisie matematycznym działaniu danej odwracalnej transformacji. Jeśli określony zbiór transformacji symetrii tworzy grupę, to mówimy o grupie symetrii.

Na strukturę modeli matematycznych mających zastosowanie w innych naukach składają się zwykle pewne elementy stałe, zadające ich ramy, oraz obiekty matematyczne wybierane z określonych zbiorów za pomocą równań i warunków. Od symetrii takiego modelu żądamy spełnienia dwóch wymagań: zachowania elementów stałych modelu oraz transformowania rozwiązań równań modelu w *inne*, na ogół, rozwiązania tych samych równań.

Zastosujemy wyłożoną ideę symetrii do modeli tensorowych bazujących na geometrii afinicznej. Elementami stałymi takich modeli są: struktura afiniczna oraz zadane dodatkowe wyposażenie przestrzeni, takie jak metryka czy orientacja. Obiektami wybieranymi jako rozwiązania równań są pola tensorowe określonych typów. Rozważymy najpierw symetrie stałych elementów.

19 Symetrie przestrzeni z iloczynem skalarnym

Jeśli jedynym elementem stałym modelu jest przestrzeń afiniczna $(\mathcal{M}, V)$, bez żadnej dodatkowej struktury, to jej dowolny automorfizm afiniczny stanowi jej symetrię, a zbiór tych transformacji tworzy grupę symetrii $AF(\mathcal{M})$. Załóżmy teraz, że oprócz struktury afinicznej istotnym elementem modelu jest również wybrana transformacja dualności p-wektorów i q-form nad V, czyli wybrana n-forma ω $(n = \dim V)$. Przyjmijmy naturalny, określony w punkcie 16, sposób transformowania się tej formy, jako stałego pola, pod działaniem automorfizmu afinicznego (f, A). Korzystając z zapisu tej transformacji w dowolnej bazie otrzymujemy natychmiast:

$$T_f\omega = (\det A)^{-1}\omega\,.$$

Jeśli więc forma ω transformująca się za pomocą odwzorowania T_f ma być niezmiennym elementem modelu, to grupa symetrii zawęża się do afinicznego rozszerzenia grupy $SL(V)$. Istnieje jednak inna możliwość. Określmy mianowicie nową transformację pól tensorowych pod działaniem automorfizmu (f, A), kładąc

$$T^{\mathrm{d}}_f = (\det A)\,T_f\,.$$

Łatwo sprawdzić, że odwzorowanie $f \mapsto T^{\mathrm{d}}_f$ również, podobnie jak T_f, jest homomorficzne:

$$T^{\mathrm{d}}_{g\circ f} = T^{\mathrm{d}}_g\,T^{\mathrm{d}}_f\,,$$

więc zachowuje własności grupowe automorfizmów. Ponadto transformacje T_f^{d} zachowują kombinacje liniowe i kontrakcje oraz pochodną:

$$T_f^{\mathrm{d}}(\alpha t + \beta s) = \alpha T_f^{\mathrm{d}} t + \beta T_f^{\mathrm{d}} s\,, \quad T_f^{\mathrm{d}}(Ct) = C(T_f^{\mathrm{d}} t)\,, \quad T_f^{\mathrm{d}}(\partial t) = \partial(T_f^{\mathrm{d}} t)\,,$$

ale stosowanie ich do iloczynu tensorowego ulega modyfikacji:

$$T_f^{\mathrm{d}}(t \otimes u) = (T_f^{\mathrm{d}} t) \otimes (T_f u) = (T_f t) \otimes (T_f^{\mathrm{d}} u)\,,$$

a tensor δ nie jest inwariantem transformacji T_f^{d}. Jeśli przyjąć, że forma ω transformuje się zgodnie z nowym przepisem, to

$$T_f^{\mathrm{d}}\, \omega = \omega\,,$$

więc ω jest zachowywana pod działaniem wszystkich automorfizmów.

Jeśli przestrzeń $\mathcal{M}$ jest wyposażona w zadaną, niezdegenerowaną metrykę biliniową g, to uznamy automorfizm afiniczny (f, A) za jej symetrię, gdy

$$T_f g = g\,, \quad \text{czyli} \quad (A^{-1})^{\mathrm{T}} \otimes (A^{-1})^{\mathrm{T}} g = g\,,$$

co oznacza, że f jest izometrią wewnętrzną przestrzeni $\mathcal{M}$. Zbiór tych symetrii tworzy afiniczne rozszerzenie liniowej grupy izometrii. Metryka kontrawariantna $\hat{g}$ związana jest z metryką g warunkiem $C(\hat{g} \otimes g) = \delta$, gdzie C jest kontrakcją drugich wskaźników. Z własności transformacji T_f mamy więc

$$T_f \hat{g} = \hat{g}\,.$$

Założymy teraz dodatkowo, że przestrzeń $\mathcal{M}$ jest rzeczywista, a metryka g – symetryczna. Wybierzmy jedną z dwu orientacji tej przestrzeni jako orientację dodatnią i oznaczmy przez e antysymetryczną n-formę jednoznacznie związaną z tą orientacją warunkiem $e = n! e_1 \wedge \ldots \wedge e_n$, gdzie $(e_1, \ldots, e_n)$ jest dowolną dodatnio zorientowaną bazą ortonormalną. Przyjmujemy T_f^{d} za transformację tej formy pod działaniem izometrii f, i wtedy, tak jak w ogólnym przypadku,

$$T_f^{\mathrm{d}} e = e\,.$$

Zwróćmy uwagę, że dla izometrii $\det A = \pm 1$, więc teraz T_f^{d} oprócz poprzednich własności spełnia także

$$T_f(t \otimes u) = (T_f^{\mathrm{d}} t) \otimes (T_f^{\mathrm{d}} u)\,.$$

20 Symetrie modeli tensorowych

Uwzględniamy teraz elementy zmienne modeli, określane za pomocą równań. W modelach tensorowych elementami tymi są pewne pola tensorowe, a równania mają postać warunków równości pól utworzonych z tych pól podstawowych. Rozważania przeprowadzimy na przykładzie równań elektrostatyki oraz magnetostatyki. Częścią zadaną modelu jest trójwymiarowa przestrzeń euklidesowa, której symetrie to afiniczne rozszerzenia transformacji ortogonalnych – szczególny przypadek struktury omówionej w ostatnim punkcie. Podstawowe zmienne pola to pole skalarne $\rho(X)$ (gęstość ładunku) oraz pola wektorowe $J(X)$, $E(X)$ i $B(X)$ (odpowiednio: gęstość prądu, pole elektryczne i pole magnetyczne). Równania elektrostatyki i magnetostatyki to:

$$\begin{aligned} &\nabla\cdot E = 4\pi\rho\,, &&\nabla\times E = 0\,,\\ &\nabla\cdot B = 0\,, &&\nabla\times B = 4\pi J\,. \end{aligned}$$

Chcemy sprawdzić, czy symetrie przestrzeni są również symetriami tych równań. Zakładamy, że pod działaniem izometrii (f,A) pola ρ i J transformują się zgodnie z transformacją T_f, co ma związek z ich pochodzeniem od elementarnych ładunków. Przypomnijmy konstrukcję różniczkowych wyrażeń obecnych w tych równaniach. Oznaczmy $C=E$ lub B i zapiszmy w notacji wskaźnikowej:

$$\nabla\cdot C = \nabla_i C^i\,, \quad (\nabla\times C)^i = \hat{g}^{ij} e_{jkl}\hat{g}^{kr}\nabla_r C^l\,.$$

Pierwsze z tych wyrażeń otrzymuje się więc z $\nabla_i C^j$ przez kontrakcję, a drugie przez pomnożenie tensorowe przez $\hat{g}\otimes e\otimes\hat{g}$ i trzy kontrakcje. Korzystając z własności transformacji T_f i T_f^{d} oraz z własności transformacyjnych $\hat{g}$ i e dostajemy więc

$$\begin{aligned} T_f(\nabla\cdot C) &= \nabla\cdot(T_f C)\,, \quad & T_f^{\mathrm{d}}(\nabla\cdot C) &= \nabla\cdot(T_f^{\mathrm{d}} C)\,,\\ T_f(\nabla\times C) &= \nabla\times(T_f^{\mathrm{d}} C)\,, \quad & T_f^{\mathrm{d}}(\nabla\times C) &= \nabla\times(T_f C)\,. \end{aligned}$$

Stosujemy teraz do pierwszego równania elektrostatyki oraz drugiego równania magnetostatyki operację T_f, a do pozostałych równań operację T_f^{d}. To prowadzi do równoważnych równań

$$\begin{aligned} &\nabla\cdot(T_f E) = 4\pi(T_f\rho)\,, \quad \nabla\times(T_f E) = 0\,,\\ &\nabla\cdot(T_f^{\mathrm{d}} B) = 0\,, \quad \nabla\times(T_f^{\mathrm{d}} B) = 4\pi(T_f J)\,. \end{aligned}$$

Widzimy więc że pola ρ, J, E, B są rozwiązaniami równań elektro- i magnetostatyki wtedy, i tylko wtedy, gdy również pola $T_f\rho$, $T_f J$, $T_f E$, $T_f^{\mathrm{d}} B$ są rozwiązaniami tych samych równań. Tak więc, przy tym sposobie określenia transformowania się pól, izometrie są symetriami modelu. Łatwo się też przekonać, że próba zamiany transformacji $T^{\mathrm{d}}_f B$ na $T_f B$ prowadzi do złamania tej symetrii. Ponadto, otrzymana symetria nie jest konieczną własnością modeli tensorowych, pozostawiamy czytelnikowi jako ćwiczenie zaproponowanie równań, dla których nie będzie się dało uzgodnić transformacji pól.

21 Symetrie modeli w przestrzeni Minkowskiego

Rozważmy teraz afiniczną przestrzeń Minkowskiego i przykład teorii zbudowanej na jej bazie. Izometrie tej przestrzeni, zwane ***transformacjami Poincaré***, tworzą grupę $\mathcal{P}$ będącą afinicznym rozszerzeniem grupy Lorentza (dyskutowanej w punkcie 5, §16). Grupa Lorentza rozpada się na cztery podzbiory nie mające spójnego połączenia, a charakteryzowane wartością $\det A$ (± 1) oraz ortochronicznością lub jej brakiem (↑↓). Dla transformacji Lorentza A oznaczmy przez $\operatorname{sgn} A$ liczbę określoną jako

$$\operatorname{sgn} A = \begin{cases} +1 & \text{gdy } A \in \mathcal{L}^{\uparrow}\,, \\ -1 & \text{gdy } A \in \mathcal{L}^{\downarrow}\,. \end{cases}$$

Odwzorowanie $\mathcal{L} \ni A \mapsto \operatorname{sgn} A \in \{+1, -1\}$ jest homomorficzne. Dla transformacji Poincaré (f, A) możemy teraz określić dwie dodatkowe, obok T_f i T^{d}_f, transformacje pól tensorowych:

$$T^{\mathrm{s}}_f := \operatorname{sgn} A\, T_f\,, \qquad T^{\mathrm{sd}}_f := \operatorname{sgn} A\, T^{\mathrm{d}}_f\,.$$

Łatwo przekonać się, że transformacje te spełniają wszystkie własności i równania wypisane w punkcie 19 dla transformacji T^{d}_f, z wyjątkiem tych, które zawierają formy ω lub e. Wszystkie transformacje T_f, T^{d}_f, T^{s}_f i T^{sd}_f mogą być użyte dla określenia symetrii modelu tensorowego.

Rozważymy jako przykład równanie ruchu cząstki naładowanej w polu elektromagnetycznym. Omawialiśmy je w przykładzie (i), p. 17, §21:

$$m\frac{d^2X^i(\tau)}{d\tau^2} = eF^i{}_j\big(X(\tau)\big)\frac{dX^j(\tau)}{d\tau}\,,$$

gdzie τ jest czasem własnym na trajektorii. Poddajmy przestrzeń transformacji izometrycznej (f, A). Krzywa $X(\tau)$ przejdzie wtedy w krzywą $\tilde{X}(\tau) = f\big(X(\tau)\big)$. Wyliczamy wektor styczny do tej krzywej:

$$\begin{aligned} \frac{d\tilde{X}}{d\tau} &= \lim_{\Delta\to 0}\frac{f\big(X(\tau+\Delta)\big) - f\big(X(\tau)\big)}{\Delta} = \\ &= A\lim_{\Delta\to 0}\frac{X(\tau+\Delta) - X(\tau)}{\Delta} = A\frac{dX(\tau)}{d\tau}\,. \end{aligned}$$

Wektor $d\tilde{X}/d\tau$ jest więc unormowanym wektorem czasopodobnym, przy tym jest on skierowany w przyszłość gdy $A \in \mathcal{L}^{\uparrow}$, ale skierowany w przeszłość w przeciwnym przypadku. W tym drugim przypadku należy więc zmienić kierunek wzrastania parametru wzdłuż krzywej. Kładziemy zatem

$$X'(\tau) = \tilde{X}\big((\operatorname{sgn} A)\tau\big)$$

i dostajemy

$$\frac{dX'(\tau)}{d\tau} = \operatorname{sgn} A \; \frac{d\tilde{X}(s)}{ds}\Big|_{s=(\operatorname{sgn} A)\tau} = (\operatorname{sgn} A)A \, \frac{dX(s)}{ds}\Big|_{s=(\operatorname{sgn} A)\tau} .$$

Wektor $dX'(\tau)/d\tau$ jest teraz w każdym przypadku unormowany i skierowany w przyszłość, więc τ jest czasem własnym na $X'(\tau)$. Dla drugiej pochodnej dostajemy

$$\frac{d^2X'(\tau)}{d\tau^2} = \frac{d^2\tilde{X}(s)}{d^2s}\Big|_{s=(\operatorname{sgn} A)\tau} = A \, \frac{d^2X(s)}{d^2s}\Big|_{s=(\operatorname{sgn} A)\tau} .$$

Możemy teraz zbadać symetrię równania ruchu. Działamy na nie operatorem A i korzystamy z tożsamości

$$(T_fF)^i{}_j\big(\tilde{X}(\tau)\big) = A^i{}_k A^{-1l}{}_j F^k{}_l\big(X(\tau)\big) .$$

Dostajemy równoważną postać równania

$$m\frac{d^2\tilde{X}^i(\tau)}{d\tau^2} = e(T_fF)^i{}_j\big(\tilde{X}(\tau)\big)\frac{d\tilde{X}^j(\tau)}{d\tau} .$$

Podstawiamy tutaj $(\operatorname{sgn} A)\tau$ w miejsce τ i przepisujemy równanie przy użyciu $X'(\tau)$. Jeśli przyjąć dla pola F sposób transformowania się:

$$F' = T^{\mathrm{s}}_f F ,$$

to równanie przyjmie postać

$$m\frac{d^2X'^i(\tau)}{d\tau^2} = eF'^i{}_j\big(X'(\tau)\big)\frac{dX'^j(\tau)}{d\tau} .$$

Transformacje Poincaré są więc symetriami równania ruchu, przy założeniu powyższej transformacji pola elektromagnetycznego. (Otrzymana tutaj transformacja pola elektromagnetycznego jest zgodna, przy ograniczeniu się do izometrii przestrzeni $L(e_1, e_2, e_3)$, z transformacjami dla E i B otrzymanymi w punkcie 20. Dla pola B należy przy tym uwzględnić, że powstaje ono przez dualność w przestrzeni $L(e_1, e_2, e_3)$ ze współrzędnych pola F, co można odczytać ze związku podanego w przykładzie (i), p. 17, §21.)

22 Pseudotensory

Tensory i pola tensorowe występujące w modelach i teoriach fizycznych w taki sposób, który wymaga transformowania ich zgodnie z operacjami T^{d}_f, T^{s}_f lub T^{sd}_f, nazywane bywają ***pseudotensorami***. Terminologia ta jest rozpowszechniona w literaturze fizycznej, szczególnie starszej, ale nie jest zbyt dobrze dostosowana do nowoczesnego pojęcia tensora i na ogół nie spotyka się jej w nowszych podręcznikach czystej matematyki. Podkreślmy raz jeszcze: wszystkie obiekty, o których tu mówimy, są *tensorami* w ujęciu przyjętym w niniejszym podręczniku, a wybór jednej z *czynnych* transformacji T_f, T^{d}_f itd. wiąże się z symetriami modeli, w których występują. Wyjaśnijmy, dla uzyskania pełnego obrazu, związek tak rozumianego pojęcia pseudotensora z bardziej tradycyjnym, spotykanym często w podręcznikach fizyki.

Przypomnijmy, że najbardziej naturalną transformacją czynną tensora (pola tensorowego) jest T_f, jako otrzymana przez rozszerzenie reguły transformacyjnej wektorów łączących dwa punkty. Pamiętamy też, że pod działaniem transformacji T_f współrzędne pola tensorowego zmieniają się tak, jak przy *odwrotnej* transformacji biernej, czyli odpowiedniej zmianie afinicznego układu odniesienia. Załóżmy więc, że

$$e'_i = A^{-1}e_i = e_j A^{-1j}{}_i\,, \quad O' = O - a\,,$$

co prowadzi do biernej transformacji współrzędnych

$$t'^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X}') = A^{i_1}{}_{k_1}\ldots A^{-1l_1}{}_{j_1}\ldots t^{k_1\cdots}{}_{l_1\ldots}\big(\mathbf{A}^{-1}(\mathbf{X}'-\mathbf{a})\big)\,,$$

identycznej ze zmianą czynną pod działaniem T_f, $f = (a, A)_O$. W ujęciu klasycznym funkcje $t^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X})$ transformujące się w taki sposób przy powyższej zmianie układu odniesienia określają pole tensorowe. Wypiszmy teraz transformację współrzędnych pola s przy transformacji czynnej T^{d}_f:

$$s'^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X}') = \det A\; A^{i_1}{}_{k_1}\ldots A^{-1l_1}{}_{j_1}\ldots s^{k_1\cdots}{}_{l_1\ldots}\big(\mathbf{A}^{-1}(\mathbf{X}'-\mathbf{a})\big)\,,$$

gdzie $s' = T^{\mathrm{d}}_f s$. W klasycznym ujęciu nadaje się tym związkom taki sam status, jak w poprzednim przypadku, i nazywa polem pseudotensorowym zespół funkcji przypisanych układom afinicznym i transformujących się między układami za pomocą tych relacji. Wadą tej interpretacji jest brak związku takiego układu tablic funkcji z jednym prostym, niezależnym od układu, pojęciem geometrycznym (w przeciwieństwie do poprzedniego przypadku tensora „czystego”). Ograniczenia tej interpretacji ujawniają się jeszcze bardziej, gdy w przypadku przestrzeni Minkowskiego chcemy ją zastosować do pozostałych dwu transformacji, T^{s}_f i T^{sd}_f. Analogiczne do powyższych związki będą w tych przypadkach miały formę odpowiednio:

$$u'^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X}') = \operatorname{sgn} A\ A^{i_1}{}_{k_1} \ldots A^{-1 l_1}{}_{j_1} \ldots u^{k_1\cdots}{}_{l_1\ldots}\big(\mathbf{A}^{-1}(\mathbf{X}' - \mathbf{a})\big)$$

oraz

$$v'^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X}') = \operatorname{sgn} A \det A\ A^{i_1}{}_{k_1} \ldots A^{-1 l_1}{}_{j_1} \ldots v^{k_1\cdots}{}_{l_1\ldots}\big(\mathbf{A}^{-1}(\mathbf{X}' - \mathbf{a})\big)\,,$$

gdzie jednak teraz transformacja A nie jest dowolnym automorfizmem liniowym, lecz musi być wybrana spośród izometrii stowarzyszonej przestrzeni wektorowej (czyli spośród transformacji Lorentza) – dla innych odwzorowań wielkość sgn A traci sens. A więc zinterpretowanie tablic funkcji $u^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X})$ i $v^{i_1\cdots}{}_{j_1\ldots}(\mathbf{X})$ wraz z ich prawami transformacyjnymi jako nowych typów *pseudotensorów* wymaga rezygnacji ze znacznej części inwariancji języka algebry i ograniczenia się tylko do pewnej grupy transformacji (a w konsekwencji także tylko do pewnej klasy baz).

VI

UZUPEŁNIENIA

§24 Dodatkowe konstrukcje algebraiczne w przestrzeniach wektorowych

1 Bazy w przestrzeniach nieskończenie wymiarowych

Omawiając podstawy teorii przestrzeni wektorowych stwierdziliśmy, że każda przestrzeń posiada bazę, oraz wskazaliśmy, że fakt ten jest konsekwencją ogólniejszego twierdzenia 8, §8, którego jednak nie udowodniliśmy w ogólnym przypadku. W tym punkcie podajemy ten dowód. Przeprowadzenie dowodu musimy poprzedzić uzupełnieniem wiadomości z teorii zbiorów.

Rzadko używanym, ale niezbędnym w pewnych dowodach, aksjomatem teorii mnogości jest tak zwany pewnik wyboru. Stwierdza on, że z każdego zbioru dowolnej zadanej rodziny zbiorów można wybrać po jednym elemencie. Pewnik ten jest zbędny jak długo mamy do czynienia wyłącznie ze skończonymi rodzinami, ale w przypadku ogólnym jest niezależny od pozostałych aksjomatów. Jego równoważną postacią, na gruncie pozostałych aksjomatów, jest lemat Kuratowskiego-Zorna, którym posłużymy się w zapowiadanym dowodzie. Równoważności nie będziemy dowodzić, musimy natomiast rozszerzyć zakres pojęć dla jego sformułowania.

Określiliśmy w punkcie 4, §2, relację częściowego porządku oraz relację porządku. Niech relacja $\preceq$ będzie częściowym porządkiem w przestrzeni X (związek $x \preceq y$ będziemy czytać: x poprzedza y). Mówimy, że podzbiór $Y \subseteq X$ jest ***łańcuchem*** w X, jeśli Y jest zbiorem uporządkowanym relacją $\preceq$ (tj. spośród każdych dwóch elementów zbioru Y jeden poprzedza drugi). Element $y_0 \in X$ nazywamy ***majorantą*** zbioru Y w X, jeśli wszystkie elementy zbioru Y poprzedzają y_0. Jeśli $m \in X$ nie poprzedza żadnego elementu przestrzeni X z wyjątkiem siebie samego, to m nazywamy ***elementem maksymal-***

nym w X. W ogólnym przypadku przestrzeń nie musi posiadać elementu maksymalnego ani łańcuch w tej przestrzeni nie musi mieć majoranty. Z drugiej strony, jeśli takie elementy istnieją, to nie muszą być jedyne. Następujące twierdzenie określa szczególny przypadek.

Twierdzenie 1 (Lemat Kuratowskiego-Zorna)**.** *Jeśli każdy niepusty łańcuch w przestrzeni częściowo uporządkowanej X ma majorantę, to X posiada element maksymalny.*

Dowód twierdzenia 8, §8. Niech X będzie przestrzenią wszystkich podzbiorów $S \subseteq A$ złożonych z wektorów liniowo niezależnych, które zawierają F, $F \subseteq S$. Relacja zawierania jest częściowym porządkiem w tej przestrzeni. Niech Y będzie dowolnym niepustym łańcuchem w X. Suma mnogościowa $U \subseteq V$ wszystkich zbiorów będących elementami Y zawiera oczywiście każdy zbiór rodziny Y, a ponadto, co wykażemy poniżej, U jest rodziną wektorów liniowo niezależnych, czyli $U \in X$. Stąd U jest majorantą łańcucha Y, więc na mocy lematu Kuratowskiego-Zorna istnieje element maksymalny $E \in X$. Oznacza to, że wektory rodziny E są liniowo niezależne, natomiast dla dowolnego $x \in A$ wektory rodziny $E \cup \{x\}$ nie są liniowo niezależne. Stąd na mocy twierdzenia 5, §8, rodzina E generuje A, a zatem generuje również V. Rodzina E jest bazą V spełniającą warunki tezy.

Aby uzupełnić brakujący krok dowodu wybierzmy skończony podzbiór $\{x_1, \ldots, x_k\} \subseteq U$. Oznacza to, że istnieją zbiory $S_i \in Y$, $i = 1, \ldots, k$ takie, że $x_i \in S_i$. Mamy wykazać, że wektory $x_1, \ldots, x_k$ są liniowo niezależne. Ponieważ Y jest łańcuchem, to porównując zbiory $S_1, \ldots, S_k$ parami znajdziemy wśród nich element maksymalny S_{i_0}. Wszystkie wektory x_i są jego elementami, więc są liniowo niezależne, co kończy dowód. □

2 Przestrzeń wektorowa z zadaną bazą

Wiemy już, że każda przestrzeń wektorowa posiada bazę. Konstrukcję, którą przeprowadzimy w tym punkcie, można uważać za rozwiązanie problemu odwrotnego do znalezienia bazy.

Niech E będzie dowolnym zbiorem, a $\mathbb{K}$ dowolnym ciałem. Pokażemy, że istnieje przestrzeń wektorowa V nad ciałem $\mathbb{K}$, której baza może być w naturalny sposób utożsamiona ze zbiorem E (za pomocą bijektywnego odwzorowania).

Przypomnijmy, że zbiór wszystkich funkcji $E \mapsto \mathbb{K}$ tworzy przestrzeń wektorową. Niech V będzie podzbiorem tej przestrzeni złożonym z funkcji znikających prawie wszędzie, to jest poza skończoną liczbą punktów. Łatwo przekonać się, że zbiór ten jest zamknięty względem operacji kombinacji liniowej, więc stanowi przestrzeń wektorową nad $\mathbb{K}$. Dla każdego elementu $e \in E$ określamy wektor e'

przestrzeni V za pomocą przepisu

$$e' : E \mapsto \mathbb{K}\,, \quad e'(x) = \begin{cases} 1 & x = e \\ 0 & x \neq e\,. \end{cases}$$

Zbiór tak otrzymanych wektorów oznaczamy E'. Jest oczywiste, że każda funkcja z przestrzeni V jest jednoznaczną kombinacją liniową wektorów z E'. Zbiór E' jest obrazem zbioru E względem naturalnej bijekcji $E \ni e \mapsto e' \in E'$, więc może być z nim utożsamiony, co kończy zapowiedzianą konstrukcję.

3 Zewnętrzna suma prosta przestrzeni wektorowych

Niech V, V_1 i V_2 będą przestrzeniami wektorowymi nad wspólnym ciałem $\mathbb{K}$. Mówimy, że V jest ***zewnętrzną sumą prostą przestrzeni*** V_1 ***i*** V_2, gdy $V = V_1' \oplus V_2'$ oraz istnieją kanoniczne izomorfizmy $V_i \ni x_i \mapsto x_i' \in V_i'$, $i = 1, 2$.

Łatwo przekonać się, że dla danych przestrzeni V_1 i V_2 wszystkie sposoby skonstruowania ich zewnętrznej sumy prostej dają przestrzenie naturalnie izomorficzne i odpowiednie składowe podprzestrzenie przechodzą w siebie pod działaniem izomorfizmu. Utożsamiając przestrzenie V_i' z przestrzeniami V_i piszemy więc $V = V_1 \oplus V_2$.

Najprostszym sposobem skonstruowania sumy prostej przestrzeni V_1 i V_2 jest przyjęcie, że $V_1 \oplus V_2$ jako zbiór jest iloczynem kartezjańskim $V_1 \times V_2$, na którym zadajemy dodawanie i mnożenie przez elementy ciała za pomocą przepisów:

$$(x_1, x_2) + (y_1, y_2) := (x_1 + x_2, y_1 + y_2)\,,$$
$$\lambda(x_1, x_2) := (\lambda x_1, \lambda x_2)\,.$$

Wektorem zerowym w nowej przestrzeni jest para $(0, 0)$. Sprawdzenie postulatów przestrzeni wektorowej jest prostym ćwiczeniem. Podprzestrzenie V_1' i V_2' są określone przez:

$$V_1' = \{(x_1, 0) \mid x_1 \in V_1\}\,, \quad V_2' = \{(0, x_2) \mid x_2 \in V_2\}\,,$$

a kanoniczne izomorfizmy, o których mowa w definicji, mają postać:

$$V_1 \ni x_1 \mapsto (x_1, 0) \in V_1'\,, \quad V_2 \ni x_2 \mapsto (0, x_2) \in V_2'\,.$$

Elementy zewnętrznej sumy prostej $V_1 \oplus V_2$ notuje się niekiedy jako $x_1 \oplus x_2$. Notacja ta jest szczególnie wygodna, gdy V_i są przestrzeniami euklidesowymi lub unitarnymi, z iloczynami skalarnymi notowanymi jako $(.\,,.)_i$. Przez ich zewnętrzną ortogonalną sumę prostą rozumiemy wtedy przestrzeń $V_1 \oplus V_2$ wyposażoną w dodatnio określony iloczyn skalarny zdefiniowany przez

$$(x_1 \oplus x_2, y_1 \oplus y_2) = (x_1, y_1)_1 + (x_2, y_2)_2\,.$$

Pojęcie sumy prostej uogólnia się w oczywisty sposób na sumę dowolnej skończonej liczby przestrzeni, a także, przy użyciu pojęcia nieskończonej sumy prostej podprzestrzeni (patrz przykład (iii) w punkcie 3, §12), na dowolną nieskończoną rodzinę przestrzeni:

$$V = \bigoplus_{i \in I} V_i \,,$$

gdzie I jest dowolnym zbiorem.

4 Algebry. Algebra tensorów

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$. Załóżmy, że przestrzeń V wyposażono w dodatkowe działanie wewnętrzne, które będziemy nazywać ***mnożeniem*** i oznaczać:

$$V \times V \ni (x, y) \mapsto xy \in V \,.$$

Otrzymaną strukturę $(V, \cdot)$ nazywamy ***algebrą***, jeśli odwzorowanie określające mnożenie jest biliniowe, czyli

$$(\alpha x)y = \alpha(xy) = x(\alpha y) \,, \quad (x + y)z = xz + yz \,, \quad x(y + z) = xy + xz \,.$$

Algebrę nazywamy łączną, gdy mnożenie jest łączne, a ***algebrą z jedynką***, gdy mnożenie posiada element neutralny.

Rozważymy przykład ilustrujący struktury określone w bieżącym i poprzednim punkcie. Dla danej przestrzeni V budujemy zewnętrzną sumę prostą jej przestrzeni tensorowych:

$$\mathcal{T}(V) = \bigoplus_{p,\, q=0}^{\infty} \mathcal{T}^p_q(V) \,.$$

Mnożenie tensorowe rozszerza się w naturalny sposób, przez biliniowość, na przestrzeń $\mathcal{T}(V)$: jeśli

$$T = \bigoplus_{p,q=0}^{\infty} t^p_q \,, \quad S = \bigoplus_{p,q=0}^{\infty} s^p_q \,, \qquad t^p_q, s^p_q \in \mathcal{T}^p_q(V) \,,$$

(tylko skończone liczby wyrazów są różne od zera), to

$$T \otimes S = \bigoplus_{p,q=0}^{\infty} \sum_{\substack{p'+p''=p \\ q'+q''=q}} t^{p'}_{q'} \otimes s^{p''}_{q''} \,,$$

przy czym iloczyn tensorowy w przypadku, gdy przynajmniej jeden czynnik ma walencję $\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right]$, jest mnożeniem tensora przez liczbę. Struktura $(\mathcal{T}(V), \otimes)$ jest łączną algebrą z jedynką, którą nazywa się algebrą tensorów. Podobnie określa się przestrzenie

$$\bigwedge(V) = \bigoplus_{p=0}^{\infty} \bigwedge^p(V)\,, \qquad \bigwedge(V^*) = \bigoplus_{q=0}^{\infty} \bigwedge^q(V^*)$$

(jeśli wymiar przestrzeni V jest skończony, to sumy są ograniczone warunkami $p, q \leq \dim V$). Struktury $(\bigwedge(V), \wedge)$, i $(\bigwedge(V^*), \wedge)$, z działaniem $\wedge$ rozszerzonym analogicznie jak w przypadku mnożenia tensorowego, są również łącznymi algebrami z jedynką.

5 Struktura zespolona

Niech V będzie zespoloną przestrzenią wektorową. Ograniczmy zakres liczb, przez które mnożymy wektory, do liczb rzeczywistych. Łatwo sprawdzić, że otrzymujemy w ten sposób rzeczywistą przestrzeń wektorową, która jako zbiór pokrywa się z V. Oznaczymy tę przestrzeń przez $V_{\mathbb{R}}$.

Jeśli $A : V \mapsto V$ jest operatorem liniowym lub antyliniowym na V, to A jest w obu przypadkach operatorem liniowym na $V_{\mathbb{R}}$; w szczególności, id jest operatorem identycznościowym zarówno na V jak i $V_{\mathbb{R}}$. Dla rzeczywistej liczby λ operator λA jest iloczynem operatora A przez skalar λ zarówno w przypadku, gdy A rozpatrujemy jako operator na V, jak i gdy rozważamy go jako operator na $V_{\mathbb{R}}$. Z drugiej strony, operator iA, jako operator na V, jest iloczynem operatora A przez liczbę i, jednak jako operator na $V_{\mathbb{R}}$ jest on liniowo niezależny od A. W szczególności, $(i\,\mathrm{id})$ jest liniowym operatorem na $V_{\mathbb{R}}$, niezależnym od operatora jednostkowego, i spełniającym

$$(i\,\mathrm{id})^2 = -\,\mathrm{id}\ .$$

Stawiamy teraz problem odwrotny: kiedy przestrzeń rzeczywista może być otrzymana w powyższy sposób z pewnej przestrzeni zespolonej? Widzieliśmy, że warunkiem koniecznym jest istnienie operatora, którego kwadrat jest równy $-\,\mathrm{id}$. Okazuje się, że warunek ten jest również wystarczający. Mianowicie, niech teraz V będzie przestrzenią rzeczywistą. Mówimy, że operator liniowy $J : V \mapsto V$ zadaje na niej ***strukturę zespoloną***, gdy

$$J^2 = -\,\mathrm{id}\ .$$

Twierdzenie 2. *Niech operator J zadaje strukturę zespoloną na rzeczywistej przestrzeni V. Oznaczmy przez V_J zbiór V z określonym na nim mnożeniem* *przez liczby zespolone za pomocą przepisu:*

$$(\mu + i\nu)x := (\mu \operatorname{id} + \nu J)x\,, \qquad \mu, \nu \in \mathbb{R}\,,$$

oraz z dodawaniem elementów przejętym od przestrzeni V. Wtedy V_J jest zespoloną przestrzenią wektorową, oraz

$$(V_J)_{\mathbb{R}} = V\,.$$

Dowód. Własności przestrzeni wektorowej nie zawierające skalarów są spełnione na mocy tych samych własności dla przestrzeni V. Prawa rozdzielności mnożenia wektorów przez liczby i dodawania liczb lub wektorów są oczywiste. Wykazujemy jedynie łączność mnożenia przez liczby:

$$\begin{aligned}(\mu + i\nu)[(\alpha + i\beta)x] &= (\mu \operatorname{id} + \nu J)(\alpha \operatorname{id} + \beta J)x = \\ &= [(\mu\alpha - \nu\beta)\operatorname{id} + (\mu\beta + \nu\alpha)J]x = [(\mu + i\nu)(\alpha + i\beta)]x\,,\end{aligned}$$

co zamyka dowód pierwszego stwierdzenia tezy. Drugie stwierdzenie wynika wprost z definicji mnożenia przez liczby w przestrzeni V_J. □

Wprost z definicji mnożenia przez liczby zespolone w przestrzeni V_J mamy $ix = Jx$, co pokazuje, że ostatnie twierdzenie istotnie jest odwróceniem procedury omówionej na początku tego punktu. Należy jednak zwrócić uwagę, że odwrócenie to nie jest jednoznaczne: dla różnych operatorów J i J' zadających na V struktury zespolone, przestrzenie V_J i $V_{J'}$ są różnymi przestrzeniami wektorowymi, choć są równe jako zbiory, a także $(V_J)_{\mathbb{R}} = (V_{J'})_{\mathbb{R}}$.

Załóżmy, że J definiuje na V strukturę zespoloną. Na przestrzeni V_J, jak na każdej przestrzeni zespolonej, można wprowadzić (na wiele sposobów) operator sprzężenia K. W obecnym kontekście definicja operatora sprzężenia ma następującą równoważną postać: K jest operatorem liniowym na V takim, że

$$K^2 = \operatorname{id}\,, \qquad KJ + JK = 0\,,$$

– drugi warunek, jak łatwo sprawdzić, odpowiada za antyliniowość (wrócimy do niego w ogólniejszym kontekście w następnym punkcie). Wprowadzamy operatory $P_K^{\pm} := \frac{1}{2}(\operatorname{id} \pm K)$, które łącznie tworzą rozkład jedności na V, więc

$$V = V_K^+ \oplus V_K^-\,, \quad V_K^{\pm} := P_K^{\pm} V\,.$$

Ponadto $JP_K^{\pm} = P_K^{\mp} J$, więc odwzorowanie $J_K^+ : V_K^+ \mapsto V_K^-$, $J_K^+ x = Jx$, jest izomorfizmem rzeczywistych przestrzeni wektorowych. Wykażemy teraz, że możliwe jest odwrócenie przeprowadzonego wnioskowania.

Twierdzenie 3. *Niech V będzie rzeczywistą przestrzenią wektorową i niech*

$$V = V^+ \oplus V^-$$

będzie rozkładem na sumę prostą izomorficznych podprzestrzeni, oraz niech $J_+ : V^+ \mapsto V^-$ będzie dowolnym izomorfizmem. Definiujemy operatory liniowe J i K na V przepisem:

$$\begin{aligned} \forall x \in V_+ : \qquad & Jx = J_+ x\,, & Kx = x\,, \\ \forall x \in V_- : \qquad & Jx = -(J_+)^{-1} x\,, & Kx = -x\,. \end{aligned}$$

Wtedy J zadaje strukturę zespoloną na V, a K jest sprzężeniem na V_J. Zbiór wektorów E jest bazą przestrzeni rzeczywistej V^+ wtedy, i tylko wtedy, gdy jest bazą zespolonej przestrzeni V_J złożoną z wektorów spełniających warunek $Ke = e$. Wówczas $E \cup JE$ jest bazą rzeczywistej przestrzeni V.

Dowód. Należy sprawdzić własności $K^2 = -J^2 = \mathrm{id}$ i $JK + KJ = 0$ osobno na V^+ i V^-, co jest prostym ćwiczeniem, które pozostawiamy dla czytelnika. Jeśli E jest bazą V^+, to dla każdego $e \in E$ mamy $Ke = e$. Ponadto, wtedy JE jest bazą V^-, bo J odwzorowuje izomorficznie V^+ na V^-. Stąd każdy wektor $x \in V$ ma jednoznaczny rozkład $x = \sum_{i=1}^k (\alpha_i e_i + \beta_i J e_i)$, gdzie $e_i \in E$, $\alpha_i, \beta_i \in \mathbb{R}$. Zgodnie z definicją mnożenia przez liczby w V_J stwierdzenie to jest równoważne istnieniu jednoznacznego rozkładu w V_J: $x = \sum_{i=1}^k (\alpha_i + i\beta_i) e_i$, więc E jest bazą V_J. Odwrotnie, niech E będzie bazą V_J złożoną z wektorów spełniających $Ke = e$. Każdy wektor $x \in V_J$ ma jednoznaczny rozkład $x = \sum_{i=1}^k \lambda_i e_i$, $\lambda_i \in \mathbb{C}$. Jeśli przy tym $x \in V^+$, to $x = Kx = \sum_{i=1}^k \overline{\lambda_i} e_i$. Z jednoznaczności rozkładu wszystkie liczby λ_i muszą być rzeczywiste, więc E jest bazą V^+. □

Jako natychmiastowy wniosek z tego twierdzenia i dyskusji je poprzedzającej otrzymujemy:

Wniosek 4. *Na rzeczywistej przestrzeni o skończonym wymiarze można wprowadzić strukturę zespoloną wtedy, i tylko wtedy, gdy ten wymiar jest liczbą parzystą. Wymiary przestrzeni V i V_J łączy związek* $\dim V_J = \frac{1}{2} \dim V$.

6 Odwzorowania liniowe na przestrzeni ze strukturą zespoloną

Wiemy już z dyskusji przeprowadzonej na początku poprzedniego punktu, że jeśli A jest operatorem liniowym lub antyliniowym na V_J, to jest on operatorem liniowym na V. Twierdzenie odwrotne nie jest prawdziwe, ale zawsze istnieje rozkład dany następującym twierdzeniem.

Twierdzenie 5. *Niech $A : V \mapsto V$ będzie operatorem liniowym.*

(i) A jest operatorem liniowym na V_J wtedy, i tylko wtedy, gdy

$$AJ - JA = 0\,.$$

(ii) A jest operatorem antyliniowym na V_J wtedy, i tylko wtedy, gdy

$$AJ + JA = 0\,.$$

(iii) Istnieje jednoznaczny rozkład $A = A_{J-} + A_{J+}$, gdzie A_{J-} (A_{J+}) jest operatorem liniowym (odpowiednio: antyliniowym) na V_J. Operatory składowe wyrażają się formułami:

$$A_{J-} = \tfrac{1}{2}(A - JAJ)\,, \quad A_{J+} = \tfrac{1}{2}(A + JAJ)\,.$$

Dowód. Operator liniowy A będzie operatorem liniowym (lub operatorem antyliniowym) na V_J, gdy spełniona będzie równość $A(ix) = i(Ax)$ (lub odpowiednio: $A(ix) = -i(Ax)$) dla wszystkich wektorów. Na mocy definicji mnożenia wektorów przez liczby zespolone oznacza to, że $AJ = \pm JA$, co dowodzi punktów (i) i (ii). Dla dowodu punktu (iii) załóżmy najpierw, że podany rozkład istnieje. Wtedy na mocy punktów (i) i (ii) mamy $JAJ = -A_{J-} + A_{J+}$, a stąd łatwo otrzymać formuły na operatory $A_{J\mp}$. Pozostaje dowieść, że operator A_{J-} jest liniowy, a operator A_{J+} – antyliniowy, na V_J. Ten fakt wynika jednak bezpośrednio z warunku definicyjnego J, gdyż $2A_{J\mp}J = AJ \pm JA = \pm 2JA_{J\mp}$. □

Niech A będzie dowolnym operatorem liniowym na skończenie wymiarowej przestrzeni V z zadaną strukturą zespoloną J i wybranym sprzężeniem K na V_J. Niech dalej rozkład $V = V^+ \oplus V^-$ oraz baza $(e_1, \ldots, e_n)$ przestrzeni V_J będą związane z tymi operatorami tak, jak w twierdzeniu 3. Wtedy macierze operatorów J i A w bazie $(e_1, \ldots, e_n, Je_1, \ldots, Je_n)$ mają postać blokową

$$\mathbf{J} = \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}\,, \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_{++} & \mathbf{A}_{+-} \\ \mathbf{A}_{-+} & \mathbf{A}_{--} \end{pmatrix}\,, \quad \text{więc} \quad \mathbf{JAJ} = \begin{pmatrix} -\mathbf{A}_{--} & \mathbf{A}_{-+} \\ \mathbf{A}_{+-} & -\mathbf{A}_{++} \end{pmatrix}\,.$$

Posługując się tą postacią mamy

$$[A_{J\mp}]e_i = \frac{1}{2}\big(A \mp (JAJ)\big)e_i = e_j\tfrac{1}{2}[(A_{++})^j{}_i \pm (A_{--})^j{}_i] + Je_j\tfrac{1}{2}[(A_{-+})^j{}_i \mp (A_{+-})^j{}_i]\,.$$

Biorąc pod uwagę, że $Je_j = ie_j$ w przestrzeni V_J, otrzymujemy macierze operatorów $A_{J\mp}$ w bazie $(e_1, \ldots, e_n)$:

$$\mathbf{A}_{J\mp} = \frac{1}{2}\Big\{[\mathbf{A}_{++} \pm \mathbf{A}_{--}] + i[\mathbf{A}_{-+} \mp \mathbf{A}_{+-}]\Big\}\,.$$

Operator A jest operatorem liniowym w V_J wtedy, i tylko wtedy, gdy $\mathbf{A}_{--} = \mathbf{A}_{++}$ i $\mathbf{A}_{-+} = -\mathbf{A}_{+-}$ oraz jest operatorem antyliniowym w V_J wtedy, i tylko wtedy, gdy $\mathbf{A}_{--} = -\mathbf{A}_{++}$ i $\mathbf{A}_{-+} = \mathbf{A}_{+-}$.

7 Kompleksyfikacja

Odwołajmy się teraz do analizy dowolnego operatora sprzężenia przeprowadzonej w przykładzie (i), p. 19, §10. Jej wynik możemy teraz ująć tak: operatory sprzężenia K na przestrzeni zespolonej V są we wzajemnie jednoznacznej odpowiedniości z rozkładami $V_{\mathbb{R}} = R \oplus iR$, przy czym R jest zbiorem wektorów $x \in V$ spełniających $Kx = x$. Kompleksyfikacja, którą teraz omówimy, jest operacją odwrotną do takiego rozkładu przestrzeni $V_{\mathbb{R}}$.

Niech V będzie dowolną rzeczywistą przestrzenią wektorową. Tworzymy zewnętrzną sumę prostą $V \oplus V$. Na tej przestrzeni istnieje kanoniczny wybór struktury zespolonej J oraz operatora K definiującego sprzężenie na V_J. Definiujemy:

$$J(x \oplus y) = (-y) \oplus x\,, \quad K(x \oplus y) = x \oplus (-y)\,.$$

Łatwo sprawdzić, że istotnie J i K spełniają odpowiednie warunki oraz podprzestrzenie $(V \oplus V)^+ = V \oplus \{0\}$ i $(V \oplus V)^- = \{0\} \oplus V$ mają takie znaczenie jak w twierdzeniu 3. Zespoloną przestrzeń $(V \oplus V)_J$ oznaczamy $V^{\mathbb{C}}$ i nazywamy ***kompleksyfikacją przestrzeni*** V. Zgodnie z twierdzeniem 3 każda baza przestrzeni V jest również bazą przestrzeni $V^{\mathbb{C}}$, więc przestrzenie te mają równe wymiary. Każdy element $x \oplus y$ przestrzeni $V \oplus V$ można zapisać jako

$$x \oplus y = x \oplus 0 + J(y \oplus 0)\,.$$

Korzystając z kanonicznego izomorfizmu przestrzeni V i podprzestrzeni $V \oplus \{0\}$ możemy więc wektor $x \oplus y$, jako wektor przestrzeni $V^{\mathbb{C}}$, notować:

$$x \oplus y = x + iy\,,$$

co przypomina konwencje wprowadzone dla liczb zespolonych. Analogię rozciągniemy dalej: wektor $z = x + iy$ będziemy nazywać ***zespolonym wektorem***, a sprzężenie K będziemy notować kreską nad symbolem, wprowadzimy również ***część rzeczywistą*** i ***część urojoną wektora***:

$$z = x + iy\,, \quad \overline{z} = x + i(-y) \equiv x - iy\,, \quad \mathfrak{Re}\, z = x\,, \quad \mathfrak{Im}\, z = y\,.$$

Niech V i W będą dowolnymi przestrzeniami rzeczywistymi, a $V^{\mathbb{C}}$ i $W^{\mathbb{C}}$ ich kompleksyfikacjami. Każde liniowe odwzorowanie $A : V \mapsto W$ można rozszerzyć do liniowego odwzorowania $A^{\mathbb{C}} : V^{\mathbb{C}} \mapsto W^{\mathbb{C}}$ za pomocą przepisu:

$$A^{\mathbb{C}}(x + iy) = Ax + iAy\,.$$

Nazwiemy to odwzorowanie ***kompleksyfikacją odwzorowania*** A. Odwrotnie, niech $B : V^{\mathbb{C}} \mapsto W^{\mathbb{C}}$ będzie dowolnym odwzorowaniem liniowym. Definiujemy

część rzeczywistą $\mathfrak{Re}\, B$ oraz ***część urojoną*** $\mathfrak{Im}\, B$ ***odwzorowania*** B jako odwzorowania liniowe $V \mapsto W$ zadane przepisami:

$$(\mathfrak{Re}\, B)x = \mathfrak{Re}(Bx)\,, \quad (\mathfrak{Im}\, B)x = \mathfrak{Im}(Bx)\,, \quad x \in V\,.$$

Korzystając z liniowości B na $V^{\mathbb{C}}$ mamy $B(x+iy) = Bx+iBy$, więc przy użyciu powyższych definicji dostajemy

$$B(x+iy) = \big[(\mathfrak{Re}\, B)x - (\mathfrak{Im}\, B)y\big] + i\big[(\mathfrak{Re}\, B)y + (\mathfrak{Im}\, B)x\big]\,.$$

Do wyliczenia części rzeczywistej i urojonej wektora Bz stosuje się regułę przypominającą mnożenie liczb zespolonych. Z tej reguły łatwo też odczytać, że odwzorowanie liniowe B jest kompleksyfikacją pewnego odwzorowania $A : V \mapsto W$ wtedy, i tylko wtedy, gdy $\mathfrak{Im}\, B = 0$, i w tym przypadku $B = (\mathfrak{Re}\, B)^{\mathbb{C}}$. Odwzorowanie spełniające ten warunek nazwiemy ***odwzorowaniem rzeczywistym***.

Wybierzmy w przestrzeniach $V^{\mathbb{C}}$ oraz $W^{\mathbb{C}}$ rzeczywiste bazy $(e_1, \ldots, e_n)$ i $(f_1, \ldots, k_k)$ odpowiednio (a więc bazy przestrzeni V i W) i oznaczmy macierz odwzorowania B w tych bazach przez $\mathbf{B}$. Mamy wtedy

$$Be_i = f_j B^j{}_i = f_j\, \mathfrak{Re}\, B^j{}_i + i f_j\, \mathfrak{Im}\, B^j{}_i\,,$$

a stąd

$$(\mathfrak{Re}\, B)e_i = f_j\, \mathfrak{Re}\, B^j{}_i\,, \quad (\mathfrak{Im}\, B)e_i = f_j\, \mathfrak{Im}\, B^j{}_i\,,$$

czyli $\mathfrak{Re}\, \mathbf{B}$ i $\mathfrak{Im}\, \mathbf{B}$ są odpowiednio macierzami odwzorowań $\mathfrak{Re}\, B$ i $\mathfrak{Im}\, B$ względem pary baz (e_i) i (f_i). Odwzorowanie liniowe $B : V^{\mathbb{C}} \mapsto W^{\mathbb{C}}$ jest więc rzeczywiste wtedy, i tylko wtedy, gdy jego macierz w rzeczywistych bazach jest rzeczywista.

Wprowadźmy teraz dodatkowo do rozważań symetryczną metrykę g w przestrzeni V. Każdą taką metrykę można rozszerzyć jednoznacznie do hermitowskiej metryki $g^{\mathbb{C}}$ na $V^{\mathbb{C}}$ za pomocą warunków:

$$g^{\mathbb{C}}(x_1 + iy_1, x_2 + iy_2) = g(x_1, x_2) + g(y_1, y_2) + i[g(x_1, y_2) - g(y_1, x_2)]\,.$$

Zostawiamy jako ćwiczenie dla czytelnika wykazanie półtoraliniowości i hermitowskości $g^{\mathbb{C}}$. W bazach rzeczywistych macierze obu metryk są identyczne, więc metryki g i $g^{\mathbb{C}}$ mają wspólną sygnaturę. Metrykę $g^{\mathbb{C}}$ nazywamy ***(hermitowską) kompleksyfikacją metryki*** g. Niech metryki będą niezdegenerowane. Jeśli $A \in \mathcal{L}(V, V)$ i $B \in \mathcal{L}(V^{\mathbb{C}}, V^{\mathbb{C}})$, to przez A^* i B^* oznaczamy sprzężenia tych operatorów względem odpowiednich metryk. Korzystając z definicji kompleksyfikacji operatora i metryki dostaje się w tym przypadku

$$g^{\mathbb{C}}\big(x_1 + iy_1, A^{\mathbb{C}}(x_2 + iy_2)\big) = g^{\mathbb{C}}\big((A^*)^{\mathbb{C}}(x_1 + iy_1), x_2 + iy_2\big)\,,$$

czyli

$$\left(A^{\mathbb{C}}\right)^* = \left(A^*\right)^{\mathbb{C}}.$$

Ograniczmy się teraz do metryk dodatnio określonych. Powyższa relacja oznacza zachowywanie pod działaniem kompleksyfikacji własności operatorów względem sprzężenia. Na przykład: operator symetryczny w przestrzeni euklidesowej staje się rzeczywistym operatorem hermitowskim w przestrzeni unitarnej, a operator ortogonalny – rzeczywistym operatorem unitarnym. Zalecamy teraz czytelnikowi przeanalizowanie dyskusji postaci kanonicznej operatorów ortogonalnych przeprowadzonej w punkcie 13, §17, w świetle omówionych struktur (powrócimy do wykorzystania kompleksyfikacji do analizy struktury operatorów liniowych w punktach 6 i 7, §25).

8 Przestrzenie ilorazowe

Niech V będzie przestrzenią nad dowolnym ciałem, a W jej dowolną podprzestrzenią. Przestrzenie V i W są, na mocy definicji przestrzeni wektorowych, grupami abelowymi ze względu na dodawanie wektorów. Przypomnijmy, że w przypadku grup abelowych każda podgrupa jest podgrupą niezmienniczą, więc zbiór ilorazowy V/W ma naturalną strukturę grupy. Elementami tej grupy są zbiory

$$[x]_W = x + W = \{x + z \mid z \in W\},$$

w których rozpoznajemy podprzestrzenie afiniczne przestrzeni V traktowanej jak przestrzeń afiniczna. Oprócz dodawania tych elementów można określić też ich mnożenie przez elementy ciała $\mathbb{K}$, i otrzymujemy wtedy strukturę $(V/W, +, \cdot\,)$ z działaniami określonymi przez

$$[x]_W + [y]_W = [x + y]_W,$$
$$\lambda[x]_W = [\lambda x]_W$$

– łatwo sprawdzić, że również definicja mnożenia przez liczbę jest poprawna (nie zależy od wyboru elementu w klasie $[x]$). Proste rachunki pokazują, że postulaty przestrzeni wektorowej zawierające mnożenie przez liczby są również spełnione dla tak otrzymanej struktury. Przestrzeń V/W jest więc z tymi działaniami przestrzenią wektorową; nazywamy ją ***przestrzenią ilorazową***. Wektorem zerowym tej przestrzeni jest $0 + W = W$, więc mówiąc obrazowo przestrzeń V/W powstaje przez utożsamienie wszystkich elementów podprzestrzeni W. W szczególności $V/\{0\} = V$.

Twierdzenie 6. *Niech V_i $(i = 1, 2)$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$, a $W_i \subseteq V_i$ – ich podprzestrzeniami, i niech ponadto $A \in \mathcal{L}(V_1, V_2)$. Przepis*

$$[A][x]_{W_1} = [Ax]_{W_2}$$

poprawnie definiuje odwzorowanie liniowe $[A] : V_1/W_1 \mapsto V_2/W_2$ *wtedy, i tylko wtedy, gdy* $AW_1 \subseteq W_2$. *Jeśli ten warunek jest spełniony, to*

$$\begin{array}{llcl} (i) & [A] \text{ jest injekcją} & \iff & A^{-1}(W_2) = W_1\,, \\ (ii) & [A] \text{ jest surjekcją} & \iff & V_2 = \operatorname{Im} A + W_2\,. \end{array}$$

Dowód. Koniecznym i wystarczającym warunkiem poprawności określenia odwzorowania $[A]$ jest, by równość $[x]_{W_1} = [y]_{W_1}$ pociągała $[Ax]_{W_2} = [Ay]_{W_2}$. Poprzednik oznacza $x - y \in W_1$, a następnik – $A(x-y) \in W_2$, więc postulowana implikacja to relacja zawierania $AW_1 \subseteq W_2$, co należało wykazać. Jeśli warunek jest spełniony, to odwzorowanie $[A]$ jest liniowe na mocy liniowości A i definicji działań w V_i/W_i. Mamy

$$\operatorname{Ker}[A] = \{[x] \mid Ax \in W_2\} = \{[x] \mid x \in A^{-1}(W_2)\}\,,$$
$$\operatorname{Im}[A] = \{[Ax] \mid x \in V_1\} = \{[y] \mid y \in \operatorname{Im} A\}\,,$$

skąd wynikają punkty (i) i (ii), szczegóły pozostawiamy jako ćwiczenie (przy dowodzie punktu (i) należy wziąć pod uwagę, że warunek konsystencji definicji można zapisać równoważnie jako $W_1 \subseteq A^{-1}(W_2)$, więc wystarczy wykazać inkluzję przeciwną). □

Wniosek 7. *Jeśli* $V = W \oplus W'$, *to odwzorowanie liniowe*

$$W' \ni x' \mapsto [x']_W \in V/W$$

jest izomorfizmem. W szczególności:

$$\dim V = \dim W + \dim(V/W)\,.$$

Dowód. Podstawiamy za A w ostatnim twierdzeniu odwzorowanie $I : W' \mapsto V$, $Ix' = x'$, i określamy $[I] : W' \mapsto V/W$; dalsze szczegóły dyskusji pozostawiamy jako ćwiczenie. □

9 Przykłady

(i) Kanoniczny izomorfizm $V^*/W^{\perp^*} \mapsto W^*$, gdzie $W \subseteq V$
Definiujemy odwzorowanie

$$I : V^* \mapsto W^*\,, \qquad (I\varphi)(x) = \varphi(x) \quad (x \in W)$$

i zauważamy, że $\operatorname{Ker} I = W^{\perp^*}$. Zastosowanie wyniku twierdzenia 6 pokazuje zatem, że odwzorowanie

$$[I] : V^*/W^{\perp^*} \mapsto W^*\,, \quad [I][\varphi] = I\varphi$$

jest poprawnie określone i injektywne. Aby wykazać jego surjektywność, należy wykazać surjektywność odwzorowania I. W tym celu dokonujemy rozkładu $V = W \oplus W'$, gdzie W' jest dowolnie wybraną podprzestrzenią, o której mówi twierdzenie 3, §12. Dla dowolnie wybranej formy $\psi \in W^*$ definiujemy $\psi_V \in V^*$ za pomocą przepisu $\psi_V(x) = \psi(x_W)$, gdzie $x = x_W + x_{W'}$. Stąd dla $x \in W$ mamy $(I\psi_V)(x) = \psi_V(x) = \psi(x)$, więc $I\psi_V = \psi$, co należało dowieść.

(ii) Przestrzeń $\mathcal{L}(V_1/W_1, V_2/W_2)$ jako przestrzeń ilorazowa
Zbiór odwzorowań liniowych $V_1 \mapsto V_2$ spełniających warunek $AW_1 \subseteq W_2$ jest zamknięty ze względu na tworzenie kombinacji liniowych, więc jest podprzestrzenią przestrzeni $\mathcal{L}(V_1, V_2)$; oznaczymy tę podprzestrzeń $\mathcal{L}(V_1, V_2 : W_1 \mapsto W_2)$. Odwzorowanie

$$\mathcal{L}(V_1, V_2 : W_1 \mapsto W_2) \ni A \mapsto [A] \in \mathcal{L}(V_1/W_1, V_2/W_2)\,,$$

gdzie $[A]$ jest odwzorowaniem określonym w twierdzeniu 6, jest odwzorowaniem liniowym. Jego jądrem jest podprzestrzeń $\mathcal{L}(V_1, V_2 : V_1 \mapsto W_2)$, a obrazem cała przeciwdziedzina (dla wykazania tej ostatniej własności można posłużyć się dowolnym uzupełnieniem $V_i = W_i \oplus W_i'$ i skorzystać z wniosku 7). Stąd odwzorowanie

$$\begin{gathered}\mathcal{L}(V_1, V_2 : W_1 \mapsto W_2)/\mathcal{L}(V_1, V_2 : V_1 \mapsto W_2) \mapsto \mathcal{L}(V_1/W_1, V_2/W_2)\,,\\ A + \mathcal{L}(V_1, V_2 : V_1 \mapsto W_2) \mapsto [A]\,,\end{gathered}$$

jest kanonicznym izomorfizmem przestrzeni wektorowych. W szczególności, dla $W_2 = \{0\}$ jest $\mathcal{L}(V_1, V_2 : V_1 \mapsto \{0\}) = \{0\}$, więc dostajemy izomorfizm:

$$\mathcal{L}(V_1, V_2 : W_1 \mapsto \{0\}) \mapsto \mathcal{L}(V_1/W_1, V_2)\,.$$

Jeśli przy tym $A \mapsto [A]$, to $\mathrm{Ker}[A] = \{0\}$ wtedy, i tylko wtedy, gdy $\mathrm{Ker}\, A = W_1$.

(iii) Macierz operatora w przestrzeni ilorazowej
Niech $A \in \mathcal{L}(V, V)$ i niech W będzie podprzestrzenią inwariantną operatora A. Wtedy $[A] : V/W \mapsto V/W$ jest poprawnie określonym operatorem w przestrzeni ilorazowej. Uzupełniamy W do rozkładu $V = W \oplus W'$ i wybieramy bazę przestrzeni V zgodną z tym rozkładem: $(e_1, \dots, e_k)$ jest bazą W, a $(e_{k+1}, \dots, e_n)$ – bazą W'. W przestrzeni V/W mamy wtedy bazę $([e_{k+1}]_W, \dots, [e_n]_W)$. Jeśli $(A^i{}_j)$ jest macierzą operatora A w bazie $(e_1, \dots, e_n)$, to dla $i > k$ mamy

$$[A][e_i]_W = [Ae_i]_W = \sum_{j=k+1}^{n} [e_j]_W A^j{}_i\,.$$

Jeśli oznaczyć przez $[\mathbf{A}]$ macierz operatora $[A]$ w bazie $([e_{k+1}]_W, \ldots, [e_n]_W)$, to macierz operatora A w bazie $(e_1, \ldots, e_n)$ ma postać

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_W & \mathbf{B} \\ 0 & [\mathbf{A}] \end{pmatrix} .$$

10 Uniwersalny model iloczynu tensorowego

Dla przestrzeni wektorowych $V_1, \ldots, V_s$ o skończonym wymiarze, nad dowolnym ciałem $\mathbb{K}$, najbardziej uniwersalnym spośród poznanych przez nas modeli iloczynu tensorowego jest model odwzorowań wieloliniowych. Przy użyciu pojęcia przestrzeni ilorazowej skonstruujemy teraz model stosujący się do *dowolnych* przestrzeni.

Tworzymy zbiór $V_1 \times \ldots \times V_s$. Dzięki dyskusji przeprowadzonej w punkcie 2 wiemy, że istnieje przestrzeń U nad ciałem $\mathbb{K}$, której bazę można utożsamić z tym zbiorem, a więc każdy wektor u przestrzeni U ma jednoznaczne przedstawienie

$$u = \sum_{i=1}^{k} \lambda_i (x_{i1}, \ldots, x_{is}) ,$$

gdzie $(x_{i1}, \ldots, x_{is})$, $i = 1, \ldots, k$, jest różnowartościowym ciągiem elementów iloczynu kartezjańskiego. Oznaczmy przez U_0 zbiór utworzony z wektorów postaci

$$(x_1, \ldots, x_{j-1}, \lambda x_j + \lambda' x_j', x_{j+1}, \ldots, x_s) - \\ -\lambda\,(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_s) - \lambda'\,(x_1, \ldots, x_{j-1}, x_j', x_{j+1}, \ldots, x_s) .$$

Konstruujemy powłokę liniową $L(U_0)$ tego zbioru, definiujemy przestrzeń wektorową $V = U/L(U_0)$ oraz odwzorowanie

$$\otimes : V_1 \times \ldots \times V_s \mapsto V , \quad (x_1, \ldots, x_s) \mapsto x_1 \otimes \ldots \otimes x_s := [(x_1, \ldots, x_s)]_{L(U_0)} .$$

Twierdzenie 8. *Odwzorowanie $\otimes$ określone w dyskusji poprzedzającej twierdzenie spełnia warunki definicyjne iloczynu tensorowego, więc $V = V_1 \otimes \ldots \otimes V_s$.*

Dowód. Z konstrukcji przestrzeni V wynika wieloliniowość odwzorowania $\otimes$ oraz fakt generowania całej przestrzeni V przez iloczyny $x_1 \otimes \ldots \otimes x_s$. Na mocy twierdzenia 7, §18, wystarczy wykazać spełnienie warunku (i) z tego twierdzenia. Wybierzmy dowolną przestrzeń W nad ciałem $\mathbb{K}$; mamy pokazać, że odwzorowanie $\mathcal{L}(V, W) \ni A \mapsto A \circ \otimes \in \mathcal{L}(V_1, \ldots, V_s; W)$ jest izomorficzne. Jeśli $A \circ \otimes = 0$, to $A(x_1 \otimes \ldots \otimes x_s) = 0$ dla wszystkich wektorów iloczynowych. Stąd jądro rozważanego odwzorowania jest trywialne, więc odwzorowanie jest injektywne. Niech teraz $\tilde{A}$ będzie dowolnym odwzorowaniem ze zbioru $\mathcal{L}(V_1, \ldots, V_s; W)$.

Definiujemy odwzorowanie liniowe $A_U : U \mapsto W$ kładąc dla wektorów bazowych $A_U\big((x_1,\dots,x_s)\big) = \tilde{A}(x_1,\dots,x_s)$. Z wieloliniowości odwzorowania $\tilde{A}$ mamy $A_U L(U_0) = \{0\}$, więc na mocy wyników punktu 8 istnieje odwzorowanie:

$$A \equiv [A_U] \in \mathcal{L}(U/L(U_0), W)\,, \quad [A_U][(x_1,\dots,x_s)]_{L(U_0)} = A_U\big((x_1,\dots,x_s)\big)\,,$$
$$\text{czyli} \qquad A(x_1 \otimes \ldots \otimes x_s) = \tilde{A}(x_1,\dots,x_s)\,.$$

Stąd $\tilde{A} = A \circ \otimes$, więc obraz rozważanego odwzorowania jest całą przeciwdziedziną, co kończy dowód jego bijektywności. □

§25 Postać kanoniczna operatora liniowego

1 Rozkład operatora na sumę prostą operatorów o względnie pierwszych wielomianach charakterystycznych

W tym paragrafie zakładamy stale, że wymiar rozważanych przestrzeni jest skończony,

$$\dim V < \infty\,.$$

Badając zagadnienie własne dla operatorów liniowych wyróżniliśmy klasę operatorów, nazwanych operatorami diagonalizowalnymi, które rozkładają się na sumę prostą operatorów proporcjonalnych do operatorów identycznościowych. W ogólnym przypadku taki rozkład nie jest możliwy, ale może być zastąpiony rozkładem prostym

$$A = A_1 \oplus \ldots \oplus A_s$$

na operatory A_i ogólniejszego, ale wyróżnionego typu. Otrzymamy ten rozkład w tym punkcie. W dalszych punktach zajmiemy się wyborem baz, w których te składowe operatory przedstawiają się szczególnie prosto. W rozważaniach tego rozdziału dużą rolę odgrywać będą wielomiany charakterystyczne operatorów. Wygodniej będzie posługiwać się wielomianem $\det(\lambda\,\mathrm{id} - A)$, który jest równy z dokładnością do znaku wielomianowi charakterystycznemu, a dla którego współczynnik przy najwyższej potędze λ jest równy 1.

Przypomnijmy najpierw, że jeśli operator A rozkłada się tak, jak w powyższym wzorze, to wielomian charakterystyczny operatora A rozkłada się na iloczyn wielomianów charakterystycznych operatorów A_i. Jeśli operator A jest diagonalizowalny, to $\det(\lambda\,\mathrm{id}_i - A_i) = (\lambda - \lambda_i)^{n_i}$. Dwumiany $(\lambda - \lambda_i)$ są wielomianami pierwszymi. Z drugiej strony wiemy, że każdy wielomian rozkłada się na potęgi wielomianów pierwszych (choć niekoniecznie dwumiennych), i to jednoznacznie z dokładnością do kolejności czynników i mnożenia ich przez liczby. Powstaje pytanie, czy istnieje taki rozkład operatora A na sumę prostą operatorów A_i, że wielomiany charakterystyczne tych operatorów są odpowiednimi potęgami czynników pierwszych tego rozkładu. Uzyskamy poniżej twierdzącą odpowiedź. Poprzedzimy sformułowanie twierdzenia pomocniczym rezultatem.

Lemat 1. *Niech A będzie operatorem na przestrzeni V nad ciałem $\mathbb{K}$, a $Q(\lambda)$ wielomianem pierwszym o współczynniku przy najwyższej potędze równym* 1. *Jeśli $Q(A)^m = 0$ dla pewnego m, to dla pewnego $k \in \mathbb{N}$:*

$$\det(\lambda \,\mathrm{id} - A) = Q(\lambda)^k \,.$$

Dowód. Wielomiany dwóch zmiennych $\lambda^l - \mu^l$ dzielą się przez $\lambda - \mu$ w każdym ciele (jak w przykładzie (vii), p. 5, §13), więc

$$Q(\lambda)^m - Q(\mu)^m = (\lambda - \mu) W(\lambda, \mu) \,,$$

gdzie $W(\lambda, \mu)$ jest wielomianem w obu argumentach. Po podstawieniu za zmienną μ operatora A, a za zmienną λ operatora $\lambda \,\mathrm{id}$, dostajemy wielomian zmiennej λ o współczynnikach operatorowych. Stąd na mocy założenia mamy

$$Q(\lambda)^m \,\mathrm{id} = (\lambda \,\mathrm{id} - A) W(\lambda \,\mathrm{id}, A) \,,$$

więc dostajemy równość wielomianów nad ciałem:

$$Q(\lambda)^{m \dim V} = \det(\lambda \,\mathrm{id} - A) \det[W(\lambda \,\mathrm{id}, A)] \,.$$

Z jednoznaczności rozkładu wielomianu na wielomiany pierwsze każdy z wielomianów po prawej stronie ostatniej równości musi być proporcjonalny do pewnej potęgi wielomianu $Q(\lambda)$, przy czym współczynnik przy najwyższej potędze jest równy 1 w wielomianie $\det(\lambda \,\mathrm{id} - A)$, co kończy dowód. □

Twierdzenie 2. *Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni wektorowej V nad ciałem $\mathbb{K}$. Niech*

$$\det(\lambda \,\mathrm{id} - A) = Q_1(\lambda)^{n_1} \ldots Q_s(\lambda)^{n_s}$$

będzie rozkładem wielomianu na czynniki pierwsze, przy czym współczynnik przy najwyższej potędze jest w każdym z wielomianów $Q_i(\lambda)$ równy 1. *Wtedy istnieje dokładnie jeden rozkład*

$$A = A_1 \oplus \ldots \oplus A_s$$

taki, że dla każdego $i = 1, \ldots, s$ jest $\det(\lambda \,\mathrm{id} - A_i) = Q_i(\lambda)^{n_i}$. Rozkład ten powstaje przez położenie

$$V_i = \operatorname{Ker} Q_i(A)^{n_i}$$

i określenie A_i jako zawężenie operatora A do operatora na V_i.

Dowód. Pokażemy najpierw, że operatory A_i określone w ostatniej części twierdzenia tworzą rozkład o żądanej własności, a do jego jednoznaczności wrócimy na końcu dowodu. Oznaczmy $P(\lambda) := \det(\lambda \operatorname{id} - A)$ oraz $R_i(\lambda) := \prod_{j \neq i} Q_j(\lambda)^{n_j}$, $i = 1, \ldots, s$. Łatwo sprawdzić następujące własności tych wielomianów

$$Q_i^{n_i} R_i = P\,, \quad R_i R_j = \prod_{k \neq i,j} Q_k^{n_k} P \quad \text{dla} \quad i \neq j\,,$$
$$\sum_i R_i S_i = 1 \quad \text{dla pewnych wielomianów } S_i\,.$$

Trzecia własność wynika z faktu, że wielomiany R_i nie mają wspólnego podzielnika (przykład (iv) w punkcie 15, §3). Wprowadźmy operatory $T_i = R_i(A)S_i(A)$. Ponieważ na mocy twierdzenia Cayley-Hamiltona $P(A) = 0$ (przykład (vi) w punkcie 5, §13), to mamy

$$\sum_i T_i = \operatorname{id}\,, \quad T_i T_j = 0 \quad \text{dla} \quad i \neq j\,,$$

więc operatory T_i tworzą rozkład jedności. Ponieważ A komutuje z T_i, to przestrzenie $V_i = T_i V$ są inwariantne względem operatora A. Zacieśnienie A do operatora na V_i oznaczamy A_i i dostajemy rozkład A na sumę prostą. Musimy pokazać, że $V_i = \operatorname{Ker} Q_i(A)^{n_i}$, oraz że $Q_i(\lambda)^{n_i}$ jest wielomianem charakterystycznym operatora A_i. Na mocy znikania $P(A)$ mamy

$$Q_i(A)^{n_i} T_i = 0\,, \quad \text{więc} \quad V_i \subseteq \operatorname{Ker} Q_i(A)^{n_i}\,.$$

Odwrotnie, jeśli $Q_i(A)^{n_i} x = 0$, to z definicji T_j i R_j wynika, że $T_j x = 0$ dla wszystkich $j \neq i$, więc $x \in V_i$. Pokazaliśmy, że istotnie $V_i = \operatorname{Ker} Q_i(A)^{n_i}$. Dla operatora A_i mamy $Q_i(A_i)^{n_i} = 0$, więc z lematu $\det(\lambda \operatorname{id}_i - A_i) = Q_i(\lambda)^{k_i}$ dla pewnej naturalnej liczby k_i. Ale iloczyn tych wielomianów dla wszystkich operatorów składowych musi dać wielomian $P(\lambda)$, więc $k_i = n_i$.

Pozostała do wykazania jednoznaczność rozkładu. Załóżmy, że mamy dany rozkład spełniający warunek twierdzenia, i oznaczmy podprzestrzeń działania operatora A_i przez V_i. Jeśli $x_i \in V_i$, to $Q_i(A)^{n_i} x = Q_i(A_i)^{n_i} x = 0$ na mocy twierdzenia Cayley-Hamiltona, więc $V_i \subseteq \operatorname{Ker} Q_i(A)^{n_i}$. Ale sumy proste podprzestrzeni po każdej ze stron tej relacji zawierania dają całą przestrzeń V, więc $V_i = \operatorname{Ker} Q_i(A)^{n_i}$, co kończy dowód jednoznaczności. □

Otrzymanie podprzestrzeni rozkładu, o którym mówi ostatnie twierdzenie, i lepsze zrozumienie uzyskanej struktury uzyskamy przy użyciu następującego rezultatu.

Twierdzenie 3. *Dla każdego operatora $B \in \mathcal{L}(V,V)$ podprzestrzenie* $\operatorname{Ker} B^q$*, $q \in \mathbb{N}$, tworzą ciąg niemalejący ze względu na q,*

$$\operatorname{Ker} B^q \subseteq \operatorname{Ker} B^{q+1}\,.$$

Jeśli q_0 jest najmniejszą liczbą naturalną, dla której zachodzi

$$\operatorname{Ker} B^{q_0} = \operatorname{Ker} B^{q_0+1}$$

(taka liczba istnieje dla każdej przestrzeni skończenie wymiarowej), to dla każdego $q \geq q_0$:

$$\operatorname{Ker} B^q = \operatorname{Ker} B^{q_0}\,.$$

Dowód. Jeśli $B^q x = 0$, to również $B^{q+1}x = 0$, co dowodzi pierwszego stwierdzenia. Załóżmy, że istnieje liczba q_0 i wybierzmy dowolną liczbę naturalną $q \geq q_0$. Jeśli wektor x spełnia $B^{q+1}x = 0$, to $B^{q-q_0}x \in \operatorname{Ker} B^{q_0+1} = \operatorname{Ker} B^{q_0}$, więc $B^q x = 0$. Stąd dla każdego $q \geq q_0$ mamy $\operatorname{Ker} B^{q+1} \subseteq \operatorname{Ker} B^q$. Łącząc ten wynik z obowiązującą zawsze inkluzją przeciwną dostajemy drugie stwierdzenie. □

Korzystając z tego wyniku, podprzestrzenie rozkładu operatora A na operatory A_i można więc otrzymać wyliczając ściśle rosnące ciągi podprzestrzeni

$$\operatorname{Ker} Q_i(A), \operatorname{Ker} Q_i(A)^2, \ldots, \operatorname{Ker} Q_i(A)^{q_i} = \operatorname{Ker} Q_i(A)^{q_i+1}\,, \quad i = 1, \ldots, s\,.$$

W dalszych rozważaniach możemy więc ograniczyć się do operatorów, których wielomiany charakterystyczne są potęgami pojedynczego wielomianu pierwszego.

2 Dalszy rozkład operatora na sumę prostą operatorów z wektorem cyklicznym

Pokażemy w tym punkcie, że każdy ze składowych operatorów A_i może być dalej rozłożony na sumę prostą operatorów, z których każdy ma w swojej przestrzeni działania wektor cykliczny (przykład (ii) w punkcie 7, §12). Rozkład ten nie będzie jednoznaczny, ale pewne istotne jego charakterystyki *będą* inwariantne. Ponieważ możemy rozważać każdy z operatorów A_i z osobna, dla uproszczenia notacji oznaczymy rozważany operator przez A.

Załóżmy najpierw, że A jest operatorem w przestrzeni V, z wektorem cyklicznym f, zatem wektory $(f, Af, A^2 f, \ldots, A^{\dim V - 1} f)$ tworzą bazę przestrzeni V.

Jeśli więc wektory $(e_i(\lambda))$ tworzą bazę przestrzeni wielomianów stopnia mniejszego lub równego $\dim V - 1$, to wektory $(e_i(A)f)$ stanowią bazę przestrzeni V. W szczególności, niech ponadto $\det(\lambda\,\mathrm{id} - A) = Q(\lambda)^r$, gdzie Q jest wielomianem stopnia m. Wtedy każdy z następujących dwóch zespołów wektorów tworzy bazę:

$$f, Af, \ldots, A^{rm-1}f\,,$$
$$Q(A)^i A^j f\,, \quad i = 0, \ldots, r-1\,, \quad j = 0, \ldots, m-1\,,$$

– drugi zbiór wielomianów łączy się z pierwszym, po odpowiednim uszeregowaniu, za pomocą nieosobliwej macierzy trójkątnej. Możemy teraz sformułować pomocniczy wynik.

Lemat 4. *Jeśli operator A w przestrzeni V posiada wektor cykliczny f oraz* $\det(\lambda\,\mathrm{id} - A) = Q(\lambda)^r$*, gdzie Q jest wielomianem stopnia m, to dla każdego $p = 1, \ldots, r$ jest*

$$\operatorname{Ker} Q(A)^p = \operatorname{Im} Q(A)^{r-p}$$

oraz wektory $Q(A)^i A^j f$*,* $i = r-p, \ldots, r-1$*,* $j = 0, \ldots, m-1$*, tworzą bazę tej podprzestrzeni, więc*

$$\dim \operatorname{Ker} Q(A)^p = pm\,.$$

Stąd r jest minimalną liczbą, dla której $Q(A)^r = 0$.

Dowód. Działając operatorem $Q(A)^p$ na dowolny wektor rozłożony w bazie i przyrównując rezultat do zera, dostajemy warunek na przynależność wektora do podprzestrzeni $\operatorname{Ker} Q(A)^p$:

$$0 = Q(A)^p \sum_{i=0}^{r-1} \sum_{j=0}^{m-1} \alpha_{ij} Q(A)^i A^j f = \sum_{i=0}^{r-p-1} \sum_{j=0}^{m-1} \alpha_{ij} Q(A)^{i+p} A^j f\,,$$

czyli $\alpha_{ij} = 0$ dla $i = 0 \ldots, r-p-1$, $j = 0, \ldots, m-1$, co pokazuje, że zespół wektorów podany w tezie jest bazą podprzestrzeni $\operatorname{Ker} Q(A)^p$. Ten sam zespół powstaje w wyniku zadziałania na wektory bazowe operatorem $Q(A)^{r-p}$, więc jest również bazą podprzestrzeni $\operatorname{Im} Q(A)^{r-p}$. Pozostałe stwierdzenia tezy są teraz oczywiste. □

Jesteśmy teraz przygotowani do sformułowania i udowodnienia głównego rezultatu tego punktu.

Twierdzenie 5. *Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni V i niech $\det(\lambda\,\mathrm{id}-A) = Q(\lambda)^n$, gdzie Q jest wielomianem pierwszym stopnia m. Wtedy prawdziwe są następujące stwierdzenia.*

(i) Istnieje rozkład

$$V = V_1 \oplus \ldots \oplus V_k\,, \quad A = A_1 \oplus \ldots \oplus A_k$$

taki, że każdy z operatorów A_i posiada wektor cykliczny f_i. Dla każdego rozkładu spełniającego ten warunek jest $\det(\lambda\,\mathrm{id}_i - A_i) = Q(\lambda)^{r_i}$, i jeśli podprzestrzenie są uszeregowane tak, że ciąg liczb r_i jest nierosnący, to ten ciąg nie zależy od rozkładu. Oznaczmy przez $\#(r)$ liczbę wyrazów tego ciągu równych r. Wtedy

$$\begin{aligned} m\,\#(r) &= \dim\operatorname{Ker} Q(A)^r - \dim\operatorname{Ker} Q(A)^{r-1} \\ &\quad - \left(\dim\operatorname{Ker} Q(A)^{r+1} - \dim\operatorname{Ker} Q(A)^r\right). \end{aligned}$$

(ii) Rozkład, o którym mówi punkt (i), z nierosnącym ciągiem (r_i), można skonstruować za pomocą następującej procedury rekurencyjnej. Załóżmy, że znaleziono wektory $f_1, \ldots, f_l$, odpowiednie podprzestrzenie $V_1, \ldots, V_l$ i operatory $A_1, \ldots, A_l$, przy czym nie uzyskano jeszcze rozkładu całej przestrzeni. Wtedy liczba r_{l+1} jest najmniejszą liczbą naturalną spełniającą warunek

$$V = V_1 \oplus \ldots \oplus V_l + \operatorname{Ker} Q(A)^{r_{l+1}}\,,$$

a wektor f_{l+1} może być dowolnie wybrany ze zbioru

$$\Delta_{l+1} \equiv \operatorname{Ker} Q(A)^{r_{l+1}} \setminus \left[V_1 \oplus \ldots \oplus V_l + \operatorname{Ker} Q(A)^{r_{l+1}-1}\right].$$

Podprzestrzeń V_{l+1} jest generowana działaniem potęg operatora A na ten wektor, a operator A_{l+1} jest zawężeniem operatora A do operatora na V_{l+1}.

Dowód. (i) Istnienie rozkładu wynika z konstrukcji, którą udowodnimy w punkcie (ii). Wielomian charakterystyczny operatora A jest iloczynem wielomianów charakterystycznych operatorów A_i, skąd wynika stwierdzenie o postaci tych ostatnich. Jednoznaczność nierosnącego ciągu liczb r_i wynika ze wzoru na $\#(r)$, który teraz wykażemy. Niech dla każdego $i = 1, \ldots, k$ wektor f_i będzie wektorem cyklicznym operatora A_i. Wtedy na podstawie lematu 4 bazę podprzestrzeni $\operatorname{Ker} Q(A)^p$ tworzą wektory

$$\begin{aligned} &Q(A)^i A^j f_l\,, \quad l = 1, \ldots, k\,, \quad j = 0, \ldots, m-1\,, \\ &\qquad i = \begin{cases} 0, \ldots, r_l - 1\,, & \text{gdy } r_l \le p\,, \\ r_l - p, \ldots, r_l - 1\,, & \text{gdy } r_l > p\,. \end{cases} \end{aligned}$$

Zliczając wszystkie wektory w tym zbiorze dostajemy

$$\dim \operatorname{Ker} Q(A)^p = m\biggl(\sum_{r=1}^{p} r\,\#(r) + p \sum_{r=p+1}^{n} \#(r)\biggr)\,,$$

więc $$\dim \operatorname{Ker} Q(A)^p - \dim \operatorname{Ker} Q(A)^{p-1} = m \sum_{r=p}^{n} \#(r)\,,$$

a stąd natychmiast wynika wzór na $\#(r)$.

(ii) Przyjmujemy założenie indukcyjne jak w punkcie (ii) twierdzenia, przy czym włączamy też wartość $l = 0$, co pozwoli zapoczątkować procedurę rekurencyjną. Znajdujemy liczbę r_{l+1} spełniającą podany warunek i wybieramy wektor f_{l+1} (zbiór Δ_{l+1} jest niepusty, gdyż w przeciwnym razie liczba r_{l+1} nie byłaby minimalna). Ponieważ w poprzednim kroku rekurencji uzyskaliśmy pierwszą z poniższych równości

$$V = V_1 \oplus \ldots \oplus V_{l-1} + \operatorname{Ker} Q(A)^{r_l} = V_1 \oplus \ldots \oplus V_l + \operatorname{Ker} Q(A)^{r_l}\,,$$

a druga równość jest wówczas oczywista, to $r_{l+1} \leq r_l$ na mocy konstrukcji. Działając na wektor f_{l+1} potęgami operatora A generujemy podprzestrzeń V_{l+1} i tworzymy zacieśnienie A_{l+1}. Ponieważ $Q(A)^{r_{l+1}} f_{l+1} = 0$ i $Q(A)^{r_{l+1}-1} f_{l+1} \neq 0$, to r_{l+1} jest najmniejszą taką liczbą, że $Q(A_{l+1})^{r_{l+1}} = 0$. Stąd na mocy lematów 1 i 4 mamy $\det(\lambda \operatorname{id}_{l+1} - A_{l+1}) = Q(\lambda)^{r_{l+1}}$.

Ostatni krok dowodu to wykazanie, że podprzestrzenie $V_1 \oplus \ldots \oplus V_l$ i V_{l+1} są liniowo niezależne. Załóżmy, że dla wektorów $x_i \in V_i$, $i = 1, \ldots, l+1$, jest

$$x_1 + \ldots + x_{l+1} = 0\,.$$

Wystarczy pokazać, że $x_{l+1} = 0$. Dowodzimy nie wprost: zakładamy, że $x_{l+1} \neq 0$. Wektor x_{l+1} ma przedstawienie $x_{l+1} = P(A) f_{l+1}$, gdzie P jest wielomianem. Rozłóżmy ten wielomian zgodnie z przepisem $P = P' Q^p$, gdzie P' nie dzieli się przez Q oraz $p < r_{l+1}$ (na mocy konstrukcji f_{l+1} i założenia nie wprost). Wielomiany P' i Q nie mają wspólnego podzielnika, więc dla pewnych wielomianów M i N jest $P'M + QN = 1$. Działając na założone równanie operatorem $Q(A)^{r_{l+1}-p-1} M(A)$ dostajemy więc

$$y_1 + \ldots y_l + Q(A)^{r_{l+1}-1} f_{l+1} = 0\,,$$

gdzie $y_i \in V_i$. Działamy dalej na to równanie operatorem $Q(A)$ i dostajemy

$$Q(A) y_1 + \ldots + Q(A) y_l = 0\,,$$

skąd na mocy liniowej niezależności podprzestrzeni $V_1, \ldots, V_l$ znikać musi każdy ze składników lewej strony oddzielnie. Na mocy lematu 4 mamy stąd

$y_i = Q(A)^{r_i - 1} z_i$, $z_i \in V_i$, więc przedostatnie wyeksponowane równanie możemy teraz zapisać jako

$$Q(A)^{r_{l+1}-1}\big(w + f_{l+1}\big) = 0\,, \quad w \in V_1 \oplus \ldots \oplus V_l\,.$$

Wnioskujemy stąd, że wektor f_{l+1} jest sumą wektora $(-w)$ i wektora z podprzestrzeni $\operatorname{Ker} Q(A)^{r_{l+1}-1}$, co przeczy jego konstrukcji i kończy dowód twierdzenia. □

Łącznym wynikiem twierdzeń 2 i 5 jest rozkład dowolnego operatora A w przestrzeni nad dowolnym ciałem na sumę prostą operatorów,

$$V = V_1 \oplus \ldots \oplus V_t\,, \qquad A = A_1 \oplus \ldots \oplus A_t\,,$$

przy czym każda z podprzestrzeni V_i ma wektor cykliczny f_i względem A, a wielomian charakterystyczny każdego z operatorów A_i jest potęgą wielomianu pierwszego. Wielkość podprzestrzeni działania operatorów składowych i postać ich wielomianów charakterystycznych nie zależą od sposobu uzyskania rozkładu. W każdej bazie zgodnej z uzyskanym rozkładem macierz operatora ma postać blokowo diagonalną. Ostatnim krokiem będzie teraz wybór szczególnych baz w każdej z podprzestrzeni składowych. Istnieje tu kilka możliwości, które kolejno omówimy.

3 Bazy Jordana

W tym punkcie zakładamy, że wielomian charakterystyczny operatora A rozkłada się na czynniki dwumianowe. W szczególności spełnia to założenie każdy operator w przestrzeni zespolonej. Przy tym założeniu wszystkie czynniki pierwsze w wielomianie charakterystycznym mają postać $Q_i(\lambda) = \lambda - \lambda_i$, gdzie λ_i są wartościami własnymi. Podprzestrzeń złożoną ze wszystkich wektorów spełniających $(A - \lambda_i \operatorname{id})^j x = 0$ przy jakiejkolwiek potędze j nazywamy ***uogólnioną podprzestrzenią własną operatora** A **do wartości własnej*** λ_i.

Rozważmy operator składowy A_l na V_l z wektorem cyklicznym f_l i wielomianem charakterystycznym $\det(\lambda \operatorname{id} - A_l) = (\lambda - \lambda_l)^{r_l}$, gdzie λ_l jest odpowiadającą mu wartością własną operatora A (która może być wspólna dla kilku operatorów A_l). W rozważanym przypadku baza wektorów $Q(A)^i A^j f_l$ upraszcza się do $(A - \lambda_l \operatorname{id}_l)^i f_l$, $i = 0, \ldots, r_l - 1$. Oznaczmy

$$e_j = (A - \lambda_l \operatorname{id})^{r_l - j} f_l\,, \qquad j = 1, \ldots, r_l\,.$$

Łatwo sprawdzić, że działanie operatora A_l w bazie $(e_1, \ldots, e_{r_l})$ ma postać

$$A_l e_1 = \lambda_l e_1\,, \qquad A_l e_j = e_{j-1} + \lambda_l e_j\,, \quad j = 2, \ldots, r_l\,,$$

więc macierz operatora A_l w tej bazie ma postać ***klatki Jordana***:

$$\mathbf{A}_l = \begin{pmatrix} \lambda_l & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_l & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_l & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_l & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_l \end{pmatrix} \equiv \lambda_l \mathbf{1} + \mathbf{E}\,.$$

Wybierając w podobny sposób bazy wszystkich podprzestrzeni V_l dostajemy macierz blokowo diagonalną z klatkami Jordana na diagonali. W szczególności: każda macierz zespolona może być sprowadzona nieosobliwą transformacją podobieństwa do postaci jordanowskiej.

4 Uogólnienia baz Jordana

Rozważmy teraz problem rozkładu operatora działającego w przestrzeni nad dowolnym ciałem $\mathbb{K}$, w którym wielomiany pierwsze mogą być stopnia większego niż jeden. Niech A_l będzie składowym operatorem z wektorem cyklicznym f_l i niech $\det(\lambda\,\mathrm{id}_l - A_l) = Q_l(\lambda)^{r_l}$, gdzie Q_l jest wielomianem pierwszym,

$$Q_l(\lambda) = \lambda^m - \sum_{j=1}^{m} a_{lj}\lambda^{m-j}\,, \qquad m > 1\,.$$

Wtedy nie istnieje baza Jordana, ale można określić uogólnienia konstrukcji z ostatniego punktu. Oznaczmy wektory bazowe przez $e_{ij} = Q(A)^{r_l - i} A^{m-j} f_l$, $i = 1, \dots, r_l$, $j = 1, \dots, m$. Działanie operatora A_l w tej bazie ma postać

$$\begin{aligned} A_l e_{11} &= \sum_{j=1}^{m} a_{lj} e_{1j}\,, \\ A_l e_{i1} &= e_{(i-1)m} + \sum_{j=1}^{m} a_{lj} e_{ij} \quad \text{dla} \quad i \geq 2\,, \\ A_l e_{ij} &= e_{i(j-1)} \quad \text{dla} \quad j \geq 2\,. \end{aligned}$$

Uszeregujmy wektory bazowe e_{ij} zgodnie z porządkiem leksykograficznym par (i, j). Macierz operatora A_l ma w tej bazie postać blokową

$$\mathbf{A}_l = \begin{pmatrix} \mathbf{\Lambda}_l & \mathbf{E}_{m1} & 0 & \dots & 0 & 0 \\ 0 & \mathbf{\Lambda}_l & \mathbf{E}_{m1} & \dots & 0 & 0 \\ 0 & 0 & \mathbf{\Lambda}_l & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathbf{\Lambda}_l & \mathbf{E}_{m1} \\ 0 & 0 & 0 & \dots & 0 & \mathbf{\Lambda}_l \end{pmatrix}, \quad \mathbf{\Lambda}_l = \begin{pmatrix} a_{l1} & 1 & & & \\ a_{l2} & 0 & 1 & & \\ \vdots & & \ddots & \ddots & \\ a_{l(m-1)} & & & 0 & 1 \\ a_{lm} & & & & 0 \end{pmatrix},$$

a jedyny niezerowy element macierzy $\mathbf{E}_{m1}$ to jedynka w lewym dolnym rogu. W szczególnym przypadku $m=1$ macierz ma postać jordanowską, z wartością własną równą a_{l1}.

5 Bazy cykliczne

Przy oznaczeniach poprzedniego punktu niech

$$Q_l(\lambda)^{r_l} = \lambda^{mr_l} - \sum_{j=1}^{mr_l} b_{lj}\lambda^{mr_l-j}\,.$$

Przyjmijmy za wektory bazowe $e_j = A^{mr_l-j}f_l$, $j=1,\dots,mr_l$. Działanie operatora A_l w tej bazie ma postać

$$A_l e_1 = \sum_{j=1}^{m} b_{lj}e_j\,, \qquad A_l e_j = e_{j-1} \quad \text{dla} \quad j\geq 2\,,$$

więc jego macierz w uporządkowanej bazie $(e_1,\dots,e_{mr_l})$ jest równa

$$\mathbf{A}_l = \begin{pmatrix} b_{l1} & 1 & & & \\ b_{l2} & 0 & 1 & & \\ \vdots & & \ddots & \ddots & \\ b_{l(mr_l-1)} & & & 0 & 1 \\ b_{l(mr_l)} & & & & 0 \end{pmatrix}.$$

6 Rozkład kanoniczny operatora w przestrzeni rzeczywistej

Jeśli A jest operatorem w rzeczywistej przestrzeni V, to wielomiany Q_l są dwumianami lub nierozkładalnymi trójmianami kwadratowymi, i szczególny przypadek procedur omówionych wyżej ma tutaj zastosowanie. Istnieje jednak w tym przypadku alternatywna metoda, prowadząca w naturalny sposób – poprzez kompleksyfikację – do innych kanonicznych baz.

Niech $A^{\mathbb{C}} : V^{\mathbb{C}} \mapsto V^{\mathbb{C}}$ będzie kompleksyfikacją operatora A. Ponieważ $A^{\mathbb{C}}$ jest rzeczywistym operatorem, to $\overline{\det(\lambda\,\mathrm{id} - A^{\mathbb{C}})} = \det(\overline{\lambda}\,\mathrm{id} - A^{\mathbb{C}})$. Stąd (przykład (ix) w punkcie 9, §4):

$$\det(\lambda\,\mathrm{id} - A^{\mathbb{C}}) = P_1(\lambda)P_2(\lambda)P_3(\lambda)\,, \quad \text{gdzie}$$
$$P_1(\lambda) = (\lambda-\lambda_1)^{n_1}\dots(\lambda-\lambda_k)^{n_k}\,, \quad P_2(\lambda) = (\lambda-\overline{\lambda_1})^{n_1}\dots(\lambda-\overline{\lambda_k})^{n_k}\,,$$
$$\lambda_i = \mu_i + i\nu_i\,, \quad \mu_i,\nu_i\in\mathbb{R}\,,\ \nu_i>0\,,\ i=1,\dots,k\,,$$
$$P_3(\lambda) = (\lambda-\mu_{k+1})^{n_{k+1}}\dots(\lambda-\mu_{s-k})^{n_{s-k}}\,, \quad \mu_i\in\mathbb{R}\,,\ i=k+1,\dots,s-k\,.$$

Zastosowanie wyniku twierdzenia 2 daje, w szczególności, rozkład

$$V^{\mathbb{C}} = W_1 \oplus W_2 \oplus W_3\,, \quad A^{\mathbb{C}} = B_1 \oplus B_2 \oplus B_3\,,$$

jednoznaczny przy nałożeniu warunku $\det(\lambda \operatorname{id} - B_i) = P_i(\lambda)$, $i = 1, 2, 3$. Sprzężenie jest antyizomorfizmem przestrzeni $V^{\mathbb{C}}$, więc $\overline{V^{\mathbb{C}}} = V^{\mathbb{C}}$, a przy tym operator $A^{\mathbb{C}}$ jest rzeczywisty, $\overline{A^{\mathbb{C}}} = A^{\mathbb{C}}$, więc dostajemy również

$$V^{\mathbb{C}} = \overline{W_1} \oplus \overline{W_2} \oplus \overline{W_3}\,, \quad A^{\mathbb{C}} = \overline{B_1} \oplus \overline{B_2} \oplus \overline{B_3}\,.$$

Ponieważ $\det(\lambda \operatorname{id} - \overline{B_i}) = \overline{\det(\overline{\lambda} \operatorname{id} - B_i)}$, to mamy $\det(\lambda \operatorname{id} - \overline{B_1}) = P_2(\lambda)$, $\det(\lambda \operatorname{id} - \overline{B_2}) = P_1(\lambda)$, $\det(\lambda \operatorname{id} - \overline{B_3}) = P_3(\lambda)$. Z jednoznaczności rozkładu mamy więc

$$W_2 = \overline{W_1}\,, \quad B_2 = \overline{B_1}\,, \quad \overline{W_3} = W_3\,, \quad \overline{B_3} = B_3\,.$$

Stąd $\mathfrak{Re}\, W_3$ jest podprzestrzenią przestrzeni V, a B_3 jest kompleksyfikacją operatora A zacieśnionego do operatora na tej podprzestrzeni. Stąd też w przestrzeni W_3 można wybrać rzeczywiste bazy kanoniczne dla tej części operatora A. W szczególności można wybrać bazy jordanowskie złożone z wektorów rzeczywistych; postać kanoniczna tej części operatora utworzona będzie z rzeczywistych klatek Jordana. Inaczej ma się rzecz z operatorami B_1 i B_2. Podprzestrzenie W_1, $\overline{W_1}$ nie są inwariantne względem sprzężenia i nie ma w nich wektorów rzeczywistych. Jeśli jednak wybrać bazy kanoniczne dla operatora B_1, to ich sprzężenia dadzą bazy kanoniczne operatora $\overline{B_1}$; wspólnie te bazy utworzą bazę $\{e_1, \ldots, e_m, \overline{e_1}, \ldots, \overline{e_m}\}$ przestrzeni $W_1 \oplus \overline{W_1}$. Ta przestrzeń jest inwariantna względem sprzężenia, i można wybrać jej bazę złożoną z części rzeczywistych wektorów $\{e_1, \ldots, e_m, ie_1, \ldots, ie_m\}$. Operator $B_1 \oplus \overline{B_1}$ jest kompleksyfikacją operatora A zacieśnionego do operatora na podprzestrzeni $\mathfrak{Re}(W_1 \oplus \overline{W_1})$. Baza $\{\mathfrak{Re}\, e_1, \ldots, \mathfrak{Re}\, e_m, -\mathfrak{Im}\, e_1, \ldots -\mathfrak{Im}\, e_m\}$, po odpowiednim uszeregowaniu, jest rzeczywistą kanoniczną bazą dla tej części operatora A.

7 Uogólnione bazy Jordana w przestrzeni rzeczywistej

Dokończymy tu konstrukcję z poprzedniego punktu dla operatora $B_1 \oplus \overline{B_1}$, wybierając dla operatora B_1 bazę jordanowską. Rozkładamy operator B_1 zgodnie z twierdzeniem 2, a następnie każdy ze składowych operatorów zgodnie z twierdzeniem 5. Wystarczy teraz rozważyć tylko jeden ze składowych operatorów B_{1l}, z wektorem cyklicznym f_l. Konstruujemy bazę Jordana tak, jak w punkcie 3 i wyliczamy:

$$\begin{aligned} &e_j = (B_1 - \lambda_l \operatorname{id})^{r_l - j} f_l\,, \quad j = 1, \ldots, r_l\,, \qquad \lambda_l = \mu_l + i\nu_l\,, \quad \nu_l > 0\,, \\ &A^{\mathbb{C}} e_1 = B_1 e_1 = \lambda_l e_1\,, \qquad A^{\mathbb{C}} e_j = B_1 e_j = e_{j-1} + \lambda_l e_j\,, \quad j = 2, \ldots, r_l\,. \end{aligned}$$

Oznaczamy $e_{j1} = \mathfrak{Re}\, e_j$, $e_{j2} = -\,\mathfrak{Im}\, e_j$, i biorąc pod uwagę rzeczywistość operatora $A^{\mathbb{C}}$ mamy teraz

$$\begin{aligned} Ae_{11} &= \mathfrak{Re}\left(A^{\mathbb{C}}e_1\right) = \mu_l e_{11} + \nu_l e_{12}\,, \\ Ae_{12} &= -\,\mathfrak{Im}\left(A^{\mathbb{C}}e_1\right) = -\nu_l e_{11} + \mu_l e_{12}\,, \\ Ae_{j1} &= \mathfrak{Re}\left(A^{\mathbb{C}}e_j\right) = e_{(j-1)1} + \mu_l e_{j1} + \nu_l e_{j2}\,, \\ Ae_{j2} &= -\,\mathfrak{Im}\left(A^{\mathbb{C}}e_j\right) = e_{(j-1)2} - \nu_l e_{j1} + \mu_l e_{j2}\,. \end{aligned}$$

W bazie (e_{jk}) uszeregowanej leksykograficznie dostajemy klatkę o postaci blokowej

$$\mathbf{A}_l = \begin{pmatrix} \mathbf{\Lambda}_l & \mathbf{1}_2 & 0 & \dots & 0 & 0 \\ 0 & \mathbf{\Lambda}_l & \mathbf{1}_2 & \dots & 0 & 0 \\ 0 & 0 & \mathbf{\Lambda}_l & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & \mathbf{\Lambda}_l & \mathbf{1}_2 \\ 0 & 0 & 0 & \dots & 0 & \mathbf{\Lambda}_l \end{pmatrix}, \quad \mathbf{\Lambda}_l = \begin{pmatrix} \mu_l & -\nu_l \\ \nu_l & \mu_l \end{pmatrix}.$$

8 Przykłady

(i) Rozkład Jordana

Operator A działający w pięciowymiarowej przestrzeni zespolonej ma w bazie $(e_1, \dots, e_5)$ macierz

$$\mathbf{A} = \begin{pmatrix} 2+i & -i & 0 & 1+i & 1 \\ 0 & 3+2i & i & 0 & -2+2i \\ 0 & 0 & 1 & 0 & 0 \\ -2 & 1+i & 0 & -1 & -1+i \\ 0 & -1+3i & -1 & 0 & -2-i \end{pmatrix}.$$

Szukamy bazy, w której macierz operatora przyjmuje postać Jordana.

Wyliczamy $\det(\lambda\,\mathrm{id} - A) = (\lambda-1)^3(\lambda-i)^2$ oraz:

$$\begin{gathered} \mathrm{rk}(\mathbf{A}-\mathbf{1}) = 3\,, \quad \mathrm{rk}(\mathbf{A}-\mathbf{1})^2 = 2\,, \quad \mathrm{rk}(\mathbf{A}-\mathbf{1})^3 = 2\,, \\ \mathrm{rk}(\mathbf{A}-i\mathbf{1}) = 4\,, \quad \mathrm{rk}(\mathbf{A}-i\mathbf{1})^2 = 3\,, \quad \mathrm{rk}(\mathbf{A}-i\mathbf{1})^3 = 3\,. \end{gathered}$$

Pamiętając o związku wymiaru obrazu operatora z wymiarem jego jądra i stosując wzór z punktu (i) twierdzenia 5 znajdujemy wymiary i liczby klatek Jordana, które znajdą się w rozkładzie: dla $\lambda = 1$ po jednej klatce wymiaru 1 i 2, dla $\lambda = i$ jedna klatka wymiaru 2.

$\lambda = 1$
Wyliczamy bazy podprzestrzeni $\mathrm{Ker}(\mathbf{A}-\mathbf{1})^2$ i $\mathrm{Ker}(\mathbf{A}-\mathbf{1})$, wybieramy wektor $\mathbf{e}'_2$ w różnicy tych podprzestrzeni, oznaczamy $\mathbf{e}'_1 = (\mathbf{A}-\mathbf{1})\mathbf{e}'_2$, po czym znajdujemy wektor $\mathbf{e}'_3$ należący do $\mathrm{Ker}(\mathbf{A}-\mathbf{1})$, liniowo niezależny od $\mathbf{e}'_1$ i $\mathbf{e}'_2$. Otrzymujemy:

$$\mathbf{e}'_1 = \begin{pmatrix} 0 \\ i \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{e}'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}'_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

$\lambda = i$
W podobny sposób znajdujemy wektor $\mathbf{e}'_5$ należący do różnicy $\mathrm{Ker}(\mathbf{A}-i\mathbf{1})^2$ i $\mathrm{Ker}(\mathbf{A}-i\mathbf{1})$, po czym oznaczamy $\mathbf{e}'_4 = (\mathbf{A}-i\mathbf{1})\mathbf{e}'_5$. Otrzymujemy:

$$\mathbf{e}'_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1+i \\ 0 \end{pmatrix}, \quad \mathbf{e}'_5 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 1+2i \end{pmatrix}.$$

Budujemy macierz przejścia $\boldsymbol{\beta}$ ustawiając kolejno kolumny $(\mathbf{e}'_i)$. Otrzymujemy bazę $e'_i = e_j \beta^j{}_i$, w której operator A przyjmuje postać kanoniczną:

$$\boldsymbol{\beta} = \begin{pmatrix} 0 & 0 & -1 & 1 & 0 \\ i & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1+i & 0 \\ -1 & 0 & 0 & 0 & 1+2i \end{pmatrix}, \quad \mathbf{A}' = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & i & 1 \\ 0 & 0 & 0 & 0 & i \end{pmatrix}.$$

(ii) Uogólniona baza Jordana w przestrzeni rzeczywistej
Operator A działający w czterowymiarowej przestrzeni rzeczywistej ma w bazie (e_1, e_2, e_3, e_4) macierz

$$\mathbf{A} = \begin{pmatrix} -1 & -4 & -1 & 0 \\ 4 & -1 & 0 & 1 \\ 1 & 0 & -3 & 4 \\ 0 & -1 & -4 & -3 \end{pmatrix}.$$

Szukamy bazy, w której macierz operatora przyjmuje uogólnioną postać Jordana.

Operator $A^{\mathbb{C}}$ ma w tej samej bazie tę samą macierz; znajdujemy:

$$\begin{aligned}\det(\lambda\,\mathrm{id}-A^{\mathbb{C}}) &= (\lambda^2+4\lambda-12)^2+(8\lambda+16)^2 = \\ &= \big((\lambda+2)^2-16\big)^2+64(\lambda+2)^2 = \\ &= [(\lambda+2)^2+16]^2 = [\lambda-(-2+4i)]^2[\lambda-(-2-4i)]^2\,.\end{aligned}$$

Wystarczy rozpatrzyć $\lambda_1 = -2+4i$. Wyliczamy $\dim\mathrm{Ker}(\mathbf{A}-\lambda_1\mathbf{1}) = 1$, $\dim\mathrm{Ker}(\mathbf{A}-\lambda_1\mathbf{1})^2 = 2$. Bez dalszych rachunków można stwierdzić, że wymiar przestrzeni $\mathrm{Ker}(\mathbf{A}-\lambda_1\mathbf{1})^3$ jest równy 2, gdyż suma prosta tej przestrzeni i przestrzeni do niej sprzężonej nie może mieć wymiaru większego od 4. Wybieramy wektor $\mathbf{e}'_2$ w różnicy $\mathrm{Ker}(\mathbf{A}-\lambda_1\mathbf{1})^2\setminus\mathrm{Ker}(\mathbf{A}-\lambda_1\mathbf{1})$ i wyliczamy $\mathbf{e}'_1 = (\mathbf{A}-\lambda_1\mathbf{1})\mathbf{e}'_2$. Dostajemy:

$$\mathbf{e}'_1 = \begin{pmatrix} i\\ 1\\ i\\ -1\end{pmatrix}, \quad \mathbf{e}'_2 = \begin{pmatrix} i\\ 1\\ 0\\ 0\end{pmatrix}, \quad \text{czyli} \qquad \begin{aligned} e'_1 &= ie_1+e_2+ie_3-e_4\,,\\ e'_2 &= ie_1+e_2\,.\end{aligned}$$

Wektory $e'_{i1} = \mathfrak{Re}\,e'_i$ oraz $e'_{i2} = -\,\mathfrak{Im}\,e'_i$ tworzą bazę całej przestrzeni V. Dostajemy bazę

$$(e'_{11}, e'_{12}, e'_{21}, e'_{22}) = (e_2-e_4, -e_1-e_3, e_2, -e_1)\,,$$

w której macierz operatora A ma postać

$$\mathbf{A}' = \begin{pmatrix} -2 & -4 & 1 & 0\\ 4 & -2 & 0 & 1\\ 0 & 0 & -2 & -4\\ 0 & 0 & 4 & -2\end{pmatrix}.$$

(iii) Funkcje operatorów na przestrzeni zespolonej
W punkcie 8, §13, rozważaliśmy szeregi potęgowe operatorów na zespolonych przestrzeniach wektorowych. Postać jordanowska operatora pozwala na proste wyliczenie tak określonych funkcji operatorowych.

Niech $f(x) = \sum_{k=0}^{\infty} c_k x^k$ będzie zespolonym szeregiem o promieniu zbieżności R, a A – liniowym operatorem w przestrzeni zespolonej V. Pytamy czy istnieje operator $f(A)$, a jeśli tak, to jak go przedstawić w zwarty sposób.

Zauważamy na wstępie, że jeśli dokonamy jordanowskiego rozkładu operatora A, to wystarczy rozwiązać problem dla jednego operatora składowego A_l, którego macierz w bazie Jordana ma postać klatki Jordana $\lambda\mathbf{1}+\mathbf{E}$ o wymiarze r. Macierz operatora $f(A_l)$ jest więc równa $f(\lambda\mathbf{1}+\mathbf{E})$, jeśli ten szereg jest zbieżny. Macierze $\mathbf{1}$ i $\mathbf{E}$ komutują, więc dla wyliczenia potęgi $(\lambda\mathbf{1}+\mathbf{E})$ można zastosować wzór Newtona. Wyliczamy:

$$\mathbf{E}^0 = \mathbf{1}\,, \qquad \mathbf{E}^j = \begin{pmatrix} 0 & \mathbf{1}_{r-j} \\ 0 & 0 \end{pmatrix}\,, \quad j = 1, \dots, r-1\,, \qquad \mathbf{E}^r = 0\,,$$

więc każda potęga macierzy $\mathbf{E}$ różna od macierzy zerowej określa inne elementy wyniku, i szereg można sprowadzić do postaci

$$f(\lambda\mathbf{1}+\mathbf{E}) = \sum_{j=0}^{r-1} \mathbf{E}^j \sum_{k=j}^{\infty} c_k \binom{k}{j} \lambda^{k-j}\,.$$

Pojawiające się tu szeregi mają ten sam promień zbieżności, co szereg określający $f(\lambda)$, i wyrażają się przez jego pochodne. Ostatecznie dla $|\lambda| < R$ mamy:

$$f(\lambda\mathbf{1}+\mathbf{E}) = \sum_{j=0}^{r-1} \frac{f^{(j)}(\lambda)}{j!}\,\mathbf{E}^j = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \dots & \frac{1}{(r-1)!}f^{(r-1)}(\lambda) \\ 0 & f(\lambda) & \dots & \frac{1}{(r-2)!}f^{(r-2)}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(\lambda) \end{pmatrix}.$$

Zwróćmy uwagę na formalne podobieństwo pierwszej równości do rozwinięcia Taylora dla funkcji zmiennej zespolonej oraz na fakt, że macierz $f(\lambda\mathbf{1}+\mathbf{E})$ jest trójkątna, z diagonalną częścią równą $f(\lambda)\mathbf{1}$.

Pozostawimy jako ćwiczenie dla czytelnika wykazanie, że dla rzeczywistej uogólnionej klatki Jordana (punkt 7) dostaje się postać podobną do skrajnej prawej strony powyższego wzoru, w której wszystkie elementy są blokami 2×2, przy czym jako argument funkcji $f^{(i)}$ występuje macierz Λ; dla $\mu^2 + \nu^2 < R^2$ funkcje $f^{(i)}(\Lambda)$ są bezwzględnie zbieżne.

W przypadku zespolonym uzyskany wynik prowadzi do następującego wniosku. Niech A będzie operatorem na zespolonej przestrzeni wektorowej, a $f(A)$ – operatorem określonym szeregiem potęgowym. Wtedy:

$$\begin{aligned} &\text{jeśli} \qquad \det(\lambda\,\mathrm{id} - A) = (\lambda-\lambda_1)^{n_1} \dots (\lambda-\lambda_s)^{n_s}\,, \\ &\text{to} \qquad \det\big(\lambda\,\mathrm{id} - f(A)\big) = \big(\lambda - f(\lambda_1)\big)^{n_1} \dots \big(\lambda - f(\lambda_s)\big)^{n_s}\,. \end{aligned}$$

(iv) Wielomian minimalny

Niech A będzie operatorem liniowym w skończenie wymiarowej przestrzeni V nad dowolnym ciałem $\mathbb{K}$. Rozważmy zbiór wszystkich wielomianów Q, dla których $Q(A) = 0$ – mówimy, że Q anihiluje A. Niech M będzie wśród wielomianów anihilujących A, o współczynniku przy najwyższej potędze równym jeden, wielomianem najniższego stopnia. To określa ten wielomian jednoznacznie: gdyby M'

spełniał te same warunki, to $M' - M$ byłby wielomianem niższego stopnia, anihilującym A, co przeczy założeniu. Wielomian M nazywamy ***wielomianem minimalnym operatora*** A. Podobne rozumowanie pokazuje, że każdy wielomian anihilujący A dzieli się przez M: w przeciwnym razie reszta z podzielenia go przez M byłaby wielomianem stopnia niższego niż M, anihilującym A. Stąd, jeśli wielomian charakterystyczny ma rozkład na czynniki pierwsze:

$$\det(\lambda \,\mathrm{id} - A) = Q_1(\lambda)^{n_1} \dots Q_s(\lambda)^{n_s}\,, \quad \text{to} \quad M(\lambda) = Q_1(\lambda)^{m_1} \dots Q_s(\lambda)^{m_s}$$

dla pewnych liczb $m_i \leq n_i$. Niech $A = A_1 \oplus \dots \oplus A_s$ będzie rozkładem określonym twierdzeniem 2. Posługując się wynikiem twierdzenia 5 pokazuje się, że m_i jest ilorazem wymiaru największej spośród podprzestrzeni z wektorem cyklicznym uzyskanych w tym ostatnim twierdzeniu dla operatora A_i przez stopień wielomianu Q_i.

Przyjęty w tej książce kierunek rozważań można odwrócić: niezależna dyskusja własności wielomianu minimalnego może być początkiem alternatywnej metody otrzymania postaci kanonicznej operatora liniowego.

§26 Formy bi-afiniczne. Kwadryki

1 Formy bi-afiniczne

Niech $(\mathcal{M}, V)$ będzie przestrzenią afiniczną nad ciałem $\mathbb{K}$, a h odwzorowaniem $h : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{K}$. Nazwiemy h ***formą bi-afiniczną***, jeśli h jest formą afiniczną w każdym ze swych argumentów, to jest dla dowolnych punktów P, Q oraz wektorów x, y mamy

$$h(P + x, Q) = h(P, Q) + k(x, Q)\,, \quad h(P, Q + y) = h(P, Q) + l(P, y)\,,$$

gdzie dla każdego P, Q funkcje $V \ni x \mapsto k(x, Q) \in \mathbb{K}$ oraz $V \ni y \mapsto l(P, y) \in \mathbb{K}$ są formami liniowymi. Korzystając z tych warunków w dwu różnych kolejnościach dostajemy stąd

$$h(P+x, Q+y) = h(P, Q) + l(P, y) + k(x, Q+y) = h(P, Q) + k(x, Q) + l(P+x, y)\,,$$

więc

$$k(x, Q + y) - k(x, Q) = l(P + x, y) - l(P, y)\,.$$

Lewa strona tej równości nie zależy od P i jest liniowa w x, a prawa strona nie zależy od Q i jest liniowa w y. Stąd obie są niezależne od P i od Q i przedstawiają formę biliniową w x i y; oznaczmy ją $g(x, y)$. Zbieramy wyniki w następującym twierdzeniu.

Twierdzenie 1. *Funkcja $h : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{K}$ jest formą bi-afiniczną wtedy, i tylko wtedy, gdy dla dowolnych punktów P i Q oraz wektorów x i y zachodzi*

$$h(P + x, Q + y) = h(P, Q) + k(x, Q) + l(P, y) + g(x, y)\,,$$

gdzie $g : V \times V \mapsto \mathbb{K}$ jest formą biliniową, a formy $k : V \times \mathcal{M} \mapsto \mathbb{K}$, $l : \mathcal{M} \times V \mapsto \mathbb{K}$ są liniowe w swoich wektorowych argumentach oraz afiniczne w argumentach punktowych, przy czym

$$k(x, Q + y) = k(x, Q) + g(x, y)\,, \quad l(P + x, y) = l(P, y) + g(x, y)\,.$$

Formę g nazwiemy ***częścią biliniową*** formy bi-afinicznej. Jeśli wybrać punkt odniesienia O, to formę bi-afiniczną możemy zapisać jako

$$h(X, Y) = h(O, O) + k(\overrightarrow{OX}, O) + l(O, \overrightarrow{OY}) + g(\overrightarrow{OX}, \overrightarrow{OY})\,.$$

Równość ta może być również wykorzystana do zadania formy bi-afinicznej. Rzeczywiście, ustalmy punkt odniesienia O, wybierzmy stałą c, dwie formy liniowe ψ i φ oraz formę biliniową g na V. Wtedy łatwo pokazać, że przepis

$$h(X, Y) = c + \psi(\overrightarrow{OX}) + \varphi(\overrightarrow{OY}) + g(\overrightarrow{OX}, \overrightarrow{OY})$$

określa formę bi-afiniczną, przy czym $c = h(O, O)$ oraz przy oznaczeniach ostatniego twierdzenia $\psi(x) = k(x, O)$, $\varphi(y) = l(O, y)$.

Korzystając z równoważnego określenia afiniczności odwzorowania możemy formę bi-afiniczną określić jako funkcję, która dla dowolnych kombinacji barycentrycznych $\sum \mu_i X_i$, $\sum \nu_i Y_i$ spełnia warunek

$$h\Big(\sum_i \mu_i X_i, \sum_j \nu_j Y_j\Big) = \sum_{i,j} \mu_i \nu_j h(X_i, Y_j)\,.$$

Formę bi-afiniczną nazywamy ***symetryczną***, gdy $h(X, Y) = h(Y, X)$ dla wszystkich punktów. Przy oznaczeniach ostatniego twierdzenia mamy:

Twierdzenie 2. *Następujące warunki są równoważne (zakładamy* $\operatorname{char}\mathbb{K} \neq 2$*):*

(i) *forma h jest symetryczna;*

(ii) *forma g jest symetryczna oraz $l(P, x) = k(x, P)$ dla wszystkich wektorów i punktów;*

(iii) *dla każdych punktów X i Y zachodzi:*

$$h(X, Y) = 2h\big(\tfrac{1}{2}X + \tfrac{1}{2}Y, \tfrac{1}{2}X + \tfrac{1}{2}Y\big) - \tfrac{1}{2}h(X, X) - \tfrac{1}{2}h(Y, Y)\,,$$

czyli h jest wyznaczona przez swoją formę kwadratową

$$\mathcal{M} \ni X \mapsto h(X, X) \in \mathbb{K}\,.$$

Dowód. Łatwo sprawdzić, że dla dowolnej formy bi-afinicznej h i dowolnego punktu P zachodzi

$$\begin{aligned} 2g(x,y) &= h(P+x,P+y) + h(P-x,P-y) - 2h(P,P)\,,\\ 2[k(x,P) - l(P,x)] &= h(P+x,P-x) - h(P-x,P+x)\,, \end{aligned}$$

więc z punktu (i) wynika (ii). Wynikanie odwrotne jest oczywiste. Przez zastosowanie warunku definicyjnego formy bi-afinicznej w postaci podanej przed twierdzeniem uzyskuje się dla dowolnej formy bi-afinicznej h równość prawej strony równania z punktu (iii) do $\frac{1}{2}[h(X,Y)+h(Y,X)]$. Stąd wynika równoważność punktów (i) i (iii). $\square$

Jeśli oprócz punktu odniesienia wybrać bazę $(e_1,\dots,e_n)$ i wprowadzić odpowiednie współrzędne punktu, $X = O + X^i e_i$, to forma kwadratowa symetrycznej formy bi-afinicznej ma postać

$$h(X,X) = c + 2\varphi_i X^i + g_{ij} X^i X^j\,,$$

jest więc zadana dowolnym wielomianem stopnia ≤ 2 we współrzędnych afinicznych X^i.

2 Klasyfikacja symetrycznych form bi-afinicznych

W dwóch kolejnych punktach chcemy rozwiązać problem klasyfikacji symetrycznych form bi-afinicznych, podobny do analogicznego zagadnienia dla symetrycznych form biliniowych. Zaczniemy od rezultatu pomocniczego. Jeśli V będzie przestrzenią euklidesową, to mówiąc o prostopadłości czy też o ortogonalnym uzupełnieniu $W^\perp$ podprzestrzeni W, będziemy zawsze mieć na myśli ortogonalność względem dodatnio określonego iloczynu definiującego przestrzeń euklidesową.

Lemat 3. *Niech φ będzie formą liniową, a g metryką na przestrzeni wektorowej V. Warunek*

$$\varphi = g(b,.)$$

ma rozwiązanie ze względu na wektor b wtedy, i tylko wtedy, gdy $\operatorname{Ker} g \subseteq \operatorname{Ker}\varphi$. *Jeśli b_0 spełnia ten warunek, to zbiór wszystkich rozwiązań dany jest przez $b_0 + \operatorname{Ker} g$. Jeśli V jest przestrzenią euklidesową, to istnieje dokładnie jeden wektor b_0 spełniający warunek i prostopadły do* $\operatorname{Ker} g$ *względem iloczynu skalarnego.*

Dowód. Jądro $\operatorname{Ker} g$ można uzupełnić przestrzenią W tak, aby $V = \operatorname{Ker} g \oplus W$, przy czym jeśli V jest przestrzenią euklidesową, to wybieramy $W = (\operatorname{Ker} g)^\perp$.

Na W metryka g jest niezdegenerowana, więc na podstawie twierdzenia 3, §16, istnieje dokładnie jeden wektor $b_0 \in W$ taki, że $\varphi_W = g_W(b_0, .\,)$, czyli $\varphi(x) = g(b_0, x)$ dla każdego $x \in W$. Jeśli $\operatorname{Ker} g \subseteq \operatorname{Ker} \varphi$, to równocześnie $\varphi(y) = 0 = g(b_0, y)$ dla $y \in \operatorname{Ker} g$. Stąd $\varphi(x) = g(b_0, x)$ dla wszystkich $x \in V$. Wynikanie odwrotne jest oczywiste. Wektor b jest innym wektorem spełniającym ogólny warunek lematu wtedy, i tylko wtedy, gdy $g(b - b_0, x) = 0$ dla wszystkich $x \in V$, czyli $b - b_0 \in \operatorname{Ker} g$. Stąd też w przypadku euklidesowym b_0 jest jedynym rozwiązaniem należącym do $(\operatorname{Ker} g)^\perp$. □

Niech h będzie symetryczną formą bi-afiniczną, z częścią biliniową g i częścią afiniczno-liniową $l(P, x)$. Z prawa transformacyjnego funkcji $l(P, x)$ przy zmianie punktu P wynika, że jeśli $\operatorname{Ker} g \subseteq \operatorname{Ker} l(P, .\,)$ dla pewnego punktu P, to ta relacja jest też spełniona dla wszystkich punktów P. Scharakteryzujemy formy w zależności od spełnienia tej relacji.

Twierdzenie 4. *Niech h będzie symetryczną formą bi-afiniczną na przestrzeni $\mathcal{M}$ ($\operatorname{char} \mathbb{K} \neq 2$).*

(i) *Jeśli $\operatorname{Ker} g \subseteq \operatorname{Ker} l(P, .\,)$, to punkty O spełniające równanie*

$$l(O, .\,) = 0$$

tworzą niepustą podprzestrzeń afiniczną $O_0 + \operatorname{Ker} g$, a wartość $h(O, O)$ nie zależy od wyboru O w tej podprzestrzeni. Przy takim wyborze punktu O jest

$$h(X, X) = g(\overrightarrow{OX}, \overrightarrow{OX}) + h(O, O)\,.$$

(ii) *Jeśli $\operatorname{Ker} g \nsubseteq \operatorname{Ker} l(P, .\,)$, to punkty O spełniające równanie*

$$h(O, O) = 0$$

tworzą niepusty zbiór. Przy takim wyborze punktu O jest

$$h(X, X) = g(\overrightarrow{OX}, \overrightarrow{OX}) + 2l(O, \overrightarrow{OX})\,.$$

Forma $l(P, .)$ jest niezerowa dla każdego punktu P.

(ii)′ *Jeśli $\operatorname{Ker} g \nsubseteq \operatorname{Ker} l(P, .\,)$ i $\mathcal{M}$ jest przestrzenią euklidesową z iloczynem skalarnym na stowarzyszonej przestrzeni wektorowej oznaczonym przez $(.\,, .\,)$, to istnieje dokładnie jeden wektor $a \in \operatorname{Ker} g \setminus \{0\}$ taki, że*

$$h(O, O) = 0 \quad \textit{oraz} \quad l(O, .\,) = -\frac{1}{2}(a, .\,)$$

dla pewnego punktu O. Wszystkie punkty O, dla których spełnione są te warunki, tworzą podprzestrzeń afiniczną $O_0 + \operatorname{Ker} g \cap L(a)^\perp$. Postać formy pozostaje jak w punkcie (ii).

Dowód. (i) Jeśli P jest dowolnie wybranym punktem, to z afiniczności l w lewym argumencie mamy $l(O,.\,) = l(P,.\,) + g(\overrightarrow{PO},.\,)$, więc szukane punkty O są rozwiązaniami równania $-l(P,.\,) = g(\overrightarrow{PO},.\,)$. Na mocy lematu 3 wszystkie punkty będące rozwiązaniem tego równania tworzą zbiór $O_0 + \operatorname{Ker} g$. Z bi-afiniczności h łatwo teraz wykazać postać $h(X,X)$, a z niej niezależność wartości $h(O,O)$ od O w podprzestrzeni $O_0 + \operatorname{Ker} g$ jest natychmiast widoczna.

(ii) Niech a będzie wektorem w $\operatorname{Ker} g \setminus \operatorname{Ker} l(P,.)$ takim, że $l(P,a) = -\frac{1}{2}$. Łatwo sprawdzić, że dla $O = P + h(P,P)a$ jest $h(O,O) = 0$. Ostatnie stwierdzenie tezy wynika bezpośrednio z założenia punktu.

(ii)$'$ Dla dowolnie wybranego punktu P istnieje dokładnie jeden wektor b taki, że $2l(P,.\,) = (-b,.\,)$. Rozłóżmy $b = a + b^\perp$, $a \in \operatorname{Ker} g$, $b^\perp \in (\operatorname{Ker} g)^\perp$. Na mocy lematu istnieje punkt Q, dla którego $2g(\overrightarrow{PQ},.\,) = (b^\perp,.\,)$, więc mamy $2l(Q,.\,) = 2l(P,.\,) + 2g(\overrightarrow{PQ},.\,) = -(a,.\,)$. Wektor a jest niezerowy, gdyż w przeciwnym razie $l(Q,.\,) = 0$, co przeczyłoby założeniu. Przy wyborze punktu $O_0 = Q + h(Q,Q)\|a\|^{-2}a$ para (O_0,a) spełnia oba warunki punktu (ii)$'$.

Niech teraz $(O_0 + c, a')$ będzie inną parą spełniającą te warunki. Wtedy

$$-\frac{1}{2}(a' - a,.\,) = l(O_0 + c,.\,) - l(O_0,.\,) = g(c,.\,)\,,$$

więc $a' - a \perp \operatorname{Ker} g$, co jest możliwe tylko gdy $a' = a$ (bo równocześnie z założenia $a' - a \in \operatorname{Ker} g$). Stąd $c \in \operatorname{Ker} g$. Uwzględniamy teraz drugi warunek $h(O_0 + c, O_0 + c) = 0$, który przy wykorzystaniu bi-afiniczności formy h i własności pary (O_0,a) daje $(a,c) = 0$, więc $c \in L(a)^\perp$. Łącznie mamy $c \in \operatorname{Ker} g \cap L(a)^\perp$. Odwrotnie, łatwo sprawdzić, że wszystkie punkty w podprzestrzeni $O_0 + \operatorname{Ker} g \cap L(a)^\perp$ spełniają warunki tezy, co kończy dowód. □

3 Postać kanoniczna symetrycznej formy bi-afinicznej na przestrzeni rzeczywistej

Ograniczamy się teraz do przestrzeni rzeczywistych. Łącząc wynik ostatniego twierdzenia poprzedniego punktu z rezultatami uzyskanymi dla symetrycznych form biliniowych dostajemy twierdzenie o postaci kanonicznej.

Twierdzenie 5. *Niech h będzie symetryczną formą bi-afiniczną w przestrzeni rzeczywistej $\mathcal{M}$ o wymiarze n. Wtedy istnieje układ afiniczny, w którym forma przyjmuje jedną z następujących postaci kanonicznych:*

$$(i) \qquad h(X,X) = \sum_{i=1}^{p}(X^i)^2 - \sum_{i=p+1}^{p+q}(X^i)^2 - \mu\,;$$

$$(ii) \qquad h(X,X) = \sum_{i=1}^{p}(X^i)^2 - \sum_{i=p+1}^{p+q}(X^i)^2 - X^n\,, \quad p+q<n\,.$$

Jeśli $\mathcal{M}$ jest przestrzenią euklidesową, to istnieje ortonormalny układ afiniczny, w którym forma przyjmuje jedną z następujących postaci kanonicznych:

$$(i)' \qquad h(X,X) = \sum_{i=1}^{p} \lambda_i (X^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X^i)^2 - \mu \quad \lambda_i > 0;$$

$$(ii)' \qquad h(X,X) = \sum_{i=1}^{p} \lambda_i (X^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X^i)^2 - \nu X^n \,, \quad \lambda_i > 0 \,,$$
$$\nu > 0 \,, \quad p+q < n \,.$$

Postać kanoniczna formy (włącznie z wartościami swobodnych parametrów) nie zależy od wyboru układu afinicznego w danej klasie układów (wszystkich lub ortonormalnych).

Dowód. Niech h będzie dowolną formą. Wybierzmy punkt centralny O układu afinicznego tak, jak w poprzednim twierdzeniu. Wybieramy bazy oddzielnie w każdym z czterech przypadków poprzedniego twierdzenia: (i) w przestrzeni ogólnej lub euklidesowej, (ii) lub (ii)$'$. W przypadku (i) ogólnym wybieramy bazę kanoniczną formy g, a w przypadku (i) euklidesowym – ortonormalną bazę, w której g przyjmuje postać diagonalną. W przypadku (ii) wybieramy bazę kanoniczną formy g zacieśnionej do podprzestrzeni $\operatorname{Ker} l(O,.\,)$ i uzupełniamy wektorem a (wprowadzonym w dowodzie punktu (ii) poprzedniego twierdzenia) jako ostatnim w bazie. W przypadku (ii)$'$ wybieramy bazę ortonormalną podprzestrzeni $\operatorname{Ker} l(O,.\,)$, w której g zacieśniona do tej podprzestrzeni przyjmuje postać diagonalną, i uzupełniamy wektorem $a/\|a\|$ jako ostatnim w bazie (gdzie a jest wektorem, o którym mówi punkt (ii)$'$ poprzedniego twierdzenia). Łatwo sprawdzić, że w takim układzie afinicznym forma h przyjmuje jedną z postaci określonych w tezie, przy czym jeśli h spełnia założenie punktu (i) poprzedniego twierdzenia, to przyjmuje postać (i) lub (i)$'$, a jeśli spełnia założenie punktu (ii) lub (ii)$'$ – postać (ii) lub (ii)$'$ odpowiednio. Zachodzi też wynikanie przeciwne: formy o postaci (i) lub (i)$'$ spełniają założenie punktu (i) poprzedniego twierdzenia, a formy o postaci (ii) lub (ii)$'$ – założenie punktu (ii) lub (ii)$'$ odpowiednio. Stąd postacie kanoniczne (i) i (ii), oraz (i)$'$ i (ii)$'$, wykluczają się wzajemnie. Ponadto, liczby p, q oraz λ_i są niezmiennikami formy g. Dalej, w przypadku (i) (lub (i)$'$) $\mu = -h(O,O)$, a w przypadku (ii)$'$ $\nu = \|a\|$, więc zgodnie z wynikami poprzedniego twierdzenia liczby te również są charakterystykami samej formy h, a nie układu, w którym forma została sprowadzona do postaci kanonicznej. $\square$

Prostym wnioskiem z tego twierdzenia jest następujące twierdzenie o podobieństwie form.

Twierdzenie 6. *Niech h i h' będą symetrycznymi formami bi-afinicznymi na rzeczywistej przestrzeni afinicznej $\mathcal{M}$.*

(i) Formy h i h' mają takie same postacie kanoniczne ((i) lub (ii) w poprzednim twierdzeniu) wtedy, i tylko wtedy, gdy istnieje automorfizm afiniczny $f : \mathcal{M} \mapsto \mathcal{M}$ taki, że

$$\forall X, Y \in \mathcal{M} : \quad h'(X, Y) = h\big(f(X), f(Y)\big) .$$

(ii) Jeśli $\mathcal{M}$ jest przestrzenią euklidesową, to formy h i h' mają takie same postacie kanoniczne właściwe dla przestrzeni euklidesowej ((i)′ lub (ii)′ w poprzednim twierdzeniu) wtedy, i tylko wtedy, gdy istnieje izometria f tej przestrzeni o takiej własności, jak w punkcie (i).

Dowód. Niech forma h przyjmuje postać kanoniczną w układzie $(O, (e_i))$, przy czym jeśli $\mathcal{M}$ jest przestrzenią euklidesową, to ten układ jest ortonormalny. To oznacza, że jeśli $X = O + X^i e_i$, to $h(X, X) = P(\mathbf{X})$, gdzie P jest wielomianem kwadratowym stojącym po prawej stronie jednej z równości w poprzednim twierdzeniu. Niech ponadto automorfizm afiniczny f i układ $(O', (e'_i))$ będą związane, przy ustalonym układzie nieprimowanym, warunkami

$$O' = f^{-1}(O) , \quad e'_i = A^{-1} e_i ,$$

gdzie A jest częścią liniową f – łatwo przekonać się, że związek jest wzajemnie jednoznaczny, przy czym w przestrzeni euklidesowej f jest izometrią wtedy, i tylko wtedy, gdy układ primowany jest ortonormalny. Na mocy własności odwzorowania afinicznego mamy

$$f(X) = f(O' + X'^i e'_i) = O + X'^i e_i , \quad \text{więc} \quad h\big(f(X), f(X)\big) = P(\mathbf{X}') .$$

Jeśli teraz $h'(X, X) = h\big(f(X), f(X)\big)$, to $h'(X, X) = P(\mathbf{X}')$, czyli h' przyjmuje w układzie $(O', (e'_i))$ taką postać, jak h w układzie $(O, (e_i))$. Odwrotnie, jeśli h i h' przyjmują odpowiednio w układach $(O, (e_i))$ i $(O', (e'_i))$ postać kanoniczną określoną tym samym wielomianem kwadratowym P (przy czym w przestrzeni euklidesowej układy są ortonormalne), to $h'(X, X) = P(\mathbf{X}') = h\big(f(X), f(X)\big)$, co kończy dowód. □

4 Algorytm sprowadzenia formy do postaci kanonicznej

Naszkicujemy teraz prosty sposób praktycznego sprowadzenia formy do postaci kanonicznej. Niech forma h dana będzie w pewnym układzie afinicznym $(O, (e_k))$ funkcją

$$h(X, X) = \sum_{i,j=1}^{n} g_{ij} X^i X^j + \sum_{i=1}^{n} \varphi_i X^i + \rho .$$

Szukamy układu $(O', (e'_k))$, w którym forma przyjmuje postać kanoniczną. Wykonujemy następujące kroki.

(i) Wybieramy bazę $(\tilde{e}_1, \ldots, \tilde{e}_n) = (e_1, \ldots, e_n)\tilde{\boldsymbol{\beta}}$, która jest bazą kanoniczną formy g lub – w przypadku euklidesowym – bazą ortonormalną, w której forma g ma postać diagonalną. Wtedy po zamianie współrzędnych $\mathbf{X} = \tilde{\boldsymbol{\beta}}\tilde{\mathbf{X}}$ mamy

$$h(X,X) = \sum_{i=1}^{p} \lambda_i (\tilde{X}^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (\tilde{X}^i)^2 + \sum_{i=1}^{n} \tilde{\varphi}_i \tilde{X}^i + \rho\,, \quad \lambda_i > 0\,,$$

gdzie w przypadku kanonicznym $\lambda_i = 1$, $i = 1, \ldots, p+q$.

(ii) Dokonujemy translacji:

$$\tilde{X}^i = X'^i - \tilde{\varphi}_i/2\lambda_i\,,\ i = 1, \ldots, p\,, \quad \tilde{X}^j = X'^j + \tilde{\varphi}_j/2\lambda_j\,,\ j = p+1, \ldots, p+q\,,$$

co prowadzi do wyeliminowania z $h(X,X)$ wyrazów liniowych w X'^i dla $i = 1, \ldots, p+q$. Dostajemy:

$$h(X,X) = \sum_{i=1}^{p} \lambda_i (X'^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X'^i)^2 + \sum_{i=p+q+1}^{n} \tilde{\varphi}_i \tilde{X}^i - \mu\,.$$

(iii) Jeśli $\tilde{\varphi}_{p+q+1} = 0$, …, $\tilde{\varphi}_n = 0$, to uzyskaliśmy już postać kanoniczną typu (i) (lub (i)$'$) w układzie afinicznym o bazie $(e'_i) = (\tilde{e}_i)$ i punkcie centralnym

$$O' = O - \sum_{i=1}^{p} \frac{\tilde{\varphi}_i}{2\lambda_i} \tilde{e}_i + \sum_{i=p+1}^{p+q} \frac{\tilde{\varphi}_i}{2\lambda_i} \tilde{e}_i\,.$$

(iv) Jeśli dla pewnego $k > p+q$ jest $\tilde{\varphi}_k \neq 0$, to załóżmy bez ograniczenia ogólności, że $\tilde{\varphi}_n \neq 0$ (uzyskamy to zawsze zamieniając ewentualnie $\tilde{e}_k$ i $\tilde{e}_n$ miejscami). W przypadku ogólnej przestrzeni wystarczy teraz położyć:

$$X'^j = \tilde{X}^j\,,\ j = p+q+1, \ldots, n-1\,, \quad X'^n = -\sum_{i=p+q+1}^{n} \tilde{\varphi}_i \tilde{X}^i + \mu\,,$$

aby uzyskać postać kanoniczną typu (ii). W przypadku przestrzeni euklidesowej, aby zachować ortonormalność, musimy żądać, by macierz części jednorodnej transformacji była ortogonalna. Kładziemy

$$\begin{aligned} X'^n &= -\sum_{i=p+q+1}^{n} \frac{\tilde{\varphi}_i}{\nu} \tilde{X}^i + \frac{\mu}{\nu}\,, & \nu &= \Big(\sum_{j=p+q+1}^{n} (\tilde{\varphi}_j)^2\Big)^{1/2} \\ X'^j &= \sum_{k=p+q+1}^{n} \gamma^j{}_k \tilde{X}^k\,, & j &= p+q+1, \ldots, n-1\,, \end{aligned}$$

przy czym wiersze $\boldsymbol{\gamma}^j{}_*$ muszą być unormowane, wzajemnie ortogonalne i ortogonalne do wiersza $\frac{1}{\nu}\begin{pmatrix}\tilde{\varphi}_{p+q+1} & \dots & \tilde{\varphi}_n\end{pmatrix}$ względem naturalnego iloczynu skalarnego w $\mathbb{R}^{n-p-q}$ (zauważmy, że ten wybrany wiersz jest już unormowany). Można to osiągnąć wyliczając ortonormalną bazę przestrzeni rozwiązań równania $\sum_{i=p+q+1}^{n} \tilde{\varphi}_i\, y^i = 0$. W nowych współrzędnych forma ma postać (ii)$'$.

Odwróćmy teraz transformację wykonaną w bieżącym punkcie (w każdym z dwóch przypadków, ogólnym i euklidesowym) i zapiszmy ją jako

$$\tilde{X}^i = \sum_{j=p+q+1}^{n} \beta'^i{}_j X'^j + \tilde{a}^i\,, \qquad i = p+q+1, \dots, n\,.$$

Uwzględniając również transformację wykonaną w punkcie (ii) wnioskujemy zatem, że w przypadku opisanym w bieżącym punkcie postać kanoniczna została uzyskana w układzie afinicznym:

$$e'_i = \tilde{e}_i\,, \quad i = 1, \dots, p+q\,, \qquad e'_j = \sum_{k=p+q+1}^{n} \tilde{e}_k \beta'^k{}_j\,, \quad j = p+q+1, \dots, n\,,$$
$$O' = O - \sum_{i=1}^{p} \frac{\tilde{\varphi}_i}{2\lambda_i}\tilde{e}_i + \sum_{i=p+1}^{p+q} \frac{\tilde{\varphi}_i}{2\lambda_i}\tilde{e}_i + \sum_{i=p+q+1}^{n} \tilde{a}^i \tilde{e}_i\,.$$

5 Kwadryki

Podzbiór Q przestrzeni afinicznej $\mathcal{M}$ nazywamy ***kwadryką***, jeśli istnieje symetryczna forma bi-afiniczna h na tej przestrzeni taka, że

$$Q = \{X \in \mathcal{M} \mid h(X,X) = 0\}\,.$$

Twierdzenie 7. *Równanie kwadryki w przestrzeni rzeczywistej $\mathcal{M}$ może być sprowadzone przez wybór układu afinicznego do jednej z trzech postaci:*

$(i)_0$ $$\sum_{i=1}^{p}(X^i)^2 - \sum_{i=p+1}^{p+q}(X^i)^2 = 0\,, \quad p \le q\,;$$

$(i)_1$ $$\sum_{i=1}^{p}(X^i)^2 - \sum_{i=p+1}^{p+q}(X^i)^2 = 1\,;$$

(ii) $$X^n = \sum_{i=1}^{p}(X^i)^2 - \sum_{i=p+1}^{p+q}(X^i)^2\,, \quad p \ge q\,, \quad p+q < n\,.$$

Jeśli $\mathcal{M}$ jest przestrzenią euklidesową, to istnieje ortonormalny układ afiniczny, w którym równanie kwadryki przyjmuje jedną z następujących postaci:

$$(i)'_0 \qquad \sum_{i=1}^{p} \lambda_i (X^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X^i)^2 = 0\,, \quad \lambda_i > 0\,, \quad p \le q\,;$$

$$(i)'_1 \qquad \sum_{i=1}^{p} \lambda_i (X^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X^i)^2 = 1\,, \quad \lambda_i > 0\,;$$

$$(ii)' \qquad X^n = \sum_{i=1}^{p} \lambda_i (X^i)^2 - \sum_{i=p+1}^{p+q} \lambda_i (X^i)^2\,, \quad \lambda_i > 0\,, \quad p \ge q\,, \quad p+q < n\,.$$

Dowód. Formę h można sprowadzić do jednej z postaci kanonicznych podanych w twierdzeniu 5. Załóżmy najpierw, że forma ma postać (i) lub (i)$'$. Jeśli $\mu = 0$, to równanie kwadryki ma postać (i)$_0$ lub (i)$'_0$ odpowiednio (spełnienie warunku $p \le q$ uzyskamy przez ewentualną zmianę kolejności współrzędnych). Jeśli $\mu \ne 0$, to dzielimy równanie przez μ; w przypadku euklidesowym uzyskujemy w wyniku równanie kwadryki o postaci (i)$'_1$; w przypadku ogólnym dodatkowo przeskalowujemy wektory bazy tak, aby uzyskać równanie w postaci (i)$_1$. Niech teraz forma ma postać (ii) lub (ii)$'$. W pierwszym przypadku daje to równanie kwadryki typu (ii), a w drugim, po podzieleniu przez ν, równanie typu (ii)$'$ – w obu tych przypadkach spełnienie dodatkowego warunku $p \ge q$ uzyskamy przez ewentualną zmianę znaku X^n. □

Rozważmy szczególne przypadki kanonicznych równań kwadryk (i)$_0$, (i)$_1$, (ii) określone przez wartość $p = 0$. Równanie (i)$_0$ definiuje wtedy podprzestrzeń afiniczną $X^1 = 0, \ldots, X^q = 0$, równanie (i)$_1$ nie posiada rozwiązań, a równanie (ii) definiuje podprzestrzeń afiniczną $X^n = 0$. Łatwo też sprawdzić, że jeśli $p \ge 1$, to kwadryka jest zbiorem niepustym, nie będącym podprzestrzenią afiniczną – pozostawiamy wykazanie tego faktu jako ćwiczenie. Obserwacja ta będzie pomocna przy dowodzie następującego twierdzenia.

Twierdzenie 8. *Jeśli kwadryka Q jest niepustym zbiorem nie będącym podprzestrzenią afiniczną, to symetryczna forma bi-afiniczna h zadająca tę kwadrykę zgodnie z warunkiem definicyjnym jest wyznaczona tym warunkiem z dokładnością do niezerowego czynnika.*

Dowód. Zakładamy, że wybrano układ, w którym równanie kwadryki ma postać (i)$_0$, (i)$_1$ lub (ii) z poprzedniego twierdzenia, przy czym $p \ge 1$ (zgodnie z uwagą poprzedzającą twierdzenie). Załóżmy, że ta sama kwadryka może być zadana

równaniem
$$h(X,X) \equiv \sum_{i,j=1}^{n} g_{ij} X^i X^j + \sum_{i=1}^{n} \varphi_i X^i + c = 0\,.$$

Rozpatrzymy osobno przypadki (i) i (ii). Niech równanie kwadryki ma jedną z postaci (i), przy czym $p \geq 1$. Każde rozwiązanie tego równania przechodzi w inne rozwiązanie tego samego równania przy zmianie znaku dowolnie wybranej współrzędnej X^i, więc to samo musi dotyczyć rozwiązań równania $h(X,X) = 0$. Biorąc pod uwagę, że istnieją takie rozwiązania równań (i), dla których wartość wszystkich współrzędnych jest różna od zera, łatwo sprawdzić, że ogranicza to postać formy h do $h(X,X) = \sum_{i=1}^n g_{ii}(X^i)^2 + c$. Równania (i) są w szczególności spełnione dla wszystkich wartości współrzędnych $X^{p+1}, \ldots, X^n$, jeśli dowolnie wybrać $i \leq p$, położyć $X^j = 0$ dla $j \leq p$, $j \neq i$, a $(X^i)^2$ wyliczyć z tego równania. Jeśli podstawimy tak otrzymane wartości zmiennych $X^1, \ldots, X^p$ do $h(X,X)$, to otrzymana funkcja zmiennych $X^{p+1}, \ldots, X^n$ musi być funkcją zerową. Zmieniając wybór wskaźnika i uzyskujemy jednoznaczność h do czynnika różnego od zera; szczegóły pozostawiamy jako ćwiczenie.

Niech teraz równanie kwadryki Q ma postać (ii). Przedstawia ono X^n jako funkcję pozostałych współrzędnych, więc jeśli użyć tego równania do wyeliminowania zmiennej X^n z funkcji $h(X,X)$, to uzyskana funkcja zmiennych $X^1, \ldots, X^{n-1}$ musi być funkcją zerową. Jeśli $p \geq 1$, to tak otrzymana funkcja jest kombinacją jednomianów stopnia mniejszego lub równego cztery; wszystkie współczynniki tej kombinacji muszą się więc zerować. Warunki te znów określają formę z dokładnością do niezerowego czynnika. □

6 Przykłady

(i) Postać kanoniczna równania kwadryki
W rzeczywistej przestrzeni czterowymiarowej, we współrzędnych afinicznych $X = O + X^i e_i$ równanie kwadryki Q ma postać

$$\begin{aligned}(X^1)^2 + 8(X^3)^2 - 2X^1X^2 + 6X^1X^3 - 2X^1X^4 - 4X^2X^3 &+ \\ + 4X^2X^4 - 8X^3X^4 - 4X^1 + 6X^2 - 14X^3 + 2X^4 + 7 &= 0\,.\end{aligned}$$

Chcemy sprowadzić to równanie do postaci kanonicznej.

Oznaczmy lewą stronę równania przez $h(X,X)$. Sprowadzamy najpierw do postaci kanonicznej część biliniową h. Otrzymujemy standardową metodą transformację

$$X^1 = \tilde{X}^1 + \tilde{X}^2 - 2\tilde{X}^3 + 2\tilde{X}^4\,, \quad X^2 = \tilde{X}^2 + \tilde{X}^3 + \tilde{X}^4\,, \quad X^3 = \tilde{X}^3\,, \quad X^4 = \tilde{X}^4\,,$$

która po podstawieniu również do pozostałej części $h(X,X)$ daje

$$h(X,X) = (\tilde{X}^1)^2 - (\tilde{X}^2)^2 - 4\tilde{X}^1 + 2\tilde{X}^2 + 7 = (\tilde{X}^1 - 2)^2 - (\tilde{X}^2 - 1)^2 + 4\,.$$

Podstawienie $\tilde{X}^2 - 1 = 2X'^1$, $\tilde{X}^1 - 2 = 2X'^2$ sprowadza równanie do postaci kanonicznej

$$(X'^1)^2 - (X'^2)^2 = 1\,.$$

Uzupełniając transformację podstawieniami $\tilde{X}^3 = X'^3$, $\tilde{X}^4 = X'^4$, otrzymujemy ostatecznie $\mathbf{X} = \boldsymbol{\beta}\mathbf{X}' + \mathbf{a}$, $(e'_1 \dots e'_4) = (e_1 \dots e_4)\boldsymbol{\beta}$, $O' = O + a^i e_i$, gdzie

$$\boldsymbol{\beta} = \begin{pmatrix} 2 & 2 & -2 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{a} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

(ii) Postać kanoniczna równania kwadryki w przestrzeni euklidesowej
W układzie ortonormalnym $(O, (e_i))$ w czterowymiarowej przestrzeni euklidesowej równanie kwadryki ma postać

$$-3(X^1)^2 + 3(X^3)^2 + (X^4)^2 + 8X^1X^3 + 2X^1X^4 + 4X^3X^4 + \sqrt{6}X^2 - 6X^4 = 0\,.$$

Szukamy układu ortonormalnego, w którym równanie ma postać kanoniczną, właściwą dla przestrzeni euklidesowej.

Oznaczmy lewą stronę równania przez $h(X,X)$. Sprowadzamy najpierw do postaci diagonalnej część biliniową h za pomocą transformacji ortogonalnej. Znajdujemy wartości własne macierzy formy biliniowej: 6, -5 i podwójną 0, i konstruujemy ortogonalną macierz przejścia. Dostajemy:

$$X^1 = \frac{1}{\sqrt{6}}\tilde{X}^1 + \frac{2}{\sqrt{5}}\tilde{X}^2 + \frac{1}{\sqrt{30}}\tilde{X}^4\,, \qquad X^2 = \tilde{X}^3\,,$$
$$X^3 = \frac{2}{\sqrt{6}}\tilde{X}^1 - \frac{1}{\sqrt{5}}\tilde{X}^2 + \frac{2}{\sqrt{30}}\tilde{X}^4\,, \qquad X^4 = \frac{1}{\sqrt{6}}\tilde{X}^1 - \frac{5}{\sqrt{30}}\tilde{X}^4\,.$$

W nowych współrzędnych mamy

$$h(X,X) = 6\Big(\tilde{X}^1 - \frac{1}{2\sqrt{6}}\Big)^2 - 5(\tilde{X}^2)^2 + \sqrt{6}\tilde{X}^3 + \sqrt{30}\tilde{X}^4 - \frac{1}{4}\,.$$

Szukamy takiej ortogonalnej transformacji

$$\begin{pmatrix} \tilde{X}'^3 \\ \tilde{X}'^4 \end{pmatrix} = \boldsymbol{\gamma} \begin{pmatrix} \tilde{X}^3 \\ \tilde{X}^4 \end{pmatrix}, \quad \text{aby} \quad \sqrt{6}\tilde{X}^3 + \sqrt{30}\tilde{X}^4 = -\nu\tilde{X}'^4\,, \quad \nu > 0\,.$$

Znajdujemy

$$\tilde{X}'^3 = -\frac{\sqrt{5}}{\sqrt{6}}\tilde{X}^3 + \frac{1}{\sqrt{6}}\tilde{X}^4\,, \quad \tilde{X}'^4 = -\frac{1}{\sqrt{6}}\tilde{X}^3 - \frac{\sqrt{5}}{\sqrt{6}}\tilde{X}^4\,, \quad \nu = 6\,.$$

Uzyskujemy

$$h(X,X) = 6\Big(\tilde{X}^1 - \frac{1}{2\sqrt{6}}\Big)^2 - 5(\tilde{X}^2)^2 - 6\Big(\tilde{X}'^4 + \frac{1}{24}\Big)\,.$$

Kładziemy teraz

$$X'^1 = \tilde{X}^1 - \frac{1}{2\sqrt{6}}\,, \quad X'^2 = \tilde{X}^2\,, \quad X'^3 = \tilde{X}'^3\,, \quad X'^4 = \tilde{X}'^4 + \frac{1}{24}\,.$$

Eliminując pośrednie zmienne dostajemy $\mathbf{X} = \boldsymbol{\beta}\mathbf{X}' + \mathbf{a}$, gdzie

$$\boldsymbol{\beta} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{1}{6\sqrt{5}} & -\frac{1}{6} \\ 0 & 0 & -\sqrt{\frac{5}{6}} & -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{5}} & \frac{1}{3\sqrt{5}} & -\frac{1}{3} \\ \frac{1}{\sqrt{6}} & 0 & -\frac{\sqrt{5}}{6} & \frac{5}{6} \end{pmatrix}\,, \quad \mathbf{a} = \frac{1}{144}\begin{pmatrix} 13 \\ \sqrt{6} \\ 26 \\ 7 \end{pmatrix}\,.$$

W nowym układzie ortonormalnym $O' = O + a^i e_i$, $e'_i = e_j \beta^j{}_i$ równanie kwadryki ma żądaną postać:

$$X'^4 = (X'^1)^2 - \frac{5}{6}(X'^2)^2\,.$$

ZADANIA

§1 Zbiory i zdania

1 Oprócz funktorów $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$ można określić inne dwuargumentowe funktory logiczne za pomocą odpowiednich tabelek. Wykazać, że jest 16 różnych funktorów, oraz że wszystkie dają się wyrazić za pomocą funktorów $\wedge$, $\vee$ i negacji $\sim$ (np. zdanie $p \Leftrightarrow q$ ma zawsze taką samą wartość logiczną, jak $(p \wedge q) \vee (\sim p \wedge \sim q)$).

2 Wykazać, że następujące formuły są prawami rachunku zdań.

$$\begin{aligned}
&\text{(a) } (p \Rightarrow q) \Leftrightarrow [(p \vee q) \Leftrightarrow q]\,, \ (p \Rightarrow q) \Leftrightarrow [p \Leftrightarrow (p \wedge q)]\,,\\
&\text{(b) } \big[p_1 \wedge \ldots \wedge p_n \Rightarrow q\big] \Leftrightarrow \big[p_1 \Rightarrow \big[p_2 \Rightarrow \big[\ldots\big[p_n \Rightarrow q\big]\ldots\big]\big]\big]\,.
\end{aligned}$$

3 Udowodnić związki:

$$\begin{aligned}
&\text{(a) } A \subseteq B \Leftrightarrow (A \cup B = B) \Leftrightarrow (A \cap B = A) \Leftrightarrow (A \setminus B = \varnothing)\,,\\
&\text{(b) } (A \cup B) \cap (A \cup C) = A \cup (B \cap C)\,, \quad \text{(c) } A \setminus (A \setminus B) = A \cap B\,.
\end{aligned}$$

4 Określamy różnicę symetryczną zbiorów: $A \vartriangle B := (A \setminus B) \cup (B \setminus A)$. Wykazać, że

$$\begin{aligned}
&(A \vartriangle B) \vartriangle C =\\
&\qquad = (A \cap B \cap C) \cup (A \cap B^{\complement} \cap C^{\complement}) \cup (B \cap C^{\complement} \cap A^{\complement}) \cup (C \cap A^{\complement} \cap B^{\complement})\,,\\
&\text{a stąd} \qquad (A \vartriangle B) \vartriangle C = (B \vartriangle C) \vartriangle A = (C \vartriangle A) \vartriangle B.
\end{aligned}$$

5 Rozważamy podzbiory przestrzeni $X \times Y$. Znaleźć tożsamości typu praw rozdzielności dla następujących zbiorów:

$$\text{(a) } (A \cup B) \times C\,, \quad \text{(b) } (A \cap B) \times C\,.$$

6 Wykazać tożsamości dla podzbiorów przestrzeni $X \times Y$:

$$\text{(a) } A \times C = (A \times Y) \cap (X \times C)\,, \quad \text{(b) } (A \times C)^{\complement} = (A^{\complement} \times Y) \cup (X \times C^{\complement})\,,$$
$$\text{(c) } (A \times C)^{\complement} = (A^{\complement} \times C) \cup (A \times C^{\complement}) \cup (A^{\complement} \times C^{\complement})\,.$$

7 Następujące zdanie jest twierdzeniem teorii zbiorów (w ramach jednego z jej ujęć): nie istnieją x_i takie, że $x_1 \in x_2 \in \ldots \in x_n \in x_1$. Korzystając z tego wykazać, że definicja $(x, y) = \{x, \{x, y\}\}$ spełnia warunek nakładany na pary uporządkowane.

8 Znaleźć zbiory liczb rzeczywistych dane wyrażeniami

$$\text{(a) } \bigcup_{n \in \mathbb{N}} \langle -1 + 1/n, 1 - 1/n \rangle\,, \qquad \text{(b) } \bigcup_{n \in \mathbb{N}} (-1 + 1/n, 1 - 1/n)\,,$$
$$\text{(c) } \bigcap_{n \in \mathbb{N}} \langle -1/n, 1/n \rangle\,, \qquad \text{(d) } \bigcap_{n \in \mathbb{N}} (-1/n, 1/n)\,,$$

§2 Relacje. Odwzorowania

9 W dziedzinie odwzorowania $f : X \mapsto Y$ określamy relację $x_1 \sim x_2 \Leftrightarrow f(x_1) = f(x_2)$. Wykazać, że jest to relacja równoważności. Zapisać zbiór ilorazowy przy użyciu pojęć obrazu i przeciwobrazu.

10 W zbiorze $\mathbb{R} \times \mathbb{R} \setminus \{(0, 0)\}$ określamy relację:
$(x_1, y_1)\mathcal{R}(x_2, y_2) \Leftrightarrow x_1 y_2 = x_2 y_1$. Wykazać, że jest to relacja równoważności, otrzymać jej klasy równoważności i przedstawić je graficznie na płaszczyźnie kartezjańskiej.

11 Rozważamy funkcję $f : \mathbb{R} \mapsto \mathbb{R}$, $f(x) = x^3 - 3x$.

(a) Przedstawić na osiach wykresu funkcji obrazy i przeciwobrazy przedziałów w $\mathbb{R}$, zidentyfikować ich rodzaje.

(b) Znaleźć graficznie wszystkie zbiory $X \subseteq \mathbb{R}$ złożone co najwyżej z trzech przedziałów rozłącznych (skończonych lub nie) takie, że zacieśnienie f do X jest bijekcją X na $\mathbb{R}$.

12 Niech $f : X \mapsto Y$. Wykazać, że

(a) dla każdego zbioru $A \subseteq X$ jest $A \subseteq f^{-1}(f(A))$, oraz dla każdego zbioru $B \subseteq Y$ jest $f(f^{-1}(B)) \subseteq B$,

(b) f jest injekcją wtedy, i tylko wtedy, gdy dla każdego zbioru $A \subseteq X$ zachodzi $f^{-1}(f(A)) = A$,

(c) f jest surjekcją wtedy, i tylko wtedy, gdy dla każdego zbioru $B \subseteq Y$ zachodzi $f(f^{-1}(B)) = B$.

13 Mając dane odwzorowanie $f : X \mapsto Y$ określamy $F : P(X) \mapsto P(Y)$, $F(A) = f(A)$. Wykazać, że F jest injekcją (surjekcją) wtedy, i tylko wtedy, gdy f jest injekcją (odpowiednio: surjekcją).

14 Jeśli $f : X \mapsto Y$ jest bijekcją, to dla $B \subseteq Y$ symbol $f^{-1}(B)$ wydaje się niejednoznaczny, gdyż może być rozumiany jako przeciwobraz zbioru B w odwzorowaniu f, lub obraz zbioru B w odwzorowaniu f^{-1}. Wykazać, że te dwa zbiory są równe, więc trudność nie powstaje.

15 Dane są odwzorowania $f_i : X_i \mapsto Y_i$ $(i = 1, \ldots, k)$. Definiujemy odwzorowanie

$$f_1 \times \ldots \times f_k : X_1 \times \ldots \times X_k \mapsto Y_1 \times \ldots \times Y_k \,,$$
$$(f_1 \times \ldots \times f_k)(x_1, \ldots, x_k) = (f_1(x_1), \ldots, f_k(x_k)) \,.$$

Wykazać, że

(a) $f_1 \times \ldots \times f_k$ jest injekcją (surjekcją) wtedy, i tylko wtedy, gdy wszystkie f_i są injekcjami (surjekcjami).

(b) $(g_1 \times \ldots \times g_k) \circ (f_1 \times \ldots \times f_k) = (g_1 \circ f_1) \times \ldots \times (g_k \circ f_k)$.

16 Wykazać, że zbiór $\mathbb{R}$ jest równej mocy z każdym przedziałem otwartym.

17 Wykazać, że zbiór $\mathbb{N} \times \mathbb{N}$ jest równej mocy z $\mathbb{N}$.

§3 Działania, grupa, ciało

18 Niech zbiór X będzie liniowo uporządkowany relacją $\leq$. Określamy w tym zbiorze działanie $x \circ y = \max\{x, y\}$. Sprawdzić, czy działanie jest łączne. Jaki jest warunek istnienia elementu neutralnego? Czy istnieją wtedy elementy odwrotne?

19 W zbiorze X zadajemy działanie przepisem $x \circ y = x$. Sprawdzić, czy to działanie jest łączne. Co można wywnioskować o zbiorze X, jeśli to działanie ma w nim element neutralny?

20 Wykazać, że zbiór potęgowy $P(X)$ tworzy z działaniem różnicy symetrycznej grupę przemienną.

21 Rozważamy grupę ilorazową $\mathbb{R}^*/\mathbb{Q}^*$ multiplikatywnych grup liczbowych. Wykazać, że dla każdego $\epsilon > 0$ zachodzi $\mathbb{R}^*/\mathbb{Q}^* = \{x\mathbb{Q}^* \mid x \in (0, \epsilon)\}$.

22 Niech X_r będzie zbiorem złożonym z liczb o postaci $a + br$, gdzie r jest ustaloną liczbą niewymierną, a a, b są dowolnymi liczbami wymiernymi.

(a) Kiedy $X_r = X_{r'}$?

(b) Wykazać, że każdy ze zbiorów X_r jest podgrupą grupy addytywnej $\mathbb{R}$.

(c) Znaleźć wszystkie liczby r, dla których zbiory $X_r^* = X_r \setminus \{0\}$ są podgrupami grupy multiplikatywnej $\mathbb{R}^*$. Znaleźć najprostszy zbiór takich liczb r, że grupy X_r^* są różne między sobą i wyczerpują wszystkie grupy tego typu.

(d) Wykazać, że zbiory X_r, gdzie r przyjmuje wartości określone warunkami poprzedniego punktu, tworzą ciała (podciała liczb rzeczywistych ze zwykłymi działaniami).

23 Wykazać, że zbiór $\mathbb{R} \times \mathbb{R}^*$ z działaniem $(a, b) \circ (a', b') = (a + ba', bb')$ tworzy grupę; oznaczmy ją $\mathbb{R} * \mathbb{R}^*$. Wykazać, że odwzorowanie $h : \mathbb{R} * \mathbb{R}^* \mapsto \mathbb{R}^*$, $h(a, b) = b$, jest homomorfizmem tej grupy na grupę multiplikatywną $\mathbb{R}^*$. Wykazać, że grupa niezmiennicza $\operatorname{Ker} h$ jest izomorficzna z grupą addytywną $\mathbb{R}$, a grupa ilorazowa $\mathbb{R} * \mathbb{R}^* / \operatorname{Ker} h$ jest izomorficzna z grupą $\mathbb{R}^*$.

24 Znaleźć podgrupę grupy $\mathbb{R} * \mathbb{R}^*$ z poprzedniego zadania generowaną podzbiorem złożonym ze wszystkich elementów, których kwadrat jest elementem jednostkowym. Sprawdzić, czy ta podgrupa jest przemienna.

25 Niech $P_1, \ldots, P_s$ będą niezerowymi wielomianami w dowolnym ciele $\mathbb{K}$, i załóżmy, co nie ogranicza ogólności, że stopień wielomianu P_s jest nie większy od stopni wszystkich pozostałych wielomianów. Podobnie jak w przykładzie (iv) w punkcie 15, §3, dokonujemy dzielenia $P_i = W_i P_s + P_i'$, $i = 1, \ldots, s-1$, i kładziemy $P_s = P_{s'}$. Wykazać, że wielomian Q jest wspólnym podzielnikiem wielomianów $P_1, \ldots, P_s$ wtedy, i tylko wtedy, gdy jest wspólnym podzielnikiem niezerowych spośród wielomianów $P_1', \ldots, P_s'$. Otrzymujemy metodę znalezienia maksymalnego podzielnika ciągu wielomianów.

26 Dla wielomianów P_i w ciele liczb rzeczywistych znaleźć:
– największy wspólny podzielnik – jeśli istnieje,
– wielomiany Q_i takie, że $\sum_i P_i Q_i = 1$ – jeśli taki podzielnik nie istnieje.

(a) $P_1(x) = x^5 + x^4 + x^3 + x^2 + x + 1, \quad P_2(x) = x^3 + x^2 + x + 1\,;$

(b) $P_1(x) = x^5 + 6x^2\,, \quad P_2(x) = x^4 + 7x\,, \quad P_3(x) = x^3 + 8\,.$

§4 Liczby zespolone

27 Przedstawić dane liczby zespolone w jawnej postaci $a + bi$.

$$\text{(a)}\ \frac{3+5i}{2-7i}\,, \quad \text{(b)}\ 3e^{i\pi/3}\,, \quad \text{(c)}\ e^{i5\pi/12}\,, \quad \text{(d)}\ \left(e^{i\pi/6}\right)^{35}\,, \quad \text{(e)}\ (1+i)^{11}\,.$$

28 Przedstawić każdą z liczb $e^{i\varphi} \pm e^{i\psi}$ w postaci $x\omega$, gdzie $x \in \mathbb{R}$ i $|\omega| = 1$.

29 Wyliczyć $\mathfrak{Re}(z^n)$ i $\mathfrak{Im}(z^n)$ dla $z = a + ib$.

30 Rozwiązać równania ($w \in \mathbb{C}$ jest zadane)

$$\text{(a)}\ z^2 + (-3+2i)z - 7 + 11i = 0\,, \quad \text{(b)}\ \left(z - \frac{1}{z}\right)^2 + (2\sin w)^2 = 0\,,$$
$$\text{(c)}\ (2-3i)z^2 + (11+16i)z - 11 + 23i = 0\,, \quad \text{(d)}\ \sinh z = \sinh w\,.$$

31 Wykazać tożsamość: $|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2|z_1|^2 + 2|z_2|^2$.

32 Znaleźć wszystkie pierwiastki trzeciego i czwartego stopnia z 1 rozwiązując równania $z^3 - 1 = 0$ i $z^4 - 1 = 0$ (nie korzystać ze znanych wartości funkcji trygonometrycznych).

33 Niech $z = x + iy$. Wykazać rozkłady

$$\cosh z = \cosh x \cos y + i \sinh x \sin y\,,$$
$$\sinh z = \sinh x \cos y + i \cosh x \sin y\,.$$

Znaleźć analogiczne rozkłady dla $\cos z$ i $\sin z$.

34 Znaleźć wszystkie pierwiastki 8-ego stopnia z 1 w postaci $x + iy$. Przyporządkować każdemu z nich odpowiednią postać eksponencjalną. Które z tych pierwiastków są pierwiastkami pierwotnymi?

35 Znaleźć wszystkie pierwiastki wielomianu $W(z) = 1 + z + z^2 + \ldots + z^n$ i rozłożyć go na czynniki pierwsze.

36 Wykazać, że jeśli $\varepsilon_1, \ldots, \varepsilon_n$ są wszystkimi pierwiastkami n-tego stopnia z 1, a w jest dowolnym pierwiastkiem n-tego stopnia z z, to wszystkie pierwiastki n-tego stopnia z z można przedstawić jako $w\varepsilon_1, \ldots, w\varepsilon_n$.

37 Przy oznaczeniach poprzedniego zadania wykazać tożsamości:

$$\text{(a)}\ \sum_i \varepsilon_i = 0\,, \quad \sum_{i_1<i_2} \varepsilon_{i_1}\varepsilon_{i_2} = 0\,, \quad \ldots\,, \quad \sum_{i_1<\ldots<i_{n-1}} \varepsilon_{i_1}\ldots\varepsilon_{i_{n-1}} = 0\,,$$

$$\varepsilon_1\ldots\varepsilon_n = (-1)^{n-1}\,,$$

$$\text{(b)}\ \sum_{i=1}^{n} (\varepsilon_i)^k = 0\,,\ k = 1,\ldots,n-1\,, \qquad \text{(c)}\ \sum_{i=0}^{n-1} (\varepsilon_k)^i = 0\,,\ \varepsilon_k \neq 1\,,$$

$$\text{(d)}\ (\varepsilon_n - \varepsilon_1)\ldots(\varepsilon_n - \varepsilon_{n-1}) = n\overline{\varepsilon_n}\,.$$

38 Wykazać, że jeśli ε jest pierwiastkiem pierwotnym n-tego stopnia z 1 i $n = mk$, to ε^m jest pierwiastkiem pierwotnym k-tego stopnia z 1.

39 Zwijając sumę $1 + z + z^2 + \ldots + z^n$ dla $z = e^{i\varphi}$ wykazać, że

$$\sin\varphi + \sin 2\varphi + \ldots + \sin n\varphi = \frac{\sin\left(\frac{n+1}{2}\varphi\right)\sin\left(\frac{n}{2}\varphi\right)}{\sin\left(\frac{1}{2}\varphi\right)}\,,$$

$$1 + \cos\varphi + \ldots + \cos n\varphi = \frac{\sin\left(\frac{n+1}{2}\varphi\right)\cos\left(\frac{n}{2}\varphi\right)}{\sin\left(\frac{1}{2}\varphi\right)} = \frac{\sin\left[(n+\frac{1}{2})\varphi\right]}{2\sin\left(\frac{1}{2}\varphi\right)} + \frac{1}{2}\,.$$

40 Dla różnowartościowego ciągu trzech liczb zespolonych z_1, z_2, z_3 oznaczmy $d(z_1, z_2, z_3) = \dfrac{z_1 - z_2}{z_1 - z_3}$. Wykazać, że jeśli $d(z_1, z_2, z_3)$ jest liczbą rzeczywistą, to liczby z_1, z_2, z_3 leżą na jednej prostej na płaszczyźnie Gaussa, a w przeciwnym przypadku leżą na jednym okręgu.

41 Dla różnowartościowego ciągu czterech liczb zespolonych z_1, z_2, z_3, z_4 oznaczmy $\mu(z_1, z_2, z_3, z_4) = \dfrac{(z_1 - z_2)(z_3 - z_4)}{(z_1 - z_4)(z_3 - z_2)}$. Wykazać, że μ jest liczbą rzeczywistą wtedy, i tylko wtedy, gdy te liczby leżą na jednym okręgu lub na jednej prostej na płaszczyźnie Gaussa.

§5 Grupy odwzorowań. Permutacje

42 Niech f i g będą odwzorowaniami uzwarconej płaszczyzny zespolonej $\overline{\mathbb{C}}$ na siebie danymi przez $f(z) = z^{-1}$, $g(z) = 1 - z$. Wykazać, że zbiór $\{f, g\}$ generuje sześcioelementową podgrupę homografii. Znaleźć wszystkie jej podgrupy.

43 Niech h_λ oznacza dla każdej liczby $\lambda \in \mathbb{C}^*$ odwzorowanie uzwarconej płaszczyzny zespolonej dane przez $h_\lambda(z) = \lambda z$. Wykazać, że zbiór odwzorowań $\{h_\lambda \mid \lambda \in \mathbb{C}^*\} \cup \{f, g\}$ (f i g jak w poprzednim zadaniu) generuje grupę homografii.

44 Wykazać, że wszystkie odwzorowania z poprzedniego zadania zachowują wartość funkcji μ określonej w zadaniu 41 (tj. jeśli k jest dowolnym z tych odwzorowań, to $\mu(k(z_1), k(z_2), k(z_3), k(z_4)) = \mu(z_1, z_2, z_3, z_4)$). Wyciągnąć wniosek, że pod działaniem homografii zbiór figur złożony z okręgów i prostych przechodzi w siebie. Kiedy obrazem jest okrąg, a kiedy prosta?

45 Wyliczyć złożenia i odwrotne w grupie permutacji.

$$\text{(a)}\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}^{-1}, \quad \text{(b)}\ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 6 & 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 6 & 3 & 1 & 4 & 5 \end{pmatrix}^{-1},$$

$$\text{(c)}\ (a, b_1, \ldots, b_m, c, d_1, \ldots, d_n)(a, c)\,,$$
$$\text{(d)}\ (a, d_1, \ldots, d_n)(c, b_1 \ldots, b_m)(a, c)\,.$$

46 Znaleźć rozwiązania równań.

$$\text{(a)}\ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix} \pi \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix},$$
$$\text{(b)}\ \left[\pi \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 1 & 2 & 5 & 4 \end{pmatrix}\right]^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 1 & 6 & 2 & 5 \end{pmatrix}.$$

47 Wyliczyć znak permutacji.

$$\text{(a)}\ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 7 & 9 & 3 & 4 & 8 & 1 & 6 & 5 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & 2 & \ldots & n \\ k & k+1 & \ldots & k-1 \end{pmatrix}.$$

48 Wykazać, że $\quad \operatorname{sgn} \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ n & n-1 & \ldots & 2 & 1 \end{pmatrix} = (-1)^{n(n-1)/2}.$

49 Wykazać, że każdy z następujących zbiorów permutacji generuje grupę S_n.

$$\text{(a)}\ \big\{(1, 2), \ldots, (1, n)\big\}, \quad \text{(b)}\ \big\{(1, 2), (2, 3), \ldots, (n-1, n)\big\},$$
$$\text{(c)}\ \big\{(1, 2), (1, \ldots, n)\big\}.$$

50 Wykazać, że jeśli c jest cyklem n-wyrazowym, to dla każdego $k \in \mathbb{Z}$ rozkład permutacji c^k na cykle rozłączne ma postać $c^k = d_1 \circ \ldots \circ d_m$, gdzie m jest największym wspólnym podzielnikiem liczb n i k, a każdy cykl d_i ma n/m wyrazów. Wystarczy rozważać $k = 0, 1, \ldots, n-1$.

51 Jeśli w poprzednim zadaniu n i k są względem siebie pierwsze, to $c^k = d$ jest cyklem. Odwrotnie, wykazać że w tym przypadku dla każdego cyklu n-wyrazowego d istnieje dokładnie jeden cykl c spełniający to równanie.

52 Przedstawić następujące permutacje w postaci iloczynów cykli rozłącznych.

$$\text{(a) } (a_1, \ldots, a_{12})^8\,, \quad \text{(b) } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 12 & 11 & 8 & 2 & 10 & 4 & 9 & 13 & 3 & 1 & 7 & 5 & 6 \end{pmatrix}^6,$$

53 Wykazać, że jeśli permutacje spełniają związek $p^{-1} q\, p = r$, to rozkłady q i r na cykle rozłączne mają tyle samo czynników o tej samej długości.

54 Znaleźć wszystkie podgrupy grupy S_3. Ile jest podgrup niezmienniczych tej grupy?

55 Wykazać, że podzbiór $G = \{\mathrm{id}, (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ grupy permutacji S_4 jest jej niezmienniczą podgrupą. Wyliczyć klasy permutacji tworzące grupę ilorazową S_4/G i wykazać, że ta grupa jest izomorficzna z grupą S_3.

56 Niech μ będzie funkcją określoną w zadaniu 41, H – grupą znalezioną w zadaniu 42, a G – grupą określoną w zadaniu 55. Wykazać następujące stwierdzenia. Dla każdej permutacji $\pi \in S_4$ zachodzi $\mu(z_{\pi^{-1}(1)}, \ldots, z_{\pi^{-1}(4)}) = F_\pi(\mu(z_1, \ldots, z_4))$, gdzie $F_\pi \in H$. Odwzorowanie $S_4 \ni \pi \mapsto F_\pi \in H$ jest surjektywnym homomorfizmem grup, którego jądrem jest G. Wyciągnąć wniosek, że H jest izomorficzna z S_3. Odwrotnie, jeśli stwierdzić niezależnie izomorficzność H i S_3, to otrzymujemy inną metodą wyniki poprzedniego zadania.

§6 Macierze

57 Wyliczyć wynik działań na macierzach.

$$\text{(a) } \begin{pmatrix} 1 & 3 & -2 & 1 \\ 3 & -1 & 2 & 0 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 2 & 7 & 6 \\ -2 & 4 & 1 \end{pmatrix}, \quad \text{(b) } \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 3+5i \\ 4-2i \end{pmatrix},$$

(c) $\begin{pmatrix} i & 1-i & 2+i \\ 0 & -1+3i & 2-3i \end{pmatrix}^{\dagger} \begin{pmatrix} 0 & 3-i & i \\ 3i & 4 & 1-i \end{pmatrix}$, (d) $X^{\dagger}Y$, $YX^{\dagger}$, $XY^{\dagger}$,

gdzie $X^{\mathrm{T}} = \begin{pmatrix} 3-2i & 1+i & 5 \end{pmatrix}$, $Y^{\mathrm{T}} = \begin{pmatrix} i & 3-i & 5+i \end{pmatrix}$.

58 Wykazać, że każdą macierz kwadratową o elementach w ciele o charakterystyce różnej od 2 można jednoznacznie rozłożyć na sumę macierzy symetrycznej i macierzy antysymetrycznej. Podobnie, każdą zespoloną macierz kwadratową A można jednoznacznie przedstawić w postaci $A = B + iC$, gdzie B i C są hermitowskie. Wyliczyć odpowiednie rozkłady dla macierzy

$$\begin{pmatrix} 2 & -5 & 4 \\ 1 & 3 & 5 \\ 8 & -7 & 9 \end{pmatrix}, \quad \begin{pmatrix} 1+i & 2+3i & i \\ 3-i & 2+3i & 2 \\ 4i & 2-i & 1+5i \end{pmatrix}.$$

59 Wykazać, że iloczyn macierzy trójkątnych jest macierzą trójkątną.

60 Oznaczmy przez $T(n, \mathbb{K})$ zbiór macierzy trójkątnych wymiaru n, o wyrazach na diagonali różnych od zera, a przez $D(n, \mathbb{K})$ podzbiór złożony z macierzy diagonalnych.

(a) Wykazać, że $T(n, \mathbb{K})$ jest podgrupą grupy $GL(n, \mathbb{K})$.

(b) Wykazać, że $D(n, \mathbb{K})$ jest podgrupą grupy $T(n, \mathbb{K})$, oraz że odwzorowanie $h : T(n, \mathbb{K}) \mapsto D(n, \mathbb{K})$, którego działanie polega na wyzerowaniu elementów pozadiagonalnych przy zachowaniu diagonali, jest surjektywnym homomorfizmem.

(c) Z jakich macierzy złożona jest grupa $\operatorname{Ker} h$? Przeanalizować izomorfizm grupy $T(n, \mathbb{K}) / \operatorname{Ker} h$ z grupą $D(n, \mathbb{K})$ (określony standardową konstrukcją – patrz przykład (iii) w punkcie 9, §3).

61 Macierz $A \in \mathbb{K}^{n \times n}$ spełnia równanie $A^2 = \mathbf{1}$. Czy wynika stąd, że (a) $A = \mathbf{1}$ lub $A = -\mathbf{1}$? (b) A jest diagonalna? Podać dowody lub kontrprzykłady.

62 Wykazać, że nie istnieją macierze kwadratowe A, B, dla których zachodziłby związek $AB - BA = \mathbf{1}$.

63 Wykazać, że zbiór macierzy rzeczywistych 2×2 przemiennych z macierzą $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, wyposażony w mnożenie i dodawanie, tworzy ciało izomorficzne z ciałem liczb zespolonych.

64 Niech macierze $A, B \in \mathbb{K}^{n\times n}$ spełniają równanie $AB = 0$. Przy jakich warunkach można stąd wywnioskować, że $A = 0$ lub $B = 0$? Podać przykład pary macierzy niezerowych, które spełniają podane równanie.

65 Wykazać bez użycia pojęcia wyznacznika, że jeśli macierz kwadratowa $A \neq \lambda\mathbf{1}$ spełnia równanie kwadratowe typu $A^2 + \alpha A + \beta\mathbf{1} = 0$, to odwrotna macierz A^{-1} istnieje wtedy, i tylko wtedy, gdy $\beta \neq 0$, i wyliczyć ją w tym przypadku.

66 Zastosować wynik poprzedniego zadania do macierzy $A = \mathbf{1} + XY^{\mathrm{T}}$, gdzie X i Y są niezerowymi kolumnami.

67 Zakładamy, że dla macierzy $A \in \mathbb{K}^{m\times n}$ i $B \in \mathbb{K}^{n\times m}$ istnieje macierz $C = (\mathbf{1}_m + AB)^{-1}$. Przekształcając równanie $(\mathbf{1}_m + AB)C = \mathbf{1}_m$ wykazać, że wtedy istnieje również macierz $(\mathbf{1}_n + BA)^{-1}$ i wyrazić ją za pomocą macierzy A, B i C oraz jednostkowej (uogólnienie wyniku poprzedniego zadania).

§7 Wyznaczniki. Macierz odwrotna

68 Obliczyć wyznaczniki.

$$\text{(a)} \begin{vmatrix} i & 1 & -i \\ 1+i & i & 1-i \\ i & -1 & -i \end{vmatrix}, \quad \text{(b)} \begin{vmatrix} 1 & -2 & 3 & -4 \\ 2 & 4 & 3 & 5 \\ 3 & 5 & 2 & 4 \\ -5 & 4 & -3 & 2 \end{vmatrix}, \quad \text{(c)} \begin{vmatrix} 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 2 \\ 0 & 5 & 3 & 1 \\ 0 & 0 & 2 & 4 \end{vmatrix}.$$

69 Znaleźć rozwiązania warunków:

$$\text{(a)} \begin{vmatrix} z & 2+i \\ 1+3i & 1-i \end{vmatrix} \in \mathbb{R}, \quad \text{(b)} \begin{vmatrix} z & 1 & 1-i \\ -1 & 1 & -i \\ -1-4i & 1 & z \end{vmatrix} = 0\,.$$

70 Wykazać, że wyznacznik macierzy $\mathbf{1}+XY^{\mathrm{T}}$, gdzie X i Y są kolumnami, jest równy $1 + X^{\mathrm{T}}Y$.

71 Określamy kwadratową macierz zespoloną

$$E = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & 1 \\ \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n \\ \varepsilon_1^2 & \varepsilon_2^2 & \dots & \varepsilon_n^2 \\ \dots & \dots & \dots & \dots \\ \varepsilon_1^{n-1} & \varepsilon_2^{n-1} & \dots & \varepsilon_n^{n-1} \end{pmatrix},$$

gdzie $\varepsilon_1, \ldots, \varepsilon_n$ są wszystkimi pierwiastkami n-tego stopnia z 1.

(a) Wykazać, że macierz E jest unitarna.

(b) Znaleźć liniowy związek wyznaczników macierzy E i $\overline{E}$. Wyciągnąć wniosek o wartości wyznacznika macierzy E – z dokładnością do znaku, który zależy od kolejności pierwiastków.

(c) Uzupełnić brakującą informację o znaku dla pierwiastków uszeregowanych w kolejności $1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}$, gdzie $\varepsilon = \exp[i2\pi/n]$.

72 Wykazać, że jeśli macierz kwadratowa ma postać blokową $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, gdzie bloki A i B nie są kwadratowe, to jej wyznacznik jest równy 0.

73 Niech $A \in \mathbb{K}^{n\times n}$, $k \in \{1, \ldots, n-1\}$ i niech $I, J \subseteq \{1, \ldots, n\}$ będą k-elementowymi podzbiorami. Oznaczmy przez A_{IJ} macierz powstającą z elementów macierzy A położonych na przecięciu wszystkich wierszy o numerach ze zbioru I i wszystkich kolumn o numerach ze zbioru J (macierz jest utworzona z zachowaniem kolejności kolumn i wierszy). Oznaczmy ponadto $I' = \{1, \ldots, n\} \setminus I$ oraz $|I| = \sum_{i\in I} i$. Udowodnić następujące uogólnienie rozwinięcia Laplace'a

$$\det A = \sum_J (-1)^{|I|+|J|} \det A_{IJ} \det A_{I'J'} = \sum_J (-1)^{|I|+|J|} \det A_{JI} \det A_{J'I'}\,,$$

gdzie sumowanie przebiega po wszystkich k-elementowych podzbiorach J zbioru $\{1, \ldots, n\}$.

74 Dla $n = m + k$ dane są dowolne macierze $A \in \mathbb{K}^{m\times n}$, $B \in \mathbb{K}^{k\times n}$, $C \in \mathbb{K}^{n\times m}$, $D \in \mathbb{K}^{n\times k}$, $M \in \mathbb{K}^{m\times k}$, $N \in \mathbb{K}^{k\times m}$. Wykazać, że zachodzą następujące równości dla wyznaczników w postaci blokowej:

$$\begin{vmatrix} A \\ B \end{vmatrix} = \begin{vmatrix} A+MB \\ B \end{vmatrix} = \begin{vmatrix} A \\ B+NA \end{vmatrix}\,, \quad \begin{vmatrix} C & D \end{vmatrix} = \begin{vmatrix} C & D+CM \end{vmatrix} = \begin{vmatrix} C+DN & D \end{vmatrix}\,.$$

75 Wykazać, że dla dowolnych macierzy $M \in \mathbb{K}^{m\times k}$, $N \in \mathbb{K}^{k\times m}$ zachodzi równość $\det(\mathbf{1}_m + MN) = \det(\mathbf{1}_k + NM)$.

76 Wyliczyć macierz odwrotną.

$$\text{(a)} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}, \ \text{(b)} \begin{pmatrix} 1+i & 1+2i & 1+3i \\ 0 & 1+i & 1+2i \\ 0 & 0 & 1+i \end{pmatrix}, \ \text{(c)} \begin{pmatrix} 0 & -1 & -2 & -3 \\ -3 & 0 & 1 & 2 \\ -2 & 3 & 0 & 1 \\ -1 & 2 & 3 & 0 \end{pmatrix}.$$

77 Wykazać, że następujące układy równań są układami Cramera. Rozwiązać je stosując wzory Cramera.

$$\text{(a)}\ \begin{pmatrix} 1 & -2 & -3 \\ -3 & 5 & 7 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ -5 \\ 3 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & 2 & 0 & 5 \\ 2 & 1 & 3 & 0 \\ 0 & 3 & 4 & 1 \\ 1 & 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 2 \end{pmatrix},$$

$$\text{(c)}\ \begin{pmatrix} i & 1 & 1-i \\ 1+i & i & 1 \\ 2 & 3-i & -3i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2+i \\ 1+3i \\ i \end{pmatrix}.$$

78 Macierze kwadratowe $A, B \in \mathbb{K}^{n\times n}$ związane są warunkami: $B_{ij} = \mu_i A_{ij} \nu_j$, $i, j = 1, \ldots, n$. Jaki związek spełniają ich macierze odwrotne, jeśli istnieją?

79 Dane są macierze $A \in \mathbb{K}^{m\times m}$ i $B \in \mathbb{K}^{n\times n}$. Wykazać wzór na wyznacznik blokowy:

$$\begin{vmatrix} A_{11}B & A_{12}B & \ldots & A_{1m}B \\ A_{21}B & A_{22}B & \ldots & A_{2m}B \\ \ldots & \ldots & \ldots & \ldots \\ A_{m1}B & A_{m2}B & \ldots & A_{mm}B \end{vmatrix} = (\det A)^n (\det B)^m .$$

80 Uzupełnić dowód punktu (iii) twierdzenia 10, §7.

81 Wykazać, że wyznacznik macierzy, jako wielomian względem jej elementów, jest wielomianem pierwszym.

§8 Podstawowe pojęcia

82 Sprawdzić, czy podzbiory zbioru macierzy $\mathbb{K}^{n\times n}$ złożone z macierzy spełniających następujące warunki tworzą podprzestrzenie wektorowe:

(a) $A^2 = 0$;
(b) $\det A = 0$;
(c) $\operatorname{Tr} A = 0$;
(d) macierze trójkątne;
(e) $AM = MA$ dla ustalonej macierzy M.

Znaleźć bazy odpowiednich podprzestrzeni (w punkcie (e) dla $M = E_{ij}$ z bazy kanonicznej).

83 Sprawdzić, które ze zbiorów macierzy jednokolumnowych X z przestrzeni $\mathbb{K}^{n\times 1}$ tworzą podprzestrzenie.

(a) $X^{\mathrm{T}}X = 0$;

(b) $AX = B$, gdzie A i B są zadanymi macierzami;

(c) $Y^{\mathrm{T}}X = 0$, gdzie Y jest zadaną kolumną.

84 Wykazać, że zbiór wielomianów $f_i(x) = (x-a)^i$, $i = 0, \dots, n$, jest bazą przestrzeni wielomianów zespolonych stopnia mniejszego lub równego n. Znaleźć macierz przejścia do tej bazy z bazy $e_i(x) = x^i$, oraz macierz do niej odwrotną.

85 Zbiór odwzorowań $\mathbb{N} \ni k \mapsto c_k \in \mathbb{R}$, czyli nieskończonych ciągów rzeczywistych, tworzy przestrzeń wektorową. Sprawdzić, czy następujące podzbiory tych ciągów są podprzestrzeniami wektorowymi:

(a) ciągi zbieżne;

(b) takie, dla których $\displaystyle\sum_{k=1}^{\infty} |c_k| < \infty$;

(c) ciągi o skończonej liczbie wyrazów różnych od zera.

Czy zbiór ciągów, które mają tylko jeden wyraz równy 1, a pozostałe zerowe, tworzy bazę którejś z przestrzeni?

86 Rozważamy przestrzeń wektorową funkcji zespolonych na przedziale $\langle 0, 2\pi\rangle \subseteq \mathbb{R}$. Wykazać, że zbiór funkcji $f_k(x) = \exp[ikx]$, $k \in \mathbb{Z}$, tworzy w tej przestrzeni układ wektorów liniowo niezależnych, który jednak nie jest bazą (w sensie algebraicznym).

87 Niech F będzie przestrzenią wszystkich funkcji na zbiorze X o wartościach w ciele $\mathbb{K}$.

(a) Wykazać, że zbiór takich funkcji f, że zbiór $f(X)$ jest skończony, tworzy podprzestrzeń przestrzeni F.

(b) Wykazać, że zbiór funkcji, z których każda nie znika tylko w skończonej liczbie punktów, jest podprzestrzenią przestrzeni z poprzedniego punktu. Znaleźć bazę tej podprzestrzeni.

Przy jakim warunku te dwie podprzestrzenie pokrywają się?

88 Niech V będzie przestrzenią nad ciałem liczb zespolonych.

(a) Wykazać, że jeśli zawęzić zakres liczb do $\mathbb{R}$, to V jest z niezmienionymi działaniami rzeczywistą przestrzenią wektorową; oznaczmy ją $V_{\mathbb{R}}$.

(b) Jak najprościej skonstruować bazę przestrzeni $V_{\mathbb{R}}$, jeśli dana jest baza $(e_1, \dots, e_n)$ przestrzeni V?

Przedyskutować oba punkty w przypadku przestrzeni macierzowej.

89 Dane są rozkłady wektorów w bazie $(e_1, \ldots, e_n)$ przestrzeni rzeczywistej lub zespolonej. Sprawdzić, czy wektory $(f_1, \ldots, f_n)$ również tworzą bazę; jeśli tak, to znaleźć rozkład wektora x w tej bazie.

$$\text{(a)}\quad \begin{aligned} f_1 &= e_1\,, \\ f_2 &= \alpha e_1 + e_2\,, \\ f_3 &= \alpha e_2 + e_3\,, \\ f_4 &= \alpha e_3 + e_4\,, \\ x &= e_1 - 2e_2 + 3e_3 - 4e_4\,. \end{aligned} \qquad \text{(b)}\quad \begin{aligned} f_1 &= 2e_1 + 3e_2 - e_3\,, \\ f_2 &= e_1 - 2e_2 + 3e_3\,, \\ f_3 &= 5e_1 + 2e_2 + 2e_3\,, \\ x &= e_1 - e_2 + 2e_3\,. \end{aligned}$$

$$\text{(c)}\quad \begin{aligned} f_1 &= (1+2i)e_1 + (2+3i)e_2\,, \\ f_2 &= (2-2i)e_1 + (3-2i)e_2\,, \\ x &= e_1 - ie_2\,. \end{aligned}$$

§9 Układy równań liniowych

90 Znaleźć ogólne rozwiązania, jeśli istnieją, układów równań określonych podanymi macierzami dołączonymi. Jeśli macierz zawiera parametr, to rozwiązać w zależności od wartości parametru.

$$\text{(a)}\ \begin{pmatrix} 2 & -1 & 3 & 5 & 7 \\ 5 & 2 & 7 & 8 & 9 \\ 2 & 8 & 2 & -4 & -10 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} i & 1-i & 2-3i & 5+i \\ 2-i & -3 & -3+i & 2-i \\ -3-4i & 2+7i & 7+3i & 7-4i \end{pmatrix},$$

$$\text{(c)}\ \begin{pmatrix} 3 & 2 & -3 \\ 4 & 3 & 1 \\ 5 & 4 & 2 \end{pmatrix}, \quad \text{(d)}\ \begin{pmatrix} 4 & -1 & -2 \\ -3 & 4 & 7 \\ -2 & 7 & 12 \\ 5 & 2 & 3 \end{pmatrix}, \quad \text{(e)}\ \begin{pmatrix} 2 & 3 & -1 & 4 & 2 \\ 3 & 2 & 3 & 1 & 5 \\ 1 & 4 & -5 & 7 & -1 \end{pmatrix},$$

$$\text{(f)}\ \begin{pmatrix} 3 & 2 & -1 & 4 \\ -2 & -5 & 7 & -2 \\ 4 & -1 & \lambda & \lambda+1 \end{pmatrix}, \quad \text{(g)}\ \begin{pmatrix} 1 & 2 & 4 & 3 & 1 & 0 \\ 3 & 1 & 2 & 4 & 3 & 5 \\ 5 & 6 & 7 & 3 & 1 & 1 \\ 6 & 9 & 8 & -1 & -2 & -3 \end{pmatrix},$$

$$\text{(h)}\ \begin{pmatrix} 1 & -2 & c & 2c \\ -2 & 1 & -2 & -c \\ c & -2 & 1 & c+1 \end{pmatrix}, \quad \text{(i)}\ \begin{pmatrix} \alpha & 1 & \alpha^2 & \alpha \\ 1 & \alpha & 1 & 1 \\ \alpha^2 & 1 & \alpha^2 & \alpha \end{pmatrix}, \quad \text{(j)}\ \begin{pmatrix} \alpha & 1 & \alpha \\ 1 & \alpha & \alpha \\ \alpha & \alpha & \alpha \end{pmatrix},$$

$$\text{(k)}\ \begin{pmatrix} 2 & 3 & 1 & 2 & -5 \\ 2 & 3 & -2 & 4 & -3 \\ 4 & 6 & 4 & \lambda & -7 \end{pmatrix}, \quad \text{(l)}\ \begin{pmatrix} 2 & 1 & 0 & 3 & 1 & 2 \\ 3 & 3 & 2 & 4 & 3 & 1 \\ 2 & 7 & 8 & 1 & 7 & -6 \\ 4 & 1 & 2 & 3 & 0 & 5 \end{pmatrix}.$$

91 Ciąg kolumn $X_1, \dots, X_n$ ustawiono w macierz. Znaleźć maksymalną rodzinę liniowo niezależnych spośród tych kolumn i wyrazić pozostałe jako ich kombinacje liniowe.

$$\text{(a)} \begin{pmatrix} 2 & 3 & 1 & 1 & 6 \\ 3 & 2 & 4 & 2 & 7 \\ -1 & 2 & -4 & 3 & 4 \\ 2 & -1 & 5 & 4 & 5 \end{pmatrix}, \text{(b)} \begin{pmatrix} 5 & 3 & 7 & 5 & 2 \\ 7 & 4 & 10 & 8 & 2 \\ 1 & 2 & 0 & 4 & -4 \end{pmatrix}, \text{(c)} \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix},$$

$$\text{(d)} \begin{pmatrix} -2 & 3 & 0 & 10 \\ 1 & -2 & -1 & -7 \\ 4 & -3 & 6 & -8 \\ -5 & 4 & -7 & 11 \\ 0 & 1 & 2 & 4 \end{pmatrix}, \text{(e)} \begin{pmatrix} 2 & 2 & 3 & 5 & 1 \\ 1 & -1 & 0 & 2 & 2 \\ 3 & -9 & -1 & 1 & 7 \\ 5 & -8 & 7 & 0 & 3 \\ 1 & -2 & -2 & 3 & 4 \end{pmatrix}, \text{(f)} \begin{pmatrix} 3 & 1 & 3 \\ 1 & -2 & 8 \\ 2 & 1 & 1 \\ 1 & 3 & -7 \end{pmatrix}.$$

92 Wyliczyć macierze odwrotne metodą Gaussa.

$$\text{(a)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} i & 0 & 0 \\ 0 & 1 & i \\ 1-i & 1 & 1 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

93 Wykazać, że rząd macierzy $A \in \mathbb{K}^{m \times n}$ jest równy 1 wtedy, i tylko wtedy, gdy istnieją kolumny $X \in \mathbb{K}^m$ i $Y \in \mathbb{K}^n$ takie, że $A = XY^{\mathrm{T}}$.

94 Dane są dwa układy równań $AX = B$ i $A'X' = B'$, gdzie $X, X' \in \mathbb{K}^n$. Znaleźć warunek konieczny i wystarczający na to, by ogólne rozwiązania tych układów były identyczne (jako zbiory).

95 Znaleźć wszystkie wielomiany rzeczywiste P stopnia ≤ 3, dla których $P(1) = a$, $P(0) = b$ i $P(-1) = c$, gdzie a, b, c są zadanymi liczbami rzeczywistymi.

§10 Odwzorowania liniowe

96 Odwzorowanie liniowe $A \in \mathcal{L}(V, W)$, przy wybranych bazach $(e_1, \dots, e_m)$ i $(f_1, \dots, f_n)$ przestrzeni V i W odpowiednio, zadaje poniższe związki. Odczytać jego macierz w zadanych bazach. Znaleźć współrzędne rozkładu wektora $y = Ax$ w bazie (f_i) (w obu punktach $n = 3$, a $m = 4, 2$ odpowiednio).

$$\text{(a)}\ \begin{aligned} Ae_1 &= f_1 + 3f_2 - 5f_4\,,\\ Ae_2 &= -2f_1 + f_2 + 5f_3\,,\\ Ae_3 &= 3f_1 + 2f_2 - f_3 - f_4\,,\\ x &= 2e_1 + e_2 - 4e_3\,. \end{aligned} \qquad \text{(b)}\ \begin{aligned} Ae_1 &= (3-i)f_1 - if_2\,,\\ Ae_2 &= -if_1 + (4+5i)f_2\,,\\ Ae_3 &= if_1 - f_2\,,\\ x &= (3-i)e_1 + ie_2\,. \end{aligned}$$

97 Przy oznaczeniach poprzedniego zadania dana jest macierz odwzorowania A w wybranych bazach. Znaleźć bazę jądra i bazę obrazu tego odwzorowania, oraz macierze przejścia do baz, w których przyjmuje ono postać kanoniczną.

$$\text{(a)}\ \begin{pmatrix} 3 & 1 & 5 & 7 & -3\\ 1 & 2 & 0 & 2 & 1\\ 2 & 5 & -1 & 1 & 6\\ 4 & 6 & 2 & 5 & 5 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & i & 0 & i\\ i & 2 & 3i & 1\\ 1+i & 1-i & 2+2i & i\\ -i & 3 & 2i & 1 \end{pmatrix}.$$

98 Dane jest działanie operatora liniowego na wektory bazy $(e_1, \dots, e_n)$ oraz macierz przejścia do bazy $(e'_1, \dots, e'_n)$. Wyliczyć macierz operatora w bazie (e'_i).

$$\text{(a)}\ \begin{aligned} Ae_1 &= 3e_1 - 2e_2 + 5e_3\,,\\ Ae_2 &= -e_1 + 7e_2 + 2e_3\,,\\ Ae_3 &= e_1 + e_2 - e_3\,, \end{aligned} \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 & 3 & -1\\ 0 & 1 & 0\\ 0 & 0 & -1 \end{pmatrix},$$

$$\text{(b)}\ \begin{aligned} Ae_1 &= -ie_1 + e_2\,,\\ Ae_2 &= 2ie_1 - e_2\,, \end{aligned} \qquad \boldsymbol{\beta} = \begin{pmatrix} i & i\\ 1 & 2+i \end{pmatrix}.$$

99 W przestrzeniach rzeczywistych V i W wybrano bazy (e_1, e_2, e_3, e_4) i (f_1, f_2, f_3) odpowiednio. Odwzorowanie liniowe $A \in \mathcal{L}(V, W)$ ma w tych bazach macierz

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 5 & 4\\ 2 & 1 & 3 & 2\\ 5 & 4 & 9 & 8 \end{pmatrix}.$$

Znaleźć wszystkie wektory x spełniające równanie $Ax = f_1 + 2f_2 + \lambda f_3$ (w zależności od wartości parametru λ).

100 W trójwymiarowej przestrzeni zespolonej V wybrano bazę (e_1, e_2, e_3) oraz operator działający na wektory tej bazy według przepisu:

$$Ae_1 = e_1 + ie_2\,, \quad Ae_2 = ie_1 + e_3\,, \quad Ae_3 = z(z-3+i)e_1 - 4ie_2 + 3e_3\,,$$

gdzie z jest zespolonym parametrem. Znaleźć bazę jądra oraz bazę obrazu operatora A w zależności od wartości parametru.

101 Operator A ma w bazie $(e_1, \ldots, e_n)$ macierz $\mathbf{A}$. Jakiej operacji należy poddać tę macierz, aby otrzymać macierz operatora A w bazie $(e_{\pi(1)}, \ldots, e_{\pi(n)})$, gdzie π jest permutacją?

102 Operator A działa na wektory danej bazy zgodnie z przepisem $Ae_i = e_{\pi(i)}$, gdzie π jest permutacją. Jaka jest macierz tego operatora w bazie (e_i)? Wyliczyć jego wyznacznik.

103 Operator A działa w n-wymiarowej przestrzeni V. Wykazać, że dla każdego niezerowego wektora x w tej przestrzeni istnieje niezerowy wielomian W stopnia mniejszego lub równego n, dla którego $W(A)x = 0$. Jaki jest związek minimalnego stopnia takiego wielomianu z wymiarem podprzestrzeni generowanej działaniem potęg operatora A na ten wektor (zgodnie z przykładem (ii), p. 7, §12)?

104 Niech przestrzeń V nad dowolnym ciałem ma skończony wymiar. Wykazać, że podzbiór przestrzeni $\mathcal{L}(V, V)$ złożony ze wszystkich operatorów o zerowym śladzie jest podprzestrzenią. Znaleźć bazę tej podprzestrzeni.

105 Niech $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ będzie nieosobliwą macierzą rzeczywistą. Określamy operator T działający w przestrzeni $\mathbb{R}^{2\times 2}$ zgodnie z regułą: $T(X) = A^{-1}XA$. Wyliczyć macierz tego operatora w bazie kanonicznej uszeregowanej leksykograficznie, oraz jego wyznacznik.

106 Niech niezerowa macierz $A \in \mathbb{R}^{2\times 2}$ ma zerowy wyznacznik. Określamy operator S działający w przestrzeni rzeczywistych, symetrycznych macierzy 2×2 zgodnie z regułą $S(X) = A^{\mathrm{T}}XA$. Wyliczyć obraz i jądro tego operatora, znaleźć bazy tych podprzestrzeni.

§11 Grupy operatorowe. Orientacja

107 W bazie (e_1, e_2, e_3) przestrzeni rzeczywistej dane są rozkłady dwóch ciągów wektorów. Wykazać, że każdy z nich tworzy bazę i znaleźć ich względną orientację (bez konieczności wyliczenia macierzy przejścia między tymi bazami).

$$\begin{aligned} f_1 &= e_1 + 4e_2 + 3e_3\,, & g_1 &= 2e_1 + e_2 + 3e_3\,,\\ f_2 &= 2e_1 + 2e_2 + e_3\,, & g_2 &= e_1 + 3e_2 + 5e_3\,,\\ f_3 &= 3e_1 + 5e_2\,, & g_3 &= 7e_1 + 2e_2 + 2e_3\,. \end{aligned}$$

108 Znaleźć ciągłą krzywą w grupie $GL(3,\mathbb{R})$ łączącą podaną macierz z macierzą jednostkową lub z macierzą diagonalną o wyrazach $(-1,1,1)$ na diagonali.

$$\text{(a)}\ \begin{pmatrix} 4 & 2 & 1\\ 3 & 0 & 2\\ 4 & 3 & 1 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & 3 & 6\\ 1 & 4 & 7\\ 1 & 5 & 9 \end{pmatrix}.$$

109 Wykazać, że grupa $SL(n,\mathbb{R})$ jest spójna. Wykazać, że relacja w zbiorze baz rzeczywistej przestrzeni n-wymiarowej określona warunkiem, by macierz przejścia była elementem $SL(n,\mathbb{R})$, jest relacją równoważności.

110 Niech π będzie permutacją z grupy S_n. Kiedy bazy $(e_1, \dots, e_n)$ i $(e_{\pi(1)}, \dots, e_{\pi(n)})$ przestrzeni rzeczywistej mają jednakową orientację?

§12 Sumy proste przestrzeni i operatorów. Operatory rzutowe

111 Sprawdzić, czy dla podprzestrzeni przestrzeni wektorowej zachodzi tożsamość $(W + W') \cap U = W \cap U + W' \cap U$. Odpowiedź uzasadnić.

112 W bazie $(e_1, \dots, e_n)$ przestrzeni rzeczywistej dane są rozkłady wektorów. Znaleźć bazy podprzestrzeni $V_1 = L(a_1, \dots, a_k)$, $V_2 = L(b_1, \dots, b_m)$, $V_1 \cap V_2$ oraz bazę $V_1 + V_2$ zgodną z pewnym rozkładem $V_1 + V_2 = V_1 \cap V_2 \oplus W_1 \oplus W_2$, gdzie $V_i = V_1 \cap V_2 \oplus W_i$, $i = 1, 2$.

$$\text{(a)}\quad \begin{aligned} a_1 &= e_1 + e_2 + e_3 + e_4 + e_5\,,\\ a_2 &= 6e_1 + e_2 - e_3 + 6e_4 + 3e_5\,, \end{aligned} \qquad \begin{aligned} b_1 &= e_1 + 3e_2 + 2e_3 - e_4 + 4e_5\,,\\ b_2 &= 5e_1 + 2e_2 - e_3 + 3e_4 + 5e_5\,,\\ b_3 &= 3e_1 - 4e_2 - 5e_3 + 5e_4 - 3e_5\,; \end{aligned}$$

$$\text{(b)}\quad \begin{aligned} a_1 &= e_1 + 2e_2 + 3e_3 - e_4\,,\\ a_2 &= 2e_1 + e_2 - 3e_3 + 3e_4\,,\\ a_3 &= -e_1 + e_2 + 2e_3 + e_4\,, \end{aligned} \qquad \begin{aligned} b_1 &= 2e_1 + e_2 + e_3 - 2e_4\,,\\ b_2 &= 4e_1 + 5e_2 + 3e_3 + 2e_4\,,\\ b_3 &= 8e_1 + 7e_2 - 3e_3 + 7e_4\,. \end{aligned}$$

113 W bazie $(e_1, \dots, e_4)$ przestrzeni rzeczywistej V dane są rozkłady wektorów $(f_1, \dots, f_4)$. Wykazać, że $V = L(f_1, f_2) \oplus L(f_3, f_4)$. Wyliczyć macierze operatorów rozkładu identyczności związanego z tym rozbiciem przestrzeni w bazie (e_i).

$$\begin{aligned} f_1 &= e_1 + 3e_2 + 2e_3 + e_4 \,, \quad & f_2 &= 2e_1 + e_2 + 2e_3 + e_4 \,, \\ f_3 &= 3e_1 - e_2 - 2e_3 + e_4 \,, \quad & f_4 &= 2e_2 - 2e_3 \,. \end{aligned}$$

114 Niech $\mathbb{R}_0^{n\times n}$ oznacza podprzestrzeń przestrzeni $\mathbb{R}^{n\times n}$ złożoną z macierzy o zerowym śladzie. Wykazać, że $\mathbb{R}^{n\times n} = \mathbb{R}_0^{n\times n} \oplus L(\mathbf{1})$.

115 Niech A będzie operatorem w dowolnej przestrzeni V.

(a) Wykazać, że następujące warunki są równoważne: (i) $\operatorname{Ker} A^2 = \operatorname{Ker} A$, (ii) $\operatorname{Ker} A \cap \operatorname{Im} A = \{0\}$, (iii) istnieje podprzestrzeń W taka, że $V = \operatorname{Ker} A \oplus W$ i $AW \subseteq W$.

(b) Wykazać, że wówczas $\operatorname{Im} A \subseteq W$, a jeśli $\dim V < \infty$, to $\operatorname{Im} A = W$.

116 Niech A będzie operatorem w przestrzeni V nad dowolnym ciałem i niech będzie dany rozkład $V = W_1 \oplus W_2$. Odpowiedni rozkład wektora oznaczmy $x = x_1 + x_2$.

(a) Wykazać, że odwzorowania $A_{ji} : W_i \mapsto W_j$, $A_{ji}(y) = (Ay)_j$ są liniowe.

(b) Wybrać bazę przestrzeni V zgodną z jej przyjętym rozkładem na sumę prostą. Znaleźć związek macierzy odwzorowań A_{ji} z macierzą operatora A.

117 Działanie operatora A na wektory bazy $(e_1, \dots, e_n)$ zadane jest przepisem: $Ae_1 = 0$, $Ae_k = e_{k-1}$, $k = 2, \dots, n$. Znaleźć wszystkie podprzestrzenie niezmiennicze tego operatora. Wykazać, że nie istnieje rozkład tego operatora na sumę prostą dwóch operatorów.

118 Wykazać, że dla operatora A w dowolnej przestrzeni następujące warunki są równoważne: (i) A nie ma nietrywialnych podprzestrzeni inwariantnych; (ii) każdy niezerowy wektor w przestrzeni jest wektorem cyklicznym operatora A.

119 Niech $A \in \mathcal{L}(V, V)$ i $\dim V < \infty$. Nie odwołując się do twierdzenia Cayley-Hamiltona wykazać następujące stwierdzenia:

(a) istnieje wielomian Q taki, że $Q(A) = 0$;

(b) stopień wielomianu w poprzednim stwierdzeniu może być ograniczony wymiarem przestrzeni;

(c) jeśli $M(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \ldots + a_0$ jest wielomianem najniższego stopnia dla którego $M(A) = 0$, to warunkiem koniecznym i wystarczającym istnienia A^{-1} jest $a_0 \neq 0$.

120 Wykazać, że jeśli operatory $P_1, \ldots, P_s$ spełniają warunki $P_i P_j = \delta_{ij} P_i$, $i, j = 1, \ldots, s$, to wspólnie z operatorem $P_{s+1} = \mathrm{id} - \sum_{i=1}^{s} P_i$ tworzą rozkład identyczności.

121 Wykazać, że każde dwa niezerowe, różne operatory rzutowe (w przestrzeni nad dowolnym ciałem) są liniowo niezależne.

122 Niech P i Q będą różnymi, niezerowymi operatorami rzutowymi. Pytamy, kiedy ich kombinacja liniowa $R = \alpha P + \beta Q$ jest również operatorem rzutowym. Jeśli co najmniej jeden ze współczynników kombinacji jest równy zero, to na mocy poprzedniego zadania $R = 0$ lub $R = P$ lub $R = Q$. Wykazać, że następujące przypadki opisują wszystkie rozwiązania z niezerowymi współczynnikami: (i) jeśli $PQ = QP = 0$, to $R = P + Q$; (ii) jeśli $PQ = QP = P$, to $R = Q - P$; (iii) jeśli $PQ = QP = Q$, to $R = P - Q$; (iv) jeśli $PQ + QP = P + Q$, to $R = \alpha P + (1 - \alpha) Q$, $\alpha \in \mathbb{K} \setminus \{0, 1\}$.

123 Niech A będzie automorfizmem przestrzeni V i niech W będzie jego inwariantną podprzestrzenią. Oznaczmy przez A_W zacieśnienie operatora A do operatora na przestrzeni W.

(a) Wykazać, że A_W jest injektywny, ale nie musi być surjektywny (gdy W ma nieskończony wymiar). Podać odpowiedni kontrprzykład.

(b) Wykazać, że jeśli W ma skończony wymiar, to A_W jest automorfizmem.

124 Wykazać, że jeśli podprzestrzeń W_1 w zadaniu 116 jest inwariantna względem A, to dla każdego wielomianu P zachodzi $[P(A)]_{ii} = P(A_{ii})$, $i = 1, 2$.

125 Wykazać, że jeśli operator w skończenie wymiarowej przestrzeni spełnia $A^k = 0$ dla pewnej naturalnej liczby k, to istnieje baza, w której jego macierz jest trójkątna, z zerami na diagonali. W szczególności, jedyną wartością własną jest 0 oraz $\mathrm{Tr}\, A = 0$.

§13 Zagadnienie własne operatora liniowego

126 Dana jest macierz operatora A w bazie (e_1, e_2, e_3) przestrzeni rzeczywistej. Wyliczyć wartości i wektory własne tego operatora. Jeśli operator jest diagonalizowalny, to znaleźć macierz przejścia do bazy złożonej z wektorów własnych. Podać związek baz i macierz operatora w nowej bazie.

$$\text{(a)} \begin{pmatrix} 1 & -8 & 2 \\ -5 & 7 & 1 \\ 2 & 4 & 4 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} 0 & -4 & 6 \\ -3 & 4 & -9 \\ 5 & -10 & 13 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} -56 & -105 & 108 \\ 9 & 16 & -18 \\ -21 & -40 & 40 \end{pmatrix}.$$

127 Operator A ma w bazie $(e_1, \ldots, e_n)$ przestrzeni rzeczywistej lub zespolonej macierz o elementach $A^i{}_j = \delta^{n-i+1}{}_j \lambda_j$. Wyliczyć wartości i wektory własne tego operatora. Sprawdzić, kiedy operator jest diagonalizowalny i podać w tym przypadku transformację do bazy własnej.

128 Operator A w przestrzeni zespolonej ma w bazie $(e_1, \ldots, e_n)$ macierz

$$\mathbf{A} = \begin{pmatrix} z_1 & z_2 & \ldots & z_{n-1} & z_n \\ z_2 & z_3 & \ldots & z_n & z_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ z_{n-1} & z_n & \ldots & z_{n-3} & z_{n-2} \\ z_n & z_1 & \ldots & z_{n-2} & z_{n-1} \end{pmatrix} .$$

Wyliczyć jego działanie na wektory $f(\varepsilon) = \sum_{i=1}^n \varepsilon^{i-1} e_i$, gdzie ε jest n-tym pierwiastkiem z jedynki. Korzystając z wyniku wyliczyć wszystkie wektory własne operatora. Kiedy operator jest diagonalizowalny?

129 Niech A będzie operatorem w skończenie wymiarowej przestrzeni nad dowolnym ciałem, spełniającym warunek $A^2 = 0$.

(a) Znaleźć spektrum tego operatora. Czy A może być diagonalizowalny (poza przypadkiem $A = 0$)?

(b) Niech $(e_1, \ldots, e_m)$ będą takimi wektorami, że wektory $(Ae_1, \ldots, Ae_m)$ tworzą bazę $\operatorname{Im} A$. Uzupełniamy ten ciąg wektorów do bazy jądra wektorami $f_1, \ldots, f_k$. Wykazać, że ciąg wektorów $(Ae_1, \ldots, Ae_m, f_1, \ldots, f_k, e_1, \ldots, e_m)$ tworzy bazę i wyliczyć macierz operatora A w tej bazie.

130 Wykazać, że jeśli operator działający w trójwymiarowej przestrzeni nad dowolnym ciałem ma podprzestrzeń własną o wymiarze 2 do wartości własnej będącej dwukrotnym pierwiastkiem wielomianu charakterystycznego, to jest diagonalizowalny.

131 Wykazać, że jeśli co najmniej jedna z przestrzeni W_i w zadaniu 116 jest inwariantna względem operatora A i cała przestrzeń ma skończony wymiar, to wielomiany charakterystyczne odpowiednich operatorów spełniają związek $P_A = P_{A_{11}} P_{A_{22}}$.

132 Wykazać, że dla każdego operatora w skończenie wymiarowej przestrzeni zespolonej istnieje baza, w której jego macierz jest trójkątna.

133 Niech D będzie operatorem różniczkowania na przestrzeni funkcyjnej $\mathcal{C}^\infty_\mathbb{R}$ (patrz przykład (ii) w punkcie 2, §10). Znaleźć wszystkie wektory własne operatorów D i D^2. Wykazać, że w żadnym z tych dwóch przypadków ich powłoka nie jest całą przestrzenią.

134 Niech A będzie operatorem działającym w przestrzeni nad dowolnym ciałem, a P – dowolnym wielomianem na tym ciele. Jeśli λ jest wartością własną operatora A, to $P(\lambda)$ jest wartością własną operatora $P(A)$. Wykazać, że w przestrzeni zespolonej o skończonym wymiarze zachodzi również twierdzenie odwrotne: jeśli μ jest wartością własną operatora $P(A)$, to $\mu = P(\lambda)$ dla pewnej wartości własnej λ operatora A. Ostatnie twierdzenie nie zachodzi w przestrzeni rzeczywistej; podać odpowiedni kontrprzykład.

135 Wykazać, że jeśli operator w dwuwymiarowej przestrzeni zespolonej ma tylko jedną wartość własną, to albo jest operatorem proporcjonalnym do identycznościowego, albo istnieje baza, w której jego macierz ma postać $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

136 Niech funkcja zespolona (lub rzeczywista) na osi rzeczywistej dana będzie jako suma szeregu o nieskończonym promieniu zbieżności. Wykazać, że jeśli operator w dwuwymiarowej przestrzeni zespolonej (lub rzeczywistej) ma w bazie (e_1, e_2) macierz $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, to macierz operatora $f(A)$ ma w tej samej bazie postać $\begin{pmatrix} f(\lambda) & f'(\lambda) \\ 0 & f(\lambda) \end{pmatrix}$, gdzie f' jest pochodną funkcji f. Jak wynik zmieni się, jeśli element macierzy operatora A równy 1 zastąpić dowolnym parametrem μ?

137 Dany jest operator B w dwuwymiarowej przestrzeni zespolonej. Znaleźć wszystkie operatory A będące rozwiązaniami równania $f(A) = B$ (jeśli istnieją dla danego B), gdzie (a) $f = \exp$, (b) $f = \sin$.

138 Niech A będzie operatorem diagonalizowalnym w przestrzeni rzeczywistej lub zespolonej, o rozkładzie spektralnym $A = \sum_{i=1}^{s} \lambda_i P_i$, i niech λ_1 będzie wartością własną o maksymalnym module (lub jedną z takich wartości). Wykazać, że wtedy istnieje granica $\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} (A/\lambda_1)^k = P_1$.

§14 Iloczyny skalarne

139 Wykazać, że jeśli w przestrzeni z iloczynem skalarnym podprzestrzenie W i U spełniają $U \subseteq W$, to $W^\perp \subseteq U^\perp$.

140 Jeśli iloczyn skalarny na skończenie wymiarowej przestrzeni V jest niezdegenerowany, to dla każdej podprzestrzeni $W \subseteq V$ zachodzi $(W^\perp)^\perp = W$.

Wykazać ten fakt w przypadku, gdy W jest niezdegenerowana względem iloczynu (dyskusję ogólnego przypadku można znaleźć w punkcie 11, §19).

141 Metryka g na przestrzeni V ma w bazie $(e_1,\dots,e_n)$ podaną niżej macierz. Znaleźć jądro tej metryki, jego bazę, oraz bazę podprzestrzeni W takiej, że $V = W \oplus \operatorname{Ker} g$. Wyliczyć macierz metryki w bazie zgodnej z tym rozkładem.

$$\text{(a)}\ \begin{pmatrix} 0 & 3 & -2 \\ -3 & 0 & 5 \\ 2 & -5 & 0 \end{pmatrix}, \ \text{(b)}\ \begin{pmatrix} -8 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}, \ \text{(c)}\ \begin{pmatrix} 3 & 1 & 6 & 0 \\ 1 & -5 & 2 & -4 \\ 6 & 2 & 12 & 0 \\ 0 & -4 & 0 & -3 \end{pmatrix}.$$

142 Dana jest macierz $\mathbf{g}$ metryki symplektycznej g w bazie $(e_1,\dots,e_4)$. Znaleźć macierz przejścia do bazy, w której metryka przyjmuje postać kanoniczną.

$$\text{(a)}\ \begin{pmatrix} 0 & 1 & -2 & 3 \\ -1 & 0 & 4 & -5 \\ 2 & -4 & 0 & 3 \\ -3 & 5 & -3 & 0 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 0 & -2 & 1 & 1 \\ 2 & 0 & -3 & 0 \\ -1 & 3 & 0 & 4 \\ -1 & 0 & -4 & 0 \end{pmatrix}.$$

143 Dana jest macierz metryki w bazie $(e_1,\dots,e_4)$ oraz układ wektorów (f_1, f_2). Sprawdzić, czy podprzestrzeń $L(f_1, f_2)$ jest niezdegenerowana. Jeśli tak, to znaleźć dowolną bazę (f_3, f_4) jej ortogonalnego dopełnienia i macierz metryki w bazie $(f_1,\dots,f_4)$.

$$\text{(a)}\ \begin{pmatrix} 2 & 1 & 4 & 0 \\ 1 & 3 & 0 & 2 \\ 4 & 0 & 3 & 1 \\ 0 & 2 & 1 & 3 \end{pmatrix}, \qquad \text{(b)}\ \begin{pmatrix} 1 & 2 & -1 & -2 \\ 2 & 3 & 2 & 1 \\ -1 & 2 & 4 & 3 \\ -2 & 1 & 3 & 0 \end{pmatrix},$$

$$f_1 = e_1\,,\ f_2 = e_2\,; \qquad f_1 = e_1 + e_3\,,\ f_2 = e_2 - e_4\,.$$

144 Niech g będzie dowolną formą biliniową lub półtoraliniową na przestrzeni V, a $(f_1,\dots,f_k)$ – dowolnym ciągiem wektorów w V. Wykazać, że jeśli macierz $(g(f_i, f_j))$ jest nieosobliwa, to wektory (f_i) są liniowo niezależne.

145 Niech g będzie iloczynem hermitowskim na zespolonej przestrzeni V. Określamy funkcje $g_+(x,y) = \mathfrak{Re}[g(x,y)]$, $g_-(x,y) = \mathfrak{Im}[g(x,y)]$.

(a) Wykazać, że jako funkcje na $V_{\mathbb{R}} \times V_{\mathbb{R}}$ (zadanie 88) są one metrykami biliniowymi na $V_{\mathbb{R}}$, przy tym g_+ jest metryką symetryczną, a g_- – metryką symplektyczną.

(b) Niech wymiar przestrzeni będzie skończony i niech $(e_1, \ldots, e_n)$ będzie bazą V, w której metryka g przyjmuje postać kanoniczną. Podać macierze metryk $g_\pm$ w bazie przestrzeni $V_\mathbb{R}$, skonstruowanej tak, jak w zadaniu 88.

§15 Przestrzenie ortogonalne i hermitowskie

146 W bazie $(e_1, \ldots, e_n)$ przestrzeni rzeczywistej forma kwadratowa metryki symetrycznej ma podaną postać. Wyliczyć macierz przejścia do bazy $(e'_1, \ldots, e'_n)$, w której forma przyjmuje postać kanoniczną. Podać jej macierz w nowej bazie.
(a) $g(x,x) = (x^1)^2 + (x^2)^2 + 9(x^4)^2 + 2x^1x^2 + 6x^1x^4 + 2x^2x^3$;
(b) $g(x,x) = x^1x^2 - 2x^2x^3 + 4x^1x^3$;
(c) $g(x,x) = (x^1)^2 + 4(x^2)^2 + (x^3)^2 - 4x^1x^2 + 2x^1x^3 - 3x^2x^3 - 4x^3x^4$;
(d) $g(x,x) = 4(x^1)^2 + (x^2)^2 + (x^4)^2 - 4x^1x^2 + 4x^1x^4 - 3x^2x^4$.

147 Podobnie jak w poprzednim zadaniu, dla formy hermitowskiej w przestrzeni zespolonej.
(a) $g(x,x) = \overline{x^1}x^1 + 2\overline{x^2}x^2 + i\overline{x^1}x^2 - i\overline{x^2}x^1 + (1+i)\overline{x^2}x^3 + (1-i)\overline{x^3}x^2$;
(b) $g(x,x) = (1-i)\overline{x^1}x^2 + (1+i)\overline{x^2}x^1 + i\overline{x^1}x^3 - i\overline{x^3}x^1 + \overline{x^2}x^3 + \overline{x^3}x^2$.

148 Forma kwadratowa metryki g w bazie (e_1, e_2, e_3) trójwymiarowej przestrzeni rzeczywistej ma postać $g(x,x) = (x^1 - x^2)^2 + (x^2 - x^3)^2 + (x^3 - x^1)^2$. Czy wynika stąd, że sygnatura metryki wynosi $(3,0,0)$? Zastosować do tej formy kwadratowej standardowy algorytm sprowadzania do postaci kanonicznej. Jaka jest przyczyna rozbieżności?

149 Wykazać, że jeśli rekurencyjny przepis w twierdzeniu o ortogonalizacji Grama-Schmidta w żadnym kroku nie prowadzi do równości $g(e_k, e_k) = 0$, to wektory $(f_1, \ldots, f_n)$ spełniają założenia twierdzenia, więc postępowanie rekurencyjne jest poprawne. Ponadto, twierdzenie może być stosowane do przestrzeni $L(f_1, \ldots, f_n)$, więc cała przestrzeń może mieć wymiar większy niż n.

150 Wykazać, że rezultat procedury Grama-Schmidta zależy od kolejności ortogonalizowanych wektorów. Podać przykład metryki i układu wektorów, dla których możliwość zastosowania procedury zależy od tej kolejności.

151 W bazie (e_i) przestrzeni rzeczywistej lub zespolonej dana jest macierz metryki symetrycznej lub, odpowiednio, hermitowskiej. Przeprowadzić ortogonalizację Grama-Schmidta tej bazy, podać macierz przejścia do nowej bazy i macierz metryki w nowej bazie.

$$(a)\ \begin{pmatrix} 2 & -1 & 5 \\ -1 & 0 & -2 \\ 5 & -2 & 3 \end{pmatrix},\ (b)\ \begin{pmatrix} 1 & 1+i & 1 \\ 1-i & 3 & 1-3i \\ 1 & 1+3i & 2 \end{pmatrix},\ (c)\ \begin{pmatrix} 3 & -2 & 5 \\ -2 & 1 & -1 \\ 5 & -1 & 0 \end{pmatrix}.$$

152 W pewnej bazie trójwymiarowej przestrzeni rzeczywistej dane są formy kwadratowe metryk symetrycznych zależnych od parametru λ. Stosując kryterium Sylwestra wyliczyć, dla jakich wartości parametru te metryki są dodatnio określone (jeśli takie wartości istnieją).

(a) $g(x,x) = 5(x^1)^2 + \lambda(x^2)^2 + (x^3)^2 + 2x^1x^2 + 4x^1x^3 - 6x^2x^3$;

(b) $g(x,x) = 2(x^1)^2 + (x^2)^2 + 2\lambda x^1x^2 - 2x^1x^3 + 2x^2x^3$.

153 Sprawdzić, dla jakich wartości parametru λ istnieje baza (e_i) przestrzeni euklidesowej złożona z unormowanych wektorów, dla których $e_i \cdot e_j = \lambda$ dla wszystkich $i \neq j$.

154 W bazie ortonormalnej (e_i) przestrzeni euklidesowej lub unitarnej dane są rozkłady układu wektorów. Wykazać, że te układy są liniowo niezależne i przeprowadzić dla nich ortogonalizację Grama-Schmidta, a następnie otrzymane wektory unormować.

$$(a)\quad \begin{aligned} f_1 &= e_1 + e_2 \\ f_2 &= e_1 - 2e_2 \\ f_3 &= 3e_1 - e_2 - e_3 \end{aligned} \qquad (b)\quad \begin{aligned} f_1 &= e_1 + e_3 \\ f_2 &= e_1 - 2e_2 - e_3 \\ f_3 &= -e_1 + e_4 \\ f_4 &= e_1 + e_2 + e_3 + e_4 \end{aligned}$$

$$(c)\quad \begin{aligned} f_1 &= ie_1 + (1-i)e_2 \\ f_2 &= e_1 + ie_2 - ie_3 \\ f_3 &= ie_2 + e_3 \end{aligned} \qquad (d)\quad \begin{aligned} f_1 &= ie_1 + e_3 \\ f_2 &= e_1 + ie_2 + e_3 - ie_4\,, \\ f_3 &= e_1 + ie_4 \end{aligned}$$

155 Wykazać, że dla każdych dwóch wektorów x i y w przestrzeni euklidesowej lub unitarnej spełniona jest tożsamość $\|x+y\|^2+\|x-y\|^2 = 2\|x\|^2+2\|y\|^2$.

156 Wykazać, że jeśli $(f_1, \ldots, f_k)$ jest dowolnym układem ortonormalnym w przestrzeni unitarnej lub euklidesowej, to dla każdego wektora x zachodzi następująca nierówność Bessela: $\|x\|^2 \geq \sum_{i=1}^{k} |(f_i, x)|^2$. Równość w tej nierówności zachodzi dla wszystkich wektorów wtedy, i tylko wtedy, gdy dany układ jest bazą.

157 Dla wektora x każdego z trzech typów w przestrzeni Minkowskiego (czasowego, świetlnego i przestrzennego) znaleźć podprzestrzeń $L(x)^{\perp}$ i kanoniczną postać metryki zacieśnionej do tej podprzestrzeni.

158 Wykazać, że twierdzenie 9, §15, pozostaje słuszne dla wszystkich wektorów kauzalnych.

159 Znaleźć wszystkie możliwe postacie kanoniczne metryki na dwu- i trójwymiarowych podprzestrzeniach przestrzeni Minkowskiego. Dla każdej z tych podprzestrzeni W znaleźć podprzestrzeń $W^{\perp}$ i postać kanoniczną metryki na tym ortogonalnym dopełnieniu.

160 Wykazać, że trzy wektory świetlne w przestrzeni Minkowskiego są liniowo niezależne wtedy, i tylko wtedy, gdy każde dwa z nich są liniowo niezależne.

161 Wykazać, że nierówność z lematu 10, §15, obowiązuje dla wszystkich trójek wektorów kauzalnych x, y, z takich, że $(x \cdot y)(y \cdot z)(z \cdot x) \neq 0$. Co oznacza to założenie?

162 Wykazać, że twierdzenie 11, §15, można rozszerzyć na wszystkie wektory kauzalne, jeśli określoną tam relację uogólnić w następujący sposób:

$$x \sim y \quad \Longleftrightarrow \quad (t \cdot x)(t \cdot y) > 0 \quad \text{dla pewnego wektora czasowego } t\,.$$

Dlaczego poprzednia definicja nie może być bezpośrednio zastosowana w tym ogólniejszym przypadku?

163 Wykazać, że w przestrzeni Minkowskiego istnieje baza (l_1, l_2, l_3, l_4) złożona z wektorów świetlnych, dla których $l_i \cdot l_j = 1$ dla $i \neq j$.

§16 Odwzorowania liniowe przestrzeni z iloczynem skalarnym

164 W bazie (e_1, e_2, e_3) dana jest forma kwadratowa metryki symetrycznej lub hermitowskiej g, oraz macierz operatora A. Wyliczyć macierz operatora A^* w tej samej bazie.

(a) $$g(x,x) = -2(x^2)^2 + 2x^1x^2 + 4x^1x^3\,, \qquad \mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 3 & -2 \\ 2 & 1 & -1 \end{pmatrix},$$

(b) $$\begin{aligned} g(x,x) = &- i\overline{x^1}x^2 + i\overline{x^2}x^1 + (1+i)\overline{x^1}x^3 \\ &+ (1-i)\overline{x^3}x^1 + \overline{x^2}x^3 + \overline{x^3}x^2\,, \end{aligned} \qquad \mathbf{A} = \begin{pmatrix} -i & 0 & 0 \\ 0 & 0 & i \\ 0 & -1 & 0 \end{pmatrix}.$$

165 Niech (V, g) będzie przestrzenią o dowolnym wymiarze z niezdegenerowanym iloczynem skalarnym, A – operatorem w tej przestrzeni. Operator sprzężony zadajemy warunkiem $g(A^*x, y) = g(x, Ay)$ dla wszystkich wektorów x, y, który jednak może nie mieć rozwiązania ze względu na A^*.

(a) Wykazać, że jeśli A^* istnieje, to jest określony jednoznacznie.

(b) Niech przestrzeń $\mathcal{S}$ funkcji Schwartza będzie wyposażona w iloczyn skalarny jak w przykładzie (i), p. 11, §15. Wybieramy funkcję $\chi \in \mathcal{S}$ i dla każdej funkcji $f \in \mathcal{S}$ określamy $Af = f(0)\chi$. Wykazać, że A jest operatorem liniowym, który nie posiada operatora sprzężonego.

166 Niech rzeczywista przestrzeń V ma przeliczalną bazę $(e_1, e_2, \ldots)$. Zadajemy iloczyn skalarny g oraz operator A kładąc $g(e_i, e_j) = \delta_{ij}$ i $Ae_i = e_{i+1}$. Wykazać, że iloczyn g jest niezdegenerowany, a operator A jest względem tego iloczynu izometrią wewnętrzną, ale nie jest bijektywny.

167 Wykazać, że jeśli A jest bijektywną izometrią wewnętrzną przestrzeni (o dowolnym wymiarze) z niezdegenerowanym iloczynem skalarnym, to operator A^* istnieje i jest równy A^{-1}.

168 Wykazać, że jeśli istnieje A^* (w przestrzeni z niezdegenerowanym iloczynem skalarnym), to zachodzi równość: $\operatorname{Ker} A^* = (\operatorname{Im} A)^\perp$.

169 Wykazać, że jeśli W jest podprzestrzenią inwariantną operatora A i istnieje A^* (w przestrzeni z niezdegenerowanym iloczynem skalarnym), to $W^\perp$ jest podprzestrzenią inwariantną operatora A^*.

170 Wykazać, że jeśli λ jest wartością własną operatora A w skończenie wymiarowej przestrzeni z niezdegenerowanym iloczynem skalarnym, to λ w przypadku biliniowym, a $\overline{\lambda}$ w przypadku półtoraliniowym, jest wartością własną operatora A^*.

171 Wykazać, że jeśli W jest skończenie wymiarową podprzestrzenią inwariantną izometrii wewnętrznej A, to $W^\perp$ jest również inwariantna względem A.

172 Niech A będzie bijektywną izometrią wewnętrzną przestrzeni V z biliniowym lub półtoraliniowym niezdegenerowanym iloczynem skalarnym g. Wykazać następujące stwierdzenia dotyczące zagadnienia własnego izometrii.

(a) Następujące warunki są równoważne: (i) $Ax = \lambda x$; (ii) $A^*x = \lambda^{-1}x$. W szczególności, $\operatorname{Ker}(A - \mathrm{id}) = \operatorname{Ker}(A^* - \mathrm{id}) = [\operatorname{Im}(A - \mathrm{id})]^\perp$.

(b) Jeśli wymiar przestrzeni jest skończony, to następujące warunki są równoważne: (i) λ jest wartością własną A; (ii) λ^{-1} w przypadku biliniowym, a $\overline{\lambda}^{-1}$ w przypadku półtoraliniowym, jest wartością własną A.

(c) Jeśli μ i ν są wartościami własnymi A takimi, że $\mu\nu \neq 1$ w przypadku biliniowym ($\bar{\mu}\nu \neq 1$ w przypadku półtoraliniowym), to odpowiednie podprzestrzenie własne są ortogonalne. W szczególności jeśli $\mu^2 \neq 1$ w przypadku biliniowym ($|\mu| \neq 1$ w przypadku półtoraliniowym), to metryka zacieśniona do podprzestrzeni własnej jest zerowa.

(d) Jeśli W jest sumą podprzestrzeni własnych A, to $W^\perp$ jest podprzestrzenią inwariantną A.

173 Wykazać, że jeśli właściwa, ortochroniczna transformacja Lorentza ma wektor własny do wartości własnej innej niż ± 1, to istnieje baza Minkowskiego (e_0, e_1, e_2, e_3) taka, że podprzestrzenie $L(e_0, e_3)$ oraz $L(e_1, e_2)$ są względem tej transformacji inwariantne. Transformacja jest złożeniem szczególnej transformacji Lorentza wzdłuż e_3 i izometrii właściwej w przestrzeni $L(e_1, e_2)$ (czyli obrotu – patrz §17).

174 Niech (V, g) i $(\tilde{V}, \tilde{g})$ będą skończenie wymiarowymi przestrzeniami z niezdegenerowanymi iloczynami skalarnymi o tych samych własnościach symetrii. Wykazać, że dla każdego odwzorowania $A \in \mathcal{L}(V, \tilde{V})$ warunek

$$\tilde{g}(y, Ax) = g(A^* y, x) \quad \text{dla wszystkich } x \in V\,,\ y \in \tilde{V}$$

określa jednoznacznie odwzorowanie $A^* \in \mathcal{L}(\tilde{V}, V)$. Definicja ta jest uogólnieniem przypadku $(\tilde{V}, \tilde{g}) = (V, g)$. Które z własności sprzężenia operatorów pozostają w mocy?

175 Niech $\mathcal{A}$ będzie zbiorem operatorów działających w przestrzeni unitarnej V, zamkniętym ze względu na sprzężenie (tj. jeśli $A \in \mathcal{A}$, to $A^* \in \mathcal{A}$).

(a) Wykazać, że jeśli $W \subseteq V$ jest podprzestrzenią inwariantną względem wszystkich operatorów z $\mathcal{A}$, to również $W^\perp$ jest taką podprzestrzenią.

(b) Wykazać, że następujące warunki są równoważne:

(i) w V nie ma nietrywialnej (różnej od V i $\{0\}$) podprzestrzeni inwariantnej względem wszystkich operatorów z $\mathcal{A}$;

(ii) jeśli B jest operatorem w V takim, że $[A, B] = 0$ dla każdego $A \in \mathcal{A}$, to B jest proporcjonalny do operatora identycznościowego.

176 Niech V będzie przestrzenią unitarną lub euklidesową, a A – operatorem w V.

(a) Wykazać że zbiór liczb $\big\{\|Ax\| \mid x \in V, \|x\| = 1\big\}$ jest ograniczony. Określamy wielkość $\|A\|$, jako supremum tego zbioru. Wykazać, że $\|Ax\| \leq \|A\|\|x\|$ dla każdego wektora x oraz, że $\|A\| = \sup\limits_{x \in V} \dfrac{\|Ax\|}{\|x\|}$.

(b) Wykazać, że ta definicja określa normę na przestrzeni $\mathcal{L}(V,V)$, tj. spełnione są następujące własności:

(i) $\|A\| \geq 0$, przy czym $\|A\| = 0 \Leftrightarrow A = 0$,

(ii) $\|\lambda A\| = |\lambda|\|A\|$,

(iii) $\|A+B\| \leq \|A\| + \|B\|$.

Ponadto, norma ta ma własność:

(iv) $\|AB\| \leq \|A\|\|B\|$.

(c) Wykazać, że $\|A\| = \sup\limits_{x,y\in V} \dfrac{|(x,Ay)|}{\|x\|\|y\|}$ (w przypadku euklidesowym znak wartości bezwzględnej można pominąć), oraz dalsze własności tej normy:

(v) $\|A^*\| = \|A\|$,

(vi) $\|A^*A\| = \|A\|^2$.

§17 Operatory normalne w przestrzeni unitarnej i euklidesowej

177 Dana jest macierz operatora symetrycznego w ortonormalnej bazie przestrzeni euklidesowej. Znaleźć macierz przejścia do ortonormalnej bazy własnej tego operatora. Podać jego macierz w nowej bazie.

$$\text{(a)} \begin{pmatrix} -1 & 3 & -15 \\ 3 & -9 & -5 \\ -15 & -5 & 15 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} -5 & -1 & 2 \\ -1 & -5 & -2 \\ 2 & -2 & -2 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} -1 & 2 & -6 \\ 2 & -4 & -3 \\ -6 & -3 & 4 \end{pmatrix},$$

$$\text{(d)} \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix}, \text{(e)} \begin{pmatrix} 0 & a & b & c \\ a & 0 & c & b \\ b & c & 0 & a \\ c & b & a & 0 \end{pmatrix}, \text{(f)} \begin{pmatrix} -4 & 3 & 2 & 4 \\ 3 & 4 & 6 & 12 \\ 2 & 6 & -1 & 8 \\ 4 & 12 & 8 & 11 \end{pmatrix}.$$

178 Dana jest macierz operatora w bazie ortonormalnej przestrzeni unitarnej. Sprawdzić, jakiego szczególnego typu jest ten operator i znaleźć macierz przejścia do ortonormalnej bazy własnej tego operatora. Podać jego macierz w nowej bazie.

$$\text{(a)} \begin{pmatrix} 2 & -2i & 1+i \\ 2i & 5 & -2+2i \\ 1-i & -2-2i & 3 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} 1 & -i & 2+i \\ i & 1 & -1+2i \\ 2-i & -1-2i & 5 \end{pmatrix},$$

$$\text{(c)}\ \frac{1}{5}\begin{pmatrix} 3i & 4i \\ 4 & -3 \end{pmatrix}, \quad \text{(d)} \begin{pmatrix} 1+i & -i \\ i & 1+i \end{pmatrix}, \quad \text{(e)} \begin{pmatrix} \cos\vartheta & \sin\vartheta\, e^{-i\varphi} \\ \sin\vartheta\, e^{i\varphi} & -\cos\vartheta \end{pmatrix},$$

$$\text{(f)} \begin{pmatrix} 0 & \sqrt{3}\overline{w} & 0 & 0 \\ \sqrt{3}w & 0 & 2\overline{w} & 0 \\ 0 & 2w & 0 & \sqrt{3}\overline{w} \\ 0 & 0 & \sqrt{3}w & 0 \end{pmatrix}, \ |w| = 1 , \quad \text{(g)} \ \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & i & 0 & 1 \\ 1 & 0 & -i & 0 \\ 0 & 1 & 0 & i \\ i & 0 & -1 & 0 \end{pmatrix}.$$

179 Wykazać, że jeśli $(e_1, \ldots, e_k)$ jest ortonormalną bazą podprzestrzeni W przestrzeni unitarnej lub euklidesowej V, to działanie operatora ortogonalnego rzutu na podprzestrzeń W można przedstawić w postaci $P_W x = \sum_{i=1}^{k} (e_i, x) e_i$.

180 W przestrzeni unitarnej lub euklidesowej dane są unormowane wektory x i y. Wektory są liniowo niezależne, ale poza tym dowolnie wzajemnie położone (nie zakładamy ortogonalności). Oznaczamy przez P_x i P_y operatory ortogonalnych rzutów na podprzestrzeni $L(x)$ i $L(y)$ odpowiednio. Wyliczyć widmo samosprzężonego operatora $\mu P_x + \nu P_y$ (μ, ν są liczbami rzeczywistymi).

181 Wykazać, że jeśli A jest samosprzężonym operatorem w przestrzeni unitarnej, to można tak dobrać dodatnią liczbę α, by operator $\alpha^2 \operatorname{id} - A^2$ był nieujemny (silna nierówność w definicji operatora dodatniego zastąpiona słabą). Wykazać, że wtedy istnieje dokładnie jeden nieujemny operator B taki, że operator U określony równaniem $\alpha U = A + iB$ jest unitarny. Czy istnieje minimalna liczba α, dla której ta konstrukcja jest możliwa?

182 Wykazać, że każdy operator samosprzężony w przestrzeni unitarnej może być przedstawiony jako kombinacja liniowa dwóch operatorów unitarnych, a dowolny operator – jako kombinacja czterech takich operatorów. Dokonać takiego rozkładu dla operatora, który w bazie ortonormalnej (e_1, e_2) ma macierz

$$\mathbf{A} = \begin{pmatrix} 1 - 5i & 1 + 4i \\ -1 & -2 - i \end{pmatrix}.$$

183 Dana jest forma kwadratowa symetrycznej formy biliniowej g w bazie ortonormalnej przestrzeni euklidesowej $(e_1, \ldots, e_n)$; iloczyn skalarny notujemy $x \cdot y$. Znaleźć ortonormalną bazę, w której ta forma ma postać diagonalną. Podać postać formy kwadratowej w nowej bazie.

(a) $g(x, x) = \sum_{i \neq j} x^i x^j$;

(b) $g(x, x) = (x^1)^2 + (x^2)^2 - (a \cdot x)^2, \quad \|a\| = 1, \ a^3 \neq \pm 1 \ (n = 3)$;

(c) $g(x, x) = a(x^1 + x^3)^2 + b(x^2 + x^4)^2 + c(x^1 x^2 + x^2 x^3 + x^3 x^4 + x^4 x^1)$, $c \neq 0 \ (n = 4)$.

184 Dane są formy kwadratowe w bazie $(e_1, \dots, e_n)$ pary form symetrycznych w przestrzeni rzeczywistej. Wykazać, że jedna z form jest dodatnio określona i wyliczyć macierz przejścia do bazy, w której forma dodatnio określona ma postać kanoniczną, a druga z form – postać diagonalną. Podać macierze form w nowej bazie.

$$\begin{aligned}
&\text{(a)} \quad g_1(x,x) = 2(x^1)^2 + 2(x^2)^2 + 2(x^3)^2 + 2x^1x^2 + 2x^1x^3 + 2x^2x^3\,,\\
&\phantom{\text{(a)}} \quad g_2(x,x) = -7(x^1)^2 + 2(x^2)^2 + (x^3)^2 + 2x^1x^2 - 4x^1x^3 + 2x^2x^3\,;\\
&\text{(b)} \quad g_1(x,x) = 2(x^1)^2 + 6(x^2)^2 + 5(x^3)^2 + 6x^1x^2 - 2x^1x^3 - 8x^2x^3\,,\\
&\phantom{\text{(b)}} \quad g_2(x,x) = (x^1)^2 - 20(x^2)^2 - 4(x^3)^2 - 12x^1x^2 + 32x^1x^3 + 48x^2x^3\,;\\
&\text{(c)} \quad g_1(x,x) = -2(x^1)^2 + 20(x^2)^2 + 7(x^3)^2 - 2x^1x^2 + 2x^1x^3 + 26x^2x^3\,,\\
&\phantom{\text{(c)}} \quad g_2(x,x) = 2(x^1)^2 + 5(x^2)^2 + 2(x^3)^2 + 2x^1x^2 - 2x^1x^3 + 4x^2x^3\,.
\end{aligned}$$

185 Dana jest macierz operatora ortogonalnego w ortonormalnej bazie przestrzeni euklidesowej. Znaleźć macierz przejścia do bazy ortonormalnej, o tej samej orientacji (jeśli to możliwe), w której operator przyjmuje postać kanoniczną. Podać jego macierz w nowej bazie.

$$\text{(a)}\ \frac{1}{3}\begin{pmatrix} -2 & -2 & 1 \\ 1 & -2 & -2 \\ -2 & 1 & -2 \end{pmatrix}\,, \quad \text{(b)}\ \frac{1}{5}\begin{pmatrix} 0 & 3 & 4 \\ 0 & -4 & 3 \\ 5 & 0 & 0 \end{pmatrix}\,, \quad \text{(c)}\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}\,.$$

186 W dwuwymiarowej przestrzeni euklidesowej odbicie R jest zadane działaniem $Re_1 = e_1$, $Re_2 = -e_2$, gdzie (e_1, e_2) jest ortonormalną, dodatnio zorientowaną bazą. Podobne warunki spełniają odbicie R' i baza (e'_1, e'_2). Znaleźć macierz przejścia z bazy nieprimowanej do bazy primowanej, dla której złożenie odbić $R'R$ jest obrotem o kąt φ.

187 Dana jest macierz operatora A w bazie ortonormalnej przestrzeni unitarnej lub euklidesowej. Znaleźć macierze operatorów jego rozkładu polarnego $A = U|A|$ (w tej samej bazie).

$$\text{(a)}\ \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\,, \quad \text{(b)}\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 2i & 1 \\ -i & 2 & i \\ 2\sqrt{2} & \sqrt{2}i & 0 \end{pmatrix}\,.$$

188 Rozkład polarny nieosobliwego operatora w przestrzeni unitarnej (lub euklidesowej) ma postać $A = U|A|$. Wykazać, że istnieje analogiczny, jednoznaczny rozkład w postaci $A = |A|'U'$. Wyrazić operatory $|A|'$ i U' za pomocą operatorów $|A|$ i U.

189 Wykazać, że jeśli A jest operatorem w przestrzeni unitarnej lub euklidesowej, to jego norma $\|A\|$ (zadanie 176) jest równa maksymalnej wartości własnej operatora $|A|$, oraz $\|Ax\| = \|A\|\,\|x\|$ wtedy, i tylko wtedy, gdy x jest wektorem własnym $|A|$ do tej wartości własnej.

190 Niech W i W' będą nietrywialnymi podprzestrzeniami przestrzeni euklidesowej. Jeśli $W \cap W' = \{0\}$, to możemy określić kąt między tymi podprzestrzeniami: $\measuredangle(W, W') = \inf\{\measuredangle(x, x') \mid x \in W\,, x' \in W'\} \in \langle 0, \pi/2\rangle$. Wykazać, że:

(a) $\cos\big[\measuredangle(W, W')\big] = \|P'P\| = \lambda$, gdzie P i P' są operatorami ortogonalnego rzutu na W i W' odpowiednio, a λ^2 jest maksymalną wartością własną operatora $PP'P$;

(b) jeśli $\lambda = 0$, to $W \perp W'$ i $\measuredangle(W, W') = \pi/2$;

(c) jeśli $\lambda > 0$, to $\measuredangle(W, W') = \measuredangle(z, z')$, gdzie $PP'Pz = \lambda^2 z$ (poza tym dowolny), a $z' = P'z$.

Jeśli warunek liniowej niezależności W i W' nie jest spełniony, to po prawej stronie definicji kąta zastępujemy te podprzestrzenie podprzestrzeniami $W \cap (W \cap W')^\perp$ i $W' \cap (W \cap W')^\perp$ odpowiednio. Jeśli którakolwiek z tych przestrzeni jest równa $\{0\}$, to powiemy, że kąt jest zerowy.

191 Wykazać, że metoda sprowadzania operatora ortogonalnego w trójwymiarowej przestrzeni euklidesowej do postaci kanonicznej, omówiona w przykładzie (ii), p.15, §17, może być rozszerzona (po niewielkich modyfikacjach) na wszystkie operatory normalne w takiej przestrzeni.

192 Dana jest macierz operatora antysymetrycznego w ortonormalnej bazie przestrzeni euklidesowej. Znaleźć macierz przejścia do ortonormalnej bazy o tej samej orientacji, w której operator przyjmuje postać kanoniczną. Podać jego macierz w nowej bazie.

$$\text{(a)}\ \begin{pmatrix} 0 & 4 & 0 \\ -4 & 0 & 3 \\ 0 & -3 & 0 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}, \quad \text{(c)}\ \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix}.$$

§18 Iloczyn tensorowy

193 Wykazać, że każdy półtoraliniowy iloczyn skalarny, dla którego relacja ortogonalności jest symetryczna, można przedstawić w postaci: (liczba o module jeden) $\times$ (iloczyn hermitowski).

194 Niech W i U będą podprzestrzeniami skończenie wymiarowej przestrzeni V. Wykazać następujące stwierdzenia:
(a) jeśli $W \subseteq U$, to $U^{\perp^*} \subseteq W^{\perp^*}$;
(b) jeśli $V = W \oplus U$, to $V^* = W^{\perp^*} \oplus U^{\perp^*}$.

195 Niech $\dim V_\alpha = n_\alpha < \infty$, $\alpha = 1, \ldots, r$, i niech $Z \in V_1 \otimes \ldots \otimes V_r$. Zastosować wynik przykładu (i) w punkcie 16, §18, dla sformułowania kryterium, by $Z = Z' \otimes x$, gdzie $Z' \in V_1 \otimes \ldots \otimes V_{r-1}$, $x \in V_r$. Przez indukcję rozszerzyć do kryterium, by Z był wektorem iloczynowym.

196 Sprawdzić, czy wektor przestrzeni iloczynowej podany w formie rozkładu w bazie iloczynowej jest wektorem iloczynowym. Jeśli tak, to przedstawić go w postaci iloczynu tensorowego wektorów.

$$
\begin{aligned}
&\text{(a)} && -6e_1 \otimes f_1 \otimes g_1 + 18e_1 \otimes f_1 \otimes g_2 - 10e_1 \otimes f_2 \otimes g_1 + 30e_1 \otimes f_2 \otimes g_2 + \\
& && + 3e_2 \otimes f_1 \otimes g_1 - 9e_2 \otimes f_1 \otimes g_2 + 5e_2 \otimes f_2 \otimes g_1 - 15e_2 \otimes f_2 \otimes g_2\,, \\
&\text{(b)} && -15e_1 \otimes f_1 \otimes g_1 + 60e_1 \otimes f_1 \otimes g_2 + 5e_1 \otimes f_2 \otimes g_1 - 20e_1 \otimes f_2 \otimes g_2 + \\
& && - 6e_2 \otimes f_1 \otimes g_1 + 24e_2 \otimes f_1 \otimes g_2 + 2e_2 \otimes f_2 \otimes g_1 - 8e_2 \otimes f_2 \otimes g_2\,.
\end{aligned}
$$

197 Wykazać, że jeśli przestrzenie V i W mają skończony wymiar, to przestrzeń dualną do przestrzeni $\mathcal{L}(V, W)$ można w naturalny sposób utożsamić z przestrzenią $\mathcal{L}(W, V)$. Podać przepis na funkcję dualności.

198 Niech (e_i), (f_j) i (h_k) będą bazami skończenie wymiarowych przestrzeni V, W i U odpowiednio, a (e^i) niech będzie bazą dualną do (e_i). Dane są rozkłady wektorów

$$
\begin{aligned}
Y &= 3f_1 \otimes (e^1 - 2e^2 + 5e^3) + (f_2 - f_3) \otimes e^2 \in W \otimes V^*\,, \\
Z &= (-5e_1 + 2e_2 + e_3) \otimes h_1 + (-4e_2 + e_3) \otimes h_2 \in V \otimes U\,.
\end{aligned}
$$

Pomnożyć tensorowo $Y \otimes Z$, a następnie wykonać w tym wektorze kontrakcję względem pary przestrzeni V, V^*.

199 Niech $\mathbf{A} = (A_{ij})$ będzie macierzą operatora $A \in \mathcal{L}(V, V)$ w bazie (e_i), a $\mathbf{B}$ – macierzą operatora $B \in \mathcal{L}(W, W)$ w bazie (f_j).

(a) Wykazać, że macierz operatora $A \otimes B$ w bazie $(e_i \otimes f_j)$ uporządkowanej leksykograficznie ma postać blokową

$$\begin{pmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \dots & A_{1n}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \dots & A_{2n}\mathbf{B} \\ \dots & \dots & \dots & \dots \\ A_{n1}\mathbf{B} & A_{n2}\mathbf{B} & \dots & A_{nn}\mathbf{B} \end{pmatrix}.$$

(b) Wykorzystać wynik zadania 79 dla wyliczenia wyznacznika operatora $A \otimes B$ (dowód alternatywny w stosunku do przykładu (v) w punkcie 16, §18).

(c) Wyliczyć ślad operatora $A \otimes B$ jako funkcję śladów operatorów A i B.

200 W unitarnych przestrzeniach V i W działają jednoparametrowe, ciągłe grupy unitarnych operatorów $S(t)$ i $T(t)$ odpowiednio, generowane samosprzężonymi operatorami A i B odpowiednio. Wykazać, że operatory $S(t) \otimes T(t)$ działające w unitarnej przestrzeni $V \otimes W$ tworzą jednoparametrową, ciągłą grupę unitarnych operatorów. Wyliczyć generator tej grupy.

201 Niech g i g' będą biliniowymi metrykami na skończenie wymiarowych przestrzeniach rzeczywistych V i V' odpowiednio (każda z metryk może być symplektyczna lub symetryczna – nie zakładamy, że mają te same własności symetrii).

(a) Wykazać, że metryki te wyznaczają jedyną metrykę G na $V \otimes V'$ taką, że $G(x \otimes x', y \otimes y') = g(x, y)g'(x', y')$ dla wszystkich wektorów iloczynowych. Jaką symetrię posiada G?

(b) Wyliczyć macierz metryki G w bazie $(e_i \otimes e'_j)$ w funkcji macierzy metryk g i g' w bazach (e_i) i (e'_j) odpowiednio. W szczególności, jeśli $\dim V = \dim V'$, to wyliczyć macierz metryki G zacieśnionej do podprzestrzeni rozpiętej wektorami $(e_i \otimes e'_i)$.

(c) Wyrazić niezmienniki metryki G przez niezmienniki metryk g i g'.

(d) Wykazać, że jeśli metryki symetryczne g i g' na rzeczywistej przestrzeni V $(V' = V)$ są dodatnio określone i mają macierze (g_{ij}) i (g'_{ij}) w bazie (e_i), to metryka g'', której macierz w tej samej bazie jest dana przez $g''_{ij} = g_{ij}g'_{ij}$, jest również dodatnio określona.

(e) Rozważyć wszystkie poprzednie punkty dla przypadku metryk hermitowskich w przestrzeniach zespolonych. Czy można łączyć przypadek hermitowskiej metryki g z biliniową metryką g'?

202 Niech (V_i, g_i), $i = 1, \dots, r$, będą przestrzeniami z niezdegenerowanymi iloczynami skalarnymi. Wszystkie iloczyny są biliniowe, lub wszystkie są hermitowskie. Posługując się wynikami poprzedniego zadania wykazać, że przestrzeń $V_1 \otimes \dots \otimes V_r$ jest wyposażona w naturalny niezdegenerowany iloczyn skalarny. Wykazać, że jeśli $A_i \in \mathcal{L}(V_i, V_i)$, to $(A_1 \otimes \dots \otimes A_r)^* = A_1^* \otimes \dots \otimes A_r^*$. (Wyniki tego zadania uogólniają odpowiednie konstrukcje dla przypadku przestrzeni euklidesowych lub unitarnych, omówione w punkcie 19, §18.)

203 Przestrzeń $V \otimes V^*$ może być kanonicznie utożsamiona z przestrzenią $\mathcal{L}(V, V)$. Jaka operacja tensorowa odpowiada w tym utożsamieniu operacji śladu operatora?

204 Wykazać, że definicja iloczynu tensorowego odwzorowań liniowych zadaje naturalne utożsamienie przestrzeni $\mathcal{L}(V_1, W_1) \otimes \dots \otimes \mathcal{L}(V_r, W_r)$ z przestrzenią $\mathcal{L}(V_1 \otimes \dots \otimes V_r, W_1 \otimes \dots \otimes W_r)$.

§19 Tensory

205 Dany jest rozkład tensora w bazie (e_i) oraz macierz przejścia do bazy (e'_i). Znaleźć rozkład tensora w nowej bazie.

(a) $$\begin{aligned} t &= e_2 \otimes (-e_1 + 3e_3) \otimes (e^1 - 3e^3) \\ &\quad + (e_1 - e_3) \otimes e_2 \otimes (e^2 + 2e^3)\,, \end{aligned} \qquad \boldsymbol{\beta} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix},$$

(b) $$\begin{aligned} t &= e_1 \otimes (e^2 + 3e^3) \otimes e^1 \\ &\quad + (e_1 - e_3) \otimes e^2 \otimes e^3\,, \end{aligned} \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 0 & 1 \\ -1 & 2 & 0 \end{pmatrix}.$$

206 Wyliczyć kontrakcje $C^2_1 t$ i $C^2_2 t$ dla podanych tensorów.

(a) $t = e_1 \otimes (5e_2 - 7e_3) \otimes (3e^2 + e^3) \otimes e^3 - e_2 \otimes e_3 \otimes (3e^2 - e^3) \otimes e^2\,,$

(b) $t = (e_1 + 4e_3) \otimes (5e_1 - 2e_2 - 4e_3) \otimes (e^2 - 2e^3) \otimes (e^1 + e^2)\,.$

207 Dany jest tensor jako forma wieloliniowa, wyrażona przez współrzędne jej argumentów przy ustalonej bazie (e_i) przestrzeni V. Zapisać rozkład tego tensora w bazie iloczynowej otrzymanej z tej bazy.

(a) $t(\varphi, \psi, x) = (\varphi_1 - 3\varphi_3)\psi_2(2x^1 - x^2) + \varphi_2(-3\psi_1 + \psi_3)x^2\,,$

(b) $t(\varphi, x, y) = \varphi_2(-x^2 + 5x^3)y^1 + (-\varphi_1 + \varphi_3)x^1 y^3\,.$

208 Przy ustalonej bazie (e_i) przestrzeni V podane są, na różne sposoby, tensory t i s oraz metryka symetryczna g. Dana jest również permutacja $\pi \in S_n$. Wykonać wskazane operacje tensorowe. W punktach (e) i (f) stosujemy konwencję o wspólnym uszeregowaniu wszystkich wskaźników.

(a) $t = e^1 \otimes (e^2 - 3e^3) \otimes e^2 + 3(2e^1 - e^3) \otimes (3e^1 - e^2) \otimes e^3\,,$
$g(x,x) = 2(x^1)^2 - (x^2)^2 - 10(x^3)^2 - 10x^1x^3 + 6x^2x^3\,,\quad$ wyliczyć $C^1_2 \hat{g}^1_2 t\,;$

(b) $t(\varphi,\psi) = \varphi_2(3\psi_1 - 2\psi_3) + (\varphi_1 - \varphi_3)\psi_2\,,$
$s(x,y) = 4x^2(y^1 - 5y^2) + (x^1 + 3x^3)y^3\,,$
rozłożyć $C^2_1(t \otimes s)$ w bazie iloczynowej;

(c) $t^1{}_{23} = 2\,,\ t^1{}_{31} = -3\,,\ t^2{}_{12} = -1\,,\ t^2{}_{33} = 5\,,\ t^3{}_{23} = 6\,,\ t^3{}_{33} = 4\,,$

$t^i{}_{jk} = 0$ w pozostałych przypadkach, $\quad \mathbf{g} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 3 \end{pmatrix}\,,$

wyliczyć współrzędne wektora $C^2_1 \hat{g}^1_1 t\,;$

(d) $t = (e_1 - e_2) \otimes (e_1 - e_3)\,,\quad s = (2e_2 + e_3) \otimes e_2\,,\quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}\,,$
wyliczyć $P^\pi(t \otimes s)\,;$

(e) $t^1{}_2{}^3 = 2\,,\ t^1{}_3{}^1 = -1\,,\ t^1{}_3{}^2 = 2\,,\ t^2{}_2{}^2 = 2\,,\ t^2{}_2{}^3 = 5\,,\ t^3{}_1{}^1 = 3\,,$
$t^3{}_2{}^3 = -4\,,\ t^3{}_3{}^3 = 1\,,\ t^i{}_j{}^k = 0\quad$ w pozostałych przypadkach,
$g(x,x) = (x^1)^2 + 2(x^2)^2 + (x^3)^2 + 2x^1x^2\,,\quad$ wyliczyć $t^{ij}{}_k$ oraz $t^{ij}{}_i\,;$

(f) $t_{ii} = 0\,,\ t_{12} = t_{21} = t_{34} = t_{43} = a\,,\ t_{13} = t_{31} = t_{24} = t_{42} = b\,,$
$t_{14} = t_{41} = t_{23} = t_{32} = c\,,\ g_{ii} = 0\,,\ g_{ij} = 1$ dla $i \neq j\,,$ wyliczyć $t^{ij}\,.$

209 Dany jest tensor $t \in \mathcal{T}^1_1(V)$ oraz niezdegenerowana metryka biliniowa $g \in \mathcal{T}^0_2(V)$. Tensor t określa operator liniowy w V. Za pomocą operacji tensorowych przedstawić tensor określający operator do niego sprzężony względem metryki g.

210 Wykazać, że jądro metryki g jest tym samym, co jądro odwzorowania liniowego $g \in \mathcal{L}(V, V^*)$ (czy też g^{T}) określanego przez tę metrykę (zgodnie z dyskusją przeprowadzoną w punkcie 8, §19).

211 Niech $A \in \mathcal{L}(V,V)$, $\dim V = n$. Operator $A^{\otimes k}$ jest k-krotnym iloczynem tensorowym operatora A. Wykazać, że $\det\left(A^{\otimes k}\right) = (\det A)^{kn^{k-1}}$.

212 Niech V będzie przestrzenią unitarną i oznaczmy przez V_n przestrzeń unitarną powstającą jako n-krotny iloczyn tensorowy tej przestrzeni, a dla $n = 0$ niech $V_0 = \mathbb{C}$. Dla każdego $z \in V$ definiujemy odwzorowania $a_n(z) \in \mathcal{L}(V_n, V_{n-1})$ $(n = 1, 2, \ldots)$ i $a_n^*(z) \in \mathcal{L}(V_n, V_{n+1})$ $(n = 0, 1, \ldots)$ za pomocą przepisów $(n \geq 1)$

$$\begin{aligned}
&a_n(z)\,(x_1 \otimes \ldots \otimes x_n) = \sqrt{n}\,(z, x_1)\,x_2 \otimes \ldots \otimes x_n\,,\\
&a_n^*(z)\,(x_1 \otimes \ldots \otimes x_n) = \sqrt{n+1}\;z \otimes x_1 \otimes \ldots \otimes x_n\,,\\
&a_0^*(z)\alpha = \alpha z\,.
\end{aligned}$$

Określamy ponadto operator działający w przestrzeni V zgodnie z przepisem $P(y, x)z = (y, z)x$. Wykazać następujące własności tych odwzorowań (w każdej relacji pełny zakres n, dla których obie strony są zdefiniowane):

(a) $a_n^*(\mu y + \nu z) = \mu a_n^*(y) + \nu a_n^*(z)\,, \quad a_n(\mu y + \nu z) = \overline{\mu} a_n(y) + \overline{\nu} a_n(z)\,,$

(b) $a_n^*(z) = [a_{n+1}(z)]^*\,,$

(c) $a_{n-1}^*(x)a_n(y) = n\,P(y, x) \otimes \mathrm{id}_{n-1}\,, \quad a_{n+1}(x)a_n^*(y) = (n+1)\,(x, y)\,\mathrm{id}_n\;.$

Sprzężenie po prawej stronie w punkcie (b) należy rozumieć w rozszerzonym sensie wprowadzonym w zadaniu 174.

213 Niech (V, g) będzie przestrzenią z niezdegenerowanym iloczynem skalarnym. Przestrzeń $\mathcal{T}_0^p(V)$ jest wyposażona w naturalny, niezdegenerowany iloczyn skalarny (zadanie 202). Wykazać, że wszystkie operatory permutacji P^π, $\pi \in S_p$, są wewnętrznymi izometriami przestrzeni względem tego iloczynu.

§20 Tensory symetryczne i antysymetryczne

214 Wyprowadzić wzór na rekurencyjną symetryzację form, analogiczny do wzoru na antysymetryzację z przykładu (iii), p. 5, §20.

215 Wykazać inną postać formuł z poprzedniego zadania i cytowanego w nim przykładu:

$$\mathcal{A}[x_1 \otimes \ldots \otimes x_n] = \frac{1}{n}\sum_{i=1}^{n}(-1)^{i-1} x_i \otimes \mathcal{A}[x_1 \otimes \ldots \not{x}_i \ldots \otimes x_n]\,,$$

$$\mathcal{S}[x_1 \otimes \ldots \otimes x_n] = \frac{1}{n}\sum_{i=1}^{n} x_i \otimes \mathcal{S}[x_1 \otimes \ldots \not{x}_i \ldots \otimes x_n]\,,$$

gdzie symbol $\not{x}_i$ oznacza, że wektor x_i został opuszczony w ciągu wektorów w iloczynie.

216 Wykazać, że przy założeniach zadania 213 operatory symetryzacji i antysymetryzacji działające w $\mathcal{T}_0^p(V)$ są samosprzężone.

217 Niech H będzie operatorem w przestrzeni $\mathcal{T}_0^k(V)$. Definiujemy działający w tej samej przestrzeni operator $\mathcal{S}(H) = \dfrac{1}{k!}\displaystyle\sum_{\pi\in S_k} P^\pi H P^{\pi^{-1}}$. Wykazać następujące własności tego operatora.

(a) Dla każdej permutacji $\sigma \in S_k$ jest $\big[\mathcal{S}(H), P^\sigma\big] = 0$. Stąd podprzestrzenie $\mathcal{S}\mathcal{T}_0^k(V)$ i $\mathcal{A}\mathcal{T}_0^k(V)$ są inwariantne względem $\mathcal{S}(H)$.

(b) Zachodzą związki: $\mathcal{S}(H)\mathcal{S} = \mathcal{S}H\mathcal{S}$, $\mathcal{S}(H)\mathcal{A} = \mathcal{A}H\mathcal{A}$.

(c) Jeśli przestrzeń $\mathcal{T}_0^k(V)$ jest wyposażona w iloczyn skalarny tak, jak w zadaniu 213, to $\mathcal{S}(H)^* = \mathcal{S}(H^*)$.

(d) Jeśli $H = H_1 \otimes H_2$, gdzie H_1 jest operatorem na $\mathcal{T}_0^r(V)$, a H_2 – na $\mathcal{T}_0^{k-r}(V)$, przy czym $\big[H_1, P^\sigma\big] = 0$ dla każdej permutacji $\sigma \in S_r$ i $\big[H_2, P^\rho\big] = 0$ dla każdej permutacji $\rho \in S_{k-r}$, to $\mathcal{S}(H) = \binom{k}{r}^{-1}\displaystyle\sum_{\pi} P^\pi H P^{\pi^{-1}}$, gdzie sumowanie przebiega po permutacjach spełniających warunki: $1 \le \pi(1) < \ldots < \pi(r) \le k$ i $1 \le \pi(r+1) < \ldots < \pi(k) \le k$.

(e) Jeśli $H = H_1 \otimes \mathrm{id}_{k-1}$, $H_1 \in \mathcal{L}(V,V)$, $\mathrm{id}_{k-1} \in \mathcal{L}(\mathcal{T}_0^{k-1}(V), \mathcal{T}_0^{k-1}(V))$, to

$$\mathcal{S}(H) = \frac{1}{k}\Big(H_1 \otimes \mathrm{id} \otimes \ldots \otimes \mathrm{id} + \mathrm{id} \otimes H_1 \otimes \mathrm{id} \otimes \ldots \otimes \mathrm{id} + \ldots + \mathrm{id} \otimes \ldots \otimes \mathrm{id} \otimes H_1\Big).$$

218 Przejmujemy oznaczenia zadania 212, i niech ponadto $\mathcal{S}_n$ i $\mathcal{A}_n$ oznaczają operatory symetryzacji i antysymetryzacji działające w przestrzeni V_n, przy czym przyjmujemy $\mathcal{S}_0 = \mathcal{A}_0 = \mathrm{id}$.

(a) Wykazać, że

$$\begin{aligned} &a_n(z)\mathcal{S}_n = \mathcal{S}_{n-1}a_n(z)\mathcal{S}_n\,, && a_n(z)\mathcal{A}_n = \mathcal{A}_{n-1}a_n(z)\mathcal{A}_n\,,\\ &\mathcal{S}_{n+1}a_n^*(z) = \mathcal{S}_{n+1}a_n^*(z)\mathcal{S}_n\,, && \mathcal{A}_{n+1}a_n^*(z) = \mathcal{A}_{n+1}a_n^*(z)\mathcal{A}_n\,. \end{aligned}$$

(b) Wprowadzamy oznaczenia $V_n^s = \mathcal{S}_n V_n$, $V_n^a = \mathcal{A}_n V_n$ oraz

$$\begin{aligned} b_n &: V_n^s \mapsto V_{n-1}^s\,, & b_n Z &= a_n Z\,,\\ b_n^* &: V_n^s \mapsto V_{n+1}^s\,, & b_n^* Z &= \mathcal{S}_{n+1}a_n^* Z\,,\\ d_n &: V_n^a \mapsto V_{n-1}^a\,, & d_n Z &= a_n Z\,,\\ d_n^* &: V_n^a \mapsto V_{n+1}^a\,, & d_n^* Z &= \mathcal{A}_{n+1}a_n^* Z\,. \end{aligned}$$

Wykazać, że spełnione są związki $b_n^*(z) = [b_{n+1}(z)]^*$, $d_n^*(z) = [d_{n+1}(z)]^*$ oraz

$$\begin{aligned} &b_{n-1}(x)b_n(y) - b_{n-1}(y)b_n(x) = 0\,, && n \ge 2\,,\\ &b_{n+1}(x)b_n^*(y) - b_{n-1}^*(y)b_n(x) = (x,y)\,\mathrm{id}_n\,, && n \ge 1\,,\\ &b_1(x)b_0^*(y) = (x,y)\,\mathrm{id}_0\,, \end{aligned}$$

$$\begin{aligned}
&d_{n-1}(x)d_n(y) + d_{n-1}(y)d_n(x) = 0\,, && n \geq 2\,,\\
&d_{n+1}(x)d_n^*(y) + d_{n-1}^*(y)d_n(x) = (x,y)\,\mathrm{id}_n\,, && n \geq 1\,,\\
&d_1(x)d_0^*(y) = (x,y)\,\mathrm{id}_0\,.
\end{aligned}$$

219 Niech układ wektorów $\{e_1, \ldots, e_n\}$ będzie bazą przestrzeni V.

(a) Wykazać, że wektory $\mathcal{S}(e_{i_1} \otimes \ldots \otimes e_{i_p})$, gdzie $1 \leq i_1 \leq \ldots \leq i_p \leq n$, tworzą bazę przestrzeni $\mathcal{S}[\mathcal{T}_0^p(V)]$.

(b) Wykazać, że wymiar przestrzeni $\mathcal{S}[\mathcal{T}_0^p(V)]$ jest równy liczbie $d(n,p)$ różnych podziałów liczby p na sumę: $k_1 + \ldots + k_n = p$.

(c) Wykazać, że $\dim \mathcal{S}[\mathcal{T}_0^p(V)] = \binom{n-1+p}{p}$.

220 Znaleźć bazę anihilatora podanego p-wektora. Jeśli ten p-wektor jest prosty, to przedstawić go w formie iloczynu zewnętrznego wektorów.

(a) $5e_1 \wedge e_2 + 29e_1 \wedge e_3 + 3e_1 \wedge e_4 - 3e_2 \wedge e_3 - e_2 \wedge e_4 - 4e_3 \wedge e_4\,,$

(b)
$$\begin{aligned}
&-5e_1 \wedge e_2 \wedge e_3 + 22e_1 \wedge e_2 \wedge e_4 + 2e_1 \wedge e_2 \wedge e_5 + e_1 \wedge e_3 \wedge e_4 +\\
&+ e_1 \wedge e_3 \wedge e_5 - 4e_1 \wedge e_4 \wedge e_5 + 7e_2 \wedge e_3 \wedge e_4 + 2e_2 \wedge e_3 \wedge e_5 -\\
&\qquad - 6e_2 \wedge e_4 \wedge e_5 + e_3 \wedge e_4 \wedge e_5\,,
\end{aligned}$$

(c) $2e_1 \wedge e_2 - 4e_3 \wedge e_4\,.$

221 Wykazać, że podany 3-wektor można przedstawić w postaci $s \wedge x$, gdzie 2-wektor s ma trywialny anihilator. Wyliczyć ten rozkład.

(a)
$$\begin{aligned}
&-4e_1 \wedge e_2 \wedge e_3 + 5e_1 \wedge e_2 \wedge e_4 - e_1 \wedge e_2 \wedge e_5 + 2e_1 \wedge e_3 \wedge e_4 +\\
&\qquad + 3e_2 \wedge e_3 \wedge e_4 - e_3 \wedge e_4 \wedge e_5\,,
\end{aligned}$$

(b) $-5e_1 \wedge e_2 \wedge e_3 + 10e_1 \wedge e_2 \wedge e_5 + e_1 \wedge e_3 \wedge e_5 + 3e_2 \wedge e_4 \wedge e_5\,.$

222 Wybieramy n-formę $\omega = n! e^1 \wedge \ldots \wedge e^n$ i związany z nią n-wektor $\hat{\omega}$ ($n = \dim V$). Znaleźć obiekt (p-wektor lub p-formę) dualny do danego, przedstawić go w jawnie prostej postaci.

(a) $\varphi = 3e^1 - 2e^2 + 5e^3 - e^4, \quad n = 4,$

(b) $x = 4e_1 - 7e_2 + e_4 - e_5, \quad n = 5,$

(c) $t = (3e_1 - e_2 + 4e_3 + 5e_4) \wedge (e_1 - 3e_3 + 6e_4), \quad n = 4,$

(d) $s = (7e^2 - 3e^3 + e^4) \wedge (2e^1 + 9e^2 - e^3), \quad n = 4.$

223 Niech $A \in \mathcal{L}(V,V)$ i niech $(A^i{}_j)$ będzie jego macierzą w bazie (e_i). Rozważamy operatory $\bigwedge^k(A)$, $k \leq n = \dim V$, określone w przykładzie (vi), punkt 8, §20. Wykazać następujące stwierdzenia.

(a) Macierz operatora $\bigwedge^k(A)$ w bazie $e_{i_1} \wedge \ldots \wedge e_{i_k}$, $i_1 < \ldots < i_k$, uporządkowanej leksykograficznie jest równa $k!A^{[j_1}{}_{i_1} \ldots A^{j_k]}{}_{i_k}$ (również $j_1 < \ldots < j_k$).

(b) $A^{[i_1}{}_{i_1} \ldots A^{i_k]}{}_{i_k} = \mathrm{Tr}\left[\bigwedge^k(A)\right]$ (po lewej stronie zwężenie po wszystkich wskaźnikach zgodnie z konwencją Einsteina, bez warunku porządku jak w poprzednim punkcie).

(c) Wyrażenie po lewej stronie równości w poprzednim punkcie jest sumą minorów głównych (o przekątnej wzdłuż diagonali) wymiaru k macierzy operatora A w dowolnej bazie.

(d) $\det(A - \lambda\,\mathrm{id}) = \sum_{k=0}^{n} (-\lambda)^{n-k}\,\mathrm{Tr}\left[\bigwedge^k(A)\right]$.

(e) $\det\left[\bigwedge^k(A)\right] = (\det A)^{\binom{n-1}{k-1}}$.

224 Dla dwuformy F w przestrzeni Minkowskiego, zapisanej tak jak w przykładzie (iv), p. 16, §20, wyliczyć niezmienniki ($F^{ij}F_{ij}$ oraz ${}^*F^{ij}F_{ij}$) za pomocą wielkości E^i i B^i. Kiedy F jest prosta?

225 W trójwymiarowej przestrzeni euklidesowej określamy operator $L_n x = n \times x$, gdzie n jest unormowanym wektorem. Wykazać, że operatory $R_n(\varphi) = \exp[\varphi L_n]$ tworzą jednoparametrową grupę obrotów wokół osi wskazywanej przez wektor n.

226 Niech $(e_1, \ldots, e_n)$ będzie dowolną bazą przestrzeni V, a ω – dowolną niezerową n-formą. Wykazać, że wektory bazy dualnej są proporcjonalne do form dualnych do $n-1$-wektorów utworzonych z wektorów danej bazy.

227 Wykazać, że jeśli a i b są wektorami w trójwymiarowej przestrzeni euklidesowej, to równanie (ze względu na wektor x) $a \times x = b$ ma rozwiązanie wtedy, i tylko wtedy, gdy $a \cdot b = 0$. Znaleźć w tym przypadku ogólne rozwiązanie.

228 Wykazać następującą tożsamość dla wektorów w trójwymiarowej przestrzeni euklidesowej: $(a_1 \times a_2) \cdot (b_1 \times b_2) = (a_1 \cdot b_1)(a_2 \cdot b_2) - (a_1 \cdot b_2)(a_2 \cdot b_1)$.

229 Niech (e_1, e_2, e_3) będzie dowolną bazą trójwymiarowej przestrzeni euklidesowej i oznaczmy przez $(\tilde{e}_1, \tilde{e}_2, \tilde{e}_3)$ bazę powstającą przez podniesienie wskaźników w bazie dualnej do danej. Baza ta spełnia $\tilde{e}_i \cdot e_j = \delta_{ij}$ i również bywa nazywana bazą dualną do danej.

(a) Znaleźć przedstawienie drugiej bazy za pomocą wektorów pierwszej bazy. Odwrotnie, wykazać, że macierz przejścia z bazy $(\tilde{e}_i)$ do (e_i) jest macierzą Grama bazy (e_i).

(b) Wykazać, że $\tilde{e}_1 \cdot (\tilde{e}_2 \times \tilde{e}_3) = [e_1 \cdot (e_2 \times e_3)]^{-1}$.

(c) Wykazać, że każdy wektor ma rozkład $x = \sum_{i=1}^{3} (x \cdot \tilde{e}_i)\, e_i$, a także analogiczny z bazami zamienionymi rolami.

230 Niech a_1, a_2 będą liniowo niezależnymi wektorami w trójwymiarowej przestrzeni euklidesowej. Znaleźć macierz przejścia z bazy $(a_1, a_2, a_1 \times a_2)$ do bazy do niej dualnej (w sensie poprzedniego zadania).

§21 Przestrzenie afiniczne

231 Wykazać, że jeśli $X_1, \ldots, X_k$ i $Y_1, \ldots, Y_k$ są dowolnymi ciągami punktów w przestrzeni afinicznej, a π jest dowolną permutacją z S_k, to zachodzi równość $\displaystyle\sum_{i=1}^{k} \overrightarrow{X_i Y_i} = \sum_{i=1}^{k} \overrightarrow{X_i Y_{\pi(i)}}$.

232 Dane są ciągi: punktów $X_1, \ldots, X_m$ i wektorów $p_1, \ldots, p_m$. Tworzymy 2-wektor $M(Z) = \displaystyle\sum_{i=1}^{m} \overrightarrow{ZX_i} \wedge p_i$ zależny od punktu Z i oznaczamy $p = \displaystyle\sum_{i=1}^{m} p_i$.

(a) Wykazać, że $M(Z + x) = M(Z) - x \wedge p$. Na jakich zbiorach punktów Z 2-wektor $M(Z)$ przyjmuje stałą wartość?

(b) Znaleźć warunek konieczny i wystarczający istnienia rozwiązań równania $M(Z) = 0$.

233 Wykazać, że zarówno sympleksy, jak i równoległościany, są zbiorami wypukłymi, tj. wraz z każdymi dwoma punktami należącymi do zbioru, należą do niego wszystkie punkty odcinka łączącego te punkty.

234 Dane są współrzędne w układzie odniesienia $(O, (e_i))$ punktów przestrzeni afinicznej. Wyznaczyć (i) równania parametryczne i (ii) układy równań określające powłoki afiniczne tych punktów.

$$\text{(a)} \begin{pmatrix} 1 \\ -1 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 2 \\ 9 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \\ -3 \\ -5 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ 4 \\ 6 \end{pmatrix}; \quad \text{(b)} \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix};$$

$$\text{(c)} \begin{pmatrix} 2 \\ 4 \\ -4 \\ -7 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ -2 \\ -2 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \\ -4 \\ -3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ -8 \\ -13 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \\ -2 \\ 2 \end{pmatrix};$$

235 Niech $Z = \sum_{i=0}^{m} \alpha_i X_i$ oraz $X_i = \sum_{k=0}^{n} \gamma_{ki} Y_k$ będą kombinacjami afinicznymi. Wykazać, że wówczas $Z = \sum_{k=0}^{n} \Big(\sum_{i=0}^{m} \gamma_{ki} \alpha_i \Big) Y_k$, gdzie współczynniki rozkładu spełniają warunek definicyjny kombinacji afinicznej.

236 Niech $(O_0, \ldots, O_n)$ będzie barycentrycznym układem odniesienia, i niech $X_i = \sum_{j=0}^{n} \lambda_{ji} O_j$, $i = 0, \ldots, m$.

(a) Wykazać, że wymiar powłoki afinicznej $A(X_0, \ldots, X_m)$ jest równy $\mathrm{rk}(\lambda_{ij}) - 1$. W szczególności, dane punkty są afinicznie niezależne wtedy, i tylko wtedy, gdy $\mathrm{rk}(\lambda_{ij}) = m + 1$.

(b) Wykazać, że punkty (X_i) tworzą układ barycentryczny wtedy, i tylko wtedy, gdy $m = n$ i macierz (λ_{ij}) jest nieosobliwa. Znaleźć w tym przypadku transformację odwrotną.

237 Niech R będzie równoległościanem w n-wymiarowej przestrzeni, opartym na liniowo niezależnych wektorach $x_1, \ldots, x_n$ wyprowadzonych z punktu P. Wykazać, że R jest przecięciem zbiorów punktów leżących pomiędzy parami równoległych hiperpłaszczyzn; liczba tych par jest równa n.

238 W n-wymiarowej przestrzeni wektorowej wybrano element objętości $\omega = n!e^1 \wedge \ldots \wedge e^n$. Wyliczyć objętość równoległościanu opartego na wektorach $x_1, \ldots, x_n$.

$$
\text{(a)} \quad \begin{aligned} x_1 &= 4e_1 - 3e_2 + 2e_3\,, \\ x_2 &= e_1 + 2e_2 + 7e_3\,, \\ x_3 &= -5e_1 + e_2 + 2e_3\,, \end{aligned} \qquad\qquad \text{(b)} \quad \begin{aligned} x_1 &= e_1 + 4e_2 + 5e_3 - e_4\,, \\ x_2 &= 2e_1 + e_4\,, \\ x_3 &= 3e_2 + 2e_3 + e_4\,, \\ x_4 &= e_1 + e_2\,. \end{aligned}
$$

239 W przestrzeni n-wymiarowej wybrano p-wektor prosty $\hat{\omega}_W$ oraz wektory $y_1, \ldots, y_p$. Wykazać, że te wektory stanowią bazę podprzestrzeni W zadanej wybranym p-wektorem i wyliczyć p-wymiarową objętość opartego na nich równoległościanu, wyznaczoną wybranym p-wektorem.

$$
\begin{aligned}
\text{(a)} \quad & \hat{\omega}_W = 2!(2e_1 - 5e_2 + 3e_4) \wedge (7e_2 + 2e_3 - 4e_4)\,, \\
& y_1 = 2e_1 + 2e_2 + 2e_3 - e_4\,,\ y_2 = 6e_1 - 8e_2 + 2e_3 + 5e_4\,, \\
\text{(b)} \quad & \hat{\omega}_W = 3!(e_1 + e_3 - 5e_4) \wedge (2e_2 + 3e_3) \wedge (4e_1 - e_3 + e_4)\,, \\
& y_1 = 9e_1 - e_3 - 3e_4\,,\ y_2 = 4e_1 + 2e_2 + 2e_3 + e_4\,, \\
& y_3 = 13e_1 + 4e_2 + 4e_3 - 2e_4\,.
\end{aligned}
$$

240 Dane są równania dwóch podprzestrzeni afinicznych, $\mathcal{N}$ i $\mathcal{R}$. Znaleźć przecięcie i powłokę afiniczną tych podprzestrzeni w postaci równań parametrycznych. Wymiar całej przestrzeni wynosi 5 w pierwszych dwóch punktach i 3 w trzecim.

(a) $\mathcal{N}: \quad \varphi^1(\overrightarrow{OX}) = 3\,, \quad \varphi^2(\overrightarrow{OX}) = 1\,,$

$\quad \varphi^1 = e^1 + 3e^2 - 2e^3 + 4e^4\,, \quad \varphi^2 = 2e^1 + e^2 + e^3 + 3e^4\,,$

$\mathcal{R}: \quad \psi^1(\overrightarrow{OX}) = 1\,, \quad \psi^2(\overrightarrow{OX}) = 3\,,$

$\quad \psi^1 = 4e^1 + e^2 + 2e^3 - e^4 + e^5\,, \quad \psi^2 = -e^1 + 3e^2 - 3e^3 + 8e^4 - e^5\,.$

(b) $\mathcal{N}: \quad \overrightarrow{OX} = a + t^1 f_1 + t^2 f_2\,, \quad t^i \in \mathbb{R}\,, \quad a = e_4 - 2e_5\,,$

$\quad f_1 = -2e_1 + e_2 + 6e_3 + 4e_4 - e_5\,,\ f_2 = e_1 + 3e_2 - 3e_3 + 12e_4 - 3e_5\,,$

$\mathcal{R}: \quad \varphi^1(\overrightarrow{OX}) = 9\,,\ \varphi^2(\overrightarrow{OX}) = 5\,,\ \varphi^1 = 3e^1 + 3e^2 + e^3 - 2e^5\,,\ \varphi^2 = e^4\,.$

(c) $\mathcal{N}: \quad \overrightarrow{OX} = a + tf\,, \quad t \in \mathbb{R}\,, \quad a = 5e_1 - e_2 + e_3\,,\ f = 4e_1 + e_2 - e_3\,,$

$\mathcal{R}: \quad \overrightarrow{OX} = b + s^1 k_1 + s^2 k_2\,, \quad s^i \in \mathbb{R}\,, \quad b = e_1 + 5e_2\,,$

$\quad k_1 = 8e_1 - 9e_2 + 7e_3\,, \quad k_2 = 2e_1 - 5e_2 + 4e_3\,.$

241 W przestrzeni afinicznej dany jest punkt Q oraz podprzestrzenie afiniczne $P+W$ i $P'+W'$, do których Q nie należy. Znaleźć wszystkie proste przechodzące przez punkt Q oraz przecinające (a) podprzestrzeń $P+W$; (b) obie podprzestrzenie. Czy każdy z tych problemów ma zawsze rozwiązanie?

242 Dane są dwie proste $P_i + L(a_i)$, $i = 1, 2$. Znaleźć wszystkie punkty O takie, że równania parametryczne tych prostych mogą być zapisane w postaci $\overrightarrow{OX} = b + ta_1$ oraz $\overrightarrow{OX} = -b + ta_2$ odpowiednio. Jaki zbiór tworzą te punkty?

243 W trójwymiarowej przestrzeni afinicznej dane są dwie nierównoległe, nieprzecinające się proste. Wybieramy punkt odniesienia O tak, aby ich równania miały postać: $\overrightarrow{OX_1} = b + t_1 a_1$, $\overrightarrow{OX_2} = -b + t_2 a_2$, $t_i \in \mathbb{R}$ (zob. poprzednie zadanie). Dany jest też punkt P nie należący do żadnej z tych prostych, o wektorze wodzącym $\overrightarrow{OP} = \lambda b + u_1 a_1 + u_2 a_2$. Znaleźć punkty Y_1 i Y_2 odpowiednio na pierwszej i drugiej prostej takie, że punkty P, Y_1, Y_2 leżą na jednej prostej. Wykazać, że rozwiązanie istnieje wtedy, i tylko wtedy, gdy $\lambda \neq \pm 1$.

244 Niech $(\mathcal{M}, g)$ będzie czasoprzestrzenią Minkowskiego, i niech będzie dana prosta $P + L(u')$, gdzie u' jest jednostkowym wektorem czasowym skierowanym w przyszłość (prosta opisuje ruch inercjalny). Niech u będzie innym takim wektorem. Dla każdego przestrzennego wektora z oznaczmy $|z| = \sqrt{|z^2|}$.

(a) Wektor przesunięcia wzdłuż prostej rozłożyć na składową czasową i przestrzenną z punktu widzenia obserwatora, dla którego u wskazuje kierunek czasowy, tj. znaleźć rozkład $t'u' = tu + y$, gdzie $u \cdot y = 0$. Wynikowi nadać postać $y = vt$, $t = \gamma t'$, wyliczyć wektor v i wykazać, że $|v| < 1$ oraz $\gamma = (1 - |v|^2)^{-1/2}$. (Zadanie opisuje tzw. lorentzowskie wydłużenie czasu.)

(b) Prosta $(P+w')+L(u')$ jest równoległa do danej, przy czym zakładamy, że $w' \cdot u' = 0$. Znaleźć taki wektor w, że $w \cdot u = 0$ i $(P+w)+L(u') = (P+w')+L(u')$. Wykazać, że $w = w' + \gamma w \cdot v\, u'$, a stąd $|w'|^2 = |w_\perp|^2 + \gamma^2 |w_\parallel|^2$, gdzie $w_\perp$ i $w_\parallel$ są częściami wektora w odpowiednio prostopadłą i równoległą do v. (Zadanie opisuje tzw. lorentzowskie skrócenie długości.)

§22 Przestrzeń afiniczna euklidesowa

245 Wyliczyć odległość podprzestrzeni afinicznych $P + L(x_1, \ldots, x_k)$ i $Q + L(y_1, \ldots, y_l)$, oraz punkty tych podprzestrzeni najbliżej siebie położone. Współrzędne punktów i wektorów podane są w ortonormalnym układzie afinicznym $(O, (e_i))$.

$$\text{(a) } P : \begin{pmatrix} 3 \\ -2 \\ 1 \\ 4 \end{pmatrix}, \ k = 0, \ Q : \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \ y_1 : \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \ y_2 : \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

$$\text{(b) } P : \begin{pmatrix} -13 \\ -10 \\ 2 \\ 5 \end{pmatrix}, \ x_1 : \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \ Q : \begin{pmatrix} 3 \\ 9 \\ -10 \\ -2 \end{pmatrix}, \ y_1 : \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \ y_2 : \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}.$$

$$\text{(c) } P : \begin{pmatrix} -9 \\ 1 \\ 0 \\ -5 \\ 2 \end{pmatrix}, \ x_1 : \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \ x_2 : \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \ Q : \begin{pmatrix} 10 \\ 1 \\ 0 \\ 3 \\ 2 \end{pmatrix}, \ y_1 : \begin{pmatrix} 1 \\ 2 \\ -2 \\ 0 \\ 0 \end{pmatrix}, \ y_2 : \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \\ 3 \end{pmatrix}.$$

246 W n-wymiarowej przestrzeni euklidesowej wybrano element objętości $e = n! e_1 \wedge \ldots \wedge e_n$, gdzie (e_i) jest dowolną dodatnio zorientowaną bazą ortonormalną. Wyliczyć p-wymiarową objętość p-wymiarowego równoległościanu opartego na wektorach $x_1, \ldots, x_p$. Kolumny współrzędnych tych wektorów w bazie (e_i) ustawione są w podaną macierz.

$$\text{(a)}\ \begin{pmatrix}1&1&2\\2&3&1\\4&5&2\end{pmatrix},\quad \text{(b)}\ \begin{pmatrix}2&-1\\0&4\\1&0\end{pmatrix},\quad \text{(c)}\ \begin{pmatrix}2&0&3\\1&1&0\\-2&1&1\\0&1&-2\end{pmatrix},\quad \text{(d)}\ \begin{pmatrix}2&3\\2&1\\1&1\\0&1\end{pmatrix}.$$

247 Przy warunkach i oznaczeniach zadania 242, ale postawionego w przestrzeni euklidesowej, znaleźć punkt O przy dodatkowym warunku $b \perp a_1, a_2$ (wymiar przestrzeni dowolny). Jak wyrażają się wtedy najbliżej położone punkty prostych i odległość między nimi?

248 Dane są równania dwóch prostych w trójwymiarowej euklidesowej przestrzeni afinicznej: $\overrightarrow{OX} \times a = k$ oraz $\overrightarrow{OY} \times b = l$. Znaleźć punkty na tych prostych najbliżej siebie położone i ich odległość. Podstawić w otrzymanych wzorach dane (w ortonormalnej, dodatnio zorientowanej bazie (e_1, e_2, e_3)): $a = 2e_2 + e_3$, $k = 3e_1 - 3e_2 + 6e_3$, $b = 3e_1 - e_3$, $l = -2e_1 + e_2 - 6e_3$.

249 Znaleźć warunek na to, aby proste zadane ogólnymi równaniami jak w poprzednim zadaniu, leżały na jednej płaszczyźnie. Znaleźć ich punkt przecięcia, jeśli istnieje.

250 Dane są dwie proste jak w poprzednim zadaniu oraz niezerowy wektor c. Znaleźć punkty P i Q odpowiednio na pierwszej i drugiej prostej takie, że $\overrightarrow{PQ} = c$. Podać warunek konieczny i wystarczający istnienia rozwiązania.

251 W trójwymiarowej afinicznej przestrzeni euklidesowej dane są cztery punkty P, Q, R, S nie leżące na jednej płaszczyźnie.

(a) Znaleźć prostą $\mathcal{L}$ przechodzącą przez punkty P i Q oraz płaszczyznę $\mathcal{N}$ przechodzącą przez punkty R i S, do której prosta $\mathcal{L}$ jest równoległa.

(b) Znaleźć odległość $\mathcal{L}$ od $\mathcal{N}$. Wyliczyć ją dla danych: $P = O + \hat{\imath} - 3\hat{\jmath} + \hat{k}$, $Q = O + 4\hat{\imath} - \hat{\jmath} + 2\hat{k}$, $R = O + 5\hat{\imath} + \hat{k}$, $S = O + 2\hat{\imath} - 2\hat{\jmath} + 2\hat{k}$, gdzie O jest pewnym punktem odniesienia, a $(\hat{\imath}, \hat{\jmath}, \hat{k})$ dodatnio zorientowaną bazą ortonormalną.

252 W trójwymiarowej afinicznej przestrzeni euklidesowej dwie nierównoległe płaszczyzny zadane są równaniami: $\overrightarrow{OX} \cdot b_1 = \mu_1$ oraz $\overrightarrow{OX} \cdot b_2 = \mu_2$. Znaleźć równanie (i) wektorowe i (ii) parametryczne ich przecięcia.

253 Dane są dwie płaszczyzny jak w poprzednim zadaniu oraz wektor a. Znaleźć parametryczną reprezentację wszystkich punktów Y o tej własności, że Y leży na pierwszej płaszczyźnie, a $Y + a$ – na drugiej. Jaki twór geometryczny uzyskujemy?

254 Kąt między podprzestrzeniami afinicznymi przestrzeni euklidesowej utożsamiamy z kątem między ich przestrzeniami kierunkowymi (zadanie 190). Wykazać, że w trójwymiarowej przestrzeni kąt między dwoma płaszczyznami jest równy mniejszemu z kątów utworzonych przez kierunki prostopadłe do płaszczyzn, a kąt między płaszczyzną a prostą, jest uzupełnieniem do $\pi/2$ mniejszego z kątów utworzonych przez kierunek prostopadły do płaszczyzny i przestrzeń kierunkową prostej.

255 Dana jest prosta $P+L(x)$ euklidesowej przestrzeni afinicznej o dowolnym wymiarze.

(a) Znaleźć wszystkie podprzestrzenie afiniczne (dowolnego wymiaru) tworzące z tą prostą zadany, niezerowy kąt.

(b) W przestrzeni trójwymiarowej znaleźć płaszczyznę przechodzącą przez punkty P oraz $Q \notin P+L(x)$, tworzącą z prostą maksymalny kąt.

256 W trójwymiarowej afinicznej przestrzeni euklidesowej dane są równania dwóch nierównoległych płaszczyzn $\overrightarrow{OX} \cdot a = \alpha$ oraz $\overrightarrow{OX} \cdot b = \beta$.

(a) Znaleźć równania wszystkich płaszczyzn zawierających przecięcie danych dwóch.

(b) Znaleźć równania wszystkich płaszczyzn, które tworzą równe kąty z tymi płaszczyznami.

(c) Znaleźć równania płaszczyzn, które spełniają warunki (a) i (b) jednocześnie.

257 Dane są równania trzech płaszczyzn w trójwymiarowej afinicznej przestrzeni euklidesowej: $\overrightarrow{OX} \cdot a_i = \mu_i$, przy czym wektory a_1, a_2, a_3 są liniowo niezależne. Znaleźć geometryczną interpretację wektorów bazy dualnej do (a_1, a_2, a_3) (patrz zad. 229). Uwzględnić również wyniki dwóch następnych zadań.

258 Znaleźć punkt przecięcia płaszczyzn określonych w poprzednim zadaniu oraz jego odległość od płaszczyzny o równaniu $\overrightarrow{OY} \cdot (a_1 + a_2 + a_3) = \nu$. Wyliczyć wektor wodzący szukanego punktu i odległość dla następujących danych: $a_1 = \hat{\imath} + \hat{\jmath} - 2\hat{k}$, $a_2 = -\hat{\imath} + 3\hat{\jmath} + 2\hat{k}$, $a_3 = \hat{\imath} - \hat{k}$, $\mu_1 = 1$, $\mu_2 = 3$, $\mu_3 = -1$ $\nu = -2$ ($(\hat{\imath}, \hat{\jmath}, \hat{k})$ – ortonormalna, dodatnio zorientowana baza).

259 Wyliczyć objętość tetraedru zamkniętego pomiędzy czterema płaszczyznami określonymi w poprzednim zadaniu (ogólnie i dla szczególnych danych).

260 W trójwymiarowej afinicznej przestrzeni euklidesowej wybieramy początek układu O, punkt P i prostą o równaniu $\overrightarrow{OX} \times a = k$. Znaleźć równanie prostej przechodzącej przez punkt P i przecinającej prostopadle zadaną prostą.

Wyliczyć odległość znalezionej prostej od punktu O. Wykonać rachunki dla danych: $a = 3\hat{\jmath} - 4\hat{k}$, $k = \hat{\imath} + 4\hat{\jmath} + 3\hat{k}$, $\overrightarrow{OP} = \hat{\imath} - \hat{\jmath} + 2\hat{k}$, gdzie $(\hat{\imath}, \hat{\jmath}, \hat{k})$ jest pewną ortonormalną, dodatnio zorientowaną bazą.

261 W trójwymiarowej afinicznej przestrzeni euklidesowej dane są trzy proste $O + L(n_i)$ $(i = 1, 2, 3)$ nie leżące w jednej płaszczyźnie. Znaleźć ogólne równania wszystkich płaszczyzn, które przecinają każdą z tych prostych w odległości d od punktu O. Ile jest tych płaszczyzn? Podać odległość punktu O od każdej z tych płaszczyzn.

262 Dana jest macierz dołączona sprzecznego układu równań. Znaleźć jego przybliżone rozwiązanie według metody najmniejszych kwadratów.

$$\text{(a)}\ \begin{pmatrix} 1 & 0 & -1 & 3 \\ 2 & 1 & 0 & 5 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & 6 \\ 2 & 1 & 2 & 0 \\ -1 & 1 & -1 & 6 \end{pmatrix}$$

§23 Odwzorowania afiniczne

263 Wykazać, że jeśli $(O_0, \ldots, O_n)$ jest układem barycentrycznym w przestrzeni $\mathcal{M}$, a $(Z_0, \ldots, Z_n)$ – dowolnym ciągiem punktów w przestrzeni $\mathcal{N}$, to odwzorowanie f zadane przez $f(O_k) = Z_k$ $(k = 0, \ldots, n)$ rozszerza się jednoznacznie do afinicznego odwzorowania $f : \mathcal{M} \mapsto \mathcal{N}$.

264 Wykazać, że jeśli podprzestrzenie afiniczne $\mathcal{N}_1, \mathcal{N}_2 \subseteq \mathcal{M}$ są równoległe, to dla każdego odwzorowania afinicznego $f : \mathcal{M} \mapsto \mathcal{M}'$ podprzestrzenie $f(\mathcal{N}_1), f(\mathcal{N}_2)$ są też równoległe.

265 Wykazać, że obrazem równoległościanu względem automorfizmu afinicznego jest równoległościan, a obrazem sympleksu – sympleks.

266 Wykazać, że każdą translację można przedstawić w postaci złożenia $(0, -\mathrm{id})_{O+a}(0, -\mathrm{id})_O$, względem dowolnego punktu O.

267 Niech przestrzeń wektorowa (skończenie wymiarowa) stowarzyszona z przestrzenią afiniczną $\mathcal{M}$ będzie wyposażona w dowolną niezdegenerowaną metrykę biliniową g, a f niech będzie afinicznym endomorfizmem tej przestrzeni o części liniowej A. Wykazać następujące uogólnienia twierdzenia 16, §23.

(a) Jeśli podprzestrzeń $\mathrm{Im}(A - \mathrm{id})$ jest niezdegenerowana, to istnieje dokładnie jeden wektor x taki, że $x \in [\mathrm{Im}(A - \mathrm{id})]^\perp$ i $f = (x, A)_O$ dla pewnego punktu odniesienia O. Wszystkie punkty, dla których to zachodzi, tworzą podprzestrzeń afiniczną $O+\mathrm{Ker}(A-\mathrm{id})$. Wykazać, że ta podprzestrzeń składa się z wszystkich punktów Y, dla których $f(Y)=t_x(Y)$.

(b) Jeśli f jest izometrią wewnętrzną i podprzestrzeń $\mathrm{Ker}(A-\mathrm{id})$ jest niezdegenerowana, to teza twierdzenia 16, §23, pozostaje w mocy w niezmienionej postaci. Translacja t_x zachowuje w tym przypadku podprzestrzeń $O+\mathrm{Ker}(A-\mathrm{id})$.

268 Niech f będzie endomorfizmem afinicznym euklidesowej przestrzeni afinicznej. Znaleźć zbiór punktów O przestrzeni, dla których wartość $\|\overrightarrow{Of(O)}\|$ osiąga minimum.

269 Dane są dwie nieprzecinające się, nierównoległe proste $\mathcal{L}_1$ i $\mathcal{L}_2$ w trójwymiarowej rzeczywistej przestrzeni afinicznej. Wykazać, że odwzorowania afiniczne f spełniające $f(\mathcal{L}_i)=\mathcal{L}_i$, $i=1,2$, tworzą podgrupę grupy automorfizmów, zależną od czterech parametrów. Sprawdzić, do jakiego zbioru redukuje się ta podgrupa, jeśli przestrzeń jest wyposażona w strukturę euklidesową i zażądać, aby f były izometriami.

270 Odbiciem w afinicznej przestrzeni euklidesowej nazwiemy każde odwzorowanie afiniczne $(0,R)_O$, gdzie R jest odbiciem jednego z ortogonalnych kierunków w stowarzyszonej euklidesowej przestrzeni wektorowej.

(a) Wykazać, że każda translacja w przestrzeni euklidesowej może być przedstawiona jako złożenie dwóch odbić.

(b) Wykazać, że każda izometria afiniczna przestrzeni euklidesowej może być otrzymana jako złożenie pewnej liczby odbić.

271 Wykazać, że każda izometria dwuwymiarowej afinicznej przestrzeni euklidesowej jest albo translacją, albo translacją z odbiciem kierunku prostopadłego do kierunku translacji, albo obrotem wokół pewnego punktu.

272 Niech operator $A\neq\mathrm{id}$ będzie obrotem właściwym w trójwymiarowej przestrzeni euklidesowej. Znaleźć jawne przedstawienie izometrii afinicznej $(y,A)_X$ w formie ruchu śrubowego (jak we wniosku 17, §23).

273 Niech $(\mathcal{M},t,g)$ będzie czasoprzestrzenią Galileusza, przy oznaczeniach przykładu (iii) w punkcie 4, §23. Dane są dwie proste $P+L(u)$ i $P'+L(u')$, gdzie $\tau(u)=\tau(u')=1$ (proste opisują ruch obserwatorów inercjalnych, wektory u i u' odmierzają jednostkowy upływ czasu).

(a) Wykazać, że każdy punkt czasoprzestrzeni (zdarzenie) można jednoznacznie przedstawić w postaci $X=P+tu+y$ oraz $X=P'+t'u'+y'$, gdzie $y,y'\in S$ (t i y są odpowiednio czasem i położeniem w przestrzeni zdarzenia X

z punktu widzenia pierwszego obserwatora, i podobnie dla drugiego). Podobnie, $\overrightarrow{PP'} = bu + a$, $a \in S$.

(b) Wykazać, że $v := u' - u \in S$ (względna prędkość). Wyrazić t i y przez t', y', v, b i a. Otrzymane równania podają związek opisów czasoprzestrzeni z punktu widzenia dwóch obserwatorów.

§24 Dodatkowe konstrukcje algebraiczne w przestrzeniach wektorowych

274 Przejmujemy warunki i oznaczenia zadania 174. Konstruujemy zewnętrzną sumę prostą $W = V \oplus \tilde{V}$ wyposażoną w iloczyn skalarny zadany przez $G(x \oplus \tilde{x}, y \oplus \tilde{y}) = g(x,y) + \tilde{g}(\tilde{x},\tilde{y})$, oraz operator $B \in \mathcal{L}(W,W)$ o działaniu $B(x \oplus \tilde{x}) = 0 \oplus Ax$. Wykazać, że operator sprzężony do B względem iloczynu skalarnego G ma postać $B^*(x \oplus \tilde{x}) = A^*\tilde{x} \oplus 0$, gdzie A^* jest sprzężeniem operatora A w sensie omówionym w zadaniu 174.

275 Przejmujemy oznaczenia zadań 212 i 218 i konstruujemy ortogonalną sumę prostą przestrzeni unitarnych $V_F = \bigoplus_{n=0}^{\infty} V_n$ z dodatnio określonym iloczynem skalarnym $\Big(\bigoplus_{n=0}^{\infty} x_n, \bigoplus_{n=0}^{\infty} y_n\Big) = \sum_{n=0}^{\infty} (x_n, y_n)_n$ (zgodnie z definicją sumy prostej tylko skończona liczba wektorów w sumach po lewej stronie jest różna od zera) oraz jej podprzestrzenie $V_F^s = \bigoplus_{n=0}^{\infty} V_n^s$ i $V_F^a = \bigoplus_{n=0}^{\infty} V_n^a$. Definiujemy operatory $a(x), a^*(x) \in \mathcal{L}(V_F, V_F)$ kładąc

$$a(x) \bigoplus_{n=0}^{\infty} Z_n = \bigoplus_{n=1}^{\infty} a_n(x) Z_n \quad \text{i} \quad a^*(x) \bigoplus_{n=0}^{\infty} Z_n = \bigoplus_{n=0}^{\infty} a_n^*(x) Z_n$$

(w drugim przypadku po prawej stronie składnik z przestrzeni V_0 jest zerowy). W analogiczny sposób definiujemy operatory $b(x), b^*(x) \in \mathcal{L}(V_F^s, V_F^s)$ oraz $d(x), d^*(x) \in \mathcal{L}V_F^a, V_F^a)$. Niech $[A,B]$ oznacza komutator operatorów; wyrażenie $[A,B]_+ := AB + BA$ nazywa się antykomutatorem.

Wykazać związki: $a^*(x) = [a(x)]^*$, $b^*(x) = [b(x)]^*$, $d^*(x) = [d(x)]^*$ oraz

$$[b(x), b(y)] = 0\,, \qquad [d(x), d(y)]_+ = 0\,,$$
$$[b(x), b^*(y)] = (x,y)\,\mathrm{id}\,, \qquad [d(x), d^*(y)]_+ = (x,y)\,\mathrm{id}\,,$$

276 Niech skończenie wymiarowa przestrzeń V będzie algebrą z działaniem $(x, y) \mapsto xy$. Korzystając z kanonicznych izomorfizmów różnych przestrzeni zbudowanych przy użyciu V wykazać, że mnożenie jest zadane jednoznacznie za pomocą tensora z przestrzeni $\mathcal{T}_2^1(V)$.

277 Wykazać, że trójwymiarowa przestrzeń euklidesowa jest algebrą z działaniem iloczynu wektorowego. Jaki tensor zadaje to działanie zgodnie z wynikiem poprzedniego zadania? Czy ta algebra posiada jedynkę?

278 Wykazać, że przestrzeń operatorów $\mathcal{L}(V, V)$ tworzy algebrę z działaniem składania odwzorowań.

279 Wykazać, że przestrzeń $\mathcal{F}(X)$ zespolonych funkcji na zbiorze X jest kompleksyfikacją analogicznej przestrzeni funkcji rzeczywistych. Przedyskutować dla tego przykładu struktury omówione w punkcie 7, §24.

280 Niech V będzie skończenie wymiarową przestrzenią unitarną z iloczynem $(.\,,.)$. Definiujemy odwzorowanie $\{.\,,.\} : V_{\mathbb{R}} \times V_{\mathbb{R}} \mapsto \mathbb{R}$, $\{x, y\} = \mathfrak{Im}(x, y)$.

(a) Wykazać, że $\{.\,,.\}$ jest niezdegenerowaną formą symplektyczną na $V_{\mathbb{R}}$.

(b) Operator $A \in \mathcal{L}(V_{\mathbb{R}}, V_{\mathbb{R}})$ rozkładamy na sumę operatora liniowego i antyliniowego na V: $A = A_- + A_+$ (jak w twierdzeniu 5, §24, z podstawieniem za J mnożenia przez i). Wykazać, że operatorem sprzężonym do A względem formy symplektycznej $\{.\,,.\}$ jest $A' = A_-^* - A_+^*$ (sprzężenia operatorów względem iloczynu skalarnego $(.\,,.)$, dla operatora A_+ zgodnie z przykładem (v), p. 8, §16).

(c) Wykazać, że następujące warunki są równoważne: (i) A jest odwzorowaniem symplektycznym (tj. izometrią formy $\{.\,,.\}$); (ii) $A_-^* A_- = A_+^* A_+ + \mathrm{id}$, $A_-^* A_+ = A_+^* A_-$; (iii) $A_- A_-^* = A_+ A_+^* + \mathrm{id}$, $A_- A_+^* = A_+ A_-^*$.

§25 Postać kanoniczna operatora liniowego

281 Wykazać, że operator (działający w przestrzeni nad dowolnym ciałem) jest diagonalizowalny wtedy, i tylko wtedy, gdy jego wielomian minimalny (określony w przykładzie (iv), p. 8, §25) ma rozkład $M(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_s)$, gdzie $\lambda_i \neq \lambda_j$ dla $i \neq j$.

282 Wykazać, że jeśli A jest operatorem na przestrzeni zespolonej, to istnieje jednoznaczny rozkład $A = D + N$, gdzie D jest operatorem diagonalizowalnym, a N – nilpotentnym (tj. $N^k = 0$ dla pewnego $k \in \mathbb{N}$), oraz $[D, N] = 0$.

283 Operator działający w przestrzeni zespolonej ma w zadanej bazie podaną niżej macierz. Znaleźć macierz przejścia do jego bazy jordanowskiej, podać macierz operatora w tej bazie.

$$\begin{pmatrix} 1+2i & 0 & 0 & -i & 2 & -2 \\ 0 & 1 & 0 & 0 & -1 & -i \\ 0 & 0 & 1 & 0 & i & 0 \\ i & 0 & 0 & 1 & 1 & 0 \\ 0 & -i & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

284 Podobnie jak w poprzednim zadaniu, dla operatora w przestrzeni rzeczywistej i uogólnionej bazy jordanowskiej.

$$\begin{pmatrix} 3 & -1 & -1 & -2 \\ 0 & 4 & 3 & -2 \\ 4 & -7 & -2 & 4 \\ 3 & -1 & 0 & 3 \end{pmatrix}$$

§26 Formy bi-afiniczne. Kwadryki

285 Wykazać, że rzeczywista kwadryka jest albo zbiorem spójnym albo, jeśli nie jest spójna, składa się z dwóch zbiorów spójnych. Jaki typ kanoniczny odpowiada drugiemu przypadkowi?

286 Wypisać wszystkie postacie kanoniczne kwadryk w przestrzeniach euklidesowych dwu- i trójwymiarowych. Które z nich mają dwie gałęzie?

287 Wykazać, że przecięcie kwadryki z podprzestrzenią afiniczną jest kwadryką w tej podprzestrzeni lub zbiorem pustym.

288 Niech h będzie symetryczną formą bi-afiniczną z niezdegenerowaną częścią biliniową g, i niech A będzie operatorem liniowym w stowarzyszonej przestrzeni wektorowej. Wykazać, że istnieje dokładnie jedna para endomorfizmów afinicznych f i f^* takich, że A jest częścią liniową f i zachodzi związek $h(X, f(Y)) = h(f^*(X), Y)$ dla wszystkich punktów X i Y.

289 Znaleźć zbiór punktów w trójwymiarowej euklidesowej przestrzeni afinicznej równoodległych od (a) dwóch różnych prostych, (b) prostej i płaszczyzny nie zawierającej tej prostej. Wykazać, że te zbiory są albo podprzestrzeniami, albo kwadrykami. Znaleźć kanoniczną postać w przypadku kwadryk nie będących podprzestrzeniami.

WSKAZÓWKI I ODPOWIEDZI

§1 Zbiory i zdania

2 (b) – dowód przez indukcję względem n.

3 Stosować prawa logiki do form zdaniowych φ_A, φ_B, φ_C.

5 Stosować prawa logiki do form zdaniowych $\varphi_A(x)$, $\varphi_B(x)$, $\varphi_C(y)$.

$$\text{(a) } (A \cup B) \times C = (A \times C) \cup (B \times C)\,,$$
$$\text{(b) } (A \cap B) \times C = (A \times C) \cap (B \times C)\,.$$

7 Dowód nie wprost: założyć, że $\{x, \{x, y\}\} = \{x', \{x', y'\}\}$ i $x \neq x'$.

8 (a) $(-1, 1)$, (b) $(-1, 1)$, (c) $\{0\}$, (d) $\{0\}$.

§2 Relacje. Odwzorowania

9 $X/\sim = \{f^{-1}(\{y\}) \mid y \in f(X)\}$.

10 Zwrócić uwagę, że w danym zbiorze w parach (x, y) jedna z liczb może być równa zeru. Klasami równoważności są proste na płaszczyźnie kartezjańskiej przechodzące przez $(0, 0)$, bez tego punktu.

12 Dla dowodu wynikania „w lewo” położyć (b) $A = \{x\}$, (c) $B = Y$.

13 Skorzystać z wyniku poprzedniego zadania.

14 Zapisać formę $f(x) \in B$ równoważnie jako $\exists y : y \in B \wedge f(x) = y$.

16 Rozważyć odwzorowanie $f : \mathbb{R} \mapsto \mathbb{R}$, $f(x) = a + b\frac{x}{\sqrt{x^2+1}}$.

17 Wykazać, że liczba elementów (i, j) zbioru $\mathbb{N} \times \mathbb{N}$ takich, że $i + j \leq n$, wynosi $\frac{(n-1)n}{2}$. Wykazać na tej podstawie, że odwzorowanie $f : \mathbb{N} \times \mathbb{N} \mapsto \mathbb{N}$, $f(i, j) = \frac{(i+j-2)(i+j-1)}{2} + i$ jest bijekcją.

§3 Działania, grupa, ciało

18 Działanie jest łączne. Elementem neutralnym jest element minimalny, jeśli taki istnieje. Tylko element neutralny ma odwrotny.

19 Działanie jest łączne. Jeśli istnieje element neutralny, to X jest jednoelementowy.

20 Wykorzystać wynik zadania 4.

22 (a) $r' = a + br$, $a \in \mathbb{Q}$, $b \in \mathbb{Q}^*$. (c) X_r^* jest podgrupą grupy R^* wtedy, i tylko wtedy, gdy $r^2 = a + br$, $a, b \in \mathbb{Q}$. Grupy X_r^*, gdzie r jest pierwiastkiem z liczby naturalnej będącej iloczynem liczb pierwszych parami różnych, wyczerpują grupy tego typu.

24 Oprócz elementu jednostkowego istnieje jednoparametrowa rodzina elementów, których kwadrat daje element jednostkowy. Rozważać wszystkie możliwe iloczyny takich elementów, aż do zamknięcia się zbioru. Podgrupa nie jest przemienna.

25 W podanym kroku w ciągu wielomianów $P_1', \ldots, P_s'$ istnieje co najmniej jeden różny od zerowego – wielomian P_s'. Jeśli pozostałe wielomiany P_i' są równe zeru, to P_s' jest maksymalnym wspólnym podzielnikiem. W przeciwnym przypadku przechodzimy do następnego kroku rekurencyjnego.

26 (a) Wielomian $P(x) = x+1$ jest największym wspólnym podzielnikiem. (b) Wielomiany nie mają wspólnych podzielników; $Q_1(x) = 0$, $Q_2(x) = x^2/8$, $Q_3(x) = (-x^3 + 1)/8$.

§4 Liczby zespolone

27 (a) $\frac{-29}{53} + \frac{31}{53}i$, (b) $\frac{3}{2} + \frac{3\sqrt{3}}{2}i$, (c) $\frac{\sqrt{6}-\sqrt{2}}{4} + \frac{\sqrt{6}+\sqrt{2}}{4}i$, (d) $\frac{\sqrt{3}}{2} - \frac{1}{2}i$, (e) $-32+32i$.

28 Przedstawić liczby $e^{i\varphi}$ i $e^{i\psi}$ za pomocą liczb $e^{i(\varphi+\psi)/2}$ oraz $e^{\pm i(\varphi-\psi)/2}$.

29 Skorzystać ze wzoru Newtona.

30 (a) $z = 5 - 3i, -2 + i$, (b) $z = \pm e^{iw}, \pm e^{-iw}$, (c) $z = 3 - 4i, -1 - i$, (d) $z = w + 2k\pi i, -w + (2k+1)\pi i$, $k \in \mathbb{Z}$.

33 Skorzystać ze wzorów na funkcje sumy argumentów.

34 ± 1, $\pm i$, $(1 \pm i)/\sqrt{2} = e^{\pm i\pi/4}$, $(-1 \pm i)/\sqrt{2} = e^{\pm i3\pi/4}$.

35 Zsumować wyrazy wielomianu.

37 (a) Rozłożyć wielomian $z^n - 1$ na czynniki. (b), (c) Użyć jawnej postaci pierwiastków. (d) Dla $\varepsilon_n = 1$ wykorzystać wynik zadania 35. Pozostałe tożsamości wynikają z tej szczególnej przez pomnożenie obu stron przez odpowiedni czynnik.

39 Skorzystać z wyniku zadania 28.

40 Liczby z_1, z_2 wyznaczają prostą. Rozważyć warunek, aby z_3 leżała na tej prostej. Wykazać dalej, że $d(z_1, z_2, z_3) \in \mathbb{R} \Leftrightarrow \overline{z_1}z_2 + \overline{z_2}z_3 + \overline{z_3}z_1 \in \mathbb{R}$. Jeśli to nie zachodzi, to układ równań $|z - z_1| = |z - z_2| = |z - z_3|$ ma jednoznaczne rozwiązanie ze względu na z.

41 Punkty z_1, z_2, z_4 leżą albo na jednej prostej, albo na jednym okręgu. W pierwszym przypadku zauważyć, że $\mu(z_1, z_2, z_3, z_4) = d(z_1, z_2, z_4)d(z_3, z_4, z_2)$ i wykorzystać wynik poprzedniego zadania. W drugim przypadku dla $i = 1, 2, 4$ mamy $z_i = z_0 + r\omega_i$, gdzie $r \in \mathbb{R}$, $|\omega_i| = 1$. Położyć $\omega_3 = (z_3 - z_0)/r$ i wykazać, że $\mu(z_1, z_2, z_3, z_4) \in \mathbb{R} \Leftrightarrow |\omega_3| = 1$.

§5 Grupy odwzorowań. Permutacje

42 Składać odwzorowania aż do zamknięcia się zbioru wszystkich otrzymanych odwzorowań; powstaje grupa $\{\mathrm{id}, f, g, f\circ g, g\circ f, f\circ g\circ f\}$. Grupa ma trzy podgrupy dwuelementowe i jedną trójelementową.

43 Wykazać, że translacja $t_\mu(z) = z+\mu$ jest generowana przez podane odwzorowania. Dla homografii $z\mapsto \frac{\alpha z+\beta}{\gamma z+\delta}$ rozpatrzyć oddzielnie przypadki $\gamma=0$ i $\gamma\neq 0$. Wykazać, że w drugim przypadku homografia może być zapisana jako $z\mapsto \frac{\lambda}{z+\mu}+\nu$, gdzie $\lambda\neq 0$.

44 Dla rozstrzygnięcia czym jest obraz – okręgiem czy prostą, zauważyć, że prosta sięga do nieskończoności.

45 (a) $\left(\begin{smallmatrix}1&2&3&4\\4&1&2&3\end{smallmatrix}\right)$, (b) $\left(\begin{smallmatrix}1&2&3&4&5&6\\3&5&6&1&4&2\end{smallmatrix}\right)$, (c) $(a, d_1,\ldots,d_n)(c, b_1\ldots,b_m)$, (d) $(a, b_1,\ldots,b_m, c, d_1,\ldots,d_n)$.

46 (a) $\left(\begin{smallmatrix}1&2&3&4&5\\4&2&3&5&1\end{smallmatrix}\right)$, (b) $\left(\begin{smallmatrix}1&2&3&4&5&6\\2&1&5&4&6&3\end{smallmatrix}\right)$.

47 (a) $+1$, (b) $(-1)^{(n-1)(k-1)}$.

49 Otrzymać w każdym z trzech przypadków wszystkie transpozycje.

50 Jeśli $c = (a_1,\ldots,a_n)$ i oznaczyć $a_i = a_j$ gdy $i = j \mod n$, to $c^k = \left(\begin{smallmatrix}a_1&\cdots&a_n\\a_{k+1}&\cdots&a_{k+n}\end{smallmatrix}\right)$. Cykl w rozkładzie tej permutacji zapoczątkowany od dowolnego elementu a_i z górnej linii zamyka się, gdy l jest minimalną naturalną liczbą, dla której $a_{i+lk} = a_i$.

51 Skorzystać z własności par n, k liczb wzajemnie pierwszych: istnieją liczby całkowite p, q, takie, że $pn + qk = 1$.

52 (a) $(a_1, a_9, a_5)(a_2, a_{10}, a_6)(a_3, a_{11}, a_7)(a_4, a_{12}, a_8)$,
(b) $(1, 5)(12, 10)(2, 13, 9)(11, 6, 3)(7, 4, 8)$.

53 Jeśli $q = (a_1\ldots a_m)\ldots(a_k\ldots a_n)$ (cykle rozłączne), to przedstawić $p = \left(\begin{smallmatrix}b_1\ldots b_n\\a_1\ldots a_n\end{smallmatrix}\right)$.

54 $\{\mathrm{id}, (1, 2)\}$, $\{\mathrm{id}, (1, 3)\}$, $\{\mathrm{id}, (2, 3)\}$, $\{\mathrm{id}, (1, 2, 3), (1, 3, 2)\}$; tylko ostatnia z nich jest niezmiennicza.

55 Skorzystać z wyniku zadania 53. Grupa ilorazowa składa się z warstw G, $(1, 2)G$, $(1, 3)G$, $(2, 3)G$, $(1, 2, 3)G$, $(1, 3, 2)G$. Stąd widoczny izomorfizm.

56 Wykazać, że $F_{(1,3)} = F_{(2,4)} = f$ oraz $F_{(1,4)} = F_{(2,3)} = g$. Permutacje tu występujące generują grupę S_4, a funkcje f, g generują H. Dalej posłużyć się homomorficznością związku $\pi\to F_\pi$ (dla jego wykazania wygodnie jest wprowadzić pomocnicze oznaczenie $w_i = z_{\sigma^{-1}(i)}$, skąd $z_{(\pi\circ\sigma)^{-1}(i)} = w_{\pi^{-1}(i)}$).

§6 Macierze

57
$$\text{(a)}\ \begin{pmatrix}-4&19&9\\8&17&17\\-8&-6&-10\\2&7&6\end{pmatrix},\ \text{(b)}\ \begin{pmatrix}2+4i\\5-3i\end{pmatrix},\ \text{(c)}\ \begin{pmatrix}0&-1-3i&1\\9-3i&-10i&-5-i\\-9+6i&13+7i&6+3i\end{pmatrix},$$
$$\text{(d)}\ 25+4i,\ \begin{pmatrix}-2+3i&1+i&5i\\11+3i&2-4i&15-5i\\13+13i&6-4i&25+5i\end{pmatrix},\ \begin{pmatrix}-2-3i&11-3i&13-13i\\1-i&2+4i&6+4i\\-5i&15+5i&25-5i\end{pmatrix}.$$

58 Napisać równanie macierzowe na szukany rozkład i posłużyć się operacją transpozycji lub sprzężenia.

$$\begin{pmatrix} 2 & -5 & 4 \\ 1 & 3 & 5 \\ 8 & -7 & 9 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 6 \\ -2 & 3 & -1 \\ 6 & -1 & 9 \end{pmatrix} + \begin{pmatrix} 0 & -3 & -2 \\ 3 & 0 & 6 \\ 2 & -6 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1+i & 2+3i & i \\ 3-i & 2+3i & 2 \\ 4i & 2-i & 1+5i \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 5+4i & -3i \\ 5-4i & 4 & 4+i \\ 3i & 4-i & 2 \end{pmatrix} + i\,\frac{1}{2}\begin{pmatrix} 2 & 2+i & 5 \\ 2-i & 6 & -1 \\ 5 & -1 & 10 \end{pmatrix}.$$

60 (a) Macierz $A \in T(n, \mathbb{K})$ przedstawić w postaci blokowej $A = \left(\begin{smallmatrix} B & C \\ 0 & a \end{smallmatrix}\right)$, gdzie $B \in T(n-1, \mathbb{K})$, i dowodzić przez indukcję szukając macierzy odwrotnej w postaci $A^{-1} = \left(\begin{smallmatrix} B' & C' \\ 0 & a' \end{smallmatrix}\right)$. (b) Wykazać, że diagonala iloczynu macierzy trójkątnych jest złożona z iloczynów elementów diagonalnych mnożonych macierzy.

61 (a) nie, (b) nie.

62 Posłużyć się operacją śladu.

63 Izomorfizm $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \mapsto a + ib$.

64 Gdy jedna z macierzy jest nieosobliwa.

65 $A^{-1} = -\beta^{-1}(A + \alpha\mathbf{1})$.

66 $A^{-1} = \mathbf{1} - (1 + Y^{\mathrm{T}}X)^{-1}XY^{\mathrm{T}}$.

67 Pomnożyć równanie przez B z lewej i A z prawej.

$$(\mathbf{1}_n + BA)^{-1} = \mathbf{1}_n - B(\mathbf{1}_m + AB)^{-1}A\,.$$

Sprawdzić iloczyn z macierzą $1 + BA$ w obu porządkach.

§7 Wyznaczniki. Macierz odwrotna

68 (a) $4i$, (b) 336, (c) -40.

69 (a) $z = a(1+i) + 7i$, $a \in \mathbb{R}$, (b) $z = -2 + i, 1 - 2i$.

70 Dowód przez indukcję. Rozbić ostatnią kolumnę na sumę dwóch kolumn.

71 Wykorzystać własności pierwiastków otrzymane w zadaniach do §4. (b) $\overline{\det E} = (-1)^{(n-2)(n-1)/2} \det E$. (c) Wyznacznik Vandermonde'a. Wykorzystując wynik zadania 28 wykazać, że $\det E = \prod_{n \geq k > j \geq 1} r_{kj} w_{kj}$, $r_{kj} = 2\sin\left[\frac{\pi(k-j)}{n}\right]$, $w_{kj} = i\exp\left[i\frac{\pi(k+j-2)}{n}\right]$. Wystarczy teraz rozważyć iloczyny czynników w_{kj}. $\det E = (-i)^{(n+2)(n-1)/2}$. (Jako produkt uboczny dostajemy tożsamość: iloczyn czynników r_{kj} jest równy 1.)

73 Sprowadzić zadanie do przypadku, gdy $I = \{1, \ldots, k\}$. Przedstawić wszystkie kolumny w postaci sumy dwóch kolumn: jedna z zerami na pierwszych k pozycjach, druga z zerami na pozostałych pozycjach. Skorzystać z liniowości wyznacznika względem każdej z kolumn i z wyniku poprzedniego zadania.

74 Dowód (i) wprost z podstawowych własności wyznaczników, lub (ii) z twierdzenia Cauchy'ego przez pomnożenie macierzy początkowych wyznaczników przez macierze $\begin{pmatrix} \mathbf{1}_m & M \\ 0 & \mathbf{1}_k \end{pmatrix}$ lub $\begin{pmatrix} \mathbf{1}_m & 0 \\ N & \mathbf{1}_k \end{pmatrix}$ z odpowiednich stron.

75 Zastosować wynik poprzedniego zadania do macierzy $\begin{pmatrix} \mathbf{1}_m & -M \\ N & \mathbf{1}_k \end{pmatrix}$.

76

$$\text{(a)}\ \tfrac{1}{3}\begin{pmatrix} -2 & -4 & 3 \\ -2 & 11 & -6 \\ 3 & -6 & 3 \end{pmatrix},\ \text{(b)}\ \tfrac{1}{4}\begin{pmatrix} 2-2i & -4+2i & 1+i \\ 0 & 2-2i & -4+2i \\ 0 & 0 & 2-2i \end{pmatrix},\ \text{(c)}\ \tfrac{1}{24}\begin{pmatrix} -5 & -7 & -1 & -1 \\ -1 & -5 & 7 & 1 \\ -1 & 1 & -5 & 7 \\ -7 & 1 & 1 & -5 \end{pmatrix}.$$

77

$$\text{(a)}\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \\ -10 \end{pmatrix},\ \text{(b)}\ \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix},\ \text{(c)}\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \tfrac{1}{5}\begin{pmatrix} 3+19i \\ -30+15i \\ 36+23i \end{pmatrix}.$$

78 $(B^{-1})_{ij} = \nu_i^{-1}(A^{-1})_{ij}\mu_j^{-1}$.

79 Dowodzić przez indukcję względem m. Założyć, że $A_{mm} \neq 0$ (każdy nieosobliwy przypadek można sprowadzić do tej postaci) i odejmując wielokrotność ostatniej blokowej kolumny od pozostałych kolumn wyzerować ostatni blokowy wiersz, z wyjątkiem bloku na diagonali.

80 Przy oznaczeniach takich, jak w dowodzie twierdzenia, należy przerzucić ograniczenia w zakresie wskaźników i_s, j_s na wskaźniki k_s, l_s.

81 Dowód indukcyjny względem wymiaru macierzy, z wykorzystaniem rozwinięcia Laplace'a.

§8 Podstawowe pojęcia

82 (a) nie; (b) nie; (c) tak, baza złożona z wektorów bazy kanonicznej E_{ij}, $i \neq j$, oraz $E_{ii} - E_{(i+1)(i+1)}$, $i = 1, \ldots, n-1$; (d) tak, baza E_{ij}, $i \leq j$; (e) tak, dla $M = E_{ij}$ baza złożona z wektorów E_{kl}, $k \neq j$, $l \neq i$, oraz wektora $E_{ii} + E_{jj}$.

83 (a) ogólnie: nie; dla $\mathbb{K} = \mathbb{R}$: tak, trywialną; (b) tak wtedy, i tylko wtedy, gdy $B = 0$; (c) tak.

84 Rozwijając potęgę wyrazić f_i przez e_j. Podobnie w przeciwnym kierunku. Zauważyć, że z wzajemnej odwrotności otrzymanych macierzy otrzymujemy dla $l < k$ tożsamość $\sum_{i=l}^{k}(-1)^i\binom{k}{i}\binom{i}{l} = 0$.

85 (a), (b), (c) tak. Podany zbiór tworzy bazę przestrzeni z punktu (c).

86 Dla dowodu liniowej niezależności wykorzystać wyznacznik Vandermonde'a.

87 (b) baza złożona z funkcji $e_x(y) = 0$ dla $y \neq x$ i $e_x(x) = 1$ $(x \in X)$. Podprzestrzenie pokrywają się wtedy, i tylko wtedy, gdy X jest zbiorem skończonym (i są wtedy równe F).

88 (b) $(e_1, \ldots, e_n, ie_1, \ldots, ie_n)$ jest bazą $V_{\mathbb{R}}$.

89 Wektory tworzą bazę w każdym punkcie.

$$\text{(a)}\ x = (1 + 2\alpha + 3\alpha^2 + 4\alpha^3)f_1 - (2 + 3\alpha + 4\alpha^2)f_2 + (3 + 4\alpha)f_3 - 4f_4\,,$$
$$\text{(b)}\ x = \tfrac{1}{7}(f_1 + 5f_2)\,,\ \text{(c)}\ x = -\tfrac{5}{13}(3 + 2i)f_1 + \tfrac{4}{13}(-2 + 3i)f_2\,.$$

§9 Układy równań liniowych

90 t, s – parametry.

$$\text{(a) } \tfrac{1}{9}\begin{pmatrix}23\\-17\\0\\0\end{pmatrix}+t\begin{pmatrix}-13\\1\\9\\0\end{pmatrix}+s\begin{pmatrix}-2\\1\\0\\1\end{pmatrix},\quad \text{(b) } \begin{pmatrix}16\\10-5i\\0\end{pmatrix}+t\begin{pmatrix}-4+5i\\-2+5i\\1\end{pmatrix},$$

$$\text{(c) sprzeczny,}\quad \text{(d) } \tfrac{1}{13}\begin{pmatrix}-1\\22\end{pmatrix},\quad \text{(e) } \tfrac{1}{5}\begin{pmatrix}11\\-4\\0\\0\end{pmatrix}+t\begin{pmatrix}-11\\9\\5\\0\end{pmatrix}+s\begin{pmatrix}1\\-2\\0\\1\end{pmatrix},$$

$$\text{(f) } \tfrac{1}{11}\begin{pmatrix}7\\17\\11\end{pmatrix}+t\begin{pmatrix}-9\\19\\11\end{pmatrix},\ t=0 \text{ gdy } \lambda\neq 5,\quad \text{(g) } \tfrac{1}{5}\begin{pmatrix}10\\-11\\3\\0\\0\end{pmatrix}+t\begin{pmatrix}-5\\11\\-8\\5\\0\end{pmatrix}+s\begin{pmatrix}-5\\8\\-4\\0\\5\end{pmatrix},$$

$$\text{(h) } c\neq 1,7:\ \tfrac{1}{7-c}\begin{pmatrix}c-4\\c^2-5c-2\\3\end{pmatrix},\ c=1:\ \begin{pmatrix}0\\-1\\0\end{pmatrix}+t\begin{pmatrix}-1\\0\\1\end{pmatrix},\ c=7:\ \text{sprzeczny,}$$

$$\text{(i) } \alpha\neq 0,1,(-1\pm i\sqrt{3})/2:\ \tfrac{1}{\alpha^2+\alpha+1}\begin{pmatrix}0\\\alpha\\\alpha+1\end{pmatrix},\ \alpha=0:\ \begin{pmatrix}1\\0\\0\end{pmatrix}+t\begin{pmatrix}-1\\0\\1\end{pmatrix},$$

$$\alpha=1:\ \begin{pmatrix}1\\0\\0\end{pmatrix}+t\begin{pmatrix}-1\\0\\1\end{pmatrix}+s\begin{pmatrix}-1\\1\\0\end{pmatrix},\ \alpha=(-1\pm i\sqrt{3})/2:\ \text{sprzeczny,}$$

$$\text{(j) } \alpha=0:\ \begin{pmatrix}0\\0\end{pmatrix},\ \alpha=1:\ \begin{pmatrix}1\\0\end{pmatrix}+t\begin{pmatrix}-1\\1\end{pmatrix},\ \alpha\neq 0,1:\ \text{sprzeczny,}$$

$$\text{(k) } 3\lambda-8\neq 0:\ \tfrac{1}{3(3\lambda-8)}\begin{pmatrix}0\\-13\lambda\\-6\lambda+42\\39\end{pmatrix}+t\begin{pmatrix}-3\\2\\0\\0\end{pmatrix},\ 3\lambda-8=0:\ \text{sprzeczny,}$$

$$\text{(l) } \tfrac{1}{10}\begin{pmatrix}16\\-12\\-1\\0\\0\end{pmatrix}+t\begin{pmatrix}-1\\-1\\1\\1\\0\end{pmatrix}+s\begin{pmatrix}2\\-14\\3\\0\\10\end{pmatrix}.$$

91 (a) X_1, X_2, X_4 lin. niezal., $X_3 = 2X_1 - X_2$, $X_5 = X_1 + X_2 + X_4$; (b) X_1, X_2, X_4 lin. niezal., $X_3 = 2X_1 - X_2$, $X_5 = 2X_1 - X_2 - X_4$; (c) przypadek $a = b$ trywialny; $a = -3b \neq 0$: X_1, X_2, X_3 lin. niezal., $X_4 = -X_1 - X_2 - X_3$; $a \neq b, -3b$: wszystkie lin. niezal., (d) X_1, X_2 lin. niezal., $X_3 = 3X_1 + 2X_2$, $X_4 = X_1 + 4X_2$; (e) X_1, X_2, X_3 lin. niezal., $X_4 = 3X_1 + X_2 - X_3$, $X_5 = 2X_1 - X_3$; (f) X_1, X_2 lin. niezal., $X_3 = 2X_1 - 3X_2$.

92
$$\text{(a) } \begin{pmatrix}2&-1&-1&-1\\-1&1&0&1\\-1&0&1&1\\1&0&0&-1\end{pmatrix},\quad \text{(b) } \tfrac{1}{2}\begin{pmatrix}-2i&0&0\\2&1+i&1-i\\2i&-1-i&1+i\end{pmatrix},\quad \text{(c) } \begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}.$$

94 Jeśli oba układy są sprzeczne, to są zawsze równoważne. Jeśli oba układy są niesprzeczne, to warunkiem koniecznym i wystarczającym ich równoważności jest $\operatorname{rk}\begin{pmatrix}A & B\end{pmatrix} = \operatorname{rk}\begin{pmatrix}A & B\\ A' & B'\end{pmatrix} = \operatorname{rk}\begin{pmatrix}A' & B'\end{pmatrix}$.

95 $P(x) = \frac{a}{2}(x + x^2) + b(1 - x^2) + \frac{c}{2}(-x + x^2) + t(x - x^3)$, $t \in \mathbb{R}$.

§10 Odwzorowania liniowe

96
$$\text{(a) } \mathbf{A}=\begin{pmatrix}1&-2&3\\3&1&2\\0&5&-1\\-5&0&-1\end{pmatrix},\ \mathbf{y}=\begin{pmatrix}-12\\-1\\9\\-6\end{pmatrix},\qquad \text{(b) } \mathbf{A}=\begin{pmatrix}3-i&-i&i\\-i&4+5i&-1\end{pmatrix},\ \mathbf{y}=\begin{pmatrix}9-6i\\-6+i\end{pmatrix}.$$

97 Odwzorowanie przyjmuje postać kanoniczną w bazach $(e'_i) = (e_i)\boldsymbol{\beta}$, $(f'_i) = (f_i)\boldsymbol{\gamma}$:

(a) $\boldsymbol{\beta} = \begin{pmatrix} 1&0&0&-2&-1\\ 0&1&0&1&-1\\ 0&0&0&1&0\\ 0&0&1&0&1\\ 0&0&0&0&1 \end{pmatrix}$, $\boldsymbol{\gamma} = \begin{pmatrix} 3&1&7&0\\ 1&2&2&0\\ 2&5&1&0\\ 4&6&5&1 \end{pmatrix}$, $\operatorname{Ker} A = L(e'_4, e'_5)$, $\operatorname{Im} A = L(f'_1, f'_2, f'_3)$,

(b) $\boldsymbol{\beta} = \begin{pmatrix} 1&0&0&-1\\ 0&1&0&-i\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix}$, $\boldsymbol{\gamma} = \begin{pmatrix} 1&i&i&0\\ i&2&1&0\\ 1+i&1-i&i&0\\ -i&3&1&1 \end{pmatrix}$, $\operatorname{Ker} A = L(e'_4)$, $\operatorname{Im} A = L(f'_1, f'_2, f'_3)$.

98 (a) $\begin{pmatrix} 4&-12&-3\\ -2&1&1\\ -5&-17&4 \end{pmatrix}$, (b) $\frac{1}{2}\begin{pmatrix} 5-7i&15-3i\\ -1+5i&-7+5i \end{pmatrix}$.

99 Dla $\lambda = -1$: $x = 3e_1 - 4e_2 + t(e_1 + e_2 - e_3) + s(2e_2 - e_4)$, $t, s \in \mathbb{R}$; dla $\lambda \neq -1$ równanie sprzeczne.

100 $z = 2+i, 1-2i$: $\operatorname{Ker} A = L(4e_1 - 3e_2 + e_3)$, $\operatorname{Im} A = L(e_1 + ie_2, ie_1 + e_3)$; w pozostałych przypadkach $\operatorname{Ker} A = \{0\}$, $\operatorname{Im} A = V$.

101 Przy oznaczeniu $e'_k = e_{\pi(k)}$ jest ${A'^i}_j = {A^{\pi(i)}}_{\pi(j)}$.

102 ${A^i}_j = {\delta^i}_{\pi(j)}$, $\det A = \operatorname{sgn} \pi$.

104 Dla znalezienia bazy użyć izomorfizmu z przestrzenią macierzową.

105 Macierz operatora ma postać blokową $\frac{1}{\det A}\begin{pmatrix} dA^{\mathrm{T}} & -bA^{\mathrm{T}}\\ -cA^{\mathrm{T}} & aA^{\mathrm{T}} \end{pmatrix}$, a jego wyznacznik jest równy 1.

106 Jeśli $\det A = 0$, to kolumny macierzy A są liniowo zależne, więc macierz ta ma postać $A = BC^{\mathrm{T}}$, gdzie B i C są kolumnami dwuwyrazowymi (ten ostatni wniosek jest zawsze słuszny tylko dla macierzy 2×2). Obraz jest jednowymiarową podprzestrzenią napiętą na wektorze CC^{T}. Jeśli $B = \begin{pmatrix} b_1\\ b_2 \end{pmatrix}$, to bazę jądra tworzą macierze $\begin{pmatrix} 2b_2 & -b_1\\ -b_1 & 0 \end{pmatrix}$ oraz $\begin{pmatrix} 0 & -b_2\\ -b_2 & 2b_1 \end{pmatrix}$.

§11 Grupy operatorowe. Orientacja

107 Bazy (e_i) i (f_i) są zorientowane jednakowo, a bazy (e_i) i (g_i) – przeciwnie, więc bazy (f_i) i (g_i) mają przeciwną orientację.

110 Gdy $\operatorname{sgn} \pi = +1$.

§12 Sumy proste przestrzeni i operatorów. Operatory rzutowe

111 Nie.

112 Kolejno bazy V_1, V_2, $V_1 \cap V_2$, W_1, W_2:

(a) (a_1, a_2), (b_1, b_2), $(b_1 - b_2) = (2a_1 - a_2)$, (a_1), (b_1);

(b) (a_1, a_2, a_3), (b_1, b_2, b_3), $(b_1, b_3) = (a_1 - a_3, 2a_1 + 3a_2)$, (a_1), (b_2).

113 $\mathbf{P} = \frac{1}{4}\begin{pmatrix} 10&3&3&-21\\ -10&-1&-1&27\\ 4&2&2&-6\\ 2&1&1&-3 \end{pmatrix}$, $\mathbf{1} - \mathbf{P}$.

115 (a) Wykazać, że $\operatorname{Ker} A \cap \operatorname{Im} A = A \operatorname{Ker} A^2$. Jeśli (ii) spełniony, to przedstawić $V = \operatorname{Ker} A \oplus \operatorname{Im} A \oplus W'$. (b) Dla $\dim V < \infty$ wykorzystać związek wymiarów jądra i obrazu.

116 (b) $\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$.

117 Niech W będzie podprzestrzenią inwariantną, a e_m – wektorem o najwyższym numerze występującym w rozkładach wektorów tej podprzestrzeni. Wykazać, że wówczas $W = L(e_1, \ldots, e_m)$.

118 Każdy wektor generuje podprzestrzeń inwariantną operatora, dla której jest cykliczny.

119 (a) Zauważyć, że $\dim \mathcal{L}(V, V) < \infty$. (b) Dowód przez indukcję względem wymiaru przestrzeni. Zbudować na dowolnym niezerowym wektorze x podprzestrzeń niezmienniczą operatora A i niech W będzie wielomianem minimalnego stopnia, dla którego $W(A)x = 0$, zgodnie z wynikiem zadania 103. Jeśli stopień wielomianu W jest równy m, to $\dim \operatorname{Ker} W(A) \geq m$, $\dim \operatorname{Im} W(A) \leq \dim V - m$. Rozważyć A na $\operatorname{Im} W(A)$.

121 Składać warunek $\alpha P + \beta Q = 0$ z operatorami P i Q.

122 Wykazać, że $P(R^2 - R) - (R^2 - R)P = \beta(\alpha + \beta - 1)(PQ - QP)$, więc warunkiem koniecznym jest, aby $PQ - QP = 0$ lub $\alpha + \beta - 1 = 0$. Jeśli $PQ = QP$, to wykazać, że PQ jest operatorem rzutowym, i skorzystać z wyniku poprzedniego zadania w równaniach $P(R^2 - R) = 0$, $Q(R^2 - R) = 0$. Jeśli $\alpha + \beta = 1$, to wyeliminować współczynnik β z równania $R^2 - R = 0$.

123 (a) W przestrzeni o bazie $\{e_k\}$, $k \in \mathbb{Z}$, rozważyć operator zadany przepisem $Te_k = e_{k+1}$.

125 Wykazać, że $\operatorname{Ker} A \neq \{0\}$ i uzupełnić $V = \operatorname{Ker} A \oplus W$. Stosując wynik poprzedniego zadania dokończyć dowód przez indukcję.

§13 Zagadnienie własne operatora liniowego

126 (c) operator niediagonalizowalny, podprzestrzenie własne: $L(2e_1 + e_3)$ – do wartości własnej -2, $L(15e_1 - 3e_2 + 5e_3)$ – do wartości 1.

$$\text{(a) } \boldsymbol{\beta} = \begin{pmatrix} 2 & 2 & -8 \\ 1 & 1 & 11 \\ -1 & 8 & 4 \end{pmatrix}, \ \mathbf{A} = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 11 \end{pmatrix}, \ \text{(b) } \boldsymbol{\beta} = \begin{pmatrix} 2 & -3 & 2 \\ 1 & 0 & -3 \\ 0 & 1 & 5 \end{pmatrix}, \ \mathbf{A} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 21 \end{pmatrix}.$$

127 Dwuwymiarowe podprzestrzenie $V_j = L(e_j, e_{n-j+1})$, $j < (n+1)/2$, są inwariantne; operator jest diagonalizowalny wtedy, i tylko wtedy, gdy jest diagonalizowalny w każdej z nich. Dla n nieparzystych podprzestrzeń $L(e_{(n+1)/2})$ jest podprzestrzenią własną do wartości własnej $\lambda_{(n+1)/2}$.

Diagonalizowalność w podprzestrzeni V_j jest równoważna spełnieniu warunku: albo $\lambda_j = \lambda_{n-j+1} = 0$, i wtedy operator jest zerowy na tej podprzestrzeni, albo $\lambda_j \lambda_{n-j+1} \neq 0$ w przypadku zespolonym oraz $\lambda_j \lambda_{n-j+1} > 0$ w przypadku rzeczywistym, i wtedy przy oznaczeniu $\mu_j = \sqrt{\lambda_j \lambda_{n-j+1}}$:

$$e'_j = e_j + (\lambda_j/\mu_j)e_{n-j+1}, \quad e'_{n-j+1} = -(\mu_j/\lambda_j)e_j + e_{n-j+1},$$
$$Ae'_j = \mu_j e'_j, \quad Ae'_{n-j+1} = -\mu_j e'_{n-j+1}.$$

128 Wektory $f(\varepsilon)$, gdy ε przebiega zbiór wszystkich pierwiastków, tworzą bazę (patrz zadanie 71), oraz $Af(\varepsilon) = \lambda(\varepsilon)f(\bar{\varepsilon})$, gdzie $\lambda(\varepsilon) = \sum_{i=1}^n z_i\varepsilon^{i-1}$. Stąd $f(1)$, $f(-1)$ (dla parzystego n) są wektorami własnymi, a podprzestrzeń rozpięta pozostałymi wektorami jest inwariantna względem A i na niej sytuacja jest jak w poprzednim zadaniu. Operator jest diagonalizowalny wtedy, i tylko wtedy, gdy dla każdego ε jest albo $\lambda(\varepsilon) = \lambda(\bar{\varepsilon}) = 0$, albo obie liczby są różne od zera. W szczególności zachodzi to w dwóch przypadkach: (i) $\det A \neq 0$, (ii) wszystkie liczby z_i są rzeczywiste – wtedy $\lambda(\bar{\varepsilon}) = \overline{\lambda(\varepsilon)}$.

129 (b) $\mathbf{A} = \left(\begin{smallmatrix} 0 & \mathbf{1}_m \\ 0 & 0 \end{smallmatrix}\right)$.

130 Zauważyć, że w danych warunkach wielomian charakterystyczny ma dodatkowy pierwiastek.

132 Dowód przez indukcję względem wymiaru przestrzeni. W zadaniu 116 podstawić za W_1 dowolną podprzestrzeń własną operatora (istnieje co najmniej jedna) i skorzystać z założenia indukcyjnego dla operatora A_{22}.

133 Oznaczmy funkcje: $e_\kappa(x) = \exp(\kappa x)$, $f_k^1(x) = \cos(kx)$, $f_k^2(x) = \sin(kx)$, gdzie $\kappa \in \mathbb{R}$, $k \in \mathbb{R}_+$. Podprzestrzenie $L(e_\kappa)$ są własne dla D do wartości κ i dla D^2 do wartości κ^2. Dodatkowo podprzestrzenie $L(f_k^1, f_k^2)$ są własne dla D^2 do wartości $-k^2$. Przecięcie sumy tych przestrzeni z przestrzenią wielomianów jest równe $L(e_0)$.

134 Jeśli $P(A)x = \mu x$, to rozważyć problem na podprzestrzeni $W = L(x, Ax, A^2x, \ldots)$. Przestrzeń ta jest inwariantna względem A i na niej $P(A_W) = \mu\,\mathrm{id}$. Twierdzenie nie zachodzi dla rzeczywistego operatora, który w pewnej bazie ma macierz $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

135 W przypadku jednowymiarowej podprzestrzeni własnej baza przestrzeni (e_1, e_2), gdzie e_2 jest dowolnym wektorem spoza tej podprzestrzeni, a e_1 – odpowiednio dobranym wektorem własnym.

136 Wykazać, że $\left(\begin{smallmatrix} \lambda & \mu \\ 0 & \lambda \end{smallmatrix}\right)^n = \left(\begin{smallmatrix} \lambda^n & n\mu\lambda^{n-1} \\ 0 & \lambda^n \end{smallmatrix}\right)$.

137 Jeśli $B \neq b\,\mathrm{id}$, to istnieje baza, w której macierz operatora ma jedną z postaci kanonicznych: (i) $\mathbf{B} = \left(\begin{smallmatrix} b_1 & 0 \\ 0 & b_2 \end{smallmatrix}\right)$, $b_1 \neq b_2$, lub (ii) $\mathbf{B} = \left(\begin{smallmatrix} b & 1 \\ 0 & b \end{smallmatrix}\right)$. Z koniecznego warunku $[A, B] = 0$ dostaje się w tych przypadkach (i) $\mathbf{A} = \left(\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right)$, lub (ii) $\mathbf{A} = \left(\begin{smallmatrix} a & c \\ 0 & a \end{smallmatrix}\right)$ odpowiednio. Rozwiązanie jest określone przez warunki na liczby a_1, a_2 lub a, c odpowiednio (jeśli mają rozwiązanie) (i) $f(a_i) = b_i$, $i = 1, 2$, oraz (ii) $f(a) = b$, $cf'(a) = 1$.

Jeśli $B = b\,\mathrm{id}$, to A ma jedną wartość własną. Jeśli A jest diagonalizowalny, to $A = a\,\mathrm{id}$, $f(a) = b$, jeśli ten warunek ma rozwiązanie. Jeśli A nie jest diagonalizowalny, to w pewnej bazie $\mathbf{A} = \left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right)$. Rozwiązanie określone warunkami (jeśli mają rozwiązanie) $f(a) = b$, $f'(a) = 0$. W tym przypadku baza, w której A przyjmuje postać kanoniczną jest dowolna.

138 Dla $i = 2, \ldots, s$ jest $\lambda_i/\lambda_1 \neq 1$ i $|\lambda_i/\lambda_1| \leq 1$, więc suma $\sum_{k=0}^{N-1} (\lambda_i/\lambda_1)^k$ jest ograniczona ze względu na N.

§14 Iloczyny skalarne

141

$$\text{(a)}\ e'_1 = e_1,\ e'_2 = e_2,\ e'_3 = 5e_1 + 2e_2 + 3e_3,\ \begin{pmatrix} 0 & 3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{(b)}\ e'_1 = e_1,\ e'_2 = e_2,\ e'_3 = -e_1 + 3e_2 - 5e_3,\ \begin{pmatrix} -8 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{(c)}\ e'_1 = e_1,\ e'_2 = e_2,\ e'_3 = -e_1 + 3e_2 - 4e_4,\ e'_4 = -2e_1 + e_3,\ \begin{pmatrix} 3 & 1 & 0 & 0 \\ 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

142

$$\text{(a)}\ \begin{pmatrix} 1 & 4 & 0 & -1 \\ 0 & 2 & 1 & -\frac{3}{5} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} \end{pmatrix},\ \begin{pmatrix} 0 & \mathbf{1}_2 \\ -\mathbf{1}_2 & 0 \end{pmatrix} \quad \text{(b)}\ \begin{pmatrix} -\frac{1}{2} & \frac{3}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & \frac{1}{11} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{11} \end{pmatrix},\ \begin{pmatrix} 0 & \mathbf{1}_2 \\ -\mathbf{1}_2 & 0 \end{pmatrix}.$$

143

$$\text{(a)}\ \begin{matrix} f_3 = 2e_1 - e_3 - e_4, \\ f_4 = \frac{1}{5}(4e_2 - e_3 - 6e_4), \end{matrix} \quad \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 3 \end{pmatrix},$$

$$\text{(b)}\ \begin{matrix} f_3 = e_1 + 2e_2 - 8e_4, \\ f_4 = -2e_2 + e_3 + 5e_4, \end{matrix} \quad \begin{pmatrix} 3 & 3 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & 0 & 21 & -21 \\ 0 & 0 & -21 & 18 \end{pmatrix}.$$

145 (b) Jeśli $\mathbf{g}$ jest macierzą metryki g w bazie kanonicznej, to macierze metryk $g_\pm$ w bazie skonstruowanej tak, jak w odpowiedzi do zadania 88, są równe $\begin{pmatrix} \mathbf{g} & 0 \\ 0 & \mathbf{g} \end{pmatrix}$ i $\begin{pmatrix} 0 & \mathbf{g} \\ -\mathbf{g} & 0 \end{pmatrix}$ odpowiednio.

§15 Przestrzenie ortogonalne i hermitowskie

146

$$\text{(a)}\ \begin{pmatrix} 1 & -1 & -1 & -3 \\ 0 & 1 & 1 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & 1 \\ 1 & -\sqrt{2} & -1 \\ 0 & \frac{1}{2\sqrt{2}} & 0 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$\text{(c)}\ \begin{pmatrix} 1 & 1 & 3 & 8 \\ 0 & 1 & 1 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{(d)}\ \begin{pmatrix} \frac{1}{2} & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

147

$$\text{(a)}\ \begin{pmatrix} 1 & -i & \frac{-1+i}{\sqrt{2}} \\ 0 & 1 & \frac{-1-i}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{(b)}\ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1+i}{2} & \frac{1-i}{2} \\ \frac{1}{\sqrt{2}} & \frac{1-i}{2} & \frac{-1-i}{2} \\ 0 & 1 & 0 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

151

$$\text{(a)}\ \begin{pmatrix} 1 & \frac{1}{2} & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -9 \end{pmatrix},\ \text{(b)}\ \begin{pmatrix} 1 & -1-i & 1-2i \\ 0 & 1 & 2i \\ 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix},$$

$$\text{(c)}\ \begin{pmatrix} 1 & \frac{2}{3} & 3 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

152 (a) $\lambda > 58$, (b) nie istnieją.
153 Zastosować kryterium Sylwestra; $1 > \lambda > -1/(\dim V - 1)$.
154 Kolumny współrzędnych w bazie (e_i) wektorów otrzymanych układów:

$$\text{(a)}\ \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},\ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ \text{(b)}\ \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ \tfrac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ -1 \\ 0 \end{pmatrix},\ \tfrac{1}{2\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 1 \\ 3 \end{pmatrix},\ \tfrac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix},$$

$$\text{(c)}\ \tfrac{1}{\sqrt{3}}\begin{pmatrix} i \\ 1-i \\ 0 \end{pmatrix},\ \tfrac{1}{2\sqrt{6}}\begin{pmatrix} 3+i \\ 1+2i \\ -3i \end{pmatrix},\ \tfrac{1}{2\sqrt{2}}\begin{pmatrix} 1-i \\ 1 \\ 2+i \end{pmatrix},$$

$$\text{(d)}\ \tfrac{1}{\sqrt{2}}\begin{pmatrix} i \\ 0 \\ 1 \\ 0 \end{pmatrix},\ \tfrac{1}{2\sqrt{3}}\begin{pmatrix} 1-i \\ 2i \\ 1+i \\ -2i \end{pmatrix},\ \tfrac{1}{4\sqrt{3}}\begin{pmatrix} 3-i \\ 1+i \\ 1+3i \\ -1+5i \end{pmatrix}.$$

156 Wyliczyć kwadrat normy wektora $x - \sum_{i=1}^{k}(f_i, x)f_i$.

157 Jeśli x jest czasowy lub przestrzenny, to wybrać bazę Minkowskiego tak, aby x był proporcjonalny do jednego z wektorów bazy. Jeśli x jest świetlny, a e_0 jest dowolnym unormowanym wektorem czasowym, to znaleźć unormowany wektor e_3 leżący w $L(e_0, x)$ i prostopadły do e_0. Uzupełnić do bazy Minkowskiego i wykazać, że wektory x, e_1, e_2 tworzą bazę przestrzeni $L(x)^\perp$.

158 Jeśli x czasowy, a y świetlny, to dowód pozostaje niezmieniony (przy czym wykluczona jest liniowa zależność). Jeśli oba świetlne, to nierówność słaba oczywista. Stwierdzenie o równości wynika z nierówności Schwarza w podprzestrzeni $L(t)^\perp$, dla dowolnie wybranego czasowego t.

159 W postaci kanonicznej metryki na dowolnej podprzestrzeni nie mogą pojawić się pary $(0,0)$, $(1,0)$, $(1,1)$ (poprzednie zadanie). Dla dowolnej bazy Minkowskiego wektor $e_0 - e_3$ jest świetlny. Postacie kanoniczne podprzestrzeni i ich ortogonalnych dopełnień: $(1,-1,-1)^\perp = (-1)$, $(-1,-1,-1)^\perp = (1)$, $(-1,-1,0)^\perp = (0)$, $(1,-1)^\perp = (-1,-1)$, $(-1,-1)^\perp = (1,-1)$, $(-1,0)^\perp = (-1,0)$ (ale $W \neq W^\perp$).

160 Jeśli l_i, $i = 1,2,3$, są wektorami świetlnymi i $\alpha_1 l_1 + \alpha_2 l_2 + \alpha_3 l_3 = 0$, to przenieść jeden ze składników kombinacji liniowej na drugą stronę i wyliczyć kwadrat obu stron otrzymanego równania.

161 W dowodzie użyć pomocniczego wektora t niezależnego od x. Założenie oznacza, że w trójce wektorów nie występują dwa liniowo zależne wektory świetlne.

162 Wykazać niezależność warunku od wektora czasowego t oraz przechodniość relacji.

163 Sprowadzić do postaci kanonicznej formę kwadratową zadaną tym warunkiem.

§16 Odwzorowania liniowe przestrzeni z iloczynem skalarnym

164

$$(a)\ \frac{1}{4}\begin{pmatrix} -8 & 8 & 0 \\ -14 & 16 & 0 \\ 13 & -2 & 4 \end{pmatrix}, \quad (b)\ \frac{1}{2}\begin{pmatrix} 2+i & 2+i & i \\ -1+i & -3+i & -1-3i \\ -1 & -1-2i & 1 \end{pmatrix}.$$

165 (a) Konsekwencja niezdegenerowania iloczynu skalarnego. (b) Nie istnieje funkcja h, dla której $\int \overline{h(x)} f(x) dx = f(0)$.

170 Posłużyć się wielomianem charakterystycznym i przedstawieniem macierzowym operatorów.

171 Jeśli W jest skończenie wymiarową inwariantną podprzestrzenią izometrii wewnętrznej A, to A_W jest izomorfizmem, więc $AW = W$. Zastosować definicję izometrii do wektora $x \in W$ i $y \in W^\perp$.

172 Wszystkie stwierdzenia są konsekwencjami połączonych rezultatów wcześniejszych zadań.

173 Wynik jest konsekwencją poprzedniego zadania. Jeśli $Ax = \lambda x$, $\lambda \neq \pm 1$, to istnieje y, dla którego $Ay = \lambda^{-1} y$ i oba wektory są świetlne. Podprzestrzenie $L(x, y)$ i $L(x, y)^\perp$ są inwariantne. Dobierając kombinacje liniowe wektorów x, y i uzupełniając wektorami ortogonalnymi można tak wybrać bazę Minkowskiego, aby $L(x, y) = L(e_0, e_3)$ i $L(x, y)^\perp = L(e_1, e_2)$. Ponieważ A jest ortochroniczna, to $\lambda > 0$. Stąd A zacieśniona do $L(e_0, e_3)$ jest właściwa, więc również zacieśniona do $L(e_1, e_2)$ jest właściwa.

174 Uogólnione sprzężenie jest operacją liniową (lub antyliniową). Jeśli użyć tych samych oznaczeń dla przestrzeni V i $\tilde{V}$ zamienionych rolami, to $(A^*)^* = A^*$. Jeśli $(\hat{V}, \hat{g})$ jest następną przestrzenią o własnościach takich, jak dwie poprzednie, i $B \in \mathcal{L}(\tilde{V}, \hat{V})$, to $(BA)^* = A^* B^*$.

175 (a) Wynik jest konsekwencją wcześniejszych zadań. (b) Jeśli W jest podprzestrzenią inwariantną i $\{P, P'\}$ jest ortogonalnym rozkładem identyczności związanym z rozkładem $V = W \oplus W^\perp$, to $AP = PAP$ i $AP' = P'AP'$, skąd $AP = PA$. Odwrotnie, jeśli $B \neq \alpha\, \mathrm{id}$, to B ma podprzestrzeń własną $W \neq V$. Jeśli przy tym $[A, B] = 0$, to $AW \subseteq W$.

176 (a) Niech (e_i) będzie dowolną bazą ortonormalną w V. Korzystając z własności normy w V wykazać, że $\|Ax\| \leq \sum_i \|Ae_i\|\,\|x\|$ dla dowolnego wektora x. (b) (iii) Wykazać, że jeśli f, g są rzeczywistymi funkcjami na dowolnym zbiorze X, to zachodzi $\sup (f + g)(X) \leq \sup f(X) + \sup g(X)$. (c) Korzystać z nierówności Schwarza, a następnie podstawić $x = Ay$.

§17 Operatory normalne w przestrzeni unitarnej i euklidesowej

177

$$\text{(a)}\ \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{14}} & \frac{3}{\sqrt{35}} \\ \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{35}} \\ 0 & \frac{2}{\sqrt{14}} & \frac{-5}{\sqrt{35}} \end{pmatrix},\ \begin{pmatrix} -10 & 0 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 25 \end{pmatrix},\ \text{(b)}\ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix},\ \begin{pmatrix} -6 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{(c)}\ \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{6}{\sqrt{70}} & \frac{2}{\sqrt{14}} \\ \frac{-2}{\sqrt{5}} & \frac{3}{\sqrt{70}} & \frac{1}{\sqrt{14}} \\ 0 & \frac{5}{\sqrt{70}} & \frac{-3}{\sqrt{14}} \end{pmatrix},\ \begin{pmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 9 \end{pmatrix},\ \text{(d)}\ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{2\sqrt{3}} & \frac{-1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \\ 0 & 0 & \frac{3}{2\sqrt{3}} & \frac{-1}{2} \end{pmatrix},\ \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&4 \end{pmatrix},$$

$$\text{(e)}\ \tfrac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix},\ \begin{pmatrix} a+b+c & 0 & 0 & 0 \\ 0 & a-b-c & 0 & 0 \\ 0 & 0 & b-c-a & 0 \\ 0 & 0 & 0 & c-a-b \end{pmatrix},$$

$$\text{(f)}\ \begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{-1}{\sqrt{35}} & \frac{-2}{\sqrt{105}} & \frac{1}{\sqrt{30}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{35}} & \frac{-6}{\sqrt{105}} & \frac{3}{\sqrt{30}} \\ 0 & \frac{5}{\sqrt{35}} & \frac{-4}{\sqrt{105}} & \frac{2}{\sqrt{30}} \\ 0 & 0 & \frac{7}{\sqrt{105}} & \frac{4}{\sqrt{30}} \end{pmatrix},\ \begin{pmatrix} -5 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}.$$

178 Operatory (a), (b), (e) i (f) są hermitowskie, operatory (c) i (g) – unitarne, a operator (d) jest normalny, żadnego z wcześniejszych typów.

$$\text{(a)}\ \begin{pmatrix} \frac{2i}{\sqrt{5}} & \frac{-1-i}{\sqrt{35}} & \frac{1}{\sqrt{7}} \\ \frac{1}{\sqrt{5}} & \frac{2-2i}{\sqrt{35}} & \frac{2i}{\sqrt{7}} \\ 0 & \frac{5}{\sqrt{35}} & \frac{1-i}{\sqrt{7}} \end{pmatrix},\ \begin{pmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&8 \end{pmatrix},\ \text{(b)}\ \begin{pmatrix} \frac{i}{\sqrt{2}} & \frac{-2-i}{\sqrt{14}} & \frac{1}{\sqrt{7}} \\ \frac{1}{\sqrt{2}} & \frac{1-2i}{\sqrt{14}} & \frac{i}{\sqrt{7}} \\ 0 & \frac{2}{\sqrt{14}} & \frac{2-i}{\sqrt{7}} \end{pmatrix},\ \begin{pmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&7 \end{pmatrix},$$

$$\text{(c)}\ \frac{1}{\sqrt{82+6\sqrt{41}}}\begin{pmatrix} 3+\sqrt{41} & -4(1+i) \\ 4(1-i) & 3+\sqrt{41} \end{pmatrix},\ \frac{1}{10}\begin{pmatrix} -3+\sqrt{41}+(3+\sqrt{41})i & 0 \\ 0 & -3-\sqrt{41}+(3-\sqrt{41})i \end{pmatrix},$$

$$\text{(d)}\ \frac{1}{\sqrt{2}}\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix},\ \begin{pmatrix} i & 0 \\ 0 & 2+i \end{pmatrix},\ \text{(e)}\ \begin{pmatrix} \cos\frac{\vartheta}{2} & -\sin\frac{\vartheta}{2}e^{-i\varphi} \\ \sin\frac{\vartheta}{2}e^{i\varphi} & \cos\frac{\vartheta}{2} \end{pmatrix},\ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\text{(f)}\ \frac{1}{2\sqrt{2}}\begin{pmatrix} -1 & \sqrt{3}\bar\omega & -\sqrt{3}\bar\omega^2 & \bar\omega^3 \\ \sqrt{3}\omega & -1 & -\bar\omega & \sqrt{3}\bar\omega^2 \\ -\sqrt{3}\omega^2 & -\omega & 1 & \sqrt{3}\bar\omega \\ \omega^3 & \sqrt{3}\omega^2 & \sqrt{3}\omega & 1 \end{pmatrix},\ \begin{pmatrix} -3&0&0&0 \\ 0&-1&0&0 \\ 0&0&1&0 \\ 0&0&0&3 \end{pmatrix},$$

$$\text{(g)}\ \tfrac{1}{2}\begin{pmatrix} -1+i & 1-i & 0 & 0 \\ i & i & i & i \\ 0 & 0 & -1+i & 1-i \\ -1 & -1 & 1 & 1 \end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & -1-i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 0 & -1+i \end{pmatrix}.$$

180 $\left\{\frac{1}{2}\Big(\mu+\nu\pm\sqrt{(\mu-\nu)^2+4|(x,y)|^2\mu\nu}\Big),0\right\}$, bez ostatniej wartości jeśli wektory x, y rozpinają całą przestrzeń.

181 Dla dowodu pierwszego stwierdzenia i odpowiedzi na ostatnie pytanie wykorzystać rozkład spektralny $A=\sum_i \lambda_i P_i$ oraz $\mathrm{id}=\sum_i P_i$; wziąć pod uwagę oba równania $U^*U=\mathrm{id}$, $UU^*=\mathrm{id}$. $B=(\alpha^2\,\mathrm{id}-A^2)^{1/2}$.

182 Wykorzystać wynik poprzedniego zadania.

$$A = \tfrac{3}{2}(U_1 + U_1^*) + 3i(U_2 + U_2^*),$$
$$\mathbf{U}_1 = \tfrac{1}{3\sqrt{5}} \begin{pmatrix} \sqrt{5}+4i & -2+2\sqrt{5}i \\ 2-2\sqrt{5}i & -2\sqrt{5}+i \end{pmatrix}, \quad \mathbf{U}_2 = \tfrac{1}{6} \begin{pmatrix} -5+i & 3+i \\ 1+3i & -1+5i \end{pmatrix}.$$

183

$$\text{(a) } e'_k = \tfrac{1}{\sqrt{k(k+1)}}(e_1 + \ldots + e_k - ke_{k+1}),\ k = 1, \ldots, n-1,$$
$$e'_n = \tfrac{1}{\sqrt{n}}(e_1 + \ldots + e_n), \quad g(x,x) = -\sum_{i=1}^{n-1}(x'^i)^2 + (n-1)(x'^n)^2,$$
$$\text{(b) } e'_1 = \tfrac{1}{\sqrt{2(1+a^3)}}(a^1 e_1 + a^2 e_2) + \sqrt{\tfrac{1+a^3}{2}}\, e_3,$$
$$e'_2 = \tfrac{1}{\sqrt{2(1-a^3)}}(a^1 e_1 + a^2 e_2) - \sqrt{\tfrac{1-a^3}{2}}\, e_3,\ e'_3 = \tfrac{1}{\sqrt{1-(a^3)^2}}(a^2 e_1 - a^1 e_2),$$
$$g(x,x) = a^3\big[-(x'^1)^2 + (x'^2)^2\big] + (x'^3)^2,$$
$$\text{(c) ozn.: } a + b = \alpha,\ a - b = \rho\cos\varphi,\ c = \rho\sin\varphi,\ \rho > 0,$$
$$e'_1 = \tfrac{1}{\sqrt{2}}\cos\tfrac{\varphi}{2}\,(e_1 + e_3) + \tfrac{1}{\sqrt{2}}\sin\tfrac{\varphi}{2}\,(e_2 + e_4),$$
$$e'_2 = -\tfrac{1}{\sqrt{2}}\sin\tfrac{\varphi}{2}\,(e_1 + e_3) + \tfrac{1}{\sqrt{2}}\cos\tfrac{\varphi}{2}\,(e_2 + e_4),$$
$$e'_3 = \tfrac{1}{\sqrt{2}}(e_1 - e_3),\ e'_4 = \tfrac{1}{\sqrt{2}}(e_2 - e_4),\ g(x,x) = (\alpha+\rho)(x'^1)^2 + (\alpha-\rho)(x'^2)^2.$$

184

$$\text{(a) } \begin{pmatrix} 0 & \frac{-1}{2\sqrt{3}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{\sqrt{3}}{2} & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}, \text{ (b) } \begin{pmatrix} \frac{2}{\sqrt{26}} & \frac{-34}{\sqrt{91}} & \frac{-4}{\sqrt{14}} \\ \frac{-3}{\sqrt{26}} & \frac{25}{\sqrt{91}} & \frac{5}{\sqrt{14}} \\ 0 & \frac{13}{\sqrt{91}} & \frac{4}{\sqrt{14}} \end{pmatrix}, \begin{pmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 10 \end{pmatrix},$$
$$\text{(c) } \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{4}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\sqrt{3} & 0 \\ 0 & \frac{5}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

185

$$\text{(a) } \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}, \text{ (b) } \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{38}} & \frac{-3}{\sqrt{19}} \\ 0 & \frac{-6}{\sqrt{38}} & \frac{-1}{\sqrt{19}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{38}} & \frac{-3}{\sqrt{19}} \end{pmatrix}, \begin{pmatrix} \frac{-9}{10} & \frac{-\sqrt{19}}{10} & 0 \\ \frac{\sqrt{19}}{10} & \frac{-9}{10} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\text{(c) } \begin{pmatrix} \frac{-1}{2} & 0 & \frac{\sqrt{2}+1}{\sqrt{8}} & \frac{-\sqrt{2}+1}{\sqrt{8}} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\ \frac{1}{2} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\ \frac{1}{2} & 0 & \frac{\sqrt{2}-1}{\sqrt{8}} & \frac{-\sqrt{2}-1}{\sqrt{8}} \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

186

$$\begin{pmatrix} \cos\frac{\varphi}{2} & -\sin\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix}$$

187
$$(a)\ |\mathbf{A}| = \sqrt{\tfrac{2}{5}}\left(\begin{smallmatrix} 2 & 1 \\ 1 & 3 \end{smallmatrix}\right),\ \mathbf{U} = \tfrac{1}{\sqrt{10}}\left(\begin{smallmatrix} 1 & 3 \\ 3 & -1 \end{smallmatrix}\right),$$
$$(b)\ |\mathbf{A}| = \begin{pmatrix} 2 & i & 0 \\ -i & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ \mathbf{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i & 1 \\ 0 & 1 & i \\ \sqrt{2} & 0 & 0 \end{pmatrix}.$$

188
$$U' = U,\ |A|' = U|A|U^*.$$

189 Rozważyć rozkład spektralny $A^*A = \sum_{i=1}^{s} \lambda_i P_i$. Jeśli λ_1 jest maksymalną wartością, to dla dowolnego wektora x zachodzi $\|Ax\|^2 \leq \lambda_1 \|x\|^2$, przy czym równość zachodzi dla wektorów własnych do wartości λ_1. Stąd $\|A\| = \sqrt{\lambda_1}$.

190 (a) Dla $(P'x, Py) \geq 0$ wykazać: $\frac{(x,P'Py)}{\|x\|\|y\|} \leq \frac{(P'x,Py)}{\|P'x\|\|Py\|} = \frac{(P'x,(P'P)Py)}{\|P'x\|\|Py\|}$, a stąd $\|P'P\| \leq \cos\measuredangle(W',W) \leq \|P'P\|$ (wykorzystać zadanie 176). Wykorzystać wynik poprzedniego zadania. (c) Wykazać, że $\|z'\|^2 = (z,z') = \lambda^2\|z\|^2$.

192
$$(a)\ \begin{pmatrix} 0 & \frac{4}{5} & \frac{-3}{5} \\ 1 & 0 & 0 \\ 0 & \frac{-3}{5} & \frac{-4}{5} \end{pmatrix},\ \begin{pmatrix} 0 & -5 & 0 \\ 5 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ (b)\ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \end{pmatrix},\ \begin{pmatrix} 0 & -\sqrt{3} & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$(c)\ \begin{pmatrix} 0 & \frac{-5}{\sqrt{70}} & \frac{3}{\sqrt{14}} \\ \frac{1}{\sqrt{5}} & \frac{6}{\sqrt{70}} & \frac{2}{\sqrt{14}} \\ \frac{-2}{\sqrt{5}} & \frac{3}{\sqrt{70}} & \frac{1}{\sqrt{14}} \end{pmatrix},\ \begin{pmatrix} 0 & -\sqrt{14} & 0 \\ \sqrt{14} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

§18 Iloczyn tensorowy

193 Wzorować się na dyskusji przeprowadzonej w przykładzie (iii), 7, §18.

196 (a) $(-2e_1 + e_2) \otimes (3f_1 + 5f_2) \otimes (g_1 - 3g_2)$,
(b) $(5e_1 + 2e_2) \otimes (-3f_1 + f_2) \otimes (g_1 - 4g_2)$.

197 Dualność można określić jako $\langle A, B\rangle = \mathrm{Tr}[AB]$.

198 $C(Y \otimes Z) = (-12f_1 + 2f_2 - 2f_3) \otimes h_1 + (39f_1 - 4f_2 + 4f_3) \otimes h_2$.

199 (c) $\mathrm{Tr}[A \otimes B] = (\mathrm{Tr}\, A)(\mathrm{Tr}\, B)$.

200 Jeśli operatory A i B generują odpowiednio grupy $S(t)$ i $T(t)$, to generatorem grupy $S(t) \otimes T(t)$ jest operator $A \otimes \mathrm{id} + \mathrm{id} \otimes B$.

201 (a) Istnienie metryki wykazuje się tak, jak w przypadku euklidesowym, p. 19, §18. Metryka G jest symetryczna gdy obie metryki g i g' są wspólnie symetryczne lub symplektyczne, a symplektyczna – gdy jedna z nich jest symetryczna, a druga symplektyczna. (b) $G_{ik,jl} = g_{ij}g'_{kl}$. W szczególnym przypadku $G_{ii,jj} = g_{ij}g'_{ij}$. (c) Oznaczmy $n = \dim V$ i $n' = \dim V'$ oraz $r = \dim[\mathrm{Ker}\, g]$ i $r' = \dim[\mathrm{Ker}\, g']$. Wówczas $\dim[\mathrm{Ker}\, G] = nn' - (n-r)(n'-r')$, co wyznacza postać kanoniczną metryki G w przypadku, gdy jest ona symplektyczna. Jeśli liczba jedynek i minus jedynek w postaciach kanonicznych metryk symetrycznych g i g' jest dana przez (p,q) i (p',q') odpowiednio, to dla metryki G te liczby wynoszą $(pp' + qq', pq' + qp')$, a jeśli obie metryki są symplektyczne, to te liczby są dane przez $(\frac{1}{2}(n-r)(n'-r'), \frac{1}{2}(n-r)(n'-r'))$. (d) Wykorzystać punkt (b). (e) Metryka G jest hermitowska, wyniki punktów (b) i (d) oraz punktu (c) w części mówiącej o przypadku metryk symetrycznych pozostają w mocy.

203 Kontrakcja.

204 Operacja iloczynu tensorowego odwzorowań liniowych jest odwzorowaniem wieloliniowym z przestrzeni $\mathcal{L}(\mathcal{L}(V_1, W_1), \ldots, \mathcal{L}(V_r, W_r); \mathcal{L}(V, W))$, gdzie $V = V_1 \otimes \ldots \otimes V_r$, $W = W_1 \otimes \ldots \otimes W_r$. Ogólnie, jeśli E_i jest bazą V_i, a F_i – bazą W_i, to baza przestrzeni $\mathcal{L}(V_i, W_i)$ jest utworzona z odwzorowań, które wybranemu wektorowi z E_i przypisują wybrany wektor z F_i, a na pozostałych wektorach bazy E_i zerują się. Wystarczy pokazać, że iloczyny odwzorowań tego typu tworzą bazę przestrzeni $\mathcal{L}(V, W)$.

§19 Tensory

205

$$\text{(a)} \quad t = (-2e'_1 + e'_2 + e'_3) \otimes (-5e'_1 + 3e'_3) \otimes (2e'^1 + 4e'^2) + \\ + (2e'_1 - e'_3) \otimes (-2e'_1 + e'_2 + e'_3) \otimes (-e'^2 + 2e'^3)\,,$$

$$\text{(b)} \quad t = (2e'_1 + e'_2) \otimes (-3e'^1 + 6e'^2 + e'^3) \otimes (e'^1 - e'^2 + 3e'^3) + \\ + e'_1 \otimes e'^3 \otimes (-e'^1 + 2e'^2)\,.$$

206

$$\text{(a)} \quad C^2_1 t = 8e_1 \otimes e^3 + e_2 \otimes e^2\,, \quad C^2_2 t = -7e_1 \otimes (3e^2 + e^3)\,,$$

$$\text{(b)} \quad C^2_1 t = 6(e_1 + 4e_3) \otimes (e^1 + e^2)\,, \quad C^2_2 t = 3(e_1 + 4e_3) \otimes (e^2 - 2e^3)\,.$$

207

$$\text{(a)} \quad t = (e_1 - 3e_3) \otimes e_2 \otimes (2e^1 - e^2) + e_2 \otimes (-3e_1 + e_3) \otimes e^2\,,$$

$$\text{(b)} \quad t = e_2 \otimes (-e^2 + 5e^3) \otimes e^1 + (-e_1 + e_3) \otimes e^1 \otimes e^3\,.$$

208

$$\text{(a)} \quad \hat{g}^1_2 t = -e_2 \otimes e^1 \otimes e^2 + (2e_1 - e_3) \otimes (2e^1 - e^3) \otimes e^3\,,$$

$$C^1_2 \hat{g}^1_2 t = -3e^1 + e^3\,,$$

$$\text{(b)} \quad C^2_1 (t \otimes s) = -3e_2 \otimes e^3 + 4(e_1 - e_3) \otimes (e^1 - 5e^2)\,,$$

$$\text{(c)} \quad \hat{g}^1_1 t = -\tfrac{1}{2}(3e_1 - e_3) \otimes e_2 \otimes e^2 + (-e_2 + e_3) \otimes (2e_1 + 6e_3) \otimes e^3 + \\ + \tfrac{1}{2}(-e_1 + 2e_2 - e_3) \otimes (-3e_1 \otimes e^1 + 5e_2 \otimes e^3 + 4e_3 \otimes e^3)\,,$$

$$C^2_1 \hat{g}^1_1 t = -2e_1 - 5e_2 + 6e_3\,,$$

$$\text{(d)} \quad P^\pi (t \otimes s) = (e_1 - e_3) \otimes e_2 \otimes (e_1 - e_2) \otimes (2e_2 + e_3)\,,$$

$$\text{(e)} \quad t^{13}{}_1 = 1,\ t^{21}{}_1 = -2,\ t^{22}{}_1 = 2,\ t^{31}{}_1 = 6,\ t^{32}{}_1 = -3,\ t^{13}{}_2 = 3,\ t^{21}{}_2 = -4, \\ t^{22}{}_2 = 4,\ t^{31}{}_2 = 6,\ t^{32}{}_2 = -3,\ t^{11}{}_3 = -2,\ t^{12}{}_3 = 2,\ t^{21}{}_3 = -5,\ t^{22}{}_3 = 5, \\ t^{31}{}_3 = 4,\ t^{32}{}_3 = -4,\ t^{33}{}_3 = 1,\ \text{pozostałe} = 0, \quad t^{i1}{}_i = 0,\ t^{i2}{}_i = 0,\ t^{i3}{}_i = 2\,,$$

(f) ozn. $\gamma = -\frac{2}{9}(a+b+c)$; $\quad t^{ii} = \gamma\,,\ t^{12} = t^{21} = t^{34} = t^{43} = a + \gamma\,,$
$t^{13} = t^{31} = t^{24} = t^{42} = b + \gamma\,,\ t^{14} = t^{41} = t^{23} = t^{32} = c + \gamma\,.$

209 W notacji wskaźnikowej: $\ {t^*}^i{}_j = \hat{g}^{il} g_{jk} t^k{}_l$.

210 Zauważyć, że $(g^{\mathrm{T}}x)(y) = g(x, y)$.

211 Dowód przez indukcję, przy wykorzystaniu wyniku przykładu (v), p.16, §18.

212 (b) Warunek definicyjny operatora sprzężonego wystarczy sprawdzić na wektorach iloczynowych.

213 Wystarczy sprawdzić zachowanie iloczynu skalarnego wektorów iloczynowych.

§20 Tensory symetryczne i antysymetryczne

214

$$t_{(i_1\ldots i_{p+1})} = \tfrac{1}{p+1}\left(t_{(i_1\ldots i_p)i_{p+1}} + \textstyle\sum_{s=1}^{p} t_{(i_1\ldots i_{s-1}i_{p+1}i_{s+1}\ldots i_p)i_s}\right) = $$
$$= \tfrac{1}{p+1}\textstyle\sum_{s=1}^{p+1} t_{(i_1\ldots i_{s-1}i_{s+1}\ldots i_{p+1})i_s}\,.$$

215 Wykazać, że prawa strona pierwszej tożsamości zmienia znak przy przestawieniu miejscami dowolnej pary wektorów. Następnie zastosować do niej operator antysymetryzacji i wykorzystać fakt, że antysymetryzację wewnątrz zakresu działania innego operatora antysymetryzacji można pominąć. Analogicznie w przypadku drugiej tożsamości.

216 Wykorzystać wynik zadania 213.

217 (d) Wykazać, że dla dowolnej permutacji μ istnieje taka permutacja ν, że $\nu(\{1,\ldots,r\}) = \{1,\ldots,r\}$ i permutacja $\pi = \mu\circ\nu$ jest w klasie określonej w tym punkcie zadania. (e) Szczególny przypadek poprzedniego punktu.

218 (a) Dla wykazania własności operatorów $a_n(z)$ oznaczmy $Y = \sum_\pi \operatorname{sgn}\pi\,(z, x_{\pi^{-1}(1)})x_{\pi^{-1}(2)}\otimes\ldots\otimes x_{\pi^{-1}(n)}$ i niech σ będzie permutacją liczb $\{2,\ldots,n\}$. Wykazać, że $P^\sigma Y = \operatorname{sgn}\sigma Y$ (numerujemy wektory w iloczynach kolejno liczbami $2,\ldots,n$); analogicznie przy zastąpieniu znaków permutacji jedynkami. Dla wykazania tożsamości dla operatorów a_n^* wykorzystać sprzężenie. (b) Wykorzystać w dowodach tożsamości udowodnione w punkcie (a) oraz w zadaniu 215. Dla operatorów $d_n(z)$ wykazać ciągi równości (w pierwszym ciągu $n\geq 2$, w pozostałych $n\geq 1$):

$$d_{n-1}(x)d_n(y)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] = a_{n-1}(x)a_n(y)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] =$$
$$= a_{n-1}(x)\tfrac{1}{\sqrt{n}}\textstyle\sum_i(-1)^{i-1}(y,z_i)\mathcal{A}_{n-1}[z_1\otimes\ldots\not{z}_i\ldots\otimes z_n] =$$
$$= \frac{1}{\sqrt{n(n-1)}}\textstyle\sum_{i<j}(-1)^{i+j}[(x,z_i)(y,z_j) - (x,z_j)(y,z_i)]\times$$
$$\times\,\mathcal{A}_{n-2}[z_1\otimes\ldots\not{z}_i\ldots\not{z}_j\ldots\otimes z_n]\,,$$

$$d_{n+1}(x)d_n^*(y)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] = a_{n+1}(x)\mathcal{A}_{n+1}a_n^*(y)[z_1\otimes\ldots\otimes z_n] =$$
$$= (x,y)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] - \textstyle\sum_i (x,z_i)\mathcal{A}_n[z_1\otimes\ldots\otimes z_{i-1}\otimes y\otimes z_{i+1}\otimes\ldots\otimes z_n],$$

$$d_{n-1}^*(y)d_n(x)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] = \mathcal{A}_n\, a_{n-1}^*(y)a_n(x)\mathcal{A}_n[z_1\otimes\ldots\otimes z_n] =$$
$$= \textstyle\sum_i (x,z_i)\mathcal{A}_n[z_1\otimes\ldots z_{i-1}\otimes y\otimes z_{i+1}\ldots\otimes z_n]\,.$$

Analogicznie dla operatorów b_n, z zastąpieniem wszystkich znaków permutacji jedynkami. Zauważyć ponadto, że $b_0^*(x) = d_0^*(x) = a_0^*(x)$, $b_1(y) = d_1(y) = a_1(y)$.

219 (a) Argumentacja podobna jak w przypadku antysymetrycznym. (c) Dowód dwoma metodami: (i) indukcyjny: wykazać wzór rekurencyjny $d(n+1,p) = \sum_{k=0}^p d(n,k)$ oraz $d(1,p) = 1$; (ii) kombinatoryczny: zauważyć, że $d(n,p)$ jest liczbą sposobów, na jakie spośród $n-1+p$ uporządkowanych liniowo miejsc można wybrać $n-1$, w których znajdą się linie podziału, a na pozostałych miejscach będą jedynki.

220
$$\text{(a)}\quad (3e_1 - e_2 - 4e_3)\wedge(5e_1 + 3e_3 + e_4)\,,$$
$$\text{(b)}\quad (-4e_1 - 6e_2 + e_3)\wedge(-e_1 - 2e_2 + e_4)\wedge(e_1 + 7e_2 + e_5)\,,$$
$$\text{(c)}\quad \text{anihilator zerowy}\,.$$

221
$$\text{(a)}\quad (e_1\wedge e_2 + e_3\wedge e_4)\wedge(2e_1 + 3e_2 - 4e_3 + 5e_4 - e_5)\,,$$
$$\text{(b)}\quad (10e_1\wedge e_2 + e_1\wedge e_3 + 3e_2\wedge e_4)\wedge(5e_2 + e_5)\,.$$

222
$$\text{(a)}\quad \hat{^*}\varphi = -(e_1 + 3e_4)\wedge(e_2 - 2e_4)\wedge(e_3 + 5e_4)\,,$$
$$\text{(b)}\quad {}^*x = \tfrac{1}{16}(7e^1 + 4e^2)\wedge e^3\wedge(-e^1 + 4e^4)\wedge(e^1 + 4e^5)\,,$$
$$\text{(c)}\quad {}^*t = (3e^1 + 13e^2 + e^3)\wedge(-6e^1 - 13e^2 + e^4)\,,$$
$$\text{(d)}\quad \hat{^*}s = (e_1 + 2e_3 + 6e_4)\wedge(e_2 + 9e_3 + 20e_4)\,.$$

223 (c) Przy ustalonym ciągu $i_1 < \ldots < i_k$ wyrażenie $k!A^{[i_1}{}_{i_1}\ldots A^{i_k]}{}_{i_k} = \sum_{\pi\in S_k}\operatorname{sgn}\pi A^{i_{\pi(1)}}{}_{i_1}\ldots A^{i_{\pi(k)}}{}_{i_k}$ jest minorem głównym o numerach wierszy i kolumn danych liczbami tego ciągu. (e) Wykorzystać wyniki zadań 211 i 81 dla wykazania, że lewa strona jest równa $\alpha(\det A)^m$. Podstawić $A = \lambda\,\mathrm{id}$.

225 Wybrać bazę (e_1, e_2, e_3) ortonormalną, dodatnio zorientowaną, dla której $n = e_3$. Znaleźć macierze operatorów L_n i $R_n(\varphi)$ w tej bazie.

226 Wykazać, że $e^i = [\omega(e_1,\ldots,e_n)]^{-1}\omega(e_1,\ldots,e_{i-1},\cdot,e_{i+1},\ldots,e_n)$. Forma ta jest proporcjonalna do formy ${}^*(e_1\wedge\ldots\wedge e_{i-1}\wedge e_{i+1}\wedge\ldots\wedge e_n)$.

227 Wyliczyć iloczyn wektorowy obu stron równania z odpowiednio dobranym wektorem. Znaleźć szczególne rozwiązanie x_0 spełniające dodatkowy warunek $a\cdot x_0 = 0$. Ogólne rozwiązanie: $x = \frac{b\times a}{\|a\|^2} + ta$, $t\in\mathbb{R}$.

229 (a) $\tilde{e}_1 = [e_1 \cdot (e_2 \times e_3)]^{-1} e_2 \times e_3$, oraz cyklicznie dla pozostałych. Relacja pomiędzy bazami (e_i) i $(\tilde{e}_i)$ jest symetryczna, więc analogiczne wzory obowiązują z bazami zamienionymi rolami.

230 $\frac{1}{\|a_1 \times a_2\|^2} \begin{pmatrix} \|a_2\|^2 & -a_1 \cdot a_2 & 0 \\ -a_2 \cdot a_1 & \|a_1\|^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 \cdot a_1 & a_1 \cdot a_2 & 0 \\ a_2 \cdot a_1 & a_2 \cdot a_2 & 0 \\ 0 & 0 & \|a_1 \times a_2\|^2 \end{pmatrix}^{-1}.$

§21 Przestrzenie afiniczne

231 Wprowadzić dowolny punkt odniesienia.

232 (a) M przyjmuje stałą wartość na prostych $Z + L(p)$. (b) $M(Z) \wedge p = 0$.

234

$$\text{(a) (i) } \overrightarrow{OX} = \overrightarrow{OP} + \textstyle\sum_{i=1}^{3} t^i f_i \,, \ \overrightarrow{OP} = e_1 - e_2 + 7e_3 + e_4 \,,$$
$$f_1 = 2e_1 + e_2 - 5e_3 + 8e_4 \,, \ f_2 = 4e_1 + 2e_2 - 10e_3 - 6e_4 \,,$$
$$f_3 = 2e_1 + e_2 + 3e_3 - 5e_4 \,, \quad \text{(ii) } \varphi(\overrightarrow{OX}) = 3 \,, \ \varphi = e^1 - 2e^2 \,,$$
$$\text{(b) (i) } \overrightarrow{OX} = \overrightarrow{OP} + t^1 f_1 + t^2 f_2 \,, \ \overrightarrow{OP} = 3e_1 - 4e_2 \,, \ f_1 = 2e_1 + 7e_2 - 2e_3 \,,$$
$$f_2 = -2e_1 + 5e_2 - 3e_3 \,, \quad \text{(ii) } \varphi(\overrightarrow{OX}) = -73 \,, \ \varphi = -11e^1 + 10e^2 + 24e^3 \,,$$
$$\text{(c) (i) } \overrightarrow{OX} = \overrightarrow{OP} + t^1 f_1 + t^2 f_2 \,, \ \overrightarrow{OP} = 2e_1 + 4e_2 - 4e_3 - 7e_4 \,,$$
$$f_1 = e_1 + 2e_3 + 5e_4 \,, \ f_2 = 4e_1 + e_2 + 4e_4 \,,$$
$$\text{(ii) } \varphi(\overrightarrow{OX}) = -24 \,, \ \psi(\overrightarrow{OX}) = 26 \,, \ \varphi = 2e^1 - 8e^2 - e^3 \,, \ \psi = 8e^2 + 5e^3 - 2e^4 \,.$$

236 (b) Sprawdzić, że wtedy $\sum_{i=0}^{n} \lambda_{ij}^{-1} = 1$ i $O_j = \sum_{i=0}^{n} \lambda_{ij}^{-1} X_i$.

237 Oznaczmy przez Ω_i zbiór punktów leżących pomiędzy lub na parze hiperpłaszczyzn $P + W_i$ i $P + x_i + W_i$, gdzie $W_i = L(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. Wtedy $\{P + \sum_{i=1}^{n} \lambda_i x_i \mid \lambda_i \in \langle 0, 1 \rangle\} = \bigcap_{i=1}^{n} \Omega_i$.

238 (a) 121. (b) 23.

239 (a) 2. (b) 1.

240

$$\text{(a) przecięcie: } \overrightarrow{OX} = a + t^1 f_1 + t^2 f_2 \,, \ \text{powłoka: } \overrightarrow{OX} = a + \textstyle\sum_{i=1}^{4} t^i f_i \,, \ t^i \in \mathbb{R} \,,$$
$$a = e_1 - e_3 - e_5 \,, \ f_1 = -e_1 + e_2 + e_3 + e_5 \,, \ f_2 = -2e_1 + e_3 + e_4 + 7e_5 \,,$$
$$f_3 = e_5 \,, \ f_4 = e_1 + 3e_3 - 10e_5 \,, \qquad \text{(b) przecięcie: } \overrightarrow{OX} = b + t'^1 f'_1 \,,$$
$$\text{powłoka: } \overrightarrow{OX} = a + \textstyle\sum_{i=1}^{4} t'^1 f'_i \,, \ t'^i \in \mathbb{R} \,, \ b = a + f_1 \,,$$
$$f'_1 = -3f_1 + f_2 = 7e_1 - 21e_3 \,, \ f'_2 = e_2 - 3e_3 \,, \ f'_3 = 2e_3 + e_4 \,, \ f'_4 = f_1 \,,$$

(c) przecięcie: puste, powłoka: cała przestrzeń,

prosta jest równoległa do 2-płaszczyzny.

241 $\overrightarrow{QP} \notin W$, $\overrightarrow{QP'} \notin W'$; szukane proste mają postać $Q + L(a)$, gdzie
(a) $a \in [L(\overrightarrow{QP}) + W] \setminus W$ (ten zbiór jest niepusty),
(b) $a \in \Big(\big[L(\overrightarrow{QP}) + W\big] \cap \big[L(\overrightarrow{QP'}) + W'\big] \Big) \setminus (W \cup W')$ (może być pusty).

242 Szukane punkty tworzą podprzestrzeń afiniczną $\left(\frac{1}{2}P_1+\frac{1}{2}P_2\right)+L(a_1,a_2)$ (wyrażenie w pierwszym nawiasie jest kombinacja afiniczną).

243 Jeśli proste są nierównoległe i nie przecinają się, to wektory a_1, a_2, b tworzą bazę, więc każdy punkt P ma rozkład jak w temacie. Rozwiązać równanie $\overrightarrow{PY_1} \wedge \overrightarrow{PY_2} = 0$. $\overrightarrow{OY_1} = b + \frac{2u_1}{1+\lambda}\, a_1$, $\overrightarrow{OY_2} = -b + \frac{2u_2}{1-\lambda}\, a_2$.

244 (a) Wykazać, że $t = t'\,(u'\cdot u)$, $v = (u'\cdot u)^{-1}u' - u$, $v^2 = (u'\cdot u)^{-2} - 1$, $-1 < v^2 \leq 0$. (b) Wykazać, że $w = w' + \lambda u'$, wykorzystać związki $u' = \gamma(u+v)$, $1 + \gamma^2|v|^2 = \gamma^2$.

§22 Przestrzeń afiniczna euklidesowa

245 Punkty $P' = P + \sum t^i x_i$, $Q' = Q + \sum s^j y_j$ są najbliżej położone, gdy $\overrightarrow{P'Q'}$ jest prostopadły do wszystkich wektorów x_i, y_j; oznaczmy przez d ich odległość.

$$(a)\ d = \sqrt{3}\,,\ P' = P\,,\ Q' = Q + 2y_1 - y_2 : \begin{pmatrix} 3 \\ -1 \\ 2 \\ 3 \end{pmatrix},$$

$$(b)\ d = \sqrt{79}\,,\ P' = P + 11x_1 : \begin{pmatrix} -2 \\ 12 \\ -9 \\ 5 \end{pmatrix},\ Q' = Q + y_2 : \begin{pmatrix} 3 \\ 10 \\ -8 \\ -2 \end{pmatrix},$$

$$(c)\ d = 4\sqrt{23}\,,\ P' = P + 2x_1 + r(x_1+x_2) : \begin{pmatrix} -7 \\ 3 \\ -2 \\ -5 \\ 6 \end{pmatrix} + r\begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \\ 3 \end{pmatrix},$$

$$Q' = Q - y_1 + r(y_1+y_2) : \begin{pmatrix} 9 \\ -1 \\ 2 \\ 3 \\ 2 \end{pmatrix} + r\begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \\ 3 \end{pmatrix},\ r \in \mathbb{R}\,.$$

246 (a) 3, (b) 9, (c) $3\sqrt{35}$, (d) $3\sqrt{3}$.

247 Oznaczmy $n_i = a_i/\|a_i\|$ oraz $\theta = \measuredangle(n_1, n_2)$. Jeśli n_1, n_2 są liniowo niezależne, to oznaczmy $\tilde{n}_1 = (\sin\theta)^{-2}[n_1 - \cos\theta\, n_2]$, $\tilde{n}_2 = (\sin\theta)^{-2}[n_2 - \cos\theta\, n_1]$. Wtedy $O = (\frac{1}{2}P_1 + \frac{1}{2}P_2) + \frac{1}{2}(\overrightarrow{P_1P_2}\cdot\tilde{n}_1)n_1 + \frac{1}{2}(\overrightarrow{P_2P_1}\cdot\tilde{n}_2)n_2$ oraz $2b = \overrightarrow{P_2P_1} - \overrightarrow{P_2P_1}\cdot\tilde{n}_1\, n_1 - \overrightarrow{P_2P_1}\cdot\tilde{n}_2\, n_2$. Jeśli n_1, n_2 są liniowo zależne, to $O \in \frac{1}{2}P_1 + \frac{1}{2}P_2 + L(n_1)$, $2b = \overrightarrow{P_2P_1} - \overrightarrow{P_2P_1}\cdot n_1\, n_1$. Punkty najbliżej położone to w obu przypadkach $O \pm b$.

248 Jeśli a i b są liniowo niezależne, to wektory wodzące punktów najbliżej położonych mają rozkłady $\overrightarrow{OX_0} = \alpha a + \beta b + \mu a \times b$, $\overrightarrow{OY_0} = \alpha a + \beta b + \nu a \times b$. Po wstawieniu tych rozkładów do równań wylicza się współczynniki: $\alpha = \gamma(a\times b)\cdot l$, $\beta = -\gamma(a\times b)\cdot k$, $\mu = \gamma k\cdot b$, $\nu = -\gamma l\cdot a$, gdzie $\gamma = \|a\times b\|^{-2}$. Stąd $\overrightarrow{Y_0X_0} = \frac{l\cdot a + k\cdot b}{\|a\times b\|^2}\, a\times b$, $\|\overrightarrow{Y_0X_0}\| = \frac{|l\cdot a + k\cdot b|}{\|a\times b\|}$. Jeśli $b = a$, to dla każdego punktu X istnieje punkt $Y = X + d$, gdzie $d = \frac{a\times(l-k)}{\|a\|^2}$, położony w minimalnej od niego odległości równej $\|d\| = \frac{\|l-k\|}{\|a\|}$. Dla podanych wartości:

$$\overrightarrow{OX_0} = \tfrac{1}{49}(43a + 51b + 3a \times b) = \tfrac{1}{49}(147e_1 + 95e_2 - 26e_3),$$
$$\overrightarrow{Y_0X_0} = -\tfrac{1}{49}a \times b = \tfrac{1}{49}(2e_1 - 3e_2 + 6e_3), \quad \|\overrightarrow{Y_0X_0}\| = \tfrac{1}{7}.$$

249 Rozwiązanie dwoma metodami. (i) Zastosować wynik poprzedniego zadania. Warunek konieczny i wystarczający aby powłoka afiniczna była płaszczyzną: $a \cdot l + b \cdot k = 0$. Proste przecinają się w jednym punkcie wtedy i tylko wtedy, gdy ten warunek jest spełniony i wektory a, b są liniowo niezależne; wówczas $X_0 = Y_0$ z poprzedniego zadania jest punktem przecięcia. (ii) Wykazać, że jeśli istnieje punkt Z spełniający oba równania prostych, to warunek jest spełniony. Odwrotnie, niech warunek będzie spełniony. Załóżmy dodatkowo, że wektory a, b są liniowo niezależne (w przeciwnym razie proste są równoległe). Wykazać, że są wtedy następujące dwie możliwości. (a) $b \cdot k \neq 0$ i wtedy wektory k i l są liniowo niezależne oraz $\overrightarrow{OZ} \cdot k = \overrightarrow{OZ} \cdot l = 0$. W tym przypadku $\overrightarrow{OZ} = \frac{k \times l}{b \cdot k}$. (b) $b \cdot k = a \cdot l = 0$ i wtedy k i l są proporcjonalne do $a \times b$. Znaleźć stąd punkt przecięcia i porównać z wynikiem pierwszej metody.

250 Wykazać, że punkt P jest miejscem przecięcia dwóch prostych (jeśli istnieje) jak w poprzednim zadaniu, z zastąpieniem $l \to l - c \times b$.

251 (a) $\mathcal{L} = P + L(\overrightarrow{PQ})$, $\mathcal{N} = R + L(\overrightarrow{RS}, \overrightarrow{PQ})$. (b) $\frac{|\overrightarrow{PR} \cdot (\overrightarrow{RS} \times \overrightarrow{PQ})|}{\|\overrightarrow{RS} \times \overrightarrow{PQ}\|} = \frac{1}{\sqrt{13}}$.

252 Przecięcie jest prostą o równaniu (i) $\overrightarrow{OY} \times (b_1 \times b_2) = \mu_2 b_1 - \mu_1 b_2$.

253 Wykazać, że punkty Y tworzą prostą będącą przecięciem płaszczyzn jak w poprzednim zadaniu, z zastąpieniem $\mu_2 \to \mu_2 - a \cdot b_2$.

254 W przypadku dwóch nierównoległych płaszczyzn $P_i + L(a_i)^\perp$, $i = 1, 2$, zadanie redukuje się do podprzestrzeni $(L(a_1)^\perp \cap L(a_2)^\perp)^\perp = L(a_1, a_2)$, z którą każda z podprzestrzeni $L(a_i)^\perp$ ma tylko jeden kierunek wspólny. Znaleźć wektory x_i takie, że $L(a_1, a_2) \cap L(a_i)^\perp = L(x_i)$ i wyliczyć cosinus kąta między nimi. Jeśli płaszczyzny są równoległe, to kąt jest zerowy. W przypadku płaszczyzny $P + L(a)^\perp$ i prostej $Q + L(b)$ takich, że a, b są liniowo niezależne i nie są prostopadłe mamy $L(a)^\perp \cap L(b) = \{0\}$. Wybrać bazę ortonormalną e_1, e_2 podprzestrzeni $L(a)^\perp$ tak, aby $b \in L(e_1, a)$. Wykazać, że wektor $b \cdot e_1 e_1$ tworzy minimalny kąt z wektorem b. Przypadki liniowo zależnych lub prostopadłych a, b są oczywiste.

255 (a) Wykazać, że kąt między prostą $P + L(x)$ a podprzestrzenią $Q + W$ jest równy kątowi $\measuredangle(x, x_\parallel)$, gdzie $x = x_\parallel + x_\perp$ jest rozkładem na część należącą do W i część prostopadłą do W odpowiednio. Wykazać stąd dalej, że każda podprzestrzeń tworząca z prostą $P + L(x)$ kąt α może być przedstawiona w postaci $Q + [L(y) \oplus W']$, gdzie y jest dowolnym wektorem tworzącym kąt α z x, a W' jest dowolną podprzestrzenią przestrzeni $L(x, y)^\perp$. (b) $P + L(\overrightarrow{PQ}, x \times \overrightarrow{PQ})$ (wykorzystać wynik (a)).

256

(a) $\overrightarrow{OX} \cdot (\mu a + \nu b) = \mu\alpha + \nu\beta\,,\ (\mu,\nu) \in \mathbb{R}^2 \setminus (0,0)\,,$

(b) $\overrightarrow{OX} \cdot \left[\mu\left(\frac{a}{\|a\|} \pm \frac{b}{\|b\|}\right) + \nu a \times b\right] = \rho\,,\ (\mu,\nu) \in \mathbb{R}^2 \setminus (0,0)\,,\ \rho \in \mathbb{R}\,,$

(c) $\overrightarrow{OX} \cdot \left[\frac{a}{\|a\|} \pm \frac{b}{\|b\|}\right] = \left[\frac{\alpha}{\|a\|} \pm \frac{\beta}{\|b\|}\right]\,.$

258 Jeśli $(\tilde{a}_i)$ jest bazą dualną do (a_i) (w sensie zadania 229), to wektor wodzący punktu przecięcia ma postać $\sum_{i=1}^{3} \mu_i \tilde{a}_i = -2\hat{\imath} + \hat{\jmath} - \hat{k}$, a odległość wynosi $\frac{|\mu+\mu_2+\mu_3-\nu|}{\|a_1+a_2+a_3\|} = \frac{5\sqrt{2}}{6}$.

259 Tetraedr jest oparty na wektorach $(\nu - \mu_1 - \mu_2 - \mu_3)\tilde{a}_i$, jego objętość wynosi $\frac{1}{3!}\left|\frac{(\nu-\mu_1-\mu_2-\mu_3)^3}{a_1\cdot(a_2\times a_3)}\right| = \frac{125}{24}$.

260 Równanie szukanej prostej: $\overrightarrow{OY} \times b = l$, gdzie $b = a \times (k - \overrightarrow{OP} \times a) = -3(4\hat{\jmath} + 3\hat{k})$, $\quad l = \overrightarrow{OP} \times b = 3(11\hat{\imath} + 3\hat{\jmath} - 4\hat{k})$; jej odległość od punktu O: $\|\overrightarrow{OP}_\perp\| = \frac{\|l\|}{\|b\|} = \frac{\sqrt{146}}{5}$.

261 Osiem płaszczyzn $O + d_1 + L(d_2 - d_1, d_3 - d_1)$, gdzie $d_i = \pm d n_i$ (znaki nieskorelowane). Odległość wynosi $\|\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_3\|^{-1}$, gdzie $(\tilde{d}_i)$ jest bazą dualną do (d_i).

262 (a) $\frac{1}{10}\begin{pmatrix} 34 \\ -19 \\ 2 \end{pmatrix}$. (b) $\frac{1}{47}\begin{pmatrix} 118 \\ 178 \\ -211 \end{pmatrix}$.

§23 Odwzorowania afiniczne

263 Należy pokazać, że odwzorowanie określone jako $f(\sum \lambda_i O_i) = \sum \lambda_i Z_i$ jest odwzorowaniem afinicznym. Wykorzystać jedno z kryteriów afiniczności oraz wynik zadania 235.

266 Podane złożenie jest translacją t_{2a}.

267 (a) Przy tak zmienionym sformułowaniu twierdzenia dowód pozostaje prawie niezmieniony. (b) Z założeń wynika, że $\mathrm{Ker}(A - \mathrm{id}) = [\mathrm{Im}(A - \mathrm{id})]^\perp$ oraz podprzestrzeń $\mathrm{Im}(A - \mathrm{id})$ jest niezdegenerowana (wykorzystać wyniki zadania 172 (a) oraz przykładu (v), p. 11, §19). Po tej obserwacji dowód jak dla twierdzenia 16, §23.

268 Wykorzystać wynik poprzedniego zadania; przy użytej notacji szukany zbiór to $O + \mathrm{Ker}(A - \mathrm{id})$, a minimalna wartość $\|\overrightarrow{Of(O)}\|$ jest równa $\|x\|$.

269 Niech $\mathcal{L}_i = O_i + L(a_i)$ i oznaczmy $a_3 = \overrightarrow{O_1O_2}$. Wtedy szukaną grupę tworzą odwzorowania o postaci $(-\mu_1 a_1, A)_{O_1}$, $Aa_1 = \lambda_1 a_1$, $Aa_2 = \lambda_2 a_2$, $\lambda_i \in \mathbb{R}^*$, $Aa_3 = \mu_1 a_1 + \mu_2 a_2 + a_3$, $\mu_i \in \mathbb{R}$. W przypadku izometrii w przestrzeni euklidesowej punkty O_i można wybrać tak, aby $a_3 \perp a_1, a_2$. Wówczas $\mu_i = 0$, $\lambda_i = \pm 1$, przy czym w ostatniej równości znaki są skorelowane i grupa składa się z dwóch elementów gdy $a_1 \cdot a_2 \neq 0$, a w przeciwnym przypadku znaki są nieskorelowane i grupa składa się z czterech elementów.

270 (a) Jeśli R jest odbiciem kierunku wektora x, to $t_x = (0, R)_O (0, R)_{O-\frac{x}{2}}$.
(b) Proste zastosowanie analogicznego wyniku dla izometrii liniowych oraz punktu (a).

271 Wykorzystać postać kanoniczną izometrii, z uwzględnieniem możliwych postaci kanonicznych jej części liniowej.

272 Wykazać, że operatorem ortogonalnego rzutu na podprzestrzeń prostopadłą do wektora własnego operatora A jest $P = [\mathrm{Tr}(\mathrm{id} - A)]^{-1}(\mathrm{id} - A^*)(\mathrm{id} - A)$. Stąd postać kanoniczna:
$(y, A)_X = (y - Py, A)_O$, gdzie $O \in X + \frac{(\mathrm{id} - A^*)y}{\mathrm{Tr}(\mathrm{id} - A)} + \mathrm{Ker}(A - \mathrm{id})$.

273 (b) $t = t' + b$, $y = y' + t'v + a$.

§24 Dodatkowe konstrukcje algebraiczne w przestrzeniach wektorowych

275 Wykorzystać wyniki zadań 212 i 218; część argumentacji podobna jak w poprzednim zadaniu.

276 $\mathcal{L}(V, V; V) \mapsto \mathcal{L}(V \otimes V, V) \mapsto (V \otimes V)^* \otimes V \mapsto V^* \otimes V^* \otimes V$

277 Szukany tensor to $e^i{}_{jk} = g^{il} e_{ljk}$. Algebra nie ma jedynki.

280 (a) Wykorzystać wynik zadania 145. (c) Rozłożyć złożenia AA' i $A'A$ na część liniową i antyliniową.

§25 Postać kanoniczna operatora liniowego

282 Istnienie rozkładu jest wnioskiem z twierdzenia 2. Jeśli dany jest rozkład, o którym mówi zadanie, to zawężenie operatora do podprzestrzeni własnych operatora D daje rozkład z twierdzenia; stąd jednoznaczność.

283
$$\begin{pmatrix} 0 & -i & 0 & -i & 2i & 1 \\ 1 & 0 & -i & 0 & 0 & 0 \\ -i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2i & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1+i & 1 \\ 0 & 0 & 0 & 0 & 0 & 1+i \end{pmatrix}.$$

284
$$\begin{pmatrix} 0 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & -3 & 1 & 0 \\ 3 & 2 & 0 & 1 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 3 & 2 \end{pmatrix}.$$

§26 Formy bi-afiniczne. Kwadryki

285 Drugi przypadek zachodzi tylko dla typu kanonicznego $(i)_1$ z $p = 1$ (tw. 7, §26).

288 Ponieważ $\mathrm{Ker}\, g = \{0\}$, to istnieje dokładnie jeden punkt O taki, że $h(X, X) = h(O, O) + g(\overrightarrow{OX}, \overrightarrow{OX})$. Wtedy jedynym rozwiązaniem warunków zadania jest $f = (0, A)_O$, $f^* = (0, A^*)_O$, gdzie A^* jest operatorem sprzężonym do A względem iloczynu g.

289 (a) Dane są proste $P+L(a)$ i $P'+L(a')$, przy czym można założyć, że P i P' są punktami na tych prostych najbliżej położonymi, oraz $\|a\|=\|a'\|=1$. Wybierzmy punkt odniesienia $O=\frac{1}{2}P+\frac{1}{2}P'$. Szukane punkty X wyznaczone są równaniem $2\overrightarrow{PP'}\cdot\overrightarrow{OX}=(a\cdot\overrightarrow{OX})^2-(a'\cdot\overrightarrow{OX})^2$. (i) Jeśli proste przecinają się, to $\overrightarrow{PP'}=0$ i wtedy rozwiązaniem są dwie płaszczyzny $O+L(a\pm a')^\perp$. (ii) Jeśli proste są równoległe, to rozwiązaniem jest płaszczyzna $O+L(\overrightarrow{PP'})^\perp$. (iii) Jeśli proste nie przecinają się i nie są równoległe, to wybierzmy znak a' tak, aby baza $(a,a',\overrightarrow{PP'})$ była dodatnio zorientowana. Wtedy wektory
$e_1=\frac{a+a'}{2\sqrt{1+a\cdot a'}}+\frac{a-a'}{2\sqrt{1-a\cdot a'}}$, $e_2=\frac{a+a'}{2\sqrt{1+a\cdot a'}}-\frac{a-a'}{2\sqrt{1-a\cdot a'}}$, $e_3=\frac{\overrightarrow{PP'}}{\|\overrightarrow{PP'}\|}$
tworzą bazę ortonormalną dodatnio zorientowaną. W układzie $(O,(e_i))$ kwadryka przyjmuje postać kanoniczną $X^3=\frac{\sqrt{1-(a\cdot a')^2}}{2\|\overrightarrow{PP'}\|}\big[(X^1)^2-(X^2)^2\big]$.
(b) Dana jest prosta $P+L(a)$ i płaszczyzna $Q+L(b)^\perp$, przy czym można założyć, że P i Q są punktami najbliżej położonymi, oraz $\|a\|=\|b\|=1$. Szukane punkty wyznaczone są równaniem $(\overrightarrow{PX})^2-(a\cdot\overrightarrow{PX})^2-(b\cdot\overrightarrow{QX})^2=0$. Jeśli wybrać punkt odniesienia $O=\frac{1}{2}P+\frac{1}{2}Q$, to przyjmuje ono postać $2\overrightarrow{PQ}\cdot\overrightarrow{OX}+(\overrightarrow{OX})^2-(a\cdot\overrightarrow{OX})^2-(b\cdot\overrightarrow{OX})^2=0$. (i) Jeśli prosta nie przecina płaszczyzny, to wybierzmy bazę ortonormalną: $e_2=a$, $e_3=-\overrightarrow{PQ}/\|\overrightarrow{PQ}\|$, $e_1=e_2\times e_3$. W układzie $(O,(e_i))$ kwadryka ma postać kanoniczną $X^3=(X^1)^2/(2\|\overrightarrow{PQ}\|)$. (ii) Jeśli prosta przecina płaszczyznę, to $P=Q=O$. Wybierzmy znak b tak, aby $a\cdot b>0$. Wektory $e_2=\frac{a-b}{\sqrt{2(1-a\cdot b)}}$, $e_3=\frac{a+b}{\sqrt{2(1+a\cdot b)}}$, $e_1=e_2\times e_3$ tworzą bazę ortonormalną (jeśli $a=b$, to e_2 wybieramy dowolnie spośród wektorów prostopadłych do e_3; podobnie dla $a=-b$). Równanie uzyskuje postać kanoniczną $(X^1)^2+a\cdot b\,(X^2)^2-a\cdot b\,(X^3)^2=0$.

LITERATURA

Ogólne wprowadzenie do algebry:

- Z. Opial, *Algebra wyższa* (wyd. 9), Państwowe Wydawnictwo Naukowe, Warszawa 1976;
- A. I. Kostrykin, *Wstęp do algebry*, t. 1–3, Państwowe Wydawnictwo Naukowe, Warszawa 2004-2005;
- A. Mostowski, M. Stark, *Elementy algebry wyższej* (wyd. 9), Państwowe Wydawnictwo Naukowe, Warszawa 1977;
- S. Lang, *Algebra*, Państwowe Wydawnictwo Naukowe, Warszawa 1984.

Algebra liniowa i geometria:

- A. I. Kostrykin, J. I. Manin, *Algebra liniowa i geometria*, Państwowe Wydawnictwo Naukowe, Warszawa 1993,
- I. M. Gelfand, *Wykłady z algebry liniowej* (wyd. 3), Państwowe Wydawnictwo Naukowe, Warszawa 1977;
- J. Gancarzewicz, *Algebra liniowa i jej zastosowania* (wyd. 2), Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2009;
- A. Białynicki-Birula, *Algebra liniowa z geometrią*, (wyd. 2), Państwowe Wydawnictwo Naukowe, Warszawa 1979;
- A. Staruszkiewicz, *Wykłady dla fizyków*, Naukowe Koło Fizyków UJ, Kraków 1993;
- J. Komorowski, *Od liczb zespolonych do tensorów, spinorów, algebr Liego i kwadryk*, Państwowe Wydawnictwo Naukowe, Warszawa 1978;

- S. H. Friedberg, A. J. Insel, L. E. Spence, *Linear Algebra*, Prentice Hall, Upper Saddle River 1997;
- S. Lang, *Linear algebra*, Addison-Wesley, Reading 1972.

Zbiory zadań:

- I. W. Proskuriakov, *Sbornik zadacz po liniejnoj algebrie*, Izdatielstwo Nauka, Moskwa 1984;
- *Zbiór zadań z algebry: praca zbiorowa*, red. A. I. Kostrykin, Państwowe Wydawnictwo Naukowe, Warszawa 1995;
- H. Arodź, K. Rościszewski, *Algebra i geometria analityczna w zadaniach*, Znak, Kraków 2005;
- D. Faddeev, I. Sominsky, *Problems in higher algebra*, Mir Publishers, Moscow.

W niniejszym podręczniku zabrakło miejsca na wprowadzenie w przedmiot spinorów. Dla fizyków szczególnie godnym polecenia wprowadzeniem w tę tematykę są początkowe rozdziały pierwszej pozycji:

- R. Penrose, W. Rindler, *Spinors and Spacetime*, vol. 1, Cambridge University Press, Cambridge 1986;
- J. Łopuszański, *Rachunek spinorów*, Państwowe Wydawnictwo Naukowe, Warszawa 1985.

INDEKS