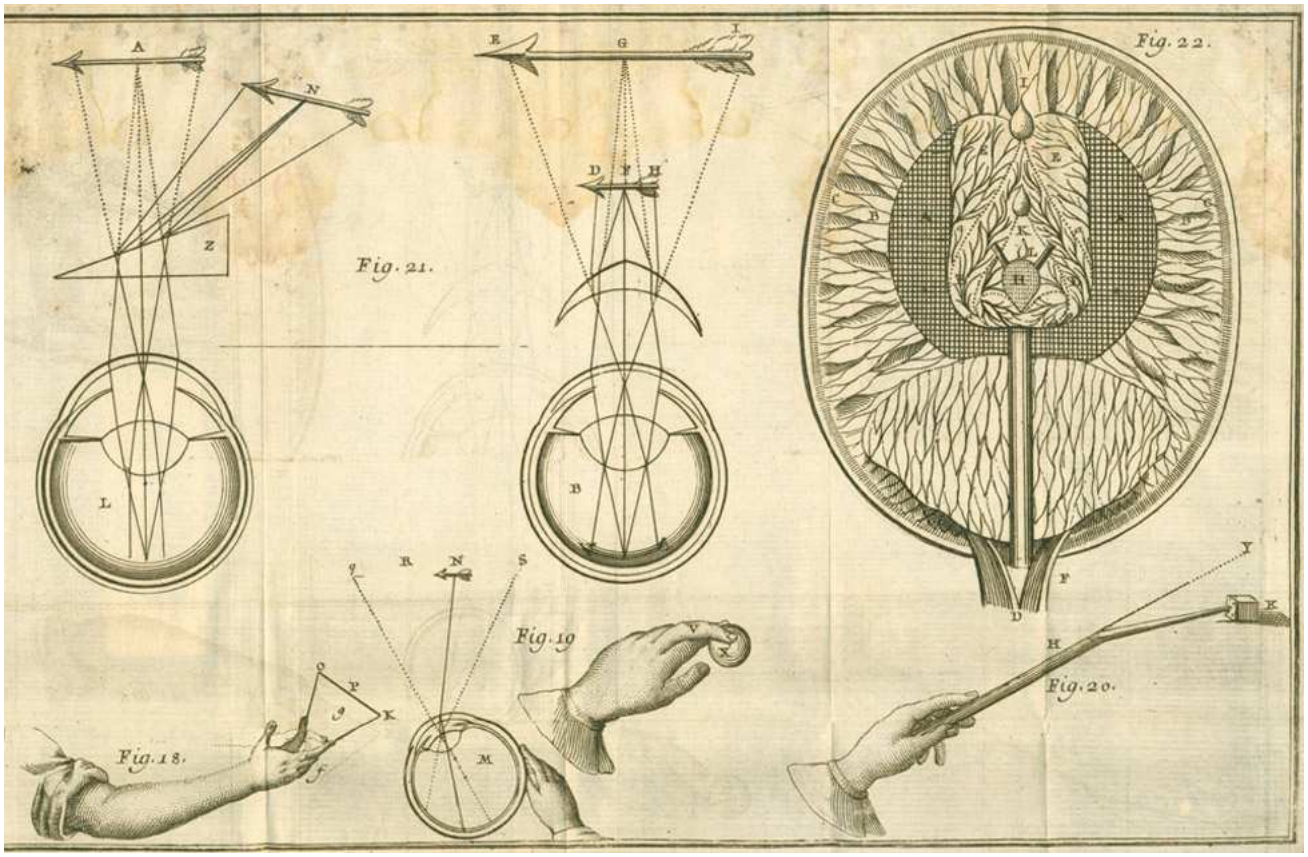
GEOMETRIA Z ALGEBRĄ LINIOWĄ I

(notatki do wykładu i opowiadania o matematyce)

Arkadiusz Męcel



Kartezjusz, *La Geometrie*, 1637, źródło: `https://www.gutenberg.org/ebooks/26400`

Materiał realizowany według skryptu *Wykłady z Algebry Liniowej* doc. dr. Tadeusza Koźniewskiego
`https://mimuw.edu.pl/~amecel/alglin.html`
ostatnia aktualizacja: 19 stycznia 2024

Spis treści

Notatki nie są skryptem. Czym są?

Poniższe notatki pochodzą z mojego wykładu z Geometrii z Algebrą Liniową na kierunku matematyka prowadzonym przez Wydział Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego. Wykład jest dwusemestralny i ma dwie części. Pierwsza część złożona z piętnastu cotygodniowych wykładów w trakcie pierwszego semestru ma charakter wstępny – omawiane są układy równań liniowych o współczynnikach w ciele, liczby zespolone, przestrzenie i przekształcenia liniowe, funkcjonały i wyznacznik.

Notatki podzielone są na części odpowiadające 15 tygodniom zajęć. Każda część podzielona jest na moduły, ale nie każda z części złożona jest ze wszystkich wymienionych niżej modułów. Na początku każdej części Czytelnik znajdzie datę ostatniej aktualizacji. Oto te moduły.

- **Wykład** – treść zasadnicza „czerpiąca bezpośrednio" (miejscami całe fragmenty) ze skryptu wydziałowego dra T. Koźniewskiego[1], rozszerzająca te treści o intuicje, przykłady i rozmaite zadania.
- **Zadania do samodzielnej pracy** – zestaw zadań ilustrujących umiejętności wymagane zarówno do zaliczenia przedmiotu na ocenę dostateczną (oznaczonych ♠), jak i do uzyskania biegłości w prowadzeniu bardziej złożonych rozumowań, koniecznej do uzyskania oceny bardzo dobrej.
- **Uzupełnienie** – zawiera treści ściśle związane z tematem – takie, które chętnie dodałbym do wykładu, gdybym tylko miał czas. Poziomem trudności treści te nie odbiegają od wykładu. Często omawiane są przykłady, na które w bardzo skondensowanym harmonogramie nie ma miejsca.
- **Dodatek** – poruszam tu tematy niekiedy bardziej zaawansowane, o charakterze „gwiazdkowym", odnoszące się do ważnych motywów pojawiających się na konkretnym wykładzie. Zagadnienia te mają charakter algebraiczny, ale nie zawsze czysto algebroliniowy (a na przykład: kombinatoryczny).
- **Trivia** – tematy luźno związane z wykładem, ale moim zdaniem ładne i mające charakter popularyzatorski. Czasem mogą nawet dotyczyć zastosowań. Oglądanie takich rzeczy pomaga niektórym podtrzymać wiarę w to, że studiowanie „czystej" matematyki ma głęboki estetyczny sens.
- **Coda** – fragment ten wskazywać będzie szerszy kontekst pojęciowy i historyczny pojęć pojawiających się na wykładzie. Na wykładzie sygnalizujemy wiele fundamentalnych koncepcji, które poznawać będą Państwo przez całe studia. Będę się starał krótko wskazywać na osoby i wydarzenia związane z rozwojem danego pojęcia. Podstawowym materiałem będzie piękna książka Johna Stillwella *Mathematics and Its History*, wydanej w 2010 roku przez Springera.

Ważne zastrzeżenie – **notatki (zwłaszcza te notatki) to nie jest skrypt**, czyli tekst realizujący obowiązujący sylabus. Skrypt zawiera to, co kluczowe dla uczestnika kursu – treści, z których będzie egzaminowany. Poniższe notatki zawierają także to, co mi w uczeniu tego kursu pomaga i co mi się w nim podoba. Obok ścisłych rozumowań ważne są dla mnie intuicje i motywacje. Ten tekst próbuje je zebrać i wciąż się zmienia – gdy coś mi nie wychodzi, gdy próbuję coś poprawić, gdy dostaję uwagi, gdy czytam innych autorów, gdy coś nowego mi się spodoba. Chciałbym, aby to był żywy, „mówiony" tekst. Zapraszam Czytelnika do krytycznej lektury oraz do wskazywania mi znalezionych błędów. Dziękuję za już wskazane i przepraszam, że pozwalam sobie nie wymieniać wszystkich osób, które je znajdowały.

I jeszcze jedno — **notatki nie są książką kucharską**. Chociaż staram się dać Czytelnikowi przegląd ciekawych typów zadań, z pewnością nie chodzi o to, by w notatkach znalazły się wszystkie możliwe typy. Studiowanie matematyki jest, trzymając się wspomnianej analogii, niczym pójście do szkoły kulinarnej. Nie tylko uczymy się przepisów na wytrawne dania, ale pytamy: dlaczego te przepisy działają? Jak powstają? Jak łączyć smaki, jak dobierać składniki, skąd je brać? Jednym słowem — chcemy wykształcić samodzielnych matematyków — otwartych, dociekliwych, twórczych i pewnych swoich umiejętności.

[1] T. Koźniewski, *Wykłady z algebry liniowej I*, Uniwersytet Warszawski, Warszawa 2008.

Algebra liniowa z geometrią?

> Wykład GAL I poświęcony jest rozwiązywaniu układów równań liniowych nad ciałem. Omówiona będzie definicja ciała i jego podstawowe własności. Skupimy się na badaniu ciała liczb rzeczywistych i zespolonych. Pokażemy jak opisywać zbiory rozwiązań wprowadzając na nich strukturę przestrzeni liniowej. Podstawowym narzędziem będą macierze opisujące zarówno układy równań jak i przekształcenia liniowe.

Powyższe streszczenie wykładu zaczerpnięte z wydziałowego sylabusa należy uzupełnić bardziej poglądowym wyjaśnieniem nawiązującym do nazwy przedmiotu, który zamierzamy studiować. Co rozumieć będziemy przez algebrę, a co przez geometrię? Co oznacza pojęcie algebry „liniowej"? Czy odnosić się będziemy do treści szkolnych? Na czym polegać będzie nowość ujęcia, które zaprezentujemy? W jaki sposób przedmiot ów wprowadzi nas w studiowanie matematyki? Pełne odpowiedzi na te pytania przekraczają ramy tego wprowadzenia. Możemy jednak poczynić kilka pożytecznych, miejmy nadzieję, uwag.

Z perspektywy absolwenta szkoły przedmiot nasz wydawać się może uogólnieniem geometrii analitycznej (pod podobną nazwą wykładany był zresztą – przy nieco innym rozkładzie akcentów – w latach powojennych). Będziemy mówić o obiektach geometrycznych opisywanych przy pomocy warunków algebraicznych. Będą to początkowo rozwiązania układów równań liniowych. Weźmy, nawiązując do wiedzy szkolnej, problem określenia wzajemnego położenia dwóch prostych leżących na płaszczyźnie kartezjańskiej, na podstawie opisujących ich równań (np. w postaci ogólnej). Trzy możliwe konfiguracje to:

- para prostych przecinających się w dokładnie jednym punkcie,
- para prostych równoległych,
- para prostych tożsamych (ta sama prosta może być opisana przez różne równania!).

Sytuacjom tym odpowiada interpretacja algebraiczna:

- układ równań ułożony z równań opisujących poszczególne proste ma dokładnie jedno rozwiązanie,
- układ równań ułożony z równań opisujących poszczególne proste nie ma rozwiązań,
- układ równań ułożony z równań opisujących poszczególne proste ma nieskończenie wiele rozwiązań.

Podczas zajęć szybko zrozumiemy, że przedstawiona wyżej odpowiedniość jest w istocie swego rodzaju utożsamieniem rzeczywistości algebraicznej i geometrycznej. Wspólną płaszczyzną pojęciową dla tego utożsamienia będzie pojęcie przestrzeni liniowej. Samo zaś utożsamienie dokonane zostanie przy pomocy aparatu pojęciowego związanego z operacjami na przestrzeniach liniowych. Co więcej, przestrzenie liniowe są pojęciem tak ogólnym i wygodnym w użyciu, że pozwolą na dostrzeganie innych, mniej oczywistych powiązań pomiędzy różnymi działami matematyki. Ciekawe co Czytelnik powiedziałby słysząc, że w odpowiednio dobranym kontekście funkcja sinus może być prostopadła do funkcji cosinus (a nawet i do samej siebie!) albo, że liniowo niezależne – cokolwiek to znaczy – są przy odpowiednich definicjach dwa podzbiory jego znajomych. Takie algebraiczne konteksty pojawią się np. w analizie i kombinatoryce.

Innymi słowy, chodzić nam będzie nie tylko o rozwiązywanie układów równań, ale o umiejętność wyabstrahowania „struktury" tych rozwiązań i dostrzegania jej w rozmaitych matematycznych sytuacjach. Tym, co charakteryzuje naukowe podejście do uprawiania matematyki jest w tym kontekście nie tylko dążenie do znalezienia ogólnej struktury, ale też próba opisu „wszystkich możliwych struktur" z dokładnością do pewnych warunków. Przykład tego rodzaju opisu widzimy wyżej: układ dwóch prostych na płaszczyźnie prowadzić może tylko do trzech na swój sposób istotnie różnych konfiguracji. W języku algebry liniowej niezmiennikiem takich „istotnie różnych" konfiguracji będzie m.in. pojęcie wymiaru.

Podejście, o którym tu piszemy nazwać można podejściem klasycznym. Badamy grupę obiektów i próbujemy znaleźć cechy charakterystyczne, które by je odróżniały (na przykład odróżniliśmy rodziny par prostych na płaszczyźnie poprzez opis możliwej liczby ich punktów przecięcia). W kontekście geometrii analitycznej ojcem tego podejścia był R. Descartes (Kartezjusz), stąd też często (w szkole?) mówimy o kartezjańskim układzie współrzędnych. Podejście to bardzo rozwinęło badania geometryczne (do tej pory rozumiane głównie w stylu starożytnych „Elementów" Euklidesa – który to znany jest absolwentowi szkoły i nazywany jest mianem geometrii elementarnej) najpierw poprzez odkrycie rachunku różniczkowego i całkowego przez Newtona i Leibniza, które doprowadziło do powstania geometrii różniczkowej, potem przez prace dotyczące współrzędnych jednorodnych (tzw. współrzędne barycentryczne) poczynione na początku XIX wieku przez Möbiusa i Plückera, następnie przez rozwój geometrii rzutowej dzięki pracom Monge'a i Ponceleta, i wreszcie przez prace Łobaczewskiego i Bolyai, którzy „odkryli" geometrie nieeuklidesowe. Narodziny dziedziny, którą będziemy się zajmować przypisuje się różnym uczonym, zwłaszcza jednak nauczycielowi gimnazjalnemu H. Grassmannowi (prace od roku 1844). Był to czas rozwoju podejścia aksjomatycznego do geometrii, które w tej i innych dziedzinach zrewolucjonizowało w XX wieku Królową Nauk. Jest ono podstawą wykładu uniwersyteckiego matematyki i współczesnych badań.

Warto pamiętać, że przy całej abstrakcji jaka towarzyszy wykładowi geometrii z algebrą liniową, mowa jest wciąż o geometrii, a więc w najszerszym sensie w jakim rozumiemy to pojęcie – nauce o przestrzeni i jej własnościach. Zarówno wysiłki Kartezjusza, jak i Ponceleta, Łobaczewskiego, Riemanna, Kleina, Hilberta, Minkowskiego czy wielkich geometrów XX wieku miały w sobie zawsze silne pragnienie zrozumienia świata, w którym żyjemy i budowania przekonujących modeli matematycznych dla opisu jego działania. Zetkniemy się z tym trudnym zjawiskiem w drugim semestrze mówiąc o przestrzeniach afinicznych.

Prawdziwą rewolucją w myśleniu (bardzo charakterystyczną dla współczesnej matematyki) jest próba opisu obiektów nie w języku ich wewnętrznej struktury, ale w języku operacji, które na nich wykonujemy. Dla przykładu: kwadrat i okrąg odróżnić można w sposób „klasyczny" mówiąc dość nieprecyzyjnie (jakkolwiek byśmy nie lubili postulatów Euklidesa) o kątach czy wierzchołkach, lub bardziej formalnie: podając równania je opisujące (i tłumacząc kiedy równania opisują obiekty różne – i co to znaczy), ale można je także odróżnić bardzo elegancko poprzez obserwację, że tylko skończenie wiele izometrii płaszczyzny (na przykład obrotów) w siebie pozostawia kwadrat w miejscu, podczas gdy okrąg „jest niezmienniczy" (znowu ważne pojęcie) przy działaniu nieskończenie wielu różnych izometrii. To nowe podejście do matematyki polegające na badaniu obiektów przez „niezmienniki operacji" pochodzi, w kontekście geometrii układów równań i ich przekształceń od matematyków angielskich A. Cayleya, J.J. Sylvestera i G. Salmona. Do ustalenia głębszych związków pomiędzy teorią przekształceń a geometrią przyczynił się tzw. program z Erlangen słynnego matematyka niemieckiego Felixa Kleina sformułowany w 1872 roku.

Nasze podejście pójdzie jednak jeszcze dalej. W powojennej matematyce rewolucji dokonały prace warszawskiego matematyka Samuela Eilenberga i cały ruch nazwany później „New Math" reprezentowany przez środowisko francuskich badaczy, zwanych bourbakistami, którzy zaproponowali jeszcze ogólniejsze pojmowanie matematyki w języku diagramów. Teoria ta, zwaną teorią kategorii, jest obecnie językiem urzędowym całej algebraiczno–geometrycznej strony matematyki. Jednym z zaskakujących przejawów tego podejścia, które zobaczymy na naszym przedmiocie jedynie w zalążku, jest istnienie `naturalnych` odpowiedniości, które zamazują rozróżnienie pomiędzy obiektami, a działającymi na nich przekształceniami. W języku matematycznym odpowiedniości takie nazywamy dualnościami. Dla przykładu: obiektom geometrycznym przypiszemy ich przekształcenia, i odwrotnie. Na jednych i drugich wprowadzimy te same operacje. Okaże się, że matematycznie nie ma sensu ich rozróżnianie! W dostrzeżeniu tych dualności podstawowym narzędziem będą niepozorne na pierwszy rzut oka tablice liczbowe zwane macierzami.

Podsumowując ten przydługi wstęp: droga od matematyki szkolnej, zajmującej się głównie rozwiązywaniem konkretnych zadań dotyczących struktury danych obiektów (np. wzajemne położenie dwóch konkretnych prostych na płaszczyźnie) do współczesnej geometrii algebraicznej czy innych dziedzin matematyki prowadzi przez trzy progi: klasyfikowanie „wszystkich konfiguracji" obiektów w języku opartym na współrzędnych, klasyfikowanie obiektów będących „niezmiennikami operacji" bez użycia współrzędnych, ustalanie „dualności" pomiędzy obiektami i operacjami na nich, także w obrębie różnych działów matematyki (ważne hasło na przyszłość dla aspirującego matematyka: `odpowiedniość Galois`). Życzę Czytelnikowi udanego wejścia w piękną przygodę z matematyczną abstrakcją dzięki algebrze liniowej z geometrią.

Rozdział 1

Macierze i układy równań liniowych

1.1 Wykład pierwszy

Wykład z geometrii z algebrą liniową zaczynamy omówieniem teorii układów równań liniowych o współczynnikach w ciele. Czym jest ciało dowiemy się na kolejnym wykładzie. W tym momencie chcemy raczej uporządkować i usystematyzować język dotyczący samych układów równań i zbiorów ich rozwiązań. Zbiór liczb rzeczywistych oznaczamy przez $\mathbb{R}$, zaś zbiór liczb wymiernych – przez $\mathbb{Q}$.

Definicja 1.1: Równanie liniowe

Równanie liniowe o zmiennych $x_1, \ldots, x_n$ i o współczynnikach w zbiorze K, to wyrażenie postaci

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \text{ gdzie } a_1, a_2, \ldots, a_n, b \in K.$$

Rozwiązanie powyższego równania to ciąg $(s_1, s_2, \ldots, s_n)$, gdzie $s_1, s_2, \ldots, s_n \in K$ taki, że

$$a_1 \cdot s_1 + a_2 \cdot s_2 + \ldots + a_n \cdot s_n = b.$$

Przykłady

- Dla $K = \mathbb{R}$, równaniem liniowym zmiennej x jest $\sqrt{2}x = \pi$. Natomiast definicji powyższej nie spełniają równania $2x + 1 = 3$ lub $2x = x$.
- Dla $K = \mathbb{Q}$, równaniem liniowym o zmiennych x_1, x_2, x_3 jest: $\frac{1}{2}x_1 + x_2 + 0 \cdot x_3 = 0$, ale nie jest nim równanie $\sqrt{2}x_1 = 2$.

Aby powyższa definicja miała sens, przyjmujemy, że K to $\mathbb{R}$ lub $\mathbb{Q}$. Przyjmujemy także **konwencję**: zapisując równanie liniowe pomijamy składniki, w których zmienna przemnożona jest przez 0, np. równanie $2x_1 + 0 \cdot x_2 + x_3 = 0$ zapisujemy w postaci $2x_1 + x_3 = 0$, a równanie $0x_1 + 0x_2 = 1$ w postaci: $0 = 1$.

Definicja 1.2: Układ równań liniowych

Układ m równań liniowych o zmiennych $x_1, \ldots, x_n$ i o współczynnikach rzeczywistych, to ciąg m równań liniowych o współczynnikach rzeczywistych postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \qquad (\heartsuit)$$

Rozwiązanie powyższego układu równań liniowych to ciąg elementów $(s_1, \ldots, s_n)$ ze zbioru $\mathbb{R}$, który jest rozwiązaniem każdego z m równań liniowych tego układu.

Ważne jest zwrócenie uwagi na duży stopień dowolności, jaki daje ta definicja. Nic nie stoi na przeszkodzie by układ składał się z pojedynczego równania, albo zawierał identyczne równania. Oto kilka przykładów:

$$\begin{cases} 2x_1 + 3x_2 + x_3 &= 1 \\ -\frac{1}{2}x_2 + x_3 &= 0 \end{cases}, \qquad \begin{cases} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 1 \\ 0x_1 + 0x_2 &= 3 \end{cases}, \qquad \begin{cases} 0x_1 + 0x_2 + 0x_3 &= 0 \\ 0x_1 + 0x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{cases}.$$

Przykładem rozwiązania pierwszego z powyższych układów jest trójka $(-3, 2, 1)$. Drugi układ nie ma rozwiązań, ponieważ trzecie jego równanie nie ma rozwiązań. Trzeci układ ma choćby rozwiązanie $(0, 0, 0)$. Można sprawdzić, że dowolna trójka $(-s, s, 0)$, gdzie $s \in \mathbb{R}$, jest rozwiązaniem trzeciego układu.

Warto w tym miejscu poczynić uwagę terminologiczną o charakterze formalnym.

Definicja 1.3: Układ

Niech T, X będą dowolnymi zbiorami. Wówczas UKŁADEM ELEMENTÓW zbioru X o wskaźnikach przebiegających zbiór T będziemy nazywali dowolną funkcję $x : T \to X$ określoną na zbiorze T o wartościach w zbiorze X. Wartość tej funkcji w punkcie $t \in T$ nazwiemy ELEMENTEM tego układu o WSKAŹNIKU t, ozn. x_t.

Definicja powyższa jest formalnym wyrażeniem myśli zawartej wyżej. Gdy rozważamy układ m równań, wówczas zbiór T równy jest $\{1, 2, \ldots, m\}$, a X jest zbiorem m równań liniowych o n zmiennych. Wprowadzenie funkcji $x : T \to M$ daje nam to, że możemy mówić o pierwszym, drugim, trzecim, ..., m-tym równaniu tego układu. A to, że poszczególne równania mogą być zerowe, identyczne, bez rozwiązań — to jest inna kwestia. Na etapie konstruowania układu równań kwestie te nie mają znaczenia.

Definicja 1.4: Typy układów równań liniowych

Układ równań liniowych postaci (♡) nazywamy:

- JEDNORODNYM, jeśli $b_i = 0$, dla każdego $i = 1, 2, \ldots, m$,
- NIESPRZECZNYM, jeśli zbiór rozwiązań układu (♡) jest niepusty,
- SPRZECZNYM, jeśli zbiór rozwiązań układu (♡) jest pusty.

Układ jednorodny powstały z układu (♡) przez przyjęcie $b_i = 0$, dla każdego $i = 1, 2, \ldots, m$, nazywamy UKŁADEM JEDNORODNYM ODPOWIADAJĄCYM układowi (♡).

Wśród powyższych trzech przykładów, jednorodny jest trzeci układ. Niesprzeczne są — pierwszy i trzeci, a sprzecznym jest drugi z wymienionych układów. Układem jednorodnym odpowiadającym pierwszemu z powyższych trzech układów jest

$$\begin{cases} 2x_1 + 3x_2 + x_3 &= 0 \\ -\frac{1}{2}x_2 + x_3 &= 0 \end{cases}.$$

Definicja 1.5: Układy równoważne

Układy równań liniowych o tym samym zbiorze zmiennych i o tym samym zbiorze współczynników nazwiemy RÓWNOWAŻNYMI, jeśli mają one identyczne zbiory rozwiązań.

Przykład. Poniższe układy równań liniowych o zmiennych x_1, x_2 o współczynnikach w $\mathbb{R}$ są równoważne:

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 2 \end{cases}, \qquad \begin{cases} x_1 = 1 \\ x_2 = -1 \end{cases}.$$

Naszym celem jest opis rozwiązań układów równań liniowych o współczynnikach rzeczywistych. Metoda polega na zastępowaniu układu innym – równoważnym, i z jakiegoś powodu prostszym do rozwiązania.

Definicja 1.6: Operacje na równaniach liniowych

Dane sa równania liniowe U, U' o zmiennych $x_1, \ldots, x_n$ i o współczynnikach w $\mathbb{R}$:

$$U: \quad a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \qquad U': \quad a'_1x_1 + a'_2x_2 + \ldots + a'_nx_n = b'.$$

Określamy równanie liniowe $U + U'$ oraz dla każdego $\lambda \in \mathbb{R}$ określamy równanie liniowe λU:

$$\begin{aligned} U + U': &\quad (a_1 + a'_1)x_1 + (a_2 + a'_2)x_2 + \ldots + (a_n + a'_n)x_n = b + b', \\ \lambda U: &\quad \lambda \cdot a_1x_1 + \lambda \cdot a_2x_2 + \ldots + \lambda \cdot a_nx_n = \lambda \cdot b. \end{aligned}$$

Przykład. Jeśli U jest równaniem $x_1 + x_2 - x_3 = 2$ oraz U' jest równaniem $3x_1 - x_2 + x_3 = 0$, to równanie $U + U'$ ma postać $4x_1 = 2$, a równanie $2U$ ma postać $2x_1 + 2x_2 - 2x_3 = 4$.

Obserwacja 1.1

Jeśli $(s_1, \ldots, s_n)$ jest rozwiązaniem obydwu równań liniowych U_1 oraz U_2, gdzie $s_1, \ldots, s_n \in \mathbb{R}$, to jest też rozwiązaniem każdego równania liniowego postaci $\lambda_1U_1 + \lambda_2U_2$, gdzie $\lambda_1, \lambda_2 \in \mathbb{R}$.

Dowód. Niech $U_1 : a_1x_1 + \ldots + a_nx_n = b$ oraz $U_2 : a'_1x_1 + \ldots + a'_nx_n = b'$. Mamy wówczas

$$(\lambda_1a_1 + \lambda_2a'_1)s_1 + \ldots + (\lambda_1a_n + \lambda_2a'_n)s_n = \lambda_1b + \lambda_2b'.$$

W szczególności $(s_1, \ldots, s_n)$ jest rozwiązaniem równania $\lambda_1U_1 + \lambda_2U_2$. □

Twierdzenie 1.1

Poniższe operacje zamieniają układ równań liniowych U o współczynnikach rzeczywistych w układ równoważny:

(1) dodanie do równania innego równania pomnożonego przez liczbę rzeczywistą,

(2) zamiana dwóch równań miejscami,

(3) pomnożenie równania przez liczbę rzeczywistą różną od zera.

Dowód. Załóżmy, że układ U ma m równań. Przez U_i oznaczamy i-te równanie układu U, dla $1 \leqslant i \leqslant m$.

Zaczniemy od dowodu dla operacji (3). Niech U' powstaje z U przez przemnożenie równania U_i przez $\lambda \neq 0$. Twierdzimy, że każde rozwiązanie $(s_1, \ldots, s_n)$ układu U jest też rozwiązaniem układu U'. Rzeczywiście, j-te równanie U' jest dla $j \neq i$, j-tym równaniem układu U, więc $(s_1, \ldots, s_n)$ jest jego rozwiązaniem. Natomiast i-te równanie U' to równanie λU_i. Zgodnie z Obserwacją 1.1, skoro $(s_1, \ldots, s_n)$ jest rozwiązaniem U_i, to jest też rozwiązaniem λU_i. A zatem $(s_1, \ldots, s_n)$ jest rozwiązaniem U'.

Z drugiej strony, każde rozwiązanie $(r_1, \ldots, r_n)$ układu U' jest rozwiązaniem układu U. Rzeczywiście, za wyjątkiem i-tego równania, wszystkie równania U to równania U', więc $(r_1, \ldots, r_n)$ jest ich rozwiązaniem. Natomiast i-te równanie układu U powstaje z i-tego równania układu U' przez przemnożenie go przez λ^{-1}. A zatem zgodnie z Obserwacją 1.1 $(r_1, \ldots, r_n)$ jest rozwiązaniem tego równania. A zatem $(r_1, \ldots, r_n)$ jest rozwiązaniem układu U. Pokazaliśmy, że każde rozwiązanie U jest rozwiązaniem U' i odwrotnie – każde rozwiązanie U' jest rozwiązaniem U. A zatem układy te są równoważne.

Wykazaliśmy, że operacja (3) przeprowadza układ równań liniowych w układ równoważny. Teza dla operacji (2) jest oczywista. Pozostała analiza operacji (1). Niech układ U' powstaje z U przez dodanie do i-tego równania U_i równania U_j przemnożonego przez $a \in \mathbb{R}$. Wtedy U powstaje z U' przez dodanie do i-tego równania $U_i + aU_j$ równania U_j przemnożonego przez $-a \in \mathbb{R}$. Na mocy Obserwacji 1.1 zbiory rozwiązań układów U oraz U' są zatem identyczne. □

Definicja 1.7: Operacje elementarne na układzie równań liniowych

Operacje (1)-(3) nazywać będziemy OPERACJAMI ELEMENTARNYMI NA UKŁADZIE U.

Do jakich postaci chcemy dochodzić w ramach takiego „upraszczania" przy pomocy operacji elementarnych? Do takich, gdzie jedne zmienne można wyrazić za pomocą innych. Tak jest w przypadku układu po prawej stronie — równoważnego, jak się okazuje, układowi po lewej.

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 - x_2 - x_3 - x_4 &= 0 \\ x_1 - x_2 - x_3 - x_4 &= 0 \end{cases}, \qquad \begin{cases} x_1 &= \frac{1}{2} \\ x_2 + x_3 + x_4 &= \frac{1}{2} \\ 0 &= 0 \end{cases}.$$

Oto operacje elementarne przeprowadzające pierwszy układ w drugi: dodanie drugiego równania do pierwszego równania, odjęcie od trzeciego równania drugiego równania, przemnożenie pierwszego równania przez $\frac{1}{2}$, odjęcie od drugiego równania pierwszego równania, przemnożenie drugiego równania przez -1:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 - x_2 - x_3 - x_4 &= 0 \\ x_1 - x_2 - x_3 - x_4 &= 0 \end{cases} \longrightarrow \begin{cases} 2x_1 &= 1 \\ x_1 - x_2 - x_3 - x_4 &= 0 \\ x_1 - x_2 - x_3 - x_4 &= 0 \end{cases}$$

$$\longrightarrow \begin{cases} 2x_1 &= 1 \\ x_1 - x_2 - x_3 - x_4 &= 0 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0 \end{cases}$$

$$\longrightarrow \begin{cases} x_1 &= \frac{1}{2} \\ x_1 - x_2 - x_3 - x_4 &= 0 \\ 0 &= 0 \end{cases}$$

$$\longrightarrow \begin{cases} x_1 &= \frac{1}{2} \\ -x_2 - x_3 - x_4 &= -\frac{1}{2} \\ 0 &= 0 \end{cases}$$

$$\longrightarrow \begin{cases} x_1 &= \frac{1}{2} \\ x_2 + x_3 + x_4 &= \frac{1}{2} \\ 0 &= 0 \end{cases}$$

Zbiór rozwiązań powyższych układów składa się z ciągów postaci $\left(\frac{1}{2}, -s_3 - s_4 + \frac{1}{2}, s_3, s_4\right)$, gdzie s_3, s_4 są dowolnymi liczbami rzeczywistymi. Bez trudu odczytaliśmy to rozwiązanie z układu uzyskanego poprzez operacje elementarne. Wyraziliśmy po prostu zmienne x_1, x_2 za pomocą stałych oraz zmiennych x_3, x_4.

Definicja 1.8: Rozwiązanie ogólne układu równań liniowych

Niech U oraz U' będą równoważnymi układami równań liniowych o n zmiennych (niewiadomych). Przypuśćmy, że układ U' MOŻNA PRZEPISAĆ w postaci:

$$\begin{cases} x_{j_1} &= c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n + d_1 \\ \vdots & \\ x_{j_k} &= c_{k1}x_1 + c_{k2}x_2 + \ldots + c_{kn}x_n + d_k \end{cases}$$

przy czym $1 \leqslant j_1 < \ldots < j_k \leqslant n$ oraz zmienne $x_{j_1}, \ldots, x_{j_k}$ nie występują po prawej stronie (to znaczy: stoją przy nich współczynniki zerowe, czyli $c_{ij} = 0$, dla $j = j_1, \ldots, j_k$). Mówimy też, że:

- układ U' JEST ROZWIĄZANIEM OGÓLNYM (zadaje rozwiązanie ogólne) układu U,
- w rozwiązaniu tym $x_{j_1}, \ldots, x_{j_k}$ nazywamy ZMIENNYMI ZALEŻNYMI,
- pozostałe x_i nazywamy ZMIENNYMI NIEZALEŻNYMI, albo PARAMETRAMI.

Potrzebna jest chwila uwagi, by zrozumieć powyższą definicję. Zdefiniowaliśmy te postaci układów równań, z których łatwo jest odczytywać rozwiązania, a które uzyskiwać chcemy z dowolnych układów poprzez stosowanie operacji elementarnych. Oto przykłady układów, będących rozwiązaniami ogólnymi.

$$\begin{cases} x_1 = 2 \\ x_2 = 1 \\ x_3 = 4 \end{cases}, \quad \begin{cases} x_1 = \frac{1}{2} \\ x_2 + x_3 + x_4 = \frac{1}{2} \end{cases}, \quad \begin{cases} x_1 + x_3 - 2x_5 = 0 \\ x_2 + x_5 = 0 \\ x_4 + x_5 = 0 \end{cases}.$$

Przejdziemy teraz do systematycznego opisu rozwiązywania układów równań liniowych o współczynnikach rzeczywistych. W tym celu wprowadzamy fundamentalne dla całego wykładu pojęcie macierzy.

Definicja 1.9: Macierz

MACIERZĄ ROZMIARU $m \times n$ lub inaczej – macierzą o m wierszach i n kolumnach o WYRAZACH ze zbioru X nazywamy tablicę:

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix}$$

gdzie $d_{ij} \in X$ dla $1 \leqslant i \leqslant m$ oraz $1 \leqslant j \leqslant n$.

Przy tym wprowadzamy następujące nazewnictwo.

- Rzędy poziome macierzy D nazywamy WIERSZAMI (uporządkowanymi od góry do dołu), a dokładniej — wyrazy $d_{i1}, d_{i2}, \dots, d_{in}$ tworzą i-ty wiersz macierzy D.
- Rzędy pionowe macierzy D nazywamy KOLUMNAMI (uporządkowanymi od lewej do prawej), a dokładniej – wyrazy $d_{1j}, d_{2j}, \dots, d_{mj}$ tworzą j-tą kolumnę macierzy D.
- Macierz D o wyrazach d_{ij}, gdzie $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$ oznaczamy w skrócie $D = [d_{ij}]$.
- Zbiór wszystkich macierzy $m \times n$ o wyrazach ze zbioru X oznaczamy jako $M_{m\times n}(X)$.

O macierzy można myśleć też jako o funkcji ze zbioru par $\{1, \dots, m\} \times \{1, \dots, n\}$ do zbioru X przypisującej każdej parze (i, j) (dla odpowiednich indeksów) element $d_{ij} \in X$, który nazywać będziemy WYRAZEM i, j macierzy D.

Oto przykład macierzy D o 3 wierszach i 5 kolumnach o wyrazach w $\mathbb{R}$, czyli elementu zbioru $M_{3\times 5}(\mathbb{R})$. W drugim wierszu i czwartej kolumnie tej macierzy znajduje się wyraz 1, czyli jeśli $D = [d_{ij}]$, to $d_{24} = 1$.

$$\begin{bmatrix} 1 & 2 & 3 & 0 & -1 \\ \sqrt{2} & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 1 & 3 & 7 & 9 \end{bmatrix}$$

Z drugiej strony zobaczmy przykład macierzy o 5 wierszach i 3 kolumnach, czyli elementu zbioru $M_{5\times 3}(\mathbb{R})$:

$$\begin{bmatrix} 0 & 1 & -1 \\ 2 & 3 & 5 \\ 1 & 1 & 1 \\ 0 & \sqrt{2} & 3 \\ 9 & 0 & 1 \end{bmatrix}.$$

Naszym celem jest przypisanie układom równań pewnych macierzy. Jak się okaże, zakodujemy za ich pomocą proces wykonywania operacji elementarnych na układach równoważnych i wskażemy typy macierzy reprezentujących układy równań liniowych, z których łatwo odczytać informację o istnieniu rozwiązań tych układów oraz łatwo odczytywać same rozwiązania.

Definicja 1.10: Macierz układu równań liniowych

Każdemu układowi m równań liniowych o n zmiennych i współczynnikach w K postaci:

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases}$$

możemy przypisać macierz rozmiaru $m \times (n+1)$ o wyrazach w K postaci:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

Nazywamy ją MACIERZĄ (ALBO MACIERZĄ ROZSZERZONĄ) UKŁADU U. Macierz powstałą przez pominięcie ostatniej kolumny w macierzy układu U będziemy nazywać MACIERZĄ WSPÓŁCZYNNIKÓW UKŁADU U.

Rozważmy kilka przykładów macierzy i odpowiadających im układów równań liniowych:

układ	macierz rozszerzona	macierz współczynników
$\begin{cases} 2x_1 + 3x_2 + x_3 = 1 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}$	$\left[\begin{array}{ccc\|c} 2 & 3 & 1 & 1 \\ 0 & -\frac{1}{2} & 1 & 0 \end{array}\right]$	$\left[\begin{array}{ccc} 2 & 3 & 1 \\ 0 & -\frac{1}{2} & 1 \end{array}\right]$
$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \\ 0x_1 + 0x_2 = 3 \end{cases}$	$\left[\begin{array}{cc\|c} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 3 \end{array}\right]$	$\left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{array}\right]$
$\begin{cases} 0x_1 + 0x_2 + 0x_3 = 0 \\ 0x_1 + 0x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}$	$\left[\begin{array}{ccc\|c} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right]$	$\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{array}\right]$

Przyjrzyjmy się teraz macierzom układów będących rozwiązaniami ogólnymi układów równań liniowych, opisanymi w Definicji 1.8. Oto przykłady.

układ	macierz rozszerzona
$\begin{cases} x_1 = 2 \\ x_2 = 1 \\ x_3 = 4 \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right]$
$\begin{cases} x_1 = \frac{1}{2} \\ x_2 + x_3 + x_4 = \frac{1}{2} \end{cases}$	$\left[\begin{array}{cccc\|c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 1 & \frac{1}{2} \end{array}\right]$
$\begin{cases} x_1 + x_3 - 2x_5 = 0 \\ x_2 + x_5 = 0 \\ x_4 + x_5 = 0 \end{cases}$	$\left[\begin{array}{ccccc\|c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$

W powyższych przykładach pokreślono kolorem współczynniki odpowiadające zmiennym zależnym. Widać, że w każdym kolejnym wierszu współczynniki te znajdują się w kolumnach o większym indeksie.

Definicja 1.11: Operacje elementarne na wierszach

Niech $A \in M_{m\times n}(\mathbb{R})$. Następujące transformacje macierzy A nazywamy OPERACJAMI ELEMENTARNYMI NA WIERSZACH:

(1) dodanie do wiersza innego wiersza przemnożonego przez liczbę rzeczywistą,

(2) zamiana dwóch wierszy miejscami,

(3) pomnożenie wiersza przez liczbę różną od zera.

Analogicznie definiuje się OPERACJE ELEMENTARNE NA KOLUMNACH macierzy A.

Widzimy zatem, że operacjom elementarnym na układzie U odpowiadają operacje elementarne na wierszach macierzy tego układu. Proces znajdowania rozwiązania ogólnego układu U będzie polegał na kolejnym upraszczaniu macierzy tego układu, przy zastosowaniu operacji elementarnych na wierszach.

Przykłady operacji elementarnych na wierszach macierzy i notacja, którą będziemy stosować:

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 2 & -1 & -2 & 1 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix} \xrightarrow{w_2-2\cdot w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix}$$

$$\xrightarrow{w_3+w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{bmatrix}$$

$$\xrightarrow{w_2\leftrightarrow w_3} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & 2 & 0 & 3 & 1 \\ 0 & -5 & 0 & 3 & 1 \end{bmatrix}$$

$$\xrightarrow{2\cdot w_2} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & 4 & 0 & 6 & 2 \\ 0 & -5 & 0 & 3 & 1 \end{bmatrix}$$

Definicja 1.12: Macierz w postaci schodkowej

Mówimy, że macierz A jest w POSTACI SCHODKOWEJ, jeśli A spełnia następujące warunki:

- każdy wiersz zerowy (tzn. złożony z samych zer) znajduje się poniżej każdego wiersza niezerowego (czyli jeśli i-ty wiersz tej macierzy jest niezerowy, to j-ty wiersz jest zerowy tylko gdy $j > i$),
- dla każdego $i > 1$ pierwszy licząc od lewej niezerowy wyraz w i-tym wierszu znajduje się w kolumnie stojącej na prawo od pierwszego niezerowego wyrazu $(i-1)$-szego wiersza.

Rozważmy, za skryptem, kilka prostych przykładów ilustrujących nową definicję: dwie pierwsze macierze są w postaci schodkowej, dwie kolejne nie. Czy Czytelnik widzi owe „schodki”?

$$\begin{bmatrix} 0 & 4 & 7 & 2 \\ 0 & 0 & 1 & 12 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 5 & 1 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 8 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 12 \end{bmatrix}.$$

Definicja 1.13: Macierz w postaci zredukowanej

Mówimy, że macierz jest w ZREDUKOWANEJ POSTACI SCHODKOWEJ (krócej: w postaci zredukowanej), jeśli jest w postaci schodkowej oraz w każdym niezerowym wierszu pierwszy niezerowy wyraz wynosi 1 i jest jedynym niezerowym wyrazem w swojej kolumnie.

Przykłady:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right], \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 1 & \frac{1}{2} \end{array}\right], \quad \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right], \quad \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right].$$

Jeśli układ ma macierz w postaci schodkowej zredukowanej, to są takie kolumny gdzie powstaje „schodek", powiedzmy kolumny (od lewej) $j_1 < j_2 < \ldots < j_k$ i zmienna z kolumny j_i raz jeden występuje w całym układzie właśnie w i-tym równaniu. Co więcej stoi przy niej współczynnik 1. A zatem jest to macierz, z której łatwo uzyskać rozwiązanie ogólne układu równań, o którym mówiliśmy wcześniej. Chyba, że... jedna z tych kolumn to ostatnia kolumna. Wtedy sytuacja jest nieco inna. Czy Czytelnik widzi jaka? Pokażemy teraz kluczową obserwację: do uzyskania postaci schodkowej i schodkowej zredukowanej wystarczy stosować na wierszach macierzy odpowiednie operacje elementarne.

Twierdzenie 1.2

Każdą macierz $A \in M_{m\times n}(\mathbb{R})$ można:

(i) za pomocą operacji elementarnych typu (1) i (2) sprowadzić do postaci schodkowej,

(ii) za pomocą operacji elementarnych typu (1), (2) i (3) sprowadzić do postaci zredukowanej.

Dowód. Pokażemy jedynie (i). Dowód (ii) będzie ćwiczeniem.

Stosujemy zasadę indukcji matematycznej względem liczby wierszy macierzy[1]. Dla $m = 1$ twierdzenie jest oczywiste. Każda macierz o jednym wierszu jest w postaci schodkowej.

Załóżmy, że twierdzenie jest udowodnione dla macierzy o co najwyżej $m-1$ wierszach. Niech A będzie macierzą rozmiaru $m \times n$ o wyrazach rzeczywistych. Jeśli A jest macierzą zerową (to znaczy każdy jej wyraz jest zerem), to oczywiście jest schodkowa. Załóżmy więc, że A nie jest macierzą zerową. Niech s będzie indeksem pierwszej niezerowej kolumny macierzy A. Wybierzmy takie r, że $a_{rs} \neq 0$.

$$A = \begin{bmatrix} 0 & \ldots & 0 & a_{1s} & * & \ldots & * \\ 0 & \ldots & 0 & a_{2s} & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{rs} & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{ms} & * & \ldots & * \end{bmatrix}$$

Zamieniamy miejscami wiersze: pierwszy i r-ty (operacja typu (2)). Za pomocą operacji (1) zerujemy[2] wszystkie, poza pierwszym, wyrazy w s-tej kolumnie (od i-tego wiersza odejmujemy pierwszy przemnożony przez $\frac{a_{is}}{a_{rs}}$). Otrzymujemy w ten sposób macierz $A' = [a'_{ij}]$, w której kolumny o indeksach $1, \ldots, s-1$ są zerowe oraz $a'_{1s} \neq 0$ i $a'_{is} = 0$, dla $i > 1$.

$$A' = \begin{bmatrix} 0 & \ldots & 0 & a_{rs} & * & \ldots & * \\ 0 & \ldots & 0 & a_{2s} - \frac{a_{2s}}{a_{rs}} \cdot a_{rs} & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{1s} - \frac{a_{1s}}{a_{rs}} \cdot a_{rs} & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & a_{ms} - \frac{a_{ms}}{a_{rs}} \cdot a_{rs} & * & \ldots & * \end{bmatrix} = \begin{bmatrix} 0 & \ldots & 0 & a_{rs} & * & \ldots & * \\ 0 & \ldots & 0 & 0 & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & * & \ldots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & * & \ldots & * \end{bmatrix}$$

Niech $A'' \in M_{(m-1)\times n}(\mathbb{R})$ będzie macierzą otrzymaną z A' przez usunięcie pierwszego wiersza. Stosujemy do A'' założenie indukcyjne. Otrzymujemy macierz, która wraz z pierwszym wierszem macierzy A' tworzy macierz A''' w postaci schodkowej. Oczywiście A''' jest uzyskana z A przez ciąg operacji typu (1) i (2). □

[1] Jeśli Czytelnik nie spotkał się dotąd z zasadą indukcji, można rozumować nieco inaczej: zakładamy, że m jest najmniejszą liczbą naturalną taką, że macierzy o m wierszach nie można za pomocą operacji elementarnych sprowadzić do postaci schodkowej. Rozumowanie przedstawione w dowodzie wyżej pokazuje, że dojdziemy do sprzeczności z tym założeniem.

[2] Krok ten nazywany jest często ELIMINACJĄ i pochodzi od niego nazwa algorytmu opisującego uzyskiwanie postaci schodkowej (a także zredukowanej), jak również wielu równoważnych mu rozkładów macierzy.

Metodę rozwiązywania układów równań liniowych opisaną w powyższym wniosku nazywamy METODĄ ELIMINACJI GAUSSA rozwiązywania układów równań. Zobaczmy jak działa na konkretnym przykładzie.

$$\begin{cases} x_1 + 2x_2 - x_3 - x_4 = 0 \\ 2x_1 - x_2 - 2x_3 + x_4 + x_5 = 0 \\ -x_1 + x_3 + 4x_4 + x_5 = 0 \end{cases}$$

Za pomocą operacji wierszowych dokonujemy eliminacji w pierwszej kolumnie.

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 2 & -1 & -2 & 1 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix} \xrightarrow{w_2-2w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ -1 & 0 & 1 & 4 & 1 \end{bmatrix} \xrightarrow{w_3+w_1} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{bmatrix}$$

Teraz zajmujemy się drugą kolumną.

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 2 & 0 & 3 & 1 \end{bmatrix} \xrightarrow{w_3+\frac{5}{2}w_2} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & \frac{21}{2} & \frac{7}{2} \end{bmatrix}$$

Uzyskaliśmy macierz układu (równoważnego wyjściowemu) w postaci schodkowej. Teraz doprowadzamy ją do postaci zredukowanej. W tym celu wykonujemy operacje służące eliminacji w kolumnach, w których znajdują się wyrazy wiodące poszczególnych wierszy.

$$\begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & \frac{21}{2} & \frac{7}{2} \end{bmatrix} \xrightarrow{\frac{2}{21}\cdot w_3} \begin{bmatrix} 1 & 2 & -1 & -1 & 0 \\ 0 & -5 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix} \xrightarrow[w_1+w_3]{w_2-3w_3} \begin{bmatrix} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix}$$

Wreszcie, operacje eliminujące wyrazy w drugiej kolumnie prowadzą do uzyskania postaci zredukowanej.

$$\begin{bmatrix} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix} \xrightarrow{-\frac{1}{5}\cdot w_2} \begin{bmatrix} 1 & 2 & -1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix} \xrightarrow{w_1-2w_2} \begin{bmatrix} 1 & 0 & -1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \end{bmatrix}$$

Wniosek 1.1

Każdy niesprzeczny układ równań liniowych ma rozwiązanie ogólne, przy czym:

- aby znaleźć rozwiązanie układu U wystarczy macierz tego układu sprowadzić do zredukowanej postaci schodkowej elementarnymi operacjami na wierszach,
- jeśli otrzymana macierz nie zawiera wiersza postaci $0 \ldots 0\,1$, to można z niej odczytać rozwiązanie ogólne układu U,
- jeśli otrzymana macierz zawiera wiersz postaci $0 \ldots 0\,1$, to układ U jest sprzeczny.

Dowód. Sprowadzamy macierz układu U do zredukowanej postaci schodkowej. Jeśli otrzymana macierz ma wiersz $0 \ldots 0\,1$, to odpowiadający jej układ równań zawiera równanie $0x_1 + 0x_2 + \ldots + 0x_n = 1$, więc układ ten (i równoważny z nim układ U) jest sprzeczny. Jeśli otrzymana macierz w postaci schodkowej zredukowanej nie zawiera wiersza $0 \ldots 0\,1$, to odpowiadający jej układ równań ma (po pominięciu ewentualnych równań postaci $0x_1 + 0x_2 + \ldots + 0x_n = 0$) postać

$$\begin{cases} x_{j_1} + a_{1(j_1+1)}x_{1(j_1+1)} + \ldots + a_{1n}x_n = b_1 \\ x_{j_2} + a_{2(j_2+1)}x_{1(j_2+1)} + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ x_{j_k} + a_{k(j_k+1)}x_{1(j_k+1)} + \ldots + a_{kn}x_n = b_k \end{cases}$$

przy czym $j_1 < j_2 < \ldots < j_k$ oraz dla każdego $1 \leqslant s \leqslant k$ mamy $a_{ij_s} = 0$, dla wszystkich i. Stąd układ ten można przepisać do postaci

$$\begin{cases} x_{j_1} = -a_{1(j_1+1)}x_{1(j_1+1)} - \ldots - a_{1n}x_n + b_1 \\ x_{j_2} = -a_{2(j_2+1)}x_{1(j_2+1)} - \ldots - a_{2n}x_n + b_2 \\ \vdots \\ x_{j_k} = -a_{k(j_k+1)}x_{1(j_k+1)} - \ldots - a_{kn}x_n + b_k \end{cases}$$

stanowiącej rozwiązanie ogólne układu U. □

1.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Rozpoznawanie rozwiązań układu równań liniowych)
   Które z poniższych ciągów $(-1,1,1,-1)$, $(2,3,1,4)$, $(4,-3,2,1)$, $(4,0,-3,\frac{1}{2})$ są rozwiązaniami poniższego układu równań liniowych o współczynnikach rzeczywistych?
   $$\begin{cases} 3x_1 + 2x_2 + 4x_3 + 2x_4 &= 1 \\ 7x_1 + 5x_2 + 9x_3 + 4x_4 &= 3 \\ 5x_1 - 3x_2 + 7x_3 + 4x_4 &= 1 \end{cases}$$

2. (♠ Wyznaczanie rozwiązania ogólnego. Operacje elementarne na wierszach układu)
   Znajdź rozwiązanie ogólne poniższego układu równań, wskazując ciąg równoważnych układów równań, uzyskiwanych za pomocą elementarnych operacji. Opisz dokładnie te operacje.
   $$\begin{cases} 2x_1 + 3x_2 + x_3 + 2x_4 &= 1 \\ 4x_1 + 6x_2 + 3x_3 + 2x_4 &= 3 \\ 6x_1 + 9x_2 + 5x_3 + 2x_4 &= 5 \end{cases}$$

3. (♠ Zapisywanie zbioru rozwiązań układu równań liniowych w zależności od parametrów)
   Rozpatrzmy układ równań liniowych o współczynnikach rzeczywistych postaci
   $$\begin{cases} x_1 + 3x_2 - x_3 + 3x_4 &= 1 \\ 2x_1 + 7x_2 + x_3 + 6x_4 &= 4 \end{cases}$$
   Znajdź rozwiązanie ogólne tego układu. Wskaż zmienne zależne i niezależne (parametry).
   Wypisz wszystkie czwórki (s_1, s_2, s_3, s_4) będące rozwiązaniami tego układu.

4. Niech ciągi $(1,2,3,-1)$ oraz $(3,6,9,-3)$ będą rozwiązaniami pewnego układu równań liniowych U o współczynnikach rzeczywistych. Wykaż, że ciąg $(0,0,0,0)$ również jest rozwiązaniem układu U.

5. Załóżmy, że ciągi $(1,2,3,4)$ oraz $(2,0,0,1)$ są rozwiązaniami pewnego układu równań liniowych o współczynnikach rzeczywistych. Wykaż, że układ ten ma nieskończenie wiele rozwiązań.

6. Znajdź rozwiązania ogólne poniższych układów równań. Wskaż zmienne zależne i parametry.
   (a) $x_1 + x_2 + \ldots + x_n = 0$,
   (b) $x_1 = x_2 = \ldots = x_n$.

7. Niech $n \geqslant 3$. Rozwiąż, w zależności od n, układ równań:
   $$\begin{cases} x_1 + x_2 = 0 \\ x_i + x_{i+1} + x_{i+2} = 0 \quad (1 \leqslant i \leqslant n-2) \\ x_{n-1} + x_n = 0. \end{cases}$$

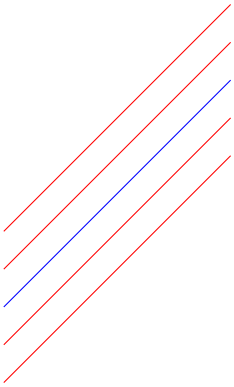
8. (♠ Wyznaczanie postaci schodkowej i zredukowanej macierzy)
   Wyznacz zredukowaną postać schodkową macierzy rzeczywistej:
   $$\begin{bmatrix} 3 & 1 & 1 & 1 \\ 2 & -1 & -1 & 2 \\ -2 & 2 & 4 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ -2 & 3 & 4 \\ 5 & 7 & 1 \\ 3 & 4 & -1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 1 & -1 \\ 5 & -1 & 1 & 2 \\ 7 & 8 & 1 & -7 \\ 1 & -1 & 1 & 2 \end{bmatrix}.$$

9. (♠ Macierz układu równań liniowych, kryterium rozwiązywalności układu równań liniowych)
   Wyznacz wartości parametru $a \in \mathbb{R}$, dla których poniższy układ równań jest niesprzeczny.
   $$\begin{cases} ax_1 + x_2 - 2x_3 + x_4 = a \\ x_1 + ax_2 + x_3 + ax_4 = 3 \\ x_1 + 2x_2 + x_3 + 2x_4 = 2 \end{cases}$$

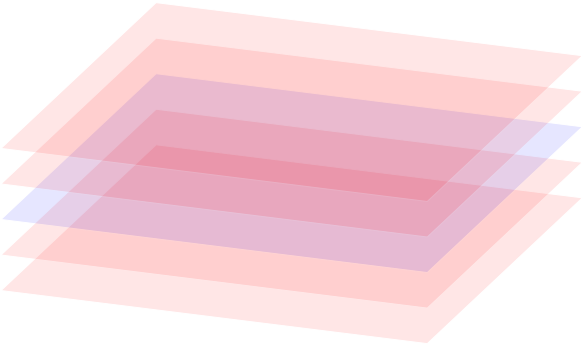
10. Macierz współczynników pewnego równań liniowych o współczynnikach rzeczywistych ma dwie identyczne kolumny. Wykaż, że układ ten jest albo sprzeczny, albo ma nieskończenie wiele rozwiązań.

1.3 Uzupełnienie. Wstępne uwagi o geometrii układów równań

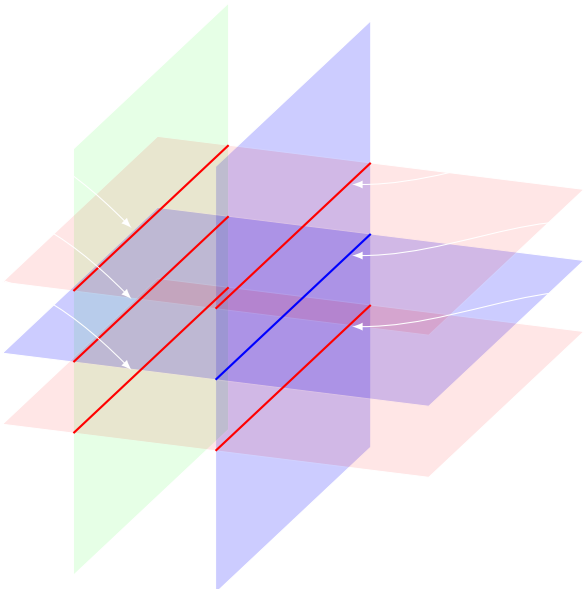
Jednym z celów naszego wykładu jest nadanie struktury geometrycznej obiektom, które (przynajmniej na poziomie intuicji) nie mają wiele wspólnego z geometrią. Rozważmy na przykład rodzinę równań U_t o zmiennych x_1, x_2 i współczynnikach w $\mathbb{R}$ postaci: $-x_1+x_2 = t$, gdzie $t \in \mathbb{R}$. Zgodnie ze szkolną geometrią analityczną rozwiązania (s_1, s_2) równania U_t można interpretować na płaszczyźnie kartezjańskiej jako współrzędne punktów stanowiących prostą przechodzącą przez punkty $(-t, 0)$ oraz $(0, t)$. Co istotne, dla różnych t uzyskujemy różne proste – wszystkie jednak równoległe do prostej $-x_1+x_2 = 0$ (stąd np. układ równań $-x_! + x_2 = 0, -x_1 + x_2 = 1$ jest sprzeczny).



Oto inna rodzina równań liniowych o współczynnikach w $\mathbb{R}$, o zmiennych x_1, x_2, x_3, postaci $x_3 = t$, dla $t \in \mathbb{R}$. W trójwymiarowej przestrzeni zbiory rozwiązań tego układu złożone są z trójek postaci (r, s, t), gdzie r, s to dowolne liczby rzeczywiste. Są to zbiory stanowiące układ równoległych płaszczyzn postaci:



Rozważmy wreszcie układ równań o zmiennych x_1, x_2, x_3 postaci $x_3 = t, x_2 = s$, gdzie $t, s \in \mathbb{R}$. Zbiory rozwiązań reprezentowane są przez przecięcia nieidentycznych i nierównoległych płaszczyzn, czyli proste



Przedstawimy teraz dwa proste fakty dotyczące zbiorów rozwiązań układów równań o n zmiennych o współczynnikach rzeczywistych, wiążące rozwiązania układu równań liniowych z odpowiadającym mu układem jednorodnym. Fakty te będą w przyszłości elementem dowodu tw. Kroneckera-Capellego.

Obserwacja 1.2

Jeśli $(s_1, s_2, \ldots, s_n)$ oraz $(s'_1, s'_2, \ldots, s'_n)$ są rozwiązaniami układu równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases}, \tag{1.1}$$

to ciąg

$$(s_1 - s'_1, s_2 - s'_2, \ldots, s_n - s'_n)$$

jest rozwiązaniem układu równań:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \tag{1.2}$$

Dowód. Ciągi $(s_1, s_2, \ldots, s_n)$ oraz $(s'_1, s'_2, \ldots, s'_n)$ sa rozwiązaniami każdego z równań

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i,$$

gdzie $i = 1, 2, \ldots, m$. Pisząc wprost mamy:

$$\begin{aligned} a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n &= b_i \\ a_{i1}s'_1 + a_{i2}s'_2 + \ldots + a_{in}s'_n &= b_i, \end{aligned}$$

czyli po odjęciu stronami widzimy, że $(s_1 - s'_1, s_2 - s'_2, \ldots, s_n - s'_n)$ spełnia każde z m równań układu (1.2):

$$a_{i1}(s_1 - s'_1) + a_{i2}(s_2 - s'_2) + \ldots + a_{in}(s_n - s'_n) = 0,$$

co oznacza, że jest to rozwiązanie całego układu (1.2). □

Obserwacja 1.3

Załóżmy, że $(s_1, \ldots, s_n)$ jest rozwiązaniem układu równań (1.1). Wówczas każde rozwiązanie układu (1.1) jest postaci:

$$(s_1 + u_1, \ldots, s_n + u_n) \qquad (\diamondsuit),$$

gdzie $(u_1, \ldots, u_n)$ jest rozwiązaniem układu (1.2).

Używając (na razie poglądowo) interpretacji geometrycznej można powiedzieć, że zbiór rozwiązań układu niejednorodnego U reprezentowany jest przez zbiór równoległy do zbioru rozwiązań układu jednorodnego U', odpowiadającego układowi U – zbiór zawierający dowolne rozwiązanie układu U.

Dowód. Weźmy rozwiązanie $(s_1, \ldots, s_n)$ układu (1.1) i rozwiązanie $(u_1, \ldots, u_n)$ układu (1.2). Wówczas ciąg $(s_1 + u_1, \ldots, s_n + u_n)$ jest rozwiązaniem układu (1.1), bo spełnia dowolne z jego równań:

$$a_{i1}(s_1 + u_1) + \ldots + a_{in}(s_n + u_n) = (a_{i1}s_1 + \ldots + a_{in}s_n) + (a_{i1}u_1 + \ldots + a_{in}u_n) = b_i + 0 = b_i.$$

Pozostaje wykazać, że dowolne rozwiązanie $(s'_1, \ldots, s'_n)$ układu (1.1) można przedstawić w postaci $(\diamondsuit)$, dla pewnego rozwiązania $(u_1, \ldots, s_n)$ układu (1.2). Zgodnie z poprzednią Obserwacją wystarczy wziąć:

$$u_1 = s'_1 - s_1, \quad \ldots \quad , u_n = s'_n - s_n$$

i dostajemy $(s'_1, \ldots, s'_n) = (s_1 + (s'_1 - s_1), \ldots, s_n + (s'_n - s_n))$. □

Dla zilustrowania powyższych faktów rozważmy układ równań liniowych o współczynnikach w $\mathbb{R}$ postaci:

$$\begin{cases} x+2y+3x &= 1 \\ 4x+5y+6z &= 1 \\ 7x+8y+9z &= 1 \end{cases} \qquad (\dagger)$$

Zgodnie z procedurą opisaną wyżej do opisania rozwiązań tego układu potrzebujemy dowolne jego rozwiązanie, na przykład $(s_1, s_2, s_3) = (-1, 1, 0)$ oraz wszystkie rozwiązania układu postaci:

$$\begin{cases} x+2y+3x &= 0 \\ 4x+5y+6z &= 0 \\ 7x+8y+9z &= 0 \end{cases} \qquad (\dagger)$$

Można sprawdzić, że rozwiązania tego układu są postaci $(z, -2z, z)$, gdzie $z \in \mathbb{R}$. W związku z tym każde rozwiązanie układu $(\dagger)$ ma postać $(-1+z, 1-2z, z)$, gdzie $z \in \mathbb{R}$.

Jaka jest zatem geometryczna natura naszego problemu? Chodziło o wyznaczenie miejsca geometrycznego przecięcia trzech płaszczyzn opisanych równaniami $x+2y+3z = 1$, $4x+5y+6z = 1$ oraz $7x+8y+9z = 1$. Najpierw, wyznaczyliśmy miejsce geometryczne przecięcia równoległych do tych płaszczyzn postaci: $x+2y+3z = 0$, $4x+5y+6z = 0$ oraz $7x+8x+9z = 0$ (to jest prosta $(z, -2z, z)$, gdzie $z \in \mathbb{R}$, którą oznaczaliśmy na niebiesko na wcześniejszych rysunkach). Później musieliśmy sprawdzić, czy każdą z tych płaszczyzn da się przesunąć równolegle tak, by przecięcie spełniało trzy wyjściowe równania. Nie zawsze musi się to udać. Trójka płaszczyzn może nie mieć punktów wspólnych (jak to wygląda?).

Interpretacja geometryczna układu równań zaprezentowana wyżej stanowi swego rodzaju **wierszowy obraz** tego układu – patrzymy na każde równanie osobno, interpretujemy jego rozwiązania jako podzbiory przestrzeni trójwymiarowej i szukamy części wspólnej tych zbiorów. Przejdźmy teraz do odrobinę mniej intuicyjnego, **kolumnowego obrazu** opisującego układ $(\dagger)$. Przepiszmy ten te układ do postaci:

$$x\begin{bmatrix}1\\4\\7\end{bmatrix} + y\begin{bmatrix}2\\5\\8\end{bmatrix} + z\begin{bmatrix}3\\6\\9\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}. \qquad (\clubsuit)$$

Widzimy, że naszym celem jest teraz znalezienie takich liczb rzeczywistych x, y, z, aby suma pewnych „skalarnych wielokrotności" wektorów $(1,4,7)$,$(2,5,8)$, $(3,6,9)$ stała się wektorem $(1,1,1)$. Ta nowa interpretacja geometryczna jest, co może być uznane za zaskoczenie, niezwykle istotna. Dlaczego? Przede wszystkim dlatego, że pozwala nam dużo powiedzieć o strukturze rozwiązań układu $(\dagger)$, o czym przekonamy się dalej. Po drugie, interpretacja ta jest wartościowa, bo pozwala nam zobaczyć na ile istotny jest wybór tego, a nie innego układu współczynników. Zmiana wektora $(1,1,1)$ na inny może całkowicie zmienić odpowiedź na pytanie czy dany układ ma rozwiązanie, czy nie. Na przykład układ postaci:

$$x\begin{bmatrix}1\\4\\7\end{bmatrix} + y\begin{bmatrix}2\\5\\8\end{bmatrix} + z\begin{bmatrix}3\\6\\9\end{bmatrix} = \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

nie może mieć rozwiązania, bo wektor $(0,0,1)$ „nie leży" na tej samej płaszczyźnie w przestrzeni, co wektory $(1,4,7)$, $(2,5,8)$. $(3,6,9)$. To nie jest oczywiste, bo nie widać wcale, że te trzy wektory muszą leżeć na płaszczyźnie i to takiej, w której nie ma wektora $(0,0,1)$. Jeśli jednak przepiszemy w języku wektorowym operacje wykonane na wyjściowym układzie, wówczas okaże się, że jest to jasne. Zauważmy, że układ

$$\begin{cases} x &= 1-2y-3z \\ y+2z &= 1 \\ 0 &= 0 \end{cases},$$

równoważny układowi $(\dagger)$ przepisuje się w języku „wektorów kolumnowych" jako:

$$x\begin{bmatrix}1\\0\\0\end{bmatrix} + y\begin{bmatrix}2\\1\\0\end{bmatrix} + z\begin{bmatrix}3\\2\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix}. \qquad (\spadesuit)$$

To, co jest kluczowe to fakt, że lewe strony warunków $(\clubsuit)$ oraz $(\spadesuit)$ opisują ten sam zbiór wektorów (tego nie widać „na wierszach"). Teraz jest już jasne, że $(0,0,1)$ nie jest jego elementem.

Badanie **kombinacji liniowych** wektorów (definicja na kolejnych wykładach) postaci $xv_1 + yv_2 + zv_3$ oraz ich geometrycznej natury jest centralnym zagadnieniem naszego wykładu. Aby to zrozumieć, na razie na poziomie wprowadzonych na wykładzie pojęć, prześledźmy rozwiązanie następującego zadania.

Zadanie. Ciągi $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$ są rozwiązaniami pewnego jednorodnego układu równań liniowych o współczynnikach rzeczywistych. Opisz wszystkie rozwiązania tego układu zakładając, że jego macierz po sprowadzeniu do postaci schodkowej ma trzy schodki.

W kontekście tego zadania rozważmy trzy pytania:

(1) Które jednorodne równania liniowe mają wśród rozwiązań ciągi $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$? Które ich **układy** zadają macierz o 3 schodkach?

(2) Jakie rozwiązania musi mieć każdy znaleziony **układ**, poza tymi dwoma?

(3) A jakich rozwiązań nie może mieć żaden znaleziony wyżej **układ**?

Odpowiedź na pytanie pierwsze jest prosta: takie równania jednorodne można po prostu wypisać. Są to równania postaci $a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = 0$ takie, że:

$$\begin{cases} a_1 + 2a_2 + 3a_3 + 4a_4 + 5a_5 &= 0 \\ 2a_1 + a_4 &= 0 \end{cases}. \quad (\ddagger)$$

Wniosek: zbiór **wszystkich równań liniowych jednorodnych**, które mają (nie tylko) rozwiązania $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$ ma *strukturę* zbioru rozwiązań układu równań[3] , czyli opisany jest przez piątki (a_1,a_2,a_3,a_4,a_5) spełniające $(\ddagger)$.

Wystarczy wziąć 3 rozwiązania układu $(\ddagger)$ postaci

$$(*,*,*,*,*), \quad (0,*,*,*,*), \quad (0,0,*,*,*)$$

i dostajemy układ 3 równań tworzący macierz o 3 schodkach, będącą macierzą układu, którego rozwiązaniami są między innymi $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$. Należy pokazać, że **każdy wybrany w ten sposób układ ma ten sam zbiór rozwiązań!** Oczywiście, są i inne układy o współczynnikach spełniających $(\ddagger)$, które mają ten sam zbiór rozwiązań, ale są i takie, które mają *więcej* rozwiązań, np. każdy układ złożony z pojedynczego równania jednorodnego spełniającego warunki z $(\ddagger)$.

Przechodzimy do pytania drugiego. Do zbioru rozwiązań **dowolnego** układu równań jednorodnych, który ma rozwiązania $(1,2,3,4,5)$ oraz $(2,0,0,1,0)$ należą też:

(a) każda *skalarna wielokrotność* $\lambda_1 \cdot (1,2,3,4,5) = (\lambda_1 \cdot 1, \lambda_1 \cdot 2, \lambda_1 \cdot 3, \lambda_1 \cdot 4, \lambda_1 \cdot 5)$,

(b) każda *skalarna wielokrotność* $\lambda_2 \cdot (2,0,0,1,0) = (\lambda_2 \cdot 2, \lambda_2 \cdot 0, \lambda_2 \cdot 0, \lambda_2 \cdot 1, \lambda_2 \cdot 0)$,

(c) każda *suma* $\lambda_1 \cdot (1,2,3,4,5) + \lambda_2 \cdot (2,0,0,1,0)$ tych *skalarnych wielokrotności*, czyli

$$(\lambda_1 \cdot 1 + \lambda_2 \cdot 2, \lambda_1 \cdot 2 + \lambda_2 \cdot 0, \lambda_1 \cdot 3 + \lambda_2 \cdot 0, \lambda_1 \cdot 4 + \lambda_2 \cdot 1, \lambda_1 \cdot 5 + \lambda_2 \cdot 0).$$

Uwaga: poza $\lambda_1 = \lambda_2 = 0$, nigdy nie mamy $\lambda_1 \cdot (1,2,3,4,5) = \lambda_2 \cdot (2,0,0,1,0)$.

A jakich rozwiązań nie może mieć żaden wyznaczony wcześniej układ $(\ddagger)$? Macierz każdego rozważanego układu ma po sprowadzeniu do postaci schodkowej 3 schodki, a zatem rozwiązanie *zależy* od dwóch parametrów. Oto delikatniejsze pytanie – czy dowolny taki układ może mieć rozwiązania poza zbiorem:

$$\{\lambda_1 \cdot (1,2,3,4,5) + \lambda_2 \cdot (2,0,0,1,0)\,,\ \lambda_1, \lambda_2 \in \mathbb{R}\}?$$

Zagadnienia, które tu szkicujemy należą już do serca naszego wykładu. Widzimy, że na razie brakuje nam precyzyjnego języka pozwalającego na opisanie zachodzących wyżej zjawisk. Język ten otrzymamy wraz z pojęciem przestrzeni liniowej nad ciałem oraz związanymi z nim pojęciami układów liniowo niezależnych. Geometrycznie rzecz biorąc budować będziemy teorię pojęcia zwanego wymiarem przestrzeni.

[3] W języku dalszej części uzupełnienia (kolejne dwie strony) – jest to przy odpowiedniej interpretacji zbiór wektorów prostopadłych (w standardowym sensie) jednocześnie do $(1,2,3,4,5)$ oraz do $(2,0,0,1,0)$ w przestrzeni pięciowymiarowej.

Powiedzieliśmy kilka słów o relacji między układami równań, a odpowiadającymi im równaniami jednorodnymi, wysławiając ją na przykładach równoległości pewnych prostych i płaszczyzn. Powiedzmy też kilka słów, choćby na poziomie intuicji, o wątkach geometrycznych wiążących układy równań z prostopadłością.

Rozważmy równanie liniowe o zmiennych x_1, x_2 o współczynnikach w $\mathbb{R}$ postaci

$$x_1 + 2x_2 = 0.$$

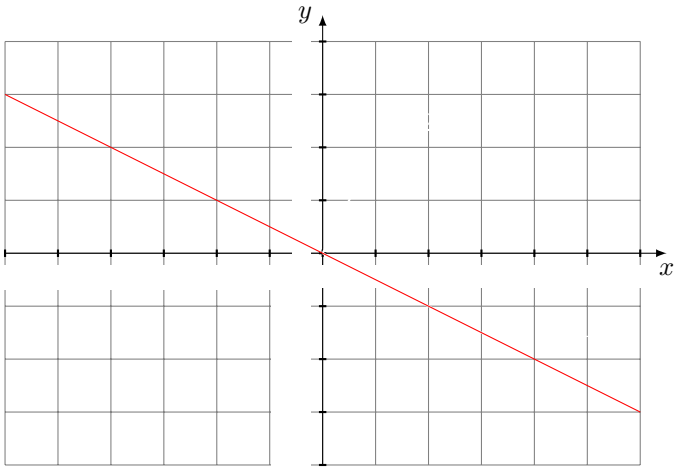
Zapiszmy je w następujący sposób:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

Cóż zrobiliśmy? W istocie wykonaliśmy iloczyn macierzy współczynników tego układu przez macierz zawierającą zmienne. Nie mówiliśmy o definicji iloczynu macierzy, ponieważ przez najbliższe wykłady nie będzie nam ona potrzebna. Wystarczy nam jednak pozostać w tym momencie na poziomie intuicji. Układ równań $x_1 + x_3 = 0, -2x_1 + x_3 = 0$ można zapisać w tej konwencji w postaci:

$$\begin{bmatrix} 1 & 0 & 1 \\ -2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

W powyższej konwencji wiersze i kolumny macierzy traktować będziemy jako wektory. Weźmy dowolne rozwiązanie pierwszego układu, na przykład $(4, -2)$. Traktując to rozwiązanie jako współrzędne wektora[4] na płaszczyźnie kartezjańskiej możemy zauważyć, że wektor ten wyznacza (przynajmniej w szkolnym rozumieniu) kierunek prostopadły do wektora o współrzędnych równych kolejnym współczynnikom naszego równania, czyli $(1, 2)$. Co więcej, zbiór rozwiązań naszego równania wyznacza zbiór wszystkich współrzędnych wektorów prostopadłych do wektora $(1, 2)$ (początek w punkcie $(0, 0)$, koniec w punkcie (2t,-t), dla $t \in R$). Bez trudu potwierdzimy tą obserwację np. za pomocą twierdzenia Pitagorasa.

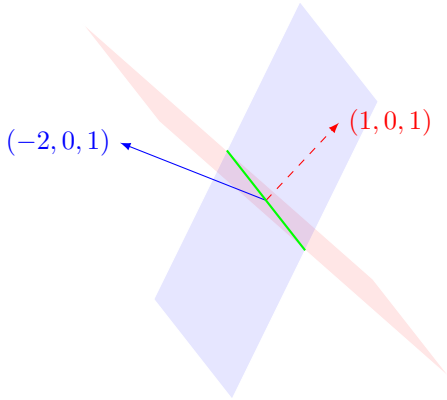


Okazuje się, że dla dowolnego niezerowego wektora (a, b), zbiór współrzędnych wektorów (x_1, x_2) prostopadłych do niego wyraża się warunkiem

$$ax_1 + bx_2 = \begin{bmatrix} a & b \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

Do niedawna jeszcze w języku szkolnym wyrażenie $ax_1 + ab_2$ nazywało się iloczynem skalarnym wektorów (a, b) oraz (x_1, x_2). W drugim semestrze będziemy dużo mówić o iloczynie skalarnym, nie tylko w prostym ujęciu pokazanym tutaj. Okaże się, że parze wektorów można przypisać wiele różnych „iloczynów skalarnych". Domyślać się jednak możemy, że obserwacja poczyniona wyżej nie może ograniczać się do jednego równania. Przejdźmy do drugiego układu. Czy on również opisuje wektory prostopadłe?

[4]Tu od razu stawiamy zastrzeżenie: na wykładzie czwartym pojęcie „wektora" uzyska abstrakcyjne znaczenie – będzie to element przestrzeni liniowej. Wektorami będą mogły być macierze, ciągi, wielomiany, funkcje, zbiory itd. W tym miejscu mówimy o wektorach stosując rozumienie szkolne, gdzie wektor jest rozumiany jako para uporządkowana punktów. Pierwszy element tej pary nazywamy „początkiem" wektora, a drugi – jego końcem. Jeśli punkty nasze znajdują się w układzie współrzędnych, to każdemu wektorowi przypisujemy też współrzędne będące ciągiem liczb powstającym przez odjęcie odpowiednio współrzędnych końca tego wektora od współrzędnych jego początku. Utożsamia się przy tym wektory o identycznych współrzędnych. Do szkolnej interpretacji wektora o „początku" i „końcu" wrócimy w drugim semestrze.



Można sprawdzić, że wektory o współrzędnych (a, b, c) oraz (x_1, x_2, x_3) są prostopadłe[5] (w szkolnym sensie, który będziemy kiedyś nazywali „standardowym") wtedy i tylko wtedy, gdy $ax_1 + bx_2 + cx_3 = 0$. W naszym przypadku płaszczyzna $x_1 + x_3 = 0$ zawiera „końce" wszystkich wektorów (o początkach w punkcie $(0, 0, 0)$) prostopadłych do wektora $(1, 0, 1)$, natomiast płaszczyzna $-2x_1 + x_3 = 0$ zawiera „końce" wszystkich wektorów prostopadłych do wektora $(-2, 0, 1)$. Prosta $(0, t, 0), t \in \mathbb{R}$, będąca rozwiązaniem tego układu, zawiera „końce" wektorów prostopadłych jednocześnie do $(-2, 0, 1)$ oraz $(1, 0, -1)$.

Chciałoby się wygłosić następujący morał: zbiór rozwiązań układu równań jednorodnych „opisuje" zbiór prostopadły do wszystkich wektorów znajdujących się w wierszach macierzy współczynników. Dopóki nie dysponujemy precyzyjnym nazewnictwem algebry liniowej i afinicznej może być nam trudno pójść dalej. Warto zwrócić jednak uwagę na jeszcze jedną rzecz. Oto przykład dwóch dualnych zdań, wziętych z geometrii szkolnej:

- „Punkty A, B leżą na prostej c".
- „Proste a, b przecinają się w punkcie C".

Dualność to rodzaj matematycznej symetrii. Nawiązujemy tu do języka geometrii, ale w dualności chodzi o to, że mamy dwa typy obiektów i jakiegoś rodzaju relacje wiążące te obiekty. W naszym przykładzie obiekty to „punkty" i „proste", zaś relacje to „leżenie na" i „przecinanie się". Czytelnik z pewnością widzi pary dualnych obiektów i dualnych operacji. W języku geometrii formalnej mówi się nawet o relacji incydencji[6] – punkt jest incydentny do prostej (leży na niej) oraz prosta jest incydentna do punktu (przechodzi przez niego). Są teorie geometryczne oparte o tzw. „aksjomaty incydencji". W dualności „chodzi o to", że obiekty i operacje na nich grają niekiedy symetryczne role i to do tego stopnia, że stają sie „jakby" nierozróżnialne. Precyzyjnie wyraża to język tak zwanej teorii kategorii, który poznamy pod koniec roku. Powyższy przykład możemy natomiast odnieść wprost do algebry liniowej formułując zdania:

- „Ciągi (a, b), (a', b') spełniają równanie liniowe C postaci $cx_1 + dx_2 = 0$".
- „Równania $ax_1 + bx_2 = 0, a'x_1 + b'x_2 = 0$ spełnione są jednocześnie przez ciąg (c, d)".

Algebraicznie obydwa zdania napisane wyżej wyrażają się dokładnie tymi samymi formułami:

$$\begin{cases} ac + bd & = 0 \\ a'c + b'd & = 0 \end{cases}.$$

Ową nierozróżnialność widać właśnie w układzie równań napisanym wyżej. Opisuje on dualne sytuacje – leżenie punktów na prostej i przecinanie się prostych w punkcie. Tego typu dualność zapisana jest w każdym (jednorodnym) równaniu liniowym. Równanie $2x_1 + 3x_2 + 5x_3 = 0$ można interpretować tak jako równanie płaszczyzny (w przestrzeni) prostopadłej do wektora $(2, 3, 5)$. Jeśli jednak odwrócimy pytanie i zapytamy: do jakich wektorów prostopadła jest płaszczyzna opisana powyższym równaniem, to te wektory opisywać będą pewną prostą. Będzie złożona z punktów przestrzeni o współrzędnych $(2a, 3b, 5c)$. Zagadnienia te formalizować będziemy stopniowo, póki co powstrzymując się zarówno od użycia terminu „wektor" (na kilka wykładów), jak i terminów „punkt", „prosta", „płaszczyzna" oraz ‚równoległość" i „prostopadłość", którym nadamy precyzyjny sens w drugim semestrze. W dodatkach będziemy wracać do intuicyjnego rozumienia tych obiektów, celem ilustracji niektórych abstrakcyjnych pojęć i wyników.

[5] Patrz np. tekst dr. Michała Krycha `https://www.mimuw.edu.pl/~krych/chemia/2016-2017/ch10-11_geoman.pdf`.
[6] Patrz np. `http://www.math.uni.wroc.pl/~ivanov/Logika4.pdf`.

1.4 Dodatek. Nieliniowe układy równań

Pokazaliśmy, że rozwiązywanie układów równań liniowych polega na zastępowaniu ich przez inne układy o tym samym zbiorze rozwiązań. Zachowanie zbioru rozwiązań gwarantuje stosowanie operacji elementarnych. Czy podobna metoda da się stosować dla innych typów układów równań? Rozważmy układ:

$$\begin{cases} x^2+yz+x &= 0 \\ y^2+xz+y &= 0 \\ z^2+xy+z &= 0 \end{cases}.$$

Układ ten nie jest układem równań liniowych, ale wydaje się, że nic nie stoi na przeszkodzie by wykonywać na nim wszystkie opisane wcześniej operacje elementarne. Czy ich stosowanie zmienia zbiór rozwiązań? Po odjęciu drugiego równania od pierwszego oraz trzeciego równania od drugiego dostajemy (co chyba jasne) układ równoważny postaci:

$$\begin{cases} (x-y)(x+y-z+1) &= 0 \\ (y-z)(y+z-x+1) &= 0 \\ z^2+xy+z &= 0 \end{cases}.$$

Chwila skrupulatnego rozważania tych warunków (nic trudnego) prowadzi do trójek (x,y,z) rozwiązań:

$$(0,0,0),(-1,0,0),(0,-1,0),(0,0,-1),(-1/2,-1/2,-1/2)\,.$$

Czy rozwiążemy w ten sposób każdy układ równań wielomianowych (na razie rozumianych intuicyjnie)?
Czy ma sens mówienie o macierzy współczynników tego układu? Jak ją określić?
Czy jest jakiś analog postaci schodkowej macierzy związanych z takimi układami?
Czy są jakieś specyficzne dla wielomianów „operacje elementarne", które nie zmieniają zbioru rozwiązań?
Czym są „zmienne niezależne" w przypadku istnienia nieskończenie wielu rozwiązań układu równań?

Konkretne odpowiedzi na powyższe pytania związane są z działem algebry zwanym teorią pierścieni oraz odkryciem Buchbergera z 1965 roku dotyczącym wyznaczania tak zwanych baz Gröbnera. Kto by chciał poczytać więcej o tym algorytmie i jego związku z algorytmem Gaussa, może zajrzeć pod poniższe adresy:

`https://www.math.usm.edu/perry/Research/GaussToGroebner.pdf`
`https://www.theoremoftheday.org/MathsStudyGroup/Buchberger.pdf`

W tym miejscu damy Czytelnikowi jedną tylko intuicję, która pozwala zobaczyć w nieco innym świetle sam algorytm Gaussa i sens wyznaczania postaci schodkowej (stąd też obecność tego tematu w tym miejscu). Załóżmy, że ograniczamy się do rozważania układów równań postaci:

$$\begin{cases} F_1(x,y) &= 0 \\ & \cdots \\ F_s(x,y) &= 0 \end{cases},$$

gdzie $F_i(x,y)$ są, dla $1 \leqslant i \leqslant s$, wielomianami zmiennych x,y, czyli formalnymi sumami jednomianów postaci $c\cdot x^m y^n$, $m,n \geqslant 0$, $c\in\mathbb{R}$, zaś $m,n\in\mathbb{Z}$ (liczby całkowite!). Jeśli każdy z takich jednomianów spełnia warunek $m+n=1$, otrzymamy po prostu układy równań liniowych (o dwóch zmiennych).

W przypadku układu równań liniowych zmienne oznaczaliśmy symbolami $x_1, x_2, \ldots$ Oznacza to – choć na wykładzie nie zwracaliśmy na to uwagi – że zmienne tworzą **zbiór uporządkowany**. Macierz współczynników układu, postać schodkowa, opis rozwiązań – wszystko to zależy od **kolejności zmiennych**, jaką przyjmujemy (choć nie zmienia się natura geometryczna zbioru rozwiązań). Idea jest następująca: w przypadku wielomianów porządkować będziemy nie tylko same zmienne, ale cały zbiór jednomianów.

Zamiast x pisać będziemy x_1 oraz zamiast y pisać będziemy x_2. Chcemy przez to podkreślić, że x jest w porządku zmiennych „wcześniejsza". Wprowadzamy następnie zasadę porządkowania jednomianów. Przez stopień $deg(x_1^m x_2^n)$ jednomianu $x_1^m x_2^n$ rozumiemy sumę $m+n$. Ustalamy zasadę, że zawsze „wcześniejszy" (czyli większy) jest jednomian wyższego stopnia. A zatem $x_1^4x_2^2 > x_1^5$, $x_1x_2^6 > x_1^2x_2$ itd. Jeśli natomiast stopnie dwóch jednomianów są identyczne, wcześniejszy jest jednomian z wyższą potęgą przy wcześniejszej zmiennej. A więc, na przykład, $x_1^4x_2^3 > x_1^3x_2^4$, chociaż $deg(x_1^4x_2^3) = deg(x_1^3x_2^4) = 7$.

Układ równań napisany na początku naszego dodatku jest, przy założeniu porządku $x > y > z$, napisany tak, że kolejne składniki są coraz mniejsze w „porządku jednomianowym". Algorytm Buchbergera proponuje – do pewnego stopnia – wykonywanie podobnej procedury, co algorytm Gaussa. Chodzi o zastępowanie jednych wielomianów innymi i dążenie do prostszego układu. Możemy sobie już wyobrazić macierz układu równań wielomianowych. Co z operacjami elementarnymi? Wprowadza się (pozornie) dodatkową „operację elementarną" na dwóch wielomianach – tak zwany S-wielomian. Jak go określić?

Dla dwóch jednomianów $x_1^{m_1}x_2^{n_1}$ oraz $x_1^{m_2}x_2^{n_2}$ rozważa się tzw. najmniejszą wspólną wielokrotność (zbieżność użytego nazewnictwa z tym występującym w teorii podzielności liczb całkowitych nie jest przypadkowa). Nazwiemy ją $NWW(x_1^{m_1}x_2^{n_1}, x_1^{m_2}x_2^{n_2})$ i jest to po prostu jednomian $x_1^{\max\{m_1,m_2\}}x_2^{\max\{n_1,n_2\}}$.

Dla każdego wielomianu f przez $LT(f)$ oznaczamy największy jednomian – w określonym wyżej porządku – który występuje w f. Na przykład $LT(2x^3y^4 + y^6) = 2x^3y^4$ (uwaga – nie samo x^3y^4, ale $2x^3y^4$).

Dla wielomianów f, g zmiennych x, y, przez $S(f, g)$ rozumiemy wielomian postaci:

$$S(f,g) = \frac{NWW(LT(f), LT(g))}{LT(f_1)} \cdot f_1 - \frac{NWW(LT(f), LT(g))}{LT(f_2)} \cdot f_2.$$

Przykład. Rozważmy przecięcie dwóch elips (proszę uwierzyć mi na słowo, że są to elipsy) postaci:

$$\begin{cases} 2x^2 + y^2 - 4x - 4y + 3 &= 0 \\ x^2 + 3y^2 - 2x - 12y + 9 &= 0 \end{cases}.$$

Niech $f = 2x^2+y^2-4x-4y+3$ oraz $g = x^2+3y^2-2x-12y+9$. Wówczas przyjmując $x > y$ i respektując opisany wyżej porządek jednomianowy dostajemy $LT(f) = 2x^2, LT(g) = x^2$, a zatem widzimy, że:

$$S(f,g) = \frac{-5}{2}y^2 + 10y - \frac{15}{2}.$$

Czy Czytelnik widzi co się stało? Jednomian wiodący S-wielomianu $S(f, g)$ jest mniejszy – w porządku jednomianowym – niż jednomian wiodący każdego z wielomianów f, g. Czy widać tu analogię do uzyskiwania postaci schodkowej przez algorytm Gaussa, zwany często algorytmem **eliminacji** Gaussa? Czy widać, że wykonanie S-wielomianu na dwóch równaniach liniowego układu równań i zastąpienie jednego z równań owym S-wielomianem (dwóch zmiennych – biorąc pod uwagę nasze definicje – ale po odpowiednim uogólnieniu – dowolnym) jest równoważne ze zwykłą operacją elementarną rozważaną na wykładzie?

Czy Czytelnik widzi w jaki sposób stosowanie S-wielomianu prowadzi do zastąpienia wyjściowego układu równań wielomianowych układem prostszym? Jak należy kontynuować ten algorytm i do jakich rozwiązań może on prowadzić? Poszukiwanie odpowiedzi polecam już Państwu. Polecam wskazane wyżej źródła.

Ps. Rozważane elipsy przecinają się w czterech punktach. Dowód wymaga jedynie matematyki szkolnej.

* * *

Zapomniałbym wspomnieć: baz Gröbnera używa się do naprawdę ciekawych rzeczy (choćby w robotyce). Polecam dwa anglojęzyczne źródła (im szybciej Państwo zaczną czytać matematykę w tym języku, tym lepiej). Na razie mogą być one dla Państwa trudne w lekturze, ale za jakiś czas będzie można wrócić do tej lektury. A może warto przejść się na konsultacje i poprosić o wyjaśnienie trudniejszych fragmentów?

- Antoine Nectoux, *Map colouring and Gröbner Bases*
- Elizabeth Arnold, Stephen Lucas, and Laura Taalman, *Gröbner Basis Representations of Sudoku*

Aby znaleźć te teksty wystarczy wpisać tytuły w Google.

Polecam również dwugodzinny wykład prof. Przemysława Koprowskiego (Uniwersytet Śląski) dotyczący baz Gröbnera i ich zastosowań (https://youtu.be/vdmyrbNqRlY). Jest to znakomite, ścisłe wprowadzenie do tej tematyki, z licznymi przykładami zastosowań.

1.5 Trivia. Kwadraty magiczne

Pierwsza trivia dotyczy pięknego tematu, mianowicie kwadratów magicznych. Jest wiele poważnych tekstów, zarówno popularyzatorskich, jak i naukowych dotyczących tych obiektów i problemów z nimi związanych. Znane były one od starożytności, a fascynowały między innymi samego Eulera, który poświęcił im kilka swoich artykułów. Obecnie można znaleźć wiele ich tłumaczeń w tym na język angielski. np.

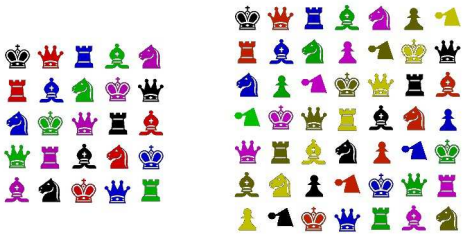
- E530 (http://eulerarchive.maa.org/docs/translations/E530.pdf),
- E795[7] (https://arxiv.org/pdf/math/0408230.pdf).

Czym więc są owe kwadraty? Może zamiast definicji podam problem 36 oficerów, badany przez Eulera. Mamy mianowicie 36 oficerów o sześciu różnych stopniach, wziętych z sześciu różnych oddziałów i próbujemy ustawić ich w kwadracie tak, by w każdym wierszu i każdej kolumnie tego kwadratu stało sześciu oficerów z innego oddziału i różnych stopni.

Oto stosowny obrazek pokazujący oficerów, ustawionych na razie zgodnie z przynależnością do oddziału.

Źródło: http://www.ams.org/publicoutreach/feature-column/fcarc-latinii1

Zachęcam do eksperymentu i próby dokonania żądanego ustawienia. Okaże się, że nie bardzo chce się to udać. Eulerowi też to nie wychodziło. A problem był o tyle frustrujący, że analogiczne problemy 25 i 49 oficerów daje się gładko rozwiązać:



Źródło: http://www.ams.org/publicoutreach/feature-column/fcarc-latinii1

Euler rozpoznaje w badanym zagadnieniu problem algebraiczny, znany zresztą wcześniej. Załóżmy, że każdemu oficerowi nadamy plakietkę a^b, gdzie a oznaczać będzie stopień, a b – numer oddziału, a potem zapomnimy o numerach oddziału i popatrzymy tylko na stopnie, to rozkład wyżej ma postać:

7	6	5	4	3	2	1
5	4	3	2	1	7	6
3	2	1	7	6	5	4
1	7	6	5	4	3	2
6	5	4	3	2	1	7
4	3	2	1	7	6	5
2	1	7	6	5	4	3

Dla Eulera kwadratem magicznym jest tablica liczb rozmiaru $n \times n$ taka, że w każdym wierszu, każdej kolumnie i na obydwu przekątnych sumy wpisanych liczb są równe. A jakie liczby wpisujemy? W przypadku wyżej: siedem zestawów od 1 do 7. Czasem rozważa się kwadraty, w które wpisuje się kolejne liczby naturalne (tzw. normalne kwadraty magiczne), a czasem półmagiczne (bez warunku na przekątne) itd. Problem oficerów rozwiązany został dopiero w 1901 roku przez matematyka-amatora Gastona Tarry'ego.

[7] Pewnie ciekawi Państwa co znaczą te liczby? W latach 1910-1913 szwedzki matematyk Gustav Eneström dokonał gruntownych badań nad dziełami Eulera i z nieprzebranych archiwów publikacji oraz zbiorów notatek wyłonił 866 pozycji: książek, artykułów i istotnych listów, którym przydzielił symbole od E1 do E866.

Aby przybliżyć się nieco do materiału z wykładu zajmijmy się magicznymi kwadratami rozmiaru 3×3 o wyrazach rzeczywistych. Możemy je oczywiście utożsamiać z macierzami. Oto przykłady:

$$\begin{bmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{bmatrix}, \quad \begin{bmatrix} -1 & 5/2 & 0 \\ 3/2 & 1/2 & -1/2 \\ 1 & -3/2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix}.$$

Ile jest macierzy magicznych? Powyższe przykłady sugerują, że jest ich nieskończenie wiele. Zauważmy, że wyznaczenie macierzy magicznej równoważne jest problemowi rozwiązania układu równań liniowych. Aby rozstrzygnąć czy macierz:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

reprezentuje kwadrat magiczny należy nałożyć następujące warunki na jej wyrazy:

$$\begin{cases} a_{21} + a_{22} + a_{23} - a_{11} - a_{12} - a_{13} = 0 \\ a_{31} + a_{32} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{11} + a_{21} + a_{31} - a_{11} - a_{12} - a_{13} = 0 \\ a_{12} + a_{22} + a_{32} - a_{11} - a_{12} - a_{13} = 0 \\ a_{13} + a_{23} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{11} + a_{22} + a_{33} - a_{11} - a_{12} - a_{13} = 0 \\ a_{31} + a_{22} + a_{13} - a_{11} - a_{12} - a_{13} = 0 \end{cases}.$$

czyli rozwiązać układ równań, którego macierz ma aż 7 wierszy i 9 kolumn odpowiadających zmiennym

$$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$$

postaci

$$\left[\begin{array}{ccccccccc|c} -1 & -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{array}\right].$$

Idę o zakład, że „schodkowanie" tej macierzy nie wydałoby się Państwu przyjemnością. Od ilu parametrów zależą rozwiązania tego układu? Nie bardzo widać rozwiązanie. Jest jednak sprytniejsza droga: skorzystamy z twierdzenia o rozwiązaniach układów jednorodnych i niejednorodnych. Proszę zauważyć, że jeśli mamy macierz magiczną $[a_{ij}]$ taką, że suma wyrazów w każdym wierszu, kolumnie na przekątnych to $3S$, to macierz o wyrazach $a_{ij} - S$ również jest magiczna, a nawet 0-magiczna, bo suma wyrazów w każdym jej wierszu, kolumnie i na przekątnych to 0. Intuicja podpowiada zatem, że liczba parametrów potrzebna do opisania wszystkich macierzy magicznych jest o 1 większa niż liczba parametrów służących do opisu macierzy 0-magicznych. Zobaczmy, że opis macierzy t-magicznych wygląda przejrzyściej. Można go dokonać przez rozwiązanie układu o macierzy:

$$\left[\begin{array}{ccccccccc|c} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & t \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & t \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & t \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & t \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & t \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & t \end{array}\right].$$

Znacznie łatwiej analizować powyższą macierz 8×9, czy ogólnie macierz rozmiaru $(2n+2) \times n^2$, uzyskiwaną dla układu opisującego macierze t-magiczne rozmiaru n. Tu już coś widać. Suma pierwszych trzech wierszy minus suma dwóch kolejnych daje wiersz szósty, więc po wyschodkowaniu jest co najmniej jeden wiersz zerowy... i jak się okazuje nie ma innych. Rozwiązanie zależy od dwóch parametrów. A zatem wszystkie macierze magiczne rozmiaru 3×3 opisać można za pomocą trzech parametrów[8].

[8] Czytelnik zechce zauważyć, że (posługując się geometryczną intuicją z Uzupełnienia) twierdzimy, że zbiory macierzy t-magicznych rozmiaru 3×3 stanowią rodzinę równoległych płaszczyzn w trójwymiarowej przestrzeni macierzy magicznych.

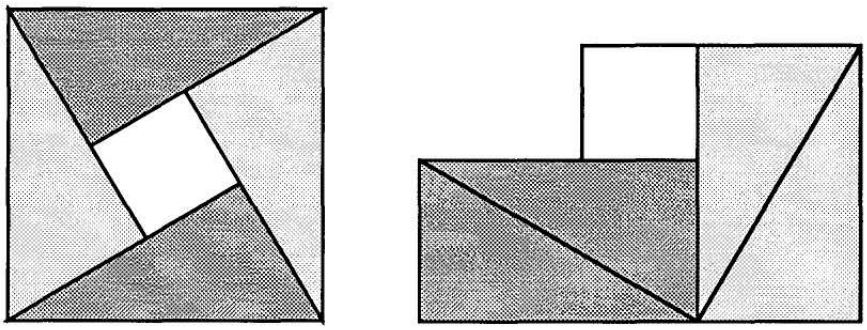
1.6 Coda. Eliminacja i podstawienie, czyli historia upraszczania

Gdyby chcieć wskazać jedną podstawową koncepcję przewijającą się przez całą metodologię nauki, to jest nią zasada mówiąca, że „Nie należy mnożyć bytów ponad potrzebę". Często przypisuje się ją Ockhamowi, ale podobne zasady „ekonomii myślenia" formułowali już Arystoteles, Platon, a także filozofowie średniowieczni. W wyjaśnianiu zjawisk należy dążyć do prostoty, opierać się na najmniejszej możliwej liczbie pojęć i założeń. Wnioski należy wyciągać z możliwie małej liczby racjonalnych przesłanek.

W matematyce zasada dążenia do prostoty wiąże się z wieloma zagadnieniami, a na pierwszym wykładzie tego kursu widzieliśmy ją przede wszystkim w opisie podstawowej metody rozwiązywania układów równań, polegającej na zastępowaniu jednych układów innymi — równoważnymi, o prostszej postaci. Drogą w tym kierunku jest eliminacja zmiennych z równań, którą osiągamy poprzez sprowadzanie postaci macierzy tego układu do postaci schodkowej. Warto powiedzieć, że tego rodzaju podejście obecne jest w historii algebry — nie tylko zresztą akademickiej, ale także tej, którą poznawaliście Państwo przez lata.

W szkole mieli Państwo okazję rozwiązywać wiele typów równań algebraicznych, począwszy od liniowych, przez kwadratowe, wielomianowe, wykładnicze, logarytmiczne czy trygonometryczne. Poznawali Państwo również podstawowe techniki rozwiązywania układów równań, bez ogólnej metody. Czy na studiach poszerzać będziemy arsenał umiejętności w tym zakresie? Do pewnego stopnia tak będzie, choć podstawowe techniki będą dalej istotne: grupowanie, wyciąganie wspólnego czynnika przed nawias, zwijanie do iloczynu, korzystanie z wzorów skróconego mnożenia, wzorów dwumianowych, później także metody rachunku różniczkowego itd. Kluczowe jest dla nas zrozumienie podstawowej koncepcji rozwiązywania równań — PODSTAWIANIA. Algorytm Gaussa przedstawiony na wykładzie jest przypadkiem ogólniejszego sposobu myślenia, którego początki sięgają starożytnych problemów „algebry geometrycznej".

Pierwsze formuły algebraiczne jakie poznajemy w szkole mają wszystkie, bez wyjątku pochodzenie geometryczne. Za pomocą wzorów zapisujemy obwody i pola podstawowych figur, oraz poprzez rozmaite techniki podziału figur na mniejsze, uzyskujemy nowe wzory — choćby na pole równoległoboku, czy w znacznie głębszym sensie — na pole koła. Za pomocą pól ilustrujemy wzory skróconego mnożenia na kwadrat sumy czy różnicy, a nawet być może zdarzyło się Państwu widzieć „obrazkowy dowód" twierdzenia Pitagorasa. Oto jedna z wielu możliwych wersji pochodząca od hinduskiego uczonego Bhaskary.



Rys. 1. Dowód twierdzenia Pitagorasa. Źródło: R. Nielsen: *Proofs without words.*

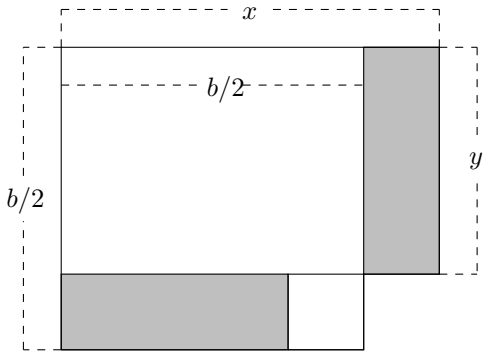
Problemy sprowadzające się do rozwiązywania równań znane były już w starożytnej Mezopotamii. Matematyka Babilończyków miała dwa źródła. Pierwszym było prowadzenie ksiąg rachunkowych, od początku istotnych dla funkcjonowania systemu biurokratycznego wczesnych dynastii rządzących tym obszarem 4000 lat temu. Termin *księgi* jest oczywiście umowny, bowiem teksty zapisywane były na tabliczkach glinianych. Drugim źródłem były problemy geometryczne związane głównie z zagadnieniami podziału terenu. Wiele starych tabliczek glinianych pochodzących z okresu 2000-1700 p.n.e. zawierają rozległe listy tego, co dziś nazwalibyśmy równaniami kwadratowymi, których celem było znalezienie takich wielkości jak długość czy szerokość prostokąta. Przykład takiego problemu pochodzi[9] z tzw. tabliczki YBC 4663.

Dane są: suma długości i szerokości prostokąta: $6\frac{1}{2}$ oraz pole prostokąta: $7\frac{1}{2}$. Wyznaczyć długość i szerokość tego prostokąta. Skryba opisuje detalicznie kroki w celu uzyskania rozwiązania. Oto one.

[9] Źródło: Victor J. Katz, *Stages in the history of algebra with implications for teaching.*

1. Przepołowić $6\frac{1}{2}$ otrzymując $3\frac{1}{4}$.
2. Podnieść uzyskaną liczbę do kwadratu uzyskując $10\frac{9}{16}$.
3. Od uzyskanego pola odjąć dane pole prostokąta $7\frac{1}{2}$, uzyskując $3\frac{1}{16}$.
4. Z uzyskanej liczby wyciągnąć pierwiastek: $1\frac{3}{4}$.
5. Stwierdzić, że długość prostokąta wynosi $3\frac{1}{4} + 1\frac{3}{4} = 5$, podczas gdy szerokość to $3\frac{1}{4} - 1\frac{3}{4} = 1\frac{1}{2}$.

Nie bez powodu opisujemy rozwiązanie językiem książki kucharskiej, ponieważ tak w istocie uczono się matematyki w starożytnej Mezopotamii — przez akumulację rozwiązanych zadań, a nie przez abstrahowanie i wyciąganie ogólnych zależności. O co chodzi w tym rozwiązaniu? Skryba miał pewnie przed oczami następujący obrazek, przy założeniu, że szukane wielkości to x, y, a znane są $x+y = b$ oraz $xy = c$.



Rys. 2. Problem z tabliczki YBC 4663.

Czy Czytelnik widzi, że kluczem tej metody jest porównanie pól kwadratu o boku $b/2$ i wyjściowego prostokąta, przez znalezienie w nich (odpowiednich miejscach) wspólnego szarego prostokąta tak, że różnicą pól jest również pole kwadratu, i to liczby $(x-y)/2$? Obrazek powyższy ilustruje więc wzory:

$$x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}, \quad y = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

Wzory te nie były oczywiście znane Babilończykom, jak również żadna ogólna metoda rozwiązywania takich zadań. Znana była niebanalna procedura postępowania przy konkretnych danych liczbowych.

Zupełnie inaczej wyglądała sytuacja w starożytnej Grecji, w której manipulacje algebraiczne wciąż wykonywane były w oparciu o obiekty geometryczne, w oparciu jednak o wyraźnie sformułowane aksjomaty. Oto przykład z *Elementów* Euklidesa — Twierdzenie II-5, znane nam jako wzór na kwadrat sumy.

> Jeśli podzielimy na dwie równe części, a także na dwie nierówne części, to prostokąt zawarty w nierównych częściach wraz z kwadratem linii łączącej punkty przecięcia jest równy kwadratowi połowy tej prostej.

Jeśli pomyślimy o „nierównych segmentach" jako o x oraz y, a o długości początkowego odcinka jako b, wówczas twierdzenie zdaje się twierdzić, że:

$$xy + \left(\frac{x-y}{2}\right)^2 = \left(\frac{x+y}{2}\right)^2$$

i w ten sposób rozwiązać można układ równań $x + y = b$ oraz $xy = c$. Dokładniej, jeśli PODSTAWIMY c zamiast xy orz b zamiast $x + y$, orzymamy

$$\left(\frac{x-y}{2}\right)^2 = \left(\frac{b}{2}\right)^2 - c \iff \frac{x-y}{2} = \sqrt{\left(\frac{b}{2}\right)^2 - c} \quad \text{skąd} \quad x = \frac{x+y}{2} + \frac{x-y}{2} = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

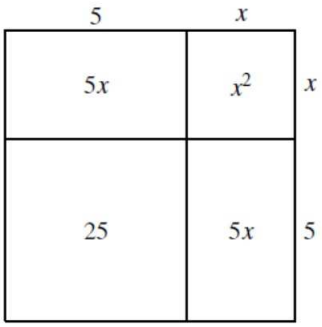
Stulecia później, arabscy matematycy cytować będą powyższy wynik uzasadniając własne algorytmiczne rozwiązanie równań kwadratowych. Sam Euklides ogólnych rozwiązań równań nie formułuje. Dowód jest czysto geometryczny. Idea *zamiany zmiennych* pochodząca z *przesuwania pól* pojawi się później.

O co więc chodzi z tym upraszczaniem i podejściem historycznym? Termin *algebra* pochodzi z tytułu arabskiego traktatu z początków IX wieku autorstwa perskiego uczonego Muhammada ibn Musa al-Chwarizmiego, bibliotekarza Bagdadzkiego *Domu Nauki*. Tytuł ten – *Zasady redukcji i przenoszenia* — ukrywa w sobie słowo *al-ğabr* oznaczające w języku arabskim *uzupełnianie*, *przenoszenie*. ale też *bilansowanie* (co miało praktyczne znaczenie). Dosłownie chodzi o operację przenoszenia wielkości z jednej strony na drugą, np. o zamianę równania $x^2 = 40x - 4x^2$ poprzez „al-ğabr" w równanie $5x^2 = 40x$. Przez redukcję (arab. *al-Muqabala*) autor rozumiał natomiast odjęcie tej samej dodatniej wielkości od obydwu stron równania, np. $x^2 + 5 = 40x + 4x^2$ zamieniamy na $5 = 40x + 3x^2$.

Ogólna metoda rozwiązywania równań kwadratowych, opracowana przez al-Khwarizmiego, prezentowana jest podobnie jak w matematyce greckiej w sposób geometryczny. Aby rozwiązać równanie

$$x^2 + 10x = 39$$

metodą arabską, przedstawiamy x^2 jako pole kwadratu o boku x, a $10x$ jako sumę pól dwóch prostokątów rozmiaru $5 \times x$. Dodatkowy kwadrat o polu 25 „uzupełnia" całą konfigurację do kwadratu o boku równym $x + 5$ o polu $25 + 39$, ponieważ 39 jest wartością wyrażenia $x^2 + 10x$. Stąd pole dużego kwadratu równe jest 64, czyli długość boku o długości $x + 5$ równa jest 8. Otrzymujemy zatem rozwiązanie $x = 3$.



Oczywiście matematyka arabska (ani grecka) nie uznawała liczb ujemnych, więc nie widziała też dodatkowego rozwiązania $x = -13$ tego równania. Ta konieczność unikania współczynników ujemnych w równaniach komplikowania rozważania algebraiczne. Nie było bowiem jednego ogólnego równania kwadratowego, ale aż trzy jego typy odpowiadające odpowiedniemu rozłożeniu dodatnich współczynników:

$$x^2 + ax = b, \quad x^2 = ax + b, \quad x^2 + b = ax.$$

Kluczowe jest jednak to, że owo zwijanie do kwadratu z algebraicznego punktu widzenia służy uproszczeniu równania przez podstawienie $t = x + 5$. Tego typu zamiana zmiennych przekształca równanie $x^2 + 10x = 39$ w $t^2 = 64$, a zatem w równanie prostsze. Podobną funkcję wzory skróconego mnożenia odgrywają w równaniach wyższego stopnia.

Przez kolejne stulecia także równania wyższych stopnia rozwiązywano przez podstawienie i eliminację. Cardano (powiemy o nim więcej za jakiś czas) wychodząc od podobnego geometrycznego ujęcia, od równania

$$x^3 + ax^2 + bx + c = 0$$

przechodzi za pomocą podstawienia

$$x = y - \frac{a}{3}$$

do równania typu

$$y^3 = py + q.$$

Dokonując kolejnej liniowej zamiany zmiennych postaci

$$y = u + v,$$

uzyskuje po lewej stronie:

$$(u^3 + v^3) + 3uv(u + v) = 3uvy + (u^3 + v^3),$$

które to wyrażenie równe jest prawej stronie wcześniejszego równania wtedy i tylko wtedy, gdy:

$$3uv = p,$$
$$u^3 + v^3 = q.$$

Eliminując v uzyskujemy równanie kwadratowe zmiennej u^3 postaci

$$u^3 + \left(\frac{p}{3u}\right)^3 = q,$$

o rozwiązaniach

$$\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}.$$

Rozumując symetrycznie, uzyskujemy te same wartości v^3. Skoro $u^3 + v^3 = q$, to jeden z pierwiastków równy jest u^3, a drugi — v^3. Bez straty ogólności można przyjąć

$$u^3 = \frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}, \quad v^3 = \frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3},$$

uzyskując

$$y = u + v = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}.$$

Jeżeli wyobrazimy sobie teraz, ze Cardano dokonywał tych manipulacji w języku geometrycznym, rozumiemy jak bardzo konieczne było przejście od geometrii do języka wyrażeń algebraicznych (zrobił to François Viète). A zatem ponownie — wzór skróconego mnożenia miał na celu dokonanie sensownego podstawienia i elimicjacji zmiennych. Podobne przykłady można wskazać dla równań wyższych stopni.

Zanim przejdziemy do układów równań, zajrzyjmy jeszcze do XVIII wieku i do wielkiego Eulera[10]. Wkrótce poznacie Państwo twierdzenie udowodnione przez Gaussa na przełomie XVIII i XIX wieku mówiące, że każdy wielomian o współczynnikach rzeczywistych można rozłożyć na czynniki liniowe lub kwadratowe. Kilkadziesiąt lat wcześniej wieku ten był otwarty, a niektórzy wątpili w jego prawdziwość, wśród nich sam Leibniz. Co gorsze, Nicolas Bernoulli twierdził, ze znalazł kontrprzykład, czyli wielomian

$$x^4 - 4x^3 + 2x^2 + 4x + 4.$$

Eulerowi udało się jednak znaleźć odpowiedni rozkład. W liście do Goldbacha z 1742 roku znalazł następujące czynniki kwadratowe wspomnianego wyżej wielomianu:

$$x^2 - \left(2 + \sqrt{4 + 2\sqrt{7}}\right) x + \left(1 + \sqrt{4 + 2\sqrt{7}} + \sqrt{7}\right),$$

$$x^2 - \left(2 - \sqrt{4 + 2\sqrt{7}}\right) x + \left(1 - \sqrt{4 + 2\sqrt{7}} + \sqrt{7}\right).$$

Fakt znalezienia tego rozkładu można chyba umieścić gdzieś pomiędzy cudem, a kuglarstwem. Euler był wielkim rachmistrzem. Właściwą intuicję daje wzór dwumianowy. Wspomniany wielomian przypomina bowiem rozwinięcie dwumianowe wyrażenia $(x-1)^4$. Dokładniej:

$$(x-1)^4 - (x^4 - 4x^3 + 2x^2 + 4x + 4) = 4x^2 - 8x - 3.$$

Pierwiastki uzyskanego wielomianu kwadratowego to $1 \pm \frac{\sqrt{7}}{2}$. Mamy więc:

$$x^4 - 4x^3 + 2x^2 + 4x + 4 = (x-1)^4 - 4\left((x-1) - \frac{\sqrt{7}}{2}\right)\left((x-1) + \frac{\sqrt{7}}{2}\right).$$

Podstawiając $u = (x-1)^2$, uzyskujemy

$$x^4 - 4x^3 + 2x^2 + 4x + 4 = u^2 - 4u + 7.$$

Tak dochodzimy do rezultatu i rozumiemy już, że chodziło w istocie o wymyślenie zwykłego podstawienia typu $x = t + 1$ sprowadzającego problem do kwestii rozłożenia na czynniki wielomianu

$$t^4 - 4t^2 + 7.$$

[10] O historii tej przeczytać można także w szerszym kontekście w książce *Euler. The Master of Us All* Wiliama Dunhama (Dolciani Mathematical Expositions Volume: 22; MAA1999).

Historia tak szerokiego tematu jak rozwiązywanie układów równań liniowych jest szalenie skomplikowana i w zasadzie nie jest możliwe choćby krótkie jej nakreślenie. Osobom zainteresowanym stosunkowo przystępnym tekstem polecam publicznie dostępny[11] artykuł *How ordinary elimination became Gaussian elimination* autorstwa Josepha Greara (Historia Mathematica 38 (2011), 163-218). Pozwolę sobie przytoczyć tu krótki wstęp, odwołując się również do wspomnianej we wstępie książki Stillwella, wybierając fragmenty przydatne dla ogólnego spojrzenia na matematykę. Niecierpliwym polecam tekst: https://www.fuw.edu.pl/ kostecki/histmat.pdf.

W istocie, ogólna teoria eliminacji dostała podwaliny w starożytnych Chinach w czasach dynastii Han (206 p.n.e. – 220 n.e.). Jest ona przedstawiona w słynnych Dziewięciu rozdziałach sztuki matematycznej, a jej ostateczna redakcja została dokonana przez Liu Hui w III wieku. W przeciwieństwie do greckiej matematyki, która budowała teorie wychodząc ze zbiorów aksjomatów, matematyka chińska szukała najbardziej ogólnych możliwych metod rozwiązywania zadań, niekoniecznie ubierając je w ogólne teorie.

Oczywiście nie pisano o układzie m równań z n niewiadomymi. Pokazano jednak na dostatecznej liczbie reprezentatywnych równań ogólną metodę odejmowania odpowiedniej wielokrotności niezerowego równania od pozostałych równań tak, by otrzymać układ w postaci trójkątnej (schodkowej). Ten typ rachunku bardzo odpowiadał urządzeniu upracowanemu do wykonywania kolejnych manipulacji na współczynnikach — analogicznych do tych, które wykonujemy na macierzach. Około XII wieku matematycy Chińscy opracowali podstawy metody wzmiankowanej w rozdziale o bazach Gröbnera, pozwalająca na eliminację zmiennej y z układu równań wielomianowych:

$$a_0(x)y^m + a_1(x)y^{m-1} + \ldots + a_m(x) = 0, \quad b_0(x)y^m + b_1(x)y^{m-1} + \ldots + b_m(x) = 0,$$

gdzie $a_i(x)$ oraz $b_j(x)$ są wielomianami zmiennej x. Oczywiście chodzi o przemnożenie pierwszego równania przez $b_0(x)$ i odjęcie drugiego przemnożonego przez $a_0(x)$. Uzyskujemy w ten sposób równanie stopnia $m-1$. Nietrudno w ten sposób zamienić wyjściowy układ na prostszy — zamiast wielomianów stopnia m o zmiennej y (o współczynnikach wielomianowych) mamy dwa równania stopnia $m-1$. W ten sposób możemy redukować (indukcyjnie) aż do uzyskania wielomianów zależnych tylko od x.

Jak się okazuje, opisana wyżej metoda odkryta została ponownie w XVII wieku na Zachodzie, przy okazji rozważania problemu znajdowania przecięć krzywych. Doprowadziło to do sformułowania metody eliminacji, którą niemal wiek później przeniesiono na grunt równań liniowych.

Jak to często bywa w historii matematyki, metodę eliminacji prowadzącą do rozwiązywania układów równań, którą tradycyjnie przypisuje się Gaussowi[12], sam Książe Matematyków uważał za dobrze znany folklor (podobnie myślał o wielu innych tematach, na przykład o kwaternionach) — wiemy już czemu (była ogólniejsza teoria). Idea eliminacji pojawiała się w opublikowanych wbrew woli Newtona notatkach do prowadzonego w latach 1673-1687 wykładu z algebry w Cambridge, wydanych uroczyście w 1707 roku jako *Arithmetica Universalis*. Newton uważał, że publikowanie takich tekstów po 20 latach (a dalej tłumaczenie ich na różne języki) może prowadzić kogoś do wniosku, że są to jego najnowsze wyniki! Tymczasem w Anglii wydano w latach 1650-1750 kilkadziesiąt podręczników do algebry. Klasyczna złośliwość historii sprawiła, że ów podręcznik do algebry stał się jednym z najszerzej znanych i wpływowych dzieł matematycznych Newtona. Wielki Euler krytykował metodę eliminacji jako „niepolecaną", jego następca Legendre nazywał ją „zwyczajną". Gauss potraktował eliminację jak znane wszystkim narzędzie.

Pojęcie macierzy zawdzięczamy Brytyjczykom: Arthurowi Cayleyowi i Jamesowi Josephowi Sylvesterowi. Drugi z nich użył tego pojęcia po raz pierwszy w 1850 roku. Pierwszy natomiast, siedem lat później, napisał pierwszą rozprawę o macierzach (*Treatise on the Theory of Matrices*). Jak się nietrudno domyśleć oznacza to, że eliminacja przypisywana wcześniejszym autorom odbyła się bez użycia pojęcia macierzy.

Kto zatem, i dlaczego, przypisał Gaussowi autorstwo metody eliminacji? Stało się to głównie za sprawą rozwoju maszyn liczących. Kluczowe było to, że algorytmy opisane w pracy Gaussa były po prostu niezwykle użyteczne dla wykorzystania praprzodków komputerów. Stosowano je, z nielicznymi modyfikacjami jeszcze do II Wojny Światowej. Notacja „klamrowa" wprowadzona przez Gaussa podkreślała kolejność zmiennych i była bardzo wygodna. Pojęcie „algorytmu eliminacji Gaussa" zastosował po raz pierwszy, jak się wydaje, Alan Turing, który przez dwa tygodnie rozwiązywał, z pomocą „komputera biurkowego" układ 18 równań w roku 1946. Po II WŚ, pojęcie to stopniowo wchodziło do programów nauczania.

[11] Zarówno na ArXiV: https://arxiv.org/pdf/0907.2397.pdf, jak i na stronach Elselviera.

[12] Gauss, C. F., *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium*, 1809.

Rozdział 2

Działania i ich własności. Ciała

2.1 Wykład drugi

Na ostatnim wykładzie rozważaliśmy układy równań liniowych o współczynnikach rzeczywistych. Są to ciągi równań postaci $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$, gdzie $a_1, a_2, \ldots, a_n, b \in \mathbb{R}$. Sprawdzenie, że $(s_1, \ldots, s_n)$ jest rozwiązaniem równania wyżej dokonywaliśmy przez wykonanie **działań** dodawania i mnożenia w $\mathbb{R}$ postaci $a_1 \cdot s_1 + a_2 \cdot s_2 + \ldots + a_n \cdot s_n$, żądając, by wynikiem było b. Na tym[1] wykładzie spróbujemy odpowiedzieć na pytanie: czy zamiast zbioru $\mathbb{R}$ z działaniami $+$ oraz $\cdot$ można rozważać układy równań liniowych, gdzie współczynnikami są **inne zbiory** X z **innymi działaniami** dodawania i mnożenia $\boxplus, \boxtimes$ tak, by metoda eliminacji Gaussa dalej działała? Na pewno nie mamy tu pełnej dowolności. Przecież choćby równanie liniowe $4x = 2$ nie ma rozwiązania w zbiorze liczb całkowitych.

Definicja 2.1: Działanie

Niech X będzie zbiorem niepustym. Przez X^n rozumieć będziemy zbiór ciągów postaci

$$(x_1, x_2, \ldots, x_n), \qquad \text{gdzie} \qquad x_i \in X, \text{ dla } 1 \leqslant i \leqslant n.$$

Działaniem n-argumentowym na zbiorze X nazywamy każdą funkcję $\omega : X^n \to X$.

Najczęściej rozważanymi działaniami są działania dwuargumentowe. Oto ich przykłady.

zbiór X	działanie ω
liczby rzeczywiste/wymierne/całkowite/naturalne	dodawanie/mnożenie
liczby rzeczywiste	$a \boxplus b = a + b + ab$
zbiór podzbiorów danego zbioru	suma/część wspólna
zbiór funkcji ze zbioru X na zbiór X	złożenie

Kluczowym elementem definicji działania jest żądanie, by nie wyprowadzało ono poza zbiór, na którym jest określone. A więc na przykład odejmowanie nie jest działaniem w zbiorze dodatnich liczb całkowitych, bo $1 - 3 \notin \mathbb{Z}_+$.

Definicja 2.2: Łączność i przemienność

Niech $*$ będzie działaniem dwuargumentowym na zbiorze X. Mówimy, że $*$ jest:

- łączne, jeśli dla każdych $a, b, c \in X$ mamy $(a * b) * c = a * (b * c)$,
- przemienne, jeśli dla każdych $a, b \in X$ mamy $a * b = b * a$.

Przyjrzyjmy się ponownie kilku przykładom.

- działanie $a \boxplus b = a^2 + b^2$ na zbiorze $\mathbb{R}$ nie jest łączne, bo $(1 \boxplus 2) \boxplus 3 = 34$, zaś $1 \boxplus (2 \boxplus 3) = 170$.
- złożenie w zbiorze bijekcji (czyli odwzorowań 1-1 i „na") zbioru $\mathbb{R}$ nie jest przemienne.

[1] Ostatnia aktualizacja: 11.10.2023 r.

Definicja 2.3: Ciało

Ciałem nazywamy piątkę $(K, \boxplus, \boxtimes, 0, 1)$, gdzie K jest zbiorem przynajmniej dwuelementowym z wyróżnionymi elementami $0 \neq 1$, zwanymi zerem i jedynką, zaś $\boxplus$, $\boxtimes$ są dwuargumentowymi działaniami zwanymi dodawaniem i mnożeniem, spełniającymi następujące aksjomaty ciała:

1)	$(a \boxplus b) \boxplus c = a \boxplus (b \boxplus c)$	$\forall_{a,b,c \in K}$	łączność dodawania
2)	$a \boxplus b = b \boxplus a$	$\forall_{a,b \in K}$	przemienność dodawania
3)	$a \boxplus 0 = a = 0 \boxplus a$	$\forall_{a \in K}$	własność elementu 0
4)	$a \boxplus b = 0 = b \boxplus a$	$\forall_{a \in K} \exists_{b \in K}$	element przeciwny
5)	$(a \boxtimes b) \boxtimes c = a \boxtimes (b \boxtimes c)$	$\forall_{a,b,c \in K}$	łączność mnożenia
6)	$a \boxtimes b = b \boxtimes a$	$\forall_{a,b \in K}$	przemienność mnożenia
7)	$a \boxtimes 1 = 1 \boxtimes a = a$	$\forall_{a \in K}$	własność elementu 1
8)	$a \boxtimes b = b \boxtimes a = 1$	$\forall_{a \in K \setminus \{0\}} \exists_{b \in K}$	odwrotność dla $a \neq 0$
9)	$a \boxtimes (b \boxplus c) = (a \boxtimes b) \boxplus (a \boxtimes c)$	$\forall_{a,b,c \in K}$	rozdzielność $\boxtimes$ wzgl. $\boxplus$

Teoria ciał jest bardzo szeroką dziedziną algebry abstrakcyjnej i w trakcie studiów będziecie Państwo poznawać różne nowe jej aspekty. W ramach naszego wykładu skupimy się na podstawowych przykładach.

- Piątka $(\mathbb{R}, +, \cdot, 0, 1)$, jest ciałem, gdzie $+, \cdot$ – standardowe dodawanie i mnożenie liczb rzeczywistych.
- Piątka $(\mathbb{Q}, +, \cdot, 0, 1)$ jest ciałem, gdzie $+, \cdot$ – standardowe dodawanie i mnożenie liczb wymiernych.
- Piątka $(\mathbb{Z}, +, \cdot, 0, 1)$ nie jest ciałem, bo żaden niezerowy element poza $-1, 1$ nie ma elementu odwrotnego. Warto odnotować, że wszystkie inne aksjomaty poza (8) są przez $(\mathbb{Z}, +, \cdot, 0, 1)$ spełnione[2]

Z aksjomatów ciała wyprowadzać można rozmaite ich własności. W tym miejscu przedstawimy dwie.

Obserwacja 2.1

Niech K będzie ciałem. Dla każdego elementu $a \in K$ istnieje dokładnie jeden element przeciwny do a. Dla każdego niezerowego elementu $b \in K$ istnieje dokładnie jeden element odwrotny do b.

Dowód. Wykażemy jedynie jednoznaczność elementu przeciwnego. Drugą część dowodzi się analogicznie. Załóżmy, że dla pewnych elementów x, x' ciała K dla dowolnego $a \in K$ mamy $a \boxplus x = 0 = a \boxplus x'$. Korzystamy dalej z aksjomatów ciała i powyższej równości:

$$\begin{aligned} x &= x \boxplus 0 && \text{(aksjomat 3 – wł. elementu 0)} \\ &= x \boxplus (a \boxplus x') && \text{(równość wyżej)} \\ &= (x \boxplus a) \boxplus x' && \text{(aksjomat 1 – łączność } \boxplus\text{)} \\ &= x' \boxplus (a \boxplus x) && \text{(aksjomat 2 – przemienność } \boxplus\text{)} \\ &= x' \boxplus 0 && \text{(równość wyżej)} \\ &= x'. && \text{(aksjomat 3 – wł. elementu 0)} \end{aligned}$$

□

Obserwacja 2.2

Niech K będzie ciałem. Wówczas jeśli $a, b \in K$ oraz $ab = 0$, to $a = 0$ lub $b = 0$.

Dowód. Niech $x, y \in K$. Wyprowadzimy kolejne wnioski z aksjomatów ciała.

(i) Jeśli $x + y = x$, to $-x + (x + y) = (-x + x) + y = 0 + y = y = -x + x = 0$.

(ii) Mamy $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$. Zatem na mocy (i) mamy $0 \cdot x = 0$.

Jeśli $a \neq 0$, to na mocy (ii) mamy $0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = b$. □

[2]Formalny dowód ma charakter wywodzący się z aksjomatyki liczb naturalnych i może być omówiony na wstępie do matematyki. Uzasadnienie przemienności dodawania liczb naturalnych wymaga zastosowania na przykład zasady indukcji.

W dalszym ciągu, o ile nie prowadzi to do nieporozumień, wprowadzamy następujące umowy: dodawanie i mnożenie w ciele K oznaczamy odpowiednio jako $+$ oraz $\cdot$, przy czym znak mnożenia może być pomijany. Przez a^n rozumiemy wynik n-krotnego przemnożenia przez siebie elementu a. Przyjmujemy też:

oznaczenie	definicja
$-a$	element odwrotny do a ze względu na $+$
a^{-1}	element odwrotny do a ze względu na $\cdot$
$a-b$	element postaci $a+(-b)$
$\frac{a}{b}$	element postaci $a\cdot(b^{-1})$

Wszystkie pojęcia zdefiniowane na poprzednim wykładzie dla układów równań o współczynnikach w $\mathbb{R}$ (równanie liniowe, układy równoważne, rozwiązania ogólne, macierz układu, operacje elementarne na wierszach, macierze w postaci schodkowej i zredukowanej) przenoszą się na układy o współczynnikach w dowolnym ciele K. Podobnie, opisana na poprzednim wykładzie metoda eliminacji Gaussa stosuje się do układów równań liniowych o współczynnikach w dowolnym ciele K. Mianowicie zachodzi twierdzenie:

Twierdzenie 2.1

Niech K będzie ciałem. Każdą macierz $A = M_{m\times n}(K)$ można za pomocą operacji elementarnych (1)-(2) na wierszach doprowadzić do postaci schodkowej. Każdą macierz $A = M_{m\times n}(K)$ można za pomocą operacji elementarnych (1)-(3) na wierszach doprowadzić do postaci zredukowanej. Każdy niesprzeczny układ równań liniowych o współczynnikach w K ma rozwiązanie ogólne. Aby je znaleźć wystarczy sprowadzić macierz tego układu do postaci schodkowej zredukowanej elementarnymi operacjami na wierszach, a następnie z otrzymanej macierzy otrzymać rozwiązanie ogólne.

Warto prześledzić dowody z poprzedniego wykładu. by zobaczyć w jaki sposób własności liczb rzeczywistych można w nich zastąpić przez kolejne aksjomaty ciała. Kluczowym elementem, który się wyłoni jest istnienie elementu przeciwnego oraz odwrotnego, niezbędnych m.in. do wykonania eliminacji Gaussa.

Celem tego wykładu nie jest systematyczny wykład teorii ciał (dotkniemy tego zagadnienia w dodatkach) ale omówienie najważniejszych przykładów ciał. Szczególnie dużo miejsca poświęcimy dwóm: ciału $\mathbb{Z}_p$ reszt z dzielenia przez liczbę pierwszą p oraz ciału $\mathbb{C}$ liczb zespolonych. Na końcu wykładu powiemy o wzajemnych związkach pomiędzy ciałami, pozwalającymi na budowanie większej liczby przykładów.

Twierdzenie 2.2

Niech p będzie liczbą pierwszą. Rozważmy zbiór reszt z dzielenia przez p:

$$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}.$$

Określamy na zbiorze $\mathbb{Z}_p$ działania dodawania $+_p$ i mnozenia $\cdot_p$ modulo p:

- $a +_p b$ to reszta z dzielenia przez p liczby $a+b$,
- $a \cdot_p b$ to reszta z dzielenia przez p liczby $a\cdot b$.

Wyróżniamy też elementy $0, 1$ w $\mathbb{Z}_p$. Wówczas $(\mathbb{Z}_p, +_p, \cdot_p, 0, 1)$ jest ciałem.

Przykłady ilustrujące powyżej zdefiniowane działania:

- $1 +_3 2 = 0$ w ciele $\mathbb{Z}_3$, więc 1 i 2 są elementami wzajemnie przeciwnymi w $\mathbb{Z}_3$,
- $2 \cdot_5 3 = 1$ w ciele $\mathbb{Z}_5$, a zatem w tym ciele $\frac{1}{2} = 3$.

Dowód Twierdzenia 2.2 jest żmudnym, ale rutynowym ćwiczeniem z elementarnej teorii liczb. Dotyczy to zwłaszcza sprawdzenia aksjomatów (1)-(7) oraz (9). Wynikają one z dwóch faktów:

- aksjomaty ciała (1)-(7) oraz (9) są spełnione przez zbiór liczb całkowitych ze standardowymi działaniami dodawania i mnożenia oraz elementami 0, 1,
- jeśli przez $[x]_n$ oznaczymy resztę z dzielenia przez n liczby całkowitej x, to dla dowolnych całkowitych a, b mamy:

$$[a]_n \ +_n \ [b]_n = [a+b]_n, \qquad [a]_n \ \cdot_n \ [b]_n = [ab]_n.$$

Istnienie elementu odwrotnego do każdego elementu niezerowego w $\mathbb{Z}_p$, czyli spełnianie aksjomatu (8), wymaga wykorzystania następującego naturalnego faktu.

Obserwacja 2.3

Niech p będzie liczbą pierwszą, zaś $a, b > 1$ niech będą liczbami całkowitymi. Wówczas jeśli p jest dzielnikiem ab, to p jest dzielnikiem a lub p jest dzielnikiem b

Fakt ten wynika z twierdzenia o istnieniu i jednoznaczności rozkładu liczby całkowitej (różnej od $-1, 0, 1$) na czynniki pierwsze. Nie jest to banalne twierdzenie, jeśli startujemy od szkolnej definicji liczby pierwszej, czyli liczby całkowitej większej od 1, której jedynymi dzielnikami są 1 i ona sama.

Twierdzenie 2.3: Istnienie odwrotności rezerowych elementów w $\mathbb{Z}_p$

Niech $0 \neq r \in \mathbb{Z}_p$. Wówczas istnieje $s \in \mathbb{Z}_p$, takie że $r \cdot_p s = 1$.

Dowód. Rozważmy zbiór wszystkich liczb postaci $\{0, r, 2r, 3r, \ldots, (p-1)r\}$. Liczy on p elementów. Zauważmy, że żadne dwa z tych elementów nie dają tej samej reszty z dzielenia przez p. Istotnie, gdyby liczby xr oraz yr dawały te same reszty z dzielenia przez p, dla pewnych $x, y \in \mathbb{Z}_p$, wówczas liczba $xr - yr = (x-y)r$ byłaby podzielna przez p. Mamy jednak $0 < r < p$, co oznacza, zgodnie z Obserwacją 2.3, że to liczba $x - y$ jest podzielna przez p. Jednak również $-p < x - y < p$, skąd $x - y = 0$. Widzimy zatem, że zbiór $\{0, r, 2r, 3r, \ldots, (p-1)r\}$ składa się z p liczb dających parami różne reszty z dzielenia przez p. Skoro jednak wszystkich możliwych reszt jest p, to istnieje element $s \in \mathbb{Z}_p$, taki że liczba rs daje resztę 1 z dzielenia przez p. Oznacza to, że s jest odwrotnością elementu r modulo p, czyli $r \cdot_p s = 1$. □

Zauważmy, że powyższy dowód nie mówi jak wskazywać odwrotność dla danego niezerowego elementu ciała $\mathbb{Z}_p$ — jest więc dowodem niekonstruktywnym (inaczej: egzystencjalnym). W uzupełnieniu powiemy jak wskazywać odwrotności, wykorzystując tzw. Lemat Bezout oraz algorytm Euklidesa.

Zobaczmy jak wyglądają tabelki ciał $\mathbb{Z}_2$ oraz $\mathbb{Z}_3$.

$+_2$	0	1
0	0	1
1	1	0

$\cdot_2$	0	1
0	0	0
1	0	1

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot_3$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Spróbujmy rozwiązać układ równań liniowych o współczynnikach w ciele $\mathbb{Z}_3$:

$$\begin{cases} x_1 + x_2 = 2 \\ 2x_1 + x_2 = 1 \end{cases} \quad \text{o macierzy} \quad \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 1 & 1 \end{array}\right] \in M_{3\times 2}(\mathbb{Z}_3).$$

Odejmujemy pierwszy wiersz przemnożony przez 2. Jaką on ma postać? Otóż jest to wiersz postaci $2\,2\,|1$, bo działania wykonujemy modulo 3. A zatem po tej operacji mamy (to samo, co po dodaniu wierszy):

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 2 & 0 \end{array}\right].$$

Cóż więc pozostaje? Przemnożyć drugi wiersz przez... odwrotność 2, czyli 2 (bo $2 \cdot_3 2 = 1$). A zatem mnożymy drugi wiersz przez 2 i odejmujemy od pierwszego. Dostajemy:

$$\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 0 \end{array}\right].$$

A zatem rozwiązaniem tego układu jest para $(x_1, x_2) = (2, 0)$. Ten układ był tak prosty, że to rozwiązanie byłoby widoczne i bez sprowadzania macierzy do postaci schodkowej. **Układ równań liniowych o współczynnikach w ciele skończonym ma zawsze skończenie wiele rozwiązań!** Rozwiązania ogólne mogą być jednak nadal reprezentowane przez zmienne zależne i niezależne. Np. zbiorem rozwiązań równania $x_1 + x_2 = 0$ o współczynnikach w ciele skończonym $\mathbb{Z}_p$ są wszystkie pary $\{(-t, t) \,|\, t \in \mathbb{Z}_p\}$, a więc równanie to ma p rozwiązań. Równanie $x_1 + x_2 + x_3 = 0$ ma p^2 rozwiązań w ciele $\mathbb{Z}_p$.

Przejdziemy teraz do drugiego fundamentalnego przykładu ciała, liczb zespolonych.

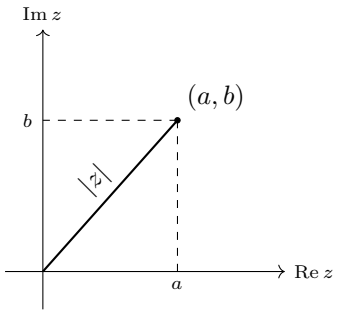
Definicja 2.4: Ciało liczb zespolonych

Ciało liczb zespolonych to piątka $(\mathbb{R}^2, \oplus, \otimes, (0,0), (1,0))$, oznaczana przez $\mathbb{C}$, którego elementami są wszystkie uporządkowane pary liczb rzeczywistych, i w którym działania $\oplus, \otimes$ określone są za pomocą działań $+$ oraz $\cdot$ w $\mathbb{R}$ wzorami:

$$(a,b) \oplus (c,d) = (a+c, b+d), \quad (a,b) \otimes (c,d) = (ac-bd, ad+bc).$$

Dla dowolnej liczby zespolonej $z = (a,b)$ wprowadzamy oznaczenia:

- a nazywamy częścią rzeczywistą liczby z i oznaczamy ją przez $\operatorname{Re} z$,
- b nazywamy częścią urojoną liczby z i oznaczamy $\operatorname{Im} z$,
- liczbę $\sqrt{a^2+b^2}$ nazywamy modułem liczby z i oznaczamy jako $|z|$.



Rys. 1. Płaszczyzna zespolona. Część rzeczywista, urojona oraz moduł liczby zespolonej.

Odwołamy się teraz do niezwykle ważnej, geometrycznej interpretacji. Liczby zespolone to pary punktów i możliwe jest reprezentowanie liczb zespolonych na tzw. płaszczyźnie zespolonej. Oś odpowiadającą części rzeczywistej liczby zespolonej oznaczamy $\operatorname{Re} z$, a oś odpowiadającą części urojonej oznaczamy jako $\operatorname{Im} z$. Moduł $|z|$ liczby zespolonej z intepretujemy jako odległość (euklidesową) punktu z od punktu $(0,0)$.

Zobaczmy kilka przykładowych działań w ciele $\mathbb{C}$.

- $(0,1) \oplus (1,0) = (1,1)$,
- $(2,1) \otimes (2,-1) = (2 \cdot 2 - 1 \cdot (-1), 2 \cdot (-1) + 1 \cdot 2) = (5,0)$.

Przyporządkowanie $a \mapsto (a,0)$ zadaje utożsamienie zbioru liczb rzeczywistych z podzbiorem zbioru liczb zespolonych złożonym ze wszystkich liczb postaci $(r,0)$. Przy tym utożsamieniu działania na liczbach rzeczywistych odpowiadają działaniom na ich odpowiednikach w zbiorze liczb zespolonych. Mamy bowiem:

$$(a,0) \oplus (a',0) = (a+a',0), \quad (a,0) \otimes (a',0) = (aa'-0, 0+0) = (aa',0).$$

W tym przyporządkowaniu:

- liczbę postaci $(a,0)$ będziemy zapisywać jako a, dla każdego $a \in \mathbb{R}$,
- liczbę $(0,1)$ oznaczać będziemy jako i.

Używając tych oznaczeń mamy na przykład:

$$(a,b) = (a,0) \oplus (0,b) = (a,0) \oplus (b,0) \otimes (0,1) = a + bi, \quad i^2 = (0,1) \otimes (0,1) = (-1,0) = -1.$$

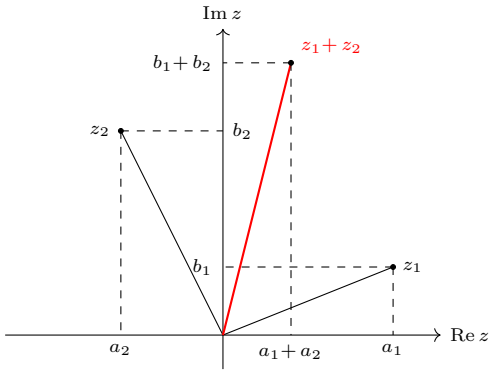
Widzimy więc, że liczbę $z = (a,b)$ zapisywać możemy w postaci ogólnej (algebraicznej)

$$z = a + bi,$$

przyjmując umowę, że jeśli $a + bi = c + di$, to $a = c$ oraz $b = d$.

W przyjętej konwencji dodawanie i mnożenie liczb zespolonych staje się bardziej zrozumiałe i pozwala na opuszczenie oznaczeń $\oplus, \otimes$. Zauważmy, że dodawanie liczb zespolonych $z_1 = a_1 + b_1 i$ oraz $z_2 = a_2 + b_2 i$ przypomina dodawanie wektorów. Mamy

$$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i.$$



Rys. 2. Interpretacja geometryczna dodawania liczb zespolonych.

Dodawanie i mnożenie liczb zespolonych w postaci ogólnej sprowadzają się do wykonywania operacji algebraicznych, uwzględniających zasadę: $i^2 = -1$, czyli np.

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Aby podać geometryczną interpretację mnożenia, wprowadzimy pojęcie argumentu liczby zespolonej.

Definicja 2.5: Postać trygonometryczna liczby zespolonej

Niech $z = a + bi \neq 0$ będzie liczbą zespoloną, zaś $\theta \in \mathbb{R}$ — miarą łukową kąta między półprostą o początku $(0, 0)$ przechodzącą przez $(1, 0)$, a półprostą o początku $(0, 0)$, przechodzącą przez z.

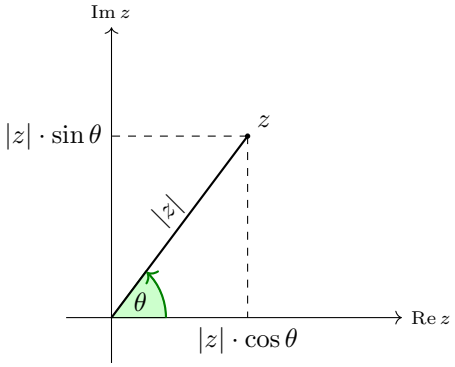
- Liczbę θ nazywamy ARGUMENTEM liczby z i oznaczamy przez $\arg z$. Argument liczby zespolonej $z \neq 0$ jest więc wyznaczony z dokładnością do całkowitej wielokrotności 2π. Liczbie 0 przypisujemy moduł 0 i dowolny argument $\theta \in \mathbb{R}$.
- Korzystając ze szkolnej definicji funkcji trygonometrycznych i z twierdzenia Pitagorasa mamy:

$$\cos\theta = \frac{a}{\sqrt{a^2 + b^2}} = \frac{a}{|z|}, \quad \sin\theta = \frac{b}{\sqrt{a^2 + b^2}} = \frac{b}{|z|}.$$

Stąd $a = |z| \cos\theta$ oraz $b = |z| \sin\theta$, a więc

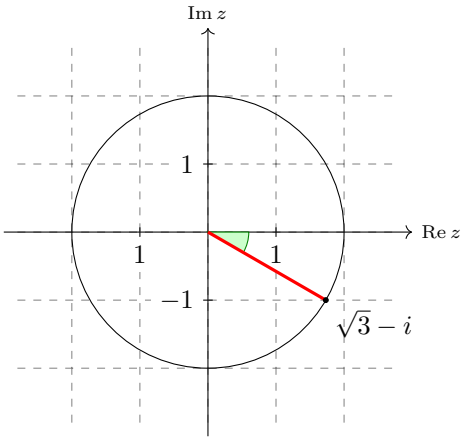
$$z = |z|(\cos\theta + i \cdot \sin\theta).$$

Jest to POSTAĆ TRYGONOMETRYCZNA LICZBY ZESPOLONEJ $z \neq 0$.



Rys. 3. Argument liczby zespolonej.

Przykład. Znajdziemy postać trygonometryczną liczby $z = \sqrt{3} - i$.



Rys. 4. Moduł i argument liczby $\sqrt{3} - i$.

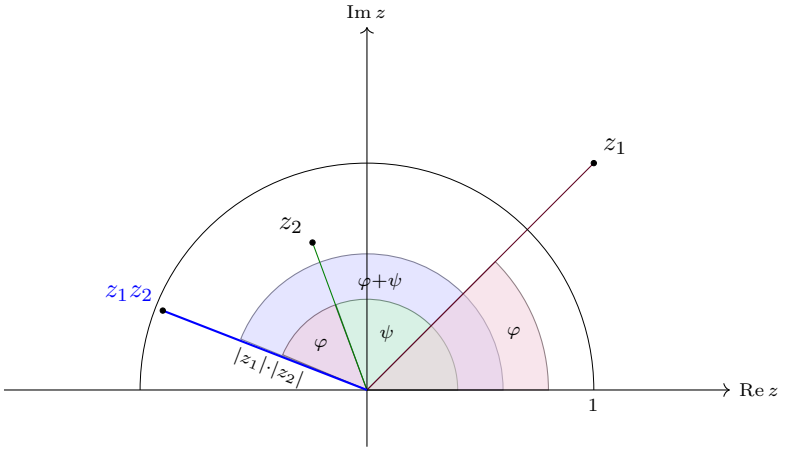
Mamy $|z| = \sqrt{\sqrt{3}^2 + (-1)^2} = 2$ oraz:

$$z = 2\left(\frac{\sqrt{3}}{2} - \frac{1}{2}\cdot i\right) = 2\left(\cos\frac{11\pi}{6} + i\cdot\sin\frac{11\pi}{6}\right) = 2\left(\cos\frac{-\pi}{6} + i\cdot\sin\frac{-\pi}{6}\right).$$

Wniosek 2.1

Dla niezerowych liczb zespolonych w, z zachodzą równości

$$\arg(z\cdot w) = \arg(z) + \arg(w), \quad |zw| = |z|\cdot|w|.$$



Rys. 5. Mnożenie liczb zespolonych z_1 i z_2 w interpretacji geometrycznej.

Dowód. Następujące obliczenie pokazuje jak zachowuje się argument przy mnożeniu liczb zespolonych z, w danych w postaciach trygonometrycznych:

$$\begin{aligned} z\cdot w &= |z|(\cos\varphi + i\sin\varphi)\cdot|w|(\cos\psi + i\sin\psi) = \\ &= |z||w|(\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\sin\varphi\cos\psi + \cos\varphi\sin\psi)) = \\ &= |z||w|(\cos(\varphi+\psi) + i\sin(\varphi+\psi)). \end{aligned}$$

Pokażmy jeszcze, że $|z|\cdot|w| = |zw|$. Niech $z = a + bi, w = c + di \in \mathbb{C}$. Wówczas

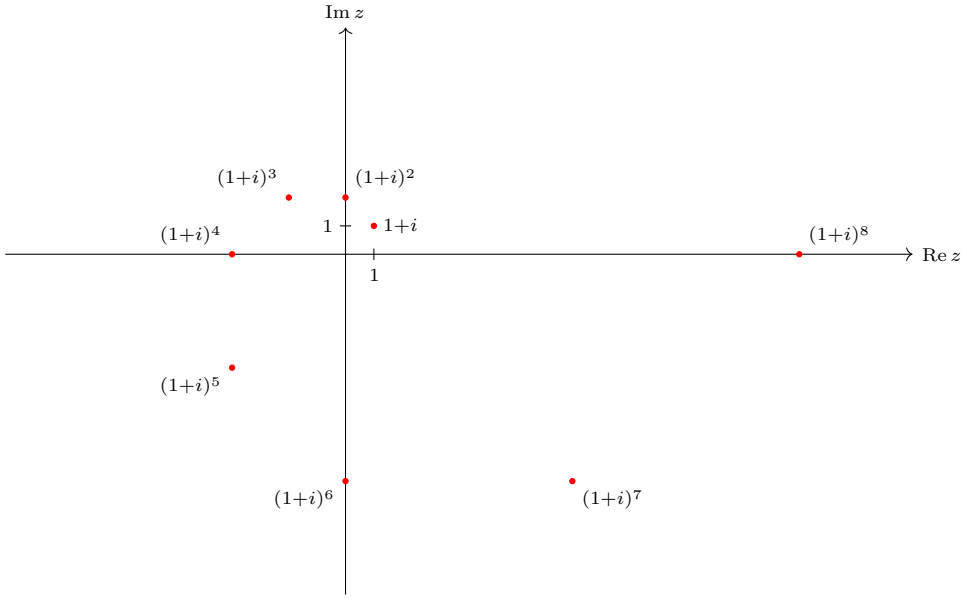
$$\begin{aligned} |z\cdot w| &= |(a+bi)(c+di)| = |(ac-bd) + (bc+ad)i| = \\ &= \sqrt{(ac-bd)^2 + (bc+ad)^2} = \sqrt{a^2c^2 + b^2d^2 + b^2c^2 + a^2d^2} = \\ &= \sqrt{(a^2+b^2)(c^2+d^2)} = |z|\cdot|w|. \end{aligned}$$

□

Wniosek 2.2: Wzór de Moivre'a, 1730

Niech $z = |z|(\cos\varphi + i \cdot \sin\varphi)$. Wówczas dla każdej dodatniej liczby całkowitej n mamy:

$$z^n = |z|^n(\cos(n \cdot \varphi) + i \cdot \sin(n \cdot \varphi)).$$



Rys. 6. Liczba $(1+i)^k$ ma argument $\frac{k\pi}{4}$ oraz moduł $\sqrt{2}^k$.

Nie uzasadniliśmy tymczasem, że $\mathbb{C}$ jest ciałem. To, że aksjomaty ciała są istotnie spełnione w zasadzie sprowadza się do manipulacji algebraicznych oraz własności liczb rzeczywistych (dowód łączności mnożenia czy rozdzielności). Odnotujmy tylko, że dla każdego $(a, b) \neq (0, 0)$ mamy:

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right).$$

W ramach naszych zajęć ważna będzie jeszcze jedna definicja, związana z liczbami zespolonymi.

Definicja 2.6: Sprzężenie liczby zespolonej

iech $z = a + bi$ będzie liczbą zespoloną. Sprzężeniem liczby zespolonej z nazywamy liczbę $a - bi$, oznaczaną przez $\overline{z}$. Na płaszczyźnie zespolonej punkt $\overline{z}$ jest obrazem z w symetrii względem osi $\mathrm{Re}(z)$. W szczególności $|z| = |\overline{z}|$ oraz $\arg(z) = -\arg(\overline{z})$.

Jako ćwiczenie pozostawiamy następującą obserwację, którą wykorzystamy na kolejnym wykładzie.

Obserwacja 2.4

Dla dowolnych $z, z_1, z_2 \in \mathbb{C}$:

(a) $z + \overline{z} = 2\,\mathrm{Re}(z)$ oraz $z \cdot \overline{z} = |z|^2$,

(b) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ oraz $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$,

(c) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$ oraz gdy $|z_2| \neq 0$ mamy też $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$,

(d) $|z_1 + z_2| \leqslant |z_1| + |z_2|$ oraz $||z_1| - |z_2|| \leqslant |z_1 - z_2|$.

2.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wykonywanie działań w ciałach reszt modulo p)
   Rozwiąż trzykrotnie równanie liniowe
   $$4x + 3 = 0$$
   zakładając, że współczynniki tego równania pochodzą kolejno z ciała: $\mathbb{Z}_5$, $\mathbb{Z}_7$ oraz $\mathbb{Z}_{13}$.

2. (♠ Rozwiązywanie układów równań liniowych o współczynnikach w $\mathbb{Z}_p$)
   Znajdź rozwiązanie ogólne następującego układu równań liniowych o współczynnikach w $\mathbb{Z}_5$:
   $$\begin{cases} 2x_1 + 3x_2 + x_3 + 4x_4 &= 1 \\ 3x_1 + x_2 + 2x_3 + 4x_4 &= 2 \\ 3x_1 + 3x_2 + x_3 + 3x_4 &= 1. \end{cases}$$

3. Niech p będzie liczbą pierwszą. Rozważmy układ równań o współczynnikach w ciele $\mathbb{Z}_p$ postaci
   $$\begin{cases} 5x + 3y = 4 \\ 3x + 6y = 1 \end{cases}$$
   Rozstrzygnij, dla jakich p układ ten nie ma rozwiązań/ma dokładnie jedno rozwiązanie/ma więcej niż jedno rozwiązanie?

4. (♠ Wykonywanie działań w ciele liczb zespolonych; część rzeczywista i urojona)
   Wyznacz część rzeczywistą i część urojoną liczb zespolonych:
   $$(2+i)(2-i)+(2+3i)(3+4i), \quad (3+i)^3-(3-i)^3, \quad \frac{(5+i)(3+5i)}{2i}, \quad \frac{(1+3i)(8-i)}{(2+i)^2}, \quad \frac{(1-i)^4-i}{(1+i)^4+i}.$$

5. (♠ Wykorzystanie jednoznaczności części rzeczywistej i urojonej)
   Znajdź wszystkie liczby zespolone z, które są rozwiązaniami równań (wyznacz $\mathrm{Re}(z)$ oraz $\mathrm{Im}(z)$):
   $$(1+i)z^2 + (3-5i)z - 6 = 0, \quad 2z + 3\overline{z} - \mathrm{Re}(z) + 2\,\mathrm{Im}(z) = 8 - 3i, \quad |z| + 3\overline{z} = 2 + 6i.$$

6. (♠ Rozwiązywanie układów równań liniowych o współczynnikach zespolonych)
   Znajdź rozwiązanie ogólne układu równań liniowych, którego macierz rozszerzona ma postać
   $$\left[\begin{array}{cccc|c} 1-i & i & 2 & -i & 1+i \\ 1+i & 1 & 2i & 1+2i & 1-i \\ i & 0 & -1+i & i & 0 \end{array}\right]$$

7. (♠ Wyznaczanie postaci trygonometrycznej i stosowanie wzoru Moivre'a)
   Wyznacz części rzeczywiste i urojone liczb zespolonych:
   $$(\sqrt{3}-i)^{32}, \quad (1+i\sqrt{3})^{150}, \quad \frac{(1+i\sqrt{3})^{27}}{(1-i)^{26}}, \quad \left(\frac{1-i\sqrt{3}}{1+i}\right)^{12}.$$

8. W zależności od $n \in \mathbb{N}$ oraz dla tych $x \in \mathbb{R}$, dla których to jest możliwe, wyznacz postać trygonometryczną liczby
   $$z = \frac{(1 + i\cos(x) + \sin(x))^n}{(1 - i\cos(x) + \sin(x))^n}.$$

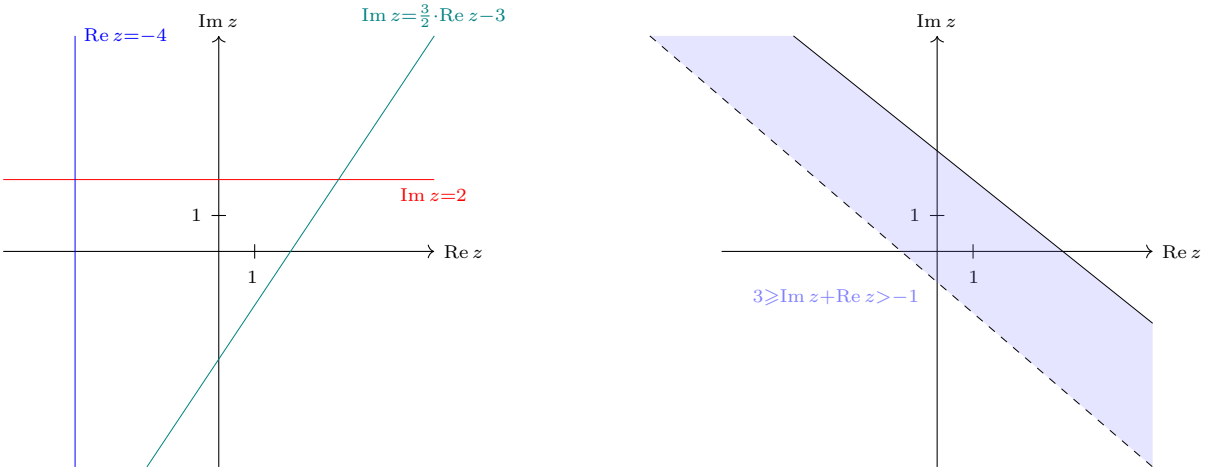
9. (♠ Szkicowanie prostych podzbiorów płaszczyzny zespolonej)
   Naszkicuj na płaszczyźnie zbiory liczb $z \in \mathbb{C}$ spełniających warunki:
   $$\begin{array}{cccc} \mathrm{Im}(iz) < 0, & \mathrm{Re}(1+i)z \geqslant 1, & \mathrm{Im}(1+i)z^2 < 0, & \mathrm{Im}(z^2) < 0, \\ \mathrm{Im}(iz^4+2) \geqslant 0, & \mathrm{Im}(z^3) < \mathrm{Re}(z^3), & |z+3-i| > 3, & 2 \leqslant |z| < |z-2| < 4. \end{array}$$

10. Liczba zespolona z spełnia warunek $|z| < 1$. Wykaż, że $|z^2 - z + i| < 3$.

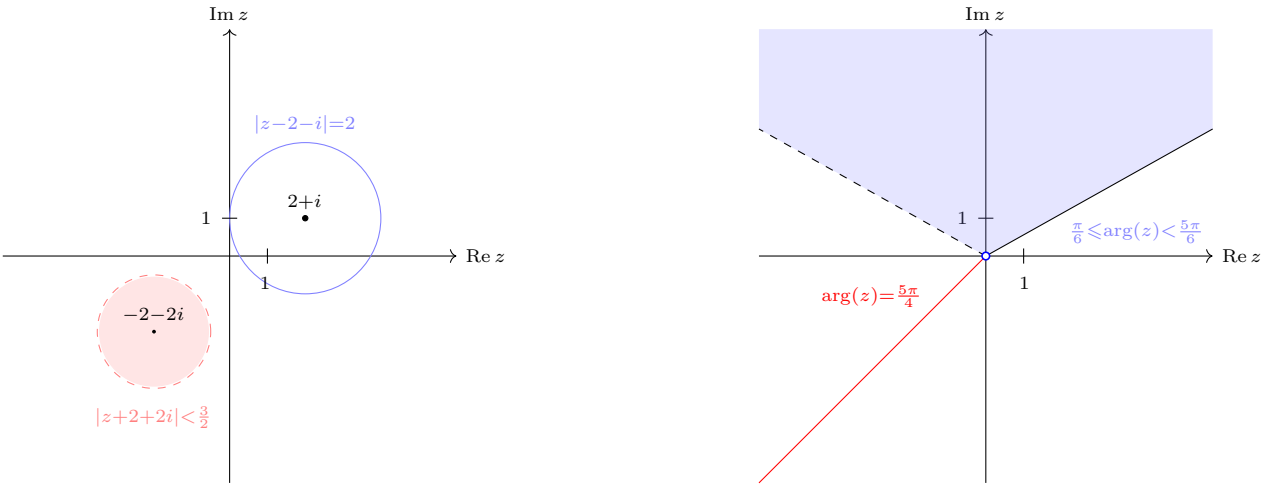
11. Liczby zespolone z_1, z_2 spełniają warunek $|z_1| = |z_2| = 1$. Wykaż, że
    $$\frac{z_1 + z_2}{1 + z_1 z_2} \in \mathbb{R}.$$

2.3 Uzupełnienie. Geometria płaszczyzny zespolonej

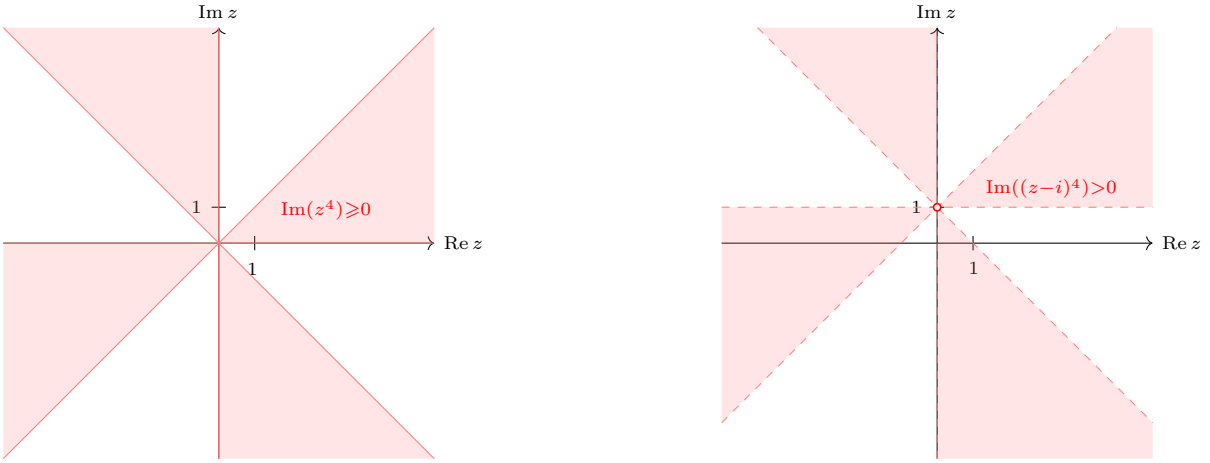
Ciało liczb zespolonych można rozumieć jako płaszczyznę z dodatkową strukturą algebraiczną. Struktura ta pozwala na wyrażenie w języku algebraicznym głębokich zależności geometrycznych. Zaczniemy od opisu zbiorów znanych z geometrii szkolnej.



Rys. 7. Każdy punkt z płaszczyzny zespolonej ma współrzędne $(\operatorname{Re} z, \operatorname{Im} z)$. Przechodząc do zapisu $z = x + yi$, współrzędne te wynoszą (x, y). Wyznaczanie prostych i zbiorów ograniczonych przez półpłaszczyzny wygląda więc podobnie jak w szkole.

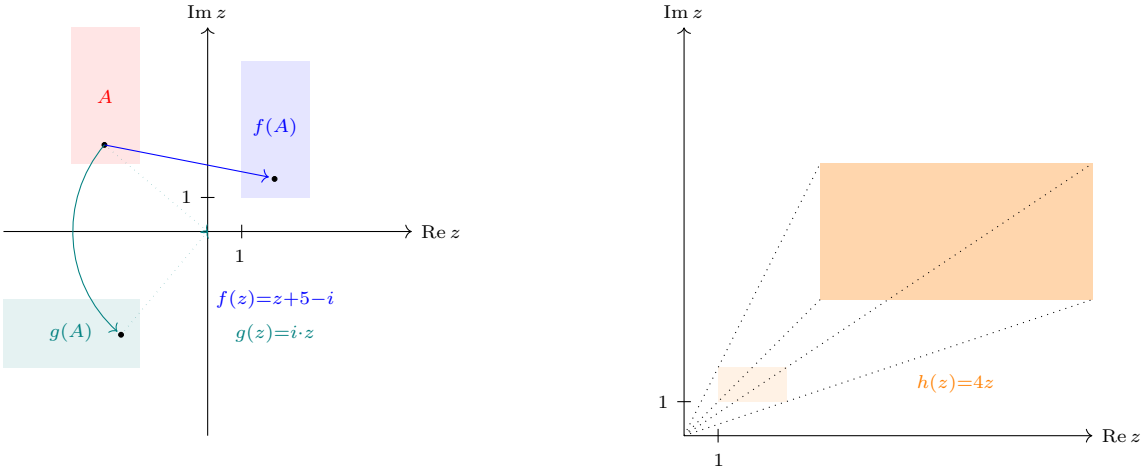


Rys. 8. Okrąg o środku w punkcie $z = a + bi$ i promieniu r opisany jest równaniem $|z - a - bi| = r$.
Półprosta bez początku w punkcie $z = 0$ i o nachyleniu θ złożona jest z punktów z spełniających warunek $\arg z = \theta$.



Rys. 9. Zbiory opisane wyżej wyznaczamy zgodnie z postacią trygonometryczną. Jeśli $z = |z|(\cos\theta + i\sin\theta)$ to $z^4 = |z|^4(\cos 4\theta + i\sin 4\theta)$. Skoro więc $\arg(z^4) \in [0, \pi]$, to $\arg(z) \in [0, \frac{\pi}{4}] \cup [\frac{\pi}{2}, \frac{3\pi}{4}] \cup [\pi, \frac{5\pi}{4}] \cup [\frac{3\pi}{2}, \frac{7\pi}{4}]$.

W prostych funkcjach zespolonych rozpoznać można izometrie i podobieństwa płaszczyzny.

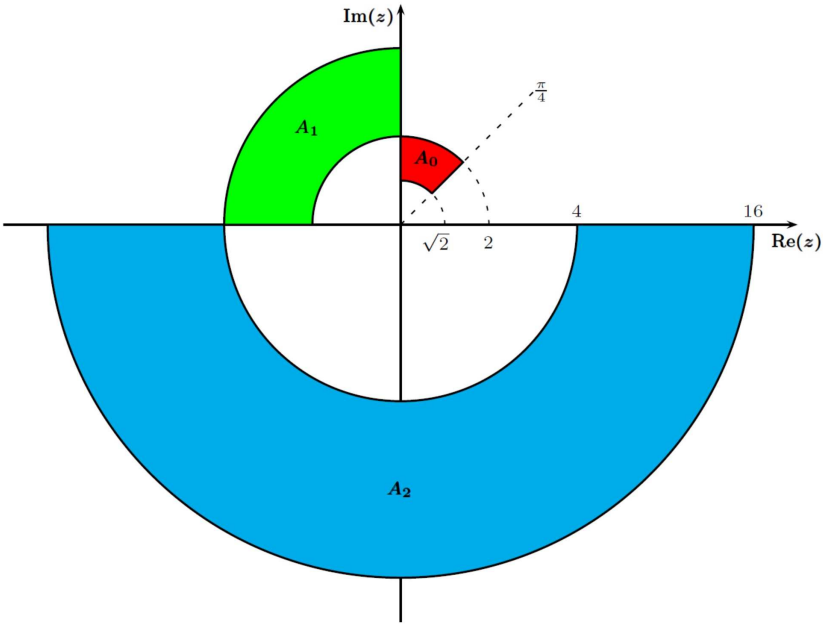


Rys. 10. Przesunięcie równoległe o wektor (a, b) opisane jest funkcją $z \mapsto z + a + bi$. Obrót wokół zera o kąt θ opisany jest wzorem $z \mapsto (cos\theta + i \cdot \sin\theta) \cdot z$. Jednokładność o środku 0 i skali $t > 0$ można otrzymać poprzez przekształcenie $z \mapsto t \cdot z$.

* * *

Jeszcze inną sytuację opisuje zadanie z kolokwium z roku 2019. Definiujmy podzbiory $\mathbb{C}$ postaci:

$$A_0 = \{z \in \mathbb{C} : \sqrt{2} \leqslant |z| \leqslant 2,\ \tfrac{\pi}{4} \leqslant \arg(z) \leqslant \tfrac{\pi}{2}\},$$
$$A_i = \{zz' :\ z, z' \in A_{i-1}\} \quad \text{dla } i > 0.$$



Rys. 11. Zbiory A_0, A_1, A_2, przy czym dla czytelności zastosowano skalę logarytmiczną.

Dla $i > 1$ jeśli zbiór A_{i-1} jest zbiorem takich $z \in \mathbb{C}$, że $r_1 \leqslant |z| \leqslant r_2$ oraz $\phi_1 \leqslant \arg(z) \leqslant \phi_2$, to A_i jest zbiorem takich $z \in \mathbb{C}$, że $r_1^2 \leqslant |z| \leqslant r_2^2$ oraz $2\phi_1 \leqslant \arg(z) \leqslant 2\phi_2$. W szczególności

$$A_1 = \{z \in \mathbb{C} :\ 2 \leqslant |z| \leqslant 4, \pi/2 \leqslant \arg(z) \leqslant \pi\} \quad \text{oraz} \quad A_2 = \{z \in \mathbb{C} :\ 4 \leqslant |z| \leqslant 16, \pi \leqslant \arg(z) \leqslant 2\pi\}.$$

* * *

Liczby zespolone można z powodzeniem wykorzystywać w rozwiązywaniu nietrywialnych zadań z geometrii elementarnej, w tym zadań olimpijskich. Zainteresowanego Czytelnika odsyłam choćby do

- J. Jaszuńska, *Liczby zespolone w geometrii*, Delta 11/2010, `https://www.deltami.edu.pl/temat/matematyka/geometria/planimetria/2010/11/29/0905k25.pdf`
- E. Chen, *Bashing Geometry with Complex Numbers*, `https://web.evanchen.cc/handouts/cmplx/en-cmplx.pdf`.

2.4 Uzupełnienie. Arytmetyka modularna i ciała skończone

Na wykładzie poznaliśmy ciała $\mathbb{Z}_p$ reszt z dzielenia przez p, gdzie p są liczbami pierwszymi. Pierwszą część uzupełnienia poświęcimy bardziej systematycznemu, choć elementarnemu wprowadzeniu do arytmetyki „modulo n" (modularnej), uzupełniając brakujące aspekty twierdzenia mówiącego, że $\mathbb{Z}_p$ jest ciałem. W drugiej części uzupełnienia przyjrzymy się problemowi istnienia innych ciał skończonych.

Aby prowadzić formalne rozumowania dotyczące liczb naturalnych i całkowitych należałoby cofnąć się aż do ich aksjomatyki, czego tu nie robimy. Za intuicyjnie jasne przyjmujemy następujące stwierdzenie (będące wnioskiem z aksjomatów Peano liczb naturalnych, o czym mówi się na wstępie do matematyki).

Obserwacja 2.5: Zasada minimum

W każdym niepustym podzbiorze złożonym z nieujemnych liczb całkowitych (naturalnych, ozn. $\mathbb{N}$) znajduje się element najmniejszy.

Natychmiastowym wnioskiem z zasady minimum jest następująca obserwacja.

Twierdzenie 2.4: O dzieleniu z resztą

Niech a, b będą dodatnimi liczbami całkowitymi, przy czym $b \geqslant a$ oraz $a \neq 0$. Wówczas istnieją liczby q, r, przy czym $q \geqslant 1$ oraz $0 \leqslant r < a$ takie, że:

$$b = qa + r.$$

Liczbę r nazywamy resztą z dzielenia b przez a, ozn. $r = [b]_a$.

Dowód. Rozważmy zbiór $A = \{b - na \,|\, n \in \mathbb{N}\} \cap \mathbb{N}$, czyli zbiór złożony z tych liczb naturalnych postaci:

$$b - a, \quad b - 2a, \quad b - 3a, \quad \dots$$

które są nieujemne. Skoro $b \geqslant a$, to zbiór A jest niepusty i posiada element najmniejszy. Nazwijmy ten element $b - ta$, dla pewnego $t \in \mathbb{N}$. Twierdzimy, że $b - ta < a$. W przeciwnym bowiem razie mielibyśmy

$$b - ta - a \geqslant 0 \quad \Longrightarrow \quad b - (t+1)a \geqslant 0.$$

To jednak przeczyłoby wyborowi t jako takiej liczby, dla której $b - ta$ jest najmniejszym elementem A. W rezultacie biorąc za q liczbę t oraz za r liczbę $b - ta$ dostajemy

$$b = qa + r = ta + (b - ta),$$

gdzie $0 \leqslant b - ta < a$ oraz $q = t \geqslant 1$. □

Powyższe twierdzenie pokazuje, że reszta z dzielenia dodatniej liczby całkowitej przez dodatnią liczbę całkowitą jest dobrze określona. Bez trudu można tą definicję przedłużyć na liczby ujemne. Dla przykładu, resztą z dzielenia -1 przez 3 jest 2, bo $-1 = (-1) \cdot 3 + 2$. Czasem mówi się też o ujemnych resztach z dzielenia przez n (np. w tzw. twierdzeniu Wilsona), mając jednak zawsze na myśli elementy przeciwne do odpowiednich elementów w $\mathbb{Z}_n$. I tak np. „reszta -2" z dzielenia przez 5 to po prostu reszta 3.

Definicja 2.7: Relacja przystawania (kongruencji) liczb całkowitych

Mówimy, że liczby całkowite a, b PRZYSTAJĄ MODULO m, co oznaczamy jako $a \equiv b \,(\mathrm{mod}\ m)$, jeśli a oraz b dają tą samą resztę z dzielenia przez m. Równoważnie – zachodzi warunek $[a]_m = [b]_m$.

Przystawanie modulo m sformułować można w języku podzielności. Przypomnijmy, że liczba całkowita m jest dzielnikiem liczby całkowitej n, jeśli istnieje liczba całkowita d taka, że $n = d \cdot m$, ozn. $m \,|\, n$.

Twierdzenie 2.5

Niech k będzie dodatnią liczbą całkowitą oraz niech n, m będą liczbami całkowitymi. Niech r, s będą odpowiednio resztami z dzielenia liczby n oraz m przez k, tzn. $r = n]_k, s = [m]_k$. Wówczas:

(i) liczba $n - m$ jest podzielna przez k wtedy i tylko wtedy, gdy reszty r oraz s są równe, tzn.

$$k \,|\, n - m \iff [n]_k = [m]_k,$$

(ii) reszta z dzielenia liczby $n + m$ przez k jest taka sama jak reszta z dzielenia liczby $r + s$ przez k, tzn.

$$[n + m]_k = [[n]_k + [m_k]]_k,$$

(iii) reszta z dzielenia liczby $n \cdot m$ przez k jest taka sama, jak reszta z dzielenia liczby $r \cdot s$ przez k, tzn.

$$[n \cdot m]_k = [[n]_k \cdot [m_k]]_k.$$

Rozwiązanie. Dowód części (i) wymaga skorzystania z twierdzenia o dzieleniu z resztą. Zacznijmy od założenia, że $r = s$. Wówczas:

$$n - m = xk + r - (yk + r) = (x - y)k,$$

a zatem $n - m$ jest wielokrotnością liczby k.

Odwrotnie, załóżmy, że $n - m$ jest liczbą podzielną przez k. Załóżmy najpierw, że $r \geqslant s$. Wówczas

$$n - m = (x - y)k + r - s,$$

gdzie $k > r - s \geqslant 0$. Liczba $r - s$ jest zatem resztą z dzielenia $n - m$ przez k. Skoro jednak $n - m$ jest liczbą podzielną przez k, mamy $r - s = 0$, czyli $r = s$.

Załóżmy teraz, że $r < s$. Skoro $n - m$ jest liczbą podzielną przez k, to liczba $m - n$ jest również podzielna przez k. Dalej rozumujemy analogicznie jak w poprzednim punkcie.

Dowodzimy punkty (ii), (iii). Mamy:

$$n + m = (x + y)k + r + s, \quad n \cdot m = (xk + r)(yk + s) = (xy + x + y)k + rs.$$

Widzimy, że $n + m - r - s$ oraz $nm - rs$ są podzielne przez k. Korzystając z (i) dostajemy tezę. ■

Wniosek 2.3

Warunki $a \equiv b \,(\mathrm{mod}\ m)$ oraz $b \equiv c \,(\mathrm{mod}\ m)$ implikują warunek $a \equiv c \,(\mathrm{mod}\ m)$.

Wniosek 2.4

Warunki $a \equiv b \,(\mathrm{mod}\ m)$ oraz $c \equiv d \,(\mathrm{mod}\ m)$ implikują

$$a + c \equiv b + d \,(\mathrm{mod}\ m), \quad ac \equiv bd \,(\mathrm{mod}\ m).$$

Uzyskaliśmy zatem również uzasadnienie postulowanych na wykładzie równości

$$[a]_n \ +_n \ [b]_n = [a + b]_n, \quad [a]_n \ \cdot_n \ [b]_n = [ab]_n.$$

W ten sposób to, że zbiór $\mathbb{Z}_n$ z działaniami dodawania i mnożenia modulo n oraz wyróżnionymi elementami 0, 1 spełnia warunki (1)-(7) oraz (9) w definicji ciała sprowadziliśmy do analogicznego stwierdzenia dla liczb całkowitych $\mathbb{Z}$ ze zwykłymi działaniami dodawania, mnożenia oraz wyróżnionymi elementami 0, 1. Tego, że $\mathbb{Z}$ posiada te własności nie dowodzimy.

Pozostał nam dowód lematu Bezout, mówiącego, że dla niezerowej liczby całkowitej a oraz dowolnej liczby całkowitej b istnieją takie liczby całkowite x, y, że $ax + by = NWD(a, b)$. Aby ten fakt uzasadnić, przypomnijmy definicję największego wspólnego dzielnika.

Definicja 2.8: Największy wspólny dzielnik

Niech a, b liczbami całkowitymi, z których co najmniej jedna jest niezerowa. Przez $NWD(a, b)$ oznaczamy największy wspólny dzielnik liczb a, b, czyli taką liczbę całkowitą n, że:

- n jest dzielnikiem zarówno a, jak i b,
- jeśli liczba m jest również dzielnikiem zarówno a, jak i b, to $m \leqslant n$.

Dowodzimy Lemat Bezout, ograniczając się do przypadku $a, b > 0$. Rozważmy zbiór L wszystkich liczb postaci

$$ax + by,$$

gdzie x, y są liczbami całkowitymi oraz $ax + by > 0$. Zbiór jest niepusty (zawiera choćby $a + b$). Zgodnie z zasadą minimum istnieje zatem **najmniejsza dodatnia** liczba postaci $ax+by$. Nazwijmy tą liczbę d.

Twierdzimy, że $d = NWD(a, b)$. Oczywiście $NWD(a, b)$ jest dzielnikiem d. Dzielnik jest nie większy niż liczba dodatnia, którą dzielimy, a zatem $NWD(a, b) \leqslant d$. Jeżeli pokażemy, że d jest zarówno dzielnikiem a, jak i b, to dowód będzie zakończony. Wykorzystamy założenie, że d jest najmniejszym elementem zbioru L.

Załóżmy, wbrew temu co oczekujemy, że d nie jest dzielnikiem a. Zatem na mocy twierdzenia o dzieleniu z resztą istnieje liczba $0 < r < d$ oraz $k \geqslant 1$ taka, że $a = kd + r$. To oznacza, że $r = a - kd$, co jest niemożliwe, bo przecież:

$$r = a - kd = a - k(ax + by) = a(1 - lkx) - kby,$$

jest również elementem zbioru L, i to mniejszym niż d, sprzeczność. A zatem d jest dzielnikiem a. Analogicznie pokazujemy, że d jest dzielnikiem b. A zatem d rzeczywiście jest wspólnym dzielnikiem a oraz b, co oznacza, że $d \leqslant NWD(a, b)$. Dowód jest zakończony.

Wykazaliśmy Lemat Bezout. Warto zaznaczyć, że bezpośrednim wnioskiem z tego rezultatu jest możliwość wyznaczania odwrotności niezerowego r w $\mathbb{Z}_p$. Istotnie, $NWD(p, r) = 1$, a zatem istnieją liczby całkowite m, s, że

$$mp + rs = 1.$$

W szczególności rs daje resztę 1 z dzielenia przez p, czyli reszta z dzielenia s przez p jest odwrotnością r w ciele $\mathbb{Z}_p$. A jak wyznaczyć to s, mając zadane p oraz r? Można to zrobić algorytmicznie za pomocą algorytmu Euklidesa. Podstawą tego algorytmu jest następująca obserwacja.

Obserwacja 2.6

Niech a, b będą niezerowymi liczbami całkowitymi i niech $r > 0$ będzie resztą z dzielenia a przez b. Wówczas:

$$NWD(a, b) = NWD(b, r).$$

Dowód. Liczba r jest postaci $a - kb$, czyli $NWD(a, b)$ jest dzielnikiem b oraz r, a więc jest też dzielnikiem $NWD(b, r)$. W szczególności $NWD(a, b) \leqslant NWD(b, r)$.

Z drugiej strony $a = kb+r$, więc $NWD(b, r)$ jest dzielnikiem $NWD(a, b)$. Zatem $NWD(b, r) \leqslant NWD(a, b)$. □

Przykład zastosowania:

$$NWD(391, 323) = NWD(323, 68) = NWD(68, 51) = NWD(51, 17) = 17.$$

Mamy bowiem $391 = 1 \cdot 323 + 68$, $323 = 4 \cdot 68 + 51$, $68 = 1 \cdot 51 + 17$ i wreszcie $51 = 4 \cdot 17$.

Oto natomiast sam algorytm Euklidesa.

Obserwacja 2.7

Niech a, b będą niezerowymi liczbami całkowitymi. Rozważmy ciągi liczb całkowitych $q_0, q_1, q_2, q_3, \ldots$ oraz $r_!, r_2, r_3, \ldots$ spełniający warunki:

$$\begin{aligned} a &= b \cdot q_0 + r_1, \quad 0 \leqslant r_1 < b \\ b &= r_1 \cdot q_1 + r_2, \quad 0 \leqslant r_2 < r_1 \\ r_1 &= r_2 \cdot q_2 + r_3, \quad 0 \leqslant r_3 < r_2 \\ r_2 &= r_3 \cdot q_3 + r_4, \quad 0 \leqslant r_4 < r_3, \\ &\ldots \end{aligned}$$

Niech n będzie taką liczbą, że $r_n \neq 0$ oraz $r_{n+1} = 0$. Wówczas $NWD(a, b) = r_n$.

Dowód. Teza wynika natychmiast z poprzedniej obserwacji. Mamy:

$$NWD(a, b) = NWD(b, r_1) = NWD(r_1, r_2) = \ldots = NWD(r_{n-1}, r_n).$$

Liczba $r_n \neq 0$ jest dzielnikiem r_{n-1} (bo $r_{n+1} = 0$), czyli oczywiście $NWD(r_{n-1}, r_n) = r_n$. □

Prawdziwość algorytmu Euklidesa daje nam możliwość przedstawiania $NWD(a, b)$ jako kombinacji liniowej dodatnich liczb a, b, a więc stanowi realizację lematu Bezout. Rzeczywiście, jeśli $x = qy + r$, dla pewnych liczb całkowitych x, y, q, r, to r jest kombinacją liniową liczb x, y. A zatem r_1 jest kombinacją liniową a, b, dalej r_2 jest kombinacją liniową a, r_1, czyli w istocie jest kombinacją liniową liczb a oraz kombinacji liniowej a oraz b. A zatem r_2 również jest kombinacją liniową a, b, itd.

Przykład. Przedstawimy w postaci kombinacji liniowej liczb 391 oraz 323 liczbę 17.

$$17 = 68 - 1 \cdot 51 = 68 - (323 - 4 \cdot 68) = -323 + 5 \cdot 68 = -323 + 5 \cdot (391 - 323) = 5 \cdot 391 + (-6) \cdot 323.$$

Ostatnia obserwacja pozwala nam widzieć, że znając p oraz r możemy wyznaczyć n oraz s spełniające $np + rs = 1$, postępując jak wyżej. A zatem nie tylko umiemy udowodnić istnienie odwrotności modulo p, ale też umiemy ją wyznaczać. Nie jest to, jak widać, zupełnie oczywiste. Warto również powiedzieć, że Obserwacja 2.2 jest bezpośrednim wnioskiem z Lematu Bezout. Zachęcam do dowodu tego faktu.

* * *

Tematyka ciał o skończenie wielu elementach (skończonych) nie zamyka się jedynie w ciałach reszt modulo p. Aby się o tym przekonać pokażemy, że istnieje ciało czteroelementowe o elementach $\{0, 1, a, b\}$. Spróbujemy to jednak zrobić w taki sposób, by dyskusja miała możliwie uniwersalny charakter i pozwalała wprowadzać nowe pojęcia. Po przedstawieniu konstrukcji sformułujemy bez dowodów kilka ogólnych wyników. Powiedzmy jeszcze tyle, że ciała skończone są niezwykle istotnymi obiektami, zarówno w samej matematyce, jak i w zastosowaniach, zwłaszcza w kryptografii i teorii kodów.

Zaczniemy zupełnie naiwnie. Mamy zbiór n elementowy i chcemy na nim określać jakieś dwuargumentowe działania. Ile jest tych działań? Które są łączne? Które są przemienne? Jak wygodnie określać te działania? Spójrzmy na tabelkę pewnego działania dwuargumentowego $*$ na zbiorze $\{a, b, c\}$:

$*$	a	b	c
a	b	c	b
b	a	c	b
c	c	a	a

Oczywiście działanie to nie jest przemienne – tabelka musiałaby mieć określoną symetrię. Mamy na przykład $a * b = c$ oraz $b * a = a$. Trudniej dostrzec brak łączności, ale po chwili widać, że $a * (b * c) = a * b = c$, podczas gdy $(a * b) * c = c * c = a$. Warto powiedzieć, że istnieje stosunkowo efektywna procedura sprawdzania czy tabelka danego działania dwuargumentowego opisuje działanie łączne, zwana testem łączności Lighta z 1949 roku (można poczytać hasło *Light's associativity test* na Wikipedii).

Nietrudno policzyć liczbę działań dwuargumentowych na zbiorze n elementowym, a także liczbę działań przemiennych, czy działań, w których jest element neutralny. Liczby te równe są odpowiednio:

$$n^{(n^2)}, \quad n^{\frac{n(n+1)}{2}}, \quad n^{(n-1)^2+1}.$$

Osobom, które byłyby zainteresowane tabelkami działań i zliczaniem niedużych tabelek o określonych własnościach polecam całkowicie elementarny tekst *Associative Operations on a Three-Element Set* autorstwa F. Diego oraz K. Jónsdóttir (dostępny on-line). Na koniec tego wstępu powiem tylko, że parę $(X, \circ)$, gdzie $\circ$ jest łącznym działaniem dwuargumentowym nazywamy **półgrupą**. Półgrup jest bardzo dużo. Istotnie różnych (nie chcę się tu wgłębiać w to, co to znaczy) półgrup 3-elementowych jest 18. Półgrup czteroelementowych, istotnie różnych, jest już 126, a pięcioelementowych – 1160 (patrz https://oeis.org/A001423).

W przypadku zliczania ciał skończonych potężnym narzędziem jest żądanie łączności aż dwóch działań dwuargumentowych, związanych prawem rozdzielności. Podstawowe obserwacje niezbędne do dalszej pracy zebrane są w następującym stwierdzeniu (dowód jest analogiczny do przedstawionych na wykładzie).

Obserwacja 2.8

Niech K będzie ciałem.

- Dla każdych $x, y, z \in K$ równość $x + y = x + z$ implikuje $y = z$.
- Dla $x \neq 0$ oraz $y, z \in K$ równość $xy = xz$ implikuje $y = z$.

W języku algebraicznym mówimy, że dodawanie i mnożenie w ciele są „skracalne". Wniosek z tych obserwacji jest następujący: w tabelkach opisujących działania w ciele skończonym K w każdym wierszu i w każdej kolumnie występować muszą wszystkie elementy z ciała – każdy dokładnie raz. Jeśli wiemy dodatkowo, że dla każdego $a \in K$ zachodzą równości:

$$a + 0 = a, \quad a \cdot 0 = 0, \quad a \cdot 1 = 1,$$

to wnioskujemy, że jest jedynie jedno ciało dwuelementowe i trzyelementowe – nie da się bowiem wypełnić tabelek dodawania i mnożenia ciał dwu i trzyelementowych inaczej, niż wypełnione są tabelki przedstawione w zasadniczej części wykładu. W przypadku tabelki ciała czteroelementowego napotkamy jednak na pewien problem. Oto tabelki działań w tym ciele, wypełnione zgodnie z wiedzą przedstawioną wyżej.

+	0	1	a	b
0	0	1	a	b
1	1			
a	a			
b	b			

·	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a		
b	0	b		

W rzeczywistości, tabelkę mnożenia wypełnić można jeszcze dokładniej. Zauważmy bowiem, że wobec istnienia odwrotności dowolnego elementu niezerowego w ciele oraz łączności mnożenia, dla dowolnych $a \neq 1$ oraz $b \neq 0$ mamy $a \cdot b \neq b$. W przeciwnym bowiem razie mielibyśmy $a = abb^{-1} = bb^{-1} = 1$. A zatem $a \cdot b = 1$, nie ma innej możliwości uzupełnienia trzeciego wiersza tabelki. W szczególności mamy:

+	0	1	a	b
0	0	1	a	b
1	1			
a	a			
b	b			

·	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

.

Co natomiast z dodawaniem? Tutaj użyjemy nieco innego podejścia, związanego z pojęciem podciała. Przypomnijmy, że każde ciało zawiera element 1 oraz wszystkie elementy postaci:

$$m \cdot 1 = \underbrace{1 + 1 + \ldots + 1}_{m}.$$

Zauważmy jednak, że jeśli mamy ciało n elementowe i $m > n$, to z zasady szufladkowej Dirichleta wynika, że pewne dwie liczby powyżej są równe. Na przykład dla ciała czteroelementowego pewne dwa z elementów:

$$1, \quad 1+1, \quad 1+1+1, \quad 1+1+1+1, \quad 1+1+1+1+1$$

są równe. Które dwa? Na pewno żadne sąsiednie dwa, bowiem z Obserwacji 2.8 mamy $1 = 1+1 \Rightarrow 0 = 1$. Gdyby było $1 + 1 + 1 = 0$ oraz $1 + 1 \neq 0$, to nasze ciało miałoby za mało elementów, bowiem 1 oraz $a = 1 + 1$ byłyby przeciwne, a więc czwarty element (niezerowy) musiałby być przeciwny do samego siebie, tzn. $b + b = 0$. Ale $b + b = b(1 + 1)$ byłoby iloczynem dwóch niezerowych elementów, sprzeczność. A zatem musi być $1 + 1 = 0$. Co więcej, mamy:

$$(1 + 1)(1 + 1) = 1 + 1 + 1 + 1 = 0.$$

Oczywiście gdyby $1 + 1 + 1 + 1 = (1 + 1)(1 + 1) = 0$, to $1 + 1 = 0$ (to jest bardzo ważny argument potrzebny niżej). Mamy więc $1+1 = a+a = b+b = 0$. co pozwala do końca uzupełnić powyższe tabelki.

+	0	1	a	b
0	0	1	a	b
1	1	0		
a	a		0	
b	b			0

·	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

.

Teraz możemy już zauważyć, że $a + 1 \neq a$, czyli $a + 1 = b$. Analogicznie $b + 1 = a$, co daje:

+	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

·	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

.

Pokazaliśmy, że istnieje dokładnie jedno ciało 4-elementowe. Zauważmy ciekawą rzecz – zachodzą równości

$$a^2 + a + 1 = 0, \quad b^2 + b + 1 = 0.$$

Dlaczego nas to interesuje? Otóż wyrażenie x^2+x+1 nie jest równe 0, dla żadnego $x \in \mathbb{Z}_2$. Używając języka wykładu można powiedzieć, że uzyskane ciało czteroelementowe jest rozszerzeniem ciała $\mathbb{Z}_2$ o pierwiastki wielomianu, który w tym ciele pierwiastków nie ma. Po zakończeniu kolejnego wykładu pokażemy, że nie jest to przypadek. Póki co odnotujmy fundamentalną obserwację wynikającą z naszych rozważań.

Obserwacja 2.9

Dla każdego ciała skończonego K istnieje liczba pierwsza p taka, że w ciele tym zachodzi równość

$$\underbrace{1 + 1 + \ldots + 1}_{p} = 0.$$

W takim przypadku mówimy, że ciało K ma CHARAKTERYSTYKĘ równą p.

Warto odnotować, że również ciała nieskończone mogą mieć charakterystykę dodatnią (parz np. liczby p-adyczne w Dodatku lub ciało funkcji wymiernych $\mathbb{Z}_p(x)$ o współczynnikach w $\mathbb{Z}_p$). Na koniec sformułuję twierdzenie, które można udowodnić zupełnie elementarnie (zachęcam).

Twierdzenie 2.6

Każde ciało skończone ma p^n elementów, gdzie p jest pewną liczbą pierwszą, a n jest pewną liczbą całkowitą dodatnią. Dla każdej pary p, n istnieje ciało p^n-elementowe.

Fakt wymagający w dowodzie nieco większej „technologii algebraicznej” mówi, że dla każdej pary p, n istnieje tylko jedna tabelka ciała p^n-elementowego. W tym momencie się na razie zatrzymamy.

2.5 Dodatek. Liczby p-adyczne

Następująca równość może być na pierwszy rzut oka zupełnie niezrozumiała:

$$-1 = 6 + 6 \cdot 7 + 6 \cdot 7^2 + 6 \cdot 7^3 + 6 \cdot 7^4 + \dots$$

Opisuje ona rozwinięcie 7−adyczne liczby -1 w tzw. ciele 7-adycznym $\mathbb{Q}_7$. Rozwinięcie to ma również zaskakującą notację, bowiem zapis 7-adyczny – inaczej niż zapis dziesiętny liczby rzeczywistej – odczytujemy od prawej do lewej strony tak, że mamy na przykład

$$-1 = \dots 6666.$$

Skąd się biorą takie rozwinięcia i czym jest ciało $\mathbb{Q}_7$ oraz ciała liczb p-adycznych $\mathbb{Q}_p$, gdzie p jest liczbą pierwszą? Konstrukcja ciał $\mathbb{Q}_p$ ma swoje źródło w próbie określenia odległości w zbiorze liczb wymiernych, mającej kompletnie inne własności niż wartość bezwzględna. Owa alternatywna definicja pozwala na zdefiniowane alternatywnych ciągów Cauchy'ego liczb wymiernych, a stąd już tylko krok do ciała $\mathbb{Q}_p$.

Na wykładzie z analizy matematycznej rozważa się dwie konstrukcje liczb rzeczywistych pochodzące od Dedekinda i Cantora. Pierwsza oparta jest na przekrojach, a druga na ciągach Cauchy'ego. Obydwie pochodzą z lat 70' XIX wieku. Wcześniejsze próby rygorystycznego opisu liczb rzeczywistych, podejmowane nawet w XVI wieku (Stevin), a potem intensywnie w wieku XIX (Bolzano, Hamilton, Weierstrass) były jedynie częściowo rygorystyczne. W XX wieku doczekaliśmy się wielu alternatywnych konstrukcji, bardziej lub mniej elementarnych[3]. Z naszego punktu widzenia szczególnie ciekawa jest konstrukcja Cantora. Pomaga nam ona dobrze rozumieć ideę rozwinięcia dziesiętnego liczby rzeczywistej, definiowanego w oparciu o zbieżność ciągów złożonych z przybliżeń tej liczby za pomocą liczb wymiernych.

W dużym skrócie, konstrukcja Cantora jest oparta na obserwacji, że każda liczba rzeczywista a jest granicą ciągu (q_n) liczb wymiernych. Co więcej, dowolne takie dwa ciągi zbieżne (q_n) oraz (q'_n) są zbieżne do tej samej granicy wtedy i tylko wtedy, gdy $|q_n - q'_n| \xrightarrow[n\to\infty]{} 0$. Kluczowym narzędziem są ciągi Cauchy'ego liczb wymiernych, czyli takie ciągi (q_n), że dla każdego $\epsilon > 0$ istnieje $k_0 \in \mathbb{N}$ takie, że $|q_n - q_m| < \epsilon$, o ile tylko $n, m > k_0$. Na czym polega? Bierzemy zbiór C wszystkich ciągów Cauchy'ego liczb wymiernych i mówimy, że dwa elementy zbioru C postaci (q_n) oraz (q'_n) są równoważne, o ile $|q_n - q'_n| \xrightarrow[n\to\infty]{} 0$, co oznaczamy jako $(q_n) \sim (q'_n)$. A zatem zbiór C rozbija się na podzbiory postaci $[(q_n)]$, z których każdy złożony jest równoważnych sobie ciągów Cauchy'ego (oczywiście zbiór $[(q_n)]$ zawiera (q_n)).

Konstrukcja Cantora liczb rzeczywistych polega na utożsamieniu każdej każdej takiej liczby z pewnym podzbiorem $[(q_n)]$. Na zbiorze tych (rozłącznych) podzbiorów definiujemy **działania** dwuargumentowe[4] $\oplus, \otimes$:

$$[(q_n)] \oplus [(q'_n)] = [(q_n + q'_n)], \quad [(q_n)] \otimes [(q'_n)] = [(q_n \cdot q'_n)],$$

gdzie $+, \cdot$ są dodawaniem i mnożeniem w $\mathbb{Q}$. Trzeba pokazać, że działania te są dobrze określone, tzn. że dla każdego ciągu (t_n) w zbiorze $[(q_n)]$ oraz dla każdego ciągu w zbiorze (t'_n) w $[(q'_n)]$ mamy:

$$[(t_n) + (t'_n)] = [(q_n) + (q'_n)], \quad [(t_n t'_n)] = [(q_n q'_n)].$$

Gdy to zrobimy, pozostaje uzasadnić, że spełnione są wszystkie aksjomaty ciała. Na przykład dla dowolnego ciągu Cauchy'ego (q_n) nierównoważnego z ciągiem zerowym trzeba znaleźć ciąg Cauchy'ego (q'_n) taki, że $(q_n q'_n)$ jest ciągiem należącym do $[(1)]$, gdzie $[(k)]$ jest ciągiem stałym, o każdym wyrazie równym k (w ten sposób widzimy, że $\mathbb{Q}$ jest podciałem tak określonego ciała $\mathbb{R}$). W rezultacie $\oplus$ oraz $\otimes$ staną się znanym nam dodawaniem i mnożeniem liczb rzeczywistych.

Co to wszystko ma wspólnego z algebrą liniową i ciałami liczb p-adycznych? Z pewnością konstrukcja działań na zbiorach ciągów, prowadzących do pojęcia ciała, jest ciekawa sama w sobie. Natomiast kluczowy jest dla nas element, którego nie porusza się na początku studiów: definicja ciągu Cauchy'ego na zbiorze liczb wymiernych oparta jest o pojęcie odległości w $\mathbb{Q}$: odległość liczb $p, q \in Q$ równa jest $|p - q|$. Jak się okazuje, nie jest to jedyny „sensowny" sposób określania odległości w $\mathbb{Q}$ (przy czym „sensowny" oznacza, że odległość jako funkcja określona na zbiorze par liczb wymiernych spełnia pewne naturalne warunki – jest tzw. normą, o czym kilka słów będzie pod koniec). Okazuje się, że można wprowadzić inną odległość, związaną z tzw. normą p-adyczną.

[3] I. Weiss: *The Real Numbers - a survey of constructions*, https://arxiv.org/pdf/1506.03467.pdf.

[4] Proszę porównać to z definicją działań dodawania i mnożenia modulo n. Można przyjąć, że definiujemy je na podzbiorach zbioru liczb całkowitych złożonych z liczb dających takie same reszty modulo n. W omawianym przypadku mamy zbiory ciągów Cauchy'ego. Mówiąc ogólniej, mamy tu do czynienia z definiowaniem działań na tzw. zbiorze klas równoważności.

Definicja 2.9

Niech p będzie dowolną liczbą pierwszą, zaś z – niezerową liczbą całkowitą.

- Określamy $v_p(z) = \max\{n \in \mathbb{Z} : p^n \mid z\}$, zwane WYKŁADNIKIEM p-ADYCZNYM z.
- Jeśli $x = \frac{a}{b}$, gdzie $a, b \in \mathbb{Z}$, $b \neq 0$, to określamy $v_p(x) = v_p(a) - v_p(b)$.
- NORMĄ p-ADYCZNĄ nazywamy funkcję $|.|_p : \mathbb{Q} \to [0, \infty)$ określoną wzorem:

$$|x|_p = \begin{cases} p^{-v_p(x)} & , x \neq 0 \\ 0 & , x = 0. \end{cases}.$$

- ODLEGŁOŚCIĄ p-ADYCZNĄ LICZB WYMIERNYCH x, y nazywamy liczbę wymierną $|x - y|_p$.

Zobaczmy kilka przykładów:

$$|2|_2 = 2^{-v_2(2)} = \frac{1}{2}, \quad |3|_2 = 2^{-v_2(3)} = 1, \quad |4|_2 = 2^{-v_2(4)} = \frac{1}{4}, , \quad |-\frac{128}{7}|_2 = 2^{-v_2(2^7)+v_2(-7)} = 2^{-7} = \frac{1}{128}.$$

Co to znaczy, że dwie liczby są „blisko" w normie p-adycznej? Jeśli ograniczymy się do liczb całkowitych, to możemy zauważyć, że są one tym bliżej siebie, im wyższą potęgę p dzielą wspólnie. Dla przykładu:

$$|81 - 1|_3 = |80|_3 = 1, \quad |81 - 6|_3 = |75|_3 = \frac{1}{3}, \quad |81 - 27|_3 = |54|_3 = \frac{1}{27}, \quad |81 - 80|_3 = |1|_3 = 1.$$

Czym jest $\mathbb{Q}_p$? I jak opisywać jego elementy? Okazuje się, że jeśli w konstrukcji Cantora liczb rzeczywistych zastąpimy wartość bewzględną przez odległość p-adyczną i powiemy, że ciąg liczb wymiernych (q_n) jest ciągiem Cauchy'ego, jeśli dla każdego $\epsilon > 0$ istnieje $k_0 \in \mathbb{N}$ takie, że $|q_n - q_m|_p < \epsilon$, o ile tylko $n, m > k_0$, to powtarzając opisaną wcześniej konstrukcję dodawania i mnozenia podzbiorów złożonych z (p-adycznych) ciągów Cauchy'ego dostaniemy własnie ciało liczb p-adycznych.

Podobnie jak liczby rzeczywiste można opisywać przy pomocy rozwinięć, np. dziesiętnego, tak liczby p-adyczne przedstawia się za pomocą tzw. rozwinięcia p-adycznego. Podstawą są następujące rezultaty[5]

Twierdzenie 2.7

Dla $0 < m \in \mathbb{Z}$ niech $d_{-m}, \ldots, d_0, d_1, \ldots$ będą nieujemnymi liczbami całkowitymi mniejszymi niż p, przy czym $d_{-m} > 0$. Rozważmy szereg:

$$d_{-m}p^{-m} + d_{-m+1}p^{-m+1} + \ldots + d_0 + d_1 p + d_2 p^2 + \ldots \qquad (*).$$

Wówczas:

- sumy częściowe szeregu $(*)$ tworzą ciąg Cauchy'ego w $\mathbb{Q}$ (względem $|.|_p$),
- dla każdego elementu $A \in \mathbb{Q}_p$ istnieje dokładnie jeden reprezentujący go ciąg Cauchy'ego (A_i), którego wyrazami są sumy częściowe szeregu typu $(*)$. Definiujemy też $|A|_p = \lim_{n\to\infty} |A_n|_p$.

Wyrażając wynik wyżej w języku teorii podzielności, mamy:

Każdy element $A \in \mathbb{Q}_p$ o własności $|A|_p \leqslant 1$ (liczby całkowite) reprezentowany jest przez dokładnie jeden ciąg Cauchy'ego (A_i) o wyrazach całkowitych spełniający $A = [(A_i)]$ oraz:

(i) $0 \leqslant A_i < p^i$, dla $i = 1, 2, \ldots$

(ii) $A_i \equiv A_{i+1} \bmod p^i$, dla $i = 1, 2, \ldots$.

[5]Rozszerzoną wersję tej opowieści znajdą Państwo pod adresem: `https://mimuw.edu.pl/~amecel/2021l/gal21/p-adyczne1.pdf`, a także w kolejnym wykładzie z tej serii: `https://mimuw.edu.pl/~amecel/2021l/gal21/p-adyczne2.pdf`.

Liczbie p-adycznej przypisujemy rozwinięcie, które po przecinku może mieć jedynie skończenie wiele elementów, ale na lewo od przecinka może mieć ich nieskończenie wiele. Zobaczmy przykłady.

Przykład 1. Liczba 320 ma w ciele $\mathbb{Q}_7$ rozwinięcie $6 \cdot 7^2 + 3 \cdot 7^1 + 5 \cdot 7^0 = 635$, ponieważ mamy ciąg spełniający warunki wyżej:

- $A_1 = 5$ oraz $320 - A_1 \equiv 0 \bmod 7$,
- $A_2 = 3 \cdot 7^1 + 5 \cdot 7^0$ oraz $320 - A_2 \equiv 0 \bmod 7^2$
- $A_n = 6 \cdot 7^2 + 3 \cdot 7^1 + 5 \cdot 7^0$, oraz $320 - A_n \equiv 0 \bmod 7^n$, dla $n \geqslant 3$.

Przykład 2. Poniższy ciąg liczb całkowitych A_n, $n \geqslant 1$, jest ciągiem Cauchy'ego w normie p-adycznej:

$$A_n = (p-1) + (p-1)p + (p-1)p^2 + \ldots + (p-1)p^{n-1}$$

reprezentującym liczbę... -1. Aby to sprawdzić proszę prześledzić wartości $|A_n + 1|_p$. Stąd np. rozwinięcie liczby -1 w zapisie 3-adycznym to ...222.

Przykład 3. Zauważmy, że w normie 3-adycznej mamy:

$$-\frac{3}{2} = \frac{3}{1-3} = 3(1 + 3 + 3^2 + \ldots)$$

Stąd w $\mathbb{Q}_3$ mamy $-\frac{3}{2} = \ldots111$.

Przykład 4. W ciałach p-adycznych nie zawsze istnieją pierwiastki. Jak to wyjaśnić? Wyobraźmy sobie na przykład, że istnieje liczba $x \in \mathbb{Q}_5$, spełniająca równanie $x^2 = 2$. To by oznaczało (pomijamy dla uproszczenia cyfry po przecinku), że liczba ta ma rozwinięcie $a_0 + a_1 5 + a_2 5^2 + \ldots$ oraz $x^2 = 2$, więc po przemnożeniu tego szeregu przez siebie i porównując uzyskane rozwinięcia widzimy, że a_0^2 musi dawać resztę 2 modulo 5 (dlaczego), co jest niemożliwe. Można pokazać, że liczba p-adyczna $x = a_0 + a_1 p + a_2 p^2 + \ldots$ jest rozwiązaniem równania $x^2 = m$ (również dla $m < 0$), wtedy i tylko wtedy, gdy ciąg:

$$A_n = a_0 + a_1 p + \ldots + a_n p^{n-1}$$

jest ciągiem rozwiązań kongruencji $X^2 \equiv m \bmod p^n$.

Ciała p-adyczne mają wiele bardzo ciekawych własności, a uprawiana na nich analiza czy geometria zupełnie nie przypomina tego, czego będziecie się Państwo uczyć (i jest ciekawa). A jednak od ponad 100 lat stanowią one bardzo ważne narzędzie współczesnej matematyki. Jeden z medalistów Fieldsa z 2018 roku, Peter Scholze, otrzymał to najbardziej prestiżowe dla matematyka wyróżnienie właśnie za rozwój geometrii p-adycznej. Zachęcam do obejrzenia filmu z prezentacją Laureata: `https://youtu.be/yEVlCZTqht8`.

Warto zwrócić uwagę na jeszcze jedno, głębokie i zaskakujące zagadnienie. Po przeczytaniu powyższego dodatku, ktoś mógłby słusznie zapytać: czy skoro z $\mathbb{Q}$ można konstruować, za pomocą ciągów Cauchy'ego liczby rzeczywiste i ciała p-adyczne, to czy można w podobny sposób konstruować inne ciała? Może są jeszcze jakieś inne odległości na $\mathbb{Q}$ pozwalające na uzyskanie takiej konstrukcji? Odpowiedź jest negatywna. Mówi o tym twierdzenie Ostrowskiego z 1916 roku. Okazuje się, że każda norma[6] na $\mathbb{Q}$ (od której pochodzi odległość) jest „równoważna" normie zadanej przez wartość bezwzględną lub normę p-adyczną! Równoważność norm ma natomiast miejsce wtedy, gdy każdy ciąg Cauchy'ego względem jednej normy jest ciągiem Cauchy'ego także względem drugiej normy. To wynik bardzo przemawiający do wyobraźni. W celu znalezienia (bardzo przystępnego) dowodu tego faktu odsyłam do tekstu prof. Tomaszewskiego, dostępnego on-line lub pod adresem `https://mimuw.edu.pl/~amecel/2021l/gal21/ostr.pdf`.

Dodam jeszcze, że twórcą/odkrywcą (niepotrzebne skreślić) liczb p-adycznych był Kurt Hensel, a pionierskie badania w tej dziedzinie prowadził na przełomie XIX i XX wieku. Wielkie znaczenie tych ciał zrozumiane zostało po raz pierwszy dzięki niezwykle ważnemu rezultatowi teorioliczbowemu: zasadzie lokalno-globalnej Hassego z 1921 roku. Wspomnimy o tym wyniku w drugim semestrze. Są również elementarne zastosowania normy p-adycznej, np. następujące ciekawe twierdzenie Monsky'ego[7] z 1970 roku: *Kwadratu nie można podzielić na nieparzystą liczbę trójkątów o równych polach.*

[6] W kontekście ciała K normą nazywamy dowolną funkcję $f : K \to \mathbb{R}_+ \cup \{0\}$ taką, że zachodzą trzy następujące warunki: (i) $f(x) = 0$ wtedy i tylko wtedy, gdy $x = 0$, (ii) $f(xy) = f(x)f(y)$, (iii) $f(x + y) \leqslant f(x) + f(y)$, dla dowolnych $x, y \in K$. Odległość $d : K \times K \to \mathbb{R}_+$ to funkcja zadana wzorem $d(x, y) = f(x - y)$. Dla $f(x) = |x|_p$ mamy $d(x, y) = |x - y|_p$.

[7] Patrz np. `http://math.uchicago.edu/~may/REU2019/REUPapers/Sablan.pdf` lub *Monsky's theorem.* Wikipedia.

2.6 Trivia. Równania językowe

Wspomnieliśmy juz o tym, że równania liniowe, których współczynniki nie są w ciele, ale np. w zbiorze liczb całkowitych, mogą nie mieć rozwiązania. Warto pociągnąć ten wątek, przyglądając się jak dalece ogólna może być definicja działania i równania liniowego oraz jak mało oczywistych dla nas własności może ona zachowywać. Rozważymy następujący, egzotyczny przykład struktury algebraicznej z działaniami dodawania i mnożenia. Będą to... zbiory słów, czyli w skrócie: słowniki.

Aby mieć słownik, trzeba mieć najpierw słowa. Niech $\Sigma_{a,b}$ będzie zbiorem złożonym ze wszystkich słów złożonych z liter a, b, np. $a, aaa, abaa, bbba$ oraz słowa pustego ϵ. W $\Sigma_{a,b}$ wprowadzamy działanie dwuargumentowe, tzw. **konkatenację**, $\cdot$ postaci $w_1 \cdot w_2 = w_1 w_2$, np.

$$aba \cdot bb = ababb, \qquad \epsilon \cdot abb = abb.$$

Rozważamy zbiór $X = P(\Sigma_{a,b})$ złożony z podzbiorów (także nieskończonych!) zbioru $\Sigma_{a,b}$, np.

$$\{\epsilon, a, ab\}, \quad \{a, aa, aaa, aaaa, \ldots\}, \quad , \{ab, abab, ababab, \ldots\}.$$

Elementy zbioru X nazywać będziemy słownikami. W $P(\Sigma_{a,b})$ wprowadzamy działania dwuargumentowe:

- $+$ oznaczające sumę mnogościową zbiorów, np. $\{aba, bb, ab\} + \{a, aa, bb\} = \{aba, bb, ab, a, aa\}$,
- $\cdot$ oznaczające zbiór powstający przez konkatenację wszystkich wyrazów z pierwszego zbioru ze wszystkimi elementami z drugiego. Innymi słowy, dla dowolnych słowników A, B, zbiór $A \cdot B$ złożony jest ze słów postaci $a \cdot b$, gdzie $a \in A, b \in B$. Np.

$$\{aba, bb, ab\} \cdot \{a, aa, bb\} = \{abaa, abaaa, ababb, bba, bbaa, bbbb, aba, abbb\}$$

Rozważamy równania liniowe o współczynnikach w $P(\Sigma_{a,b})$, np. równanie o zmiennych x_1, x_2:

$$\{a, aa\} \cdot x_1 + \{bb\} \cdot x_2 = \{ab, aab, bbb, aaab\},$$

którego **rozwiązaniem są pary elementów $P(\Sigma_{a,b})$**. W tym przypadku: $x_1 = \{b, ab\}$, $x_2 = \{b\}$.

Widzimy, że od strony formalnej cała opisana konstrukcja mieści się w definicji równania liniowego i jego rozwiązania. Równania liniowe, między innymi takie, jak wyżej, nazywa się **równaniami językowymi**. Rozwiązywanie tych równań w niczym nie przypomina znanych nam metod. Dlaczego?

- Pierwszy problem to konieczność określenia strony, z której piszemy współczynniki. Równania:

  $$\{a\} \cdot x_1 = \{abaa\}, \quad x_1 \cdot \{a\} = \{abaa\}$$

  mają różne rozwiązania! Przyczyna – nieprzemienność działania $\cdot$.

- Zauważmy, że w zbiorze słowników $P(\Sigma_{a,b})$ nie ma elementów *przeciwnych* i *odwrotnych*. Mając układ:

  $$\begin{cases} \{a\} \cdot x_1 = \{abaa\} \\ \{a\} \cdot x_1 = \{aba\}. \end{cases}$$

  nie sprowadzimy jego *macierzy* do postaci schodkowej lub zredukowanej, co utrudnia sprawdzanie kiedy jest on sprzeczny!

- Układy jednorodne nie mają sensu, bo $\{\epsilon\} \cup \{w\} = \{w\}$, ale $\{\epsilon\} \cdot \{w\} \neq \{\epsilon\}$.

- Trudno kontrolować zbiory rozwiązań. Rozwiązanie równania wyżej było proste, bo współczynnikami były skończone słowniki. Proszę jednak pomyśleć na przykład o równaniu postaci:

  $$X = \{a^n b^n \mid n \geqslant 1\} \cdot X \cdot \{b^n a^n \mid n \geqslant 1\} \cup \{\epsilon\},$$

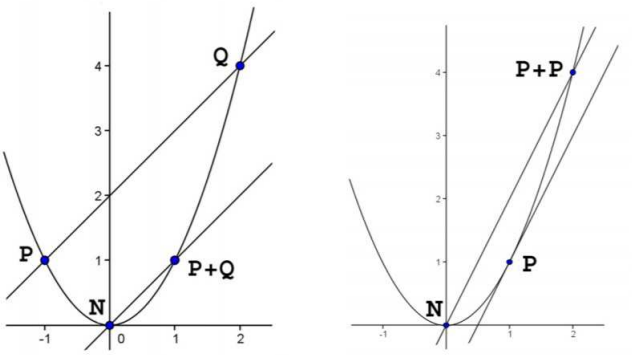
  gdzie ϵ jest słowem pustym.

Na potrzeby ilustracji problemów z nieprzemiennością i brakiem odwrotności operacji mnożenia (a to nie jedyne problemy, jak widać wyżej) dokonuję tu i tak dużego uproszczenia tego tematu. Czytelnik zainteresowany szczegółami może zapoznać się z prezentacją Michala Kunca pt. Language Equations (dostępna online) lub z monografią *Language Equations* autorstwa Ernsta Leissa (biblioteka IMPAN). Wcześniej jednak warto przejść/przeczytać wykład kursowy z Języków, automatów i obliczeń.

2.7 Trivia. Ciało na paraboli

Dziwna może się wydawać definicja ciała liczb zespolonych, gdy patrzymy na nią z punktu widzenia działania na parach liczb. Przypomnę jednak uwagę z wykładu: zbiór $\mathbb{R}^2$ z działaniami dodawania i mnożenia „po współrzędnych" (tzn. z mnożeniem $(a,b)\otimes(c,d)=(ac,bd)$) nie jest ciałem! Ale nie o tym jest ten dodatek. Chciałbym w nim wspomnieć o mało znanym przykładzie ciała, w którym działania dodawania i mnożenia wyglądają bardziej intuicyjnie niż w $\mathbb{C}$, ale zdefiniowane są na dość nietypowym obiekcie.

Rozważamy mianowicie parabolę $\mathcal{P}$ o równaniu $y=x^2$. Pokażemy najpierw, że można wprowadzić na niej strukturę przemiennego dodawania. Niech $N=(0,0)$. Określamy sumę $P\oplus Q$ dwóch punktów paraboli $\mathcal{P}$ jako drugi punkt przecięcia paraboli oraz prostej równoległej do prostej PQ przechodzącej przez punkt N. Jeśli $P=Q$, to zastępujemy prostą PQ przez prostą styczną do paraboli w punkcie P. Poniższe (zapożyczone) rysunki pokazują odpowiednio działania postaci $(-1,1)\oplus(2,4)=(1,1)$ oraz $(1,1)\oplus(1,1)=(2,4)$.



Źródło: Franz Lemmermeyer. *Pell Conics. An Alternative Approach to Elementary Number Theory.* http://www.rzuser.uni-heidelberg.de/~hb3/pell.html

Weryfikacja przypuszczenia, że wprowadzone działanie dwuargumentowe $\oplus$ zadaje na $\mathcal{P}$ działanie łączne z elementem neutralnym N wymaga odrobiny wiedzy szkolnej i wytrwałości. Wyznaczmy wzory na dodawanie dwóch punktów $P_1=(x_1,x_1^2)$ oraz $P_2=(x_2,x_2^2)$. Jeśli $P_1\neq P_2$, to prosta przechodząca przez P_1 oraz P_2 ma współczynnik kierunkowy postaci: $m=(x_2^2-x_1^2)/(x_2-x_1)=x_1+x_2$. Prosta równoległa do prostej P_1P_2 przechodząca przez punkt $N=(0,0)$ ma równanie postaci $y=mx$. Aby znaleźć jej drugi punkt przecięcia z parabolą potrzebujemy rozwiązać równanie $mx=x^2$. To zaś daje nam dwa punkty przecięcia: $N=(0,0)$ oraz $R=(m,m^2)$. A zatem na mocy naszej definicji działania $\oplus$ mamy:

$$(x_1,x_1^2)\oplus(x_2,x_2^2)=(x_1+x_2,(x_1+x_2)^2).$$

Powyższa formuła pozostaje prawdziwa także gdy $P_1=P_2$.

Ktoś powie: *to w zasadzie nic ciekawego*. Matematyk od razu widzi, że zdefiniowane działanie to „właściwie" (nie znamy pojęcia izomorfizmu) dodawanie liczb rzeczywistych. W podobny sposób na hiperboli $xy=1$ wprowadzić można działanie mnożenia, które jest „izomorficzną kopią" mnożenia liczb rzeczywistych. Czy umielibyście Państwo zaproponować geometryczną konstrukcję mnożenia punktów na hiperboli? Aby dowiedzieć się więcej o podobnych konstrukcjach na stożkowych (i nie tylko) zachęcam do lektury książki, do której odsyłam pod powyższym obrazkiem. Znajdziecie tam Państwo łagodny wstęp do teorii punktów wymiernych na krzywych i ich zastosowań. W latach 90' zaawansowane metody teorii krzywych eliptycznych zaowocowały dowodem Wielkiego Twierdzenia Fermata przez Andrew Wilesa. Polecam poglądowy wykład Wilesa i następujący po nim inspirujący wywiad mówiący o życiu zawodowego matematyka. Adres: `https://www.youtube.com/watch?v=uQgcpzKA5jk`.

To jednak nie koniec opowieści. Parabola ma tę szczególną cechę, że można na niej wprowadzić nie tylko strukturę tzw. grupy algebraicznej, ale i strukturę ciała. Spróbujmy więc określić działanie mnożenia. Niech $I=(1,1)$. Dla punktów $P,Q\in\mathcal{P}$ definiujemy mnożenie w następujący sposób: prosta PQ przecina oś OY w punkcie A oraz prosta IA przecina parabolę $\mathcal{P}$ w punkcie $B:=P\star Q$ (proszę to dopracować).

Czy potraficie Państwo napisać algebraiczną formułę opisującą działanie $\star$? Czy potraficie Państwo pokazać, za pomocą tych wzorów, że $(\mathcal{P},\oplus,\star,N,I)$ jest ciałem? Okazuje się, że wcale nie potrzeba tu algebry. Zarówno łączność mnożenia, jak i rozdzielność mnożenia względem dodawania można udowodnić geometrycznie. Czy ktoś z Państwa potrafiłby to zrobić? Fakt ten, co zaskakujące, pochodzi z 2003 roku.

2.8 Coda. O kształtowaniu się pojęcia liczby

Liczby zespolone, a zwłaszcza ich trudny do uchwycenia przez stulecia aspekt bycia pierwiastkami z liczb ujemnych (których przecież *nie powinno* być!), każe przyjrzeć się nieco szerzej całej gamie przykładów historycznych sytuacji, w której jakimś liczbom odmawiano prawa do istnienia, do czasu gdy stojące za nimi koncepcje zostały nie tylko ugruntowane teoretycznie, ale i gdy ich użyteczność stała się w zasadzie ewidentna. Wielkie i historyczne kontrowersje dotyczyły nie tylko liczb zespolonych, o których historii powiemy w dalszej części, ale również innych rodzin liczb — niewymiernych, ujemnych, a nawet zera. Pogłębienie rozumienia koncepcji liczby znamionowało zawsze istotny przełom w matematyce.

Zacznijmy od liczb naturalnych. Czy kiedykolwiek ich nie akceptowano? Z pewnością małe liczby — owszem, ale czy również duże? Największa liczba mająca samodzielną nazwę u starożytnych Greków to *myrias*, czyli 10 000. Stąd największa wówczas — wciąż obecna w kulturze nazwa: miriady miriad — odnosi się do liczby stu milionów. Badacze starych dialektów oraz języków ludów pierwotnych zauważają, że budowa liczebników ma swoiście addytywną strukturę — w odróżnieniu od stosowanej przez nas notacji pozycyjnej, opartej na potęgowaniu. I tak w papuaskim plemieniu Wedau, liczba 2 ma nazwę *ruag'a*, a liczba 4 — *ruag'a ma-ruag'a*, czyli $2+2$. Liczba 5 to ura-i-ga, a liczba 9 — *ura-g'ela-ruag'a-mu-ruga'a*, czyli 5 i $2+2$. Nawet zapis łaciński cyfr rzymskich nie kojarzy nam się z wielkimi liczbami. Imperium Cezara interesujące się głównie praktycznymi zastosowaniami, wyrażało milion jako *dieces centena milia*, czyli dziesięć setek tysięcy. Samo słowo „milion" pojawiło się dopiero w trzynastowiecznej Francji.

Czy duże liczby nie istniały więc w starożytności? Istniały, choć nie znalazły wykorzystania. Były raczej ucieleśnieniem koncepcji — liczby nie są ograniczone. Archimedes oraz Diofantos dali początki notacji potęgowej, mówiąc o potęgach *miriady miriad* (osiągając liczbę $10^{8\cdot 10^{16}}$, czyli miriadę miriad podniesioną do miriady miriad, która to jest podniesiona do potęgi o wykładniku miriady miriad). Inaczej sytuacja wyglądała na dalekim Wschodzie. Już starożytny Sanskryt zawiera nazwy potęg dziesiątki aż do 10^{12}. Późniejsze teksty hinduistyczne i buddyjskie rozszerzały zakres nazwanych liczb najpierw do 10^{421}, a później nawet do $10^{10\cdot 2^{122}}$, nazwanej liczbą niewypowiadalną. Teksty tych kultur jeszcze przed początkiem naszej ery dzieliły liczby na liczby na trzy kategorie: policzalne (najmniejsze, pośrednie i najwyższe), niepoliczalne (prawie, istotnie i niepoliczalnie niepoliczalne) oraz nieskończone (prawie, istotnie i nieskończenie).

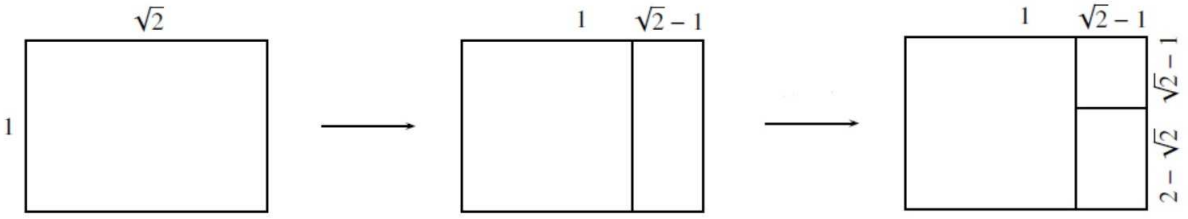
A co z samą nieskończonością? Temu zagadnieniu należałoby poświęcić osobny dodatek. Arystoteles wyróżniał nieskończoność potencjalną oraz aktualną. Ta pierwsza oznacza, że jakąkolwiek podamy wielkość geometryczną czy liczbową, zawsze znajdzie się wielkość (liczba) od niej większa. Tego rodzaju nieskończoność akceptowana była od starożytności. Nie zakładała ona istnienia nieskończonego bytu, a jedynie nieustającą możliwość powiększania lub pomniejszania, i tak stosowały ją choćby *Elementy* Euklidesa. Również nowożytne przejścia graniczne pozostają w istocie w obszarze nieskończoności potencjalnej.

Nieskończoność aktualna, to nieskończoność dokonana, to jest taka, w której nieskończenie wiele kroków jest wykonanych w jednym działaniu, np. poprzez wzięcie nieskończonego zbioru liczb naturalnych. Tego typu nieskończoność odrzucano zarówno w starożytności, jak i w epokach nowożytnych w zasadzie aż do XIX wieku, do badań Cantora. Wiązała się ona z lękiem przed paradoksami, począwszy od słynnego paradoksu Zenona, do współczesnych paradoksalnych rozkładów czy innych „pułapek" teorii mnogości. Stąd nawet XX-wieczny wielki matematyk Henri Poincare powie, że „Następne pokolenia potraktują teorię zbiorów jako chorobę, z której udało im się wyleczyć". Znany nam symbol nieskończoności ∞ pochodzi od Wallisa (1655). Symbole $\aleph_0$ oraz $\mathfrak{c}$ związane z liczbami kardynalnymi pochodzą od Cantora.

Można by przypuszczać, że po dodatnich liczbach naturalnych służących do zliczania, naturalnym etapem rozwoju będą liczby ujemne, ale stało się inaczej. W istocie, wcześniej rozważane były ułamki (podobnie jest zresztą w szkole). Jak czytaliśmy w dodatku do wykładu pierwszego, już Babilończycy posługiwali się połówkami, ćwiartkami czy trzecimi częściami całości. Również zapis hieroglificzny stosowany w starożytnym Egipcie dopuszczał odwrotności liczb całkowitych, wraz z przekonaniem, że każdy ułamek można zapisać jako sumę takich odwrotności (i odpowiedniej liczby całości), czego oczywiście nikt wówczas nie dowodził — zrobił to dopiero Fibonacci w 1202 roku. Idea egipskiego zapisu ułamków przeniknęła do świata greckiego, gdzie przestrzegał go jak się wydaje sam Euklides. Egipcjanie mieli zresztą specjalne tablice rozkładu liczb $2/n$ na ułamki, a dla ułamka $2/3$ istniał osobny hieroglif. To, co jest interesujące z punktu widzenia historii rozwoju pojęcia liczby, to wstrząsające odkrycie, że nie wszystkie wielkości są proporcjonalne, które zachwiało filozoficznymi podstawami szkoły Pitagorejczyków.

Dlaczego dla greckiego świata odkrycie wielkości nieproporcjonalnych było aż tak wielkim szokiem? Pitagorejczycy wierzyli w spajającą świat *harmonię*, która ma być dla ludzi — jak pisze prof. Kordos w swoich *Wykładach z historii matematyki* — motywacją życia i pełnią człowieczeństwa. Pitagorejczycy wyróżnili dziedziny, w których harmonia jest najbardziej widoczna: muzykę, astronomię, arytmetykę i geometrię (późniejsze *quadrivium* stanowiące wyższy stopień kształcenia średniowiecznego, obok niższego *trivium* — gramatyki, retoryki i dialektyki). Oni to odkryli, że jeśli długości dwu napiętych jednakową siłą strun mają się jak 1 do 2, to wydają one harmonijne brzmienie. Podobnie z innymi stosunkami długości: 2:3, czy 3:4, które dały podstawy do wyróżnienia interwałów: oktawy, kwarty i kwinty. Tego typu głębokie związki budziły niezwykłe zainteresowanie. Koncepcja powiązania wielkości proporcjonalnych ze zjawiskami natury była tak atrakcyjna i narzucająca się, że powstała hipoteza, że w liczbach należy szukać istoty harmonii. Również z matematycznego, czy raczej — geometrycznego punktu widzenia, proporcje były w centrum rozważań pitagorejskich. Stąd rozbudowana teoria figur podobnych, skumulowana wokół twierdzenia Talesa. Z punktu widzenia arytmetyki (wciąż uprawianej geometrycznie) — kluczem był algorytm Euklidesa. Co ciekawe nie dotyczył on w istocie liczb całkowitych, ale *współmiernych*.

Wiemy dobrze, że długość boku kwadratu i jego przekątnej nie jest współmierna — znamy różne dowody tego faktu. Szczególnie interesujący polega na następującej obserwacji. Rozważmy prostokąt o bokach długości $a = 1$ oraz $b = \sqrt{2}$ (czy inaczej — jeden z boków to bok kwadratu, a drugi – ma długość jego przekątnej). Reprezentujemy odejmowanie mniejszej liczby od większej poprzez obcięcie z wyjściowego prostokąta kwadratu o krótszym boku. Zatem w dwóch krokach otrzymamy prostokąt o bokach długości $\sqrt{2}-1$ oraz $\sqrt{2}-2 = \sqrt{2}(\sqrt{2}-1)$, mający taki sam kształt, jak wyjściowy prostokąt, przy czym dłuższy bok jest teraz pionowy, a krótszy poziomy. Z tego wynika, że opisany proces nigdy się nie skończy, a przecież dla liczb wymiernych jest inaczej! Oczywiście liczby niewymierne $\sqrt{2}$ i $3\sqrt{2}$ są współmierne.



Przejdźmy teraz do liczb ujemnych. Rozumiejąc już, że w ujęciu starożytnych Greków liczby stanowiły głównie reprezentację obiektów geometrycznych, nietrudno rozumieć, że nie stosowano liczb ujemnych. Długości, pola, objętości musiały być dodatnie. Około trzeciego wieku aleksandryjski matematyk Diofantos napisał wielkie i ważne dzieło *Arytmetyka*, w którym przedstawił zbiór problemów wraz z zaczątkami symboliki służącej do ich rozwiązywania. W jednym z rozwiązań Diofantos zapisze równanie, które dziś czytalibyśmy jako $4 = 4x + 20$, a które nazwie *absurdalnym*.

W istocie prawa działań na liczbach ujemnych zostały sformułowane już w VII wieku przez hinduskiego uczonego Brahmaguptę. Sięgając jeszcze głębiej w przeszłość, ślady liczb ujemnych znaleźć można w starożytnych chińskich dziełach, zawierających ciekawą koncepcję (dydaktyczną) oznaczania liczb dodatnich kolorem czerwonym, a ujemnych — kolorem czarnym. Wszystko to wprowadzano w kontekście obliczeń finansowych i podatkowych, w których liczby czarne bilansowały czerwone. Kwota sprzedaży była czerwona (otrzymujemy pieniądze), a która wydana na zakup była czarna (trzeba wydać). Bilans był dodatni, a deficyt — ujemny. Koncepcja ta pochodziła również z astronomii, gdzie przybliżano liczby z góry i z dołu, Przybliżenia z góry traktowano jako silne, a z dołu — jako słabe. Również Brahmagupta używał nie tylko liczb ujemnych, ale i specjalnego oznaczenia dla liczb ujemnych, a także liczby zero, wraz z odpowiednimi prawami działań. Co ciekawe, matematyka arabska, choć świadoma była osiągnięć hinduskich, odrzucała liczby ujemne, podobnie jak matematyka europejskiego Średniowiecza.

Nawet szesnastowieczni uczeni rozwiązujący równania wielomianowe metodą geometryczną, nie akceptowali do końca używania współczynników ujemnych, choć stosowali odpowiednią symbolikę. John Wallis (1616-1703) oswoił nieco liczby ujemne poprzez wprowadzenie osi liczbowej. Jednak współczesny mu Kartezjusz — autor układu współrzędnych, zaniedbywał współrzędne ujemne. Dopiero rozwój rachunku wektorowego, począwszy od Galileusza, da liczbom ujemnym solidną pozycję w matematyce. Rozmaite wątpliwości pozostaną jednak na długi czas. Nie rozumiano choćby znaczenia iloczynu $(-1) \times (-1)$.

Jeszcze w 1770 roku Euler „dowodzi" w swojej *Algebrze*, że $\sqrt{-2}\cdot\sqrt{-3} = \sqrt{6}$. Między Leibnizem, Eulerem, Bernoullim i d'Alembertem zaistniał poważny spór o to czy $\log(-x)$ jest tym samym, co $\log(x)$. W ten sposób doszło do pewnego rozdwojenia: liczb ujemnych (i zespolonych) używano jako narzędzi formalnych do uzyskiwania rozwiązań, choć niekoniecznie przypisywano im niezależny byt. Jeszcze w 1758 roku brytyjski matematyk Maseres napisze, że liczby ujemne *zaciemniają całą naukę o równaniach i zaciemniają rzeczy, które w swej naturze są zbyt oczywiste i proste.*

Zanim przejdziemy do liczb zespolonych, dodajmy jeszcze kilka zdań o zerze, które przybyło do Europy wraz z Fibonaccim (który jako syn kupca podróżował po basenie Morza Śródziemnego poznając matematykę hinduską oraz arabską) i popularyzowanym przezeń zapisem dziesiętnym. Pierwotnie zero funkcjonowało jako symbol pozwalający odróżnić 1 od 10 lub 100. Mimo rozbieżności zdań historyków co do tego kto pierwszy używał go w takim kontekście, wiadomo na pewno, że symbol ten wystąpił w hinduskiej notacji w roku 876 (już w formie zera, a nie kropki, czy innego symbolu). Jeśli chodzi o zero jako liczbę, rozważane było ono już przez Brahmaguptę, który starał się jednocześnie określić prawa działań na liczbach, i jako jeden z pierwszych natknął się na problem dzielenia przez zero, określając przy tym absurdalne prawa w rodzaju $0/0 = 0$. Również późniejsi uczeni hinduscy zmagali się z tym wyzwaniem, aż do żyjącego w XII wieku Bhaskary, który próbował przypisać ilorazowi $n/0$ wartość nieskończoną, oczywiście łącząc to z językiem religijnym. Sformułuje jednak poprawnie prawa $0^2 = 0$ oraz $\sqrt{0} = 0$.

Wielkim osiągnięciem matematyki arabskiej, a zwłaszcza Khwarizmiego było opisanie indyjskiego systemu pozycyjnego, oraz praw działań pisemnych, przejętych później przez Fibonacciego i Europejczyków (nie bez kontrowersji — zwyciężył pragmatyzm i szybkość, z jaką dokonywano rachunków). Równiez w Europie nie od razu akceptowano zero. Ufundowana na grecko-rzymskiej filozofii, i na dziełach Arystotelesa, nauka europejska była raczej przeciwna idei nicości. Poza tym zero łatwo był pomylić z cyfrą 6 lub 9, co powodowało nawet czasowe wyłączanie go z użycia (na przykład we Florencji na początku XIV wieku). Ważnym osiągnięciem, obecnym już u Claviusa (1608) było jednakże rozważanie zera jako współczynnika w równaniu wielomianowym i świadomość, że jego rozwiązanie może być uzyskane przez rozkład na czynniki liniowe. Twierdzenie Bezout w istocie zdaje się pochodzić od Kartezjusza, a sformułowane jest w jego dziele *Geometria* z roku 1637. W roku 1657 wspomniany już Wallis zadeklaruje, że zero nie jest liczbą, a wspomniany już symbol nieskończoności wprowadzi właśnie dla oznaczenia wyniku $1/0$.

Słynny filozof i matematyki Alfred North Whitehead napisze w roku 1911 w swoim *Wstępie do Matematyki*, że „nikt nie idzie do sklepu, aby kupić zero ryb". Doda jednak zaraz, że użycie zera jest na nas wymuszone potrzebami utartych już schematów myślenia. Być może nie utarły się one jednak całkowicie. W 2000 roku ludzkość witała „nowe millenium", choć w istocie tak trzecie tysiąclecie, jak i XXI wiek rozpoczęły się 1 stycznia 2001 roku, z uwagi na brak roku zerowego.

* * *

Co z liczbami zespolonymi? Definicja, którą oglądaliśmy na wykładzie pochodzi z roku 1833 i jest zasługą Hamiltona — odkrywcy kwaternionów. Była ona swego rodzaju ukoronowaniem 300 lat wysiłków matematyków, którzy od 1545 roku mierzyli się z konsekwencjami wyników opisanych w dziele *Ars Magna* (`Wielka Sztuka`) autorstwa (jak mówią historycy – wybitnego uczonego, ale oszusta) Girolamo Cardano. Wyjawiona w nim była metoda rozwiązywania równań wielomianowych stopnia trzeciego i czwartego zakładająca konieczność wyciągania pierwiastków z liczb ujemnych, a która opracowana została w połowie XVI wieku przez del Ferro, Targaglię oraz Cardano. Ich rozwiązanie równania trzeciego stopnia postaci

$$y^3 = py + q.$$

Historia ta warta jest opowiedzenia. Pierwszą osobą, która znalazła rozwiązanie powyższego równania był boloński profesor Scipione del Ferro (1465-1526). Jego ojciec Floriano pracował w przemyśle drukarskim i już w wieku młodzieńczym Scipione miał dostęp do wielu klasycznych prac, w tym oczywiście do „Liber Quadratorum" (Księga Kwaddratów) Fibobacciego. Czytelnika zaskoczy być może fakt, ze nie zachowały się żadne prace del Ferro. Wydaje się, że obawiał się wyzwania i ewentualnej utraty pozycji na uniwersytecie w Bolonii. Trzymał więc swoją pracę w ukryciu, dzieląc się jedynie z najbliższymi uczniami. Zachował przy tym notatnik, w którym zapisał wszystkie największe osiągnięcia. Po śmierci, jego zięć Annibale della Nove — sam matematyk i były uczeń del Ferro, odziedziczył notatki stryja, wraz z pozycją na uniwersytecie. Same notatki pozostały jednak w ukryciu aż do roku 1543, gdy della Nove odwiedzili dwaj bardzo znani matematycy: Gerolamo Cardano i Lodovico Ferrari, poszukujący metody del Ferro. Skąd o niej wiedzieli?

Kilka lat wcześniej w Wenecji głośno było o matematyku samouku Nicolo Targaglii, chętnie uczestniczącym w pojedynkach matematycznych. Do jednego z tych pojedynków stanął z Tartaglią niejako Fior — jeden z uczniów del Ferro, niezbyt pojętny, jak podają relacje. Każda ze stron podała drugiej 30 równań do rozwiązania. Wszystkie dotyczyły równania wielomianowego stopnia 3. Co o nich wówczas wiedziano?

Matematycy wiedzieli wówczas, że rozwiązanie ogólne równania trzeciego stopnia może być zredukowane do rozwiązania jednego z dwóch typów równań:

$$x^3 + mx = n, \quad x^3 = mx + n, \quad \text{gdzie } m, n > 0.$$

Dlaczego dwa typy? Nie uznawano wówczas liczb ujemnych i przekształceń równoważnych z ich użyciem. Fior wiedział jak rozwiązywać tylko pierwszy z wymienionych wyżej typów. Zadania Fiora dotyczyły zatem jedynie tej klasy równań, podczas gdy zadania Tartaglii były bardzo różnorodne. W trakcie pojedynku, 13 lutego nad ranem, Tartaglia odkrył metodę rozwiązywania równań pierwszego typu i mecz wygrał, co dało mu wielką sławę i zainteresowanie słynnego lombardzkiego matematyka Cardano.

Cardano poprosił Tartaglię w 1539 roku o wyjawienie metody rozwiązywania tych równań i obiecał dochowania tajemnicy i nieujawniania metody. W 1540 roku asystent Cardano Lodovico Ferrari opracował metodę redukcji równań czwartego stopnia do równań sześciennych, co motywowało dodatkowo do złamania obietnicy. Jak to zrobić? Rozwiązaniem okazała się wizyta u zięcia del Ferro — wspomnianego już della Nove w 1543 roku. Cardano uznał, ze to właśnie del Ferro odkrył jako pierwszy metodę rozwiązania równań stopnia 3 i poczuł się zwolniony z tajemnicy danej Cardano.

W 1545 roku Cardano opublikował *Artis Magnae, Sive de Regulis Algebraicis Liber Unus* (Księga Pierwsza o Wielkiej Sztuce, lub o Zasadach Algebry, stanowiące obok słynnego "De revolutionibus" Kopernika i "De human corporis fabrica" Vasaliusa, jedno z trzech najważniejszych traktatów naukowych wczesnego renesansu. Pierwsze wydania tych trzech dzieł miały miejsce w latach 1543-1545.

Wielkość dzieła Cardano oparta była na kilku czynnikach. Kojarzymy przede wszystkim pierwsze bezpośrednie wprowadzenie do języka matematyki pierwiastków z liczb ujemnych, zwanych później liczbami zespolonymi. Istotą rewolucji nie były wówczas jednak same liczby zespolone (których występowanie bagatelizował sam autor, nie potrafiąc im przypisać żadnego fizycznego znaczenia), ponieważ liczby te wcale nie posłużyły Cardano do rozwiązywania równań stopnia 3.

Z uwagi na to, że ówcześnie nie stosowano w przekształceniach algebraicznych liczb ujemnych, Cardano zmuszony był rozpatrywać aż trzynaście rozmaitych klas równań stopnia 3. Rozwiązanie żadnej z tych klas nie wymagało użycia liczb zespolonych. Liczby zespolone wspomniane są przy rozwiązywaniu klasycznego problemu poszukiwania liczb, których suma równa jest 10, a iloczyn równy jest 40. Cardano wprowadził również koncepcję pierwiastka wielokrotnego wielomianu, między innymi rozpatrując liczbę -2 jako dwukrotny pierwiastek równania $x^3 = 12x + 16$.

Jak rozwiązywano równanie $y^3 = py+q$? Dokonując kolejnej liniowej zamiany zmiennych postaci $y = u+v$, uzyskuje się po lewej stronie:

$$(u^3 + v^3) + 3uv(u + v) = 3uvy + (u^3 + v^3),$$

które to wyrażenie równe jest prawej stronie wcześniejszego równania wtedy i tylko wtedy, gdy:

$$3uv = p,$$
$$u^3 + v^3 = q.$$

Eliminując v uzyskujemy równanie kwadratowe zmiennej u^3 postaci

$$u^3 + \left(\frac{p}{3u}\right)^3 = q,$$

o rozwiązaniach

$$\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}.$$

Rozumując symetrycznie, uzyskujemy te same wartości v^3. Skoro $u^3 + v^3 = q$, to jeden z pierwiastków równy jest u^3, a drugi — v^3. Bez straty ogólności można przyjąć

$$u^3 = \frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}, \quad v^3 = \frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3},$$

uzyskując

$$y = u + v = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}.$$

Jeżeli wyobrazimy sobie teraz, ze Cardano dokonywał tych manipulacji w języku geometrycznym, rozumiemy jak bardzo konieczne było przejście od geometrii do języka wyrażeń algebraicznych, dokonane przez Viete'a ponad 100 lat później. Co jest jednak kluczowe? Uzyskane rozwiązanie wymaga użycia liczb zespolonych w przypadku, gdy $(q/2)^2 - (p/3)^3 < 0$. Nie jest możliwe pominięcie tego rozwiązania, gdyż wielomian stopnia 3 zawsze posiada choćby jeden pierwiastek rzeczywisty. Stąd formuła Cardano stawia problem wyciągnięcia liczby rzeczywistej z wyrażenia postaci:

$$\sqrt[3]{a + b\sqrt{-1}} + \sqrt[3]{a - b\sqrt{-1}}.$$

Cardano nie zmierzył się z tym problemem w swojej *Ars Magna* z 1545 roku. Wspomniał wprawdzie o liczbach zespolonych w kontekście równania kwadratowego, ale uznał je za „tak subtelne, jak bezużyteczne". Problemem powyższym poważnie zajął się dopiero Rafael Bombelli w 1572 roku, w podręczniku, którego popularność trwała aż do czasów Leibniza i Eulera. To jemu zawdzięczamy symbol $\sqrt{-1}$ oraz redukcję wyrażenia postaci $\sqrt[3]{a + b\sqrt{-1}}$ do postaci algebraicznej $c + d\sqrt{-1}$. I rzeczywiście, rozpatrując równanie

$$x^3 = 15x + 4$$

uzyskujemy z wzoru Cardano rozwiązanie

$$x = \sqrt[3]{2 + 11\sqrt{-1}} + \sqrt[3]{2 - 11\sqrt{-1}},$$

a czysto bezpośrednie sprawdzenie pozwala znalezienie pierwiastka $x = 4$. Bombelli przypuszczał, że dwa składniki rozwiązania x są postaci $2 + n\sqrt{-1}$ oraz $2 - n\sqrt{-1}$. Podnosząc tę równość do potęgi 3 i używając formalizmu $(\sqrt{-1})^2 = -1$, uzyskał

$$\sqrt[3]{2 + 11\sqrt{-1}} = 2 + \sqrt{-1}, \quad \sqrt[3]{2 - 11\sqrt{-1}} = 2 - \sqrt{-1}.$$

Termin *liczby urojone*, przypisywany liczbom zespolonym, pochodzi od twórcy geometrii analitycznej — Kartezjusza i ma już pierwsze konotacje geometryczne. Reprezentację geometryczną sugerował zresztą już John Wallis w połowie wieku XVII. Oznaczenie $i = \sqrt{-1}$ pochodzi od Leonharda Eulera. On również wykorzystywał postać trygonometryczną, a także przedstawiał pierwiastki równania $z^n = 1$ jako wierzchołki wielokąta foremnego. Jednocześnie, przywołując znów książkę prof. Kordosa, sam Euler w swojej *Algebrze* odnosił się z dystansem do liczb zespolonych, pisząc:

> Pierwiastki kwadratowe z liczb ujemnych nie są zerem, ani nie są ujemne, ani dodatnie. Stąd wynika, ze pierwiastki te nie mogą się znajdować wśród możliwych liczb. W konsekwencji są to niemożliwe liczby. I tak dochodzimy do pojęcia liczb na ogół zwanych urojonymi lub wyobrażalnymi dlatego, że istnieją one tylko w wyobraźni.

Reprezentacja geometryczna liczb zespolonych pochodzi od Caspara Wessela (1797) i jest zbliżona do rachunku na wektorach. Pojęcie płaszczyzny zespolonej, zwanej inaczej płaszczyzną Arganda, związane jest z geometryczną interpretacją Jeana-Roberta Arganda z 1806 roku. Algebraiczną definicję, którą posłużyliśmy się na wykładzie podał w 1831 roku Rowan Hamilton. Pojęcie *liczb zespolonych* wprowadził w tym samym roku Gauss, zdecydowanie oponując przeciwko terminowi *liczb urojonych*.

Teoria liczb zespolonych znalazła wcześnie ważne zastosowania praktyczne, między innymi w równaniach hydrodynamiki, dzięki pracom d'Alemberta (1752). Równania ta zostały dogłębnie zrozumiane i opracowane przez Cauchy'ego oraz Riemanna w połowie XIX wieku, którzy to zajmowali się już funkcjami zespolonymi w perspektywie szeregów potęgowych i rachunku różniczkowego. O tych zagadnieniach dowiecie się Państwo najpierw na Analizie, a potem na dedykowanym przedmiocie — Funkcjach Analitycznych. Wraz z rozwojem fizyki, liczby zespolone zagościły m.in. w teorii fal elektromagnetycznych czy w mechanice kwantowej, między innymi w słynnym równaniu Schrödingera. A co dalej? Kształtowanie się rozumienia pojęcia liczby rozwijało się i przed, i obok, i po zaakceptowaniu liczb zespolonych, gdy budowano teorie kwaternionów, teorię liczb hiperzespolonych, liczb algebraicznych, porządną teorię liczb rzeczywistych i dwudzieste skomplikowane teorie, choćby leżące u podstaw analizy niestandardowej.

Rozdział 3

Wielomiany i funkcje wielomianowe. Równania wielomianowe

3.1 Wykład trzeci

Na poprzednim wykładzie wprowadziliśmy pojęcie ciała i skupiliśmy się omówieniu dwóch istotnych przykładów: ciała reszt z dzielenia przez p oraz ciała liczb zespolonych. Celem tego wykładu[1] jest przedstawienie kilku uwag dotyczących funkcji o wartościach w tych ciałach. Zagadnienie to jest w ogólności niezwykle szerokie, natomiast mając na względzie program kolejnych wykładów, ograniczymy się jedynie do tzw. funkcji wielomianowych. Zacznijmy od pojęcia wielomianu o współczynnikach w ciele.

Definicja 3.1: Wielomian

WIELOMIANEM zmiennej x o współczynnikach w ciele K nazywamy wyrażenie:

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,$$

gdzie n jest nieujemną liczbą całkowitą oraz $a_0, a_1, \ldots, a_n \in K$. Utożsamiamy przy tym takie napisy, jeśli różnią się o składniki postaci $0 \cdot x^i$ oraz jeśli różnią się kolejnością składników.

Elementy a_i nazywamy WSPÓŁCZYNNIKAMI wielomianu. Zbiór wielomianów o współczynnikach z ciała K oznaczamy przez $K[x]$. Jeśli wszystkie współczynniki wielomianu w są równe zero, to piszemy $w = 0$, a wielomian w nazywamy wówczas WIELOMIANEM ZEROWYM.

Innymi słowy, wielomiany zmiennej x o współczynnikach w ciele K utożsamiać można z ciągami nieskończonymi (a_i), dla $i \in \mathbb{N}$, w których $a_i \neq 0$ tylko dla skończenie wielu i. Jest to definicja nieco inna niż ta znana ze szkoły, w której utożsamialiśmy wielomiany (jako wyrażenia algebraiczne) i funkcje wielomianowe. Wprowadźmy od razu to rozróżnienie.

Definicja 3.2: Funkcja wielomianowa

Niech $F = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \in K[x]$. Funkcję $f : K \to K$ daną wzorem

$$f(s) = a_0 + a_1s + a_2s^2 + \ldots + a_ns^n$$

nazwiemy FUNKCJĄ WIELOMIANOWĄ odpowiadającą wielomianowi F.

Zauważmy, że z punktu widzenia podejścia funkcyjnego, dwie funkcje wielomianowe f, g są równe, jeśli istnieją takie wielomiany $F = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, $G = b_0 + b_1x + b_2x^2 + \ldots + b_mx^m \in K[x]$, że dla każdego $s \in K$ mamy

$$f(s) = a_0 + a_1s + a_2s^2 + \ldots + a_ns^n = b_0 + b_1s + b_2s^2 + \ldots + b_ms^m = g(s).$$

[1] Ostatnia aktualizacja: 13.10.2023 r.

To kryterium równości funkcji wielomianowych nie daje podstawy do łatwego rozstrzygnięcia, czy wielomiany F, G są równe. Wymaga to uzasadnienia. W szkole następuje zatarcie pomiędzy pojęciami wielomianu i funkcji wielomianowej tak, że w zasadzie pod pojęciem wielomianu umieszcza się w istocie funkcję wielomianową i formułuje się tzw. TWIERDZENIE O WIELOMIANACH RÓWNYCH, mówiące że równość dwóch funkcji wielomianowych o współczynnikach w zbiorze $\mathbb{R}$ pociąga za sobą równość odpowiadających im wielomianów. Rezultat ten jest prawdziwy dla każdego ciała nieskończonego, czego tu nie dowodzimy. Oto przykład świadczący o tym, że wielomiany o współczynnikach w ciele skończonym nie mają tej własności.

Wielomian $x^2 + x \in \mathbb{Z}_2[x]$ jest niezerowy, ale dla każdego $s \in \mathbb{Z}_2$ wyrażenie $s^2 + s$ równe jest 0. A zatem funkcja wielomianowa $f : \mathbb{Z}_2 \to \mathbb{Z}_2$ zadana wzorem $f(x) = x^2 + x$ jest funkcją zerową. A zatem znając jedynie zbiór wartości funkcji wielomianowej $w(s)$ nie zawsze rozpoznamy wielomian $w \in K[x]$, przynajmniej gdy K jest nad ciałem skończonym. W drugim semestrze mówiąc o wielomianach i funkcjach wielomianowych (wielu zmiennych) pokażemy, że nad ciałem nieskończonym taki problem nie zachodzi.

Wniosek 3.1: Równość wielomianów

Jeśli $w = a_0+a_1x+a_2x^2+\ldots+a_nx^n$ oraz $v = b_0+b_1x+b_2x^2+\ldots+b_mx^m$, dla pewnych nieujemnych liczb całkowitych n, m, to następujące warunki są równoważne:

- $w = v$ jako elementy $K[x]$.
- $m = n$ oraz $a_i = b_i$, dla każdego $1 \leqslant i \leqslant n$.

Powyższy wniosek łatwo przełożyć na język ciągów nieskończonych o skończenie wielu niezerowych wyrazach ze zbioru K. Dwa wielomiany, widziane jako ciągi, są równe, gdy są równe jako ciągi.

Definicja 3.3: Stopień wielomianu

Niech $f = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \in K[x]$, gdzie K jest ciałem. STOPNIEM WIELOMIANU f, ozn $\deg(f)$, nazywamy:

- największe takie i, że $a_i \neq 0$, o ile f nie jest wielomianem zerowym.
- $-\infty$, jeśli f jest wielomianem zerowym.

Przykłady. Mamy $\deg f = 4, \deg p = 7, \deg h = 2$, przy czym

$$f = 1 - 2x + 7x^3 + 5x^4 \in \mathbb{R}[x], \quad p = 2t^7 - \sqrt{2}t^3 - 99 \in \mathbb{R}[t], \quad h = 9i + (5 - i)z + (2 + 7i)z^2 \in \mathbb{C}[z].$$

W zbiorze $K[x]$ określamy działania 2-argumentowe dodawania i mnożenia, pochodzące od działań w K.

Definicja 3.4: Suma i iloczyn wielomianów

Dla wielomianów $f = a_0 + a_1x + \ldots + a_nx^n, g = b_0 + b_1x + \ldots + b_mx^m$ ze zbioru $K[x]$ określamy:

- sumę $f + g$ wielomianów f, g daną wzorem

  $$f + g = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots$$

  Innymi słowy współczynnik wielomianu $f + g$ stojący przy x^i równy jest $a_i + b_i$.

- iloczyn $f \cdot g$ wielomianów f, g dany wzorem

  $$f \cdot g = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \ldots + a_nb_mx^{n+m}.$$

  Innymi słowy współczynnik wielomianu $f \cdot g$ stojący przy x^i równy jest $\sum_{j=0}^{i} a_jb_{i-j}$.

Jako ćwiczenie pozostawiamy następujące własności stopnia, związane z wprowadzonymi operacjami[2]:

$$\deg(f+g) \leqslant \max(\deg(f), \deg(g)), \quad \deg(fg) = \deg(f) + \deg(g). \qquad (\heartsuit)$$

Definicja 3.5: Pierwiastek wielomianu, równanie wielomianowe

PIERWIASTKAMI WIELOMIANU $f \in K[x]$ (inaczej: miejscami zerowymi) nazywamy takie $s \in K$, że funkcja wielomianowa odpowiadająca wielomianowi s przyjmuje w s wartość 0, tzn.

$$f(s) = 0.$$

Jeśli $f \in K[x]$ jest wielomianem stopnia n, to równanie $f = 0$ nazywamy RÓWNANIEM WIELOMIANOWYM STOPNIA n o współczynnikach w K.

Definicja 3.6: Pierwiastek stopnia n

Niech K będzie ciałem i niech $w \in K$. Pierwiastki wielomianu

$$x^n = w,$$

o ile istnieją, nazywamy PIERWIASTKAMI STOPNIA n z liczby w.

Przykłady.

- Równanie $x^2 - 2 = 0$ nie ma rozwiązań w ciele $\mathbb{Z}_3$, ponieważ kwadrat żadnej liczby całkowitej nie daje reszty 2 z dzielenia przez 3. A zatem w ciele $\mathbb{Z}_3$ nie ma pierwiastka stopnia 2 z liczby 3.
- Równanie $x^2 - 1 = 0$ ma dwa pierwiastki w ciele $\mathbb{Z}_3$, czyli $x_1 = 1$ oraz $x_2 = 2$. Oznacza to, że pierwiastek kwadratowy nie jest funkcją jednoznaczną w ciele $\mathbb{Z}_3$.
- Są dokładnie 3 liczby zespolone, które podniesione do potęgi 3 dają 2:

  $$\sqrt[3]{2}(\cos 0 + i \sin 0), \quad \sqrt[3]{2}(\cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3}), \quad \sqrt[3]{2}(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3}).$$

  Są dokładnie 4 liczby zespolone, które podniesione do potęgi 4 dają i:

  $$\cos \frac{\pi}{8} + i \sin \frac{\pi}{8}, \quad \cos \frac{5\pi}{8} + i \sin \frac{5\pi}{8}, \quad \cos \frac{9\pi}{8} + i \sin \frac{9\pi}{8}, \quad \cos \frac{13\pi}{8} + i \sin \frac{13\pi}{8}.$$

Wnioskiem z wzoru Moivre'a są również następujące formuły.

Wniosek 3.2: Wzory na pierwiastki zespolone stopnia n

Jeśli $0 \neq w = |w|(\cos\theta + i \sin\theta) \in \mathbb{C}$, to pierwiastkami stopnia n z liczby w są liczby postaci:

$$\sqrt[n]{|w|}\left(\cos \frac{\theta + 2\pi k}{n} + i \sin \frac{\theta + 2\pi k}{n}\right), \quad \text{dla } k = 0, 1, \ldots, n-1.$$

W szczególności pierwiastki stopnia n z 1 to liczby:

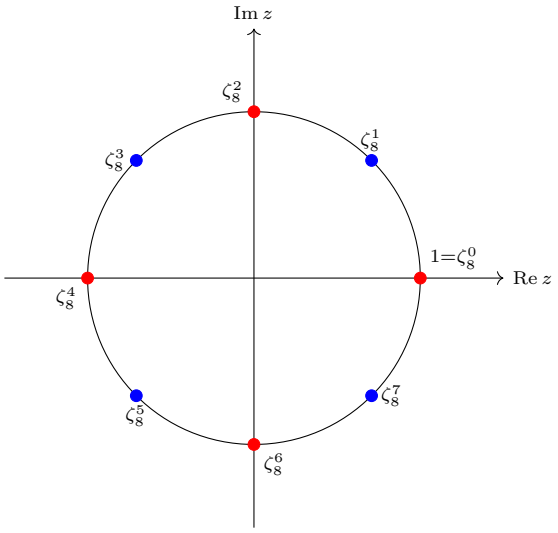
$$\zeta_n^k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \quad \text{dla } k = 0, 1, \ldots, n-1.$$

Nietrudno pokazać, że powyższe listy pierwiastków zespolonych stopnia n z liczby $0 \neq w \in \mathbb{C}$ są pełne. Jedyna liczba zespolona, która w n-tej potędze daje liczbę w to taka, której moduł równy jest $\sqrt[n]{|w|}$, a której argument po przemnożeniu przez n równy jest, z dokładnością do 2π, liczbie $\arg w$.

[2]Trzeba tu dodać trzy zastrzeżenia. Pierwsze – te działania mają sens także, gdy f, g są zerowe, przy naturalnych umowach typu $\max(-\infty, 1) = 1$, $-\infty + n = \infty$, $-\infty + -\infty = -\infty$. Druga – równość w pierwszej nierówności zachodzi wtedy (ale nie tylko wtedy), gdy $\deg(f) \neq \deg(g)$. Trzecia – jeśli K nie jest ciałem, wówczas tożsamość dla iloczynu trzeba zmodyfikować. Np. dla wielomianów $f, g \in \mathbb{Z}_4[x]$ postaci: $f = 2x, g = 1 + 2x$ mamy $\deg(f) = \deg(g) = 1$, ale $\deg(fg) = 1$.

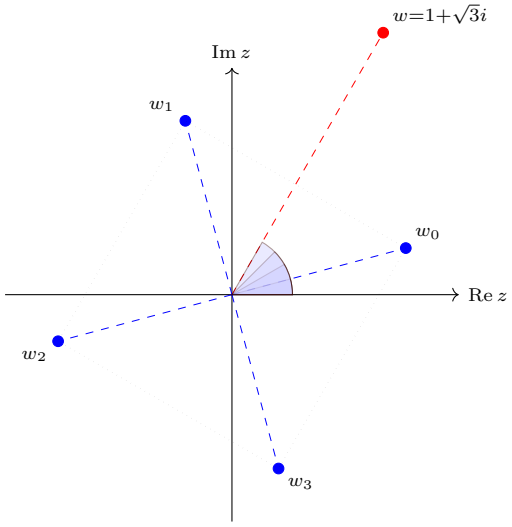
Definicja 3.7: Pierwiastek pierwotny

Mówimy, że liczba $z \in \mathbb{C}$ jest PIERWIASTKIEM PIERWOTNYM stopnia n z 1, jeśli z jest pierwiastkiem stopnia n z 1, ale nie jest pierwiastkiem z 1 stopnia m, gdzie $m < n$.



Rys. 1. Pierwiastki stopnia 8 z 1. Na niebiesko zaznaczono pierwiastki pierwotne stopnia 8 z 1.

Nietrudno widzieć, że pierwiastki stopnia n z liczby zespolonej $w \neq 0$ o argumencie θ interpretować można jako wierzchołki n-kąta foremnego.



Rys. 2. Pierwiastki stopnia 4 z liczby $1 + \sqrt{3}i$ mają moduły równe $\sqrt[4]{2}$ oraz argumenty równe odpowiednio $\frac{\pi}{12}, \frac{4\pi}{12}, \frac{7\pi}{12}, \frac{10\pi}{12}$.

Zajmiemy się teraz tematem rozkładu wielomianu na czynniki mniejszych stopni. W szkole poznali Państwo kluczowy rezultat Bezout, który formułujemy teraz w większej ogólności[3].

Twierdzenie 3.1: Bezout

Niech $f \in K[x]$. Następujące warunki są równoważne.

(1) Element $s \in K$ jest pierwiastkiem wielomianu f.

(2) Istnieje $g \in K[x]$ taki, że $f = (x - s)g$.

[3]Warto poczytać teksty dr. Bartłomieja Bzdęgi w *Delcie* o olimpijskich zastosowaniach wielomianów, m.in. tw. Bezout: `https://deltami.edu.pl/temat/matematyka/algebra/2019/11/29/Twierdzenie_B_zouta/`.

Dowód. Dowód korzysta z następującej obserwacji. Dla każdego $s \in K$ mamy

$$x^n - s^n = (x-s)(x^{n-1} + x^{n-2}s + \ldots + xs^{n-2} + s^{n-1}).$$

W szczególności, jeśli $f \in K[x]$ jest wielomianem stopnia $n > 0$, to istnieje wielomian $q \in K[x]$ stopnia $n-1$ taki, że

$$f = (x-s)q + f(s).$$

Jeśli s jest pierwiastkiem f, to mamy $f(s) = 0$ i rozkład powyżej przybiera postać $f = (x-s)q$, jak w punkcie (2). Implikacja z (2) do (1) jest jasna. Jeśli $f = (x-s)g$, dla pewnego $g \in K[x]$, to $f(s) = (s-s)g(s) = 0$. □

Prostym zastosowaniem twierdzenia Bezout jest twierdzenie mówiące, że wielomian stopnia $n > 0$ o współczynnikach w ciele ma nie więcej niż n pierwiastków (dowód to indukcja ze względu na stopień – zachęcam do jego przeprowadzenia na przykład na ćwiczeniach). Twierdzenie to jest prawdziwe w nieco ogólniejszym kontekście, tzw. przemiennych dziedzin (np. liczby całkowite).

Przejdziemy teraz do omówienia ważnej klasy ciał. Wielomiany stopnia $n > 0$ o współczynnikach w tych ciałach mają zawsze n pierwiastków.

Definicja 3.8: Ciało algebraicznie domknięte

Jeśli każdy wielomian stopnia większego od 0 o współczynnikach z ciała K ma w ciele K pierwiastek, to K nazywamy CIAŁEM ALGEBRAICZNIE DOMKNIĘTYM.

Oczywiście $\mathbb{R}$, ani tym bardziej $\mathbb{Q}$ nie jest algebraicznie domknięte, ponieważ wielomian $x^2 + 1$ nie ma pierwiastków w tych ciałach. Z twierdzenia Bezout wynika także łatwo, że żadne ciało skończone nie jest algebraicznie domknięte.

Twierdzenie 3.2

Niech K będzie ciałem. Następujące warunki są równoważne.

(1) Ciało K jest algebraicznie domknięte.

(2) Każdy wielomian stopnia > 0 o współczynnikach z K rozkłada się nad K na czynniki stopnia 1 (to znaczy: jest iloczynem wielomianów stopnia 1 o współczynnikach z K).

Dowód. Pokażemy, że ze zdania (1) wynika zdanie (2). Zakładamy zatem, że K jest ciałem algebraicznie domkniętym. Przy pomocy dowodu indukcyjnego pokażemy, że każdy wielomian stopnia > 0 rozkłada się nad K na współczynniki liniowe.

Krok bazowy indukcji jest jasny – każdy wielomian stopnia 1 da się rozłożyć na iloczyn czynników liniowych. Załóżmy prawdziwość naszego założenia dla wielomianów stopnia $n-1$. Niech f będzie wielomianem stopnia n. Ciało K jest algebraicznie domknięte, więc f ma pierwiastek $c \in K$. Stąd

$$f(x) = (x-c) \cdot g(x),$$

dla pewnego wielomianu $g \in K[x]$, na mocy twierdzenia Bezout. Wielomian g jest zatem stopnia $n-1$, więc z założenia indukcyjnego g jest iloczynem czynników stopnia 1 o współczynnikach z K. Stąd f jest iloczynem czynników stopnia 1 o współczynnikach z K.

Na odwrót: jeśli f jest iloczynem wielomianów stopnia 1 o współczynnikach w K, to każdy taki czynnik stopnia 1 ma pierwiastek w K, będący też pierwiastkiem wielomianu f. A zatem (2) implikuje (1). □

Wskazywanie ciał algebraicznie domkniętych zwykle jest skomplikowane, poza przypadkiem poznanym przez nas na ostatnim wykładzie. Jedną z kluczowych motywacji rozważania ciała liczb zespolonych jest następujący wynik.

Twierdzenie 3.3: Zasadnicze Twierdzenie Algebry Gaussa, 1799

Ciało $\mathbb{C}$ jest algebraicznie domknięte.

W tym momencie nie mamy narzędzi do przedstawienia dowodu tego twierdzenia[4]. Stosunkowo elementarny dowód będą Państwo (teoretycznie) w stanie przeprowadzić po pierwszym semestrze zajęć z Analizy. Na wyższych latach studiów poznacie Państwo krótkie (m.in. algebraiczne) dowody tego rezultatu. Przedstawimy rezultat korzystający z ZTA dotyczący rozkładów wielomianów rzeczywistych na czynniki.

Obserwacja 3.1

Niech w będzie wielomianem stopnia większego od 0 o współczynnikach rzeczywistych, traktowanych jako liczby zespolone. Wówczas jeśli liczba zespolona z jest pierwiastkiem wielomianu w, to również liczba $\overline{z}$, sprzężona do z, jest pierwiastkiem wielomianu w.

Dla przykładu, można sprawdzić, że wielomian z^2-2z+5 o współczynnikach rzeczywistych ma pierwiastek zespolony $1+2i$. Jak się okazuje, pierwiastkiem tego wielomianu jest również $1-2i$ i mamy:

$$z^2-2z+5=(z-1-2i)(z-1+2i).$$

Z drugiej strony wielomian z $\mathbb{C}[x]$, którego współczynniki nie są zespolone, nie musi (i nie spełnia) własności wyżej. Wielomian $x-i$ ma pierwiastek w $\mathbb{C}$ równy i, natomiast $-i$ nie jest jego pierwiastkiem.

Dowód. Istotnie, niech s będzie pierwiastkiem wielomianu $w=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$, gdzie $a_n,\ldots,a_0\in\mathbb{R}$. Korzystamy z tego, że sprzężenie liczby rzeczywistej jest tą liczbą oraz z formuł

$$\overline{z_1+z_2}=\overline{z_1}+\overline{z_2},\quad \overline{z_1z_2}=\overline{z_1}\cdot\overline{z_2},$$

prawdziwych dla dowolnych $z_1,z_2\in\mathbb{C}$. Mamy:

$$0=\overline{0}=\overline{a_ns^n+\ldots+a_1s+a_0}=\overline{a_n}\overline{s^n}+\ldots+\overline{a_1s}+\overline{a_0}=a_n(\overline{s})^n+\ldots+a_1\overline{s}+a_0.$$

Wniosek 3.3

Każdy wielomian w stopnia większego od 0 o współczynnikach rzeczywistych rozkłada się na iloczyn wielomianów stopnia pierwszego i stopnia drugiego o współczynnikach rzeczywistych.

Dowód. Jeśli wielomian w jest stopnia pierwszego lub drugiego, to nie ma czego dowodzić. Dalsze rozumowanie jest indukcją ze względu na stopień n wielomianu. Załóżmy, że $\deg(w)>2$. Jeśli $r_0\in\mathbb{R}$ jest pierwiastkiem w, to z lematu Bezout $w=(x-r_0)g$, gdzie $g\in\mathbb{R}[x]$ i teza wynika z założenia indukcyjnego zastosowanego do wielomianu g. Jeśli w nie ma pierwiastków rzeczywistych, to postępowanie jest następujące. Bierzemy pierwiastek $z_0\in\mathbb{C}\setminus\mathbb{R}$ wielomianu w, który musi istnieć na mocy ZTA. Wówczas na mocy poprzedniej obserwacji $\overline{z_0}$ też jest także pierwiastkiem w. Skoro z_0 i $\overline{z_0}$ to różne pierwiastki w dostajemy, że iloczyn $(x-z_0)(x-\overline{z_0})\in\mathbb{R}[x]$ jest dzielnikiem stopnia 2 wielomianu w. □

Powyższe twierdzenie nie ma zastosowania do ciała $\mathbb{Q}$. Wielomian $x^4+2\in\mathbb{Q}[x]$ nie ma żadnego rozkładu na czynniki niższego stopnia niż 4 w $\mathbb{Q}[x]$. Warto w kontekście ciała $\mathbb{Q}$ pamiętać szkolne twierdzenie o wymiernych pierwiastkach wielomianu (i umieć je udowodnić, co jest raczej nietrudnym ćwiczeniem). Problem rozkładalności wielomianu na czynniki (nierozkładalne) jest ważnym zagadnieniem, nie tylko na naszym przedmiocie czy w różnych działach matematyki, ale także w jej (praktycznych) zastosowaniach.

Należy podkreślić, że jedną z podstawowych motywacji historycznych rozważania liczb zespolonych było rozwiązywanie równań wielomianowych stopni 3, 4 oraz wyższych. W XVI wieku opracowane zostały wzory i metody wyznaczania pierwiastków wielomianów stopni 3 i 4, których znajomość nie należy do umiejętności wymaganych w tym kursie. Kwestią nieistnienia ogólnych wzorów na pierwiastki wyższych stopni (używających współczynników wielomianu i podstawowych operacji algebraicznych) rozstrzygnięty został dopiero w pierwszej połowie XIX wieku przez Ruffiniego, Abeta i Galois, czemu poświęcony jest między innymi kurs Algebry 2 (wymagający między innymi narzędzi teorii grup).

[4]Istnieje dowód oparty o wyniki algebry liniowej, ale wymaga znajomości szeregu pojęć i faktów z drugiego semestru.

3.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wyznaczanie pierwiastków stopnia n, wskazywanie pierwiastków pierwotnych z jedynki)
   Znajdź wszystkie pierwiastki, wyznacz ich postaci ogólne (nie tylko trygonometryczne)

   (a) stopnia 4 z liczby -4,
   (b) stopnia 4 z liczby $-\sqrt{3}+3i$,
   (c) stopnia 6 z liczby $-27i$,
   (d) stopnia 3 z liczby $5+5i$.

   Znajdź liczbę pierwiastków pierwotnych stopnia 12 z jedynki i wypisz ich postaci trygonometryczne.

2. Znajdź sumę oraz iloczyn wszystkich pierwiastków stopnia n z jedynki oraz iloczyn wszystkich pierwiastków pierwotnych stopnia n z 1.

3. Niech $\{1, z_1, \ldots, z_{2022}\}$ będzie zbiorem zespolonych pierwiastków stopnia 2023 z jedynki. Wykaż, że
   $$(1-z_1)(1-z_2)(1-z_3)\ldots(1-z_{2022}) = 2023.$$
   Jaka jest interpretacja geometryczna tej równości?

4. (♠ Rozkład wielomianu o współczynnikach rzeczywistych na czynniki stopnia $\leqslant 2$)
   Dla każdego z poniższych wielomianów znajdź jego rozkład na czynniki będące wielomianami rzeczywistymi stopnia $\leqslant 2$.

   (a) x^4-2x^2+4,
   (b) x^6+x^2,
   (c) x^7-x,
   (d) x^4+4,
   (e) $x^7+8x^4+4x^3+32$,
   (f) x^6+27,
   (g) x^6-x^3+1.

5. Znajdź rozkład wielomianu[5] rzeczywistego $x^4-4x^3+2x^2+4x+4$ na czynniki rzeczywiste stopnia 2.
   Wskazówka: skorzystaj z podstawienia $x=t+1$.

6. (♠ Znajdowanie pierwiastków równań wielomianowych o współczynnikach w ciele $\mathbb{Z}_p$)
   Znajdź pierwiastki wielomianów

   - $x^5-x \in \mathbb{Z}_5[x]$,
   - $x^3-1 \in \mathbb{Z}_7[x]$,
   - $x^3-2 \in \mathbb{Z}_7[x]$.

7. Znajdź wszystkie wspólne pierwiastki wielomianów $10x^{15}+9x^2+1$ oraz $10x^{15}+8x^2+2$ o współczynnikach w ciele $\mathbb{Z}_{19}$.

8. (♠ Wyznaczanie pierwiastków zespolonych prostych równań wielomianowych)
   Rozłóż na czynniki liniowe wielomiany o współczynnikach zespolonych

   (a) $z^2+5-12i$,
   (b) $(1+i)z^3+(3-5i)z^2-6z$,
   (c) z^4+2z^2+2.

9. (♠ Sprzężenie pierwiastka nierzeczywistego wielomianu z $\mathbb{R}[x]$ jest też pierwiastkiem)
   Wyznacz wszystkie zespolone pierwiastki wielomianu $z^4-6z^3+18z^2-30z+25$, wiedząc że jednym z nich jest $2-i$.

10. Liczba $\cos\phi+i\cdot\sin\phi$ spełnia równanie $z^n+a_1z^{n-1}+\ldots+a_n=0$, gdzie $a_1, a_2, \ldots, a_n \in \mathbb{R}$. Wykaż, że zachodzi równość
    $$a_1\sin\phi+a_2\sin 2\phi+\ldots+a_n\sin n\phi=0.$$

[5] W XVIII wieku, przed dowodem Zasadniczego Twierdzenia Algebry autorstwa Gaussa, N. Bernoulli twierdził, że jest to kontrprzykład do twierdzenia o rozkładzie na czynniki liniowe i kwadratowe. Rozkład znalazł jednak w 1742 r. L. Euler.

3.3 Uzupełnienie. Kilka faktów o pierwiastkach wielomianów

W kontekście rozwiązywania równań wielomianowych i zastosowań warto przypomnieć/uogólnić kilka faktów szkolnych, opowiadając o nich w nieco szerszym kontekście.

Twierdzenie 3.4: Wzory Viete'a

Niech K będzie ciałem oraz niech $x_1, x_2, x_3, \ldots, x_n$ będą pierwiastkami wielomianu

$$f = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in K[x],$$

gdzie $a_n \neq 0$ (tzn. $\deg f = n > 0$). Wówczas zachodzą równości[a]:

$$\begin{cases} \sum_{i=1}^{n} x_i & = x_1 + x_2 + \ldots + x_n & = \frac{-a_{n-1}}{a_n} \\ \sum_{1\leqslant i<j\leqslant n} x_i x_j & = x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n & = \frac{a_{n-2}}{a_n} \\ \sum_{1\leqslant i<j<k\leqslant n} x_i x_j x_k & = x_1x_2x_3 + x_1x_2x_4 + \ldots + x_{n-2}x_{n-1}x_n & = \frac{-a_{n-3}}{a_n} \\ & & \vdots \\ x_1x_2x_3\ldots x_n & & = (-1)^n \frac{a_0}{a_n}. \end{cases}$$

[a] Nie wypisujemy dokładnie równości dotyczących współczynników stojących przy kolejnych potęgach. Po lewych stronach stoją tzw. elementarne wielomiany symetryczne stopnia n od pierwiastków wielomianu.

Dowód. Indukcja ze względu na n. Dla $n = 1$ dowód jest oczywisty. Załóżmy, że dla każdego wielomianu stopnia n, wzory te są prawdziwe. Rozważmy wielomian stopnia $n+1$, o pierwiastkach: $x_1, x_2, \ldots, x_n, x_{n+1}$. Zgodnie z twierdzeniem Bezout istnieje wielomian $g(x)$ (którego wiodący współczynnik to 1), taki, że: $f(x) = a_{n+1} \cdot (x - x_1) \cdot g(x)$. Wielomian g jest stopnia n, i jego pierwiastkami są $x_2, x_3, \ldots, x_n, x_{n+1}$.Więcej pierwiastków, zgodnie z udowodnionym wcześniej faktem, mieć nie może. Zatem są to jego wszystkie pierwiastki. Z założenia indukcyjnego mamy zatem:

$$g(x) = x^n - (x_2 + x_3 + \ldots + x_{n+1})x^{n-1} + \ldots + (-1)^{n+1}(x_2x_3\ldots x_{n+1}).$$

Wymnażając g w takiej postaci przez $a_{n+1} \cdot (x - x_1)$ dostajemy tezę. □

Przyjrzyjmy się nietrywialnemu zadaniu wykorzystującemu wzory Viete'a i Obserwację 3.1.

Zadanie. Wielomian $w(x) = x^4 + ax^3 + bx^2 + cx + d$ ma współczynniki rzeczywiste oraz pierwiastki nierzeczywiste z_1, z_2, z_3, z_4. Wiadomo, że $z_1z_2 = 13 + i$ oraz $z_3 + z_4 = 3 + 4i$. Wyznacz a, b, c, d.

Rozwiązanie. Wielomian w jest stopnia 4, a zatem z_1, z_2, z_3, z_4 są wszystkimi jego pierwiastkami.

Skoro $z_1 \notin \mathbb{R}$, to jedna z liczb z_2, z_3, z_4 musi być sprzężonym do z_1 pierwiastkiem w. Jednak dla każdego $z \in \mathbb{C}$ mamy $z\overline{z} = |z|^2 \in \mathbb{R}$, a zatem $\overline{z_1} \neq z_2$, bo $z_1z_2 \notin \mathbb{R}$. Analogicznie $\overline{z_3} \neq z_4$. Stąd $\{\overline{z_1}, \overline{z_2}\} = \{z_3, z_4\}$ (nieco dokładniej: po podzieleniu w przez $(z - z_1)(z - \overline{z_1})$ mamy wielomian, którego pierwiastkami są $z_2, \overline{z_2}$). A zatem mamy:

$$z_3z_4 = \overline{z_1z_2} = 13 - i \quad \text{oraz} \quad z_1 + z_2 = \overline{z_3 + z_4} = 3 - 4i.$$

Ze wzorów Viete'a dostajemy zatem:

$$\begin{aligned} a &= -(z_1 + z_2 + z_3 + z_4) = -(3 + 4i + 3 - 4i) = -6 \\ b &= z_1z_2 + z_1z_3 + z_1z_4 + z_2z_3 + z_2z_4 + z_3z_4 = \\ &= (z_1 + z_2)(z_3 + z_4) + z_1z_2 + z_3z_4 = (3 + 4i)(3 - 4i) + (13 + i) + (13 - i) = 51 \\ c &= -(z_1z_2z_3 + z_1z_2z_4 + z_1z_3z_4 + z_2z_3z_4) = \\ &= -((z_1z_2)(z_3 + z_4) + (z_1 + z_2)(z_3z_4)) = -((13 + i)(3 + 4i) + (3 - 4i)(13 - i)) = -70 \\ d &= z_1z_2z_3z_4 = (13 + i)(13 - i) = 170. \end{aligned}$$

■

Poświęcimy teraz trochę miejsca wielomianom o współczynnikach całkowitych[6]. Przypomnijmy najpierw szkolny rezultat.

Twierdzenie 3.5: O wymiernych pierwiastkach wielomianu o współczynnikach w $\mathbb{Z}$

Załóżmy, że liczby $a_0, a_1, \ldots, a_n$ są całkowite, $a_n \neq 0$, $n \neq 1$ oraz

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

dla pewnej liczby x. Jeśli $x = \frac{p}{q}$ i liczby całkowite p, q są względnie pierwsze, to

$$p \,|\, a_0 \quad \text{oraz} \quad q \,|\, a_n.$$

Dowód. Pomnóżmy równość

$$a_0 + a_1 \frac{p}{q} + a_2 \left(\frac{p}{q}\right)^2 + \ldots + a_n \left(\frac{p}{q}\right)^n = 0$$

przez q^n otrzymując:

$$a_0 q^n + a_p q^{n-1} + a_2 p^2 q^{n-1} + \ldots + a_{n-1} p^{n-1} q + a_n p^n = 0.$$

Zauważmy, że:

$$q \,|\, a_n p^n = -q(a_0 q^{n-1} + a_1 p q^{n-2} + a_2 p q^{n-3} + \ldots + a_{n-1} p^{n-1}),$$

a ponieważ liczby p, q są względnie pierwsze, to p^n, q też są względnie pierwsze, a z tego wynika, że q jest dzielnikiem a_n. W ten sam sposób widzimy, że $p \,|\, a_0 p^n$, co oznacza, że p jest dzielnikiem a_0. □

Powyższe twierdzenie mówi między innymi o całkowitych pierwiastkach wielomianów o współczynnikach całkowitych. Może ono służyć do wyznaczania rozkładów wielomianów (jak w szkole). Warto jednak przyjrzeć się tej własności w nieco ogólniejszym kontekście, rozszerzając nieco pojęcie liczby całkowitej.

Rozważmy liczby zespolone $z_1 = a + bi$, $z_2 = c + di$ oraz $z_3 = e + fi$, gdzie $a, b, c, d, e, f \in \mathbb{Z}$ i załóżmy, że:

$$(a + bi) = (c + di)(e + fi).$$

Oczywiście mamy stąd, że $|z_1| = |z_2| \cdot |z_3|$, a więc $a^2 + b^2$ jest wielokrotnością $c^2 + d^2$ oraz $e^2 + f^2$. A zatem jeśli rozważymy liczby zespolone, których części: rzeczywista i urojona są liczbami całkowitymi, wówczas zachodzić się zdają pewne związki pomiędzy rozkładem tych liczb na czynniki, a rozkładem na czynniki kwadratów ich modułów – czyli zwykłych liczb całkowitych. Ta prosta obserwacja ma, jak się okazuje, niezwykle daleko idące konsekwencje. Zacznijmy od banalnego zastosowania. Rozwiążmy zadanie.

Zadanie. Niech a, b oraz n będą liczbami naturalnymi. Udowodnić, że istnieją liczby całkowite x, y, dla których zachodzi równość

$$(a^2 + b^2)^n = x^2 + y^2.$$

Jest to zadanie egzaminacyjne (na ocenę celującą) z książki *Algebra liniowa 1 Kolokwia i egzaminy*, autorstwa M Gewerta i Z. Skoczylasa, Nie jestem pewien czy Czytelnik od razu wskazałby rozwiązanie bez użycia liczb zespolonych. Tymczasem stosując argumentację podaną wyżej naszą równość przepisujemy do postaci:

$$(a + bi)^n (a - bi)^n = (x + yi)(x - yi).$$

A zatem rozwiązanie, to $x = \mathrm{Re}(a + bi)^n$ oraz $y = \mathrm{Im}(a + bi)^n$. Są to oczywiście liczby naturalne.

Zadanie. Rozwiązać w liczbach zespolonych równanie: $z^4 - 6z^3 + 18z^2 - 30z + 25 = 0$.

W tym przypadku możliwe są różne podejścia, na przykład korzystając z twierdzenia z wykładu można argumentować, że istnieją $a, b, c, d \in \mathbb{R}$, że:

$$z^4 - 6z^3 + 18z^2 - 30z + 25 = (z^2 + az + b)(z^2 + cz + d).$$

[6]Opieram się o tekst dr. Michała Krycha: .Po co komu wymierność?, dostępny pod adresem: `https://www.mimuw.edu.pl/~krych/odczyty/18-09-13-warszawa.pdf`.

Trzeba zatem rozwiązać układ równań:

$$a+c=-6, \quad b+d+ac=18, \quad ad+bc=-30, \quad 25=bd.$$

W tym przypadku akurat kładąc $b=d=5$ dostajemy układ $a+c=-6, ac=8, 5(a+c)=-30$, co daje $a^2+6a+8=(a+4)(a+2)=0$, czyli $a=-4, c=-2$ (lub odwrotnie). A zatem:

$$z^4-6z^3+18z^2-30z+25=(z^2-2z+5)=0.$$

Widać, że te równania kwadratowe „da" się dalej rozwiązać. Udało się. Czy można było to zrobić inaczej? Może, ale będzie trzeba trochę „gdybać". Spróbujmy. „Gdyby" istniał pierwiastek postaci $z=a+bi$, gdzie a, b są **całkowite**, również $a-bi$ byłby pierwiastkiem, a więc mielibyśmy $25=(a^2+b^2)z_3z_4$, gdzie z_3, z_4 to pozostałe pierwiastki. „Gdyby" jeszcze z_3, z_4 również miały całkowite części rzeczywiste i urojone, wówczas a^2+b^2 byłaby dzielnikiem 25. To teoretycznie nie musi się zdarzyć (nie mamy narzędzi, by stwierdzić czy tak musi być), ale niewiele jest liczb całkowitych a, b, że a^2+b^2 dzieli 25, więc warto „pogdybać". Może jednym z pierwiastków jest liczba „całkowita" $a+bi$ postaci:

$$\pm 1, \quad \pm i, \quad \pm 5, \quad \pm 5i, \quad \pm 2 \pm i, \quad \pm 1 \pm 2i, \quad \pm 5 \pm 5i.$$

Nietrudno sprawdzić, że wśród tych liczb właśnie liczby $2 \pm i$ oraz $1 \pm 2i$ są rozwiązaniami naszego równania. To rodzi rozmaite domysły: czy przypadkiem nie mamy tu do czynienia z jakąś wersją twierdzenia o pierwiastku całkowitym/wymiernym wielomianu o współczynnikach całkowitych? Tak rzeczywiście jest i wiąże się to z faktem, że podzbiór liczb zespolonych postaci

$$\mathbb{Z}[i]=\{a+bi \,|\, a, b \in \mathbb{Z}\},$$

czyli tzw. **pierścień liczb całkowitych Gaussa**, ma jednoznaczność rozkładu na czynniki pierwsze. Co to znaczy? Czym są te czynniki? Czytelnik zainteresowany bliższym poznaniem tych liczb i związanej z nimi teorii podzielności zechce zajrzeć do:

M. Krych: *Skąd się wzięła liczba i*, `https://smp.uph.edu.pl/msn/34/krych.pdf`.

W tekście tym znajdują się informacje nie tylko o liczbach całkowitych Gaussa, ale też o tzw. **liczbach całkowitych Eisensteina** $\mathbb{Z}[\omega_3]$, złożonych z liczb postaci $a+b\omega_3$, gdzie $a, b \in \mathbb{Z}$ oraz ω_3 jest nierzeczywistym pierwiastkiem stopnia 3 z 1. Również w tym zbiorze zachodzi teoria podzielności, a nawet twierdzenie o jednoznacznym rozkładzie... Dlaczego ten jednoznaczny rozkład jest tak istotny?

Prawie 400 lat temu Pierre de Fermat stwierdził, że znalazł „niezwykły dowód" następującego twierdzenia:

Twierdzenie 3.6: Wielkie Twierdzenie Fermata

Równanie diofantyczne: $x^n+y^n=z^n$, gdzie x, y, z, n są niezerowymi liczbami całkowitymi, nie ma rozwiązań, dla $n>2$.

Niestety, Fermat nie był w stanie przedstawić rozwiązania, ponieważ swoje odkrycie zapisał na marginesie kopii starożytnego dzieła *Arytmetyka* Diofantosa. Stwierdził jedynie, że *margines jest zbyt mały, by pomieścić dowód*. Notatka Fermata stała się jedną z wielu nieudowodnionych obserwacji, zostawionych kolejnym pokoleniom. Jak się okazało, wiele przypuszczeń Fermata zostało z czasem rozstrzygniętych. Jedną z osób, która poświęciła im sporo miejsca był sam Euler. Nie był on jednak w stanie pokazać ogólnego dowodu powyższego rezultatu. Z trudem znalazł niełatwe uzasadnienie dla $n=3$ (używając liczb zespolonych, o czym można przeczytać w tekście dr. Krycha). Problem stał się jednym z najsłynniejszych w historii matematyki, a także źródłem rozwoju licznych jej dziedzin. Twierdzenie Fermata zostało udowodnione dopiero w 1994 roku przez Andrew Wilesa.

Przez stulecia szukano błyskotliwych, krótkich dowodów hipotezy Fermata. Jedna z takich nieudanych prób warta jest jednak przypomnienia, ponieważ dała początek rozwojowi współczesnej teorii liczb. Cofnijmy się do roku 1847. Problem Fermata był już wówczas jednym z największych wyzwań matematycznych. Centrum matematycznego świata wciąż jeszcze leżało w Paryżu (niedługo potem trafić miało do Getyngi, a potem za Ocean Atlantycki). Francuska Akademia Nauk oferowała (od 31 lat) złoty medal i nagrodę 3000 franków za rozwiązanie problemu Fermata. Na posiedzeniu 1 marca, z propozycją dowodu wystąpił znany matematyk Gabriel Lamé. Twierdził, że znalazł cudowne rozwiązanie, bardzo krótkie. Idea dowodu była rzeczywiście niezwykle prosta, a dla jej przedstawienia oprzemy się na następującej obserwacji (związanej bezpośrednio z treścią wykładu).

Obserwacja 3.2

Niech $n \geqslant 1$ będzie liczbą całkowitą oraz niech $\zeta = \cos\frac{2pi}{n} + i \cdot \sin\frac{2\pi}{n}$ będzie pierwiastkiem pierwotnym stopnia n z 1. Wówczas

$$z^n - 1 = \prod_{k=0}^{n-1}(z - \zeta^k) = (z-1)(z-\zeta)(z-\zeta^2)\ldots(z-\zeta^{n-1}).$$

Dowód jest oczywisty, bowiem dla $0 \leqslant k \leqslant n-1$ liczba ζ^k podniesiona do potęgi n równa jest 1, zaś wszystkie te liczby są parami różne.

Wróćmy do argumentu Lame. Przedstawmy $x^n + y^n = z^n$ jako iloczyn n czynników „całkowitych" na dwa sposoby. Jak? Weźmy $\zeta \in \mathbb{C}$ takie, że $\zeta^n = 1$, $\zeta \neq 1$ oraz n – nieparzyste. Dostaniemy wówczas:

$$x^n + y^n = (x+y)\left(x+\zeta y\right)\left(x+\zeta^2 y\right)\cdots\left(x+\zeta^{n-1}y\right) = z \cdot z \cdot \ldots \cdot z$$

Formuła ta wynika natychmiast z powyższej obserwacji, dla $z = -x/y$. Co tu widzimy? Lamé wysnuwa stąd wniosek, że $x+y$ oraz z mają wspólny dzielnik, co prowadziłoby do sprzeczności. Opiera się przy tym na „względnej pierwszości" czynników uzyskanego rozkładu (można się do niej ograniczyć).

Nie tylko Lamé był niezwykle przejęty zaproponowanym dowodem. Również Cauchy wystąpił i stwierdził, że od dłuższego czasu pracuje nad dowodem, w zasadzie opartym na analogicznych obserwacjach. Obydwoje część „zasługi" oddawali Josephowi Liouvillowi, który zasugerował im rozważanie liczb zespolonych w kontekście problemu Fermata. Paradoksalnie, to właśnie Liouville zwrócił uwagę na pewien problem. Zaproponowany wyżej rozkład wyrażenia $x^n + y^n$ na „czynniki względnie pierwsze" dokonuje się w zbiorze liczb $\mathbb{Z}[\zeta]$ postaci:

$$a_1 + a_2\zeta + a_3\zeta^2 + \ldots + a_{n-1}\zeta^{n-1}, a_i \in \mathbb{Z}.$$

Nie ma gwarancji, że w zbiorze tym zachodzi jednoznaczność rozkładu na czynniki. Gdyby jej nie było, wówczas wyciągnięcie wniosku, że każdy czynnik $x^n + y^n$ jest n-tą potęgą nie jest możliwe... Do tego momentu Czytelnik ma prawo być już poważnie zniecierpliwiony: ani nie powiedzieliśmy czym jest „całkowitość" w $\mathbb{C}$, ani czym są czynniki pierwsze, nierozkładalne czy względnie pierwsze w $\mathbb{Z}[\zeta]$. Jeśli tak jest, to być może osiągnąłem swój cel. Dokładny opis tego problemu przekracza ów skromny dodatek, ale jest absolutnie w zasięgu. Proszę jedynie o kontynuowanie lektury przy bardziej kompetentnym źródle: artykule prof. Balcerzyka i dr. Szurka: „*Nieco historii matematyki w wykładzie algebry*":`http://www.deltami.edu.pl/temat/matematyka/2016/05/30/1981-05-Fermat.pdf`.

Na koniec warto dodać jeszcze jeden komentarz dotyczący wielomianów $x^n - 1$, które pojawiły się po drodze. Można zapytać: jak wyglądają rozkłady tych wielomianów na iloczyny wielomianów niższych (dodatnich) stopni o współczynnikach całkowitych? Nie jest to proste zagadnienie. Gdybyśmy rozkładali na wielomiany o współczynnikach rzeczywistych, to nie ma żadnego problemu, bo $\overline{\zeta^k} = \zeta^{n-k}$, co oznacza, że $x^n - 1$ ma czynniki liniowe postaci $(x-1)$ (zawsze), $(x+1)$ (jeśli n jest liczbą parzystą) oraz (dla $n > 2$) czynniki kwadratowe (o współczynnikach rzeczywistych) postaci

$$x^2 - 2x\cos\frac{2k\pi}{n} + 1 = (x-\zeta^k)(x-\zeta^{n-k}),$$

gdzie $k \notin \{0, \frac{n}{2}\}$. Jeśli jednak ograniczymy się tylko do czynników w $\mathbb{Z}[x]$, to np. dla $n = 15$ mamy:

$$x^{15} - 1 = (x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)(x^8-x^7+x^5-x^4+x^3-x+1).$$

Czym są cztery czynniki, występujące w tym rozkładzie? Są to tak zwane: pierwszy, trzeci, piąty i piętnasty WIELOMIAN CYKLOTOMICZNY. Z definicji, pierwiastkami n-tego wielomianu cyklotomicznego są pierwiastki pierwotne stopnia n z jedynki. Krótko mówiąc: stopnie tych wielomianów wynoszą odpowiednio: $1, 2, 4, 8$, bowiem wśród pierwiastków stopnia 15 z 1 jest jeden pierwiastek pierwotny stopnia 1, dwa pierwiastki pierwotne stopnia 2, cztery pierwiastki pierwotne stopnia 5 oraz 8 pierwiastków pierwotnych stopnia 15. Jak się okazuje, wielomiany cyklotomiczne są nierozkładalne na iloczyn wielomianów niższych (dodatnich) stopni o współczynnikach w $\mathbb{Z}$.

Temat wielomianów cyklotomicznych jest niezwykle pięknym i ciekawym fragmentem teorii wielomianów będącym na styku algebry i teorii liczb. Czytelnika zainteresowanego tym zagadnieniem odsyłam do tekstu prof. Andrzeja Nowickiego: `https://www-users.mat.umk.pl//~anow/imperium/wlm12.pdf`.

3.4 Dodatek. Jednoznaczność rozkładu na czynniki w $K[x]$

Jeszcze w szkole mogliśmy zauważyć podobieństwo pomiędzy teorią podzielności w zbiorze liczb całkowitych oraz w zbiorze wielomianów (o współczynnikach rzeczywistych). Mówimy bowiem o podzielności, algorytmie dzielenia z resztą, a także o największym wspólnym dzielniku czy rozkładzie na czynniki.

Twierdzenie 3.7: O dzieleniu z resztą

Niech f, g będą wielomianami o współczynnikach z ciała K. Załóżmy ponadto, że g nie jest wielomianem zerowym. Wówczas istnieją wielomiany q i r takie, że:

$$g = q \cdot g + r, \quad \deg(r) < \deg(g). \tag{3.1}$$

Ponadto wielomiany q, r są wyznaczone jednoznacznie.

Twierdzenie to jest skądinąd dobrze znane – choć zapewne nie dla ciał (i warto pod tym kątem obserwować dowód) – ale jest ono ważne jako fakt typu: *dla pewnych obiektów istnieją inne obiekty i są one wyznaczone jednoznacznie*. Warto dobrze zrozumieć to zagadnienie, zwłaszcza problem owej **jednoznaczności**.

Dowód. Pokażemy najpierw, że dla każdej pary $f, g \in K[x]$ takiej, że $g \neq 0$ istnieje para q, r taka, że zachodzi (3.1). Później wykażemy, że wielomiany q, r są wyznaczone jednoznacznie.

Niech $f, g \in K[x]$, $g \neq 0$. Zauważmy, że jeśli $\deg(g) > \deg(f)$, to za szukane wielomiany q, r można wziąć $q = 0$ oraz $r = f$. Wówczas oczywiście $\deg(r) = \deg(f) < \deg(g)$.

Załóżmy dalej, że $\deg(f) \geqslant \deg(g)$. Dowód istnienia wielomianów q, r spełniających (3.1) jest indukcją ze względu na $\deg(f)$. Z założenia $\deg(f) \geqslant 0$, a zatem w bazowym kroku indukcji rozważamy sytuację, gdy $\deg(f) = 0$. Skoro $g \neq 0$, to $\deg(g) = 0$ i wystarczy wziąć $q = f/g$ oraz $r = 0$. Wtedy $\deg(r) < \deg(g)$.

Przechodzimy wreszcie do kroku indukcyjnego. Niech $n = \deg(f)$ oraz $m = \deg(g) \leqslant n$. Niech:

$$f = a_0 + a_1 x + \ldots + a_n x^n, \quad g = b_0 + b_1 x + \ldots + b_m x^m.$$

Definiujemy **nowy wielomian** (dla niektórych tylko nowy, bo przecież Czytelnicy dodatku do wykładu pierwszego zobaczą natychmiast, że to jest właśnie S-wielomian $S(f, g)$ – tylko dla jednej zmiennej):

$$\tilde{f} = b_m \cdot f - a_n x^{n-m} \cdot g.$$

Nietrudno widzieć, że odejmujemy od siebie dwa wielomiany stopnia n o tym samym współczynniku wiodącym. Wielomian $\tilde{f}$ ma zerowy współczynnik przy x^n, a zatem $\deg(\tilde{f}) \leqslant n-1$. Z założenia indukcyjnego zastosowanego do $\tilde{f}$ wynika, że istnieją wielomiany $\tilde{q}$ oraz $\tilde{r}$ takie, że $\tilde{f} = g\tilde{q} + \tilde{r}$, $\deg(g) > \deg(\tilde{r})$ oraz:

$$b_m f - a_n x^{n-m} g = g\tilde{q} + \tilde{r}.$$

W szczególności mamy też (tu się w sposób istotny wykorzystuje założenie, że K jest ciałem!):

$$f = g\left(\frac{a_n x^{n-m} + \tilde{q}}{b_m}\right) + \frac{\tilde{r}}{b_m}.$$

A zatem dla pary wielomianów f, g definiujemy szukane wielomiany q, r jako $q := \frac{a_n x^{n-m} + \tilde{q}}{b_m}$ oraz $r := \frac{\tilde{r}}{b_m}$. Oczywiście spełnione jest założenie $\deg(g) > \deg(r) = \deg(\tilde{r})$. Krok indukcyjny jest zatem zakończony.

Pozostaje pokazać jednoznaczność istnienia wielomianów q, r spełniających (3.1) dla danej pary wielomianów f, g. Będzie to (jak zwykle w takich problemach) rozumowanie nie wprost. Załóżmy, że dla pewnej pary f, g wielomianów istnieją wielomiany q, q', r, r' takie, że $\deg(g) > \deg(r)$, $\deg(g) > \deg(r')$ oraz

$$f = qg + r = q'g + r'.$$

Wynika stąd, że:

$$(q - q')g = r' - r.$$

Załóżmy (wbrew tezie o jednoznaczności rozkładu f), że $q \neq q'$. Mamy zatem $\deg(q - q')g \geqslant \deg(g)$. W rezultacie $\deg(r - r') \geqslant \deg(g)$. Mamy jednak $\deg(r' - r) \leqslant \max\{\deg(r'), \deg(r)\} < \deg(g)$ (założenie o stopniach r, r' i g). Otrzymaliśmy sprzeczność. A zatem musi zachodzić równość $q = q'$. Wtedy jednak zachodzi także równość $r = r'$. Dowód jednoznaczności przedstawienia (3.1) jest zatem zakończony. $\square$

Kluczem do dalszych zagadnień jest pojęcie największego wspólnego dzielnika wielomianów. W tym celu wprowadzimy kilka intuicyjnych określeń.

Definicja 3.9

Niech $f = a_n x^n + \ldots + a_1 x + a_0 \in K[x]$, gdzie $\deg(f) = n \geqslant 0$.

- Wielomian f nazwiemy UNORMOWANYM, jeśli $a_n = 1$.
- Powiemy, że $0 \neq g \in K[x]$ DZIELI f, jeśli istnieje $h \in K[x]$ taki, że
$$f = g \cdot h.$$
W takim przypadku mówimy również, że g jest DZIELNIKIEM f, co zapisujemy w postaci $f \,|\, g$.
- Wielomian unormowany $d \in K[x]$ nazwiemy największym wspólnym dzielnikiem[a] (NWD) układu wielomianów niezerowych $p_1, \ldots, p_n$ jeśli:
  - $d \,|\, p_i$, dla każdego $i = 1, \ldots, n$,
  - jeśli dla pewnego $h \in K[x]$ mamy $h \,|\, p_i$, dla każdego i, to $h \,|\, d$.
- Jeśli niezerowe $p_1, \ldots, p_n \in K[x]$ spełniają $NWD(p_1, \ldots, p_n) = 1$, to nazywamy je WZGLĘDNIE PIERWSZYMI.
- Jeśli $\deg(f) \geqslant 1$, to wielomian f nazwiemy NIEROZKŁADALNYM, jeśli nie istnieją (niestałe) wielomiany $g, h \in K[x]$, $\deg(f), \deg(h) \geqslant 1$ takie, że $f = g \cdot h$. Wielomian stopnia $\geqslant 1$, który nie jest nierozkładalny nazywamy ROZKŁADALNYM.

[a] W ramach wykładów z Algebry I dowiecie się Państwo, że jest to pojęcie określone z dokładnością do relacji stowarzyszenia (nie będziemy tego wyjaśniać), a więc niekoniecznie ograniczone do wielomianów unormowanych

Rozkładalność wielomianu zależy oczywiście od ciała współczynników. Na wykładzie pokazaliśmy, że każdy wielomian $f \in \mathbb{R}[x]$ stopnia co najmniej 3 jest rozkładalny. Tymczasem wielomian $x^3 + 2$ nie jest rozkładalny jako element $\mathbb{Q}[x]$. Po odpowiednim rozszerzeniu współczynników, na przykład w ciele $\mathbb{Q}[\sqrt[3]{2}]$, wielomian ten można już jednak rozłożyć na czynniki stopnia $\geqslant 2$ postaci: $x - \sqrt[3]{2}$ oraz $x^2 + \sqrt[3]{2}x + \sqrt[3]{2}x^2$.

Pierwszym krokiem do zrozumienia rozkładów wielomianów jest rezultat będący wersją lematu Bezout.

Obserwacja 3.3: Lemat Bezout

Niech K bedzie ciałem. Dla dowolnych niezerowych $p_1, \ldots, p_n \in K[x]$ istnieją wielomiany $q_1, \ldots, q_n \in K[x]$ takie, że
$$q_1 p_1 + \ldots + q_n p_n = NWD(p_1, \ldots, p_n).$$

Dowód. Niech
$$I = \{q_1 p_1 + \ldots + q_n p_n \,|\, q_i \in K[x]\} \subseteq K[x].$$
Niech d będzie wielomianem unormowanym najmniejszego możliwego stopnia należącym do I. Pokażemy, że jest to NWD wielomianów $p_1, p_2, \ldots, p_n$. Potrzeba zatem sprawdzić dwa warunki.

Zacznijmy od pokazania, że $d \,|\, p_i$, dla każdego i. Gdyby któryś z wielomianów p_i nie był podzielny przez d, to korzystając z twierdzenia o dzieleniu z resztą mamy $p_i = h_i d + r_i$, gdzie $\deg(r_i) < \deg(d)$. Ale skoro $d \in K$, to dla pewnych $q_1', \ldots, q_n' \in K[x]$ mamy
$$r_i = p_i - h_i d = p_i - (q_1' p_1 + \ldots + q_n' p_n),$$
zatem $r_i \in I$. Sprzeczność z wyborem d. Zatem d dzieli wszystkie p_i.

Druga część dowodu to pokazanie, że wspólny dzielnik wielomianów p_i jest dzielnikiem d. To jest jednak oczywiste. Zauważmy, że NWD układu p_i jest dzielnikiem każdego elementu I, a więc i jest dzielnikiem elementu d. Zatem d jest rzeczywiście NWD układu $p_1, \ldots, p_n$. □

Warto odnotować, że istotne jest założenie o tym, że wielomiany mają współczynniki w ciele. W $\mathbb{Z}[x]$ największy wspólny dzielnik 2 oraz x to 1, ale nie istnieją wielomiany $f, g \in \mathbb{Z}[x]$, że $1 = 2f(x) + xg(x)$.

Twierdzenie 3.8: O jednoznaczności rozkładu wielomianów o współczynnikach w ciele

Niech p będzie wielomianem stopnia $\geqslant 1$ w $K[x]$, gdzie K – ciało. Wówczas p można zapisać w postaci:

$$p = a \cdot q_1 \cdot \ldots \cdot q_k, \qquad (\diamondsuit)$$

gdzie a jest współczynnikiem wiodącym p oraz $q_1, \ldots, q_k$ są unormowanymi nierozkładalnymi wielomianami w $K[x]$. Co więcej, rozkład ów jest jednoznaczny z dokładnością do porządku występowania czynników.

Dowód. Dowód ma dwie części. Pierwsza to uzasadnienie istnienia rozkładu ($\diamondsuit$), a druga to dowód jego jednoznaczności.

Zacznijmy od wyrazenia istnienia rozkładu ($\diamondsuit$). Rozumowanie to indukcja ze względu na stopień p. Jeśli p jest nierozkładalny, a w szczególności, jeśli p jest stopnia 1, to $p = a \cdot q$, gdzie a jest wiodącym współczynnikiem p i jest jasne, że q jest unormowany i nierozkładalny.

Możemy zatem przejść do kroku indukcyjnego i jednocześnie założyć, że p jest rozkładalny. W takim przypadku $p = p_1 p_2$, dla pewnych $p_1, p_2 \in K[x]$, przy czym $\deg(p) > \deg(p_i) \geqslant 1$ praz z założenia indukcyjnego:

$$p_1 = a_1 \cdot q_1 \ldots q_l, \quad p_2 = a_2 \cdot q_{l+1} \ldots q_k,$$

gdzie q_i są unormowane i nierozkładalne oraz a_i są współczynnikami wiodącymi w p_i. W szczególności

$$p = a_1 a_2 \cdot q_1 \ldots q_l \cdot q_{l+1} \ldots q_k,$$

jest rozkładem typu ($\diamond$).

Aby udowodnić jednoznaczność, wykażemy najpierw pewną obserwację. Zauważmy mianowicie, że jeśli $f \in K[x]$ jest nierozkładalny oraz $f \mid gh$, gdzie $g, h \in K[x]$, to $f \mid g$, lub $f \mid h$,. Innymi słowy, element f jest **pierwszy** w $K[x]$. Dowód wymaga Lematu Bezout. Gdyby f nie dzielił f, to wobec nierozkładalności f mielibyśmy $NWD(f, g) = 1$. W szczególności, Lemat Bezout gwarantowałby istnienie $a, b \in K[x]$ takich, że

$$af + bg = 1 \quad \Rightarrow h = afh + bgh.$$

Stąd oczywiście $f \mid h$.

Aby pokazać jednoznaczność rozkładu ($\diamondsuit$) niech:

$$p = a \cdot q_1 \ldots q_k = a \cdot q_1' \ldots q_r'$$

przedstawia dwa rozkłady p na czynniki nierozkładalne. Na mocy obserwacji wyżej, skoro q_1 dzieli $q_1' \ldots q_r'$, to q_1 dzieli jeden z czynników q_i'. Skoro jednak obydwa te wielomiany są unormowane i nierozkładalne, to $q_1 = q_i'$. Skracając te dwa czynniki i powtarzając en proces aż wszystkie $q_1, \ldots, q_k$ zostaną skrócone prowadzi nas do tezy. □

Uwaga. Uzyskany rezultat mówi, że $K[x]$ jest tzw. dziedziną z jednoznacznością rozkładu (UFD). Na Algebrze I pokażemy także, że $\mathbb{Z}[x]$ też jest UFD, choć droga dowodu jest tam nieco bardziej skomplikowana. Czytelnik może zastanowić się czym dowód powyższy różni się od dowodu twierdzenia o istnieniu i jednoznaczności rozkładu na czynniki pierwsze w zbiorze liczb całkowitych.

Związek jest, jak się okazuje, dość głęboki, bowiem gdy uda się precyzyjnie określić czym są w pierścieniu elementy nierozkładalne i elementy pierwsze, jak je utożsamiać, oraz jak sformułować pojęcie dziedziny z jednoznacznością rozkładu, wówczas okazuje się, że gdy R jest taką dziedziną, to również wielomiany o współczynnikach w R mają jednoznaczny rozkład. Te zagadnienia wymagają jednak głębszego wejścia w teorię pierścieni przemiennych bez dzielników zera oraz ich tak zwanych ciał ułamków.

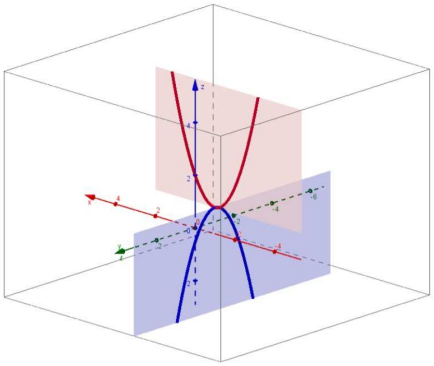
3.5 Trivia. Wykres funkcji zespolonej?

W przeciwieństwie do funkcji zmiennej rzeczywistej, której przebieg zmienności prezentujemy często graficznie na płaszczyźnie kartezjańskiej, prezentowanie „wykresów" funkcji zmiennej zespolonej nie jest czytelne z uwagi na to, że wymagałoby operowania w przestrzeni czterowymiarowej (zbiór argumentów ma dwie współrzędne i zbiór wartości ma dwie współrzędne). Z uwagi jednak na to, że często interesuje nas rozwiązanie równania $f(z) = 0$, wprowadza się różne ciekawe metody wizualizacji tego problemu. Jedną z nich są tzw. krzywe bliźniacze, wprowadzone w jednym z podręczników licealnych (!) w USA w latach 50' przez Howarda Fehra (to były początki „New Math" w nauczaniu – kto by chciał przeczytać więcej polecam artykuł: *New Thinking in School Mathematics* słynnego matematyka J. Dieudonné'a).

Rozważmy funkcję $f(z) = z^2 + 2z + 2$. Nietrudno sprawdzić, że rozwiązaniami równania $f(z) = 0$ są liczby zespolone $-1 \pm i$. Jak to zobaczyć na „wykresie"? Niech $z = x + iy$. Wówczas:

$$f(z) = f(x + iy) = (x^2 - y^2 + 2x + 2) + (2y(x + 1))i.$$

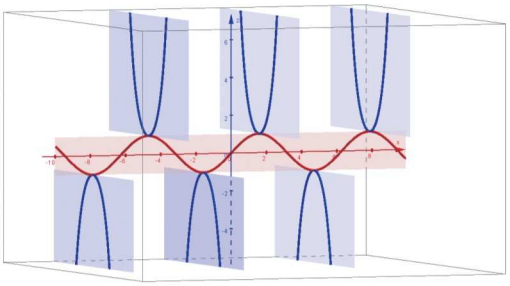
Skoro interesuje nas jedynie prezentacja warunku dotyczącego rzeczywistej wartości funkcji f (chodzi nam o wartość zero), to pomyślmy dla jakich x, y powyższa funkcja przyjmuje jedynie wartości rzeczywiste? Oczywiście dla $y = 0$ lub $x = -1$. Na płaszczyźnie $y = 0$ wartości $f(z)$ dane są przez $f(x) = x^2 + 2x + 2$, $x \in \mathbb{R}$, co reprezentowane jest przez dobrze znaną parabolę. W płaszczyźnie $x = -1$, prostopadłej do płaszczyzny $y = 0$, funkcja nasza ma postać $f(-1 + yi) = -y^2 + 1, y \in \mathbb{R}$. Oto stosowny obrazek:

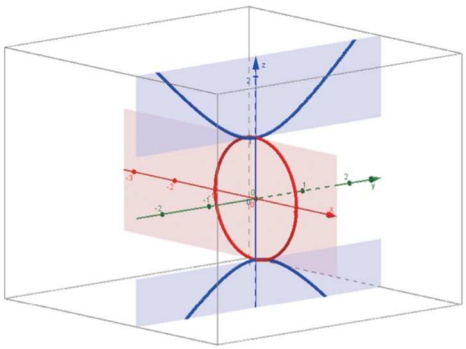


Rys. 1. Krzywe bliźniacze funkcji $f(z) = z^2 + 2z + 1$ w przestrzeni $\mathbb{R}^3$ to $x^2 + 2x + 2$ w płaszczyźnie $y = 0$ oraz $-y^2 + 1$ w płaszczyźnie $x = -1$. Źródło: Wiggins H., Harding A., Engelbrecht J.: *Visualising Complex Polynomials: A Parabola Is but a Drop in the Ocean of Quadratics*. J. Math. Research 10 (2018)

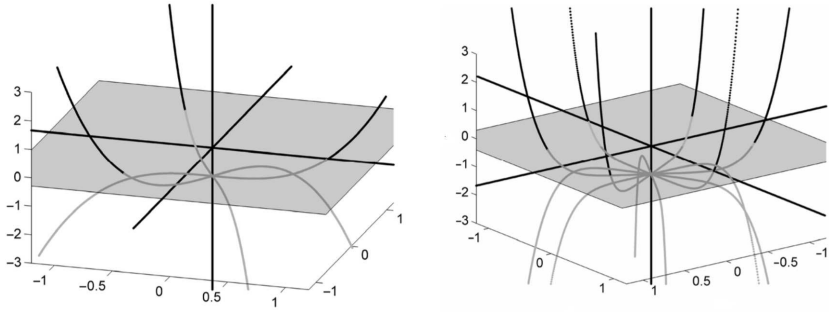
Co ten obrazek nam w zasadzie mówi? Otóż pokazuje nam on fragment czterowymiarowego wykresu funkcji $f(z)$ – ten mianowicie, na którym wartości funkcji są jedynie liczbami rzeczywistymi. Te wartości rzeczywiste reprezentowane są na osi OZ. Inaczej mówiąc: każda z powyższych dwóch krzywych ma punkty o trzech współrzędnych: (x, y, z). Pierwsze dwie współrzędne „kodują" punkt $x+iy$ z dziedziny funkcji f, zaś współrzędna z zawiera wartość rzeczywistą funkcji $f(z)$. A zatem zgodnie z intuicją: czerwona parabola nie ma punktu o współrzędnej $z = 0$, natomiast niebieska parabola – owszem: przecina płaszczyznę $z = 0$ w punktach $(-1, -1)$ oraz $(-1, 1)$. Te punkty reprezentują oczywiście liczby zespolone $-1 \pm i$.

Podobnego typu obrazki generować można dla innych funkcji, korzystając niekiedy z postaci trygonometrycznej lub wykładniczej liczb zespolonych. Ciekaw jestem czy Czytelnik potrafiłby powiedzieć jakimi równaniami opisane są krzywe bliźniacze dla funkcji $f(z) = \sin(z)$ oraz dla krzywej postaci $y^2 + z^2 = 1$?





Pouczająco wyglądają także obrazki prezentujące rozwiązania równań $z^3 = 1$ lub $z^6 = 1$.

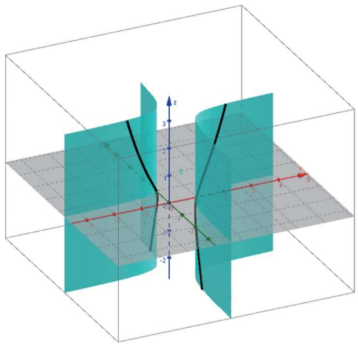


Rys. 3. Krzywe bliźniacze funkcji $f(z) = z^3 - 1$ oraz $f(z) = z^6 - 1$. Źródło: Harding A., Engelbrecht J.: *Sibling curves and complex roots 2: Looking ahead.* International Journal of Mathematical Education in Science and Technology, 38 (2017), 975-985.

Dla „funkcji kwadratowych" (zmiennej zespolonej) postaci $f(z) = az^2 + bz + c$, gdzie $a, b, c \in \mathbb{C}$, $a \neq 0$, pokazuje się, że zachodzić musi jedna z wykluczających się dwóch sytuacji:

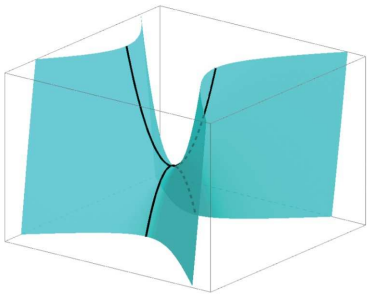
- dwie krzywe bliźniacze przecinają się – i wtedy są dwiema parabolami,
- dwie krzywe bliźniacze nie przecinają się – i stanowią gałęzie hiperboli.

Druga sytuacja zachodzi np. dla $f(z) = (z-1)(z-i)$, gdzie krzywe bliźniacze dane są wzorem $y = \frac{x-1}{2x-1}$.



Rys. 4. Krzywe bliźniacze funkcji $f(z) = (z-1)(z-i)$. Źródło jw.

Nie ma w tym rozróżnieniu, jak się okazuje, nic dziwnego. Jak pokazać, że taka klasyfikacja ma miejsce? Zasadniczo chodzi o sprowadzenie funkcji do postaci kanonicznej i rozważanie jedynie jej – jako geometrycznie „istotnej" dla kształtu krzywych bliźniaczych. Przez użycie przesunięcia, skalowania i obrotu można, bez straty ogólności, rozważać jedynie krzywe bliźniacze dla równania $f(z) = z^2 + c$, dla pewnej liczby zespolonej c. Przyjmując $z = x + iy$ dostajemy, że krzywe bliźniacze zawsze leżą, z dokładnością do skalowania, przesunięcia czy obrotu, na tak zwanej **hiperboloidzie parabolicznej** $z = x^2 - y^2$.



Rys. 5. Krzywe bliźniacze funkcji $f(z) = z^2 - 1$ na paraboloidzie hiperbolicznej $x^2 - y^2 = z$.

Należałoby, rzecz jasna, uściślić co znaczy stwierdzenie, że przesunięcie, skalowanie i obrót nie zmieniają „istoty geometrycznej" rozważanego problemu? Dlaczego wystarczyło rozważać jedynie krzywe bliźniacze równania $f(z) = z^2 + c$? Znacznie ogólniejsze i bardziej szczegółowe wyjaśnienie otrzymacie Państwo w drugim semestrze, gdy rozważać będziemy tak zwaną afiniczną klasyfikację hiperpowierzchni stopnia 2.

Animację ukazującą zmianę położenia krzywych bliźniaczych na paraboloidzie hiperbolicznej dla rodziny funkcji o równaniach $f(z)z^2 + 2z + (1 + ki)$ w zakresie $-2 \leqslant k \leqslant 2$ znajdziecie Państwo pod adresem: `https://cardanogroup.files.wordpress.com/2014/08/sibling-animation.gif`.

3.6 Coda. Wokół rozkładu na czynniki wielomianów i ich funkcji

Po dość rozbudowanych dodatkach traktujących o rozkładalności, warto powiedzieć kilka słów o ogólnej koncepcji stojącej za rozkładem, umieszczając ją w odpowiednim kontekście historycznym. W powyższych rozważaniach wspomnieliśmy bowiem tak o istnieniu, jak i o jednoznaczności rozkładu liczb całkowitych (różnych od $-1, 0, 1$) na czynniki pierwsze, o analogicznej własności wielomianów o współczynnikach w ciele oraz o przedziwnych własnościach rozkładów w pierścieniach typu $\mathbb{Z}[\epsilon_n]$, gdzie ϵ_n jest pierwiastkiem stopnia n z 1 (dla $n = 2$ otrzymujemy pierścień liczb całkowitych Gaussa, dla $n = 3$ — liczby Eisensteina, a dla większych n — liczby rozważane w kontekście Wielkiego Twierdzenia Fermata). Rozważania te mają już bardzo współczesny charakter. Poprzedziło je jednak kilka stuleci rozważań nad problemami bardziej chyba dla nas namacalnymi. Wyznaczyły one nurt ważnej koncepcji matematycznej.

Wspomnieliśmy w poprzednim dodatku o przełomie, jaki dokonał Clavius, a po nim Kartezjusz, kojarząc związek rozkładalności na czynniki z istnieniem rozwiązań równań wielomianowych. Zamiast stosować metody geometryczne do rozwiązywania równania typu $x^2 = 9x + 70$, można przenieść wszystkie wyrazy na jedną stronę i przedstawić równanie $x^2 - 9x - 70 = 0$ w postaci równoważnej: $(x - 14)(x + 5) = 0$, uzyskując dwa rozwiązania. Kluczowa koncepcja polega na (niełatwej często) redukcji złożonego problemu do układu problemów łatwych: równanie stopnia drugiego sprowadzamy do dwóch łatwych równań stopnia pierwszego. Tego typu sposób działania ma fundamentalne znaczenie. Nie bez przesady będzie stwierdzenie, że na kursie algebry liniowej wielomiany potrzebować będziemy właśnie po to (w drugim semestrze), aby skomplikowaną naturę przekształcenia geometrycznego „rozkładać" na składowe mające bardzo czytelną interpretację geometryczną. Zwrócimy więc jedynie uwagę na kilka wątków, w których rozkład na czynniki liniowe odkrywał ważną rolę. Zaczniemy jednak od kwestii zasadniczej, czyli twierdzenia Bezout, należącego w istocie do Kartezjusza.

Aż do XVII wieku teoria równań wielomianowych była, jak wspominaliśmy w rozdziale o wzorach skróconego mnożenia, zagadnieniem rozważanym w języku geometrycznym. *Geometria* Kartezjusza z roku 1637 zawierała ona dwa istotne wzbogacenia dotychczasowej teorii i notacji, wprowadzonej częściowo już w XVI wieku przez Viete'a, a mianowicie czytelną notację wykładniczą: x^3, x^4, x^5 itd. (choć nie x^2, które pozostało jako xx aż do XVIII wieku) i właśnie owo twierdzenie Bezout. Zobaczmy nieco szerszy kontekst.

Gdy mówimy o rozkładzie $w(x) = (x - c) \cdot v(x)$, to mamy na myśli, że po wymnożeniu $(x - c) \cdot v(x)$ otrzymamy wielomian, który ma takie same współczynniki, jak $w(x)$. Dlaczego to jest ważne? Chcemy bowiem korzystać z następującej implikacji: jeśli $w(x), v(x) \in K[x]$ oraz mamy **rozkład wielomianu** na czynniki: $h(x) = w(x) \cdot v(x)$, to dla każdego $s \in K$ mamy **rozkład wartości funkcji wielomianowej**

$$h(s) = w(s) \cdot v(s).$$

Cóż to za szaleństwo, czy to nie jest oczywiste? Dla wielomianów o współczynnikach w ciele, istotnie jest to prawda. Przyjęcie współczynników innego typu może to zaburzyć. Rozważmy ważny przykład.

W 1843 roku Rowan Hamilton dokonał odkrycia kwaternionów, czyli liczb zapisywanych w postaci

$$a + xi + yj + zk,$$

gdzie $a, x, y, z \in \mathbb{R}$ oraz gdzie (przy mnożeniu tych liczb) spełnione są następujące reguły:

$$ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j, \quad i^2 = j^2 = k^2 = -1.$$

Co ciekawe, dodawanie (po współrzędnych) i mnożenie kwaternionów są łączne (Hamilton pokazał to w roku 1844, po raz pierwszy używając tego terminu) oraz rozdzielne, dzielenie przez niezerowy element jest zawsze możliwe, dodawanie jest przemienne. Mnożenie kwaternionów jest jednak nieprzemienne. Stąd z punktu widzenia struktury, kwaterniony tworzą swego rodzaju ciało nieprzemienne (tak były zresztą nazywane, dziś mówimy raczej o pierścieniach z dzieleniem). Zbiór kwaternionów oznaczamy przez $\mathbb{H}$. Rozważmy iloczyn wielomianów o współczynnikach w $\mathbb{H}$, postaci

$$h(x) = (x - i)(x - j).$$

Jeśli wymnożymy go zgodnie z tradycyjnymi zasadami mnożenia wielomianów, wypisanymi w definicji z wykładu, uzyskamy

$$h(x) = x^2 - ix - xj + ij \stackrel{?}{=} x^2 - (i + j)x + ij.$$

W ten sposób poznajemy współczynniki wielomianu $h(x) \in \mathbb{H}[x]$. Podstawiając jednak do odpowiadającej mu funkcji wielomianowej kwaternion $s = i$, uzyskamy szokujący wynik

$$h(i) = i^2 - (i + j)i + ij = 2k \neq 0.$$

Dlaczego tak się stało? Przecież „twierdzenie Bezout" wyraźnie mówi, że wielomian $h(x)$ ma mieć pierwiastek i, skoro ma dzielnik postaci $x - i$.

Problem leży w samej definicji mnożenia wielomianów, który ma w istocie postać $x^2 - ix - xj + ij$ i nie można, jak się okazuje zapisać $xj = jx$, jeśli chce się zachować dobre własności funkcji wielomianowych wypisane wyżej. Innymi słowy skalary nie są konieczne przemienne ze zmiennymi. Trzeba się umówić czym jest xj, i jak się okazuje, sensowne jest przyjąć raczej $xj = -jx$. Co to za wielomiany?! O tym więcej tu nie powiemy. Powiemy natomiast dokładniej o jakie dobre własności chodzi.

Rozważmy przyporządkowanie działające w następujący sposób: dla każdego $a \in K$ rozważamy funkcję $v_a : K[x] \to K$, która przyporządkowuje wielomianowi postaci $w(x) = r_0 + r_1x + r_2x^2 + \ldots + r_nx^n \in K[x]$ wartość odpowiadającej mu funkcji wielomianowej w punkcie a, a więc element

$$v_a(w) = r_0 + r_1a + r_2a^2 + \ldots + r_na^n \in K.$$

Funkcja ta nazywa się EWALUACJĄ WIELOMIANU w punkcie a. Przyzwyczailiśmy się do pewnych własności ewaluacji, na przykład do następujących. Dla każdego $a \in K$ mamy (tzw. homomorfizm pierścieni):

$$v_a(w + w') = v_a(w) + v_a(w'), \quad v_a(w \cdot w') = v_a(w) \cdot v_a(w').$$

Jak się jednak okazuje, tak być nie musi, jeśli zbiór współczynników wielomianu nie jest przemienny, czego przykład mamy wyżej! Czytelnik może odczuwać pewną konsternację, dochodząc do tej konkluzji. Wydaje mi się jednak ważne, by pokazać, że stwierdzenie „funkcja wielomianowa iloczynu to iloczyn funkcji wielomianowych" ma głębokie podłoże.

Proszę zauważyć, że ewaluacja może zdecydowanie ułatwić wykonywanie rachunków. Oto przykład.

Zadanie. Rozważmy zbiór $\mathcal{A} = \{\varepsilon_0, \ldots, \varepsilon_{11}\}$ pierwiastków zespolonych stopnia 12 z 1. Uzasadnij, że

$$(\sqrt{3} + i - \varepsilon_0) \cdot (\sqrt{3} + i - \varepsilon_1) \cdot \ldots \cdot (\sqrt{3} + i - \varepsilon_{11}) = 2^{12} - 1.$$

Proszę zauważyć, że pierwiastki stopnia 12 z 1 można wyznaczyć, wyliczając ich postaci ogólne. W ten sposób możliwe jest policzenie powyższego iloczynu przez wymnożenie 12 nawiasów, odpowiednio grupując czynniki zawierające pierwiastki sprzężone. Znacznie łatwiej jest jednak zauważyć, że wobec rozkładu

$$x^{12} - 1 = (x - \varepsilon_0)(x - \varepsilon_1) \cdot \ldots \cdot (x - \varepsilon_{11})$$

możemy wyznaczyć żądany iloczyn, poprzez wyznaczenie wartości funkcji wielomianowej odpowiadającej wielomianowi $x^{12} - 1$ w punkcie $s = \sqrt{3} + i$. Innymi słowy, mamy:

$$(\sqrt{3} + i)^{12} - 1 = (\sqrt{3} + i - \varepsilon_0) \cdot (\sqrt{3} + i - \varepsilon_1) \cdot \ldots \cdot (\sqrt{3} + i - \varepsilon_{11}) = 2^{12} - 1.$$

Nie tylko przemienność mnożenia w zbiorze współczynników, ale i inna kwestia wchodzi w grę. Chcemy bowiem z tego, że $0 = h(s) = w(s) \cdot v(s)$ wnioskować, że $w(s) = 0$ lub $v(s) = 0$. Bez tej własności rozwiązywanie równań będzie często trudne. Wystarczy odejść niedaleko od definicji ciała, by znaleźć zbiory współczynników nie mające takiej własności. Rozważmy zbiór reszt z dzielenia przez 6, czyli $\mathbb{Z}_6$. Jego definicja jest analogiczna, jak dla ciał $\mathbb{Z}_p$, ale nie jest to jednak ciało, ponieważ nie każdy element $\mathbb{Z}_6$ ma odwrotność — choćby element 2. Mamy wręcz $2 \cdot 3 = 0$, co dla ciał nie może mieć miejsca. Wystarczy zobaczyć, że zwykły wielomian $x^2 + 5x \in \mathbb{Z}_6[x]$ o współczynnikach w pierścieniu reszt z dzielenia przez 6 ma więcej niż dwa pierwiastki, bowiem

$$x(x + 5) = (x + 2)(x + 3) = x^2 + 5x,$$

co dla ciał nie jest możliwe. Gdy zaburzymy przemienność, sytuacja ulega znacznemu pogorszeniu. Dla wielomianu o współczynnikach kwaternionowych $w(x) = x^2 + 1 \in \mathbb{H}[x]$ mamy $w(i) = w(j) = w(k) = 0$, a w istocie wielomian ten ma nieskończenie wiele pierwiastków! Stąd zbiór współczynników ma ogromne znaczenie dla rozkładów na czynniki i dla stowarzyszonych z wielomianami funkcji wielomianowych.

Przejdźmy do ważnych zastosowań historycznych rozmaitych aspektów rozkładu wielomianów na czynniki.

Pierwszy wydawać się może zaskakujący, dotyczy bowiem teorii liczb i znany jest pod nazwą małego twierdzenia Fermata. Twierdzenie to pochodzi z roku 1640 i mówi, że jeśli p jest liczbą pierwszą oraz n jest dodatnią liczbą całkowitą względnie pierwszą z p, to liczba $n^{p-1}-1$ jest podzielna przez p, lub równoważnie — liczba $n^p - n$ jest podzielna przez p.

Twierdzenie to stało się współcześnie jednym z podstawowych narzędzi kryptograficznych, stąd wydaje się ciekawe wspomnienie, że Fermata w istocie interesowało to kiedy liczba $2^m - 1$ ma dzielniki pierwsze. W istocie, największe znajdowane dziś liczby pierwsze są tej postaci (tzw. liczby pierwsze Mersenne'a). Z punktu widzenia teorii rozkładu wielomianów, małe twierdzenie Fermata mówi, że w ciele $\mathbb{Z}_p$ wielomian $x^p - x$ rozkłada się na iloczyn czynników liniowych:

$$x^p - x = \prod_{a \in \mathbb{Z}_p} (x - a),$$

czyli mówiąc inaczej — każdy element tego ciała jest jego pierwiastkiem (a jednak nie jest to wielomian zerowy). Faktoryzacja wielomianów nad ciałami skończonymi ma fundamentalne znaczenie w kryptografii. Do jej zastosowań należy wyznaczanie tzw. dyskretnego logarytmu, niezbędnego przy konstrukcji szyfrowania klucza publicznego. Idee znajdowania tego typu rozkładów pochodziły od Legendre'a i Gaussa, a w XX wieku rozwinięte zostały przez Berlekampa, czy współcześnie przez Cantora i Zassenhausa.

Drugie spojrzenie również pochodzi z pierwszej połowy XVII wieku, gdy Fermat i Kartezjusz niezależnie budowali podstawy geometrii analitycznej. Motywacją było badanie krzywych opisanych przez równania. Krzywym przypisywano stopień: proste to krzywe stopnia 1, krzywe stożkowe (okrąg, parabola, hiperbola) to krzywe stopnia 2, tzw. koniki to krzywe stopnia 3, w tym na przykład krzywa $y^2 = x^3$, niebędąca wykresem funkcji. Podstawowym celem było osiągnięcie następującego rezultatu, zwanego twierdzeniem Bezout: krzywa stopnia m przecina zawsze krzywą stopnia n w nie więcej niż mn punktach. Dla przykładu: prosta przecina okrąg w nie więcej niż dwóch punktach, ale już dwie elipsy mogą się przeciąć w czterech punktach, podobnie okrąg i parabola. Dlaczego twierdzenie to było tak istotne?

Paradygmat dla prac Kartezjusza stanowiła starożytna konstrukcja Menaechmusa liczby $\sqrt[3]{2}$ za pomocą przecięcia paraboli i hiperboli. Uczony ten, współczesny Aleksandrowi Wielkiemu, był faktycznym twórcą teorii stożkowych, czyli krzywych uzyskiwanych przez cięcie stożka za pomocą płaszczyzny (będziemy o nich mówić w drugim semestrze). Przypisujemy mu zdanie „Nie ma królewskiej drogi do geometrii". On też zaproponował proste rozwiązanie problemu podwojenia sześcianu (który chciano uzyskać za pomocą cyrkla i linijki, co nie jest możliwe, i co wiemy dopiero od wieku XIX-tego). W języku algebraicznym chodzi o przecięcie ze sobą paraboli $y = \frac{1}{2}x^2$ oraz hiperboli o równaniu $xy = 1$. W ten sposób mamy:

$$x \cdot \frac{1}{2}x^2 = 1 \iff x^3 = 2.$$

W okolicach roku 1620 Kartezjusz zorientował się, że dowolne rozwiązanie równania wielomianowego stopnia 3 lub 4 może skonstruować poprzez przecinanie krzywych stopnia 2. Co więcej, w swoim fundamentalnym dziele *Geometria* (1637), wskazał on krzywą stopnia 3, zwaną parabolą kartezjańską, której przecięcie z odpowiednim okręgiem dawało rozwiązania dowolnego równania wielomianowego stopnia 5 lub 6. Oczywiście dało mu to niezachwianą wiarę w możliwość uogólnienia tych rezultatów. Jak się okazało, stosowne uogólnienie nie było banalne i odpowiednie konstrukcje rozwiązań wielomianów stopnia n ustalono dopiero około roku 1750. Co to wszystko ma wspólnego z rozkładem?

Problem wyznaczania liczby przecięć krzywych sprowadzić można do problemu rozwiązywania równań wielomianowych stopnia n. W prostych przypadkach jest to oczywiste. Badając liczbę przecięć elipsy $x^2 + 2y^2 = 1$ oraz paraboli $y = x^2$ należy jedynie wstawić w pierwszym równaniu x^2 w miejsce y i rozwiązać równanie stopnia 4. Otrzymamy zatem nie więcej niż 4 punkty przecięcia.

Czasem proste rozdzielenie zmiennych nie jest zupełnie jasne, zwłaszcza gdy w grę wchodzą tzw. punkty wielokrotne czy punkty w nieskończoności. Niemniej jednak dowód zasadniczego twierdzenia algebry oraz rozwój metod geometrii rzutowej (i związanych z nimi tzw. wielomianów jednorodnych) pozwolił ustalić, że takie rozdzielenie zmiennych jest w istocie zawsze możliwe. Rezultat ten uzyskano ostatecznie dopiero pod koniec XIX wieku, wykorzystując teorię wyznaczników.

Co natomiast można powiedzieć o historii samego dowodu zasadniczego twierdzenia algebry? Rezultat z wykładu mówiący, że wielomian o współczynnikach rzeczywistych mający pierwiastek zespolony $z = a + bi$ ma również pierwiastek sprzężony $\overline{z} = a - bi$ pochodzi od d'Alemberta (1746). Wraz z nim uzyskano oczywiście postawiony przez nas wniosek: zasadnicze twierdzenie algebry jest dla wielomianów rzeczywistych równoważne możliwości rozłożenia każdego takiego wielomianu (dodatniego stopnia) na iloczyn czynników rzeczywistych stopnia 1 lub 2. W ten sposób twierdzenie to formułowano przez większość XVIII stulecia, co pozwoliło unikać wspomnień o (wciąż podejrzanym) $\sqrt{-1}$ oraz zezwalało na użytek metod analizy funkcji rzeczywistych.

Dowody zasadniczego twierdzenia algebry, zarówno te proponowane przez d'Alemberta, jak i pierwotny dowód Gaussa, miały luki, które naprawione zostały dopiero pod koniec XIX wieku i miały charakter analityczny. Jakie było podejście algebraiczne? Główne nurty pochodziły od Eulera, w którego czasach obok ZTA najsłynniejszym problemem było zagadnienie znalezienia wzorów na pierwiastki wielomianów wyższych stopni (ostatecznie rozstrzygnięte negatywnie, dla stopnia $\geqslant 5$ przez Abela, Ruffiniego i Galois), Euler zajął się problemem znanym mu skądinąd z badań nad wielomianem $x^n - 1$. W tym celu wprowadził jednostkę urojoną i zaczął badać wyrażenia postaci $(\cos\theta + i\sin\theta)$ oraz ich potęgi. To właśnie Euler w istocie jako pierwszy sformułował w pełnej wersji znany nam już wzór Moivre'a. W 1749 roku wyznaczył wzory na pierwiastki stopnia n-tego z liczby zespolonej i postulował, że zbiór liczb zespolony jest zamknięty na branie pierwiastków — co jest własnością nieznaną ani liczbom naturalnym, ani całkowitym, wymiernym czy rzeczywistym. Pomijamy w tym miejscu ogromne zastosowania, jakie wniosły prace Eulera do analizy, w tym do przedstawień funkcji w postaci szeregów.

Z punktu widzenia rozkładu na czynniki liniowe, Euler ustalił tożsamość

$$x^n - 1 = (x - \omega_0)(x - \omega_1)(x - \omega_2)\cdots(x - \omega_{n-1}),$$

gdzie $\omega_0, \ldots, \omega_{n-1}$ są pierwiastkami stopnia n z 1.

Eulerowi udało się udowodnić, że każdy wielomian rzeczywistych stopnia $n \leqslant 6$ ma dokładnie n pierwiastków zespolonych. W tym samym roku 1749 podjął próbę ogólnego dowodu, opartą o rozkład wielomianu unormowanego stopnia 2^n na iloczyn wielomianów stopni 2^{n-1}. Ideą była sztuczka znana już z prac Cardano mówiąca, że przez odpowiednie podstawienie można pozbyć się z wielomianu stopnia n wyrazu przy potędze $n - 1$. Planował też udowodnić istnienie rozkładu:

$$x^{2m} + Ax^{2m-2} + Bx^{2m-3} + \ldots = (x^m + tx^{m-1} + gx^{m-2} + \ldots)(x^m - tx^{m-1} + hx^{m-2} + \ldots).$$

Twierdził, że współczynniki A, B będą funkcjami wymiernymi od $A, B, \ldots, t$, ale ogólny przypadek uzasadnił jedynie w formie szkicu, podważonego przez Lagrange'a (zachodziła obawa, że niektóre funkcje wymierne są postaci $0/0$). Dowody przedstawiane przez kolejnych autorów (Laplace, Gauss, Argand) zawierały wciąż luki. Drugie podejście Gaussa z roku 1816 uznane zostało jednak za całkowicie poprawne, i stąd przypisujemy mu pierwszeństwo. Rozumowanie jest w pełni algebraiczne, za wyjątkiem użycia szczególnego przypadku twierdzenia o wartości średniej, które dla wielomianów uzasadnił Bolzano (1817), a dla funkcji ciągłych — Weierstrass (1874). W 1849 roku, Gauss przedstawił dowód ZTA dla wielomianów zespolonych. Późniejsi autorzy, w tym Frobenius, uznawali zasługi Eulera, którego od pełnego dowodu dzieliła nieumiejętność wykazania, że wielomian rzeczywisty nieparzystego stopnia ma pierwiastek.

Na zakończenie warto dodać, że rozkłady na czynniki mają ważne odniesienia teorioliczbowe, zaobserwowane przez arabskich komentatorów dzieł Diofantosa. Chodzi o tożsamość

$$(a_1^2 + b_1^2)(a_2^2 + b_2^2) = (a_1a_2 \pm b_1b_2)^2 + (b_1a_2 \pm a_1b_2)^2.$$

Rezultat ten łatwo interpretować w języku kwadratów modułów liczb zespolonych, czyli

$$|a_1 + ib_1| \cdot |a_2 + ib_2| = |(a_1 + ib_1)(a_2 + ib_2)|.$$

Analogiczna równość nie jest możliwa dla trzech zmiennych. Innymi słowy, nie istnieje tożsamość postaci:

$$(a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2) = A^2 + B^2 + C^2,$$

gdzie A, B, C są kombinacjami a_i, b_i, c_i o współczynnikach całkowitych. Odkrycie przez Hamiltona kwaternionów sprawiło, że sformułował stosowny wzór dla czterech zmiennych, pochodzący zresztą od Eulera. W 1898 r. Hurwitz pokazał, że tego typu tożsamości uzyskać można jedynie dla $n = 1, 2, 4, 8$ zmiennych.

Rozdział 4

Przestrzenie liniowe

4.1 Wykład czwarty

Wykład ten poświęcony będzie pojęciu przestrzeni liniowej nad ciałem. Jest to fundamentalne pojęcie dla całego naszego wykładu i jedno z najważniejszych w całej matematyce. Mówić będziemy o strukturze określonej jednocześnie na dwóch typach obiektów: wektorach i skalarach. Struktura ta jest w swojej istocie „geometryczna", choć odnaleźć ją można w bardzo odległych z pozoru dziedzinach matematyki.

Definicja 4.1: Przestrzeń liniowa nad ciałem K

Przestrzenią liniową nad ciałem $(K, +, \cdot, 0, 1)$ nazywamy zbiór V, wraz z:

- odwzorowaniem: $\oplus : V \times V \longrightarrow V$, zwanym dodawaniem wektorów,
- odwzorowaniem: $\otimes : K \times V \longrightarrow V$, zwanym mnożeniem wektora przez skalar,
- wyróżnionym elementem Θ w V zwanym wektorem zerowym,

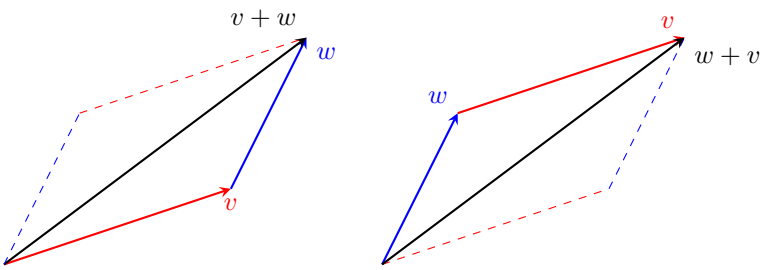
przy czym spełnione są następujące aksjomaty przestrzeni liniowej:

1)	$\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma$	$\forall_{\alpha,\beta,\gamma \in V}$	łączność dodawania wektorów
2)	$\alpha \oplus \beta = \beta \oplus \alpha$	$\forall_{\alpha,\beta \in V}$	przemienność dodawania wektorów
3)	$\alpha \oplus \Theta = \alpha$	$\forall_{\alpha \in V}$	Θ jest elem. neutralnym $\oplus$
4)	$\alpha \oplus \gamma = \Theta$	$\forall_{\alpha \in V} \exists_{\gamma \in V}$	istnienie wekt. przeciwnego
5)	$1 \otimes \alpha = \alpha$	$\forall_{\alpha \in V}$	mnożenie wektora przez 1
6)	$(a \cdot b) \otimes \alpha = a \otimes (b \otimes \alpha)$	$\forall_{\alpha,\beta,\gamma \in V}$	zgodność $\cdot$ z $\otimes$
7)	$(a + b) \otimes \alpha = (a \otimes \alpha) \oplus (b \otimes \alpha)$	$\forall_{\alpha \in V}, \forall_{a,b \in K}$	rozdzielność $\otimes$ względem $+$
8)	$a \otimes (\alpha \oplus \beta) = (a \otimes \alpha) \oplus (a \otimes \beta)$	$\forall_{\alpha,\beta \in V}, \forall_{a \in K}$	rozdzielność $\otimes$ względem $\oplus$

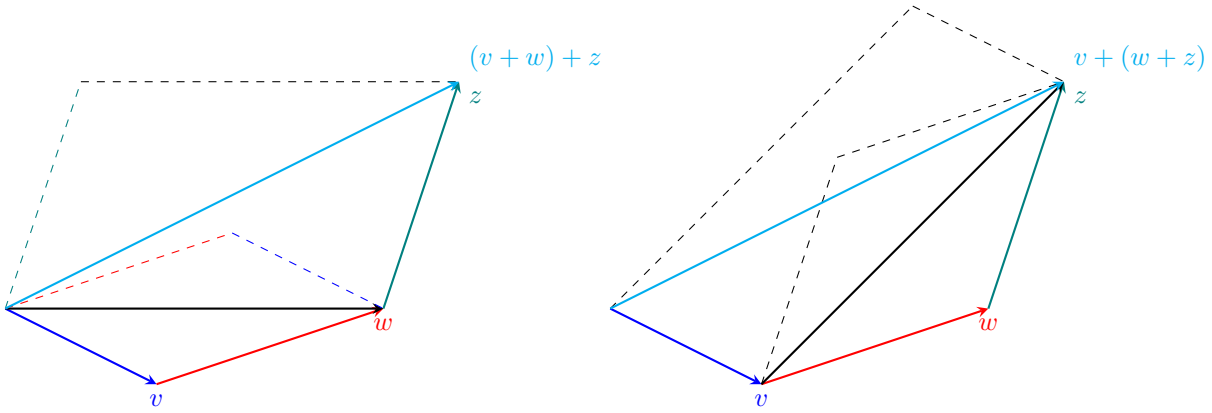
Elementy przestrzeni liniowej V nazywamy wektorami.

Jak widać, w definicji wystąpiło mnóstwo oznaczeń, zwłaszcza odnośnie działań. W dalszej części wszystkie symbole dodawania $\oplus, +$ będą zamienione na $+$ oraz wszystkie symbole mnożenia $\cdot, \otimes$ będą pomijane.

Zilustrujmy aksjomaty poprzez szkolne intuicje, w myśl których wektor wyznacza kierunek przesunięcia (np. działanie siły), a dodawanie wektorów umożliwia składanie przesunięć (np. wypadkowa układu sił). Dla wektorów v, w, wektor $v + w$ wyznacza przekątna równoległoboku wyznaczonego przez v oraz w.



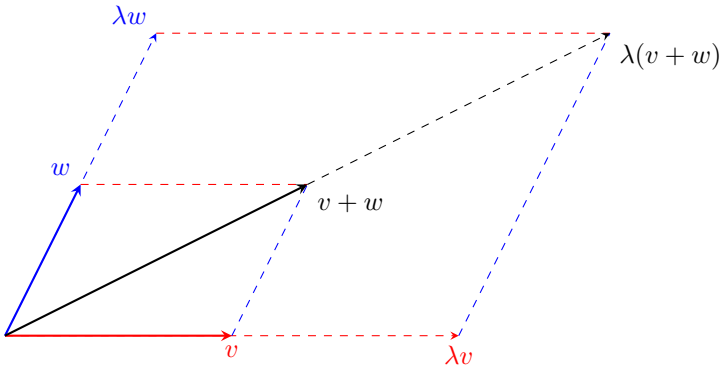
Powyższa ilustracja tłumaczy aksjomat (2) przemienności dodawania wektorów. Oto interpretacja szkolna aksjomatu (1) o łączności dodawania.



Należy pamiętać, że powyższe ilustracje mają charakter komentarza ilustrującego sensowność akscjomatów. W formułowanej przez nas definicji przestrzeni liniowej nie występuje choćby pojęcie punktu, a więc i początku lub końca wektora, które wprowadzimy na gruncie przestrzeni afinicznych w drugim semestrze.

Wektor przeciwny do v, czyli $-v$, reprezentuje kierunek przesunięcia przeciwny do v tak, by złożenie przesunięć o wektor v oraz $-v$ było identycznością, czyli przesunięciem o wektor 0.

Mnożenie przez skalar interpretujemy jako skalowanie (jednokładność). Szczególne znaczenie ma aksjomat (8) rozdzielności mnożenia przez skalar względem dodawania wektorów. Czy Czytelnik widzi w nim abstrakcyjną formę (bez definiowania czym jest długość czy proporcja wektorów) twierdzenia Talesa?



Powyższe aksjomaty, choć nie będą dotyczyły jedynie (a na starcie: w zasadzie w ogóle) wektorów w takim sensie, w jakim poznaliśmy je w szkole, to będą zachowywać własności, jakie mają wektory. Może to się wydawać niejasne, więc użyjmy następującej analogii: załóżmy, że zapomnieliśmy czym są „szkolne wektory", i jakie mają geometryczne własności, ale zapisaliśmy sobie na wszelki wypadek kilka kluczowych informacji, z których „chcemy odzyskać" tę wiedzę. Tak można interpretować aksjomaty (również aksjomaty ciała, poznane na wykładzie drugim oraz na Analizie) — zapominamy czym jest choćby prosta rzeczywista $\mathbb{R}$ (a raczej — zapominamy o modelu, którym się posługiwaliśmy mówiąc o niej) i zostawiamy jedynie kluczowe własności. Za ich pomocą budujemy pewną teorię, która nie tylko pozwoli nam odzyskać wiedzę ilustrowaną poprzez model, ale pozwoli wyprowadzić nowe własności natury geometrycznej.

Ale czym są w końcu wektory? To jest kluczowe — wektory stanowić mogą elementy dowolnego zbioru, który wraz z odpowiednimi działaniami spełnia listę aksjomatów wypisanych wyżej. Jak się okaże dalej, mogą to być nie tylko ciągi, ale i macierze, wielomiany, funkcje, podzbiory i wiele innych obiektów. Przy odpowiedniej interpretacji, wszystkie one mają te naturalne własności, które przypisywaliśmy wektorom „szkolnym" tak, że na zbiorach tych uprawiamy w istocie geometrię.

Zanim wyprowadzimy podstawowe własności przestrzeni liniowych musimy zapoznać się z dostateczną liczbą przykładów, przekonujących nas do zasadności wprowadzenia tak abstrakcyjnej definicji. Przykłady te pochodzą, jak się okazuje, z wielu różnych gałęzi matematyki.

Definicja 4.2: Przestrzeń współrzędnych K^n

Niech K^n oznacza zbiór wszystkich ciągów n-elementowych o wyrazach z ciała K, tzn.:

$$K^n = \{(x_1, x_2, \ldots, x_n) \,|\, x_i \in K, i = 1, 2, \ldots, n\}.$$

Wyraz x_i w ciągu $(x_1, x_2, \ldots, x_n)$ nazywamy i-TĄ WSPÓŁRZĘDNĄ tego wektora.

Wprowadzamy działania w K^n. Dla dowolnych $x_1, \ldots, x_n, y_1, \ldots, y_n, a \in K$ definiujemy:

- dodawanie wektorów: $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$,
- mnożenie wektora przez skalar: $a \cdot (x_1, x_2, \ldots, x_n) = (a \cdot x_1, a \cdot x_2, \ldots, a \cdot x_n)$.

Przykłady działań w przestrzeni K^n:

- dla $K = \mathbb{Z}_3$ i $n = 4$ mamy np. $(1, 2, 1, 2) + (0, 2, 2, 1) = (1, 1, 0, 0), \quad 2 \cdot (2, 2, 1, 1) = (1, 1, 0, 0)$,
- dla $K = \mathbb{C}$ i $n = 2$ mamy np. $(1, i) + i(i, 0) = (1, i) + (-1, 0) = (0, i)$.

Definicja 4.3: Przestrzeń liniowa macierzy rozmiaru $m \times n$ o wyrazach w ciele K

Niech $M_{m\times n}(K)$ oznacza zbiór wszystkich macierzy $m \times n$ o wyrazach z ciała K.

- Sumą macierzy $[a_{ij}]$ oraz $[b_{ij}]$ z $M_{m\times n}(K)$ nazywamy macierz $[c_{ij}] \in M_{m\times n}(K)$, której wyrazy spełniają warunek $c_{ij} = a_{ij} + b_{ij}$:

$$\begin{bmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{bmatrix} + \begin{bmatrix} & \vdots & \\ \cdots & b_{ij} & \cdots \\ & \vdots & \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \cdots & a_{ij} + b_{ij} & \cdots \\ & \vdots & \end{bmatrix}.$$

- Iloczynem macierzy $[d_{ij}] \in M_{m\times n}(K)$ przez skalar $c \in K$ nazywamy macierz $[c \cdot d_{ij}]$:

$$c \cdot \begin{bmatrix} & \vdots & \\ \cdots & d_{ij} & \cdots \\ & \vdots & \end{bmatrix} = \begin{bmatrix} & \vdots & \\ \cdots & c \cdot d_{ij} & \cdots \\ & \vdots & \end{bmatrix}.$$

Wektorem zerowym w przestrzeni liniowej $M_{m\times n}(K)$ jest MACIERZ ZEROWA rozmiarów $m \times n$.

Przykładowo, w przestrzeni liniowej $M_{2\times 3}(\mathbb{Z}_5)$ (ponownie $\oplus$ i $\otimes$ zastępujemy symbolami $+$ oraz $\cdot$):

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} + 2 \cdot \begin{bmatrix} 4 & 4 & 0 \\ 0 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 & 0 \\ 0 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 3 & 1 \end{bmatrix}.$$

Definicja 4.4: Przestrzeń liniowa wielomianów o współczynnikach w ciele K

Niech $K[x]$ będzie zbiorem wszystkich wielomianów zmiennej x o współczynnikach w ciele K, czyli

$$K[x] = \{a_0 + a_1 x + \ldots + a_n x^n \,|\, n = \mathbb{N} \cup \{0\}, a_0, a_1, \ldots, a_n \in K\}.$$

Dodawanie i mnożenie przez skalar pochodzą od omawianych tydzień wcześniej operacji na wielomianach. Wektorem zerowym w przestrzeni liniowej $K[x]$ jest wielomian zerowy.

Zauważmy, że w definicji przestrzeni liniowej wielomianów rozróżniamy działanie mnożenia skalara przez wielomian oraz działanie mnożenia wielomianów. Drugie z tych działań nie jest częścią definicji przestrzeni liniowej (choć będzie nam w różnych sytuacjach potrzebne). Struktura przestrzeni liniowej V z dodatkowym działaniem mnożenia wektorów (zgodnym z działaniami w V) nazywa się algebrą.

Definicja 4.5: Przestrzeń liniowa ciągów nieskończonych o wyrazach w ciele K

Oznaczmy przez K^∞ zbiór wszystkich ciągów o wyrazach z ciała K, to znaczy:

$$K^\infty = \{(x_i) \,|\, x_i \in K, i = 1, 2, \ldots\}.$$

Ciągi $x = (x_i)$ oraz $y = (y_i)$ dodajemy i mnożymy przez skalary według zasady:

$$(x \oplus y)_i = x_i + y_i, \quad (a \otimes x)_i = a \cdot x_i.$$

Wektorem zerowym w przestrzeni liniowej K^∞ jest ciąg, którego wszystkie wyrazy są zerem w K.

Przykładowo, z równości:

$$\frac{1}{n} + (-1)\frac{1}{n+1} = \frac{1}{n(n+1)},$$

zachodzącej dla każdej dodatniej liczby całkowitej n mamy równość w $\mathbb{Q}^\infty$ postaci

$$\left(\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots\right) - \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right) = \left(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \ldots\right).$$

Definicja 4.6: Przestrzeń liniowa funkcji ze zbioru X do ciała K

Niech $F(X, K)$ będzie zbiorem wszystkich funkcji z danego niepustego zbioru X do ciała K. Dla $f, g \in F(X, K)$ i dla $a \in K$ funkcje $f + g$ oraz af określone są warunkami:

$$(f + g)(x) = f(x) + g(x), \quad (af)(x) = af(x).$$

Wektor zerowy przestrzeni liniowej $F(X, K)$ to funkcja stale równa 0.

Podajmy jeszcze jeden przykład o dużym znaczeniu w kombinatoryce.

Definicja 4.7: Przestrzeń podzbiorów zbioru niepustego X nad ciałem $\mathbb{Z}_2$

Niech X będzie zbiorem niepustym, zaś $P(X)$ – niech będzie zbiorem podzbiorów zbioru X. Na zbiorze $P(X)$ określamy strukturę przestrzeni liniowej nad ciałem $\mathbb{Z}_2$.

Operacja $\triangle$ dodawania wektorów określona jest w sposób następujący dla dowolnych $A, B \in P(X)$ jako ich tzw. różnica symetryczna $A \triangle B = A \cup B \setminus (A \cap B)$.

Dla każdego $A \in P(X)$ definiujemy mnożenie wektora A przez skalar (jeden z dwóch w $\mathbb{Z}_2$):

- $0 \otimes A = \emptyset$ – zbiór pusty
- $1 \otimes A = A$.

Odnotujmy kilka istotnych własności wynikających z aksjomatów przestrzeni liniowej.

Obserwacja 4.1. *W każdej przestrzeni liniowej V nad ciałem K zachodzi:*

(a) dla każdego $\alpha \in V$ istnieje tylko jeden taki wektor $\delta \in V$, że $\alpha \oplus \delta = 0$. Wektor ten oznaczamy $-\alpha$ i nazywamy WEKTOREM PRZECIWNYM *do α.*

(b) $0 \otimes \alpha = 0$, dla każdego $\alpha \in V$ oraz $a \otimes 0 = 0$, dla każdego $a \in K$.

(c) Jeśli $\alpha \in V$ oraz $a \in K$, to $a \otimes \alpha = 0$, to $a = 0$ lub $\alpha = 0$.

(d) $-\alpha = (-1) \otimes \alpha$, dla każdego $\alpha \in V$.

Dowód. Zacznijmy od (a). Jeśli $\alpha \oplus \delta_1 = 0$ oraz $\alpha \oplus \delta_2 = 0$, to:

$$\delta_1 = \delta_1 \oplus 0 = \delta_1 \oplus (\alpha \oplus \delta_2) = (\delta_1 + \alpha) + \delta_2 = 0 + \delta_2 = \delta_2.$$

Dowodzimy (b). W ciele K mamy $0 + 0 = 0$. Stąd $0 \otimes \alpha = (0 + 0) \otimes \alpha = 0 \otimes \alpha \oplus 0 \otimes \alpha$. Dodając do obu stron tej równości wektor $-0 \otimes \alpha$ otrzymujemy $0 = a \otimes \alpha$. Dowód $a \otimes 0 = 0$ jest analogiczny. Uzasadnijmy teraz (c). Jeśli $a \neq 0$, to

$$\alpha = 1 \otimes \alpha = a^{-1}a \otimes \alpha = a^{-1} \otimes 0 = 0.$$

Dowód (d) pozostawiamy jako ćwiczenie. □

Podobnie jak w przypadku ciał, podstawowym narzędziem do uzyskiwania kolejnych przykładów przestrzeni liniowych jest pojęcie podprzestrzeni liniowej.

Definicja 4.8: Podprzestrzeń przestrzeni liniowej

Niepusty podzbiór $W \subset V$ nazywamy PODPRZESTRZENIĄ PRZESTRZENI LINIOWEJ V jeśli dla każdych $\alpha, \beta \in W$ oraz każdego $a \in K$ zachodzi:

(i) $\alpha + \beta \in W$,

(ii) $a \cdot \alpha \in W$.

W każdej przestrzeni liniowej V podzbiór $\{0\}$, złożony tylko z wektora zerowego, jest jej podprzestrzenią. Nazywamy ją PODPRZESTRZENIĄ ZEROWĄ. Mówimy, że przestrzeń liniowa V jest PRZESTRZENIĄ ZEROWĄ, jeśli składa się tylko z wektora zerowego.

UWAGA: Podprzestrzeń przestrzeni liniowej jest przestrzenią liniową (z działaniami pochodzącymi z V, w tym z odziedziczonym wektorem zerowym). **Wektor zerowy należy do każdej podprzestrzeni!**

Przejdźmy do kluczowego przykładu podprzestrzeni, który już poznaliśmy:

Obserwacja 4.1

Rozpatrzmy jednorodny układ równań liniowych o współczynnikach w ciele K.

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases} \quad (*)$$

Zbiór wszystkich rozwiązań układu U jest podprzestrzenią przestrzeni liniowej K^n.

Dowód. Niech $W \subseteq K^n$ będzie zbiorem rozwiązań układu $(*)$. Niech $(s_1, \ldots, s_n)$, $(r_1, \ldots, r_n) \in W$. Należy pokazać, że do W należą także wektory:

$$(s_1, \ldots, s_n) + (r_1, \ldots, r_n) = (s_1 + r_1, \ldots, s_n + r_n)$$

oraz, że dla każdego $a \in K$ do zbioru W należą również wektory:

$$a \cdot (r_1, r_2, \ldots, r_n) = (ar_1, ar_2, \ldots, ar_n).$$

Wystarczy sprawdzić, że wektory te spełniają każde równanie układu $(*)$. Rzeczywiście, dla każdego $1 \leqslant i \leqslant m$ mamy $a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n = 0$ oraz $a_{i1}r_1 + a_{12}r_2 + \ldots + a_{in}r_n = 0$, a zatem

$$a_{i1}(s_1 + r_1) + a_{i2}(s_2 + r_2) + \ldots + a_{in}(s_n + r_n) = 0.$$

Widzimy więc, że $(s_1, \ldots, s_n) + (r_1, \ldots, r_n) \in W$. Podobnie, dla każdego $a \in K$:

$$a \cdot (a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n) = a_{i1}as_1 + a_{i2}as_2 + \ldots + a_{in}as_n = 0.$$

A zatem $a \cdot (s_1, s_2, \ldots, s_n) \in W$, co oznacza, że W jest podprzestrzenią K^n. □

Przykład. Rozwiązaniami układu jednorodnego o współczynnikach w $\mathbb{R}$:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_1 - x_2 - x_3 - x_4 &= 0 \end{cases}$$

jest zbiór wektorów $(0, -s - t, s, t)$, gdzie $s, t \in \mathbb{R}$. Każdy element tego zbioru można zapisać w postaci:

$$s(0, -1, 1, 0) + t(0, -1, 0, 1)$$

gdzie $s, t \in \mathbb{R}$. Tu jest fundamentalna nowość, która jest sercem „geometrycznej strony" naszych rozważań – zbiór rozwiązań możemy teraz zapisywać jako zbiór sum wektorów z odpowiedniej przestrzeni liniowej.

Obserwacja 4.2

Dla każdej liczby naturalnej m niech $K_{\leqslant m}[x]$ oznacza zbiór wszystkich wielomianów stopnia co najwyżej m w $K[x]$. Jest to podprzestrzeń $K[x]$.

Interpretując zbiór wielomianów $K[x]$ jako podzbiór K^∞ złożony z ciągów mających jedynie skończenie niezerowych wyrazów możemy zauważyć, że $K[x]$ można w istocie traktować jako podprzestrzeń K^∞. Oczywiście przykładów podprzestrzeni K^∞ można wskazać więcej.

Na ćwiczeniach omawiać Państwo będą, w ramach rozmaitych przykładów, szereg podprzestrzeni w przestrzeni $\mathbb{R}^\infty$ oraz $F(\mathbb{R}, \mathbb{R})$, mających związki z analizą. Warto zwrócić uwagę na kilka z nich.

Obserwacja 4.3

W przestrzeni ciągów $\mathbb{R}^\infty$ wskazać można bardzo wiele podprzestrzeni, np.:

- ciągi mające skończenie wiele niezerowych wyrazów,
- ciągi ograniczone,
- ciągi zbieżne,
- ciągi $(x_i)_{i=1}^\infty$ spełniające $\sum_{i=1}^{\infty} x_i^2 < \infty$.
- ciągi $(x_i)_{i=1}^\infty$ spełniające określone rekurencje liniowe, np. $x_{n+2} = x_{n+1} + x_n$.

Obserwacja 4.4

Przykłady podprzestrzeni w przestrzeni funkcji $F(K, K)$:

- funkcje parzyste, spełniające równanie $f(x) = f(-x)$, dla $x \in K$,
- funkcje nieparzyste, spełniające równanie $f(x) = -f(-x)$, dla $x \in K$,
- nad $\mathbb{R}$ (i nie tylko): funkcje ograniczone, monotoniczne itd.
- funkcje będące rozwiązaniami równania Cauchy'ego[a], tzn. dla każdych $x, y \in K$:

$$f(x + y) = f(x) + f(y),$$

[a] To słynne równanie funkcyjne rozważane dla funkcji rzeczywistych badane było przez wielkich matematyków, jak Cauchy, Darboux, d' Alembert i inni. Przy niewielu dodatkowych założeniach można pokazać, że jego rozwiązaniami są jedynie funkcje postaci $f(x) = ax$, dla $a \in \mathbb{R}$. Do tych „drobnych" dodatkowych założeń należą: ciągłość (Cauchy, 1821), ciągłość w punkcie (Darboux, 1875), monotoniczność lub ograniczoność na dowolnym przedziale (Darboux, 1880). W 1905 roku Georg Hamel pokazał, używając aksjomatu wyboru, że bez przyjęcia tego typu założeń o regularności wskazać można znacznie bardziej skomplikowane i egzotyczne funkcje spełniające równanie Cauchy'ego.

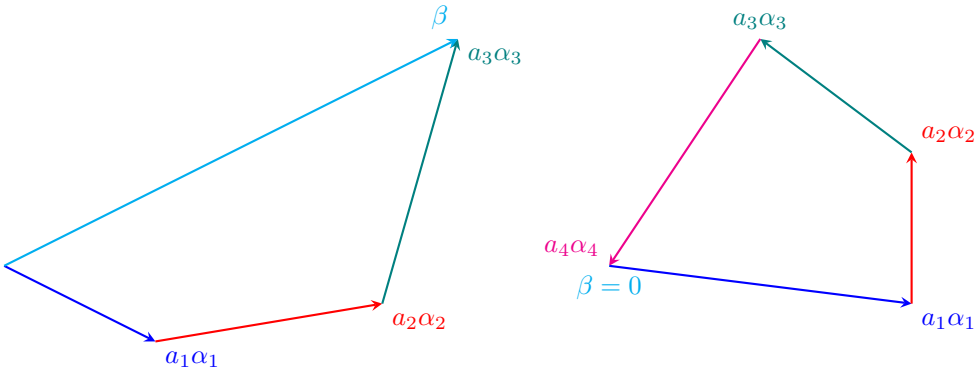
Powiemy teraz o bardzo ważnym typie konstrukcji związanych z podprzestrzeniami. Chodzi o sytuację, gdy mamy przestrzeń liniową i szukamy takiej jej podprzestrzeni, która zawierałaby z góry określone przez nas wektory – wybranej przy tym możliwie oszczędnie.

Definicja 4.9: Kombinacja liniowa

Niech V będzie przestrzenią liniową nad ciałem K. KOMBINACJĄ LINIOWĄ układu wektorów $\alpha_1, \ldots, \alpha_k$ o współczynnikach $a_1, \ldots, a_k \in K$ nazywamy wektor:

$$\beta = a_1\alpha_1 + \ldots + a_k\alpha_k = \sum_{i=1}^{k} a_i\alpha_i.$$

Na poziomie intuicji geometrycznej wektor β jest kombinacją liniową $\sum_{i=1}^{k} a_i\alpha_i$ układu wektorów, gdy złożenie przesunięć o wektory $a_1\alpha, a_2\alpha_2, \ldots, a_n\alpha_n$ jest równe przesunięciu o wektor β. W szczególności, gdy $\beta = 0$ oznacza to, że złożenie tych przesunięć jest identycznością, czyli przesunięciem o wektor zerowy.



Przykłady.

- W przestrzeni $V = \mathbb{R}^4$ kombinacją liniową wektorów

  $$(2, 1, -3, 4), (0, 2, 5, 1), (7, 4, 3, 2)$$

  ze współczynnikami $2, -1, 1$ jest wektor

  $$2 \cdot (2, 1, -3, 4) - 1 \cdot (0, 2, 5, 1) + 1 \cdot (7, 4, 5, 2) = (11, 4, -8, 9).$$

- W przestrzeni funkcji $F(\mathbb{R}, \mathbb{R})$ kombinacją liniową wektorów $\sin(x)$ oraz $\cos(x)$ o współczynnikach $\frac{1}{\sqrt{2}}$ oraz $-\frac{1}{\sqrt{2}}$ jest funkcja

  $$\frac{1}{\sqrt{2}}\sin(x) - \frac{1}{\sqrt{2}}\cos(x) = \sin(x - \frac{\pi}{4}).$$

- Wektor $(0, 3, 1) \in \mathbb{R}^3$ nie jest kombinacją liniową wektorów $(0, 1, 1), (-1, 0, 1)$, bo założenie, że

  $$(0, 3, 1) = a(0, 1, 1) + b(-1, 0, 1)$$

  prowadzi do układu równań $0 = -b$, $3 = a$, $1 = a + b$, który nie ma rozwiązań.

- Wektory $(1, 1, -2), (1, 0, -1) \in \mathbb{R}^3$ są rozwiązaniami równania jednorodnego $x_1 + x_2 + x_3 = 0$, skąd wynika, że każda ich kombinacja liniowa $a(1, 1, -2) + b(1, 0, -1) = (a + b, a, -2a - b)$ jest również rozwiązaniem tego równania.

- Jeśli $\beta_1, \beta_2, \ldots, \beta_r \in K^n$ są rozwiązaniami układu liniowych równań jednorodnych U, to również

  $$a_1\beta_1 + a_2\beta_2 + \ldots + a_r\beta_r$$

  są rozwiązaniami tego układu, dla dowolnych układów skalarów $a_1, a_2, \ldots, a_r \in K$.

Obserwacja 4.5

Niech $\alpha_1, \ldots, \alpha_k$ będą wektorami przestrzeni liniowej V nad K. Jeśli wektory β, γ są kombinacjami liniowymi wektorów $\alpha_1, \ldots, \alpha_k$, to wektory $\beta + \gamma$ oraz $a\beta$, dla każdego $a \in K$, również są kombinacjami liniowymi wektorów $\alpha_1, \ldots, \alpha_k$.

Dowód. Niech $a_1, \ldots, a_k$ oraz $b_1, \ldots, b_k$ będą elementami ciała K oraz niech

$$\beta = a_1\alpha_! + \ldots + a_k\alpha_k, \quad \gamma = b_1\alpha_1 + \ldots + b_k\alpha_k.$$

Wówczas

$$\beta + \gamma = (a_1 + b_1)\alpha_1 + \ldots + (a_k + b_k)\alpha_k$$

oraz dla każdego $a \in K$ mamy

$$a\beta = aa_1\alpha_1 + \ldots + aa_k\alpha_k.$$

□

Definicja 4.10: Podprzestrzeń rozpięta na układzie wektorów

Niech V będzie przestrzenią liniową nad ciałem K i niech $\alpha_1, \ldots, \alpha_k \in V$. Wówczas przez

$$\text{lin}(\alpha_1, \ldots, \alpha_k)$$

oznaczamy zbiór wszystkich kombinacji liniowych wektorów $\alpha_1, \ldots, \alpha_k$.

Poprzednią obserwację możemy teraz wyrazić w następujący sposób.

Obserwacja 4.6

Zbiór $\text{lin}(\alpha_1, \ldots, \alpha_k)$ jest podprzestrzenią przestrzeni V. Podprzestrzeń ta jest najmniejszą podprzestrzenią V (względem inkluzji) zawierającą wektory $\alpha_1, \ldots, \alpha_k$.

Dowód. Z Obserwacji 4.5 wynika, że $\text{lin}(\alpha_1, \ldots, \alpha_k)$ jest podprzestrzenią w V. Niech W będzie dowolną podprzestrzenią zawierającą wektory $\alpha_1, \ldots, \alpha_k$. Z definicji podprzestrzeni W zawiera każdą kombinację liniową wektorów $\alpha_1, \ldots, \alpha_k$, czyli każdy wektor z $\text{lin}(\alpha_1, \ldots, \alpha_k)$. Stąd $\text{lin}(\alpha_1, \ldots, \alpha_k) \subseteq W$. □

Definicja 4.11: Układ wektorów rozpinający podprzestrzeń

Niech $\alpha_1, \ldots, \alpha_k$ będzie układem wektorów w V. Wówczas podprzestrzeń liniową $\text{lin}(\alpha_1, \ldots, \alpha_k)$ nazywamy PRZESTRZENIĄ ROZPIĘTĄ NA UKŁADZIE $\alpha_1, \ldots, \alpha_k$. Mówimy, że układ $\alpha_1, \ldots, \alpha_k$ ROZPINA PRZESTRZEŃ V, jeśli $V = \text{lin}(\alpha_1, \ldots, \alpha_k)$, to znaczy każdy wektor z V jest kombinacją liniową wektorów $\alpha_1, \ldots, \alpha_k$.

Wracając do przykładu układu równań liniowych rozważanego wyżej:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_1 - x_2 - x_3 - x_4 &= 0 \end{cases}$$

możemy zauważyć, że zbiór rozwiązań tego układu jest podprzestrzenią $\mathbb{R}^4$ rozpiętą przez wektory $(0, -1, 1, 0), (0, -1, 0, 1)$, czyli jest to zbiór

$$\text{lin}((0, -1, 1, 0), (0, -1, 0, 1)).$$

Zauważmy, że zbiór rozwiązań powyższego układu można zapisać na wiele innych sposobów. Można wziąć np. rozwiązania $(0,0,-1,1)$ oraz $(0,-2,1,1)$ i zauważyć, że zbiór rozwiązań powyższego układu jest równy:

$$\text{lin}((0,0,-1,1),(0,-2,1,1)).$$

Co więcej, nic nie stoi na przeszkodzie, by rozważyć zbiór wszystkich kombinacji liniowych postaci:

$$s(0,0,-1,1)+t(0,-1,1,0)+r(0,-1,0,1), \quad s,t,r\in\mathbb{R}$$

i jest to również zbiór rozwiązań układu powyżej! Innymi słowy mamy równości:

$$\text{lin}((0,-1,1,0),(0,-1,0,1))=\text{lin}((0,0,-1,1),(0,-2,1,1))=\text{lin}((0,0,-1,1),(0,-1,1,0),(0,-1,0,1)).$$

Przy badaniu przestrzeni rozpiętych na układach wektorów w K^n użyteczna jest prosta obserwacja.

Obserwacja 4.7

Niech $A, A'\in M_{m\times n}(K)$ oraz niech

- $\alpha_1,\ldots,\alpha_m$ – wiersze macierzy A traktowane jako wektory w K^n,
- $\alpha'_1,\ldots,\alpha'_m$ – wiersze macierzy A' traktowane jako wektory w K^n.

Jeśli założymy, że A' może być otrzymana z A za pomocą ciągu operacji elementarnych na wierszach, to wynika stąd, że

$$\text{lin}(\alpha_1,\ldots,\alpha_m)=\text{lin}(\alpha'_1,\ldots,\alpha'_m).$$

Zanim pokażemy dowód, przedstawmy przykład zaczerpnięty z przestrzeni $\mathbb{R}^4$. Weźmy układ wektorów postaci

$$(2,1,1,1),(4,2,2,0),(0,0,0,1)$$

z przestrzeni $\mathbb{R}^4$. Wektory te traktować możemy jako wiersze macierzy o czterech kolumnach:

$$\begin{bmatrix}2&1&1&1\\4&2&2&0\\0&0&0&1\end{bmatrix}.$$

Po dodaniu dwukrotności trzeciego wiersza do drugiego wiersza, a następnie po przemnożeniu pierwszego wiersza przez 2 otrzymujemy macierz

$$\begin{bmatrix}4&2&2&2\\4&2&2&2\\0&0&0&1\end{bmatrix},$$

a zatem dostajemy równość:

$$\text{lin}((2,1,1,1),(4,2,2,0),(0,0,0,1))=\text{lin}((4,2,2,2),(4,2,2,2),(0,0,0,1)).$$

Zauważmy teraz, że układ rozpinający $\text{lin}((2,1,1,1),(4,2,2,0),(0,0,0,1))$ można pomniejszyć. Po odjęciu pierwszego wiersza od drugiego, a następnie po zamianie drugiego i trzeciego wiersza mamy:

$$\begin{bmatrix}4&2&2&2\\0&0&0&1\\0&0&0&0\end{bmatrix},$$

A zatem możemy napisać (i bez macierzy mogliśmy):

$$\text{lin}((2,1,1,1),(4,2,2,0),(0,0,0,1))=\text{lin}((4,2,2,2),(0,0,0,1)).$$

Czytelnik zechce zauważyć, że powyższa podprzestrzeń nie może być rozpięta przez jeden wektor – to by wymagało, by po pewnej liczbie operacji na wierszach możliwe było osiągnięcie macierzy o dwóch wierszach zerowych. Dlaczego nie jest to możliwe?

Dowód. Wystarczy pokazać tezę w przypadku, gdy A' powstaje z A przez wykonanie pojedynczej operacji elementarnej na wierszach. Wykażemy tezę jedynie w najtrudniejszym przypadku. Pokazujemy mianowicie, że dla dowolnych $1 \leqslant i, j \leqslant m$ oraz dowolnego $a \in K$ mamy:

$$\lin(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_m) = \lin(\alpha_1, \ldots, \alpha_i, \ldots, a \cdot \alpha_i + \alpha_j, \ldots, \alpha_m).$$

Weźmy $\beta \in \lin(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_m)$. Istnieją $b_1, b_2, \ldots, b_m \in K$, że:

$$\begin{aligned}\beta &= b_1\alpha_1 + b_2\alpha_2 + \ldots + b_i\alpha_i + \ldots + b_j\alpha_j + \ldots + b_m\alpha_m = \\ &= b_1\alpha_1 + b_2\alpha_2 + \ldots + (b_i - a \cdot b_j)\alpha_i + \ldots + b_j(a\alpha_i + \alpha_j) + \ldots + b_m\alpha_m.\end{aligned}$$

Zatem $\beta \in \lin(\alpha_1, \ldots, \alpha_i, \ldots, a \cdot \alpha_i + \alpha_j, \ldots, \alpha_m)$.

Weźmy teraz $\gamma \in \lin(\alpha_1, \ldots, \alpha_i, \ldots, a \cdot \alpha_i + \alpha_j, \ldots, \alpha_m)$. Istnieją $c_1, c_2, \ldots, c_m \in K$, że:

$$\begin{aligned}\gamma &= c_1\alpha_1 + c_2\alpha_2 + \ldots + c_i\alpha_i + \ldots + c_j(a \cdot \alpha_i + \alpha_j) + \ldots + c_m\alpha_m = \\ &= c_1\alpha_1 + c_2\alpha_2 + \ldots + (c_i + a \cdot c_j)\alpha_i + \ldots + c_j\alpha_j + \ldots + c_m\alpha_m.\end{aligned}$$

Zatem $\gamma \in \lin(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_m)$. □

* * *

Pamiętajmy, że rozważać można przykłady podprzestrzeni rozpiętych na układach wektorów w innych przestrzeniach niż K^n, np. $K[x], K^\infty, M_{n\times m}(K)$ czy $F(K, K)$ i wtedy korzystać należy raczej z definicji wprowadzonych wcześniej, nie zaś z metody opisanej wyżej. Dla przykładu, w przestrzeni $\mathbb{R}[x]$ mamy

$$7x^2 + 4x - 3 \in \lin(4x^2 + x, x^2 - 2x + 3),$$

bowiem

$$7x^2 + 4x - 3 = 2(4x^2 + x) - 1 \cdot (x^2 - 2x + 3).$$

Nietrudno też widzieć, że na przykład zachodzi równość:

$$\lin(7x^2 + 4x - 3, 9x - 12) = \lin(4x^2 + x, x^2 - 2x + 3).$$

Rzeczywiście, mamy

$$\lin(7x^2 + 4x - 3, 9x - 12) \subseteq \lin(4x^2 + x, x^2 - 2x + 3),$$

ponieważ:

- wektor $7x^2 + 4x - 3 = 2 \cdot (4x^2 + x) - 1 \cdot (x^2 - 2x + 3)$ należy do $\lin(4x^2 + x, x^2 - 2x + 3)$,
- wektor $9x - 12 = 1 \cdot (4x^2 + x) - 4 \cdot (x^2 - 2x + 3)$ należy do $\lin(4x^2 + x, x^2 - 2x + 3)$,
- każda kombinacja liniowa wektorów $7x^2 + 4x - 3$ oraz $9x - 12$ jest, zgodnie z Obserwacją 4.6, kombinacją liniową wektorów $4x^2 + x, x^2 - 2x + 3$, czyli należy do $\lin(4x^2 + x, x^2 - 2x + 3)$.

Podobnie dowodzimy przeciwną inkluzję:

$$\lin(7x^2 + 4x - 3, 9x - 12) \supseteq \lin(4x^2 + x, x^2 - 2x + 3),$$

ponieważ

- wektor $4x^2 + x = 1 \cdot (9x - 12) + 4 \cdot (x^2 - 2x + 3)$ należy do $\lin(7x^2 + 4x - 3, 9x - 12)$,
- wektor $x^2 - 2x + 3 = 2 \cdot (4x^2 + x) - 1 \cdot (7x^2 + 4x - 3)$ należy do $\lin(7x^2 + 4x - 3, 9x - 12)$,
- każda kombinacja liniowa wektorów $4x^2 + x$ oraz $x^2 - 2x + 3$ jest, zgodnie z Obserwacją 4.6, kombinacją liniową wektorów $7x^2 + 4x - 3, 9x - 12$, czyli należy do $\lin(7x^2 + 4x - 3, 9x - 12)$.

Pojęcie przestrzeni liniowej to pierwszy krok w kierunku uzyskania nowej geometrycznej perspektywy na rozmaite obiekty matematyczne. Na kolejnym wykładzie zastanowimy się nad fundamentalnym problemem: ile elementów z przestrzeni liniowej rozpiętej przez n wektorów musimy znać, aby przestrzeń ta była wyznaczona jednoznacznie oraz jakie własności mają takie „minimalne układy rozpinające".

4.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Sprawdzanie czy dana trójka spełnia aksjomaty przestrzeni liniowej)
   Niech $X = \{x \in \mathbb{R} : x > 0\}$. Zdefiniujmy dodawanie $\oplus$ elementów zbioru X oraz mnożenie $\odot$ elementów zbioru X przez liczby rzeczywiste wzorami $x \oplus y = xy$ oraz $\lambda \odot x = x^\lambda$. Wykaż, że trójka $(X, \oplus, \odot)$ jest, przy odpowiednim wyborze wektora zerowego, przestrzenią liniową nad ciałem $\mathbb{R}$.

2. (♠ Sprawdzanie czy podziór przestrzeni liniowej spełnia warunki z definicji podprzestrzeni)
   Dla każdego z poniższych podzbiorów $\mathbb{R}^2$ sprawdź, czy spełnia on warunek (i) oraz czy spełnia on warunek (ii) z definicji podprzestrzeni.

   (a) $\{(x_1, x_2) : x_1, x_1 \in \mathbb{Z}\}$,
   (b) $\{(x_1, x_2) : x_1 = 0 \text{ lub } x_2 = 0\}$,
   (c) $\{(x_1, x_2) : |x_1| - |x_2| = 1\}$,
   (d) $\{(x_1, x_2) : x_1^2 + x_2^2 = 2x_1x_2\}$.

3. Rozważmy przestrzeń liniową $V = \mathbb{R}^n$ i niech W będzie podzbiorem V składającym się z wektorów $(x_1, \ldots, x_n)$, takich że

   (a) $x_n = 0$,
   (b) $x_1 + \ldots + x_n = 1$,
   (c) $x_1 + \ldots + x_n = 0$,
   (d) $x_1 + \ldots + x_n \geqslant 0$,
   (e) $x_i = x_{n+1-i}$, dla $i = 1, 2, \ldots, n$.

   W którym z powyższych przypadków W jest podprzestrzenią V?

4. Dla jakich wartości $s \in \mathbb{R}$ następujący podzbiór $W \subseteq \mathbb{R}^4$ jest podprzestrzenią?

$$W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 + 2x_3 + 3x_4 = s^2 - s - 2,\ x_1 + s^2x_3^2 + x_4 = x_3^2.\}$$

5. Rozważmy podzbiór $\mathcal{S}$ przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$, złożony z takich funkcji $f(x)$, dla których istnieją $a, b \in \mathbb{R}$, że $f(x) = a\sin(x + b)$. Czy jest to podprzestrzeń przestrzeni liniowej $F(\mathbb{R}, \mathbb{R})$?

6. Niech V będzie przestrzenią liniową nad ciałem $\mathbb{R}$ złożoną z ciągów o wyrazach rzeczywistych, czyli: $V = \{(a_i)_{i\in\mathbb{N}} : a_i \in \mathbb{R}\}$. Niech $W = \{(a_i) : \forall_{i\in\mathbb{N}}\ a_{i+1} \leqslant a_i\} \subset V$. Czy W jest podprzestrzenią liniową przestrzeni V?

7. Niech W_1, W_2 będą podprzestrzeniami przestrzeni liniowej V. Wykaż, że zbiór $W_1 \cup W_2$ jest podprzestrzenią przestrzeni V wtedy i tylko wtedy, gdy $W_1 \subset W_2$ lub $W_2 \subset W_1$.

8. (♠ Sprawdzanie czy wektor jest kombinacją liniową innych)

   (a) Czy wektor $(2, 1, 2, 1)$ należy do $\text{lin}((1, 2, 0, 2), (1, 0, 3, 1), (0, 1, 0, 2), (1, 1, 2, 0)) \subseteq \mathbb{R}^4$?
   (b) Czy wektor $(1, 1, 1, 1)$ należy do $\text{lin}((1, 1, 0, 0), (0, 1, 1, 0), (0, 0, 1, 1)) \subset \mathbb{Z}_2^4$?
   (c) Czy wielomian $x - x^3$ należy do $\text{lin}(x^2, 2x + x^2, x + x^3) \subset R[x]$?
   (d) Czy macierz $\begin{bmatrix} -2 & -4 \\ -2 & -6 \end{bmatrix}$ należy do $\text{lin}\left(\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}\right) \subset M_{2\times2}(\mathbb{Q})$?
   (e) Czy funkcja $\cos(3x)$ należy do $\text{lin}(1, \sin(x), \sin^2(x), \sin^3(x)) \subset F(\mathbb{R}, \mathbb{R})$?

9. (♠ Opis zbioru rozwiązań jednorodnego układu równań liniowych jako podprzestrzeni K^n rozpiętej na układzie wektorów) Zapisz zbiór rozwiązań poniższego układu równań jako podprzestrzeń $\text{lin}(v_1, v_2, v_3)$ przestrzeni $\mathbb{R}^5$

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 & = 0 \\ 2x_1 - x_2 - 2x_3 + x_4 + x_5 & = 0 \\ -x_1 + x_3 + 4x_4 + x_5 & = 0 \end{cases}.$$

10. W przestrzeni liniowej V dane są wektory u, v, w. Czy zachodzi równość

$$\text{lin}(u, v, w) = \text{lin}(u + v, v + w, w + u)?$$

4.3 Uzupełnienie. Kombinacje liniowe i układy równań

Jedną z przestrzeni liniowych poznanych na wykładzie jest przestrzeń macierzy o m wierszach i n kolumnach o wyrazach z ciała K. Nietrudno zauważyć, że dodawanie macierzy lub mnożenie ich przez skalar są w zasadzie identyczne z operacjami wprowadzonymi w przestrzeni $K^{m\cdot n}$. Aby to unaocznić weźmy na przykład sumę macierzy w $M_{2\times 3}(\mathbb{Q})$ oraz sumę wektorów w $\mathbb{Q}^6$ postaci:

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}+\begin{bmatrix}6&5&4\\3&2&1\end{bmatrix}=\begin{bmatrix}7&7&7\\7&7&7\end{bmatrix},\quad (1,2,3,4,5,6)+(6,5,4,3,2,1)=(7,7,7,7,7,7).$$

Wkrótce poznamy język, który pozwoli nam powiedzieć, że z punktu widzenia „struktury" przestrzeni liniowych przestrzenie $M_{2\times 3}(\mathbb{Q})$ oraz $\mathbb{Q}^6$ w zasadzie niczym się nie różnią – są IZOMORFICZNE. Dlaczego więc rozróżniamy te dwie przestrzenie? Macierze okazały się wygodnym narzędziem do badania układów równań. Jak niedługo zobaczymy, są one również wygodnym narzędziem do badania przekształceń pomiędzy przestrzeniami liniowymi. Jest jeden przypadek, gdy utożsamienie wektorów z macierzami wykonać można bez żadnych dodatkowych umów: gdy rozważamy macierze o jednym wierszu lub jednej kolumnie. Zajmiemy się teraz drugą sytuacją.

Zapiszmy równań liniowych nad $\mathbb{R}$ za pomocą operacji w $M_{3\times 1}(\mathbb{R})$:

$$\begin{cases}x\quad -z&=0\\2x+y&=0\\3x+y+z&=0\end{cases}\quad\Rightarrow\quad x\cdot\begin{bmatrix}1\\2\\3\end{bmatrix}+y\cdot\begin{bmatrix}0\\1\\1\end{bmatrix}+z\cdot\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Widzimy zatem, że rozwiązanie układu równań sprowadza się do sprawdzenia, czy pewna macierz rozmiaru 3×1 jest kombinacją liniową pewnych trzech macierzy ze współczynnikami x,y,z. Jest jednak jasne, że w istocie jest to zagadnienie równoważne z przedstawieniem wektora $(0,0,0)\in\mathbb{R}^3$ jako kombinacji liniowej wektorów $(1,2,3),(0,1,1),(-1,0,1)$. Często mówimy nawet, że wektory te zapisane zostały w równaniu wyżej w notacji kolumnowej. A zatem w dalszym ciągu często dokonywać będziemy utożsamienia elementów K^n oraz przestrzeni macierzy $M_{1\times n}(K)$ oraz $M_{n\times 1}(K)$ mówiąc przy tym, że wektor $v\in K^n$ zapisujemy w formie kolumnowej v^T lub wierszowej v.

Rozwiązywanie układów równań przez poszukiwanie kombinacji liniowych nie przyspieszy samego procesu rozwiązywania (dalej stosować będziemy metodę Gaussa), ale pozwoli nam zadać kilka istotnych pytań. Wróćmy do układu wyżej i zapytajmy: czy jeśli zamienimy wektor $(0,0,0)$ na dowolny inny, układ pozostanie niesprzeczny? A zatem: czy dowolny wektor $(a,b,c)\in\mathbb{R}^3$ jest kombinacją liniową wektorów $(1,2,3),(0,1,1),(-1,0,1)$? Zupełnie wprost: czy dla każdych a,b,c istnieją x,y,z takie, że

$$x\cdot\begin{bmatrix}1\\2\\3\end{bmatrix}+y\cdot\begin{bmatrix}0\\1\\1\end{bmatrix}+z\cdot\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}a\\b\\c\end{bmatrix}?$$

Skąd mamy wiedzieć coś takiego i jak wyznaczyć x,y,z? Okazuje się, że nie jest to trudne. W języku kombinacji liniowych nasze pytanie brzmi: czy dowolny wektor z $\mathbb{R}^3$ jest kombinacją liniową wektorów $(1,2,3),(0,1,1),(-1,0,1)$? W skrócie, pytamy o prawdziwość równości:

$$\operatorname{lin}((1,2,3),(0,1,1),(-1,0,1))=\mathbb{R}^3.$$

Czy to może być prawda? Nietrudno się przekonać, że tak jest: twierdzenie wykazane na wykładzie mówi, że wpisując powyższe trzy wektory w wiersze możemy wykonywać operacje wierszowe i przekonać się, że ciągiem operacji elementarnych na wierszach macierz

$$\begin{bmatrix}1&2&3\\0&1&1\\-1&0&1\end{bmatrix}$$

sprowadzić można do macierzy:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

Jest natomiast jasne, że $\operatorname{lin}((1,0,0),(0,1,0),(0,0,1))=\mathbb{R}^3$, bo dla każdego $(a,b,c)\in\mathbb{R}^3$ mamy $(a,b,c)=a(1,0,0)+b(0,1,0)+c(0,0,1)$.

A zatem odpowiedzieliśmy na pytanie o rozwiązywalność dowolnego układu niejednorodnego o pewnej konkretnej macierzy współczynników. A jak wygląda rozwiązanie dla konkretnych a, b, c? Zobaczmy nasz układ w jeszcze innej postaci:

$$x \cdot \begin{bmatrix}1\\2\\3\end{bmatrix} + y \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} + z \cdot \begin{bmatrix}-1\\0\\1\end{bmatrix} = a \begin{bmatrix}1\\0\\0\end{bmatrix} + b \begin{bmatrix}0\\1\\0\end{bmatrix} + c \begin{bmatrix}0\\0\\1\end{bmatrix}$$

Przypomnijmy, że jeśli zaczniemy wykonywać jednocześnie te same operacje na wierszach następujących macierzy:

$$\begin{bmatrix}1 & 0 & -1\\2 & 1 & 0\\3 & 1 & 1\end{bmatrix}, \quad \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix},$$

wówczas kombinacje liniowe ich kolumn ze współczynnikami x, y, z oraz a, b, c będą nadal równe! Zobaczmy to. Wykonajmy dwie operacje na obydwu macierzach:

$$\begin{bmatrix}1 & 0 & -1\\2 & 1 & 0\\3 & 1 & 1\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & -1\\0 & 1 & 2\\3 & 1 & 1\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & -1\\0 & 1 & 2\\0 & 1 & 4\end{bmatrix}, \quad \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & 0\\-2 & 1 & 0\\0 & 0 & 1\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 0 & 0\\-2 & 1 & 0\\-3 & 0 & 1\end{bmatrix}$$

Mamy:

$$x \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} + y \cdot \begin{bmatrix}0\\2\\1\end{bmatrix} + z \cdot \begin{bmatrix}-1\\2\\4\end{bmatrix} = a \begin{bmatrix}1\\-2\\-3\end{bmatrix} + b \begin{bmatrix}0\\1\\0\end{bmatrix} + c \begin{bmatrix}0\\0\\1\end{bmatrix}$$

Czy Czytelnik widzi, że kontynuując proces schodkowania macierzy wyjściowego układu równań dojdziemy w końcu do postaci pozwalającej wyznaczyć x, y, z za pomocą a, b, c? Kontynuujmy eliminację, tym razem zapisując już macierze obok siebie:

$$\begin{bmatrix}1 & 0 & -1 & 1 & 0 & 0\\0 & 1 & 2 & -2 & 1 & 0\\0 & 1 & 4 & -3 & 0 & 1\end{bmatrix} \longrightarrow \left[\begin{array}{ccc|ccc}1 & 0 & -1 & 1 & 0 & 0\\0 & 1 & 2 & -2 & 1 & 0\\0 & 0 & 6 & -1 & -1 & 1\end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc}1 & 0 & -1 & 1 & 0 & 0\\0 & 1 & 2 & -2 & 1 & 0\\0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6}\end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc}1 & 0 & 0 & \frac{5}{6} & -\frac{1}{6} & \frac{1}{6}\\0 & 1 & 0 & -\frac{10}{6} & \frac{4}{6} & \frac{2}{6}\\0 & 0 & 1 & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6}\end{array}\right]$$

A zatem mamy:

$$x \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} + y \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} + z \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = a \begin{bmatrix}\frac{5}{6}\\-\frac{5}{3}\\-\frac{1}{6}\end{bmatrix} + b \begin{bmatrix}-\frac{1}{6}\\\frac{2}{3}\\-\frac{1}{6}\end{bmatrix} + c \begin{bmatrix}\frac{1}{6}\\\frac{1}{3}\\\frac{1}{6}\end{bmatrix}$$

Po uproszczeniu:

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}\frac{5}{6}a - \frac{1}{6}b + \frac{1}{6}c\\-\frac{5}{3}a + \frac{2}{3}b + \frac{1}{3}c\\-\frac{1}{6}a - \frac{1}{6}b + \frac{1}{6}c\end{bmatrix}.$$

Wracając do wyjściowego problemu widzimy, że rozwiązaniem układu:

$$\begin{cases} x \quad - z & = a\\ 2x + y & = b\\ 3x + y + z & = c \end{cases}$$

jest trójka:

$$\left(\frac{5}{6}a - \frac{1}{6}b + \frac{1}{6}c, \quad -\frac{5}{3}a + \frac{2}{3}b + \frac{1}{3}c, \quad -\frac{1}{6}a - \frac{1}{6}b + \frac{1}{6}c\right).$$

Jeśli Czytelnik dotrwał do tego momentu, to gratuluję: odwróciliśmy właśnie wspólnie pierwszą macierz. Nie wiemy na razie co to znaczy, ale sam termin „odwrócenia" powinien rodzić jasne skojarzenia. Rozpisaliśmy ustalony wektor jako kombinację liniową trzech zadanych z góry wektorów. Nie zawsze będzie to jednak możliwe. Proszę zauważyć, że gdyby zamiast wektorów $(1,2,3), (0,1,1), (-1,0,1)$ szukać kombinacji liniowych wektorów: $(1,2,3), (2,4,6), (-1,0,1)$, to nie każdy wektor $\mathbb{R}^3$ byłby ich kombinacją liniową. Inaczej mówiąc $\text{lin}((1,2,3), (2,4,6), (-1,0,1)) \neq \mathbb{R}^3$. Sprawom tym przyjrzymy się już na następnym wykładzie.

4.4 Dodatek. Ciało jako przestrzeń liniowa nad podciałem

Pojęcie wielomianu o współczynnikach w ciele K oraz pojęcie pierwiastka wielomianu, pozwalają na istotne wzbogacenie naszego zasobu przykładów ciał. Rozważmy następującą sytuację. W zbiorze liczb rzeczywistych wybieramy wszystkie liczby postaci:

$$\{a + b\sqrt{2} \,|\, a, b \in \mathbb{Q}\}.$$

Zbiór ten oznaczamy jako $\mathbb{Q}(\sqrt{2})$. Czytelnik zechce zauważyć, że $a+b\sqrt{2} = c+d\sqrt{2}$ wtedy i tylko wtedy, gdy $a = c$ oraz $b = d$ (wynika to z niewymierności liczby $\sqrt{2}$). Co więcej, wprowadzenie w powyższym zbiorze działań dodawania i mnożenia liczb rzeczywistych prowadzi do zauważenia, że wniosku, że zbiór ten jest ciałem. Rzeczywiście, dla dowolnych $a + b\sqrt{2}$ oraz $c + d\sqrt{2}$ należących do $\mathbb{Q}(\sqrt{2})$ liczby

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}, \quad (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$$

są elementami $\mathbb{Q}(\sqrt{2})$. Odrobina wysiłku, między innymi zauważenie równości:

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

pokazuje, że $\mathbb{Q}(\sqrt{2})$ jest ciałem. Jest to przykład tzw. podciała ciała $\mathbb{R}$. Przejdźmy do ogólnej definicji.

Definicja 4.12: Podciało i rozszerzenie ciał

Mówimy, że piątka $(K', +', \cdot', 0', 1')$ jest PODCIAŁEM ciała $(K, +, \cdot, 0, 1)$ jeśli K' jest podzbiorem ciała K, $0' = 0$, $1' = 1$ oraz działania $+'$ i $\cdot'$ powstają przez ograniczenie działań $+, \cdot$ określonych na $K \times K$ do zbioru $K' \times K'$.

Parę $K' \subset K$, gdzie K' jest podciałem ciała K nazywamy ROZSZERZENIEM CIAŁ.

Najbardziej znanym podciałem ciała liczb rzeczywistych jest ciało liczb wymiernych $\mathbb{Q}$ ze zwykłymi działaniami dodawania, mnożenia oraz z wyróżnionymi elementami 0 i 1. Innym przykładem podciała jest $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \,|\, a, b \in \mathbb{Q}\}$. Jest to tak zwane rozszerzenie kwadratowe ciała $\mathbb{Q}$ o $\sqrt{2}$. Aby zdefiniować czym jest owo rozszerzenie powiedzmy kilka słów o podciałach ustalonego ciała.

Obserwacja 4.8

Rozważmy dowolną rodzinę podciał K_t ciała L, gdzie $t \in T$. Wówczas część wspólna wszystkich ciał K_t jest podciałem ciała K.

Obserwacja 4.9

Dla każdego podciała K ciała L oraz podzbioru S zbioru L istnieje najmniejsze podciało $K(S)$ ciała L, które zawiera jednocześnie ciało K oraz zbiór S.

Oto przykłady podciał ciała liczb rzeczywistych, utworzone w oparciu o powyższe obserwacje:

- ciała $\mathbb{Q}(\sqrt{p})$, gdzie p jest liczbą pierwszą,
- ciała $\mathbb{Q}(\sqrt{p}, \sqrt{q})$, gdzie p, q są liczbami pierwszymi,
- najmniejsze ciało zawierające $\mathbb{Q}$ i pierwiastki z wszystkich liczb pierwszych,
- ciała typu $\mathbb{Q}(\pi)$, $\mathbb{Q}(\sqrt{2}, \pi)$, $\mathbb{Q}(\pi, \pi^2, \pi^3, \ldots)$ itd.

Jak ująć powyższe zagadnienia w języku przestrzeni liniowych? Otóż należy zdać sobie sprawę, że jeśli ciało K jest podciałem ciała L, to ciało L traktować można jako przestrzeń liniową nad ciałem K. Zobaczmy kilka przykładów.

- Ciało $\mathbb{C}$ jest przestrzenią liniową nad ciałem $\mathbb{R}$. A zatem liczby zespolone traktujemy jako wektory, które mnożymy jedynie przez liczby rzeczywiste (zapominamy o tym, że liczby zespolone można też mnożyć). Zauważmy, że biorąc w tym ujęciu wektory $1, i \in \mathbb{C}$ widzimy, że każda liczba zespolona jest kombinacją liniową wektorów α oraz β, czyli

  $$z = a \cdot 1 + b \cdot i, \quad a, b \in \mathbb{R}$$

  Innymi słowy: $\mathbb{C} = \text{lin}(1, i) = \mathbb{R}(i)$.

- Ciało $\mathbb{Q}(\sqrt{2})$ jest przestrzenią liniową nad ciałem $\mathbb{Q}$. Bardzo podobnie jak wyżej widzimy, że $\mathbb{Q}(\sqrt{2}) = \text{lin}(1, \sqrt{2})$, przy czym teraz skalarami sa liczby wymierne, a wektorami – liczby postaci

  $$a \cdot 1 + b \cdot \sqrt{2}, \quad a, b \in \mathbb{Q}.$$

- Ciało czteroelementowe wprowadzone w uzupełnieniu do wykładu pierwszego zawiera ciało $\mathbb{Z}_2$ jako podciało. Co więcej, ciało to jest w istocie postaci $\mathbb{Z}_2(\zeta)$, gdzie ζ jest pierwiastkiem wielomianu $x^2 + x + 1 \in \mathbb{Z}_2[x]$.

- Zupełnie innym przykładem jest ciało $\mathbb{Q}$ traktowane jako podciało ciała $\mathbb{R}$. Można liczby rzeczywiste traktować jako wektory, a liczby wymierne jako skalary. Nie jest jednak możliwe wskazanie takiego skończonego (ani nawet przeliczalnego – to zrozumieją Państwo na wstępie do matematyki) układu wektorów $r_1, r_2, \ldots, r_n$ takiego, by $\mathbb{R}$ było równe $\text{lin}(r_1, r_2, \ldots, r_n)$. Jak się okazuje wiąże się to z tym, że istnieją liczby rzeczywiste, które nie są pierwiastkami wielomianów rzeczywistych.

Definicja 4.13: Rozszerzenie algebraiczne

Niech $K \subset L$ będą ciałami. Powiemy, że element $a \in L$ jest algebraiczny nad ciałem K, jeśli istnieje wielomian $f \in K[x]$ taki, że $f(a) = 0$. Jeśli element $a \in L$ nie jest ALGEBRAICZNY nad ciałem K, wówczas element ten nazywamy PRZESTĘPNYM nad ciałem K.

Powiemy, że para ta jest ROZSZERZENIEM ALGEBRAICZNYM CIAŁ, jeśli każdy element ciała L jest pierwiastkiem pewnego wielomianu o współczynnikach ciała K. Rozszerzenie $K \subset L$ nazywamy PRZESTĘPNYM, jeśli nie jest ono algebraiczne.

Zobaczmy kilka przykładów.

- Liczby rzeczywiste $\sqrt{2}$, i, $\sqrt[3]{3}$, $\sqrt{1+\sqrt{2}}$ są algebraiczne nad $\mathbb{Q}$, bowiem są pierwiastkami wielomianów wymiernych $x^2 - 2, x^2 + 1, x^3 - 3, x^4 - 2x^2 - 1$.

- Liczby rzeczywiste π, e są przestępne nad $\mathbb{Q}$ (chociaż dowód nie jest łatwy).

- Liczba π jest algebraiczna nad $\mathbb{R}$ – jest pierwiastkiem wielomianu $x - \pi$.

Rozważania dotyczące rozszerzeń algebraicznych leżą u podstaw nowoczesnej teorii równań, a także teorii liczb. Wymagają one aparatu tzw. teorii pierścieni, którą poznacie Państwo na Algebrze I. Na Algebrze II poznacie Państwo następujący podstawowy rezultat.

Twierdzenie 4.1

Niech $K \subseteq L$ będzie rozszerzeniem ciał. Niech $\alpha \in L$ będzie elementem algebraicznym nad K i niech $f \in K[x]$ będzie wielomianem nierozkładalnym stopnia n takim, że $f(\alpha) = 0$. Wówczas ciało $K(\alpha)$ jest przestrzenią liniową nad ciałem K oraz

$$K(\alpha) = \text{lin}(1, \alpha, \alpha^2, \ldots, \alpha^{n-1}).$$

Innymi słowy każdy element ciała $K(\alpha)$ jest postaci:

$$\beta = b_0 + b_1\alpha + \ldots + b_{n-1}\alpha^{n-1}, \quad \text{gdzie } b_0, b_1, \ldots, b_{n-1} \in L.$$

W szczególności, jeśli L jest ciałem skończonym, to $|K(\alpha)| = |K|^n$, gdzie $|X|$ — moc zbioru X.

Z historycznego punktu widzenia ważną rolę wśród rozszerzeń grają tzw. ROZSZERZENIA KWADRATOWE, a więc takie rozszerzenia $K \subseteq K(\alpha)$, że $\alpha \notin K$ jest rozwiązaniem pewnego równania wielomianowego stopnia 2 o współczynnikach w ciele K. Podstawowym przykładem są tu ciała $\mathbb{Q}(\sqrt{p})$, gdzie p jest liczbą pierwszą. Liczba $\sqrt{p}$ jest pierwiastkiem wielomianu $x^2 - p \in \mathbb{Q}[x]$.

Rozszerzenia kwadratowe wiążą się ze starożytnym zagadnieniem tzw. liczb konstruowalnych (nad ciałem $\mathbb{Q}$), czyli takich długości odcinków, które można skontruować za pomocą cyrkla i linijki, mając do dyspozycji odcinek długości 1 (czyli też wszystkie odcinki długości $n \in \mathbb{N}$). Zagadnienie to pytało między innymi czy można za pomocą cyrkla i linijki[1]:

- dokonać trysekcji dowolnego kąta, a więc np. czy można skonstruować kąt o mierze $20°$,
- skonstruować odcinek o tej własności, że sześcian, którego krawędzią jest ten odcinek ma objętość 2,
- skonstruować siedmiokąt foremny?

Jak się okazuje, opisane problemy dotyczą liczb algebraicznych (nad ciałem $\mathbb{Q}$). Liczba $\cos 20°$ jest, jak się okazuje, pierwiastkiem wielomianu $4x^3 - 3x - \frac{1}{2}$. Liczba $\sqrt[3]{2}$ jest pierwiastkiem wielomianu $x^3 - 2$. Liczba $\cos\frac{2\pi}{7}$ jest, jak się okazuje, pierwiastkiem wielomianu

$$64x^7 - 112x^5 + 56x^3 - 7x - 1.$$

Nie jest to zupełnie elementarny wynik, ale żadna z powyższych liczb nie jest konstruowalna. Co więcej, zachodzą następujące twierdzenie, udowodnione w wieku XIX-tym.

Twierdzenie 4.2

Liczby konstruowalne nad $\mathbb{Q}$ tworzą podciało ciała liczb rzeczywistych. Liczba rzeczywista x jest konstruowalna nad $\mathbb{Q}$ wtedy i tylko wtedy, gdy istnieje ciąg rozszerzeń kwadratowych:

$$\mathbb{Q} \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq K_n$$

takich, że $x \in K_n$.

Mówiąc nieco nieprecyzyjnie, każda liczba niewymierna jest „iterowaną niewymiernością kwadratową". Na czym to polega? Tak, jak rozważamy np. „niewymierność kwadratową" postaci $1 + \sqrt{2}$ należącą do ciała $\mathbb{Q}(\sqrt{2})$, tak można rozważać element postaci $a + b\sqrt{3}$, gdzie $a, b \in \mathbb{Q}(\sqrt{2})$, na przykład element postaci:

$$(2 + 3\sqrt{2}) + (3 - \sqrt{2}) \cdot \sqrt{3}.$$

Liczba ta nie jest pierwiastkiem żadnego równania kwadratowego stopnia 2 o współczynnikach wymiernych, ale jest pierwiastkiem wielomianu stopnia 2o współczynnikach w $\mathbb{Q}(\sqrt{3})$ postaci:

$$x^2 - (4 + 6\sqrt{2})x + (-11 + 6\sqrt{2}),$$

a więc można powiedzieć, że przywołana liczba jest „dwukrotnie iterowaną" niewymiernością kwadratową i jest konstruowalna. Odpowiednim ciągiem rozszerzeń kwadratowych jest tu:

$$\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

Więcej rachunków i przykładów, między innymi wyjaśnienia klasycznych problemów konstruowalności zarysowanych wyżej, znajdą Państwo w podręczniku prof. Guzickiego, natomiast warto wspomnieć, że zarysowana tu tematyka jest fragmentem tzw. teorii Galois związanej z zagadnieniem rozwiązywalności równań wielomianowych stopnia $n > 4$ przez tak zwane pierwiastniki. Tematy te poznają Państwo na Algebrze II (zachęcam do wyboru tego przedmiotu). Do rozważania powyższych tematów potrzebne są zarówno elementy teorii ciał i przestrzeni liniowych, jak również elementy teorii grup i pierścieni, które poznają Państwo na II roku studiów.

[1] Więcej o tym zagadnieniu przeczytać można w rozdziale 13. podręcznika prof. Wojciecha Guzickiego *Geometria analityczna. Rozszerzony program matematyki w liceum*, wyd. Omega 2022.

4.5 Trivia. Kody samokorekcyjne.

Nie sposób opisać wszystkich zastosowań przestrzeni liniowych. W naszym wykładzie zajmować się będziemy w dużej mierze przestrzenią współrzędnych K^n, jej podprzestrzeniami, rozmaitymi opisami tych podprzestrzeni itd. Warto przekonać się od razu, że te podstawowe i bardzo elementarne przestrzenie mają ważne zastosowania. Opowieść przedstawiona poniżej mówi o przykładach tzw. kodów liniowych.

W 1948 roku Claude Shannon, amerykański inżynier i matematyk, wydał artykuł „A Mathematical Theory of Communitacion", który uważany jest za początek tzw. teorii informacji i teorii kodowania. Podstawowym celem jest efektywne i wiarygodne przesyłanie komunikatów w niekooperacyjnym (być może wrogim) środowisku. Aby być **efektywne** – komunikaty nie mogą wymagać nadawania przez zbyt długi czas lub zbyt duży koszt. Aby transmisja była **wiarygodna** potrzebne jest by otrzymywany sygnał przypominał ten wyemitowany, przynajmniej w ramach pewnej z góry określonej tolerancji. Wysiłki matematyków poszły w dwóch kierunkach. Shannon, ojciec teorii informacji, studiował osiągalne ograniczenia komunikacyjne głównie metodami analitycznymi i probabilistycznymi. Jego kolega – Richard Hamming, pracował nad poprawianiem kodów pierwszych komputerów i stosował głównie metody algebraiczne.

Informacja nadana ze źródła trafia do „przewodu", „przestrzeni" „kanału", którym podróżuje do odbiorcy. Nasz model komunikacji oparty jest o założenie, że informacja poddana jest zgodnie z naszą wolą pewnej strukturze u źródła oraz pewnej metodzie odczytu u odbiorcy, ale nie mamy żadnej kontroli nad przestrzenią pomiędzy nadawcą, a odbiorcą. W ten sposób wiadomość ulec może zniekształceniu. Prostym przykładem jest rozmowa w bardzo głośnej kawiarni, pisanie książki, która ma być odczytana lata później. Jest też wiele sposobów radzenia sobie z możliwymi zaburzeniami przekazu. Osobę, której nie dosłyszałem mogę poprosić o powtórzenie, a w przypadku znalezienia zniszczonego manuskryptu mogę próbować poszukiwać innej jego kopii. Tu jednak zaburzone są: efektywność (*Ile razy mam powtarzać*?!) i wiarygodność (*może nie ma innego manuskryptu, a może obydwa są fałszywe*?).

Zakłady Bell Telephone Laboratories w latach 50-tych XX wieku

Shannon i Hamming, a także wielu innych ojców teorii komunikacji, pracowali dla Bell Telephone Laboratories. Byli szczególnie zainteresowani radzeniem sobie z błędami, które powstają gdy wiadomość podróżuje kablem telefonicznym i zostanie zniekształcona przez uderzenie pioruna lub przez nałożenie się na siebie dwóch rozmów. Komunikacja w przestrzeni kosmicznej zaburzana jest przez atmosferę ziemską i aktywność słoneczną. Podczas misji Galileo, gdy padła jedna z anten sondy, naukowcy przeprogramowali komputer pokładowy sondy tak, by w sposób bardziej intensywny przetwarzał kod wysyłany na Ziemię i w ten sposób byli w stanie odzyskać część pierwotnej efektywności przekazu wiadomości. Dyski twarde naszych komputerów wyposażone są w CRC, czyli *Cyclic Redundancy Check*, z uwagi na konieczność wykrywania zaburzeń w przechowywaniu danych wystawionych na działanie promieni gamma czy interferencji magnetycznej. Gdy Phillips wprowadził technologię płyt CD reklamował ją jako niewrażliwą na wiele typów zniszczenia – nawet z porysowanej (nieznacznie) płyty jesteśmy (byliśmy?) w stanie odczytywać informacje. Jest to zasługa teorii kodowania. Można podać wiele więcej przykładów.

Informację można zapisać na wiele sposobów. Używamy w tym celu najczęściej słów zbudowanych z liter określonego alfabetu. W informatyce najczęściej są to bity, a więc ciągi zer i jedynek. **Kodowanie wiadomości polega na dodaniu do niej pewnego dodatkowego zestawu bitów** służącego do jej odczytania w sytuacji, gdy wiemy, że wystąpić może błąd. Można tego dokonywać na wiele sposobów.

Załóżmy, że chcecie Państwo przesłać Komuś wiadomość złożoną z trzech liter ze zbioru $\{0,1\}$ postaci $v = abc$. Między emiterem a odbiornikiem wiadomość może ulec zniekształceniu i dojdzie do Kogoś niewłaściwe słowo. Czy ów Ktoś zdoła wykryć taki błąd i odczytać poprawną wiadomość, jeśli wiemy na przykład, że błąd zwykle nie dotyczy więcej niż jednej litery?

Do opisu rozwiązania zastosujemy algebrę liniową. W tym celu zakłada się, że zakodowana wiadomość, którą przesyłamy, jest podprzestrzenią przestrzeni liniowej. Kodem liniowym długości n nad ciałem F nazywamy podprzestrzeń przestrzeni F^n. Zakodowane słowa to wektory.

Najpierw naiwne rozwiązanie problemu. Dla każdego 0 w planowanej wiadomości, wysyłamy dwa zera. Podobnie dla jedynek. A zatem jeśli oryginalna wiadomość miała na przykład postać [010], to zakodujemy ją jako [00 : 11 : 00]. A zatem nasz kod to element przestrzeni $(\mathbb{Z}_2)^6$. Czy możemy traktować go jako podprzestrzeń? Zauważmy, że $\mathrm{lin}((0,0,1,1,0,0))$ to po prostu $\{(0,0,0,0,0,0),(0,0,1,1,0,0)\}$, a więc niezerowe elementy tej podprzestrzeni tworzą zakodowaną wiadomość. Odbiorca jest w stanie stwierdzić czy jest ona poprawna, a więc czy wiadomość przyszła z jednym błędem (nie rozważamy tutaj dla uproszczenia innych sytuacji). Dla przykładu: jeśli wiadomość jest postaci [00 : 11 : 10] odbiorca wie, że jest błąd w trzecim segmencie wiadomości. Może zatem wydedukować, że oryginalna wiadomość miała postaci [010] lub [011]. Naiwne podejście ma dwie wady: po pierwsze przesyłamy dwa razy więcej danych, niż potrzeba, a po drugie – odbiorca nie ma dość informacji, by poprawić błąd, który wykrył.

Naiwne rozwiązanie problemu niemożności naprawienia (pojedynczego) błędu w transmisji jest proste: wysyłać trzy razy więcej danych. A więc na przykład oryginalną wiadomość postaci [010] przesłać możemy jako [000 : 111 : 000]. Jeśli odbiorca otrzyma, powiedzmy, wiadomość postaci [000 : 111 : 010] to wie już nie tylko, że błąd wystąpił w trzecim segmencie ale też, że w oryginalnej wiadomości ten segment miał postać 000. Widzimy jednak, że nie jest to efektywne przesyłanie danych. Oto inna propozycja, pochodząca od Hamminga.

Jeśli chcemy wysłać wiadomość postaci $[c_1 : c_2 : c_3]$, gdzie $c_1, c_2, c_3 \in \{0,1\}$, to wysyłamy ciąg złożony z pięciu znaków postaci: $[c_1 : c_2 : c_3 : c_1 + c_2 : c_2 + c_3]$, przy czym operacje dodawania wykonujemy nad ciałem $\mathbb{Z}_2$, czyli $c_1 + c_2$ jest równe 0 lub 1 w zależności od składników c_1, c_2. Okazuje się, że w tym kodowaniu jesteśmy w stanie wykryć nawet dwa błędy, a jeśli jest tylko jeden – to możemy go naprawić. Zobaczmy przykłady.

Wysyłamy [100], a wiec po zakodowaniu dostajemy słowo [10010]. Załóżmy, że wystąpi dokładnie jeden błąd przy transmisji i otrzymamy jedną z wiadomości: [00010], [01010], [10110], [10000], [10011]. Czy Czytelnik widzi, że w każdym wypadku możemy nie tylko wykryć błąd, ale i go naprawić? W pierwszym przypadku [00010] nie spełnia na czwartej współrzędnej warunku $c_1 + c_2 = 1$, ale spełnia na piątej warunek $c_2 + c_3 = 0$. A zatem skoro jest dokładnie jeden błąd, to c_2, c_3 są przesłane dobrze, a błędny jest przekaz c_1. Oczywiście umiemy też wykryć wiadomość poprawnie odebraną.

Czy wykrywanie pojedynczego błędu w ogóle może kogoś interesować? Nie tylko może, ale jest powszechne. Nie ma dwóch numerów kont, które różniłyby się tylko jedną lub dwiema cyframi. Jeśli wysyłając przelew pomylimy się o jedną lub dwie cyfry w numerze konta, to przelew zostanie odrzucony. Kod Hamminga stosuje się dla wiadomości dowolnej długości. Do zakodowania słowa długości n potrzeba $2n - 1$ znaków (oczywiście chodzi o słowo zerojedynkowe).

Być może Czytelnik nie dostrzega jeszcze żadnej wielkiej „matematyki" w tej opowieści, ale zapewniam, że dzieje się tak tylko dlatego, że niemal zmuszam się do unikania wprowadzania jakiejkolwiek terminologii, a dzieje się tu bardzo dużo. Mówiąc o kodach wspomnielibyśmy zaraz o odległości Hamminga, problemie pakowania sfer, macierzach generujących, wielomianach kodujących słowa itd. Zainteresowanych odsyłam do bardzo ciekawych notatek J. Halla z teorii kodowania (polecam zwłaszcza wstępny rozdział – kolejne mogą być za trudne na razie – tylko na razie) dostępnych pod adresem:

https://users.math.msu.edu/users/jhall/classes/CODENOTES/CODING-NOTES.HTML

Kto by chciał poczytać (w języku polskim) więcej o kodach, szyfrach i ogólnie o teorii informacji, czy też przekonać się wszechstronnym występowaniu kodowania, np. w numerach PESEL, ISBN, IBAN, polecam tekst dr. Grzegorza Szkibiela „Wstęp do teorii informacji i kodowania", dostępny online.

4.6 Coda. O kształtowaniu się pojęcia wektora

Pojęcie wektora kształtowało się w nauce przez stulecia[2] i proces ten miał istotny wpływ na jej współczesny język. Nie chodzi jedynie o matematykę, ale też astronomię, fizykę, chemię, informatykę, ekonomię czy nauki techniczne. Słowo *wektor* pochodzi od łacińskiego *vectus*, znaczącego dosłownie „przewóz".

Historycznie rzecz biorąc intucje wektorowe związane były najpierw przede wszystkim z reprezentacją sił działających na obiekt za pomocą skierowanych odcinków oraz z obserwacją, że składanie tych sił spełnia tzw. prawo równoległoboku. Idee te wysłowił bezpośrednio już Arystoteles w czasach antycznych, w dziele *Questiones Mechanicae*. Dzieło to znali autorzy renesansowi, łącznie z Galileuszem, nie zawsze doceniając znaczenie samej reguły, a nawet nie uwzględniając jej wcale w swoich badaniach[3]. Dyskusję w kierunku wysłowienia tej reguły rozpoczną dopiero siedemnastowieczni autorzy tacy jak Fermat, Hobbes czy Mersenne, głównie w oparciu o próbę zrozumienia praw optyki (odbicia i załamania) Kartezjusza. Należy jednak pamiętać, że uczeni ci nie określali pojęcia wektora. Formułowali jedynie pewne obserwacje w języku geometrii. Podejście w zasadzie „istotowo wektorowe" stosuje dopiero Newton w *Philosophiae naturalis principia mathematica* (1687), gdzie prawo równoległoboku jest sformułowane wprost, wciąż jednak jednak bez użycia wektorów a jedynie w oparciu o geometrię Euklidesa[4].

Samo pojęcie wektora stosowane było najpierw w astronomii, w kontekście, w jakim dziś rozumiemy pojęcie *wektora wodzącego* (mówiąc mało precyzyjnie chodzi o wektor o ustalonym początku i końcu poruszającym się według pewnych zasad, na przykład po okręgu, elipsie itd.) i pojawiło się po raz pierwszy w 1704 roku. Pojęcie to (rayon vecteur) stosują również[5] Laplace w swoim *Traktacie o mechanice niebieskiej* (1798) oraz Andre Ampère w *Traktacie o matematycznej teorii zjawisk elektrodynamicznych* z roku 1826. Wcześniej pojęcia tego używał też de la Lande w słynnej *Wielkiej encyklopedii francuskiej* (1776).

W geometrii pojęcie wektora pojawiło się wraz z geometryczną interpratacją pojęcia liczby zespolonej, proponowaną już przez Wessela (1797) i Arganda (1806). Podejście to poznaliśmy w ujęciu zaproponowanym przez Hamiltona, jest jednak pewne, że już Gauss posługiwał się nim swobodnie na początku XIX-tego wieku. W starych podręcznikach (np. autorstwa Webera z 1925 roku) znajdziemy zresztą informację, że reprezentacja wektorowa liczb zespolonych pochodzi od Gaussa. Również w geometrii używano pojęcia *rayon vecteur*: robił to zarówno Möbius (1827) w swoim rachunku barycentrycznym, jak i Cauchy, we wstępie do ważnego traktatu *Leçons sur les applications de calcul infinitésimal* (1826).

Do około 1830 roku liczby zespolone były w zasadzie reprezentowane jako wektory, choć nazewnictwo to wprowadził Hamilton i w to w kontekście swojego największego odkrycia — kwaternionów. Jednym z istotnych problemów matematyki początku XIX-tego stulecia było przeniesienie teorii liczb zespolonych w trzy wymiary, tak by na trójkach postaci $a+bi+cj$, gdzie i, j są pierwiastkami z -1, okreslić mnożenie, mające sens geometryczny i porządne własności algebraiczne (łączność, przemienność, rozdzielność itd.).

W 1837 roku Hamilton publikuje długą i niezwykle ważną pracę interpretującą liczby zespolone jako uporządkowane pary liczb rzeczywistych, wprowadzając znaną nam zasadę mnożenia owych par. Jednocześnie rozpoczyna poszukiwania „teorii trójek", wspomnianej już wyżej. Jak się okazało jest to głęboki problem, który doprowadza w 1843 roku do odkrycia kwaternionów, czyli liczb zapisywanych w postaci

$$a + xi + yj + zk,$$

gdzie $a, x, y, z \in \mathbb{R}$ oraz gdzie spełnione są następujące reguły:

$$ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j, \quad i^2 = j^2 = k^2 = -1.$$

Co ciekawe, dodawanie (po współrzędnych) i mnożenie kwaternionów są łączne (Hamilton pokazał to w roku 1844, po raz pierwszy używając tego terminu) oraz rozdzielne, dzielenie przez niezerowy element jest zawsze możliwe, dodawanie jest przemienne. Mnożenie kwaternioów jest jednak nieprzemienne.

---

[2] Skrótowe opisy historii pojęć matematycznych znaleźć można w *Earliest Uses of Some Words of Mathematics* w ramach serwisu MacTutor History of Mathematics Archive: `https://mathshistory.st-andrews.ac.uk/Miller/mathword/`.

[3] David Marshall Miller, *The Parallelogram Rule from Pseudo-Aristotle to Newton*, Archive for History of Exact Sciences 71 (2017), 157-191, `https://www.jstor.org/stable/24913240`.

[4] Sformułowanie w języku angielskim: `https://archive.org/details/100878576/page/84/mode/2up?view=theater`

[5] Gregory H. Moore, *The axiomatization of linear algebra*, Historia Mathematica 22 (1995), 262-303, `https://www.sciencedirect.com/science/article/pii/S0315086085710257`.

Odkrycie kwaternionów miało znaczenie nie tylko naukowe, ale też filozoficzne. Polowa XIX-tego wieku to czasy, gdy Bolyai i Łobaczewski odkrywali pierwsze modele geometrii nieeuklidesowych. Hamilton pokazał, że możliwe jest stworzenie systemu algebraicznego, który łamie jedną z podstawowych reguł (przemienność mnożenia) mając sens także w kontekście naturalnych operacji geometrycznych. Takich systemów powstało więcej (najpierw rozważano tzw. liczby hiperzespolone, a później tzw. algebry).

W 1846 roku Hamilton publikuje pracę, w której wprowadza pojęcie *skalara* i *wektora*, mając na myśli „rzeczywistą" i „urojoną" część swoich kwaternionów. To znaczy: częścią skalarną kwaternionu $a + bi + cj + dk$ jest a, zaś częścią wektorową: $bi + cj + dk$. Jak się okazało, prawa mnożenia kwaternionów o części skalarnej zero miały ważne znaczenie w geometrii. W fizyce, fundamentalne znaczenie miało odkrycie przez Hamiltona tzw. operatora nabla, uczytenilnione jeszcze przez Taita do postaci

$$\nabla = i\frac{d}{dx} + j\frac{d}{dy} + k\frac{d}{dz}.$$

W istocie, pierwotne sformułowanie słynnych równań elektrodynamiki klasycznej Maxwella z 1865 roku miało postać kwaternionową[6]. To równania Maxwella i próba ich uczytelnienia doprowadziły do nowoczesnego ujęcia pojęcia wektora i stało się to w pierwszej kolejności za sprawą fizyków.

W 1881 roku J.W. Gibbs opublikował pierwszą połowę *Elementów analizy wektorowej*, gdzie wektor opisywany jest zarówno za pomocą pojedynczego symbolu, jak i za pomocą wspólrzędnych. Gibbs sformułował, niezależnie od Grassmanna, o którym powiemy niżej, podstawowe działania na wektorach, a także pojęcia iloczynu skalarnego i iloczynu wektorowego. Prace Gibbsa (oraz Taita) rozwinął następnie Heaviside, dodając do niej wyniki Grassmanna. Widzimy zatem, że istotne motywacje do badania wektorów i operacji na nich nie pochodziły w XIX-tym wieku z matematyki, ale właśnie z fizyki. Podobnie było na początku wieku XX-tego, gdy sformułowana została teoria względności Einsteina.

Wracając do matematyki, można powiedzieć, że pod koniec wieku XIX-tego i na początku XX-tego pojęcie wektora traktowano wciąż geometrycznie – jako odcinek skierowany $\overrightarrow{AB}$ lub jako formalną różnicę punktów $B - A$. Fizycy traktowali wektor jako obiekt (np. moment lub siłę) mający kierunek i długość. Zakładano też, że wektory mają nie więcej niż trzy współrzędne. Istniał już jednak od ponad 50 lat grunt pod ogólne podejście, pochodzący od mało znanego ówczesnym matematykom uczonego — Grassmanna.

Za twórcę pojęcia przestrzeni liniowej[7], uważa się właśnie Hermanna Gunthera Grassmanna, urodzonego w Szczecinie w roku 1809. Był jednym z dwanaściorga dzieci. Sam wziął ślub po czterdziestym roku życia, mając ich siedmioro. Trzy lata życia Grassmann spędził w Berlinie, studiując teologię i filologię. Nie miał żadnego wykształcenia matematycznego i nigdy nie zajmował pozycji na uniwersytecie (chociaż starał się o nie wielokrotnie). Życie spędził jako nauczyciel gimnazjalny. Zmarł w roku 1877, nigdy nie otrzymawszy uznania jako twórca jakiejkolwiek teorii matematycznej. Jego prace matematycy odkryli później.

Rozważania Grassmanna zawierały idee przekraczające epokę, w której żył. Zanim przejdziemy do algebry liniowej warto wspomnieć, że w dziedzinie arytmetyki, już w 1861 roku Grassmann zdefiniował operacje arytmetyczne w zbiorze liczb naturalnych za pomocą pojęcia indukcji i dowiódł takie ich własności, jak przemienność, łączność, rozdzielność. Zdołał więc przewidzieć założenia teorii Peano czy Dedekinda opublikowanych niemal 30 lat później. Autorzy ci są wyraźnie zainspirowani wpływem Grassmanna, choć nie nie pomijają trudności w lekturze filozoficznego (miejscami nieprecyzyjnego) języka jego prac.

Najważniejszym dziełem Grassmanna była *Ausdehnungslehre* z 1844 roku, czyli Teoria Rozszerzeń, w zasadzie niezauważona aż do czasu publikacji dzieł zebranych Grassmanna pomiędzy rokiem 1894 oraz 1911, i to mimo tego, że autor przesyłał swoje prace między innymi do Möbiusa, Gaussa, Kummera, Cauchy'ego. Z ostatnim łączył go zresztą wieloletni spór o pierwszeństwo wyników, nierozstrzygnięty przed Francuską Akademią Nauk. Grassmann zdefiniował pojęcie kombinacji liniowej, podprzestrzeni, liniowej niezależności, podprzestrzeni rozpinającej, wymiaru (w tym sumy i przecięcia podprzestrzeni) oraz rzutu na podprzestrzenie. Otrzymał również wzory na zamianę współrzędnych przy zmianie bazy w postaci iloczynu operacji elementarnych (fakty te będziemy stopniowo poznawać). Wprowadził również pojęcia, które dały początek iloczynowi zewnętrznemu oraz iloczynowi skalarnemu.

[6] *On the Notation of Maxwell's Field Equations*, `http://www.zpenergy.com/downloads/Orig_maxwell_equations.pdf`.

[7] Tekst na podstawie artykułów D. Fearnley-Sander, *Hermann Grassmann and the Creation of Linear Algebra*, The American Mathematical Monthly , Vol. 86, No. 10 (1979), str. 809-817 oraz W. Więsław, *Drogi i manowce początków algebry*, Szkoła Matematyki Poglądowej, https://smp. uph.edu.pl/msn/15/16-26.pdf.

Prace Grassmanna nie zostały na początku zauważone. Stosunkowo niewielką zmianę wniosły prace Peano, który w 1887 roku zaczął rozważać n-tki (wektory o n współrzędnych) wraz z operacjami dodawania i mnożenia przez skalar. To właśnie Peano, inspirowany pracami Grassmanna, wprowadził pojęcie systemu liniowego (obecnie przestrzeni liniowej) za pomocą aksjomatów (w swoim trzecim podejściu do tego tematu) w 1898 roku. Należy jednak podkreślić, że nawet Peano pisał o wektorach jedynie w kontekście geometrycznym. Robił to mimo tego, że to właśnie on był autorem dowodu istnienia rozwiązania układu n liniowych równań różniczkowych pierwszego rzędu o n zmiennych. Oczywiście Peano zajmował się przestrzeniami liniowymi nad $\mathbb{R}$, nie znając ogólnej teorii ciał. Jednym z głównych nowych pomysłów Peano było zrozumienie, że wielomiany jednej zmiennej rzeczywistej, a także wielomiany ograniczonego stopnia, tworzą przestrzenie liniowe. Właściwe ujęcie prac Grassmana znajdą dopiero wielcy geometrzy różniczkowie początku XX wieku, przede wszystkim Henri Cartan oraz Henri Poincare.

Alternatywne podejście do aksjomatyzacji pojęcia wektora zaproponował Gaston Darboux. W 1875 roku opublikował pracę analizującą różne dowody prawa składania sił statycznych (np. prawo równoległoboku), rozpoczynając od dowodu Daniela Bernoulliego z 1726 roku. Celem Darboux było uzyskanie uzasadnień tych prac zawartych wewnątrz geometrii i ustalenia jakie założenia wymagane są do tego, by prawa te zachodziły. Zaproponował cztery takie aksjomaty (których w tym miejscu nie wysławiamy). Prace Darboux podjęli młodzi matematycy Schimmack i Hamel, którzy badali między innymi formalną zależność tych aksjomatów. Jak się okazało, niezależność czwartego aksjomatu Darboux od trzech wcześniejszych wymagała znalezienia nieciągłej funkcji rzeczywistej f, spełniającej dla dowolnych $x, y \in \mathbb{R}$ równanie funkcyjne Cauchy'ego $f(x+y) = f(x) + f(y)$. Hamel znalazł przykład takiej funkcji i jego praca doktorstwa opublikowana została przez samego Hilberta w *Mathematische Annalen* w 1905 roku.

Odkrycia Hamela miały fundamentalne znaczenie dla teorii zbiorów, bowiem wymagały nowego wówczas rezultatu Zermelo mówiącego, że każdy zbiór można dobrze uporządkować. Wyniki Hamela, o których wspomnimy w komentarzach do kolejnych wykładów, wymagały skonstruowania bazy przestrzeni liczb rzeczywistych traktowanych jako przestrzeń liniowa nad ciałem liczb wymiernych. Istnienie takiej bazy wymaga aksjomatu wyboru, o czym powiemy w dodatku do kolejnego wykładu. To, co jest na ten moment istotne w podejściu Hamela, to zauważenie, że same liczby rzeczywiste traktować można jak wektory nad ciałem skalarów $\mathbb{Q}$. Co więcej, problem stwierdzania czy liczba rzeczywista jest skończoną kombinacją liniową (o współczynnikach w $\mathbb{Q}$) innych liczb rzeczywistych ma głębokie zastosowania.

Pod koniec Pierwszej Wojny Światowej sytuacja przestrzeni liniowych była następująca — ogólne pojęcie przestrzeni liniowej nad $\mathbb{R}$ znane było, ale nie powszechnie, we Włoszech, wśród spadkobierców Peano. Jako jeszcze mniej znane pojęcie, przestrzenie te znano we Francji i Niemczech za sprawą prac Darboux, a potem Schimmacka i Hamela. Pojęcie to doprowadziło do powstania „baz Hamela" badanych intensywnie w kontekście analizy i teorii zbiorów (np. przez Wacława Sierpińskiego). Kluczowym momentem dla przyjęcia przestrzeni liniowych jako pełnoprawnych obiektów matematycznych były prace Hahna, Banacha i Wienera, związane z tzw. unormowanymi przestrzeniami liniowymi. Na nasz użytek powiedzmy, że chodzi o takie przestrzenie liniowe, gdzie można za pomocą pewnej funkcji (zwanej normą) wprowadzić odległość. Pojęcie normy wektora poznamy w drugim semestrze (w ograniczonym kontekście).

W 1922 roku wiedeński matematyki Hans Hann sformułował pojęcie unormowanej przestrzeni liniowej i przedstawił 21 przykładów przestrzeni tego typu. Wszystkie one były przestrzeniami funkcji, co miało przełomowe znaczenie. Przykładem były badane już wcześniej przez Schura przestrzenie ciągów nieskończonych (i ich przekształceń). Hahn zajmował się też układami równań liniowych w tych przestrzeniach.

Niezależnie od Hahna, pojęcie przestrzeni unormowanej wprowadził w 1922 roku Stefan Banach[8]. Prace Banacha były o tyle przełomowe, że wprowadzały na dobre metodę aksjomatyczną do analizy, Przestrzenie Banacha określone były najpierw za pomocą 13 aksjomatów, określających w istocie rzeczywistą przestrzeń liniową. Banach cytował, w ramach przykładów, Grassmanna, prace Hamiltona, teorie wektorów Peano itd. Druga grupa aksjomatów dotyczyła normy, a trzecia – pojęcia zupełności. Prace Banacha wywołały pozytywne reakcje wielkich matematyków, między innymi Norberta Wienera i Maurice'a Frecheta, a pojęcie przestrzeni liniowej trafiło na Międzynarodowy Kongres Matematyków.

W międzyczasie do gry wkroczyło największe nazwisko matematyki początku XX-tego wieku — Dawida Hilberta. W roku 1904 Hilbert opublikował pracę dotyczącą liniowych równań całkowych, gdzie badano między innymi ciągi liczb rzeczywistych, których (nieskończona) suma kwadratów była skończona. Po-

[8]Chodzi o tzw. zupełne unormowane przestrzenie liniowe, zwane przestrzeniami Banacha.

dejście geometryczne do tych badań zaproponowali między innymi Schmidt (1908) oraz Riesz (1913), badający między innymi układy równań liniowych o nieskończenie wielu zmiennych. Badając strukturę rozwiązań tych układów, Riesz wprowadził pojęcie przestrzeni Hilberta. W 1927 roku pojęcie to zostało sformułowane aksjomatycznie przez von Neumanna, w celu zbudowania matematycznych mechaniki kwantowej Heisenberga i Schrödingera. Już wcześniej, pojęcie przestrzeni liniowej aksjomatyzował dla potrzeb zbudowania matematycznej teorii względności Hermann Weyl (1918). Pojęcie wektora i przestrzeni liniowej było zatem dobrze umotywowane przez teorie fizyczne. Czy istniało jakieś algebraiczne źródło?

Algebraiczne źródło pojęcia przestrzeni liniowej wywodzi się z prac grupy niemieckich matematków, Dirichleta, Kummera, Kroneckera, Dedekinda i Webera, związanych z próbą dowodu Wielkiego Twierdzenia Fermata Dedekind wprowadził pojęcie ideału, czyli podzbioru A w zbiorze B (np. w $\mathbb{C}$) zamkniętego na dodawanie, odejmowanie i mnożenie przez element z B. Dedekind sformułował również pojęcie modułu, mając na myśli podzbiór M zbioru $\mathbb{C}$ zamknięty na dodawanie i odejmowanie. Dedekind wprowadził notację $a \equiv b \bmod M$ mając na myśli $a - b \in M$. Pojęcie to w latach 70-tych XIX-tego wieku było bardzo ogólne — obejmowało bowiem ideały Dedekinda, a naśladowało przy tym teorię kongruencji Gaussa, uogólniając jednak relację przystawania modulo z pojedynczej liczby całkowitej do całego zbioru.

Dedekind zorientował się, że istnieje związek pomiędzy badanymi przez niego liczbami algebraicznymi Ω, np. postaci $a + b\sqrt{2} + c\sqrt{-3} + d\sqrt{5}$, gdzie $a, b, c, d \in \mathbb{Q}$, a pojęciem „bazy" i „liniowej niezależności". Z obecnej perspektywy można rozumieć, że Dedekind umiał pokazać, że Ω jest przestrzenią skończonego wymiaru nad $\mathbb{Q}$ mimo, że pojęcie to jeszcze nie funkcjonowało. Prace Dedekinda dotyczyły też sytuacji, gdy współczynniki były całkowite, a nie wymierne, co wyprowadza nas z algebry liniowej w kierunku tzw. teorii pierścieni. Dedekind współpracował ściśle z Heinrichem Weberem, z którym rozszerzył w 1882 roku pojęcie modułu do kontekstu funkcyjnego, definiując obiekt nazywany dziś modułem nad pierścieniem wielomianów $\mathbb{C}[z]$ i badając takie moduły, mające skończoną „bazę". Dekadę później, w 1892 roku, Weber zunifikował rozmaite pojęcia ciała (algebraiczne ciało liczbowe, algebraiczne ciało funkcyjne, ciało skończone) i sformułował abstrakcyjną definicją, znaną do dziś. Prace te podjął w 1910 roku Ernst Steinitz. Pojęcie modułu zwróciło uwagę wielkich algebraików, m.in. Hilberta i Noether, którzy użyli go do zbudowania podstaw teorii pierścieni oraz ich ideałów. Po 1945 roku rozważania te nabrały nowego kontekstu w świetle teorii kategorii.

Podejście algebraiczne zostało dojrzale ukształtowane w przełomowym podręczniku *Moderne Algebra* van der Waerdena w latach 1930-1931. Po kilku latach książka ta trafiła z Niemiec do Ameryki, a nowoczesne podejście do algebry, uwzględniające przestrzenie liniowe, trafiło do najsłynniejszego przedwojennego podręcznika algebry — *Przeglądu algebry wspołczesnej* Birkhoffa i Mac Lane'a (1941). W nauczaniu akademickim spopularyzował je ważny podręcznik Mirsky'ego z 1955 roku[9] W Polsce pojęcie przestrzeni liniowej upowszechnił w nauczaniu akademickim Profesor Andrzej Mostowski z Uniwersytetu Warszawskiego. Kolejne wydania podręcznika, zwłaszcza pisane wspólne z Marcelim Starkiem, były w zasadzie podstawą wykładu akademickiego przez niemal pół wieku. Po czasie dołączyła do nich wspaniała *Algebra liniowa z geometrią* Profesora Andrzeja Białynickiego-Biruli, która była podstawą do opracowania obecnego programu nauczania tego przedmiotu na naszym Wydziale. Warto tu przywołać fragment recenzji tej ostatniej pozycji z Wiadomości Matematycznych (1976), autorstwa Profesora Narkiewicza

> *Nowy program studiów matematycznych zlikwidował wykładaną tradycyjnie na I roku geometrię analityczną, łącząc ten przedmiot z algebrą liniową. Recenzowana książka jest pierwszym podręcznikiem powstałego w ten sposób przedmiotu, dopasowanym ściśle do wymogów programowych. W istocie swej jest to podręcznik algebry liniowej w klasycznym ujęciu, z dodaniem elementów teorii przestrzeni afinicznych i przekształceń afinicznych. Czytelnik, przyzwyczajony do tradycyjnej geometrii analitycznej, nie znajdzie jej tu wcale. Jedynie jej ślad przewija się tu i ówdzie w zadaniach. Taki jest los tej archaicznej dyscypliny, zdaniem recenzenta, w pełni zasłużony.*

Podsumowując, widzimy jak skomplikowane są dzieje pojęć matematycznych. Powyższy tekst przedstawia i tak jedynie wierzchołek góry lodowej, zarzucając Czytelnika nazwiskami wielkich matematyków, których poznawanie zajmie większość studiów. Warto jednak rozumieć, że uporządkowana i sterylnie wręcz wyglądająca teoria ma korzenie dotykające niemal każdej dziedziny matematyki — i nie tylko matematyki, ale także fizyki czy astronomii. Historia wektorów jest znacznie starsza niż historia teorii, która je opisuje.

[9] `https://mathshistory.st-andrews.ac.uk/Extras/Mirsky_books/`.

Rozdział 5

Liniowo niezależne układy wektorów. Baza przestrzeni liniowej

5.1 Wykład piąty

Na ostatnim wykładzie wprowadziliśmy pojęcie przestrzeni liniowej V nad ciałem K, której elementy nazywamy wektorami, wraz z operacjami dodawania wektorów i mnożenia wektorów przez skalar z ciała K. Podstawowy problem, który chcemy dziś[1] rozstrzygnąć jest następujący. Dany jest skończony układ[2] wektorów $\alpha_1, \ldots, \alpha_n$ wektorów w V (mogą to być np. rozwiązania jednorodnego układu równań, dla $V = K^n$). Jaka jest geometryczna struktura podprzestrzeni $\lin(\alpha_1, \ldots, \alpha_n)$? Jest jasne, że poniższe dwie podprzestrzenie $\mathbb{R}^3$, choć rozpięte na układach dwóch wektorów, są „diametralnie różne":

$$W_1 = \lin((1,1,1),(2,2,2)), \quad W_2 = \lin((1,1,1),(1,2,1)).$$

Pierwszą z tych podprzestreni można przedstawić w postaci $W_1 = \lin((1,1,1))$. Drugiej nastomiast nie można przedstawić w postaci $\lin(\alpha)$, gdzie $\alpha \in \mathbb{R}^3$. To jest jasne, bo wektory $(1,1,1),(1,2,1)$ nie są proporcjonalne. Gdy liczba wektorów rozpinających podprzestrzeń wzrasta, analiza robi się bardziej skomplikowana i sam test proporcjonalności jest niewystarczający. Narzędziem właściwym dla rozstrzygnięcia tego problemu jest fundamentalne dla całej matematyki pojęcie liniowej niezależności układu wektorów.

Definicja 5.1: Liniowo zależny i liniowo niezależny układ wektorów (skończony)

Niech V będzie przestrzenią liniową nad ciałem K i niech 0_V będzie wektorem zerowym w V.

- Układ wektorów $\beta_1, \ldots, \beta_m$ przestrzeni V nad ciałem K nazwiemy LINIOWO ZALEŻNYM, jeśli istnieją elementy $a_1, \ldots, a_m$ ciała K, nie wszystkie równe 0, spełniające:

$$a_1\beta_1 + \ldots + a_m\beta_m = 0_V.$$

- Układ wektorów $\alpha_1, \ldots, \alpha_m$ przestrzeni V nazwiemy LINIOWO NIEZALEŻNYM, jeśli nie jest liniowo zależny. Równoważnie — układ ten jest liniowo niezależny, gdy dla dowolnych skalarów $a_1, \ldots, a_m \in K$ zachodzi implikacja

$$a_1\alpha_1 + \ldots + a_m\alpha_m = 0_V \quad \Longrightarrow \quad a_1 = \ldots = a_m = 0.$$

  Pusty układ wektorów uważamy za liniowo niezależny.

Przykład 1. Układ złożony z jednego niezerowego wektora α jest liniowo niezależny, bowiem z równości $a\alpha = 0_V$ wynika, że $a = 0$ lub $\alpha = 0_V$. Skoro $\alpha \neq 0_V$, to $a = 0$.

Przykład 2. Układ wektorów $\alpha_1, \ldots, \alpha_n$ zawierający wektor zerowy — powiedzmy α_n jest liniowo zależny, bo mamy $0 \cdot \alpha_1 + \ldots + 0 \cdot \alpha_{n-1} + 1 \cdot 0_V = 0_V$. Podobnie pokazujemy, że układ zawierający choćby dwa identyczne (czy też proporcjonalne) wektory jest liniowo zależny.

[1] Ostatnia aktualizacja: 30.12.2022 r.
[2] Używając słowa *układ* dopuszczamy możliwość, że pewne wektory wchodzące w jego skład powtarzają się.

Przykład 3. Układ $(1,0,0),(2,0,0),(3,0,0),(4,0,0),(5,0,0)$ jest liniowo zależny w $\mathbb{R}^3$, bo:

$$1(1,0,0)+1(2,0,0)+1(3,0,0)+1(4,0,0)+(-2)(5,0,0)=(0,0,0),$$

oraz

$$\operatorname{lin}((1,0,0),(2,0,0),(3,0,0),(4,0,0),(5,0,0))=\operatorname{lin}((1,0,0)).$$

Przykład 4. Układ

$$\alpha_1=\begin{bmatrix}1&1\\0&1\end{bmatrix},\quad \alpha_2=\begin{bmatrix}1&1\\1&0\end{bmatrix},\quad \alpha_3=\begin{bmatrix}0&0\\1&1\end{bmatrix}$$

jest liniowo niezależny w przestrzeni liniowej $V=M_{2\times 2}(\mathbb{R})$, bowiem dla dowolnych $a,b,c\in\mathbb{R}$:

$$a\alpha_1+b\alpha_2+c\alpha_3=0_V \quad\Longleftrightarrow\quad a\begin{bmatrix}1&1\\0&1\end{bmatrix}+b\begin{bmatrix}1&1\\1&0\end{bmatrix}+c\begin{bmatrix}0&0\\1&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix},$$

co oznacza, że

$$a\alpha_1+b\alpha_2+c\alpha_3=\begin{bmatrix}a+b&a+b\\b+c&a+c\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

A zatem mamy:

$$\begin{cases}a+b&=0\\b+c&=0\\a+c&=0\end{cases}. \qquad (\dagger)$$

W rezultacie dostajemy $a=-b=c=-c=0$. *Uwaga.* Ten sam układ macierzy $\alpha_1,\alpha_2,\alpha_3$ traktowanych jako elementy $M_{2\times 2}(\mathbb{Z}_2)$ jest liniowo zależny, bowiem układ $(\dagger)$ ma niezerowe rozwiązanie $(1,1,1)$ w $\mathbb{Z}_2^3$.

Poniższy przykład zostanie szczegółowo omówiony podczas ćwiczeń.

Obserwacja 5.1

Niech $0\neq A=[a_{ij}]\in M_{m\times n}(K)$ będzie w postaci schodkowej oraz niech $\alpha_1,\ldots,\alpha_r\in K^n$ – niezerowe wiersze macierzy A. Wówczas układ $\alpha_1,\ldots,\alpha_r$ jest liniowo niezależny.

Dowód to indukcja po liczbie niezerowych wierszy r macierzy A. Krok bazowy: układ złożony z jednego niezerowego wektora jest liniowo niezależny. Przejdźmy do kroku indukcyjnego. Rozważmy macierz A w postaci schodkowej o r niezerowych wierszach $\alpha_1,\ldots\alpha_r$. Jeśli dla pewnych $\lambda_1,\ldots,\lambda_r\in K$ mamy:

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\ldots+\lambda_r\alpha_r=(0,\ldots,0), \qquad (\diamondsuit)$$

to niech pierwszy niezerowy wyraz w wierszu α_1 stoi na k-tym miejscu.

$$\begin{bmatrix}0&\ldots&0&a_{1k}&\ldots&a_{1n}\\0&\ldots&0&0&\ldots&a_{2n}\\\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\0&\ldots&0&0&\ldots&a_{rn}\end{bmatrix}$$

Suma k–tych współrzędnych wektorów $\lambda_1\alpha_1,\ldots,\lambda_r\alpha_r$ równa jest k-tej współrzędnej wektora zerowego, czyli $\lambda_1a_{1k}+\lambda_2a_{2k}+\ldots+\lambda_ra_{rk}=0$. Jednak $a_{2k}=\ldots=a_{rk}=0$, bo A jest schodkowa. Co więcej, $a_{1k}\neq 0$. A zatem mamy $\lambda_1a_{1k}=0$, czyli $\lambda_1=0$. A zatem w równości $(\diamondsuit)$ dostajemy: $\lambda_2\alpha_2+\ldots+\lambda_r\alpha_r=(0,\ldots,0)$. Skoro $\alpha_2,\ldots,\alpha_r$ są kolejnymi wierszami macierzy schodkowej, to z założenia indukcyjnego wektory te tworzą układ liniowo niezależny, czyli mamy $\lambda_2=\lambda_3=\ldots=\lambda_r=0$. Pokazaliśmy, że $\lambda_1\alpha_1+\ldots+\lambda_r\alpha_r=(0,\ldots,0)$ implikuje $\lambda_1=\ldots=\lambda_r=0$.

Widzimy zatem, że ostatni przykład pozwala rozwiązać następujące zagadnienie w przestrzeni K^n: dana jest podprzestrzeń $W=\operatorname{lin}(\beta_1,\ldots,\beta_m)$ w K^n. Znajdź układ liniowo niezależny $\alpha_1,\ldots,\alpha_r$ taki, że $W=\operatorname{lin}(\alpha_1,\ldots,\alpha_r)$. Rozwiązanie jest takie: traktujemy wektory $\beta_1,\ldots,\beta_m$ jako wiersze macierzy $A\in M_{m\times n}(K)$ i doprowadzamy A do postaci schodkowej. Zgodnie z powyższą obserwacją niezerowe wiersze $\alpha_1,\ldots,\alpha_r$ macierzy A' są liniowo niezależne. Co więcej, na poprzednim wykładzie pokazaliśmy, że wiersze macierzy A' rozpinają tę samą podprzestrzeń K^n, co wiersze macierzy A. Widzimy więc, że problem jest rozwiązany, bo A' ma r niezerowych wierszy i $m-r$ wierszy zerowych, oraz:

$$\operatorname{lin}(\alpha_1,\ldots,\alpha_r)=\operatorname{lin}(\alpha_1,\ldots,\alpha_r,\underbrace{0,\ldots,0}_{m-r})=\operatorname{lin}(\beta_1,\ldots,\beta_m).$$

W dalszych rozważaniach wektor 0_V oznaczamy po prostu za pomocą symbolu 0.

Obserwacja 5.2

Niech V będzie przestrzenią liniową nad ciałem K i niech $\beta_1, \ldots, \beta_k \in V$. Następujące warunki są równoważne:

- układ $\beta_1, \ldots, \beta_k$ jest liniowo zależny,
- jeden z wektorów $\beta_1, \ldots, \beta_k$ jest kombinacją liniową pozostałych.

Intuicja jest następująca: liniowo zależny układ rozpinający jest „nadmiarowy" – można go „uszczuplić" do podukładu, który rozpina tę samą podprzestrzeń. Należy też zauważyć delikatność założenia: nie twierdzimy, że każdy wektor w układzie liniowo zależnym musi być kombinacją liniową pozostałych. Twierdzimy tylko, że w układzie liniowo zależnym istnieje taki wektor.

Przykład. Układ $(1,0,0), (2,0,0), (1,1,1)$ jest liniowo zależny w $\mathbb{R}^3$, bo

$$2(1,0,0) + (-1)(2,0,0) + 0(1,1,1) = (0,0,0)$$

ale

- $(1,1,1)$ **nie jest kombinacją liniową** $(1,0,0), (2,0,0)$,
- $(1,0,0) = \frac{1}{2}(2,0,0) + 0(1,1,1)$.
- $(2,0,0) = 2(1,0,0) + 0(1,1,1)$.

Dowód. Przypuśćmy, że układ wektorów $\beta_1, \ldots, \beta_k$ jest liniowo zależny. Istnieją zatem $a_1, \ldots, a_k \in K$, nie wszystkie równe 0, że $a_1\beta_1 + \ldots + a_k\beta_k = 0$. Po ewentualnym przenumerowaniu wektorów możemy zakładać, że $a_1 \neq 0$ (tu nie ma żadnego oszustwa – proszę się nad tym chwilę zastanowić). Wtedy:

$$a_1\beta_1 = -a_2\beta_2 - \ldots - a_k\beta_k,$$

czyli

$$\beta_1 = -\frac{a_2}{a_1}\beta_2 - \frac{a_3}{a_1}\beta_3 - \ldots - \frac{a_k}{a_1}\beta_k.$$

Zatem β_1 jest kombinacją liniową pozostałych wektorów układu.

Na odwrót: jeśli jeden z wektorów układu jest kombinacją liniową pozostałych, to po ewentualnym przenumerowaniu możemy zakładać, że $\beta_1 = b_2\beta_2 + \ldots + b_k\beta_k$. Wtedy $\beta_1 - b_2\beta_2 - \ldots - b_k\beta_k = 0$, przy czym współczynnik przy β_1 jest równy 1, a więc jest niezerowy. Stąd układ $\beta_1, \ldots, \beta_k$ jest liniowo zależny. □

Powyższe stwierdzenie sugeruje następujący, kluczowy wniosek.

Wniosek 5.1

Jeśli wektor β jest kombinacją liniową wektorów $\beta_1, \ldots, \beta_n$, to

$$\mathrm{lin}(\beta, \beta_1, \ldots, \beta_n) = \mathrm{lin}(\beta_1, \ldots, \beta_n).$$

W szczególności, jeśli $V = \mathrm{lin}(\beta_1, \ldots, \beta_n)$, to z układu $\beta_1, \ldots, \beta_n$ wybrać można liniowo niezależny podukład $\alpha_1, \ldots, \alpha_k$ taki, że

$$V = \mathrm{lin}(\alpha_1, \ldots, \alpha_k) = \mathrm{lin}(\beta_1, \ldots, \beta_n).$$

Przyjmujemy też $\{0\} = \mathrm{lin}(\emptyset)$ (aby objąć $V = \{0\}$).

Dowód. Wystaczy udowodnić pierwszą część tezy. W tym celu dowodzimy dwie inkluzje. Z jednej strony, skoro dla pewnych $a_1, \ldots, a_n \in K$ mamy $\beta = a_1\beta_1 + \ldots + a_n\beta_n$, to także $\beta = 0 \cdot \beta + a_1\beta_1 + \ldots + a_n\beta_n$, czyli $\mathrm{lin}(\beta, \beta_1, \ldots, \beta_n) \subseteq \mathrm{lin}(\beta_1, \ldots, \beta_n)$. Przeciwna inkluzja jest natomiast oczywista. □

Pokazaliśmy, że każda przestrzeń rozpięta na skończonym układzie wektorów jest rozpięta przez pewien układ liniowo niezależny. Zobaczmy teraz, że przestrzeni rozpiętej przez układ liniowo niezależny $\mathcal{B}$ nie można rozpiąć przez układ liniowo niezależny $\mathcal{C} \supsetneq \mathcal{B}$.

Obserwacja 5.3

Niech $\alpha_1, \ldots, \alpha_k$ będzie układem liniowo niezależnym w przestrzeni V i niech wektor $\beta \in V$. Następujące warunki są równoważne:

(a) $\beta \in \text{lin}(\alpha_1, \ldots, \alpha_k)$,

(b) układ $\alpha_1, \ldots, \alpha_k, \beta$ jest liniowo zależny.

Dowód. Jeśli $\beta \in \text{lin}(\alpha_1, \ldots, \alpha_k)$, to układ $\alpha_1, \ldots, \alpha_k, \beta$ jest liniowo zależny na mocy Obserwacji 5.2. Na odwrót: przypuśćmy, że układ $\alpha_1, \ldots, \alpha_k, \beta$ jest liniowo zależny. Istnieją wówczas $b, a_1, \ldots, a_k \in K$, nie wszystkie równe 0, spełniające $b\beta + a_1\alpha_1 + \ldots + a_k\alpha_k = 0$. Rozważamy dwa przypadki.

- Przypadek 1: $b = 0$. Wówczas $a_1\alpha_1 + \ldots + a_k\alpha_k = 0$, przy czym pewne a_i jest niezerowe, co przeczy liniowej niezależności układu $\alpha_1, \ldots, \alpha_k$.
- Przypadek 2: $b \neq 0$. Mamy $\beta = -\frac{a_1}{b}\alpha_1 - \ldots - \frac{a_k}{b}\alpha_k$, czyli $\beta \in \text{lin}(\alpha_1, \ldots, \alpha_k)$.

□

Definicja 5.2: Baza (skończona) przestrzeni liniowej

Układ $\alpha_1, \ldots, \alpha_k$ wektorów przestrzeni V nazywamy BAZĄ PRZESTRZENI V, jeśli spełnia on następujące dwa warunki:

(a) układ $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależny,

(b) układ $\alpha_1, \ldots, \alpha_k$ ROZPINA V, to znaczy $V = \text{lin}(\alpha_1, \ldots, \alpha_k)$.

Wniosek 5.1 oznacza, że każda przestrzeń rozpięta na skończonym układzie wektorów posiada bazę. Celem kolejnego wykładu będzie pokazanie, że dwie skończone bazy przestrzeni liniowej są równoliczne[3], co doprowadzi nas do pojęcia wymiaru przestrzeni liniowej.

Przykład 1. W przestrzeni K^n rozważmy układ wektorów $\epsilon_1, \ldots, \epsilon_n$, zdefiniowany w następujący sposób, dla $i = 1, \ldots, n$:

$$\epsilon_i = (a_1, \ldots, a_n), \qquad \text{gdzie } a_j = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

Na przykład dla $n = 3$ mamy $\epsilon_1 = (1, 0, 0), \epsilon_2 = (0, 1, 0), \epsilon_3 = (0, 0, 1)$.

Układ $\epsilon_1, \ldots, \epsilon_n$ jest bazą przestrzeni K^n, zwaną BAZĄ STANDARDOWĄ przestrzeni K^n. Rzeczywiście:

- układ $\epsilon_1, \ldots, \epsilon_n$ jest liniowo niezależny, bowiem jeśli $a_1\epsilon_1 + \ldots + a_n\epsilon_n = (0, \ldots, 0)$, to zgodnie z działaniami w K^n mamy: $(a_1, a_2, \ldots, a_n) = (0, \ldots, 0)$. A zatem $a_1 = 0, a_2 = 0, \ldots, a_n = 0$.
- układ $\epsilon_1, \ldots, \epsilon_n$ rozpina K^n, Rzeczywiście, dowolny wektor $(x_1, x_2, \ldots, x_n)$ należy do $\text{lin}(\epsilon_1, \ldots, \epsilon_n)$, bo $(x_1, \ldots, x_n) = x_1(1, 0, ...) + x_2(0, 1, 0, ...) + \ldots + x_n(0, 0, \ldots, 1)$.

Przykład 2. Niech $V = (x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 + 2x_2 - x_3 = 0\}$, czyli $(x_1, x_2, x_3) \in V$ wtedy i tylko wtedy, gdy $x_1 = -2x_2 + x_3$. Wektory w V są zatem postaci:

$$(-2x_2 + x_3, x_2, x_3) = (2x_2, x_2, 0) + (x_3, 0, x_3) = x_2(-2, 1, 0) + x_3(1, 0, 1).$$

Stąd $V = \text{lin}((-2, 1, 0), (1, 0, 1))$. Wektory $(-2, 1, 0)$, $(1, 0, 1)$ są oczywiście liniowo niezależne (bo jeśli $a(-2, 1, 0) + b(1, 0, 1) = (0, 0, 0)$, to łatwo widzieć, że $a = b = 0$), a zatem układ ten jest bazą V.

Przykład 3. Niech $W = \text{lin}((1, 2, 1), (0, 1, 1), (1, 3, 2))$ będzie podprzestrzenią $\mathbb{R}^3$. Jest to przestrzeń rozpięta przez 3 wektory, ale nie jest to „oszczędny" układ. Wektor $(1, 3, 2)$ jest kombinacją liniową $(1, 2, 1), (0, 1, 1)$. A zatem układ $(1, 2, 1), (0, 1, 1), (1, 3, 2)$ nie jest bazą W. Jest nią natomiast układ $(1, 2, 1), (0, 1, 1)$, układ $(1, 2, 1), (1, 3, 2)$ i wiele innych.

[3] To nie wynika z Wniosku 5.1 – dlaczego?

Twierdzenie 5.1

Niech $\alpha_1, \ldots, \alpha_k$ będzie układem wektorów przestrzeni V. Wówczas następujące warunki są równoważne:

(1) układ $\alpha_1, \ldots, \alpha_k$ jest bazą przestrzeni V,

(2) każdy wektor $\alpha \in V$ można przedstawić w sposób jednoznaczny jako kombinację liniową układu $\alpha_1, \ldots, \alpha_k$.

Dowód. Zaczniemy od uzasadnienia implikacji (1) $\Rightarrow$ (2). Niech $\alpha_1, \ldots, \alpha_k$ będzie bazą przestrzeni V. Wówczas $V = \mathrm{lin}(\alpha_1, \ldots, \alpha_k)$, więc każdy $\alpha \in V$ jest kombinacją układu $\alpha_1, \ldots, \alpha_k$. Pozostaje wykazać jednoznaczność. Gdyby:

$$\alpha = a_1\alpha_1 + \ldots + a_k\alpha_k = a'_1\alpha_1 + \ldots + a'_k\alpha'_k,$$

dla pewnych $a_1, \ldots, a_k, a'_1, \ldots, a'_k \in K$, to mielibyśmy:

$$(a_1 - a'_1)\alpha_1 + \ldots + (a_k - a'_k)\alpha_k = 0.$$

Z liniowej niezależności wektorów $\alpha_1, \ldots, \alpha_k$ wynikałoby zatem, że $a_1 - a'_1 = \ldots = a_k - a'_k = 0$. A zatem rozkład każdego $\alpha \in V$ jest jednoznaczny.

Dowodzimy odwrotną implikację, (2) $\Rightarrow$ (1). Przypuśćmy, że każdy wektor $\alpha \in V$ można jednoznacznie przedstawić jako kombinację układu $\alpha_1, \ldots, \alpha_k$. Wykażemy, że $\alpha_1, \ldots, \alpha_k$ jest bazą. Oczywiście skoro każdy wektor z V jest kombinacją liniową wektorów $\alpha_1, \ldots, \alpha_k$, to układ ten rozpina V. A zatem warunek (b) z definicji bazy jest spełniony. Pozostaje pokazać, że układ $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależny. Przypuśćmy, że $a_1\alpha_1 + \ldots + a_k\alpha_k = 0$, dla pewnych $a_1, \ldots, a_k \in K$. Wówczas mamy:

$$a_1\alpha_1 + \ldots + a_k\alpha_k = 0\alpha_1 + \ldots + 0\alpha_k = 0,$$

a skoro także 0 ma jednoznaczny rozkład w V, to $a_1 = 0, a_2 = 0, \ldots, a_k = 0$, co dowodzi liniowej niezależności $\alpha_1, \ldots, \alpha_k$. □

Definicja 5.3

Niech V będzie przestrzenią liniową nad ciałem K i niech układ $\alpha_1, \ldots, \alpha_k$ będzie bazą przestrzeni V. WSPÓŁRZĘDNYMI WEKTORA $\alpha \in V$ W BAZIE $\alpha_1, \ldots, \alpha_k$ nazywamy układ elementów $a_1, \ldots, a_k$ ciała K spełniających

$$\alpha = a_1\alpha_1 + \ldots + a_k\alpha_k.$$

Przykłady:

- Wektor $(1, 2, 1)$ ma współrzędne $1, 2, 1$ w bazie standardowej przestrzeni $\mathbb{R}^3$.
- Wektor $(1, 2, 1)$ ma współrzędne $-1, 1, 1$ w bazie $\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ przestrzeni $\mathbb{R}^3$, bo $(1, 2, 1) = -1(1, 0, 0) + 1(1, 1, 0) + 1(1, 1, 1)$.
- Układ $(1, 0)$, $(2, 0)$ nie jest bazą $V = \mathrm{lin}((1, 0))$, bo mamy $(1, 0) = 1 \cdot (1, 0) = 1 \cdot (2, 0) + (-1) \cdot (1, 0)$.

Definicja 5.4

Mówimy, że układ $\alpha_1, \ldots, \alpha_k$ wektorów przestrzeni V jest

- MAKSYMALNYM UKŁADEM LINIOWO NIEZALEŻNYM, jeśli $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależny i każdy większy układ – to znaczy taki układ wektorów przestrzeni V, który zawiera $\alpha_1, \ldots, \alpha_k$ jako podukład właściwy – jest liniowo zależny,
- MINIMALNYM UKŁADEM ROZPINAJĄCYM V, jeśli $\alpha_1, \ldots, \alpha_k$ rozpina V i żaden mniejszy układ – to znaczy żaden podukład właściwy układu $\alpha_1, \ldots, \alpha_k$ nie rozpina V.

Przykłady.

- Układ $(1,0,0),(1,1,0)$ nie jest maksymalnym układem liniowo niezależnym $\mathbb{R}^3$. Jest to bowiem podukład właściwy układu liniowo niezależnego $(1,0,0),(1,1,0),(1,1,1)$.
- Układ $(1,0,0),(2,0,0)$ nie jest minimalnym układem rozpinającym $V = \mathrm{lin}((1,0,0),(2,0,0)) \subset \mathbb{R}^3$, bo układ $(1,0,0)$ jest jego podukładem właściwym i $V = \mathrm{lin}((1,0,0))$.

Twierdzenie 5.2

Niech $\alpha_1, \ldots, \alpha_k$ będzie układem wektorów przestrzeni liniowej V. Następujące warunki są równoważne.

(i) $\alpha_1, \ldots, \alpha_k$ jest bazą przestrzeni V,

(ii) $\alpha_1, \ldots, \alpha_k$ jest maksymalnym układem liniowo niezależnym w V,

(iii) $\alpha_1, \ldots, \alpha_k$ jest minimalnym układem rozpinającym V.

Dowód. Zaczniemy od implikacji $(i) \Rightarrow (ii)$. Niech $\alpha_1, \ldots, \alpha_k$ będzie bazą przestrzeni V. Gdyby układ $\alpha_1, \ldots, \alpha_k$ nie był maksymalny, to istniałby taki wektor $\alpha \in V$, że układ $\alpha_1, \ldots, \alpha_k, \alpha$ byłby liniowo niezależny. Wówczas jednak, zgodnie z Obserwacją 5.3, wektor α nie mógłby należeć do $\mathrm{lin}(\alpha_1, \ldots, \alpha_k)$. To jest jednak niemożliwe, bo $V = \mathrm{lin}(\alpha_1, \ldots, \alpha_k)$, zgodnie z definicją bazy. Zatem układ $\alpha_1, \ldots, \alpha_k$ jest maksymalnym niezależnym liniowo układem w V.

Dowód implikacji $(ii) \Rightarrow (i)$. Skoro $\alpha_1, \ldots, \alpha_k$ jest maksymalnym układem liniowo niezależnym w V, to jest on w szczególności liniowo niezależny. Do pokazania, że $\alpha_1, \ldots, \alpha_k$ jest bazą V pozostaje wykazać, że $V = \mathrm{lin}(\alpha_1, \ldots, \alpha_k)$. Jednak z maksymalności tego układu wynika, że dla każdego wektora $\alpha \in V$ układ $\alpha_1, \ldots, \alpha_k, \alpha$ jest liniowo zależny. W szczególności z Obserwacji 5.3 wynika, że α jest kombinacją liniową wektorów $\alpha_1, \ldots, \alpha_k$. A zatem istotnie $V = \mathrm{lin}(\alpha_1, \ldots, \alpha_k)$.

Dowód implikacji $(i) \Rightarrow (iii)$. Niech $\alpha_1, \ldots, \alpha_k$ będzie (ponownie) bazą V. Gdyby $\alpha_1, \ldots, \alpha_k$ nie był minimalnym układem rozpinającym V, to zawierałby podukład właściwy rozpinający V. Weźmy jednak dowolny wektor spośród $\alpha_1, \ldots, \alpha_k$ nie należący do tego podukładu. Jest on kombinacją liniową elementów tego podukładu, bo podukład ten rozpina (ponoć) przestrzeń V. Daje to sprzeczność z liniową niezależnością układu $\alpha_1, \ldots, \alpha_k$.

Dowód implikacji $(iii) \Rightarrow (i)$. Mamy minimalny układ $\alpha_1, \ldots, \alpha_k$ rozpinający przestrzeń V. Pokażemy, że jest on bazą tej przestrzeni. Wystarczy pokazać liniową niezależność tego układu. Gdyby układ ten był liniowo zależny, to któryś z $\alpha_1, \ldots, \alpha_k$ byłby liniową kombinacją pozostałych, na mocy Obserwacji 5.2. Np. (po ewentualnym przenumerowaniu)

$$\alpha_k = b_1\alpha_1 + \ldots + b_{k-1}\alpha_{k-1}.$$

Wówczas jednak

$$V = \mathrm{lin}(\alpha_1, \ldots, \alpha_{k-1}),$$

bo dla dowolnego $\alpha \in V$ istniałyby $a_1, \ldots, a_k \in K$, że:

$$\begin{aligned}\alpha &= a_1\alpha_1 + \ldots + a_k\alpha_k = \\ &= a_1\alpha_! + \ldots + a_k\underbrace{(b_1\alpha_1 + \ldots + b_{k-1}\alpha_{k-1})}_{\alpha_k} = \\ &= (a_1 + a_kb_1)\alpha_1 + \ldots + (a_{k-1} + a_kb_{k-1})\alpha_{k-1}.\end{aligned}$$

Jest to jednak sprzeczne z założeniem, że $\alpha_1, \ldots, \alpha_k$ jest minimalnym układem rozpinającym V. A zatem układ $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależny. □

Pojęcia podprzestrzeni rozpiętej na układzie wektorów, liniowej niezależności oraz bazy rozszerzyć można również w umiejętny sposób na nieskończone układy wektorów. Zanim zajmiemy się tymi konstrukcjami pokażemy, że każde dwie bazy przestrzeni rozpiętej przez skończony układ wektorów są równoliczne i wprowadzimy pojęcie wymiaru takich przestrzeni. Kluczowym narzędziem pomocniczym (lematem) będzie twierdzenie o wymianie, udowodnione przez Steinitza w 1910 roku.

5.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Rozstrzyganie czy układ wektorów jest liniowo niezależny — zadania 1–4)
   Dla jakich wartości parametrów $s, t \in \mathbb{R}$ wektory
   $$(5, 7, s, 2), \quad (1, 3, 2, 1), \quad (2, 2, 4, t)$$
   przestrzeni liniowej $\mathbb{R}^4$ tworzą układ liniowo niezależny?

2. Sprawdź czy układ wielomianów $w_1 = 1 + t, w_2 = 1 + t^2, w_3 = t + t^2$ jest liniowo niezależny w przestrzeni liniowej $\mathbb{R}[t]$. Czy odpowiedź zmienia się, jeśli układ ten rozpatrujemy w przestrzeni liniowej $\mathbb{Z}_2[t]$?

3. Wykaż, że funkcje $\sin(x)$ oraz $\cos(x)$ tworzą liniowo niezależny układ wektorów przestrzeni $F(\mathbb{R}, \mathbb{R})$.

4. Niech
   $$A_1 = \begin{bmatrix} a & 2a \\ 2 & 3a \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 \\ a & 3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 2a \\ a+1 & a+2 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 1 & a+1 \\ 2 & 2a+1 \end{bmatrix}.$$
   Dla jakich $a \in \mathbb{R}$ układ $\{A_1, A_2, A_3, A_4\}$ jest liniowo niezależny w przestrzeni liniowej $\mathrm{M}_2(\mathbb{R})$?

5. (♠ Rozstrzyganie czy układ wektorów jest liniowo niezależny)
   Wektory $\alpha_1, \alpha_2, \alpha_3$ tworzą układ liniowo niezależny w przestrzeni liniowej V nad ciałem $\mathbb{R}$. Rozpatrzmy wektor $\beta = -\alpha_1 + 7\alpha_2 - 3\alpha_3$. Czy wektory $\alpha_1, \alpha_2, \beta$ tworzą układ liniowo niezależny?

6. Niech $\alpha_1, \alpha_2, \ldots, \alpha_k$ będzie liniowo niezależnym układem wektorów przestrzeni V nad ciałem K. Czy układ $\beta_1, \ldots, \beta_k$ jest liniowo niezależny, jeśli:
   (a) $\beta_1 = \alpha_1$ oraz $\beta_i = \alpha_i + \alpha_{i-1}$, dla $i = 2, 3, \ldots, k$,
   (b) $\beta_i = \alpha_1 + \ldots + \alpha_i$, dla $i = 1, 2, \ldots, k$,
   (c) $\beta_i = \alpha_i + \alpha_{i+1}$, dla $i = 1, 2, \ldots, k-1$ oraz $\beta_k = \alpha_k + \alpha_1$.

7. Układ wektorów $v_1, \ldots, v_n$ przestrzeni liniowej V jest liniowo niezależny. Wykaż, że jeśli dla pewnego $w \in V$ układ $v_1 + w, \ldots, v_n + w$ jest liniowo zależny, to $w \in \mathrm{lin}(v_1, \ldots, v_n)$.

8. Wykaż, że istnieje nieskończony podzbiór wektorów przestrzeni $\mathbb{R}^n$, taki że każde n wektorów w tym podzbiorze tworzy układ liniowo niezależny.

9. (♠ Wybór bazy jako maksymalnego układu liniowo niezależnego)
   Z podanego podzbioru S przestrzeni wektorowej V wybierz maksymalny (względem inkluzji) podzbiór liniowo niezależny, gdy:
   - $V = \mathbb{R}^4$ oraz $S = \{(3, 2, 1, 1), (5, 0, 2, 3), (4, 1, 4, 5), (4, 1, -1, -1)\}$.
   - $V = \mathbb{C}^3$ oraz $S = \{(2, 1, 4), (3, 5, -1), (3, -2, 13), (7, 7, 7), (-4, -9, 6)\}$.
   - $V = (\mathbb{Z}_3)^3$ oraz $S = \{(1, 1, 1), (1, 2, 0), (0, 1, 2), (0, 0, 2), (2, 1, 2), (2, 2, 1)\}$.

10. (♠ Znajdowanie współrzędnych wektora w danej bazie — zadania 10, 11)
    Układ wektorów α, β tworzy bazę przestrzeni liniowej $\mathbb{R}^2$. Czy układ wektorów $\alpha + \beta, \alpha - \beta$ również tworzy bazę tej przestrzeni? Jeśli tak, to znajdź współrzędne wektorów α oraz β w tej bazie.

11. Rozpatrzmy następujące wektory przestrzeni $\mathbb{R}^3$:
    $$\alpha_1 = (3, 2, 1), \quad \alpha_2 = (7, 3, 1), \quad \alpha_3 = (4, 2, 1),$$
    $$\beta_1 = (0, 2, 1), \quad \beta_2 = (1, 1, 2), \quad \beta_3 = (1, 0, 0).$$
    (a) Wykaż, że $\alpha_1, \alpha_2, \alpha_3$ jest bazą przestrzeni $\mathbb{R}^3$. Dla $i = 1, 2, 3$ znajdź współrzędne β_i w tej bazie.
    (b) Podaj przykład takiej bazy przestrzeni $\mathbb{R}^3$, że wektor β_1 ma w niej współrzędne $1, 2, -1$ a wektor β_2 ma współrzędne $0, 0, 1$.
    (c) Czy istnieje taka baza przestrzeni $\mathbb{R}^3$, w której wektor β_1 ma współrzędne $1, 1, 0$, wektor β_2 ma współrzędne $0, 0, 1$, a wektor β_3 ma współrzędne $1, 1, 1$?

12. Niech $v_1 = (1, 2, 0), v_2 = (0, 1, 2), v_3 = (2, 0, 1)$. Dla jakich liczb pierwszych p układ $\{v_1, v_2, v_3\}$ jest bazą przestrzeni $\mathbb{Z}_p^3$? Ile jest baz przestrzeni $\mathbb{Z}_p^3$ (kolejność wektorów w bazie nie ma znaczenia)?

5.3 Uzupełnienie. Wielomiany ograniczonego stopnia i ich bazy

Z punktu widzenia analizy matematycznej, a także zastosowań matematyki, warto przyjrzeć się pewnym bazom przestrzeni wielomianów[4], przy czym na razie ograniczymy się do przestrzeni $K_{\leqslant n}[x]$ złożonej z wielomianów stopnia nie większego niż n. Rozpoczniemy od następującej obserwacji.

Obserwacja 5.4

Skończony układ wielomianów o parami różnych stopniach jest liniowo niezależny w $K[x]$.

Dowód. Niech $p_1, p_2, \ldots, p_m \in K[x]$ oraz niech $\deg(p_i) = d_i$. Po ewentualnym przenumerowaniu możemy założyć, że $d_1 > d_2 > \ldots > d_m$. Załóżmy, że dla pewnych $t_1, t_2, \ldots, t_m \in K$ mamy

$$t_1 p_1 + t_2 p_2 + \ldots + t_m p_m = 0. \qquad (\spadesuit)$$

Skoro $\deg(p_1) = d_1$, to niech ax^{d_1} będzie wyrazem najwyższego stopnia w p_1, dla pewnego $a \neq 0$. Skoro $d_1 > d_2 > \ldots > d_m$, to $t_1 a x^{d_1}$ jest jedynym wyrazem stopnia d_1 w wielomianie ($\spadesuit$), będącym zarazem wielomianem zerowym. Zatem $t_1 a x^{d_1} = 0$, skąd $t_1 a = 0$. Skoro $a \neq 0$, to $t_1 = 0$. A zatem została nam kombinacja liniowa wielomianów $t_2 p_2 + \ldots + t_m p_m = 0$, dla której możemy powtórzyć powyższy argument, co oznacza, że $t_i = 0$, dla $i = 2, 3, \ldots, n$. □

Nietrudno widzieć, że układ $\{1, x, x^2, \ldots, x^n\}$ jest bazą przestrzeni wielomianów stopnia nie większego niż n. Jak już wiemy, układ ten jest liniowo niezależny. Jasne jest jednak, że każdy wielomian stopnia nie większego niż n jest liniową kombinacją powyższych wielomianów. Na kolejnym wykładzie pokażemy, że każda przestrzeń liniowa rozpięta na skończonym układzie wektorów ma równoliczne bazy. Warto uświadomić sobie siłę tego twierdzenia, nawet wtedy gdy chcemy je odnieść do baz $K_{\leqslant n}[x]$. Pokażmy pewien szczególny przypadek.

Wniosek 5.2

Niech $p_0, p_1, \ldots, p_n \in K_{\leqslant n}[x]$ będą wielomianami odpowiednio stopni $0, 1, 2, \ldots, n$. Wówczas układ

$$\{p_0, p_1, \ldots, p_n\}$$

jest bazą przestrzeni liniowej $K_{\leqslant n}[x]$.

Dowód. Na mocy wcześniejszej obserwacji wystarczy pokazać, że $\operatorname{lin}(p_0, \ldots, p_n) = K_{\leqslant n}[x]$. Dowód jest indukcją ze względu na n. Dla $n = 0$ teza jest oczywista, bo wielomian stopnia 0 jest niezerowy. Załóżmy, że układ wielomianów $\{p_0, \ldots, p_k\}$ stopni od 0 do k rozpina $K_{\leqslant k}[x]$ i weźmy dowolny wielomian p_{k+1} stopnia $k+1$ należący do $K_{\leqslant k+1}[x]$. Pokażmy, że $\operatorname{lin}(p_0, \ldots, p_k, p_{k+1}) = K_{\leqslant k+1}[x]$. Weźmy dowolny wielomian w taki, że $\deg(w) \leqslant k+1$. Jeśli $\deg w < k+1$, to w jest z założenia indukcyjnego kombinacją liniową wielomianów $p_0, \ldots, p_k$. Jeśli $\deg(w) = k+1$, to istnieje takie $a \in K$, że

$$w - a \cdot p_{k+1}$$

jest wielomianem stopnia mniejszego niż $k+1$. A zatem $w - a \cdot p_{k+1} \in \operatorname{lin}(p_0, \ldots, p_k)$, co oznacza, że $w \in \operatorname{lin}(p_0, \ldots, p_k, p_{k+1})$. Zatem $\operatorname{lin}(p_0, \ldots, p_k, p_{k+1}) = K_{\leqslant k+1}[x]$, co kończy dowód. □

Wniosek 5.3

Niech $a \in K$. Układ wielomianów

$$\{1, (x-a), (x-a)^2, \ldots, (x-a)^n\}$$

jest bazą przestrzeni $K_{\leqslant n}[x]$.

[4]Na podstawie jednego z najlepszych podręczników do algebry jaki znam, autorstwa Keitha Nicholsona, udostępnionego przez Autora do ogólnego użytku. Jest to skarbnica wiedzy o zastosowaniach algebry liniowej i po prostu świetny tekst: `https://lyryx.com/linear-algebra-applications/`.

Wiemy, że jeśli mamy bazę przestrzeni liniowej, to każdy wektor zapisuje się jednoznacznie jako kombinacja liniowa elementów bazowych. Oznacza to, że każdy wielomian $f \in K_{\leqslant n}[x]$ może być przedstawiony w postaci:

$$f = a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots + a_n(x-a)^n.$$

Jak się okazuje, współczynniki a_i mają duże znaczenie w analizie. Są to bowiem współczynniki tzw. wielomianu Taylora funkcji f (dla nas to będzie funkcja wielomianowa, a na analizie to będzie funkcja różniczkowalna odpowiednią liczbę razy). Oczywiście z twierdzenia Bezout wynika, że $a_0 = f(a)$, czyli a_0 jest wartością funkcji wielomianowej odpowiadającej wielomianowi f. Czym są wyższe współczynniki?

Czytelnik znający wzór na pochodną złożenia (lub iloczynu) bez trudu sprawdzi, że pochodna funkcji wielomianowej $f(x)$ odpowiadająca powyższemu wielomianowi równa jest:

$$f^{(1)}(x) = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + \ldots + na_n(x-a)^{n-1}.$$

Oznacza to, że $f^{(1)}(a) = a_1$. Jeśli przez $f^{(n)} = f^{(1)}(f^{(n-1)})$ oznaczymy n-tą pochodną wielomianu f, wówczas biorąc $f = f^{(0)}$ (zerowa pochodna) otrzymujemy następujący wniosek.

Wniosek 5.4

Jeśli $f(x)$ jest funkcją wielomianową stopnia n, to

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Przykład. Przedstawmy $f(x) = 5x^3 + 10x + 2$ w bazie $\{1, x-1, (x-1)^2\}$. Kolejne pochodne funkcji wielomianowej $f(x)$ to:

$$f^{(1)}(x) = 15x^2 + 10, \quad f^{(2)}(x) = 30x, \quad f^{(3)}(x) = 30.$$

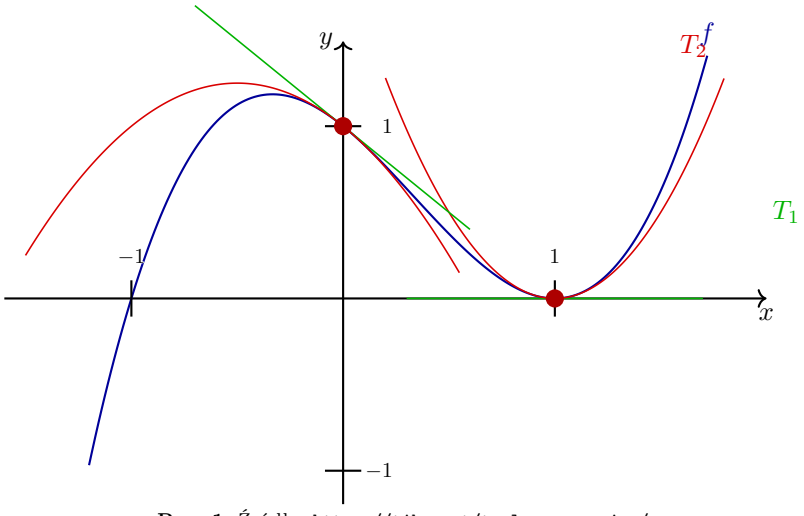
Stąd szukane przedstawienie ma postać:

$$f(x) = f(1) + \frac{f^{(1)}(1)}{1!}(x-1) + \frac{f^{(2)}(1)}{2!}(x-1)^2 + \frac{f^{(3)}(1)}{3!}(x-1)^3 = 17 + 25(x-1) + 15(x-1)^2 + 5(x-1)^3.$$

Wzór wyznaczony wyżej ma ogromne znaczenie w analizie, bowiem umożliwi Państwu badanie własności funkcji różniczkowalnych $n+1$ razy w sposób ciągły, poprzez zapisanie ich (w otoczeniu punktu x_0) w postaci:

$$f(x_0 + h) = w_n(x_0) + r(x_0., h),$$

gdzie $w_n(x)$ będzie n-tym wielomianem Taylora funkcji f w punkcie x_0, zaś $r(x_0, h)$ – tak zwaną resztą Peano we wzorze Taylora i w dalszej konsekwencji pozwoli na rozwijanie funkcji w szeregi. Oto przykład takiej sytuacji. Pewną funkcję f, różniczkowalną tyle razy ile trzeba, przybliżamy w punktach 0 oraz 1 Dla każdego z tych punktów wyznaczamy odpowiednie wielomiany Taylora stopnia 1 oraz 2. Ich wartości w tych punktach są oczywiście takie same, jak wartości funkcji f. W niewielkim otoczeniu tych punktów wielomiany te przybliżają z odpowiednio „kontrolowalną" dokładnością przebieg funkcji f.



Rys. 1. Źródło: `https://tikz.net/taylor_expansion/`.

Drugi ważny przykład dotyczy do pewnego stopnia innej sytuacji. Tym razem nie rozważamy wielomianów różnych stopni, ale szczególny układ $n+1$ wielomianów stopnia n, zwanych wielomianami Lagrange'a, który będzie stanowił bazę $K_{\leqslant n}[x]$. Załóżmy, że dany jest układ parami różnych elementów ciała K $a_0, a_1, \ldots, a_n$. Okazuje się, że dla każdego takiego układu argumentów znajdziemy bazę przestrzeni $K_{\leqslant n}[x]$ złożoną z wielomianów $\delta_0, \delta_1, \ldots, \delta_n$ taką, że dla każdego $f \in K_{\leqslant n}[x]$ mamy:

$$f = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \ldots + f(a_n) \cdot \delta_n.$$

Dla przykładu, dla układu trzech punktów a_0, a_1, a_2 będzie to baza:

$$\delta_0 = \frac{(x-a_1)(x-a_2)}{(a_0-a_1)(a_0-a_2)}, \quad \delta_1 = \frac{(x-a_0)(x-a_2)}{(a_1-a_0)(a_1-a_2)}, \quad \delta_2 = \frac{(x-a_0)(x-a_1)}{(a_2-a_0)(a_2-a_1)}.$$

Oto przykład. Dany jest wielomian $f(x) = x^2 - 2x + 1$ i wiemy, że $f(-1) = 4, f(0) = 1$ oraz $f(1) = 0$. Wielomiany określone wyżej mają dla $a_0 = -1, a_1 = 0, a_2 = 1$ postać:

$$\delta_0 = \frac{(x-0)(x-1)}{(-1-0)(-1-1)} = \frac{1}{2}(x^2-x), \quad \delta_1 = \frac{(x+1)(x-1)}{(0+1)(0-1)} = -(x^2-1), \quad \delta_2 = \frac{(x+1)(x-0)}{(1+1)(1-0)} = \frac{1}{2}(x^2+x).$$

Mamy też:

$$f(x) = 4 \cdot \frac{1}{2}(x^2 - x) + 1 \cdot (-1)(x^2 - 1) + 0 \cdot \frac{1}{2}(x^2 + x) = x^2 - 2x + 1.$$

Ktoś zapyta – po co nam takie rozwinięcia? Warto zauważyć, że każdy z trójki wielomianów $\delta_0, \delta_1, \delta_2$ ma tę własność, że $\delta_i(a_j) = 0$, dla $i \neq j$ oraz $\delta_i(a_i) = 1$. Innymi słowy, postulowane przez nas twierdzenie mówi, że jak z góry zadamy wartości $f(a_0), \ldots, f(a_n)$, to znajdziemy wielomian stopnia nie większego niż n, który w punktach $a_0, \ldots, a_n$ ma dokładnie te zadane z góry wartości. Na przykład chcąc, aby dla $a_0 = -1, a_1 = 0, a_2 = 1$ pewna funkcja kwadratowa przyjmowała wartości $10, 15, 30$, wystarczy rozpatrzyć wielomian

$$f = 10\delta_0 + 15\delta_1 + 30\delta_2.$$

W ten sposób skończone układy $n+1$ punktów na płaszczyźnie można **interpolować** wielomianami z $K_{\leqslant n}[x]$. Przejdźmy do twierdzenia[5], które wyjaśnia powyższą sytuację.

Twierdzenie 5.3

Załóżmy, że $a_0, a_1, \ldots, a_n$ są parami różne. Rozważmy zbiór wielomianów Lagrange'a $\delta_0, \delta_1, \ldots, \delta_n$ postaci:

$$\delta_k = \frac{\prod_{i \neq k}(x - a_i)}{\prod_{i \neq k}(a_k - a_i)}, \qquad \text{dla } k = 0, 1, 2, \ldots, n.$$

Wówczas układ $\delta_0, \ldots, \delta_n$ jest bazą $K_{\leqslant n}[x]$. Co więcej, dla każdego wielomianu $f \in K_{\leqslant n}[x]$ mamy:

$$f = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \ldots + f(a_n) \cdot \delta_n.$$

Dowód. W wielomianie δ_k licznik jest iloczynem wyrażeń liniowych $x - a_0, x - a_1, \ldots, x - a_n$ z pominięciem czynnika $x - a_k$. Podobnie dla mianownika. Widzimy więc, że wielomian δ_k przyjmuje wartości 0, dla wszystkich a_i różnych od a_k. Natomiast $\delta_k(a_k) = 1$.

Układ $\delta_0, \ldots, \delta_n$ jest liniowo niezależny. Rzeczywiście, jeśli $r_0 \cdot \delta_0 + r_1 \cdot \delta_1 + \ldots + r_n \cdot \delta_n = 0$, dla pewnych $r_1, \ldots, r_n \in K$, wówczas skoro porównujemy funkcje wielomianowe, mamy z prawej strony funkcję, która dla każdego x przyjmuje wartość zero. Wartość funkcji po lewej stronie dla a_0 równa jest $r_0\delta_0(a_0) + r_1\delta_1(a_0) + \ldots + r_n\delta_n(a_0) = r_0$, a zatem $r_0 = 1$. Analogicznie pokazujemy, że $r_2 = \ldots = r_n = 0$.

Układ $\delta_0, \ldots, \delta_n$ rozpina $K_{\leqslant n}[x]$. Istotnie, weźmy dowolny wielomian $f \in K_{\leqslant n}[x]$ i rozważmy kombinację liniową postaci $w = f(a_0) \cdot \delta_0 + f(a_1) \cdot \delta_1 + \ldots + f(a_n) \cdot \delta_n$. Wielomian w jest stopnia nie większego niż n i ma dla $a_0, \ldots, a_n$ takie same wartości, jak wielomian f. To znaczy, że wielomian $f - w$ jest stopnia nie większego niż n i ma $n+1$ pierwiastków. Musi być zatem, zgodnie z twierdzeniami z wykładu drugiego (wielomian stopnia n ma nie więcej niż n pierwiastków) wielomianem zerowym. Zatem $w = f$. □

[5] W tekście źródłowym jest ono sformułowane ogólniej i dowód jest identyczny jak poniżej. Źródło zakłada twierdzenie o równoliczności baz. My korzystamy z wiedzy o liczbie pierwiastków wielomianów stopnia n.

5.4 Dodatek. Permutacje w przestrzeni macierzy

Obok budowania ogólnej teorii przestrzeni liniowej i stosowania jej w przestrzeni współrzędnych K^n (niedługo zrozumiemy dlaczego ta przestrzeń jest tak ważna), dla lepszego zrozumienia geometrii jej podprzestrzeni (rozwiązań jednorodnych układów równań) warto zainteresować się podprzestrzeniami w przestrzeni macierzy. Motywacja jest następująca: za jakiś czas dokonamy utożsamienia macierzy z pewnymi przekształceniami przestrzeni liniowych. Podprzestrzenie przestrzeni macierzy odpowiadać będą zatem pewnym szczególnym typom przekształceń i ich badanie będzie miało duże znaczenie. Zacznijmy od przypomnienia podstawowego przykładu omówionego na wykładzie.

Definicja 5.5: Baza jedynek macierzowych w $M_{m\times n}(K)$

W przestrzeni liniowej macierzy rozmiaru $m\times n$ o wyrazach w ciele K rozważmy zbiór $m\cdot n$ macierzy E_{ij}, dla $1\leqslant i\leqslant m$ oraz $1\leqslant j\leqslant n$, takich że w i-tym wierszu i j-tej kolumnie macierzy E_{ij} stoi wyraz 1, a na pozostałych miejscach – 0. Innymi słowy jeśli E_{ij} jest macierzą o wyrazach a_{kl}, to:

$$a_{kl}=\begin{cases}1, & \text{gdy } (i,j)=(k,l),\\ 0, & \text{gdy } (i,j)\neq(k,l)\end{cases}.$$

Wówczas układ $\{E_{ij}, 1\leqslant i\leqslant m, 1\leqslant j\leqslant n\}$ jest bazą przestrzeni $M_{m\times n}(K)$.

Dla przykładu, w przestrzeni $M_{2\times 3}(K)$ zbiór jedynek macierzowych ma postać:

$$E_{11}=\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix},\quad E_{12}=\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix},\quad E_{13}=\begin{bmatrix}0&0&1\\0&0&0\end{bmatrix},$$

$$E_{21}=\begin{bmatrix}0&0&0\\1&0&0\end{bmatrix},\quad E_{22}=\begin{bmatrix}0&0&0\\0&1&0\end{bmatrix},\quad E_{23}=\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}.$$

Naśladując dowód z wykładu pokazujący, że baza standardowa jest bazą przestrzeni K^n pokazujemy, że układ jedynek macierzowych rozmiaru $m\times n$ jest bazą $M_{m\times n}$. Oczywiście wskazać można więcej baz przestrzeni macierzy. Z punktu widzenia zastosowań ciekawym aspektem jest rozpinanie podprzestrzeni na pewnych szczególnych typach macierzy. Przykładem takiej sytuacji jest dość trudne zadanie, które pojawiło się na kolokwium w 2019 roku. Dotyczyło ono podprzestrzeni $M_{n\times n}(K)$ rozpiętej przez tak zwane macierze permutacyjne. Powiedzmy więcej o permutacjach zbioru n-elementowego i tych macierzach.

Definicja 5.6: Permutacje i macierze permutacyjne

Przez S_n oznaczać będziemy zbiór wszystkich bijekcji (funkcji różnowartościowych i „na") zbioru n-elementowego $\{1,2,\ldots,n\}$. Funkcje te nazywamy PERMUTACJAMI zbioru n-elementowego.

Dla dowolnej permutacji $\sigma\in S_n$ określamy macierz $P_\sigma\in M_{n\times n}(K)$, której i-ta kolumna stanowi $\sigma(i)$-ty wektor bazy standardowej przestrzeni K^n. Podzbiór $M_{n\times n}(K)$ złożony z macierzy permutacyjnych również oznaczamy przez S_n.

Permutacje są obiektem fundamentalnym nie tylko w kombinatoryce, ale także w algebrze i teorii grup. Szczególnie widoczne stanie się to na wykładzie z teorii grup w ramach Algebry 1. Również ujęcie macierzowe ma jednak duże znaczenie.

Aby definicja była jasna ustalmy, że macierz P_{id} rozmiaru $n\times n$, która w i-tej kolumnie ma i-ty element bazy standardowej przestrzeni K^n to macierz permutacji identycznościowej. Obok zapiszmy macierz P_σ, gdzie permutacja σ polega jedynie na zamianie 1 z 2 oraz pozostawieniu na miejscu liczb od 3 do n:

$$P_{id}=\begin{bmatrix}1&0&0&\cdots&0&0&0\\0&1&0&\cdots&0&0&0\\0&0&1&\cdots&0&0&0\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\0&0&0&\cdots&1&0&0\\0&0&0&\cdots&0&1&0\\0&0&0&\cdots&0&0&1\end{bmatrix},\quad P_\sigma=\begin{bmatrix}0&1&0&\cdots&0&0&0\\1&0&0&\cdots&0&0&0\\0&0&1&\cdots&0&0&0\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\0&0&0&\cdots&1&0&0\\0&0&0&\cdots&0&1&0\\0&0&0&\cdots&0&0&1\end{bmatrix}.$$

Przykład. Rozważmy permutację $\sigma \in S_4$, czyli funkcję $\sigma : \{1,2,3,4\} \longrightarrow \{1,2,3,4\}$ daną wzorem:

$$\sigma(1) = 3, \quad \sigma(2) = 2, \quad \sigma(3) = 4. \quad \sigma(4) = 1.$$

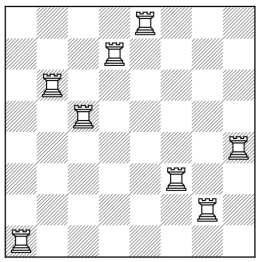
Notacja *tabelkowa* (będzie o niej mowa w przyszłości):

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$$

Macierz P_σ odpowiadająca tej permutacji ma postać:

$$P_\sigma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Z punktu widzenia kombinatoryki interesująca może być dla Państwa następująca obrazowa intuicja. Rozważmy takie rozstawienie wież na szachownicy (a_{ij}) rozmiaru $n \times n$, by w każdym wierszu i kolumnie znajdowała się dokładnie jedna wieża. Permutacji $\sigma \in S_n$ odpowiada jedno z $n!$ różnych rozstawień.



W powyższym przykładzie dla macierzy $[a_{ij}]$ rozmiaru 8×8 wieże rozstawione są na miejscach:

$$a_{\sigma(1)1}, a_{\sigma(2)2}, a_{\sigma(3)4}, a_{\sigma(4)4}, a_{\sigma(5)5}, a_{\sigma(6)6}, a_{\sigma(7)7}, a_{\sigma(8)8},$$

czyli:

$$a_{81}, a_{32}, a_{43}, a_{24}, a_{15}, a_{66}, a_{77}, a_{58}.$$

Tak jak w przestrzeni K^n czy $K[x]$ rozpinać można podprzestrzenie na różnych układach wektorów, tak i można pytać – jak wygląda podprzestrzeń $M_n(K)$ rozpiętych przez wszystkie macierze permutacyjne? Jak się okazuje jest to podprzestrzeń złożona z tzw. **macierzy półmagicznych**. To nie jest trywialne zadanie i w 2019 roku było ono motywem przewodnim ostatniego (dość trudnego) zadania na kolokwium. Oto ono.

Zadanie. Niech $I_n \in M_{n\times n}(\mathbb{Q})$ będzie macierzą, której jedyne niezerowe wyrazy znajdują się na przekątnej i są równe 1. Zbiór $\mathcal{P}_n$ zawiera wszystkie macierze powstałe z I_n przez wykonanie dowolnie wielu operacji elementarnych zamiany wierszy (w tym I_n). Niech $V = \text{lin}(\mathcal{P}_n)$. Wykazać, że:

(a) jeśli $A \in V$, to suma wyrazów w każdym wierszu i w każdej kolumnie macierzy A jest taka sama,

(b) jeśli W_0 jest podprzestrzenią $M_{n\times n}(\mathbb{Q})$ złożoną z macierzy o zerowej sumie wyrazów w każdym wierszu i w każdej kolumnie oraz jeśli $F_{ij} \in W_0$ jest taką macierzą, która na pozycjach $(i,j),(n,n)$ ma 1, na pozycjach $(i,n),(n,j)$ ma -1 oraz na pozostałych pozycjach ma 0, to

$$W_0 = \text{lin}(F_{ij}, 1 \leqslant i,j \leqslant n-1),$$

(c) jeśli macierz $R_{(i,j,n)}$ powstaje z I_n przez dwie kolejno wykonane operacje elementarne: zamianę wiersza i-tego z n-tym, a następnie zamianę wiersza j-tego z n-tym, dla $1 \leqslant i < j < n$, oraz jeśli macierz $S_{(i,n)}$ powstaje z I_n przez zamianę k-tego i n-tego wiersza, dla $1 \leqslant k < n$, to układ macierzy:

$$\{R_{(i,j,n)}, 1 \leqslant i < j \leqslant n-1\} \cup \{S_{(k,n)}, 1 \leqslant k \leqslant n-1\} \cup \{I_n\}$$

jest bazą V. W szczególności $\dim(V) = (n-1)^2 + 1$.

Rozwiązanie. Dowód (a). Jest jasne, że każda macierz ze zbioru $\mathcal{P}_n$ posiada dokładnie jeden niezerowy element w każdym wierszu i w każdej kolumnie, który jest równy 1 (można to udowodnić na przykład przez indukcję względem liczby operacji elementarnych zamiany wierszy wykonanych na I_n). A zatem każda macierz z tego zbioru ma jednakową sumę wyrazów w każdym wierszu i w każdej kolumnie. Wynosi ona 1.

Zauważmy teraz, że jeśli pewne dwie macierze $A, B \in M_{n\times n}(\mathbb{Q})$ mają tę własność, że suma wyrazów w każdym wierszu i w każdej kolumnie każdej z tych macierzy jest jednakowa i wynosi odpowiednio s_A oraz s_B, to macierz $A+B$ oraz macierze aA, gdzie $a \in \mathbb{Q}$, też mają tę własność, że suma wyrazów w każdym wierszu i w każdej kolumnie jest jednakowa, i wynosi ona odpowiednio $s_A + s_B$ oraz as_A. A zatem dowolna kombinacja liniowa macierzy o tej własności też ma tę własność. W szczególności $V = \operatorname{lin}(\mathcal{P}_n)$ złożona jest z **macierzy półmagicznych** (nie wiemy jednak czy każda macierz półmagiczna należy do $\operatorname{lin}(\mathcal{P}_n)$ – to pokażemy dalej).

Dowód (b). Macierz F_{ij} ma postać (przykład dla $i, j > 1$):

$$F_{ij} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & -1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & -1 & \cdots & 1 \end{bmatrix},$$

gdzie na czerwono podkreślone zostały i-ty wiersz i j-ta kolumna. Niech $Z = (z_{ij}) \in M_{n\times n}(\mathbb{Q})$ będzie macierzą, w której sumy wyrazów w każdym wierszu i każdej kolumnie wynoszą 0. Wówczas dla każdego $1 \leqslant i \leqslant n$ kombinacja liniowa

$$z_{i1}F_{i1} + z_{i2}F_{i2} + \ldots + z_{i,n-1}F_{i,n-1}$$

równa jest (zgodnie z założeniem o sumie wyrazów w i-tym wierszu $z_{i1} + z_{i2} + \ldots + z_{i,n-1} + z_{i,n} = 0$):

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{i1} & z_{i2} & \ldots & z_{i,n-1} & -z_{i1} - z_{i2} - \ldots - z_{i,n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -z_{i1} & -z_{i2} & \ldots & -z_{i,n-1} & z_{i1} + z_{i2} + \ldots + z_{i,n-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{i1} & z_{i2} & \ldots & z_{i,n-1} & z_{i,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -z_{i1} & -z_{i2} & \ldots & -z_{i,n-1} & -z_{i,n} \end{bmatrix}$$

Oznacza to, że macierz postaci

$$\sum_{i=1}^{n-1} z_{i1}F_{i,1} + z_{i2}F_{i,2} + \ldots + z_{i,n-1}F_{i,n-1}$$

można zapisać jako:

$$\begin{bmatrix} z_{11} & \ldots & z_{1,n} \\ 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \\ -z_{11} & \ldots & -z_{1,n} \end{bmatrix} + \begin{bmatrix} 0 & \ldots & 0 \\ z_{21} & \ldots & z_{2,n} \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \\ -z_{21} & \ldots & -z_{2,n} \end{bmatrix} + \ldots + \begin{bmatrix} 0 & \ldots & 0 \\ 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ z_{n-1,1} & \ldots & z_{n-1,n} \\ -z_{n-1,1} & \ldots & -z_{n-1,n} \end{bmatrix} = Z.$$

W ostatniej równości korzystamy z założenia, że suma wyrazów każdej z kolumn jest zerowa, a więc mamy na przykład równość

$$-z_{11} - z_{21} - \ldots - z_{n-1,1} = z_{n1}.$$

W rezultacie macierz Z jest kombinacją liniową macierzy F_{ij} postaci:

$$Z = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} z_{ij}F_{ij}.$$

Pokazaliśmy zatem, że każdy element podprzestrzeni W_0 jest kombinacją liniową macierzy F_{ij}.

Dowód (c). Znowu warto się przyjrzeć jakie permutacje opisują macierze $R_{(i,j,n)}$ oraz $S_{(k,n)}$. Pierwsza z nich umieszcza w j-tej kolumnie i-ty wektor standardowy, w n-tej kolumnie umieszcza j-ty wektor standardowy oraz n-ty wektor standardowy umieszcza w i-tej kolumnie (pozostałe są na swoich pozycjach). Podobnie macierz $S_{(k,n)}$ powstaje przez zamianę k-tej oraz n-tej kolumny. Oto ilustracja.

$$R_{(i,j,n)} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \end{bmatrix}, \quad S_{(k,n)} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \end{bmatrix},$$

Zauważmy, że jeśli $i \neq j$ oraz $i, j < n$, a także jeśli $k < n$, wówczas mamy:

$$F_{ij} = I_n - S_{(i,n)} - S_{(j,n)} + R_{(i,j,n)}, \quad F_{k,k} = I_n - S_{(k,n)}. \qquad (\star)$$

Rzeczywiście F_{ij} równe jest (można też to policzyć rozkładając wszystko na jedynki macierzowe):

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} - \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \end{bmatrix} - \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \end{bmatrix}.$$

Zauważmy też, że macierze $R_{(i,j,n)}$ oraz $S_{(k,n)}$ należą do $\mathcal{P}_n$. A zatem na mocy punktów (a) i (b) mamy:

$$W_0 \overset{(b)}{=} \operatorname{lin}(F_{ij}) \overset{(\star)}{\subseteq} \operatorname{lin}(\{R_{(i,j,n)}\} \cup \{S_{(k,n)}\} \cup \{I_n\}) \subseteq \operatorname{lin}(\mathcal{P}_n) = V \overset{(a)}{\subseteq} W,$$

gdzie W jest podprzestrzenią wszystkich macierzy, które mają taką samą sumę wyrazów w każdym wierszu i w każdej kolumnie (oczywiście $W_0 \subseteq W$).

Zauważmy jednak, że jeśli S jest macierzą, której suma elementów w każdym wierszu i każdej kolumnie równa jest t, to $S - tI_n \in W_0$. Każda macierz z W jest w rezultacie, na mocy punktu (b), kombinacją liniową macierzy F_{ij} oraz I_n. A zatem $W = \operatorname{lin}(F_{ij} \cup \{I_n\})$. Stąd:

$$W = \operatorname{lin}(F_{ij} \cup \{I_n\}) \subseteq \operatorname{lin}(\{R_{(i,j,n)}\} \cup \{S_{(k,n)}\} \cup \{I_n\}) \subseteq V \subseteq W \quad \Rightarrow V = W.$$

Układ $\{F_{ij}, 1 \leqslant i, j < n\}$ jest liniowo niezależny. Wynika to natychmiast z dowodu (b), gdzie opisaliśmy `każdą` liniową kombinację wektorów F_{ij}. Jeśli macierz Z z punktu (b) jest zerowa, to wszystkie z_{ij} są zerowe. Skoro więc $I_n \notin W_0$, to widzimy, że układ $\{F_{ij}, 1 \leqslant i, j < n\}$ tworzy bazę W_0, a układ

$$\{F_{ij} \cup I_n, 1 \leqslant i, j < n\}$$

tworzy bazę W. Dostajemy zatem $\dim(V) = \dim(W) = (n-1)^2 + 1$. Jednak zbiór

$$\{R_{(i,j,n)}, 1 \leqslant i < j \leqslant n-1\} \cup \{S_{(k,n)}, 1 \leqslant k \leqslant n-1\} \cup \{I_n\}$$

rozpina $W = V$ oraz ma $(n-1)^2 + 1$ elementów. Jest to zatem minimalny układ rozpinający $W = V$, a więc także baza przestrzeni V. Dowód jest zakończony. *Uwaga.* Można też pokazać wprost, że układ $\{R_{(i,j,n)}, 1 \leqslant i < j \leqslant n-1\} \cup \{S_{(k,n)}, 1 \leqslant k \leqslant n-1\} \cup \{I_n\}$ jest liniowo niezależny, patrząc na ostatnie wiersze i ostatnie kolumny dowolnych ich kombinacji liniowych. W ten sposób w ogóle nie musielibyśmy korzystać z twierdzenia o wymiarze (mówiącego, że każde dwie bazy są równoliczne).

Uzasadniony wynik pozwala już łatwo na udowodnienie ważnego twierdzenia Birkhoffa mówiącego, że zbiór tzw. wypukłych (kombinacje liniowe nad $\mathbb{R}$ o współczynnikach nieujemnych i sumie 1) kombinacji macierzy permutacyjnych równy jest dokładnie zbiorowi macierzy podwójnie stochastycznych. Twierdzenie to mówi, że w przestrzeni (afinicznej) wymiaru $n!$ wielościan o wierzchołkach w macierzach permutacyjnych złożony jest z macierzy półmagicznych o wyrazach nieujemnych i sumie w każdym wierszu i kolumnie równej 1. O szczegółach jeszcze napiszemy – to temat ważny m.in. w analizie i probabilistyce.

5.5 Trivia. Wektory przynależności do klubów

Wiele ciekawych zastosowań algebry liniowej odnaleźć można w kombinatoryce[6]. Weźmy zbiór n elementowy, na przykład złożony z mieszkańców pewnego miasta. Załóżmy dalej, że w mieście tym są pewne kluby, na przykład Wisła i Cracovia. Każdemu z tych klubów przypisujemy wektor v_1, v_2 o n współrzędnych w zbiorze $\{0, 1\}$ w następujący sposób: i-ta współrzędna wektora v_l jest równa

- 1, jeśli i-ty mieszkaniec naszego kraju jest kibicem tego klubu
- 0, jeśli i-ty mieszkaniec naszego kraju nie jest kibicem tego klubu,

Gdyby miasto to miało $n = 4$ mieszkańców $1, 2, 3, 4$, z których 1 i 4 są kibicami Wisły, a $2, 3$ nimi nie są, wówczas mielibyśmy $v_1 = (1, 0, 0, 1)$.

Wprowadzamy operację na parach wektorów przynależności, która ma czytelną interpretację kombinatoryczną. Biorąc dwa wektory $x = (x_1, \ldots, x_n)$ oraz $y = (y_1, \ldots, y_n)$, reprezentujące przynależność do grona kibiców pewnych dwóch klubów i rozważając operację:

$$\langle x,\, y\rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

widzimy, że wartość $\langle x,\, y\rangle$ równa jest po prostu liczbie kibiców, wspierających jednocześnie obydwa kluby, a liczba $\langle x, x\rangle$ to liczba kibiców Wisły. Przykładowo, jeśli wektor kibiców Wisły ma postać $(1, 0, 0, 1)$, a wektor kibiców Cracovii ma postać $(0, 0, 0, 1)$, to

$$\langle (1, 0, 0, 1), (0, 0, 0, 1)\rangle = 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 = 1,$$

a zatem tylko jeden kibic należy jednocześnie do grona sympatyków Wisły i Cracovii.

Być może Czytelnik nie dostrzega jeszcze do czego mogłyby nam się tu przydać kombinacje liniowe czy liniowa niezależność. Sformalizujmy nasze pojęcia, przechodząc do przestrzeni współrzędnych nad dowolnym ciałem. Aby to zrobić, wprowadzimy tak zwany **wektor przynależności** do klubu w mieście o n mieszkańcach, który będzie elementem przestrzeni liniowej K^n nad ciałem K.

Dokładniej, niech m będzie liczbą klubów $C_1, C_2, \ldots, C_m$ w mieście o n mieszkańcach $\{1, 2, \ldots, n\}$. Dla każdego $i = 1, 2, \ldots, m$, niech $n_i \in K^n$ będzie wektorem kodującym przynależność mieszkańców do i-tego klubu, to znaczy $v_i = (v_{i1}, \ldots, v_{in})$, gdzie:

$$v_{ij} = \begin{cases} 1, & j \in C_i \\ 0, & j \notin C_i \end{cases}.$$

W przykładach wyżej mamy $m = 2$, C_1 to klub Wisła, C_2 to klub Cracovia, a $n = 4$. Dla wektorów przynależności do klubów C_i oraz C_j, czyli $v_i = (v_{i1}, \ldots, v_{in})$ oraz $v_j = (w_{j1}, \ldots, w_{jn})$ należących do K^n, definiujemy operację:

$$\langle v_i, v_j\rangle = \langle (v_{i1}, \ldots, v_{in}), (v_{j1}, \ldots, v_{jn})\rangle = v_{i1}v_{j1} + v_{i2}v_{j2} + \ldots + v_{in}v_{jn} \in K.$$

Zauważmy, że dla $K = \mathbb{R}$ wartość powyższej liczby równa jest liczbie wspólnych członków obydwu klubów. Dla $K = \mathbb{Z}_2$ dostajemy jedynie informacje, czy liczby te są parzyste, czy nieparzyste. Czytelnikowi zostawiamy uzasadnienie jeszcze jednej ważnej własności wprowadzonej przez nas operacji (wynika ona wprost z definicji). Jeśli wektory v, w, z należą do K^n, wówczas dla dowolnych $c_1, c_2 \in K$ mamy:

$$\langle v, c_1w + c_2z\rangle = c_1\langle v, w\rangle + c_2\langle v, z\rangle. \qquad (\dagger)$$

Obserwacja 5.5: Problem Parzystkowa

W mieście o n mieszkańcach i m klubach, w którym każdy klub ma nieparzystą liczbę członków i każde dwa kluby mają parzystą liczbę wspólnych członków, zachodzi $m \leqslant n$.

[6]Na podstawie: Y. Zhao: Linear algebra tricks for the Putnam, `yufeizhao.com/olympiad/putnam_linear_algebra.pdf`.

Dowód. Rozważmy wektory przynależności $v_1, \ldots, v_m \in \mathbb{Z}_2^n$. Zauważmy, że skoro wektor v_i ma nieparzyście wiele niezerowych współrzędnych, to $\langle v_i, v_i \rangle = 1$. Co więcej, dla każdych $i \neq j$ mamy $\langle v_i, v_j \rangle = 0$, ponieważ dowolne dwa kluby mają parzystą liczbę wspólnych członków. Twierdzimy, że wynika stąd, że wektory $v_1, \ldots, v_m$ są liniowo niezależne w $\mathbb{Z}_2^n$. Istotnie, jeśli istnieją $c_1, \ldots, c_m \in \mathbb{Z}_2$, takie że

$$c_1 v_1 + c_2 v_2 + \ldots + c_m v_m = 0,$$

to mamy również, na mocy (†):

$$0 = \langle v_i, 0 \rangle = \langle v_i, c_1 v_1 + c_2 v_2 + \ldots + c_m v_m \rangle = c_1 \langle v_i, v_1 \rangle + c_2 \langle v_i, v_2 \rangle + \ldots + c_m \langle v_i, v_m \rangle = c_i.$$

Stąd układ wektorów $v_1, \ldots, v_m$ jest liniowo niezależny. Jednak układ m wektorów w przestrzeni K^n jest liniowo niezależny jedynie, gdy $m \leqslant n$. Dlaczego? W zasadzie wykażemy to dopiero na następnym wykładzie, ale dla przestrzeni K^n to jest jasne — jeśli $m > n$, to wpisując wektory $v_1, \ldots, v_m$ w wiersze macierzy rozmiaru $m \times n$ wiemy, że wykonywanie operacji elementarnych zamienia wiersze na kombinacje liniowe innych wierszy, a zatem po uzyskaniu postaci schodkowej tej macierzy, która musi mieć nie więcej niż n wierszy, jeden z $m > n$ wierszy macierzy musiał się wyzerować. To jednak oznacza, że wektor zerowy jest kombinacją liniową wierszy $v_1, \ldots, v_m$ o dodatnich współczynnikach, co jest niemożliwe, jeśli są one liniowo niezależne. □

W dowodzie skorzystaliśmy z obserwacji, która mówi, że m wektorów w przestrzeni K^n jest liniowo niezależne, jedynie gdy $m \leqslant n$ (nie jest to jednak warunek dostateczny). Zobaczmy jeszcze jedno zastosowanie tej obserwacji, tym razem pracując z wektorami przynależności w przestrzeni $\mathbb{R}^n$.

Obserwacja 5.6: Nierówność Fishera

Niech k będzie dodatnią liczbą całkowitą. W mieście o n mieszkańcach utworzono m klubów, przy czym każde dwa kluby mają dokładnie k wspólnych członków. Wykaż, że $m \leqslant n$.

Idea dowodu jest podobna, choć tym razem wektory $v_1, \ldots, v_m$ reprezentujące kluby $C_1, \ldots, C_m$ należeć będą do $\mathbb{R}^n$. W takim przypadku $\langle v_i, v_i \rangle$ opisuje liczbę członków klubu C_i, a liczba $\langle v_i, v_j \rangle$ opisuje liczbę wspólnych członków klubów C_i oraz C_j, czyli k.

Dowód. Rozumując podobnie jak poprzednio stwierdzamy, że jeśli $v_1, \ldots, v_m$ są liniowo niezależne, to $m \leqslant n$. Przypuśćmy, że układ ten nie jest liniowo niezależny. Istnieją zatem $c_1, \ldots, c_m \in \mathbb{R}$, że

$$c_1 v_1 + \ldots + c_m v_m = 0$$

Niech $v_i = (v_{i1}, \ldots, v_{in})$. Zatem korzystając z tego, że $\langle v, w \rangle = \langle w, v \rangle$ oraz korzystając wielokrotnie z (†), uzyskujemy

$$\begin{aligned}
\langle c_1 v_1 + \ldots + c_m v_m, c_1 v_1 + \ldots + c_m v_m \rangle &= \sum_{i=1}^{m} c_i^2 \langle v_i, v_i \rangle + 2 \sum_{i<j} c_i c_j \langle v_i, v_j \rangle = \\
&= \sum_{i=1}^{m} c_i^2 \langle v_i, v_i \rangle + 2k \sum_{i<j} c_i c_j = \\
&= \sum_{i=1}^{m} c_i^2 (\langle v_i, v_i \rangle - k) + \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j = \\
&= \sum_{i=1}^{m} c_i^2 (\langle v_i, v_i \rangle - k) + \left(\sum_{i=1}^{n} c_i \right)^2.
\end{aligned}$$

Skoro dowolne dwa kluby mają dokładnie k wspólnych członków, to każdy klub ma ich co najmniej k. Stąd $\langle v_i, v_i \rangle \geqslant k$. W szczególności wszystkie wyrazy końcowej sumy są nieujemne, a stąd muszą być równe 0. Skoro pewne c_i jest niezerowe, to pewne $\langle v_i, v_i \rangle - k$ jest równe 0, a więc klub C_i ma dokładnie k członków. Stąd jednak wynika, zgodnie z założeniem twierdzenia, że wszystkie kluby zawierają wszystkich k członków klubu C_i, a poza tym są rozłączne (zawierają unikatowych mieszkańców). Stąd wynika natychmiast, że klubów tych jest nie więcej niż mieszkańców, czyli n.

□

5.6 Coda. Kombinacje, czyli o przestrzeni barw

Pojęcia przestrzeni liniowej, kombinacji, podprzestrzeni rozpiętej na układzie i wreszcie układów liniowo zależnych i niezależnych oraz bazy — to spora dawka abstrakcji na przestrzeni dwóch tygodni zajęć. Warto poświęcić moment na kilka przyjaznych intuicji, niezupełnie zresztą odległych od rzeczywistości, mających bowiem głębokie podłoże historyczne, związane także z początkami algebry liniowej[7].

Mówiąc w skrócie chodzi o mieszanie kolorów. Kolory możemy traktować jak wektory i uzyskiwać rozmaite ich kombinacje. Oto przykład kilku „kombinacji liniowych" koloru zielonego i czerwonego.

Rys. 2. Kombinacje liniowe $az + br$, dla $a = \frac{k}{10}$ oraz $b = 1 - a$, dla $a = \frac{10}{10}, \frac{9}{10}, \frac{8}{10}, \ldots, \frac{2}{10}, \frac{1}{10}, \frac{0}{10}$.

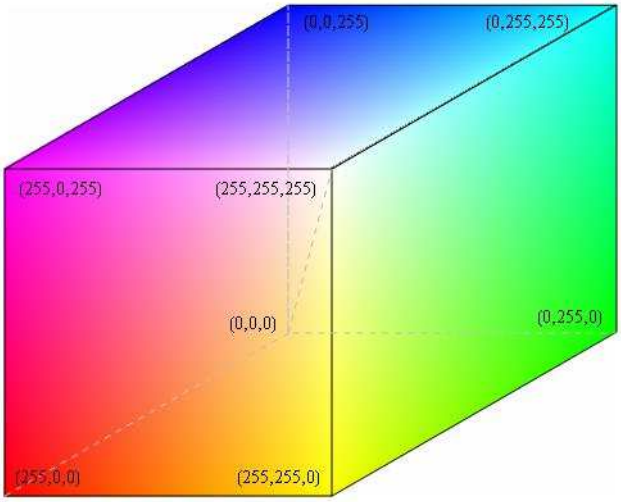
Czy kombinacją zielonego i czerwonego nie jest po prostu kolor żółty? Tak, ale zółty jest w istocie sumą $z+r$. Jak to wyjaśnić? Mowa tu mianowicie o modelu RGB „przestrzeni kolorów", opartego o wyróżnienie trzech barw podstawowych: czerwonej, zielonej oraz niebieskiej, przypisanie im wektorów o współrzędnych

$$r = (255, 0, 0), \quad z = (0, 255, 0), \quad n = (0, 0, 255)$$

lub — jeśli ktoś woli notację w zapisie szesnastkowym (heksagonalnym) — FF0000, 00FF00, 0000FF, i rozważanie wszystkich kombinacji liniowych tych wektorów postaci:

$$ar + bz + cn,$$

ale tylko w zakresie $a, b, c \in [0.1]$. Oto reprezentacja przestrzeni barw RGB.



Rys. 3. Paleta barw RGB, źródło: `https://www.whydomath.org/node/wavlets/imagebasics.html`.

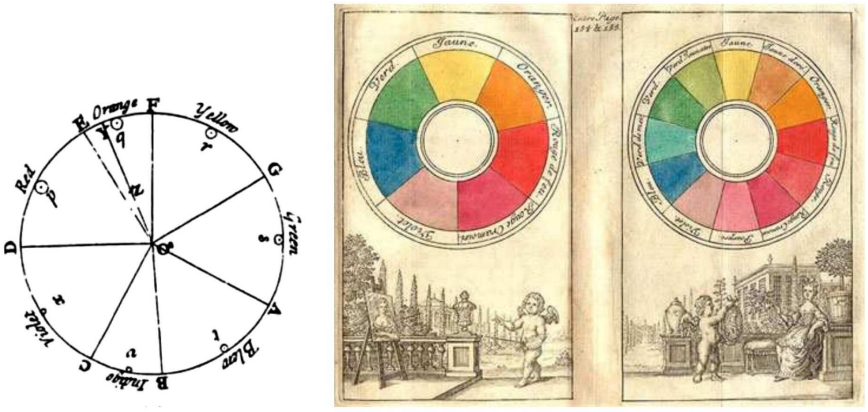
Wierzchołek sześcianu o współrzędnych $(255, 255, 255)$ odpowiada wektorowi który w przestrzeni barw RGB oznacza kolor biały. Wektor $(0, 0, 0)$ oznacza kolor czarny. Kolory mieszane przez nas wyżej „leżą" na przekątnej niewidocznej dla nas podstawy łączącej: czerwony i zielony wierzchołek. Natomiast wektor $z + r$ odpowiada wierzchołkowi $(255, 255, 0)$, widocznemu na naszym rysunku jako wierzchołek żółty.

Nasz model posiada wiele wad, odróżniających go od przestrzeni liniowej i nic dziwnego — teoria kolorów i ich postrzegania ma olbrzymią historię i bardzo liczne modele (dziś głównie nieliniowe). Mimo wszystko można coś powiedzieć i przy okazji wzmocnić swoje intuicje z algebry liniowej. Po pierwsze: mnożenie wektora przez skalar prowadzi do zmiany jego nasycenia. Podstawowy kolor określony jest przez kolory podstawowe. Zerowe nasycenie oznacza zawsze kolor czarny (a w innych modelach — biały). Oto przykład dla skalarnych wielokrotności wektora r , postaci $a \cdot r$, dla $a \in \{\frac{0}{10}, \frac{1}{10}, \frac{2}{10}, \ldots, \frac{8}{10}, \frac{9}{10}, \frac{10}{10}\}$.

[7] Zachęcam też do lektury tekstu: `https://scholar.harvard.edu/files/schwartz/files/lecture17-color.pdf`

W naszym modelu nie istnieje wektor $1,5r$. Liniowa suma dwóch kolorów reprezentowanych jak wyżej nie musi być kolorem. Czy ta teoria ma sens? Przecież mamy doświadczenie mówiące, że każde kolory można zmieszać. Na czym więc polega mieszanie kolorów? Pytania tego typu stawiał już w starożytności Platon i jego uczeń Arystoteles, a nawet i wcześniejszy autorzy: Philolaus ze szkoły pitagorejskiej, Plutarch, Empedokles, a nawet Demokryt, uważany za twórcę teorii atomicznej budowy świata. Chodziło oczywiście o wyróżnienie barw podstawowych, mających również zastosowania sakralne — białego (leukhyn), czarnego (melan), czerwonego (erydron), zielonego (khloron) oraz uzyskiwania innych kolorów jako ich mieszanin. Dla filozofów greckich podstawowe kolory reprezentowały raczej fundamentalne własności materii – a ich mieszanie miało odwzorowywać złożone własności obiektów mających uzyskany kolor.

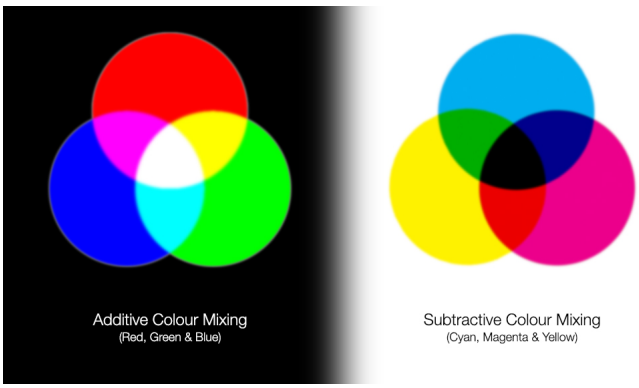
Z naukowego punktu widzenia przełomem były badania Newtona rozpoczęte w roku 1665. W dziele *Optyka* Newton wyjaśnia zjawisko znane ludzkości od zawsze pod symbolem tęczy — światło białe rozszczepia się na siedem rozróżnialnych barw, które reprezentować można na tzw. okręgu barw.



Rys. 4. Tarcza kolorów Newtona (Optyka, 1704) wzorowana na muzycznym kole kwintowym służącym do zmiany tonacji. Zauważcie Państwo, że każdy kolor jest mieszaniną barw. Obok popularyzująca je reprodukcja malarza Claude Bouteta (1708),

Słynny uczony nie tylko dokonał rozszczepienia światła widzialnego za pomocą pryzmatu. To jest nieco bardziej skomplikowane i istotne dla naszych intuicji. Newton rozumiał, że widzialne (białe) światło jest kombinacją fal różnej długości. W swoim przełomowym dziele *Optyka* zaprzeczył powszechnej opinii uczonych, jakoby to zanieczyszczenie pryzmatu powodowało tęczę. Użył w tym celu systemu soczewek, luster i pryzmatów, dzięki którym odizolował światło czerwone i skierował je na pryzmat, który wypuścił światło czerwone. Następnie rozbił światło widzialne na kolory tęczy i skierował je z powrotem na pryzmat, uzyskując z powrotem światło widzialne. Gorące dyskusje wokół teorii Newtona trwały 200 lat.

Co to ma wszystko wspólnego z algebrą liniową? Aby to zrozumieć, trzeba sięgnąć znowu do historii. Demokryt twierdził, iż obserwowane obiekty wysyłają do oka „atomy" wywołujące obraz. Teoria Newtona rozszczepiania światła spotkała się z dużym sprzeciwem m.in. Goethego, który nie przyjmował idei uzyskiwania bieli z barw chromatycznych, myśląc wyłączenie o syntezie subtraktywnej, czyli efektach odbicia światła od powierzchni pokrytych mieszaninami barwników, pochłaniających fale o różnych długościach (nakładanie się tych efektów powoduje wzrost udziału czerni). Różnice między syntezą subtraktywną i addytywną stała się znane dzięki pracom Tobiasa Mayera, autora pracy *De Affi nitate Colorum Commentatio* z 1772 roku. Zobaczmy na obrazku czym różni się subtraktywne i addytywne mieszanie.



Rys. 5. Subtraktywne i addytywne mieszanie kolorów w systemach RGB i CMYK.

Aby wylądować wreszcie w algebrze liniowej brakuje nam postaci Grassmanna, dla którego teoria koloru była również osobistą pasją i jedną z motywacji do rozpatrywania przestrzeni liniowych. Skąd to się wzięło? W 1802 roku Thomas Young sformułował tzw. teorię trójchromatyczną, która miała wyjaśniać widzenie kolorów dzięki obecności w siatkówce człowieka i innych naczelnych trzech różnych fotoreceptorów absorbujących światło w różnych zakresach długości. Już wcześniej formułowano empirycznie zasady mieszania barw, głównie w oparciu o syntezę subtraktywną. Odkrycie Newtona było przełomem w tym sensie, że za widzenie barw odpowiadają cechy światła, a nie widzianych obiektów. Young wchodzi do tej opowieści jako fizyk i co ważniejsze — lekarz. Uważał on, że jednakowe pobudzenie trzech rodzajów włókiem wywołuje wrażenie bieli (co wspiera rozumienie addytywne — kolokwialnie mówiąc kolory nie składają się przy widzeniu do czarnego – czarny jest wtedy, gdy nie patrzymy). Teorię Younga wsparł w połowie lat 50. tych XX wieku niemiecki fizjolog, fizyk i filozof Hermann von Helmholtz. Uważał on, że reaktywność światłoczułych receptorów zależy od długości fali i jest największa wówczas, gdy długość fali odpowiada barwom podstawowym: czerwonej, zielonej i fioletowej. Helmholtz zdefiniował też trzy stosowane do dziś cechy otrzymywanych kolorów: barwa (hue), nasycenie (saturation) oraz jasność (value), dając źródło dla tzw. przestrzeni barw HSV. Według tego modelu wszystkie barwy wywodzą się ze światła białego, gdzie część widma jest odbita a pozostała pochłonięta przez oświetlane przedmioty.

Wszystkie te idee wsparte zostały wreszcie przez matematykę. W 1853 roku Hermann Grassmann wydał pracę pt. *Teoria mieszania kolorów*. Zaciekawiony zależnościami mieszania kolorów typu:

$$\text{czerwony} + \text{niebieski} = \text{fioletowy}$$

stwierdził, że gdy dysponujemy takim równaniem, można do każdej jego strony dodać kolor i wciąż uzyskamy prawdziwą równość! Np. do obydwu stron równości wyżej można dodać żółty i dostać:

$$\text{czerwony} + \text{niebieski} + \text{żółty} = \text{fioletowy} + \text{żółty}$$

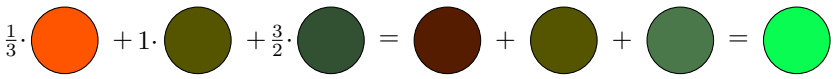
Prawo to, nam się kojarzące z łącznością, zwane jest w kolorymetrii **trzecim prawem Grassmanna**. Dlaczego to jest takie ważne? Grassmann postulował, że barwa mieszaniny zależy jedynie od barw podstawowych jej składników, a nie od ich składu widmowego oraz, że można wybierać różne „bazy kolorów podstawowych", za pomocą których można uzyskać dowolne kolory. Oto **pierwsze prawo Grassmanna**.

Każda dowolnie wybrana barwa może być określona za pomocą trzech liniowo niezależnych barw.
Inaczej: każde cztery barwy są liniowo zależne.[8]

Czytelnik ewidentnie zobaczy w tych sformułowaniach algebrę liniową. Model Grassmanna został oczywiście poddany wielu ulepszeniom i krytyce, ale na nasz użytek może dobrze opisywać intuicje liniowej niezależności i kombinacji liniowych. Można zastosować intuicyjne analogie uznając, że chcemy uzyskiwać kolory jako kombinacje liniowe innych (niekoniecznie podstawowych, tak jak w przestrzeni liniowej są różne bazy). Oto przykład takiej sytuacji w modelu RGB.

Rozważmy zbiór liniowych kombinacji wektorów $\text{lin}(v_1 = (255, 85, 0), v_2 = (85, 85, 0), v_3 = (50, 80, 50))$.

Oto przykład kombinacji liniowej $\frac{1}{3}v_1 + v_2 + \frac{3}{2}v_3$ tych kolorów:



Czy trzy wymienione kolory tworzą bazę przestrzeni kolorów? Gdybyśmy mogli stosować ujemne nasycenie (a to się robi) – to owszem tak. Są to trzy wektory liniowo niezależne (żaden nie jest kombinacją pozostałych) i gdybyśmy mieli możliwość stosowania kombinacji liniowych z ujemnymi współczynnikami, wówczas wygenerowalibyśmy z nich dowolne kolory. Wszystko to jedynie zbiór intuicji – ważnych jednak z praktycznego punktu widzenia. Dziś teoria kolorów jest znacznie bardziej skomplikowana i stanowi osobną dziedzinę wiedzy (fizyki i chemii). Warto pamiętać, że ma ona duży styk także z algebrą liniową.

[8] Prawa te zostaly ustalone dla części środkowej siatkówki oka człowieka. Prawo drugie i trzecie dotyczy również zwierząt. Prawo pierwsze zachowuje słuszność tylko w postaci uogólnionej, ponieważ maksymalna liczba barw liniowo niezależnych może być większa lub mniejsza od trzech. Istnieją ludzie i zwierzęta, dla których rejestrowana maksymalna liczba liniowo niezależnych barw wynosi dwa lub jeden, co powoduje, że niektórzy widzą więcej barw, inni mniej, a jeszcze inni rejestrują tylko szarości. Źródło: `https://slownikzprepressu.weebly.com/grassmanna-prawa.html`.

Rozdział 6

Wymiar przestrzeni liniowej

6.1 Wykład szósty

Na ostatnim wykładzie wprowadzone zostało pojęcie liniowej niezależności układu wektorów w przestrzeni liniowej. Dla przestrzeni liniowych, które można przedstawić w postaci

$$V = \operatorname{lin}(\beta_1, \ldots, \beta_n), \qquad (\heartsuit)$$

dla pewnego układu $\beta_1, \ldots, \beta_n$ wektorów w V pokazaliśmy, że z układu tego wybrać można podukład liniowo niezależny, który dalej rozpina V. Innymi słowy, z dowolnego układu rozpinającego podprzestrzeń można wybrać bazę. Można jednak zapytać: jaka będzie zależność pomiędzy bazami wybieranymi z różnych układów rozpinających przestrzeń liniową V? Dziś pokażemy[1], że dla przestrzeni rozpiętych na skończonym układzie wektorów każde dwie bazy są równoliczne. Pozwoli to na określenie pojęcia wymiaru przestrzeni liniowej oraz wyróżnienia klasy przestrzeni liniowych nazywanych skończenie wymiarowymi. Powiemy też kilka słów o przestrzeniach nieskończonego wymiaru, czyli takich, których nie można przedstawić w postaci ($\heartsuit$). Przy tej okazji uogólnimy pojęcia układu rozpinającego podprzestrzeń, układu liniowo niezależnego i bazy na dowolne (niekoniecznie skończone) układy wektorów.

Podczas wykładu powoływać się będziemy często na Obserwacje 5.2 oraz 5.3 z poprzedniego wykładu. Pierwsza z nich mówi, że układ wektorów jest liniowo zależny wtedy i tylko wtedy, gdy pewien jego element jest kombinacją liniową pozostałych. Druga natomiast mówi, że po dołączeniu do układu liniowo niezależnego dodatkowego wektora uzyskamy układ, który jest liniowo niezależny wtedy i tylko wtedy, gdy dołączony wektor nie jest kombinacją liniową wyjściowego liniowo niezależnego układu wektorów.

Czas na rezultat prowadzący w kierunku twierdzenia o istnieniu wymiaru przestrzeni liniowej.

Twierdzenie 6.1: Steinitz (1910)

Jeśli układ wektorów $\alpha_1, \ldots, \alpha_k$ leżących w przestrzeni

$$V = \operatorname{lin}(\beta_1, \ldots, \beta_m)$$

jest liniowo niezależny, to:

(a) $k \leqslant m$,

(b) z układu $\beta_1, \ldots, \beta_m$ można wybrać taki podukład $\beta_{i_1}, \ldots, \beta_{i_{m-k}}$, że:

$$\operatorname{lin}(\beta_1, \ldots, \beta_m) = \operatorname{lin}(\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_{m-k}}).$$

Twierdzenie powyższe mówi o tym, że liczność rozpinających układów liniowo niezależnych respektuje porządek wyznaczony przez inkluzję. Jeśli układ k wektorów rozpina przestrzeń liniową, to nie można w niej „zmieścić" układu złożonego z więcej niż k liniowo niezależnych wektorów (punkt (a)). Co więcej, odpowiedni fragment dowolnego układu rozpinającego daną przestrzeń liniową można zastąpić dowolnym równolicznym układem liniowo niezależnym zawartym w tej przestrzeni (punkt (b)). Stąd też rezultat ten nazywa się w wielu źródłach **twierdzeniem o wymianie**.

[1] Ostatnia aktualizacja: 16.11.2022 r.

Dowód. Pokazujemy najpierw punkt (a). Dla $m = 1$ twierdzenie jest oczywiste, bo jeśli niezerowe $\alpha_1, \ldots, \alpha_k$ leżą w $\lin(\beta_1)$, to każdy z nich jest postaci $a_i \cdot \beta_1$, gdzie $a_1, \ldots, a_k$ to niezerowe elementy K, zaś układ

$$a_1\beta_1, a_2\beta_1, \ldots, a_k\beta_1$$

jest liniowo niezależny tylko dla $k = 1$.

Przechodzimy do kroku indukcyjnego. Mamy układ liniowo niezależny $\alpha_1, \ldots, \alpha_k$ w przestrzeni liniowej $V = \lin(\beta_1, \ldots, \beta_m)$. Po ewentualnym przenumerowaniu[2] $\beta_1, \ldots, \beta_m$ weźmy takie a_{ij}, dla $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant m$, że $a_{11} \neq 0$ oraz:

$$\begin{aligned} \alpha_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \ldots + a_{1m}\beta_m, \\ &\ldots \\ \alpha_k &= a_{k1}\beta_1 + a_{k2}\beta_2 + \ldots + a_{km}\beta_m. \end{aligned}$$

Teraz **poprawiamy** podukład $\{\alpha_2, \ldots, \alpha_k\}$ układu $\{\alpha_1, \ldots, \alpha_k\}$ do takiego układu $\{\gamma_2, \ldots, \gamma_k\}$, który jest rozpięty tylko przez $\beta_2, \ldots, \beta_m$. Dokładniej, określamy dla $i = 2, 3, \ldots, k$ określamy układ wektorów $\gamma_2, \ldots, \gamma_k \subseteq \lin(\beta_2, \ldots, \beta_m)$:

$$\gamma_i = \alpha_i - \frac{a_{i1}}{a_{11}}\alpha_1 = \underbrace{\boldsymbol{a_{i1}}\beta_1 + a_{i2}\beta_2 + \ldots + a_{im}\beta_m}_{\alpha_i} - \frac{\boldsymbol{a_{i1}}}{\boldsymbol{a_{11}}} \underbrace{(\boldsymbol{a_{11}}\beta_1 + a_{12}\beta_2 + \ldots + a_{1m}\beta_m)}_{\alpha_1}.$$

Nowy układ składa się z $k-1$ wektorów i każdy jest rzeczywiście kombinacją jedynie wektorów postaci $\beta_2, \ldots, \beta_m$ (po powyższym rozpisaniu γ_i przy β_1 stoi 0). Przekonajmy się natomiast, że wektory $\gamma_2, \ldots, \gamma_k$ są liniowo niezależne. Istotnie, gdybyśmy dla pewnych $c_2, \ldots, c_k \in K$, nie wszystkich równych 0, mieli:

$$c_2\gamma_2 + \ldots + c_k\gamma_k = 0 \Leftrightarrow c_2\left(\alpha_2 - \frac{a_{21}}{a_{11}}\alpha_1\right) + c_3\left(\alpha_3 - \frac{a_{31}}{a_{11}}\alpha_1\right) + \ldots + c_k\left(\alpha_k - \frac{a_{k1}}{a_{11}}\alpha_1\right) = 0,$$

czyli równoważnie:

$$-\frac{c_2a_{21} + c_3a_{31} + \ldots + c_ka_{k1}}{a_{11}}\alpha_1 + c_2\alpha_2 + \ldots + c_k\alpha_k = 0.$$

Skoro jednak układ $c_2 = 0, \ldots, c_k = 0$, sprzeczność. Zatem układ $\gamma_2, \ldots, \gamma_k$ jest liniowo niezależny.

Podsumujmy: układ $k-1$ liniowo niezależnych wektorów $\gamma_2, \ldots, \gamma_k$ zawarty jest w przestrzeni rozpiętej przez $m-1$ wektorów postaci $\lin(\beta_2, \ldots, \beta_m)$. Z założenia indukcyjnego mamy więc $k - 1 \leqslant m - 1$. A zatem $k \leqslant m$.

Dowodzimy punkt (b). Mamy liniowo niezależne $\alpha_1, \ldots, \alpha_k \in \lin(\beta_1, \ldots, \beta_m)$. Niech $\beta_{i_1}, \ldots, \beta_{i_s}$ będzie najdłuższym podukładem w $\beta_1, \ldots, \beta_m$ zawierającym $\alpha_1, \ldots, \alpha_k$, że układ

$$\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s}$$

jest liniowo niezależny (na mocy punktu (a) takie s istnieje). W szczególności, dla każdego $1 \leqslant j \leqslant m$, dłuższy układ wektorów

$$\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s}, \beta_j$$

jest już liniowo zależny. Na mocy Obserwacji 5.3 mamy zatem:

$$\beta_j \in \lin(\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s}).$$

W szczególności $\lin(\beta_1, \ldots, \beta_m) \subseteq \lin(\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s})$. Oczywiście wszystkie α_i są kombinacjami liniowymi β_j więc $\lin(\beta_1, \ldots, \beta_m) = \lin(\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s})$. Z udowodnionego już punktu (a) wynika, że

$$k + s \leqslant m,$$

a więc

$$s \leqslant m - k.$$

Stąd dołączając do układu $\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s}$ dowolne $m - k - s$ wektorów $\gamma_1, \ldots, \gamma_{m-k-s}$ spośród β_i, dla $i \neq i_1, \ldots, i_s$, otrzymujemy układ spełniający

$$\lin(\beta_1, \ldots, \beta_m) = \lin(\alpha_1, \ldots, \alpha_k, \beta_{i_1}, \ldots, \beta_{i_s}, \gamma_1, \ldots, \gamma_{m-k-s}).$$

□

[2] Czy Czytelnik widzi, dlaczego? W układzie liniowo niezależnym żaden wektor nie może być zerowy, czyli $\alpha_1 \neq 0$.

Wniosek 6.1: O liczbie wektorów rozpinających podprzestrzeń

(a) Jeśli W jest podprzestrzenią przestrzeni $V = \lin(\beta_1, \ldots, \beta_m)$, to w W istnieje taki układ liniowo niezależny $\alpha_1, \ldots, \alpha_k$, $k \leqslant m$, że $W = \lin(\alpha_1, \ldots, \alpha_k)$.

(b) Jeśli $\lin(\alpha_1, \ldots, \alpha_k) = \lin(\alpha'_1, \ldots, \alpha'_l)$ i układy $\alpha_1, \ldots, \alpha_k$ oraz $\alpha'_1, \ldots, \alpha'_l$ są liniowo niezależne, to $k = l$.

Dowód. Weźmy najdłuższy liniowo niezależny układ w W (to ma sens, bo wszystkie mają długość $\leqslant m$). Niech to będzie układ $\alpha_1, \ldots, \alpha_k$. Pokażemy, że:

$$W = \lin(\alpha_1, \ldots, \alpha_k).$$

Dowodzimy, że mają miejsce dwie inkluzje. Jedna z nich: $\lin(\alpha_1, \ldots, \alpha_k) \subseteq W$, jest oczywista, bo skoro wektory $\alpha_1, \ldots, \alpha_k$ należą do W, to każda ich kombinacja liniowa też (bo W to podprzestrzeń). Dowodzimy teraz, że: $\lin(\alpha_1, \ldots, \alpha_k) \supseteq W$. Weźmy dowolny wektor $\alpha \in W$. Układ $\alpha_1, \ldots, \alpha_k, \alpha$ jest dłuższy niż układ $\alpha_1, \ldots, \alpha_k$, więc jest liniowo zależny. Ponownie korzystamy z implikacji $(b) \Rightarrow (a)$ w dowodzie Obserwacji 5.3, otrzymując $\alpha \in \lin(\alpha_1, \ldots, \alpha_k)$. Wobec dowolności α otrzymujemy drugą inkluzję.

Dowód (b). Skoro $\lin(\alpha_1, \ldots, \alpha_k) = \lin(\alpha'_1, \ldots, \alpha'_l)$ i układy $\alpha_1, \ldots, \alpha_k$ oraz $\alpha'_1, \ldots, \alpha'_l$ są liniowo niezależne, to wobec $\alpha_1, \ldots, \alpha_k \in \lin(\alpha'_1, \ldots, \alpha'_l)$ mamy $k \leqslant l$. Z drugiej strony mamy przecież także symetryczne należenie: $\alpha'_1, \ldots, \alpha'_l \in \lin(\alpha_1, \ldots, \alpha_k)$, czyli z twierdzenia Steinitza: $l \leqslant k$. A zatem $k = l$. □

Twierdzenie 6.2

Jeśli przestrzeń liniowa V posiada bazę złożoną z n wektorów, to każda baza przestrzeni V jest złożona z n wektorów.

Dowód. Jeśli $\alpha_1, \ldots, \alpha_k$ oraz $\alpha'_1, \ldots, \alpha'_l$ są bazami przestrzeni V, to $\lin(\alpha_1, \ldots, \alpha_k) = \lin(\alpha'_1, \ldots, \alpha'_l) = V$. Układy te są liniowo niezależne, a zatem na mocy Wniosku mamy $k = l$. □

Definicja 6.1: Wymiar przestrzeni liniowej

Mówimy, że przestrzeń liniowa V jest n WYMIAROWA, jeśli V posiada bazę złożoną z n wektorów. Piszemy wówczas

$$\dim V = n$$

i liczbę n nazywamy WYMIAREM PRZESTRZENI V. Przyjmujemy też $\dim\{0\} = 0$.

Mówimy, że przestrzeń liniowa V jest SKOŃCZENIE WYMIAROWA, jeśli V jest n wymiarowa dla pewnego $n \in \mathbb{N} \cup \{0\}$. Jeśli V nie jest skończenie wymiarowa, to V nazywamy NIESKOŃCZENIE WYMIAROWĄ i piszemy $\dim V = \infty$.

Podajmy kilka ważnych przykładów.

- Oczywiście $\dim K^n = n$, o czym świadczy choćby baza standardowa.
- Jeśli $V = M_{m\times n}(K)$, to baza przestrzeni V złożona jest (na przykład) z macierzy E_{ij}, które poza wyrazem w i-tym wierszu i j-tej kolumnie, równym 1, mają same wyrazy zerowe. Nietrudno zatem widzieć, że $\dim M_{m\times n}(K) = m \cdot n$.
- Niech $V = (x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1 + 2x_2 - x_3 = 0\}$. Wówczas $\dim(V) = 2$, bo V ma bazę postaci $\{(-2, 1, 0), (1, 0, 1)\}$.
- Dla każdego n układ n wektorów $x, x^2, \ldots, x^n$ przestrzeni $K[x]$ jest liniowo niezależny. A zatem przestrzeń ta nie może być skończenie wymiarowa. Gdyby jej wymiar wynosił k, to na mocy twierdzenia Steinitza każdy układ liniowo niezależny w $K[x]$ musiałby liczyć nie więcej niż k wektorów. A zatem $\dim K[x] = \infty$. Podobnie nietrudno pokazać, że $\dim K^\infty = \infty$.

Wyznaczając wymiar podprzestrzeni rozpiętych na układach wektorów w K^n korzystać będziemy często z Obserwacji 4.7 oraz 5.1. Zobaczmy przykład. Dana jest podprzestrzeń

$$V = \mathrm{lin}((1,2,0,1,0),(0,1,1,1,1),(2,2,3,0,3),(1,3,1,2,1)) \subseteq K^5.$$

Wykonując operacje elementarne na wierszach macierzy rozmiaru 4×5, której wiersze stanowią układy rozpinające powyższą podprzestrzeń mamy: Przekształcamy macierz układu wektorów rozpinających V przy pomocy operacji elementarnych na wierszach:

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 0 & 3 \\ 1 & 3 & 1 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & -2 & 3 & -2 & 3 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 5 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & -1 & -2 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 5 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Na mocy Obserwacji 4.7 mamy

$$V = \mathrm{lin}((1,0,-2,-1,-2),(0,1,1,1,1),(0,0,5,0,5)).$$

Jeśli K jest np. ciałem $\mathbb{Q}$, to układ ten jest na mocy Obserwacji 5.1 liniowo niezależny, a zatem jest bazą V i $\dim V = 3$. Jeśli zaś założymy, że $K = \mathbb{Z}_5$, to $V = \mathrm{lin}((1,0,3,-1,3),(0,1,1,1,1))$ i $\dim V = 2$.

Wniosek 6.2

Podprzestrzeń przestrzeni rozpiętej na skończonym układzie wektorów jest skończenie wymiarowa. Jeśli W jest podprzestrzenią V i $\dim V = n$, to $\dim W \leqslant n$.

Dowód. Niech $W \subseteq V = \mathrm{lin}(\beta_1, \ldots, \beta_m)$. Wówczas $W = \mathrm{lin}(\alpha_1, \ldots, \alpha_k)$ dla pewnego układu liniowo niezależnego $\alpha_1, \ldots, \alpha_k$ na mocy wniosku. Układ $\alpha_1, \ldots, \alpha_k$ jest bazą przestrzeni W, więc W jest skończenie wymiarowa. Jeśli $\dim V = n$, to dla każdej bazy $\gamma_1, \ldots, \gamma_n$ przestrzeni V układ $\alpha_1, \ldots, \alpha_k$ jest zawartym w $\mathrm{lin}(\gamma_1, \ldots, \gamma_n)$, a więc $k \leqslant n$, z twierdzenia Steinitza. □

Poniższe wnioski wynika w sposób oczywisty z przedstawionych wyżej rozumowań

Wniosek 6.3

Niech V będzie przestrzenią skończenie wymiarową. Wówczas:

(a) Każdy liniowo niezależny układ wektorów V można, dołączając pewną liczbę wektorów, uzupełnić do bazy przestrzeni V,

(b) Z każdego układu $\beta_1, \ldots, \beta_m$ rozpinającego V można wybrać bazę podprzestrzeni V,

(c) Jeśli $\dim(V) = k$ i $\alpha_1, \ldots, \alpha_k$ jest liniowo niezależnym układem wektorów przestrzeni V, to $\alpha_1, \ldots, \alpha_k$ jest bazą przestrzeni K.

(d) Niech W będzie podprzestrzenią przestrzeni V. Wówczas $\dim W \leqslant \dim V$. Przy tym jeśli zachodzi $\dim W = \dim V$, to $W = V$.

Dowód. Ad (a). Niech $\alpha_1, \ldots, \alpha_k$ będzie układem liniowo niezależnym wektorów przestrzeni V. Wówczas najdłuższy układ liniowo niezależny postaci $\alpha_1, \ldots, \alpha_k, \gamma_1, \ldots, \gamma_s$ jest bazą przestrzeni V.

Ad (b). Najdłuższy spośród liniowo niezależnych podukładów układu $\beta_1, \ldots, \beta_m$ jest bazą przestrzeni V.

Ad (c). Układ $\alpha_1, \ldots, \alpha_k$ jest maksymalnym układem liniowo niezależnym w V, więc jest bazą.

Ad (d). Na mocy wcześniejszego wniosku $\dim W \leqslant \dim V$. Jeśli $\dim W = \dim V$, to każda baza przestrzeni W jest też bazą przestrzeni V, więc $W = V$. □

Na koniec kilka uwag i przykładów dotyczących liniowej niezależności przestrzeni, których nie można przedstawić jako $\mathrm{lin}(\alpha_1, \ldots, \alpha_n)$, a więc przestrzeni nieskończenie wymiarowych.

Definicja 6.2

Niech $X = \{\alpha_t\}_{t\in T}$ będzie dowolnym układem wektorów przestrzeni V. Wówczas przez $\lin(X)$ oznaczamy zbiór wszystkich kombinacji liniowych SKOŃCZONYCH PODUKŁADÓW układu X. To znaczy:

$$\beta \in \lin(X) \iff \beta = \sum_{i=1}^{k} a_i \alpha_{t_i}, \text{ dla pewnych } a_1, \ldots, a_k \in K, \alpha_{t_1}, \ldots, \alpha_{t_k} \in X.$$

Jeśli $V = \lin(X)$ to mówimy, że układ X ROZPINA V i przestrzeń V JEST ROZPIĘTA na X.
Dla układu pustego $X = \emptyset$ przyjmujemy $\lin(X) = \{0\}$.

Przykłady:

- $V = \lin(V)$,
- $K[x] = \lin(1, x, x^2, x^3, \ldots)$,
- problem: „wypisać" najmniejszy taki zbiór X, by $\mathbb{R} = \lin(X)$, gdzie $\mathbb{R}$ – przestrzeń nad $\mathbb{Q}$.

Definicja 6.3

Układ $X = \{\alpha_i\}_{i\in T}$ wektorów przestrzeni V nazywamy LINIOWO NIEZALEŻNYM, jeśli każdy jego SKOŃCZONY PODUKŁAD jest liniowo niezależny.

Przykłady (większość to całkiem ciekawe ćwiczenia):

(a) układ $\{1, x, x^2, x^3, \ldots, \}$ jest liniowo niezależny w $K[x]$,

(b) układ ciągów $a_1 = (1, 0, 0, \ldots), a_2 = (0, 1, 0, \ldots), a_3 = (0, 0, 1, \ldots), \ldots$ jest liniowo niezależny w K^∞,

(c) układ ciągów $\{(1, t, t^2, t^3, \ldots), t \in (0, 1)\}$ jest liniowo niezależny w $\mathbb{R}^\infty$,

(d) układ $\{\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \sqrt{11}, \ldots\} = \{\sqrt{p}, p \in P\}$, gdzie P - zbiór liczb pierwszych, jest liniowo niezależny w przestrzeni $\mathbb{R}$ nad ciałem $\mathbb{Q}$,

(e) układ $\{\sin(x), \sin^2(x), \sin^3(x), \ldots\} = \{\sin(x)^n, n \in \mathbb{N}_+\}$ jest liniowo niezależny w przestrzeni $F(\mathbb{R}, \mathbb{R})$.

Definicja 6.4

Układ $X = \{\alpha_i\}_{i\in T}$ wektorów przestrzeni V nazywamy BAZĄ, jeśli jest on liniowo niezależny oraz $\lin(X) = V$.

Warto odnotować, że z układów (a)-(e) tylko układ (a) jest bazą $K[x]$. Rzeczywiście, dowolny wielomian jest kombinacją liniową skończenie wielu elementów z układu

$$\{1, x, x^2, x^3, \ldots, \},$$

a układ ten jest liniowo niezależny: jeśli $a_1 \cdot x^{i_1} + a_2 \cdot x^{i_2} + \ldots + a_n x^{i_n}$ jest wielomianem zerowym, to oczywiście $a_1 = \ldots = a_n$. Pozostałe układy są wprawdzie liniowo niezależne, ale nie stanowią bazy.

Przyjrzyjmy się bliżej przykładowi (b). Ciągu

$$(1, 1, \ldots) \in K^\infty,$$

którego wszystkie wyrazy są równe 1 nie można przedstawić jako kombinacji liniowej elementów postaci a_i. Zasadniczy problem powyższego przykładu polega nie na wskazywaniu dużych układów liniowo niezależnych w K^∞, ale na tym, że te duże układy wskazane „wprost" są za małe by rozpinać całe K^∞. Nawet jeśli dorzucimy do układu w (b) wektor złożony z samych jedynek, nie będziemy mieli bazy. Można znajdować kolejne wektory, które nie należą do przestrzeni rozpiętej przez ten poszerzony układ. Nawet gdybyśmy połączyli rodzinę w (b) z rodziną w (c) (dla $K = \mathbb{R}$), wciąż nie uzyskamy bazy.

W ramach wykładu nie pokazujemy dowodu twierdzenia mówiącego, że każda przestrzeń liniowa (nieskończonego wymiaru) posiada bazę. Zagadnienie to rozpatrzymy jednak w uzupełnieniu.

6.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Znajdowanie bazy i wymiaru przestrzeni rozwiązań układu równań liniowych)
   Znajdź bazę i wymiar przestrzeni rozwiązań układu równań liniowych w $\mathbb{R}^4$ postaci
   $$\begin{cases} x_1 + x_2 + 3x_3 + 2x_4 & = 0 \\ 2x_1 + x_2 + 7x_3 + 5x_4 & = 0 \\ 5x_1 + 4x_2 + 16x_3 + 11x_4 & = 0 \end{cases}.$$

2. Niech $X = \mathbb{Z}_{13}^4$ oraz
   $$U = \{(x_1, x_2, x_3, x_4) \in X : x_1 + x_2 + x_4 = 0 \text{ oraz } x_3 - x_4 = 0\}.$$
   Wyznacz bazę i wymiar przestrzeni U. Ile wektorów ma ta przestrzeń liniowa?

3. (♠ Znajdowanie bazy i wymiaru przestrzeni rozpiętej na układzie wektorów)
   Niech W będzie podprzestrzenią w $\mathbb{R}^5$ postaci:
   $$W = \operatorname{lin}((10, 3, 9 + s, 1, 2 - s), (4, 1, 6, 1, 1), (2, 1, -1, -1, -2)).$$
   Znajdź $\dim W$ w zależności od $s \in \mathbb{R}$.

4. (♠ Dopełnianie układu wektorów do bazy)
   Dopełnij wektory $v_1 = (1, 2, 3, -2, -4)$ oraz $v_2 = (6, 4, -5, -4, -1)$ do bazy przestrzeni rozwiązań układu równań liniowych zadanego macierzą
   $$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 & 1 & 0 \end{array}\right].$$

5. Znajdź bazę i wymiar przestrzeni liniowej
   $$V = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathrm{M}_2(\mathbb{C}) : a_{11} + a_{12} + a_{21} + a_{22} = 0 \right\}$$
   nad ciałem $\mathbb{C}$. Podaj współrzędne macierzy
   $$A = \begin{bmatrix} 2 + i & -1 \\ -2 & 1 - i \end{bmatrix} \in V$$
   w znalezionej bazie.

6. Wyznacz bazę i wymiar podprzestrzeni przestrzeni $F(\mathbb{R}, \mathbb{R})$ złożonej z funkcji wielomianowych f spełniających warunek $f(1) = f(2) = 0$.

7. W przestrzeni $\mathbb{C}^\infty$ rozważmy wszystkie ciągi $(x_n) = (x_1, x_2, x_3, \ldots)$ spełniające dla każdego $n \geqslant 1$ warunek
   $$x_{n+2} = 2x_{n+1} - 2x_n$$
   Wykaż, że ciągi te tworzą podprzestrzeń $\mathbb{C}^\infty$, wyznacz jej wymiar oraz wskaż bazę tej podprzestrzeni złożoną z ciągów geometrycznych.

8. Dana jest przestrzeń liniowa V wymiaru n oraz jej podprzestrzenie W_1, W_2, przy czym
   $$0 < \dim W_1, \dim W_2 < n.$$
   Wykaż, że istnieje element $\alpha \in V$, taki że $\alpha \notin W_1$ oraz $\alpha \notin W_2$. Wykaż dalej, że istnieje baza przestrzeni V, taka że żaden z jej elementów nie jest zawarty ani w W_1, ani w W_2.

9. Wykaż, że $\dim V = \infty$ wtedy i tylko wtedy, gdy istnieje ciąg wektorów $v_1, v_2, \ldots$ przestrzeni V taki, że dla każdego n układ $\{v_1, \ldots, v_n\}$ jest liniowo niezależny.

10. Niech $\dim V = \infty$ oraz niech $\mathcal{A} = \{\alpha_i\}_{i \in \mathbb{N}}$ będzie bazą przestrzeni V. Dla każdej liczby $n \in \mathbb{N}$ niech
    $$V_n = \operatorname{lin}(\alpha_1, \ldots, \alpha_n).$$
    Wykaż, że dla każdej skończenie wymiarowej podprzestrzeni $W \subset V$ istnieje $n \in \mathbb{N}$, dla którego W jest podprzestrzenią przestrzeni V_n.

6.3 Uzupełnienie. Każda przestrzeń liniowa ma bazę

Celem tego uzupełnienia jest przybliżenie Czytelnikowi zagadnienia istnienia bazy w dowolnej przestrzeni liniowej. O genezie tego problemu opowiemy innym razem – dotyczy ona istotnych obiektów, zwanych bazami Hamela. Twierdzenie, które chcemy udowodnić pochodzi od Haussdorffa i w pełnej ogólności pokazane zostało w 1932 roku. Rozważania prowadzimy na bazie tekstu[3] K. Conrada.

Twierdzenie 6.3: Haussdorff, 1932

Każda przestrzeń liniowa V nad ciałem K ma bazę.

Podstawowa idea polega na konstrukcji bazy jako maksymalnego liniowo niezależnego zbioru, nawiązując w ten sposób do równoważności stwierdzonej przez nas dla baz przestrzeni skończonego wymiaru. Aby zrozumieć dobrze o co chodzi musimy powiedzieć kilka słów o porządkach i elementach maksymalnych. Dlaczego to podejście zadziała i jakie są wyzwania?

Klasycznym przykładem, który obrazuje złożoność rozważanego przez nas problemu jest ciało liczb rzeczywistych traktowane jako przestrzeń liniowa nad ciałem liczb wymiernych. Jest to oczywiście przestrzeń nieskończonego wymiaru. Przykładowym nieskończonym układem liniowo niezależnym jest zbiór pierwiastków ze wszystkich liczb pierwszych $\{\sqrt{2}, \sqrt{3}, \sqrt{5}, \dots, \}$. Nietrudno się o tym przekonać pokazując, że $\sqrt{p}$ nie jest kombinacją liniową elementów $\sqrt{q}$, dla $q < p$, tzn. nie istnieją liczby wymierne a_q takie, że:

$$\sum_{q<p, q\in\mathbb{P}} a_q \cdot \sqrt{q} = \sqrt{p}.$$

Nietrudno się jednak przekonać, że wypisany zbiór liniowo niezależny nie jest bazą $\mathbb{R}$ nad $\mathbb{Q}$. Żadna liczba postaci $\sqrt[3]{p}$ nie jest kombinacją liniową (o współczynnikach w $\mathbb{Q}$) pierwiastków z liczb pierwszych. A nawet gdyby rozważać zbiór pierwiastków dowolnego stopnia ze wszystkich liczb pierwszych – to również nie jest maksymalny zbiór liniowo niezależny – nie należy do niego choćby liczba π czy e. Widzimy więc, że wypisanie maksymalnego zbioru liniowo niezależnego „wprost" jest w zasadzie niemożliwe. W dodatku zobaczymy przykład nieprzeliczalnego podzbioru $\mathbb{R}$, który jest liniowo niezależny nad $\mathbb{Q}$ i który też nie jest bazą. Co więcej, wydaje się, że możemy wystartować z rozłącznych nieskończonych zbiorów liniowo niezależnych i „nadbudowywać" na nich różne większe zbiory liniowo niezależne. Jak owe „nadbudowane" zbiory miałyby się do potencjalnej bazy? Potrzebne są pewne narzędzia, by doprecyzować te problemy.

Udowodnimy następujący rezultat, korzystając z rezultatów ze wstępu do matematyki.

Twierdzenie 6.4

Niech V będzie niezerową przestrzenią liniową i niech $\mathcal{S}$ będzie zbiorem złożonym z liniowo niezależnych podzbiorów w V. Wówczas $\mathcal{S}$ zawiera PODZBIÓR MAKSYMALNY ze względu na inkluzję – to znaczy taki podzbiór $M \in \mathcal{S}$, że nie istnieje zbiór $N \in \mathcal{S}$, który zawiera M jako podzbiór właściwy. Zbiór M jest bazą przestrzeni V.

Kluczowym elementem jest kwestia istnienia owego maksymalnego zbioru. Jeśli taki zbiór istnieje, to pokazanie, że jest on bazą jest niemal identyczne do przypadku skończenie wymiarowego. Otóż, jeśli M jest maksymalnym elementem $\mathcal{S}$, to rozpatrując podprzestrzeń $W = \mathrm{lin}(M)$ pokażemy, że $W = V$, co pokaże, że M jest bazą. Gdyby M nie rozpinał V, wówczas $W \neq V$ i można wskazać wektor $v \in V$ taki, ze $v \notin W$. W szczególności M jest podzbiorem właściwym zbioru $M \cup \{v\}$. Pokażemy jednak, że $M \cup \{v\}$ jest układem liniowo niezależnym, co przeczyć będzie maksymalności M postulowanej w tezie twierdzenia.

Aby pokazać, że $M \cup \{v\}$ jest liniowo niezależny, załóżmy przeciwnie, że dla pewnego skończonego podzbioru $\{v_1, \dots, v_k\}$ zbioru $M \cup \{v\}$ mamy

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k = 0,$$

[3] https://kconrad.math.uconn.edu/blurbs/zorn1.pdf

przy czym $c_i \in K$ nie są wszystkie zerowe. Skoro elementy M tworzą układ liniowo niezależny, to także układ $\{v_1, \ldots, v_k\}$ jest liniowo niezależny i koniecznie jednym z elementów v_i musi być v. Możemy więc przenumerować elementy układu $v_1, \ldots, v_k$ tak, by założyć $v_k = v$. Co więcej, musimy mieć także $c_k \neq 0$. Mamy też $k \geqslant 2$, bo inaczej $c_1 v = 0$, co jest niemożliwe, bo $v \neq 0$ oraz $c_1 \neq 0$. A zatem mamy:

$$v = -\frac{c_1}{c_k} v_1 - \frac{c_2}{c_k} v_2 - \ldots - \frac{c_{k-1}}{c_k} v_{k-1}.$$

A zatem $v \in \lin(v_1, \ldots, v_{k-1}) \subseteq W$. Ale zakładaliśmy, że $v \notin W$, więc $M \cup \{v\}$ okazuje się zbiorem liniowo niezależnym, co przeczy maksymalności M w zbiorze $\mathcal{S}$. A zatem $W = \lin(M) = V$ i M jest bazą V.

Dlaczego istnieje element maksymalny w $\mathcal{S}$? Nie unikniemy w tym miejscu użycia rezultatu, znanego jako lemat Kuratowskiego-Zorna, stanowiącego zwieńczenie podstawowego kursu ze wstępu do matematyki.

Twierdzenie 6.5: Lemat Kuratowskiego-Zorna

Niech $\mathcal{S}$ będzie niepustym zbiorem częściowo uporządkowanym. Jeśli każdy liniowo uporządkowany podzbiór (łańcuch) w $\mathcal{S}$ ma ograniczenie górne w $\mathcal{S}$, wówczas $\mathcal{S}$ zawiera element maksymalny.

Aby zrozumieć lemat Kuratowskiego-Zorna, potrzebujemy czterech pojęć: częściowego porządku w zbiorze, zbioru liniowo uporządkowanego, ograniczenia górnego oraz elementu maksymalnego.

Definicja 6.5: Relacja częściowego porządku

Mówimy, że RELACJA $\leqslant$ na zbiorze $\mathcal{S}$ (czyli podzbiór zbioru $\mathcal{S} \times \mathcal{S}$) jest CZĘŚCIOWYM PORZĄDKIEM, jeśli spełnione są następujące warunki:

- zwrotność — dla każdego $s \in \mathcal{S}$ mamy $s \leqslant s$,
- antysymetryczność — dla każdych $s, s' \in \mathcal{S}$ jeśli $s \leqslant s'$ oraz $s' \leqslant s$, to $s = s'$,
- przechodniość — dla każdych $s, s', s'' \in \mathcal{S}$ jeśli $s \leqslant s'$ oraz $s' \leqslant s''$, to $s \leqslant s''$.

Mówimy, że relacja $\leqslant$ jest LINIOWYM PORZĄDKIEM, jeśli dla każdych $s, s' \in \mathcal{S}$ mamy $s \leqslant s'$ lub $s' \leqslant s$. Jeśli $\leqslant$ jest częściowym porządkiem w zbiorze $\mathcal{S}$, a po ograniczeniu do niepustego podzbioru $\mathcal{T} \subseteq \mathcal{S}$ jest ona porządkiem liniowym, wówczas podzbiór $\mathcal{T}$ nazywamy ŁAŃCUCHEM w $\mathcal{S}$.

Zobaczmy trzy przykłady, z których trzeci jest kluczowy dla naszych rozważań.

- Porządek liniowy $\leqslant$ w $\mathbb{R}$ – czyli zwykła relacja nierówności pomiędzy liczbami rzeczywistymi jest porządkiem liniowym.
- W zbiorze dodatnich liczb całkowitych $\mathbb{Z}_+$ wprowadzamy relację częściowego porządku $a \leqslant b$ postaci $a \,|\, b$, czyli RELACJĘ PODZIELNOŚCI. Oczywiście jest to częściowy porządek, ale nie jest to porządek liniowy, bowiem (choćby) 2 nie dzieli 3, ani 3 nie dzieli 2. Dla każdej liczby pierwszej p wskazać możemy łańcuch $\{p^n \,|\, n \in \mathbb{N}_+\} = \{p, p^2, p^3, \ldots\}$.
- Niech $\mathcal{S}$ będzie podzbiorem zbioru wszystkich podzbiorów $P(X)$ niepustego zbioru X. Wprowadzamy RELACJĘ INKLUZJI w $\mathcal{S}$ postaci $A \leqslant B$ wtedy i tylko wtedy, gdy $A \subseteq B$. Gdy X jest zbiorem co najmniej dwuelementowym oraz $\mathcal{S} = P(X)$, to relacja ta nie jest relacją liniowego porządku, bowiem podzbiory jednoelementowe (różne) są nieporównywalne.

Definicja 6.6: Ograniczenie górne

Niech $\leqslant$ będzie relacją częściowego porządku w zbiorze $\mathcal{S}$ oraz niech $\mathcal{T}$ będzie podzbiorem $\mathcal{S}$. Powiemy, że element $s \in \mathcal{S}$ jest OGRANICZENIEM GÓRNYM zbioru $\mathcal{T}$, jeśli dla każdego $t \in \mathcal{T}$ mamy $t \leqslant s$.

Oczywiście ograniczenie górne nie musi należeć do zbioru $\mathcal{T}$. W zbiorze $\mathbb{R}$ z relacją liniowego porządku $\leqslant$ element 1 jest ograniczeniem górnym zbioru $(0, 1)$, ale do niego nie należy.

Definicja 6.7: Element maksymalny i element największy

Niech $\leqslant$ będzie relacją częściowego porządku w zbiorze $\mathcal{S}$. Powiemy, że element $x \in \mathcal{S}$ jest

- MAKSYMALNY, jeśli dla każdego $y \in \mathcal{S}$ takiego, że $y \geqslant x$ mamy $y = x$,
- NAJWIĘKSZY, jeśli dla każdego $y \in \mathcal{S}$ mamy $x \geqslant y$.

Kluczowe jest zauważenie, że element maksymalny w zbiorze $\mathcal{S}$ nie musi być większy od każdego innego elementu tego zbioru (czyli największy), ale jedynie od elementów, z którymi można go porównać – to właśnie znaczy, że jest maksymalny. Być może niewiele Państwo jeszcze mieli do czynienia z nieliniowymi porządkami i rozróżnienie to wydawać się może sztuczne. Ma ono jednak wielkie znaczenie.

Prostym przykładem takiej sytuacji jest choćby następująca sytuacja kombinatoryczna: w gronie 5 osób każdy ma pewną liczbę znajomych i można wprowadzić relację częściowego porządku przez inkluzję na owych zbiorach znajomych (podzbiorach zbioru pięcioelementowego). Jeśli założymy, że nie istnieje osoba, która zna wszystkich, to może się zdarzyć, że każda z 5 osób ma 4-osobowy zbiór znajomych i każdy z tych czterech podzbiorów jest maksymalny, tzn. nie zawiera się w zbiorze znajomych żadnej innej osoby (wystarczy przykładowo założyć, że znamy wszystkich, tylko nie samych siebie).

Oczywiście, gdyby pewna osoba jednak znała wszystkich, to jej zbiór znajomych byłby nie tylko maksymalnym, ale i największym podzbiorem znajomych w tej grupie. W tym uzupełnieniu, elementy największe nas nie interesują. Pierwsze skrzypce grają natomiast elementy maksymalne.

Z naszej perspektywy kluczowa jest sytuacja, gdy rozważamy podzbiór $\mathcal{S}$ zbioru podzbiorów przestrzeni liniowej V złożony z podzbiorów liniowo niezależnych. Dla przykładu, dla $V = \mathbb{R}^3$ zbiór:

$$B_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$$

jest elementem maksymalnym w $\mathcal{S}$, to znaczy – nie istnieje układ liniowo niezależny B_2, który zawiera B_1 oraz pewien element, który nie należy do B_1. Wskazać można wiele innych elementów maksymalnych $\mathcal{S}$ (czyli innych baz $\mathbb{R}^3$). A czym jest liniowo uporządkowany zbiór baz? Zróbmy krok do tyłu.

Aby zrozumieć jak działa Lemat Kuratowskiego-Zorna warto przyjrzeć się relacji częściowego porządku inkluzji w zbiorze podzbiorów $P(X)$ zbioru niepustego X. Czym jest liniowo uporządkowany ciąg elementów $P(X)$? Jest to na przykład (czy to jedyna możliwość?) ciąg wstępujących podzbiorów $A_1, A_2, A_3, \ldots$ spełniający warunek:

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq$$

Czy umiemy wskazać w $P(X)$ ograniczenie górne podzbioru $\{A_1, A_2, A_3, \ldots\}$? Tak, jest to nieskończona suma tych zbiorów

$$A = A_1 \cup A_2 \cup A_3 \cup \ldots = \bigcup_{i=1}^{\infty} A_i.$$

Zauważmy, że zbiór A nie jest żadnym ze zbiorów A_i, a jednak jest zawsze podzbiorem $P(X)$. To jest prawdą niezależnie od tego, czy łańcuch jest „wstępujący” (czy może być inaczej?) i ponumerowany liczbami naturalnymi (a jeśli jest nieprzeliczalny?). Sumę rodziny zbiorów można zdefiniować niezależnie od tego jaki zbiór ją indeksuje. W tym przypadku lemat Kuratowskiego-Zorna można wysłowić prościej.

Wniosek 6.4

Niech $\mathcal{S}$ będzie niepustą rodziną podzbiorów zbioru X uporządkowaną przez inkluzję. Jeśli dla każdego łańcucha $\{A_t\}_{t\in T}$ elementów $\mathcal{S}$ wiadomo, że

$$\bigcup_{t\in T} A_t \in \mathcal{S},$$

to w $\mathcal{S}$ istnieje element maksymalny.

Przykład. Niech $\mathbb{R}_+$ będzie zbiorem liczb rzeczywistych dodatnich. Zbiór ten można rozbić na sumę dwóch rozłącznych, niepustych podzbiorów, z których każdy jest zamknięty na dodawanie.

Rozwiązanie. Wprowadzamy porządek częściowy na zbiorze par (A, B), gdzie A, B są rozłącznymi podzbiorami $\mathbb{R}_+$, każdy domknięty ze względu na dodawanie, tzn. jeśli $a_1, a_2 \in A$, to $a_1 + a_2 \in A$, podobnie dla zbioru B. Nazwijmy ten zbiór par rozłącznych podzbiorów $\mathbb{R}_+$ przez $\mathcal{S}$. Jest to zbiór niepusty. Możemy wybrać na przykład $A = \mathbb{N} \setminus \{0\}$ oraz $B = \{n \cdot \pi \mid n \in \mathbb{N} \setminus \{0\}\}$. W tym przypadku $A \cup B \neq \mathbb{R}_+$.

Powiemy, że $(A, B) \leqslant (A', B')$, dla pewnych par $(A, B), (A', B') \in \mathcal{S}$, jeśli $A \subseteq A'$ oraz $B \subseteq B'$. Jest to oczywiście porządek częściowy w $\mathcal{S}$. Chcemy użyć lematu Kuratowskiego-Zorna i wybrać w $\mathcal{S}$ element maksymalny ze względu na relację $\leqslant$. Czego nam brakuje? Trzeba pokazać, że każdy łańcuch elementów w $\mathcal{S}$ ma ograniczenie górne. Niech A_t, B_t będą, dla $t \in T$ i pewnego zbioru T, takimi podzbiorami $\mathbb{R}_+$, że dla każdego t mamy $(A_t, B_t) \in \mathcal{S}$ oraz rodziny $\{A_t\}_{t\in T}$, $\{B_t\}_{t\in T}$ są łańcuchami. Zauważmy, że zbiory

$$A = \bigcup_{t\in T} A_t, \quad B = \bigcup_{t\in T} B_i$$

są rozłączne. Jesli bowiem $x \in A \cap B$, to z definicji sumy zbiorów dla pewnych $s, t \in T$ mamy $x \in A_s \cap B_t$. Mamy jednak $A_t \subseteq A_s$ lub $A_s \subseteq A_t$ (bo $\{A_t\}$ to łańcuch). A zatem $x \in A_s \cap B_s$ lub $x \in A_t \cap B_t$, co jest niemożliwe, bo $(A_s, B_s), (A_t, B_t) \in \mathcal{S}$, czyli zbiory A_s oraz B_s są rozłączne. Podobnie $A_t \cap B_t = \emptyset$.

Czy zbiory A, B są zamknięte na dodawanie? Oczywiście tak. Dowodzimy to podobnie jak wyżej. Czy widzimy, że dla każdego $t \in T$ zachodzi warunek $(A_t, B_t) \leqslant (A, B)$? Skoro tak, to element (A, B) jest ograniczeniem górnym łańcucha $\{(A_t, B_t)\}, t \in T$. Spełnione są zatem założenia lematu Kuratowskiego-Zorna i w zbiorze $\mathcal{S}$ istnieje element maksymalny (X, Y).

Chcemy teraz pokazać, że $X \cup Y = \mathbb{R}^+$. Gdyby istniał element $r \in \mathbb{R}^+ \setminus (X \cup Y)$, to bierzemy $X \cup \{r\}$ i domykamy ze względu na dodawanie dostając X'. Innymi słowy – X' jest maksymalnym podzbiorem $\mathbb{R}_+$ zawierającym X oraz r i zamkniętym na dodawanie. Czy taki zbiór istnieje? Owszem – na mocy Lematu Kuratowskiego-Zorna (proszę to sprawdzić). Bierzemy też Y' jako domknięcie addytywne $Y \cup \{r\}$. Twierdzimy, że jeden ze zbiorów $X' \cap Y$ lub $X \cap Y'$ jest pusty. Gdyby było inaczej, to mielibyśmy elementy $x_0, x_1 \in X, y_0, y_1 \in Y, m, n \geqslant 1$ takie, że

$$X' \ni x_0 + nr = y_0 \in Y \quad \text{oraz} \quad Y' \ni y_1 + mr = x_1 \in X.$$

Musimy mieć $m, n \geqslant 1$, bo $X \cap Y = \emptyset$. Wtedy jednak

$$mx_0 - nx_1 = m(y_0 - nr) - n(x_1 - mr) = my_0 + ny_1 \in X \cap Y.$$

Ale $X \cap Y = \emptyset$, sprzeczność. Zatem jeden ze zbiorów $X' \cap Y$ lub $X \cap Y'$ jest pusty. Ale to oznacza, że para (X', Y) lub (X, Y') jest ściśle większa niż (X, Y), sprzeczność z maksymalnością (X, Y). Dowód jest zakończony. Co ciekawe wiedząc, że istnieje baza $\mathbb{R}$ nad $\mathbb{Q}$ można wskazać jednolinijkowe uzasadnienie.

Przejdźmy do pokazania, że każda przestrzeń liniowa ma bazę. Pokażemy teraz, że w zbiorze $\mathcal{S}$ podzbiorów liniowo niezależnych dowolnej przestrzeni liniowej V istnieje element maksymalny. Tego nam jedynie brakuje, by pokazać, że każda przestrzeń liniowa posiada bazę. Czy zbiór $\mathcal{S}$ spełnia warunki Lematu Kuratowskiego-Zorna? Zakładamy, że $V \neq 0$, więc $\mathcal{S} \neq 0$. Na mocy wniosku wystarczy pokazać, że dla dowolnego łańcucha $\{B_t\}, t \in T$ liniowo niezależnych podzbiorów B_t przestrzeni V również zbiór $B = \bigcup_t B_t$ jest liniowo niezależny. To jest jednak oczywiste. Biorąc dowolny skończony układ elementów $v_1, \ldots, v_n$ zbioru B wiemy, że istnieją elementy $B_{t_1}, \ldots, B_{t_k}$ łańcucha B_t, że każdy v_i należy do jednego ze zbiorów B_{t_i}. Skoro zbiory te tworzą skończony łańcuch, to istnieje $N \in \{1, \ldots, k\}$ takie[4], że $v_1, \ldots, v_k \in B_{i_N}$. A zatem $v_1, \ldots, v_n$ jest podukładem układu liniowo niezależnego B_{i_N} — czyli jest to układ liniowo niezależny. A zatem B jest układem liniowo niezależnym. Zgodnie z Lematem Kuratowskiego-Zorna w $\mathcal{S}$ istnieje element maksymalny i jak pokazaliśmy wcześniej – jest to baza przestrzeni V.

Widzimy teraz zaskakującą wymowę Lematu Kuratowskiego-Zorna. Rezultat ten ma ogromne znaczenie dla wielu obszarów matematyki. Wysłowiony przez Kazimierza Kuratowskiego w 1922 roku, a potem przez Zorna w roku 1935 pozwolił uprościć wiele rozumowań opartych wcześniej o tzw. aksjomat wyboru lub twierdzenie o dobrym porządku, z którymi jest równoważny. Również twierdzenie o istnieniu bazy dowolnej przestrzeni liniowej równoważne jest lematowi Kuratowskiego-Zorna. W toku studiów poznają Państwo wiele głębokich zastosowań tego fundamentalnego rezultatu, nie tylko w algebrze.

[4]Fakt: jeśli $\{s_1, \ldots, s_n\}$ jest skończonym łańcuchem, to istnieje s_i takie, że $s_j \leqslant s_i$, dla $j \neq i$. Dowód można sformalizować przez indukcję, albo wyobrazić sobie za pomocą algorytmu sortowania bąbelkowego.

6.4 Dodatek. Nieprzeliczalne układy. Algebraiczna niezależność.

Poniższe zadanie ilustruje jak skomplikowana jest struktura ciała $\mathbb{R}$ traktowanego jako przestrzeń liniowa (nieskończonego wymiaru) nad ciałem $\mathbb{Q}$. Wypisanie wprost bazy $\mathbb{R}$ nad $\mathbb{Q}$ nie jest możliwe. Można jednak wypisać równoliczny z $\mathbb{R}$ układ liniowo niezależny. Za udostępnienie materiału dziękuję dr. Ł. Kubatowi.

Zadanie. Ustawmy liczby wymierne $\mathbb{Q}$ w ciąg $(q_n)_{n\in\mathbb{N}}$. Dla dowolnego $t\in\mathbb{R}$ niech

$$a(t)=\sum_{n\in N(t)}\frac{1}{n!},\qquad \text{gdzie } N(t)=\{n\in\mathbb{N}: q_n<t\}.$$

(1) Sprawdź, że szereg definiujący $a(t)$ jest zbieżny (czyli definicja jest poprawna).

(2) Wykaż, że dla dowolnych $s,t\in\mathbb{R}$ zachodzi $s\neq t\Longrightarrow a(s)\neq a(t)$.

(3) Udowodnij, że zbiór $A=\{a(t): t\in\mathbb{R}\}\subseteq\mathbb{R}$ jest liniowo niezależny nad $\mathbb{Q}$.

Rozwiązanie. Zauważmy, że

$$a(t)=\sum_{n\in N(t)}\frac{1}{n!}<\sum_{n=0}^{\infty}\frac{1}{n!}=e,$$

co dowodzi (1). Gdy $s,t\in\mathbb{R}$ spełniają $s<t$, to $(s,t)\cap\mathbb{Q}\neq\varnothing$. Istnieje więc takie $n\in\mathbb{N}$, że $s<q_n<t$. Zatem $a(s)<a(s)+\frac{1}{n!}<a(t)$, co dowodzi (2). Aby udowodnić (3) załóżmy, dla dowodu nie wprost, że

$$q_1a(t_1)+\cdots+q_ka(t_k)=0 \tag{$*$}$$

dla pewnych $t_1,\dots,t_k\in\mathbb{R}$ spełniających $t_1>\cdots>t_k$, gdzie liczby $q_1,\dots,q_k\in\mathbb{Q}$ nie są wszystkie równe zeru. Wśród równości typu $(*)$ możemy wybrać najkrótszą, czyli taką, w której k jest najmniejsze. Oczywiście musi być $k\geqslant 2$. Ponadto, dzięki minimalności k, koniecznie $q_1,\dots,q_k\neq 0$. Mnożąc $(*)$ przez stosowną liczbę naturalną możemy założyć, że $q_1,\dots,q_k\in\mathbb{Z}$. Dla $t\in\mathbb{R}$ oraz $m\in\mathbb{N}$ niech

$$L_m(t)=\{n\in N(t): n\leqslant m\}\qquad\text{oraz}\qquad R_m(t)=\{n\in N(t): n>m\}.$$

Mnożąc $(*)$ przez $m!$ otrzymujemy $L(m)=-R(m)$, gdzie

$$L(m)=q_1\Big(\sum_{n\in L_m(t_1)}\frac{m!}{n!}\Big)+\cdots+q_k\Big(\sum_{n\in L_m(t_k)}\frac{m!}{n!}\Big),$$
$$R(m)=q_1\Big(\sum_{n\in R_m(t_1)}\frac{m!}{n!}\Big)+\cdots+q_k\Big(\sum_{n\in R_m(t_k)}\frac{m!}{n!}\Big).$$

Oczywiście $L(m)\in\mathbb{Z}$. Ponadto

$$\begin{aligned}|R(m)|&\leqslant|q_1|\Big(\sum_{n\in R_m(t_1)}\frac{m!}{n!}\Big)+\cdots+|q_k|\Big(\sum_{n\in R_m(t_k)}\frac{m!}{n!}\Big)\\&\leqslant|q_1|\Big(\sum_{n>m}\frac{m!}{n!}\Big)+\cdots+|q_k|\Big(\sum_{n>m}\frac{m!}{n!}\Big)\\&\leqslant\frac{|q_1|+\cdots+|q_k|}{m+1}\Big(\sum_{n>m}\frac{1}{(n-m)!}\Big)\\&\leqslant\frac{|q_1|+\cdots+|q_k|}{m+1}e.\end{aligned}$$

Wynika stąd, że gdy m jest duże, to $|R(m)|<1$. Skoro $R(m)=-L(m)\in\mathbb{Z}$, to musi zachodzić równość $L(m)=R(m)=0$. Ponieważ zbiór $(t_2,t_1)\cap\mathbb{Q}$ jest nieskończony, to znajdziemy takie $m\in\mathbb{N}$ by jednocześnie $|R(m)|<1$ (wtedy, jak wiemy, $L(m)=R(m)=0$) oraz $t_2<q_m<t_1$. W tej sytuacji mamy $m\in L_m(t_1)\setminus(L_m(t_2)\cup\cdots\cup L_m(t_k))$. Zatem równość $L(m)=0$ implikuje

$$-q_1=q_1\Big(\sum_{\substack{n\in L_m(t_1)\\ n\neq m}}\frac{m!}{n!}\Big)+q_2\Big(\sum_{n\in L_m(t_2)}\frac{m!}{n!}\Big)+\cdots+q_k\Big(\sum_{n\in L_m(t_k)}\frac{m!}{n!}\Big). \tag{$**$}$$

Obie strony równania $(**)$ są liczbami całkowitymi. Ponadto prawa strona jest podzielna przez m. Zatem także q_1 jest podzielne przez m. W takim razie musi być $q_1=0$, gdyż m można wybrać tak, by $m>|q_1|$. Uzyskana sprzeczność ($q_1=0$) prowadzi do wniosku, że zbiór A jest liniowo niezależny nad $\mathbb{Q}$. $\square$

Uwaga. Dowodzi się, że choć zbiór A jest tej samej mocy co $\mathbb{R}$, to nie rozpina on $\mathbb{R}$ nad $\mathbb{Q}$. Można także wykazać (patrz J. von Neumann, *Ein System algebraisch unabhängiger Zahlen*, Math. Ann. **99** (1928), pp. 134–141), że liczby postaci

$$b(t) = \sum_{n=0}^{\infty} \frac{2^{2^{[nt]}}}{2^{2^{n^2}}} \qquad \text{dla } t > 0$$

($[x]$ oznacza część całkowitą liczby $x \in \mathbb{R}$)

są nie tylko liniowo niezależne nad $\mathbb{Q}$, ale nawet `algebraicznie niezależne` nad $\mathbb{Q}$, tzn. dla dowolnego $n \geqslant 1$, dowolnych $0 < t_1 < \cdots < t_n$ oraz dowolnego wielomianu zmiennych $x_1, \ldots, x_n$, czyli dla pewnego $0 \neq f \in \mathbb{Q}[x_1, \ldots, x_n]$ zachodzi $f(b(t_1), \ldots, b(t_n)) \neq 0$.

Przykład ilustrujący algebraiczną zależność. Liczby $\sqrt{\pi}$ oraz $2\pi + 1$ są liniowo niezalezne nad $\mathbb{Q}$, ale są algebraicznie zależne, ponieważ wielomian $2x^2 - y - 1 \in \mathbb{Q}[x, y]$ zeruje się dla $x = \sqrt{\pi}$ oraz $y = 2\pi + 1$.

Prof. J. Mycielski pokazał nastepujace twierdzenie (*Algebraic independence and measure*, Fund. Math **61** (1967), pp. 165–169) dla dowolnego doskonałego podzbioru $\mathbb{R}$ (tzn. niepustego, domkniętego oraz bez punktów izolowanych), patrz: `http://matwbn.icm.edu.pl/ksiazki/fm/fm61/fm61117.pdf`.

Twierdzenie 6.6

Każdy doskonały podzbiór zbioru liczb rzeczywistych zawiera doskonały podzbiór, który jest algebraicznie niezależny nad $\mathbb{Q}$.

Pojęcie algebraicznej niezależności elementów $\mathbb{R}$ nad $\mathbb{Q}$, a także elementów $\mathbb{C}$ nad $\mathbb{Q}$, związane jest ściśle z pojęciem liczb algebraicznych i przestępnych, z którymi spotkaliście się Państwo (lub spotkacie) na Analizie Matematycznej. Pojęcie to jest bardzo subtelne i tajemnicze. Przestępność liczby π została udowodniona po raz pierwszy właśnie dzięki badaniu algebraicznej niezależności (dowód dla e dokonał elementarnymi metodami analitycznymi Hermite w 1873 roku). Zachodzi mianowicie następujący rezultat.

Twierdzenie 6.7: Lindemann-Weierstrass, 1885

Jeśli $\alpha_1, \ldots, \alpha_n$ są liczbami algebraicznymi liniowo niezależnymi nad $\mathbb{Q}$, to liczby $e^{\alpha_1}, \ldots, e^{\alpha_n}$ są algebraicznie niezależne nad $\mathbb{Q}$.

Aby zrozumieć jakie są związki tego wyniku z przestępnością odnotujmy inne, równoważne sformułowanie.

Twierdzenie 6.8: Baker, 1966

Jeśli $\alpha_1, \ldots, \alpha_n$ są parami różnymi liczbami algebraicznymi, to liczby $e^{\alpha_1}, \ldots, e^{\alpha_n}$ są liniowo niezależne nad ciałem liczb algebraicznych $\overline{\mathbb{Q}}$ (algebraiczne domknięcie $\mathbb{Q}$ w $\mathbb{C}$).

Jak można stosować to twierdzenie? Jeśli α jest niezerową liczbą algebraiczną to zbiór $\{0, \alpha\}$ zawiera różne elementy algebraiczne, więc zbiór $\{e^0, e^\alpha\}$, czyli $\{1, e^\alpha\}$ jest liniowo niezależny nad ciałem liczb algebraicznych, w szczególności e^α nie jest algebraiczna. Gdy udowodnimy przestępność liczby e możemy z niej łatwo wywnioskować przestępność liczby π, korzystając ze słynnej tożsamości algebraicznej Eulera $e^{\pi i} + 1 = 0$. Dokładniej, gdyby π była liczbą algebraiczną to πi również, a wtedy przestępna musi być, na mocy poprzedniego argumentu liczba $e^{\pi i} = -1$, co jest niemożliwe. Zatem π jest przestępna. Prosty wariant tego argumentu pokazuje również, że dla niezerowej liczby algebraicznej α liczby $\sin(\alpha), \cos(\alpha), \operatorname{tg}(\alpha)$ i ich hiperboliczne odpowiedniki są liczbami przestępnymi.

Dowód Twierdzenia Lindemanna jest skomplikowany ale o zagadnieniach tego typu i szeregu innych rezultatów dotyczących liczb przestępnych: `http://www.math.leidenuniv.nl/~evertse/dio15-4.pdf`.

6.5 Trivia. Podział prostokąta na kwadraty, czyli intuicja miary.

Problem – Zadanie. Prostokąt R o bokach długości 1 oraz x, gdzie x jest liczbą niewymierną, nie może być złożony ze skończenie wielu kwadratów.

Załóżmy przeciwnie, że takie rozcięcie prostokąta o rozmiarach $1 \times x$ jest możliwe. Dzielimy go na kwadraty $Q_1, \ldots, Q_n$, gdzie s_i jest długością boku każdego z kwadratów Q_i, dla $1 \leqslant i \leqslant n$. UWAGA: to wcale nie muszą (nie mogą wszystkie) być liczby wymierne! Potraktujemy te liczby jako... wektory!

Rozważać będziemy ciało $\mathbb{R}$ jako przestrzeń liniową nad ciałem $\mathbb{Q}$. Mówiliśmy już kilkakrotnie, że to jest dość niezwykła, nieskończenie wymiarowa przestrzeń, kryjąca wiele niespodzianek. Niech $V \subseteq \mathbb{R}$ będzie podprzestrzenią rozpiętą przez liczby $s_1, \ldots, s_n$. Czyli: V to zbiór kombinacji liniowych (o współczynnikach w $\mathbb{Q}$) tego układu liczb. Skoro (jak twierdzimy) możliwy jest podział prostokąta R na sumę kwadratów o bokach s_i, to mamy $1, x \in \mathrm{lin}(s_1, s_2, \ldots, s_n)$, bo $1, x$ są po prostu sumami pewnych s_i.

Teraz pojawia się sprytna (ale jakże często stosowana w matematyce) sztuczka. Określamy funkcję $f : V \to \mathbb{R}$ spełniającą warunki $f(1) = 1$, $f(x) = -1$ oraz taką, że dla każdych $x, y \in V$ mamy $f(x + y) = f(x) + f(y)$ oraz $f(qx) = qf(x)$, dla $q \in \mathbb{Q}$. Czy taka funkcja istnieje? Przecież to musiałoby być jedno z tych dziwnych rozwiązań równania Cauchy'ego postawionego na poprzednim wykładzie! Mamy tu funkcję nie z $\mathbb{R}$ do $\mathbb{R}$, tylko z V do $\mathbb{R}$. Dziwne, prawda? Funkcja taka jednak istnieje.

Owszem, skoro $1, x$ są liniowo niezależne nad $\mathbb{Q}$ (a to łatwo sprawdzić), to układ ten możemy na mocy tw. Steinitza dopełnić do bazy $1, x, b_3, \ldots, b_k$ przestrzeni V (niekoniecznie $k = n$, bo może niektóre s_i to kombinacje liniowe pozostałych?). Kładziemy dalej $f(1) = 1$, $f(x) = -1$ oraz $f(b_i) = 0$, dla $i = 3, 4, \ldots$ Następnie mając funkcję f określoną na samej tylko bazie V bierzemy dowolny wektor $v \in V$ i rozpisujemy go (jednoznacznie!) w bazie $1, x, b_3, \ldots, b_k$ w postaci $v = a_1 + a_2x + a_3b_3 + \ldots + a_kb_k$, gdzie $a_i \in \mathbb{Q}$. Definiujemy teraz $f(v) := a_1f(1) + a_2f(x) + a_3f(b_3) + \ldots + a_kf(b_k) = a_1 - a_2$. Zachęcam każdego do sprawdzenia, że teraz nasza funkcja spełnia warunki $f(x + y) = f(x) + f(y)$ oraz $qf(x) = f(qx)$. Tego typu funkcje wprowadzimy niedługo na wykładzie w większej ogólności. Na razie ograniczamy się do powyższych wyjaśnień. Zauważmy też, że absolutnie nie pojawił się **wzór na** f (w zwykłym sensie).

Rozważmy teraz, dla każdego prostokąta A o bokach a, b, gdzie $a, b \in V$, liczbę $v(A) = f(a)f(b)$. Jeśli prostokąt R rozmiarów $1 \times x$ byłby złożony z kwadratów $Q_1, \ldots, Q_n$, to ze wzoru na sumę pól mamy:

$$v(R) = v(Q_1) + v(Q_2) + \ldots + v(Q_n).$$

Jak to jest jednak możliwe, skoro $v(R) = f(1)f(x) = -1$, zaś $v(Q_i) = f(s_i)^2 \geqslant 0$, dla wszystkich i? To jest sprzeczność. A zatem R nie może być rozcięty na kwadraty $Q_1, \ldots, Q_n$. Problem rozwiązany.

Rozumowanie to może budzić wiele pytań. Wydaje się, że jest ono przesadnie skomplikowane i wymaga jakiejś strasznej maszynerii. Dlaczego to było konieczne? Oczywiście problemem jest fakt, że postulowany podział kwadratu jest stosunkowo dowolny, liczba składników jest duża, nie muszą to być kwadraty o bokach wymiernej długości – mimo wszystko mamy prawo być zaskoczeni. Użyliśmy poważnej technologii z wykładu, a nawet przemyciliśmy po cichu pojęcie przekształcenia liniowego. To powinno zastanowić.

Rzecz jasna istnieje drugie dno całego tego problemu. Tak naprawdę korzystaliśmy tu po cichu z własności addytywności pola na płaszczyźnie. Nie definiowaliśmy zbyt ściśle co oznacza rozbicie na kwadraty itd. To oczywiście można doprecyzować. Ale jest pewien ogólniejszy problem. Nietrudno pokazać, że dowolny wielokąt na płaszczyźnie można pociąć na części, z których ułoży się prostokąt (a nawet kwadrat) – mówimy, że dowolny wielokąt jest **równoważny przez pocięcie** z prostokątem. Stąd sposób obliczenia pola dowolnego wielokąta jest wyznaczony jednoznacznie. Pytanie: **czy można dowolny wielościan pociąć na skończoną liczbę mniejszych wielościanów, z których ułoży się prostopadłościan**?

To zaskakujące pytanie było jednym z tzw. 23 problemów Hilberta ogłoszonych w 1900 jako najpoważniejsze problemy matematyczne na nadchodzący XX wiek. Przynajmniej cztery są otwarte do dziś. Problem, który rozważamy rozwiązano jako jeden z pierwszych. Jeszcze w tym samym roku Max Dehn udowodnił, że istnieją pary wielościanów... które nie są równoważne przez pocięcie! Dowód ten jest zrozumiały dla wytrwałych i opisany bardzo przejrzyście w artykule prof. Marka Kordosa „Pole i objętość" na łamach czasopisma Delta (jest dostępny online – wystarczy wyszukać na stronie `http://www.deltami.edu.pl/`).

6.6 Coda. O kształtowaniu się pojęcia wymiaru

W trakcie studiów poznają Państwo szereg podstawowych koncepcji matematycznych, które rozważane będą w coraz ogólniejszym kontekście na kolejnych przedmiotach. Należą do nich chociażby koncepcje liczby, przestrzeni, ciągłości, nieskończoności, niezależności oraz właśnie wymiaru (a także wiele innych). Na kursie algebry liniowej budujemy dość prostą, ale niezwykle użyteczną teorię przestrzeni liniowych, zbudowaną wokół pojęcia wektora i pewnych podstawowych aksjomatów motywowanych geometrycznie. W ten sposób możliwe jest mówienie o pewnych zjawiskach mających podłoże geometryczne i możliwe jest określenie choćby pojęcia wymiaru. Nie jest to pojęcie banalne, a jego historia jest bardzo ciekawa.

Czytelnik pytać może — czy wystarczy mówić o wymiarze przestrzeni liniowych? A jeśli chcemy rozważać ogólniejsze przestrzenie, choćby przestrzenie metryczne lub topologiczne, rozmaitości, czy fraktale — jak dla tych obiektów definiować wymiar? Historycznie rzecz biorąc właściwej dynamiki problem ten nabiera w wieku XIX, wraz z odkryciem przez Cantora, że prosta i płaszczyzna są w istocie równoliczne oraz wraz ze znalezieniem przez Peano ciągłego odwzorowania odcinka na kwadrat. Co było wcześniej?

W świecie Pitagorejczyków wymiary miały znaczenie filizoficzne i do pewnego stopnia religijne. Sam Arystoteles pisze w swoim dziele *O niebie*:

> Wielkość, jeśli jest podzielna w jedną stronę, jest linią, jeśli w dwie strony – powierzchnią, a jeśli w trzy – ciałem. Poza nimi nie ma innej wielkości, ponieważ istnieją tylko trzy wymiary, a to, co jest podzielne w trzech kierunkach, jest podzielne we wszystkich [...] Bo, jak mówią pitagorejczycy, wszechświat i wszystko, co się w nim znajduje, jest określone przez liczbę trzy, ponieważ początek, środek i koniec dają liczbę wszechświata, a liczba, którą podają, jest triadą. I tak, wziąwszy te trzy z natury jako (że tak powiem) jej prawa, dalej używamy liczby trzy w kulcie Bogów.

Także u Euklidesa występują obiekty wymiaru 1, 2 czy 3, a obiektom trójmymiarowym poświęca się sporo miejsca. Jednym ze szczytowych wszakże osiągnięć *Elementów* i starożytnej matematyki jest klasyfikacja wielościanów foremnych. Poza trzema wymiarami, ani Arystoteles, ani Euklides czy Ptolemeusz nie widzieli innych możliwych opcji. Ten stan rzeczy zachował się w zasadzie aż do wieku XVIII. I tak dla przykładu Kepler argumentować będzie, że liczba wymiarów równa 3 istnieje na chwałę Trójcy Świętej, Wallis w swoim wykładzie algebry z roku 1686 napisze, że wyżej wymiarowa przestrzeń jest to ‚potwór z natury, mniej możliwy niż chimera czy centaur", a trójwymiarowości przestrzeni dowodzić będzie próbował (bez powodzenia) w swoim doktoracie sam Immanuel Kant (1747).

Skąd wzięła się potrzeba wyjścia poza trzy wymiary? Z jednej strony, kluczowym elementem było przeniesienie rozważań geometrycznych do układu współrzędnych, czyli osiągnięcie Kartezjusza. W ten sposób teoria krzywych czy powierzchni, stała się teorią równań z wieloma niewiadomymi. Na różne sposoby badano intuicję „stopni swobody", czyli liczby niezależnych parametrów potrzebnych do opisu ruchu punktu znajdującego się na obiekcie, którego wymiar rozważamy. Także z naszego punktu widzenia — rozwiązanie układu równań jednorodnych zależne od k parametrów opisuje podprzestrzeń wymiaru k, i z każdego rozwiązania można „dotrzeć do innego" przy użyciu k liniowo niezależnych rozwiązań.

Na nasze potrzeby wystarczy stwierdzić, że już w połowie XVIII wieku zorientowano się, że do rozwiązywania niektórych zagadnień (choćby pochodzących z mechaniki) nie wystarczy rozważanie nie więcej niż trzech parametrów. Idea ta wyrażona została wprost przez Jeana d'Alemberta w Wielkiej Encyklopedii Francuskiej — pomniku Oświecenia redagowanym wspólnie z Diderotem. We wpisie „Wymiar" wpisana była idea reprezentowania praw mechaniki w czterech wymiarach, z których czwartym był czas. Podobną ideę wyraził w 1797 roku jeden z najważniejszych po odejściu Eulera matematyków w Europie — Joseph-Louis Lagrange w swoich *Mécanique Analytique* (1788) oraz *Théorie des Fonctions Analytiques* (1797).

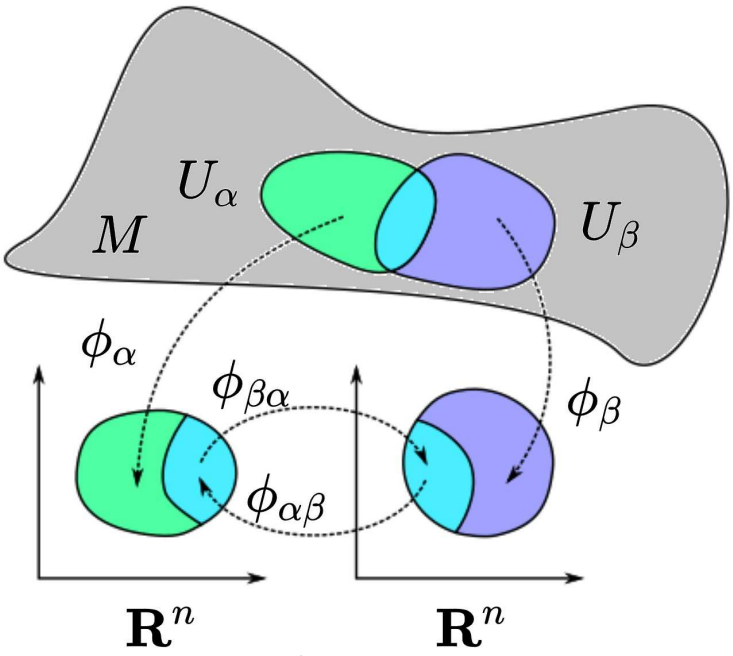
Kolejna ciekawa intuicja czwartego wymiaru pochodziła od Ferdynanda Mobiusa, jednego z twórców geometrii rzutowej, który stwierdził, że symetryczne do siebie figury trójwymiarowe można by na siebie nałożyć, gdyby istniał czwarty wymiar (opowiemy o tym w drugim semestrze — idea jest taka, że tak jak do nałożenia na siebie obiektów symetrycznych na płaszczyźnie typu $\S 2$ potrzeba wyjścia w trzeci wymiar, tak i podobnie jest np. z prawą i lewą dłonią lub lewym czy prawym butem — jeśli wyjdziemy w czwarty wymiar). Współrzędne jednorodne (barycentryczne), które poznamy w kolejnym semestrze, również wymagały w wersji przestrzennej używania czterech współrzędnych.

Kolejni matematycy używający punktów o więcej niż trzech współrzędnych, wywodzili się nie tylko z geometrii (a jest tu wiele ciekawych wątków, choćby Julius Plücker i współrzędne jednorodne w przestrzeni rzutowej, Ludwig Schläfli i wielościany foremne w czwartym wymiarze, Hamilton i kwaterniony...), ale przede wszystkim z analizy. W 1847 roku Cauchy ogłosił, że „nazywać będziemy zbiór n zmiennych punktem analitycznym, a równanie lub układ równań — miejscem analitycznym". Zasadnicze jednak i ostateczne przejście do wyżej wymiarowej geometrii dokonało się za sprawą słynnego wykładu habilitacyjnego Riemanna z 1854 roku „O hipotezach leżących u podstaw geometrii". Wykład ten był jednym z najważniejszych wydarzeń w historii matematyki, wciąż mającym na nią wielki wpływ.

Co było tak istotnego w wykładzie Riemanna? Gdy w 1915 roku Einstein zmienił dzięki ogólnej teorii względności nasze rozumienie wszechświata, sformułował pojęcie czterowymiarowej czasoprzestrzeni, która zagina się i zakrzywia w reakcji na koncentrację masy lub energii. Jest więc obiektem zakrzywionym — jak bardzo? To gigantyczny problem współczesnej matematyki i fizyki (choćby teoria superstrun, przewiduje, że do unifikacji teorii względności i teorii kwantowej należy rozważać 11-wymiarową czasoprzestrzeń). Pytanie brzmi jednak — skąd możemy wiedzieć, że znajdujemy się na zakrzywionej czasoprzestrzeni, będąc w jej środku? Skąd wiemy, że jest zakrzywiona? To pytanie pięknie. spopularyzowano. Abbott we *Flatlandii* (1884) pyta nas jak rozumieją trzeci wymiar „płaszczaki" żyjące na płaszczyźnie?

Ktoś mógłby powiedzieć: bez trudu umiemy odróżnić czy żyjemy na obiekcie zakrzywionym, czy nie, bo przecież mamy równania opisujące różne obiekty. Równanie $x_1^2 + x_2^2 + x_3^2 = 1$ opisuje sferę w przestrzeni trójwymiarowej i nikt nie wątpi, że nie jest to równanie liniowe (dokładniej powiemy o tym w drugim semestrze). Co więcej, sfera, gdy przyjrzymy się jej z bliska, wygląda jak zwykła powierzchnia dwuwymiarowa, a stąd rozsądne jest przypisywanie jej właśnie wymiaru 2. Jest to obiekt dwuwymiarowy w przestrzeni trójwymiarowej. Tak o niej myślimy. Sfera jest przykładem tzw. rozmaitości dwuwymiarowej, tak jak okrąg — jednowymiarowej na płaszczyźnie. Pojęcie rozmaitości pochodzi od Riemanna.

Nie mamy tu narzędz, by dokładnie opowiedzieć o rozmaitościach, ale idea jest następująca: aby o pewnym obiekcie zawartym w (powiedzmy) $\mathbb{R}^N$ powiedzieć, że jest n-wymiarową rozmaitością, potrzebujemy mieć sposób rysowania map obszarów (otwartych) tej rozmaitości. Sfera jest dwuwymiarowa, bo zataczając wokół dowolnego znajdującego się na niej punktu okrąg, możemy uzyskany wycinek sfery przekształcić („zmapować") jednoznacznie w dysk leżący na płaszczyźnie. Oto intuicyjny obrazek owych map.



Rys 1. Źródło: Wikipedia.

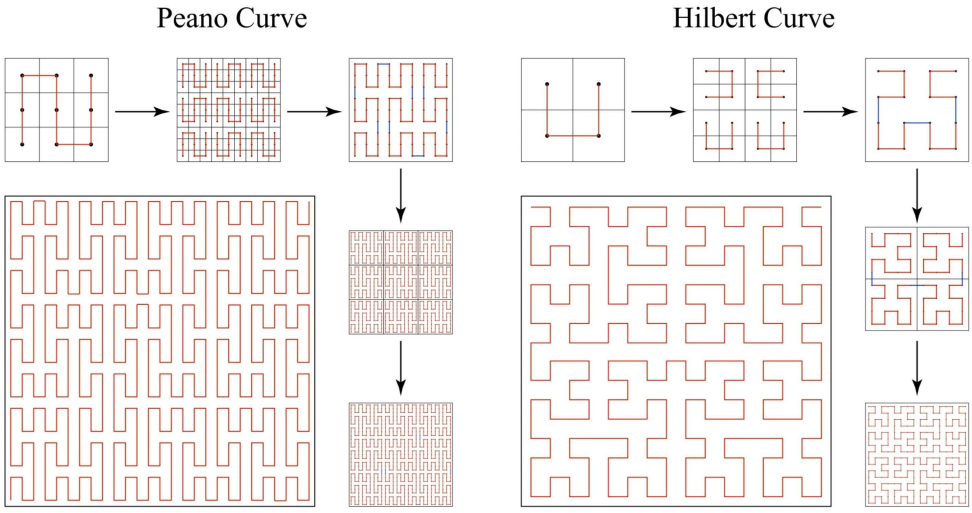
Formalnie są to różniczkowalne i wzajemnie jednoznaczne przekształcenia tych obszarów na odpowiednie otwarte podzbiory przestrzeni $\mathbb{R}^n$. Muszą one być zgodne, co oznacza, że dla dwóch obszarów U_α oraz U_β i dwóch map ϕ_α, ϕ_β, mapy $\phi_\alpha^{-1} \circ \phi_\beta$ oraz $\phi_\beta^{-1} \circ \phi_\alpha$ muszą również mieć własności map — być różniczkowalne i wzajemnie jednoznaczne. Co to wszystko znaczy, dowiedzą się Państwo na Analizie i Topologii.

Założenie, które przyjmujemy w tym podejściu jest takie, że jesteśmy w stanie widzieć nasz obiekt, np. sferę, jakby od zewnątrz, czyli — zanurzamy go w coś większego i opisujemy np. równaniami czy funkcjami. Jednak, jeśli zaczniemy rozmawiać o kształcie całego wszechświata, lub po prostu obiektu, poza który nie możemy wyjść, to jak mierzyć jego krzywiznę? Na to pytanie próbowali odpowiedzieć Gauss i Riemann.

Riemann był uczniem Gaussa, u którego w 1851 roku przygotował rozprawę o teorii zmiennych zespolonych, mających później stać się podstawą tzw. powierzchni riemannowskich. Gauss opisywał to osiągnięcie jako wzniosłe i niezwykłe płodne. To na prośbę Gaussa Riemann przygotowywał swój wykład inauguracyjny. Celem było właśnie sformułowanie użytecznej definicji miary krzywizny przestrzeni. Teoria ta była rozwijana przez Gaussa dla przestrzeni dwuwymiarowej. Wykazał on, że jedna zmienna potrzebna jest do opisu krzywizny w otoczeniu punktu w przestrzeni dwuwymiarowej (tzw. krzywizna Gaussa). Riemann rozwinął to pojęcie na przestrzenie wyższych wymiarów. Wykazał, że do opisu krzywizny w przestrzeni trójwymiarowej potrzeba sześciu zmiennych (tzw. metryka riemannowska), a w czterowymiarowej przestrzeni — dwudziestu zmiennych. Ogólny obiekt opisany przez Riemanna — tensor krzywizny, jest właśnie podstawą i głównym narzędziem ogólnej teorii Einsteina. O kwestiach tych będziecie Państwo się uczyć na geometrii różniczkowej. Bez algebry liniowej i jej zaawansowanych narzędzi teoria ta nie ma racji bytu.

Z naszej perspektywy ważne jest to, że pierwsza część wykładu Riemanna zawierała zdefiniowaną w sposób jawny i klarowny przestrzeń n-wymiarową. Idee Riemanna zdecydowanie wyprzedzały swoje czasy i zapewne jedynie Gauss był w stanie docenić ich głębię. Niemniej jednak napływające z różnych źródeł matematyki koncepcje wielowymiarowej przestrzeni sprawiły, że już pod koniec XIX wieku pojawiła się mnogość książek, wspomnień i dzieł omawiających i popularyzujących wyższe wymiary. W 1895 roku wielki Henri Poincare pisał, że „geometria n wymiarów bada rzeczywistość; nikt w to nie wątpi".

Z matematycznego punktu widzenia dopiero co wyemancypowane pojęcie czekał ogromny kryzys, i to nie w czterech czy więcej wymiarach, ale już w wymiarze 1 czy 2. Wspomniane już wyżej odkrycia Cantora (1878) i Peano (1890) sprawiły, że stosowane dotychczas intuicyjne definicje wymiaru przestały był sensowne.



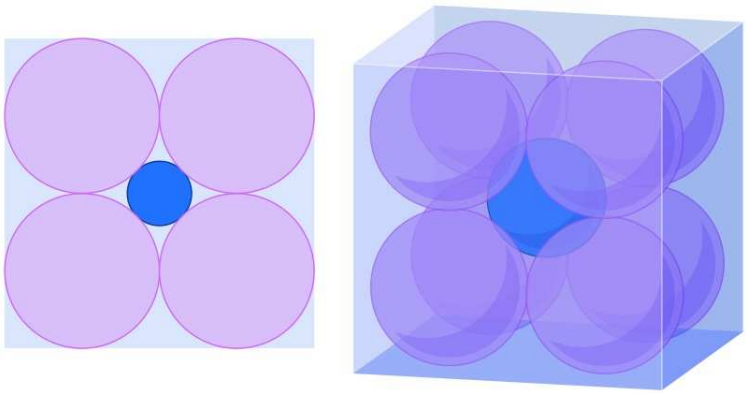
Rys 2. Źródło: `https://galileo-unbound.blog/2023/03/08/a-short-history-of-hyperspace/`.

Głębsze zrozumienie przykładów wyżej otrzymają Państwo na analizie i topologii. Problem jest jasny — jak zdefiniować krzywą wymiaru 1, by definicja nie obejmowała kwadratu? Pod koniec XIX wieku pojawiło się pytanie czy istnieje parametryczna reprezentacja kwadratu na odcinek, która byłaby jednocześnie ciągła i wzajemnie jednoznaczna (czyli homeomorfizm — pojęcie, które poznają Państwo na II roku). Zapytano ogólniej — kiedy iloczyn kartezjański n-kopii odcinka jednostkowego $[0,1]$, oznaczany przez I^n, jest homeomorficzny z kostką I^m, dla $m \neq n$. Oczekiwano, że nie jest to możliwe i pomiędzy rokiem 1890 i 1910 pojawiło się wiele fałszywych dowodów tego faktu. Poprawny przedstawił dopiero Brouwer w 1911.

Zauważmy, że zbiory typu I^n nie są przestrzeniami liniowymi, więc rezultat ten dał przekonanie, że powinna istnieć funkcja, przypisująca tzw. przestrzeniom topologicznym liczbę zwaną wymiarem. Kroki w tym kierunku poczynił najpierw Poincare w roku 1912, wywodząc z intuicji „oddzielania" obiektów wyżej wymiarowych obiektami niżej wymiarowymi ideę indukcyjnej definicji wymiaru. Idea opierała się na obserwacji, że obiekty trójwymiarowe oddzielać można dwuwymiarowymi (np. dwa rozłączne wielo-

ściany za pomocą sfer), obiekty dwumiarowe krzywymi itd. Stosowną definicję podał Brouwer w 1913 r.

W przestrzeniach wysokich wymiarów odkryć można wiele nieintuicyjnych zjawisk, które rozważać będziecie Państwo na wyższych latach studiów (a co dopiero w wymiarze nieskończonym). Oto stosunkowo prosty przykład. Umieśćmy 2^n sfer o promieniu 1 wewnątrz n-wymiarowej kostki, której bok ma długość 4. W środku kostki umieśćmy kolejną sferę, styczną zewnętrznie do czterech już umieszczonych.

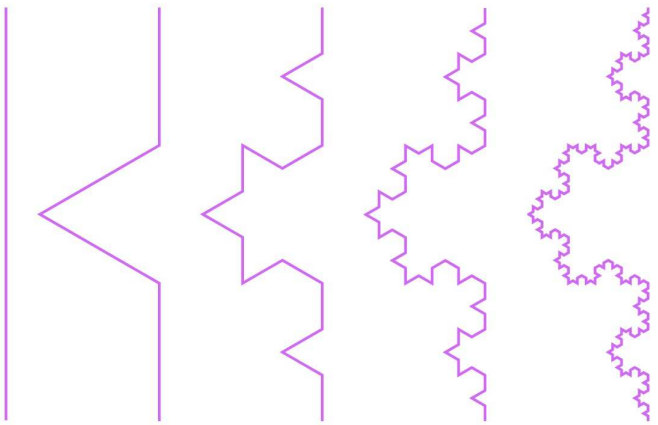


Rys 3. Źródło: `https://www.quantamagazine.org/a-mathematicians-guided-tour-through-high-dimensions-20210913/`.

Gdy wymiar n rośnie, rośnie również promień niebieskiej sfery — wynosi on $\sqrt{n} - 1$. Nietrudno więc stwierdzić, że dla $n \geqslant 10$ promień tej kuli będzie większy niż długość krawędzi n-wymiarowej kostki, która go zawiera. Innymi słowy sfera ta wystawać będzie poza kostkę! Wydaje się to dziwne.

W międzyczasie wspomniany już wcześniej Felix Haussdorff, sformułował w 1918 roku teorię wymiaru samopodobieństwa, przypisującą niektórym zbiorom wymiar ułamkowy, a nie tylko całkowity. Mandelbrot spopularyzował te zbiory w latach osiemdziesiątych jako fraktale. Pomysł był prosty, a podamy jedynie jego intuicję — obiekt wymiaru samopodobieństwa d to taki, który poddany jednokładności o skali k, zmienia miarę o czynnik k^d. Co to znaczy? Na razie niewiele wiemy o pojęciu miary, ale oto intuicja. Dla zwykłych obiektów typu odcinek czy kwadrat — wymiar samopodobieństwa działa tak jak zwykły wymiar. Odcinek powiększony trzykrotnie zwiększa długość 3^1 razy, a kwadrat po jednokładności w skali 3 zmienia pole na 3^2 razy większe. Sześcian po jednokładności o skali 3 zmienia objętość na 3^3 razy większą.

Rozważmy jednak np. tak zwaną krzywą Kocha, która konstruujemy poprzez kolejne iteracje. Zaczynamy od odcinka, z którego usuwamy środkową trzecią część (długości 1/3) i zastępujemy ją dwoma odcinkami o długości równej usuniętemu fragmentowi. Następnie z każdego z czterech powstałych w ten sposób odcinków usuwamy środkową trzecią część i zastępujemy dwiema. Powyższą procedurę kontynuujemy "w nieskończoność" (co da się formalnie opisać — ale to na razie zostawmy, podobne konstrukcje np. zbioru Cantora czy dywanu Sierpińskiego poznacie Państwo na topologii). Jaki ta krzywa ma wymiar?

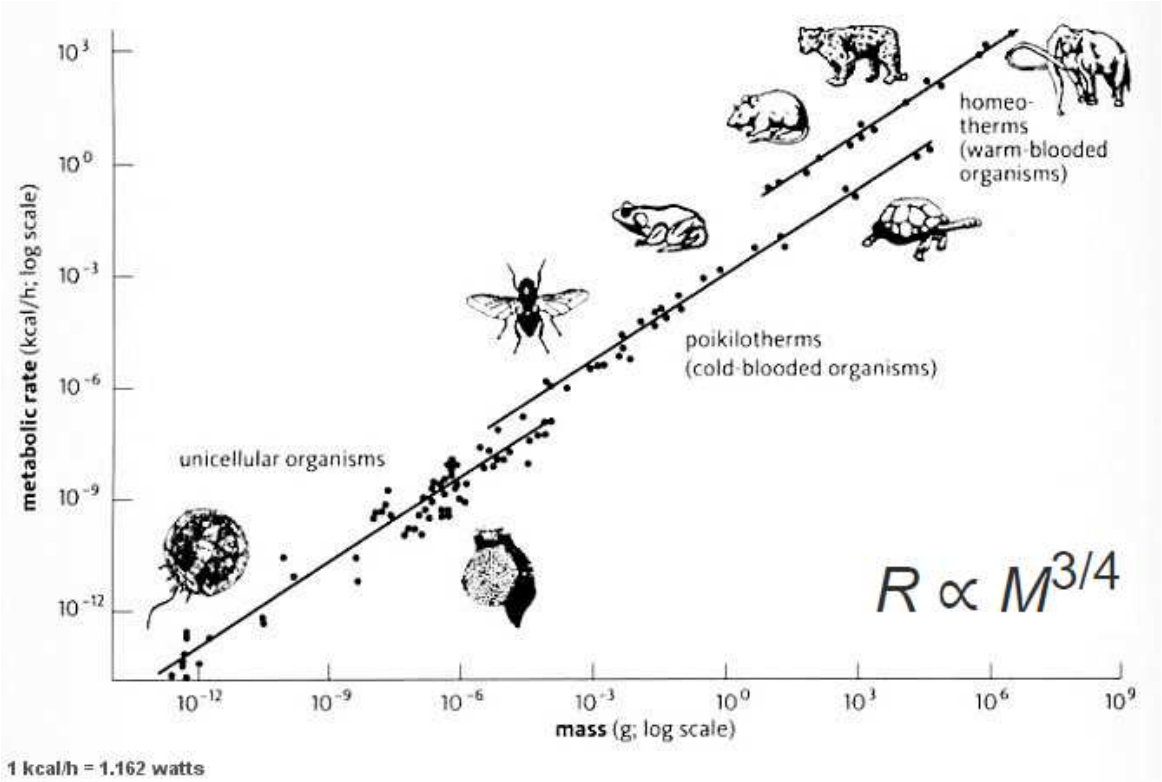


Rys 4. Pierwsze kilka przybliżeń krzywej Kocha. Źródło jak wyżej.

Uzyskana krzywa ma tę własność, że jeśli powiększymy ją trzykrotnie, otrzymamy cztery kopie wyjściowego obiektu. To znaczy, że wymiar Haussdorffa d tej krzywej spełnia równość $3^d = 4$, a więc $d = \log_3 4$.

Czytelnik może uważać podobną konstrukcję jedynie za ciekawostkę. Okazuje się jednak, że ma ona głębokie znaczenie dla współczesnego rozumienia wielu ważnych zależności w przyrodzie, finansach, statystyce i wielu innych dziedzinach.

Świetnym przykładem jest tak zwane prawo Kleibera z roku 1930, które mówi — w największym skrócie, że dla większości wyższych kręgowców metabolizm jest proporcjonalny do masy ciała podniesionej do potęgi 3/4. Co ciekawe, Kleiber uzyskał tę obserwację na podstawie szerokich empirycznych studiów rozmaitych gatunków, a mimo to uzyskany przezeń wykładnik 3/4 nie był intuicyjny. Dlaczego?



Rys 5. Prawo Kleibera.

Intuicyjnie rzecz biorąc, metabolizm, czyli zużycie energii, pochodzi głównie z potrzeby ogrzewania się, a na ogrzewanie główny wpływ ma powierzchnia ciała osobnika. Im większa powierzchnia — tym więcej tracimy ciepła. Im większa masa — tym więcej produkujemy ciepła. Przy takim wyjaśnieniu wydawałoby się, że wykładnik powinien wynosić 2/3, ponieważ (po wzięciu logarytmów) tyle wynosi stosunek między zmianą powierzchni, a zmianą objętości obiektu trójwymiarowego. Jedną z prób wyjaśnienia, dlaczego te wykładniki się różnią jest koncepcja fraktalnej budowy naszych organów krążenia — obejmujących płuca, układ żył i tętnic itd. Nosi ona znamiona struktury samopodobnej, która zdaje się wyjaśniać skąd pojawia się 3/4. Wykładnik ten pojawia się zresztą przy badaniu wielu innych pozornie niepowiązanych ze sobą zjawisk, nie tylko w biologii. Zainteresowanych tym tematem odsyłam do kanonicznego artykułu Jamesa Browna, Briana Enquista i Geoffrey Westa z 1999 roku: The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms: `https://www.santafe.edu/research/results/working-papers/the-fourth-dimension-of-life-fractal-geometry-and-`.

* * *

Pojęcie wymiaru przekroczyło dawno geometrię. Ujęcie prezentowane przez nas w duchu algebry liniowej dopuszcza o myśleniu o liczbach zespolonych, jako dwuwymiarowej przestrzeni liniowej nad ciałem liczb rzeczywistych, a o liczbach rzeczywistych — jako nieskończenie wymiarowej przestrzeni nad ciałem liczb wymiernych. To ujęcie uogólnione zostało na liczne struktury, takie jak pierścienie, grupy, a także na obiekty kombinatoryczne. Kluczem do sformułowania poprawnej algebraicznej definicji wymiaru jest zastanowienie się jakie będzie on miał własności ze względu na wzajemne położenie obiektów, które mierzy. Podobnie jak w algebrze liniowej, w żadnym ujęciu nie wyobrażamy sobie, by zbiór niskowymiarowy zawierał jako podzbiór zbiór wyżej wymiarowy. Chcemy mieć jakiś sensowny sposób przekształcania na siebie zbiorów o tym samym wymiarze. Chcemy wiedzieć jak wymiar może zachowywać się przy braniu części wspólnej. Niektóre z tych zagadnień, odniesionych do przestrzeni liniowych, rozważamy na kolejnym wykładzie. Inne — te dotyczące przekształceń, poznamy za kilka tygodni mówiąc o izomorfizmach.

Rozdział 7

Rząd macierzy
Twierdzenie Kroneckera-Capellego

7.1 Wykład siódmy

Do tej pory sporo miejsca poświęciliśmy abstrakcyjnym pojęciom przestrzeni liniowej, podprzestrzeni, układów liniowo niezależnych, bazy i wymiaru. Dziś chcielibyśmy odnieść te rezultaty do teorii układów równań liniowych, która w tym nowym języku otrzymuje bardzo elegancką postać. Nie chodzi przy tym jedynie o układy równań – w istocie celem najbliższych dwóch wykładów będzie przyjrzenie się strukturze podprzestrzeni przestrzeni współrzędnych K^n. Gdy bowiem dysponujemy już pewnymi nowymi obiektami, próbujemy formułować rozmaite wyniki klasyfikujące te przestrzenie. Celem dzisiejszego wykładu[1] jest zastosowanie pojęć wprowadzonych ostatnio do opisu podprzestrzeni przestrzeni liniowej K^n. Kluczowe będzie rozumienie następującej równoważności.

Obserwacja 7.1

Niech $\alpha_1, \ldots, \alpha_n$ będzie układem wektorów w przestrzeni K^m, gdzie

$$\alpha_1 = (a_{11}, \ldots, a_{m1}), \ldots, \alpha_n = (\alpha_{1n}, \ldots, \alpha_{mn}).$$

Wówczas następujące warunki są równoważne:

- wektor $\beta = (b_1, \ldots, b_m)$ jest kombinacją liniową wektorów $\alpha_1, \ldots, \alpha_n$,
- istnieją $s_1, s_2 \ldots, s_n \in K$ takie, że

$$s_1 \cdot \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + s_2 \cdot \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + s_n \cdot \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

- układ równań liniowych zmiennych $x_1, \ldots, x_n$ nad ciałem K o macierzy rozszerzonej

$$A_u = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]$$

  ma rozwiązanie (jest niesprzeczny).

Dowód wynika z definicji operacji dodawania wektorów i mnożenia przez skalar w przestrzeni liniowej macierzy. Ważny wniosek: kolumny dowolnej macierzy rozpinają podprzestrzeń złożoną z wektorów, które są rozwiązaniami układów równań, których macierzą współczynników jest ta macierz.

[1] Ostatnia aktualizacja: 22.11.2022 r.

Kluczowym na tym wykładzie będzie pojęcie rzędu macierzy. Oparte jest ono o prawdziwość następującego zaskakującego (w pewnym sensie) rezultatu.

Twierdzenie 7.1

Niech $A \in M_{m\times n}(K)$ oraz niech:

- $w(A) = \dim \operatorname{lin}(\alpha_1, \ldots, \alpha_m)$, gdzie $\alpha_1, \ldots, \alpha_m \in K^n$ są wierszami macierzy A,
- $k(A) = \dim \operatorname{lin}(\beta_1, \ldots, \beta_n)$, gdzie $\beta_1, \ldots, \beta_n \in K^m$ są kolumnami macierzy A.

Wówczas $\boldsymbol{w(A) = k(A)}$. Innymi słowy: dla każdej macierzy A maksymalna liczba liniowo niezależnych wierszy macierzy A jest równa maksymalnej liczbie liniowo niezależnych kolumn macierzy A.

Przykład:

$$A = \begin{bmatrix} 1 & 0 & 1 & 1 & 2 & 3 & 1 & 2 \\ 4 & 1 & 3 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} \in M_{2\times 8}(\mathbb{R}).$$

- $w(A) = \dim(\operatorname{lin}((1,0,1,1,2,3,1,2),(4,1,3,1,0,0,1,1))) = 2$,
- $k(A) = \dim(\operatorname{lin}((1,4),(0,1),(1,3),(1,1),(2,0),(3,0),(1,1),(2,1))) = 2$.

Dowód. Przypomnijmy (Obserwacja 4.7), że jeśli A' jest macierzą schodkową otrzymaną z A elementarnymi operacjami na wierszach oraz jeśli $\alpha'_1, \ldots, \alpha'_r$ to wszystkie niezerowe wiersze macierzy A', wówczas:

(i) $\operatorname{lin}(\alpha_1, \ldots, \alpha_m) = \operatorname{lin}(\alpha'_1, \ldots, \alpha'_r)$,

(ii) $\alpha'_1, \ldots, \alpha'_r$ jest bazą przestrzeni rozpiętej przez wiersze macierzy A.

Widzimy zatem, że $w(A) = r$. Załóżmy, że w macierzy schodkowej A' pierwsze niezerowe wyrazy w wierszach $\alpha'_1, \ldots, \alpha_r{}'$ znajdują się odpowiednio w kolumnach o indeksach $s_1 < s_2 < \ldots < s_r$. Pokażemy, że $\beta_{s_1}, \ldots, \beta_{s_r}$ stanowią bazę $\operatorname{lin}(\beta_1, \ldots, \beta_n)$. A zatem trzeba dowieść, że wektory te są liniowo niezależne oraz, że rozpinają podprzestrzeń kolumnową macierzy A.

Niech $a_1, \ldots, a_r \in K$ oraz $a_1\beta_{s_1} + \ldots + a_r\beta_{s_r} = 0$. Niech układ $\beta'_{s_1}, \ldots, \beta'_{s_r}$ powstaje z $\beta_{s_1}, \ldots, \beta_{s_r}$ przez wykonanie na A elementarnej operacji σ na wierszach $\begin{bmatrix} \beta_1 & \cdots & \beta_n \end{bmatrix} \xrightarrow{\sigma} \begin{bmatrix} \beta'_1 & \cdots & \beta'_n \end{bmatrix}$ to

$$a_1\beta'_{s_1} + \ldots + a_r\beta'_{s_r} = 0.$$

Czy to widać? Przy operacji elementarnej następuje albo zamiana współrzędnych wszystkich powyższych wektorów, albo przemnożenie współrzędnych każdego z powyższych wektorów przez stałą, albo dodanie do współrzędnych o numerze j współrzędnych o numerze i przemnożonych przez stałą. Ilustracja (dodanie do j-tego wiersza i-tego przemnożonego przez a):

$$a_1 \begin{bmatrix} b_{1s_1} \\ b_{2s_1} \\ \vdots \\ b_{is_1} \\ \vdots \\ b_{js_1} + a \cdot b_{is_1} \\ \vdots \\ b_{ms_1} \end{bmatrix} + \ldots + a_r \begin{bmatrix} b_{1s_r} \\ b_{2s_r} \\ \vdots \\ b_{is_r} \\ \vdots \\ b_{js_r} + a \cdot b_{is_r} \\ \vdots \\ b_{ms_r} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 + a \cdot 0 \\ \vdots \\ 0 \end{bmatrix}$$

Wniosek: układ $\beta_{s_1}, \ldots, \beta_{s_r}$ jest liniowo niezależny wtedy i tylko wtedy, gdy $\beta'_{s_1}, \ldots, \beta'_{s_r}$ jest liniowo niezależny. Wykonajmy operacje elementarne na wierszach A aż dostaniemy postać zredukowaną A''.

I teraz **kluczowy argument** całego rozumowania: kolumna s_i-ta β''_{s_i} macierzy zredukowanej A'' to i-ty wektor bazy standardowej ϵ_i przestrzeni K^m, czyli $a_1\epsilon_1 + \cdots + a_r\epsilon_r = 0$ Stąd $a_1 = a_2 = \ldots = a_r = 0$. A zatem $\beta_{s_1}, \ldots, \beta_{s_r}$ to układ liniowo niezależny. Czy jest to układ rozpinający $\operatorname{lin}(\beta_1, \ldots, \beta_n)$?

Niech β będzie dowolną kolumną macierzy A. Szukamy $a_1, \ldots, a_r$ takich, że $a_1\beta_{s_1} + \ldots + a_r\beta_{s_r} = \beta$. Dostajemy układ równań liniowych o macierzy rozszerzonej:

$$U = \left[\beta_{s_1} \quad \cdots \quad \beta_{s_r} \quad | \quad \beta\right] \qquad (*)$$

Sprowadzenie macierzy U do postaci zredukowanej U'' odbywa się przy pomocy tych samych operacji, które sprowadzają A do A'', a więc pierwsze r kolumn U'' to pierwsze r wektorów bazy standardowej.

$$\left[\beta_{s_1} \quad \cdots \quad \beta_{s_r} \quad | \quad \beta\right] \xrightarrow{\cdots} \left[\epsilon_1 \quad \cdots \quad \epsilon_r \quad | \quad \beta''\right] = U''$$

Analogicznie jak w rozumowaniu wyżej mamy:

$$a_1\epsilon_1 + \cdots + a_r\epsilon_r = \beta''.$$

Ale β'' jest kolumną macierzy A'' (bo była kolumną A), więc ma tylko pierwsze r niezerowych współrzędnych, co oznacza, że układ $(*)$ ma rozwiązanie. Zatem układ $\beta_{s_1}, \ldots, \beta_{s_r}$ rozpina $\operatorname{lin}(\beta_1, \ldots, \beta_n)$. A zatem jest to baza tej przestrzeni i ostatecznie $\dim \operatorname{lin}(\beta_1, \ldots, \beta_n) = k(A) = r$. □

Definicja 7.1: Rząd macierzy

Rzędem macierzy $A \in M_{m\times n}(K)$ nazywamy liczbę $\dim \operatorname{lin}(\alpha_1, \ldots, \alpha_m) = \dim \operatorname{lin}(\beta_1, \ldots, \beta_n)$, gdzie $\alpha_1, \ldots, \alpha_m \in K^n$ są wierszami macierzy A, zaś $\beta_1, \ldots, \beta_n \in K^m$ są kolumnami macierzy A. Rząd macierzy oznaczamy przez $\boldsymbol{r(A)}$.

Wniosek 7.1

Rząd macierzy A równy jest liczbie niezerowych wierszy po doprowadzeniu A do postaci schodkowej.

Przykłady:

$$r\begin{bmatrix}1 & 1 & 1\\ 2 & 2 & 2\\ 3 & 3 & 3\end{bmatrix} = 1, \quad r\begin{bmatrix}1 & 1\\ 1 & 2\\ 3 & 1\\ 0 & 0\end{bmatrix} = 2, \quad r\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} = 0.$$

Wykonywanie operacji elementarnych na wierszach nie zmienia rzędu macierzy. Podobnie jest oczywiście z **elementarnymi operacjami na kolumnach macierzy**: dodaniem do kolumny innej kolumny pomnożonej przez stałą, zamianą kolumn, przemnożeniem kolumny przez niezerowy skalar. Stosowanie obydwu typów operacji, zarówno wierszowych jak i kolumnowych, może uprościć wyznaczanie rzędu.

Przykład. Dla $n > 1$ policzyć rząd macierzy $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ postaci:

$$A = \begin{bmatrix}-n+1 & 1 & \ldots & 1 & 1\\ 1 & -n+1 & \ldots & 1 & 1\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 1 & 1 & \ldots & -n+1 & 1\\ 1 & 1 & \ldots & 1 & -n+1\end{bmatrix}.$$

Wykonujemy następujące operacje. 1. Do ostatniego wiersza dodajemy (kolejno) wszystkie pozostałe wiersze: 2. Odejmujemy ostatnią kolumnę (kolejno) od każdej z pozostałych kolumn: W ten sposób:

$$\begin{bmatrix}-n+1 & 1 & \ldots & 1 & 1\\ 1 & -n+1 & \ldots & 1 & 1\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 1 & 1 & \ldots & -n+1 & 1\\ 0 & 0 & \ldots & 0 & 0\end{bmatrix} = \begin{bmatrix}-n & 0 & \ldots & 0 & 1\\ 0 & -n & \ldots & 0 & 1\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \ldots & -n & 1\\ 0 & 0 & \ldots & 0 & 0\end{bmatrix}.$$

Otrzymaliśmy macierz w postaci schodkowej, która ma dokładnie $n-1$ niezerowych wierszy. A zatem rząd wyjściowej macierzy A równy jest $n-1$.

Poniższe twierdzenie jest jednym z głównych rezultatów tego wykładu.

Twierdzenie 7.2: Kroneckera-Capellego

Niech U będzie układem równań liniowych o współczynnikach w ciele K postaci:

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n & = b_1 \\ & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m. \end{cases}$$

o macierzy współczynników A oraz rozszerzonej macierzy współczynników A_u. Wówczas:

(a) Układ U ma rozwiązanie wtedy i tylko wtedy, gdy $r(A) = r(A_u)$,

(b) Przestrzeń rozwiązań układu jednorodnego odpowiadającego układowi U ma wymiar $n - r(A)$

(c) Jeśli α jest rozwiązaniem układu U, a W jest przestrzenią rozwiązań układu jednorodnego odpowiadającego układowi U, to zbiór rozwiązań układu U jest postaci

$$\alpha + W = \{\alpha + \beta, \ | \beta \in W\}.$$

Przypomnijmy pojęcia użyte w twierdzeniu. Macierze A oraz A_u układu wyżej to odpowiednio:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad A_u = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

Układ jednorodny odpowiadający układowi U to układ jednorodny o macierzy A.

Punkt (c) wykazaliśmy już w uzupełnieniu do pierwszego wykładu (Obserwacja 1.3). Punkty (a) i (b) opisują problem rozwiązywalności i „rozmiaru" zbioru rozwiązań układu równań liniowych w języku wymiaru. Również te fakty są dla nas w zasadzie intuicyjnie jasne. Wiemy bowiem, że układ równań może okazać się sprzeczny jedynie, gdy w wyniku sprowadzania macierzy A_u do postaci zredukowanej pojawi się wiersz postaci $[00\ldots0|1]$. Nietrudno będzie nam formalnie pokazać, na podstawie posiadanej już wiedzy, że sytuacja ta może wystąpić jedynie, gdy $r(A) < r(A_u)$.

Również punkt (b) jest intuicyjnie jasny. Nie pokazaliśmy jeszcze tego w sposób formalny, ale na podstawie wielu przykładów podejrzewamy, że wymiar przestrzeni rozwiązań jednorodnego układu równań równy jest liczbie zmiennych niezależnych tego układu. Ta zaś równa jest liczbie wszystkich zmiennych pomniejszonej o liczbę zmiennych zależnych. Wszystkich zmiennych jest n, zaś zmiennych zależnych jest tyle, co schodków macierzy A po sprowadzeniu jej, za pomocą elementarnych operacji wierszowych, do postaci zredukowanej. Tych schodków jest, jak wiemy, $r(A)$. W dowodzie opiszemy konstrukcję bazy zbioru rozwiązań układu U. Wcześniej jednak zobaczmy przykłady.

Przykład. Porównajmy układy równań o macierzy współczynników $A \in M_{2\times3}(\mathbb{R})$ oraz macierzach rozszerzonych A_u oraz B_u:

$$A_u = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{array}\right], \quad B_u = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 4 \end{array}\right], \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}.$$

Macierze współczynników obydwu tych układów mają rząd 1. Macierz A_u ma również rząd 1, natomiast macierz B_u ma rząd 2. Pierwszy układ jest w sposób oczywisty niesprzeczny. Również zgodnie z punktem (a) powyższego twierdzenia układ ten ma rozwiązanie, bo rząd macierzy współczynników tego układu jest równy rzędowi macierzy rozszerzonej. W drugim przypadku rząd macierzy współczynników jest mniejszy — i układ niejednorodny nie ma rozwiązań. Jak znajdujemy rozwiązanie układu o macierzy A_u? Rozwiązujemy odpowiadający mu układ jednorodny o macierzy (współczynników) A. Na mocy punktu (b) twierdzenia wyżej wiemy, że wymiar W przestrzeni rozwiązań układu równań jednorodnych równy jest 3 (liczba niewiadomych) -1 (rząd macierzy współczynników), czyli 2. Wreszcie, na mocy punktu 3, możemy stwierdzić, że po wyznaczeniu $W = \text{lin}((1,-1,0),(1,0,-1))$ wystarczy wziąć dowolne rozwiązanie układu o macierzy A_u, na przykład $\alpha = (1,0,0)$ i widzimy, że zbiór rozwiązań całego układu niejednorodnego o macierzy rozszerzonej A_u ma postać podzbioru $\mathbb{R}^3$ (ale nie podprzestrzeni) postaci:

$$(1,0,0) + \text{lin}((1,-1,0),(1,0,-1)).$$

Ważny przykład. Niech $(a_1, \ldots, a_n) \in K^n$, gdzie $a_1 \neq 0$. Wówczas zbiór rozwiązań równania

$$a_1x_1 + \ldots + a_nx_n = 0$$

jest podprzestrzenią wymiaru $n-1$ postaci:

$$\operatorname{lin}\left((-\frac{a_2}{a_1}, 1, 0, \ldots, 0), \ldots, (-\frac{a_n}{a_1}, 0, 0, \ldots, 1)\right).$$

Rzeczywiście, sprowadzamy macierz układu do postaci zredukowanej:

$$(a_1 \quad a_2 \quad \ldots \quad a_n) \sim \begin{pmatrix} 1 & \frac{a_2}{a_1} & \ldots & \frac{a_n}{a_1} \end{pmatrix}.$$

A zatem rozwiązanie ogólne tego układu zadane jest przez:

$$x_1 = -\frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3 - \ldots - \frac{a_n}{a_1}x_n.$$

Zmienne $x_2, \ldots, x_n$ stanowią $n-1$ parametrów i wszystkie rozwiązania tego układu są postaci:

$$(-\frac{a_2}{a_1}t_2 - \ldots - \frac{a_n}{a_1}t_n, \underbrace{t_2, t_3, \ldots t_n}_{\text{parametry}}), \qquad \text{gdzie } t_2, t_3, \ldots, t_n \in K.$$

W szczególności rozwiązania te są postaci:

$$t_2(-\frac{a_2}{a_1}, 1, 0, \ldots, 0) + t_3(-\frac{a_3}{a_1}, 0, 1, \ldots, 0) + \ldots + t_n(-\frac{a_n}{a_1}, 0, 0, \ldots, 1).$$

Dowodzimy twierdzenie Kroneckera-Capellego. Punkt (a) jest jasny, na mocy Obserwacji 7.1. Weźmy $\alpha_1, \ldots, \alpha_n, \beta \in K^n$, które są kolumnami macierzy A_u. Wówczas mamy ciąg równoważnych stwierdzeń:

$$\begin{aligned} x_1, \ldots, x_n \text{ jest rozwiązaniem układu } U &\iff x_1\alpha_1 + \ldots + x_n\alpha_n = \beta \\ &\iff \beta \in \operatorname{lin}(\alpha_1, \ldots, \alpha_n) \\ &\iff \operatorname{lin}(\alpha_1, \ldots, \alpha_n) = \operatorname{lin}(\alpha_1, \ldots, \alpha_n, \beta), \\ &\iff dim\operatorname{lin}(\alpha_1, \ldots, \alpha_n) = \dim\operatorname{lin}(\alpha_1, \ldots, \alpha_n, \beta) \\ &\iff r(A) = r(A_u). \end{aligned}$$

Niech U' będzie układem jednorodnym odpowiadającym układowi U. Dowód punktu (b) polega na zauważeniu, że jeśli $r(A) = r$, to macierz A' uzyskana z A przez sprowadzenie do postaci schodkowej ma dokładnie r niezerowych wierszy, a zatem postać ogólna rozwiązania tego układu ma $n-r$ parametrów. Załóżmy, że zmienne zależne to $x_{j_1}, \ldots, x_{j_r}$, a owe parametry to zmienne $x_{t_1}, \ldots, x_{t_{n-r}}$. Innymi słowy, rozwiązanie ogólne układu U' ma, dla pewnych $c_{ij} \in K$, postać:

$$\begin{cases} x_{j_1} & = c_{11}x_{t_1} + c_{12}x_{t_2} + \ldots + c_{1,n-r}x_{t_{n-r}} \\ & \vdots \\ x_{j_r} & = c_{r1}x_{t_1} + c_{r2}x_{t_2} + \ldots + c_{r,n-r}x_{t_{n-r}} \end{cases} \qquad (*)$$

Rozważmy układ $n-r$ wektorów $\alpha_1, \ldots, \alpha_{n-r}$ takich, że α_j jest rozwiązaniem powyższego układu powstałym przez wstawienie za j-ty parametr 1, a za pozostałe parametry – zera. Innymi słowy, jeśli $\alpha_j = (a_{j1}, \ldots, a_{jn})$, to

$$a_{jt_1} = 0, \quad \ldots \quad , a_{jt_j} = 1, \quad \ldots \quad , a_{jt_{n-k}} = 0.$$

Oczywiście wektory $\alpha_1, \ldots, \alpha_{n-r}$ istnieją, bo układ $(*)$ ma jednoznaczne rozwiązane dla każdego z góry zadanego układu parametrów. Twierdzimy, że układ ten jest bazą zbioru rozwiązań układu U'.

Dowodzimy, że układ $\alpha_1, \ldots, \alpha_{n-r}$ jest liniowo niezależny. Istotnie, jeśli dla $b_1, \ldots, b_{n-r} \in K$ mamy $b_1\alpha_1 + \ldots + b_{n-r}\alpha_{n-r} = 0$, to dla i-tych współrzędnych $a_{1i}, \ldots, a_{n-r,i}$ wektorów $\alpha_1, \ldots, \alpha_{n-r}$ mamy

$$b_1a_{1i} + \ldots + b_{n-r}a_{n-r,i} = 0.$$

Zatem biorąc i równe kolejno $t_1, \ldots t_{n-r}$ dostajemy $b_1 = \ldots = b_{n-r} = 0$.

Układ $\alpha_1, \ldots, \alpha_{n-r}$ rozpina zbiór rozwiązań układu $(*)$. Jest bowiem jasne, że wektor

$$\alpha = (a_1, \ldots, a_n)$$

spełnia układ $(*)$ wtedy i tylko wtedy, gdy każda z jego współrzędnych jest kombinacją liniową współrzędnych $a_{t_1}, \ldots, a_{t_n}$. Inaczej mówiąc α jest rozwiązaniem układu U' wtedy i tylko wtedy, gdy

$$\alpha = a_{t_1}\alpha_1 + \ldots + a_{t_{n-r}}\alpha_{n-r}.$$

Zatem zbiór rozwiązań układu U' równy jest $\operatorname{lin}(\alpha_1, \ldots, \alpha_{n-r})$, co kończy dowód (b) i całego twierdzenia.

Przykład. Rozwiązanie ogólne układu U' zmiennych x_1, x_2, x_3, x_4 o współczynnikach w $\mathbb{R}$ postaci:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_1 - x_2 + x_3 + x_4 &= 0 \end{cases}$$

ma $4 - 2$ parametry x_3, x_4 (bo rząd macierzy U' to 2). Każde rozwiązanie U' jest postaci:

$$(-s - t, 0, s, t) = s(-1, 0, 1, 0) + t(-1, 0, 0, 1), \quad s, t \in \mathbb{R},$$

i powstaje przed odpowiedni wybór s, t na współrzędnych $t_1 = 3, t_2 = 4$. W zatem w tym przykładzie:

$$\alpha_1 = (a_{11}, a_{12}, a_{13}, a_{14}) = (-1, 0, 1, 0), \quad \alpha_2 = (a_{21}, a_{22}, a_{23}, a_{24}) = (-1, 0, 0, 1).$$

Definicja 7.2: Podprzestrzeń opisana układem równań

Jeśli $V \subseteq K^n$ jest przestrzenią rozwiązań jednorodnego układu równań liniowych U, to mówimy, że przestrzeń V jest OPISANA UKŁADEM U.

Możemy teraz sformułować ważny wniosek stanowiący twierdzenie klasyfikacyjne.

Wniosek 7.2

Każda podprzestrzeń V przestrzeni K^n jest opisana pewnym jednorodnym układem równań liniowych U. Jeśli $\dim V = k$, to można tak dobrać ten układ U, by składał się z $n - k$ równań. Dla $\dim V = k$ oraz $i < n - k$ nie istnieje złożony z i równań układ równań liniowych opisujący V.

Dowód tego twierdzenia zawiera w sobie istotny algorytm opisu podprzestrzeni za pomocą układu równań. Ponownie przeprowadzimy jego ilustrację najpierw na przestrzeni rozpiętej przez pojedynczy wektor.

Ważny przykład. Niech $\alpha = (a_1, \ldots, a_n) \in K^n$, gdzie $a_1 \neq 0$. Wówczas przestrzeń $\operatorname{lin}(\alpha)$ opisana jest przez układ równań postaci:

$$\begin{cases} -\frac{a_2}{a_1}x_1 + x_2 &= 0 \\ -\frac{a_3}{a_1}x_1 + x_3 &= 0 \\ & \vdots \\ -\frac{a_n}{a_1}x_1 + x_n &= 0 \end{cases}.$$

Czytelnik dostrzega z pewnością związek pomiędzy wypisanymi równaniami, a wektorami uzyskanymi wcześniej jako rozwiązania równania $a_1x_1 + \ldots + a_nx_n = 0$. Ta dualność nie jest przypadkowa. W istocie, rozwiązać należy ten sam układ, przy czym tym razem szukamy takich n-tek współczynników $(t_1, \ldots, t_n)$, aby równanie liniowe

$$t_1x_1 + \ldots + t_nx_n = 0$$

miało z góry zadane rozwiązanie $a_1, \ldots, a_n$. Zbiór wszystkich takich $(t_1, \ldots, t_n)$ jest podprzestrzenią w przestrzeni K^n, której bazą są wektory $(-\frac{a_2}{a_1}, 1, 0, \ldots, 0), \ldots, (-\frac{a_n}{a_1}, 0, 0, \ldots, 1)$.

Analogiczna sytuacja jest dla układów równań. Jeśli $(s_{11}, \ldots, s_{1n}), \ldots, (s_{m1}, \ldots, s_{mn})$ są rozwiązaniami układu

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n & = 0 \\ \cdots & \\ a_{k1}x_1 + \ldots + a_{kn}x_n & = 0 \end{cases},$$

to $(a_{11}, \ldots, a_{1n}), \ldots, (a_{k1}, \ldots, a_{kn})$ są rozwiązaniami układu:

$$\begin{cases} s_{11}x_1 + \ldots + s_{1n}x_n & = 0 \\ \cdots & \\ s_{m1}x_1 + \ldots + s_{mn}x_n & = 0 \end{cases}.$$

Jeśli $(s_{11}, \ldots, s_{1n}), \ldots, (s_{m1}, \ldots, s_{mn})$ jest bazą przestrzeni V rozwiązań układu o macierzy schodkowej

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix},$$

to na mocy tw. Kroneckera-Capellego $m = n - k$, co więcej wektory $(a_{11}, \ldots, a_{1n}), \ldots, (a_{k1}, \ldots, a_{kn})$ są liniowo niezależne i są rozwiązaniami układu o macierzy rzędu m postaci

$$\begin{bmatrix} s_{11} & s_{12} & \ldots & s_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \ldots & s_{mn} \end{bmatrix}.$$

Ponownie na mocy tw. Kroneckera-Capellego, powyższy układ ma przestrzeń rozwiązań wymiaru

$$n - m = n - (n - k) = k.$$

Czyli jest to $\operatorname{lin}((a_{11}, \ldots, a_{1n}), \ldots, (a_{k1}, \ldots, a_{kn}))$ – przestrzeń wymiaru k.

Wykazaliśmy zatem, że jeśli $\alpha_1, \ldots, \alpha_k$ jest bazą przestrzeni $V \subseteq K^n$ oraz $A \in M_{k\times n}(K)$ jest macierzą o wierszach $\alpha_1, \ldots, \alpha_k$, to przestrzeń V można opisać układem dowolnych $n - k$ równań, których współczynniki tworzą bazę przestrzeni rozwiązań układu danego macierzą A. Równań tych nie może być oczywiście mniej, bowiem przestrzeń rozwiązań układu o mniej niż $n-k$ równaniach ma wymiar większy niż $n - (n - k)$, na mocy twierdzenia Kroneckera-Capellego.

Przykład. Rozważmy $V = \operatorname{lin}((1, 2, 0, 1, 0), (0, 0, 1, 1, 1)) \subseteq \mathbb{R}^5$. Rozwiązania układu równań o macierzy

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

stanowi przestrzeń współczynników wszystkich równań liniowych, których rozwiązania zawierają V. Wybierając różne bazy tej przestrzeni dostajemy różne (ale równoważne) układy równań opisujące V, na przykład dla bazy

$$(-2, 1, 0, 0, 0), (-1, 0, -1, 1, 0), (0, 0, -1, 0, 1)$$

mamy następujący układ opisujący V:

$$\begin{cases} -2x_1 + x_2 & = 0 \\ -x_1 - x_3 + x_4 & = 0 \\ -x_3 + x_5 & = 0 \end{cases}.$$

Wniosek 7.3

Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A) = n$.

Dowód. Układ U ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A)$ oraz $\dim W = 0$, co jest równoważne $r(A_u) = r(A) = n$. □

Dokonaliśmy zatem klasyfikacji podprzestrzeni w K^n i umiemy je wyrażać zarówno jako przestrzenie rozpięte przez dany układ wektorów, jak i jako przestrzenie opisane przez określony układ równań.

Zobaczmy nietrywialne zastosowanie pojecia rzędu i twierdzenia Kroneckera-Capellego, rozwiązując następujące zadanie z kolokwium z roku 2021 (jeden z podpunktów).

Zadanie. Niech K będzie ciałem oraz niech $n > 1$. Podzbiór $\mathcal{S}$ w przestrzeni $M_{n\times n}(K)$ złożony jest ze wszystkich macierzy $S = (s_{ij})$ o wyrazach ze zbioru $\{0, 1\}$, spełniających warunki

$$s_{ii} = 0, \text{ dla } i = 1, 2, \ldots, n \quad \text{oraz} \quad s_{ij} + s_{ji} = 1, \text{ dla dowolnych } 1 \leqslant i < j \leqslant n.$$

Niech $K = \mathbb{Z}_2$. Rozstrzygnij czy $\mathcal{S}$ jest podprzestrzenią $M_{n\times n}(K)$ oraz wyznacz $\dim \operatorname{lin}(\mathcal{S})$.

Rozwiązanie. Zauważmy, że macierz zerowa nie należy do $\mathcal{S}$, więc nie jest to podprzestrzeń $M_{n\times n}(K)$.

Pokażemy, że dla dowolnego ciała K wymiar przestrzeni $\operatorname{lin}(\mathcal{S}) \subset M_{n\times n}(K)$ to

$$\frac{n(n-1)}{2} + 1.$$

Zaczniemy od intuicji, a potem przedstawimy formalny dowód. Każda macierz z $\operatorname{lin}(\mathcal{S})$ jest, niezależnie od ciała, wyznaczona jednoznacznie przed określenie czy poszczególne wyrazy pod przekątną równe są 0 czy 1 oraz przez określenie dowolnego wyrazu nad przekątną. Istotnie, wyrazy dowolnej macierzy $A = (a_{ij}) \in \mathcal{S}$ spełniają, dla $i \neq j$, warunki:

$$a_{12} + a_{21} = a_{ij} + a_{ji} = 1,$$

a zatem równości $c_{ji} = c_{ij} - c_{12} - c_{21}$, dla $j > i$, zachodzą dla każdej macierzy $C = (c_{ij})$ z $\operatorname{lin}(\mathcal{S})$. Intuicyjnie zatem układ opisujący macierze z $\operatorname{lin}(\mathcal{S})$ zależy od $\frac{n(n-1)}{2} + 1$ parametrów i to jest właśnie wymiar $\operatorname{lin}(\mathcal{S})$.

Przejdźmy do dowodu (przykładowego, bo wychodząc z powyższych intuicji można wskazać bazę $\operatorname{lin}(\mathcal{S})$). Weźmy $C \in \operatorname{lin}(\mathcal{S})$. Mamy:

$$C = a_1 S_1 + a_2 S_2 + \ldots + a_k S_k,$$

gdzie $a_1, \ldots, a_k \in K$ oraz $S_1, \ldots S_k \in \mathcal{S}$. A zatem wyrazy macierzy C spełniają warunki

$$c_{ij} + c_{ji} = a_1 + a_2 + \ldots + a_k.$$

Oznacza to, że dla $c = a_1 + a_2 + \ldots + a_k$ wyrazy c_{ij} macierzy C są rozwiązaniami układu U_c postaci:

$$U_c : \begin{cases} c_{11} & = 0 \\ \quad \vdots & \\ c_{nn} & = 0 \\ c_{12} + c_{21} & = c \\ \quad \vdots & \\ c_{n-1,n} + c_{n,n-1} & = c. \end{cases}$$

Innymi słowy, układ U_c jest złożony z:

- n równań postaci $c_{11} = 0, \ldots, c_{nn} = 0$,
- $\frac{n(n-1)}{2}$ równań postaci $c_{ij} + c_{ji} = c$.

Rząd macierzy jednorodnego układu U_0 wynosi $n + \frac{n(n-1)}{2}$, bo każda z n^2 niewiadomych występuje tylko w jednym równaniu. Stąd wymiar przestrzeni rozwiązań W_0 układu U_0 to, zgodnie z tw. Kroneckera-Capellego

$$\dim W_0 = n^2 - n - \frac{n(n-1)}{2} = \frac{n(n-1)}{2}.$$

Wiemy z wykładu, że zbiór wszystkich macierzy spełniających niejednorodny układ U_1 ma tę własność, że różnica dowolnych dwóch macierzy z tego zbioru spełnia układ jednorodny U_0, czyli jest w W_0. Co więcej, jeśli dla $c \neq 0$ macierz M spełnia U_c, to macierz $c^{-1} \cdot M$ spełnia U_1. Oznacza to, że biorąc bazę W_0 oraz dowolną macierz $C \in \operatorname{lin}(\mathcal{S}) \setminus W_0$ dostajemy bazę $\operatorname{lin}(\mathcal{S})$. W szczególności:

$$\dim \operatorname{lin}(\mathcal{S}) = \dim W_0 + 1 = \frac{n(n-1)}{2} + 1.$$

■

7.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Obliczanie rzędów macierzy) Wyznacz rząd macierzy rzeczywistych:

$$\begin{bmatrix} 1 & -1 & 0 & 3 \\ 2 & -3 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 4 & 0 & -2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -2 & -1 & 3 & 1 \\ 2 & 0 & -5 & 3 & -4 \\ -3 & -1 & -2 & 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1+i & -1 & 3-2i \\ 0 & 2+3i & -1 \\ 0 & 5-i & 9 \\ -1 & 8 & 7i \end{bmatrix}.$$

   Wyznacz rzędy macierzy rzeczywistych, w zależności od parametrów $a, b, c \in \mathbb{R}$ oraz $s, t \in \mathbb{R}$:

$$\begin{bmatrix} a & -b & 1 \\ b & a & 1 \\ 1 & 1 & c \end{bmatrix}, \quad \begin{bmatrix} 10 & -1 & -1 & 3 \\ 2s & -3 & 2 & 1 \\ 4 & 2 & t+3 & 3 \\ 0 & -2 & 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 3 & 4 & 2 & 1 & 2 \\ 2 & 3 & 1 & 4 & 6 \\ 1 & 2 & t^2-2t & 7 & 10 \\ 4 & 5 & 3 & -t & -2 \end{bmatrix}$$

2. (♠ Znajdowanie wymiaru przestrzeni rozwiązań układu jednorodnego)
   Dla każdego $t \in \mathbb{R}$ niech V_t będzie przestrzenią rozwiązań układu równań o współczynnikach w $\mathbb{R}$

$$\begin{cases} x_1 - x_2 + 2x_4 = 0 \\ x_1 + 2x_2 + 3x_3 + (2-t)x_4 = 0 \\ tx_1 + tx_2 + tx_3 + tx_4 = 0 \end{cases}.$$

   Dla każdego $t \in \mathbb{R}$ znajdź wymiar przestrzeni V_t.

3. Niech V będzie podprzestrzenią przestrzeni $\mathbb{R}^5$ opisaną układem równań

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 + x_5 = 0 \\ 2x_1 + 4x_2 - 2x_3 + 2x_4 + ax_5 = 0 \end{cases}.$$

   Wyznacz wymiar V w zależności od parametru a.

4. Znajdź bazę i wymiar przestrzeni rozwiązań układu $n+1$ równań liniowych o $2n$ zmiennych i współczynników rzeczywistych:

$$\begin{cases} x_1 + x_2 + \ldots + x_n = 0 \\ x_2 + x_3 + \ldots + x_{n+1} = 0 \\ \qquad\qquad \ddots \\ x_{n+1} + x_{n+2} + \ldots + x_{2n} = 0 \end{cases}.$$

5. (♠ Opisywanie podprzestrzeni rozpiętej na układzie wektorów układem równań)
   Dla każdej z poniższych podprzestrzeni $W \subset \mathbb{R}^n$ znaleźć opisujący ją układ równań liniowych.
   a) $W = \mathrm{lin}((3,1,2,-1),(4,2,1,5),(5,5,4,3)) \subset \mathbb{R}^4$,
   b) $W = \mathrm{lin}((4,1,2,-3),(2,3,1,-9),(2,-1,1,3),(6,4,3,-12)) \subset \mathbb{R}^4$,
   c) $W = \mathrm{lin}((5,1,9,0,2),(5,2,-2,5,-1),(4,1,5,1,1)) \subset \mathbb{R}^5$.

6. Niech $W = \mathrm{lin}((7,9,6,8),(11,u,12,u+1),(2,1,3,2),(3,-4,9,2)) \subset \mathbb{R}^4$. Dla jakich $u \in \mathbb{R}$ można podprzestrzeń W opisać jednym równaniem liniowym?

7. Niech macierz B powstaje z macierzy A poprzez wykreślenie s wierszy i t kolumn. Wykaż, że

$$r(A) \leqslant s + t + r(B).$$

8. Niech A, B będą macierzami rozmiaru $m \times n$ o wyrazach z ciała K. Wykaż, że

$$r(A+B) \leqslant r(A) + r(B).$$

9. Wykaż, że jeśli $A = [a_{ij}] \in M_{3\times 3}(\mathbb{R})$ jest niezerową macierzą spełniającą warunek $a_{ij} = -a_{ji}$, dla każdych i, j, to $r(A) = 2$.

10. Niech $n \geqslant 1$ oraz $a_1, \ldots, a_n \in \mathbb{R}$. Rozważmy macierz A rozmiaru $n! \times n$, której wierszami są wszystkie możliwe permutacje ciągu $a_1, \ldots, a_n$. Wyznacz możliwe wartości liczby $r(A)$.

7.3 Uzupełnienie. Rekurencje liniowe. Wzór Bineta.

Czy istnieje odpowiednik pojęcia jednorodnego układu równań liniowych w przypadku przestrzeni nieskończonego wymiaru? Choć często nie myślimy o tym w ten sposób – jest pojęcie bliskie tej intuicji.

Definicja 7.3

Liniowym jednorodnym równaniem rekurencyjnym rzędu k (krócej: rekurencją rzędu k) o współczynnikach nad ciałem K nazywamy równanie postaci:

$$x_{n+k} = c_1 x_{n+k-1} + c_2 x_{n+k-2} + \ldots + c_k x_n, \tag{7.1}$$

gdzie $c_1, \ldots, c_k \in K$. Rozwiązaniem powyższej rekurencji jest dowolny ciąg $(s_0, s_1, \ldots) \in K^\infty$ spełniający równości (7.1) dla każdego $n \geqslant 0$, nazywany ciągiem rekurencyjnym rzędu k.

Nie sposób opisać znaczenia rekurencji dla różnych działów matematyki, zwłaszcza dla matematyki dyskretnej, teorii funkcji tworzących, teorii szeregów (np. kryterium wymierności funkcji dającej się rozpisać w szereg), liniowych równań różniczkowych itd. ale nas rekurencje interesują jako swego rodzaju układy nieskończenie wielu jednorodnych równań o zmiennych ze zbioru $X = \{x_0, \ldots\}$, przy czym w każdym z równań występuje skończona (i jednakowa) liczba zmiennych.

Nietrudno uświadomić sobie, że zbiór rozwiązań rekurencji liniowej (7.1) tworzy podprzestrzeń liniową w K^∞. Jeśli ciągi $a = (a_1, a_2, \ldots)$ oraz $b = (b_1, b_2, \ldots)$ są rozwiązaniami (7.1), to także ciągi $a+b$ oraz λa są w sposób oczywisty jej rozwiązaniami, dla każdego $\gamma \in K$. Co ciekawego można powiedzieć o przestrzeni rozwiązań rekurencji rzędu k? Na ten temat można opowiedzieć kilka semestralnych wykładów, ale ograniczmy się do kilku uwag na temat najsłynniejszej zapewne rekurencji postaci:

$$x_{n+2} = x_{n+1} + x_n$$

Jednym z rozwiązań tej rekurencji jest słynny ciąg Fibonacciego. Wystarczy określić $s_0 = 0, s_1 = 1$. W istocie: każde rozwiązanie powyższej rekurencji jest wyznaczone jednoznacznie przez pierwsze dwa elementy. Naśladując język wprowadzony na pierwszym wykładzie: przez operacje elementarne na układzie zadanym przez rekurencję wszystkie równości sprowadzić można do równań postaci $x_m = f_m(s_0, s_1)$, gdzie f_m jest pewną liniową zależnością wiążącą s_0 i s_1. A więc s_0 i s_1 są parametrami w „rozwiązaniu ogólnym" tego układu. Wszystkie te intuicje można przerobić na formalne rozumowanie i pokazać, że wymiar podprzestrzeni W opisanej rekurencją $x_{n+2} = x_{n+1} + x_n$ wynosi 2. Przykładowa baza W to

$$(0, 1, 1, 2, 3, ...), \quad (1, 0, 1, 1, 2, \ldots).$$

Jeden z jej elementów to oczywiście ciąg Fibonacciego. Zachodzi następujące twierdzenie, będące (przynajmniej intuicyjnie) uogólnieniem powyższych obserwacji.

Twierdzenie 7.3

Zbiór rozwiązań rekurencji liniowej rzędu k ma wymiar k.

Jak wiadomo ciąg Fibonacciego można opisać ogólnym wzorem postaci:

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

To dość niezwykłe, że tak prosta rekurencja opisywalna jest w sposób jawny tak skomplikowanym wzorem. Czy aby na pewno skomplikowanym? Okazuje się, że jego pochodzenie można łatwo wyjaśnić narzędziami algebry liniowej, choć rozumowanie, które pokażę niżej można jeszcze zdecydowanie uprościć. Jak wiemy przestrzeń rozwiązań rekurencji $x_{n+2} = x_{n+1} + x_n$ jest dwuwymiarowa i powyżej podaliśmy przykładową jej bazę. Ale może istnieje „ładniejsza" baza, pozwalająca opisywać elementy tej przestrzeni? Pomysł polega na poszukiwaniu bazy złożonej z ciągów postaci

$$\alpha = (a^0, a^1, a^2, \ldots), \quad \beta = (b^0, b^1, b^2, \ldots),$$

dla pewnych $a, b \in K$, które należą do W i są liniowo niezależne. Nie dla każdej rekurencji postaci (7.1) da się taką bazę znaleźć, ale w przypadku rozważanej przez nas rekurencji rzędu 2 jest to możliwe (to się wiąże z zagadnieniami, które dokładniej omawiać będziemy podczas rozważania teorii endomorfizmów i wyprowadzania twierdzenia Jordana). Przekonajmy się wyznaczając tę bazę.

Wyznaczymy szukaną bazę α, β przestrzeni W jakby „od tyłu", zapisując ciąg $F = (F_0, F_1, F_2, \ldots)$ w tej bazie. A zatem mamy równanie

$$F = c_1\alpha + c_2\beta \in K^\infty.$$

W szczególności mamy układ równań

$$\begin{aligned} c_1 + c_2 &= F_0 \\ c_1 a + c_2 b &= F_1 \\ c_1 a^2 + c_2 b^2 &= F_2 \\ c_1 a^3 + c_2 b^3 &= F_3 \\ &\vdots, \end{aligned}$$

czyli w rezultacie otrzymujemy równości

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 a + c_2 b &= 1 \\ c_1 a^2 + c_2 b^2 &= 1 \\ c_1 a^3 + c_2 b^3 &= 2 \\ &\vdots, \end{aligned}$$

A zatem $c_1 = -c_2$, czyli $c_1 a - c_1 b = 1, c_1 a^2 - c_1 b^2 = 1$, czyli $c_1(a-b) = c_1(a-b)(a+b) = 1$, czyli $a + b = 1$ itd. Ostatecznie:

$$c_1 = \frac{1}{\sqrt{5}}, \quad c_2 = -\frac{1}{\sqrt{5}}, \quad a = \frac{1+\sqrt{5}}{2}, \quad b = \frac{1-\sqrt{5}}{2}.$$

A zatem wzór na F_n pochodzi od rozpisania n-tej współrzędnej F jako kombinacji liniowej elementów bazowych α, β, czyli $F_n = c_1\alpha^n + c_2\beta^n$. Pozostaje oczywiście sprawdzić, że tak uzyskane ciągi spełniają założenia, to znaczy: należą do W i są liniowo niezależne. Pierwsza obserwacja jest jasna, bo:

$$\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} = \left(\frac{1+\sqrt{5}}{2}\right)^{n-1}\left(\frac{1+\sqrt{5}}{2}\right)^{2} = \left(\frac{1+\sqrt{5}}{2}\right)^{n-1}\left(1+\frac{1+\sqrt{5}}{2}\right) = \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} + \left(\frac{1+\sqrt{5}}{2}\right)^{n},$$

czyli $a^{n+1} = a^n + a^{n-1}$. Podobnie pokazujemy, że $b^{n+1} = b^n + b^{n-1}$. To, że ciągi α, β są liniowo niezależne to łatwe ćwiczenie. A zatem ciągi α, β są rzeczywiście bazą W i zachodzi wzór na ciąg Fibonacciego[2].

Czytelnika nie do końca przekonanego skąd rzeczywiście wzięły się w rozwiązaniu liczby $(1 \pm \sqrt{5})/2$ polecam nieco ogólniejsze spojrzenie. Rozważmy ciąg (x_n) z K^n spełniający warunki

$$x_{n+2} + a x_{n+1} + b x_n = 0.$$

Interesują nas niezerowe ciągi geometryczne określone wzorem $t_n = q^n$ spełniające to równanie. A zatem iloraz $q \neq 0$ tych ciągów spełnia równanie $q^{n+2} + a q^{n+1} + b q^n = 0$, czyli

$$q^2 + aq + b = 0.$$

Równanie $x^2 + ax + b = 0$ nazywane jest zwykle RÓWNANIEM CHARAKTERYSTYCZNYM równania rekurencyjnego $x_{n+2} + a x_{n+1} + b x_n = 0$. W zależności od tego, czy równanie charakterystyczne ma jedno, czy dwa rozwiązania, postać ogólna ciągów spełniających wyjściowe równanie jest inna. Jeśli równanie to ma dwa różne rozwiązania p oraz q, to wyjściową rekurencję spełnia ciąg $x_n = c \cdot p^n + d \cdot q^n$, gdzie współczynniki c, d wyznaczamy z układu równań utworzonego przez wstawienie $n = 0$ i $n = 1$ do rozwiązania ogólnego. Czytelnik domyśla się zapewne, że również liniowa rekurencja rzędu k ma równanie charakterystyczne rzędu k i odpowiednio skomplikowane zbiory rozwiązań. Zainteresowanych tym tematem oraz elementarnymi zastosowaniami ciągów rekurencyjnych zachęcam do lektury tekstu prof. Wojciecha Guzickiego: „Równania rekurencyjne": `https://www.mimuw.edu.pl/~guzicki/materialy/Rekurencja.pdf`.

[2] Wzór ten, zwany też wzorem Bineta, znany był już w XVIII wieku Bernoullemu, Eulerowi czy de Moivre'owi.

7.4 Uzupełnienie. Kilka uwag o prostopadłości

Zbiór rozwiązań liniowego równania jednorodnego $a_1x_1 + a_2x_2 = 0$ w $\mathbb{R}^2$ to prosta przechodząca przez punkt $(0,0)$, o ile tylko $(a_1, a_2) \neq (0,0)$. To jest, proszę zauważyć, twierdzenie Kroneckera-Capellego. Tylko gdy $(a_1, a_2) \neq (0,0)$, macierz tego układu ma rząd 1, a zatem i zbiór rozwiązań ma wymiar 1, co interpretujemy geometrycznie używając pojęcia „prostej". Z punktu widzenia geometrii elementarnej ważne może być stwierdzenie, że zbiór wektorów (x_1, x_2) spełniających powyższe równanie odpowiada zbiorowi wektorów (x_1, x_2), które są **prostopadłe**[3] do wektora (a_1, a_2). Kiedyś w szkole uczono (ale teraz już nie), że prostopadłość wektorów równoważna jest temu, że ich iloczyn skalarny równy jest 0. Natomiast wyrażenie $a_1x_1 + a_2x_2$ opisuje właśnie ów elementarny iloczyn skalarny wektorów (a_1, a_2) oraz (x_1, x_2). Podobnie zdefiniowany elementarny iloczyn skalarny w przestrzeni $\mathbb{R}^3$ pozwala stwierdzić, że jeśli tylko $(a_1, a_2, a_3) \neq (0,0,0)$, to zbiór rozwiązania równania $a_1x_2 + a_2x_2 + a_3x_3 = 0$ jest płaszczyzną przechodzącą przez punkt $(0,0,0)$ i prostopadłą do wektora (a_1, a_2, a_3). Używam określenia „elementarny". bowiem w drugim semestrze pojęcie iloczynu skalarnego zdefiniujemy w sposób aksjomatyczny, dla dowolnej przestrzeni liniowej nad $\mathbb{R}$, a po pewnych umowach i osłabieniu nazwy „iloczyn skalarny" na „funkcjonał dwuliniowy", badać będziemy choćby prostopadłość wektorów także w przestrzeniach nad innymi ciałami. A zatem okaże się, że prostopadłe mogą być grafy, funkcje zespolone, macierze, wielomiany (to akurat będzie dość ważne już na pierwszym roku Analizy) itd.

Możemy na razie nie martwić się abstrakcją, a zastanowić czym miałaby być prostopadłość układu wektorów w znanej przestrzeni K^n. Powiemy mianowicie, że dwa wektory $(a_1, \ldots, a_n)$ oraz $(b_1, \ldots, b_n)$ są prostopadłe, ozn. $(a_1, \ldots, a_n) \perp (b_1, \ldots, b_n)$, jeśli $a_1b_1 + a_2b_2 + \ldots + a_nb_n = 0$. Zbiór wektorów w K^n prostopadłych do ustalonego zbioru wektorów X oznaczać będziemy przez $X^\perp$. Innymi słowy:

$$X^\perp = \{v \in K^n : v \perp x, \forall_{x \in X}\}.$$

Proszę zauważyć, że dla każdego podzbioru $X \in K^n$ zbiór $X^\perp$ jest podprzestrzenią w K^n. Istotnie, jeśli $(v_1, v_2, \ldots, v_n)$ oraz $(w_1, \ldots, w_n)$ są prostopadłe do dowolnego wektora $(x_1, \ldots, x_n) \in X$, to są do niego prostopadłe również wektory $(v_1 + w_1, \ldots, v_n + w_n)$ oraz $(av_1, \ldots, av_n)$. po prostu dlatego, że

$$(v_1 + w_1)x_1 + \ldots + (v_n + w_n)x_n = v_1x_1 + \ldots + v_nx_n + w_1x_1 + \ldots + w_nx_n = 0.$$

Zauważmy dalej, że jeśli $X = \mathrm{lin}(\alpha_1, \ldots, \alpha_n)$, to podprzestrzeń $X^\perp$ opisuje zbiory współczynników wszystkich równań liniowych, których rozwiązaniami są wektory z X. Z drugiej strony: jeśli mamy jednorodny układ równań liniowych o macierzy, której wierszami są wektory $\alpha_1, \ldots, \alpha_n$, to zbiór rozwiązań tego układu równy jest... $\mathrm{lin}(\alpha_1, \ldots, \alpha_n)^\perp$. Czy Czytelnik widzi dualność, którą tu otrzymujemy? Czy Czytelnik widzi co robi tu Twierdzenie Kroneckera-Capellego? Mówi ono po prostu, że dla podprzestrzeni V przestrzeni liniowej K^n mamy (to się w ogólności zepsuje dla pewnych K i pewnych „niestandardowych prostopadłości", ale dla tej „elementarnej" – tzw. standardowej wersji to zawsze jest prawda):

$$(V^\perp)^\perp = V.$$

Zachęcam Czytelnika, by zastanowił się nad innymi konsekwencjami naszkicowanych tu definicji. Oczywiście (choć będą „wyjątki") $\dim V^\perp = n - \dim V$. Oczywiście, jeśli $V \subseteq W$ są podprzestrzeniami K^n, to $W^\perp \subseteq V^\perp$. Jakże istotne jest to odwrócenie kolejności pomiędzy podzbiorami i przestrzeniami prostopadłymi! Jest to bodaj najprostszy przykład odpowiedniości Galois, o której napiszę dalej. To jeszcze nie koniec. Nie wnikając w geometrię warto zauważyć, że prostopadłość jest swego rodzaju „lepszą liniową niezależnością". W przypadku skończonego układu wektorów liniowa niezależność nie wynika z tego, że dowolne dwa elementy układu są liniowo niezależne. Przyjmuje się natomiast następującą definicję.

Definicja 7.4

Układ wektorów $X \subseteq K^n$ nazwiemy PROSTOPADŁYM (albo ORTOGONALNYM), jeśli $\alpha \perp \beta$, dla każdych $\alpha, \beta \in X$. Układ prostopadły będący bazą przestrzeni V nazywamy BAZĄ PROSTOPADŁĄ (albo ORTOGONALNĄ) przestrzeni V (względem naszego „standardowego" iloczynu skalarnego).

Przykładem bazy prostopadłej jest oczywiście baza standardowa, oczywiście niejedynym. Czytelnikowi zostawiam następujące proste ćwiczenie: dowolny układ prostopadły złożony z niezerowych wektorów jest liniowo niezależny! Konsekwencje tego faktu są bardzo ciekawe, ale na razie nie będziemy eksplorować wątków geometrycznych. Zachęcam Czytelnika do poszukiwania prostopadłości w naszych rozważaniach.

[3] Osoby zainteresowane elementarnymi dowodami tych własności odnoszącymi się do twierdzeń szkolnych zachęcam do zajrzenia do wykładu dra Michała Krycha: *Elementy geometrii analitycznej*, dostępnego na stronie: `https://www.mimuw.edu.pl/~krych/chemia/2016-2017`.

7.5 Dodatek. Odpowiedniość Galois i Nullstellensatz Hilberta

Aby jeszcze lepiej i głębiej zrozumieć dlaczego twierdzenie Kroneckera-Capellego jest istotne, zdefiniujmy dwie operacje $\mathcal{R}$ oraz $\mathcal{W}$ na podzbiorach w K^n.

- Dla podzbioru $S \subseteq K^n$ przez $\mathcal{W}(S) \subseteq K^n$ rozumiemy zbiór złożony z n współczynników każdego takiego jednorodnego równania n zmiennych, którego **rozwiązaniem jest każdy element** z S. Innymi słowy, wektor $(a_1, \ldots, a_n) \in K^n$ należy do $\mathcal{W}(S)$ jeśli dla każdego $(s_1, \ldots, s_n) \in S$: zachodzi
$$a_1 s_1 + a_2 s_2 + \ldots + a_n s_n = 0.$$
- Dla podzbioru $T \subseteq K^n$ przez $\mathcal{R}(T) \subseteq K^n$ rozumiemy **zbiór rozwiązań wszystkich jednorodnych równań** liniowych n zmiennych, których n-tki współczynników należą do T. Innymi słowy, wektor $(s_1, \ldots, s_n)$ należy do $\mathcal{R}(T)$ jeśli dla każdego $(a_1, \ldots, a_n) \in T$ zachodzi równość:
$$a_1 s_1 + a_2 s_2 + \ldots + a_n s_n = 0.$$

Na przykład $(1, 1, -1) \in \mathcal{W}((2, 1, 3))$, ponieważ $1 \cdot 2 + 1 \cdot 1 + (-1) \cdot 3 = 0$, czyli $(2, 1, 3)$ jest rozwiązaniem równania $1 \cdot x_1 + 1 \cdot x_2 + (-1) \cdot x_3 = 0$. Są oczywiście inne elementy $\mathcal{W}(2, 1, 3)$, na przykład (-2, -2, 2). Weźmy jednak odwrotną sytuację: biorę wektor $(1, 1, -1)$ i interesuje mnie jakiś element $\mathcal{R}((1, 1, -1))$. Oczywiście – jednym z nich jest $(2, 1, 3)$, ale nie jedynym. Co to wszystko znaczy? Po co te komplikacje?

Nietrudno widzieć, że mamy dwie zależności (jest ich więcej):
$$\mathcal{W}(\mathcal{R}(S)) \supseteq S, \quad \mathcal{R}(\mathcal{W}(T)) \supseteq T.$$
Pierwsza z nich mówi, że każdy wektor jest elementem (czasami niejedynym, stąd inkluzja) zbioru rozwiązań równania, którego jest rozwiązaniem, a druga mówi, że jeśli równanie ma określone rozwiązanie, to rozwiązanie to jest jego rozwiązaniem (niekoniecznie jedynym, więc znowu jest inkluzja). Brzmi to niemal banalnie, ale interesujące jest to, że zależności te nie dotyczą jedynie równań liniowych! Zauważmy, że jeśli $S_1 \subseteq S_2$, to $\mathcal{W}(S_1) \supseteq \mathcal{W}(S_2)$, podobnie dla operacji $\mathcal{R}$. Wszystko, co powiedzieliśmy na dzisiejszym wykładzie można w zasadzie streścić prostym i eleganckim stwierdzeniem, że jeśli S jest podprzestrzenią liniową przestrzeni K^n – niezależnie czy rozumianą jako przestrzeń współczynników czy przestrzeń rozwiązań, to mamy $\mathcal{R}(S) = S^\perp$ oraz $\mathcal{W}(S) = S^\perp$, czyli:
$$\mathcal{W}(\mathcal{R}(S)) = S, \quad \mathcal{R}(\mathcal{W}(S)) = S. \qquad (*)$$
Możemy też wrócić do wyjściowego przykładu i zapisać wyrażone w nim postulaty w nowym języku. Chcemy znaleźć układ $n-1$ równań, którego zbiorem rozwiązań jest **dokładnie** $\mathrm{lin}(\alpha) \neq 0$. Rzeczywiście:

- $\mathcal{W}(\mathrm{lin}(\alpha))$ jest przestrzenią $n-1$ wymiarową,
- dla $n-1$ liniowo niezależnych elementów $r_1, \ldots, r_{n-1}$ z $\mathcal{W}(\mathrm{lin}(\alpha))$ mamy $\mathcal{R}(r_1, \ldots, r_{n-1}) = \mathrm{lin}(\alpha)$.

Innymi słowy szukane przez nas $n-1$ równań będzie miało współczynniki będące bazą $\mathcal{W}(\mathrm{lin}(\alpha))$.

Operacje tego typu, co $\mathcal{W}$ i $\mathcal{R}$ „rozsiane są" po całej matematyce. Rozważmy jeden ważny przykład:

- podzbiorom X ciała K przyporządkowujemy zbiór $\mathcal{W}(X)$ wszystkich wielomianów o współczynnikach z $K[x]$, których pierwiastkami są wszystkie elementy zbioru X,
- każdemu zbiorowi wielomianów W można przypisać zbiór $\mathcal{R}(W)$ jego wspólnych pierwiastków w K

Przykład. Niech $K = \mathbb{C}$. Wówczas:

- $W(\{-i, i\})$ to zbiór wszystkich wielomianów podzielnych przez $(x^2 + 1)$
- $W(\{0, -i, i\})$ — zbiór wszystkich wielomianów podzielnych przez $x(x^2 + 1)$.

Zauważmy też, że zbiór $\{-i, i\}$ jest zbiorem wspólnych rozwiązań istotnie różnych zbiorów wielomianów, na przykład zbioru wielomianów podzielnych przez $(x^2 + 1)^5$.

Dokładny opis zbioru wielomianów $\mathcal{W}(\mathcal{R}(W))$ nie jest banalny! Jest on treścią słynnego Nullstellensatz – twierdzenia Hilberta o zerach z 1893 roku, które jest uogólnieniem Zasadniczego Twierdzenia Algebry i punktem wyjścia geometrii algebraicznej. Powiedzmy kilka słów o tym twierdzeniu, bez wchodzenia w techniczne detale. Ograniczymy się jedynie do pokazania w jaki sposób twierdzenie to uogólnia Twierdzenie Kroneckera-Capellego. Chodzi mianowicie o rozwiązywanie układów równań, ale wielomianowych. Wychodzimy od następującej sytuacji. Mamy wielomiany $f_1, f_2, \ldots, f_m$ i chcemy coś powiedzieć o zbiorze rozwiązań układu $f_1 = 0, f_2 = 0, \ldots, f_m = 0$. I nie chodzi nam tylko o wielomiany w $K[x]$ tak, by zbiór wspólnych rozwiązań leżał w K. Chodzi nam o tzw. wielomiany n zmiennych i (znowu) o podzbiory K^n.

Definicja 7.5

WIELOMIANEM ZMIENNYCH $x_1, \ldots, x_n$ O WSPÓŁCZYNNIKACH Z CIAŁA K nazywamy wyrażenie postaci:

$$\sum_{i_1,\ldots,i_n} a_{i_1\ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n},$$

gdzie $i_1, \ldots, i_n$ są liczbami całkowitymi nieujemnymi (suma ta brana jest po wszystkich możliwych układach liczb całkowitych nieujemnych), elementy $a_{i_1 i_2 \ldots i_n} \in K$, przy czym zakładamy, że suma ta jest skończona, czyli współczynniki $a_{i_1 i_2 \ldots i_n}$ są różne od 0 tylko dla skończonej liczby indeksów $i_1, \ldots, i_n$. Zbiór wszystkich wielomianów zmiennych $x_1, \ldots, x_n$ o współczynnikach w ciele K oznaczamy $K[x_1, \ldots, x_n]$.

Rozważmy kilka przykładów dla przyswojenia sobie wprowadzonej notacji.

- W wielomianie $w \in \mathbb{R}[x_1, x_2]$ postaci $w = x_1^2 + 4x_1x_2 + 3x_2 + 5$ mamy $a_{20} = 1$, $a_{11} = 4$, $a_{01} = 3, a_{00} = 5$ oraz $a_{i_1 i_2} = 0$ dla pozostałych par i_1, i_2.
- W wielomianie $g \in \mathbb{Q}[x_1, x_2, x_3]$ postaci $g = 7x_1x_2^2x_3^7 - 3x_1x_3^4 + 14x_2^5x_3$ mamy $a_{127} = 7$, $a_{104} = -3, a_{051} = 14$ oraz $a_{i_1 i_2 i_3} = 0$, dla pozostałych i_1, i_2, i_3.

Uwaga. W świetle powyższej definicji wyrażenie postaci x_2x_1 nie jest wielomianem zmiennych x_1, x_2 o współczynnikach w żadnym ciele K, czyli nie należy do $K[x_1, x_2]$. Jest to natomiast wielomian zmiennych x_2, x_1, czyli element $K[x_2, x_1]$. Wyrażenie $x_1x_2 + x_2x_1$ nie ma sensu ani jako element $K[x_1, x_2]$, ani jako element $K[x_2, x_1]$. W przyszłości poznacie Państwo obiekty zwane NIEPRZEMIENNYMI WIELOMIANAMI zmiennych $x_1, \ldots, x_n$, które oznacza się jako $K\langle x_1, \ldots, x_n\rangle$ i nazywa czasem algebrą wolną o generatorach $x_1, \ldots, x_n$ nad ciałem K.

Definicja 7.6

STOPNIEM WIELOMIANU $\sum_{i_1,\ldots,i_n} a_{i_1,\ldots,i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \in K[x_1, \ldots, x_n]$ nazywamy największą z liczb $i_1 + i_2 + \ldots + i_n$, dla których $a_{i_1\ldots i_n} \neq 0$. Stopień wielomianu f oznaczamy $\deg f$. Jeśli f jest WIELOMIANEM ZEROWYM – to znaczy $a_{i_1\ldots i_n} = 0$, dla wszystkich $i_1, \ldots, i_n$, to piszemy $\deg f = -\infty$.

Definicja 7.7

SUMĄ WIELOMIANÓW $\sum_{i_1,\ldots,i_n} a_{i_1\ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ oraz $\sum_{i_1,\ldots,i_n} b_{i_1\ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ nazywamy wielomian $\sum_{i_1,\ldots,i_n} c_{i_1\ldots i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ taki, że $c_{i_1\ldots i_n} = a_{i_1\ldots i_n} + b_{i_1\ldots i_n}$, dla każdych $i_1, \ldots, i_n$.

Dla wielomianów w i g z wcześniejszych przykładów mamy $\deg w = 2$, $\deg g = 10$. Sumą wielomianów $2x_1^2x_2 + 6x_1x_2 - 5x_1$ oraz $7x_1^5 - 2x_1x_2 + 5x_1$ jest wielomian $7x_1^5 + 2x_1^2x_2 + 4x_1x_2$.

Szczególnymi typami wielomianów są wielomiany liniowe, to znaczy wielomiany stopnia 1, np.

$$x_1, x_1 + \ldots + x_n, \quad 2x_1 + x_3 - x_4.$$

Rozwiązywanie jednorodnych układów równań liniowych jest z tej perspektywy szczególnym przypadkiem rozwiązywania wielomianowych układów równań. Ich rozwiązaniami są podprzestrzenie liniowe. Rozwiązaniami układów równań wielomianowych są tzw. zbiory algebraiczne. Powiemy o nich więcej na zakończenie drugiego semestru. Istotne jest to, że znamy wiele zbiorów algebraicznych, np. zbiór zer wielomianu dwóch zmiennych $x_1^2 - x_2^2$ to dwie przecinające się proste, a zbiór rozwiązań równania $x_1^2 - x_2$ to parabola (są też sfery, walce, hiperboloidy itd.). Badanie układów równań wielomianowych to punkt wyjścia wielkiego działu matematyki – geometrii algebraicznej. O czym jest zatem[4] Twierdzenie Hilberta? Zacznijmy od „stosunkowo prostej" sytuacji.

[4]Na motywach tekstu prof. Andrzeja Nowickiego: Afiniczne zbiory algebraiczne (do znalezienia na stronie Profesora) oraz tekstu Hilbert's Nullstellensatz w *The Princeton Companion to Mathematics.*

Weźmy element $a = (a_1, \ldots, a_n) \in K^n$ i zastanówmy się jak może wyglądać zbiór wielomianów n zmiennych, które zerują się na a, czyli $\mathcal{W}(\{a\})$. W przypadku wielomianów jednej zmiennej jest to po prostu zbiór

$$(x-a)f, \quad \text{gdzie} \quad f \in K[x].$$

W przypadku wielomianów wielu zmiennych wnioskować należy, że w $\mathcal{W}(\{a\})$ jest każdy wielomian postaci:

$$(x_1 - a_1)f_1 + (x_2 - a_2)f_2 + \ldots + (x_n - a_n)f_n,$$

gdzie $f_1, \ldots, f_n \in K[x_1, \ldots, x_n]$ są dowolne. Można, jak się okazuje pokazać, że powyższy zbiór jest w istocie całym $\mathcal{W}(\{a\})$. Nie jest to bardzo trudne. Trudniej rozwiązywać problem odwrotny.

Twierdzenie Hilberta o zerach podejmuje następujący problem: załóżmy, że startujemy od pewnego zbioru wielomianów n zmiennych nad k, wyznaczamy wszystkie punkty w K^n, na których mogą się one zerować, a potem dla tych punktów wyznaczamy wszystkie wielomiany, które się na nich zerują. Co dostajemy? Innymi słowy, jeśli $X \subseteq K[x_1, \ldots, x_n]$, to czym jest $\mathcal{W}(\mathcal{R}(X))$?

Nawet w przypadku wielomianów jednej zmiennej możemy otrzymać nietrywialne odpowiedzi, jak widzieliśmy wyżej. Warto założyć chociaż, że ciało K jest algebraicznie domknięte, żeby nie martwić się zbiorami pustymi. Załóżmy, że X jest skończonym zbiorem złożonym z wielomianów $f_1, \ldots, f_m$. Zbiór $\mathcal{R}(X)$ to zbiór jego wspólnych pierwiastków. Czym jest teraz $\mathcal{W}(\mathcal{R}(X))$? Jakie jeszcze wielomiany zerują się na tym samym zbiorze, co wielomiany $f_1, \ldots, f_m$? Na pewno są to wielomiany postaci:

$$f_1 g_1 + \ldots + f_m g_m,$$

gdzie $g_1, \ldots, g_m$ są dowolnymi wielomianami z $K[x_1, \ldots, x_n]$. Co jeszcze? Czasem coś jeszcze, bo np.

$$(x-1) \in \mathcal{W}(\mathcal{R}(\{(x-1)^2\})).$$

Okazuje się, że jest to jedyny rodzaj „niespodzianki", o czym mówi słynny wynik Hilberta.

Twierdzenie 7.4: Hilberta o zerach

Niech K będzie ciałem algebraicznie domkniętym oraz niech $f_1, \ldots, f_m$ należą do $K[x_1, \ldots, x_n]$. Wówczas jeśli $h \in \mathcal{W}(\mathcal{R}(f_1, \ldots, f_m))$ (tzn. funkcja wielomianowa odpowiadająca h zeruje się na podzbiorze K^n, będącym częścią wspólną zbiorów zer funkcji wielomianowych odpowiadających $f_1, \ldots, f_m$), to istnieje $r \in \mathbb{Z}_+$ oraz wielomiany $g_1, \ldots, g_m$, że:

$$h^r = f_1 g_1 + \ldots + f_m g_m.$$

W ten sposób wyróżnione zostają zbiory wielomianów X, dla których

$$\mathcal{W}(\mathcal{R}(X)) = X.$$

Zbiory te są bowiem w odpowiedniości ze zbiorami rozwiązań wielomianowych układów równań nad ciałem algebraicznie domkniętym, w podobny sposób jak w zależności tej są podprzestrzenie liniowe ze zbiorami rozwiązań jednorodnych liniowych układów równań. Dla zainteresowanych zostawiam jedynie hasło: ideał radykalny. Porównajmy tę sytuację z twierdzeniem Kroneckera-Capellego. Czy Czytelnik widzi, porównując z $(*)$, że jest ono w jakimś sensie szczególnym przypadkiem twierdzenia Hilberta?

Olimpijczykom znany może być następujący fakt żyjący pod nazwą „kombinatoryczne Nullstellensatz".

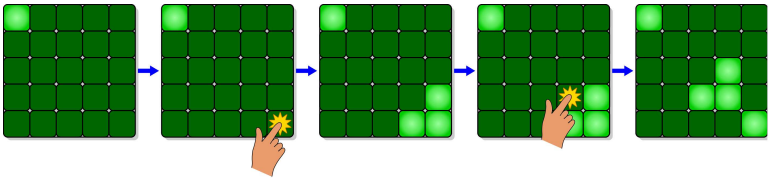
Twierdzenie 7.5

Niech p będzie niezerowym wielomianem zmiennych $x_1, \ldots, x_n$ stopnia $\sum_{i=1}^n m_i$, w którym współczynnik przy $x_1^{m_1} \cdots x_n^{m_n}$ jest różny od zera. Wówczas dla dowolnych zbiorów $S_1, \ldots, S_n$ zawartych w $\mathbb{R}$ spełniających warunki $|S_i| > m_i$, dla $1 \leqslant i \leqslant n$, istnieją takie $c_i \in S_i$, że $p(c_1, \ldots, c_n) \neq 0$.

Zainteresowanych tym twierdzeniem i jego ładnymi elementarnymi aspektami odsyłam na przykład do artykułu Jacka Dymela „O zastosowaniach Combinatorial Nullstellensatz", dostępnego na stronach Delty: `http://www.deltami.edu.pl/temat/matematyka/algebra/2017/06/16/2017-07-delta-dymel.pdf`.

7.6 Trivia. Lights Out

W 1995 roku Tiger Electronics wydało grę *Lights Out*. Gra składa się z tablicy rozmiaru 5 na 5 złożonej z 25 przycisków, z których każdy jest w jednym z dwóch stanów: włączony (wtedy przycisk jest podświetlony) lub wyłączony. Po rozpoczęciu gry włącza się losowa konfiguracja przycisków. Naciśnięcie dowolnego przycisku spowoduje przełączenie tego przycisku, a także jego sąsiadów (ale nie po przekątnej).



Zadaniem jest wyłączenie wszystkich świateł, najlepiej za pomocą jak najmniejszej liczby ruchów.

W 1998 roku Marlow Anderson oraz Todd Fell użyli metod algebry liniowej[5] do pokazania, że nie wszystkie konfiguracje prowadzą do rozwiązania oraz, że dla każdej rozwiązywalnej konfiguracji istnieją dokładnie cztery strategie (nie mające zbędnych ruchów). Idea jest prosta: każdą z tablic reprezentować można jednoznacznie jako element przestrzeni $M_5(\mathbb{Z}_2)$. W ten sposób wciśnięcie klawisza w i-tym wierszu i j-tej kolumnie odpowiada dodawaniu pewnej macierzy $t_{ij} \in M_5(\mathbb{Z}_2)$, na przykład (dla ilustracji wyżej):

$$\begin{bmatrix} 1&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{bmatrix} + \begin{bmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&1\\ 0&0&0&1&1 \end{bmatrix} + \begin{bmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&1&0\\ 0&0&1&1&1\\ 0&0&0&1&0 \end{bmatrix} = \begin{bmatrix} 1&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&1&0\\ 0&0&1&1&0\\ 0&0&0&0&1 \end{bmatrix}.$$

Dla dowolnych macierzy $x, y \in M_5(\mathbb{Z}_2)$ mamy $x + y = y + x$ oraz $x + x = 0$, więc widzimy, że kolejność wciskania przycisków nie ma znaczenia dla strategii oraz żadnego przycisku nie trzeba wciskać więcej niż raz. Załóżmy, że wyjściową konfigurację świateł opisuje macierz b. Problem wyłączenia świateł w b równoważny jest zatem pytaniu: czy b należy do $\mathrm{lin}(t_{ij}, 1 \leqslant i, j \leqslant 5)$? A zatem jest to w istocie problem rozwiązania układu równań — w tym przypadku o 25 niewiadomych (o współczynnikach w $\mathbb{Z}_2$). Nietrudno sprawdzić, że macierz $A \in M_{25\times 25}(\mathbb{Z}_2)$ tego układu ma 25 bloków rozmiaru 5×5 postaci:

$$A = \begin{bmatrix} C & I_5 & 0 & 0 & 0\\ I_5 & C & I_5 & 0 & 0\\ 0 & I_5 & C & I_5 & 0\\ 0 & 0 & I_5 & C & I_5\\ 0 & 0 & 0 & I_5 & C \end{bmatrix}, \quad \text{gdzie} \quad C = \begin{bmatrix} 1&1&0&0&0\\ 1&1&1&0&0\\ 0&1&1&1&0\\ 0&0&1&1&1\\ 0&0&0&1&1 \end{bmatrix} \quad \text{oraz} \quad I_5 = \begin{bmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{bmatrix}.$$

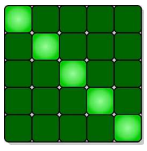
Widzimy zatem, że sprawdzając, czy można „zgasić światła" w macierzy b rozwiązujemy układ równań $Ax = b$, gdzie b jest wektorem o 25 współrzędnych złożonym z kolejnych wyrazów macierzy b. Dla przykładu, gdy b jest macierzą po prawej na ilustracji wyżej, to wektor b ma postać:

$$b = (1,0,0,0,0,\ 0,0,0,0,0,\ 0,0,0,1,0,\ 0,0,1,1,0,\ 0,0,0,0,1).$$

Wektor $x \in \mathbb{Z}_2^{25}$ wskazuje które przyciski mamy wcisnąć, aby ze zgaszonej macierzy dostać b.

Jak można policzyć (najlepiej przy pomocy komputera), rząd macierzy $A \in M_{25\times 25}(\mathbb{Z}_2)$ równy jest 23. Oznacza to, że liczba rozwiązywalnych konfiguracji gry równa jest 2^{23}. Prawdpodobieństwo, że losowa konfiguracja jest rozwiązywalna równe jest $2^{23}/2^{25} = 1/4$.

Odnotujmy dwie ciekawe sytuacje: pierwsza, gdy rozwiązanie układu $Ax = 0$ zwraca strategie, które nie zmieniają oświetlenia na tablicy oraz druga, gdy rozwiązanie układu $Ax = x$ zwraca te strategie, w których wciśnięcie odpowiednich przycisków sprawia, że jedynie wciśnięte przyciski są zapalone. Przykładem tej sytuacji jest wciśnięcie przycisków na przekątnej, startując od pustej tablicy. Jest jeszcze 31 innych.



[5] Wyjaśnienie tego szkicu i ładny opis problemu można znaleźć w rozdziale 24. pięknej książki „*Permutation puzzles*", dostępnej pod adresem `http://www.sfu.ca/~jtmulhol/math302/notes/permutation-puzzles-book.pdf` (strona autora).

7.7 Coda. Bardzo wstępnie o twierdzeniach klasyfikacyjnych

Twierdzenie Kroneckera-Capellego przeprowadza klasyfikację podprzestrzeni przestrzeni liniowej K^n w języku układów równań. Każda taka podprzestrzeń wymiaru k może być opisana za pomocą układu równań liniowych złożonego z $n-k$ elementów, którego macierz ma rząd $n-k$. Tu w istocie zawarta jest bardzo głęboka treść, której jeszcze nie wysłowiliśmy. Twierdzenie to mówi w istocie, że jeśli rozważymy podprzestrzeń wymiaru k i opiszemy ją poprzez układ $n-k$ równań zapisany w postaci ogólnej:

$$\begin{cases} x_{j_1} & = c_{11}x_{t_1} + c_{12}x_{t_2} + \ldots + c_{1,k}x_{t_k} \\ & \vdots \\ x_{j_{n-k}} & = c_{n-k,1}x_{t_1} + c_{n-k,2}x_{t_2} + \ldots + c_{n-k,k}x_{t_k} \end{cases}$$

to dokonując w tych równaniach **liniowej zamiany zmiennych** postaci:

$$\begin{cases} y_1 & = x_{j_1} - (c_{11}x_{t_1} + c_{12}x_{t_2} + \ldots + c_{1,k}x_{t_k}) \\ & \vdots \\ y_{n-k} & = x_{j_{n-k}} - (c_{n-k,1}x_{t_1} + c_{n-k,2}x_{t_2} + \ldots + c_{n-k,k}x_{t_k}) \\ y_{n-k+1} & = 0 \\ & \vdots \\ y_n & = 0 \end{cases},$$

to w owym układzie współrzędnych nasza podprzestrzeń opisana jest równaniami $y_1 = 0, \ldots, y_{n-k} = 0$.

Oto przykład: weźmy dwuwymiarową podprzestrzeń W w $\mathbb{R}^3$ opisaną równaniem $x_1 - x_2 + x_3 = 0$. Określenie to mówi, że podprzestrzeń W złożona jest z wektorów

$$v = x_1 \cdot (1,0,0) + x_2 \cdot (0,1,0) + x_3(0,0,1),$$

których współrzędne w bazie standardowej przestrzeni $\mathbb{R}^3$ spełniają równanie wyżej. A co by się stało, gdybyśmy rozważali zbiory wektorów, których współrzędne spełniają równania liniowe, ale nie są to współrzędne w bazie standardowej? Jakie? Możemy mianowicie rozważyć liniową zamianę zmiennych postaci $y_1 = x_1 - x_2 + x_3$ oraz $y_2 = 0$, $y_3 = 0$. Wówczas nasza podprzestrzeń opisana jest równanie $y_1 = 0$. Liniowa zamiana zmiennych oznacza, że teraz rozważamy współrzędne z innej bazie. Jakiej? Weźmy bazę zawierającą jako pierwszy wektor $(1,-1,1)$, a dwa pozostałe — to elementy bazy W, na przykład: $(1,-1,1), (1,1,0), (1,0,-1)$. Innymi słowy, rozważając wektory postaci $y_1(1,-1,1) + y_2(1,1,0) + y_3(1,0,-1)$ stwierdzamy łatwo, że wektory te należą do W wtedy i tylko wtedy, gdy $y_1 = 0$.

Twierdzenia klasyfikacyjne obecne są w całej nauce. Być może klasycznym dziełem w historii nauki pokazującym jej dążenie do klasyfikowania obiektów w zależności od tego ile wspólnych cech posiadają, była *Systema Naturae* Karola Linneusza z 1735 roku. W tym rozumieniu „charakteryzacja" ukazuje po prostu „wspólny charakter" — w tym przypadku chodzi o klasyfikację biologiczną organizmów. Twierdzenia klasyfikacyjne matematyki, idą nieco dalej. Przechodzą one bowiem do języka niezmienników i funkcji.

Dla dwóch krzywych opisanych równaniem kwadratowym, na przykład elipsy i hiperboli, nie wystarczy powiedzieć, że posiadają one pewne charakterystyczne własności. Wymagamy, aby istniała transformacja/przekształcenie określonego typu, które przeprowadzi jedną krzywą w drugą. Innymi słowy, interesuje nas, by obiekty były identyczne modulo pewna operacja. Być może Czytelnika zaciekawi fakt, że w zależności od przyjętego typu przekształcenia, okrąg i hiperbola mogą być równoważne, lub nie. Powiemy o tym w drugim semestrze: nie istnieje izometria lub izomorfizm afiniczny przeprowadzający elipsę w hiperbolę, ale istnieje przekształcenie rzutowe, które tego dokonuje.

Na najbliższych wykładach poznamy pojęcie przekształcenia liniowego i izomorfizmu przestrzeni liniowych. Będzie to pierwsza kluczowa klasa przekształceń, która będzie utożsamiała przestrzenie liniowe, które uważamy jako takie same. Udowodnimy też, że każda skończenie wymiarowa przestrzeń liniowa wymiaru n nad ciałem K jest izomorficzna z przestrzenią K^n. Ten właśnie izomorfizm będzie formalnie zapisywał w sobie liniową zamianę zmiennych za pomocą pewnej macierzy odwracalnej. Wkrótce dowiemy się co to wszystko znaczy. Przedstawmy kilka przykładów z historii matematyki.

Wielkim osiągnięciem była klasyfikacja wielościanów foremnych w przestrzeni trójwymiarowej — jest ich tylko 5. Ten wynik dawał starożytnym silny impuls filozoficzny do pewnej konstukcji światopoglądowej — mówiącej, że wszystko na świecie ma pierwotną przyczynę — Arché, Nie było zgody, gdzie jest ona umieszczona. Jedni widzieli ją w wodzie, inni w bezkresie, inni w powietrzym inni w ogniu. Później próbowano scalać te koncepcje i u podstaw rzeczywistości widziano kilka podstawowych *elementów*, działających przeciwstawnie, których wzajemne oddziaływanie miało być źródłem zmiany. Inna koncepcja pochodziła od Demokryta, który wszystkie rzeczy materialne postrzegał jako stworzone z małych, niepodzielnych cząstek (atomów). Różne proporcje ich połączeń miały prowadzić do różnorodności rzeczy.

Wielościany foremne nazywany czasem bryłami platońskimi właśnie dlatego, że słynny ateński filozof uważał, że są one budulcem owych „podstawowych elementów" czy „atomów". Ogień miał być zbudowany z czworościanów foremnych, ziemia z sześcianów, powietrze z ośmiościanów, woda z dwudziestościanów. Dwudziestościan reprezentować miał żywioł niebieski. Arystoteles nazywał go eterem. Kierunek ten nie pochodził jedynie od uczonych greckich. Niezależnie formułowany był w kulturze starożytnych Chin, Japonii, czy Indii. W czasach nowożytnych koncepcje żywiołów obecne były w badaniach alchemicznych, którego przedmiotem zainteresowania były metale i ich „transmutacje", których celem miało być otrzymanie złota lub odkrycie kamienia filozoficznego. Rozwój filozofii przyrody, badań empirycznych i aparatury w XVII wieku doprowadził do skupienia się na bardziej ogólnych celach poznawczych i przyczyniło się stopniowo do sformułowania idei pierwiastka chemicznego. Na koncepcji pięciu żywiołów opierał się jeszcze Kartezjusz — twórca geometrii analitycznej.

Z czego składa się matematyczne twierdzenie klasyfikacyjne? Ma ono dwa elementy:

- Listę normalnych (kanonicznych) form, reprezentantów danej relacji równoważności — w przypadku rozważanej przez nas teorii są to równania postaci, na przykład $x_1 = 0, x_2 = 0, \ldots, x_{n-k} = 0$.
- Twierdzenie klasyfikacyjne stwierdzający, że każdy obiekt jest równoważny do jednej zz form normalnych (kanonicznych). W naszym przypadku chodzi o stwierdzenie, że dla każdej przestrzeni wymiaru k istnieje liniowa zamiana zmiennych (izomorfizm liniowy) taka, że w odpowiednim układzie współrzędnym podprzestrzeń opisać można dokładnie jednym z równań kanonicznych. Co więcej twierdzenie to orzeka, że każdych dwóch różnych form kanonicznych nie można przekształcić na siebie — w naszym przypadku — za pomocą liniowej zamiany zmiennych.

Zarówno na GALu, jak i na kolejnych przedmiotach poznają Państwo szereg pięknych twierdzeń klasyfikacyjnych. Są one bardzo często celem całego kursu, i często to są właśnie „twierdzenia z nazwiskiem". W drugim semestrze zmierzać będziemy do następujących twierdzeń klasyfikacyjnych:

- twierdzenie Jordana o klasyfikacji endomorfizmów przestrzeni liniowej skończonego wymiaru nad ciałem algebraicznie domkniętym,
- twierdzenie Cartana o rozkładzie dowolnej izometrii liniowej n-wymiarowej przestrzeni liniowej na złożenie nie więcej niż n symetrii prostopadłych, zawierające w sobie klasyfikację izometrii płaszczyzny i przestrzeni trójwymiarowej, dokonaną już wieki wcześniej,
- twierdzenie spektralne o ortogonalnej diagonalizacji, pochodzące w pierwszych wersjach od Kartezjusza i Fermata, w kolejnych od twierdzenie Eulera (twierdzenie o osiach głównych), aż do ogólnych rozważań Lagrange'a i Jacobiego, skumulowanych w pracach Cauchy'ego i Sylvestera, które współczesną formę przybrały dzięki Frobeniusowi[6]
- twierdzenia Sylvestera o inercji, o klasyfikaującego rzeczywiste formy dwuliniowe, rozróżniające między innymi czterowymiarową przestrzeń euklidesową i czasoprzestrzeń,
- twierdzenie klasyfikujące kwadryki — powierzchnie stopnia 2 w afinicznej przestrzeni euklidesowej.

Na innych wykładach poznawać będziecie Państwo twierdzenia klasyfikujące grupy przemienne, powierzchnie wyższych stopni (zarówno od strony algebraicznej jak i różniczkowej), rozmaitości na analizie i topologii i wiele innych. Twierdzenia te ukazują jedynie wierzchołek współczesnej konstrukcji matematyki, w której odkrywamy coraz mocniej, jak twierdzenia klasyfikujące obiekty w różnych „kategoriach" mają się do siebie. Będziemy do tego tematu wracać. Morał na dziś brzmi: kluczowym elementem działalności matematycznej jest klasyfikowanie. Wielkie problemy matematyki to między innymi problemy klasyfikacyjne, np. rozwiązana przez Perelmana (w roku 2002) słynna Hipoteza Poincarégo z roku 1904.

[6] L.A. Steen, *Highlight in the history of spectral theory*, The American Mathematical Monthly, Vol. 80, No. 4 (Apr. 1973), pp. 359-381, `https://www.jstor.org/stable/2319079`.

Rozdział 8

Operacje na podprzestrzeniach

8.1 Wykład ósmy

Dziś zobaczymy w jaki sposób poznane metody pozwalają nam poruszać się w świecie podprzestrzeni przestrzeni liniowej, w szczególności w jaki sposób pozwalają na przypisanie parze czy też całej rodzinie podprzestrzeni dwóch naturalncych obiektów – sumy i części wspólnej. Szczególną rolę grają też rozkłady przestrzeni na sumy proste podprzestrzeni. Zacznijmy od powrotu do przestrzeni K^n i zobaczmy w jaki sposób mając dane jej podprzestrzenie konstruować możemy nowe, związane z nimi podprzestrzenie.

- Jeśli V_1, V_2 są podprzestrzeniami w K^n opisanymi układami równań U_1 oraz U_2, to podprzestrzeń

  $$V_1 \cap V_2$$

  jest opisana układem równań złożonym ze wszystkich równań z U_1 oraz wszystkich równań z U_2. Na przykład, jeśli $W_1, W_2 \subseteq \mathbb{R}^3$ opisane są odpowiednio układami $x_1 + x_2 + x_3 = 0, x_1 - x_3 = 0$ oraz $x_1 - x_2 = 0$, to największa podprzestrzeń zawierająca elementy zarówno z W_1, jak i W_2, czyli właśnie $W_1 \cap W_2$, opisana jest układem równań:

  $$\begin{cases} x_1 + x_2 + x_3 & = 0 \\ x_1 - x_3 & = 0 \\ x_1 - x_2 & = 0 \end{cases}.$$

- Jeśli $V_1 = \operatorname{lin}(\alpha_1, \ldots, \alpha_n)$ oraz $V_2 = \operatorname{lin}(\beta_1, \ldots, \beta_m)$ są podprzestrzeniami przestrzeni V, to podprzestrzeń:

  $$W = \operatorname{lin}(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$$

  złożona jest ze wszystkich wektorów postaci $\alpha + \beta$, gdzie $\alpha \in V_1, \beta \in V_2$ jest najmniejszą podprzestrzenią w V, która zawiera jednocześnie V_1 oraz V_2. Innymi słowy, jest to przestrzeń $\operatorname{lin}(V_1 \cup V_2)$.

  Na przykład, jeśli $V_1 = \operatorname{lin}((1,0,1)), V_2 = \operatorname{lin}((1,0,2)) \subseteq \mathbb{R}^3$, to najmniejsza podprzestrzeń $\mathbb{R}^3$ zawierająca te dwie podprzestrzenie to $\operatorname{lin}((1,0,1),(1,0,2))$.

Definicja 8.1: Suma podprzestrzeni

Niech $X, Y \subseteq V$. Przez $X + Y$ oznaczać będziemy zbiór

$$\{x + y \,|\, x \in X, y \in Y\}.$$

Jeśli X, Y są podprzestrzeniami w przestrzeni V, to $X + Y$ też jest podprzestrzenią przestrzeni V zwaną SUMĄ PODPRZESTRZENI X i Y.

Również pojęcie części wspólnej podprzestrzeni przestrzeni K^n przenosi się na dowolne przestrzenie liniowe. Następującą obserwację pozostawiamy jako kolejne proste ćwiczenie.

Obserwacja 8.1

Część wspólna $X \cap Y$ podprzestrzeni liniowej V jest podprzestrzenią liniową. Nazywamy ją ILOCZYNEM, PRZECIĘCIEM (lub CZĘŚCIĄ WSPÓLNĄ) podprzestrzeni.

Zauważmy zatem, że z wraz dowolnymi dwiema podprzestrzeniami V_1, V_2 przestrzeni V rozważać można dwie podprzestrzenie: $V_1 \cap V_2$ – będącą ich największą wspólną podprzestrzenią oraz $V_1 + V_2$ – najmniejszą przestrzeń liniową, której V_1, V_2 są podprzestrzeniami. Kluczowa dla zrozumienia związku w opisanym układzie podprzestrzeni jest następująca formuła.

Twierdzenie 8.1: Formuła Grassmanna, 1844

Niech V_1, V_2 będą skończenie wymiarowymi podprzestrzeniami przestrzeni V. Wówczas podprzestrzenie $V_1 \cap V_2$ oraz $V_1 + V_2$ też są skończenie wymiarowe i zachodzi:

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Dowód. Niech $\gamma_1, \ldots, \gamma_m$ będzie bazą przestrzeni $V_1 \cap V_2$. Uzupełniamy ja, na mocy twierdzenia Steinitza, do bazy $\gamma_1, \ldots, \gamma_m, \alpha_1, \ldots, \alpha_k$ przestrzeni V_1 oraz do bazy $\gamma_1, \ldots, \gamma_m, \beta_1, \ldots, \beta_l$ przestrzeni V_2. Wykażemy, że układ $\gamma_1, \ldots, \gamma_m, \alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l$ jest bazą przestrzeni $V_1 + V_2$. Oczywiście układ ten rozpina tą przestrzeń. Pozostaje wykazać jego liniową niezależność.

Przypuśćmy, że $c_1\gamma_1 + \ldots + c_m\gamma_m + a_1\alpha_1 + \ldots + a_k\alpha_k + b_1\beta_1 + \ldots + b_l\beta_l = 0$. Stąd wynika, że

$$c_1\gamma_1 + \ldots + c_m\gamma_m + a_1\alpha_1 + \ldots + a_k\alpha_k = -(b_1\beta_1 + \ldots + b_l\beta_l) \in V_1. \tag{8.1}$$

Zauważmy jednak, że jeśli $b_1\beta_1 + b_2\beta_2 + \ldots + b_l\beta_l \in V_1$, to kombinacja ta należy w istocie do iloczynu $V_1 \cap V_2$, a więc jest równa pewnej kombinacji postaci $c_1'\gamma_1 + \ldots + c_m'\gamma_m$. Wtedy jednak

$$b_1\beta_1 + \ldots + b_l\beta_l - c_1'\gamma_1 - \ldots - c_m'\gamma_m = 0,$$

co z liniowej niezależności układu $\beta_1, \ldots, \beta_l, \gamma_1, \ldots, \gamma_m$ implikuje, że $b_1 = b_2 = \ldots = b_l = c_1' = \ldots = c_m' = 0$.

Powyższy argument oznacza, że we wzorze (8.1) całą kombinację liniową $-(b_1\beta_1 + \ldots + b_l\beta_l)$ możemy zastąpić przez 0. Mamy zatem:

$$c_1\gamma_1 + \ldots + c_m\gamma_m + a_1\alpha_1 + \ldots + a_k\alpha_k = 0.$$

To jest jednak kombinacja wektorów bazowych z V_1, co oznacza, że $c_1 = \ldots = c_m = a_1 = \ldots = a_k = 0$. Istotnie więc

$$\dim(V_1 + V_2) = m + k + l = (m + k) + (m + l) - m = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

□

Definicja 8.2: Suma prosta podprzestrzeni

eśli dla pewnych podprzestrzeni X, Y przestrzeni liniowej V każdy wektor $\alpha \in V$ da się przedstawić jednoznacznie w postaci sumy wektorów $x \in X$ oraz $y \in Y$ to mówimy, że V jest SUMĄ PROSTĄ podprzestrzeni X, Y, ozn. $V = X \oplus Y$.

Przykład 1. Zgodnie z opisem poczynionym wyżej mamy: $\mathbb{R}^2 = \text{lin}(1, 1) \oplus \text{lin}(1, -1)$.

Przykład 2. Każdy ciąg zbieżny o wyrazach rzeczywistych można w jednoznaczny sposób przedstawić jako sumę ciągu stałego i ciągu zbieżnego do 0. A zatem podprzestrzeń $\mathcal{C}$ przestrzeni $\mathbb{R}^\infty$ złożona z ciągów zbieżnych jest sumą prostą podprzestrzeni złożonej z ciągów stałych (jednowymiarowa) i podprzestrzeni $\mathcal{C}_0$ złożonej z ciągów zbieżnych do zera (nieskończenie wymiarowa – dlaczego?).

Przykład 3. Podzbiory macierzy symetrycznych i antysymetrycznych w $M_{n\times n}(K)$ są podprzestrzeniami i dla ciała charakterystyki[1] $\neq 2$ ich suma to całe $M_{n\times n}(K)$. Dla każdego $M \in M_{n\times n}(K)$ mamy bowiem:

$$M = \frac{M + M^T}{2} + \frac{M - M^T}{2}.$$

Kolejny przykład zaczerpniemy z przestrzeni funkcji rzeczywistych.

Przykład 4. Każdą funkcję $f : \mathbb{R} \to \mathbb{R}$ można przedstawić w sposób jednoznaczny jako sumę funkcji parzystej i nieparzystej, bo mamy rozkład:

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

Dlaczego rozkłady przedstawione z Przykładach 3 i 4 są jednoznaczne i uzyskane sumy podprzestrzeni są sumami prostymi? Dowód opiera się o następującą ogólną obserwację.

Obserwacja 8.2

Niech V_1, V_2 będą podprzestrzeniami przestrzeni V. Następujące warunki są równoważne:

- $V = V_1 \oplus V_2$,
- $V = V_1 + V_2$ i $V_1 \cap V_2 = \{0\}$.

Dowód. Załóżmy najpierw, że $V = V_1 \oplus V_2$. Wówczas dla każdego $\alpha \in V$ mamy $\alpha = \alpha_1 + \alpha_2$, gdzie $\alpha_1 \in V_1, \alpha_2 \in V_2$, więc $V = V_1 + V_2$. Gdyby $V_1 \cap V_2 \neq \{0\}$, to istniałby niezerowy wektor $\alpha \in V_1 \cap V_2$. Co więcej, można by było przedstawić ten wektor na dwa sposoby jako sumę wektora z V_1 i V_2, mianowicie: $\alpha = \alpha + 0 = 0 + \alpha$. Jest to sprzeczne z jednoznacznością w definicji sumy prostej.

Na odwrót: przypuśćmy, że $V = V_1 + V_2$ oraz $V_1 \cap V_2 = \{0\}$. Musimy wykazać, że każdy wektor $\alpha \in V$ daje się jednoznacznie przedstawić jako suma wektora z V_1 i wektora z V_2. Z równości $V = V_1 + V_2$ wynika, że istnieją $\alpha_1 \in V_1$ oraz $\alpha_2 \in V_2$ spełniające $\alpha = \alpha_1 + \alpha_2$. Gdyby to przedstawienie nie było jednoznaczne, to zachodziłoby równanie $\alpha = \alpha_1' + \alpha_2'$, dla pewnych $\alpha_1' \in V_1$, $\alpha_2' \in V_2$, przy czym $\alpha_1 \neq \alpha_1'$ (równoważnie: $\alpha_2 \neq \alpha_2'$). Wtedy jednak

$$0 \neq \alpha_1 - \alpha_1' = \alpha_2' - \alpha_2 \in V_1 \cap V_2,$$

co jest sprzeczne z $V_1 \cap V_2 = \{0\}$. □

Poniższy wniosek odnosi się bezpośrednio do formuły Grassmanna.

Wniosek 8.1

Niech V_1, V_2 będą podprzestrzeniami skończenie wymiarowej przestrzeni V. Załóżmy, że $V_1 \cap V_2 = \{0\}$. Wówczas następujące warunki są równoważne:

- $V = V_1 \oplus V_2$,
- $\dim V = \dim V_1 + \dim V_2$,
- jeśli $\mathcal{A}$ jest bazą V_1 oraz $\mathcal{B}$ jest bazą V_2, to $\mathcal{A} \cup \mathcal{B}$ jest bazą V.

Na koniec omówimy definicje uogólniające pojęcie sumy i iloczynu podprzestrzeni, a także pojęcie sumy prostej na dowolną rodzinę podprzestrzeni. Wyjściowa sytuacja jest następująca: dana jest rodzina podprzestrzeni $\{V_t\}_{t\in T}$ przestrzeni V, gdzie T może być zbiorem nieskończonym (np. zbiorem liczb naturalnych lub rzeczywistych). Interesuje nas znalezienie najmniejszej podprzestrzeni V, której podprzestrzeniami są wszystkie elementy rozważanej rodziny oraz znalezienie największej podprzestrzeni, będącej jednocześnie podprzestrzenią wszystkich podprzestrzeni należących do rozważanej rodziny.

[1] Przypomnienie: ciało ma charakterystykę p jeśli $\underbrace{1 + \ldots + 1}_{p} = 0$. Charakterystyka może być liczbą pierwszą lub zerem.

Definicja 8.3

Niech $\{V_t\}_{t\in T}$ będzie rodziną podprzestrzeni przestrzeni V. Wówczas określamy zbiór

$$\sum_{t\in T} V_i = \{\alpha_{t_1} + \ldots + \alpha_{t_r} \,|\alpha_{t_i} \in V_{t_i}, t_1, \ldots, t_r \in T, r \in \mathbb{N}\},$$

zwany SUMĄ RODZINY PODPRZESTRZENI $\{V_t\}_{t\in T}$. W przypadku $T = \{1, \ldots, n\}$ piszemy:

$$\sum_{i=1}^{n} V_i = V_1 + \ldots + V_n.$$

Przykłady:

- $K[x] = \sum_{n\in\mathbb{N}} K_{\leqslant n}[x]$.
- $\mathbb{R}^3 = \mathrm{lin}((1,0,0),(0,1,0)) + \mathrm{lin}((1,0,0),(1,1,0)) + \mathrm{lin}((1,1,0),(0,1,0))$.
- Przestrzeń $\mathrm{lin}((1,0,1),(1,0,2)) \subseteq \mathbb{R}^3$ jest sumą rodziny podprzestrzeni indeksowanej (na przykład) liczbami niewymiernymi postaci:

$$V_t = \mathrm{lin}(1,0,t), \quad t \in T = \mathbb{R} \setminus \mathbb{Q}.$$

Obserwacja 8.3

Niech $\{V_t\}_{t\in T}$ będzie rodziną podprzestrzeni przestrzeni V. Wówczas

$$\sum_{t\in T} V_i = \mathrm{lin}\left(\bigcup_{t\in T} V_t\right).$$

W szczególności, suma rodziny podprzestrzeni V jest podprzestrzenią V. Jest to najmniejsza podprzestrzeń w V zawierająca wszystkie $\{V_t\}_{t\in T}$.

Wówczas podprzestrzeń

$$\bigcap_{t\in T} V_t$$

nazywamy ILOCZYNEM, PRZECIĘCIEM lub CZĘŚCIĄ WSPÓLNĄ rodziny $\{V_t\}_{t\in T}$.

Rozważmy następujący przykład: niech $V_{[0,1]} \subseteq F(\mathbb{R},\mathbb{R})$ będzie podprzestrzenią złożoną ze wszystkich funkcji, które przyjmują wartość zero na zbiorze $[0,1]$. Rozważmy też, dla $x \in [0,1]$, podprzestrzenie $F(\mathbb{R},\mathbb{R})$ postaci $V_x = \{f \in F(R,R) : f(x) = 0\}$. Wówczas:

$$V_{[0,1]} = \bigcap_{x\in[0,1]} V_x$$

Czytelnik znający zasadę włączeń i wyłączeń, pozwalającą wyznaczyć moc sumy skończenie wielu zbiorów skończonych, patrząc na formułę Grassmanna może dojść do przekonania, że zachodzić musi jej uogólnienie na przypadek wymiaru sumy trzech lub więcej składników. Jest to niestety nieprawda. W szczególności, jeśli V_1, V_2, V_3 są podprzestrzeniami V, to $\dim(V_1 + V_2 + V_3)$ nie jest równy:

$$\dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) + \dim(V_1 \cap V_2 \cap V_3),$$

Czytelnik zechce sprawdzić to dla $V_t = \mathrm{lin}(1,t)$, gdzie $t = 0,1,2$. Prawidłowe uogólnienie formuły Grassmanna nie jest proste, o ile suma podprzestrzeni nie spełnia jakiegoś dodatkowego warunku. Najistotniejszym przykładem takiego warunku jest oczywiście bycie sumą prostą. Co to oznacza?

Definicja 8.4

Mówimy, że przestrzeń V jest SUMĄ PROSTĄ RODZINY PODPRZESTRZENI $\{V_t\}_{t\in T}$ jeśli każdy wektor $\alpha \in V$ daje się przedstawić jednoznacznie jako suma

$$\alpha_{t_1} + \ldots + \alpha_{t_r},$$

gdzie $\alpha_{t_i} \in V_{t_i}$, dla pewnych parami różnych $t_i \in T$. Wówczas piszemy:

$$V = \bigoplus_{t\in T} V_t,$$

a w przypadku, gdy $T = \{1, \ldots, n\}$ po prostu

$$V = V_1 \oplus \ldots \oplus V_n.$$

Przykłady (zachęcam do samodzielnego uzasadnienia):

- $\mathbb{R}^4 = \mathrm{lin}(1,0,0,0) \oplus \mathrm{lin}(0,1,0,0) \oplus \mathrm{lin}((1,1,1,0),(0,0,0,1))$.
- Jeśli $\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}$ jest bazą przestrzeni V, to:
$$V = \mathrm{lin}(\alpha_1) \oplus \mathrm{lin}(\alpha_2) \oplus \ldots \oplus \mathrm{lin}(\alpha_n).$$
- Jeśli $\mathcal{A}$ jest bazą V oraz $\mathcal{A} = \mathcal{A}_1 \cup \ldots \cup \mathcal{A}_n$ jest rozbiciem $\mathcal{A}$ na parami rozłączne podzbiory, to:
$$V = \bigoplus_{i=1}^{n} \mathrm{lin}(\mathcal{A}_i).$$

Jak się okazuje, jeśli przestrzeń skończenie wymiarowa V spełnia $V = V_1 \oplus \ldots \oplus V_n$, to

$$\dim V = \sum_{i=1}^{n} V_i.$$

Dowód wymaga uzasadnienia następującej obserwacji.

Obserwacja 8.4

Niech $\{V_t\}_{t\in T}$ będzie rodziną podprzestrzeni przestrzeni V. Wówczas $V = \bigoplus_{t\in T} V_t$ wtedy i tylko wtedy, gdy $V = \sum_{t\in T} V_t$ oraz dla każdych $t_0, t_1, \ldots, t_k \in T$ zachodzi: $V_{t_0} \cap \sum_{i=1}^{k} V_{t_i} = \{0\}$.

W szczególności aby suma algebraiczna była sumą prostą $V = V_1 \oplus V_2 \oplus V_3$ nie wystarczy, aby mieć

$$V = V_1 + V_2 + V_3, \quad \text{oraz} \quad V_1 \cap V_2 \cap V_3 = \{0\}.$$

W szczególności aby suma algebraiczna była sumą prostą $V = V_1 \oplus V_2 \oplus V_3$ nie wystarczy, aby mieć

$$V = V_1 + V_2 + V_3, \quad \text{oraz} \quad V_1 \cap V_2 \cap V_3 = \{0\}.$$

Dla przykładu weźmy podprzestrzenie $\mathbb{R}^2$ postaci:

$$V_1 = \mathrm{lin}(1,0), \quad V_2 = \mathrm{lin}(1,1), \quad V_3 = \mathrm{lin}(0,1).$$

Właściwe uogólnienie Obserwacji 8.2 wymaga zastąpienia warunku $V_1 \cap V_2 \cap V_3 = \{0\}$ układem warunków:

$$V_1 \cap (V_2 + V_3) = \{0\}, \quad V_2 \cap (V_1 + V_3) = \{0\}, \quad V_3 \cap (V_1 + V_2) = \{0\}.$$

Rozkłady na sumy proste mają wielkie znaczenie dla lepszego zrozumienia wykładu w drugim semestrze, choć nie są one, z uwagi na brak miejsca i godzin wykładowych, szerzej omówione w skrypcie, na którym się opieramy. Nietrudno się o tym przekonać próbując uogólnić przykład podany na zakończenie zasadniczej części wykładu. Gdy zapoznamy się (a zajmie nam to pozostałą część semestru) z językiem niezbędnym do badania przekształceń liniowych (zarówno macierzowym, jak i szkicowo – diagramowym), wówczas przyjdzie czas na badanie niezmienników przekształceń liniowych. Wiele z nich łatwiej będzie zrozumieć wiążąc z przekształceniami liniowymi rozkłady na sumy proste, związane z tzw. przestrzeniami niezmienniczymi. Póki co zajmiemy się jednak inną ważną strukturą związaną z podprzestrzeniami.

8.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Znajdowanie baz i wymiarów sum i przecięć podprzestrzeni). Niech V_1 i v_2 będą następującymi podprzestrzeniami przestrzeni $\mathbb{R}^n$. Znajdź bazy i wymiary przestrzeni $V_1 + V_2$ oraz $V_1 \cap V_2$.
   (a) $V_1 = \mathrm{lin}((2,1,3,4),(3,9,3,9),(-1,7,-3,1)), V_2 = \mathrm{lin}((1,-3,3,0),(2,5,3,5),(1,8,0,5))$,
   (b) $V_1 = \mathrm{lin}((3,2,1,0),(4,3,0,2),(1,2,2,-3))$, zaś V_2 jest opisana układem równań:
   $$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 0 \\ 3x_1 + 5x_2 + x_3 - 5x_4 = 0 \end{cases}.$$
   (c) V_1, V_2 są opisane układami równań liniowych, odpowiednio U_1, U_2:
   $$U_1 : \begin{cases} 2x_1 + x_2 - x_3 + 4x_4 = 0 \\ 3x_1 - x_2 + 2x_3 + x_4 = 0 \end{cases}, \quad U_2 : \begin{cases} -x_1 + 2x_2 - 5x_3 + 3x_4 = 0 \\ 2x_1 - 4x_2 + 10x_3 - 6x_4 = 0 \end{cases}.$$
2. (♠ Rozstrzyganie kiedy suma podprzestrzeni jest prosta).
   (a) Niech $V_1 \subset \mathbb{R}^3$ będzie podprzestrzeniami opisanymi równaniami $x_1 + 2x_2 - x_3 = 0$ i niech $V_2 = \mathrm{lin}((2,-t+2,4),(2s,6,-8)) \subset \mathbb{R}^3$. Dla jakich wartości paramerów $s,t \in \mathbb{R}$ zachodzi $\mathbb{R}^3 = V_1 \oplus V_2$?
   (b) Niech V_1, V_2 będą podprzestrzenimi przestrzeni $\mathbb{R}^4$, przy czym $V_1 = \mathrm{lin}((1,1,1,2),(2,0,1,3),(0,2,1,1))$ oraz V_2 jest opisana układem równań $x_1 + x_2 - x_3 = 0, x_2 + tx_4 = 0$. Znajdź wszystkie takie $t \in \mathbb{R}$, że $\mathbb{R}^4 = V_1 \oplus V_2$.
   (c) Niech $A = \mathrm{lin}(-2,1,0,-3),(2,-1,1,3))$. Znajdź takie podprzestrzenie A i B przestrzeni $\mathbb{R}^4$, by $\mathbb{R}^4$ było sumą prostą A i V, a nie było sumą prostą podprzestrzeni B i V ani A i B.
   (d) Niech $A = \mathrm{lin}((1,2,3,4),(4,3,2,1),(2,3,4,5)) \subseteq \mathbb{R}^4$. Znajdź takie podprzestrzenie $B, C \subseteq \mathbb{R}^4$, że $\mathbb{R}^4 = A \oplus B = B \oplus C = C \oplus A$ lub wykaż, że takie podprzestrzenie nie istnieją.
3. Dla podprzestrzeni V_1 przestrzeni V znaleźć taką podprzestrzeń V_2, aby $V_1 \oplus V_2 = V$, jeśli
   (a) $V = M_{n\times n}(K)$, $V_1 = \{[a_{ij}] \in M_{n\times n}(K) : a_{ij} = 0.\text{ dla } i > j\}$.
   (b) V — ciągi zbieżne o wyrazach w $\mathbb{R}$, V_1 — ciągi stałe.
   (c) $V = F(\mathbb{R},\mathbb{R})$, $V_1 = \{f \in V : f(0) = f(1) = 0\}$.
4. (♠ Wyciąganie prostych wniosków z formuły Grassmanna).
   (a) Czy istnieje przestrzeń liniowa V wymiaru 7 zawierająca podprzestrzenie W_1, W_2 takie, że $\dim W_1 = 4$, $\dim W_2 = 5$, $\dim(W_1 \cap W_2) = 1$?
   (b) Niech W_1 i W_2 będą podprzestrzeniami liniowymi przestrzeni liniowej V oraz $\dim V = 5$, $\dim W_1 = \dim W_2 = 4$. Czy wymiar przestrzeni $W_1 \cap W_2$ może być równy 2?
   (c) Czy istnieją podprzestrzenie V_1 i V_2 przestrzeni $\mathbb{R}^7$, że $\dim(V_1 \cap V_2) = 2$ i $\dim V_1 = \dim V_2 = 5$?
   (d) Niech $V_1, V_2 \subset \mathbb{R}^6$ będą podprzestrzeniami wymiaru 5. Czy możliwe jest, aby $\dim(V_1 \cap V_2) = 1$?
   (e) W przestrzeni $\mathbb{R}^{11}$ dane są podprzestrzenie V, W, przy czym $\dim V = 6$ oraz $\dim W = 8$. Czy przestrzeń $V \cap W$ może mieć wymiar 5?
   (f) Dane są podprzestrzenie liniowe $V \subseteq W \subseteq \mathbb{R}^6$, przy czym $\dim V = 5$ oraz $W \neq \mathbb{R}^6$. Czy wynika z tego, że $V = W$?
5. Załóżmy, że $U, W \neq \{0\}$ są podprzestrzeniami przestrzeni liniowej V. Przypuśćmy, że istnieje taka funkcja $f : V \to \mathbb{R}$, że $f(u) < f(w)$, dla dowolnych niezerowych wektorów $u \in U$ oraz $w \in W$. Uzasadnij, że $\dim U + \dim W \leqslant \dim V$.
6. Załóżmy, że U, W skończenie wymiarowymi podprzestrzeniami przestrzeni liniowej V. Pokaż, że:
   - gdy $\dim U \leqslant \dim W$ oraz $\dim(U + W) = \dim(U \cap W) + 1$, to $U \subseteq W$.
   - gdy $\dim U < \dim W$ oraz $\dim(U + W) = \dim(U \cap W) + 2$, to $U \subseteq W$.
7. Niech V_1, V_2 będą n wymiarowymi podprzestrzeniami skończenie wymiarowej przestrzeni liniowej V. Załóżmy, że układ $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ jest bazą V_1 oraz układ $\{\beta_1, \beta_2, \ldots, \beta_n\}$ jest bazą V_2. Wykaż, że układ $\{\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n\}$ jest bazą $V_1 + V_2$ wtedy i tylko wtedy, gdy $V_1 \cap V_2 = \{0\}$. Wykaż, że istnieje podprzestrzeń W przestrzeni V taka, że $V_1 \oplus W = V_2 \oplus W = V$.

8.3 Uzupełnienie. Wstęp do przestrzeni ilorazowych

Kończymy wstępne rozważania nad strukturą przestrzeni liniowych. W dodatku tym omówimy, na razie na bardzo elementarnym poziomie, strukturę zwaną przestrzenią ilorazową. O czym mowa? Zacznijmy od prostej sytuacji w $\mathbb{R}^3$ i od układu złożonego z dwóch równań postaci:

$$\begin{cases} x_1 & = a, \\ x_2 + x_3 & = b. \end{cases}$$

gdzie $a, b \in \mathbb{R}$ są parametrami. Zbiór rozwiązań tego układu to zbiór $\{(a, b-s, s), s \in \mathbb{R}\}$. W języku twierdzenia Kroneckera-Capellego jest to zbiór $(a, b, 0) + \mathrm{lin}((0, -1, 1))$, gdzie $W = \mathrm{lin}(0, -1, 1)$ jest przestrzenią rozwiązań jednorodnego układu równań (o trzech niewiadomych) $x_1 = 0, x_2 + x_3 = 0$.

Popatrzmy na uzyskane zbiory rozwiązań, dla różnych a, b. Geometrycznie można o nich myśleć jak o prostych przechodzących przez punkt $(a, b, 0)$ i równoległych do prostej o wektorze kierunkowym $(0, -1, 1)$. Załóżmy, że interesuje nas geometria nie całej przestrzeni $\mathbb{R}^3$, ale właśnie tych różnych prostych. Każda z nich opisana jest wektorem $(a, b, 0)$ i w tym sensie każdą z nich można traktować jako wektor! Co więcej, dlaczego by nie dodawać tych prostych? Ich „suma", czyli $(a_1, b_1, 0) + (a_2, b_2, 0)$ to prosta przechodząca przez $(a_1 + a_2, b_1 + b_2, 0)$ równoległa do $\mathrm{lin}(0, -1, 1)$. A zatem na zbiorze różnych zbiorów rozwiązań układów niejednorodnych, o jednym i tym samym układzie jednorodnym, można jak się wydaje wprowadzić pojęcie przestrzeni liniowej. Poświęćmy kilka uwag pojęciom znajdującym się dookoła tej idei.

Definicja 8.5: Układ liniowo niezależny modulo podprzestrzeń

Niech U będzie podprzestrzenią przestrzeni liniowej V nad ciałem K. Powiemy, że układ wektorów $v_1, \ldots, v_n \in V$ jest LINIOWO NIEZALEŻNY MODULO U, jeśli dla $a_1, \ldots, a_n \in K$ mamy

$$a_1 v_1 + \ldots + a_n v_n \in U \Rightarrow a_1 = \ldots = a_n = 0.$$

Powiemy, że układ liniowo niezależny modulo U jest BAZĄ V MODULO U, jeśli jest maksymalnym liniowo niezależnym układem modulo U.

Oczywiście liniowa niezależność układu modulo podprzestrzeń zerowa jest zwykłą liniową niezależnością rozważaną na wykładzie. Co więcej, jeśli $U' \subseteq U$ jest podprzestrzenią, to układ wektorów jest liniowo niezależny modulo U tylko wtedy, gdy jest liniowo niezależny modulo U'. W szczególności każdy układ liniowo niezależny modulo U jest układem liniowo niezależnym w V w myśl definicji z wykładu czwartego.

Przykład 1. Niech $V = \mathbb{R}^4$ oraz $U = \{(x_1, x_2, x_3, x_4) \in V : x_1 + 2x_4 = 0, x_1 + 2x_2 - 2x_3 + 4x_4 = 0\}$. Wówczas przykładem bazy przestrzeni V modulo U jest np. para wektorów

$$v_1 = (1, 0, 0, 0), v_2 = (0, 1, 0, 0).$$

Dlaczego? Zauważmy, że jeśli liniowa kombinacja $av_1 + bv_2 = (a, b, 0, 0)$ należy do U, to musi spełniać obydwa równania $x_1 + 2x_4 = 0$ oraz $x_1 + 2x_2 - 2x_3 + 4x_4 = 0$, czyli $a = 0$ oraz $a + 2b = 0$, co daje oczywiście $a = b = 0$, A zatem $\{v_1, v_2\}$ to układ liniowo niezależny modulo U. Dlaczego jest to układ maksymalny? Gdyby dało się go rozszerzyć przy pomocy wektora v_3, to warunek $a_1 v_1 + a_2 v_2 + a_3 v_3 \in U$ oraz $v_1, v_2 \notin U$ implikują $v_3 \in U$, co daje sprzeczność z liniową niezależnością układu $\{v_1, v_2, v_3\}$ modulo U, bowiem można przyjąć $a_1 = 0, a_2 = 0, a_3 = 1$.

Czy Czytelnik widzi, że dla każdego układu niejednorodnego, któremu odpowiada układ jednorodny o zbiorze rozwiązań U istnieją $a, b \in K$, że $(av_1 + bv_2) + U$ jest zbiorem rozwiązań tego układu niejednorodnego? Zbiory $(1, 0, 0, 0) + U$ oraz $(0, 1, 0, 0) + U$ to rozwiązania odpowiednio układów postaci:

$$\begin{cases} x_1 + 2x_4 & = 1 \\ x_1 + 2x_2 - 2x_3 + 4x_4 & = 0 \end{cases}, \quad \begin{cases} x_1 + 2x_4 & = 0 \\ x_1 + 2x_2 - 2x_3 + 4x_4 & = 2. \end{cases}.$$

Podkreślmy raz jeszcze: wektory $(1, 0, 0, 0)$, $(0, 1, 0, 0)$ są rozwiązaniami układów niejednorodnych, wiec nie mogą należeć do U, które jest zbiorem rozwiązań odpowiadającego tym układom układu jednorodnego!

Przykład 2. Niech $X = \mathbb{R}[x]$ oraz

$$U = \{f \in X : f(0) = f(1) = 0\}, \qquad V = \{f \in X : f(0) = f(1)\}.$$

Dokonujemy tu cichego utożsamienia funkcji wielomianowych nad $\mathbb{R}$ z wielomianami. Z uwagi na to, że ciało $\mathbb{R}$ jest nieskończone, jest to możliwe. Zarówno U, jak i W są oczywiście nieskończonego wymiaru. Istotnie, mają one postać:

$$U = \{x(x-1)f(x) : f \in K[x]\}, \quad V = \{a + x(x-1)f(x) : a \in \mathbb{R} \text{ oraz } f \in K[x]\}.$$

Układem liniowo niezależnym modulo U jest choćby $\{1, x\}$. Czy można go rozszerzyć? Jeśli tak, to wielomian rozszerzający ten układ musi być stopnia co najmniej drugiego (układ liniowo niezależny modulo U musi być liniowo niezależny, patrz komentarz wyżej). Jednak dla każdego wielomianu f stopnia większego niż 1 istnieje kombinacja liniowa $a_1 + a_2x + a_3f$ taka, że powstały wielomian jest podzielny przez $x(x-1)$. Zachęcam Czytelnika do wykazania takiej tezy.

Układy liniowo niezależne i bazy modulo podprzestrzeń są istotne z punktu widzenia sum prostych. Jeśli zastanawiamy się jak mogą wyglądać wszystkie możliwe bazy modulo podprzestrzeń, to okazuje się, że muszą to być bazy dopełnień prostych tej podprzestrzeni do całej przestrzeni Mówi o tym poniższa uwaga.

Obserwacja 8.5

Niech $U \subsetneq V$ i niech $v_1, \ldots, v_n$ będzie układem liniowo niezależnym modulo U. Wówczas następujące warunki są równoważne:

(1) układ $v_1, \ldots, v_n$ jest bazą V modulo U,

(2) $U \oplus \operatorname{lin}(v_1, \ldots, v_n) = V$.

Dowód. Pokażemy tezę jedynie w przypadku, gdy V jest skończenie wymiarowa, choć jest ona zawsze prawdziwa. Załóżmy (1) i niech $u_1, \ldots, u_m$ będzie bazą U. Nasza teza: $v_1, \ldots, v_n, u_1, \ldots, u_m$ jest bazą V. Zacznijmy od liniowej niezależności tego układu. Weźmy $a_1, \ldots, a_n, b_1, \ldots, b_m \in K$ takie, że:

$$a_1v_1 + \ldots + a_nv_n + b_1u_1 + \ldots + b_mu_m = 0.$$

Wówczas $a_1v_1 + \ldots + a_nv_n = -(b_1u_1 + \ldots + b_mu_m) \in U$, a skoro $v_1, \ldots, v_n$ to układ liniowo niezależny modulo U, to $a_1 = \ldots = a_n = 0$. Zatem w wypisanej wyżej kombinacji mamy tylko $b_1u_1 + \ldots + b_mu_m = 0$, co oznacza, ze $b_1 = \ldots = b_m = 0$, skoro $u_1, \ldots, u_m$ są bazą U. A zatem układ $v_1, \ldots, v_n, u_1, \ldots, u_m$ jest liniowo niezależny. Pokażmy, że układ ten rozpina V. Weźmy $w \in V$. Twierdzimy mianowicie, ze istnieją $c_1, \ldots, c_n$ takie, że

$$w - c_1v_1 - \ldots - c_nv_n \in U.$$

W przeciwnym bowiem razie, układ $w, v_1, \ldots, v_n$ jest liniowo niezależny modulo U. Istotnie, gdyby dla pewnych $d, d_1, \ldots, d_n \in K$, nie wszystkich równych zero było $dw + d_1v_1 + \ldots + d_nv_n \in U$, to albo $d \neq 0$ i wtedy $w - d_1v_1 + \ldots - d_nv_n \in U$, albo $d = 0$ i któryś z $d_i \neq 0$, co oznacza, że układ $v_1, \ldots, v_n$ jest liniowo zależny modulo U, sprzeczność. A zatem, wbrew założeniu, że $v_1, \ldots, v_n$ jest bazą V modulo U znaleźliśmy zawierający ją w sposób właściwy układ $w, v_1, \ldots, v_n$, również liniowo niezależny modulo U. A zatem istotnie istnieją $c_1, \ldots, c_n$ takie, że $w - c_1v_1 - \ldots - c_nv_n \in U$, czyli w jest kombinacją wektorów $v_1, \ldots, v_n, u_1, \ldots, u_m$.

Pokazaliśmy, że $v_1, \ldots, v_n, u_1, \ldots, u_m$ jest bazą V. A zatem

$$U + \operatorname{lin}(v_1, \ldots, v_n) = \operatorname{lin}(\operatorname{lin}(u_1, \ldots, u_m) \cup \operatorname{lin}(v_1, \ldots, v_m)) = \operatorname{lin}(v_1, \ldots, v_n, u_1, \ldots, u_m) = V$$

oraz oczywiście $\operatorname{lin}(v_1, \ldots, v_n) \cap \operatorname{lin}(u_1, \ldots, u_m) = 0$, bo jeśli jakiś wektor w należy do części wspólnej, to

$$w = e_1v_1 + \ldots + e_nv_n = f_1u_1 + \ldots + f_mu_m,$$

czyli $e_1v_1 + \ldots + e_nv_n - f_1u_1 + \ldots - f_mu_m = 0$, co wobec faktu, że $v_1, \ldots, v_n, u_1, \ldots, u_m$ jest bazą V oznacza, że $e_1 = \ldots = e_n = f_1 = \ldots = f_m = 0$.

Pokazaliśmy zatem, że $U \oplus \operatorname{lin}(v_1, \ldots, v_n) = V$. Dowód drugiej implikacji zostawiamy jako ćwiczenie. □

Widzimy zatem, że każda baza przestrzeni V modulo U odpowiada pewnemu dopełnieniu prostemu U do V, i odwrotnie. Konstrukcja przestrzeni ilorazowej wskazuje jeden „kanoniczny” obiekt, który będzie „izomorficzny” (identyczny, co do struktury – patrz kolejny wykład) z każdym z tych dopełnień.

Definicja 8.6: Warstwa podprzestrzeni

Niech W będzie podprzestrzenią przestrzeni V nad ciałem K i niech $\alpha \in V$. Zbiór

$$\alpha + W = \{\alpha + \gamma \,|\, \gamma \in W\}$$

nazywamy WARSTWĄ PODPRZESTRZENI W w przestrzeni V.

Warstwa jest, jak widzimy, abstrakcyjnym odpowiednikiem zbioru rozwiązań niejednorodnego układu równań. Celem jest, jak wspominaliśmy, wprowadzenie struktury przestrzeni liniowej na zbiorze warstw (tak, jak wprowadziliśmy ją w początkowym przykładzie na zbiorze prostych równoległych do danej).

Fundamentalne obserwacje dotyczące warstw zawarte są w następującej uwadze[2].

Obserwacja 8.6

Niech W będzie podprzestrzenią przestrzeni V.

(i) $\alpha + W = \beta + W \iff \alpha - \beta \in W$,

(ii) Dla warstw $v + W$ oraz $v' + W$ określamy sumę warstw $+$ oraz iloczyn $\cdot$ skalara z ciała K przez warstwę:

$$(v + W) + (v' + W) = (v + v') + W, \quad a \cdot (v + W) = av + W.$$

Działania te są dobrze określone, tzn. jeśli $v + W = v' + W$, to dla każdego $v'' \in V$ oraz dla każdego $a \in K$ mamy: $(v+W)+(v''+W) = (v'+W)+(v''+W)$ oraz $a\cdot(v+W) = a\cdot(v'+W)$.

(iii) zbiór $V/W = \{\alpha + W \,|\, \alpha \in V\}$ z działaniami dodawania i mnożenia przez skalar określonymi wyżej oraz z warstwą $0 + W$ tworzy przestrzeń liniową nad ciałem K,

(iv) Jeśli $\mathcal{A} = \{v_1, \ldots, v_n\}$ jest bazą W oraz $\mathcal{B} = \{y_1, \ldots, y_m\}$ ma tę własność, że $\{y_1+W, \ldots, y_m+W\}$ to baza V/W, wówczas $\mathcal{A} \cap \mathcal{B} = \emptyset$ oraz $\mathcal{A} \cup \mathcal{B}$ jest bazą V. W szczególności, jeśli V jest skończenie wymiarowa, to V/W też jest skończenie wymiarowa i $\dim V = \dim W + \dim(V/W)$.

Definicja 8.7: Przestrzeń ilorazowa

Przestrzeń V/W określoną w poprzedniej uwadze nazywamy PRZESTRZENIĄ ILORAZOWĄ przestrzeni V przez podprzestrzeń W.

Dowód. Dowodzimy kolejne punkty.

- Zacznijmy od (i). Załóżmy, że $v + W = v' + W$. Skoro $v \in v + W$, to $v \in v' + W$. Stąd istnieje $w \in W$ takie, że $v = v' + w$. Stąd $v - v' \in W$. W drugą stronę, załóżmy, że $v - v' \in W$. Bez straty ogólności wystarczy pokazać, że $v \in v' + W$. Niech $w = v - v' \in W$. Wówczas $v = v' + w$, a stąd $v \in v' + W$, co pokazuje $v + W \subseteq v' + W$. Drugie zawieranie pokazujemy w sposób analogiczny.

- Dowodzimy (ii). Pokażmy, że dodawanie warstw jest dobrze zdefiniowane. Warunek jest symetryczny, więc wystarczy pokazać zawieranie $(v + W) + (v'' + W) \subseteq (v' + W) + (v'' + W)$. Niech $u \in (v + W) + (v'' + W) = (v + v'') + W$. Istnieje $w \in W$, że $u = (v + v'') + w$. Chcemy pokazać, że $u \in (v' + W) + (v'' + W) = (v' + w'') + W$. Skoro $v + W = v' + W$, to mamy element $w' = v - v' \in V$. A zatem $v = v' + w'$. Stąd: $u = (u + v'') + w = ((v' + w') + v'') + w = (v' + v'') + (w' + w)$. A zatem rzeczywiście $u \in (v' + v'') + W$, skoro $w' + w \in W$. W rezultacie $(v + v'') + W = (v' + v'') + W$, czyli dodawanie warstw jest dobrze określone.

  Niech teraz $a \in K$. Aby pokazać, że mnożenie warstwy przez skalar jest dobrze określone, wystarczy pokazać, że: $a \cdot (v + W) \subseteq a \cdot (v' + W)$. Niech $u \in a \cdot (v + W) = av + W$. Mamy $u = av + w$, dla

[2] Dowód jest w zasadzie tłumaczeniem tekstu S. Caneza: *Notes on quotient spaces*, do wyszukania w Sieci.

pewnego $w \in W$. Ponownie oznaczmy $w' = v - v' \in W$. Zatem $u = av + w = a(v' + w') + w = av' + (aw' + w) \in av' + W$, bowiem $aw' + w \in W$, bo W jest podprzestrzenią. Zatem zbiory

$$a \cdot (v + W), \quad a \cdot (v' + W)$$

są równe i mnożenie warstwy przez skalar jest dobrze określone.

- Sprawdzenie, że V/W spełnia aksjomaty przestrzeni liniowej sprowadza się w większości przypadków (łączność, przemienność, rozdzielność) do skorzystania z tego, że V jest liniowa. Sprawdzimy jedynie istnienie elementu zerowego i przeciwnego.

  Twierdzimy, że $W = 0 + W$ jest zerem w V/W. Istotnie, niech $v + W \in V/W$. Wówczas:

  $$(v + W) + W = (v + 0) + W = v + W, \quad W + (v + W) = (0 + v) + W = v + W,$$

  co załatwia sprawę. Oczywiście biorąc $v + W \in V/W$ widzimy, że warstwą przeciwną jest $-v + W$.

- Dowód (iv) przypomina uzasadnienie Obserwacji 8.5. Weźmy element v z bazy $\mathcal{A} \subset W$. Wówczas mamy $v + W = W$, czyli jest to warstwa zerowa. Nie może ona należeć do żadnej bazy V/W, czyli $\mathcal{A} \cap \mathcal{B} = \emptyset$.

  Pokażmy, że $\mathcal{A} \cup \mathcal{B}$ rozpina V. Niech $v \in V$. Zatem $v + W \in V/W$ i istnieją skalary $t_1, \ldots, t_m$ takie, że

  $$v + W = t_1(y_1 + W) + \ldots + t_m(y_m + W) = (t_1 y_1 + \ldots + t_m y_m) + W.$$

  Na mocy (i) mamy $v - t_1 y_1 + \ldots + t_m y_m \in W$, a więc istnieją $s_1, \ldots, s_n$, ze:

  $$v - t_1 y_1 + \ldots + t_m y_m = s_1 v_1 + \ldots + s_n v_n.$$

  Widzimy zatem, że $v \in \operatorname{lin}(\mathcal{A} \cup \mathcal{B})$.

  Pokażmy wreszcie, że $\mathcal{A} \cup \mathcal{B}$ to zbiór liniowo niezależny. Załóżmy, że dla pewnych $s_1, \ldots, s_n$ oraz $t_1, \ldots, t_m$ z ciała K mamy $s_1 v_1 + \ldots + s_n t_n + t_1 y_1 + \ldots + t_m y_m = 0$. Zatem $s_1 v_1 + \ldots + s_n v_n = -(t_1 y_1 + \ldots + t_m y_m) \in W$, a skoro $v_1, \ldots, v_n$ to baza W, dostajemy $s_1 = \ldots = s_n = 0$. A zatem mamy $t_1 y_1 + \ldots + t_m y_m = 0$. To oznacza, że $t_1(y_1 + W) + \ldots + t_m(y_m + W) = 0 + W$. Ale $y_1 + W, \ldots, y_m + W$ to baza V/W, czyli $t_1 = \ldots = t_m = 0$, co kończy dowód.

□

Zachęcam Czytelnika do pokazania w analogiczny sposób wariantu tezy postawionej w (iv): jeśli $\mathcal{C}$ jest bazą V taką, że $\mathcal{A} \subseteq \mathcal{C}$ jest bazą W, to układ $\{v + W, v \in \mathcal{C} \setminus \mathcal{A}\}$ jest bazą V/W.

Czytelnik zechce zauważyć, że dwa przykłady rozważane wcześniej pokazują, że:

$$\mathbb{R}^4/U = \operatorname{lin}((1,0,0,0) + U, (0,1,0,0) + U),$$

gdzie $U = \{(x_1, x_2, x_3, x_4) \in V : x_1 + 2x_4 = 0, x_1 + 2x_2 - 2x_3 + 4x_4 = 0\}$, czyli $\dim \mathbb{R}^4/U = 2$. Również w drugim z rozważanych przykładów, czyli $X = \mathbb{R}[x]$ oraz $U = \{f \in X : f(0) = f(1) = 0\}$, $V = \{f \in X : f(0) = f(1)\}$ postulowaliśmy w istocie, że:

$$\dim X/U = 2, \quad \dim X/V = 1, \quad \dim V/U = 1.$$

Definicja 8.8: Kowymiar

Niech $U \subseteq V$. Liczbę

$$\operatorname{codim} U := \dim V/U$$

nazywamy KOWYMIAREM przestrzeni U.

8.4 Dodatek. Krata podprzestrzeni przestrzeni liniowej

Omówione na wykładzie operacje sumy i części wspólnej podprzestrzeni przestrzeni V dają nam lepsze zrozumienie zależności pomiędzy podprzestrzeniami. Jest to jednak z konieczności spojrzenie „lokalne". Brakuje nam bowiem języka do zrozumienia struktury kombinatorycznej zbioru wszystkich podprzestrzeni przestrzeni liniowej V. Na to wyzwanie odpowiada tzw. TEORIA KRAT, o której powiemy kilka słów.

Zacznijmy od następującego problemu. Niech V będzie skończenie wymiarową przestrzenią liniową nad ciałem K oraz niech $U_1, \ldots, U_n$ będzie układem podprzestrzeni przestrzeni V. Niech $\mathcal{L}(U_1, \ldots, U_n)$ oznacza zbiór wszystkich podprzestrzeni, które można uzyskać startując z układu $U_1, \ldots, U_n$ i używając dowolnie wiele razy operacji sumy i przecięcia podprzestrzeni. Jak wygląda zbiór $\mathcal{L}$ i jaką ma moc?

Oczywiście dla $n = 1$ oraz $n = 2$ moc $\mathcal{L}$ nie przekracza odpowiednio 1 oraz 4. Rzeczywiście, suma lub iloczyn dowolnej podprzestrzeni z samą sobą jest jej równy. Jeśli zaś U_1, U_2 są podprzestrzeniami V, to:

$$\mathcal{L}(U_1, U_2) = \{U_1 \cap U_2, U_1, U_2, U_1 + U_2\}.$$

Dlaczego? Startując od U_1, U_2 i wykonując operację $+$ oraz $\cap$ dostajemy:

$$U_1 + U_1 = U_1, \quad U_2 + U_2 = U_2, \quad U_1 + U_2, \quad U_1 \cap U_1 = U_1, \quad U_2 \cap U_2 = U_2, \quad U_1 \cap U_2.$$

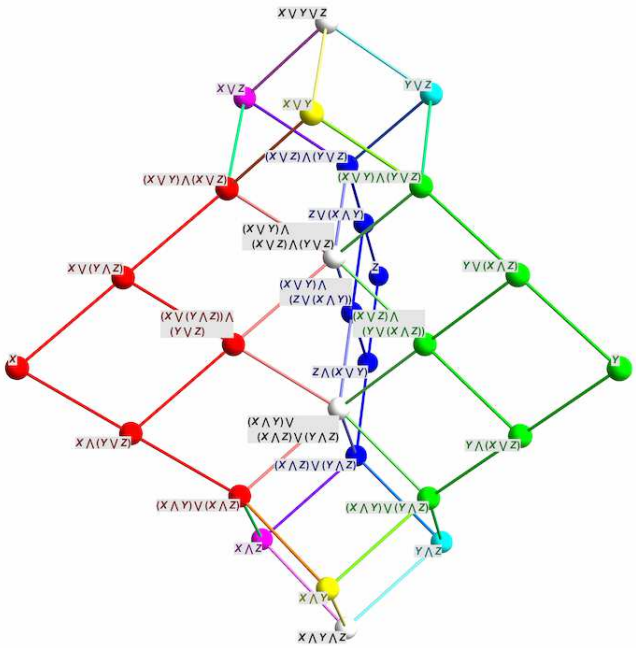
Niech $P = U_1 \cap U_2$ oraz $S = U_1 + U_2$. Ponawiamy stosowanie operacji $+$ oraz $\cap$ dostając:

$$U_i + P = U_i, \quad U_i \cap P = P, \quad U_i + S = S, \quad U_i \cap S = U_i, \quad P + S = S, \quad P \cap S = P.$$

Oczywiście nietrudno podać przykład sytuacji, gdy czwórka U_1, U_2, P, S zawiera parami różne podprzestrzenie. Jak wygląda sytuacja, gdy startujemy od trzech podprzestrzeni? Problem ten należy do klasyki.

Twierdzenie 8.2: Dedekind, 1900

Dla przestrzeni liniowej V oraz trójki jej podprzestrzeni X, Y, Z zbiór $\mathcal{L}(X, Y, Z)$ może mieć nie więcej niż 28 elementów, które zaprezentować można na następującym diagramie



Rys. 1. Krata 28-elementowa generowana w sposób wolny (jako krata modularna) przez podprzestrzenie X, Y, Z.
Źródło wraz z ładną wizualizacją 3D pod adresem:
`https://blogs.ams.org/visualinsight/2016/01/01/free-modular-lattice-on-3-generators/`

Przykład sytuacji, gdy $|\mathcal{L}(X, Y, Z)| = 28$ ma miejsce na przykład w niezupełnie banalnej sytuacji, gdy

$$V = \mathbb{R}^8, \quad X = \mathrm{lin}(\epsilon_2, \epsilon_4, \epsilon_5, \epsilon_8), \quad Y = \mathrm{lin}(\epsilon_2, \epsilon_3, \epsilon_6, \epsilon_7), \quad Z = \mathrm{lin}(\epsilon_1, \epsilon_4, \epsilon_6, \epsilon_7 + \epsilon_8).$$

Dedekind pokazał również, że dla $n \geqslant 4$ istnieje przestrzeń liniowa V oraz podprzestrzenie $U_1, \ldots, U_n$ takie, że $\mathcal{L}(U_1, \ldots, U_n)$ jest zbiorem nieskończonym. Jeden z poważnych i trudnych problemów teorii krat mający odniesienie do współczesnej matematyki[3], polega na zrozumieniu „geometrycznej" struktury „generowanej" przez cztery podprzestrzenie[4] znajdujące się „w położeniu ogólnym".

Spróbujmy objaśnić rysunek przedstawiony na poprzedniej stronie oraz przekonać Czytelnika, że problem podprzestrzeni ma pewne ogólne metody i wyniki niezwykle istotne w matematyce. Przede wszystkim widzimy na rysunku pewien graf, czy też diagram, którego wierzchołkami są podprzestrzenie. Jak rozumieć krawędzie? Otóż dwie podprzestrzenie $W_1, W_2 \in \mathcal{L}(X, Y, Z)$ połączone są krawędzią, jeśli $W_1 \subseteq W_2$ oraz nie istnieje element $W \in \mathcal{L}(X, Y, Z)$ taki, że $W_1 \subsetneq W \subsetneq W_2$. Innymi słowy, jeśli określimy porządek w zbiorze $\mathcal{L}(X, Y, Z)$ przez relację inkluzji, to dwie podprzestrzenie połączone są krawędzią, to jedna położona jest w tym porządku „bezpośrednio nad" drugą. Takie diagramy (tzw. diagramy Hassego) można rysować dla każdego zbioru P z częściowym porządkiem $\leqslant$ (parę $(P, \leqslant)$ nazywamy też POSETEM).

Wreszcie, wyjaśnijmy znaczenie symboli $\vee, \wedge$, które oznaczają tzw. `operacje kratowe`.

Definicja 8.9: Krata

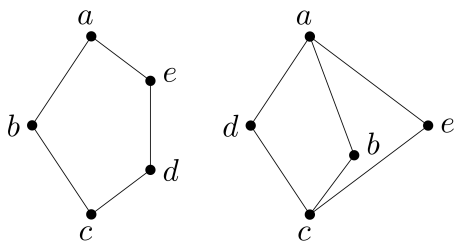
Niech A będzie niepustym zbiorem oraz $\vee, \wedge$ będą działaniami dwuargumentowymi w A. Mówimy, że trójka $(A, \vee, \wedge)$ jest KRATĄ, jeżeli dla dowolnych $x, y, z \in A$ następujące warunki:

1. $x \vee x = x, \quad x \wedge x = x,$
2. $(x \vee y) \vee z = x \vee (y \vee z), \quad (x \wedge y) \wedge z = x \wedge (y \wedge z),$
3. $x \vee y = y \vee x, \quad x \wedge y = y \wedge x,$
4. $(x \vee y) \wedge y = y, \quad (x \wedge y) \vee y = y.$

W każdej kracie spełniona jest równoważność $x \vee y = y \iff x \wedge y = x$. Relacja $\leqslant$, zdefiniowana na A za pomocą równoważności $x \leqslant y \iff x \vee y = y$, jest częściowym porządkiem, w którym każda para x, y ma ograniczenie górne $x \vee y$ oraz ograniczenie dolne $x \wedge y$. Z kratą związany jest więc porządek.

Przykłady:

- Jeśli $(A, \vee, \wedge)$ jest kratą i $B \subseteq A$, to jeśli B jest zamknięty na operacje $\vee, \wedge$ (tzn. dla dowolnych $x, y \in B$ mamy $x \vee y, x \wedge y \in B$), to B nazywamy podkratą kraty $(A, \vee, \wedge)$.
- Niech X będzie zbiorem oraz $P(X)$ – zbiorem jego podzbiorów. Wówczas zbiór $P(X)$ z działaniami $\vee$ – sumy zbiorów oraz $\wedge$ – części wspólnej zbiorów jest kratą.
- Niech V będzie przestrzenią liniową nad ciałem K, zaś $S(V)$ – zbiorem wszystkich podprzestrzeni przestrzeni V. Wówczas $S(V)$ z operacjami $\vee$ sumy podprzestrzeni oraz $\wedge$ – części wspólnej jest kratą. Jeśli $U_1, \ldots, U_n$ należą do $S(V)$, wówczas przez $\mathcal{L}(U_1, \ldots, U_n)$ oznaczamy część wspólną wszystkich podkrat $S(V)$, zawierających $U_1, \ldots, U_n$. Jest to, co łatwo pokazać, krata.
- Dwiema niezwykle istotnymi przykładami krat są tak zwany PIECIOKĄT i DIAMENT, czyli kraty na zbiorze pięcioelementowym $A = \{a, b, c, d, e\}$ reprezentowane za pomocą następujących diagramów:



Rys. 2. Kraty pięcioelementowe. Źródło: Wikipedia.

[3]Patrz artykuł Gian-Carlo Roty, *Ten Mathematics Problems I will never solve*, 1997: `https://www.degruyter.com/document/doi/10.1515/dmvm-1998-0215/html`

[4]Jest też tzw. problem czterech podprzestrzeni, badany m.in. przez Gelfanda, związany z tzw. teorią reprezentacji. Aby zrozumieć jak wiąże się z zagadnieniem Dedekinda, polecam: `https://golem.ph.utexas.edu/category/2015/09/the_free_modular_lattice_on_3.html` oraz `https://www.sciencedirect.com/science/article/pii/S0024379504002575`.

Pierwsza z nich, zwana pięciokątem lub kratą N_5 to krata, w której spełnione sa relacje:

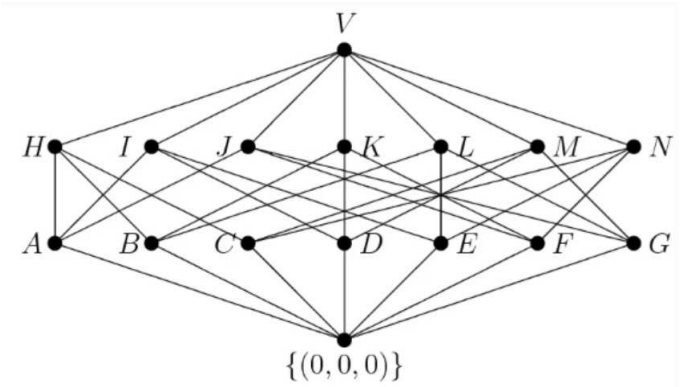
$$c \leqslant x \leqslant a, \text{ dla dowolnego x}, \quad d \wedge b = e \wedge b = c, \quad d \vee b = e \vee b = a$$

Diament lub krata M_3 to krata, w której spełnione są relacje:

$$c \leqslant a \leqslant x, \text{ dla dowolnego x}, \quad x \wedge y = c \text{ oraz } x \vee y = a, \text{ dla dowolnych } x \neq y \text{ w zbiorze } \{b.d.e\}.$$

Warto zauważyć, że $M_3 = S(\mathbb{Z}_2^2)$ jest kratą podprzestrzeni dwuwymiarowej przestrzeni liniowej nad ciałem skończonym $\mathbb{Z}_2$. Zachęcam Czytelnika do narysowania krat y podprzestrzeni w $\mathbb{Z}_3^2$.

- Krata podprzestrzeni przestrzeni $\mathbb{Z}_2^3$ jest bardziej skomplikowana i jej diagram ma postać:



Rys. 3. Krata podprzestrzeni przestrzeni trójwymiarowej $V = \mathbb{Z}_2^3$. Źródło: https://link.springer.com/article/10.1007/s00500-019-03866-y

- Zbiór liczb całkowitych $\mathbb{Z}$ z operacjami $\vee = NWW$ oraz $\wedge = NWD$ jest kratą, której odpowiada porządek częściowy wyznaczony przez podzielność.

Gdy rozważamy nową abstrakcyjną strukturę zawsze zastanawiamy się nad tym czy można ją realizować za pomocą szczególnych typów struktur (poprzez tzw. reprezentacje). Na przykład: czy pięciokąt może być kratą podprzestrzeni przestrzeni liniowej? Czy może być kratą podzbiorów pewnego zbioru? Odpowiedzi na te pytania prowadzą do niezwykle istotnych własności algebraicznych w teorii grup, modułów czy algebr. Powiedzmy o dwóch najbardziej znanych, nawiązujących do arytmetyki.

Definicja 8.10: Krata rozdzielna i krata modularna

Powiemy, że A jest KRATĄ ROZDZIELNĄ (DYSTRYBUTYWNĄ), jeśli dla dowolnych $x, y, z \in A$ mamy:

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z), \quad (x \wedge y) \vee z = (x \vee z) \wedge (y \vee z). \qquad (\spadesuit)$$

Mówimy, że krata A jest MODULARNA, jeśli warunki rozdzielności ($\spadesuit$) zachodzą dla takich trójek x, y, z, dla których zachodzi warunek $x \leqslant z$ (gdzie $\leqslant$ jest porządkiem wyznaczonym przez kratę A).

Dla każdego zbioru X, krata $(P(X), \cup, \cap)$ jest kratą rozdzielną. Podkrata kraty rozdzielnej też jest zawsze rozdzielna. Co więcej, ważne twierdzenie Birkhoffa-Stone'a z 1934 roku mówi, że krata jest rozdzielna wtedy i tylko wtedy, gdy jest izomorficzna[5] z pewną podkratą kraty $(P(X), \cup, \cap)$, dla pewnego zbioru X. Zauważmy, że krata podprzestrzeni przestrzeni liniowej wymiaru większego niż 1 nie jest nigdy rozdzielna. Aby to zrozumieć wystarczy popatrzeć na $S(V)$ dla $V = \mathbb{R}^2$ i rozważyć $x = \text{lin}((1, 0))$, $y = \text{lin}((0, 1)), z = \text{lin}((1, 1))$. Jak się natomiast okazuje, krata podprzestrzeni jest zawsze modularna.

Ciekawe, że pięciokąt i diament stanowią niezwykle istotne obiekty dla stwierdzania czy krata jest rozdzielna lub modularna. Żadna z tych dwóch krat nie jest, jak się okazuje, rozdzielna, Okazuje się, że krata jest rozdzielna wtedy i tylko wtedy, gdy żadna z jej podkrat nie zawiera ani diamentu, ani pięciokąta. Krata jest modularna wtedy i tylko wtedy, gdy nie ma podkraty zawierającej pięciokąt. Czytelnika zainteresowanego dowodami tych rezultatów oraz innymi ciekawostkami odsyłam choćby do artykułu dr Małgorzaty Jastrzębskiej *O pewnych kratach testowych* w czasopiśmie Delta (12/2013): https://www.deltami.edu.pl/temat/matematyka/algebra/2013/12/31/kraty.pdf.

[5]Nie mówimy tu czym jest izomorfizm krat – powiedzmy, że diagram dowolnej kraty rozdzielnej jest diagramem pewnej podkraty w kracie $(P(X), \cup, \cap)$, dla pewnego zbioru X. Szczegóły znaleźć można w dowolnym wykładzie algebry uniwersalnej lub teorii krat, na przykład w https://math.uwb.edu.pl/~mariusz/share/classes/tk/teoria_krat-w.pdf.

8.5 Trivia. Cykle i rozcięcia w grafach

W tym dodatku pokażemy ciekawe zastosowanie algebry liniowej w kombinatoryce, związane z tak zwanymi przestrzeniami cykli i rozcięć w grafach. Ustalmy kilka pojęć wstępnych.

Definicja 8.11: Graf niezorientowany (prosty)

Niech X będzie skończonym zbiorem niepustym, E zaś niech będzie podzbiorem zbioru par nieuporządkowanych zbioru X. Parę $G = (X, E)$ nazwiemy GRAFEM NIEZORIENTOWANYM o zbiorze wierzchołków $X = V(G)$ i zbiorze krawędzi $E = E(G)$. Jeśli $\{a, b\} \in E(G)$ to mówimy, że między wierzchołkami a, b grafu G jest KRAWĘDŹ oraz mówimy, że wierzchołki a, b sąsiadują ze sobą. Dodatkowo:

- STOPNIEM WIERZCHOŁKA $x \in V(G)$ w grafie G nazywamy liczbę $\deg(x)$ krawędzi, których jeden z elementów równy jest x,
- PODGRAFEM grafu $G = (X, E)$ nazywamy graf (X', E'), że $X' \subseteq X$ oraz $E' \subseteq E$, przy czym jeśli $\{a, b\} \in E'$, to $a, b \in X'$,
- ŚCIEŻKĄ nazywamy ciąg wierzchołków $x_0, x_1, \ldots, x_n$, taki że dla każdego $k \in \{0, 1, \ldots, n-1\}$ wierzchołki x_k oraz x_{k+1} są sąsiadami,
- DROGĄ nazywamy ciąg krawędzi $e_1 = \{x_0, x_1\}, e_2 = \{x_1, x_2\}, \ldots, e_n = \{x_{n-1}, x_n\}$, których zbiór wierzchołków $x_0, \ldots, x_n$ tworzy ścieżkę,
- CYKLEM nazywamy drogę zamkniętą, czyli taką, w której $x_0 = x_n$,

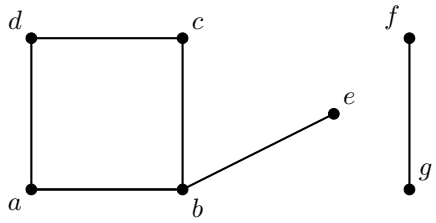
Co więcej, graf nazywamy:

- SPÓJNYM, jeśli każde dwa jego wierzchołki łączy ścieżka,
- ACYKLICZNYM, jeśli nie zawiera on cykli (jako podgrafów),
- DRZEWEM, jeśli jest spójny i acykliczny.

Mówimy też, że podgraf G' grafu G jest jego SPÓJNĄ SKŁADOWĄ, jeśli jest on spójny i nie jest zawarty w sposób właściwy w żadnym podgrafie spójnym grafu G.

Przykład. Rozważmy graf $G = (V(G), E(G))$, gdzie

- zbiór wierzchołków $V(G) = \{a, b, c, d, e, f, g\}$,
- zbiór krawędzi $E(G) = \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}, \{b, e\}, \{f, g\}\}$.



W tym grafie mamy:

- $\deg(e) = \deg(f) = \deg(g) = 1$,
- $\deg(a) = \deg(c) = \deg(d) = 2$,
- $\deg(b) = 3$.

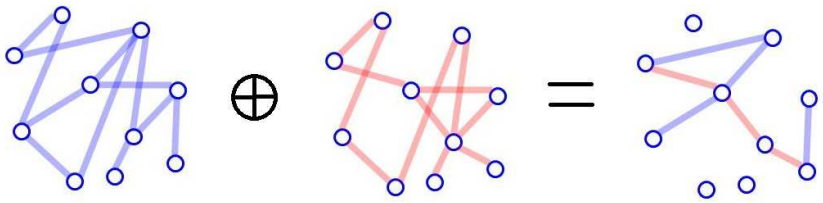
Przykładowa ścieżka w G to: a, b, c, przykładowa droga: $\{a, b\}, \{b, e\}$, przykładowy cykl — $\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}$. Graf G nie jest spójny. Jego spójnymi składowymi są podgrafy $G' = (V', E')$ oraz $G'' = (V'', E'')$, gdzie

$$V' = \{a, b, c, d, e\}, E' = \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}, \{b, e\}\}, \quad V'' = \{f, g\}, E'' = \{\{f, g\}\}.$$

Wynalazek grafu może nam się kojarzyć głównie z Eulerem (słusznie) i z problemem mostów w Królewcu (ponownie słusznie), ale z punktu widzenia nauki znaczenie tego pojęcia ukazał Kirchoff, zajmujący się badaniem obwodów i sieci elektrycznych. W tym ujęciu interesują nas zwykle grafy z krawędziami zorientowanymi (wskazującymi kierunek przepływu prądu) i z przypisanymi wagami (różne przewody mogą mieć różne właściwości jako przewodniki). Z naszego punktu widzenia ważne jest to, że przy konstruowaniu takiej sieci, napięcie powinno się zbalansować, czyli nie gromadzimy nadmiernego natężenia w żadnym punkcie. Krótko mówiąc: wzdłuż każdego cyklu tego obwodu suma spadków napięć powinna wynosić zero. Wydaje się więc, że chcąc zaplanować taką sieć trzeba sprawdzić każdy cykl w ilustrującym ją grafie G. Jak się jednak okazuje, nie jest to konieczne. Okazuje się, że graf ten spełnia drugie prawo Kirchoffa wtedy i tylko wtedy, gdy G spełnia to prawo na pewnym podzbiorze cykli, zwanym bazą przestrzeni cykli. Czym jest ta przestrzeń liniowa? Zacznijmy od przestrzeni krawędziowej.

Definicja 8.12: Przestrzeń krawędziowa grafu

Określamy przestrzeń liniową $P(E)$ nad ciałem $\mathbb{Z}_2$, zwaną PRZESTRZENIĄ KRAWĘDZIOWĄ grafu $G = (V(G), E(G))$, jak w przypadku przestrzeni liniowej podzbiorów zbioru niepustego. Zbiorem wektorów jest $P(E)$ — zbiór podzbiorów zbioru krawędzi $E(G)$ grafu G, a sumą $X \oplus Y$ dwóch podzbiorów X i Y należących do $E(G)$ jest różnica symetyrczna tych zbiorów $X \Delta Y = (X \cup Y) \setminus (X \cap Y)$. Mnożenie przez skalar jest zdefiniowane jako $1 \cdot X = X$ oraz $0 \cdot X = \emptyset$



Różnica symetryczna dwóch podzbiorów krawędzi grafu

Widzimy więc, że przestrzeń krawędziowa jest pewnym typem przestrzeni podzbiorów. Singletony zawierające pojedyncze wierzchołki grafu G tworzą bazę przestrzeni $P(E)$. Zauważmy, że operacja różnicy symetrycznej zachowuje się dobrze na innych wprowadzonych przez nas strukturach. Kluczowa obserwacja jest taka, że różnica symetryczna dwóch cykli jest cyklem. W ten sposób dochodzimy do definicji.

Definicja 8.13: Przestrzeń cykli grafu

Niech $G = (V(G), E(G))$ będzie grafem. Podprzestrzeń $C(E)$ przestrzeni $P(E)$ rozpiętą przez wszystkie cykle nazywamy przestrzenią cykli grafu G. Wymiar $C(E)$ nazywamy LICZBĄ CYKLOMATYCZNĄ grafu G.

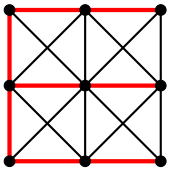
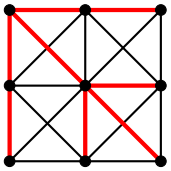
O znaczeniu przestrzeni cykli świadczy podstawowa własność szczególnych i najstarszych ich typów — cykli Eulerowskich. Jak się okazuje, różnica symetryczna dwóch takich cykli również jest cyklem Eulerowskim, czyli przechodzącym przez każdą krawędź dokładnie raz. Jak wiadomo, graf F ma cykl Eulerowski, jeśli każdy jego wierzchołek ma parzysty stopień. Różnica symetryczna dwóch grafów, których wierzchołki mają parzyste stopnie oczywiście ma również tę własność.

Skoro dysponujemy przestrzenią cykli można się zastanawiać nad jej bazą. Jest to zagadnienie wymagające odrobinę wiedzy z teorii grafów. Do podstawowych pojęć należy drzewo (las) rozpinające oraz fundamentalne cykle.

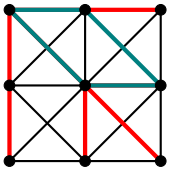
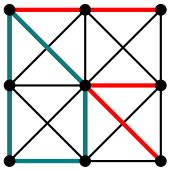
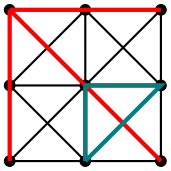
Definicja 8.14

Podgraf, który zawiera wszystkie wierzchołki grafu spójnego G i jest drzewem, nazywamy DRZEWEM ROZPINAJĄCYM graf G. Jeśli G nie jest spójny, wówczas dowolną sumę rozłączną grafów rozpinających spójne składowe tego grafu, nazywamy LASEM ROZPINAJĄCYM G.

Oto przykłady dwóch drzew rozpinających grafu o 9 wierzchołkach (na czerwono). Możemy też przyjąć, że suma rozłączna tych dwóch drzew stanowi las rozpinający graf o 18 wierzchołkach i dwóch składowych.

Niech L będzie lasem rozpinającym grafu G. Wówczas dodanie dowolnej krawędzi e z G nie należącej do L utworzy dokładnie jeden cykl, zwany CYKLEM FUNDAMENTALNYM C_e związanym z lasem rozpinającym L. Cykl C_e jest wyznaczony jednoznacznie, ponieważ biorąc końce x, y krawędzi e wiemy, że w lesie L jest dokładnie jedna droga między x, a y. Oto przykłady cykli fundamentalnych dla lasu L na prawym grafie.

Nietrudno widzieć, że jeśli L jest lasem rozpinającym grafu G, to każdy cykl w G ma wspólną krawędź z dopełnieniem L. Gdyby bowiem cykl nie miał wspólnej krawędzi z dopełnieniem L, to byłby zawarty w L, co przeczy acykliczności L. Okazuje się, że zachodzi następujące twierdzenie.

Twierdzenie 8.3

Zbiór cykli fundamentalnych dowolnego lasu rozpinającego L graf G stanowi bazę podprzestrzeni cykli $C(E)$. W szczególności $\dim C(E) = m + n - c$, gdzie m jest liczbą krawędzi w grafie G, n jest liczbą wierzchołków, a c jest liczbą składowych spójnych.

Dowód. Oznaczmy przez C_e cykl fundamentalny powstały przez dopisanie krawędzi e do lasu L. Rozważmy układ

$$\{C_e, e \in E(G) \setminus E(L)\}$$

Po pierwsze zauważmy, że układ ten jest liniowo niezależny, ponieważ graf $C_{e_1} \Delta C_{e_2} \Delta \ldots \Delta C_{e_k}$ zawiera krawędzie $e_1, \ldots, e_k$. Innymi słowy — C_e jest jedynym elementem tego układu zawierającym e.

Z drugiej strony układ ten rozpina przestrzeń cykli. Istotnie, jeśli $H \in C(E)$, to bierzemy wszystkie krawędzie $e_1, \ldots, r_k$ z $V(H)$, które nie są w $V(L)$ i rozważamy $H \Delta C_{e_1} \Delta C_{e_2} \Delta \ldots \Delta C_{e_k}$. Rezultat dalej jest w przestrzeni cykli, i jest to podgraf L, ponieważ każda krawędź $e \notin E(L)$ została usunięta przez różnicę symetryczną z C_e. Ale element przestrzeni cykli nie może być podgrafem drzewa, o ile nie jest pusty. Stąd $H = C_{e_1} \Delta C_{e_2} \Delta \ldots \Delta C_{e_k}$. □

Czytelnik może się zastanawiać po co nam taka osobliwa przestrzeń? Istnieje sporo powodów, które należą do bardziej zaawansowanej matematyki — jak choćby zastosowania w teorii grup homologii kompleksów symplicjalnych w topologii algebraicznej, czy już wspomniane prawo Kirchoffa. Jest również niezwykłe zastosowanie w samej teorii grafów, które wspominamy bez dowodu.

Przypomnijmy, że GRAFEM PLANARNYM nazywamy graf, który można narysować na płaszczyźnie w taki sposób, by żadne dwie krawędzie się nie przecinały. Klasyczne kryterium planarności pochodzi od polskiego matematyka Kazimierza Kuratowskiego i wskazuje podgrafy, których graf planarny zawierać nie może. Poniższe kryterium dotyczy natomiast samej przestrzeni cykli.

Twierdzenie 8.4: Mac Lane 1937, O'Neil 1973

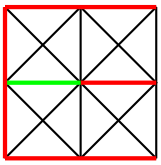
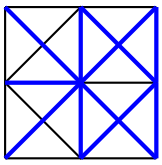
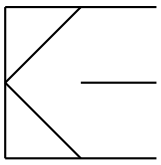
Skończony niezorientowany graf G jest planarny wtedy i tylko wtedy, gdy każda krawędź grafu G jest składnikiem dokładnie dwóch wektorów bazy $C(E)$, czyli dwóch fundamentalnych cykli.

Powiedzieliśmy więc coś o podprzestrzeni cykli, ale skoro tematem wykładu była suma prosta, to pokażemy naturalne dopełnienie proste przestrzeni cykli, czyli tzw. przestrzeń rozcięć grafu.

Definicja 8.15: Przestrzeń rozcięć grafu

Niech $G = (V, E)$. Podprzestrzeń $R(E)$ przestrzeni $P(E)$ rozpiętą przez wszystkie rozcięcia, czyli podzbiory $P(E)$, których usunięcie rozspaja graf, nazywamy PRZESTRZENIĄ ROZCIĘĆ grafu G.

Jeśli z lasu L usuniemy dowolną krawędź, to w (odpowiadającej jej spójnej składowej) powstają dwa rozłączne zbiory wierzchołków V_1, V_2. Zbiór wszystkich krawędzi G takich, że koniec jest w V_1, a drugi w V_2 tworzy rozcięcie, które nazywamy ROZCIĘCIEM FUNDAMENTALNYM związanym z lasem L. Oto przykłady (na niebiesko) rozcięć fundamentalnych, dla jednego z drzew wskazanych wyżej, z którego usunięto krawędź (zieloną. Spójne grafy po usunięciu rozcięcia fundamentalnego są po prawej.

To, co jest ciekawe w tym dodatkowym obiekcie to fakt, że zachodzi równość $P(E) = C(E) \oplus R(E)$. Biorąc pod uwagę konstrukcję rozcięć fundamentalnych widzimy, że wymiar $R(E)$ równy jest liczbie krawędzi w dowolnym lesie rozpinającym. Istnienie powyższej sumy prostej dowodzi się na różne sposoby, ale najbardziej popularny korzysta z następującego rezultatu.

Obserwacja 8.7

Każdy cykl i rozcięcie w grafie G mają parzystą liczbę wspólnych krawędzi.

Dowód. Rozważmy rozcięcie S w spójnym grafie G (to oczywiście wystarczy do dowodu). Załóżmy, że usunięcie S ze zbioru krawędzi G rozbija zbiór wierzchołków na dwie rozłączne podzbiory V_1 i V_2. Niech C będzie cyklem w G. Jeśli wszystkie wierzchołki C leżą w jednym ze zbiorów V_1 lub V_2, to wszystkie krawędzie C są różne od krawędzi w S, co oznacza, że w tym przypadku cykl S i rozcięcie S mają 0 wspólnych krawędzi — czyli liczbę parzystą.

Jeśli pewne wierzchołki C są w V_1, a pewne w V_2, to przechodząc cykl przechodzimy między zbiorami V_1 i V_2. Skoro cykl jest drogą zamkniętą, liczba krawędzi przejścia między V_1 i V_2 musi być parzysta. Każdemu przejściu z jednego zbioru do drugiego odpowiadać musi przejście z powrotem. □

Rozumowanie pokazujące, że przestrzeń krawędzi grafu niezorientowanego jest sumą prostą przestrzeni cykli i krawędzi można przeprowadzić na wiele sposobów, ale wskażemy uniwersalna drogę — ważną w całej algebrze liniowej. Już w jednym z poprzednich wykładów mówiliśmy o naiwnym ujęciu prostopadłości, pochodzącym od swego rodzaju uogólnienia iloczynu skalarnego. To uogólnienie może pójść bardzo daleko, o czym świadczy poniższa definicja.

Definicja 8.16: Forma dwuliniowa dla grafów i podzbiorów

Niech $P(E)$ będzie przestrzenią krawędzi grafu G i weźmy wektory $v_1 = a_1e_1 + \ldots + a_me_m$ oraz $v_2 = b_1e'_1 + \ldots + b_me'_m$, gdzie a_i, b_j należą do $\mathbb{Z}_2$, oraz $e_i, e'_j \in E$. Określamy:

$$\langle v_1, v_2 \rangle = a_1b_1 + a_2b_2 + \ldots + a_mb_m.$$

Warunek $\langle v_1, v_2 \rangle = 0$ spełniony jest zawsze, gdy zbiory v_1, v_2 mają parzystą liczbę wspólnych krawędzi. Z twierdzenia powyżej nietrudno wywnioskować, że w istocie przestrzeń cykli jest „prostopadła" do przestrzeni rozcięć grafu. Dokładniej, dla dowolnej podprzestrzeni $W \subset P(E)$ określić można

$$W^{\perp} = \{v \in P(E) : \langle v, w \rangle = 0, \forall_{w \in W}\}.$$

Czytelnikowi pozostawiam pokazanie, że w rozważanym przypadku mamy $C(E)^{\perp} = R(E)$ oraz, że wynika stąd równość $P(E) = C(E) \oplus R(E)$. Z pewnością w przyszłym semestrze będą Państwo mieli więcej narzędzi do uzasadniania takich wyników.

8.6 Coda. Jednoznaczność daje wyniki o nieistnieniu

Na wykładzie po raz kolejny pojawiło się pojęcie, które w swojej naturze zawiera koncepcję jednoznaczności przedstawienia pewnego elementu za pomocą innych. Przestrzeń liniowa V jest sumą prostą podprzestrzeni V_1 oraz V_2, jeśli każdy wektor v z przestrzeni V można zapisać jednoznacznie właśnie za pomocą sumy wektorów $v_1 \in V_1$ oraz $v_2 \in V_2$. O sile tej jednoznaczności przekonamy się wielokrotnie, nie tylko w kontekście sumy prostej dwóch, ale i większej liczby podprzestrzeni. Jaka koncepcja matematyczna stoi za tym pojęciem? Byla ona już wspominana — jednoznaczność oznacza, że pewne konfiguracje obiektów matematycznych nie są możliwe do uzyskania. Tego typu argument ma centralne znaczenie w matematyce.

Z jednoznaczności rozkładu liczby całkowitej na czynniki pierwsze wynika, że nie istnieje para liczb całkowitych m, n, spełniających równość $2m^2 = n^2$. Taka para nie istnieje, ponieważ liczba całkowita ma jeden rozkład na czynniki pierwsze. Równość $2m^2 = n^2$ oznacza, że liczba 2 wchodzi do rozkładu liczby całkowitej $2m^2$ nieparzyście wiele razy, a do rozkładu liczby n^2 — parzyście wiele razy. Jednoznaczność tego zabrania. Po obydwu stronach stać muszą liczby, które dzielą się przez tą samą potęgę liczby 2.

Tego typu dowodów można formułować więcej. Wspominaliśmy o jednoznaczności rozkładu na czynniki wielomianów o współczynnikach w ciele. I ponownie można z tego wyciągać rozmaite wnioski — choćby taki (stosunkowo przyziemny), że funkcja rzeczywista $\sqrt{x}$ nie jest wymierna — nie można jej przedstawić jako ilorazu rzeczywistych funkcji wielomianowych. Czy Czytelnik umie to ściśle uzasadnić? Dlaczego nie zakładamy, że ciało jest dowolne, skoro twierdzenie o jednoznaczności jest prawdziwe nad dowolnym ciałem? Utożsamienie wielomianów z funkcjami wielomianowymi nie jest prawdziwe dla ciał skończonych. Funkcja $\sqrt{x} : \mathbb{Z}_2 \to \mathbb{Z}_2$ jest identycznością, czyli funkcją wielomianową $\sqrt{x} = x$.

W tym dodatku opowiemy o bardzo słynnym zagadnieniu algebraicznym, którego elementarne rozwiązanie dostarcza pojęcie sumy prostej. Chodzi o rozstrzygnięcie, czy w przestrzeni $\mathbb{R}^3$ można wprowadzić strukturę mnożenia wektorów w taki sposób, aby uzyskać strukturę ciała nad $\mathbb{R}$, podobnie jak liczby zespolone $\mathbb{C}$ traktować można jako przestrzeń $\mathbb{R}^2$ z odpowiednio zdefiniowanym dodawaniem i mnożeniem. Pytanie to miało duże znaczenie w wieku XIX-tym, zwłaszcza dla Rowana Hamiltona, który to właśnie dostrzegł strukturę liczb zespolonych jako par liczb rzeczywistych z określonym działaniem.

O co chodzi? Dobre opisuje to tekst prof. Zbigniewa Marciniaka w Delcie pt. *Dlaczego w przestrzeni trójwymiarowej nie ma przyzwoitego mnożenia?* (`https://www.deltami.edu.pl/media/articles/1996/04/delta-1996-04-dlaczego-w-przestrzeni-trojwymiarowej-nie-ma-przyzwoitego_yKYRwmA.pdf`). Profesor w różnych miejscach w Delcie podaje dwa różne dowody. Ja chciałbym jednak przywołać taki, który wydaje mi się jeszcze bardziej naturalny i korzysta z pojęcia sumy prostej.

Przypomnijmy o co chodzi. O każdej liczbie zespolonej $a + bi \in \mathbb{C}$ możemy myśleć jako o parze liczb rzeczywistych (a, b) z działaniami dodawania i mnożenia:

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Dodawanie nietrudno uogólnić na przestrzeń trójwymiarową nad $\mathbb{R}$, czyli:

$$(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Pytanie brzmi: czy $\mathbb{R}^3$ można wyposażyć w działanie mnożenia, które z tej przestrzeni robiłoby ciało? Może nawet bylibyśmy gotowi zrezygnować z przemienności mnożenia, jak w przypadku kwaternionów, o których można myśleć jak o elementach $\mathbb{R}^4$ z odpowiednim dodawaniem i mnożeniem. Czytelnik zechce sprawdzić, że zwykłe mnożenie po współrzędnych nie jest dobre, ponieważ w żadnym ciele iloczyn elementów niezerowych nie może być zerowy, a tu mielibyśmy $(1, 0, 0) \cdot (0, 1, 0) = (0, 0, 0)$.

Zakładać będziemy, że D jest ciałem, które jest jednocześnie n-wymiarową przestrzenią liniową nad $\mathbb{R}$. Dla $\lambda \in \mathbb{R}$ oraz $1 \in \mathbb{D}$ pisać będziemy po prostu $\lambda \cdot 1 = \lambda \in D$. Przytoczone rozumowanie pochodzi z książki Mateja Bresara *Introduction to Noncommutative Algebra*.

Obserwacja 8.8

Dla każdego $s \in D$ istnieje $\lambda \in \mathbb{R}$, że $s^2 + \lambda s \in \mathbb{R}$.

Zanim zobaczymy dowód zauważmy, że stwierdzenie to oznacza, że każdy element D spełnia równanie kwadratowe o współczynnikach w $\mathbb{R}$.

Dowód. Zauważmy, że jeśli $\dim D = n$, to układ $n+1$ elementów $1, s, \ldots, s^n$ jest liniowo zależny. Oznacza to, że istnieje wielomian $f(x) \in \mathbb{R}[x]$ stopnia co najwyżej n, że $f(s) = 0$. Załóżmy, że wiodący współczynnik f równy jest 1. Wiemy z wykładu, że $f(x)$ rozkłada się na iloczyn czynników liniowych i kwadratowych w $\mathbb{R}[x]$:

$$f(x) = (x-\lambda_1)\ldots(x-\lambda_r)(x^2+a_1x+b_1)\ldots(x^2+a_sx+b_s),$$

gdzie $\lambda_i, a_i, b_i \in \mathbb{R}$. Skoro $f(s) = 0$, to

$$(x-\alpha_1)\ldots(x-\alpha_r)(x^2+a_1s+b_1)\ldots(x^2+a_sx+b_s) = 0.$$

Skoro D jest ciałem (nawet nieprzemiennym), to jeden z czynników musi być równy 0, a zatem s jest pierwiastkiem wielomianu kwadratowego o współczynnikach w $\mathbb{R}$. □

Czytelnika może boleć, że nagle do gry wkracza wynika wymagający zasadniczego twierdzenia algebry. To prawda, ponieważ chcemy coś wykazać dla przestrzeni dowolnego wymiaru n. Gdyby ograniczyć się do $n = 3$, wówczas trzeba tylko wiedzieć, że każdy wielomian rzeczywisty stopnia 3 ma pierwiastek, a tu ZTA nie potrzeba. Przechodzimy do kluczowego argumentu.

Obserwacja 8.9

Rozważmy podzbiór ciała D postaci:

$$V = \{v \in D : v^2 \in \mathbb{R}, v^2 \leqslant 0\}.$$

Wówczas V jest podprzestrzenią D (nad $\mathbb{R}$) oraz

$$D = \mathbb{R} \oplus V.$$

Ów rozkład będzie miał fundamentalne znaczenie dla dalszej klasyfikacji. Naszym celem będzie później pokazanie, że wymiar przestrzeni V wynosić może jedynie 0, 1 lub 3.

Dowód. Zacznijmy od następującej obserwacji. Jeśli $s \in D \setminus V$ spełnia jednocześnie $s^2 \in \mathbb{R}$, to $s^2 > 0$, a stąd $s^2 = \lambda^2$, dla pewnego $\lambda \in \mathbb{R}$. Stąd $(s-\lambda)(s+\lambda) = 0$, skąd $s = \pm\lambda \in \mathbb{R}$.

Jest jasne, że $\mathbb{R} \cap V = \{0\}$ oraz, że V jest zamknięty na mnożenie przez liczby rzeczywiste tzn. jeśli $s \in D$ oraz $\lambda \in \mathbb{R}$, wówczas $\lambda s \in V$. Sprawdzimy teraz, że jeśli $u, v \in V$, to $u+v \in V$. Istotnie, możemy założyć, że u, v są liniowo niezależne (inaczej to oczywiste). Twierdzimy, że wówczas układ $1, u, v$ również jest liniowo niezależny. Rzeczywiście, jeśli dla pewnych $a, b, c \in \mathbb{R}$ mamy $au = bv + c$, to podniesieniu obydwu stron do kwadrat dostajemy $bcv \in \mathbb{R}$, skąd $b = 0$ lub $c = 0$, skąd $a = b = c = 0$. Jak wykazać teraz, że $u + v \in V$? Skoro $u + v \in D$, to istnieją $a, b \in \mathbb{R}$, że:

$$(u+v)^2 + a(u+v) \in \mathbb{R}, \quad (u-v)^2 + b(u-v) \in \mathbb{R}.$$

Z drugiej strony:

$$(u+v)^2 + (u-v)^2 = 2u^2 + 2v^2 \in \mathbb{R}.$$

Porównując te równości, uzyskujemy $a(u+v) + b(u-v) \in \mathbb{R}$. Skoro jednak $u, v, 1$ są liniowo niezależne, to $a+b = a-b = 0$, skąd $a = b = 0$, więc $u+v \in V$. Stąd V jest podprzestrzenią D.

Pozostało pokazać, że $D = \mathbb{R}+V$. Weźmy $s \in D \setminus \mathbb{R}$. Zgodnie z poprzednią obserwacją mamy $s^2 + \lambda s \in \mathbb{R}$, dla pewnego $\lambda \in \mathbb{R}$. Ponownie używając argumentu z pierwszego akapitu mamy $s + \frac{\lambda}{2} \in V$. Stąd

$$s = -\frac{\lambda}{2} + (s + \frac{\lambda}{2}) \in \mathbb{R} + V.$$

□

W tym momencie możemy do końca nie widzieć jakie zalety ma wydzielenie jednowymiarowego składnika prostego w naszym ciele D. Wnioskować chcemy, że jeśli D ma wymiar powyżej 2, to ma wymiar co najmniej 4 (a dokładniej – zawiera kwaterniony), a więc nie jest wymiaru 3. Oto dokładne sformułowanie.

Obserwacja 8.10

Jeśli $\dim D > 2$, to istnieje liniowo niezależny układ wektorów $i, j, k \in V$, że

$$i^2 = j^2 = k^2 = -1, \quad ij = -ij = k, \quad ki = -ik = j, \quad jk = -kj = i.$$

Jak się okazuje, tu wkracza niebanalny pomysł. Rozważamy nowe działanie, dla dowolnych elementów $u, v \in V$ definiujemy

$$u \circ v = uv + vu.$$

Tu jest podkreślony potencjalnie nieprzemienny charakter D. Zauważmy, że :

$$u \circ v = (u+v)^2 - u^2 - v^2 \in \mathbb{R}, \text{ a jeśli } v \neq 0, \text{ to } v \circ v = 2v^2 \neq 0.$$

Dowód. Uzasadnienie poprzedniego faktu wraz z formułą Grassmanna daje nam, że $\dim V = \dim D - 1 = n - 1 > 1$. Możemy zatem wybrać dwa liniowo niezależne wektory $u, v \in V$. Niech

$$u := w - \frac{w \circ v}{v \circ v} v.$$

Łatwo sprawdzić, że $u \neq 0$ oraz $u \circ v = 0$. Niech:

$$i := \frac{1}{\sqrt{-u^2}} u, \quad j := \frac{1}{\sqrt{-v^2}} v, \quad k = ij.$$

Bezpośredni rachunek pozwala sprawdzić, że zachodzą równości postulowane w tezie. Co więcej, dla dowolnych $a, b, c, d \in \mathbb{R}$ mamy:

$$(a + bi + cj + dk)(a - bi - cj - dk) = a^2 + b^2 + c^2 + d^2,$$

co oznacza, że $1, i, j, k$ są liniowo niezależne. □

Zostawiłem Czytelnikowi sprawdzenie dokładnych rachunków, ale może wypada powiedzieć, że w drugim semestrze wektory typu $w - \frac{w \circ v}{v \circ v} v$ będą rzutami prostopadłymi wektora w na przestrzeń prostopadłą do $\mathrm{lin}(v)$. W tym sensie zamiast $\circ$ stać będzie iloczyn skalarny. Krótko mówiąc wykazaliśmy, że jeśli dany jest układ liniowo niezależny, który zawiera wektory z V, to ma on co najmniej 3 elementy. To już nam pokazuje, że w $D = \mathbb{R}^3$ nie ma struktury ciała z kompatybilnym z dodawaniem po współrzędnych mnożeniu. Teraz wykażemy, że jeśli D ma mieć takie *dobre mnożenie* (niekonieczne przemienne), to $\dim D = 0, 1, 3$. W istocie, D „jest” (z dokładnoscią do *izomorfizmu algebr*) jedną z *algebr*: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$.

Twierdzenie 8.5: Frobenius (1878)

Przestrzeń liniowa D skończonego wymiaru nad $\mathbb{R}$ z mnożeniem, które spełnia wraz z dodawaniem wektorów aksjomaty ciała, ewentualnie poza przemiennością (tzw. *algebra z dzieleniem* nad $\mathbb{R}$), jest wymiaru 1, 2 lub 4.

Dowód. Niech $\dim D = n > 4$. Niech i, j, k będą elementami uzyskanymi w poprzednim twierdzeniu. Skoro $\dim V > n - 1 > 3$, to istnieje $v \notin \mathrm{lin}(i, j, k)$. Rozważmy element:

$$e := v + \frac{i \circ v}{2} i + \frac{j \circ v}{2} j + \frac{k \circ v}{2} k.$$

Zauważmy, że jest to element niezerowy (bo inaczej v jest kombinacją liniową i, j, k) oraz mamy:

$$i \circ e = j \circ e = k \circ e = 0.$$

Z pierwszych dwóch równości mamy zatem $i \circ e = ie + ei = 0 = j \circ e = je + ej$, czyli:

$$ie = -ei, \quad je = -ej \Rightarrow -iej = eij, \quad \text{oraz} \quad je = -ej \Rightarrow ije = -iej.$$

Stąd $eij = -iej = ije$. Natomiast z równości $k \circ e = 0$ wynika, że $ke = -ek$, czyli wstawiając ij =k mamy $ije = -eij$. To jednaj oznacza, że $eij = 0$, co nie jest możliwe. Sprzeczność z założeniem, że $n > 4$. □

Dowód był dość subtelny, ale wynik jest wysoce niebanalny. Widzimy jak istotne znaczenie miało wydzielenie składnika prostego V. Tego typu myślenie jest charakterystyczne dla zaawansowanej algebry – wydzielić „duży” fragment, który lepiej rozumiemy i w nim prowadzić właściwe rozumowanie.

Rozdział 9

Przekształcenia liniowe. Przestrzenie izomorficzne

9.1 Wykład dziewiąty

Ostatni wykład pokazał nam jak rzeczywistość algebraiczną można oglądać w rzeczywistości geometrycznej, i odwrotnie. Algebraicznemu układowi równań liniowych przypisaliśmy geometrycznie opisywalny zbiór rozwiązań, ale i odwrotnie – pokazaliśmy, że dla każdego zbioru możemy wskazać odpowiedni układ równań jednorodnych. Tego typu ODPOWIEDNIOŚCI są kluczowe w naszym spojrzeniu. Drugą, po układach równań – znacznie ogólniejszą klasą obiektów algebraicznych prowadzącą do rozważania konfiguracji geometrycznych, będą przekształcenia liniowe, którymi zaczniemy się dziś zajmować.

Definicja 9.1: Przekształcenie liniowe

Niech V, W będą przestrzeniami liniowymi nad ciałem K. Funkcję $\phi : V \to W$ nazwiemy PRZEKSZTAŁCENIEM LINIOWYM, jeśli dla dowolnych $\alpha, \beta \in V$ oraz dla każdego $a \in K$ zachodzi:

(i) $\phi(\alpha + \beta) = \phi(\alpha) + \phi(\beta)$,

(ii) $\phi(a \cdot \alpha) = a \cdot \phi(\alpha)$. (W szczególności: $\phi(0_V) = 0_W$, czyli zero przechodzi w zero.)

Zwróćmy uwagę na to, że w powyższych warunkach z lewej strony mamy do czynienia z dodawaniem i mnożeniem w przestrzeni V, a po prawej – z dodawaniem i mnożeniem w przestrzeni W.

Kilka przykładów przekształceń liniowych (zachęcam do sprawdzenia)

(a) $\phi : \mathbb{R}^2 \to \mathbb{R}^3$ postaci $\phi((a, b)) = \left(\frac{a+b}{2}, \frac{a+b}{2}, a\right)$,

(b) $\phi : F(\mathbb{R}, \mathbb{R}) \to F(\mathbb{R}, \mathbb{R})$ przyporządkowujące funkcji f funkcję parzystą $\phi(f)$ daną wzorem

$$(\phi(f))(x) = \frac{f(x) + f(-x)}{2},$$

(c) OBROT o kąt θ, czyli $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ dany wzorem $\phi((x_1, x_2)) = (x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta)$.

(d) Odwzorowanie $d : K[x] \to K[x]$, zwane POCHODNĄ, zadane wzorem

$$\phi(a_0 + a_1 x + \ldots + a_n x^n) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}.$$

(e) Odwzorowanie $tr : M_{n\times n}(K) \to K$ przypisujące macierzy $A = (a_{ij})$ sumę elementów na przekątnej $a_{11} + \ldots + a_{nn}$, zwane ŚLADEM.

(f) Przekształcenie $\phi : K^\infty \to K^\infty$ dane wzorem $\phi(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$.

(g) Niech $(\mathbb{R}_+, \boxplus, \boxtimes, 1)$ będzie przestrzenią liniową nad $\mathbb{R}$, gdzie $x \boxplus y = xy$ oraz $a \boxtimes x = x^a$, dla $a \in \mathbb{R}$. Przekształcenie $l : \mathbb{R}_+ \to \mathbb{R}$ dane wzorem $l(x) = \ln(x)$ jest liniowe.

Definicja 9.2: Ważne klasy przekształceń liniowych

- Przekształcenie $\phi : V \to W$ nazwiemy ZEROWYM, jeśli dla każdego $\alpha \in V$ mamy $\phi(\alpha) = 0$.
- Przekształcenie $\phi : V \to W$ nazwiemy IDENTYCZNOŚCIĄ, jeśli dla każdego $\alpha \in V$ mamy $\phi(\alpha) = \alpha$.
- Przekształcenie liniowe $f : V \to V$ przestrzeni liniowej w siebie dane wzorem $f(\alpha) = a\alpha$ nazywamy HOMOTETIĄ (albo jednokładnością) o skali a.
- Jeśli $V = V_1 \oplus V_2$, to dla każdego $\alpha \in V$ istnieją jednoznacznie wyznaczone $\alpha_1 \in V_1$ oraz $\alpha_2 \in V_2$, że $\alpha = \alpha_1 + \alpha_2$. Definiujemy:
  - RZUT $\phi : V \to V$ przestrzeni V na V_1 wzdłuż V_2 dany wzorem $\phi(\alpha) = \alpha_1$.
  - SYMETRIĘ $\psi : V \to V$ przestrzeni V względem V_1 wzdłuż V_2 daną wzorem $\psi(\alpha) = \alpha_1 - \alpha_2$.

Przekształcenia liniowe to jedyne funkcje pomiędzy przestrzeniami liniowymi, które ZACHOWUJĄ KOMBINACJE LINIOWE.

Obserwacja 9.1

Następujące warunki są równoważne:

(i) $\phi : V \to W$ jest przekształceniem liniowym,

(ii) dla każdych $a_1, \ldots, a_k \in K$ oraz $\alpha_1, \ldots, \alpha_k \in V$ zachodzi

$$\phi(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_k\alpha_k) = a_1\phi(\alpha_1) + a_2\phi(\alpha_2) + \ldots + a_k\phi(\alpha_k).$$

Dowód. Indukcja ze względu na k. Dla $k = 2$ teza wynika z definicji przekształcenia liniowego. Najpierw korzystamy z tego, że ϕ zachowuje dodawanie, a potem mnożenie przez skalar.

$$\phi(a_1\alpha_1 + a_2\alpha_2) = \phi(a_1\alpha_1) + \phi(a_2\alpha_2) = a_1\phi(\alpha_1) + a_2\phi(\alpha_2).$$

Niech $k > 2$. Z definicji przekształcenia liniowego mamy:

$$\phi(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_k\alpha_k) = \phi(a_1\alpha_1 + a_2\alpha_2 + \ldots + a_{k-1}\alpha_{k-1}) + a_k\phi(\alpha_k).$$

Korzystając z założenia indukcyjnego dostajemy implikację (i) $\Rightarrow$ (ii). Odwrotna implikacja jest oczywista. □

Wniosek 9.1

Jeśli $\phi : V \to W$ jest przekształceniem liniowym. Wówczas:

- jeśli A jest podprzestrzenią V, to $\phi(A)$ jest podprzestrzenią W,
- jeśli B jest podprzestrzenią W, to $\phi^{-1}(B)$ jest podprzestrzenią V.

Dowód. Pokażemy, że $\phi(A)$ jest podprzestrzenią W. Jeżeli $\beta_1, \beta_2 \in \phi(A)$, to istnieją $\alpha_1, \alpha_2 \in A$ takie, że $\beta_1 = \phi(\alpha_1)$ oraz $\beta_2 = \phi(\alpha_2)$. Skoro $\alpha_1 + \alpha_2 \in A$, to $\phi(\alpha_1 + \alpha_2) \in \phi(A)$. A zatem z definicji przekształcenia liniowego mamy $\phi(\alpha_1) + \phi(\alpha_2) \in \phi(A)$. Podobnie pokazujemy, że jeśli $\lambda \in K$ oraz $\beta = \phi(\alpha)$, dla pewnego $\alpha \in A$, to oczywiście $\lambda\alpha \in A$, czyli $\phi(\lambda\alpha) = \lambda\phi(\alpha) \in \phi(A)$. A zatem $\phi(A)$ jest podprzestrzenią W. Analogicznie pokazujemy, że $\phi^{-1}(B)$ jest podprzestrzenią V. □

Szczególnie istotna jest sytuacja, gdy mowa jest o obrazie całej przestrzeni V zawartym w V oraz o przeciwobrazie przestrzeni zerowej zawartym w V. Zgodnie z powyższym wnioskiem są to podprzestrzenie.

Definicja 9.3: Jądro i obraz przekształcenia liniowego

Niech $\phi : V \to W$ będzie przekształceniem liniowym.

- Jądrem przekształcenia ϕ nazywamy zbiór $\ker(\phi) = \{\alpha \in V \,|\, \phi(\alpha) = 0\} \subseteq V$.
- Obrazem przekształcenia ϕ nazywamy zbiór $\mathrm{im}(\phi) = \{\phi(\alpha) \,|\, \alpha \in V\} \subseteq W$.

Oczywiście jądro i obraz są podprzestrzeniami, odpowiednio $\ker(\phi) = \phi^{-1}(\{0\})$ oraz $\mathrm{im}(\phi) = \phi(V)$. Zobaczmy kilka przykładów.

- Jeśli $\phi : \mathbb{R}^3 \to \mathbb{R}$ dane jest wzorem
$$\phi((x_1, x_2, x_3)) = x_1 + x_2,$$
to $\ker(\phi) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = 0\}$.
- Jeśli $\psi : \mathbb{R} \to \mathbb{R}^3$ dane jest wzorem
$$\phi(x) = (x, x, x),$$
to $\mathrm{im}(\phi) = \mathrm{lin}((1, 1, 1))$.
- Niech $\phi : V \to V$ będzie homotetią, przy czym $\phi(v) = av$, dla każdego $v \in V$ oraz pewnego ustalonego $a \in K$. Wówczas:
$$\ker(\phi) = \begin{cases} \{0\}, & a \neq 0 \\ V, & a = 0 \end{cases}, \quad \mathrm{im}(\phi) = \begin{cases} V, & a \neq 0 \\ \{0\}, & a = 0 \end{cases}.$$
- Niech $\phi : V \to V$ będzie rzutem na V_1 wzdłuż V_2. Wówczas:
$$\ker(\phi) = V_2, \quad \mathrm{im}(\phi) = V_1.$$
- Niech $\phi : \mathbb{R}[x] \to \mathbb{R}[x]$ jest pochodną, to:
$$\ker(\phi) = \{w \in \mathbb{R}[x] \,:\, \deg(w) \leqslant 0\}, \quad \mathrm{im}(\phi) = \mathbb{R}[x].$$

Kluczowy przykład

Niech $a_{ij} \in K$, gdzie $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$. Niech $\phi : K^n \to K^m$ będzie przekształceniem liniowym postaci:
$$f(x_1, \ldots, x_n) = (a_{11}x_1 + \ldots + a_{1n}x_n, \ldots, a_{m1}x_1 + \ldots + a_{mn}x_n).$$
Wówczas $\ker(f) \subseteq K^n$ jest zbiorem rozwiązań jednorodnego układu równań postaci:
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ \qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases}$$

A przestrzeń $\mathrm{im}(f) \subseteq K^m$? Jest to w istocie przestrzeń kolumnowa macierzy A. Mamy:
$$\mathrm{lin}(f(\epsilon_1), \ldots, f(\epsilon_n)) = \mathrm{lin}((a_{11}, \ldots, a_{m1}), \ldots, (a_{1n}, \ldots, a_{mn})).$$

Rzeczywiście, $\mathrm{im}(f)$ jest rozpięta przez obrazy wektorów bazy standardowej przestrzeni K^n. co jest przykładem ogólnej prawidłowości, wynikającej natychmiast z poprzedniego wniosku.

Wniosek 9.2: Obraz jest rozpięty przez obrazy układu rozpinającego

Niech $V = \mathrm{lin}(\alpha_1, \ldots, \alpha_n)$ oraz niech $f : V \to W$ będzie przekształceniem liniowym. Wówczas
$$\mathrm{im}(\phi) = \mathrm{lin}(\phi(\alpha_1), \ldots, \phi(\alpha_n)).$$

Zauważmy, że z tw. Kroneckera-Capellego wynika ważny wniosek dotyczący dowolnego przekształcenia liniowego $f : K^n \to K^m$.

$$n = \dim(K^n) = \dim \ker(f) + \dim \operatorname{im}(f).$$

Powyższa obserwacja dotycząca wymiaru obrazu i jądra przekształcenia zadanego wzorem (†) uogólnia się na dowolne przekształcenia liniowe przestrzeni skończonego wymiaru (nie jest tak naprawdę uogólnienie, ale to zrozumiemy później). Odnotujmy, że we wzorze powyżej $\dim \operatorname{im}(f)$ równe jest rzędowi macierzy rozważanego układu. Stąd bierze nazwę poniższa definicja.

Definicja 9.4: Rząd przekształcenia liniowego

Wymiar przestrzeni $\operatorname{im}(\phi)$ nazywamy RZĘDEM PRZEKSZTAŁCENIA, ozn. $r(\phi)$.

Przejdziemy teraz do twierdzenia będącego uogólnieniem twierdzenia Kroneckera-Capellego.

Twierdzenie 9.1: Wymiary obrazu i jądra przekształcenia liniowego

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Wówczas:

$$\dim V = \dim \ker(\phi) + \dim \operatorname{im}(\phi).$$

Dowód. Pokażemy najpierw następujący istotny fakt.

Obserwacja 9.2

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Niech U będzie taką podprzestrzenią przestrzeni V, że $V = \ker(\phi) \oplus U$. Niech $\alpha_1, \ldots, \alpha_k$ będzie bazą przestrzeni U. Wówczas układ $\phi(\alpha_1), \ldots, \phi(\alpha_k)$ jest bazą przestrzeni $\operatorname{im}(\phi)$.

Pokażemy, że dla każdego dopełnienia prostego przestrzeni $\ker(\phi)$ i każdej jego bazy $\alpha_1, \ldots, \alpha_k$ zbiór wektorów $\phi(\alpha_1), \ldots, \phi(\alpha_k)$ rozpina $\operatorname{im}(\phi)$. Następnie pokażemy, że układ ten jest liniowo niezależny.

Niech $\beta \in \operatorname{im}(\phi)$. Chcemy pokazać, że $\beta \in \operatorname{lin}(\phi(\alpha_1), \ldots, \phi(\alpha_k))$. Wiadomo, że $\beta = \phi(\alpha) = \phi(\alpha' + \alpha'')$, gdzie $\alpha' \in \ker(\phi)$oraz $\alpha'' \in U$. A zatem $\alpha'' = a_1\alpha_1 + \ldots + a_k\alpha_k$, dla pewnych $a_1, \ldots, a_k \in K$. Zatem:

$$\begin{aligned}\beta = \phi(\alpha) &= \phi(\alpha' + \alpha'') \\ &= \phi(\alpha') + \phi(a_1\alpha_1 + \ldots + a_k\alpha_k) \\ &= 0 + a_1\phi(\alpha_1) + \ldots + a_k\phi(\alpha_k) \\ &\in \operatorname{lin}(\phi(\alpha_1), \ldots, \phi(\alpha_k)).\end{aligned}$$

Dowodzimy liniowej niezależności tego układu. Przypuśćmy, że $a_1\phi(\alpha_1) + \ldots + a_k\phi(\alpha_k) = 0$. Wówczas $\phi(a_1\alpha_1 + \ldots + a_k\alpha_k) = 0$, a zatem $a_1\alpha_1 + \ldots + a_k\alpha_k \in \ker(\phi)$. Ale przecież $\alpha_1, \ldots, \alpha_k$ jest bazą U. A zatem $a_1\alpha_1 + \ldots + a_k\alpha_k \in \ker(\phi) \cap U = \{0\}$. A szczególności $a_1\alpha_1 + \ldots + a_k\alpha_k = 0$, czyli $a_1 = \ldots = a_k = 0$, bo $\alpha_1, \ldots, \alpha_k$ jest bazą U. Układ $\phi(\alpha_1), \ldots, \phi(\alpha_k)$ jest zatem liniowo niezależny.

Dowód twierdzenia jest teraz natychmiastowy. Na mocy twierdzenia o wymiarze sumy prostej mamy:

$$\dim(V) = \dim \ker(\phi) + \dim(U) = \dim \ker(\phi) + k = \dim \ker(\phi) + \dim \operatorname{im}(\phi).$$

□

Rezultat nasz ma sens również w przypadku, gdy V jest przestrzenią nieskończonego wymiaru. Dowód wymaga pewnej modyfikacji, ale w rezultacie okazuje się, że jeśli $\phi : V \to W$ jest liniowe i $\dim(V) = \infty$, to wymiary przestrzeni $\ker(\phi)$ oraz $\operatorname{im}(\phi)$ nie mogą być jednocześnie skończenie wymiarowe.

Patrząc na formułę wiążącą wymiar jądra i obrazu przekształcenia liniowego warto pochylić się dłużej nad przypadkami, gdy $\dim \ker(\phi) = \{0\}$ oraz, gdy $\dim \operatorname{im}(\phi) = \dim W$.

Definicja 9.5: Monomorfizm, epimorfizm, izomorfizm

Przekształcenie liniowe $\phi : V \to W$ nazywamy:

- MONOMORFIZMEM, gdy ϕ jest różnowartościowe, tzn. $\phi(\alpha) = \phi(\beta) \Rightarrow \alpha = \beta$, dla $\alpha, \beta \in V$.
- EPIMORFIZMEM, gdy jest „na", tzn. gdy dla każdego $\gamma \in W$ istnieje $\alpha \in V$ takie, że $\phi(\alpha) = \gamma$.
- IZOMORFIZMEM, gdy ϕ jest różnowartościowe i „na" (to znaczy, gdy ϕ jest bijekcją).

Przykłady

- Przekształcenie liniowe $\phi : \mathbb{R} \to \mathbb{R}^3$ dane wzorem $\phi(x) = (x, x, x)$, jest monomorfizmem, ale nie jest epimorfizmem.
- Przekształcenie liniowe $\phi : \mathbb{R}^3 \to \mathbb{R}$ dane wzorem $\phi(x, y, z) = x$ jest epimorfizmem, ale nie jest monomorfizmem.
- Przekształcenie liniowe $\phi : \mathbb{R}^4 \to M_{2\times 2}(\mathbb{R})$ dane poniższym wzorem jest izomorfizmem:

$$\phi((x_1, x_2, x_3, x_4)) = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

Obserwacja 9.3

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Wówczas:

- ϕ jest monomorfizmem wtedy i tylko wtedy, gdy $\ker(\phi) = \{0\}$,
- ϕ jest epimorfizmem wtedy i tylko wtedy, gdy $\operatorname{im}(\phi) = W$.

Dowód. Tylko pierwsza równoważność wymaga dowodu. Jeśli ϕ jest monomorfizmem oraz dla pewnego $\alpha \in V$ mamy $\phi(\alpha) = 0$, to skoro $\phi(0) = 0$, z różnowartościowości ϕ wynika, że $\alpha = 0$. A zatem $\ker(\phi) = 0$. Na odwrót: jeśli $\ker(\phi) = \{0\}$ oraz dla pewnych $\alpha, \beta \in V$ mamy $\phi(\alpha) = \phi(\beta)$, to z liniowości $\phi(\alpha - \beta) = 0$. Skoro $\ker(\phi) = \{0\}$, to $\alpha - \beta = 0$, czyli $\alpha = \beta$. W szczególności ϕ to monomorfizm. □

Wniosek 9.3

Niech $\phi : V \to W$ będzie przekształceniem liniowym, przy czym $\dim(V), \dim(W) < \infty$. Wówczas:

- jeśli ϕ jest monomorfizmem, to $\dim V \leqslant \dim W$,
- jeśli ϕ jest epimorfizmem, to $\dim W \leqslant \dim V$,
- jeśli ϕ jest izomorfizmem, to $\dim W = \dim V$.

Co ważne, tezę punktu trzeciego można dla przestrzeni skończenie wymiarowych łatwo odwrócić.

Wniosek 9.4

iech $\phi : V \to W$ będzie przekształceniem liniowym i niech $\dim V = \dim W < \infty$. Wówczas następujące warunki są równoważne:

(a) ϕ jest monomorfizmem,

(b) ϕ jest epimorfizmem,

(c) ϕ jest izomorfizmem.

Definicja 9.6: Izomorfizm przestrzeni liniowych

Mówimy, że przestrzenie V i W nad ciałem K są IZOMORFICZNE, jeśli istnieje izomorfizm $\phi : V \to W$. Oznaczenie: $V \simeq W$.

To właśnie izomorfizm przestrzeni liniowych jest pojęciem, które mówi o tym, że jakieś dwie przestrzenie są „jednakowe" z punktu widzenia algebry liniowej, czyli mają tą samą strukturę. Co to znaczy jednakowe? Przytoczymy teraz kilka rezultatów, które o tym mówią.

Twierdzenie 9.2

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

(a) ϕ jest izomorfizmem,

(b) ϕ przeprowadza każdą bazę przestrzeni V na bazę przestrzeni W,

(c) ϕ przeprowadza pewną bazę przestrzeni V na bazę przestrzeni W.

Dowód. Pokażemy tezę w sytuacji, gdy $\dim(V) < \infty$. Ogólny przypadek uzasadnia się analogicznie, ale notacja jest bardziej uciążliwą. Dowodzimy implikacje $(i) \Rightarrow (ii)$, $(ii) \Rightarrow (iii)$, $(iii) \Rightarrow (i)$.

Wiemy już, że dla dowolnego przekształcenia liniowego $\phi : V \to W$ i każdej podprzestrzeni U w W takiej, że $V = \ker(\phi) \oplus U$ przekształcenie ϕ przeprowadza bazę przestrzeni U na bazę przestrzeni $\mathrm{im}(\phi)$. Jeśli ϕ jest izomorfizmem, to $\mathrm{im}(\phi) = W$, a także $\ker(\phi) = \{0\}$, więc $V = U$. Zatem ϕ przeprowadza bazę przestrzeni V na bazę przestrzeni W. Pokazaliśmy $(i) \Rightarrow (ii)$. Implikacja $(ii) \Rightarrow (iii)$ jest oczywista.

Przypuśćmy, że pewne przekształcenie liniowe ϕ przeprowadza bazę $\alpha_1, \ldots, \alpha_n$ przestrzeni V na bazę $\beta_1, \ldots, \beta_n$ przestrzeni W. Pokażemy, że ϕ jest izomorfizmem czyli, że $\ker(\phi) = \{0\}$ oraz $\mathrm{im}(\phi) = W$.

Jeśli $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n \in \ker(\phi)$, to $0 = \phi(\alpha) = a_1\beta_1 + \ldots + a_n\beta_n$, co wobec liniowej niezależności układu $(\beta_1, \ldots, \beta_n)$ oznacza, że $a_1 = \ldots = a_n = 0$. A zatem $\alpha = 0$. A zatem wobec dowolności wyboru α mamy $\ker(\phi) = \{0\}$.

Weźmy $\beta \in W$ i niech $\beta = a_1\beta_1 + \ldots + a_n\beta_n$. Wówczas $\beta = \phi(\alpha)$, dla pewnego $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n$. A zatem z dowolności wyboru β mamy $W = \mathrm{im}(\phi)$. □

Rezultat ten pozwala nam udowodnić kluczowy wniosek.

Wniosek 9.5: Izomorfizm z K^n

Niech V, W będą przestrzeniami skończonego wymiaru nad ciałem K. Wówczas następujące warunki są równoważne:

(i) $V \simeq W$,

(ii) $\dim V = \dim W$,

W konsekwencji, mamy izomorfizm $V \simeq K^{\dim V}$.

Implikacja $(i) \Rightarrow (ii)$ została pokazana już wcześniej w oparciu o formułę

$$\dim V = \dim \ker(\phi) + \dim \mathrm{im}(\phi),$$

gdzie $\phi : V \to W$ jest izomorfizmem. Mamy bowiem $\ker(\phi) = 0$ oraz $\mathrm{im}(\phi) = W$. A zatem $\dim V = \dim W$.

Implikacja $(ii) \Rightarrow (i)$ wymaga dodatkowego rezultatu. Nie mamy bowiem jak dotąd metody definiowania przekształceń liniowych inaczej niż wzorami. Jak pokazać, że jakieś przekształcenie liniowe istnieje? Oto kluczowy rezultat.

Twierdzenie 9.3: O jednoznaczności na bazie

Niech V, W będą przestrzeniami liniowymi nad ciałem K. Niech $\alpha_1, \ldots, \alpha_n$ będzie bazą przestrzeni V, zaś $\beta_1, \ldots, \beta_n$ niech będzie dowolnym układem wektorów przestrzeni W. Wówczas istnieje **dokładnie jedno** takie przekształcenie liniowe $\phi : V \to W$, że

$$\phi(\alpha_1) = \beta_1, \quad \phi(\alpha_2) = \beta_2, \quad \ldots, \quad \phi(\alpha_n) = \beta_n. \qquad (*)$$

Stwierdzenie to będziemy przywoływać często mówiąc w skrócie, że **przekształcenie liniowe zadane jest w sposób jednoznaczny na bazie**.

Zobaczmy jak wygląda uzasadnienie implikacji (ii) $\Rightarrow$ (i) wniosku powyżej. Niech $\alpha_1, \ldots, \alpha_n$ będzie bazą V oraz niech $\beta_1, \ldots, \beta_n$ będzie bazą W. Definiujemy $\phi(V)$ warunkiem $\phi(\alpha_i) = \beta_i$, dla $1 \leqslant i \leqslant n$. Wiadomo, że takie przekształcenie istnieje dla każdego układu wektorów W równolicznego z bazą $\alpha_1, \ldots, \alpha_n$. Takie przekształcenie ϕ, które wybraliśmy, musi być jednak izomorfizmem, bo przeprowadza bazę V na bazę W.

Pokażmy teraz twierdzenie o jednoznacznym definiowaniu przekształcenia liniowego na bazie.

Dowód. Idea dowodu jest prosta. Pokażemy najpierw, że istnieje przekształcenie ϕ spełniające $(*)$, a następnie pokażemy, że jest tylko jedno takie przekształcenie.

Dla każdego $\alpha \in V$ istnieją $a_1, \ldots, a_n \in K$, że $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n$ (czyli: współrzędne α w bazie $\alpha_1, \ldots, \alpha_n$). Oznacza to, że poniższe przekształcenie ϕ jest dobrze określone:

$$\phi(\alpha) = a_1\beta_1 + \ldots + a_n\beta_n.$$

Podstawiając za α kolejne α_i, dla $1 \leqslant i \leqslant n$, dostajemy oczywiście $\phi(\alpha_i) = \beta_i$, bo jedyną niezerową współrzędną wektora α_i w bazie $\alpha_1, \ldots, \alpha_i$ jest i-ta współrzędna równa 1. A zatem ϕ spełnia $(*)$. Dlaczego jest liniowe? Jeśli $\alpha' = a'_1\alpha_1 + \ldots + a'_n\alpha_n$, dla pewnych współrzędnych $a'_1, \ldots, a'_n$, to

$$\phi(\alpha + \alpha') = (a_1 + a'_1)\beta_1 + \ldots + (a_n + a'_n)\beta_n = a_1\beta_1 + \ldots + a_n\beta_n + a'_1\beta_1 + \ldots + a'_n\beta_n = \phi(\alpha) + \phi(\alpha').$$

Podobnie pokazujemy, że dla każdego $a \in K$ mamy $\phi(a\alpha) = a\phi(\alpha)$. A zatem ϕ jest liniowe.

Załóżmy, że istnieje przekształcenie liniowe $\psi : V \to W$ takie, że $\psi(\alpha_i) = \beta_i$, dla $1 \leqslant i \leqslant n$. Wówczas dla każdego $\alpha \in V$ mamy:

$$\begin{aligned}\psi(\alpha) &= \psi(a_1\alpha_1 + \ldots + a_n\alpha_n) = a_1\psi(\alpha_1) + \ldots + a_n\psi(\alpha_n) = \\ &= a_1\beta_1 + \ldots + a_n\beta_n = \\ &= a_1\phi(\alpha_1) + \ldots + a_n\phi(\alpha_n) = \phi(a_1\alpha_1 + \ldots + a_n\alpha_n) = \phi(\alpha).\end{aligned}$$

A zatem przekształcenie liniowe ϕ spełniające $(*)$ jest wyznaczone jednoznacznie. □

Warto jest dokładnie przemyśleć i zrozumieć konsekwencje zaprzeczenia założeń powyższego faktu.

- W przypadku, gdy układ $\alpha_1, \ldots, \alpha_n$ jest liniowo niezależny, ale nie rozpina V, dla ustalonego układu $\beta_1, \ldots, \beta_n$ w przestrzeni W może istnieć wiele różnych przekształceń liniowych spełniających $(*)$.

  Weźmy $\alpha_1 = (1, 0, 0), \alpha_2 = (0, 1, 0), \beta_1 = (1, 0, 0), \beta_2 = (0, 1, 0)$. Układ $((1, 0, 0), (0, 1, 0))$ jest liniowo niezależny w $\mathbb{R}^3$, ale nie rozpina tej przestrzeni. Istnieje więc więcej niż jedno przekształcenie liniowe $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ takie, że $\phi(\alpha_1) = \beta_1$ oraz $\phi(\alpha_2) = \beta_2$. Jednym z nich jest identyczność, a drugim symetria względem płaszczyzny $\mathrm{lin}((1, 0, 0), (0, 1, 0))$ wzdłuż $\mathrm{lin}(0, 0, 1)$,

- Jeśli $\alpha_1, \ldots, \alpha_n$ nie jest liniowo niezależny, to dla pewnego układu $\beta_1, \ldots, \beta_n$ może nie istnieć przekształcenie liniowe $\phi : V \to W$ spełniające $(*)$.

  Weźmy $\alpha_1 = (1, 0), \alpha_2 = (0, 1), \alpha_3 = (1, 1)$ oraz $\beta_1 = \beta_2 = \beta_3 = (1, 0, 0)$. Teraz układ $\alpha_1, \alpha_2, \alpha_3$ rozpina $\mathbb{R}^2$, ale jest liniowo zależny. Gdyby istniało takie przekształcenie liniowe $\phi : \mathbb{R}^2 \to \mathbb{R}^3$, że $\phi(\alpha_1) = \phi(\alpha_2) = \phi(\alpha_3)$, to z liniowości ϕ mielibyśmy

  $$(0, 0, 0) = \phi((0, 0)) = \phi(\alpha_1 + \alpha_2 - \alpha_3) = \phi(\alpha_1) + \phi(\alpha_2) - \phi(\alpha_3) = (1, 0, 0).$$

Zauważmy, że nie postawiliśmy żadnych warunków odnośnie układu wektorów $\{\beta_i\}$ – poza równolicznością z bazą $\alpha_1, \ldots, \alpha_n$. Wektory te mogą być nawet wszystkie równe wektorowi zerowemu przestrzeni W.

Wniosek 9.6

Każde przekształcenie liniowe $\phi : K^n \to K^m$ jest wyznaczone jednoznacznie przez swoje wartości na bazie standardowej $\epsilon_1, \ldots, \epsilon_n$ przestrzeni K^n.

W szczególności, biorąc macierz $A = (a_{ij}) \in M_{m\times n}(K)$ i przypisując i-temu wektorowi standardowemu ϵ_i w K^n i-tą kolumnę macierzy A, czyli

$$\phi(\epsilon_i) = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

otrzymujemy wzór na ϕ postaci:

$$\begin{aligned} \phi((x_1, \ldots, x_n)) &= \phi(x_1(1, 0, \ldots, 0) + x_2(0, 1, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 1)) = \\ &= \phi((x_1\epsilon_1 + \ldots x_n\epsilon_n)) = x_1\phi(\epsilon_1) + \ldots + x_n\phi(\epsilon_n) = \\ &= x_1(a_{11}, \ldots, a_{m1}) + \ldots + x_n(a_{1n}, \ldots, a_{mn}) = \\ &= (a_{11}x_1 + \ldots a_{1n}x_n,\ a_{21}x_1 + \ldots + a_{2n}x_n,\ \ldots\ ,\ a_{m1}x_1 + \ldots + a_{mn}x_n). \end{aligned}$$

Przykład. Niech $f : \mathbb{R}^2 \to \mathbb{R}^4$ zadane będzie na bazie standardowej warunkami:

$$f((1, 0)) = (1, 1, 2, 1), \quad f((0, 1)) = (0, 3, 1, -2).$$

Wówczas z liniowości f mamy:

$$\begin{aligned} f((x, y)) = f(x(1, 0) + y(0, 1)) = x \cdot f((1, 0)) + y \cdot f((0, 1)) = x(1, 1, 2, 1) + y(0, 3, 1, -2) = \\ = (x, x + 3y, 2x + y, x - 2y). \end{aligned}$$

A zatem pokazaliśmy, że każda przestrzeń n-wymiarowa nad ciałem K jest izomorficzna z przestrzenią K^n. Izomorfizmy te można uzyskać na wiele sposobów, o czym powiemy następnym razem.

* * *

Na koniec poczyńmy drobną uwagę dotyczącą geometrycznej motywacji dla rozważania przekształceń liniowych. Rozważmy przekształcenie liniowe $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ dane wzorem:

$$\phi((x_1, x_2, x_3)) = (2x_1, 2x_2, 4x_3).$$

Z geometrycznego punktu widzenia przekształcenie to nie jest ani obrotem, ani symetrią czy rzutem, ale biorąc rozkład:

$$\mathbb{R}^3 = \text{lin}((1, 0, 0), (0, 1, 0)) \oplus \text{lin}(0, 0, 1) \qquad (*)$$

widzimy, że ϕ ograniczone do każdego ze składników jest na nim homotetią, bo:

- dla każdego $v \in \text{lin}((1, 0, 0), (0, 1, 0))$ mamy $\phi(v) = 2v$,
- dla każdego $w \in \text{lin}((0, 0, 1))$ mamy $\phi(w) = 4w$.

Idea jest zatem następująca: znając ϕ na każdym ze składników prostych, „wiemy co robi" ϕ na całej przestrzeni liniowej! To nie przypadek, ale zwiastun wielkiej i ważnej teorii, którą zajmiemy się w kolejnym semestrze. Kluczową kwestią jest wskazywanie rozkładów takich, jak $(*)$, dla innych przekształceń liniowych. Na razie jednak zajmiemy się zbudowaniem podstaw, zwłaszcza opisem przekształceń liniowych pomiędzy przestrzeniami skończonego wymiaru w języku macierzowym.

9.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Znajdowanie wzoru na przekształcenie liniowe zadane na bazie)
   Znajdź wzory na przekształcenie liniowe $\phi : \mathbb{R}^n \to \mathbb{R}^m$ zadane następującymi warunkami:
   (a) $\phi : \mathbb{R}^2 \to \mathbb{R}^3, \phi((3,1)) = (4,5,-1), \phi((7,2)) = (-3,0,5)$,
   (b) $\phi : \mathbb{R}^3 \to \mathbb{R}^2, \phi((1,2,1)) = (7,2), \phi((3,2,4)) = (20,17), \phi((5,1,2)) = (17,12)$,
   (c) $\phi : \mathbb{R}^3 \to \mathbb{R}^3, \phi((1,0,1)) = (5,1,3), \phi((0,1,1)) = (2,3,4), \phi((1,0,0)) = (6,7,7)$

2. Czy istnieje (jeśli tak, podaj przykład) przekształcenie liniowe $\phi : \mathbb{R}^4 \to \mathbb{R}^3$ spełniające warunki:
   (a) $\ker(\phi) = \{(x_1,x_2,x_3,x_4) : x_1 - x_2 + 6x_3 + 2x_4 = 0\}$, $\operatorname{im}(\phi) = \operatorname{lin}((2,3,1))$.
   (b) $\ker(\phi) = \operatorname{lin}((1,0,3,3))$, $\operatorname{im}(\phi) = \{(x_1,x_2,x_3) : 4x_1 + 5x_2 - x_3 = 0\}$.
   (c) $\ker(\phi) = \operatorname{lin}((1,1,1,1),(1,1,1,0))$, $\operatorname{im}(\phi) = \operatorname{lin}((1,1,1),(1,1,0))$.

3. Niech $\mathcal{A} = \{(1,2,2),(1,2,1),(1,1,2)\}$. Niech $V = \operatorname{lin}(\alpha_1,\alpha_2)$ oraz $W = \operatorname{lin}(\alpha_3)$. Znajdź wzór na $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ będący symetrią $\mathbb{R}^3$ względem podprzestrzeni V wzdłuż podprzestrzeni W.

4. (♠ Znajdowanie bazy oraz wymiaru obrazu i jądra przekształcenia liniowego) Dla każdego z poniższych przekształceń znajdź bazę i wymiar jego obrazu oraz bazę i wymiar jego jądra.
   (a) $\phi : \mathbb{R}^3 \to \mathbb{R}^2, \phi((x_1,x_2,x_3)) = (2x_1 + x_2 - 3x_3, x_1 + 4x_2 + 2x_3)$,
   (b) $\phi : \mathbb{R}^3 \to \mathbb{R}^4, \phi((x_1,x_2,x_3)) = (4x_1 + 3x_2 + 5x_3, x_1 + 2x_2 + x_3, 2x_1 - x_2 + 3x_3, 6x_1 + 7x_2 + 7x_3)$.

5. (♠ Stwierdzanie, kiedy przekształcenie liniowe jest monomorfizmem, epimorfizmem, izomorfizmem).
   Dla jakich wartości parametru $r \in \mathbb{R}$ przekształcenie
   (a) $\phi : \mathbb{R}^3 \to \mathbb{R}^5, \phi((x_1,x_2,x_3)) = (x_1 + x_2 + 2x_3, 2x_1 + x_2 + x_3, x_1 + 3x_2 + rx_3, 5x_1 + 3x_2 + 4x_3, x_1 + 2x_2 + 5x_3)$ jest monomorfizmem?
   (b) $\phi : \mathbb{R}^4 \to \mathbb{R}^3, \phi((x_1,x_2,x_3,x_4)) = (4x_1 + x_2 + rx_3 + x_4, 3x_1 + 2x_2 + x_3 + x_4, 2x_1 + 3x_2 + 3x_3 + x_4)$ jest epimorfizmem?
   (c) $\phi : \mathbb{R}^4 \to \mathbb{R}^4, \phi((x_1,x_2,x_3,x_4)) = (5x_1 - x_2 + rx_3 + 5x_4, 2x_1 - 3x_2 - 6x_3 + rx_4, 3x_1 + 2x_2 + x_3 + 4x_4, x_1 + 5x_2 + 7x_3 + 3x_4)$ jest izomorfizmem?

6. (♠ Stosowanie wzoru na sumę wymiarów jąder i obrazu).
   - Czy istnieje przekształcenie liniowe $\phi : \mathbb{R}^5 \longrightarrow \mathbb{R}^5$ takie, że $\operatorname{im}\phi = \ker\phi$?
   - Czy istnieje przekształcenie liniowe $\phi : K^{10} \longrightarrow K^{11}$, że $\dim\ker\phi = 2$, $\dim\ker\operatorname{im}\phi = 9$?
   - Czy istnieje epimorfizm $\phi : M_{3\times 2}(\mathbb{R}) \to \mathbb{R}_{\leqslant 6}[x]$?
   - Dane są przekształcenia liniowe $\phi_1, \phi_2 : \mathbb{R}^5 \longrightarrow \mathbb{R}^3$. Czy musi istnieć niezerowy wektor $\alpha \in \mathbb{R}^5$ taki, że $\phi_1(\alpha) = \phi_2(\alpha)$?

7. Niech $T : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ będzie przekształceniem określonym wzorem
   $$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}.$$
   Wykaż, że T jest przekształceniem liniowym. Znajdź wymiary obrazu i jądra tego przekształcenia.

8. Niech $\mathbb{R}_{\leqslant n}[x]$ będzie podprzestrzenią przestrzeni wielomianów złożoną z wielomianów stopnia $\leqslant n$. Niech $f : \mathbb{R}_{\leqslant n}[x] \to \mathbb{R}_{\leqslant n}[x]$ będzie funkcją zadaną wzorem $f(w(x)) = w(x+1) - w(x)$. Wykaż, że f jest przekształceniem liniowym, znajdź bazy i wymiary jądra oraz obrazu f.

9. Niech V będzie przestrzenią liniową i niech $\phi_1, \phi_2 : V \to V$ będą przekształceniami liniowymi.
   (a) Udowodnij, że $\ker\phi_1 \cap \ker\phi_2 \subseteq \ker\phi_1 + \ker\phi_2$. Kiedy zachodzi równość?
   (b) Niech $\operatorname{im}\phi_1 + \operatorname{im}\phi_2 = V = \ker\phi_1 + \ker\phi_2$. Wykaż, że $\operatorname{im}\phi_1 \cap \operatorname{im}\phi_2 = \{0\} = \ker\phi_1 \cap \ker\phi_2$.

10. Niech $V_0, \ldots, V_{n+1}$ będą przestrzeniami liniowymi nad ciałem K, przy czym $V_0 = V_{n+1} = \{0\}$. Niech $f_i : V_i \to V_{i+1}$, dla $i = 0, \ldots, n$ będzie ciągiem przekształceń liniowych, takim że $\operatorname{im} f_i = \ker f_{i+1}$. Wykaż, że
    $$\sum_{i=1}^{n} (-1)^i \dim V_i = 0$$

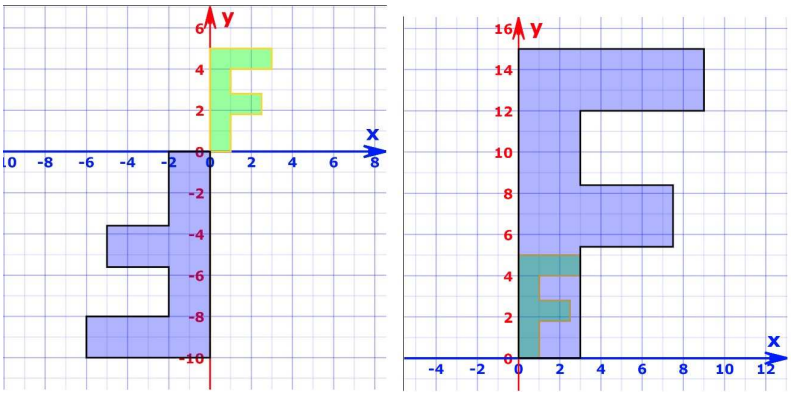
9.3 Uzupełnienie. Geometria przekształceń liniowych

Aby przekonać się, że przekształcenia liniowe mają sporo wspólnego z podstawowymi przekształceniami geometrycznymi płaszczyzny znanymi ze szkoły, rozważmy funkcje $f : \mathbb{R}^2 \to \mathbb{R}^2$ postaci:

$$f(x, y) = (ax + by, cx + dy),$$

gdzie $a, b, c, d \in \mathbb{R}$. Jak wiemy tylko funkcje zadane tymi wzorami mogą opisywać przekształcenia liniowe z K^2 do K^2. Jeśli spojrzymy na owe funkcje[1] z punktu widzenia geometrii analitycznej, a więc jako funkcje z (płaszczyzny kartezjańskiej) $\mathbb{R}^2$ do $\mathbb{R}^2$, wówczas wśród f znajdują się między innymi:

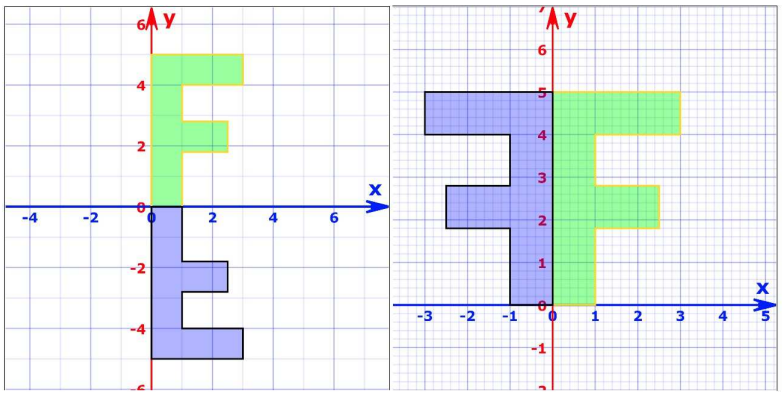
- jednokładność o skali λ i środku $(0, 0)$ postaci $f(x, y) = (\lambda x, \lambda y)$



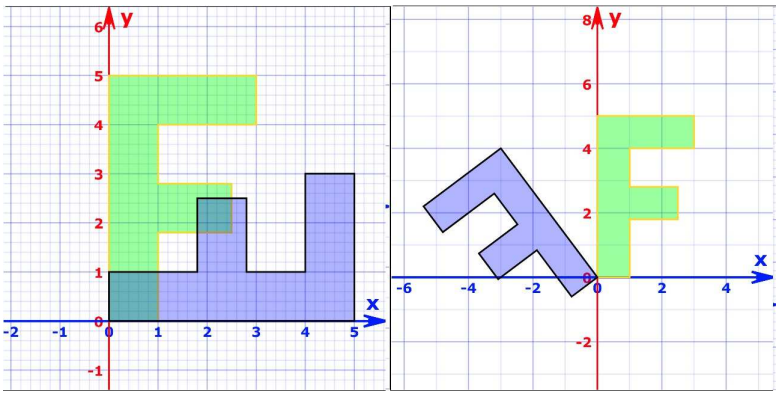
Rys. 1. Jednokładność o skalach odpowiednio $\lambda = -2$ (z lewej) oraz $\lambda = 3$.

- symetria prostopadła względem prostej $y = \mathrm{tg}(\theta/2)$ dana wzorem

$$f(x, y) = (\cos\theta \cdot x + \sin\theta \cdot y, \sin\theta \cdot x - \cos\theta \cdot y).$$



Rys. 2. Symetria prostopadła względem prostej $x = 0$ (z lewej) oraz $y = 0$ (z prawej).

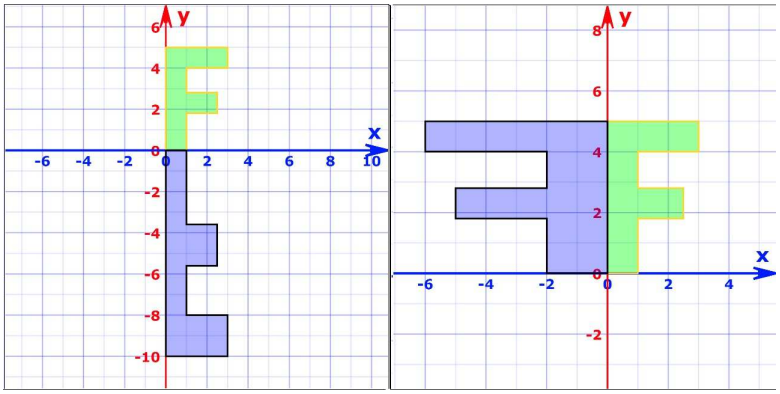


Rys. 3. Symetria prostopadła względem prostej $y = x$ (z lewej) dana wzorem $(x, y) \mapsto (y, x)$ oraz względem prostej $y = -3x$ (z prawej) dana wzorem $(x, y) \mapsto (-\frac{4}{5}x - \frac{3}{5}y, -\frac{3}{5}x + \frac{4}{5}y)$.

[1] Rysunki uzyskane za pomocą portalu https://www.mathsisfun.com/algebra/matrix-transform.html, gdzie można samodzielnie poeksperymentować — choćby po to, by przekonać się, że nie dostajemy tylko izometrii czy podobieństw.

- powinowactwa prostokątne o osiach OX lub OY oraz skali λ, czyli przekształcenia dane wzorami:

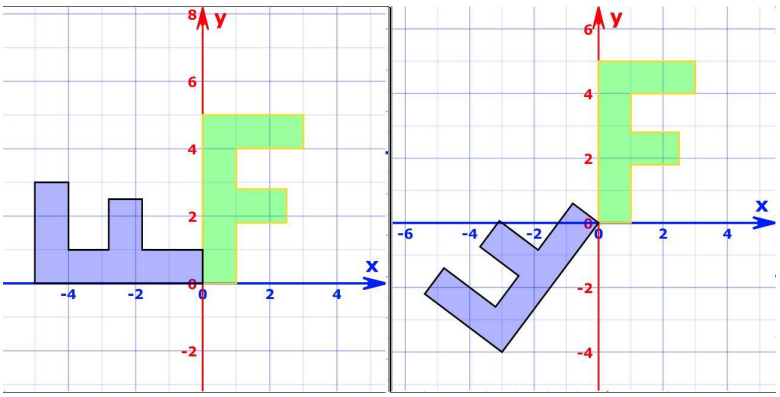
$$f(x,y) = (x, \lambda y), \quad g(x,y) = (\lambda x, y).$$



Rys. 4. Powinowactwa względem osi OX oraz OY o skali $\lambda = -2$.

- obroty o kąt θ względem punktu $(0,0)$ dane wzorem:

$$f(x,y) = (\cos\theta x - \sin\theta y, \sin\theta x + \cos\theta y).$$



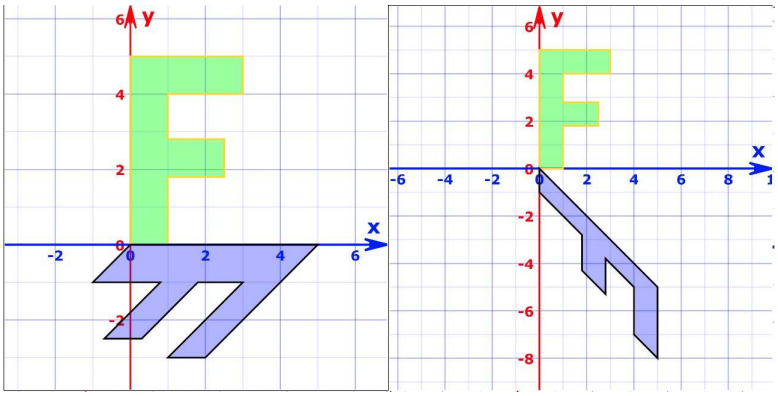
Rys. 5. Obroty o kąty $\pi/2$ oraz .... ?

- złożenia wyżej wymienionych w pewnej kolejności, np. dla przekształceń danych wzorami:

$$f(x,y) = (x+y, y), \quad g(x,y) = (y, -x),$$

mamy:

$$f(g(x,y)) = (-x+y, -x), \quad g(f(x,y)) = (y, -x-y)$$



Rys. 6. Funkcja $(x,y) \mapsto (-x+y, -x)$ (z lewej) oraz $(x,y) \mapsto (y, -x-y)$ (z prawej).

Zachęcam do zabawy apletem `https://www.mathsisfun.com/algebra/matrix-transform.html`. Należy oczywiście pamiętać, że obrazem przestrzeni wymiaru 2 przy przekształceniu liniowym nie musi być przestrzeń dwuwymiarowa, co można zobaczyć np. dla odwzorowania $(x,y) \mapsto (x+y, 2x+2y)$. W drugim semestrze w odwzorowaniach opisanych wyżej rozpoznamy przekształcenia afiniczne płaszczyzny $\mathbb{R}^2$, które przeprowadzają punkt $(0,0)$ na $(0,0)$. Można je utożsamić z przekształceniami liniowymi. Póki co — zostańmy przy intuicjach pamiętając, że budujemy ogólną teorię „geometryczną", nie tylko dla K^n.

9.4 Uzupełnienie. Układy współrzędnych i układy równań

Pokazaliśmy właśnie izomorfizm przestrzeni skończenie wymiarowej V nad ciałem K z przestrzenią współrzędnych $K^{\dim V}$. Jak wspomnieliśmy, takich izomorfizmów jest wiele i polegają one na wyborze bazy $\mathcal{A}$ przestrzeni V i zadaniu funkcji $\phi_{\mathcal{A}}$ postaci:

$$V \ni \alpha \xrightarrow{\phi_{\mathcal{A}}} (a_1, \ldots, a_n) \in K^n,$$

gdzie $n = \dim V$ oraz $a_1, \ldots, a_n$ są współrzędnymi α w bazie $\mathcal{A}$. Zobaczmy dwa przykłady.

- Przestrzeń wielomianów $K[x]_{\leqslant 3}$ stopnia nie większego niż 3 nad ciałem K jest wymiaru 4. Wybierzmy dwie bazy tej przestrzeni postaci:

  $$\mathcal{A} = (1, x, x^2, x^3), \quad \mathcal{B} = (1, x+1, (x+1)^2, (x+1)^3).$$

  Mamy zatem dwa izomorfizmy $\phi_{\mathcal{A}}, \phi_{\mathcal{B}} : K[x]_{\leqslant 3} \to K^4$ zadane warunkami:

  $$\begin{array}{llll} \phi_{\mathcal{A}}(1) = (1,0,0,0), & \phi_{\mathcal{A}}(x) = (0,1,0,0), & \phi_{\mathcal{A}}(x^2) = (0,0,1,0), & \phi_{\mathcal{A}}(x^3) = (0,0,0,1), \\ \phi_{\mathcal{B}}(1) = (1,0,0,0), & \phi_{\mathcal{B}}(x+1) = (0,1,0,0), & \phi_{\mathcal{B}}((x+1)^2) = (0,0,1,0), & \phi_{\mathcal{A}}((x+1)^3) = (0,0,0,1). \end{array}$$

  Biorąc np. wielomian $w(x) = x^3 + 1 = (x+1)^3 - 3(x-1)^2 + 3(x+1) \in K[x]$ mamy:

  $$\phi_{\mathcal{A}}(w(x)) = (1,0,0,1), \quad \phi_{\mathcal{B}}(w(x)) = (1,3,-3,1).$$

  Zauważmy, że przy ustalonym utożsamieniu, np. przy $\phi_{\mathcal{A}}$ możemy opisywać podprzestrzenie $K[x]_{\leqslant 3}$ jako zbiory rozwiązań układów równań o czterech zmiennych, utożsamianych ze współrzędnymi wektorów $\phi_{\mathcal{A}}(w)$, dla $w \in K[x]_{\leqslant 3}$. W ten sposób możemy więc uznać, że dowolna podprzestrzeń przestrzeni skończonego wymiaru może być opisana układem równań liniowych, o ile wybierzemy wcześniej układ współrzędnych, które to współrzędne traktować będziemy dalej jako zmienne.

  Oto przykład. Rozważmy równanie o czterech zmiennych postaci $x_1 - x_4 = 0$. Przy izomorfizmie $\phi_{\mathcal{A}}$ można uznać, że równanie to opisuje te wielomiany w $K[x]_{\leqslant 3}$, których współrzędne $\phi_{\mathcal{A}} = (x_1, x_2, x_3, x_4)$ spełniają rownanie $x_1 - x_4 = 0$. A zatem są to wielomany postaci $a + bx + cx^2 + ax^3$, dla dowolnych $a, b, c \in K$.

  Jeśli jednak rozważymy izomorfizm $\phi_{\mathcal{B}}$, wówczas należy dowolny wielomian $w \in K[x]_{\leqslant 3}$ przedstawić w bazie $\mathcal{B}$ i szukać wielomianów, które przy 1 oraz $(x+1)^3$ mają te same współczynniki. Będzie to oczywiście inny zbiór wielomianów niż te, uzyskane przy izomorfizmie $\phi_{\mathcal{A}}$. To zagadnienie będzie dla nas ważne w drugim semestrze i rozważać je będziemy w znacznie większej ogólności.

- Przestrzeń $W \subseteq K^4$ rozwiązań układu $x_1 + x_2 + x_3 + x_4 = 0$ jest trójwymiarowa i biorąc jej bazy:

  $$\mathcal{A} = ((1,-1,0,0), (1,0,-1,0), (1,0,0,-1)), \quad \mathcal{B} = ((-1,1,0,0), (0,-1,1,0), (0,0,-1,1))$$

  możemy zadań przykładowe dwa izomorfizmy $\phi_{\mathcal{A}}, \phi_{\mathcal{B}} : W \to K^3$ warunkami:

  $$\begin{array}{lll} \phi_{\mathcal{A}}((1,-1,0,0)) = (1,0,0), & \phi_{\mathcal{A}}((1,0,-1,0)) = (0,1,0), & \phi_{\mathcal{A}}((1,0,0,-1)) = (0,0,1), \\ \phi_{\mathcal{B}}((-1,1,0,0)) = (1,0,0), & \phi_{\mathcal{B}}((0,-1,1,0)) = (0,1,0), & \phi_{\mathcal{B}}((0,0,-1,1)) = (0,0,1). \end{array}$$

Definicja 9.7: Układ współrzędnych

Izomorfizmy n-wymiarowej przestrzeni V nad V nad K na przestrzeń K^n nazywamy UKŁADAMI WSPÓŁRZĘDNYCH w V. UKŁADEM WSPÓŁRZĘDNYCH ZWIĄZANYCH Z BAZĄ $(v_1, \ldots, v_n)$ w V nazywamy izomorfizm $\sigma : V \to K^n$ przeprowadzający v_j na j-ty wektor bazy standardowej ϵ_j.

W szczególności, biorąc dowolną bazę $\mathcal{A}$ w K^n można zadać na tej przestrzeni układ współrzędnych odpowiadający tej bazie. Warto jednak patrzeć na to ogólnie: tzw. współrzędne wektora v przestrzeni skończenie wymiarowej V w bazie $\mathcal{A}$, które już jakiś czas rozważamy, to nic innego jak obraz tego wektora w odpowiednim układzie współrzędnych wyznaczonym przez tę bazę, czyli $\phi_{\mathcal{A}}(v)$. Widzimy zatem, że mamy różne sposoby zadawania współrzędnych na przestrzeniach skończonego wymiaru. Jest to istotne, jak się przekonamy w drugim semestrze, choćby dlatego, że przekształcenia liniowe mogą wyznaczać pewne układy współrzędnych, w których geometria przekształcenia liniowego jest szczególnie dobrze widoczna.

9.5 [Dodatek. Przekształcenia w nieskończonym wymiarze]

(w przygotowaniu...)

9.6 [Trivia. Kiedy przekształcenie jest liniowe?]

Stwierdzenie „przekształcenie liniowe" jest pewnym skrótem myślowym. Precyzyjniej byłoby myśleć, że przekształcenie ϕ zdefiniowane na wykładzie jest K-liniowe co oznacza, że jest przekształceniem pomiędzy przestrzeniami liniowymi określonymi nad tym samym ciałem K. Rozważmy następujący przykład.

Przykład 1. Rozważmy $\mathbb{C}$ jako dwuwymiarową przestrzeń liniową nad ciałem $\mathbb{R}$. Wówczas przekształcenie $\phi : \mathbb{C} \to \mathbb{C}$ zadane wzorem

$$\phi(z) = \overline{z}$$

jest $\mathbb{R}$-liniowe. Rzeczywiście, dla dowolnych liczb zespolonych z_1, z_2, oraz dla dowolnego $a \in \mathbb{R}$ mamy:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{az_1} = a\overline{z_1}.$$

Natomiast sprzężenie nie jest przekształceniem $\mathbb{C}$-liniowym, czyli przekształceniem pomiędzy $\mathbb{C}^1$ oraz $\mathbb{C}^1$. Rzeczywiście, nie jest prawdą, że dla dowolnych $a \in \mathbb{C}$ oraz $z \in \mathbb{C}$ mamy $\overline{az} = a\overline{z}$, bowiem $\overline{az} = \overline{a} \cdot \overline{z}$.

Powyższy przykład pokazuje także, że w definicji przekształcenia liniowego warunek (2) nie wynika z warunku (1). Łatwo również sprawdzić, że poniższa funkcja $f : \mathbb{R}^2 \to \mathbb{R}^2$ spełnia (2), ale nie spełnia (1):

$$f(x, y) = \begin{cases} (x, 0), & \text{dla } y = 0 \\ (0, 0), & \text{dla } y \neq 0. \end{cases}.$$

(ciąg dalszy w przygotowaniu...)

9.7 [Notka historyczna. Skąd się wzięły morfizmy?]

(ciąg dalszy w przygotowaniu...)

Rozdział 10

Macierz przekształcenia liniowego. Mnożenie macierzy

10.1 Wykład dziesiąty

Wprowadziliśmy na ostatnim wykładzie definicję przekształceń liniowych i pokazaliśmy, że każda przestrzeń liniowa wymiaru n nad ciałem K jest izomorficzna z przestrzenią K^n. Oznacza to, że każdemu przekształceniu liniowemu przestrzeni skończenie wymiarowych przypisać można przekształcenie liniowe opisane pewną macierzą. Jak się okaże, przypisanie to można dokonać na różne sposoby.

Definicja 10.1: Mnożenie macierzy

Niech $A \in M_{1\times n}(K)$ oraz $B \in M_{n\times 1}(K)$ będą postaci:

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}.$$

Wówczas iloczynem $A \cdot B$ macierzy A, B nazywamy macierz rozmiaru 1×1, której jedyny wyraz ma postać:

$$\sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n.$$

Iloczynem macierzy $A = [a_{ij}] \in M_{m\times k}(K)$ oraz $B = [b_{ij}] \in M_{k\times n}(K)$ nazywamy taką macierz $C = [c_{ij}] \in M_{m\times n}(K)$ rozmiarów $m \times n$, że dla każdych i, j mamy:

$$c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{ik} b_{kj}$$

Innymi słowy: wyraz w i-tym wierszu i j-tej kolumnie macierzy C to jedyny wyraz macierzy będącej iloczynem i-tego wiersza macierzy A (rozmiaru $1 \times k$) i j-tego wiersza macierzy B (rozmiaru $k \times 1$).

Przykład. Jeśli

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

to nie istnieje iloczyn macierzy postaci BA, zaś

$$AB = \begin{bmatrix} 1\cdot 1 + 0\cdot 2 + 2\cdot 1 & 1\cdot 3 + 0\cdot 1 + 2\cdot 0 & 1\cdot 1 + 0\cdot 0 + 2\cdot 0 & 1\cdot 0 + 0\cdot 0 + 2\cdot 1 \\ 1\cdot 1 + 3\cdot 2 + 1\cdot 1 & 1\cdot 3 + 3\cdot 1 + 1\cdot 0 & 1\cdot 1 + 3\cdot 0 + 1\cdot 0 & 1\cdot 0 + 3\cdot 0 + 1\cdot 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 1 & 2 \\ 8 & 6 & 1 & 1 \end{bmatrix}.$$

Ważna motywacja. Niech $\phi : K^n \to K^m$ będzie przekształceniem liniowym postaci:

$$\phi(x_1, \ldots, x_n) = (a_{11}x_1 + \ldots + a_{1n}x_n,\ a_{21}x_1 + \ldots + a_{2n}x_n,\ \ldots,\ a_{m1}x_1 + \ldots + a_{mn}x_n).$$

Wówczas wartości tego przekształcenia liniowego możemy wyznaczać za pomocą formuły:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix}.$$

Warto zauważyć, że w powyższej interpretacji kolumny rozważanej macierzy $[a_{ij}]$ są po prostu obrazami wektorów bazy standardowej przestrzeni K^n przy przekształceniu ϕ, a więc:

$$\phi(\epsilon_1) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \ldots, \quad \phi(\epsilon_n) = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Wiemy jednocześnie, że przekształcenie liniowe $\phi : V \to W$ przestrzeni skończonego wymiaru określić można jednoznacznie na dowolnej bazie $\mathcal{A}$ przestrzeni V. Możemy więc przyjrzeć się obrazom elementów bazy $\mathcal{A}$ i odczytywać ich współrzędne w (dowolnej) bazie przestrzeni W.

Definicja 10.2: Macierz przekształcenia liniowego

Niech V, W będą przestrzeniami liniowymi nad ciałem K i niech $\phi : V \to W$ będzie przekształceniem liniowym. Niech

- $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ będzie (uporządkowaną) bazą przestrzeni V,
- $\mathcal{B} = (\beta_1, \ldots, \beta_m)$ będzie (uporządkowaną) bazą przestrzeni W.

Macierzą przekształcenia ϕ w bazach $\mathcal{A}$, $\mathcal{B}$ nazywamy taką macierz $A = (a_{ij}) \in M_{m\times n}(K)$, że dla każdego $1 \leqslant j \leqslant n$:

$$\phi(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \ldots + a_{mj}\beta_m = \sum_{i=1}^{m} a_{ij}\beta_i,$$

Taką macierz A oznaczamy $M(\phi)^{\mathcal{B}}_{\mathcal{A}}$. Innymi słowy:

w j-tej kolumnie macierzy $M(\phi)^{\mathcal{B}}_{\mathcal{A}}$ stoją współrzędne wektora $\phi(\alpha_j)$ w bazie $\mathcal{B}$.

Przykład. Niech $\phi : \mathbb{R}^3 \to \mathbb{R}^2$ będzie dane wzorem

$$\phi((x_1, x_2, x_3)) = (2x_1 + x_2 - x_3, x_1 - x_2 + x_3).$$

Rozważmy teraz bazy:

- $\mathcal{A} = (\alpha_1 = (1, 0, 1), \alpha_2 = (0, 1, 2), \alpha_3 = (2, 1, 0))$ – baza przestrzeni $\mathbb{R}^3$,
- $\mathcal{B} = (\beta_1 = (0, 1), \beta_2 = (2, 0))$ – baza przestrzeni $\mathbb{R}^2$.

Wówczas:

$$\begin{aligned} \phi(\alpha_1) = (1, 2) &= 2 \cdot \beta_1 + \frac{1}{2} \cdot \beta_2, \\ \phi(\alpha_2) = (-1, 1) &= 1 \cdot \beta_1 - \frac{1}{2} \cdot \beta_2, \\ \phi(\alpha_3) = (5, 1) &= 1 \cdot \beta_1 + \frac{5}{2} \cdot \beta_2 \end{aligned}$$

Możemy zatem przypisać przekształceniu ϕ macierz, która w kolumnach będzie miała współrzędne obrazów kolejnych wektorów z bazy $\mathcal{A}$, ale współrzędne te będą w bazie $\mathcal{B}$. Macierz tą oznaczać będziemy jako: $M(\phi)^{\mathcal{B}}_{\mathcal{A}}$, tzn.

$$M(\phi)^{\mathcal{B}}_{\mathcal{A}} = \begin{bmatrix} 2 & 1 & 1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{5}{2} \end{bmatrix}.$$

Oczywiście mamy też

$$M(\phi)^{st}_{st} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix},$$

bowiem

$$\begin{aligned} \phi((1,0,0)) = (2,1) &= 2\cdot(1,0) + 1\cdot(0,1), \\ \phi((0,1,0)) = (1,-1) &= 1\cdot(1,0) - 1\cdot(0,1), \\ \phi((0,0,1)) = (-1,1) &= -1\cdot(1,0) + 1\cdot(0,1) \end{aligned}$$

W kolumnach macierzy $M(\phi)^{st}_{st}$ stoją wektory zawierające współrzędne obrazów wektorów z bazy standardowej w $\mathbb{R}^3$, a te współrzędne są w bazie standardowej $\mathbb{R}^2$.

Ogólnie, jeśli $f : K^n \to K^m$ jest przekształceniem liniowym zadanym wzorem

$$f((x_1,\ldots,x_n)) = (a_{11}x_1 + \ldots a_{1n}x_n,\ a_{21}x_1 + \ldots + a_{2n}x_n,\ \ldots,\ a_{m1}x_1 + \ldots + a_{mn}x_n).,$$

wówczas

$$M(\phi)^{st}_{st} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

Po co nam takie macierze przekształceń liniowych? Dobrą motywacją może być następujący przykład. Załóżmy, że mamy bazę

$$\mathcal{A} = ((1,0,1),(2,0,-1),(5,1,3))$$

przestrzeni $\mathbb{R}^3$ i rozważmy przekształcenie $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ o następującej macierzy $M(\phi)^{\mathcal{A}}_{\mathcal{A}}$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Zgodnie z definicją, w pierwszej kolumnie są współrzędne wektora $\phi((1,0,1))$ w bazie $\mathcal{A}$, w drugiej kolumnie są współrzędne wektora $\phi((2,0,-1))$ w bazie $\mathcal{A}$, zaś w trzeciej kolumnie są współrzędne wektora $(5,1,3)$ w bazie $\mathcal{A}$, czyli:

$$\begin{aligned} \phi((1,0,1)) &= 1\cdot(1,0,1) + 0\cdot(2,0,-1) + 0\cdot(5,1,3) \\ \phi((2,0,-1)) &= 0\cdot(1,0,1) - 1\cdot(2,0,-1) + 0\cdot(5,1,3) \\ \phi((5,1,3)) &= 0\cdot(1,0,1) + 0\cdot(2,0,-1) - 1\cdot(5,1,3) \end{aligned}$$

Czy Czytelnik widzi, że z tej konkretnej postaci macierzy możemy odczytać informację, że ϕ jest w istocie symetrią względem $\mathrm{lin}((1,0,1))$ wzdłuż $\mathrm{lin}((2,0,-1),(5,1,3))$? Tak właśnie jest! Z drugiej strony zupełnie nie „widać" tego rozważając tylko wzór tego przekształcenia liniowego. Nietrudno go wyznaczyć.

Skoro $\phi((1,0,1)) + \phi((2,0,-1)) = \phi((3,0,0)) = (1,0,1) - (2,0,-1) = (-1,0,2)$, to

$$\phi((1,0,0)) = \left(-\frac{1}{3}, 0, \frac{2}{3}\right).$$

Stąd

$$\phi((0,0,1)) = \phi((1,0,1)) - \phi((1,0,0)) = (1,0,1) - \left(-\frac{1}{3},0,\frac{2}{3}\right) = \left(\frac{4}{3},0,\frac{1}{3}\right).$$

Natomiast

$$\phi((0,1,0)) = \phi((5,1,3)) - \phi((5,0,0)) - \phi((0,0,3)) = (-5,-1,-3) - \left(-\frac{5}{3},0,\frac{10}{3}\right) - \left(\frac{12}{3},0,\frac{3}{3}\right) = \left(-\frac{22}{3},-1,-\frac{22}{3}\right).$$

W szczególności wzór przekształcenia ϕ to:

$$\phi((x_1,x_2,x_3)) = \left(-\frac{1}{3}x_1 - \frac{22}{3}x_2 + \frac{4}{3}x_3, -x_2, \frac{2}{3}x_1 - \frac{22}{3}x_2 + \frac{1}{3}x_3\right).$$

Czytelnik zechce sprawdzić, że po wstawieniu do uzyskanego wzoru kolejno wektorów $(1,0,1)$, $(2,0,-1)$ oraz $(5,1,3)$ otrzymamy odpowiednio wektory $(1,0,1)$, $(-2,0,1)$, $(-5,-1,-3)$. Czy znając jedynie wzór przekształcenia ϕ lub macierz $M(\phi)^{st}_{st}$ umielibyśmy stwierdzić, że jest to „w istocie" pewna symetria?

Dalsze przykłady.

- Niech $\mathrm{id} : K^n \to K^n$ będzie identycznością, zaś $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ niech będzie bazą K^n. Wówczas:

$$M(\mathrm{id})^{st}_{st} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

  Macierz tę nazywać będziemy MACIERZĄ IDENTYCZNOŚCIOWĄ (rozmiaru n), ozn. I (lub I_n, gdy chcemy podkreślić jej rozmiar). Zauważmy jednak, że w innych niż standardowe bazach, macierze przekształcenia id nie muszą być wcale identycznościowe.

- Niech $\phi_a : K^n \to K^n$ będzie jednokładnością o skali a. W bazach standardowych macierz ϕ_a to:

$$M(\phi_a)^{st}_{st} = a \cdot I_n.$$

- Niech $V = W \oplus U$ i niech
  - $\phi : V \to V$ będzie **rzutem** na W wzdłuż U.
  - ψ będzie **symetrią** względem W wzdłuż U

  Niech $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ będzie taką bazą przestrzeni V, że
  - $(\alpha_1, \ldots, \alpha_k)$ (dla pewnego $1 < k < n$) jest bazą przestrzeni W,
  - $(\alpha_{k+1}, \ldots, \alpha_n)$ jest bazą przestrzeni U.

  Wówczas macierz $M(\phi)^{\mathcal{A}}_{\mathcal{A}}$ ma w pierwszych k kolumnach pierwsze k wektorów bazy standardowej K^n, zaś dalej kolumny zerowe, natomiast macierz $M(\psi)^{\mathcal{A}}_{\mathcal{A}}$ ma w pierwszych k kolumnach pierwsze k wektorów, zaś dalej $n-k$ wektorów przeciwnych do wektorów z bazy standardowej K^n:

$$M(\phi)^{\mathcal{A}}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}, \quad M(\psi)^{\mathcal{A}}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & -1 \end{bmatrix},$$

Wykonywanie przekształcenia liniowego pomiędzy przestrzeniami skończenie wymiarowymi można realizować za pomocą mnożenia macierzy tego przekształcenia przez odpowiedni wektor współrzędnych.

Obserwacja 10.1

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Niech $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ będzie bazą przestrzeni V oraz $(\beta_1, \ldots, \beta_m)$ niech będzie bazą przestrzeni W. Jeśli

- $a_1, \ldots, a_n$ są współrzędnymi wektora α w bazie $\mathcal{A}$,
- $b_1, \ldots, b_m$ są współrzędnymi wektora $\phi(\alpha)$ w bazie $\mathcal{B}$, to

$$M(\phi)^{\mathcal{B}}_{\mathcal{A}} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

Dowód. Niech $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = [a_{ij}]$, dla $1 \leqslant i \leqslant m$ orz $1 \leqslant j \leqslant m$. Wówczas z definicji tej macierzy mamy:

$$\phi(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \ldots + a_{mj}\beta_m.$$

Stą dla dowolnego $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n$ mamy:

$$\begin{aligned}\phi(\alpha) &= \phi(a_1\alpha_1 + \ldots + a_n\alpha_n)\\ &= a_1\phi(\alpha_1) + \ldots + a_n\phi(\alpha_n)\\ &= a_1(a_{11}\beta_1 + a_{21}\beta_2 + \ldots + a_{m1}\beta_m) + \ldots + a_n(a_{1n}\beta_1 + a_{2n}\beta_2 + \ldots + a_{mn}\beta_m)\\ &= (a_{11}a_1 + \ldots + a_{1n}a_n)\beta_1 + \ldots + (a_{m1}a_1 + \ldots + a_{mn}a_n)\beta_m.\end{aligned}$$

Zapisaliśmy więc $\phi(\alpha)$ w bazie $\mathcal{B}$. Z drugiej strony wiemy, że współrzędne te równe są $b_1, \ldots, b_m$. A zatem rzeczywiście b_i powstaje przez przemnożenie i-tego wiersza macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ przez macierz mającą kolumnę $a_1, \ldots, a_n$, czyli $b_i = a_{i1}a_1 + \ldots + a_{in}a_n$. □

Warto odnotować wniosek dotyczący tzw. MACIERZY ZAMIANY WSPÓŁRZĘDNYCH.

Wniosek 10.1

Jeśli $\mathcal{A}, \mathcal{B}$ są bazami przestrzeni V i $C = M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}}$, gdzie $\mathrm{id} = \mathrm{id}_V$ jest identycznością na V, to dla każdego $\alpha \in V$: jeśli $a_1, \ldots, a_n$ są współrzędnymi α w bazie $\mathcal{A}$, zaś $b_1, \ldots, b_n$ są jego współrzędnymi w bazie $\mathcal{B}$, to:

$$C \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Aby zrozumieć lepiej rolę mnożenia macierzy w teorii przekształceń liniowych odwołamy się do pojęcia znanego jeszcze ze szkoły — złożenia funkcji. Na poziomie zupełnie podstawowym poziomie pojęcie złożenia uświadomi nam dlaczego definicja macierzy wygląda tak, jak została określona. Zrozumiemy dlaczego iloczynu niektórych macierzy nie można określić (podobnie jak nie każde dwie funkcje można złozyć). Gdy wejdziemy w temat głębiej okaże się, że tłumacząc iloczyn macierzy na macierz złożenia pewnych przekształceń liniowych uzyskamy naturalne zrozumienie własności algebraicznych mnożenia macierzy.

Definicja 10.3: Złożenie przekształceń liniowych

Niech $\phi : V \to W$ oraz $\psi : W \to Z$ będą przekształceniami liniowymi przestrzeni nad ciałem K. ZŁOŻENIEM PRZEKSZTAŁCEŃ ϕ i ψ nazywamy odwzorowanie $\psi \circ \phi : V \to Z$ zadane wzorem:

$$(\psi \circ \phi)(v) = \psi(\phi(v)).$$

Oczywiście $\phi : V \to W$ oraz $\psi : W \to Z$ są przekształceniami liniowymi, to także $\psi \circ \phi$ jest przekształceniem liniowym, bo dla dowolnych $\alpha, \beta \in V$ mamy:

$$(\psi \circ \phi)(\alpha + \beta) = \psi(\phi(\alpha + \beta)) = \psi(\phi(\alpha) + \phi(\beta)) = \psi(\phi(\alpha)) + \psi(\phi(\beta)) = (\psi \circ \phi)(\alpha) + (\psi \circ \phi)(\beta).$$

Powyższa definicja jest niezwykle istotna. W matematyce nieustannie spotykamy się z problemem składania wielu przekształceń i badaniem jak wyglądają „rozkłady". Podstawową trudnością jest fakt, że składanie przekształceń liniowych nie jest przemienne. Np. dla $\phi, \psi : \mathbb{R}^2 \to \mathbb{R}^2$ zadanych wzorami:

$$\phi(x, y) = (x, 0), \quad \psi(x, y) = (x, x)$$

mamy

$$(\psi \circ \phi)(x, y) = \psi(x, 0) = (x, x), \quad (\phi \circ \psi)(x, y) = \phi(x, x) = (x, 0).$$

Twierdzenie 10.1: Składanie przekształceń, a mnożenie ich macierzy

Jeśli V, W, Z są przestrzeniami liniowymi nad K z bazami odpowiednio $\mathcal{A}, \mathcal{B}, \mathcal{C}$, oraz $\phi : V \to W$, $\psi : W \to Z$ są przekształceniami liniowymi, to:

$$M(\psi \circ \phi)^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}}$$

Dowód. Rozważmy następujące bazy odpowiednio przestrzeni V, W, Z:

$$\mathcal{A} = (\alpha_1, \ldots, \alpha_n), \quad \mathcal{B} = (\beta_1, \ldots, \beta_m), \quad \mathcal{C} = (\gamma_1, \ldots, \gamma_k).$$

Niech też dane będą macierze:

$$M(\phi)^{\mathcal{B}}_{\mathcal{A}} = (a_{ij}), \quad M(\psi)^{\mathcal{C}}_{\mathcal{B}} = (b_{ij}), \quad M(\psi \circ \phi)^{\mathcal{C}}_{\mathcal{A}} = (c_{ij})$$

dla odpowiednich zakresów i, j w każdej z macierzy. Z definicji macierzy przekształceń liniowych $\phi, \psi, \psi\circ\phi$:

$$\begin{aligned} \phi(\alpha_j) &= a_{1j} \cdot \beta_1 + a_{2j} \cdot \beta_2 + \ldots + a_{mj} \cdot \beta_m, \\ \psi(\beta_l) &= b_{1l} \cdot \gamma_1 + b_{2l} \cdot \gamma_2 + \ldots + b_{kl} \cdot \gamma_k, \\ (\psi \circ \phi)(\alpha_j) &= c_{1j} \cdot \gamma_1 + c_{2j} \cdot \gamma_2 + \ldots + c_{kj} \cdot \gamma_k. \end{aligned}$$

Z definicji złożenia oraz liniowości ψ mamy jednak:

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) = \psi(\phi(\alpha_j)) &= \psi(a_{1j} \cdot \beta_1 + a_{2j} \cdot \beta_2 + \ldots + a_{mj} \cdot \beta_m) = \\ &= a_{1j} \cdot \psi(\beta_1) + a_{2j} \cdot \psi(\beta_2) + \ldots + a_{mj} \cdot \psi(\beta_m). \end{aligned}$$

Rozkładamy każdy z wektorów $\psi(\beta_l)$ w bazie $\mathcal{C}$:

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) = \psi(\phi(\alpha_j)) = a_{1j} \cdot (b_{11} \cdot \gamma_1 + b_{21} \cdot \gamma_2 + \ldots + b_{k1} \cdot \gamma_k) + \\ + a_{2j} \cdot (b_{12} \cdot \gamma_1 + b_{22} \cdot \gamma_2 + \ldots + b_{k2} \cdot \gamma_k) + \cdots \\ + a_{mj} \cdot (b_{1m} \cdot \gamma_1 + b_{2m} \cdot \gamma_2 + \ldots + b_{km} \cdot \gamma_k). \end{aligned}$$

Grupujemy teraz wszystkie wyrazy stojące przy wektorach z bazy $\mathcal{C}$:

$$\begin{aligned} (\psi \circ \phi)(\alpha_j) = \psi(\phi(\alpha_j)) = (a_{1j}b_{11} + a_{2j}b_{12} + \ldots + a_{mj}b_{1m})\gamma_1 + \\ + (a_{1j}b_{21} + a_{2j}b_{22} + \ldots + a_{mj}b_{2m})\gamma_2 + \cdots \\ + (a_{1j}b_{k1} + a_{2j}b_{k2} + \ldots + a_{mj}b_{km})\gamma_k. \end{aligned}$$

A zatem wyraz c_{ij}, stojący przy wektorze γ_i w powyższym przedstawieniu, równy jest

$$a_{1j}b_{i1} + a_{2j}b_{i2} + \ldots + a_{mj}b_{im},$$

czyli jest on *iloczynem* i-tego wiesza $M(\psi)^{\mathcal{C}}_{\mathcal{B}}$ i j-tej kolumny $M(\phi)^{\mathcal{B}}_{\mathcal{A}}$. Uzyskaliśmy tezę. □

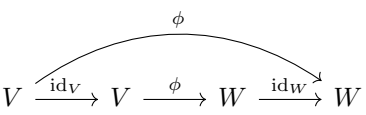
Odnotujmy ważny wniosek, fundamentalny dla naszych rozważań.

Wniosek 10.2: Zmiana baz w macierzy przekształcenia

Jeśli $\phi : V \to W$ jest przekształceniem liniowym, zaś $\mathcal{A}, \mathcal{A}'$ są bazami przestrzeni V oraz jeśli $\mathcal{B}, \mathcal{B}'$ są bazami przestrzeni W, to:

$$M(\phi)^{\mathcal{B}'}_{\mathcal{A}'} = M(\mathrm{id}_W)^{\mathcal{B}'}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}} \cdot M(\mathrm{id}_V)^{\mathcal{A}}_{\mathcal{A}'}. \qquad (*)$$

Dowód. Mamy równość $\phi = \mathrm{id}_W \circ \phi \circ \mathrm{id}_V$, którą wyrazić można na diagramie (o nich w dodatku):



a zatem $M(\phi)^{\mathcal{B}'}_{\mathcal{A}'} = M(\mathrm{id}_W \circ \phi \circ \mathrm{id}_V)^{\mathcal{B}'}_{\mathcal{A}'} = M(\mathrm{id}_W)^{\mathcal{B}'}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}} \cdot M(\mathrm{id}_V)^{\mathcal{A}}_{\mathcal{A}'}$. □

Przykład. Niech

$$U = \lin((10, 14, -4)) \subseteq \mathbb{R}^3 \quad \text{oraz} \quad W = \lin((0, 1, 0), (0, 1, 1))$$

tak, że $\mathbb{R}^3 = U \oplus V$. Wyznaczmy wzór na symetrię $\mathbb{R}^3$ względem U i wzdłuż W. Nazwijmy tą symetrię ϕ. Innymi słowy, szukamy $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33} \in \mathbb{R}$ takich, że:

$$\phi((x_1, x_2, x_3)) = (a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_2 + a_{12}x_2 + a_{23}x_3, a_{31}x_1 + a_{32}x_2 + a_{33}x_3).$$

Powyższy warunek jest równoważny temu, że:

$$\phi((1, 0, 0)) = (a_{11}, a_{21}, a_{31}), \quad \phi((0, 1, 0)) = (a_{12}, a_{22}, a_{32}), \quad \phi((0, 0, 1)) = (a_{13}, a_{23}, a_{33}).$$

Innymi słowy szukamy macierzy:

$$M(\phi)^{st}_{st} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Oczywiście zgodnie z definicją symetrii mamy:

$$\begin{aligned} \phi((10, 14, -4)) = (10, 14, -4) &= 1 \cdot (10, 14, -4) + 0 \cdot (0, 1, 0) + 0 \cdot (0, 1, 1) \\ \phi((0, 1, 0)) = -(0, 1, 0) &= 0 \cdot (10, 14, -4) + (-1) \cdot (0, 1, 0) + 0 \cdot (0, 1, 1) \\ \phi((0, 1, 1)) = -(0, 1, 1) &= 0 \cdot (10, 14, -4) + 0 \cdot (0, 1, 0) + (-1) \cdot (0, 1, 1) \end{aligned}$$

Nietrudno z tych warunków wywnioskować, że

$$\phi((0, 0, 1)) = \phi((0, 1, 1)) - \phi((0, 1, 0)), \quad \phi(1, 0, 0) = \frac{1}{10}\phi((10, 14, -4)) - \frac{7}{5}\phi((0, 1, 0)) + \frac{2}{5}\phi((0, 0, 1)),$$

ale naszym celem jest zaprezentowanie metody wykorzystującej formułę $(*)$.

Weźmy bazę $\mathcal{A} = ((10, 14, -4), (0, 0, 1), (0, 1, 1))$. Z warunków zapisanych wyżej wynika, że:

$$M(\phi)^{\mathcal{A}}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Możemy teraz skorzystać z formuły$M(\phi)^{st}_{st} = M(\mathrm{id})^{st}_{\mathcal{A}} \cdot M(\phi)^{\mathcal{A}}_{\mathcal{A}} \cdot M(\mathrm{id})^{\mathcal{A}}_{st}$.

Wyznaczmy macierze $M(\mathrm{id})^{st}_{\mathcal{A}}$ oraz $M(\mathrm{id})^{\mathcal{A}}_{st}$. Dla każdego $v \in \mathbb{R}^3$ mamy $\mathrm{id}(v) = v$, czyli

$$\begin{aligned} \mathrm{id}((10, 14, -4)) = (10, 14, -4) &= 10 \cdot (1, 0, 0) + 14 \cdot (0, 1, 0) - 4 \cdot (0, 0, 1) \\ \mathrm{id}((0, 1, 0)) = (0, 1, 0) &= 0 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 0 \cdot (0, 0, 1) \\ \mathrm{id}((0, 1, 1)) = (0, 1, 1) &= 0 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 1 \cdot (0, 0, 1). \end{aligned}$$

Nietrudno trudniej jest wyznaczyć współrzędne wektorów bazy standardowej w bazie $\mathcal{A}$:

$$\begin{aligned} \mathrm{id}((1, 0, 0)) = (1, 0, 0) &= \frac{1}{10} \cdot (10, 14, -4) - \frac{9}{5} \cdot (0, 1, 0) + \frac{2}{5} \cdot (0, 1, 1) \\ \mathrm{id}((0, 1, 0)) = (0, 1, 0) &= 0 \cdot (10, 14, -4) + 1 \cdot (0, 1, 0) + 0 \cdot (0, 1, 1) \\ \mathrm{id}((0, 0, 1)) = (0, 0, 1) &= 0 \cdot (10, 14, -4) + (-1) \cdot (0, 1, 0) + 1 \cdot (0, 1, 1). \end{aligned}$$

W szczególności $M(\mathrm{id})^{st}_{\mathcal{A}} = \begin{bmatrix} 10 & 0 & 0 \\ 14 & 1 & 1 \\ -4 & 0 & 1 \end{bmatrix}$, $M(\mathrm{id})^{\mathcal{A}}_{st} = \begin{bmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{9}{5} & 1 & -1 \\ \frac{2}{5} & 0 & 1 \end{bmatrix}$.

W rezultacie

$$M(\phi)^{st}_{st} = M(\mathrm{id})^{st}_{\mathcal{A}} \cdot M(\phi)^{\mathcal{A}}_{\mathcal{A}} \cdot M(\mathrm{id})^{\mathcal{A}}_{st} = \begin{bmatrix} 10 & 0 & 0 \\ 14 & 1 & 1 \\ -4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{10} & 0 & 0 \\ -\frac{9}{5} & 1 & -1 \\ \frac{2}{5} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{14}{5} & -1 & 0 \\ -\frac{4}{5} & 0 & -1 \end{bmatrix}.$$

Otrzymaliśmy więc szukany wzór $\phi((x_1, x_2, x_3)) = (x_1, \frac{14}{5}x_1 - x_2, -\frac{4}{5}x_1 - x_3)$.

Następnym razem przyjrzymy się macierzom izomorfizmów w kontekście istnienia przekształcenia odwrotnego oraz mnożenia macierzy. Wprowadzimy przy tym ważny obiekt – macierze odwracalne.

Wprowadziliśmy mnożenie macierzy i nadaliśmy mu od razu kontekst związany ze składaniem przekształceń liniowych. Ma to znaczenie także dla przejrzystości dowodów niektórych własności mnożenia macierzy.

1. **Mnożenie macierzy nie jest przemienne**. Idea jest taka, że po prostu składanie funkcji nie jest koniecznie przemienne, a nawet nie zawsze można wykonać złożenie funkcji z dowolnej strony. Warto pomyśleć o tym na przykładzie dwóch przekształceń ϕ, ψ przestrzeni $\mathbb{R}^2$, danych w bazach standardowych macierzami:
$$A = M(\phi)^{st}_{st} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = M(\psi)^{st}_{st} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$
Czytelnik może rozpoznać w przekształceniach ϕ, ψ omawiane już w dodatku o geometrii przekształceń liniowych: symetrię względem $\mathrm{lin}(1,0)$ wzdłuż $\mathrm{lin}(0,1)$ oraz obrót o $90°$. Mamy też:
$$BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$
Wyniki okazały się różne — i nic dziwnego. Złożenie $\psi \circ \phi$ opisane w bazach standardowych macierzą BA to odbicie symetryczne względem prostej $y = x$. Natomiast złożenie $\phi \circ \psi$ opisane w bazach standardowych macierzą AB to odbicie symetryczne względem prostej $y = -x$.

2. **Iloczyn niezerowych macierzy może być macierzą zerową**. Znowu nas to nie dziwi. Zdarza się choćby, że dwukrotnie wykonane przekształcenie staje się zerowe, choćby:
$$\phi(x, y) = (0, x) \quad \Rightarrow \quad (\phi \circ \phi)(x, y) = (0, 0).$$
Odpowiada to mnożeniu przez siebie odpowiedniej macierzy w bazie standardowej:
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

3. **Nie zachodzi prawo skracania** to znaczy, jeśli dla pewnych macierzy A, B, C mamy równość $AB = AC$ oraz $A \neq 0$, to może zajść sytuacja $B \neq C$ (chyba, że A jest odwracalna, o czym powiemy następnym razem). Istotnie, jeśli założymy na przykład, że A jest macierzą rzutu na przestrzeń jednowymiarową równą $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, to łatwo widzieć, że można dobrać do niej różne przekształcenia, choćby dane macierzami:
$$B = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

4. **Mnożenie macierzy jest łączne**, to znaczy: jeśli dane są macierze A, B, C oraz iloczyny AB i BC mają sens (do tego konieczne są odpowiednie rozmiary), to mamy:
$$(AB)C = A(BC).$$
Fakt ten wynika na przykład z tego, że dla dowolnych trzech funkcji f, g, h takich, że istnieją złożenia $f \circ g$ oraz $g \circ h$ mamy równość, dla każdego x:
$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x))) = f((g \circ h)(x)) = (f \circ (g \circ h))(x).$$
Skoro dla macierzy A, B, C można dobrać przekształcenia liniowe f, g, h, dla których są one macierzami w bazach standardowych, to z argumentów wyżej wnioskujemy równość $(AB)C = A(BC)$.

   Oczywiście własność łączności można dowodzić również wprost, przerachowując tożsamości postaci:
$$\sum_j a_{ij} \left(\sum_k b_{jk} c_{kl} \right) = \sum_{j,k} a_{ij} b_{jk} c_{kl} = \sum_k \left(\sum_j a_{ij} b_{jk} \right) c_{kl}.$$

5. **Mnożenie macierzy jest rozdzielne ze względu na dodawanie**, tzn. jeśli można wykonać odpowiednie mnożenia na macierzach A, B, C, mamy:
$$A(B + C) = AB + AC, \quad (A + B)C = AC + BC.$$
Zachęcam Czytelnika do podania funkcyjnego uzasadnienia, naśladującego dowód łączności.

10.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Wyznaczanie macierzy przekształcenia w bazie)
   Znajdź macierz przekształcenia liniowego ϕ w bazach $\mathcal{A}$, $\mathcal{B}$:
   - $\phi : \mathbb{R}^2 \to \mathbb{R}^4$, $\phi((x_1, x_2)) = (3x_1 + x_2, x_1 + 5x_2, -x_1 + 4x_2, 2x_1 + x_2)$,
     $\mathcal{A} = \{(3,1),(4,2)\}$, $\mathcal{B} = \{(1,0,1,0),(0,1,1,1),(0,1,2,3),(0,0,0,1)\}$,
   - $\phi : \mathbb{R}^3 \to \mathbb{R}^3$, $\phi((x_1, x_2, x_3)) = (4x_1 + x_2 + x_3, 3x_1 + 2x_2 + x_3, 3x_1 + 2x_2 + x_3)$,
     $\mathcal{A} = \{(3,1,1),(1,0,0),(5,1,0)\}$, $\mathcal{B} = \{(3,4,5),(4,1,1),(2,0,1)\}$.
   - $\phi : \mathbb{R}^4 \to \mathbb{R}^2$, $\phi((x_1, x_2, x_3, x_4)) = (5x_1 - 2x_2 + 3x_3 - x_4, 3x_1 + 4x_2 + 6x_4)$,
     $\mathcal{A} = \{(2,1,0,1),(1,0,3,1),(2,1,1,3),(3,1,2,1)\}$, $\mathcal{B} = \{(5,2),(3,1)\}$.

2. (♠ Wyznaczanie wzoru przekształcenia liniowego mając macierz w bazach)
   Dane są bazy $\mathcal{A}, \mathcal{B}$ przestrzeni $\mathbb{R}^3$: $\mathcal{A} = \{(3,1,1),(1,0,0),(5,1,0)\}$, $\mathcal{B} = \{(3,4,5),(4,1,1),(2,0,1)\}$.
   Znajdź wzór przekształcenia $\psi : \mathbb{R}^3 \to \mathbb{R}^3$ wiedząc, że:
   $$M(\psi)^{\mathcal{B}}_{\mathcal{A}} = \begin{bmatrix} 1 & 1 & 4 \\ 2 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix}.$$

3. Dla przekształcenia $\phi : \mathbb{R}^3 \to \mathbb{R}^2$ danego wzorem $\phi((x_1, x_2, x_3)) = (x_1 - x_2 + 2x_3, 3x_1 + x_2 + x_3)$ znajdź takie bazy $\mathcal{A}$ przestrzeni $\mathbb{R}^3$ oraz $\mathcal{B}$ przestrzeni $\mathbb{R}^2$, że $M(\phi)^{\mathcal{B}}_{\mathcal{A}} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$.

4. (♠ Baza jądra i obrazu przekształcenia liniowego danego macierzą w bazach)
   Niech $\mathcal{A} = ((1,1,0),(1,0,1),(0,1,0))$ oraz $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{bmatrix}$. Znajdź bazy jądra i obrazu przekształceń ϕ, ψ, gdzie $M(\phi)^{\mathcal{A}}_{st} = A^T$ oraz $M(\psi)^{st}_{\mathcal{A}} = A$.

5. (♠ Wyznaczanie macierzy złożenia w bazach) Dane są następujące bazy przestrzeni $\mathbb{R}^3, \mathbb{R}^4, \mathbb{R}^2$:
   $\mathcal{A} = \{(1,2,0),(3,5,0),(6,4,1)\}$, $\mathcal{B} = \{(1,0,1,0),(0,1,1,1),(0,1,1,0),(1,1,1,1)\}$, $\mathcal{C} = \{(5,4),(4,3)\}$.
   Przy tym dane jest również przekształcenie $\phi : \mathbb{R}^3 \to \mathbb{R}^4$ o wzorze
   $$\phi((x_1, x_2, x_3)) = (3x_1 - x_2 - 2x_3, 3x_1 + 4x_2 + x_3, 5x_1 + 2x_3, x_1 + x_2 + x_3)$$
   oraz przekształcenie ψ takie, że $M(\psi)^{\mathcal{C}}_{\mathcal{B}} = \begin{bmatrix} 0 & 2 & 1 & 1 \\ 2 & 5 & 0 & 3 \end{bmatrix}$. Znajdź $M(\psi \circ \phi)^{\mathcal{C}}_{\mathcal{A}}$ oraz $M(\psi \circ \phi)^{st}_{st}$

6. (♠ Zmienianie współrzędnych wektorów przy zmianie bazy)
   Niech $\phi : V \to W$ oraz $\psi : W \to Z$ będą przekształceniami liniowymi i niech
   $$M(\phi)^{\mathcal{B}}_{\mathcal{A}} = \begin{bmatrix} 2 & 1 & 4 & 5 \\ 1 & 0 & 4 & 3 \end{bmatrix}, \quad M(\psi)^{\mathcal{C}}_{\mathcal{B}} = \begin{bmatrix} 3 & 1 \\ 2 & 5 \\ 0 & 1 \end{bmatrix}$$
   w pewnych bazach $\mathcal{A}, \mathcal{B}, \mathcal{C}$ przestrzeni V, W, Z. Niech $\alpha \in V$ ma w bazie $\mathcal{A}$ współrzędne $1, -1, 3, -2$.
   Znajdź współrzędne wektora $\phi(\alpha)$ w bazie $\mathcal{B}$ oraz współrzędne wektora $(\psi \circ \phi)(\alpha)$ w bazie $\mathcal{C}$.

7. Niech $f\colon M_{2\times 2}(\mathbb{C}) \to M_{2\times 2}(\mathbb{C})$ będzie przekształceniem danym wzorem
   $$f(X) = AX - XA, \qquad \text{gdzie } A = \begin{bmatrix} 1 & i \\ 1 & 1 \end{bmatrix} \in \mathrm{M}_2(\mathbb{C}).$$
   Wykaż, że f jest przekształceniem liniowym. Wyznacz $M^{\mathcal{B}}_{\mathcal{B}}(f)$, gdzie $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, jest bazą złożoną z jedynek macierzowych. Znajdź bazy i wymiary przestrzeni $\ker f$ oraz $\operatorname{im} f$.

8. Niech $A, B \in M_{n\times n}(K)$. Udowodnij, że jeśli $AB = 0$, to $r(A) + r(B) \leqslant n$.

9. Udowodnij, że jeśli $A \in M_{n\times k}(K)$ oraz $B \in M_{k\times m}(K)$, to $r(AB) \leqslant \min\{r(A), r(B)\}$.

10. Niech $A, B, C \in M_{n\times n}(\mathbb{C})$ przy czym $r(A) = r$ oraz $AB = AC$. Jaki jest największy możliwy rząd macierzy $B - C$?

11. Wykaż, że jeśli macierz $A \in M_{n\times n}(K)$ spełnia $AC = CA$, dla każdej macierzy $C \in M_{n\times n}(K)$, to $A = aI$, dla pewnego $a \in K$.

10.3 Uzupełnienie. Diagramy przekształceń liniowych

Gdy rozważamy złożenia przekształceń liniowych, zapis „funkcyjny" jest często niezmiernie kłopotliwy. Często uniemożliwia on widzenie całej struktury złożeń tych przekształceń i odniesień pomiędzy odpowiednimi przestrzeniami. Aby radzić sobie jakoś z tym zjawiskiem przekształcenia (i obiekty, które one ze sobą wiążą) reprezentujemy często przy pomocy **diagramów**. Jest to spojrzenie charakterystyczne dla nowoczesnej matematyki, czerpiącej silnie z tak zwanej teorii kategorii (powiemy o niej w dalszych wykładach). Naszym celem jest zwrócenie uwagi na tak zwane MYŚLENIE DIAGRAMOWE.

Czym więc są owe diagramy? Zacznijmy od przykładu. Fakt istnienia złożenia przekształceń $\phi : V \to W$ oraz $\psi : W \to Z$ postaci $\psi \circ \phi : V \to Z$ opisujemy na diagramie w następujący sposób:

$$\begin{array}{ccc} V & \xrightarrow{\phi} & W \\ & \searrow^{\psi\circ\phi} & \downarrow{\scriptstyle \psi} \\ & & Z \end{array}$$

Zauważmy, że w konkretnym kontekście macierzy przekształcenia liniowego możemy interpretować mnożenie macierzy i istnienie iloczynu własnie w analogiczny sposób. Jeśli $A \in M_{p\times q}(K)$ oraz $B \in M_{r\times s}(K)$, to określając przekształcenia liniowe $\phi : K^q \to K^p$ oraz $\psi : K^s \to K^r$ wzorami (czyli poprzez macierz standardową): $A = M(\phi)^{st}_{st}$ oraz $B = M(\psi)^{st}_{st}$, to istnienie przekształcenia $\psi \circ \phi$ zależy od tego, czy $p = s$. Tylko wtedy istnieje iloczyn BA, który będzie macierzą w bazie standardowej przekształcenia $\psi \circ \phi$

$$\begin{array}{ccc} K^q & \xrightarrow{A} & K^p \\ & \searrow^{BA} & \downarrow{\scriptstyle B} \\ & & K^r \end{array}$$

Ogólniej DIAGRAMEM PRZEKSZTAŁCEŃ LINIOWYCH nazywać będziemy graf skierowany, którego wierzchołki etykietowane są przestrzeniami liniowymi (lub nie – jeśli mowa o dowolnych przestrzeniach), a krawędzie – przekształceniami liniowymi pomiędzy nimi. Przekształcenie liniowe postaci $f : V \to W$ oznaczamy zatem $V \xrightarrow{f} W$. Jeśli w diagramie występuje podgraf typu $V \xrightarrow{f} W \xrightarrow{g} Z$ to znaczy, że istnieje złożenie $g \circ f$. Podstawowym diagramem jest CIĄG, czyli diagram postaci:

$$V_1 \xrightarrow{\phi_1} V_2 \xrightarrow{\phi_2} V_3 \xrightarrow{\phi_3} \cdots \xrightarrow{\phi_{n-2}} V_{n-1} \xrightarrow{\phi_{n-1}} V_n \quad (\star) \ .$$

Złożenie $\phi_{n-1} \circ \ldots \circ \phi_1$ nazwiemy ZŁOŻENIEM wzdłuż ciągu $(\star)$. Powiemy, że diagram przekształceń jest PRZEMIENNY, jeśli dla dowolnych dwóch wierzchołków V, W tego diagramu, złożenia wzdłuż dowolnych dwóch ciągów tego diagramu o początkach w V i końcach w W są sobie równe (jako przekształcenia).

Dla przykładu, poniższy diagram jest przemienny, o ile $\phi_2 \circ \phi_1 = \psi_2 \circ \psi_1$.

$$\begin{array}{ccc} V & \xrightarrow{\phi_1} & W \\ \downarrow{\scriptstyle \psi_1} & & \downarrow{\scriptstyle \phi_2} \\ X & \xrightarrow{\psi_2} & Z \end{array}$$

Równość tę można intepretować w języku macierzowym. Jeśli dla macierzy $A \in M_{q\times p}(K), B \in M_{s\times q}(K)$, $C \in M_{r\times p}(K), D \in M_{s\times r}(K)$ mamy równość $BA = DC$, to traktując je jako macierze standardowe odpowiednich przekształceń liniowych przestrzeni skończonego wymiaru możemy zapisać diagram

$$\begin{array}{ccc} K^p & \xrightarrow{A} & K^q \\ \downarrow{\scriptstyle C} & & \downarrow{\scriptstyle B} \\ K^r & \xrightarrow{D} & K^s \end{array}$$

Podobnie, wracając do ciągu $(\star)$, możemy powiedzieć, że jeśli V_i są przestrzeniami skończonego wymiaru k_i, to rozważając macierz $A_i \in M_{k_{i+1}\times k_i}(K)$ oraz $A = M_{k_n\times k_1}(K)$ możemy napisać diagram przemienny ilustrujący równość $A = A_{n-1}A_{n-2}\ldots A_3A_2A_1$.

$$V_1 \xrightarrow{A_1} V_2 \xrightarrow{A_2} V_3 \xrightarrow{A_3} \cdots \xrightarrow{A_{n-2}} V_{n-1} \xrightarrow{A_{n-1}} V_n \quad (\star)$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overset{A}{\searrow}$$

Zobaczymy teraz przykład zadania z kolokwium z 2022 roku, które ma zarówno rozwiązanie w języku klasycznych kursowych pojęć algebry liniowych, jak i w języku diagramów. Oto ono.

Zadanie. Niech $A, B \in M_{n\times n}(K)$ oraz niech $\phi, \psi : K^n \to K^n$ będą przekształceniami liniowymi takimi, że $A = M(\phi)^{st}_{st}$ oraz $B = M(\psi)^{st}_{st}$. Wykaż, że:

(a) jeśli $CA = B$ oraz $DB = A$, dla pewnych $C, D \in M_{n\times n}(K)$, to $\ker(\phi) = \ker(\psi)$,

(b) jeśli $AE = B$ oraz $BF = A$, dla pewnych $E, F \in M_{n\times n}(K)$, to $\mathrm{im}(\phi) = \mathrm{im}(\psi)$.

(c) jeśli $\ker(\phi) = \ker(\psi)$, to istnieją $C, D \in M_{n\times n}(K)$ takie, że $CA = B$ oraz $DB = A$.

Autorką diagramowego rozwiązania tego zadania jest dr Agnieszka Bojanowska-Jackowska. Jest ono dostępne pod adresem `https://mimuw.edu.pl/~amecel/galkol/GALI-22-23-kol2r5ABJ.pdf`.

Główny pomysł jest następujący: w przestrzeni K^n wybieramy bazę standardową, więc oznaczamy przekształcenia liniowe tak, jak ich macierze w bazach standardowych. I tak na przykład w punkcie (a) mamy z założenia dwa przemienne diagramy przekształceń liniowych.

Pierwszy diagram ilustruje równość $CA = B$. Ten diagram ilustruje równość $DB = A$. Patrząc na pierwszy diagram, dla każdego wektora $v \in K^n$ mamy

$$Av = 0 \Rightarrow CAv = Bv = 0,$$

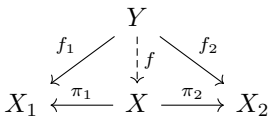
czyli $\ker(\phi) \subseteq \ker(\psi)$. Analogiczne rozumowanie w odniesieniu do drugiego diagramu prowadzi do zawierania $\ker(\psi) \subseteq \ker(\phi)$, co daje $\ker(\phi) = \ker(\psi)$.

* * *

Diagramy to ważne narzędzia w definiowaniu nowych przekształceń (i nie tylko) przy pomocy starych. Przekonamy się o tym wkrótce. Drugim ważnym aspektem związanym z diagramami są tzw. własności uniwersalne. Oto przykład takiej własności. Zachęcam do próby dowodu tego faktu.

Obserwacja 10.2

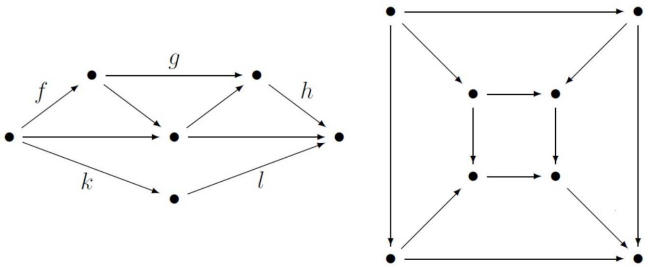
Niech X będzie przestrzenią liniową nad ciałem K wraz z podprzestrzeniami X_1, X_2. Określamy też epimorfizmy $\pi_1 : X \to X_1$ oraz $\pi_2 : X \to X_2$. Wówczas $X = X_1 \oplus X_2$ wtedy i tylko wtedy, gdy dla każdej przestrzeni liniowej Y (nad K) oraz przekształceń liniowych $f_1 : Y \to X_1$ oraz $f_2 : Y \to X_2$ istnieje **dokładnie jedno przekształcenie** $f : Y \to X$ takie, że przemienny jest diagram:



Nie jest naszym celem tworzenie jakiejkolwiek teorii w oparciu o język diagramów przemiennych. Przedmiotem tym zajmuje się teoria kategorii, o której jeszcze będziemy w przyszłości mówić. Na ten moment chcemy mieć po prostu swobodę korzystania z formułowania definicji czy obserwacji w języku podobnym do powyższego. Będziemy to oczywiście robić z pewną dozą ostrożności.

Zobaczmy kilka zadań, które obrazują na czym może polegać owe myślenie diagramowe.

Zadanie 1. Rozważmy następujący diagram przekształceń (a nawet funkcji, morfizmów...).



Cztery wewnętrzne *trójkąty* na diagramie po lewej są przemienne. Pokaż, że również zewnętrze jest przemienne, czyli $h \circ g \circ f = l \circ k$. Cztery wewnętrzne *trapezy* na diagramie po prawej są przemienne. Pokaż, że jeśli wewnętrzny kwadrat jest przemienny, to zewnętrzny też. A odwrotnie?

Zadanie 2. *Załóżmy, że w diagramie przemiennym odwzorowań liniowych*

$$\begin{array}{ccccccccc} U_1 & \xrightarrow{r_1} & U_2 & \xrightarrow{r_2} & U_3 & \xrightarrow{r_3} & U_4 & \xrightarrow{r_4} & U_5 \\ \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \downarrow{\scriptstyle f_4} & & \downarrow{\scriptstyle f_5} \\ V_1 & \xrightarrow{s_1} & V_2 & \xrightarrow{s_2} & V_3 & \xrightarrow{s_3} & V_4 & \xrightarrow{s_4} & V_5 \end{array}$$

wiersze są ciągami dokładnymi. Wykazać, że gdy f_1, f_2, f_4, f_5 są izomorfizmami, to f_3 jest izomorfizmem.

O co tu chodzi? Co oznacza, że wiersze są CIĄGAMI DOKŁADNYMI? Z definicji: ciąg

$$A \xrightarrow{f} B \xrightarrow{g} C$$

jest dokładny, jeśli $\mathrm{im}(f) = \ker(g)$. Pojęcie to ma olbrzymie znaczenie w algebrze i topologii, a zwłaszcza w tak zwanej algebrze homologicznej (jeszcze o nim powiemy). Mówiąc, że górny (i dolny) wiersz omawianego diagramu jest dokładny mamy na myśli to, że wszystkie poniższe ciągi

$$U_1 \xrightarrow{r_1} U_2 \xrightarrow{r_2} U_3, \quad U_2 \xrightarrow{r_2} U_3 \xrightarrow{r_3} U_4, \quad U_3 \xrightarrow{r_3} U_4 \xrightarrow{r_4} U_5$$

są dokładne (analogicznie w dolnym wierszu). Jest to szczególny przypadek tzw. lematu o piątce.

Zadanie 3. *Kolumny w poniższym diagramie przemiennym odwzorowań liniowych są ciągami dokładnymi (dokładność złożenia $0 \longrightarrow A \xrightarrow{r} B$ oznacza, że r to monomorfizm – czy to widać?)*:

$$\begin{array}{ccccccccc} & & 0 & & 0 & & 0 & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & U_1 & \longrightarrow & U_2 & \longrightarrow & U_3 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & V_1 & \xrightarrow{f} & V_2 & \xrightarrow{g} & V_3 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & W_1 & \longrightarrow & W_2 & \longrightarrow & W_3 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & 0 & & 0 & & 0 & & \end{array}$$

Pokazać, że gdy pierwszy i trzeci wiersz są dokładne oraz $g \circ f = 0$, to drugi wiersz jest także dokładny. Czy z dokładności pierwszego i drugiego wiersza wynika dokładność trzeciego?

Sama myśl o tym, że zaczniemy tu wypisywać jakieś macierze, mnożyć je czy reprezentować złożenia w pewnych bazach jest natychmiast traktowana jako szaleństwo. Tymczasem w matematyce mnóstwo jest takich zagadnień, jak wyżej. Jedne układy przekształceń są powiązane z innymi i bada się problem „przenoszenia się" (indukowania) własności jednego ciągu przekształceń na inny. Więcej podobnych zadań znajdą Państwo w zbiorze dr. Kubata. Przedstawione zadania dotyczą przestrzeni i przekształceń liniowych, ale także w wielu innych kontekstach można mówić o przekształceniach między rozmaitymi obiektami, odrywając się stosunkowo od ich natury, a jedynie ukazując strukturę zależności między nimi.

Rozdział 11

Izomorfizmy i macierze odwracalne Przestrzeń przekształceń liniowych

11.1 Wykład jedenasty

Od dwóch wykładów mówimy o przekształceniach liniowych, czyli funkcjach pomiędzy przestrzeniami liniowymi (nad ustalonym ciałem) zachowującymi kombinacje liniowe. Jeśli przekształcenie takie jest bijekcją, wówczas mówimy, że jest ono izomorfizmem, a przestrzenie, między którymi działa są przestrzeniami izomorficznymi. Z punktu widzenia algebry liniowej są to przestrzenie o identycznej strukturze. Pokazaliśmy, że w przypadku przestrzeni skończonego wymiaru niezmiennikiem odróżniającym przestrzenie, z dokładnością do izomorfizmu, jest właśnie wymiar przestrzeni liniowej.

Na ostatnim wykładzie powiedzieliśmy o składaniu przekształceń liniowych i odpowiadającej mu operacji mnożenia macierzy. Sugerowaliśmy, że badanie macierzy przekształcenia w różnych bazach umożliwia lepsze zrozumienie jego geometrycznej natury. Również (i nie tylko) kluczowe własności algebraiczne – bycie izomorfizmem, monomorfizmem czy epimorfizmem odczytać można w języku mnożenia macierzy i w języku złożeń. Podstawowej intuicji dostarczają tu izomorfizmy. Teoria funkcji podpowiada bowiem, że dla każdej bijekcji istnieje odwrotna bijekcja, a więc taka, że złożenie bijekcji z odwrotną do niej jest identycznością. Okazuje się, że dla bijekcji-izomorfizmów, bijekcja odwrotna jest również izomorfizmem.

Twierdzenie 11.1

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

(1) ϕ jest izomorfizmem,

(2) istnieje takie przekształcenie liniowe $\psi : W \to V$, że:

$$\psi \circ \phi = \mathrm{id}_V \text{ oraz } \phi \circ \psi = \mathrm{id}_W . \qquad (*)$$

Dowód. Niech ϕ będzie izomorfizmem. Określamy $\psi : W \to V$ warunkiem $\psi(\beta) = \alpha$, gdzie $\phi(\alpha) = \beta$. Jeśli $\psi(\beta_1) = \alpha_1$ oraz $\psi(\beta_2) = \alpha_2$, to $\phi(\alpha_1) = \beta_1$, $\phi(\alpha_2) = \beta_2$. Z liniowości ϕ mamy $\phi(\alpha_1 + \alpha_2) = \beta_1 + \beta_2$. A zatem

$$\psi(\beta_1 + \beta_2) = \alpha_1 + \alpha_2 = \psi(\beta_1) + \psi(\beta_2).$$

Analogicznie sprawdzamy $\psi(c\beta) = c\psi(\beta)$, dla każdych $\beta \in W$, $c \in K$. Zatem ψ jest liniowe i $\psi \circ \phi = \mathrm{id}_V$ oraz $\phi \circ \psi = \mathrm{id}_W$. Stąd (1) $\Rightarrow$ (2).

Przechodzimy do implikacji (2) $\Rightarrow$ (1). Weźmy $\alpha, \beta \in V$ i niech $\phi(\alpha) = \phi(\beta)$. Wówczas

$$\alpha = \mathrm{id}_V(\alpha) = (\psi \circ \phi)(\alpha) = \psi(\phi(\alpha)) = \psi(\phi(\beta)) = (\psi \circ \phi)(\beta) = \mathrm{id}_V(\beta) = \beta.$$

Zatem ϕ jest różnowartościowe. Mamy $\phi \circ \psi = \mathrm{id}_W$, a więc dla każdego $\gamma \in W$ mamy

$$\gamma = \mathrm{id}_W(\gamma) = (\phi \circ \psi)(\gamma) = \phi(\psi(\gamma)).$$

A więc $\gamma = \phi(\psi(\gamma))$, czyli ϕ jest „na". □

Powyższy dowód pokazuje, że może być tylko jedno ψ spełniające warunek $(*)$. Co więcej, ψ to izomorfizm.

Definicja 11.1: Przekształcenie odwrotne

Jeśli dla przekształcenia liniowego $\phi : V \to W$ istnieje przekształcenie liniowe $\psi : W \to V$ spełniające $(*)$, to ψ nazywamy PRZEKSZTAŁCENIEM ODWROTNYM do ϕ i oznaczamy przez ϕ^{-1}.

Jak widzimy ϕ^{-1} istnieje wtedy i tylko wtedy, gdy ϕ jest izomorfizmem. W języku złożeń wysłowić można także własności monomorfizmów i epimorfizmów. Dowody tych własności wynikają bezpośrednio z dowodu wyżej (a osobno zapisane są też w skrypcie w postaci Wniosku 4.15 na str. 57).

Wniosek 11.1

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Wówczas

- ϕ jest monomorfizmem wtedy i tylko wtedy, gdy istnieje takie przekształcenie liniowe $\psi : W \to V$, że $\psi \circ \phi = \mathrm{id}_V$.
- ψ jest epimorfizmem wtedy i tylko wtedy, gdy istnieje takie przekształcenie liniowe $\psi : W \to V$, że $\phi \circ \psi = \mathrm{id}_W$.

Odczytajmy powyższe rezultaty w języku macierzy przekształceń liniowych. Przypomnijmy kluczowy rezultat wiążący składanie przekształceń i mnożenie macierzy.

Twierdzenie 11.2

Jeśli V, W, Z są przestrzeniami liniowymi nad K z bazami odpowiednio $\mathcal{A}, \mathcal{B}, \mathcal{C}$, oraz $\phi : V \to W$, $\psi : W \to Z$ są przekształceniami liniowymi, to: $M(\psi \circ \phi)^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}}$.

Załóżmy teraz, że ϕ jest izomorfizmem przestrzeni wymiaru n. Aby istniało przekształcenie ψ, które złożone z nim daje identyczność na przestrzeni V, zachodzić musi:

$$M(\mathrm{id})^{\mathcal{C}}_{\mathcal{A}} = M(\psi)^{\mathcal{C}}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}}.$$

W szczególności, jeśli $\mathcal{A} = \mathcal{C}$ otrzymujemy:

$$I_n = M(\mathrm{id})^{\mathcal{A}}_{\mathcal{A}} = M(\psi)^{\mathcal{A}}_{\mathcal{B}} \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}}.$$

Obydwie powyższe obserwacje prowadzą do ważnych wniosków i nowych definicji.

Definicja 11.2: Macierz odwrotna

Powiemy, że macierz $B \in M_{n\times n}(K)$ jest ODWROTNA do macierzy $A \in M_{n\times n}(K)$, jeśli

$$AB = BA = I_n.$$

Macierz odwrotną do macierzy A oznaczamy, o ile istnieje, jako A^{-1}. Macierz, która ma odwrotną nazywamy MACIERZĄ ODWRACALNĄ.

Uwaga 1. Dla każdej przestrzeni liniowej V wymiaru n oraz jej baz $\mathcal{X}$, $\mathcal{Y}$ mamy

$$M(\mathrm{id}_V)^{\mathcal{Y}}_{\mathcal{X}} \cdot M(\mathrm{id}_V)^{\mathcal{X}}_{\mathcal{Y}} = M(\mathrm{id}_V)^{\mathcal{Y}}_{\mathcal{Y}} = I_n = M(\mathrm{id}_V)^{\mathcal{X}}_{\mathcal{X}} = M(\mathrm{id}_V)^{\mathcal{X}}_{\mathcal{Y}} \cdot M(\mathrm{id}_V)^{\mathcal{Y}}_{\mathcal{X}}.$$

Uwaga 2. Jeśli $A_1, \ldots, A_k \in M_{n\times n}(K)$ są odwracalne, to również $A_1 \cdot \ldots \cdot A_k \in M_{n\times n}(K)$ jest odwracalna, bo na mocy łączności mnożenia macierzy:

$$(A_1 \cdot \ldots \cdot A_k) \cdot (A_k^{-1} \cdot \ldots \cdot A_1^{-1}) = (A_1 \cdot \ldots \cdot A_{k-1}) \cdot A_k \cdot A_k^{-1} \cdot (A_{k-1}^{-1} \cdot \ldots \cdot A_1^{-1}) = (A_1 \cdot \ldots \cdot A_{k-1}) \cdot (A_{k-1}^{-1} \cdot \ldots \cdot A_1^{-1}) = I_n.$$

Twierdzenie 11.3

Niech $\phi : K^n \to K^n$ będzie przekształceniem liniowym. Następujące warunki są równoważne:

(i) ϕ jest izomorfizmem,

(ii) macierz $M(\phi)_{st}^{st}$ jest odwracalna,

(iii) dla dow. baz $\mathcal{A}, \mathcal{B}$ przestrzeni K^n macierz $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ jest odwracalna.

Dowód. Jeśli ϕ jest izomorfizmem oraz $\psi = \phi^{-1}$, to biorąc $M(\psi)_{st}^{st}$ mamy:

$$M(\phi)_{st}^{st} \cdot M(\psi)_{st}^{st} = M(\phi \circ \psi)_{st}^{st} = M(\phi \circ \phi^{-1})_{st}^{st} = M(\mathrm{id})_{st}^{st} = I_n.$$

Analogicznie $M(\psi)_{st}^{st} \cdot M(\phi)_{st}^{st} = I_n$, co daje (i) $\Rightarrow$ (ii). Jeśli $A = M(\phi)_{st}^{st}$ jest odwracalna i $AB = I_n$, to niech $\psi : K^n \to K^n$ będzie zadane warunkiem $M(\psi)_{st}^{st} = B$. Wówczas

$$M(\phi \circ \psi)_{st}^{st} = M(\phi)_{st}^{st} \cdot M(\psi)_{st}^{st} = A \cdot B = I_n = M(id)_{st}^{st}.$$

Zatem $\phi \circ \psi = \mathrm{id}$. Analogicznie z $BA = I_n$ mamy $\psi \circ \phi = \mathrm{id}$. Zatem ϕ to izomorfizm i mamy (ii) $\Rightarrow$ (i).

Równoważność (i) oraz (ii) implikuje, że macierz odwrotna jest jednoznacznie wyznaczona, jeśli istnieje.

Implikacja (iii) $\Rightarrow$ (ii) jest oczywista. Implikacja odwrotna wynika z rozkładu $M(\phi)_{\mathcal{A}}^{\mathcal{B}} = M(\mathrm{id})_{st}^{\mathcal{B}} \cdot M(\phi)_{st}^{st} \cdot M(\mathrm{id})_{\mathcal{A}}^{st}$. Macierz $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ jest więc odwracalna, jako iloczyn macierzy odwracalnych. Stąd (ii) $\Rightarrow$ (iii). $\square$

Wniosek 11.2

Jeśli $A, B \in M_{n\times n}(K)$ spełniają warunek $AB = I_n$, to $B = A^{-1}$.

W ostatnim wniosku kluczowe jest założenie, że A, B są rozmiaru $n \times n$. Oczywiście mamy $\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = [1]$, ale żaden z czynników nie jest macierzą odwracalną.

Znając związek pomiędzy izomorfizmami a macierzami odwracalnymi przechodzimy do dwóch zagadnień.

1. Opis wszystkich macierzy izomorfizmów przestrzeni n-wymiarowej.
2. Wyznaczanie macierzy odwrotnej do danej (o ile to możliwe).

Definicja 11.3

Niech $\mathcal{A}, \mathcal{B}$ będą bazami przestrzeni V. Macierz $M(\mathrm{id}_V)_{\mathcal{A}}^{\mathcal{B}}$ nazywamy MACIERZĄ ZAMIANY (TRANSFORMACJI) WSPÓŁRZĘDNYCH z $\mathcal{A}$ do $\mathcal{B}$.

Twierdzenie 11.4

Niech $A \in M_{n\times n}(K)$. Następujące warunki są równoważne:

(i) A jest macierzą zamiany współrzędnych w K^n,

(ii) A jest macierzą odwracalną,

(iii) przekształcenie liniowe $\phi : K^n \to K^n$ zadane warunkiem $M(\phi)_{st}^{st} = A$ jest izomorfizmem,

(iv) $r(A) = n$.

Dowód. Równoważność warunków (ii) oraz (iii) pokazaliśmy wyżej. Mamy $r(A) = \dim \operatorname{im} \phi = n$, więc w sposób oczywisty (iii) jest równoważne (iv) (na mocy równości $n = \dim \ker \phi + \dim \operatorname{im} \phi$). Również implikacja (i) $\Rightarrow$ (iii) została uzasadniona wyżej. Pozostaje więc wykazać (iii) $\Rightarrow$ (i). To jest jednak jasne, bowiem jeśli przez $\mathcal{A}$ oznaczymy zbiór kolumn macierzy A, to $\mathcal{A}$ jest bazą K^n (izomorfizm przeprowadza bazę na bazę). Oznacza to, że $A = M(\mathrm{id})^{st}_{\mathcal{A}}$, co kończy dowód. □

Wyznaczanie macierzy odwrotnej do macierzy $A \in M_{n\times n}(K)$ można przeprowadzić, startując od macierzy, której pierwsze n kolumn to kolejne kolumny macierzy A, a kolejne n kolumn to kolejne kolumny macierzy I_n. Jeśli za pomocą elementarnych operacji wierszowych sprowadzimy taką macierz do postaci, w której pierwsze n kolumn to kolejne kolumny macierzy I_n, to n kolejnych kolumn powstałej macierzy to kolejne kolumny macierzy A^{-1}. Schematycznie algorytm przedstawia się następująco:

$$\left[\begin{array}{c|c} A & I_n \end{array}\right] \longrightarrow \left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right].$$

Uzasadnienie: rozważmy równanie

$$AX = I_n,$$

gdzie $A \in M_{n\times n}(K)$ jest dana natomiast $X \in M_{n\times n}(K)$ – szukana. Wówczas i-ta kolumna macierzy X jest rozwiązaniem układu równań o macierzy rozszerzonej

$$[A \,|\, \epsilon_i],$$

gdzie ϵ_i to i-ty wektor bazy standardowej K^n. Innymi słowy algorytm opisany wyżej jest w istocie algorytmem jednoczesnego rozwiązania n układów równań takich, jak wyżej.

Przykład. Wyznaczmy macierz odwrotną do macierzy

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \in M_{2\times 2}(\mathbb{R}).$$

Mamy

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -6 & -3 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} \end{array}\right].$$

Rzeczywiście więc:

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{6} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

* * *

Skoro mówimy już o operacjach elementarnych to czas uzyskać świadomość, że wykonywanie operacji elementarnych również wiąże się z mnożeniem macierzy. Pomoże to nam uzyskać drugie kryterium odwracalności (a także kolejny efektywny algorytm odwracania) macierzy. Rozważmy pewne przykłady.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \end{bmatrix} \quad , \quad \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b & a & c \\ e & d & f \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 0 & x \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a & b & c \\ xd & xe & xf \end{bmatrix} \quad , \quad \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & xb & c \\ d & xe & f \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a & b & c \\ d+a & e+b & f+c \end{bmatrix} \quad , \quad \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b & c+a \\ d & e & f+d \end{bmatrix}.$$

Widzimy, że **przemnażanie z lewej strony** macierzy o wyrazach a, b, c, d, e, f przez pewne macierze dokonuje na niej odpowiedniej operacji elementarnej na wierszach. Podobnie można, przez **przemnażanie z prawej strony**, uzyskać analogiczne operacje na kolumnach (za pomocą tych samych macierzy).

Powyższe obserwacje prowadzą do następującej definicji.

Definicja 11.4

Niech n, i, j będą liczbami naturalnymi spełniającymi $1 \leqslant i, j \leqslant n$, $i \neq j$ i niech a, c będą elementami ciała K, przy czym $c \neq 0$. Definiujemy następujące macierze $E_{ij}^n(a), T_{ij}^n, I_i^n(c)$ należące do $M_{n\times n}(K)$:

- $E_{ij}^n(a) = [a_{st}] \in M_{n\times n}(K)$, gdzie

$$a_{st} = \begin{cases} a, & \text{gdy } s = i, t = j \\ 1, & \text{gdy } s = t \\ 0, & \text{w pozostałych przypadkach,} \end{cases}$$

- $T_{ij}^n = [a_{st}] \in M_{n\times n}(K)$, gdzie

$$a_{st} = \begin{cases} 1, & \text{gdy } s = t \neq i, j \\ 1, & \text{gdy } s = i, t = j \text{ lub } s = j, t = i \\ 0, & \text{w pozostałych przypadkach,} \end{cases}.$$

- $I_i^n(c) = [a_{st}] \in M_{n\times n}(K)$, gdzie

$$a_{st} = \begin{cases} c, & \text{gdy } s = t = i \\ 1, & \text{gdy } s = t \neq i \\ 0, & \text{w pozostałych przypadkach.} \end{cases},$$

Macierze $E_{ij}^n(a), T_{ij}^n, I_i^n(c)$ nazywamy **macierzami operacji elementarnych**.

Przykłady (a dalej ogólny fakt będący łatwym ćwiczeniem):

$$E_{24}^5(a) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & a & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_{35}^5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad I_1^5(c) = \begin{bmatrix} c & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Obserwacja 11.1

Dla każdej macierzy $A \in M_{m\times n}(K)$ macierz

- $E_{ij}^m(x) \cdot A$ powstaje z A przez dodanie do i-tego wiersza j-tego wiersza pomnożonego przez x,
- $A \cdot E_{ij}^n(x)$ powstaje z A przez dodanie do j-tej kolumny i-tej kolumny pomnożonej przez x,
- $T_{ij}^m \cdot A$ powstaje z A przez przestawienie i-tego i j-tego wiersza,
- $A \cdot T_{ij}^n$ powstaje z A przez przestawienie i-tej i j-tej kolumny,
- $I_i^m(y) \cdot A$ powstaje z A przez pomnożenie i-tego wiersza przez y,
- $A \cdot I_i^n(y)$ powstaje z A przez pomnożenie i-tej kolumny przez y.

Wniosek 11.3

Dla każdej macierzy $A \in M_{m\times n}(K)$ istnieje macierz

- $P \in M_{m\times m}(K)$, będąca iloczynem macierzy typu $E_{ij}^m(x)$, T_{ij}^m, że PA jest schodkowa,
- $Q \in M_{m\times m}(K)$, będąca iloczynem macierzy typu $E_{ij}^m(x), T_{ij}, I_i(y)$, że QA jest zredukowana.

Dowód pozostawiamy Czytelnikowi – wynika on natychmiast z tego, że każdą macierz można sprowadzić do postaci schodkowej i schodkowej zredukowanej operacjami odpowiedniego typu.

Przykład: postacią zredukowaną macierzy

$$A = \begin{bmatrix} 6 & 6 & -2 \\ -1 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

jest macierz jednostkowa I_3, dokładniej:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{6} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -6 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = I_3.$$

Zauważmy, że wyznaczyliśmy w istocie macierz odwrotną A^{-1} do macierzy A. Jest to iloczyn macierzy operacji elementarnych przeprowadzających macierzy A w swoją postać zredukowaną I_3. Wyjaśnienie wynika z następującej obserwacji, formalizującej znaną nam intuicję – działanie operacji elementarnej można „odwrócić" za pomocą operacji tego samego typu.

Obserwacja 11.2

Dla każdej macierzy $S \in M_{n\times n}(K)$ jednego z typów $E^n_{ij}(a), T^n_{ij}, I^n_i(c)$ istnieje macierz S' tego samego typu taka, że:

$$S'S = SS' = I.$$

Dowód. Istotnie, łatwo sprawdzić, że

- jeśli $S = E^n_{ij}(a)$, to $S' = E^n_{ij}(-a)$,
- jeśli $S = T^n_{ij}$, to $S' = T^n_{ij}$,
- jeśli $S = I^n_i(c)$, to $S' = I^n_i(c^{-1})$

□

W wielu zastosowaniach, które poznamy w przyszłym semestrze, wygodnie jest wykonywać operacje wierszowe i analogiczne operacje kolumnowe na macierzy mającej tyle samo wierszy i kolumn. Jeśli, dla przykładu, na macierzy A rozmiaru 3×3 wykonać chcemy jednocześnie operację dodania do drugiego wiersza pierwszego wiersza przemnożonego przez 5 oraz dodanie do drugiej kolumny pierwszej kolumny przemnożonej przez 5, to wykonujemy iloczyn

$$E^3_{21}(5) \cdot A \cdot E^3_{12}(5) = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot A \cdot \begin{bmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Macierze $E^3_{21}(5)$ i $E^3_{12}(5)$ powstają przez zamianę wierszy na kolumny. Motywuje to następującą definicję.

Definicja 11.5: Macierz transponowana

Macierzą transponowaną macierzy $A \in M_{m\times n}(K)$ nazywamy macierz $A^T \in M_{n\times m}(K)$, której kolejne kolumny są kolejnymi wierszami macierzy A. Jeśli macierz A spełnia warunek $A = A^T$, to nazywamy ją macierzą symetryczną.

Przykłady:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^T = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & 4 \end{bmatrix}.$$

Obserwacja 11.3

Dla macierzy $A \in M_{m\times n}(K)$ mamy:

$$(A^T)^T = A, \quad r(A) = r(A^T), \quad (AB)^T = B^T A^T, \quad (A^T)^{-1} = (A^{-1})^T,$$

jeśli A posiada macierz odwrotną.

Zauważmy, że macierze T_{ij}^n oraz $I_i^n(c)$ są symetryczne. Natomiast $E_{ij}^n(a)^T = E_{ji}^n(a)$.

Wniosek 11.4

Niech $A' \in M_{n\times n}(K)$ będzie macierzą otrzymaną z A przez sprowadzenie do zredukowanej postaci schodkowej za pomocą elementarnych operacji na wierszach. Wówczas następujące warunki są równoważne:

- macierz A jest odwracalna
- $A' = I$.

W szczególności następujące warunki są równoważne:

- macierz A jest odwracalna,
- A jest iloczynem macierzy typu $E_{ij}(x)$, T_{ij}, $I_i(y)$, gdzie $y \neq 0$.

Wskazówka. Niech $A' = W_r \cdot \ldots \cdot W_1 \cdot A$, gdzie W_i – macierze operacji elementarnych. Wyznacz A.

Wniosek 11.5

Jeśli $A \in M_{m\times n}(K), B \in M_{m\times m}(K)$ oraz $C \in M_{n\times n}(K)$, przy czym $r(B) = m$ i $r(C) = n$, to:

$$r(BAC) = r(A).$$

Wniosek ten jest oczywiście jasny z punktu widzenia teorii przekształceń liniowych. Macierze B oraz C są macierzami izomorfizmów, powiedzmy $\phi_B : \mathbb{R}^m \to \mathbb{R}^m$, $\phi_C : \mathbb{R}^n \to \mathbb{R}^n$, stąd wymiar obrazu przekształcenia $\psi_A : \mathbb{R}^n \to \mathbb{R}^m$ zadanego macierzą A jest taki sam jak wymiar obrazu złożenia $\phi_B \circ \psi_A \circ \phi_C$.

* * *

Powiemy na koniec o strukturze przestrzeni przekształceń liniowych pomiędzy ustalonymi przestrzeniami liniowymi V, W nad ciałem K. Na poprzednich wykładach powiedzieliśmy sporo o złożeniach przekształceń liniowych i ich interpretacji w języku mnożenia macierzy. Określmy operacje dodawania przekształceń i mnożenia ich przez skalary.

Definicja 11.6: Działania na przekształceniach liniowych

Niech V, W będą przestrzeniami liniowymi nad K i niech $\phi, \psi : V \to W$ będą przekształceniami liniowymi.

- Sumą ϕ i ψ nazywamy odwzorowanie $\phi + \psi : V \to W$ zadane wzorem:

$$(\phi + \psi)(v) = \phi(v) + \psi(v), \text{ dla każdego } v \in V,$$

- Iloczynem ϕ przez skalar $a \in K$ nazywamy odwzorowanie $a \cdot \phi : V \to W$ postaci:

$$(a \cdot \phi)(v) = a \cdot \phi(v), \text{ dla każdego } v \in V.$$

Przykład: jeśli $V = \mathbb{R}^3$ oraz $W = \mathbb{R}^2$, to dla przekształceń liniowych $f, g : V \to W$ zadanych wzorami:

$$f((x_1, x_2, x_3)) = (x_1 + x_3, x_1 - x_2 + x_3), \quad g((x_1, x_2, x_3)) = (0, 2x_1)$$

mamy:

$$(f + g)((x_1, x_2, x_3)) = (x_1 + x_3, 3x_1 - x_2 + x_3), \quad (2 \cdot g)((x_1, x_2, x_3)) = (0, 4x_1).$$

Oczywiście jeśli $\phi, \psi : V \to W$ są przekształceniami liniowymi przestrzeni liniowych nad ciałem K oraz jeśli $a \in K$, to funkcje $\phi + \psi$ oraz $a \cdot \phi$ traktowane jako elementy $F(V, W)$ są przekształceniami liniowymi.

Definicja 11.7: Przestrzeń przekształceń liniowych ustalonych przestrzeni liniowych

Niech V, W będą przestrzeniami liniowymi nad K. Przestrzeń liniową wszystkich przekształceń liniowych $\phi : V \to W$ będziemy oznaczać symbolem[a] $L(V, W)$. Zerem tej przestrzeni liniowej jest przekształcenie zerowe.

[a] Często stosuje się także ogólniejszą notację: $\mathrm{Hom}(V, W)$.

W przypadku przestrzeni V, W skończonego wymiaru opis przestrzeni $L(V, W)$ jest nietrudny. Zobaczmy najpierw przykład. Jeśli rozważymy przestrzeń $L(K^3, K^2)$, to jej wymiar wynosi 6, ponieważ bez trudu jesteśmy w stanie wskazać bazę tej przetrzeni. Jest ona złożona z sześciu przekształceń liniowych postaci:

$$\phi_1((x_1, x_2, x_3)) = (x_1, 0), \quad \phi((x_1, x_2, x_3)) = (x_2, 0), \quad \phi_3((x_1, x_2, x_3)) = (x_3, 0),$$

$$\phi_4((x_1, x_2, x_3)) = (0, x_1), \quad \phi((x_1, x_2, x_3)) = (0, x_2), \quad \phi_3((x_1, x_2, x_3)) = (0, x_3).$$

Funkcje te tworzą układ liniowo niezależny. Rzeczywiście, gdyby dla pewnych $a_1, \ldots, a_6$ funkcja

$$a_1\phi_1 + a_2\phi_2 + \ldots + a_6\phi_6$$

była tożsamościowo równa zero (czyli $(0, 0)$), to w szczególności przyjmowałaby ona taki wektor dla każdego wektora bazy standardowej $\epsilon_1 = (1, 0, 0), \epsilon_2 = (0, 1, 0), \epsilon_3 = (0, 0, 1)$. Mielibyśmy zatem:

$$(a_1\phi_1 + a_2\phi_2 + \ldots + a_6\phi_6)(1, 0, 0) = a_1(1, 0) + a_4(0, 1) = (0, 0),$$

czyli $a_1 = a_4 = 0$. Analogicznie dowodzimy, że $a_2 = a_5 = 0$ oraz $a_3 = a_6 = 0$.

Wiemy, że każde przekształcenie liniowe $\phi : K^3 \to K^2$ jest postaci:

$$\phi((x_1, x_2, x_3) = (a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_1 + a_{22}x_2 + a_{23}x_3).$$

Zatem:

$$\phi = a_{11}\phi_1 + a_{12}\phi_2 + a_{13}\phi_3 + a_{21}\phi_4 + a_{22}\phi_5 + a_{23}\phi_6.$$

A zatem funkcje te rozpinają $L(K^3, K^2)$.

Twierdzenie 11.5

Niech V, W będą skończenie wymiarowymi przestrzeniami liniowymi, przy czym $\dim V = n$ oraz $\dim W = m$. Ma miejsce izomorfizm przestrzeni liniowych:

$$L(V, W) \simeq M_{m \times n}(K),$$

Dokładniej, biorąc dowolne ustalone bazy $\mathcal{A}$ przestrzeni V oraz $\mathcal{B}$ przestrzeni W możemy sformułować izomorfizm przestrzeni liniowych $\Psi : L(V, W) \longrightarrow M_{m \times n}(K)$ dany wzorem

$$\Psi(\phi) = M(\phi)^{\mathcal{B}}_{\mathcal{A}}.$$

W szczególnośći $\dim L(V, W) = m \cdot n$.

Dowód. Zauważmy najpierw, że Ψ jest przekształceniem liniowym. Weźmy dwa przekształcenia liniowe $\phi, \psi : V \to W$. Niech $\mathcal{A} = (\alpha_1, \ldots, \alpha_n)$ oraz $\mathcal{B} = (\beta_1, \ldots, \beta_m)$. Załóżmy, że istnieją $a_{1j}, \ldots, a_{mj}$ oraz $b_{1j}, \ldots, b_{mj}$ takie, że:

$$\phi(\alpha_j) = a_{1j}\beta_1 + \ldots + a_{mj}\beta_m, \quad \psi(\alpha_j) = b_{1j}\beta_1 + \ldots + b_{mj}\beta_m.$$

Wówczas

$$(\phi+\psi)(\alpha_j) = \phi(\alpha_j)+\psi(\alpha_j) = a_{1j}\beta_1+\ldots+a_{mj}\beta_m+b_{1j}\beta_1+\ldots+b_{mj}\beta_m = (a_{1j}+b_{1j})\beta_1+\ldots+(a_{mj}+b_{mj})\beta_m.$$

A zatem współrzędna wektora $(\phi + \psi)(\alpha_j)$ w bazie $\mathcal{B}$ przy wektorze β_i to $a_{ij} + b_{ij}$. Krótko mówiąc

$$\Psi(\phi + \psi) = M(\phi + \psi)^{\mathcal{B}}_{\mathcal{A}} = M(\phi)^{\mathcal{B}}_{\mathcal{A}} + M(\psi)^{\mathcal{B}}_{\mathcal{A}} = \Psi(\phi) + \Psi(\psi).$$

Analogicznie dowodzimy, że $M(\lambda\phi)^{\mathcal{B}}_{\mathcal{A}} = \lambda \cdot M(\phi)^{\mathcal{B}}_{\mathcal{A}}$, dla $\lambda \in K$. A zatem Ψ jest przekształceniem liniowym.

Pozostaje pokazać, że Ψ jest izomorfizmem. Oczywiście Ψ jest różnowartościowe, bo każde przekształcenie liniowe jest jednoznacznie określone na bazie (zaś kolumny $M(\phi)^{\mathcal{B}}_{\mathcal{A}}$ to obrazy wektorów z bazy $\mathcal{A}$).

Odwzorowanie Ψ to również suriekcja: dla dowolnej macierzy $X = [x_{ij}] \in M_{m\times n}(K)$ można określić przekształcenie liniowe $\phi \in L(V, W)$, zadane na bazie $\mathcal{A}$ (w sposób jednoznaczny) warunkiem

$$\alpha_j \mapsto x_{1j}\beta_1 + \ldots + x_{mj}\beta_m.$$

Stąd Ψ jest liniową bijekcją, czyli izomorfizmem przestrzeni liniowych. □

Zwróćmy uwagę, że wiele naturalnych funkcji reprezentować będziemy w postaci macierzy. Są wśród nich choćby funkcje permutujące kolejność współrzędnych. Przekształcenie liniowe $\phi : K^n \to K^n$ zadane wzorem

$$\phi((x_1, x_2, \ldots, x_n)) = (x_2, x_3, \ldots, x_n, x_1)$$

przechodzi przy powyższym izomorfizmie (dla $\mathcal{A} = \mathcal{B} = st$) na macierz

$$M(\phi)^{st}_{st} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

W drugim semestrze szczególnie istotne będzie dla nas badanie przestrzeni $L(V, V)$, gdzie V jest przestrzenią skończonego wymiaru. Przestrzeń ta, izomorficzna z przestrzenią liniową $M_{n\times n}(K)$ wyposażona jest w dodatkowe działanie składania przekształceń, które czyni ją tak zwaną **algebrą łączną**. Nie wchodzimy tu w szczegóły. Sygnalizujemy jednak, że napotkaliśmy ważny przykład przestrzeni liniowej, w której w sposób sensowny można wprowadzić łączne mnożenie wektorów. Mnożenie to, podobnie jak mnożenie macierzy, nie jest przemienne i własnościami algebraicznymi odbiega zdecydowanie od mnożenia w ciele.

11.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Znajdowanie macierzy odwrotnej)
   Dla każdej z poniższych macierzy znajdź macierz odwrotną:
$$A_1 = \begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 9 & 8 \\ 2 & 7 & 8 \\ 1 & 3 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 5 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 7 & 9 & 3 & 4 \\ 4 & 6 & 2 & 3 \end{bmatrix}$$

2. Dla każdej z poniższych macierzy rozmiaru $n \times n$ znajdź macierz odwrotną:
$$\begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

3. (♠ Rozstrzyganie czy złożenie jest monomorfizmem/epimorfizmem/izomorfizmem)
   - Przekształcenia liniowe $\phi : \mathbb{R}^4 \longrightarrow \mathbb{R}^5$ i $\psi : \mathbb{R}^5 \longrightarrow \mathbb{R}^4$ spełniają $r(\phi) = 4 = r(\psi)$. Czy wynika stąd, że przekształcenie $\psi \circ \phi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ jest izomorfizmem?
   - Niech $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ oraz $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ będą przekształceniami liniowymi. Czy przekształcenie liniowe $\psi \circ \phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ będące ich złożeniem może być monomorfizmem?
   - Czy istnieją przekształcenia liniowe $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$, $g : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ takie, że zachodzi równość $f \circ g = \mathrm{id}_{\mathbb{R}^3}$?
   - Niech:
$$M(\phi)^{\mathcal{B}}_{st} = \begin{bmatrix} 2 & 3 & 1 & 4 \\ 1 & 2 & 1 & 3 \end{bmatrix}, \quad M(\psi_t)^{st}_{st} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ -1 & 1 \\ 1 & t \end{bmatrix},$$
     gdzie $\mathcal{B} = ((0,1),(1,-1))$. Dla jakich t złożenie $\phi \circ \psi_t$ jest izomorfizmem?

4. Niech $A \in M_{4\times 2}(\mathbb{R})$ oraz $B \in M_{2\times 4}(\mathbb{R})$ będą takimi macierzami, że:
$$AB = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$
   Znajdź macierz BA (wskazówka: podziel macierz AB na cztery bloki).

5. Przedstaw macierz $A = \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix} \in M_2(\mathbb{R})$ jako iloczyn macierzy operacji elementarnych.

6. Niech $\phi : V \to V$ będzie takim przekształceniem liniowym, ze $\phi \circ \phi = \phi$. Wykaż, że istnieją takie podprzestrzenie V_1, V_2 w V, że ϕ jest rzutem na V_1 wzdłuż V_2.

7. Niech $\phi : V \to V$ będzie takim przekształceniem liniowym, że $\phi \circ \phi = \mathrm{id}$, gdzie V jest przestrzenią liniową nad ciałem K, w którym $1 + 1 \neq 0$ (np. $K = \mathbb{R}$ lub $\mathbb{C}$ lub $\mathbb{Q}$). Wykaż, że istnieją takie podprzestrzenie V_1, V_2 w V, że ϕ jest symetrią względem V_1 wzdłuż V_2.

8. Niech V, W będą przestrzeniami liniowymi, zaś niech $\phi \in L(V,W)$ oraz $\psi \in L(W,W)$ spełniają $\psi \circ \phi = 0$ oraz $\mathrm{im}(\mathrm{id}_W - \psi) \subseteq \mathrm{im}(\phi)$. Wykaż, że $W = \mathrm{im}\,\phi \oplus \mathrm{im}\,g$.

9. Niech V będzie skończenie wymiarową przestrzenią liniową. Załóżmy, że $\phi, \psi \in L(V,V)$ spełniają $\phi \circ \psi = \psi \circ \phi$, oraz że $\phi - \psi$ jest monomorfizmem. Wykaż, że $\ker(\phi \circ \psi) = \ker\phi \oplus \ker\psi$.

10. Dane są przestrzenie liniowe V, W nad ciałem K, przy czym $\dim V = n$, $\dim W = 2$. Dane jest również $\phi \in L(V,V)$. Definiujemy funkcję $\Phi : L(V,W) \to L(V,W)$ wzorem $\Phi(\psi) = \psi \circ \phi$. Wykaż, że Ψ jest przekształceniem liniowym. Czy jeśli ϕ jest izomorfizmem, to również Φ jest izomorfizmem?

11.3 Uzupełnienie. Funkcjonały i przestrzeń sprzężona

W tym uzupełnieniu omówimy szczególnie ważny typ przestrzeni przekształceń, mający fundamentalne znaczenie w algebrze liniowej i wielu innych dziedzinach matematyki.

Definicja 11.8: Funkcjonał, przestrzeń sprzężona

FUNKCJONAŁEM LINIOWYM (albo FORMĄ LINIOWĄ) na przestrzeni liniowej V nad ciałem K nazywamy przekształcenie liniowe $\phi : V \to K$. Zbiór

$$V^* = L(V, K)$$

funkcjonałów liniowych na przestrzeni V nazywamy PRZESTRZENIA SPRZĘŻONĄ (DUALNĄ) do V.

Przykłady.

- Wiemy już z wcześniejszych wykładów, że dla każdego elementu $\phi \in (K^n)^*$ istnieją $a_1, \ldots, a_n \in K$ takie, że ϕ zadana jest wzorem

$$\phi((x_1, \ldots, x_n)) = a_1x_1 + \ldots + a_nx_n.$$

- Przekształcenie $\mathrm{tr} \in (M_{n\times n}(\mathbb{R}))^*$ zwane ŚLADEM, zadane wzorem

$$\mathrm{tr}([a_{ij}]) = a_{11} + \ldots + a_{nn}.$$

- Niech X będzie niepustym zbiorem, K – ciałem oraz $F(X, K)$ – przestrzenią funkcji z W do K. Niech $x_0 \in X$. Wówczas odwzorowanie $\mathrm{ev}_0 : F(X, K) \to K$ zadane wzorem:

$$\mathrm{ev}_0(f) = f(0)$$

  jest funkcjonałem liniowym na przestrzeni $F(X, K)$. Ten niezwykle istotny funkcjonał nazywany EWALUACJĄ W PUNKCIE 0.

Obserwacja 11.4

Jeśli V jest przestrzenią skończenie wymiarową, to $V \simeq V^*$.

Dowód. Niech $\dim V = n$. Wówczas $V^* = L(V, K)$ jest również wymiaru n, jako przestrzeń izomorficzna z $M_{1\times n}(K)$. Dwie przestrzenie tego samego, skończonego wymiaru, są izomorficzne. □

Obserwacja 11.5

Niech $\mathcal{A} = (v_1, \ldots, v_n)$ będzie bazą przestrzeni V i niech $f_i : V \to K$ będzie jedynym funkcjonałem liniowym takim, że:

$$f_i(v_j) = \begin{cases} 1, & \text{jeśli } i = j \\ 0, & \text{jeśli } i \neq j. \end{cases} \qquad (\star)$$

Wówczas:

(a) $v = f_1(v)v_1 + f_2(v)v_2 + \ldots + f_n(v)v_n$, czyli f_i przyporządkowuje wektorowi jego i-tą współrzędną w bazie $\mathcal{A}$,

(b) dla dowolnego $f \in V^*$ mamy $f = f(v_1)f_1 + f(v_2)f_2 + \ldots + f(v_n)f_n$, i jest to przedstawienie jednoznaczne,

(c) układ funkcjonałów $\mathcal{A}^* = (f_1, \ldots, f_n)$ jest bazą V^* i wartość funkcjonału $f \in V^*$ na wektorze v_j jest j-tą współrzędną tego funkcjonału w bazie $\mathcal{A}^*$.

Przykłady.

- Dla bazy przestrzeni $\mathbb{R}^3$ postaci

$$v_1 = (1,1,1), \quad v_2 = (1,1,0), \quad v_3 = (1,0,0)$$

układ funkcjonałów f_i określony warunkami wyżej istnieje i ma postać:

$$f_1((x_1,x_2,x_3)) = x_3, \quad f_2((x_1,x_2,x_3)) = x_2 - x_3, \quad f_3((x_1,x_2,x_3)) = x_1 - x_2.$$

Na przykład dla $\alpha = (10,5,2)$ otrzymujemy:

$$f_1(\alpha) = 2, f_2(\alpha) = 3, f_3(\alpha) = 5 \quad \text{oraz} \quad (10,5,2) = 2(1,1,1) + 3(1,1,0) + 5(1,0,0).$$

- **Przykład.** Współrzędne funkcjonału $\phi \in (\mathbb{R}^3)^*$, gdzie

$$\phi((x_1,x_2,x_3)) = 2x_1 - 9x_2 + 5x_3$$

w bazie sprzężonej do $\mathcal{A} = ((1,1,1),(5,1,1),(1,1,3))$ wynoszą:

$$\phi((1,1,1)) = -2, \quad \phi((5,1,1)) = 6, \quad \phi((1,1,3)) = 8.$$

Dowód. Pierwszy punkt jest oczywisty. Istotnie, jeśli $v = a_1v_1 + \ldots + a_nv_n$, to obkładając tę równość z obydwu stron funkcjonałem f_i dostajemy: $f_i(v) = a_1f_i(v_1) + a_2f_i(v_2) + \ldots + a_nf_i(v_n)$. Tylko element $f_i(v_i)$ sumy po prawej jest niezerowy, z definicji f_i. A zatem $f_i(v) = a_i$.

Punkty (b) i (c) postulują, że $(f_1,\ldots,f_n)$ jest bazą V^*. Sprawdźmy najpierw, że układ ten jest liniowo niezależny. Załóżmy, że istnieją takie $a_1,\ldots,a_n \in K$, że

$$a_1f_1 + \ldots + a_nf_n = 0,$$

przy czym 0 po prawej stronie intepretujemy jako funkcjonał zerowy! A zatem $a_1f_1 + \ldots + a_nf_n$ jest funkcjonałem, który dowolny wektor posyła na zero. Z drugiej strony biorąc element v_i bazy $\mathcal{A}$ mamy:

$$(a_1f_1 + \ldots + a_nf_n)(v_i) = a_1f_1(v_i) + \ldots + a_nf_n(v_i) = a_i.$$

A zatem $a_i = 0$, dla każdego $1 \leqslant i \leqslant n$. A zatem $(f_1,\ldots,f_n)$ jest układem liniowo niezależnym.

Zobaczmy teraz, że $(f_1,\ldots,f_n)$ rozpina V^*. Niech $f \in V^*$. Twierdzimy, że:

$$f = f(v_1)f_1 + f(v_2)f_2 + \ldots + f(v_n)f_n.$$

Aby stwierdzić czy dwa przekształcenia są identyczne wystarczy to sprawdzić na dowolnej bazie V, na przykład na $(v_1,\ldots,v_n)$. Wówczas rzeczywiście:

$$(f(v_1)f_1 + f(v_2)f_2 + \ldots + f(v_n)f_n)(v_i) = f(v_1)f_1(v_i) + \ldots + f(v_n)f_n(v_i) = f(v_i)\cdot 1.$$

□

Definicja 11.9: Baza dualna

Bazę $\mathcal{A}^*$ zdefiniowaną wyżej wzorem $(\star)$ nazywamy BAZĄ DUALNĄ (SPRZĘŻONĄ) do bazy $\mathcal{A}$.

Przykład. Niech

$$\alpha_1 = (1,3), \alpha_2 = (2,7)$$

będzie bazą przestrzeni $\mathbb{R}^2$. Weźmy

$$\alpha_1^*(x_1,x_2) = 7x_1 - 2x_2 \quad \text{oraz} \quad \alpha_2^*(x_1,x_2) = -3x_1 + 1x_2.$$

Wówczas, jak we wzorze $(\star)$ mamy: $\alpha_1^*(\alpha_1) = 1$, $\alpha_1^*(\alpha_2) = 0$, $\alpha_2^*(\alpha_1) = 0$, $\alpha_2^*(\alpha_2) = 1$. Zauważmy, że jeśli wpiszemy współczynniki funkcjonałów w macierz (w kolumny), a wektory z wyjściowej bazy wpiszemy w wiersze macierzy, dostaniemy zależność:

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Problem wyznaczania bazy dualnej jest problemem rozwiązania układu równań danego warunkami z $(\star)$. Macierze: mająca w wierszach wektory z $\mathcal{A}$ oraz: mająca w kolumnach wektory z $\mathcal{A}^*$ są **odwrotne**.

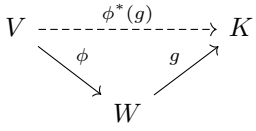
Bazą sprzężoną do bazy standardowej przestrzeni K^n jest baza złożona z funkcjonałów postaci

$$f_i(x_1, \ldots, x_n) = x_i,$$

dla $1 \leqslant i \leqslant n$. Bazę tę oznaczać będziemy przez st^*.

Definicja 11.10: Przekształcenie sprzężone

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Przekształceniem sprzężonym do ϕ nazywamy przekształcenie $\phi^* : W^* \to V^*$ określone wzorem $\phi^*(g) = g \circ \phi$. Innymi słowy jest to takie przekształcenie, które bierze funkcjonał g z W^* i przeprowadza go na funkcjonał $\phi^*(g) : V \to K$ tak, że następujący diagram jest przemienny dla każdego $g \in W^*$:



Przykład przekształcenia sprzężonego. Rozważmy $\psi : \mathbb{R}^3 \to \mathbb{R}^2$ dane wzorem:

$$\psi((x_1, x_2, x_3)) = (2x_1 + 3x_2 + x_3, 5x_1 - x_2 - 2x_3)).$$

Wówczas dla funkcjonału $f : \mathbb{R}^2 \to \mathbb{R}$ zadanego wzorem:

$$f((y_1, y_2)) = 3y_1 - 2y_2$$

funkcjonał $\psi^*(f) : \mathbb{R}^3 \to \mathbb{R}$ jest zadany wzorem:

$$\begin{aligned} \psi^*(f)((x_1, x_2, x_3)) &= (f \circ \psi)((x_1, x_2, x_3)) = f(\psi((x_1, x_2, x_3)) = \\ &= f((2x_1 + 3x_2 + x_3, 5x_1 - x_2 - 2x_3)) = \\ &= 3(2x_1 + 3x_2 + x_3) - 2(5x_1 - x_2 - 2x_3) = \\ &= -4x_1 + 11x_2 + 7x_3. \end{aligned}$$

A jak wygląda wzór przekształcenia ψ^*? Jak je zapisać? Można np. $\psi^*(y_1\epsilon_1^* + y_2\epsilon_2^*) = a_1\epsilon_1^* + a_2\epsilon_2^* + a_3\epsilon_3^*$

Twierdzenie 11.6

Niech $\mathcal{A}, \mathcal{B}$ będą bazami przestrzeni V oraz W. Niech $\phi : V \to W$. Wówczas:

$$M(\phi)_{\mathcal{A}}^{\mathcal{B}} = (M(\phi^*)_{\mathcal{B}^*}^{\mathcal{A}^*})^T,$$

gdzie $\mathcal{A} = (v_1, \ldots, v_n)$ jest bazą V oraz $\mathcal{B} = (w_1, \ldots, w_m)$ jest bazą W.

Dowód. Niech $(v_1^*, \ldots, v_n^*)$ będzie bazą dualną do $\mathcal{A}$ oraz $(w_1^*, \ldots, w_m^*)$ będzie bazą dualną do $\mathcal{B}$. Z definicji macierzy przekształcenia liniowego mamy, że i-ty wyraz j tej kolumny macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ jest i-tą współrzędną wektora $\phi(v_j)$ w bazie $\mathcal{B}$. Z definicji bazy dualnej $\mathcal{B}^*$ wiadomo, że ta współrzędna wynosi $a_{ij} = w_i^*(\phi(v_j))$. Po transpozycji macierzy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$ wyraz a_{ij} staje się i-tym wyrazem j-tego wiersza macierzy transponowanej. Odpowiedni wyraz b_{ji} macierzy $M(\phi^*)_{\mathcal{B}^*}^{\mathcal{A}^*}$ jest j-tą współrzędną i-tej kolumny tej macierzy, a więc to j-ta współrzędna wektora $\phi^*(w_i^*)$ w bazie $\mathcal{A}^*$. Ale $\phi^*(w_i^*) = w_i^* \circ \phi$. W rezultacie j-ta współrzędna tego funkcjonału w bazie $\mathcal{A}^*$ to $w_i^*(\phi(v_j))$, co należało pokazać. □

Wniosek 11.6

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Wówczas:

(a) ϕ jest monomorfizmem $\Leftrightarrow$ gdy ϕ^* jest epimorfizmem,

(b) ϕ jest epimorfizmem $\Leftrightarrow$ gdy ϕ^* jest monomorfizmem.

Pokazujemy dowód dla przestrzeni skończonego wymiaru. Dla dowolnych przestrzeni rezultat ten wymaga pewnika wyboru. Dowód można znaleźć w skrypcie dra Strojnowskiego oraz notatkach dra Koźniewskiego.

Dowód. Wiemy już, że $r(\phi) = r(\phi^*)$, a zatem na mocy twierdzenia o sumie wymiarów jądra i obrazu przekształcenia mamy równoważności:

$$\phi \text{ jest monomorfizmem} \iff r(\phi) = \dim V \iff r(\phi^*) = \dim V^* \iff \phi^* \text{ jest epimorfizmem.}$$

oraz równoważności:

$$\phi \text{ jest epimorfizmem} \iff r(\phi) = \dim W \iff r(\phi^*) = \dim W^* \iff \phi^* \text{ jest monomorfizmem.}$$

□

Czym jest przestrzeń sprzężona do przestrzeni sprzężonej, czyli V^{**}? Jest to przestrzeń złożona z funkcjonałów $\Omega : V^* \to K$. A zatem każdemu funkcjonałowi z V^* przypisujemy element z K.

Definicja 11.11: Ewaluacja

Niech $\phi : V \to K$ będzie funkcjonałem liniowym. Ewaluacją funkcjonału ϕ w wektorze α nazywamy przekształcenie $\mathrm{ev}_\alpha \in V^*$ zadane wzorem $\mathrm{ev}_\alpha(\phi) = \phi(\alpha)$.

Jest zupełnie jasne, że ev_α jest przekształceniem liniowym, dla każdego $\alpha \in V$. Co więcej, każda ewaluacja jest elementem V^{**}. A zatem każdemu wektorowi z V przypisaliśmy w naturalny sposób element V^{**}. Ma to ważne skutki w przypadku skończenie wymiarowym.

Twierdzenie 11.7

Niech V będzie przestrzenią skończonego wymiaru. Przekształcenie $e : V \to V^{**}$ zadane wzorem

$$e(\alpha) = \mathrm{ev}_\alpha$$

jest izomorfizmem przestrzeni liniowych.

Dowód. Oczywiście e jest przekształceniem liniowym. Skoro V oraz V^{**} są tego samego wymiaru wystarczy pokazać, że e jest monomorfizmem. Załóżmy, że $\alpha \in \ker(e)$. Wówczas ev_α jest elementem zerowym w V^{**}, czyli dla każdego $\phi : V \to K$ mamy $\mathrm{ev}_\alpha(\phi) = 0$. Z definicji e oznacza to, że $\phi(\alpha) = 0$, dla każdego $\phi \in V^*$. A zatem wektor α ma tę własność, że ewaluowany na każdym funkcjonale liniowym jest zerem. Jedyny element z V o tej własności to 0, a więc $\ker(e) = \{0\}$. □

Przykład. Pokażemy, że $K[x]^* \simeq K[[x]]$, gdzie $K[[x]]$ jest zbiorem nieskończonych sum formalnych postaci: $\sum_{i=1}^{\infty} a_i x^i$, gdzie $a_i \in K$. Przestrzeń $K[[x]]$ nazywamy **szeregami formalnymi** nad ciałem K.

Dowód. Niech $\phi : K[x]^* \to K[[x]]$ będzie określone wzorem:

$$\phi(f) = \sum_{i=1}^{\infty} f(x_i) x^i.$$

A więc współczynnikiem przy x^i w szeregu $\phi(f)$ jest wartość funkcjonału f na x^i. Oczywiście $\phi(af+bg) = a\phi(f)+b\phi(g)$, dla dowolnych $f, g \in K[x]^*$ oraz $a, b \in K$. Jest to więc przekształcenie liniowe. Nietrudno też widzieć, że ϕ ma trywialne jądro. Tylko przekształcenie zerowe ma tę własność, że $f(x^i) = 0$, dla każdego i. Co więcej, ψ jest surjekcją, bo dla $w = \sum_{i=1}^{\infty} a_i x^i \in K[[x]]$ bierzemy $f \in K[x]^*$ taki, że $f(x^i) = a_i$ (określamy funkcjonał na bazie $K[x]$, więc taki f istnieje). Oczywiście $w = \phi(f)$, wiec ψ jest izomorfizmem. □

Nie mamy do dyspozycji teorii liczb kardynalnych, ale w jej języku zachodzą równości: $\dim K[x] = \omega$, zaś $\dim K[[x]] \geqslant 2^\omega$, zależnie od mocy ciała K. A zatem $K[x] \not\simeq K[x]^*$. Dowodzi się następujące twierdzenie.

Twierdzenie 11.1. *Jeżeli V jest przestrzenią liniową nad ciałem K i $\dim(V) = \infty$, to $\dim V^* = |K|^{\dim V}$.*

Dowód można przeczytać w skrypcie dr. Strojnowskiego:
https://www.mimuw.edu.pl/~stroa/Gal_Dodatki/Sprzezone.pdf.

Trudno być może dostrzec od razu motywacje geometryczne jakie stoją za pojęciem przestrzeni sprzężonej. Takie motywacje dostaniecie Państwo na analizie matematycznej, ale nie tylko... Nawet odnosząc się do fizyki (szkolnej?) wiemy, że siła reprezentowana jest najczęściej przez wektor F. Zwykle interesuje nas pytanie: co „robi" siła gdy przesuwa się obiekt z punktu A do punktu B. Jeśli q jest pozycją obiektu w przestrzeni, to F działa na q dając liczbę zwaną pracą... A więc o pracy można myśleć jak o funkcjonale... Można by tu dużo mówić, ale na razie powiedzmy tylko, że przestrzenie sprzężone są ważne.

Na koniec przyjrzyjmy się jeszcze jednej konstrukcji związanej z przestrzenią sprzężoną, która da nam ciekawą intuicję związaną z twierdzeniem Kroneckera-Capelli, którą od początku jakoś tu agituję.

Definicja 11.2. *Niech U będzie podprzestrzenią V. Anihilatorem podprzestrzeni U w V^*, oznaczamy przez $Ann(U)$ nazwiemy zbiór wszystkich funkcjonałów na V, które znikają na U, czyli:*

$$Ann(U) = \{f \in V^* \mid f(u) = 0, \text{ dla każdego } u \in U\}.$$

Nietrudno widzieć, że $Ann(U)$ to podprzestrzeń liniowa.

Twierdzenie 11.3. *Istnieje izomorfizm $Ann(U) \simeq (V/U)^*$. A zatem możemy identyfikować funkcjonały na V/U z elementami $Ann(U)$.*

Dowód. Niech $f \in Ann(U)$. Jest to liniowy funkcjonał na V, który znika na U. A zatem możemy określić funkcjonał liniowy f' na V/U dany wzorem:

$$f'(v + U) = f(v).$$

Innymi słowy, f' posyła warstwę $v + U$ na skalar $f(v)$. Zobaczmy, że jest to przekształcenie dobrze określone. Załóżmy, że $v + U = v' + U$. Musimy sprawdzić, że $f'(v + U) = f'(v' + U)$. Istotnie, skoro $v + U = v' + U$, to $v - v' \in U$, a zatem:

$$0 = f(v - v') = f(v) - f(v').$$

Jest jasne, że f' jest liniowe. Pokażmy, że to izomorfizm. Weźmy najpierw $\phi \in \ker(f')$. A zatem $f'(\phi)$ jest funkcjonałem zerowym na V/U. A zatem:

$$0 = f'(\phi)(v + U) = \phi(v), \text{ dla każdego } v \in V.$$

A zatem ϕ jest funkcjonałem zerowym. W szczególności jest to element zerowy $Ann(U)$. Zatem $\ker(f') = \{0\}$. Aby pokazać, że f' to surjekcja, weźmy $g \in (V/U)^*$. Określamy element $f \in V^*$ jako:

$$f(v) = g(v + U), \text{ dla każdego } v \in V.$$

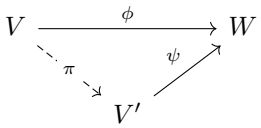
Twierdzimy, że f należy tak naprawdę do $Ann(U)$. Istotnie, jeśli $u \in U$, to $g(u + U) = g(U) = 0$, skoro U jest elementem zerowym V/U oraz g jest liniowe. Zatem $f(v) = 0$, czyli $f \in Ann(U)$. Z definicji f' wynika, że $f'(f) = g$, więc f' jest surjekcją. □

11.4 Dodatek. Faktoryzacje i przekształcenia ilorazowe

Jakiś czas temu w notatkach do wykładu pojawiła się definicja przestrzeni ilorazowej V/W, gdzie $W \subseteq V$ jest podprzestrzenią V. Wspomnieliśmy wówczas o formule $\dim V/W = \dim V - \dim W$. Jak ją wyprowadzić, do czego używa się takich przestrzeni, skąd taka nazwa i dlaczego warto o nich powiedzieć coś właśnie teraz, gdy zajmujemy się przekształceniami liniowymi? Kluczem do sprawy jest pojęcie faktoryzowania się przekształcenia liniowego, na swój sposób odwrotne do pojęcia złożenia. Oto definicja.

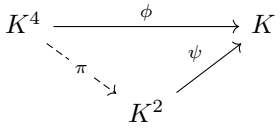
Definicja 11.12: Faktoryzowanie się przekształcenia przez inne

Niech $\phi : V \to W$ będzie przekształceniem liniowym. Powiemy, że ϕ FAKTORYZUJE SIĘ PRZEZ PRZEKSZTAŁCENIE liniowe $\psi : V' \to W$, jeśli istnieje przekształcenie $\pi : V \to V'$, że $\phi = \psi \circ \pi$, czyli gdy następujący diagram jest przemienny.



Oczywiście faktoryzacja zawsze jest możliwa, jeśli weźmiemy $V' = V$ oraz $\psi = \phi$. Sprawa jest jednak nieco ciekawsza. Popatrzmy najpierw na przykłady:

- Rozważmy przekształcenie $\phi : K^4 \to K$ dane wzorem $\phi((x_1, x_2, x_3, x_4)) = x_4$. Jest ono oczywiście liniowe. Przekształcenie to faktoryzuje się przez $\psi' : K^2 \to K$ dane wzorem $\psi(y_1, y_2) = y_2$. Istotnie, jeśli $\pi : K^4 \to K^2$ dane jest wzorem: $\pi(z_1, z_2, z_3, z_4) = (z_2, z_4)$, to mamy: $\psi(\pi((x_1, x_2, x_3, x_4))) = \psi((x_2, x_4)) = x_4 = \phi((x_1, x_2, x_3, x_4))$. Mamy więc:

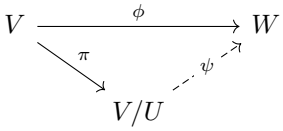


- Rozważmy podprzestrzeń C przestrzeni $\mathbb{R}^\infty$ złożoną z wszystkich ciągów zbieżnych i rozważmy przekształcenie $\phi : C \to \mathbb{R}$ dane wzorem $\phi((x_1, \ldots)) = \lim_{n\to\infty} x_n$. Jest to oczywiście przekształcenie liniowe. Czy znajdziemy dla niego jakąś faktoryzację? Może ktoś uzna to za trywialne – ale owszem, jesteśmy w stanie to zrobić. Rozważmy podprzestrzeń D wszystkich ciągów stałych. Bierzemy teraz przekształcenie $\psi : D \to \mathbb{R}$ dane wzorem $\psi((x, x, x, ...)) = x$. Czy widzimy, że ϕ faktoryzuje się przez ψ? Jak wygląda przekształcenie π? I skąd wiedzieliśmy, żeby szukać właśnie ciągów stałych?

Kluczem jest pojęcie przestrzeni ilorazowej. Zauważmy, że jeśli $W \subseteq V$, to mamy naturalne przekształcenie $\pi : V \to V/W$ zadane wzorem: $\pi(\alpha) = \alpha + W$ (przyporządkowujemy wektorowi jego warstwę). Jest to dobrze określone przekształcenie liniowe. Zachodzi następujące twierdzenie.

Twierdzenie 11.8: Własność uniwersalna przestrzeni ilorazowej

Niech V będzie przestrzenią liniową, zaś U – jej podprzestrzenią. Wówczas dla każdego przekształcenia liniowego $\phi : V \to W$ takiego, że $\ker(\phi)$ zawiera U, istnieje dokładnie jedno przekształcenie liniowe $\psi : V/U \to W$ takie, że $\phi = \pi \circ \psi$, czyli następujący diagram jest przemienny.



W szczególności dowolne przekształcenie liniowe $\phi : V \to W$ faktoryzuje się przez odpowiednie przekształcenie $\psi : V/ker(\phi) \to W$ dane wzorem: $\psi(\alpha + \ker(\phi)) = \phi(\alpha)$.

Dowód. Określamy przekształcenie liniowe ψ na dowolnej warstwie $\alpha + U$ wektora $\alpha \in V$:

$$\psi(\alpha + U) = \phi(\alpha).$$

Przekształcenie to jest dobrze określone na V/U, bo dla $\alpha, \alpha' \in U$ takich, że $\alpha + U = \alpha' + U$ mamy

$$\phi(\alpha) = \phi(\alpha') \iff \phi(\alpha - \alpha') = 0 \iff \alpha - \alpha' \in \ker(\phi).$$

Skoro jednak $\alpha + U = \alpha' + U$, to $\alpha - \alpha' \in U$. Skoro zaś $U \subseteq \ker(\phi)$, to $\phi(\alpha - \alpha') = 0$, czyli $\phi(\alpha) = \phi(\alpha)'$.

Liniowość ψ jest prostą konsekwencją liniowości ϕ oraz działań w V/U:

$$\psi((\alpha + U) + (\alpha' + U)) = \psi((\alpha + \alpha' + U) = \phi(\alpha + \alpha') = \phi(\alpha) + \phi(\alpha') = \psi(\alpha + U) + \psi(\alpha' + U).$$

$$\psi(\lambda(\alpha + U)) = \psi(\lambda\alpha + U) = \phi(\lambda\alpha) = \lambda\phi(\alpha) = \lambda\psi(\alpha + U).$$

Oczywiście co najwyżej jedna funkcja spełniać może warunek $\phi = \pi \circ \psi$, co kończy dowód.

□

Z powyższego rezultatu płyną następujące wnioski, zwane twierdzeniami o izomorfizmie.

Twierdzenie 11.9: Pierwsze twierdzenie o izomorfizmie

Niech V, W będą przestrzeniami liniowymi nad ciałem K oraz niech $\phi : V \to W$ będzie przekształceniem liniowym. Wówczas przekształcenie $\psi : V/\ker(\phi) \to \mathrm{im}(\phi)$ dane wzorem

$$\psi(\alpha + \ker(\phi)) = \phi(v).$$

jest izomorfizmem przestrzeni liniowych $V/\ker(\phi)$ oraz $\mathrm{im}(\phi)$.

Twierdzenie 11.10: Drugie twierdzenie o izomorfizmie

Niech V będzie przestrzenią liniową, zaś $U, W \subseteq V$ będą jej podprzestrzeniami. Wówczas mamy izomorfizm przestrzeni liniowych:

$$U/(U \cap W) \simeq (U + W)/W.$$

Twierdzenie 11.11: Trzecie twierdzenie o izomorfizmie

Niech V będzie przestrzenią liniową, W — będzie podprzestrzenią V, zaś U — podprzestrzenią W. Wówczas W/U jest podprzestrzenią V/U i mamy izomorfizm przestrzeni liniowych:

$$(V/U)/(W/U) \simeq V/W.$$

Twierdzenia te mają spore znacznie na przykład w teorii Jordana lub teorii iloczynów tensorowych. Zostawiam Czytelnikowi ich dowody, między innymi sprawdzenie, że przekształcenie liniowe ψ określone w pierwszym twierdzeniu tak, jak we własności uniwersalnej, jest rzeczywiście izomorfizmem. To łatwe ćwiczenie. Warto natomiast wspomnieć jeszcze o wyniku, zwanym twierdzeniem o odpowiedniości.

Twierdzenie 11.12: Twierdzenie o odpowiedniości

Niech V będzie przestrzenią liniową oraz W jej podprzestrzenią. Przez $S(X)$ oznaczmy zbiór wszystkich podprzestrzeni przestrzeni liniowej X. Wówczas ma miejsce bijekcja

$$S(V/W) \longleftrightarrow \{U \in S(V) : W \subseteq U \subseteq V\}$$

polegająca na przypisaniu podprzestrzeni U spełniającej warunek $W \subseteq U \subseteq V$ podprzestrzeni U/W przestrzeni V/W. Bijekcja ta zachowuje sumy i przecięcia podprzestrzeni (jest izomorfizmem krat).

Rozdział 12

Wyznacznik macierzy kwadratowej

12.1 Wykład dwunasty

Zakończyliśmy podstawową część wykładu dotyczącą przestrzeni i przekształceń liniowych. Kolejne dwa wykłady poświęcimy najbardziej zapewne znanemu (z tych niebanalnych) pojęciu algebraicznemu — wyznacznikowi. Czytelnik mógł o nim słyszeć w kontekście rozwiązywania układów równań. Pojęcie wyznacznika można określić i badać w sposób czysto algebraiczny, co jest naszym celem na ten wykład. Niektóre istotne motywacje wygodnie jest wysławiać także w języku geometrii, o czym powiemy później.

Definicja 12.1: Macierze kwadratowe

Dla każdego całkowitego $n \geqslant 1$ zbiór $M_{n\times n}(K)$ macierzy o n wierszach i n kolumnach nazywamy zbiorem MACIERZY KWADRATOWYCH ROZMIARU $\boldsymbol{n}$ i oznaczamy przez $M_n(K)$.

Z punktu widzenia przekształceń liniowych macierze kwadratowe są macierzami przekształceń liniowych pomiędzy przestrzeniami tego samego (skończonego) wymiaru. W szczególności, są to macierze przekształceń liniowych przestrzeni liniowej do niej samej, którymi zajmiemy się w drugim semestrze.

Definicja 12.2

Niech $A = [a_{ij}] \in M_n(K)$, gdzie $n > 1$. Dla każdej pary $1 \leqslant i, j \leqslant n$ określamy macierz postaci

$$A_{ij} \in M_{n-1}(K)$$

otrzymaną z macierzy A przez skreślenie odpowiednio i-tego wiersza i j-tej kolumny.

Przykład: dla macierzy $A \in M_3(\mathbb{R})$ postaci

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 5 \end{bmatrix}$$

mamy:

$$A_{11} = \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \quad A_{31} = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

Intuicja jest następująca: chcemy określić funkcję $\det : M_n(K) \to K$, przypominającą „funkcję objętości". Objętość często liczymy według formuł postaci: „długość/pole podstawy razy wysokość". Ogólnie objętość obiektu w przestrzeni n wymiarowej wyznaczać chcemy poprzez znajomosc objętości $n-1$ oraz 1-wymiarowej. Okazuje się, że aby określić wyznacznik macierzy $A \in M_n(K)$ potrzebna jest znajomość:

- wyrazów macierzy w pierwszej kolumnie: $a_{11}, a_{21}, \ldots, a_{n1}$,
- wyznaczników macierzy rozmiaru $n-1$ postaci $A_{11}, A_{21}, \ldots, A_{n1}$.

Definicja 12.3: Wyznacznik — rozwinięcie Laplace'a względem pierwszej kolumny

Definiujemy funkcję $\det : M_n(K) \to K$ w sposób rekurencyjny

- Dla $n = 1$ kładziemy $\det : M_1(K) \to K$, gdzie $\det(A) = a$, dla $A = [a]$.
- Dla $A = [a_{ij}] \in M_n(K)$ określamy $\det : M_n(K) \to K$ znając $\det : M_{n-1}(K) \to K$ wzorem:

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} + \ldots + (-1)^{n+1} a_{n1} \det A_{n1} = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}.$$

Funkcję $\det : M_n(K) \to K$ nazywamy WYZNACZNIKIEM. Czasem zamiast pisać $\det A$, piszemy $|A|$.

Od razu warto dodać pewne doprecyzowanie. W zasadzie definiujemy ciąg funkcji – formalnie należałoby być może pisać (ale nikt tego nie robi) $\det\limits_n : M_n(K) \to K$. Przy takiej konwencji mielibyśmy

$$\det_n A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det_{n-1} A_{i1}.$$

Zacznijmy od kilku przykładów dla małych n.

- Dla $A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \in M_2(\mathbb{R})$ mamy $A_{11} = [2], A_{21} = [4]$, czyli:

$$\det A = (-1)^{1+1} \cdot 1 \cdot \det A_{11} + (-1)^{2+1} \cdot 3 \cdot \det A_{21} = 1 \cdot 2 - 3 \cdot 4 = -10.$$

  Ogólnie dla macierzy $A = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}$ mamy $|A| = x_1 y_2 - x_2 y_1$.

- Dla $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 5 & 3 \end{bmatrix} \in M_3(\mathbb{R})$ mamy $a_{11} = 4, a_{21} = 0, a_{31} = 0, \quad |A_{11}| = 11, |A_{21}| = 0, |A_{31}| = 0,$

  czyli

$$\det A = 4 \cdot 11 - 0 \cdot 0 + 0 \cdot 0 = 44.$$

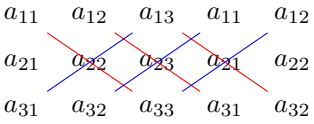
  Ogólnie dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

  mamy:

$$|A| = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} =$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}.$$

  Formułę tą można uzyskać korzystając z tzw. metody Sarrusa, polegającej na wypisaniu obok macierzy A dwóch pierwszych jej kolumn. Wówczas trzy składniki powyższej sumy występujące ze znakiem $+$ uzyskujemy przez iloczyn wyrazów połączonych na czerwono (rys. niżej), a trzy składniki ze znakiem $-$ uzyskujemy przez wymnożenie wyrazów połączonych na niebiesko.



Na koniec wstępu jeszcze jedna fundamentalna kwestia. **Wyznacznik nie jest funkcją liniową!** Mamy np.

$$4 = \det \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \neq \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 2.$$

Definicja 12.4: Macierze trójkątne i diagonalne

Niech $A = [a_{ij}] \in M_n(K)$. Zbiór wyrazów $\{a_{11}, a_{22}, \ldots, a_{nn}\}$ nazywamy PRZEKĄTNĄ lub DIAGONALĄ macierzy A. Powiemy, że A jest:

- GÓRNOTRÓJKĄTNA, jeśli $a_{ij} = 0$, dla $i > j$
  (pod przekątną macierzy A stoją wyrazy zerowe),
- DOLNOTRÓJKĄTNA, jeśli $a_{ij} = 0$, dla $j > i$
  (czyli nad przekątną macierzy A stoją wyrazy zerowe),
- DIAGONALNA, jeśli jest jednocześnie górnotrójkątna i dolnotrójkątna.

Przykłady macierzy odpowiednio: górnotrójkątnej, dolnotrójkątnej, diagonalnej w $M_2(K)$:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Odnotujmy też, że każda macierz w postaci schodkowej jest górnotrójkątna.

Obserwacja 12.1

Jeśli $A \in M_n(K)$ jest macierzą górnotrójkątną (na przykład: macierzą w postaci schodkowej), to jej wyznacznik równy jest iloczynowi wyrazów na przekątnej.

Dowód. Dowodzimy tezę przez indukcję. Dla $n = 1$ wynika ona wprost z definicji. Weźmy zatem macierz $A = [a_{ij}]$ rozmiaru $n \times n$ oraz zauważmy, że $a_{21} = \ldots = a_{n1} = 0$, a zatem z definicji wyznacznika mamy

$$\det A = (-1)^{1+1} \cdot a_{11} \cdot \det A_{11}.$$

Macierz A_{11} powstaje z A przez usunięcie pierwszego wiersza i kolumny. Skoro A jest górnotrójkątna, to również A_{11} jest górnotrójkątna. A zatem z założenia indukcyjnego $\det A_{11} = a_{22} \cdot \ldots \cdot a_{nn}$. □

Czytelnik mógłby spytać czy to, że liczymy wyznacznik „za pomocą pierwszej kolumny", to znaczy – mnożąc (z odpowiednim znakiem) wyrazy pierwszej kolumny przez wyznaczniki odpowiednich macierzy rzeczywiście zależy od wyboru pierwszej kolumny? Okazuje się, że tak nie jest. Sposób liczenia wyznacznika opisany w definicji oparty o korzystanie z wyznaczników macierzy mniejszego rodzaju oraz wyrazów pierwszej kolumny zwany jest obliczaniem wyznacznika za pomocą rozwinięcia Laplace'a względem pierwszej kolumny. Można pytać czy możliwe jest obliczenie wyznacznika za pomocą rozwinięcia względem innych kolumn, albo nawet wierszy? Innymi słowy, czy mając macierz $A \in M_n(K)$ oraz znając:

- wyrazy macierzy w i-tym wierszu (odp. j-tej kolumnie): $a_{i1}, a_{i2}, \ldots, a_{in}$ (odp. $a_{1j}, a_{2j}, \ldots, a_{nj}$),
- wyznaczniki macierzy rozmiaru $n-1$ postaci $A_{i1}, A_{i2}, \ldots, A_{in}$ (odp. $A_{1j}, A_{2j}, \ldots, A_{nj}$)

mamy na przykład formuły

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}? \qquad (\dagger)$$

Okazuje się, że tak jest i pokażemy to w dalszej części wykładu. Bezpośrednie uzasadnienie można znaleźć w skrypcie wydziałowym. Przykładowym skutkiem (†) jest dowód poniższej obserwacji.

Obserwacja 12.2

Jeśli $A \in M_n(K)$ jest dolnotrójkątna, to $\det A$ równy jest iloczynowi wyrazów na przekątnej.

Dowód. Dowodzimy tezę przez indukcję. Dla $n = 1$ wynika ona wprost z definicji. Weźmy zatem macierz $A = [a_{ij}]$ rozmiaru $n \times n$ oraz zauważmy, że $a_{12} = \ldots = a_{1n} = 0$, a zatem z definicji wyznacznika oraz z wzoru (†) mamy (rozwinięcie względem pierwszego wiersza): $\det A = (-1)^{1+1} \cdot a_{11} \cdot \det A_{11}$. Skoro A jest dolnotrójkątna, to również A_{nn} taka jest. A zatem z założenia indukcyjnego $\det A_{11} = a_{22} \cdot \ldots \cdot a_{nn}$. □

W dalszym ciągu posługujemy się jednak nadal jedynie definicją wyznacznika za pomocą rozwinięcia Laplace'a względem pierwszej kolumny. Do wykonywania rachunków kluczowe są kolejne obserwacje, dotyczące zmiany wyznacznika przy wykonywaniu operacji elementarnych na wierszach lub kolumnach.

Twierdzenie 12.1: Wyznacznik, a operacje elementarne

Rozważmy funkcję $\det : M_n(K) \to K$. Wówczas

(1) jeśli macierz A' została otrzymana z macierzy A przez dodanie do pewnego wiersza skalarnej wielokrotności innego wiersza, wówczas $\det(A') = \det(A)$,

(2) przestawienie wierszy zmienia znak wyznacznika, tzn. jeśli macierz A' została otrzymana z macierzy A przez zamianę miejscami dwóch wierszy, to $\det(A') = -\det(A)$.

(3) jeśli macierz A' została otrzymana z macierzy A przez pomnożenie pewnego wiersza przez c, to $\det(A') = c \cdot \det(A)$.

Łącząc powyższe twierdzenie z obserwacją mówiącą, że wyznacznik macierzy górnotrójkątnej jest iloczynem wyrazów na przekątnej dostajemy prosty algorytm liczenia wyznacznika (inaczej niż z rozwinięcia Laplace'a) poprzez „schodkowanie": aby policzyć $|A|$ sprowadzamy macierz A do postaci schodkowej A' operacjami typu (1) i (2). Wyznacznik A' to po prostu iloczyn wyrazów na przekątnej. Mamy też $|A| = (-1)^k \cdot |A'|$, gdzie k oznacza liczbę operacji typu (2) użytych przy sprowadzaniu A do A'.

Przykład (kolokwium, 2019). Obliczyć $\det A$, gdzie $A \in M_n(\mathbb{R})$ jest postaci:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & 1 & 1 & \dots & 1 & 2 \\ 1 & 1 & 3 & 1 & \dots & 1 & 3 \\ 1 & 1 & 1 & 4 & \dots & 1 & 4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \dots & n-1 & n-1 \\ 1 & 2 & 3 & 4 & \dots & n-1 & n \end{bmatrix}.$$

Rozwiązanie. Załóżmy, że $n > 1$. Odejmujemy pierwszy wiersz macierzy A od pozostałych i mamy

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 2 & 0 & \dots & 0 & 2 \\ 0 & 0 & 0 & 3 & \dots & 0 & 3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & n-2 & n-2 \\ 0 & 1 & 2 & 3 & \dots & n-2 & n-1 \end{bmatrix}$$

Następnie od ostatniego wiersza odejmujemy wiersze $2, 3, \dots, n-1$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 2 & 0 & \dots & 0 & 2 \\ 0 & 0 & 0 & 3 & \dots & 0 & 3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & n-2 & n-2 \\ 0 & 0 & 0 & 0 & \dots & 0 & a \end{bmatrix}$$

gdzie $a = (n-1) - 1 - 2 - 3 - \dots - (n-2) = (n-1) - (n-1)(n-2)/2 = -(n-1)(n-4)/2$. Powyższe operacje nie zmieniają wyznacznika, a uzyskana macierz jest (górno)trójkątna; mamy więc

$$\det(A) = (n-2)! \cdot (-(n-1)(n-4))/2 = -(n-1)! \cdot (n-4)/2.$$

Argumenty prowadzące do (†) sprowadzają się do pokazania, że funkcja det określona za pomocą rozwinięcia Laplace'a względem pierwszej kolumny ma jako jedyna pewne własności, jako funkcja określona na wierszach macierzy A. Właśności te mają istotne podłoże geometryczne, do którego wrócimy później.

Definicja 12.5: Funkcje jednorodne i addytywne względem wierszy macierzy

Powiemy, że funkcja $\phi : M_n(K) \to K$, jest

- JEDNORODNA WZGLĘDEM k-TEGO WIERSZA, jeśli dla każdej $A \in M_n(K)$ oraz każdego $c \in K$ mamy $\phi(A') = c \cdot \phi(A)$, gdzie A' powstaje z A przez pomnożenie k-tego wiersza przez c,
- ADDYTYWNA WZGLĘDEM k-TEGO WIERSZA, jeśli dla każdej trójki macierzy $A, B, C \in M_n(K)$ takiej, że:
  - k-ty wiersz macierzy C to suma k-tego wiersza A oraz k-tego wiersza B,
  - l-te macierzy A, B, C są identyczne, dla $l \neq k$.

  zachodzi $\phi(C) = \phi(A) + \phi(B)$.

Przykład. Rozważmy macierze A, B, C rozmiaru 3×3 o takich samych pierwszych i drugich wierszach oraz takie, że trzeci wiersz macierzy C jest sumą trzeciego wiersza macierzy A oraz trzeciego wiersza macierzy B:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 4 & 2 \end{bmatrix}.$$

Nietrudno sprawdzić (choćby na mocy wzoru Sarrusa), że:

$$\det C = \det A + \det B.$$

Przykład ten sugeruje addytywność wyznacznika macierzy rozmiaru 3×3 względem trzeciego wiersza. Pokażmy ogólny rezultat (mówi on, że wyznacznik jest tzw. *funkcją wieloliniową* na wierszach macierzy).

Twierdzenie 12.2

Funkcja $\det : M_n(K) \to K$ jest jednorodna i addytywna względem każdego wiersza.

Dowód. Indukcja ze względu na n. Dla $n = 1$ – jasne. Weźmy $A = [a_{ij}] \in M_n(K)$. Mnożymy k-ty wiersz A przez $c \in K$ dostając B. Wówczas:

- $B_{k1} = A_{k1}$,
- dla $j \neq k$ każda z macierzy B_{j1} powstaje z A_{j1} przez pomnożenie pewnego wiersza przez stałą. Z założenia indukcyjnego $\det B_{j1} = c \det A_{j1}$.

Zatem:

$$\begin{aligned} \det B &= (-1)^{1+1} a_{11} \det B_{11} + \ldots + (-1)^{k+1} c a_{k1} \det B_{k1} + \ldots + (-1)^{n+1} a_{n1} \det B_{n1} = \\ &= (-1)^{1+1} a_{11} c \det A_{11} + \ldots + (-1)^{k+1} c a_{k1} \det A_{k1} + \ldots + (-1)^{n+1} a_{n1} c \det A_{n1} = c \cdot \det A. \end{aligned}$$

Pokazujemy teraz addytywność zdefiniowanej na początku funkcji det. Chcemy pokazać, że:

$$\underbrace{\det \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} + y_{k1} & \ldots & x_{kn} + y_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}}_{Z} = \underbrace{\det \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} & \ldots & x_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}}_{X} + \underbrace{\det \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ y_{k1} & \ldots & y_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}}_{Y}$$

gdzie $X, Y, Z \in M_n(K)$ różnią się tylko k-tymi wierszami — k-ty wiersz Z jest sumą k-tych wierszy X, Y.

Zauważmy, że:

- dla każdego $1 \leqslant k \leqslant n$ macierze powstające z Z, X, Y przez usunięcie k-tego wiersza i pierwszej kolumny są równe, tzn. $Z_{k1} = Y_{k1} = X_{k1}$,
- dla $j \neq k$ macierze Z_{j1}, Y_{j1}, X_{j1} różnią się tylko k-tym wierszem, przy czym k-ty wiersz Z_{j1} jest sumą k-tych wierszy Y_{j1} oraz X_{j1}. Z założenia indukcyjnego mamy zatem

$$\det Z_{j1} = \det X_{j1} + \det Y_{j1}, \qquad \text{dla } j \neq k.$$

Stąd:

$$\begin{aligned}\det Z &= (-1)^{1+1} a_{11} \det Z_{11} + \ldots + (-1)^{k+1}(x_{k1} + y_{k1}) \det Z_{k1} + \ldots + (-1)^{n+1} a_{n1} \det Z_{n1} = \\ &= \sum_{j \neq k} (-1)^{j+1} a_{j1} (\det X_{j1} + \det Y_{j1}) + (-1)^{k+1} x_{k1} \det X_{k1} + y_{k1} \det Y_{k1} = \det X + \det Y\end{aligned}$$

□

Wniosek 12.1

Jeśli $A \in M_n(K)$ oraz A ma zerowy wiersz, to $\det(A) = 0$.

Dowód. Jeśli k-ty wiersz macierzy A jest zerowy, to z addytywności wyznacznika względem k-tego wiersza mamy $\det A = \det A + \det A$ (w dowodzie wyżej przyjmyjemy $X = Y = Z = A$). □

Obserwacja 12.3

Jeśli dwa sąsiednie wiersze macierzy $A \in M_n(K)$ są identyczne, dla $n \geqslant 2$, wówczas $\det A = 0$.

Dowód. Dowód to indukcja ze względu na n. Dla $n = 2$ teza jest oczywiście prawdziwa. Załóżmy, że $n > 2$ oraz identyczne są i-ty oraz $i+1$-ty wiersz macierzy $A = [a_{ij}]$, czyli dla $1 \leqslant k \leqslant n$ mamy $a_{ik} = a_{i+1,k}$.

$$A = \begin{bmatrix} \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \ldots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix}.$$

Zauważmy, że $A_{i1} = A_{i+1,1}$ oraz dla $k \neq i, i+1$ macierze A_{k1} mają dwa identyczne wiersze. Z założenia indukcyjnego mamy zatem $\det A_{k1} = 0$. Zatem $\det A$ równy jest:

$$\sum_{k \neq i, i+1} (-1)^{k+1} a_{k1} \det A_{k1} + (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+1+1} a_{i+1,1} \det A_{i+1,1} = (-1)^{i+1}(1-1) a_{i1} \det A_{i1} = 0.$$

□

Przykład, ilustrujący krok indukcyjny w dowodzie wyżej. Rozważmy macierz rozmiaru 4×4 o dwóch identycznych wierszach i załóżmy, że obserwacja jest prawdziwa dla macierzy 3×3. Weźmy macierz o dwóch identycznych wierszach i policzmy jej wyznacznik

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

Mamy zatem (stosujemy notację $|A|$ alternatywną do $\det A$):

$$|A| = 1 \cdot |A_{11}| - 5 \cdot |A_{21}| + 5 \cdot |A_{31}| - 9 \cdot |A_{41}| = 1 \cdot \underbrace{\begin{vmatrix} 6 & 7 & 8 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_{0} - 5 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_{20} + 5 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 8 & 7 & 1 \end{vmatrix}}_{20} - 9 \cdot \underbrace{\begin{vmatrix} 2 & 3 & 4 \\ 6 & 7 & 8 \\ 6 & 7 & 8 \end{vmatrix}}_{0}.$$

Wykażemy kolejne własności det, dążąc do dowodu (†). Wynikać ono będzie z następującego faktu. Proszę się na ten moment nie martwić, że nie jest jasne o czym mówi „slogan".

Twierdzenie 12.3: Slogan – Wyznacznik jest *jedyną* formą n-liniową antysymetryczną

Dla każdego $n \geqslant 1$ istnieje dokładnie jedna funkcja $\phi : M_n(K) \to K$, taka, że:

(1) Dla każdego $1 \leqslant k \leqslant n$ funkcja ϕ jest jednorodna względem k-tego wiersza.

(2) Dla każdego $1 \leqslant k \leqslant n$ funkcja ϕ jest addytywna względem k-tego wiersza.

(3) $\phi(A) = 0$, jeśli A ma identyczne dwa sąsiednie wiersze.

(4) $\phi(I_n) = 1$.

Obserwacja 12.4

Funkcja $\phi = \det$ spełnia warunki (1)-(4).

Dowód. Spełnianie warunków (1), (2) zapewnia Twierdzenie 12.2. Warunek (3) wynika stąd, że wyznacznik macierzy A jest niezerowy wtedy i tylko przy zamianie wierszy jest dalej niezerowy, a przy dwóch identycznych sąsiednich wierszach jest zerowy (Obserwacja 12.3). Oczywiście $\det(I_n) = 1$. □

Twierdzimy, że żadnej innej funkcji niż det spełniającej warunki (1)-(4) nie ma. Idea dowodu polega na pokazaniu, że funkcja spełniająca warunki (1)-(4) musi spełniać założenia z Twierdzenia 12.1. Następnie pokażemy, że ϕ jest jednoznacznie określona na macierzach operacji elementarnych co po dalszym rozumowaniu (i dzięki uzyskanym niedawno własnościom macierzy operacji elementarnych) pozwoli stwierdzić, że dla każdej macierzy może ona przyjmować tylko jedną wartość. Zaczynamy od badania własności funkcji spełniających (1)-(4) przy operacjach elementarnych. W szczególności dowodzić będziemy w kolejnych obserwacjach dalsze własności funkcji $\phi = \det$ (pokazując jednocześnie, że innych ϕ niż det nie ma).

Obserwacja 12.5

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Jeśli $C' \in M_n(K)$ powstaje z C przez zamianę dwóch sąsiednich wierszy, to $\phi(C) = -\phi(C')$.

Dowód. Niech wiersze macierzy C mają postać $w_1, \ldots, w_n$. Niech C' powstaje z C przez zamianę wiersza k-tego i $k+1$-wszego. Na mocy własności (2) i (3) funkcji ϕ:

$$\underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_k + w_{k+1} \\ w_k + w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_{0} = \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{0} + \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_{k+1} \\ w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_{0} + \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ w_{k+1} \\ \vdots \\ w_n \end{bmatrix}}_{\phi(C)} + \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_{k+1} \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(C')}$$

□

Obserwacja 12.6

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Jeśli macierz $C \in M_n(K)$ ma dwa identyczne wiersze, to $\phi(C) = 0$.

Uzasadnienie: za pomocą skończenie wielu operacji zamiany wierszy możemy zamienić C w macierz C' o dwóch sąsiednich wierszach równych. Z poprzedniej obserwacji mamy $\phi(C) = \pm\phi(C') = 0$.

Obserwacja 12.7

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Niech B będzie macierzą otrzymaną z macierzy A w wyniku dodania do wiersza l-tego wiersza k-tego pomnożonego przez $a \in K$. Wówczas: $\phi(B) = \phi(A)$.

Schemat uzasadnienia, korzystający z poprzednich wyników:

$$\underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_l + aw_k \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(B)} = \phi\begin{bmatrix} w_1 \\ \vdots \\ w_l \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix} + \phi\begin{bmatrix} w_1 \\ \vdots \\ aw_k \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix} = \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_l \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{\phi(A)} + a \cdot \underbrace{\phi\begin{bmatrix} w_1 \\ \vdots \\ w_k \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix}}_{0}.$$

Obserwacja 12.8: Wyznacznik macierzy operacji elementarnych

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. det, o czym już wiemy). Niech M będzie macierzą operacji elementarnej oraz $A \in M_n(K)$. Wówczas:

$$\phi(MA) = \begin{cases} \phi(A), & \text{dla M dodającej do wiersza skalar razy inny wiersz,} \\ -\phi(A), & \text{dla M zamieniającej dwa wiersze miejscami,} \\ c \cdot \phi(A), & \text{dla M mnożącej pewien wiersz przez } c \neq 0. \end{cases}$$

W szczególności dla $A = I_n$ mamy

$$\phi(M) = \begin{cases} 1, & \text{dla M dodającej do wiersza skalar razy inny wiersz,} \\ -1, & \text{dla M zamieniającej dwa wiersze miejscami,} \\ c, & \text{dla M mnożącej pewien wiersz przez } c \neq 0. \end{cases}$$

W każdym z opisanych przypadków zachodzi równość

$$\phi(MA) = \phi(M) \cdot \phi(A). \qquad (\diamondsuit)$$

Wynik ten wynika natychmiast z interpretacji operacji elementarnych w języku mnożenia macierzy.

Kolejny wynik pomocniczy ma jednocześnie fundamentalne znaczenie dla całego wykładu.

Twierdzenie 12.4: Macierz jest odwracalna $\iff$ macierz ma niezerowy wyznacznik

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Dla każdej macierzy $A \in M_n(K)$ równoważne są warunki:

- $\phi(A) \neq 0$,
- $r(A) = n$,
- A jest odwracalna.

Aby pokazać ten rezultat przypomnijmy, że jeśli A' jest postacią schodkową zredukowaną macierzy A, to istnieją macierze operacji elementarnych $M_1, \ldots, M_s$ takie, że:

$$A' = M_1 M_2 M_3 \ldots M_s A.$$

Dowód. Stosując wiele razy Obserwację 12.7 i formułę ($\diamondsuit$) mamy:

$$\phi(A') = \phi(M_1 M_2 M_3 \dots M_s A) = \phi(M_1)\phi(M_2)\phi(M_3)\dots\phi(M_s)\phi(A).$$

W rezultacie $\phi(A) \neq 0 \Leftrightarrow \phi(A') \neq 0$ (bo ϕM_i są zawsze niezerowe). Skoro A jest kwadratowa, to są dwie możliwości:

- $A' = I$,
- A' ma zerowy wiersz.

Twierdzimy, że w pierwszym przypadku $\phi(A') = 1$, a w drugim: $\phi(A') = 0$. Rzeczywiście, jeśli A' ma zerowy wiersz, to dodając do tego wiersza inny wiersz dostajemy macierz A'' o dwóch identycznych wierszach. A zatem $\phi(A'') = 0$, zgodnie z Obserwacją 12.5. Jednocześnie $\phi(A') = \phi(A'')$ (gdyż A'' powstaje przez dodanie wiersza A' do innego), czyli $\phi(A') = 0$. W rezultacie: $\phi(A) \neq 0 \Leftrightarrow A' = I$. Na poprzednim wykładzie pokazywaliśmy natomiast, że $A' = I \Leftrightarrow A$ jest odwracalna. □

Twierdzenie 12.5: Wzór Cauchy'ego

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Niech $A, B \in M_n(K)$. Wówczas:

$$\phi(AB) = \phi(A) \cdot \phi(B).$$

Dowód wzoru Cauchy'ego rozbija się na dwa przypadki.

- Przypadek 1. Macierz AB nie jest odwracalna. Zgodnie z Twierdzeniem 12.4 mamy $\phi(AB) = 0$. Oznacza to, że $\phi(A) = 0$ lub $\phi(B) = 0$. Inaczej na mocy Twierdzenia 12.4 macierze A, B byłyby odwracalne, a z nimi i AB, bo jak wiadomo $(AB) \cdot B^{-1}A^{-1} = I$.
- Przypadek 2. Załóżmy, że AB jest odwracalna. Zgodnie z wynikiem wyżej $r(AB) = n$, a wiemy z wcześniejszych wykładów[1], że to oznacza, że $r(A) = n$ oraz $r(B) = n$. Z Twierdzenia 12.4 mamy $\phi(A) \neq 0$ oraz $\phi(B) \neq 0$, więc $\phi(AB) \neq 0$. W szczególności postacią schodkową zredukowaną A oraz B jest I. Mówiąc inaczej: istnieją macierze operacji elementarnych $M_1, \dots, M_s$ oraz $N_1, \dots, N_t$ takie, że

  $$I = M_1 M_2 M_3 \dots M_s A, \quad I = N_1 N_2 N_3 \dots N_t B.$$

  Zatem $A = M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1}$, $\quad B = N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}$. Ale M_i^{-1} oraz N_j^{-1} to macierze operacji elementarnych, więc z Obserwacji 12.7, a dokładniej formuły ($\diamondsuit$):

  $$\begin{aligned}\phi(AB) &= \phi(M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1} N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}) = \\ &= \phi(M_s^{-1}) \cdot \phi(M_{s-1}^{-1}) \cdot \dots \cdot \phi(M_1^{-1}) \cdot \phi(N_t^{-1}) \cdot \phi(N_{t-1}^{-1}) \cdot \dots \cdot \phi(N_1^{-1}) = \\ &= \phi(M_s^{-1} M_{s-1}^{-1} \dots M_1^{-1}) \cdot \phi(N_t^{-1} N_{t-1}^{-1} \dots N_1^{-1}) = \phi(A)\phi(B).\end{aligned}$$

Pozostało dokończyć uzasadnienie Twierdzenia 12.3. Mianowicie twierdzimy, że wartość funkcji ϕ spełniającej $(1)-(4)$ (w tym, jak wiemy, funkcji $\phi = \det$) jest jednoznacznie wyznaczona, dla każdej macierzy $A \in M_n(K)$. Rzeczywiście:

- Jeśli A nie jest odwracalna, to $\phi(A) = 0$, zgodnie z Twierdzeniem 12.4.
- Jeśli A jest odwracalna to algorytm Gaussa podaje jednoznaczny, najkrótszy możliwy ciąg operacji elementarnych pozwalających na sprowadzenie A do postaci zredukowanej I. Niech macierze tych operacji to $M_1, \dots, M_s$. Na mocy wzoru Cauchy'ego:

  $$1 = \phi(I) = \phi(M_s)\phi(M_{s-1}) \dots \phi(M_1)\phi(A).$$

  Zatem gdy A jest odwracalna, to

  $$\det \phi = (\phi(M_s)\phi(M_{s-1}) \dots \phi(M_1))^{-1},$$

  gdzie $M_1, \dots, M_s$ jest jednoznacznie wyznaczonym ciągiem macierzy operacji elementarnych. Skoro znamy $\phi(M_i)$, to $\phi(A)$ jest wyznaczona jednoznacznie, co kończy dowód Twierdzenia 12.3.

Odnotujmy ważny wniosek ze wzoru Cauchy'ego, kluczowy do dowodu własności (†).

[1] Przykładowe argumenty: (1) AB to macierz izomorfizmu będącego złożeniem przekształceń o macierzach A oraz B, co oznacza, że A to macierz monomorfizmu, a B – macierz epimorfizmu. Ale te przekształcenia działają pomiędzy przestrzeniami wymiaru n, więc to izomorfizmy., że skoro $A, B \in M_n(K)$. Zatem A, B są odwracalne. (2) Mamy $r(AB) \leqslant \min\{r(A), r(B)\}$.

Obserwacja 12.9

Załóżmy, że funkcja $\phi : M_n(K) \to K$ spełnia warunki (1)-(4) (np. $\phi = \det$, o czym już wiemy). Dla każdej $A \in M_n(K)$ mamy $\phi(A) = \phi(A^T)$.

Dowód. Korzystamy z faktu, że $r(A) = r(A^T)$. Rozważamy dwa przypadki.

- Jeśli $r(A) < n$, to $r(A^T)$, czyli obie macierze nie są odwracalne i ich wartości na ϕ (w szczególności – wyznaczniki) są równe 0.
- Jeśli $r(A) = n$, to A rozkłada się na iloczyn macierzy operacji elementarnych

  $$A = M_1 M_2 \dots M_s.$$

  Zatem zgodnie ze wzorem $(XY)^T = Y^T X^T$ mamy:

  $$A^T = M_s^T M_{s-1}^T \dots M_1^T.$$

  Łatwo sprawdzić, że dla każdej macierzy operacji elementarnej M mamy

  $$\phi(M) = \phi(M^T).$$

  Rzeczywiście, dla macierzy operacji typu (2) i (3) po prostu mamy $M = M^T$. Co do macierzy operacji (1) to przecież M^T jest również macierzą operacji typu (1), a wszystkie te macierze mają wartość ϕ (a więc też wyznacznika) równą 1, zgodnie z Obserwacją 12.7. Zatem z twierdzenia Cauchy'ego:

  $$\phi(A) = \phi(M_1)\phi(M_2)\dots\phi(M_s) = \phi(M_s^T)\phi(M_{s-1}^T)\dots\phi(M_1^T) = \phi(A^T).$$

□

Pozostało uzasadnić wzór (†) mówiący, że obliczanie wyznacznika za pomocą rozwinięcia względem dowolnego wiersza i dowolnej kolumny daje ten sam wynik (przy założeniu, że wiemy, że fakt ten zachodzi dla macierzy mniejszego rozmiaru i nie martwimy się jak policzyć wyznaczniki macierzy typu A_{ij}). Argumenty są następujące:

- Analogicznie jak w dowodach Twierdzenia 12.2 oraz Obserwacji 12.3 pokazujemy, że funkcje postaci $d_k : M_n(K) \to K$ określone przez rozwinięcie względem k-tej kolumny spełniają warunki (1)-(4), a zatem zgodnie z Twierdzeniem 12.3, funkcje z $M_n(K) \to K$ zadające rozwinięcia względem poszczególnych kolumn są równe.
- Korzystając z Obserwacji 12.9 zauważamy, że funkcja $w_k : M_n(K) \to K$ określona przez rozwinięcie względem k-tego wiersza równa jest funkcji określonej przez rozwinięcie względem k-tej kolumny. Istotnie, z założenia:

  $$\sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A_{ik})^T$$

  Wyrazy $a_{i1}, \dots, a_{in}$ to wyrazy i-tej kolumny macierzy A^T, zaś $(A_{ik})^T$ powstają z usunięcia z A^T k-tego wiersza i n-tej kolumny. A zatem po prawej stronie znajduje się $\det A^T = \det A$.

Wzór (†) jest zatem uzasadniony, z dokładnością do prostych powtórzeń dowodów. Wniosek jest następujący. Licząc wyznaczniki konkretnych macierzy możemy korzystać zamiennie z różnych rozwinięć: jeśli na przykład sprowadzimy przez rozwinięcie względem drugiej kolumny obliczenie wyznacznika macierzy 4×4 do obliczenia czterech wyznaczników macierzy 3×3, to każdy z tych czterech wyznaczników możemy liczyć za pomocą innego rozwinięcia – możemy korzystać zarówno z rozwinięć na wierszach i na kolumnach. Dla ścisłości – całe rozumowanie powyższe należy rozumieć indukcyjnie: najpierw stwierdzamy równość wszystkich (dostępnych) rozwinięć dla macierzy 2×2, a dalej korzystając z niej i równości (†), możemy postulować równość rozwinięć dla macierzy 3×3 rozumiejąc, że licząc w ramach tych różnych rozwinięć wyznaczniki macierzy typu A_{ij}, możemy już korzystać z dowolnego (dostępnego) rozwinięcia.

Dlaczego zaś posługiwaliśmy się funkcjami ϕ i własnościami (1)-(4)? Ta bowiem wyjątkowość wyznacznika wśród funkcji o tej włąsności ma dogłębne znaczenie geometryczne. Jego idea brzmi – z dokładnością do skalara, istnieje jedna (liniowa) funkcja objętości. O tym więcej napiszę w uzupełnieniu.

* * *

Na koniec opowemy o podstawowym zastosowaniu wyznacznika – historycznie rzecz biorąc – źródłowym dla jego powstania, a więc o rozwiązywaniu układów równań.

Rozważamy układ U złożony n równań liniowych z n niewiadomymi o współczynnikach w ciele K:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n. \end{cases} \qquad (\diamondsuit).$$

Wiemy już, że możemy ten układ zapisać w postaci iloczynu macierzy współczynników oraz wektora o współrzędnych złożonych ze zmiennych tak, by wynikiem była macierz o kolumnie z wyrazami b_i:

$$A \cdot X = B \qquad (\spadesuit),$$

gdzie

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Równanie typu $(\spadesuit)$ jest przykładem **równania macierzowego**. Rozmaite problemy algebraiczne można formułować w języku tych równań, a ich rozwiązywanie bywa trudne z uwagi na nieprzemienność mnożenia macierzy oraz to, że nie są one zawsze odwracalne (mogą one dotyczyć też macierzy prostokątnych). Rozpocznijmy od fundamentalnej obserwacji, wynikającej z Twierdzenia 12.4.

Obserwacja 12.10

Następujące warunki są równoważne:

- układ $(\diamondsuit)$ ma dokładnie jedno rozwiązanie,
- $\det A \neq 0$,
- macierz A jest odwracalna.

Gdy zachodzi dowolny z powyższych warunków, to $X = A^{-1}B$.

Obserwacja ta pozwala sformułować **metodę macierzową rozwiązywania układów równań**, których macierz współczynników ma niezerowy wyznacznik.

Zobaczmy przykład. Rozważmy układ równań o współczynnikach w $\mathbb{R}$ postaci:

$$\begin{cases} x_1 + x_2 + x_3 & = 6 \\ \quad 2x_2 + 5x_3 & = -4 \\ 2x_1 + 5x_2 - x_3 & = 27 \end{cases}.$$

Równanie macierzowe równoważne powyższemu układowi ma postać:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}.$$

Wyznaczamy teraz macierz odwrotną do macierzy A współczynników. Możemy to zrobić korzystając z algorytmu przedstawionego na poprzednich wykładach, to znaczy: za pomocą elementarnych operacji na wierszach sprowadzić macierz $[A \,|\, I]$ do macierzy $[I \,|\, A^{-1}]$, uzyskując:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix}^{-1} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}.$$

Znamy metodę wyznaczania macierzy odwrotnej przy pomocy operacji elementarnych na wierszach. Pokażemy teraz metodę opartą o wyznacznik i tzw. macierz stowarzyszoną (inaczej: dołączoną).

Definicja 12.6: Macierz stowarzyszona

Załóżmy, że $A \in M_n(K)$. Macierzą stowarzyszoną z A definiujemy następująco:

$$\operatorname{adj}(A) = [(-1)^{i+j} \det(A_{ij})]^T = \begin{bmatrix} (-1)^{1+1}|A_{11}| & (-1)^{1+2}|A_{12}| & \dots & (-1)^{1+n}|A_{1n}| \\ (-1)^{2+1}|A_{21}| & (-1)^{2+2}|A_{22}| & \dots & (-1)^{2+n}|A_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1}|A_{n1}| & (-1)^{n+2}|A_{n2}| & \dots & (-1)^{n+n}|A_{nn}| \end{bmatrix}^T.$$

Przykład.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \operatorname{adj}(A) = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Zauważmy też, że w powyższym przypadku:

$$\operatorname{adj}(A) \cdot A = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = |A| \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Twierdzenie 12.6

Zachodzi równość $\operatorname{adj} A \cdot A = \det(A) \cdot I_n$. W szczególności, jeśli A jest macierzą odwracalną, to

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj}(A).$$

Dowód. Niech $A = [a_{ij}]$, dla $1 \leqslant i, j \leqslant n$. Mnożymy $\operatorname{adj}(A)$ przez A, czyli

$$\begin{bmatrix} (-1)^{1+1} \det A_{11} & (-1)^{2+1} \det A_{21} & \dots & (-1)^{n+1} \det A_{n1} \\ (-1)^{1+2} \det A_{12} & (-1)^{2+2} \det A_{22} & \dots & (-1)^{2+n} \det A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1} \det A_{1n} & (-1)^{n+2} \det A_{2n} & \dots & (-1)^{n+n} \det A_{nn} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Wymnóżmy i-ty wiersz $\operatorname{adj}(A)$ oraz j-tą kolumnę w A. Mamy:

$$(-1)^{i+1} \det A_{1i} \cdot a_{1j} + (-1)^{i+2} \det A_{2i} \cdot a_{2j} + \ldots + (-1)^{n+i} \det A_{ni} \cdot a_{nj}, \qquad (\dagger)$$

To wyrażenie wygląda prawie jak wzór na wyznacznik w rozwinięciu Laplace'a względem i-tej kolumny macierzy A z tym, że zamiast wyrazów z i-tej kolumny macierzy A w poszczególnych składnikach pojawiają się wyrazy z j-tej kolumny. Możemy jednak powiedzieć, że $(\dagger)$ to wyznacznik macierzy D_{ij} powstaje z A przez zastąpienie j-tej kolumny kolumną i-tą (wystarczy policzyć wyznacznik D_{ij} rozwijając względem i-tej kolumny. Zauważmy jednak, że jeśli $i \neq j$, to D_{ij} ma dwie identyczne kolumny. czyli:

$$\det D_{ij} = \begin{cases} \det A, & \text{dla } i = j \\ 0, & \text{dla } i \neq j. \end{cases}.$$

W rezultacie:

$$\operatorname{adj}(A) \cdot A = \begin{bmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det A \end{bmatrix},$$

co kończy dowód. □

Jesteśmy również gotowi do sformułowania wzorów pozwalających na uzyskanie rozwiązania równania $(\diamondsuit)$ w przypadku, gdy jest ono jedyne.

Twierdzenie 12.7: Wzory Cramera

Niech U będzie układem n równań liniowych z n niewiadomymi

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n. \end{cases} \qquad (\diamondsuit).$$

o macierzy współczynników $A \in M_n(K)$ i kolumnie wyrazów wolnych $B \in M_{n\times 1}(K)$. Załóżmy, że $\det A \neq 0$. Wówczas układ U ma dokładnie jedno rozwiązanie $s_1, \ldots, s_n$, przy czym dla każdego i mamy

$$s_i = \frac{\det G_i}{\det A},$$

gdzie G_i jest macierzą powstałą z A przez zastąpienie i-tej kolumny kolumną B.

Zobaczmy, dla przykładu, układ równań nad $\mathbb{Q}$ postaci

$$\begin{cases} x + y &= 2 \\ x - y &= 0 \end{cases}.$$

Jeśli A jest macierzą współczynników tego układu to to zgodnie z definicją G_i oraz wzorami Cramera:

$$|A| = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}, \quad |G_1| = \begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix}, \quad |G_2| = \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} \quad \Rightarrow \quad x = \frac{|G_1|}{|A|} = 1, \quad y = \frac{|G_2|}{|A|} = 1.$$

Dowód. Jak wiemy z metody macierzowej, aby rozwiązać równanie $AX = B$ powstałe z równania $(\diamondsuit)$ należy wykonać następujące mnożenie:

$$A^{-1}A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad \Rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A^{-1} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Ale macierz A^{-1} ma wyrazy c_{ij} postaci:

$$(-1)^{j+i}\frac{\det A_{ji}}{\det A},$$

czyli jeśli G_i to macierz powstała z A przez zamianę i-tej kolumny na B, to argumentując podobnie jak w poprzednim dowodzie widzimy, że iloczyn i-tego wiersza macierzy A^{-1} przez kolumnę B równy jest

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det A} \begin{bmatrix} (-1)^{1+1}\det A_{11} & (-1)^{2+1}\det A_{21} & \ldots & (-1)^{n+1}\det A_{n1} \\ (-1)^{1+2}\det A_{12} & (-1)^{2+2}\det A_{22} & \ldots & (-1)^{2+n}\det A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{1+n}\det A_{1n} & (-1)^{2+n}\det A_{2n} & \ldots & (-1)^{n+n}\det A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

czyli istatnia równość ma postać:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \dfrac{\sum_{s=1}^{n}(-1)^{s+1}\det A_{s1}b_s}{\det A} \\ \vdots \\ \dfrac{\sum_{s=1}^{n}(-1)^{s+n}\det A_{sn}b_s}{\det A} \end{bmatrix} = \begin{bmatrix} \dfrac{\sum_{s=1}^{n}(-1)^{s+1}\det (G_1)_{s1}b_s}{\det A} \\ \vdots \\ \dfrac{\sum_{s=1}^{n}(-1)^{s+n}\det (G_n)_{sn}b_s}{\det A} \end{bmatrix} = \begin{bmatrix} \dfrac{\det G_1}{\det A} \\ \vdots \\ \dfrac{\det G_n}{\det A} \end{bmatrix}.$$

□

Wzory Cramera są uzasadnione, co zamyka przegląd podstawowych pojęć związanych z wyznacznikiem. W drugim semestrze wielokrotnie przekonamy się o potędze tego pojęcia, nie tylko w kontekście metod algorytmicznych, ale też geometrycznych, co wstępnie szkicujemy w uzupełnieniu.

12.2 Zadania do samodzielnej pracy

Zadania ilustrujące umiejętności niezbędne do zaliczenia przedmiotu oznaczone są symbolem ♠.
Przy tych zadaniach dodany jest opis umiejętności, które sprawdzają.

1. (♠ Obliczanie wyznacznika za pomocą rozwinięcia Laplace'a) Oblicz wyznaczniki macierzy:

$$A_1 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 3 & 1 \\ 1 & 0 & 2 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 4 & 3 & 6 & 0 \\ 7 & 9 & 2 & 8 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 2 & 1 & 1 & 2 \\ 3 & 4 & 5 & 4 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

2. (♠ Obliczanie wyznacznika przez sprowadzenie do postaci trójkątnej) Oblicz wyznaczniki:

$$\begin{vmatrix} 3 & 4 & 2 & 2 \\ 4 & 5 & 6 & 5 \\ 2 & 3 & 6 & 0 \\ 8 & 7 & 7 & 8 \end{vmatrix}, \quad \begin{vmatrix} 36 & 60 & 72 & 37 \\ 43 & 71 & 78 & 34 \\ 44 & 69 & 73 & 32 \\ 30 & 50 & 65 & 38 \end{vmatrix}, \quad \begin{vmatrix} 35 & 59 & 71 & 52 \\ 42 & 70 & 77 & 54 \\ 43 & 68 & 72 & 52 \\ 29 & 49 & 65 & 50 \end{vmatrix}.$$

3. (♠ Obliczanie wyznacznika za pomocą rozwinięcia Laplace'a i indukcji matematycznej)
Oblicz wyznaczniki następujących macierzy rozmiaru $n \times n$:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}, \quad A_5 = \begin{bmatrix} 1 & i & 0 & \cdots & 0 & 0 \\ i & 1 & i & \cdots & 0 & 0 \\ 0 & i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & i \\ 0 & 0 & 0 & \cdots & i & 1 \end{bmatrix}, \quad A_6 = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 2 & \cdots & 2 & 2 \\ 1 & 2 & 3 & \cdots & 3 & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & 3 & \cdots & n-1 & n-1 \\ 1 & 2 & 3 & \cdots & n-1 & n \end{bmatrix}.$$

4. Oblicz wyznaczniki następujących macierzy $n \times n$ w zależności od parametrów $s, t \in \mathbb{R}$

$$C_1 \begin{bmatrix} 0 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & s & s & \dots & s \\ 1 & s & 0 & s & \dots & s \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & s & s & s & \dots & s \\ 1 & s & s & s & \dots & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 2 & t & 0 & 0 & \dots & 0 \\ 0 & 2 & t & 0 & \dots & 0 \\ 0 & 0 & 2 & t & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & t \\ t & 0 & 0 & 0 & \dots & 2 \end{bmatrix}, \quad C_3 = \begin{bmatrix} s & t & t & t & \dots & t \\ t & s & t & t & \dots & t \\ t & t & s & t & \dots & t \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t & t & t & t & \dots & t \\ t & t & t & t & \dots & s \end{bmatrix}.$$

5. Czy poniższy wyznacznik jest niezerowy?

$$\begin{vmatrix} 102495 & 550429 & 873298 & 660697 \\ 370628 & 909093 & 127450 & 925601 \\ 835044 & 601178 & 624655 & 263392 \\ 663780 & 487252 & 292276 & 593107 \end{vmatrix}$$

6. Wyrazami macierzy kwadratowej A należącej do zbioru $M_4(\mathbb{R})$ są tylko liczby -2 oraz 1 (dowolnie ustawione). Pokaż, że wyznacznik macierzy A jest liczbą całkowitą podzielną przez 27.

7. Korzystając z własności macierzy $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ uzasadnić, że wyrazy $F_0, F_1, F_2, \dots$ ciągu Fibonacciego spełniają dla każdego $n \in \mathbb{N}$ równość $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$.

8. Wykaż, że jeśli w macierzy $n \times n$ na przecięciu k wierszy i l kolumn znajdują się same zera, przy czym $k + l > n$, to wyznacznik tej macierzy jest równy 0.

9. Oblicz wyznaczniki następujących macierzy (por. dodatek o wyznacznikach macierzy blokowych)

$$C_1 = \begin{bmatrix} 7 & 9 & 8 & 2 \\ 4 & 6 & 7 & 0 \\ 8 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 0 & 3 & 6 & 7 \\ 7 & 9 & 2 & 8 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 4 & 1 & 2 & 9 \\ 5 & 0 & 1 & 7 \\ 2 & 1 & 1 & 8 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad C_4 = \begin{bmatrix} 5 & 4 & 7 & 1 \\ 9 & 7 & 8 & 3 \\ 0 & 0 & 8 & 7 \\ 0 & 0 & 6 & 5 \end{bmatrix}.$$

10. (♠ Wyznacznik iloczynu i wyznacznik macierzy odwrotnej)
Oblicz $\det(A \cdot B)$, $\det(A^7)$, $\det(A^3 \cdot B^{-1})$ dla poniższych macierzy:

$$A = \begin{bmatrix} 6 & 1 & 0 & 4 \\ 2 & 0 & 0 & 0 \\ 7 & 0 & 0 & 1 \\ 6 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 5 & 6 & 6 \\ 1 & 5 & 9 & 7 \end{bmatrix}.$$

11. Niech A będzie macierzą rozmiaru 3×3 nad ciałem $\mathbb{Q}$. Wiadomo, że $\det(A) = -1$. Oblicz $\det(2A)$.

12. Niech suma kolumn macierzy $A \in M_n(K)$ będzie wektorem zerowym. Czy $\det(A) = 0$?

13. Dana jest macierz $A \in M_{6\times 6}(\mathbb{R})$ o wyznaczniku $\det(A) = 7$. Załóżmy, ze wszystkie współczynniki A są całkowite. Czy macierz $21A^{-1}$ też musi mieć wszystkie wyrazy całkowite?

14. Macierz $A \in M_{n\times n}(\mathbb{R})$ ma w każdym wierszu dokładnie jeden niezerowy wyraz, równy x i w każdej kolumnie dokładnie jeden niezerowy wyraz. Ile wynosi $|\det(A)|$?

15. Czy istnieją macierze $A \in M_{4\times 3}(\mathbb{R})$ i $B = M_{3\times 4}(\mathbb{R})$ takie, że $\det(AB) = 1$?

16. Pewna macierz $A \in M_{n\times n}(\mathbb{R})$ spełnia równanie $A^2 + A = I_n$. Wykaż, że macierz A jest odwracalna.

17. Niech $n \geqslant 1$. Załóżmy, że macierz $A \in \mathrm{M}_n(\mathbb{C})$ spełnia $A^T \cdot A = I$ oraz $\det A < 0$. Oblicz $\det(I + A)$.

18. (♠ Wyznacznikowe kryterium odwracalności i wzór na macierz odwrotną) Dla jakich wartości parametru $s \in \mathbb{R}$ poniższa macierz

$$A = \begin{bmatrix} 2 & 5 & 3 \\ 1 & s & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

jest odwracalna? Dla każdego takiego s znaleźć A^{-1}.

19. (♠ Stosowanie wzorów Cramera) Przedyskutuj rozwiązywalność następującego układu równań, w zależności od parametru $m \in \mathbb{R}$:

$$3x + z = 1, \quad mx - my + z = -m, \quad x + my + z = 3.$$

Dla tych m, dla których istnieją rozwiązania, wyznacz je.

20. Załóżmy, że układ n równań liniowych U o n zmiennych i o współczynnikach całkowitych ma dla dowolnej liczby pierwszej p dokładnie jedno rozwiązanie w ciele $\mathbb{Z}_p$ (współczynniki traktujemy modulo p). Czy układ równań U posiada rozwiązanie, którego wszystkie współrzędne są całkowite?

21. Załóżmy, że $A \in M_{n\times n}(\mathbb{Z})$ oraz $\det A = \pm 1$. Wykaż, że $A^{-1} \in M_{n\times n}(\mathbb{Z})$. Czy to prawda, gdy $\det A = 2$?

22. Niech $n \geqslant 1$. Załóżmy, że macierz $A = [a_{ij}] \in \mathrm{M}_n(\mathbb{R})$ spełnia $a_{ij} = 1$ dla $i \neq j$ oraz $a_{11}, \ldots, a_{nn} \geqslant 2$. Udowodnij, że $\det A \geqslant n + 1$.

23. Jeśli a_1, a_2 są różnymi liczbami zespolonymi, wówczas prosta przechodząca przez te punkty opisana jest przez liczby zespolone z spełniające równanie:

$$\begin{vmatrix} z & \overline{z} & 1 \\ a_1 & \overline{a_1} & 1 \\ a_2 & \overline{a_2} & 1 \end{vmatrix} = 0.$$

24. Wykaż, że jeśli a_1, a_2, a_3 są parami różnymi punktami na płaszczyźnie zespolonej, to okrąg przechodzący przez te punkty złożony jest z punktów z spełniających równość:

$$\begin{vmatrix} z\overline{z} & z & \overline{z} & 1 \\ a_1\overline{a_1} & a_1 & \overline{a_1} & 1 \\ a_2\overline{a_2} & a_2 & \overline{a_2} & 1 \\ a_3\overline{a_3} & a_3 & \overline{a_3} & 1 \end{vmatrix} = 0.$$

12.3 Objętość, orientacja i wzór permutacyjny na wyznacznik

Zdefiniowaliśmy na ostatnim wykładzie wyznacznik, czyli funkcję zachowującą się dobrze przy wykonywaniu operacji elementarnych oraz przy mnożeniu macierzy. Funkcję, która w łatwy sposób pozwala sprawdzić odwracalność macierzy oraz które (już niezbyt łatwo i efektywnie) pozwala bez użycia operacji elementarnych odwracać macierze i rozwiązywać układy równań. W tym materiale przyjrzymy się, na razie zupełnie wstępnie, geometrycznemu kontekstowi pojęcia wyznacznika. Powiemy przy tym także o ważnych jego własnościach algebraicznych, zwłaszcza wieloliniowości i antysymetryczności. Zaczniemy od definicji mającej na ten moment charakter bardziej poglądowy niż formalny (w następnym semestrze rozważać ją będziemy we właściwym kontekście).

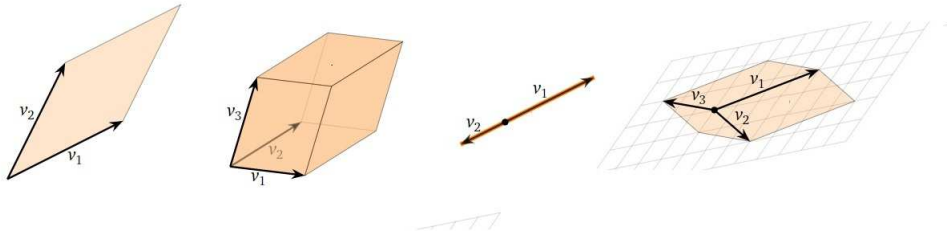
Definicja 12.7

RÓWNOLEGŁOŚCIANEM rozpiętym na wektorach $v_1, \ldots, v_n \in \mathbb{R}^n$ nazywamy podzbiór

$$R(v_1, \ldots, v_n) = \{a_1 v_1 + \ldots + a_n v_n \,|\, 0 \leqslant a_1, \ldots, a_n \leqslant 1\}.$$

Zauważmy, że powyższa definicja zakłada, że każdy równoległościan zawiera wektor zerowy. W drugim semestrze rozszerzymy tę definicję na odpowiednie podzbiory (euklidesowych) przestrzeni afinicznych. Łatwo widzieć, że dla przestrzeni niskich wymiarów równoległościany utożsamiać można z obiektami znanymi z geometrii. Na przykład równoległościan $R(v_1)$ utożsamiać można, dla niezerowego wektora v_1, z odcinkiem, zaś $R(v_1, v_2)$ utożsamić można dla nieproporcjonalnych v_1, v_2 – z równoległobokiem. Nierówności $0 \leqslant a_1, \ldots, a_n \leqslant 1$ mają sens dla ciała $\mathbb{R}$, choć można to założenie na różne sposoby osłabiać.

Poniżej przedstawionych jest kilka ilustracji równoległościanów, nawiązujących do definicji szkolnych. Pierwsze dwa (od lewej) rozpięte są przez układy liniowo niezależne, a kolejne dwa (odpowiednio w przestrzeniach dwu- i trójwymiarowej) rozpięte są przez układy liniowo zależne. Czym różnią się te sytuacje?



Dlaczego formułujemy taką definicję, zwłaszcza skoro dotyczy ona jedynie przestrzeni nad ciałem $\mathbb{R}$? Powodów algebraicznych jest kilka, ale nam chodzi o podkreślenie następującego zjawiska.

Obserwacja 12.11

Obraz równoległościanu przy przekształceniu liniowym przestrzeni liniowej $\mathbb{R}^n$ w siebie jest równoległościanem. Innymi słowy, jeśli $\phi : \mathbb{R}^n \to \mathbb{R}^n$ to $R(\phi(v_1), \ldots, \phi(v_n)) = \phi(R(v_1, \ldots, v_n))$.

Dowód. Rzeczywiście, jeśli $\phi : \mathbb{R}^n \to \mathbb{R}^n$, to obrazem wektora $a_1 v_1 + \ldots + a_n v_n \in R(v_1, \ldots, v_n)$ jest

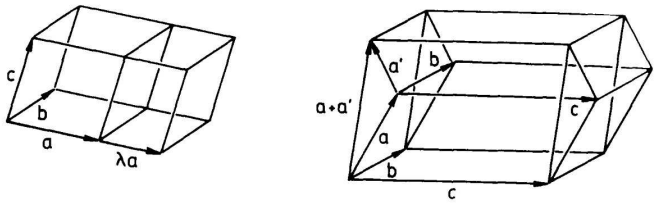
$$a_1 \phi(v_1) + \ldots + a_n \phi(v_n) \in R(\phi(v_1), \ldots, \phi(v_n)).$$

A zatem $\phi(R(v_1, \ldots, v_n)) \subseteq R(\phi(v_1), \ldots, \phi(v_n))$. Z drugiej strony, biorąc układ $0 \leqslant c_1, \ldots, c_n \leqslant 1$ widzimy, że $c_1\phi(v_1) + \ldots + c_n\phi(v_n)$ jest obrazem wektora $c_1 v_1 + \ldots + c_n v_n$ przy ϕ, należącego do $R(v_1, \ldots, v_n)$, zatem dostajemy drugą inkluzję $R(\phi(v_1), \ldots, \phi(v_n)) \subseteq \phi(R(v_1, \ldots, v_n))$. □

Widzimy zatem, że w przypadku przestrzeni nad ciałem $\mathbb{R}$ podprzestrzenie liniowe nie są jedynym „typem podzbiorów" zachowywanych przy przekształceniach liniowych (jak wiemy – obraz podprzestrzeni przy przekształceniu liniowym jest podprzestrzenią). Z punktu widzenia geometrii równoległościany mają znaczenie w teorii zbiorów wypukłych. Z analitycznego czy topologicznego punktu widzenia – są to obiekty służące do (trywializując nieco temat) „opisu przez przybliżenia" bardziej skomplikowanych

zbiorów. Podstawowym pojęciem opisującym równoległościany jest tzw. objętość (czy raczej: miara). Stanowi ono uogólnienie szkolnych pojęć długości, pola czy objętości. Poznamy je w drugim semestrze.

Spójrzmy na ilustrację równoległościanów $R(\lambda \cdot a, b, c)$ oraz $R(a + a', b, c)$, gdzie $a, b, c \in \mathbb{R}^3$ oraz $\lambda > 0$



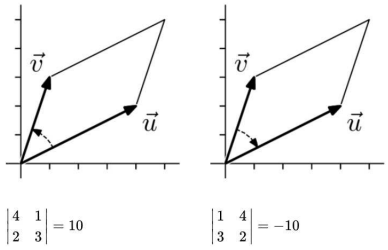
Rysunek 2. Źródło: K. Spindler, Abstract Algebra with Applications

Stosując znane fakty z geometrii szkolnej można pokazać, że „objętość" równoległościanu $R(\lambda \cdot a, b, c)$ równa jest λ razy „objętość" $R(a, b, c)$, zaś „objętość" $R(a+a', b, c)$ równa jest sumie „objętości" $R(a, b, c)$ oraz $R(a', b, c)$. Dlaczego stosujemy cudzysłów? Czy są różne „objętości"? Na ten moment naszym celem jest wyabstrahowanie własności „objętości" traktowanej jako funkcja na układzie wektorów i przyjrzenie się funkcjom z $M_n(K)$ do K mającym analogiczne własności. Ograniczymy się do konwencji, w której wektory rozpinające równoległościan są wierszami macierzy rozmiaru $n \times n$, a badane funkcje „o własnościach objętości" zachowują się w odpowiedni sposób. Jak się okaże (po porządnym wysiłku) w zasadzie istnieje tylko jedna taka funkcja. Skutki tego faktu wykraczają daleko poza algebrę liniową.

Przypomnijmy jeden z przykładów z poprzedniego wykładu. Mamy:

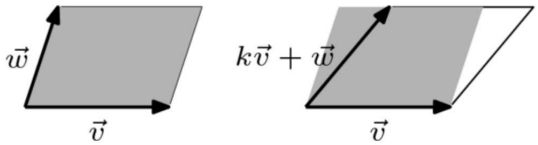
$$\det \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} = (-1)^{1+1} \cdot 1 \cdot \det A_{11} + (-1)^{2+1} \cdot 3 \cdot \det A_{21} = 1 \cdot 2 - 3 \cdot 4 = -10.$$

Ten wynik może budzić niepokój. Właśnie dowiedzieliśmy się, że funkcja mająca grać rolę objętości przyjmuje nad $\mathbb{R}$ ujemne wartości. To może wydawać się dziwne, ale za jakiś czas to się wyjaśni. Wiązać się to będzie z tzw. orientacją układu wektorów. Na razie intuicji niech dostarczy poniższy obrazek związany z pojęciem tzw. pola skierowanego.



Rysunek 3. Źródło: Jim Hefferon, `https://hefferon.net/linearalgebra/`.

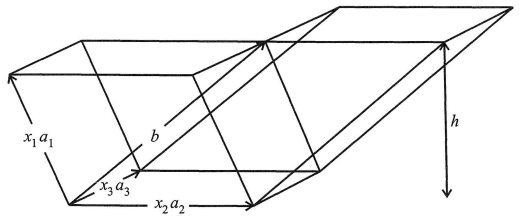
Geometrycznie mówiąc macierz diagonalna o niezerowych wyrazach na diagonali reprezentuje równoległościan o prostopadłych krawędziach. Sugeruje to, że wyznacznik macierzy diagonalnej jest iloczynem wyrazów na jej przekątnej. Tak istotnie jest, i to nie tylko dla macierzy diagonalnej, jak zdążyliśmy się już przekonać. Również wykonywanie operacji elementarnych na wierszach macierzy można interpretować w języku geometrycznym. Zachęcam w tym celu do kontemplacji poniższego obrazka:



Rysunek 4. Źródło: Jim Hefferon, `https://hefferon.net/linearalgebra/`

Nawet wzory Cramera wyglądające dość abstrakcyjnie i niebędące przydatne z punktu widzenia rachunków mają ciekawą interpretację geometryczną w przestrzeni trójwymiarowej. Niech $A \in M_3(\mathbb{R})$ będzie macierzą odwracalną o kolumnach a_1, a_2, a_3 i rozważmy wektor $b \in \mathrm{lin}(a_1, a_2, a_3)$ tak, że układ $Ax = b$ ma dokładnie jedno rozwiązanie. Innymi słowy, istnieją jednoznacznie wyznaczone $x_1, x_2, x_3 \in \mathbb{R}$ takie,

że $x_1a_1 + x_2a_2 + x_3a_3 = b$. Przyjmijmy, dla uproszczenia, że $\det A > 0$, $x_1, x_2, x_3 > 0$ i rozważmy równoległościany $R = R(x_1a_1, x_2a_2, x_3a_3)$ oraz $R_1 = R(b, x_2a_2, x_3a_3)$:



Rysunek 5. Źródło: A Geometric Interpretation of Cramer's Rule, Gregory Conner and Michael Lundquist

Równoległościan $R(x_2a_2, x_3a_3)$ traktować możemy jako wspólną podstawę obydwu tych równoległościanów. Mają one również wspólną wysokość opuszczoną na tę podstawę. W tym momencie nie mamy narzędzi by to ściśle uzasadnić, ale geometrycznie sprawa jest oczywista: ściany w R, R_1 równoległe do wspólnej podstawy leżą w równoległej do niej płaszczyźnie rozpiętej przez wektory a_2, a_3 i przesuniętej względem podstawy o wektor x_1a_1 (to „przesunięcie" sprawia, że jest to płaszczyzna afiniczna, o czym będziemy się uczyć w przyszłym semestrze). Stąd, na mocy naszej (intuicyjnej na razie) interpretacji wyznacznika jako objętości (z dokładnością do wartości bezwzględnej, ale korzystamy z $x_1, x_2, x_3 > 0$):

$$\det[x_1a_1 \;\; x_2a_2 \;\; x_3a_3] = \det[b \;\; x_2a_2 \;\; x_3a_3].$$

Stąd korzystając z własności wyznacznika mamy $x_1x_2x_3 \det[a_1 \;\; a_2 \;\; a_3] = x_2x_3 \det[b \;\; a_2 \;\; a_3]$. A zatem:

$$x_1 = \frac{\det[b \;\; a_2 \;\; a_3]}{\det[a_1 \;\; a_2 \;\; a_3]}.$$

Ten dodatek poświęcimy algernatywnej definicji wyznacznika nazywanej wzorem permutacyjnym (Leibniza). Idea jest prosta – zamiast skomplikowanej rekurencyjnej definicji dostajemy otwarty wzór, pozwalający zresztą wykazać łatwo wiele podstawowych, znanych nam już własności wyznacznika. Kłopot stanowi dość skomplikowana postać wzoru, wymagająca wstępnych wyjaśnień i stosunkowo złożonej notacji. Zapewne nie będą Państwo często liczyć wyznacznika właśnie za pomocą tej formuły. Wielokrotnie jednak definiować będziemy w drugim semestrze funkcje, właśnie w oparciu o wyznaczniki. Będą to najpierw pewne niezmienniki przekształceń liniowych, a później też niezmienniki przestrzeni liniowych wyposażonych w dodatkową strukturę (i przekształceń liniowych zachowujących te dodatkowe struktury). Rozumienie własności tych funkcji ma fundamentalne znaczenie w algebrze linowej. Nie odejdziemy więc zupełnie od kontekstu geometrycznego. Nieśmiałe aluzje do pojęcia „objętości ze znakiem" nabiorą już dziś nowego sensu. Z jednej strony domykamy więc teorię kluczowego dla nas pojęcia – z drugiej zaś przygotowujemy grunt do niezwykle intensywnej pracy, która dopiero przez nami.

Przypomnijmy najpierw notację kolumnową, którą będziemy dziś stosować. Niech $A_1, \ldots, A_n$ będą kolumnami macierzy $A = [a_{ij}] \in M_n(K)$. Wówczas pisać będziemy

$$A = [A_1, \ldots, A_n].$$

Kolumny macierzy identycznościowej oznaczamy (kolejno) przez $E_1, \ldots, E_n$. Przypomnijmy definicje, które podane zostały już w jednym z dodatków do wcześniejszych wykładów.

Definicja 12.8: Permutacja

PERMUTACJĄ zbioru n-elementowego nazwiemy dowolną bijekcję $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Zbiór takich bijekcji oznaczamy przez S_n. Każdej permutacji $\sigma \in S_n$ odpowiada MACIERZ PERMUTACJI:

$$E_\sigma = [E_{\sigma(1)}, E_{\sigma(2)}, \ldots, E_{\sigma(n)}].$$

Liczbę $\det E_\sigma \in \{1, -1\}$ nazywamy ZNAKIEM PERMUTACJI σ, ozn. $\operatorname{sgn}(\sigma)$.

Przykład. Rozważmy permutację $\sigma \in S_4$ daną wzorem:

$$\sigma(1) = 3, \quad \sigma(2) = 2, \quad \sigma(3) = 4. \quad \sigma(4) = 1.$$

Notacja *tabelkowa* oraz postać macierzowa tej permutacji mają odpowiednio postać:

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}, \quad E_\sigma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

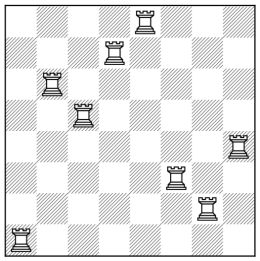
Znak tej permutacji liczymy sprawdzając liczbę zamian kolumn potrzebną do uzyskania macierzy identycznościowej – parzysta liczba zamian oznacza wyznacznik równy 1, zaś nieparzysta liczba zamian – wyznacznik równy -1 (łatwo pokazać, że możliwości te wykluczają się dla $1+1 \neq 0$). Zatem $\operatorname{sgn}(\sigma) = 1$.

Czytelnik może znać nieco inne definicje znaku permutacji, oparte o pojęcia nieporządków, rozkłady na cykle itd. Ten czysto kombinatoryczny aspekt można w naszych rozważaniach pominąć dysponując bogatą maszynerią dotyczącą wyznacznika. Odnotujmy jednak choćby jedną zasadniczą kwestię.

Elementy zbioru S_n można składać (tak, jak składa się funkcje). tzn. jeśli $\sigma, \rho \in S_n$ to $\rho \circ \sigma \in S_n$. Można sprawdzić, że wraz z permutacją identycznościową $\mathrm{id} \in S_n$ zbiór S_n tworzy GRUPĘ, tzn. $(S_n, \circ, \mathrm{id})$ jest zbiorem z działaniem dwuargumentowym $\circ$ spełniającym (w sposób oczywisty) następujące warunki:

- działanie $\circ$ jest łączne,
- dla każdego $\sigma \in S_n$ mamy $\sigma \circ \mathrm{id} = \mathrm{id} \circ \sigma = \sigma$,
- dla każdego $\sigma \in S_n$ istnieje $\rho \in S_n$ takie, że $\sigma \circ \rho = \rho \circ \sigma = \mathrm{id}$.

Z punktu widzenia teorii wyznaczników przydatna będzie dla Państwa następująca obrazowa intuicja. Rozważmy takie rozstawienie wież na szachownicy (a_{ij}) rozmiaru $n \times n$, by w każdym wierszu i kolumnie znajdowała się dokładnie jedna wieża. Permutacji $\sigma \in S_n$ odpowiada jedno z $n!$ różnych rozstawień.



W powyższym przykładzie dla $n = 8$ wieże rozstawione są na miejscach:

$$a_{\sigma(1)1}, a_{\sigma(2)2}, a_{\sigma(3)4}, a_{\sigma(4)4}, a_{\sigma(5)5}, a_{\sigma(6)6}, a_{\sigma(7)7}, a_{\sigma(8)8},$$

czyli:

$$a_{81}, a_{32}, a_{43}, a_{24}, a_{15}, a_{66}, a_{77}, a_{58}.$$

Jesteśmy gotowi do wprowadzenia tytułowego wzoru.

Twierdzenie 12.8

Dla macierzy $A = [a_{ij}] \in M_n(K)$ zachodzi WZÓR PERMUTACYJNY:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}.$$

Przykład 1. Dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

rozmiaru 2×2 mamy dwie możliwe permutacje kolumn reprezentowane przez następujące rozstawienia wież:

Dokładniej, $S_2 = \{\sigma_1, \sigma_2\}$, gdzie $\sigma_1(1) = 1, \sigma_1(2) = 2$ oraz $\sigma_2(1) = 2, \sigma_2(2) = 1$. Oczywiście widzimy, że $\operatorname{sgn}(\sigma_1) = 1, \operatorname{sgn}(\sigma_2) = -1$, skąd

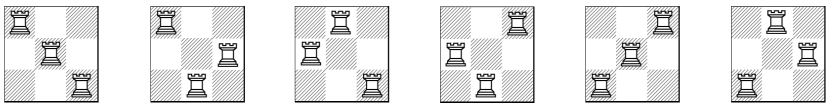
$$\det A = 1 \cdot a_{11}a_{22} + (-1)a_{21}a_{12}.$$

Przykład 2. Dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

rozmiaru 3×3 mamy:

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} + a_{31}a_{12}a_{23}.$$



Ważne: jeśli dla $\sigma \in S_n$ *jakaś wieża stoi na zerze*, tzn. $a_{\sigma(i)i} = 0$, dla pewnego i, to odpowiedniego składnika $(= 0)$ nie wliczamy do obliczania wyznacznika. Popatrzmy na dwa kolejne przykłady.

Przykład 3. Dla macierzy

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

mamy, podobnie jak wyżej,

$$a_{\sigma(1)1}a_{\sigma(2)2}a_{\sigma(3)3}a_{\sigma(4)4}a_{\sigma(5)5} \neq 0 \Leftrightarrow \sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 4 & 1 \end{bmatrix} \Rightarrow \det(A) = -1.$$

Przykład 4. Dla macierzy górnotrójkątnej

$$\begin{bmatrix} a_{11} & * & * & \dots & * \\ 0 & a_{22} & * & \dots & * \\ 0 & 0 & a_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

$$a_{\sigma(1)1}a_{\sigma(2)2} \dots a_{\sigma(n)n} \neq 0 \Leftrightarrow \sigma(i) = i, \text{ dla } i = 1, 2, \dots, n. \Rightarrow \det A = a_{11}a_{22} \dots a_{nn}.$$

* * *

Dowodzimy formułę permutacyjną. Będziemy korzystać z tego, że wyznacznik jest jedyną funkcją $M_n(K) \to K$ spełniającą warunki: (1) jednorodność ze względu na k-tą kolumnę, (2) addytywność względem k-tej kolumny, (3) funkcja zeruje się, jeśli macierz ma identyczne dwie (sąsiednie) kolumny, (4) przyjmuje wartość 1 na macierzy I. Innymi słowy:

$$\det[A_1, \dots, A_{k-1}, B + C, A_{k+1}, \dots, A_n] = \det[A_1, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n] + \det[A_1, \dots, A_{k-1}, C, A_{k+1}, \dots, A_n],$$
$$\det[A_1, \dots, A_{k-1}, aC, A_{k+1}, \dots, A_n] = a \cdot \det[A_1, \dots, A_{k-1}, C, A_{k+1}, \dots, A_n].$$

Fakty te wynikają natychmiast z tego, że $\det X = \det X^T$.

Mamy też, dla i-tej kolumny macierzy A:

$$\begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix} = a_{1i} \cdot \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_{ni} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = a_{1i}E_1 + \dots + a_{ni}E_n.$$

Niech $A = [a_{ij}]$. Mamy:

$$\det[A_1, \ldots, A_n] = \det[a_{11}E_1 + \ldots + a_{n1}E_n, \ldots, a_{n1}E_1 + \ldots + a_{nn}E_n].$$

Korzystamy teraz z addytywności i jednorodności względem pierwszej kolumny dostając:

$$\begin{aligned}\det[A_1, \ldots, A_n] &= \det[a_{11}E_1 + \ldots + a_{n1}E_n, \ldots, a_{n1}E_1 + \ldots + a_{nn}E_n] = \\ &= a_{11}\det[E_1, a_{12}E_1 + \ldots + a_{n2}E_n, \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ a_{21}\det[E_2, a_{12}E_1 + \ldots + a_{n2}E_n, \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ \ldots + \\ &+ a_{n1}\det[E_n, a_{12}E_1 + \ldots + a_{n2}E_n, \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n].\end{aligned}$$

Teraz dla dwóch z otrzymanych n składników korzystamy z liniowości względem drugiej kolumny:

$$\begin{aligned}\det[A_1, \ldots, A_n] &= \det[a_{11}E_1 + \ldots + a_{n1}E_n, \ldots, a_{n1}E_1 + \ldots + a_{nn}E_n] = \\ &+ a_{11}a_{12}\det[E_1, E_1, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ a_{11}a_{22}\det[E_1, E_2, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ \ldots + \\ &+ a_{11}a_{n2}\det[E_1, E_n, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ a_{21}a_{12}\det[E_1, E_1, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ a_{21}a_{22}\det[E_1, E_2, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ \ldots + \\ &+ a_{21}a_{n2}\det[E_1, E_n, , \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n] + \\ &+ \ldots + \\ &+ a_{n1}\det[E_n, a_{12}E_1 + \ldots + a_{n2}E_n, \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n].\end{aligned}$$

Tą samą procedurę wykonujemy dla pozostałych $n-2$ składników, dostając n^2 składników postaci:

$$\det[A_1, \ldots, A_n] = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{i_1 1}a_{i_2 2}\det[E_{i_1}, E_{i2}, \ldots, a_{1n}E_1 + \ldots + a_{nn}E_n].$$

Teraz dla każdego z n^2 składników korzystamy z liniowości względem trzeciej kolumny, co da nam n^3 składników – i tak dalej aż otrzymamy przedstawienie w postaci n^n składników:

$$\det[A_1, \ldots, A_n] = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\ldots\sum_{i_n=1}^{n} a_{i_1 1}a_{i_2 2}\ldots a_{i_n n}\det[E_{i_1}, E_{i_2}, \ldots, E_{i_n}].$$

Zauważmy, że z tych n^n składników tylko $n!$ może być niezerowych – te, gdzie $\det[E_{i_1}, E_{i_2}, \ldots, E_{i_n}] \neq 0$. Tymczasem z własności (3) dostajemy:

$$\det[E_{i_1}, E_{i_2}, \ldots, E_{i_n}] = \begin{cases} 0, & \text{gdy } i_k \text{ nie są parami różne,} \\ \operatorname{sgn}(\sigma) & \text{dla } \sigma \in S_n : \sigma(k) = i_k. \end{cases}.$$

Na koniec powiemy o intuicjach geometrycznych, które rozwiniemy w drugim semestrze.

Definicja 12.9

Niech V będzie przestrzenią $\mathbb{R}^n$. Mówimy, że bazy

$$\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}, \quad \mathcal{B} = \{\beta_1, \ldots \beta_n\}$$

przestrzeni V są:

- ZGODNIE ZORIENTOWANE, jeśli $\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} > 0$,
- PRZECIWNIE ZORIENTOWANE, jeśli $\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} < 0$.

Przykład. Bazy

$$\mathcal{A} = \{(3,2),(7,4)\}, \quad \mathcal{B} = \{(1,2),(1,0)\}$$

są zgodnie zorientowane, bo:

$$\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} > 0.$$

Obserwacja 12.12

Zgodne zorientowanie jest relacją równoważności w zbiorze wszystkich baz w $\mathbb{R}^n$.

Dowód. Należy sprawdzić, że zgodne zorientowanie jest relacją zwrotną, symetryczną i przechodnią.

- **Zwrotność**. Jeśli $\mathcal{A}$ jest bazą $\mathbb{R}^n$ to

  $$M(\mathrm{id})_{\mathcal{A}}^{\mathcal{A}} = I \Rightarrow \det I = 1,$$

  czyli bazy $\mathcal{A}$ oraz $\mathcal{A}$ są zgodnie zorientowane.

- **Symetryczność**. Jeśli $\mathcal{A}, \mathcal{B}$ są bazami $\mathbb{R}^n$ to

  $$M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} = (M(\mathrm{id})_{\mathcal{B}}^{\mathcal{A}})^{-1},$$

  czyli

  $$\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} > 0 \Rightarrow \det M(\mathrm{id})_{\mathcal{B}}^{\mathcal{A}} > 0.$$

- **Przechodniość**. Jeśli $\mathcal{A}, \mathcal{B}, \mathcal{C}$ są bazami $\mathbb{R}^n$ to

  $$M(\mathrm{id})_{\mathcal{A}}^{\mathcal{C}} = M(\mathrm{id})_{\mathcal{B}}^{\mathcal{C}} \cdot M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}},$$

  czyli

  $$\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} > 0, \det M(\mathrm{id})_{\mathcal{B}}^{\mathcal{C}} > 0 \Rightarrow \det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{C}} > 0.$$

□

Uwaga: dla każdej bazy $\mathcal{A} = (\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$ przestrzeni $\mathbb{R}^n$ oraz bazy $\mathcal{A}' = (\alpha_2, \alpha_1, \alpha_3, \ldots, \alpha_n)$ powstałej przez zamianę kolejności wektorów na pierwszych dwóch współrzędnych mamy:

- bazy $\mathcal{A}$ oraz $\mathcal{A}'$ są przeciwnie zorientowane,
- każda baza $\mathbb{R}^n$ jest zgodnie zorientowana z $\mathcal{A}$ lub $\mathcal{A}'$.

Definicja 12.10: Orientacja rzeczywistej przestrzeni liniowej

Rodzinę wszystkich baz zgodnie zorientowanych z pewną bazą przestrzeni $\mathbb{R}^n$ nazywamy ORIENTACJĄ przestrzeni $\mathbb{R}^n$. Mówimy, że przestrzeń $\mathbb{R}^n$ jest ZORIENTOWANA, jeśli wybrana jest jedna z jej (dwóch) orientacji. W przestrzeni zorientowanej mówimy, że jej baza $\mathcal{A}$ jest DODATNIO (UJEMNIE) ZORIENTOWANA, jeśli zorientowana zgodnie (przeciwnie) z wybraną orientacją przestrzeni V.

Czytelnik może zastanawiać się co ten powrót do geometrii ma wspólnego z permutacjami? Tu warto spojrzeć na równoległościany i naszą, intuicyjną na razie, „objętość". Co się stanie, jeśli w $\mathbb{R}^n$ weźmiemy prostopadłościan $R(v_1, \ldots, v_n)$ i dla $\sigma \in S_n$ rozważymy równoległościan

$$R(v_{\sigma(1)} \ldots, v_{\sigma(n)})?$$

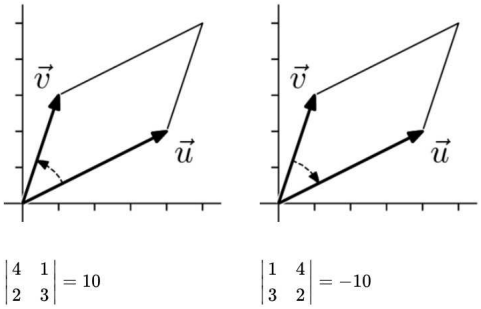
Okazuje się, że prostopadłościany te będą miały, z dokładnością do znaku, tą samą n-wymiarową objętość. Co więcej, jeśli zechcemy rozważać „objętość ze znakiem", czyli np. „skierowane pole" to tak wprowadzone obiekty dla wyżej wymienionych równoległościanów różnić się będą jedynie znakiem permutacji σ. Oto sugestywny, znany nam już przykład.

Przykład. Bazy

$$\mathcal{A} = \{(4,2),(1,3)\}, \quad \mathcal{B} = \{(1,3),(4,2)\}$$

są przeciwnie zorientowane, bo:

$$\det M(\mathrm{id})_{\mathcal{A}}^{\mathcal{B}} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} < 0.$$



Jeszcze raz przypominamy ten rysunek. Źródło: Jim Hefferon, `https://hefferon.net/linearalgebra/`.

W tym momencie wypada zakończyć. Czytelnika zainteresowanego głębszymi zastosowaniami permutacji w teorii wyznaczników zapraszam w uzupełnieniu do poznania słynnego twierdzenia Cauchy'ego-Bineta[2], będącego uogólnieniem formuły Cauchy'ego, przy czym interesuje nas policzenie wyznacznika iloczynu AB, gdzie AB jest macierzą kwadratową, podczas gdy A, B nie są kwadratowe!

* * *

Zakończyliśmy nasze rozważania i pierwszy semestr algebry liniowej. Zebraliśmy narzędzia potrzebne do badania następującego problemu: jak bardzo „zmieniają" się podprzestrzenie gdy działamy na całą przestrzeń przekształceniem liniowym? Innymi słowy: czy są takie podprzestrzenie, z których to przekształcenie nas nie wyprowadza (na przykład dla obrotu w $\mathbb{R}^2$ takich nietrywialnych podprzestrzeni nie ma), czy są takie, do których ograniczenie zamienia przekształcenie w homotetię? Czy zawsze da się rozbić przestrzeń na sumę prostą podprzestrzeni takich, że na każdej przekształcenie zachowuje się jak homotetia? I dlaczego kogoś to by miało interesować? A to dopiero początek.

Poznawszy teorię endomorfizmów zajmiemy się geometrią – najpierw poznamy podstawy geometrii afinicznej, a później zajmiemy się kluczowym pojęciem algebry liniowej, mającym niewyobrażalne wprost znaczenie – ortogonalnością. Konieczne będzie swobodne korzystanie ze wszystkich poznanych dotąd metod. Ciąg dalszy nastąpi.

[2]Patrz `https://mimuw.edu.pl/~amecel/2021l/gal21/galII+21wc.html`, wykład szósty, str. 29-24. Zadania dotyczące tego twierdzenia: `https://www.mimuw.edu.pl/~amecel/2017z/galj/[01.19]gal.pdf`.

12.4 Uzupełnienie. Macierze blokowe i wyznacznik

Szczególne znaczenie dla rozważań w drugim semestrze mieć będą macierze w postaciach blokowych. Teoria wyznaczników tych macierzy jest niezwykle bogata i prowadzi do ładnych rezultatów dotyczących między innymi rzędu macierzy.

Definicja 12.11: Macierz w postaci blokowej

Niech $A \in M_n(K)$ oraz $n = n_1 + n_2 + \ldots + n_k$, dla pewnych całkowitych $n_1, \ldots, n_k, k > 0$. Niech macierz $D_{ij} \in M_{n_i \times n_j}(K)$, zwana dalej BLOKIEM A względem podziału $n = n_1 + n_2 + \ldots + n_k$), powstaje z macierzy A przez:

- usunięcie wszystkich wierszy poza wierszami o indeksach $n_1 + \ldots + n_{i-1} + 1, \ldots, n_1 + \ldots + n_{i-1} + n_i$,
- usunięcie wszystkich kolumn poza kolumnami o indeksach $n_1 + \ldots + n_{j-1} + 1, \ldots, n_1 + \ldots + n_{j-1} + n_j$,

przy czym przyjmujemy $n_0 = 0$. Wówczas mówimy, że macierz A jest w POSTACI BLOKOWEJ (D_{ij}) (względem rozbicia $n = n_1 + n_2 + \ldots + n_k$), co oznaczamy często w następujący sposób:

$$A = \left[\begin{array}{c|c|c|c} D_{11} & D_{12} & \ldots & D_{1k} \\ \hline D_{21} & D_{22} & \ldots & D_{2k} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline D_{k1} & D_{k2} & \ldots & D_{kk} \end{array}\right] \quad \text{lub prościej} \quad A = \begin{bmatrix} D_{11} & D_{12} & \ldots & D_{1k} \\ D_{21} & D_{22} & \ldots & D_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ D_{k1} & D_{k2} & \ldots & D_{kk} \end{bmatrix}.$$

Wiemy już, że wykonywanie operacji elementarnych można wyrazić w języku mnożenia macierzy. W szczególności, korzystając wiemy, że dla macierzy $A \in M_{n \times m}(K)$ rzędu r istnieje macierz P, będąca iloczynem macierzy operacji elementarnych ma wierszach taka, że macierz PA jest schodkowa zredukowana. Wiemy przy tym, że wszystkie niezerowy wyrazy macierzy PA znajdują się w pewnych r wierszach. Co więcej, wykonując operacje elementarne na kolumnach, których iloczyn stanowi macierz Q, możemy ją doprowadzić do postaci blokowej

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = P \cdot A \cdot Q,$$

gdzie macierz po lewej ma bloki rozmiarów $r \times r$ — jest to macierz identycznościowa I_r, zaś pozostałe bloki są macierzami zerowymi odpowiednich rozmiarów. To oczywiście nie jest dziwne z punktu widzenia przekształceń liniowych. Jeśli przekształcenie liniowe $\phi : K^m \to K^n$ ma w bazach standardowych macierz A, to biorąc bazę $\mathcal{A}$ przestrzeni K^m, której ostatnie $n - r$ elementów stanowi baza jądra ϕ oraz bazę $\mathcal{B}$, której pierwsze r wektorów stanowią obrazy pierwszych r wektorów bazy $\mathcal{A}$, wówczas

$$M(\phi)^{\mathcal{B}}_{\mathcal{A}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} = M(\phi)^{\mathcal{B}}_{st} \cdot A \cdot M(\phi)^{st}_{\mathcal{A}}.$$

Głębsze zrozumienie istoty macierzy blokowych nabędziemy rozważając teorię podprzestrzeni własnych i podprzestrzeni niezmienniczych. W tym momencie skupmy się głównie na aspekcie rachunkowym.

Definicja 12.12

Niech A będzie taka, jak w poprzedniej definicji. Bloki D_{ii} nazywamy BLOKAMI DIAGONALNYMI. Co więcej, macierz A nazywamy:

- BLOKOWO GÓRNOTRÓJKĄTNĄ, jeśli istnieje rozbicie $n = n_1 + \ldots + n_k$ na dodatnie składniki takie, że postać blokowa (D_{ij}) macierzy A względem tego rozbicia spełnia $D_{ij} = 0$, dla $i > j$,
- BLOKOWO DOLNOTRÓJKĄTNĄ, jeśli istnieje rozbicie $n = n_1 + \ldots + n_k$ na dodatnie składniki takie, że postać blokowa (D_{ij}) macierzy A względem tego rozbicia spełnia $D_{ij} = 0$, dla $i < j$,
- BLOKOWO DIAGONALNĄ, jeśli jest jednocześnie blokowo górnotrójkątna i blokowo-dolnotrójkątna, dla pewnego podziału $n = n_1 + \ldots + n_k$.

Powyższa notacja może wydawać się nieco niespójna z notacją przyjętą w definicji wyznacznika, ale zwykle nie będziemy używać oznaczenia D_{ij} poza powyższą definicją. Macierze blokowe będą odgrywały olbrzymią rolę w naszych rozważaniach w przyszłym semestrze. W tym momencie pokażemy jedynie następujący ważny fakt, stanowiący uogólnienie Obserwacji 12.1 oraz 12.2, (dla rozbicia $n = 1+1+\ldots+1$ macierz blokowo górnotrójkątna to po prostu macierz górnotrójkątna, analogicznie dla dolnotrójkątnej).

Obserwacja 12.13: Wyznacznik macierzy blokowo-trójkątnej

Niech A będzie macierzą blokowo górnotrójkątną lub blokowo dolnotrójkątną o blokach diagonalnych $D_{11}, \ldots, D_{kk}$. Wówczas $\det A = \det D_{11} \cdot \det D_{22} \cdot \ldots \det D_{kk}$.

Dowód. Pokażmy najpierw tezę dla macierzy blokowo górnotrójkątnej rozmiaru $n \times n$ postaci:

$$X = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}.$$

Rozumowanie jest indukcją ze względu na rozmiar n macierzy X. Oczywiście teza zachodzi dla $n = 2$ i macierzy o czterech blokach rozmiarów 1×1. Niech $n > 2$. Niech pierwsza kolumna macierzy A ma wyrazy $a_{11}, a_{21}, \ldots, a_{k1}$. Liczymy wyznacznik przez rozwinięcie względem pierwszej kolumny, otrzymując

$$\det X = (-1)^{1+1} a_{11} \det X_{11} + \ldots + (-1)^{k+1} a_{k1} \det X_{k1}.$$

Rzeczywiście, kolejne $n - k$ składników rozwinięcia zawiera czynnik x_{j1}, który dla $j > k$ równy jest zero. Zauważmy też, że X_{i1} są, dla $1 \leqslant i \leqslant k$ macierzami blokowo-górnotrójkątnymi postaci

$$X_{i1} = \begin{bmatrix} A_{i1} & * \\ 0 & D \end{bmatrix}.$$

A zatem zgodnie z założeniem indukcyjnym

$$\det X_{i1} = \det A_{i1} \cdot \det D.$$

W ten sposób uzyskujemy krok indukcyjny, bowiem:

$$\det X = (-1)^{1+1} a_{11} \det A_{11} \cdot \det D + \ldots + (-1)^{k+1} a_{k1} \det A_{k1} \cdot \det D = \det A \cdot \det D.$$

Dla macierzy blokowo-górnotrójkątnej o więcej niż 2 blokach rozumowanie jest prostą indukcją ze względu na liczbę bloków. Zauważmy bowiem, że macierz o $k > 1$ blokach diagonalnych $D_{11}, \ldots, D_{kk}$ traktować można jako macierz o dwóch blokach diagonalnych: D_{11} oraz bloku, którego blokami diagonalnymi są $D_{22}, \ldots, D_{kk}$. Rozumowanie dla macierzy blokowo dolnotrójkątnych wynika natomiast natychmiast z tego, że wyznacznik nie zmienia się przy transponowaniu, a transpozycja macierzy blokowo górnotrójkątnej jest macierzą blokowo dolnotrójkątną. □

* * *

Rozważmy macierze blokowe M_1, M_2 postaci:

$$M_1 = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix},$$

gdzie odpowiednie bloki typu A, B, C, D mają te same rozmiary. Oczywiście mamy

$$M_1 + M_2 = \begin{bmatrix} A_1 + A_2 & B_1 + B_2 \\ C_1 + C_2 & D_1 + D_2 \end{bmatrix}.$$

Można natomiast zapytać czy macierz $M_1 \cdot M_2$ ma postać blokową, o ile istnieje. Jak się okazuje, o ile iloczyny odpowiednich bloków są dobrze zdefiniowane, mamy:

$$M_1 M_2 = \begin{bmatrix} A_1 A_2 + B_1 C_2 & A_1 B_2 + B_1 D_2 \\ C_1 A_2 + D_1 C_2 & C_1 B_2 + D_1 D_2 \end{bmatrix}.$$

Nie będziemy dowodzić tego faktu (jest to dość uciążliwy, ale standardowy rachunek). Przekonajmy się jednak o jego elegancji i przydatności. Po pierwsze, pozwala on natychmiastowo uzasadnić wzór wyżej. Mamy bowiem:

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \cdot \begin{bmatrix} I & B \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}.$$

Wyznaczniki kolejnych macierzy równe są zatem $\det D, 1, \det A$ i teza wynika z wzoru Cauchy'ego.

Twierdzenie 12.9: Tożsamość Sylvestera

Niech I_n będzie macierzą identycznościową rozmiaru $n \times n$ oraz niech $X \in M_{m\times n}(K)$ oraz niech $Y \in M_{n\times m}(K)$. Wówczas:

$$\det(I_m + XY) = \det(I_n + YX).$$

Zauważmy, że mamy tu równość wyznaczników macierzy różnych rozmiarów. Jak pokazać ten fakt? Mamy mianowicie:

$$\begin{bmatrix} I_n & -Y \\ X & I_m \end{bmatrix} \cdot \begin{bmatrix} I_n & Y \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} I_n \cdot I_n - Y \cdot 0 & I_n \cdot Y - Y \cdot I_m \\ X \cdot I_n + I_m \cdot 0 & X \cdot Y + I_m \cdot I_m \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ X & XY + I_m \end{bmatrix}.$$

Z drugiej strony mamy:

$$\begin{bmatrix} I_n & Y \\ 0 & I_m \end{bmatrix} \cdot \begin{bmatrix} I_n & -Y \\ X & I_m \end{bmatrix} = \begin{bmatrix} I_n \cdot I_m + Y \cdot X & I_n \cdot (-Y) + Y \cdot I_m \\ 0 \cdot I_n + I_m \cdot X & 0 \cdot (-Y) + I_m \cdot I_m \end{bmatrix} = \begin{bmatrix} I_n + YX & 0 \\ X & I_m \end{bmatrix}.$$

W formułach powyższych występują macierze blokowo-górnotrójkątne i blokowo-dolnotrójkątne, których wyznaczniki umiemy już liczyć. Mamy zatem, na mocy wzoru Cauchy'ego:

$$\begin{vmatrix} I_n & -Y \\ X & I_m \end{vmatrix} \cdot 1 = |XY + I_m|, \quad 1 \cdot \begin{vmatrix} I_n & -Y \\ X & I_m \end{vmatrix} = |I_n + YX|.$$

Dostajemy zatem tezę. Czytelnika zainteresowanego różnymi rozwinięciami tego faktu oraz jego zastosowaniami np. w analizie numerycznej odsyłam do artykułu S. Paszkowskiego *Tożsamości wyznacznikowe i przykłady ich zastosowań*[3]

Rozważmy nieco inny przykład zwany czasem nierównością Sylvestera.

Twierdzenie 12.10

Niech $A \in M_{p\times n}(K), B \in M_{n\times q}(K)$. Wówczas:

$$r(A) + r(B) \leqslant r(AB) + n.$$

Czytelnik pewnie widział już nierówność

$$r(A + B) \leqslant r(A) + r(B).$$

Powyższa nierówność daje jednak nieco większe możliwości i ma rozmaite zastosowania. Dowodzimy ją korzystając z tożsamości:

$$\begin{bmatrix} I_n & 0 \\ -A & I_p \end{bmatrix} \cdot \begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix} \cdot \begin{bmatrix} I_n & -B \\ 0 & I_p \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & -AB \end{bmatrix}.$$

Tym razem powa jest o rzędzie, więc należy zauważyć, że powyższa równość stwierdza rzecz następującą:

$$r \begin{bmatrix} I_n & B \\ A & 0 \end{bmatrix} = r \begin{bmatrix} I_n & 0 \\ 0 & -AB \end{bmatrix} = r(AB).$$

Równość powyższa wynika z faktu, że poniższe dwie macierze mają wyznaczniki 1, czyli są odwracalne i mają rzędy równe $n + p$ (mnożenie przez macierz odwracalną nie zmienia rzędu, bo jest „tym samym" co składanie przekształcenia liniowego z izomorfizmem z odpowiedniej strony, patrz też Wniosek 11.5):

$$\begin{bmatrix} I_n & 0 \\ -A & I_p \end{bmatrix}, \quad \begin{bmatrix} I_n & -B \\ 0 & I_p \end{bmatrix}.$$

Twierdzenia Sylvestera maja spore znacznie choćby w kombinatoryce czy tzw. algebraicznej teorii grafów.

[3] https://wydawnictwa.ptm.org.pl/index.php/matematyka-stosowana/article/download/1348/1287.

Do dalszych rozważań potrzebny nam będzie następujący Lemat.

Lemat. Niech I_n będzie macierzą identycznościową rozmiaru $n \times n$. Dla dowolnych dodatnich liczb całkowitych n, m mamy

$$\begin{vmatrix} 0 & I_n \\ I_m & 0 \end{vmatrix} = (-1)^{mn}.$$

Uzasadnienie. Ustalmy n i prowadźmy indukcję względem m. Dla $m = 1$ teza jest jasna. Stosujemy rozwinięcie względem pierwszej kolumny i mamy:

$$\begin{vmatrix} 0 & I_n \\ 1 & 0 \end{vmatrix} = (-1)^{n+1+1} \cdot |I_n| = (-1)^{n+2} \cdot |I_n| = (-1)^{n+2} \cdot 1 = (-1)^{1 \cdot n}.$$

Przechodzimy do kroku indukcyjnego. Niech $m > 1$. Ponownie liczymy wyznacznik względem pierwszej kolumnie, w której jedyny niezerowy element równy 1 znajduje się w $n + 1$-wszym wierszy i pierwszej kolumnie. Po usunięciu $n + 1$-wszego wiersza i pierwszej kolumny dostajemy ponownie macierz o blokach I_{m-1} oraz I_n. Mamy zatem:

$$\begin{vmatrix} 0 & I_n \\ I_m & 0 \end{vmatrix} = (-1)^{n+1+1} \cdot \begin{vmatrix} 0 & I_n \\ I_{m-1} & 0 \end{vmatrix}.$$

A zatem z założenia indukcyjnego widzimy, że ostatni wyznacznik to $(-1)^{(m-1)n}$. Dostajemy zatem wynik

$$(-1)^{n+1+1} \cdot (-1)^{(m-1)n} = (-1)^{n+1+1+mn-n} = (-1)^{2+mn} = (-1)^{mn}.$$

Przydatność powyższego lematu wynika z tego, że za pomocą macierzy blokowych rozważanych w tym lemacie można wykonać „mieszanie" bloków w macierzach mających taką postać. Zobaczmy kilka konkretnych zastosowań.

Przykład 1. Dla macierzy blokowych $A \in M_{n\times n}(K)$, $B = M_{n\times m}(K)$, $C = M_{m\times m}(K)$, $D = M_{m\times n}(K)$:

$$\begin{bmatrix} 0 & A \\ D & C \end{bmatrix} \cdot \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ D & C \end{bmatrix}.$$

A zatem korzystając z wzoru Cauchy'ego mamy zatem

$$\begin{vmatrix} 0 & A \\ D & C \end{vmatrix} \cdot \begin{vmatrix} 0 & I_n \\ I_m & 0 \end{vmatrix} = \begin{vmatrix} A & 0 \\ D & C \end{vmatrix} \Longrightarrow \begin{vmatrix} 0 & A \\ D & C \end{vmatrix} \cdot (-1)^{m+n} = |A| \cdot |C| \Longrightarrow \begin{bmatrix} 0 & A \\ D & C \end{bmatrix} = (-1)^{mn} \cdot |A| \cdot |C|.$$

Przykład 2. Dla macierzy blokowych $A \in M_{m\times p}(K)$, $B = M_{m\times q}(K)$, $C = M_{n\times p}(K)$, $D = M_{n\times q}(K)$, przy czym spełniony jest warunek $m + n = p + q$. Wówczas:

$$\begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} 0 & I_p \\ I_q & 0 \end{bmatrix} = \begin{bmatrix} D & C \\ B & A \end{bmatrix}.$$

Tym razem biorąc wyznaczniki obydwu stron dostajemy równość:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = (-1)^{mn+pq} \cdot \begin{vmatrix} D & C \\ B & A \end{vmatrix}.$$

W szczególności, jeśli A, B, C, D są macierzami kwadratowymi rozmiaru m, to liczba $mn + pq$ równa jest $2m^2$, czyli wyznaczniki macierzy po zamianie bloków są różne.

Przykład 3. Niech $A, B, C, D \in M_n(\mathbb{K})$. Wówczas jeśli A jest macierzą odwracalną, to:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B|.$$

Gdy istnieje A^{-1}, wówczas mamy:

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}.$$

Wyznacznik macierzy blokowo-dolnotrójkątnej o blokach diagonalnych I równy jest 1, natomiast wyznacznik macierzy blokowo-górnotrójkątnej o blokach A oraz $D - CA^{-1}B$ równy jest $|A| \cdot |D - CA^{-1}B|$.

Czytelnika może dziwić, że wychodzi tak skomplikowana równość, a nie na przykład wynik typu $|AD-BC|$ czy $|AD-CB|$. Okazuje się, że żaden z tych wzorów nie musi mieć miejsca. Wystarczy rozważyć choćby macierze A, B, C, D postaci:

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Wzór ten można uczytelnić, w przypadku, gdy $AC = CA$. Wymaga to jednak zastosowania narzędzi, które same w sobie mają duże znaczenie dla zrozumienia pojęcia wyznacznika.

Obserwacja 12.14

Dla dowolnej macierzy $A = (a_{ij}) \in M_n(K)$ funkcja $K \ni x \mapsto \det(A + xI)$ jest wielomianem stopnia n, którego współczynnik wiodący (tzn. współczynnik przy x^n) jest równy 1.

Teza tego faktu wynika natychmiast z rozwinięcia Laplace'a. Dla $n = 1$ nasza macierz ma postać $a + x$, czyli jest wielomianem stopnia 1 i ma przy x współczynnik 1. Niech $n > 1$. Rozwijamy wyznacznik macierzy $A + xI$ względem pierwszego wiersza. Wyrazami tego wiersza są $a_{11} + x, a_{12}, \ldots, a_{1n}$. Zgodnie z założeniem indukcyjnym wyznacznik $\det A_{11}$ jest wielomianem stopnia $n-1$ o współczynniku wiodącym równym 1. A zatem $(-1)^{1+1} \cdot (a + x) \cdot \det A_{11}$ jest wielomianem stopnia n o współczynniku wiodącym 1. Zauważmy, że składniki postaci $(-1)^{1+j} a_{1j} \det A_{1j}$ są wielomianami stopnia $n - 1$, dla $j \neq 1$. Istotnie, każda z macierzy A_{1j} po pewnej permutacji kolumn ma postać $B_j + xI_{n-1}$. A zatem zgodnie z założeniem indukcyjnym wyznacznik tej macierzy jest wielomianem stopnia $n - 1$. A zatem $\det(A + xI)$ jest sumą wielomianu stopnia n o współczynniku wiodącym 1 oraz wielomianów stopnia $n - 1$, co kończy dowód.

Wniosek 12.2

Jeśli $A, B, C, D \in M_n(\mathbb{R})$ oraz jeśli $AC = CA$, to wyznacznik macierzy blokowej rozmiaru $2n \times 2n$ o takich blokach ma postać:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|.$$

Czytelnik zechce sprawdzić, że wyniku $|AD-CB|$ nie można zamienić na $|AD-BC|$, z uwagi na macierz o blokach

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Dowód. Rozważamy osobno przypadki, gdy $\det A \neq 0$ oraz, gdy $\det A = 0$.

- Niech A – odwracalna. Wówczas korzystając z przemienności A oraz C mamy

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B| = \underbrace{|AD - ACA^{-1}B| = |AD - CB|}_{AC=CA}.$$

- Niech $\det A = 0$. Weźmy d takie, że $|A + \epsilon I| \neq 0$, dla $0 < \epsilon < d$. W istocie, skoro funkcja $x \mapsto \det(A + xI)$ jest wielomianem, to jest funkcją ciągłą (jesteśmy nad $\mathbb{R}$). A zatem dla dowolnego zera tego wielomianu istnieje takie jego otoczenie, gdzie jest on niezerowy. Wtedy mamy

  $$(A + \epsilon I)C = C(A + \epsilon I),$$

  czyli z przypadku rozważanego wyżej:

  $$\begin{vmatrix} A + \epsilon I & B \\ C & D \end{vmatrix} = |(A + \epsilon I)D - CB|.$$

  Funkcja $\det(A + xI)$ jest ciągła (jesteśmy nad $\mathbb{R}$), wiec biorąc $\epsilon \to 0$ dostajemy tezę.

□

Czy wniosek sformułowany wyżej rzeczywiście wymaga $K = \mathbb{R}$? Warto pomyśleć.

12.5 Dodatek. Interpolacja i wyznacznik Vandermonde'a

Przyjrzyjmy się następującemu zagadnieniu, interesującemu również z punktu widzenia algebry.

Zadanie interpolacyjne Lagrange'a polega na znalezieniu dla danej funkcji $f : \mathbb{R} \to \mathbb{R}$ wielomianu $P_n \in \mathbb{R}[x]$ stopnia nie wyższego niż n, którego wartości w $n+1$ z góry zadanych parami różnych punktach $x_0 \ldots, x_n$ są takie same, jak wartości interpolowanej funkcji, tzn.

$$P_n(x_i) = f(x_i), \qquad \text{dla } i = 0, 1, \ldots, n.$$

Zamiast $\mathbb{R}$ można rozważać dowolne ciało charakterystyki 0. Zachodzi następujące twierdzenie.

Twierdzenie 12.11

Zadanie interpolacyjne Lagrange'a ma dokładnie jedno rozwiązanie. Mianowicie konstruując funkcje pomocnicze:

$$p_i(x) = \prod_{\substack{j=0 \\ j\neq i}} \frac{x - x_j}{x_i - x_j}. \quad i = 0, 1, 2, \ldots, n,$$

określamy rozwiązanie zadania interpolacyjnego wzorem:

$$P_n(x) = f(x_0)p_0(x) + f(x_1)p_1(x) + \ldots + f(x_n)p_n(x) = \sum_{i=0}^{n} f(x_i) \prod_{\substack{j=0 \\ j\neq i}} \frac{x - x_j}{x_i - x_j}. \qquad (\heartsuit)$$

Przykład: $f : \mathbb{R} \to \mathbb{R}$ spełniają warunki $f(0) = 1$, $f(1) = 3$, $f(3) = 2$, $f(4) = 1$. Wielomiany $p_0(x)$, $p_1(x)$, $p_2(x)$, $p_3(x)$ wymienione w twierdzeniu wyżej są wówczas postaci:

$$\frac{(x-1)(x-3)(x-4)}{(0-1)(0-3)(0-4)}, \quad \frac{(x-0)(x-3)(x-4)}{(1-0)(1-3)(1-4)}, \quad \frac{(x-0)(x-1)(x-4)}{(3-0)(3-1)(3-4)}, \quad \frac{(x-0)(x-1)(x-3)}{(4-0)(4-1)(4-3)}.$$

Dowód. Nietrudno widzieć, że $p_i(x)$ to wielomiany stopnia n takie, że[4]:

$$p_i(x_j) = \begin{cases} 1, & \text{dla } i = j \\ 0, & \text{dla } i \neq j. \end{cases}$$

Stąd $P_n(x)$ jest wielomianem stopnia co najwyżej n przyjmującym w punktach x_i wartości $f(x_i)$, czyli jest rozwiązaniem problemu interpolacyjnego. Z drugiej strony z twierdzenia Bezout wiadomo, że wielomian taki jest jednoznaczny. Istotnie, gdyby jakiś wielomian P'_n stopnia nie większego od n również spełniał zadanie interpolacyjne, wówczas $P_n(x) - P'_n(x)$ jest wielomianem stopnia n o $n+1$ pierwiastkach $x_0, \ldots, x_n$, co implikuje, że $P_n(x) = P'_n(x)$. □

Powyższe twierdzenie zawiera pewien zaskakujący element: definicję wielomianów p_i. Dlaczego mają one taką właśnie postać? Podpowiedzi dostarczają nam wzory Cramera zastosowane do układu $n+1$ równań, w którym niewiadomymi są współczynniki wielomianu $P_n = a_0 + a_1x + a_2x^2 + \ldots + a_{n-1}x^{n-1} + a_nx^n$, spełniającego dla pewnych z góry zadanych $f(x_0), \ldots, f(x_n)$ warunki:

$$\begin{cases} P_n(x_0) = a_0 + a_1x_0 + a_2x_0^2 + \ldots + a_{n-1}x_0^{n-1} + a_nx_0^n & = f(x_0) \\ P_n(x_1) = a_0 + a_1x_1 + a_2x_1^2 + \ldots + a_{n-1}x_1^{n-1} + a_nx_1^n & = f(x_1) \\ & \vdots \\ P_n(x_n) = a_0 + a_1x_n + a_2x_n^2 + \ldots + a_{n-1}x_n^{n-1} + a_nx_n^n & = f(x_n). \end{cases}.$$

Wiemy już, że powyższy układ ma jednoznaczne rozwiązanie $(a_0, a_1, \ldots, a_n)$, a więc jego macierz współczynników jest odwracalna. Policzmy wyznacznik tej macierzy – zwanej macierzą Vandermonde'a.

$$\Delta(x_0, \ldots, x_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^{n-1} & x_1^n \\ 1 & x_2 & x_1^2 & \ldots & x_1^{n-1} & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} & x_n^n \end{vmatrix}.$$

[4] Mówiąc dokładniej, co zresztą za chwilę pokażemy, układ wielomianów p_i jest bazą przestrzeni wielomianów stopnia nie większego niż n, zaś współrzędne dowolnego innego wielomianu $f \in \mathbb{R}_n[X]$ w tej bazie to $f(x_1), f(x_2), \ldots, f(x_n)$. To powinno się kojarzyć z bazą dualną. Rzeczywiście, rozważając liniowo niezależne funkcjonały $f \xrightarrow{\mu_i} f(x_i)$ w $(\mathbb{R}_n[X])^*$ widzimy, że stanowią one bazę dualną do zaprezentowanego wyżej układu wielomianów $p_i \in \mathbb{R}_n[X]$.

Twierdzenie 12.12

Zachodzi równość:

$$\Delta(x_0,\dots,x_n)=\prod_{0\leqslant i<j\leqslant n}(x_j-x_i).$$

Czy Czytelnik widzi, że wyznacznik Vandermonde'a pojawia się w określeniu wielomianów $p_i(x)$? Formuła (♡) nie ujawnia współczynników wielomianu interpolacyjnego, ale teraz widzimy, że mogą być one wyznaczone z wzorów Cramera. Widzimy tu duże podobieństwo do iloczynu postaci $A^{-1}b$, rozważanego w dowodzie wzorów Cramera. Kluczowy wniosek jest tu taki: wyznaczenie $\Delta(x_0,\dots,x_n)$ zapewnia egzystencjalny dowód istnienia wielomianu interpolacyjnego, bez „zgadywania" wielomianów p_i.

Dowód. Indukcja ze względu na liczbę n. Dla $n=1$ mamy: $\det\begin{bmatrix}1 & x_0\\ 1 & x_1\end{bmatrix}=x_1-x_0$. Niech $n>1$. Idea jest taka, by rozbić macierz Vandermonde'a na iloczyn macierzy i skorzystać ze wzoru Cauchy'ego i założenia indukcyjnego. Dokładniej, wystarczy pokazać, że:

$$\Delta(x_0,\dots,x_n)=(x_1-x_0)(x_2-x_0)\cdots(x_n-x_0)\cdot\Delta(x_1,\dots,x_n). \qquad (\spadesuit)$$

Bierzemy macierz Vandermonde'a i odejmujemy pierwszy wiersz od pozostałych. Zrobimy to za pomocą mnożenia macierzy, żeby Czytelnik mógł się przekonać, że nie tylko macierze operacji elementarnych wykonują pewne operacje na wierszach macierzy, przez które pomnożyliśmy je z prawej strony:

$$\begin{bmatrix}1&0&0&\dots&0\\-1&1&0&\dots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\-1&0&0&\dots&1\end{bmatrix}\cdot\begin{bmatrix}1&x_0&x_0^2&\dots&x_0^{n-1}&x_0^n\\1&x_1&x_1^2&\dots&x_1^{n-1}&x_1^n\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\1&x_n&x_n^2&\dots&x_n^{n-1}&x_n^n\end{bmatrix}=\begin{bmatrix}1&x_0&x_0^2&\dots&x_0^{n-1}&x_0^n\\0&x_1-x_0&x_1^2-x_0^2&\dots&x_1^{n-1}-x_0^{n-1}&x_1^n-x_0^n\\\vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&x_n-x_0&x_n^2-x_0^2&\dots&x_n^{n-1}-x_0^{n-1}&x_n^n-x_0^n\end{bmatrix}.$$

Macierz po prawej jest dolnotrójkatna i ma wyznacznik równy 1, co zgadza się z formułą Cauchy'ego i obserwacją mówiącą, że ciąg operacji typu (1) nie zmienia wyznacznika. Idźmy dalej. Uzyskana macierz jest blokowo górnotrójkątna, a zatem mamy (inaczej mówiąc: rozwijając względem pierwszej kolumny):

$$\Delta(x_0,\dots,x_n)=\det\begin{bmatrix}x_1-x_0&x_1^2-x_0^2&\dots&x_1^{n-1}-x_0^{n-1}&x_1^n-x_0^n\\\vdots&\vdots&\ddots&\vdots&\vdots\\x_n-x_0&x_n^2-x_0^2&\dots&x_n^{n-1}-x_0^{n-1}&x_n^n-x_0^n\end{bmatrix}$$

Teraz przedstawimy powyższą macierz w postaci iloczynu trzech macierzy Po pierwsze wyciągamy wspólne czynniki x_i-x_0 z każdego wiersza i korzystając ze wzorów skróconego mnożenia mamy:

$$\begin{bmatrix}x_1-x_0&0&\dots&0\\0&x_2-x_0&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&x_n-x_0\end{bmatrix}\cdot\begin{bmatrix}1&x_1+x_0&\dots&\sum_{i=0}^{n-1}x_1^{n-1-i}x_0^i\\1&x_2+x_0&\dots&\sum_{i=0}^{n-1}x_2^{n-1-i}x_0^i\\\vdots&\vdots&\ddots&\vdots\\1&x_n+x_0&\dots&\sum_{i=0}^{n-1}x_n^{n-1-i}x_0^i\end{bmatrix}.$$

Macierz po prawej wygląda nieco nieprzyjemnie, ale to się zmieni po rozbiciu jej na następujący iloczyn:

$$\begin{bmatrix}1&x_1&x_1^2&\dots&x_1^{n-1}\\1&x_2&x_2^2&\dots&x_2^{n-1}\\\vdots&\vdots&\vdots&\ddots&\vdots\\1&x_n&x_2^2&\dots&x_n^{n-1}\end{bmatrix}\cdot\begin{bmatrix}1&x_0&x_0^2&\dots&x_0^{n-1}\\0&1&x_1&\dots&x_0^{n-2}\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\dots&1\end{bmatrix}.$$

A zatem pokazaliśmy, że wyznacznik Vandermonde'a $\Delta(x_0,\dots,x_n)$ równy jest w istocie iloczynowi wyznaczników trzech macierzy:

- macierzy diagonalnej o wyrazach x_i-x_0,
- macierzy Vandermonde'a o wyznaczniku $\Delta(x_1,\dots,x_n)$,
- macierzy górnotrójkątnej mającej jedynki na przekątnej.

A zatem z formuły Cauchy'ego mamy (♠), □

W trakcie studiów wielokrotnie napotkają Państwo (mam nadzieję) powyższą macierz i wyznacznik.

12.6 Notka historyczna. Na początku był wyznacznik...

Historia wyznacznika jest niezwykle skomplikowana, ale warto przytoczyć kilka ogólnych uwag[5]. Pierwszy był zdecydowanie wyznacznik, służący przede wszystkim do rozwiązywania układów równań. Historycy matematyki mówią, że już starożytni matematycy chińscy mieli pierwsze intuicje wyznacznika, a w słynnym dziele *Dziewięć rozdziałów o sztuce matematycznej*. Słynny japoński matematyk Seki Shinsuke Kowa, lub Takakazu (1642-1708), odkrył rozwinięcia wyznacznikowe (przez nazywane rozwinięciami Laplace'a), oczywiście w celu rozwiązywania układów równań. W Europie metodę zwaną wzorami Cramera wykorzystywał już de facto dla układów dwóch równań Cardano w swojej *Ars Magna* (1545). Nie dotarł on wpradzie do pojęcia wyznacznika. W Europie najwcześniejsze użycie funkcji nazywanych wyznacznikami znajdujemy w korespondencji pomiędzy Leibnizem i De L'Hospitalem z 1693 roku. Listy te nie zostały opublikowane aż do połowy wieku XIX-tego i nie miały znaczenia dla rozwoju teorii układów równań. Nie znaczy to oczywiście, że nie rozwiazywano układów równań, i to nawe takich, które w istocie wykorzystywały macierz Vandermonde'a. Robili to już kilkasiedziesiąt lat wcześniej Newton i de Moivre.

Pierwsze wyniki dotyczące wyznacznika spisał w 1730 roku Maclaurin, dowodząc między innymi metodę Cramera dla macierzy 2x2 oraz 3x3, a także wskazując jak należy ją uogólnić. Sam Cramer sformułował ogólnie tą zasadę w roku 1750. Chodziło oczywiście o rozwiązanie zagadnienia interpolacyjnego, a dokładnie o opis krzywej płaskiej przechodzącej przez określoną liczbę punktów. Należy odnotować, że Cramer nie podał ogólnego dowodu owej zasady. Od tego czasu regularne stały się prace o wyznaczniku (jesteśmy 200 lat przed definicjami macierzy autorstwa Sylvestera). W 1764 roku Bezout sformułował nowe metody liczenia wyznaczników. Podążył za nim Vandermonde (1771) oraz Laplace (1772), który studiując orbity planet próbował ocenić rozwiązywalność układów równań liniowych bez wyznaczania rozwiązań, właśnie za pomocą „rezultanty"', czyli dzisiejszego wyznacznika. Laplace sformułował również metodę rozwinięć.

Obok zastosowań algebraicznych, pojawiły się również wątki geometryczne. W 1773 roku Lagrange pokazał, że czworościan o wierzchołkach w punktach $(0,0,0),(x,y,z),(x',y',z'),(x'',y'',z'')$ ma objętość:

$$\frac{1}{6}(z(x'y''-y'x'')+z'(yx''-xy'')+z''(xy'-yx').$$

Termin „wyznacznik" pochodzi od Gauusa ze wspominanych już kilkakrotnie *Disquisitiones arithmeticae* (1801). Gauss zajmował się (co i my uczynimy w drugim semestrze) formami kwadratowymi, sformułował mnożenie macierzy (choć myślał o tym jako o składaniu przekształceń, a nie jako o algebrze „tablic liczb".

W sensie współczesnym pojęcia wyznacznika używał Cauchy. W roku 1812 opublikował pierwszą stosunkowo kompletną teorię wyznaczników, obejmującą również nowe wyniki dotyczące minorów i macierzy dołączonej, a także oczywiście twierdzenie Cauchy'ego o wyznaczniku iloczynu. Posłużyła mu ona do położenia podwalin pod teorię podobieństwa macierzy i wielomianu charakterystycznego (pewien ważny wyznacznik), którą poznamy w drugim semestrze. To również Cauchy udowodnił jedno z ważniejszych twierdzeń II semestru – twierdzenie spektralne o diagonalizowalności rzeczywistej macierzy symetrycznej.

Teoria wyznaczników stała się szeroko znana dzięki trzytomowej monografii Jacobiego z 1841 roku. Dzieło to dopuszczało szereg uogólnień, między innymi możliwość by wyznacznik wiązał ze sobą funkcje. W tym samym roku Cayley wprowadził notację wyznacznika, standardową przez kolejne stulecie. Gdy sformułowa została teoria macierzy (1850 – Sylvester, 1858 – Cayley), twiedzenia o wyznacznikach byly już standardem. Pierwszy w świecie podręcznik algebry liniowej napisał nikt inny tylko Charles Lutwidge Dodgson, znany jako autor książek dla dzieci i piszący jako Lewis Carroll. W 1867 roku, dwa lata po publikacji Alicji w Krainie Czarów, wydany zostaje podręcznik[6] *Elementary Treatise on Determinants, with their application to simultaneous linear equations and algebraical geometry*, gdzie udowodnione jest między innymi twierdzenie zwane przez nas (niesłusznie) twierdzeniem Kroneckera-Capellego. W kontekście literatury warto wspomnieć, że już w 1879 w Paryżu ukazała się ponad 600-stronicowa monografia o wyznacznikach w języku polskim, autorstwa Mariana Baranieckiego Jest ona dostępna online[7] Dzieło to poważane było nawet przez światowej klasy matematyków. Definicję aksjomatyczną wyznacznika, którą poznamy na kolejnym wykładzie, sformułował po raz pierwszy Weierstrass w roku 1903. Wraz z notatkami Kroneckera teoria wyznaczników ujęta w języku macierzy była już dojrzałą i ugruntowaną dziedziną algebry i analizy. Teoria ta uzyskała następnie w 1920 pięciotomową monografię Muira, wraz z pełnym przeglądem historycznym.

[5] Czynię to tłumacząc w zasadzie fragmenty odpowiedniego tekstu z MacTutor History of Mathematics Archive

[6] `https://archive.org/details/anelementarytre03carrgoog`

[7] `https://rcin.org.pl/impan/dlibra/publication/3422/edition/18584/content`.