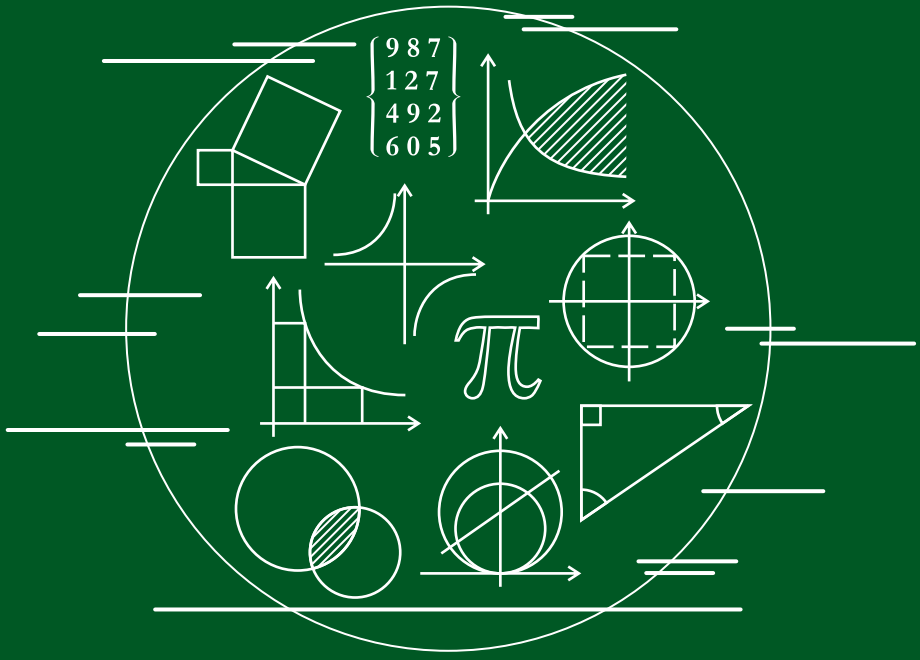
ALGEBRA LINIOWA Z GEOMETRIĄ ANALITYCZNĄ 1

Skrypt dla studentów
pierwszego semestru
matematyki stosowanej

Mateusz Woronowicz



Politechnika Białostocka

Mateusz Woronowicz

ALGEBRA LINIOWA Z GEOMETRIĄ ANALITYCZNĄ 1

Skrypt dla studentów pierwszego semestru matematyki stosowanej

Politechnika Białostocka

OFICYNA WYDAWNICZA POLITECHNIKI BIAŁOSTOCKIEJ
BIAŁYSTOK 2022

Recenzent:
dr Karol Pryszczepko

Redaktor naukowy dyscypliny matematyka:
prof. dr hab. inz. Zbigniew Bartosiewicz

Korekta językowa:
mgr Agnieszka Polecka

Skład, opracowanie graficzne:
dr Mateusz Woronowicz

Okładka:
Marcin Dominów

ISBN 978-83-67185-31-8 (eBook)
DOI: 10.24427/978-83-67185-31-8

Oficyna Wydawnicza Politechniki Białostockiej
ul. Wiejska 45C, 15-351 Białystok
e-mail: oficyna.wydawnicza@pb.edu.pl
www.pb.edu.pl

Spis treści

Wstęp

Niniejszy skrypt został napisany w celu ułatwienia nauki algebry linowej oraz geometrii analitycznej studentom pierwszego roku matematyki. Zakres tematyczny prezentowanych w nim treści jest zgodny z sylabusem przedmiotu Algebra liniowa z geometrią analityczną I wykładanego na Wydziale Informatyki Politechniki Białostockiej w pierwszym semestrze studiów na kierunku matematyka stosowana. Zdaniem autora, opracowanie to może służyć również jako dodatkowa pomoc dydaktyczna zainteresowanym matematyką studentom informatyki oraz informatyki i ekonometrii; szczególnie tym, którzy chcieliby zrozumieć teoretyczne podstawy algebry liniowej oraz geometrii analitycznej.

Kolejność omawianych w tym skrypcie zagadnień zgodna jest z kolejnością tematów realizowanych w ramach wykładu prowadzonego przez jego autora. Tożsame z wykładem są również treści składające się na tę książkę. Liczne, w pełni rozwiązane przykłady stanowią ilustrację oraz uzupełnienie prezentowanej w niej teorii. W zamyśle autora mają one ułatwić studentom przygotowanie się do ćwiczeń ściśle powiązanych ze wspomnianym wykładem.

Forma prezentacji materiału związana jest z dotychczasowymi dydaktycznymi doświadczeniami autora, z których wynika, że abstrakcyjna matematyka często sprawia studentom znaczne trudności. Dlatego ważną rolę w niniejszej publikacji odgrywają liczne uwagi odnoszące się do zasadniczych matematycznych treści, które bardziej zaawansowani Czytelnicy mogą pominąć, oraz wspomniane już wyżej przykłady. Istotnym celem przyświecającym powstaniu niniejszego skryptu było stworzenie możliwie samowystarczalnych materiałów dydaktycznych. Z tego względu zawarto w nim kompletne dowody niemal wszystkich omawianych twierdzeń oraz wiążących się z nimi wniosków, stwierdzeń i lematów; pominięto jedynie te, które – zdaniem autora – są nieadekwatne do etapu studiów głównych adresatów niniejszej publikacji.

Na kształt tej książki duży wpływ wywarły także doświadczenia studenckie jej autora. W tym kontekście szczególnie wartościowe okazały się wykłady z algebry liniowej dr. hab. Ryszarda R. Andruszkiewicza, Profesora Uniwersytetu w Białymstoku opublikowane w Andruszkiewicz (2005a, 2007). Częściowo na ich podstawie powstały czysto algebraiczne rozdziały niniejszego skryptu. Pierwotnymi, ogólnymi źródłami inspiracji dla części skryptu poświęconej geometrii analitycznej były natomiast materiały dydaktyczne Góra (n.d.) autorstwa dr. Michała Góry z Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie oraz podręcznik z geometrii analitycznej Kącki i i in. (1975) napisany przez prof. dr. hab. Edwarda Kąckiego, doc. dr. hab. Danutę Sadowską oraz prof. dr. hab. Lucjana Siewierskiego.

Używane w tej książce symbole $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{P}$, $\mathbb{N}$ i $\mathbb{N}_0$ oznaczają odpowiednio zbiory wszystkich liczb: wymiernych, całkowitych pierwszych, naturalnych (rozumianych jako dodatnie liczby całkowite) i całkowitych nieujemnych. Wszystkie inne oznaczenia są zgodne z powszechnie przyjętymi normami lub zostaną wyjaśnione później.

Autor pragnie wyrazić swą wdzięczność dr. Karolowi Pryszczepce za wnikliwą recenzję niniejszej książki oraz dr. Krzysztofowi Piekarskiemu, za życzliwą pomoc w rozwiązaniu kilku istotnych problemów związanych z jej komputerowym składem.

Białystok, czerwiec 2022 Mateusz Woronowicz

Rozdział 1

Podstawowe pojęcia algebraiczne

1.1 Dwuargumentowe działanie w niepustym zbiorze

Definicja 1.1. Dwuargumentowym działaniem w niepustym zbiorze A nazywamy dowolną funkcję przekształcającą zbiór $A \times A$ w zbiór A. Jeżeli $\star$ jest działaniem w A oraz $a_1, a_2 \in A$, to $\star\big((a_1, a_2)\big)$ nazywamy wynikiem działania $\star$ na parze (a_1, a_2).

Uwaga 1.1. Ponieważ przedmiotem naszych rozważań będą głównie działania dwuargumentowe, to nie będziemy za każdym razem używać określenia „dwuargumentowe"; będziemy mówić krótko: działanie. Zamiast notacji prefiksowej $\star\big((a_1, a_2)\big)$ (symbol działania znajduje się przed argumentami działania) stosować będziemy znaną z wcześniejszych etapów edukacji zwyczajową notację infiksową $a_1 \star a_2$ (symbol działania znajduje się między argumentami działania).

Przykład 1.1. Dodawanie jest działaniem w zbiorze liczb naturalnych $\mathbb{N}$ (bo mając dwie dowolne liczby naturalne możemy je do siebie dodać i otrzymana suma będzie liczbą naturalną). Odejmowanie nie jest działaniem w $\mathbb{N}$, gdyż dla liczb 1 i 2 mamy wprawdzie $1, 2 \in \mathbb{N}$ oraz odejmowanie liczb jest zawsze wykonalne, ale $1 - 2 = -1 \notin \mathbb{N}$. Oczywiście odejmowanie jest działaniem w zbiorze liczb całkowitych $\mathbb{Z}$. Zauważmy jeszcze, że dzielenie nie jest działaniem w zbiorze $\mathbb{R}$, bo $0 \in \mathbb{R}$ i dzielenie przez 0 jest niewykonalne.

Przykład 1.2. Niech $A = \{1, 2, 3, \ldots, 2020\}$. Wówczas wzór $a_1 \star a_2 = 1987$ określa działanie w zbiorze A. Innymi słowy, funkcja $\star\colon A \times A \to A$ dana wzorem $a_1 \star a_2 = 1987$ dla wszystkich $a_1, a_2 \in A$, jest (dwuargumentowym) działaniem w A.

Uwaga 1.2. Jeżeli n jest liczbą naturalną, zaś $A = \{a_1, a_2, \ldots, a_n\}$ jest zbiorem n-elementowym, to dowolne dwuargumentowe działanie $\star$ w zbiorze A można opisać za pomocą tabelki:

$\star$	a_1	a_2	a_3	$\cdots$	a_i	$\cdots$	a_j	$\cdots$	a_n
a_1	$a_1 \star a_1$	$a_1 \star a_2$	$a_1 \star a_3$	$\cdots$	$a_1 \star a_i$	$\cdots$	$a_1 \star a_j$	$\cdots$	$a_1 \star a_n$
a_2	$a_2 \star a_1$	$a_2 \star a_2$	$a_2 \star a_3$	$\cdots$	$a_2 \star a_i$	$\cdots$	$a_2 \star a_j$	$\cdots$	$a_2 \star a_n$
a_3	$a_3 \star a_1$	$a_3 \star a_2$	$a_3 \star a_3$	$\cdots$	$a_3 \star a_i$	$\cdots$	$a_3 \star a_j$	$\cdots$	$a_3 \star a_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
a_i	$a_i \star a_1$	$a_i \star a_2$	$a_i \star a_3$	$\cdots$	$a_i \star a_i$	$\cdots$	$a_i \star a_j$	$\cdots$	$a_i \star a_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
a_j	$a_j \star a_1$	$a_j \star a_2$	$a_j \star a_3$	$\cdots$	$a_j \star a_i$	$\cdots$	$a_j \star a_j$	$\cdots$	$a_j \star a_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
a_n	$a_n \star a_1$	$a_n \star a_2$	$a_n \star a_3$	$\cdots$	$a_n \star a_i$	$\cdots$	$a_n \star a_j$	$\cdots$	$a_n \star a_n$

w której dla dowolnych $i, j \in \{1, 2, \ldots, n\}$, wynik $a_i \star a_j$ działania $\star$ na parze (a_i, a_j) wpisany jest na przecięciu i-tego wiersza z j-tą kolumną.

Definicja 1.2. Mówimy, że dwuargumentowe działanie $\star$ w niepustym zbiorze A:

(i) jest przemienne, gdy $a_1 \star a_2 = a_2 \star a_1$ dla wszystkich $a_1, a_2 \in A$;
(ii) jest łączne, gdy $(a_1 \star a_2) \star a_3 = a_1 \star (a_2 \star a_3)$ dla wszystkich $a_1, a_2, a_3 \in A$;
(iii) posiada element neutralny, gdy istnieje $e \in A$ takie, że $a \star e = e \star a = a$ dla każdego $a \in A$.

Stwierdzenie 1.1. Każde dwuargumentowe działanie w niepustym zbiorze posiada co najwyżej jeden element neutralny.

Dowód. Niech $A \neq \emptyset$ i niech $\star \colon A \times A \to A$. Załóżmy, że e_1 oraz e_2 są elementami neutralnymi działania $\star$. Wtedy $e_1 \star e_2 = e_2$, bo e_1 jest elementem neutralnym działania $\star$, oraz $e_1 \star e_2 = e_1$, bo e_2 jest elementem neutralnym działania $\star$. Zatem $e_2 = e_1$.

Definicja 1.3. Niech e będzie elementem neutralnym działania $\star$ w zbiorze A i niech $a \in A$. Mówimy, że element a jest odwracalny w A względem działania $\star$, gdy istnieje $x \in A$ takie, że $a \star x = x \star a = e$. Jeżeli takie x istnieje, to x nazywamy elementem odwrotnym do a względem działania $\star$.

Stwierdzenie 1.2. Jeżeli a jest elementem odwracalnym względem łącznego działania $\star$ w zbiorze A, to istnieje dokładnie jeden element odwrotny do a względem działania $\star$.

Dowód. Istnienie co najmniej jednego elementu odwrotnego do a względem działania $\star$ wynika wprost z odwracalności a w A względem tego działania. Niech e będzie elementem neutralnym działania $\star$. Załóżmy, że $a \star x = x \star a = e$ oraz $a \star y = y \star a = e$ dla pewnych $x, y \in A$. Wtedy $y = e \star y = (x \star a) \star y = x \star (a \star y) = x \star e = x$.

Uwaga 1.3. Jeżeli symbol dwuargumentowego działania w niepustym zbiorze A posiadającego element neutralny przypomina znak $+$, to w stosunku do elementu x opisanego w Definicji 1.3 zamiast określenia „odwrotny” używamy określenia „przeciwny”.

1.2 Pojęcie grupy

Definicja 1.4. Niech m i n będą liczbami naturalnymi. Systemem algebraicznym nazywamy układ postaci:

$$(A, \star_1, \star_2, \ldots, \star_m, e_1, e_2, \ldots, e_n),$$

gdzie A jest niepustym zbiorem, $\star_1, \star_2, \ldots, \star_m$ są dwuargumentowymi działaniami w A oraz $e_1, e_2, \ldots, e_n$ są wyróżnionymi elementami zbioru A.

Definicja 1.5. Grupą nazywamy system algebraiczny $(G, \star, e)$ spełniający układ warunków:

(G1) działanie $\star$ jest łączne;
(G2) e jest elementem neutralnym działania $\star$;
(G3) każdy element zbioru G jest odwracalny w G względem działania $\star$.

Uwaga 1.4. Ze Stwierdzenia 1.2 wynika, że jeżeli system algebraiczny $(G, \star, e)$ jest grupą, to dla dowolnego $g \in G$ istnieje dokładnie jeden element $h \in G$ taki, że $g \star h = h \star g = e$. Element ten oznaczamy symbolem g^{-1} (lub $-g$, gdy symbol działania „przypomina" dodawanie).

Definicja 1.6. Grupę $(G, \star, e)$ nazywamy przemienną (lub abelową), wówczas gdy działanie $\star$ jest przemienne.

Uwaga 1.5. Określenie „abelowa" pochodzi od nazwiska norweskiego matematyka Nielsa Henrika Abela żyjącego w latach 1802-1829, który jako pierwszy prowadził systematyczne badania wykorzystujące wiedzę z zakresu grup przemiennych.

Uwaga 1.6. Często używa się również mniej formalnych, lecz bardzo naturalnych i nieprowadzących do nieporozumień sformułowań: „Zbiór G jest grupą względem działania $\star$" lub „Zbiór G wraz z działaniem $\star$ tworzy grupę". Oznaczają one, że spełniona jest koniunkcja następujących warunków:

(i) $\star$ jest dwuargumentowym łącznym działaniem w niepustym zbiorze G;
(ii) zbiór G zawiera pewien element e, który jest elementem neutralnym działania $\star\colon G \times G \to G$;
(iii) dla każdego $g \in G$, istnieje $h \in G$ takie, że $g \star h = h \star g = e$.

Przykład 1.3. Zbiór $\mathbb{Z}$ tworzy grupę abelową względem dodawania (formalnie: system algebraiczny $(\mathbb{Z}, +, 0)$ jest grupą abelową).

Przykład 1.4. System algebraiczny $\big(\{-1, 1\}, \cdot, 1\big)$ jest grupą abelową.

Przykład 1.5. Dowolny zbiór jednoelementowy $\{a\}$ rozważany wraz z (jedynym możliwym) działaniem $\star$ zdefiniowanym przez równość $a \star a = a$ tworzy grupę abelową. Taką grupę nazywamy grupą trywialną.

Przykład 1.6. Niech X będzie dowolnym niepustym zbiorem i niech $S(X)$ oznacza zbiór wszystkich bijekcji zbioru X. Niech ponadto $\circ$ oznacza składanie przekształceń w $S(X)$. Wtedy odwzorowanie identycznościowe id_X zbioru X należy $S(X)$ oraz $g \circ f \in S(X)$ dla dowolnych $f, g \in S(X)$. Zatem $\big(S(X), \circ, \mathrm{id}_X\big)$ jest systemem algebraicznym. Ponadto działanie $\circ$ jest łączne, id_X jest jego elementem neutralnym oraz dla dowolnego $f \in S(X)$ istnieje w $S(X)$ element odwrotny do f względem $\circ$ (to wszystko wynika bezpośrednio z przedmiotu Logika i Teoria Mnogości). Zatem $\big(S(X), \circ, \mathrm{id}_X\big)$ jest grupą. Jeżeli $|X| = 1$, to $S(X) = \{\mathrm{id}_X\}$, więc $\big(S(X), \circ, \mathrm{id}_X\big)$ jest grupą abelową (por. Przykład 1.5). Jeśli $|X| = 2$, to $X = \{x_1, x_2\}$ dla pewnych różnych elementów x_1 i x_2 oraz $S(X) = \{\mathrm{id}_X, f\}$, przy czym $x_1 \overset{f}{\mapsto} x_2$ oraz $x_2 \overset{f}{\mapsto} x_1$. Zatem również w tym przypadku grupa $\big(S(X), \circ, \mathrm{id}_X\big)$ jest abelowa. Niech teraz $|X| > 2$. Istnieją wówczas parami różne elementy x_1, x_2, x_3 zbioru X. Rozważmy funkcje $f, g \in X^X$ określone następująco:

$$f(x) = \begin{cases} x_2, & \text{gdy } x = x_1 \\ x_1, & \text{gdy } x = x_2 \\ x, & \text{gdy } x \in X \setminus \{x_1, x_2\} \end{cases}$$

oraz

$$g(x) = \begin{cases} x_2, & \text{gdy } x = x_1 \\ x_3, & \text{gdy } x = x_2 \\ x_1, & \text{gdy } x = x_3 \\ x, & \text{gdy } x \in X \setminus \{x_1, x_2, x_3\} \end{cases}.$$

Bezpośrednio z określenia funkcji f i g wynika, że $f, g \in S(X)$. Ponadto $(g \circ f)(x_3) = g\big(f(x_3)\big) = g(x_3) = x_1$ oraz $(f \circ g)(x_3) = f\big(g(x_3)\big) = f(x_1) = x_2$, więc $g \circ f \neq f \circ g$. Zatem w tym przypadku działanie $\circ$ nie jest przemienne i, w konsekwencji, grupa $\big(S(X), \circ, \mathrm{id}_X\big)$ nie jest abelowa.

1.2.1 Grupa permutacji

Definicja 1.7. Grupę $\big(S(X), \circ, \mathrm{id}_X\big)$ opisaną w Przykładzie 1.6, nazywamy grupą symetryczną zbioru X. Jeżeli $|X| = n$ dla pewnego $n \in \mathbb{N}$ oraz $X = \{x_1, x_2, \ldots, x_n\}$, to będziemy pisali krótko: S_n, zamiast: $S(\{x_1, x_2, \ldots, x_n\})$. Grupę $\big(S_n, \circ, \mathrm{id}\big)$, gdzie id oznacza funkcję identycznościową określoną na zbiorze $\{x_1, x_2, \ldots, x_n\}$, będziemy wówczas nazywali grupą permutacji zbioru n-elementowego.

Uwaga 1.7. W celu uproszczenia notacji grupę permutacji zbioru n-elementowego będziemy zazwyczaj oznaczali tym samym symbolem, co zbiór wszystkich permutacji zbioru n-elementowego, tzn. symbolem S_n. W algebrze bardzo często dokonuje się takich skrótów – praktycznie zawsze, gdy z kontekstu jasno wynika z jakim działaniem w rozważanym zbiorze mamy do czynienia. Inną sytuacją, w której pra-

wie zawsze unika się zapisywania grupy w postaci systemu algebraicznego są abstrakcyjne rozważania o dowolnej grupie, w których nie ma potrzeby wprowadzania wymyślnych oznaczeń działania. Piszemy wtedy krótko: „Niech G będzie grupą". Działanie związane z tą grupą domyślnie oznacza się standardową kropką $\cdot$, którą się często pomija (zamiast pisać: $g_1 \cdot g_2$, pisze się po prostu : $g_1 g_2$). Element neutralny tego działania oznacza się zazwyczaj symbolem e lub 1 (chociaż oczywiście nie musi być to liczba jeden!). Taki sposób przedstawienia grupy nazywamy zapisem multiplikatywnym (od łacińskiego słowa *multiplicare* oznaczającego mnożenie). W sytuacji gdy pojawia się informacja: „G jest grupą abelową", domyślnym oznaczeniem działania związanego z tą grupą jest $+$ (nie pomija się go w zapisach). Element neutralny takiej grupy oznacza się symbolem 0. Taki sposób przedstawienia grupy nazywamy zapisem addytywnym (od łacińskiego słowa *addere* oznaczającego dodawanie). Oczywiście istnieją wyjątki od tych nieformalnych reguł – abstrakcyjne grupy abelowe rozważa się niekiedy w zapisie multiplikatywnym.

Uwaga 1.8. Niech $n \in \mathbb{N}$. W rozważaniach dotyczących permutacji zbioru n-elementowego nie ma żadnego znaczenia ani jaki n-elementowy zbiór permutujemy, ani w jaki sposób będziemy oznaczali jego elementy. Wygodnie jednak przyjąć umowę, że rozważanym zbiorem zawsze jest $X_n = \{1, 2, 3, \ldots, n\}$. Wtedy dowolną permutację $\sigma \in S_n$ możemy zapisać w formie tablicy:

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & i & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(i) & \ldots & \sigma(n) \end{pmatrix}.$$

Taki zapis umożliwia wygodne składanie i odwracanie permutacji (zob. Przykład 1.8).

Definicja 1.8. Dla dowolnej liczby naturalnej $n \geq 2$ oraz dowolnej permutacji $\sigma \in S_n$ określamy zbiór:

$$X_\sigma = \big\{ i \in X_n \colon \sigma(i) \neq i \big\}.$$

Stwierdzenie 1.3. Dla dowolnej liczby naturalnej $n \geq 2$ oraz dowolnej permutacji $\sigma \in S_n$ zachodzi równość $\sigma(X_\sigma) = X_\sigma$.

Dowód. Weźmy dowolne $i \in X_\sigma$. Wtedy $\sigma(i) \neq i$, więc z różnowartościowości funkcji σ wynika, że $\sigma\big(\sigma(i)\big) \neq \sigma(i)$. Zatem $\sigma(i) \in X_\sigma$, skąd $\sigma(X_\sigma) \subseteq X_\sigma$. Ponadto $|X_\sigma| < \infty$, więc powołując się ponownie na iniektywność σ otrzymujemy żądaną równość.

Wniosek 1.1. W zapisie permutacji w postaci tablicy (1.3) można pomijać punkty stałe, tzn. takie $i \in X_n$, że $\sigma(i) = i$.

Przykład 1.7. Jeżeli $\sigma = \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7 \\ 1\,4\,2\,3\,5\,6\,7 \end{pmatrix} \in S_7$, to permutacje σ równie dobrze możemy zapisać w skróconej postaci $\sigma = \begin{pmatrix} 2\,3\,4 \\ 4\,2\,3 \end{pmatrix}$.

Uwaga 1.9. Gdy mamy do czynienia ze skróconą notacją $\sigma = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix}$, to nie widać, w której konkretnej grupie rozważamy permutację σ. Wiadomo jedynie, że $n \geq 4$ (bo $X_\sigma = \{2,3,4\}$ i mamy umowę, że $X_n = \{1,2,\ldots,n\}$) Nie stanowi to jednak problemu, bo permutację σ możemy w naturalny sposób „włożyć" do każdej grupy S_n, gdzie $n \geq 4$ – wszystkie punkty zbioru X_n oprócz $2,3$ i 4 będą punktami stałymi permutacji σ. Przy analizowaniu własności permutacji σ istotny jest jedynie zbiór X_σ. Z tego powodu nazywa się go dziedziną istotną permutacji σ. Oczywiście, gdy chcemy wyraźnie podkreślić, że permutację $\sigma = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix}$ rozważamy w grupie S_7, to piszemy $\sigma = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix} \in S_7$.

Przykład 1.8. W grupie S_7 rozważmy permutacje $\sigma = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix}$ i $\delta = \begin{pmatrix} 1 & 3 & 5 & 6 \\ 6 & 5 & 3 & 1 \end{pmatrix}$. Wówczas:

(i) $\delta \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 2 & 5 & 3 & 1 \end{pmatrix}$. Technicznie rzecz biorąc, wynik ten uzyskujemy w następujący sposób: (1) Zauważamy, że $X_\sigma = \{2,3,4\}$ oraz $X_\delta = \{1,3,5,6\}$. (2) Dla każdego $i \in X_\sigma \cup X_\delta$ obliczamy $(\delta \circ \sigma)(i) = \delta\big(\sigma(i)\big)$. (3) W pierwszym wierszu tablicy opisującej permutację $\delta \circ \sigma$ wypisujemy w porządku rosnącym wszystkie $i \in X_\sigma \cup X_\delta$, dla których $(\delta \circ \sigma)(i) \neq i$ (tj. wypisujemy w porządku rosnącym wszystkie elementy zbioru $X_{\delta\circ\sigma}$). (4) W drugim wierszu tablicy opisującej permutację $\delta \circ \sigma$ wypisujemy kolejno wartości $(\delta \circ \sigma)(i)$ odpowiadające liczbom i z pierwszego wiersza.
Oczywiście możemy także wykonać powyższe działanie, stosując pełny zapis permutacji. Wtedy, dla każdego $i \in X_7$ obliczamy kolejno $(\delta \circ \sigma)(i)$ oraz zapisujemy $\delta \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 4 & 2 & 5 & 3 & 1 & 7 \end{pmatrix}$.

(ii) $\sigma^{-1} = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 2 \end{pmatrix}$. Technicznie rzecz ujmując, wynik ten można uzyskać w następujący sposób: (1) zamieniamy miejscami wiersze tablicy opisującej permutację σ, (2) kolumny nowo uzyskanej tablicy porządkujemy w taki sposób, aby liczby w pierwszym wierszu były zapisane w porządku rosnącym.

Definicja 1.9. Niech $b \in \mathbb{N}$. Inwersją permutacji $\sigma \in S_n$ nazywamy dwuelementowy podzbiór $\{i,j\}$ zbioru X_n taki, że $i < j$ oraz $\sigma(i) > \sigma(j)$. Zbiór wszystkich inwersji permutacji σ oznaczamy symbolem I_σ.

Przykład 1.9. Dla permutacji $\sigma = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 2 & 3 \end{pmatrix} \in S_7$ mamy $I_\sigma = \big\{\{2,3\},\{2,4\}\big\}$.

Definicja 1.10. Niech $n \geq 2$ będzie liczbą naturalną i niech $i < j$ będą elementami zbioru X_n. Symbolem (i,j) oznaczamy permutację, która zamienia miejscami ele-

menty i, j oraz nie zamienia pozostałych elementów zbioru X_n. Taką permutację nazywamy transpozycją.

Uwaga 1.10. Zachowując wszystkie oznaczenia wprowadzone powyżej otrzymujemy, że $X_{(i,j)} = \{i, j\}$ oraz:

$$(i,j) = \begin{pmatrix} 1\ 2\ \ldots\ i-1\ \ i\ \ i+1\ \ldots\ \ j-1\ \ j\ \ j+1\ \ldots\ n \\ 1\ 2\ \ldots\ i-1\ \ j\ \ i+1\ \ldots\ \ j-1\ \ i\ \ j+1\ \ldots\ n \end{pmatrix}. \tag{1.2.1}$$

Zatem $I_{(i,j)}$ jest sumą rozłącznych zbiorów:

$$A = \big\{\{i,i+1\},\{i,i+2\},\{i,i+3\},\ldots,\{i,j-1\},\{i,j\}\big\}$$

oraz

$$B = \big\{\{i+1,j\},\{i+2,j\},\{i+3,j\},\ldots,\{j-1,j\}\big\}.$$

Ponadto $|A| = j-i$ oraz $|B| = (j-1)-i$, więc:

$$|I_{(i,j)}| = (j-i)+(j-1)-i = 2(j-i)-1.$$

Bezpośrednią konsekwencją powyższych obserwacji jest następujący

Wniosek 1.2. Każda transpozycja posiada nieparzystą liczbę wszystkich inwersji.

Definicja 1.11. Niech $n \in \mathbb{N}$. Znakiem permutacji $\sigma \in S_n$ nazywamy liczbę $\operatorname{sgn}(\sigma)$ określoną za pomocą wzoru:

$$\operatorname{sgn}(\sigma) = (-1)^{|I_\sigma|}.$$

Definicja 1.12. Niech $n \in \mathbb{N}$. Mówimy, że permutacja $\sigma \in S_n$ jest parzysta, gdy $\operatorname{sgn}(\sigma) = 1$. W przeciwnym razie permutację σ nazywamy nieparzystą.

Bezpośrednią konsekwencją Wniosku 1.2 jest kolejny

Wniosek 1.3. Każda transpozycja jest permutacją nieparzystą.

Twierdzenie 1.1. Niech $n \in \mathbb{N}$. Dla dowolnych $\sigma, \delta \in S_n$ zachodzi wzór:

$$\operatorname{sgn}(\delta \circ \sigma) = \operatorname{sgn}(\delta) \cdot \operatorname{sgn}(\sigma). \tag{1.2.2}$$

Dowód. Niech $\mathfrak{A}$ oznacza rodzinę wszystkich dwuelementowych podzbiorów zbioru X_n. Rozważmy dowolne $\alpha \in S_n$ oraz dowolne $A \in \mathfrak{A}$. Wówczas $A = \{i, j\}$ dla pewnych różnych elementów i, j zbioru X_n, oraz $\frac{\alpha(j)-\alpha(i)}{j-i} = \frac{-\big(\alpha(i)-\alpha(j)\big)}{-(i-j)} = \frac{\alpha(i)-\alpha(j)}{i-j}$. Możemy więc zdefiniować liczbę:

$$\operatorname{sgn}_\alpha(A) = \operatorname{sgn}\left(\frac{\alpha(i)-\alpha(j)}{i-j}\right).$$

(Można ją nazwać znakiem zbioru $A = \{i, j\}$ względem permutacji α zdefiniowanym jako znak ułamka $\frac{\alpha(i)-\alpha(j)}{i-j}$). Zauważmy, że A jest inwersją permutacji α wtedy i tylko wtedy, gdy $\operatorname{sgn}_\alpha(A) = -1$, co oznacza, że:

$$I_\alpha = \{A \in \mathfrak{A} \colon \operatorname{sgn}_\alpha(A) = -1\}.$$

Stąd oraz na mocy Definicji 1.11 otrzymujemy, że:

$$\operatorname{sgn}(\alpha) = \prod_{A \in \mathfrak{A}} \operatorname{sgn}_\alpha(A). \tag{1.2.3}$$

Rozważmy dowolne $\sigma, \delta \in S_n$ oraz funkcję $\mathfrak{A} \ni A = \{i, j\} \overset{F_\sigma}{\longmapsto} \{\sigma(i), \sigma(j)\} \in \mathfrak{A}$. Wprost z definicji funkcji F_σ wynika, że jest ona iniektywna (bo σ jest bijekcją!). Ponadto $|\mathfrak{A}| < \infty$, bo $|X_n| < \infty$. Zatem F_σ jest bijekcją. Stąd oraz na mocy wzoru (1.2.3) uzyskujemy wzór:

$$\operatorname{sgn}(\delta) = \prod_{A \in \mathfrak{A}} \operatorname{sgn}_\delta\left(F_\sigma(A)\right). \tag{1.2.4}$$

Dalej,

$$\operatorname{sgn}_{\delta\circ\sigma}(A) = \operatorname{sgn}\left(\frac{(\delta\circ\sigma)(i) - (\delta\circ\sigma)(j)}{i-j}\right) = \operatorname{sgn}\left(\frac{\delta(\sigma(i)) - \delta(\sigma(j))}{i-j}\right) =$$

$$\operatorname{sgn}\left(\frac{\delta(\sigma(i)) - \delta(\sigma(j))}{\sigma(i)-\sigma(j)} \cdot \frac{\sigma(i)-\sigma(j)}{i-j}\right) = \operatorname{sgn}\left(\frac{\delta(\sigma(i)) - \delta(\sigma(j))}{\sigma(i)-\sigma(j)}\right) \cdot$$

$$\operatorname{sgn}\left(\frac{\sigma(i)-\sigma(j)}{i-j}\right) = \operatorname{sgn}_\delta\left(F_\sigma(A)\right) \cdot \operatorname{sgn}_\sigma(A).$$

W ten sposób otrzymujemy wzór:

$$\operatorname{sgn}_{\delta\circ\sigma}(A) = \operatorname{sgn}_\delta\left(F_\sigma(A)\right) \cdot \operatorname{sgn}_\sigma(A). \tag{1.2.5}$$

Na mocy wzorów (1.2.3), (1.2.5) i (1.2.4) uzyskujemy ostatecznie, że $\operatorname{sgn}(\delta\circ\sigma) = \prod_{A\in\mathfrak{A}} \operatorname{sgn}_{\delta\circ\sigma}(A) = \prod_{A\in\mathfrak{A}} \left(\operatorname{sgn}_\delta\left(F_\sigma(A)\right) \cdot \operatorname{sgn}_\sigma(A)\right) = \prod_{A\in\mathfrak{A}} \operatorname{sgn}_\delta\left(F_\sigma(A)\right) \cdot \prod_{A\in\mathfrak{A}} \operatorname{sgn}_\sigma(A) = \operatorname{sgn}(\delta) \cdot \operatorname{sgn}(\sigma)$.

Wniosek 1.4. Niech $n \in \mathbb{N}$. Wówczas $\operatorname{sgn}(\sigma) = \operatorname{sgn}\left(\sigma^{-1}\right)$ dla każdego $\sigma \in S_n$.

Dowód. Ze wzoru (1.2.2) wynika, że $\operatorname{sgn}\left(\sigma\circ\sigma^{-1}\right) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}\left(\sigma^{-1}\right)$. Ponadto $\sigma\circ\sigma^{-1} = \mathrm{id}$ oraz $I_{\mathrm{id}} = \emptyset$, więc $\operatorname{sgn}(\mathrm{id}) = 1$ na mocy Definicji 1.11. Wobec tego $\operatorname{sgn}(\sigma) \cdot \operatorname{sgn}\left(\sigma^{-1}\right) = 1$, skąd $\operatorname{sgn}(\sigma) = \operatorname{sgn}\left(\sigma^{-1}\right)$.

Z Twierdzenia 1.1 przez prostą indukcję wynika ponadto następujący

Wniosek 1.5. Niech $n \in \mathbb{N}$. Dla dowolnych $r \in \mathbb{N}$ oraz $\sigma_1, \sigma_2 \ldots, \sigma_r \in S_n$ zachodzi wzór:

$$\operatorname{sgn}(\sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_r) = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2) \cdot \ldots \cdot \operatorname{sgn}(\sigma_r).$$

1.3 Podzielność w $\mathbb{Z}$

Definicja 1.13. Mówimy, że liczba całkowita b dzieli liczbę całkowitą a, co zapisujemy $b \mid a$, wówczas gdy istnieje liczna całkowita c taka, że $a = b \cdot c$. W przeciwnym przypadku mówimy, że b nie dzieli a. Fakt ten zapisujemy symbolicznie: $b \nmid a$.

Definicja 1.14. Częścią całkowitą liczby rzeczywistej x nazywamy liczbę $\lfloor x \rfloor$ określoną równością:

$$\lfloor x \rfloor = \max\{y \in \mathbb{Z} \colon y \leq x\}.$$

Przykład 1.10. $\left\lfloor \frac{4}{3} \right\rfloor = 1$, $\left\lfloor -\frac{4}{3} \right\rfloor = -2$.

Twierdzenie 1.2 (o dzieleniu z resztą). Dla dowolnej liczby całkowitej a oraz dowolnej niezerowej liczby całkowitej b istnieje dokładnie jedna para liczb całkowitych (q,r) taka, że $0 \leq r < |b|$ oraz $a = bq + r$.

Dowód. Niech $a, b \in \mathbb{Z}$, przy czym $b \neq 0$. Załóżmy najpierw, że $b > 0$. Wtedy dla $q = \left\lfloor \frac{a}{b} \right\rfloor$ i $r = a - bq$ otrzymujemy, że $q \leq \frac{a}{b} < q + 1$. Zatem $bq \leq a < bq + b$ i w konsekwencji $0 \leq a - bq < b$, czyli $0 \leq r < |b|$. Przypuśćmy teraz, że $b < 0$. Definiujemy wówczas $q = -\left\lfloor \frac{a}{|b|} \right\rfloor$ i $r = a - bq$. Ponieważ $-q \leq \frac{a}{|b|} < -q + 1$ oraz $b = -|b|$, to $bq \leq a < bq - b$. Zatem $0 \leq a - bg < -b$, czyli $0 \leq r < |b|$. W ten sposób wykazaliśmy istnienie żądanej pary liczb (q,r). Pozostało udowodnić jej jednoznaczność. W tym celu załóżmy, że dla liczb całkowitych q_1, q_2, r_1, r_2 zachodzi: $0 \leq r_1 < |b|$, $0 \leq r_2 < |b|$, $a = bq_1 + r_1$ i $a = bq_2 + r_2$. Dwa ostatnie warunki implikują, że $r_2 - r_1 = b(q_1 - q_2)$, skąd $|b| \mid r_2 - r_1$. Ponadto na mocy dwóch pierwszych warunków uzyskujemy oszacowanie $-|b| < r_2 - r_1 < |b|$, więc $r_2 - r_1 = 0$ i w konsekwencji $r_2 = r_1$ oraz $b(q_1 - q_2) = 0$. Ponieważ $b \neq 0$, to ostatnia równość oznacza, że $q_2 = q_1$.

Definicja 1.15. Niech a i b będą liczbami całkowitymi takimi, że $a^2 + b^2 > 0$. Liczbę naturalną d nazywamy największym wspólnym dzielnikiem liczb a i b wówczas, gdy spełniony jest układ warunków:

(i) $d \mid a$ i $d \mid b$;
(ii) jeżeli $d' \in \mathbb{N}$ oraz $d' \mid a$ i $d' \mid b$, to $d' \leq d$.

Największy wspólny dzielnik liczb całkowitych a i b oznaczamy przez $\mathrm{NWD}(a,b)$.

Uwaga 1.11. Warunek $a^2+b^2>0$ jaki spełniają liczby całkowite a i b w Definicji 1.15 oznacza, że co najmniej jedna spośród tych liczb jest niezerowa, co formalnie można zapisać również w postaci: $a\neq 0$ lub $b\neq 0$. Ponieważ $k\mid 0$ dla dowolnej liczby całkowitej k, to dla $a=0$ i $b=0$ napis $\mathrm{NWD}(a,b)$ pozbawiony jest sensu.

Uwaga 1.12. W oparciu o Zasadę maksimum, na mocy której w każdym niepustym ograniczonym podzbiorze zbioru liczb naturalnych istnieje liczba największa, dowodzi się, że dla dowolnych liczb całkowitych a i b takich, że $a^2+b^2>0$, istnieje ich największy wspólny dzielnik. W praktyce największy wspólny dzielnik dwóch liczb całkowitych oblicza się, stosując znany z wcześniejszych etapów edukacji Algorytm Euklidesa lub rozkład na czynniki pierwsze. Przy rozwiązywaniu zadań wykorzystuje się też często równość $\mathrm{NWD}(a,b)=\mathrm{NWD}\left(|a|,|b|\right)$ zachodzącą dla wszystkich $a,b\in\mathbb{Z}$ takich, że $a^2+b^2>0$.

Stwierdzenie 1.4. Niech a,b,c będą liczbami całkowitymi takimi, że $a^2+b^2>0$. Równanie $ax+by=c$ posiada rozwiązanie w zbiorze liczb całkowitych wtedy i tylko wtedy, gdy $\mathrm{NWD}(a,b)\mid c$.

Dowód. Oznaczmy $d=\mathrm{NWD}(a,b)$. Załóżmy, że istnieją $x,y\in\mathbb{Z}$ takie, że $ax+by=c$. Ponieważ $d\mid a$ i $d\mid b$, to $d\mid ax+by$, czyli $d\mid c$.

Przypuśćmy teraz, że $d\mid c$. Wtedy $c=dw$ dla pewnego $w\in\mathbb{Z}$. Niech $A=\{au+bv\colon u,v\in\mathbb{Z}\}\cap\mathbb{N}$. Ponieważ $a,b\in\mathbb{Z}$ i $a^2+b^2>0$, to $A\neq\emptyset$. Stąd oraz na mocy Zasady minimum, w zbiorze A istnieje element najmniejszy d'. W szczególności $d'=ax_0+by_0$ dla pewnych $x_0,y_0\in\mathbb{Z}$. Załóżmy nie wprost, że $d'\nmid a$. Istnieją wówczas $k\in\mathbb{Z}$ oraz $r\in\{1,2,\ldots d'-1\}$ takie, że $a=kd'+r$. Zatem $r=a-kd'=(1-kx_0)a-kby_0=a(1-kx_0)+b(-ky_0)\in A$. Ale $r<d'$, sprzeczność. Stąd $d'\mid a$. Analogicznie pokazuje się, że $d'\mid b$. Oznacza to, że d' jest wspólnym dzielnikiem liczb a i b. Wobec tego $d'\leq d$. Z drugiej strony, $d\mid a$ i $d\mid b$, więc $d\mid ax_0+by_0=d'$. Stąd $d\leq d'$. Wobec tego $d'=d$, czyli $ax_0+by_0=d$. Mnożąc obustronnie ostatnią równość przez w i podstawiając $x=wx_0$ i $y=wy_0$ otrzymujemy $ax+by=c$.

Twierdzenie 1.3. Niech a,b,c będą liczbami całkowitymi takimi, że $a^2+b^2>0$ oraz $\mathrm{NWD}(a,b)\mid c$. Wówczas istnieją $x_0,y_0\in\mathbb{Z}$ takie, że $ax_0+by_0=c$ oraz wszystkie rozwiązania równania $ax+by=c$ w liczbach całkowitych opisane są wzorami:

$$\begin{cases} x=x_0+\frac{b}{\mathrm{NWD}(a,b)}t \\ y=y_0-\frac{a}{\mathrm{NWD}(a,b)}t \end{cases}, \tag{1.3.1}$$

przy czym t przebiega cały zbiór $\mathbb{Z}$.

Dowód. Istnienie $x_0,y_0\in\mathbb{Z}$ spełniających równanie $ax+by=c$ wynika wprost ze Stwierdzenia 1.4. Bezpośredni rachunek pokazuje, że dla dowolnego $t\in\mathbb{Z}$, liczby całkowite x, y opisane wzorami $(1.3.1)$ spełniają równanie $ax+by=c$. Rozważmy teraz dowolne $X,Y\in\mathbb{Z}$ takie, że $aX+bY=c$. Wtedy $aX+bY=ax_0+by_0$, skąd

$a(X-x_0)=b(y_0-Y)$. Zatem $\frac{a}{\mathrm{NWD}(a,b)}(X-x_0)=\frac{b}{\mathrm{NWD}(a,b)}(y_0-Y)$. Ponadto liczby $\frac{a}{\mathrm{NWD}(a,b)}$ i $\frac{b}{\mathrm{NWD}(a,b)}$ są względnie pierwsze, więc z Zasadniczego twierdzenia arytmetyki wynika, że $\frac{b}{\mathrm{NWD}(a,b)} \mid X-x_0$. Wobec tego $X-x_0=\frac{b}{\mathrm{NWD}(a,b)}t$ dla pewnego $t\in\mathbb{Z}$ i konsekwencji $X=x_0+\frac{b}{\mathrm{NWD}(a,b)}t$. Analogicznie pokazuje się, że $Y=y_0-\frac{a}{\mathrm{NWD}(a,b)}k$ dla pewnego $k\in\mathbb{Z}$. Ponadto $c=aX+bY=ax_0+\frac{ab}{\mathrm{NWD}(a,b)}t+by_0-\frac{ba}{\mathrm{NWD}(a,b)}k=(ax_0+by_0)+\frac{ab}{\mathrm{NWD}(a,b)}(t-k)=c+\frac{ab}{\mathrm{NWD}(a,b)}(t-k)$. Jeżeli $a\neq 0$ i $b\neq 0$, to $k=t$, więc rozwiązanie (X,Y) jest opisane wzorami (1.3.1). Jeśli $a=0$, to $Y=\frac{c}{b}=y_0$ oraz X jest dowolną liczbą całkowitą. Zatem $Y=y_0-\frac{0}{\mathrm{NWD}(0,b)}t=y_0-\frac{a}{\mathrm{NWD}(a,b)}t$ i $X=x_0+t=x_0+\frac{b}{\mathrm{NWD}(0,b)}t=x_0+\frac{b}{\mathrm{NWD}(a,b)}t$, skąd rozwiązanie (X,Y) jest opisane wzorami (1.3.1). Analogicznie pokazuje się, że gdy $a\neq 0$ i $b=0$, to rozwiązanie (X,Y) także opisane jest wzorami (1.3.1).

Twierdzenie 1.4 (Podstawowe twierdzenie arytmetyki). Niech a, b oraz c będą niezerowymi liczbami całkowitymi. Jeżeli $c \mid ab$ i $\mathrm{NWD}(c,a)=1$, to $c \mid b$.

Dowód. Załóżmy, że $c \mid ab$ i $\mathrm{NWD}(c,a)=1$. Wtedy $ab=kc$ dla pewnego $k\in\mathbb{Z}$ oraz Stwierdzenie 1.4 implikuje istnienie takich $x,y\in\mathbb{Z}$, że $ax+cy=1$. Zatem $b=abx+bcy=kcx+bcy=c(kx+by)$, skąd $b \mid c$.

1.4 Relacja kongruencji na $\mathbb{Z}$

Definicja 1.16. Niech $m\in\mathbb{N}$. Dla wszystkich $a,b\in\mathbb{Z}$ definiujemy:

$$a\equiv b \pmod{m} \Leftrightarrow m \mid b-a.$$

Tak określoną relację $\cdot\equiv\cdot \pmod{m}\subseteq\mathbb{Z}\times\mathbb{Z}$ nazywamy kongruencją na zbiorze $\mathbb{Z}$ o module m. Napis $a\equiv b \pmod{m}$ czytamy: a przystaje do b modulo m.

Przykład 1.11. $12\equiv 20 \pmod{8}$, bo $8 \mid 20-12=8$.

Stwierdzenie 1.5. Niech $m\in\mathbb{N}$. Dla dowolnych $a,b,c,d\in\mathbb{Z}$ zachodzi:

(i) $a\equiv a \pmod{m}$;
(ii) jeżeli $a\equiv b \pmod{m}$, to $b\equiv a \pmod{m}$;
(iii) jeżeli $a\equiv b \pmod{m}$ i $b\equiv c \pmod{m}$, to $a\equiv c \pmod{m}$;
(iv) $ac\equiv bc \pmod{m} \Leftrightarrow a\equiv b \pmod{\frac{m}{\mathrm{NWD}(m,c)}}$;
(v) jeżeli $a\equiv b \pmod{m}$ i $c\equiv d \pmod{m}$, to:

(a) $a+c\equiv b+d \pmod{m}$,
(b) $a-c\equiv b-d \pmod{m}$,
(c) $a\cdot c\equiv b\cdot d \pmod{m}$.

Dowód. Weźmy dowolne $a,b,c,d \in \mathbb{Z}$.

(i). Ponieważ $a-a=0$ i $m \mid 0$, to $a \equiv a \pmod{m}$.

(ii). Załóżmy, że $a \equiv b \pmod{m}$. Wtedy $m \mid b-a$, więc istnieje $k \in \mathbb{Z}$ takie, że $b-a=km$. Zatem $a-b=(-k)m$. Wobec tego dla $l=-k$ otrzymujemy, że $l \in \mathbb{Z}$ oraz $a-b=lm$, skąd $b \equiv a \pmod{m}$.

(iii). Jeżeli $a \equiv b \pmod{m}$ i $b \equiv c \pmod{m}$, to istnieją $k,l \in \mathbb{Z}$ takie, że $b-a=km$ oraz $c-b=lm$, więc $c-a=(c-b)+(b-a)=lm+km=(l+k)m$ i w konsekwencji $a \equiv c \pmod{m}$.

(iv). Niech $D=\mathrm{NWD}(m,c)$. Wtedy $D \mid m$ i $D \mid c$, więc istnieją $l,t \in \mathbb{Z}$ takie, że $m=lD$ i $c=tD$.

Przypuśćmy najpierw, że $ac \equiv bc \pmod{m}$. Wówczas $m \mid bc-ac=(b-a)c$, skąd $(b-a)c=km$ dla pewnego $k \in \mathbb{Z}$. Zatem $(b-a)tD=klD$, skąd $(b-a)t=kl$. Wobec tego $l \mid (b-a)t$. Ponadto z określenia liczby D wynika, że $\mathrm{NWD}(l,t)=1$, więc z Twierdzenia 1.4 wynika, że $l \mid b-a$, czyli $\frac{m}{D} \mid b-a$. Oznacza to, że $a \equiv b \pmod{\frac{m}{D}}$.

Na odwrót. Załóżmy teraz, że $a \equiv b \pmod{\frac{m}{D}}$. Wtedy $\frac{m}{D} \mid b-a$, więc $b-a=h \cdot \frac{m}{D}$ dla pewnego $h \in \mathbb{Z}$. Zatem $(b-a)D=hm$, skąd $(b-a)Dt=(ht)m$, czyli $(b-a)c=(ht)m$. Wobec tego $m \mid (b-a)c=bc-ac$, co oznacza, że $ac \equiv bc \pmod{m}$.

(v). Załóżmy, że $a \equiv b \pmod{m}$ i $c \equiv d \pmod{m}$. Wtedy $m \mid b-a$ i $m \mid d-c$, więc istnieją $k,l \in \mathbb{Z}$ takie, że $b-a=km$ oraz $d-c=lm$. Zatem $(b+d)-(a+c)=(b-a)+(d-c)=km+lm=(k+l)m$, czyli $a+c \equiv b+d \pmod{m}$. W ten sposób udowodniliśmy punkt (a). Analogicznie dowodzi się punkty (b) i (c).

Wniosek 1.6. Z punktów $(i)-(iii)$ powyższego stwierdzenia wynika, że dla dowolnego $m \in \mathbb{N}$, relacja $\cdot \equiv \cdot \pmod{m} \subseteq \mathbb{Z} \times \mathbb{Z}$ jest zwrotna, symetryczna i przechodnia. Jest więc ona relacją równoważności.

1.5 Pojęcie pierścienia

Definicja 1.17. Pierścieniem nazywamy system algebraiczny $(R,+,\cdot,0,1)$ spełniający układ następujących warunków:

(R1) $(R,+,0)$ jest grupą abelową;
(R2) $a \cdot b = b \cdot a$ dla wszystkich $a,b \in R$;
(R3) $a \cdot (b+c) = a \cdot b + a \cdot c$ dla wszystkich $a,b,c \in R$;
(R4) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ dla wszystkich $a,b,c \in R$;
(R5) $1 \cdot a = a$ dla każdego $a \in R$.

Uwaga 1.13. Grupę $(R,+,0)$ nazywamy grupą addytywną pierścienia i oznaczamy ją symbolem R^+.

Definicja 1.18. Element b pierścienia R nazywamy elementem przeciwnym do elementu a tego pierścienia wówczas, gdy $a+b=0$ (czyli gdy b jest elementem przeciwnym do a w grupie R^+).

Definicja 1.19. Element a pierścienia R nazywamy odwracalnym, gdy istnieje $b \in R$ takie, że $a \cdot b = 1$. Jeśli taki element b istnieje, to nazywamy go elementem odwrotnym do a.

Przykład 1.12. System algebraiczny $(\mathbb{Z},+,\cdot,0,1)$ jest pierścieniem.

Przykład 1.13. System algebraiczny $\big(\{a\},*,*,a,a\big)$, gdzie a jest dowolnym przedmiotem, zaś $*$ działaniem wyznaczonym przez wzór $a*a=a$ jest pierścieniem. Pierścień ten nazywamy pierścieniem zerowym i oznaczamy symbolem $\{0\}$.

Definicja 1.20. Zbiór wszystkich elementów odwracalnych pierścienia R oznaczamy symbolem R^*.

Stwierdzenie 1.6. Dla dowolnego pierścienia R, zbiór R^* rozważany wraz z mnożeniem pierścienia R jest grupą abelową.

Dowód. Ponieważ $1 \in R^*$, to $R^* \neq \emptyset$. Weźmy dowolne $a,b \in R^*$. Wtedy istnieją $a^{-1}, b^{-1} \in R$ oraz $(a \cdot b)(b^{-1} \cdot a^{-1}) = a \cdot \big(b \cdot b^{-1}\big) \cdot a^{-1} = a \cdot 1 \cdot a^{-1} = a \cdot a^{-1} = 1$, więc $ab \in R^*$. Zatem $\cdot$ jest działaniem w zbiorze R^*. To działanie jest oczywiście łączne i przemienne oraz 1 jest jego elementem neutralnym. Ponieważ $a^{-1} \cdot a = 1$, to $a^{-1} \in R^*$.

Przykład 1.14. $\mathbb{Z}^* = \{-1,1\}$.

1.5.1 Pierścień $\mathbb{Z}_n$ reszt modulo n

Przykład 1.15. Niech $n \geq 2$ będzie liczbą naturalną. Dla dowolnej liczby całkowitej x symbolem $[x]_n$ oznaczamy resztę z dzielenia x przez n. Niech ponadto $\mathbb{Z}_n = \{0,1,\ldots,n-1\}$. W zbiorze $\mathbb{Z}_n$ wprowadzamy dwa działania $\oplus_n$ i $\odot_n$ określone za pomocą wzorów:

$$a \oplus_n b = [a+b]_n \quad \text{oraz} \quad a \odot_n b = [a \cdot b]_n,$$

dla wszystkich $a,b \in \mathbb{Z}_n$. Pokażemy, że system algebraiczny $(\mathbb{Z}_n, \oplus_n, \odot_n, 0, 1)$ jest pierścieniem. W tym celu rozważmy dowolne $a,b,c \in \mathbb{Z}_n$. Bezpośrednio z określenia działań $\oplus_n$ i $\odot_n$ wynika ich przemienność oraz, że 0 i 1 są kolejno elementami neutralnymi tych działań. Nietrudno zauważyć, że jeżeli $a \neq 0$, to $n-a \in \mathbb{Z}_n$ oraz

$a \oplus_n (n-a) = 0$. Jeśli natomiast $a = 0$, to $a \oplus_n a = 0$. Dalej, z Twierdzenia 1.2 wynika istnienie określonych jednoznacznie $k,l,t,h \in \mathbb{Z}$ takich, że $(a+b)+c = \big(kn+(a \oplus_n b)\big)+c = kn+\big((a \oplus_n b)+c\big) = kn+\big(ln+((a \oplus_n b) \oplus_n c)\big) = (k+l)n+\big((a \oplus_n b) \oplus_n c\big)$ i $a+(b+c) = (t+h)n+\big(a \oplus_n (b \oplus_n c)\big)$. Ponieważ $(a+b)+c = a+(b+c)$, to powołując się ponownie na Twierdzenie 1.2 otrzymujemy stąd równość $(a \oplus_n b) \oplus_n c = a \oplus_n (b \oplus_n c)$. W pełni analogicznie uzasadnia się równość $(a \odot_n b) \odot_n c = a \odot_n (b \odot_n c)$. Powołując się ponownie na Twierdzenie 1.2, że $a \cdot (b+c) = a \cdot \big(Kn+(b \oplus_n c)\big) = aKn + a \cdot (b \oplus_n c) = aKn + Ln + \big(a \odot_n (b \oplus_n c)\big) = (aK+L)n + \big(a \odot_n (b \oplus_n c)\big)$ oraz $a \cdot b + a \cdot c = Tn + (a \odot_n b) + Hn + (a \odot_n c) = (T+H)n + \big((a \odot_n b) + (a \odot_n c)\big) = (T+H)n + Mn + \big((a \odot_n b) \oplus_n (a \odot_n c)\big) = (T+H+M)n + \big((a \odot_n b) \oplus_n (a \odot_n c)\big)$ dla pewnych $K,L,T,H,M \in \mathbb{Z}$. Ponadto $a \cdot (b+c) = a \cdot b + a \cdot c$, więc powołując się ponownie na Twierdzenie 1.2 otrzymujemy stąd $a \odot_n (b \oplus_n c) = (a \odot_n b) \oplus_n (a \odot_n c)$. W ten sposób pokazaliśmy, że spełnione są wszystkie warunki opisane w Definicji 1.17. Zatem system algebraiczny $(\mathbb{Z}_n, \oplus_n, \odot_n, 0, 1)$ jest pierścieniem. Nazywamy go pierścieniem reszt modulo n.

Uwaga 1.14. W teorii pierścieni obowiązują naturalne odpowiedniki Uwag 1.6 i 1.7 dotyczących teorii grup, tj. w Definicji 1.17 dokonujemy nieformalnego utożsamienia zbioru R z całym pierścieniem (w domyśle znamy działania określone w zbiorze R i elementy wyróżnione). Wtedy możemy używać bardzo naturalnych sformułowań typu: „Niech a będzie elementem pierścienia R", „R wraz z działaniami $+$ i $\cdot$ jest pierścieniem", itp. Na przykład, kiedy mówimy: „Niech a będzie elementem pierścienia $\mathbb{Z}_6$", to mamy na myśli, że a jest elementem zbioru $\{0,1,2,3,4,5\}$ rozważanego wraz z działaniami $\oplus_6$ i $\odot_6$, co odpowiada formalnemu rozważeniu systemu algebraicznego $(\mathbb{Z}_6, \oplus_6, \odot_6, 0, 1)$.

Uwaga 1.15. Przypomnijmy, że relacje równoważności pozwalają utożsamiać elementy będące ze sobą w relacji. Niech $n \geq 2$ będzie liczbą naturalną. Z Twierdzenia o dzieleniu z resztą i określenia pierścienia $\mathbb{Z}_n$ wynika, że dla każdego $k \in \mathbb{Z}$ istnieje dokładnie jedno $r \in \mathbb{Z}_n$ takie, że $k \equiv r \pmod{n}$ ($r = [k]_n$). Oznacza to, że każdą liczbę całkowitą k możemy utożsamić względem relacji $\cdot \equiv \cdot \pmod{n}$ z dokładnie jednym elementem pierścienia $\mathbb{Z}_n$ będącym resztą z dzielenia liczby k przez n.

Uwaga 1.16. Pierścień $\mathbb{Z}_n$ można również określić dla $n = 1$. Otrzymamy wówczas pierścień zerowy opisany w Przykładzie 1.13. Nie ma on jednak żadnych interesujących własności.

Przykład 1.16. W pierścieniu $\mathbb{Z}_9$ elementem odwrotnym do 5 jest 2, gdyż $5 \odot_9 2 = 1$. Fakt ten można zapisać w następujący sposób: $5^{-1} = 2$ w $\mathbb{Z}_9$. Elementem przeciwnym do 5 w pierścieniu $\mathbb{Z}_9$ jest 4, bo $5 \oplus_9 4 = 0$. Możemy ten fakt zapisać również tak: $-5 = 4$ w $\mathbb{Z}_9$. Można używać też następujących sformułowań: „2 jest elementem odwrotnym do 5 modulo 9" oraz „4 jest elementem przeciwnym do 5 modulo 9".

Stwierdzenie 1.7. Niech $n \geq 2$ będzie liczbą naturalną i niech $a \in \mathbb{Z}_n$. Wówczas a jest elementem odwracalnym pierścienia $\mathbb{Z}_n$ wtedy i tylko wtedy, gdy $\mathrm{NWD}(a,n) = 1$.

Dowód. Ze Stwierdzenia 1.4 wynika, że równość $\mathrm{NWD}(a,n) = 1$ równoważna jest istnieniu $x, y \in \mathbb{Z}$ takich, że $ax + ny = 1$.

Jeśli takie x i y istnieją, to istnieją także określone jednoznacznie $k \in \mathbb{Z}$ i $b \in \{0, 1, \ldots, n-1\}$ takie, że $x = kn + b$. Zatem $ab = 1 - ny - akn = 1 - n(y - ak)$, skąd wynika, że $a \odot_n b = 1$ czyli, że a jest elementem odwracalnym w $\mathbb{Z}_n$.

Na odwrót. Jeżeli a jest elementem odwracalnym w $\mathbb{Z}_n$, to $a \odot_n c = 1$ dla pewnego $c \in \mathbb{Z}_n$. Stąd oraz na mocy określenia działania $\odot_n$, $ac = 1 + tn$ dla pewnego $t \in \mathbb{Z}$. Zatem $ac + n(-t) = 1$. Ponadto $c, t \in \mathbb{Z}$, więc wystarczy przyjąć $x = c$ oraz $y = -t$.

Stwierdzenie 1.7 pozwala w prosty sposób weryfikować odwracalność dowolnego elementu a pierścienia $\mathbb{Z}_n$. Natomiast Twierdzenie 1.3 wraz z poznanym w szkole średniej Algorytmem Euklidesa i własnościami relacji kongruencji $\cdot \equiv \cdot \pmod{n}$ (por. Uwaga 1.15) pozwala w efektywny sposób znaleźć odwrotność elementu $a \in \mathbb{Z}_n$, o ile jest on odwracalny. Jest to zilustrowane w poniższym przykładzie.

Przykład 1.17. Rozważmy element 16 pierścienia $\mathbb{Z}_{75}$. Ponieważ $16 = 2^4$ i $2 \nmid 75$, to $\mathrm{NWD}(16, 75) = 1$. Stąd oraz na mocy Stwierdzenia 1.7 otrzymujemy, że $16 \in \mathbb{Z}_{75}^*$ (tzn. 16 jest elementem odwracalnym pierścienia $\mathbb{Z}_{75}$). Znajdziemy teraz odwrotność 16 w $\mathbb{Z}_{75}$. W tym celu wyznaczymy $\mathrm{NWD}(75, 16)$, korzystając ze znanego Algorytmu Euklidesa:

$$\begin{gathered} 75 = 16 \cdot 4 + 11 \\ 16 = 11 \cdot 1 + 5 \\ 11 = 5 \cdot 2 + \underline{1} \\ 5 = 1 \cdot 5 + 0. \end{gathered}$$

Zatem $\mathrm{NWD}(75, 16) = 1$ (co nie jest zaskoczeniem wobec początkowych obserwacji). Dalej możemy postąpić na dwa sposoby:

(i) Z powyższych rachunków wynika, że $1 = 11 - 5 \cdot 2 = (75 - 16 \cdot 4) - (16 - 11) \cdot 2 = 75 - 16 \cdot 6 + 11 \cdot 2 = 75 - 16 \cdot 6 + (75 - 16 \cdot 4) \cdot 2 = 75 \cdot 3 - 16 \cdot 14$. Zatem $75 \cdot 3 + 16 \cdot (-14) = 1$, skąd wynika, że $(x_0, y_0) = (3, -14)$ jest rozwiązaniem równania $75x + 16y = 1$. Na mocy Twierdzenia 1.3 otrzymujemy, że wszystkimi całkowitymi rozwiązaniami tego równania są pary (x, y) takie, że:

$$\begin{cases} x = 3 + 16t \\ y = -14 - 75t \end{cases},$$

gdzie $t \in \mathbb{Z}$. Wśród tych rozwiązań szukamy teraz takiego, że $y \in \mathbb{Z}_{75}$. Takie rozwiązanie uzyskamy dla $t = -1$. Jest nim $(X, Y) = (-13, 61)$. Stąd $75 \cdot (-13) + 16 \cdot 61 = 1$. Zatem $16 \cdot 61 = 1 + 13 \cdot 75$ i w konsekwencji $16 \odot_{75} 61 = 1$. Wobec tego $16^{-1} = 61$ w $\mathbb{Z}_{75}$.

(*ii*) Z rachunków związanych z Algorytmem Euklidesa i własności kongruencji $\cdot \equiv \cdot \pmod{75}$ wynika, że $1 \equiv 11 - 5 \cdot 2 \equiv (75 - 16 \cdot 4) - (16 - 11) \cdot 2 \equiv 0 - 16 \cdot 6 + 11 \cdot 2 \equiv -16 \cdot 6 + (-16 \cdot 4) \cdot 2 \equiv -16 \cdot 14 \equiv 16 \cdot (-14) \equiv 16 \cdot 61 \pmod{75}$. Zatem $16 \odot_{75} 61 = 1$, czyli $16^{-1} = 61$ w $\mathbb{Z}_{75}$.

1.6 Pojęcie ciała

Definicja 1.21. Pierścień $(K, +, \cdot, 0, 1)$ nazywamy ciałem, gdy $|K| \geq 2$ oraz dla każdego $a \in K \setminus \{0\}$, istnieje $b \in K$ takie, że $a \cdot b = 1$.

W świetle Definicji 1.17 i 1.19, ciało możemy zdefiniować też w następujący sposób:

Definicja 1.22. Niezerowy pierścień K, w którym każdy niezerowy element jest odwracalny nazywamy ciałem.

Przykład 1.18. Przykładami ciał są $(\mathbb{Q}, +, \cdot, 0, 1)$ i $(\mathbb{R}, +, \cdot, 0, 1)$. Ponieważ $2 \notin \mathbb{Z}^*$, to pierścień $(\mathbb{Z}, +, \cdot, 0, 1)$ nie jest ciałem.

Wprost z Definicji 1.21 i 1.22 oraz Przykładu 1.18 uzyskujemy następujący

Wniosek 1.7. Każde ciało jest pierścieniem. Istnieją pierścienie niebędące ciałami.

Przykład 1.19. Można łatwo wykazać, że zbiór $\mathbb{Q}\left(\sqrt{2}\right) = \left\{a + b\sqrt{2} \colon a, b \in \mathbb{Q}\right\}$ rozważany wraz ze standardowymi działaniami $+$ i $\cdot$ jest ciałem.

Twierdzenie 1.5. Niech $n \geq 2$ będzie liczbą naturalną. Pierścień $\mathbb{Z}_n$ jest ciałem wtedy i tylko wtedy, gdy n jest liczbą pierwszą.

Dowód. Załóżmy nie wprost, że $n \notin \mathbb{P}$ oraz pierścień $\mathbb{Z}_n$ jest ciałem. Wtedy $n = rs$ dla pewnych liczb naturalnych $1 < r, s < n$ (bo $n > 1$). W szczególności $r \in \mathbb{Z}_n \setminus \{0\}$, więc r jest elementem odwracalnym w $\mathbb{Z}_n$. Stąd oraz na mocy Stwierdzenia 1.7, $\mathrm{NWD}(r, n) = 1$. Ale $r \mid n$, więc $\mathrm{NWD}(r, n) = r > 1$, sprzeczność.

Na odwrót. Przypuśćmy, że $n \in \mathbb{P}$. Wtedy $|\mathbb{Z}_n| \geq 2$. Rozważmy dowolne $a \in \mathbb{Z}_n \setminus \{0\}$. Wówczas $\mathrm{NWD}(a, n) = 1$, więc a jest elementem odwracalnym w $\mathbb{Z}_n$ na mocy Stwierdzenia 1.7. Wobec tego pierścień $\mathbb{Z}_n$ jest ciałem.

Przykład 1.20. W zbiorze $\mathbb{R} \times \mathbb{R}$ wprowadzamy dwa dwuargumentowe działania:

$$(a, b) \boxplus (c, d) = (a + c, b + d)$$

oraz

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

Wtedy system algebraiczny $\big(\mathbb{R}\times\mathbb{R},\boxplus,\boxdot,(0,0),(1,0)\big)$ jest ciałem. Istotnie, nietrudno zauważyć, że system algebraiczny $\big(\mathbb{R}\times\mathbb{R},\boxplus,(0,0)\big)$ jest grupą abelową (bo $(\mathbb{R},+,0)$ jest grupą abelową oraz działanie $\boxplus$ jest określone „po współrzędnych"). Ponadto, dla dowolnych $a,b,c,d,x,y\in\mathbb{R}$ zachodzą równości: $(c,d)\boxdot(a,b)=(ca-db,cb+da)=(ac-bd,ad+bc)=(a,b)\boxdot(c,d)$, $(x,y)\boxdot\big((a,b)\boxplus(c,d)\big)=(x,y)\boxdot(a+c,b+d)=\big(x(a+c)-y(b+d),x(b+d)+y(a+c)\big)=\big((xa-yb)+(xc-yd),(xb+ya)+(xd+yc)\big)=(xa-yb,xb+ya)\boxplus(xc-yd,xd+yc)=(x,y)\boxdot(a,b)\boxplus(x,y)\boxdot(c,d)$, $\big((a,b)\boxdot(c,d)\big)\boxdot(x,y)=(ac-bd,ad+bc)\boxdot(x,y)=\big((ac-bd)x-(ad+bc)y,(ac-bd)y+(ad+bc)x\big)=\big(a(cx-dy)-b(cy+dx),a(cy+dx)+b(cx-dy)\big)=(a,b)\boxdot(cx-dy,cy+dx)=(a,b)\boxdot\big((c,d)\boxdot(x,y)\big)$ oraz $(1,0)\boxdot(a,b)=(1\cdot a-0\cdot b,1\cdot b+0\cdot a)=(a,b)$. Z dotychczasowych obserwacji wynika więc, że rozważany system algebraiczny jest pierścieniem. Jasne jest, że pierścień ten ma więcej niż jeden element (ma on nieskończoną ilość elementów!). Jeśli $(a,b)\neq(0,0)$ i $u,v\in\mathbb{R}$, to $(a,b)\boxdot(u,v)=(1,0)$ wtedy i tylko wtedy, gdy $(au-bv,av+bu)=(1,0)$, czyli gdy $au-bv=1$ oraz $av+bu=0$. Jeżeli $a\neq 0$, to z drugiej równości otrzymujemy, że $v=\frac{-bu}{a}$. Podstawiając to pierwszej równości uzyskujemy $u=\frac{a}{a^2+b^2}$. Stąd $v=\frac{-b}{a^2+b^2}$. Zatem $(u,v)=\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)$. Jeśli natomiast $a=0$, to $b\neq 0$ i postępując analogicznie, ponownie uzyskamy, że $(u,v)=\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)$. Wobec tego $(a,b)^{-1}=\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)$.

Definicja 1.23. Ciało skonstruowane w Przykładzie 1.20 nazywamy ciałem liczb zespolonych i oznaczamy symbolem $\mathbb{C}$. Elementy tego ciała nazywamy liczbami zespolonymi.

Własności liczb zespolonych zostaną szczegółowo omówione w ramach następnego rozdziału.

Rozdział 2

Liczby zespolone i ich własności

2.1 Postać algebraiczna liczby zespolonej

Uwaga 2.1. Na mocy określenia liczb zespolonych (zob. Definicja 1.23 i Przykład 1.20), dla wszystkich $a,b,c,d \in \mathbb{R}$ mamy $(a,b) = (c,d)$ wtedy i tylko wtedy, gdy $a = c$ i $b = d$, skąd w szczególności uzyskujemy, że $(a,0) = (b,0)$ wtedy i tylko wtedy, gdy $a = b$. Ponadto z określenia działań w ciele $\mathbb{C}$ wynika, że prawdziwe są równości $(a,0) \boxplus (b,0) = (a+b,0)$, $(a,0) \boxdot (b,0) = (a \cdot b,0)$ oraz $-(a,0) = (-a,0)$. Ponadto, jeśli $a \neq 0$, to $(a,0)^{-1} = \left(a^{-1},0\right)$. Dlatego, dla dowolnej liczby rzeczywistej x można dokonać utożsamienia: $x \equiv (x,0) \in \mathbb{C}$. Możemy wówczas pisać $\mathbb{R} \subseteq \mathbb{C}$.

Definicja 2.1. Jednostką urojoną nazywamy liczbę zespoloną $(0,1)$, którą oznaczamy symbolem i.

Uwaga 2.2. Korzystając z utożsamienia opisanego w Uwadze 2.1 otrzymujemy bardzo ważny wzór:

$$i^2 = -1. \tag{2.1.1}$$

Istotnie, $i^2 = (0,1) \boxdot (0,1) = (-1,0) \equiv -1$. Ponadto dla dowolnych $a,b \in \mathbb{R}$ mamy $(a,b) = (a,0) \boxplus (0,b) = (a,0) \boxplus (b,0) \boxdot (0,1) = (a,0) \boxplus (b,0) \boxdot i$. Z tego powodu dokonujemy utożsamienia $(a,b) \equiv a+bi$. W ten sposób otrzymujemy tzw. postać algebraiczną $a+bi$ liczby zespolonej (a,b). Ponieważ mnożenie liczb zespolonych jest przemienne, to stosując wprowadzone dotychczas utożsamienia uzyskujemy równość $a+bi = a+ib$ i w konsekwencji również $a+ib$ nazywa się postacią algebraiczną liczby zespolonej (a,b).

Uwaga 2.3. W praktyce, przy zapisywaniu działań na liczbach zespolonych, używa się standardowych oznaczeń tych działań, tj.: „+”, „−”, „·”, „:” (gdyż jest to wygodne i na ogół nie prowadzi do nieporozumień). Symbol „·” często się pomija. Symbolu „:” nie używa się dla liczb zespolonych zapisanych w postaci par uporządkowanych; zamiast $(a,b) : (c,d)$ pisze się $(a,b) \cdot (c,d)^{-1}$.

Wniosek 2.1. Dodawanie, odejmowanie i mnożenie liczb zespolonych zapisanych w postaci algebraicznej wykonujemy dokładnie tak samo jak dodawanie i odejmowanie wielomianów zmiennej i. Przy mnożeniu liczb zespolonych zapisanych w postaci algebraicznej pamiętamy dodatkowo o równości $(2.1.1)$, której zastosowanie pozwoli nam przedstawić wynik w postaci algebraicznej. Dzielenie liczb zespolonych zostanie omówione w dalszej części rozdziału.

Przykład 2.1. $\big((5+3i)+(-2-i)\big)\cdot(1+2i)-(7+5i)=(3+2i)\cdot(1+2i)-(7+5i)=(-1+8i)-(7+5i)=(-8+3i)$.

Definicja 2.2. Liczbą sprzężoną z liczbą zespoloną $z=a+bi$ nazywamy liczbę zespoloną $\overline{z}$ określoną równością $\overline{z}=a-bi$.

Przykład 2.2. $\overline{3+2i}=3-2i$, $\overline{1-i}=1+i$.

Definicja 2.3. Modułem liczby zespolonej $z=a+bi$ nazywamy liczbę rzeczywistą $\sqrt{a^2+b^2}$. Moduł liczby zespolonej z oznaczamy symbolem $|z|$.

Przykład 2.3. $|3+4i|=\sqrt{3^2+4^2}=\sqrt{25}=5$.

Stwierdzenie 2.1. Dla dowolnej liczby zespolonej z zachodzi równość:

$$z\cdot\overline{z}=|z|^2.$$

Dowód. Rozważmy dowolne $z\in\mathbb{C}$. Wtedy $z=a+bi$ dla pewnych $a,b\in\mathbb{R}$. Stąd $z\cdot\overline{z}=(a+bi)(a-bi)=a^2+b^2=|z|^2$.

Stwierdzenie 2.1 znajduje praktyczne zastosowanie w rachunkach. Świadczy o tym poniższy

Przykład 2.4. $\frac{3-2i}{2+i}=\frac{(3-2i)\cdot\overline{2+i}}{(2+i)\cdot\overline{2+i}}=\frac{(3-2i)(2-i)}{|2+i|^2}=\frac{4-7i}{2^2+1^2}=\frac{4-7i}{5}$.

Uwaga 2.4. W powyższym przykładzie, tak jak w przypadku rachunków w zbiorze liczb rzeczywistych, kreska ułamkowa oznacza dzielenie. Dzielenie przez $2+i$ oznacza mnożenie przez element odwrotny do $2+i$, który możemy wyznaczyć w oparciu o wzory otrzymane pod koniec Przykładu 1.20. Mamy mianowicie $\frac{3-2i}{2+i}=(3-2i)\cdot(2+i)^{-1}=(3+2i)\cdot\left(\frac{2}{2^2+1^2}+\frac{-1}{2^2+1^2}\cdot i\right)=\frac{(3-2i)(2-i)}{5}=\frac{4-7i}{5}$.

Definicja 2.4. Częścią rzeczywistą liczby zespolonej $z=a+bi$ nazywamy liczbę rzeczywistą a. Liczbę rzeczywistą b nazywamy częścią urojoną liczby zespolonej z. Część rzeczywistą i część urojoną liczby zespolonej z oznaczamy odpowiednio przez $\mathrm{re}(z)$ i $\mathrm{im}(z)$.

Przykład 2.5. $\mathrm{re}(5+2i)=5$ oraz $\mathrm{im}(5+2i)=2$.

Stwierdzenie 2.2. Dla dowolnej liczby naturalnej $n\geq 2$ oraz dowolnych $z,z_1,z_2,\ldots,z_n\in\mathbb{C}$ i $w\in\mathbb{C}\setminus\{0\}$ zachodzą następujące równości:

(i) $\overline{\sum_{k=1}^n z_k}=\sum_{k=1}^n\overline{z_k}$;
(ii) $\overline{\prod_{k=1}^n z_k}=\prod_{k=1}^n\overline{z_k}$;
(iii) $\overline{z^n}=(\overline{z})^n$;
(iv) $\overline{\left(\frac{z}{w}\right)}=\frac{\overline{z}}{\overline{w}}$;
(v) $|\prod_{k=1}^n z_k|=\prod_{k=1}^n|z_k|$;

(vi) $|z^n| = |z|^n$;
(vii) $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$;
(viii) $\left|\sum_{k=1}^{n} z_k\right| \leq \sum_{k=1}^{n} |z_k|$.

Dowód. (i). Dla $n = 2$ mamy $z_1 = a_1 + b_1 i$ oraz $z_2 = a_2 + b_2 i$, gdzie a_1, a_2, b_1, b_2 są pewnymi liczbami rzeczywistymi, więc $\overline{z_1 + z_2} = \overline{(a_1 + a_2) + (b_1 + b_2)i} = (a_1 + a_2) - (b_1 + b_2)i = (a_1 - b_1 i) + (a_2 - b_2 i) = \overline{a_1 + b_1 i} + \overline{a_2 + b_2 i} = \overline{z_1} + \overline{z_2}$. Stosując prostą indukcję uzyskujemy stąd tezę.
(ii). Udowodnimy tezę dla $n = 2$. Dla pozostałych n wynika ona przez prostą indukcję. Istnieją $a_1, a_2, b_1, b_2 \in \mathbb{R}$ takie, że $z_1 = a_1 + b_1 i$ oraz $z_2 = a_2 + b_2 i$. Stąd $\overline{z_1 \cdot z_2} = \overline{(a_1 + b_1 i) \cdot (a_2 + b_2 i)} = \overline{(a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i} = (a_1 a_2 - b_1 b_2) - (a_1 b_2 + a_2 b_1)i = (a_1 - b_1 i) \cdot (a_2 - b_2 i) = \overline{a_1 + b_1 i} \cdot \overline{a_2 + b_2 i} = \overline{z_1} \cdot \overline{z_2}$.
(iii). Wynika wprost z (ii) (wystarczy podstawić $z_1 = z_2 = \ldots = z_n = z$).
(iv). Z (ii) wynika, iż $\overline{z} = \overline{w \cdot \frac{z}{w}} = \overline{w} \cdot \overline{\left(\frac{z}{w}\right)}$. Ponadto $\overline{w} \neq 0$, gdyż $w \neq 0$, więc po obustronnym podzieleniu ostatniej równości przez $\overline{w}$ otrzymujemy tezę.
(v). Udowodnimy tezę tylko dla $n = 2$, gdyż dla pozostałych n wynika ona przez prostą indukcję. Ze Stwierdzenia 2.1 oraz (ii) otrzymujemy, że $|z_1 \cdot z_2|^2 = z_1 \cdot z_2 \cdot \overline{z_1 \cdot z_2} = z_1 \cdot z_2 \cdot \overline{z_1} \cdot \overline{z_2} = z_1 \cdot \overline{z_1} \cdot z_2 \cdot \overline{z_2} = |z_1|^2 \cdot |z_2|^2 = (|z_1| \cdot |z_2|)^2$. Zatem $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.
(vi). Wynika wprost z (v) (wystarczy podstawić $z_1 = z_2 = \ldots = z_n = z$).
(vii). Na mocy (v) uzyskujemy, że $|z| = \left|w \cdot \frac{z}{w}\right| = |w| \cdot \left|\frac{z}{w}\right|$. Ponadto $|w| \neq 0$, bo $w \neq 0$. Stąd, po obustronnym podzieleniu ostatniej równości przez $|w|$, otrzymujemy tezę.
$(viii)$. Udowodnimy tezę dla $n = 2$. Dla pozostałych n wynika ona przez prostą indukcję. Jeżeli $z_1 + z_2 = 0$, to teza jest oczywista. Niech dalej $z_1 + z_2 \neq 0$. Wówczas $|z_1 + z_2| > 0$. Ponadto $1 = \operatorname{re}\left(\frac{z_1}{z_1+z_2} + \frac{z_2}{z_1+z_2}\right) = \operatorname{re}\left(\frac{z_1}{z_1+z_2}\right) + \operatorname{re}\left(\frac{z_2}{z_1+z_2}\right) \leq \left|\frac{z_1}{z_1+z_2}\right| + \left|\frac{z_2}{z_1+z_2}\right| = \frac{|z_1|}{|z_1+z_2|} + \frac{|z_2|}{|z_1+z_2|}$. Stąd, po obustronnym pomnożeniu przez $|z_1 + z_2|$, uzyskujemy tezę.

2.2 Postać trygonometryczna liczby zespolonej

Uwaga 2.5. Niech $z \in \mathbb{C} \setminus \{0\}$. Istnieją wówczas $x, y \in \mathbb{R}$ takie, że $x^2 + y^2 > 0$ oraz $z = x + iy$. Liczbę zespoloną z możemy traktować jako punkt (x, y) płaszczyzny, którego odległość od początku układu współrzędnych wynosi $\sqrt{x^2 + y^2} = |z|$. Literą φ oznaczmy miarę kąta skierowanego jaki tworzy wektor $\overrightarrow{Oz}$ z osią OX w orientacji płaszczyzny przeciwnej do kierunku ruchu wskazówek zegara (czyli tak jak do tej pory było to przyjmowane w szkole). Wówczas $\varphi \in [0, 2\pi)$, $\cos\varphi = \frac{x}{|z|}$ oraz $\sin\varphi = \frac{y}{|z|}$. Otrzymujemy stąd wzór:

$$z = |z|(\cos\varphi + i\sin\varphi). \tag{2.2.1}$$

Definicja 2.5. Równość (2.2.1) określa tak zwaną postać trygonometryczną niezerowej liczby zespolonej z.

Uwaga 2.6. Nie rozważa się postaci trygonometrycznej liczby zespolonej 0, gdyż nie można jednoznacznie przypisać jej kąta φ określonego w Uwadze 2.5.

Definicja 2.6. Liczbę φ opisaną w Uwadze 2.5 nazywamy argumentem głównym liczby zespolonej z i oznaczamy ją symbolem $\mathrm{Arg}(z)$.

Przykład 2.6. Niech $z = 1+i$. Wówczas $|z| = \sqrt{1^2+1^2} = \sqrt{2}$, $\mathrm{Arg}(z) = \frac{\pi}{4}$ oraz $z = \sqrt{2}\cdot\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$.

Definicja 2.7. Każdą liczbę rzeczywistą ψ postaci $\psi = \varphi + 2k\pi$, gdzie $k \in \mathbb{Z}$, nazywamy argumentem niezerowej liczby zespolonej z opisanej wzorem (2.2.1).

W świetle poniższej uwagi, nie jest nadużyciem wprowadzenie jednego oznaczenia $\arg(z)$ na dowolny argument liczby zespolonej z.

Uwaga 2.7. Zachowując wszystkie wprowadzone dotychczas oznaczenia otrzymujemy, że $z = |z|\big(\cos(\varphi+2k\pi) + i\sin(\varphi+2k\pi)\big)$ dla każdego $k \in \mathbb{Z}$. Ponadto, jeśli $r, \alpha \in \mathbb{R}$ spełniają warunki $r > 0$ oraz $z = r(\cos\alpha + i\sin\alpha)$, to $|z| = |r|\cdot|(\cos\alpha + i\sin\alpha)| = r\cdot\sqrt{\cos^2\alpha+\sin^2\alpha} = r$. Stąd oraz na mocy wzoru (2.2.1), $\cos\alpha = \cos\varphi$ i $\sin\alpha = \sin\varphi$. Wobec tego istnieje $k \in \mathbb{Z}$ takie, że $\alpha = \varphi + 2k\pi$.

Uwaga 2.8. Dla dowolnych $z_1, z_2 \in \mathbb{C}\setminus\{0\}$, zapis $\arg(z_1) = \arg(z_2)$ będziemy interpretowali w taki sposób, że liczby rzeczywiste $\arg(z_1)$ i $\arg(z_2)$ różnią się jedynie o całkowitą wielokrotność liczby 2π.

Stwierdzenie 2.3. Dla dowolnej liczby naturalnej $n \geq 2$ oraz dowolnych $z, z_1, z_2, \ldots, z_n, w \in \mathbb{C}\setminus\{0\}$ i $\alpha \in \mathbb{R}$ zachodzą następujące równości:

(i) $\arg\left(\prod_{k=1}^n z_k\right) = \sum_{k=1}^n \arg(z_k)$;
(ii) $(\cos\alpha + i\sin\alpha)^n = \cos(n\alpha) + i\sin(n\alpha)$;
(iii) $\arg\left(\frac{z}{w}\right) = \arg(z) - \arg(w)$;
(iv) $\arg(\bar{z}) = \arg\left(z^{-1}\right) = -\arg(z)$;
(v) $\arg(-z) = \pi + \arg(z)$.

Dowód. (i). Dla $n = 2$ oznaczmy $\psi_1 = \arg(z_1)$ oraz $\psi_2 = \arg(z_2)$. Wówczas $z_1 = |z_1|\cdot(\cos\psi_1 + i\sin\psi_1)$ oraz $z_2 = |z_2|\cdot(\cos\psi_2 + i\sin\psi_2)$. Zatem $z_1\cdot z_2 = |z_1|\cdot|z_2|\cdot\big((\cos\psi_1\cos\psi_2 - \sin\psi_1\sin\psi_2) + i(\cos\psi_1\sin\psi_2 + \sin\psi_1\cos\psi_2)\big) = |z_1\cdot z_2|\big(\cos(\psi_1+\psi_2) + i\sin(\psi_1+\psi_2)\big)$, skąd $\arg(z_1\cdot z_2) = \arg(z_1) + \arg(z_2)$. Załóżmy indukcyjnie, że rozważana równość zachodzi dla pewnej liczby naturalnej $n \geq 2$. Rozważmy dowolne $z_1, z_2, \ldots, z_n, z_{n+1} \in \mathbb{C}\setminus\{0\}$. Wówczas $\arg\left(\prod_{k=1}^{n+1} z_k\right) = \arg\left(\left(\prod_{k=1}^n z_k\right)\cdot z_{n+1}\right) = \arg\left(\prod_{k=1}^n z_k\right) + \arg(z_{n+1}) = \left(\sum_{k=1}^n \arg(z_k)\right) + \arg(z_{n+1}) = \sum_{k=1}^{n+1}\arg(z_k)$. Zasada indukcji matematycznej kończy dowód.

(ii). Podstawiając $z_1 = z_2 = \ldots = z_n = \cos\alpha + i\sin\alpha$ we wzorze (i) natychmiast uzyskujemy tezę.
(iii). Ponieważ $z = w \cdot \frac{z}{w}$, to z (i) wynika, że $\arg(z) = \arg(w) + \arg\left(\frac{z}{w}\right)$. Przenosząc $\arg(w)$ na drugą stronę tej równości, otrzymujemy tezę.
(iv). Ponieważ $z \cdot \overline{z} \in \mathbb{R}$, to $\arg(z \cdot \overline{z}) = 0$. Stąd oraz na mocy (i), $\arg(z) + \arg(\overline{z}) = 0$ i w konsekwencji $\arg(\overline{z}) = -\arg(z)$. Ponadto z (iii) wynika, że $\arg\left(z^{-1}\right) = \arg\left(\frac{1}{z}\right) = \arg(1) - \arg(z) = 0 - \arg(z) = -arg(z)$.
(v). Ponieważ $\pi = \arg(-1)$, to na mocy (i) otrzymujemy, że $\arg(-z) = \arg((-1) \cdot z) = \arg(-1) + \arg(z) = \pi + \arg(z)$.

Uwaga 2.9. Równość dana w punkcie (ii) powyższego stwierdzenia nosi nazwę wzoru de Moivre'a.

2.3 Pierwiastkowanie liczb zespolonych

Definicja 2.8. Niech $n \in \mathbb{N}$. Pierwiastkiem zespolonym n-tego stopnia z liczby zespolonej z nazywamy każdą liczbę zespoloną ω taką, że $\omega^n = z$.

Uwaga 2.10. Łatwo zauważyć, że dla każdego $n \in \mathbb{N}$ jedynym pierwiastkiem zespolonym n-tego stopnia z 0 jest 0.

Dla wszystkich niezerowych liczb zespolonych zachodzi następujące

Twierdzenie 2.1. Niech $n \in \mathbb{N}$ i niech $z \in \mathbb{C} \setminus \{0\}$. Istnieje wówczas dokładnie n różnych pierwiastków zespolonych n-tego stopnia z liczby z i są nimi liczby zespolone ω_k opisane wzorem:

$$\omega_k = \sqrt[n]{|z|} \cdot \left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right),$$

gdzie $\varphi = \mathrm{Arg}(z)$ oraz k przebiega zbiór $\{0, 1, \ldots, n-1\}$.

Dowód. Wprost ze wzoru de Moivre'a (por. Uwaga 2.9) wynika, że dla każdego $k \in \{0, 1, \ldots, n-1\}$, liczba ω_k jest pierwiastkiem zespolonym stopnia n z liczby z. Aby uzasadnić, że liczby zespolone $\omega_0, \omega_1 \ldots, \omega_{n-1}$ są parami różne, weźmy dowolne $k, l \in \{0, 1, \ldots, n-1\}$ i załóżmy, że $\omega_k = \omega_l$. Wówczas $\frac{\varphi+2k\pi}{n} = \frac{\varphi+2l\pi}{n} + 2h\pi$ dla pewnego $h \in \mathbb{Z}$ (por. Uwaga 2.7), skąd $k - l = hn$. Ale $|k-l| < n$, więc $h = 0$ i w konsekwencji $k = l$. W ten sposób pokazaliśmy, że liczby zespolone $\omega_0, \omega_1 \ldots, \omega_{n-1}$ są parami różnymi pierwiastkami zespolonymi n-tego stopnia z liczby z. Pozostało wykazać, że nie istnieją inne pierwiastki zespolone stopnia n z liczby z. W tym celu załóżmy, że ω jest takim pierwiastkiem. Wtedy $\omega \neq 0$, gdyż $z \neq 0$. Zatem $\omega = |\omega| \cdot (\cos\psi + i\sin\psi)$ dla pewnego $\psi \in \mathbb{R}$. Powołując się ponownie na wzór de Moivre'a otrzymujemy stąd $z = \omega^n = |\omega|^n \cdot \big(cos(n\psi) + i\sin(n\psi)\big)$. Wobec tego

$|z| = |\omega|^n$ oraz $n\psi = \varphi + 2t\pi$ dla pewnego $t \in \mathbb{Z}$. Ponadto $t = sn + k$ dla pewnych określonych jednoznacznie $s \in \mathbb{Z}$ oraz $k \in \{0, 1, \ldots, n-1\}$, więc $\psi = \frac{\varphi + 2k\pi}{n} + 2s\pi$. Ostatecznie otrzymujemy więc, że $\omega = \omega_k$.

Przykład 2.7. Wyznaczymy wszystkie pierwiastki zespolone stopnia 4 z liczby zespolonej $z = \frac{-1}{2} + \frac{\sqrt{3}}{2}i$. Najpierw przedstawiamy liczbę z w postaci trygonometrycznej. W tym celu obliczamy $|z| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2} = \sqrt{\frac{1}{4} + \frac{3}{4}} = 1$, a następnie wyznaczamy argument główny $\varphi \in [0, 2\pi)$ liczby z:

$$\begin{cases} \cos\varphi = \frac{-1}{2} \\ \sin\varphi = \frac{\sqrt{3}}{2} \end{cases} \Leftrightarrow \varphi = \frac{2\pi}{3}.$$

Zatem $z = 1 \cdot \left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right)$. Stąd oraz na mocy Twierdzenia 2.1 otrzymujemy, że

$$\omega_k = \sqrt[4]{1} \cdot \left(\cos\frac{\frac{2\pi}{3} + 2k\pi}{4} + i\sin\frac{\frac{2\pi}{3} + 2k\pi}{4}\right),$$

dla $k \in \{0, 1, 2, 3\}$, czyli:

$$\omega_k = \cos\frac{\pi + 3k\pi}{6} + i\sin\frac{\pi + 3k\pi}{6},$$

dla $k \in \{0, 1, 2, 3\}$. Stąd $\omega_0 = \cos\frac{\pi}{6} + i\sin\frac{\pi}{6} = \frac{\sqrt{3}}{2} + \frac{1}{2}i$, $\omega_1 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}$, $\omega_2 = \cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6} = \cos\left(\pi + \frac{\pi}{6}\right) + i\sin\left(\pi + \frac{\pi}{6}\right) = -\cos\frac{\pi}{6} - i\sin\frac{\pi}{6} = -\frac{\sqrt{3}}{2} - \frac{1}{2}i$ oraz $\omega_3 = \cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3} = \cos\left(2\pi - \frac{\pi}{3}\right) + i\sin\left(2\pi - \frac{\pi}{3}\right) = \cos\frac{\pi}{3} - i\sin\frac{\pi}{3} = \frac{1}{2} - \frac{\sqrt{3}}{2}i$.

Definicja 2.9. Niech $n \in \mathbb{N}$. Dla każdego $k \in \{0, 1, \ldots, n-1\}$, k-ty pierwiastek zespolony stopnia n z 1 oznaczamy symbolem ε_k, tj. $\varepsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}$.

Definicja 2.10. Niech $n \in \mathbb{N}$. Liczbę zespoloną ω nazywamy pierwiastkiem pierwotnym stopnia n z jedności wówczas, gdy $\omega^n = 1$ oraz $\omega^m \neq 1$ dla każdej liczby naturalnej $m < n$.

Twierdzenie 2.2. Niech n będzie liczbą naturalną. Liczba zespolona ω jest pierwiastkiem pierwotnym stopnia n z jedności wtedy i tylko wtedy, gdy $\omega = \varepsilon_k$ dla pewnego $k \in \{0, 1, \ldots, n-1\}$ takiego, że $\mathrm{NWD}(k, n) = 1$.

Dowód. Załóżmy, że ω jest pierwiastkiem pierwotnym stopnia n z jedności. Wtedy $\omega^n = 1$ i $\omega^m \neq 1$ dla każdej liczby naturalnej $m < n$. Na mocy Twierdzenia 2.1 oraz Definicji 2.9 otrzymujemy, że $\omega = \varepsilon_k$ dla pewnego $k \in \{0, 1, \ldots, n-1\}$. Załóżmy nie wprost, że $\mathrm{NWD}(k, n) > 1$. Istnieje wówczas liczba naturalna $s > 1$ taka,

że $k = sl$ oraz $n = sr$ dla pewnych $l \in \mathbb{Z}$ oraz $r \in \mathbb{N}$. W szczególności wynika stąd, że $r < n$. Ponadto, na mocy wzoru de Moivre'a, $\omega^r = {\varepsilon_k}^r = \cos\frac{2kr\pi}{n} + i\sin\frac{2kr\pi}{n} = \cos\frac{2lsr\pi}{sr} + i\sin\frac{2lsr\pi}{sr} = \cos(2l\pi) + i\sin(2l\pi) = \cos(2\pi) + i\sin(2\pi) = 1$, sprzeczność. Zatem $\mathrm{NWD}(k,n) = 1$.

Na odwrót. Przypuśćmy, że $\omega = \varepsilon_k$ dla pewnego $k \in \{0,1,\ldots,n-1\}$ takiego, że $\mathrm{NWD}(k,n) = 1$. Jasne jest, że $\omega^n = 1$. Jeżeli $n = 1$, to $k = 0$ i $\omega = \varepsilon_0 = 1$ jest pierwiastkiem pierwotnym stopnia pierwszego z jedności. Niech dalej $n > 1$. Wówczas $k \neq 0$. Załóżmy nie wprost, że istnieje liczba naturalna $m < n$ taka, że $\omega^m = 1$. Stąd oraz na mocy wzoru de Moivre'a otrzymujemy, że $1 = \cos\frac{2km\pi}{n} + i\sin\frac{2km\pi}{n}$. Zatem $\frac{2km\pi}{n} = 2t\pi$ dla pewnego $t \in \mathbb{Z}$. Wobec tego $km = nt$, skąd $k \mid nt$. Ale $\mathrm{NWD}(k,n) = 1$, więc $k \mid t$. Zatem $t = kh$ dla pewnego $h \in \mathbb{Z}$ i w konsekwencji $km = khn$. Ponieważ $k \neq 0$, to ostatnia równość równoważna jest równości $m = hn$. Stąd $n \mid m$, co jest niemożliwe gdyż $m < n$, sprzeczność. Wobec tego $\omega^m \neq 1$ dla każdej liczby naturalnej $m < n$. Zatem ω jest pierwiastkiem pierwotnym stopnia n z jedności.

Zespolone pierwiastki kwadratowe z liczby zespolonej z można wyznaczać także w oparciu o następujące

Twierdzenie 2.3. Niech a oraz b będą dowolnymi liczbami rzeczywistymi i niech

$$\omega = \begin{cases} \sqrt{a} & \text{, gdy } a \geq 0 \text{ oraz } b = 0 \\ \sqrt{-a}\cdot i & \text{, gdy } a < 0 \text{ oraz } b = 0 \\ \sqrt{\frac{\sqrt{a^2+b^2}+a}{2}} + \mathrm{sgn}(b)\cdot\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}}\cdot i & \text{, gdy } b \neq 0. \end{cases}$$

Wówczas $\{-\omega,\omega\} = \{z \in \mathbb{C}\colon z^2 = a+bi\}$.

Dowód. Jeżeli $a \geq 0$ oraz $b = 0$, to $\omega^2 = \left(\sqrt{a}\right)^2 = a = a + 0\cdot i = a + bi$. Jeśli $a < 0$ i $b = 0$, to $\omega^2 = \left(\sqrt{-a}\cdot i\right)^2 = -a\cdot(-1) = a + 0\cdot i = a + bi$. Załóżmy teraz, że $b \neq 0$. Oznaczając $x = \sqrt{\frac{\sqrt{a^2+b^2}+a}{2}}$ oraz $y = \mathrm{sgn}(b)\cdot\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}}$, otrzymujemy $\omega = x + yi$. Ponadto $\omega^2 = \left(x^2 - y^2\right) + 2xyi$ i bezpośrednie sprawdzenie pokazuje, że $x^2 - y^2 = a$ oraz $2xy = b$. Zatem również w tym przypadku $\omega^2 = a + bi$. Ponieważ $(-\omega)^2 = (-1\cdot\omega)^2 = (-1)^2\omega^2 = \omega^2$, to $\{-\omega,\omega\} \subseteq \{z \in \mathbb{C}\colon z^2 = a+bi\}$. Aby wykazać inkluzję przeciwną zauważmy, że jeśli $z^2 = a + bi$, to $z^2 = \omega^2$, więc $0 = z^2 - \omega^2 = (z-\omega)\cdot(z+\omega)$ i w konsekwencji $z = \omega$ lub $z = -\omega$. Wobec tego $\{z \in \mathbb{C}\colon z^2 = a+bi\} \subseteq \{-\omega,\omega\}$ i ostatecznie $\{-\omega,\omega\} = \{z \in \mathbb{C}\colon z^2 = a+bi\}$.

Ze szkoły średniej wiadomo, że równanie kwadratowe $ax^2 + bx + c = 0$ o współczynnikach rzeczywistych posiada rozwiązanie w liczbach rzeczywistych wtedy i tylko wtedy, gdy $\Delta \geq 0$, przy czym $\Delta = b^2 - 4ac$. Takie ograniczenie nie obowiązuje w przypadku równań kwadratowych rozwiązywanych nad ciałem $\mathbb{C}$. Mamy bowiem następujące

Stwierdzenie 2.4. Dla dowolnych $a \in \mathbb{C} \setminus \{0\}$ oraz $b, c \in \mathbb{C}$ równanie:

$$az^2 + bz + c = 0 \tag{2.3.1}$$

posiada pierwiastek zespolony. Ponadto $\left\{\frac{-b-\omega}{2a}, \frac{-b+\omega}{2a}\right\}$, gdzie ω jest dowolnym zespolonym pierwiastkiem kwadratowym z liczby Δ określonej równością $\Delta = b^2 - 4ac$, jest zbiorem wszystkich rozwiązań tego równania.

Dowód. Ponieważ $a \neq 0$, to równanie (2.3.1) możemy przekształcić do równoważnej postaci $a\left(z^2 + \frac{b}{a}z + \frac{c}{a}\right) = 0$. Zatem równanie (2.3.1) równoważne jest równaniu $z^2 + \frac{b}{a}z + \frac{c}{a} = 0$. Ponadto $z^2 + \frac{b}{a}z + \frac{c}{a} = \left(z + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{4ac}{4a} = \left(z + \frac{b}{2a}\right)^2 - \frac{b^2-4ac}{4a^2} = \left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}$ oraz z Twierdzenia 2.1 wynika istnienie takiego $\omega \in \mathbb{C}$, że $\omega^2 = \Delta$, więc $\left(z + \frac{b}{2a}\right)^2 - \frac{\omega^2}{4a^2} = \left(z + \frac{b}{2a} - \frac{\omega}{2a}\right) \cdot \left(z + \frac{b}{2a} + \frac{\omega}{2a}\right)$ i w konsekwencji równanie (2.3.1) równoważne jest równaniu $\left(z - \frac{-b+\omega}{2a}\right) \cdot \left(z - \frac{-b-\omega}{2a}\right) = 0$, skąd natychmiast otrzymujemy tezę.

Przykład 2.8. W ciele $\mathbb{C}$ rozwiążemy równanie $z^2 + (-8+i)z + 17 - i = 0$. Ponieważ $\Delta = (-8+i)^2 - 4 \cdot 1 \cdot (17-i) = 64 - 1 - 16i - 68 + 4i - 5 - 12i$, to z Twierdzenia 2.3 wynika, że $\Delta = \left(\pm(2-3i)\right)^2$. Zatem pierwiastkami rozważanego równania są $z_1 = \frac{8-i-2+3i}{2} = \frac{6+2i}{2} = 3+i$ oraz $z_2 = \frac{8-i+2-3i}{2} = \frac{10-4i}{2} = 5-2i$.

Uwaga 2.11. W świetle Twierdzenia 2.1, symbol pierwiastka powinien być używany wyjątkowo ostrożnie w kontekście liczb zespolonych. Ponieważ istnieją dokładnie dwa pierwiastki kwadratowe z niezerowej liczby zespolonej, to bez obaw możemy używać go np. podczas rozwiązywania równań kwadratowych. Rozwiązując powyższy przykład moglibyśmy napisać $\sqrt{\Delta} = \pm(2-3i)$ zamiast $\Delta = \left(\pm(2-3i)\right)^2$. Niektórzy matematycy symbolem $\sqrt[n]{z}$ oznaczają zbiór wszystkich pierwiastków n-tego stopnia z liczby zespolonej z. Przy takiej notacji $\sqrt[4]{1} = \{-1, 1, -i, i\}$ oraz ma sens napis $i \in \sqrt[4]{1}$.

Twierdzenie 2.4. Dla dowolnych $n \in \mathbb{N}$ i $\varphi \in \mathbb{R}$, wszystkie rozwiązania równania:

$$n \cdot \arg(z) = \varphi \tag{2.3.2}$$

w liczbach zespolonych są postaci $z = z_{r,k}$, gdzie:

$$z_{r,k} = r \cdot \left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right), \tag{2.3.3}$$

przy czym $r \in (0, \infty)$ oraz $k \in \{0, 1, \ldots, n-1\}$. W szczególności, $r = |z_{r,k}|$. Ponadto $z_{r,rk} = z_{s,sl}$ wtedy i tylko wtedy, gdy $s = r$ i $l = k$ dla wszystkich $r, s \in (0, \infty)$ oraz $k, l \in \{0, 1, \ldots, n-1\}$.

Dowód. Dla dowolnych $r \in (0,\infty)$ i $k \in \{0,1,\ldots,n-1\}$ otrzymujemy, że $n \cdot \arg(z_{r,k}) = n \cdot \frac{\varphi+2k\pi}{n} = \varphi + 2k\pi$. Stąd oraz na mocy Uwagi 2.8, liczba zespolona $z_{r,k}$ spełnia równanie (2.3.2). Zatem każda spośród liczb zespolonych opisanych wzorem (2.3.3) jest rozwiązaniem równania (2.3.2).

Załóżmy, że z_0 jest rozwiązaniem równania (2.3.2). Pokażemy, że istnieją $r \in (0,\infty)$ i $k \in \{0,1,\ldots,n-1\}$ takie, że $z_0 = z_{r,k}$. Ponieważ dziedziną równania (2.3.2) jest $\mathbb{C} \setminus \{0\}$, to $z_0 \neq 0$. Z Uwagi 2.5 wynika więc istnienie takich $r_0 \in (0,\infty)$ oraz $\varphi_0 \in [0,2\pi]$, że $r_0 = |z_0|$, $\varphi_0 = \arg(z_0)$ i $z_0 = r_0 \cdot (\cos\varphi_0 + i\sin\varphi_0)$. Stąd oraz na mocy wzoru de Moivre'a, ${z_0}^n = {r_0}^n \cdot (\cos(n\varphi_0) + i\sin(n\varphi_0))$. Ponadto $n\varphi_0 = \varphi$, bo z_0 spełnia równanie (2.3.2). Zatem ${z_0}^n = {r_0}^n \cdot (\cos\varphi + i\sin\varphi)$, skąd wynika, że z_0 jest zespolonym pierwiastkiem n-tego stopnia z liczby ${r_0}^n \cdot (\cos\varphi + i\sin\varphi)$. Twierdzenie 2.1 implikuje więc istnienie takiego $k_0 \in \{0,1,\ldots,n-1\}$, że $z_0 = r_0 \cdot \left(\cos\frac{\varphi+2k_0\pi}{n} + i\sin\frac{\varphi+2k_0\pi}{n}\right)$. Zatem $z_0 = z_{r,k}$ dla $k = k_0$ i $r = r_0$. W ten sposób wykazaliśmy, że każde rozwiązanie równania (2.3.2) jest postaci (2.3.3).

Z przeprowadzonego wyżej rozumowania wynika w szczególności, że $r = |z_{r,k}|$ (co można zweryfikować również za pomocą bezpośredniego rachunku). Weźmy dowolne $k,l \in \{0,1,\ldots,n-1\}$ i $r,s \in (0,\infty)$. Jeżeli $l = k$ i $s = r$, to oczywiście $z_{r,k} = z_{s,l}$. Przypuśćmy teraz, że $z_{r,k} = z_{s,l}$. Wówczas $r = |z_{r,k}| = |z_{s,l}| = s$ oraz $\arg(z_{r,k}) = \arg(z_{s,l})$. Uwaga 2.7 implikuje więc istnienie takiego $h \in \mathbb{Z}$, że $\frac{\varphi+2k\pi}{n} = \frac{\varphi+2l\pi}{n} + 2h\pi$. Zatem $k - l = hn$. Ponadto $|k-l| < n$, więc $h = 0$ i ostatecznie $k = l$.

Przykład 2.9. W oparciu o Twierdzenie 2.4 oraz własności modułu, sprzężenia i argumentu liczby zespolonej omówione w Stwierdzeniach 2.2 i 2.3 oraz Uwadze 2.8 rozwiążemy w ciele $\mathbb{C}$ równanie:

$$z^3 = (\overline{z})^6. \tag{2.3.4}$$

Łatwo zauważyć, że $z = 0$ jest rozwiązaniem powyższego równania. Znajdziemy teraz wszystkie pozostałe rozwiązania tego równania. W tym celu rozważmy układ warunków:

$$\begin{cases} z \neq 0 \\ z^3 = (\overline{z})^6 \end{cases} \Leftrightarrow \begin{cases} z \neq 0 \\ \left|z^3\right| = \left|(\overline{z})^6\right| \\ \arg\left(z^3\right) = \arg\left((\overline{z})^6\right) \end{cases} \Leftrightarrow \begin{cases} z \neq 0 \\ |z|^3 = |\overline{z}|^6 \\ 3\arg(z) = 6\arg(\overline{z}) \end{cases} \Leftrightarrow$$

$$\begin{cases} z \neq 0 \\ |z|^3 = |z|^6 \\ 3\arg(z) = -6\arg(z) \end{cases} \Leftrightarrow \begin{cases} 1 = |z|^3 \\ 9\arg(z) = 0 \end{cases} \Leftrightarrow \begin{cases} |z| = 1 \\ z = z_{k,r} \\ z_{k,r} = r \cdot \left(\cos\frac{2k\pi}{9} + i\sin\frac{2k\pi}{9}\right) \\ r > 0 \\ k \in \{0,1,\ldots,8\} \end{cases} \Leftrightarrow$$

$$\begin{cases} z = z_{k,1} \\ z_{k,1} = \cos\frac{2k\pi}{9} + i\sin\frac{2k\pi}{9} \\ k \in \{0,1,\ldots,8\} \end{cases} \Leftrightarrow z \in \left\{ \cos\frac{2k\pi}{9} + i\sin\frac{2k\pi}{9} : k \in \{0,1,\ldots,8\} \right\}.$$

Wobec tego równanie $(2.3.4)$ posiada dokładnie 10 rozwiązań i wszystkimi jego rozwiązaniami są: 0, 1, $\cos\frac{2\pi}{9} + i\sin\frac{2\pi}{9}$, $\cos\frac{4\pi}{9} + i\sin\frac{4\pi}{9}$, $-i$, $\cos\frac{8\pi}{9} + i\sin\frac{8\pi}{9}$, $\cos\frac{10\pi}{9} + i\sin\frac{10\pi}{9}$, $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$, $\cos\frac{14\pi}{9} + i\sin\frac{14\pi}{9}$ oraz $\cos\frac{16\pi}{9} + i\sin\frac{16\pi}{9}$.

2.4 Interpretacja geometryczna dodawania i mnożenia liczb zespolonych

Uwaga 2.12. Dowolną liczbę zespoloną $z = x + iy$ możemy traktować jak punkt (x,y) płaszczyzny lub wektor $[x,y]$ zaczepiony w początku układu współrzędnych, tj. wektor o początku w punkcie $(0,0)$ i końcu w punkcie (x,y). Przy drugiej spośród tych interpretacji, dodawanie liczb zespolonych można geometrycznie utożsamić z dodawaniem odpowiadających im wektorów zaczepionych w początku układu współrzędnych na płaszczyźnie. Ze wzoru $(2.2.1)$ oraz punktu (i) Stwierdzenia 2.3 wynika, że aby pomnożyć dwie niezerowe liczby zespolone należy pomnożyć ich moduły oraz dodać ich argumenty. W szczególności, jeśli $z_0 \in \mathbb{C} \setminus \{0\}$, to funkcja $f\colon \mathbb{C} \to \mathbb{C}$ dana wzorem $f(z) = z_0 \cdot z$ dla każdego $z \in \mathbb{C}$, jest złożeniem obrotu o kąt miary $\mathrm{Arg}(z_0)$ z jednokładnością o środku w początku układu współrzędnych i skali $|z_0|$.

Przykład 2.10. Wiedząc, że $z_1 = 3 + 2i$ oraz $z_3 = -1 - 4i$ są przeciwległymi wierzchołkami kwadratu na płaszczyźnie zespolonej, wyznaczymy pozostałe wierzchołki tego kwadratu. Zauważmy, że możemy to uczynić „przesuwając" rozważany kwadrat w taki sposób, aby jego środek znalazł się w początku kartezjańskiego układu współrzędnych, obracając punkty z_1 oraz z_4 o kąt $\frac{\pi}{2}$, a następnie „przesuwając" kwadrat na jego pierwotne „miejsce". Niech z_0 będzie środkiem rozważanego kwadratu, zaś z_2 oraz z_4 jego wierzchołkami znajdującymi się odpowiednio w drugiej i czwartej ćwiartce kartezjańskiego układu współrzędnych. Wówczas wspomniane „przesuwanie" kwadratu po płaszczyźnie zespolonej opisuje funkcja $f\colon \mathbb{C} \to \mathbb{C}$ określona wzorem $f(z) = z - z_0$, którą geometrycznie (tzn. w kartezjańskim układzie współrzędnych) interpretujemy jako translację o wektor $[\mathrm{re}(z_0), \mathrm{im}(z_0)]$. Ponieważ $z_0 = \frac{1}{2} \cdot (z_1 + z_3) = \frac{1}{2} \cdot (2 - 2i) = 1 - i$, to $f(z) = z - 2 + 2i$ oraz $f^{-1}(z) = z + 2 - 2i$ dla każdego $z \in \mathbb{C}$. Ponadto z Uwagi 2.12 wynika, że obrót o kąt $\frac{\pi}{2}$ opisany jest funkcją $g\colon \mathbb{C} \to \mathbb{C}$ daną wzorem $g(z) = iz$. Niech $F = f^{-1} \circ g \circ f$. Wówczas $F(z) = i(z-4)$ dla każdego $z \in \mathbb{C}$, $z_2 = F(z_1)$ oraz $z_4 = F(z_3)$. Zatem $z_2 = i(-1+2i) = -2 - i$ oraz $z_3 = i(-5-4i) = 4 - 5i$.

Rozdział 3

Pierścień wielomianów

3.1 Określenie wielomianu

Definicja 3.1. Dla dowolnego niezerowego pierścienia R, symbolem $R[x]$ oznaczamy zbiór wszystkich nieskończonych ciągów $f = (f_0, f_1, f_2, \ldots)$ takich, że $f_i \in R$ dla każdego $i \in \mathbb{N}_0$ oraz $0 = f_j = f_{j+1} = f_{j+2} = \ldots$ dla pewnego $j \in \mathbb{N}_0$. Elementy zbioru $R[x]$ nazywamy wielomianami zmiennej x o współczynnikach z pierścienia R.

Uwaga 3.1. Przyjmujemy umowę, że jeśli wielomian oznaczony jest literą g, to $g = (g_0, g_1, \ldots)$, tzn. $g_0, g_1, \ldots$ są kolejnymi współczynnikami wielomianu g.

Uwaga 3.2. Dla dowolnych $f, g \in R[x]$ mamy:

$$f = g \Leftrightarrow f_i = g_i \text{ dla każdego } i \in \mathbb{N}_0. \tag{3.1.1}$$

Definicja 3.2. Wielomian $(0,0,\ldots)$ nazywamy zerowym, zaś wielomian $(1,0,0,\ldots)$ nazywamy jedynkowym. Wielomian zerowy oznaczamy krótko przez 0, natomiast wielomian jedynkowy przez 1. Wielomian f taki, że $0 = f_1 = f_2 = \ldots$ nazywamy stałym.

Definicja 3.3. Wyrazem wolnym wielomianu f nazywamy współczynnik f_0.

Definicja 3.4. Stopniem niezerowego wielomianu $f \in R[x]$ nazywamy największą nieujemną liczbę całkowitą n taką, że $f_n \neq 0$. Współczynnik f_n nazywamy najstarszym współczynnikiem wielomianu f. Dla $f = 0$ przyjmujemy, że stopień f wynosi $-\infty$. Stopień wielomianu f oznaczamy symbolem $\mathrm{st}(f)$.

Uwaga 3.3. Dla dowolnego $f \in R[x]$:

$$\mathrm{st}(f) = \begin{cases} \max\{n \in \mathbb{N}_0 \colon f_n \neq 0\} & \text{, gdy } f \neq 0 \\ -\infty & \text{, gdy } f = 0 \end{cases}.$$

Uwaga 3.4. Zakładamy, że $-\infty < m$ oraz $(-\infty) + m = (-\infty) + (-\infty) = -\infty$ dla każdego $m \in \mathbb{N}_0$.

Twierdzenie 3.1. Niech R będzie niezerowym pierścieniem. System algebraiczny $\big(R[x], \oplus, \odot, 0, 1\big)$ gdzie:

$$f \oplus g = (f_0 + g_0, f_1 + g_1, f_2 + g_2, \ldots) \tag{3.1.2}$$

oraz

$$f \odot g = (f_0 g_0, f_0 g_1 + f_1 g_0, f_0 g_2 + f_1 g_1 + f_2 g_0, \ldots) \tag{3.1.3}$$

jest pierścieniem.

Dowód. Rozważmy dowolne $f, g, h \in R[x]$. Istnieją wówczas $u, v \in \mathbb{N}_0$ takie, że $f_i = 0$ dla każdego $i > u$ oraz $g_j = 0$ dla każdego $j > v$. Stąd $f_i + g_i = 0$ dla każdego $i > \max\{u, v\}$. Zatem działanie $\oplus$ jest poprawnie zdefiniowane (tj. $f \oplus g \in R[x]$). Wprost z określenia tego działania wynika, że jest ono łączne i przemienne oraz wielomian zerowy 0 jest jego elementem neutralnym. Ponadto wielomian $F = (-f_0, -f_1, -f_2, \ldots)$ jest elementem przeciwnym do f względem działania $\oplus$. Zatem system algebraiczny $\big(R[x], \oplus, 0\big)$ jest grupą abelową. Zauważmy, że wprost ze wzoru (3.1.3) otrzymujemy równości:

$$(f \odot g)_n = \sum_{i=1}^{n} f_i g_{n-i} = \sum_{i+j=n} f_i g_j \tag{3.1.4}$$

zachodzące dla każdego $n \in \mathbb{N}_0$. Aby uzasadnić poprawność określenia mnożenia $\odot$ weźmy dowolne $n \in \mathbb{N}_0$ spełniające nierówność $n > u + v$ oraz dowolne $i, j \in \mathbb{N}_0$ takie, że $i + j = n$. Jeżeli $i > u$, to $f_i = 0$, skąd $f_i g_j = 0$; jeśli zaś $i \leq u$, to $j = n - i \geq n - u > u + v - u = v$, więc $g_j = 0$. Wobec tego $f_i g_j = 0$. Stąd dla $n > u + v$ uzyskujemy, że $(f \odot g)_n = \sum_{i+j=n} f_i g_j = 0$. Zatem $f \odot g \in R[x]$. Przemienność mnożenia $\odot$ oraz równość $1 \odot f = f$ wynikają wprost ze wzoru (3.1.4). Dalej,

$$(f \odot (g \oplus h))_n = \sum_{i+j=n} f_i (g \oplus h)_j = \sum_{i+j=n} f_i (g_j + h_j) = \sum_{i+j=n} (f_i g_j + f_i h_j) =$$

$$\sum_{i+j=n} f_i g_j + \sum_{i+j=n} f_i h_j = (f \odot g)_n + (f \odot h)_n,$$

więc $f \odot (g \oplus h) = f \odot g \oplus f \odot h$. Wobec tego mnożenie $\odot$ jest rozdzielne względem dodawania $\oplus$. Ponieważ:

$$((f \odot g) \odot h)_n = \sum_{i+j=n} (f \odot g)_i h_j =$$

$$\sum_{i+j=n} \sum_{s+t=i} (f_s g_t) h_j = \sum_{s+t+j=n} (f_s g_t) h_j = \sum_{s+t+j=n} f_s (g_t h_j)$$

oraz

$$(f \odot (g \odot h))_n = \sum_{s+k=n} f_s (g \odot h)_k = \sum_{s+k=n} \sum_{t+j=k} f_s (g_t h_j) = \sum_{s+t+j=n} f_s (g_t h_j),$$

to $f \odot (g \odot h) = (f \odot g) \odot h$, co oznacza, że mnożenie wielomianów jest łączne.

Definicja 3.5. Dla dowolnego niezerowego pierścienia R, pierścień $\big(R[x],\oplus,\odot,0,1\big)$ nazywamy pierścieniem wielomianów zmiennej x o współczynnikach z pierścienia R.

Uwaga 3.5. Jak zwykle, w celu uproszczenia zapisu, sumę oraz iloczyn wielomianów f i g będziemy zapisywali odpowiednio $f+g$ oraz $f\cdot g$, przy czym kropkę będziemy często pomijać.

Wniosek 3.1. Dla dowolnych wielomianów $f,g\in R[x]$:

(i) $\mathrm{st}(f+g)\leqslant \max\big\{\mathrm{st}(f),\mathrm{st}(g)\big\}$;
(ii) jeżeli $st(f)<st(g)$, to $st(f+g)=st(g)$;
(iii) jeżeli $f_i=0$ dla każdego $i\in\mathbb{N}$, to $(f_0,0,0,\ldots)\cdot(g_0,g_1,\ldots)=(f_0g_0,f_0g_1,\ldots)$.
(iv) jeżeli $f_1=1$ oraz $f_i=0$ dla każdego $i\in\mathbb{N}_0\setminus\{1\}$, to $(0,1,0,0,\ldots)\cdot(g_0,g_1,\ldots)=(0,g_0,g_1,\ldots)$.
(v) $\mathrm{st}(f\cdot g)\leq \mathrm{st}(f)+\mathrm{st}(g)$.

Dowód. Uzasadnienie punktów (i) oraz (ii) wynika wprost z określenia dodawania wielomianów. Punkty (iii) oraz (iv) są bezpośrednią konsekwencją definicji mnożenia wielomianów. Aby uzasadnić punkt (v) zauważmy, że jeśli $f\neq 0$, $g\neq 0$ oraz $n=\mathrm{st}(f)$ i $m=\mathrm{st}(g)$, to z dowodu Twierdzenia 3.1 wynika, że $(f\cdot g)_i=0$ dla każdego $i>n+m$. Stąd oraz na mocy wzoru (3.1.4) uzyskujemy żądaną nierówność.

Definicja 3.6. Element a pierścienia R nazywamy regularnym, wówczas gdy dla każdego $b\in R$ warunek $a\cdot b=0$ implikuje równość $b=0$.

Stwierdzenie 3.1. Niech R będzie pierścieniem i niech $f,g\in R[x]\setminus\{0\}$ będą takie, że najstarszy współczynnik wielomianu f lub najstarszy współczynnik wielomianu g jest elementem regularnym w R. Wtedy $\mathrm{st}(f\cdot g)=\mathrm{st}(f)+\mathrm{st}(g)$ oraz najstarszy współczynnik wielomianu $f\cdot g$ jest iloczynem najstarszych współczynników wielomianów f i g.

Dowód. Niech $n=st(f)$, $m=st(g)$ i niech $i,j\in\mathbb{N}_0$ będą takie, że $i+j=n+m$. Jeżeli $i<n$, to $j=n+m-i>m$, więc $g_j=0$ oraz $f_ig_j=0$. Jeśli natomiast $i>n$, to $f_i=0$, skąd $f_ig_j=0$. Ze wzoru (3.1.4) wynika więc, że $(f\cdot g)_{n+m}=f_ng_m\neq 0$. Stąd oraz na mocy punktu (v) Wniosku 3.1 otrzymujemy tezę.

Uwaga 3.6. Niech R będzie niezerowym pierścieniem. Wprost z określenia działań w pierścieniu wielomianów $R[x]$ wynika, że elementy pierścienia R możemy traktować jak wielomiany stałe, tzn. możemy dokonać utożsamienia $a\equiv(a,0,0,\ldots)$ dla dowolnego $a\in R$. Przy takim utożsamieniu mamy inkluzję $R\subseteq R[x]$.

Uwaga 3.7. Wprowadzając oznaczenie $x=(0,1,0,0,\ldots)$ i powołując się na punkt (iv) Wniosku 3.1 przez prostą indukcję otrzymujemy, że $x^n=(0,0,\ldots,0,\overset{n}{1},0,\ldots)$ dla każdego $n\in\mathbb{N}_0$. Jeśli więc $f\in R[x]$ i $\mathrm{st}(f)=n>0$, to $(f_k,0,0,\ldots)\cdot x^k=(0,\ldots,0,\overset{k}{f_k}$

$,0,\ldots)$ dla każdego $k \in \{1,2,\ldots,n\}$, skąd $f = (f_0,0,0,\ldots) + (0,f_1,0,\ldots) + \ldots + (0,0,\ldots,f_n,0,\ldots) \equiv f_0 + f_1x + f_2x^2 + \ldots + f_nx^n$. Ponadto, jeżeli $\mathrm{st}(f) = 0$, to $f \equiv f_0$. Zatem dla dowolnego $f \in R[x] \setminus \{0\}$ mamy utożsamienie:

$$f \equiv f_0 + f_1x + f_2x^2 + \ldots + f_nx^n.$$

Zastępując znak utożsamienia „$\equiv$" znakiem równości, uzyskujemy naturalną notację dla wielomianów, której używamy także dla wielomianu zerowego. Przy jej zastosowaniu otrzymujemy, że wielomiany $f,g \in R[x]$ są równe wtedy i tylko wtedy, gdy $\mathrm{st}(f) = \mathrm{st}(g)$ oraz $f_i = g_i$ dla każdego $i \in \{0,1,\ldots,\mathrm{st}(f)\}$.

Definicja 3.7. Wartością wielomianu $f = f_0 + f_1x + \ldots + f_nx^n \in R[x]$ w punkcie $a \in R$ nazywamy element $f(a)$ pierścienia R określony następująco $f(a) = f_0 + f_1a + \ldots + f_na^n$.

Uwaga 3.8. Wykorzystując wiadomości z zakresu teorii pierścieni wykraczające poza zakres tematyczny pierwszego semestru studiów, można udowodnić, że dla dowolnych $a \in R$ i $w_1,w_2,\ldots,w_n \in R[x]$ prawdziwe są równości:

(i) $(w_1 \cdot w_2 \cdot \ldots \cdot w_n)(a) = w_1(a) \cdot w_2(a) \cdot \ldots \cdot w_n(a)$;
(ii) $(w_1 + w_2 + \ldots + w_n)(a) = w_1(a) + w_2(a) + \ldots + w_n(a)$.

Definicja 3.8. Pierwiastkiem wielomianu $f \in R[x]$ nazywamy taki element a pierścienia R, że $f(a) = 0$.

Definicja 3.9. Wielomiany $f,g \in R[x]$ nazywamy równymi funkcyjnie, wówczas gdy $f(a) = g(a)$ dla każdego $a \in R$.

Przykład 3.1. W pierścieniu $\mathbb{Z}_5[x]$ wielomiany x oraz x^5 są równe funkcyjnie, ale nie są równe.

Wniosek 3.2. Istnieją pieścienie, nad którymi wielomiany nie mogą być traktowane jak funkcje wielomianowe. W szczególności założenie o równości stopni porównywanych wielomianów przy zastosowaniu szkolnej notacji jest istotne (zob. Uwaga 3.7).

3.2 Dzielenie wielomianów

Definicja 3.10. Niech R będzie niezerowym pierścieniem. Mówimy, że w pierścieniu $R[x]$ wykonalne jest dzielenie z resztą przez wielomian $f \in R[x]$, jeśli dla każdego wielomianu $g \in R[x]$ istnieje dokładnie jedna para (q,r) wielomianów $q,r \in R[x]$ taka, że $g = q \cdot f + r$ i $st(r) < st(f)$. Wówczas dla dowolnego $g \in R[x]$ wielomiany q i r nazywamy odpowiednio niepełnym ilorazem oraz resztą z dzielenia wielomianu g przez wielomian f.

Uwaga 3.9. Z powyższej definicji wynika, że nie istnieje pierścień R taki, że w pierścieniu $R[x]$ wykonalne jest dzielenie przez wielomian zerowy (bo $\mathrm{st}(0) = -\infty$).

Lemat 3.1. Jeżeli a jest elementem odwracalnym pierścienia R, to a jest elementem regularnym tego pierścienia.

Dowód. Z przyjętego założenia wynika istnienie takiego $c \in \mathbb{R}$, że $c \cdot a = 1$. Weźmy dowolne $b \in R$. Załóżmy, że $a \cdot b = 0$. Wtedy $b = 1 \cdot b = (c \cdot a) \cdot b = c \cdot (a \cdot b) = 0$. Zatem a jest elementem regularnym w R.

Twierdzenie 3.2. Niech R będzie niezerowym pierścieniem i niech $f \in R[x]$. W pierścieniu $R[x]$ wykonalne jest dzielenie z resztą przez wielomian f wtedy i tylko wtedy, gdy najstarszy współczynnik wielomianu f jest elementem odwracalnym w pierścieniu R.

Dowód. Niech n oraz a oznaczają odpowiednio stopień i najstarszy współczynnik wielomianu f.

Załóżmy, że w pierścieniu $R[x]$ wykonalne jest dzielenie z resztą przez wielomian f. Jeśli element a nie jest regularny w R, to istnieje $b \in R$ takie, że $b \neq 0$ i $b \cdot a = 0$. Ale wtedy $st(bf) < n$, $bf = b \cdot f + 0 = 0 \cdot f + bf$ i $(b,0) \neq (0,bf)$ oraz $st(0) < n$, sprzeczność. Wobec tego a jest elementem regularnym w R. Na mocy przyjętego założenia, istnieją $q, r \in R[x]$ takie, że $x^n = q \cdot f + r$ i $st(r) < n$. W szczególności wynika stąd, że $q \neq 0$. Powołując się więc na Stwierdzenie 3.1 uzyskujemy, że $\mathrm{st}(q \cdot f) = \mathrm{st}(q) + \mathrm{st}(f) = \mathrm{st}(q) + n > \mathrm{st}(r)$. Stąd oraz na mocy punktu (ii) Wniosku 3.1, $n = \mathrm{st}(x^n) = \mathrm{st}(q \cdot f + r) = \mathrm{st}(q \cdot f) = \mathrm{st}(q) + n$ i w konsekwencji $\mathrm{st}(q) = 0$. Zatem $q \in R \setminus \{0\}$. Powołując się ponownie na Stwierdzenie 3.1 uzyskujemy, że najstarszym współczynnikiem wielomianu $q \cdot f$ jest qa. Ponieważ najstarszym współczynnikiem wielomianu x^n jest 1, $x^n = q \cdot f + r$ i $st(r) < n$, to $qa = 1$. Zatem $a \in R^*$.

Na odwrót. Przypuśćmy, że $a \in R^*$. Wtedy $ab = 1$ dla pewnego $b \in R$. Załóżmy nie wprost, że pewien wielomian z $R[x]$ nie jest podzielny z resztą przez wielomian f. Istnieje wówczas wielomian $g \in R[x]$ najniższego stopnia m niepodzielny z resztą przez f. Jeśli $m < n$, to $g = 0 \cdot f + g$, skąd g jest podzielny z resztą przez f w $R[x]$, sprzeczność. Zatem $m \geq n$. Niech c oznacza najstarszy współczynnik wielomianu g. Ponadto definiujemy wielomian $h = g - cbx^{m-n}f$. Na mocy Lematu 3.1 i Stwierdzenia 3.1 otrzymujemy, że $\mathrm{st}(cbx^{m-n}f) = m - n + n = m$ oraz $cba = c \cdot 1 = c$ jest najstarszym współczynnikiem wielomianu $cbx^{m-n}f$. Zatem $\mathrm{st}(h) < m$. Z minimalności m wynika istnienie takich $q_1, r \in R[x]$, że $h = q_1 \cdot f + r$ i $st(r) < n$. Stąd $g = (cbx^{m-n} + q_1) \cdot f + r$, co oznacza, że wielomian g jest podzielny z resztą przez wielomian f w $R[x]$, sprzeczność. Pozostało wykazać jednoznaczność reszty i niepełnego ilorazu. W tym celu weźmy dowolne $q_1, q_2, r_1, r_2 \in R[x]$ i załóżmy, że $\mathrm{st}(r_1), \mathrm{st}(r_2) < n$ oraz $q_1 \cdot f + r_1 = q_2 \cdot f + r_2$. Wówczas $(q_1 - q_2) \cdot f = r_2 - r_1$. Załóżmy nie wprost, że $q_1 - q_2 \neq 0$. Ze Stwierdzenia 3.1 wynika wówczas, że $\mathrm{st}((q_1 - q_2) \cdot f) = \mathrm{st}(q_1 - q_2) + \mathrm{st}(f) = \mathrm{st}(q_1 - q_2) + n$. Ponadto $\mathrm{st}(r_2 - r_1) < n$, więc

$\mathrm{st}(q_1 - q_2) = -\infty$, czyli $q_1 - q_2 = 0$. Otrzymana sprzeczność oznacza, że $q_1 = q_2$ i w konsekwencji $r_1 = r_2$.

Uwaga 3.10. Znany ze szkoły algorytm dzielenia wielomianów z resztą jest dobry dla dowolnego pierścienia wielomianów.

Definicja 3.11. Wielomian $f \in R[x]$ nazywamy unormowanym, gdy jego najstarszy współczynnik jest równy 1 (tzn. jest jedynką pierścienia R) .

Bezpośrednią konsekwencją Twierdzenia 3.2 jest następujący

Wniosek 3.3. W pierścieniu wielomianów $R[x]$ nad dowolnym niezerowym pierścieniem R wykonalne jest dzielenie z resztą przez wielomiany unormowane.

Twierdzenie 3.3 (Bézout). Niech a będzie elementem niezerowego pierścienia R i niech $f \in R[x]$. Wówczas reszta z dzielenia wielomianu f przez dwumian $x - a$ równa jest $f(a)$.

Dowód. Na mocy Wniosku 3.3 otrzymujemy, że $f = q \cdot (x - a) + r$ dla pewnych $q, r \in R[x]$ takich, że $\mathrm{st}(r) < 1$. Zatem $r \in R$. Stąd oraz na mocy Uwagi 3.8, $f(a) = q(a) \cdot (a - a) + r = r$. Wobec tego $r = f(a)$.

Definicja 3.12. Niech R będzie niezerowym pierścieniem i niech $f, g \in R[x]$. Mówimy, że wielomian f dzieli wielomian g w pierścieniu $R[x]$, co zapisujemy: $f \mid g$, gdy istnieje wielomian $h \in R[x]$ taki, że $g = f \cdot h$.

Ważną konsekwencją Twierdzenia Bézouta jest następujący

Wniosek 3.4. Niech a będzie elementem niezerowego pierścienia R i niech $f \in R[x]$. Wówczas $x - a \mid f$ wtedy i tylko wtedy, gdy $f(a) = 0$.

Uwaga 3.11. Znany ze szkoły średniej algorytm dzielenia wielomianów można stosować dla wielomianów nad dowolnym niezerowym pierścieniem, o ile w rozważanym pierścieniu wielomianów wykonalne jest dzielenie z resztą przez podany wielomian.

3.3 Zasadnicze twierdzenie algebry

Twierdzenie 3.4 (Zasadnicze twierdzenie algebry). Każdy wielomian dodatniego stopnia o współczynnikach zespolonych posiada pierwiastek zespolony.

Uwaga 3.12. Kompletny dowód powyższego twierdzenia jest bardzo długi i zdecydowanie wykracza poza zakres wiedzy, którą można było zdobyć na obecnym etapie studiów. Dlatego został on pominięty. Warty podkreślenia jest fakt, że wszystkie

znane obecnie dowody Zasadniczego twierdzenia algebry wykorzystują aparat analizy matematycznej; również zespolonej. Uwaga ta dotyczy także dowodów nazywanych algebraicznymi. Takie zjawisko nie jest częste w algebrze, która jest jedną z najbardziej autonomicznych dziedzin matematyki. Dowody zasadniczego twierdzenia algebry dostępne są np. w Andruszkiewicz (2005b) oraz (Leja, 1979, s. 105).

Współczesna algebra jest dziedziną tak dalece rozwiniętą, że dziś już nie sposób wskazać twierdzenie, które można by nazwać zasadniczym. Niemniej, ze względu na tradycję oraz doniosłość Twierdzenia 3.4 w historii matematyki, wciąż nazywa się je Zasadniczym twierdzeniem algebry.

Rozdział 4

Pojęcie macierzy. Wyznacznik macierzy i jego własności

4.1 Określenie macierzy oraz wyznacznika

Definicja 4.1. Niech K będzie ciałem i niech $m,n \in \mathbb{N}$. Funkcję:

$$A\colon \{1,2,\ldots,m\} \times \{1,2,\ldots,n\} \to K$$

nazywamy $m \times n$-macierzą nad ciałem K. Dla dowolnych $(i,j) \in \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}$ element $A\big((i,j)\big)$ ciała K nazywamy (i,j)-tym wyrazem macierzy A i oznaczamy go przez $[A]_{ij}$ lub a_{ij}. Przy zastosowaniu drugiej spośród wymienionych notacji, macierz A zapisuje się w postaci następującej tablicy o m wierszach (rzędy poziome) i n kolumnach (rzędy pionowe):

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}. \tag{4.1.1}$$

Stosuje się także uproszczoną notację $A = [a_{ij}]_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}}$ lub nawet $A = [a_{ij}]_{ij}$, gdy wymiary macierzy są wyraźnie określone wcześniej.

Gdy A jest $(n \times n)$-macierzą, to A nazywamy macierzą kwadratową stopnia n. Zbiór wszystkich $(m \times n)$-macierzy nad ciałem K oznaczamy symbolem $\mathrm{M}_{m\times n}(K)$. Dla zbioru wszystkich macierzy kwadratowych stopnia n nad ciałem K stosujemy oznaczenie $\mathrm{M}_n(K)$.

Definicja 4.2. Dla dowolnej macierzy $A = [a_{ij}]_{\substack{i=1,2,\ldots,m \\ j=1,2,\ldots,n}}$ określamy macierz transponowaną $A^T = [a_{ji}]_{\substack{j=1,2,\ldots,n \\ i=1,2,\ldots,m}}$ (macierz A^T powstaje z macierzy A poprzez wzajemną zamianę miejscami wierszy i kolumn).

Definicja 4.3. Wyznacznikiem macierzy kwadratowej $A = [a_{ij}]_{ij}$ stopnia n nad ciałem K nazywamy element $\det(A)$ ciała K określony wzorem:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}. \tag{4.1.2}$$

Uwaga 4.1. Wyznacznik macierzy $A = [a_{ij}]_{ij} \in \mathrm{M}_n(K)$ oznaczamy także w następujący sposób:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{4.1.3}$$

Stwierdzenie 4.1. Dla dowolnych elementów a,b,c,d ciała K prawdziwa jest równość:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Dowód. Oznaczmy $a_{11} = a$, $a_{12} = b$, $a_{21} = c$ i $a_{22} = d$. Stosując wzór (4.1.2) dla macierzy $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ otrzymujemy, że $\det(A) = \sum_{\sigma \in S_2} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)}$. Ponadto grupa S_2 zawiera dokładnie dwie permutacje: id oraz $(1,2)$. Pierwsza z nich jest parzysta, druga zaś nieparzysta. Stąd $\det(A) = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$, czyli $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Stwierdzenie 4.2 (Wzór Sarrusa). Dla dowolnych elementów $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$ ciała K prawdziwa jest równość:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} + a_{21} \cdot a_{32} \cdot a_{13} + a_{12} \cdot a_{23} \cdot a_{31} -$$

$$a_{31} \cdot a_{22} \cdot a_{13} - a_{21} \cdot a_{12} \cdot a_{33} - a_{32} \cdot a_{23} \cdot a_{11}.$$

Dowód. Wszystkimi elementami grupy S_3 są: $\sigma_0 = \mathrm{id}$, $\sigma_1 = (1,2)$, $\sigma_2 = (1,3)$, $\sigma_3 = (2,3)$, $\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ oraz $\sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. Ponadto:

$$\operatorname{sgn}(\sigma_i) = \begin{cases} 1 & \text{, dla } i \in \{0,4,5\} \\ -1 & \text{, dla } i \in \{1,2,3\} \end{cases}.$$

Stąd oraz na mocy wzoru (4.1.2) i własności działań w ciele otrzymujemy tezę.

Stwierdzenie 4.3. Dla dowolnej liczby naturalnej $n > 1$ oraz dowolnych elementów $a_{11}, a_{12}, \dots, a_{1n}, a_{22}, a_{23}, \dots, a_{2n}, a_{23}, a_{24}, \dots, a_{2n}, \dots, a_{nn}$ ciała K zachodzi równość:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{23} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = \prod_{i=1}^{n} a_{ii}.$$

Dowód. Lewą stronę powyższej równości oznaczmy przez W. Rozważmy dowolne $\sigma \in S_n \setminus \{\mathrm{id}\}$. Niech $m = \max X_\sigma$. Wówczas dla każdego $j \in \mathbb{N}$ takiego, że $m+j \in X_n$ zachodzi równość $\sigma(m+j) = m+j$. Zatem $\sigma(m) < m$ i w konsekwencji $a_{m\sigma(m)} = 0$. Stąd oraz na mocy wzoru (4.1.2) uzyskujemy równość $W = \mathrm{sgn}(\mathrm{id}) \cdot \prod_{i=1}^{n} a_{ii} = \prod_{i=1}^{n} a_{ii}$.

4.2 Własności wyznaczników

Stwierdzenie 4.4. Dla dowolnej macierzy kwadratowej stopnia n nad ciałem K zachodzi równość $\det\left(A^T\right) = \det(A)$.

Dowód. Niech $A = [a_{ij}]_{ij}$ i niech $B = [b_{ij}]_{ij} = A^T$. Wówczas $b_{ij} = a_{ji}$ dla wszystkich $i, j \in \{1,2,\ldots,n\}$. Stąd oraz na mocy wzoru (4.1.2) otrzymujemy, że:

$$\det\left(A^T\right) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^{n} b_{i\sigma(i)} = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{\sigma(i)i}. \tag{4.2.1}$$

Ponadto:

$$\left\{\left(\sigma(i), i\right) : i \in \{1,2,\ldots,n\}\right\} = \left\{\left(i, \sigma^{-1}(i)\right) : i \in \{1,2,\ldots,n\}\right\}$$

oraz $\mathrm{sgn}\left(\sigma^{-1}\right) = \mathrm{sgn}(\sigma)$ (zob. Wniosek 1.4), więc:

$$\det\left(A^T\right) = \sum_{\sigma \in S_n} \mathrm{sgn}\left(\sigma^{-1}\right) \cdot \prod_{i=1}^{n} a_{i\sigma^{-1}(i)}.$$

Ponieważ funkcja $S_n \ni \sigma \longmapsto \sigma^{-1} \in S_n$ jest bijekcją, to ostatnią równość możemy zapisać w równoważnej postaci:

$$\det\left(A^T\right) = \sum_{\sigma \in S_n} \mathrm{sgn}\left(\sigma\right) \cdot \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Zatem $\det\left(A^T\right) = \det(A)$.

Jak się wkrótce okaże, wyznacznik macierzy kwadratowej A stopnia n nad ciałem K posiada wiele własności związanych z pewnymi operacjami na jej wierszach i kolumnach. Ponadto dla wszystkich $i, j \in \{1,2,\ldots,n\}$, i-ty wiersz macierzy A jest i-tą kolumną macierzy A^T oraz j-ta kolumna macierzy A jest j-tym wierszem macierzy A^T. Stąd oraz na mocy Stwierdzenia 4.4 możemy wysnuć następujący

Wniosek 4.1. Każda własność wyznacznika sformułowana w języku wierszy pozostaje prawdziwa po przetłumaczeniu na język kolumn i na odwrót.

Bezpośrednią konsekwencją powyższego Stwierdzenia 4.4 oraz wzoru (4.2.1) jest również następujący

Wniosek 4.2. Dla dowolnej macierzy kwadratowej A stopnia n nad ciałem K zachodzi równość:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{\sigma(i)i}.$$

Stwierdzenie 4.5. Niech A będzie macierzą kwadratową stopnia n nad ciałem K i niech $\delta \in S_n$. Jeżeli B jest macierzą powstałą z macierzy A wskutek przestawienia wierszy (kolumn) macierzy A opisanego permutacją δ, to $\det(B) = \operatorname{sgn}(\delta) \cdot \det(A)$.

Dowód. W świetle Wniosku 4.1 wystarczy udowodnić stwierdzenie w wersji dotyczącej wierszy. Niech $A = [a_{ij}]_{ij}$ i niech $B = [b_{ij}]_{ij}$. Wówczas $b_{ij} = a_{\delta(i)j}$ dla wszystkich $i, j \in \{1, 2, \ldots, n\}$. Zatem:

$$\det(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} b_{i\sigma(i)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{\delta(i)\sigma(i)}.$$

Ponadto:

$$\Big\{ \big(\delta(i), \sigma(i)\big) : i \in \{1, 2, \ldots, n\} \Big\} = \Big\{ \big(i, (\delta^{-1} \circ \sigma)(i)\big) : i \in \{1, 2, \ldots, n\} \Big\}$$

oraz z Twierdzenia 1.1 wynika, że $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\delta) \cdot \operatorname{sgn}\big(\delta^{-1} \circ \sigma\big)$, więc:

$$\det(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i(\delta^{-1} \circ \sigma)(i)} = \operatorname{sgn}(\delta) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}\big(\delta^{-1} \circ \sigma\big) \cdot \prod_{i=1}^{n} a_{i(\delta^{-1} \circ \sigma)(i)}.$$

Ponieważ funkcja $S_n \ni \sigma \longmapsto \delta^{-1} \circ \sigma \in S_n$ jest bijekcją, otrzymujemy stąd:

$$\det(B) = \operatorname{sgn}(\delta) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i\sigma(i)} = \operatorname{sgn}(\delta) \cdot \det(A).$$

Ponieważ transpozycja jest permutacją nieparzystą, to powyższe stwierdzenie implikuje natychmiast następujący

Wniosek 4.3. Jeżeli macierz B powstaje z macierzy kwadratowej A stopnia $n \geq 2$ nad ciałem K wskutek zamiany miejscami dwóch dowolnie ustalonych wierszy (kolumn) macierzy A, to $\det(B) = -\det(A)$.

Stwierdzenie 4.6. Jeżeli B jest macierzą powstałą z macierzy kwadratowej A stopnia n nad ciałem K wskutek pomnożenia pewnego wiersza (kolumny) macierzy A przez ustalony element a ciała K, to $\det(B) = a \cdot \det(A)$.

Dowód. Niech $s \in \{1, 2, \ldots n\}$. Załóżmy, że B jest macierzą powstałą z macierzy $A \in \mathrm{M}_n(K)$ poprzez pomnożenie s-tego wiersza macierzy A przez element $a \in K$.

Wówczas $b_{sj} = a \cdot a_{sj}$ dla każdego $j \in \{1,2,\ldots,n\}$ oraz $b_{ij} = a_{ij}$ dla wszystkich $i, j \in \{1,2,\ldots,n\}$ takich, że $i \neq s$. Stąd oraz na mocy wzoru (4.1.2) otrzymujemy, że $\det(B) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot \big(a \cdot a_{s\sigma(s)}\big) \cdot \ldots \cdot a_{n\sigma(n)} = a \cdot \det(A)$. W świetle Wniosku 4.1 uzyskujemy stąd, że prawdziwa jest również część stwierdzenia dotycząca kolumn.

Bezpośrednią konsekwencją powyższego stwierdzenia jest następujący

Wniosek 4.4. Jeżeli A jest macierzą kwadratową stopnia n nad ciałem K posiającą zerowy wiersz (zerową kolumnę), to $\det(A) = 0$.

Definicja 4.4. Dla dowolnej liczby naturalnej n, zbiór wszystkich parzystych permutacji zbioru n-elementowego oznaczamy symbolem A_n.

Uwaga 4.2. $A_n = \{\sigma \in S_n \colon \operatorname{sgn}(\sigma) = 1\}$.

Stwierdzenie 4.7. Jeżeli macierz kwadratowa A stopnia $n \geq 2$ nad ciałem K posiada dwa identyczne wiersze (dwie identyczne kolumny), to $\det(A) = 0$.

Dowód. Na mocy Wniosku 4.1 wystarczy rozważyć sytuację, w której macierz A ma dwa identyczne wiersze. Niech $A = [a_{ij}]_{ij}$. Załóżmy, że $k, l \in \{1,2,\ldots,n\}$, $k < l$ oraz k-ty i l-ty wiersz macierzy A są identyczne. Wtedy:

$$a_{kj} = a_{lj} \text{ dla każdego } j \in \{1,2,\ldots,n\}. \tag{4.2.2}$$

Rozważmy permutację $\alpha = (k,l)$. Wtedy $\alpha \in S_n \setminus A_n$ oraz $\alpha \circ \alpha = \mathrm{id}$. Stąd oraz na mocy Twierdzenia 1.1 otrzymujemy, że $A_n \ni f \mapsto f \circ \alpha \in S_n \setminus A_n$ jest poprawnie określoną bijekcją. Zatem:

$$\sum_{g \in S_n \setminus A_n} \operatorname{sgn}(g) \cdot \prod_{i=1}^{n} a_{ig(i)} = -\sum_{f \in A_n} \prod_{i=1}^{n} a_{i(f\circ\alpha)(i)}.$$

Ponadto dla dowolnego $f \in A_n$ jest $(f \circ \alpha)(k) = f(l)$, $(f \circ \alpha)(l) = f(k)$ oraz $(f \circ \alpha)(t) = f(t)$, o ile $t \in X_n \setminus \{k,l\}$. Stąd oraz na mocy (4.2.2):

$$\prod_{i=1}^{n} a_{i(f\circ\alpha)(i)} = a_{1f(1)} \cdot \ldots \cdot a_{kf(l)} \cdot \ldots \cdot a_{lf(k)} \cdot \ldots \cdot a_{nf(n)} = \prod_{i=1}^{n} a_{if(i)},$$

dla każdego $f \in A_n$. Wobec tego:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{f \in A_n} \operatorname{sgn}(f) \cdot \prod_{i=1}^{n} a_{if(i)} + \sum_{g \in S_n \setminus A_n} \operatorname{sgn}(g) \cdot \prod_{i=1}^{n} a_{ig(i)}$$

$$= \sum_{f \in A_n} \prod_{i=1}^{n} a_{if(i)} - \sum_{f \in A_n} \prod_{i=1}^{n} a_{if(i)} = 0.$$

Stwierdzenie 4.8. Niech $A = [a_{ij}]_{ij}$ będzie macierzą kwadratową stopnia n nad ciałem K i niech $r \in \{1,2,\ldots,n\}$. Jeżeli dla każdego $j \in \{1,2,\ldots,n\}$, $a_{rj} = \sum_{k=1}^{s} x_{kj}$, gdzie $x_{1j}, x_{2j}, \ldots, x_{sj} \in K$, to:

$$\det(A) = \sum_{k=1}^{s} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r-11} & a_{r-12} & \ldots & a_{r-1n} \\ x_{k1} & x_{k2} & \ldots & x_{kn} \\ a_{r+11} & a_{r+12} & \ldots & a_{r+1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Dowód. Dla każdego $k \in \{1,2,\ldots,s\}$, wyznacznik występujący po prawej stronie powyższej sumy oznaczmy przez W_k. Wówczas:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot \left(\sum_{k=1}^{s} x_{k\sigma(r)} \right) \cdot \ldots \cdot a_{n\sigma(n)}$$

$$= \sum_{k=1}^{s} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot x_{k\sigma(r)} \cdot \ldots \cdot a_{n\sigma(n)} = \sum_{k=1}^{s} W_k.$$

Uwaga 4.3. W świetle Wniosku 4.1, prawdziwa jest także wersja Stwierdzenia 4.8, w której każdy wyraz r-tej kolumny macierzy A jest sumą s składników.

Stwierdzenie 4.9. Niech A będzie macierzą kwadratową stopnia $n \geq 2$ nad ciałem K, niech r oraz s będą różnymi elementami zbioru $\{1,2,\ldots,n\}$ i niech $a \in K$. Niech ponadto B będzie macierzą powstałą z macierzy A wskutek dodania do wyrazów r-tego wiersza (r-tej kolumny) macierzy A odpowiednich wyrazów s-tego wiersza (s-tej kolumny) macierzy A pomnożonych przez a. Wówczas $\det(B) = \det(A)$.

Dowód. Wniosek 4.1 implikuje, że wystarczy udowodnić stwierdzenie w wersji dotyczącej wierszy. Niech C będzie macierzą powstałą z macierzy A poprzez zastąpienie jej r-tego wiersza wierszem s-tym. Ze Stwierdzenia 4.7 wynika wówczas, że $\det(C) = 0$. Stąd oraz odpowiednio na mocy Stwierdzeń 4.8 i 4.6, $\det(B) = \det(A) + a \cdot \det(C) = \det(A)$.

4.3 Operacje elementarne na macierzy

Definicja 4.5. Operacjami elementarnymi na wierszach (kolumnach) $m \times n$-macierzy A nad ciałem K nazywamy następujące czynności:

(OM1) pomnożenie i-tego wiersza (j-tej kolumny) przez niezerowy element a ciała K. Przy wykonywaniu tej operacji każdy wyraz i-tego wiersza (j-tej kolumny) mnożymy przez a, natomiast wszystkie pozostałe wyrazy macierzy A pozostawiamy bez zmian. Operację tę oznaczamy symbolicznie przez: $a \cdot w_i$ ($a \cdot k_j$).

(OM2) Zamiana miejscami i-tego wiersza (i-tej kolumny) macierzy A z wierszem j-tym (z j-tą kolumną), dla $j \neq i$. Wykonując tę operacje zamieniamy miejscami wyłącznie wiersze (kolumny) o numerach i oraz j. Operację tę oznaczamy symbolem $w_i \leftrightarrow w_j$ ($k_i \leftrightarrow k_j$).

(OM3) Dodanie do i-tego wiersza (i-tej kolumny) macierzy A j-tego wiersza (j-tej kolumny) macierzy A pomnożonego (pomnożonej) przez element a ciała K, dla $j \neq i$. Przy wykonywaniu tej operacji nie zmieniamy pozostałych wierszy (kolumn) macierzy A. Operację tę oznaczamy symbolicznie przez $w_i + a \cdot w_j$ ($k_i + a \cdot k_j$).

Uwaga 4.4. Przy obliczaniu wyznaczników macierzy kwadratowych stopnia $n \geq 4$ wykorzystuje się operacje elementarne opisane w Definicji 4.5 w połączeniu z własnościami wyznaczników zawartymi w Stwierdzeniach 4.3 - 4.9.

Przykład 4.1.

$$\begin{vmatrix} 3 & -5 & 2 & 1 \\ 0 & 1 & -2 & 4 \\ 1 & -1 & 2 & -3 \\ 2 & 3 & 4 & 5 \end{vmatrix} \overset{\substack{w_1 - 3w_3 \\ w_4 - 2w_3}}{=} \begin{vmatrix} 0 & -2 & -4 & 10 \\ 0 & 1 & -2 & 4 \\ 1 & -1 & 2 & -3 \\ 0 & 5 & 0 & 11 \end{vmatrix} \overset{w_1 \leftrightarrow w_3}{=} - \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & -2 & -4 & 10 \\ 0 & 5 & 0 & 11 \end{vmatrix} =$$

$$-2 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & -1 & -2 & 5 \\ 0 & 5 & 0 & 11 \end{vmatrix} \overset{\substack{w_3 + w_2 \\ w_4 - 5w_2}}{=} -2 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & -4 & 9 \\ 0 & 0 & 10 & -9 \end{vmatrix} \overset{w_4 + w_3}{=} -2 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & -4 & 9 \\ 0 & 0 & 6 & 0 \end{vmatrix}$$

$$\overset{w_3 \leftrightarrow w_4}{=} 2 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -4 & 9 \end{vmatrix} = 12 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -4 & 9 \end{vmatrix} \overset{w_4 + 4w_3}{=} 12 \begin{vmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{vmatrix}$$

$$= 12 \cdot 1 \cdot 1 \cdot 1 \cdot 9 = 108.$$

4.4 Metoda Laplace'a obliczania wyznaczników

Definicja 4.6. Niech $A = [a_{rs}]_{rs}$ będzie macierzą kwadratową stopnia $n \geq 2$ nad ciałem K i niech $i, j \in \{1, 2, \ldots, n\}$. Symbolem A_{ij} oznaczamy macierz powstałą z macierzy A poprzez wykreślenie z macierzy A i-tego wiersza oraz j-tej kolumny.

Uwaga 4.5. Jeżeli:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} \in \mathrm{M}_n(K), \tag{4.4.1}$$

to

$$A_{ij} = \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \\ a_{i-11} & \dots & a_{i-1j-1} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i+11} & \dots & a_{i+1j-1} & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj-1} & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \in \mathrm{M}_{n-1}(K). \tag{4.4.2}$$

Twierdzenie 4.1 (Laplace). Dla dowolnej macierzy kwadratowej $A = [a_{rs}]_{rs}$ stopnia $n \geq 2$ nad ciałem K oraz dowolnych $i, j \in \{1, 2, \dots, n\}$ zachodzą równości:

$$\det(A) = \sum_{t=1}^{n} (-1)^{i+t} \cdot a_{it} \cdot \det(A_{it}) \tag{4.4.3}$$

oraz

$$\det(A) = \sum_{t=1}^{n} (-1)^{t+j} \cdot a_{tj} \cdot \det(A_{tj}). \tag{4.4.4}$$

Dowód. Rozważamy dowolną macierz kwadratową $A = [a_{rs}]_{rs}$ stopnia $n \geq 2$ nad ciałem K. W świetle Wniosku 4.1 wystarczy udowodnić wzór (4.4.3). Ustalmy więc dowolne $i \in \{1, 2, \dots, n\}$. Dla każdego $t \in \{1, 2, \dots, n\}$ definiujemy macierz:

$$A[i,t] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1t-1} & a_{1t} & a_{1t+1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1t-1} & a_{i-1t} & a_{i-1t+1} & \cdots & a_{i-1n} \\ 0 & 0 & \cdots & 0 & a_{it} & 0 & \cdots & 0 \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1t-1} & a_{i+1t} & a_{i+1t+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nt-1} & a_{nt} & a_{nt+1} & \cdots & a_{nn} \end{bmatrix} \in \mathrm{M}_n(K).$$

Ponieważ:

$$a_{it} = \underbrace{0+0+\ldots+0}_{t-1} + a_{it} + \underbrace{0+0+\ldots+0}_{n-t}$$

dla każdego $t \in \{1, 2, \dots, n\}$, to każdy wyraz i-tego wiersza macierzy A określonej w (4.4.1) jest sumą n skalarów. Spełnione są więc założenia Stwierdzenia 4.8, skąd:

$$\det(A) = \sum_{t=1}^{n} \det\left(A[i,t]\right). \tag{4.4.5}$$

Weźmy dowolne $t \in \{1,2,\ldots,n\}$. Wykonując na macierzy $A[i,t]$ kolejno $n-i$ operacji elementarnych:

$$w_i \leftrightarrow w_{i+1}, w_{i+1} \leftrightarrow w_{i+2}, \ldots, w_{n-1} \leftrightarrow w_n,$$

uzyskujemy macierz:

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,t-1} & a_{1t} & a_{1\,t+1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1\,1} & a_{i-1\,2} & \cdots & a_{i-1\,t-1} & a_{i-1\,t} & a_{i-1\,t+1} & \cdots & a_{i-1\,n} \\ a_{i+1\,1} & a_{i+1\,2} & \cdots & a_{i+1\,t-1} & a_{i+1\,t} & a_{i+1\,t+1} & \cdots & a_{i+1\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n\,t-1} & a_{nt} & a_{n\,t+1} & \cdots & a_{nn} \\ 0 & 0 & \cdots & 0 & a_{it} & 0 & \cdots & 0 \end{bmatrix} \in \mathrm{M}_n(K). \tag{4.4.6}$$

Na mocy Wniosku 4.3 otrzymujemy więc, że:

$$\det(B) = (-1)^{n-i} \cdot \det\left(A[i,t]\right). \tag{4.4.7}$$

Dalej, po wykonaniu na macierzy B kolejno $n-t$ operacji elementarnych:

$$k_t \leftrightarrow k_{t+1}, k_{t+1} \leftrightarrow k_{t+2}, \ldots, k_{n-1} \leftrightarrow k_n,$$

otrzymujemy macierz:

$$C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\,t-1} & a_{1\,t+1} & \cdots & a_{1n} & a_{1t} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{i-1\,1} & a_{i-1\,2} & \cdots & a_{i-1\,t-1} & a_{i-1\,t+1} & \cdots & a_{i-1\,n} & a_{i-1\,t} \\ a_{i+1\,1} & a_{i+1\,2} & \cdots & a_{i+1\,t-1} & a_{i+1\,t+1} & \cdots & a_{i+1\,n} & a_{i+1\,t} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n\,t-1} & a_{n\,t+1} & \cdots & a_{nn} & a_{nt} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & a_{it} \end{bmatrix} \in \mathrm{M}_n(K). \tag{4.4.8}$$

Powołując się ponownie na Wniosek 4.3 uzyskujemy stąd równość:

$$\det(C) = (-1)^{n-i} \cdot \det(B). \tag{4.4.9}$$

Podstawiając (4.4.7) do (4.4.9) uzyskujemy:

$$\det(C) = (-1)^{2n-(i+t)} \cdot \det\left(A[i,t]\right) = (-1)^{-(i+t)} \cdot \det\left(A[i,t]\right),$$

czyli:

$$\det(A[i,t]) = (-1)^{i+t} \cdot \det(C). \tag{4.4.10}$$

Zauważmy, że $[C]_{nt} = 0$ dla każdego $t \in \{1,2,\ldots,n-1\}$. Jeśli więc $\sigma \in S_n$ i $\sigma(n) \neq n$, to istnieje $t \in \{1,2,\ldots,n-1\}$ takie, że $n = \sigma(t)$ i w konsekwencji $[C]_{\sigma(t)t} = 0$. Ponadto funkcja:

$$\begin{pmatrix} 1 & 2 & \ldots & n-1 \\ \delta(1) & \delta(2) & \ldots & \delta(n-1) \end{pmatrix} \overset{F}{\longmapsto} \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ \delta(1) & \delta(2) & \ldots & \delta(n-1) & n \end{pmatrix}$$

jest bijekcją zbioru S_{n-1} na zbiór $\Omega_n = \{\mu \in S_n \colon \mu(n) = n\}$ oraz $I_{F(\delta)} = I_\delta$. Stąd oraz na mocy Wniosku 4.2:

$$\det(C) = \sum_{\sigma \in \Omega_n} \operatorname{sgn}(\sigma) \cdot [C]_{nn} \cdot \prod_{l=1}^{n-1} [C]_{\sigma(l)l} = [C]_{nn} \cdot \sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma) \cdot \prod_{l=1}^{n-1} [C]_{\sigma(l)l}.$$

Ponadto $[C]_{nn} = a_{it}$, $\sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma) \cdot \prod_{l=1}^{n-1} [C]_{\sigma(l)l} = \det C_{nn}$ oraz z (4.4.8) i (4.4.2) wynika, że $C_{nn} = A_{it}$. Zatem:

$$\det(C) = a_{it} \cdot \det(A_{it}). \tag{4.4.11}$$

Podstawiając (4.4.11) do (4.4.10) uzyskujemy, że $\det(A[i,t]) = (-1)^{i+t} \cdot a_{it} \cdot \det(A_{it})$. Stąd oraz na mocy (4.4.10) otrzymujemy tezę.

Uwaga 4.6. Stosowanie wzoru (4.4.3) przy obliczaniu wyznacznika macierzy kwadratowej A stopnia n nad ciałem K nazywa się stosowaniem rozwinięcia Laplace'a względem i-tego wiersza macierzy A. Analogicznie, stosowanie wzoru (4.4.4) przy obliczaniu wyznacznika macierzy A nazywa się stosowaniem rozwinięcia Laplace'a względem j-tej kolumny macierzy A.

Wniosek 4.5. Niech $A = [a_{rs}]_{rs}$ będzie macierzą kwadratową stopnia $n \geq 2$ nad ciałem K. Wówczas dla dowolnych $i, j \in \{1,2,\ldots,n\}$ takich, że $i \neq j$ zachodzą równości:

(i) $\sum_{k=1}^{n} a_{ik} \cdot (-1)^{j+k} \cdot \det(A_{jk}) = 0$;
(ii) $\sum_{k=1}^{n} a_{ki} \cdot (-1)^{k+j} \cdot \det(A_{kj}) = 0$.

Dowód. (i). Weźmy dowolne różne $i, j \in \{1,2,\ldots,n\}$. Ze Stwierdzenia 4.7 wynika, że zastępując j-ty wiersz macierzy A jej i-tym wierszem otrzymujemy macierz B taką, że $\det(B) = 0$. Stąd oraz na mocy Twierdzenia 4.1 otrzymujemy, że $\sum_{k=1}^{n} (-1)^{j+k} \cdot b_{jk} \cdot \det(B_{jk}) = 0$. Ponadto $b_{jk} = a_{ik}$ oraz $B_{jk} = A_{jk}$ dla każdego $k \in \{1,2,\ldots,n\}$, więc zachodzi żądana równość.

(ii). Dowód przebiega analogicznie jak w punkcie (i).

Definicja 4.7. Niech $A = [a_{ij}]_{ij}$ będzie macierzą kwadratową stopnia $n \geq 2$ nad ciałem K i niech $i, j \in \{1, 2, \ldots, n\}$. Element $(-1)^{i+j} \cdot \det(A_{ij})$ ciała K nazywamy dopełnieniem algebraicznym wyrazu a_{ij} macierzy A.

Przykład 4.2. Stosując rozwinięcie Laplace'a względem trzeciego wiersza macierzy $A = \begin{bmatrix} 3 & 1 & 2 & 0 \\ 0 & 1 & 2 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix}$ oraz stosując wzór Sarrusa obliczymy jej wyznacznik. Mamy:

$$\det(A) = \begin{vmatrix} 3 & 1 & 2 & 0 \\ 0 & 1 & 2 & 4 \\ 1 & 0 & 2 & 0 \\ 2 & 3 & 0 & 1 \end{vmatrix} = (-1)^{3+1} \cdot 1 \cdot \begin{vmatrix} 1 & 2 & 0 \\ 1 & 2 & 4 \\ 3 & 0 & 1 \end{vmatrix} + (-1)^{3+2} \cdot 0 \cdot \begin{vmatrix} 3 & 2 & 0 \\ 0 & 2 & 4 \\ 2 & 0 & 1 \end{vmatrix} +$$

$$(-1)^{3+3} \cdot 2 \cdot \begin{vmatrix} 3 & 1 & 0 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{vmatrix} + (-1)^{3+4} \cdot 0 \cdot \begin{vmatrix} 3 & 1 & 2 \\ 0 & 1 & 2 \\ 2 & 3 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 0 \\ 1 & 2 & 4 \\ 3 & 0 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 3 & 1 & 0 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{vmatrix} =$$

$(1 \cdot 2 \cdot 1 + 1 \cdot 0 \cdot 0 + 2 \cdot 4 \cdot 3) - (3 \cdot 2 \cdot 0 + 1 \cdot 2 \cdot 1 + 0 \cdot 4 \cdot 1) + 2 \cdot \big((3 \cdot 1 \cdot 1 + 0 \cdot 3 \cdot 0 + 1 \cdot 4 \cdot 2) - (2 \cdot 1 \cdot 0 + 0 \cdot 1 \cdot 1 + 3 \cdot 4 \cdot 3)\big) = 24 + 2 \cdot (-25) = -26.$

Uwaga 4.7. W praktyce, przy obliczaniu wyznaczników często łączy się wszystkie poznane metody i własności wyznacznika. W szczególności, przed zastosowaniem metody Laplace'a, warto za pomocą operacji elementarnych wyzerować maksymalną ilość wyrazów wiersza lub kolumny, względem którego lub której chcemy zastosować rozwinięcie Laplace'a. Pozwala to na skrócenie i uproszczenie niezbędnych do wykonania rachunków.

Rozdział 5

Rachunek macierzowy

5.1 Podstawowe operacje na macierzach

5.1.1 Dodawanie i odejmowanie macierzy

Definicja 5.1. Niech $m,n \in \mathbb{N}$ i niech K będzie ciałem. Sumą macierzy $A,B \in \mathrm{M}_{m\times n}(K)$ nazywamy taką macierz $C \in \mathrm{M}_{m\times n}(K)$, że $[C]_{ij} = [A]_{ij} + [B]_{ij}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$. Sumę macierzy A i B oznaczamy standardowo przez $A+B$. Mamy więc $[A+B]_{ij} = [A]_{ij} + [B]_{ij}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$.

Uwaga 5.1. Zapisując $m \times n$-macierze w notacji $[a_{ij}]_{ij}$ oraz $[b_{ij}]_{ij}$ otrzymujemy, że $[a_{ij}]_{ij} + [b_{ij}]_{ij} = [a_{ij} + b_{ij}]_{ij}$.

Przykład 5.1.
$$\begin{bmatrix} 5 & 10 & -2 \\ 4 & 2 & 0 \end{bmatrix} + \begin{bmatrix} -4 & 0 & 1 \\ 3 & -2 & 4 \end{bmatrix} = \begin{bmatrix} 5+(-4) & 10+0 & -2+1 \\ 4+3 & 2+(-2) & 0+4 \end{bmatrix} = \begin{bmatrix} 1 & 10 & -1 \\ 7 & 0 & 4 \end{bmatrix}.$$

Uwaga 5.2. Rozważmy zbiór wszystkich $m \times n$-macierzy nad dowolnym ciałem K. Ponieważ w ciele K dodawanie jest łączne i przemienne, to z Definicji 5.1 wynika, że łączne i przemienne jest dodawanie macierzy. Ponadto macierz $\Theta_{m\times n} \in \mathrm{M}_{m\times n}(K)$, której wszystkie wyrazy są równe 0 jest elementem neutralnym dodawania macierzy. Łatwo również zauważyć, że jeśli $A = [a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$, to $[-a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$ jest elementem przeciwnym do A względem dodawania macierzy. Macierz $[-a_{ij}]_{ij}$ oznaczamy przez $-A$ i nazywamy ją macierzą przeciwną do macierzy A. W szczególności przez różnicę $A-B$ macierzy $A,B \in \mathrm{M}_{m\times n}(K)$ rozumiemy sumę macierzy A i $-B$.

Bezpośrednią konsekwencją obserwacji poczynionych w Uwadze 5.2 jest następujące

Stwierdzenie 5.1. Dla dowolnych liczb naturalnych m i n oraz dowolnego ciała K zbiór $\mathrm{M}_{m\times n}(K)$ wszystkich $m \times n$-macierzy nad ciałem K tworzy grupę abelową względem dodawania macierzy.

5.1.2 Mnożenie macierzy przez skalar

Uwaga 5.3. W sytuacji gdy ciało K rozważamy wraz z jakąś inną strukturą algebraiczną z nim związaną, elementy ciała K nazywamy skalarami.

Definicja 5.2. Niech $m,n \in \mathbb{N}$ i niech K będzie ciałem. Iloczynem macierzy $A \in \mathrm{M}_{m\times n}(K)$ przez skalar $a \in K$ nazywamy macierz $C \in \mathrm{M}_{m\times n}(K)$ taką, że $[C]_{ij} = a \cdot [A]_{ij}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$. Macierz tę oznaczamy symbolem $a \cdot A$.

Przykład 5.2. $2 \cdot \begin{bmatrix} 5 & 10 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 2\cdot 5 & 2\cdot 10 \\ 2\cdot 4 & 2\cdot 2 \end{bmatrix} = \begin{bmatrix} 10 & 20 \\ 8 & 4 \end{bmatrix}$.

Wprost z Definicji 5.1 i 5.2 wynika następujące

Stwierdzenie 5.2. Niech $m,n \in \mathbb{N}$ i niech K będzie ciałem. Dla dowolnych $A,B \in \mathrm{M}_{m\times n}(K)$ oraz $a,b \in K$ zachodzą równości:

(i) $a \cdot (A+B) = a \cdot A + a \cdot B$;
(ii) $(a+b) \cdot A = a \cdot A + b \cdot A$;
(iii) $(ab) \cdot A = a \cdot (b \cdot A)$.

5.1.3 Mnożenie macierzy

Uwaga 5.4. Iloczyn macierzy A przez macierz B określa się tylko wówczas, gdy liczba kolumn macierzy A równa jest liczbie wierszy macierzy B. Można więc pomnożyć $m \times n$-macierz przez $n \times s$-macierz. Sposób w jaki się to robi opisany jest w Definicji 5.3.

Definicja 5.3. Niech m,n oraz s będą liczbami naturalnymi i niech K będzie ciałem. Iloczynem macierzy $A \in \mathrm{M}_{m\times n}(K)$ i $B \in \mathrm{M}_{n\times s}(K)$ nazywamy taką macierz $C \in \mathrm{M}_{m\times s}(K)$, że $[C]_{ij} = \sum_{k=1}^{n} [A]_{ik} \cdot [B]_{kj}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,s\}$. Iloczyn macierzy A i B oznaczamy standardowo przez $A \cdot B$. Mamy więc $[A \cdot B]_{ij} = \sum_{k=1}^{n} [A]_{ik} \cdot [B]_{kj}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,s\}$.

Uwaga 5.5. Gdy mamy do czynienia z zapisem $A = [a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$, $B = [b_{ij}]_{ij} \in \mathrm{M}_{n\times s}(K)$ oraz $C = A \cdot B$, to $C = [c_{ij}]_{ij} \in \mathrm{M}_{m\times s}(K)$ i dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2\ldots,s\}$ zachodzi $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Przykład 5.3. Rozważmy macierze $A = \begin{bmatrix} 1 & 4 & 5 \\ 3 & 4 & 2 \end{bmatrix}$ oraz $B = \begin{bmatrix} 3 & 0 & 2 & 5 \\ 0 & 1 & 4 & 0 \\ 1 & 1 & 0 & 2 \end{bmatrix}$ nad ciałem $\mathbb{R}$.
Ponieważ A jest 2×3-macierzą i B jest 3×4-macierzą, to nie istnieje iloczyn $B \cdot A$ oraz istnieje iloczyn $A \cdot B$ i jest on 2×4-macierzą. Mianowicie:

$$A \cdot B = \begin{bmatrix} 8 & 9 & 18 & 15 \\ 11 & 6 & 22 & 19 \end{bmatrix},$$

gdyż: $8 = [A \cdot B]_{11} = \sum_{k=1}^{3} [A]_{1k} \cdot [B]_{k1} = 1 \cdot 3 + 4 \cdot 0 + 5 \cdot 1 = 3 + 5,$

$9 = [A \cdot B]_{12} = \sum_{k=1}^{3} [A]_{1k} \cdot [B]_{k2} = 1 \cdot 0 + 4 \cdot 1 + 5 \cdot 1 = 4 + 5,$

$18 = [A \cdot B]_{13} = \sum_{k=1}^{3} [A]_{1k} \cdot [B]_{k3} = 1 \cdot 2 + 4 \cdot 4 + 5 \cdot 0 = 2 + 16,$

$15 = [A \cdot B]_{14} = \sum_{k=1}^{3} [A]_{1k} \cdot [B]_{k4} = 1 \cdot 5 + 4 \cdot 0 + 5 \cdot 2 = 5 + 10,$

$11 = [A \cdot B]_{21} = \sum_{k=1}^{3} [A]_{2k} \cdot [B]_{k1} = 3 \cdot 3 + 4 \cdot 0 + 2 \cdot 1 = 9 + 2,$

$6 = [A \cdot B]_{22} = \sum_{k=1}^{3} [A]_{2k} \cdot [B]_{k2} = 3 \cdot 0 + 4 \cdot 1 + 2 \cdot 1 = 4 + 2,$

$22 = [A \cdot B]_{23} = \sum_{k=1}^{3} [A]_{2k} \cdot [B]_{k3} = 3 \cdot 2 + 4 \cdot 4 + 2 \cdot 0 = 6 + 16,$

$19 = [A \cdot B]_{24} = \sum_{k=1}^{3} [A]_{2k} \cdot [B]_{k4} = 3 \cdot 5 + 4 \cdot 0 + 2 \cdot 2 = 15 + 4.$

Wniosek 5.1. Niech $m, n \in \mathbb{N}$ i niech K będzie ciałem. Mnożenie macierzy jest działaniem w zbiorze $\mathrm{M}_{m \times n}(K)$ wtedy i tylko wtedy, gdy $m = n$.

Bezpośrednią konsekwencją Definicji 5.2 i 5.3 oraz łączności mnożenia w dowolnym ciele jest następujące

Stwierdzenie 5.3. Niech m, n, s, r będą liczbami naturalnymi i niech K będzie ciałem. Wówczas dla wszystkich $A \in \mathrm{M}_{m \times n}(K)$, $B \in \mathrm{M}_{n \times s}(K)$ oraz $a \in K$ zachodzą równości:

$$a \cdot (A \cdot B) = (a \cdot A) \cdot B = A \cdot (a \cdot B).$$

Stwierdzenie 5.4. Niech m, n, s, r będą liczbami naturalnymi i niech K będzie ciałem. Wówczas dla wszystkich $A \in \mathrm{M}_{m \times n}(K)$, $B \in \mathrm{M}_{n \times s}(K)$ oraz $C \in \mathrm{M}_{s \times r}(K)$ zachodzi równość:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Dowód. Ponieważ $A \cdot B \in \mathrm{M}_{m \times s}(K)$ i $B \cdot C \in \mathrm{M}_{n \times r}(K)$, to $(A \cdot B) \cdot C, A \cdot (B \cdot C) \in \mathrm{M}_{m \times r}(K)$. Ponadto, dla wszystkich $i \in \{1, 2, \ldots, m\}$ oraz $j \in \{1, 2, \ldots, r\}$, otrzymujemy, że:

$$\big[(A\cdot B)\cdot C\big]_{ij}=\sum_{k=1}^{s}[A\cdot B]_{ik}\cdot[C]_{kj}=\sum_{k=1}^{s}\left(\sum_{t=1}^{n}[A]_{it}\cdot[B]_{tk}\right)\cdot[C]_{kj}=$$

$$\sum_{k=1}^{s}\sum_{t=1}^{n}\big([A]_{it}\cdot[B]_{tk}\big)\cdot[C]_{kj}=\sum_{k=1}^{s}\sum_{t=1}^{n}[A]_{it}\cdot\big([B]_{tk}\cdot[C]_{kj}\big)=\sum_{t=1}^{n}\sum_{k=1}^{s}[A]_{it}\cdot\big([B]_{tk}\cdot[C]_{kj}\big)=$$

$$\sum_{t=1}^{n}[A]_{it}\cdot\sum_{k=1}^{s}[B]_{tk}\cdot[C]_{kj}=\sum_{t=1}^{n}[A]_{it}\cdot[B\cdot C]_{tj}=\big[A\cdot(B\cdot C)\big]_{ij}.$$

Stwierdzenie 5.5. Niech m,n,s,r będą liczbami naturalnymi i niech K będzie ciałem. Wówczas dla wszystkich $A,B\in \mathrm{M}_{m\times n}(K)$, $C\in \mathrm{M}_{n\times s}(K)$ oraz $D\in \mathrm{M}_{r\times m}(K)$ zachodzą równości:

(i) $(A+B)\cdot C=A\cdot C+B\cdot C$;
(ii) $D\cdot(A+B)=D\cdot A+D\cdot B$.

Dowód. (i). Jasne jest, że $(A+B)\cdot C, A\cdot C+B\cdot C\in \mathrm{M}_{m\times s}(K)$. Ponadto, dla wszystkich $i\in\{1,2,\ldots,m\}$ oraz $j\in\{1,2,\ldots,s\}$, otrzymujemy, że: $\big[(A+B)\cdot C\big]_{ij}=\sum_{k=1}^{n}[A+B]_{ik}\cdot[C]_{kj}=\sum_{k=1}^{n}\big([A]_{ik}+[B]_{ik}\big)\cdot[C]_{kj}=\sum_{k=1}^{n}[A]_{ik}\cdot[C]_{kj}+[B]_{ik}\cdot[C]_{kj}=\sum_{k=1}^{n}[A]_{ik}\cdot[C]_{kj}+\sum_{k=1}^{n}[B]_{ik}\cdot[C]_{kj}=[A\cdot C]_{ij}+[B\cdot C]_{ij}=[A\cdot C+B\cdot C]_{ij}$.
(ii). Dowód przebiega analogicznie jak w punkcie (i).

Stwierdzenie 5.6. Niech m,n,s będą liczbami naturalnymi i niech K będzie ciałem. Wówczas dla wszystkich $A\in \mathrm{M}_{m\times n}(K)$ i $B\in \mathrm{M}_{n\times s}(K)$ zachodzi równość:

$$(A\cdot B)^T=B^T\cdot A^T.$$

Dowód. Jasne jest, że $A\cdot B\in \mathrm{M}_{m\times s}(K)$, $A^T\in \mathrm{M}_{n\times m}(K)$ i $B^T\in \mathrm{M}_{s\times n}(K)$. Zatem $(A\cdot B)^T, B^T\cdot A^T\in \mathrm{M}_{s\times m}(K)$. Ponadto dla wszystkich $i\in\{1,2,\ldots,s\}$ oraz $j\in\{1,2,\ldots,m\}$ uzyskujemy, że $\big[(A\cdot B)^T\big]_{ij}=[A\cdot B]_{ji}=\sum_{k=1}^{n}[A]_{jk}\cdot[B]_{ki}=\sum_{k=1}^{n}[B]_{ki}\cdot[A]_{jk}$
$=\sum_{k=1}^{n}\big[B^T\big]_{ik}\cdot\big[A^T\big]_{kj}=\big[B^T\cdot A^T\big]_{ij}$.

Stwierdzenie 5.7. Dla dowolnych liczb naturalnych m i n oraz dowolnych $m\times n$-macierzy A i B nad ciałem K równość:

$$A\cdot\begin{bmatrix}x_1\\x_2\\\vdots\\x_m\end{bmatrix}=B\cdot\begin{bmatrix}x_1\\x_2\\\vdots\\x_m\end{bmatrix}\tag{5.1.1}$$

zachodząca dla wszystkich $x_1,x_2,\ldots,x_n\in K$ implikuje, że $A=B$.

Dowód. Załóżmy, że równość (5.1.1) zachodzi dla wszystkich $x_1, x_2, \ldots, x_n \in K$. Weźmy dowolne $j \in \{1,2,\ldots,n\}$. Niech $x_j = 1$ oraz $x_k = 0$ dla każdego $k \in \{1,2,\ldots,n\} \setminus \{j\}$. Z określenia mnożenia macierzy wynika wówczas, że dla każdego $i \in \{1,2,\ldots,m\}$, i-tymi wierszami macierzy $m \times 1$-macierzy:

$$A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \text{ i } B \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

są odpowiednio $[A]_{ij}$ oraz $[B]_{ij}$. Stąd oraz na mocy dowolności wyboru elementów $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$ otrzymujemy, że $[A]_{ij} = [B]_{ij}$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$, czyli $A = B$.

5.1.4 Mnożenie macierzy kwadratowych

Uwaga 5.6. Niech $n \in \mathbb{N}$ i niech K będzie ciałem. Na mocy Wniosku 5.1 otrzymujemy, że mnożenie macierzy jest działaniem w zbiorze $\mathrm{M}_n(K)$. Łatwo zauważyć, że macierz:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 1 \end{bmatrix} \in \mathrm{M}_n(K)$$

jest elementem neutralnym tego działania. Oznaczamy ją symbolem I_n.

Definicja 5.4. Niech $n \in \mathbb{N}$ i niech λ będzie elementem ciała K. Macierz $\lambda \cdot I_n$ nazywamy macierzą skalarną stopnia n nad ciałem K (wyznaczoną przez λ). Macierz skalarną stopnia n nad ciałem K wyznaczoną przez 1, czyli macierz I_n, nazywamy macierzą jednostkową stopnia n nad ciałem K.

Bezpośrednią konsekwencją powyższej definicji oraz Stwierdzenia 5.3 jest następujący

Wniosek 5.2. Niech $n \in \mathbb{N}$ i niech K będzie ciałem. Wówczas $(\lambda \cdot I_n) \cdot A = A \cdot (\lambda \cdot I_n)$ dla wszystkich $\lambda \in K$ oraz $A \in \mathrm{M}_n(K)$.

Definicja 5.5. Niech $n > 1$ będzie liczbą naturalną, niech $i, j \in \{1,2,\ldots n\}$ i niech K będzie ciałem. Symbolem E_{ij} oznaczamy taką macierz $X \in \mathrm{M}_n(K)$, że $[X]_{ij} = 1$ oraz $[X]_{kl} = 0$ dla wszystkich $k, l \in \{1,2,\ldots n\}$ takich, że $(k,l) \neq (i,j)$.

Przykład 5.4. Jeśli K jest ciałem i $E_{32} \in \mathrm{M}_4(K)$, to $E_{32} = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{bmatrix}$.

Bezpośrednią konsekwencją określenia mnożenia macierzy jest poniższe

Stwierdzenie 5.8. Niech $n \in \mathbb{N} \setminus \{1\}$, $i, j, k, l \in \{1, 2, \ldots n\}$, niech K będzie ciałem i niech $E_{ij}, E_{kl}, A \in \mathrm{M}_n(K)$. Wówczas:

(i) $A = \sum_{i=1}^{n} \sum_{j=1}^{n} [A]_{ij} \cdot E_{ij}$;

(ii) $E_{ij} \cdot E_{kl} = \begin{cases} E_{il} \text{ , gdy } j = k \\ \Theta_n \text{ , gdy } j \neq k \end{cases}$,

gdzie Θ_n oznacza kwadratową macierz zerową stopnia n nad ciałem K.

Stwierdzenie 5.9. Niech $n \in \mathbb{N}$ i niech K będzie ciałem. Mnożenie macierzy jest przemiennym działaniem w zbiorze $\mathrm{M}_n(K)$ wtedy i tylko wtedy, gdy $n = 1$.

Dowód. Jeżeli $n = 1$, to dla dowolnych $A, B \in \mathrm{M}_n(K)$ istnieją $a, b \in K$ takie, że $A = [a]$ oraz $B = [b]$, skąd $B \cdot A = [b] \cdot [a] = [ba] = [ab] = [a] \cdot [b] = A \cdot B$. Na odwrót. Załóżmy, że $n > 1$. Niech $E_{11}, E_{12} \in \mathrm{M}_n K$. Wtedy $E_{11} \cdot E_{12} = E_{12}$ oraz $E_{12} \cdot E_{11} = \Theta_n$. Zatem $E_{11} \cdot E_{12} \neq E_{12} \cdot E_{11}$.

5.1.5 Twierdzenie Cauchy'ego

Twierdzenie 5.1 (Cauchy). Niech $n \in \mathbb{N}$ i niech K będzie ciałem. Dla wszystkich $A, B \in \mathrm{M}_n(K)$ zachodzi równość $\det(A \cdot B) = \det(A) \cdot \det(B)$.

Dowód. Rozważmy dowolne $A, B \in \mathrm{M}_n(K)$ oraz $i, j \in \{1, 2, \ldots, n\}$. Wtedy:

$$[A \cdot B]_{ij} = \sum_{t=1}^{n} [A]_{it} \cdot [B]_{tj}. \tag{5.1.2}$$

Niech κ_j oznacza j-tą kolumnę macierzy $A \cdot B$, zaś α_j - j-tą kolumnę macierzy A. Na mocy (5.1.2) otrzymujemy wówczas:

$$\kappa_j = \begin{bmatrix} \sum_{t=1}^{n} [A]_{1t} \cdot [B]_{tj} \\ \sum_{t=1}^{n} [A]_{2t} \cdot [B]_{tj} \\ \vdots \\ \sum_{t=1}^{n} [A]_{nt} \cdot [B]_{tj} \end{bmatrix} = \sum_{t=1}^{n} [B]_{tj} \cdot \alpha_t. \tag{5.1.3}$$

Wobec tego:

$$A \cdot B = \left[\sum_{t_1=1}^{n} [B]_{t_1 1} \cdot \alpha_{t_1}, \sum_{t_2=1}^{n} [B]_{t_2 2} \cdot \alpha_{t_2}, \cdots, \sum_{t_n=1}^{n} [B]_{t_n n} \cdot \alpha_{t_n} \right]. \tag{5.1.4}$$

Stąd oraz kolejno na mocy Stwierdzenia 4.8, Uwagi 4.3 i Stwierdzenia 4.6 zastosowanych n-krotnie uzyskujemy, że:

$$\det(A \cdot B) = \sum_{t_1=1}^{n} \sum_{t_3=1}^{n} \ldots \sum_{t_n=1}^{n} \prod_{s=1}^{n} [B]_{t_s s} \cdot \det\left[\alpha_{t_1} \alpha_{t_2}, \ldots, \alpha_{t_n}\right]. \tag{5.1.5}$$

Ponadto ze Stwierdzenia 4.7 wynika, że we wzorze (5.1.5) tylko te składniki sumy mogą być niezerowe, dla których $\left|\{t_1, t_2, \ldots, t_n\}\right| = n$, czyli takie, że $\{t_1, t_2, \ldots, t_n\} = \left\{\sigma(1), \sigma(2), \ldots, \sigma(n)\right\}$ dla pewnego $\sigma \in S_n$. Zatem:

$$\det(A \cdot B) = \sum_{\sigma \in S_n} \prod_{s=1}^{n} [B]_{\sigma(s)s} \cdot \det\left[\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(n)}\right]. \tag{5.1.6}$$

Ponadto $\det\left[\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \ldots, \alpha_{\sigma(n)}\right] = \operatorname{sgn}(\sigma) \cdot \det\left[\alpha_1, \alpha_2, \ldots, \alpha_n\right] = \operatorname{sgn}(\sigma) \cdot \det A$, na mocy Stwierdzenia 4.5, więc równość (5.1.6) zapisuje się w równoważnej postaci:

$$\det(A \cdot B) = \sum_{\sigma \in S_n} \prod_{s=1}^{n} [B]_{\sigma(s)s} \cdot \operatorname{sgn}(\sigma) \cdot \det(A), \tag{5.1.7}$$

a stąd:

$$\det(A \cdot B) = \det(A) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{s=1}^{n} [B]_{\sigma(s)s} = \det(A) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{s=1}^{n} \left[B^T\right]_{s\sigma(s)}$$

$$= \det(A) \cdot \det\left(B^T\right).$$

Ponadto $\det\left(B^T\right) = \det(B)$ na mocy Stwierdzenia 4.4, więc ostatecznie $\det(A \cdot B) = \det(A) \cdot \det(B)$.

Definicja 5.6. Niech $n \in \mathbb{N}$ i niech K będzie ciałem. Macierzą dopełnień macierzy $A \in \mathrm{M}_n(K)$ nazywamy macierz $\mathscr{D}(A) = \left[(-1)^{i+j} \cdot \det(A_{ij})\right]_{ij} \in \mathrm{M}_n(K)$.

Stwierdzenie 5.10. Dla dowolnej macierzy kwadratowej A stopnia n nad ciałem K zachodzą równości:

$$A \cdot \mathscr{D}(A)^T = \mathscr{D}(A)^T \cdot A = \det(A) \cdot I_n.$$

Dowód. Weźmy dowolne $i, j \in \{1, 2, \ldots, n\}$. Wówczas: $\left[A \cdot \mathscr{D}(A)^T\right]_{ij} = \sum_{k=1}^{n} [A]_{ik} \cdot \left[\mathscr{D}(A)^T\right]_{kj} = \sum_{k=1}^{n} [A]_{ik} \cdot \left[\mathscr{D}(A)\right]_{jk} = \sum_{k=1}^{n} [A]_{ik} \cdot (-1)^{k+j} \cdot \det\left(A_{jk}\right)$, więc odpowiednio z Twierdzenia 4.1 oraz punktu (i) Wniosku 4.5 otrzymujemy, że:

$$\left[A \cdot \mathscr{D}(A)^T\right]_{ij} = \begin{cases} \det(A) & \text{, dla } j = i, \\ 0 & \text{, dla } j \neq i. \end{cases}$$

Stąd $A \cdot \mathscr{D}(A)^T = \det(A) \cdot I_n$. Równość $\mathscr{D}(A)^T \cdot A = \det(A) \cdot I_n$ uzasadnia się analogicznie.

5.1.6 Macierz odwrotna

Definicja 5.7. Macierz kwadratową stopnia n nad ciałem K nazywamy odwracalną, gdy jest ona elementem odwracalnym względem mnożenia macierzy w zbiorze $\mathrm{M}_n(K)$, czyli gdy istnieje $X \in \mathrm{M}_n(K)$ takie, że $A \cdot X = X \cdot A = I_n$. Jeżeli taka macierz X istnieje, to nazywamy ją macierzą odwrotną do A i oznaczamy symbolem A^{-1}.

Twierdzenie 5.2. Macierz kwadratowa stopnia n nad ciałem K jest odwracalna wtedy i tylko wtedy, gdy $\det(A) \neq 0$. Ponadto, jeśli $\det(A) \neq 0$, to $A^{-1} = \det(A)^{-1} \cdot \mathscr{D}(A)^T$.

Dowód. Załóżmy, że macierz $A \in \mathrm{M}_n(K)$ jest odwracalna. Istnieje wówczas $X \in \mathrm{M}_n(K)$ takie, że $A \cdot X = I_n$. Stąd oraz na mocy Twierdzenia 5.1, $1 = \det(I_n) = \det(A \cdot X) = \det(A) \cdot \det(X)$, skąd $\det(A) \neq 0$.

Przypuśćmy teraz, że $\det(A) \neq 0$. Ze Stwierdzeń 5.10 i 5.3 oraz punktu (iii) Stwierdzenia 5.2 wynika wówczas, że $A \cdot \left(\det(A)^{-1} \cdot \mathscr{D}(A)^T\right) = \left(\det(A)^{-1} \cdot \left(A \cdot \mathscr{D}(A)^T\right)\right. = \det(A)^{-1} \cdot \left(\det(A) \cdot I_n\right) = I_n$. Analogicznie, $\left(\det(A)^{-1} \cdot \mathscr{D}(A)^T\right) \cdot A = I_n$. Zatem macierz A jest odwracalna oraz $A^{-1} = \det(A)^{-1} \cdot \mathscr{D}(A)^T$.

Definicja 5.8. Macierz kwadratową A stopnia n nad ciałem K nazywamy osobliwą, gdy $\det(A) = 0$. Gdy $\det(A) \neq 0$, to mówimy, że macierz A jest nieosobliwa.

Wniosek 5.3. Macierz kwadratową stopnia n nad ciałem K jest nieosobliwa wtedy i tylko wtedy, gdy jest ona odwracalna.

Stwierdzenie 5.11. Dla dowolnych macierzy kwadratowych A i X stopnia n nad ciałem K następujące warunki są równoważne:

(i) $A^{-1} = X$;
(ii) $A \cdot X = I_n$;
(iii) $X \cdot A = I_n$.

Dowód. $(i) \Rightarrow (ii)$. Oczywiste.
$(ii) \Rightarrow (iii)$. Z Twierdzenia 5.1 wynika, że $1 = \det(I_n) = \det(A \cdot X) = \det(A) \cdot \det(X)$, skąd $\det(A) \neq 0$. Z Twierdzenia 5.2 wynika więc istnienie A^{-1}. Zatem $A^{-1} = A^{-1} \cdot I_n = A^{-1} \cdot (A \cdot X) = \left(A^{-1} \cdot A\right) \cdot X = I_n \cdot X = X$ i w konsekwencji $X \cdot A = I_n$.
$(iii) \Rightarrow (i)$. Analogicznie jak w dowodzie poprzedniej implikacji uzasadnia się istnienie A^{-1}. Ponadto $A^{-1} = I_n \cdot A^{-1} = (X \cdot A) \cdot A^{-1} = X$, więc $A \cdot X = I_n$.

Przykład 5.5. Nad ciałem $\mathbb{R}$ rozważmy macierz $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Wtedy $\det(A) = (-1)^{3+3} \cdot 2 \cdot \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} = 2 \cdot (1^2 - 2 \cdot 0) = 2 \neq 0$, więc z Twierdzenia 5.2 wynika, że macierz A^{-1} istnieje oraz $A^{-1} = \det(A)^{-1} \cdot \mathscr{D}(A)^T$. Oznaczmy $d_{ij} = \left[\mathscr{D}(A)\right]_{ij}$ dla wszystkich $i, j \in \{1, 2 \ldots, n\}$. Wtedy:

$$d_{11} = (-1)^{1+1} \cdot \det(A_{11}) = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2, \; d_{12} = (-1)^{1+2} \cdot \det(A_{12}) = -\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = -4,$$

$$d_{13} = (-1)^{1+3} \cdot \det(A_{13}) = \begin{vmatrix} 2 & 1 \\ 0 & 0 \end{vmatrix} = 0, \; d_{21} = (-1)^{2+1} \cdot \det(A_{21}) = -\begin{vmatrix} 0 & 3 \\ 0 & 2 \end{vmatrix} = 0,$$

$$d_{22} = (-1)^{2+2} \cdot \det(A_{22}) = \begin{vmatrix} 1 & 3 \\ 0 & 2 \end{vmatrix} = 2, \; d_{23} = (-1)^{2+3} \cdot \det(A_{23}) = -\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = 0,$$

$$d_{31} = (-1)^{3+1} \cdot \det(A_{31}) = \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} = -3, \; d_{32} = (-1)^{3+2} \cdot \det(A_{32}) = -\begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} = 6,$$

$d_{33} = (-1)^{3+3} \cdot \det(A_{33}) = \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} = 1$. Zatem $\mathscr{D}(A) = \begin{bmatrix} 2 & -4 & 0 \\ 0 & 2 & 0 \\ -3 & 6 & 1 \end{bmatrix}$, skąd $A^{-1} = \det(A)^{-1} \cdot \mathscr{D}(A)^T = \frac{1}{2} \cdot \begin{bmatrix} 2 & 0 & -3 \\ -4 & 2 & 6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ -2 & 1 & 3 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$.

Stwierdzenie 5.12. Niech $n \in \mathbb{N}$, niech K będzie ciałem i niech $A, B \in \mathrm{M}_n(K)$. Jeżeli macierze A i B są odwracalne, to odwracalna jest również macierz $A \cdot B$. Ponadto $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Dowód. Wystarczy zauważyć, że $(A \cdot B) \cdot \left(B^{-1} \cdot A^{-1}\right) = A \cdot \left(B \cdot B^{-1}\right) \cdot A^{-1} = A \cdot I_n \cdot A^{-1} = A \cdot A^{-1} = I_n$ i powołać się na Stwierdzenie 5.11.

5.1.7 Odwracanie macierzy za pomocą operacji elementarnych

Uwaga 5.7. Niech A będzie macierzą kwadratową stopnia n nad ciałem K. Wprost z określenia mnożenia macierzy wynika, że wykonanie operacji elementarnej na wierszach macierzy A tożsame jest z pomnożeniem macierzy A z lewej strony przez macierz powstałą z macierzy jednostkowej I_n stopnia n, na której wykonano tę samą operację elementarną. Przypuśćmy, że $\det(A) \neq 0$. Stosując operacje elementarne na wierszach macierzy A można sprowadzić ją wówczas do macierzy I_n (uzasadnienie tego faktu wymaga wiedzy z zakresu liniowej niezależności wektorów, która zostanie zaprezentowana w dalszych rozdziałach tej książki - zob. też Uwaga 9.4). Istnieją więc macierze $X_1, X_2, \ldots, X_s \in M_n(K)$ takie, że $X_s \cdot \ldots \cdot X_2 \cdot X_1 \cdot A = I_n$. Stąd $A^{-1} = X_s \cdot \ldots \cdot X_2 \cdot X_1$ i w konsekwencji $A^{-1} = X_s \cdot \ldots \cdot X_2 \cdot X_1 \cdot I_n$. Wobec tego macierz A^{-1} powstaje wskutek wykonania na wierszach macierzy I_n tych samych operacji elementarnych, które zostały wykonane na macierzy A przy sprowadzaniu jej do macierzy I_n.

Poniższy przykład pokazuje praktyczne zastosowanie Uwagi 5.7 przy wyznaczaniu macierzy odwrotnej do nieosobliwej macierzy A.

Przykład 5.6. Niech A będzie macierzą z Przykładu 5.5. Wtedy $\det(A) \neq 0$, więc istnieje macierz A^{-1}. Wyznaczymy ją, wykonując operacje elementarne na wierszach macierzy blokowej $[A|I_n]$, które sprowadzą macierz A do macierzy I_n i tym samym, na mocy Uwagi 5.7, sprowadzą one macierz I_n do macierzy A^{-1}. Mamy:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \begin{array}{c} w_2 - 2w_1 \\ \equiv \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -6 & -2 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \begin{array}{c} \frac{1}{2} \cdot w_3 \\ \equiv \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -6 & -2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \end{array}\right] \begin{array}{c} w_1 - 3w_3 \\ w_2 + 6w_3 \\ \equiv \end{array} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & -2 & 1 & 3 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \end{array}\right],$$

$$\text{więc } A^{-1} = \left[\begin{array}{ccc} 1 & 0 & -\frac{3}{2} \\ -2 & 1 & 3 \\ 0 & 0 & \frac{1}{2} \end{array}\right].$$

Istnieje również dualna metoda odwracania nieosobliwych macierzy za pomocą operacji elementarnych.

Uwaga 5.8. Niech A będzie nieosobliwą macierzą kwadratową stopnia n nad ciałem K. Wówczas wykonanie operacji elementarnej na kolumnach macierzy A prowadzi do macierzy, którą można otrzymać również wskutek pomnożenia macierzy A z prawej strony przez macierz powstałą z macierzy I_n w wyniku wykonania na niej tej samej operacji elementarnej. Wykorzystując operacje elementarne na kolumnach macierzy A można sprowadzić ją wówczas do macierzy I_n (podobnie jak w Uwadze 5.7, uzasadnienie tego faktu wykracza nieco poza omówiony dotychczas zakres wiadomości; zob. ew. Uwaga 9.4). Istnieją więc macierze $X_1, X_2, \ldots, X_s \in M_n(K)$ takie, że $A \cdot X_1 \cdot X_2 \cdot \ldots \cdot X_s = I_n$. Zatem $A^{-1} = X_1 \cdot X_2 \cdot \ldots \cdot X_s$, skąd $A^{-1} = I_n \cdot X_1 \cdot X_2 \cdot \ldots \cdot X_s$. Wobec tego macierz A^{-1} powstaje wskutek wykonania na kolumnach macierzy I_n tych samych operacji elementarnych, które zostały wykonane na macierzy A przy sprowadzaniu jej do macierzy I_n.

Rozdział 6

Układy równań liniowych

6.1 Wiadomości wstępne

Definicja 6.1. Układem m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K nazywamy koniunkcję równań postaci:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = b_m \end{cases}, \tag{6.1.1}$$

gdzie $a_{ij}, b_i \in K$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$. Dla tych wszystkich i, j, elementy a_{ij} ciała K nazywamy współczynnikami układu równań $(6.1.1)$, zaś elementy b_i ciała K nazywamy wyrazami wolnymi tego układu.

Definicja 6.2. Układ równań postaci $(6.1.1)$ nazywamy jednorodnym, gdy $b_1 = b_2 = \ldots = b_m = 0$.

Definicja 6.3. Dla układu równań postaci $(6.1.1)$ macierze:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \in \mathrm{M}_{m\times n}(K),\ b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathrm{M}_{m\times 1}(K)$$

$$\text{oraz } A_u = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix} \in \mathrm{M}_{m\times(n+1)}(K)$$

nazywamy odpowiednio: macierzą (współczynników) układu $(6.1.1)$, kolumną wyrazów wolnych i macierzą uzupełnioną tego układu. Macierz A_u zapisuje się często w tzw. postaci blokowej:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right] \in \mathrm{M}_{m\times(n+1)}(K)$$

oznaczanej krótko przez $[A|b]$.

Uwaga 6.1. Wprost z określenia mnożenia macierzy wynika, że układ (6.1.1) można zapisać w tak zwanej postaci macierzowej:

$$Ax = B, \tag{6.1.2}$$

gdzie $x = [\, x_1 \; x_2 \; \ldots \; x_n \,]^T$ oraz A i B są opisane w Definicji 6.3.

Definicja 6.4. Mówimy, że równanie:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b \tag{6.1.3}$$

jest kombinacją liniową równań układu (6.1.1), gdy istnieją takie $c_1, c_2, \ldots, c_m \in K$, że po pomnożeniu stronami i-tego równania układu (6.1.1) przez c_i dla każdego $i \in \{1,2,\ldots,m\}$ oraz dodaniu stronami otrzymanych równań, uzyskamy równanie (6.1.3), czyli gdy:

$$b = \sum_{i=1}^{m} c_i b_i \text{ oraz } a_j = \sum_{i=1}^{m} c_i a_{ij} \text{ dla każdego } j \in \{1,2,\ldots,n\}. \tag{6.1.4}$$

Elementy $c_1, c_2, \ldots, c_m$ ciała K nazywamy wówczas współczynnikami kombinacji liniowej (6.1.3) układu równań liniowych (6.1.1).

Definicja 6.5. Rozwiązaniem układu równań (6.1.1) nazywamy każdy ciąg $(\xi_1, \xi_2, \ldots, \xi_n)$ elementów ciała K taki, że zastąpienie niewiadomych $x_1, x_2, \ldots, x_n$ tego układu odpowiednimi elementami $\xi_1, \xi_2, \ldots, \xi_n$ prowadzi do równości prawdziwych.

Uwaga 6.2. W świetle powyższej definicji otrzymujemy, że ciąg $(\xi_1, \xi_2, \ldots, \xi_n)$ elementów ciała K jest rozwiązaniem układu równań (6.1.1) wtedy i tylko wtedy, gdy $A\xi = B$ dla $\xi = [\, \xi_1 \; \xi_2 \; \ldots \; \xi_n \,]^T$ oraz macierzy A i B określonych w Definicji 6.3.

Wprost z Definicji 6.4 i 6.5 wynika następujące

Stwierdzenie 6.1. Każde rozwiązanie układu równań (6.1.1) jest rozwiązaniem dowolnego równania będącego kombinacją liniową równań tego układu.

Wprowadzimy teraz fundamentalne dla rozważanej teorii pojęcie równoważności układów równań liniowych.

Definicja 6.6. Układy (U_1) i (U_2) równań liniowych z n-niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K nazywamy równoważnymi, co zapisujemy symbolicznie $(U_1) \sim (U_2)$, gdy każde równanie układu (U_1) jest pewną kombinacją liniową równań układu (U_2) i odwrotnie.

Bezpośrednią konsekwencją Definicji 6.6 i Stwierdzenia 6.1 jest poniższe:

Twierdzenie 6.1. Równoważne układy równań liniowych posiadają ten sam zbiór rozwiązań.

Definicja 6.7. Układ równań liniowych z n-niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K nazywamy sprzecznym, gdy równanie $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 1$ jest kombinacją liniową równań tego układu.

Ponieważ $0 \neq 1$ w dowolnym ciele K, to równanie $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 1$ nie posiada rozwiązania w K. Stąd oraz na mocy Definicji 6.7 i Stwierdzenia 6.1 otrzymujemy natychmiast następujące

Stwierdzenie 6.2. Sprzeczny układ równań liniowych nie posiada rozwiązania.

Udowodnimy teraz techniczny

Lemat 6.1. Niech i-te równanie układu równań:

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n = b'_1 \\ a'_{21}x_1 + a'_{22}x_2 + \ldots + a'_{2n}x_n = b'_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{s1}x_1 + a'_{s2}x_2 + \ldots + a'_{sn}x_n = b'_s \end{cases} \tag{6.1.5}$$

nad ciałem K będzie kombinacją liniową równań układu (6.1.1) o współczynnikach $c_{i1}, c_{i2}, \ldots, c_{im}$ dla każdego $i \in \{1, 2, \ldots, s\}$. Jeżeli równanie (6.1.3) jest kombinacją liniową równań układu (6.1.5) o współczynnikach $c_1, c_2, \ldots, c_s$, to równanie (6.1.3) jest kombinacją liniową równań układu (6.1.1) o współczynnikach $\sum_{i=1}^{s} c_i c_{i1}, \sum_{i=1}^{s} c_i c_{i2}, \ldots, \sum_{i=1}^{s} c_i c_{im}$.

Dowód. Z (6.1.4) uzyskujemy natychmiast, że dla każdego $i \in \{1, 2, \ldots, s\}$ zachodzą równości: $a'_{ij} = \sum_{t=1}^{m} c_{it} a_{tj}$ dla każdego $j \in \{1, 2, \ldots, n\}$, oraz $b'_i = \sum_{t=1}^{m} c_{it} b'_t$. Ponadto $a_j = \sum_{i=1}^{s} c_i a'_{ij}$ dla każdego $j \in \{1, 2, \ldots, n\}$, oraz $b = \sum_{i=1}^{s} c_i b'_i$, więc dla każdego $j \in \{1, 2, \ldots, n\}$ otrzymujemy, że:

$$a_j = \sum_{i=1}^{s} \left(c_i \sum_{t=1}^{m} c_{it} a_{tj} \right) = \sum_{i=1}^{s} \sum_{t=1}^{m} c_i c_{it} a_{tj} = \sum_{t=1}^{m} \sum_{i=1}^{s} c_i c_{it} a_{tj} = \sum_{t=1}^{m} \left(\sum_{i=1}^{s} c_i c_{it} \right) a_{tj}$$

oraz

$$b = \sum_{i=1}^{s} \left(c_i \sum_{t=1}^{m} c_{it} b'_t \right) = \sum_{i=1}^{s} \sum_{t=1}^{m} c_i c_{it} b'_t = \sum_{t=1}^{m} \sum_{i=1}^{s} c_i c_{it} b'_t = \sum_{t=1}^{m} \left(\sum_{i=1}^{s} c_i c_{it} \right) b'_t,$$

co w świetle (6.1.4) daje tezę.

Wniosek 6.1. Jeżeli równanie $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = a$, gdzie $a \in K^*$, jest kombinacją liniową równań układu (6.1.1), to układ ten jest sprzeczny.

Dowód. Jako (6.1.3) rozważmy równanie $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 1$. Wówczas (6.1.3) jest kombinacją liniową równania $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = a$ o współczynniku a^{-1}, więc z Definicji 6.7 oraz Lematu 6.1 wynika, że aby otrzymać tezę, w (6.1.5) wystarczy rozważyć układ równań złożony wyłącznie z równania $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = a$.

Ważnymi, bezpośrednimi konsekwencjami Lematu 6.1 są dwa poniższe twierdzenia.

Twierdzenie 6.2. Jeżeli (U_1) i (U_2) są równoważnymi układami równań liniowych nad ciałem K z n niewiadomymi $x_1, x_2, \ldots, x_n$, to układ (U_1) jest sprzeczny wtedy i tylko wtedy, gdy sprzeczny jest układ (U_2).

Twierdzenie 6.3. Niech $(U_1), (U_2)$ oraz (U_3) będą układami równań liniowych nad ciałem K z n niewiadomymi $x_1, x_2, \ldots, x_n$. Wtedy:

(i) $(U_1) \sim (U_1)$;
(ii) jeżeli $(U_1) \sim (U_2)$, to $(U_2) \sim (U_1)$;
(iii) jeżeli $(U_1) \sim (U_2)$ i $(U_2) \sim (U_3)$, to $(U_1) \sim (U_3)$.

6.2 Metody rozwiązywania układów równań liniowych

6.2.1 Operacje elementarne na układzie równań liniowych

Rozwiązanie układu równań liniowych polega na znalezieniu wszystkich rozwiązań tego układu. W tym kontekście pomocne są tzw. operacje elementarne wykonywane na danym układzie m równań linowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ rozważanym nad ciałem K (zob. (6.1.1)). Są to następujące czynności:

(OU.1) Pomnożenie i-tego równania układu przez dowolny niezerowy element a ciała K, oznaczane przez $a \cdot r_i$ (przy wykonywaniu tej operacji zmianie może ulec wyłącznie i-te równanie).
(OU.2) Zamiana miejscami równań o numerach i oraz j dla $j \neq i$. Operację tę oznaczamy symbolicznie przez $r_i \leftrightarrow r_j$ (przy jej wykonywaniu nie modyfikujemy w żaden sposób pozostałych równań układu ani zmieniamy ich kolejności).
(OU.3) Dodanie do i-tego równania układu równania j-tego pomnożonego przez dowolny element a ciała K. Operację tę oznaczamy przez $r_i + a \cdot r_j$ (przy wykonywaniu tej operacji zmianie może ulec wyłącznie i-te równanie).
(OU.4) Wykreślenie powtarzających się kopii pewnego równania (zostawiamy dokładnie jeden egzemplarz powtarzającego się równania).
(OU.5) Wykreślenie równań postaci $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_s = 0$ dla $s \in \{1, 2, \ldots, n\}$ (czyli równań tożsamościowych) w przypadku, gdy $m > 1$ (tj. gdy układ składa się z co najmniej dwóch równań).

(OU.6) Zamiana kolejności niewiadomych x_i oraz x_j dla $j \neq i$ wraz ze stojącymi przy nich współczynnikami w każdym równaniu rozważanego układu, oznaczana symbolicznie jako $x_i \leftrightarrow x_j$ (przy zastosowaniu tej operacji s-te równanie:

$$a_{s1}x_1 + \ldots + a_{si}x_i + \ldots + a_{sj}x_j + \ldots + a_{sn}x_n = b_s$$

rozważanego układu przyjmuje postać:

$$a_{s1}x_1 + \ldots + a_{sj}x_j + \ldots + a_{si}x_i + \ldots + a_{sn}x_n = b_s$$

dla każdego $s \in \{1,2,\ldots,m\}$).

Wniosek 6.2. Z Definicji 6.6 wynika, że jeśli układ (U') równań liniowych powstaje z układu (U) wskutek wykonania którejkolwiek operacji elementarnej $(OU.1) - (OU.6)$, to $(U') \sim (U)$.

Z Wniosku 6.2 oraz Twierdzeń 6.1, 6.2 i 6.3 przez prostą indukcję wynika następujące

Twierdzenie 6.4. Jeżeli układ równań (U') równań liniowych powstaje z układu (U) wskutek wykonania skończonej liczby operacji elementarnych, to $(U') \sim (U)$. W szczególności układy (U') i (U) mają wówczas ten sam zbiór rozwiązań oraz układ (U') jest sprzeczny wtedy i tylko wtedy, gdy sprzeczny jest układ (U).

Twierdzenie 6.5. Układ $(6.1.1)$ m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K, w którym $a_{ij} \neq 0$ dla pewnych $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$, po ewentualnej permutacji nazw niewiadomych $x_1, x_2, \ldots x_n$ równoważny jest układowi równań:

$$\begin{cases} x_1 & & & + c_{1\,s+1}x_{s+1} & + c_{1n}x_n = & d_1 \\ & x_2 & & + c_{2\,s+1}x_{s+1} & + c_{2n}x_n = & d_2 \\ & & x_3 & + c_{3\,s+1}x_{s+1} & + c_{3n}x_n = & d_3 \\ & & \ddots & \vdots \quad \vdots & \vdots \quad \vdots \quad \vdots \quad \vdots & \\ & & & x_s + c_{3s\,s+1}x_{s+1} & + c_{sn}x_n = & d_s \\ & & & & 0 = & d_{s+1} \end{cases}, \tag{6.2.1}$$

gdzie $s \leq n$.

Dowód. Załóżmy, że $(i,j) \neq (1,1)$. Wówczas na układzie $(6.1.1)$ wykonujemy kolejno operacje elementarne $x_1 \leftrightarrow x_j$, $r_1 \leftrightarrow r_i$ oraz $a_{ij}^{-1} \cdot r_1$, a następnie permutujemy nazwy niewiadomych za pomocą transpozycji $(1,j)$ (tj. przemianowujemy niewiadomą x_j na x_1, zaś niewiadomą x_1 na x_j). W ten sposób otrzymujemy układ postaci:

$$\begin{cases} x_1 + a'_{12}x_2 + a'_{13}x_3 + \ldots + a'_{1n}x_n = b'_1 \\ a'_{21}x_1 + a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n = b'_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \ddots \vdots \quad \vdots \quad \vdots \quad \vdots \\ a'_{m1}x_1 + a'_{m2}x_2 + a'_{m3}x_3 + \ldots + a'_{mn}x_n = b'_m \end{cases}. \tag{6.2.2}$$

Jeśli $i = 1$ oraz $j \neq 1$, to pomijamy operację $r_1 \leftrightarrow r_i$. Jeżeli $i \neq 1$ oraz $j = 1$ to nie wykonujemy operacji $x_1 \leftrightarrow x_j$ ani nie permutujemy nazw niewiadomych. Jeśli zaś $i = j = 1$, to w celu otrzymania układu (6.2.2) wykonujemy na układzie (6.1.1) wyłącznie operację $a_{ij}^{-1} \cdot r_1$. Następnie na układzie (6.2.2) wykonujemy kolejno następujące operacje elementarne: $r_2 - a'_{21} \cdot r_1, r_3 - a'_{31} \cdot r_1, \ldots, r_m - a'_{m1} \cdot r_1$. W ten sposób wyeliminujemy niewiadomą x_1 z równań $r_2, r_3, \ldots, r_m$, otrzymując układ postaci:

$$\begin{cases} x_1 + a''_{12}x_2 + a''_{13}x_3 + \ldots + a''_{1n}x_n = b''_1 \\ a''_{22}x_2 + a''_{23}x_3 + \ldots + a''_{2n}x_n = b''_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \ddots \vdots \quad \vdots \quad \vdots \quad \vdots \\ a''_{m2}x_2 + a''_{m3}x_3 + \ldots + a''_{mn}x_n = b''_m \end{cases}. \tag{6.2.3}$$

Jeżeli istnieją $i \in \{2, 3, \ldots, m\}$ oraz $j \in \{2, 3, \ldots, n\}$ takie, że $a'_{ij} \neq 0$, to powtarzamy opisaną wyżej procedurę dla układu:

$$\begin{cases} a''_{22}x_2 + a''_{23}x_3 + \ldots + a''_{2n}x_n = b''_2 \\ a''_{32}x_2 + a''_{33}x_3 + \ldots + a''_{3n}x_n = b''_3 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \ddots \vdots \quad \vdots \quad \vdots \quad \vdots \\ a''_{m2}x_2 + a''_{m3}x_3 + \ldots + a''_{mn}x_n = b''_m \end{cases} \tag{6.2.4}$$

i zastępujemy równania $r_2, r_3, \ldots, r_m$ układu (6.2.3) kolejno równaniami $r_1, r_2, \ldots, r_{n-1}$ nowo otrzymanego układu równoważnego układowi (6.2.4). Po wykonaniu skończenie wielu takich kroków otrzymamy układ postaci:

$$\begin{cases} x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \ldots + \alpha_{1s}x_s + \alpha_{1\,s+1}x_{s+1} + \ldots + \alpha_{1n}x_n = \beta_1 \\ \alpha_{22}x_2 + \alpha_{23}x_3 + \ldots + \alpha_{2s}x_s + \alpha_{2\,s+1}x_{s+1} + \ldots + \alpha_{2n}x_n = \beta_2 \\ \alpha_{33}x_3 + \ldots + \alpha_{3s}x_s + \alpha_{3\,s+1}x_{s+1} + \ldots + \alpha_{3n}x_n = \beta_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_s + \alpha_{s\,s+1}x_{s+1} + \ldots + \alpha_{sn}x_n = \beta_s \\ 0 = \beta_{s+1} \end{cases}, \tag{6.2.5}$$

gdzie $s \leq n$. Wykonując kolejno operacje elementarne: $r_1 - \alpha_{1s} \cdot r_s, r_2 - \alpha_{2s} \cdot r_s, \ldots, r_{s-1} - \alpha_{s-1\,s} \cdot r_s$ na układzie (6.2.5) wyeliminujemy niewiadomą x_s z $s-1$ pierwszych równań tego układu. Następnie przy pomocy równania r_{s-1} w ten sam sposób eliminujemy niewiadomą x_{s-1} z każdego spośród równań $r_1, r_2, \ldots, r_{s-2}$. Powtarzając tę procedurę skończoną liczbę razy uzyskamy więc układ postaci (6.2.1). Równoważność układów (6.1.1)i (6.2.1)jest konsekwencją Wniosku 6.2.

6.2.2 Metoda eliminacji Gaussa

Metoda eliminacji Gaussa jest metodą rozwiązywania układów równań liniowych nad dowolnym ciałem opartą na Twierdzeniu 6.5 i jego dowodzie. Mianowicie, po ewentualnej permutacji nazw niewiadomych układu (6.1.1), w którym $a_{ij} \neq 0$ dla pewnych $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$, i sprowadzeniu go przy pomocy operacji elementarnych do postaci (6.2.1) zauważamy, że mamy do rozważenia następujące przypadki:

(i) $d_{s+1} \neq 0$. Wtedy układ (6.2.1) jest sprzeczny, więc z Twierdzenia 6.4 wynika, że sprzeczny jest także wyjściowy układ (6.1.1). Stwierdzenie 6.2 implikuje więc, że wówczas układ (6.1.1) nie posiada rozwiązania.

(ii) $d_{s+1} = 0$ i $s = n$. Wówczas układ (6.1.1) posiada dokładnie jedno rozwiązanie, którym jest $(d_1, d_2, \ldots, d_n)$. Piszemy wówczas:

$$\begin{cases} x_1 = d_1 \\ x_2 = d_2 \\ \vdots \\ x_n = d_n \end{cases}.$$

(iii) $d_{s+1} = 0$ oraz $s < n$. Wszystkimi rozwiązaniami układu (6.2.1) są wówczas ciągi $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ takie, że $\alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_n$ są dowolnymi elementami ciała K, zaś dla każdego $i \in \{1,2,\ldots,s\}$, $\alpha_i = d_i - c_{ik+1}\alpha_{k+1} - c_{ik+2}\alpha_{k+2} - \ldots - c_{in}\alpha_n$. Piszemy wtedy:

$$\begin{cases} x_1 = d_1 - c_{1k+1}\alpha_{k+1} - c_{1k+2}\alpha_{k+2} - \ldots - c_{1n}\alpha_n \\ x_2 = d_2 - c_{2k+1}\alpha_{k+1} - c_{2k+2}\alpha_{k+2} - \ldots - c_{2n}\alpha_n \\ \vdots \\ x_s = d_1 - c_{sk+1}\alpha_{k+1} - c_{sk+2}\alpha_{k+2} - \ldots - c_{sn}\alpha_n \\ x_{s+1} = \alpha_{s+1} \\ x_{s+2} = \alpha_{s+2} \\ \vdots \\ x_n = \alpha_n. \end{cases} \text{, gdzie } \alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_n \in K$$

W szczególności, układ (6.1.1) posiada wówczas więcej niż jedno rozwiązanie. Dowolne elementy $\alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_n$ ciała K nazywamy parametrami. Używana jest również uproszczona wersja powyższego zapisu:

$$\begin{cases} x_1 = d_1 - c_{1\,k+1}x_{k+1} - c_{1\,k+2}x_{k+2} - \ldots - c_{1\,n}x_n \\ x_2 = d_2 - c_{2\,k+1}x_{k+1} - c_{2\,k+2}x_{k+2} - \ldots - c_{2\,n}x_n \\ \vdots \\ x_s = d_1 - c_{s\,k+1}x_{k+1} - c_{s\,k+2}x_{k+2} - \ldots - c_{s\,n}x_n \\ x_{s+1}, x_{s+2}, \ldots, x_n \in K \end{cases}.$$

Zauważmy, że z rozumowania przedstawionego w dowodzie Twierdzenia 6.5 wynika, że przy zastosowaniu metody eliminacji Gaussa liczba równań układu (6.2.1) nie jest większa niż liczba równań układu (6.1.1). Otrzymujemy stąd następujący

Wniosek 6.3. Jeżeli liczba m równań układu (6.1.1) jest większa od liczby n niewiadomych, to układ ten nie może mieć dokładnie jednego rozwiązania.

Uwaga 6.3. Jeżeli układ (6.1.1) nie posiada rozwiązania, to z Twierdzenia 6.4 wynika, że rozwiązania nie ma także układ (6.2.1), skąd $d_{s+1} \neq 0$. Zatem układ (6.2.1) jest wówczas sprzeczny. Powołując się ponownie na Twierdzenie 6.4 otrzymujemy więc, że również układ (6.1.1) jest sprzeczny.

Bezpośrednią konsekwencją powyższej uwagi oraz Stwierdzenia 6.2 jest następujące

Twierdzenie 6.6. Układ równań liniowych jest sprzeczny wtedy i tylko wtedy, gdy nie posiada on rozwiązania.

Uwaga 6.4. Ponieważ i-te równanie układu (6.1.1) wzajemnie jednoznacznie odpowiada i-temu wierszowi macierzy uzupełnionej A_u tego układu, to pewnym operacjom elementarnym na równaniach układu (6.1.1) wzajemnie jednoznacznie odpowiadają pewne operacje elementarne na macierzy A_u uzupełnionej tego układu (zob. Definicja 4.5). Mianowicie:

(i) operacja $(OU.1)$ odpowiada wierszowemu wariantowi operacji $(OM.1)$;
(ii) operacja $(OU.2)$ odpowiada wierszowemu wariantowi operacji $(OM.2)$;
(iii) operacja $(OU.3)$ odpowiada wierszowemu wariantowi operacji $(OM.3)$;
(iv) operacja $(OU.6)$ odpowiada kolumnowemu wariantowi operacji $(OM.2)$.

Ponadto operacje $(OU.4)$ i $(OM.5)$ można w naturalny sposób zaadaptować do rozważań związanych z macierzą A_u. Operacja $(OU.4)$ odpowiada wtedy wykreśleniu powtarzających się kopii pewnego wiersza macierzy A_u, zaś operacja $(OU.5)$ dopowiada wykreśleniu zerowych wierszy tej macierzy.

Powyższe utożsamienie układu równań liniowych z macierzą uzupełnioną tego układu pozwala nieco zaoszczędzić czas potrzebny na zapis rachunków związanych z rozwiązywaniem rozważanego układu metodą eliminacji Gaussa. Poszczególne etapy rozwiązania będziemy oddzielali symbolem $\sim$, nad którym będziemy zapisywali wykonywane operacje elementarne. Wybór takiego oznaczenia jest naturalny w kontekście Uwagi 6.4 oraz Twierdzenia 6.4 i Definicji 6.6.

Przykład 6.1. Stosując metodę eliminacji Gaussa oraz obserwacje odnotowane w ramach Uwagi 6.4, rozwiążemy nad ciałem $\mathbb{R}$ układ równań:

$$\begin{cases} 3x+2y+z=0 \\ 2x+2y+4z+2t=2 \\ x+y+2z+t=1 \\ 7x+5y+4z+t=1 \end{cases}. \tag{6.2.6}$$

Macierzą uzupełnioną tego układu jest:

$$A_u = \left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 0 \\ 2 & 2 & 4 & 2 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 7 & 5 & 4 & 1 & 1 \end{array}\right].$$

Mamy więc:

$$\left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 0 \\ 2 & 2 & 4 & 2 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 7 & 5 & 4 & 1 & 1 \end{array}\right] \begin{array}{l} w_2-2w_3 \\ w_4-2w_1 \end{array} \sim \left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{array}\right] \sim \left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 0 \\ 1 & 1 & 2 & 1 & 1 \end{array}\right]$$

$$\overset{w_1\leftrightarrow w_2}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 1 \\ 3 & 2 & 1 & 0 & 0 \end{array}\right] \overset{w_2-3w_1}{\sim} \left[\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 1 \\ 0 & -1 & -5 & -3 & -3 \end{array}\right] \overset{w_1+w_2}{\sim}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & -3 & -2 & -2 \\ 0 & -1 & -5 & -3 & -3 \end{array}\right] \overset{(-1)\cdot w_2}{\sim} \left[\begin{array}{cccc|c} 1 & 0 & -3 & -2 & -2 \\ 0 & 1 & 5 & 3 & 3 \end{array}\right].$$

Zatem rozwiązanie układu (6.2.6) opisane jest następującym układem warunków:

$$\begin{cases} x=3z+2t-2 \\ y=-5z-3t+3 \\ z,t\in\mathbb{R} \end{cases}.$$

W szczególności układ (6.2.6) posiada nieskończenie wiele rozwiązań zależnych od dwóch parametrów rzeczywistych z i t.

6.2.3 Wzory Cramera, metoda wyznacznikowa

Układ równań liniowych nad dowolnym ciałem, w którym liczba równań równa jest liczbie niewiadomych oraz wyznacznik macierzy tego układu jest niezerowy, można rozwiązać, stosując tzw. wzory Cramera, które szczegółowo omówimy w ostatniej części niniejszego rozdziału. W tym celu załóżmy, że $m = n$ dla układu (6.1.1). Wtedy układ ten przyjmuje postać:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n = b_n \end{cases}, \tag{6.2.7}$$

gdzie $a_{ij}, b_i \in K$ dla wszystkich $i, j \in \{1, 2, \ldots, n\}$. Dla każdego $i \in \{1, 2, \ldots, n\}$, symbolem A_i oznaczmy macierz powstałą z macierzy:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \in \mathrm{M}_n(K)$$

układu (6.2.7) w wyniku zastąpienia i-tej kolumny macierzy A kolumną wyrazów wolnych $\begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}^T$ tego układu. Mamy więc:

$$A_1 = \begin{bmatrix} b_1 & a_{12} & \ldots & a_{1n} \\ b_2 & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \ldots & a_{nn} \end{bmatrix}, A_2 = \begin{bmatrix} a_{11} & b_1 & \ldots & a_{1n} \\ a_{21} & b_2 & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n a_{n1} & b_2 & \ldots & a_{nn} \end{bmatrix}, \ldots,$$

$$A_n = \begin{bmatrix} a_{11} & a_{12} & \ldots & b_1 \\ a_{21} & a_{22} & \ldots & b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & b_n \end{bmatrix}.$$

Definicja 6.8. Wyznacznikiem głównym układu (6.2.7) nazywamy wyznacznik macierzy A tego układu i oznaczamy go literą W. Dla każdego $i \in \{1, 2, \ldots, n\}$, symbolem W_i oznaczać będziemy wyznacznik macierzy A_i, który nazywamy wyznacznikiem stowarzyszonym z niewiadomą x_i (mówi się też nieformalnie: wyznacznik ze względu na x_i).

Poniższe twierdzenie opisuje tzw. metodę wyznacznikową rozwiązywania układów n równań liniowych z n niewiadomymi nad dowolnym ciałem, związaną ze wzorami Cramera.

Twierdzenie 6.7 (Cramer). Jeżeli wyznacznik główny W układu (6.2.7) jest różny od zera, to układ ten posiada dokładnie jedno rozwiązanie i jest ono dane wzorami Cramera:

$$x_1 = \frac{W_1}{W},\ x_2 = \frac{W_2}{W},\ \ldots,\ x_n = \frac{W_n}{W}. \tag{6.2.8}$$

Jeżeli $W = 0$ oraz $W_i \neq 0$ dla pewnego $i \in \{1,2,\ldots,n\}$, to układ (6.2.7) nie posiada rozwiązania.

Dowód. Weźmy dowolne $i \in \{1,2,\ldots,n\}$. Mnożąc j-te równanie (6.2.7) przez dopełnienie algebraiczne wyrazu a_{ji} macierzy współczynników tego układu (zob. Definicja 4.7) dla każdego $j \in \{1,2,\ldots,n\}$ uzyskujemy układ równań:

$$\begin{cases} \sum_{t=1}^{n} a_{1t} \cdot (-1)^{1+i} \cdot \det(A_{1i}) \cdot x_t = (-1)^{1+i} \cdot b_1 \cdot \det(A_{1i}) \\ \sum_{t=1}^{n} a_{2t} \cdot (-1)^{2+i} \cdot \det(A_{2i}) \cdot x_t = (-1)^{2+i} \cdot b_2 \cdot \det(A_{2i}) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{t=1}^{n} a_{nt} \cdot (-1)^{n+i} \cdot \det(A_{ni}) \cdot tx_{=}(-1)^{n+i} \cdot b_n \cdot \det(A_{ni}) \end{cases}.$$

Po dodaniu stronami wszystkich równań powyższego układu otrzymujemy równanie:

$$\sum_{k=1}^{n} \sum_{t=1}^{n} a_{kt} \cdot (-1)^{k+i} \cdot \det(A_{ki}) \cdot x_t = \sum_{k=1}^{n} (-1)^{k+i} \cdot b_k \cdot \det(A_{ki}). \tag{6.2.9}$$

Dalej, na mocy punktu (ii) Wniosku 4.5 otrzymujemy, że:

$$\sum_{t=1}^{n} a_{kt} \cdot (-1)^{k+i} \cdot \det(A_{ki}) = 0 \text{ dla każdego } t \in \{1,2,\ldots,n\} \setminus \{i\},$$

więc równanie (6.2.9) równoważne jest równaniu:

$$\sum_{k=1}^{n} (-1)^{k+i} \cdot a_{ki} \cdot \det(A_{ki}) \cdot x_i = \sum_{k=1}^{n} (-1)^{k+i} \cdot b_k \cdot \det(A_{ki}). \tag{6.2.10}$$

Ponadto ze wzoru (4.4.3) oraz Definicji 6.8 wynika, że:

$$\sum_{k=1}^{n} (-1)^{k+i} \cdot b_k \cdot \det(A_{ki}) = W_i \tag{6.2.11}$$

oraz

$$\sum_{k=1}^{n} (-1)^{k+i} \cdot a_{ki} \cdot \det(A_{ki}) = W. \tag{6.2.12}$$

Podstawiając (6.2.12) i (6.2.11) do (6.2.10) uzyskujemy równanie:

$$W \cdot x_i = W_i. \tag{6.2.13}$$

Stąd oraz na mocy Stwierdzenia 6.1 i dowolności wyboru $i \in \{1,2,\ldots,n\}$ wnioskujemy, że jeśli ciąg $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ elementów ciała K jest rozwiązaniem układu (6.2.7):

$$W \cdot \alpha_i = W_i \text{dla każdego } i \in \{1,2,\ldots,n\}. \tag{6.2.14}$$

Niemożliwa jest więc wówczas sytuacja, w której $W = 0$ oraz $W_i \neq 0$ dla pewnego $i \in \{1,2,\ldots,n\}$. Zatem jeżeli $W = 0$ oraz istnieje $i \in \{1,2,\ldots,n\}$ takie, że $W_i \neq 0$, to układ (6.2.7) nie posiada rozwiązania. Jeśli natomiast $W \neq 0$, to z (6.2.14) wynika, że $\alpha_i = \frac{W_i}{W}$ dla każdego $i \in \{1,2,\ldots,n\}$. W ten sposób pokazaliśmy, że jeżeli $W \neq 0$, to układ równań (6.2.7) może posiadać co najwyżej jedno rozwiązanie i jest ono wówczas opisane wzorami (6.2.8). Pozostało udowodnić, że warunek $W \neq 0$ implikuje, że ciąg $\left(\frac{W_1}{W}, \frac{W_2}{W}, \ldots, \frac{W_n}{W}\right)$ jest rozwiązaniem układu (6.2.7). Z (6.2.11) wynika, że dla każdego $i \in \{1,2,\ldots,n\}$ jest:

$$\sum_{j=1}^{n} a_{ij} \frac{W_j}{W} = W^{-1} \cdot \sum_{j=1}^{n} a_{ij} \cdot W_j = W^{-1} \cdot \sum_{j=1}^{n} \left(a_{ij} \cdot \sum_{k=1}^{n} b_k \cdot (-1)^{k+j} \cdot \det(A_{kj}) \right) =$$

$$W^{-1} \cdot \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ij} \cdot b_k \cdot (-1)^{k+j} \cdot \det(A_{kj}) = W^{-1} \cdot \sum_{k=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot b_k \cdot (-1)^{k+j} \cdot \det(A_{kj})$$

$$= W^{-1} \cdot \sum_{k=1}^{n} \left(b_k \cdot \sum_{j=1}^{n} a_{ij} \cdot (-1)^{k+j} \cdot \det(A_{kj}) \right).$$

Ponadto, odpowiednio na mocy wzoru (4.4.4) oraz punktu (i) Wniosku 4.5 uzyskujemy, że $\sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} \cdot \det(A_{ij}) = W$ i $\sum_{j=1}^{n} a_{ij} \cdot (-1)^{k+j} \cdot \det(A_{kj}) = 0$ dla $k \neq i$. Wobec tego:

$$\sum_{j=1}^{n} a_{ij} x_j = W^{-1} \cdot b_i \cdot W = b_i \text{ dla każdego } i \in \{1,2,\ldots,n\}.$$

Zatem wzory (6.2.8) rzeczywiście opisują rozwiązanie układu (6.2.7).

Uwaga 6.5. Twierdzenie Cramera nic nie mówi o przypadku, w którym $W = W_1 = W_2 = \ldots = W_n = 0$. Układ (6.2.7) może wówczas nie mieć rozwiązania lub posiadać więcej niż jedno rozwiązanie. Aby to zweryfikować, należy zastosować inną metodę rozwiązywania układu równań liniowych, np. omówioną wcześniej metodę eliminacji Gaussa.

Definicja 6.9. Układ równań postaci (6.2.7), w którym $W \neq 0$ nazywamy układem Cramera.

W świetle Wniosku 4.4, Twierdzenie Cramera implikuje natychmiast następujący

Wniosek 6.4. Jednorodny układ Cramera z n niewiadomymi nad ciałem K posiada dokładnie jedno rozwiązanie $(0,0,\ldots,0) \in K^n$.

Przykład 6.2. Stosując wzory Cramera rozwiążemy nad ciałem $\mathbb{R}$ układ równań:

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 14 \\ 4x_1 + 3x_2 - x_3 = 7 \\ x_1 - x_2 + x_3 = 2 \end{cases}. \tag{6.2.15}$$

Ponieważ $W = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 3 & -1 \\ 1 & -1 & 1 \end{vmatrix} = 3 - 12 - 2 - (9 + 8 + 1) = -11 - 18 = -29 \neq 0$, to z Twierdzenia Cramera wynika, że układ $(6.2.15)$ posiada dokładnie jedno rozwiązanie i jest ono dane wzorami Cramera $(6.2.8)$. Dalej,

$$W_1 = \begin{vmatrix} 14 & 2 & 3 \\ 7 & 3 & -1 \\ 2 & -1 & 1 \end{vmatrix} = 42 - 21 - 4 - (18 + 14 + 14) = 17 - 18 - 28 = -29,$$

$$W_2 = \begin{vmatrix} 1 & 14 & 3 \\ 4 & 7 & -1 \\ 1 & 2 & 1 \end{vmatrix} = 7 + 24 - 14 - (21 + 56 - 2) = 17 - 75 = -58,$$

$$W_3 = \begin{vmatrix} 1 & 2 & 14 \\ 4 & 3 & 7 \\ 1 & -1 & 2 \end{vmatrix} = 6 - 4 \cdot 14 + 14 - (3 \cdot 14 + 16 - 7) = 6 - 6 \cdot 14 - 16 + 7 = -3 - 84 = -87.$$

Zatem:

$$\begin{cases} x_1 = \frac{-29}{-29} = 1 \\ x_2 = \frac{-58}{-29} = 2 \\ x_3 = \frac{-87}{-29} = 3 \end{cases}.$$

Uwaga 6.6. Zauważmy, że w powyższym przykładzie, po wyznaczeniu W, w celu znalezienia rozwiązania układu $(6.2.15)$ wystarczy policzyć tylko dwa spośród wyznaczników W_1, W_2 i W_3, zastosować związane z nimi wzory Cramera $(6.2.8)$, a następnie ostatnią niewiadomą wyznaczyć z dowolnego równania układu $(6.2.15)$, podstawiając wyznaczone wcześniej niewiadome.

Rozdział 7

Przestrzenie liniowe

7.1 Pojęcie przestrzeni liniowej

Definicja 7.1. Lewostronnym dwuargumentowym działaniem zewnętrznym w niepustym zbiorze X nad niepustym zbiorem Y nazywamy każdą funkcję $\circ\colon Y \times X \to X$.

Uwaga 7.1. Dla wszystkich $x \in X$ i $y \in Y$ będziemy pisali $y \circ x$ zamiast $\circ\big((y,x)\big)$.

Przykład 7.1. Niech $p \in \mathbb{P}$ i niech $m \leq n$ będą liczbami naturalnymi. Wówczas funkcja $\circ\colon \mathbb{Z}_{p^n} \times \mathbb{Z}_{p^m} \to \mathbb{Z}_{p^m}$ określona za pomocą wzoru:

$$y \circ x = \underbrace{x \oplus_{p^m} x \oplus_{p^m} \ldots \oplus_{p^m} x}_{y}$$

dla wszystkich $y \in \mathbb{Z}_{p^n}$ i $x \in \mathbb{Z}_{p^m}$, jest lewostronnym dwuargumentowym działaniem zewnętrznym w zbiorze $\mathbb{Z}_{p^m}$ nad zbiorem $\mathbb{Z}_{p^n}$.

W ramach poniższej definicji wprowadzimy najbardziej fundamentalne pojęcie Algebry liniowej.

Definicja 7.2. Mówimy, że grupa abelowa $(V, \boxplus, \Theta)$ jest przestrzenią liniową nad ciałem K względem działania $\circ\colon K \times V \to V$, gdy dla wszystkich $\lambda, \mu \in K$ oraz $v, w \in V$ spełniony jest układ warunków:

(L1) $(\lambda + \mu) \circ v = \lambda \circ v \boxplus \mu \circ v$;
(L2) $\lambda \circ (v \boxplus w) = \lambda \circ v \boxplus \lambda \circ w$;
(L3) $(\lambda \cdot \mu) \circ v = \lambda \circ (\mu \circ v)$;
(L4) $1 \circ v = v$.

Elementy zbioru V nazywamy wówczas wektorami, zaś elementy ciała K – skalarami. Ponadto lewostronne dwuargumentowe działanie zewnętrzne $\circ$ w zbiorze V nad ciałem K nazywamy (lewostronnym) mnożeniem przez skalary.

Uwaga 7.2. Symbole $+$ i $\cdot$ w powyższej definicji oznaczają odpowiednio dodawanie i mnożenie w ciele K. Ponadto 1 oznacza jedynkę ciała K, czyli element neutralny mnożenia w K.

Uwaga 7.3. Napis „$\lambda \circ v \boxplus \mu \circ v$" interpretujemy jako „$(\lambda \circ v) \boxplus (\mu \circ v)$" – jak zwykle przyjmujemy umowę, że działania „podobne" do kropki wiążą silniej niż działania „podobne"do plusa (lub zawierające w sobie symbol $+$).

Uwaga 7.4. Przestrzenie liniowe nad ciałem K nazywa się też przestrzeniami wektorowymi nad ciałem K albo krótko: K-przestrzeniami (gdy z kontekstu jasno wynika, że K jest ciałem).

Uwaga 7.5. Jeżeli V jest przestrzenią liniową nad ciałem K, to równości dane w punktach $(L1)$-$(L4)$ Definicji 7.2 zachodzące dla wszystkich $\lambda, \mu \in K$ oraz $v, w \in V$ nazywamy kolejno: rozdzielnością mnożenia zewnętrznego względem dodawania w ciele K, rozdzielnością mnożenia zewnętrznego względem dodawania w przestrzeni liniowej V, łącznością mieszaną (jest określenie nieformalne, ale oddające istotę sprawy) oraz unitarnością.

Uwaga 7.6. Mówiąc, że V jest przestrzenią liniową nad ciałem K mamy na myśli, że: V jest niepustym zbiorem z wyróżnionym elementem Θ, istnieje działanie $\boxplus\colon V \times V \to V$ takie, że system algebraiczny $(V, \boxplus, \Theta)$ jest grupą abelową oraz istnieje działanie $\circ\colon K \times V \to V$ takie, że spełnione są warunki $(L1)$.-$(L4)$. Definicji 7.2. W sytuacji, gdy dany zbiór rozważa się w kontekście teorii przestrzeni liniowych wraz ze standardowymi dla niego działaniami, zazwyczaj nie opisuje się tych działań. Analogiczna uwaga odnosi się do elementu wyróżnionego Θ.

7.2 Przykłady przestrzeni liniowych

Podamy teraz kilka najbardziej typowych przykładów przestrzeni liniowych.

Przykład 7.2. Z omówionych w Rozdziale 5. własności operacji na macierzach wynika, że dla dowolnych $m, n \in \mathbb{N}$ oraz dowolnego ciała K, zbiór $\mathrm{M}_{m\times n}(K)$ wszystkich $(m \times n)$-macierzy nad ciałem K jest przestrzenią liniową nad K (oczywiście domyślnymi działaniami są tu standardowe dodawanie macierzy i standardowe mnożenie macierzy z lewej strony przez skalar; natomiast domyślnie rozważanym elementem wyróżnionym jest $m \times n$-macierz zerowa).

Przykład 7.3. Z określenia i własności działań na wielomianach (zob. Rozdział 3.) wynika, że dla dowolnego ciała K, zbiór $K[x]$ wszystkich wielomianów zmiennej x nad ciałem K jest K-przestrzenią.

W kontekście naszych dalszych rozważań szczególnie istotny okaże się poniższy

Przykład 7.4. Niech K będzie ciałem, niech $n \in \mathbb{N}$ i niech:

$$K^n = \big\{[x_1, x_2, \ldots, x_n] \colon x_i \in K \text{ dla każdego } i \in \{1, 2, \ldots, n\}\big\}.$$

Wtedy zbiór K^n rozważany wraz z działaniami $\boxplus\colon K^n \times K^n \to K^n$ i $\circ\colon K \times K^n \to K^n$ określonymi punktowo, tzn. za pomocą wzorów:

$$[a_1, a_2, \ldots, a_n] \boxplus [b_1, b_2, \ldots, b_n] = [a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n]$$

oraz

$$a \circ [a_1, a_2, \ldots, a_n] = [a \cdot a_1, a \cdot a_2, \ldots, a \cdot a_n],$$

jest K-przestrzenią.

Przykład 7.5. Dla dowolnego ciała K, zbiór:

$$K^{\infty} = \left\{ [x_1, x_2, \ldots] \colon x_i \in K \text{ dla każdego } i \in \mathbb{N} \right\}$$

rozważany wraz z działaniami $\boxplus$ i $\circ$ określonymi punktowo jest K-przestrzenią liniową.

Przykład 7.6. Zbiór rozwiązań jednorodnego układu (U) m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K jest przestrzenią liniową nad K. Ponieważ n-elementowy $(0, 0 \ldots, 0)$ ciąg elementów ciała K jest rozwiązaniem układu (U), to zbiór Ω wszystkich rozwiązań tego układu można utożsamić z niepustym podzbiorem przestrzeni liniowej K^n opisanej w Przykładzie 7.4. Przy takim utożsamieniu rozważmy dowolne $[\alpha_1, \alpha_2, \ldots, \alpha_n], [\beta_1, \beta_2, \ldots, \beta_n] \in \Omega$ i $\lambda \in K$. Niech $i \in \{1, 2, \ldots, m\}$ i niech $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0$ będzie i-tym równaniem układu (U). Wówczas $a_{i1}(\alpha_1 + \beta_1) + a_{i2}(\alpha_2 + \beta_2) + \ldots + a_{in}(\alpha_n + \beta_n) = (a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n) + (a_{i1}\beta_1 + a_{i2}\beta_2 + \ldots + a_{in}\beta_n) = 0 + 0 = 0$ i $a_{i1}(\lambda\alpha_1) + a_{i2}(\lambda\alpha_2) + \ldots + a_{in}(\lambda\alpha_n) = \lambda\,(a_{i1}\alpha_1 + a_{i2}\alpha_2 + \ldots + a_{in}\alpha_n) = \lambda \cdot 0 = 0$. Wobec tego $[\alpha_1, \alpha_2, \ldots, \alpha_n] \boxplus [\beta_1, \beta_2, \ldots, \beta_n] \in \Omega$ i $\lambda \circ [\alpha_1, \alpha_2, \ldots, \alpha_n] \in \Omega$. Zatem zbiór Ω jest przestrzenią liniową nad ciałem ze względu na działania $\boxplus$ oraz $\circ$ opisane w Przykładzie 7.4.

Przykład 7.7. Zbiór wszystkich równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K rozważany wraz ze standardowym dodawaniem równań stronami, standardową operacją mnożenia równań z lewej strony przez skalary oraz zerem Θ określonym jako równanie tożsamościowe $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$ jest przestrzenią wektorową nad ciałem K.

Przykład 7.8. Grupa trywialna $\left(\{a\}, \boxplus, a\right)$ jest przestrzenią liniową nad dowolnym ciałem względem działania $\circ \colon K \times \{a\} \to \{a\}$ danego wzorem $\lambda \circ a = a$ dla każdego $\lambda \in K$. Taką K-przestrzeń nazywamy zerową lub trywialną.

Przykład 7.9. Podzbiór K ciała L nazywamy podciałem ciała L wówczas, gdy $0, 1$ należą do K oraz K jest ciałem ze względu na wszystkie działania określone w L ograniczone do K. Jeśli K jest podciałem ciała L, to L jest w naturalny sposób K-przestrzenią ($K \times L \ni (\lambda, v) \overset{\circ}{\longmapsto} \lambda \cdot v \in L$, gdzie $\cdot$ oznacza mnożenie w ciele L). Na przykład, $\mathbb{R}$ jest przestrzenią liniową nad $\mathbb{Q}$. W szczególności każde ciało L jest w naturalny sposób przestrzenią liniową nad samym sobą.

Przykład 7.10. Niech X będzie dowolnym niepustym zbiorem, zaś K – dowolnym ciałem. Wówczas zbiór K^X wszystkich funkcji przekształcających zbiór X w ciało K

rozważany wraz z działaniami $\boxplus\colon K^X \times K^X \to K^X$ i $\circ\colon K \times K^X \to K^X$ określonymi punktowo, tzn. za pomocą wzorów:

$$(f \boxplus g)(x) = f(x) + g(x) \text{ dla każdego } x \in X$$

oraz

$$(\lambda \circ f)(x) = \lambda \cdot f(x) \text{ dla każdego } x \in X,$$

jest przestrzenią liniową nad ciałem K. Aby się o tym przekonać, rozważmy dowolne $f, g, h \in K^X$ oraz $\lambda, \mu \in K$. Wprost z określenia działań $\boxplus$ i $\circ$ wynika, że są one zdefiniowane poprawnie (tzn., że $f \boxplus g \in K^X$ oraz $\lambda \circ f \in K^X$). Dalej, dla dowolnego $x \in X$ uzyskujemy, że $(g \boxplus f)(x) = g(x) + f(x) = f(x) + g(x) = (f \boxplus g)(x)$ (skorzystaliśmy tu z przemienności dodawania $+\colon K \times K \to K$), więc $g \boxplus f = f \boxplus g$. Dla funkcji $\Theta\colon X \to K$ określonej wzorem $\Theta(x) = 0$ dla każdego $x \in X$ otrzymujemy, że $(f \boxplus \Theta)(x) = f(x) + \Theta(x) = f(x) + 0 = f(x)$ dla każdego $x \in X$. Stąd oraz na mocy uzasadnionej wcześniej przemienności dodawania $\boxplus$, $f \boxplus \Theta = \Theta \boxplus f = f$. Niech $F\colon X \to K$ będzie funkcją daną wzorem $F(x) = -f(x)$ dla każdego $x \in X$. Wówczas $(f \boxplus F)(x) = f(x) + F(x) = f(x) + \big(-f(x)\big) = 0 = \Theta(x)$ dla każdego $x \in X$. Powołując się więc na uzasadnioną wcześniej przemienność dodawania $\boxplus$ uzyskujemy, że $f \boxplus F = F \boxplus f = \Theta$. Ponadto $\big((f \boxplus g) \boxplus h\big)(x) = (f \boxplus g)(x) + h(x) = \big(f(x) + g(x)\big) + h(x) = f(x) + \big(g(x) + h(x)\big) = f(x) + \big(g \boxplus h\big)(x) = \big(f \boxplus (g \boxplus h)\big)(x)$ dla każdego $x \in X$, skąd $f \boxplus (g \boxplus h) = (f \boxplus g) \boxplus h$. Zatem $\big(K^X, \boxplus, \Theta\big)$ jest grupą abelową. Dalej, dla każdego $x \in X$ zachodzą równości: $\big((\lambda + \mu) \circ f\big)(x) = (\lambda + \mu) \cdot f(x) = \lambda \cdot f(x) + \mu \cdot f(x) = \big(\lambda \circ f\big)(x) + \big(\mu \circ f\big)(x) = \big(\lambda \circ f \boxplus \mu \circ f\big)(x)$, $\big(\lambda \circ (f \boxplus g)\big)(x) = \lambda \cdot (f \boxplus g)(x) = \lambda \cdot \big(f(x) + g(x)\big) = \lambda \cdot f(x) + \lambda \cdot g(x) = (\lambda \circ f)(x) + (\lambda \circ g)(x) = (\lambda \circ f \boxplus \lambda \circ g)(x)$, $\big((\lambda \cdot \mu) \circ f\big)(x) = (\lambda \cdot \mu) \cdot f(x) = \lambda \cdot \big(\mu \cdot f(x)\big) = \lambda \cdot \big(\mu \circ f\big)(x) = \big(\lambda \circ (\mu \circ f)\big)(x)$ oraz $(1 \circ f)(x) = 1 \cdot f(x) = f(x)$. Wynika stąd kolejno, że $(\lambda + \mu) \circ f = \lambda \circ f \boxplus \mu \circ f$, $\lambda \circ (f \boxplus g) = \lambda \circ f \boxplus \lambda \circ g$, $(\lambda \cdot \mu) \circ f = \lambda \circ (\mu \circ f)$ i $1 \circ f = f$. Zatem rzeczywiście spełnione są wszystkie aksjomaty przestrzeni liniowej podane w Definicji 7.2.

7.3 Operacje na wektorach

Od tej pory dodawanie w ciele K oraz dodawanie w K-przestrzeni V oznaczać będziemy tym samym symbolem $+$. Również znaczenie symbolu 0 będzie zależało od kontekstu (raz 0 będzie oznaczać element neutralny dodawania w ciele K, a innym razem element neutralny dodawania w V).

Uwaga 7.7. Niech V będzie przestrzenią liniową nad ciałem K. Ponieważ V jest grupą ze względu na dodawanie wektorów, to dla każdego $v \in V$ istnieje dokładnie

jeden wektor $w \in V$ taki, że $v + w = 0$. Wektor w nazywamy wektorem przeciwnym do v i oznaczamy symbolem $-v$.

Stwierdzenie 7.1. Niech V będzie przestrzenią liniową nad ciałem K. Wówczas dla dowolnych $v, w, u \in V$ oraz $\lambda \in K$:

(i) $v + w = v + u \Rightarrow w = u$;
(ii) $0 \circ v = 0$;
(iii) $\lambda \circ 0 = 0$;
(iv) $(-1) \circ v = -v$;
(v) $(-\lambda) \circ v = \lambda \circ (-v) = -(\lambda \circ v)$;
(vi) jeżeli $\lambda \circ v = 0$, to $\lambda = 0$ lub $v = 0$.

Dowód. Weźmy dowolne $v, w, u \in V$ i $\lambda \in K$.

(i). Załóżmy, że $v + w = v + u$. Wtedy $-v + (v + w) = -v + (v + u)$, więc z łączności dodawania w V uzyskujemy $(-v + v) + w = (-v + v) + u$. Zatem $0 + w = 0 + u$, czyli $w = u$.

(ii). Zauważmy, że $0 \circ v = (0 + 0) \circ v = 0 \circ v + 0 \circ v$, na mocy punktu $(L1)$ Definicji 7.2, oraz $0 \circ v = 0 \circ v + 0$. Zatem $0 \circ v + 0 = 0 \circ v + 0 \circ v$. Stąd oraz na mocy punktu (i), $0 \circ v = 0$.

(iii). Ponieważ $\lambda \circ 0 + 0 = \lambda \circ 0 = \lambda \circ (0 + 0) = \lambda \circ 0 + \lambda \circ 0$ na mocy punktu $(L2)$ Definicji 7.2, to z punktu (i) wynika, że $\lambda \circ 0 = 0$.

(iv). Zauważmy, że $(-1) \circ v + v = (-1) \circ v + 1 \circ v = (-1 + 1) \circ v = 0 \circ v$ na mocy punktu $(L1)$ Definicji 7.2. Stąd oraz na mocy punktu (ii), $(-1) \circ v + v = 0$. Zatem $-v = (-1) \circ v$.

(v). Ponieważ $(-\lambda) \circ v + \lambda \circ v = (-\lambda + \lambda) \circ v = 0 \circ v$ na mocy punktu $(L1)$ Definicji 7.2, to z punktu (ii) wynika, że $(-\lambda) \circ v + \lambda \circ v = 0$. Zatem $(-\lambda) \circ v = -(\lambda \circ v)$. Dalej, $(-\lambda) \circ v = (\lambda \cdot (-1)) \circ v = \lambda \circ ((-1) \circ v)$. Ponadto $(-1) \circ v = -v$ na mocy (iv), więc $(-\lambda) \circ v = \lambda \circ (-v)$.

(vi). Załóżmy, że $\lambda \circ v = 0$ i $\lambda \neq 0$. Wtedy istnieje w K element λ^{-1} oraz $0 = \lambda^{-1} \circ 0 = \lambda^{-1} \circ (\lambda \circ v) = (\lambda^{-1} \cdot \lambda) \circ v = 1 \circ v = v$ odpowiednio na mocy punktów punktu (iii) niniejszego stwierdzenia oraz punktów $(L3)$ i $(L4)$ Definicji 7.2.

Uwaga 7.8. Dowód punktu (vi) powyższego twierdzenia opiera się na tautologii:

$$\big(p \Rightarrow (q \vee r)\big) \Leftrightarrow \big((p \wedge \neg q) \Rightarrow r\big)$$

rachunku zdań.

Wniosek 7.1. Niech V będzie przestrzenią liniową nad ciałem K. Wówczas dla dowolnych $n \in \mathbb{N}$, $v_1, v_2, \ldots, v_n \in V$ oraz $\lambda, \lambda_1, \lambda_2, \ldots, \lambda_n \in K$:

(i) $-(v_1+v_2+\ldots+v_n)=(-v_1)+(-v_2)+\ldots+(-v_n)$;
(ii) $\lambda\circ(v_1+v_2+\ldots+v_n)=\lambda\circ v_1+\lambda\circ v_2+\ldots+\lambda\circ v_n$;
(iii) $\lambda\circ(\lambda_1\circ v_1+\lambda_2\circ v_2+\ldots+\lambda_n\circ v_n)=(\lambda\cdot\lambda_1)\circ v_1+(\lambda\cdot\lambda_2)\circ v_2+\ldots+(\lambda\cdot\lambda_n)\circ v_n$.

Dowód. (i). Powołując się na punkt (iv) Stwierdzenia 7.1 oraz punkt $(L2)$ Definicji 7.2 otrzymujemy, że $-(v_1+v_2+\ldots+v_n)=(-1)\circ(v_1+v_2)=(-1)\circ v_1+(-1)\circ v_2=(-v_1)+(-v_2)$. Stosując prostą indukcję uzyskujemy tezę.

(ii). Teza wynika z punktu $(L2)$ Definicji 7.2 przez prostą indukcję.

(iii). Teza wynika wprost z punktu (ii) oraz punktu $(L3)$ Definicji 7.2

7.4 Podprzestrzeń przestrzeni liniowej

7.4.1 Określenie podprzestrzeni przestrzeni liniowej

Definicja 7.3. Niech V będzie przestrzenią liniową nad ciałem K. Niepusty podzbiór U zbioru V nazywamy podprzestrzenią K-przestrzeni liniowej V, gdy dla wszystkich $u,u_1,u_2\in U$ oraz $\lambda\in K$ spełniona jest koniunkcja warunków:

(P1) $u_1+u_2\in U$;
(P2) $\lambda\circ u\in U$.

W teoretycznych rozważaniach często pomocna jest poniższa charakteryzacja podprzestrzeni przestrzeni liniowej.

Stwierdzenie 7.2. Niepusty podzbiór U przestrzeni liniowej V nad ciałem K jest podprzestrzenią w V wtedy i tylko wtedy, gdy $\lambda_1\circ u_1+\lambda_2\circ u_2\in U$ dla wszystkich $\lambda_1,\lambda_2\in K$ oraz $u_1,u_2\in U$.

Dowód. Załóżmy, że U jest podprzestrzenią K-przestrzeni V. Weźmy dowolne $\lambda_1,\lambda_2\in K$ i $u_1,u_2\in U$. Z $(P2)$ wynika wówczas, że $\lambda_1\circ u_1\in U$ oraz $\lambda_2\circ u_2\in U$, więc na mocy $(P1)$ otrzymujemy, że $\lambda_1\circ u_1+\lambda_2\circ u_2\in U$.

Na odwrót. Przypuśćmy, że niepusty podzbiór U K-przestrzeni V dla wszystkich $\lambda_1,\lambda_2\in K$ i $u_1,u_2\in U$ spełnia warunek $\lambda_1\circ u_1+\lambda_2\circ u_2\in U$. Weźmy dowolne $u,u_1,u_2\in U$ oraz dowolne $\lambda\in K$. Wtedy $u_1+u_2=1\circ u_1+1\circ u_2\in U$ oraz $\lambda\circ u=\lambda\circ u+0\circ u\in U$, skąd wynika, że spełnione są warunki $(P1)$ i $(P2)$. Zatem U jest podprzestrzenią w V.

Uwaga 7.9. Niech U będzie podprzestrzenią przestrzeni liniowej V nad ciałem K. Ponieważ $U\neq\emptyset$, to istnieje $u\in U$. Stąd oraz na mocy $(P2)$ i punktu (ii) Stwierdzenia 7.1, $0=0\circ u\in U$. Ponadto z $(P2)$ i punktu (iv) Stwierdzenia 7.1 otrzymujemy, że $-u\in U$ dla każdego $u\in U$.

Uwaga 7.10. Podprzestrzeń U przestrzeni liniowej V nad ciałem K jest przestrzenią liniową nad K. Istotnie, $U \neq \emptyset$ oraz z warunków $(P1)$ i $(P2)$ wynika, że zbiory wartości funkcji $+_{|U\times U}$ oraz $\circ_{|K\times U}$ są podzbiorami w U. Ponadto zbiór V rozważany wraz z działaniami $+\colon V\times V\to V$ oraz $\circ\colon K\times V\to V$ jest K-przestrzenią, więc zbiór U rozważany wraz z działaniami $+_{|U\times U}$ i $\circ_{|K\times U}$ spełnia wszystkie warunki dane w Definicji 7.2.

Przykład 7.11. Dla dowolnej przestrzeni liniowej V nad ciałem K, $\{0\}$ oraz V są podprzestrzeniami w V.

Wniosek 7.2. Każda niezerowa przestrzeń linowa V nad ciałem K posiada co najmniej dwie podprzestrzenie: $\{0\}$ i V.

Przykład 7.12. Z Przykładu 7.6 wynika, że zbiór rozwiązań jednorodnego układu m równań z n niewiadomymi jest podprzestrzenią przestrzeni liniowej K^n.

Definicja 7.4. Podprzestrzeń U przestrzeni liniowej V nad ciałem K taką,że $U \subsetneq V$ nazywamy podprzestrzenią właściwą.

Przykład 7.13. Podzbiór $U = \{[x,0]\colon x\in\mathbb{R}\}$ jest właściwą nietrywialną podprzestrzenią przestrzeni liniowej $\mathbb{R}^2$ nad ciałem $\mathbb{R}$. Istotnie, $U\neq\emptyset$, $U\neq\{[0,0]\}$, $U\neq\mathbb{R}^2$ oraz dla wszystkich $x,y,z\in\mathbb{R}$ mamy: $[x,0]+[y,0]=[x+y,0]\in U$ oraz $z\circ[x,0]=[zx,0]\in U$.

7.4.2 Podprzestrzeń generowana i jej własności

Twierdzenie 7.1. Część wspólna dowolnej niepustej rodziny podprzestrzeni przestrzeni liniowej V nad ciałem K jest podprzestrzenią w V.

Dowód. Rozważmy dowolną niepustą rodzinę $\mathfrak{A}$ podprzestrzeni przestrzeni liniowej V nad ciałem K i oznaczmy $U_0=\bigcap\mathfrak{A}$. Z Uwagi 7.9 wynika, że $0\in U_0$, więc $U_0\neq\emptyset$. Weźmy dowolne $\lambda,\mu\in K$ oraz $a,b\in U_0$. Wtedy $a,b\in U$ dla każdego $U\in\mathfrak{A}$, więc ze Stwierdzenia 7.2 wynika, że $\lambda\circ a+\mu\circ b\in U$ dla każdego $U\in\mathfrak{A}$. Zatem $\lambda\circ a+\mu\circ b\in U_0$. Powołując się ponownie na Stwierdzenie 7.2 otrzymujemy stąd, że U_0 jest podprzestrzenią K-przestrzeni V.

Twierdzenie 7.2. Dla dowolnego podzbioru X przestrzeni liniowej V nad ciałem K istnieje najmniejsza względem inkluzji podprzestrzeń K-przestrzeni V zawierająca zbiór X.

Dowód. Niech $\mathfrak{A}$ oznacza rodzinę wszystkich podprzestrzeni K-przestrzeni V zawierających zbiór X. Wówczas $\mathfrak{A}\neq\emptyset$, bo $V\in\mathfrak{A}$. Niech $U_0=\bigcap\mathfrak{A}$. Z Twierdzenia 7.1 wynika wówczas, że U_0 jest podprzestrzenią w V. Ponadto $X\subseteq U_0$. Rozważmy dowolną podprzestrzeń W K-przestrzeni V taką, że $X\subseteq W$. Wtedy $W\in\mathfrak{A}$,

więc $U_0 \subseteq W$. Wobec tego U_0 jest najmniejszą względem inkluzji podprzestrzenią K-przestrzeni V zawierającą zbiór X.

Możemy teraz wprowadzić ważne pojęcie podprzestrzeni generowanej przez podzbiór przestrzeni liniowej.

Definicja 7.5. Niech X będzie dowolnym podzbiorem przestrzeni liniowej V nad ciałem K. Najmniejszą względem inkluzji podprzestrzeń K-przestrzeni V zawierającą zbiór X nazywamy podprzestrzenią generowaną przez zbiór X w V lub podprzestrzenią w V rozpiętą na zbiorze X. Oznaczamy ją przez $\mathrm{lin}_K(X)$ albo $\mathrm{lin}(X)$, gdy z kontekstu jasno wynika nad jakim ciałem rozważane są przestrzenie liniowe. Jeżeli $X = \{x_1, x_2, \ldots, x_n\}$ dla pewnego $n \in \mathbb{N}$, to będziemy pisać $\mathrm{lin}(x_1, x_2, \ldots, x_n)$ zamiast $\mathrm{lin}\big(\{x_1, x_2, \ldots, x_n\}\big)$.

Uwaga 7.11. Z powyższej definicji wynika natychmiast, że:

(i) $\mathrm{lin}(\emptyset) = \{0\}$;
(ii) $\mathrm{lin}(V) = V$;
(iii) jeżeli $X \subseteq Y \subseteq V$, to $\mathrm{lin}(X) \subseteq \mathrm{lin}(Y)$.

Stwierdzenie 7.3. Niech V będzie przestrzenią liniową nad ciałem K, niech $v \in V$ i niech $X \subseteq V$. Wówczas $v \in \mathrm{lin}(X)$ wtedy i tylko wtedy, gdy $\mathrm{lin}\big(X \cup \{v\}\big) = \mathrm{lin}(X)$.

Dowód. Jeżeli $v \in \mathrm{lin}(X)$, to $X \cup \{v\} \subseteq \mathrm{lin}(X)$, skąd $\mathrm{lin}\big(X \cup \{v\}\big) \subseteq \mathrm{lin}(X)$. Inkluzja przeciwna wynika wprost z punktu (iii) Uwagi 7.11. Jeśli natomiast $\mathrm{lin}\big(X \cup \{v\}\big) = \mathrm{lin}(X)$, to $v \in \mathrm{lin}(X)$, bo $v \in \mathrm{lin}\big(X \cup \{v\}\big)$.

Definicja 7.6. Niech $n \in \mathbb{N}$ i niech $v_1, v_2, \ldots, v_n$ będą wektorami przestrzeni liniowej V nad ciałem K. Mówimy, że wektor $v \in V$ jest kombinacją liniową wektorów $v_1, v_2, \ldots, v_n$, gdy istnieją skalary $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ takie, że $v = \sum_{i=1}^{n} \lambda_i v_i$. Skalary $\lambda_1, \lambda_2, \ldots, \lambda_n$ nazywamy wówczas współczynnikami tej kombinacji liniowej.

Stwierdzenie 7.4. Niech X będzie dowolnym niepustym podzbiorem przestrzeni liniowej V nad ciałem K. Wtedy $\mathrm{lin}(X)$ jest zbiorem wszystkich kombinacji liniowych wszystkich skończonych podzbiorów zbioru X, tzn.:

$$\mathrm{lin}(X) = \left\{ \sum_{i=1}^{n} \lambda_i \circ x_i \colon n \in \mathbb{N} \wedge \lambda_1, \lambda_2 \ldots, \lambda_n \in K \wedge x_1, x_2, \ldots, x_n \in X \right\}. \qquad (7.4.1)$$

Dowód. Niech W oznacza zbiór występujący po prawej stronie równości (7.4.1). Ponieważ $x = 1 \circ x \in W$ dla każdego $x \in X$, to $X \subseteq W$. W szczególności wynika stąd, że $W \neq \emptyset$. Weźmy dowolne $a, b \in W$ oraz $\lambda, \mu \in K$. Istnieją wówczas $m, n \in \mathbb{N}$ oraz $\alpha_1, \alpha_2 \ldots, \alpha_m, \beta_1, \beta_2 \ldots, \beta_n \in K$ i $a_1, a_2 \ldots, a_m, b_1, b_2, \ldots, b_n \in X$ takie, że $a = \sum_{i=1}^{m} \alpha_i \circ a_i$ oraz $b = \sum_{j=1}^{n} \beta_j \circ b_j$. Zatem $\lambda \circ a + \mu \circ b = \lambda \circ \sum_{i=1}^{m} \alpha_i \circ a_i + \mu \circ \sum_{j=1}^{n} \beta_j \circ b_j = \sum_{i=1}^{m} (\lambda \cdot \alpha_i) \circ a_i + \sum_{j=1}^{n} (\mu \cdot \beta_j) \circ b_j$. Oznaczając $s = m + n$ oraz $x_i = a_i$, $\lambda_i =$

$\lambda \cdot \alpha_i$, $x_{m+j} = b_j$ i $\lambda_{m+j} = \lambda \cdot \beta_j$ dla wszystkich $i \in \{1,2,\ldots,m\}$ oraz $j \in \{1,2,\ldots,n\}$ otrzymujemy, że $s \in \mathbb{N}$, $x_t \in X$ i $\lambda_t \in K$ dla każdego $t \in \{1,2,\ldots,s\}$ oraz $\lambda \circ a + \mu \circ b = \sum_{t=1}^{s} \lambda_t \circ x_t \in W$. Stąd oraz na mocy Stwierdzenia 7.2, W jest podprzestrzenią K-przestrzeni V zawierającą zbiór X. Dalej, rozważmy dowolną podprzestrzeń U K-przestrzeni V taką, że $X \subseteq U$. Niech $n \in \mathbb{N}$ i niech $x_1, x_2, \ldots, x_n$ oznaczają teraz dowolne elementy X. Wówczas $x_1, x_2, \ldots, x_n \in U$, więc ze Stwierdzenia 7.2 przez prostą indukcję wynika, że $\sum_{i=1}^{n} \lambda_i \circ x_i \in U$ dla wszystkich $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Wobec tego $W \subseteq U$. Stąd oraz na mocy dowolności wyboru U uzyskujemy, że $W = \mathrm{lin}(X)$.

Bezpośrednią konsekwencją powyższego twierdzenia oraz własności działań na wektorach jest następujący

Wniosek 7.3. Niech V będzie przestrzenią liniową nad ciałem K i niech $v, v_1, \ldots, v_n \in V$. Wówczas:

(i) $\mathrm{lin}(v) = \{\lambda \circ v \colon \lambda \in K\}$;
(ii) $\mathrm{lin}(v_1, v_2, \ldots, v_n) = \{\sum_{i=1}^{n} \lambda_i \circ v_i \colon \lambda_1, \lambda_2, \ldots, \lambda_n \in K\}$.

Definicja 7.7. Niech $I \neq \emptyset$ i niech $\{V_i \colon i \in I\}$ będzie rodziną podprzestrzeni przestrzeni liniowej V nad ciałem K. Podprzestrzeń $\mathrm{lin}(\bigcup_{i \in I} V_i)$ nazywamy sumą algebraiczną podprzestrzeni rodziny $\{V_i \colon i \in I\}$ i oznaczamy przez $\sum_{i \in I} V_i$.

Uwaga 7.12. Gdy $I = \{1,2,\ldots,n\}$ dla pewnego $n \in \mathbb{N}$, to często pisze się $\bigcup_{i=1}^{n} V_i$ zamiast $\bigcup_{i \in I} V_i$ oraz $\sum_{i=1}^{n} V_i$ zamiast $\sum_{i \in I} V_i$.

Twierdzenie 7.3. Niech $I \neq \emptyset$ i niech $\{X_i \colon i \in I\}$ będzie rodziną podzbiorów przestrzeni liniowej V nad ciałem K. Wtedy $\mathrm{lin}\left(\bigcup_{i \in I} X_i\right) = \sum_{i \in I} \mathrm{lin}(X_i)$.

Dowód. Ponieważ $X_i \subseteq \mathrm{lin}(X_i)$ dla każdego $i \in I$, to $\bigcup_{i \in I} X_i \subseteq \bigcup_{i \in I} \mathrm{lin}(X_i)$ i w konsekwencji $\bigcup_{i \in I} X_i \subseteq \mathrm{lin}\left(\bigcup_{i \in I} \mathrm{lin}(X_i)\right)$, czyli $\bigcup_{i \in I} X_i \subseteq \sum_{i \in I} \mathrm{lin}(X_i)$. Stąd $\mathrm{lin}\left(\bigcup_{i \in I} X_i\right) \subseteq \sum_{i \in I} \mathrm{lin}(X_i)$. Udowodnimy teraz inkluzję przeciwną. W tym celu weźmy dowolne $i_0 \in I$. Wówczas $X_{i_0} \subseteq \bigcup_{i \in I} X_i$, więc $\mathrm{lin}\left(X_{i_0}\right) \subseteq \mathrm{lin}\left(\bigcup_{i \in I} X_i\right)$. Zatem $\bigcup_{i \in I} \mathrm{lin}\left(X_i\right) \subseteq \mathrm{lin}\left(\bigcup_{i \in I} X_i\right)$. Stąd $\mathrm{lin}\left(\bigcup_{i \in I} \mathrm{lin}\left(X_i\right)\right) \subseteq \mathrm{lin}\left(\bigcup_{i \in I} X_i\right)$, czyli $\sum_{i \in I} \mathrm{lin}\left(X_i\right) \subseteq \mathrm{lin}\left(\bigcup_{i \in I} X_i\right)$.

Ponieważ $X = \bigcup_{x \in X} \{x\}$ dla dowolnego niepustego zbioru X, to powyższe twierdzenie implikuje następujący

Wniosek 7.4. Niech X będzie dowolnym niepustym podzbiorem przestrzeni liniowej V nad ciałem K. Wtedy $\mathrm{lin}(X) = \sum_{x \in X} \mathrm{lin}(x)$.

Twierdzenie 7.4. Niech $I \neq \emptyset$ i niech $\{V_i \colon i \in I\}$ będzie rodziną podprzestrzeni przestrzeni liniowej V nad ciałem K. Wówczas:

$$\sum_{i \in I} V_i = \left\{ \sum_{j=1}^{n} v_{i_j} \colon n \in \mathbb{N} \wedge v_{i_j} \in V_{i_j} \text{ dla każdego } i_1, i_2, \ldots, i_n \in I \right\}. \tag{7.4.2}$$

Dowód. Niech W oznacza zbiór występujący po prawej stronie równości (7.4.2). Wtedy $W \neq \emptyset$, bo $0 \in W$. Weźmy dowolne $a, b \in W$ i $\alpha, \beta \in K$. Istnieją wówczas $n \in \mathbb{N}$, $i_1, i_2, \ldots, i_n \in I$ oraz $w_{i_1}, u_{i_1} \in V_{i_1}$, $w_{i_2}, u_{i_2} \in V_{i_2}, \ldots$, $w_{i_n}, u_{i_n} \in V_{i_n}$ takie, że $a = \sum_{j=1}^{n} w_{i_j}$ oraz $b = \sum_{j=1}^{n} u_{i_j}$. Dla każdego $j \in \{1, 2, \ldots, n\}$ definiujemy $v_{i_j} = \alpha \circ w_{i_j} + \beta \circ u_{i_j}$. Ponieważ V_{i_j} jest podprzestrzenią w V dla każdego $j \in \{1, 2, \ldots, n\}$, to ze Stwierdzenia 7.2 wynika, że $v_{i_j} \in V_{i_j}$ dla każdego $j \in \{1, 2, \ldots, n\}$. Zatem $\alpha \circ a + \beta \circ b = \sum_{j=1}^{n} v_{i_j} \in W$. Powołując się ponownie na Stwierdzenie 7.2 otrzymujemy stąd, że W jest podprzestrzenią K-przestrzeni V. Ponadto wprost z określenia W wynika, że $\bigcup_{i \in I} V_i \subseteq W$. Rozważmy dowolną podprzestrzeń U K-przestrzeni V taką, że $\bigcup_{i \in I} V_i \subseteq U$. Z punktu $(P2)$ Definicji 7.3 przez prostą indukcję wynika wówczas, że $\sum_{j=1}^{s} v_{i_j} \in U$ dla wszystkich $s \in \mathbb{N}$, $i_1, i_2, \ldots, i_s \in I$ oraz $v_{i_1} \in V_{i_s}$, $v_{i_2} \in V_{i+2}, \ldots$, $v_{i_s} \in V_{i_s}$. Zatem $W \subseteq U$. Wobec tego $W = \operatorname{lin}(\bigcup_{i \in I} V_i)$, co w świetle Definicji 7.7 oznacza, że $W = \sum_{i \in I} V_i$.

Bezpośrednią konsekwencją powyższego twierdzenia jest następujący

Wniosek 7.5. Niech V będzie przestrzenią liniową nad ciałem K, niech $n \in \mathbb{N}$ i niech $V_1, V_2, \ldots, V_n$ będą podprzestrzeniami w V. Wówczas:

$$\sum_{i=1}^{n} V_i = \left\{ \sum_{i=1}^{n} v_i \colon v_i \in V_i \text{ dla każdego } i \in \{1, 2, \ldots, n\} \right\}.$$

Przykład 7.14. Niech K będzie ciałem i niech $n \in \mathbb{N}$. Weźmy dowolne $a \in K^n$. Wówczas $a = [a_1, a_2, \ldots, a_n]$ dla pewnych $a_1, a_2, \ldots, a_n \in K$. Dla każdego $i \in \{1, 2, \ldots, n\}$ definiujemy $\varepsilon_i = [0, 0, \ldots, 0, \overset{i}{1}, 0 \ldots 0]$. Wówczas $a = \sum_{i=1}^{n} a_i \circ \varepsilon_i$. Zatem:

$$K^n = \operatorname{lin}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n).$$

7.5 Operacje elementarne na układach wektorów

Definicja 7.8. Niech $v_1, v_2, \ldots, v_n$ będą wektorami przestrzeni liniowej V nad ciałem K. Ciąg $(v_1, v_2, \ldots, v_n)$ nazywamy układem wektorów $v_1, v_2, \ldots, v_n$.

Uwaga 7.13. Niech $n \geq 2$ będzie liczbą naturalną i niech i oraz j będą różnymi elementami zbioru $\{1, 2, \ldots, n\}$. Wyróżniamy następujące operacje elementarne na układzie wektorów $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K:

(O1.) zamiana miejscami wektora v_i z wektorem v_j oznaczana przez $v_i \leftrightarrow v_j$;
(O2.) pomnożenie wektora v_i przez dowolny niezerowy skalar $\lambda \in K$ oznaczane przez $\lambda \circ v_i$;
(O3.) dodanie do wektora v_i wektora v_j pomnożonego przez dowolny skalar $\lambda \in K$ oznaczane przez $v_i + \lambda \circ v_j$.

Nietrudno zauważyć, że każda z powyższych operacji jest odwracalna.

Twierdzenie 7.5. Niech n będzie liczbą naturalną większą od 1. Jeżeli układ wektorów $(w_1, w_2, \ldots, w_n)$ powstaje wskutek wykonania kolejno skończonej liczby operacji elementarnych na układzie $(v_1, v_2, \ldots, v_n)$ wektorów przestrzeni liniowej V nad ciałem K, to $\lin(v_1, v_2, \ldots, v_n) = \lin(w_1, w_2, \ldots, w_n)$.

Dowód. Ponieważ operacje $(O1.)$-$(O3.)$ są odwracalne, to wystarczy wykazać inkluzję $\lin(w_1, w_2, \ldots, w_n) \subseteq \lin(v_1, v_2, \ldots, v_n)$. Niech s oznacza liczbę wykonanych kolejno operacji elementarnych. Dla $s = 0$ teza jest oczywista. Załóżmy, że $s = 1$. Jeśli wykonana została operacja $(O1.)$, to $w_i = v_j$, $w_j = v_i$ oraz $w_t = v_t$ dla każdego $t \in \{1, 2, \ldots, n\} \setminus \{i, j\}$. Żądana równość wynika więc z równości zbiorów $\{v_1, \ldots, v_i \ldots, v_j, \ldots, v_n\} = \{v_1, \ldots, v_j \ldots, v_i, \ldots, v_n\}$. Jeżeli wykonana została operacja $(O2.)$, to $w_j = v_j$ dla każdego $j \in \{1, 2, \ldots, n\} \setminus \{i\}$ oraz $w_i = \lambda \circ v_i$ dla pewnego $\lambda \in K \setminus \{0\}$. Zatem $w_t \in \lin(v_1, v_2, \ldots, v_n)$ dla każdego $t \in \{1, 2, \ldots, n\}$, skąd $\lin(w_1, w_2, \ldots, w_n) \subseteq \lin(v_1, v_2, \ldots, n)$. Przypuśćmy teraz, że została wykonana operacja $(O3.)$. Wtedy $w_t = v_t$ dla każdego $t \in \{1, 2, \ldots, n\} \setminus \{i\}$ oraz $w_i = v_i + \lambda \circ v_j$ dla pewnego $\lambda \in K$. Zatem $w_k \in \lin(v_1, v_2, \ldots, v_n)$ dla każdego $k \in \{1, 2, \ldots, n\}$, skąd $\lin(w_1, w_2, \ldots, w_n) \subseteq \lin(v_1, v_2, \ldots, v_n)$. Dla pozostałych $s \in \mathbb{N}$ teza wynika przez prostą indukcję.

Przykład 7.15. Pokażemy, że dla dowolnych wektorów v_1, v_2 przestrzeni liniowej V nad ciałem K zachodzi równość $\lin(3v_1 + 4v_2, 5v_1 + 7v_2) = \lin(v_1, v_2)$. W poniższych rachunkach symbole w_1 i w_2 za każdym razem oznaczają zapisane kolejno elementy zbioru rozpinającego podprzestrzeń. Mamy:

$$\lin(3v_1 + 4v_2, 5v_1 + 7v_2) \overset{w_2 - 2 \circ w_1}{=} \lin(3v_1 + 4v_2, -v_1 - v_2) \overset{w_1 + 3 \circ w_2}{=} \lin(v_2, -v_1 - v_2)$$

$$\overset{w_2 + w_1}{=} \lin(v_2, -v_1) \overset{(-1) \circ w_2}{=} \lin(v_2, v_1) \overset{w_1 \leftrightarrow w_2}{=} \lin(v_1, v_2).$$

Rozdział 8

Baza i wymiar przestrzeni liniowej

8.1 Liniowa niezależność wektorów

Definicja 8.1. Niech $n \in \mathbb{N}$. Mówimy, że układ wektorów $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K jest liniowo niezależny, wówczas gdy dla wszystkich $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ warunek $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$ implikuje, że $\lambda_i = 0$ dla każdego $i \in \{1, 2, \ldots, n\}$.

Uwaga 8.1. Warunek $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$ można wysłowić w następujący sposób: kombinacja liniowa wektorów $v_1, v_2, \ldots, v_n$ o współczynnikach $\lambda_1, \lambda_2, \ldots, \lambda_n$ jest zerowa. Jeśli wszystkie współczynniki tej kombinacji są równe zero, to mówimy, że jest ona trywialna. Zatem układ $(v_1, v_2, \ldots, v_n)$ wektorów K-przestrzeni V jest liniowo niezależny, gdy każda zerowa kombinacja liniowa wektorów $v_1, v_2, \ldots, v_n$ jest trywialna.

Przykład 8.1. Zachowując oznaczenia z Przykładu 7.14, łatwo zauważyć, że dla dowolnego ciała K i dowolnej liczby naturalnej n, ciąg $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ jest liniowo niezależnym układem wektorów przestrzeni liniowej K^n.

Definicja 8.2. Niech $n \in \mathbb{N}$. Mówimy, że układ wektorów $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K jest liniowo zależny, wówczas gdy nie jest on liniowo niezależny, czyli gdy istnieją $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ takie, że $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$ oraz $\lambda_{i_0} \neq 0$ dla pewnego $i_0 \in \{1, 2, \ldots n\}$.

Uwaga 8.2. Innymi słowy, układ wektorów $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K jest liniowo zależny, gdy istnieje nietrywialna zerowa kombinacja liniowa wektorów $v_1, v_2, \ldots, v_n$.

Przykład 8.2. Niech V będzie przestrzenią liniową nad ciałem K. Wówczas układ (0) jest liniowo zależny, bo $1 \circ 0 = 0$ i $1 \neq 0$ w ciele K.

Przykład 8.3. Niech $n \in \mathbb{N}$. Dla dowolnych wektorów $v_1, v_2, \ldots, v_n$ przestrzeni liniowej V nad ciałem K układ wektorów $(0, v_1, v_2, \ldots, v_n)$ jest liniowo zależny. Istotnie, kombinacja liniowa $1 \circ 0 + 0 \circ v_1 + 0 \circ v_2 + \ldots + 0 \circ v_n$ jest zerowa i nietrywialna.

Przykład 8.4. Dla dowolnych wektorów $v_1, v_2, \ldots, v_n$ przestrzeni liniowej V nad ciałem K takich, że $v_i = v_j$ dla pewnych różnych $i, j \in \{1, 2, \ldots, n\}$, układ $(v_1, v_2, \ldots, v_n)$ jest liniowo zależny, gdyż $0 \circ v_1 + \ldots + 0 \circ v_{i-1} + 1 \circ v_i + 0 \circ v_{i+1} + \ldots + 0 \circ v_{j-1} + (-1) \circ v_j + 0 \circ v_{j+1} + \ldots + 0 \circ v_n = 0$ jest nietrywialną zerową kombinacją liniową wektorów $v_1, v_2, \ldots, v_n$.

Uwaga 8.3. Niech $n \in \mathbb{N}$. Łatwo zauważyć, że układ wektorów $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K jest liniowo niezależny wtedy i tylko wtedy, gdy dla dowolnej permutacji $\sigma \in S_n$, liniowo niezależny jest układ $(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)})$.

Obserwacja odnotowana w powyższej uwadze uzasadnia wprowadzenie następującej definicji.

Definicja 8.3. Niech $n \in \mathbb{N}$. Mówimy, że n-elementowy podzbiór $\{v_1, v_2, \ldots, v_n\}$ przestrzeni liniowej V nad ciałem K jest liniowo niezależny, gdy liniowo niezależny jest układ wektorów $(v_1, v_2, \ldots, v_n)$.

Uwaga 8.4. Zbiór pusty uznajemy za liniowo niezależny.

Przykład 8.5. Niech $w, v_1, v_2, \ldots, v_n$ będą wektorami przestrzeni liniowej V nad ciałem K takimi, że układ $(w, v_1, v_2, \ldots, v_n)$ jest liniowo niezależny. Z Przykładu 8.4 wynika wówczas, że układ $(w, w, v_1, v_2, \ldots, v_n)$ jest liniowo zależny. Ponieważ $\{w, w, v_1, v_2, \ldots, v_n\} = \{w, v_1, v_2, \ldots, v_n\}$ i $(w, v_1, v_2, \ldots, v_n)$ jest liniowo niezależnym układem wektorów, to zbiór $\{w, w, v_1, v_2, \ldots, v_n\}$ jest liniowo niezależny. Dlatego pojęć układu oraz zbioru liniowo niezależnego nie można używać zamiennie.

Lemat 8.1. Niech $n \in \mathbb{N}$. Jeżeli n-elementowy podzbiór przestrzeni liniowej V nad ciałem K jest liniowo niezależny, to liniowo niezależny jest także każdy niepusty podzbiór tego zbioru.

Dowód. Niech $X = \{v_1, v_2, \ldots, v_n\}$ będzie n-elementowym podzbiorem w V, niech $s \leq n$ będzie liczbą naturalną i niech $Y = \{v_{i_1}, v_{i_2}, \ldots, v_{i_s}\}$ będzie s-elementowym podzbiorem w X. Załóżmy, że zbiór Y nie jest liniowo niezależny. Istnieją wówczas $\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_s} \in K$ takie, że $\sum_{j=1}^{s} \lambda_{i_j} \circ v_{i_j} = 0$ oraz $\lambda_{i_t} \neq 0$ dla pewnego $t \in \{1, 2, \ldots, s\}$. Uzupełniając ciąg $(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_s})$ $n-s$ zerami $0 \in K$ uzyskujemy ciąg $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ taki, że $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$ oraz $\lambda_k \neq 0$ dla pewnego $k \in \{1, 2, \ldots, n\}$. Zatem zbiór X nie jest liniowo niezależny.

Definicja 8.4. Podzbiór X przestrzeni liniowej V nad ciałem K nazywamy liniowo niezależnym, gdy każdy jego skończony podzbiór jest liniowo niezależny. W przeciwnym razie mówimy, że podzbiór X jest liniowo zależny w V.

Bezpośrednią konsekwencją Lematu 8.1 i Definicji 8.4 jest następujące

Stwierdzenie 8.1. Każdy podzbiór liniowo niezależnego zbioru wektorów dowolnej przestrzeni liniowej jest zbiorem liniowo niezależnym.

Przykład 8.6. Dla dowolnej liczby naturalnej $n > 1$ i dowolnego ciała K zbiór

$$\left\{E_{ij} \colon i, j \in \{1, 2, \ldots, n\}\right\}$$

jest liniowo niezależnym podzbiorem w $\mathrm{M}_n(K)$ (por. Definicja 5.5).

Przykład 8.7. Zbiór $\{1, x, x^2, \ldots\}$ jest liniowo niezależnym podzbiorem przestrzeni liniowej $K[x]$ wielomianów zmiennej x o współczynnikach z ciała K.

Przykład 8.8. Pokażemy, że $A = \{\sin, \cos, 1\}$ jest liniowo niezależnym podzbiorem przestrzeni liniowej $\mathbb{R}^{\mathbb{R}}$ (por. Przykład 7.10). W tym celu rozważmy dowolne $\alpha, \beta\gamma \in \mathbb{R}$ i załóżmy, że $\alpha \circ \sin + \beta \circ \cos + \gamma \circ 1 = \Theta$. Wówczas $(\alpha \circ \sin + \beta \circ \cos + \gamma \circ 1)(x) = \Theta(x)$, czyli $\alpha \cdot \sin(x) + \beta \cdot \cos(x) + \gamma = 0$ dla każdego $x \in \mathbb{R}$. Podstawiając za x kolejno $\frac{\pi}{2}$, 0 i $\frac{\pi}{4}$ otrzymujemy:

$$\begin{cases} \alpha + \gamma = 0 \\ \beta + \gamma = 0 \\ \frac{\sqrt{2}}{2}\alpha + \frac{\sqrt{2}}{2}\beta + \gamma = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha + \gamma = 0 \\ \beta - \alpha = 0 \\ \frac{\sqrt{2}}{2}\alpha + \frac{\sqrt{2}}{2}\beta + \gamma = 0 \end{cases} \Leftrightarrow \begin{cases} \gamma = -\alpha \\ \beta = \alpha \\ \left(\sqrt{2} - 1\right)\alpha = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \alpha = 0 \\ \beta = 0 \\ \gamma = 0 \end{cases}.$$

Zatem A jest liniowo niezależnym podzbiorem w $\mathbb{R}^{\mathbb{R}}$.

Twierdzenie 8.1. Niech n będzie liczbą naturalną większą od 1. Jeżeli układ wektorów $(w_1, w_2, \ldots, w_n)$ powstaje wskutek wykonania kolejno skończonej liczby operacji elementarnych na układzie $(v_1, v_2, \ldots, v_n)$ wektorów przestrzeni liniowej V nad ciałem K, to układ $(w_1, w_2, \ldots, w_n)$ jest liniowo niezależny wtedy i tylko wtedy, gdy liniowo niezależny jest układ $(v_1, v_2, \ldots, v_n)$.

Dowód. Ponieważ operacje elementarne na układach wektorów są odwracalne to wystarczy wykazać, że liniowa niezależność układu $(v_1, v_2, \ldots, v_n)$ implikuje liniową niezależność układu $(w_1, w_2, \ldots, w_n)$. Niech s oznacza liczbę wykonanych kolejno operacji elementarnych. Dla $s = 0$ teza jest oczywista. Załóżmy, że $s = 1$. Jeśli wykonana została operacja $(O1.)$ (zob. Uwaga 7.13), to $w_i = v_j$, $w_j = v_i$ oraz $w_t = v_t$ dla każdego $t \in \{1, 2, \ldots, n\} \setminus \{i, j\}$. Teza wynika więc z Uwagi 8.3. Jeżeli wykonana została operacja $(O2.)$, to $w_j = v_j$ dla każdego $j \in \{1, 2, \ldots, n\} \setminus \{i\}$ oraz $w_i = \lambda \circ v_i$ dla pewnego $\lambda \in K \setminus \{0\}$. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Załóżmy, że $\sum_{i=1}^{n} \lambda_i \circ w_i = 0$. Wtedy $\lambda_1 \circ v_1 + \ldots + (\lambda_i \cdot \lambda) \circ v_i + \ldots + \lambda_n \circ v_n = 0$, więc na mocy liniowej niezależności układu $(v_1, v_2, \ldots, v_n)$ otrzymujemy, że $\lambda_1 = \ldots = \lambda_{i-1} = \lambda_{i+1} = \ldots = \lambda_n = 0$ oraz $\lambda_i \circ \lambda = 0$. Ponadto $\lambda \neq 0$, więc $\lambda_i = 0$. Zatem każda zerowa kombinacja wektorów $w_1, w_2, \ldots, w_n$ jest trywialna, co oznacza liniową niezależność układu $(w_1, w_2, \ldots, w_n)$. Przypuśćmy teraz, że została wykonana operacja $(O3.)$. Bez utraty ogólności możemy wówczas przyjąć, że $i = 1$ oraz $j = 2$. Wtedy $w_t = v_t$ dla każdego $t \in \{1, 2, \ldots, n\} \setminus \{1\}$ oraz $w_1 = v_1 + \lambda \circ v_2$ dla pewnego $\lambda \in K$. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Załóżmy, że $\sum_{i=1}^{n} \lambda_i \circ w_i = 0$. Wtedy $\lambda_1 \circ v_1 + (\lambda_2 + \lambda_1 \cdot \lambda) \circ v_2 + \lambda_3 \circ v_3 + \ldots + \lambda_n \circ v_n = 0$, więc liniowa niezależność układu $(v_1, v_2, \ldots, v_n)$ implikuje, że $\lambda_1 = \lambda_3 = \ldots = \lambda_n = 0$ oraz $\lambda_2 + \lambda_1 \cdot \lambda = 0$. Za-

tem również $\lambda_2 = 0$ i w konsekwencji układ $(w_1, w_2, \ldots, w_n)$ jest liniowo niezależny. Dla pozostałych $s \in \mathbb{N}$ teza wynika przez prostą indukcję.

Bezpośrednią konsekwencją Twierdzenia 8.1 i Twierdzenia 7.5 jest następujące

Twierdzenie 8.2. Niech $v_1, v_2, \ldots, v_n$ będą parami różnymi wektorami przestrzeni liniowej V nad ciałem K i niech $w_1, w_2, \ldots, w_n$ będą parami różnymi wektorami tej przestrzeni takimi, że układ $(w_1, w_2, \ldots, w_n)$ powstaje wskutek wykonania na układzie $(v_1, v_2, \ldots, v_n)$ kolejno skończonej liczby operacji elementarnych. Wówczas zbiór $\{w_1, w_2, \ldots, w_n\}$ jest bazą K-przestrzeni V wtedy i tylko wtedy, gdy zbiór $\{v_1, v_2, \ldots, v_n\}$ jest bazą tej przestrzeni.

Stwierdzenie 8.2. Niech X będzie liniowo niezależnym podzbiorem przestrzeni liniowej V nad ciałem K. Dla dowolnego $v \in V$ równoważne są warunki:

(i) $v \in \text{lin}(X)$;
(ii) $v \in X$ albo zbiór $X \cup \{v\}$ jest liniowo zależny.

Dowód. $(i) \Rightarrow (ii)$. Ze Stwierdzenia 7.4 wynika istnienie $n \in \mathbb{N}$, $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ oraz parami różnych $x_1, x_2, \ldots x_n \in X$ takich, że $v = \sum_{i=1}^{n} \lambda_i \circ x_i$. Zatem:

$$1 \circ v + (-\lambda_1) \circ x_1 + (-\lambda_2) \circ x_2 + \ldots + (-\lambda_n) \circ x_n = 0.$$

Jeśli więc $v \notin X$, to układ $(v, x_1, x_2, \ldots, x_n)$ jest liniowo zależny i w konsekwencji zbiór $X \cup \{v\}$ jest liniowo zależny. W ten sposób wykazaliśmy, że jeżeli $v \in \text{lin}(X)$, to $v \in X$ lub zbiór $X \cup \{v\}$ jest liniowo zależny. Ponieważ X jest liniowo niezależnym podzbiorem w V, to niemożliwe jest aby $v \in X$ i zbiór $X \cup \{v\}$ był liniowo zależny (bo wtedy $X \cup \{v\} = X$). Zatem warunek (i) implikuje warunek (ii).

$(ii) \Rightarrow (i)$. Jeżeli $v \in X$, to $v \in \text{lin}(X)$, gdyż $X \subseteq \text{lin}(X)$. Załóżmy teraz, że zbiór $X \cup \{v\}$ jest liniowo zależny. Z liniowej niezależności zbioru X wynika wówczas, że $v \notin X$ oraz istnieją $n \in \mathbb{N}$ i $x_1, x_2, \ldots, x_n \in X$ takie, że zbiór $\{v, x_1, x_2, \ldots, x_n\}$ jest liniowo zależny. Istnieją więc $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_n \in K$ takie, że $\lambda_0 \circ v + \sum_{i=1}^{n} \lambda_i \circ x_i = 0$ oraz $\lambda_j \neq 0$ dla pewnego $j \in \{0, 1, \ldots, n\}$. Jeżeli $\lambda_0 = 0$, to $\sum_{i=1}^{n} \lambda_i \circ x_i = 0$ oraz istnieje $i \in \{1, 2, \ldots, n\}$ takie, że $\lambda_i \neq 0$. Przeczy to liniowej niezależności zbioru X. Zatem $\lambda_0 \neq 0$, skąd $\lambda_0 \in K^*$ oraz $v = \sum_{i=1}^{n} \left(\lambda_0^{-1} \cdot \lambda_i\right) \circ x_i \in \text{lin}(X)$ (por. Stwierdzenie 7.3).

8.2 Baza przestrzeni liniowej

Definicja 8.5. Bazą przestrzeni liniowej nazywamy jej każdy maksymalny liniowo niezależny podzbiór.

Uwaga 8.5. Przez maksymalny liniowo niezależny podzbiór przestrzeni liniowej V nad ciałem K rozumiemy taki liniowo niezależny zbiór $X \subseteq V$, że dla dowolnego liniowo niezależnego podzbioru Y K-przestrzeni V inkluzja $X \subseteq Y$ implikuje równość $Y = X$.

Twierdzenie 8.3. Każdy liniowo niezależny podzbiór X_0 przestrzeni liniowej V nad ciałem K można rozszerzyć do bazy tej przestrzeni.

Dowód. Rozważmy dowolny liniowo niezależny podzbiór X_0 przestrzeni liniowej V nad ciałem K. Niech $\mathfrak{A}$ będzie rodziną wszystkich liniowo niezależnych podzbiorów w V zawierających zbiór X_0. Wtedy $\mathfrak{A} \neq \emptyset$, bo $X_0 \in \mathfrak{A}$. Ponadto zbiór $\mathfrak{A}$ jest częściowo uporządkowany przez inkluzję „$\subseteq$”. Rozważmy dowolny niepusty łańcuch $\mathfrak{L} \subseteq \mathfrak{A}$ i oznaczmy $Y_0 = \bigcup \mathfrak{L}$. Ponieważ $X_0 \subseteq Y$ dla każdego $Y \in \mathfrak{L}$, to $X_0 \subseteq Y_0$. Weźmy dowolne $v_1, v_2, \ldots, v_n \in Y_0$. Istnieją wówczas $Y_1, Y_2, \ldots Y_n \in \mathfrak{L}$ takie, że $v_i \in Y_i$ dla każdego $i \in \{1, 2, \ldots, n\}$. Ponadto $\mathfrak{L}$ jest łańcuchem, więc istnieje $j \in \{1, 2, \ldots, n\}$ takie że $Y_i \subseteq Y_j$ dla każdego $i \in \{1, 2, \ldots, n\}$. Zatem $v_1, v_2, \ldots, v_n \in Y_j$. Stąd oraz na mocy liniowej niezależności zbioru Y_j otrzymujemy, że zbiór $\{v_1, v_2, \ldots, v_n\}$ jest liniowo niezależny. Wobec tego Y_0 jest zbiorem liniowo niezależnym. Zatem $Y_0 \in \mathfrak{A}$. Ponadto $Y \subseteq Y_0$ dla każdego $Y \in \mathfrak{L}$. Stąd Y_0 jest majorantą (czyli ograniczeniem górnym) łańcucha $\mathfrak{L}$ w zbiorze $\mathfrak{A}$. Spełnione są więc założenia Lematu Kuratowskiego-Zorna i w konsekwencji w zbiorze $\mathfrak{A}$ istnieje element maksymalny X. W szczególności X jest bazą K-przestrzeni V zawierającą zbiór X_0.

Uwaga 8.6. Powyższe twierdzenie oznacza, że dla każdego liniowo niezależnego podzbioru X_0 przestrzeni liniowej V nad ciałem K istnieje baza X tej przestrzeni taka, że $X_0 \subseteq X$.

Ponieważ $\emptyset$ jest zbiorem liniowo niezależnym (zob. Uwaga 8.4), to z powyższego twierdzenia wynika natychmiast następujące

Twierdzenie 8.4. Każda przestrzeń liniowa posiada bazę.

Następne twierdzenie zawiera bardzo ważną charakteryzację bazy przestrzeni liniowej.

Twierdzenie 8.5. Podzbiór X przestrzeni liniowej V nad ciałem K jest bazą tej przestrzeni wtedy i tylko wtedy, gdy X jest zbiorem liniowo niezależnym generującym K-przestrzeń V.

Dowód. Przypuśćmy najpierw, że X jest bazą K-przestrzeni V. Wówczas zbiór X jest liniowo niezależny. Załóżmy nie wprost, że $\mathrm{lin}(X) \subsetneq V$. Istnieje wówczas $v \in V \setminus \mathrm{lin}(X)$. W szczególności wynika stąd, że $v \notin X$, więc Stwierdzenie 8.2 implikuje, że zbiór $X \cup \{v\}$ jest liniowo niezależny. Ale $X \subsetneq X \cup \{v\}$, wbrew maksymalności liniowo niezależnego zbioru X, sprzeczność.

Załóżmy teraz, że zbiór X jest liniowo niezależny i $\mathrm{lin}(X) = V$. Rozważmy dowolny liniowo niezależny podzbiór Y przestrzeni V taki, że $X \subseteq Y$. Jeżeli $Y \neq X$, to

istnieje $v \in Y \setminus X$ i z liniowej niezależności zbioru Y wynika liniowa niezależność zbioru $X \cup \{v\}$. Stąd oraz na mocy Stwierdzenia 8.2 otrzymujemy, że $v \notin \lin(X)$, czyli $v \notin V$, co jest niemożliwe, sprzeczność. Zatem $Y = X$ i w konsekwencji X jest bazą K-przestrzeni liniowej V.

Przykład 8.9. Z Przykładu 8.6, punktu (i) Stwierdzenia 5.8 oraz Twierdzenia 8.5 wynika, że dla dowolnej liczby naturalnej $n > 1$ oraz dowolnego ciała K zbiór $\big\{E_{ij} \colon i, j \in \{1, 2, \ldots, n\}\big\}$ jest bazą K-przestrzeni $\mathrm{M}_n(K)$.

Przykład 8.10. Niech K będzie ciałem. Weźmy dowolne $f \in K[x]$. Istnieją wówczas $n \in \mathbb{N}$ oraz $a_0, a_1, \ldots, a_n \in K$ takie, że $f = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$. Zatem $f \in \lin(1, x, x^2, \ldots)$ i w konsekwencji $K[x] = \lin(1, x, x^2, \ldots)$. Stąd oraz na mocy Przykładu 8.7 i Twierdzenia 8.5, zbiór $\{1, x, x^2, \ldots\}$ jest bazą K-przestrzeni liniowej $K[x]$.

Przykład 8.11. Z Przykładu 7.14 wynika, że dla dowolnej liczby naturalnej n zachodzi równość $K^n = \lin(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$. Stąd oraz na mocy Przykładu 8.1 i Twierdzenia 8.5 uzyskujemy, że $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ jest bazą przestrzeni K^n.

Definicja 8.6. Niech K będzie ciałem i niech $n \in \mathbb{N}$. Bazę $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ przestrzeni liniowej K^n nazywamy bazą kanoniczną tej przestrzeni.

Stwierdzenie 8.3. Niech $n \in \mathbb{N}$ i niech $v_1, v_2, \ldots, v_n$ będą wektorami przestrzeni liniowej V nad ciałem K. Jeżeli X jest maksymalnym liniowo niezależnym podzbiorem zbioru $\{v_1, v_2, \ldots, v_n\}$, to X jest bazą przestrzeni liniowej $\lin(v_1, v_2, \ldots, v_n)$.

Dowód. Niech $W = \lin(v_1, v_2, \ldots, v_n)$. Wtedy $\lin(X) \subseteq W$, gdyż $X \subseteq \{v_1, v_2, \ldots, v_n\}$. Weźmy dowolne $i \in \{1, 2, \ldots, n\}$. Jeżeli $v_i \in X$, to oczywiście $v_i \in \lin(X)$. Jeśli natomiast $v_i \notin X$, to z maksymalności liniowo niezależnego podzbioru X w zbiorze $\{v_1, v_2, \ldots, v_n\}$ wynika liniowa zależność zbioru $X \cup \{v_i\}$. Stąd oraz na mocy Twierdzenia 8.2, $v_i \in \lin(X)$. Zatem $\{v_1, v_2, \ldots, v_n\} \subseteq \lin(X)$, skąd $W \subseteq \lin(X)$ i w konsekwencji $\lin(X) = W$. Twierdzenie 8.5 implikuje więc, że X jest bazą przestrzeni liniowej $\lin(v_1, v_2, \ldots, v_n)$.

8.3 Baza uporządkowana

Twierdzenie 8.6. Niech $n \in \mathbb{N}$ i niech $v_1, v_2, \ldots, v_n$ będą parami różnymi wektorami przestrzeni liniowej V nad ciałem K. Wówczas zbiór $\{v_1, v_2, \ldots, v_n\}$ jest bazą K-przestrzeni V wtedy i tylko wtedy, gdy każdy wektor $v \in V$ można jednoznacznie zapisać w postaci $v = \lambda_1 \circ v_1 + \lambda_2 \circ v_2 + \ldots + \lambda_n \circ v_n$ dla pewnych $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$.

Dowód. Załóżmy najpierw, że zbiór $\{v_1, v_2, \ldots, v_n\}$ jest bazą K-przestrzeni V. Wtedy $V = \lin(v_1, v_2, \ldots, v_n)$, więc dla dowolnego $v \in V$, istnieją $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ takie, że

$v = \sum_{i=1}^{n} \lambda_i \circ v_i$ (zob. punkt (ii) Wniosku 7.3). Rozważmy dowolne $\mu_1, \mu_2, \ldots, \mu_n \in K$. Przypuśćmy, że $v = \sum_{i=1}^{n} \mu_i \circ v_i$. Wówczas $0 = v - v = \sum_{i=1}^{n} \lambda_i \circ v_i - \sum_{i=1}^{n} \mu_i \circ v_i = \sum_{i=1}^{n} (\lambda_i - \mu_i) \circ v_i$, więc z liniowej niezależności zbioru $\{v_1, v_2, \ldots, v_n\}$ wynika, że $\lambda_i - \mu_i = 0$, czyli $\mu_i = \lambda_i$ dla każdego $i \in \{1, 2, \ldots, n\}$.

Na odwrót. Załóżmy, że każdy wektor $v \in V$ można jednoznacznie zapisać w postaci $v = \lambda_1 \circ v_1 + \lambda_2 \circ v_2 + \ldots + \lambda_n \circ v_n$ dla pewnych $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Wówczas $V = \mathrm{lin}(v_1, v_2, \ldots, v_n)$. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Przypuśćmy, że $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$. Ponieważ $\sum_{i=1}^{n} 0 \circ v_i = 0$, to z przyjętego założenia wynika, że zbiór $\{v_1, v_2, \ldots, v_n\}$ jest liniowo niezależny. Zatem jest on bazą K-przestrzeni V.

Dla skończenie wymiarowych przestrzeni liniowych, powyższe twierdzenie motywuje następującą definicję bazy uporządkowanej.

Definicja 8.7. Niech $n \in \mathbb{N}$ i niech $\{v_1, v_2, \ldots, v_n\}$ będzie n-elementową bazą przestrzeni liniowej V nad ciałem K. Wówczas układ wektorów $(v_1, v_2, \ldots, v_n)$ nazywamy bazą uporządkowaną K-przestrzeni V. Jeżeli $v = \sum_{i=1}^{n} \lambda_i \circ v_i$, gdzie $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$, to ciąg $(\lambda_1, \lambda_2, \ldots \lambda_n)$ elementów ciała K nazywamy ciągiem współrzędnych wektora v w bazie uporządkowanej $(v_1, v_2, \ldots, v_n)$. Dla każdego $i \in \{1, 2, \ldots, n\}$ skalar λ_i nazywamy i-tą współrzędną wektora v w bazie uporządkowanej $(v_1, v_2, \ldots, v_n)$.

Uwaga 8.7. Ciąg współrzędnych $(\lambda_1, \lambda_2, \ldots \lambda_n)$ wektora v przestrzeni liniowej V nad ciałem K w bazie uporządkowanej $(v_1, v_2, \ldots, v_n)$ możemy wzajemnie jednoznacznie utożsamić z wektorem $[\lambda_1, \lambda_2, \ldots \lambda_n]$ przestrzeni liniowej K^n.

Definicja 8.8. Niech $n \in \mathbb{N}$ i niech $(\lambda_1, \lambda_2, \ldots \lambda_n)$ będzie ciągiem współrzędnych wektora v w bazie uporządkowanej $(v_1, v_2, \ldots, v_n)$ przestrzeni liniowej V nad ciałem K. Wektor $[\lambda_1, \lambda_2, \ldots \lambda_n]$ przestrzeni K^n nazywamy wówczas wektorem współrzędnych wektora v w bazie uporządkowanej $(v_1, v_2, \ldots, v_n)$.

Przykład 8.12. Bazami uporządkowanymi przestrzeni liniowej $\mathrm{M}_2(\mathbb{R})$ są np. $\mathscr{B}_1 = (E_{11}, E_{12}, E_{21}, E_{22})$ oraz $\mathscr{B}_2 = (E_{11}, E_{21}, E_{12}, E_{22})$. Niech $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Wtedy $A = 1 \cdot E_1 1 + 2 \cdot E_{12} + 3 \cdot E_{21} + 4 \cdot E_{22}$ oraz $A = 1 \cdot E_1 1 + 3 \cdot E_{21} + 2 \cdot E_{12} + 4 \cdot E_{22}$, więc $[1, 2, 3, 4] \in \mathbb{R}^4$ jest wektorem współrzędnych macierzy A w bazie $\mathscr{B}_1$, zaś $[1, 3, 2, 4] \in \mathbb{R}^4$ jest wektorem współrzędnych macierzy A w bazie $\mathscr{B}_2$.

8.4 Wymiar przestrzeni liniowej

Twierdzenie 8.7 (Steinitza o wymianie). Niech n i s będą nieujemnymi liczbami całkowitymi. Jeżeli B jest bazą przestrzeni liniowej V nad ciałem K, $|B| = n$ oraz X jest s-elementowym liniowo niezależnym podzbiorem w V, to $s \leq n$ oraz istnieje $(n-s)$-elementowy podzbiór Y zbioru B taki, że zbiór $X \cup Y$ jest bazą K-przestrzeni V.

Dowód. Przeprowadzimy dowód indukcyjny względem s. Jeżeli $s = 0$, to $X = \emptyset$ i wystarczy przyjąć $Y = B$. Załóżmy indukcyjnie, że twierdzenie jest prawdziwe dla wszystkich nieujemnych liczb całkowitych r nie większych niż pewna nieujemna liczba całkowita s. Rozważmy sytuację, w której $|X| = s+1$. Wówczas $X = \{x_1, x_2, \ldots, x_{s+1}\}$ dla pewnych parami różnych $x_1, x_2, \ldots, x_{s+1} \in V$ oraz układ $(x_1, x_2, \ldots, x_s)$ jest liniowo niezależny. Z założenia indukcyjnego wynika więc nierówność $s \leq n$ oraz istnienie $(n-s)$-elementowego podzbioru Y_0 zbioru B takiego, że $\{x_1, x_2, \ldots, x_s\} \cup Y_0$ jest bazą K-przestrzeni V. Przypuśćmy, że $s = n$. Wówczas $Y_0 = \emptyset$, skąd wynika, że $\{x_1, x_2, \ldots, x_s\}$ jest bazą K-przestrzeni V. Ale wtedy $x_{s+1} \in \mathrm{lin}(x_1, x_2, \ldots, x_s)$ na mocy Twierdzenia 8.5. Stwierdzenie 8.2 implikuje więc, liniową zależność zbioru $\{x_{s+1}\} \cup \{x_1, x_2, \ldots, x_s\} = X$, sprzeczność. Wobec tego $s < n$, czyli $s+1 \leq n$. W szczególności wynika stąd istnienie $b_1, b_2, \ldots, b_{n-s} \in B$ takich, że $Y_0 = \{b_1, b_2, \ldots, b_{n-s}\}$. Ponieważ $\{x_1, x_2, \ldots, x_s, b_1, b_2, \ldots, b_{n-s}\}$ jest bazą K-przestrzeni V oraz zbiór X jest liniowo niezależny, to Twierdzenie 8.5 wraz ze Stwierdzeniem 8.2 implikują istnienie takich $\lambda_1, \lambda_2, \ldots, \lambda_s, \mu_1, \mu_2, \ldots, \mu_{n-s} \in K$, że $x_{s+1} = \sum_{i=1}^{s} \lambda_i \circ x_i + \sum_{i=1}^{n-s} \mu_i \circ b_i$ oraz $\mu_i \neq 0$ dla pewnego $i \in \{1, 2, \ldots, n-s\}$. Bez utraty ogólności możemy przyjąć, że $\mu_1 \neq 0$. Wówczas układ $(x_1, x_2, \ldots, x_{s+1}, b_2, \ldots, b_{n-s})$ powstaje z liniowo niezależnego układu wektorów $(x_1, x_2, \ldots, x_s, b_1, b_2, \ldots, b_{n-s})$ wskutek wykonania na nim kolejno skończonej liczby operacji elementarnych. Stąd oraz na mocy Twierdzenia 8.2, zbiór $X \cup \big(Y \setminus \{b_1\}\big)$ jest bazą K-przestrzeni V. Ponadto $\big|Y_0 \setminus \{b_1\}\big| = |Y_0| - 1 = (n-s) - 1 = n - (s+1)$, więc wystarczy przyjąć $Y = Y_0 \setminus \{b_1\}$.

Wniosek 8.1. Jeżeli B jest bazą przestrzeni liniowej V nad ciałem K i $|B| = n$ dla pewnego $n \in \mathbb{N}_0$, to każda baza tej przestrzeni zawiera dokładnie n-elementów.

Dowód. Rozważmy dowolną bazę A K-przestrzeni V. Załóżmy nie wprost, że $|A| > n$. Istnieje wówczas $n+1$-elementowy podzbiór zbioru A i jest on oczywiście liniowo niezależnym podzbiorem w V. Z Twierdzenia 8.7 wynika więc, że $n+1 \leq n$, sprzeczność. Zatem $|A| \leq n$. Ponadto A jest bazą K-przestrzeni V oraz B jest n-elementowym liniowo niezależnym podzbiorem w V, więc powołując się ponownie na Twierdzenie 8.7 uzyskujemy, że $|B| \leq |A|$, czyli $n \leq |A|$. Zatem $|A| = n$.

Uwaga 8.8. Zauważmy, że zbiory X i Y opisane w Twierdzeniu 8.7 są rozłączne. Istotnie, zachowując wszystkie pozostałe oznaczenia ze wspomnianego twierdzenia i powołując się na Wniosek 8.1 otrzymujemy, że $|X| = s$, $|Y| = n - s$ oraz $|X \cup Y| = |B| = n$. Ponadto, ze Wstępu do teorii mnogości wiadomo, że $|X \cup Y| = |X| + |Y| - |X \cap Y|$. Zatem $|X \cap Y| = 0$, czyli $X \cap Y = \emptyset$.

Wniosek 8.2. Jeżeli w przestrzeni liniowej V nad ciałem K istnieje nieskończony liniowo niezależny podzbiór X, to każda baza B tej przestrzeni jest nieskończona.

Dowód. Załóżmy nie wprost, że w V istnieje baza B taka, że $|B| < \infty$. Ponieważ $|X| = \infty$, to istnieje podzbiór Y zbioru X taki, że $|Y| = |B| + 1$. Stąd oraz na mocy Twierdzenia 8.7, $|B| + 1 < |B|$, sprzeczność.

Uwaga 8.9. Opierając się na arytmetyce liczb kardynalnych, Wniosek 8.1 można uogólnić na przypadek nieskończony i w ten sposób wzmocnić tezę podaną we wniosku 8.2. Można mianowicie udowodnić, że w dowolnej przestrzeni liniowej, każda baza ma tę samą liczbę elementów.

Powyższa uwaga pozwala wprowadzić pojęcie wymiaru przestrzeni liniowej.

Definicja 8.9. Wymiarem przestrzeni liniowej V nad ciałem K nazywamy moc jej dowolnej bazy (czyli liczbę elementów w dowolnej bazie). Tę liczbę kardynalną oznaczamy przez $\dim_K V$ lub krótko: $\dim V$, gdy z kontekstu jasno wynika, że przestrzeń V rozważana jest nad ciałem K.

Uwaga 8.10. Notacji $\dim_K V = \infty$ lub $\dim V = \infty$ używamy, gdy chcemy podkreślić, że wymiar K-przestrzeni liniowej V jest nieskończony i jednocześnie nie ma potrzeby dokładnego jego określenia.

Przykład 8.13. Ponieważ dla dowolnego ciała K, $\emptyset$ jest bazą K-przestrzeni $\{0\}$, to $\dim_K\{0\} = 0$.

Przykład 8.14. Niech a będzie dowolnym niezerowym elementem ciała K. Wówczas zbiór $\{a\}$ jest liniowo niezależny nad K oraz $\operatorname{lin}_K(a) = K$. Zatem $\{a\}$ jest bazą przestrzeni liniowej K nad K. Stąd $\dim_K K = 1$.

Przykład 8.15. Ponieważ $\{1, i\}$ jest bazą przestrzeni liniowej $\mathbb{C}$ nad ciałem $\mathbb{R}$, to $\dim_{\mathbb{R}} \mathbb{C} = 2$. Z Przykładu 8.14 wynika, że $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Przykład 8.16. Z Przykładu 8.10 wynika, że $\dim K[x] = \infty$ dla dowolnego ciała K. Dokładniej, opierając się na arytmetyce liczb kardynalnych można uzasadnić, że $\dim K[x] = \aleph_0$, gdzie $\aleph_0 = |\mathbb{N}|$.

Przykład 8.17. Z Przykładu 8.11 wynika, że dla dowolnej liczby naturalnej n oraz dowolnego ciała K, $\dim K^n = n$.

Przykład 8.18. Z Przykładu 7.9 wynika, że $\mathbb{R}$ jest $\mathbb{Q}$-przestrzenią liniową. Ponadto Ferdinand Lindemann w roku 1882 udowodnił, że nie istnieje $f \in \mathbb{Q}[x]$ takie, że $f \neq 0$ oraz $f(\pi) = 0$. Zatem $\{\pi, \pi^2, \pi^3, \ldots\}$ jest nieskończonym podzbiorem w $\mathbb{R}$ liniowo niezależnym nad $\mathbb{Q}$, skąd $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.

Przykład 8.19. Dla dowolnej liczby naturalnej n i dowolnego ciała K, $\dim \mathrm{M}_n(K) = n^2$. Istotnie, dla $n = 1$ K-przestrzeń $\mathrm{M}_n(K)$ możemy utożsamić z K-przestrzenią K, więc z Przykładu 8.14 wynika, że $\dim \mathrm{M}_n(K) = 1 = 1^2$. Jeżeli $n > 1$, to na mocy Przykładu 8.9 otrzymujemy, że $\dim \mathrm{M}_n(K) = \left|\left\{E_{ij} \colon i, j \in \{1, 2, \ldots, n\}\right\}\right| = \left|\{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}\right| = n^2$.

Bezpośrednią konsekwencją powyższego Stwierdzenia 8.3 jest następujący

Wniosek 8.3. Dla dowolnych wektorów $v_1, v_2, \ldots, v_n$ przestrzeni liniowej V nad ciałem K, $\dim \mathrm{lin}_K(v_1, v_2, \ldots, v_n) \leq n$. Ponadto $\dim \mathrm{lin}_K(v_1, v_2, \ldots, v_n) = n$ wtedy i tylko wtedy, gdy układ $(v_1, v_2, \ldots, v_n)$ jest liniowo niezależny.

Możemy teraz z łatwością udowodnić ważne kryterium pozwalające weryfikować, czy n-elementowy podzbiór skończenie wymiarowej przestrzeni liniowej wymiaru n jest jej bazą. Mamy bowiem następujące

Twierdzenie 8.8. Niech n będzie liczbą naturalną i niech V będzie n-wymiarową przestrzenią liniową nad ciałem K. Dla dowolnych $v_1, v_2, \ldots, v_n \in V$ następujące warunki są równoważne:

(i) zbiór $\{v_1, v_2, \ldots, v_n\}$ jest bazą K-przestrzeni V;
(ii) $\{v_1, v_2, \ldots, v_n\}$ jest n-elementowym zbiorem liniowo niezależnym nad K;
(iii) $\mathrm{lin}_K(v_1, v_2, \ldots, v_n) = V$.

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

$(ii) \Rightarrow (iii)$. Wynika natychmiast z Twierdzeń 8.4, 8.7 i 8.5.

$(iii) \Rightarrow (i)$. Załóżmy, że $\mathrm{lin}_K(v_1, v_2, \ldots, v_n) = V$. Wtedy $\dim \mathrm{lin}_K(v_1, v_2, \ldots, v_n) = \dim_K V = n$, więc z Wniosku 8.3 wynika, że układ $(v_1, v_2, \ldots, v_n)$ jest liniowo niezależny. Zatem zbiór $\{v_1, v_2, \ldots, v_n\}$ jest baza K-przestrzeni V.

Poniższe twierdzenie opisuje ważne, zgodne z intuicją własności wymiaru podprzestrzeni skończenie wymiarowej przestrzeni liniowej.

Twierdzenie 8.9. Jeżeli W jest podprzestrzenią skończenie wymiarowej przestrzeni liniowej V nad ciałem K, to $\dim_K W \leq \dim_K V$. Ponadto $\dim_K W = \dim_K V$ wtedy i tylko wtedy, gdy $W = V$.

Dowód. Niech B_V i B_W oznaczają odpowiednio bazy K-przestrzeni V i W. Wtedy $|B_V| = \dim_K V < \infty$, więc $|B_W| < \infty$ na mocy Wniosku 8.2. Twierdzenie 8.7 implikuje więc, że $\dim_K W \leq \dim_K V$.

Jasne jest, że równość K-przestrzeni V i W implikuje równość ich wymiarów nad K. Załóżmy teraz, że $\dim_K W = \dim_K V$. Z Twierdzenia 8.7 wynika wówczas istnienie takiego podzbioru Y bazy B_V, że $B_W \cup Y$ jest bazą K-przestrzeni V i $|Y| = \dim_K V - \dim_K W = 0$. Zatem $Y = \emptyset$ i w konsekwencji B_W jest bazą V. Stąd $V = \mathrm{lin}(B_W) = W$.

Twierdzenie 8.10. Niech W i U będą skończenie wymiarowymi podprzestrzeniami przestrzeni liniowej V nad ciałem K. Wówczas skończenie wymiarowe są również podprzestrzenie $W + U$ i $W \cap U$. Ponadto:

$$\dim(W + U) = \dim W + \dim U - \dim(W \cap U). \tag{8.4.1}$$

Dowód. Ponieważ $W \cap U$ jest podprzestrzenią skończenie wymiarowej K-przestrzeni W, to z Twierdzenia 8.9 wynika, że $\dim(W \cap U) < \infty$. Niech B oznacza bazę K-przestrzeni $W \cap U$. Twierdzenie 8.7 implikuje istnienie zbiorów $X \subseteq W$ i $Y \subseteq U$ takich, że $B \cup X$ jest bazą K-przestrzeni W, $B \cup Y$ jest bazą K-przestrzeni U, oraz $|X| = |B_W| - |B|$ i $|Y| = |B_U| - |B|$, przy czym B_W i B_U oznaczają odpowiednio bazy K-przestrzeni W i U. Zatem:

$$\dim(W \cap U) = |B|,\ \dim W = |B_W| = |X| + |B| \text{ i } \dim U = |B_U| = |Y| + |B|. \quad (8.4.2)$$

Wystarczy więc pokazać, że:

$$\dim(W + U) = |B| + |X| + |Y|. \quad (8.4.3)$$

Na mocy Wniosku 7.5 uzyskujemy, że:

$$W + U = \operatorname{lin}(B \cup X) + \operatorname{lin}(B \cup Y) = \operatorname{lin}\big((B \cup X) \cup (B \cup Y)\big) = \operatorname{lin}(B \cup X \cup Y). \quad (8.4.4)$$

Jeżeli $X = \emptyset$, to $W + U = \operatorname{lin}(B \cup Y) = U$ i $|X| = 0$, więc równość (8.4.3) jest prawdziwa na mocy (8.4.2). Jeśli $Y = \emptyset$, to $W + U = \operatorname{lin}(B \cup X) = U$ i $|Y| = 0$. Wtedy, powołując się na (8.4.2) ponownie otrzymujemy (8.4.3). Niech dalej $X \neq \emptyset$ i $Y \neq \emptyset$. Istnieją wówczas $m, n \in \mathbb{N}$ takie, że $|X| = m$ i $|Y| = n$. Istnieją więc także $x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \in V$ takie, że $X = \{x_1, x_2, \ldots, x_m\}$ oraz $Y = \{y_1, y_2, \ldots, y_n\}$. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_n \in K$ i oznaczmy $x = \sum_{i=1}^{m} \lambda_i \circ x_i$ oraz $y = \sum_{i=1}^{n} \mu_i \circ y_i$.

Przypuśćmy, że $B = \emptyset$. Wtedy $|B| = 0$ oraz warunek $x + y = 0$ implikuje, że $x = -y \in \operatorname{lin}(X) \cap \operatorname{lin}(Y) \subseteq W \cap U = \operatorname{lin}(B) = \{0\}$, co wobec liniowej niezależności zbioru X oznacza, że $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$ i w konsekwencji również $\mu_1 = \mu_2 = \ldots = \mu_n = 0$. Zatem układ $(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n)$ jest liniowo niezależny. Ponadto $B \cup X \cup Y = \{x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n\}$. Stąd oraz na mocy (8.4.4) zbiór $B \cup X \cup Y$ jest bazą K-przestrzeni $W + U$ taką, że $|B \cup X \cup Y| = 0 + m + n$, co oznacza, że zachodzi równość (8.4.3).

Załóżmy teraz, że $B \neq \emptyset$. Wtedy $|B| = r$ dla pewnego $r \in \mathbb{N}$, więc istnieją $b_1, b_2 \ldots, b_r \in V$ takie, że $B = \{b_1, b_2, \ldots, b_r\}$. Weźmy dowolne $\beta_1, \beta_2 \ldots, \beta_r \in K$ i oznaczmy $b = \sum_{i=1}^{r} \beta_i \circ b_i$. Przypuśćmy, że $b + x + y = 0$. Wtedy $y = -(b + x) \in \operatorname{lin}(B \cup X) \cap \operatorname{lin}(Y) = W \cap \operatorname{lin}(Y) \subseteq W \cap U = \operatorname{lin}(B)$. Zatem $y = \sum_{i=1}^{r} \alpha_i \circ b_i$, czyli $\sum_{i=1}^{n} \mu_i \circ y_i = \sum_{i=1}^{r} \alpha_i \circ b_i$. Stąd $0 = \sum_{i=1}^{r} \alpha_i \circ b_i - \sum_{i=1}^{n} \mu_i \circ y_i = \sum_{i=1}^{r} \alpha_i \circ b_i + \sum_{i=1}^{n} (-\mu_i) \circ y_i$, więc z liniowej niezależności zbioru $B \cup Y$ wynika, że $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$ oraz $\mu_1 = \mu_2 = \ldots \mu_n = 0$. Stąd $b = 0$ oraz $y = 0$. Zatem także $x = 0$, co wobec liniowej niezależności zbioru X oznacza, że $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$. Wobec tego układ $(b_1, b_2, \ldots, b_r, x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n)$ jest liniowo niezależny. Ponadto $B \cup X \cup Y = \{b_1, b_2, \ldots, b_r, x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n\}$. Stąd oraz na mocy (8.4.4) zbiór $B \cup X \cup Y$ jest bazą K-przestrzeni $W + U$ taką, że $|B \cup X \cup Y| = r + m + n$, co oznacza, że zachodzi równość (8.4.3).

Przykład 8.20. Dla podprzestrzeni $W = \operatorname{lin}\big([0,1,3,0],[1,1,1,0],[1,2,4,0]\big)$ oraz $U = \operatorname{lin}\big([2,0,2,4],[0,0,1,0]\big)$ przestrzeni liniowej $\mathbb{R}^4$ wyznaczymy bazy i wymiary podprzestrzeni $W+U$ i $W\cap U$. Wyznaczymy najpierw bazy przestrzeni W i U. Zauważmy, że dowolnej operacji elementarnej wykonanej na układzie wektorów:

$$\big([0,1,3,0],[1,1,1,0],[1,2,4,0]\big)$$

wzajemnie jednoznacznie odpowiada analogiczna operacja elementarna wykonywana na wierszach macierzy:

$$\begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 0 \end{bmatrix}.$$

Mamy:

$$\begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 0 \end{bmatrix} \overset{w_3-(w_1+w_2)}{\equiv} \begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \overset{w_2-w_1}{\equiv} \begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Stąd oraz na mocy Twierdzenia 7.5, $W = \operatorname{lin}\big([0,1,3,0],[1,0,-2,0]\big)$. Weźmy dowolne $\alpha,\beta \in K$. Załóżmy, że $\alpha \circ [0,1,3,0] + \beta \circ [1,0,-2,0] = [0,0,0,0]$. Wtedy $[\beta,\alpha,3\alpha-2\beta,0] = [0,0,0,0]$, więc $\alpha = \beta = 0$. Zatem wektory $[0,1,3,0]$ oraz $[1,0,-2,0]$ są liniowo niezależne. Stąd oraz na mocy Twierdzenia 8.5 zbiór:

$$B_W = \{[0,1,3,0],[1,0,-2,0]\}$$

jest bazą W nad K. Zatem $\dim W = 2$. Ponieważ:

$$\begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix} \overset{\frac{1}{2}\cdot w_1}{\equiv} \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \overset{w_1-w_2}{\equiv} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

to $U = \operatorname{lin}\big([1,0,0,2],[0,0,1,0]\big)$. Ponadto warunek:

$$\alpha \circ [1,0,0,2] + \beta \circ [0,0,1,0] = [0,0,0,0]$$

implikuje równość $[\alpha,0,\beta,2\alpha] = [0,0,0,0]$, z której wynika, że $\alpha = \beta = 0$. Zatem wektory $[1,0,0,2]$ i $[0,0,1,0]$ są liniowo niezależne i w konsekwencji zbiór $B_U = \{[1,0,0,2],[0,0,1,0]\}$ jest bazą K-przestrzeni U oraz $\dim U = 2$.

Wyznaczymy teraz bazę i wymiar K-przestrzeni $W+U$. Na mocy wcześniejszych rozważań oraz Twierdzenia 7.3 uzyskujemy, że:

$$W+U = \operatorname{lin}\big([0,1,3,0],[1,0,-2,0]\big) + \operatorname{lin}\big([1,0,0,2],[0,0,1,0]\big) =$$

$$= \operatorname{lin}\big([0,1,3,0],[1,0,-2,0],[1,0,0,2],[0,0,1,0]\big).$$

Ponadto:

$$\begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 0 & -2 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} w_1 - 3 \cdot w_4 \\ w_2 + 2 \cdot w_4 \\ \equiv \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} w_3 - w_2 \\ \equiv \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} \frac{1}{2} \cdot w_3 \\ \equiv \end{matrix}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \text{ skąd:}$$

$$W + U = \operatorname{lin}(\varepsilon_2, \varepsilon_1, \varepsilon_4, \varepsilon_3) = \operatorname{lin}(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) = \mathbb{R}^4.$$

Zatem $\dim(W+U) = 4$. Z Twierdzenia 8.10 wynika więc, że $\dim(W \cap U) = \dim(W+U) - \dim W - \dim U = 4 - 2 - 2 = 0$. Wobec tego bazą K-przestrzeni U jest $\emptyset$ (i $U = \{0\}$).

Bazę i wymiar K-przestrzeni $W \cap U$ można wyznaczyć także nie odwołując się do Twierdzenia 8.10. Istotnie, weźmy dowolne $v \in W \cap U$. Wtedy $v \in W$ oraz $v \in U$. Ponieważ $W = \operatorname{lin}\big([0,1,3,0],[1,0,-2,0]\big)$ i $U = \operatorname{lin}\big([1,0,0,2],[0,0,1,0]\big)$, to istnieją $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in K$ takie, że $v = \lambda_1 \circ [0,1,3,0] + \lambda_2 \circ [1,0,-2,0]$ i $v = \lambda_3 \circ [1,0,0,2] + \lambda_4 \circ [0,0,1,0]$. Zatem $\lambda_1 \circ [0,1,3,0] + \lambda_2 \circ [1,0,-2,0] = \lambda_3 \circ [1,0,0,2] + \lambda_4 \circ [0,0,1,0]$, czyli $\lambda_1 \circ [0,1,3,0] + \lambda_2 \circ [1,0,-2,0] - \lambda_3 \circ [1,0,0,2] - \lambda_4 \circ [0,0,1,0] = [0,0,0,0]$, co oznacza, że $[\lambda_2 - \lambda_3, \lambda_1, 3\lambda_1 - 2\lambda_2 - \lambda_4, -2\lambda_3] = [0,0,0,0]$. Otrzymujemy stąd jednorodny układ równań liniowych:

$$\begin{cases} \lambda_2 - \lambda_3 = 0 \\ \lambda_1 = 0 \\ 3\lambda_1 - 2\lambda_2 - \lambda_4 = 0 \\ -2\lambda_3 = 0 \end{cases}. \tag{8.4.5}$$

Łatwo sprowadzić go do równoważnej postaci:

$$\begin{cases} \lambda_1 = 0 \\ \lambda_3 = 0 \\ \lambda_3 = 0 \\ \lambda_4 = 0 \end{cases}.$$

Zatem $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0,0,0,0)$ jest jedynym rozwiązaniem układu (8.4.5). Wobec tego $v = [0,0,0,0]$, skąd $W \cap U \subseteq \{0\}$. Po uwzględnieniu oczywistej inkluzji przeciwnej otrzymujemy więc równość $W \cap U = \{0\}$. Zatem bazą K-przestrzeni $W \cap U$ jest $\emptyset$ i w konsekwencji $\dim(W \cap U) = 0$.

Rozdział 9

Rząd macierzy

9.1 Rząd wierszowy oraz rząd kolumnowy macierzy

Uwaga 9.1. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$. Wówczas wiersze macierzy A możemy w naturalny sposób traktować jak wektory przestrzeni liniowej K^n, zaś kolumny macierzy A możemy analogicznie utożsamić z wektorami przestrzeni K^m, przy czym te wektory będziemy zazwyczaj utożsamiali z $(m\times 1)$-macierzami; tzn., dla każdego $j \in \{1,2,\ldots,n\}$ dokonujemy utożsamienia j-tej kolumny macierzy A z wektorem $[a_{1j},a_{2j},\ldots,a_{mj}] \in K^m$, który z kolei utożsamiamy z macierzą:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in M_{m\times 1}(K).$$

Macierze tej postaci nazywa się często wektorami kolumnowymi przestrzeni K^m.

Definicja 9.1. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$. Rzędem wierszowym macierzy A nazywamy wymiar przestrzeni liniowej:

$$\operatorname{lin}_K\big([a_{11},a_{12},\ldots,a_{1n}],[a_{21},a_{22},\ldots,a_{2n}],\ldots,[a_{m1},a_{m2},\ldots,a_{mn}]\big).$$

Rzędem kolumnowym macierzy A nazywamy wymiar przestrzeni liniowej:

$$\operatorname{lin}_K\left(\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}\right).$$

Rząd wierszowy i rząd kolumnowy macierzy A oznaczamy odpowiednio przez $r_w(A)$ i $r_k(A)$.

Uwaga 9.2. Z powyższej definicji oraz Stwierdzenia 8.3 wynika, że liczba $r_w(A)$ równa jest maksymalnej liczbie liniowo niezależnych wierszy macierzy A (rozumianych jako wektory przestrzeni K^n). Analogicznie, liczba $r_k(A)$ równa jest maksymalnej liczbie liniowo niezależnych kolumn macierzy A (rozumianych jako wektory przestrzeni K^m).

Omówimy teraz szereg własności rzędu wierszowego i rzędu kolumnowego macierzy oraz związki między tymi pojęciami pozwalające zdefiniować rząd macierzy (zob. Definicja 9.2). Na początku zauważmy, że bezpośrednią konsekwencją Definicji 9.1, Uwagi 9.2 oraz Twierdzeń 8.5 i 8.2 jest następujący

Lemat 9.1. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A \in M_{m\times n}(K)$. Rząd wierszowy (kolumnowy) macierzy A nie ulega zmianie po zastosowaniu dowolnej operacji elementarnej na wierszach (kolumnach) macierzy A.

Lemat 9.2. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$. Niech ponadto $\lambda \in K$ oraz i_0, j_0 będą różnymi elementami zbioru $\{1,2,\ldots m\}$. Jeżeli B jest macierzą powstałą z macierzy wskutek wykonania na wierszach macierzy A operacji elementarnej $w_{i_0} + \lambda \cdot w_{j_0}$, to $r_k(B) = r_k(A)$.

Dowód. Bez utraty ogólności możemy przyjąć, że $i = 1$ oraz $j = 2$. Jeżeli $A = \Theta$, to teza jest oczywista. Niech dalej $A \neq \Theta$. Wtedy co najmniej jedna kolumna macierzy A jest niezerowa, więc $r_k(A) \in \{1,2,\ldots,n\}$. Oznaczmy $r = r_k(A)$. Z Lematu 9.1 wynika, że bez zmniejszania ogólności możemy przyjąć, iż liniowo niezależnych jest r pierwszych kolumn macierzy A. Wykażemy, że wówczas liniowo niezależny jest układ wektorów:

$$\left(\begin{bmatrix} a_{11}+\lambda\cdot a_{21} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12}+\lambda\cdot a_{22} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1r}+\lambda\cdot a_{2r} \\ a_{2r} \\ \vdots \\ a_{mr} \end{bmatrix} \right). \tag{9.1.1}$$

W tym celu rozważmy dowolne $x_1, x_2, \ldots, x_r \in K$ i załóżmy, że:

$$\sum_{j=1}^{r} x_j \circ \begin{bmatrix} a_{1j}+\lambda\cdot a_{2j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Wtedy:

$$\begin{cases} \sum_{j=1}^{r} (a_{1j}+\lambda\cdot a_{2j})\cdot x_j = 0 \\ \sum_{j=1}^{r} a_{2j}\cdot x_j = 0 \\ \vdots \\ \sum_{j=1}^{r} a_{mj}\cdot x_j = 0 \end{cases}.$$

Wykonując na powyższym układzie równań operację elementarną $r_1 - \lambda \cdot r_2$ uzyskujemy układ:

$$\begin{cases} \sum_{j=1}^{r} a_{1j} \cdot x_j = 0 \\ \sum_{j=1}^{r} a_{2j} \cdot x_j = 0 \\ \vdots \\ \sum_{j=1}^{r} a_{mj} \cdot x_j = 0 \end{cases},$$

który można zapisać w postaci zerowej kombinacji liniowej r pierwszych kolumn macierzy A:

$$\sum_{j=1}^{r} x_j \circ \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Liniowa niezależność tych kolumn implikuje więc, że $x_1 = x_2 = \ldots = x_r = 0$, co dowodzi liniowej niezależności układu wektorów (9.1.1). Zatem $r_k(B) \geq r_k(A)$. Ponieważ macierz B powstaje z macierzy A wskutek wykonania na wierszach macierzy B operacji elementarnej $w_1 + (\lambda) \cdot w_2$, to powtarzając powyższe rozumowanie uzyskujemy, że $r_k(A) \geq r_k(B)$. Wobec tego $r_k(B) = r_k(A)$.

Ponieważ $r_w(A) = r_k\left(A^T\right)$ dla dowolnej macierzy $A \in M_{m\times n}(K)$, to prawdziwy jest także lemat dualny do Lematu 9.2:

Lemat 9.3. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A \in M_{m\times n}(K)$. Niech ponadto $\lambda \in K$ oraz i, j będą różnymi elementami zbioru $\{1,2,\ldots n\}$. Jeżeli B jest macierzą powstałą z macierzy wskutek wykonania na kolumnach macierzy A operacji elementarnej $k_i + \lambda \cdot k_j$, to $r_w(B) = r_w(A)$.

Lemat 9.4. Niech K będzie ciałem i niech $m,n > 1$ będą liczbami naturalnymi. Jeżeli macierz $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$ spełnia warunki: $a_{uv} \neq 0$ dla pewnych $u \in \{1,2,\ldots,m\}$ i $v \in \{1,2,\ldots,n\}$, oraz $a_{iv} = a_{uj} = 0$ dla wszystkich $i \in \{1,2,\ldots,m\}$, $j \in \{1,2,\ldots,n\}$ takich, że $i \neq u$ oraz $j \neq v$, to $r_w(A) = 1 + r_w(A_{uv})$ i $r_k(A) = 1 + r_k(A_{uv})$.

Dowód. Ponieważ $r_k(A) = r_w\left(A^T\right)$, to wystarczy wykazać, że $r_w(A) = 1 + r_w(A_{uv})$. Oznaczmy $r = r_w(A_{uv})$. Jeżeli $r = 0$, to $A_{uv} = \Theta$, więc tylko u-ty wiersz macierzy A jest niezerowy. Zatem $r_w(A) = 1$ i żądana równość zachodzi. Załóżmy teraz, że $r > 0$. Istnieje wówczas r liniowo niezależnych wierszy $w_1, w_2, \ldots, w_r$ macierzy A_{uv} takich, że każdy wiersz macierzy A_{uv} jest pewną kombinacją liniową wierszy $w_1, w_2, \ldots, w_r$. Dla każdego $i \in \{1,2,\ldots,r\}$ symbolem ω_i oznaczmy wiersz macierzy A powstały z wiersza w_i poprzez dopisanie 0 w jego w v-tej kolumnie. Niech ponadto ω_{r+1} oznacza u-ty wiersz macierzy A. Weźmy dowolne $\lambda_1, \lambda_2, \ldots \lambda_{r+1} \in K$. Załóżmy, że $\sum_{t=1}^{r+1} \lambda_t \circ \omega_t = \Theta \in M_{1\times n}(K)$. Wówczas $\lambda_{r+1} \cdot a_{uv} = 0$ oraz $\sum_{t=1}^{r} \lambda_t \circ w_t = \Theta \in M_{1\times(n-1)}(K)$. Na mocy pierwszej spośród uzyskanych równości otrzymujemy, że $\lambda_{r+1} = 0$. Druga zaś, wraz z liniową niezależnością wierszy $w_1, w_2, \ldots, w_r$ implikuje, że $\lambda_1 = \lambda_2 = \ldots = \lambda_r = 0$. Zatem wiersze $\omega_1, \omega_2, \ldots, \omega_{r+1}$ są liniowo niezależne. Rozważmy dowolny wiersz ω macierzy A o numerze różnym od u. Niech ω'

będzie wierszem macierzy A_{uv} powstałym z wiersza ω wskutek wykreślenia 0 znajdującego się w v-tej kolumnie wiersza ω. Istnieją wówczas $\mu_1, \mu_2, \ldots, \mu_r \in K$ takie, że $\omega' = \sum_{i=1}^{r} \mu_i \circ w_i$. Zatem $\omega = \sum_{i=1}^{r} \mu_i \circ \omega_i$. Wobec tego każdy wiersz macierzy A jest pewną kombinacją liniową liniowo niezależnych wierszy $\omega_1, \omega_2, \ldots, \omega_{r+1}$ tej macierzy. Ostatecznie otrzymujemy więc, że $r_w(A) = r + 1 = 1 + r_w(A_{uv})$.

Możemy teraz udowodnić, fundamentalny dla prezentowanej w ramach niniejszego rozdziału teorii, związek między rzędem wierszowym, a rzędem kolumnowym dowolnej macierzy dany w poniższym twierdzeniu.

Twierdzenie 9.1. Niech K będzie ciałem i niech $m, n \in \mathbb{N}$. Wówczas $r_k(A) = r_w(A)$ dla dowolnej macierzy $A \in M_{m\times n}(K)$.

Dowód. Niech $A = [a_{ij}]_{ij}$. Jeżeli $A = \Theta_{m\times n}$, to teza jest oczywista. Dla $A \neq \Theta_{m\times n}$ przeprowadzimy dowód indukcyjny względem liczby m wierszy macierzy A. Dla $m = 1$ otrzymujemy, że $A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{bmatrix}$ i $a_{ij} \neq 0$ dla pewnego $j \in \{1, 2, \ldots, n\}$. Zatem $\dim \mathrm{lin}_K\left([a_{11}, a_{12}, \ldots, a_{1n}]\right) = 1$ oraz $\dim \mathrm{lin}_K(a_{11}, a_{12}, \ldots, a_{1n}) = 1$, skąd $r_w(A) = 1$ i $r_k(A) = 1$, skąd $r_k(A) = r_w(A)$. Wobec tego teza twierdzenia jest prawdziwa dla $m = 1$. Rozważmy dowolną liczbę naturalną $m > 1$ i załóżmy, że żądana równość jest prawdziwa dla wszystkich liczb naturalnych mniejszych od m. Ponieważ $A \neq \Theta_{m\times n}$, to istnieją $u \in \{1, 2, \ldots, m\}$ oraz $v \in \{1, 2, \ldots, n\}$ takie, że $a_{uv} \neq 0$. Jeżeli $n = 1$, to $r_w(A) = r_k(A^T)$, więc z założenia indukcyjnego wynika, że $r_k(A^T) = r_w(A^T)$, czyli $r_w(A) = r_k(A)$. Przypuśćmy teraz, że $n > 1$. Niech $B = [b_{ij}]_{ij}$ będzie macierzą powstałą wskutek wykonania na wierszach macierzy A operacji elementarnych: $w_i - \frac{a_{iv}}{a_{uv}} \cdot w_u$ dla kolejnych elementów i zbioru $\{1, 2, \ldots, m\} \setminus \{u\}$. Wówczas $r_w(B) = r_w(A)$ na mocy Lematu 9.1. Ponadto, z Lematu 9.2 wynika, że $r_k(B) = r_k(A)$. Dalej, niech $C = [c_{ij}]_{ij}$ będzie macierzą powstałą wskutek wykonania na kolumnach macierzy B operacji elementarnych: $k_j - \frac{b_{uj}}{a_{uv}} \cdot k_v$ dla kolejnych elementów j zbioru $\{1, 2, \ldots, n\} \setminus \{v\}$. Powołując się ponownie na Lemat 9.1 uzyskujemy wówczas równość $r_k(C) = r_k(B)$, więc $r_w(C) = r_w(B)$ na mocy Lematu 9.3. Ponieważ macierz C spełnia założenia Lematu 9.4, to $r_k(C) = 1 + r_k(C_{uv})$ i $r_w(C) = 1 + r_w(C_{uv})$. Ponadto $r_k(C_{uv}) = r_w(C_{uv})$ na mocy założenia indukcyjnego, skąd $r_w(A) = r_w(B) = r_w(C) = 1 + r_w(C_{uv}) = 1 + r_k(C_{uv}) = r_k(C) = r_k(B) = r_k(A)$. Zasada indukcji matematycznej kończy dowód.

9.2 Rząd macierzy i jego własności

W świetle Twierdzenia 9.1 poprawna jest poniższa definicja rzędu macierzy.

Definicja 9.2. Rzędem macierzy A nazywamy wspólną wartość rzędu wierszowego $r_w(A)$ oraz rzędu kolumnowego $r_k(A)$ tej macierzy. Rząd macierzy A oznaczamy przez $r(A)$.

Uwaga 9.3. W celu uproszczenia notacji, w przypadku konkretnej macierzy zapisanej w formie tablicy, zazwyczaj pomija się nawiasy występujące po r.

Bezpośrednią konsekwencją Twierdzenia 9.1 oraz Lematów 9.1-9.3 jest następujące

Twierdzenie 9.2. Rząd dowolnej macierzy nie ulega zmianie po wykonaniu dowolnej operacji elementarnej na wierszach lub kolumnach tej macierzy.

Kolejną natychmiastową konsekwencją Twierdzenia 9.1 oraz równości $r_w(A) = r_k\left(A^T\right)$ jest:

Twierdzenie 9.3. Dla dowolnej macierzy A, $r(A) = r(A^T)$.

Dwa następne twierdzenia wraz z Twierdzeniem 9.2 odgrywają kluczową rolę przy obliczaniu rzędu konkretnych macierzy.

Twierdzenie 9.4. Niech K będzie ciałem i niech $m,n > 1$ będą liczbami naturalnymi. Jeżeli macierz $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$ spełnia warunki: $a_{uv} \neq 0$ dla pewnych $u \in \{1,2,\ldots,m\}$ i $v \in \{1,2,\ldots,n\}$, oraz $a_{iv} = 0$ dla każdego $i \in \{1,2,\ldots,m\} \setminus \{u\}$, to $r(A) = 1 + r(A_{uv})$.

Dowód. Niech B będzie macierzą powstałą wskutek wykonania na kolumnach macierzy A operacji elementarnych: $k_j - \frac{a_{uj}}{a_{uv}} \cdot k_v$ dla kolejnych elementów j zbioru $\{1,2,\ldots,n\} \setminus \{v\}$. Wtedy $B_{uv} = A_{uv}$ oraz z Twierdzenia 9.2 wynika, że $r(B) = r(A)$. Ponadto macierz B spełnia założenia Lematu 9.4. Stąd oraz na mocy Twierdzenia 9.1 otrzymujemy, że $r(A) = r(B) = r_k(B) = 1 + r_k(B_{uv}) = 1 + r_k(A_{uv}) = 1 + r(A_{uv})$.

Z Twierdzeń 9.4 i 9.3 wynika, że prawdziwe jest również:

Twierdzenie 9.5. Niech K będzie ciałem i niech $m,n > 1$ będą liczbami naturalnymi. Jeżeli macierz $A = [a_{ij}]_{ij} \in M_{m\times n}(K)$ spełnia warunki: $a_{uv} \neq 0$ dla pewnych $u \in \{1,2,\ldots,m\}$ i $v \in \{1,2,\ldots,n\}$, oraz $a_{uj} = 0$ dla każdego $j \in \{1,2,\ldots,n\}$, takiego, że $j \neq v$, to $r(A) = 1 + r(A_{uv})$.

Przykład 9.1. W oparciu o zaprezentowaną wyżej teorię wyznaczymy rząd macierzy:

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 7 & 4 & 14 & 11 \\ 2 & 3 & 4 & 5 \end{bmatrix} \in M_{3\times 4}(\mathbb{R}).$$

Mamy:

$$r\begin{bmatrix} 1 & 3 & 2 & 4 \\ 7 & 4 & 14 & 11 \\ 2 & 3 & 4 & 5 \end{bmatrix} \overset{\substack{k_3-2\cdot k_1 \\ k_4-k_2}}{=} r\begin{bmatrix} 1 & 3 & 0 & 1 \\ 7 & 4 & 0 & 7 \\ 2 & 3 & 0 & 2 \end{bmatrix} \overset{k_4-k_1}{=} r\begin{bmatrix} 1 & 3 & 0 & 0 \\ 7 & 4 & 0 & 0 \\ 2 & 3 & 0 & 0 \end{bmatrix} = r\begin{bmatrix} 1 & 3 \\ 7 & 4 \\ 2 & 3 \end{bmatrix} \overset{w_3-w_1}{=}$$

$$r\begin{bmatrix}1&3\\7&4\\1&0\end{bmatrix} = 1 + r\begin{bmatrix}3\\4\end{bmatrix} = 1+1 = 2.$$

9.3 Rząd macierzy kwadratowej a jej wyznacznik

Poniższe twierdzenie ilustruje ważny związek wyznacznika macierzy kwadratowej z jej rzędem.

Twierdzenie 9.6. Niech K będzie ciałem, niech $n \in \mathbb{N}$ i niech $A \in \mathrm{M}_n(K)$. Wówczas $r(A) = n$ wtedy i tylko wtedy, gdy $\det(A) \neq 0$.

Dowód. Niech $A = [a_{ij}]_{ij}$. Oznaczmy $\kappa_j = \begin{bmatrix} a_{1j} & a_{2j} & \dots & a_{nj} \end{bmatrix}^T$ dla każdego $j \in \{1,2,\dots,n\}$. Załóżmy najpierw, że $r(A) = n$. Wtedy zbiór $\{\kappa_1, \kappa_2, \dots, \kappa_n\}$ tworzy bazę przestrzeni K^n. Dla każdego $j \in \{1,2,\dots,n\}$ istnieją więc $x_{1j}, x_{2j}, \dots, x_{nj} \in K$ takie, że $\sum_{i=1}^n x_{ij} \circ \kappa_i = \varepsilon_j$, gdzie ε_j oznacza j-ty kolumnowy wektor bazy kanonicznej przestrzeni K^n. Stąd:

$$\varepsilon_j = \begin{bmatrix} \sum_{t=1}^n x_{tj} a_{1t} \\ \sum_{t=1}^n x_{tj} a_{2t} \\ \vdots \\ \sum_{t=1}^n x_{tj} a_{nt} \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^n a_{1t} x_{tj} \\ \sum_{t=1}^n a_{2t} x_{tj} \\ \vdots \\ \sum_{t=1}^n a_{nt} x_{tj} \end{bmatrix} \tag{9.3.1}$$

dla każdego $j \in \{1,2,\dots,n\}$. Niech $X = [x_{ij}]_{ij}$. Wtedy $X \in \mathrm{M}_n(K)$ oraz z (9.3.1) wynika, że $A \cdot X = I_n$. Z Twierdzenia Cauchy'ego (zob. Twierdzenie 5.1) wynika więc, że $\det(A) \cdot \det(X) = \det(A \cdot X) = \det(I_n) = 1$, skąd $\det(A) \neq 0$.

Na odwrót. Przypuśćmy teraz, że $\det(A) \neq 0$. Istnieje wówczas macierz $X = [x_{ij}]_{ij} \in \mathrm{M}_n(K)$ taka, że $A \cdot X = I_n$ (zob. Twierdzenie 5.2). Zatem dla każdego $j \in \{1,2,\dots,n\}$ zachodzi wzór (9.3.1). Stąd $\sum_{i=1}^n x_{ij} \circ \kappa_i = \varepsilon_j$ dla każdego $j \in \{1,2,\dots,n\}$. Ponadto $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ jest bazą n-wymiarowej przestrzeni liniowej K^n, więc $\mathrm{lin}(\kappa_1, \kappa_2, \dots, \kappa_n) = K^n$. Stąd oraz na mocy Twierdzenia 8.8 uzyskujemy, że zbiór $\{\kappa_1, \kappa_2, \dots, \kappa_n\}$ jest bazą przestrzeni liniowej K^n. Wobec tego $r(A) = n$.

Uwaga 9.4. Na podstawie powyższego twierdzenia i drugiej części jego dowodu można uzasadnić tezę podaną w Uwadze 5.7, na mocy której każdą nieosobliwą macierz można sprowadzić przy pomocy operacji elementarnych na wierszach do macierzy jednostkowej. Istotnie, jeśli $A = [a_{ij}]_{ij} \in \mathrm{M}_n(K)$, $\det(A) \neq 0$ oraz $\omega_i = [a_{i1}, a_{i2}, \dots, a_{in}]$ dla każdego $i \in \{1,2,\dots,n\}$, to z Twierdzenia 9.6 wynika, że zbiór $\{\omega_1, \omega_2, \dots, \omega_n\}$ jest bazą przestrzeni liniowej K^n. Dla każdego $i \in \{1,2,\dots,n\}$ istnieją więc $x_{i1}, x_{i2}, \dots, x_{in} \in K$ takie, że $\varepsilon_i = \sum_{j=1}^n x_{ij} \omega_j$. Dla każdego $i \in \{1,2,\dots,n\}$ wystarczy więc wykonać kolejno następujące operacje elementarne: $x_{ii} \cdot w_i$, $w_i + x_{i1} \cdot w_1$, $w_i + x_{i2} \cdot w_2$,..., $w_i + x_{ii-1} \cdot w_{i-1}$, $w_i + x_{ii+1} \cdot w_{i+1}$,..., $w_i + x_{in} \cdot w_n$.

Analogicznie uzasadnia się, że przy zastosowaniu operacji elementarnych na kolumnach, każda nieosobliwa macierz może zostać sprowadzona do macierzy jednostkowej.

9.4 Pojęcie minora i jego związek z rzędem macierzy

Definicja 9.3. Niech K będzie ciałem i niech m,n i s będą liczbami naturalnymi takimi, że $s \leq \min\{m,n\}$. Minorem stopnia s macierzy $A \in \mathrm{M}_{m\times n}(K)$ nazywamy wyznacznik macierzy kwadratowej stopnia s powstałej z macierzy A wskutek wykreślenia $m-s$ wierszy oraz $n-s$ kolumn.

Twierdzenie 9.7. Rząd dowolnej niezerowej macierzy równy jest maksymalnemu stopniowi jej niezerowego minora.

Dowód. Niech K będzie ciałem, niech $m,n \in \mathbb{N}$ i niech $A \in \mathrm{M}_{m\times n}(K) \setminus \{\Theta_{m\times n}\}$. Niech ponadto r i s oznaczają odpowiednio rząd macierzy A oraz maksymalny stopień niezerowego minora macierzy A. Istnieje wówczas r liniowo niezależnych wierszy macierzy A. Wykreślając pozostałe wiersze tej macierzy uzyskujemy macierz kwadratową $B \in \mathrm{M}_{r\times n}(K)$ taką, że $r(B) = r$. Z Twierdzenia 9.1 wynika więc, że istnieje r liniowo niezależnych kolumn macierzy B. Po wykreśleniu pozostałych kolumn tej macierzy otrzymujemy macierz $C \in \mathrm{M}_r(K)$ taką, że $r(C) = r$. Stąd oraz na mocy Twierdzenia 9.6, $\det(C) \neq 0$. Zatem $\det(C)$ jest niezerowym minorem stopnia r macierzy A i w konsekwencji $r \leq s$.

Pozostało udowodnić nierówność przeciwną. W tym celu rozważmy dowolną macierz kwadratową $D \in \mathrm{M}_s(K)$ powstałą z macierzy A wskutek wykreślenia $m-s$ wierszy oraz $n-s$-kolumn taką, że $\det(D) \neq 0$. Wtedy $r(D) = s$ na mocy Twierdzenia 9.6. Niech E będzie macierzą powstałą z macierzy A poprzez wykreślenie tych samych wierszy co dla macierzy D. Wtedy $E \in \mathrm{M}_{s\times n}(K)$, więc z Twierdzenia 9.1 wynika, że $r(E) \leq s$. Ponadto wszystkie kolumny macierzy D są liniowo niezależne, skąd $r(E) \geq s$. Zatem $r(E) = s$. Wobec tego $r \geq s$.

Przykład 9.2. W oparciu o Twierdzenie 9.7 wyznaczymy rząd macierzy rzeczywistej:

$$A = \begin{bmatrix} 2 & 3 & -1 & 1 \\ 4 & 2 & 0 & 5 \\ 0 & 4 & -2 & -3 \end{bmatrix}.$$

Ponieważ $A \in \mathrm{M}_{3\times 4}(\mathbb{R})$, to $r(A) \leq \min\{3,4\} = 3$. Badamy więc minory stopnia 3. Wszystkich takich minorów jest $\binom{4}{3} = \binom{4}{1} = \frac{4!}{1!\cdot(4-1)!} = \frac{4!}{3!} = 4$. Ponieważ:

$$\begin{vmatrix} 2 & 3 & -1 \\ 4 & 2 & 0 \\ 0 & 4 & -2 \end{vmatrix} = -8-16+0-0+24 = 0, \quad \begin{vmatrix} 2 & 3 & 1 \\ 4 & 2 & 5 \\ 0 & 4 & -3 \end{vmatrix} = -12+16+0+36-40 = 0.$$

$$\begin{vmatrix} 2 & -1 & 1 \\ 4 & 0 & 5 \\ 0 & -2 & -3 \end{vmatrix} = 0-8+0-0-12+20 = 0, \quad \begin{vmatrix} 3 & -1 & 1 \\ 2 & 0 & 5 \\ 4 & -2 & -3 \end{vmatrix} = 0-4-20-0-6+30 = 0,$$

to wszystkie minory stopnia 3 macierzy A są zerowe. Zatem $r(A) \leq 2$ i badamy minory stopnia 2 macierzy A. Wszystkich takich minorów jest $\binom{4}{2} = \frac{4!}{2!\cdot(4-2)!} = \frac{4!}{4} = 3! = 6$. Ponieważ:

$$\begin{vmatrix} 2 & 3 \\ 4 & 2 \end{vmatrix} = 4 - 12 = -8 \neq 0,$$

to $r(A) = 2$.

9.5 Twierdzenie Kroneckera-Capellego

Twierdzenie 9.8 (Kronecker & Capelli). Układ:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{9.5.1}$$

m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K posiada rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A)$ dla $A = [a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$ oraz $A_u = [A|b] \in \mathrm{M}_{m\times(n+1)}(K)$, gdzie $b = \begin{bmatrix} b_1 & b_2 & \ldots & b_m \end{bmatrix}^T$. Ponadto układ ten posiada dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy $r(A_u) = r(A) = n$.

Dowód. Niech κ_j oznacza j-tą kolumnę macierzy A dla każdego $j \in \{1, 2, \ldots, n\}$. Z przemienności mnożenia w ciele wynika wówczas, że ciąg $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ elementów ciała K jest rozwiązaniem układu (9.5.1) wtedy i tylko wtedy, gdy zachodzi równość:

$$\sum_{j=1}^{n} \lambda_j \circ \kappa_j = b. \tag{9.5.2}$$

Stąd oraz na mocy punktu (ii) Wniosku 7.3 otrzymujemy, że układ (9.5.1) posiada rozwiązanie wtedy i tylko wtedy, gdy $b \in \mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n)$, co wobec Stwierdzenia 7.3 oznacza równość $\mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n, b) = \mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n)$. Ponieważ $\mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n) \subseteq \mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n, b)$, to w świetle Twierdzenia 8.9 powyższa równość podprzestrzeni jest tożsama z równością ich wymiarów, czyli $r(A_u) = r(A)$.

Na mocy powyższych rozważań otrzymujemy, że jeżeli $r(A_u) = r(A) = n$, to b jest wektorem n-wymiarowej przestrzeni liniowej $V = \mathrm{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n)$. Twierdzenia 8.8 i 8.6 implikują więc istnienie dokładnie jednego ciągu $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ elementów

ciała K takiego, że zachodzi równość (9.5.2). Ciąg ten jest więc jedynym rozwiązaniem układu (9.5.1).

Przypuśćmy teraz, że układ (9.5.1) ma dokładnie jedno rozwiązanie i jest nim $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Z udowodnionej już części twierdzenia wynika wówczas, że $r(A_u) = r(A) = \dim \operatorname{lin}_K(\kappa_1, \kappa_2, \ldots, \kappa_n)$. W świetle Twierdzenia 8.8 wystarczy więc wykazać liniową niezależność układu $(\kappa_1, \kappa_2, \ldots, \kappa_n)$. W tym celu rozważmy dowolne $\gamma_1, \gamma_2, \ldots, \gamma_n \in K$ i załóżmy, że $\sum_{j=1}^{n} \gamma_j \circ \kappa_j = 0$. Ponieważ $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ jest rozwiązaniem układu (9.5.1), to również $\sum_{j=1}^{n} \lambda_j \circ \kappa_j = 0$, skąd $\sum_{j=1}^{n} (\lambda_j + \gamma_j) \circ \kappa_j = 0$ i w konsekwencji $(\lambda_1 + \gamma_1, \lambda_2 + \gamma_2, \ldots, \lambda_n + \gamma_n)$ jest rozwiązaniem układu (9.5.1). Przyjęte założenie implikuje więc, że $(\lambda_1 + \gamma_1, \lambda_2 + \gamma_2, \ldots, \lambda_n + \gamma_n) = (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Zatem $\lambda_j + \gamma_j = \lambda_j$, czyli $\gamma_j = 0$ dla każdego $j \in \{1, 2, \ldots, n\}$. Wobec tego układ $(\kappa_1, \kappa_2, \ldots, \kappa_n)$ jest liniowo niezależny.

Przykład 9.3. Korzystając z Twierdzenia Kroneckera-Capellego oraz poznanych metod obliczania rzędu macierzy, określimy liczbę rozwiązań układu równań:

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 3 \\ 2x_1 - 3x_2 - x_3 + x_4 = -1 \\ x_1 + 7x_2 - x_4 = 4 \\ x_1 - x_2 + 2x_3 - ax_4 = 1 \end{cases} \tag{9.5.3}$$

nad ciałem $\mathbb{R}$ w zależności od parametru rzeczywistego a. Mamy:

$$A = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & -3 & -1 & 1 \\ 1 & 7 & 0 & -1 \\ 1 & -1 & 2 & -a \end{bmatrix} \text{ i } A_u = \left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 3 \\ 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 1 & -1 & 2 & -a & 1 \end{array}\right] \tag{9.5.4}$$

oraz:

$$r(A_u) = r \left[\begin{array}{cccc|c} 3 & 2 & 1 & 0 & 3 \\ 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 1 & -1 & 2 & -a & 1 \end{array}\right] \begin{matrix} w_1 - w_2 \\ w_4 - w_3 \\ = \end{matrix} r \left[\begin{array}{cccc|c} 1 & 5 & 2 & -1 & 4 \\ 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 0 & -8 & 2 & 1-a & -3 \end{array}\right]$$

$$\begin{matrix} w_2 - 2w_1 \\ w_3 - w_1 \\ = \end{matrix} r \left[\begin{array}{cccc|c} 1 & 5 & 2 & -1 & 4 \\ 0 & -13 & -5 & 3 & -9 \\ 0 & 2 & -2 & 0 & 0 \\ 0 & -8 & 2 & 1-a & -3 \end{array}\right] = 1 + r \left[\begin{array}{ccc|c} -13 & -5 & 3 & -9 \\ 2 & -2 & 0 & 0 \\ -8 & 2 & 1-a & -3 \end{array}\right] \begin{matrix} \\ \\ \frac{1}{2} w_2 \\ = \end{matrix}$$

$$1 + r \left[\begin{array}{ccc|c} -13 & -5 & 3 & -9 \\ 1 & -1 & 0 & 0 \\ -8 & 2 & 1-a & -3 \end{array}\right] \begin{matrix} w_1 - 5w_2 \\ w_3 + 2w_2 \\ = \end{matrix} 1 + r \left[\begin{array}{ccc|c} -18 & 0 & 3 & -9 \\ 1 & -1 & 0 & 0 \\ -6 & 0 & 1-a & -3 \end{array}\right] =$$

$$2+r\begin{bmatrix}-18 & 3 & | & -9\\ -6 & 1-a & | & -3\end{bmatrix}\overset{-\frac{1}{3}w_1}{=}2+r\begin{bmatrix}6 & -1 & | & 3\\ -6 & 1-a & | & -3\end{bmatrix}\overset{w_3+w_1}{=}2+r\begin{bmatrix}6 & -1 & | & 3\\ 0 & -a & | & 0\end{bmatrix}$$

$$\overset{k_3-\frac{1}{2}k_1}{=}2+r\begin{bmatrix}6 & -1 & | & 0\\ 0 & -a & | & 0\end{bmatrix}.$$

Po wykonaniu takich samych operacji elementarnych na wierszach macierzy A uzyskujemy, że:

$$r(A)=2+r\begin{bmatrix}6 & -1\\ 0 & -a\end{bmatrix}. \tag{9.5.5}$$

Zatem $r(A_u)=r(A)$ dla każdego $a\in\mathbb{R}$. Ponadto dla $a\neq 0$ otrzymujemy, że $r(A_u)=r(A)=3+r\begin{bmatrix}6\end{bmatrix}=3+1=4$, zaś dla $a=0$, $r(A_u)=r(A)=2+r\begin{bmatrix}6 & -1\end{bmatrix}=2+1=3$. Stąd oraz na mocy Twierdzenia 9.8 i Uwagi 9.5, układ (9.5.3):

- posiada dokładnie jedno rozwiązanie dla $a\in\mathbb{R}\setminus\{0\}$;
- posiada nieskończenie wiele rozwiązań zależnych od $4-3=1$ parametru dla $a=0$.

Uwaga 9.5. Jeżeli przy oznaczeniach Twierdzenia 9.8 zachodzi $r(A_u)=r(A)=s$ i $s<n$, to układ (9.5.1) posiada więcej niż jedno rozwiązanie, przy czym s niewiadomych jest zależnych od $n-s$ pozostałych niewiadomych, które są dowolnymi elementami ciała K. W szczególności, jeśli $|K|=\infty$, to układ (9.5.1) posiada nieskończenie wiele rozwiązań zależnych od $n-s$ parametrów. Istotnie, po wyznaczeniu s liniowo niezależnych wierszy macierzy A_u, wykreślamy wszystkie pozostałe wiersze tej macierzy. W powstałej w ten sposób macierzy A'_u wyznaczamy s liniowo niezależnych kolumn i zapisujemy układ równań liniowych, którego macierzą uzupełnioną jest A'_u. Następnie przenosimy na prawą stronę wszystkie niewiadome o numerach pozostałych $n-s$ kolumn. Możemy teraz zastosować wzory Cramera do wyznaczenia pozostałych niewiadomych, traktując przeniesione na prawą stronę niewiadome jak dowolnie ustalone elementy ciała K (wyznacznik główny takiego układu jest niezerowy na mocy liniowej niezależności wyznaczonych wcześniej s kolumn).

Inną metodą znalezienia rozwiązania układu (9.5.1) jest wykonanie na macierzy A'_u operacji elementarnych w sposób opisany w algorytmie związanym z metodą eliminacji Gaussa. Wtedy macierz A'_u sprowadzi się do $s\times n$-macierzy macierzy, której pierwszych s-kolumn utworzy macierz jednostkową stopnia s i możemy odczytać rozwiązanie (9.5.1) w sposób, w jaki robiliśmy to rozwiązując układy równań liniowych metodą eliminacji Gaussa.

Przykład 9.4. Nad ciałem $\mathbb{R}$ rozważmy układ równań:

$$\begin{cases}3x_1+2x_2+x_3=3\\ 2x_1-3x_2-x_3+x_4=-1\\ x_1+7x_2-x_4=4\\ x_1-x_2+2x_3=1\end{cases}. \tag{9.5.6}$$

Układ ten otrzymujemy z układu (9.5.3), podstawiając $a = 0$. Z Przykładu 9.3 wynika więc, że układ (9.5.6) posiada nieskończenie wiele rozwiązań zależnych od jednego parametru. Pokażemy w jaki sposób można rozwiązać ten układ na podstawie metod omówionych w Uwadze 9.5. Najpierw należy oczywiście wyznaczyć rzędy macierzy uzupełnionej A_u rozważanego układu i macierzy A współczynników tego układu. Na podstawie (9.5.4) i (9.5.5) otrzymujemy, że:

$$A = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & -3 & -1 & 1 \\ 1 & 7 & 0 & -1 \\ 1 & -1 & 2 & 0 \end{bmatrix} \text{ i } A_u = \begin{bmatrix} 3 & 2 & 1 & 0 & 3 \\ 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 1 & -1 & 2 & 0 & 1 \end{bmatrix}$$

oraz $r(A_u) = r(A) = 3$. Zatem można wskazać maksymalnie trzy liniowo niezależne wiersze macierzy A_u. Ponieważ:

$$\begin{vmatrix} 2 & -3 & -1 \\ 1 & 7 & 0 \\ 1 & -1 & 2 \end{vmatrix} = 28 + 1 + 0 + 7 + 6 - 0 = 42 \neq 0, \tag{9.5.7}$$

to z równości $r(A_u) = 3$ i Twierdzenia 9.7 wynika, że wiersze w_2, w_3 i w_4 macierzy A_u są liniowo niezależne. Po wykreśleniu wiersza w_1 uzyskujemy macierz:

$$A'_u = \begin{bmatrix} 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 1 & -1 & 2 & 0 & 1 \end{bmatrix}.$$

Zapisujemy układ równań liniowych, którego macierzą uzupełnioną jest A'_u:

$$\begin{cases} 2x_1 - 3x_2 - x_3 + x_4 = -1 \\ x_1 + 7x_2 - x_4 = 4 \\ x_1 - x_2 + 2x_3 = 1 \end{cases}. \tag{9.5.8}$$

Powołując się na (9.5.7) i Twierdzenie 9.7 wnioskujemy, że kolumny k_1, k_2 i k_3 macierzy A'_u są liniowo niezależne. Układ (9.5.8) przekształcamy więc do postaci:

$$\begin{cases} 2x_1 - 3x_2 - x_3 = -1 - x_4 \\ x_1 + 7x_2 = 4 + x_4 \\ x_1 - x_2 + 2x_3 = 1 \end{cases}. \tag{9.5.9}$$

Traktując niewiadomą x_4 jak dowolnie ustalony parametr, na mocy (9.5.7) otrzymujemy, że (9.5.9) jest układem Cramera, dla którego $W = 42$. Ponadto:

$$W_1 = \begin{vmatrix} -1-x_4 & -3 & -1 \\ 4+x_4 & 7 & 0 \\ 1 & -1 & 2 \end{vmatrix} = -14-14x_4+0+4+x_4+7+24+6x_4-0 = 21-7x_4,$$

$$W_2 = \begin{vmatrix} 2 & -1-x_4 & -1 \\ 1 & 4+x_4 & 0 \\ 1 & 1 & 2 \end{vmatrix} = 16+4x_4-1+0+4+x_4+2+2x_4-0 = 21+7x_4$$

oraz

$$W_3 = \begin{vmatrix} 2 & -3 & -1-x_4 \\ 1 & 7 & 4+x_4 \\ 1 & -1 & 1 \end{vmatrix} = 14+1+x_4-12-3x_4+7+7x_4+3+8+2x_4 = 21+7x_4,$$

skąd:

$$\begin{cases} x_1 = \frac{1}{2} - \frac{1}{6}x_4 \\ x_2 = \frac{1}{2} + \frac{1}{6}x_4 \\ x_3 = x_2 \\ x_4 \in \mathbb{R} \end{cases}. \tag{9.5.10}$$

Jak wynika z Uwagi 9.5, w celu rozwiązania układu (9.5.6), zamiast stosować wzory Cramera dla układu (9.5.9), możemy przekształcić macierz A'_u w następujący sposób:

$$A'_u = \left[\begin{array}{cccc|c} 2 & -3 & -1 & 1 & -1 \\ 1 & 7 & 0 & -1 & 4 \\ 1 & -1 & 2 & 0 & 1 \end{array}\right] \overset{w_1 \leftrightarrow w_3}{\sim} \left[\begin{array}{cccc|c} 1 & -1 & 2 & 0 & 1 \\ 1 & 7 & 0 & -1 & 4 \\ 2 & -3 & -1 & 1 & -1 \end{array}\right] \overset{\substack{w_2 - w_1 \\ w_3 - 2w_1}}{\sim}$$

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 0 & 1 \\ 0 & 8 & -2 & -1 & 3 \\ 0 & -1 & -5 & 1 & -3 \end{array}\right] \overset{\substack{w_1 - w_3 \\ w_2 + 8w_3}}{\sim} \left[\begin{array}{cccc|c} 1 & 0 & 7 & -1 & 4 \\ 0 & 0 & -42 & 7 & -21 \\ 0 & -1 & -5 & 1 & -3 \end{array}\right] \overset{\substack{\frac{-1}{7}w_2 \\ (-1)w_3}}{\sim}$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & 7 & 0 & -1 & 4 \\ 0 & 0 & 6 & -1 & 1 & 3 \\ 0 & 1 & 5 & -1 & 1 & 3 \end{array}\right] \overset{w_2 \leftrightarrow w_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 7 & -1 & 1 & 4 \\ 0 & 1 & 5 & -1 & 1 & 3 \\ 0 & 0 & 6 & -1 & 1 & 3 \end{array}\right] \overset{\frac{1}{6}w_3}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 7 & -1 & 1 & 4 \\ 0 & 1 & 5 & -1 & 1 & 3 \\ 0 & 0 & 1 & -\frac{1}{6} & 1 & \frac{1}{2} \end{array}\right]$$

$$\overset{\substack{w_1 - 7w_3 \\ w_2 - 5w_3}}{\sim} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & \frac{1}{6} & 1 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{6} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{6} & 1 & \frac{1}{2} \end{array}\right],$$

skąd uzyskujemy rozwiązanie dane w (9.5.10).

Rozdział 10

Odwzorowania liniowe

10.1 Określenie odwzorowania liniowego

Definicja 10.1. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Odwzorowaniem liniowym K-przestrzeni V w K-przestrzeń W nazywamy odwzorowanie $f\colon V \to W$ spełniające dla wszystkich $x, y \in V$ oraz $\lambda \in K$ koniunkcję warunków:

(i) $f(x+y) = f(x) + f(y)$;
(ii) $f(\lambda \circ x) = \lambda \circ f(x)$.

Uwaga 10.1. Warunek (i) nazywa się addytywnością, zaś warunek (ii) – jednorodnością.

Uwaga 10.2. Odwzorowania liniowe nazywane są także przekształceniami liniowymi lub homomorfizmami liniowymi. Nie należy ich mylić z funkcjami liniowymi rozpatrywanymi w szkole średniej (zob. Przykład 10.2).

Postępując analogicznie jak w dowodzie Stwierdzenia7.2 można wykazać następujące

Stwierdzenie 10.1. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Odwzorowanie $f\colon V \to W$ jest liniowe wtedy i tylko wtedy, gdy $f(\lambda_1 \circ v_1 + \lambda_2 \circ v_2) = \lambda_1 \circ f(v_1) + \lambda_2 \circ f(v_2)$ dla wszystkich $\lambda_1, \lambda_2 \in K$ i $v_1, v_2 \in V$.

Przykład 10.1. Pokażemy, że odwzorowanie $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ określone jest wzorem:

$$f\big([x,y,z]\big) = [x+2y-z, x+y+2z, 2x+3y+z]$$

jest przekształceniem liniowym. W tym celu weźmy dowolne $[x,y,z], [a,b,c] \in \mathbb{R}^3$ oraz $\lambda \in \mathbb{R}$. Wtedy $f\big([x,y,z]+[a,b,c]\big) = f\big([x+a, y+b, z+c]\big) = \big[(x+a)+2(y+b)-(z+c), (x+a)+(y+b)+2(z+c), 2(x+a)+3(y+b)+(z+c)\big] = [x+2y-z, x+y+2z, 2x+3y+z] + [a+2b-c, a+b+2c, 2a+3b+c] = f\big([x,y,z]\big) + f\big([a,b,c]\big)$ i $f\big(\lambda \circ [x,y,z]\big) = f\big([\lambda x, \lambda y, \lambda z]\big) = \big[\lambda x + 2\lambda y - \lambda z, \lambda x + \lambda y + 2\lambda z, 2\lambda x + 3\lambda y + \lambda z\big] = \big[\lambda(x+2y-z), \lambda(x+y+2z), \lambda(2x+3y+z)\big] = \lambda \circ [x+2y-z, x+y+2z, 2x+3y+z] = \lambda \circ f\big([x,y,z]\big)$. Zatem odwzorowanie f jest addytywne i jednorodne, co oznacza, że jest ono przekształceniem liniowym.

10.2 Podstawowe własności odwzorowań liniowych

Stwierdzenie 10.2. Niech V, W i U będą przestrzeniami liniowymi nad ciałem K. Jeżeli $f\colon V \to W$ oraz $g\colon W \to U$ są odwzorowaniami liniowymi, to złożenie $g \circ f$ funkcji f z funkcją g jest odwzorowaniem liniowym przekształcającym K-przestrzeń V w K-przestrzeń U.

Dowód. W celu uniknięcia konfliktu oznaczeń, w tym dowodzie mnożenie wektorów z lewej strony przez skalary będzie oznaczane przez $\bullet$. Wprost z określenia składania funkcji wynika, że $g \circ f\colon V \to U$. Weźmy dowolne $\lambda_1, \lambda_2 \in K$ oraz $v_1, v_2 \in V$. Na mocy Stwierdzenia 10.1 otrzymujemy wówczas, że $(g \circ f)(\lambda_1 \bullet v_1 + \lambda_2 \bullet v_2) = g\Big(f\big(\lambda_1 \bullet v_1 + \lambda_2 \bullet v_2\big)\Big) = g\big(\lambda_1 \bullet f(v_1) + \lambda_2 \bullet f(v_2)\big) = \lambda_1 \bullet g\big(f(v_1)\big) + \lambda_2 \bullet g\big(f(v_2)\big) = \lambda_1 \bullet (g \circ f)(v_1) + \lambda_2 \bullet (g \circ f)(v_2)$. Powołując się ponownie na Stwierdzenie 10.1 uzyskujemy więc, że $g \circ f$ jest odwzorowaniem liniowym.

Poniższe stwierdzenie zawiera podstawowe własności odwzorowań liniowych.

Stwierdzenie 10.3. Niech V i W będą przestrzeniami liniowymi nad ciałem K i niech $f\colon V \to W$ będzie odwzorowaniem liniowym. Wówczas:

(i) $f(0) = 0$;
(ii) $f(-v) = -f(v)$ dla każdego $v \in V$;
(iii) $f(v_1 - v_2) = f(v_1) - f(v_2)$ dla wszystkich $v_1, v_2 \in V$;
(iv) $f\left(\sum_{i=1}^{n} \lambda_i \circ v_i\right) = \sum_{i=1}^{n} \lambda_i \circ f(v_i)$ dla wszystkich $n \in \mathbb{N}$, $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ oraz $v_1, v_2, \ldots, v_n \in V$;
(v) dla dowolnych $v_1, v_2, \ldots, v_n$ wektorów przestrzeni V, liniowa niezależność układu $\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$ wektorów przestrzeni W implikuje liniową niezależność układu $(v_1, v_2, \ldots, v_n)$ wektorów przestrzeni V;
(vi) jeżeli U jest podprzestrzenią K-przestrzeni V, to $f(U)$ jest podprzestrzenią K-przestrzeni W;
(vii) jeżeli Y jest podprzestrzenią K-przestrzeni W, to $f^{-1}(Y)$ jest podprzestrzenią K-przestrzeni V;
(viii) jeżeli $X \subseteq V$, to $f\big(\mathrm{lin}(X)\big) = \mathrm{lin}\big(f(X)\big)$.

Dowód. (i). Ponieważ $f(0) + 0 = f(0) = f(0+0) = f(0) + f(0)$, to z punktu (i) Stwierdzenia 7.1 wynika, że $f(0) = 0$.

(ii). Weźmy dowolne $v \in V$. Podstawiając $\lambda = -1$ w punkcie (ii) Definicji 10.1 i powołując się na punkt (iv) Stwierdzenia 7.1 otrzymujemy $f(-v) = f\big((-1) \circ c\big) = (-1) \circ f(v) = -f(v)$.

(iii). Teza wynika natychmiast ze Stwierdzenia 10.1 przy zastosowaniu podstawienia $\lambda_1 = 1$ i $\lambda_2 = -1$, punktu (L4) Definicji 7.2 oraz z punktu (ii) niniejszego stwierdzenia.

(iv). Teza wynika ze Stwierdzenia 10.1 przez prostą indukcję.

(v). Weźmy dowolne $v_1, v_2, \ldots, v_n \in V$. Załóżmy, że układ $\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$ jest liniowo niezależny. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Załóżmy, że $\sum_{i=1}^{n} \lambda_i \circ v_i = 0$. Wtedy $0 = f(0) = f\big(\sum_{i=1}^{n} \lambda_i \circ v_i\big) = \sum_{i=1}^{n} \lambda_i \circ f(v_i)$ odpowiednio na mocy punktów (i) oraz (iv) niniejszego stwierdzenia. Liniowa niezależność układu $\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$ implikuje więc, że $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$. Wobec tego układ $(v_1, v_2, \ldots, v_n)$ jest liniowo niezależny.

(vi). Załóżmy, że U jest podprzestrzenią K-przestrzeni V. Wtedy $0 \in U$, więc $f(U) \neq \emptyset$. Jasne jest, że $f(U) \subseteq W$. Weźmy dowolne $w_1, w_2 \in f(U)$ oraz dowolne $\lambda_1, \lambda_2 \in K$. Istnieją wówczas $v_1, v_2 \in U$ takie, że $w_1 = f(v_1)$ i $w_2 = f(v_2)$. Stąd oraz na mocy Stwierdzeń 10.1 i 7.2 uzyskujemy, że $\lambda_1 \circ w_1 + \lambda_2 \circ w_2 = \lambda_1 \circ f(v_1) + \lambda_2 \circ f(v_2) = f(\lambda_1 \circ v_1 + \lambda_2 \circ v_2) \in f(U)$. Powołując się ponownie na Stwierdzenie 7.2 otrzymujemy stąd, że $f(U)$ jest podprzestrzenią w W.

(vii). Przypuśćmy, że Y jest podprzestrzenią K-przestrzeni W. Wtedy $0 \in W$. Ponadto $0 = f(0)$ na mocy punktu (i) niniejszego stwierdzenia, więc $0 \in f^{-1}(Y)$. Zatem $f^{-1}(Y) \neq \emptyset$. Jasne jest, że $f^{-1}(Y) \subseteq V$. Weźmy dowolne $v_1, v_2 \in f^{-1}(Y)$ oraz dowolne $\lambda_1, \lambda_2 \in K$. Wówczas $f(v_1) \in Y$ oraz $f(v_2) \in Y$, więc ze Stwierdzenia 7.2 i Stwierdzenia 10.1 wynika, że $f(\lambda_1 \circ v_1 + \lambda_2 \circ v_2) = \lambda_1 \circ f(v_1) + \lambda_2 \circ f(v_2) \in Y$, skąd $\lambda_1 \circ v_1 + \lambda_2 \circ v_2 \in f^{-1}(Y)$. Powołując się ponownie na Stwierdzenie 7.2 otrzymujemy stąd, że $f^{-1}(Y)$ jest podprzestrzenią w V.

$(viii)$. Załóżmy, że $X \subseteq V$. Jeżeli $X = \emptyset$, to $\operatorname{lin}(X) = \{0\}$ oraz $f(X) = \emptyset$, więc teza wynika z punktu (i) niniejszego stwierdzenia. Niech dalej $X \neq \emptyset$. Weźmy dowolne $w \in W$. Załóżmy najpierw, że $w \in f\big(\operatorname{lin}(X)\big)$. Istnieje wówczas $v \in \operatorname{lin}(X)$ takie, że $w = f(v)$. Ponadto $v = \sum_{i=1}^{n} \lambda_i \circ x_i$ dla pewnych $n \in \mathbb{N}$, $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ oraz $x_1, x_2, \ldots, x_n \in X$ (zob. Stwierdzenie 7.4), więc z punktu (iv) niniejszego stwierdzenia wynika, że $w = f(v) = f\big(\sum_{i=1}^{n} \lambda_i \circ x_i\big) = \sum_{i=1}^{n} \lambda_i \circ f(x_i)$. Zatem $w \in \operatorname{lin}\big(f(X)\big)$. Wobec tego $f\big(\operatorname{lin}(X)\big) \subseteq \operatorname{lin}\big(f(X)\big)$. Aby wykazać inkluzję przeciwną załóżmy teraz, że $w \in \operatorname{lin}\big(f(X)\big)$. Wtedy $w = \sum_{i=1}^{m} \alpha_i \circ y_i$ dla pewnych $m \in \mathbb{N}$, $\alpha_1, \alpha_2, \ldots, \alpha_n \in K$ oraz $y_1, y_2, \ldots, y_m \in f(X)$. Ponadto dla każdego $i \in \{1, 2, \ldots, m\}$ istnieje $v_i \in X$ takie, że $y_i = f(v_i)$. Stąd $w = \sum_{i=1}^{m} \alpha_i \circ f(v_i) = f\big(\sum_{i=1}^{m} \alpha_i \circ v_i\big) \in f\big(\operatorname{lin}(X)\big)$. Zatem $\operatorname{lin}\big(f(X)\big) \subseteq f\big(\operatorname{lin}(X)\big)$ i w konsekwencji $f\big(\operatorname{lin}(X)\big) = \operatorname{lin}\big(f(X)\big)$.

Uwaga 10.3. Stwierdzenie odwrotne do stwierdzenia podanego w ramach punktu (v) Stwierdzenia 10.3 nie jest prawdziwe. Aby się o tym przekonać wystarczy rozważyć funkcję $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ daną wzorem $f\big([x,y]\big) = [x,0]$ dla wszystkich $x, y \in \mathbb{R}$. Bezpośrednie sprawdzenie pokazuje, że f jest odwzorowaniem liniowym. Ponadto w przestrzeni $\mathbb{R}^2$, $\big([1,0],[0,1]\big)$ jest liniowo niezależny zaś układ $\Big(f\big([1,0]\big), f\big([0,1]\big)\Big) = \big([1,0],[0,0]\big)$ jest liniowo zależny.

Implikację daną w punkcie (v) Stwierdzenia 10.3 można odwrócić wówczas, gdy funkcja f jest różnowartościowa (zob. Stwierdzenie 10.5).

Przykład 10.2. Niech $a,b \in \mathbb{R}$ i niech $f\colon \mathbb{R} \to \mathbb{R}$ będzie funkcją daną wzorem $f(x) = ax+b$ dla każdego $x \in \mathbb{R}$. Jeżeli $b = 0$, to bezpośrednie sprawdzenie pokazuje, że funkcja f spełnia warunki (i) oraz (ii) Definicji 10.1, więc f jest wówczas odwzorowaniem liniowym. Jeśli natomiast $b \neq 0$, to $f(0) = b \neq 0$, więc funkcja f nie jest odwzorowaniem liniowym na mocy punktu (i) Stwierdzenia 10.3. Zatem „szkolna" funkcja liniowa $f(x) = ax+b$ jest odwzorowaniem liniowym wtedy i tylko wtedy, gdy $b = 0$.

10.3 Jądro i obraz przekształcenia liniowego

Definicja 10.2. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Jądrem odwzorowania liniowego $f\colon V \to W$ nazywamy zbiór:

$$\ker(f) = \big\{v \in V\colon f(v) = 0\big\}.$$

Uwaga 10.4. Wprost z powyższej definicji wynika, że $\ker(f) = f^{-1}\big(\{0\}\big)$. Ponadto $\{0\}$ jest podprzestrzenią K-przestrzeni W, więc na mocy punktu (vii) Stwierdzenia 10.3 otrzymujemy, że $\ker(f)$ jest podprzestrzenią K-przestrzeni liniowej V.

Definicja 10.3. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Obrazem odwzorowania liniowego $f\colon V \to W$ nazywamy zbiór:

$$\operatorname{im}(f) = \big\{f(v)\colon v \in V\big\}.$$

Uwaga 10.5. Wprost z powyższej definicji wynika, że $\operatorname{im}(f) = f(V)$. Ponadto V jest podprzestrzenią K-przestrzeni V, więc z punktu (vi) Stwierdzenia 10.3 wynika, że $\operatorname{im}(f)$ jest podprzestrzenią K-przestrzeni W.

Twierdzenie 10.1. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Jeżeli $\dim V < \infty$ oraz $f\colon V \to W$ jest odwzorowaniem liniowym, to $\dim\operatorname{im}(f) < \infty$ oraz $\dim V = \dim\ker(f) + \dim\operatorname{im}(f)$.

Dowód. Załóżmy, że $\dim V < \infty$ oraz $f\colon V \to W$ jest odwzorowaniem liniowym. Z Uwagi 10.4 wynika, że $\ker(f)$ jest podprzestrzenią w V. Stąd oraz na mocy Twierdzenia 8.9, $\dim\ker(f) < \infty$. Niech X będzie bazą K-przestrzeni $\ker(f)$. Wtedy $|X| = \dim\ker(f)$. Ponadto z Twierdzenia 8.7 wynika istnienie takiego podzbioru Y zbioru V, że $|Y| = \dim V - |X|$ oraz zbiór $X \cup Y$ jest bazą K-przestrzeni V. W szczególności wynika stąd, że $|Y| < \infty$. Jeżeli $Y = \emptyset$, to $|X| = \dim V$, czyli $\dim\ker(f) = \dim V$. Z Twierdzenia 8.9 wynika wówczas, że $V = \ker(f)$, co oznacza, że $f(v) = 0$ dla każdego $v \in V$. Zatem $im(f) = \{0\}$, skąd $\dim\operatorname{im}(f) = 0 < \infty$ i w konsekwencji $\dim V =$

$\dim\ker(f)+\dim\operatorname{im}(f)$. Niech dalej $Y\neq\emptyset$. Istnieje wówczas $n\in\mathbb{N}$ takie, że $|Y|=n$. Istnieją więc $v_1,v_2,\ldots,v_n\in V$ takie, że $Y=\{v_1,v_2,\ldots,v_n\}$. Ponadto $\operatorname{im}(f)=f(V)=f\big(\operatorname{lin}(X\cup Y)\big)=\operatorname{lin}\big(f(X\cup Y)\big)=\operatorname{lin}\big(f(X)\cup f(Y)\big)=\operatorname{lin}\big(f(X)\big)+\operatorname{lin}\big(f(Y)\big)=f\big(\operatorname{lin}(X)\big)+\operatorname{lin}\big(f(Y)\big)=f\big(\ker(f)\big)+\operatorname{lin}\big(f(Y)\big)=\{0\}+\operatorname{lin}\big(f(Y)\big)=\operatorname{lin}\big(f(Y)\big)$ na mocy punktu $(viii)$ Stwierdzenia 10.3 oraz Twierdzenia 7.3. Zatem $\dim\operatorname{im}(f)=\dim\operatorname{lin}\big(f(Y)\big)=\dim\operatorname{lin}\big(f(v_1),f(v_2),\ldots,f(v_n)\big)$. Weźmy dowolne $\lambda_1,\lambda_2,\ldots,\lambda_n\in K$. Załóżmy, że $\sum_{i=1}^n\lambda_i\circ f(v_i)=0$. Wtedy $f\left(\sum_{i=1}^n\lambda_i\circ v_i\right)=0$, na mocy punktu (iv) Stwierdzenia 10.3, więc $\sum_{i=1}^n\lambda_i\circ v_i\in\ker(f)$. Ale $\ker(f)=\operatorname{lin}(X)$. Jeżeli $X=\emptyset$, to $\ker(f)=\{0\}$, skąd $\sum_{i=1}^n\lambda_i\circ v_i=0$. Z liniowej niezależności zbioru Y wynika więc, że $\lambda_1=\lambda_2=\ldots=\lambda_n=0$. Jeśli natomiast $X\neq\emptyset$, to $\sum_{i=1}^n\lambda_i\circ v_i=\sum_{i=1}^m\mu_i\circ x_i$ dla pewnych $m\in\mathbb{N}$, $\mu_1,\mu_2,\ldots,\mu_m\in K$ oraz $x_1,x_2,\ldots,x_m\in X$. Stąd $\sum_{i=1}^n\lambda_i\circ v_i+\sum_{i=1}^m(-\mu_i)\circ x_i=0$. Liniowa niezależność zbioru $X\cup Y$ implikuje więc, że $\lambda_1=\lambda_2=\ldots=\lambda_n=0$ oraz $\mu_1=\mu_2=\ldots=\mu_m=0$. Zatem układ $\big(f(v_1),f(v_2),\ldots,f(v_n)\big)$ jest liniowo niezależny. Stąd oraz na mocy Wniosku 8.3 otrzymujemy, że $\dim\operatorname{im}(f)=n=|Y|$. Wobec tego $\dim V=|X|+|Y|=\dim\ker(f)+\dim\operatorname{im}(f)$.

Przykład 10.3. Dla odwzorowania liniowego f określonego w Przykładzie 10.1 wyznaczymy $\ker(f)$, $\operatorname{im}(f)$ oraz ich bazy i wymiary. Weźmy dowolne $x,y,z\in\mathbb{R}$. Wówczas:

$$[x,y,z]\in\ker(f)\Leftrightarrow f\big([x,y,z]\big)=[0,0,0]\Leftrightarrow[x+2y-z,x+y+2z,2x+3y+z]=[0,0,0] \tag{10.3.1}$$

Rozwiązanie powyższego równania równoważne jest rozwiązaniu układu równań:

$$\begin{cases} x+2y-z=0\\ x+y+2z=0\\ 2x+3y+z=0 \end{cases}. \tag{10.3.2}$$

Powyższy układ rozwiązujemy np. metodą eliminacji Gaussa (metoda Cramera nie da żadnego rezultatu, bo wyznacznik główny tego układu jest równy zero). Mamy:

$$\left[\begin{array}{ccc|c}1&2&-1&0\\1&1&2&0\\2&3&1&0\end{array}\right]\overset{w_3-(w_1+w_2)}{\equiv}\left[\begin{array}{ccc|c}1&2&-1&0\\1&1&2&0\\0&0&0&0\end{array}\right]\overset{w_2-w_1}{\equiv}\left[\begin{array}{ccc|c}1&2&-1&0\\0&-1&3&0\end{array}\right]\overset{w_1+2w_2}{\equiv}$$

$$\equiv\left[\begin{array}{ccc|c}1&0&5&0\\0&-1&3&0\end{array}\right].$$

Zatem układ (10.3.2) równoważny jest układowi warunków:

$$\begin{cases} x=-5z\\ y=3z\\ z\in\mathbb{R} \end{cases}.$$

Zatem układ (10.3.2) ma nieskończenie wiele rozwiązań i opisane są one warunkiem:

$$(x,y,z) = (-5t,3t,t), \tag{10.3.3}$$

gdzie t przebiega zbiór $\mathbb{R}$. Z (10.3.1) – (10.3.3) wynika więc, że:

$$\ker(f) = \big\{[-5t,3t,t] : t \in \mathbb{R}\big\} = \big\{t \circ [-5,3,1] : t \in \mathbb{R}\big\} = \operatorname{lin}\big([-5,3,1]\big).$$

Zatem zbiór $\big\{[-5,3,1]\big\}$ jest bazą $\mathbb{R}$-przestrzeni $\ker(f)$ i $\dim\ker(f) = 1$. Stąd oraz na mocy Twierdzenia 10.1, $\dim\operatorname{im}(f) = \dim\mathbb{R}^3 - \dim\ker(f) = 3-1 = 2$. Wobec tego $\operatorname{im}(f) = f\left(\mathbb{R}^3\right) = \operatorname{lin}(v_1,v_2)$, dla pewnych liniowo niezależnych wektorów v_1, v_2 przestrzeni $\mathbb{R}^3$ (oczywiście nie są one określone jednoznacznie!). Ponadto:

$$f\left(\mathbb{R}^3\right) = \big\{[x+2y-z, x+y+2z, 2x+3y+z] : x,y,z \in \mathbb{R}\big\} =$$

$$\big\{x \circ [1,1,2] + y \circ [2,1,3] + z \circ [-1,2,1] : x,y,z \in \mathbb{R}\big\} = \operatorname{lin}\big([1,1,2],[2,1,3],[-1,2,1]\big).$$

Opisane wyżej wektory v_1 i v_2 można wyznaczyć, opierając się na metodzie omówionej na początku Przykładu 8.20, tzn. wykonując operacje elementarne na wierszach macierzy:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{bmatrix}.$$

Ponieważ:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{bmatrix} \overset{\substack{w_2 - 2w_1 \\ w_3 + w_1}}{\equiv} \begin{bmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & 3 & 3 \end{bmatrix} \overset{\substack{w_1 + w_2 \\ w_3 + 3w_2}}{\equiv} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \overset{(-1)w_2}{\equiv} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

to $\operatorname{im}(f) = \operatorname{lin}\big([1,0,1],[0,1,1]\big)$. Stąd oraz na mocy Twierdzenia 8.8, zbiór:

$$\big\{[1,0,1],[0,1,1]\big\}$$

jest bazą $\mathbb{R}$-przestrzeń $\operatorname{im}(f)$.

Uwaga 10.6. Liniową niezależność zbioru $\big\{[1,0,1],[0,1,1]\big\}$ można dodatkowo zweryfikować bezpośrednio, np. wyznaczając rząd macierzy:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

(nie ma jednak takiej konieczności). Mamy wówczas:

$$r\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}=2,\ \text{bo}\ \begin{vmatrix}1&0&1\\0&1&1\\0&0&0\end{vmatrix}=0\ \text{i}\ \begin{vmatrix}1&0\\0&1\end{vmatrix}=1\neq 0$$

(por. Twierdzenie 9.7).

Taką dodatkową weryfikację liniowej niezależności wektorów rozpinających daną skończenie generowaną przestrzeń przeprowadza się czasem w bardziej skomplikowanych rachunkowo zadaniach, gdyż pozwala ona wykryć przynajmniej część ewentualnych błędów rachunkowych, które można popełnić szukając minimalnego zbioru generującego tę przestrzeń (czyli szukając bazy tej przestrzeni).

10.4 Szczególne typy odwzorowań liniowych

Definicja 10.4. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Odwzorowanie liniowe $f\colon V\to W$ nazywamy:

- monomorfizmem (liniowym), gdy funkcja f jest iniektywna (tzn. różnowartościowa);
- epimorfizmem (liniowym), gdy funkcja f jest surjektywna (tzn. typu „na");
- izomorfizmem (liniowym), gdy f jest jednocześnie monomorfizmem i epimorfizmem;
- endomorfizmem (liniowym), gdy $W=V$;
- automorfizmem (liniowym), gdy f jest jednocześnie endomorfizmem i izomorfizmem.

Przykład 10.4. Odwzorowanie $f\colon\mathbb{R}^2\to\mathbb{R}^3$ dane wzorem $f\big([x,y]\big)=[x,y,0]$ jest monomorfizmem liniowym.

Przykład 10.5. Funkcja $f\colon\mathbb{R}^2\to\operatorname{lin}_{\mathbb{R}}\big([1,0]\big)$ dana wzorem $f\big([x,y]\big)=[x,0]$ jest epimorfizmem liniowym.

Przykład 10.6. Niech $n\in\mathbb{N}$ i niech V będzie n-wymiarową przestrzenią liniową nad ciałem K. Niech ponadto $B=(v_1,v_2,\ldots,v_n)$ będzie bazą uporządkowaną K-przestrzeni V. Z Twierdzenia 8.6 i Definicji 8.7 wynika wówczas, że dowolny wektor $v\in V$ można w jednoznaczny sposób zapisać w postaci $v=\sum_{i=1}^{n}\lambda_i\circ v_i$, gdzie $\lambda_1,\lambda_2,\ldots,\lambda_n\in K$. Bezpośrednie sprawdzenie pokazuje, że funkcja:

$$V\ni v=\sum_{i=1}^{n}\lambda_i\circ v_i\overset{f}{\longmapsto}[\lambda_1,\lambda_2,\ldots,\lambda_n]\in K^n$$

jest izomorfizmem liniowym przestrzeni V na przestrzeń K^n.

Definicja 10.5. Mówimy, że przestrzeń liniowa V nad ciałem K jest izomorficzna z K-przestrzenią liniową W, gdy istnieje izomorfizm liniowy $f\colon V \to W$. Piszemy wówczas $V \cong W$.

Uwaga 10.7. Niech V, W oraz U będą przestrzeniami liniowymi nad ciałem K. Łatwo zauważyć, że:

(i) $V \cong V$;
(ii) jeżeli $V \cong W$, to $W \cong V$;
(iii) jeżeli $V \cong W$ i $W \cong U$, to $V \cong U$.

Przykład 10.7. Niech $n \in \mathbb{N}$ i niech V będzie n-wymiarową przestrzenią liniową nad ciałem K. Na mocy Przykładu 10.6 otrzymujemy, że $V \cong K^n$. Stąd oraz na mocy Uwagi 10.7 otrzymujemy, że wszystkie n-wymiarowe K-przestrzenie są izomorficzne.

Uwaga 10.8. Izomorficzne przestrzenie liniowe mają dokładnie te same własności algebraiczne. Dlatego można je ze sobą utożsamiać.

Poniższe stwierdzenie zawiera ważną charakteryzację monomorfizmów liniowych:

Stwierdzenie 10.4. Niech V i W będą przestrzeniami liniowymi nad ciałem K. Odwzorowanie liniowe $f\colon V \to W$ jest monomorfizmem wtedy i tylko wtedy, gdy $\ker(f) = \{0\}$.

Dowód. Załóżmy najpierw, że f jest monomorfizmem. Inkluzja $\{0\} \subseteq \ker(f)$ wynika wprost z punktu (i) Stwierdzenia 10.3. Aby wykazać inkluzję przeciwną, weźmy dowolne $v \in \ker(f)$. Wtedy $f(v) = 0$. Ponadto $f(0) = 0$ na mocy punktu (i) Stwierdzenia 10.3, więc iniektywność funkcji f implikuje, że $v = 0$. Zatem $\ker(f) \subseteq \{0\}$ i w konsekwencji $\ker(f) = \{0\}$.

Przypuśćmy teraz, że $\ker(f) = \{0\}$. Weźmy dowolne $v_1, v_2 \in V$. Załóżmy, że $f(v_1) = f(v_2)$. Wtedy $f(v_1) - f(v_2) = 0$. Stąd oraz na mocy punktu (iii) Stwierdzenia 10.3, $f(v_1 - v_2) = 0$, czyli $v_1 - v_2 \in \ker(f)$. Ale $\ker(f) = \{0\}$, więc $v_1 - v_2 = 0$, skąd $v_1 = v_2$. Zatem odwzorowanie liniowe f jest iniektywne, czyli f jest monomorfizmem.

Stwierdzenie 10.5. Niech V i W będą przestrzeniami liniowymi nad ciałem K i niech $f\colon V \to W$ będzie monomorfizmem. Wówczas dla dowolnych $v_1, v_2, \ldots, v_n \in V$, następujące warunki są równoważne:

(i) układ $(v_1, v_2, \ldots, v_n)$ jest liniowo niezależny;
(ii) układ $\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$ jest liniowo niezależny.

Dowód. Załóżmy, że układ $(v_1, v_2, \ldots, v_n)$ jest liniowo niezależny. Weźmy dowolne $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Przypuśćmy, że $\sum_{i=1}^n \lambda_i \circ f(v_i) = 0$. Wtedy $f\left(\sum_{i=1}^n \lambda_i \circ v_i\right) = 0$, skąd $\sum_{i=1}^n \lambda_i \circ v_i \in \ker(f)$. Ponadto $\ker(f) = \{0\}$ na mocy Stwierdzenia 10.4, więc $\sum_{i=1}^n \lambda_i \circ v_i = 0$ i z przyjętego założenia wynika, że $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$. Zatem układ $\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$ jest liniowo niezależny. Implikacja przeciwna jest bezpośrednią konsekwencją punktu (v) Stwierdzenia 10.3.

10.5 Konstruowanie odwzorowań liniowych

Twierdzenie 10.2. Niech $n \in \mathbb{N}$ i niech $\{e_1, e_2, \ldots, e_n\}$ będzie bazą n-wymiarowej przestrzeni liniowej V nad ciałem K. Niech ponadto $w_1, w_2, \ldots, w_n$ będą dowolnymi wektorami K-przestrzeni liniowej W. Istnieje wówczas dokładnie jedno odwzorowanie liniowe $f\colon V \to W$ takie, że $f(e_i) = w_i$ dla każdego $i \in \{1,2,\ldots,n\}$, i jest ono opisane wzorem:

$$f\left(\sum_{i=1}^{n} \lambda_i \circ e_i\right) = \sum_{i=1}^{n} \lambda_i \circ w_i, \tag{10.5.1}$$

dla wszystkich $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$. Ponadto:

(i) f jest monomorfizmem wtedy i tylko wtedy, gdy układ $(w_1, w_2, \ldots, w_n)$ wektorów K-przestrzeni W jest liniowo niezależny;
(ii) f jest epimorfizmem wtedy i tylko wtedy, gdy $W = \mathrm{lin}(w_1, w_2, \ldots, w_n)$;
(iii) f jest izomorfizmem wtedy i tylko wtedy, gdy $(w_1, w_2, \ldots, w_n)$ jest bazą uporządkowaną K-przestrzeni W.

Dowód. Weźmy dowolne $v, v' \in V$. Uzasadnimy najpierw, że wzór (10.5.1) określa odwzorowanie liniowe $f\colon V \to W$ takie, że $f(e_i) = w_i$ dla każdego $i \in \{1,2,\ldots,n\}$. Ponieważ zbiór $\{e_1, e_2, \ldots, e_n\}$ jest bazą K-przestrzeni V, to z Twierdzenia 8.5 wynika istnienie takich $\lambda_1, \lambda_2, \ldots, \lambda_n, \lambda_1', \lambda_2', \ldots, \lambda_n' \in K$, że $v = \sum_{i=1}^n \lambda_i \circ e_i$ oraz $v' = \sum_{i=1}^n \lambda_i' \circ e_i$. Stąd $f(v+v') = f\left(\sum_{i=1}^n \lambda_i \circ e_i + \sum_{i=1}^n \lambda_i' \circ e_i\right) = f\left(\sum_{i=1}^n \lambda_i \circ e_i + \lambda_i' \circ e_i\right) = f\left(\sum_{i=1}^n (\lambda_i + \lambda_i') \circ e_i\right) = \sum_{i=1}^n (\lambda_i + \lambda_i') \circ w_i = \sum_{i=1}^n \lambda_i \circ w_i + \lambda_i' \circ w_i = \sum_{i=1}^n \lambda_i \circ w_i + \sum_{i=1}^n \lambda_i' \circ w_i = f\left(\sum_{i=1}^n \lambda_i \circ e_i\right) + f\left(\sum_{i=1}^n \lambda_i' \circ e_i\right) = f(v) + f(v')$. Ponadto dla dowolnego $\alpha \in K$, $f(\alpha \circ v) = f\left(\alpha \circ \sum_{i=1}^n \lambda_i \circ e_i\right) = f\left(\sum_{i=1}^n \alpha \circ (\lambda_i \circ e_i)\right) = f\left(\sum_{i=1}^n (\alpha \cdot \lambda_i) \circ e_i\right) = \sum_{i=1}^n (\alpha \cdot \lambda_i) \circ w_i = \sum_{i=1}^n \alpha \circ (\lambda_i \circ w_i) = \alpha \circ \sum_{i=1}^n \lambda_i \circ w_i = \alpha \circ f\left(\sum_{i=1}^n \lambda_i \circ e_i\right) = \alpha \circ f(v)$. Zatem wzór (10.5.1) określa odwzorowanie liniowe $f\colon V \to W$. Wprost z (10.5.1) wynika, że $f(e_i) = w_i$ dla każdego $i \in \{1,2,\ldots,n\}$. Aby wykazać jednoznaczność przekształcenia f, rozważmy dowolne odwzorowanie liniowe $g\colon V \to W$ takie, że $g(e_i) = w_i$ dla każdego $i \in \{1,2,\ldots,n\}$. Wtedy $g(e_i) = w_i = f(e_i)$ dla każdego $i \in \{1,2,\ldots,n\}$, więc punkt (iv) Stwierdzenia 10.3 implikuje równość $g(v) = f(v)$. Stąd oraz na mocy dowolności wyboru $v \in V$ otrzymujemy, że $g = f$.

Pozostało wykazać równoważności dane w (i)-(iii):

(i). Przypuśćmy, że f jest monomorfizmem. Ponieważ $w_i = f(e_i)$ dla każdego $i \in \{1,2,\ldots,n\}$ oraz z założenia $(e_1, e_2, \ldots, e_n)$ jest liniowo niezależny, to liniowa niezależność układu $(w_1, w_2, \ldots, w_n)$ jest konsekwencją Stwierdzenia 10.5. Na odwrót. Załóżmy, że układ $(w_1, w_2, \ldots, w_n)$ jest liniowo niezależny. Jeżeli $v \in \ker(f)$, to $0 = f(v) = f\left(\sum_{i=1}^n \lambda_i \circ e_i\right) = \sum_{i=1}^n \lambda_i \circ w_i$, więc $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$ i w konsekwencji $v = 0$. Zatem $\ker(f) = \{0\}$, co w świetle Stwierdzenia 10.4 oznacza, że f jest monomorfizmem.

(ii). Teza jest oczywista na mocy równości $\mathrm{im}(f) = \mathrm{lin}(w_1, w_2, \ldots, w_n)$, która wy-

nika natychmiast z (10.5.1) i Twierdzenia 8.6.

(iii). Teza jest bezpośrednią konsekwencją punktów (i) oraz (ii) i Twierdzenia 8.5.

Przykład 10.8. Niech $m,n \in \mathbb{N}$ i niech K będzie ciałem. Opiszemy wszystkie odwzorowania liniowe przestrzeni K^n w przestrzeń K^m. Ponieważ zbiór $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ jest bazą przestrzeni K^n, to dowolny wektor $[x_1, x_2, \ldots, x_n]$ przestrzeni K^n zapisuje się w postaci $[x_1, x_2, \ldots, x_n] = \sum_{j=1}^{n} x_j \circ \varepsilon_j$. Stąd oraz na mocy Twierdzenia 10.2, ogólna postać wzoru odwzorowania liniowego przestrzeni K^n w przestrzeń K^m jest następująca:

$$f\big([x_1, x_2, \ldots, x_n]\big) = \sum_{j=1}^{n} x_j \circ \omega_j, \tag{10.5.2}$$

gdzie $\omega_1, \omega_2, \ldots, \omega_n$ są dowolnie ustalonymi wektorami w K^m. Ponadto dla każdego $j \in \{1,2,\ldots,n\}$ istnieją $a_{1j}, a_{2j}, \ldots, a_{mj} \in K$ takie, że $\omega_j = [a_{1j}, a_{2j}, \ldots, a_{mj}]$. Stąd oraz na mocy (10.5.2) uzyskujemy wzór:

$$f\big([x_1, x_2, \ldots, x_n]\big) = \left[\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j\right]. \tag{10.5.3}$$

Definicja 10.6. Wzór postaci (10.5.3) nazywamy wzorem analitycznym odwzorowania liniowego f przestrzeni K^n w przestrzeń K^m.

Przykład 10.9. Przekształcenie liniowe z Przykładu 10.1 określone jest za pomocą wzoru analitycznego.

Uwaga 10.9. Zachowując wszystkie oznaczenia z Przykładu 10.8 otrzymujemy, że dla wszystkich $x_1, x_2, \ldots, x_n \in K$ zachodzi równość:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix}.$$

Stąd oraz na mocy wzoru (10.5.3) przekształcenie liniowe f przestrzeni K^n w przestrzeń K^m jest określone jednoznacznie przez macierz $A = [a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$. Istotne związki odwzorowań liniowych i macierzy zostaną bardziej szczegółowo omówione w dalszej części niniejszego rozdziału.

Przykład 10.10. Przekształcenie liniowe z Przykładu 10.1 jest jednoznacznie określone przez macierz:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}.$$

Przykład 10.11. W oparciu o Twierdzenia 10.1, 10.2 i 8.7, skonstruujemy odwzorowanie liniowe $F\colon \mathbb{R}^3 \to \mathbb{R}^4$ spełniające warunki $\ker F = \operatorname{lin}\big([1,-1,2]\big)$ oraz $\operatorname{im} F = \operatorname{lin}\big([1,2,1,-1],[3,1,2,0]\big)$.

Ponieważ $\dim \mathbb{R}^3 = 3 < \infty$, $\dim \operatorname{lin}\big([1,-1,2]\big) = 1$, $\dim \operatorname{lin}\big([1,2,1,-1],[3,1,2,0]\big) = 2$ oraz $3 = 1+2$, to z Twierdzenia 10.1 wynika, że spełniony jest warunek konieczny istnienia takiego odwzorowania. Niech $v_0 = [1,-1,2]$, $w_1 = [1,2,1,-1]$ i $w_2 = [3,1,2,0]$. Z Twierdzenia 8.7 wynika istnienie takich $v_1, v_2 \in \mathbb{R}^3$, że zbiór $\{v_0, v_1, v_2\}$ jest bazą przestrzeni $\mathbb{R}^3$. Ponieważ:

$$r \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 3,$$

to wystarczy przyjąć $v_1 = \varepsilon_2$ i $v_2 = \varepsilon_3$. Powołując się na Twierdzenie 10.2 otrzymujemy, że funkcja $f\colon \mathbb{R}^3 \to \mathbb{R}^4$ dana wzorem:

$$f(x \circ v_0 + y \circ v_1 + z \circ v_2) = y \circ w_1 + z \circ w_2 \tag{10.5.4}$$

dla wszystkich $x, y, z \in \mathbb{R}$, jest jedynym odwzorowaniem liniowym przestrzeni $\mathbb{R}^3$ w przestrzeń $\mathbb{R}^4$ spełniającym warunki $f(v_0) = 0$, $f(v_1) = w_1$ i $f(v_2) = w_2$. W szczególności wynika stąd, że $v_0 \in \ker(f)$. Ponadto z (10.5.4) wynika, że $\operatorname{im}(f) = \operatorname{lin}(w_1, w_2)$. Zatem $\dim \operatorname{im}(f) = 2$. Twierdzenie 10.1 implikuje więc, że $\dim \ker(f) = 1$. Wobec tego $\ker(f) = \operatorname{lin}(v_0)$. Wystarczy więc przyjąć $F = f$.

Uwaga 10.10. Przy użyciu tych samych technik i oznaczeń można pokazać, że wzór:

$$g(x \circ v_0 + y \circ \varepsilon_1 + z \circ \varepsilon_2) = y \circ w_1 + z \circ w_2$$

dla wszystkich $x, y, z \in \mathbb{R}$, poprawnie definiuje odwzorowanie liniowe $g\colon \mathbb{R}^3 \to \mathbb{R}^4$ takie, że $\ker(g) = \operatorname{lin}(v_0)$ oraz $\operatorname{im}(g) = \operatorname{lin}(w_1, w_2)$. Ponadto $f(\varepsilon_1) = f(v_0 + v_1 - 2v_2) = f(v_0) + f(v_1) - 2f(v_2) = 0 + w_1 + w_2 = w_1 + w_2$ i $g(\varepsilon_1) = w_1$, więc $f(\varepsilon_1) \neq g(\varepsilon_1)$ i w konsekwencji $g \neq f$. Zatem odwzorowanie liniowe F spełniające warunki podane w Przykładzie 10.11 nie jest wyznaczone jednoznacznie.

10.6 Przestrzeń odwzorowań liniowych

Definicja 10.7. Zbiór wszystkich odwzorowań liniowych przestrzeni liniowej V nad ciałem K w K-przestrzeń W oznaczamy przez $\mathscr{L}_K(V,W)$. Jeśli $W = V$, to zamiast $\mathscr{L}_K(V,V)$ stosujemy oznaczenie $\operatorname{End}_K(V)$.

W sytuacjach niebudzących wątpliwości stosuje się również nieco uproszczoną notację: $\mathscr{L}(V,W)$ oraz $\operatorname{End}(V)$.

Stwierdzenie 10.6. Dla dowolnych przestrzeni liniowych V oraz W nad ciałem K zbiór $\mathscr{L}_K(V,W)$ rozważany wraz z działaniami $+\colon \mathscr{L}_K(V,W)\times\mathscr{L}_K(V,W)\to\mathscr{L}_K(V,W)$ oraz $\bullet\colon K\times\mathscr{L}_K(V,W)\to\mathscr{L}_K(V,W)$ określonymi dla wszystkich $f,g\in\mathscr{L}_K(V,W)$ i $\lambda\in K$ za pomocą wzorów:

$$(f+g)(v)=f(v)+g(v)\text{ dla każdego }v\in V \tag{10.6.1}$$

oraz

$$(\lambda\bullet f)(v)=\lambda\circ f(v)\text{ dla każdego }v\in V, \tag{10.6.2}$$

jest przestrzenią liniową nad ciałem K.

Dowód. Weźmy dowolne $f,g\in\mathscr{L}_K(V,W)$, $\lambda,\mu\in K$ oraz $v,v'\in V$. Wtedy $(f+g)(v+v')=f(v+v')+g(v+v')=\big(f(v)+f(v')\big)+\big(g(v)+g(v')\big)=\big(f(v)+g(v)\big)+\big(f(v')+g(v')\big)=(f+g)(v)+(f+g)(v')$ oraz $(f+g)(\mu\circ v)=f(\mu\circ v)+g(\mu\circ v)=\mu\circ f(v)+\mu\circ g(v)=\mu\circ\big(f(v)+g(v)\big)=\mu\circ(f+g)(v)$, więc $f+g\in\mathscr{L}_K(V,W)$. Ponadto $(\lambda\bullet f)(v+v')=\lambda\circ f(v+v')=\lambda\circ\big(f(v)+f(v')\big)=\lambda\circ f(v)+\lambda\circ f(v')=(\lambda\bullet f)(v)+(\lambda\bullet f)(v')$ i $(\lambda\bullet f)(\mu\circ v)=\lambda\circ f(\mu\circ v)=\lambda\circ\big(\mu\circ f(v)\big)=(\lambda\cdot\mu)\circ f(v)=(\mu\cdot\lambda)\circ f(v)=\mu\circ\big(\lambda\circ f(v)\big)=\mu\circ(\lambda\bullet f)(v)$, skąd $\lambda\bullet f\in\mathscr{L}_K(V,W)$. Zatem wzory (10.6.1) i (10.6.2) określają poprawnie zdefiniowane działania. Analogicznie jak w Przykładzie 7.10 pokazuje się, że są spełnione wszystkie aksjomaty przestrzeni liniowej.

Poniższe stwierdzenie opisuje rozważaną standardowo bazę przestrzeni liniowej $\mathscr{L}_K(V,W)$.

Stwierdzenie 10.7. Niech $(v_1,v_2,\ldots,v_n)$ oraz $(w_1,w_2,\ldots,w_m)$ będą bazami uporządkowanymi odpowiednio przestrzeni liniowych V i W nad ciałem K. Niech ponadto $I=\{1,2,\ldots,m\}$ oraz $J=\{1,2,\ldots,n\}$. Wówczas zbiór:

$$\mathscr{B}=\big\{\phi_{ij}\colon i\in I\wedge j\in J\big\} \tag{10.6.3}$$

odwzorowań liniowych $\phi_{ij}\colon V\to W$ określonych dla wszystkich $i\in I$ oraz $j\in J$ przez warunki:

$$\phi_{ij}(v_k)=\begin{cases} w_i & \text{, gdy } k=j\\ 0 & \text{, gdy } k\in J\setminus\{j\}\end{cases}, \tag{10.6.4}$$

jest bazą przestrzeni liniowej $\mathscr{L}_K(V,W)$.

Dowód. Wprost z Twierdzenia 10.2 wynika, że $\mathscr{B}\subseteq\mathscr{L}_K(V,W)$. W celu wykazania liniowej niezależności zbioru $\mathscr{B}$, dla wszystkich $i\in I$ oraz $j\in J$, rozważmy dowolne $a_{ij}\in K$ i przypuśćmy, że $\sum_{i\in I}\sum_{j\in J}a_{ij}\bullet\phi_{ij}=\Theta$, gdzie Θ oznacza zero przestrzeni liniowej $\mathscr{L}_K(V,W)$. Wówczas dla dowolnie ustalonego $k\in J$ otrzymujemy, że $\sum_{i\in I}\sum_{j\in J}a_{ij}\circ\phi_{ij}(v_k)=0$. Z (10.6.4) wynika więc, że $\sum_{i\in I}a_{ik}\circ w_i=0$. Ponadto $(w_1,w_2,\ldots,w_m)$ jest bazą K-przestrzeni W, skąd $a_{ik}=0$ dla każdego $i\in I$.

Wobec dowolności wyboru elementu $k \in J$ oznacza to, iż $a_{ij} = 0$ dla wszystkich $i \in I$ oraz $j \in J$. Zatem $\mathscr{B}$ jest liniowo niezależnym podzbiorem w $\mathscr{L}_K(V,W)$. W świetle Twierdzenia 8.5 pozostało udowodnić, że $\mathscr{L}_K(V,W) \subseteq \lin_K(\mathscr{B})$. Weźmy więc dowolne $\varphi \in \mathscr{L}_K(V,W)$. Dla każdego $k \in J$ istnieją wówczas określone jednoznacznie $b_{1k}, b_{2k}, \ldots, b_{mk} \in K$ takie, że $\varphi(v_k) = \sum_{i\in I} b_{ik} \circ w_i$. Stąd oraz na mocy (10.6.4), dla dowolnego $k \in J$ i odwzorowania liniowego Φ określonego równością $\Phi = \sum_{i\in I}\sum_{j\in J} b_{ij} \bullet \phi_{ij}$ otrzymujemy, że $\Phi(v_k) = \sum_{i\in I}\sum_{j\in J} b_{ij} \circ \phi_{ij}(v_k) = \sum_{i\in I} b_{ik} \circ w_i$. Zatem $\Phi(v_k) = \varphi(v_k)$ dla każdego $k \in J$, co w świetle Twierdzenia 10.3 oznacza równość $\varphi = \Phi$. Ponadto wprost z określenia funkcji Φ wynika, że $\Phi \in \lin_K(\mathscr{B})$, więc $f \in \lin_K(\mathscr{B})$ i ostatecznie $\mathscr{L}_K(V,W) \subseteq \lin_K(\mathscr{B})$.

Bezpośrednią konsekwencją powyższego stwierdzenia jest następujący

Wniosek 10.1. Dla dowolnych niezerowych skończenie wymiarowych przestrzeni liniowych V i W nad ciałem K, $\dim_K \mathscr{L}(V,W) = \dim_K V \cdot \dim_K W$.

Definicja 10.8. Bazę $\mathscr{B}$ zdefiniowaną w (10.6.3) przy pomocy warunków (10.6.4) nazywamy bazą przestrzeni liniowej $\mathscr{L}(V,W)$ indukowaną przez bazy uporządkowane $(v_1, v_2, \ldots, v_n)$ oraz $(w_1, w_2, \ldots, w_m)$ odpowiednio przestrzeni liniowych V i W nad ciałem K.

10.7 Macierz odwzorowania liniowego

Definicja 10.9. Niech $m, n \in \mathbb{N}$ i niech $X = (v_1, v_2, \ldots, v_n)$ oraz $Y = (w_1, w_2, \ldots, w_m)$ będą uporządkowanymi bazami odpowiednio n-wymiarowej przestrzeni liniowej V nad ciałem K oraz m-wymiarowej K-przestrzeni W. Niech ponadto $f\colon V \to W$ będzie odwzorowaniem liniowym. Macierz $[a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$ taką, że $f(v_k) = \sum_{i=1}^{m} a_{ik} \circ w_i$ dla każdego $k \in \{1, 2, \ldots, n\}$ nazywamy macierzą odwzorowania liniowego f w bazach X i Y. Oznaczamy ją przez M_{XY}^{f}.

Uwaga 10.11. Poprawność i jednoznaczność określenia macierzy M_{XY}^{f} wynika natychmiast z Twierdzenia 8.6.

Definicja 10.10. Macierz M_{XX}^{f} endomorfizmu f niezerowej, skończenie wymiarowej przestrzeni liniowej V nad ciałem K w bazach uporządkowanych X i X nazywamy krótko macierzą endomorfizmu f w bazie X.

Przykład 10.12. Zachowując wszystkie oznaczenia wprowadzone w Przykładzie 10.8 otrzymujemy, że $[a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$ jest macierzą odwzorowania liniowego f określonego wzorem (10.5.3) w bazach kanonicznych.

Przykład 10.13. Niech $\mathbb{R}_2[x] = \{ax^2+bx+c\colon a,b,c \in \mathbb{R}\}$. Bezpośrednie sprawdzenie pokazuje, że $\mathbb{R}_2[x]$ jest przestrzenią liniową nad ciałem $\mathbb{R}$. Wyznaczymy macierz odwzorowania liniowego $f\colon \mathrm{M}_2(\mathbb{R}) \to \mathbb{R}_2[x]$ danego wzorem:

$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ax^2+(b+3c)x+2d$$

w bazach:

$$X = \left(\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\right) \text{ i } Y = (x^2+1, x+1, 2).$$

Mamy:

$$f\left(\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}\right) = 3x^2+5x+4 = 3\cdot(x^2+1)+5\cdot(x+1)+(-2)\cdot 2,$$

$$f\left(\begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}\right) = x^2+4 = 1\cdot(x^2+1)+0\cdot(x+1)+\tfrac{3}{2}\cdot 2,$$

$$f\left(\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\right) = 2x^2+3x+2 = 2\cdot(x^2+1)+3\cdot(x+1)-\tfrac{3}{2}\cdot 2,$$

$$f\left(\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}\right) = x^2-4x+2 = 1\cdot(x^2+1)+(-4)\cdot(x+1)+\tfrac{5}{2}\cdot 2,$$

skąd:

$$M^f_{XY} = \begin{bmatrix} 3 & 1 & 2 & 1 \\ 5 & 0 & 3 & -4 \\ -2 & \frac{3}{2} & -\frac{3}{2} & \frac{5}{2} \end{bmatrix}.$$

Twierdzenie 10.3. Niech $m,n \in \mathbb{N}$ i niech $f\colon V \to W$ będzie odwzorowaniem liniowym n-wymiarowej przestrzeni liniowej V nad ciałem K w K-przestrzeń W wymiaru m. Jeżeli $[x_1,x_2,\ldots,x_n] \in K^n$ jest wektorem współrzędnych wektora $v \in V$ w bazie uporządkowanej $X = (v_1,v_2,\ldots,v_n)$ oraz $[y_1,y_2,\ldots,y_m] \in K^m$ jest wektorem współrzędnych wektora $f(v) \in W$ w bazie uporządkowanej $Y = (w_1,w_2,\ldots,w_m)$ i $A = \mathrm{M}^f_{XY}$, to:

$$A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}. \tag{10.7.1}$$

Ponadto $\dim_K \operatorname{im}(f) = r(A)$.

Dowód. Dla każdego $j \in \{1,2,\ldots,n\}$ istnieją określone jednoznacznie $y_{1j},y_{2j},\ldots,y_{mj} \in K$ takie, że $f(v_j) = \sum_{i=1}^{m} y_{ij} \circ w_i$. Zatem $A = [y_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$. Ponadto $f(v) =$

$f\left(\sum_{j=1}^{n} x_j \circ v_j\right) = \sum_{j=1}^{n} x_j \circ f(v_j) = \sum_{j=1}^{n} x_j \circ \sum_{i=1}^{m} y_{ij} \circ w_i = \sum_{j=1}^{n} \sum_{i=1}^{m} x_j \circ (y_{ij} \circ w_i) = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j \cdot y_{ij}) \circ w_i = \sum_{i=1}^{m} \sum_{j=1}^{n} (y_{ij} \cdot x_j) \circ w_i$. Wobec tego $y_i = \sum_{i=1}^{m} \sum_{j=1}^{n} y_{ij} \cdot x_j$ dla dowolnie ustalonego $i \in \{1,2,\ldots,m\}$. Ponadto wprost z określenia mnożenia macierzy wynika, że y_i jest i-tym wyrazem $m \times 1$-macierzy występującej po lewej stronie wzoru (10.7.1), co dowodzi prawdziwości tego wzoru.

Dalej, $\operatorname{im}(f) = \operatorname{lin}_k\big(f(v_1), f(v_2), \ldots, f(v_n)\big)$. Stąd oraz na mocy pierwszej części dowodu uzyskujemy, że:

$$\operatorname{im}(f) = \operatorname{lin}_K \left(\sum_{i=1}^{m} y_{i1} \circ w_i, \sum_{i=1}^{m} y_{i2} \circ w_i, \ldots, \sum_{i=1}^{m} y_{in} \circ w_i \right). \tag{10.7.2}$$

Ponadto Przykład 10.6 implikuje istnienie izomorfizmu $\varphi\colon W \to K^m$ takiego, że $\varphi(w_i) = \varepsilon_i$ dla każdego $i \in \{1,2,\ldots,m\}$. Z (10.7.2) oraz punktu (*viii*) Stwierdzenia 10.3 wynika więc, że $\varphi\big(\operatorname{im}(f)\big) = \operatorname{lin}_K\big(\sum_{i=1}^{m} y_{i1} \circ \varepsilon_i, \sum_{i=1}^{m} y_{i2} \circ \varepsilon_i, \ldots, \sum_{i=1}^{m} y_{in} \circ \varepsilon_i\big)$. Zatem $\dim_K \varphi\big(\operatorname{im}(f)\big) = \dim_K \operatorname{lin}\big([y_{11}, y_{21}, \ldots, y_{m1}], \ldots, [y_{1n}, y_{2n}, \ldots, y_{mn}]\big) = r(A)$. Ponadto $\varphi\big(\operatorname{im}(f)\big) \cong \operatorname{im}(f)$, więc $\dim_K \operatorname{im}(f) = \dim_K \varphi\big(\operatorname{im}(f)\big)$, czyli $\dim_K \operatorname{im}(f) = r(A)$.

Wniosek 10.2. Wymiar przestrzeni rozwiązań jednorodnego układu m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K wynosi $n - r(A)$, gdzie A jest macierzą współczynników tego układu.

Dowód. Z Przykładu 7.12 wynika, że zbiór Ω rozwiązań jednorodnego układu:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = 0 \end{cases}, \tag{10.7.3}$$

m równań liniowych z n niewiadomymi $x_1, x_2, \ldots, x_n$ nad ciałem K jest podprzestrzenią w K^n. Dla tego układu otrzymujemy oczywiście, że $A = [a_{ij}]_{ij} \in \mathrm{M}_{m\times n}(K)$. Z własności działań na macierzach wynika, że funkcja $f\colon K^n \to K^m$ określona wzorem:

$$f\big([x_1, x_2, \ldots, x_n]\big) = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{10.7.4}$$

dla wszystkich $[x_1, x_2, \ldots, x_n] \in K^n$, jest odwzorowaniem liniowym. Niech X i Y oznaczają odpowiednio bazy kanoniczne przestrzeni K^n i K^m. Wprost z określenia przekształcenia liniowego f wynika wówczas, że $A = \mathrm{M}_{XY}^{f}$ oraz $[\xi_1, \xi_2, \ldots, \xi_n] \in \ker(f)$ wtedy i tylko wtedy, gdy ciąg $(\xi_1, \xi_2, \ldots, \xi_n)$ jest rozwiązaniem układu (10.7.3). Zatem przy naturalnym i wzajemnie jednoznacznym utożsamieniu ciągu

$(\xi_1, \xi_2, \ldots, \xi_n)$ z wektorem $[\xi_1, \xi_2, \ldots, \xi_n]$ przestrzeni K^n otrzymujemy równość $\Omega = \ker(f)$. Ponadto $\dim_K \operatorname{im}(f) = r(A)$ i $\dim K^n = \dim\ker(f) + \dim\operatorname{im}(f)$ odpowiednio na mocy Twierdzeń 10.3 i 10.1. Zatem $\dim\Omega = n - r(A)$.

Przykład 10.14. Wyznaczymy jednorodny układ równań liniowych nad ciałem K, którego przestrzenią rozwiązań jest $U = \operatorname{lin}_{\mathbb{R}}\big([0,1,2,4],[1,2,3,5]\big)$. Ponieważ:

$$\begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} = -1 \neq 0,$$

to z Twierdzenia 9.7 i określenia rzędu macierzy wynika więc, że $\dim U = 2$. Ponadto:

$$r\begin{bmatrix} 0 & 1 & 2 & 4 \\ 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 1 + r\begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} = 2 + r\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = 2 + 2 = 4,$$

więc $\big\{[0,1,2,4],[1,2,3,5],[0,0,1,0],[0,0,0,1]\big\}$ jest bazą przestrzeni $\mathbb{R}^3$. Oznaczmy $e_1 = [0,1,2,4]$, $e_2 = [1,2,3,5]$, $e_3 = [0,0,1,0]$ i $e_4 = [0,0,0,1]$. Z Twierdzenia 10.2 wynika, że funkcja $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ określona dla wszystkich $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ za pomocą wzoru:

$$f\left(\sum_{i=1}^{4} \lambda_i e_i\right) = \lambda_3 \circ e_3 + \lambda_4 \circ e_4,$$

jest endomorfizmem liniowym. Wprost z określenia funkcji f wynika, że $U \subseteq \ker(f)$ oraz $\dim\operatorname{im}(f) = \dim\operatorname{lin}(e_3, e_4) = 2$. Stąd oraz na mocy Twierdzenia 10.1, $\dim\ker(f) = 2$. Z Twierdzenia 8.9 wynika więc, że $U = \ker(f)$. Wyznaczymy teraz wzór analityczny przekształcenia liniowego f. Niech $\mathscr{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ Wtedy:

$$f(\varepsilon_1) = f(e_2 - 2 \circ e_1 + e_3 + 3 \circ e_4) = e_3 + 3 \circ e_4 = \varepsilon_3 + 3 \circ \varepsilon_4,$$

$$f(\varepsilon_2) = f(e_1 - 2 \circ e_3 - 4 \circ e_4) = -2\varepsilon_3 - 4\varepsilon_4,$$

$$f(\varepsilon_3) = \varepsilon_3 \text{ oraz } f(\varepsilon_4) = \varepsilon_4.$$

Ponieważ $[x,y,z,t] = x \circ \varepsilon_1 + y \circ \varepsilon_2 + z \circ \varepsilon_3 + t \circ \varepsilon_4$ dla wszystkich $x, y, z, t \in \mathbb{R}$, to:

$$f\big([x,y,z,t]\big) = [0,0,x-2y+z-t,3x-4y+z+t]$$

jest wzorem analitycznym endomorfizmu liniowego f, skąd:

$$\mathrm{M}^f_{\mathscr{B}\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 3 & -4 & 0 & 1 \end{bmatrix}.$$

Ponadto z Twierdzenia 10.3 wynika, że dla $A = \mathrm{M}^f_{\mathscr{B}\mathscr{B}}$ oraz wszystkich $x,y,z,t \in \mathbb{R}$ zachodzi równość:

$$f\big([x,y,z,t]\big) = A \cdot \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}.$$

Zatem U jest przestrzenią rozwiązań jednorodnego układu równań liniowych:

$$\begin{cases} x - 2y + z = 0 \\ 3x - 4y + t = 0 \end{cases}$$

nad ciałem $\mathbb{R}$.

Twierdzenie 10.4. Niech X, Y oraz Z będą kolejno bazami uporządkowanymi przestrzeni liniowych V, W i U nad ciałem K. Jeżeli $f\colon V \to W$ i $g\colon W \to U$ są odwzorowaniami liniowymi oraz $h = g \circ f$, to $M^h_{XZ} = \mathrm{M}^g_{YZ} \cdot \mathrm{M}^f_{XY}$.

Dowód. Niech $n = \dim_K V$, $m = \dim_K W$, $s = \dim_K U$, $A = \mathrm{M}^f_{XY}$, $B = \mathrm{M}^g_{YZ}$ i $C = M^h_{XZ}$. Weźmy dowolne $v \in \mathbb{V}$. Niech $[x_1, x_2, \ldots, x_n] \in K^n$ będzie wektorem współrzędnych wektora v w bazie X i niech $[y_1, y_2, \ldots, y_m] \in K^m$ będzie wektorem współrzędnych wektora $f(v)$ w bazie Y. Z Twierdzenia 10.3 wynika wtedy, że zachodzi równość (10.7.1). Niech $[z_1, z_2, \ldots, z_s] \in K^s$ będzie wektorem współrzędnych wektora $g\big(f(v)\big)$ w bazie Z. Powołując się ponownie na Twierdzenie 10.3 otrzymujemy wówczas równości:

$$B \cdot \left(A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_s \end{bmatrix} \tag{10.7.5}$$

oraz

$$C \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_s \end{bmatrix}. \tag{10.7.6}$$

Ponadto:

$$B \cdot \left(A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = (B \cdot A) \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

więc z (10.7.5) wynika, że:

$$(B\cdot A)\cdot\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}z_1\\z_2\\\vdots\\z_s\end{bmatrix}.$$

Stąd oraz na mocy (10.7.6), dowolności wyboru wektora $v\in V$ i Stwierdzenia 5.7 otrzymujemy, że $C=B\cdot A$.

10.8 Izomorfizm przestrzeni liniowych $\mathscr{L}_K(V,W)$ i $\mathrm{M}_{m\times n}(K)$

Twierdzenie 10.5. Niech $X=(v_1,v_2,\ldots,v_n)$ oraz $Y=(w_1,w_2,\ldots,w_m)$ będą bazami uporządkowanymi odpowiednio przestrzeni liniowych V i W nad ciałem K. Wówczas odwzorowanie $\varphi\colon \mathscr{L}_K(V,W)\to \mathrm{M}_{m\times n}(K)$ określone wzorem:

$$\varphi(f)=\mathrm{M}_{XY}^f \text{ dla każdego } f\in\mathscr{L}_K(V,W) \tag{10.8.1}$$

jest izomorfizmem liniowym.

Dowód. Poprawność określenia funkcji φ wynika wprost z Uwagi 10.11. Weźmy dowolne $f,g\in\mathscr{L}_K(V,W)$, $\lambda\in K$, oraz $k\in\{1,2,\ldots,n\}$. Oznaczmy $A=\mathrm{M}_{XY}^f$ i $B=\mathrm{M}_{XY}^g$. Wtedy ε_k jest wektorem współrzędnych wektora v_k w bazie X oraz $A=[a_{ij}]_{ij}\in\mathrm{M}_{m\times n}(K)$ i $B=[b_{ij}]_{ij}\in\mathrm{M}_{m\times n}(K)$. Z Twierdzenia 10.3 wynika więc, że $(f+g)(v_k)=f(v_k)+g(v_k)=\sum_{i=1}^m a_{ik}\circ w_i+\sum_{i=1}^m b_{ik}\circ w_i=\sum_{i=1}^m(a_{ik}+b_{ik})\circ w_i$. Stąd oraz na mocy Definicji 10.9 i Uwagi 10.11 otrzymujemy, że $A+B$ jest macierzą odwzorowania liniowego $f+g$ w bazach X i Y. Zatem $\varphi(f+g)=\mathrm{M}_{XY}^{f+g}=\mathrm{M}_{XY}^f+\mathrm{M}_{XY}^g=\varphi(f)+\varphi(g)$. Analogicznie, $(\lambda\bullet f)(v_k)=\lambda\circ f(v_k)=\lambda\circ\sum_{i=1}^m a_{ik}\circ w_i=\sum_{i=1}^m\lambda\circ(a_{ik}\circ w_i)=\sum_{i=1}^m(\lambda\cdot a_{ik})\circ w_i$, więc $\lambda\cdot A$ jest macierzą przekształcenia liniowego $\lambda\bullet f$ w bazach X i Y. Wobec tego $\varphi(\lambda\bullet f)=\mathrm{M}_{XY}^{\lambda\bullet f}=\lambda\cdot\mathrm{M}_{XY}^f=\lambda\cdot\varphi(f)$. Wobec tego φ jest odwzorowaniem liniowym K-przestrzeni $\mathscr{L}_K(V,W)$ w K-przestrzeń $\mathrm{M}_{m\times n}(K)$.

Załóżmy, że $f\in\ker(\varphi)$. Wtedy $\varphi(f)=\Theta_{m\times n}$, czyli $\mathrm{M}_{XY}^f=\Theta_{m\times n}$. Stąd $f(v_k)=0$ dla każdego $k\in\{1,2,\ldots,n\}$ i w konsekwencji $f=\Theta$. Zatem $\ker\varphi=\{\Theta_{m\times n}\}$, co w świetle Twierdzenia oznacza, że φ jest monomorfizmem.

Weźmy dowolne $C=[c_{ij}]_{ij}\in\mathrm{M}_{m\times n}(K)$. Rozważmy funkcję F określoną równością $F=\sum_{i=1}^m\sum_{j=1}^n c_{ij}\bullet\phi_{ij}$, gdzie $\{\phi_{ij}\colon i\in\{1,2,\ldots,m\}\wedge j\in\{1,2,\ldots,n\}\}$ jest bazą K-przestrzeni $\mathscr{L}_K(V,W)$ indukowaną przez bazy X i Y. Wtedy $f\in\mathscr{L}_K(V,W)$ oraz dla każdego $k\in\{1,2,\ldots,n\}$ otrzymujemy, że $F(v_k)=\sum_{i=1}^m\sum_{j=1}^n c_{ij}\circ\phi_{ij}(v_k)=\sum_{i=1}^m c_{ik}\circ w_i$, skąd $C=\mathrm{M}_{XY}^F=\varphi(F)$. Wobec tego φ jest również epimorfizmem i w konsekwencji φ jest izomorfizmem K-przestrzeni $\mathscr{L}_K(V,W)$ na K-przestrzeń $\mathrm{M}_{m\times n}(K)$.

Twierdzenie 10.6. Niech X oraz Y będą bazami uporządkowanymi odpowiednio przestrzeni liniowych V i W nad ciałem K i niech $f \in \mathscr{L}_K(V,W)$. Wówczas f jest izomorfizmem liniowym wtedy i tylko wtedy, gdy M_{XY}^f jest macierzą odwracalną.

Dowód. Niech $n = \dim_K V$ i $m = \dim_K W$. Załóżmy najpierw, że f jest izomorfizmem. Wówczas $m = n$ oraz istnieje $g \in \mathscr{L}_K(W,V)$ takie, że $g = f^{-1}$. W szczególności, $g \circ f = \mathrm{id}_V$, więc $\mathrm{M}_{XX}^{g\circ f} = I_n$. Stąd oraz na mocy Twierdzenia 10.4, $\mathrm{M}_{YX}^g \cdot \mathrm{M}_{XY}^f = I_n$. Odwracalność macierzy M_{XY}^f jest więc konsekwencją Stwierdzenia 5.11.

Przypuśćmy teraz, że macierz M_{XY}^f jest odwracalna i oznaczmy $A = \mathrm{M}_{XY}^f$. Wtedy A jest macierzą kwadratową stopnia $n = m$ oraz istnieje macierz $B \in \mathrm{M}_n(K)$ taka, że $A \cdot B = B \cdot A = I_n$. Ponadto z Twierdzenia 10.5 wynika istnienie takiego $h \in \mathscr{L}_K(W,V)$, że $B = \mathrm{M}_{YX}^h$. Stąd oraz na mocy Twierdzenia 10.4, $\mathrm{M}_{XX}^{h\circ f} = I_n$ oraz $\mathrm{M}_{YY}^{f\circ h} = I_n$. Zatem $h \circ f = \mathrm{id}_V$ i $f \circ h = \mathrm{id}_W$, co oznacza, że $h = f^{-1}$. Zatem f jest izomorfizmem liniowym.

Bezpośrednią konsekwencją pierwszej części dowodu Twierdzenia 10.6 jest następujący

Wniosek 10.3. Niech X i Y będą bazami uporządkowanymi odpowiednio przestrzeni liniowych V oraz W nad ciałem K. Jeżeli $f\colon V \to W$ jest izomorfizmem liniowym, $A = \mathrm{M}_{XY}^f$ i $g = f^{-1}$, to $A^{-1} = \mathrm{M}_{YX}^g$.

10.9 Macierz przejścia

Definicja 10.11. Niech X oraz X' będą uporządkowanymi bazami przestrzeni liniowej V nad ciałem K. Macierzą przejścia od bazy X do bazy X' nazywamy macierz odwzorowania identycznościowego na V w bazach X' i X. Macierz przejścia od bazy X do bazy X' oznaczamy przez $P_{X\to X'}$.

Z powyższej definicji wynika, że:

$$P_{X\to X'} = \mathrm{M}_{X'X}^{\mathrm{id}_V}, \text{ gdzie } \mathrm{id}_V(v) = v \text{ gdzie dla każdego } v \in V. \tag{10.9.1}$$

Przykład 10.15. W przestrzeni liniowej $V = \mathrm{M}_2(\mathbb{R})$ rozważmy bazę uporządkowaną $X' = (E_{11}, E_{12}, E_{21}, E_{22})$, zaś w przestrzeni liniowej $W = \mathbb{R}_2[x]$ bazę uporządkowaną $Y' = (x^2, x, 1)$. Niech ponadto X i Y będą bazami uporządkowanymi tych przestrzeni określonymi w Przykładzie 10.13. Wtedy:

$$P_{X'\to X} = \mathrm{M}^{\mathrm{id}_V}_{XX'} = \begin{bmatrix} 3 & 1 & 2 & 1 \\ 2 & 3 & 0 & 2 \\ 1 & -1 & 1 & -2 \\ 2 & 2 & 1 & 1 \end{bmatrix}, \; P_{X\to X'} = (P_{X'\to X})^{-1} = \frac{1}{6}\begin{bmatrix} 3 & 6 & 3 & -9 \\ -4 & -2 & 0 & 8 \\ -1 & -8 & -3 & 11 \\ 3 & 0 & -3 & -3 \end{bmatrix},$$

$$P_{Y'\to Y} = \mathrm{M}^{\mathrm{id}_W}_{YY'} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \text{ oraz } P_{Y\to Y'} = (P_{Y'\to Y})^{-1} = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$

Bezpośrednią konsekwencją wzoru (10.9.1) oraz Twierdzenia 10.4 jest następujący

Wniosek 10.4. Niech X, X' oraz X'' będą bazami uporządkowanymi przestrzeni liniowej V nad ciałem K. Wówczas:

$$P_{X''\to X} = P_{X''\to X'} \cdot P_{X'\to X}.$$

Wniosek 10.5. Jeżeli X oraz X' są uporządkowanymi bazami przestrzeni liniowej V nad ciałem K, to macierz $P_{X\to X'}$ jest odwracalna oraz $P_{X'\to X} = (P_{X\to X'})^{-1}$.

Dowód. Niech $n = \dim_K(V)$. Wówczas $P_{X\to X} = I_n$. Na mocy Wniosku 10.4 otrzymujemy więc, że $I_n = P_{X\to X'} \cdot P_{X'\to X}$, skąd wynika odwracalność macierzy $P_{X\to X'}$ oraz równość $P_{X'\to X} = (P_{X\to X'})^{-1}$ (por. Stwierdzenie 5.11).

10.9.1 Zmiana baz a macierz odwzorowania liniowego

Twierdzenie 10.7. Niech V oraz W będą niezerowymi, skończenie wymiarowymi przestrzeniami liniowymi nad ciałem K, niech X i X' będą uporządkowanymi bazami przestrzeni V i niech Y oraz Y' będą uporządkowanymi bazami przestrzeni W. Niech ponadto $f\colon V \to W$ będzie odwzorowanie liniowym. Wtedy:

$$M^f_{X'Y'} = P_{Y'\to Y} \cdot M^f_{XY} \cdot P_{X\to X'}. \tag{10.9.2}$$

Dowód. Ponieważ $f \circ \mathrm{id}_V = f$, to z (10.9.1) oraz Twierdzenia 10.4 wynika, że:

$$M^f_{XY} \cdot P_{X\to X'} = \mathrm{M}^f_{X'Y} \tag{10.9.3}$$

Ponadto $\mathrm{id}_W \circ f = f$, więc powołując się ponownie na (10.9.1) i Twierdzenie 10.4 uzyskujemy:

$$P_{Y'\to Y} \cdot \mathrm{M}^f_{X'Y} = M^f_{X'Y'}. \tag{10.9.4}$$

Podstawiając (10.9.3) do (10.9.4) otrzymujemy (10.9.2).

Przykład 10.16. Niech f będzie odwzorowaniem liniowym określonym w Przykładzie 10.13. Na mocy Twierdzenia 10.7 i Przykładu 10.15 otrzymujemy wówczas:

$$M^f_{X'Y'} = P_{Y'\to Y}\cdot M^f_{XY}\cdot P_{X\to X'} = \begin{bmatrix} 1&0&0&0\\ 0&1&3&0\\ 0&0&0&2 \end{bmatrix}.$$

Oczywiście ten sam wynik uzyskamy, wyznaczając macierz $M^f_{X'Y'}$ z definicji (tzn. w sposób podany w Przykładzie 10.13).

Bezpośrednią konsekwencją Twierdzenia 10.7 oraz Wniosku 10.5 jest następujące

Stwierdzenie 10.8. Niech X oraz X' będą uporządkowanymi bazami przestrzeni liniowej V nad ciałem K i niech f będzie endomorfizmem tej przestrzeni. Niech ponadto $P = P_{X\to X'}$. Wówczas:

$$\mathrm{M}^f_{X'X'} = P^{-1}\cdot \mathrm{M}^f_{XX}\cdot P.$$

10.9.2 Podobieństwo macierzy

Definicja 10.12. Niech A i B będą macierzami kwadratowymi stopnia n nad ciałem K. Mówimy, że macierz A jest podobna do macierzy B, gdy istnieje nieosobliwa macierz P nad ciałem K taka, że $B = P^{-1}\cdot A\cdot P$. Piszemy wówczas $A\sim B$.

Stwierdzenie 10.9. Dla dowolnej liczby naturalnej n oraz dowolnego ciała K, relacja podobieństwa macierzy jest relacją równoważności na zbiorze $\mathrm{M}_n(K)$.

Dowód. Weźmy dowolne $A,B,C\in \mathrm{M}_n(K)$. Wówczas $A = I_n^{-1}\cdot A\cdot I_n$. Zatem relacja $\sim$ jest zwrotna. Przypuśćmy, że $A\sim B$. Wtedy $B = P^{-1}\cdot A\cdot P$ dla pewnego $P\in \mathrm{M}_n(K)$ takiego, że $\det P\neq 0$. Zatem $A = \left(P^{-1}\right)^{-1}\cdot B\cdot P^{-1}$, czyli $B\sim A$. Zatem relacja $\sim$ jest symetryczna. Załóżmy dodatkowo, że $B\sim C$. Istnieje wówczas $D\in \mathrm{M}_n(K)$ takie, że $\det D\neq 0$ oraz $C = D^{-1}\cdot B\cdot D$. Stąd $C = D^{-1}\cdot(P^{-1}\cdot A\cdot P)\cdot D = \left(D^{-1}\cdot P^{-1}\right)\cdot A\cdot (P\cdot D) = (P\cdot D)^{-1}\cdot A\cdot (P\cdot D)$. Wobec tego relacja $\sim$ jest przechodnia i ostatecznie uzyskujemy, że $\sim$ jest relacją równoważności.

Uwaga 10.12. Uzasadniona wyżej symetria relacji $\sim$ uzasadnia poprawność sformułowania: „Macierze kwadratowe A i B stopnia n nad ciałem K są podobne".

Twierdzenie 10.8. Macierze kwadratowe A i B stopnia n nad ciałem K są podobne wtedy i tylko wtedy, gdy są one macierzami pewnego endomorfizmu f przestrzeni K^n w pewnych bazach tej przestrzeni. Ponadto, jeśli macierze A i B są podobne, to $\det(B) = \det(A)$ oraz $r(B) = r(A)$.

Dowód. Jeżeli A i B są macierzami pewnego endomorfizmu f przestrzeni K^n w pewnych bazach tej przestrzeni, to $A\sim B$ na mocy Stwierdzenia 10.8.

Załóżmy teraz, że $A \sim B$. Istnieje wówczas macierz $P = [p_{ij}]_{ij} \in \mathrm{M}_n(K)$ taka, że $\det P \neq 0$ oraz $B = P^{-1} \cdot A \cdot P$. Z Twierdzenia Cauchy'ego (zob. Twierdzenie 5.1) wynika więc, że $det(B) = \det(A)$. Dalej, niech $\mathscr{B}$ będzie bazą kanoniczną przestrzeni liniowej K^n. Twierdzenie 10.5 implikuje istnienie endomorfizmu f przestrzeni K^n takiego, że $\mathrm{M}^f_{\mathscr{B}\mathscr{B}} = A$. Ponadto $r(A) = \dim_K \mathrm{im}(f)$ na mocy Twierdzenia 10.3. Dla każdego $j \in \{1, 2, \ldots, n\}$ oznaczmy $\rho_j = \left[p_{1j}, p_{2j}, \ldots, p_{nj}\right]$. Ponieważ $\det P \neq 0$, to z Twierdzenia 9.6 wynika, że $r(P) = n$. Zatem $\mathscr{B}' = (\rho_1, \rho_2, \ldots, \rho_n)$ jest uporządkowaną bazą przestrzeni K^n. Wobec tego $P = P_{\mathscr{B} \to \mathscr{B}'}$. Stąd oraz na mocy Stwierdzenia 10.8, $B = \mathrm{M}^f_{\mathscr{B}'\mathscr{B}'}$. Twierdzenie 10.3 implikuje więc, że $r(B) = \dim_K \mathrm{im}(f)$, co wobec wcześniej uzyskanej równości $r(A) = \dim_K \mathrm{im}(f)$ oznacza, że $r(B) = r(A)$.

Rozdział 11

Elementy geometrii analitycznej w przestrzeni

Od tej pory słowa „wektor" będziemy używać również w znanym ze szkoły średniej kontekście geometrycznym. Podobnie jak w przypadku dwuwymiarowym, każdy element $[x,y,z]$ przestrzeni liniowej $\mathbb{R}^3$ możemy traktować jako:

(i) punkt o współrzędnych x,y,z; piszemy wówczas (x,y,z) zamiast $[x,y,z]$;
(ii) wektor zaczepiony w punkcie $(0,0,0)$, zwany wektorem wodzącym o współrzędnych x,y,z;
(iii) reprezentant zbioru wszystkich wektorów zaczepionych w dowolnym punkcie, które mają tę samą długość, ten sam kierunek i ten sam zwrot co wektor wodzący o współrzędnych x,y,z. Zbiór ten nazywamy wektorem swobodnym wyznaczonym przez $[x,y,z]$.

Punkty przestrzeni $\mathbb{R}^3$ będziemy oznaczali tradycyjnie wielkimi literami alfabetu łacińskiego: $A,B,C,\ldots$, przy czym litera O przypisana będzie zawsze punktowi $(0,0,0)$. W szczególności, zapis $P=(x,y,z)$ oznacza, że P jest punktem o współrzędnych x,y,z. Łatwo zauważyć, że dowolny punkt $P=(x,y,z)$ wyznacza dokładnie jeden wektor wodzący. Jest nim oczywiście wektor wodzący o współrzędnych x,y,z oznaczany przez $\overrightarrow{OP}$. Chcąc podkreślić, że element $[x,y,z]$ przestrzeni $\mathbb{R}^3$ traktujemy jako wektor wodzący wyznaczony przez punkt $P=(x,y,z)$ stosujemy notację: $\overrightarrow{OP}=[x,y,z]$.

Powołując się na wiadomości ze Wstępu do logiki i teorii mnogości, możemy odnotować fakt, że zbiór wszystkich wektorów swobodnych przestrzeni $\mathbb{R}^3$ jest zbiorem ilorazowym zbioru wszystkich wektorów tej przestrzeni względem relacji równoważności utożsamiającej wektory o tej samej długości oraz tym samym kierunku i zwrocie. W szczególności wynika stąd, że każdy wektor swobodny jest określony jednoznacznie przez swój dowolny element, a więc także przez należący do niego dokładnie jeden wektor wodzący. Dlatego często utożsamia się wektor swobodny z jego dowolnym reprezentantem i na ogół nie prowadzi to do nieporozumień.

Aby zaznaczyć, że element $[x,y,z]$ przestrzeni $\mathbb{R}^3$ rozważamy jako wektor w znaczeniu geometrycznym, stosujemy notację $\vec{v}=[x,y,z]$.

Powołując się na określenie działań w $\mathbb{R}$-przestrzeni liniowej $\mathbb{R}^3$, dla dowolnych $\vec{v}=[v_1,v_2,v_3]\in\mathbb{R}^3$ i $\vec{u}=[u_1,u_2,u_3]\in\mathbb{R}^3$ oraz $\lambda\in\mathbb{R}$ definiujemy:

$$\vec{v}+\vec{u}=[v_1+u_1,v_2+u_2,v_3+u_3]$$

oraz

$$\lambda\circ\vec{v}=[\lambda v_1,\lambda v_2,\lambda v_3].$$

Zauważmy, że dowolne dwa punkty $A=(a_1,a_2,a_3)$ i $B=(b_1,b_2,b_3)$ przestrzeni $\mathbb{R}^3$ wyznaczają dokładnie dwa wektory swobodne o reprezentantach: $[b_1-a_1,b_2-a_2,b_3-a_3]$ oraz $[a_1-b_1,a_2-b_2,a_3-b_3]$. Wektor $[b_1-a_1,b_2-a_2,b_3-a_3]$ nazywamy wektorem o początku w punkcie $A=(a_1,a_2,a_3)$ oraz końcu w punkcie $B=(b_1,b_2,b_3)$ i oznaczamy symbolem $\overrightarrow{AB}$.

Ponieważ $\overrightarrow{AB}+\overrightarrow{BA}=[0,0,0]$, to wektor $\overrightarrow{BA}$ nazywamy wektorem przeciwnym do wektora $\overrightarrow{AB}$ i stosujemy notację: $\overrightarrow{BA}=-\overrightarrow{AB}$ (co jest zgodne również z motywacją geometryczną, gdyż wektory $\overrightarrow{AB}$ i $\overrightarrow{BA}$ mają przeciwne zwroty). Wektor zerowy $[0,0,0]$ oznaczamy krótko przez $\vec{0}$.

11.1 Długość wektora

Definicja 11.1. Długością wektora $\vec{v}=[x,y,z]$ przestrzeni $\mathbb{R}^3$ nazywamy liczbę $\|\vec{v}\|$ określoną wzorem:

$$\|\vec{v}\|=\sqrt{x^2+y^2+z^2}. \tag{11.1.1}$$

Zanim udowodnimy podstawowe własności długości wektora, wykażemy ważną nierówność Cauchy'ego-Buniakowskiego-Schwarza potrzebną do wykazania tzw. nierówności trójkąta (zob. punkt (iv) Stwierdzenia 11.1).

Twierdzenie 11.1 (Cauchy, Buniakowski, Schwarz). Niech n będzie liczbą naturalną. Wówczas:

$$\left(\sum_{i=1}^{n}a_ib_i\right)^2\leq\left(\sum_{i=1}^{n}a_i^2\right)\cdot\left(\sum_{i=1}^{n}b_i^2\right), \tag{11.1.2}$$

dla wszystkich liczb rzeczywistych $a_1,a_2,\ldots,a_n,b_1,b_2,\ldots,b_n$. Ponadto w $(11.1.2)$ zachodzi równość wtedy i tylko wtedy, gdy istnieje liczba rzeczywista λ taka, że $a_i=\lambda b_i$ dla każdego $i\in\{1,2,\ldots,n\}$.

Dowód. Jeżeli $\sum_{i=1}^{n}a_i^2=0$, to $a_1=a_2=\ldots=a_n=0$, więc w $(11.1.2)$ zachodzi równość i wystarczy przyjąć $\lambda=0$.

Załóżmy teraz, że $\sum_{i=1}^{n}a_i^2>0$ i rozważmy wielomian:

$$f=\sum_{i=1}^{n}(a_ix-b_i)^2\in\mathbb{R}[x]. \tag{11.1.3}$$

Wtedy $\mathrm{st}(f)=2$ oraz $f(c)\geq 0$ dla każdego $c\in\mathbb{R}$, więc trójmian kwadratowy f ma co najwyżej jeden pierwiastek. Zatem wyróżnik Δ_f wielomianu f spełnia nierówność $\Delta_f\leq 0$. Ponieważ $f=\left(\sum_{i=1}^{n}a_i^2\right)x^2-2\left(\sum_{i=1}^{n}a_ib_i\right)x+\sum_{i=1}^{n}b_i^2$, to:

$$\frac{\Delta_f}{4} = \left(\sum_{i=1}^{n} a_i b_i\right)^2 - \left(\sum_{i=1}^{n} a_i^2\right) \cdot \left(\sum_{i=1}^{n} b_i^2\right). \tag{11.1.4}$$

Wobec tego nierówność (11.1.2) równoważna jest prawdziwej nierówności $\Delta_f \leq 0$. W szczególności wynika stąd, że jeżeli w (11.1.2) zachodzi równość, to $\Delta_f = 0$. Wówczas wielomian f posiada dokładnie jeden pierwiastek x_0. Zatem na mocy (11.1.3) otrzymujemy, że $b_i = a_i x_0$ dla każdego $i \in \{1,2,\ldots,n\}$, więc wystarczy przyjąć $\lambda = x_0$.

Pozostało wykazać, że istnienie liczby rzeczywistej λ takiej, że $a_i = \lambda b_i$ dla każdego $i \in \{1,2,\ldots,n\}$, implikuje, że w (11.1.2) zachodzi równość. Przypuśćmy więc, że taka liczba λ istnieje. Jeżeli $\lambda b_i = 0$ dla każdego $i \in \{1,2,\ldots,n\}$, to $a_i = 0$ dla każdego $i \in \{1,2,\ldots,n\}$ i żądana równość zachodzi. Jeśli natomiast $\lambda b_j \neq 0$ dla pewnego $j \in \{1,2,\ldots,n\}$, to $a_j \neq 0$. Wtedy $\sum_{i=1}^{n} a_i^2 > 0$, więc trójmian kwadratowy f określony w (11.1.3) przyjmuje postać $f = \left(\sum_{i=1}^{n} a_i^2\right) \cdot (x-\lambda)^2$. Zatem λ jest jego jedynym pierwiastkiem. Stąd $\Delta_f = 0$, co wobec (11.1.4) oznacza, że w (11.1.2) zachodzi równość.

Omówimy teraz podstawowe własności długości wektora:

Stwierdzenie 11.1. Dla dowolnych wektorów $\vec{u}$ i $\vec{v}$ przestrzeni $\mathbb{R}^3$ oraz dowolnej liczby rzeczywistej α:

(i) $\|\vec{v}\| \geq 0$;
(ii) $\|\vec{v}\| = 0$ wtedy i tylko wtedy, gdy $\vec{v} = \vec{0}$;
(iii) $\|\lambda \circ \vec{v}\| = |\lambda| \cdot \|\vec{v}\|$;
(iv) $\|\vec{v} + \vec{u}\| \leq \|\vec{v}\| + \|\vec{u}\|$.

Dowód. Własności (i) oraz (ii) wynikają wprost ze wzoru (11.1.1). Aby uzasadnić dwie pozostałe własności, rozważmy dowolne wektory $\vec{u} = [a,b,c]$ i $\vec{v} = [x,y,z]$ przestrzeni $\mathbb{R}^3$ oraz dowolną liczbę rzeczywistą λ. Wtedy $\|\lambda \circ \vec{v}\| = \big\|\lambda \circ [x,y,z]\big\| = \big\|[\lambda x, \lambda y, \lambda z]\big\| = \sqrt{(\lambda x)^2 + (\lambda y)^2 + (\lambda z)^2} = \sqrt{\lambda^2 \cdot (x^2+y^2+z^2)} = \sqrt{\lambda^2} \cdot \sqrt{x^2+y^2+z^2} = |\lambda| \cdot \|\vec{v}\|$, co uzasadnia własność (iii). Dalej, $\|\vec{v}+\vec{u}\|^2 = (x+a)^2 + (y+b)^2 + (z+c)^2 = (x^2+y^2+z^2) + (a^2+b^2+c^2) + 2 \cdot (xa+yb+zc) = \|\vec{v}\|^2 + \|\vec{u}\|^2 + 2 \cdot (xa+yb+zc)$ oraz $(\|\vec{v}\| + \|\vec{u}\|)^2 = \|\vec{v}\|^2 + \|\vec{u}\|^2 + 2 \cdot \|\vec{v}\| \cdot \|\vec{u}\|$. Ponadto $\|\vec{v}\| = \sqrt{x^2+y^2+z^2}$, $\|\vec{u}\| = \sqrt{a^2+b^2+c^2}$ oraz na mocy Twierdzenia 11.1 uzyskujemy, że $xa+yb+zc \leq |xa+yb+zc| = \sqrt{(xa+yb+zc)^2} \leq \sqrt{(x^2+y^2+z^2) \cdot (a^2+b^2+c^2)} = \sqrt{x^2+y^2+z^2} \cdot \sqrt{a^2+b^2+c^2}$. Zatem $\|\vec{v}+\vec{u}\|^2 \leq (\|\vec{v}\| + \|\vec{u}\|)^2$, skąd wynika nierówność dana w (iv).

11.2 Iloczyn skalarny

Definicja 11.2. Iloczynem skalarnym wektorów $\vec{v}$ i $\vec{u}$ przestrzeni $\mathbb{R}^3$ nazywamy liczbę rzeczywistą $\langle\vec{v}|\vec{u}\rangle$ określoną wzorem:

$$\langle\vec{v}|\vec{u}\rangle = \|\vec{v}\| \cdot \|\vec{u}\| \cdot \cos\varphi, \tag{11.2.1}$$

gdzie φ jest kątem między wektorami $\vec{v}$ i $\vec{u}$, przy czym przyjmujemy, że $\varphi \in [0,\pi]$. Kąt ten oznacza się przez $\measuredangle(\vec{v},\vec{u})$.

Stwierdzenie 11.2. Dla dowolnych wektorów $\vec{v} = [v_1, v_2, v_3]$ i $\vec{u} = [u_1, u_2, u_3]$ przestrzeni $\mathbb{R}^3$:

$$\langle\vec{v}|\vec{u}\rangle = v_1u_1 + v_2u_2 + v_3u_3. \tag{11.2.2}$$

Dowód. Jeżeli $\vec{u} = \vec{0}$ lub $\vec{v} = \vec{0}$, to teza jest bezpośrednią konsekwencją wzorów (11.2.1) oraz (11.1.1).

Załóżmy teraz, że $\vec{u} \neq \vec{0}$ i $\vec{v} \neq \vec{0}$. Oznaczmy $\varphi = \measuredangle(\vec{v},\vec{u})$. Przypuśćmy najpierw, że $\vec{v} = \lambda \circ \vec{u}$ dla pewnego $\lambda \in \mathbb{R} \setminus \{0\}$. Wtedy:

$$\cos\varphi = \begin{cases} 1 & \text{, gdy } \lambda > 0 \\ -1 & \text{, gdy } \lambda < 0 \end{cases},$$

skąd $\lambda = |\lambda| \cdot \cos\varphi$. Na mocy wzoru (11.2.1) oraz punktu (iii) Stwierdzenia 11.3 uzyskujemy więc, że $\langle\vec{v}|\vec{u}\rangle = \|\lambda \circ \vec{u}\| \cdot \|\vec{u}\| \cdot \cos\varphi = |\lambda|\|\vec{u}\|^2 \cdot \cos\varphi = \lambda \cdot (u_1^2 + u_2^2 + u_3^2) = (\lambda u_1)u_1 + (\lambda u_2)u_2 + (\lambda u_3)u_3 = v_1u_1 + v_2u_2 + v_3u_3$. Niech dalej $\vec{v} \neq \lambda \circ \vec{u}$ dla każdego $\lambda \in \mathbb{R} \setminus \{0\}$. Wówczas $\varphi \in (0,\pi)$, więc istnieje trójkąt rozpięty na wektorach $\vec{v}$ i $\vec{u}$. Trzeci bok tego trójkąta ma długość $\|\vec{u} - \vec{v}\|$. Niech $\vec{w} = \vec{u} - \vec{v}$. Na mocy znanego ze szkoły średniej Twierdzenia cosinusów zastosowanego dla trójkąta o bokach $\vec{v}, \vec{u}$ i $\vec{w}$ otrzymujemy, że $\|\vec{w}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2 \cdot \|\vec{u}\| \cdot \|\vec{v}\| \cdot \cos\varphi$. Ponadto $\|\vec{w}\|^2 = \|\vec{v} - \vec{u}\|^2 = (v_1 - u_1)^2 + (v_2 - u_2)^2 + (v_3 - u_3)^2 = (v_1^2 + v_2^2 + v_3^2) + (u_1^2 + u_2^2 + u_3^2) - 2(v_1u_1 + v_2u_2 + v_3u_3) = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2(v_1u_1 + v_2u_2 + v_3u_3)$. Zatem $\|\vec{u}\| \cdot \|\vec{v}\| \cdot \cos\varphi = v_1u_1 + v_2u_2 + v_3u_3$, co wobec (11.2.1) oznacza żądaną równość.

Uwaga 11.1. Wzór (11.2.2) uogólnia się na przypadek przestrzeni $\mathbb{R}^n$, gdzie n jest dowolną liczbą naturalną. Określa on wówczas tzw. standardowy iloczyn skalarny w przestrzeni $\mathbb{R}^n$.

Bezpośrednią konsekwencją wzorów (11.2.1) i (11.2.2) oraz określenia kąta φ jest następujący

Wniosek 11.1. Dla dowolnych wektorów $\vec{v} = [v_1, v_2, v_3]$ i $\vec{u} = [u_1, u_2, u_3]$ przestrzeni $\mathbb{R}^3$ zachodzi wzór:

$$\left|\measuredangle(\vec{u},\vec{v})\right| = \arccos \frac{v_1u_1 + v_2u_2 + v_3u_3}{\sqrt{v_1^2 + v_2^2 + v_3^2} \cdot \sqrt{u_1^2 + u_2^2 + u_3^2}}, \tag{11.2.3}$$

gdzie $\left|\measuredangle(\vec{u},\vec{v})\right|$ oznacza miarę kąta $\measuredangle(\vec{u},\vec{v})$.

Poniższe stwierdzenie zawiera zestawienie ważnych własności iloczynu skalarnego wektorów.

Stwierdzenie 11.3. Dla dowolnych wektorów $\vec{u}$, $\vec{v}$, $\vec{w}$ przestrzeni $\mathbb{R}^3$ oraz dowolnej liczby rzeczywistej λ:

(i) $\langle\vec{u}|\vec{v}\rangle = \langle\vec{v}|\vec{u}\rangle$;
(ii) $\langle\vec{v}|\vec{v}\rangle = \|\vec{v}\|^2$;
(iii) $\langle\lambda\circ\vec{v}|\vec{u}\rangle = \langle\vec{v}|\lambda\circ\vec{u}\rangle = \lambda\cdot\langle\vec{u}|\vec{v}\rangle$;
(iv) $\langle\vec{v}+\vec{w}|\vec{u}\rangle = \langle\vec{v}|\vec{u}\rangle + \langle\vec{w}|\vec{u}\rangle$;
(v) $\langle\vec{v}|\vec{u}+\vec{w}\rangle = \langle\vec{v}|\vec{u}\rangle + \langle\vec{v}|\vec{w}\rangle$;
(vi) $|\langle\vec{v}|\vec{u}\rangle| \leq \|\vec{v}\|\cdot\|\vec{u}\|$;
(vii) $\vec{v}\perp\vec{u}$ wtedy i tylko wtedy, gdy $\langle\vec{v}|\vec{u}\rangle = 0$, przy czym zapis $\vec{v}\perp\vec{u}$ oznacza, że wektor $\vec{v}$ jest prostopadły do wektora $\vec{u}$.

Dowód. Równości (i), (iii) oraz (iv) są bezpośrednią konsekwencją wzoru (11.2.2). Ten sam wzór wraz z (11.1.1) implikuje także równość (ii). Nierówność (vi) wynika wprost z Twierdzenia 11.1 zastosowanego do wzorów (11.1.1) i (11.2.2). Równoważność dana w (vii) jest natychmiastową konsekwencją wzoru (11.2.1) oraz faktu, że jedynym kątem φ należącym do przedziału $[0,\pi]$ takim, że $\cos\varphi = 0$ jest $\varphi = \frac{\pi}{2}$.

Uwaga 11.2. Własność (vi) nazywa się warunkiem ortogonalności (czyli prostopadłości) wektorów $\vec{u}$ i $\vec{v}$.

11.3 Iloczyn wektorowy

Definicja 11.3. Iloczynem wektorowym wektorów $\vec{v} = [v_1, v_2, v_3]$ i $\vec{u} = [u_1, u_2, u_3]$ przestrzeni $\mathbb{R}^3$ nazywamy wektor $\vec{v}\times\vec{u}$ określony wzorem:

$$\vec{v}\times\vec{u} = [v_2u_3 - v_3u_2, v_3u_1 - v_1u_3, v_1u_2 - v_2u_1]. \tag{11.3.1}$$

Bezpośrednie sprawdzenie pokazuje, że wzór (11.3.1) można wyrazić przy użyciu symbolicznego wyznacznika:

$$\vec{v}\times\vec{u} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ v_1 & v_2 & v_3 \\ u_1 & u_2 & u_3 \end{vmatrix}, \tag{11.3.2}$$

gdzie $\vec{i} = [1,0,0]$, $\vec{j} = [0,1,0]$ oraz $\vec{k} = [0,0,1]$.

Definicja 11.4. Wektory $\vec{i} = [1,0,0]$, $\vec{j} = [0,1,0]$ oraz $\vec{k} = [0,0,1]$ nazywamy wersorami.

Uwaga 11.3. W przeciwieństwie do iloczynu skalarnego, iloczyn wektorowy definiuje się wyłącznie w przestrzeni $\mathbb{R}^3$.

Stwierdzenie 11.4. Dla dowolnych wektorów $\vec{v}$ i $\vec{u}$ przestrzeni $\mathbb{R}^3$ zachodzi wzór:

$$\|\vec{v}\times\vec{u}\| = \|\vec{v}\|\cdot\|\vec{u}\|\cdot\sin\measuredangle(\vec{v},\vec{u}). \tag{11.3.3}$$

Dowód. Rozważmy dowolne wektory $\vec{v}=[v_1,v_2,v_3]$ oraz $\vec{u}=[u_1,u_2,u_3]$ i oznaczmy $\varphi = \measuredangle(\vec{v},\vec{u})$. Na mocy (11.3.1), (11.2.2) i (11.2.1) otrzymujemy wówczas, że $\|\vec{v}\times\vec{u}\|^2 = (v_2u_3-v_3u_2)^2+(v_3u_1-v_1u_3)^2+(v_1u_2-v_2u_1)^2 = v_2^2u_3^2+v_3^2u_2^2+v_3^2u_1^2+v_1^2u_3^2+v_1^2u_2^2+v_2^2u_1^2-2(v_2u_3v_3u_2+v_3u_1v_1u_3+v_1u_2v_2u_1) = \big((v_1^2+v_2^2+v_3^2)\cdot(u_1^2+u_2^2+u_3^2)-(v_1^2u_1^2+v_2^2u_2^2+v_3^2u_3^2)\big)-2(v_2u_3v_3u_2+v_3u_1v_1u_3+v_1u_2v_2u_1) = (v_1^2+v_2^2+v_3^2)\cdot(u_1^2+u_2^2+u_3^2)-\big((v_1^2u_1^2+2v_1u_1v_2u_2+v_2^2u_2^2)+2(v_1u_1+v_2u_2)(v_3u_3)+(v_3u_3)^2\big) = \|\vec{v}\|^2\cdot\|\vec{u}\|^2-(v_1u_1+v_2u_2+v_3u_3)^2 = \|\vec{v}\|^2\cdot\|\vec{u}\|^2-\langle\vec{v}|\vec{u}\rangle^2 = \|\vec{v}\|^2\cdot\|\vec{u}\|^2-\|\vec{u}\|^2\cdot\|\vec{v}\|^2\cdot\cos^2\varphi = \|\vec{v}\|^2\cdot\|\vec{u}\|^2\cdot\big(1-\cos^2\varphi\big) = \|\vec{v}\|^2\cdot\|\vec{u}\|^2\cdot\sin^2\varphi$. Ponadto $\sin\varphi\geq 0$, bo $\varphi\in[0,\pi]$. Wobec tego $\|\vec{v}\times\vec{u}\| = \|\vec{v}\|\cdot\|\vec{u}\|\cdot\sin\varphi$.

Podstawowe własności iloczynu wektorowego oraz jego związki z iloczynem skalarnym opisane są w poniższym stwierdzeniu.

Stwierdzenie 11.5. Dla dowolnych wektorów $\vec{u}$, $\vec{v}$, $\vec{w}$ przestrzeni $\mathbb{R}^3$ oraz dowolnej liczby rzeczywistej λ:

(i) $\vec{u}\times\vec{v} = -(\vec{v}\times\vec{u})$;
(ii) $(\lambda\circ\vec{u})\times\vec{v} = \vec{u}\times(\lambda\circ\vec{v}) = \lambda\circ(\vec{u}\times\vec{v})$;
(iii) $(\vec{u}+\vec{v})\times\vec{w} = \vec{u}\times\vec{w}+\vec{v}\times\vec{w}$;
(iv) $\vec{u}\times(\vec{v}+\vec{w}) = \vec{u}\times\vec{v}+\vec{u}\times\vec{w}$;
(v) jeżeli $\vec{u}=\vec{0}$ lub $\vec{v}=\vec{0}$, to $\vec{u}\times\vec{v}=\vec{0}$;
(vi) jeżeli $\vec{u}\neq\vec{0}$ i $\vec{v}\neq\vec{0}$, to $\vec{u}\parallel\vec{v}$ wtedy i tylko wtedy, gdy $\vec{u}\times\vec{v}=\vec{0}$, przy czym zapis $\vec{u}\parallel\vec{v}$ oznacza, że wektory $\vec{u}$ i $\vec{v}$ są równoległe;
(vii) $(\vec{v}\times\vec{u})\perp\vec{v}$ oraz $(\vec{v}\times\vec{u})\perp\vec{u}$;
(viii) $\vec{u}\times(\vec{v}\times\vec{w}) = \langle\vec{u}|\vec{w}\rangle\circ\vec{v}-\langle\vec{u}|\vec{v}\rangle\circ\vec{w}$;
(ix) $\langle\vec{u}|\vec{v}\times\vec{w}\rangle = \langle\vec{u}\times\vec{v}|\vec{w}\rangle$;
(x) $\|\vec{v}\times\vec{u}\|\leq\|\vec{v}\|\cdot\|\vec{u}\|$.

Dowód. Własności (i)-(v) wynikają natychmiast ze wzoru (11.3.2) i własności wyznacznika opisanych we Wniosku 12 oraz Stwierdzeniach 4.6 i 4.8 (można je też oczywiście udowodnić w oparciu o bezpośrednie rachunki związane ze wzorem (11.3.1)). Rozważmy dowolne $\vec{v},\vec{u},\vec{w}\in\mathbb{R}^3$ i $\lambda\in\mathbb{R}$.

Aby udowodnić własność (vi) przypuśćmy, że $\vec{u}\neq\vec{0}$ i $\vec{v}\neq\vec{0}$. Jeżeli $\vec{u}\parallel\vec{v}$, to równoległe są także odpowiadające im wektory zaczepione w punkcie O. Istnieje więc $\lambda\in\mathbb{R}^*$ taka, że $\vec{u}=\lambda\circ\vec{v}$. Stąd oraz na mocy wzoru (11.3.1), $\vec{u}\times\vec{v}=\vec{0}$. Jeśli natomiast $\vec{u}\times\vec{v}=\vec{0}$, to $\|\vec{u}\times\vec{v}\|=0$, więc ze wzoru (11.3.3) i określenia kąta φ wynika, że $\varphi=0$ lub $\varphi=\Pi$. Zatem $\vec{u}\parallel\vec{v}$.

Z punktu (vi) Stwierdzenia 11.3 wynika, że aby uzasadnić punkt (vii) niniejszego stwierdzenia wystarczy uzasadnić, że $\langle \vec{v} \times \vec{u} | \vec{v} \rangle = 0$ i $\langle \vec{v} \times \vec{u} | \vec{u} \rangle = 0$ dla $\vec{v} = [v_1, v_2, v_3]$ oraz $\vec{u} = [u_1, u_2, u_3]$. Mamy: $\langle \vec{v} \times \vec{u} | \vec{v} \rangle = \big\langle [v_2u_3 - v_3u_2, v_3u_1 - v_1u_3, v_1u_2 - v_2u_1] \big| [v_1, v_2, v_3] \big\rangle = (v_2u_3v_1 - v_3u_2v_1) + (v_3u_1v_2 - v_1u_3v_2) + (v_1u_2v_3 - v_2u_1v_3) = (v_1v_2u_3 - v_1v_2u_3) + (v_2v_3u_1 - v_2v_3u_1) + (v_1v_3u_2 - v_1v_3u_2) = 0$. Ponadto $\langle \vec{v} \times \vec{u} | \vec{u} \rangle = \langle -(\vec{u} \times \vec{v}) | \vec{u} \rangle = -\langle \vec{u} \times \vec{v} | \vec{u} \rangle = 0$ odpowiednio na mocy punktu (i) niniejszego stwierdzenia, przeprowadzonych wyżej rachunków oraz punktu (iii) Stwierdzenia 11.3.

Niech dodatkowo $\vec{w} = [w_1, w_2, w_3]$. Na mocy wzoru (11.3.1) otrzymujemy wówczas, że $\vec{u} \times (\vec{v} \times \vec{w}) = [u_1, u_2, u_3] \times [v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1] = \big(u_2(v_1 w_2 - v_2w_1) - u_3(v_3w_1 - v_1w_3)\big) \circ \vec{i} + \big(u_3(v_2w_3 - v_3w_2) - u_1(v_1w_2 - v_2w_1)\big) \circ \vec{j} + \big(u_1(v_3w_1 - v_1w_3) - u_2(v_2w_3 - v_3w_2)\big) \circ \vec{k}$. Ponadto ze wzoru (11.2.2) wynika, że $\langle \vec{u} | \vec{w} \rangle \circ \vec{v} - \langle \vec{u} | \vec{v} \rangle \circ \vec{w} = (u_1w_1 + u_2w_2 + u_3w_3) \circ \vec{v} - (u_1v_1 + u_2v_2 + u_3v_3) \circ \vec{w} = \big(u_2(w_2v_1 - w_1v_2) - u_3(w_1v_3 - w_3v_1)\big) \circ \vec{i} + \big(u_3(w_3v_2 - w_2v_3) - u_1(w_2v_1 - w_1v_2)\big) \circ \vec{j} + \big(u_1(w_1v_3 - w_3v_1) - u_2(w_3v_2 - w_2v_3)\big) \circ \vec{k}$, więc zachodzi równość dana w $(viii)$.

Powołując się ponownie na wzory (11.3.1) i (11.2.2) uzyskujemy, że $\langle \vec{u} | \vec{v} \times \vec{w} \rangle = \big\langle [u_1, u_2, u_3] \big| [v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1] \big\rangle = u_1(v_2w_3 - v_3w_2) + u_2(v_3w_1 - v_1w_3) + u_3(v_1w_2 - v_2w_1) = (u_2v_3 - u_3v_2)w_1 + (u_3v_1 - u_1v_3)w_2 + (u_1v_2 - u_2v_1)w_3 = \big\langle [u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1] \big| [w_1, w_2, w_3] \big\rangle = \langle \vec{u} \times \vec{v} | \vec{w} \rangle$, co uzasadnia własność (ix).

Nierówność dana w (x) jest bezpośrednią konsekwencją wzoru (11.3.3) oraz nierówności $0 \leq \sin \measuredangle(\vec{v}, \vec{u}) \leq 1$.

Iloczyn wektorowy $\vec{u} \times \vec{v}$ dwóch niezerowych i niewspółliniowych wektorów $\vec{u}$ i $\vec{v}$ przestrzeni $\mathbb{R}^3$ ma taką własność, że orientacja trójki wektorów $\vec{u}$, $\vec{v}$ i $\vec{u} \times \vec{v}$ jest zgodna z orientacją układu współrzędnych O_{xyz}.

Przypomnijmy, że rozróżniamy dwie orientacje układu współrzędnych O_{xyz}: dodatnią i ujemną. Układ O_{xyz} o orientacji dodatniej nazywamy prawoskrętnym, zaś układ O_{xyz} o orientacji ujemnej nazywamy lewoskrętnym. Przypomnijmy, że orientacja układu współrzędnych O_{xyz} zależy od wzajemnego położenia osi O_x, O_y i O_z tego układu. Jeżeli wyprostowany kciuk prawej dłoni umieścimy w ten sposób, aby wskazywał dodatnią część osi 0_z, a zgięte pozostałe palce wskażą kierunek obrotu od osi O_x do osi O_y, to mamy wówczas do czynienia z układem prawoskrętnym. Jeśli zaś wyprostowany kciuk prawej dłoni wskazuje dodatnią część osi O_z, a zgięte pozostałe palce wskażą kierunek obrotu od osi O_y do osi O_x, to mamy do czynienia z układem lewoskrętnym.

11.4 Iloczyn mieszany

Definicja 11.5. Iloczynem mieszanym uporządkowanej trójki wektorów $\vec{u}$, $\vec{v}$, $\vec{w}$ przestrzeni $\mathbb{R}^3$ nazywamy liczbę $(\vec{u}, \vec{v}; \vec{w})$ określoną wzorem:

$$(\vec{u},\vec{v};\vec{w}) = \langle \vec{u}\times\vec{v}|\vec{w}\rangle. \tag{11.4.1}$$

Bezpośrednie sprawdzenie pokazuje, że iloczyn mieszany uporządkowanej trójki dowolnych wektorów $\vec{u}=[u_1,u_2,u_3]$, $\vec{v}=[v_1,v_2,v_3]$ i $\vec{w}=[w_1,w_2,w_3]$ przestrzeni $\mathbb{R}^3$ wyraża się wzorem:

$$(\vec{u},\vec{v};\vec{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}. \tag{11.4.2}$$

Własności iloczynu mieszanego wymienione są w poniższym stwierdzeniu.

Stwierdzenie 11.6. Dla dowolnych wektorów $\vec{a}$, $\vec{u}$, $\vec{v}$, $\vec{w}$ przestrzeni $\mathbb{R}^3$ oraz dowolnej liczby rzeczywistej λ:

(i) $(\vec{u},\vec{v};\vec{w}) = -(\vec{v},\vec{u};\vec{w}) = (\vec{v},\vec{w};\vec{u})$;
(ii) $\lambda\cdot(\vec{u},\vec{v};\vec{w}) = (\lambda\circ\vec{u},\vec{v};\vec{w}) = (\vec{u},\lambda\circ\vec{v};\vec{w}) = (\vec{u},\vec{v};\lambda\circ\vec{w})$;
(iii) $(\vec{a}+\vec{u},\vec{v};\vec{w}) = (\vec{a},\vec{v};\vec{w}) + (\vec{u},\vec{v};\vec{w})$;
(iv) $(\vec{u},\vec{a}+\vec{v};\vec{w}) = (\vec{u},\vec{a};\vec{w}) + (\vec{u},\vec{v};\vec{w})$;
(v) $(\vec{u},\vec{v};\vec{a}+\vec{w}) = (\vec{u},\vec{v};\vec{a}) + (\vec{u},\vec{v};\vec{w})$;
(vi) $\left|(\vec{u},\vec{v};\vec{w})\right| \le \|\vec{u}\|\cdot\|\vec{v}\|\cdot\|\vec{w}\|$.

Dowód. Własności (i)-(v) wynikają natychmiast ze wzoru $(11.4.2)$ oraz własności wyznaczników opisanych w Stwierdzeniach 4.6, 4.5 i 4.8. Ponadto $\left|(\vec{u},\vec{v};\vec{w})\right| = \left|\langle\vec{u}\times\vec{v}|\vec{w}\rangle\right| \le \|\vec{u}\times\vec{v}\|\cdot\|\vec{w}\| \le \|\vec{u}\|\cdot\|\vec{v}\|\cdot\|\vec{w}\|$ odpowiednio na mocy punktu (vi) Stwierdzenia 11.3 oraz punktu (x) Stwierdzenia 11.5.

Stwierdzenie 11.7. Trzy niezerowe wektory przestrzeni $\mathbb{R}^3$ są współpłaszczyznowe wtedy i tylko wtedy, gdy ich iloczyn mieszany wynosi zero.

Dowód. Niech U oznacza podprzestrzeń przestrzeni liniowej $\mathbb{R}^3$ rozpiętą na dowolnie ustalonych niezerowych wektorach $\vec{u}=[u_1,u_2,u_3]$, $\vec{v}=[v_1,v_2,v_3]$ i $\vec{w}=[w_1,w_2,w_3]$ przestrzeni $\mathbb{R}^3$. Wówczas wektory $\vec{u}$, $\vec{v}$ i $\vec{w}$ nie leżą w jednej płaszczyźnie wtedy i tylko wtedy, gdy $\dim_{\mathbb{R}} U = 3$ czyli, gdy wyznacznik dany w $(11.4.2)$ jest niezerowy, co oznacza, że $(\vec{u},\vec{v};\vec{w}) \neq 0$ (z punktu (i) Stwierdzenia 11.6 wynika, że sposób uporządkowania wektorów $\vec{u}$, $\vec{v}$ i $\vec{w}$ jest nieistotny w kontekście naszych rozważań związanych z tym stwierdzeniem).

11.4.1 Zastosowania geometryczne iloczynu wektorowego oraz iloczynu mieszanego wektorów

Ze wzoru $(11.3.3)$ oraz poznanych w szkole średniej wiadomości z zakresu obliczania pól czworokątów wynika natychmiast poniższe:

Stwierdzenie 11.8. Pole P równoległoboku rozpiętego przez wektory $\vec{v}$ i $\vec{u}$ przestrzeni $\mathbb{R}^3$ wyraża się wzorem:

$$P = \|\vec{v} \times \vec{u}\|. \tag{11.4.3}$$

Bezpośrednią konsekwencją powyższego stwierdzenia jest następujące

Stwierdzenie 11.9. Pole P trójkąta w przestrzeni $\mathbb{R}^3$ rozpiętego przez wektory $\vec{v}$ i $\vec{u}$ wyraża się wzorem:

$$P = \frac{1}{2} \cdot \|\vec{v} \times \vec{u}\|. \tag{11.4.4}$$

Przykład 11.1. Wyznaczymy długość h wysokości trójkąta o wierzchołkach w punktach $A = (2,2,2)$, $B = (3,3,3)$ i $C = (4,5,6)$ opuszczonej z wierzchołka C.

Niech P oznacza pole trójkąta o wierzchołkach A, B i C. Wówczas $P = \frac{1}{2} \cdot \|\overrightarrow{AB}\| \cdot h$ oraz, na mocy wzoru (11.4.4), $P = \frac{1}{2} \cdot \|\overrightarrow{AB} \times \overrightarrow{AC}\|$, więc:

$$h = \frac{\|\overrightarrow{AB} \times \overrightarrow{AC}\|}{\|\overrightarrow{AB}\|}. \tag{11.4.5}$$

Dalej, $\overrightarrow{AB} = [1,1,1]$ i $\overrightarrow{AC} = [2,3,4]$, więc:

$$\overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{vmatrix} = (4-3) \circ \vec{i} + (2-4) \circ \vec{j} + (3-2) \circ \vec{k} = [1,-2,1],$$

skąd:

$$\|\overrightarrow{AB} \times \overrightarrow{AC}\| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6} \tag{11.4.6}$$

oraz

$$\|\overrightarrow{AB}\| = \sqrt{3}. \tag{11.4.7}$$

Podstawiając (11.4.6) i (11.4.7) do (11.4.5) uzyskujemy, że $h = \frac{\sqrt{6}}{\sqrt{2}} = \sqrt{2}$.

Definicja 11.6. Równoległościanem nazywamy wielościan o trzech parach boków równoległych.

Stwierdzenie 11.10. Objętość V równoległościanu rozpiętego na wektorach $\vec{u}$, $\vec{v}$ i $\vec{w}$ wyraża się wzorem:

$$V = \big|(\vec{u}, \vec{v}; \vec{w})\big|. \tag{11.4.8}$$

Dowód. Bez utraty ogólności możemy przyjąć, że podstawa równoległościanu jest rozpięta na wektorach $\vec{u}$ i $\vec{v}$ (por. punkt (i) Stwierdzenia 11.6). Niech P oznacza jej pole. Ze Stwierdzenia 11.8 wynika wówczas, że $P = \|\vec{u} \times \vec{v}\|$. Niech h oznacza

długość wysokości rozważanego równoległościanu opuszczonej z wierzchołka znajdującego się najbliżej wspólnego punktu zaczepienia wektorów $\vec{u},\vec{v}$ oraz $\vec{w}$ i niech α będzie kątem nachylenia wektora $\vec{w}$ do płaszczyzny podstawy. Wtedy $h = \|\vec{w}\| \cdot \sin\alpha$. Niech $\beta = \measuredangle(\vec{u}\times\vec{v},\vec{w})$. Jeżeli $\beta \in \left[0,\frac{\pi}{2}\right)$, to $\alpha = \frac{\pi}{2} - \beta$, skąd $h = \|\vec{w}\| \cdot \sin\left(\frac{\pi}{2} - \beta\right) = \|\vec{w}\| \cdot \cos\beta$. Zatem $V = P \cdot h = \|\vec{v}\times\vec{u}\| \cdot \|\vec{w}\| \cdot \cos\beta = \langle \vec{u}\times\vec{v}|\vec{w}\rangle = (\vec{u},\vec{v};\vec{w})$. Jeśli natomiast $\beta \in \left(\frac{\pi}{2},\pi\right]$, to $\alpha = \frac{\pi}{2} - (\pi - \beta)$, więc $h = \|\vec{w}\| \cdot \sin\left(\frac{\pi}{2} - (\pi-\beta)\right) = \|\vec{w}\| \cdot \cos(\pi-\beta) = \|\vec{w}\| \cdot (\cos\pi\cdot\cos\beta + \sin\pi\cdot\sin\beta) = -\|\vec{w}\| \cdot \cos\beta$. Wobec tego, w tym przypadku, $V = -(\vec{u},\vec{v};\vec{w})$. Ostatecznie otrzymujemy więc, że $V = \left|(\vec{u},\vec{v};\vec{w})\right|$.

Ze Stwierdzeń 11.9 i 11.10 oraz z poznanego w szkole średniej wzoru na objętość ostrosłupa wynika następujące

Stwierdzenie 11.11. Objętość V czworościanu rozpiętego na wektorach $\vec{u}$, $\vec{v}$ i $\vec{w}$ wyraża się wzorem:

$$V = \frac{1}{6} \cdot \left|(\vec{u},\vec{v};\vec{w})\right|. \tag{11.4.9}$$

Przykład 11.2. Obliczymy długość h wysokości czworościanu o wierzchołkach $A = (3,0,0)$, $B = (0,4,0)$, $C = (0,0,5)$ i $D = (2,3,4)$ opuszczonej z wierzchołka D.

Niech V oznacza objętość rozważanego czworościanu. Ze wzoru (11.4.9) wynika wówczas, że $V = \frac{1}{6} \cdot \left|\left(\overrightarrow{DA},\overrightarrow{DB},\overrightarrow{DC}\right)\right|$. Ponadto $V = \frac{1}{3} \cdot P_{\triangle ABC} \cdot h$, gdzie $P_{\triangle ABC}$ oznacza pole trójkąta o wierzchołkach A, B i C, oraz ze Stwierdzenia 11.9 wynika, że $P_{\triangle ABC} = \frac{1}{2} \cdot \|\overrightarrow{AB}\times\overrightarrow{AC}\|$. Zatem:

$$h = \frac{\left|\left(\overrightarrow{DA},\overrightarrow{DB},\overrightarrow{DC}\right)\right|}{\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\|}. \tag{11.4.10}$$

Dalej $\left(\overrightarrow{DA},\overrightarrow{DB},\overrightarrow{DC}\right) = \begin{vmatrix} -1 & 3 & 4 \\ 2 & -1 & 4 \\ 2 & 3 & -1 \end{vmatrix} = -1+24+24+8+6+12 = 73$ na mocy wzoru (11.4.2), skąd:

$$\left|\left(\overrightarrow{DA},\overrightarrow{DB},\overrightarrow{DC}\right)\right| = 73. \tag{11.4.11}$$

Ponadto $\overrightarrow{AB}\times\overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -3 & 4 & 0 \\ -3 & 0 & 5 \end{vmatrix} = (20-0)\circ\vec{i} + (0+15)\circ\vec{j} + (0+12)\circ\vec{k} = [20,15,12]$, więc:

$$\left\|\overrightarrow{AB}\times\overrightarrow{AC}\right\| = \sqrt{20^2+15^2+12^2} = \sqrt{769}. \tag{11.4.12}$$

Podstawiając (11.4.11) oraz (11.4.12) do (11.4.10) otrzymujemy, że $h = \frac{73\sqrt{769}}{769}$.

11.5 Równania płaszczyzny w przestrzeni $\mathbb{R}^3$

11.5.1 Ogólne równanie płaszczyzny

Rozważmy dowolną płaszczyznę Π w przestrzeni $\mathbb{R}^3$. Niech $P_0 = (x_0, y_0, z_0)$ będzie jej dowolnym punktem, zaś $\vec{\eta} = [A, B, C]$ – dowolnym niezerowym wektorem przestrzeni $\mathbb{R}^3$ prostopadłym do Π. Zauważmy, że płaszczyznę Π możemy traktować jako zbiór punktów $P = (x, y, z)$ takich, że $\overrightarrow{P_0P} \perp \vec{\eta}$. Z punktu (vii) Stwierdzenia 11.3 wynika wówczas, że:

$$\Pi = \Big\{ [x, y, z] \in \mathbb{R}^3 : \big\langle [x - x_0, y - y_0, z - z_0] \big| [A, B, C] \big\rangle = 0 \Big\}.$$

Ponadto $\big\langle [x - x_0, y - y_0, z - z_0] \big| [A, B, C] \big\rangle = A(x - x_0) + B(y - y_0) + C(z - z_0)$, więc dla $D = -(Ax_0 + By_0 + Cz_0)$ otrzymujemy, że płaszczyzna Π opisana jest równaniem:

$$Ax + By + Cz + D = 0, \tag{11.5.1}$$

gdzie A, B, C i D są pewnymi liczbami rzeczywistymi takimi, że $A^2 + B^2 + C^2 > 0$.

Na odwrót. Rozważmy teraz dowolne liczby rzeczywiste A, B, C i D takie, że $A^2 + B^2 + C^2 > 0$, oraz podzbiór Π przestrzeni liniowej $\mathbb{R}^3$ określony warunkiem danym w (11.5.1). Wówczas $\vec{\eta} = [A, B, C]$ jest niezerowym wektorem. Weźmy dowolne $P_0 = (x_0, y_0, z_0) \in \Pi$. Wtedy $D = -(Ax_0 + By_0 + Cz_0)$, więc $Ax + By + Cz + D = \big\langle [x - x_0, y - y_0, z - z_0] \big| [A, B, C] \big\rangle$, co oznacza, że dowolny punkt $P = (x, y, z) \in \Pi$ spełnia warunek $\overrightarrow{P_0P} \perp \vec{\eta}$. Wobec tego Π jest płaszczyzną w $\mathbb{R}^3$.

Powyższe rozważania stanowią przesłankę do sformułowania następującej definicji.

Definicja 11.7. Równanie (11.5.1), gdzie $A, B, C, D \in \mathbb{R}$ i $A^2 + B^2 + C^2 > 0$ nazywamy ogólnym równaniem płaszczyzny w przestrzeni $\mathbb{R}^3$ w kartezjańskim układzie współrzędnych.

Poniższy wniosek jest bezpośrednią konsekwencją rozważań przeprowadzonych w związku z omówieniem ogólnego równania płaszczyzny:

Wniosek 11.2. Niech $\vec{\eta} = [A, B, C]$ będzie niezerowym wektorem przestrzeni $\mathbb{R}^3$ i niech Π będzie płaszczyzną opisaną równaniem $Ax + By + Cz + D = 0$. Wówczas wektor $\vec{\eta}$ jest prostopadły do płaszczyzny Π.

Definicja 11.8. Wektor prostopadły do płaszczyzny nazywamy wektorem normalnym płaszczyzny.

Powyższa definicja wraz z Wnioskiem 11.8 implikuje natychmiast:

Wniosek 11.3. Wektor $[A, B, C]$ jest wektorem normalnym płaszczyzny opisanej równaniem (11.5.1).

Uwaga 11.4. Niech Π będzie płaszczyzną w przestrzeni $\mathbb{R}^3$ opisaną za pomocą wzoru (11.5.1). Wówczas:

(i) początek O układu współrzędnych O_{xyz} jest punktem płaszczyzny Π wtedy i tylko wtedy, gdy $D = 0$;
(ii) płaszczyzna Π jest równoległa do osi O_x wtedy i tylko wtedy, gdy $A = 0$;
(iii) płaszczyzna Π jest równoległa do osi O_y wtedy i tylko wtedy, gdy $B = 0$;
(iv) płaszczyzna Π jest równoległa do osi O_z wtedy i tylko wtedy, gdy $C = 0$;
(v) płaszczyzna Π jest prostopadła do osi O_z wtedy i tylko wtedy, gdy $A = 0$ i $B = 0$;
(vi) płaszczyzna Π jest prostopadła do osi O_x wtedy i tylko wtedy, gdy $B = 0$ i $C = 0$;
(vii) płaszczyzna Π jest prostopadła do osi O_y wtedy i tylko wtedy, gdy $A = 0$ i $C = 0$;
(viii) płaszczyzna Π zawiera oś O_x wtedy i tylko wtedy, gdy $A = 0$ i $D = 0$;
(ix) płaszczyzna Π zawiera oś O_y wtedy i tylko wtedy, gdy $B = 0$ i $D = 0$;
(x) płaszczyzna Π zawiera oś O_z wtedy i tylko wtedy, gdy $C = 0$ i $D = 0$.

11.5.2 Równanie płaszczyzny przechodzącej przez punkt i prostopadłej do wektora

Z rozważań dotyczących ogólnego równania płaszczyzny wynika natychmiast, że równanie:

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0, \tag{11.5.2}$$

gdzie A, B i C są ustalonymi liczbami rzeczywistymi takimi, że $A^2+B^2+C^2>0$ opisuje płaszczyznę przechodzącą przez punkt $P_0=(x_0,y_0,z_0)$ i prostopadłą do niezerowego wektora $\vec{\eta}=[A,B,C]$.

Stąd oraz na mocy Wniosku 11.2 otrzymujemy natychmiast następujący

Wniosek 11.4. Przy dowolnie ustalonych liczbach rzeczywistych A, B, C, D, x_0, y_0 i z_0 takich, że $A^2+B^2+C^2>0$ równanie (11.5.2) opisuje płaszczyznę przechodząca przez punkt $P=(x_0,y_0,z_0)$ i równoległą do płaszczyzny opisanej równaniem (11.5.1).

Uwaga 11.5. Jeżeli w równaniu (11.5.2) rozważymy wszystkie liczby rzeczywiste A, B i C spełniające warunek $A^2+B^2+C^2>0$, to otrzymamy równanie opisujące pęk płaszczyzn przechodzących przez punkt $P=(x_0,y_0,z_0)\in\mathbb{R}^3$.

11.5.3 Równanie płaszczyzny przechodzącej przez trzy niewspółliniowe punkty

Stwierdzenie 11.12. Niech $P_1=(x_1,y_1,z_1)$, $P_1=(x_2,y_2,z_2)$ oraz $P_1=(x_3,y_3,z_3)$ będą niewspółliniowymi punktami przestrzeni $\mathbb{R}^3$. Wówczas:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0 \tag{11.5.3}$$

jest równaniem płaszczyzny przechodzącej przez punkty P_1, P_2 i P_3.

Dowód. Niech Π będzie płaszczyzną przechodzącą przez punkty P_1, P_2 oraz P_3. Oznaczmy $\vec{v} = \overrightarrow{P_1P_2}$, $\vec{u} = \overrightarrow{P_1P_3}$ oraz $\vec{\eta} = \vec{v} \times \vec{u}$. Wówczas Π jest płaszczyzną przechodzącą przez punkt P_2 i prostopadłą do wektora $\vec{\eta}$. Ponadto:

$$\vec{\eta} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = [A, B, C],$$

gdzie:

$$A = (y_2 - y_1)(z_3 - z_1) - (y_3 - y_1)(z_2 - z_1), \tag{11.5.4}$$

$$B = (z_2 - z_1)(x_3 - x_1) - (x_2 - x_1)(z_3 - z_1), \tag{11.5.5}$$

$$C = (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1). \tag{11.5.6}$$

Zatem płaszczyzna Π opisana jest równaniem:

$$A(x - x_2) + B(y - y_2) + C(z - z_2) = 0. \tag{11.5.7}$$

Dalej, niech W oznacza wyznacznik występujący po lewej stronie równania (11.5.3). Po wykonaniu kolejno operacji elementarnych $w_1 - w_3$, $w_2 - w_3$ i $w_4 - w_3$ związanych z obliczaniem tego wyznacznika, zastosowaniu rozwinięcia Laplace'a względem ostatniej kolumny macierzy z nim związanej, a następnie wykonaniu operacji elementarnej $(-1) \cdot w_2$ i zastosowaniu wzorów (11.5.4)-(11.5.6) otrzymujemy, że:

$$W = \begin{vmatrix} x - x_2 & y - y_2 & z - z_2 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = A(x - x_2) + B(y - y_2) + C(z - z_2). \tag{11.5.8}$$

Stąd oraz na mocy wzoru (11.5.7) otrzymujemy, że równanie (11.5.3) jest równaniem płaszczyzny przechodzącym przez punkty P_1, P_2 i P_3.

Stwierdzenie 11.13. Cztery punkty $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, $P_3 = (x_3, y_3, z_3)$ i $P_4 = (x_4, y_4, z_4)$ przestrzeni $\mathbb{R}^3$ są współpłaszczyznowe wtedy i tylko wtedy, gdy:

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0. \tag{11.5.9}$$

Dowód. Oznaczmy przez W wyznacznik występujący po lewej stronie równania (11.5.9). Po wykonaniu kolejno operacji elementarnych $w_1 - w_4$, $w_2 - w_4$, $w_3 - w_4$ oraz stosując rozwinięcie Laplace'a względem ostatniej kolumny macierzy związanej z wyznacznikiem W otrzymujemy, że:

$$W = \begin{vmatrix} x_1 - x_4 & y_1 - y_4 & z_1 - z_4 \\ x_2 - x_4 & y_2 - y_4 & z_2 - z_4 \\ x_3 - x_4 & y_3 - y_4 & z_3 - z_4 \end{vmatrix}.$$

Zatem $W \neq 0$ wtedy i tylko wtedy, gdy wektory $\overrightarrow{P_4P_1}$, $\overrightarrow{P_4P_2}$ i $\overrightarrow{P_4P_3}$ o wspólnym początku w punkcie P_4 są liniowo niezależne, co oznacza, że punkty P_1, P_2, P_3 i P_4 nie leżą w jednej płaszczyźnie.

11.5.4 Równanie odcinkowe płaszczyzny

Załóżmy, że niezerowe punkty P_1, P_2 oraz P_3 przestrzeni $\mathbb{R}^3$ leżą odpowiednio na osiach O_x, O_y i O_z układu współrzędnych O_{xyz}. Wtedy $P_1 = (a,0,0)$, $P_2 = (0,b,0)$ oraz $P_3 = (0,0,c)$ dla pewnych $a,b,c \in \mathbb{R} \setminus \{0\}$. Ze Stwierdzenia 11.12 wynika, że:

$$\begin{vmatrix} x & y & z & 1 \\ a & 0 & 0 & 1 \\ 0 & b & 0 & 1 \\ 0 & 0 & c & 1 \end{vmatrix} = 0 \tag{11.5.10}$$

jest równaniem płaszczyzny Π przechodzącej przez punkty P_1, P_2 i P_3. Stąd oraz na mocy wzorów (11.5.4)-(11.5.8) otrzymujemy, że równanie (11.5.10) równoważne jest równaniu:

$$bcx + ac(y-b) + abz = 0,$$

które z kolei można zapisać w równoważnej postaci:

$$bcx + acy + abz = abc.$$

Ponieważ $a \neq 0$, $b \neq 0$ i $c \neq 0$, to powyższe równanie równoważne jest równaniu:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \tag{11.5.11}$$

Jako, że niezerowe liczby rzeczywiste a,b,c występujące w powyższym równaniu niosą informację na temat punktów przecięcia płaszczyzn Π z osiami układu współrzędnych, to naturalna jest poniższa

Definicja 11.9. Równanie (11.5.11) nazywamy odcinkowym równaniem płaszczyzny.

Rozważmy równanie ogólne (11.5.1) płaszczyzny Π opisanej równaniem odcinkowym (11.5.11). Wówczas $D \neq 0$ (bo płaszczyzna Π przecina osie układu współrzędnych w niezerowych punktach), więc równanie (11.5.1) można zapisać równoważnie jako równanie:

$$\frac{A}{D}x + \frac{B}{D}y + \frac{C}{D}z = -1.$$

Powyższe równanie sprowadza się do równoważnej postaci:

$$\frac{x}{-\frac{D}{A}} + \frac{y}{-\frac{D}{B}} + \frac{z}{-\frac{D}{C}} = 1.$$

W ten sposób wykazaliśmy następujące

Stwierdzenie 11.14. Jeżeli płaszczyzna Π opisana jest jednocześnie równaniami (11.5.1) i (11.5.11), to $a = -\frac{D}{A}$, $b = -\frac{D}{B}$ oraz $ac = -\frac{D}{C}$.

11.5.5 Równanie normalne płaszczyzny

Niech Π będzie płaszczyzną opisaną równaniem (11.5.1), w którym $D \neq 0$. Wtedy z początku układu współrzędnych można poprowadzić prostopadle do płaszczyzny Π niezerowy wektor $\vec{v}$. Niech $\delta = \|\vec{v}\|$ i niech α, β oraz γ będą miarami kątów, które tworzy wektor $\vec{v}$ kolejno z osiami O_x, O_y i O_z układu współrzędnych. Rozważmy równanie:

$$x\cos\alpha + y\cos\beta + z\cos\gamma - \delta = 0. \tag{11.5.12}$$

Ponieważ $\vec{v} \neq \vec{0}$, to $\cos^2\alpha + \cos^2\beta + \cos^2\gamma > 0$. Zatem (11.5.12) jest równaniem pewnej płaszczyzny Π'. Pokażemy, że $\Pi' = \Pi$. Niech $\vec{u} = [A,B,C]$, gdzie A, B i C są współczynnikami równania (11.5.1) opisującego płaszczyznę Π. Wtedy $A^2 + B^2 + C^2 > 0$, więc $\vec{u} \neq \vec{0}$. Ponadto Wniosek 11.2 implikuje, że wektor $\vec{u}$ jest prostopadły do płaszczyzny Π. Stąd oraz na mocy określenia wektora $\vec{v}$ otrzymujemy, że $\vec{u} \parallel \vec{v}$. Zatem $\vec{v} = \lambda \circ \vec{u}$ dla pewnego $\lambda \in \mathbb{R} \setminus \{0\}$. Niech $P_0 = (x_0, y_0, z_0)$ będzie końcem wektora $\vec{v}$. Wtedy $\vec{v} = [x_0, y_0, z_0]$ i $P_0 \in \Pi$, więc z (11.5.1) wynika, że $D = -(Ax_0 + By_0 + Cz_0)$, czyli $-D = \langle \vec{u}|\vec{v}\rangle = \langle \vec{u}|\lambda \circ \vec{u}\rangle = \lambda \cdot \langle \vec{u}|\vec{u}\rangle = \lambda \cdot \|\vec{u}\|^2$ (por. punkty (ii) oraz (iii) Stwierdzenia 11.3). Stąd oraz na mocy punktu (iii) Stwierdzenia 11.1, $\frac{-D}{\|\vec{u}\|} = \lambda \cdot \|\vec{u}\| = \operatorname{sgn}(\lambda) \cdot |\lambda| \cdot \|\vec{u}\| = \operatorname{sgn}(\lambda) \cdot \|\lambda \circ \vec{u}\| = \operatorname{sgn}(\lambda) \cdot \|\vec{v}\| = \operatorname{sgn}(\lambda) \cdot \delta$. Zatem dla $\varepsilon = \operatorname{sgn}(\lambda)$ otrzymujemy, że $\frac{D}{\|\vec{u}\|} = -\varepsilon \cdot \delta$, skąd $-\delta = \frac{\varepsilon \cdot D}{\|\vec{u}\|} < 0$. Wobec tego:

$$\varepsilon = \begin{cases} 1 & \text{, gdy } D < 0 \\ -1 & \text{, gdy } D > 0 \end{cases}. \tag{11.5.13}$$

Stąd, jeżeli $D < 0$, to wektory $\vec{v}$ oraz $\vec{u}$ mają ten sam zwrot i w konsekwencji α, β oraz γ są również miarami kątów, które tworzy wektor $\vec{u}$ odpowiednio z osiami O_x,

O_y i O_z. Zatem wówczas $\cos\alpha = \frac{A}{\|\vec{u}\|}$, $\cos\beta = \frac{B}{\|\vec{u}\|}$ oraz $\cos\gamma = \frac{C}{\|\vec{u}\|}$. Jeśli natomiast $D > 0$, to powołując się ponownie na (11.5.13) uzyskujemy, że zwroty wektorów $\vec{u}$ i $\vec{v}$ są przeciwne, więc miary kątów, które tworzy wektor $\vec{u}$ z osiami O_x, O_y i O_z wynoszą kolejno $\pi-\alpha$, $\pi-\beta$ i $\pi-\gamma$. Ponadto $\frac{A}{\|u\|} = \cos(\pi-\alpha) = cos\pi\cdot\cos\alpha + \sin\pi\cdot\sin\alpha = -\cos\alpha$ i analogicznie $\frac{B}{\|\vec{u}\|} = -\cos\beta$ oraz $\frac{C}{\|\vec{u}\|} = -\cos\gamma$. Wobec tego $\cos\alpha = \frac{-A}{\|\vec{u}\|}$, $\cos\beta = \frac{-B}{\|\vec{u}\|}$ oraz $\cos\gamma = \frac{-C}{\|\vec{u}\|}$.

Ponadto $\|\vec{u}\| = \sqrt{A^2+B^2+C^2}$, więc mnożąc obustronnie równanie (11.5.1) przez liczbę:

$$\kappa = \frac{\varepsilon}{\sqrt{A^2+B^2+C^2}} \tag{11.5.14}$$

otrzymujemy równanie (11.5.12), skąd $\Pi' = \Pi$.

Odnotujmy jeszcze fakt, że z powyższych rozważań wynika, iż:

$$\cos\alpha = \pm\frac{A}{\sqrt{A^2+B^2+C^2}},\ \cos\beta = \pm\frac{B}{\sqrt{A^2+B^2+C^2}} \text{ i } \cos\gamma = \pm\frac{C}{\sqrt{A^2+B^2+C^2}}.$$

Zatem zachodzi równość:

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1. \tag{11.5.15}$$

Definicja 11.10. Niech Π będzie płaszczyzną opisaną równaniem (11.5.1), w którym $D \neq 0$, niech δ oznacza długość wektora poprowadzonego prostopadle do płaszczyzny Π z początku układu współrzędnych i niech α, β oraz γ będą miarami kątów, które tworzy wspomniany wektor kolejno z osiami O_x, O_y i O_z układu współrzędnych. Wówczas równanie (11.5.12) nazywamy równaniem normalnym płaszczyzny Π, zaś liczbę κ określoną wzorem (11.5.14) – czynnikiem normującym ogólne równanie płaszczyzny Π.

Uwaga 11.6. Przy zapisywaniu równania normalnego konkretnej płaszczyzny zazwyczaj nie podaje się miar kątów α, β i γ, lecz wartości cosinusów tych kątów. Mając więc informację, że $-\frac{3}{7}x - \frac{2}{7}y + \frac{6}{7}z - \frac{39}{7} = 0$ jest normalnym równaniem płaszczyzny Π należy pamiętać, że $\alpha = \arccos\left(-\frac{3}{7}\right)$, $\beta = \arccos\left(-\frac{2}{7}\right)$ oraz $\gamma = \arccos\frac{6}{7}$.

Uwaga 11.7. Równość (11.5.15) pozwala łatwo zweryfikować, czy ogólne równanie $Ax+By+Cz-D=0$ płaszczyzny, w którym $A,B,C \in [-1,1]$ oraz $D>0$, jest jednocześnie jej równaniem normalnym. Wystarczy bowiem sprawdzić, czy $A^2+B^2+C^2=1$. W szczególności wynika stąd, że równanie $-\frac{3}{7}x - \frac{2}{7}y + \frac{6}{7}z - \frac{39}{7} = 0$ podane w Uwadze 11.6 rzeczywiście jest równaniem normalnym pewnej płaszczyzny Π.

11.5.6 Równanie parametryczne płaszczyzny

Rozważmy dowolną płaszczyznę Π, dwa dowolne należące do niej punkty $P_0 = (x_0, y_0, z_0)$ i $P = (x, y, z)$ oraz dwa dowolne niewspółliniowe wektory $\vec{u} = [u_1, u_2, u_3]$ i $\vec{v} = [v_1, v_2, v_3]$ równoległe do płaszczyzny Π. Wówczas, w przestrzeni liniowej $\mathbb{R}^3$ mamy, że $\overrightarrow{P_0P} \in \text{lin}(\vec{u}, \vec{v})$, więc istnieją $t, k \in \mathbb{R}$ takie, że $\overrightarrow{P_0P} = t \circ \vec{u} + k \circ \vec{v}$. Ponadto $\overrightarrow{P_0P} = [x - x_0, y - y_0, z - z_0]$, więc równanie:

$$[x, y, z] = [x_0 + tu_1 + kv_1, y_0 + tu_2 + kv_2, z_0 + tu_3 + kv_3], \tag{11.5.16}$$

gdzie t oraz k przebiegają zbiór $\mathbb{R}$, opisuje płaszczyznę przechodzącą przez punkt $P_0 = (x_0, y_0, z_0)$ i równoległą do niewspółliniowych wektorów $\vec{u} = [u_1, u_2, u_3]$ oraz $\vec{v} = [v_1, v_2, v_3]$. Równanie (11.5.16) zapisuje się często w równoważny sposób jako układ równań:

$$\begin{cases} x = x_0 + tu_1 + ku_2 \\ y = y_0 + tu_1 + ku_2 \\ z = z_0 + tu_1 + ku_2 \end{cases}, \tag{11.5.17}$$

gdzie $t, k \in \mathbb{R}$.

Definicja 11.11. Wzór (11.5.16) nazywamy parametrycznym równaniem płaszczyzny przechodzącej przez punkt $P_0 = (x_0, y_0, z_0)$ i równoległej do niewspółliniowych wektorów $\vec{u} = [u_1, u_2, u_3]$ oraz $\vec{v} = [v_1, v_2, v_3]$ (parametrami są t oraz k). Równania układu (11.5.17) nazywamy równaniami parametrycznymi współrzędnych punktów tej płaszczyzny.

11.6 Odległość punktu od płaszczyzny w przestrzeni $\mathbb{R}^3$

Stwierdzenie 11.15. Odległość $d(P_0, \Pi)$ punktu $P_0 = (x_0, y_0, z_0)$ od płaszczyzny Π opisanej równaniem (11.5.1) wyraża się wzorem:

$$d(P_0, \Pi) = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}. \tag{11.6.1}$$

Dowód. Rozważmy punkt $P = (x, y, z)$ płaszczyzny Π taki, że wektor $\overrightarrow{PP_0}$ jest do niej prostopadły. Oznaczmy $\vec{\rho} = \overrightarrow{PP_0}$ oraz $\vec{\eta} = [A, B, C]$. Ponieważ $\vec{\rho} \perp \Pi$, $\vec{\eta} \perp \Pi$ oraz $\vec{\eta} \neq \vec{0}$, to $\vec{\rho} = \lambda \circ \vec{\eta}$ dla pewnego $\lambda \in \mathbb{R}$. Ponadto $D = -(Ax + By + Cz)$, $\vec{\rho} = [x_0 - x, y_0 - y, z_0 - z]$, $\|\vec{\eta}\| = \sqrt{A^2 + B^2 + C^2}$ i $\|\vec{\rho}\| = d(P_0, \Pi)$, więc $|Ax_0 + By_0 + Cz_0 + D| = |A(x_0 - x) + B(y_0 - y) + C(z_0 - z)| = |\langle \vec{\eta} | \vec{\rho} \rangle| = |\langle \vec{\eta} | \lambda \circ \vec{\eta} \rangle| = |\lambda \cdot \langle \vec{\eta} | \vec{\eta} \rangle| = |\lambda| \cdot \langle \vec{\eta} | \vec{\eta} \rangle = |\lambda| \cdot \|\vec{\eta}\|^2 = |\lambda| \cdot \|\vec{\eta}\| \cdot \sqrt{A^2 + B^2 + C^2} = \|\lambda \circ \vec{\eta}\| \cdot \sqrt{A^2 + B^2 + C^2} = \|\vec{\rho}\| \cdot \sqrt{A^2 + B^2 + C^2} = d(P_0, \Pi) \cdot \sqrt{A^2 + B^2 + C^2}$, co jest równoważne równości (11.6.1).

Bezpośrednią konsekwencją powyższego stwierdzenia oraz równości (11.5.15) jest następujący

Wniosek 11.5. Odległość punktu $P_0 = (x_0, y_0, z_0)$ od płaszczyzny Π opisanej równaniem (11.5.12) wyraża się wzorem:

$$d(P_0, \Pi) = |x_0 \cos\alpha + y_0 \cos\beta + z_0 \cos\gamma - \delta|. \tag{11.6.2}$$

11.7 Wzajemne położenie dwóch płaszczyzn

Stwierdzenie 11.16. Niech Π_1 oraz Π_2 będą płaszczyznami opisanymi odpowiednio równaniami ogólnymi $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$. Wówczas:

(i) $\Pi_1 \parallel \Pi_2$ wtedy i tylko wtedy, gdy $[A_2, B_2, C_2] = \lambda \circ [A_1, B_1, C_1]$ dla pewnego $\lambda \in \mathbb{R}$;
(ii) $\Pi_1 \perp \Pi_2$ wtedy i tylko wtedy, gdy $\big\langle [A_1, B_1, C_1] \big| [A_2, B_2, C_2] \big\rangle = 0$.

Dowód. Oznaczmy $\vec{v}_i = [A_i, B_i, C_i]$ dla $i \in \{1, 2\}$. Na mocy Wniosku 11.3 otrzymujemy wówczas, że $\vec{v}_i \perp \Pi_i$ dla $i \in \{1, 2\}$. Zatem $\Pi_1 \parallel \Pi_2$ wtedy i tylko wtedy, gdy $\vec{v}_1 \parallel \vec{v}_2$, co oznacza, że w przestrzeni liniowej $\mathbb{R}^3$ zachodzi $\vec{v}_2 \in \mathrm{lin}_{\mathbb{R}}(\vec{v}_1)$, co z kolei równoważne jest istnieniu takiej $\lambda \in \mathbb{R}$, że $\vec{v}_2 = \lambda \circ \vec{v}_1$. W ten sposób udowodniliśmy punkt (i). Aby uzasadnić punkt (ii) zauważmy, że $\Pi_1 \perp \Pi_2$ wtedy i tylko wtedy, gdy $\vec{v}_1 \perp \vec{v}_2$, co wobec punktu (vii) Stwierdzenia 11.3 oznacza, że $\langle \vec{v}_1 | \vec{v}_2 \rangle = 0$.

Bezpośrednią konsekwencją powyższego stwierdzenia oraz rozważań związanych z normalnym równaniem płaszczyzny (11.5.12) jest następujący

Wniosek 11.6. Dla płaszczyzn Π_1 oraz Π_2 opisanych odpowiednio równaniami normalnymi $x\cos\alpha_1 + y\cos\beta_1 + z\cos\gamma_1 - \delta_1 = 0$ i $x\cos\alpha_2 + y\cos\beta_2 + z\cos\gamma_2 - \delta_2 = 0$ następujące warunki są równoważne:

(i) $\Pi_1 \parallel \Pi_2$;
(ii) $[\cos\alpha_1, \cos\beta_1, \cos\gamma_1] = \pm 1 \circ [\cos\alpha_2, \cos\beta_2, \cos\gamma_2]$;
(iii) $\alpha_2 \in \{\alpha_1, \pi - \alpha_1\}$, $\beta_2 \in \{\beta_1, \pi - \beta_1\}$ i $\gamma_2 \in \{\gamma_1, \pi - \gamma_1\}$.

Definicja 11.12. Kątem między dwiema płaszczyznami nazywamy kąt między ich wektorami normalnymi.

Na mocy powyższej definicji oraz Wniosku 11.1 uzyskujemy natychmiast następujące

Stwierdzenie 11.17. Cosinus kąta ostrego φ między płaszczyznami Π oraz Π_2 opisanymi odpowiednio równaniami ogólnymi $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$ wyraża się wzorem:

$$\cos\varphi = \frac{|A_1A_2 + B_1B_2 + C_1C_2|}{\sqrt{A_1^2 + B_1^2 + C_1^2} \cdot \sqrt{A_2^2 + B_2^2 + C_2^2}}. \tag{11.7.1}$$

Stwierdzenie 11.18. Niech Π_1 oraz Π_2 będą płaszczyznami opisanymi odpowiednio równaniami ogólnymi $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$. Wówczas $\Pi_2 = \Pi_1$ wtedy i tylko wtedy, gdy istnieje $\lambda \in \mathbb{R}$ taka, że $[A_2, B_2, C_2, D_2] = \lambda \circ [A_1, B_1, C_1, D_1]$.

Dowód. Równość $\Pi_2 = \Pi_1$ oznacza, że równania $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$ są równoważne. Ponadto $A_i^2 + B_i^2 + C_i^2 > 0$ dla $i \in \{1,2\}$, więc równoważność wspomnianych równań tożsama jest z istnieniem $\lambda \in \mathbb{R}$ takiej, że $[A_2, B_2, C_2, D_2] = \lambda \circ [A_1, B_1, C_1, D_1]$.

Stwierdzenie 11.19. Odległość $d(\Pi_1, \Pi_2)$ między dwiema równoległymi płaszczyznami Π_1 oraz Π_2 opisanymi kolejno równaniami normalnymi $x\cos\alpha_1 + y\cos\beta_1 + z\cos\gamma_1 - \delta_1 = 0$ i $x\cos\alpha_2 + y\cos\beta_2 + z\cos\gamma_2 - \delta_2 = 0$ wyraża się wzorem:

$$d(\Pi_1, \Pi_2) = \begin{cases} \delta_1 + \delta_2 & \text{, gdy } \operatorname{sgn}(\cos\alpha_1) \neq \operatorname{sgn}(\cos\alpha_2) \\ |\delta_1 - \delta_2| & \text{, gdy } \operatorname{sgn}(\cos\alpha_1) = \operatorname{sgn}(\cos\alpha_2) \end{cases}. \tag{11.7.2}$$

Dowód. Niech $\vec{v}_i$ będzie wektorem poprowadzonym z początku układu współrzędnych prostopadle do płaszczyzny Π_i dla $i \in \{1,2\}$. Wtedy $\|\vec{v}_i\| = \delta_i$ oraz α_i jest kątem jaki tworzy wektor $\vec{v}_i$ z osią O_x układu współrzędnych, dla $i \in \{1,2\}$ (zob. Definicja 11.10 i poprzedzający ją komentarz). Ponadto $\vec{v}_2 \parallel \vec{v}_1$, więc warunki $\operatorname{sgn}(\cos\alpha_1) \neq \operatorname{sgn}(\cos\alpha_2)$ i $\operatorname{sgn}(\cos\alpha_1) = \operatorname{sgn}(\cos\alpha_2)$ oznaczają, że zwroty wektorów $\vec{v}_1$ oraz $\vec{v}_2$ są odpowiednio przeciwne i zgodne, co uzasadnia wzór (11.7.2).

Przykład 11.3. Wyznaczymy równanie ogólne $Ax + By + Cz + D = 0$ płaszczyzny Π przechodzącej przez punkt $P_0 = (2,-1,3)$ prostopadłej do dwóch płaszczyzn Π_1 oraz Π_2 opisanych odpowiednio równaniami ogólnymi $5x + 2y - 7z + 1$ i $-3x + y - 2z - 4 = 0$.

1. sposób. Oznaczmy $\vec{\eta} = [A,B,C]$, $\vec{\eta}_1 = [5,2,-7]$ i $\vec{\eta}_2 = [-3,1,-2]$. Ponieważ $\vec{\eta} \perp \Pi$, $\vec{\eta}_1 \perp \Pi_1$, $\vec{\eta}_2 \perp \Pi_2$ oraz $\Pi \perp \Pi_1$ i $\Pi \perp \Pi_2$, to $\vec{\eta} \perp \vec{\eta}_1$ oraz $\vec{\eta} \perp \vec{\eta}_2$. Przyjmijmy więc $\vec{\eta} = \vec{\eta}_1 \times \vec{\eta}_2$ (por. punkt (vii) Stwierdzenia 11.5). Zatem:

$$\vec{\eta} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 5 & 2 & -7 \\ -3 & 1 & -2 \end{vmatrix} = (-4+7) \circ \vec{i} + (21+10) \circ \vec{j} + (5+6) \circ \vec{k} = [3,31,11],$$

skąd $A = 3$, $B = 31$ i $C = 11$. Ponadto $P_0 \in \Pi$, więc ze wzoru (11.5.2) wynika, że płaszczyzna Π opisana jest równaniem:

$$3(x-2) + 31(y+1) + 11(z-3) = 0. \tag{11.7.3}$$

Wobec tego $D = -6+31-33 = -8$. Ostatecznie otrzymujemy stąd, że:

$$3x+31y+11z-8=0$$

jest szukanym równaniem ogólnym płaszczyzny Π.

2. sposób. Ponieważ $P_0 \in \Pi$, to ze wzoru (11.5.2) wynika, że równanie płaszczyzny Π jest postaci:

$$A(x-2)+B(y+1)+C(z-3)=0, \text{ przy czym } A^2+B^2+C^2>0. \tag{11.7.4}$$

Ponadto $\Pi \perp \Pi_1$ oraz $\Pi \perp \Pi_2$, więc Stwierdzenie 11.16 implikuje, że:

$$\big\langle [A,B,C] \big| [5,2,-7] \big\rangle = 0 \text{ i } \big\langle [A,B,C] \big| [-3,1,-2] \big\rangle = 0.$$

Stąd oraz na mocy wzoru (11.2.2) otrzymujemy układ warunków:

$$\begin{cases} (x-2)A+(y+1)B+(z-3)C=0 \\ 5A+2B-7C=0 \\ -3A+B-2C=0 \end{cases}. \tag{11.7.5}$$

Załóżmy, że liczby rzeczywiste x, y i z występujące w równaniu (11.7.5) są współrzędnymi dowolnie ustalonego punktu płaszczyzny Π. Układ (11.7.5) możemy wówczas traktować jak jednorodny układ równań liniowych z niewiadomymi A, B i C spełniającymi warunek $A^2+B^2+C^2>0$. Ponieważ szukana płaszczyzna istnieje, to układ ten posiada niezerowe rozwiązanie. Z Twierdzenia Cramera wynika więc, że wyznacznik główny W tego układu równy jest 0. Ponadto:

$$W = \begin{vmatrix} x-2 & y+1 & z-3 \\ 5 & 2 & -7 \\ -3 & 1 & -2 \end{vmatrix}, \tag{11.7.6}$$

więc $-4(x-2)+5(z-3)+21(y+1)+6(z-3)+10(y+1)+7(x-2)=0$, czyli:

$$3(x-2)+31(y+1)+11(z-3)=0. \tag{11.7.7}$$

Porównując powyższe równanie z pierwszym równaniem układu (11.7.5) uzyskujemy, że $A=3$, $B=31$ oraz $C=11$. Ponadto z równania (11.7.7) wynika, że $D=-8$, więc $3x+31y+11z-8=0$ jest szukanym równaniem ogólnym płaszczyzny Π.

Przykład 11.4. Wyznaczymy równanie ogólne $Ax+By+Cz+D=0$ płaszczyzny Π przechodzącej przez punkty $P_0=(2,-1,3)$ i $P_1=(-5,2,4)$ oraz prostopadłej do płaszczyzny Π' opisanej równaniem $-3x+5y-7z+1=0$. W świetle Stwierdzenia 11.16 i wzoru (11.2.2) warunek $\Pi \perp \Pi'$ implikuje, że:

$$-3A+5B-7C=0. \tag{11.7.8}$$

Ponieważ $P_0 \in \Pi$, to powołując się na wzór (11.5.2) otrzymujemy, że zachodzi równość (11.7.4). Ponadto $P_1 \in \Pi$, więc współrzędne punktu P_1 spełniają równanie (11.7.4) opisujące płaszczyznę Π. Zatem:

$$-7A+3B+C=0. \tag{11.7.9}$$

Traktując liczby rzeczywiste x, y i z występujące w równaniu (11.7.4) jak współrzędne dowolnie ustalonego punktu płaszczyzny Π i rozważając to równanie wraz z równaniami (11.7.8) oraz (11.7.9) otrzymujemy następujący jednorodny układ równań linowych:

$$\begin{cases} (x-2)A+(y+1)B+(z-3)C=0 \\ -3A+5B-7C=0 \\ -7A+3B+C=0 \end{cases} \tag{11.7.10}$$

z niewiadomymi A, B i C spełniającymi warunek $A^2+B^2+C^2>0$. Ponieważ szukana płaszczyzna istnieje, to układ ten posiada niezerowe rozwiązanie. Powołując się na Twierdzenia Cramera otrzymujemy stąd, że wyznacznik główny tego układu jest równy 0, czyli:

$$\begin{vmatrix} x-2 & y+1 & z-3 \\ -3 & 5 & -7 \\ -7 & 3 & 1 \end{vmatrix} = 0, \tag{11.7.11}$$

skąd $5(x-2)-9(z-3)+49(y+1)-105(z-3)+3(y+1)+21(x-2)=0$, co oznacza, że:

$$26(x-2)+52(y+1)-114(z-3)=0. \tag{11.7.12}$$

Porównując równanie (11.7.12) z pierwszym równaniem układu (11.7.5) otrzymujemy, że $A=26$, $B=52$ oraz $C=-114$. Ponadto $D=-2\cdot 26+52+3\cdot 114=342$ na mocy (11.7.7). Wobec tego $26x+52y-114z+342=0$ jest szukanym równaniem ogólnym płaszczyzny Π.

Przykład 11.5. Wyznaczymy ogólne równanie $Ax+By+Cz+D=0$ płaszczyzny Π równoległej do płaszczyzny Π' opisanej równaniem $-3x+2y-6z+4=0$ takiej, że $d(\Pi,\Pi')=5$. Najpierw sprowadzimy równanie ogólne płaszczyzny Π' do postaci normalnej. Ponieważ $4>0$, to czynnikiem normującym równanie $-3x+2y-6z+4=0$ jest $\kappa'=\frac{-1}{\sqrt{(-3)^2+2^2+(-6)^2}}=\frac{-1}{\sqrt{49}}=-\frac{1}{7}$. Zatem:

$$\frac{3}{7}x-\frac{2}{7}y+\frac{6}{7}z-\frac{4}{7}=0 \tag{11.7.13}$$

jest równaniem normalnym płaszczyzny Π'. W szczególności wynika stąd, że $\delta'=\frac{4}{7}$ jest odległością płaszczyzny Π' od początku układu współrzędnych. Niech δ oznacza odległość płaszczyzny Π od początku układu współrzędnych. Ponieważ $d(\Pi,\Pi')=5>\frac{4}{7}=\delta'$, to możliwe są dwa przypadki:

(i). płaszczyzna Π leży po przeciwnej stronie początku układu współrzędnych co płaszczyzna Π'. Wtedy $d(\Pi,\Pi') = \delta' + \delta$ oraz zwroty wektorów $\vec{v}$ i $\vec{v'}$ poprowadzonych z początku układu współrzędnych prostopadle odpowiednio do płaszczyzn Π oraz Π' są przeciwne. Zatem $\delta = \frac{39}{7}$ oraz znaki cosinusów kątów nachylenia tych wektorów do osi O_x są przeciwne. Wobec tego płaszczyzna Π opisana jest równaniem normalnym:

$$-\frac{3}{7}x - \frac{2}{7}y + \frac{6}{7}z - \frac{39}{7} = 0. \tag{11.7.14}$$

(ii). płaszczyzna Π leży po tej samej stronie początku układu współrzędnych co płaszczyzna Π'. Wówczas $d(\Pi,\Pi') = \delta' + \delta$ oraz znaki cosinusów kątów nachylenia opisanych w punkcie (i) wektorów $\vec{v}$ i $\vec{v'}$ do osi O_x są zgodne, skąd $\delta = \frac{31}{7}$ i w konsekwencji płaszczyzna Π jest opisana równaniem normalnym:

$$\frac{3}{7}x - \frac{2}{7}y + \frac{6}{7}z - \frac{31}{7} = 0. \tag{11.7.15}$$

Równania (11.7.14) i (11.7.14) są w szczególności równaniami ogólnymi płaszczyzny.

Przykład 11.6. Wyznaczymy równanie płaszczyzny Π równoodległej od płaszczyzn Π_1 oraz Π_2 opisanych odpowiednio równaniami $2x+3y-4z+5=0$ oraz $-6x-9y+12z+3=0$. Czynnikiem normującym równanie $2x+3y-4z+5=0$ płaszczyzny Π_1 jest $\kappa_1 = \frac{-1}{\sqrt{2^2+3^2+(-4)^2}} = -\frac{1}{29}$. Zauważmy, że po obustronnym podzieleniu równania $-6x-9y+12z+3=0$ przez -3 otrzymujemy równanie $2x+3y-4z-1=0$. Czynnikiem normującym tego równania jest $\kappa_2 = \frac{1}{29}$. Zatem:

$$-\frac{2}{29}x - \frac{3}{29}y + \frac{4}{29}z - \frac{5}{29} = 0 \tag{11.7.16}$$

oraz

$$\frac{2}{29}x + \frac{3}{29}y - \frac{4}{29}z - \frac{1}{29} = 0 \tag{11.7.17}$$

są równaniami normalnymi odpowiednio płaszczyzn Π_1 oraz Π_2. Stąd oraz na mocy Wniosku 11.6, $\Pi_1 \parallel \Pi_2$. Zatem płaszczyzna Π istnieje. Dalej, z (11.7.16) i (11.7.17) wynika, że odległość δ_1 płaszczyzny Π_1 od początku układu współrzędnych wynosi $\frac{5}{29}$, zaś odległość δ_2 płaszczyzny Π_2 od początku układu współrzędnych wynosi $\frac{1}{29}$. Ponadto $d(\Pi_1,\Pi_2) = \delta_1 + \delta_2 = \frac{6}{29}$ na mocy Stwierdzenia 11.19. W szczególności wynika stąd, że płaszczyzny Π_1 oraz Π_2 leżą po przeciwnych stronach początku układu współrzędnych. Ponieważ płaszczyzna Π jest równoodległa od płaszczyzn Π_1 oraz Π_2, to znajduje się ona po tej samej stronie początku układu współrzędnych co płaszczyzna Π_1 oraz jej odległość δ od początku układu współrzędnych jest mniejsza niż δ_1. Powołując się ponownie na Stwierdzenie 11.19 otrzymujemy więc, że $d(\Pi,\Pi_1) = |\delta - \delta_1| = \delta_1 - \delta$. Ponadto $d(\Pi,\Pi_1) = \frac{1}{2}\cdot d(\Pi_1,\Pi_2) = \frac{3}{29}$,

skąd $\delta = \delta_1 - d(\Pi, \Pi_1) = \frac{5}{29} - \frac{3}{29} = \frac{2}{29}$.Wobec tego płaszczyzna Π opisana jest równaniem normalnym:

$$-\frac{2}{29}x - \frac{3}{29}y + \frac{4}{29}z - \frac{2}{29} = 0.$$

11.8 Równania prostej w przestrzeni $\mathbb{R}^3$

11.8.1 Równanie parametryczne prostej

Rozważmy dowolną jednowymiarową podprzestrzeń V przestrzeni liniowej $\mathbb{R}^3$. Wtedy $V = \mathrm{lin}_{\mathbb{R}}([a,b,c])$ dla pewnych $a,b,c \in \mathbb{R}$ takich, że $a^2+b^2+c^2 \neq 0$. Podprzestrzeń V możemy geometrycznie interpretować jako prostą l_0 przechodzącą przez punkty $(0,0,0)$ i (a,b,c). W szczególności l_0 jest prostą przechodzącą przez punkt $(0,0,0)$ i równoległą do niezerowego wektora $\vec{\eta} = [a,b,c]$ opisaną równaniem:

$$[x,y,z] = \lambda \circ [a,b,c], \tag{11.8.1}$$

gdzie λ przebiega zbiór $\mathbb{R}$. Zatem prosta l równoległa do wektora $\vec{\eta}$ przechodząca przez punkt $P_0 = (x_0,y_0,z_0)$ jest obrazem prostej prostej l_0 względem translacji T o wektor $\overrightarrow{OP_0} = [x_0,y_0,z_0]$. Wobec tego prosta l opisana jest równaniem:

$$[x,y,z] = [x_0,y_0,z_0] + \lambda \circ [a,b,c], \tag{11.8.2}$$

gdzie λ przebiega zbiór $\mathbb{R}$.

Na odwrót. Rozważmy podzbiór k przestrzeni $\mathbb{R}^3$ opisany warunkiem (11.8.2) i oznaczmy $k_0 = T^{-1}(k)$. Wówczas zbiór k_0 opisany jest warunkiem (11.8.1), więc geometrycznie k_0 jest prostą przechodzącą przez punkt $(0,0,0)$ równoległą do wektora $[a,b,c]$. Stąd k również jest prostą równoległą do wektora $[a,b,c]$. Ponadto wprost ze wzoru (11.8.2) wynika, że (x_0,y_0,z_0) jest punktem prostej k.

Równanie (11.8.2) równoważne jest układowi równań:

$$\begin{cases} x = x_0 + \lambda a \\ y = y_0 + \lambda b \\ z = z_0 + \lambda c \end{cases}, \tag{11.8.3}$$

z rzeczywistym parametrem λ opisującemu kolejne współrzędne punktów należących do prostej przechodzącej przez punkt $P_0 = (x_0,y_0,z_0)$ równoległej do niezerowego wektora $[a,b,c]$ przestrzeni $\mathbb{R}^3$.

Powyższe obserwacje motywują następującą definicję.

Definicja 11.13. Równanie (11.8.2) nazywamy parametrycznym równaniem prostej l przechodzącej przez punkt $P_0 = (x_0, y_0, z_0)$ równoległej do niezerowego wektora $\vec{\eta} = [a, b, c]$ przestrzeni $\mathbb{R}^3$, zaś równania układu (11.8.3) określamy mianem równań parametrycznych współrzędnych punktów tej prostej. Wektor $\vec{\eta}$ nazywamy wektorem kierunkowym prostej l, natomiast jego współrzędne – współczynnikami kierunkowymi prostej l.

11.8.2 Równania kierunkowe prostej

Załóżmy, że (x, y, z) jest dowolnym punktem prostej l przechodzącej przez punkt $P_0 = (x_0, y_0, z_0)$ równoległej do niezerowego wektora $\vec{\eta} = [a, b, c]$. Z (11.8.3) wynikają wówczas równości $x - x_0 = \lambda a$, $y - y_0 = \lambda b$ oraz $z - z_0 = \lambda c$, gdzie λ jest pewną liczbą rzeczywistą. Ponieważ $\vec{\eta} \neq \vec{0}$, to ma sens co najmniej jeden spośród ilorazów $\frac{x-x_0}{a}$, $\frac{y-y_0}{b}$ i $\frac{z-z_0}{c}$. Ponadto wszystkie spośród tych ilorazów, które mają sens są równe λ. Wobec tego:

(i) jeżeli $a \neq 0$, $b \neq 0$ i $c \neq 0$, to prosta l opisana jest równaniami:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}; \tag{11.8.4}$$

(ii) jeżeli $a = 0$ oraz $b \neq 0$ i $c \neq 0$, to prosta l jest prostopadła do osi O_x i jest opisana układem równań:

$$\begin{cases} x = x_0 \\ \frac{y-y_0}{b} = \frac{z-z_0}{c} \end{cases}; \tag{11.8.5}$$

(iii) jeżeli $a \neq 0$ oraz $b = 0$ i $c \neq 0$, to prosta l jest prostopadła do osi O_y i jest opisana układem równań:

$$\begin{cases} y = y_0 \\ \frac{x-x_0}{a} = \frac{z-z_0}{c} \end{cases}; \tag{11.8.6}$$

(iv) jeżeli $a \neq 0$ oraz $b \neq 0$ i $c = 0$, to prosta l jest prostopadła do osi O_z i jest opisana układem równań:

$$\begin{cases} z = z_0 \\ \frac{x-x_0}{a} = \frac{y-y_0}{b} \end{cases}; \tag{11.8.7}$$

(v) jeżeli $a = 0$ oraz $b = 0$ i $c \neq 0$, to prosta l jest równoległa do osi O_z i jest opisana układem równań:

$$\begin{cases} x = x_0 \\ y = y_0 \\ z = z_0 + c\lambda \end{cases}, \tag{11.8.8}$$

gdzie λ przebiega zbiór $\mathbb{R}$;

(vi) jeżeli $a = 0$ oraz $b \neq 0$ i $c = 0$, to prosta l jest równoległa do osi O_y i jest opisana układem równań:

$$\begin{cases} x = x_0 \\ y = y_0 + b\lambda \\ z = z_0 \end{cases}, \tag{11.8.9}$$

gdzie λ przebiega zbiór $\mathbb{R}$;

(vii) jeżeli $a \neq 0$ oraz $b = 0$ i $c = 0$, to prosta l jest równoległa do osi O_x i jest opisana układem równań:

$$\begin{cases} x = x_0 + a\lambda \\ y = y_0 \\ z = z_0 \end{cases}, \tag{11.8.10}$$

gdzie λ przebiega zbiór $\mathbb{R}$.

Definicja 11.14. Równania (11.8.4) nazywamy równaniami kierunkowymi prostej przechodzącej przez punkt $P_0 = (x_0, y_0, z_0)$ równoległej do wektora $\vec{\eta} = [a, b, c]$ takiego, że $abc \neq 0$.

Uwaga 11.8. O równaniach kierunkowych prostej przechodzącej przez punkt $P_0 = (x_0, y_0, z_0)$ równoległej do niezerowego wektora $\vec{\eta} = [a, b, c]$ mówi się również przy braku założenia niezerowości wszystkich współrzędnych wektora $\vec{\eta}$. Przyjmuje się wówczas umowę, że zerowość mianownika dowolnego ułamka występującego w (11.8.4) pociąga za sobą zerowość licznika tego ułamka, co w praktyce prowadzi do któregoś spośród układów równań (11.8.5)–(11.8.10). W szczególności równania $z = z_0 + \lambda c$, $y = y_0 + \lambda b$ i $x = x_0 + \lambda a$ z rzeczywistym parametrem λ równoważne są odpowiednio warunkom $z \in \mathbb{R}$, $y \in \mathbb{R}$ oraz $x \in \mathbb{R}$.

11.8.3 Równanie prostej przechodzącej przez dwa punkty

Rozważmy dwa dowolne punkty $P_1 = (x_1, y_1, z_1)$ i $P_2 = (x_2, y_2, z_2)$ przestrzeni $\mathbb{R}^3$ i oznaczmy $\vec{\eta} = \overrightarrow{P_1P_2}$. Wtedy $\vec{\eta} \neq \vec{0}$, bo $P_2 \neq P_1$. Ponadto $\vec{\eta} = [x_2 - x_1, y_2 - y_1, z_2 - z_1]$, więc na mocy rozważań związanych z Definicją 11.13 otrzymujemy, że równanie:

$$[x, y, z] = [x_1, y_1, z_1] + \lambda \circ \big[x_2 - x_1, y_2 - y_1, z_2 - z_1\big], \tag{11.8.11}$$

gdzie λ przebiega zbiór liczb rzeczywistych, opisuje prostą l przechodzącą przez punkt P_1 równoległą do wektora $\vec{\eta}$. Zauważmy, że podstawiając $\lambda = 1$ w (11.8.11) uzyskujemy $[x, y, z] = [x_2, y_2, z_2]$. Zatem również P_2 jest punktem prostej l. Otrzymujemy stąd natychmiast następujące

Stwierdzenie 11.20. Prosta l przechodząca przez dwa punkty $P_1 = (x_1, y_1, z_1)$ i $P_2 = (x_2, y_2, z_2)$ opisana jest równaniem (11.8.11).

Łatwo zauważyć, że równanie (11.8.11) równoważne jest każdemu spośród równań:

$$[x,y,z] = [x_1,y_1,z_1] + \lambda \circ \big[x_1 - x_2, y_1 - y_2, z_1 - z_2\big],$$

$$[x,y,z] = [x_2,y_2,z_2] + \lambda \circ \big[x_2 - x_1, y_2 - y_1, z_2 - z_1\big],$$

$$[x,y,z] = [x_2,y_2,z_2] + \lambda \circ \big[x_1 - x_2, y_1 - y_2, z_1 - z_2\big],$$

gdzie λ przebiega zbiór $\mathbb{R}$.

Ponadto równanie (11.8.11) równoważne jest układowi równań:

$$\begin{cases} x - x_1 = \lambda(x_2 - x_1) \\ y - y_1 = \lambda(y_2 - y_1) \\ z - z_1 = \lambda(z_2 - z_1) \end{cases}.$$

Stąd, przy zachowaniu umowy z Uwagi 11.8, równanie (11.8.11) prostej przechodzącej przez dwa punkty $P_1 = (x_1,y_1,z_1)$ i $P_2 = (x_2,y_2,z_2)$ zastępuje się często równaniami:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

11.8.4 Postać krawędziowa prostej

Rozważmy dwie nierównoległe płaszczyzny Π_1 oraz Π_2 opisane odpowiednio równaniami ogólnymi $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$. Częścią wspólną tych płaszczyzn jest wówczas prosta l opisana układem równań:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}. \tag{11.8.12}$$

Definicja 11.15. Układ równań (11.8.12) nazywamy postacią krawędziową prostej będącej częścią wspólną dwóch nierównoległych płaszczyzn opisanych równaniami ogólnymi $A_1x + B_1y + C_1z + D_1 = 0$ i $A_2x + B_2y + C_2z + D_2 = 0$.

Następne stwierdzenie ilustruje ważny związek między postacią krawędziową prostej a współczynnikami kierunkowymi tej prostej.

Stwierdzenie 11.21. Jeżeli układ równań (11.8.12) jest postacią krawędziową prostej l oraz $\vec{\eta} = [a,b,c]$, gdzie:

$$a = \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix},\ b = \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix} \text{ i } c = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}, \tag{11.8.13}$$

to $\vec{\eta}$ jest wektorem kierunkowym tej prostej.

Dowód. Niech $\vec{\eta}_i = [A_i, B_i, C_i]$ i niech Π_i będzie płaszczyzną opisaną równaniem ogólnym $A_i x + B_i y + C_i z + D_i = 0$ dla $i \in \{1,2\}$. Ze wzoru (11.3.2) wynika, że $\vec{\eta}_1 \times \vec{\eta}_2 = \vec{\eta}$. Odpowiednio na mocy punktu (vii) Stwierdzenia 11.5 i Wniosku 11.3 otrzymujemy, że $\vec{\eta} \perp \vec{\eta}_1$, $\vec{\eta} \perp \vec{\eta}_2$, $\eta_1 \perp \Pi_1$ oraz $\eta_2 \perp \Pi_2$. Ponadto płaszczyzny Π_1 oraz Π_2 nie są równoległe, więc równoległe nie są także niezerowe wektory $\vec{\eta}_1$ i $\vec{\eta}_2$. Stąd oraz na mocy punktu (vi) Stwierdzenia 11.5, $\vec{\eta} \neq \vec{0}$. Wobec tego $\vec{\eta} \parallel \Pi_1$ i $\vec{\eta} \parallel \Pi_2$, skąd $\vec{\eta} \parallel l$ (bo $l = \Pi_1 \cap \Pi_2$). Oznacza to, iż $\vec{\eta}$ jest wektorem kierunkowym prostej l.

11.8.5 Równanie parametryczne prostej a jej postać krawędziowa

Załóżmy, że znana jest postać krawędziowa prostej l i dana jest ona układem (11.8.12). Wyznaczymy równanie parametryczne tej prostej. Niech $\vec{\eta} = [a,b,c]$ będzie wektorem kierunkowym prostej l. Współrzędne tego wektora obliczmy, stosując Stwierdzenie 11.21. W szczególności wynika stąd, że co najmniej jeden z wyznaczników danych w (11.8.13) jest niezerowy (bo $\vec{\eta} \neq \vec{0}$). Bez utraty ogólności możemy przyjąć, że jest to wyznacznik c. Weźmy dowolne $z_0 \in \mathbb{R}$ i podstawy $z = z_0$ w (11.8.12). Ten układ równań przyjmuje wówczas postać:

$$\begin{cases} A_1 x + B_1 y = -(C_1 z_0 + D_1) \\ A_2 x + B_2 y = -(C_2 z_0 + D_2) \end{cases}. \tag{11.8.14}$$

Ponieważ $c \neq 0$ i wyznacznik główny układu (11.8.14) równy jest c, to z Twierdzenia Cramera wynika, że układ ten ma dokładnie jedno rozwiązanie (x_0, y_0). Zatem $P_0 = (x_0, y_0, z_0)$ jest punktem prostej l. Wobec tego prosta l opisana jest równaniem (11.8.2).

Na odwrót. Przypuśćmy, że prosta l opisana jest równaniem (11.8.2). Ponieważ wektor $[a,b,c]$ jest niezerowy to mamy do rozważenia trzy przypadki:

(i). $a \neq 0$, $b \neq 0$ i $c \neq 0$. Z rozważań związanych z równaniami (11.8.4) wynika wówczas, że równanie (11.8.2) można zastąpić układem równań:

$$\begin{cases} \frac{x-x_0}{a} = \frac{y-y_0}{b} \\ \frac{y-y_0}{b} = \frac{z-z_0}{c} \end{cases}, \tag{11.8.15}$$

który można przekształcić do układu równań postaci (11.8.12), przy czym w każdym z dwóch równań przekształconego układu spełniony będzie warunek niezerowości przynajmniej jednego współczynnika.

(ii). Dwie spośród liczb a, b i c są niezerowe. Bez utraty ogólności możemy przyjąć wówczas, że $a \neq 0$, $b \neq 0$ i $c = 0$. Równanie (11.8.2) równoważne jest wtedy układowi równań (11.8.7), który również można zapisać w postaci (11.8.12) i w każdym z dwóch równań przekształconego układu przynajmniej jeden współczynnik będzie

niezerowy.

(iii). Jedna spośród liczb a, b i c jest niezerowa. Bez utraty ogólności możemy przyjąć, że $c \neq 0$. Wówczas równanie (11.8.2) równoważne jest układowi równań (11.8.8). Stąd oraz na mocy Uwagi 11.8 otrzymujemy, że układ (11.8.8) można sprowadzić do układu (11.8.12) przy zachowaniu warunku, iż w każdym z dwóch równań przekształconego układu co najmniej jeden współczynnik jest niezerowy.

11.9 Wzajemne położenie dwóch prostych

Stwierdzenie 11.22. Proste l_1 oraz l_2 opisane równaniami parametrycznymi $[x,y,z] = [x_1,y_1,z_1] + \lambda \circ [a_1,b_1,c_1]$ i $[x,y,z] = [x_2,y_2,z_2] + \lambda \circ [a_2,b_2,c_12]$ są:

(i) równoległe wtedy i tylko wtedy, gdy $[a_2,b_2,c_2] = \lambda \circ [a_1,b_1,c_1]$ dla pewnego $\lambda \in \mathbb{R}$;
(ii) prostopadłe wtedy i tylko wtedy, gdy $a_1a_1 + b_1b_2 + c_1c_2 = 0$;
(iii) współpłaszczyznowe wtedy i tylko wtedy, gdy:

$$\begin{vmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0. \tag{11.9.1}$$

Dowód. Niech $\vec{\eta}_i = [a_i,b_i,c_i]$ dla $i \in \{1,2\}$. Ponieważ $l_i \parallel \vec{\eta}_i$ dla $i \in \{1,2\}$, to:

(i). $l_1 \parallel l_2$ wtedy i tylko wtedy, gdy $\vec{\eta}_1 \parallel \vec{\eta}_2$, czyli gdy istnieje $\lambda \in \mathbb{R}$ taka, że $[a_2,b_2,c_2] = \lambda \circ [a_1,b_1,c_1]$;

(ii). $l_1 \perp l_2$ wtedy i tylko wtedy, gdy $\vec{\eta}_1 \perp \vec{\eta}_2$, co wobec punktu (vii) Stwierdzenia 11.3 oznacza, że $\langle \vec{\eta}_1 | \vec{\eta}_2 \rangle = 0$, czyli $a_1a_1 + b_1b_2 + c_1c_2 = 0$.

Pozostało udowodnić punkt (iii) niniejszego stwierdzenia. Załóżmy najpierw, że proste l_1 i l_2 są współpłaszczyznowe. Wówczas $l_1 \parallel l_2$ albo proste l_1 i l_2 przecinają się w pewnym punkcie. W pierwszym przypadku udowodniony wcześniej punkt (i) implikuje istnienie takiej $\lambda \in \mathbb{R}$, że $[a_2,b_2,c_2] = \lambda \circ [a_1,b_1,c_1]$, więc równość (11.9.1) wynika ze Stwierdzeń 4.6 i 4.7. Przypuśćmy teraz, że zachodzi druga ze wspomnianych wyżej możliwości i oznaczmy $\vec{v} = [x_1 - x_2, y_1 - y_2, z_1 - z_2]$. Istnieją wówczas $\lambda_1, \lambda_2 \in \mathbb{R}$ takie, że $\vec{v} = \lambda_1 \circ \vec{\eta}_1 + \lambda_2 \circ \vec{\eta}_2$. Wobec tego równość (11.9.1) uzyskamy po wykonaniu na macierzy A związanej z rozważanym wyznacznikiem operacji elementarnych $w_1 - \lambda_1 \cdot w_2$ i $w_1 - \lambda_2 \cdot w_3$ i powołaniu się na Stwierdzenie 4.6. Na odwrót. Załóżmy, że zachodzi równość (11.9.1). Jeżeli $\vec{\eta}_1 \parallel \vec{\eta}_2$, to $l_1 \parallel l_2$, więc proste l_1 i l_2 są współpłaszczyznowe. Jeśli natomiast $\vec{\eta}_1 \nparallel \vec{\eta}_2$, to $r(A) = 2$, skąd $\vec{v} = \mu_1 \circ \vec{\eta}_1 + \mu_2 \circ \vec{\eta}_2$ dla pewnych $\mu_1, \mu_2 \in \mathbb{R}$. Stąd oraz na mocy określenia wektorów $\vec{v}$, $\vec{\eta}_1$ i $\vec{\eta}_2$ otrzymujemy, że istnieją $\xi_1, \xi_2 \in \mathbb{R}$ takie, że $[x_1,y_1,z_1] + \xi_1 \circ [a_1,b_1,c_1] =$

$[x_2,y_2,z_2]+\xi_2\circ[a_2,b_2,c_2]$. Wobec tego proste l_1 i l_2 przecinają się w pewnym punkcie P_0 i w konsekwencji są współpłaszczyznowe.

Bezpośrednią konsekwencją dowodu punktu (iii) Stwierdzenia 11.22 jest następujący

Wniosek 11.7. Nierównoległe proste l_1 oraz l_2 opisane równaniami parametrycznymi $[x,y,z]=[x_1,y_1,z_1]+\lambda\circ[a_1,b_1,c_1]$ i $[x,y,z]=[x_2,y_2,z_2]+\lambda\circ[a_2,b_2,c_2]$ przecinają się wtedy i tylko wtedy, gdy zachodzi równość (11.9.1).

Definicja 11.16. Kątem między prostymi l_1 i l_2 opisanymi równaniami parametrycznymi $[x,y,z]=[x_1,y_1,z_1]+\lambda\circ[a_1,b_1,c_1]$ i $[x,y,z]=[x_2,y_2,z_2]+\lambda\circ[a_2,b_2,c_2]$ nazywamy kąt ostry między wektorami równoległymi do tych prostych.

Stwierdzenie 11.23. Cosinus kąta ostrego φ między prostymi l_1 i l_2 opisanymi równaniami parametrycznymi $[x,y,z]=[x_1,y_1,z_1]+\lambda\circ[a_1,b_1,c_1]$ i $[x,y,z]=[x_2,y_2,z_2]+\lambda\circ[a_2,b_2,c_2]$ wyraża się wzorem:

$$\cos\varphi=\frac{|a_1a_2+b_1b_2+c_1c_2|}{\sqrt{a_1^2+b_1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}. \tag{11.9.2}$$

Dowód. Niech $\psi=\big|\measuredangle(\vec{\eta}_1,\vec{\eta}_2)\big|$, gdzie $\vec{\eta}_1=[a_1,b_1,c_1]$ i $\vec{\eta}_2=[a_2,b_2,c_2]$. Wtedy:

$$\varphi=\begin{cases}\psi & \text{, gdy } 0<\psi\leq\frac{\pi}{2}\\ \pi-\psi & \text{, gdy } \frac{\pi}{2}<\psi<\pi\end{cases}. \tag{11.9.3}$$

Jeśli więc $\psi\in\big(0,\frac{\pi}{2}\big)$, to $\cos\psi=\cos\varphi$. Dla $\psi\in\big(\frac{\pi}{2},\pi\big)$ otrzymujemy natomiast, że $\cos\psi=\cos(\pi-\varphi)=-\cos\varphi$. Ponieważ $\cos\varphi\geq 0$, to $\cos\varphi=|\cos\psi|$. Ponadto:

$$\cos\psi=\frac{a_1a_2+b_1b_2+c_1c_2}{\sqrt{a_1^2+b1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}$$

na mocy Wniosku 11.1, oraz mianownik powyższego ułamka jest zawsze dodatni. Zatem:

$$\cos\varphi=|\cos\psi|=\frac{|a_1a_2+b_1b_2+c_1c_2|}{\sqrt{a_1^2+b_1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}.$$

11.10 Wzajemne położenie prostej i płaszczyzny

Stwierdzenie 11.24. Niech l będzie prostą opisaną równaniem parametrycznym danym w (11.8.2) i niech Π będzie płaszczyzną opisaną równaniem ogólnym (11.5.1). Wówczas:

(i) $l \parallel \Pi$ wtedy i tylko wtedy, gdy $Aa+Bb+Cc=0$;
(ii) $l \perp \Pi$ wtedy i tylko wtedy, gdy $[A,B,C] = \lambda \circ [a,b,c]$ dla pewnego $\lambda \in \mathbb{R}$.

Dowód. Oznaczmy $\vec{\Gamma} = [A,B,C]$ i $\vec{\eta} = [a,b,c]$. Ponieważ $\vec{\Gamma} \perp \Pi$ oraz $\vec{\eta} \parallel l$, to:

(i). $l \parallel \Pi$ wtedy i tylko wtedy, gdy $\vec{\Gamma} \perp \vec{\eta}$, czyli gdy $Aa+Bb+Cc = \langle \vec{\Gamma} | \vec{\eta} \rangle = 0$;

(ii). $l \perp \Pi$ wtedy i tylko wtedy, gdy $\vec{\Gamma} \parallel \vec{\eta}$, co oznacza istnienie takiej $\lambda \in \mathbb{R}$, że $\vec{\Gamma} = \lambda \circ \vec{\eta}$.

Definicja 11.17. Kątem nachylenia prostej do płaszczyzny nazywamy kąt ostry, jaki tworzy prosta wraz ze swoim rzutem na płaszczyznę.

Stwierdzenie 11.25. Niech φ będzie kątem nachylenia prostej l opisanej równaniem parametrycznym (11.8.2) do płaszczyzny Π opisanej równaniem ogólnym (11.5.1). Wówczas:

$$\sin\varphi = \frac{|Aa+Ba+Cc|}{\sqrt{A^2+B^2+C^2}\cdot\sqrt{a^2+b^2+c^2}}. \tag{11.10.1}$$

Dowód. Niech $\vec{\Gamma} = [A,B,C]$ i niech $\vec{\eta} = [a,b,c]$. Wtedy $\measuredangle(\vec{\Gamma},\vec{\eta}) = \frac{\pi}{2} - \varphi$, skąd $\sin\varphi = \cos\measuredangle(\vec{\Gamma},\vec{\eta})$. Teza jest więc konsekwencją Wniosku 11.1.

Wniosek 11.8. Prosta l opisana równaniem parametrycznym (11.8.2) leży na płaszczyźnie Π opisanej równaniem ogólnym (11.5.1) wtedy i tylko wtedy, gdy $Aa+Bb+Cc=0$ oraz $Ax_0+By_0+Cz_0+D=0$.

Dowód. Prosta l leży na płaszczyźnie Π wtedy i tylko wtedy, gdy jej kąt φ nachylenia do tej płaszczyzny wynosi 0 (co oznacza, że $l \parallel \Pi$) oraz punkt (x_0,y_0,z_0), przez który przechodzi prosta l, należy do płaszczyzny Π. Drugi spośród wymienionych warunków oznacza, że $Ax_0+By_0+Cz_0+D=0$, zaś pierwszy w świetle wzoru (11.10.1) jest równoważny równości $Aa+Bb+Cc=0$.

Stwierdzenie 11.26. Niech:

$$\begin{cases} A_0x+B_0y+C_0z+D_0=0 \\ Ax+By+Cz+D=0 \end{cases} \tag{11.10.2}$$

będzie postacią krawędziową prostej l. Wówczas równanie:

$$A_0x+B_0y+C_0z+D_0+t(Ax+By+Cz+D)=0 \tag{11.10.3}$$

z rzeczywistym parametrem t opisuje pęk wszystkich płaszczyzn Π_t wyznaczonych przez prostą l oprócz płaszczyzny Π opisanej równaniem ogólnym $Ax+By+Cz+D=0$.

Dowód. Niech Π_0 oznacza płaszczyznę opisaną równaniem ogólnym $A_0x+B_0y+C_0z+D_0=0$, zaś Π' – płaszczyznę opisaną równaniem ogólnym $A'x+B'y+C'z+$

$D' = 0$ należącą do pęku płaszczyzn wspomnianego w powyższym stwierdzeniu. Wtedy $l = \Pi_0 \cap \Pi'$, więc układ równań (11.10.2) równoważny jest układowi:

$$\begin{cases} A_0x + B_0y + C_0z + D_0 = 0 \\ A'x + B'y + C'z + D' = 0 \end{cases}. \tag{11.10.4}$$

Zatem drugie równanie powyższego układu jest pewną kombinacją liniową równań układu (11.10.2). Istnieją więc $\lambda_0, \lambda \in \mathbb{R}$ takie, że $[A', B', C', D'] = \lambda_0 \circ [A_0, B_0, C_0, D_0] + \lambda \circ [A, B, C, D]$. Jeżeli $\lambda_0 = 0$, to równanie $A'x + B'y + C'z + D' = 0$ równoważne jest równaniu $Ax + By + Cz + D = 0$ i w konsekwencji opisuje ono płaszczyznę Π nienależącą do rozważanego pęku płaszczyzn, sprzeczność. Wobec tego $\lambda_0 \neq 0$, skąd $\lambda_0^{-1} \circ [A', B', C', D'] = [A_0, B_0, C_0, D_0] + (\lambda \cdot \lambda_0^{-1}) \circ [A, B, C, D]$. Ponadto równanie $A'x + B'y + C'z + D' = 0$ równoważne jest równaniu $\lambda_0^{-1}A'x + \lambda_0^{-1}B'y + \lambda_0^{-1}C'z + \lambda_0^{-1}D' = 0$, więc dla $t = \lambda \cdot \lambda_0^{-1}$ otrzymujemy, że równanie $A_0x + B_0y + C_0z + D_0 + t(Ax + By + Cz + D) = 0$ opisuje płaszczyznę Π'. W ten sposób pokazaliśmy, że równanie dowolnej płaszczyzny Π' należącej do pęku płaszczyzn wyznaczonego przez prostą l, różnych od płaszczyzny Π, równoważne jest równaniu (11.10.3), w którym $t = \lambda \cdot \lambda_0^{-1}$ czyli, że $\Pi' = \Pi_t$ dla $t = \lambda \cdot \lambda_0^{-1}$.

Na odwrót. Weźmy dowolne $t \in \mathbb{R}$. Wykonanie operacji elementarnej $r_1 - t \cdot r_2$ na układzie równań:

$$\begin{cases} A_0x + B_0y + C_0z + D_0 + t(Ax + By + Cz + D) = 0 \\ Ax + By + Cz + D = 0 \end{cases}$$

sprowadza go do równoważnej postaci (11.10.2). Ponadto równanie (11.10.3) równoważne jest równaniu $(A_0 + tA)x + (B_0 + tB)y + (C_0 + tC)z + (D_0 + tD) = 0$, które jest równaniem ogólnym pewnej płaszczyzny Π_t, gdyż $(A_0 + tA)^2 + (B_0 + tB)^2 + (C_0 + tC)^2 > 0$ jako, że równania układu (11.10.2) opisują dwie przecinające się płaszczyzny (por. Stwierdzenie 11.16). Zatem płaszczyzna Π_t opisana równaniem (11.10.3) należy do pęku płaszczyzn wyznaczonych przez prostą l różnych od płaszczyzny Π.

11.11 Punkt przecięcia prostej z płaszczyzną

Załóżmy, że prosta l opisana równaniem parametrycznym (11.8.2) przecina się w pewnym punkcie P z płaszczyzną Π opisaną równaniem ogólnym (11.5.1). Pokażemy w jaki sposób wyznaczyć współrzędne x_P, y_P i z_P punktu P. Ponieważ $P \in l$, to istnieje $\lambda \in \mathbb{R}$ taka, że współrzędne x_P, y_P i z_P punktu P spełniają układ warunków:

$$\begin{cases} x_P = x_0 + \lambda a \\ y_P = y_0 + \lambda b \\ z_P = z_0 + \lambda c \end{cases}. \tag{11.11.1}$$

Ponadto $P \in \Pi$, więc współrzędne x_P, y_P i z_P punktu P spełniają ogólne równanie płaszczyzny Π. Mamy więc:

$$A(x_0 + \lambda a) + B(y_0 + \lambda b) + C(z_0 + \lambda c) + D = 0. \tag{11.11.2}$$

Ponieważ prosta l przecina płaszczyznę Π, to ze Stwierdzenia 11.24 wynika, że $Aa + Bb + Cc \neq 0$. Stąd oraz na mocy (11.11.2) uzyskujemy, że:

$$\lambda = \frac{Ax_0 + By_0 + Cz_0 + D}{Aa + Bb + Cc}. \tag{11.11.3}$$

Podstawiając (11.11.3) do (11.11.1) otrzymujemy szukane współrzędne punktu P.

Jeżeli prosta l dana jest w postaci krawędziowej (11.8.12), to aby znaleźć współrzędne punktu $P = (x_P, y_P, z_P)$ przecięcia prostej l z płaszczyzną Π wystarczy rozwiązać układ równań:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \\ Ax + By + Cz + D = 0 \end{cases}. \tag{11.11.4}$$

Przykład 11.7. Wyznaczymy punkt przecięcia $P_0 = (x_0, y_0, z_0)$ prostej l opisanej równaniami kierunkowymi $\frac{x-1}{3} = \frac{y+1}{2} = \frac{z}{4}$ z płaszczyzną Π, określoną równaniem ogólnym $2x - 3y + 5z - 1 = 0$. Ponieważ $2 \cdot 3 + (-3) \cdot 2 + 5 \cdot 4 = 20 \neq 0$, to z punktu (i) Stwierdzenia 11.24 wynika, że możliwe jest wyznaczenie punktu P_0. W tym celu, w oparciu o równania kierunkowe prostej l, wyznaczymy najpierw równania parametryczne współrzędnych punktów tej prostej. Każdemu punktowi $P = (x, y, z)$ prostej l wzajemnie jednoznacznie odpowiada liczba rzeczywista λ taka, że $\frac{x-1}{3} = \lambda$, $\frac{y+1}{2} = \lambda$ i $\frac{z}{4} = \lambda$, skąd otrzymujemy, że:

$$\begin{cases} x = 1 + 3\lambda \\ y = -1 + 2\lambda \\ z = 4\lambda \end{cases}, \tag{11.11.5}$$

gdzie $\lambda \in \mathbb{R}$. Ponieważ $\{P_0\} = l \cap \Pi$, to podstawiając (11.11.5) do równania płaszczyzny Π wyznaczymy wartość λ_0 parametru λ, dla której $x = x_0$, $y = y_0$ i $z = z_0$ i tym samym znajdziemy szukany punkt P_0. Mamy:

$$2(1 + 3\lambda) - 3(-1 + 2\lambda) + 5(0 + 4\lambda) - 1 = 0,$$

skąd $20\lambda = -4$, czyli $\lambda = -\frac{1}{5}$. Wobec tego szukaną wartością parametru λ jest $\lambda_0 = -\frac{1}{5}$. Zatem $x_0 = 1+3\lambda_0 = \frac{2}{5}$, $y_0 = -1+2\lambda_0 = -\frac{7}{5}$ oraz $z_0 = 4\lambda_0 = -\frac{4}{5}$. Ostatecznie otrzymujemy więc, że $P_0 = \left(\frac{2}{5}, -\frac{7}{5}, -\frac{4}{5}\right)$.

Przykład 11.8. Wyznaczymy odległość d punktu $P_0 = (-2,1,4)$ od prostej $l : \frac{x+3}{4} = \frac{y-1}{-2} = \frac{z-2}{3}$. Zauważmy, że odległość d równa jest odległości między punktem P_0 a punktem przecięcia $P_1 = (x_1, y_1, z_1)$ prostej l z prostopadłą do niej płaszczyzną Π przechodzącą przez punkt P_0. W szczególności Π jest płaszczyzną przechodzącą przez punkt P_0 prostopadłą do wektora kierunkowego prostej l, czyli do wektora $\vec{\eta} = (4,-2,3)$. Jest więc ona opisana równaniem $4(x+2)-2(y-1)+3(z-4) = 0$. Otrzymujemy stąd równanie ogólne płaszczyzny Π:

$$4x - 2y + 3z - 2 = 0. \tag{11.11.6}$$

Analogicznie jak w Przykładzie 11.7 wyznaczymy teraz współrzędne x_1, y_1 i z_1 punktu P_1. Z równań kierunkowych prostej l uzyskujemy równanie parametryczne tej prostej:

$$[x,y,z] = [-3+4\lambda, 1-2\lambda, 2+3\lambda]. \tag{11.11.7}$$

Z (11.11.6) i (11.11.7) wynika więc, że $4(-3+4\lambda)-2(1-2\lambda)+3(2+3\lambda)-2=0$. Zatem $29\lambda = 10$, skąd $\lambda = \frac{10}{29}$. Ponieważ P_1 jest punktem prostej l opisanej równaniem (11.11.7), to $[x_1,y_1,z_1] = \left[-3+4\cdot\frac{10}{29}, 1-2\cdot\frac{10}{29}, 2+3\cdot\frac{10}{29}\right] = \left[-\frac{47}{29}, \frac{9}{29}, \frac{88}{29}\right]$. Stąd $P_1 = \left(-\frac{47}{29}, \frac{9}{29}, \frac{88}{29}\right)$. Wobec tego:

$$d = \left\|\overrightarrow{P_0P_1}\right\| = \sqrt{\left(-\frac{47}{29}+2\right)^2 + \left(\frac{9}{29}-1\right)^2 + \left(\frac{88}{29}-4\right)^2} =$$

$$\sqrt{\frac{11^2+20^2+28^2}{29^2}} = \frac{\sqrt{121+400+784}}{29} = \frac{\sqrt{1305}}{29}.$$

Rozdział 12

Elementy geometrii analitycznej na płaszczyźnie

W oparciu o zaprezentowaną w poprzednim rozdziale teorię związaną z geometrią analityczną w przestrzeni trójwymiarowej, przeanalizujemy teraz analogiczne zagadnienia w przypadku dwuwymiarowym. W niektórych przypadkach będą one przypomnieniem wiedzy zdobytej w szkole średniej, w innych zaś jej uzupełnieniem.

Jak zostało wspomniane w Uwadze 11.3, w przestrzeni dwuwymiarowej nie określa się iloczynu wektorowego i w konsekwencji również iloczynu mieszanego. W przypadku dwuwymiarowym definiuje się natomiast iloczyn skalarny. Jest on określony wzorem (11.2.1), przy czym długość wektora określa się analogicznie jak w (11.1.1). Wynika stąd, że pozostają w mocy wszystkie udowodnione wcześniej własności długości wektora oraz iloczynu skalarnego. W szczególności zachodzi więc wzór analogiczny do (11.2.2) (por. Uwaga 11.1).

12.1 Równania prostej na płaszczyźnie $\mathbb{R}^2$

12.1.1 Równanie ogólne prostej

Równanie ogólne prostej na płaszczyźnie:

$$Ax + By + C = 0, \text{ gdzie } A^2 + B^2 > 0 \tag{12.1.1}$$

wyprowadza się analogicznie jak równanie ogólne płaszczyzny w przestrzeni. Rozważamy mianowicie dowolną prostą l, jej dowolny punkt $P_0 = (x_0, y_0)$ oraz dowolny niezerowy wektor $\vec{\eta} = [A, B]$ prostopadły do l, a następnie zauważamy, że punkt $P = (x, y)$ należy do prostej l wtedy i tylko wtedy, gdy $\overrightarrow{P_0P} \perp \vec{\eta}$ i korzystamy z własności iloczynu skalarnego; w szczególności uzyskujemy, iż $C = -(Ax_0 + By_0)$. Analogicznie uzasadnia się też, że równanie (12.1.1) przy dowolnie ustalonych $A, B, C \in \mathbb{R}$ spełniających warunek $A^2 + B^2 > 0$ opisuje pewną prostą. Ponadto $\vec{\eta} = [A, B]$ jest wektorem normalnym prostej l.

12.1.2 Równanie prostej przechodzącej przez punkt i prostopadłej do wektora

Na mocy powyższych uwag otrzymujemy natychmiast, że dla dowolnych liczb rzeczywistych x_0, y_0, A i B takich, że $A^2+B^2>0$, równanie:

$$A(x-x_0)+B(y-y_0)=0 \tag{12.1.2}$$

opisuje prostą przechodzącą przez punkt $P_0=(x_0,y_0)$ prostopadłą do niezerowego wektora $\vec{\eta}=[A,B]$.

12.1.3 Wzajemne położenie prostych

Analogicznie jak Stwierdzenie 11.16 dowodzi się następujące

Stwierdzenie 12.1. Niech l_1 oraz l_2 będą prostymi opisanymi odpowiednio równaniami ogólnymi $A_1x+B_1y+C_1=0$ i $A_2x+B_2y+C_2=0$. Wówczas:

(i) $l_1 \parallel l_2$ wtedy i tylko wtedy, gdy $[A_2,B_2]=\lambda\circ[A_1,B_1]$ dla pewnego $\lambda\in\mathbb{R}$;
(ii) $l_1\perp l_2$ wtedy i tylko wtedy, gdy $\big\langle [A_1,B_1]\big|[A_2,B_2]\big\rangle=0$.

Uwaga 12.1. Ponieważ warunek istnienia liczby rzeczywistej λ takiej, że $[A_2,B_2]=\lambda\circ[A_1,B_1]$ równoważne jest liniowej zależności wektorów $[A_1,B_1]$ i $[A_2,B_2]$ w przestrzeni liniowej $\mathbb{R}^2$, to można zastąpić go równoważnym warunkiem:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}=0.$$

Bezpośrednią konsekwencją Stwierdzenia 12.1 oraz Uwagi 12.1 jest następujący

Wniosek 12.1. Proste l_1 i l_2 opisane odpowiednio równaniami ogólnymi $A_1x+B_1y+C_1=0$ i $A_2x+B_2y+C_2=0$ przecinają się w jednym punkcie wtedy i tylko wtedy, gdy:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}\neq 0. \tag{12.1.3}$$

12.1.4 Pęk prostych

Definicja 12.1. Właściwym pękiem prostych wyznaczonym przez punkt P_0 nazywamy zbiór wszystkich prostych na płaszczyźnie przechodzących punkt P_0.

Niewłaściwym pękiem prostych wyznaczonym przez prostą l nazywamy zbiór wszystkich prostych na płaszczyźnie równoległych do prostej l.

Stwierdzenie 12.2. Niech l_1 oraz l_2 będą nierównoległymi prostymi opisanymi odpowiednio równaniami ogólnymi $A_1x+B_1y+C_1=0$ i $A_2x+B_2y+C_2=0$. Równanie:

$$\lambda_1(A_1x+B_1y+C_1)+\lambda_2(A_2x+B_2y+C_2)=0, \tag{12.1.4}$$

w którym λ_1 i λ_2 są rzeczywistymi parametrami spełniającymi warunek $\lambda_1^2+\lambda_2^2>0$, opisuje wówczas właściwy pęk prostych wyznaczony przez punkt przecięcia prostych l_1 oraz l_2.

Dowód. Ponieważ $l_1 \nparallel l_2$, to istnieje punkt $P_0=(x_0,y_0)$ taki, że $\{P_0\}=l_1\cap l_2$. W szczególności:

$$\begin{cases} A_1x_0+B_1y_0+C_1=0 \\ A_2x_0+B_2y_0+C_2=0 \end{cases}. \tag{12.1.5}$$

Zauważmy najpierw, że przy dowolnie ustalonych $\lambda_1,\lambda_2\in\mathbb{R}$ takich, że $\lambda_1^2+\lambda_2^2>0$, równanie (12.1.4) możemy zapisać w postaci:

$$(\lambda_1A_1+\lambda_2A_2)x+(\lambda_1B_1+\lambda_2B_2)y+(\lambda_1C_1+\lambda_2C_2)=0. \tag{12.1.6}$$

Ponadto $(\lambda_1A_1+\lambda_2A_2)^2+(\lambda_1B_1+\lambda_2B_2)^2>0$, bo $\lambda_1^2+\lambda_2^2>0$ i wektory $[A_1,B_1]$ oraz $[A_2,B_2]$ są liniowo niezależne (por. Stwierdzenie 12.1 i Uwaga 12.1). Stąd oraz na mocy (12.1.5), równanie (12.1.6) opisuje prostą przechodzącą przez punkt P_0. W ten sposób wykazaliśmy, że dla dowolnie ustalonych $\lambda_1,\lambda_2\in\mathbb{R}$ spełniających warunek $\lambda_1^2+\lambda_2^2>0$, równanie (12.1.4) opisuje prostą przechodzącą przez punkt przecięcia prostych l_1 i l_2.

Na odwrót. Przypuśćmy, że prosta l opisana równaniem ogólnym $Ax+By+C=0$ należy do pęku prostych wyznaczonego przez punkt P_0. Wtedy $Ax_0+By_0+C=0$, więc układ równań:

$$\begin{cases} A_1x+B_1y=-C_1 \\ A_2x+B_2y=-C2 \end{cases},$$

którego jedynym rozwiązaniem jest (x_0,y_0), równoważny jest układowi równań:

$$\begin{cases} A_1x+B_1y=-C_1 \\ A_2x+B_2y=-C2 \\ Ax+By=-C \end{cases}.$$

Zatem ostatnie równanie powyższego układu jest nietrywialną kombinacją liniową dwóch pozostałych równań. Istnieją więc $\lambda_1,\lambda_2\in\mathbb{R}$ takie, że $\lambda_1^2+\lambda_2^2>0$ oraz $A=\lambda_1\circ A_1+\lambda_2\circ A_2$, $B=\lambda_1\circ B_1+\lambda_2\circ B_2$ i $C=\lambda_1\circ C_1+\lambda_2\circ C_2$, co oznacza, że prosta l opisana jest równaniem (12.1.4).

Stwierdzenie 12.3. Niech l_1, l_2 i l_3 będą parami różnymi prostymi, które opisane są odpowiednio równaniami ogólnymi $A_1x+B_1y+C_1=0$, $A_2x+B_2y+C_2=0$ i $A_3x+B_3y+C_3=0$.

(i) Jeżeli proste l_1, l_2 i l_3 przecinają się w jednym punkcie, to:

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0. \tag{12.1.7}$$

(ii) Jeżeli dwie spośród prostych l_1, l_2 i l_3 przecinają się w jednym punkcie, to równość (12.1.7) implikuje, że wszystkie te proste przecinają się w jednym punkcie.

Dowód. (i). Z przyjętego założenia wynika, że prosta l_3 należy do pęku prostych wyznaczonego przez punkt przecięcia prostych l_1 i l_2. Stwierdzenie 12.2 implikuje więc istnienie takich $\lambda_1, \lambda_2 \in \mathbb{R}$, że $\lambda_1^2 + \lambda_2^2 > 0$ oraz $A = \lambda_1 A_1 + \lambda_2 A_2$, $B = \lambda_1 B_1 + \lambda_2 B_2$ i $C = \lambda_1 C_1 + \lambda_2 C_2$. Zatem, po wykonaniu operacji elementarnych $w_3 - \lambda_1 w_1$ i $w_3 - \lambda_2 w_2$ na macierzy związanej z wyznacznikiem zapisanym po lewej stronie równości (12.1.7) uzyskujemy, że:

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = \begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

(ii). Bez utraty ogólności możemy założyć, że $P_0 = (x_0, y_0)$ jest jedynym punktem wspólnym prostych l_1 i l_2. Na mocy Wniosku 12.1 otrzymujemy wówczas, że wektory $[A_1, B_1]$ i $[A_2, B_2]$ przestrzeni liniowej $\mathbb{R}^2$ są liniowo niezależne. Zatem liniowo niezależne są również wektory $[A_1, B_1, C_1]$ i $[A_2, B_2, C_2]$ przestrzeni liniowej $\mathbb{R}^3$. Jeśli więc zachodzi równość (12.1.7), to istnieją $\lambda_1, \lambda_2 \in \mathbb{R}$ takie, że $\lambda_1^2 + \lambda_2^2 > 0$ oraz $[A_3, B_3, C_3] = \lambda_1 \circ [A_1, B_1, C_1] + \lambda_2 \circ [A_2, B_2, C_2]$. Zatem $A_3 = \lambda_1 A_1 + \lambda_2 A_2$, $B_3 = \lambda_1 B_1 + \lambda_2 B_2$ i $C_3 = \lambda_1 C_1 + \lambda_2 C_2$, skąd $A_3 x_0 + B_3 y_0 + C_3 = \lambda_1 (A_1 x_0 + B_1 y_0 + C_1) + \lambda_2 (A_2 x_0 + B_2 y_0 + C_2) = 0 + 0 = 0$. Zatem prosta l_3 przechodzi przez punkt P_0. Ponadto $l_3 \neq l_1$ i $l_3 \neq l_2$, więc P_0 jest jedynym punktem wspólnym prostych l_1, l_2 i l_3.

Stwierdzenie 12.4. Jeżeli φ jest kątem ostrym między nierównoległymi prostymi l_1 i l_2 opisanymi odpowiednio równaniami ogólnymi $A_1 x + B_1 y + C_1 = 0$ oraz $A_2 x + B_2 y + C_2 = 0$, to:

$$\cos\varphi = \frac{|A_1 A_2 + B_1 B_2|}{\sqrt{A_1^2 + B_1^2} \cdot \sqrt{A_2^2 + B_2^2}} \text{ i } \sin\varphi = \frac{|A_1 B_2 - A_2 B_1|}{\sqrt{A_1^2 + B_1^2} \cdot \sqrt{A_2^2 + B_2^2}}.$$

Dowód. Niech $\psi = \left|\measuredangle(\vec{\eta}_1, \vec{\eta}_2)\right|$, gdzie $\vec{\eta}_1 = [A_1, B_1]$ i $\vec{\eta}_2 = [A_2, B_2]$. Ponieważ $\vec{\eta}_i \perp l_i$ dla $i \in \{1, 2\}$, to φ spełnia zależność opisaną w (11.9.2). Analogicznie jak w dowodzie Stwierdzenia 11.23 uzasadnia się więc równość $\cos\varphi = |\cos\psi|$. Adaptując wzór (11.2.3) do rozważanego przypadku dwuwymiarowego wnioskujemy więc, że:

$$\cos\varphi = |\cos\psi| = \frac{|A_1A_2 + B_1B_2|}{\sqrt{A_1^2+B_1^2}\cdot\sqrt{A_2^2+B_2^2}}.$$

Dalej, $\sin^2\varphi = 1-\cos^2\varphi = 1 - \frac{|A_1A_2+B_1B_2|^2}{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)} = \frac{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)-(A_1A_2+B_1B_2)^2}{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)} = \frac{A_1^2A_2^2+A_1^2B_2^2+A_2^2B_1^2+B_1^2B_2^2-A_1^2A_2^2-2A_1A_2B_1B_2-B_1^2B_2^2}{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)} = \frac{A_1^2B_2^2-2A_1A_2B_1B_2+A_2^2B_1^2}{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)} = \frac{(A_1B_2-A_2B_1)^2}{(A_1^2+B_1^2)\cdot(A_2^2+B_2^2)}$.
Ponadto $\varphi \in \left(0,\frac{\pi}{2}\right]$, więc $\sin\varphi \geq 0$ i w konsekwencji $\sin\varphi = |\sin\varphi| = \sqrt{\sin^2\varphi} = \frac{|A_1B_2-A_2B_1|}{\sqrt{A_1^2+B_1^2}\cdot\sqrt{A_2^2+B_2^2}}$.

Wniosek 12.2. Pole równoległoboku rozpiętego przez wektory $\vec{v} = [v_1, v_2]$ i $\vec{u} = [u_1, u_2]$ przestrzeni $\mathbb{R}^2$ wyraża się wzorem:

$$P = \left|\det\begin{bmatrix} v_1 & v_2 \\ u_1 & u_2 \end{bmatrix}\right|. \tag{12.1.8}$$

Dowód. Niech l oraz k będą prostymi prostopadłymi odpowiednio do wektorów $\vec{v}$ i $\vec{u}$ przechodzącymi przez punkt będący wspólnym początkiem tych wektorów. Wówczas miara kąta między prostymi l i k równa jest mierze kąta między wektorami $\vec{v}$ i $\vec{u}$. Oznaczmy ją przez φ. Ze Stwierdzenia 12.4 wynika wówczas, że $|v_1u_2 - v_2u_1| = \sqrt{v_1^2+v_1^2}\cdot\sqrt{u_2^2+u_2^2}\cdot\sin\varphi$. Ponadto $\det\begin{bmatrix} v_1 & v_2 \\ u_1 & u_2 \end{bmatrix} = v_1u_2 - v_2u_1$, $\|\vec{v}\| = \sqrt{v_1^2+v_1^2}$, $\|\vec{u}\| = \sqrt{u_1^2+u_1^2}$ oraz ze szkoły średniej wiadomo, że pole P rozważanego równoległoboku wyraża się wzorem $P = \|\vec{v}\|\cdot\|\vec{u}\|\cdot\sin\varphi$, więc wzór (12.1.8) jest prawdziwy.

Bezpośrednią konsekwencją powyższego wniosku jest następujący

Wniosek 12.3. Pole trójkąta rozpiętego przez wektory $\vec{v} = [v_1, v_2]$ i $\vec{u} = [u_1, u_2]$ przestrzeni $\mathbb{R}^2$ wyraża się wzorem:

$$P = \frac{1}{2}\cdot\left|\det\begin{bmatrix} v_1 & v_2 \\ u_1 & u_2 \end{bmatrix}\right|.$$

Analogicznie jak Stwierdzenie 11.15 dowodzi się następujące

Stwierdzenie 12.5. Odległość $d(P_0, l)$ punktu $P_0 = (x_0, y_0)$ od prostej l opisanej równaniem ogólnym $Ax + By + C = 0$ wyraża się wzorem:

$$d(P_0, l) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}}. \tag{12.1.9}$$

Uwaga 12.2. Niech l_1 i l_2 będą nierównoległymi prostymi opisanymi odpowiednio równaniami ogólnymi $A_1x + B_1y + C_1 = 0$ oraz $A_2x + B_2y + C_2 = 0$. Niech ponadto k_i będzie dwusieczną kąta o mierze φ_i dla $i \in \{1,2\}$. Rozważmy dowolne $i \in \{1,2\}$ oraz

dowolny punkt $P=(x,y,z)$ prostej k_i. Wtedy $d(P,l_1)=d(P,l_2)$, więc ze Stwierdzenia 12.5 wynika, że:

$$\frac{|A_1x+B_1y+C_1|}{\sqrt{A_1^2+B_1^2}}=\frac{|A_2x+B_2y+C_2|}{\sqrt{A_2^2+B_2^2}}.$$

Wobec tego, korzystając z powyższej równości możemy wyznaczyć równania dwusiecznych k_1 i k_2.

12.1.5 Równanie kierunkowe prostej

Rozważmy prostą l opisaną równaniem ogólnym (12.1), w którym $B\neq 0$. Równanie to jest wówczas równoważne równaniu:

$$y=ax+b, \tag{12.1.10}$$

w którym $b=-\frac{A}{B}$ i $b=-\frac{C}{B}$.

Definicja 12.2. Równanie (12.1.10) nazywamy równaniem kierunkowym prostej. Współczynnik a występujący w równaniu (12.1.10) określamy mianem współczynnika kierunkowego prostej.

Uwaga 12.3. Załóżmy, że prosta l opisana jest równaniem kierunkowym (12.1.10). Przecina ona oś O_x w punkcie $P_1=\left(-\frac{b}{a},0\right)$, zaś oś O_y w punkcie $P_2=(0,b)$. Niech α będzie miarą kąta jaki tworzy prosta l z osią O_x. Wtedy $\operatorname{tg}\alpha=\operatorname{sgn}(a)\cdot\frac{\|\overrightarrow{OP_2}\|}{\|\overrightarrow{OP_1}\|}=\operatorname{sgn}(a)\cdot\frac{|b|}{|\frac{b}{a}|}=\operatorname{sgn}(a)\cdot|a|=a$.

Stwierdzenie 12.6. Niech l_1 i l_2 będą prostymi opisanymi odpowiednio równaniami kierunkowymi $y=a_1x+b_1$ oraz $y=a_2x+b_2$. Wówczas:

(i) $l_1\parallel l_2$ wtedy i tylko wtedy, gdy $a_1=a_2$;
(ii) $l_1\perp l_2$ wtedy i tylko wtedy, gdy $a_1\cdot a_2=-1$.

Dowód. (i). Zauważmy, że $l_1\parallel l_2$ wtedy i tylko wtedy, gdy proste l_1 i l_2 są nachylone do osi O_x pod tym samym kątem α, co wobec Uwagi 12.3 oznacza, że $\operatorname{tg}\alpha=a_1$ i $\operatorname{tg}\alpha=a_2$, czyli $a_1=a_2$.

(ii). Równaniami ogólnymi prostych l_1 oraz l_2 są odpowiednio $-a_1x+y-b_1=0$ i $-a_2x+y-b_2=0$. Stąd oraz na mocy punktu (ii) Stwierdzenia 12.1 otrzymujemy, że $l_1\perp l_2$ wtedy i tylko wtedy, gdy $\langle[-a_1,1]\big|[-a_2,1]\rangle=0$, czyli gdy $a_1a_2+1=0$, co oznacza równość $a_1a_2=-1$.

Stwierdzenie 12.7. Jeżeli nieprostopadłe proste l_1 i l_2 opisane odpowiednio równaniami kierunkowymi $y=a_1x+b_1$ oraz $y=a_2x+b_2$ przecinają się w jednym punkcie

pod kątem ostrym φ, to:

$$\operatorname{tg}\varphi = \left|\frac{a_2 - a_1}{1 + a_1 a_2}\right|. \tag{12.1.11}$$

Dowód. Równaniami ogólnymi prostych l_1 oraz l_2 są odpowiednio $-a_1x + y - b_1 = 0$ i $-a_2x + y - b_2 = 0$. Stąd oraz na Stwierdzenia 12.4 otrzymujemy, że:

$$\cos\varphi = \frac{|a_1a_2 + 1|}{\sqrt{a_1^2+1}\cdot\sqrt{a_2^2+1}} \text{ i } \sin\varphi = \frac{|a_2 - a_1|}{\sqrt{a_1^2+1}\cdot\sqrt{a_2^2+1}}.$$

Zatem:

$$\operatorname{tg}\varphi = \frac{\sin\varphi}{\cos\varphi} = \frac{|a_2 - a_1|}{|1 + a_1a_2|} = \left|\frac{a_2 - a_1}{1 + a_1a_2}\right|.$$

Stwierdzenie 12.8. Dla dowolnej liczby rzeczywistej a, równanie:

$$y - y_0 = a(x - x_0) \tag{12.1.12}$$

opisuje prostą przechodząca przez punkt $P_0 = (x_0, y_0)$ nachyloną do osi O_x pod kątem $\alpha = \operatorname{arctg} a$.

Dowód. Łatwo zauważyć, że powyższe równanie równoważne jest równaniu $y = ax + (ax_0 + y_0)$ oraz że podstawiając w nim $x = x_0$ otrzymujemy $y = y_0$. Zatem (11.4.4) jest równaniem prostej l przechodzącej przez punkt P_0 nachylonej do osi O_x pod kątem $\alpha = \operatorname{arctg} a$.

Uwaga 12.4. Traktując współczynnik a równania (12.1.12) jak rzeczywisty parametr otrzymujemy równanie opisujące wszystkie proste przechodzące przez punkt $P_0 = (x_0, y_0)$, które nie są równoległe do osi O_y. Istotnie, ze Stwierdzenia 12.8 wynika, że dla każdego $a \in \mathbb{R}$ równanie (12.1.12) opisuje prostą przechodzącą przez punkt P_0, która nie jest równoległa do osi O_y. Ponadto jeśli l jest prostą, która nie jest równoległa do osi O_y i przechodzącą przez punkt P_0, to prosta l opisana jest równaniem kierunkowym $y = a'x + b'$, gdzie a' i b' są liczbami rzeczywistymi takimi, że $y_0 = a'x_0 + b'$. Zatem $b' = y_0 - a'x_0$, skąd otrzymujemy, że $y = a'(x - x_0) + y_0$ jest równaniem prostej l.

Ponadto prosta przechodząca przez punkt P_0 i równoległa do osi O_y opisana jest równaniem $x = x_0$.

12.1.6 Równanie prostej przechodzącej przez dwa punkty

Stwierdzenie 12.9. Prosta l przechodząca przez dwa punkty $P_1 = (x_1, y_1)$ i $P_2 = (x_2, y_2)$ opisana jest równaniem:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0, \tag{12.1.13}$$

które równoważne jest równaniu:

$$(y_2 - y_1)(x - x_1) = (x_2 - x_1)(y - y_1). \tag{12.1.14}$$

Dowód. Ponieważ:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} \begin{matrix} w_1 - w_2 \\ w_3 - w_2 \\ = \end{matrix} \begin{vmatrix} x - x_1 & y - y_1 & 0 \\ x_1 & y_1 & 1 \\ x_2 - x_1 & y_2 - y_1 & 0 \end{vmatrix} = - \begin{vmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix},$$

to:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0 \Longleftrightarrow (y_2 - y_1)(x - x_1) = (x_2 - x_1)(y - y_1).$$

Zatem równania (12.1.13) i (12.1.14) są równoważne. Ponieważ równanie (12.1.14) przekształca się do równoważnej postaci:

$$(y_2 - y_1)x + \big(-(x_2 - x_1)\big)y + \big(-(y_2 - y_1)x_1 + (x_2 - x_1)y_1\big) = 0$$

oraz $(y_2 - y_1)^2 + \big(-(x_2 - x_1)\big)^2 > 0$ (bo $P_2 \neq P_1$), to opisuje ono prostą. Bezpośrednie sprawdzenie pokazuje, że współrzędne punktów P_1 i P_2 spełniają równanie (12.1.14). Ostatecznie otrzymujemy więc, że równania (12.1.13) oraz (12.1.14) opisują prostą przechodzącą przez punkty P_1 i P_2.

Uwaga 12.5. Jeżeli $x_2 \neq x_1$, to równanie (12.1.14) można zapisać w postaci kierunkowej:

$$y = \frac{y_2 - y_1}{x_2 - x_1} x + \left(y_1 - x_1 \cdot \frac{y_2 - y_1}{x_2 - x_1} \right).$$

12.1.7 Równanie odcinkowe prostej

Równanie odcinkowe:

$$\frac{x}{a} + \frac{y}{b} = 1 \tag{12.1.15}$$

prostej przecinającej osie O_x i O_y w punktach $(a,0)$ oraz $(0,b)$ wyprowadza się w oparciu o równanie (12.1.13) w sposób analogiczny jak równanie odcinkowe płaszczyzny nieprzechodzącej przez początek układu współrzędnych (zob. s. 149).

Sprowadzając równanie (12.1.15) do postaci kierunkowej otrzymujemy równanie:

$$y = -\frac{b}{a}x + b.$$

12.1.8 Równanie normalne prostej

Niech l będzie prostą opisaną równaniem ogólnym (12.1.1), w którym $C \neq 0$. Z początku układu współrzędnych można wówczas poprowadzić niezerowy wektor $\vec{v}$, który jest prostopadły do prostej l. Niech $\gamma = \|\vec{v}\|$ i niech α oraz β będą miarami kątów, które tworzy wektor $\vec{v}$ odpowiednio z osiami O_x i O_y układu współrzędnych. Analogicznie jak przy okazji przeprowadzonych wcześniej rozważań na temat normalnego równania płaszczyzny (zob. (11.5.12)) uzasadnia się, że równanie:

$$x\cos\alpha + y\cos\beta - \gamma = 0, \tag{12.1.16}$$

zwane równaniem normalnym prostej, opisuje prostą l oraz że obustronne pomnożenie równania ogólnego (12.1.1) prostej l przez tzw. czynnik normujący:

$$\kappa = \frac{\varepsilon}{\sqrt{A^2 + B^2}},$$

gdzie:

$$\varepsilon = \begin{cases} 1 & \text{, gdy } C < 0 \\ -1 & \text{, gdy } C > 0 \end{cases},$$

sprowadza to równanie do postaci normalnej (12.1.16). Analogicznie uzasadnia się też, że $\cos^2\alpha + \cos^2\beta = 1$.

12.1.9 Równanie parametryczne prostej

Analogicznie jak w omówionym wcześniej przypadku trójwymiarowym uzasadnia się, że prosta przechodząca przez punkt $P_0 = (x_0, y_0)$ równoległa do niezerowego wektora $\vec{\eta} = [a, b]$ opisana jest równaniem parametrycznym:

$$[x, y] = [x_0, y_0] + \lambda \circ [a, b] \tag{12.1.17}$$

z rzeczywistym parametrem λ (por. (11.8.2)). Powyższe równanie równoważne jest układowi równań:

$$\begin{cases} x = x_0 + \lambda a \\ y = y_0 + \lambda b \end{cases}, \tag{12.1.18}$$

gdzie $\lambda \in \mathbb{R}$.

Załóżmy, że $a \neq 0$. Z (12.1.18) wynika wówczas, że dla dowolnie ustalonego punktu $P = (x,y)$ prostej l istnieje $\lambda \in \mathbb{R}$ taka, że $\frac{x-x_0}{a} = \lambda$ oraz $y - y_0 = \lambda b$. Jeśli $b \neq 0$, to $\frac{y-y_0}{b} = \lambda$ i w konsekwencji $\frac{x-x_0}{a} = \frac{y-y_0}{b}$, skąd $y = \frac{b}{a}x + \frac{b}{a} \cdot (y_0 - x_0)$ jest równaniem kierunkowym prostej l. Jeżeli $b = 0$, to $y = y_0$ jest równaniem kierunkowym prostej l. Otrzymujemy stąd następujący

Wniosek 12.4. Jeżeli prosta jest opisana równaniem parametrycznym (12.1.17), w którym $a \neq 0$, to współczynnikiem kierunkowym tej prostej jest $\frac{b}{a}$.

Przykład 12.1. Załóżmy, że promień światła poruszający się wzdłuż prostej l_1 opisanej równaniem $5x + 3y - 26 = 0$ pada na zwierciadło płaskie umieszczone prostopadle do płaszczyzny O_{xy} w taki sposób, że prosta O_x leży w jego płaszczyźnie. Wyznaczymy równanie ogólne prostej l_2 opisującej drogę promienia odbitego. Przypomnijmy, że zgodnie z fizycznym prawem odbicia światła, kąt odbicia promienia światła jest równy kątowi jego padania. Ponadto każdy z tych kątów zdefiniowany jest jako kąt między promieniem światła a normalną (czyli prostą prostopadłą) wystawioną w punkcie padania tego promienia. Niech β oznacza wspólną miarę tych kątów, zaś α_i – miarę kąta nachylenia prostej l_i do osi O_x, dla $i \in \{1,2\}$. Wówczas $\operatorname{tg}\alpha_1 = -\frac{5}{3}$, gdyż $y = -\frac{5}{3}x + 2$ jest równaniem kierunkowym prostej l_1. Ponadto β jest kątem ostrym, więc $\operatorname{tg}\alpha_2 = \operatorname{tg}\left(\frac{\pi}{2} - \beta\right) = \operatorname{ctg}\beta = -\operatorname{tg}\left(\frac{\pi}{2} + \beta\right) = -\operatorname{tg}\alpha_1 = \frac{5}{3}$. Wobec tego prosta l_2 opisana jest równaniem kierunkowym $y = \frac{5}{3}x + b$ dla pewnego $b \in \mathbb{R}$. W celu wyznaczenia liczby b znajdziemy najpierw punkt $P_0 = (x_0, 0)$ przecięcia prostej l_1 z prostą O_x, który jest również punktem wspólnym prostych l_1 i l_2. Mamy $0 = -\frac{5}{3}x_0 + 2$, więc $x_0 = \frac{6}{5}$. Zatem $0 = \frac{5}{3} \cdot \frac{6}{5} + b$, skąd $b = -2$. Wobec tego $y = \frac{5}{3}x - 2$ jest równaniem kierunkowym prostej l_2. Zatem $-5x + 3y + 2 = 0$ jest ogólnym równaniem prostej l_2.

Rozdział 13

Podstawowe metody algebry liniowej w kryptografii

W dwóch poprzednich rozdziałach Czytelnik mógł zaobserwować proste, konkretne zastosowanie kilku najbardziej podstawowych, abstrakcyjnych pojęć z zakresu algebry liniowej w geometrii. W ostatnim rozdziale niniejszego skryptu krótko zaprezentujemy, w jaki sposób arytmetyka modularna i rachunek macierzowy mogą zostać wykorzystane do szyfrowania i deszyfrowania wiadomości. W naszych kryptograficznych rozważaniach będziemy używali standardowego 26-literowego alfabetu łacińskiego (tzn. bez polskich znaków, zob. np. alfabet języka angielskiego). Kolejnym literom tego alfabetu możemy wzajemnie jednoznacznie przypisać w porządku rosnącym liczby całkowite z zakresu od 0 do 25. W ten sposób uzyskuje się tzw. numeryczne odpowiedniki tekstu, na których można wykonywać działania z wykorzystaniem arytmetyki pierścienia $\mathbb{Z}_{26}$. Zostanie to wyjaśnione przy okazji omówienia szyfru afinicznego oraz blokowego szyfru afinicznego.

W całym niniejszym rozdziale zamiast symboli $\oplus_{26}$ i $\odot_{26}$ będziemy używali standardowych oznaczeń dodawania i mnożenia, przy czym kropkę $\cdot$ będziemy często pomijać. W szczególności, znaczenie symboli $+$ i $\cdot$ będzie zależne od kontekstu.

13.1 Szyfr afiniczny

Stwierdzenie 13.1. Dla dowolnych $a, b \in \mathbb{Z}_{26}$ funkcja $f\colon \mathbb{Z}_{26} \to \mathbb{Z}_{26}$ określona wzorem:

$$f(x) = ax + b \text{ dla każdego } x \in \mathbb{Z}_{26}, \tag{13.1.1}$$

jest bijekcją wtedy i tylko wtedy, gdy $a \in \mathbb{Z}_{26}^*$.

Dowód. Rozważmy funkcje g i h przekształcające $\mathbb{Z}_{26}$ w $\mathbb{Z}_{26}$ określone za pomocą wzorów $g(x) = ax$ i $h(x) = x + b$ dla każdego $x \in \mathbb{Z}_{26}$. Wówczas $f = h \circ g$ oraz h jest bijekcją, więc $g = h^{-1} \circ f$. Zatem bijektywność funkcji f równoważna jest bijektywności funkcji g. Jeżeli $a \in \mathbb{Z}_{26}^*$, to funkcja $\rho\colon \mathbb{Z}_{26} \to \mathbb{Z}_{26}$ określona wzorem $\rho(x) = a^{-1}x$ dla każdego $x \in \mathbb{Z}_{26}$ jest funkcją odwrotną do g, więc g jest bijekcją. Załóżmy teraz, że funkcja g jest bijektywna. Istnieje wówczas $c \in \mathbb{Z}_{26}$ takie, że $1 = g(c)$. Oznacza to, iż $ac = 1$, czyli $a \in \mathbb{Z}_{26}^*$.

Definicja 13.1. Bijekcję postaci $(13.1.1)$ nazywamy afiniczną funkcją szyfrującą.

Przykład 13.1. Używając afinicznej funkcji szyfrującej f określonej wzorem $f(x) = 3x + 5$ zaszyfrujemy słowo algebra. Odnotujmy najpierw fakt, że ze Stwierdzenia

1.7 wynika, że $3 \in \mathbb{Z}_{26}^*$, gdyż $3 \in \mathbb{P}$ i $3 \nmid 26$. Stwierdzenie 13.1 implikuje więc, że f jest bijekcją i w konsekwencji f jest poprawnie określoną afiniczną funkcją szyfrującą. Numerycznym odpowiednikiem słowa algebra jest ciąg $(0,11,6,4,1,17,0)$, więc ciąg $\big(f(0),f(11),f(6),f(4),f(1),f(17),f(0)\big)$ jest numerycznym odpowiednikiem szukanego szyfrogramu (czyli zaszyfrowanego tekstu). Pamiętając o wykonywaniu wszystkich rachunków w pierścieniu $\mathbb{Z}_{26}$ otrzymujemy: $f(0) = 5$, $f(11) = 3 \cdot 11 + 5 = 7 + 5 = 12$, $f(6) = 23$, $f(4) = 17$, $f(1) = 8$ i $f(17) = 4$. Numerycznym odpowiednikiem szukanego szyfrogramu jest więc ciąg $(5,12,23,17,8,4,5)$. Otrzymujemy stąd szyfrogram FMXRIEF (szyfrogramy zapisuje się tradycyjnie wielkimi literami, zaś tekst jawny – małymi).

Uwaga 13.1. Deszyfrowanie tekstu tajnego (tj. szyfrogramu) uzyskanego za pomocą znanej afinicznej funkcji szyfrującej f polega na wyznaczeniu funkcji odwrotnej do f i zastosowaniu dla szyfrogramu i funkcji f^{-1} procedury opisanej w powyższym przykładzie. Jeśli funkcja f dana w $(13.1.1)$ jest afiniczną funkcją szyfrującą, to łatwo zauważyć, że $f^{-1}(x) = a^{-1}(x-b)$ dla każdego $x \in \mathbb{Z}_{26}$.

13.2 Blokowy szyfr afiniczny

Teoria macierzy uprawiana jest nie tylko nad ciałami. Można mianowicie rozważać macierze nad pierścieniem niebędącym ciałem (a więc w szczególności nad pierścieniem $\mathbb{Z}_{26}$). Dodawanie, odejmowanie oraz mnożenie takich macierzy oraz mnożenie macierzy z lewej strony przez elementy z pierścienia, określone są wówczas tak samo jak dotychczas. Analogicznie określa się także operacje elementarne na takich macierzach – jedyną różnicą jest konieczność zastąpienia niezerowego skalara elementem odwracalnym pierścienia przy operacji $(OM1)$. Przy uwzględnieniu tej modyfikacji, w mocy pozostają również definicja wyznacznika macierzy kwadratowej nad dowolnym pierścieniem oraz wszelkie poznane dotychczas własności wyznacznika, w tym Twierdzenie Cauchy'ego (zob. Twierdzenie 5.1). Dla macierzy kwadratowych nad dowolnym pierścieniem nieco ogólniejsze staje się sformułowane w Twierdzeniu 5.2 kryterium odwracalności takich macierzy. Podobnie jak wspomniane twierdzenie można bowiem udowodnić następujące

Twierdzenie 13.1. Macierz kwadratowa A nad pierścieniem R jest odwracalna wtedy i tylko wtedy, gdy $\det(A) \in R^*$.

W oparciu o powyższe twierdzenie oraz materiał wykraczający poza zakres tematyczny niniejszego skryptu dowodzi się poniższe stwierdzenie, będące naturalnym uogólnieniem Stwierdzenia 13.1.

Stwierdzenie 13.2. Niech R będzie pierścieniem, niech $n \in \mathbb{N}$ i niech:

$$R^n = \big\{[x_1\ x_2\ \ldots\ x_n]^T : x_1, x_2, \ldots, x_n \in R\big\}.$$

Niech ponadto $A \in \mathrm{M}_n(R)$ i niech $b \in R^n$. Wówczas funkcja $F\colon R^n \to R^n$ określona wzorem:

$$F(x) = Ax + b \text{ dla każdego } x \in R^n, \tag{13.2.1}$$

jest bijekcją wtedy i tylko wtedy, gdy A jest macierzą odwracalną.

Definicja 13.2. Blokową szyfrującą funkcją afiniczną nazywa się każdą bijekcję $F\colon {\mathbb{Z}_{26}}^n \to {\mathbb{Z}_{26}}^n$ postaci (13.2.1), gdzie n jest ustaloną liczbą naturalną.

Przykład 13.2. Niech:

$$A = \begin{bmatrix} 23 & 7 & 1 \\ 5 & 2 & 0 \\ 1 & 18 & 1 \end{bmatrix} \in \mathrm{M}_3(\mathbb{Z}_{26}) \text{ i } b = \begin{bmatrix} 5 \\ 0 \\ 19 \end{bmatrix} \in {\mathbb{Z}_{26}}^3. \tag{13.2.2}$$

Bezpośredni rachunek (np. z wykorzystaniem wzoru Sarrusa) pokazuje, że $\det(A) = 21$. Ponieważ $\mathrm{NWD}(21,26) = 1$, to ze Stwierdzenia 1.7 wynika, że $21 \in \mathbb{Z}_{26}^*$. Stąd oraz na mocy Twierdzenia 13.1 otrzymujemy, że macierz A jest odwracalna. Stwierdzenie 13.2 implikuje więc, że funkcja $F\colon {\mathbb{Z}_{26}}^3 \to {\mathbb{Z}_{26}}^3$ określona wzorem $F = Ax + b$ dla każdego $x \in {\mathbb{Z}_{26}}^3$ jest blokową szyfrującą funkcją afiniczną. Przy jej wykorzystaniu, zaszyfrujemy wiadomość Ala ma kota. Najpierw dzielimy ten tekst jawny na bloki długości trzy i zapisujemy numeryczny odpowiednik tak przygotowanego tekstu:

$$\begin{array}{ccc} ala & mak & ota \\ (0,11,0) & (12,0,10) & (14,19,0) \end{array}$$

(gdyby ostatni blok miał długość mniejszą niż trzy, tzn. składałby się z jednej lub dwóch liter, uzupełnilibyśmy go dowolnymi literami do bloku długości trzy). Następnie dla każdego $x \in \left\{[0\ 11\ 0]^T, [12\ 0\ 10]^T, [14\ 19\ 0]^T\right\}$ obliczamy $F(x)$, uzyskując numeryczny odpowiednik szyfrogramu i szukany szyfrogram:

$$\begin{array}{ccc} (4,22,9) & (5,8,15) & (8,4,11) \\ EWJ & FIP & IEL \end{array}.$$

Uwaga 13.2. Aby deszyfrować tekst tajny uzyskany przy wykorzystaniu znanej blokowej szyfrującej funkcji afinicznej F należy postępować analogicznie jak w Uwadze 13.1. W szczególności, jeżeli funkcja F opisana jest wzorem (13.2.1), to $F^{-1}(x) = A^{-1}(x-b)$ dla każdego $x \in {\mathbb{Z}_{26}}^n$, przy czym macierz A^{-1} wyznacza się dowolną spośród poznanych wcześniej metod zaadaptowanych do uwag poczynionych przed Twierdzeniem 13.1.

Bibliografia

Andruszkiewicz, R. R. (2005a). Wykłady z algebry liniowej i. Białystok: Wydawnictwo Uniwersytetu w Białymstoku.

Andruszkiewicz, R. R. (2005b). Wykłady z algebry ogólnej i. Białystok: Wydawnictwo Uniwersytetu w Białymstoku.

Andruszkiewicz, R. R. (2007). Wykłady z algebry liniowej ii. Białystok: Wydawnictwo Uniwersytetu w Białymstoku.

Góra, M. (n.d.). Algebra liniowa. Available at `https://home.agh.edu.pl/~gora/algebra/Wyklad12.pdf` (2021/01/23).

Kącki, E., i in. (1975). Geometria analityczna w zadaniach. Warszawa: PWN.

Leja, F. (1979). Funkcje zespolone. Warszawa: PWN.

Politechnika Białostocka