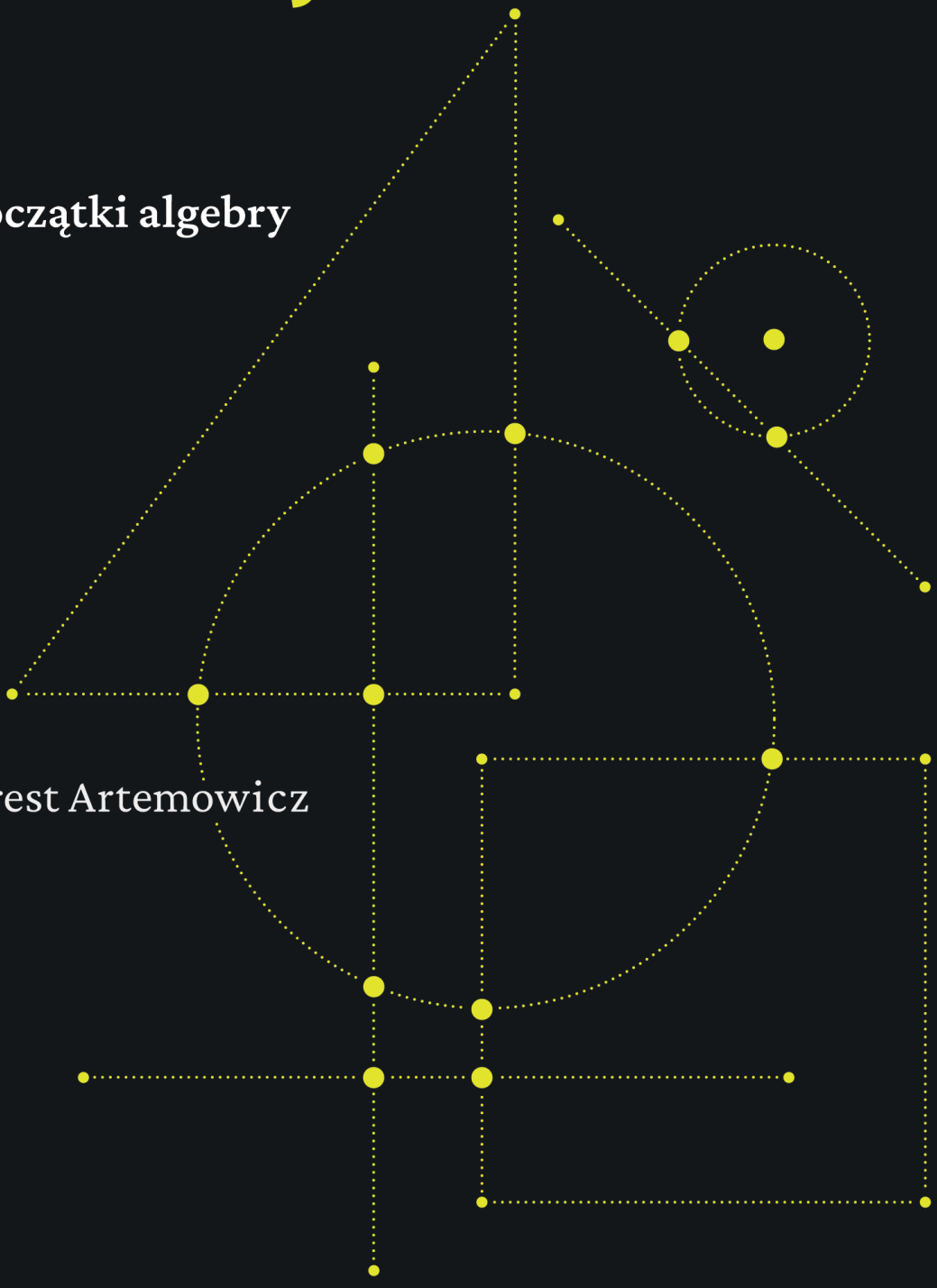
matematyka

Początki algebry



Orest Artemowicz

Kraków 2022

Politechnika Krakowska
im. Tadeusza Kościuszki

Początki algebry

matematyka

Początki algebry

Orest Artemowicz

Kraków 2022

Tekst został opublikowany w ramach projektu „Programowanie doskonałości – PK XXI 2.0. Program rozwoju Politechniki Krakowskiej na lata 2018-2022".
Dofinansowanie z Europejskiego Funduszu Społecznego: 18,048,774.96 PLN

Edycja online
eISBN 978-83-67188-35-7

20 ark. wyd.

Wydawnictwo PK, ul. Skarżyńskiego 1, 31-866 Kraków; 12 628 37 25, fax 12 628 37 60
wydawnictwo@pk.edu.pl
www.wydawnictwo.pk.edu.pl
Adres korespondencyjny: ul. Warszawska 24, 31-155 Kraków

Spis treści

MAMIE, z wdzięcznością

Przedmowa

Niniejszy podręcznik poświęcony podstawom współczesnej algebry jest w pierwszej kolejności adresowany do studentów kierunków informatycznych i matematycznych.

Algebra jest dziedziną matematyki o korzeniach starożytnych. Podstawy współczesnej algebry zostały zbudowane w drugiej połowie XIX w. i w pierwszej połowie XX w. Jeśli na początku XX w. algebra była jeszcze postrzegana jako bardzo abstrakcyjna dziedzina matematyki, to w epoce komputerowej znalazła wiele bardzo ważnych zastosowań (na przykład w informatyce czy w teorii informacji, w sposobach kodowania informacji w technologiach cyfrowych i, w szczególności, w kryptologii).

Podręcznik zawiera klasyczne podstawy struktur algebraicznych (półgrup, grup, pierścieni, ciał), które są niezbędne w modelowaniu wielu badanych własności obiektów przyrody czy działalności człowieka. Rozpatrywane pojęcia są ilustrowane licznymi przykładami. Ostatni rozdział zawiera informacje o strukturze ciał skończonych i faktoryzacji wielomianów nad nimi (co jest niezbędne w licznych zastosowaniach).

Jedna z merytorycznych metod nauczania bazuje na samodzielnym rozwiązywaniu zadań (co powoduje szybsze zrozumienie podstawowych idei przedmiotu), listy zadań są zamieszczone w każdym podrozdziale.

Używamy ogólnie przyjętych oznaczeń, których lista jest podana na końcu tej książki. Zaznaczmy, że w podręczniku często wykorzystujemy

dwa symbole: „■” zwykle oznacza część podrozdziału, uwagę lub definicję, a „□” koniec dowodu każdego twierdzenia matematycznego.

Proponowany materiał został ułożony na podstawie treści wykładów dla studentów kierunków „informatyka” i „matematyka” prowadzonych przez autora na Politechnice Krakowskiej i na Politechnice Śląskiej. W celu pogłębienia wiedzy proponujemy Czytelnikowi inne pozycje zawarte w bibliografii. Zakładamy, że Czytelnik zna elementarne podstawy liczb zespolonych, wielomianów
i macierzy.

Autor serdecznie dziękuje szanownym recenzentom: doktorowi hab. inż. Waldemarowi Hołubowskiemu, profesorowi Politechniki Śląskiej, oraz doktorowi hab. Andriyowi Panasyukowi, adiunktowi Uniwersytetu Warmińsko-Mazurskiego, za bardzo wnikliwe recenzje.

Kraków, czerwiec 2021 r. *Autor*

■ Dla ułatwienia studiowania materiału podręcznika przez Czytelnika umieszczamy też

Alfabet grecki

A, α – (alfa),	B, β – (beta),	Γ, γ – (gamma),
Δ, δ – (delta),	E, ε – (epsilon),	Z, ζ – (zeta),
H, η – (eta),	Θ, θ – (teta),	I, ι – (jota),
K, κ – (kappa),	Λ, λ – (lambda),	M, μ – (mi),
N, ν – (ni),	Ξ, ξ – (ksi),	O, o – (omikron),
Π, π – (pi),	P, ρ – (ro),	Σ, σ – (sigma),
T, τ – (tau),	Υ, υ – (ypsilon),	Φ, φ – (fi),
X, χ – (chi),	Ψ, ψ – (psi),	Ω, ω – (omega).

Rozdział 1

Wprowadzenie

1.1. Symbolika zbiorowa

Teoria zbiorów jako część matematyki zaczęła się od G. Cantora[1].

* * *

■ W 1872 r. G. Cantor zdefiniował pojęcie zbioru, według którego *zbiór jest dowolnym ogółem określonych obiektów naszej intuicji czy intelektu, które są rozróżnialne i rozpatrywane jako całość*. Istotne w tym pojęciu jest to, że ogół jest rozpatrywany jako jedna całość. Właśnie takie meritum tego pojęcia odzwierciedlają takie jego synonimy jak „rodzina", „szereg", „układ", „kolekcja", „kompania", „stado", „klasa" itd. Ponieważ definicja G. Cantora wyłącza z matematyki zbiory, przyroda elementów których jest „źle (czy może nie do końca precyzyjnie) określona" (na przykład zbiór twierdzeń, które będą udowodnione w przyszłości, czy zbiór typów komputerów, które będą stworzone przez najbliższe sto lat) oraz, oprócz tego, zakłada się wymaganie rozróżnialności elementów (w szczególności między sobą), to w świecie matematyki od czasu do czasu toczy się dyskusja co do poprawności tej definicji.

* * *

■ W 1939 r. N. Bourbaki[2] w pierwszym wydaniu swojej „Teorii zbiorów" zaproponował taką definicję: *zbiór składa się z elementów, które mają pewne własności i przebywają w pewnych relacjach między sobą lub*

[1] George Ferdinand Ludwig Phillip Cantor (1845–1918)

[2] Pseudonim grupy (głównie francuskich) matematyków, założonej na początku lat 30. XX wieku na uniwersytecie Ècole Normale Superiéure w Paryżu, „Association des collaborateurs de Nicolas Bourbaki"

z elementami innych zbiorów. Ale takie podejście również ma słabe punkty i powoduje pewne niezadowolenie, szczególnie u niematematyków.
W naszym kursie pojęcie zbioru jest pojęciem pierwotnym i niedefiniowalnym, z którym należy obchodzić się ostrożnie. Dalej będą rozpatrywane tylko takie *zbiory, których elementy są wystarczająco dokładnie określone, niezmienne i możemy zawsze stwierdzić, czy ten lub inny element należy do danego zbioru.*

Przykładami takich zbiorów są:
— zbiór $\mathbb{N}$ liczb naturalnych: $0, 1, 2, \ldots$;
— zbiór $\mathbb{N}^*$ liczb naturalnych niezerowych $1, 2, \ldots$;
— zbiór $\mathbb{Z}$ liczb całkowitych $0, \pm 1, \pm 2, \ldots$;
— zbiór liter w języku polskim;
— zbiór liter w alfabecie łacińskim;
— zbiór działań w pewnym programie komputerowym;
— zbiór wszystkich słów w języku komputerowym „Pascal" itd.

Niektóre z wyżej określonych zbiorów są wystarczająco skomplikowane. W celu ich badania wygodnie jest abstrahować od ich konkretnej przyrody, co będziemy robić dalej. W tym celu opiszmy niezbędny początkowy roboczy zestaw terminów i oznaczeń.

Najczęściej zbiory będziemy oznaczać dużymi literami alfabetu łacińskiego

$$A,\ B,\ C,\ \ldots,\ X,\ Y,\ Z.$$

Według G. Cantora każdy zbiór składa się z pewnych przedmiotów (=obiektów), które są nazywane *członkami* lub *elementami* zbioru. Przy tym elementy, które tworzą dany zbiór, mogą mieć różną przyrodę.

Na przykład możemy mówić o zbiorze rzeczy wszystkich mieszkańców Krakowa albo o zbiorze wszystkich obiektów we Wszechświecie, albo o zbiorze roślin w Tatrach itd.

Z terminem „zbiór" jest nieoddzielnie związany termin „należeć" (lub równoważnie „być elementem", „leżeć w" itd.).

Na przykład żuraw należy do rodziny ptaków; student należy do zbioru tych, którzy (w szczególności) pragną jak najszybciej złożyć wszystkie egzaminy przed wakacjami letnimi; okrąg jest zbiorem (=miejscem geometrycznym) punktów płaszczyzny, równooddalonych od stałego punktu, zwanego środkiem okręgu; pani Olga Tokarczuk jest wybitną pisarką polską; trójkąt jest elementem zbioru figur geometrycznych.

Jeśli obiekt *a należy* do zbioru A (obiekt a jest *elementem* zbioru A; obiekt *a wchodzi w* zbiór A; *a leży* w A), to zapisujemy

$$a \in A.$$

Jeśli zaś element *a nie należy* do zbioru A (element *a nie mieści się w* A czy *nie jest elementem* zbioru A), to zapisujemy

$$a \notin A.$$

Symbol „$\in$" jest nazywany *symbolem przynależności*. W celu skrócenia zamiast „$x_1 \in A$ oraz $x_2 \in A$ itd. oraz $x_n \in A$" będziemy krótko zapisywać, że

$$x_1, x_2, \ldots, x_n \in A.$$

Jeśli zbiór składa się ze skończonej liczby elementów, to jest nazywany *zbiorem skończonym*; w innym przypadku taki zbiór jest nazywany *nieskończonym*.

Ważne jest, aby prawidłowo (poprawnie) zadawać zbiór. Temu służy parę sposobów, wśród których są takie:

1) *Zbiór możemy zadawać przez przeliczanie jego elementów.*

Na przykład

- $\bullet_1$ zbiór Θ, składający się z liter alfabetu łacińskiego $a, b, c, \ldots, z$, zapisujemy w postaci $\Theta = \{a, b, c, \ldots, z\}$;
- $\bullet_2$ $A = \{0, 2, 4, 6, 8\}$ jest zbiorem parzystych liczb całkowitych dodatnich mniejszych niż 10;
- $\bullet_3$

$$K = \{\text{Australia, Ameryka Północna, Ameryka Południowa, Afryka, Antarktyda, Eurazja}\}$$

jest zbiorem wszystkich kontynentów planety Ziemia;

- $\bullet_4$ $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ jest zbiorem liczb całkowitych.

Sposób zadania zbioru przez przeliczanie jego elementów nie zawsze jest efektywny (w przypadku gdy zbiór zawiera wystarczająco dużo elementów, takie określenie zbioru potrzebuje wystarczająco dużo wysiłku).

2) *Zbiór A możemy zdefiniować przez określenie własności ograniczającej, która wyróżnia elementy zbioru A wśród elementów większego zbioru U*; w taki sposób zadany zbiór zwykle zapisujemy w postaci

$$A = \{x \in U \mid \mathcal{P}(x)\}$$

i mówimy, że zbiór „A składa się z tych elementów x ze zbioru U, dla których zachodzi własność czy warunek $\mathcal{P}(x)$".

Na przykład

- $\bullet_1$ $A = \{x \in \mathbb{Q} \mid x^2 = x + 2\}$ jest zbiorem tych liczb wymiernych, które są rozwiązaniami równania $x^2 = x + 2$;
- $\bullet_2$ zbiorem liczb parzystych całkowitych jest

$$2\mathbb{Z} = \{z \in \mathbb{Z} \mid \text{liczba } z \text{ jest podzielna przez 2 z resztą zerową}\}.$$

3) *Nowe zbiory możemy też tworzyć za pomocą działań na zbiorach*, o których opowiemy niżej.

* * *

■ **Działania na zbiorach.** Niech niżej A oraz B będą zbiorami. Jeśli każdy element zbioru A jest też elementem zbioru B, to mówimy, że A jest *podzbiorem* zbioru B (lub równoważnie: zbiór A *zawiera się* w zbiorze B; zbiór B *zawiera* A; zbiór A jest *częścią* zbioru B; zbiór A *wchodzi* w B; B jest *nadzbiorem* zbioru A). Wtedy zapisujemy, że $A \subseteq B$ lub $B \supseteq A$. Jeśli A jest podzbiorem zbioru B oraz B ma element, który nie należy do zbioru A, to mówimy, że A jest podzbiorem *właściwym* zbioru B i zapisujemy

$$A \subset B$$

(lub $A \subsetneq B$) albo $B \supset A$ (lub $B \supsetneq A$). Jeśli A posiada taki element a, że $a \notin B$, to zapisujemy $A \not\subseteq B$ (czyli A *nie jest podzbiorem* w zbiorze B lub A *nie zawiera się* w B).

Na przykład

- $\bullet_1$ zbiór parzystych liczb całkowitych $2\mathbb{Z} \subset \mathbb{Q}$ jest podzbiorem właściwym w zbiorze liczb wymiernych $\mathbb{Q}$ (innymi słowy, każda parzysta liczba całkowita jest wymierna);
- $\bullet_2$ $\mathbb{Z} \not\subseteq 2\mathbb{Z}$ znaczy, że zbiór liczb całkowitych $\mathbb{Z}$ nie zawiera się w zbiorze parzystych liczb całkowitych $2\mathbb{Z}$ (czyli nie każda liczba całkowita jest parzysta lub, co znaczy to samo, istnieją nieparzyste liczby całkowite);
- $\bullet_3$

$$\{\sqrt{2}, 3, -5, 0, 1/2\} \subsetneq \{3, \sqrt{2}, 1/2, 0, -6, \bullet, -5, 7\};$$

- $\bullet_4$ $\mathbb{Q} \subset \mathbb{R}$ – zbiór liczb wymiernych $\mathbb{Q}$ jest podzbiorem właściwym w zbiorze liczb rzeczywistych $\mathbb{R}$ (czyli każda liczba wymierna jest rzeczywistą, lecz znajdzie się liczba rzeczywista, która nie jest wymierną).

■ Będziemy mówić, że zbiory A oraz B są *równe* (*jednakowe*), jeśli jednocześnie $A \subseteq B$ oraz $B \subseteq A$; wtedy będziemy zapisywać, że

$$A = B.$$

Dualnie zbiory A i B *nie są równymi* (są *niejednakowymi* lub *różnymi*), jeśli $A \nsubseteq B$ lub $B \nsubseteq A$ (zapisujemy $A \neq B$).

Na przykład $2\mathbb{Z} \neq \mathbb{Z}$, $\mathbb{Z} \neq \mathbb{Q}$ itd.

■ Bardzo wygodne jest założenie, że istnieje zbiór, który nie posiada żadnego elementu. Taki zbiór będziemy oznaczać symbolem

$$\varnothing$$

i nazywać zbiorem *pustym*.

Zbiór pusty możemy scharakteryzować w taki sposób. Niech A będzie dowolnym zbiorem, a własność $\mathcal{P}(x)$ oznacza, że $x \neq x$. Wtedy

$$\{x \in A \mid x \neq x\} = \varnothing.$$

Zatem zbiór pusty $\varnothing$ jest podzbiorem każdego zbioru A, czyli

$$\varnothing \subseteq A.$$

Każdy zbiór niepusty A zawsze zawiera co najmniej dwa podzbiory: pusty $\varnothing$ oraz A, które są nazywane podzbiorami *trywialnymi*. Jest zrozumiałe, że każdy element $a \in A$ określa podzbiór jednoelementowy $\{a\}$ w zbiorze A. Zbiór A jest nazywany podzbiorem *niewłaściwym* zbioru A.

Dalej wprowadzamy pewne działania nad dowolnymi zbiorami A i B.

■ *Połączeniem* (*unią* lub *sumą mnogościową*) zbiorów A i B (oznaczamy przez $A \cup B$) jest nazywany zbiór tych obiektów, które są elementami zbioru A lub zbioru B, czyli

$$A \bigcup B = \{x \mid x \in A \text{ lub } x \in B\}.$$

Przykłady 1.1.1.

(**1**) Jeśli $A = \{1, 2, 3, 4, 5\}$ oraz $B = \{2, 4, \hbar, 6, 7\}$, to $A \cup B = \{1, 2, 3, 4, \hbar, 5, 6, 7\}$.

(**2**) Jeśli $A = \{\bullet, \oslash, \P, \clubsuit, \natural, \diamondsuit\}$ oraz $B = \{\bullet, \oslash, \P, ©, \spadesuit, \clubsuit\}$, to

$$A \bigcup B = \{\bullet, \oslash, \P, ©, \natural, \diamondsuit, \spadesuit, \clubsuit\}.$$

■ Jeśli elementami zbioru X są zbiory, to taki zbiór X będziemy nazywać *rodziną* zbiorów.

■ Niech Λ będzie dowolną (skończoną lub nieskończoną) rodziną zbiorów. Wtedy *unia* (lub *suma mnogościowa*) zbiorów rodziny Λ składa się ze wszystkich tych i tylko tych elementów, które należą do choćby jednego zbioru X z rodziny Λ, i jest oznaczana symbolem

$$\bigcup_{X\in\Lambda} X.$$

■ *Przekrojem* (*częścią wspólną*, *iloczynem teoriomnogościowym* lub *przecięciem*) zbiorów A i B (oznaczamy przez $A \cap B$) jest nazywany zbiór tych obiektów, które jednocześnie należą do zbioru A i do zbioru B, czyli

$$A\bigcap B = \{x \mid x \in A \text{ oraz } x \in B\}.$$

Przykłady 1.1.2.

(**1**) Jeśli $A = \{\oplus, \star, 1, 2, 3, 6, 7, 8\}$ oraz $B = \{2, 3, 7, \uplus, \vee\}$, to $A \cap B = \{2, 3, 7\}$.

(**2**) Niech $A = \{\bullet, \oslash, ¶, \clubsuit, \natural, \diamondsuit\}$ oraz $B = \{\bullet, \oslash, ¶, ©, \spadesuit, \clubsuit\}$. Wtedy $A \cap B = \{\bullet, \oslash, ¶, \clubsuit\}$.

■ Zbiory A i B są nazywane *rozłącznymi*, jeśli

$$A\bigcap B = \varnothing.$$

■ Niech Λ będzie dowolną (skończoną lub nieskończoną) rodziną zbiorów. Wtedy *przecięcie rodziny zbiorów* Λ składa się dokładnie z tych elementów, które jednocześnie należą do wszystkich zbiorów X z rodziny Λ, i oznacza się przez

$$\bigcap_{X\in\Lambda} X.$$

■ *Dopełnienie względne* zbioru B do zbioru A (oznaczamy przez $A \setminus B$ i krótko mówimy „A minus B") jest podzbiorem tych elementów ze zbioru A, które nie są elementami zbioru B, czyli

$$A \setminus B = \{x \in A \mid x \notin B\}.$$

Mówimy też, że $A \setminus B$ jest *różnicą* zbiorów A i B.

Przykłady 1.1.3.

(**1**) Różnica $\mathbb{Z} \setminus 2\mathbb{Z}$ jest zbiorem nieparzystych liczb całkowitych.

(**2**) $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ jest zbiorem niezerowych liczb rzeczywistych.

(**3**) Niech $A = \{1, 2, 3, 7, 5, \bullet, \star\}$ oraz $B = \{8, 2, 3, \star\}$. Wtedy $A \setminus B = \{1, 5, 7, \bullet\}$.

■ Zbiór wszystkich podzbiorów danego zbioru A jest nazywany *booleanem* (i oznaczany symbolem $\mathcal{B}(A)$ lub $\mathcal{P}(A)$) lub *potęgą* (i oznaczany symbolem 2^A) zbioru A, czyli

$$\mathcal{B}(A) = \mathcal{P}(A) = 2^A = \{B \mid B \subseteq A\}.$$

Przykłady 1.1.4.

(**1**) $\mathcal{B}(\varnothing) = \{\varnothing\}$, $\mathcal{B}(\{1, 2\}) = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$.

(**2**) Jeśli A jest zbiorem jednoelementowym, to $2^A = \{\varnothing, A\}$ jest zbiorem dwuelementowym.

Zachodzą takie własności działań na zbiorach.

Twierdzenie 1.1.5. *Niech A, B, C będą podzbiorami zbioru U. Wtedy są spełnione takie równości:*

(1) $A \cap B = B \cap A$ *(przemienność przecięcia);*
(2) $A \cup B = B \cup A$ *(przemienność sumy mnogościowej);*
(3) $(A \cap B) \cap C = A \cap (B \cap C)$ *(łączność przecięcia);*
(4) $(A \cup B) \cup C = A \cup (B \cup C)$ *(łączność sumy mnogościowej);*
(5) $A \cap A = A$ *(idempotentność przecięcia);*
(6) $A \cup A = A$ *(idempotentność sumy mnogościowej);*
(7) $A \cap U = A$*;*
(8) $A \cup \varnothing = A$.

Dowód. (1) Niech x będzie dowolnym elementem z przecięcia $A \cap B$. Wtedy $x \in A$ oraz $x \in B$ lub, co jest tym samym, $x \in B$ oraz $x \in A$. To znaczy, że $x \in B \cap A$, czyli $A \cap B \subseteq B \cap A$. Odwrotnie, jeśli $z \in B \cap A$, to $z \in B$ oraz $z \in A$. Zatem $z \in A$ oraz $z \in B$, a więc $z \in A \cap B$. Wnosimy, że $B \cap A \subseteq A \cap B$. Na podstawie definicji równości zbiorów otrzymujemy, że

$$B \bigcap A = A \bigcap B.$$

Własności (2)-(8) można udowodnić podobnie (zostawiamy to Czytelnikowi jako ćwiczenia). □

* * *

■ **Iloczyn kartezjański zbiorów.** Niech $A, A_1, \ldots, A_n$ będą pewnymi zbiorami niepustymi ($n \in \mathbb{N}^*$). Zbiór $\{a_1, a_2, \ldots, a_n\}$ elementów $a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n$, ujętych w ustalonej kolejności (czyli a_1 ze zbioru A_1 jest pierwszym elementem, a_2 ze zbioru A_2 jest drugim elementem, ..., a_n ze zbioru A_n jest ostatnim elementem) jest nazywany *n-ką uporządkowaną* i oznaczany przez

$$(a_1, \ldots, a_n).$$

■ Zbiór wszystkich n-ek uporządkowanych

$$(a_1, \ldots, a_n), \text{ gdzie } a_1 \in A_1, \ldots, a_n \in A_n,$$

jest nazywany iloczynem *kartezjańskim* zbiorów $A_1, \ldots, A_n$ i oznaczany przez

$$A_1 \times \cdots \times A_n,$$

czyli

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) \mid a_i \in A_i \ (i = 1, \ldots, n)\}.$$

■ Zbiór

$$A^n = \underbrace{A \times \cdots \times A}_{n \text{ czynników}} = \{(a_1, \ldots, a_n) \mid \ a_i \in A \ (i = 1, \ldots, n)\}$$

jest nazywany *n-tą potęgą kartezjańską* zbioru A. n-Ka uporządkowana postaci

$$(a_1, \ldots, a_n)$$

jest też nazywana *wierszem* długości n. Dość często n-ka uporządkowana elementów $a_1 \in A_1, \ldots, a_n \in A_n$ jest zapisywana w postaci

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

i nazywana *kolumną* wysokości n. Element a_i ($i = 1, \ldots, n$) przy tym jest nazywany *i-tym współrzędnym* (lub *i-tą składową*) wiersza

$$(a_1, \ldots, a_i, \ldots, a_n)$$

czy kolumny

$$\begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix}.$$

■ Dwie n-ki uporządkowane $(a_1, \ldots, a_n)$ oraz $(b_1, \ldots, b_n)$ są nazywane *równymi* (co zapisujemy w postaci

$$(a_1, \ldots, a_n) = (b_1, \ldots, b_n) \text{ lub } \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}),$$

jeśli zachodzą równości

$$a_1 = b_1, \ldots, a_n = b_n.$$

Jeśli $n = 2$, to 2-ka uporządkowana (a_1, a_2) jest nazywana *parą uporządkowaną* z pierwszym elementem a_1 i z drugim (=kolejnym) elementem a_2.

■ **Kwantyfikatory.** Korzystne bywają kwantyfikatory: *kwantyfikator ogólności* „$\forall$" oraz *kwantyfikator istnienia* „$\exists$" – symbole, które będziemy czasem stosować w celu skrócenia naszych notatek:

- kwantyfikator *ogólności* „$\forall$" będziemy wykorzystywać we wzorach postaci

  $$\forall_{x \in X} \ : \varphi(x),$$

  które czytamy na jeden ze sposobów: „dla każdego elementu $x \in X$ zachodzi własność $\varphi(x)$" albo „dla dowolnego $x \in X$ jest spełniony warunek $\varphi(x)$", albo „ dla wszystkich $x \in X$ jest spełnione $\varphi(x)$" itd.;
- kwantyfikator *istnienia* (czy *szczególności*) „$\exists$" będziemy wykorzystywać w formułach postaci

  $$\exists_{x \in X} \ : \varphi(x),$$

  co czytamy na jeden ze sposobów: „dla pewnego elementu $x \in X$ jest spełnione $\varphi(x)$" albo „istnieje taki element $x \in X$, że zachodzi $\varphi(x)$", albo „dla pewnego $x \in X$ spełnia się $\varphi(x)$" itd.

Przykłady 1.1.6.

(**1**) Notację „$\forall_{x\in\mathbb{N}} : x > 2$” czytamy jako „dla wszystkich liczb naturalnych x, które są większe od liczby 2”.

(**2**) Zdanie „znajdzie się nieujemna liczba wymierna x” krótko możemy zapisać w takiej postaci „$\exists_{x\in\mathbb{Q}} : x \geqslant 0$”.

Ćwiczenia 1.1.7.

(**1**) Znaleźć $A \cup B$, $A \cap B$, $A \setminus B$ oraz $B \setminus A$ dla zbiorów A i B, jeśli:
(*a*) $A = \{x, y, z, t\}$ oraz $B = \{x, t\}$;
(*b*) $A = \{\{x, y\}, z, t\}$ oraz $B = \{x, t\}$;
(*c*) $A = \{\{x, y\}, y, \{z, t\}, z\}$ oraz $B = \{\{y\}, y, \{z\}\}$;
(*d*) $A = \{x, y, \{z, t\}\}$ oraz $B = \{a, x, z\}$;
(*e*) $A = \{\{x, \{y\}, z, t\}\}$ oraz $B = \{a, \{x\}, z\}$;
(*f*) $A = \{x \in \mathbb{Q} \mid x < 5\}$ oraz $B = \{x \in \mathbb{Q} \mid x \geqslant 5\}$;
(*g*) $A = \{x \in \mathbb{N} \mid x < -1\}$ oraz $B = \{x \in \mathbb{N} \mid x = 3\}$;
(*h*) $A = \{x \in \mathbb{R} \mid x > 5\}$ oraz $B = \{x \in \mathbb{Z} \mid x > 5\}$;
(*k*) $A = \{x \in \mathbb{Z} \mid x < 2\}$ oraz $B = \{z \in \mathbb{Z} \mid x < 1\}$.
(**2**) Udowodnić, że:
(*a*) $A \setminus (B \setminus C) = (A \setminus B) \cup (A \cap C)$;
(*b*) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$;
(*c*) $(A \setminus B) \cup C = ((A \cup C) \setminus B) \cup (B \cap C)$;
(*d*) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
(*e*) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
(*f*) $\varnothing \cap A = \varnothing$;
(g) $A \cup (A \cap B) = A$.
(**3**) Udowodnić, że:
(*a*) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$;
(*b*) $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.
(**4**) Traktując punkty na płaszczyźnie jako uporządkowane pary (a, b) liczb rzeczywistych a oraz b, znaleźć iloczyny kartezjańskie $A \times B$ i $B \times A$ (oraz naszkicować ich rysunki), jeśli:
(*a*) $A = \{x \in \mathbb{R} \mid 0 < x \leqslant 1\}$ oraz $B = \{x \in \mathbb{R} \mid 1 < x \leqslant 3\}$;
(*b*) $A = \{x \in \mathbb{R} \mid -1 < x < 1\}$ oraz $B = \{x \in \mathbb{R} \mid 0 < x \leqslant 1\}$;
(*c*) $A = \{x \in \mathbb{R} \mid -1 < x < 2 \text{ lub } 3 < x \leqslant 5\}$ oraz $B = \{x \in \mathbb{R} \mid 2 < x \leqslant 3 \text{ lub } 5 < x \leqslant 7\}$;
(*d*) $A = \{x \in \mathbb{R} \mid x < 2 \text{ oraz } x > 2\}$ oraz $B = \{x \in \mathbb{R} \mid x^2 > 0\}$.

Uwagi. Jeszcze w 200 roku p.n.e. starożytny grecki matematyk Appolonius[3] wykorzystywał dwie liczby do oznaczenia punktu na płaszczyźnie. Kartezjusz[4] oraz P. de Fermat[5] pierwsi rozwinęli tę metodę (która później została nazwana *geometrią analityczną*) i uczynili ją dostępną dla wszystkich. W końcu XIX w. G. Cantor wprowadził pojęcie iloczynu kartezjańskiego.

---

[3] Appolonius z Pergi (240–190 r. p.n.e.)
[4] René Descartes (1596–1650)
[5] Pierre de Fermat (1601–1665)

Matematyk perski Al-Chwarizmi[(6)] niewiadomą w równaniach nazywał słowem „recz". Kartezjusz jako pierwszy użył litery x do oznaczenia niewiadomej.

[(6)] Muhamed ibn Musu al-Chwarizmi (780–850)

1.2. Wstęp do arytmetyki liczb całkowitych

■ **Indukcja matematyczna.** Indukcja matematyczna jest jednym z głównych instrumentów wykorzystywanych w celu dowodów twierdzeń w arytmetyce.

`Pewnik indukcji.` Niech X będzie dowolnym zbiorem. Jeśli są spełnione własności:

1) $0 \in X$,
2) dla dowolnego n zachodzi implikacja

$$n \in X \Rightarrow n+1 \in X,$$

to $\mathbb{N} \subseteq X$.

■ Z tego pewnika wynikają trzy następne ważne własności:

1°. (`zasada indukcji`) Dla każdej liczby naturalnej $n \in \mathbb{N}$ przez $\mathcal{P}(n)$ oznaczmy zagadnienie, które oznacza, że własność $\mathcal{P}$ zachodzi dla danego $n \in \mathbb{N}$. Jeśli są spełnione dwa warunki:
   i) (*warunek początkowy*) twierdzenie $\mathcal{P}(0)$ jest poprawne;
   ii) (*krok indukcji*) dla każdego $n \in \mathbb{N}$ ze spełnienia warunku $\mathcal{P}(n)$ wynika, że warunek $\mathcal{P}(n+1)$ też zachodzi;
   to teza $\mathcal{P}(n)$ jest poprawna dla wszystkich $n \in \mathbb{N}$;

2°. (`zasada maksymalności` lub `zasada maksimum`) Każdy niepusty podzbiór ograniczony S zbioru liczb naturalnych $\mathbb{N}$ posiada największą liczbę naturalną;

3°. (`zasada minimalności` lub `zasada minimum`) Każdy niepusty podzbiór S zbioru liczb naturalnych $\mathbb{N}$ posiada najmniejszą liczbę naturalną.

* * *

■ **Podzielność liczb całkowitych.** Niech a, b będą dowolnymi liczbami całkowitymi. Liczba całkowita b jest nazywana *dzielnikiem* (lub *czynnikiem*) liczby a, jeśli $a = bt$ dla pewnego $t \in \mathbb{Z}$. Przy tym mówimy, że „a jest podzielne przez b z resztą zerową" lub „a jest podzielne przez b bez reszty" lub, co jest tym samym, piszemy $b \mid a$ (i mówimy „b dzieli

a"). Jeśli b nie jest dzielnikiem liczby a, to zapisujemy $b \nmid a$. Przypomnijmy też, że $|a|$ jest wartością bezwzględną liczby rzeczywistej a, czyli

$$|a| = \begin{cases} a, & \text{gdy } a \geqslant 0, \\ -a, & \text{gdy } a < 0. \end{cases}$$

Lemat 1.2.1. *Dla dowolnych liczb całkowitych a, b, c, d, m, n zachodzą następujące własności:*

(1) $a \mid a$ *(zwrotność);*
(2) $a \mid 0$*;*
(3) $(\pm 1) \mid a$*;*
(4) *jeśli* $0 \mid a$*, to* $a = 0$*;*
(5) *jeśli* $b \mid a$ *oraz* $c \mid b$*, to* $c \mid a$ *(przechodniość);*
(6) *jeśli* $c \mid a$*, to* $c \mid ab$*;*
(7) *jeśli* $c \mid a$ *oraz* $c \mid b$*, to* $c \mid (a \pm b)$*;*
(8) $a \mid b$ *wtedy i tylko wtedy, gdy* $a \mid (-b)$*;*
(9) *jeśli* $c \mid a$ *oraz* $c \mid (a \pm b)$*, to* $c \mid (\pm b)$*;*
(10) *jeśli* $c \mid a$ *oraz* $c \nmid b$*, to* $c \nmid (a \pm b)$*;*
(11) *jeśli* $b \mid a$*, to* $bc \mid ac$*;*
(12) *jeśli* $a \mid c$ *oraz* $b \mid d$*, to* $ab \mid cd$*;*
(13) *jeśli* $c \mid a$ *oraz* $c \mid b$*, to* $c \mid (ma \pm nb)$*;*
(14) *jeśli* $b \mid a$ *oraz* $a \mid b$*, to* $a = \pm b$*;*
(15) *jeśli* $bc \mid ac$ *oraz* $c \neq 0$*, to* $b \mid a$*;*
(16) *jeśli* $b \mid a$ *oraz* $|b| > |a|$*, to* $a = 0$*.*

Dowód. (1) Rzeczywiście, $a = a \cdot 1$, a więc $a \mid a$.

(2) Z równości $0 = 0 \cdot a$ wnosimy, że $a \mid 0$.

(3) Skoro $a = (\pm 1) \cdot (\pm a)$, to $(\pm 1) \mid a$.

(4) Niech $0 \mid a$, czyli znajdzie się taka liczba całkowita k, że $a = 0 \cdot k$. Wtedy $a = 0$.

(5) Ponieważ $b \mid a$ (to znaczy $a = bk_1$ dla pewnego $k_1 \in \mathbb{Z}$) oraz $c \mid b$ (czyli $b = ck_2$ dla pewnego $k_2 \in \mathbb{Z}$), to

$$a = (ck_2)k_1 = c(k_1 k_2),$$

a zatem $c \mid a$.

(6) Załóżmy, że $c \mid a$. Wtedy istnieje taka liczba całkowita k, że $a = ck$, a stąd

$$ab = (ck)b = c(kb)$$

oraz $kb \in \mathbb{Z}$, co daje, że $c \mid ab$.

(7) Z warunku $c \mid a$ oraz $c \mid b$ wnosimy, że $a = ck_1$ oraz $b = ck_2$ dla pewnych $k_1, k_2 \in \mathbb{Z}$, a więc

$$a \pm b = ck_1 \pm ck_2 = c(k_1 \pm k_2),$$

czyli $c \mid (a \pm b)$.

(8) Niech $a \mid b$. Wtedy $b = ak$ dla pewnej liczby całkowitej k, skąd $(-b) = a(-k)$ oraz $a \mid (-b)$. Odwrotnie, jeśli $a \mid (-b)$, to $-b = as$ dla pewnej liczby całkowitej s. Zatem

$$b = a(-s) \text{ oraz } a \mid b.$$

(9) Z relacji $c \mid a$ oraz $c \mid (a \pm b)$ wynika, że $a = ck_1$ oraz $a \pm b = ck_2$ dla pewnych liczb całkowitych k_1, k_2. Wtedy

$$\pm b = (a \pm b) - a = ck_2 - ck_1 = c(k_2 - k_1),$$

gdzie $k_2 - k_1 \in \mathbb{Z}$, a stąd $c \mid (\pm b)$.

(10) Z warunku $c \mid a$ otrzymujemy, że $a = ck$ dla pewnej liczby całkowitej k. Niech $c \nmid b$. Udowodnimy nie wprost, że $c \nmid (a \pm b)$. W tym celu załóżmy, że $c \mid (a \pm b)$. To znaczy, że istnieje taka liczba całkowita l, że $a \pm b = cl$, i na tej podstawie

$$\pm b = (a \pm b) - a = cl - ck = c(l - k).$$

Zatem $c \mid b$, co przeczy założeniu. Wnosimy, że $c \nmid (a \pm b)$.

(11) Jeśli $b \mid a$, to $a = bs$ dla pewnej liczby całkowitej s i wtedy

$$ac = (bs)c = b(sc) = b(cs) = (bc)s,$$

a więc $bc \mid ac$.

(12) Z założenia wynika, że $c = at$ oraz $d = br$ dla pewnych $t, r \in \mathbb{Z}$ i na tej podstawie

$$cd = (at)(br) = ((at)b)r =$$
$$= (a(tb))r = (a(bt))r = ((ab)t)r = (ab)(tr),$$

czyli $ab \mid cd$.

(13) Niech $c \mid a$, $c \mid b$ oraz $m, n \in \mathbb{Z}$. Wtedy $a = ck_1$, $b = ck_2$ dla pewnych liczb całkowitych k_1, k_2. Mnożąc równości przez m oraz n, odpowiednio, oraz sumując (odejmując) je, otrzymujemy

$$\begin{gathered} ma \pm nb = m(ck_1) \pm n(ck_2) = (mc)k_1 \pm (nc)k_2 = \\ = (cm)k_1 \pm (cn)k_2 = c(mk_1) \pm c(nk_2) = c(mk_1 \pm nk_2), \end{gathered}$$

czyli $c \mid ma \pm nb$.

(14) Jak wyżej, z $a \mid b$ oraz $b \mid a$ na podstawie definicji wnosimy, że $b = ak_1$ oraz $a = bk_2$ dla pewnych liczb całkowitych k_1, k_2. Zatem

$$b = (bk_2)k_1 = b(k_2k_1),$$

skąd $b(1 - k_1k_2) = 0$. Jeśli teraz $b = 0$, to $a = bk_2 = 0$ i teza zachodzi. Dlatego rozpatrzmy inną ewentualność, czyli $b \neq 0$. Wtedy $1 - k_1k_2 = 0$, a to znaczy, że $k_i = \pm 1$ $(i = 1, 2)$. Udowodniliśmy, że

$$a = \pm b.$$

(15) W rzeczy samej, jeśli $bc \mid ac$, to istnieje taka liczba całkowita l, że $ac = (bc)l$, a stąd

$$ac = b(cl) = b(lc) = (bl)c$$

i wtedy $(a - bl)c = 0$. Ponieważ $c \neq 0$, to $a - bl = 0$, a zatem $b \mid a$.

(16) Z warunku $b \mid a$ otrzymujemy, że $a = bl$ dla pewnego $l \in \mathbb{Z}$. Wtedy

$$|a| = |bl| = |b||l|.$$

Ponieważ $|b| > |a|$, to obowiązkowo zachodzi, że $l = 0 = a$. □

Przykłady 1.2.2.

(**1**) Liczba całkowita 5 jest dzielnikiem liczby 155, bo $155 = 5 \cdot 31$; zatem $5 \mid 155$ lub, co jest równoważne, liczba 155 jest krotna liczbie 5.

(**2**) Liczba 3 nie jest czynnikiem liczby -217 (czyli $3 \nmid (-217)$), bo $-217 = 3 \cdot (-72) - 1$.

Twierdzenie 1.2.3 (twierdzenie o dzieleniu z resztą dla liczb całkowitych)**.** *Dla dowolnych liczb całkowitych a oraz b, gdzie $b \neq 0$, istnieje dokładnie jedna para liczb całkowitych q, r taka, że*

$$a = bq + r \text{ oraz } 0 \leqslant r < |b|.$$

Dowód. 1) *Istnienie.* Załóżmy, że $b > 0$. Rozpatrzmy zbiór

$$S = \{a - bq \mid q \in \mathbb{Z} \text{ oraz } a - bq \geqslant 0\}.$$

Ponieważ $a - b(-a^2) \geqslant 0$, to S jest podzbiorem niepustym w zbiorze liczb naturalnych $\mathbb{N}$. Na podstawie zasady minimum S posiada liczbę najmniejszą; niech to będzie $r_0 = a - bq_0$ dla pewnej liczby całkowitej q_0. Jeśli $r_0 \geqslant b$, to

$$a - b(q_0 + 1) = a - bq_0 - b = r_0 - b \geqslant 0,$$

a więc $r_0 - b \in S$ oraz $r_0 - b < r_0$, co jest sprzeczne z minimalnością liczby r_0. Zatem $r_0 < b$. Przypadek $b < 0$ jest podobny (zostawiamy Czytelnikowi do samodzielnego udowodnienia).

2) *Jednoznaczność.* W rzeczy samej, jeśli $a = bq + r$ oraz $a = bq_1 + r_1$ dla pewnych liczb całkowitych q, q_1, r, r_1 takich, że $0 \leqslant r, r_1 < |b|$, to

$$0 = a - a = bq - bq_1 + r - r_1,$$

a zatem $r - r_1 = b(q_1 - q)$, gdzie $|b| > |r - r_1|$. Z lematu 1.2.1(16) otrzymujemy, że $r_1 = r$. Wtedy $b(q_1 - q) = 0$, gdzie $b \neq 0$ na podstawie założenia, a więc $q = q_1$. □

■ Liczby, o których mówi się w twierdzeniu 1.2.3, są nazywane tak: q – *ilorazem niepełnym* (lub *cząstką niepełną*), a r – *resztą* z dzielenia a przez b. Jeśli $r = 0$, to mówią, że *a jest podzielne przez b bez reszty* (lub *a jest krotne b*) i wtedy q jest nazywane *ilorazem* (pełnym) (lub *cząstką* (pełną)) z dzielenia a przez b.

Przykłady 1.2.4.

(**1**) Przy dzieleniu liczby $a = -357$ przez liczbę $b = 5$ reszta jest równa $r = 3$, a iloraz niepełny $q = -72$, bo

$$-357 = 5 \cdot (-72) + 3.$$

Zwracamy uwagę na to, że $357 = 5 \cdot 71 + 2$, czyli $r = 2$ jest resztą z dzielenia 357 przez 5, a $q = 71$ jest ilorazem niepełnym.

(**2**) Wykażmy, że kwadrat nieparzystej liczby całkowitej z w wyniku dzielenia przez 8 daje resztę równą 1. Rzeczywiście, na podstawie twierdzenia 1.2.3 mamy $z = 2q+r$ dla pewnych liczb całkowitych q, r, gdzie $0 \leqslant r < 2$. Skoro liczba z jest nieparzysta, to $r = 1$. Wtedy

$$z^2 = (2q + 1)^2 = 4q^2 + 4q + 1.$$

Liczba $q^2 + q = q(q + 1)$ jest iloczynem dwóch kolejnych liczb całkowitych, a więc jedna z liczb q lub $q + 1$ jest parzysta. Zatem $8 \mid (4q^2 + 4q)$ oraz 1 jest resztą z dzielenia z^2 przez 8.

* * *

■ **Największy wspólny dzielnik dwóch liczb całkowitych.** Niech a, b, d będą liczbami całkowitymi. Jeśli $d \mid a$ oraz $d \mid b$, to liczba d jest nazywana *dzielnikiem wspólnym* liczb a oraz b. Największy wśród ich wspólnych dzielników jest nazywany *największym wspólnym dzielnikiem* liczb a oraz b (i oznaczany symbolem (a, b) lub $\mathrm{NWD}(a, b)$).

Lemat 1.2.5. *Dla liczb całkowitych a, b, q, r, k, d zachodzą następujące własności:*

(1) *jeśli $b \mid a$, to $\mathrm{NWD}(a, b) = |b|$;*

(2) *jeśli $a = bq + r$, to $\mathrm{NWD}(a, b) = \mathrm{NWD}(r, b)$;*

(3) *jeśli k jest wspólnym dzielnikiem liczb $a \neq 0$ oraz b, to*

$$\mathrm{NWD}\left(\frac{a}{k}, \frac{b}{k}\right) = \frac{\mathrm{NWD}(a, b)}{|k|};$$

(4) *jeśli $\mathrm{NWD}(a, b) = d$, to*

$$\mathrm{NWD}\left(\frac{a}{d}, \frac{b}{d}\right) = 1;$$

(5) *jeśli $d \mid a$, $d \mid b$ oraz*

$$\mathrm{NWD}\left(\frac{a}{d}, \frac{b}{d}\right) = 1,$$

to $|d| = \mathrm{NWD}(a, b)$.

Dowód. (1) Oczywiście, że $|b|$ dzieli $\mathrm{NWD}(a,b)$. Natomiast $\mathrm{NWD}(a,b) \mid b$ na podstawie definicji, a zatem $\mathrm{NWD}(a,b)$ dzieli $|b|$. Dalej z definicji największego wspólnego dzielnika i lematu 1.2.1(14) otrzymujemy, że $|b| = \mathrm{NWD}(a,b)$.

(2) Niech $\mathrm{NWD}(a,b) = d$ oraz $\mathrm{NWD}(r,b) = t$. Jeśli $a = bq + r$, to $r = a - bq$ i na mocy lematu 1.2.1(7) mamy, że $d \mid r$, a zatem $d \mid t$. Ponieważ $t \mid b$ oraz $t \mid r$, to $t \mid a$ i wtedy $t \mid d$. Wnioskujemy, że $d = t$.

(3) Zanotujmy, że $k \neq 0$. Oznaczmy

$$d = \mathrm{NWD}(a,b) \text{ oraz } u = \mathrm{NWD}\left(\frac{a}{k}, \frac{b}{k}\right).$$

Wtedy $k \mid d$ (przekonać się samodzielnie). Skoro $a = ds_1$ i $b = ds_2$ dla pewnych $s_1, s_2 \in \mathbb{Z}$, to

$$\frac{a}{k} = \frac{d}{k} \cdot s_1 \quad \text{oraz} \quad \frac{b}{k} = \frac{d}{k} \cdot s_2,$$

a stąd $\frac{d}{k} \mid u$. Zatem $u = \frac{d}{k} \cdot l$ dla pewnego $l \in \mathbb{Z}$, a więc $uk = dl$. Jako wniosek

$$u|k| = d|l| \quad \text{oraz} \quad d \mid (u|k|).$$

Następnie skoro

$$u \mid \frac{a}{k} \text{ oraz } u \mid \frac{b}{k},$$

to

$$\frac{a}{k} = uk_1 \text{ oraz } \frac{b}{k} = uk_2$$

dla pewnych liczb całkowitych k_1, k_2, a zatem $(u|k|) \mid a$ oraz $(u|k|) \mid b$. Lecz wtedy $(u|k|) \mid d$. Na tej podstawie z lematu 1.2.1(14) otrzymujemy, że $d = u|k|$.

(4) Wynika z własności (3), jeśli wziąć $k = d$.

(5) Wynika z własności (3). □

Twierdzenie 1.2.6 (o istnieniu NWD)**.** *Największy wspólny dzielnik* $\mathrm{NWD}(a,b)$ *dwóch liczb całkowitych* a, b*, z których przynajmniej jedna jest niezerowa, zawsze istnieje i jest dokładnie jedoznacznie określony.*

Dowód. Istnienie. Załóżmy, że $b \neq 0$. Na podstawie twierdzenia o dzieleniu z resztą istnieje taki ciąg kolejnych dzieleń:

$$\begin{array}{rcll} a & = & bq_1 + r_1, & \text{gdzie } 0 < r_1 < |b|, \quad (\epsilon_1) \\ b & = & r_1q_2 + r_2, & \text{gdzie } 0 < r_2 < r_1, \quad (\epsilon_2) \\ r_1 & = & r_2q_3 + r_3, & \text{gdzie } 0 < r_3 < r_2, \quad (\epsilon_3) \\ & \vdots & & \vdots \qquad\qquad \vdots \\ r_{n-3} & = & r_{n-2}q_{n-1} + r_{n-1}, & \text{gdzie } 0 < r_{n-1} < r_n, \quad (\epsilon_{n-1}) \\ r_{n-2} & = & r_{n-1}q_n + r_n, & \text{gdzie } 0 < r_n < r_{n-1}, \quad (\epsilon_n) \\ r_{n-1} & = & r_nq_{n+1} + r_{n+1}, & \text{gdzie } r_{n+1} = 0, \quad (\epsilon_{n+1}) \end{array}$$

przy czym $n \in \mathbb{N}^*$, $r_i, q_i \in \mathbb{Z}$ $(i = 1, \ldots, n+1)$.

Rzeczywiście, ponieważ

$$|b| = r_1 > r_2 > \cdots ,$$

to znajdzie się taka liczba naturalna $n+1$, że reszta r_{n+1} jest równa 0, lecz $r_n \neq 0$. Łańcuch $(\epsilon_1) - (\epsilon_{n+1})$ kolejnych dzieleń jest nazywany *algorytmem Euklidesa*[(7)].

Chcemy udowodnić, że *ostatnia niezerowa reszta w łańcuchu* $(\epsilon_1) - (\epsilon_{n+1})$ *jest równa* $\mathrm{NWD}(a, b)$, czyli

$$\mathrm{NWD}(a, b) = r_n.$$

W rzeczy samej, z (ϵ_{n+1}) wynika, że $r_n \mid r_{n-1}$, a stąd, biorąc pod uwagę (ϵ_n) oraz lemat 1.2.1(13), otrzymujemy, że $r_n \mid r_{n-2}$. Rozumując podobnie dalej, przez skończoną liczbę kroków otrzymujemy, że $r_n \mid a$ oraz $r_n \mid b$.

Teraz wykażmy, że r_n jest największym wśród wspólnych dzielników liczb a i b. W tym celu załóżmy, że d jest dowolnym wspólnym dzielnikiem liczb a oraz b. Wtedy z równości (ϵ_1) w wyniku lematu 1.2.1(13) wyciągamy, że $d \mid r_1$. Znowu z (ϵ_2) i lematu 1.2.1(13) mamy $d \mid r_2$. Podobnymi rozumowaniami przez skończoną liczbę kroków otrzymujemy, że $d \mid r_n$. To znaczy, że $r_n = \mathrm{NWD}(a, b)$.

Jednoznaczność wynika na podstawie twierdzenia 1.2.3. □

(7) Euklides (325–265 r. p.n.e.)

Przykład 1.2.7.

Znajdźmy $d = \mathrm{NWD}(-2585, 7985)$, stosując algorytm Euklidesa. Oczywiście, że $\mathrm{NWD}(-2585, 7985) = \mathrm{NWD}(2585, 7985)$. Dalej:

1) dzielimy z resztą większą liczbę 7985 przez mniejszą liczbę 2585; otrzymujemy

$$7985 = 2585 \cdot 3 + 230,$$

2) dzielimy 2585 przez 230; dostajemy

$$2585 = 230 \cdot 11 + 55,$$

3) dzielimy 230 przez 55; otrzymujemy

$$230 = 55 \cdot 4 + 10,$$

4) dzielimy 55 przez 10; mamy

$$55 = 10 \cdot 5 + 5,$$

5) dzielimy 10 przez 5; otrzymujemy

$$10 = 5 \cdot 2 + 0.$$

Zatem ostatnia niezerowa reszta

$$d = \mathrm{NWD}(-2585, 7985) = 5$$

jest największym wspólnym dzielnikiem.

* * *

■ Wnioski z algorytmu Euklidesa.

Zachodzi taki

Lemat 1.2.8. *Dla dowolnych liczb całkowitych a oraz b, gdzie $b \neq 0$, istnieją takie liczby całkowite u oraz v, że*

$$\mathrm{NWD}(a, b) = au + bv. \tag{1.1}$$

Dowód. Niech

$$\mathcal{E} = \{au + bv \mid u, v \in \mathbb{Z}\}.$$

Wtedy zbiór $\mathcal{E}$ jest niepusty i posiada najmniejszą liczbę dodatnią d_0, gdzie $d_0 = au_0 + bv_0$ dla pewnych liczb całkowitych u_0, v_0. Na podstawie twierdzenia o dzieleniu z resztą $a = d_0q + r$ dla pewnego $r \in \mathbb{Z}$, gdzie $0 \leqslant r < d_0$. Wtedy

$$r = a - d_0q = a - (au_0 + bv_0)q = a(1 - u_0q) + b(-v_0q) \in \mathcal{E},$$

a zatem $r = 0$. To znaczy, że $d_0 \mid a$. Podobnie możemy udowodnić, że $d_0 \mid b$.

Zostało nam przekonać się, że d_0 jest największym wśród dzielników wspólnych liczb a oraz b. Zatem załóżmy, że t jest dowolnym dzielnikiem wspólnym liczb a i b. Wtedy z lematu 1.2.1(6) dostajemy

$$t \mid au \text{ oraz } t \mid bv.$$

Na podstawie lematu 1.2.1(7) wnosimy, że t dzieli $au + bv$, czyli $t \mid d_0$. To oznacza, że $d_0 = \mathrm{NWD}(a, b)$. □

Z lematu 1.2.8 otrzymujemy

Wniosek 1.2.9. *Jeśli $a, b, c \in \mathbb{Z}$ oraz c jest wspólnym dzielnikiem liczb a i b, to $c \mid \mathrm{NWD}(a, b)$.*

Lemat 1.2.10. *Największy wspólny dzielnik dwóch liczb całkowitych a i b jest równy 1 wtedy i tylko wtedy, gdy znajdą się takie liczby całkowite u oraz v, że*

$$au + bv = 1.$$

Dowód. ($\Rightarrow$) Wynika z lematu 1.2.8, bo $\mathrm{NWD}(a, b) = 1$.

($\Leftarrow$) Załóżmy, że dla liczb całkowitych a, b istnieją $u, v \in \mathbb{Z}$ takie, że

$$au + bv = 1.$$

Jedna z liczb a, b jest niezerowa. Niech d będzie dowolnym wspólnym dzielnikiem liczb a i b. Wtedy $d \mid a$ oraz $d \mid b$ i za lematem 1.2.1(13) mamy, że $d \mid 1$, co oznacza, że $d = \pm 1$. Zatem $\mathrm{NWD}(a, b) = 1$. □

■ Przedstawienie $\mathrm{NWD}(a, b)$ w postaci (1.1) jest nazywane *liniowym*, a liczby u, v są nazywane *współczynnikami Bezouta*[8]. Wykorzystując notacje z twierdzenia 1.2.6, łatwo przekonać się, że

$$\begin{aligned} r_i &= au_i + bv_i, \text{ gdzie} \\ u_0 &= v_1 = 1, \\ u_1 &= v_0 = 0, \\ u_{i+1} &= u_{i-1} - q_i u_i, \\ v_{i+1} &= v_{i-1} - q_i v_i \ (0 \leqslant i \leqslant n). \end{aligned}$$

[8] Étienne Bézout (1730–1883)

Zatem $u = u_n$, $v = v_n$ (gdzie u, v są współczynnikami, o których mówi się w lemacie 1.2.8). Opisany algorytm obliczania $\mathrm{NWD}(a, b)$, u oraz v jest nazywany *rozszerzonym algorytmem Euklidesa.*

Przykład 1.2.11.

Znajdźmy liniowe przedstawienie $\mathrm{NWD}(a, b)$, gdzie $a = 2585$ oraz $b = 7985$. Stosując algorytm Euklidesa (patrz przykład 1.2.7), obliczamy:

- $5 = 55 - 10 \cdot 5$;
- $10 = 230 - 55 \cdot 4$;
- $55 = 2585 - 230 \cdot 11$;
- $230 = 7985 - 2585 \cdot 3$,

a stąd

$$5 = 55 - 10 \cdot 5 = 55 - (230 - 55 \cdot 4) \cdot 5 = 55 \cdot (1 + 4 \cdot 5) + 230 \cdot (-5) =$$
$$= (2585 - 230 \cdot 11) \cdot 21 + 230 \cdot (-5) = 230 \cdot ((-11) \cdot 21 - 5) + 2585 \cdot 21 =$$
$$230 \cdot (-236) + 2585 \cdot 21 = (7985 - 2585 \cdot 3) \cdot (-236) + 2585 \cdot 21 =$$
$$= 7985 \cdot (-236) + 2585 \cdot (236 \cdot 3 + 21) = 7985 \cdot (-236) + 2585 \cdot 729.$$

Zatem

$$u = -236, \quad v = 729.$$

Wniosek 1.2.12. *Jeśli a, b, m są niezerowymi liczbami całkowitymi, to*

$$\mathrm{NWD}(am, bm) = \mathrm{NWD}(a, b)|m|.$$

Dowód. Niech

$$w = \mathrm{NWD}(am, bm) \text{ oraz } d = \mathrm{NWD}(a, b).$$

Wtedy $md \mid am$ oraz $md \mid bm$, a więc $(|m|d) \mid w$. Natomiast $d = au + bv$ dla pewnych $u, v \in \mathbb{Z}$, a więc $md = amu + bmv$, na podstawie czego $w \mid (|m|d)$. Z lematu 1.2.1(14) otrzymujemy, że $w = d|m|$. □

* * *

■ **Liczby pierwsze. Podstawowe twierdzenie arytmetyki.** Liczba całkowita p jest nazywana *pierwszą*, jeśli $p \geqslant 2$ oraz wszystkie jej dzielniki całkowite tworzą zbiór $\{\pm p, \pm 1\}$. Liczba naturalna, która nie jest pierwsza i różni się od $0, 1$, jest nazywana *złożoną.*

Lemat 1.2.13. *Niech $a, b, a_1, \ldots, a_n \in \mathbb{Z}$ oraz p będzie liczbą pierwszą. Wtedy zachodzą następujące własności:*

(1) *jeśli $p \nmid a$, to* $\mathrm{NWD}(a, p) = 1$;

(2) *jeśli* $p \mid ab$, *to* $p \mid a$ *lub* $p \mid b$;
(3) *jeśli* $p \mid a_1 \cdots a_n$, *to* $p \mid a_i$ *dla pewnego* i $(1 \leqslant i \leqslant n)$.

Dowód. (1) Niech $d = \mathrm{NWD}(a, p)$. Wtedy $d \mid p$, a zatem $d = 1$ lub $d = p$. Ponieważ $d \mid a$ oraz $p \nmid a$, to $d = 1$.

(2) Nie wprost. Załóżmy, że $p \mid ab$ oraz p nie dzieli ani a, ani b. Stosując lemat 1.2.10, na podstawie części (1) mamy

$$pu_1 + av_1 = 1 \text{ oraz } pu_2 + bv_2 = 1$$

dla pewnych $u_1, v_1, u_2, v_2 \in \mathbb{Z}$ i wtedy iloczyn stronami tych równości daje

$$p^2u_1u_2 + pbu_1v_2 + apv_1u_2 + abv_1v_2 = 1.$$

Skoro $p \mid ab$, to w wyniku lematu 1.2.1(7) otrzymujemy, że $p \mid 1$, co daje sprzeczność. Zatem $p \mid a$ lub $p \mid b$.

(3) Wynika na podstawie części (2). □

Twierdzenie 1.2.14 (arytmetyki podstawowe)**.** *Każda dodatnia liczba całkowita, która różni się od* 1*, ma dokładnie jednoznaczne rozłożenie w iloczyn liczb pierwszych* (*z dokładnością do kolejności czynników w tym rozłożeniu*).

Dowód. Istnienie. Niech a będzie liczbą całkowitą, przy czym $a > 1$. Jeśli a jest liczbą pierwszą, to teza zachodzi. Jeśli zaś a nie jest liczbą pierwszą, to

$$a = a_1 \cdot a_2$$

dla pewnych liczb całkowitych a_1 i a_2, gdzie $1 < a_i < a$ $(i = 1, 2)$. Stosując rozumowania indukcyjne do liczb a_i, przez skończoną liczbę kroków otrzymamy, że a jest iloczynem skończonej liczby liczb pierwszych.

Jednoznaczność. Wykażmy teraz, że jeśli

$$a = p_1 \cdots p_n \text{ oraz } a = q_1 \cdots q_m \ (n, m \in \mathbb{N}^*)$$

są iloczynami liczb pierwszych $p_1, \ldots, p_n$ oraz $q_1, \ldots, q_m$, to $n = m$ i dla każdej liczby naturalnej k $(1 \leqslant k \leqslant n)$ znajdzie się taka liczba naturalna s $(1 \leqslant s \leqslant m)$, że $p_k = q_s$.

Stosujemy indukcję względem liczby czynników pierwszych n. Jeśli $n = 1$, to teza zachodzi. Załóżmy, że teza zachodzi dla wszystkich $i \leqslant n-1$. Ponieważ

$$a = (\prod_{i=1}^{n-1} p_i) \cdot p_n,$$

to

$$p_n \mid \prod_{j=1}^{m} q_j.$$

Wtedy na mocy lematu 1.2.13 istnieje taka liczba całkowita s ($1 \leqslant s \leqslant m$), że $p_n \mid q_s$ i na tej podstawie $p_n = q_s$. Zatem

$$\prod_{i=1}^{n-1} p_i = q_1 \cdots q_{s-1} q_{s+1} \cdots q_m$$

i możemy zastosować założenie indukcji. Wnosimy, że $n-1 = m-1$ oraz dla każdego i ($1 \leqslant i \leqslant n-1$) istnieje taka liczba naturalna j, że $j \neq s$, $1 \leqslant j \leqslant m$ oraz $p_i = q_j$, a zatem teza zachodzi. □

■ Ponieważ

$$z = \pm|z|,$$

to łącząc jednakowe czynniki pierwsze w rozkładzie dodatniej liczby całkowitej $|z|$, otrzymujemy następny

Wniosek 1.2.15. *Każdą liczbę całkowitą z, która różni się od $-1, 0, 1$, możemy dokładnie jednym sposobem przedstawić w* `postaci kanonicznej`

$$z = \pm p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s},$$

gdzie $s, \alpha_i \in \mathbb{N}^*$ $(i = 1, \ldots, s)$, $p_1, p_2, \ldots, p_s$ są różnymi liczbami pierwszymi takimi, że

$$p_1 < p_2 < \cdots < p_s.$$

Przykład 1.2.16.

Znajdźmy rozkład kanoniczny liczby 3360. Liczba 3360 jest parzysta, a więc $3360 = 2 \cdot 1680$. Podobnie

$$1680 = 2 \cdot 840,\ 840 = 2 \cdot 420,\ 420 = 2 \cdot 210,\ 210 = 2 \cdot 105.$$

Liczba 105 jest podzielna przez kolejną liczbę pierwszą 3, czyli $105 = 3 \cdot 35$. W końcu $35 = 5 \cdot 7$, gdzie 5 oraz 7 są liczbami pierwszymi. Ten ciąg dzieleń możemy uprościć i zapisać w postaci:

3360	2
1680	2
840	2
420	2
210	2
105	3
35	5
7	7
1	

Zatem

$$3360 = 2^5 \cdot 3 \cdot 5 \cdot 7$$

jest rozłożeniem kanonicznym liczby 3360.

* * *

■ **Twierdzenie Euklidesa o liczbach pierwszych.** Zachodzi następny fakt, który jest wiadomy już od ponad 2300 lat.

Twierdzenie 1.2.17. *Zbiór wszystkich liczb pierwszych $\mathcal{P}$ jest nieskończony.*

Dowód. Nie wprost. Załóżmy, że $\mathcal{P}$ jest zbiorem skończonym, czyli $\mathcal{P} = \{p_1, p_2, \ldots, p_n\}$ dla pewnych liczb pierwszych takich, że

$$p_1 = 2 < p_2 = 3 < p_3 < \cdots < p_n.$$

Niech

$$N = p_1 \cdots p_n + 1.$$

Na mocy podstawowego twierdzenia arytmetyki liczba N jest podzielna przez pewną liczbę pierwszą q. Wtedy $q = p_i$ dla pewnego i $(1 \leqslant i \leqslant n)$. W wyniku lematu 1.2.1 liczba pierwsza q dzieli liczbę

$$N - p_1 \cdots p_n,$$

czyli $q \mid 1$, a to nie jest możliwe. Otrzymana sprzeczność oznacza, że moc

$$\mathrm{card}(\mathcal{P}) = \infty$$

jest nieskończona. □

* * *

■ **Sito Eratostenesa**[(9)]**.** Podstawą metody zaproponowanej przez Eratostenesa jest takie spostrzeżenie: *z podstawowego twierdzenia arytmetyki wynika, że każda liczba złożona n ma co najmniej jeden dzielnik pierwszy p taki, że*

$$p \leqslant \sqrt{n}.$$

W rzeczy samej, niech $n = a \cdot b$ dla pewnych liczb naturalnych a, b większych niż 1. Możemy założyć, że $a \leqslant b$. Wtedy

$$a^2 \leqslant a \cdot b = n,$$

a więc $a \leqslant \sqrt{n}$. Z tego otrzymujemy taką metodę (wiadomą pod nazwą *sita Eratostenesa*): `jeśli` $n > 1$ `jest liczbą całkowitą, to albo` n `jest liczbą pierwszą, albo istnieje liczba piersza` p `taka, że`

$$p \leqslant \sqrt{n} \quad \texttt{oraz} \quad p \mid n.$$

Przykład 1.2.18.

Znajdźmy rozłożenie kanoniczne liczby 7102. Skoro liczba 7102 jest parzysta, to $7102 = 3551 \cdot 2$. Ponieważ

$$59 < \sqrt{3551} < 60,$$

to jedna z liczb pierwszych

$$2, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59$$

dzieli 3551 lub liczba 3551 jest pierwsza. W rzeczy samej, za pomocą obliczeń bezpośrednich przekonujemy się, że liczba 3551 nie jest podzielna przez żadną z liczb pierwszych

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47.$$

Lecz $3551 = 53 \cdot 67$ oraz liczby $53,\ 67$ są pierwsze. Zatem otrzymujemy takie rozłożenie kanoniczne

$$7102 = 2 \cdot 53 \cdot 67.$$

* * *

(9) Eratostenes (276–194 p.n.e.)

■ **Liczby względnie pierwsze.** Liczby całkowite a oraz b, spełniające warunek

$$\mathrm{NWD}(a,b)=1,$$

są nazywane *względnie pierwszymi*.

Na przykład łatwo zauważyć, że $\mathrm{NWD}(27,1351)=1$, czyli liczby 27 oraz 1351 są względnie pierwsze.

Lemat 1.2.19. *Dla dowolnych liczb całkowitych a,b,c i dowolnych nieujemnych liczb całkowitych n,m zachodzą własności:*

(1) *jeśli* $\mathrm{NWD}(a,b)=1$, *to* $\mathrm{NWD}(ac,b)=\mathrm{NWD}(b,c)$;
(2) *jeśli* $\mathrm{NWD}(a,b)=\mathrm{NWD}(a,c)=1$, *to* $\mathrm{NWD}(a,bc)=1$;
(3) *jeśli* $c \mid ab$ *oraz* $\mathrm{NWD}(b,c)=1$, *to* $c \mid a$;
(4) *jeśli* $b \mid a$, $c \mid a$ *oraz* $\mathrm{NWD}(b,c)=1$, *to* $bc \mid a$;
(5) *jeśli* $\mathrm{NWD}(a,b)=1$, *to* $\mathrm{NWD}(a^n,b^m)=1$;
(6) *jeśli* $\mathrm{NWD}(a^s,b^t)=1$ *dla pewnych* $s,t\in\mathbb{N}^*$, *to* $\mathrm{NWD}(a,b)=1$;
(7) *jeśli p oraz q są różnymi liczbami pierwszymi, to* $\mathrm{NWD}(p^n,q^m)=1$.

Dowód. (1) Niech $d=\mathrm{NWD}(ac,b)$. Znajdą się takie liczby całkowite k_1,k_2,u,v, że

$$b=dk_1,\ ac=dk_2 \text{ oraz } au+bv=1.$$

Mnożąc pierwszą równość przez vc, a drugą przez u, otrzymujemy $bvc=dk_1vc$ oraz $auc=dk_2u$ i na tej podstawie

$$c=auc+bvc=dk_2u+dk_1vc=d(k_2u+k_1vc),$$

czyli $d \mid c$. Wnosimy, że $d \mid \mathrm{NWD}(b,c)$. Z kolei $\mathrm{NWD}(b,c)$ dzieli i liczbę ac, i liczbę b, co daje, że $\mathrm{NWD}(b,c) \mid d$. Na podstawie lematu 1.2.1(14) wnioskujemy, że

$$\mathrm{NWD}(ac,b)=d=\mathrm{NWD}(b,c).$$

(2) Skoro $\mathrm{NWD}(a,b)=1$ i $\mathrm{NWD}(a,c)=1$, to istnieją takie liczby całkowite u_1,v_1,u_2,v_2, że

$$au_1+bv_1=1 \text{ oraz } au_2+cv_2=1.$$

Mnożąc otrzymane równości stronami, otrzymujemy

$$\begin{aligned}1&=a^2u_1u_2+acu_1v_2+abu_2v_1+bcv_1v_2=\\&=a(au_1u_2+cu_1v_2+bu_2v_1)+bc(v_1v_2),\end{aligned}$$

czyli $\mathrm{NWD}(a, bc) = 1$ na podstawie lematu 1.2.10.

(3) Z założenia $c \mid ab$, czyli $ab = ck$ dla pewnego $k \in \mathbb{Z}$, oraz $\mathrm{NWD}(b, c) = 1$, czyli $bu + cv = 1$ dla pewnych liczb całkowitych u, v. Mnożąc pierwszą równość przez u, otrzymujemy $cku = abu$; natomiast druga równość daje $bu = 1 - cv$. Łącząc wyniki, otrzymujemy

$$cku = a(bu) = a(1 - cv)$$

lub równoważnie $a = c(ku + av)$, czyli $c \mid a$.

(4) Skoro $b \mid a$ oraz $c \mid a$ i $\mathrm{NWD}(b, c) = 1$, to istnieją takie liczby całkowite s, l, u, v, że $a = bs$, $a = cl$ oraz $bu + cv = 1$. Mnożąc ostatnią równość przez sl, otrzymujemy

$$a(ul + vs) = aul + avs = bsul + clvs = (bu + cv)sl = sl.$$

Lecz wtedy

$$a^2 = a \cdot a = bs \cdot cl = bc(sl) = bca(ul + vs),$$

a stąd $a = bc(ul + vs)$, czyli $bc \mid a$.

(5) Niech $d = \mathrm{NWD}(a^n, b^m)$. Jeśli $n = 0$ lub $m = 0$, to $d = 1$. W wyniku tego załóżmy, że n, m są dodatnimi liczbami całkowitymi. Załóżmy nie wprost, że $d \neq 1$. Na podstawie twierdzenia 1.2.14 znajdzie się taka liczba pierwsza p, że $p \mid d$ i, jako wniosek, $p \mid a^n$ oraz $p \mid b^m$. Za lematem 1.2.13(3) mamy $p \mid a$ oraz $p \mid b$. Zatem p dzieli $\mathrm{NWD}(a, b)$, co daje sprzeczność. To znaczy, że $\mathrm{NWD}(a^n, b^m) = 1$.

(6) Jeśli $\mathrm{NWD}(a^s, b^t) = 1$ dla pewnych $s, t \in \mathbb{N}^*$, to za lematem 1.2.10 znajdą się takie liczby całkowite $u, v \in \mathbb{Z}$, że $a^s u + b^t v = 1$ lub, co jest równoważne,

$$a(a^{s-1}u) + b(b^{t-1}v) = 1,$$

czyli $\mathrm{NWD}(a, b) = 1$ na podstawie lematu 1.2.10.

(7) Wynika z części (5). □

* * *

■ **Najmniejsza wspólna wielokrotność dwóch liczb całkowitych.** Liczba całkowita k, która jest jednocześnie podzielna przez a oraz b, jest

nazywana ich *wspólną krotnością*. Najmniejsze nieujemne ze wszystkich wspólnych krotnych liczb a i b jest nazywane ich *najmniejszym wspólnym wielokrotnym* (i oznaczane przez $[a, b]$ lub $\mathrm{NWW}(a, b)$).

Lemat 1.2.20. *Jeśli a, b są liczbami całkowitymi, gdzie $a, b \neq 0$, to*

$$\mathrm{NWW}(a, b) = \frac{|a| \cdot |b|}{\mathrm{NWD}(a, b)}.$$

Dowód. Niech $d = \mathrm{NWD}(a, b)$. Skoro $a = dk_1$ i $b = dk_2$ pewnych liczb całkowitych k_1, k_2, to

$$\frac{ab}{d} = ak_2 = bk_1,$$

a więc a dzieli $\frac{|a||b|}{d}$ oraz b dzieli $\frac{|a||b|}{d}$. To, że liczba $\frac{|a||b|}{d}$ jest najmniejsza wśród wspólnych krotności liczb a i b zostawiamy Czytelnikowi do samodzielnego udowodnienia. □

Przykład 1.2.21.

Znajdźmy najmniejsze wspólne wielokrotne liczb 2585 oraz 7985. Biorąc pod uwagę wyniki obliczeń z przykładu 1.2.11, wnosimy, że

$$\mathrm{NWW}(2585, 7985) = \frac{2585 \cdot 7985}{5} = 517 \cdot 7985 = 4128245.$$

Ćwiczenia 1.2.22.

(**1**) Znaleźć $\mathrm{NWD}(a, b)$, jeśli:
(a) $a = 31605$ oraz $b = -12915$;
(b) $a = 1402$ oraz $b = -689$;
(c) $a = 7525$ oraz $b = -3337$;
(d) $a = 3655$ oraz $b = 663$;
(e) $a = 772$ oraz $b = 64$;
(f) $a = 774$ oraz $b = 62$;
(g) $a = 140$ oraz $b = 335$;
(h) $a = 170$ oraz $b = 425$;
(i) $a = 666$ oraz $b = 324$;
(j) $a = 287$ oraz $b = 14$;
(k) $a = -827$ oraz $b = -131$;
(l) $a = -2352$ oraz $b = -268$;
(m) $a = -29049$ oraz $b = -2047$;
(n) $a = 213$ oraz $b = 94$;
(o) $a = 3211$ oraz $b = 7163$;
(p) $a = 15088$ oraz $b = 4554$;

(q) $a = 13699$ oraz $b = 1349$;
(r) $a = 354$ oraz $b = 66$;
(s) $a = 871$ oraz $b = 3627$;
(t) $a = 2159$ oraz $b = 221$;
(u) $a = 21567$ oraz $b = 5005$.
(**2**) Znaleźć NWW(a, b), jeśli:
(a) $a = -31605$ oraz $b = 12915$;
(b) $a = -1402$ oraz $b = 689$;
(c) $a = -7525$ oraz $b = 3337$;
(d) $a = 3655$ oraz $b = 663$;
(e) $a = 772$ oraz $b = 64$;
(f) $a = 774$ oraz $b = 66$;
(g) $a = 770$ oraz $b = 15$;
(h) $a = 279$ oraz $b = 372$;
(i) $a = 178$ oraz $b = -381$;
(j) $a = 213$ oraz $b = 94$.
(**3**) Znaleźć rozłożenie kanoniczne liczby:
(a) 31605;
(b) 12915;
(c) 689;
(d) 3655;
(e) 663;
(f) 1402;
(g) 7525;
(h) 3337;
(i) 13699;
(j) 15088;
(k) 354;
(l) 871.
(**4**) Udowodnić, że:
(a) iloczyn trzech kolejnych liczb całkowitych jest podzielny przez 6;
(b) suma kwadratów dwóch kolejnych liczb całkowitych przy dzieleniu przez 4 daje resztę równą 1;
(c) $10 \mid (16 \cdot 23^{23} - 142 \cdot 23^{32})$;
(d) $6 \mid (n^3 - n)$ dla każdego $n \in \mathbb{Z}$.

Uwagi. Teoria liczb (jak i geometria) są najstarszymi rozdziałami matematyki, podstawy których są zawarte w *Elementach* Euklidesa, opublikowanych jeszcze ponad 2300 lat temu, oraz w *Arytmetyce* Diofantosa[(10)].

W roku 1364 król Kazimierz Wielki[(11)] założył uniwersytet w Krakowie (=Akademię Krakowską), a w roku 1402 pewien krakowianin sfinansował założenie katedry matematyki. Podręcznik akademicki (faktycznie pierwszy polski z teorii liczb) *Arytmetyka liczb całkowitych* z 1620 r. opublikował J. Brożek[(12)].

---

(10) Diofantos (200/214–284/298 n.e.)
(11) Kazimierz III Wielki (1310–1370)
(12) Jan Brożek (1585–1652)

Początkowa klasyczna matematyczna terminologia polska w dużym stopniu jest dziełem J. Śniadeckiego[13].

[13] Jan Śniadecki (1756–1830)

1.3. Relacje binarne

■ Do opisania dowolnego związku między różnymi rzeczami (przedmiotami czy istotami) wykorzystujemy pojęcie „relacja". Rozpatrzmy je z algebraicznego punktu widzenia. Wszędzie dalej ograniczamy się z powodu naszych potrzeb do relacji odnoszących się tylko do pary obiektów (czyli *relacji binarnych*).

■ Niech X i Y będą dowolnymi zbiorami. Trójka uporządkowana

$$(X, Y, \mathcal{R})$$

jest nazywana *relacją binarną* (lub *dwuczłonową*) między zbiorem X oraz zbiorem Y, jeśli

$$\mathcal{R} \subseteq X \times Y.$$

Z każdą relacją $(X, Y, \mathcal{R})$ są związane takie zbiory:

- X jest *dziedziną* relacji $(X, Y, \mathcal{R})$;
- Y jest *przeciwdziedziną* relacji $(X, Y, \mathcal{R})$;
- $D(\mathcal{R}) = \{x \in X \mid (x, y) \in \mathcal{R} \text{ dla pewnego } y \in Y\}$ jest *rzutem* relacji $(X, Y, \mathcal{R})$ na zbiór X;
- $\operatorname{Im} \mathcal{R} = \{y \in Y \mid (x, y) \in \mathcal{R} \text{ dla pewnego } x \in X\}$ jest *obrazem* relacji $(X, Y, \mathcal{R})$;
- $\mathcal{R}$ jest *wykresem* relacji $(X, Y, \mathcal{R})$.

■ Jeśli $\mathcal{R} \subseteq X \times X$, to zamiast trójki uporządkowanej $(X, X, \mathcal{R})$ będziemy krótko pisać $(X, \mathcal{R})$ i nazywać taką parę uporządkowaną *relacją binarną* na zbiorze X. `Wszędzie dalej w celu skrócenia zamiast terminu ,,relacja binarna'' będziemy używać krótszego terminu ,,relacja''.`

■ **Konwencja.** Zamiast $(X, \mathcal{R})$ będziemy częściej pisać $\mathcal{R}$, jeżeli z kontekstu wiemy, o którym zbiorze X mówimy, i mówić (dopuszczając pewną wolność terminologiczną), że $\mathcal{R}$ jest *relacją na zbiorze* X.

■ Jeśli $x \in X$, to przez $\mathcal{R}(x)$ będziemy oznaczać podzbiór

$$\{y \in X \mid (x, y) \in \mathcal{R}\},$$

który jest nazywany *klasą równoważności* (*klasą abstrakcji* lub *warstwą*) relacji $\mathcal{R}$ z *reprezentantem* x.

■ Dla notacji takiej, że elementy $x, y \in X$ są związane relacją $\mathcal{R}$ będziemy używać równoważnych oznaczeń:

- $(x, y) \in \mathcal{R}$, czyli para elementów x, y jest w relacji $\mathcal{R}$;
- $x\mathcal{R}y$, czyli element x jest związany z elementem y przez relację $\mathcal{R}$;
- $x \backsim_{\mathcal{R}} y$, czyli x jest równoważny elementowi y względem relacji $\mathcal{R}$;
- $y \in \mathcal{R}(x)$, co znaczy, że element y należy do tej samej klasy abstrakcji co i element x (względem relacji $\mathcal{R}$).

Przykłady 1.3.1.

(**1**) Jeśli $A = \{1, 2, 3, 4\}$, $\mathcal{R} = \{(x, y) \mid x, y \in A, x \text{ dzieli } y \text{ oraz } x \leqslant 3\}$, to mamy relację

$$\mathcal{R} = \{(1,1), (1,2), (1,3), (1,4), (2,2), (2,4), (3,3)\}$$

na zbiorze A.

(**2**) Ponieważ $\varnothing \subseteq X \times X$, to $(X, \varnothing)$ jest relacją nazywaną *relacją pustą* na zbiorze X.

■ Na każdym zbiorze niepustym X zawsze możemy rozpatrywać takie dwie relacje:

- *tożsamościową* (X, I_X) z wykresem
$$I_X = \{(x, x) \mid x \in X\};$$
- *uniwersalną* (X, U_X) z wykresem
$$U_X = \{(x, y) \mid x, y \in X\}.$$

* * *

■ **Działania na relacjach.** Niech $\mathcal{R}$ będzie relacją na zbiorze X. Wtedy:

- relacja *odwrotna* do relacji $\mathcal{R}$ jest określana jako
$$\mathcal{R}^{-1} = \{(y, x) \mid (x, y) \in \mathcal{R}\};$$
- jeśli $\mathcal{R}_1$ jest relacją na zbiorze X_1, a X jest podzbiorem w X_1 oraz
$$\mathcal{R} = \mathcal{R}_1 \bigcap (X \times X),$$
to relacja $\mathcal{R}$ jest nazywana *zawężeniem* (lub *ograniczeniem*) relacji $(X_1, \mathcal{R}_1)$ do podzbioru X (ograniczenie $\mathcal{R}$ relacji $\mathcal{R}_1$ do zbioru X często oznaczane jest przez $\mathcal{R}_{1|X}$);

- jeśli $(X, \mathcal{R})$ oraz $(X, \mathcal{S})$ są relacjami na zbiorze X, to relacja $(X, \mathcal{R} \cap \mathcal{S})$ jest nazywana ich *przecięciem*, a relacja $(X, \mathcal{R} \cup \mathcal{S})$ – ich *sumą*;
- relacja $(X, \mathcal{R} \circ \mathcal{S})$ jest nazywana *złożeniem* (*kompozycją* lub *iloczynem*) relacji $(X, \mathcal{R})$ i $(X, \mathcal{S})$, jeśli

$$\mathcal{R} \circ \mathcal{S} = \{(x, y) \in X \times X \mid \text{istnieje takie } z \in X, \text{ że } x\mathcal{R}z \text{ oraz } z\mathcal{S}y\};$$

- mówią, że z relacji $(X, \mathcal{R})$ *wynika* relacja $(X, \mathcal{S})$ (lub relacja $\mathcal{S}$ jest *rozszerzeniem* relacji $\mathcal{R}$), jeśli $\mathcal{R} \subseteq \mathcal{S}$. Zatem $\mathcal{S} = \mathcal{R}$ w tym i tylko tym przypadku, gdy $\mathcal{S} \subseteq \mathcal{R}$ oraz $\mathcal{R} \subseteq \mathcal{S}$.

Przykłady 1.3.2.

(**1**) Niech

$$\mathcal{A} = \{(x, y) \in \mathbb{Q}^2 \mid x \geqslant y\} \text{ oraz } \mathcal{B} = \{(x, y) \in \mathbb{Q}^2 \mid x \neq y\}.$$

Wtedy zauważamy, że

$$\begin{array}{lcl}
\mathcal{A}^{-1} & = & \{(x, y) \in \mathbb{Q}^2 \mid y \geqslant x\}, \\
\mathcal{B}^{-1} & = & \{(x, y) \in \mathbb{Q}^2 \mid y \neq x\}, \\
\mathcal{A} \cap \mathcal{B} & = & \{(x, y) \in \mathbb{Q}^2 \mid x > y\}, \\
\mathcal{A} \cup \mathcal{B} & = & \{(x, y) \in \mathbb{Q}^2 \mid x \geqslant y \text{ lub } x \neq y\} = \mathbb{Q}^2 = U_{\mathbb{Q}}, \\
\mathcal{A} \circ \mathcal{B} & = & \{(x, y) \in \mathbb{Q}^2 \mid \text{ istnieje taki element } z \in \mathbb{Q}, \text{ że } x \geqslant z \text{ oraz } z \neq y\}.
\end{array}$$

(**2**) Jeśli Λ jest ogółem ludzi, zamieszkujących naszą planetę, a $x\mathcal{R}y$ oznacza, że „x jest mężem y", gdzie $x, y \in \Lambda$, to $y\mathcal{R}^{-1}x$ oznacza, że „y jest żoną x".

Niech $a\mathcal{S}b$ oznacza, że „a jest ojcem b". Wtedy $x\mathcal{R}^{-1} \circ \mathcal{S}y$ oznacza, że „x jest matką lub macochą y", bo istnieje takie $z \in \Lambda$, że „x jest żoną z", a „z jest ojcem y".

(**3**) Jeśli $\mathcal{T}$ jest relacją „być bratem", a $\mathcal{V}$ jest relacją „być jednym z rodziców" na zbiorze Λ (z poprzedniego przykładu), to $\mathcal{T} \circ \mathcal{V}$ jest relacją oznaczającą „być bratem jednego z rodziców" w zbiorze Λ.

Twierdzenie 1.3.3 (własności algebraiczne działań na relacjach)**.** *Niech $\mathcal{A}, \mathcal{B}, \mathcal{C}$ będą relacjami na zbiorze X. Wtedy zachodzą następujące własności:*

(1) $(\mathcal{A}^{-1})^{-1} = \mathcal{A}$;
(2) $\mathcal{A} \circ I_X = \mathcal{A} = I_X \circ \mathcal{A}$;
(3) $\mathcal{A} \circ \varnothing = \varnothing = \varnothing \circ \mathcal{A}$;
(4) $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} = \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$;
(5) $(\mathcal{A} \circ \mathcal{B})^{-1} = \mathcal{B}^{-1} \circ \mathcal{A}^{-1}$;
(6) *w przypadku ogólnym relacje $\mathcal{A}$ oraz $\mathcal{B}$ nie są przemienne, czyli $\mathcal{A} \circ \mathcal{B} \neq \mathcal{B} \circ \mathcal{A}$;*

(7) *jeśli* $\mathcal{A} \subseteq \mathcal{B}$, *to* $\mathcal{A}^{-1} \subseteq \mathcal{B}^{-1}$.

Dowód. Niech dalej $x, y, z \in X$.

(1) W rzeczy samej, jeśli $x(\mathcal{A}^{-1})^{-1}y$, to $y\mathcal{A}^{-1}x$, a więc $x\mathcal{A}y$. To znaczy, że

$$(\mathcal{A}^{-1})^{-1} \subseteq \mathcal{A}.$$

Odwrotnie, z $x\mathcal{A}y$ wynika, że $y\mathcal{A}^{-1}x$, a stąd $x(\mathcal{A}^{-1})^{-1}y$, czyli

$$\mathcal{A} \subseteq (\mathcal{A}^{-1})^{-1}.$$

(2) Z warunku $x\mathcal{A} \circ I_X y$ wynika, że $x\mathcal{A}z$ oraz $zI_X y$ dla pewnego elementu $z \in X$, a to jest możliwe, gdy $z = y$. Zatem $\mathcal{A} \circ I_X \subseteq \mathcal{A}$. Jeśli $x\mathcal{A}y$, to $x\mathcal{A}y$ oraz $yI_{\mathcal{A}}y$, co daje, że $x\mathcal{A} \circ I_X y$, czyli $\mathcal{A} \subseteq \mathcal{A} \circ I_X$. Zatem

$$\mathcal{A} \circ I_X = \mathcal{A}.$$

Inna równość ma podobny dowód (proponujemy Czytelnikowi zbudować go samodzielnie).

(3) Jeśli $x\mathcal{A} \circ \varnothing y$, to znajdzie się taki element $z \in X$, że $x\mathcal{A}z$ oraz $z\varnothing y$, a to nie jest możliwe. Zatem

$$\mathcal{A} \circ \varnothing \subseteq \varnothing.$$

Dowód odwrotnego zawierania jest oczywisty.

W podobny sposób możemy udowodnić drugą równość.

(4) Rzeczywiście, jeśli $x(\mathcal{A}\circ\mathcal{B})\circ\mathcal{C}y$, to znajdzie się taki element $z \in X$, że $x\mathcal{A} \circ \mathcal{B}z$ oraz $z\mathcal{C}y$. Z $x\mathcal{A} \circ \mathcal{B}z$ wynika, że $x\mathcal{A}w$ oraz $w\mathcal{B}z$ dla pewnego elementu $w \in X$. Wtedy z $w\mathcal{B}z$ oraz $z\mathcal{C}y$ otrzymujemy, że $w\mathcal{B} \circ \mathcal{C}y$. Dalej z $x\mathcal{A}w$ i $w\mathcal{B} \circ \mathcal{C}y$ mamy $x\mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})y$, czyli $(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C} \subseteq \mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})$. Podobnym sposobem z $x\mathcal{A} \circ (\mathcal{B} \circ \mathcal{C})y$ dostajemy, że $x(\mathcal{A} \circ \mathcal{B}) \circ \mathcal{C}y$ i na tej podstawie teza zachodzi.

(5) Jeśli $x(\mathcal{A} \circ \mathcal{B})^{-1}y$, to $y\mathcal{A} \circ \mathcal{B}x$, a więc istnieje taki element $z \in X$, że $y\mathcal{A}z$ oraz $z\mathcal{B}x$. To znaczy, że $x\mathcal{B}^{-1}z$ oraz $z\mathcal{A}^{-1}y$, co powoduje, że $x\mathcal{B}^{-1} \circ \mathcal{A}^{-1}y$, czyli

$$(\mathcal{A} \circ \mathcal{B})^{-1} \subseteq \mathcal{B}^{-1} \circ \mathcal{A}^{-1}.$$

Dowód odwrotnego zawierania jest podobny (co zostawiamy Czytelnikowi).

(6) Zostawiamy Czytelnikowi do samodzielnego znalezienia takie relacje $\mathcal{A}$ i $\mathcal{B}$.

(7) Mamy

$$x\mathcal{A}^{-1}y \Rightarrow y\mathcal{A}x \Rightarrow y\mathcal{B}x \Rightarrow x\mathcal{B}^{-1}y.$$

□

* * *

■ **Klasyfikacja relacji oraz ich własności.** Niech $\mathcal{R}$ będzie relacją na zbiorze X. Relacja $\mathcal{R}$ jest nazywana:

- *zwrotną* na X, jeśli dla każdego elementu $x \in X$ zachodzi związek $x\mathcal{R}x$. Innymi słowy $\mathcal{R}$ jest zwrotna na zbiorze X, jeśli

$$I_X \subseteq \mathcal{R};$$

- *symetryczną* na X, jeśli dla dowolnych elementów $x, y \in X$ z warunku $x\mathcal{R}y$ wynika, że $y\mathcal{R}x$. To oznacza, że relacja $\mathcal{R}$ jest symetryczna na X, jeśli

$$\mathcal{R} \subseteq \mathcal{R}^{-1};$$

- *antysymetryczną* na zbiorze X, jeśli dla dowolnych elementów $x, y \in X$ z tego, że $x\mathcal{R}y$ oraz $y\mathcal{R}x$ wynika, że $x = y$. W języku działań to znaczy, że przecięcie

$$\mathcal{R} \cap \mathcal{R}^{-1} = I_X;$$

- *przechodnią* (lub *tranzytywną*) na zbiorze X, jeśli dla dowolnych elementów $x, y, z \in X$ ze związków $x\mathcal{R}y$ oraz $y\mathcal{R}z$ wynika, że $x\mathcal{R}z$. Zatem $\mathcal{R}$ jest tranzytywne na X, jeśli

$$\mathcal{R} \circ \mathcal{R} \subseteq \mathcal{R}.$$

Lemat 1.3.4. *Relacja $\mathcal{R}$ jest symetryczna na zbiorze X wtedy i tylko wtedy, gdy $\mathcal{R} = \mathcal{R}^{-1}$.*

Dowód. ($\Rightarrow$) Z definicji wynika, że $\mathcal{R} \subseteq \mathcal{R}^{-1}$. Oprócz tego wyżej udowodniono (patrz twierdzenie 1.3.3(7)), że $\mathcal{R}^{-1} \subseteq (\mathcal{R}^{-1})^{-1}$, czyli w wyniku twierdzenia 1.3.3(1) mamy $\mathcal{R}^{-1} \subseteq \mathcal{R}$. Zatem $\mathcal{R} = \mathcal{R}^{-1}$.

($\Leftarrow$) Wynika z równości $\mathcal{R} = \mathcal{R}^{-1}$ i definicji relacji symetrycznej. □

Nietrudno też przekonać się, że sprawdza się takie

Twierdzenie 1.3.5. *Niech $\mathcal{R}$ oraz $\mathcal{S}$ będą relacjami na zbiorze X. Wtedy zachodzą następujące własności:*

(1) *jeśli $\mathcal{R}$ i $\mathcal{S}$ są zwrotne na X, to $\mathcal{R}\cup\mathcal{S}$, $\mathcal{R}\cap\mathcal{S}$, $\mathcal{R}^{-1}$, $\mathcal{R}\circ\mathcal{S}$ również są zwrotne na X;*

(2) *jeśli $\mathcal{R}$ i $\mathcal{S}$ są symetryczne na X, to $\mathcal{R}\cap\mathcal{S}$, $\mathcal{R}\cup\mathcal{S}$ oraz $\mathcal{R}^{-1}$ też są symetryczne na X;*

(3) *jeśli $\mathcal{R}$ i $\mathcal{S}$ są symetryczne na X, to ich iloczyn $\mathcal{R}\circ\mathcal{S}$ jest symetryczny na X w tym i tylko tym przypadku, gdy $\mathcal{R}\circ\mathcal{S}=\mathcal{S}\circ\mathcal{R}$;*

(4) *jeśli $\mathcal{R}$ i $\mathcal{S}$ są antysymetryczne na X, to relacje $\mathcal{R}\cap\mathcal{S}$ oraz $\mathcal{R}^{-1}$ także są antysymetryczne na X;*

(5) *jeśli $\mathcal{R}$ i $\mathcal{S}$ są przechodnie na X, to relacje $\mathcal{R}\cap\mathcal{S}$ i $\mathcal{R}^{-1}$ również są przechodnie na X.*

□

Przykład 1.3.6.

Na zbiorze Λ wszystkich ludzi:

- zwrotną jest relacja „być podobnym do", „być nie starszym", lecz relacje „być siostrą" oraz „być starszym" nie są zwrotnymi;
- symetrycznymi są relacje „być sąsiadem", „być podobnym do", „być krewnym", ale relacje „być siostrą" i „być młodszym" nie są symetrycznymi;
- przechodnimi są relacje „być spadkobiercą", „być bratem", „być siostrą", lecz „być krewnym" oraz „być ojcem" nie są przechodnimi.

■ Relacja $\mathcal{R}$ jest nazywana *relacją równoważności* na zbiorze X, jeśli jednocześnie jest zwrotna, symetryczna i przechodnia na X.

Przykłady 1.3.7.

(**1**) Na iloczynie kartezjańskim $\mathbb{N}\times\mathbb{N}$ rozpatrzmy relację „$\sim$":

$$(a,b)\sim(c,d)\Leftrightarrow a+d=b+c.$$

Wykażmy, że „$\sim$" jest relacją równoważności $\mathbb{N}\times\mathbb{N}$. Niech $x,y,z,t,u,v\in\mathbb{N}$. Skoro $x+y=y+x$, to $(x,y)\sim(x,y)$, czyli relacja „$\sim$" jest zwrotna. Z warunku $(x,y)\sim(z,t)$ wynika, że $x+t=y+z$, a więc $z+y=t+x$, czyli $(z,t)\sim(x,y)$ i relacja „$\sim$" jest symetryczna. Teraz załóżmy, że $(x,y)\sim(z,t)$ oraz $(z,t)\sim(u,v)$. Wtedy

$$x+t=y+z \quad\text{oraz}\quad z+v=t+u.$$

Sumując obie równości stronami, otrzymujemy, że

$$x+t+z+v=y+z+t+u,$$

a stąd $x+v=y+u$. To znaczy, że $(x,y)\sim(u,v)$ oraz „$\sim$" jest relacją przechodnią.

Zatem „$\sim$" jest relacją równoważności na zbiorze $\mathbb{N}\times\mathbb{N}$.

(**2**) Na zbiorze liczb wymiernych $\mathbb{Q}$ rozpatrzmy relację ρ, określoną tak:

$$x\rho y \Leftrightarrow \text{ istnieje takie } a\in\mathbb{Q},\ \text{że } x=y+a.$$

Wtedy $x=x+0$ dla każdego $x\in\mathbb{Q}$, co znaczy, że relacja ρ jest zwrotna.

Niech $x,y\in\mathbb{Q}$ oraz $x\rho y$. Wtedy $x=y+a$ dla pewnego $a\in\mathbb{Q}$, a więc $y=x+(-a)$ dla pewnego elementu $(-a)\in\mathbb{Q}$. To oznacza, że ρ jest symetryczne. Jeśli teraz $x,y,z\in\mathbb{Q}$, $x\rho y$ oraz $y\rho z$, to $x=y+a$ i $y=z+b$ dla pewnych $a,b\in\mathbb{Q}$. Stąd wynika, że

$$x=y+a=(z+b)+a=z+(b+a).$$

Skoro $a+b\in\mathbb{Q}$, to $x\rho z$ oraz ρ jest przechodnie.

Zatem ρ jest relacją równoważności na zbiorze $\mathbb{Q}$.

Łatwo sprawdzić, że zachodzi następny

Lemat 1.3.8 (kryterium równości klas równoważności)**.** *Niech ρ będzie relacją równoważności na zbiorze niepustym X oraz $a,b\in X$. Wtedy*

$$\rho(a)=\rho(b)$$

w tym i tylko tym przypadku, gdy $a\rho b$.

□

* * *

■ **Rozbicie zbioru i relacja równoważności.** *Rozbiciem* (lub *podziałem*) zbioru $X\neq\varnothing$ jest nazywana rodzina zbiorów

$$\{X_\alpha\}_{\alpha\in\Lambda}$$

(gdzie Λ jest pewnym zbiorem indeksów), która spełnia następujące cztery warunki:

(a) $X_\alpha\neq\varnothing$ dla każdego $\alpha\in\Lambda$;

(b) $X_\alpha\subseteq X$ dla każdego $\alpha\in\Lambda$;

(c)

$$X_\alpha\bigcap X_\beta=\varnothing \text{ lub } X_\alpha=X_\beta$$

dla dowolnych $\alpha,\beta\in\Lambda$;

(d)

$$\bigcup_{\alpha\in\Lambda}X_\alpha=X.$$

Przykłady 1.3.9.

(**1**) Niech $A = \{1, 2, 3, 4, 5\}$. Wtedy rodzina $\{\{1, 2\}, \{3\}, \{4, 5\}\}$ jest rozbiciem zbioru A.

(**2**) Zbiór studentów pierwszego roku (każdego) wydziału na uniwersytecie jest dzielony na ogół grup akademickich, które tworzą rozbicie tego zbioru.

(**3**) Zbiór liczb całkowitych $\mathbb{Z}$ możemy podzielić na dwie klasy w taki sposób: klasa liczb parzystych $2\mathbb{Z}$ i klasa liczb nieparzystych $1 + 2\mathbb{Z}$. Z warunków $2 \in 2\mathbb{Z}$ i $3 \in 1 + 2\mathbb{Z}$ wynika, ze każda z klas jest niepustą. Także $2\mathbb{Z} \subseteq \mathbb{Z}$ (słownie, każda liczba parzysta całkowita jest całkowitą), $1 + 2\mathbb{Z} \subseteq \mathbb{Z}$ (słownie, każda liczba nieparzysta całkowita jest całkowitą) oraz

$$2\mathbb{Z} \bigcap (1 + 2\mathbb{Z}) = \varnothing$$

(to znaczy, że nie istnieje liczba całkowita, która jednocześnie jest parzysta i nieparzysta). W końcu

$$\mathbb{Z} = 2\mathbb{Z} \bigcup (1 + 2\mathbb{Z})$$

(czyli dowolna liczba całkowita jest parzysta lub nieparzysta). Zatem rodzina $\{2\mathbb{Z}, 1 + 2\mathbb{Z}\}$ jest rozbiciem zbioru liczb całkowitych $\mathbb{Z}$.

Kolejne dwa twierdzenia charakteryzują związki między rozbiciami danego zbioru a relacjami równoważności na tym zbiorze.

Twierdzenie 1.3.10 (o rozbiciu generowanym przez relację równoważności)**.** *Niech ρ będzie relacją równoważności na zbiorze niepustym X. Wtedy rodzina klas równoważności (względem relacji ρ)*

$$\mathcal{P}_\rho = \{\rho(x) \mid x \in X\}$$

tworzy rozbicie zbioru X (i mówimy, że relacja równoważności ρ *generuje* rozbicie $\mathcal{P}_\rho$ zbioru X).

Dowód. Niech ρ będzie relacją równoważności na zbiorze X. Skoro ρ jest zwrotne na X, to $x\rho x$ dla każdego $x \in X$, a stąd $x \in \rho(x)$. To znaczy, że $\rho(x) \neq \varnothing$. Ponadto na podstawie definicji $\rho(x) \subseteq X$.

Chcemy udowodnić, że dla dowolnych $x, y \in X$ mamy

$$\rho(x) \bigcap \rho(y) = \varnothing \text{ lub } \rho(x) = \rho(y).$$

W tym celu załóżmy, że

$$z \in \rho(x) \bigcap \rho(y)$$

dla pewnych $x, y, z \in X$. Wtedy $z \in \rho(x)$ (czyli $z\rho x$) oraz $z \in \rho(y)$ (czyli $z\rho y$). W wyniku symetryczności relacji ρ z pierwszego związku wynika,

że $x\rho z$. A stąd na podstawie przechodniości relacji ρ dostajemy, że $x\rho y$, czyli $x \in \rho(y)$. Teraz dla dowolnego $u \in \rho(x)$ otrzymujemy, że $u\rho x$. Z powodu i w związku z $x\rho y$ mamy $u\rho y$, czyli $u \in \rho(y)$, a więc $\rho(x) \subseteq \rho(y)$. Podobnie możemy udowodnić i odwrotne zawieranie $\rho(y) \subseteq \rho(x)$ (co zostawiamy Czytelnikowi). Zatem

$$\rho(x) = \rho(y).$$

Mamy także $x \in \rho(x)$, a więc $X \subseteq \cup_{x\in X}\rho(x)$. Jako wniosek wynika z tego, że

$$X = \bigcup_{x\in X} \rho(x).$$

□

Twierdzenie 1.3.11 (o relacji równoważności indukowanej przez rozbicie)**.** *Jeśli*

$$\mathcal{P} = \{X_\alpha \mid \alpha \in \Lambda\}$$

jest rozbiciem zbioru X, a $\rho_\mathcal{P}$ jest relacją na zbiorze X taką, że

$$a\rho_\mathcal{P} b \Leftrightarrow \text{ znajdzie się taki indeks } \alpha \in \Lambda, \text{ że } a, b \in X_\alpha,$$

to $\rho_\mathcal{P}$ jest relacją równoważności na zbiorze X (i wtedy mówimy, że rozbicie $\mathcal{P}$ *indukuje* relację równoważności $\rho_\mathcal{P}$ na zbiorze X).

Dowód. Niech

$$\mathcal{P} = \{X_\alpha\}_{\alpha\in\Lambda}$$

będzie rozbiciem zbioru X. Wtedy

$$X = \bigcup_{\alpha\in\Lambda} X_\alpha,$$

a zatem dla każdego elementu $x \in X$ znajdzie się taki indeks $\alpha \in \Lambda$, że $x \in X_\alpha$. Lecz wtedy także $x, x \in X_\alpha$, czyli $x\rho_\mathcal{P} x$ oraz $\rho_\mathcal{P}$ jest zwrotne na zbiorze X.

Jeśli x, y są takimi elementami ze zbioru X, że $x\rho_\mathcal{P} y$, a więc $x, y \in X_\beta$ dla pewnego $\beta \in \Lambda$, to także $y, x \in X_\beta$, czyli $y\rho_\mathcal{P} x$. Zatem $\rho_\mathcal{P}$ jest symetryczne na zbiorze X.

Niech x, y, z będą takimi elementami z X, że $x\rho_{\mathcal{P}}y$ (czyli $x, y \in X_\beta$ dla pewnego indeksu β) oraz $y\rho_{\mathcal{P}}z$ (czyli $y, z \in X_\gamma$ dla pewnego indeksu γ). Wtedy $y \in X_\beta \cap X_\gamma$, a zatem

$$X_\beta \bigcap X_\gamma \neq \varnothing.$$

Biorąc pod uwagę, że $\mathcal{P}$ jest rozbiciem zbioru X, otrzymujemy równość $X_\beta = X_\gamma$. Zatem $x, z \in X_\beta$ oraz $x\rho_{\mathcal{P}}z$, czyli $\rho_{\mathcal{P}}$ jest przechodnie.

Wnosimy, że $\rho_{\mathcal{P}}$ jest relacją równoważności na zbiorze X. □

Kolejne dwa twierdzenia pokazują, że relacja równoważności, która jest indukowana przez rozbicie, jest określona jednoznacznie. Podobnie rozbicie, które jest generowane przez daną relację równoważności, też jest określone dokładnie jednoznacznie.

Twierdzenie 1.3.12. *Jeśli relacja równoważności ρ generuje rozbicie*

$$\mathcal{P}_\rho = \{\rho(x) \mid x \in X\}$$

zbioru X, to relacja równoważności, która jest indukowa przez rozbicie $\mathcal{P}_\rho$, jest równa ρ.

Dowód. Niech ρ będzie relacją równoważności na zbiorze X, generującą rozbicie $\mathcal{P}_\rho$ zbioru X. Załóżmy, że rozbicie $\mathcal{P}_\rho$ indukuje relację równoważności ρ^* na zbiorze X. Jeśli $(x, y) \in \rho$, to $x, y \in \rho(x)$. Lecz $\rho(x) \in \mathcal{P}_\rho$, a więc $x\rho^*y$. Zatem $\rho \subseteq \rho^*$. Odwrotnie, jeśli $(x, y) \in \rho^*$, to istnieje taki podzbiór $A \subseteq X$, że $A \in \mathcal{P}_\rho$ oraz $x, y \in A$. Skoro A jest klasą równoważności względem ρ, to $x\rho y$. Zatem $\rho^* \subseteq \rho$. Z tych informacji wynika, że

$$\rho = \rho^*.$$

□

Twierdzenie 1.3.13. *Jeśli rozbicie $\mathcal{P}$ zbioru X indukuje relację równoważności $\rho_{\mathcal{P}}$ na zbiorze X, to rozbicie zbioru X, które jest generowane przez relację ρ, jest równe $\mathcal{P}$.*

Dowód. Niech $\mathcal{P}$ będzie rozbiciem zbioru X, indukującym relacje równoważności $\rho_{\mathcal{P}}$ na zbiorze X, a relacja $\rho_{\mathcal{P}}$ generuje rozbicie $\mathcal{P}^*$ zbioru X. Wtedy podobnie (jak w twierdzeniu 1.3.12) możemy sprawdzić, że

$$\mathcal{P} = \mathcal{P}^*$$

(co zostawiamy Czytelnikowi jako ćwiczenie). □

■ Jeśli ρ jest relacją równoważności na X, to rozbicie zbioru X, które jest generowane przez relację ρ, będziemy oznaczać przez

$$X/\rho$$

i nazywać *zbiorem ilorazowym* zbioru X względem relacji równoważności ρ (krótko X *modulo* ρ lub X *przez* ρ).

Przykłady 1.3.14.

(**1**) (**Zbiór liczb całkowitych jako zbiór ilorazowy**) Rozpatrzmy relację równoważności „$\sim$” na zbiorze $\mathbb{N} \times \mathbb{N}$ (z przykładu 1.3.7 (1)). Wtedy klasa równoważności $(a,b)_\sim$ z reprezentantem $(a,b) \in \mathbb{N} \times \mathbb{N}$ jest zbiorem

$$(a,b)_\sim = \{(c,d) \in \mathbb{N}^2 \mid (c,d) \sim (a,b)\}$$

oraz dla dowolnego $a \in \mathbb{N}$ mamy:

$$\begin{array}{rcrcrcl} (0,0)_\sim & = & \{(c,d) \in \mathbb{N}^2 \mid c+0 = d+0\} & = & \{(x,x) \mid x \in \mathbb{N}\} & = & (a,a)_\sim, \\ (0,1)_\sim & = & \{(c,d) \in \mathbb{N}^2 \mid c+1 = d+0\} & = & \{(x,x+1) \mid x \in \mathbb{N}\} & = & (a,a+1)_\sim, \\ (1,0)_\sim & = & \{(c,d) \in \mathbb{N}^2 \mid c+0 = d+1\} & = & \{(x+1,x) \mid x \in \mathbb{N}\} & = & (a+1,a)_\sim. \end{array}$$

Rozumując podobnie, wnosimy, że dla dowolnych $a, b \in \mathbb{N}$:

$$(a,b)_\sim = \begin{cases} (a-b,0)_\sim, & \text{gdy } a > b \\ (0,b-a)_\sim, & \text{gdy } a < b. \end{cases}$$

Biorąc $(a,b)_\sim \leqslant (c,d)_\sim$ wtedy i tylko wtedy, gdy $a+d \leqslant b+c$, przekonujemy się, że dla $a \in \mathbb{N}^*$

$$(0,a)_\sim < (0,0)_\sim, (a,0)_\sim > (0,0)_\sim.$$

To zezwala traktować zbiór ilorazowy

$$\mathbb{N} \times \mathbb{N}/\sim = \{(0,0)_\sim, (a,0)_\sim, (0,a)_\sim \mid a \in \mathbb{N}\}$$

jako zbiór liczb całkowitych $\mathbb{Z}$ (przy tym

$$\begin{array}{rcl} (0,0)_\sim & = & 0 \in \mathbb{Z}, \\ (a,0)_\sim & = & a \in \mathbb{N}, \\ (0,a)_\sim & = & -a \in \{-a \mid a \in \mathbb{N}\}). \end{array}$$

Zbiór $\mathbb{N} \times \mathbb{N}/\sim$ jest rozbiciem zbioru $\mathbb{N} \times \mathbb{N}$.

(**2**) Rozpatrzmy relację równoważności ρ (z przykładu 1.3.7 (2)) na zbiorze liczb wymiernych $\mathbb{Q}$. Ze związków

$$\rho(0) = \{x \in \mathbb{Q} \mid x = 0 + a \text{ dla pewnego } a \in \mathbb{Q}\} = \{x \in \mathbb{Q} \mid x = 0 + x\} = \mathbb{Q}$$

oraz

$$\rho(b) = \{x \in \mathbb{Q} \mid x = b + a \text{ dla pewnego } a \in \mathbb{Q}\ \} = \{x \in \mathbb{Q} \mid x = b + (x - b)\} = \mathbb{Q}$$

wnosimy, że $\rho(b) = \rho(0)$ dla dowolnego elementu $b \in \mathbb{Q}$. To oznacza, że zbiór ilorazowy

$$\mathbb{Q}/\rho = \{\mathbb{Q}\}$$

jest jednoelementowym rozbiciem zbioru liczb wymiernych $\mathbb{Q}$.

* * *

Przytoczmy jeszcze dwa ważne przykłady rozbić.

■ **Klasy reszt modulo** n**.** Niech n będzie ustaloną liczbą dodatnią całkowitą[(14)] oraz

$$n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}.$$

Wyznaczmy relację kongruencji modulo „$\equiv_n$" (lub krótko „$\equiv$") modulo n na zbiorze liczb całkowitych $\mathbb{Z}$ w taki sposób:

$$a \equiv b \pmod{n} \quad \Leftrightarrow \quad n \mid (a - b)$$

lub równoważnie

$$a \equiv b \pmod{n} \quad \Leftrightarrow \quad a - b \in n\mathbb{Z}.$$

Przy tym liczby a i b są nazywane *kongruentnymi* (albo *równymi*, albo *przystającymi*) modulo n.

Lemat 1.3.15. *Relacja „$\equiv_n$" jest relacją równoważności na zbiorze $\mathbb{Z}$ liczb całkowitych.*

Dowód. Niech a, b, c będą dowolnymi liczbami całkowitymi. Skoro $a - a = 0$ dla każdej liczby całkowitej a i $n \mid 0$, to $a \equiv a \pmod{n}$, a więc „$\equiv_n$" jest zwrotne. Jeśli $a \equiv b \pmod{n}$, czyli $n \mid (a - b)$, to $a - b = nk$ dla pewnej liczby całkowitej k, a stąd

$$b - a = n(-k),$$

(14) Warunek $n > 0$ jest spowodowany tym, że $(-n)\mathbb{Z} = n\mathbb{Z}$ oraz $0\mathbb{Z} = \{0\}$.

czyli $n \mid (b-a)$ oraz $b \equiv a \pmod{n}$. Zatem „$\equiv_n$" jest relacją symetryczną. Załóżmy, że $a \equiv b \pmod{n}$ oraz $b \equiv c \pmod{n}$, czyli $n \mid (a-b)$ oraz $n \mid (b-c)$. Wtedy znajdą się takie liczby całkowite k_1 oraz k_2, że $a-b = nk_1$ i $b-c = nk_2$. Sumując obie równości stronami, otrzymujemy

$$a - c = (a - b) + (b - c) = nk_1 + nk_2 = n(k_1 + k_2),$$

czyli $n \mid (a-c)$ lub, co jest tym samym, $a \equiv c \pmod{n}$. To oznacza, że „$\equiv_n$" jest przechodnie.

Zatem udowodniliśmy, że „$\equiv_n$" jest relacją równoważności na zbiorze liczb całkowitych $\mathbb{Z}$.

□

Wniosek 1.3.16. *Niech $a \in \mathbb{Z}$. Na zbiorze liczb całkowitych $\mathbb{Z}$ klasa równoważności $\overline{a}$ z reprezentantem a (względem „$\equiv_n$") składa się ze wszystkich liczb całkowitych postaci*

$$a + nk, \ \textit{gdzie } k \in \mathbb{Z}.$$

Dowód. W rzeczy samej, jeśli $b \equiv a \pmod{n}$, to istnieje taka liczba całkowita k, że $b = a + nk$. Odwrotnie, jeśli $c = a + ns$ dla pewnego $s \in \mathbb{Z}$, to $c - a = ns$, a więc $n \mid (c-a)$, czyli

$$c \equiv a \pmod{n}.$$

□

Konwencja. Dalej klasę równoważności relacji „$\equiv_n$" na zbiorze liczb całkowitych $\mathbb{Z}$ z reprezentantem a będziemy oznaczać na jeden z równoważnych sposobów:

$$\overline{a}, \ [a], \ [a]_n \text{ lub } \ a_n$$

i nazywać *klasą reszt* liczby a modulo n (lub *klasą kongruentności* liczby a modulo n), czyli

$$\overline{a} = \{a + nk \mid k \in \mathbb{Z}\} = \{a, a \pm n, a \pm 2n, \ldots\}.$$

Wniosek 1.3.17. *Niech n będzie dodatnią liczbą całkowitą. Wtedy zbiór*

$$\{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

jest rozbiciem zbioru liczb całkowitych $\mathbb{Z}$; w szczególności

$$\mathbb{Z} = \overline{0} \bigcup \overline{1} \bigcup \cdots \bigcup \overline{n-1}. \tag{1.2}$$

Dowód. Oczywiście, że $\overline{k} \subseteq \mathbb{Z}$ dla każdej liczby całkowitej k. Skoro $k-k = 0$ oraz $n \mid 0$, to $k \in \overline{k}$, a więc $\overline{k} \neq \varnothing$. Na mocy twierdzenia o dzieleniu z resztą dla każdej liczby całkowitej a znajdą się takie liczby całkowite q, r, że $a = n \cdot q + r$, gdzie $0 \leqslant r \leqslant n-1$, a więc $a = r + nq$. Stąd wynika, że $a \in \overline{r}$. Zatem zachodzi (1.2).

Niech k i s będą takimi różnymi liczbami całkowitymi, że $0 \leqslant k < s \leqslant n-1$. Jeśli $u \in \overline{k} \cap \overline{s}$, to $u = k + nl_1$ oraz $u = s + nl_2$ dla pewnych liczb całkowitych l_1 i l_2. Wtedy

$$0 = u - u = (k - s) + n(l_1 - l_2),$$

a stąd otrzymujemy, że $n \mid (s-k)$, co nie jest możliwe, bo $0 < s-k < n$. Zatem

$$\overline{k} \bigcap \overline{s} = \varnothing$$

dla różnych liczb k, s takich, że $0 \leqslant k < s \leqslant n-1$. □

Wniosek 1.3.18 (kryterium równości klas reszt)**.** *Niech $r_1, r_2 \in \mathbb{Z}$. Dwie klasy reszt $\overline{r}_1$ oraz $\overline{r}_2$ modulo n są równe wtedy i tylko wtedy, gdy*

$$n \mid (r_1 - r_2). \tag{1.3}$$

Dowód. ($\Rightarrow$) Niech $\overline{r}_1 = \overline{r}_2$. Ponieważ $r_1 - r_1 = 0$ oraz $n \mid 0$, to $r_1 \in \overline{r}_1$. Zatem $r_1 \in \overline{r}_2$ i, biorąc pod uwagę definicję, $r_1 = r_2 + nk$ dla pewnego $k \in \mathbb{Z}$, skąd otrzymujemy (1.3).

($\Leftarrow$) Jeśli $n \mid (r_1 - r_2)$, to istnieje taka liczba całkowita s, że $r_1 - r_2 = ns$. Stąd $r_1 = r_2 + ns$ (czyli $r_1 \in \overline{r}_2$). Oprócz tego $r_1 \in \overline{r}_1$, a zatem

$$\overline{r}_1 \bigcap \overline{r}_2 \neq \varnothing.$$

Z wniosku 1.3.17 wynika, że $\overline{r}_1 = \overline{r}_2$. □

Zatem mamy następujący

Wniosek 1.3.19. *Rozbicie zbioru $\mathbb{Z}$ liczb całkowitych, które jest generowane przez relację równoważności „$\equiv_n$", składa się z n parami różnych elementów:*

$$\mathbb{Z}/\equiv_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

□

Konwencja. Dalej zbiór $\{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ klas reszt modulo n będziemy krótko oznaczać przez

$$\mathbb{Z}_n.$$

Przykład 1.3.20.

Niech $n = 3$. Wtedy $\mathbb{Z}_3 = \{\overline{0}, \overline{1}, \overline{2}\}$, gdzie $\overline{0}$ składa się z liczb całkowitych krotnych 3; $\overline{1}$ jest zbiorem liczb całkowitych, które przy dzieleniu przez 3 dają resztę 1; a $\overline{2}$ składa się z liczb całkowitych dzielących się przez 3 z resztą równą 2.

* * *

■ **Liczby wymierne.** Niech $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ będzie zbiorem niezerowych liczb całkowitych. Na iloczynie kartezjańskim $\mathbb{Z} \times \mathbb{Z}^*$ rozpatrzmy relację ρ, określoną przez regułę

$$(x, y)\rho(u, v) \Leftrightarrow x \cdot v = y \cdot u \tag{1.4}$$

dla liczb $x, u \in \mathbb{Z}$ oraz $y, v \in \mathbb{Z}^*$.

Lemat 1.3.21. *ρ jest relacją równoważności na zbiorze $\mathbb{Z} \times \mathbb{Z}^*$.*

Dowód. Skoro $ab = ba$ dla dowolnych elementów $a \in \mathbb{Z}$ i $b \in \mathbb{Z}^*$, to $(a, b)\rho(a, b)$. Jeśli $(a, b)\rho(c, d)$ dla pewnych $a, c \in \mathbb{Z}$ oraz $b, d \in \mathbb{Z}^*$, to $ad = bc$, a stąd

$$cb = da,$$

czyli $(c, d)\rho(a, b)$ i relacja ρ jest symetryczna.

Załóżmy teraz, że $(a,b)\rho(c,d)$ oraz $(c,d)\rho(e,f)$ dla pewnych $a,c,e \in \mathbb{Z}$ oraz $b,d,f \in \mathbb{Z}^*$. Wtedy $ad = bc$ oraz $cf = de$. Mnożąc obie równości stronami, otrzymujemy, że $adcf - bcde = 0$, skąd

$$dc(af - be) = 0.$$

Ponieważ $d \neq 0$, to na podstawie ostatniej równości otrzymujemy, że

$$c(af - be) = 0.$$

Jeśli $c = 0$, to $ad = b0 = 0$ oraz $0 = 0f = de$, a zatem $a = 0$ i $e = 0$. Wtedy $af = 0 = be$ oraz $(a,b)\rho(e,f)$. Jeśli zaś $c \neq 0$, to mamy $af - be = 0$, w wyniku czego

$$(a,b)\rho(e,f),$$

czyli ρ jest relacją przechodnią. Podsumowując, wnosimy, że ρ jest relacją równoważności na zbiorze $\mathbb{Z} \times \mathbb{Z}^*$. □

Konwencja. Zbiór ilorazowy $(\mathbb{Z}\times\mathbb{Z}^*)/\rho$ będziemy oznaczać przez $\mathbb{Q}$, a klasę równoważności $\rho((a,b))$ z reprezentantem (a,b) przez

$$\frac{a}{b}$$

i nazywać ją *liczbą wymierną* lub *ułamkiem* z *licznikiem* a oraz *mianownikiem* b.

Na przykład

$$\tfrac{1}{2} = \{(c,d) \in \mathbb{Z} \times \mathbb{Z}^* \mid d = 2c\} = \\ = \{\tfrac{1}{2}, \tfrac{2}{4}, \tfrac{3}{6}, \ldots\} = \tfrac{2}{4} = \tfrac{3}{6} = \cdots = \tfrac{n}{2n},$$

gdzie $n \in \mathbb{Z}^*$.

Wniosek 1.3.22. *Klasa równoważności $\rho((a,b))$ z reprezentantem $(a,b) \in \mathbb{Z} \times \mathbb{Z}^*$ względem relacji ρ na zbiorze $\mathbb{Z} \times \mathbb{Z}^*$ (patrz (1.4)) składa się z par liczb całkowitych postaci*

$$(az, bz),$$

gdzie $z \in \mathbb{Z}^*$.

* * *

■ **Zbiory uporządkowane.** Niech $\mathcal{R}$ będzie relacją na zbiorze X. Relacja $\mathcal{R}$ jest nazywana relacją *porządku* (relacją *częściowego porządku* lub *porządkiem częściowym*) na zbiorze X, jeśli ona jest jednocześnie zwrotna, antysymetryczna i przechodnia. Jeśli $\mathcal{R}$ jest relacją częściowego porządku na zbiorze X, to mówią, że X jest zbiorem *częściowo uporządkowanym* względem $\mathcal{R}$ (lub równoważnie $(X, \mathcal{R})$ jest zbiorem częściowo uporządkowanym). Zbiór częściowo uprządkowany $(X, \mathcal{R})$ jest nazywany *liniowo uporządkowanym*, jeśli relacja $\mathcal{R}$ jest *spójną*, czyli dla dowolnych elementów $x, y \in X$ zachodzi $x = y$ albo $x\mathcal{R}y$, albo $y\mathcal{R}x$; w tym przypadku $\mathcal{R}$ jest nazywane *liniowym porządkiem* na zbiorze X.

■ Podzbiór A zbioru częściowo uporządkowanego $(X, \mathcal{R})$ jest nazywany:

- *łańcuchem*, jeśli $a = b$ albo $a\mathcal{R}b$, albo $b\mathcal{R}a$ dla dowolnych elementów $a, b \in A$;
- *ograniczonym z góry*, jeśli istnieje taki element $m \in X$, że $a\mathcal{R}m$ dla wszystkich elementów $a \in A$;
- *ograniczonym z dołu*, jeśli istnieje taki element $M \in X$, że $M\mathcal{R}a$ dla wszystkich elementów $a \in A$.

Przykłady 1.3.23.

(**1**) Relacja „być mniejszym lub równym" (krótko „$\leqslant$") na zbiorze liczb naturalnych $\mathbb{N}$ jest porządkiem liniowym. W rzeczy samej, $x \leqslant x$ dla każdej liczby $x \in \mathbb{N}$ (zwrotność); oraz jeśli $x \leqslant y$ i $y \leqslant x$ dla pewnych $x, y \in \mathbb{N}$, to $x = y$ (antysymetryczność). Oprócz tego z $x \leqslant y$ i $y \leqslant z$ dla liczb naturalnych x, y, z zawsze wynika, że $x \leqslant z$ (przechodniość).

Zanotujmy, że relacja „być mniejszym" (krótko „$<$") nie jest porządkiem na zbiorze liczb naturalnych $\mathbb{N}$, bo dla żadnej liczby naturalnej x nie zachodzi, że $x < x$ (a to znaczy, że relacja „$<$" nie jest zwrotna).

(**2**) Relacja „być dzielnikiem" na zbiorze liczb naturalnych $\mathbb{N}$ jest relacją częściowego porządku. W rzeczy samej, $x \mid x$ dla każdego $x \in \mathbb{N}$ (zwrotność). Jeśli $y \mid x$ oraz $x \mid y$, to za lematem 1.2.1(14) $x = y$ (antysymetryczność). Także z $y \mid x$ oraz $z \mid y$ dla liczb naturalnych x, y, z zawsze wynika, że $z \mid x$ (przechodniość). Zatem „$\mid$" jest porządkiem częściowym na zbiorze liczb naturalnych $\mathbb{N}$.

(**3**) Niech $\mathcal{B}(A)$ będzie zbiorem wszystkich podzbiorów pewnego zbioru niepustego A. Wtedy dla dowolnych podzbiorów $X, Y \subseteq A$ mamy $X \subseteq X$ (zwrotność relacji „$\subseteq$") oraz z $X \subseteq Y$ i $Y \subseteq X$ wynika, że $X = Y$ (antysymetryczność). W końcu, jeśli $X \subseteq Y$ i $Y \subseteq Z$, to łatwo wyciągamy, że $X \subseteq Z$ (przechodniość). Zatem „$\subseteq$" jest relacją częściowego porządku na booleanie $\mathcal{B}(A)$.

Konwencja. Dalej relację częściowego porządku na dowolnym zbiorze niepustym X najczęściej będziemy oznaczać przez „$\leqslant$" (chociaż to niekoniecznie musi być relacja nierówności dla liczb). Oznaczenie $x < y$ określa, że $x \leqslant y$ oraz $x \neq y$.

■ Niech $(X, \leqslant)$ będzie zbiorem częściowo uporządkowanym, A będzie jej podzbiorem. Element $a \in X$ jest nazywany:

- *największym* w zbiorze X, jeśli $x \leqslant a$ dla wszystkich elementów $x \in X$;
- *maksymalnym* w X, jeśli z $a \leqslant x$ dla $x \in X$ wynika, że $x = a$;
- *najmniejszym* w X, jeśli $a \leqslant x$ dla wszystkich elementów $x \in X$;
- *minimalnym* w zbiorze X, jeśli z $x \leqslant a$ dla $x \in X$ wynika, że $x = a$;
- *majorantą* podzbioru A, jeśli $x \leqslant a$ dla każdego $x \in A$;
- *minorantą* podzbioru A, jeśli $a \leqslant x$ dla każdego elementu $x \in A$;
- *kresem dolnym* zbioru A, jeśli a jest minorantą zbioru A oraz $a_1 \leqslant a$ dla każdej innej minoranty a_1 zbioru A (w tym przypadku zapisujemy $a = \inf A$);
- *kresem górnym* podzbioru A, jeśli a jest majorantą zbioru A oraz $a \leqslant a_1$ dla każdej innej majoranty a_1 zbioru A (w tym przypadku piszemy $a = \sup A$).

Przykłady 1.3.24.

(**1**) Na zbiorze liczb naturalnych niezerowych $\mathbb{N}^*$ rozpatrzmy relację „$\leqslant$”, określoną przez regułę

$$n \leqslant m \Leftrightarrow n \mid m.$$

Wtedy „$\leqslant$” jest relacją częściowego porządku. Lecz ani $2 \leqslant 3$, ani $3 \leqslant 2$ nie zachodzi. To znaczy, że „$\leqslant$” nie jest porządkiem liniowym na zbiorze $\mathbb{N}^*$. Podzbiór

$$A = \{3^k \mid k \in \mathbb{N}^*\}$$

jest łańcuchem w $\mathbb{N}^*$ (względem rozpatrywanej relacji).

(**2**) Niech

$$\mathbb{R}^\infty = \{(a_1, a_2, \ldots, a_n, \ldots) \mid a_i \in \mathbb{R}\ (i \in \mathbb{N}^*)\}$$

będzie zbiorem ciągów nieskończonych, elementami których są liczby rzeczywiste. Reguła

$$(a_1, a_2, \ldots, a_n, \ldots) \leqslant (b_1, b_2, \ldots, b_n, \ldots) \Leftrightarrow \forall_{n \in \mathbb{N}^*} : a_n \leqslant b_n$$

określa relację częściowego porządku na zbiorze $\mathbb{R}^\infty$. Lecz ta relacja „$\leqslant$” nie jest relacją porządku liniowego na zbiorze $\mathbb{R}^\infty$, bo ani $(0, 1, 0, 1, \ldots) \leqslant (1, 0, 1, 0, \ldots)$, ani $(1, 0, 1, 0, \ldots) \leqslant (0, 1, 0, 1, \ldots)$ nie zachodzą.

(**3**) Zbiór $(\mathbb{Q}, \leqslant)$ jest liniowo uporządkowany (względem zwykłej nierówności „$\leqslant$” dla liczb), lecz nie posiada ani minimalnych, ani maksymalnych elementów. Natomiast zbiór $(\mathbb{N}, \leqslant)$ posiada element minimalny 1, lecz nie posiada żadnego elementu maksymalnego.

(**4**) Niech $(X, \leqslant)$ będzie zbiorem częściowo uporządkowanym oraz $n \in \mathbb{N}^*$. Na n-tej potędze kartezjańskiej X^n rozpatrzmy relację „$\preceq$”:

$$(x_1, \ldots, x_n) \preceq (y_1, \ldots, y_n) \Leftrightarrow \text{albo } (x_1, \ldots, x_n) = (y_1, \ldots, y_n), \text{ albo znajdzie się taki indeks } s \in \{1, \ldots, n\}, \text{ że } x_1 = y_1, \ldots, x_{s-1} = y_{s-1}, x_s < y_s.$$

Wtedy „$\preceq$" jest relacją porządku na zbiorze X^n (która jest nazywana *porządkiem leksykograficznym*).

(**5**) Równość „$=$" jest relacją równoważności oraz jednocześnie relacją porządku na każdym zbiorze niepustym X.

■ Jeśli $(X, \leqslant)$ jest zbiorem częściowo uporządkowanym oraz $\varnothing \neq A \subseteq X$, to:

(a) jeśli m jest elementem największym zbioru A, to:
  (a_1) m jest jedynym elementem największym zbioru A,
  (a_2) m jest kresem górnym zbioru A;

(b) jeśli M jest elementem najmniejszym zbioru A, to:
  (b_1) M jest jedynym elementem najmniejszym zbioru A,
  (b_2) M jest kresem dolnym zbioru A.

Przykłady 1.3.25.

(**1**) Niech

$$A = \{r \in \mathbb{R} \mid r < \pi\} \subseteq \mathbb{R}.$$

Wtedy podzbiór A jest ograniczony z góry liczbą $\pi \in \mathbb{R}$, lecz nie posiada elementu maksymalnego. Liczba π jest kresem górnym zbioru A.

(**2**) Niech $X \neq \varnothing$ oraz $\varnothing \neq A \subseteq 2^X$. Łatwo zauważyć, że

$$\begin{array}{rcl} \inf A & = & \bigcap\{Y \mid Y \in A\}, \\ \sup A & = & \bigcup\{Y \mid Y \in A\}. \end{array}$$

(**3**) Niech

$$X = \{(-\infty, -n), (-3, 3), (n, +\infty) \mid n \in \mathbb{N}^*\}$$

będzie zbiorem częściowo uporządkowanym względem zawierania „$\subseteq$". Wtedy $(-3, 3)$ jest jednocześnie jedynym minimalnym i jedynym maksymalnym elementem zbioru X.

* * *

■ **Pewnik wyboru.** Niech X_α będzie zbiorem niepustym dla każdego $\alpha \in \Lambda$. Wtedy istnieje taka funkcja *wyboru*

$$f : \Lambda \to \bigcup_{\alpha \in \Lambda} X_\alpha,$$

że $f(\alpha) \in X_\alpha$ dla dowolnego $\alpha \in \Lambda$.

W inny sposób pewnik wyboru brzmi tak: *dla każdej rodziny* $\{X_\alpha\}_{\alpha \in \Lambda}$ *zbiorów niepustych i parami rozłącznych istnieje zbiór* X*, który z każdym* X_α *(*$\alpha \in \Lambda$*) ma dokładnie jeden element wspólny.*

Pewnik wyboru jest równoważny z takim lematem z 1922 r.

Lemat 1.3.26 (Kuratowskiego[15]-Zorna[16]). *Zbiór częściowo uporządkowany, którego każdy łańcuch posiada kres górny (odpowiednio kres dolny), zawiera element maksymalny (odpowiednio minimalny).*

Pewnik wyboru jest równoważny z następnym twierdzeniem.

Twierdzenie 1.3.27 (Tarskiego[17]). *Iloczyn kartezjański dowolnej rodziny zbiorów niepustych jest niepusty.*

Ćwiczenia 1.3.28.

(**1**) Dla każdego zbioru X i dla każdej relacji $\mathcal{R} \subseteq X \times X$ sprawdzić, czy $\mathcal{R}$ jest relacją równoważności; jeżeli jest, to znaleźć klasy równoważności, jeśli:
(a) $\mathbb{N}^*$ jest zbiorem dodatnich liczb naturalnych, gdzie $x\mathcal{R}y \Leftrightarrow 3 \mid (x-y)$;
(b) $X = \{1, 2, \ldots, 25\}$, gdzie $x\mathcal{R}y \Leftrightarrow 5 \mid (x^2 - y^2)$;
(c) $X = \mathbb{N} \times \mathbb{N}$, gdzie $(r,s)\mathcal{R}(t,u) \Leftrightarrow r+u = s+t$;
(d) $X = \mathbb{N} \times \mathbb{N}^*$, gdzie $(a,b)\mathcal{R}(c,d) \Leftrightarrow ad = bc$;
(e) $X = \mathbb{R}^2$, gdzie $(x_1, y_1)\mathcal{R}(x_2, y_2) \Leftrightarrow {x_1}^2 + {y_1}^2 = {x_2}^2 + {y_2}^2$;
(f) $X = \mathbb{Z}$, gdzie $x\mathcal{R}y \Leftrightarrow x^4 \leqslant y^4$;
(g) $X = \{1, 2, 3, 4\}$, gdzie $x\mathcal{R}y \Leftrightarrow x^4 \leqslant y^4$;
(h) $X = \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x - y = 5$.
(**2**) Udowodnić dla relacji $\mathcal{A}, \mathcal{B}, \mathcal{C}$ na zbiorze X, że:
(a) $(\mathcal{A} \cap \mathcal{B}) \circ \mathcal{C} = (\mathcal{A} \circ \mathcal{C}) \cup (\mathcal{B} \circ \mathcal{C})$;
(b) $\mathcal{A} \circ \mathcal{A}^{-1} \supseteq I_X$ oraz $\mathcal{A}^{-1} \circ \mathcal{A} \supseteq I_X$;
(c) $(\mathcal{A} \cap \mathcal{B}) \circ \mathcal{C} \subseteq (\mathcal{A} \circ \mathcal{C}) \cap (\mathcal{B} \circ \mathcal{C})$;
(d) $(\mathcal{A} \cap \mathcal{B})^{-1} = \mathcal{A}^{-1} \cap \mathcal{B}^{-1}$;
(e) $(\mathcal{A} \cup \mathcal{B})^{-1} = \mathcal{A}^{-1} \cup \mathcal{B}^{-1}$;
(f) jeśli $\mathcal{A} \subseteq \mathcal{B}$, to $\mathcal{A} \circ \mathcal{C} \subseteq \mathcal{B} \circ \mathcal{C}$ i $\mathcal{C} \circ \mathcal{A} \subseteq \mathcal{C} \circ \mathcal{B}$.
(**3**) Sprawdzić, jakie własności posiada relacja $\mathcal{R}$, jeśli:
(a) $\mathcal{R} \subseteq \mathbb{N} \times \mathbb{N}$, gdzie $x\mathcal{R}y \Leftrightarrow 2 \mid (x+y)$;
(b) $\mathcal{R} \subseteq \mathbb{N} \times \mathbb{N}$, gdzie $x\mathcal{R}y \Leftrightarrow \{x \neq 0$ oraz $x \mid y\}$;
(c) $\mathcal{R} \subseteq \mathbb{Z} \times \mathbb{Z}$, gdzie $x\mathcal{R}y \Leftrightarrow \{x = 20$ oraz $y = 31\}$;
(d) $\mathcal{R} \subseteq \mathbb{N} \times \mathbb{N}$, gdzie $x\mathcal{R}y \Leftrightarrow x \cdot y = 6$;
(e) $\mathcal{R} \subseteq \mathbb{N}^{*3} \times \mathbb{N}^{*3}$, gdzie $(x,y,z)\mathcal{R}(a,b,c) \Leftrightarrow \{x = a$ oraz $y = c$ oraz $z = b\}$;
(f) $\mathcal{R} \subseteq \mathbb{Q}^2 \times \mathbb{Q}^2$, gdzie $(x,y)\mathcal{R}(a,b) \Leftrightarrow x \cdot b = y \cdot a$;
(g) $\mathcal{R} \subseteq \mathbb{Q} \times \mathbb{Q}$, gdzie $a\mathcal{R}b \Leftrightarrow |a-2| = |b+2|$;
(h) $\mathcal{R} \subseteq \mathbb{Q} \times \mathbb{Q}$, gdzie $x\mathcal{R}y \Leftrightarrow y = x+3$;
(k) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow e^x = 3e^y$;
(l) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x - y \in \mathbb{N}$;
(m) $X = \{1, 2, \ldots, 25\}$, gdzie $x\mathcal{R}y \Leftrightarrow 5 \mid (x^2 - y^2)$;
(n) $\mathcal{R} \subseteq \mathbb{Z} \times \mathbb{Z}$, gdzie $x\mathcal{R}y \Leftrightarrow |x| + |y| \neq 7$;
(o) $\mathcal{R} \subseteq \mathbb{Z} \times \mathbb{Z}$, gdzie $x\mathcal{R}y \Leftrightarrow |x| + |y| = 7$;
(p) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x^2 = y^5$;
(q) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x^2 \neq y^5$;
(r) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x^2 + yr \leqslant 25$;
(s) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow |x + y + 1| = 1$;

---

[15] Kazimierz Kuratowski (1896–1980)
[16] Max August Zorn (1906–1993)
[17] Alfred Tarski (1901–1983)

(t) $\mathcal{R} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$, gdzie $(x_1, x_2)\mathcal{R}(y_1, y_2) \Leftrightarrow x_1 = y_2$;
(u) $\mathcal{R} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$, gdzie $(x_1, x_2)\mathcal{R}(y_1, y_2) \Leftrightarrow x_1 = y_2$;
(v) $\mathcal{R} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$, gdzie $(x_1, x_2)\mathcal{R}(y_1, y_2) \Leftrightarrow x_1 = y_1$;
(w) $\mathcal{R} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$, gdzie $(x_1, x_2)\mathcal{R}(y_1, y_2) \Leftrightarrow x_1 = x_2$;
(z) $\mathcal{R} \subseteq \mathbb{R} \times \mathbb{R}$, gdzie $x\mathcal{R}y \Leftrightarrow x - y \notin \mathbb{Q}$.
(**4**) Na zbiorze X są określone relacje $\mathcal{R}$ oraz $\mathcal{S}$:
(a) $X = \mathbb{N}$, $\mathcal{R} = \{(1,2), (2,2), (2,3), (2,4), (6,6)\}$ oraz $\mathcal{S} = \{(1,2), (2,2), (2,3), (6,8), (11,3)\}$;
(b) $X = \mathbb{R}$, $\mathcal{R} = \{(x,y) \mid x + 2y - 2 = 0\}$ oraz $\mathcal{S} = \{(x,y) \mid x - 3y - 3 = 0\}$.
Znaleźć $\mathcal{R} \circ \mathcal{S}$, $\mathcal{S} \circ \mathcal{R}$, $(\mathcal{R} \circ \mathcal{S})^{-1}$, $(\mathcal{S} \circ \mathcal{R})^{-1}$, $\mathcal{R}^{-1}$, $\mathcal{S}^{-1}$, $\mathcal{R} \cup \mathcal{S}$ oraz $\mathcal{R} \cap \mathcal{S}$.

Uwagi. Matematyk i mechanik francuski J. Lagrange[18] pierwszy wykorzystał termin „równoważność" w jednej swojej pracy z 1773 r. Symbolu „$\equiv$" dla oznaczenia kongruencji pierwszy użył matematyk niemiecki C. Gauss[19] w słynnej pracy *Disquisitiones Arithmeticae* z 1801 r. Właśnie ta data jest uważana za datę narodzin współczesnego pojęcia „relacja równoważności". Podstawy współczesnej terminologii stosowanej do relacji zostały opracowane wspólnie przez matematyków angielskich B. Russela[20] oraz A. Whiteheada[21] w 1910 r.

[18] Joseph Louis de Lagrange (Giuseppe Lodovico) (1736–1813)
[19] Carl Friedrich Gauss (1777–1855)
[20] Bertrand Russel (1872–1970)
[21] Alfred North Whitehead (1861–1947)

1.4. Odwzorowania

Pojęcie odwzorowania (=funkcji) jest jednym z podstawowych we współczesnej matematyce.

■ Niech X i Y będą dowolnymi zbiorami niepustymi. *Odwzorowaniem* zbioru X w zbiór Y nazywamy regułę f, za pomocą której każdemu elementowi x ze zbioru X zestawiamy dokładnie jeden element y ze zbioru Y. Przy tym:

- zbiór X nazywamy *dziedziną* odwzorowania f (i zapisujemy $X = D(f)$ lub $X = D_f$);
- zbiór Y nazywamy *przeciwdziedziną* odwzorowania f;
- element y nazywamy *obrazem* elementu x (i zapisujemy $y = f(x)$ lub $f : x \mapsto y$); mówimy również, że y jest *wartością* funkcji f w punkcie x;

■ Odwzorowanie f zbioru X w zbiór Y oznaczamy przez

$$f : X \to Y$$

lub

$$X \xrightarrow{f} Y$$

(i mówimy, że f *odwzoruje zbiór* X *w zbiór* Y). W analizie odwzorowanie f zbioru X w zbiór Y jest nazywane *funkcją* określoną na zbiorze X o wartościach w zbiorze Y. Z każdym odwzorowaniem $f : X \to Y$ są związane takie zbiory:

- $\operatorname{Im} f = \{f(x) \mid x \in X\}$ jest *obrazem* odwzorowania f (lub *zbiorem wartości funkcji* f);
- $f^{-1}(y) = \{x \in X \mid f(x) = y\}$ jest *przeciwobrazem* elementu $y \in Y$ względem f;
- jeśli $W \subseteq Y$, to zbiór

$$f^{-1}(W) = \{x \in X \mid f(x) \in W\}$$

jest nazywany *przeciwobrazem* podzbioru W względem f. Wtedy $f^{-1}(\{y\}) = f^{-1}(y)$ dla $y \in Y$.

Przykłady 1.4.1.

(**1**) Niech $X = \{a, b, c, d\}$, $Y = \{1, 2, 3\}$ oraz

$$\begin{cases} f : X \to Y, \\ f(a) = 1, \\ f(b) = 3, \\ f(d) = 2, \end{cases} \quad \begin{cases} g : X \to Y, \\ g(a) = 2, \\ g(b) = 1, \\ g(c) = 2, \\ g(d) = 3, \end{cases}$$

$$\begin{cases} h : X \to Y, \\ h(a) = 1, \\ h(b) = 1, \\ h(c) = 2, \\ h(d) = 2, \end{cases} \quad \begin{cases} k : Y \to X, \\ k(1) = a, \\ k(2) = d, \\ k(3) = b. \end{cases}$$

Reguła f nie jest odwzorowaniem, bo nie istnieje obraz $f(c)$ dla elementu c z dziedziny X. Reguła $g : X \to Y$ jest odwzorowaniem, przy czym $\operatorname{Im} g = Y$ oraz $f^{-1}(2) = \{a, c\}$. Odwzorowaniami są też $h : X \to Y$ i $k : Y \to X$, przy czym

$$\operatorname{Im} k = \{a, b, d\} \neq X, \ \operatorname{Im} h = \{1, 2\} \neq Y, \ h^{-1}(3) = \varnothing, \ k^{-1}(c) = \varnothing.$$

(**2**) Reguła $f : \mathbb{R} \to \{-1, 0, 1\}$, gdzie

$$f(m) = \operatorname{sgn} m \quad \text{oraz} \quad \operatorname{sgn} m = \begin{cases} 1, & \text{gdy } m > 0, \\ 0, & \text{gdy } m = 0, \\ \text{-1}, & \text{gdy } m < 0 \end{cases}$$

jest odwzorowaniem (nazywanym *znakiem* liczby).

■ Często regułę (czy odwzorowanie) $f : X \to Y$ taką, że

$$f(x) = y \in Y$$

dla elementów $x \in X$ oznaczają na jeden ze sposobów:

$$f : \begin{cases} X \to Y, \\ x \mapsto y \ (x \in X), \end{cases}$$

albo

$$\begin{cases} f : X \to Y, \\ f(x) = y \ (x \in X), \end{cases}$$

albo krótko

$$f : X \ni x \mapsto y \in Y.$$

■ Funkcja (=odwzorowanie) jest szczególnym przypadkiem relacji. A mianowicie, *odwzorowaniem* $f : X \to Y$ określonym na zbiorze X o wartościach w zbiorze Y nazywamy taką relację $f \subseteq X \times Y$, że dla każdego elementu $x \in X$ znajdzie się dokładnie jeden element $y \in Y$ taki, że $(x, y) \in f$.

■ Niech X, Y, Z oraz W będą zbiorami niepustymi. Wtedy odwzorowania $f : X \to Y$ i $g : Z \to W$ nazywa się *równymi* (i oznacza przez $f = g$), jeśli są równe odpowiednio ich dziedziny, ich przeciwdziedziny (czyli $X = Z$ oraz $Y = W$) i w każdym punkcie x z dziedziny X oba odwzorowania przyjmują równe wartości:

$$f(x) = g(x).$$

■ Odwzorowanie $f : X \to Y$ jest nazywane:

- *iniektywnym* (równoważnie *odwzorowaniem wzajemnie jednoznacznym w*), jeśli różne elementy x_1, x_2 z dziedziny X są odwzorowane w różne elementy $f(x_1), f(x_2)$ z przeciwdziedziny Y, czyli krótko

  $$\forall_{x_1,x_2 \in X} : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2);$$

  równoważnie też mówi się, że odwzorowanie $f : X \to Y$ jest *iniektywne*, jeśli dla dowolnych $x_1, x_2 \in X$ zachodzi implikacja

  $$\forall_{x_1,x_2 \in X} : f(x_1) = f(x_2) \Rightarrow x_1 = x_2;$$

- *suriektywnym* (równoważnie *odwzorowaniem „na”* lub *„typu na”*), jeśli dla każdego elementu y z przeciwdziedziny Y znajdzie się przeciwobraz x w dziedzinie X taki, że $y = f(x)$; krótko

  $$\forall_{y \in Y} \exists_{x \in X} : y = f(x);$$

- *bijektywnym* (równoważnie *odwzorowaniem wzajemnie jednoznacznym na*), jeśli jest jednocześnie iniektywne i suriektywne.

Przykłady 1.4.2.

(**1**) Odwzorowanie $g : X \to Y$ z przykładu 1.4.1(1) jest suriektywne, lecz nie jest iniektywne; odwzorowanie $h : X \to Y$ nie jest ani iniektywne, ani suriektywne; a odwzorowanie $k : Y \to X$ jest iniektywne, lecz nie jest suriektywne.

(**2**) Reguła $l : \{a, b, c, d\} \to \{1, 2, 3, 4\}$ taka, że

$$\begin{array}{rcl} l(a) & = & 1, \\ l(b) & = & 4, \\ l(c) & = & 2, \\ l(d) & = & 3, \end{array}$$

jest bijekcją.

(**3**) Odwzorowania

$$\alpha : \mathbb{R} \ni x \mapsto \sin x \in \mathbb{R},$$

oraz

$$\beta : \mathbb{R} \ni x \mapsto \sin x \in [-1, 1]$$

nie są równymi, bo ich przeciwdziedziny $\mathbb{R}$ i $[-1, 1]$ są różne.

■ Odwzorowanie $f : X \to Y$ jest suriektywne wtedy i tylko wtedy, gdy $\operatorname{Im} X = Y$.

* * *

■ **Złożenie odwzorowań.** Niech X, Y, Z oraz W będą zbiorami niepustymi.

Lemat 1.4.3. *Niech $f : X \to Y$ oraz $g : Y \to Z$ będą odwzorowaniami. Wtedy reguła $g \circ f$ taka, że*

$$(g \circ f)(x) = g(f(x)) \tag{1.5}$$

dla dowolnego $x \in X$ określa odwzorowanie $g \circ f : X \to Z$ zbioru X w zbiór Z.

Dowód. Niech x będzie dowolnym elementem zbioru X. Wtedy wartość $f(x) \in Y$ istnieje i jest określona dokładnie jednoznacznie. Ponieważ $g : Y \to Z$ jest odwzorowaniem, to $g(f(x)) \in Z$ istnieje i jest określone jednoznacznie. Zatem dla każdego elementu x jego obraz

$$(g \circ f)(x) = g(f(x))$$

istnieje i jest określony dokładnie jednoznacznie. To znaczy, że $g \circ f : X \to Z$ jest odwzorowaniem. □

■ Niech $f : X \to Y$ i $g : Y \to Z$ będą odwzorowaniami. *Złożeniem* $g \circ f$ (=*kompozycją* lub *iloczynem* lub, co jest równoważne, *superpozycją*)

odwzorowań f i g jest nazywane odwzorowanie $g \circ f : X \to Z$ zbioru X w zbiór Z takie, że zachodzi (1.5) dla każdego elementu $x \in X$. Z definicji wynika, że złożenie $f \circ g$ odwzorowań g i f istnieje w tym przypadku, gdy $X = Z$; jeśli $X \neq Z$, to złożenie $f \circ g$ nie istnieje.

Przykłady 1.4.4.

(**1**) Niech $g : X \to Y$ oraz $k : Y \to X$ będą odwzorowaniami z przykładu 1.4.1(1). Wtedy złożenie $k \circ g : X \to X$ istnieje oraz

$$\begin{array}{rcl} (k \circ g)(a) & = & k(g(a)) = k(2) = d, \\ (k \circ g)(b) & = & k(g(b)) = k(1) = a, \\ (k \circ g)(c) & = & k(g(c)) = k(2) = d, \\ (k \circ g)(d) & = & k(g(d)) = k(3) = b. \end{array}$$

Odwzorowanie $g \circ k : Y \to Y$ też istnieje (proponujemy Czytelnikowi zbudować go samodzielnie).

(**2**) Jeśli $k : Y \to X$ jest odwzorowaniem z przykładu 1.4.1(1), a $l : X \to Z$ jest odwzorowaniem z przykładu 1.4.2(2), to złożenie $l \circ k : Y \to Z$ istnieje, a złożenie $k \circ l$ nie istnieje.

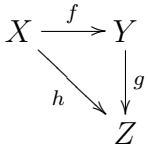
(**3**) Niech $f : \mathbb{R} \to \mathbb{R}$, gdzie $f(x) = \sin x$, a $g : \mathbb{R} \to \mathbb{R}$, gdzie $g(x) = 2^x$. Wtedy oba złożenia $g \circ f : \mathbb{R} \to \mathbb{R}$ i $f \circ g : \mathbb{R} \to \mathbb{R}$ istnieją oraz dla każdego elementu $x \in \mathbb{R}$ zachodzą równości

$$\begin{array}{rcccl} (g \circ f)(x) & = & g(f(x)) = g(\sin x) & = & 2^{\sin x}, \\ (f \circ g)(x) & = & f(g(x)) = f(2^x) & = & \sin 2^x. \end{array}$$

■ Jeśli $f : X \to Y$ i $g : Y \to Z$ są odwzorowaniami, to

$$D(g \circ f) = D(f) = X \text{ oraz } \operatorname{Im}(g \circ f) \subseteq \operatorname{Im} g \subseteq Z.$$

■ Fakt, że odwzorowanie $h : X \to Z$ jest złożeniem odwzorowań $f : X \to Y$ i $g : Y \to Z$, możemy zilustrować graficznie w postaci takiego diagramu *przemiennego*:



■ Odwzorowanie postaci $f : X \to X$ jest nazywane *przekształceniem* zbioru X (czyli odwzorowanie f jest nazywane przekształceniem zbioru X, jeśli X jest jednocześnie dziedziną i przeciwdziedziną odwzorowania f). Przekształcenie $i_X : X \to X$ zbioru niepustego X jest nazywane *tożsamościowym* (=*jednostkowym*) odwzorowaniem, jeśli

$$i_X(x) = x$$

dla każdego elementu $x \in X$. Zaznaczmy, że odwzorowanie jednostkowe i_X jest bijektywne. Zamiast symbolu i_X czasem równoważnie będziemy też używać symbolu id_X.

Lemat 1.4.5. *Niech X i Y będą zbiorami niepustymi, a $f : X \to Y$ będzie odwzorowaniem. Wtedy*

$$f \circ \mathrm{id}_X = f \text{ oraz } \mathrm{id}_Y \circ f = f.$$

Dowód. W rzeczy samej, złożenie $f \circ \mathrm{id}_X$ istnieje oraz $f \circ \mathrm{id}_X : X \to Y$. Oprócz tego dla każdego argumentu $x \in X$ wartość

$$(f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x).$$

Zatem $f = f \circ \mathrm{id}_X$. Inna równość ma podobny dowód. □

Twierdzenie 1.4.6 (łączność złożenia)**.** *Złożenie odwzorowań jest podporządkowane prawu łączności, czyli jeśli mamy odwzorowania $f : X \to Y$, $g : Y \to Z$ oraz $h : Z \to W$, to*

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

Dowód. Skoro $\mathrm{Im}\, g \subseteq Z$ i $D(h) = Z$, to istnieje złożenie $h \circ g : Y \to W$. Z powodu tego, że $\mathrm{Im}\, f \subseteq Y$ wnosimy, że istnieje $(h \circ g) \circ f : X \to W$. Podobnie, z warunków $\mathrm{Im}\, f \subseteq Y$ i $D(g) = Y$ wynika, że istnieje złożenie $g \circ f : X \to Z$. I skoro $\mathrm{Im}(g \circ f) \subseteq Z$, to istnieje też odwzorowanie $h \circ (g \circ f) : X \to W$. Oprócz tego dla każdego elementu $x \in X$ mamy

$$\begin{gathered}((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = \\ = h(g(f(x))) = h((g \circ f)(x)) = (h \circ (g \circ f))(x),\end{gathered}$$

a więc odwzorowania $(h \circ g) \circ f$ oraz $h \circ (g \circ f)$ są równe. □

* * *

■ **Złożenie odwzorowań iniektywnych.** Niech X, Y i Z będą zbiorami niepustymi.

Lemat 1.4.7. *Jeśli złożenie dwóch odwzorowań iniektywnych istnieje, to jest odwzorowaniem iniektywnym.*

Dowód. Niech $f : X \to Y$ oraz $g : Y \to Z$ będą odwzorowaniami iniektywnymi, dla których istnieje złożenie $g \circ f : X \to Z$. Jeśli x_1, x_2 są

różnymi elementami z dziedziny X odwzorowania f, to w wyniku iniektywności odwzorowania f ich obrazy są różne, czyli $f(x_1) \neq f(x_2)$. Ponieważ $f(x_1), f(x_2) \in Y$, Y jest dziedziną odwzorowania iniektywnego g, to elementy $f(x_1), f(x_2)$ mają różne obrazy względem g, czyli

$$(g \circ f)(x_1) = g(f(x_1)) \neq g(f(x_2)) = (g \circ f)(x_2).$$

To oznacza, że złożenie $g \circ f$ jest odwzorowaniem iniektywnym. □

* * *

■ **Złożenie odwzorowań suriektywnych.** Niech X, Y i Z będą zbiorami niepustymi.

Lemat 1.4.8. *Jeśli złożenie dwóch odwzorowań suriektywnych istnieje, to jest odwzorowaniem suriektywnym.*

Dowód. Niech $f : X \to Y$ i $g : Y \to Z$ będą odwzorowaniami suriektywnymi, dla których istnieje złożenie $g \circ f : X \to Z$. Jeśli z jest dowolnym elementem z przeciwdziedziny Z odwzorowania g, to w wyniku suriektywności g istnieje taki element $y \in Y$, że $z = g(y)$. Podobnie z suriektywności f otrzymujemy, że istnieje takie $x \in X$, że $y = f(x)$. Wtedy

$$z = g(y) = g(f(x)) = (g \circ f)(x),$$

a więc $g \circ f$ jest odwzorowaniem suriektywnym. □

Z udowodnionych wyżej dwóch lematów wyciągamy taki

Wniosek 1.4.9. *Jeśli złożenie dwóch odwzorowań bijektywnych istnieje, to jest odwzorowaniem bijektywnym.*

* * *

■ **Odwzorowanie odwrotne.** Niech X i Y będą zbiorami niepustymi. *Odwrotnym* do odwzorowania $f : X \to Y$ jest nazywane takie odwzorowanie $g : Y \to X$, że

$$g \circ f = \mathrm{id}_X \text{ oraz } f \circ g = \mathrm{id}_Y$$

(w tym przypadku zapisujemy, że $g = f^{-1}$ oraz $f = g^{-1}$).

Przykład 1.4.10.

Niech $\mathbb{R}_+ = \{r \in \mathbb{R} \mid r > 0\}$ będzie zbiorem liczb rzeczywistych dodatnich,

$$f : \mathbb{R} \ni x \mapsto 2^x \in \mathbb{R}_+ \quad \text{oraz} \quad g : \mathbb{R}_+ \ni x \mapsto \log_2 x \in \mathbb{R}.$$

Wtedy dla każdego $x \in \mathbb{R}$ oraz każdego $y \in \mathbb{R}_+$ mamy

$$\begin{array}{rclcccl} (g \circ f)(x) & = & g(f(x)) = g(2^x) = \log_2(2^x) & = & x & = & \mathrm{id}_{\mathbb{R}}(x), \\ (f \circ g)(y) & = & f(g(y)) = g(\log_2 y) = 2^{\log_2 y} & = & y & = & \mathrm{id}_{\mathbb{R}_+}(y), \end{array}$$

a więc $f^{-1} = g$ oraz $f = g^{-1}$.

Twierdzenie 1.4.11 (o istnieniu odwzorowania odwrotnego)**.** *Odwzorowanie $f : X \to Y$ ma odwzorowanie odwrotne do siebie wtedy i tylko wtedy, gdy jest bijektywne.*

Dowód. ($\Rightarrow$) Załóżmy, że odwzorowanie $f : X \to Y$ ma odwrotne do siebie odwzorowanie $g : Y \to X$, czyli

$$g \circ f = \mathrm{id}_X \ \text{ oraz } f \circ g = \mathrm{id}_Y .$$

Najpierw przekonajmy się, że f jest iniektywne. Istotnie, załóżmy, że x_1, x_2 są takimi elementami z X, że $f(x_1) = f(x_2)$. Wtedy otrzymujemy

$$\begin{array}{c} x_1 = \mathrm{id}_X(x_1) = (g \circ f)(x_1) = g(f(x_1)) = \\ = g(f(x_2)) = (g \circ f)(x_2) = \mathrm{id}_X(x_2) = x_2, \end{array}$$

a więc f jest iniektywne.

Teraz skoro g jest odwzorowaniem, to dla każdego elementu $y \in Y$ jego obraz $g(y)$ istnieje i leży w zbiorze X. Oznaczmy go przez x. Wtedy

$$f(x) = f(g(y)) = (f \circ g)(y) = \mathrm{id}_Y(y) = y$$

oraz x jest przeciwobrazem elementu y względem odwzorowania f; zatem f jest suriektywne. Wnosimy, że f jest odwzorowaniem bijektywnym.

($\Leftarrow$) Niech teraz odwzorowanie $f : X \to Y$ będzie bijektywne. Skonstruujmy odwrotne do niego. W tym celu rozpatrzmy regułę g taką, że

$$g(y) = x \ (\text{gdzie } y \in Y, x \in X) \Leftrightarrow y = f(x).$$

Skoro f jest odwzorowaniem suriektywnym, to każdemu elementowi $y \in Y$ odpowiada pewien jego przeciwobraz $x \in X$ względem g, przy czym

z iniektywności odwzorowania f wynika, że ten przeciwobraz jest dokładnie jednoznacznie określony dla każdego elementu $y \in Y$. Wnosimy, że g jest odwzorowaniem. Polecamy Czytelnikowi samodzielnie przekonać się, że g jest odwrotne do odwzorowania f. □

Wniosek 1.4.12. *Jeśli $f : X \to Y$ jest odwzorowaniem bijektywnym, to odwrotne odwzorowanie $f^{-1} : Y \to X$ istnieje, jest bijektywne oraz*

$$(f^{-1})^{-1} = f.$$

□

Uwaga 1.4.13. Jeśli $\mathcal{B} = \{B_\alpha \mid \alpha \in \Lambda\}$ jest rodziną zbiorów B_α, gdzie Λ jest pewnym zbiorem indeksów, to *iloczynem kartezjańskim* rodziny zbiorów $\mathcal{B}$ jest nazywany zbiór

$$\prod_{\alpha\in\Lambda} B_\alpha = \{f : \Lambda \to \bigcup_{\alpha\in\Lambda} B_\alpha \mid f \text{ jest odwzorowaniem takim, że } f(\alpha) \in B_\alpha,\ \forall_{\alpha\in\Lambda}\}.$$

Jeśli $B_\alpha = B$ dla wszystkich $\alpha \in \Lambda$, to iloczyn kartezjański $\prod_{\alpha\in\Lambda} B_\alpha$ krótko jest oznaczany przez B^Λ i nazywany *potęgą kartezjańską* zbioru B.

* * *

■ **Zbiory przeliczalne.** Zbiór, który jest skończony lub równoliczny ze zbiorem liczb naturalnych $\mathbb{N}$, jest nazywany *przeliczalnym*. Zbiór, który nie jest przeliczalny, jest nazywany *nieprzeliczalnym*.

Na przykład zbiór $\{\frac{1}{n} \mid n \in \mathbb{N}^*\}$ jest przeliczalny.

Ćwiczenia 1.4.14.

(**1**) Jeśli $f : X \to X$ jest odwzorowaniem, to sprawdzić, czy f jest iniektywne, suriektywne, bijektywne. Znaleźć obraz $\operatorname{Im} f$, jeśli:
(a) $X = \mathbb{R}$ oraz $f(x) = 3^x$;
(b) $X = \mathbb{R}$ oraz $f(x) = x^5$;
(c) $X = \mathbb{Q}$ oraz $f(x) = \frac{3x}{x^4+1}$;
(d) $X = \mathbb{R}$ oraz $f(x) = 3^x + 2x$;
(e) $X = \mathbb{R}$ oraz $f(x) = \begin{cases} 2x \ln|x|, & \text{gdy } x \neq 0, \\ 0, & \text{gdy } x = 0; \end{cases}$
(f) $X = \mathbb{Q}$ oraz $f(x) = \begin{cases} \frac{2x}{x+3}, & \text{gdy } x \neq -3, \\ 2, & \text{gdy } x = -3; \end{cases}$
(g) $X = \mathbb{Q}$ oraz $f(x) = x^3 - 2x + 3$.
(**2**) Sprawdzić, czy $f : X \to X$ jest odwzorowaniem bijektywnym? Znaleźć (jeżeli istnieje) f^{-1} dla:

(a) $X = \mathbb{Z} \times \mathbb{Z}$ oraz $f(a,b) = (-a,-b)$;
(b) $X = \mathbb{Z} \times \mathbb{Z}$ oraz $f(a,b) = (a+b, a-b)$;
(c) $X = (0,+\infty)$ oraz $f(x) = \frac{1}{x}$;
(d) $X = \mathbb{Q}$ oraz $f(x) = x - 1$.
(**3**) Sprawdzić, czy $f : \mathbb{R} \to \mathbb{R}$ jest odwzorowaniem bijektywnym? Znaleźć (jeśli istnieje) f^{-1} dla:
(a) $f(x) = 3x + 5$;
(b) $f(x) = x^3 - 3$;
(c) $f(x) = (x-3)^2$;
(d) $f(x) = \sqrt[5]{x} + 3$;
(e) $f(x) = x^3 + 1$.
(**4**) Niech $f : \mathbb{R} \to \mathbb{R}$, gdzie $f(x) = x^2 - 3x$, $g : \mathbb{R} \to \mathbb{R}$, gdzie $g(x) = \frac{1}{x^2+1}$, oraz $h : \mathbb{R} \to \mathbb{R}$, gdzie $h(x) = x^3$. Znaleźć (jeśli istnieje) złożenie:
(a) $g \circ f, f \circ g$;
(b) $f \circ g, g \circ f$;
(c) $g \circ h, h \circ g$;
(d) $f \circ g \circ h, f \circ h \circ g$;
(e) $g \circ h \circ f$;
(f) $g \circ f \circ h$;
(g) $f \circ f, h \circ h, g \circ g$.

Uwagi. Współczesną definicję funkcji wprowadził P. Dirichlet[(22)] w 1837 r.; podstawy współczesnej terminologii zostały opracowane przez N. Bourbaki.

[(22)] Peter Lejeune Dirichlet (1805–1859)

Rozdział 2

Podstawowe struktury algebraiczne

2.1. Działania algebraiczne

Pojęcie działania algebraicznego jest jednym z głównych we współczesnej algebrze.

■ Niech X będzie zbiorem niepustym. *Działaniem algebraicznym* (lub *wewnętrznym*) na zbiorze X jest nazywana reguła „$*$”, która każdej parze uporządkowanej (x, y) elementów $x, y \in X$ dokładnie jednoznacznie przyporządkowuje element $z \in X$; przy tym element z jest nazywany *iloczynem* elementów x, y i oznaczany przez $z = x * y$.

■ Zatem każde odwzorowanie postaci $X \times X \to X$ jest działaniem algebraicznym na zbiorze niepustym X i na odwrót.

■ Dalej, mówiąc o działaniach algebraicznych, będziemy krótko mówić o działaniach.

■ Do oznaczania działań często są wykorzystywane symbole: $+$, $\cdot$, $\circ$, $\dotplus$, $\times$, $\div$, $\divideontimes$, $*$, $\star$, $\oplus$, $\ominus$, $\otimes$, $\odot$, $\circledcirc$, $\boxplus$, $\boxdot$, $\intercal$, $\diamond$ itd.

■ Zapis $(G, *)$ oznacza, że zbiór G jest rozpatrywany względem reguły czy działania „$*$” lub też że na zbiorze G rozpatrujemy regułę „$*$”.

Przykłady 2.1.1.

(**1**) Dodawanie „$+$” i mnożenie „$\cdot$” liczb są działaniami algebraicznymi na każdym ze zbiorów $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ oraz $\mathbb{C}$. Odejmowanie liczb „$-$” także jest działaniem algebraicznym na zbiorach liczb całkowitych $\mathbb{Z}$, liczb wymiernych $\mathbb{Q}$, liczb rzeczywistych $\mathbb{R}$ oraz liczb zespolonych $\mathbb{C}$, lecz na zbiorze liczb naturalnych $\mathbb{N}$ odejmowanie nie jest algebraiczne, bo $0, 1 \in \mathbb{N}$, a $0 - 1 = -1 \notin \mathbb{N}$.

(**2**) W zbiorze klas reszt $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ modulo 2 możemy wprowadzić dwa działania (dodawanie „+" i mnożenie „·", które jak łatwo sprawdzić są algebraiczne) klas reszt w taki sposób:

$$\begin{array}{ll} \overline{0}+\overline{0}=\overline{0}, & \overline{0}\cdot\overline{0}=\overline{0}, \\ \overline{0}+\overline{1}=\overline{0}, & \overline{0}\cdot\overline{1}=\overline{0}, \\ \overline{1}+\overline{0}=\overline{1}, & \overline{1}\cdot\overline{0}=\overline{0}, \\ \overline{1}+\overline{1}=\overline{0}, & \overline{1}\cdot\overline{1}=\overline{1}. \end{array}$$

Takie działania (w przypadku zbiorów skończonych) będziemy czasem zapisywać w postaci tabelek Cayleya[(1)]:

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

(**3**) Niech $\mathbb{K}$ będzie zbiorem wszystkich punktów płaszczyzny. Każdej parze uporządkowanej punktów (A, B) jednoznacznie odpowiada punkt M, będący środkiem odcinka AB; takie przyporządkowanie $(A, B) \mapsto M$ określa działanie algebraiczne na zbiorze $\mathbb{K}$.

(**4**) Niech X będzie zbiorem niepustym. Wtedy różnica symetryczna „$\div$", gdzie

$$A \div B = (A \setminus B) \bigcup (B \setminus A)$$

dla zbiorów $A, B \in 2^X$, określa działanie algebraiczne na zbiorze 2^X.

■ Działanie algebraiczne „$*$", określone na zbiorze niepustym X, jest nazywane:

- *przemiennym* (lub *komutatywnym*), jeśli dla dowolnych elementów $x, y \in X$ zachodzi równość

$$x * y = y * x;$$

- *łącznym*, jeśli dla dowolnych elementów $x, y, z \in X$ zachodzi

$$(x * y) * z = x * (y * z).$$

■ Niech na zbiorze niepustym X będzie określone działanie algebraiczne „$*$". Wtedy:

(1) Arthur Cayley (1821–1895)

- element $e \in X$ jest nazywany *neutralnym* (względem działania „$*$"), jeśli dla każdego elementu $a \in X$ zachodzą równości

$$a * e = a = e * a;$$

- element $e \in X$ jest nazywany *idempotentem* lub *elementem idempotentnym* (względem „$*$"), jeśli

$$e * e = e;$$

- jeśli e jest elementem neutralnym w X (względem „$*$") oraz $x, y \in X$, to element y jest nazywany *odwrotnym do elementu* x, jeśli

$$y * x = e = x * y$$

(jeśli $y \in X$ jest odwrotny do elementu x, to zapisujemy $y = x^{-1}$; i wtedy element x nazywamy *odwracalnym* w zbiorze X);

- będziemy mówić, że w zbiorze X zachodzi *prawo skracania* dla działania „$*$", jeśli dla dowolnych elementów $a, b, c \in X$ z równości $a*b = a*c$ wynika, że $b = c$, a z równości $b * a = c * a$ wynika, że $b = c$.

Przykłady 2.1.2.

(**1**) Dla liczb wymiernych $a, b \in \mathbb{Q}$ określmy iloczyn „$\perp$" według reguły

$$a \perp b = \frac{a+b}{2}.$$

Sprawdźmy, jakie własności posiada działanie „$\perp$". Połowa sumy dwóch liczb wymiernych zawsze istnieje i jest określona dokładnie jednoznacznie, a więc „$\perp$" jest działaniem algebraicznym w $\mathbb{Q}$.

(a) (przemienność „$\perp$") Dla dowolnych elementów $a, b \in \mathbb{Q}$ mamy

$$a \perp b = \frac{a+b}{2} = \frac{b+a}{2} = b \perp a.$$

(b) Jeśli a, b, c są dowolnymi elementami z $\mathbb{Q}$, to

$$(a \perp b) \perp c = \frac{a+b}{2} \perp c = \frac{\frac{a+b}{2} + c}{2} = \frac{a+b+2c}{4}$$

oraz

$$a \perp (b \perp c) = a \perp \frac{b+c}{2} = \frac{a + \frac{b+c}{2}}{2} = \frac{2a+b+c}{4}.$$

Z porównania tych wyników wyłania się hipoteza, że „$\perp$" nie jest łączne. W rzeczy samej, dla trójki liczb $a = 1$, $b = 0$ oraz $c = 0$ otrzymujemy

$$(1 \perp 0) \perp 0 = \frac{1}{4} \neq \frac{1}{2} = 1 \perp (0 \perp 0),$$

a więc działanie „$\perp$” nie jest łączne.

(c) Znajdźmy element $e \in \mathbb{Q}$ taki, że $e \perp a = a = a \perp e$ dla wszystkich liczb wymiernych a. W tym celu rozwiązujemy układ równań

$$\begin{cases} e + a = 2a, \\ \forall_{a \in \mathbb{Q}}, \end{cases}$$

i otrzymujemy, że $e = a$, gdzie a jest dowolną liczbą wymierną. Z tych rozumowań wynika, że w zbiorze liczb wymiernych $\mathbb{Q}$ względem działania „$\perp$” nie istnieje element neutralny (taki element musi być jedyny, jeśli istnieje (patrz twierdzenie 2.1.3)).

(d) Jeśli $a, b, c \in \mathbb{Q}$, to z

$$\frac{a+b}{2} = \frac{a+c}{2}$$

wynika, że $b = c$ (czyli dla „$\perp$” zachodzi prawo skracania).

(**2**) Zbadajmy, jakie własności ma działanie „$\oplus$” na zbiorze liczb rzeczywistych $\mathbb{R}$, gdzie

$$a \oplus b = a + b + 1$$

dla elementów $a, b, c \in \mathbb{R}$. Skoro suma trzech liczb rzeczywistych jest zawsze określona dokładnie jednoznacznie, to „$\oplus$” jest działaniem algebraicznym na zbiorze $\mathbb{R}$.

(a) (przemienność działania „$\oplus$”) Jeśli $a, b \in \mathbb{R}$ są dowolnymi elementami, to

$$a \oplus b = a + b + 1 = b + a + 1 = b \oplus a.$$

(b) (łączność działania „$\oplus$”) Dla dowolnych elementów $a, b, c \in \mathbb{R}$ mamy

$$(a \oplus b) \oplus c = (a + b + 1) \oplus c = a + b + 1 + c + 1 = a + b + c + 2 =$$

$$= a + (b + c + 1) + 1 = a + (b \oplus c) + 1 = a \oplus (b \oplus c).$$

(c) Znajdźmy element $e \in \mathbb{R}$ taki, że $a \oplus e = a$ dla wszystkich $a \in \mathbb{R}$. W tym celu rozwiązujemy równanie $a + e + 1 = a$ (względem niewiadomego e), z którego wynika, że $e = -1$ jest elementem neutralnym w zbiorze $\mathbb{R}$ (względem „$\oplus$”).

(d) Dlatego, żeby znaleźć idempotenty $x \in \mathbb{R}$ względem działania „$\oplus$”, rozwiązujemy równanie $a \oplus a = a$ lub równoważnie $a + a + 1 = a$, skąd otrzymujemy, że $a = -1$ jest dokładnie jedynym idempotentem w $\mathbb{R}$.

(e) Niech $a \in \mathbb{R}$. Znajdźmy (jeśli istnieje) element odwrotny do a. Niech $x \in \mathbb{R}$ będzie elementem spełniającym warunek $x \oplus a = -1$. Wtedy $x + a + 1 = -1$ lub w postaci równoważnej $x = -a - 2$. Zatem

$$a^{-1} = -a - 2$$

jest odwrotnym do elementu a w zbiorze $\mathbb{R}$.

(f) Jeśli $a, b, c \in \mathbb{R}$, to z równości

$$a + b + 1 = a + c + 1$$

wynika, że $b = c$ (czyli dla działania „$\oplus$” zachodzi prawo skracania w zbiorze $\mathbb{R}$).

■ Zwykle w przypadku ogólnym (=abstrakcyjnym) rozpatrywane działanie jest nazywane *mnożeniem* i oznaczane przez „$\cdot$”. Lecz często dla oznaczenia działania algebraicznego jest także używany symbol „$+$”; wtedy

takie działanie jest nazywane *dodawaniem*. Istnieją pewne równoległości w nazewnictwie elementów szczególnych w zależności od tego, jak są nazywane działania algebraiczne. Zilustrujmy to w poniższej tabelce.

Notacja multiplikatywna $(X, \cdot)$	Notacja addytywna $(X, +)$
działanie „$\cdot$" jest nazywane *mnożeniem*, a element $z = x \cdot y$ – *iloczynem* elementów $x, y \in X$; elementy x i y przy tym są nazywane *czynnikami*;	działanie „$+$" jest nazywane *dodawaniem*, a element $z = x + y$ – *sumą* elementów $x, y \in X$; przy tym elementy x i y są nazywane *składnikami*;
element neutralny $e \in X$ względem działania „$\cdot$" jest nazywany *jednością* lub *elementem jednostkowym*; i oznaczany przez $e = 1$;	element neutralny $e \in X$ względem działania „$+$" jest nazywany *zerem* lub *elementem zerowym*; i oznaczany przez $e = 0$;
element odwrotny do elementu $a \in X$ zapisujemy w postaci a^{-1};	zamiast terminu *odwrotny* do elementu $a \in X$ używamy terminu *przeciwny* element do $a \in X$, który oznaczamy przez $-a$;
jeśli $n \in \mathbb{Z}$, to *n-ta potęga* elementu $a \in X$ jest określona tak: $a^n = \begin{cases} \underbrace{a \cdots a}_{n \text{ czynników}}, & \text{gdy } n > 0, \\ 1, & \text{gdy } n = 0, \\ \underbrace{a^{-1} \cdots a^{-1}}_{\lvert n \rvert \text{ czynników}}, & \text{gdy } n < 0. \end{cases}$	jeśli $n \in \mathbb{Z}$, to *n-ta krotna* elementu $a \in X$ jest określone następującym sposobem: $na = \begin{cases} \underbrace{a + \cdots + a}_{n \text{ składników}}, & \text{gdy } n > 0, \\ 0, & \text{gdy } n = 0, \\ \underbrace{(-a) + \cdots + (-a)}_{\lvert n \rvert \text{ składników}}, & \text{gdy } n < 0. \end{cases}$

■ Z przytoczonych powyżej przykładów wynika, że nie wszystkie działania mają element neutralny i nie każdy element ma odwrotny względem pewnego działania algebraicznego. Lecz zawsze dla każdego działania algebraicznego zachodzą takie dwie podstawowe własności.

Twierdzenie 2.1.3. *Jeśli X jest zbiorem niepustym, na którym zostało określone działanie algebraiczne „$*$", to w zbiorze X istnieje nie więcej niż jeden element neutralny (względem działania „$*$").*

Dowód. Załóżmy, że w zbiorze X istnieją elementy neutralne względem działania „$*$", na przykład e_1 oraz e_2. Wtedy z neutralności elementu e_1 wynika, że $e_1 * e_2 = e_2$, a z neutralności e_2 otrzymujemy $e_1 * e_2 = e_1$. W wyniku tego wnosimy, że $e_1 = e_2$. □

Twierdzenie 2.1.4. *Niech na zbiorze niepustym X będzie określone łączne działanie algebraiczne „$*$”. Jeśli w X istnieje element neutralny e (względem działania „$*$”), to każdy element $a \in X$ ma co najwyżej jeden element odwrotny do siebie.*

Dowód. Niech $a \in X$. Załóżmy, że w zbiorze X istnieją elementy b_1, b_2, będące odwrotnymi do a. Wtedy

$$\begin{array}{rcccl} b_1 * a & = & e & = & a * b_1, \\ b_2 * a & = & e & = & a * b_2, \end{array}$$

a więc

$$b_1 = b_1 * e = b_1 * (a * b_2) = (b_1 * a) * b_2 = e * b_2 = b_2,$$

czyli $b_1 = b_2$. □

■ Niech X będzie zbiorem niepustym, na którym są określone dwa działania algebraiczne: „$*$” oraz „$\circ$”. Będziemy mówić, że działanie „$*$”:

- jest *rozdzielne lewostronnie* względem działania „$\circ$”, jeśli dla dowolnych elementów $a, b, c \in X$ zachodzi równość

$$a * (b \circ c) = (a * b) \circ (a * c);$$

- jest *rozdzielne prawostronnie* względem działania „$\circ$”, jeśli dla dowolnych elementów $a, b, c \in X$ zachodzi równość

$$(b \circ c) * a = (b * a) \circ (c * a);$$

- jest *rozdzielne obustronnie* (krótko *rozdzielne*) względem działania „$\circ$”, jeśli dla dowolnych $a, b, c \in X$ są spełnione równości

$$a * (b \circ c) = (a * b) \circ (a * c) \text{ oraz } (b \circ c) * a = (b * a) \circ (c * a).$$

Przykłady 2.1.5.

(**1**) Jeśli a, b, c są dowolnymi liczbami naturalnymi (odpowiednio całkowitymi, wymiernymi, rzeczywistymi czy zespolonymi), to

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \text{ oraz } (b + c) \cdot a = (b \cdot a) + (c \cdot a),$$

czyli mnożenie liczb „$\cdot$" jest rozdzielne względem dodawania liczb „$+$".

(**2**) Na dwuelementowym zbiorze $X = \{a, b\}$ reguły „$*$" i „$\intercal$" określają dwa działania algebraiczne:

$$\begin{array}{rcl} a * b & = & a, \\ b * a & = & b, \\ a \intercal b & = & b, \\ b \intercal a & = & a. \end{array}$$

Wtedy dla dowolnych elementów $x, y, z \in X$ przekonujemy się, że

$$\begin{array}{rcl} x \intercal (y * z) & = & x \intercal y = y = y * z = (x \intercal y) * (x \intercal z), \\ (y * z) \intercal x & = & x = x * x = (y \intercal x) * (z \intercal x), \end{array}$$

a to znaczy, że działanie "$\intercal$" jest rozdzielne względem „$*$".

Ćwiczenia 2.1.6.

(**1**) Sprawdzić, czy reguła „$*$" jest algebraiczna na zbiorze X. Jeśli „$*$" jest działaniem algebraicznym na X, to czy jest przemienne, łączne, czy posiada element neutralny. Znaleźć wszystkie idempotenty w X. Jeśli „$*$" ma element neutralny w X, to znaleźć wszystkie elementy odwracalne w X, jeśli:
(a) $X = \mathbb{Q}_+$ jest zbiorem dodatnich liczb wymiernych oraz $x * y = \sqrt{xy}$;
(b) X jest dowolnym zbiorem niepustym oraz $x * y = x$;
(c) $X = \mathbb{Z}$ oraz $x * y = \sqrt{x^2 + y - 1}$;
(d) $X = \mathbb{R}$ oraz $x * y = \sqrt{x + y + xy}$;
(e) $X = \mathbb{Z}$ oraz $x * y = \sqrt{x + y + 1}$;
(f) $X = \mathbb{N} \times \mathbb{N}$ oraz $(a, b) * (c, d) = (a + d, c + b)$;
(g) $X = \mathbb{R} \times \mathbb{R}$ oraz $(a, b) * (c, d) = (ac, bd)$;
(h) $X = \mathbb{Q} \times \mathbb{Q}^*$, gdzie $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ oraz $(a, b) * (c, d) = (\frac{a}{d}, \frac{c}{d})$;
(i) $X = \mathbb{N}$ oraz $x * y = x^y$;
(j) $X = \mathbb{Z}$ oraz $x * y = \mathrm{NWD}(x, y)$;
(k) $X = \mathbb{Z}$ oraz $x * y = y - x$;
(l) $X = \mathbb{Z}$ oraz $x * y = \mathrm{NWW}(x, y)$;
(m) $X = \mathbb{Q}$ oraz $x * y = \frac{a}{b} + \frac{b}{a}$;
(n) $X = \mathbb{Z}$ oraz $x * y = 2xy$;
(o) $X = \mathbb{Q}$ oraz $x * y = \frac{xy}{2}$;
(p) $X = \mathbb{R}$ oraz $x * y = \frac{x+y}{2}$;
(r) $X = \mathbb{Q} \setminus \{0\}$ oraz $x * y = \frac{x}{y}$.

(**2**) Niech $F(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ jest odwzorowaniem}\}$. Znaleźć element neutralny (względem złożenia odwzorowań), zawierający się w zbiorze $X \subseteq F(\mathbb{R})$, jeśli:
(a) X jest zbiorem funkcji ciągłych;
(b) X jest zbiorem funkcji ciągłych f takich, że $f(x) = x$ dla wszystkich $x \in [-1, 1]$;
(c) X jest zbiorem wielomianów stopni parzystych;
(d) X jest zbiorem wielomianów stopni nieparzystych;
(e) X jest zbiorem takich funkcji ciągłych f, że $f(0) = 0$.

Uwagi. Abstrakcyjne pojęcie działania algebraicznego zostało wyodrębnione w pracach algebraistów angielskich w latach 30.–50. XIX wieku.

2.2. Półgrupy i monoidy

■ Para $(G, *)$ jest nazywana *półgrupą*, jeśli:
$0_1)$ $G \neq \varnothing$,
1) „$*$” jest działaniem algebraicznym na zbiorze G;
2) „$*$” jest działaniem łącznym na zbiorze G.

■ Jeśli $(G, *)$ jest półgrupą z elementem neutralnym e, to G jest nazywana *monoidem*. Ponadto, jeżeli

$$x * x = x$$

dla każdego $x \in G$, to półgrupa (odpowiednio monoid) G jest nazywana (nazywany) *idempotentną* (odpowiednio *idempotentnym*). Półgrupa (odpowiednio monoid) G jest nazywana (nazywany) *przemienną* (odpowiednio *przemiennym*), jeśli

$$x * y = y * x$$

dla wszystkich $x, y \in G$.

Przykłady 2.2.1.

(**1**) Zbiór liczb naturalnych $\mathbb{N}$ jest monoidem przemiennym (względem działania dodawania liczb) z elementem neutralnym 0, a zbiór niezerowych liczb naturalnych $\mathbb{N}^*$ jest półgrupą przemienną.

(**2**) Jeśli $X \neq \varnothing$, a

$$\mathbb{M}(X) = \{f : X \to X \mid f \text{ jest odwzorowaniem}\},$$

to na podstawie lematu 1.4.3 i twierdzenia 1.4.6 wnosimy, że $\mathbb{M}(X)$ jest monoidem (względem złożenia „$\circ$”) z elementem neutralnym id_X (często $\mathbb{M}(X)$ jest nazywana *półgrupą symetryczną* zbioru X).

(**3**) Jeśli $X \neq \varnothing$, to 2^X jest przemiennym monoidem idempotentnym (względem działania przecięcia podzbiorów „$\cap$”), gdyż

$$\begin{array}{lcr} (A \bigcap B) \bigcap C & = & A \bigcap (B \bigcap C), \\ A \bigcap B & = & B \bigcap A, \\ A \bigcap A & = & A \end{array}$$

dla dowolnych $A, B, C \in 2^X$. Oprócz tego $A \cap X = A$, czyli X jest elementem neutralnym (względem działania „$\cap$”). Podobnie $(2^X, \cup)$ jest przemiennym monoidem idempotentnym z elementem neutralnym $\varnothing$, ponieważ

$$\begin{array}{lcr} (A \bigcup B) \bigcup C & = & A \bigcup (B \bigcup C), \\ A \bigcup B & = & B \bigcup A, \\ A \bigcup A & = & A, \\ A \bigcup \varnothing & = & A \end{array}$$

dla dowolnych $A, B, C \in 2^X$.

(**4**) (**Półgrupa słów**) Niech X będzie zbiorem niepustym, elementy którego będziemy nazywać *literami*. Każdy skończony ciąg liter

$$x_1 \cdots x_m,$$

gdzie $x_i \in X$ $(i = 1, \ldots, m;\ m \in \mathbb{N}^*)$, będziemy nazywać *słowem* nad alfabetem X. Słowo nieposiadające żadnej litery będziemy nazywać *pustym* (lub *spacją*). Niech $W = W(X)$ będzie zbiorem wszystkich słów niepustych nad alfabetem X. Na tym zbiorze zdefiniujmy iloczyn dwóch słów $a, b \in W(X)$ jako słowo ab (czyli litery a i b przy mnożeniu zapisujemy obok bez spacji między nimi). Tak określone mnożenie słów jest łączne; a zatem W jest półgrupą (która jest nazywana *półgrupą słów* nad alfabetem X).

(**5**) (**Półgrupa relacji**) Z twierdzenia 1.3.3 (2) oraz (4) wynika, że zbiór wszystkich relacji na zbiorze niepustym X tworzy monoid.

(**6**) (**Monoid macierzy kwadratowych**) Niech

$$M_n(\mathbb{R}) = \left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \mid a_{ij} \in \mathbb{R} \right\}$$

jest monoidem względem mnożenia macierzy (elementem neutralnym, którego jest macierz jednostkowa).

■ Jeśli G jest półgrupą z jednością e, to przyjmujemy, że

$$g^0 = e$$

dla każdego $g \in G$.

Lemat 2.2.2. *Jeśli g jest elementem półgrupy $(G, *)$ oraz $m, n \in \mathbb{N}^*$, to:*
(1) $g^m * g^n = g^{m+n} = g^{n+m} = g^n * g^m$;
(2) $(g^m)^n = g^{mn}$.

Dowód. (1) W rzeczy samej, z powodu łączności działania „$*$” otrzymujemy

$$g^m * g^n = \underbrace{(g * \cdots * g)}_{m \text{ czynników}} * \underbrace{(g * \cdots * g)}_{n \text{ czynników}} = \underbrace{g * \cdots * g}_{m+n \text{ czynników}} = g^{m+n} =$$
$$= g^{n+m} = \underbrace{g * \cdots * g}_{n+m \text{ czynników}} = \underbrace{(g * \cdots * g)}_{n \text{ czynników}} * \underbrace{(g * \cdots * g)}_{m \text{ czynników}} = g^n * g^m.$$

(2) Dowód jest podobny jak dla części (1). □

■ Półgrupa G jest nazywana *monogeniczną*, jeśli istnieje taki element $g \in G$, że każdy element $a \in G$ możemy przedstawić w postaci

$$a = g^n,$$

gdzie $n = n(a) \in \mathbb{N}$. Element g przy tym jest nazywany *generatorem* (lub *elementem generującym*) półgrupę G.

Przykład 2.2.3.

Półgrupa addytywna liczb naturalnych niezerowych $(\mathbb{N}^*, +)$ jest monogeniczna z generatorem 1, bo

$$n = \underbrace{1 + \cdots + 1}_{n \text{ składników}}$$

dla każdej liczby naturalnej niezerowej n.

■ Niech $(G, *)$ będzie półgrupą (odpowiednio monoidem). Wtedy H jest nazywane *podpółgrupą* (odpowiednio *podmonoidem*) w G, jeśli:
(0_1) $H \neq \varnothing$;
(0_2) H jest podzbiorem w zbiorze G;
(1) $(H, *)$ jest półgrupą (odpowiednio monoidem) względem tego samego działania „$*$”, co jest określone na zbiorze G (lecz zawężonego do podzbioru H).

■ Łatwo sprawdzić, że zachodzi takie „kryterium podpółgrupy (odpowiednio podmonoidu)”: zbiór H jest podpółgrupą (odpowiednio podmonoidem) w półgrupie (odpowiednio monoidzie) $(G, *)$ wtedy i tylko wtedy, gdy:
(0_1) $H \neq \varnothing$;
(0_2) H jest podzbiorem zbioru G;
(1) $a * b \in H$ dla dowolnych elementów $a, b \in H$ (w przypadku monoidu, oprócz tego konieczne jest, żeby element neutralny $e \in H$).

Przykłady 2.2.4.

(**1**) Zbiór parzystych liczb naturalnych niezerowych $2\mathbb{N}^*$ jest podpółgrupą w półgrupie $(\mathbb{N}^*, +)$, a zbiór $2\mathbb{N}$ jest podmonoidem w $\mathbb{N}$.

(**2**) Niech $(G, *)$ będzie monoidem (z jednością e),

$$\mathbb{L}(G) = \{L_g \mid L_g(x) = g * x \text{ dla elementów } x \in G, \text{gdzie } g \in G\}.$$

Wtedy $\mathrm{id}_G = L_e \in \mathbb{L}(G)$. Jeśli $L_g, L_h \in \mathbb{L}(G)$ dla pewnych $g, h \in G$, to dla dowolnego elementu $x \in G$ otrzymujemy

$$(L_g \circ L_h)(x) = L_g(L_h(x)) = L_g(h * x) = g * (h * x) = (g * h) * x = L_{g*h}(x),$$

czyli

$$L_g \circ L_h = L_{g*h} \in \mathbb{L}(G).$$

Poza tym złożenie odwzorowań „o" jest łączne na zbiorze $\mathbb{L}(G)$. Zatem $\mathbb{L}(G)$ jest monoidem będącym podmonoidem w monoidzie $\mathbb{M}(G)$.

■ Jeśli $(G, *)$ oraz $(H, \cdot)$ są półgrupami, to odwzorowanie $f : G \to H$ jest nazywane:

- *homomorfizmem* półgrup, jeśli zachodzi równość
$$f(g_1 * g_2) = f(g_1) \cdot f(g_2) \tag{2.1}$$
dla dowolnych elementów $g_1, g_2 \in G$;
- *monomorfizmem* półgrup, jeśli f jest odwzorowaniem iniektywnym i zachodzi (2.1) dla dowolnych elementów $g_1, g_2 \in G$;
- *epimorfizmem* półgrup, jeśli f jest odwzorowaniem suriektywnym i zachodzi (2.1) dla dowolnych elementów $g_1, g_2 \in G$;
- *izomorfizmem* półgrup, jeśli f jest odwzorowaniem bijektywnym i zachodzi (2.1) dla dowolnych elementów $g_1, g_2 \in G$.

W podobny sposób są definiowane homomorfizm (monomorfizm, epimorfizm, izomorfizm) monoidów z dodatkowym warunkiem, że $f(e_G) = e_H$ dla elementów neutralnych $e_G \in G$ oraz $e_H \in H$.

■ Półgrupy G i H są nazywane *izomorficznymi*, jeśli istnieje pewien izomorfizm półgrup postaci $f : G \to H$ (i wtedy piszą, że $G \cong H$).

Lemat 2.2.5. *Niech $(G, *)$ i $(H, \cdot)$ będą półgrupami. Jeśli $f : G \to H$ jest homomorfizmem półgrup, to jego obraz*

$$\operatorname{Im} f = \{f(g) \mid g \in G\}$$

jest podpółgrupą w H.

Dowód. Obraz $\operatorname{Im} f$ jest podzbiorem niepustym w H. Niech α, β będą dowolnymi elementami z $\operatorname{Im} f$. Wtedy $\alpha = f(a)$, $\beta = f(b)$ dla pewnych $a, b \in G$ oraz

$$\alpha \cdot \beta = f(a) \cdot f(b) = f(a * b) \in \operatorname{Im} f.$$

Zatem $\operatorname{Im} f$ jest podpółgrupą w H. □

■ Jeśli $f : G \to H$ jest monomorfizmem półgrup, to jego obraz $\operatorname{Im} f$ jest izomorficzny z półgrupą G. Dlatego w tym przypadku też mówimy, że f jest *włożeniem* półgrupy G w półgrupę H lub G *wkłada się izomorficznie* w półgrupę H (lub krótko G *wkłada się* w H).

Twierdzenie 2.2.6. *Każda półgrupa* (*odpowiednio monoid*) $(G, *)$ *jest izomorficzna podpółgrupie* (*odpowiednio izomorficzny podmonoidowi*) *półgrupy symetrycznej* (*odpowiednio monoidu symetrycznego*) $\mathbb{M}(G)$ *przekształceń zbioru* G.

Dowód. (1) Niech $(G, *)$ będzie monoidem (z jednością e). Wtedy odwzorowanie (patrz przykład 2.2.4(2))

$$\Theta : G \ni g \mapsto L_g \in \mathbb{L}(G)$$

jest homomorfizmem półgrup, bo dla dowolnych elementów $g, h \in G$ mamy

$$\Theta(g * h) = L_{g*h} = L_g \circ L_h = \Theta(g) \circ \Theta(h).$$

Jeśli $\Theta(g) = \Theta(h)$, to $L_g = L_h$, a więc

$$g = g * e = L_g(e) = L_h(e) = h * e = h,$$

czyli Θ jest odwzorowaniem iniektywnym. Oprócz tego Θ jest suriekcją oraz $\Theta(e) = \mathrm{id}_G$. Zatem Θ jest izomorfizmem monoidów G i $\mathbb{L}(G)$. Z udowodnionego wyżej też wynika, że $\mathbb{L}(G)$ jest podmonoidem w $\mathbb{M}(G)$.

(2) Załóżmy, że $(G, *)$ jest półgrupą bez jedności. Rozpatrzmy zbiór $G' = G \cup \{e\}$, gdzie $e \notin G$ oraz zachodzą równości $g * e = g = e * g$ dla każdego $g \in G$. Wtedy $(G, *)$ jest podpółgrupą w półgrupie G' i istnieje izomorfizm półgrup

$$\Theta_1 : G' \ni a \mapsto L_a \in \mathbb{L}(G'),$$

przy czym $\Theta_1(G)$ jest podpółgrupą w $\mathbb{L}(G')$. □

Lemat 2.2.7. *Niech* $(G, *)$ *i* $(H, \cdot)$ *będą półgrupami, a* $\phi : G \to H$ *będzie homomorfizmem półgrup. Wtedy zachodzą następujące własności:*

(1) *jeśli* ϕ *jest epimorfizmem, to* $\phi(e)$ *jest jednością półgrupy* H *dla elementu jednostkowego* e *półgrupy* G*;*

(2) *jeśli* ϕ *jest epimorfizmem i dla elementu* $a \in G$ *istnieje element odwrotny* $a^{-1} \in G$*, to* $\phi(a^{-1}) = \phi(a)^{-1}$.

Dowód. (1) Niech h będzie dowolnym elementem z H. Wtedy $h = \phi(g)$ dla pewnego elementu $g \in G$ oraz

$$h \cdot \phi(e) = \phi(g) \cdot \phi(e) = \phi(g * e) = \phi(g) = h,$$
$$\phi(e) \cdot h = \phi(e) \cdot \phi(g) = \phi(e * g) = \phi(g) = h,$$

czyli $\phi(e)$ jest jednością półgrupy H.

(2) Niech e będzie jednością półgrupy G oraz załóżmy, że istnieje element odwrotny $a^{-1} \in G$ dla pewnego elementu $a \in G$. Wtedy

$$a * a^{-1} = e = a^{-1} * a,$$

a zatem

$$\phi(a) \cdot \phi(a^{-1}) = \phi(a * a^{-1}) = \phi(e) =$$
$$= \phi(a^{-1} * a) = \phi(a^{-1}) \cdot \phi(a).$$

Stąd łatwo otrzymujemy, że $\phi(a^{-1}) = \phi(a)^{-1}$. □

Twierdzenie 2.2.8. *Każda przemienna półgrupa idempotentna* $(G, *)$ *wkłada się w półgrupę* $(2^G, \cap)$.

Dowód. Niech $g \in G$ oraz $A_g = \{g * x \mid x \in G\}$. Wtedy dla dowolnych elementów $g, h \in G$ otrzymujemy, że $A_g, A_h \in 2^G$ oraz

$$A_{g*h} = \{(g * h) * x \mid x \in G\} \subseteq A_g \bigcap A_h.$$

Natomiast jeśli $t \in A_g \cap A_h$, to $t = g * x$ oraz $t = h * y$ dla pewnych elementów $x, y \in G$. W wyniku przemienności i idempotentności półgrupy G otrzymujemy

$$(g * h) * x = h * (g * x) = h * t =$$
$$= h * (h * y) = (h * h) * y = h * y = t,$$

czyli $t \in A_{g*h}$, a więc $A_g \cap A_h \subseteq A_{g*h}$. Jako wniosek mamy

$$A_g \bigcap A_h = A_{g*h}.$$

Stąd też wynika, że $\{A_g \mid g \in G\}$ jest podpółgrupą w $(2^G, \cap)$, co powoduje, że odwzorowanie

$$\theta : G \ni g \mapsto A_g \in \{A_g \mid g \in G\}$$

jest homomorfizmem półgrup. Odwzorowanie θ jest suriekcją. Jeśli $g, h \in G$ i $\theta(g) = \theta(h)$, to $A_g = A_h$. Z równości $g * g = g$ oraz $h * h = h$ wyciągamy, że $g \in A_h$ oraz $h \in A_g$, a więc $g = h * x$ i $h = g * y$ dla pewnych elementów $x, y \in G$. Na tej podstawie wnosimy, że

$$h = g * y = (h * x) * y = h * x * y,$$
$$g = h * x = (g * y) * x = g * x * y.$$

Zatem

$$h = h * x * y = (g * y) * (x * y) = g * x * y = g,$$

czyli θ jest odwzorowaniem iniektywnym. Wnosimy, że G wkłada się w półgrupę $(2^G, \cap)$. □

■ Przez $\mathcal{I} = \{0, 1\}$ będziemy oznaczać półgrupę względem mnożenia „$\cdot$" określonego równościami

$$\begin{aligned} 0 \cdot 0 &= 0, \\ 1 \cdot 0 &= 0, \\ 0 \cdot 1 &= 0, \\ 1 \cdot 1 &= 1. \end{aligned}$$

Zaznaczmy bez dowodu, że zachodzi

Wniosek 2.2.9. *Przemienna półgrupa idempotentna $(G, *)$ wkłada się w pewną potęgę kartezjańską półgrupy $\mathcal{I}$.*

Ćwiczenia 2.2.10.

(**1**) Niech $F(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ jest funkcją}\}$ oraz $X \subseteq F(\mathbb{R})$. Sprawdzić, czy X jest półgrupą względem złożenia odwzorowań, jeśli:
(a) X jest zbiorem wielomianów stopni parzystych;
(b) X jest zbiorem wielomianów stopnia 1 z najwyższym współczynnikiem równym 1;
(c) X jest zbiorem wielomianów stopnia 1;
(d) $X = \{f \in F(\mathbb{R}) \mid f(-x) = f(x), \text{ gdy } x \in [-1, 1]\}$;
(f) $X = \{f \in F(\mathbb{R}) \mid f(-x) = -f(x), \text{ gdy } x \in [-1, 1]\}$.
(**2**) Niech X będzie zbiorem niepustym oraz $a * b = b$ dla wszystkich $a, b \in X$. Udowodnić, że $(X, *)$ jest półgrupą idempotentną.
(**3**) Na zbiorze liczb rzeczywistych $\mathbb{R}$ rozpatrzmy regułę „$*$" taką, że $a * b = \max\{a, b\}$. Udowodnić, że $(\mathbb{R}, *)$ jest półgrupą przemienną.
(**4**) Niech $n \in \mathbb{N}^*$ oraz $X_n = \{nx + 1 \mid x \in \mathbb{N}\}$. Udowodnić, że X_n jest półgrupą względem mnożenia liczb.
(**5**) Udowodnić, że zbiór macierzy nieosobliwych (odpowiednio osobliwych) stopnia $n \geqslant 2$ tworzy półgrupę względem działania mnożenia macierzy.
(**6**) Udowodnić, że półgrupa $G = \{-1, 1\}$ z działaniem mnożenia liczb nie jest izomorficzna z półgrupą $X = \{a, b\}$ z taką tabelką mnożenia:

$\cdot$	a	b
a	a	b
b	b	b

(**7**) Na zbiorze X^4, gdzie $X \neq \varnothing$, rozpatrzmy regułę „$*$" taką, że $(a,b,c,d) * (x,y,z,t) = (a,b,z,t)$. Udowodnić, że $(X,*)$ jest półgrupą oraz $\alpha^2 = \alpha$ dla każdego $\alpha \in X^4$;
(**8**) Udowodnić, że ogół wszystkich idempotentów półgrupy przemiennej tworzy jej podpółgrupę.
(**9**) Sprawdzić, czy $\varphi : G \to H$ jest homomorfizmem półgrup, jeśli:
(a) $G = (\mathbb{R}^n, +)$, $H = (\mathbb{R}, +)$ oraz $\varphi(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i$;
(b) $G = (\{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid 1, a \in \mathbb{Q}\}, \cdot)$, $H = (\mathbb{Q}, +)$ oraz $\varphi(\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}) = a$;
(c) G jest półgrupą słów nad alfabetem $\{a, b\}$, $G = H$ oraz $\varphi(x) = x^3$;
(d) $G = (\mathbb{Q}, \cdot)$, $H = (\mathbb{Z}, \cdot)$ oraz $\varphi(q) = [q]$ jest częścią całkowitą liczby wymiernej q.
(**10**) Udowodnić, że każda półgrupa skończona posiada idempotent.

Uwagi. Termin „półgrupa" wprowadzono do algebry w 1904 r. w książce *J.-A. de Séguier. Élements de la Théorie des Groupes Abstraits*. Pierwsza publikacja naukowa o półgrupach należy do A.K. Suszkiewicza(2). Rozwój teorii półgrup na początku toczył się (w pewnym sensie) „równolegle" do rozwoju teorii grup, chociaż i z pewnym opóźnieniem w czasie.

(2) Anton Kazimirowicz Suszkiewicz (1889–1961)

2.3. Grupy i ich własności elementarne

Pojęcie grupy jest jednym z centralnych we współczesnej algebrze.

■ Zbiór G jest nazywany *grupą* (względem działania „$*$”), jeśli są spełnione następujące warunki:
0_1) $G \neq \varnothing$;
0_2) „$*$” jest działaniem algebraicznym na zbiorze G;
1) działanie „$*$” jest łączne na G, czyli

$$\forall_{a,b,c\in G} : (a * b) * c = a * (b * c);$$

2) istnieje element neutralny $e \in G$ względem działania „$*$”, czyli

$$\exists_{e\in G}\forall_{a\in G} : a * e = a = e * a;$$

3) każdy element $a \in G$ ma odwrotny do siebie w G względem „$*$”, czyli

$$\forall_{a\in G}\exists_{a^{-1}\in G} : a * a^{-1} = e = a^{-1} * a.$$

■ Zbiór G jest nazywany grupą *przemienną* (*komutatywną* lub *abelową*[3]) względem działania „$*$”, jeśli zachodzą warunki 0_1), 0_2), 1), 2), 3) oraz, oprócz tego,
4) działanie „$*$” jest przemienne, czyli

$$\forall_{a,b\in G} : a * b = b * a.$$

■ Skrót, że para $(G, *)$ jest grupą, oznacza, że zbiór G tworzy grupę względem działania „$*$”.

■ Jeśli warunek 4) w grupie $(G, *)$ nie jest spełniony, to grupa G jest nazywana *nieabelową*.

■ Grupa G jest nazywana *skończoną*, jeśli składa się ze skończonej liczby elementów; w innym przypadku grupa G jest nazywana *nieskończoną*. Z definicji grupy i twierdzeń 2.1.3 oraz 2.1.4 otrzymujemy następny

[3] Niels Henrik Abel (1802–1829)

Wniosek 2.3.1.

(a) *W każdej grupie* $(G, *)$ *istnieje dokładnie jeden element neutralny.*

(b) *W dowolnej grupie* $(G, *)$ *każdy element* $a \in G$ *ma dokładnie jeden odwrotny do siebie element* $a^{-1} \in G$.

□

Przykłady 2.3.2.

(**1**) Każdy ze zbiorów $\mathbb{Z}$, $\mathbb{Q}$ oraz $\mathbb{R}$ jest grupą abelową względem działania dodawania liczb „+". Zbiór liczb naturalnych $\mathbb{N}$ nie jest grupą względem „+", bo, na przykład, liczba naturalna 1 nie posiada liczby przeciwnej do siebie -1 w zbiorze $\mathbb{N}$.

(**2**) Zbiory $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ oraz $\mathbb{C}$ nie są grupami względem działania mnożenia liczb „·", ponieważ liczba 0 nie posiada elementu odwrotnego do siebie. Niech $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ oraz $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Wtedy każdy ze zbiorów $\mathbb{Q}^*$, $\mathbb{R}^*$ oraz $\mathbb{C}^*$ jest grupą względem mnożenia liczb „·".

(**3**) Para $(\mathbb{R}, \diamond)$ jest grupą abelową z elementem neutralnym -2, gdzie $a \diamond b = a + b + 2$ dla elementów $a, b \in \mathbb{R}$. Przekonajmy się o tym.

(3_0) W rzeczy samej, „$\diamond$" jest działaniem algebraicznym określonym na zbiorze niepustym $\mathbb{R}$.

(3_1) Jeśli a, b, c są dowolnymi liczbami rzeczywistymi, to

$$\begin{array}{rcl} (a \diamond b) \diamond c & = & (a + b + 2) \diamond c = (a + b + 2) + c + 2 = a + b + c + 4, \\ a \diamond (b \diamond c) & = & a \diamond (b + c + 2) = a + (b + c + 2) + 2 = a + b + c + 4, \end{array}$$

a więc działanie „$\diamond$" jest łączne.

(3_2) Jeśli a, b są dowolnymi liczbami rzeczywistymi, to

$$a \diamond b = a + b + 2 = b + a + 2 = b \diamond a,$$

czyli działanie „$\diamond$" jest przemienne.

(3_3) Znajdźmy taki element $e \in \mathbb{R}$, że $e \diamond a = a = a \diamond e$. W tym celu rozwiążmy nieskończony układ równań

$$\begin{cases} e + a + 2 = a, \\ \forall_{a \in \mathbb{R}} \end{cases}$$

z niewiadomym e. Z tego układu otrzymujemy, że $e = -2$. Zatem $e = -2$ jest elementem neutralnym w $\mathbb{R}$ względem „$\diamond$".

(3_4) Niech a będzie dowolnym elementem z $\mathbb{R}$. Znajdźmy takie $x \in \mathbb{R}$, że $a \diamond x = -2 = x \diamond a$, czyli $a + x + 2 = -2$, a stąd $x = -a - 4 \in \mathbb{R}$. To oznacza, że

$$a^{-1} = -a - 4 \in \mathbb{R}.$$

Zatem $(\mathbb{R}, \diamond)$ jest grupą abelową.

Twierdzenie 2.3.3. *Niech* $(G, *)$ *będzie grupą (z elementem neutralnym* e*) oraz* $a, b, c \in G$. *Wtedy zachodzą następujące własności:*

(1) *w grupie* G *zachodzi prawo skracania dla działania „*$*$*", czyli z równości* $a * b = a * c$ *wynika, że* $b = c$ *oraz z równości* $b * a = c * a$ *wynika, że* $b = c$;

(2) $(a^{-1})^{-1} = a$ *oraz*

$$(a * b)^{-1} = b^{-1} * a^{-1};$$

(3) *każde równanie (tutaj x, y są niewiadome)*

$$a * x = b \text{ oraz } y * a = b \tag{2.2}$$

ma dokładnie jedno rozwiązanie w grupie G.

Dowód. (1) Jeśli $a * b = a * c$, to

$$\begin{gathered} b = e * b = (a^{-1} * a) * b = a^{-1} * (a * b) = \\ = a^{-1} * (a * c) = (a^{-1} * a) * c = e * c = c, \end{gathered}$$

a zatem $b = c$. Podobnym sposobem z równości $b * a = c * a$ wynika, że $b = c$.

(2) Zauważamy, że

$$a^{-1} * (a^{-1})^{-1} = e = (a^{-1})^{-1} * a^{-1}$$

oraz

$$a^{-1} * a = e = a * a^{-1},$$

skąd $(a^{-1})^{-1} = a$ po zastosowaniu prawa skracania. Skoro

$$\begin{gathered} (a * b) * (b^{-1} * a^{-1}) = ((a * b) * b^{-1}) * a^{-1} = \\ = (a * (b * b^{-1})) * a^{-1} = \\ = (a * e) * a^{-1} = a * a^{-1} = e \\ \text{oraz} \\ (b^{-1} * a^{-1}) * (a * b) = ((b^{-1} * a^{-1}) * a) * b = \\ = (b^{-1} * (a^{-1} * a)) * b = (b^{-1} * e) * b = b^{-1} * b = e, \end{gathered}$$

to

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

(3) Jeśli $a * x = b$, to

$$x = (a^{-1} * a) * x = a^{-1} * (a * x) = a^{-1} * b$$

jest dokładnie jednym rozwiązaniem pierwszego równania. Podobnie $y = b * a^{-1}$ jest rozwiązaniem drugiego równania. □

Wniosek 2.3.4. *Zbiór niepusty G, w którym jest określone łączne działanie algebraiczne „$*$” oraz równania $a*x = b$ i $y*a = b$ mają rozwiązania w G dla wszystkich $a, b \in G$, jest grupą.*

Dowód. Niech $a \in G$. Wtedy istnieje taki element $e \in G$, że $e * a = a$. Podobnie, jeśli $c \in G$ jest dowolnym elementem, to znajdzie się taki element $g \in G$, że $a * g = c$, a stąd

$$e * c = e * (a * g) = (e * a) * g = a * g = c, \tag{2.3}$$

czyli dla każdego elementu $c \in G$ zachodzi $e * c = c$. Podobnie istnieje taki element $e' \in G$, że $a * e' = a$. Jeśli c jest dowolnym elementem z G, to $h * a = c$ dla pewnego elementu $h \in G$, a zatem

$$c * e' = (h * a) * e' = h * (a * e') = h * a = c. \tag{2.4}$$

Jako wniosek z tego, biorąc $c = e'$ w równości (2.3) oraz $c = e$ w równości (2.4), otrzymujemy $e' = e*e' = e$ (czyli e jest elementem neutralnym w G względem „$*$”). Z tego, że równania (2.2) mają rozwiązania wynika, że znajdą się takie elementy $b_1, b_2 \in G$, dla których $a * b_1 = e$ oraz $b_2 * a = e$, a stąd

$$b_1 = e * b_1 = (b_2 * a) * b_1 = b_2 * (a * b_1) = b_2 * e = b_2.$$

Zatem $b_1 = b_2$ jest odwrotnym do elementu a i G jest grupą. □

Wniosek 2.3.5. *Zbiór skończony niepusty G z określonym na nim łącznym działaniem algebraicznym „$*$” jest grupą wtedy i tylko wtedy, gdy w G zachodzi prawo skracania dla „$*$”.*

Dowód. ($\Rightarrow$) Wynika z twierdzenia 2.3.3.

($\Leftarrow$) Niech $a \in G$. Wtedy

$$\phi_a : G \ni x \mapsto a * x \in G,$$

jest odwzorowaniem. Jeśli $\phi_a(x) = \phi_a(y)$ dla pewnych elementów $x, y \in G$, to $a*x = a*y$ i z powodu prawa skracania $x = y$. To oznacza, że ϕ_a jest iniekcją. Skoro zbiór G jest skończony i odwzorowanie ϕ_a jest iniektywne, to również jest bijektywne (przekonać się samodzielnie). Zatem

$$\{a * x \mid x \in G\} = \phi_a(G) = G,$$

czyli każdy element $b \in G$ możemy przedstawić w postaci $b = a * x$, a to znaczy, że równanie $b = a * x$ ma rozwiązanie w grupie G. Podobnie możemy udowodnić, że równanie $y * a = b$ posiada rozwiązanie w G. Z powodu wniosku 2.3.4 otrzymujemy, że $(G, *)$ jest grupą. □

Przykłady 2.3.6.

(**1**) Zbiór $\{-1, 1\}$, składający się z dwóch liczb całkowitych 1 oraz -1, jest grupą abelową (względem mnożenia liczb „$\cdot$").

(**2**) Grupa addytywna $(\{0\}, +)$ (nazywana *zerową*), gdzie $0 + 0 = 0$, oraz grupa multiplikatywna $(\{1\}, \cdot)$ (nazywana *jednostkową*), gdzie $1 \cdot 1 = 1$, są skończonymi grupami o jednym elemencie (które często są nazywane grupami *trywialnymi*).

(**3**) Na zbiorze klas reszt

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

modulo n wprowadźmy dodawanie „$+$": jeśli $\overline{k}, \overline{s} \in \mathbb{Z}_n$, to

$$\overline{k} + \overline{s} = \overline{k+s} = \overline{r},$$

gdzie r jest resztą z dzielenia liczby całkowitej $k + s$ przez n. Łatwo przekonać się, że $(\mathbb{Z}_n, +)$ jest grupą abelową (z elementem neutralnym $\overline{0}$). Grupa klas reszt $(\mathbb{Z}_n, +)$ składa się z n elementów.

(**4**) Niech

$$\mathbb{Q}_+ = \{r \in \mathbb{Q} \mid r > 0\}.$$

Wtedy mnożenie liczb „$\cdot$" jest działaniem algebraicznym w $\mathbb{Q}_+$, bo iloczyn dwóch dodatnich liczb wymiernych jest jednoznacznie określoną dodatnią liczbą wymierną. Łatwo przekonać się, że $(\mathbb{Q}_+, \cdot)$ jest nieskończoną grupą abelową.

(**5**) Zbiór

$$GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid \det A \neq 0\}$$

macierzy nieosobliwych tworzy grupę względem mnożenia macierzy (przekonać się samodzielnie). Grupa $GL_n(\mathbb{R})$ jest nazywana *ogólną grupą liniową* stopnia n nad ciałem liczb rzeczywistych $\mathbb{R}$.

* * *

■ **Rząd grupy oraz rząd elementu w grupie.** Niech $(G, *)$ będzie grupą (z elementem neutralnym e). Wtedy:

- moc zbioru G jest nazywana *rzędem* grupy G (co oznaczamy przez $|G|$);
- element $a \in G$ jest nazywany *elementem rzędu skończonego*, jeśli istnieje taka niezerowa liczba naturalna k, że

  $$a^k = e;$$

  najmniejsze wśród takich $k \in \mathbb{N}^*$ jest nazywane *rzędem* elementu a (co zapisujemy w postaci $|a| = k$ lub równoważnie $o(a) = k$);

- jeśli a^k nie jest równe elementowi neutralnemu e dla żadnego $k \in \mathbb{N}^*$, to będziemy mówić, że a jest elementem *rzędu nieskończonego* (co krótko oznaczamy przez $|a| = \infty$).

Lemat 2.3.7. *Niech $(G, *)$ będzie grupą z elementem neutralnym e, $a \in G$ oraz $k \in \mathbb{N}^*$. Jeśli $a^k = e$, to $|a| \leqslant k$ oraz rząd $|a|$ dzieli liczbę k.*

Dowód. Z założenia wynika, że a jest elementem rzędu skończonego. Niech $|a| = n$ dla pewnego $n \in \mathbb{N}^*$. Na mocy twierdzenia o dzieleniu z resztą (patrz twierdzenie 1.2.3) znajdzie się taka para liczb całkowitych q, r, że

$$k = qn + r \text{ oraz } 0 \leqslant r < |n| = n.$$

Stąd otrzymujemy, że

$$e = a^k = a^{qn+r} = a^{nq} * a^r = (a^n)^q * a^r = e^q * a^r = e * a^r = a^r.$$

W wyniku minimalnego wyboru liczby n wnosimy, że $r = 0$. Zatem $k = nq$ oraz $|a| = n \leqslant k$. □

Lemat 2.3.8. *Niech $(G, *)$ będzie grupą, $a \in G$ oraz $m, n \in \mathbb{Z}$. Wtedy*

$$a^m * a^n = a^{m+n} \text{ oraz } (a^m)^n = a^{mn}.$$

Dowód jest podobny jak dla lematu 2.2.2.

□

■ Dalej, jeśli nie określono, jak oznaczamy działanie w grupie G, to wykorzystujemy „najkrótsze" oznaczenie dla symbolu działania, czyli kropkę „$\cdot$" (i nazywamy takie działanie *mnożeniem*). Symbol „$\cdot$" często w celu skrócenia w notatkach opuszczamy.

■ Niech $(G, +)$ będzie grupą addytywną (z elementem zerowym 0). *Rzędem* $|g|$ (lub równoważnie $o(g)$) elementu g w grupie G jest nazywana najmniejsza liczba niezerowa naturalna n z własnością

$$ng = 0$$

(co oznaczamy przez $|g| = n$). Jeśli taka liczba $n \in \mathbb{N}^*$ nie istnieje, to będziemy mówić, że g jest *elementem rzędu nieskończonego* (i oznaczać przez $|g| = \infty$).

Przykłady 2.3.9.

(**1**) Jeśli e jest elementem neutralnym grupy G, to $e^1 = e$, a więc rząd $|e| = 1$.

(**2**) Jeśli 1 jest elementem addytywnej grupy liczb całkowitych $\mathbb{Z}$, to

$$n \cdot 1 \neq 0$$

dla każdego niezerowego $n \in \mathbb{N}$, a zatem rząd $|1| = \infty$.

(**3**) $(\mathbb{R}, +)$ jest grupą nieskończoną, bo moc $\operatorname{card} \mathbb{R} = \infty$.

(**4**) Grupa multiplikatywna $\{-1, 1\}$ składa się z dwóch elementów.

■ Niech p będzie liczbą pierwszą. Jeśli rzędy wszystkich elementów grupy G są potęgami ustalonej liczby p (czyli dla każdego elementu $g \in G$ znajdzie się taka nieujemna liczba całkowita k, że

$$g^{p^k} = e$$

jest jednością grupy G), to G jest nazywana *p-grupą*. Grupa, której wszystkie elementy mają rzędy skończone, jest nazywana *torsyjną* (lub *periodyczną*). Jeśli zaś w grupie G każdy niejednostkowy (odpowiednio niezerowy) element ma rząd nieskończony, to będziemy mówić, że G jest grupą *beztorsyjną*. Grupa, która nie jest ani torsyjną, ani beztorsyjną, jest nazywana *mieszaną*.

Przykłady 2.3.10.

(**1**) Każda grupa skończona (odpowiednio każda p-grupa) jest torsyjna.

(**2**) Grupa multiplikatywna $\{-1, 1\}$ jest 2-grupą, bo

$$(-1)^1 \neq 1,\ (-1)^2 = 1 \text{ oraz } 1^1 = 1,$$

a więc rzędy $o(-1) = 2$ oraz $o(1) = 1 = 2^0$.

(**3**) Skoro w grupie addytywnej liczb całkowitych $\mathbb{Z}$ dla każdej liczby niezerowej $z \in \mathbb{Z}$ zachodzi $nz \neq 0$ dla wszystkich $n \in \mathbb{N}^*$, to rząd $o(z) = \infty$, a więc $(\mathbb{Z}, +)$ jest grupą beztorsyjną.

(**4**) Grupa multiplikatywna liczb rzeczywistych niezerowych $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ jest mieszana, bo, na przykład, $o(-1) = 2$ oraz $o(5) = \infty$.

(**5**) Znajdźmy rzędy elementów grupy multiplikatywnej liczb wymiernych niezerowych $\mathbb{Q}^*$. Oczywiście, że $|1| = 1$, $|-1| = 2$. Jeśli $a \in \mathbb{Q} \setminus \{0, -1, 1\}$, to $a^k \neq 1$ dla każdego $k \in \mathbb{N}^*$, a zatem $|a| = \infty$. To znaczy, że grupa $(\mathbb{Q}^*, \cdot)$ jest mieszana.

Ćwiczenia 2.3.11.

(**1**) Sprawdzić, czy zbiór G jest grupą względem reguły „$*$”, jeśli:
(a) $G = \mathbb{Q}$, gdzie $x * y = x^y$;

(b) $G = \mathbb{N}$, gdzie $x * y = \mathrm{NWD}(x, y)$;
(c) $G = \{x \in \mathbb{R} \mid -1 < x < 1\}$ oraz $x * y = \frac{x+y}{1+xy}$;
(d) $G = \mathbb{R} \times \mathbb{R}$, gdzie $(x, y) * (z, t) = (x, t)$;
(e) $G = \mathbb{Q}$, gdzie $x * y = x^2 - 2xy + y^2$;
(f) $G = M_2(\mathbb{Q})$, gdzie $A * B = \frac{1}{2}(A + B)$;
(g) $G = \mathbb{R}$, gdzie $x * y = x + y - xy$;
(h) $G = \{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R}, a^2 + b^2 \neq 0\}$, gdzie „$*$" jest mnożeniem macierzy;
(i) $G = \{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R}\}$, gdzie $A * B = AB - BA$ dla $A, B \in G$;
(j) $G = \{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R}\}$, gdzie „$*$" jest mnożeniem macierzy;
(k) $G = \{\begin{bmatrix} a & b \\ \varkappa b & a \end{bmatrix} \mid a, b \in \mathbb{Q}\}$, gdzie $\varkappa$ jest ustaloną liczbą rzeczywistą, a „$*$" jest mnożeniem macierzy;
(l) $G = \{(a, b) \mid a, b \in \mathbb{Q}, a \neq 0\}$, gdzie „$*$" jest działaniem określonym wzorem $(a, b) * (c, d) = (ac, ad + b)$;
(m) $G = \mathbb{C} \setminus \{-1\}$, gdzie reguła „$*$" jest określona za pomocą wzoru $x * y = x + y + xy$.
(**2**) Udowodnić, że grupa G jest abelowa, jeśli dla dowolnych elementów $a, b \in G$ zachodzi warunek: (a) $a^2 = 1$; (b) $a^2 = a^4$; (c) $(ab)^2 = a^2b^2$.
(**3**) Udowodnić, że:
(a) zbiór $[0, 1) = \{x \in \mathbb{R} \mid 0 \leqslant x < 1\}$ względem działania „$*$", określonego przez regułę $x * y = x + y - [x + y]$, gdzie $[x]$ jest częścią całkowitą liczby rzeczywistej x, tworzy nieskończoną grupę abelową;
(b) zbiór $\mathbb{R}^3 = \{(a, b, c) \mid a, b, c \in \mathbb{R}\}$ jest nieskończoną grupą abelową względem działania „$*$", określonego przez wzór $(a, b, c) * (u, v, w) = (a + u, b + v, c + w + av)$;
(c) zbiór $\{\begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \mid 0, a, b, c \in \mathbb{R}\}$ tworzy grupę względem działania określonego przez wzór $A \circ B = A + B + \frac{1}{2}(AB - BA)$;
(d) zbiór $\{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$ jest grupą względem dodawania macierzy;
(e) zbiór $\{\begin{bmatrix} a & -3b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R}$ oraz $a^2 + b^2 > 0\}$ jest grupą względem mnożenia macierzy;
(f) zbiór $\{z \in \mathbb{C} \mid |z| = 1\}$ jest grupą względem mnożenia liczb zespolonych;
(g) zbiór $[0, 1)$ jest grupą względem działania „$\circ$", określonego przez

$$a \circ b = \begin{cases} a + b, & \text{gdy } a + b < 1, \\ a + b - 1, & \text{gdy } a + b \geqslant 1; \end{cases}$$

(h) zbiór $\mathbb{R}$ jest grupą względem działania „$*$", określonego przez regułę $a * b = ab - a - b + 2$;
(i) zbiór $\{\phi_{a,b} : \mathbb{R} \to \mathbb{R} \mid a, b \in \mathbb{R}, a \neq 0\}$ przekształceń afinicznych postaci $\phi_{a,b}(x) = ax + b$ $(x \in \mathbb{R})$ jest grupą względem działania złożenia odwzorowań;
(j) zbiór $[0, 1)$ jest grupą względem działania „$*$", określonego przez $a * b = \{a + b\}$, gdzie $\{z\}$ jest częścią ułamkową liczby rzeczywistej z;
(k) zbiór 2^X, gdzie X jest zbiorem niepustym, jest grupą względem działania różnicy symetrycznej;
(l) zbiór niepusty X jest grupą względem działania „$*$", określonego przez $x * y = x$;
(m) zbiór macierzy $\{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{R}$ oraz $a^2 + b^2 \neq 0\}$ tworzy grupę względem mnożenia macierzy;

(n) zbiór $\mathbb{R}_+ = \{r \in \mathbb{R} \mid r > 0\}$ jest grupą względem działania „$*$", określonego przez

$$a * b = \begin{cases} x, & \text{gdy } x \leqslant y, \\ y, & \text{gdy } y \leqslant x; \end{cases}$$

(o) zbiór macierzy $\{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R}\}$ jest grupą względem działania „$[-,-]$", określonego wzorem $[A, B] = AB - BA$;

(p) zbiór macierzy $\{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R}\}$ jest grupą względem działania „$\odot$", określonego przez $A \odot B = \frac{1}{2}(AB + BA)$;

(q) zbiór funkcji $\{f : \mathbb{R} \to \mathbb{R} \mid f$ jest ciągłą funkcją ściśle rosnącą, $f(0) = 0, f(1) = 1\}$ jest grupą względem złożenia odwzorowań.

(**4**) Sprawdzić, czy grupą jest zbiór:

(a) $\mathbb{R}$ względem działania „$\circ$", określonego regułą $a \circ b = a + b + ab$;

(b) $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ względem dodawania liczb;

(c) $n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$ względem dodawania liczb;

(d) $[-1, 1] \setminus \{0\}$ względem mnożenia liczb;

(e) $\mathbb{R}_+$ względem mnożenia liczb;

(f) $\{-1, 0, 1\}$ względem mnożenia liczb;

(g) $\mathbb{R}$ względem działania „$*$", określonego przez $a * b = a + b + 2$;

(h) $X = \{x \in \mathbb{R} \mid x > 1\}$ względem działania „$*$", określonego przez $a * b = ab - a - b + 2$;

(i) $\mathbb{Q} \times \mathbb{Q}^*$ względem działania „$*$", zadanego regułą $(a, b) * (c, d) = (ac - 2bd, ad + bc)$.

(**5**) Znaleźć rząd elementu x w grupie G, jeśli:

(a) $x = i$ oraz $G = \mathbb{C}^*$;

(b) $x = i$ oraz $G = \mathbb{C}$;

(c) $x = \overline{2}$ oraz $G = \mathbb{Z}_{12}$;

(d) $x = \begin{bmatrix} -1 & z \\ 0 & 1 \end{bmatrix}$ oraz $G = GL_2(\mathbb{C})$, gdzie $z \in \mathbb{C}$;

(e) $x = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i$ oraz $G = \mathbb{C}^*$.

(**6**) Niech G będzie grupą, $x \in G$ oraz $|x| = n$. Udowodnić, że jeśli $n = mk$ dla pewnych liczb naturalnych m, k, to rząd $|x^m| = k$.

(**7**) Udowodnić, że $GL_2(\mathbb{Z}_2)$ jest grupą nieabelową, wypisać wszystkie jej elementy i znaleźć ich rzędy.

(**8**) Niech G będzie grupą, a g jej elementem rzędu skończonego oraz $k \in \mathbb{N}^*$. Udowodnić, że:

(a) jeśli $|g| = mn$, gdzie m i n są względnie pierwszymi liczbami całkowitymi, to w grupie G znajdą się takie elementy u, v, że $g = uv = vu$, $|u| = m$ oraz $|v| = n$;

(b) $|g^k| = |g|$ wtedy i tylko wtedy, gdy $\mathrm{NWD}(k, |g|) = 1$.

(**9**) Udowodnić, że jeśli grupa G ma dokładnie jeden element z rzędu 2, to $zg = gz$ dla wszystkich elementów $g \in G$.

Uwagi. Jeden z założycieli współczesnej algebry, matematyk francuski E. Galois[(4)], wprowadził pojęcie grupy w 1830 r. W 1870 r. matematyk francuski C. Jordan[(5)] i w 1884 r. matematyk niemiecki F. Klein[(6)] w pojęciu grupy uwzględniali tylko pewnik o algebraiczności działania (patrz

[(4)] Evariste Galois (1811–1832)

[(5)] Camille Jordan (1838–1922)

[(6)] Felix Klein (1849–1925)

warunek 0_2) z definicji grupy). A. Cayley w 1854 r. dołączył do definicji grupy aksjomat łączności i aksjomat istnienia elementu neutralnego. W definicji grupy wprowadzonej przez A. Cayleya nie było warunku o istnieniu odwrotności. I tylko w 1882 r. matematycy niemieccy W. von Dyck[7] oraz H. Weber[8] wprowadzili współczesną definicję grupy, która została ogólnie przyjęta po publikacji książki H. Webera w 1886 r. Współczesną definicją grupy operowali już E. Hundington i E. Moore w 1902 r. Matematyk francuski A. Cauchy[9] w 1815 r. zdefiniował pojęcie rzędu elementu w grupie.

[7] Walter von Dyck (1856–1934)
[8] Heinrich Weber (1842–1913)
[9] Augustin Louis Cauchy (1789–1857)

2.4. Grupy przekształceń. Grupy permutacji

■ Niech X będzie dowolnym zbiorem niepustym,

$$\mathbb{S}(X) = \{f : X \to X \mid f \text{ jest odwzorowaniem bijektywnym}\}.$$

Odwzorowanie postaci $f : X \to X$ jest nazywane *przekształceniem* zbioru X. Każde odwzorowanie bijektywne $f : X \to X$ jest nazywane *permutacją* (lub *symetrią*) zbioru X.

Twierdzenie 2.4.1. *$\mathbb{S}(X)$ jest grupą względem złożenia odwzorowań „$\circ$”.*

Dowód. Odwzorowanie jednostkowe $\mathrm{id}_X : X \ni x \mapsto x \in X$ jest bijekcją, a więc $i_X \in \mathbb{S}(X)$ oraz $\mathbb{S}(X) \neq \varnothing$. Złożenie $f \circ g$ dwóch odwzorowań bijektywnych $g : X \to X$ i $f : X \to X$ istnieje i jest bijekcją na podstawie wniosku 1.4.9, a więc złożenie „$\circ$” jest działaniem algebraicznym w zbiorze $\mathbb{S}(X)$. Na mocy twierdzenia 1.4.6 kompozycja „$\circ$” jest łączna. Z lematu 1.4.5 wynika też, że id_X jest elementem neutralnym w $\mathbb{S}(X)$ (względem „$\circ$”). Jeśli $f \in \mathbb{S}(X)$, to na podstawie twierdzenia 1.4.11 istnieje f^{-1} i z wniosku 1.4.12 otrzymujemy, że $f^{-1} \in \mathbb{S}(X)$. □

■ Para $(\mathbb{S}(X), \circ)$ jest nazywana *grupą przekształceń* zbioru X (lub *grupą symetryczną* zbioru X).

■ Rozpatrzmy przypadek cząstkowy, gdy $X = \{1, 2, \ldots, n\}$. Wtedy grupa

$$\mathbb{S}_n = \mathbb{S}(\{1, 2, \ldots, n\})$$

jest nazywana *grupą symetryczną stopnia n* lub *grupą permutacji* zbioru $\{1, 2, \ldots, n\}$. Każdy element grupy $\mathbb{S}_n$ jest nazywany *permutacją* zbioru $\{1, 2, \ldots, n\}$ (lub *permutacją stopnia n*).

* * *

■ **Postać kanoniczna permutacji.** Jeśli $\sigma \in \mathbb{S}_n$, to

$$\begin{cases} \sigma(1) & = i_1, \\ \sigma(2) & = i_2, \\ & \vdots \\ \sigma(n) & = i_n \end{cases}$$

dla pewnych elementów $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, n\}$, przy czym z bijektywności odwzorowania $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ wynika, że elementy $i_1, i_2, \ldots, i_n$ są parami różne i

$$\{1, 2, \ldots, n\} = \{i_1, i_2, \ldots, i_n\}.$$

Dlatego zapisujemy permutację σ w postaci

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix},$$

gdzie w wierszu górnym są zawarte po kolei wszystkie elementy zbioru $X = \{1, 2, \ldots, n\}$, a w wierszu dolnym pod każdym $j \in X$ zapisujemy jego obraz $i_j = \sigma(j)$ względem σ, czyli

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

Jeszcze raz przypomnijmy, że każdy element dolnego wiersza należy do zbioru $X = \{1, 2, \ldots, n\}$ i wszystkie elementy dolnego wiersza są parami różne.

* * *

■ **Mnożenie permutacji.** Mamy permutacje $\sigma, \tau \in \mathbb{S}_n$. Wtedy złożenie odwzorowań $\sigma \circ \tau$ tradycyjnie jest nazywane *iloczynem permutacji* i oznaczane symbolem $\sigma\tau$, czyli

$$\sigma\tau(i) = \sigma(\tau(i)) \quad (\text{tutaj} \quad i \in \{1, 2, \ldots, n\}).$$

Permutacja

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix} \in \mathbb{S}_n$$

jest nazywana *odwrotną* do permutacji σ, a

$$e = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} \in \mathbb{S}_n$$

jest nazywana permutacją *jednostkową* (lub *tożsamościową*) stopnia n.

Przykład 2.4.2.

Niech

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \in \mathbb{S}_4.$$

Wtedy

$$\begin{array}{rcl} \sigma\tau(1) & = & \sigma(\tau(1)) = \sigma(4) = 3, \\ \sigma\tau(2) & = & \sigma(\tau(2)) = \sigma(3) = 4, \\ \sigma\tau(3) & = & \sigma(\tau(3)) = \sigma(2) = 1, \\ \sigma\tau(4) & = & \sigma(\tau(4)) = \sigma(1) = 2, \end{array}$$

a więc

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

Oprócz tego

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

■ Grupa $\mathbb{S}_n$ (gdzie $n \geqslant 3$) jest nieabelowa, a grupy $\mathbb{S}_1$ i $\mathbb{S}_2$ są abelowe.

■ Niech $k \leqslant n$. Permutacja $\sigma \in \mathbb{S}_n$ jest nazywana *cykliczną długości k* (lub *cyklem długości k*), jeśli istnieje taki podzbiór

$$Y = \{i_1, i_2, \ldots, i_k\} \subseteq \{1, 2, \ldots, n\},$$

że

$$\begin{array}{rcl} \sigma(i_1) & = & i_2, \\ \sigma(i_2) & = & i_3, \\ & \vdots & \\ \sigma(i_{k-1}) & = & i_k, \\ \sigma(i_k) & = & i_1 \end{array}$$

oraz $\sigma(i) = i$ dla wszystkich innych elementów $i \in \{1, 2, \ldots, n\} \setminus Y$ (wtedy krótko zapisujemy

$$\sigma = (i_1, i_2, \ldots, i_k)$$

lub $\sigma = (i_1 i_2 \ldots i_k)$). Oczywiście, że permutacja jednostkowa $e \in \mathbb{S}_n$ jest cyklem długości 1 (i na odwrót). Dla każdej permutacji $\sigma \in \mathbb{S}_n$ zbiór

$$\operatorname{supp} \sigma = \{i \mid \sigma(i) \neq i \text{ oraz } 1 \leqslant i \leqslant n\}$$

jest nazywany jej *nośnikiem.*

Na przykład $\operatorname{supp} e = \varnothing$,

$$\operatorname{supp}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix} = \{1, 3\}.$$

Lemat 2.4.3. *Jeśli* $\varphi, \psi \in \mathbb{S}_n$ *oraz*

$$\operatorname{supp}\varphi \bigcap \operatorname{supp}\psi = \varnothing$$

(wtedy permutacje φ i ψ są nazywane *niezależnymi*)*, to permutacje* φ *oraz* ψ *komutują między sobą, czyli*

$$\varphi\psi = \psi\varphi.$$

Dowód. Ćwiczenie. □

Twierdzenie 2.4.4. *Każda permutacja* $\sigma \in \mathbb{S}_n$ $(n \in \mathbb{N}^*)$ *jest cyklem lub rozkłada się w iloczyn komutujących cykli niezależnych. Takie rozłożenie permutacji* σ *jest określone dokładnie jednoznacznie* (*z dokładnością do kolejności czynników w iloczynie*).

Dowód. Niech $X = \{1, 2, \ldots, n\}$.

1) *Istnienie rozkładu.* Wybierzmy dowolny element $i \in X$ i rozpatrzmy ciąg

$$\sigma(i), \sigma^2(i), \ldots$$

Skoro σ jest przekształceniem bijektywnym zbioru skończonego X, to $\sigma^k(i) = i$ dla pewnego $k \in \mathbb{N}^*$. Zatem mamy cykl

$$\sigma_i = (i, \sigma(i), \ldots, \sigma^{k-1}(i)).$$

Jeśli $\sigma_i = \sigma$, to rozłożenie permutacji σ jest zakończone i teza zachodzi. Jeśli zaś $\sigma_i \neq \sigma$, to wybierzmy kolejny element j, tak żeby

$$j \in X \setminus \{i, \sigma(i), \ldots, \sigma^{k-1}(i)\},$$

i zbudujmy kolejny cykl

$$\sigma_j = (j, \sigma(j), \ldots, \sigma^{s-1}(j)).$$

gdzie s jest taką najmniejszą liczbą naturalną niezerową, że $\sigma^s(j) = j$. Kontynuując w podobny sposób, przez skończoną liczbę kroków otrzymamy, że zbiór X jest sumą mnogościową nośników parami niezależnych cykli $\sigma_i, \sigma_j, \ldots, \sigma_l$. Zatem

$$\sigma = \sigma_i \sigma_j \cdots \sigma_l.$$

2) *Jednoznaczność.* Załóżmy, że permutacja $\sigma \in \mathbb{S}_n$ jest przedstawiona w postaci iloczynów

$$\sigma = \tau_1 \tau_2 \cdots \tau_s \text{ oraz } \sigma = \pi_1 \pi_2 \cdots \pi_m$$

parami niezależnych cykli $\tau_1, \tau_2, \ldots, \tau_s$ oraz parami niezależnych cykli $\pi_1, \pi_2, \ldots, \pi_m$. Niech $i \in X$ będzie takim elementem, że $\sigma(i) \neq i$. Wtedy znajdą się jednoznacznie określone indeksy $u, v \ \ (1 \leqslant u \leqslant m; 1 \leqslant v \leqslant s)$ takie, że

$$\pi_u(i) \neq i \text{ oraz } \tau_v(i) \neq i.$$

Oczywiście, że $\pi_u(i) = \sigma(i) = \tau_v(i)$. Załóżmy, że

$$\pi_u^k(i) = \sigma^k(i) = \tau_v^k(i)$$

dla pewnych $k \in \mathbb{N}^*$. Wtedy

$$\sigma \pi_u^k(i) = \sigma^{k+1}(i) = \sigma \tau_v^k(i).$$

Skoro $\sigma \pi_u = \pi_u \sigma$ oraz $\sigma \tau_v = \tau_v \sigma$, to

$$\pi_u^k \sigma(i) = \sigma^{k+1}(i) = \tau_v^k \sigma(i),$$

a stąd

$$\pi_u^{k+1}(i) = \sigma^{k+1}(i) = \tau_v^{k+1}(i).$$

Ponieważ każdy cykl jest jednoznacznie określony przez swoje potęgi (to znaczy ich działaniem na zbiorze X), to $\pi_u = \sigma_v$. Stosując dalej rozumowania indukcyjne, otrzymujemy wynik. □

Przykłady 2.4.5.

(**1**) Łatwo otrzymujemy

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 3 & 2 & 5 & 6 & 4 & 8 & 7 \end{pmatrix} = (1)(23)(456)(78) = (23)(456)(78),$$

czyli cykle długości 1 w takich rozłożeniach pomijamy, bo cykle długości 1 w iloczynie nie wpływają na końcowy wynik mnożenia.

(**2**) Permutacja

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix} \in \mathbb{S}_5$$

ma rozłożenie $\sigma = (12345)$ lub, co znaczy to samo,

$$\sigma = (23451) = (34512) = (45123) = (51234).$$

■ Cykl długości 2 jest nazywany *transpozycją*.

Lemat 2.4.6. *Każda permutacja $\sigma \in \mathbb{S}_n$ $(n \geqslant 2)$ rozkłada się w iloczyn transpozycji.*

Dowód. Skoro na mocy twierdzenia 2.4.4 każda permutacja jest iloczynem parami niezależnych cykli, to wystarczy założyć, że $\sigma = (i_1 i_2 \dots i_k)$ jest cyklem. Wtedy

$$\sigma = (i_1 i_2 \dots i_k) = (i_1 i_k)(i_1 i_{k-1}) \cdots (i_1 i_3)(i_1 i_2),$$

a więc teza zachodzi. □

■ Transpozycja $\sigma = (ij) \in \mathbb{S}_n$, gdzie $i, j \in \{1, 2, \dots, n\}$, jest nazywana transpozycją *liczb sąsiednich* (lub *elementarną*), jeśli

$$j - i = \pm 1.$$

Lemat 2.4.7. *Każda transpozycja $\sigma = (ij) \in \mathbb{S}_n$ rozkłada się w iloczyn transpozycji liczb sąsiednich.*

Dowód. Dla pewności załóżmy, że $i < j$. Wtedy

$$\begin{gathered} \sigma = (ij) = \\ = (i, i+1)(i+1, i+2) \cdots (j-2, j-1)(j-1, j) \\ (j-1, j-2) \cdots (i+2, i+1)(i+1, i) \end{gathered}$$

jest iloczynem $2(j-i)-1$ transpozycji liczb sąsiednich. □

Przykłady 2.4.8.

(**1**) Jeśli

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 1 & 6 & 7 & 5 \end{pmatrix} \in \mathbb{S}_7,$$

to $\sigma = (1234)(567) = (14)(13)(12)(57)(56)$.

(**2**) Mamy

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 3 & 1 & 6 \end{pmatrix} = (12435)(6) = (12435) =$$

$$= (15)(13)(14)(12) = (12)(23)(34)(45)(34)(23)(12)(12)(23)(12)(12)(23)(34)(23)(12)(12),$$

na podstawie czego wnosimy, że rozłożenie permutacji w iloczyn transpozycji nie jest jednoznaczne.

■ Niech

$$\tau = \begin{pmatrix} 1 & 2 & \dots & k & \dots & j & \dots & n \\ i_1 & i_2 & \dots & i_k & \dots & i_j & \dots & i_n \end{pmatrix} \in \mathbb{S}_n \ (n \geqslant 2).$$

Będziemy mówić, że para liczb naturalnych i_k, i_j (w dolnym wierszu permutacji τ) tworzy *inwersję*, jeśli

$$k < j \text{ oraz } i_k > i_j.$$

■ Permutacja τ jest nazywana *parzystą* (odpowiednio *nieparzystą*), jeśli liczba wszystkich inwersji (w jej dolnym wierszu) jest parzysta (odpowiednio nieparzysta).

Lemat 2.4.9 (o transpozycji)**.** *Niech $n \geqslant 2$ będzie liczbą całkowitą. Permutacje $\sigma \in \mathbb{S}_n$ oraz $\sigma(kj)$ mają różną parzystość. Permutacje σ oraz $(kj)\sigma$ też mają różną parzystość.*

Dowód. Niech $j = k + 1$. Wtedy

$$\sigma(kj) = \begin{pmatrix} 1 & 2 & \dots & k & k+1 & \dots & n \\ i_1 & i_2 & \dots & i_k & i_{k+1} & \dots & i_n \end{pmatrix} (kj) = \begin{pmatrix} 1 & 2 & \dots & k & k+1 & \dots & n \\ i_1 & i_2 & \dots & i_{k+1} & i_k & \dots & i_n \end{pmatrix}.$$

Jeśli w permutacji σ para liczb i_k, i_{k+1} tworzy inwersję, to w permutacji $\sigma(kj)$ para liczb i_{k+1}, i_k już nie tworzy inwersji i na odwrót. W pozostałych parach liczb w dolnym wierszu tych dwóch permutacji liczby inwersji są takie same.

Teraz, jeśli k, j są dowolnymi liczbami ze zbioru $\{1, 2, \ldots, n\}$ i, na przykład, $k < j$, to według lematu 2.4.7 transpozycja (ij) jest iloczynem $2(j-i)-1$ transpozycji liczb sąsiednich. To oznacza, że liczby inwersji w permutacjach σ i $\sigma(kj)$ różnią się, przy czym ich różnica jest nieparzystą liczbą całkowitą. □

Lemat 2.4.10. *Załóżmy, że permutacja $\sigma \in \mathbb{S}_n$ jest iloczynem k transpozycji ($k \in \mathbb{N}^*$). Jeśli k jest parzyste (odpowiednio nieparzyste), to permutacja σ jest parzysta (odpowiednio nieparzysta).*

Dowód. Stosujemy rozumowania indukcyjne. Jeśli $k = 1$ (czyli w tym przypadku permutacja σ jest transpozycją), to σ jest nieparzysta. Załóżmy, że $k > 1$ oraz teza zachodzi dla $k - 1$. Wtedy teza sprawdza się również dla k, bo liczby k oraz $k - 1$ mają różną parzystość. □

Z lematów 2.4.6, 2.4.7, 2.4.9 oraz 2.4.10 wynika takie

Twierdzenie 2.4.11. *Permutacja $\sigma \in \mathbb{S}_n$ jest parzysta (odpowiednio nieparzysta) wtedy i tylko wtedy, gdy rozkłada się w iloczyn parzystej (odpowiednio nieparzystej) liczby transpozycji.*

□

Proponujemy Czytelnikowi, aby samodzielnie udowodnił następny

Lemat 2.4.12. *Rząd grupy symetrycznej $\mathbb{S}_n$ jest równy $n!$.*

□

■ Odwzorowanie $\operatorname{sgn} : \mathbb{S}_n \ni \sigma \mapsto \operatorname{sgn}\sigma \in \{-1, 1\}$, gdzie

$$\operatorname{sgn}\sigma = \begin{cases} 1, & \text{gdy permutacja } \sigma \text{ jest parzysta,} \\ -1, & \text{gdy permutacja } \sigma \text{ jest nieparzysta,} \end{cases}$$

jest nazywane *znakiem* permutacji σ.

Przykład 2.4.13.

Jeśli $\sigma \in \mathbb{S}_6$ jest permutacją z przykładu 2.4.8(1), to $\operatorname{sgn}\sigma = 1$, czyli permutacja τ jest nieparzysta.

Ćwiczenia 2.4.14.

(**1**) Udowodnić, że jeśli $\sigma \in \mathbb{S}_n$ jest cyklem długości $m \leqslant n$ oraz $\tau \in \mathbb{S}_n$, to $\tau^{-1} \circ \sigma \circ \tau$ jest cyklem długości m.

(**2**) Niech $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 2 & 3 & 1 & 7 & 5 & 4 & 6 \end{pmatrix}$ oraz $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 2 & 4 & 1 & 7 & 5 & 6 & 3 \end{pmatrix}$. Wtedy:

(a) w grupie $\mathbb{S}_8$ obliczyć $\sigma^2 \circ \tau$, $\tau \circ \sigma^2$, $\tau^{-1} \circ \sigma$, $\sigma \circ \tau^{-3}$;

(b) rozłożyć każdą z tych permutacji w iloczyn cykli, transpozycji, transpozycji liczb sąsiednich;

(c) rozwiązać równanie:

(x) $\sigma \circ x = \tau$;

(y) $x \circ \sigma^{-1} = \tau$;

(z) $\tau^{-1} \circ x = \sigma$;

(t) $\tau^2 \circ x = \sigma^3$.

(**3**) Sprawdzić, czy permutacja $\sigma \in \mathbb{S}_8$ jest parzysta, jeśli:

(a) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 3 & 2 & 1 & 6 & 5 & 4 & 8 \end{pmatrix}$;

(b) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 5 & 4 & 3 & 2 & 7 & 8 & 1 \end{pmatrix}$;

(c) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 3 & 7 & 2 & 1 & 6 & 8 & 5 \end{pmatrix}$;

(d) $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 1 & 2 & 6 & 8 & 7 & 4 & 5 \end{pmatrix}$.

Uwagi. Grupy $\mathbb{S}_2$, $\mathbb{S}_3$ oraz $\mathbb{S}_4$ w 1770 r. przebadał J. Lagrange. Matematyk włoski i doktor medycyny P. Ruffini[(10)] w 1799 r. udowodnił, że $|\mathbb{S}_n| = n!$. Funkcja $\operatorname{sgn}\sigma$ pierwszy raz pojawiła się w 1812 r. w pracy A. Cauchy'ego poświęconej wyznacznikom.

E. Galois w 1830 r. wprowadził pojęcie grupy w kontekście grupy permutacji w swojej pracy, która została opublikowana dopiero w 1846 r. Mocną inspiracją do badań grupowych był obszerny traktat o własnościach grup permutacji (w szczególności o ich aspektach geometrycznych) opublikowany przez C. Jordana w 1870 r.

(10) Paolo Ruffini (1765–1822)

2.5. Podgrupy

■ Niech G będzie grupą względem działania „$\cdot$”. Zbiór H jest nazywany *podgrupą* grupy G, jeśli H jest podzbiorem niepustym z G oraz H jest grupą względem tego samego działania „$\cdot$” co i G, lecz zawężonego do podzbioru H. Jeśli H jest podgrupą w G, to będziemy pisać $H \leqslant G$. Podgrupa H grupy G jest nazywana *właściwą* (krótko oznaczamy przez $H < G$), jeśli $H \leqslant G$ oraz $H \neq G$.

■ Skoro każda podgrupa $H \leqslant G$ jest grupą, to możemy mowić o *rzędzie* podgrupy (jako grupy) H.

Przykłady 2.5.1.

(**1**) Ponieważ zbiór $\{-1, 1\}$ jest grupą względem mnożenia liczb, to $\{-1, 1\}$ jest podgrupą rzędu 2 grupy multiplikatywnej niezerowych liczb rzeczywistych $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$.

(**2**) Podobnie $\mathbb{Z}$ jest addytywną grupą liczb całkowitych, a zatem jest podgrupą addytywnej grupy liczb rzeczywistych $\mathbb{R}$.

(**3**) Niech

$$V_4 = \left\{ e,\ (12)(34) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \right.$$

$$\left. (13)(24) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, (14)(23) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \right\}$$

będzie podzbiorem grupy symetrycznej $\mathbb{S}_4$. Łatwo przekonać się, że V_4 jest podgrupą abelową rzędu 4 w $\mathbb{S}_4$. Grupa V_4 jest nazywana *grupą Kleina*.

Twierdzenie 2.5.2 (kryterium podgrupy)**.** *Niech G będzie grupą względem działania „$\cdot$”. Zbiór H jest podgrupą grupy G wtedy i tylko wtedy, gdy są spełnione warunki:*

(0_1) $H \neq \varnothing$;

(0_2) $H \subseteq G$;

(1) *zbiór H jest domknięty względem działania „$\cdot$”, czyli dla dowolnych elementów $h_1, h_2 \in H$ zachodzi $h_1 \cdot h_2 \in H$;*

(2) *zbiór H jest domknięty względem przejścia do elementu odwrotnego, czyli dla każdego elementu $h \in H$ zachodzi $h^{-1} \in H$.*

Dowód. Niech e będzie elementem neutralnym grupy G.

($\Rightarrow$) Załóżmy, że H jest podgrupą w G. Wtedy na podstawie definicji wnosimy, że $H \neq \varnothing$, $H \subseteq G$ oraz $(H, \cdot)$ jest grupą. Skoro „$\cdot$” jest działaniem algebraicznym w zbiorze H, to $h_1 \cdot h_2 \in H$ dla dowolnych elementów

$h_1, h_2 \in H$. Oprócz tego z definicji grupy wynika, że istnieje $h^{-1} \in H$ dla każdego $h \in H$.

($\Leftarrow$) Niech H będzie podzbiorem niepustym w G. Skoro działanie „$\cdot$" jest algebraiczne na zbiorze G oraz $h_1 \cdot h_2 \in H$ dla dowolnych elementów $h_1, h_2 \in H$, to działanie „$\cdot$" jest algebraiczne na podzbiorze H. Z tego, że „$\cdot$" jest łączne na G, a H jest podzbiorem G wynika, że „$\cdot$" jest też łączne w H. Z warunku (2) wynika, że każdy element $h \in H$ ma odwrotny w H, a więc na podstawie własności (1) otrzymujemy

$$e = h \cdot h^{-1} \in H,$$

czyli H posiada element neutralny e. Biorąc pod uwagę definicję grupy, wnosimy, że $(H, \cdot)$ jest grupą. □

■ W przypadku grupy addytywnej G kryterium podgrupy jest takie

Twierdzenie 2.5.3 (kryterium podgrupy addytywnej)**.** *Niech G będzie grupą (względem dodawania „+"). Zbiór H jest podgrupą w grupie G wtedy i tylko wtedy, gdy są spełnione warunki:*
($0'_1$) $H \neq \varnothing$;
($0'_2$) $H \subseteq G$;
($1'$) $h_1 + h_2 \in H$ *dla dowolnych* $h_1, h_2 \in H$;
($2'$) $-h \in H$ *dla każdego* $h \in H$.

■ Łatwo zauważyć, że dwa warunki (1) oraz (2) możemy przepisać w jednej równoważnej postaci:

$$h_1 \cdot h_2^{-1} \in H \text{ dla dowolnych elementów } h_1, h_2 \in H.$$

■ W przypadku grupy addytywnej warunki ($1'$) oraz ($2'$) możemy przepisać w takiej postaci:

$$h_1 - h_2 \in H \text{ dla dowolnych elementów } h_1, h_2 \in H.$$

■ Jeśli H jest dowolną podgrupą grupy G, a e jest elementem neutralnym w G, to zawsze $e \in H$.

Przykłady 2.5.4.

(**1**) Każda grupa G zawsze ma podgrupy *trywialne*: podgrupę *niewłaściwą* G (czyli grupa G jest podgrupą w grupie G) oraz podgrupę *jednostkową* $\{e\}$ (składającą się dokładnie z jednego elementu neutralnego $e \in G$); analog podgrupy jednostkowej w przypadku addytywnym jest nazywany *podgrupą zerową* i oznaczany przez $\{0\}$.

(**2**) Grupa G z jednością e ma dokładnie jedną podgrupę wtedy i tylko wtedy, gdy $G = \{e\}$ jest grupą jednostkową.

(**3**) Podzbiór

$$\mathbb{Q}[\sqrt{5}] = \{a + b\sqrt{5} \mid a, b \in \mathbb{Q}\}$$

ze zbioru liczb rzeczywistych $\mathbb{R}$ jest niepusty oraz dla dowolnych elementów $a+b\sqrt{5}, c+d\sqrt{5} \in \mathbb{Q}[\sqrt{5}]$ mamy

$$\begin{array}{lcl} (a + b\sqrt{5}) + (c + d\sqrt{5}) & = & (a + c) + (b + d)\sqrt{5} \in \mathbb{Q}[\sqrt{5}], \\ -(a + b\sqrt{5}) & = & (-a) + (-b)\sqrt{5} \in \mathbb{Q}[\sqrt{5}], \end{array}$$

czyli $\mathbb{Q}[\sqrt{5}]$ jest podgrupą addytywnej grupy liczb rzeczywistych $\mathbb{R}$.

(**4**) W przypadku ogólnym nie każdy podzbiór niepusty grupy jest jej podgrupą. Na przykład podzbiór

$$B = \{a\sqrt{5} \mid a \in \mathbb{Q} \setminus \{0\}\}$$

nie jest podgrupą w $\mathbb{Q}(\sqrt{5})$, bo $a\sqrt{5} \in B$, $(-a)\sqrt{5} \in B$, lecz $a\sqrt{5} + (-a)\sqrt{5} = 0 \notin B$.

(**5**) Zbiór

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det A = 1\}$$

jest podgrupą w grupie $GL_n(\mathbb{R})$. Możemy przekonać się o tym, sprawdzając spełnienie warunków z kryterium podgrupy. Jeśli I_n jest macierzą jednostkową stopnia n, to $\det I_n = 1$, a więc $I_n \in SL_n(\mathbb{R})$ oraz $SL_n(\mathbb{R}) \neq \varnothing$. Oprócz tego $SL_n(\mathbb{R})$ jest podzbiorem w $GL_n(\mathbb{R})$. Niech $A, B \in SL_n(\mathbb{R})$. Wtedy $\det A = 1$, $\det B = 1$, a zatem na podstawie wzoru Cauchy'ego-Bineta[(11)]

$$\det(AB) = \det A \cdot \det B = 1 \cdot 1 = 1 \text{ oraz } \det(A^{-1}) = \frac{1}{\det A} = \frac{1}{1} = 1.$$

To znaczy, że $SL_n(\mathbb{R})$ jest podgrupą w ogólnej grupie liniowej $GL_n(\mathbb{R})$ stopnia n nad ciałem liczb rzeczywistych $\mathbb{R}$. Grupa $SL_n(\mathbb{R})$ jest nazywana *szczególną grupą liniową* stopnia n nad ciałem liczb rzeczywistych $\mathbb{R}$.

(**6**) Niech

$$\mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \sigma \text{ jest permutacją parzystą}\}.$$

Wtedy $e \in \mathbb{A}_n$, a więc $\mathbb{A}_n \neq \varnothing$. Mamy $\mathbb{A}_n \subseteq \mathbb{S}_n$. Jeśli σ, τ są dowolnymi elementami z $\mathbb{A}_n$, to

$$\sigma = \sigma_1 \cdots \sigma_{2n}, \ \tau = \tau_1 \cdots \tau_{2m},$$

gdzie σ_i, τ_j są transpozycjami ($i = 1, \ldots, 2n$; $j = 1, \ldots, 2m$). Zatem

$$\sigma\tau = \sigma_1 \cdots \sigma_{2n}\tau_1 \cdots \tau_{2m}$$

jest iloczynem $2n + 2m$ transpozycji, a $\sigma_i^{-1} = \sigma_i$ oraz

$$\sigma^{-1} = (\sigma_1 \cdots \sigma_{2n})^{-1} = \sigma_{2n}^{-1} \cdots \sigma_1^{-1} = \sigma_{2n} \cdots \sigma_1$$

(11) Jacques Philippe Marie Binet (1785–1856)

jest iloczynem $2n$ transpozycji. Wnosimy, że $\sigma\tau, \sigma^{-1} \in \mathbb{A}_n$ i na podstawie kryterium $\mathbb{A}_n$ jest podgrupą w $\mathbb{S}_n$. Grupa $\mathbb{A}_n$ jest nazywana *grupą alternującą* stopnia n.

(**7**) Niech

$$U(n) = \{A \in M_n(\mathbb{C}) \mid A^*A = I_n = AA^*\},$$

$$SU(n) = \{A \in M_n(\mathbb{C}) \mid \det A = 1 \text{ oraz } \ A^*A = I_n = AA^*\},$$

gdzie $A^* = \overline{A}^T = [\overline{a}_{ji}]$ jest macierzą sprzężoną i transponowaną do macierzy $A = [a_{ij}] \in M_n(\mathbb{C})$ oraz $I_n \in M_n(\mathbb{C})$ jest macierzą jednostkową. Czytelnik łatwo przekona się, że mamy łańcuchy podgrup

$$SU(n) \leqslant U(n) \leqslant GL_n(\mathbb{C}),\ SU(n) \leqslant SL_n(\mathbb{C}) \leqslant GL_n(\mathbb{C}).$$

Oczywiście, że $SU(1) = \{1\}$, $U(1) = \{z \in \mathbb{C} \mid |z| = 1\}$ oraz

$$SU(2) = \left\{ \begin{bmatrix} a & -\overline{b} \\ b & \overline{a} \end{bmatrix} \mid a, b \in \mathbb{C} \text{ oraz } |a|^2 + |b|^2 = 1 \right\}.$$

Jeśli A jest ustaloną macierzą z $SU(2)$, to

$$SU(2) = \{\sigma A \mid \sigma \in \mathbb{C} \text{ oraz } |\sigma| = 1\}.$$

Grupa $U(n)$ jest nazywana *grupą unitarną* macierzy stopnia n, a grupa $SU(n)$ – *grupą unitarną szczególną* macierzy stopnia n.

Ponadto zachodzi takie

Twierdzenie 2.5.5. *Niech $X \in M_n(\mathbb{C})$ będzie macierzą nieosobliwą. Wtedy zbiór*

$$H_{\mathbb{C}} = \{A \in M_n(\mathbb{C}) \mid \overline{A}^T X A = X\}$$

tworzy grupę, gdzie $\overline{A} = [\overline{a}_{ij}]$ jest macierzą sprzężoną z macierzą $A = [a_{ij}] \in M_n(\mathbb{C})$.

□

(**8**) Jeśli $p, q \in \mathbb{N}$ oraz $p + q \geqslant 1$, to

$$U(p,q) = \left\{ A \in M_{p+q}(\mathbb{C}) \mid \overline{A}^T \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} A = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix} \right\}$$

jest grupą (przekonać się samodzielnie).

(**9**) Mamy takie grupy liniowe

$$O(n) = U(n) \bigcap M_n(\mathbb{R}) \text{ oraz } \ SO(n) = SU(n) \bigcap M_n(\mathbb{R}).$$

Wtedy

$$SO_n(\mathbb{R}) = O_n(\mathbb{R}) \bigcap SL_n(\mathbb{R}) \text{ oraz } SU(n) = U(n) \bigcap SL_n(\mathbb{C}).$$

* * *

■ **Działania na podgrupach.** Niech $(G, \cdot)$ będzie grupą.

1) Jeśli H oraz S są podgrupami grupy G, to podzbiór

$$H \bigcup S = \{g \in G \mid g \in H \text{ lub } g \in S\}$$

jest nazywany ich *sumą mnogościową*. W przypadku ogólnym suma mnogościowa dwóch podgrup grupy G niekoniecznie będzie jej podgrupą.

Na przykład $3\mathbb{Z}$ oraz $5\mathbb{Z}$ są podgrupami addytywnej grupy liczb całkowitych $\mathbb{Z}$, lecz $3 + 5 \notin 3\mathbb{Z} \cup 5\mathbb{Z}$, a zatem w wyniku kryterium podgrupy wnosimy, że $3\mathbb{Z} \cup 5\mathbb{Z}$ nie jest podgrupą w $\mathbb{Z}$.

Lecz, jak wynika z następnego lematu, jeśli $H \subseteq S$ lub $H \supseteq S$, to $H \cup S \leqslant G$.

Lemat 2.5.6. *Jeśli H_i są takimi podgrupami grupy G $(i \in \mathbb{N}^*)$, że*

$$H_1 \leqslant H_2 \leqslant \cdots \leqslant H_n \leqslant \cdots ,$$

to ich suma mnogościowa

$$H_0 = \bigcup_{i=1}^{\infty} H_i$$

jest podgrupą w G.

Dowód. Z własności $H_i \neq \varnothing$ dla każdego $i \in \mathbb{N}$ wynika, że $H_0 \neq \varnothing$. Niech x, y będą dowolnymi elementami z H_0. Wtedy znajdą się takie $i, j \in \mathbb{N}$, że $x \in H_i$ oraz $y \in H_j$, przy czym jedna z podgrup H_i lub H_j zawiera się w innej. Załóżmy, na przykład, że $H_i \subseteq H_j$. Skoro H_i jest grupą, to

$$x^{-1} \in H_i \subseteq H_0.$$

Wtedy także $x, y \in H_j$, a więc

$$x \cdot y \in H_j \subseteq H_0.$$

Na podstawie kryterium podgrupy H_0 jest podgrupą grupy G. □

2) *Przecięciem* (lub *przekrojem*) dwóch podgrup H i S z grupy G jest zbiór

$$H \bigcap S = \{g \in G \mid g \in H \text{ oraz } g \in S\}.$$

Lemat 2.5.7. *Jeśli H oraz S są podgrupami grupy G, to ich przecięcie $H \cap S$ jest podgrupą w G.*

Dowód. Niech e będzie jednością grupy G. Skoro $e \in H$ i $e \in S$, to $e \in H \cap S$ oraz $H \cap S \neq \varnothing$. Jeśli x i y są dowolnymi elementami z $H \cap S$, to $x, y \in H$ oraz $x, y \in S$, a stąd na podstawie kryterium podgrupy $x \cdot y, x^{-1} \in H$ oraz $x \cdot y, x^{-1} \in S$. Jako wniosek

$$x \cdot y, x^{-1} \in H \bigcap S.$$

Na mocy twierdzenia 2.5.2 otrzymujemy, że $H \cap S$ jest podgrupą w G. □

3) Jeśli $H, S \leqslant G$, to ich *iloczyn* (algebraiczny) definiujemy w taki sposób:

$$H \cdot S = HS = \{h \cdot s \mid h \in H \text{ oraz } \ s \in S\}.$$

Lemat 2.5.8. *Jeśli H i S są podgrupami grupy G, to ich iloczyn HS jest podgrupą w G wtedy i tylko wtedy, gdy $HS = SH$.*

Dowód. ($\Rightarrow$) Niech $H, S \leqslant G$. Wtedy $H \leqslant HS$ oraz $S \leqslant HS$, a zatem na podstawie kryterium podgrupy $s \cdot h \in HS$ dla dowolnych elementów $h \in H$ i $s \in S$, czyli $SH \subseteq HS$. Także dla dowolnych elementów $h \in H$ i $s \in S$ z warunku $(h \cdot s)^{-1} \in HS$ wynika, że

$$(h \cdot s)^{-1} = h_1 \cdot s_1$$

dla pewnych elementów $h_1 \in H$ i $s_1 \in S$, skąd $h \cdot s = s_1^{-1} \cdot h_1^{-1} \in SH$ lub, co jest tożsame w języku zbiorów, $HS \subseteq SH$. Zatem

$$HS = SH.$$

($\Leftarrow$) Teraz załóżmy, że $HS = SH$. Jeśli $h_i \in H$ i $s_i \in S$ $(i = 1, 2)$, to

$$h_1 \cdot s_1 \cdot (h_2 \cdot s_2)^{-1} = h_1 \cdot s_1 \cdot s_2^{-1} \cdot h_2^{-1}.$$

Ponieważ $s_1 \cdot s_2^{-1} \cdot h_2^{-1} \in HS$, to

$$s_1 \cdot s_2^{-1} \cdot h_2^{-1} = h_3 \cdot s_3$$

dla pewnych elementów $h_3 \in H$, $s_3 \in S$ i wtedy

$$h_1 \cdot s_1 \cdot (h_2 \cdot s_2)^{-1} = (h_1 \cdot h_3) \cdot s_3 \in HS.$$

Na podstawie kryterium podgrupy wnioskujemy, że $HS \leqslant G$. □

Przykłady 2.5.9.

(**1**) Mamy $2\mathbb{Z} \cap 3\mathbb{Z} = 6\mathbb{Z}$ w grupie addytywnej liczb całkowitych $\mathbb{Z}$.

(**2**) Iloczyn HS dwóch podgrup $H = \{e, (12)\}$ oraz $S = \{e, (23)\}$ z grupy $\mathbb{S}_3$ nie jest jej podgrupą, bo

$$HS = \{ee, e(23), (12)e, (12)(23)\} = \{e, (23), (12), (123)\},$$
$$SH = \{ee, e(12), (23)e, (23)(12)\} = \{e, (12), (23), (132)\},$$

a więc $HS \neq SH$.

Ćwiczenia 2.5.10.

(**1**) Sprawdzić, czy HK jest podgrupą w grupie $\mathbb{S}_3$, jeśli:
(a) $H = \{e, (12)\}$ oraz $K = \{e, (13)\}$;
(b) $H = \{e, (23)\}$ oraz $K = \{e, (123), (132)\}$.
(**2**) Sprawdzić, czy H jest podgrupą w grupie G, jeśli:
(a) $H = \{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{Q}\}$ oraz $G = GL_2(\mathbb{Q})$;
(b) $H = \{\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix} \mid a \in \mathbb{R}\}$ oraz $G = \{\begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \mid a, b, c \in \mathbb{R}, ac \neq 0\}$;
(c) $G = GL_3(\mathbb{R})$ oraz $H = \{\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \mid a, b, c \in \mathbb{R}\}$;
(d) $G = M_2(\mathbb{R})$ oraz $H = \{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(e) $G = L(\mathbb{C})$, gdzie $L(\mathbb{C}) = \{f_{a,b} : \mathbb{C} \to \mathbb{C} \mid f_{a,b}(z) = az + b,$ gdzie $a, b, z \in \mathbb{C}$ oraz $a \neq 0\}$, oraz $H = L_1(\mathbb{C}) = \{f_{1,b} \in L(\mathbb{C}) \mid b \in \mathbb{C}\}$;
(f) $G = L(\mathbb{C})$ oraz $H = \{f_{a,b} \in L(\mathbb{C}) \mid a \in \mathbb{C}^*, b \in \mathbb{R}\}$;
(g) $G = \mathbb{Q}$ oraz $H = \mathbb{Z}[\sqrt{7}]$, gdzie $\mathbb{Z}[\sqrt{7}] = \{a + b\sqrt{7} \mid a, b \in \mathbb{Z}\}$;
(h) $G = \mathbb{C}^*$ oraz $H = \{z \in \mathbb{C} \mid |z| = 1\}$.
(**3**) Znaleźć wszystkie podgrupy w grupie:
(a) $\mathbb{S}_3$;
(b) V_4;
(c) $SL_2(\mathbb{Z}_2)$;
(d) $\mathbb{Z}_6$.
(**4**) Zbudować tabelkę Cayleya dla mnożenia w grupie:
(a) $GL_2(\mathbb{Z}_2)$;
(b) $GL_2(\mathbb{Z}_3)$.
(**5**) Niech H, K będą podgrupami w grupie G. Znaleźć $H \cup K$, $H \cap K$ oraz $H + K$ i sprawdzić, czy $H \cup K$ i $H + K$ są podgrupami w G, jeśli:
(a) $G = \mathbb{Z}$ oraz $H = 2\mathbb{Z}$ oraz $K = 14\mathbb{Z}$;
(b) $G = \mathbb{Z}$ oraz $H = 3\mathbb{Z}$ oraz $K = 5\mathbb{Z}$;
(c) $G = \mathbb{Z}$ oraz $H = 30\mathbb{Z}$ oraz $K = 66\mathbb{Z}$.
(**6**) Znaleźć wszystkie podgrupy w grupie:
(a) $\mathbb{Z}_5$;

(b) $\mathbb{Z}_4$;
(c) $\mathbb{Z}_6$;
(d) $\mathbb{Z}_{30}$;
(e) $\mathbb{Z}$.
(**7**) Udowodnić, że w każdej grupie G jednakowe rzędy mają takie elementy:
(a) x oraz yxy^{-1};
(b) xy oraz yx, gdzie $x, y \in G$.
(**8**) Znaleźć rząd elementu a i rząd podgrupy H w grupie G, jeśli:
(a) $a = (12)(34)$, $G = \mathbb{S}_4$ oraz $H = V_4$;
(b) $a = \begin{bmatrix} 0 & i \\ 1 & 0 \end{bmatrix}$, $G = GL_2(\mathbb{C})$ oraz $H = SL_2(\mathbb{C})$;
(c) $a = \frac{1}{2} - \frac{\sqrt{3}}{2}i$, $G = \mathbb{C}^*$ oraz $H = \{z \in \mathbb{C}^* \mid z^4 = 1\}$;
(d) $a = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, $G = GL_2(\mathbb{Z}_7)$ oraz $H = SL_2(\mathbb{Z}_7)$;
(e) $a = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$, $G = GL_2(\mathbb{C})$ oraz $H = \{\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \mid b \in \mathbb{Q}\}$;
(f) $a = \frac{4}{5} - \frac{3}{5}i$, $G = \mathbb{C}^*$ oraz $H = \mathbb{R}^*$;
(g) $a = f_{5,0}$, $G = \{f_{u,v} : \mathbb{Z}_{11} \to \mathbb{Z}_{11} \mid u, v \in \mathbb{Z}_{11}, u \neq 0\}$ oraz $H = \{f_{1,v} \in G \mid v \in \mathbb{Z}_{11}\}$.
(**9**) Udowodnić, że jeśli $|x| = |y| = 2$ oraz x, y są różnymi elementami grupy G, to $|xy| = 2$ wtedy i tylko wtedy, gdy $xy = yx$.

Uwagi. Terminu „podgrupa" po raz pierwszy użyto w pracy E. Galoisa.

2.6. Pierścienie i ich własności elementarne

Zacznijmy od definicji podstawowych.

■ Zbiór R jest nazywany *pierścieniem* (względem dodawania „+" i mnożenia „·"), jeśli:

0_1) R jest zbiorem niepustym;

0_2) na zbiorze R są określone dwa działania algebraiczne:

- (dodawanie „+")

$$R \times R \ni (a, b) \mapsto a + b \in R$$

oraz

- (mnożenie „·")

$$R \times R \ni (a, b) \mapsto a \cdot b \in R,$$

spełniające takie warunki:

1) (łączność dodawania „+")

$$\forall_{a,b,c \in R} : (a + b) + c = a + (b + c);$$

2) (istnienie elementu neutralnego względem dodawania „+")

$$\exists_{0 \in R} \forall_{a \in R} : a + 0 = a = 0 + a$$

(element 0, który jest neutralny względem „+" w R, jest nazywany *zerem* (lub *elementem zerowym*) pierścienia R);

3) (istnienie elementu przeciwnego)

$$\forall_{a \in R} \exists_{(-a) \in R} : a + (-a) = 0 = (-a) + a$$

(element $(-a)$ jest nazywany *przeciwnym* do elementu $a \in R$);

4) (przemienność dodawania „+")

$$\forall_{a,b \in R} : a + b = b + a;$$

5) (łączność mnożenia „·")

$$\forall_{a,b,c \in R} : (a \cdot b) \cdot c = a \cdot (b \cdot c);$$

6) (rozdzielność dwustronna mnożenia „$\cdot$" względem dodawania „$+$")

$$\forall_{a,b,c\in R} : a \cdot (b + c) = (a \cdot b) + (a \cdot c) \text{ oraz } (a + b) \cdot c = (a \cdot c) + (b \cdot c).$$

■ Zbiór R jest nazywany *pierścieniem przemiennym* (względem dodawania „$+$" i mnożenia „$\cdot$"), jeśli są spełnione warunki $0_1),\ 0_2),\ 1)$–$6)$ oraz

7) (przemienność mnożenia „$\cdot$")

$$\forall_{a,b\in R} : a \cdot b = b \cdot a.$$

■ Zbiór R jest nazywany *pierścieniem z jednością* 1 (albo *pierścieniem z elementem jednostkowym*, albo *pierścieniem unitarnym*) względem dodawania „$+$" i mnożenia „$\cdot$", jeśli są spełnione warunki $0_1),\ 0_2),\ 1)$–$6)$ oraz

8) (istnienie elementu neutralnego względem mnożenia „$\cdot$")

$$\exists_{1\in R}\forall_{a\in R} : a \cdot 1 = a = 1 \cdot a.$$

■ Pierścień R jest nazywany *pierścieniem komutatywnym (przemiennym) z jednością* 1 (względem dodawania „$+$" i mnożenia „$\cdot$"), jeśli zachodzą warunki $0_1),\ 0_2),\ 1)$–$8)$.

■ Jeśli zbiór R jest pierścieniem względem dodawania „$+$" i mnożenia „$\cdot$", to będziemy krótko mówić, że trójka $(R, +, \cdot)$ jest pierścieniem.

■ W pierścieniu $(R, +, \cdot)$ reguła

$$a - b = a + (-b),$$

gdzie $a, b \in R$, wyznacza nowe działanie algebraiczne „$-$", które jest nazywane *odejmowaniem*.

■ Jeśli $(A, +)$ jest dowolną grupą abelową (z elementem neutralnym 0), to biorąc

$$a \cdot b = 0$$

dla wszystkich $a, b \in A$, otrzymujemy pierścień przemienny $(A, +, \cdot)$ bez jedności (nazywany *zero-pierścieniem* na grupie A).

■ W pierścieniu R iloczyn $a \cdot b$ będziemy krótko oznaczać przez ab.

■ Jeśli $(R, +, \cdot)$ jest pierścieniem, to $(R, +)$ jest grupą abelową (nazywaną grupą *addytywną* pierścienia R i oznaczaną przez R^+).

■ Jeśli $(R, +, \cdot)$ jest pierścieniem, to $(R, \cdot)$ jest półgrupą (nazywaną półgrupą *multiplikatywną* pierścienia R).

Przykłady 2.6.1.

(**1**) Każdy ze zbiorów liczbowych $\mathbb{Z}$, $\mathbb{Q}$ oraz $\mathbb{R}$ tworzy pierścień przemienny z jednością 1 względem (zwykłych) działań dodawania „+" i mnożenia „·" liczb.

(**2**) Zbiór liczb całkowitych parzystych

$$2\mathbb{Z} = \{2a \mid a \in \mathbb{Z}\}$$

jest pierścieniem przemiennym względem dodawania „+" oraz mnożenia „·" liczb, lecz bez elementu jednostkowego.

(**3**) Zbiór $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ klas reszt modulo n, gdzie

$$\overline{k} = k + n\mathbb{Z} = \{k + na \mid a \in \mathbb{Z}\}$$

oraz $0 \leqslant k \leqslant n-1$, jest pierścieniem przemiennym z jednością $\overline{1}$ względem dwóch działań „+" i „·", określonych według takich reguł:

$$\overline{a} + \overline{b} = \overline{a+b} \text{ oraz } \overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

Biorąc pod uwagę, że $\overline{a} = \overline{r}$, gdzie r jest resztą z dzielenia a przez n, łatwo sprawdzić, że działania „+" oraz „·" są algebraiczne na zbiorze $\mathbb{Z}_n$. Pierścień

$$(\mathbb{Z}_n, +, \cdot)$$

jest nazywany pierścieniem *klas reszt modulo* n ($n \in \mathbb{N}^*$).

Jako przykład zbudujmy tabelki Cayleya dla dodawania „+" oraz mnożenia „·" w pierścieniu klas reszt $\mathbb{Z}_4 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}$ modulo 4:

+	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

·	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

(**4**) Pierścień macierzy kwadratowych stopnia n

$$M_n(\mathbb{R}) = \{[a_{ij}] \mid a_{ij} \in \mathbb{R}\}$$

jest pierścieniem z jednością I_n (względem dodawania „+" i mnożenia „·" macierzy).

(**5**) Jeśli w pierścieniu R zachodzi równość $1 = 0$, to dla każdego elementu $a \in R$ mamy, że

$$a = a \cdot 1 = a \cdot 0 = 0.$$

Zatem $R = \{0\}$ jest pierścieniem *zerowym*.

* * *

■ **Symbol sumy i jego własności.** Niech $(G, +)$ będzie grupą addytywną oraz a_i, a_{ij}, b_i będą jej elementami $(i = 1, \ldots, n; j = 1, \ldots, m)$. Wtedy sumę

$$a_1 + a_2 + \cdots + a_n$$

krótko oznaczamy przez

$$\sum_{i=1}^{n} a_i$$

(i mówimy „suma elementów a_i po i od 1 do n"); przy tym i jest nazywane *indeksem* (lub *wskaźnikiem*) *sumowania*. Wśród własności symbolu sumy są takie.

1)
$$\sum_{i=1}^{n+1} a_i = \sum_{i=1}^{n} a_i + a_{n+1}.$$

2) Wynik sumowania nie zależy od tego, jakim symbolem oznaczamy indeks sumowania:
$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j.$$

3) (addytywność) Jeśli grupa addytywna G jest abelowa, to

$$\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i.$$

4) Dla każdego indeksu s $(1 \leqslant s \leqslant n)$ zachodzi równość

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{s} a_i + \sum_{i=s+1}^{n} a_i.$$

5) W sumie podwójnej wynik sumowania nie zależy od kolejności sumowania:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}.$$

6) Jeśli $(R, +, \cdot)$ jest pierścieniem przemiennym oraz $\lambda, a_i \in R$ $(i = 1, \ldots, n)$, to zachodzi uogólniona rozdzielność mnożenia względem dodawania:

$$\sum_{i=1}^{n} \lambda a_i = \lambda \sum_{i=1}^{n} a_i,$$

i, w szczególności,

$$\sum_{i=1}^{n} \lambda = n\lambda.$$

* * *

■ **Symbol iloczynu i jego własności.** Niech $(G, \cdot)$ będzie grupą multiplikatywną oraz a_i, a_{ij}, b_i będą jej elementami $(i = 1, \ldots, n; j = 1, \ldots, m)$. Wtedy iloczyn

$$a_1 \cdot a_2 \cdots a_n$$

krótko oznaczamy przez

$$\prod_{i=1}^{n} a_i$$

(i mówimy „iloczyn elementów a_i po i od 1 do n"); przy tym i jest nazywane *indeksem* (lub *wskaźnikiem*) *mnożenia*. Własności tego symbolu są następujące.

1)

$$\prod_{i=1}^{n+1} a_i = (\prod_{i=1}^{n} a_i) \cdot a_{n+1}.$$

2) Wynik iloczynu nie zależy od tego, jaką literą oznaczamy indeks mnożenia:
$$\prod_{i=1}^{n} a_i = \prod_{j=1}^{n} a_j.$$

3) (multiplikatywność) Jeśli grupa multiplikatywna G jest abelowa, to
$$(\prod_{i=1}^{n} a_i) \cdot (\prod_{i=1}^{n} b_i) = \prod_{i=1}^{n} a_i b_i.$$

4) Dla grupy abelowej G w iloczynie podwójnym wynik mnożenia nie zależy od kolejności mnożenia:
$$\prod_{i=1}^{n}\prod_{j=1}^{m} a_{ij} = \prod_{j=1}^{m}\prod_{i=1}^{n} a_{ij}.$$

5) Dla każdego indeksu s $(1 \leqslant s \leqslant n)$ zachodzi równość
$$\prod_{i=1}^{n} a_i = (\prod_{i=1}^{s} a_i)(\prod_{i=s+1}^{n} a_i);$$

6) Jeśli $(R, +, \cdot)$ jest pierścieniem przemiennym oraz $\lambda, a_i \in R$ $(i = 1, \ldots, n)$, to
$$\prod_{i=1}^{n} \lambda a_i = \lambda^n \prod_{i=1}^{n} a_i,$$
i, w szczególności,
$$\prod_{i=1}^{n} \lambda = \lambda^n.$$

* * *

■ **Własności elementarne pierścieni.** Pierścień może posiadać pewne elementy szczególne. Niech $(R, +, \cdot)$ będzie pierścieniem. Element $a \in R$ nazywamy *dzielnikiem lewostronnym zera* (odpowiednio *dzielnikiem prawostronnym zera*), jeśli:
1) $a \neq 0$ oraz

2) istnieje taki element $b \in R$, że $b \neq 0$ oraz $ab = 0$ (odpowiednio znajdzie się takie $c \in R$, że $c \neq 0$ oraz $ca = 0$).

Element pierścienia, który jest jednocześnie lewostronnym i prawostronnym dzielnikiem zera, jest nazywany (*obustronnym*) *dzielnikiem zera.* Z definicji wynika, że zero pierścienia nie jest dzielnikiem zera. Pierścień, niezawierający dzielników zera (lewostronnych oraz prawostronnych), jest nazywany *dziedziną całkowitości* (lub *pierścieniem bez dzielników zera*).

Przykłady 2.6.2.

(**1**) Jeśli $m \in \mathbb{N} \setminus \{0, 1\}$ i m nie jest liczbą pierwszą (a więc $m = ns$ dla pewnych różnych od 1 liczb naturalnych n, s), to w pierścieniu $\mathbb{Z}_m$ mamy elementy niezerowe $\overline{n} \neq \overline{0}$ oraz $\overline{s} \neq \overline{0}$ takie, że ich iloczyn

$$\overline{n} \cdot \overline{s} = \overline{m} = \overline{0}$$

jest zerowy (czyli $\overline{n}, \overline{s}$ są dzielnikami zera w $\mathbb{Z}_m$). Na przykład $\overline{2} \cdot \overline{2} = \overline{0}$ oraz $\overline{2} \neq \overline{0}$ w pierścieniu $\mathbb{Z}_4$.

(**2**) Niech

$$C_{[-1,1]} = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ jest funkcją ciągłą na przedziale } [-1, 1]\}.$$

Wtedy $(C_{[-1,1]}, +, \cdot)$ jest pierścieniem przemiennym z zerem (=funkcją zerową)

$$\mathcal{O} : \mathbb{R} \ni x \mapsto 0 \in \mathbb{R},$$

oraz jednością (=funkcją jednostkową)

$$\mathbf{1} : \mathbb{R} \ni x \mapsto 1 \in \mathbb{R},$$

względem takich dwóch działań:

- (popunktowe dodawanie)
$$(f + g)(x) = f(x) + g(x),$$
- (popunktowe mnożenie)
$$(f \cdot g)(x) = f(x) \cdot g(x),$$

gdzie $f, g \in C_{[-1,1]}$ oraz $x \in \mathbb{R}$. Jak wiadomo, suma i iloczyn funkcji ciągłych na przedziale $[-1, 1]$ są ciągłe na $[-1, 1]$. Jeśli teraz

$$a(x) = \begin{cases} 0, & \text{gdy } x \leqslant 0, \\ x, & \text{gdy } x > 0 \end{cases} \quad \text{oraz} \quad b(x) = \begin{cases} x, & \text{gdy } x \leqslant 0, \\ 0, & \text{gdy } x > 0, \end{cases}$$

to

$$a(x) \cdot b(x) = 0$$

w każdym punkcie $x \in \mathbb{R}$. Zatem $a(x), b(x)$ są (obustronnymi) dzielnikami zera w pierścieniu $C_{[-1,1]}$ funkcji ciągłych na przedziale $[-1, 1]$.

■ Niech $(R, +, \cdot)$ będzie pierścieniem oraz $n \in \mathbb{N}$. Element niezerowy $a \in R$ jest nazywany elementem *nilpotentnym indeksu* (nilpotentności) n $(n \geqslant 1)$, jeśli

$$a^n = 0 \text{ oraz } a^{n-1} \neq 0.$$

Element zerowy 0 jest nilpotentnym indeksu 1.

■ Niech $(R, +, \cdot)$ będzie pierścieniem z jednością 1. Element $a \in R$ jest nazywany:

- *odwracalnym lewostronnie* (lub *jednością lewostronną*) w pierścieniu R, jeśli
$$a \cdot b = 1$$
dla pewnego $b \in R$;
- *odwracalnym prawostronnie* (lub *jednością prawostronną*) w pierścieniu R, jeśli
$$c \cdot a = 1$$
dla pewnego $c \in R$;
- *odwracalnym* (lub *jednością*) w pierścieniu R, jeśli
$$a \cdot b = 1 = b \cdot a$$
dla pewnego elementu $b \in R$.

Przykłady 2.6.3.

(**1**) W pierścieniu $\mathbb{Z}_6$ mamy
$$\overline{5} \cdot \overline{5} = \overline{25} = \overline{1},$$
czyli element $\overline{5}$ jest odwracalny w $\mathbb{Z}_6$ oraz $\overline{5}^{-1} = \overline{5}$.

(**2**) W pierścieniu klas reszt $\mathbb{Z}_8$ mamy

$$\overline{2}^3 = \overline{8} = \overline{0} \text{ oraz } \overline{2}^4 = \overline{4} \neq \overline{0},$$

a zatem $\overline{2}$ jest elementem nilpotentnym indeksu 3.

(**3**) Zbiór $M_n(\mathbb{Q})$ macierzy kwadratowych stopnia n o współczynnikach wymiernych jest pierścieniem (względem działań dodawania i mnożenia macierzy) z jednością

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Niech $n = 3$. Macierz

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \in M_3(\mathbb{Q})$$

jest niezerowa. Skoro

$$A^2 = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \mathcal{O}$$

oraz

$$A^3 = A^2 \cdot A \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathcal{O}$$

jest macierzą zerową, to macierz A jest elementem nilpotentnym (indeksu 3) w pierścieniu $M_3(\mathbb{Q})$.

■ Niech

$$U(R) = \{r \in R \mid r \text{ jest elementem odwracalnym w } R\},$$

gdzie R jest pierścieniem unitarnym.

Lemat 2.6.4. *Jeśli $(R, +, \cdot)$ jest pierścieniem z jednością 1, to $U(R)$ jest grupą z elementem neutralnym 1.*

Dowód. W rzeczy samej, $1 \in U(R)$. Jeśli teraz $a, b, c, d \in R$ spełniają równości

$$ab = 1 = ba \text{ i } cd = 1 = dc,$$

to

$$\begin{array}{rcl} (ac)(db) & = & ((ac)d)b = (a(cd))b = (a \cdot 1)b = ab = 1, \\ (db)(ac) & = & ((db)a)c = (d(ba))c = (d \cdot 1)c = dc = 1, \end{array}$$

czyli mnożenie „$\cdot$" jest działaniem algebraicznym w $U(R)$. To działanie jest łączne w R, a więc i w $U(R)$; i ma element neutralny 1. Na podstawie definicji $U(R)$ posiada element a^{-1} odwrotny do każdego $a \in U(R)$. Zatem $(U(A), \cdot)$ jest grupą. □

■ Grupa $U(R)$ jest nazywana *grupą multiplikatywną* (lub *grupą jedności*) pierścienia unitarnego R.

Twierdzenie 2.6.5. *Niech $(R, +, \cdot)$ będzie pierścieniem oraz $a, b, c \in R$. Wtedy zachodzą własności:*

(1) $a \cdot 0 = 0 = 0 \cdot a$;
(2) $(-a) \cdot b = a \cdot (-b) = -(ab)$;
(3) $-(-a) = a$;

(4) $(-a)(-b) = ab$;
(5) $(a-b)c = ac - bc$ *oraz* $a(b-c) = ab - ac$;
(6) *jeśli* R *jest pierścieniem przemiennym, to*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i},$$

gdzie współczynnik $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ *dla każdego* $n \in \mathbb{N}^*$;
(7) (prawo skracania w pierścieniu) *jeśli* R *jest dziedziną całkowitości i* $a \neq 0$, *to z* $ab = ac$ *(odpowiednio* $ba = ca$*) wynika, że* $b = c$;
(8) *jeśli w pierścieniu* R *z warunków* $a \neq 0$ *oraz* $ab = ac$ *(odpowiednio* $ba = ca$*) wynika, że* $b = c$, *to* R *jest dziedziną całkowitości.*

Dowód. (1) Skoro $0 = 0 + 0$, to

$$0 + a \cdot 0 = a \cdot 0 = a \cdot (0+0) = a \cdot 0 + a \cdot 0,$$

a stąd na podstawie prawa skracania w grupie addytywnej $(R, +)$ otrzymujemy $a \cdot 0 = 0$. Podobnie dostajemy, że $0 \cdot a = 0$.

(2) Z równości

$$(-a)b + ab = ((-a) + a)b = 0 \cdot b = 0 = (-(ab)) + ab$$

wynika, że $(-a)b = -(ab)$. W podobny sposób otrzymujemy

$$a(-b) = -(ab).$$

(3) Skoro

$$a + (-a) = 0 = -(-a) + (-a),$$

to na podstawie prawa skracania w grupie $(R, +)$ wnosimy, że

$$-(-a) = a.$$

(4) Rzeczywiście, korzystając z własności (2) i (3), mamy

$$(-a)(-b) = -(a(-b)) = -(-(ab)) = ab.$$

(5) W rzeczy samej, z (2) oraz rozdzielności mnożenia względem dodawania otrzymujemy

$$(a-b)c = (a+(-b))c = ac + (-b)c = ac - bc.$$

Prawostronna własność ma podobny dowód.

(6) Proponujemy Czytelnikowi samodzielnie znaleźć dowód.

(7) Jeśli R jest pierścieniem bez dzielników zera, $ab = ac$ oraz $a \neq 0$, to $a(b-c) = 0$, a więc $b - c = 0$ i $b = c$.

(8) Z równości $ab = 0$ (odpowiednio $ba = 0$) dla pewnych elementów $a, b \in R$ oraz $ab = a0$ (odpowiednio $ba = 0$) wynika, że $b = 0$, a zatem R jest dziedziną całkowitości. □

■ Niech $(R, +, \cdot)$ będzie pierścieniem (z jednością 1 i zerem 0). Będziemy mówić, że pierścień R ma *charakterystykę* n (i zapisywać $\operatorname{char} R = n$), gdzie $n \in \mathbb{N}^*$, jeśli są spełnione dwa warunki:

1)
$$n \cdot 1 = \underbrace{1 + \cdots + 1}_{n \text{ składników}} = 0;$$

2) liczba n jest najmniejsza wśród liczb naturalnych niezerowych z własnością 1).

Jeżeli takiej liczby naturalnej niezerowej n nie istnieje, to będziemy mówić, że R ma *charakterystykę* 0 (i zapisywać $\operatorname{char} R = 0$).

Przykłady 2.6.6.

(**1**) Pierścienie $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ oraz $\mathbb{C}$ mają charakterystykę zerową, bo

$$n \cdot 1 \neq 0$$

dla każdej liczby naturalnej niezerowej n.

(**2**) Pierścień klas reszt $\mathbb{Z}_n$ ma charakterystykę n, gdzie $n \in \mathbb{N}^*$. Na przykład $\operatorname{char} \mathbb{Z}_{12} = 12$.

Ćwiczenia 2.6.7.

(**1**) Niech $n \geqslant 2$. Sprawdzić, czy A jest pierścieniem względem dodawania i mnożenia macierzy, jeśli:

(*a*) $A = \{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \mid a_{ij} \in \mathbb{C} \ (i, j = 1, \ldots, n) \};$

(*b*) $A = \{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \mid a \in \mathbb{R}\}$;
(*c*) $A = \{\begin{bmatrix} a & b \\ \delta b & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$, gdzie δ jest ustaloną liczbą wymierną;
(*d*) $A = \{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(*e*) $A = \{\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \mid z, w \in \mathbb{C}\}$.

(**2**) Sprawdzić, czy A jest pierścieniem (odpowiednio pierścieniem z jednością, pierścieniem przemiennym), jeśli:
(*a*) $A = \{\begin{bmatrix} a & b \\ mb & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$ względem dodawania i mnożenia macierzy, gdzie m jest ustaloną liczbą całkowitą;
(*b*) $A = \{a + \sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Q}\}$ względem dodawania i mnożenia liczb (jak w ciele $\mathbb{R}$);
(*c*) $A = \{(a, b) \mid a, b \in \mathbb{R}\}$, gdzie „+" oraz „·" są określone według wzorów:

$$(a, b) + (c, d) = (a + c, b + d) \text{ oraz } (a, b) \cdot (c, d) = (ac + 2bd, ad + bc);$$

(*d*) $A = \mathbb{Q}[\sqrt{6}]$ względem dodawania „+" i mnożenia „·" liczb, gdzie $\mathbb{Q}[\sqrt{6}] = \{a + b\sqrt{6} \mid a, b \in \mathbb{Q}\}$.
(**3**) Znaleźć grupę multiplikatywną:
(*a*) $U(\mathbb{Z}_6)$;
(*b*) $U(\mathbb{Z}_8)$;
(*c*) $U(\mathbb{Z})$;
(*d*) $U(\mathbb{Q}[\sqrt{6}])$.
(**4**) Znaleźć elementy niebędące dzielnikami zera w pierścieniu:
(*a*) $\mathbb{Z}_9$;
(*b*) $\mathbb{Z}_{13}$;
(*c*) $M_2(\mathbb{R})$.
(**5**) Dla pierścienia nieprzemiennego A z jednością, gdzie $a, b \in A$, udowodnić, że:
(*a*) jeśli $a \in A$ jest elementem nilpotentnym, to $1 + a \in U(A)$;
(*b*) suma elementu odwracalnego i skończonej liczby elementów nilpotentnych jest elementem odwracalnym;
(*c*) $1 - ab$ jest elementem odwracalnym w pierścieniu A wtedy i tylko wtedy, gdy $1 - ba$ jest elementem odwracalnym w A.
(**6**) Sprawdzić, czy pierścień A posiada dzielniki zera, jeśli:
(*a*) $A = C_{[0,1]} = \{g : [0, 1] \to \mathbb{R} \mid \ g \text{ jest funkcją ciągłą na przedziale } [0, 1]\}$;
(*b*) $A = \{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{C}\}$.
(**7**) Znaleźć wszystkie idempotenty w pierścieniu: (*a*) $M_2(\mathbb{Q})$; (*b*) $M_2(\mathbb{Z})$.
(**8**) Dla pierścienia A udowodnić, że:
(*a*) jeśli $g^2 = g \in A$ oraz $x \in A$, to $(xg - gx)^2 = 0$;
(*b*) jeśli $x^3 = x$ dla dowolnych elementów $x \in A$, to pierścień A jest przemienny.

Uwagi. R. Dedekind[(12)] wprowadził termin „porządek", który nieco później, w 1892 r. (co zostało opublikowane w 1897 r.), D. Hilbert[(13)] zamienił na termin „pierścień (liczbowy)". Studiowanie pierścieni zaczęło się od

[(12)] Julius Wilhelm Richard Dedekind (1831–1916)
[(13)] David Hilbert (1862–1943)

badania pierścieni wielomianów i pierścienia liczb całkowitych. Pierwsza definicja aksjomatyczna pierścienia została zaproponowana przez A. Fraenkela[14] w 1914 r. Współczesne pojęcie pierścienia przemiennego pochodzi od E. Noether[15].

[14] Abraham Halevi (Adolf) Fraenkel (1891–1965)
[15] Amalie Emmy Noether (1882–1935)

2.7. Ciała i ich własności elementarne

Zaczynamy od takiej definicji.

■ Zbiór R jest nazywany *pierścieniem z dzieleniem* (względem dodawania „$+$” i mnożenia „$\cdot$”), jeśli są spełnione następujące warunki:

$T1$) $(R, +, \cdot)$ jest pierścieniem z jednością 1 (czyli zachodzą warunki 0_1), 0_2), 1)–6) oraz 8) z definicji pierścienia);

$T2$) zbiór R ma co najmniej dwa elementy (czyli moc zbioru $\operatorname{card} R \geqslant 2$);

$T3$) każdy element niezerowy $a \in R$ ma odwrotny w R, czyli

$$\forall_{a \in R \setminus \{0\}} \exists_{b \in R} : a \cdot b = 1 = b \cdot a.$$

■ Pierścień z dzieleniem jest też nazywany *ciałem*. Jeśli mnożenie „$\cdot$” jest przemienne w R, to $(R, +, \cdot)$ jest nazywane *ciałem przemiennym* (ale najczęściej krótko *ciałem*).

■ Każde ciało jest pierścieniem z jednością.

■ Nie każdy pierścień jest ciałem.

Na przykład pierścień liczb całkowitych $\mathbb{Z}$ nie jest ciałem, bo żaden element $a \in \mathbb{Z} \setminus \{0, 1, -1\}$ nie posiada odwrotnego do siebie w $\mathbb{Z}$.

■ Jeśli $(R, +, \cdot)$ jest ciałem, to

$$R^* = R \setminus \{0\}$$

jest grupą względem mnożenia „$\cdot$” (czyli $R^* = U(R)$). Grupa R^* jest nazywana *grupą multiplikatywną* ciała R.

■ Żadne ciało R nie posiada dzielników zera. Rzeczywiście, jeśli $ab = 0$ dla pewnych elementów $a, b \in R$, przy czym $a \neq 0$, to istnieje element odwrotny $a^{-1} \in R$, a zatem

$$0 = a^{-1}0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b.$$

■ Na podstawie przykładu 2.6.1(5) warunek $T2$) z definicji ciała oznacza, że w ciele R zawsze $0 \neq 1$.

Przykłady 2.7.1.

(**1**) Klasycznymi przykładami ciał (przemiennych) są ciało liczb wymiernych $\mathbb{Q}$, ciało liczb rzeczywistych $\mathbb{R}$ i ciało liczb zespolonych $\mathbb{C}$.

(**2**) Wykażmy, że

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$$

jest ciałem przemiennym (względem dodawania i mnożenia liczb). Oczywiście, że $1 = 1 + 0\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$. Jeśli $a, b, c, d \in \mathbb{Q}$, to

$$\begin{array}{lcl} (a + b\sqrt{2}) + (c + d\sqrt{2}) & = & (a + c) + (b + d)\sqrt{2} \in \mathbb{Q}[\sqrt{2}], \\ (a + b\sqrt{2})(c + d\sqrt{2}) & = & (ac + 2bd) + (ad + bc)\sqrt{2} \in \mathbb{Q}[\sqrt{2}], \end{array}$$

co oznacza, że dodawanie „+" i mnożenie „·" są działaniami algebraicznymi w $\mathbb{Q}[\sqrt{2}]$. Ponieważ $\mathbb{Q}[\sqrt{2}] \subseteq \mathbb{R}$, to działania dodawania i mnożenia są łączne, przemienne i związane prawem rozdzielności w zbiorze $\mathbb{Q}[\sqrt{2}]$. Ponadto $0 = 0 + 0\sqrt{2}$ jest elementem neutralnym względem dodawania, a $1 = 1 + 0\sqrt{2}$ jest elementem neutralnym względem mnożenia w $\mathbb{Q}[\sqrt{2}]$.

Łatwo przekonać się, że $a + b\sqrt{2} = 0$, gdzie $a, b \in \mathbb{Q}$, wtedy i tylko wtedy, gdy $a = 0$ oraz $b = 0$. Niech teraz $a + b\sqrt{2}$ będzie dowolnym elementem niezerowym z $\mathbb{Q}[\sqrt{2}]$. Wtedy $a^2 - 2b^2 \neq 0$. Znajdźmy takie $x, y \in \mathbb{Q}$, że

$$(a + b\sqrt{2})(x + y\sqrt{2}) = 1,$$

czyli

$$(ax + 2by) + (bx + ay)\sqrt{2} = 1.$$

Przepisując ostatnie równanie w postaci równoważnego układu równań liniowych (który jest układem Cramera)

$$\begin{cases} ax + 2by = 1, \\ bx + ay = 0, \end{cases}$$

obliczamy jego rozwiązanie

$$x = \frac{a}{a^2 - 2b^2},\ y = \frac{-b}{a^2 - 2b^2}.$$

Zatem

$$(a + b\sqrt{2})^{-1} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$$

oraz $\mathbb{Q}[\sqrt{2}]$ jest ciałem.

Podajmy jeszcze parę przykładów grup w pewien sposób związanych z ciałami.

Przykłady 2.7.2.

(**1**) **Inne podgrupy w $GL_2(\mathbb{F})$.** Niech $\mathbb{F}$ będzie ciałem. Zbiór

$$D_2 = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a, b \in \mathbb{F} \text{ oraz } ab \neq 0 \right\}.$$

tworzy podgrupę w $GL_2(\mathbb{F})$, która jest nazywana *grupą macierzy diagonalnych* stopnia 2 nad ciałem $\mathbb{F}$.

(**2**) Niech $\mathbb{F}$ będzie dowolnym ciałem oraz

$$T_n(\mathbb{F}) = \{[a_{ij}] \in M_n(\mathbb{F}) \mid a_{ij} = 0 \text{ dla wszystkich } i > j\}.$$

Wtedy $T_n(\mathbb{F})$ jest podgrupą w $GL_n(\mathbb{F})$, która jest nazywana *grupą macierzy trójkątnych górnych* stopnia n nad ciałem $\mathbb{F}$.

Lemat 2.7.3. *Charakterystyka każdego ciała $\mathbb{F}$ jest równa 0 lub jest liczbą pierwszą.*

Dowód. Niech 1 będzie jednością ciała $\mathbb{F}$ oraz $\operatorname{char}\mathbb{F} = n$. Załóżmy, że $n \neq 0$ oraz $n = mk$ dla pewnych liczb naturalnych m, k takich, że $1 < m, k < n$. Wtedy mamy elementy niezerowe

$$a = m \cdot 1 = \underbrace{1 + \cdots + 1}_{m \text{ składników}} \neq 0 \text{ oraz } \; b = k \cdot 1 = \underbrace{1 + \cdots + 1}_{k \text{ składników}} \neq 0,$$

lecz ich iloczyn

$$\begin{aligned} 0 = \underbrace{1 + \cdots + 1}_{n \text{ składników}} = (m \cdot k) \cdot 1 = mb = \underbrace{b + \cdots + b}_{m \text{ składników}} = \\ = \underbrace{1 \cdot b + \cdots + 1 \cdot b}_{m \text{ składników}} = (\underbrace{1 + \cdots + 1}_{m \text{ składników}}) \cdot b = a \cdot b \end{aligned}$$

jest zerowy, a więc $\mathbb{F}$ posiada dzielniki zera a, b, co jest niemożliwe. Z tych rozumowań wynika, że jeśli $\operatorname{char}\mathbb{F} \neq 0$, to $\operatorname{char}\mathbb{F}$ jest liczbą pierwszą. □

Lemat 2.7.4. *Jeśli $n \in \mathbb{N}^*$, to:*

(1) $\operatorname{char}\mathbb{Z}_n = n$;

(2) *pierścień klas reszt $\mathbb{Z}_n$ jest ciałem wtedy i tylko wtedy, gdy n jest liczbą pierwszą.*

Dowód. (1) Każdy element $\alpha \in \mathbb{Z}_n$ ma postać $\alpha = s + n\mathbb{Z}$, gdzie $s \in \mathbb{Z}$ jest takie, że $0 \leqslant s \leqslant n - 1$. Wtedy dla dowolnego $k \in \{1, 2, \ldots, n-1\}$ mamy

$$\overline{k} = k \cdot \overline{1} = \underbrace{\overline{1} + \cdots + \overline{1}}_{k \text{ składników}} \neq \overline{0}$$

oraz

$$n \cdot \overline{1} = \underbrace{\overline{1} + \cdots + \overline{1}}_{n \text{ składników}} = \overline{0}.$$

Zatem $\operatorname{char}\mathbb{Z}_n = n$.

(2) ($\Rightarrow$) Jeśli $\mathbb{Z}_n$ jest ciałem, to za lematem 2.7.3 wnosimy, że $n = \operatorname{char}\mathbb{Z}_n$ jest liczbą pierwszą.

($\Leftarrow$) Niech n będzie liczbą pierwszą oraz α będzie dowolnym elementem niezerowym z $\mathbb{Z}_n$. Wtedy $\alpha = k + n\mathbb{Z}$ dla pewnej liczby całkowitej k, gdzie $1 \leqslant k \leqslant n-1$. Skoro liczby k i n są względnie pierwsze, to z wniosku z algorytmu Euklidesa $ku + nv = 1$ dla pewnych $u, v \in \mathbb{Z}$. Zatem

$$\alpha\beta = ku + n\mathbb{Z} = 1 + n\mathbb{Z} = \overline{1}$$

dla klasy reszt $\beta = u + n\mathbb{Z}$. To znaczy, że każdy element niezerowy α ma odwrotny β w pierścieniu $\mathbb{Z}_n$, a zatem $\mathbb{Z}_n$ jest ciałem. □

■ Niech p będzie liczbą pierwszą. Ciało skończone $\mathbb{Z}_p$ dość często jest oznaczane w postaci równoważnej przez $\mathbb{F}_p$ lub przez $GF(p)$ (i nazywane prostym *ciałem Galois*). Elementy ciała $\mathbb{F}_p$ zwykle (w odróżnieniu od elementów ciała $\mathbb{Z}_p$ klas reszt modulo liczba pierwsza p) są zapisywane bez kresek u góry. Zbudujmy tabelki Cayleya dla dodawania „+” i mnożenia „·” elementów ciała $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$:

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Ciało $\mathbb{F}_2$ (elementy którego w informatyce nazywają *bitami*) jest nazywane *prostym dwójkowym* lub *binarnym*.

Ćwiczenia 2.7.5.

(**1**) Sprawdzić, czy T jest ciałem, jeśli:

(*a*) $T = \{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$;

(*b*) $T = \mathbb{Q}[\sqrt{2}]$, gdzie $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$;

(*c*) $T = \mathbb{Q}[i\sqrt{2}]$, gdzie $\mathbb{Q}[i\sqrt{2}] = \{a + bi\sqrt{2} \mid a, b \in \mathbb{Q}\}$;

(*d*) $T = \mathcal{B}(X)$, gdzie X jest zbiorem niepustym oraz $\mathcal{B}(X) = \{Y \mid Y \subseteq X\}$, względem działań „$\cup$" oraz „$\cap$";

(*e*) $T = \mathcal{B}(X)$ względem działań „$\triangle$" oraz „$\cap$", gdzie $\mathcal{B}(X) = \{Y \mid Y \subseteq X\}$, $X \neq \varnothing$ oraz $A \triangle B = (A \setminus B) \cup (B \setminus A)$ dla dowolnych $A, B \in \mathcal{B}(X)$.

(**2**) Sprawdzić, czy $(T, \oplus, \bigcirc)$ jest ciałem (przemiennym), jeśli:

(*a*) $T = \mathbb{R}$ oraz działania są określone wzorami $a \oplus b = a + b + 1$ oraz $a \bigcirc b = a + b + ab$;

(*b*) $T = \{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{C}\}$, gdzie „$\oplus$" oraz „$\bigcirc$" są działaniami dodawania i mnożenia macierzy;

(*c*) $T = \mathbb{R}$, gdzie „$\bigcirc$" jest mnożeniem liczb, a $a \oplus b = \sqrt[5]{a^5 + b^5}$;

(*d*) $T = \{f : \mathbb{R} \to \mathbb{R} \mid f(x) = a\cos x + b\sin x \text{ dla } a, b \in \mathbb{R}\}$, gdzie „$\oplus$" jest działaniem popunktowego dodawania funkcji, a działanie „$\bigcirc$" jest określone wzorem

$$(f \bigcirc g)(x) = \int_0^{2\pi} f(t)g(x - t)dt$$

dla $f, g \in T$ oraz $x \in \mathbb{R}$;

(*e*) $T = \mathbb{Z}_3[i]$, gdzie $\mathbb{Z}_3[i] = \{a + bi \mid a, b \in \mathbb{Z}_3\}$, a „$\oplus$" oraz „$\bigcirc$" są zwykłymi działaniami, które są indukowane dodawaniem i mnożeniem liczb zespolonych.

Uwagi. Liczbami wymiernymi i rzeczywistymi (jako elementami ciała) matematycy zaczęli operować jeszcze przed Euklidesem. W XIX w. zostało skonstruowane ciało liczb zespolonych. W 1801 r., po opublikowaniu słynnej pracy C. Gaussa *Disquisitiones Arithmeticae*, w matematykę weszło pojęcie ciała klas reszt $\mathbb{Z}_p$. W tej pracy C. Gauss badał również ciała cyklotomiczne. E. Galois studiował skończone rozszerzenia ciała $\mathbb{Z}_p$. W 1893 r. matematyk amerykański E. Moore[16] udowodnił, że każde ciało skończone jest izomorficzne (co z punktu widzenia algebry jest tym samym) pewnemu rozszerzeniu Galois. Definicję ciała wprowadził R. Dedekind w 1871 r. W. Weber[17] jako pierwszy zaczął aksjomatyzować teorię ciał.

[16] Eliakim Hastings Moore (1862–1932)

[17] Wilhelm Eduard Weber (1804–1891)

2.8. Podpierścienie i podciała

■ Niech $(R, +, \cdot)$ będzie pierścieniem (z jednością 1). Zbiór S jest nazywany *podpierścieniem* pierścienia R, jeśli są spełnione warunki:
$0_1)$ $S \neq \varnothing$;
$0_2)$ $S \subseteq R$;
1) $(S, +, \cdot)$ jest pierścieniem (z jednością 1), czyli S jest pierścieniem (z jednością 1) względem tych samych działań co i R, lecz zawężonych do podzbioru S.

■ Podpierścień S pierścienia R taki, że $S \neq R$ jest nazywany *właściwym*.

■ Niech $(F, +, \cdot)$ będzie ciałem. Zbiór S jest nazywany *podciałem* ciała F, jeśli są spełnione warunki:
$0_1)$ $S \neq \varnothing$;
$0_2)$ $S \subseteq F$;
1) $(S, +, \cdot)$ jest ciałem, czyli podzbiór S tworzy ciało względem tych samych działań co i F, lecz zawężonych do podzbioru S.

Przykłady 2.8.1.

(**1**) Każdy pierścień R zawiera podpierścienie *trywialne*: podpierścień *zerowy* $\{0\}$, składający się z dokładnie jednego elementu zerowego 0, oraz podpierścień *niewłaściwy* R (czyli pierścień R jest swoim podpierścieniem).

(**2**) Niech

$$K = \{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \mid a \in \mathbb{Q} \}.$$

Wtedy K jest podpierścieniem w pierścieniu macierzy $M_2(\mathbb{Q})$.

(**3**) Niech

$$\mathbb{Q}[\sqrt{7}] = \{a + b\sqrt{7} \mid a, b \in \mathbb{Q}\}.$$

Wtedy $\mathbb{Q}[\sqrt{7}]$ jest podciałem w ciele liczb rzeczywistych $\mathbb{R}$ (przekonać się samodzielnie).

Twierdzenie 2.8.2 (kryterium podpierścienia z jednością)**.** *Mamy pierścień $(R, +, \cdot)$ z jednością 1. Wtedy S jest podpierścieniem w R w tym i tylko tym przypadku, gdy są spełnione własności:*
(0_1) $S \neq \varnothing$;
(0_2) $S \subseteq R$;
(1) $a - b \in S$ *dla dowolnych elementów* $a, b \in S$;
(2) $a \cdot b \in S$ *dla dowolnych* $a, b \in S$;
(3) $1 \in S$.

Dowód. ($\Rightarrow$) Niech S będzie podpierścieniem pierścienia R z 1. Wtedy na mocy definicji S jest podzbiorem niepustym w R oraz $(S, +, \cdot)$ jest pierścieniem z jednością 1. Zatem $a \cdot b \in S$ dla dowolnych elementów $a, b \in S$. Ponadto $(S, +)$ jest grupą, a więc $a - b \in S$ na podstawie kryterium podgrupy (patrz twierdzenie 2.5.2).

($\Leftarrow$) Załóżmy, że S spełnia warunki (0_1), (0_2) oraz (1)–(3). Wtedy S jest podzbiorem niepustym w R, na którym są określone dwa działania algebraiczne (dodawanie „+" i mnożenie „·") i który posiada element jednostkowy 1. Ponieważ $S \subseteq R$ oraz dodawanie „+" jest przemienne i łączne w R, mnożenie „·" jest łączne w R i dwa działania są związane rozdzielnością w R, to podobne własności zachodzą dla elementów z S. Oprócz tego dla elementu $a \in S$ mamy

$$0 = a - a \in S \text{ oraz } -a = 0 - a \in S.$$

Zatem są spełnione wszystkie warunki z definicji pierścienia i $(S, +, \cdot)$ jest pierścieniem z jednością 1. Wnosimy, że S jest podpierścieniem w R. □

W tym przypadku, gdy nie zwracamy uwagi, czy pierścień R ma jedność, to kryterium podpierścienia jest takie.

Twierdzenie 2.8.3 (kryterium podpierścienia)**.** *Niech $(R, +, \cdot)$ będzie pierścieniem. Wtedy S jest podpierścieniem w R w tym i tylko tym przypadku, gdy są spełnione następujące warunki:*

(0_1) $S \neq \varnothing$;
(0_2) $S \subseteq R$;
(1) $a - b \in S$ *dla wszystkich elementów* $a, b \in S$;
(2) $a \cdot b \in S$ *dla wszystkich* $a, b \in S$.

□

Twierdzenie 2.8.4 (kryterium podciała)**.** *Niech $(F, +, \cdot)$ będzie ciałem. Wtedy S jest podciałem w F w tym i tylko tym przypadku, gdy są spełnione następujące własności:*

(0_1) $S \neq \varnothing$;
(0_2) $S \subseteq F$;
(1) $a - b \in S$ *dla wszystkich elementów* $a, b \in S$;
(2) $a \cdot b \in S$ *dla wszystkich* $a, b \in S$;

(3) $1 \in S$ *(lub równoważnie moc* $\operatorname{card} S \geqslant 2$*);*
(4) $a^{-1} \in S$ *dla każdego* $a \in S \setminus \{0\}$.

Dowód jest podobny do dowodu kryterium podpierścienia.

□

Przykłady 2.8.5.

(**1**) Podzbiór

$$n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$$

jest podpierścieniem pierścienia $\mathbb{Z}$ dla każdej liczby całkowitej n. W rzeczy samej, $n = n \cdot 1 \in n\mathbb{Z}$. Jeśli α, β są dowolnymi elementami z $n\mathbb{Z}$, to $\alpha = na$ i $\beta = nb$ dla pewnych $a, b \in \mathbb{Z}$. Wtedy

$$\begin{array}{rcl} \alpha - \beta & = & na - nb = n(a-b) \in n\mathbb{Z}, \\ \alpha\beta & = & (na)(nb) = n(anb) \in n\mathbb{Z}. \end{array}$$

Zatem wszystkie warunki z twierdzenia 2.8.4 zachodzą.

(**2**) Udowodnijmy, że

$$\mathbb{Q}(i) = \{a + bi \mid a, b \in \mathbb{Q}\}$$

jest ciałem, gdzie $i \in \mathbb{C}$. W tym celu na podstawie kryterium podciała wystarczy udowodnić, że $\mathbb{Q}(i)$ jest podciałem w ciele liczb zespolonych $\mathbb{C}$. Najpierw zauważmy, że $i = 0 + 1 \cdot i \in \mathbb{Q}(i)$ oraz $\mathbb{Q}(i)$ jest podzbiorem niepustym w $\mathbb{C}$. Jeśli $\alpha = a + bi$ i $\beta = c + di$, gdzie $a, b, c, d \in \mathbb{Q}$, to $a - c, b - d, ac - bd, ad + bc \in \mathbb{Q}$, a więc

$$\begin{array}{rcl} \alpha - \beta & = & (a-c) + (b-d)i \in \mathbb{Q}(i) \\ \alpha\beta & = & (ac - bd) + (ad + bc)i \in \mathbb{Q}(i). \end{array}$$

Ponadto $\alpha \neq 0$ wtedy i tylko wtedy, gdy $a^2 + b^2 \neq 0$. Zatem jeśli $\alpha \neq 0$, to

$$\alpha^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i \in \mathbb{Q}(i).$$

To znaczy, że $\mathbb{Q}(i)$ jest podciałem ciała $\mathbb{C}$. Na tej podstawie wnosimy, że $\mathbb{Q}(i)$ jest ciałem.

(**3**) Przekonajmy się, że

$$\mathbb{Z}[\frac{1}{15}] = \{\frac{a}{15^n} \mid a \in \mathbb{Z} \text{ oraz } \ n \in \mathbb{N}\}$$

jest podpierścieniem w $\mathbb{Q}$ względem działań dodawania i mnożenia liczb. Istotnie

$$1 = \frac{1}{15^0} \in \mathbb{Z}[\frac{1}{15}] \ \text{ oraz } \ 0 = \frac{0}{15} \in \mathbb{Z}[\frac{1}{15}].$$

Jeśli α, β są dowolnymi elementami z $\mathbb{Z}[\frac{1}{15}]$, to

$$\alpha = \frac{a}{15^n} \ \text{ oraz } \ \beta = \frac{b}{15^m}$$

dla pewnych $a, b \in \mathbb{Z}$ oraz $n, m \in \mathbb{N}$. Wtedy

$$\begin{array}{rcl} \alpha - \beta & = & \frac{a}{15^n} - \frac{b}{15^m} = \frac{a15^m - b15^n}{15^{n+m}} \in \mathbb{Z}[\frac{1}{15}] \\ \alpha\beta & = & \frac{a}{15^n}\frac{b}{15^m} = \frac{ab}{15^{n+m}} \in \mathbb{Z}[\frac{1}{15}]. \end{array}$$

Na podstawie twierdzenia 2.8.3 wnosimy, że $\mathbb{Z}[\frac{1}{15}]$ jest podpierścieniem (z jednością 1) w $\mathbb{Q}$.

(**4**) Łatwo zauważyć, że ciało liczb rzeczywistych $\mathbb{R}$ jest podciałem w ciele liczb zespolonych $\mathbb{C}$, a ciało liczb wymiernych $\mathbb{Q}$ – podciało w $\mathbb{R}$ oraz w $\mathbb{C}$.

(**5**) Pierścień macierzy $M_n(\mathbb{Q})$ jest podpierścieniem (z jednością I_n) w pierścieniu $M_n(\mathbb{R})$ i w pierścieniu $M_n(\mathbb{C})$.

(**6**) **Pierścień formalnych szeregów potęgowych.** Niech R będzie pierścieniem oraz X będzie symbolem. Jeśli dla formalnych szeregów potęgowych

$$f = \sum_{i=0}^{\infty} a_i X^i \text{ oraz } \ g = \sum_{i=0}^{\infty} b_i X^i$$

położyć

$$\begin{aligned} f+g &= \sum_{i=0}^{\infty} a_i X^i + \sum_{i=0}^{\infty} b_i X^i = \sum_{i=0}^{\infty} (a_i + b_i) X^i \\ fg &= (\sum_{i=0}^{\infty} a_i X^i)(\sum_{i=0}^{\infty} b_i X^i) = \sum_{k=0}^{\infty} c_k X^k, \end{aligned}$$

gdzie

$$c_k = \sum_{i+j=k} a_i b_j,$$

to, sprawdzając spełnienie warunków (przy założeniu, że $aX = Xa$ dla wszystkich elementów $a \in R$) z definicji pierścienia, przekonujemy się, że

$$R[[X]] = \{\sum_{i=0}^{\infty} a_i X^i \mid a_i \in R \ (i \in \mathbb{N})\}$$

jest pierścieniem (nazywanym *pierścieniem formalnych szeregów potęgowych jednej zmiennej X nad pierścieniem R*). Jeśli pierścień R jest przemienny i posiada jedność 1, to pierścień $R[[X]]$ też jest przemienny z 1.

Termin „formalny" oznacza, że szeregi rozpatrujemy z algebraicznego punktu widzenia (nie zwracając uwagi na ich zbieżność itd.).

(**7**) **Pierścień wielomianów**

$$R[X] = \{\sum_{i=0}^{n} a_i X^i \mid a_i \in R, \ n \in \mathbb{N} \ (i = 0, 1, \ldots n)\}$$

jest podpierścieniem w pierścieniu formalnych szeregów potęgowych $R[[X]]$.

(**8**) (**Pierścień grupowy**) Połączmy teraz własności grupy $(G, *)$ z własnościami pierścienia unitarnego $(R, +, \cdot)$ w takiej konstrukcji.

Niech $R[G]$ będzie zbiorem wszystkich formalnych sum postaci

$$\sum_{g \in G} \alpha_g g \ \ (g \in G),$$

w których jest tylko skończona liczba współczynników α_g niezerowych.

Niech dalej

$$\alpha = \sum_{g \in G} \alpha_g g, \quad \beta = \sum_{g \in G} \beta_g g.$$

Mówimy, że elementy $\alpha, \beta \in R[G]$ są *równe*, jeśli

$$\forall_{g\in G} : \alpha_g = \beta_g.$$

W zbiorze $R[G]$ wprowadźmy dwa działania według wzorów:

- dodawanie „+”

$$\alpha + \beta = \sum_{g\in G} \alpha_g g + \sum_{g\in G} \beta_g g = \sum_{g\in G} (\alpha_g + \beta_g),$$

- mnożenie „·”

$$\alpha \cdot \beta = \left(\sum_{g\in G} \alpha_g g\right) \cdot \left(\sum_{g\in G} \beta_g g\right) = \sum_{g\in G} \left(\sum_{h\in G} \alpha_h \beta_{h^{-1}*g}\right).$$

Przyjmujemy, że $r \cdot g = g \cdot r$ dla wszystkich $r \in R$ oraz $g \in G$. Jeśli e jest elementem neutralnym grupy G, a 1 jednością pierścienia R, to zakładamy, że $r \cdot e = r$ dla każdego $r \in R$ (a więc $1 \cdot e = e$). Wtedy $(R[G], +, \cdot)$ jest pierścieniem z jednością 1 (nazywanym *pierścieniem grupowym* grupy G nad pierścieniem R).

Ćwiczenia 2.8.6.

(**1**) Niech A będzie pierścieniem oraz $Z(A) = \{z \in A \mid zr = rz \text{ dla wszystkich } r \in A\}$. Udowodnić, że:
(a) $Z(A)$ jest podpierścieniem w A (podpierścień $Z(A)$ nazywamy *centrum* pierścienia A);
(b) jeśli $x^2 - x \in Z(A)$ dla dowolnych $x \in A$, to pierścień A jest przemienny.
(**2**) Sprawdzić, czy S jest podpierścieniem w A, jeśli:
(a) $A = C_{[0,1]}$ oraz $S = \{f \in C_{[0,1]} \mid f(0) = f(1)\}$;
(b) $A = C_{[0,1]}$ oraz $S = \{f \in C_{[0,1]} \mid f(0) = 0\}$;
(c) $A = C_{[0,1]}$ oraz $S = \{f \in C_{[0,1]} \mid \int_0^1 f(x)dx = 0\}$;
(d) $A = M_2(\mathbb{R})$ oraz $S = \{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(e) $A = M_2(\mathbb{R})$ oraz $S = \{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(f) $A = M_2(\mathbb{R})$ oraz $S = \{\begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(g) $A = M_2(\mathbb{R})$ oraz $S = \{\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \mid a, b \in \mathbb{R}\}$;
(h) $A = M_2(\mathbb{R})$ oraz $S = \{A \in M_2(\mathbb{R}) \mid \det A \neq 0\}$;
(i) $A = \mathbb{Q}$ oraz $S = \mathbb{Z}[\frac{1}{2}] = \{\frac{a}{2^k} \mid a \in \mathbb{Z}, k \in \mathbb{N}\}$;
(j) $A = C_{[0,1]}$ oraz $S = \{f \in C_{[0,1]} \mid \lim_{x\to\frac{1}{2}} f(x) = 0\}$;
(k) $A = C_{[0,1]}$ oraz $S = \{f \in C_{[0,1]} \mid |f(x)| \leqslant 3 \text{ dla dowolnego } x \in [0,1]\}$.
(**3**) Sprawdzić, czy T jest podciałem w ciele R, jeśli:
(a) $T = \mathbb{Q}[\sqrt{7}]$ oraz $R = \mathbb{R}$;
(b) $T = \{a + b\sqrt[3]{2} \mid a, b \in \mathbb{Q}\}$ oraz $R = \mathbb{R}$;
(c) $T = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$ oraz $R = \mathbb{R}$;
(d) $T = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Q}\}$ oraz $R = \mathbb{Q}$.

Uwagi. Podpierścienie i podciała weszły do algebry wraz z pierścieniami i ciałami.

Rozdział 3

Początki teorii grup

Grupa jest jedną z głównych struktur algebraicznych i, w szczególności, jest ważnym instrumentem w badaniu symetrii obiektów geometrycznych i fizycznych. Teoria grup powstała w wyniku poszukiwań rozwiązań równań algebraicznych. Równania liniowe i kwadratowe rozwiązywali jeszcze matematycy starożytni. Przed 1550 r. odkryto metodę rozwiązywania równań algebraicznych stopnia 3. Następnie wyniknęły różne i niezrozumiałe trudności. Stałe poszukiwania rozwiązań równań algebraicznych stopni $\geqslant 4$ doprowadziły do powstania teorii grup. Norweski matematyk N.H. Abel udowodnił w 1823 r., że w ogólnym przypadku rozwiązanie takich równań w pierwiastkach nie jest możliwe. Francuski matematyk E. Galois stał się pionierem w dziedzinie teorii grup.

Algebrą, w szczególności, zajmował się polski matematyk J.M. Hoene-Wroński[(1)], który w 1812 r. w pracy *Résolution générale des équations de tous les dégrés* przedstawił metodę rozwiązywania pewnych typów równań algebraicznych n-tego stopnia.

3.1. Grupy cykliczne

■ Niech $(G, \cdot)$ będzie grupą. Będziemy mówić, że G jest grupą *cykliczną* generowaną przez element g, jeśli

$$G = \{g^n \mid n \in \mathbb{Z}\}.$$

(1) Józef Maria Hoene-Wroński (1778–1853)

■ Jeśli G jest grupą cykliczną generowaną przez element g (lub, mówiąc inaczej, g jest *generatorem* grupy cyklicznej G), to będziemy oznaczać to przez

$$G = \langle g \rangle.$$

Jeśli G jest grupą względem dodawania „+", to G jest nazywana *cykliczną* z generatorem g, jeśli

$$G = \{ng \mid n \in \mathbb{Z}\};$$

w tym przypadku też zapisujemy, że $G = \langle g \rangle$.

Lemat 3.1.1. *Każda grupa cykliczna jest abelowa.*

Dowód. W rzeczy samej, niech $(G, \cdot)$ będzie grupą cykliczną generowaną przez element g. Wtedy dla dowolnych $z_1, z_2 \in G$ znajdą się takie liczby całkowite m_1, m_2, że $z_i = g^{m_i} \ \ (i = 1, 2)$, a zatem

$$z_1 \cdot z_2 = g^{m_1} \cdot g^{m_2} = g^{m_1+m_2} = g^{m_2+m_1} = g^{m_2} \cdot g^{m_1} = z_2 \cdot z_1.$$

□

Przykłady 3.1.2.

(**1**) Grupa jednostkowa $\{e\} = \langle e \rangle$ jest cykliczna z generatorem e. W przypadku addytywnym grupa zerowa $\{0\} = \langle 0 \rangle$ też jest cykliczna.

(**2**) Addytywna grupa liczb całkowitych $\mathbb{Z}$ jest cykliczna z generatorem 1, bo

$$\mathbb{Z} = \{n \cdot 1 \mid n \in \mathbb{Z}\} = \langle 1 \rangle,$$

gdzie

$$n \cdot 1 = \begin{cases} \underbrace{1 + \cdots + 1}_{n \text{ składników}}, & \text{gdy } n > 0, \\ 0, & \text{gdy } n = 0, \\ \underbrace{(-1) + \cdots + (-1)}_{|n| \text{ składników}}, & \text{gdy } n < 0. \end{cases}$$

Podobnie, liczba -1 też jest innym generatorem grupy $\mathbb{Z}$, czyli $\mathbb{Z} = \langle -1 \rangle$. Wnosimy, że wybór generatora grupy cyklicznej nie jest jednoznaczny.

(**3**) Oczywiście, że dla liczb $-1, 1 \in \mathbb{R}$ grupa

$$\{-1, 1\} = \langle -1 \rangle$$

jest cykliczna z generatorem -1.

(**4**) Ponieważ

$$(12)(13) = (132) \neq (123) = (13)(12),$$

to grupa $\mathbb{S}_3$ nie jest abelowa i w wyniku lematu 3.1.1 grupa $\mathbb{S}_3$ nie jest cykliczna.

(**5**) Addytywna grupa klas reszt $\mathbb{Z}_n$ modulo n jest cykliczna, bo $0_n = 0 \cdot 1_n$ i dla każdego elementu niezerowego $k_n \in \mathbb{Z}_n$ $(0 < k \leqslant n-1; k \in \mathbb{N})$ mamy

$$k_n = \underbrace{1_n + \cdots + 1_n}_{k \text{ składników}} = k \cdot 1_n.$$

To znaczy, że $\mathbb{Z}_n = \langle 1_n \rangle$ jest cykliczna.

■ Niech $(G, \cdot)$ będzie grupą. Będziemy mówić, że H jest podgrupą *cykliczną* grupy G, jeśli H jest podgrupą w G oraz $(H, \cdot)$ jest grupą cykliczną.

Twierdzenie 3.1.3. *Każda podgrupa grupy cyklicznej także jest grupą cykliczną.*

Dowód. Niech G będzie grupą cykliczną generowaną przez g, a H będzie dowolną jej podgrupą. Jeśli $H = \{e\}$ jest jednostkowa, to jest cykliczna. Załóżmy, że podgrupa H nie jest jednostkowa, a więc istnieje $e \neq h \in H$. Wtedy znajdzie się taka liczba całkowita s, że $h = g^s$. Ponieważ $h^{-1} = g^{-s} \in H$, to zbiór

$$S = \{n \in \mathbb{N}^* \mid g^n \in H\}$$

jest niepusty. Na podstawie zasady minimum wynika, że S posiada najmniejszą liczbę całkowitą dodatnią; niech to będzie $k_0 \in S$. Ponieważ H jest podgrupą, to $(g^{k_0})^m \in H$ dla każdego $m \in \mathbb{Z}$, a więc $\langle g^{k_0} \rangle \subseteq H$. Z innej strony, niech $z = g^l$ będzie dowolnym elementem z H. Na podstawie twierdzenia o dzieleniu z resztą znajdą się takie liczby całkowite q, r, że $l = qk_0 + r$, gdzie $0 \leqslant r < k_0$. Zatem

$$z = g^l = g^{qk_0+r} = (g^{k_0})^q g^r,$$

a stąd

$$g^r = (g^{k_0})^{-q} z \in H.$$

W wyniku minimalnego wyboru k_0 wnosimy, że $r = 0$, czyli

$$z = (g^{k_0})^q \in \langle g^{k_0} \rangle.$$

To znaczy, że $\langle g^{k_0} \rangle \supseteq H$. Z udowodnionego wyżej wnioskujemy, że podgrupa

$$H = \langle g^{k_0} \rangle$$

jest cykliczna. □

Lemat 3.1.4. *Rząd $|\langle g\rangle|$ grupy cyklicznej $\langle g\rangle$ generowanej przez element g jest równy rzędowi $o(g)$ generatora g, czyli $|\langle g\rangle| = o(g)$.*

Dowód. 1) Niech $n = o(g) < \infty$. Wtedy

$$g^n = e \text{ oraz } g^k \neq e$$

dla każdej dodatniej liczby całkowitej $k < n$. Jeśli

$$g^{k_1} = g^{k_2}$$

dla różnych liczb całkowitych k_1, k_2 takich, że $0 \leqslant k_1, k_2 < n$, to biorąc pod uwagę, że jedna z nich jest większa od innej (na przykład $k_2 > k_1$), otrzymujemy $g^{k_2-k_1} = e$, a więc

$$o(g) \leqslant k_2 - k_1 < n,$$

co jest sprzeczne z założeniem. Zatem elementy $e = g^0, g, \ldots, g^{n-1}$ są parami różne. Ponieważ dla każdej liczby całkowitej m znajdą się takie $q, r \in \mathbb{Z}$, że $m = nq + r$, przy czym $0 \leqslant r < n$, to

$$g^m = g^{nq+r} = (g^n)^q \cdot g^r = e \cdot g^r = g^r \in \{e = g^0, g, \ldots, g^{n-1}\},$$

czyli $\langle g\rangle \subseteq \{e = g^0, g, \ldots, g^{n-1}\}$. Natomiast

$$\{e = g^0, g, \ldots, g^{n-1}\} \subseteq \{g^s \mid s \in \mathbb{Z}\} = \langle g\rangle.$$

Zatem

$$\langle g\rangle = \{e = g^0, g, \ldots, g^{n-1}\}$$

i, jako wniosek, $|\langle g\rangle| = n$.

2) Załóżmy teraz, że $|g| = \infty$. Podobnie jak wyżej otrzymujemy, że $g^{k_1} \neq g^{k_2}$ dla różnych $k_1, k_2 \in \mathbb{N}$, a zatem zbiór

$$\{g^t \mid t \in \mathbb{Z}\}$$

jest nieskończony. Wnosimy stąd, że rząd $|\langle g\rangle| = \infty$.

3) Zostawiamy Czytelnikowi do samodzielnego rozpatrzenia przypadek, jaki jest rząd $|g|$, gdy $|\langle g\rangle| < \infty$ (odpowiednio $|\langle g\rangle| = \infty$). □

Przykłady 3.1.5.

(**1**) Niech n będzie liczbą całkowitą. Wtedy $(-n)\mathbb{Z} = n\mathbb{Z}$. Udowodnimy, że jeśli m, n są liczbami całkowitymi, to $m\mathbb{Z} \leqslant n\mathbb{Z}$ wtedy i tylko wtedy, gdy $n \mid m$. Rzeczywiście, jeśli $m\mathbb{Z} \leqslant n\mathbb{Z}$, to $m \in n\mathbb{Z}$ i $m = nt$ dla pewnej liczby całkowitej t, a więc n dzieli m. Odwrotnie, jeśli $n \mid m$, to $m = ns$ dla pewnej liczby całkowitej s, a zatem $m \in n\mathbb{Z}$. Jeśli u jest dowolnym elementem z $m\mathbb{Z}$, to $u = ml$ dla pewnego $l \in \mathbb{Z}$ i, jako wniosek, $u = n(sl) \in n\mathbb{Z}$, co ciągnie za sobą zawieranie $m\mathbb{Z} \leqslant n\mathbb{Z}$.

Zauważmy, że grupa addytywna $\mathbb{Z}$ posiada nieskończony ściśle malejący łańcuch podgrup; na przykład

$$\mathbb{Z} > 2\mathbb{Z} > \cdots > 2^k\mathbb{Z} > \cdots,$$

gdzie $k \in \mathbb{N}^*$. Z innej strony, każdy rosnący łańcuch jej podgrup

$$H_1 \leqslant H_2 \leqslant \cdots \leqslant H_n \leqslant \cdots$$

stabilizuje się przez skończoną liczbę kroków, czyli znajdzie się taki indeks $n \in \mathbb{N}$, że zachodzi równość

$$H_n = H_{n+1} = \ldots \tag{3.1}$$

Rzeczywiście, H_i jest podgrupą cykliczną w $\mathbb{Z}$, a więc $H_i = \langle h_i \rangle$ dla pewnego $h_i \in H$. Z udowodnionego powyżej $h_i \mid h_1$ dla wszystkich $i \in \mathbb{N}$. Ponieważ liczba całkowita h_1 posiada tylko skończoną liczbę parami różnych dzielników h_i, to (3.1) zachodzi dla pewnego $n \in \mathbb{N}$.

(**2**) Znajdźmy wszystkie podgrupy grupy $\mathbb{Z}_6 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}\}$. Najpierw zauważmy, że $\mathbb{Z}_6$ jest grupą cykliczną z generatorem $\overline{1}$. Niech H będzie dowolną podgrupą z $\mathbb{Z}_6$. Na podstawie twierdzenia 3.1.3 grupa H jest cykliczna, a zatem $H = \langle g \rangle$ dla pewnego elementu $g \in \mathbb{Z}_6$. Rozpatrzmy możliwe przypadki.

- Jeśli $g = \overline{0}$, to $H = \langle \overline{0} \rangle$.
- Niech $g = \overline{1}$. Wtedy mamy

$$\begin{array}{rcl} \overline{1} & \neq & \overline{0}, \\ \overline{2} & = & \overline{1} + \overline{1} \in H, \\ \overline{3} & = & \overline{1} + \overline{1} + \overline{1} \in H, \\ \overline{4} & = & \overline{1} + \overline{1} + \overline{1} + \overline{1} \in H, \\ \overline{5} & = & \overline{1} + \overline{1} + \overline{1} + \overline{1} + \overline{1} \in H, \\ \overline{0} & = & \overline{1} + \overline{1} + \overline{1} + \overline{1} + \overline{1} + \overline{1} \in H, \end{array}$$

a więc $\langle \overline{1} \rangle = \mathbb{Z}_6$.

- Jeśli zaś $g = \overline{2}$, to

$$\begin{array}{rcl} \overline{2} & \neq & \overline{0}, \\ \overline{4} & = & \overline{2} + \overline{2} \in H, \\ \overline{0} & = & \overline{2} + \overline{2} + \overline{2} \in H, \end{array}$$

czyli $\langle \overline{2} \rangle = \{\overline{0}, \overline{2}, \overline{4}\}$.

- Załóżmy, że $g = \overline{3}$. Wtedy

$$\overline{3},\ \overline{0} = \overline{3} + \overline{3} \in H$$

oraz $\langle \overline{3} \rangle = \{\overline{0}, \overline{3}\}$.

- W przypadku gdy $g = \overline{4}$, otrzymujemy

$$\begin{array}{rcl} \overline{4} & \neq & \overline{0}, \\ \overline{2} & = & \overline{4} + \overline{4} \in H, \\ \overline{0} & = & \overline{4} + \overline{4} + \overline{4} \in H \end{array}$$

oraz $\langle\overline{4}\rangle = \{\overline{0}, \overline{2}, \overline{4}\} = \langle\overline{2}\rangle$.

• W końcu, jeśli $g = \overline{5}$, to

$$\begin{array}{rcl} \overline{5} & \neq & \overline{0}, \\ \overline{4} & = & \overline{5}+\overline{5} \in H, \\ \overline{3} & = & \overline{5}+\overline{5}+\overline{5} \in H, \\ \overline{2} & = & \overline{5}+\overline{5}+\overline{5}+\overline{5} \in H, \\ \overline{1} & = & \overline{5}+\overline{5}+\overline{5}+\overline{5}+\overline{5} \in H, \\ \overline{0} & = & \overline{5}+\overline{5}+\overline{5}+\overline{5}+\overline{5}+\overline{5} \in H, \end{array}$$

i wtedy $\langle\overline{5}\rangle = \mathbb{Z}_6$.

Zatem grupa $\mathbb{Z}_6$ ma dokładnie 4 parami różne podgrupy: $\langle\overline{0}\rangle, \langle\overline{2}\rangle, \langle\overline{3}\rangle$ oraz $\mathbb{Z}_6$.

* * *

■ **Własności rzędów elementów w grupie.** Zaczniemy od takiego lematu.

Lemat 3.1.6. *Niech G będzie grupą oraz $a \in G$. Jeśli rząd $o(a) = dt$ dla pewnych $d, t \in \mathbb{N}^*$, to rząd $o(a^d) = t$.*

Dowód. Niech e będzie elementem neutralnym grupy G. Skoro

$$(a^d)^t = e \quad \Rightarrow \quad o(a^d) \mid t$$

na podstawie lematu 2.3.7 oraz $(a^d)^i \neq e$ dla wszystkich $i = 1, \ldots, t-1$, to rząd

$$o(a^d) = t.$$

□

Lemat 3.1.7. *Niech G będzie grupą, $k \in \mathbb{N}^*$ oraz $a \in G$. Jeśli rząd $o(a) = n$ dla pewnego $n \in \mathbb{N}^*$ oraz $d = \mathrm{NWD}(k, n)$, to*

$$\langle a^k \rangle = \langle a^d \rangle.$$

Dowód. Na podstawie lematu 1.2.8 zachodzi, że $d = ku + nv$ dla pewnych $u, v \in \mathbb{Z}$. Wtedy

$$a^d = (a^k)^u \cdot (a^n)^v = (a^k)^u \in \langle a^k \rangle.$$

Jako wniosek $\langle a^d \rangle \leqslant \langle a^k \rangle$. Natomiast $k = dt$ dla pewnego $t \in \mathbb{N}^*$, a więc

$$a^k = (a^d)^t \in \langle a^d \rangle,$$

co powoduje, że $\langle a^k \rangle \leqslant \langle a^d \rangle$. Zatem $\langle a^d \rangle = \langle a^k \rangle$. □

Twierdzenie 3.1.8. *Niech G będzie grupą, $k \in \mathbb{N}^*$ oraz $a \in G$. Jeśli rząd $o(a) = n$, to*

$$o(a^k) = \frac{n}{\mathrm{NWD}(k,n)}.$$

Dowód. Oznaczmy $d = \mathrm{NWD}(k,n)$. Wtedy $n = dt$ dla pewnego $t \in \mathbb{N}^*$ oraz na podstawie lematu 3.1.6 rząd $o(a^d) = t$. Zatem na mocy lematów 3.1.6, 3.1.4, 3.1.7 mamy

$$\frac{n}{d} = t = o(a^d) = |\langle a^d \rangle| = |\langle a^k \rangle| = o(a^k)$$

i teza zachodzi. □

■ **Liczba generatorów skończonej grupy cyklicznej.** Odwzorowanie $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ jest nazywane *funkcją Eulera*[(2)], jeśli

$$\varphi(n) = |\{i \mid 1 \leqslant i \leqslant n \ \text{ oraz } \ \mathrm{NWD}(i,n) = 1\}|$$

dla każdego $n \in \mathbb{N}^*$.

Lemat 3.1.9. *Niech $k \in \mathbb{N}^*$ oraz rząd $|\langle g \rangle| = n$. Element g^k jest generatorem grupy cyklicznej $\langle g \rangle$ wtedy i tylko wtedy, gdy* $\mathrm{NWD}(k,n) = 1$.

Dowód. ($\Rightarrow$) Skoro $\langle g^k \rangle = \langle g \rangle$, to $o(g^k) = o(g)$. Wtedy na podstawie twierdzenia 3.1.8 wnosimy, że $\mathrm{NWD}(k,n) = 1$.

($\Leftarrow$) Oczywiście, że $\langle g^k \rangle \subseteq \langle g \rangle$. Z twierdzenia 3.1.8 wynika, że $o(g^k) = o(g)$, a więc $\langle g^k \rangle = \langle g \rangle$. □

Zatem otrzymujemy takie

Twierdzenie 3.1.10. *Liczba parami różnych generatorów skończonej grupy cyklicznej G rzędu n jest równa $\varphi(n)$.*

Dowód. Teza wynika na podstawie lematu 3.1.9. □

Przykład 3.1.11.

Grupa cykliczna G rzędu 15 ma $\varphi(15) = 8$ parami różnych generatorów.

(2) Leonard Euler (1707–1783)

Ćwiczenia 3.1.12.

(**1**) Udowodnić, że grupa G jest cykliczna rzędu m, jeśli:
(a) $G = \mathbb{C}_m = \{z \in \mathbb{C} \mid z^m = 1\}$ oraz $m \in \mathbb{N}^*$;
(b) $G = \{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{Z}_7\}$ oraz $m = 7$;
(c) $G = \{f_{a,0} : \mathbb{Z}_{13} \to \mathbb{Z}_{13} \mid f_{a,0}(z) = az, \text{ gdzie } a \in \mathbb{Z}_{13},\ a \neq 0\}$ oraz $m = 12$;
(d) $G = U(\mathbb{Z}_5)$ oraz $m = 4$.
(**2**) Znaleźć wszystkie (parami różne) podgrupy w grupie:
(a) $\mathbb{Z}_5$;
(b) $\mathbb{Z}_4$;
(c) $\mathbb{Z}_6$;
(d) $\mathbb{Z}_{30}$;
(e) $\mathbb{Z}$.
(**3**) Niech G będzie grupą skończoną oraz $g \in G$. Udowodnić, że $G = \langle g \rangle$ jest cykliczna wtedy i tylko wtedy, gdy $|G| = |g|$.
(**4**) Sprawdzić, czy każda podgrupa skończona grupy multiplikatywnej liczb zespolonych $\mathbb{C}^*$ jest cykliczna.

Uwagi. Podstawy teorii skończonych grup abelowych założył C. Gauss w *Disquisitiones* w 1801 r., w której zwrócił uwagę, że pewne skończone grupy abelowe klas form kwadratowych nie są cykliczne.

3.2. Grupy pierwiastków zespolonych z jedności 1

■ Niech n będzie ustaloną liczbą naturalną,

$$\epsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$$

oraz

$$\mathbb{C}_n = \{\epsilon_k \mid k = 0, 1, \ldots, n-1\}.$$

Lemat 3.2.1. *Zbiór $\mathbb{C}_n = \langle\epsilon_1\rangle$ jest multiplikatywną grupą cykliczną rzędu n.*

Dowód. Oczywiście, że dla $1 \in \mathbb{C}^*$ mamy $1^n = 1$, a więc $1 \in \mathbb{C}_n$. Oprócz tego jeśli $z, w \in \mathbb{C}_n$, to $z^n = 1$ i $w^n = 1$ oraz, jako wniosek,

$$\begin{aligned}(z\cdot w)^n &= z^n\cdot w^n = 1\cdot 1 = 1,\\ (z^{-1})^n &= (z^n)^{-1} = 1^{-1} = 1.\end{aligned}$$

Na podstawie kryterium podgrupy $\mathbb{C}_n$ jest podgrupą grupy $\mathbb{C}^*$, a zatem $(\mathbb{C}_n, \cdot)$ jest grupą. W końcu z tego, że $\epsilon_k = (\epsilon_1)^k$ dla $0 \leqslant k < n$ oraz

$$1 = \epsilon_0 = (\epsilon_1)^n$$

wynika, że $\mathbb{C}_n = \langle\epsilon_1\rangle$ jest grupą cykliczną, która ma rząd n na podstawie lematu 3.1.4. □

■ Zatem $\mathbb{C}_{p^n}$ jest p-grupą cykliczną rzędu p^n dla każdej liczby pierwszej p i dowolnej liczby naturalnej n.

Wniosek 3.2.2. *Zbiór*

$$\mathbb{C}_{p^\infty} = \bigcup_{n=0}^{\infty}\mathbb{C}_{p^n}$$

jest podgrupą multiplikatywnej grupy liczb zespolonych $\mathbb{C}^*$ (tutaj p jest liczbą pierwszą).

Dowód. Ponieważ z warunku $z^{p^n} = 1$ wynika, że $z^{p^{n+1}} = 1$ dla wszystkich $n \in \mathbb{N}$, to

$$\mathbb{C}_{p^n} \subseteq \mathbb{C}_{p^{n+1}}$$

i na podstawie lematu 2.5.6 wnosimy, że $\mathbb{C}_{p^\infty}$ jest podgrupą grupy $\mathbb{C}^*$. □

■ Grupa $\mathbb{C}_{p^\infty}$ (p jest liczbą pierwszą) jest nazywana *p-grupą kwazicykliczną* (a także *grupą typu* p^∞ lub *p-grupą Prüfera*).

Twierdzenie 3.2.3. *Każda p-grupa kwazicykliczna* $\mathbb{C}_{p^\infty}$ *ma następujące własności:*

(1) $\mathbb{C}_{p^\infty}$ *jest p-grupą nieskończoną;*
(2) $\mathbb{C}_{p^\infty}$ *nie jest cykliczna;*
(3) *każda podgrupa właściwa z* $\mathbb{C}_{p^\infty}$ *jest skończona i cykliczna.*

Dowód. (1) Ćwiczenie.

(2) Jeżeli założyć, że grupa $\mathbb{C}_{p^\infty} = \langle a \rangle$ jest cykliczna z generatorem a, to znajdzie się taka liczba naturalna k, że $a \in \mathbb{C}_{p^k}$, a wtedy $\langle a \rangle \leqslant \mathbb{C}_{p^k}$, co jest sprzeczne z założeniem. Zatem grupa $\mathbb{C}_{p^\infty}$ nie jest cykliczna.

(3) Niech T będzie podgrupą właściwą z $\mathbb{C}_{p^\infty}$ oraz

$$S_T = \{s \in \mathbb{N} \mid x^{p^s} = 1, \text{ gdzie } x \in T\}.$$

Jeśli T jest nieskończona, to zbiór indeksów S_T też jest nieskończony. To oznacza, że dla każdej liczby naturalnej $n \in \mathbb{N}$ znajdzie się taka liczba naturalna $j \in S_T$, że $j > n$, a zatem $T = \mathbb{C}_{p^\infty}$, co nie jest prawdą. Wnosimy, że każda podgrupa właściwa T z $\mathbb{C}_{p^\infty}$ jest skończona, co powoduje, że w zbiorze S_T znajdzie się największa liczba $k = k(T) \in \mathbb{N}$ taka, że

$$t^{p^k} = 1$$

dla wszystkich elementów $t \in T$. Wtedy

$$T \leqslant \mathbb{C}_{p^k}$$

i na podstawie twierdzenia 3.1.3 podgrupa T jest cykliczna. □

Ćwiczenia 3.2.4.

(**1**) Udowodnić, że H jest grupą cykliczną, jeśli:
(*a*) $H = \mathbb{C}_m = \{z \in \mathbb{C} \mid z^m = 1\}$ oraz $m \in \mathbb{N}$;
(*b*) $H = U(\mathbb{Z}_5)$ oraz $m = 4$;
(*c*) $H = \{f_{a,0} : \mathbb{Z}_{13} \to \mathbb{Z}_{13} \mid f_{a,0}(z) = az,\ a \in \mathbb{Z}_{13},\ a \neq 0\}$ oraz $m = 12$;
(*d*) $H = \{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{Z}_7\}$ oraz $m = 7$.

(**2**) Sprawdzić, czy grupa G jest cykliczna, jeśli:
(*a*) $G = V_4$;

(b) $G = U(\mathbb{Z}_7)$;
(c) $G = SL_2(\mathbb{Z}_2)$;
(d) $G = \{f_{a,b} : \mathbb{Z}_7 \to \mathbb{Z}_7 \mid f_{a,b}(z) = az + b,\ a \neq 0,\ a, b \in \mathbb{Z}_7\}$;
(e) $G = Q_8$;
(f) $G = D_8$;
(g) $G = \mathbb{Z}[i]$.
(**3**) Udowodnić, że każda podgrupa skończona grupy $\mathbb{C}^*$ jest cykliczna.
(**4**) Niech G będzie grupą skończoną oraz $g \in G$. Udowodnić, że $G = \langle g \rangle$ wtedy i tylko wtedy, gdy $|G| = |g|$.

Uwagi. Na początku XVIII w. R. Cotes[3] i A. de Moivre[4] powiązali rozwiązania równania $x^n - 1 = 0$ z podziałem koła na n równych części. Grupa Prüfera została zbudowana w pracy doktorskiej E. Prüfera[5] w 1920 r.

[3] Roger Cotes (1682–1716)
[4] Abraham de Moivre (1667–1754)
[5] Ernst Paul Heinz Prüfer (1896–1934)

3.3. Warstwy. Twierdzenie Lagrange'a.

■ Niech $(G, \cdot)$ będzie grupą, a H będzie jej podgrupą. Jeśli $g \in G$, to zbiór

$$gH = \{g \cdot h \mid h \in H\}$$

jest nazywany *warstwą lewostronną grupy G względem jej podgrupy H* z *reprezentantem g*. Podobnie możemy zdefiniować *warstwę prawostronną*

$$Hg = \{h \cdot g \mid h \in H\}$$

grupy G względem H z *reprezentantem g*.

■ Warstwę gH (odpowiednio Hg) będziemy krótko oznaczać przez

$$\overline{g}.$$

■ Jeśli $(G, +)$ jest grupą addytywną, to warstwa

$$\overline{g} = g + H = \{g + h \mid h \in H\}$$

jest lewostronna, a

$$\overline{g} = H + g = \{h + g \mid h \in H\}$$

jest warstwą prawostronną grupy G względem jej podgrupy H z reprezentantem g.

■ Przypomnijmy, że zbiory X i Y są nazywane *równolicznymi*, jeśli istnieje pewne odwzorowanie bijektywne postaci $X \to Y$.

Lemat 3.3.1. *Niech G będzie grupą. Wtedy dla dowolnych elementów $g_1, g_2 \in G$ warstwy lewostronne g_1H i g_2H (odpowiednio warstwy prawostronne Hg_1 i Hg_2) są równoliczne.*

Dowód. Niech e będzie elementem neutralnym grupy G. Rozpatrzmy tylko warstwy lewostronne (dla warstw prawostronnych dowód jest prawo-symetryczny). Reguła $\varphi : g_1H \to g_2H$ taka, że

$$\varphi(g_1 \cdot h) = g_2 \cdot h$$

dla wszystkich $h \in H$ określa odwzorowanie φ. Niech $g_1 \cdot h_1$ oraz $g_1 \cdot h_2$ będą elementami z g_1H, gdzie $h_1, h_2 \in H$. Jeśli $\varphi(g_1 \cdot h_1) = \varphi(g_1 \cdot h_2)$, to $g_2 \cdot h_1 = g_2 \cdot h_2$, a stąd

$$h_1 = e \cdot h_1 = (g_2^{-1} \cdot g_2) \cdot h_1 = g_2^{-1} \cdot (g_2 \cdot h_1) = g_2^{-1} \cdot (g_2 \cdot h_2) = \\ = (g_2^{-1} \cdot g_2) \cdot h_2 = e \cdot h_2 = h_2.$$

Z tych rozumowań wynika, że φ jest iniekcją. Teraz jeśli $g_2 \cdot h$ jest dowolnym elementem z g_2H, gdzie $h \in H$, to $g_2 \cdot h = \varphi(g_1 \cdot h)$ dla pewnego $g_1 \cdot h \in g_1H$, czyli φ jest suriekcją.

Zatem odwzorowanie $\varphi : g_1H \to g_2H$ jest bijektywne. □

Lemat 3.3.2. *Niech G będzie grupą, H będzie jej podgrupą oraz $g_1, g_2 \in G$. Wtedy warstwy lewostronne g_1H i g_2H (odpowiednio warstwy prawostronne Hg_1 i Hg_2) są albo równe, albo rozłączne.*

Dowód. Zbudujmy dowód tylko dla warstw lewostronnych. Załóżmy, że

$$g_1H \bigcap g_2H \neq \varnothing$$

i z jest dowolnym elementem z tego przecięcia. Wtedy

$$z = g_1 \cdot h_1 = g_2 \cdot h_2$$

dla pewnych elementów $h_1, h_2 \in H$, a stąd

$$g_1 = (g_1 \cdot h_1) \cdot h_1^{-1} = (g_2 \cdot h_2) \cdot h_1^{-1} = g_2 \cdot (h_2 \cdot h_1^{-1}) \in g_2H.$$

Także mamy

$$g_1 \cdot h = g_2 \cdot (h_2 \cdot h_1^{-1} \cdot h) \in g_2H$$

dla każdego $h \in H$, a to implikuje, że $g_1H \subseteq g_2H$. Podobnie otrzymujemy

$$g_2 = (g_2 \cdot h_2) \cdot h_2^{-1} = (g_1 \cdot h_1) \cdot h_2^{-1} = g_1 \cdot (h_1 \cdot h_2^{-1}) \in g_1H.$$

Zatem

$$g_2 \cdot h = g_1 \cdot (h_1 \cdot h_2^{-1} \cdot h) \in g_1H,$$

czyli $g_2H \subseteq g_1H$, a więc $g_1H = g_2H$. □

Twierdzenie 3.3.3. *Warstwy lewostronne* (*odpowiednio prawostronne*) *grupy G względem jej podgrupy H tworzą rozbicie zbioru G.*

Dowód. Jak zwykle rozpatrzmy tylko warstwy lewostronne. Niech e będzie elementem neutralnym grupy G oraz $X = \{gH \mid g \in G\}$ będzie klasą wszystkich warstw lewostronnych grupy G względem jej podgrupy H. Ponieważ

$$g = g \cdot e \in gH,$$

to $gH \neq \varnothing$. Oczywiście $gH \subseteq G$ dla każdego elementu $g \in G$. Za lematem 3.3.2 różne warstwy lewostronne są rozłączne. Na koniec dla elementu $g \in G$ mamy $g \in gH$, a więc $G \subseteq \cup_{g \in G} gH$. Wyciągamy stąd wniosek, że

$$\bigcup_{g \in G} gH = G.$$

Zatem X jest rozbiciem zbioru G. □

Lemat 3.3.4. *Niech G będzie grupą, H będzie jej podgrupą oraz $g_1, g_2 \in G$. Wtedy:*

(1) *$g_1H = g_2H$ wtedy i tylko wtedy, gdy $g_1H \subseteq g_2H$;*
(2) *$g_1H = g_2H$ wtedy i tylko wtedy, gdy $g_1H \supseteq g_2H$;*
(3) *$Hg_1 = Hg_2$ wtedy i tylko wtedy, gdy $Hg_1 \subseteq Hg_2$;*
(4) *$Hg_1 = Hg_2$ wtedy i tylko wtedy, gdy $Hg_1 \supseteq Hg_2$.*

Dowód. (1) Niech e będzie elementem neutralnym grupy G.

($\Rightarrow$) Jeśli $g_1H = g_2H$, to biorąc pod uwagę definicję równości dwóch zbiorów, mamy

$$g_1H \subseteq g_2H.$$

($\Leftarrow$) Załóżmy, że zachodzi $g_1H \subseteq g_2H$. Skoro $g_1 = g_1 \cdot e \in g_1H$, to $g_1 = g_2 \cdot h$ dla pewnego $h \in H$, a więc $g_2 = g_1 \cdot h^{-1} \in g_1H$. Lecz wówczas

$$g_2 \cdot w = (g_1 \cdot h^{-1}) \cdot w = g_1 \cdot (h^{-1} \cdot w) \in g_1H$$

dla wszystkich elementów $w \in H$. A to oznacza, że $g_2H \subseteq g_1H$. Zatem $g_2H = g_1H$.

(2) Dowód jest podobny do (1), a dowody (3) i (4) są prawostronnymi analogami odpowiednio dowodów (1) i (2) (zostawiamy Czytelnikowi do samodzielnego udowodnienia). □

Bardziej użyteczny jest taki

Lemat 3.3.5 (kryterium równości warstw)**.** *Niech G będzie grupą, a H będzie jej podgrupą oraz $g_1, g_2, g \in G$. Wtedy:*
(1) *$g_1H = g_2H$ w tym i tylko tym przypadku, gdy ${g_1}^{-1}g_2 \in H$;*
(1′) *$g_1H = g_2H$ w tym i tylko tym przypadku, gdy ${g_2}^{-1}g_1 \in H$;*
(2) *$Hg_1 = Hg_2$ w tym i tylko tym przypadku, gdy $g_1{g_2}^{-1} \in H$;*
(2′) *$Hg_1 = Hg_2$ w tym i tylko tym przypadku, gdy $g_2{g_1}^{-1} \in H$;*
(3) *$gH = H$ w tym i tylko tym przypadku, gdy $g \in H$;*
(3′) *$Hg = H$ w tym i tylko tym przypadku, gdy $g \in H$.*

Dowód. (1′) Niech e będzie elementem neutralnym grupy G.

($\Rightarrow$) Załóżmy, że $g_1H = g_2H$. Ponieważ $g_1 = g_1 \cdot e \in g_1H$, to $g_1 \in g_2H$, a zatem $g_1 = g_2 \cdot h_0$ dla pewnego $h_0 \in H$, skąd otrzymujemy

$$g_2^{-1} \cdot g_1 = g_2^{-1} \cdot (g_2 \cdot h_0) = (g_2^{-1} \cdot g_2) \cdot h_0 = e \cdot h_0 = h_0 \in H.$$

($\Leftarrow$) Teraz załóżmy, że $g_2^{-1} \cdot g_1 \in H$, czyli $g_2^{-1} \cdot g_1 = h$ dla pewnego $h \in H$, i na tej podstawie

$$g_1 = e \cdot g_1 = (g_2 \cdot g_2^{-1}) \cdot g_1 = g_2 \cdot (g_2^{-1} \cdot g_1) = g_2 \cdot h \in g_2H.$$

Wtedy

$$g_1 \cdot t = (g_2 \cdot h) \cdot t = g_2 \cdot (h \cdot t) \in g_2H$$

dla wszystkich elementów $t \in H$, czyli $g_1H \subseteq g_2H$. Za lematem 3.3.4 mamy $g_1H = g_2H$.

Dowody reszty własności są podobne.

□

Lemat 3.3.6. *Jeśli H jest podgrupą grupy, to zbiory warstw lewostronnych $\{gH \mid g \in G\}$ oraz prawostronnych $\{Hg \mid g \in G\}$ są równoliczne.*

Dowód. (Szkic) Rzeczywiście

$$\varphi : \{gH \mid g \in G\} \ni gH \mapsto Hg \in \{Hg \mid g \in G\}$$

jest bijekcją. □

Na podstawie udowodnionych wyżej własności możemy wprowadzić taką definicję.

■ Moc rozbicia

$$\{gH \mid g \in G\}$$

grupy G na warstwy lewostronne względem jej podgrupy H jest nazywana *indeksem* podgrupy H w grupie G (i oznaczana przez $|G : H|$).

Twierdzenie 3.3.7 (Lagrange'a)**.** *Jeśli G jest grupą skończoną, a H jej podgrupą, to*

$$|G| = |H||G : H|$$

(*czyli rząd grupy skończonej G jest iloczynem rzędu $|H|$ podgrupy H i indeksu $|G : H|$ podgrupy H w grupie G*).

Dowód. Niech e będzie elementem neutralnym grupy G oraz $s = |H|$. Ponieważ G jest grupą skończoną, to jej rozbicie $X = \{gH \mid g \in G\}$ składa się ze skończonej liczby (parami różnych) elementów, czyli

$$X = \{eH, g_2H, \ldots, g_kH\},$$

co powoduje, że

$$G = eH \bigcup g_2H \bigcup \cdots \bigcup g_kH$$

dla pewnych $g_2, \ldots, g_k$. Zatem $|G : H| = k$. Za lematem 3.3.1 wszystkie warstwy lewostronne eH, $g_2H, \ldots, g_kH$ są parami równoliczne. Biorąc pod uwagę, że G jest grupą skończoną oraz $H = eH$, wnosimy, że każda z tych warstw lewostronnych składa się z s elementów. Skoro według lematu 3.3.2 warstwy

$$eH, g_2H, \ldots, g_kH$$

są parami rozłączne, to

$$|G| = \operatorname{card}(eH) + \sum_{i=2}^{k} \operatorname{card}(g_iH) = sk = |H||G : H|.$$

□

Wniosek 3.3.8. *Zachodzą następujące własności:*
(*a*) *rząd* $|H|$ *każdej podgrupy* H *dzieli rząd* $|G|$ *grupy skończonej* G*;*
(*b*) *rząd* $o(g)$ *każdego elementu* g *grupy skończonej* G *jest dzielnikiem jej rzędu* $|G|$*.*

Dowód. (*a*) Wynika z twierdzenia 3.3.7.
(*b*) Ponieważ na podstawie lematu 3.1.4 mamy

$$o(g) = |\langle g \rangle|$$

i w wyniku twierdzenia 3.3.7 liczba $|\langle g \rangle|$ dzieli $|G|$, to $o(g)$ jest dzielnikiem rzędu $|G|$ całej grupy G. □

Wniosek 3.3.9. *Grupa* G*, rzędem której jest liczba pierwsza* p*, jest cykliczna.*

Dowód. Skoro $|G| = p > 1$, to grupa G posiada pewien element g, który różni się od elementu neutralnego e grupy G. Wtedy $|\langle g \rangle| > 1$. Na podstawie wniosku 3.3.8 otrzymujemy, że $|\langle g \rangle|$ jest dzielnikiem liczby pierwszej p, a więc

$$|\langle g \rangle| = p = |G|.$$

Stąd łatwo wynika, że $\langle g \rangle = G$ jest cykliczna. □

Przykłady 3.3.10.

(**1**) Mamy $\mathbb{C}_4 = \langle i \rangle$ oraz $o(i) = |\mathbb{C}_4| = 4$.

(**2**) Twierdzenie odwrotne do wniosku 3.3.8 w ogólnym przypadku nie zachodzi. To znaczy, że znajdzie się grupa skończona G i dzielnik s jej rzędu $|G|$ taki, że grupa G nie posiada podgrup rzędu s (zostawiamy Czytelnikowi samodzielne znalezienie przykładu takiej grupy).

(**3**) Twierdzenie odwrotne do wniosku 3.3.9 w ogólnym przypadku też się nie sprawdza. Istotnie, grupa klas reszt $\mathbb{Z}_8$ jest cykliczna, lecz jej rząd 8 nie jest liczbą pierwszą.

(**4**) Znajdźmy warstwy lewostronne grupy $\mathbb{S}_3$ względem jej podgrupy $\langle (23) \rangle$. Skoro $\langle (23) \rangle = \{e, (23)\}$, to:
- $e\langle (23) \rangle = \{ee, e(23)\} = \{e, (23)\} = \langle (23) \rangle$;
- $(23)\langle (23) \rangle = \{(23)e, (23)(23)\} = \{(23), e\} = \langle (23) \rangle$;
- $(12)\langle (23) \rangle = \{(12)e, (12)(23)\} = \{(12), (123)\}$;
- $(13)\langle (23) \rangle = \{(13)e, (13)(23)\} = \{(13), (132)\}$;
- $(123)\langle (23) \rangle = \{(123)e, (123)(23)\} = \{(123), (12)\}$;
- $(132)\langle (23) \rangle = \{(132)e, (132)(23)\} = \{(132), (13)\}$.

Zatem mamy trzy parami różne warstwy lewostronne

$$\{e, (23)\}, \{(13), (132)\}, \{(123), (12)\}$$

grupy $\mathbb{S}_3$ względem podgrupy $\langle(23)\rangle$, a więc jej indeks $|\mathbb{S}_3 : \langle(23)\rangle| = 3$.

Zostawiamy Czytelnikowi znalezienie wszystkich warstw prawostronnych grupy $\mathbb{S}_3$ względem jej podgrupy $\langle(123)\rangle$.

Ćwiczenia 3.3.11.

(**1**) Udowodnić, że:
(a) jeśli H_1 oraz H_2 są podgrupami grupy G i $gH_1 = tH_2$ dla pewnych elementów $g, t \in G$, to $H_1 = H_2$;
(b) jeśli H jest podgrupą grupy G i $x \in G$, to zachodzą implikacje

$$y \in xH \Rightarrow x^{-1}y \in H \Rightarrow y = xh \text{ dla pewnego } h \in H.$$

(**2**) Znaleźć wszystkie warstwy lewostronne i prawostronne grupy G względem podgrupy H i znaleźć indeks $|G : H|$, jeśli:
(a) $G = \mathbb{Z}$ oraz $H = 2\mathbb{Z}$;
(b) $G = \mathbb{Z}_{12}$ oraz $H = \{\overline{0}, \overline{6}\}$;
(c) $G = \mathbb{Z}_8$ oraz $H = \{\overline{0}, \overline{2}, \overline{4}\}$;
(d) $G = \mathbb{S}_3$ oraz $H = \{e, (12)\}$;
(e) $G = GL_2(\mathbb{Z}_3)$ oraz $H = SL_2(\mathbb{Z}_3)$;
(f) $G = \{f_{a,b} : \mathbb{Z}_5 \to \mathbb{Z}_5 \mid f_{a,b}(x) = aX + b,\ a \neq 0,\ a, b \in \mathbb{Z}_5\}$ oraz $H = \{f_{a,0} \in G \mid a \in \mathbb{Z}_5^*\}$;
(g) $G = \mathbb{C}^*$ oraz $H = \mathbb{S}^1$;
(h) $G = \mathbb{R}^*$ oraz $H = \mathbb{R}_+$;
(i) $G = \mathbb{C}^*$ oraz $H = \mathbb{R}$;
(j) $G = GL_2(\mathbb{R})$ oraz $H = \{\lambda I \mid \lambda \in \mathbb{R}^*\}$, gdzie $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Uwagi. Pewne elementy twierdzenia 3.3.7 J. Lagrange opublikował w 1770 r. w pracy poświęconej poszukiwaniu pierwiastków wielomianów stopni $\geqslant 5$, które później w 1802 r. sprecyzował i rozszerzył matematyk włoski P. Abbati[6]. W postaci sformułowanej wyżej pierwszy raz to twierdzenie zostało opublikowane w podręczniku matematyka francuskiego J. Serreta[7] w 1866 r. Termin „warstwa (=coset)" po raz pierwszy znalazł się w 1910 r. w pracy matematyka amerykańskiego G. Millera[8], chociaż badanie pewnych warstw zawiera się w jeszcze jednej pracy J. Lagrange'a z 1770 r.

[6] Pietro Abbati (1768–1842)
[7] Joseph Alfred Serret (1819–1885)
[8] George Abram Miller (1863–1951)

3.4. Podgrupy normalne

■ Niech $(G, \cdot)$ będzie grupą. Mówimy, że elementy g i h są *sprzężone* w grupie G, jeśli

$$g = t^{-1} \cdot h \cdot t$$

dla pewnego elementu $t \in G$. Podgrupa H grupy G jest nazywana *normalną* w G (co zapisujemy jako $H \lhd G$), jeśli

$$gH = Hg$$

dla każdego elementu $g \in G$.

Twierdzenie 3.4.1 (kryterium normalności podgrupy)**.** *Równoważne są następujące własności:*

(1) *H jest podgrupą normalną grupy G;*
(2) *dla dowolnych sprzężonych w grupie G elementów g_1 i g_2 z warunku $g_1 \in H$ wynika, że $g_2 \in H$;*
(3) *$g^{-1}Hg = H$ dla każdego elementu $g \in G$;*
(4) *$g^{-1}Hg \subseteq H$ dla dowolnego elementu $g \in G$;*
(5) *$g^{-1}hg \in H$ dla dowolnych $g \in G$ i $h \in H$.*

Dowód. (1) $\Rightarrow$ (2) Niech $H \lhd G$ oraz g_1, g_2 będą takimi elementami z G, że $g_1 = t^{-1}g_2t$ dla pewnego $t \in G$. Jeśli $g_1 \in H$, to $g_2 \in tHt^{-1}$. Skoro $tH = Ht$, to

$$g_2 \in (tH)t^{-1} = (Ht)t^{-1} = H.$$

(2) $\Rightarrow$ (3) Mamy dowolne elementy $h \in H$ i $g \in G$. Ponieważ h i $g^{-1}hg$ są sprzężone w grupie G oraz $h \in H$, to na mocy założenia $g^{-1}hg \in H$, czyli $g^{-1}Hg \subseteq H$. Podobnie h i $(g^{-1})^{-1}hg^{-1}$ są sprzężone w G, a więc $(g^{-1})^{-1}Hg^{-1} \subseteq H$. Wtedy

$$H = g^{-1} \cdot (g^{-1})^{-1}Hg^{-1} \cdot g \subseteq g^{-1}Hg,$$

co powoduje, że $g^{-1}Hg = H$.

(3) $\Rightarrow$ (4) Ćwiczenie.

(4) $\Rightarrow$ (5) Przepisując $g^{-1}Hg \subseteq H$ w języku elementów, dla dowolnych $g \in G$, $h \in H$, dostajemy, że $g^{-1}hg \in H$.

(5) $\Rightarrow$ (1) Istotnie, skoro $g^{-1}hg \in H$ dla dowolnych $g \in G$ oraz $h \in H$, to $Hg \subseteq gH$. Według lematu 3.3.4 otrzymujemy, że $Hg = gH$. □

Przykłady 3.4.2.

(**1**) Oczywiście, że w grupie abelowej G dwa elementy g i h są sprzężone wtedy i tylko wtedy, gdy $g = h$.

(**2**) Cykle (12) i (23) są sprzężone w grupie symetrycznej $\mathbb{S}_3$, bo

$$(23) = (132)^{-1}(12)(132) = (123)(12)(132).$$

(**3**) Macierze

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \text{ oraz } \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

są sprzężone w grupie $GL_2(\mathbb{R})$, ponieważ

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}.$$

(**4**) Każda podgrupa H grupy abelowej G jest normalna, gdyż

$$aH = \{a \cdot h = h \cdot a \mid h \in H\} = Ha$$

dla każdego elementu $a \in G$.

(**5**) Grupa alternująca $\mathbb{A}_n$ stopnia n jest podgrupą normalną grupy symetrycznej $\mathbb{S}_n$. W rzeczy samej, $\mathbb{A}_n$ jest podgrupą grupy $\mathbb{S}_n$ (patrz przykład 2.5.4 (6)). Teraz o dowolnym elemencie $\sigma \in \mathbb{A}_n$ wiemy, że $\sigma = \tau_1 \cdots \tau_{2k}$ jest iloczynem parzystej liczby transpozycji $\tau_1, \ldots, \tau_{2k}$ (gdzie $k \in \mathbb{N}^*$). Niech θ będzie dowolnym elementem z $\mathbb{S}_n$. Skoro $\theta = \mu_1 \cdots \mu_l$ jest iloczynem pewnych transpozycji $\mu_1, \ldots, \mu_l$ oraz $\mu_i^{-1} = \mu_i$, to

$$\theta^{-1}\sigma\theta = \mu_l^{-1} \cdots \mu_1^{-1}\tau_1 \cdots \tau_{2k}\mu_1 \cdots \mu_l = \mu_l \cdots \mu_1\tau_1 \cdots \tau_{2k}\mu_1 \cdots \mu_l$$

jest iloczynem parzystej liczby transpozycji, a więc $\theta^{-1}\sigma\theta \in \mathbb{A}_n$. Zatem $\mathbb{A}_n \lhd \mathbb{S}_n$.

(**6**) Udowodnijmy, że $SL_n(\mathbb{R}) \lhd GL_n(\mathbb{R})$. Rzeczywiście, jak wiadomo z przykładu 2.5.4 (5), $SL_n(\mathbb{R})$ jest podgrupą w $GL_n(\mathbb{R})$. Jeśli $S \in SL_n(\mathbb{R})$ i $A \in GL_n(\mathbb{R})$, to $\det S = 1$ i wtedy, stosując wzór Cauchy'ego-Bineta,

$$\det(A^{-1}SA) = \frac{1}{\det A} \cdot \det S \cdot \det A = \frac{1}{\det A} \cdot 1 \cdot \det A = 1,$$

a zatem $A^{-1}SA \in SL_n(\mathbb{R})$.

(**7**) Niech

$$M = \{X \in GL_n(\mathbb{Q}) \mid \det X > 0\}.$$

Wtedy podgrupa $M \lhd GL_n(\mathbb{Q})$ jest normalna, bo dla dowolnych macierzy $B \in M$ i $Y \in GL_n(\mathbb{Q})$ mamy

$$\det(Y^{-1}BY) = \frac{\det B \cdot \det Y}{\det Y} = \det B > 0,$$

czyli $Y^{-1}BY \in M$.

(**8**) Jeśli $|G:H|=2$, to $H \lhd G$. W rzeczy samej, niech $g \in G \setminus H$. Wtedy za lematem 3.3.5 mamy $gH \neq H$ oraz $H \neq Hg$. Zatem gH, H są różnymi warstwami lewostronnymi, a Hg, H są różnymi warstwami prawostronnymi grupy G względem podgrupy H. Założyliśmy, że $|G:H|=2$, a więc

$$G = gH \bigcup H = Hg \bigcup H.$$

Oprócz tego

$$H \bigcap gH = \varnothing = H \bigcap Hg.$$

Zatem zbiory Hg i gH są równe. To znaczy, że $H \lhd G$.

(**9**) Podgrupa $\langle(12)\rangle$ nie jest normalna w grupie symetrycznej $\mathbb{S}_4$. Istotnie

$$\langle(12)\rangle = \{(12) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix},\ (12)^2 = e\} \text{ oraz } (13)^{-1}(12)(13) =$$
$$= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} = (23) \notin \langle(12)\rangle.$$

Z tego wynika, że $\langle(12)\rangle \not\lhd \mathbb{S}_4$.

(**10**) Grupa V_4 jest abelową podgrupą normalną w grupie permutacji $\mathbb{S}_4$ (sprawdzić samodzielnie).

Ćwiczenia 3.4.3.

(**1**) Udowodnić, że sprzężoność elementów w grupie G jest relacją równoważności w zbiorze G.

(**2**) Sprawdzić, czy podgrupa H jest normalną w grupie $G = \{\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{C}^*, b \in \mathbb{C}\}$, jeśli:

(a) $H = \{\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{C}^*\}$;

(b) $H = \{\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{R}\}$.

(**3**) Udowodnić, że H jest podgrupą normalną w grupie G, jeśli:

(a) $G = GL_n(\mathbb{R})$ oraz $H = \{A \in G \mid |\det A| = 1\}$;

(b) $G = GL_n(\mathbb{R})$ oraz $H = \{A \in G \mid \det A > 0\}$;

(c) $G = \{f_{a,b} : \mathbb{R} \to \mathbb{R} \mid f_{a,b}(X) = aX + b,\ a, b \in \mathbb{R},\ a \neq 0\}$ oraz $H = \{f_{1,b} \mid b \in \mathbb{R}\}$.

(**4**) Udowodnić, że jeśli H_1 i H_2 są podgrupami normalnymi w grupie G, to $H_1 \cap H_2 \lhd G$ i $H_1 \cdot H_2 \lhd G$.

(**5**) Znaleźć wszystkie podgrupy normalne w grupie:

(a) $\mathbb{S}_3$;

(b) $\mathbb{A}_4$;

(c) $\mathbb{S}_4$;

(d) $L(\mathbb{Z}_7) = \{f_{a,b} : \mathbb{Z}_7 \to \mathbb{Z}_7 \mid f_{a,b}(X) = aX + b,\ a, b \in \mathbb{Z}_7,\ a \neq 0\}$.

(**6**) Niech $S(\mathbb{N})$ będzie zbiorem wszystkich odwzorowań bijektywnych postaci $f : \mathbb{N} \to \mathbb{N}$. Sprawdzić, czy H jest podgrupą normalną w grupie $\mathbb{S}(\mathbb{N})$, jeśli:

(a) $H = \{\sigma \in \mathbb{S}(\mathbb{N}) \mid \sigma(2) = 2\}$;

(b) $H = \{\sigma \in \mathbb{S}(\mathbb{N}) \mid \sigma(a) = a \text{ dla } a \in \mathbb{N} \setminus \{2\}\}$.

Uwagi. C. Gauss wykorzystywał własności podgrup normalnych i grup ilorazowych w swoich badaniach w 1830 r. Termin „podgrupa normalna" należy do E. Galoisa.

3.5. Homomorfizmy grup

■ *Homomorfizmem* grupy G (z działaniem „$\cdot$”) w grupę F (z działaniem „$*$”) jest nazywane odwzorowanie $f : G \to F$ takie, że

$$f(g \cdot h) = f(g) * f(h) \tag{3.2}$$

dla dowolnych elementów $g, h \in G$.

■ Dalej e_G jest elementem neutralnym grupy G, a e_F jest elementem neutralnym grupy F. *Jądrem* $\operatorname{Ker} f$ homomorfizmu $f : G \to F$ jest nazywany zbiór takich elementów $g \in G$, że

$$f(g) = e_F.$$

Zbiór

$$\{f(x) \mid x \in G\}$$

jest nazywany *obrazem* homomorfizmu f i oznaczany przez $\operatorname{Im} f$. Oprócz tego przez

$$\operatorname{Hom}(G, F)$$

będziemy oznaczać zbiór wszystkich homomorfizmów z grupy G w grupę F.

■ Odwzorowanie $f : G \to F$ jest nazywane:

- *monomorfizmem grup*, jeśli f jest odwzorowaniem iniektywnym oraz zachodzi warunek (3.2) dla dowolnych elementów $g, h \in G$;
- *epimorfizmem grup*, jeśli f jest odwzorowaniem suriektywnym oraz zachodzi warunek (3.2) dla dowolnych elementów $g, h \in G$;
- *izomorfizmem grup*, jeśli f jest odwzorowaniem bijektywnym oraz zachodzi warunek (3.2) dla dowolnych elementów $g, h \in G$.

■ Grupy G i F są nazywane *izomorficznymi*, jeśli istnieje pewien izomorfizm postaci $f : G \to F$ (co oznaczamy przez $G \cong F$ lub $G \cong^f F$). Jeśli istnieje pewien epimorfizm postaci $f : G \to F$, to grupa F jest nazywana *obrazem homomorficznym* grupy G.

■ Homomorfizm postaci $f : G \to G$ jest nazywany *endomorfizmem*, a izomorfizm postaci $f : G \to G$ – *automorfizmem* grupy G. Przez $\operatorname{Aut} G$ oznaczamy zbiór wszystkich automorfizmów grupy G.

Lemat 3.5.1. *Jeśli odwzorowanie $f : G \to F$ jest izomorfizmem grup, to odwzorowanie odwrotne $f^{-1} : F \to G$ również jest izomorfizmem grup.*

Dowód. Niech $a \in G$ oraz $b \in F$. Przypomnijmy, że odwzorowanie odwrotne $f^{-1} : F \to G$ istnieje, jest bijektywne, i jeśli $f(a) = b$, to $f^{-1}(b) = a$. Jeśli $b_1, b_2 \in F$ są dowolnymi elementami, to w wyniku suriektywności odwzorowania f znajdą się takie $a_1, a_2 \in G$, że

$$f(a_1) = b_1, \ f(a_2) = b_2.$$

Wtedy

$$a_1 = f^{-1}(b_1), \ a_2 = f^{-1}(b_2)$$

oraz

$$\begin{gathered} f^{-1}(b_1 * b_2) = f^{-1}(f(a_1) * f(a_2)) = \\ = f^{-1}(f(a_1 \cdot a_2)) = a_1 \cdot a_2 = f^{-1}(b_1) \cdot f^{-1}(b_2). \end{gathered}$$

Zatem f^{-1} jest izomorfizmem grup. □

Twierdzenie 3.5.2. *Niech G będzie grupą (względem działania „$\cdot$”) z elementem neutralnym e_G, a F będzie grupą (względem działania „$*$”) z elementem neutralnym e_F. Dla homomorfizmu grup $f : G \to F$ są spełnione następujące własności:*

(1) $f(e_G) = e_F$;
(2) $f(x^{-1}) = f(x)^{-1}$ *dla każdego elementu* $x \in G$;
(3) $f(x^m) = f(x)^m$ *dla każdego* $x \in G$ *i dla dowolnej liczby całkowitej* m;
(4) $o(f(x)) \leqslant o(x)$ *dla każdego* $x \in G$;
(5) *rząd* $o(f(x))$ *jest dzielnikiem rzędu* $o(x)$ *dla każdego elementu* $x \in G$;
(6) *jeśli f jest izomorfizmem, to* $o(f(x)) = o(x)$ *dla każdego elementu* $x \in G$;
(7) $\operatorname{Ker} f \lhd G$;
(8) $\operatorname{Im} f$ *jest podgrupą w* F;
(9) *jeśli* $ab = ba$ *dla pewnych elementów* $a, b \in G$, *to*

$$f(a) * f(b) = f(b) * f(a);$$

(10) *f jest monomorfizmem grup w tym i tylko tym przypadku, gdy* $\operatorname{Ker} f = \{e_G\}$*;*
(11) *f jest epimorfizmem grup w tym i tylko tym przypadku, gdy* $\operatorname{Im} f = F$*;*
(12) *f jest izomorfizmem grup w tym i tylko tym przypadku, gdy* $\operatorname{Ker} f = \{e_G\}$ *oraz* $\operatorname{Im} f = F$.

Dowód. (1) Istotnie

$$e_F * f(e_G) = f(e_G) = f(e_G \cdot e_G) = f(e_G) * f(e_G),$$

a stąd na podstawie prawa skracania w grupie $(G, \cdot)$ otrzymujemy

$$e_F = f(e_G).$$

(2) Ponieważ dla dowolnego $x \in G$ zachodzą równości

$$f(x) * f(x)^{-1} = e_F = f(e_G) = f(x \cdot x^{-1}) = f(x) * f(x^{-1}),$$

to, skracając, dostajemy $f(x)^{-1} = f(x^{-1})$.
(3) Zostawiamy Czytelnikowi jako ćwiczenie.
(4) Niech $n = o(x)$. Wtedy $x^n = e_G$ oraz

$$f(x)^n = \underbrace{f(x) * \cdots * f(x)}_{n \text{ razy}} = f(\underbrace{x \cdots x}_{n \text{ razy}}) = f(x^n) = f(e_G) = e_F,$$

a zatem $o(f(x)) \leqslant n$.

(5) Załóżmy, że $m = o(f(x))$ i $n = o(x)$. Z udowodnionego w części (4) wynika, że $m \leqslant n$. Na mocy twierdzenia o dzieleniu z resztą znajdą się takie liczby całkowite q i r, że $n = mq + r$ oraz $0 \leqslant r < m$. Wtedy

$$e_F = f(x^n) = f(x)^n = f(x)^{mq+r} = (f(x)^m)^q (f(x))^r = f(x)^r,$$

a to jest możliwe pod warunkiem, że $r = 0$.

(6) Niech f będzie izomorfizmem grup, $n = o(x)$ oraz $m = o(f(x))$. Na podstawie własności (5) wnosimy, że $m \mid n$. Natomiast

$$e_F = f(x)^m = f(x^m).$$

Wtedy w wyniku lematu 3.5.1 i udowodnionej własności (1) otrzymujemy, że $e_G = f^{-1}(e_F)$ i stąd, jako wniosek,

$$e_G = f^{-1}(e_F) = f^{-1}(f(x^m)) = x^m.$$

Zatem $m \mid n$. W końcu mamy $m \mid n$ oraz $n \mid m$, co implikuje, że $m = n$.

(7) Skoro $f(e_G) = e_F$, to $e_G \in \operatorname{Ker} f$, a zatem $\operatorname{Ker} f \neq \varnothing$. Jeśli $x, y \in \operatorname{Ker} f$, to $f(x) = e_F$, $f(y) = e_F$, a stąd

$$f(x \cdot y) = f(x) * f(y) = e_F * e_F = e_F,$$

czyli $x \cdot y \in \operatorname{Ker} f$. Oprócz tego warunek $f(x) = e_F$ i udowodniona wyżej własność (2) implikują, że

$$e_F = e_F^{-1} = f(x)^{-1} = f(x^{-1}),$$

czyli $x^{-1} \in \operatorname{Ker} f$. Na podstawie kryterium podgrupy wnosimy, że $\operatorname{Ker} f$ jest podgrupą w G. Jeśli teraz $x \in \operatorname{Ker} f$ i $g \in G$, to

$$f(g^{-1} \cdot x \cdot g) = f(g^{-1}) * f(x) * f(g) = f(g)^{-1} * e_F * f(g) = e_F,$$

a zatem $g^{-1} \cdot x \cdot g \in \operatorname{Ker} f$. To oznacza, że $\operatorname{Ker} f \lhd G$.

(8) Niech x i y będą dowolnymi elementami z $\operatorname{Im} f$. Wtedy znajdą się takie elementy $a, b \in G$, że $x = f(a)$ i $y = f(b)$ oraz

$$\begin{array}{ccccc} x * y & = & f(a) * f(b) & = & f(a \cdot b), \\ x^{-1} & = & f(a)^{-1} & = & f(a^{-1}). \end{array}$$

Ponieważ $a \cdot b, a^{-1} \in G$, to $x * y,\ x^{-1} \in \operatorname{Im} f$, czyli $\operatorname{Im} f$ jest podgrupą w F.

(9) Zostawiamy Czytelnikowi do samodzielnego udowodnienia.

(10) ($\Rightarrow$) Załóżmy, że $f : G \to F$ jest monomorfizmem grup. Wtedy f jest odwzorowaniem iniektywnym, a więc z $g \neq e_G$ dla elementu $g \in G$ wynika, że

$$f(g) \neq f(e_G) = e_F.$$

Zatem $\operatorname{Ker} f = \{e_G\}$.

($\Leftarrow$) Niech $\operatorname{Ker} f = \{e_G\}$, a g_1, g_2 będą elementami z G. Chcemy udowodnić, że f jest iniekcją. Nie wprost. Załóżmy, że $f(g_1) = f(g_2)$. Wtedy

$$f(g_1 \cdot g_2^{-1}) = f(g_1) * f(g_2^{-1}) = f(g_2) * f(g_2)^{-1} = e_F,$$

czyli $g_1 \cdot g_2^{-1} \in \operatorname{Ker} f$. Stąd otrzymujemy, że $g_1 \cdot g_2^{-1} = e_G$ lub równoważnie $g_1 = g_2$. Z tych rozumowań wnosimy, że f jest odwzorowaniem iniektywnym.

(11) ($\Rightarrow$) Niech $f : G \to F$ będzie epimorfizmem grup. Wtedy dla każdego elementu $b \in F$ znajdzie się taki element $g \in G$, że $b = f(g)$. Stąd wynika, że $F \subseteq \operatorname{Im} f$ i, jako wniosek, $\operatorname{Im} f = F$.

($\Leftarrow$) Jeśli obraz $\operatorname{Im} f = F$, to dla każdego elementu $b \in F$ mamy $b \in \operatorname{Im} f$, a zatem $b = f(g)$ dla pewnego $g \in G$. Ostatecznie f jest odwzorowaniem suriektywnym.

(12) Wynika z (10) oraz (11). □

Przykłady 3.5.3.

(**1**) Niech $0 : G \to F$ będzie odwzorowaniem takim, że $0(g) = e_F$ dla wszystkich elementów $g \in G$. Wtedy

$$0(g \cdot h) = e_F = e_F * e_F = 0(g) * 0(h)$$

dla dowolnych $g, h \in G$. Homomorfizm 0 jest nazywany *trywialnym* (lub *zerowym*).

Odwzorowanie $i_G : G \ni g \mapsto g \in G$ jest automorfizmem grupy G (nazywanym *tożsamościowym* lub *jednostkowym*), a $0 : G \ni g \mapsto e \in G$ (gdzie e jest elementem neutralnym w G) jest endomorfizmem grupy G.

(**2**) Odwzorowanie $h : \mathbb{C}^* \to \mathbb{R}^*$ takie, że

$$h(z) = \frac{1}{|z|},$$

jest homomorfizmem grup, bo dla dowolnych $z_1, z_2 \in \mathbb{C}^*$ zachodzi

$$h(z_1 z_2) = \frac{1}{|z_1 z_2|} = \frac{1}{|z_1| \cdot |z_2|} = \frac{1}{|z_1|} \cdot \frac{1}{|z_2|} = h(z_1) \cdot h(z_2).$$

Z równości $h(z) = \frac{1}{|z|} = 1$ wynika, że jądro

$$\operatorname{Ker} h = \{z \in \mathbb{C}^* \mid |z| = 1\}.$$

Skoro $h(z) > 0$ dla dowolnego elementu $z \in \mathbb{C}^*$ i, oprócz tego, dla dowolnej liczby rzeczywistej r mamy $h(\frac{1}{r}) = r$, to obraz $\operatorname{Im} h = \mathbb{R}_+$ jest grupą multiplikatywną dodatnich liczb rzeczywistych.

(**3**) Znajdźmy: *a*) $\operatorname{Hom}(\mathbb{Z}_3, \mathbb{Z}_{13})$ oraz *b*) $\operatorname{Hom}(\mathbb{Z}_3, \mathbb{Z}_{12})$.

a) Niech $f : \mathbb{Z}_3 \to \mathbb{Z}_{13}$ będzie dowolnym homomorfizmem grup. Dalej przez k_n będziemy oznaczać klasę reszt $k \pmod n$. Oczywiście, że $k_n = k \cdot 1_n$, gdzie $k \in \mathbb{Z}$. Skoro

$$3f(1_3) = f(1_3 + 1_3 + 1_3) = f(0_3) = 0_{13},$$

to rząd $o(f(1_3))$ jest dzielnikiem liczby 3. Z innej strony $|\mathbb{Z}_{13}| = 13$ i $o(f(1_3))$ jest dzielnikiem liczby 13, a zatem $o(f(1_3)) = 1$ oraz $f(1_3) = 0_{13}$. Stąd

$$f(k_3) = kf(1_3) = 0_{13}$$

oraz f jest homomorfizmem zerowym. Obliczyliśmy, że $\mathrm{Hom}(\mathbb{Z}_3, \mathbb{Z}_{13}) = \{0\}$.

b) Niech $g : \mathbb{Z}_3 \to \mathbb{Z}_{12}$ będzie dowolnym homomorfizmem. Ponieważ $k_3 = k \cdot 1_3$, to wystarczy sprawdzić, jaki może być element $g(1_3)$. Oczywiście, że $o(g(1_3)) \mid 3$. Jeśli $o(g(1_3)) = 1$, to $g = g_0$ jest homomorfizmem zerowym. Jeśli zaś $o(g(1_3)) = 3$, to albo $g(1_3) = 4 \cdot 1_{12}$, albo $g(1_3) = 8 \cdot 1_{12}$.

Łatwo zauważyć, że odwzorowanie $g_a : \mathbb{Z}_3 \to \mathbb{Z}_{12}$, określone wzorem $g_a(k_3) = ak \cdot 1_{12}$ (gdzie $a = 4$ lub $a = 8$), jest homomorfizmem grup. W rzeczy samej, ponieważ $(k + m)_3 = k_3 + m_3$ dla dowolnych $k, m \in \mathbb{Z}$, to

$$g_a(k_3 + m_3) = g((k + m)_3) = a(k + m) \cdot 1_{12} = ak \cdot 1_{12} + am \cdot 1_{12} = g_a(k_3) + g_a(m_3).$$

Oprócz tego

$$\begin{array}{rcl} 2g_8(k_3) & = & 2 \cdot 8 \cdot k \cdot 1_{12} = 4 \cdot k \cdot 1_{12}, \\ 3g_8(k_3) & = & 3 \cdot 8 \cdot k \cdot 1_{12} = 0_{12} = g_0(k_3), \end{array}$$

a więc $\mathrm{Hom}(\mathbb{Z}_3, \mathbb{Z}_{12})$ jest addytywną grupą cykliczną generowaną przez element g_8 rzędu 3.

(**4**) Niech $g : GL_n(\mathbb{R}) \to \mathbb{R}^*$ będzie odwzorowaniem, które każdej nieosobliwej macierzy kwadratowej $A \in M_n(\mathbb{R})$ stopnia n przyporządkowuje jej wyznacznik $\det A \in \mathbb{R}^*$. Wtedy na podstawie wzoru Cauchy'ego-Bineta

$$g(AB) = \det(AB) = \det A \cdot \det B = g(A) \cdot g(B),$$

czyli g jest homomorfizmem grup. Z równości

$$t = \begin{vmatrix} t & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix},$$

gdzie $t \in \mathbb{R}^*$, wynika, że t ma przeciwobraz

$$\begin{bmatrix} t & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in GL_n(\mathbb{R}),$$

czyli g jest suriekcją, oraz obraz $\mathrm{Im}\, g = \mathbb{R}^*$. Teraz jeśli $A \in \mathrm{Ker}\, g$, to $\det A = 1$, a zatem $\mathrm{Ker}\, g = SL_n(\mathbb{R})$. Wnosimy, że g jest epimorfizmem grup, który nie jest izomorfizmem.

(**5**) Odwzorowanie $\alpha : \mathbb{Z} \to \mathbb{Z}$ takie, że $\alpha(n) = -n$ $(n \in \mathbb{Z})$ jest automorfizmem grupy $(\mathbb{Z}, +)$.

(**6**) Odwzorowanie $f : \mathbb{S}_n \to \mathbb{Z}_2$, gdzie

$$f(\sigma) = \begin{cases} \overline{0}, & \text{gdy permutacja } \sigma \text{ jest parzysta,} \\ \overline{1}, & \text{gdy permutacja } \sigma \text{ jest nieparzysta,} \end{cases}$$

jest homomorfizmem grup, przy czym jądro $\operatorname{Ker} f = \mathbb{A}_n$ oraz obraz $\operatorname{Im} f = \mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$.
(**7**) Grupy $(\mathbb{R}_+, \cdot)$ oraz $(\mathbb{R}, +)$ są izomorficzne, bo

$$\varphi : \mathbb{R} \ni x \mapsto e^x \in \mathbb{R}_+$$

jest izomorfizmem grup (sprawdzić samodzielnie).
(**8**) Rozpatrzmy zbiór

$$S = \left\{ \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{F} \text{ oraz } a \neq 0 \right\},$$

gdzie $\mathbb{F}$ jest ciałem. Nietrudno przekonać się, że S jest podgrupą w $GL_2(\mathbb{F})$. Ponadto odwzorowanie

$$\phi : S \ni \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \mapsto a \in \mathbb{F}^*$$

jest izomorfizmem grup S oraz $\mathbb{F}^*$.

* * *

■ **Twierdzenie Cayleya.** Zachodzi takie

Twierdzenie 3.5.4 (Cayleya)**.** *Każda grupa G jest izomorficzna z pewną podgrupą grupy przekształceń bijektywnych $\mathbb{S}(G)$ zbioru G.*

Dowód. Niech e będzie elementem neutralnym grupy G oraz $g \in G$. Odwzorowanie

$$L_g : G \ni x \mapsto gx \in G$$

jest bijekcją. Rzeczywiście

$$\begin{gathered}(L_g \circ L_{g^{-1}})(x) = L_g(L_{g^{-1}}(x)) = L_g(g^{-1}x) = (gg^{-1})x = \\ = ex = x = \mathrm{id}_G(x) = x = ex = (g^{-1}g)x = g^{-1}(gx) = \\ = L_{g^{-1}}(gx) = L_{g^{-1}}(L_g(x)) = (L_{g^{-1}} \circ L_g)(x),\end{gathered}$$

czyli $L_{g^{-1}}$ jest odwzorowaniem odwrotnym do L_g, a to oznacza, że L_g jest bijekcją dla każdego $g \in G$. Oprócz tego

$$L_{gh}(x) = (gh)x = g(hx) = g(L_h(x)) = L_g(L_h(x)) = (L_g \circ L_h)(x)$$

dla dowolnych $g, h, x \in G$. Zatem

$$L_{gh} = L_g \circ L_h.$$

Ponieważ $\mathbb{S}(G)$ jest grupą i $L_g \in \mathbb{S}(G)$, to rozpatrzmy odwzorowanie

$$\theta : G \ni g \mapsto L_g \in \mathbb{S}(G).$$

Skoro

$$\theta(gh) = L_{gh} = L_g \circ L_h = \theta(g) \circ \theta(h)$$

dla elementów $g, h \in G$, to wnosimy, że θ jest homomorfizmem grup. Jeśli $\theta(a) = \theta(b)$ dla pewnych elementów $a, b \in G$, to $L_a = L_b$, a stąd $L_a(x) = L_b(x)$ dla wszystkich $x \in G$. Przepisując ostatnią równość w postaci $ax = bx$ i domnażając ją prawostronnie przez element x^{-1}, otrzymujemy

$$a = (ax)x^{-1} = (bx)x^{-1} = b,$$

czyli odwzorowanie θ jest iniektywne.

Rozpatrzmy podzbiór

$$L = \{L_g \mid g \in G\} \subseteq \mathbb{S}(G).$$

Ponieważ $\mathrm{id}_G = L_e$, $L_{gh} = L_g \circ L_h$ oraz $(L_g)^{-1} = L_{g^{-1}}$ dla dowolnych elementów $g, h \in G$, to L jest podgrupą w $\mathbb{S}(G)$. Wtedy odwzorowanie

$$\varphi : G \ni g \mapsto L_g \in L$$

jest izomorfizmem grup (przekonać się samodzielnie). □

Wniosek 3.5.5. *Niech G będzie grupą skończoną rzędu n. Wtedy G jest izomorficzna z pewną podgrupą grupy permutacji $\mathbb{S}_n$ stopnia n.*

■ Zatem każda podgrupa skończona rzędu n wkłada się w grupę symetryczną $\mathbb{S}_n$.

* * *

■ **Automorfizmy grupy.**

Twierdzenie 3.5.6. *Niech $(G, \cdot)$ będzie grupą. Wtedy zachodzą następujące własności:*

(1)

$$\mathrm{Aut}\, G = \{\varphi : G \to G \mid \varphi \textit{ jest automorfizmem grupy } G\}$$

jest grupą względem złożenia „$\circ$”, (która jest nazywana *grupą automorfizmów* grupy G)*;*

(2) *odwzorowanie*

$$\varphi_a : G \ni g \mapsto a^{-1} \cdot g \cdot a \in G$$

jest automorfizmem grupy G (nazywanym *automorfizmem wewnętrznym* grupy G *indukowanym* przez element $a \in G$)*;*

(3)

$$\operatorname{Inn} G = \{\varphi_a \mid a \in G\}$$

jest podgrupą normalną w G*.*

Dowód. (1) Odwzorowanie tożsamościowe $\mathrm{id}_G : G \ni g \mapsto g \in G$ jest automorfizmem grupy G. Skoro złożenie przekształceń bijektywnych grupy G jest bijekcją oraz

$$\begin{aligned}(\varphi \circ \psi)(a \cdot b) &= \varphi(\psi(a \cdot b)) = \varphi(\psi(a) \cdot \psi(b)) = \\ &= \varphi(\psi(a)) \cdot \varphi(\psi(b)) = (\varphi \circ \psi)(a) \cdot (\varphi \circ \psi)(b)\end{aligned}$$

dla dowolnych $\varphi, \psi \in \operatorname{Aut} G$ oraz dowolnych $a, b \in G$, to „$\circ$” jest działaniem algebraicznym na zbiorze $\operatorname{Aut} G$. Na mocy twierdzenia 1.4.6 działanie „$\circ$” jest łączne. Oprócz tego

$$\mathrm{id}_G \circ \varphi = \varphi = \varphi \circ \mathrm{id}_G$$

dla każdego $\varphi \in \operatorname{Aut} G$, czyli id_G jest elementem neutralnym (względem „$\circ$”).

Załóżmy, że $\varphi \in \operatorname{Aut} G$. Wtedy odwzorowanie $\varphi^{-1} : G \to G$ istnieje i jest bijekcją na podstawie wniosku 1.4.12. Jeśli $a, b \in G$, to $\varphi(a) = g$, $\varphi(b) = h$ dla pewnych $g, h \in G$, a więc $a = \varphi^{-1}(g)$ oraz $b = \varphi^{-1}(h)$. Biorąc to pod uwagę, otrzymujemy, że

$$\begin{aligned}&\varphi^{-1}(g \cdot h) = \varphi^{-1}(\varphi(a) \cdot \varphi(b)) = \varphi^{-1}(\varphi(a \cdot b)) = (\varphi^{-1} \circ \varphi)(a \cdot b) = \\ &= \mathrm{id}_G(a \cdot b) = a \cdot b = \varphi^{-1}(g) \cdot \varphi^{-1}(h).\end{aligned}$$

Zatem $\varphi^{-1} \in \operatorname{Aut} G$.

(2) Ponieważ

$$\varphi_a \circ \varphi_{a^{-1}} = \mathrm{id}_G = \varphi_{a^{-1}} \circ \varphi_a,$$

to

$$\varphi_{a^{-1}} = (\varphi_a)^{-1}$$

i na podstawie twierdzenia 1.4.11 wnosimy, że φ_a jest bijekcją. Oprócz tego

$$\varphi_a(g \cdot h) = a^{-1} \cdot (g \cdot h) \cdot a = (a \cdot g \cdot a^{-1}) \cdot (a \cdot h \cdot a^{-1}) = \varphi_a(g) \cdot \varphi_a(h) \ ,$$

a więc $\varphi_a \in \operatorname{Aut} G$.

(3) Ćwiczenie. □

■ Automorfizm grupy G, który nie jest wewnętrzny, jest nazywany *zewnętrznym*.

Ćwiczenia 3.5.7.

(**1**) Znaleźć wszystkie homomorfizmy grupy G w grupę H, jeśli:
(a) $G = \mathbb{Z}_3$ oraz $H = \mathbb{Z}_2$;
(b) $G = \mathbb{Z}_2$ oraz $H = \mathbb{Z}_3$;
(c) $G = \mathbb{Z}_{37}$ oraz $H = \mathbb{Z}_{51}$;
(d) $G = \mathbb{Z}_8$ oraz $H = \mathbb{Z}_4$;
(e) $G = H = \mathbb{Z}_6$;
(f) $G = \mathbb{Z}$ oraz $H = \mathbb{Z}_4$.
(**2**) Niech G i H będą grupami. Sprawdzić, czy reguła $f : G \to H$ określa homomorfizm (odpowiednio monomorfizm, epimorfizm, izomorfizm) grup, jeśli:
(a) $G = \mathbb{C}^*$, $H = \mathbb{R}^*$ oraz $f(z) = \frac{1}{|z|^2}$;
(b) $G = \mathbb{C}^*$, $H = \mathbb{R}^*$ oraz $f(z) = |z|$;
(c) $G = GL_2(\mathbb{C})$, $H = \mathbb{C}^*$ oraz $f(A) = \det A$;
(d) $G = \mathbb{C}^*$, $H = \mathbb{R}^*$ oraz $f(z) = 2|z|$;
(e) $G = H = \mathbb{Z}$ oraz $f(n) = n^2$;
(f) $G = M_n(\mathbb{Q})$, $H = \mathbb{Q}$ oraz $f(A) = \operatorname{tr} A$ jest śladem macierzy A;
(g) $G = C_{[0,1]}$, $H = \mathbb{R}$ oraz $f(g) = \int_0^1 g(x)dx$, gdzie $C_{[0,1]}$ jest grupą abelową funkcji rzeczywistych ciągłych na odcinku $[0, 1]$;
(h) $G = H = C_{[0,2]}$ oraz $(f(g))(X) = g(X^3)$;
(i) $G = H = M_2(\mathbb{R})$ oraz $f(A) = A^T$ jest macierzą transponowaną do A;
(j) $G = GL_2(\mathbb{R})$, $H = \mathbb{R}_+$ oraz $f(A) = |\det A|$.
Znaleźć jądro $\operatorname{Ker} f$ i obraz $\operatorname{Im} f$.
(**3**) Udowodnić, że grupy G i H są izomorficzne, jeśli:
(a) $G = \mathbb{C}^*$ oraz $H = \{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R},\ a^2 + b^2 \neq 0\}$;
(b) $G = \mathbb{Z}$ oraz $H = \{2^a \mid a \in \mathbb{Z}\}$;
(c) $G = \{\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{C}\}$ oraz $H = \mathbb{C}$;
(d) $G = M_2(\mathbb{C})$ oraz $H = \mathbb{C}^4$;
(e) $G = C_{[1,3]}$ oraz $H = C_{[5,6]}$;
(f) $G = \mathbb{Q}^*$ oraz $H = \mathbb{Q} \setminus \{-1\}$, gdzie H jest grupą względem działania „$\circ$”, określonego wzorem $x \circ y = x + y + xy$.
(**4**) Sprawdzić, czy grupy G oraz H są izomorficzne, jeśli:
(a) $G = \mathbb{R}$ oraz $H = \mathbb{R}^*$;
(b) $G = \mathbb{R}^*$ oraz $H = \mathbb{R}_+$;
(c) $G = \mathbb{C}^*$ oraz $H = \mathbb{C}$;
(d) $G = M_2(\mathbb{R})$ oraz $H = GL_2(\mathbb{R})$;

(e) $G = \mathbb{Q}$ oraz $H = \mathbb{Q}^*$.
(**5**) Znaleźć grupę automorfizmów:
(a) $\operatorname{Aut}\mathbb{Q}$;
(b) $\operatorname{Aut}\mathbb{Z}_2$;
(c) $\operatorname{Aut}\mathbb{Z}_3$;
(d) $\operatorname{Aut}\mathbb{Z}_4$;
(e) $\operatorname{Aut}\mathbb{Z}$;
(f) $\operatorname{Aut}V_4$.
(**6**) Udowodnić, że grupa $\mathbb{S}_n$ wkłada się w grupę $\mathbb{S}_{n+1}$.
(**7**) Niech G będzie grupą oraz $g \in G$. Udowodnić, że:
(a) odwzorowanie $R_g : G \ni x \mapsto xg \in G$ jest bijektywne i $R_{gh} = R_h \circ R_g$ dla dowolnych $g, h \in G$;
(b) $R_g \circ L_h = L_h \circ R_g$, gdzie $L_h : G \ni x \mapsto hx \in G$;
(c) jeśli $\varphi : G \to G$ jest odwzorowaniem bijektywnym takim, że $L_g \circ \varphi = \varphi \circ L_g$ dla wszystkich $g \in G$, to $\varphi = R_h$ dla pewnego elementu $h \in G$.

Uwagi. Pojęciem homomorfizmu operował jeszcze E. Galois, ale formalnie tego nie zdefiniował. Badanie homomorfizmów i ich jąder w 1878 r. zainicjował A. Capelli[(9)]. Automorfizmy wewnętrzne i zewnętrzne jako pierwszy studiował G. Frobenius w 1901 r.

[(9)] Alfredo Capelli (1895–1910)

3.6. Grupy ilorazowe

■ Niech $(G, \cdot)$ będzie grupą (z elementem neutralnym e), a H będzie jej podgrupą normalną. Wtedy, jak udowodniono wyżej, każda warstwa lewostronna gH jest równa odpowiedniej warstwie prawostronnej Hg. Niech

$$G/H = \{gH \mid g \in G\}$$

będzie zbiorem warstw grupy G względem jej podgrupy H. Zdefiniujmy na zbiorze G/H działanie „$\cdot$" w taki sposób

$$gH \cdot tH = (g \cdot t)H$$

dla dowolnych elementów $g, t \in G$. Udowodnijmy, że to działanie jest algebraiczne na zbiorze G/H. W tym celu wystarczy przekonać się, że wynik działania „$\cdot$" nie zależy od wyboru reprezentantów w warstwach lub, co znaczy to samo, że z równości $g_1H = g_1'H$ i $g_2H = g_2'H$ wynika równość

$$(g_1 \cdot g_2)H = (g_1' \cdot g_2')H,$$

gdzie $g_1, g_2, g_1', g_2' \in G$. Istotnie, bo z

$$g_1H = g_1'H \text{ oraz } g_2H = g_2'H$$

otrzymujemy, że $g_1 = g_1' \cdot h_1$ i $g_2 = g_2' \cdot h_2$ dla pewnych $h_1, h_2 \in H$ i wtedy

$$g_1 \cdot g_2 = (g_1' \cdot h_1)(g_2' \cdot h_2) = g_1' \cdot g_2' \cdot ((g_2')^{-1} \cdot h_1 \cdot g_2' \cdot h_2),$$

gdzie $(g_2')^{-1} \cdot h_1 \cdot g_2' \cdot h_2 \in H$. Na podstawie lematu 3.3.5 wnioskujemy, że

$$(g_1 \cdot g_2)H = (g_1' \cdot g_2')H.$$

Twierdzenie 3.6.1. *Jeśli H jest podgrupą normalną grupy G, to G/H jest grupą względem działania „$\cdot$", przy czym jeśli G jest grupą abelową, to G/H również jest abelowa.*

Dowód. Jak udowodniono wyżej, działanie „$\cdot$” jest algebraiczne na zbiorze warstw lewostronnych G/H grupy G względem jej podgrupy H. Niech a, b, c będą dowolnymi elementami z G. Wtedy

$$((aH) \cdot (bH)) \cdot (cH) = (a \cdot b)H \cdot cH = ((a \cdot b) \cdot c)H =$$

$$= (a \cdot (b \cdot c))H = aH \cdot (b \cdot c)H = aH \cdot ((bH) \cdot (cH)),$$

czyli działanie „$\cdot$” jest łączne na zbiorze G/H. Jeśli e jest elementem neutralnym grupy G, to

$$aH \cdot eH = (a \cdot e)H = aH = (e \cdot a)H = eH \cdot aH,$$

a zatem $eH = H$ jest elementem neutralnym w G/H. Teraz, jeśli a^{-1} jest odwrotnym do a w grupie G, to $a \cdot a^{-1} = e = a^{-1} \cdot a$ i wtedy

$$aH \cdot a^{-1}H = (a \cdot a^{-1})H = eH = (a^{-1} \cdot a)H = a^{-1}H \cdot aH,$$

a więc

$$(aH)^{-1} = a^{-1}H.$$

Jeśli G jest grupą abelową, to $a \cdot b = b \cdot a$ dla dowolnych elementów $a, b \in G$ i na tej podstawie

$$aH \cdot bH = (a \cdot b)H = (b \cdot a)H = bH \cdot aH.$$

□

■ Jeśli $H \lhd G$, to grupa

$$G/H$$

jest nazywana *grupą ilorazową* grupy G względem jej podgrupy (normalnej) H.

■ Tutaj oznaczamy działanie w grupie G oraz działanie w grupie ilorazowej G/H tym samym symbolem „$\cdot$”. Chociaż to są różne działania, takie zachowanie jednak nie prowadzi do nieporozumień.

Przykłady 3.6.2.

(**1**) Grupa ilorazowa $\mathbb{Z}/2\mathbb{Z}$ addytywnej grupy liczb całkowitych $\mathbb{Z}$ względem podgrupy parzystych liczb całkowitych $2\mathbb{Z}$ składa się z dwóch elementów, które oznaczamy przez $\overline{0}$ i $\overline{1}$ (a zatem $\mathbb{Z}/2\mathbb{Z}$ jest 2-grupą rzędu 2).

(**2**) Podobnie grupa ilorazowa $\mathbb{Z}/n\mathbb{Z}$ składa się z n parami różnych elementów:

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}.$$

(**3**) Udowodnijmy, że

$$G/\langle e_G\rangle \cong G.$$

Istotnie odwzorowanie $\alpha : G \to G/\langle e_G\rangle$, określone przez $\alpha(g) = g\langle e_G\rangle$ $(g \in G)$, jest bijektywne i, oprócz tego, dla dowolnych elementów $g_1, g_2 \in G$ mamy

$$\alpha(g_1 \cdot g_2) = (g_1 \cdot g_2)\langle e_G\rangle = g_1\langle e_G\rangle \cdot g_2\langle e_G\rangle = \alpha(g_1) \cdot \alpha(g_2),$$

czyli α jest izomorfizmem grup. Jeśli podgrupa normalna H nie jest jednostkowa w grupie $(G, \cdot)$, to grupa ilorazowa G/H jest nazywana *właściwą*.

(**4**) Udowodnijmy, że

$$G/G \cong \langle e_G\rangle.$$

Z kryterium równości warstw wnosimy, że $g_1G = G = g_2G$ dla dowolnych $g_1, g_2 \in G$. Zatem grupa ilorazowa G/G składa się z dokładnie jednego elementu, a więc jest jednostkowa. Dlatego odwzorowanie $\beta : \langle e_G\rangle \to G/G$, określone przez regułę $\beta(e_G) = e_GG$, jest bijektywne.

(**5**) Niech H będzie podgrupą normalną grupy G, a $\pi : G \to G/H$ będzie takim odwzorowaniem, że $\pi(g) = gH$ dla każdego elementu $g \in G$. Wtedy π jest odwzorowaniem suriektywnym i dla dowolnych $g_1, g_2 \in G$ mamy

$$\pi(g_1 \cdot g_2) = (g_1 \cdot g_2)H = g_1H \cdot g_2H = \pi(g_1) \cdot \pi(g_2),$$

a zatem π jest epimorfizmem (który jest nazywany *naturalnym* lub *homomorfizmem kanonicznym*).

Twierdzenie 3.6.3 (pierwsze o izomorfizmie grup)**.** *Jeśli $f : G \to F$ jest homomorfizmem grupy $(G, \cdot)$ w grupę $(F, *)$, to mamy taki izomorfizm grup*

$$\operatorname{Im} f \cong G/\operatorname{Ker} f.$$

Dowód. Reguła $\phi : G/\operatorname{Ker} f \to \operatorname{Im} f$ taka, że

$$\phi(a \cdot \operatorname{Ker} f) = f(a)$$

$(a \in G)$ określa odwzorowanie. Ponadto dla dowolnych dwóch warstw $a_1 \cdot \operatorname{Ker} f$ oraz $a_2 \cdot \operatorname{Ker} f$, gdzie $a_1, a_2 \in G$, mamy

$$\phi((a_1 \cdot \operatorname{Ker} f) \cdot (a_2 \cdot \operatorname{Ker} f)) = \phi((a_1 \cdot a_2)\operatorname{Ker} f) =$$

$$= f(a_1 \cdot a_2) = f(a_1) * f(a_2) = \phi(a_1 \cdot \operatorname{Ker} f) * \phi(a_2 \cdot \operatorname{Ker} f),$$

a więc ϕ jest homomorfizmem grup. Jeśli warstwy

$$a_1 \cdot \operatorname{Ker} f \neq a_2 \cdot \operatorname{Ker} f$$

są różne, to na podstawie kryterium równości warstw ${a_2}^{-1} \cdot a_1 \notin \operatorname{Ker} f$, co daje

$$f(a_2)^{-1} * f(a_1) = f({a_2}^{-1}) * f(a_1) = f({a_2}^{-1} \cdot a_1) \neq e_F,$$

a stąd

$$\phi(a_1 \cdot \operatorname{Ker} f) = f(a_1) \neq f(a_2) = \phi(a_2 \cdot \operatorname{Ker} f).$$

To znaczy, że odwzorowanie ϕ jest iniektywne. Jeśli zaś z jest dowolnym elementem z obrazu $\operatorname{Im} f$, to $z = f(g)$ dla pewnego $g \in G$, a wtedy

$$z = \phi(g \cdot \operatorname{Ker} f)$$

i ϕ jest odwzorowaniem suriektywnym. Zatem ϕ jest izomorfizmem grup. □

Przykłady 3.6.4.

(**1**) Z przykładu 3.5.3 (5) na podstawie pierwszego twierdzenia o izomorfizmie grup wynika, że

$$\mathbb{R}^* \cong GL_n(\mathbb{R})/SL_n(\mathbb{R}).$$

(**2**) Mamy $\mathbb{S}_n/\mathbb{A}_n \cong \mathbb{Z}_2$ w wyniku pierwszego twierdzenia o izomorfizmie grup i przykładu 3.5.3 (7).

Twierdzenie 3.6.5 (drugie o izomorfizmie grup)**.** *Jeśli H jest podgrupą grupy $(G, \cdot)$ oraz N jest podgrupą normalną w G, to zachodzi izomorfizm grup*

$$(HN)/N \cong H/(H \bigcap N).$$

Dowód. (Szkic) Zaznaczmy, że przekrój $H \cap N$ jest podgrupą normalną w H oraz $N \leqslant NH$, a więc N jest podgrupą normalną w NH. Skoro

$$\varphi : H \ni h \mapsto hN \in (HN)/N$$

jest homomorfizmem grup, jądro $\operatorname{Ker} \varphi = H \cap N$ oraz obraz $\operatorname{Im} \varphi = (HN)/N$ (sprawdzić samodzielnie), to teza zachodzi na podstawie twierdzenia 3.6.3. □

Twierdzenie 3.6.6 (trzecie o izomorfizmie grup)**.** *Jeśli A, B są podgrupami normalnymi grupy $(G,\cdot)$ oraz $A \subseteq B$, to:*

(1) *A jest podgrupą normalną w B;*
(2) *B/A jest podgrupą normalną w G/A;*
(3) *zachodzi izomorfizm grup*

$$G/B \cong (G/A)/(B/A).$$

Dowód. (Szkic) Ponieważ

$$\varphi : G/A \ni xA \mapsto xB \in G/B \quad (x \in G)$$

jest homomorfizmem grup, $\operatorname{Ker}\varphi = B/A$ oraz $\operatorname{Im}\varphi = G/B$ (sprawdzić samodzielnie), to teza zachodzi na podstawie twierdzenia 3.6.3. □

Następne twierdzenie jest też nazywane twierdzeniem o odpowiedniości podgrup.

Twierdzenie 3.6.7. *Niech G będzie grupą, a H, S jej podgrupami, przy czym $H \lhd G$ oraz $H \leqslant S$. Załóżmy, że $L(G)$ jest zbiorem podgrup grupy G zawierających H, a $L(G/H)$ jest zbiorem wszystkich podgrup grupy ilorazowej G/H. Wtedy odwzorowanie*

$$\varphi : L(G) \ni S \mapsto S/H \in L(G/H)$$

jest bijektywne.

Dowód. Ćwiczenie. □

* * *

■ **Relacje równoważności na grupie, które są zgodne z działaniem w grupie.** Relacja równoważności „$\sim$”, która jest określona na grupie G, jest nazywana *kongruencją* w grupie $(G,\cdot)$, jeśli zachodzi implikacja

$$\forall_{a,b,c,d\in G} : a \sim b \text{ oraz } c \sim d \;\Rightarrow\; a \cdot c \sim b \cdot d.$$

Zachodzi takie

Twierdzenie 3.6.8. *Niech G będzie grupą (względem „$\cdot$”). Wtedy:*

(1) *jeśli H jest podgrupą normalną w G, to relacja „$\sim$” taka, że*

$$a \sim b \quad \Leftrightarrow \quad a^{-1}b \in H$$

dla elementów $a, b \in G$, jest kongruencją w grupie G (czyli a i b należą do jednej warstwy względem podgrupy H);

(2) *jeśli „$\backsim$” jest kongruencją w grupie G, to:*

(a) *klasa równoważności*

$$e_{\backsim} = \{g \in G \mid g \backsim e\},$$

której reprezentantem jest element neutralny e grupy G, jest podgrupą normalną w G;

(b) *zachodzi równość*

$$G/e_{\backsim} = G/\backsim.$$

Dowód. Niech $a, b, c, d, x \in G$.

(1) Załóżmy, że $a \sim b$ oraz $c \sim d$. Wtedy $a^{-1}b, c^{-1}d \in H$. Skoro $c^{-1}(a^{-1}b)c \in H$, to

$$(ac)^{-1}(bd) = c^{-1}a^{-1}bd = (c^{-1}(a^{-1}b)c) \cdot (c^{-1}d) \in H,$$

a to znaczy, że teza zachodzi.

(2) Oczywiście, że $a \sim a$. Oznaczmy klasę równoważności $e_{\backsim}$ przez H. Wtedy

$$\begin{array}{lllll} a, b \in H & \Rightarrow & a \backsim e,\ b \backsim e & \Rightarrow & \\ & \Rightarrow & ab \backsim e^2 = e & \Rightarrow & ab \in H, \\ a \in H & \Rightarrow & a^{-1} \backsim a^{-1},\ a \backsim e & \Rightarrow & \\ & \Rightarrow & a^{-1}a \backsim a^{-1}e & \Rightarrow & e \backsim a^{-1} \ \Rightarrow \ a^{-1} \in H \end{array}$$

i na podstawie kryterium H jest podgrupą w G. Skoro $x^{-1} \backsim x^{-1}$, to

$$a \backsim e,\ x \backsim x \Rightarrow ax \backsim ex = x \Rightarrow x^{-1}ax \backsim x^{-1}x = e \Rightarrow x^{-1}ax \in H,$$

a więc H jest normalna w G.

Rozpatrzmy dowolną warstwę $x_{\backsim} \in G/\backsim$. Skoro $x \backsim x$, to

$$xh \backsim xe = x \in x_{\backsim}$$

dla dowolnego $h \in H$ i na tej podstawie $xH \subseteq x_\backsim$.

Odwrotnie, jeśli $z \in x_\backsim$, to

$$z \backsim x \;\Rightarrow\; x^{-1}z \backsim x^{-1}x = e \in H \;\Rightarrow\; z \in xH,$$

a stąd $x_\backsim \subseteq xH$.

Zatem $x_\backsim = xH$ i teza zachodzi. □

* * *

■ **Homomorfizmy grup a kongruencje w grupie.** Zachodzi

Twierdzenie 3.6.9. *Jeśli $\varphi : G \to H$ jest homomorfizmem grup G i H, to relacja „$\sim$" taka,że*

$$g_1 \sim g_2 \quad \Leftrightarrow \quad \varphi(g_1) = \varphi(g_2)$$

dla elementów $g_1, g_2 \in G$ jest kongruencją w grupie G.

Dowód. Ćwiczenie. □

Ćwiczenia 3.6.10.

(**1**) Znaleźć grupy ilorazowe:
(a) $\mathbb{C}_{12}/\mathbb{C}_4$;
(b) $4\mathbb{Z}/20\mathbb{Z}$;
(c) $\mathbb{R}^*/\mathbb{R}_+$.
(**2**) Udowodnić, że zachodzą izomorfizmy grup:
(a) $C_{[-1,1]}/H \cong \mathbb{R}^2$, gdzie $H = \{f \in C_{[-1,1]} \mid f(-1) = 0 = f(1)\}$;
(b) $\mathbb{C}^*/\mathbb{R}^* \cong \mathbb{S}^1$, gdzie $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$;
(c) $\mathbb{C}^*/\mathbb{S}^1 \cong \mathbb{R}_+$;
(d) $\mathbb{C}^*/\mathbb{R}^+ \cong [0, 2\pi)$;
(e) $\mathbb{C}^*/\mathbb{C}_n \cong \mathbb{C}^*$;
(f) $GL_n(\mathbb{C})/SL_n(\mathbb{C}) \cong \mathbb{C}^*$;
(g) $GL_n(\mathbb{R})/H \cong \mathbb{C}_2$, gdzie $H = \{A \in GL_n(\mathbb{R}) \mid \det A > 0\}$;
(h) $GL_n(\mathbb{R})/H \cong \mathbb{R}_+$, gdzie $H = \{A \in GL_n(\mathbb{R}) \mid |\det A| = 1\}$;
(i) $\mathbb{R}^*/\mathbb{R}_+ \cong \{-1, 1\}$;
(j) $\mathbb{R}^*/\{-1, 1\} \cong \mathbb{R}_+$;
(k) $\mathbb{S}^1/\mathbb{C}_n \cong \mathbb{S}^1$;
(l) $\mathbb{R}^3/H \cong \mathbb{R}^2$ dla $H = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 = x_1 + 2x_2 + x_3 = x_1 - x_2 - 2x_3 = 0\}$;
(m) $\mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1$.
(**3**) Udowodnić, że:
(a) $\mathbb{Z}^3$ jest grupą względem działania „$*$", określonego przez

$$(a, b, c) * (x, y, z) = (a + (-1)^c x, b + y, c + z), \text{gdzie } \; a, b, c, x, y, z \in \mathbb{Z};$$

(b) $H = \{(a, 0, 0) \mid a \in \mathbb{Z}\} \lhd \mathbb{Z}^3$;

(*c*) $\mathbb{Z}^3/H \cong \mathbb{Z}[i]$, gdzie $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ jest grupą addytywną gaussowskich liczb całkowitych.

(**4**) Udowodnić, że $\mathbb{Q}/\mathbb{Z}$ jest grupą nieskończoną, której każdy element ma rząd skończony.

Uwagi. Ideologia pojęcia grupy ilorazowej w rzeczywistości została założona w oryginalnych badaniach C. Gaussa (w szczególności w pracy poświęconej arytmetyce modularnej w 1801 r.), lecz dopiero E. Galois w 1830 r. zdefiniował pojęcia podgrupy normalnej i grupy ilorazowej, a C. Jordan w 1873 r. wprowadził oznaczenie G/H dla grup ilorazowych.

3.7. Klasyfikacja izomorficzna grup cyklicznych

Ma miejsce takie

Twierdzenie 3.7.1. *Zachodzą następujące własności:*
(*a*) *Każde dwie nieskończone grupy cykliczne są izomorficzne.*
(*b*) *Dwie skończone grupy cykliczne są izomorficzne wtedy i tylko wtedy, gdy mają jednakowe rzędy.*

Dowód. Niech G_1 (odpowiednio G_2) będzie grupą cykliczną z działaniem „$\cdot$" (odpowiednio „$*$") generowaną przez element g_1 (odpowiednio g_2) z elementem neutralnym e_1 (odpowiednio e_2), czyli

$$G_i = \{g_i^n \mid n \in \mathbb{Z}\}$$

($i = 1, 2$). Wtedy reguła

$$\varphi(g_1^m) = g_2^m,$$

gdzie $m \in \mathbb{Z}$, zadaje odwzorowanie $\varphi : G_1 \to G_2$, które jest suriektywne. Skoro dla dowolnych elementów z_1 i z_2 z grupy G_1 istnieją takie liczby całkowite m_1 oraz m_2, że $z_i = g_i^{m_i}$ ($i = 1, 2$), to

$$\begin{aligned}&\varphi(z_1 \cdot z_2) = \varphi(g_1^{m_1} \cdot g_1^{m_2}) = \varphi(g_1^{m_1+m_2}) = \\ &= g_2^{m_1+m_2} = g_2^{m_1} * g_2^{m_2} = \varphi(z_1) * \varphi(z_2).\end{aligned}$$

Załóżmy, że z_1 i z_2 są elementami grupy G_1. Jeśli $\varphi(z_1) = \varphi(z_2)$, to otrzymujemy równości

$$g_2^{m_1} = \varphi(g_1^{m_1}) = \varphi(g_1^{m_2}) = g_2^{m_2},$$

skąd

$$g_2^{m_1-m_2} = e_2. \tag{3.3}$$

(*a*) Teraz niech G_1 i G_2 będą grupami nieskończonymi. Wtedy

$$o(g_1) = o(g_2) = \infty$$

na podstawie lematu 3.1.4. W wyniku (3.3) otrzymujemy, że $m_1 - m_2 = 0$ i, jako wniosek, $z_1 = z_2$. To oznacza, że odwzorowanie φ jest iniektywne. Wnosimy, że φ jest izomorfizmem grup.

(b) Teraz niech G_1 i G_2 będą grupami skończonymi.

($\Leftarrow$) Niech $|G_1| = |G_2|$. Na mocy lematu 3.1.4 mamy $o(g_1) = o(g_2) = s$ dla pewnego $s \in \mathbb{Z}$. Wtedy, jak ustalono w dowodzie lematu 3.1.4,

$$G_i = \{g_i^0 = e_i, g_i, \ldots, g_i^{s-1}\} \ (i = 1, 2).$$

Dlatego dalej uważamy, że w definicji reguły φ mamy $0 \leqslant m \leqslant s-1$.

Jeśli $z_1 = g_1^{m_1}$ oraz $z_2 = g_1^{m_2}$, gdzie $0 \leqslant m_1, m_2 \leqslant s-1$ są różnymi elementami i, na przykład, $m_1 > m_2$, to z równości $\varphi(z_1) = \varphi(z_2)$ wynika (3.3), skąd otrzymujemy, że $o(g_2) \leqslant m_1 - m_2 < s$, co nie jest możliwe. Zatem φ jest odwzorowaniem iniektywnym i z powyższych rozumowań wnosimy, że φ jest izomorfizmem grup G_1 oraz G_2.

($\Rightarrow$) Załóżmy, że grupy G_1 oraz G_2 są izomorficzne i $\theta : G_1 \to G_2$ jest pewnym izomorfizmem. Skoro θ jest odwzorowaniem suriektywnym, to dla każdego elementu $w \in G_2$ znajdzie się taki element $g \in G_1$, że $w = \theta(g)$. Lecz $g = g_1^s$ dla pewnego $s \in \mathbb{Z}$, a zatem

$$w = \theta(g_1^s) = \theta(g_1)^s.$$

Z tego wynika, że $G_2 \leqslant \langle \theta(g_1) \rangle$. Z innej strony $\theta(g_1) \in G_2$, a więc

$$\theta(g_1^t) = \theta(g_1)^t \in G_2$$

dla wszystkich $t \in \mathbb{Z}$, czyli $\langle \theta(g_1) \rangle \leqslant G_2$. Z tych rozumowań dostajemy $\langle \theta(g_1) \rangle = G_2$. Na mocy twierdzenia 3.5.2 (własność (6)) rzędy

$$o(g_1) = o(\theta(g_1))$$

są równe. Za lematem 3.1.4 mamy

$$|G_1| = o(g_1), \ |G_2| = o(\theta(g_1)),$$

a stąd $|G_1| = |G_2|$.

□

Wniosek 3.7.2. *Zachodzą następujące własności:*

(a) *jeśli G jest nieskończoną grupą cykliczną, to G jest izomorficzna z addytywną grupą liczb całkowitych $\mathbb{Z}$;*

(*b*) *jeśli G jest skończoną grupą cykliczną rzędu n, to mamy izomorfizm grup $G \cong \mathbb{Z}_n$.*

□

Ćwiczenia 3.7.3.

(**1**) Sprawdzić, czy zbiór liczb całkowitych $\mathbb{Z}$ jest grupą cykliczną względem reguły „$*$”, określonej przez wzór $a * b = a + b + 2$.
(**2**) Sprawdzić, czy G jest grupą cykliczną, jeśli:
(*a*) $G = V_4$;
(*b*) $G = U(\mathbb{Z}_7)$;
(*c*) $G = SL_2(\mathbb{Z}_2)$;
(*d*) $G = \{f_{a,b} : \mathbb{Z}_7 \to \mathbb{Z}_7 \mid f_{a,b}(X) = aX + b,\ a \neq 0,\ a, b \in \mathbb{Z}_7\}$;
(*e*) $G = Q_8$;
(*f*) $G = D_8$;
(*g*) $G = \mathbb{Z}[i]$.

Uwagi. Izomorfizm grup (chociaż formalnie niezdefiniowany) jest „obecny” już w słynnej pracy C. Gaussa z 1801 r.

3.8. Działanie grupy na zbiorze

■ Niech $(G, \cdot)$ będzie grupą (z elementem neutralnym e), a X będzie zbiorem niepustym. Odwzorowanie

$$* : G \times X \ni (g, x) \mapsto g * x \in X$$

jest nazywane *działaniem lewostronnym grupy G na zbiorze X*, jeśli są spełnione następujące warunki:
1)

$$\forall_{x \in X} : e * x = x;$$

2)

$$\forall_{g,h \in G} \forall_{x \in X} : g * (h * x) = (g \cdot h) * x.$$

■ Podobnie możemy zdefiniować *działanie prawostronne grupy G na zbiorze X*.

Lemat 3.8.1. *Działanie „$*$" grupy G (względem „$\cdot$") na zbiorze X indukuje relację równoważności ϱ na zbiorze X według reguły:*

$$x \varrho y \Leftrightarrow \exists_{g \in G} : g * x = y,$$

gdzie $x, y \in X$.

Dowód. Skoro $x = e * x$ dla każdego $x \in X$, to relacja ϱ jest zwrotna. Jeśli $x \varrho y$ dla pewnych $x, y \in X$, to $g * x = y$ dla pewnego $g \in G$. Wtedy

$$g^{-1} * y = g^{-1} * (g * x) = (g^{-1} \cdot g) * x = e * x = x,$$

czyli $y \varrho x$ i relacja ϱ jest symetryczna. Załóżmy, że $x \varrho y$ oraz $y \varrho z$ (czyli $g * x = y$ i $h * y = z$ dla pewnych elementów $g, h \in G$). Wtedy

$$z = h * (g * x) = (h \cdot g) * x,$$

a zatem $x \varrho z$ i relacja ϱ jest przechodnia. Udowodniliśmy, że ϱ jest relacją równoważności na zbiorze X. □

■ Z działaniem

$$* : G \times X \ni (g, x) \mapsto g * x \in X$$

grupy G na zbiorze niepustym X są związane takie pojęcia:

- zbiór X jest nazywany *G-przestrzenią*;
- zbiór $\mathrm{Ker}(*) = \{g \in G \mid g * x = x \text{ dla wszystkich } x \in X\}$ jest nazywany *jądrem* działania grupy G na zbiorze X;
- jeśli $\mathrm{Ker}(*) = \{e\}$, to działanie „$*$" grupy G jest nazywane *wiernym* (lub *efektywnym*) na zbiorze X;
- zbiór $G(x) = \{g * x \mid g \in G\}$ jest nazywany *orbitą* (lub *obszarem przechodniości*) elementu $x \in X$ (i oznaczany przez $Orb(x)$ lub $O(x)$);
- zbiór $G_x = \{g \in G \mid g*x = x\}$ jest nazywany *stabilizatorem* elementu $x \in X$ (i oznaczany przez $St(x)$);
- element $x \in X$ jest nazywany *elementem stałym działania „$*$"*, jeśli $G_x = G$;
- grupa G działa *tranzytywnie* (lub *przechodnio*) na zbiorze X, jeśli dla dowolnych elementów $x, y \in X$ znajdzie się taki element $g \in G$ (zależny od x i y), że $g * x = y$ (w tym przypadku działanie „$*$" ma dokładnie jedną orbitę, czyli $G(x) = X$ dla dowolnego elementu $x \in X$).

Twierdzenie 3.8.2 (własności działania grupy na zbiorze)**.** *Niech*

$$* : G \times X \ni (g, x) \mapsto g * x \in X$$

będzie działaniem grupy G na zbiorze X. Wtedy zachodzą następujące własności:

(1) *istnieje homomorfizm grup*

$$G \ni g \mapsto \mu_g \in \mathbb{S}(X), \tag{3.4}$$

gdzie odwzorowanie $\mu_g : X \to X$ jest zadane wzorem

$$\forall_{x \in X} : \mu_g(x) = g * x;$$

(2) *jądro* $\mathrm{Ker}(*)$ *jest podgrupą normalną w G.*

Dowód. Ćwiczenie. □

Przykłady 3.8.3.

Niech niżej $(G, \cdot)$ będzie grupą z elementem jednostkowym e oraz w pierwszych dwóch przykładach $X = G$.

(**1**) Reguła

$$* : G \times G \ni (x, g) \mapsto x * g = x \cdot g \in G$$

spełnia warunki $x * e = x \cdot e = x$ oraz $(x * g) * h = (x \cdot g) \cdot h = x \cdot (g \cdot h) = x * (g \cdot h)$ dla dowolnych $g, h, x \in G$, a zatem jest określone działanie prawostronne „$*$" grupy G na zbiorze G (czyli grupa G działa prawymi przesunięciami na zbiorze G).

• Skoro jądro

$$\mathrm{Ker}(*) = \{g \in G \mid \forall_{x \in G} : x * g = x\} = \{g \in G \mid \forall_{x \in G} : x \cdot g = x\} = \{e\}$$

jest trywialne, to działanie „$*$" jest wierne.

• Orbita elementu $x \in G$ jest równa

$$G(x) = \{x * g \mid g \in G\} = \{x \cdot g \mid g \in G\} = xG = G,$$

czyli takie działanie „$*$" jest tranzytywne na zbiorze G.

• Dla elementu $x \in G$ jego stabilizator

$$St(x) = \{g \in G \mid x * g = x\} = \{g \in G \mid x \cdot g = x\} = \{e\}$$

jest trywialny.

Jeśli $\tau_g(x) = xg$, to reguła

$$G \ni g \mapsto \tau_g \in \mathbb{S}(X)$$

określa homomorfizm grup, który jest nazywany *prawostronnym regularnym przedstawieniem grupy* G.

Jeśli $(G, +)$ jest grupą addytywną, to działanie „$*$" zadajemy przez regułę

$$* : G \times G \ni x * g \mapsto x + g \in G.$$

(**2**) Reguła

$$* : G \times G \ni (g, x) \mapsto g * x = g \cdot x \cdot g^{-1} \in G$$

ma własności

$$e * x = exe^{-1} = x$$

oraz

$$(g \cdot h) * x = (g \cdot h) \cdot x \cdot (g \cdot h)^{-1} = g \cdot (h * x) \cdot g^{-1} = g * (h * x),$$

dla dowolnych $g, h, x \in G$, a więc „$*$" jest działaniem lewostronnym grupy G na zbiorze G. To działanie ma takie własności.

• Jądro

$$\mathrm{Ker}(*) = \{g \in G \mid \forall_{x \in G} : g * x = x\} = \{g \in G \mid \forall_{x \in G} : g \cdot x \cdot g^{-1} = x\} =$$
$$= \{g \in G \mid \forall_{x \in G} : g \cdot x = x \cdot g\} = Z(G),$$

gdzie zbiór $Z(G)$ jest nazywany *centrum* grupy G.

• Orbitą elementu $x \in G$ jest

$$G(x) = \{g * x \mid g \in G\} = \{g \cdot x \cdot g^{-1} \mid g \in G\} = x^G.$$

Podzbiór x^G grupy G jest nazywany *klasą elementów sprzężonych* z elementem x (=*klasą sprzężoności* elementu x) w grupie G. Orbita x^G elementu x jest jednoelementowa (czyli $x^G = \{x\}$) wtedy i tylko wtedy, gdy $x \in Z(G)$ jest centralny.

• Stabilizatorem elementu $x \in G$ jest

$$St(x) = \{g \in G \mid g * x = x\} = \{g \in G \mid g \cdot x \cdot g^{-1} = x\} = \{g \in G \mid g \cdot x = x \cdot g\} = C_G(x).$$

Podzbiór $C_G(x)$ jest nazywany *centralizatorem* elementu x w grupie G. Łatwo zauważyć, że $C_G(x) = G$ wtedy i tylko wtedy, gdy $x \in Z(G)$ jest centralny.

Proponujemy Czytelnikowi znaleźć warunki, kiedy takie działanie jest wierne i przechodnie.

(**3**) Niech X będzie zbiorem wszystkich podgrup grupy G. Reguła

$$* : G \times X \ni (g, A) \mapsto g * A = g \cdot A \cdot g^{-1} \in X$$

spełnia warunki:

a) skoro A jest podgrupą w G, to $g \cdot A \cdot g^{-1}$ również jest podgrupą w G. Dwie podgrupy $A,\ B$ grupy G są nazywane *sprzężonymi* w grupie G, jeśli istnieje element $g \in G$ taki, że $B = g \cdot A \cdot g^{-1}$.

b) Mamy $e * A = e \cdot A \cdot e^{-1} = A$ oraz

$$(g \cdot h) * A = (g \cdot h) \cdot A \cdot (g \cdot h)^{-1} = g \cdot (h \cdot A \cdot h^{-1}) \cdot g^{-1} = g \cdot (h * A) \cdot g^{-1} = g * (h * A)$$

dla dowolnych $g, h \in G$ oraz $A \in X$, a zatem „$*$" jest działaniem lewostronnym grupy G na zbiorze X. Zachodzą takie własności.

• Jądro tego działania

$$\mathrm{Ker}(*) = \{g \in G \mid \forall_{A \in X} : g*A = A\} = \{g \in G \mid \forall_{A \in X} : g \cdot A \cdot g^{-1} = A\} = \{g \in G \mid \forall_{A \in X} : g \cdot A = A \cdot g\}$$

• orbita elementu $A \in X$

$$G(A) = \{g * A \mid g \in G\} = \{g \cdot A \cdot g^{-1} \mid g \in G\}$$

jest *klasą sprzężonych z A podgrup* grupy G.

• stabilizator elementu $A \in X$

$$G_A = \{g \in G \mid g * A = A\} = \{g \in G \mid g \cdot A \cdot g^{-1} = A\} = \{g \in G \mid g \cdot A = A \cdot g\} = N_G(A).$$

Podzbiór $N_G(A)$ jest nazywany *normalizatorem* podgrupy A w grupie G. Zachodzi równość $G_A = G$ w tym i tylko tym przypadku, gdy A jest podgrupą normalną w G.

Proponujemy Czytelnikowi znaleźć warunki, kiedy takie działanie jest wierne i przechodnie.

(**4**) Przykładem klasycznym jest takie działanie „$*$" grupy przekształceń $\mathbb{S}(X)$ na zbiorze niepustym X: jeśli $\sigma \in \mathbb{S}(X)$ oraz $x \in X$, to

$$\sigma * i = \sigma(i).$$

W szczególności takim przykładem jest działanie grupy symetrycznej $\mathbb{S}_n$ na zbiorze $X = \{1, \ldots, n\}$.

(**5**) Rozpatrzmy przypadek cząstkowy poprzedniego przykładu. Niech $X = \{1, 2, 3, 4, 5, 6\}$,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 6 & 5 & 4 \end{pmatrix} \in \mathbb{S}_6$$

oraz $H = \langle \sigma \rangle$ będzie podgrupą cykliczną w $\mathbb{S}_6$. Jak zwykle przez e oznaczamy permutację tożsamościową. Wtedy reguła

$$\mu * i = \mu(i),$$

gdzie $\mu \in \langle\sigma\rangle$ oraz $i \in X$, zadaje działanie grupy H na zbiorze X. W rzeczy samej,

$$e * i = e(i) = i$$

oraz

$$(\mu\nu) * i = \mu\nu(i) = \mu(\nu(i)) = \mu(\nu * i) = \mu * (\nu * i)$$

dla dowolnych $\mu, \nu \in H$. Skoro rząd $|\sigma| = 6$, to $\sigma^6 = e$ oraz

$$H = \{e, \sigma, \sigma^2, \sigma^3, \sigma^4, \sigma^5\}.$$

Obliczamy, że:

- $$\begin{aligned} e(1) &= 1, \\ \sigma(1) &= 2, \\ \sigma^2(1) &= \sigma(2) = 3, \\ \sigma^3(1) &= \sigma(3) = 1, \\ \sigma^4(1) &= \sigma(1) = 2, \\ \sigma^5(1) &= \sigma(2) = 3, \end{aligned}$$

  a zatem stabilizator elementu $1 \in X$ jest równy

  $$H_1 = \{\mu \in H \mid \mu(1) = 1\} = \{e, \sigma^3\};$$

- $$\begin{aligned} e(2) &= 2, \\ \sigma(2) &= 3, \\ \sigma^2(2) &= \sigma(3) = 1, \\ \sigma^3(2) &= \sigma(1) = 2, \\ \sigma^4(2) &= \sigma(2) = 3, \\ \sigma^5(2) &= \sigma(3) = 1, \end{aligned}$$

  a więc stabilizator elementu $2 \in X$ jest równy

  $$H_2 = \{\mu \in H \mid \mu(2) = 2\} = \{e, \sigma^3\};$$

- $$\begin{aligned} e(3) &= 3, \\ \sigma(3) &= 1, \\ \sigma^2(3) &= \sigma(1) = 2, \\ \sigma^3(3) &= \sigma(2) = 3, \\ \sigma^4(3) &= \sigma(3) = 1, \\ \sigma^5(3) &= \sigma(1) = 2, \end{aligned}$$

  i wtedy stabilizatorem elementu $3 \in X$ jest

  $$H_3 = \{\mu \in H \mid \mu(3) = 3\} = \{e, \sigma^3\};$$

- $$\begin{aligned} e(4) &= 4, \\ \sigma(4) &= 6, \\ \sigma^2(4) &= \sigma(6) = 4, \\ \sigma^3(4) &= \sigma(4) = 6, \\ \sigma^4(4) &= \sigma(6) = 4, \\ \sigma^5(4) &= \sigma(4) = 6, \end{aligned}$$

  czyli stabilizator elementu $4 \in X$ to

  $$H_2 = \{\mu \in H \mid \mu(4) = 4\} = \{e, \sigma^2, \sigma^4\};$$

- $$\begin{array}{rcll} e(5) & = & 5, & \\ \sigma(5) & = & 5, & \\ \sigma^k(5) & = & 5 & \text{dla dowolnego } k \in \mathbb{Z}, \end{array}$$

  a więc stabilizatorem elementu $5 \in X$ jest

  $$H_5 = \{\mu \in H \mid \mu(5) = 5\} = \{e, \sigma, \sigma^2, \sigma^3, \sigma^4, \sigma^5\} = H;$$

- $$\begin{array}{rcl} e(6) & = & 6, \\ \sigma(6) & = & 4, \\ \sigma^2(6) & = & \sigma(4) = 6, \\ \sigma^3(6) & = & \sigma(6) = 4, \\ \sigma^4(6) & = & \sigma(4) = 6, \\ \sigma^5(6) & = & \sigma(6) = 4, \end{array}$$

  a to oznacza, że stabilizatorem elementu $6 \in X$ jest

  $$H_6 = \{\mu \in H \mid \mu(6) = 6\} = \{e, \sigma^3, \sigma^4\}.$$

Z powyższych obliczeń wynika, że orbity elementów ze zbioru X są takie:

$$\begin{array}{rcccl} H(1) & = & \{\sigma(1) \mid \sigma \in H\} & = & \{1, 2, 3\} = H(2) = H(3), \\ H(5) & = & \{\sigma(5) \mid \sigma \in H\} & = & \{5\}, \\ H(4) & = & \{\sigma(4) \mid \sigma \in H\} & = & \{4, 6\} = H(6). \end{array}$$

(**6**) Niech $X = \{gH \mid g \in G\}$ będzie zbiorem warstw lewostronnych grupy G względem jej podgrupy H oraz $g, h \in G$. Biorąc

$$g * hH = (gh)H,$$

zadajemy działanie „$*$" grupy G na zbiorze X (sprawdzić samodzielnie). Oczywiście, że:

- stabilizator

  $$St(gH) = \{a \in G \mid a * gH = gH\} = \{a \in G \mid agH = gH\} = \{a \in G \mid g^{-1}ag \in H\};$$

- orbita

  $$G(gH) = \{agH \mid a \in G\} = \{xH \mid x \in G\} = X,$$

  czyli działanie „$*$" jest przechodnie;
- jądro działania „$*$"

  $$\begin{array}{c} \mathrm{Ker}(*) = \{a \in G \mid agH = gH \text{ dla wszystkich } g \in G\} = \{a \in G \mid g^{-1}ag \in H \\ \text{dla wszystkich } g \in G\} = \{a \in G \mid a \in gHg^{-1},\ \forall_{g \in G}\} = \bigcap\{g^{-1}Hg \mid g \in G\} = H_G \end{array}$$

  jest największą podgrupą normalną grupy G zawierającą się w H (=*jądro podgrupy* H w grupie G).

Zostawiamy Czytelnikowi znalezienie warunków, kiedy to działanie będzie efektywne oraz pod jakim warunkiem działanie ma elementy stałe.

(**7**) Niech G będzie grupą oraz X będzie zbiorem niepustym. Jeśli

$$g * x = x$$

dla wszystkich $g \in G$ i $x \in X$, to będziemy mówić, że grupa G *działa trywialnie na zbiorze* X. W tym przypadku stabilizator $G_x = G$ oraz orbita $G(x) = \{x\}$ dla każdego $x \in X$ (to znaczy, że wszystkie elementy tego działania są stałe). Oprócz tego jądro tego działania

$$\mathrm{Ker}(*) = \{g \in G \mid g * x = x \text{ dla wszystkich } x \in X\} = G,$$

a zatem takie działanie nie jest efektywne.

(**8**) Rozpatrzmy regułę

$$* : \mathbb{Z} \times X \ni (n, x) \mapsto n * x = (-1)^n x \in X,$$

gdzie $X = \{-7, -\sqrt{7}, 0, \sqrt{7}, 7\}$. Wtedy 0 jest elementem neutralnym grupy addytywnej $(\mathbb{Z}, +)$ oraz

$$0 * x = (-1)^0 x = x$$

oraz

$$(a + b) * x = (-1)^{a+b} x = (-1)^a \cdot ((-1)^b x) = (-1)^a \cdot (b * x) = a * (b * x)$$

dla dowolnych $a, b \in \mathbb{Z}$ i $x \in X$. Zatem „$*$" jest działaniem lewostronnym grupy $\mathbb{Z}$ na zbiorze X.

• Jądro

$$\mathrm{Ker}(*) = \{a \in \mathbb{Z} \mid \forall_{x \in X} : a * x = x\} = \{a \in \mathbb{Z} \mid \forall_{x \in X} : (-1)^a x = x\} = 2\mathbb{Z}$$

jest zbiorem liczb parzystych całkowitych.

• Orbita elementu $x \in X$ jest taka

$$G(x) = \{a * x \mid a \in \mathbb{Z}\} = \{(-1)^a x \mid a \in \mathbb{Z}\} = \{-x, x\}.$$

• Stabilizator elementu $x \in X$

$$St(x) = \{a \in \mathbb{Z} \mid a * x = x\} = \{a \in \mathbb{Z} \mid (-1)^a x = x\} = \begin{cases} \mathbb{Z}, & \text{gdy } x = 0, \\ 2\mathbb{Z}, & \text{gdy } x \neq 0. \end{cases}$$

Lemat 3.8.4. *Niech $(G, \cdot)$ będzie grupą działającą na zbiorze niepustym X. Wtedy zbiór orbit $\{G(x) \mid x \in X\}$ tworzy rozbicie zbioru X.*

Dowód. Rozpatrzmy relację $\varrho \subseteq X \times X$ taką, że

$$(x, y) \in \varrho \Leftrightarrow \exists_{g \in G} : y = g * x,$$

gdzie „$*$" jest działaniem grupy G na zbiorze X.

Na podstawie lematu 3.8.1 relacja ϱ jest relacją równoważności na zbiorze X. Klasa równoważności $\varrho(x)$ z reprezentantem $x \in X$ ma postać

$$\varrho(x) = \{y \in X \mid (x, y) \in \varrho\} = \{g * x \mid g \in G\} = G(x)$$

i jest orbitą elementu x (względem działania „$*$"). Zatem zbiór orbit $\{\varrho(x) \mid x \in X\}$ jest rozbiciem zbioru X. □

Twierdzenie 3.8.5 (Poincarégo)**.** *Załóżmy, że grupa G działa na zbiorze niepustym X oraz $x \in X$. Wtedy zachodzą następujące własności:*

(*a*) *stabilizator $St(x)$ jest podgrupą w G;*

(*b*) *moc orbity $|G(x)| = |G : St(x)|$ jest indeksem stabilizatora $St(x)$ w grupie G;*

(*c*) *jeśli grupa G jest skończona, to:*

(1) $|X| = \sum_{i=1}^{n} |G : St(x_i)|$ *dla pewnych $x_1, \ldots, x_n \in X$;*

(2)
$$|G| = |G(x)| \cdot |St(x)|$$

(liczba $|G(x)|$ jest nazywana *długością orbity* $G(x)$).

Dowód. Niech e będzie elementem neutralnym grupy G.

(*a*) Skoro $e * x = x$, to $e \in G_x$. Jeśli $g, h \in G_x$, to

$$(gh) * x = g * (h * x) = g * x = x$$

oraz

$$x = e * x = (g^{-1}g) * x = g^{-1} * (g * x) = g^{-1} * x,$$

czyli $gh, g^{-1} \in G_x$. Na podstawie kryterium G_x jest podgrupą w G.

(*b*) Odwzorowanie

$$\varphi : G(x) \ni g * x \mapsto gSt(x) \in \{gSt(x) \mid g \in G\}$$

jest bijektywne (udowodnić samodzielnie).

(*c*) Na podstawie lematu 3.8.4 wnosimy, że

$$X = \bigcup_{i=1}^{n} G(x_i)$$

dla pewnych elementów $x_1, \ldots, x_n \in X$. Zatem biorąc pod uwagę własność (*b*), otrzymujemy

$$|X| = \sum_{i=1}^{n} |G(x_i)| = \sum_{i=1}^{n} |G : St(x_i)|.$$

Ponadto z twierdzenia Lagrange'a wynika, że

$$|G| = |St(x)| \cdot |G : St(x)| = |St(x)| \cdot |G(x)|$$

dla każdego $x \in X$. □

Wniosek 3.8.6 (kryterium równości orbit)**.** *Niech „$*$" będzie działaniem grupy G na zbiorze niepustym X oraz $x, y \in X$. Wtedy orbity $G(x) = G(y)$ są równe w tym i tylko tym przypadku, gdy $x = h * y$ dla pewnego $h \in G$.*

Dowód. ($\Rightarrow$) Załóżmy, że $G(x) = G(y)$. Z tego, że $x \in G(x)$ wynika, że $x \in G(y)$, czyli $x = h * y$ dla pewnego $h \in G$.

($\Leftarrow$) Mamy $x = h * y$ dla pewnego elementu $h \in G$. Wtedy dla dowolnego elementu $g \in G$ otrzymujemy

$$g * x = g * (h * y) = (gh) * y \in G(y),$$

czyli $G(x) \subseteq G(y)$. Odwrotnie

$$h^{-1} * x = h^{-1} * (h * y) = (h^{-1}h) * y = e * y = y,$$

czyli $y \in G(x)$. Wtedy

$$g * y = g * (h^{-1} * x) = (gh^{-1}) * x \in G(x)$$

dla każdego $g \in G$, a to oznacza, że $G(y) \subseteq G(x)$. Zatem mamy równe orbity $G(y) = G(x)$. □

Wniosek 3.8.7. *Niech G będzie grupą skończoną, działającą na skończonym zbiorze niepustym X, oraz $n \in \mathbb{N}^*$. Jeśli n dzieli indeks $|G : G_x|$ dla każdego $x \in X$, to n jest dzielnikiem mocy $|X|$ zbioru X.*

Dowód. Udowodnić samodzielnie. □

Bez dowodu zaznaczmy, że zachodzi takie

Twierdzenie 3.8.8. *Niech p będzie liczbą pierwszą. Jeśli G jest p-grupą skończoną, to $|G| = p^n$ dla pewnego $n \in \mathbb{N}$.*

Na tej podstawie możemy teraz udowodnić następne

Twierdzenie 3.8.9. *Niech p będzie liczbą pierwszą. Każda p-grupa skończona G ma nietrywialne centrum $Z(G)$.*

Dowód. Jak wynika z przykładu 3.8.3(2), mamy określone działanie „$*$" grupy G na zbiorze G takie, że $g * x = g \cdot x \cdot g^{-1}$ dla dowolnych $g, x \in G$. Na mocy lematu 3.8.4 oraz przykładu 3.8.3(2) grupa

$$G = \bigcup_{i=1}^{n} x_i^G$$

jest sumą mnogościową skończonej liczby parami rozłącznych klas sprzężoności pewnych elementów $x_1, \ldots, x_n \in G$. Wiemy, że dla elementu $x \in G$ zachodzi równoważność

$$|x^G| = 1 \quad \Leftrightarrow \quad x \in Z(G).$$

Skoro element neutralny $e \in Z(G)$ jest centralny, to $|Z(G)| \geqslant 1$. Każdy element centralny $x \in Z(G)$ tworzy zbiór jednoelementowy $x^G = \{x\}$. Jeśli $Z(G) = \{x_1, \ldots, x_k\}$, to

$$|G| = |Z(G)| + \sum_{i=k+1}^{n} |x_i^G|,$$

a więc

$$p^n = |Z(G)| + p^{m_{k+1}} + \cdots + p^{m_n}$$

dla pewnych $m_{k+1}, \ldots, m_n \in \mathbb{N}^*$. Z tego wynika, że liczba p dzieli $|Z(G)|$ i na tej podstawie $Z(G)$ jest nietrywialne. □

Twierdzenie 3.8.10. *Niech P będzie p-grupą skończoną, działającą na skończonym zbiorze niepustym X. Jeśli p dzieli moc $|X|$ oraz istnieje element stały (względem tego działania), to działanie posiada co najmniej p elementów stałych.*

□

■ Podzbiór

$$C_G(A) = \{g \in G \mid \forall_{a \in A} : a \cdot g = g \cdot a\}$$

jest nazywany *centralizatorem* podgrupy A w grupie $(G, \cdot)$.

Twierdzenie 3.8.11. *Niech G będzie grupą, A jej podgrupą oraz $x \in G$. Wtedy zachodzą następujące własności:*

(1) $Z(G) \lhd G$;
(2) $C_G(A) \leqslant G$;
(3) $N_G(A) \leqslant G$;
(4) $C_G(A) \leqslant N_G(A)$;
(5) $C_G(x) \leqslant G$;
(6) $C_G(x) = G \quad \Leftrightarrow \quad x \in Z(G)$;
(7) $C_G(A) = G \quad \Leftrightarrow \quad A \leqslant Z(G)$;
(8) $N_G(A) = G \quad \Leftrightarrow \quad A \lhd G$;
(9) *jeśli* $A \lhd G$, *to* $C_G(A) \lhd G$.

Dowód. Ćwiczenie. □

Zaznaczmy też takie

Twierdzenie 3.8.12. *Jeśli $(G, \cdot)$ jest grupą, to zachodzi izomorfizm grup*

$$\operatorname{Inn} G \cong G/Z(G).$$

Dowód. (Szkic) Oczywiście, że centrum $Z(G)$ jest podgrupą normalną w G oraz

$$\chi : G/Z(G) \ni aZ(G) \mapsto \varphi_a \in \operatorname{Inn} G \quad (a \in G),$$

gdzie φ_a jest określone w twierdzeniu 3.5.6, jest izomorfizmem grup (przekonać się samodzielnie). □

Ćwiczenia 3.8.13.

(**1**) Niech G będzie grupą skończoną, a $g_1, \ldots, g_n$ będą reprezentantami (po jednym) z każdej klasy jej elementów sprzężonych. Udowodnić, że jeśli reprezentanty parami komutują, to G jest abelowa.
(**2**) Załóżmy, że grupa G działa na zbiorze niepustym X. Jeśli A jest podgrupą abelową z G, która działa tranzytywnie na zbiorze X, to $C_G(A) = A$.

(**3**) Udowodnić, że jeśli H jest podgrupą grupy G oraz $G = \bigcup\limits_{x \in G} x^{-1}Hx$, to $H = G$.

(**4**) Udowodnić, że wzór $z * n = z + n$, gdzie $z, n \in \mathbb{Z}$, zadaje działanie wierne addytywnej grupy liczb całkowitych $\mathbb{Z}$ na zbiorze $\mathbb{Z}$.
(**5**) Udowodnić, że wzór $r * (x, y) = (x + ry, y)$ określa działanie „$*$” addytywnej grupy liczb rzeczywistych $(\mathbb{R}, +)$ na płaszczyźnie XOY.
(**6**) Udowodnić, że grupa G działa wiernie na zbiorze niepustym X wtedy i tylko wtedy, gdy jądro tego działania składa się dokładnie z jednego elementu neutralnego grupy G.

(**7**) Niech G będzie grupą skończoną rzędu n, a $Z(G)$ będzie jej centrum rzędu a. Załóżmy, że b jest liczbą elementów pewnej klasy elementów sprzężonych. Udowodnić, że b jest dzielnikiem liczby $\frac{n}{a}$.
(**8**) Znaleźć wszystkie grupy G (z dokładnością do izomorfizmu), które mają dokładnie dwie klasy elementów sprzężonych.
(**9**) Niech H, K będą podgrupami grupy G oraz $x, y \in G$. *Warstwą podwójną* z reprezentantem x (względem podgrup H i K) jest nazywany podzbiór $HxK = \{hxk \mid h \in H,\ k \in K\}$. Udowodnić, że:
(a) dwie warstwy HxK i HyK są równe lub rozłączne;
(b) liczba warstw lewostronnych w HxK względem podgrupy H jest równa $|K : K \cap x^{-1}Hx|$, a liczba warstw prawostronnych w HxK względem podgrupy K jest równa $|x^{-1}Hx : K \cap x^{-1}Hx|$;
(c) $|HxK| = |K| \cdot |H : H \cap x^{-1}Kx|$;
(d) $|HxK| = |H| \cdot |K : K \cap x^{-1}Hx|$.
(**10**) Niech H będzie podgrupą właściwą grupy skończonej G. Udowodnić, że $G \neq \bigcup\limits_{g \in G} g^{-1}Hg$.
(**11**) Udowodnić, że :
(a) $H = \{\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{C},\ ac \neq 0\}$ jest podgrupą grupy $GL_2(\mathbb{C})$;
(b) każdy element z $GL_2(\mathbb{C})$ jest sprzężony z pewnym elementem z H, czyli

$$GL_2(\mathbb{C}) = \bigcup_{g \in GL_2(\mathbb{C})} g^{-1}Hg.$$

Uwagi. Termin „działanie grupy na zbiorze” powstał jako uogólnienie własności grup permutacji.

3.9. Przykłady grup przekształceń

■ Przytoczmy przykłady grup z geometrii, analizy, fizyki i mechaniki. Zaznaczmy, że grupa izomorficzna z pewną podgrupą grupy $GL_n(\mathbb{F})$, gdzie $\mathbb{F}$ jest ciałem, jest nazywana *liniową.*

* * *

■ **Transformacje liniowo-frakcyjne rozszerzonej płaszczyzny zespolonej**. Zbiór

$$\mathbb{C}\bigcup\{\infty\}$$

jest nazywany *rozszerzoną płaszczyzną zespoloną* (lub *sferą zespoloną*, lub *sferą Riemanna*[10], lub *sferą Möbiusa*), a ∞ jest nazywane *punktem nieskończenie oddalonym.* Niech $a, b, c, d \in \mathbb{C}$ oraz

$$f_{a,b,c,d} : \mathbb{C}\bigcup\{\infty\} \to \mathbb{C}\bigcup\{\infty\}$$

jest takie, że

$$\begin{cases} f_{a,b,c,d}(z) &= \frac{az+b}{cz+d}, \text{ gdy } z \in \mathbb{C}, \\ f_{a,b,c,d}(\infty) &= \infty, \text{ gdy } c = 0, \\ f_{a,b,c,d}(\infty) &= \frac{a}{c}, \text{ gdy } c \neq 0, \\ f_{a,b,c,d}(-\frac{d}{c}) &= \infty. \end{cases}$$

■ Odwzorowanie liniowo-frakcyjne

$$f_{a,b,c,d} : \mathbb{C}\bigcup\{\infty\} \to \mathbb{C}\bigcup\{\infty\},$$

gdzie $ad - bc \neq 0$, jest bijekcją (i jest nazywane *przekształceniem Möbiusa*[11] lub *homografią*).

Lemat 3.9.1. *Zbiór przekształceń liniowo-frakcyjnych*

$$LF(\mathbb{C}) = \{f_{a,b,c,d} \mid a, b, c, d \in \mathbb{C} \text{ oraz } ad - bc \neq 0\}$$

tworzy grupę.

---

[10] Georg Friedrich Bernhard Riemann (1826–1866)

[11] August Ferdinand Möbius (1790–1868)

Dowód. Odwzorowanie jednostkowe $\mathrm{id}_{\mathbb{C}\cup\{\infty\}} : \mathbb{C}\cup\{\infty\} \to \mathbb{C}\cup\{\infty\}$ jest liniowo-frakcyjne, gdyż

$$\mathrm{id}_{\mathbb{C}\bigcup\{\infty\}}(z) = \frac{1z+0}{0z+1} = f_{1,0,0,1}(z)$$

w każdym punkcie $z \in \mathbb{C}$ oraz

$$\mathrm{id}_{\mathbb{C}\bigcup\{\infty\}}(\infty) = \infty = f_{1,0,0,1}(\infty).$$

Na mocy twierdzenia 1.4.6 złożenie „$\circ$" jest działaniem łącznym z elementem neutralnym $\mathrm{id}_{\mathbb{C}\cup\{\infty\}}$. Niech

$$w = \frac{dz-b}{-cz+a}.$$

Wtedy

$$(w \circ f_{a,b,c,d})(\mu) = \mathrm{id}_{\mathbb{C}\bigcup\{\infty\}}(\mu) = (f_{a,b,c,d} \circ w)(\mu)$$

dla każdego $\mu \in \mathbb{C}\cup\infty$ oraz

$$(f_{a,b,c,d})^{-1} = w = f_{d,-b,-c,a}.$$

Zatem $LF(\mathbb{C})$ jest grupą. □

■ Zostawiamy Czytelnikowi samodzielne znalezienie kryterium równości dwóch odwzorowań $f_{a,b,c,d}$ i $f_{u,v,w,t}$ z $LF(\mathbb{C})$. Odwzorowania liniowo-frakcyjne nad każdym ciałem $\mathbb{F}$ są konstruowane podobnie, czyli analogicznie otrzymujemy, że $LF(\mathbb{F})$ jest grupą, gdzie zbiór

$$\mathbb{F}\bigcup\{\infty\}$$

jest nazywany *prostą rzutową* nad ciałem $\mathbb{F}$. Na przykład dla ciała dwójkowego $\mathbb{F}_2$ grupa $LF(\mathbb{Z}_2)$ składa się z 6 elementów:

$$LF(\mathbb{F}_2) = \{f_{1,0,0,1}, f_{1,0,1,1}, f_{1,1,1,0}, f_{1,1,0,1}, f_{1,0,1,0}, f_{0,1,1,1}\}.$$

Przykłady 3.9.2.

(**1**) (**Grupa rzeczywistych odwzorowań liniowo-frakcyjnych**) Odwzorowanie $f : \mathbb{R} \to \mathbb{R}$ jest nazywane *liniowo-frakcyjnym,* jeśli istnieją takie $a, b, c, d \in \mathbb{R}$, że

$$f = \frac{aX + b}{cX + d},\ ad - bc \neq 0\ (X \in \mathbb{R}).$$

Niech $QL(\mathbb{R})$ będzie zbiorem wszystkich odwzorowań liniowo-frakcyjnych postaci $f : \mathbb{R} \to \mathbb{R}$. Ponieważ

$$\mathrm{id}_{\mathbb{R}} = \frac{1 \cdot X + 0}{0 \cdot X + 1},$$

to $\mathrm{id}_{\mathbb{R}} \in QL(\mathbb{R})$. Oznaczmy

$$\Delta = ad - bc \text{ oraz } g = \frac{dX - b}{-cX + a}.$$

Wtedy $g \in QL(\mathbb{R})$ oraz w każdym punkcie $x \in \mathbb{R}$ mamy

$$\begin{array}{rcccl} f(g(x)) & = & \frac{a(\frac{dx-b}{-cx+a})+b}{c(\frac{dx-b}{-cx+a})+d} = \frac{adx-ab-cxb+ab}{cdx-cb-cxd+ad} = \frac{\Delta x}{\Delta} = x & = & \mathrm{id}_{\mathbb{R}}(x), \\ g(f(x)) & = & \frac{d(\frac{ax+b}{cx+d})-b}{-c(\frac{ax+b}{cx+d})+a} = \frac{dax+db-cbx-db}{-cax-cb+cax+da} = \frac{\Delta x}{\Delta} = x & = & \mathrm{id}_{\mathbb{R}}(x). \end{array}$$

Zatem $g = f^{-1}$, czyli f jest bijekcją. Oprócz tego $f^{-1} = g \in QL(\mathbb{R})$. Jeżeli teraz $h \in QL(\mathbb{R})$, gdzie

$$h = \frac{uX + v}{wX + z},$$

$u, v, w, z \in \mathbb{R}$ oraz $uz - vw \neq 0$, to

$$(f \circ h)(x) = f(h(x)) = \frac{a(\frac{ux+v}{wx+z}) + b}{c(\frac{ux+v}{wx+z}) + d} = \frac{aux + av + bwx + bz}{cux + cv + dwx + dz} = \frac{(au + bw)x + (av + bz)}{(cu + dw)x + (cv + dz)}$$

dla każdego $x \in \mathbb{R}$. Skoro

$$(au + bw)(cv + dz) - (cu + dw)(av + bz) = (ad - bc)(uz - wv) \neq 0,$$

to $f \circ h \in QL(\mathbb{R})$. Na podstawie kryterium $QL(\mathbb{R})$ jest podgrupą grupy $\mathbb{S}(\mathbb{R})$. Grupa $QL(\mathbb{R})$ jest nazywana *grupą odwzorowań (przekształceń) liniowo-frakcyjnych* zbioru liczb rzeczywistych $\mathbb{R}$.

(**2**) (**Grupa rzeczywistych odwzorowań liniowych afinicznych**) Niech $a \in \mathbb{R} \setminus \{0\}$ oraz $b \in \mathbb{R}$. Odwzorowanie

$$f : \mathbb{R} \ni X \mapsto aX + b \in \mathbb{R}$$

jest nazywane *funkcją liniową.* Niech $L(\mathbb{R})$ będzie zbiorem wszystkich funkcji liniowych postaci $f : \mathbb{R} \to \mathbb{R}$, gdzie $a \neq 0$. Wtedy dla odwzorowania tożsamościowego zachodzi

$$\mathrm{id}_{\mathbb{R}} = 1 \cdot X + 0,$$

a zatem $\mathrm{id}_{\mathbb{R}}$ jest funkcją liniową. Jeśli $x_1 \neq x_2$ są różnymi liczbami rzeczywistymi, to

$$f(x_1) = ax_1 + b \neq ax_2 + b = f(x_2),$$

a więc f jest odwzorowaniem iniektywnym. Także $b = a \cdot 0 + b = f(0)$ dla każdego elementu $b \in \mathbb{R}$, czyli f jest funkcją suriektywną. Zatem $L(\mathbb{R}) \subseteq \mathbb{S}(\mathbb{R})$. Jeśli $h, g \in L(\mathbb{R})$, to istnieją takie $c, v \in \mathbb{R} \setminus \{0\}$ oraz $d, u \in \mathbb{R}$, że

$$h = cX + d,\ g = vX + u.$$

Wtedy

$$(g \circ h)(x) = g(h(x)) = g(cx + d) = v(cx + d) + u = vcx + (vd + u),$$

gdzie $vc \neq 0$ oraz $x \in \mathbb{R}$. Zatem $g \circ h \in L(\mathbb{R})$. Kończąc, jeśli $t = c^{-1}X - c^{-1}d \in L(\mathbb{R})$, to dla $x \in \mathbb{R}$ mamy

$$\begin{array}{rcccccl} t(h(x)) & = & t(cx+d) = c^{-1}(cx+d) - c^{-1}d & = & x & = & \mathrm{id}_{\mathbb{R}}(x), \\ h(t(x)) & = & h(c^{-1}x - c^{-1}d) = c(c^{-1}x - c^{-1}d) + d & = & x & = & \mathrm{id}_{\mathbb{R}}(x), \end{array}$$

co daje $h^{-1} = t \in L(\mathbb{R})$. Na podstawie kryterium $L(\mathbb{R})$ jest podgrupą w grupie symetrycznej $\mathbb{S}(\mathbb{R})$. Grupa $L(\mathbb{R})$ jest nazywana *grupą funkcji liniowych* na zbiorze $\mathbb{R}$. Oczywiście, że $L(\mathbb{R})$ jest podgrupą w $QL(\mathbb{R})$.

(**3**) Niech

$$L_1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f = X + b \text{ dla pewnego } b \in \mathbb{R}\}.$$

Wtedy $\mathrm{id}_{\mathbb{R}} = X + 0$, a zatem $\mathrm{id}_{\mathbb{R}} \in L_1(\mathbb{R})$. Jeśli $f = X + b$ oraz $g = X - b$, to dla $x \in \mathbb{R}$

$$\begin{array}{rcl} f(g(x)) & = & (x - b) + b = x = \mathrm{id}_{\mathbb{R}}(x), \\ g(f(x)) & = & (x + b) - b = x = \mathrm{id}_{\mathbb{R}}(x), \end{array}$$

a więc $f^{-1} = g \in L(\mathbb{R})$. Jeśli teraz $h = X + r$ dla pewnej liczby rzeczywistej r, to

$$(f \circ h)(x) = f(h(x)) = f(x + r) = (x + r) + b = x + (r + b)$$

dla wszystkich $x \in \mathbb{R}$, czyli $f \circ h \in L_1(\mathbb{R})$. Z tych rozumowań wynika, że $L_1(\mathbb{R})$ jest podgrupą w grupie $QL(\mathbb{R})$.

Weźmy dowolne elementy $f \in L_1(\mathbb{R})$ oraz $g \in QL(\mathbb{R})$, gdzie

$$f = X + c,\quad g = \frac{uX + v}{wX + z}$$

dla pewnych $c, u, v, w, z \in \mathbb{R}$, przy czym $uz - wv \neq 0$. Wtedy dla wartości argumentu $x \in \mathbb{R}$ obliczamy

$$\begin{array}{l} (g^{-1} \circ f \circ g)(x) = (g^{-1} \circ f)(\frac{ux+v}{wx+z}) = g^{-1}(\frac{ux+v}{wx+z} + c) = g^{-1}(\frac{ux+v+cwx+cz}{wx+z}) = \\ = \frac{z(\frac{ux+v+cwx+cz}{wx+z})-v}{-w(\frac{ux+v+cwx+cz}{wx+z})+u} = \\ = \frac{zux+zv+zcwx+cz^2-vwx-vz}{-wux-vw-cw^2x-cwz+uwx+zu} = \frac{x(zu+zcw-vw)+cz^2}{x(-cw^2)+(zu-vw-cwz)}. \end{array}$$

Jeżeli wziąć $c = 2$, $v = 0$, $w = u = z = 1$, to $g^{-1} \circ f \circ g = \frac{3X+2}{-2X-1}$, a więc $g^{-1} \circ f \circ g \notin L_1(\mathbb{R})$. To oznacza, że podgrupa $L_1(\mathbb{R})$ nie jest normalna w $QL(\mathbb{R})$.

Jeśli $f = X + c$ i $h = aX + b$, gdzie $a, b, c \in \mathbb{R}$ oraz $a \neq 0$, to

$$(h^{-1} \circ f \circ h)(x) = (h^{-1} \circ f)(ax + b) = h^{-1}(ax + b + c) = a^{-1}(ax + b + c) - a^{-1}b = x + a^{-1}c$$

dla $x \in \mathbb{R}$, a stąd $h^{-1} \circ f \circ h \in L_1(\mathbb{R})$. Zatem $L_1(\mathbb{R})$ jest podgrupą normalną w $L(\mathbb{R})$.

Możemy podobnie zdefiniować grupy $L(\mathbb{F})$ i $L_1(\mathbb{F})$ nad dowolnym ciałem $\mathbb{F}$.

(**4**) Niech

$$L(\mathbb{Q}) = \{f_{a,b} : \mathbb{Q} \to \mathbb{Q} \mid f_{a,b} = aX + b, \text{ gdzie } a, b \in \mathbb{Q}, a \neq 0\}.$$

Znajdźmy centralizator $C_{L(\mathbb{Q})}(f_{1,-1})$ elementu $f_{-1,1} \in L(\mathbb{Q})$. Niech $f_{a,b} \in C_{L(\mathbb{Q})}(f_{1,-1})$. W każdym punkcie $x \in \mathbb{Q}$ mamy

$$f_{a,b}(f_{1,-1}(x)) = (f_{a,b} \circ f_{1,-1})(x) = (f_{1,-1} \circ f_{a,b})(x) = f_{1,-1}(f_{a,b}(x)),$$

a stąd

$$a(x-1) + b = (ax + b) - 1,$$

czyli $-a + b = b - 1$ oraz $a = 1$. Zatem

$$C_{L(\mathbb{Q})}(f_{1,-1}) = \{f_{1,b} \mid b \in \mathbb{Q}\}.$$

Podzbiór

$$L_1(\mathbb{Q}) = \{f_{1,b} \in L(\mathbb{Q}) \mid b \in \mathbb{Q}\}$$

jest podgrupą normalną w $L(\mathbb{Q})$ (dowód jest podobny jak dla przypadku ciała $\mathbb{F} = \mathbb{R}$). Znajdźmy centralizator $C_{L(\mathbb{Q})}(L_1(\mathbb{Q}))$. Niech $f_{1,b} \in L_1(\mathbb{Q})$ oraz $f_{u,v} \in C_{L(\mathbb{Q})}(L_1(\mathbb{Q}))$. Wtedy w każdym punkcie $x \in \mathbb{Q}$ mamy

$$(f_{1,b} \circ f_{u,v})(x) = (f_{u,v} \circ f_{1,b})(x)$$

lub równoważnie

$$(ux + v) + b = u(x + b) + v,$$

skąd

$$v + b = ub + v$$

dla wszystkich $b \in \mathbb{Q}$. To oznacza, że $u = 1$ oraz

$$C_{L(\mathbb{Q})}(L_1(\mathbb{Q})) = \{f_{1,v} \in L(\mathbb{Q}) \mid v \in \mathbb{Q}\} = L_1(\mathbb{Q}),$$

czyli $L_1(\mathbb{Q})$ jest grupą abelową.

Znajdźmy normalizator $N_{L(\mathbb{Q})}(L_1(\mathbb{Q}))$. Mamy elementy dowolne $f_{1,b} \in L_1(\mathbb{Q})$ oraz $\varphi \in L(\mathbb{Q})$. Wtedy

$$\varphi = \frac{uX + v}{wX + t}$$

dla pewnych $u, v, w, t \in \mathbb{Q}$, przy czym $ut - wv \neq 0$. Skoro $\varphi^{-1} \circ f_{1,b} \circ \varphi \in L_1(\mathbb{Q})$, to

$$\varphi^{-1} \circ f_{1,b} \circ \varphi = f_{1,c}$$

dla pewnego współczynnika $c \in \mathbb{Q}$. Wtedy dla $x \in \mathbb{Q}$ otrzymujemy

$$(f_{1,b} \circ \varphi)(x) = (\varphi \circ f_{1,c})(x)$$

lub równoważnie

$$\frac{ux + v}{wx + t} + b = \frac{u(x + c) + v}{w(x + c) + t}.$$

Przepiszemy ostatnią równość w postaci układu

$$\begin{cases} bw^2 & = 0, \\ bwt & = 0, \\ vwc + bt^2 - tuc & = 0, \\ ut - wv & \neq 0, \\ b & \in \mathbb{Q}. \end{cases}$$

Z tego układu wynika, że $w = 0$ oraz $t \neq 0$, a zatem $\varphi \in L(\mathbb{Q})$. Wnosimy, że $N_{QL(\mathbb{Q})}(L_1(\mathbb{Q})) \subseteq L(\mathbb{Q})$. Podobnie jak w powyższym przykładzie (**3**), wynika, że $N_{L(\mathbb{Q})}(L_1(\mathbb{Q})) = L(\mathbb{Q})$. Zatem $N_{L(\mathbb{Q})}(L_1(\mathbb{Q})) = L(\mathbb{Q})$. Ten wynik można było otrzymać w taki sposób jak w przykładzie (**3**), czyli stosując kryterium normalności podgrupy, udowodnić, że $L_1(\mathbb{Q}) \lhd L(\mathbb{Q})$.

Uwagi. Grupa $LF(\mathbb{C})$ została przebadana przez G. Möbiusa w latach 1852–56, grupa $LF(\mathbb{R})$ była studiowana przez C. Staudta[(12)] w 1847 r., a grupa $LF(\mathbb{Z}_p)$ przez C. Gaussa w 1830 r.

* * *

■ **Grupa rzutowa liniowa**. Macierz postaci

$$\lambda I_2 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \in M_2(\mathbb{F}) \tag{3.5}$$

jest nazywana *skalarną* (stopnia 2 nad ciałem $\mathbb{F}$).

Lemat 3.9.3. *Niech $\mathbb{F}$ będzie ciałem. Centrum $Z(GL_2(\mathbb{F}))$ składa się ze wszystkich macierzy skalarnych postaci* (3.5)*, gdzie $\lambda \in \mathbb{F}$.*

Dowód. Niech $a, b, c, d \in \mathbb{F}$, przy czym $ad - bc \neq 0$. Znajdźmy takie elementy $x, y, z, t \in \mathbb{F}$, że

$$xt - zy \neq 0$$

oraz

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} x & y \\ z & t \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Przepiszemy ostatnią równość w postaci układu

$$\begin{cases} bz & = & yc, \\ ay + bt & = & xb + yd, \\ cx + dz & = & za + tc, \\ ad - bc & \neq & 0, \\ a, b, c, d \in \mathbb{F}. \end{cases}$$

[(12)] Christian von Staudt (1798–1867)

Biorąc $b = c = 1$, otrzymujemy $z = y$ oraz $x = t$. Wtedy łatwo zauważyć, że $y = z = 0$. □

■ Grupa ilorazowa

$$GL_2(\mathbb{F})/Z(GL_2(\mathbb{F}))$$

jest nazywana *ogólną grupą rzutową stopnia* 2 *nad cialem* $\mathbb{F}$ i oznaczana przez $PGL_2(\mathbb{F})$.

Lemat 3.9.4. *Nad każdym ciałem* $\mathbb{F}$ *grupy* $PGL_2(\mathbb{F})$ *oraz* $LF(\mathbb{F})$ *są izomorficzne.*

Dowód. Reguła

$$\varphi : GL_2(\mathbb{F}) \ni \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto f_{a,b,c,d} \in LF(\mathbb{F})$$

jest homomorfizmem grup z jądrem $\operatorname{Ker}\varphi = Z(GL_2(\mathbb{F}))$ i obrazem $\operatorname{Im}\varphi = LF(\mathbb{F})$. Na podstawie twierdzenia 3.6.3 wnosimy, że

$$LF(\mathbb{F}) \cong GL_2(\mathbb{F})/Z(GL_2(\mathbb{F})).$$

□

■ Podobnie grupa ilorazowa

$$SL_2(\mathbb{F})/Z(SL_2(\mathbb{F}))$$

jest oznaczana przez $PSL_2(\mathbb{F})$ i nazywana *szczególną grupą rzutową stopnia* 2 *nad ciałem* $\mathbb{F}$.

■ Proponujemy Czytelnikowi samodzielnie udowodnić, że

$$PGL_2(\mathbb{Z}_3) \cong \mathbb{S}_4 \quad \text{oraz} \quad PSL_2(\mathbb{Z}_3) \cong \mathbb{A}_4.$$

■ Skoro

$$Z(SL_2(\mathbb{F})) = SL_2(\mathbb{F}) \bigcap Z(GL_2(\mathbb{F})),$$

to $Z(SL_2(\mathbb{F}))$ składa się z macierzy skalarnych zawierających się w $SL_2(\mathbb{F})$. Jeśli charakterystyka ciała $p > 2$, to centrum

$$Z(SL_2(\mathbb{F}_p)) = \{\pm I\}$$

jest grupą o dwóch elementach. Grupa $PSL_2(\mathbb{F}_p)$ jest grupą przekształceń prostej rzutowej $\mathbb{F}_p \cup \{\infty\}$.

Lemat 3.9.5. *Grupa $PSL_2(\mathbb{F})$ jest izomorficzna z podgrupą S w grupie $LF(\mathbb{F})$, składającą się z odwzorowań liniowo-frakcyjnych $f_{a,b,c,d}$, dla których $ad - bc$ jest kwadratem*[13].

□

Uwagi. Matematyk amerykański E. Moore badał grupę $PSL_2(\mathbb{F})$ nad ciałami $\mathbb{F}$ rzędu > 3 w 1893 r. Homomorfizm z lematu 3.9.4 był rozpatrywany przez E. Galoisa i wykorzystywany przez A. Cayleya w 1880 r. w badaniach własności przekształceń liniowo-frakcyjnych.

* * *

■ **Izometrie**. Zwykle *izometrią* jest nazywane takie przekształcenie przestrzeni metrycznej, które zachowuje bez zmiany odległości między jej punktami. Jak wcześniej, $\mathbb{S}(X)$ jest grupą symetryczną zbioru niepustego X.

Izometrie przestrzeni $\mathbb{R}^n$. Niech $n \geqslant 1$ oraz $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Wtedy *odległością* między wektorami $\mathbf{x}$ i $\mathbf{y}$ jest nazywana liczba rzeczywista

$$|\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

gdzie

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

■ Odwzorowanie bijektywne $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ jest nazywane *izometrią* przestrzeni $\mathbb{R}^n$, jeśli

$$|\mathbf{x} - \mathbf{y}| = |\sigma(\mathbf{x}) - \sigma(\mathbf{y})|$$

dla dowolnych wektorów $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

[13] Element a ciała $\mathbb{F}$ jest nazywany *kwadratem*, jeśli $a = b^2$ dla pewnego $b \in \mathbb{F}$.

Twierdzenie 3.9.6. *Zbiór*

$$I(\mathbb{R}^n) = \{\sigma \in \mathbb{S}(\mathbb{R}^n) \mid \sigma \textit{ jest izometrią przestrzeni } \mathbb{R}^n\}$$

jest grupą względem złożenia (która jest nazywana *grupą izometrii* przestrzeni $\mathbb{R}^n$).

Dowód. Wystarczy pokazać, że $I(\mathbb{R}^n)$ jest podgrupą w grupie symetrycznej $\mathbb{S}(\mathbb{R}^n)$. Odwzorowanie tożsamościowe jest izometrią przestrzeni $\mathbb{R}^n$. Jeśli $\sigma, \tau \in I(\mathbb{R}^n)$, to przekonujemy się, że

$$\begin{aligned}&|(\sigma \circ \tau)(\mathbf{x}) - (\sigma \circ \tau)(\mathbf{y})| = |\sigma(\tau(\mathbf{x})) - \sigma(\tau(\mathbf{y}))| = \\ &= |\tau(\mathbf{x}) - \tau(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|,\end{aligned}$$

czyli $\sigma \circ \tau \in I(\mathbb{R}^n)$. Także $\mathbf{x} = \sigma(\mathbf{a})$ oraz $\mathbf{y} = \sigma(\mathbf{b})$ dla pewnych wektorów $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ i wtedy

$$|\sigma^{-1}(\mathbf{x}) - \sigma^{-1}(\mathbf{y})| = |\mathbf{a} - \mathbf{b}| = |\sigma(\mathbf{a}) - \sigma(\mathbf{b})| = |\mathbf{x} - \mathbf{y}|,$$

a więc $\sigma^{-1} \in I(\mathbb{R}^n)$. Na podstawie kryterium 2.5.2 wnosimy, że $I(\mathbb{R}^n)$ jest podgrupą grupy symetrycznej $\mathbb{S}(\mathbb{R}^n)$. □

■ Zbiór izometrii

$$I_0(\mathbb{R}^n) = \{\sigma \in I(\mathbb{R}^n) \mid \sigma(\mathbf{0}) = \mathbf{0}\}$$

przestrzeni $\mathbb{R}^n$, względem których wektor zerowy $\mathbf{0}$ jest niezmienniczy, jest podgrupą w $I(\mathbb{R}^n)$ (udowodnić samodzielnie).

■ Niech X będzie figurą geometryczną (ciałem geometrycznym) z $\mathbb{R}^n$, a $\sigma \in I(\mathbb{R}^n)$ będzie taką izometrią, że $\sigma(X) = X$. Wtedy odwzorowanie σ jest nazywane *symetrią* figury geometrycznej (ciała geometrycznego) X. Jeśli

$$\sigma \in I_0(\mathbb{R}^n) \text{ oraz } \sigma(X) = X,$$

to σ jest nazywane *symetrią centralną* figury geometrycznej (ciała geometrycznego) X.

Izometrie płaszczyzny. Odwzorowanie bijektywne $\sigma \in \mathbb{S}(\mathbb{C})$ takie, że

$$|z_1 - z_2| = |\sigma(z_1) - \sigma(z_2)|$$

dla dowolnych $z_1, z_2 \in \mathbb{C}$ jest nazywane *izometrią* płaszczyzny $\mathbb{R}^2$ (tutaj $z_j = x_j + iy_j$ dla liczb rzeczywistych x_j, y_j $(j = 1, 2)$ oraz liczba zespolona z_j jest rozpatrywana jako punkt $(x_j, y_j) \in \mathbb{R}^2$).

Lemat 3.9.7. *Zachodzą następujące własności:*

(1) *zbiór*

$$E = \{\sigma \in \mathbb{S}(\mathbb{C}) \mid \sigma \textit{ jest izometrią płaszczyzny } \mathbb{R}^2\}$$

jest grupą (która jest nazywana *grupą izometrii* lub *grupą euklidesową* płaszczyzny $\mathbb{R}^2$)*;*

(2) *zbiór*

$$D = \{\sigma \in E \mid \arg \tfrac{z_1 - z_2}{z_1 - z_3} = \arg \tfrac{\sigma(z_1) - \sigma(z_2)}{\sigma(z_1) - \sigma(z_3)} \textit{ dla dowolnych } z_1, z_2, z_3 \in \mathbb{C}\}$$

jest podgrupą w E (która jest nazywana *grupą izometrii zgodnych* płaszczyzny $\mathbb{R}^2$)*;*

(3) *każda izometria $\sigma \in E$ ma jedną z postaci:*

(*a*) (izometria zgodna)

$$\sigma : \mathbb{C} \ni z \mapsto e^{i\alpha} z + c \in \mathbb{C},$$

(*b*) (izometria niezgodna)

$$\sigma : \mathbb{C} \ni z \mapsto e^{i\alpha} \overline{z} + c \in \mathbb{C},$$

gdzie $\overline{z}$ jest sprzężeniem zespolonym liczby z, $\alpha = \arg z$ jest argumentem liczby $z \in \mathbb{C}$ oraz $c \in \mathbb{C}$.

Dowód. Ćwiczenie. □

Izometrie liczb rzeczywistych. Odwzorowanie bijektywne $\sigma \in \mathbb{S}(\mathbb{R})$ jest nazywane *izometrią* ciała liczb rzeczywistych $\mathbb{R}$, jeśli

$$|x - y| = |\sigma(x) - \sigma(y)|$$

dla dowolnych $x, y \in \mathbb{R}$ (czyli σ jest odwzorowaniem, względem którego odległość między punktami pozostaje niezmienniczą).

Lemat 3.9.8. *Zbiór $I(\mathbb{R})$, który tworzą izometrie ciała liczb rzeczywistych $\mathbb{R}$, jest grupą względem złożenia „$\circ$”* (która jest nazywana *grupą izometrii ciała* $\mathbb{R}$).

Dowód. Ćwiczenie. □

* * *

■ **Przesunięcia.** Odwzorowanie bijektywne $\sigma \in \mathbb{S}(\mathbb{R})$ jest nazywane *przesunięciem* ciała $\mathbb{R}$, jeśli

$$x - y = \sigma(x) - \sigma(y)$$

dla dowolnych $x, y \in \mathbb{R}$.

Lemat 3.9.9. *Zachodzą następujące własności:*
(1) *zbiór*

$$T(\mathbb{R}) = \{\sigma \in \mathbb{S}(\mathbb{R}) \mid \sigma \text{ jest przesunięciem ciała } \mathbb{R}\}$$

jest grupą (która jest nazywana *grupą przesunięć* ciała $\mathbb{R}$)*;*
(2) *dla każdego przesunięcia $\sigma \in T(\mathbb{R})$ ciała $\mathbb{R}$ znajdzie się taka liczba rzeczywista $a = a(\sigma)$, że*

$$\sigma(r) = r + a \ (r \in \mathbb{R}).$$

Dowód. (1) Przekonajmy się, że $T(\mathbb{R})$ jest podgrupą grupy symetrycznej $\mathbb{S}(\mathbb{R})$. Mamy $i_{\mathbb{R}} \in T(\mathbb{R})$. Jeśli $\sigma, \tau \in T(\mathbb{R})$, to

$$\begin{aligned}&(\sigma \circ \tau)(x) - (\sigma \circ \tau)(y) = \\ &= \sigma(\tau(x)) - \sigma(\tau(y)) = \tau(x) - \tau(y) = x - y,\end{aligned}$$

a zatem $\sigma \circ \tau \in T(\mathbb{R})$. Skoro $x = \sigma(a)$ oraz $y = \sigma(b)$ dla pewnych $a, b \in \mathbb{R}$, to

$$\sigma^{-1}(x) - \sigma^{-1}(y) = a - b = \sigma(a) - \sigma(b) = x - y,$$

czyli $\sigma^{-1} \in T(\mathbb{R})$. Wnosimy, że $T(\mathbb{R}) \leqslant \mathbb{S}(\mathbb{R})$.
(2) Udowodnić samodzielnie. □

■ Przekształcenie bijektywne $\sigma \in \mathbb{S}(\mathbb{C})$ jest nazywane *przesunięciem płaszczyzny* $\mathbb{R}^2$, jeśli

$$z_1 - z_2 = \sigma(z_1) - \sigma(z_2)$$

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

Lemat 3.9.10. *Zachodzą następujące własności:*
(1) *zbiór przesunięć płaszczyzny*

$$T(\mathbb{R}^2) = \{\sigma \in \mathbb{S}(\mathbb{C}) \mid \sigma \textit{ jest przesunięciem płaszczyzny } \mathbb{R}^2\}$$

tworzy grupę (która jest nazywana *grupą przesunięć płaszczyzny* $\mathbb{R}^2$)*;*
(2) *dla każdego przesunięcia* $\sigma \in T(\mathbb{R}^2)$ *płaszczyzny* $\mathbb{R}^2$ *znajdzie się taka liczba zespolona* $c = c(\sigma)$*, że*

$$\sigma(z) = z + c \ (z \in \mathbb{C}).$$

Dowód. Ćwiczenie. □

* * *

■ **Przekształcenia afiniczne.** Niech $A \in M_n(\mathbb{F})$ będzie ustaloną macierzą nieosobliwą stopnia n nad ciałem $\mathbb{F}$, L będzie przestrzenią liniową wymiaru n nad ciałem $\mathbb{F}$, a $\mathbf{c}$ będzie ustalonym wektorem z L. Wtedy odwzorowanie postaci

$$\sigma_{A,\mathbf{c}} : L \ni \mathbf{x} \mapsto A\mathbf{x} + \mathbf{c} \in L$$

jest nazywane *przekształceniem afinicznym* przestrzeni liniowej L.

Lemat 3.9.11. *Zbiór*

$$AG_n(L) = \{\sigma \in \mathbb{S}(\mathbb{L}) \mid \sigma \textit{ jest przekształceniem afinicznym przestrzeni liniowej } L\}$$

jest grupą (która jest nazywana *grupą przekształceń afinicznych* przestrzeni liniowej L wymiaru n nad ciałem $\mathbb{F}$).

Dowód. Odwzorowanie tożsamościowe $\mathrm{id}_L \in AG_n(L)$ jest afiniczne, gdyż

$$\mathrm{id}_L(\mathbf{x}) = I_n\mathbf{x} + \mathbf{0}$$

dla każdego $\mathbf{x} \in L$, gdzie I_n (odpowiednio $\mathbf{0}$) jest macierzą jednostkową stopnia n (odpowiednio wektorem zerowym) nad ciałem $\mathbb{F}$. Złożenie odwzorowań „$\circ$" jest działaniem łącznym z elementem neutralnym id_L. Znajdźmy warunek, gdy $\sigma_{A,\mathbf{c}} = \sigma_{B,\mathbf{d}}$, gdzie $A, B \in M_n(\mathbb{F})$ są ustalonymi macierzami nieosobliwymi, a $\mathbf{c}, \mathbf{d} \in L$ są ustalonymi wektorami. Ostatnią równość przepisujemy w postaci

$$A\mathbf{x} + \mathbf{c} = B\mathbf{x} + \mathbf{d}$$

dla dowolnego $\mathbf{x} \in L$, a zatem $\mathbf{c} = \mathbf{d}$ oraz $A = B$. Teraz znajdźmy takie $X \in M_n(\mathbb{F})$ oraz $\mathbf{u} \in L$, że

$$\sigma_{A,\mathbf{c}} \circ \sigma_{X,\mathbf{u}} = \mathrm{id}_L = \sigma_{X,\mathbf{u}} \circ \sigma_{A,\mathbf{c}},$$

czyli

$$\begin{cases} AX = I_n = XA, \\ A\mathbf{u} + \mathbf{c} = \mathbf{0}, \\ X\mathbf{c} + \mathbf{u} = \mathbf{0}, \end{cases}$$

a stąd

$$X = A^{-1} \text{ oraz } \mathbf{u} = -A^{-1}\mathbf{c}.$$

Zatem istnieje przekształcenie odwrotne

$$(\sigma_{A,\mathbf{c}})^{-1} = \sigma_{A^{-1},-A^{-1}\mathbf{c}}$$

i otrzymujemy, że $AG_n(L)$ jest grupą. □

■ Jeśli $\mathbb{F}$ jest ciałem, to grupa

$$AG_1(\mathbb{F}) = \{\sigma_{a,b} : \mathbb{F} \to \mathbb{F} \mid \sigma_{a,b}(X) = aX + b,\ a, b \in \mathbb{F} \text{ oraz } a \neq 0\}$$

jest nazywana *grupą przekształceń afinicznych* ciała $\mathbb{F}$.

Uwagi. Badanie przekształceń afinicznych zostało zapoczątkowane w pracach J. Plückera[(14)] i G. Möbiusa.

* * *

■ **Podobieństwo na płaszczyźnie.** Przekształcenie płaszczyzny $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$ jest nazywane *podobieństwem*, jeśli ono odwzoruje każdy trójkąt w podobny trójkąt, czyli jeśli ABC oraz $A'B'C'$ są trójkątami podobnymi oraz

$$\sigma(A) = A',\ \sigma(B) = B' \text{ i } \sigma(C) = C',$$

to stosunki długości odcinków

$$\frac{AB}{AC} = \frac{A'B'}{A'C'}$$

są równe. Odwzorowanie $\sigma \in \mathbb{S}(\mathbb{C})$ takie, że

$$\frac{z_1 - z_2}{z_1 - z_3} = \frac{\sigma(z_1) - \sigma(z_2)}{\sigma(z_1) - \sigma(z_3)}$$

dla dowolnych $z_1, z_2, z_3 \in \mathbb{C}$ jest nazywane *podobieństwem zgodnym.*

Lemat 3.9.12. *Zachodzą następujące własności:*

(1) *zbiór*

$$DS(\mathbb{C}) = \{\sigma \in \mathbb{S}(\mathbb{C}) \mid \sigma \textit{ jest podobieństwem zgodnym}\}$$

jest grupą.

(2) *dla każdego elementu $\sigma \in DS(\mathbb{C})$ znajdą się takie liczby $a, b \in \mathbb{C}$ $(a \neq 0)$, że*

$$\sigma(z) = az + b.$$

(14) Julius Plücker (1801–1868)

Dowód. Ćwiczenie. □

■ Odwzorowanie

$$\sigma : \mathbb{C} \ni z \mapsto a\overline{z} + b \in \mathbb{C},$$

gdzie $a, b \in \mathbb{C}$ $(a \neq 0)$ są ustalonymi liczbami zespolonymi, jest nazywane *podobieństwem niezgodnym* na płaszczyźnie. Proponujemy Czytelnikowi samodzielnie udowodnić, że podobieństwo na płaszczyźnie jest podobieństwem zgodnym lub niezgodnym. Pytanie do Czytelnika: *czy podobieństwa na płaszczyźnie tworzą grupę?*

* * *

■ Grupa Heisenberga[(15)]. Niech $\mathbb{F}$ będzie ciałem. Wtedy

$$H(\mathbb{F}) = \left\{ \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \mid a, b, c \in \mathbb{F} \right\}$$

jest podgrupą w grupie $SL_3(\mathbb{F})$, która jest nazywana *grupą Heisenberga nad ciałem* $\mathbb{F}$. Grupa $SL_3(\mathbb{R})$ odgrywa ważną rolę w mechanice kwantowej.

Ćwiczenia 3.9.13.

(**1**) Niech $G = \{I_2, A, BA, BA^2, BA^3, B, B^2, B^3\}$, gdzie $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Udowodnić, że G jest grupą nieabelową.
(**2**) Znaleźć rzędy wszystkich elementów grupy Heisenberga:
(a) $H(\mathbb{Z}_2)$;
(b) $H(\mathbb{Z}_3)$.
(**3**) Udowodnić, że każdy element niejednostkowy grupy Heisenberga $H(\mathbb{R})$ ma rząd nieskończony.

[(15)] Werner Carl Heisenberg (1901–1975)

3.10. Generatory grupy

Najpierw taki

Lemat 3.10.1. *Niech $(G,\cdot)$ będzie grupą, X będzie jej podzbiorem, a $\{H_i \leqslant G \mid X \subseteq H_i \ (i \in I)\}$ rodziną podgrup grupy G. Jeśli*

$$H_0 = \bigcap_{i\in I} H_i,$$

to zachodzą własności:
(1) *$X \subseteq H_0$;*
(2) *H_0 jest podgrupą grupy G.*

Dowód. (1) Wynika z definicji przecięcia.

(2) Element neutralny e grupy G mieści się w H_i dla każdego $i \in I$, a więc $e \in H_0$. Jeśli $a, b \in H_0$, to $a, b \in H_i$ dla każdego $i \in I$, a zatem też $a^{-1} \cdot b \in H_i$. Lecz wtedy $a^{-1} \cdot b \in H_0$ oraz $H_0 \leqslant G$ na podstawie kryterium 2.5.2. □

■ Jeśli X jest podzbiorem grupy G, to przecięcie wszystkich podgrup z G, zawierających X, jest nazywane *podgrupą generowaną przez podzbiór* X i oznaczane przez $\langle X\rangle$.

■ Podgrupa $\langle X\rangle$ jest „najmniejszą" z podgrup grupy G (względem relacji „$\subseteq$"), które zawierają podzbiór X. Jeśli $G = \langle X\rangle$, to X jest nazywany *zbiorem generatorów* grupy G.

■ Jeśli moc $\operatorname{card} X = 1$ zbioru generatorów X grupy G, to G jest cykliczna. Jeśli zaś $X = \{x_1, \ldots, x_m\}$ oraz $G = \langle X\rangle$, to będziemy notować

$$G = \langle x_1, \ldots, x_m\rangle;$$

wtedy grupa G jest nazywana *skończenie generowaną* z generatorami $x_1, \ldots, x_m$. W przypadku gdy $X = \varnothing$, to z definicji wynika, że $\langle X\rangle = \{e\}$, gdzie e jest elementem neutralnym grupy G.

Lemat 3.10.2. *Jeśli X jest podzbiorem niepustym grupy $(G,\cdot)$, to*

$$\langle X\rangle = \{x_1^{\varepsilon_1} \cdots x_l^{\varepsilon_l} \mid x_1, \ldots, x_l \in X,\\ \varepsilon_1, \ldots, \varepsilon_l \in \{-1, 0, 1\} \quad \textit{oraz} \quad l \in \mathbb{N}^*\}.$$

Dowód. W rzeczy samej, skoro $\langle X\rangle \leqslant G$, to $(\langle X\rangle, \cdot)$ jest grupą i $\langle X\rangle$ posiada wszystkie możliwe skończone iloczyny generatorów i ich odwrotności. □

■ Jeśli X jest zbiorem niepustym, to każdy element z $\langle X\rangle$ ma postać

$$x_1^{\alpha_1} \cdots x_l^{\alpha_l},$$

gdzie $x_1, \ldots, x_l \in X$ oraz $\alpha_1, \ldots, \alpha_l \in \mathbb{Z}$ dla pewnego $l \in \mathbb{N}^*$.

Przykłady 3.10.3.

(**1**) Ponieważ każda permutacja $\sigma \in \mathbb{S}_n$ jest iloczynem niezależnych cykli $\mathbb{S}_n$, to zbiór wszystkich cykli generuje grupę $\mathbb{S}_n$.

(**2**) Niech $(12), (123) \in \mathbb{S}_3$. Wtedy

$$\begin{array}{rcl} e & = & (12)^0(123)^0, \\ (12) & = & (12)^1(123)^0, \\ (23) & = & (12)^1(123)^1, \\ (13) & = & (12)^1(123)^{-1}, \\ (123) & = & (12)^0(123)^1, \\ (132) & = & (12)^0(123)^{-1}, \end{array}$$

a więc $\mathbb{S}_3 = \langle (12), (123)\rangle$. Proponujemy Czytelnikowi sprawdzić, że $\mathbb{S}_3 = \langle (12), (132)\rangle = \langle (23), (123)\rangle = \langle (23), (132)\rangle$ itd. To oznacza, że zbiór generatorów grupy jest określony niejednoznacznie.

(**3**) Udowodnijmy, że $L(\mathbb{Z}) = \{f_{a,b} : \mathbb{Z} \to \mathbb{Z} \mid f_{a,b}(X) = aX + b,\ b \in \mathbb{Z},\ a = \pm 1\}$ jest 2-generowana, czyli jest generowana przez dwa elementy. Mamy $f_{1,1}^{-1} = f_{1,-1}$. Jeśli $b \in \mathbb{Z} \setminus \{0\}$, to

$$f_{1,b} = f_{1,1}^b \text{ oraz } f_{-1,b} = f_{1,1}^b \circ f_{-1,0}.$$

Jako wniosek $L(\mathbb{Z}) = \langle f_{1,1}, f_{-1,0}\rangle$ też jest 2-generowana.

Twierdzenie 3.10.4. *Grupa alternująca* $\mathbb{A}_n$ $(n \geqslant 3)$ *jest generowana przez zbiór wszystkich cykli długości* 3.

Dowód. Każda permutacja $\sigma \in \mathbb{A}_n$ jest parzysta, a zatem jest iloczynem parzystej liczby transpozycji, czyli cykli długości 2. Ponieważ

$$\begin{array}{rcl} (a,b)(a,c) & = & (a,c,b), \\ (a,d)(b,c) & = & (a,b,c)(a,d,c), \end{array}$$

to σ rozkłada się w iloczyn cykli długości 3. Ponadto cykle długości 3 są permutacjami parzystymi, a więc zawierającymi się w $\mathbb{A}_n$. □

■ Zapisujemy

$$G = \langle X \mid \varphi(X) \rangle,$$

jeśli grupa G jest generowana przez zbiór elementów X, które spełniają warunek $\varphi(X)$.

Ćwiczenia 3.10.5.

(**1**) Udowodnić, że grupa $\langle x, y \mid x^2 = y^2 = (xy)^3 = 1 \rangle$ jest izomorficzna z grupą $\mathbb{S}_3$.
(**2**) Udowodnić, że $\mathbb{S}_n = \langle (12), (23 \dots n) \rangle$.
(**3**) Udowodnić, że $\langle a, b \mid a^4 = 1,\ b^2 = 1,\ abab = 1 \rangle$ ma rząd 8.
(**4**) Udowodnić, że $H = \langle x^2 \rangle$ jest podgrupą normalną grupy $G = \langle x, y \mid x^{22} = y^{15} = 1,\ xy = yx^3 \rangle$.
(**5**) Udowodnić, że elementy x^2, $y^{-1}x^2y$ oraz yx^2y^{-1} parami komutują w grupie $G = \langle x, y \mid x^3 = y^4 = (xy)^2 = 1 \rangle$.
(**6**) Udowodnić izomorficzność grup:
(*a*) $\langle x, y \mid x^3 = y^3 = xyxy = 1 \rangle$ oraz $\mathbb{A}_4$;
(*b*) $\langle a, b \mid a^4 = b^4 = bab^2a^2 = 1 \rangle$ oraz $\langle a, b \mid a^5 = b^4 = aba^3b^3 = 1 \rangle$.
(**7**) Udowodnić, że addytywna grupa liczb wymiernych nie jest skończenie generowana.
(**8**) Niech $G = \langle A, B \rangle$, gdzie

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ oraz } B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Udowodnić, że $G = \{I_2, A, A^2, A^3, B, AB, A^2B, A^3B\}$ jest podgrupą nieabelową rzędu 8 w grupie $GL_3(\mathbb{Q})$.
(**9**) Udowodnić, że przecięcie podgrup skończonych indeksów w abelowej grupie skończenie generowanej jest zerowe.
(**10**) Udowodnić, że zbiór wszystkich elementów skończonych rzędów abelowej grupy skończenie generowanej tworzy podgrupę.
(**11**) Udowodnić, że jeśli G jest grupą, to $G = \langle G \setminus \{1\} \rangle$.
(**12**) Udowodnić, że w grupie $\mathbb{S}_4$:
(*a*) $\langle (12), (12)(34) \rangle$ jest grupą niecykliczną rzędu 4;
(*b*) $\mathbb{S}_4 = \langle (2431), (2341) \rangle$.
(**13**) Udowodnić, że $\langle \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \rangle = SL_2(\mathbb{Z}_3)$.
(**14**) Udowodnić, że jeśli X jest podzbiorem w Y, gdzie Y jest podzbiorem grupy G, to $\langle X \rangle \subseteq \langle Y \rangle$. Skonstruować przykład grupy G z różnymi podzbiorami X, Y takimi, że $\langle X \rangle = \langle Y \rangle$.
(**15**) Niech A, B, C będą podgrupami grupy G, przy czym $C \lhd G$. Przez $\overline{X}$ oznaczmy obraz podgrupy $X \leqslant G$ w grupie ilorazowej $\overline{G} = G/C$. Udowodnić, że $\overline{\langle A, B \rangle} = \langle \overline{A}, \overline{B} \rangle$.
(**16**) Udowodnić, że funkcje $\{f_0, f_1, f_2, f_3, f_4, f_5\}$ generują skończoną grupę nieabelową. Znaleźć jej rząd, jeśli $f_0(X) = X$, $f_1(X) = \frac{1}{X}$, $f_2(X) = 1 - X$, $f_3(X) = \frac{1}{X-1}$, $f_4(X) = \frac{X-1}{X}$ oraz $f_5(X) = \frac{1}{1-X}$.
(**17**) Udowodnić, że grupa skończenie generowana ma tylko skończoną liczbę podgrup ustalonego indeksu skończonego.
(**18**) Udowodnić, że zbiór X generuje grupę symetryczną $\mathbb{S}_n$, jeśli:
(*a*) $X = \{(1, 2), (2, 3), \dots, (n-1, n)\}$;
(*b*) $X = \{(1, 2), (1, 3), \dots, (1, n)\}$.
(**19**) Udowodnić, że zbiór $X = \{(123), (124), \dots, (12n)\}$ generuje grupę alternującą $\mathbb{A}_n$.
(**20**) Udowodnić, że:
(*a*) $\mathbb{Q}^* = \langle \frac{1}{p} \mid p \text{ jest dowolną liczbą pierwszą} \rangle$;
(*b*) $\mathbb{Q} = \langle \frac{1}{n} \mid n \in \mathbb{N} \rangle$;

(*c*) $\mathbb{Q} = \langle -1, p \mid p \text{ jest dowolną liczbą pierwszą} \rangle$.

(**21**) Grupa $SL_2(\mathbb{Z}) = \{A \in M_2(\mathbb{Z}) \mid \det A = 1\}$ jest nazywana *ogólną grupą modularną*.

(*a*) Udowodnić, że $\Gamma(n) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) \mid a \equiv d \equiv 1 \pmod{n}, b \equiv c \equiv 0 \pmod{n}\}$ jest grupą ($n \geqslant 2$);

(*b*) Znaleźć rzędy wszystkich elementów podgrupy $\langle A, B \rangle \leqslant SL_2(\mathbb{Z})$, gdzie $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ oraz $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Uwagi. F. Klein, autor programu erlangeńskiego, poruszył ważną kwestię badania grup przekształceń. Podstawy klasycznej teorii grup przekształceń ciągłych założył S. Lie[16].

[16] Marius Sophus Lie (1842–1899)

3.11. Sumy proste i iloczyny proste grup

■ **Własności iloczynów prostych.**

Lemat 3.11.1. *Niech* $(G_1, *)$, $(G_2, \circ)$ *będą grupami oraz*

$$G = G_1 \times G_2 = \{(g_1, g_2) \mid g_1 \in G_1, g_2 \in G_2\}.$$

Wtedy $G = G_1 \times G_2$ *jest grupą względem działania „$\cdot$”, określonego wzorem*

$$(g_1, g_2) \cdot (h_1, h_2) = (g_1 * h_1, g_2 \circ h_2),$$

gdzie $g_1, h_1 \in G_1$ *oraz* $g_2, h_2 \in G_2$.

Dowód. Zostawiamy Czytelnikowi do samodzielnego sprawdzenia spełnienie wszystkich warunków z definicji grupy. □

■ Grupa

$$G = G_1 \times G_2$$

jest nazywana (zewnętrznym) *iloczynem prostym* grup G_1 oraz G_2. W przypadku gdy używamy notacji addytywnej dla grup G_1 i G_2, grupę G oznaczamy przez

$$G_1 \oplus G_2$$

i nazywamy ją (zewnętrzną) *sumą prostą* grup G_1 oraz G_2. Podobnym sposobem definiujemy iloczyn prosty grup G_1, G_2, ... , G_n (oznaczany przez

$$\prod_{i=1}^{n} G_i = G_1 \times G_2 \times \cdots \times G_n)$$

i sumę prostą grup G_1, G_2, ... , G_n (którą oznaczamy przez

$$\bigoplus_{i=1}^{n} G_i = G_1 \oplus G_2 \oplus \cdots \oplus G_n).$$

■ Niech $(G, \cdot)$ będzie grupą, a G_1, G_2 będą jej podgrupami. Będziemy mówić, że G jest (wewnętrznym) *iloczynem prostym* podgrup G_1 i G_2 (i oznaczać przez $G = G_1 \times G_2$), jeśli są spełnione warunki:

1) G_1, G_2 są podgrupami normalnymi w G;
2) $G = G_1 \cdot G_2$, gdzie iloczyn podgrup definiujemy wzorem

$$G_1 \cdot G_2 = \{g_1 \cdot g_2 \mid g_1 \in G_1 \text{ oraz } g_2 \in G_2\};$$

3) przecięcie $G_1 \cap G_2 = \langle e \rangle$ jest podgrupą jednostkową w G.

Podobnie w przypadku addytywnym mamy taką definicję.

■ Niech $(G, +)$ będzie grupą addytywną, a G_1, G_2 będą jej podgrupami. Będziemy mówić, że G jest (wewnętrzną) *sumą prostą* podgrup G_1 i G_2 (i oznaczać przez $G = G_1 \oplus G_2$), jeśli są spełnione warunki:
1) G_1, G_2 są podgrupami normalnymi w G;
2) $G = G_1 + G_2$, gdzie suma podgrup jest zdefiniowana wzorem

$$G_1 + G_2 = \{g_1 + g_2 \mid g_1 \in G_1 \text{ oraz } g_2 \in G_2\};$$

3) $G_1 \cap G_2 = \langle 0 \rangle$ jest podgrupą zerową w G.

Twierdzenie 3.11.2. *Niech G będzie grupą, a G_1, G_2 będą jej podgrupami. Wtedy zachodzą następujące własności:*

(1) *jeśli G jest iloczynem prostym (odpowiednio sumą prostą) podgrup G_1 i G_2, to każdy element $g_1 \in G_1$ komutuje z każdym elementem $g_2 \in G_2$;*

(2) (iloczyn prosty) *grupa G jest iloczynem prostym podgrup G_1 i G_2 w tym i tylko tym przypadku, gdy każdy element $g \in G$ ma dokładnie jedno przedstawienie postaci*

$$g = g_1 \cdot g_2,$$

gdzie $g_1 \in G_1$ oraz $g_2 \in G_2$, oraz elementy z G_1 są przemienne z elementami z G_2;

(2′) (suma prosta) *grupa G jest sumą prostą podgrup G_1 i G_2 w tym i tylko tym przypadku, gdy każdy element $g \in G$ ma dokładnie jedno przedstawienie postaci*

$$g = g_1 + g_2,$$

gdzie $g_1 \in G_1$ oraz $g_2 \in G_2$, oraz elementy z G_1 są przemienne z elementami z G_2.

Dowód. (1) Niech $G = G_1 \times G_2$ będzie iloczynem prostym podgrup G_1, G_2 oraz $g = g_1 \cdot g_2 \in G$, gdzie $g_1 \in G_1$ oraz $g_2 \in G_2$. W wyniku normalności podgrupy G_2 w grupie G mamy $g_1^{-1}g_2^{-1}g_1 \in G_2$, a więc $(g_1^{-1}g_2^{-1}g_1)g_2 \in G_2$. Natomiast z normalności podgrupy G_1 w G wynika, że $g_2^{-1}g_1g_2 \in G_1$, a zatem $g_1^{-1}(g_2^{-1}g_1g_2) \in G_1$. Otrzymujemy

$$g_1^{-1}g_2^{-1}g_1g_2 \in G_1 \bigcap G_2 = \langle e \rangle,$$

a więc

$$g_1^{-1}g_2^{-1}g_1g_2 = e.$$

Na tej podstawie wnosimy, że

$$g_1g_2 = g_2g_1.$$

(2) ($\Rightarrow$) Niech $G = G_1 \times G_2$ będzie iloczynem prostym podgrup G_1 i G_2. Załóżmy, że element $g \in G$ możemy rozłożyć na dwa sposoby

$$g = g_1g_2 = g_1'g_2',$$

gdzie $g_1, g_1' \in G_1$ oraz $g_2, g_2' \in G_2$. Mnożąc ostatnią równość z lewej strony przez g_1^{-1}, a z prawej strony przez $(g_2')^{-1}$, otrzymujemy

$$g_2(g_2')^{-1} = g_1^{-1}(g_1g_2)(g_2')^{-1} = g_1^{-1}(g_1'g_2')(g_2')^{-1} = g_1^{-1}g_1'.$$

Biorąc pod uwagę, że

$$g_2(g_2')^{-1} \in G_2,\ g_1^{-1}g_1' \in G_1$$

oraz $G_1 \bigcap G_2 = \langle e \rangle$, dostajemy

$$g_2(g_2')^{-1} = e \text{ oraz } g_1^{-1}g_1' = e,$$

na podstawie czego $g_2' = g_2$ i $g_1' = g_1$.

($\Leftarrow$) Teraz załóżmy, że każdy element $g \in G$ ma dokładnie jedno rozłożenie postaci $g = g_1 \cdot g_2$, gdzie $g_1 \in G_1$ oraz $g_2 \in G_2$. Wtedy $G = G_1 \cdot G_2$. Jeśli $h \in G_1 \cap G_2$, to

$$G_1G_2 \ni eh = h = he \in G_1G_2$$

i w wyniku jedności rozłożenia $h = e$. Udowodniliśmy zatem, że

$$G_1 \bigcap G_2 = \langle e \rangle$$

jest podgrupą jednostkową. Jeśli a_i jest dowolnym elementem z G_i $(i = 1, 2)$, to $g_1^{-1} a_1 g_1 \in G_1$ i na podstawie części (1) otrzymujemy

$$\begin{gathered} g^{-1} a_1 g = (g_1 g_2)^{-1} a_1 (g_1 g_2) = \\ = (g_2^{-1} g_1^{-1}) a_1 (g_1 g_2) = g_2^{-1} (g_1^{-1} a_1 g_1) g_2 = \\ = (g_1^{-1} a_1 g_1) g_2^{-1} g_2 = g_1^{-1} a_1 g_1 \in G_1 \end{gathered}$$

oraz

$$\begin{gathered} g^{-1} a_2 g = (g_1 g_2)^{-1} a_2 (g_1 g_2) = \\ = (g_2^{-1} g_1^{-1}) a_2 (g_1 g_2) = (g_2^{-1} a_2 g_2) g_1^{-1} g_1 = \\ = g_2^{-1} a_2 g_2 \in G_2, \end{gathered}$$

czyli G_1 i G_2 są podgrupami normalnymi w G. Zatem G jest iloczynem prostym podgrup G_1 i G_2.

$(2')$ Dowód podobny jak dla części (2). □

Kolejne twierdzenie pokazuje, że zewnętrzny iloczyn prosty (odpowiednio zewnętrzną sumę prostą) możemy traktować jako wewnętrzny iloczyn prosty (odpowiednio wewnętrzną sumę prostą) i na odwrót.

Twierdzenie 3.11.3. *Zachodzą następujące własności:*

(1) *jeśli $G = G_1 \times G_2$ jest zewnętrznym iloczynem prostym grup G_1 i G_2, to G jest wewnętrznym iloczynem prostym podgrup*

$$H_1 = \{(g_1, e_2) \mid g_1 \in G_1\}$$

oraz

$$H_2 = \{(e_1, g_2) \mid g_2 \in G_2\},$$

gdzie e_i jest elementem neutralnym grupy G_i oraz grupa H_i jest izomorficzna z grupą G_i $(i = 1, 2)$;

$(1')$ *jeśli $G = G_1 \oplus G_2$ jest zewnętrzną sumą prostą podgrup G_1 i G_2, to grupa G jest wewnętrzną sumą prostą podgrup*

$$H_1 = \{(g_1, 0_2) \mid g_1 \in G_1\}$$

oraz

$$H_2 = \{(0_1, g_2) \mid g_2 \in G_2\},$$

gdzie 0_i *jest zerem grupy* G_i *oraz grupa* H_i *jest izomorficzna z grupą* G_i $(i = 1, 2)$*;*

(2) *jeśli* G *jest wewnętrznym iloczynem prostym podgrup* H_1 *oraz* H_2*, to grupa* G *jest izomorficzna z zewnętrznym iloczynem prostym grup* H_1 *oraz* H_2*;*

(2′) *jeśli* G *jest wewnętrzną sumą prostą podgrup* H_1 *oraz* H_2*, to grupa* G *jest izomorficzna z zewnętrzną sumą prostą podgrup* H_1 *oraz* H_2*.*

Dowód. (1) Niech $(G, \cdot)$ będzie zewnętrznym iloczynem prostym grup $(G_1, *)$ i $(G_2, \circ)$. Jeśli $g = (g_1, g_2) \in G$, gdzie $g_1 \in G_1$ oraz $g_2 \in G_2$, to

$$g = (g_1, g_2) = (g_1, e_2) \cdot (e_1, g_2) \in H_1 H_2,$$

a zatem $G = H_1 H_2$. Oczywiście, że

$$H_1 \bigcap H_2 = \langle (e_1, e_2) \rangle$$

jest podgrupą jednostkową w G. Niech (h_1, e_2) będzie dowolnym elementem z H_1, a (e_1, h_2) dowolnym elementem z H_2. Wtedy

$$\begin{aligned} &g^{-1} \cdot (h_1, e_2) \cdot g = (g_1, g_2)^{-1} \cdot (h_1, e_2) \cdot (g_1, g_2) = \\ &= (g_1^{-1}, g_2^{-1}) \cdot (h_1, e_2) \cdot (g_1, g_2) = (g_1^{-1} * h_1 * g_1, g_2^{-1} \circ e_2 \circ g_2) = \\ &= (g_1^{-1} * h_1 * g_1, e_2) \in H_1 \end{aligned}$$

oraz

$$\begin{aligned} &g^{-1} \cdot (e_1, h_2) \cdot g = (g_1, g_2)^{-1} \cdot (e_1, h_2) \cdot (g_1, g_2) = \\ &= (g_1^{-1}, g_2^{-1}) \cdot (e_1, h_2) \cdot (g_1, g_2) = (g_1^{-1} * e_1 * g_1, g_2^{-1} \circ h_2 \circ g_2) = \\ &= (e_1, g_2^{-1} \circ h_2 \circ g_2) \in H_2, \end{aligned}$$

a zatem H_1, H_2 są podgrupami normalnymi w G. Na podstawie tego wnosimy, że G jest wewnętrznym iloczynem prostym podgrup H_1 i H_2.

(1′) Jest addytywnym analogiem własności (1).

(2) Skoro G jest wewnętrznym iloczynem prostym podgrup H_1 i H_2, to na mocy twierdzenia 3.11.2 każdy element $g \in G$ ma dokładnie jednoznaczne przedstawienie postaci $g = h_1h_2$, gdzie $h_1 \in H_1$ oraz $h_2 \in H_2$. Wtedy reguła

$$\begin{cases} \phi : G \to H_1 \times H_2, \\ \phi(g) = \phi(h_1h_2) = (h_1, h_2) \text{ (gdzie } g \in G) \end{cases}$$

jest odwzorowaniem bijektywnym. Oprócz tego, jeśli $g' = h_1'h_2'$ jest innym elementem z grupy G, to na podstawie twierdzenia 3.11.2 mamy

$$gg' = (h_1h_2)(h_1'h_2') = (h_1h_1')(h_2h_2')$$

i wtedy

$$\begin{aligned} &\phi(gg') = \phi((h_1h_1')(h_2h_2')) = (h_1h_1', h_2h_2') = \\ &= (h_1, h_2)(h_1', h_2') = \phi(h_1h_2)\phi(h_1'h_2') = \phi(g)\phi(g'). \end{aligned}$$

Zatem ϕ jest izomorfizmem grup G oraz $H_1 \times H_2$.

(2′) Jest addytywnym analogiem własności (2). □

Przytoczmy uogólnienie pojęcia iloczynu prostego podgrup dla przypadku $n \geqslant 2$ czynników (odpowiednio składników), gdzie $n \in \mathbb{N}$.

■ Będziemy mówić, że G jest *iloczynem prostym* podgrup $G_1, G_2, \ldots, G_n$ (i zapisywać $G = G_1 \times G_2 \times \ldots \times G_n = \times_{i=1}^{n} G_i$), jeśli są spełnione następujące warunki:

1) $G_1, G_2, \ldots, G_n$ są podgrupami normalnymi w G;

2)
$$G = G_1G_2 \cdots G_n,$$

gdzie iloczyn podgrup

$$G_1G_2 \cdots G_n = \{g_1g_2 \cdots g_n \mid g_i \in G_i \ (i = 1, \ldots, n)\};$$

3)
$$G_i \bigcap (G_1 \cdots \widehat{G_i} \cdots G_n) = \langle e \rangle$$

jest podgrupą jednostkową dla każdego $i = 1, \ldots, n$, gdzie symbol $\widehat{\ }$ nad czynnikiem G_i oznacza, że czynnik G_i jest nieobecny w iloczynie.

■ Podobnie będziemy mówić, że G jest *sumą prostą* podgrup $G_1, G_2, \ldots, G_n$ (i zapisywać $G = G_1 \oplus G_2 \oplus \ldots \oplus G_n = \oplus_{i=1}^{n} G_i$), jeśli są spełnione warunki:

1) $G_1, G_2, \ldots, G_n$ są podgrupami normalnymi w G;

2)
$$G = G_1 + G_2 + \cdots + G_n,$$
gdzie suma podgrup
$$G_1 + G_2 + \cdots + G_n = \{g_1 + g_2 + \cdots + g_n \mid g_i \in G_i \ (i = 1, \ldots, n)\};$$

3)
$$G_i \bigcap (G_1 + \cdots + \widehat{G_i} + \cdots + G_n) = \langle 0 \rangle$$
jest podgrupą zerową dla każdego $i = 1, \ldots, n$, gdzie symbol $\widehat{\ }$ nad składnikiem G_i oznacza, że G_i jest nieobecny w sumie.

■ Analog twierdzenia 3.11.3 zachodzi dla dowolnej liczby $n \in \mathbb{N}^*$ czynników (odpowiednio składników).

■ W wyniku twierdzenia 3.11.3 dalej `nie będziemy robić różnicy między iloczynami prostymi zewnętrznym a wewnętrznym (podobnie z sumami prostymi)` i nie będziemy korzystać ze słów „zewnętrzny" i „wewnętrzny", gdy mówimy o sumach czy iloczynach prostych.

Przykłady 3.11.4.

(**1**) Jeśli $r \in \mathbb{R}^*$, to
$$r = \pm 1 \cdot |r| \in \{-1, 1\} \cdot \mathbb{R}_+,$$
gdzie $\{-1, 1\} \cap \mathbb{R}_+ = \langle 1 \rangle$. Z abelowości grupy multiplikatywnej niezerowych liczb rzeczywistych $\mathbb{R}^*$ wynika, że
$$\{-1, 1\},\ \mathbb{R}_+ \lhd \mathbb{R}^*.$$
Zatem $\mathbb{R}^* = \{-1, 1\} \times \mathbb{R}_+$ jest iloczynem prostym podgrup $\{-1, 1\}$ oraz $\mathbb{R}_+$.

(**2**) Niech n będzie dodatnią liczbą całkowitą oraz $G = \{A \in GL_n(\mathbb{R}) \mid \det A > 0\}$,
$$H = \{\lambda I = \begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{bmatrix} \mid \lambda \in \mathbb{R}_+\}.$$
Nietrudno sprawdzić, że G jest grupą multiplikatywną, a H jej podgrupą. Jeśli $A \in G$, to istnieje A^{-1} oraz
$$A^{-1}(\lambda I)A = \lambda A^{-1} I A = \lambda A^{-1} A = \lambda I \in H,$$

czyli H jest podgrupą normalną w G. Ponadto $SL_n(\mathbb{R})$ jest podgrupą w G. Jeśli $A \in G$ i $B \in SL_n(\mathbb{R})$, to ze wzoru Cauchy'ego-Bineta

$$\det(A^{-1}BA) = \det(A^{-1}) \det B \det A = \frac{\det A}{\det A} \cdot \det B = 1,$$

a zatem $A^{-1}BA \in SL_n(\mathbb{R})$, czyli $SL_n(\mathbb{R})$ jest podgrupą normalną w G. Niech $\lambda = \det A$, gdzie $A \in G$. Wtedy $\lambda > 0$ oraz

$$A = (\frac{1}{\sqrt[n]{\lambda}} A)(\sqrt[n]{\lambda} I) \in SL_n(\mathbb{R}) \cdot H.$$

Załóżmy, że $C \in SL_n(\mathbb{R}) \cap H$. Wtedy $\det C = 1$ oraz $C = \lambda I$ dla pewnej dodatniej liczby rzeczywistej λ, a więc $\lambda = 1$ oraz

$$SL_n(\mathbb{R}) \bigcap H = \langle I \rangle.$$

Z przytoczonych rozumowań wynika, że

$$G = SL_n(\mathbb{R}) \times H$$

jest iloczynem prostym.

(**3**) Grupa addytywna $\mathbb{Z}_{30}$ jest sumą prostą podgrup $\langle \overline{15} \rangle$, $\langle \overline{10} \rangle$ i $\langle \overline{6} \rangle$:

$$\mathbb{Z}_{30} = \langle \overline{15} \rangle \oplus \langle \overline{10} \rangle \oplus \langle \overline{6} \rangle.$$

Patrząc z innej strony, podgrupy $\langle \overline{15} \rangle$, $\langle \overline{10} \rangle$ i $\langle \overline{6} \rangle$ są podgrupami cyklicznymi odpowiednio rzędów 2, 3 i 5, a zatem $\langle \overline{15} \rangle \cong \mathbb{Z}_2$, $\langle \overline{10} \rangle \cong \mathbb{Z}_3$ oraz $\langle \overline{6} \rangle \cong \mathbb{Z}_5$. Jako wniosek

$$\mathbb{Z}_{30} \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_5$$

jest sumą prostą grup $\mathbb{Z}_2$, $\mathbb{Z}_3$ i $\mathbb{Z}_5$.

(**4**) Przekonajmy się, że grupa addytywna liczb wymiernych $\mathbb{Q}$ nie jest sumą prostą dwóch podgrup właściwych A oraz B. Nie wprost. Załóżmy, że $\mathbb{Q} = A \oplus B$. Skoro istnieją liczby wymierne $\frac{a}{b}, \frac{u}{v} \in \mathbb{Q}$ takie, że

$$\frac{a}{b} \in A \quad \text{oraz} \quad \frac{u}{v} \in B,$$

to $-\frac{a}{b} \in A$, $-\frac{u}{v} \in B$ i możemy założyć, że wszystkie liczby całkowite a, b, u, v są dodatnie. Wtedy

$$a = \underbrace{\frac{a}{b} + \cdots + \frac{a}{b}}_{b \text{ składników}} \in A, \quad u = \underbrace{\frac{u}{v} + \cdots + \frac{u}{v}}_{v \text{ składników}} \in B,$$

w wyniku czego

$$A \ni \underbrace{a + \cdots + a}_{u \text{ składników}} = au = \underbrace{u + \cdots + u}_{a \text{ składników}} \in B.$$

Skoro $au \in A \cap B = \langle 0 \rangle$, to otrzymujemy sprzeczność. Zatem $(\mathbb{Q}, +)$ nie jest sumą prostą dwóch niezerowych podgrup właściwych. Natomiast grupę $\mathbb{Q}$ możemy rozłożyć w *trywialne* sumy proste

$$\mathbb{Q} = \mathbb{Q} \oplus \langle 0 \rangle = \langle 0 \rangle \oplus \mathbb{Q}.$$

* * *

■ **Iloczyn półprosty grup.** Będziemy mówić, że grupa G jest *iloczynem półprostym* podgrup A i B (i oznaczać przez $G = A \rtimes B$), jeśli:

(a) A jest podgrupą normalną, a B jest podgrupą w G;
(b) przecięcie $A \cap B = \langle 1 \rangle$ jest podgrupą jednostkową w G;
(c) $G = AB$ jest iloczynem podgrup A oraz B.

■ Iloczyn półprosty $A \rtimes B$ dwóch grup A i B jest cząstkowym przypadkiem iloczynu prostego, gdyż $A \lhd A \rtimes B$, $A \rtimes B = AB$ oraz $A \cap B = \langle 1 \rangle$.

Przykłady 3.11.5.

(**1**) Niech $\mathbb{S}_3$ będzie grupą symetryczną stopnia $n \geqslant 3$. Wtedy $\mathbb{S}_3 = \mathbb{A}_3 \rtimes \langle (12) \rangle$ jest iloczynem półprostym (lecz nie jest iloczynem prostym) grupy alternującej $\mathbb{A}_3$ i podgrupy cyklicznej $\langle (12) \rangle$, gdzie

$$(12) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \in \mathbb{S}_3.$$

(**2**) Rozpatrzmy grupę diedra

$$D_{2n} = \langle x, y \mid x^2 = y^2 = (xy)^n = 1 \rangle \ (n \geqslant 2).$$

Jeśli wziąć $a = xy$, to

$$D_{2n} = \langle x, a \mid x^2 = a^n = 1,\ x^{-1}ax = a^{-1} \rangle$$

(sprawdzić samodzielnie). Zatem

$$D_{2n} = \langle a \rangle \rtimes \langle x \rangle.$$

Ponieważ grupa D_{2n} jest nieabelowa, to nie jest iloczynem prostym dwóch właściwych podgrup niejednostkowych.

(**3**) Niech G będzie nieskończoną grupą diedralną, czyli

$$G = \langle x, y \mid x^2 = y^2 = 1 \rangle$$

(taka grupa jest oznaczana przez D_∞). Jeśli położyć $a = xy$, to

$$G = \langle x, a \mid x^{-1}ax = a^{-1},\ x^2 = 1 \rangle.$$

Wtedy

$$G = \langle a \rangle \rtimes \langle x \rangle$$

jest iloczynem półprostym normalnej podgrupy cyklicznej $\langle a \rangle$ rzędu nieskończonego i podgrupy cyklicznej $\langle x \rangle$ rzędu 2.

(**4**) Jak wiadomo, $SL_2(\mathbb{Q}) \lhd GL_2(\mathbb{Q})$. Jeśli

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL_2(\mathbb{Q}),$$

to $\Delta = \det A \neq 0$, a więc

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \frac{a}{\Delta} & b \\ \frac{c}{\Delta} & d \end{bmatrix} \cdot \begin{bmatrix} \Delta & 0 \\ 0 & 1 \end{bmatrix}.$$

Ponieważ

$$G = \{\begin{bmatrix} \Delta & 0 \\ 0 & 1 \end{bmatrix} \mid \Delta \in \mathbb{Q} \setminus \{0\}\}$$

jest grupą multiplikatywną izomorficzną z $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ (udowodnić samodzielnie) oraz

$$\begin{bmatrix} \frac{a}{\Delta} & b \\ \frac{c}{\Delta} & d \end{bmatrix} \in SL_2(\mathbb{Q}) \text{ oraz } G \bigcap SL_2(\mathbb{Q}) = \{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\},$$

to

$$GL_2(\mathbb{Q}) \cong SL_2(\mathbb{Q}) \rtimes \mathbb{Q}^*$$

jest iloczynem półprostym. Proponujemy Czytelnikowi samodzielnie przekonać się, że podgrupa G nie jest normalna w $GL_2(\mathbb{Q})$.

(**5**) Rozpatrzmy grupę

$$G = \{\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \mid a, b \in \mathbb{R}, a \neq 0\} \leqslant GL_2(\mathbb{R}).$$

Oznaczmy

$$A = \{\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \mid b \in \mathbb{R}\} \text{ oraz } B = \{\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \mid a \in \mathbb{R}^*\}.$$

Wtedy A jest grupą multiplikatywną izomorficzną z grupą $\mathbb{R}^+$, a B jest grupą multiplikatywną izomorficzną z grupą $\mathbb{R}^*$ (udowodnić samodzielnie). Jeśli

$$X = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$$

jest dowolnym elementem grupy G, to $a \neq 0$ oraz

$$X = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \in AB.$$

Ponieważ

$$A \bigcap B = \{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\}$$

oraz $A \lhd G$, to $G = A \rtimes B$ (udowodnić samodzielnie, że $A \lhd G$, a B nie jest normalna w G). Zatem

$$G = A \rtimes B \cong \mathbb{R}^+ \rtimes \mathbb{R}^*.$$

Ćwiczenia 3.11.6.

(**1**) Udowodnić, że jeśli H_i jest podgrupą normalną w G_i $(i = 1, \ldots, n)$, to:
(a) $(H_1 \times \cdots \times H_n) \lhd (G_1 \times \cdots \times G_n)$;
(b) $(G_1 \times \cdots \times G_n)/(H_1 \times \cdots \times H_n) \cong (G_1/H_1) \times \cdots \times (G_n/H_n)$.
(**2**) Udowodnić izomorficzność grup:
(a) $\mathbb{C}^* \cong \mathbb{R}_+ \times \mathbb{S}^1$;
(b) $\mathbb{C}^* \cong \mathbb{R}_+ \times [0, 2\pi)$;
(c) $\mathbb{C} \cong \mathbb{R} \oplus \mathbb{R}$;

(*d*) $C_{[0,1]} \cong A \oplus B$, gdzie

$$A = \{a \in C_{[0,1]} \mid \int_0^1 a(t)dt = 0\} \text{ oraz } B = \{b \in C_{[0,1]} \mid b(x) = b(0) \text{ dla dowolnych } x \in [0,1]\};$$

(*e*) $\mathbb{R}^n \cong \mathbb{R}^m \oplus \mathbb{R}^{n-m}$ dla $n, m \in \mathbb{N}^*$ oraz $n > m$.
(**3**) Udowodnić, że grupa G nie jest iloczynem prostym (sumą prostą) dwóch podgrup właściwych, jeśli:
(*a*) G jest grupą cykliczną rzędu p^n, gdzie p jest liczbą pierwszą;
(*b*) $G = \mathbb{Z}$;
(*c*) $G = \mathbb{S}_3$;
(*d*) $G = \mathbb{A}_4$;
(*e*) $G = \mathbb{S}_4$.

Uwagi. W teorii grup iloczyn prosty został wprowadzony w pracach A. Cauchy'ego z 1845 r. i C. Jordana z 1870 r.

Rozdział 4

Pierścienie i ciała

Ważne miejsce we współczesnej matematyce zajmują struktury algebraiczne z dwoma działaniami (algebry, pierścienie i ich różne uogólnienia), połączonymi prawem rozdzielności.

4.1. Ideały

■ Niech $(R, +, \cdot)$ będzie pierścieniem. Zbiór I jest nazywany:
- *ideałem lewostronnym* pierścienia R, jeśli:
  1) $(I, +)$ jest podgrupą grupy addytywnej $(R, +)$;
  2_l) $ri \in I$ dla dowolnych elementów $r \in R$ oraz $i \in I$;
- *ideałem prawostronnym* pierścienia R, jeśli:
  1) $(I, +)$ jest podgrupą grupy addytywnej $(R, +)$;
  2_r) $ir \in I$ dla dowolnych elementów $r \in R$ oraz $i \in I$;
- *ideałem obustronnym* pierścienia R, jeśli:
  1) $(I, +)$ jest podgrupą grupy addytywnej $(R, +)$;
  2) $ri, ir \in I$ dla dowolnych elementów $r \in R$ oraz $i \in I$.

■ Zamiast terminu „ideał obustronny" będziemy wykorzystywać krótszy termin „ideał". Czasem ideał prawostronny i ideał lewostronny zręcznie nazywać jednym terminem *ideał jednostronny*.

Twierdzenie 4.1.1 (kryterium ideału)**.** *Niech $(R, +, \cdot)$ będzie pierścieniem.*

(a) *Zbiór I jest ideałem prawostronnym w R w tym i tylko tym przypadku, gdy są spełnione własności:*
(0_1) $I \neq \varnothing$;

(0_2) $I \subseteq R$;
(1) $a - b \in I$ *dla dowolnych elementów* $a, b \in I$;
(2_r) $ar \in I$ *dla dowolnych* $a \in I$ *oraz* $r \in R$.

(*b*) *Zbiór* I *jest ideałem lewostronnym w* R *w tym i tylko tym przypadku, gdy są spełnione własności:*
(0_1) $I \neq \varnothing$;
(0_2) $I \subseteq R$;
(1) $a - b \in I$ *dla dowolnych elementów* $a, b \in I$;
(2_l) $ra \in I$ *dla dowolnych* $a \in I$ *oraz* $r \in R$.

(*c*) *Zbiór* I *jest ideałem obustronnym w* R *w tym i tylko tym przypadku, gdy są spełnione własności:*
(0_1) $I \neq \varnothing$;
(0_2) $I \subseteq R$;
(1) $a - b \in I$ *dla dowolnych elementów* $a, b \in I$;
(2) $ra, ar \in I$ *dla dowolnych* $a \in I$ *oraz* $r \in R$.

Dowód. (*a*) ($\Rightarrow$) Niech I będzie ideałem prawostronnym w R. Wtedy własności (0_1), (0_2), (1) oraz (2_r) są konsekwencjami własności z definicji oraz kryterium podgrupy (patrz twierdzenie 2.5.2).

($\Leftarrow$) Załóżmy, że zbiór I spełnia warunki (0_1), (0_2), (1) oraz (2_r). Wtedy z twierdzenia 2.5.2 otrzymujemy, że I jest podgrupą grupy addytywnej $(R, +)$, a więc I jest ideałem prawostronnym w R.

(*b*) oraz (*c*) mają podobne dowody. □

Przykłady 4.1.2.

(**1**) Każdy pierścień R ma ideały *trywialne*: ideał *zerowy* $\{0\}$, który składa się dokładnie z jednego (zerowego) elementu 0, oraz ideał *niewłaściwy* R. Jeśli $I \neq R$, to ideał obustronny (czy jednostronny) I jest nazywany *właściwym*.

(**2**) Pierścień liczb całkowitych $\mathbb{Z}$ zawiera ideał $n\mathbb{Z}$ dla każdej liczby całkowitej n, przy czym $n\mathbb{Z} = (-n)\mathbb{Z}$ oraz $0\mathbb{Z} = \{0\}$.

(**3**) Niech

$$I = \{f \in C_{[0,1]} \mid f(0) = f(1) = 0\}.$$

Dla dowolnych elementów $f, g \in I$ oraz $r \in C_{[0,1]}$ otrzymujemy

$$\begin{array}{lcl} (f-g)(0) & = & f(0) - g(0) = 0, \\ (f-g)(1) & = & f(1) - g(1) = 0, \\ (rf)(0) & = & r(0)f(0) = r(0)0 = 0, \\ (rf)(1) & = & r(1)f(1) = r(1)0 = 0, \end{array}$$

a zatem $f-g,\ rf \in I$. Wnosimy, że I jest ideałem w pierścieniu (przemiennym) funkcji ciągłych $C_{[0,1]}$.

(**4**) Jeśli I jest ideałem jednostronnym pierścienia R, to I jest podpierścieniem w R. Odwrotna implikacja nie zachodzi w ogólnym przypadku, bo, na przykład, $\mathbb{Z}$ jest podpierścieniem w pierścieniu liczb wymiernych $\mathbb{Q}$, lecz $\frac{1}{2}\cdot 1 \notin \mathbb{Z}$, czyli $\mathbb{Z}$ nie jest ideałem w $\mathbb{Q}$.

Lemat 4.1.3. *Niech R będzie pierścieniem oraz $a \in R$. Wtedy zachodzą następujące własności:*

(1)
$$\langle a\rangle_r = aR + a\mathbb{Z} = \{ar + az \mid r \in R \text{ oraz } z \in \mathbb{Z}\}$$
jest ideałem prawostronnym w R (który jest nazywany *ideałem prawostronnym głównym* generowanym przez element a)*;*

(2)
$$\langle a\rangle_l = Ra + a\mathbb{Z} = \{ra + za \mid r \in R \text{ oraz } z \in \mathbb{Z}\}$$
jest ideałem lewostronnym w R (który jest nazywany *ideałem lewostronnym głównym* generowanym przez element a)*;*

(3)
$$\langle a\rangle = aR + Ra + RaR + a\mathbb{Z} = \{ar + ta + \sum_{i=1}^{n} u_i a w_i + az \mid r, t, u_i, w_i \in R,\quad z \in \mathbb{Z}\ (i = 1, \dots n;\ n \in \mathbb{N}^*)\}$$
jest ideałem obustronnym w R (który jest nazywany *ideałem obustronnym głównym* generowanym przez element a).

Dowód. (1) W rzeczy samej, $a = a\cdot 0 + a\cdot 1 \in \langle a\rangle_r$ oraz $\langle a\rangle_r \subseteq R$. Jeśli $\alpha, \beta \in \langle a\rangle_r$, to
$$\alpha = at_1 + an_1 \text{ oraz } \beta = at_2 + an_2$$
dla pewnych $t_1, t_2 \in R$ oraz $n_1, n_2 \in \mathbb{Z}$. Wtedy
$$\begin{aligned} \alpha - \beta &= a(t_1 - t_2) + a(n_1 - n_2) \in \langle a\rangle_r, \\ \alpha q &= a(t_1 q + n_1 q) \in \langle a\rangle_r \end{aligned}$$
dla dowolnego $q \in R$. Zatem $\langle a\rangle_r$ jest ideałem prawostronnym w R.

(2) i (3) mają podobne dowody. □

■ Dla oznaczenia ideałów głównych stosujemy podobne oznaczenie jak dla grup cyklicznych. Jeśli z kontekstu wiadomo, o co chodzi, to nie powoduje kolizji.

■ Jeśli pierścień R posiada jedność 1, to zachodzą równości

$$\begin{aligned} \langle a\rangle_r &= aR, \\ \langle a\rangle_l &= Ra, \\ \langle a\rangle &= aR + Ra + RaR = RaR. \end{aligned}$$

■ Ideał zerowy $\langle 0\rangle$ jest główny (i krótko będziemy go oznaczać przez 0).

Lemat 4.1.4. *Niech R będzie pierścieniem przemiennym z jednością 1 oraz $a \in R$. Wtedy element a jest odwracalny w R w tym i tylko tym przypadku, gdy $\langle a\rangle = R$.*

Dowód. ($\Rightarrow$) Jeśli element a jest odwracalny w R, to $au = 1$ dla pewnego elementu $u \in R$. Wtedy $1 \in \langle a\rangle$ i dlatego $r = r \cdot 1 \in \langle a\rangle$ dla każdego $r \in R$. Stąd wynika, że $R = \langle a\rangle$.

($\Leftarrow$) Jeśli $R = \langle a\rangle$, to $ua = au = 1$ dla pewnego elementu $u \in R$, a więc a jest odwracalny w R. □

■ Każde ciało ma dokładnie dwa ideały, które są trywialne.

* * *

■ **Działania na ideałach.** Niech R będzie pierścieniem, a I, J będą jego ideałami lewostronnymi (odpowiednio prawostronnymi czy obustronnymi). Wtedy:

a) *przecięciem* ideałów jest zbiór

$$I \bigcap J = \{a \in R \mid a \in I \text{ oraz } a \in J\};$$

b) *sumą* (algebraiczną) ideałów jest

$$I + J = \{i + j \mid i \in I \text{ oraz } \ j \in J\};$$

c) *iloczynem* ideałów jest

$$IJ = \{\sum_{s=1}^{n} i_s j_s \mid i_s \in I \text{ oraz } \ j_s \in J \ (s = 1, \ldots, n;\ n \in \mathbb{N}^*)\};$$

d) *sumą mnogościową* ideałów jest zbiór

$$I \bigcup J = \{a \in R \mid a \in I \text{ lub } a \in J\}.$$

Lemat 4.1.5. *Niech R będzie pierścieniem, a I, J będą jego ideałami lewostronnymi* (*odpowiednio prawostronnymi lub obustronnymi*). *Wtedy $I + J$, $I \cap J$ oraz IJ też są ideałami lewostronnymi* (*odpowiednio prawostronnymi lub obustronnymi*) *w R.*

Dowód polega na sprawdzeniu warunków z kryterium ideału, co zostawiamy Czytelnikowi.

□

Przykłady 4.1.6.

(**1**) Suma mnogościowa dwóch jednostronnych (odpowiednio obustronnych) ideałów niekoniecznie musi być ideałem jednostronnym (odpowiednio obustronnym). Na przykład $2\mathbb{Z}$ oraz $7\mathbb{Z}$ są ideałami w pierścieniu liczb całkowitych $\mathbb{Z}$, lecz ich suma mnogościowa $2\mathbb{Z} \cup 7\mathbb{Z}$ nie jest już ideałem w $\mathbb{Z}$, gdyż $2 + 7 \notin 2\mathbb{Z} \cup 7\mathbb{Z}$.

(**2**) Dla ideałów $2\mathbb{Z}$ oraz $7\mathbb{Z}$ w pierścieniu $\mathbb{Z}$ obliczamy, że

$$\begin{array}{rcl} 2\mathbb{Z} \bigcap 7\mathbb{Z} & = & 14\mathbb{Z} = 2\mathbb{Z} \cdot 7\mathbb{Z}, \\ 2\mathbb{Z} + 7\mathbb{Z} & = & \mathbb{Z}. \end{array}$$

(**3**) W pierścieniu przemiennym pojęcia ideałów lewostronnego, prawostronnego i obustronnego są równoważne. Lecz w pierścieniu, który nie jest przemienny, ideał lewostronny nie zawsze jest prawostronny i na odwrót. Na przykład

$$I = \left\{ \begin{bmatrix} a & b & c \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mid a, b, c \in \mathbb{C} \right\}$$

jest ideałem prawostronnym w pierścieniu macierzy kwadratowych $M_3(\mathbb{C})$. Jednak

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \notin I,$$

a zatem I nie jest ideałem lewostronnym w $M_3(\mathbb{C})$.

Ćwiczenia 4.1.7.

(**1**) Znaleźć wszystkie ideały lewostronne (odpowiednio prawostronne) w pierścieniu A, jeśli:

(*a*) $A = \{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbb{R} \}$;

(*b*) $A = M_2(\mathbb{Z}_2)$;

(*c*) $A = M_2(\mathbb{Z}_3)$.
(**2**) Niech A będzie pierścieniem, a I jego ideałem lewostronnym. Udowodnić, że:
(*a*) anihilator lewostronny $\mathrm{ann}_l(I) = \{r \in A \mid ri = 0 \text{ dla wszystkich } i \in I\}$ jest ideałem w A;
(*b*) anihilator prawostronny $\mathrm{ann}_r(I) = \{r \in A \mid ir = 0 \text{ dla wszystkich } i \in I\}$ jest ideałem prawostronnym w A.
(**3**) Sprawdzić, czy elementy nieodwracalne pierścienia A tworzą ideał, jeśli:
(*a*) $A = \mathbb{Z}$;
(*b*) $A = \mathbb{Z}_8$;
(*c*) $A = \mathbb{Z}_{16}$;
(*d*) $A = M_2(\mathbb{Z}_2)$.

Uwagi. R. Dedekind wprowadził pojęcia ciała, porządku i ideału. Później matematyk niemiecki D. Hilbert zamiast terminu „porządek" zaczął używać terminu „pierścień". A. Poincaré i, niezależnie, J. Wedderburn wprowadzili koncepcję ideału jednostronnego.

4.2. Homomorfizmy pierścieni

Jednym ze sposobów studiowania pierścieni jest badanie własności ich homomorfizmów.

■ Niech $(R, +, \cdot)$ oraz $(S, +, \cdot)$ będą pierścieniami. Wtedy $f : R \to S$ jest nazywane *homomorfizmem* pierścieni, jeśli:
1) f jest odwzorowaniem;
2)
$$f(r + t) = f(r) + f(t) \tag{4.1}$$
dla dowolnych $r, t \in R$;
3)
$$f(r \cdot t) = f(r) \cdot f(t) \tag{4.2}$$
dla dowolnych $r, t \in R$.

■ Niech wszędzie dalej 0_R będzie zerem pierścienia R, a 0_S będzie zerem w S. Wtedy *jądro* $\operatorname{Ker} f$ homomorfizmu f jest zbiorem elementów a pierścienia R takich, że $f(a) = 0_S$. Zbiór
$$\operatorname{Im} f = \{f(a) \mid a \in R\}$$
jest nazywany *obrazem* homomorfizmu f.

■ Homomorfizm pierścieni $f : R \to S$ jest nazywany:
- *monomorfizmem* (lub *włożeniem*), jeśli f jest iniekcją;
- *epimorfizmem*, jeśli odwzorowanie f jest suriekcją;
- *izomorfizmem*, jeśli odwzorowanie f jest bijekcją.

■ Pierścienie R oraz S są nazywane *izomorficznymi*, jeśli istnieje pewien izomorfizm postaci $f : R \to S$ (wtedy oznaczamy to przez $R \cong S$ lub $R \simeq S$). Pierścień S jest nazywany *obrazem homomorficznym* pierścienia R, jeśli istnieje pewien epimorfizm postaci $f : R \to S$.

■ Homomorfizm pierścieni postaci $f : R \to R$ jest nazywany *endomorfizmem*, a izomorfizm postaci $f : R \to R$ – *automorfizmem* pierścienia R.

Lemat 4.2.1. *Jeśli $f : R \to S$ jest izomorfizmem pierścieni R i S, to odwzorowanie odwrotne $f^{-1} : S \to R$ też jest izomorfizmem pierścieni R i S.*

Dowód nie jest trudny i zostawiamy go Czytelnikowi do samodzielnego opracowania.

□

Twierdzenie 4.2.2. *Niech R i S będą pierścieniami. Wtedy homomorfizm pierścieni $f : R \to S$ posiada następujące własności:*

(1) $f(0_R) = 0_S$;
(2) $f(-x) = -f(x)$ *dla każdego* $x \in R$;
(3) $f(nx) = nf(x)$ *dla dowolnych* $x \in R$ *oraz* $n \in \mathbb{Z}$;
(4) $f(r - t) = f(r) - f(t)$ *dla dowolnych elementów* $r, t \in R$;
(5) $f(x^n) = f(x)^n$ *dla dowolnych* $x \in R$ *oraz* $n \in \mathbb{N}^*$;
(6) $\operatorname{Ker} f$ *jest ideałem w* R;
(7) $\operatorname{Im} f$ *jest podpierścieniem w* S.

Dowód. (1) Rzeczywiście,

$$0_S + f(0_R) = f(0_R) = f(0_R + 0_R) = f(0_R) + f(0_R)$$

i na podstawie prawa skracania w grupie $(R, +)$ otrzymujemy $f(0_R) = 0_S$.

(2) Skoro $0_R = x + (-x)$, to dostajemy

$$0_S = f(0_R) = f(x + (-x)) = f(x) + f(-x),$$

a stąd

$$f(-x) = -f(x).$$

(3) Zachodzi na podstawie tego, że

$$nx = \begin{cases} \underbrace{x + \cdots + x}_{n \text{ składników}}, & \text{gdy } n > 0, \\ 0, & \text{gdy } n = 0, \\ \underbrace{(-x) + \cdots + (-x)}_{|n| \text{ składników}}, & \text{gdy } n < 0. \end{cases}$$

(4) Wynika z własności (2) i definicji homomorfizmu.

(5) Łatwo przekonać się, jeśli weźmie się pod uwagę definicję potęgi elementu.

(6) Z własności (1) wynika, że $0_R \in \operatorname{Ker} f$. Jeśli $a, b \in \operatorname{Ker} f$ oraz $r \in R$, to

$$\begin{array}{rcl} f(a-b) & = & f(a) - f(b) = 0_S - 0_S = 0_S, \\ f(ar) & = & f(a)f(r) = 0_S f(r) = 0_S, \\ f(ra) & = & f(r)f(a) = f(r)0_S = 0_S, \end{array}$$

a zatem $\operatorname{Ker} f$ jest ideałem w R.

(7) Z części (1) wynika, że $0_S \in \operatorname{Im} f$. Jeśli $\alpha, \beta \in \operatorname{Im} f$, to $\alpha = f(u)$ oraz $\beta = f(v)$ dla pewnych elementów $u, v \in R$. Wtedy

$$\begin{array}{rcl} \alpha - \beta & = & f(u) - f(v) = f(u-v) \in \operatorname{Im} f, \\ \alpha\beta & = & f(u)f(v) = f(uv) \in \operatorname{Im} f, \end{array}$$

a więc $\operatorname{Im} f$ jest podpierścieniem w S. □

Przykłady 4.2.3.

(**1**) Niech

$$M_2(\mathbb{R}) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbb{R}\}.$$

Odwzorowanie

$$f : M_2(\mathbb{R}) \ni \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto d \in \mathbb{R}$$

nie jest homomorfizmem. Istotnie, jeśli

$$A_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \in M_2(\mathbb{R}),\ A_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \in M_2(\mathbb{R}),$$

to

$$f(A_1A_2) = f\left(\begin{bmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{bmatrix}\right) = c_1b_2 + d_1d_2$$

oraz $f(A_1)f(A_2) = d_1d_2$. Jeśli teraz

$$A_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

to

$$f(A_1A_2) = 1 \neq 0 = f(A_1)f(A_2),$$

co potwierdza hipotezę.

(**2**) Niech

$$\mathbb{Q}[\sqrt{3}] = \{x + y\sqrt{3} \mid x, y \in \mathbb{Q}\} \text{ oraz } T = \{\begin{bmatrix} x & y \\ 3y & x \end{bmatrix} \mid x, y \in \mathbb{Q}\}.$$

Wtedy reguła

$$g : \mathbb{Q}[\sqrt{3}] \ni x + y\sqrt{3} \mapsto \begin{bmatrix} x & y \\ 3y & x \end{bmatrix} \in T$$

określa odwzorowanie. Ponadto dla dowolnych elementów $x, y, z, t \in \mathbb{Q}$ otrzymujemy

$$g\Big((x+y\sqrt{3})+(z+t\sqrt{3})\Big) = g\Big((x+z)+(y+t)\sqrt{3}\Big) = \begin{bmatrix} x+z & y+t \\ 3(y+t) & x+z \end{bmatrix} = \begin{bmatrix} x & y \\ 3y & x \end{bmatrix} + \\ + \begin{bmatrix} z & t \\ 3t & z \end{bmatrix} = g(x+y\sqrt{3}) + g(z+t\sqrt{3}) \text{ oraz}$$

$$g\Big((x+y\sqrt{3})(z+t\sqrt{3})\Big) = g\Big((xz+3yt)+(xt+yz)\sqrt{3}\Big) = \begin{pmatrix} xz+3yt & xt+yz \\ 3(xt+yz) & xz+3yt \end{pmatrix} = \\ = \begin{bmatrix} x & y \\ 3y & x \end{bmatrix} \begin{bmatrix} z & t \\ 3t & z \end{bmatrix} = g(x+y\sqrt{3})g(z+t\sqrt{3}),$$

czyli g jest homomorfizmem pierścieni. Jeśli $g(x+y\sqrt{3}) = g(z+t\sqrt{3})$, to

$$\begin{bmatrix} x & y \\ 2y & x \end{bmatrix} = \begin{bmatrix} z & t \\ 3t & z \end{bmatrix},$$

a więc $x = z$ i $y = t$, czyli g jest monomorfizmem. Poza tym dla dowolnej macierzy $A \in T$ znajdą się takie $u, v \in \mathbb{Q}$, że

$$A = \begin{bmatrix} u & v \\ 3v & u \end{bmatrix}$$

i wtedy

$$g(u+v\sqrt{3}) = A,$$

czyli g jest epimorfizmem. Zatem g jest izomorfizmem pierścieni oraz $\mathbb{Q}[\sqrt{3}] \cong T$.

(**3**) Niech R, S będą pierścieniami. Reguła

$$0 : R \ni r \mapsto 0 \in S$$

określa homomorfizm pierścieni, który jest nazywany *zerowym*. W szczególności

$$0_R : R \ni r \mapsto 0 \in R$$

jest endomorfizmem zerowym pierścienia R. Homomorfizm tożsamościowy

$$\mathrm{i}_R : R \ni r \mapsto r \in R$$

jest automorfizmem pierścienia R. Endomorfizmy tożsamościowy i_R oraz zerowy 0_R pierścienia R są nazywane *trywialnymi*.

■ Jeśli R i S są ciałami (z jednościami 1_R i odpowiednio 1_S) oraz odwzorowanie $f : R \to S$ spełnia warunki (4.1) oraz (4.2), to

$$f(1_R) = f(1_R^2) = f(1_R)^2 \in S,$$

a to znaczy, że $f(1_R) = 0_S$ lub $f(1_R) = 1_S$. Zatem jeśli f jest homomorfizmem niezerowym ciał, to

$$f(1_R) = 1_S.$$

Poza tym $\operatorname{Ker} f = \{0_R\}$, a więc homomorfizm niezerowy f jest monomorfizmem ciał. W ten sposób udowodniliśmy

Lemat 4.2.4. *Jeśli $f : R \to S$ jest homomorfizmem niezerowym ciał, to $f(1_R) = 1_S$ oraz f jest iniekcją.*

Ćwiczenia 4.2.5.

(**1**) Udowodnić, że:
(a) jeśli $\varphi : A \to S$ jest homomorfizmem pierścieni, to $\varphi^{-1}(K) = \{r \in A \mid \varphi(r) \in K\}$ jest ideałem jednostronnym pierścienia A dla każdego ideału jednostronnego K pierścienia S;
(b) jeśli $\varphi : A \to S$ jest epimorfizmem pierścieni oraz I jest jednostronnym (odpowiednio obustronnym) ideałem w A, to $\varphi(I) = \{\varphi(i) \mid i \in I\}$ jest jednostronnym (odpowiednio obustronnym) ideałem w S.
(**2**) Znaleźć wszystkie homomorfizmy pierścienia A w pierścień S, jeśli:
(a) $A = \mathbb{Z}$ oraz $S = 2\mathbb{Z}$;
(b) $A = S = 3\mathbb{Z}$;
(c) $A = 2\mathbb{Z}$ oraz $S = 3\mathbb{Z}$;
(d) $A = \mathbb{Z}_2$ oraz $S = \mathbb{Z}_3$;
(e) $A = S = \mathbb{Z}_2$;
(f) $A = S = \mathbb{Z}_3$;
(g) $A = \mathbb{Z}$ oraz $S = \mathbb{Q}$.
(**3**) Udowodnić, że zachodzą izomorfizmy pierścieni:
(a) $C_{[0,1]}/I \cong \mathbb{R}$, gdzie $I = \{f \in C_{[0,1]} \mid f(0) = 0\}$;
(b) $\mathbb{Z}[\frac{1}{2}]/5\mathbb{Z}[\frac{1}{2}] \cong \mathbb{Z}_5$;
(c) $D_2(\mathbb{C}) \cong \mathbb{C}^2$, gdzie $D_2(\mathbb{C}) = \{\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a, b \in \mathbb{C}\}$;
(d) $C_{[0,1]} \cong C_{[5,6]}$;
(e) $R \cong S$, gdzie $R = \{f \in C_{[-1,1]} \mid f(-1) = 0\}$ oraz $S = \{f \in C_{[-1,1]} \mid f(1) = 0\}$.
(**4**) Sprawdzić, czy ψ jest homomorfizmem pierścieni. Znaleźć jądro $\operatorname{Ker}\psi$ i obraz $\operatorname{Im}\psi$, jeśli:
(a) $\psi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ oraz $\psi(A) = X^{-1}AX$, gdzie $A \in M_n(\mathbb{C})$, a X jest ustaloną macierzą nieosobliwą z $M_n(\mathbb{C})$;
(b) $\psi : M_2(\mathbb{C}) \to \mathbb{C}$ oraz $\psi(A) = \det A$;
(c) $\psi : C_{[0,1]} \to \mathbb{R}$ oraz $\psi(f) = f(0)$;
(d) $\psi : \mathbb{C}^2 \to \mathbb{C}$ oraz $\psi((z_1, z_2)) = z_1 + z_2$.
(**5**) Znaleźć wszystkie homomorfizmy pierścieni postaci $\psi : A \to S$, jeśli:
(a) $A = \mathbb{Z}$ oraz $S = \mathbb{Z}_6$;
(b) $A = \mathbb{Z}_7$ oraz $S = \mathbb{Z}_5$;
(c) $A = S = \mathbb{Z} \times \mathbb{Z}$;
(d) $A = S = \mathbb{Z}_2 \times \mathbb{Z}_2$.
(**6**) Udowodnić, że nie istnieje niezerowy homomorfizm postaci $\phi : A \to S$, jeśli:
(a) $A = \mathbb{Z}[\sqrt{2}]$ oraz $S = \mathbb{Z}[\sqrt{5}]$;
(b) $A = \mathbb{Z}[\sqrt{2}]$ oraz $S = \mathbb{Z}[i]$.
(**7**) Niech a, b będą niezerowymi liczbami wymiernymi. Udowodnić, że ciała $\mathbb{Q}(\sqrt{a})$ oraz $\mathbb{Q}(\sqrt{b})$ są izomorficzne, jeśli $\frac{a}{b}$ jest kwadratem liczby wymiernej.
(**8**) Znaleźć wszystkie (z dokładnością do izomorfizmu) ciała, które nie posiadają podciał właściwych.

Uwagi. Użycie terminu „homomorfizm” zasugerował G. Frobenius[(1)].

[(1)] Ferdinand Georg Frobenius (1849–1917)

4.3. Pierścień ilorazowy

■ Niech $(R, +, \cdot)$ będzie pierścieniem, a I jego ideałem (obustronnym). Na zbiorze warstw

$$R/I = \{a + I \mid a \in R\}$$

pierścienia R względem ideału I zadajmy dwa działania:

- (dodawanie „+")

$$\overline{a} + \overline{b} = (a + I) + (b + I) = (a + b) + I = \overline{a + b};$$

- (mnożenie „·")

$$\overline{a} \cdot \overline{b} = (a + I) \cdot (b + I) = (a \cdot b) + I = \overline{a \cdot b}$$

dla dowolnych warstw $\overline{a} = a + I, \overline{b} = b + I \in R/I$.

Twierdzenie 4.3.1. *Niech $(R, +, \cdot)$ będzie pierścieniem z jednością 1 oraz I będzie jego ideałem. Wtedy $(R/I, +, \cdot)$ jest pierścieniem z jednością $\overline{1} = 1 + I$. Ponadto, jeżeli R jest przemienny, to R/I również.*

Dowód. Mamy $\overline{0} = 0 + I = I \in R/I$. Udowodnimy, że dodawanie „+" i mnożenie „·" są dobrze określone na zbiorze R/I, czyli są algebraiczne. Niech dalej

$$\overline{a} = a + I,\ \overline{a'} = a' + I,\ \overline{b} = b + I,\ \overline{b'} = b' + I \text{ oraz } \overline{c} = c + I$$

będą dowolnymi elementami z R/I.

Chcemy przekonać się, że z równości $\overline{a} = \overline{a'}$ oraz $\overline{b} = \overline{b'}$ wynika, że

$$\overline{a} + \overline{b} = \overline{a'} + \overline{b'} \text{ oraz } \overline{a} \cdot \overline{b} = \overline{a'} \cdot \overline{b'}.$$

W rzeczy samej, z $\overline{a} = \overline{a'}$ oraz $\overline{b} = \overline{b'}$ na podstawie kryterium równości warstw mamy $a - a', b - b' \in I$, a więc

$$a - a' = i_1 \text{ oraz } b - b' = i_2$$

dla pewnych elementów $i_1, i_2 \in I$. Wtedy $i_1 + I = I$, $i_2 + I = I$ oraz

$$\overline{a} + \overline{b} = (a + I) + (b + I) = (a' + i_1 + I) + (b' + i_2 + I) = \\ = (a' + I) + (b' + I) = \overline{a'} + \overline{b'}$$

oraz

$$\overline{a} \cdot \overline{b} = (a+I) \cdot (b+I) = (a' + i_1 + I) \cdot (b' + i_2 + I) = \\ = (a'+I) \cdot (b'+I) = \overline{a'} \cdot \overline{b'}.$$

Tak jak w przypadku grup ilorazowych (patrz twierdzenie 3.6.1) otrzymujemy, że R/I jest grupą abelową z elementem zerowym $\overline{0} = 0 + I$.

Mamy też

$$(\overline{a}\ \overline{b})\ \overline{c} = ((a+I)(b+I))(c+I) = \\ = (ab+I)(c+I) = (ab)c + I = a(bc) + I = \\ = (a+I)(bc+I) = (a+I)((b+I)(c+I)) = \overline{a}\ (\overline{b}\ \overline{c}),$$

czyli mnożenie „$\cdot$” jest łączne w R/I,

$$(\overline{a} + \overline{b})\ \overline{c} = ((a+I) + (b+I))(c+I) = \\ = ((a+b)+I)(c+I) = (a+b)c + I = (ac+bc) + I = \\ = (ac+I) + (bc+I) = (a+I)(c+I) + (b+I)(c+I) = \\ = \overline{a}\ \overline{c} + \overline{b}\ \overline{c}$$

oraz

$$\overline{a}\ (\overline{b} + \overline{c}) = (a+I)((b+I) + (c+I)) = (a+I)((b+c)+I) = \\ = a(b+c) + I = (ab+ac) + I = (ab+I) + (ac+I) = \\ = (a+I)(b+I) + (a+I)(c+I) = \overline{a}\ \overline{b} + \overline{a}\ \overline{c},$$

a więc mnożenie „$\cdot$” jest rozdzielne względem dodawania „$+$” w R/I. Poza tym

$$\overline{a} \cdot \overline{1} = (a+I)(1+I) = a \cdot 1 + I = a + I = \\ = \overline{a} = a + I = 1 \cdot a + I = (1+I)(a+I) = \overline{1} \cdot \overline{a},$$

czyli $\overline{1} = 1 + I$ jest elementem neutralnym względem mnożenia „$\cdot$” w pierścieniu ilorazowym R/I.

Załóżmy, że pierścień R jest przemienny. Wtedy

$$\overline{a} \cdot \overline{b} = (a+I)(b+I) = a \cdot b + I = b \cdot a + I = \\ = (b+I)(a+I) = \overline{b} \cdot \overline{a},$$

czyli mnożenie „$\cdot$” też jest przemienne w R/I. □

■ Jeśli I jest ideałem pierścienia R, to pierścień

$$R/I$$

jest nazywany *pierścieniem ilorazowym* pierścienia R względem ideału I.

Przykłady 4.3.2.

(**1**) Niech 0 będzie zerem pierścienia R. Wtedy pierścień ilorazowy $R/\langle 0\rangle$ jest izomorficzny z R. Rzeczywiście, odwzorowanie

$$\phi : R \ni r \mapsto r + \langle 0\rangle \in R/\langle 0\rangle$$

jest izomorfizmem pierścieni (sprawdzić samodzielnie), czyli

$$R/\langle 0\rangle \cong R.$$

Podobnie, jeśli $r \in R$, to warstwa $r + R = 0 + R = R = \overline{0}$ jest zerowa w pierścieniu ilorazowym R/R, czyli R/R składa się dokładnie z jednego (a mianowicie zerowego) elementu. Jako wniosek

$$R/R \cong \langle 0\rangle.$$

Pierścienie ilorazowe $R/\langle 0\rangle$ oraz R/R są nazywane *trywialnymi* pierścieniami ilorazowymi pierścienia R.

(**2**) Niech I będzie ideałem pierścienia R. Wtedy odwzorowanie

$$\pi : R \ni r \mapsto r + I \in R/I$$

jest epimorfizmem pierścieni (udowodnić samodzielnie). Takie odwzorowanie jest nazywane *homomorfizmem kanonicznym* (*odwzorowaniem kanonicznym* lub *epimorfizmem naturalnym*). Jeśli I jest niezerowym ideałem pierścienia R, to pierścień ilorazowy R/I jest nazywany *właściwym*.

Lemat 4.3.3. *Niech R, S będą pierścieniami, $\phi : R \to S$ będzie homomorfizmem pierścieni, I będzie ideałem w R, a K ideałem prawostronnym (odpowiednio lewostronnym lub obustronnym) w S. Wtedy zachodzą następujące własności:*

(1) *jeśli T jest podpierścieniem w R, zawierającym I, oraz T/I jest ideałem pierścienia ilorazowego R/I, to T jest ideałem w R;*
(2) $\operatorname{Ker}\phi \subseteq \phi^{-1}(K)$*;*
(3) $\phi^{-1}(K)$ *jest ideałem prawostronnym (odpowiednio lewostronnym lub obustronnym) w R.*

Dowód. (1) Jeśli $r \in R$ oraz $t \in T$ są dowolnymi elementami, to

$$\begin{aligned}(r+I)(t+I) &= rt + I \in T/I,\\ (t+I)(r+I) &= tr + I \in T/I,\end{aligned}$$

a więc $rt, tr \in T$. Biorąc pod uwagę, że T jest podpierścieniem w R, wnosimy, że T jest ideałem w R.

(2) Wynika na podstawie definicji jądra i przeciwobrazu.

(3) Niech K będzie ideałem lewostronnym pierścienia S. Skoro $\phi(0_R) = 0_S \in K$, to $0_R \in \phi^{-1}(K)$. Jeśli $\alpha, \beta \in \phi^{-1}(K)$ oraz $r \in R$, to $\phi(\alpha), \phi(\beta) \in K$, a zatem

$$\begin{aligned} \phi(\alpha - \beta) &= \phi(\alpha) - \phi(\beta) \in K, \\ \phi(r\alpha) &= \phi(r)\phi(\alpha) \in K. \end{aligned}$$

Na tej podstawie $\alpha - \beta, r\alpha \in \phi^{-1}(K)$, czyli $\phi^{-1}(K)$ jest ideałem lewostronnym w R.

Dowód w przypadku ideału prawostronnego (odpowiednio obustronnego) K jest podobny. □

Przykład 4.3.4.

Wielomian $X^2 + 1$ nie posiada pierwiastków w ciele liczb rzeczywistych $\mathbb{R}$, a więc jest nieprzywiedlny[(2)] nad ciałem $\mathbb{R}$ (udowodnić samodzielnie). Znajdźmy pierścień ilorazowy $\mathbb{R}[X]/\langle X^2 + 1\rangle$ pierścienia wielomianów $\mathbb{R}[X]$ względem ideału głównego $\langle X^2+1\rangle$. Oznaczmy $I = \langle X^2+1\rangle$. Najpierw zaznaczmy, że z twierdzenia o dzieleniu z resztą dla wielomianu $f \in \mathbb{R}[X]$ istnieją takie wielomiany $q, r \in \mathbb{R}[X]$, że

$$f = q(X^2 + 1) + r,$$

przy czym $\deg r < \deg(X^2 + 1) = 2$, czyli $r = aX + b$ dla pewnych $a, b \in \mathbb{R}$. Stąd dla wielomianu $f \in \mathbb{R}[X]$ otrzymujemy, że

$$f + I = r + q(X^2 + 1) + I = (aX + b) + (q(X^2 + 1) + I) = aX + b + I.$$

Wtedy reguła

$$\theta : \mathbb{R}[X]/I \ni f + I \mapsto ai + b \in \mathbb{C}$$

określa odwzorowanie. Sprawdźmy, że θ jest izomorfizmem pierścieni. Rzeczywiście, jeśli $f, g \in \mathbb{R}[X]$ oraz

$$\begin{aligned} f + I &= aX + b + I, \\ g + I &= cX + d + I \end{aligned}$$

są elementami z $\mathbb{R}[X]/\langle X^2 + 1\rangle$, gdzie $a, b, c, d \in \mathbb{R}$, to

$$\begin{gathered} \theta((f + I) + (g + I)) = \theta((aX + b + I) + (cX + d + I)) = \theta((a + c)X + (b + d) + I) = \\ = (a + c)i + (b + d) = (ai + b) + (ci + d) = \theta(aX + b + I) + \theta(cX + d + I) = \\ = \theta(f + I) + \theta(g + I) \end{gathered}$$

---

(2) Przypomnijmy, że wielomian niezerowy $f \in \mathbb{F}[X]$ stopnia $n > 0$ jest nazywany *przywiedlnym* (=*rozkładalnym*) nad ciałem $\mathbb{F}$, jeśli istnieją wielomiany $g, h \in \mathbb{F}[X]$ stopni dodatnich takie, że $f = g \cdot h$. Jeśli takie wielomiany nie istnieją, to f jest nazywany *nieprzywiedlnym* (=*nierozkładalnym*) nad ciałem $\mathbb{F}$.

oraz

$$\theta((f+I)(g+I)) = \theta((aX+b+I)(cX+d+I)) =$$
$$= \theta((ad+bc)X + (bd-ac) + ac(X^2+1) + I) = \theta((ad+bc)X + (bd-ac) + I) =$$
$$= (ad+bc)i + (bd-ac) = (ai+b)(ci+d) = \theta(aX+b+I)\theta(cX+d+I) = \theta(f+I)\theta(g+I),$$

czyli θ jest homomorfizmem pierścieni.

Niech z będzie dowolną liczbą zespoloną. Wtedy znajdą się takie liczby rzeczywiste u, v, że $z = ui + v$, a więc

$$\theta(uX + v + I) = ui + v = z$$

i θ jest suriektywne. Teraz załóżmy, że $\theta(f+I) = \theta(g+I)$. Stąd wynika, że $ai + b = ci + d$, a zatem $a = c$ i $b = d$. Lecz wtedy $aX + b = cX + d$ oraz

$$f + I = aX + b + I = cX + d + I = g + I,$$

czyli odwzorowanie θ jest iniekcją. Zatem θ jest izomorfizmem pierścieni oraz

$$\mathbb{R}[X]/\langle X^2 + 1\rangle \cong \mathbb{C}.$$

Wniosek 4.3.5. *Niech R będzie pierścieniem, a I, J będą jego ideałami. Jeśli $\pi : R \to R/I$ jest homomorfizmem kanonicznym, to*

$$\pi(J) = (I + J)/I.$$

Dowód nie jest trudny; proponujemy Czytelnikowi znaleźć go samodzielnie.

□

Twierdzenie 4.3.6 (pierwsze o izomorfizmie pierścieni)**.** *Jeśli $f : R \to S$ jest homomorfizmem pierścieni R oraz S, to*

$$R/\operatorname{Ker} f \cong \operatorname{Im} f.$$

Dowód. Rozpatrzmy odwzorowanie $\phi : R/\operatorname{Ker} f \to \operatorname{Im} f$, określone wzorem

$$\phi(r + \operatorname{Ker} f) = f(r) \quad (r \in R).$$

Jeśli

$$\overline{r}_1 = r_1 + \operatorname{Ker} f \text{ oraz } \overline{r}_2 = r_2 + \operatorname{Ker} f$$

są dowolnymi elementami pierścienia ilorazowego $R/\operatorname{Ker} f$, to

$$\phi(\overline{r}_1 + \overline{r}_2) = \phi((r_1 + \operatorname{Ker} f) + (r_2 + \operatorname{Ker} f)) =$$
$$= \phi(r_1 + r_2 + \operatorname{Ker} f) =$$
$$= f(r_1 + r_2) = f(r_1) + f(r_2) =$$
$$= \phi(r_1 + \operatorname{Ker} f) + \phi(r_2 + \operatorname{Ker} f) = \phi(\overline{r}_1) + \phi(\overline{r}_2)$$

oraz

$$\begin{gathered}\phi(\overline{r}_1\overline{r}_2) = \phi((r_1 + \operatorname{Ker} f)(r_2 + \operatorname{Ker} f)) = \phi(r_1r_2 + \operatorname{Ker} f) = \\ = f(r_1r_2) = f(r_1)f(r_2) = \\ = \phi(r_1 + \operatorname{Ker} f)\phi(r_2 + \operatorname{Ker} f) = \phi(\overline{r}_1)\phi(\overline{r}_2),\end{gathered}$$

a to znaczy, że ϕ jest homomorfizmem pierścieni. Jeśli $\alpha \in \operatorname{Im} f$, to $\alpha = f(t)$ dla pewnego $t \in R$. Wtedy

$$\phi(t + \operatorname{Ker} f) = f(t) = \alpha,$$

a więc odwzorowanie ϕ jest suriekcją. Załóżmy, że $\phi(\overline{r}_1) = \phi(\overline{r}_1)$. Wtedy $f(r_1) = f(r_2)$, a stąd

$$f(r_1 - r_2) = f(r_1) - f(r_2) = 0.$$

To znaczy, że $r_1 - r_2 \in \operatorname{Ker} f$. Na podstawie kryterium równości warstw dostajemy

$$\overline{r}_1 = r_1 + \operatorname{Ker} f = r_2 + \operatorname{Ker} f = \overline{r}_2,$$

czyli ϕ jest iniekcją. Zatem ϕ jest izomorfizmem pierścieni. □

Przykład 4.3.7.

Niech $\mathbb{F}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ będzie ciałem o 7 elementach. Rozpatrzmy pierścień wielomianów $\mathbb{F}_7[X]$. Odwzorowanie

$$\theta_4 : \mathbb{F}_7[X] \ni f \mapsto f(4) \in \mathbb{F}_7$$

jest homomorfizmem pierścieni. Istotnie, dla dowolnych wielomianów $f, g \in \mathbb{F}_7[X]$

$$\begin{array}{rcccccc}\theta_4(f+g) & = & (f+g)(4) = f(4) + g(4) & = & \theta_4(f) + \theta_4(g), \\ \theta_4(fg) & = & (fg)(4) = f(4)g(4) & = & \theta_4(f)\theta_4(g).\end{array}$$

Niech q będzie dowolnym elementem ciała $\mathbb{F}_7$, a $f = X + 3 + q$. Wtedy $\theta_4(f) = f(4) = 4 + 3 + q = q$, czyli odwzorowanie θ_4 jest suriekcją. Skoro $u = X^2 + X + 4$ oraz $v = X^2 + 2X$ są różnymi wielomianami i $\theta_4(u) = \theta_4(v)$, to odwzorowanie θ_4 nie jest iniektywne. Zatem jego jądro

$$\operatorname{Ker} \theta_4 = \{f \in \mathbb{F}_7[X] \mid \theta_4(f) = 0\} = \{f \in \mathbb{F}_7[X] \mid f(4) = 0\}$$

jest niezerowe. Na podstawie twierdzenia 4.3.6 wnosimy, że

$$\mathbb{F}_7[X]/\operatorname{Ker} \theta_4 \cong \operatorname{Im} \theta_4 = \mathbb{F}_7.$$

Twierdzenie 4.3.8 (drugie o izomorfizmie pierścieni)**.** *Niech R będzie pierścieniem, a I, J będą jego ideałami. Wtedy zachodzi taki izomorfizm pierścieni*

$$(I + J)/I \cong J/(I \bigcap J).$$

Dowód. Zostawiamy Czytelnikowi, aby przekonał się, że

$$\varphi : J \ni j \mapsto j + I \in (I + J)/I$$

jest homomorfizmem pierścieni z jądrem

$$\begin{gathered}\operatorname{Ker}\varphi = \{j \in J \mid \varphi(j) = 0 + I \in (I + J)/I\} = \\ = \{j \in J \mid j \in I\} = I \bigcap J\end{gathered}$$

i obrazem $\operatorname{Im}\varphi = (I + J)/I$. Reszta wynika na mocy twierdzenia 4.3.6. □

Twierdzenie 4.3.9 (trzecie o izomorfizmie pierścieni)**.** *Niech I, J będą ideałami pierścienia R, przy czym $I \leqslant J$. Wtedy zachodzi taki izomorfizm pierścieni*

$$R/J \cong (R/I)/(J/I).$$

Dowód. Łatwo zauważyć, że J/I jest ideałem w pierścieniu ilorazowym R/I. Skoro

$$\varphi : R/I \ni x + I \mapsto x + J \in R/J$$

jest homomorfizmem pierścieni, $\operatorname{Ker}\varphi = J/I$ oraz $\operatorname{Im}\varphi = R/J$ (przekonać się samodzielnie), to teza zachodzi na podstawie twierdzenia 4.3.6. □

Następne twierdzenie jest też nazywane twierdzeniem o odpowiedniości ideałów.

Twierdzenie 4.3.10. *Niech I będzie ideałem pierścienia R. Wtedy istnieje bijektywna odpowiedniość między ideałami pierścienia ilorazowego R/I a ideałami pierścienia R zawierającymi ideał I.*

Dowód. Ćwiczenie. □

■ Łatwo zauważyć, że ciało liczb wymiernych $\mathbb{Q}$ i ciało klas reszt $\mathbb{Z}_p$ modulo liczba pierwsza p nie zawierają żadnego podciała właściwego, czyli w pewnym sensie one są „najmniejszymi" ciałami. To potwierdza takie

Twierdzenie 4.3.11. *Każde ciało $\mathbb{F}$ zawiera dokładnie jedno (najmniejsze względem relacji „$\subseteq$") podciało F_0, które jest izomorficzne z:*

(1) *ciałem liczb wymiernych $\mathbb{Q}$, jeśli* $\operatorname{char}\mathbb{F} = 0$*;*

(2) *ciałem liczb klas reszt $\mathbb{Z}_p$ modulo liczby pierwszej p, jeśli* $\operatorname{char}\mathbb{F} = p$*.*

(Takie podciało F_0 jest nazywane *podciałem prostym* ciała $\mathbb{F}$).

Dowód. Niech e i 0 będą odpowiednio jednością i zerem ciała $\mathbb{F}$. Wtedy odwzorowanie $\varphi : \mathbb{Z} \to \mathbb{F}$, określone według reguły

$$\varphi(n) = ne \ (n \in \mathbb{Z}),$$

jest homomorfizmem pierścieni (sprawdzić samodzielnie). Znajdźmy jego jądro

$$\operatorname{Ker}\varphi = \{n \in \mathbb{Z} \mid \varphi(n) = 0\}.$$

(1) Jeśli $\operatorname{char}\mathbb{F} = 0$, to

$$ne = 0 \Leftrightarrow n = 0,$$

a więc $\operatorname{Ker}\varphi = \{0\}$. To oznacza, że pierścień $\mathbb{Z}$ jest izomorficzny z obrazem $\operatorname{Im}\varphi$ i ciało F zawiera podpierścień izomorficzny z pierścieniem $\mathbb{Z}$ (wtedy, jako wniosek, $\mathbb{F}$ zawiera podciało F_0 izomorficzne z ciałem liczb wymiernych $\mathbb{Q}$).

(2) Jeśli $\operatorname{char}\mathbb{F} = p$, to $\varphi(pn) = 0$ dla każdego $n \in \mathbb{Z}$. Innymi słowy, $\operatorname{Ker}\varphi = p\mathbb{Z}$ oraz

$$\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/\operatorname{Ker}\varphi \cong \operatorname{Im}\varphi$$

jest ciałem zawierającym się w $\mathbb{F}$ jako podciało. □

* * *

■ **Relacje równoważności w pierścieniu, które są zgodne z działaniami w pierścieniu.** Relacja równoważności „$\sim$", która jest określona w pierścieniu R, jest nazywana *kongruencją* w R, jeśli zachodzi implikacja

$$\forall_{a,b,c,d\in R} : a \sim b \text{ oraz } c \sim d \;\Rightarrow\; ac \sim bd \text{ oraz } a+c \sim b+d.$$

Zachodzi takie

Twierdzenie 4.3.12. *Niech R będzie pierścieniem (względem działań „$+$" oraz „$\cdot$"). Wtedy:*

(1) *jeśli I jest ideałem w R, to relacja „$\sim$" taka, że*

$$a \sim b \quad \Leftrightarrow \quad a - b \in I$$

dla elementów $a, b \in R$ jest kongruencją w pierścieniu R;

(2) *jeśli relacja „$\backsim$" jest kongruencją w pierścieniu R, to:*

(a) *klasa równoważności*

$$0_{\backsim} = \{a \in R \mid a \backsim 0\},$$

której reprezentantem jest element zerowy 0 pierścienia R, jest ideałem w R;

(b) *zachodzi równość*

$$R/0_{\backsim} = R/\backsim.$$

Dowód. Niech $a, b, c, d, x \in R$.

(1) Załóżmy, że $a \sim b$ oraz $c \sim d$. Wtedy $a-b, c-d \in I$ i na podstawie kryterium równości warstw wynika, że

$$a + I = b + I,\ c + I = d + I,$$

a stąd

$$\begin{aligned}
&a + b + I = (a + I) + (b + I) = (c + I) + (d + I) = c + d + I \Rightarrow \\
&\Rightarrow (a + b) - (c + d) \in I, \\
&ab + I = (a + I)(b + I) = (c + I)(d + I) = cd + I \Rightarrow \\
&\Rightarrow ab - cd \in I.
\end{aligned}$$

Udowodniliśmy, że „$\sim$” jest kongruencją w R.

(2) Oczywiście, że $a \backsim a$. Oznaczmy klasę równoważności $0_{\backsim}$ przez I. Z udowodnionego twierdzenia 3.6.8(2) wynika, że I jest podgrupą w grupie addytywnej $(R, +)$. Jeśli $i \in I$, to

$$i \backsim 0,\ x \backsim x \Rightarrow ix \backsim 0x = 0 \text{ oraz } xi \backsim x0 = 0 \ \Rightarrow ix,\ xi \in I.$$

Na podstawie kryterium wnosimy, że I jest ideałem w R.

Podobnie jak w dowodzie twierdzenia 3.6.8(2) otrzymujemy, że $x_{\backsim} = x + I$ i teza zachodzi. □

* * *

■ **Homomorfizmy pierścieni a kongruencje w pierścieniu.** Zachodzi

Twierdzenie 4.3.13. *Jeśli $\varphi : R \to S$ jest homomorfizmem pierścieni R i S, to relacja „$\sim$” taka, że*

$$r \sim t \quad \Leftrightarrow \quad \varphi(r) = \varphi(t)$$

dla elementów $r, t \in R$ jest kongruencją w pierścieniu R.

Dowód. W rzeczy samej, jeśli $x, y, r, t \in R$, to

$$r \sim t \text{ oraz } x \sim y \quad \Rightarrow \quad \varphi(r) = \varphi(t) \text{ oraz } \varphi(x) = \varphi(y),$$

a stąd

$$\begin{array}{lcll} \varphi(r+x) & = & \varphi(r) + \varphi(x) = \varphi(t) + \varphi(y) = \varphi(t+y) & \Rightarrow \\ & \Rightarrow & r + x \sim t + y, & \\ \varphi(rx) & = & \varphi(r)\varphi(x) = \varphi(t)\varphi(y) = \varphi(ty) & \Rightarrow \\ & \Rightarrow & rx \sim ty & \end{array}$$

i teza zachodzi. □

Ćwiczenia 4.3.14.

(**1**) Udowodnić izomorficzność ciał:
(*a*) $\mathbb{R}[X]/\langle X^2 + X + 1\rangle \cong \mathbb{C}$;
(*b*) $\mathbb{R}[X]/\langle 3X - 7\rangle \cong \mathbb{R}$;

(*c*) $\mathbb{Q}(\sqrt{3}) \cong \mathbb{Q}[X]/\langle X^2 - 3\rangle$;
(*d*) $\mathbb{R}[X]/\langle X^2 - 1\rangle \cong \mathbb{R}^2$;
(*e*) $\mathbb{Q}[X]/\langle X - 2\rangle \cong \mathbb{Q}$.
(**2**) Niech F będzie ciałem. Udowodnić, że:
(*a*) $F[X,Y]/\langle X - Y\rangle \cong F[X]$;
b) $F[X,Y]/\langle X^2 - Y^2\rangle$ nie jest dziedziną całkowitości.
(**3**) Udowodnić, że pierścień ilorazowy $\mathbb{Z}_2[X]/\langle X^3 + X\rangle$ ma 4 ideały właściwe.

Uwagi. Idea, aby traktować ciało liczb zespolonych $\mathbb{C}$ jako ciało reszt

$$\mathbb{R}[X]/\langle X^2 + 1\rangle,$$

należy do Cauchy'ego. F. Molin[(3)] zajmował się badaniem pewnej algebry ilorazowej w 1897 r.

[(3)] Fedor Eduardovich Molin (1861–1941)

4.4. Pierścienie euklidesowe i pierścienie ideałów głównych

■ Przemienna dziedzina całkowitości A z jednością 1 jest nazywana *euklidesową*, jeśli istnieje odwzorowanie

$$N : A \setminus \{0\} \to \mathbb{N},$$

spełniające takie dwa warunki:

a) $N(xy) \geqslant N(x)$ dla dowolnych elementów $x, y \in A \setminus \{0\}$;

b) dla dowolnych elementów $a \in A$ oraz $b \in A \setminus \{0\}$ znajdą się takie elementy $q, r \in A$, że

$$a = bq + r,$$

gdzie $r = 0$ lub $N(r) < N(b)$.

Takie odwzorowanie N jest nazywane *normą* (lub *normą Dedekinda-Hassego*[(4)]) pierścienia A.

Przykłady 4.4.1.

(**1**) Jeśli $N(a) = |a|$ dla każdego $a \in \mathbb{Z}$, to N jest normą pierścienia liczb całkowitych $\mathbb{Z}$. Zatem $\mathbb{Z}$ jest dziedziną euklidesową.

(**2**) Niech $\mathbb{F}$ będzie dowolnym ciałem. Pierścień wielomianów $\mathbb{F}[X]$ jest dziedziną euklidesową z normą N, określoną według wzoru

$$N(f) = \deg f$$

dla dowolnego wielomianu $f \in \mathbb{F}[X]$. Zatem mamy taki

Wniosek 4.4.2 (twierdzenie o dzieleniu z resztą dla wielomianów)**.** *Niech $\mathbb{F}$ będzie ciałem oraz $f, g \in \mathbb{F}[X]$. Jeśli $g \neq 0$, to istnieje jedna para wielomianów $q, r \in \mathbb{F}[X]$ taka, że*

$$f = qg + r \quad \textit{oraz} \quad \deg r < \deg g.$$

(**3**) (**Pierścień liczb całkowitych Gaussa**) Niech

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}.$$

Nietrudno zauważyć, że $\mathbb{Z}[i]$ jest pierścieniem przemiennym z jednością $1 = 1 + 0i$ (proponujemy Czytelnikowi przekonać się samodzielnie). Skoro $\mathbb{Z}[i] \subseteq \mathbb{C}$, to $\mathbb{Z}[i]$ jest dziedziną całkowitości. Załóżmy

$$N(a + bi) = a^2 + b^2,$$

(4) Helmut Hasse (1898–1979)

gdzie $a, b \in \mathbb{Z}$. Jeśli $c, d \in \mathbb{Z}$, $a^2 + b^2 \neq 0$ oraz $c^2 + d^2 \neq 0$, to

$$\frac{a+bi}{c+di} = \alpha + \beta i$$

dla pewnych liczb wymiernych α, β. Ponieważ

$$\begin{array}{rcl} \alpha & = & \gamma + \eta, \\ \beta & = & \delta + \psi \end{array}$$

dla $\gamma, \delta \in \mathbb{Z}$ i takich liczb wymiernych η, ψ, że

$$\begin{array}{l} |\eta| \leqslant \frac{1}{2}, \\ |\psi| \leqslant \frac{1}{2}, \end{array}$$

to

$$\begin{array}{l} (a+bi) = (c+di)(\alpha+\beta i) = (c+di)[(\gamma+\eta)+(\delta+\psi)i] = \\ = (c+di)(\gamma+\delta i) + (c+di)(\eta+\psi i) = (c+di)q + r, \end{array}$$

gdzie $q = \gamma + \delta i$, $r = (c+di)(\eta+\psi i)$. Wtedy $(c+di)(\gamma+\delta i) \in \mathbb{Z}[i]$ oraz

$$N(r) = N((c+di)(\eta+\psi i)) = (c^2+d^2)(\eta^2+\psi^2) \leqslant (c^2+d^2)(\frac{1}{4}+\frac{1}{4}) = \frac{1}{2}(c^2+d^2) < N(c+di)$$

pod warunkiem, że $(c+di)(\eta+\psi i) \neq 0$. Zatem $\mathbb{Z}[i]$ jest dziedziną euklidesową.

Lemat 4.4.3. *Niech A będzie dziedziną euklidesową z normą N. Wtedy zachodzą następujące własności:*

(1) *jeśli a, b są elementami niezerowymi z A, to*

$$\begin{cases} N(ab) = N(a), & \text{gdy } b \text{ jest odwracalny w } A, \\ N(ab) > N(a), & \text{gdy } b \text{ nie jest odwracalny w } A; \end{cases}$$

(2) *element a jest odwracalny w A w tym i tylko tym przypadku, gdy $N(a) = N(1)$.*

Dowód. Ćwiczenie. □

■ Pierścień przemienny R z jednością 1 jest nazywany *pierścieniem ideałów głównych*, jeśli dla każdego ideału $I \leqslant R$ znajdzie się taki element $a \in R$, że

$$I = \langle a \rangle.$$

Przykładowo każde ciało jest pierścieniem ideałów głównych. Jeśli R jest pierścieniem ideałów głównych i dziedziną całkowitości, to mówimy, że R jest *dziedziną ideałów głównych*.

Twierdzenie 4.4.4. *Każda dziedzina euklidesowa A jest dziedziną ideałów głównych.*

Dowód. Niech I będzie dowolnym ideałem w A. Jeśli $I = \{0\}$ jest zerowy, to jest główny. Dlatego niech I posiada pewien element niezerowy a. Rozpatrzmy zbiór

$$S = \{N(a) \mid a \in I\}.$$

Skoro $\varnothing \neq S \subseteq \mathbb{N}$, to na podstawie zasady minimum S posiada najmniejszą liczbę nieujemną, na przykład $N(a_0)$ dla pewnego elementu $a_0 \in I$. Jeśli teraz x jest dowolnym elementem z I, to znajdą się takie elementy $q, r \in A$, że

$$x = a_0 q + r,$$

gdzie $r = 0$ lub $N(r) < N(a_0)$. W wyniku minimalności $N(a_0)$ wnosimy, że $r = 0$ oraz $x = a_0 q \in a_0 A$. To oznacza, że $I \leqslant a_0 A$. Dodatkowo $a_0 \in I$, co powoduje, że $a_0 A \leqslant I$, a zatem ideał

$$I = a_0 A = \langle a_0 \rangle$$

jest główny.

□

Wniosek 4.4.5. *Zachodzą następujące własności:*

(1) *pierścień liczb całkowitych $\mathbb{Z}$ jest dziedziną ideałów głównych;*
(2) *pierścień klas reszt $\mathbb{Z}_n$ modulo n jest pierścieniem ideałów głównych;*
(3) *jeśli $\mathbb{F}$ jest ciałem, to pierścień wielomianów $\mathbb{F}[x]$ jest dziedziną ideałów głównych.*

Dowód. Części (1) oraz (3) są konsekwencjami twierdzenia 4.4.4 i przykładu 4.4.1. Własność (2) otrzymujemy w wyniku części (1). □

Przykłady 4.4.6.

(**1**) Jeśli $\mathbb{F}$ jest ciałem, a X, Y są niewiadome, to pierścień wielomianów $\mathbb{F}[X, Y]$ nie jest pierścieniem ideałów głównych, gdyż na przykład jego ideał

$$I = X\mathbb{F}[X, Y] + Y\mathbb{F}[X, Y]$$

nie jest główny.

(**2**) Pierścień wielomianów ze współczynnikami całkowitymi $\mathbb{Z}[X]$ posiada ideał, który nie jest główny. Na przykład

$$I_p = X\mathbb{Z}[X] + p\mathbb{Z}[X]$$

jest ideałem, który nie jest główny dla każdej liczby pierwszej p.

* * *

■ **Największy wspólny dzielnik.** Jeśli a, b są elementami pierścienia R, to mówimy, że element a *dzieli* element b (oznaczamy przez $a \mid b$), jeśli $b = ac$ dla pewnego $c \in R$. Relacja „|" jest zwrotna i przechodnia.

Elementy $x, y \in R$ takie, że $x \mid y$ oraz $y \mid x$ są nazywane *stowarzyszonymi* w pierścieniu R (oznaczamy to przez $x \sim y$).

Lemat 4.4.7. *Relacja stowarzyszenia „$\sim$" jest relacją równoważności w pierścieniu R.*

Dowód. Ćwiczenie. □

Lemat 4.4.8. *Niech R będzie dziedziną całkowitości z jednością oraz $x, y \in R$. Wtedy x i y są stowarzyszone w R w tym i tylko tym przypadku, gdy $x = \varepsilon y$ dla pewnego elementu odwracalnego $\varepsilon \in U(R)$.*

Dowód. ($\Leftarrow$) Oczywiście, że $y \mid x$. Skoro $y = \varepsilon^{-1}x$, to $x \mid y$.

($\Rightarrow$) Załóżmy, że $x \mid y$ oraz $y \mid x$. Jeśli $x = 0$ (odpowiednio $y = 0$), to $x = 0$ (odpowiednio $x = 0$). Wtedy $x = 1 \cdot y$ oraz $y = 1 \cdot x$ i teza zachodzi. Dlatego załóżmy, że $x \neq 0$ i $y \neq 0$. Ponieważ $x = z_1 y$ oraz $y = z_2 x$ dla pewnych $z_1, z_2 \in R$, to

$$(1 - z_1 z_2)x = 0 = (1 - z_2 z_1)y.$$

Zatem $z_1 z_2 = 1 = z_2 z_1$, co implikuje, że $x \sim y$. □

■ *Największym wspólnym dzielnikiem* elementów $a, b \in R$ będziemy nazywać element $d \in R$ (i oznaczać przez $d = \mathrm{NWD}(a, b)$) taki, że:
1) $d \mid a$ oraz $d \mid b$;
2) jeśli $c \in R$ oraz $c \mid a$ i $c \mid b$, to $c \mid d$.

Twierdzenie 4.4.9. *Niech R będzie pierścieniem euklidesowym oraz $a, b \in R$. Jeśli $b \neq 0$, to* $\mathrm{NWD}(a, b)$ *istnieje.*

Dowód. (Szkic) Podobnie jak w dowodzie twierdzenia 1.2.6, na podstawie własności pierścienia euklidesowego, wnioskujemy, że istnieją $q_i, r_i \in R$ oraz taka dodatnia liczba naturalna k $(i = 1, \ldots, k, k+1)$, że

$$\begin{array}{lcrcllcrcl}
a & = & bq_1 & + & r_1, & r_1 & \neq & 0, & \delta(r_1) & < & \delta(b), \\
b & = & r_1q_2 & + & r_2, & r_2 & \neq & 0, & \delta(r_2) & < & \delta(r_1), \\
r_1 & = & r_2q_3 & + & r_3, & r_3 & \neq & 0, & \delta(r_3) & < & \delta(r_2), \\
 & \vdots & & \vdots & & & \vdots & & & \vdots & \\
r_{k-3} & = & r_{k-2}q_{k-1} & + & r_{k-1}, & r_{k-1} & \neq & 0, & \delta(r_{k-1}) & < & \delta(r_{k-2}), \\
r_{k-2} & = & r_{k-1}q_k & + & r_k, & r_k & \neq & 0, & \delta(r_k) & < & \delta(r_{k-1}), \\
r_{k-1} & = & r_kq_{k+1} & + & r_{k+1}, & r_{k+1} & = & 0. & & &
\end{array}$$

Ten ciąg kolejnych dzieleń jest nazywany *algorytmem Euklidesa*. Ostatnia reszta niezerowa

$$r_k = \mathrm{NWD}(a, b)$$

jest największym wspólnym dzielnikiem elementów a oraz b. □

■ Największy wspólny dzielnik $\mathrm{NWD}(a, b)$ dwóch elementów $a, b \in R$ w każdej dziedzinie całkowitości R z jednością (jeśli istnieje) jest określony jednoznacznie (z dokładnością do stowarzyszoności). To znaczy, że każde dwa największe wspólne dzielniki elementów a i b są stowarzyszone.

Twierdzenie 4.4.10. *Niech R będzie dziedziną ideałów głównych oraz $a, b \in R$. Jeśli $b \neq 0$, to istnieją takie elementy $u, v \in R$, że*

$$\mathrm{NWD}(a, b) = au + bv.$$

Dowód. Skoro R jest pierścieniem ideałów głównych, to suma ideałów

$$\langle a \rangle + \langle b \rangle = \langle d \rangle$$

jest ideałem głównym generowanym przez pewien element $d \in R$. Wtedy $d = au + bv$ dla pewnych $u, v \in R$. Ponieważ $a, b \in \langle d \rangle$, to $d \mid a$ oraz $d \mid b$. Załóżmy, że element $c \in R$ jest taki, że $c \mid a$ oraz $c \mid b$. Wtedy $a = cx$ i $b = cy$ dla pewnych $x, y \in R$, a stąd

$$d = au + bv = cxu + cyv = c(xu + yv).$$

Zatem $c \mid d$. Wnosimy, że $d = \mathrm{NWD}(a, b)$. □

■ Elementy $a, b \in R$ są nazywane *względnie pierwszymi* w dziedzinie całkowitości R, jeśli $\mathrm{NWD}(a, b) = 1$.

■ W pierścieniu przemiennym R z jednością 1 element $a \in R$ jest stowarzyszony z 1 (czyli $a \sim 1$) wtedy i tylko wtedy, gdy $a \in U(R)$.

Rzeczywiście, $1 \mid a$ oraz $a \mid 1$ w tym i tylko tym przypadku, gdy $1 = ab$ dla pewnego $b \in R$, a więc $a \in U(R)$.

Wniosek 4.4.11. *Niech R będzie pierścieniem euklidesowym oraz $a, b \in R$. Wtedy elementy a i b są względnie pierwsze w tym i tylko tym przypadku, gdy $au + bv = 1$ dla pewnych $u, v \in R$.*

Dowód. ($\Rightarrow$) Wynika z twierdzenia 4.4.10.

($\Leftarrow$) Załóżmy, że $au + bv = 1$ dla pewnych $u, v \in R$ oraz największy wspólny dzielnik $\mathrm{NWD}(a, b) = d$. Skoro $d \mid a$ oraz $d \mid b$, to $d \mid 1$. Jednak $1 \mid d$, a więc $d \sim 1$ i teza zachodzi. □

Stwierdzenie 4.4.12. *Niech R będzie dziedziną całkowitości oraz $a, b \in R$. Wtedy $\langle a \rangle = \langle b \rangle$ w tym i tylko tym przypadku, gdy elementy a i b są stowarzyszone.*

Dowód. ($\Rightarrow$) Jeśli $a = 0$, to $b = 0$, a więc a i b są stowarzyszone (bo $0 = 1 \cdot 0$). Dlatego załóżmy, że $a \neq 0$, a więc i $b \neq 0$. Skoro $a = a \cdot 1 \in \langle a \rangle$, to $a \in \langle b \rangle$, a stąd $a = bx$ dla pewnego $x \in R$. Podobnymi rozumowaniami otrzymujemy $b = ay$ dla pewnego $y \in R$. Wtedy

$$a = bx = (ay)x = a(xy),$$

skąd $a(1 - xy) = 0$, co implikuje, że $yx = 1$. Zatem elementy a i b są stowarzyszone.

($\Leftarrow$) Jeśli a oraz b są stowarzyszone w R, to $b = at \in \langle a \rangle$ dla pewnego elementu $t \in R$, który jest odwracalny w R. Zatem $\langle b \rangle \leqslant \langle a \rangle$. Ponadto $a = bt^{-1} \in \langle b \rangle$, a więc $\langle a \rangle \subseteq \langle b \rangle$. Wnosimy, że $\langle a \rangle = \langle b \rangle$. □

Lemat 4.4.13. *Niech $\mathbb{F}$ będzie ciałem oraz $f, g \in \mathbb{F}[X]$. Wtedy są spełnione następujące własności:*

(1) $f \mid g \Leftrightarrow g \in \langle f \rangle$;

(2) $g \in \langle f \rangle \Leftrightarrow \langle g \rangle \subseteq \langle f \rangle$.

Dowód. (1) Istotnie, $f \mid g$ wtedy i tylko wtedy, gdy $g = fu$ dla pewnego wielomianu $u \in \mathbb{F}[X]$, a to jest równoważne z tym, że $g \in \langle f \rangle$.

(2) Jeśli $g \in \langle f \rangle$, to $gr \in \langle f \rangle$ dla każdego wielomianu $r \in \mathbb{F}[X]$. A to oznacza, że $\langle g \rangle \subseteq \langle f \rangle$. Odwrotnie, niech $\langle g \rangle \subseteq \langle f \rangle$. Skoro $g = g \cdot 1 \in \langle g \rangle$, to $g \in \langle f \rangle$.

□

■ Jeśli X jest podzbiorem pierścienia R, to przecięcie

$$\bigcap \{I \mid I \text{ jest ideałem w } R \text{ takim, że } I \supseteq X\}$$

jest ideałem w R (udowodnić samodzielnie), który jest oznaczany symbolem $\langle X \rangle$ i nazywany ideałem *generowanym* przez zbiór X. Jeśli $X = \varnothing$ jest zbiorem pustym, to $\langle \varnothing \rangle$ jest ideałem zerowym (bo $\langle \varnothing \rangle$ jest przecięciem wszystkich ideałów pierścienia R). Jeśli zaś $X = \{x_1, \ldots, x_n\}$ oraz $I = \langle X \rangle$, to zapisujemy, że $I = \langle x_1, \ldots, x_n \rangle$. W pierścieniu przemiennym R z jednością zachodzi równość

$$\langle x_1, \ldots, x_n \rangle = x_1 R + \cdots + x_n R$$

(przekonać się samodzielnie).

* * *

■ **Najmniejsza wspólna wielokrotność.** Niech a, b, c będą elementami dziedziny całkowitości R. *Najmniejszą wspólną wielokrotnością* elementów a oraz b jest nazywany taki element $m \in R$, który spełnia warunki:
1) $a \mid m$ oraz $b \mid m$;
2) jeśli $a \mid c$ oraz $b \mid c$, to $m \mid c$.
Wtedy zapisujemy $m = \mathrm{NWW}(a, b)$.

Twierdzenie 4.4.14. *Niech a, b będą elementami niezerowymi dziedziny całkowitości R. Jeśli $ab = \mathrm{NWW}(a, b) \cdot t$ dla pewengo $t \in R$, to $t = \mathrm{NWD}(a, b)$.*

Dowód. Ćwiczenie. □

Ćwiczenia 4.4.15.

(**1**) Niech R będzie dziedziną całkowitości. Udowodnić, że jeśli pierścień wielomianów $R[X]$ jest pierścieniem ideałów głównych, to R jest ciałem.
(**2**) Udowodnić, że ideał I nie jest główny w pierścieniu A, jeśli:
(a) $I = \langle X, Y\rangle$ oraz $A = \mathbb{R}[X, Y]$;
(b) $I = \langle 2, X\rangle$ oraz $A = \mathbb{Z}[X]$;
(c) $I = \langle 2, 1 - \sqrt{5}\rangle$ oraz $A = \mathbb{Z}[\sqrt{5}]$.
(**3**) Udowodnić, że $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ jest dziedziną ideałów głównych, lecz nie jest euklidesowy.
(**4**) Udowodnić, że $\mathbb{Z}\left[\sqrt{-5}\right]$ nie jest pierścieniem ideałów głównych (wskazówka: udowodnić, że ideał $\langle 2 + \sqrt{-5}, 3\rangle$ nie jest główny).
(**5**) Znaleźć NWD$(1 + 13i, 85)$ w pierścieniu $\mathbb{Z}[i]$.
(**6**) Znaleźć NWD$(18 + i, 6 - 17i)$ w pierścieniu $\mathbb{Z}[i]$.
(**7**) Udowodnić izomorficzność pierścieni $\mathbb{Z}[i] \cong \mathbb{Z}[X]/\langle X^2 + 1\rangle$.
(**8**) Znaleźć grupę multiplikatywną $U(R)$ pierścienia R, jeśli:
(a) $R = \mathbb{Z}[i]$;
(b) $R = \mathbb{Z}[e^{\frac{2\pi i}{3}}]$.
(**9**) Udowodnić, że $\mathbb{Z}[\sqrt{n}] = \{a + b\sqrt{n} \mid a, b \in \mathbb{Z}\}$ jest dziedziną euklidesową, jeśli:
(a) $n = -2$;
(b) $n = 2$;
(c) $n = 3$.

Uwagi. Dowód istnienia największego wspólnego dzielnika dwóch liczb całkowitych jest zawarty w VII księdze Euklidesa, na samym jej początku. Metoda dzielenia wielomianów z resztą, która faktycznie przywiodła do wprowadzenia pierścienia euklidesowego, została wypracowana przez P. Ruffiniego.

4.5. Ideały pierwsze i maksymalne

■ **Ideały pierwsze.** Niech R będzie pierścieniem przemiennym z jednością 1. Ideał P pierścienia R jest nazywany *pierwszym*, jeśli są spełnione takie dwa warunki:
1) $P \neq R$;
2) dla dowolnych elementów $a, b \in R$ zachodzi implikacja

$$ab \in P \quad \Rightarrow \quad a \in P \ \text{ lub } \ b \in P.$$

Lemat 4.5.1. *Niech a, b, c, d będą dowolnymi liczbami całkowitymi. Wtedy zachodzą następujące własności:*
(1) $c \mid d \Leftrightarrow d\mathbb{Z} \subseteq c\mathbb{Z}$;
(2) $|c| = |d| \Leftrightarrow d\mathbb{Z} = c\mathbb{Z}$;
(3) $m\mathbb{Z} = \mathbb{Z} \Leftrightarrow m = \pm 1$.

Dowód. Ćwiczenie (patrz dowód lematu 4.4.13). □

Przykłady 4.5.2.

(**1**) Niech p będzie liczbą pierwszą oraz $a, b \in \mathbb{Z}$. Jeśli $ab \in p\mathbb{Z}$, to $ab = pt$ dla pewnej liczby całkowitej t. Wtedy p dzieli ab i na podstawie lematu 1.2.13 mamy $p \mid a$ (czyli $a \in p\mathbb{Z}$) lub $p \mid b$ (czyli $b \in p\mathbb{Z}$). Zatem $p\mathbb{Z}$ jest ideałem pierwszym pierścienia $\mathbb{Z}$ dla każdej liczby pierwszej p.

(**2**) Niech $\mathbb{F}$ będzie ciałem oraz $p \in \mathbb{F}[X]$ będzie wielomianem nieprzywiedlnym nad ciałem $\mathbb{F}$. Z tych samych rozumowań, co w poprzednim przykładzie (1), ideał główny

$$\langle p \rangle = p\mathbb{F}[X]$$

jest ideałem pierwszym pierścienia wielomianów $\mathbb{F}[X]$.

Twierdzenie 4.5.3. *Niech R będzie pierścieniem przemiennym z jednością 1. Wtedy ideał P jest pierwszy w pierścieniu R w tym i tylko tym przypadku, gdy pierścień ilorazowy R/P jest dziedziną całkowitości.*

Dowód. Niech a, b będą dowolnymi elementami pierścienia R oraz $\overline{a} = a + P$, $\overline{b} = b + P$. Najpierw przekonajmy się, że

$$\overline{a} = \overline{0} \Leftrightarrow a \in P.$$

Istotnie, na podstawie kryterium równości warstw $\overline{a} = \overline{0}$ wtedy i tylko wtedy, gdy $a + P = P = 0 + P$, co na podstawie lematu 3.3.5 jest równoważne z tym, że $a \in P$.

($\Rightarrow$) Załóżmy, że

$$\overline{a} \neq \overline{0} \quad \text{oraz} \quad \overline{a}\,\overline{b} = \overline{0}. \tag{4.3}$$

Mamy $\overline{a} \neq \overline{0}$ wtedy i tylko wtedy, gdy $a+P \neq 0+P$, a to jest równoważne z $a \notin P$. Warunek

$$\overline{a}\,\overline{b} = \overline{0} \tag{4.4}$$

możemy przepisać w postaci równoważnej

$$ab + P = (a + P)(b + P) = \overline{a}\,\overline{b} = \overline{0} = 0 + P,$$

co na mocy lematu 3.3.5 oznacza, że $ab \in P$. Skoro P jest ideałem pierwszym w R oraz $a \notin P$, to $b \in P$ lub równoważnie

$$\overline{b} = b + P = 0 + P = \overline{0}.$$

Zatem z (4.3) wynika, że $\overline{b} = \overline{0}$, a więc pierścień ilorazowy R/P nie posiada dzielników zera (czyli R/P jest dziedziną całkowitości).

($\Leftarrow$) Niech R/P będzie dziedziną całkowitości. Wtedy z warunku (4.4) (czyli $ab \in P$) wynika, że $\overline{a} = \overline{0}$ (czyli $a \in P$) lub $\overline{b} = \overline{0}$ (czyli $b \in P$).

□

* * *

■ **Ideały maksymalne.** Niech R będzie pierścieniem przemiennym z jednością 1. Ideał M pierścienia R jest nazywany *maksymalnym*, jeśli są spełnione takie dwa warunki:

1) $M \neq R$;
2) jeśli S jest ideałem w R takim, że $M \leqslant S \leqslant R$, to

$$S = M \quad \text{lub} \quad S = R.$$

Twierdzenie 4.5.4 (o istnieniu ideału maksymalnego)**.** *Niech R będzie pierścieniem z jednością 1. Wtedy dla każdego ideału właściwego I pierścienia R istnieje ideał maksymalny M w R taki, że $I \leqslant M$.*

Dowód. Zbiór

$$\mathcal{A} = \{J \mid J \text{ jest ideałem w } R \text{ zawierającym ideał } I\}$$

jest częściowo uporządkowany względem relacji zawierania „$\subseteq$”. Jeśli mamy liniowo uporządkowany podzbiór ideałów $\{J_\lambda\} \subseteq \mathcal{A}$, to

$$1 \notin \bigcup_\lambda J_\lambda,$$

a więc

$$\bigcup_\lambda J_\lambda \in \mathcal{A}.$$

Na podstawie lematu Kuratowskiego-Zorna (patrz lemat 1.3.26) zbiór $\mathcal{A}$ posiada element maksymalny, co kończy dowód. □

Twierdzenie 4.5.5. *Niech R będzie pierścieniem przemiennym z jednością. Wtedy ideał M jest maksymalny w pieścieniu R w tym i tylko tym przypadku, gdy pierścień ilorazowy R/M jest ciałem.*

Dowód. Niech a, b będą dowolnymi elementami z R, $\overline{a} = a + M$ oraz $\overline{b} = b + M$.

($\Rightarrow$) Załóżmy, że M jest ideałem maksymalnym w R oraz $\overline{a}$ jest elementem niezerowym pierścienia ilorazowego R/M. Wtedy $a \notin M$ i, jako wniosek,

$$M + aR \neq M \text{ oraz } M \leqslant M + aR.$$

Łatwo przekonać się, że $M + aR$ jest ideałem pierścienia R. Na podstawie definicji ideału maksymalnego wnosimy, że

$$M + aR = R.$$

A zatem $1 = m + ar$ dla pewnych elementów $m \in M$ oraz $r \in R$. Stąd

$$\overline{1} = 1 + M = ar + (m + M) = ar + M = (a + M)(r + M) = \overline{a}\,\overline{r}$$

i każdy element niezerowy $\overline{a}$ posiada odwrotny do siebie $\overline{r}$ w pierścieniu ilorazowym R/M. Zatem R/M jest ciałem.

($\Leftarrow$) Niech R/M będzie ciałem. Wtedy $M \neq R$. Załóżmy, że J jest ideałem w R takim, że $M \leqslant J \leqslant R$. Są możliwe dwa przypadki: $J = M$ lub $J \neq M$. Jeśli $J \neq M$, to znajdzie się element $a \in J \setminus M$. Wtedy

$$\overline{a} \neq \overline{0} \text{ oraz } \overline{a}\,\overline{b} = \overline{1}$$

dla pewnego elementu $\overline{b} \in R/M$. Za lematem 4.1.4 mamy

$$\overline{a}(R/M) = R/M,$$

skąd otrzymujemy

$$R = aR + M.$$

To oznacza, że $J = R$. Zatem ideał M jest maksymalny w R. □

■ W wyniku twierdzeń 4.5.3 i 4.5.5 wnosimy, że każdy ideał maksymalny pierścienia przemiennego jest pierwszy. Odwrotna implikacja nie zachodzi. Świadczy o tym przykład 4.5.6 (2).

Przykłady 4.5.6.

(**1**) Znajdźmy wszystkie ideały maksymalne pierścienia liczb całkowitych $\mathbb{Z}$. W tym celu załóżmy, że M jest ideałem maksymalnym w $\mathbb{Z}$. Wtedy $M = m\mathbb{Z}$ dla pewnej niezerowej liczby całkowitej m na podstawie wniosku 4.4.5. Przekonajmy się, że m jest liczbą pierwszą. Nie wprost. Jeśli $m = \pm 1$, to $M = R$, a to nie jest możliwe. Zatem $m = m_1 m_2$ dla pewnych liczb całkowitych m_i takich, że $1 < m_i < m$ $(i = 1, 2)$. Skoro $m_1 \neq 1$, to

$$m_1\mathbb{Z} \leqslant \mathbb{Z} \text{ oraz } m_1\mathbb{Z} \neq \mathbb{Z}.$$

Podobnie z nierówności $m_1 < m$ wynika, że

$$m\mathbb{Z} \leqslant m_1\mathbb{Z} \text{ oraz } m\mathbb{Z} \neq m_1\mathbb{Z}.$$

Otrzymujemy sprzeczność z maksymalnością ideału $m\mathbb{Z}$. Z tych rozumowań wynika, że liczba m jest pierwsza. Zatem, jeśli M jest ideałem maksymalnym w pierścieniu liczb całkowitych $\mathbb{Z}$, to

$$M = m\mathbb{Z}$$

dla pewnej liczby pierwszej m. Zostawiamy Czytelnikowi, aby przekonał się samodzielnie, że jeśli p jest liczbą pierwszą, to ideał $p\mathbb{Z}$ jest maksymalny w $\mathbb{Z}$.

(**2**) Jeśli $\mathbb{Z}[X]$ jest pierścieniem wielomianów ze współczynnikami całkowitymi, to $P = X\mathbb{Z}[X]$ jest jego ideałem pierwszym, właściwym sposobem zawierającym się w ideale $X\mathbb{Z}[X] + 2\mathbb{Z}[X]$. Jeśli założyć, że

$$X\mathbb{Z}[X] + 2\mathbb{Z}[X] = \mathbb{Z}[X],$$

to otrzymujemy $1 = Xf + 2g$ dla pewnych wielomianów $f, g \in \mathbb{Z}[X]$. Stąd wynika, że $1 = 2b$, gdzie b jest współczynnikiem wolnym wielomianu g. A to nie jest możliwe. Zatem

$$X\mathbb{Z}[X] + 2\mathbb{Z}[X]$$

jest ideałem właściwym w $\mathbb{Z}[X]$. Przekonaliśmy się, że nie każdy ideał pierwszy jest maksymalny.

Wniosek 4.5.7. *Niech $\mathbb{F}$ będzie ciałem oraz $p \in \mathbb{F}[X]$. Wtedy następujące własności są równoważne:*

(1) $\mathbb{F}[X]/\langle p\rangle$ *jest ciałem;*
(2) $\mathbb{F}[X]/\langle p\rangle$ *jest dziedziną całkowitości;*
(3) *wielomian p jest nieprzywiedlny nad ciałem $\mathbb{F}$.*

Dowód. (1) $\Rightarrow$ (2) Wynika na podstawie definicji ciała i przemiennej dziedziny całkowitości.

(2) $\Rightarrow$ (3) Na mocy twierdzenia 4.5.3 ideał $P = \langle p\rangle$ jest pierwszy. Jeśli

$$p = p_1 p_2$$

dla pewnych wielomianów $p_1, p_2 \in \mathbb{F}[X]$, to w wyniku pierwszości ideału P otrzymujemy $p_1 \in \langle p\rangle$ (czyli $p \mid p_1$) lub $p_2 \in \langle p\rangle$ (czyli $p \mid p_2$). To implikuje, że wielomian p jest nieprzywiedlny nad ciałem $\mathbb{F}$.

(3) $\Rightarrow$ (1) Niech $P = \langle p\rangle$ oraz J będzie takim ideałem pierścienia $\mathbb{F}[X]$, że $P \leqslant J \leqslant \mathbb{F}[X]$. Z wniosku 4.4.5 otrzymujemy, że $J = \langle f\rangle$ dla pewnego wielomianu $f \in \mathbb{F}[X]$. Za lematem 4.4.13 wnosimy, że f dzieli p i w wyniku nieprzywiedlności wielomianu p otrzymujemy, że jego stopień $\deg f = 0$ jest zerowy (a więc $f \in \mathbb{F} \setminus \{0\}$ jest odwracalny w pierścieniu $\mathbb{F}[X]$) lub $\deg f = \deg p$ (wtedy $f = up$ dla pewnego elementu niezerowego u ciała $\mathbb{F}$). W pierwszym przypadku $J = \mathbb{F}[X]$ za lematem 4.1.4, a w drugim przypadku zachodzą równości

$$J = \langle f\rangle = \langle p\rangle = P.$$

Wnosimy, że P jest ideałem maksymalnym pierścienia $\mathbb{F}[X]$, a więc $\mathbb{F}[X]/P$ jest ciałem na podstawie twierdzenia 4.5.5.

□

Przykłady 4.5.8.

(**1**) Niech

$$p = X^2 + X + 1 \in \mathbb{F}_2[X] \text{ oraz } P = \langle X^2 + X + 1\rangle.$$

Skoro wielomian p jest nieprzywiedlny nad ciałem $\mathbb{F}_2$, to ideał P jest pierwszy. Jeśli f jest dowolnym wielomianem z $\mathbb{F}_2[X]$, to

$$f = (X^2 + X + 1)q + r,\ \deg r < \deg(X^2 + X + 1) = 2$$

dla pewnych $q, r \in \mathbb{F}_2[X]$. Wtedy $r = aX + b$ dla pewnych $a, b \in \mathbb{F}_2$, a więc warstwa

$$f + P = aX + b + (X^2 + X + 1)q + P = aX + b + P.$$

Jako wniosek ciało

$$\mathbb{F}_2[X]/P = \{0 + P, 1 + P, X + P, 1 + X + P\}$$

składa się z czterech elementów. W celu uproszczenia oznaczmy elementy tego ciała odpowiednio przez

$$0,\ 1,\ x \text{ oraz } 1 + x.$$

Skonstruujmy tabelki Cayleya działań dodawania i mnożenia w ciele $\mathbb{F}_2[X]/P$:

$+$	0	1	x	$1+x$
0	0	1	x	$1+x$
1	1	0	$1+x$	x
x	x	$1+x$	0	1
$1+x$	$1+x$	x	1	0

$\cdot$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	$1+x$	1
$1+x$	0	$1+x$	1	x

Niech teraz $t = X^2 + 1 \in \mathbb{F}_2[X]$ oraz $I = \langle X^2 + 1 \rangle$. Ponieważ $t(1) = 0$ dla elementu $1 \in \mathbb{F}_2$, to wielomian t jest przywiedlny nad ciałem $\mathbb{F}_2$ na podstawie twierdzenia Bezouta (patrz później twierdzenie 6.1.3), a więc pierścień ilorazowy $\mathbb{F}_2[X]/I$ nie jest ciałem. Z podobnych rozumowań, jak wyżej, pierścień ilorazowy $\mathbb{F}_2[X]/I$ składa się z czterech elementów

$$0 + I,\ 1 + I,\ X + I \text{ oraz } 1 + X + I,$$

które krótko będziemy oznaczać odpowiednio przez 0, 1, x i $1 + x$. Wtedy tabelki Cayleya dla dodawania i mnożenia w pierścieniu ilorazowym $\mathbb{F}_2[x]/I$ przyjmują takie postacie:

$+$	0	1	x	$1+x$
0	0	1	x	$1+x$
1	1	0	$1+x$	x
x	x	$1+x$	0	1
$1+x$	$1+x$	x	1	0

$\cdot$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	1	$1+x$
$1+x$	0	$1+x$	$1+x$	0

Ponieważ $(1 + x)^2 = 0$, to $1 + x$ jest dzielnikiem zera w pierścieniu

$$\mathbb{F}_2[X]/\langle X^2 + 1 \rangle.$$

(**2**) Niech $R[[X]]$ będzie pierścieniem formalnych szeregów potęgowych nad pierścieniem przemiennym R. Załóżmy, że dwa szeregi

$$f = \sum_{i=0}^{\infty} a_i X^i,\ g = \sum_{j=0}^{\infty} b_j X^j \in R[[X]]$$

są takie, że $fg = 1$. Wtedy otrzymujemy nieskończony układ równań

$$\begin{aligned} b_0 a_0 &= 1, \\ a_0 b_1 + a_1 b_0 &= 0, \\ a_0 b_2 + a_1 b_1 + a_2 b_0 &= 0, \\ &\vdots \\ \textstyle\sum_{i+j=k} a_i b_j &= 0, \\ &\vdots \end{aligned}$$

Przyjmując, że elementy $a_i \in R$ $(i \in \mathbb{N})$ są wiadome oraz istnieje element odwrotny $a_0^{-1} \in R$, obliczamy b_i. Z tych rozumowań otrzymujemy następny

Lemat 4.5.9. *Niech R będzie pierścieniem przemiennym z jednością 1 oraz $a_i \in R$ $(i \in \mathbb{N})$. Formalny szereg potęgowy*

$$\sum_{i=0}^{\infty} a_i X^i$$

jest odwracalny w pierścieniu $R[[X]]$ wtedy i tylko wtedy, gdy jego wyraz wolny a_0 jest odwracalny w pierścieniu R.

□

Wniosek 4.5.10. *Niech $\mathbb{F}$ będzie ciałem oraz $a_i \in \mathbb{F}$ $(i \in \mathbb{N})$. Wtedy formalny szereg potęgowy*

$$\sum_{i=0}^{\infty} a_i X^i$$

jest odwracalny w pierścieniu $\mathbb{F}[[X]]$ wtedy i tylko wtedy, gdy $a_0 \neq 0$.

□

Proponujemy Czytelnikowi samodzielnie przekonać się, że każdy ideał niezerowy I w pierścieniu $\mathbb{F}[[X]]$ ma postać $X^k\mathbb{F}[[X]]$, gdzie k jest taką najmniejszą liczbą całkowitą, że $a_k \neq 0$ oraz

$$\sum_{i=k}^{\infty} a_i X^i \in I.$$

* * *

■ **Elementy pierwsze.** Niech R będzie dziedziną całkowitości. Wtedy:

- jeśli element $z \in R$ jest niezerowy i nie jest odwracalny w R, to jest nazywany *nierozkładalnym* (=*nieprzywiedlnym*) w R, jeśli z tego, że $z = xy$ dla pewnych $x, y \in R$ wynika, że jeden z elementów x lub y jest odwracalny w R;

- element niezerowy $p \in R$ jest nazywany *pierwszym* w R, jeśli ideał główny $\langle p \rangle$ jest ideałem pierwszym w pierścieniu R.

Lemat 4.5.11. *W dziedzinie całkowitości R każdy element pierwszy p jest nierozkładalny.*

Dowód. Załóżmy, że $p = xy$ dla pewnych $x, y \in R$. Skoro $xy \in \langle p \rangle$, to $x \in \langle p \rangle$ lub $y \in \langle p \rangle$. Niech dla pewności $x \in \langle p \rangle$, a więc $x = pt$ dla pewnego $t \in R$. Wtedy $p = xy = p(ty)$, a stąd $ty = 1$, co implikuje, że element p jest nierozkładalny. □

W przypadku dziedziny ideałów głównych możemy powiedzieć więcej.

Twierdzenie 4.5.12. *Niech R będzie dziedziną ideałów głównych oraz $p \in R$. Wtedy p jest pierwszy w R w tym i tylko tym przypadku, gdy p jest nierozkładalny w R.*

Dowód. ($\Rightarrow$) Wynika z lematu 4.5.11.

($\Leftarrow$) Załóżmy, że element p jest nierozkładalny w pierścieniu R. Niech M będzie ideałem w R takim, że $\langle p \rangle \leqslant M$. Skoro $M = \langle a \rangle$ dla pewnego elementu $a \in R$, to $p = at$ dla pewnego $t \in R$. Lecz element p jest nierozkładalny w R, a to powoduje, że $a \in U(R)$ lub $t \in U(R)$, a więc $M = \langle p \rangle$ lub $M = R$. To znaczy, że ideał $\langle p \rangle$ jest maksymalny w R, a zatem jest pierwszy w R. Wnosimy, że element p jest pierwszy w pierścieniu R. □

Przykłady 4.5.13.

(**1**) Liczba całkowita 2 jest pierwsza w pierścieniu liczb całkowitych $\mathbb{Z}$. Lecz element

$$2 = (1 - i)(1 + i)$$

jest rozkładalny w pierścieniu całkowitych liczb gaussowskich $\mathbb{Z}[i]$. Skoro $\mathbb{Z}[i]$ jest dziedziną ideałów głównych (patrz przykład 4.4.1 i twierdzenie 4.4.4), to element 2 nie jest pierwszy w pierścieniu $\mathbb{Z}[i]$ na podstawie twierdzenia 4.5.12.

(**2**) Liczba 3 jest elementem pierwszym w pierścieniu $\mathbb{Z}$. Jeśli założyć, że

$$3 = (a + bi)(x + iy)$$

dla pewnych $a, b, x, y \in \mathbb{Z}$, to norma

$$9 = N(3) = (a^2 + b^2)(x^2 + y^2).$$

Wtedy $a^2 + b^2 \in \{1, 3, 9\}$ i mamy trzy przypadki:

- $$a^2 + b^2 = 1 \;\Rightarrow\; a + ib \in \{\pm 1, \pm i\} \;\Rightarrow\; a \in U(\mathbb{Z}[i]) \text{ oraz } x + yi \in \{\pm 3, \pm 3i\};$$

  zatem element 3 jest nierozkładalny, a więc jest pierwszy w pierścieniu $\mathbb{Z}[i]$;
- $a^2 + b^2 = 3$, co nie jest możliwe;
- $$a^2 + b^2 = 9 \;\Rightarrow\; a + ib \in \{\pm 3, \pm 3i\} \;\Rightarrow\; x + yi \in \{\pm 1, \pm i\};$$

  zatem element 3 jest nierozkładalny, a więc jest pierwszy w pierścieniu $\mathbb{Z}[i]$ na podstawie twierdzenia 4.5.12.

* * *

■ **Chińskie twierdzenie o resztach.** Zachodzi taki

Lemat 4.5.14. *Niech $R_1, \ldots, R_n$ będą pierścieniami przemiennymi z elementami jednostkowymi. Wtedy*

$$R = R_1 \oplus \cdots \oplus R_n = \{(r_1, \ldots, r_n) \mid r_i \in R_i \ (i = 1, \ldots, n)\}$$

jest pierścieniem przemiennym z jednością względem działań określonych przez wzory:

$$\begin{aligned} (a_1, \ldots, a_n) + (b_1, \ldots, b_n) &= (a_1 + b_1, \ldots, a_n + b_n), \\ (a_1, \ldots, a_n)(b_1, \ldots, b_n) &= (a_1 b_1, \ldots, a_n b_n), \end{aligned}$$

gdzie $a_i, b_i \in R_i$ $(i = 1, \ldots, n)$.

Dowód. Ćwiczenie. □

Nietrudno sprawdzić (patrz lemat 4.5.14), że

(*a*) $$I_s = \{(0, \ldots, 0, r_s, 0, \ldots, 0) \mid r_s \in R_s\}$$

jest ideałem pierścienia R $(s = 1, \ldots, n)$;

(*b*) $R = I_1 + \cdots + I_n$, gdzie suma

$$I_1 + \cdots + I_n = \{i_1 + \cdots + i_n \mid i_s \in I_s \ (s = 1, \ldots, n)\};$$

(*c*) przecięcie

$$I_s \bigcap (I_1 + \cdots + \widehat{I_s} + \cdots + I_n) = \langle 0 \rangle$$

jest zerowe dla każdego $s = 1, \ldots, n$ (symbol $\widehat{\ }$ nad I_s oznacza, że składnik I_s nie jest obecny w sumie).

■ Pierścień

$$R = R_1 \oplus \cdots \oplus R_n$$

jest nazywany (zewnętrzną) *sumą prostą* pierścieni $R_1, \ldots, R_n$. Z tą konstrukcją są związane takie homomorfizmy pierścieni:
α) *włożenie*

$$i : R_k \ni r_k \mapsto (0, \ldots, 0, r_k, 0, \ldots, 0) \in R;$$

β) *k-ty rzut*

$$\pi_k : R \ni (r_1, \ldots, r_k, \ldots, r_n) \mapsto r_k \in R_k,$$

gdzie $r_i \in R_i$ $(i = 1, \ldots, k, \ldots, n)$.

■ Będziemy mówić, że R jest (wewnętrzną) *sumą prostą* ideałów $I_1, \ldots, I_n$, jeśli są spełnione warunki (a), (b) oraz (c). Tak jak dla grup (patrz twierdzenie 3.11.3), pojęcia sum prostych wewnętrznej i zewnętrznej ideałów są identyczne.

Dalej nam będzie potrzebny następny

Lemat 4.5.15. *Niech R będzie pierścieniem przemiennym z jednością, a $I, I_1, \ldots, I_n$ będą jego ideałami. Jeśli $I + I_s = R$ dla wszystkich $s = 1, \ldots, n$, to*

$$I + I_1 \cdots I_n = R = I + (I_1 \bigcap \cdots \bigcap I_n).$$

Dowód. Stosując indukcję względem liczby n, wykażmy, że $R = I + I_1 \cdots I_n$. W rzeczy samej, jeśli $n = 1$, to teza zachodzi na mocy założenia. Jeśli $n = 2$, to

$$R = I + I_1 = I + I_2,$$

a zatem znajdą się takie elementy $a, b \in I$, $c_1 \in I_1$ oraz $c_2 \in I_2$, że

$$a + c_1 = 1 = b + c_2.$$

Wtedy

$$1 = 1^2 = (a + c_1)(b + c_2) = (ab + ac_2 + c_1 b) + c_1 c_2 \in I + I_1 I_2$$

i dostajemy $I + I_1 I_2 = R$.

Załóżmy, że teza zachodzi dla $n-1$, czyli z $I+I_s=R$ dla wszystkich $s=1,\ldots,n-1$ wynika, że

$$R=I+I_1\cdots I_{n-1}.$$

Wtedy

$$x+u_1=1=y+u_2$$

dla pewnych elementów $x,y\in I$, $u_1\in I_n$ oraz $u_2\in I_1\cdots I_{n-1}$. Stosując podobne rozumowanie, otrzymujemy

$$\begin{gathered}1=1^2=(x+u_1)(y+u_2)=\\=(xy+xu_2+u_1y)+u_1u_2\in I+I_1\cdots I_{n-1}I_n,\end{gathered}$$

a więc

$$I+I_1\cdots I_{n-1}I_n=R.$$

Zatem jedna z równości zachodzi. Skoro

$$I_1\cdots I_n\subseteq I_1\bigcap\cdots\bigcap I_n,$$

to zachodzi i druga równość. □

Twierdzenie 4.5.16 (chińskie o resztach)**.** *Niech R będzie pierścieniem przemiennym z jednością 1, a $I_1,\ldots,I_n$ będą jego ideałami. Jeśli $I_i+I_j=R$ dla wszystkich i,j, gdzie $1\leqslant i\neq j\leqslant n$, to odwzorowanie*

$$\phi:R\ni r\mapsto(r+I_1,\ldots,r+I_n)\in R/I_1\oplus\cdots\oplus R/I_n$$

jest epimorfizmem pierścieni z jądrem $\operatorname{Ker}\phi=I_1\cap\cdots\cap I_n$.

Dowód. Przekonajmy się, że dla dowolnych $r_1,\ldots,r_n\in R$ znajdzie się taki element $r\in R$, że

$$r_i+I_i=r+I_i$$

dla wszystkich $i=1,\ldots,n$. Stosujemy indukcję względem liczby n.

Jeśli $n=1$, to teza zachodzi. Jeśli zaś $n=2$, to z warunku $I_1+I_2=R$ wynika istnienie takich elementów $i_1\in I_1$ oraz $i_2\in I_2$, że $1=i_1+i_2$. Biorąc $x=r_1i_2+r_2i_1$, otrzymujemy

$$\begin{aligned}x-r_1&=r_1i_2+r_2i_1-r_1=r_1i_2+r_2i_1-r_1(i_1+i_2)=(r_2-r_1)i_1\in I_1,\\x-r_2&=r_1i_2+r_2i_1-r_2=r_1i_2+r_2i_1-r_2(i_1+i_2)=(r_1-r_2)i_2\in I_2,\end{aligned}$$

a zatem teza zachodzi.

Załóżmy teraz, że dla $n-1$ twierdzenie ma miejsce, czyli znajdzie się taki element $y \in R$, że $r_i + I_i = y + I_i$ dla wszystkich $i = 1, \ldots, n-1$. Skoro $R = I_i + I_n$ dla $i = 1, \ldots, n-1$, to

$$R = R^{n-1} = \prod_{i=1}^{n-1}(I_i + I_n) \leqslant (\prod_{i=1}^{n-1} I_i) + \\ + I_n \leqslant (I_1 \bigcap \cdots \bigcap I_{n-1}) + I_n \subseteq R,$$

a zatem

$$R = (I_1 \bigcap \cdots \bigcap I_{n-1}) + I_n$$

oraz $1 = j_1 + j_2$ dla pewnych elementów $j_1 \in I_1 \cap \cdots \cap I_{n-1}$ oraz $j_2 \in I_n$. Biorąc

$$z = yj_2 + r_n j_1,$$

otrzymujemy

$$\begin{array}{rcll} z - y & = & y(j_2 - 1) + r_n j_1 = (y - r_n)j_1 & \in \quad I_1 \bigcap \cdots \bigcap I_{n-1}, \\ z - r_n & = & yj_2 + r_n(j_1 - 1) = (y - r_n)j_2 & \in \quad I_n. \end{array}$$

Skoro $z - y \in I_1 \cap \cdots \cap I_{n-1} \leqslant I_i$, to

$$z - r_i = (z - y) + (y - r_i) \in I_i$$

dla wszystkich $i = 1, \ldots, n-1$. Z udowodnionego wyżej wynika, że odwzorowanie ϕ jest suriekcją. Jeśli a, b są dowolnymi elementami w R, to

$$\phi(a + b) = (a + b + I_1, \ldots, a + b + I_n) = \\ = ((a + I_1) + (b + I_1), \ldots, (a + I_n) + (b + I_n)) = \\ = (a + I_1, \ldots, a + I_n) + \\ + (b + I_1, \ldots, b + I_n) = \phi(a) + \phi(b)$$

oraz

$$\phi(ab) = (ab + I_1, \ldots, ab + I_n) = \\ = ((a + I_1)(b + I_1), \ldots, (a + I_n)(b + I_n)) = \\ = (a + I_1, \ldots, a + I_n)(b + I_1, \ldots, b + I_n) = \phi(a)\phi(b).$$

Zatem ϕ jest epimorfizmem pierścieni. □

Wniosek 4.5.17. *Niech $n = n_1 \cdots n_s$ będzie iloczynem liczb naturalnych $n_1, \ldots, n_s$ większych niż 1. Jeśli liczby $n_1, \ldots, n_s$ są parami względnie pierwsze, to pierścienie $\mathbb{Z}_n$ oraz $\mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_s}$ są izomorficzne.*

Dowód. Jak wiadomo, $\mathbb{Z}_{n_i} = \mathbb{Z}/n_i\mathbb{Z}$ dla wszystkich $i = 1, \ldots, s$. Biorąc $I_i = n_i\mathbb{Z}$, na podstawie założenia wnosimy, że $\mathrm{NWD}(n_i, n_j) = 1$ dla wszystkich $1 \leqslant i \neq j \leqslant n$, a zatem otrzymujemy, że $I_i + I_j = \mathbb{Z}$. Zostało nam tylko zastosować twierdzenie 4.5.16 i dostaniemy tezę. □

Przykład 4.5.18.

Pierścień $\mathbb{Z}_{30} = \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_5$ jest sumą prostą pierścieni $\mathbb{Z}_2$, $\mathbb{Z}_3$ oraz $\mathbb{Z}_5$.

■ Niech $m \in \mathbb{N}^*$ oraz $a, b \in \mathbb{Z}$. *Rozwiązaniem kongruencji*

$$aX \equiv b \pmod{m} \tag{4.5}$$

jest nazywana klasa liczb postaci $\{x_0 + mt \mid t \in \mathbb{Z}\}$ z reprezentantem $x_0 \in \mathbb{Z}$ taka, że x_0 spełnia kongruencję (4.5).

Wniosek 4.5.19 (chińskie twierdzenie o resztach)**.** *Jeśli $m_1, m_2, \ldots, m_n$ są parami względnie pierwszymi liczbami naturalnymi, a $b_1, b_2, \ldots, b_n$ są dowolnymi liczbami całkowitymi, to układ kongruencji*

$$\begin{cases} x & \equiv b_1 \pmod{m_1}, \\ x & \equiv b_2 \pmod{m_2}, \\ & \vdots \\ x & \equiv b_n \pmod{m_n} \end{cases} \tag{4.6}$$

posiada rozwiązanie. Jeśli x_1, x_2 są dwoma rozwiązaniami tego układu, to

$$x_1 \equiv x_2 \pmod{m_1 m_2 \cdots m_n}.$$

Dowód. (Szkic) Weźmy

$$\begin{array}{rcl} M & = & m_1 m_2 \cdots m_n, \\ M_i & = & \frac{M}{m_i} \quad (i = 1, \ldots, n) \end{array}$$

oraz niech M_i^{-1} będzie elementem odwrotnym do elementu M_i w pierścieniu klas reszt $\mathbb{Z}_{m_i}$, czyli

$$M_i M_i^{-1} \equiv 1 \pmod{m_i}.$$

Wtedy rozwiązanie ogólne układu (4.6) ma postać

$$x_0 = a_1 M_1 M_1^{-1} + \cdots + a_n M_n M_n^{-1} \pmod{M}.$$

□

Przykład 4.5.20.

Rozwiążmy układ kongruencji

$$\begin{cases} X \equiv 19 \pmod{21}, \\ X \equiv 3 \pmod{8}, \\ X \equiv 5 \pmod{11}. \end{cases}$$

Skoro liczby 21, 8, 11 są parami względnie pierwsze, to układ posiada rozwiązanie. Obliczamy:

- $M = 21 \cdot 8 \cdot 11 = 1848$;
- $M_1 = \frac{M}{m_1} = 8 \cdot 11 = 88$;
- $M_2 = \frac{M}{m_2} = 21 \cdot 11 = 231$;
- $M_3 = \frac{M}{m_3} = 21 \cdot 8 = 168$.

Skoro

$$\begin{aligned} 88 &= 21 \cdot 4 + 4, \\ 21 &= 4 \cdot 5 + 1, \\ 4 &= 1 \cdot 4 + 0, \end{aligned}$$

to

$$1 = 21 - 4 \cdot 5 = 21 - (88 - 21 \cdot 4) \cdot 5 = 88 \cdot (-5) + 21 \cdot 21,$$

a więc $88(-5) \equiv 1 \pmod{21}$. Stąd $88^{-1} \equiv 16 \pmod{21}$ oraz

$$M_1^{-1} = 16.$$

Na podstawie $231 \equiv 7 \pmod{8}$ oraz $7^2 \equiv 1 \pmod{8}$ wnosimy, że $7^{-1} \equiv 7 \pmod{8}$ oraz

$$M_2^{-1} = 7.$$

Podobnie $168 \equiv 3 \pmod{11}$ oraz $3 \cdot 4 \equiv 1 \pmod{1}$, a więc

$$M_3^{-1} = 4.$$

Zatem

$$x_0 = 19 \cdot 88 \cdot 16 + 2 \cdot 231 \cdot 7 + 5 \cdot 168 \cdot 4 \pmod{1848} = 82.$$

Ćwiczenia 4.5.21.

(**1**) Udowodnić, że $3\mathbb{Z}[i]$ oraz $(3+2i)\mathbb{Z}[i]$ są ideałami maksymalnymi w pierścieniu $\mathbb{Z}[i]$.
(**2**) Niech $\mathbb{Z}_{(p)} = \{\frac{m}{n} \mid m \in \mathbb{Z},\ n \in \mathbb{N} \text{ oraz } \mathrm{NWD}(n,p) = 1\}$, gdzie p jest liczbą pierwszą. Udowodnić, że:
(a) $\mathbb{Z}_{(p)}$ jest pierścieniem;
(b) $\mathbb{Z}_{(p)}$ ma dokładnie jeden ideał maksymalny $p\mathbb{Z}_{(p)}$.
(**3**) Udowodnić, że:
(a) $I_c = \{f \in C_{[0,1]} \mid f(c) = 0\}$ jest ideałem maksymalnym w pierścieniu $C_{[0,1]}$ dla każdego $c \in [0,1]$;
(b) jeśli M jest ideałem maksymalnym w $C_{[0,1]}$, to $M = I_c$ dla pewnego c $(0 \leqslant c \leqslant 1)$.
(**4**) Niech A będzie pierścieniem przemiennym. Udowodnić, że jeśli:
(a) każdy ideał właściwy pierścienia A jest pierwszy, to A jest ciałem;
(b) P jest ideałem pierwszym w A, to dla dowolnych ideałów I oraz J pierścienia A z warunku $I \cap J \leqslant P$ wynika, że $I \leqslant P$ lub $J \leqslant P$;
(c) X jest podzbiorem multiplikatywnym w A oraz $X \cap I = \varnothing$ dla ideału $I \leqslant A$, to znajdzie się taki ideał pierwszy $P \leqslant A$, że $I \leqslant P$ oraz $P \cap X = \varnothing$;
(d) I, J są ideałami pierwszymi w pierścieniu A oraz $I + J = A$, to $I^m + J^n = A$ dla dowolnych $m, n \in \mathbb{N}^*$;
(e) I, J są ideałami pierwszymi w pierścieniu A oraz $I + J = A$, to $IJ = I \cap J$.
(**5**) Niech $f : A \to S$ będzie homomorfizmem pierścieni, I będzie ideałem w A oraz K będzie ideałem w S. Udowodnić, że:
(a) jeśli ideał K jest pierwszy w S, to $f^{-1}(K)$ jest pierwszy w A;
(b) jeśli f jest epimorfizmem oraz K jest ideałem maksymalnym w S, to ideał $f^{-1}(K)$ jest maksymalny w pierścieniu A.
(**6**) Udowodnić, że $I = \langle n, X\rangle \leqslant \mathbb{Z}[X]$, gdzie $n \in \mathbb{Z}$, jest maksymalny w pierścieniu $\mathbb{Z}[X]$ wtedy i tylko wtedy, gdy n jest liczbą pierwszą.
(**7**) Niech A będzie pierścieniem przemiennym z ideałem maksymalnym M. Udowodnić, że $P = M[X]$ jest ideałem pierwszym w pierścieniu wielomianów $A[X]$.
(**8**) Niech A będzie pierścieniem przemiennym. Udowodnić, że jeśli P jest ideałem pierwszym w A, to $P + XA[[X]]$ oraz $P[[X]]$ są ideałami pierwszymi w pierścieniu formalnych szeregów potęgowych $A[[X]]$.
(**9**) Udowodnić, że w pierścieniu $\mathbb{Z}[\sqrt{-5}]$ są nierozkładalne takie elementy:
(a) 2;
(b) 3;
(c) $1 - \sqrt{-5}$;
(d) $1 + \sqrt{-5}$.
(**10**) Sprawdzić, czy a jest elementem pierwszym w pierścieniu R, jeśli:
(a) $a = 40 + \sqrt{7}$ oraz $R = \mathbb{Z}[\sqrt{7}]$;
(b) $a = 2$ oraz $R = \mathbb{Z}[i\sqrt{3}]$;
(c) $a = 1 + i\sqrt{3}$ oraz $R = \mathbb{Z}[i\sqrt{3}]$;
(d) $a = 3$ oraz $R = \mathbb{Z}[i\sqrt{2}]$.

Uwagi. Twierdzenie o istnieniu ideału maksymalnego w pierścieniu z jednością udowodnił W. Krull[(5)] w 1929 r. L. Kronecker[(6)] pokazał, że ciało liczb algebraicznych (stopnia skończonego) jest izomorficzne z ciałem reszt $\mathbb{Q}[X]/\langle p\rangle$, gdzie $p \in \mathbb{Q}[X]$ jest wielomianem nieprzywie-

[(5)] Wolfgang Krull (1899–1971)
[(6)] Leopold Kronecker (1823–1891)

dlnym nad ciałem liczb wymiernych $\mathbb{Q}$. W postaci oryginalnej chińskie twierdzenie o resztach należy się matematykowi chińskiemu Sun Tzu[7]. Później to twierdzenie zostało odkryte na nowo przez Qin Jinshao[8].

[7] Sun Tzu (między 3 a 5 w. p.n.e.)
[8] Qin Jinshao (1202–1261)

4.6. Przykłady pierścieni i ciał

■ **Ciało formalnych szeregów potęgowych Laurenta**[9]**.** Niech $\mathbb{F}$ będzie ciałem. Rozpatrzmy zbiór $\mathbb{F}((X))$, który składa się z formalnych szeregów postaci

$$\sum_{i=-\infty}^{\infty} a_i X^i, \text{ gdzie } a_i \in \mathbb{F},$$

przy czym prawie wszystkie współczynniki a_i ze wskaźnikami ujemnymi $i < 0$ (czyli z wyjątkiem skończonej liczby) są zerowe. Jeśli $a_i, b_i \in \mathbb{F}$, to przyjmujemy, że

$$\begin{aligned} \sum_{i=-\infty}^{\infty} a_i X^i + \sum_{i=-\infty}^{\infty} b_i X^i &= \sum_{i=-\infty}^{\infty} (a_i + b_i) X^i, \\ (\sum_{i=-\infty}^{\infty} a_i X^i)(\sum_{j=-\infty}^{\infty} b_j X^j) &= \sum_{k=-\infty}^{\infty} c_k X^k, \end{aligned}$$

gdzie

$$c_j = \sum a_i b_{j-i},$$

i sumujemy po wszystkich wskaźnikach i, dla których $a_i b_{j-i} \neq 0$. Najmniejsza liczba całkowita i, dla której $a_i \neq 0$, jest nazywana *rzędem* szeregu

$$f = \sum_{i=-\infty}^{\infty} a_i X^i.$$

Jeśli szereg

$$\sum_{i=-\infty}^{\infty} a_i X^i \in \mathbb{F}((X))$$

ma rząd k, to szereg

$$X^{-k}(\sum_{i=-\infty}^{\infty} a_i X^i)$$

jest odwracalny (przekonać się samodzielnie).

Stąd otrzymujemy takie

[9] Pierre Alphonse Laurent (1813–1854)

Twierdzenie 4.6.1. *Jeśli* $\mathbb{F}$ *jest ciałem, to zbiór* $\mathbb{F}((X))$ *tworzy ciało* (które jest nazywane *ciałem formalnych szeregów Laurenta*).

* * *

■ **Ciała skończone.** Zachodzi takie

Twierdzenie 4.6.2. *Każda dziedzina całkowitości* (*z jednością* 1)*, składająca się ze skończonej liczby elementów, jest ciałem.*

Dowód. Niech A będzie skończoną dziedziną całkowitości oraz a będzie dowolnym jej elementem niezerowym. Jeśli b i c są różnymi elementami z A, to ab i ac też są różnymi w A. To implikuje, że zbiory A i aA są równoliczne, a zatem $A = aA$. Podobnym sposobem otrzymujemy, że $A = Aa$. Skoro $1 \in A$, to

$$1 = au \text{ oraz } 1 = va$$

dla pewnych elementów $u, v \in A$. Ponadto

$$u = (va)u = v(au) = v,$$

czyli a jest elementem odwracalnym w pierścieniu A. Zatem A jest ciałem. □

Kolejne twierdzenie (przytaczamy bez dowodu), udowodnione przez Wedderburna[(10)] w 1905 r., zezwala na wyciągnięcie wniosku, że każda skończona dziedzina całkowitości jest ciałem przemiennym.

Twierdzenie 4.6.3 (Wedderburna)**.** *Każde ciało skończone jest przemienne.*

□

* * *

■ **Skończony pierścień nieprzemienny.** Niech $\mathbb{F}_2 = \{0, 1\}$ będzie ciałem o dwóch elementach. Istnieją pierścienie skończone, które nie są przemienne. Ze względu na poprzednie dwa twierdzenia takie pierścienie

(10) Joseph Henry Maclagen Wedderburn (1882–1948)

posiadają dzielniki zera. Rzeczywiście pierścień $M_2(\mathbb{F}_2)$, który składa się z 16 elementów:

$$0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},\ I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$A_5 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},\ A_6 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$A_7 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix},\ A_8 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},$$

$$A_9 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},\ A_{10} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

$$A_{11} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},\ A_{12} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

$$A_{13} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},\ A_{14} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

nie jest przemienny, bo, na przykład

$$A_1 A_2 = A_2 \neq 0 = A_2 A_1.$$

Obliczając wyznaczniki macierzy z tego pierścienia, wnosimy, że

$$GL_2(\mathbb{F}_2) = \{I, A_6, A_{10}, A_{11}, A_{12}, A_{13}\} = SL_2(\mathbb{F}_2)$$

jest grupą nieabelową rzędu 6, gdyż

$$A_6 A_{10} = A_{12} \neq A_{13} = A_{10} A_6.$$

Łatwo przekonać się, że charakterystyka $\operatorname{char} M_2(\mathbb{F}_2) = 2$.

* * *

■ **Pierścienie boolowskie.** Pierścień A jest nazywany *boolowskim*[(11)], jeśli wszystkie jego elementy są idempotentami, czyli $a^2 = a$ dla wszystkich $a \in A$. Każdy pierścień boolowski A jest przemienny oraz $2a = 0$ dla każdego elementu $a \in A$. Istotnie dla dowolnych elementów $a, b \in A$ mamy

$$a + b = (a + b)^2 = a + ab + ba + b,$$

a stąd $ab + ba = 0$. W szczególności, jeśli $b = a$, to $2a = 2a^2 = 0$. Wtedy też

$$ab + ab = 0 = ab + ba$$

i na podstawie prawa skracania $ab = ba$.

■ Pierścień $\mathbb{Z}_2$ jest boolowski. Dodatkowo

- każdy pierścień boolowski, składający się z dwóch elementów, jest izomorficzny z $\mathbb{Z}_2$;
- każdy skończony pierścień boolowski jest izomorficzny z pewną skończoną sumą prostą (patrz lemat 4.5.14) postaci

$$\mathbb{Z}_2 \oplus \cdots \oplus \mathbb{Z}_2.$$

* * *

■ **Ciała kwadratowe.** Niech d będzie taką liczbą naturalną, że $\sqrt{d} \notin \mathbb{Q}$. Rozpatrzmy podzbiór

$$\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$$

ciała liczb rzeczywistych $\mathbb{R}$. Jeśli $a_i, b_i \in \mathbb{Q}$ $(i = 1, 2)$, to

$$\begin{aligned}
&(a_1 + b_1\sqrt{d}) - (a_2 + b_2\sqrt{d}) = \\
&= (a_1 - a_2) + (b_1 - b_2)\sqrt{d} \in \mathbb{Q}[\sqrt{d}], \\
&(a_1 + b_1\sqrt{d})(a_2 + b_2\sqrt{d}) = \\
&= (a_1a_2 + db_1b_2) + (a_1b_2 + b_1a_2)\sqrt{d} \in \mathbb{Q}[\sqrt{d}].
\end{aligned}$$

Niech $a, b, c, e \in \mathbb{Q}$. Załóżmy, że $a + b\sqrt{d} = 0$. Jeśli $b \neq 0$, to

$$\sqrt{d} = -\frac{a}{b} \in \mathbb{Q},$$

(11) George Boole (1815–1864)

a to jest sprzeczne z założeniem. Zatem $b = 0$, a więc i $a = 0$. Udowodniliśmy, że

$$a + b\sqrt{d} = 0 \Leftrightarrow \begin{cases} a = 0, \\ b = 0. \end{cases}$$

Wnosimy, że

$$a + b\sqrt{d} = c + e\sqrt{d} \Leftrightarrow \begin{cases} a = c, \\ b = e. \end{cases}$$

Zaznaczmy, że $1 = 1 + 0\sqrt{d} \in \mathbb{Q}[\sqrt{d}]$.

Dalej załóżmy, że $a + b\sqrt{d} \neq 0$. Jeśli $a^2 - db^2 = 0$ oraz $b = 0$, to $a = 0$, a więc $a + b\sqrt{d} = 0$, a to jest sprzeczne z założeniem. Jeśli zaś $a^2 - db^2 = 0$ oraz $b \neq 0$, to

$$d = \frac{a^2}{b^2},$$

co też nie jest możliwe. Zatem $a^2 - db^2 \neq 0$.

Znajdźmy takie liczby wymierne x, y, że

$$(a + b\sqrt{d})(x + y\sqrt{d}) = 1,$$

co możemy przepisać równoważnie w postaci układu równań liniowych

$$\begin{cases} ax + dby &= 1, \\ bx + \ ay &= 0. \end{cases}$$

Ponieważ wyznacznik tego układu $\Delta = a^2 - db^2 \neq 0$ jest niezerowy, to mamy układ Cramera, a zatem

$$x = \frac{a}{a^2 - db^2}, \ y = \frac{-b}{a^2 - db^2}.$$

Z tego, że $a + b\sqrt{d} \neq 0$ wnosimy, że

$$(a + b\sqrt{d})^{-1} = \frac{a}{a^2 - db^2} - \frac{b}{a^2 - db^2}\sqrt{d} \in \mathbb{Q}[\sqrt{d}].$$

Udowodniliśmy, że $\mathbb{Q}[\sqrt{d}]$ jest podciałem ciała $\mathbb{R}$. Zatem $\mathbb{Q}[\sqrt{d}]$ jest ciałem (które jest nazywane *kwadratowym*).

Ćwiczenia 4.6.4.

(**1**) Skonstruować tabelki Cayleya dla dodawania i mnożenia w pierścieniu:
(*a*) $\mathbb{Z}_2[X]/\langle X^3+X+1\rangle$;
(*b*) $\mathbb{Z}_3[X]/\langle X^3+1\rangle$.

(**2**) Udowodnić, że $\mathbb{Z}_3[X]/\langle X^2+1\rangle$ jest ciałem o 9 elementach. Zbudować tabelki Cayleya dla dodawania i mnożenia.

(**3**) Niech $n \in \mathbb{N}^*$, $I_1 = X^n\mathbb{R}[X]$ oraz $I_2 = X^n\mathbb{R}[[X]]$. Udowodnić, że

$$\mathbb{R}[X]/I_1 \cong \mathbb{R}[[X]]/I_2.$$

(**4**) Znaleźć przykład niezerowego homomorfizmu pierścieni postaci $\psi : \mathbb{R}[X] \to \mathbb{R}[X]$, który nie jest monomorfizmem.

(**5**) Niech R będzie pierścieniem z jednością 1, I będzie jego ideałem oraz

$$I[[X]] = \{\sum_{i=0}^{\infty} a_i X^i \mid a_i \in I \text{ dla każdego } i \in \mathbb{N}\}.$$

Udowodnić, że $I[[X]]$ jest ideałem pierścienia $R[[X]]$. Sprawdzić, czy $I[[X]] = I \cdot R[[X]]$.

Uwagi. Pierwsze kroki w badaniu rozszerzeń kwadratowych ciała $\mathbb{Q}$ zostały poczynione przez C. Gaussa i M. Eisensteina[(12)].

[(12)] Ferdinand Gotthold Max Eisenstein (1823–1852)

4.7. Ciało kwaternionów

■ Poszukiwania uogólnień liczb zespolonych doprowadziły do wynalezienia kwaternionów.

Twierdzenie 4.7.1. *Zbiór macierzy*

$$\mathcal{H} = \{\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \mid z, w \in \mathbb{C}\}$$

tworzy ciało (względem dodawania i mnożenia macierzy).

Dowód. Udowodnimy na początek, że $\mathcal{H}$ jest podpierścieniem w $M_2(\mathbb{C})$. Rzeczywiście jest macierz jednostkowa

$$I_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -\overline{0} & \overline{0} \end{bmatrix} \in \mathcal{H}.$$

Niech $z, w, u, v \in \mathbb{C}$. Jeśli

$$A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \in \mathcal{H},\ B = \begin{bmatrix} u & v \\ -\overline{v} & \overline{u} \end{bmatrix} \in \mathcal{H},$$

to

$$A - B = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} - \begin{bmatrix} u & v \\ -\overline{v} & \overline{u} \end{bmatrix} = \begin{bmatrix} z-u & w-v \\ -\overline{(w-v)} & \overline{z-u} \end{bmatrix} \in \mathcal{H}$$

oraz

$$AB = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \begin{bmatrix} u & v \\ -\overline{v} & \overline{u} \end{bmatrix} = \begin{bmatrix} zu - w\overline{v} & zv + w\overline{u} \\ -\overline{(zv + w\overline{u})} & \overline{zu - w\overline{v}} \end{bmatrix} \in \mathcal{H}.$$

Na podstawie kryterium podpierścienia wnosimy, że $\mathcal{H}$ jest pierścieniem z jednością I_2. Zostało nam udowodnienie, że każda macierz niezerowa postaci

$$A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$$

ma macierz odwrotną w pierścieniu $\mathcal{H}$. Istotnie warunek $A \neq 0$ implikuje, że $z \neq 0$ lub $w \neq 0$, a więc

$$\det A = z\overline{z} + w\overline{w} = |z|^2 + |w|^2 > 0.$$

Zatem macierz odwrotna A^{-1} istnieje oraz nietrudno przekonać się, że

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix} \in \mathcal{H}.$$

Wnosimy, że $\mathcal{H}$ jest pierścieniem z dzieleniem. □

■ Niech $a, b, c, d \in \mathbb{R}$. Skoro każdą macierz

$$A = \begin{bmatrix} a+ib & c+id \\ -(c+id) & \overline{a+ib} \end{bmatrix} \in \mathcal{H}$$

możemy przepisać w postaci

$$A = \begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix} =$$

$$= a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

oraz

$$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}^2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}^2 = -I_2,$$

to utożsamiając

$$1 = I_2,\ \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},\ \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},\ \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix},$$

przychodzimy do przedstawienia kwaternionu w postaci algebraicznej. Mianowicie element

$$a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d,$$

gdzie a, b, c, d są liczbami rzeczywistymi oraz

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1,\ \mathbf{ij} = \mathbf{k} = -\mathbf{ji},\ \mathbf{ki} = \mathbf{j} = -\mathbf{ik},\ \mathbf{jk} = \mathbf{i} = -\mathbf{kj},$$

jest nazywany *kwaternionem w postaci algebraicznej*. Twierdzenie 4.7.1 możemy teraz przeformułować w takiej postaci.

Twierdzenie 4.7.2. *Zbiór*

$$\mathbb{H} = \{a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d \mid a, b, c, d \in \mathbb{R}, \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \\ \mathbf{ij} = \mathbf{k} = -\mathbf{ji}, \mathbf{ki} = \mathbf{j} = -\mathbf{ik}, \mathbf{jk} = \mathbf{i} = -\mathbf{kj}\}$$

tworzy ciało względem dodawania i mnożenia, określonych przez takie reguły:

- (dodawanie „+")

$$(a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d) + (a_1 + \mathbf{i}b_1 + \mathbf{j}c_1 + \mathbf{k}d_1) = \\ = (a + a_1) + \mathbf{i}(b + b_1) + \mathbf{j}(c + c_1) + \mathbf{k}(d + d_1);$$

- (mnożenie „·")

$$(a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d)(a_1 + \mathbf{i}b_1 + \mathbf{j}c_1 + \mathbf{k}d_1) = \\ = (aa_1 - bb_1 - cc_1 - dd_1) + \mathbf{i}(cd_1 - dc_1 + ab_1 + ba_1) + \\ + \mathbf{j}(db_1 - bd_1 + ac_1 + ca_1) + \mathbf{k}(bc_1 - cb_1 + ad_1 + da_1),$$

gdzie $a, a_1, b, b_1, c, c_1, d, d_1$ *są liczbami rzeczywistymi* (wszędzie dalej $\mathbb{H}$ będziemy nazywać *ciałem kwaternionów*).

■ Mnożenie nie jest działaniem przemiennym w $\mathbb{H}$.

■ Łatwo sprawdzić, że ciała $\mathcal{H}$ oraz $\mathbb{H}$ są izomorficzne (udowodnić samodzielnie).

* * *

■ **Sprzężenie.** Kwaternion

$$a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d \in \mathbb{H},$$

gdzie $a, b, c, d \in \mathbb{R}$, jest nazywany *sprzężonym* z kwaternionem $\alpha = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d \in \mathbb{H}$ (i oznaczany przez $\overline{\alpha}$), a liczba rzeczywista

$$N(\alpha) = a^2 + b^2 + c^2 + d^2$$

jest nazywana *normą* kwaternionu α. Nietrudno przekonać się (sprawdzić samodzielnie), że dla dowolnych kwaternionów $\alpha, \beta \in \mathbb{H}$ zachodzą związki:

(1) $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$;

(2) $\overline{\alpha}\ \overline{\beta} = \overline{\beta}\ \overline{\alpha}$;
(3) $N(\alpha) = \alpha\overline{\alpha}$;
(4) $N(\alpha\beta) = N(\alpha)N(\beta)$.

* * *

■ **Części urojona i rzeczywista kwaternionu.** Jeśli mamy kwaternion $\alpha = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d$, gdzie a, b, c, d są liczbami rzeczywistymi, to:

- nieujemna liczba rzeczywista

$$||\alpha|| = \sqrt{a^2 + b^2 + c^2 + d^2}$$

  jest nazywana *modułem* kwaternionu α;
- liczba rzeczywista $\operatorname{Re}\alpha = a$ jest nazywana *częścią rzeczywistą* kwaternionu α;
- element $\operatorname{Im}\alpha = \mathbf{i}b+\mathbf{j}c+\mathbf{k}d$ jest nazywany *częścią urojoną* kwaternionu α.

Ćwiczenia 4.7.3.

Udowodnić, że dla kwaternionów $\alpha, \beta \in \mathbb{H}$ zachodzą własności:
(**1**) $||\alpha\beta|| = ||\alpha|| \cdot ||\beta||$;
(**2**) $\operatorname{Re}\alpha = \frac{1}{2}(\alpha + \overline{\alpha})$;
(**3**) $\operatorname{Im}\alpha = \frac{1}{2}(\alpha - \overline{\alpha})$;
(**4**) $||\alpha + \beta|| \leqslant ||\alpha|| + ||\beta||$;
(**5**) $|\operatorname{Re}\alpha| \leqslant ||\alpha||$, $||\operatorname{Im}\alpha|| \leqslant ||\alpha||$;
(**6**) $\alpha = 0$ wtedy i tylko wtedy, gdy $||\alpha|| = 0$;
(**7**) $\overline{\operatorname{Im}\alpha} = -\operatorname{Im}\alpha$.

* * *

■ **Postać trygonometryczna kwaternionu.** *Osią główną* kwaternionu

$$\alpha = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d,$$

gdzie a, b, c, d są liczbami rzeczywistymi, jest nazywany kwaternion postaci

$$i_\alpha = \begin{cases} \mathbf{i}, & \text{gdy } \operatorname{Im}\alpha = 0, \\ \frac{\mathbf{i}b+\mathbf{j}c+\mathbf{k}d}{\sqrt{b^2+c^2+d^2}}, & \text{gdy } \operatorname{Im}\alpha \neq 0. \end{cases}$$

Twierdzenie 4.7.4. *Każdy kwaternion niezerowy $\alpha \in \mathbb{H}$ ma taką postać trygonometryczną*

$$\alpha = ||\alpha||(\cos\phi + i_\alpha \sin\phi),$$

(gdzie liczba $\phi \in \mathbb{R}$ jest nazywana *argumentem* kwaternionu α i oznaczana przez $\arg \alpha$).

Dowód. Niech $\alpha = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d$, gdzie a, b, c, d są liczbami rzeczywistymi, oraz ϕ będzie rozwiązaniem układu

$$\begin{cases} 0 \leqslant \phi \leqslant \pi, \\ \cos \phi = \frac{a}{||\alpha||}. \end{cases}$$

Ponieważ wartość bezwzględna

$$\left| \frac{a}{||\alpha||} \right| \leqslant 1,$$

to istnieje dokładnie jedno takie rozwiązanie ϕ. Wtedy $\sin \phi \geqslant 0$ oraz

$$\sin \phi = \sqrt{1 - \cos^2 \phi} = \sqrt{1 - \frac{a^2}{||\alpha||^2}} = \frac{\sqrt{b^2 + c^2 + d^2}}{||\alpha||} = \frac{||\operatorname{Im} \alpha||}{||\alpha||},$$

skąd $i_\alpha ||\alpha|| \sin \phi = \operatorname{Im} \alpha$. Zatem

$$\alpha = ||\alpha||(\cos \phi + i_\alpha \sin \phi).$$

□

* * *

■ **Analog wzoru de Moivre'a.** Proponujemy Czytelnikowi udowodnić takie twierdzenie (stosując rozumowania indukcyjne względem n).

Twierdzenie 4.7.5. *Dla każdego kwaternionu $\alpha \in \mathbb{H}$ i każdej liczby niezerowej naturalnej n zachodzi wzór*

$$(||\alpha||(\cos \phi + i_\alpha \sin \phi))^n = ||\alpha||^n (\cos n\phi + i_\alpha \sin n\phi),$$

gdzie $\phi = \arg \alpha$ jest argumentem kwaternionu α.

□

* * *

■ **Pierwiastkowanie kwaternionów.** Niech $n \in \mathbb{N}^*$. Kwaternion γ jest nazywany *pierwiastkiem stopnia n-tego* z α, jeśli

$$\gamma^n = \alpha.$$

Jak wiadomo, można znaleźć wszystkie liczby zespolone, które są pierwiastkami n-go stopnia z danej liczby zespolonej; przy czym zbiór takich pierwiastków jest skończony. Dla kwaternionu ta sytuacja jest inna: *Jeśli*

$$\alpha = ||\alpha||(\cos\phi + i_\alpha \sin\phi)$$

jest kwaternionem niezerowym oraz $\phi = \arg\alpha$*, to* `pewne` *jego pierwiastki* α_k *stopnia* n *możemy znaleźć za pomocą wzoru:*

$$\alpha_k = \sqrt[n]{||\alpha||}(\cos\frac{\phi + 2\pi k}{n} + i_\alpha \sin\frac{\phi + 2\pi k}{n}),$$

gdzie $k = 0, 1, \ldots, n-1$.

■ Pierwiastki α_k $(k = 0, 1, \ldots, n-1)$ niekoniecznie są wszystkimi pierwiastkami stopnia n z kwaternionu α (patrz ćwiczenia 4.7.7(1)).

* * *

■ **Grupa kwaternionów** Q_8**.** Niech

$$Q_8 = \{1, -1, \mathbf{i}, -\mathbf{i}, \mathbf{j}, -\mathbf{j}, \mathbf{k}, -\mathbf{k}\}$$

będzie podzbiorem w ciele kwaternionów $\mathbb{H}$. Mnożenie elementów w zbiorze Q_8 zadaje taka tabelka Cayleya:

$\cdot$	1	-1	$\mathbf{i}$	$-\mathbf{i}$	$\mathbf{j}$	$-\mathbf{j}$	$\mathbf{k}$	$-\mathbf{k}$
1	1	-1	$\mathbf{i}$	$-\mathbf{i}$	$\mathbf{j}$	$-\mathbf{j}$	$\mathbf{k}$	$-\mathbf{k}$
-1	-1	1	$-\mathbf{i}$	$\mathbf{i}$	$-\mathbf{j}$	$\mathbf{j}$	$-\mathbf{k}$	$\mathbf{k}$
$\mathbf{i}$	$\mathbf{i}$	$-\mathbf{i}$	-1	1	$\mathbf{k}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{j}$
$-\mathbf{i}$	$-\mathbf{i}$	$\mathbf{i}$	1	-1	$-\mathbf{k}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{j}$
$\mathbf{j}$	$\mathbf{j}$	$-\mathbf{j}$	$-\mathbf{k}$	$\mathbf{k}$	-1	1	$\mathbf{i}$	$-\mathbf{i}$
$-\mathbf{j}$	$-\mathbf{j}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{k}$	1	-1	$-\mathbf{i}$	$\mathbf{i}$
$\mathbf{k}$	$\mathbf{k}$	$-\mathbf{k}$	$\mathbf{j}$	$-\mathbf{j}$	$-\mathbf{i}$	$\mathbf{i}$	-1	1
$-\mathbf{k}$	$-\mathbf{k}$	$\mathbf{k}$	$-\mathbf{j}$	$\mathbf{j}$	$\mathbf{i}$	$-\mathbf{i}$	1	-1

Jak wynika z obliczeń w tabelce, zbiór Q_8 jest domknięty względem mnożenia i każdy element ma odwrotny do siebie w Q_8. Skoro $Q_8 \subseteq \mathbb{H}$, to Q_8 jest grupą (która jest nazywana *grupą kwaternionów*).

■ Pierścień (łączny lub niełączny) R jest nazywany *algebrą nad ciałem* $\mathbb{F}$ (lub krótko $\mathbb{F}$-algebrą), jeśli R jest przestrzenią liniową nad ciałem $\mathbb{F}$ oraz

$$\lambda(rt) = (\lambda r)t = r(\lambda t)$$

dla dowolnych $\lambda \in \mathbb{F}$ oraz $r, t \in R$.

Na przykład pierścień macierzy kwadratowych $M_n(\mathbb{C})$ stopnia n oraz pierścień wielomianów $\mathbb{C}[X]$ są algebrami nad ciałem liczb zespolonych $\mathbb{C}$, lecz ciało kwaternionów $\mathbb{H}$ nie jest algebrą nad ciałem $\mathbb{C}$ (ciało kwaternionów $\mathbb{H}$ jest algebrą nad ciałem liczb rzeczywistych $\mathbb{R}$).

Jeśli ciało jest algebrą, to jest nazywane *algebrą z dzieleniem.*

■ W 1878 r. G. Frobenius opublikował dowód następnego twierdzenia.

Twierdzenie 4.7.6. *Nad ciałem liczb rzeczywistych* $\mathbb{R}$ *istnieją* (*z dokładnością do izomorfizmu*) *trzy skończenie wymiarowe łączne algebry z dzieleniem:* $\mathbb{R}$, $\mathbb{C}$ *oraz* $\mathbb{H}$.

Ćwiczenia 4.7.7.

(**1**) Udowodnić, że równanie $x^2 = -1$ ma nieskończenie wiele rozwiązań w ciele kwaternionów $\mathbb{H}$.
(**2**) Niech $A, B \in SL_2(\mathbb{Z}_3)$, gdzie

$$A = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \quad \text{oraz} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Udowodnić, że grupa $\langle A, B \rangle$ jest izomorficzna z grupą kwaternionów Q_8.

Uwagi. Po interpretacji geometrycznej liczb zespolonych wynalezionej przez C. Gaussa, jako punktów na płaszczyźnie zespolonej, wynikła idea uogólnienia (czy rozszerzenia) liczb zespolonych tak, aby „nowe" liczby interpretować jako punkty w przestrzeni trójwymiarowej. Jako pierwszy ten pomysł wdrożył K. Wessel[13] w 1799 r., a później W. Hamilton w latach 1837–38. A. de Morgan[14] rozważał liczby postaci $a\xi + b\eta + c\zeta$,

[13] Caspar Wessel (1745–1818)
[14] Augustus de Morgan (1806–1871)

gdzie $a, b, c \in \mathbb{R}$. Częściowym przypadkiem tej konstrukcji jest algebra zbudowana przez C. Gravesa[15] w 1847 r. Ta algebra składa się z elementów postaci $a + be + ce^2$, gdzie $a, b, c \in \mathbb{R}$ oraz $e^2 = 1$, lecz w niej są zawarte dzielniki zera, które znalazł C. Graves. W. Hamilton w jednej swojej pracy z 1850 r. odkrył algebrę kwaternionów. Niezależnie od niego kwaterniony wynalazł też G. Grassmann[16]. A. Cayley przedstawił kwaterniony w postaci macierzowej (jak to zrobiono powyżej).

[15] Charles Graves (1812–1899)
[16] Herman Günter Grassmann (1809–1877)

4.8. Algebra oktonionów

Oktoniony coraz częściej są stosowane w opisach badań geometrycznych współczesnego świata. Jak wiadomo, w ciele liczb zespolonych $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ działania dodawania „+" i mnożenia „·" są określone według wzorów:

$$\begin{array}{rcl}(a,b)+(c,d) & = & (a+b,c+d),\\ (a,b)\cdot(c,d) & = & (ac-bd,ad+bc).\end{array}$$

Jeśli na zbiorze $\mathbb{C}\times\mathbb{C}$ w podobny sposób zdefiniować dwa działania

$$(z,t)+(z',t')=(z+z',t+t')$$

oraz

$$(z,t)\cdot(z',t')=(zz'-\overline{t'}t,t'z+t\overline{z'}),$$

to otrzymamy ciało $(\mathbb{C}\times\mathbb{C},+,\cdot)$, które jest izomorficzne z ciałem kwaternionów $\mathbb{H}$ (sprawdzić samodzielnie). Istnieje sposób, aby otrzymać „nowe" rozszerzenie ciała kwaternionów. Zaczniemy po kolei opowiadać, jak to zrealizować.

■ Rozpatrzmy zbiór

$$\mathbb{O}=\mathbb{H}\times\mathbb{H}$$

i określmy na nim takie działania:

- (dodawanie „+")

$$(h_1,h_2)+(h_1',h_2')=(h_1+h_1',h_2+h_2') \tag{4.7}$$

oraz

- (mnożenie „·")

$$(h_1,h_2)\cdot(h_1',h_2')=(h_1h_1'-\overline{h_2'}h_2,h_2'h_1+h_2\overline{h_1'}), \tag{4.8}$$

gdzie $h_1,h_1',h_2,h_2'\in\mathbb{H}$.

Taki element $(h_1,h_1')\in\mathbb{O}$ będziemy nazywać *oktonionem*.

Na podstawie (4.7) oraz (4.8) otrzymujemy

Twierdzenie 4.8.1. *Zachodzą następujące własności*[17]*:*

(1) $(\mathbb{O}, +)$ *jest grupą abelową;*

(2) $\mathbb{O}$ *jest przestrzenią liniową nad ciałem liczb rzeczywistych* $\mathbb{R}$*;*

(3) *mnożenie oktonionów „*$\cdot$*” jest rozdzielne względem ich dodawania „*$+$*”, czyli*

$$\forall_{\alpha,\beta,\gamma\in\mathbb{O}} : (\alpha+\beta)\cdot\gamma = (\alpha\cdot\gamma)+(\beta\cdot\gamma) \text{ oraz } \alpha\cdot(\beta+\gamma) = (\alpha\cdot\beta)+(\alpha\cdot\gamma);$$

(4) *mnożenie posiada element neutralny* $1 = (1, 0)$*, czyli*

$$\exists_{1\in\mathbb{O}}\forall_{\alpha\in\mathbb{O}} : \alpha\cdot 1 = \alpha = 1\cdot\alpha;$$

(5) *mnożenie oktonionów nie jest działaniem łącznym, ale zachodzą następujące własności:*

$$\forall_{\alpha,\beta\in\mathbb{O}} : \alpha(\beta\beta) = (\alpha\beta)\beta \text{ oraz } (\alpha\alpha)\beta = \alpha(\alpha\beta);$$

(6) *jeśli* (p, q) *jest elementem niezerowym z* $\mathbb{O}$*, to istnieje element odwrotny* $(p, q)^{-1}$ *do niego taki, że*

$$(p, q)^{-1} = \left(\frac{\overline{p}}{||p||^2 + ||q||^2}, \frac{-q}{||p||^2 + ||q||^2}\right).$$

Dowód. Niech dalej $(h_1, h_2), (g_1, g_2), (f_1, f_2) \in \mathbb{O}$.

(1) Ćwiczenie.

(2) Ćwiczenie.

(3) Obliczamy, że

$$\begin{gathered}((f_1, f_2) + (g_1, g_2))\cdot(h_1, h_2) = (f_1 + g_1, f_2 + g_2)\cdot(h_1, h_2) = \\ = ((f_1 + g_1)h_1 - \overline{h_2}(f_2 + g_2), h_2(f_1 + g_1) + (f_2 + g_2)\overline{h_1}) = \\ = (f_1h_1 + g_1h_1 - \overline{h_2}f_2 - \overline{h_2}g_2, h_2f_1 + h_2g_1 + f_2\overline{h_1} + g_2\overline{h_1}),\end{gathered}$$

$$\begin{gathered}(f_1, f_2)(h_1, h_2) + (g_1, g_2)(h_1, h_2) = (f_1h_1 - \overline{h_2}f_2, h_2f_1 + f_2\overline{h_1}) + \\ +(g_1h_1 - \overline{h_2}g_2, h_2g_1 + g_2\overline{h_1}) = \\ = (f_1h_1 - \overline{h_2}f_2 + g_1h_1 - \overline{h_2}g_2, h_2f_1 + f_2\overline{h_1} + h_2g_1 + g_2\overline{h_1})\end{gathered}$$

[17] Własności oktonionów opisane w twierdzeniu 4.8.1 oznaczają, że $\mathbb{O}$ jest *algebrą alternatywną* nad ciałem liczb rzeczywistych $\mathbb{R}$.

oraz

$$\begin{gathered}(f_1,f_2)((g_1,g_2)+(h_1,h_2))=(f_1,f_2)(g_1+h_1,g_2+h_2)=\\ =(f_1(g_1+h_1)-\overline{(g_2+h_2)}f_2,(g_2+h_2)f_1+f_2\overline{(g_1+h_1)})=\\ =(f_1g_1+f_1h_1-\overline{g_2}f_2-\overline{h_2}f_2,g_2f_1+h_2f_1+f_2\overline{g_1}+f_2\overline{h_1})\end{gathered}$$

$$\begin{gathered}(f_1,f_2)(g_1,g_2)+(f_1,f_2)(h_1,h_2)=(f_1g_1-\overline{g_2}f_2,g_2f_1+f_2\overline{g_1})+\\ +(f_1h_1-\overline{h_2}f_2,h_2f_1+f_2\overline{h_1})=\\ =(f_1g_1-\overline{g_2}f_2+f_1h_1-\overline{h_2}f_2,g_2f_1+f_2\overline{g_1}+h_2f_1+f_2\overline{h_1})\end{gathered}$$

a zatem mnożenie jest rozdzielne względem dodawania.

(4) Ćwiczenie.

(5) Dalej przekonujemy się, że

$$\begin{gathered}((f_1,f_2)(g_1,g_2))(h_1,h_2)=(f_1g_1-\overline{g_2}f_2,g_2f_1+f_2\overline{g_1})(h_1,h_2)=\\ =((f_1g_1-\overline{g_2}f_2)h_1-\overline{h_2}(g_2f_1+f_2\overline{g_1}),\\ h_2(f_1g_1-\overline{g_2}f_2)+(g_2f_1+f_2\overline{g_1})\overline{h_1})=\\ =(f_1g_1h_1-\overline{g_2}f_2h_1-\overline{h_2}g_2f_1-\overline{h_2}f_2\overline{g_1},\\ h_2f_1g_1-h_2\overline{g_2}f_2+g_2f_1\overline{h_1}+f_2\overline{g_1}\overline{h_1})\end{gathered}$$

oraz

$$\begin{gathered}(f_1,f_2)((g_1,g_2)(h_1,h_2))=(f_1,f_2)(\overline{g_1h_1-\overline{h_2}g_2},h_2g_1+g_2\overline{h_1})=\\ =((f_1(g_1h_1-\overline{h_2}g_2)-\overline{(h_2g_1+g_2\overline{h_1})}f_2,\\ (h_2g_1+g_2\overline{h_1})f_1+f_2\overline{(g_1h_1-\overline{h_2}g_2)})=\\ =(f_1g_1h_1-f_1\overline{h_2}g_2-\overline{h_2}\overline{g_1}f_2-\overline{g_2}h_1\overline{f_2},\\ h_2g_1f_1+g_2\overline{h_1}f_1+f_2\overline{g_1}\overline{h_1}-f_2h_2\overline{g_2}),\end{gathered}$$

skąd po porównaniu stron tych równości wysuwamy hipotezę, że mnożenie nie jest łączne w $\mathbb{O}$ (zostawiamy Czytelnikowi zbudowanie odpowiedniego kontrprzykładu).

(6) Ćwiczenie.

□

■ Elementy $1,\mathbf{i},\mathbf{j},\mathbf{k},\mathbf{p},\mathbf{q},\mathbf{r},\mathbf{l}\in\mathbb{O}$ są nazywane oktonionami *bazowymi*. Podobnie jak i kwaternion nad ciałem liczb rzeczywistych $\mathbb{R}$, oktonion $\alpha\in\mathbb{O}$ w *postaci algebraicznej* jest zapisywany w taki sposób:

$$\alpha=a_0+a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}+a_4\mathbf{p}+a_5\mathbf{q}+a_6\mathbf{r}+a_7\mathbf{l}$$

gdzie $a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7 \in \mathbb{R}$ są jego *współczynnikami* (lub *komponentami*). Wtedy dodajemy oktoniony pokomponentowo (czyli sumujemy między sobą współczynniki obok jednakowych oktonionów bazowych), a przemnażamy oktoniony jako wyrazy, pamiętając przy tym, że mnożenie jest rozdzielne względem dodawania, lecz nie jest łączne, a mnożenie jednomianów a_ie_i oraz a_je_j, gdzie $a_i, a_j \in \mathbb{R}$, jest określone regułą

$$a_ie_i \cdot a_je_j = (a_ia_j)(e_i \cdot e_j),$$

gdzie iloczyn $e_i \cdot e_j$ oktonionów bazowych $e_i, e_j \in \{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{l}\}$ jest zadany taką tablicą:

$\cdot$	1	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$\mathbf{p}$	$\mathbf{q}$	$\mathbf{r}$	$\mathbf{l}$
1	1	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$\mathbf{p}$	$\mathbf{q}$	$\mathbf{r}$	$\mathbf{l}$
$\mathbf{i}$	$\mathbf{i}$	-1	$\mathbf{k}$	$-\mathbf{j}$	$\mathbf{q}$	$-\mathbf{p}$	$-\mathbf{l}$	$\mathbf{r}$
$\mathbf{j}$	$\mathbf{j}$	$-\mathbf{k}$	-1	$\mathbf{i}$	$\mathbf{r}$	$\mathbf{l}$	$-\mathbf{p}$	$-\mathbf{q}$
$\mathbf{k}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{i}$	-1	$\mathbf{l}$	$-\mathbf{r}$	$\mathbf{q}$	$-\mathbf{p}$
$\mathbf{p}$	$\mathbf{p}$	$-\mathbf{q}$	$-\mathbf{r}$	$-\mathbf{l}$	-1	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{q}$	$\mathbf{q}$	$\mathbf{p}$	$-\mathbf{l}$	$\mathbf{r}$	$-\mathbf{i}$	-1	$-\mathbf{k}$	$\mathbf{j}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{l}$	$\mathbf{p}$	$-\mathbf{q}$	$-\mathbf{j}$	$\mathbf{k}$	-1	$-\mathbf{i}$
$\mathbf{l}$	$\mathbf{l}$	$-\mathbf{r}$	$\mathbf{q}$	$\mathbf{p}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{i}$	-1

Oprócz tego, jeśli

$$\beta = b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} + b_4\mathbf{p} + b_5\mathbf{q} + b_6\mathbf{r} + b_7\mathbf{l},$$

gdzie $b_i \in \mathbb{R}$, to

$$\alpha = \beta \quad \Leftrightarrow \quad a_i = b_i$$

dla wszystkich $i = 0, 1, \ldots, 7$.

■ Z oktonionem

$$\alpha = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} + a_4\mathbf{p} + a_5\mathbf{q} + a_6\mathbf{r} + a_7\mathbf{l} \in \mathbb{O},$$

gdzie $a_i \in \mathbb{R}$ $(i = 1, \ldots, 7)$, są związane takie pojęcia:

- oktonion

$$\overline{\alpha} = a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k} - a_4\mathbf{p} - a_5\mathbf{q} - a_6\mathbf{r} - a_7\mathbf{l}$$

jest nazywany *sprzężonym* z α;

- liczba
$$||\alpha|| = \sqrt{\alpha\overline{\alpha}}$$
jest nazywana jego *normą*;
- oktonion
$$\operatorname{Re}\alpha = \frac{\alpha + \overline{\alpha}}{2} = a_0$$
jest nazywany jego *częścią rzeczywistą*;
- oktonion
$$\operatorname{Im}\alpha = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} + a_4\mathbf{p} + a_5\mathbf{q} + a_6\mathbf{r} + a_7\mathbf{l}$$
jest nazywany jego *częścią urojoną*.

Ponadto dla oktonionów $\alpha, \beta \in \mathbb{O}$ określono:
- ich *komutator*
$$[\alpha, \beta] = \alpha\beta - \beta\alpha;$$
- *iloczyn wektorowy*
$$\alpha \times \beta = \frac{\alpha\beta - \beta\alpha}{2}.$$

W 1958 r. M. Kervaire[18] i, niezależnie, R. Bott[19] oraz J. Milnor[20] opublikowali dowód takiego twierdzenia.

Twierdzenie 4.8.2. *Każda skończenie wymiarowa algebra z dzieleniem (jako przestrzeń liniowa nad ciałem liczb rzeczywistych* $\mathbb{R}$*) ma jeden z wymiarów* 1, 2, 4 *lub* 8.

□

Ćwiczenia 4.8.3.

(**1**) Udowodnić, że dla oktonionów $\alpha = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} + a_4\mathbf{p} + a_5\mathbf{q} + a_6\mathbf{r} + a_7\mathbf{l}$ oraz $\beta \in \mathbb{O}$ zachodzą takie własności:
(*a*)
$$\alpha^{-1} = \frac{\overline{\alpha}}{||\alpha||^2};$$

[18] Michel André Kervaire (1927–2007)
[19] Raoul Bott (1923–2005)
[20] John Willard Milnor (ur. 1931)

(*b*) $\alpha^{-1}\alpha = 1 = \alpha\alpha^{-1}$;
(*c*) $\alpha\overline{\alpha} = a_0^2 - a_1^2 - a_2^2 - a_3^2 - a_4^2 - a_5^2 - a_6^2 - a_7^2$;
(*d*) $\overline{\alpha} = -\frac{1}{6}(\alpha + (\mathbf{i}\alpha)\mathbf{i} + (\mathbf{j}\alpha)\mathbf{j} + (\mathbf{k}\alpha)\mathbf{k} + (\mathbf{p}\alpha)\mathbf{p} + (\mathbf{q}\alpha)\mathbf{q} + (\mathbf{r}\alpha)\mathbf{r} + (\mathbf{l}\alpha)\mathbf{l})$;
(*e*) $||\alpha\beta|| = ||\alpha|| \cdot ||\beta||$.

Uwagi. Po publikacji W. Hamiltona A. Cayley opublikował w 1845 r. pracę o oktawach, tzn. liczbach postaci

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + x\mathbf{l} + y\mathbf{p} + z\mathbf{q} + t\mathbf{r},$$

gdzie $a, b, c, d, x, y, z, t \in \mathbb{R}$, które są dodawane i mnożone jako wielomiany, biorąc pod uwagę zależności

$$\begin{array}{rcrcrcr} \mathbf{i}^2 & = & \mathbf{j}^2 & = & \mathbf{k}^2 & = & -1, \\ \mathbf{ij} & = & -\mathbf{ji} & = & \mathbf{k}, & & \\ \mathbf{il} & = & -\mathbf{li} & = & \mathbf{p}, & & \\ \mathbf{lj} & = & -\mathbf{jl} & = & \mathbf{q}, & & \\ \mathbf{kl} & = & -\mathbf{lk} & = & \mathbf{r}. & & \end{array}$$

Stąd wynika, że

$$\begin{array}{rcrcrcl} \mathbf{k}^2 & = & \mathbf{p}^2 & = & \mathbf{q}^2 & = & \mathbf{r}^2 = -1, \\ \mathbf{iq} & = & -\mathbf{qi} & = & \mathbf{r}, & & \\ \mathbf{jp} & = & -\mathbf{pj} & = & \mathbf{r}, & & \\ \mathbf{kp} & = & -\mathbf{pk} & = & \mathbf{q}. & & \end{array}$$

Ponieważ niezależnie w 1843 r. liczby te zostały również odkryte przez J. Gravesa[(21)], brata C. Gravesa, to są nazywane *liczbami Gravesa-Cayleya*. Praca J. Gravesa nie została opublikowana.

[(21)] John Thomas Graves (1806–1870)

Rozdział 5

Ciało funkcji wymiernych

5.1. Konstrukcja ciała funkcji wymiernych

■ Niech $\mathbb{F}[X]$ będzie pierścieniem wielomianów nad ciałem $\mathbb{F}$. Na iloczynie kartezjańskim

$$\mathbb{F}[X] \times (\mathbb{F}[X] \setminus \{0\})$$

rozpatrzmy relację ρ, określoną według reguły

$$(f, g)\rho(h, t) \Leftrightarrow ft = gh$$

dla wielomianów $f, h \in \mathbb{F}[X]$ oraz $g, t \in \mathbb{F}[X] \setminus \{0\}$.

Lemat 5.1.1. *ρ jest relacją równoważności na zbiorze $\mathbb{F}[X]\times(\mathbb{F}[X]\setminus\{0\})$.*

Dowód jest podobny do dowodu lematu 1.3.21.

□

■ Klasę równoważności $\rho(f, g)$ z reprezentantem (f, g) względem relacji ρ na zbiorze $\mathbb{F}[X] \times (\mathbb{F}[X] \setminus \{0\})$ składa się z par wielomianów postaci (fu, gu), gdzie $u \in \mathbb{F}[X] \setminus \{0\}$.

Konwencja. Zbiór ilorazowy

$$(\mathbb{F}[X] \times (\mathbb{F}[X] \setminus \{0\}))/\rho$$

będziemy oznaczać przez $\mathbb{F}(X)$, a klasę równoważności $\rho(f, g)$ przez

$$\frac{f}{g}$$

i nazywać *funkcją wymierną* (lub ułamkiem z *licznikiem* f i *mianownikiem* g).

■ Dwie funkcje wymierne

$$\frac{f_1}{g_1} \quad \text{oraz} \quad \frac{f_2}{g_2},$$

gdzie $f_1, f_2, g_1, g_2 \in \mathbb{F}[X]$ oraz $g_1 \neq 0$, $g_2 \neq 0$, są nazywane *równymi* (co zapisujemy w postaci

$$\frac{f_1}{g_1} = \frac{f_2}{g_2}),$$

jeśli $f_1 g_2 = g_1 f_2$.

Twierdzenie 5.1.2. *Niech $\mathbb{F}$ będzie ciałem. Wtedy*

$$\mathbb{F}(x) = \left\{ \frac{f}{g} \mid f, g \in \mathbb{F}[X] \text{ oraz } g \neq 0 \right\}$$

jest ciałem względem działań (dodawania „+” i mnożenia „·”), określonych według wzorów:

$$\begin{array}{rcl} \frac{f_1}{g_1} + \frac{f_2}{g_2} & = & \frac{f_1 g_2 + g_1 f_2}{g_1 g_2}, \\ \frac{f_1}{g_1} \cdot \frac{f_2}{g_2} & = & \frac{f_1 f_2}{g_1 g_2} \end{array}$$

dla wielomianów $f_i, g_i \in \mathbb{F}[X]$, gdzie $g_i \neq 0$ $(i = 1, 2)$ (dlatego $\mathbb{F}(X)$ jest nazywane *ciałem funkcji wymiernych* jednej zmiennej X nad ciałem $\mathbb{F}$).

Dowód. Zostawiamy Czytelnikowi do samodzielnego sprawdzenia spełnienie warunków z definicji ciała. □

■ Rzeczywiste funkcje wymierne, tworzące ciało $\mathbb{R}(X)$, są badane w analizie matematycznej.

* * *

■ **Właściwe funkcje wymierne.** Niech $\mathbb{F}$ będzie ciałem oraz f, f_1, g, $g_1 \in \mathbb{F}[X]$, przy czym wielomiany g i g_1 są niezerowe. Funkcja wymierna

$$\frac{f}{g}$$

jest nazywana:

- *normalizowaną*, jeśli $\mathrm{NWD}(f, g) = 1$,
- *właściwą* (lub ułamek $\frac{f}{g}$ *jest właściwy*), jeśli $\deg f < \deg g$ (przyjmujemy, że wielomian zerowy jest funkcją właściwą).

■ Łatwo sprawdzić, że dwie funkcje normalizowane $\frac{f}{g}$ i $\frac{f_1}{g_1}$ są równe w tym i tylko tym przypadku, gdy $f = f_1$ i $g = g_1$.

Lemat 5.1.3. *Każda funkcja wymierna jest sumą wielomianu i właściwej funkcji wymiernej, przy czym takie przedstawienie jest dokładnie jednoznaczne.*

Dowód. Niech

$$\frac{f}{g} \in \mathbb{F}(X).$$

Na podstawie twierdzenia o dzieleniu z resztą dla wielomianów znajdą się takie wielomiany $q, r \in \mathbb{F}[X]$, że $f = gq + r$, gdzie $\deg r < \deg g$. Wtedy

$$\frac{f}{g} = \frac{gq + r}{g} = q + \frac{r}{g}.$$

Proponujemy Czytelnikowi samodzielnie sprawdzić, że takie przedstawienie jest dokładnie jednoznaczne (z dokładnością do kolejności składników w sumie). □

Przykład 5.1.4.

Niech

$$f = \frac{X^4 + X + 1}{X^3 - X + 1} \in \mathbb{R}(X)$$

będzie rzeczywistą funkcją wymierną. Ponieważ $X^4 + X + 1 = (X^3 - X + 1)X + X^2 + 1$, to

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

jest sumą wielomianu X i ułamka właściwego

$$\frac{X^2 + 1}{X^3 - X + 1}.$$

Twierdzenie 5.1.5. *Ułamki właściwe* (*nad dowolnym ciałem* $\mathbb{F}$) *tworzą pierścień bez jedności.*

Dowód. Wystarczy sprawdzić, że suma oraz iloczyn dwóch ułamków właściwych też będą ułamkami właściwymi. Niech

$$\frac{f}{g} \text{ oraz } \frac{f_1}{g_1}$$

będą ułamkami właściwymi z $\mathbb{F}(X)$, czyli

$$\deg f < \deg g \text{ oraz } \deg f_1 < \deg g_1.$$

Wtedy

$$\deg(fg_1 + gf_1) \leqslant \max\{\deg(fg_1), \deg(gf_1)\} < \deg(gg_1)$$

oraz

$$\deg(ff_1) = \deg f + \deg f_1 < \deg g + \deg g_1 = \deg(gg_1).$$

Zatem dodawanie i mnożenie ułamków właściwych są działaniami algebraicznymi. Rzecz jasna, że 1 nie jest ułamkiem właściwym. Inne własności z definicji pierścienia proponujemy Czytelnikowi sprawdzić samodzielnie. □

■ Podzbiór niepusty S pierścienia A jest nazywany *multiplikatywnym*, jeśli zero $0 \notin S$ i zachodzi implikacja

$$a, b \in S \Rightarrow a \cdot b \in S.$$

Ćwiczenia 5.1.6.

(**1**) Udowodnić, że jeśli F jest ciałem oraz

$$S = \{\sum_{i=0}^{n} a_i X^{n-i} \mid a_n \neq 0, a_i \in F\},$$

to podzbiór $S \subseteq F[X]$ jest multiplikatywny.
(**2**) Niech A będzie dziedziną całkowitości, $S \subseteq A$. Sprawdzić, czy S jest podzbiorem multiplikatywnym, jeśli:
(*a*) $A = \mathbb{Z}[X]$ oraz $S = \{X^k \mid k \in \mathbb{N}\}$;
(*b*) $A = \mathbb{Z}$ oraz $S = \{n \in \mathbb{Z} \mid \text{NWD}(n, p) = 1\}$, gdzie p jest ustaloną liczbą pierwszą.

(**3**) Niech R będzie pierścieniem przemiennym. Udowodnić, że $S \subseteq R$ jest zbiorem multiplikatywnym, jeśli:
(a) $S = 1 + I$, gdzie $1 + I = \{1 + i \mid i \in I\}$ oraz I jest ideałem pierścienia R;
(b) $S = \{a^m \mid m \in \mathbb{N}\}$, gdzie a jest elementem nienilpotentnym pierścienia R.
(**4**) Rozłożyć funkcję wymierną w sumę wielomianu i ułamka właściwego:
(a) $\frac{X-3}{X-1}$;
(b) $\frac{5X^2+X-3}{X-2}$;
(c) $\frac{X^3-X-3}{X^2+X-1}$;
(d) $\frac{X^4-2X^3+2X-3}{X^3-3X-1}$;
(e) $\frac{X^{11}}{X^3+1}$.

Uwagi. Ułamki dziesiętne zostały wprowadzone przez al Kashiego[(1)]. Idea badania funkcji algebraicznych została wywołana jeszcze przez Kartezjusza. O funkcjach algebraicznych dyskutował też E. Warning[(2)]. Ciała liczb wymiernych $\mathbb{Q}$ i liczb rzeczywistych $\mathbb{R}$ „przeglądają się" jeszcze u Euklidesa (300 r. p.n.e.). Liczby zespolone wynaleziono w XIX w. Ciało $\mathbb{Z}_p$ zostało wprowadzone przez C. Gaussa w 1801 r. Podstawy współczesnej teorii ciał (w szczególności ciał funkcji wymiernych) stworzył E. Steinitz[(3)] w 1910 r. L. Kronecker[(4)] w 1882 r. zauważył, że jeśli X jest elementem przestępnym nad ciałem $\mathbb{F}$, to ciało $\mathbb{F}(X)$ jest izomorficzne z ciałem funkcji wymiernych jednej zmiennej X nad ciałem $\mathbb{F}$.

(1) al Kashi (początek XV w.)
(2) Edward Warning (1736–1798)
(3) Ernst Steinitz (1871–1928)
(4) Leopold Kronecker (1823–1891)

5.2. Ułamki proste

■ Ułamek właściwy

$$\frac{f}{g}$$

jest nazywany *prostym*, jeśli mianownik $g = p^k$ jest pewną potęgą wielomianu nieprzywiedlnego $p \in \mathbb{F}[X]$ nad ciałem $\mathbb{F}$ ($k \geqslant 1$), przy czym $\deg f < \deg p$.

Przykłady 5.2.1.

(**1**) Rzecz jasna, że w wyniku podstawowego twierdzenia algebry ułamek prosty nad ciałem liczb zespolonych $\mathbb{C}$ ma postać

$$\frac{A}{(z+a)^n},$$

gdzie $a, A \in \mathbb{C}$ oraz $n \in \mathbb{N}^*$.

(**2**) Nad ciałem liczb rzeczywistych $\mathbb{R}$ istnieją dwa typy ułamków prostych:

- ułamek *pierwszego rodzaju*

$$\frac{A}{(X+a)^n},$$

gdzie $a, A \in \mathbb{R}$ oraz $n \in \mathbb{N}^*$;

- ułamek *drugiego rodzaju*

$$\frac{AX+B}{(X^2+pX+q)^n},$$

gdzie $A, B, p, q \in \mathbb{R}$, $n \in \mathbb{N}^*$ oraz $\Delta = p^2 - 4q < 0$.

Twierdzenie 5.2.2. *Niech $f, g_1, \ldots, g_n \in \mathbb{F}[X]$, gdzie $\mathbb{F}$ jest ciałem. Jeśli*

$$\frac{f}{g_1 \cdots g_n}$$

jest ułamkiem właściwym i wielomiany $g_1, \ldots, g_n$ są parami względnie pierwsze, to ten ułamek dokładnie jednoznacznie rozkłada się w sumę

$$\frac{f}{g_1 \cdots g_n} = \frac{f_1}{g_1} + \cdots + \frac{f_n}{g_n}$$

pewnych ułamków właściwych postaci $\frac{f_1}{g_1}, \ldots, \frac{f_n}{g_n}$.

Dowód. Istnienie rozłożenia. Stosujemy rozumowania indukcyjne względem liczby n względnie pierwszych czynników w mianowniku ułamka. Niech $n = 2$ oraz

$$\frac{f}{g_1 g_2} \in \mathbb{F}(X)$$

będzie ułamkiem właściwym, gdzie $g_1, g_2 \in \mathbb{F}[X]$ są parami względnie pierwsze. Wtedy na podstawie wniosku z algorytmu Euklidesa znajdą się wielomiany $u_1, u_2 \in \mathbb{F}[X]$ takie, że $1 = g_1 u_1 + g_2 u_2$. Zatem $f = 1 \cdot f = g_1(u_1 f) + g_2(u_2 f)$ oraz

$$\frac{f}{g_1 g_2} = \frac{f u_1}{g_2} + \frac{f u_2}{g_1}. \tag{5.1}$$

Na podstawie twierdzenia o dzieleniu z resztą istnieją takie wielomiany $q, r \in \mathbb{F}[X]$, że $f u_1 = g_2 q + r$, gdzie $\deg r < \deg g_2$. Biorąc pod uwagę, że

$$\frac{f}{g_1 g_2} = (q + \frac{r}{g_2}) + \frac{f u_2}{g_1} = \frac{r}{g_2} + \frac{q g_1 + f u_2}{g_1}, \tag{5.2}$$

gdzie $\frac{r}{g_2}$ jest ułamkiem właściwym, w wyniku (5.2) otrzymujemy

$$g_1 r + g_2(q g_1 + f u_2) = f.$$

Ponieważ

$$\deg(g_1 r) < \deg(g_1 g_2),\ \deg f < \deg(g_1 g_2),$$

to

$$\deg(g_2(q g_1 + f u_2)) < \deg(g_1 g_2),$$

a więc $\deg(g_1 q + f u_2) < \deg g_1$ oraz ułamek

$$\frac{q g_1 + f u_2}{g_1}$$

jest właściwy. Na podstawie (5.1) wnosimy, że $\frac{f}{g_1 g_2}$ jest sumą ułamków właściwych.

Załóżmy, że teza zachodzi dla dowolnego ułamka właściwego postaci

$$\frac{f}{g_1 \cdots g_{n-1}},$$

gdzie $g_1 \cdots g_{n-1}$ jest iloczynem $n-1$ parami względnie pierwszych wielomianów $g_1, \ldots, g_{n-1} \in \mathbb{F}[X]$. Niech

$$\frac{f}{g_1 \cdots g_n}$$

będzie ułamkiem właściwym, gdzie wielomiany $g_1, \ldots, g_n \in \mathbb{F}[X]$ są parami względnie pierwsze. Wtedy wielomiany

$$G_1 = g_1 \cdots g_{n-1} \text{ oraz } G_2 = g_n$$

są względnie pierwsze i, na mocy udowodnionego wyżej, ułamek $\frac{f}{G_1 G_2}$ rozkłada się w sumę dwóch ułamków właściwych

$$\frac{f_1}{G_1} \quad \text{oraz} \quad \frac{f_2}{G_2}.$$

Stosując założenie indukcyjne do ułamka $\frac{f_1}{G_1}$, otrzymujemy tezę.

Jednoznaczność rozłożenia. W tym celu załóżmy, że

$$\frac{w_1}{g_1} + \frac{w_2}{g_2} = \frac{f}{g_1 g_2} = \frac{v_1}{g_1} + \frac{v_2}{g_2},$$

gdzie $\frac{v_i}{g_i}, \frac{w_i}{g_i}$ są ułamkami właściwymi, $v_i, w_i, g_i \in \mathbb{F}[X]$ $(i = 1, 2)$. Stąd

$$\frac{w_1 - v_1}{g_1} = \frac{v_2 - w_2}{g_2},$$

a więc $(w_1 - v_1)g_2 = g_1(v_2 - w_2)$. Skoro $\mathrm{NWD}(g_1, g_2) = 1$, to g_1 dzieli $w_1 - v_1$. Lecz

$$\deg(w_1 - v_1) < \deg g_1$$

i dlatego $w_1 - v_1 = 0$. Podobne rozumowania dają $v_2 - w_2 = 0$. Załóżmy, że teza zachodzi dla ułamka właściwego

$$\frac{f}{t_1 \cdots t_{n-1}} \in \mathbb{F}(X),$$

gdzie $t_1 \cdots t_{n-1}$ jest iloczynem $n-1$ parami względnie pierwszych wielomianów $t_1, \ldots, t_{n-1} \in \mathbb{F}[X]$.

Teraz niech

$$\frac{f}{g_1 g_2 \cdots g_n} \in \mathbb{F}(X)$$

będzie ułamkiem właściwym, gdzie $g_1, g_2, \ldots, g_n$ są parami względnie pierwsze wielomiany z pierścienia $\mathbb{F}[X]$. Wtedy $g = g_1(g_2 \cdots g_n)$ jest iloczynem dwóch względnie pierwszych wielomianów g_1 oraz $g_2 \cdots g_n$. Jak udowodniono wyżej,

$$\frac{f}{g_1 \cdots g_n} = \frac{f_1}{g_1} + \frac{h}{g_2 \cdots g_n}$$

jest sumą dwóch ułamków właściwych, gdzie h jest pewnym wielomianem z $\mathbb{F}[X]$, i takie przedstawienie jest dokładnie jednoznaczne. Zostało nam zastosowanie założenia indukcyjnego do ułamka

$$\frac{h}{g_2 \ldots g_n}$$

i teza zachodzi. □

■ Każdy wielomian $g \in \mathbb{F}[X]$ ma *rozkład kanoniczny*

$$g = a_0 p_1^{k_1} \cdots p_s^{k_s},$$

gdzie a_0 jest jego współczynnikiem najwyższym, $k_1, \ldots, k_s$ są pewnymi liczbami naturalnymi niezerowymi,

$$\deg g = \sum_{i=1}^{s} k_i \deg p_i$$

oraz $p_1, \ldots, p_s$ są parami różnymi unormowanymi wielomianami nieprzywiedlnymi nad ciałem $\mathbb{F}$. W wyniku udowodnionego wyżej twierdzenia

$$a_0 \frac{f}{g} = \frac{f}{p_1^{k_1} \cdots p_s^{k_s}} = \frac{f_1}{p_1^{k_1}} + \cdots + \frac{f_s}{p_s^{k_s}} \tag{5.3}$$

jest sumą ułamków właściwych.

Twierdzenie 5.2.3. *Niech* $f, p \in \mathbb{F}[X]$, $\deg f < \deg p^k$ *oraz* $k \in \mathbb{N}^*$. *Wtedy*

$$\frac{f}{p^k} = \frac{f_1}{p} + \frac{f_2}{p^2} + \cdots + \frac{f_k}{p^k}$$

dla pewnych wielomianów $f_1, f_2, \ldots, f_k \in \mathbb{F}[X]$ *takich, że*

$$\deg f_i < \deg p \quad (i = 1, 2, \ldots, k).$$

Dowód. Z twierdzenia o dzieleniu z resztą $f = pq + r$ dla pewnych $q, r \in \mathbb{F}[X]$, gdzie $\deg r < \deg p$. Stąd

$$\frac{f}{p^k} = \frac{pq + r}{p^k} = \frac{r}{p^k} + \frac{q}{p^{k-1}},$$

gdzie $\frac{r}{p^k}$ jest ułamkiem prostym.

Dalej otrzymujemy $q = pq_1 + r_1$, gdzie $q_1, r_1 \in \mathbb{F}[X]$ oraz $\deg r_1 < \deg p$. Wtedy

$$\frac{q}{p^{k-1}} = \frac{r_1}{p^{k-1}} + \frac{q_1}{p^{k-2}}.$$

Rozumując podobnie przez skończoną liczbę kroków, otrzymamy wynik. □

Z twierdzeń 5.2.2, 5.2.3 i równości (5.3) otrzymujemy następny

Wniosek 5.2.4. *Każdy ułamek właściwy z ciała funkcji wymiernych* $\mathbb{F}(X)$ *jest sumą ułamków prostych, przy czym taki rozkład jest jedyny (z dokładnością do kolejności czynników).*

Przykład 5.2.5.

Przedstawmy rzeczywistą funkcję wymierną

$$\frac{13X^3 + 15X^2 + 3X + 20}{X(X+2)(X^2+1)}$$

w postaci sumy ułamków prostych. Ponieważ $X,\ X+2,\ X^2+1 \in \mathbb{R}[X]$ są wielomianami nieprzywiedlnymi nad ciałem liczb rzeczywistych $\mathbb{R}$, to poszukujemy rozkładu w takiej postaci

$$\frac{13X^3 + 15X^2 + 3X + 20}{X(X+2)(X^2+1)} = \frac{A}{X} + \frac{B}{X+2} + \frac{CX+D}{X^2+1},$$

gdzie współczynniki $A, B, C, D \in \mathbb{R}$ są niewiadome. Wtedy

$$\frac{13X^3+15X^2+3X+20}{X(X+2)(X^2+1)} = \frac{A(X+2)(X^2+1)+BX(X^2+1)+(CX+D)X(X+2)}{X(X+2)(X^2+1)},$$

a stąd wynika, że

$$\begin{gathered}13X^3+15X^2+3X+20 = \\ = (A+B+C)X^3+(2A+2C+D)X^2+(A+B+2D)X+2A.\end{gathered}$$

Porównując współczynniki obok jednakowych potęg zmiennej X, otrzymujemy układ równań liniowych

$$\left\{\begin{array}{rcrcrcrcl} A & + & B & + & C & & & = & 13, \\ 2A & + & & + & 2C & + & D & = & 15, \\ A & + & B & + & & & 2D & = & 3, \\ 2A & & & & & & & = & 20. \end{array}\right.$$

Wtedy $A = 10$ i układ możemy przepisać w postaci

$$\left\{\begin{array}{rcrcrcl} B & + & C & & & = & 3, \\ & & 2C & + & D & = & -5, \\ B & & & & 2D & = & -7. \end{array}\right.$$

Sprowadzając macierz rozszerzoną ostatniego układu do postaci schodkowej

$$A_b = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 2 & 1 & -5 \\ 1 & 0 & 2 & -7 \end{array}\right] \overset{w_3-w_1}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 2 & 1 & -5 \\ 0 & -1 & 2 & -10 \end{array}\right]$$

$$\overset{\frac{2w_3+w_2}{5}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 2 & 1 & -5 \\ 0 & 0 & 1 & -5 \end{array}\right],$$

zapisujemy odpowiadający jej układ schodkowy

$$\left\{\begin{array}{rcrcrcl} B & + & C & & & = & 3, \\ & & 2C & + & D & = & -5, \\ & & & & D & = & -5 \end{array}\right.$$

i obliczamy rozwiązanie

$$\left\{\begin{array}{rcr} A & = & 10, \\ B & = & 3, \\ C & = & 0, \\ D & = & -5. \end{array}\right.$$

Zatem

$$\frac{13X^3+15X^2+3X+20}{X(X+2)(X^2+1)} = \frac{10}{X} + \frac{3}{X+2} - \frac{5}{X^2+1}.$$

Ćwiczenia 5.2.6.

(**1**) Rozłożyć ułamek właściwy w sumę ułamków prostych (nad ciałem liczb rzeczywistych $\mathbb{R}$):

(a) $\frac{4X-3}{(X-1)(X-3)}$;
(b) $\frac{-X}{X^2+3X+4}$;
(c) $\frac{X^3-X-3}{X^4+2X^3+5X^2}$;
(d) $\frac{X-2}{X^2+3X}$;
(e) $\frac{X^3+4X^2+2}{2X^5+X^3}$;
(f) $\frac{X}{X^4-4X^3+10X^2-12X+9}$;
(g) $\frac{X^3-X^2-5X-2}{(X-1)^4}$;
(h) $\frac{3X^4-12X^3+18X^2-12X+7}{(X-1)^2(X^2-2X+2)^3}$.

Uwagi. Ogólne pojęcie pierścienia ułamków zostało zdefiniowane przez H. Grella[(5)] w 1926 r.

[(5)] Heinrich Grell (1903–1974)

Rozdział 6

Rozszerzenia ciał

6.1. Stopień rozszerzenia ciał

■ Niech $\mathbb{E}$ oraz $\mathbb{F}$ będą ciałami. Jeśli istnieje monomorfizm ciał $\varphi : \mathbb{F} \to \mathbb{E}$, to mówimy, że ciało $\mathbb{E}$ jest *rozszerzeniem* ciała $\mathbb{F}$ (zapisujemy $\mathbb{F} \leqslant \mathbb{E}$).

Wtedy $\varphi(\mathbb{F}) = \{\varphi(a) \mid a \in \mathbb{F}\}$ jest podciałem ciała $\mathbb{E}$ oraz ciała $\mathbb{F}$ i $\varphi(\mathbb{F})$ są izomorficzne. Zatem bez ograniczenia ogólności możemy uważać, że $\mathbb{E}$ jest *rozszerzeniem* ciała $\mathbb{F}$, jeśli $\mathbb{F}$ jest podciałem ciała $\mathbb{E}$.

■ Jeśli $\mathbb{E}$ jest rozszerzeniem ciała $\mathbb{F}$, to $\mathbb{E}$ jest przestrzenią liniową nad ciałem $\mathbb{F}$. Wtedy jej wymiar $\dim_{\mathbb{F}} \mathbb{E}$ jest nazywany *stopniem rozszerzenia* $\mathbb{F} \leqslant \mathbb{E}$ (oznaczanym przez $|\mathbb{E} : \mathbb{F}|$).

■ Jeśli stopień $|\mathbb{E} : \mathbb{F}|$ jest dodatnią liczbą całkowitą, to mówimy, że rozszerzenie $\mathbb{F} \leqslant \mathbb{E}$ jest *skończone*.

Twierdzenie 6.1.1. *Jeśli mamy łańcuch $\mathbb{G} \leqslant \mathbb{F} \leqslant \mathbb{E}$ rozszerzeń skończonych ciał $\mathbb{G}, \mathbb{F}, \mathbb{E}$, to zachodzi równość*

$$|\mathbb{E} : \mathbb{G}| = |\mathbb{E} : \mathbb{F}| \cdot |\mathbb{F} : \mathbb{G}|.$$

Dowód. Załóżmy, że $|\mathbb{E} : \mathbb{F}| = n$ oraz $|\mathbb{F} : \mathbb{G}| = m$ dla pewnych dodatnich liczb całkowitych n, m oraz $E = (e_1, e_2, \ldots, e_n)$ (odpowiednio $F = (f_1, f_2, \ldots, f_m)$) jest bazą przestrzeni liniowej $\mathbb{E}$ nad ciałem $\mathbb{F}$ (odpowiednio przestrzeni liniowej $\mathbb{F}$ nad ciałem $\mathbb{G}$). Wtedy każdy element $a \in \mathbb{E}$ możemy przedstawić w postaci kombinacji liniowej $a = \sum_{i=1}^{n} \alpha_i e_i$,

gdzie $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, oraz każdy element $\alpha_i \in \mathbb{F}$ możemy przedstawić w postaci $\alpha_i = \sum_{j=1}^{m} \gamma_{ij} f_j$ dla pewnych $\gamma_{i1}, \ldots, \gamma_{im} \in \mathbb{G}$. W konsekwencji

$$a = \sum_{i=1}^{n} \sum_{j=1}^{m} \gamma_{ij} e_i f_j,$$

czyli zbiór

$$B = \{e_i f_j \mid i = 1, \ldots, n;\ j = 1, \ldots, m\}$$

jest układem generatorów przestrzeni liniowej $\mathbb{E}$ nad ciałem $\mathbb{G}$.

Zostało nam udowodnienie, że układ B jest liniowo niezależny nad ciałem $\mathbb{G}$. W tym celu załóżmy, że

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \sigma_{ij} e_i f_j = 0 \tag{6.1}$$

dla pewnych współczynników $\sigma_{ij} \in \mathbb{G}$. Skoro (6.1) możemy przepisać w postaci

$$\sum_{i=1}^{n} (\sum_{j=1}^{m} \sigma_{ij} f_j) e_i = 0$$

oraz układ E jest liniowo niezależny nad ciałem $\mathbb{F}$, to współczynniki

$$\sum_{j=1}^{m} \sigma_{ij} f_j = 0 \quad (1 \leqslant i \leqslant n)$$

są zerowe. Wtedy z liniowej niezależności układu F nad ciałem $\mathbb{G}$ wnosimy, że wszystkie współczynniki $\sigma_{ij} = 0$. Udowodniliśmy, że B jest liniowo niezależnym układem generatorów przestrzeni liniowej $\mathbb{E}$ nad ciałem $\mathbb{G}$. To znaczy, że układ B jest bazą tej przestrzeni oraz

$$|\mathbb{E} : \mathbb{G}| = \dim_{\mathbb{G}} \mathbb{E} = |B| = mn$$

i teza zachodzi. □

Przykłady 6.1.2.

(**1**) Każde ciało $\mathbb{F}$ jest przestrzenią liniową nad ciałem $\mathbb{F}$. Jeśli α jest elementem przestrzeni liniowej $\mathbb{F}$ oraz element $1 \in \mathbb{F}$ jest jednością ciała $\mathbb{F}$ (rozpatrywany jako element przestrzeni liniowej $\mathbb{F}$), to

$$\alpha = \alpha \cdot 1.$$

Ponieważ układ $\{1\}$ jest liniowo niezależny nad ciałem $\mathbb{F}$, to jest bazą przestrzeni liniowej $\mathbb{F}$ nad ciałem $\mathbb{F}$. Zatem $|\mathbb{F} : \mathbb{F}| = 1$.

(**2**) Udowodniliśmy wcześniej, że

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$$

jest ciałem. Niech dalej $a, b, c, d \in \mathbb{Q}$ będą dowolnymi liczbami wymiernymi. Najpierw zaobserwujmy, że

$$a + b\sqrt{2} = c + d\sqrt{2} \;\Leftrightarrow\; a - c = (d - b)\sqrt{2}.$$

Jeśli $d - b \neq 0$, to

$$\mathbb{Q} \ni \frac{a - c}{d - b} = \sqrt{2} \notin \mathbb{Q},$$

a więc otrzymujemy sprzeczność. Zatem $d - b = 0$ oraz $a - c = 0$. Udowodniliśmy zatem, że

$$a + b\sqrt{2} = c + d\sqrt{2} \;\Leftrightarrow\; \begin{cases} a = c, \\ b = d. \end{cases}$$

Jeśli $a \cdot 1 + b \cdot \sqrt{2} = 0$, to $a = b = 0$, czyli układ $B = \{1, \sqrt{2}\}$ jest liniowo niezależny nad ciałem liczb wymiernych $\mathbb{Q}$. Oprócz tego każdy element $\alpha \in \mathbb{Q}[\sqrt{2}]$ ma postać

$$\alpha = x \cdot 1 + y \cdot \sqrt{2}$$

dla pewnych $x, y \in \mathbb{Q}$ i na tej podstawie B jest układem generatorów przestrzeni liniowej $\mathbb{Q}[\sqrt{2}]$ nad ciałem $\mathbb{Q}$ oraz stopień rozszerzenia $\mathbb{Q} \leqslant \mathbb{Q}[\sqrt{2}]$ jest równy

$$|\mathbb{Q}[\sqrt{2}] : \mathbb{Q}| = \dim_{\mathbb{Q}} \mathbb{Q}[\sqrt{2}] = 2.$$

(**3**) Każda liczba zespolona $z \in \mathbb{C}$ ma postać algebraiczną

$$z = x \cdot 1 + y \cdot i$$

dla pewnych liczb rzeczywistych $x, y \in \mathbb{R}$, czyli układ $B = \{1, i\}$ jest układem generatorów przestrzeni liniowej $\mathbb{C}$ nad ciałem $\mathbb{R}$. Skoro

$$x \cdot 1 + y \cdot i = 0 \;\Leftrightarrow\; \begin{cases} x = 0, \\ y = 0, \end{cases}$$

to układ B jest liniowo niezależny (a więc jest bazą przestrzeni liniowej $\mathbb{C}$ nad ciałem $\mathbb{R}$). Zatem $|\mathbb{C} : \mathbb{R}| = 2$.

■ Zatem $|\mathbb{E} : \mathbb{F}| \geqslant 1$ dla każdego rozszerzenia ciał $\mathbb{F} \leqslant \mathbb{E}$.

■ Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał $\mathbb{E}$ oraz $\mathbb{F}$. Element $\theta \in \mathbb{E}$ jest nazywany *algebraicznym* nad ciałem $\mathbb{F}$, jeśli istnieje taki wielomian niezerowy $f \in \mathbb{F}[X]$, że $f(\theta) = 0$. Jeśli element $\theta \in \mathbb{E}$ nie jest algebraiczny nad $\mathbb{F}$, to jest nazywany *transcendentnym* (=*przestępnym*) nad ciałem $\mathbb{F}$. Jeśli każdy element z $\mathbb{E}$ jest algebraiczny nad ciałem $\mathbb{F}$, to rozszerzenie $\mathbb{F} \leqslant \mathbb{E}$ jest nazywane *algebraicznym*.

Twierdzenie 6.1.3 (Bézouta[(1)])**.** *Niech $\mathbb{F}$ będzie ciałem oraz $\theta \in \mathbb{F}$. Wtedy θ jest pierwiastkiem wielomianu niezerowego $f \in \mathbb{F}[X]$ w tym i tylko tym przypadku, gdy $f = (X - \theta) \cdot g$ dla pewnego wielomianu $g \in \mathbb{F}[X]$.*

Dowód. ($\Leftarrow$) Jest oczywiste.

($\Rightarrow$) Z twierdzenia o dzieleniu z resztą dla wielomianów wynika, że istnieją takie $q, r \in \mathbb{F}[X]$, że $f = (X-\theta)\cdot q + r$ oraz $\deg r < \deg(X-\theta) = 1$. Stąd wynika, że $r \in \mathbb{F}$. Wtedy wartość $r(\theta) = r$ oraz

$$0 = f(\theta) = (\theta - \theta)q(\theta) + r(\theta) = r$$

i teza zachodzi. □

Twierdzenie 6.1.4. *Każde skończone rozszerzenie $\mathbb{F} \leqslant \mathbb{E}$ ciał $\mathbb{E}$ oraz $\mathbb{F}$ jest algebraiczne.*

Dowód. Ponieważ $|\mathbb{E} : \mathbb{F}| < \infty$, to $|\mathbb{E} : \mathbb{F}| = n$ dla pewnej dodatniej liczby całkowitej n. Jeśli $\theta \in \mathbb{E}$, to układ

$$\{\theta^0, \theta^1, \ldots, \theta^n\}$$

zawiera $n+1$ elementów, a więc jest liniowo zależny nad ciałem $\mathbb{F}$. Zatem istnieją współczynniki $a_0, a_1, \ldots, a_n \in \mathbb{E}$, wśród których jest co najmniej jeden $a_i \neq 0$ $(0 \leqslant i \leqslant n)$ niezerowy taki, że

$$a_0\theta^0 + a_1\theta^1 + \cdots + a_n\theta^n = 0 \in \mathbb{F}.$$

Z tego wynika, że istnieje wielomian niezerowy

$$f = a_0X^0 + a_1X^1 + \cdots + a_nX^n \in \mathbb{F}[X]$$

(1) Étienne Bézout (1730–1783)

spełniający warunek $f(\theta) = 0$. Udowodniliśmy zatem, że rozszerzenie skończone $\mathbb{F} \leqslant \mathbb{E}$ jest algebraiczne. □

Ćwiczenia 6.1.5.

(**1**) Udowodnić, że następujące rozszerzenia są skończone:
(*a*) $\mathbb{Q} \leqslant \mathbb{Q}(\sqrt{7})$;
(*b*) $\mathbb{Q} \leqslant \mathbb{Q}(\sqrt{5} - \sqrt{7})$;
(*c*) $\mathbb{Q} \leqslant \mathbb{Q}(i\sqrt{11})$.
(**2**) Obliczyć stopień rozszerzenia ciał:
(*a*) $|\mathbb{Q}(\sqrt[3]{5}) : \mathbb{Q}|$;
(*b*) $|\mathbb{Q}(i, 5i) : \mathbb{Q}|$;
(*c*) $|\mathbb{C} : \mathbb{R}(\sqrt{-5})|$;
(*d*) $|\mathbb{Q}(i, \sqrt{5}) : \mathbb{Q}|$.

Uwagi. Własności rozszerzeń ciał okazały się podstawowymi w teorii Galois.

6.2. Proste rozszerzenia ciał

Twierdzenie 6.2.1 (o wielomianie minimalnym)**.** *Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał oraz $\theta \in \mathbb{E}$. Jeśli element θ jest algebraiczny nad ciałem $\mathbb{F}$, to zachodzą własności:*

(1) *istnieje wielomian $m_\theta \in \mathbb{F}[X]$ spełniający warunki:*
  (*a*) *m_θ jest unormowany;*
  (*b*) *m_θ ma pierwiastek θ;*
  (*c*) *m_θ jest nieprzywiedlny nad ciałem $\mathbb{F}$;*

(2) *jeśli $g \in \mathbb{F}[X]$ oraz $g(\theta) = 0$, to m_θ dzieli g*

(wielomian m_θ, spełniający warunki (*a*), (*b*) oraz (*c*), jest nazywany *wielomianem minimalnym* elementu θ nad ciałem $\mathbb{F}$).

Dowód. (1) Skoro θ jest algebraiczny nad ciałem $\mathbb{F}$, to zbiór

$$S = \{\deg f \mid 0 \neq f \in \mathbb{F}[X] \text{ oraz } f(\theta) = 0\} \subseteq \mathbb{N}$$

jest niepusty. Na podstawie zasady minimum istnieje element minimalny s w zbiorze S, a więc znajdzie się taki wielomian

$$g = a_0X^s + a_1X^{s-1} + \cdots + a_{s-1}X + a_s \in \mathbb{F}[X],$$

że $a_0 \neq 0$ oraz $g(\theta) = 0$. Połóżmy

$$m_\theta = \frac{1}{a_0}g.$$

Wtedy $m_\theta \in \mathbb{F}[X]$ spełnia warunki (*a*) oraz (*b*). Jeśli $\deg m_\theta = 1$, to jest nieprzywiedlny. Załóżmy, że $\deg m_\theta \geqslant 2$ oraz $m_\alpha = g_1g_2$ dla pewnych wielomianów $g_1, g_2 \in \mathbb{F}[X]$ takich, że $1 \leqslant \deg g_i < \deg m_\theta$ $(i = 1, 2)$. Skoro

$$0 = m_\theta(\theta) = g_1(\theta)g_2(\theta)$$

oraz $g_1(\theta)$, $g_2(\theta)$ są elementami ciała $\mathbb{F}$, to, na przykład, $g_1(\theta) = 0$. Ponieważ $\deg g_1 < \deg m_\theta$, to otrzymujemy sprzeczność. To znaczy, że zachodzi warunek (*c*).

(2) Z twierdzenia o dzieleniu z resztą mamy $g = m_\theta q + r$ dla pewnych $q, r \in \mathbb{F}[X]$ takich, że $\deg r < \deg m_\theta$. Ponieważ

$$0 = g(\theta) = m_\theta(\theta)q(\theta) + r(\theta) = r(\theta),$$

to w wyniku minimalności wielomianu m_θ wnosimy, że reszta $r = 0$ jest zerowa i teza zachodzi. □

■ Jeśli $\mathbb{F} \leqslant \mathbb{E}$ jest rozszerzeniem ciał oraz $\theta \in \mathbb{E}$, to oznaczamy:

- $$\mathbb{F}[\theta] = \{f(\theta) \mid f \in \mathbb{F}[X]\};$$
- $$\mathbb{F}(\theta) = \left\{ \frac{f(\theta)}{g(\theta)} \mid f, g \in \mathbb{F}[X] \text{ oraz } \ g(\theta) \neq 0 \right\}.$$

Lemat 6.2.2. *Jeśli $\mathbb{F} \leqslant \mathbb{E}$ jest rozszerzeniem ciał oraz $\theta \in \mathbb{E}$, to:*
(1) *$\mathbb{F}[\theta]$ jest podpierścieniem w $\mathbb{E}$;*
(2) *$\mathbb{F}(\theta)$ jest podciałem w $\mathbb{E}$;*
(3) *$\mathbb{F}[\theta] \subseteq \mathbb{F}(\theta)$.*

Dowód. Udowodnić samodzielnie. □

■ Rozszerzenie ciał $\mathbb{F} \leqslant \mathbb{E}$ jest nazywane *prostym*, jeśli $\mathbb{E} = \mathbb{F}(\theta)$ dla pewnego elementu $\theta \in \mathbb{E}$.

Twierdzenie 6.2.3 (struktura prostego rozszerzenia ciał)**.** *Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał oraz $\theta \in \mathbb{E}$. Wtedy zachodzą własności:*
(1) *pierścień ilorazowy $\mathbb{F}[X]/\langle m_\theta \rangle$, gdzie $m_\theta \in \mathbb{F}[X]$ jest wielomianem minimalnym elementu θ nad ciałem $\mathbb{F}$, jest izomorficzny z $\mathbb{F}[\theta]$;*
(2) *jeśli element θ jest przestępny nad $\mathbb{F}$, to pierścień $\mathbb{F}[\theta]$ jest izomorficzny z pierścieniem wielomianów $\mathbb{F}[X]$;*
(3) *jeśli element θ jest algebraiczny nad $\mathbb{F}$, to pierścień $\mathbb{F}[\theta] = \mathbb{F}(\theta)$ jest ciałem oraz stopień rozszerzenia $|\mathbb{F}(\theta) : \mathbb{F}| = \deg m_\theta$;*
(4) *jeśli $\alpha \in \mathbb{F}(\theta)$, to $\deg m_\alpha$ jest dzielnikiem liczby $\deg m_\theta$.*

Dowód. (1) Odwzorowanie

$$\varphi : \mathbb{F}[X] \ni f \mapsto f(\theta) \in \mathbb{F}[\theta]$$

jest homomorfizmem pierścieni (sprawdzić samodzielnie), jego jądro

$$\operatorname{Ker} \varphi = \{f \in \mathbb{F}[X] \mid f(\theta) = 0\} = \langle m_\theta \rangle$$

oraz obraz

$$\operatorname{Im}\varphi = \{f(\theta) \mid f \in \mathbb{F}[X]\} = \mathbb{F}[\theta].$$

Na mocy pierwszego twierdzenia o izomorfizmie (patrz twierdzenie 4.3.6) pierścień $\mathbb{F}[\theta]$ oraz pierścień ilorazowy $\mathbb{F}[X]/\langle m_\theta\rangle$ są izomorficzne.

(2) Jeśli $\theta \in \mathbb{E}$ jest przestępny nad $\mathbb{F}$, to wielomian $m_\theta = 0$ jest zerowy i teza zachodzi na podstawie części (1).

(3) Jeśli $\theta \in \mathbb{E}$ jest algebraiczny nad $\mathbb{F}$, to wielomian $m_\theta \in \mathbb{F}[X]$ jest nieprzywiedlny nad $\mathbb{F}$ oraz $\mathbb{F}[X]/\langle m_\theta\rangle$ jest ciałem na podstawie wniosku 4.5.7. Zatem $\mathbb{F}[\theta]$ jest ciałem, a więc $\mathbb{F}[\theta] = \mathbb{F}(\theta)$.

Dla każdego $f \in \mathbb{F}[X]$ istnieją takie wielomiany $q, r \in \mathbb{F}[X]$, że

$$f = m_\theta q + r \text{ oraz } \deg r < k,$$

gdzie $k = \deg m_\theta$. Skoro warstwa

$$\begin{gathered} f + \langle m_\theta\rangle = r + (m_\theta q + \langle m_\theta\rangle) = r + \langle m_\theta\rangle = \\ = a_0X^0 + a_1X^1 + \cdots + a_{k-1}X^{k-1} + \langle m_\theta\rangle \end{gathered}$$

dla pewnych współczynników $a_0, a_1, \ldots, a_{k-1} \in \mathbb{F}$, to

$$\begin{gathered} \mathbb{F}[X]/\langle m_\theta\rangle = \{\sum_{i=0}^{k-1} a_iX^i + \langle m_\theta\rangle \mid a_i \in \mathbb{F} \quad (i = 0, 1, \ldots, k-1)\} = \\ = \{\sum_{i=0}^{k-1} a_i\Phi^i \mid a_i \in \mathbb{F} \quad (i = 0, 1, \ldots, k-1)\}, \end{gathered}$$

gdzie $\Phi = X + \langle m_\theta\rangle$. Jeśli

$$\psi : \mathbb{F} \ni a \mapsto a + \langle m_\theta\rangle \in \mathbb{F}[X]/\langle m_\theta\rangle,$$

to obraz $\psi(\mathbb{F})$ jest ciałem izomorficznym z ciałem $\mathbb{F}$. Oprócz tego

$$B = (\Phi^0, \Phi^1, \ldots, \Phi^{k-1})$$

jest bazą przestrzeni liniowej $\mathbb{F}[X]/\langle m_\theta\rangle$ nad ciałem $\psi(\mathbb{F})$ oraz stopień rozszerzenia

$$|\mathbb{F}[X]/\langle m_\theta\rangle : \psi(\mathbb{F})| = \dim_{\psi(\mathbb{F})}(\mathbb{F}[X]/\langle m_\theta\rangle) = k = \deg m_\theta.$$

Zatem $|\mathbb{F}(\theta) : \mathbb{F}| = k = \deg m_\theta$.

(4) Ćwiczenie. □

Przykłady 6.2.4.

(**1**) Wielomian $f = X^2 + X + 2 \in \mathbb{F}_3[X]$ jest nieprzywiedlny nad ciałem $\mathbb{F}_3 = \{0, 1, 2\}$, bo nie posiada pierwiastków w ciele $\mathbb{F}_3$. Zatem $\mathbb{F} = \mathbb{F}_3[X]/\langle f \rangle$ jest ciałem. Jeśli $\theta = X + \langle f \rangle$, to

$$\mathbb{F} = \{a_0\theta^0 + a_1\theta^1 \mid a_0, a_i \in \mathbb{F}_3\} = \{0, 1, 2, \theta+1, \theta+2, 2\theta, 2\theta+1, 2\theta+2\}$$

jest ciałem o 9 elementach z takimi tabelkami Cayleya dodawania „+" i mnożenia „·":

·	0	1	2	θ	$\theta+1$	$\theta+2$	2θ	$2\theta+1$	$2\theta+2$
0	0	0	0	0	0	0	0	0	0
1	0	1	2	θ	$\theta+1$	$\theta+2$	2θ	$2\theta+1$	$2\theta+2$
2	0	2	1	2θ	$2\theta+2$	$2\theta+1$	θ	$\theta+2$	$\theta+1$
θ	0	θ	2θ	$2\theta+1$	1	$\theta+1$	$\theta+2$	$2\theta+2$	2
$\theta+1$	0	$\theta+1$	$2\theta+2$	1	$\theta+2$	2θ	2	θ	$2\theta+1$
$\theta+2$	0	$\theta+2$	$2\theta+1$	$\theta+1$	2θ	2	$2\theta+2$	1	θ
2θ	0	2θ	θ	$\theta+2$	2	$2\theta+2$	$2\theta+1$	$\theta+1$	1
$2\theta+1$	0	$2\theta+1$	$\theta+2$	$2\theta+2$	θ	1	$\theta+1$	2	2θ
$2\theta+2$	0	$2\theta+2$	$\theta+1$	2	$2\theta+1$	θ	1	2θ	$\theta+2$

+	0	1	2	θ	$\theta+1$	$\theta+2$	2θ	$2\theta+1$	$2\theta+2$
0	0	1	2	θ	$\theta+1$	$\theta+2$	2θ	$2\theta+1$	$2\theta+2$
1	1	2	0	$\theta+1$	$\theta+2$	θ	$2\theta+1$	$2\theta+2$	2θ
2	2	0	1	$\theta+2$	θ	$\theta+1$	$2\theta+2$	2θ	$2\theta+1$
θ	θ	$\theta+1$	$\theta+2$	2θ	$2\theta+1$	$2\theta+2$	0	1	2
$\theta+1$	$\theta+1$	$\theta+2$	θ	$2\theta+1$	$2\theta+2$	2θ	1	2	0
$\theta+2$	$\theta+2$	θ	$\theta+1$	$2\theta+2$	2θ	$2\theta+1$	2	0	1
2θ	2θ	$2\theta+1$	$2\theta+2$	0	1	2	θ	$\theta+1$	$\theta+2$
$2\theta+1$	$2\theta+1$	$2\theta+2$	2θ	1	2	0	$\theta+1$	$\theta+2$	θ
$2\theta+2$	$2\theta+2$	2θ	$2\theta+1$	2	0	1	$\theta+2$	θ	$\theta+1$

(**2**) Niech $m \in \mathbb{N}^*$ będzie liczbą ustaloną. Wykażmy, że stopień rozszerzenia $|\mathbb{C} : \mathbb{Q}| = \infty$. W rzeczy samej, na podstawie twierdzenia 6.2.3 stopień rozszerzenia

$$|\mathbb{Q}(\sqrt[m]{7}) : \mathbb{Q}| = \deg(X^m - 7) = m.$$

Jeśli założyć, że $|\mathbb{C} : \mathbb{Q}| = n$ dla pewnej liczby $n \in \mathbb{N}^*$, to

$$n = |\mathbb{C} : \mathbb{Q}| = |\mathbb{C} : \mathbb{Q}(\sqrt[m]{7})| \cdot |\mathbb{Q}(\sqrt[m]{7}) : \mathbb{Q}| = |\mathbb{C} : \mathbb{Q}(\sqrt[m]{7})| \cdot m.$$

To znaczy, że każda liczba całkowita dodatnia m dzieli liczbę n, co nie jest możliwe. Zatem $|\mathbb{C} : \mathbb{Q}| = \infty$.

Twierdzenie 6.2.5. *Niech $\mathbb{F} \leqslant \mathbb{E} \leqslant \mathbb{G}$ będzie rozszerzeniem ciał. Rozszerzenia $\mathbb{F} \leqslant \mathbb{E}$ i $\mathbb{E} \leqslant \mathbb{G}$ są algebraiczne wtedy i tylko wtedy, gdy rozszerzenie $\mathbb{F} \leqslant \mathbb{G}$ jest algebraiczne.*

Dowód. ($\Leftarrow$) Jest oczywiste.

($\Rightarrow$) Załóżmy, że każde z rozszerzeń ciał $\mathbb{F} \leqslant \mathbb{E}$ i $\mathbb{E} \leqslant \mathbb{G}$ jest algebraiczne. Niech $\theta \in \mathbb{G}$. Wtedy istnieje wielomian niezerowy

$$f = a_0X^n + a_1X^{n-1} + \cdots + a_n \in \mathbb{E}[X]$$

taki, że $f(\theta) = 0$. Skoro element $a_0 \in \mathbb{E}$ jest algebraiczny nad ciałem $\mathbb{F}$, to ciało $\mathbb{F}(a_0)$ jest skończonym rozszerzeniem na podstawie twierdzenia 6.2.3. Stosując rozumowania indukcyjne w taki sposób, otrzymujemy, że rozszerzenie $\mathbb{F} \leqslant \mathbb{F}(a_0, a_1, \ldots, a_n)$ ma skończony stopień, a więc jest algebraiczne na podstawie twierdzenia 6.1.4. Ponieważ rozszerzenie $\mathbb{F}(a_0, a_1, \ldots, a_n) \leqslant \mathbb{F}(a_0, a_1, \ldots, a_n, \theta)$ ma skończony stopień w wyniku twierdzenia 6.2.3, to element θ jest algebraiczny nad ciałem $\mathbb{F}$. □

■ Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał oraz $\theta, \nu \in \mathbb{E}$. Jeśli wielomiany minimalne $m_\theta = m_\nu \in \mathbb{F}[X]$ są równe, to elementy θ oraz ν są nazywane *sprzężonymi* nad ciałem $\mathbb{F}$.

Twierdzenie 6.2.6. *Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał oraz $\theta, \nu \in \mathbb{E}$. Jeśli elementy θ i ν są sprzężone nad ciałem $\mathbb{F}$, to ciała $\mathbb{F}(\theta)$ oraz $\mathbb{F}(\nu)$ są izomorficzne.*

Dowód. Skoro $m_\theta = m_\nu \in \mathbb{F}[X]$, to

$$|\mathbb{F}(\theta) : \mathbb{F}| = \deg m_\theta = \deg m_\nu = |\mathbb{F}(\nu) : \mathbb{F}|.$$

Ponieważ na podstawie twierdzenia 6.2.3 mamy

$$\mathbb{F}(\theta) = \mathbb{F}[\theta] = \{f(\theta) \mid f \in \mathbb{F}[X]\}$$

oraz $\mathbb{F}(\nu) = \mathbb{F}[\nu]$, to odwzorowanie

$$\varphi : \mathbb{F}(\theta) \ni f(\theta) \mapsto f(\nu) \in \mathbb{F}(\nu)$$

jest izomorfizmem ciał. Rzeczywiście, z równości $\varphi(f(\theta)) = \varphi(g(\theta))$ dla pewnych $f, g \in \mathbb{F}[X]$ wynika, że $(f-g)(\nu) = 0$, a więc $m_\theta \mid f - g$. Wtedy $f - g = m_\theta \cdot h$ dla pewnego $h \in \mathbb{F}[X]$, a stąd $f(\theta) = g(\theta)$. To oznacza, że odwzorowanie φ jest iniekcją. Jeśli zachodzi $h(\nu) \in \mathbb{F}(\nu)$, to $h(\nu) = \varphi(h(\theta))$, a więc φ jest suriekcją. Pozostałe własności zostawiamy Czytelnikowi do samodzielnego udowodnienia. □

Ćwiczenia 6.2.7.

(**1**) Znaleźć wielomian minimalny $f \in \mathbb{F}[X]$, który ma pierwiastek θ, jeśli:
(a) $\mathbb{F} = \mathbb{Q}$ oraz $\theta = \sqrt{2} + \sqrt{5}$;
(b) $\mathbb{F} = \mathbb{R}$ oraz $\theta = e - \pi i$.
(**2**) Sprawdzić, dla jakich wartości $k \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ pierścień ilorazowy $\mathbb{F}_7[X]/\langle X^2 + k\rangle$ jest ciałem.
(**3**) Znaleźć stopień rozszerzenia ciał:
(a) $|\mathbb{Q}(\sqrt[3]{5}) : \mathbb{Q}|$;
(b) $|\mathbb{Q}(\sqrt[3]{2}, \sqrt{5}) : \mathbb{Q}|$;
(c) $|\mathbb{Q}(\sqrt{2}, \sqrt{5}) : \mathbb{Q}|$.
(**4**) Znaleźć wielomian minimalny $m_\theta \in \mathbb{Q}[X]$, jeśli:
(a) $\theta = \sqrt{2} + \sqrt{7}$;
(b) $\theta = \sqrt{2} + \sqrt{3}$;
(c) $\theta = \sqrt{2} + e^{\frac{2\pi i}{3}}$.

Uwagi. W 1882 r. F. Lindemann dowiódł, że liczba π jest przestępna, co oznacza, że kwadratura koła nie jest możliwa.

6.3. Ciało liczb algebraicznych

■ Liczba $\theta \in \mathbb{C}$, która jest elementem algebraicznym nad ciałem liczb wymiernych $\mathbb{Q}$, jest nazywana *liczbą algebraiczną*. Stopień $\deg m_\theta$ jej wielomianu minimalnego $m_\theta \in \mathbb{Q}[X]$ jest nazywany *stopniem* liczby algebraicznej θ. Liczba $\theta \in \mathbb{C}$, która nie jest algebraiczna, jest nazywana *transcendentną* (lub *przestępną*).

Przykłady 6.3.1.

(**1**) Jeśli $f = X - \theta \in \mathbb{F}[X]$, to $f(\theta) = 0$ dla elementu $\theta \in \mathbb{F}$. To oznacza, że każdy element $\theta \in \mathbb{F}$ jest algebraiczny nad ciałem $\mathbb{F}$.

(**2**) Skoro wielomian $f = X^2 - 5 \in \mathbb{Q}[X]$ jest niezerowy oraz $f(\sqrt{5}) = 0$, to element $\sqrt{5} \in \mathbb{R}$ jest algebraiczny nad ciałem $\mathbb{Q}$ (czyli $\sqrt{5}$ jest liczbą algebraiczną).

(**3**) Przekonajmy się, że liczba $\theta = \sqrt[3]{2} + \sqrt{5}$ jest algebraiczna. Skoro $\theta - \sqrt{5} = \sqrt[3]{2}$, to $(\theta - \sqrt{5})^3 = 2$ i wtedy

$$\theta^3 + 15\theta - 2 = (3\theta^2 + 5)\sqrt{5},$$

skąd, podnosząc do kwadratu obie strony i przekształcając, otrzymujemy

$$\theta^6 - 15\theta^4 - 4\theta^3 + 75\theta^2 - 60\theta - 121 = 0,$$

czyli liczba θ jest algebraiczna.

Dla dowodu następnego twierdzenia potrzebujemy taki

Lemat 6.3.2. *Niech $\lambda \in \mathbb{C}$. Jeśli istnieje taki wektor niezerowy $\mathbf{z} \in \mathbb{C}^n$ oraz macierz kwadratowa $A \in M_n(\mathbb{Q})$, że*

$$A\mathbf{z}^T = \lambda \mathbf{z}^T, \tag{6.2}$$

to λ jest liczbą algebraiczną stopnia $\leqslant n$.

Dowód. Rzeczywiście z (6.2) otrzymujemy taki kwadratowy jednorodny układ równań liniowych

$$(A - \lambda I)\mathbf{z}^T = \mathbf{0}.$$

Ponieważ $\mathbf{z} \neq \mathbf{0}$, to wyznacznik $\det(A - \lambda I) = 0$ jest zerowy. Lecz wielomian $\det(A - X \cdot I) \in \mathbb{Q}[X]$ jest niezerowy z pierwiastkiem $\lambda \in \mathbb{C}$, a więc liczba λ jest algebraiczna. □

Twierdzenie 6.3.3. *Zbiór*

$$\mathbb{A} = \{\alpha \in \mathbb{C} \mid \text{ liczba } \alpha \text{ jest algebraiczna}\}$$

tworzy ciało (nazywane *ciałem liczb algebraicznych*).

Dowód. Na podstawie kryterium podciała wystarczy udowodnić, że

$$\begin{array}{lll} \alpha, \beta \in \mathbb{A} & \Rightarrow & \alpha - \beta, \alpha\beta \in \mathbb{A}, \\ 0 \neq \alpha \in \mathbb{A} & \Rightarrow & \alpha^{-1} \in \mathbb{A}. \end{array}$$

Załóżmy, że wielomian minimalny elementu α ma postać

$$m_\alpha = X^n + a_1 X^{n-1} + \cdots + a_{n-1} X + a_n \in \mathbb{Q}[X].$$

a) Skoro α^{-1} jest pierwiastkiem wielomianu niezerowego

$$a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + 1 \in \mathbb{Q}[X],$$

to $\alpha^{-1} \in \mathbb{A}$.

b) Załóżmy, że $\alpha\beta \neq 0$. Rozpatrzmy zbiór

$$\begin{gathered} Z = \{\alpha^k \beta^l \mid k = 0, 1, \ldots, n-1;\ l = 0, 1, \ldots, m-1\} = \\ = \{z_1, z_2, \ldots, z_{nm}\}, \end{gathered}$$

gdzie $m = \deg m_\beta$. Wtedy:

- jeśli $k < n-1$, to
$$\alpha(\alpha^k \beta^l) \in Z;$$
- jeśli $k = n-1$, to

$$\begin{gathered} \alpha(\alpha^{n-1} \beta^l) = \alpha^n \beta^l = \\ = -(a_1 \alpha^{n-1} + a_2 \alpha^{n-2} + \cdots + a_n)\beta^l = \\ = -a_1 \alpha^{n-1} \beta^l - a_2 \alpha^{n-2} \beta^l - \cdots - a_n \beta^l \in \mathrm{Lin}_{\mathbb{Q}}(Z) \end{gathered}$$

jest kombinacją liniową elementów zbioru Z[(2)];

[(2)] Tutaj $\mathrm{Lin}_{\mathbb{Q}}(Z) = \{a_1 z_1 + \cdots + a_n z_n \mid a_i \in \mathbb{Q},\ z_i \in Z,\ (i = 1, \ldots, n;\ n \in \mathbb{N}^*)\}$ jest domknięciem liniowym zbioru Z nad ciałem liczb wymiernych $\mathbb{Q}$.

- podobnie otrzymujemy

$$\beta(\alpha^k\beta^l) \in \mathrm{Lin}_{\mathbb{Q}}(Z).$$

Jeśli teraz wektor $\mathbf{z} = (z_1, z_2, \ldots, z_{nm}) \in \mathbb{C}^{nm}$, to $(\alpha - \beta)z_i \in \mathrm{Lin}_{\mathbb{Q}}(Z)$ dla każdego $i = 1, \ldots, nm$, a więc możemy zastosować lemat 6.3.2, na podstawie którego dostajemy, że $\alpha - \beta \in \mathbb{A}$.

c) Za pomocą podobnych rozumowań, jak w części *b*), otrzymujemy

$$\alpha\beta \in \mathrm{Lin}_{\mathbb{Q}}(Z)$$

i z lematu 6.3.2 wynika, że $\alpha\beta \in \mathbb{A}$.

Zatem $\mathbb{A}$ jest podciałem w ciele liczb zespolonych $\mathbb{C}$. Wnosimy, że $\mathbb{A}$ jest ciałem. □

* * *

■ **Liczby Liouville'a.** Najpierw następne

Twierdzenie 6.3.4 (Liouville'a[3])**.** *Dla dowolnej liczby algebraicznej $\theta \in \mathbb{R}$ stopnia $n \geqslant 2$ znajdzie się dodatnia liczba całkowita c (zależna od θ), taka że nierówność*

$$\left|\theta - \frac{p}{q}\right| > \frac{c}{q^n} \tag{6.3}$$

zachodzi dla każdego ułamka $\frac{p}{q} \in \mathbb{Q}$ (tutaj $p, q \in \mathbb{Z}$ oraz $q > 0$).

Dowód. Skoro $m_\theta \in \mathbb{Q}[X]$, to istnieje takie $A \in \mathbb{Z}$, że

$$A \cdot m_\theta = a_0X^n + a_1X^{n-1} + \cdots + a_n \in \mathbb{Z}[X].$$

Z twierdzenia Bezouta wnosimy, że $A \cdot m_\theta = (X - \theta)g$ dla pewnego wielomianu $g \in \mathbb{Z}[X]$ stopnia $n - 1$. Wtedy wartość

$$0 \neq A \cdot m_\theta\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \theta\right) g\left(\frac{p}{q}\right)$$

[3] Joseph Liouville (1809–1882)

oraz wartość bezwzględna

$$\left|A \cdot m_\theta\left(\frac{p}{q}\right)\right| = \left|\frac{p}{q} - \theta\right| \left|g\left(\frac{p}{q}\right)\right|.$$

Na tej podstawie

$$\left|A \cdot m_\theta\left(\frac{p}{q}\right)\right| = \frac{|a_0 p^n + a_1 p^{n-1} q + \cdots + a_n q^n|}{q^n} \geqslant \frac{1}{q^n}.$$

a) Załóżmy, że

$$\frac{p}{q} \in (\theta - 1, \theta + 1)$$

oraz $\frac{1}{a}$ jest największą wartością funkcji $|g|$ w przedziale $(\theta - 1, \theta + 1)$. Wtedy możemy ocenić wartości w taki sposób:

$$\begin{array}{rcl} \left|g\left(\frac{p}{q}\right)\right| & \leqslant & \frac{1}{a}, \\ \left|\theta - \frac{p}{q}\right| & = & \left|A \cdot m_\theta\left(\frac{p}{q}\right)\right| \cdot \frac{1}{\left|g\left(\frac{p}{q}\right)\right|} \geqslant \frac{a}{q^n}. \end{array}$$

b) Jeśli zaś $\frac{p}{q} \notin (\theta - 1, \theta + 1)$, to $\left|\theta - \frac{p}{q}\right| \geqslant 1$, a więc

$$\left|\theta - \frac{p}{q}\right| \geqslant \frac{1}{q^n}.$$

Teraz jeśli $C = \min\{1, a\}$, to (6.3) zachodzi. □

■ Liczba $\theta \in \mathbb{R}$ taka, że dla każdego naturalnego n istnieją liczby $p, q \in \mathbb{Z}$ takie, że $q > 1$ i zachodzą nierówności

$$0 < \left|\theta - \frac{p}{q}\right| < \frac{1}{q^n}, \tag{6.4}$$

jest nazywana *liczbą Liouville'a.*

Wniosek 6.3.5 (Liouville'a)**.** *Każda liczba rzeczywista postaci*

$$L_0 = \frac{a_1}{10^{1!}} + \frac{a_2}{10^{2!}} + \cdots + \frac{a_n}{10^{n!}} + \cdots,$$

gdzie $a_i \in \mathbb{Z}$ *oraz* $1 \leqslant a_i \leqslant 9$*, jest przestępna.*

Dowód. *a*) Wykażmy, że każda liczba Liouville'a L nie jest wymierna. Nie wprost. Niech $L = \frac{a}{b}$ dla pewnych $a, b \in \mathbb{N}^*$, gdzie $b > 1$. Załóżmy, że L różni się od liczby wymiernej $\frac{p}{q}$, gdzie $q > 1$. Skoro znajdzie się taka liczba naturalna m, że $2^{m-1} > b$, to

$$\left| L - \frac{p}{q} \right| = \left| \frac{a}{b} - \frac{p}{q} \right| \geqslant \frac{1}{qb} > \frac{1}{q2^{m-1}} \geqslant \frac{1}{q^m},$$

a to jest sprzeczne z definicją liczby Liouville'a.

b) Teraz przekonajmy się, że L_0 jest liczbą Liouville'a. Dla $n \in \mathbb{N}^*$ w jakości ułamka $\frac{p}{q}$ weźmy sumę

$$\frac{p}{q} = \sum_{i=1}^{n} \frac{a_i}{10^{i!}}.$$

Wtedy $q = 10^{n!}$ oraz

$$\left| L_0 - \frac{p}{q} \right| = \sum_{i=n+1}^{\infty} \frac{a_i}{10^{i!}} \leqslant \sum_{i=n+1}^{\infty} \frac{9}{10^{i!}} < \frac{9}{10^{(n+1)!}} \sum_{j=0}^{\infty} \frac{1}{10^j} \to \frac{10}{10^{(n+1)!}},$$

gdy $n \to \infty$, i na tej podstawie otrzymujemy

$$0 < \left| L_0 - \frac{p}{q} \right| \leqslant \frac{1}{10^{n!n}},$$

a więc warunek (6.4) zachodzi.

c) Liczba L_0 jest przestępna. Nie wprost. Załóżmy, ze liczba L_0 jest algebraiczna stopnia n. Wtedy $n > 1$. Na mocy twierdzenia 6.3.4 zachodzi (6.3). Istnieje taka liczba naturalna $m \geqslant n$, że stała $c > \frac{1}{2^{m-n}}$. Skoro liczba L_0 spełnia warunek (6.4), to

$$0 < \left| L_0 - \frac{p}{q} \right| < \frac{1}{q^n} \leqslant \frac{1}{2^{m-n} q^n} < \frac{c}{q^n} < \left| L_0 - \frac{p}{q} \right|,$$

co jest sprzeczne. □

Wniosek 6.3.6. *Liczby Liouville'a tworzą zbiór nieprzeliczalny.*

Przykłady 6.3.7.

(**1**) (**twierdzenie Lindemanna**[4]) Liczba $\pi \in \mathbb{R}$ (=połowa długości okręgu o promieniu 1) jest przestępna.

(**2**) (**twierdzenie Hermite'a**[5]) Liczba Eulera $e \in \mathbb{R}$ jest przestępna.

(**3**) Pozostaje `otwarty` (do dzisiaj) problem: *udowodnić, że liczby* $e \pm \pi$ *są przestępne.* Możemy udowodnić tylko taki

Lemat 6.3.8. *Co najmniej jedna z liczb* $e + \pi$, $e - \pi$ *jest przestępna.*

Dowód. Nie wprost. Jeśli obie liczby $e + \pi$ oraz $e - \pi$ są algebraiczne, to ich suma $(e + \pi) + (e - \pi) = 2e$ jest liczbą algebraiczną na podstawie twierdzenia 6.3.3, a to jest sprzeczne z twierdzeniem Hermite'a. □

Lemat 6.3.9. *Każdy przedział otwarty* $(a, b) \subseteq \mathbb{R}$ $(a < b)$ *jest równoliczny ze zbiorem liczb rzeczywistych* $\mathbb{R}$.

Dowód. Rzeczywiście, mamy takie dwa odwzorowania bijektywne (sprawdzić samodzielnie):

$$g : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \ni x \mapsto \operatorname{tg} x \in \mathbb{R}$$

oraz

$$f : (a, b) \ni x \mapsto \frac{\pi}{b-a}(x - a) - \frac{\pi}{2} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

a więc ich złożenie $g \circ f : (a, b) \to \mathbb{R}$ jest bijekcją. □

Twierdzenie 6.3.10 (Cantora)**.** *Przedział otwarty* $(0, 1)$ *(a więc i zbiór liczb rzeczywistych* $\mathbb{R}$*) jest nieprzeliczalny.*

Dowód. Nie wprost. Załóżmy, że przedział $(0, 1)$ jest przeliczalny, a więc mamy ciąg $(0, 1) = \{x_0, x_1, \ldots, x_n, \ldots\} = \{x_n\}_{n\in\mathbb{N}}$. Jak wiadomo, każda z liczb tego ciągu ma rozwinięcie dziesiętne postaci:

$$\begin{aligned}
x_0 &= 0, x_{0,0}x_{0,1}x_{0,2}x_{0,3}x_{0,4}x_{0,5}\ldots \\
x_1 &= 0, x_{1,0}x_{1,1}x_{1,2}x_{1,3}x_{1,4}x_{1,5}\ldots \\
x_2 &= 0, x_{2,0}x_{2,1}x_{2,2}x_{2,3}x_{2,4}x_{2,5}\ldots \\
&\vdots \\
x_n &= 0, x_{n,0}x_{n,1}x_{n,2}x_{n,3}x_{n,4}x_{n,5}\ldots \\
&\vdots
\end{aligned}$$

[4] Ferdinand Lindemann (1852–1939)

[5] Charles Hermite (1822–1901)

Skoro istnieją takie cyfry a_i, że

$$a_i \notin \{0,\ 9\} \quad \text{oraz} \quad a_i \neq x_{i,i},$$

to liczba

$$z = 0, a_0 a_1 a_3 a_4 a_5 \cdots \in (0, 1)$$

różni się od każdej liczby x_i i mamy sprzeczność. Zatem zbiór $(0, 1)$ nie jest przeliczalny. □

Twierdzenie 6.3.11 (Cantora)**.** *Zbiór liczb algebraicznych $\mathbb{A}$ jest przeliczalny.*

Dowód. (Szkic) *a*) Najpierw wykażmy, że zbiór wszystkich wielomianów $f \in \mathbb{Z}[X]$ jednej zmiennej X jest przeliczalny. Istotnie zbiór wielomianów stopnia t jest równoliczny ze zbiorem

$$\underbrace{\mathbb{Z} \times \cdots \times \mathbb{Z}}_{t+1 \text{ czynników}},$$

co powoduje, że jest przeliczalny. To znaczy, że dla każdego $t \in \mathbb{N}$ wszystkie wielomiany stopnia t możemy przedstawić w postaci ciągu

$$f_{1,t}, f_{2,t}, \ldots, f_{n,t}, \ldots.$$

Wtedy wszystkie wielomiany ze zbioru $\mathbb{Z}[X]$ możemy przedstawić w postaci takiej tablicy (nieskończonej w prawo i w dół):

$$\begin{array}{cccccccc}
f_{1,1} & f_{2,1} & f_{3,1} & f_{4,1} & \cdots & f_{n,1} & \cdots \\
f_{1,2} & f_{2,2} & f_{3,2} & f_{4,2} & \cdots & f_{n,2} & \cdots \\
f_{1,3} & f_{2,3} & f_{3,3} & f_{4,3} & \cdots & f_{n,3} & \cdots \\
f_{1,4} & f_{2,4} & f_{3,4} & f_{4,4} & \cdots & f_{n,4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}.$$

Za pomocą tej tablicy konstruujemy „nowy" ciąg w taki sposób:

- najpierw wielomiany o sumie indeksów równej 2;
- dalej wielomiany o sumie indeksów równej 3;
- dalej wielomiany o sumie indeksów równej 4;
- itd.

Otrzymujemy ciąg

$$f_{1,1}, f_{1,2}, f_{2,1}, f_{1,3}, f_{2,2}, f_{3,1}, f_{1,4}, f_{2,3}, f_{3,2}, f_{4,1}, \ldots$$

Każdy z wielomianów $f_{i,j}$ tego ciągu ma skończoną liczbę pierwiastków.

b) Wypisując po kolei pierwiastki zespolone tych wielomianów (najpierw pierwiastki pierwszego wielomianu, dalej pierwiastki drugiego wielomianu itd., pomijając pierwiastki, które napotkaliśmy wcześniej), otrzymujemy ciąg pierwiastków zespolonych wszystkich wielomianów ze zbioru $\mathbb{Z}[X]$. Skoro zbiory pierwiastków

$$\{\alpha \in \mathbb{C} \mid f(\alpha) = 0, \text{ gdzie } 0 \neq f \in \mathbb{Z}[X]\}$$

oraz

$$\{\beta \in \mathbb{C} \mid g(\beta) = 0, \text{ gdzie } 0 \neq g \in \mathbb{Q}[X]\}$$

są równe, to zbiór $\mathbb{A}$ jest przeliczalny. □

Ponieważ zbiór liczb rzeczywistych jest nieprzeliczalny, to G. Cantor w 1874 r. też otrzymał taki

Wniosek 6.3.12. *Liczby przestępne istnieją.*

* * *

■ **Domkniętość algebraiczna.** Ciało $\mathbb{F}$ jest nazywane *algebraicznie domkniętym*, jeśli każdy wielomian niezerowy $f \in \mathbb{F}[X]$ ma pierwiastek w tym ciele. Jak wiadomo, ciało liczb zespolonych $\mathbb{C}$ jest algebraicznie domknięte. Zachodzi też takie

Twierdzenie 6.3.13. *Ciało liczb algebraicznych $\mathbb{A}$ jest algebraicznie domknięte.*

Dowód. Niech $f = X^n + a_1X^{n-1} + \cdots + a_n \in \mathbb{A}[X]$ będzie unormowanym wielomianem niezerowym. Wtedy $f \in \mathbb{C}[X]$, a więc ma pewien pierwiastek zespolony θ. Jeśli $\theta = 0$, to $\theta \in \mathbb{A}$. Dlatego dalej załóżmy, że $\theta \neq 0$. Skoro $a_i \in \mathbb{A}$ $(i = 1, \ldots, n)$, to istnieje wielomian

$$g_i = X^{m_i} + b_{i,1}X^{m_i-1} + \cdots + b_{i,m_i} \in \mathbb{Q}[X]$$

o współczynnikach wymiernych takich, że

$$g_i(a_i) = 0.$$

Rozpatrzmy zbiór

$$\begin{aligned} Z = \ & \{\theta^j a_1^{j_1} a_2^{j_2} \cdots a_n^{j_n} \mid i = 1, \ldots, n;\ j_i = 0, 1, \ldots, m_i - 1; \\ & j = 0, 1, \ldots, n-1\} = \{z_1, z_2, \ldots, z_{nm_1m_2\cdots m_n}\}, \end{aligned}$$

gdzie czynnik $a_i^0 = 1$, gdy współczynnik $a_i = 0$ jest zerowy. Podobnie jak w twierdzeniu 6.3.3 otrzymujemy, że $a_i z_i \in \mathrm{Lin}_{\mathbb{Q}}(Z)$, a więc wektor $\mathbf{z} = (z_1, \ldots, z_{nm_1m_2\cdots m_n}) \in \mathbb{C}^{nm_1m_2\cdots m_n}$ oraz θ spełniają warunek (6.2) dla pewnej macierzy kwadratowej $A \in M_{nm_1m_2\cdots m_n}(\mathbb{Q})$. Na mocy lematu 6.3.2 wnosimy, że $\theta \in \mathbb{A}$. □

Ćwiczenia 6.3.14.

(**1**) Udowodnić, że algebraicznymi są liczby:
(*a*) $1 + \sqrt{3}$;
(*b*) $\sqrt{3} - \sqrt{5}$;
(*c*) $\sqrt{3 - \sqrt{7}}$;
(*d*) $2 - i\sqrt{7}$;
(*e*) $\sqrt{5} + i\sqrt{11}$;
(*f*) $i\sqrt{7} - \sqrt{5}$.
(**2**) Sprawdzić, czy liczba

$$1 + \frac{1}{3!} + \frac{1}{5!} + \frac{1}{7!} + \cdots$$

jest przestępna.

Uwagi. Istnienie liczb przestępnych (patrz twierdzenie 6.3.11) wykazał matematyk francuski J. Liouville w 1844 r. Praca G. Cantora z 1874 r. (gdzie, w szczególności, udowodniono twierdzenie 6.3.10) zapoczątkowała współczesną teorię mnogości. W tej pracy G. Cantor dowiódł też, że zbiory liczb naturalnych $\mathbb{N}$ i liczb rzeczywistych $\mathbb{R}$ nie są równoliczne.

To, że ciało liczb zespolonych $\mathbb{C}$ jest algebraicznie domknięte (a to jest twierdzenie, które jest nazywane podstawowym twierdzeniem algebry), zostało bez dowodu sformułowane przez A. Girarda[(6)] w 1629 r. i udowodnione przez C. Gaussa w 1799 r.

Wniosek 6.3.5 był udowodniony przez Liouville'a w 1844 r.

[(6)] Albert Girard (1595–1632)

6.4. Ciało rozkładu wielomianu

Twierdzenie 6.4.1 (Kroneckera-Artina[7]). *Dla każdego wielomianu niezerowego $f \in \mathbb{F}[X]$ nad ciałem $\mathbb{F}$ istnieje ciało $\mathbb{E}$ będące rozszerzeniem ciała $\mathbb{F}$, w którym f ma pierwiastek.*

Dowód. Skoro f jest iloczynem wielomianów nieprzywiedlnych nad ciałem $\mathbb{F}$, to możemy (bez ograniczenia ogólności dowodu) założyć, że wielomian

$$f = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n \in \mathbb{F}[X]$$

jest nieprzywiedlny nad $\mathbb{F}$. Wtedy pierścień ilorazowy $\mathbb{E} = \mathbb{F}[X]/\langle f\rangle$ jest ciałem na podstawie wniosku 4.5.7. Ponieważ warstwa

$$\overline{f} = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n + \langle f\rangle = \overline{0}$$

jest zerowa, to

$$f(\overline{X}) = \overline{a}_0\overline{X}^n + \overline{a}_1\overline{X}^{n-1} + \cdots + \overline{a}_{n-1}\overline{X} + \overline{a}_n = \overline{0}, \tag{6.5}$$

gdzie $\overline{X} = X + \langle f\rangle \in \mathbb{E}$. Odwzorowanie

$$\varphi : \mathbb{F} \ni a \mapsto a + \langle f\rangle \in \mathbb{E}$$

jest izomorfizmem ciał (przekonać się samodzielnie), a więc możemy utożsamiać każdy element $\overline{a} = a + \langle f\rangle$ ciała $\mathbb{E}$ z elementem a ciała $\mathbb{F}$. W wyniku tego (6.5) możemy przepisać w postaci

$$f(\overline{X}) = a_0\overline{X}^n + a_1\overline{X}^{n-1} + \cdots + a_{n-1}\overline{X} + a_n = 0,$$

a zatem element $\overline{X} \in \mathbb{E}$ jest pierwiastkiem wielomianu $f \in \mathbb{F}[X]$. □

■ Niech $\mathbb{F}$ będzie ciałem oraz $X_1, X_2, \ldots, X_n$ będą niewiadomymi ($n \geqslant 2$). Wówczas przyjmujemy, że:

- $$\mathbb{F}[X_1, X_2] = (\mathbb{F}[X_1])\,[X_2]$$

  jest zbiorem wielomianów jednej zmiennej X_2 o współczynnikach z pierścienia $\mathbb{F}[X_1]$;

[7] Emil Artin (1898–1962)

•
$$\mathbb{F}[X_1,\ldots,X_{n-1},X_n] = (\mathbb{F}[X_1,\ldots,X_{n-1}])\,[X_n]$$
jest zbiorem wielomianów jednej zmiennej X_n o współczynnikach z pierścienia $\mathbb{F}[X_1,\ldots,X_{n-1}]$.

Wtedy $\mathbb{F}[X_1,X_2,\ldots,X_n]$ jest pierścieniem z jednością. Podobnie jak w rozdziale 5 możemy wprowadzić *ciało funkcji wymiernych*, zależnych od zmiennych $X_1,X_2,\ldots,X_n$, a mianowicie

$$\mathbb{F}\,(X_1,X_2,\ldots,X_n) = \left\{ \frac{f}{g} \mid f,g \in \mathbb{F}[X_1,X_2,\ldots,X_n] \text{ oraz } g \neq 0, \right\}.$$

Lemat 6.4.2. *Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem ciał $\mathbb{F}$ i $\mathbb{E}$ oraz elementy $\theta_1,\theta_2,\ldots,\theta_n \in \mathbb{E}$. Wtedy zachodzą następujące własności:*

(1) $\mathbb{F}[\theta_1,\theta_2,\ldots,\theta_n] = \{f(\theta_1,\theta_2,\ldots,\theta_n) \mid f \in \mathbb{F}[X_1,X_2,\ldots,X_n]\}$ *jest podpierścieniem w $\mathbb{E}$;*

(2)
$$\mathbb{F}(\theta_1,\theta_2,\ldots,\theta_n) = \left\{ \frac{f(\theta_1,\theta_2,\ldots,\theta_n)}{g(\theta_1,\theta_2,\ldots,\theta_n)} \mid f,g \in \mathbb{F}[X_1,X_2,\ldots,X_n] \text{ oraz} \right.$$
$$\left. g(\theta_1,\theta_2,\ldots,\theta_n) \neq 0, \right\}$$
jest podciałem w $\mathbb{E}$;

(3) $\mathbb{F}[\theta_1,\theta_2,\ldots,\theta_n] \subseteq \mathbb{F}(\theta_1,\theta_2,\ldots,\theta_n)$.

■ *Ciałem rozkładu wielomianu $f \in \mathbb{F}[X]$* stopnia $n \geqslant 1$ nad ciałem $\mathbb{F}$ jest nazywane ciało $\mathbb{L}$, które jest rozszerzeniem ciała $\mathbb{F}$ takim, że są spełnione dwa warunki:

1) wielomian f rozkłada się na czynniki liniowe nad ciałem $\mathbb{L}$, czyli istnieją takie elementy $\theta_1,\theta_2,\ldots,\theta_n \in \mathbb{L}$, że
$$f = a_0(X-\theta_1)(X-\theta_2)\cdots(X-\theta_n),$$
gdzie a_0 jest współczynnikiem najwyższym wielomianu f;

2)
$$\mathbb{L} = \mathbb{F}(\theta_1,\theta_2,\ldots,\theta_n).$$

Twierdzenie 6.4.3 (o istnieniu ciała rozkładu wielomianu)**.** *Dla każdego ciała* $\mathbb{F}$ *i każdego wielomianu* $f \in \mathbb{F}[X]$ *stopnia* $n \geqslant 1$ *istnieje rozszerzenie* $\mathbb{L}$ *ciała* $\mathbb{F}$*, które jest ciałem rozkładu wielomianu* f*.*

Dowód. Stosujemy indukcję względem stopnia n. Jeśli $n = 1$, to $f = aX + b$ dla pewnych $a, b \in \mathbb{F}$, gdzie $a \neq 0$, a więc

$$f = a(X - \theta),$$

gdzie $\theta = -a^{-1}b \in \mathbb{F}$. Zatem $\mathbb{L} = \mathbb{F}$.

Załóżmy, że teza twierdzenia zachodzi dla wielomianów stopni $n-1$ z pierścienia $\mathbb{F}[X]$. Niech $f \in \mathbb{F}[X]$ oraz $\deg f = n > 1$. Na podstawie twierdzenia 6.4.1 istnieje ciało $\mathbb{E}$, które jest rozszerzeniem ciała $\mathbb{F}$ oraz $f(\theta_1) = 0$ dla pewnego $\theta_1 \in \mathbb{E}$. To implikuje, że

$$f = (X - \theta_1) \cdot g$$

dla pewnego wielomianu $g \in \mathbb{E}[X]$. Na mocy twierdzenia 6.2.3 wnosimy, że $\mathbb{E} = \mathbb{F}(\theta_1)$. Skoro stopień $\deg g = n - 1$, to na podstawie założenia indukcyjnego istnieje ciało rozkładu $\mathbb{L} = \mathbb{E}(\theta_2, \ldots, \theta_n)$ wielomianu g, a więc

$$g = a_0(X - \theta_1) \cdots (X - \theta_n),$$

gdzie a_0 jest współczynnikiem najwyższym wielomianu g (a zatem i f). Wnosimy, że $\mathbb{L}$ jest ciałem rozkładu wielomianu f nad ciałem $\mathbb{F}$. □

Wniosek 6.4.4. *Niech* $f \in \mathbb{F}[X]$ *będzie wielomianem stopnia* $n \geqslant 1$*. Jeśli* $\mathbb{L}$ *jest ciałem jego rozkładu, to stopień rozszerzenia* $|\mathbb{L} : \mathbb{F}|$ *spełnia nierówność*

$$|\mathbb{L} : \mathbb{F}| \leqslant n!$$

oraz jest skończony.

Dowód. Stosujemy indukcję względem n. Jeśli $n = 1$, to $\mathbb{L} = \mathbb{F}(\theta_1)$ i teza wynika na podstawie twierdzenia 6.2.3.

Załóżmy, że $n > 1$ oraz dla wielomianów $g \in \mathbb{F}[X]$ stopni $n - 1$ zachodzi podobna nierówność. Skoro na podstawie twierdzenia 6.4.3 dla każdego wielomianu $f \in \mathbb{F}[X]$ stopnia n istnieje ciało rozkładu $\mathbb{L} = \mathbb{F}(\theta_1, \theta_2, \ldots, \theta_n)$, to w wyniku twierdzenia Bezouta

$$f = (X - \theta_1) \cdot g$$

dla wielomianu

$$g = a_0(X - \theta_2) \cdots (X - \theta_n) \in (\mathbb{F}(\theta_1))[X],$$

gdzie a_0 jest współczynnikiem najwyższym wielomianu f. Ponieważ $\mathbb{L}$ jest ciałem rozkładu wielomianu g nad ciałem $\mathbb{F}(\theta_1)$, to biorąc pod uwagę założenie indukcyjne na podstawie twierdzenia 6.1.1, otrzymujemy

$$|\mathbb{L} : \mathbb{F}| = |\mathbb{L} : \mathbb{F}(\theta_1)| \cdot |\mathbb{F}(\theta_1) : \mathbb{F}| \leqslant (n-1)! \cdot n = n!.$$

□

Zaznaczmy bez dowodu takie

Twierdzenie 6.4.5. *Niech $f \in \mathbb{F}[X]$ będzie wielomianem stopnia $n \geqslant 1$. Każde dwa ciała rozkładu wielomianu f nad ciałem $\mathbb{F}$ są izomorficzne.*

Przykłady 6.4.6.

(**1**) Ciało $\mathbb{Q}[\sqrt{5}]$ (które jest ciałem rozkładu wielomianu $f = X^2 - 5 \in \mathbb{Q}[X]$) i ciało $\mathbb{Q}[\sqrt{7}]$ (które jest ciałem rozkładu wielomianu $f = X^2 - 7 \in \mathbb{Q}[X]$) nie są izomorficzne. Wykażmy to nie wprost. Jeśli założyć, że istnieje pewien izomorfizm ciał $\varphi : \mathbb{Q}[\sqrt{5}] \to \mathbb{Q}[\sqrt{7}]$, to

$$\varphi(1) = a + b\sqrt{7}$$

dla pewnych współczynników $a, b \in \mathbb{Q}$. Wtedy

$$a + b\sqrt{7} = \varphi(1) = \varphi(1)^2 = (a^2 + 7b^2) + 2ab\sqrt{7},$$

a więc dostajemy układ równań

$$\begin{cases} a^2 + 7b^2 &= a, \\ 2ab &= b. \end{cases}$$

Jeśli założyć, że $b \neq 0$, to otrzymujemy $a = \frac{1}{2}$ i wtedy

$$\mathbb{Q} \ni b = \pm\frac{1}{\sqrt{28}} \notin \mathbb{Q},$$

co nie jest możliwe. Zatem $b = 0$, a więc $a(a-1) = 0$.

Niech $a = 0$. Wtedy $\varphi(1) = 0$, co powoduje, że $\varphi(\frac{u}{v}) = 0$ dla dowolnych $u, v \in \mathbb{Z}$. A z tego wynika, że $\varphi = 0$ jest zerowe. Ponieważ wobec założenia φ jest izomorfizmem, to $a = 1$, co daje $\varphi(q) = q$ dla każdego $q \in \mathbb{Q}$ (przekonać się samodzielnie). Skoro $\varphi(\sqrt{5}) = x + y\sqrt{7}$ dla pewnych $x, y \in \mathbb{Q}$, to

$$5 = \varphi(5) = \varphi(\sqrt{5})^2 = (x + y\sqrt{7})^2 = (x^2 + 7y^2) + 2xy\sqrt{7},$$

a więc mamy układ

$$\begin{cases} x^2 + 7y^2 &= 5, \\ 2xy &= 0. \end{cases}$$

Ponieważ $\pm\sqrt{5} \notin \mathbb{Q}$, to $y \neq 0$, co implikuje, że $x = 0$. Wtedy

$$\mathbb{Q} \ni y = \pm\sqrt{\frac{5}{7}} \notin \mathbb{Q}.$$

Otrzymaliśmy sprzeczność. To znaczy, że ciała $\mathbb{Q}[\sqrt{5}]$ oraz $\mathbb{Q}[\sqrt{7}]$ nie są izomorficzne.

(**2**) Skoro

$$X^2 + 1 = (X - i)(X + i) \in \mathbb{C}[X],$$

to ciałem rozkładu wielomianu $X^2 + 1$ nad ciałem liczb wymiernych $\mathbb{Q}$ jest ciało $\mathbb{Q}(-i, i) = \mathbb{Q}(i)$.

Podobnie

$$X^2 - 11 = (X - \sqrt{11})(X + \sqrt{11}) \in \mathbb{R}[X],$$

a więc ciałem rozkładu wielomianu X^2-11 nad ciałem liczb wymiernych $\mathbb{Q}$ jest ciało $\mathbb{Q}(-\sqrt{11}, \sqrt{11}) = \mathbb{Q}(\sqrt{11})$.

Ćwiczenia 6.4.7.

(**1**) Znaleźć stopień ciała rozkładu wielomianu f nad ciałem F, jeśli:
(a) $f = X^2 - 5$ oraz $F = \mathbb{Q}$;
(b) $f = X^2 - 2$ oraz $F = \mathbb{F}_3$;
(c) $f = X^3 - 5$ oraz $F = \mathbb{Q}$;
(d) $f = X^3 - 7$ oraz $F = \mathbb{R}$;
(e) $f = X^4 - 3$ oraz $F = \mathbb{Q}$;
(f) $f = X^4 + X^2 + 1$ oraz $F = \mathbb{Q}$;
(g) $f = X^4 + X^2 + 1$ oraz $F = \mathbb{F}_2$;
(h) $f = X^6 - 9$ oraz $F = \mathbb{Q}$.
(**2**) Udowodnić, że:
(a) $\sqrt{2} + \sqrt{3}$ jest pierwiastkiem wielomianu $X^4 - 10X^2 + 1 \in \mathbb{Q}[X]$;
(b) wielomian $X^4 - 10X^2 + 1$ jest wielomianem minimalnym elementu $\sqrt{2} + \sqrt{3}$ nad ciałem $\mathbb{Q}$.
(**3**) Znaleźć wszystkie wielomiany nierozkładalne stopnia n nad ciałem F, jeśli:
(a) $n = 2$ oraz $F = \mathbb{F}_2$;
(b) $n = 3$ oraz $F = \mathbb{F}_2$;
(c) $n = 4$ oraz $F = \mathbb{F}_2$;
(d) $n = 2$ oraz $F = \mathbb{F}_3$;
(e) $n = 3$ oraz $F = \mathbb{F}_3$.
(**4**) Znaleźć stopień $|F : \mathbb{Q}|$ rozszerzenia $\mathbb{Q} \leqslant F$ ciał, jeśli:
(a) $F = \mathbb{Q}(\sqrt[3]{2} + 2\sqrt[3]{4})$;
(b) $F = \mathbb{Q}(\frac{\sqrt{2}}{\sqrt[3]{2}})$;
(c) $F = \mathbb{Q}(\sqrt{2} - 3i)$;
(d) $F = \mathbb{Q}(1 - \sqrt{7})$;
(e) $F = \mathbb{Q}(\sqrt{3}, \sqrt[4]{3})$;
(f) $F = \mathbb{Q}(\sqrt[3]{2}, \sqrt{2})$;
(g) $F = \mathbb{Q}(\sqrt[4]{2}, i)$.
(**5**) Udowodnić, że jeśli $0 \neq a \in \mathbb{Z}_p$, gdzie p jest liczbą pierwszą, to wielomian $X^p - X - a$ jest nierozkładalny nad ciałem $\mathbb{Z}_p$.

Uwagi. Koncepcję ciała jako pierwsi stosowali jeszcze N. Abel i E. Galois w ich badaniach rozwiązywalności równań algebraicznych. W 1893 r.

H. Weber jako pierwszy zdefiniował aksjomatycznie pojęcie ciała. E. Steinitz[8] założył podstawy współczesnej teorii ciał w jednej ze swoich prac z 1910 r. W tej pracy z algebraicznego punktu widzenia zostało wprowadzone pojęcie różniczkowania funkcji wymiernych.

[8] Ernst Steinitz (1871–1928)

6.5. Ciała algebraicznie domknięte

■ Jak wiemy, ciało liczb algebraicznych $\mathbb{A}$ (patrz twierdzenie 7.2.1) oraz ciało liczb zespolonych $\mathbb{C}$ są algebraicznie domknięte.

Twierdzenie 6.5.1. *Następujące własności są równoważne:*
(1) *ciało* $\mathbb{F}$ *jest algebraicznie domknięte;*
(2) *każdy wielomian* $f \in \mathbb{F}[X]$ *stopnia* $n \geqslant 1$ *rozkłada się na czynniki liniowe nad ciałem* $\mathbb{F}$.

Dowód. (2) $\Rightarrow$ (1) Jest oczywiste.

(1) $\Rightarrow$ (2) Z definicji wynika, że f ma pewien pierwiastek $\theta_1 \in \mathbb{F}$ i na mocy twierdzenia Bezouta

$$f = (X - \theta_1) \cdot g$$

dla pewnego wielomianu $g \in \mathbb{F}[X]$. Skoro $\deg d = n - 1$, to łatwo otrzymujemy tezę, stosując rozumowania indukcyjne. □

Wniosek 6.5.2. *Ciało* $\mathbb{F}$ *jest algebraicznie domknięte wtedy i tylko wtedy, gdy każdy wielomian* $f \in \mathbb{F}[X]$, *który jest nieprzywiedlny nad ciałem* $\mathbb{F}$, *ma stopień równy* 1.

■ Ciało $\overline{\mathbb{F}}$ jest nazywane *domknięciem algebraicznym* ciała $\mathbb{F}$, jeśli $\overline{\mathbb{F}}$ jest algebraiczne nad $\mathbb{F}$ oraz każdy wielomian $f \in \mathbb{F}[X]$ stopnia $n \geqslant 1$ rozkłada się na czynniki liniowe nad ciałem $\overline{\mathbb{F}}$, czyli

$$f = a_0(X - \theta_1)(X - \theta_2) \cdots (X - \theta_n),$$

gdzie $\theta_1, \theta_2, \ldots, \theta_n \in \overline{\mathbb{F}}$ oraz a_0 jest współczynnikiem najwyższym wielomianu f.

Przykłady 6.5.3.

(**1**) Ciało liczb zespolonych $\mathbb{C}$ jest domknięciem algebraicznym ciała liczb rzeczywistych $\mathbb{R}$.

(**2**) Ciało liczb algebraicznych $\mathbb{A}$ jest domknięciem algebraicznym ciała liczb wymiernych $\mathbb{Q}$.

Lemat 6.5.4. *Każde ciało algebraicznie domknięte* $\mathbb{F}$ *jest nieskończone.*

Dowód. Nie wprost. Jeśli ciało algebraicznie domknięte $\mathbb{F} = \{\theta_1, \theta_2, \ldots, \theta_n\}$ jest skończone, to wielomian

$$f = (X - \theta_1)(X - \theta_2) \cdots (X - \theta_n) + 1 \in \mathbb{F}[X]$$

nie posiada w ciele $\mathbb{F}$ żadnego pierwiastka, a to jest sprzeczne z definicją ciała algebraicznie domkniętego. □

Lemat 6.5.5. *Domknięcie algebraiczne $\overline{\mathbb{F}}$ ciała $\mathbb{F}$ jest algebraicznie domknięte.*

Dowód. Jak wiadomo z twierdzenia 6.4.1, wielomian $f \in \overline{\mathbb{F}}[X]$ ma pierwiastek θ w pewnym rozszerzeniu ciała $\overline{\mathbb{F}}$. Wtedy ciało $\overline{\mathbb{F}}(\theta)$ jest rozszerzeniem algebraicznym ciała $\mathbb{F}$ i w konsekwencji na podstawie twierdzenia 6.2.5 element θ jest algebraiczny nad ciałem $\mathbb{F}$. To implikuje, że $\theta \in \overline{\mathbb{F}}$. □

Twierdzenie 6.5.6. *Każde ciało $\mathbb{F}$ wkłada się w ciało algebraicznie domknięte.*

Dowód. (Szkic) Załóżmy, że $f \in \mathbb{F}[X]$ jest wielomianem unormowanym stopnia $n \geqslant 1$ oraz X_f jest symbolem. Rozpatrzmy pierścień wielomianów

$$R = \mathbb{F}[\ldots, X_f, \ldots]$$

zależny od zmiennych ze zbioru

$$A = \{X_f \mid f \in \mathbb{F}[X] \text{ jest unormowany}\}$$

Wtedy

$$I = \sum_{X_f \in A} X_f R$$

jest jego ideałem. Jeśli założyć, że I nie jest właściwy w R, to znajdą się takie wielomiany $f_i \in \mathbb{F}[X]$ oraz $g_i \in R \ \ (i = 1, \ldots, n)$, że

$$\sum_{i=1}^{n} g_i f_i(X_{f_i}) = 1. \tag{6.6}$$

Jeśli oznaczymy

$$X_i = X_{f_i} \ \ (i = 1, \ldots, n),$$

to ewentualnie istnieją jeszcze pewne zmienne (które oznaczamy przez $X_{n+1}, \ldots, X_k$), od których zależą wielomiany $g_1, \ldots, g_n$. Zatem (6.6) możemy przepisać w postaci

$$\sum_{i=1}^{n} g_i(X_1, \ldots, X_k) f_i(X_i) = 1. \tag{6.7}$$

Istnieje takie rozszerzenie $\mathbb{F}_1$ ciała $\mathbb{F}$, które posiada elementy $\theta_1, \theta_2, \ldots, \theta_n$ takie, że $f_i(\theta_i) = 0 \quad (i = 1, \ldots, n)$. Podstawiając wartości

$$\begin{aligned} X_i &= \theta_i \ (i = 1, \ldots, n), \\ X_j &= 0 \ (i = n+1, \ldots, k) \end{aligned}$$

w (6.7), otrzymujemy sprzeczność. Zatem I jest ideałem właściwym w R. Stosując twierdzenie 4.5.4, wnosimy, że I zawiera się w pewnym ideale maksymalnym M pierścienia R. Wtedy pierścień ilorazowy $\mathbb{F}_1 = R/M$ jest ciałem, przy czym ciało $\mathbb{F}$ wkłada się w ciało $\mathbb{F}_1$. Skoro

$$f(X_f) \in I \leqslant M,$$

to obraz $X_f + M$ jest pierwiastkiem wielomianu f należącym do ciała $\mathbb{F}_1$. To znaczy, że każdy wielomian $f \in \mathbb{F}[X]$ ma pierwiastek w ciele $\mathbb{F}_1$. Stosując rozumowania indukcyjne, możemy skonstruować łańcuch rozszerzeń ciał

$$\mathbb{F} = \mathbb{F}_0 \leqslant \mathbb{F}_1 \leqslant \mathbb{F}_2 \leqslant \cdots \leqslant \mathbb{F}_m \leqslant \mathbb{F}_{m+1} \leqslant \cdots$$

takich, że każdy wielomian $f \in \mathbb{F}_m$ ma pierwiastek w ciele $\mathbb{F}_{m+1}$ $(m \in \mathbb{N})$. Wtedy suma mnogościowa

$$\mathbb{E} = \bigcup_{m \in \mathbb{N}} \mathbb{F}_m$$

jest ciałem zawierającym ciało $\mathbb{F}$. Jeśli teraz

$$g = a_o X^s + a_1 X^{s-1} + \cdots + a_s \in \mathbb{E}[X],$$

to istnieje taki indeks l, że wszystkie współczynniki $a_0, a_1, \ldots, a_s \in \mathbb{F}_l$. Wtedy $g \in \mathbb{F}_l[X]$, a więc wielomian g ma pierwiastek w ciele $\mathbb{F}_{l+1} \leqslant \mathbb{E}$. Wnosimy zatem, że ciało $\mathbb{E}$ jest domknięte algebraicznie. $\square$

Twierdzenie 6.5.7. *Niech $\mathbb{F}$ będzie podciałem ciała $\mathbb{E}$. Wtedy zachodzą następujące własności:*

(1) *jeśli ciało $\mathbb{E}$ jest algebraicznie domknięte, to zbiór*

$$\overline{\mathbb{F}} = \{\theta \in \mathbb{E} \mid \theta \text{ jest algebraiczny nad ciałem } \mathbb{F}\}$$

jest domknięciem algebraicznym ciała $\mathbb{F}$;

(2) *dla każdego ciała $\mathbb{F}$ istnieje jego domknięcie algebraiczne $\overline{\mathbb{F}}$;*

(3) *domknięcie algebraiczne $\overline{\mathbb{F}}$ ciała $\mathbb{F}$ jest jedyne (z dokładnością do izomorfizmu).*

Dowód. (1) Zbiór $\overline{\mathbb{F}}$ jest ciałem algebraicznym nad ciałem $\mathbb{F}$ i każdy wielomian $f \in \mathbb{F}[X]$ stopnia $n \geqslant 1$ rozkłada się na czynniki liniowe nad ciałem $\mathbb{E}$, czyli

$$f = a_0(X - \theta_1) \cdots (X - \theta_n),$$

gdzie a_0 jest współczynnikiem najwyższym wielomianu f. Wtedy $\theta_i \in \overline{\mathbb{F}}$ $(i = 1, \ldots, n)$. Zatem f rozkłada się na czynniki liniowe nad ciałem $\overline{\mathbb{F}}$, a więc ciało $\overline{\mathbb{F}}$ jest domknięciem algebraicznym ciała $\mathbb{F}$.

(2) Wynika z części (1) oraz twierdzenia 6.5.6.

(3) Zostawiamy bez dowodu. □

* * *

■ **Twierdzenie o elemencie prymitywnym.** Będziemy mówić, że rozszerzenie ciał $\mathbb{F} \leqslant \mathbb{E}$ jest *rozdzielcze*, jeśli dla każdego elementu $\theta \in \mathbb{E}$, który jest algebraiczny nad ciałem $\mathbb{F}$, jego wielomian minimalny $m_\theta \in \mathbb{F}[X]$ nie ma pierwiastków krotnych w domknięciu algebraicznym $\overline{\mathbb{F}}$ ciała $\mathbb{F}$.

Twierdzenie 6.5.8 (Abela)**.** *Niech $\mathbb{F} \leqslant \mathbb{E}$ będzie rozszerzeniem rozdzielczym ciał $\mathbb{F}$, $\mathbb{E}$ oraz $\alpha, \beta \in \mathbb{E}$ będą elementami algebraicznymi nad ciałem $\mathbb{F}$. Wtedy istnieje element $\theta \in \mathbb{E}$ taki, że*

$$\mathbb{F}(\alpha, \beta) = \mathbb{F}(\theta)$$

(element θ jest nazywany *elementem prymitywnym* tego rozszerzenia).

Dowód. 1) Załóżmy, że ciało $\mathbb{E}$ jest nieskończone oraz $\overline{\mathbb{F}}$ jest domknięciem algebraicznym ciała $\mathbb{F}$ takim, że $\mathbb{E} \leqslant \overline{\mathbb{F}}$. Niech

$$\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\} \text{ oraz } \mathcal{B} = \{\beta_1, \ldots, \beta_m\}$$

będą zbiorami wszystkich pierwiastków odpowiednio wielomianów minimalnych $m_\alpha, m_\beta \in \mathbb{F}[X]$, zawierającymi się w $\overline{\mathbb{F}}$, gdzie $\alpha_1 = \alpha$, $\beta_1 = \beta$. Wtedy istnieje element $\gamma \in \mathbb{F}$ taki, że

$$\theta = \alpha + \beta\gamma \quad \text{oraz} \quad \theta - \gamma\beta_i \neq \alpha_j$$

dla wszystkich indeksów $i \in \{2, \ldots, m\}$ oraz $j \in \{1, \ldots, n\}$, a więc $\mathbb{F}(\theta) \subseteq \mathbb{F}(\alpha, \beta)$.

Rozpatrzmy wielomian

$$g(X) = m_\alpha(\theta - \gamma X) \in (\mathbb{F}(\theta))[X].$$

Wtedy

$$\begin{array}{rcccl} g(\beta) & = & m_\alpha(\theta - \gamma\beta) = m_\alpha(\alpha) & = & 0, \\ g(\beta_i) & = & m_\alpha(\theta - \gamma\beta_i) & \neq & 0 \ (j = 2, \ldots, m). \end{array}$$

Zatem w wyniku rozdzielczości rozszerzenia $\beta \in \overline{\mathbb{F}}$ jest dokładnie jedynym wspólnym pierwiastkiem wielomianów $g, m_\beta \in (\mathbb{F}(\theta))[X]$, co implikuje, że

$$\deg \mathrm{NWD}(m_\alpha, m_\beta) = 1 \quad \text{oraz} \quad X - \beta = \mathrm{NWD}(g, m_\beta) \in (\mathbb{F}(\theta))[X].$$

Na tej podstawie $\alpha = \theta - \gamma\beta \in (\mathbb{F}(\theta))[X]$, skąd otrzymujemy, że $\mathbb{F}(\alpha, \beta) \subseteq \mathbb{F}(\theta)$. W końcu $\mathbb{F}(\alpha, \beta) = \mathbb{F}(\theta)$.

2) Przypadek, gdy ciało $\mathbb{F}$ jest skończone, będzie rozpatrzone później we wniosku 7.1.10. □

Wniosek 6.5.9. *Jeśli rozszerzenie ciał $\mathbb{F} \leqslant \mathbb{E}$ jest rozdzielcze oraz elementy $\alpha_1, \ldots, \alpha_n \in \mathbb{E}$ są algebraiczne nad ciałem $\mathbb{F}$, to istnieje element $\theta \in \mathbb{E}$ taki, że*

$$\mathbb{F}(\alpha_1, \ldots, \alpha_n) = \mathbb{F}(\theta).$$

Przykład 6.5.10.

Znajdźmy element prymitywny ciała $\mathbb{E} = \mathbb{Q}(\sqrt{5}, -i\sqrt{3})$. Skoro wielomiany minimalne $m_{\sqrt{5}} = m_{\pm\sqrt{5}} = X^2 - 5 \in \mathbb{Q}[X]$ oraz $m_{i\sqrt{3}} = m_{\pm i\sqrt{3}} = X^2 + 3 \in \mathbb{Q}[X]$, to biorąc $\theta = \alpha + \beta$ dla $\alpha = \sqrt{5}$ oraz $\beta = -i\sqrt{3}$ (patrz twierdzenie 6.5.8), otrzymujemy, że $\mathbb{Q}(\sqrt{5}, -i\sqrt{3}) = \mathbb{Q}(\sqrt{5} - i\sqrt{3})$.

Z następnego lematu wynika, że przykładami rozszerzeń rozdzielczych są rozszerzenia podciał ciała liczb zespolonych.

Lemat 6.5.11. *Jeśli* $\mathbb{F} \leqslant \mathbb{E} \leqslant \mathbb{C}$*, to rozszerzenie ciał* $\mathbb{F} \leqslant \mathbb{E}$ *jest rozdzielcze.*

Dowód. Jeśli element $\alpha \in \mathbb{E}$ jest algebraiczny nad ciałem $\mathbb{F}$ oraz $m_\alpha = a_0X^n + \cdots + a_{n-1}X + a_n \in \mathbb{F}[X]$ ma pierwiastek $\theta \in \mathbb{C}$ krotności $k \geqslant 2$, to na podstawie twierdzenia Bezouta $m_\alpha = (X - \theta)^k \cdot g$ dla pewnego wielomianu $g \in \mathbb{C}[X]$. Wtedy θ jest pierwiastkiem pochodnej tego wielomianu $m'_\alpha = na_0X^{n-1} + \cdots + 2a_{n-2}X + a_{n-1} \in \mathbb{F}[X]$, a to prowadzi do sprzeczności. □

Ćwiczenia 6.5.12.

(**1**) Udowodnić, że jeśli ciało $\mathbb{F}$ nie jest skończone, to każde rozszerzenie algebraiczne ciała $\mathbb{F}$ jest równoliczne z ciałem $\mathbb{F}$.
(**2**) Niech $\overline{\mathbb{Q}}$ będzie pewnym domknięciem algebraicznym ciała $\mathbb{Q}$, a S będzie podciałem w $\overline{\mathbb{Q}}$, które nie posiada $\sqrt{3}$ i jest podciałem maksymalnym z tą własnością. Jeśli F jest podciałem w $\overline{\mathbb{Q}}$ oraz $S \subseteq F$ jest rozszerzeniem skończonym, to

$$G(F, S) = \{f : F \to F \mid f \text{ jest automorfizmem ciała } F \text{ oraz } f(s) = s \text{ dla wszystkich } s \in S\}$$

jest grupą cykliczną.
(**2**) Udowodnić, że $a, b \in \mathbb{R}$ są liczbami algebraicznymi wtedy i tylko wtedy, gdy $a + bi \in \mathbb{C}$ jest liczbą algebraiczną.
(**3**) Niech A będzie rozszerzeniem algebraicznym ciała $\mathbb{F}$. Udowodnić, że jeśli każdy wielomian $f \in \mathbb{F}[X]$ stopnia $n \geqslant 1$ ma przynajmniej jeden pierwiastek w ciele A, to A jest algebraicznie domknięte.

Uwagi. Domknięcie algebraiczne $\overline{\mathbb{Q}}$ ciała liczb wymiernych $\mathbb{Q}$ jest nazywane ciałem liczb algebraicznych.

Rozdział 7

Ciała skończone

7.1. Istnienie i jedność ciał skończonych

■ **Własności funkcji Eulera.** Funkcja $f : \mathbb{N} \to \mathbb{N}$ jest nazywana *funkcją multiplikatywną*, jeśli są spełnione dwa warunki:
1) f nie jest zerowa;
2)

$$\forall_{a,b\in\mathbb{N}} \ : \mathrm{NWD}(a,b) = 1 \quad \Rightarrow \quad f(a \cdot b) = f(a) \cdot f(b).$$

Przypomnijmy, że funkcja Eulera φ dla każdej dodatniej liczby całkowitej n przyjmuje wartość $\varphi(n)$, która jest równa liczbie całkowitych liczb dodatnich mniejszych od liczby n i względnie pierwszych z liczbą n.

Lemat 7.1.1. *Funkcja Eulera φ jest multiplikatywna.*

Dowód. Niech $x \in \mathbb{Z}$. Jeśli $n \in \mathbb{N}^*$, to klasa reszt $x_n = x + n\mathbb{Z}$ jest odwracalna w pierścieniu $\mathbb{Z}_n$ wtedy i tylko wtedy, gdy $\mathrm{NWD}(x, n) = 1$. Zatem w $\mathbb{Z}_n$ istnieje $\varphi(n)$ elementów odwracalnych. Skoro

$$f : \mathbb{Z}_{mn} \ni x_{mn} \mapsto (x_m, x_n) \in \mathbb{Z}_m \oplus \mathbb{Z}_n$$

jest homomorfizmem pierścieni, to w pierścieniach $\mathbb{Z}_{mn}$ i $\mathbb{Z}_m \oplus \mathbb{Z}_n$ liczby elementów odwracalnych są równe, a więc $\varphi(mn) = \varphi(m)\varphi(n)$. □

Twierdzenie 7.1.2. *Jeśli*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} \tag{7.1}$$

jest rozkładem kanonicznym liczby $n \in \mathbb{N} \setminus \{0, 1\}$, to zachodzi wzór

$$\varphi(n) = n \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right).$$

Dowód. Rzeczywiście, ponieważ

$$\varphi(n) = \prod_{i=1}^{k} \varphi(p_i^{\alpha_i})$$

oraz

$$\varphi(p_i^{\alpha_i}) = p_i^{\alpha_i} - p_i^{\alpha_i - 1},$$

to wśród liczb $1,\ 2,\ \ldots,\ p_i^{\alpha_i}$ dokładnie liczby

$$1 \cdot p_i,\ 2 \cdot p_i,\ \ldots,\ p_i^{\alpha_i - 1} \cdot p_i$$

nie są względnie pierwsze z liczbą $p_i^{\alpha_i}$, a więc teza zachodzi. □

Twierdzenie 7.1.3. *Jeśli* $n, t \in \mathbb{N}^*$, *to*

$$\sum_{t|n} \varphi(t) = n.$$

Dowód. Dla $n = 1$ teza ma miejsce. Załóżmy, że $n > 1$ oraz n ma postać kanoniczną (7.1). Wtedy

$$t \mid n \quad \Leftrightarrow \quad t = \prod_{i=1}^{k} p_i^{\beta_i} \quad (0 \leqslant \beta_i \leqslant \alpha_i;\ \beta_i \in \mathbb{N}).$$

Zatem

$$\begin{aligned}
\sum_{t|n} \varphi(t) &= \sum_{\substack{(\beta_1, \ldots, \beta_k) \\ 0 \leqslant \beta_i \leqslant \alpha_i}} \varphi(p_1^{\beta_1})\varphi(p_2^{\beta_2}) \cdots \varphi(p_k^{\beta_k}) = \\
&= \prod_{i=1}^{k} \sum_{\beta_i=0}^{\alpha_i} \varphi(p_i^{\beta_i}) = \prod_{i=1}^{k} (\varphi(1) + \varphi(p_i) + \cdots + \varphi(p_i^{\alpha_i})) = \\
&= \prod_{i=1}^{k} (1 + (p_i - 1) + p_i(p_i - 1) + \cdots + p_i^{\alpha_i - 1}(p_i - 1)) = \\
&= \prod_{i=1}^{k} p_i^{\alpha_i} = n.
\end{aligned}$$

□

* * *

■ **Liczba elementów ciała skończonego.** Ciało $\mathbb{F}$, składające się ze skończonej liczby elementów, jest nazywane *skończonym*. Jeśli e jest jednością tego ciała oraz $k \in \mathbb{Z}$, to

$$k \cdot e = \begin{cases} \underbrace{e + e + \cdots + e}_{k \text{ składników}}, & \text{gdy } k > 0, \\ 0, & \text{gdy } k = 0, \\ \underbrace{(-e) + (-e) + \cdots + (-e)}_{|k| \text{ składników}}, & \text{gdy } k < 0. \end{cases}$$

Lemat 7.1.4. *Jeśli $\mathbb{F}$ jest ciałem skończonym, to jego moc $|\mathbb{F}| = p^n$ dla pewnej liczby pierwszej p oraz dodatniej liczby całkowitej n.*

Dowód. Skoro ciało $\mathbb{F}$ jest skończone, to jego charakterystyka $\operatorname{char}\mathbb{F} = p$ jest liczbą pierwszą p. Niech

$$F_0 = \{k \cdot e \mid k \in \mathbb{Z}\},$$

gdzie e jest jednością ciała $\mathbb{F}$. Na podstawie twierdzenia o dzieleniu z resztą istnieją takie liczby całkowite q, r, że

$$k = pq + r \text{ oraz } 0 \leqslant r \leqslant p - 1,$$

a więc

$$k \cdot e = (p \cdot e) \cdot (q \cdot e) + (r \cdot e) = r \cdot e$$

oraz

$$F_0 = \{r \cdot e \mid r = 0, 1, \ldots, p - 1\}.$$

Jeśli $r \neq 0$, to $pu + rv = 1$ dla pewnych liczb całkowitych u, v i wtedy

$$e = 1 \cdot e = (p \cdot e) \cdot (u \cdot e) + (r \cdot e) \cdot (v \cdot e) = (r \cdot e) \cdot (v \cdot e) = (v \cdot e) \cdot (r \cdot e),$$

czyli $(r \cdot e)^{-1} = v \cdot e \in F_0$. Zatem F_0 jest ciałem zawierającym się w $\mathbb{F}$ i w konsekwencji $\mathbb{F}$ jest przestrzenią liniową nad ciałem F_0. Ponieważ $|\mathbb{F}| < \infty$, to wymiar $\dim_{F_0}\mathbb{F}$ jest skończony i dlatego $\dim_{F_0}\mathbb{F} = n$ dla pewnej liczby całkowitej $n \geqslant 1$. Przestrzeń liniowa $\mathbb{F}$ ma pewną bazę

$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ i każdy jej element $a \in \mathbb{F}$ jest dokładnie jednoznacznie zapisywany w postaci kombinacji liniowej $a = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$, gdzie $\alpha_1, \alpha_2, \ldots, \alpha_n \in F_0$. Zatem możemy wnioskować, że $\mathbb{F}$ zawiera dokładnie p^n elementów.

□

■ Zachodzi izomorfizm ciał

$$F_0 \ni k \cdot e \mapsto k \cdot \overline{1} \in \mathbb{Z}_p \quad (k \in \mathbb{Z})$$

i dlatego ciało F_0 z dowodu lematu 7.1.4 możemy utożsamiać z ciałem klas reszt $\mathbb{Z}_p$ modulo liczby pierwszej p (to znaczy, że z dokładnością do izomorfizmu ciał istnieje dokładnie jedno ciało o p elementach), czyli $F_0 = \mathbb{Z}_p$ (to znaczy, że F_0 oraz $\mathbb{Z}_p$ mają takie same własności algebraiczne i z punktu widzenia algebry ciała F_0 oraz $\mathbb{Z}_p$ są identyczne). Dalej ciało skończone o p elementach będziemy oznaczać przez $\mathbb{F}_p$. Zrozumiałe, że każde ciało charakterystyki p zawiera podciało $\mathbb{F}_p$.

Przykład 7.1.5.

Jeśli $p = 2$, to $\mathbb{F}_2 = \{0, 1\}$ jest ciałem o dwóch elementach z takimi tabliczkami dodawania „+” i mnożenia „·”:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Wniosek 7.1.6. *Jeśli $\mathbb{F}$ jest ciałem skończonym o p^n elementach, gdzie n jest dodatnią liczbą całkowitą, p jest liczbą pierwszą oraz $a, b \in \mathbb{F}$, to:*

(1) $(a+b)^{p^k} = a^{p^k} + b^{p^k}$ *dla każdej nieujemnej liczby całkowitej k;*

(2) $a^{p^n} = a$.

Dowód. (1) W rzeczy samej, charakterystyka p ciała $\mathbb{F}$ jest dzielnikiem współczynnika Newtona $\binom{p}{i}$ dla każdego $i = 1, \ldots, p-1$ oraz

$$(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i},$$

skąd dostajemy, że $(a+b)^p = a^p + b^p$ i otrzymujemy tezę.

(2) Łatwo wynika z części (1). □

* * *

■ **Istnienie i jednoznaczność ciała skończonego.**

Twierdzenie 7.1.7. *Zachodzą następujące własności:*

(1) *Dla każdej liczby pierwszej p i każdej niezerowej liczby naturalnej n istnieje ciało skończone o p^n elementach;*

(2) *Każde ciało skończone o p^n elementach jest izomorficzne z ciałem rozkładu wielomianu*
$$X^{p^n} - X \in \mathbb{F}_p[X];$$

(3) *Istnieje jedno ciało skończone o p^n elementach (z dokładnością do izomorfizmu).*

Dowód. Niech $f = X^{p^n} - X \in \mathbb{F}_p[X]$.

(1) Istnieje ciało rozkładu $\mathbb{L}$ wielomianu f (patrz twierdzenie 6.4.3). Niech
$$\mathbb{F} = \{\theta \in \mathbb{L} \mid f(\theta) = 0\}.$$
Wielomian f ma w ciele $\mathbb{L}$ co najwyżej $\deg f = p^n$ pierwiastków. Skoro pochodna $f' = -1$, to wszystkie pierwiastki wielomianu f są jednokrotne, a więc $|\mathbb{F}| = p^n$. Ponieważ $0, 1 \in \mathbb{F}$ oraz $\alpha - \beta, \alpha^{-1} \in \mathbb{F}$ dla dowolnych $\alpha, \beta \in \mathbb{F}$, to $\mathbb{F}$ jest podciałem w $\mathbb{L}$. Zatem $\mathbb{F}$ jest ciałem i w konsekwencji $\mathbb{F} = \mathbb{L}$.

(2) Niech $\mathbb{S}$ będzie ciałem o p^n elementach. Jeśli $0 \neq x \in \mathbb{S}$, to na podstawie wniosku 3.3.8 rząd $o(x)$ dzieli rząd $p^n - 1$ grupy multiplikatywnej $\mathbb{S}^*$ ciała $\mathbb{S}$. To implikuje, że
$$x^{p^n - 1} = 1,$$
a więc x jest pierwiastkiem wielomianu f. Wnosimy, że każdy element ciała $\mathbb{S}$ jest pierwiastkiem wielomianu f. Zatem $\mathbb{S}$ jest ciałem rozkładu wielomianu f i teza zachodzi.

(3) Wynika z części (1) i twierdzenia 6.4.5. □

* * *

■ **Grupa multiplikatywna ciała skończonego.** Mamy takie

Twierdzenie 7.1.8. *Niech p będzie liczbą pierwszą oraz $t, n \in \mathbb{N}^*$. Jeśli t dzieli liczbę $p^n - 1$, to zachodzą następujące własności:*

(1) *liczba elementów rzędu t w grupie multiplikatywnej $\mathbb{F}_{p^n}^*$ jest równa $\varphi(t)$;*
(2) *$\mathbb{F}_{p^n}^*$ zawiera dokładnie jedną podgrupę rzędu t;*
(3) *$\mathbb{F}_{p^n}^*$ jest grupą cykliczną rzędu $p^n - 1$.*

Dowód. Niech t będzie dzielnikiem liczby p^n. Wtedy grupa $\mathbb{F}_{p^n}^*$ albo nie zawiera żadnego elementu rzędu t, albo zawiera co najmniej jeden taki element. Załóżmy, że w grupie $\mathbb{F}_{p^n}^*$ istnieje element θ rzędu t. Wtedy $\langle\theta\rangle$ jest grupą cykliczną rzędu t. Ponieważ:

- każdy element $b \in \langle\theta\rangle$ jest pierwiastkiem wielomianu $X^t - 1 \in \mathbb{F}_p[X]$ zawierającym się w ciele $\mathbb{F}_{p^n}$,
- wszystkie pierwiastki wielomianu $X^t - 1 \in \mathbb{F}_p[X]$, zawierające się w ciele $\mathbb{F}_{p^n}$, leżą w podgrupie $\langle\theta\rangle$,
- podgrupa cykliczna $\langle\theta\rangle$ ma $\varphi(t)$ generatorów (czyli $\varphi(t)$ elementów rzędu t na podstawie twierdzenia 3.1.10),

to $\mathbb{F}_{p^n}^*$ ma $\varphi(t)$ elementów rzędu t. To implikuje, że w grupie $\mathbb{F}_{p^n}^*$ zawiera się 0 lub $\varphi(t)$ elementów rzędu t. Na podstawie twierdzenia 7.1.3 wnosimy, że grupa $\mathbb{F}_{p^n}^*$ zawsze ma element rzędu t takiego, że t jest dzielnikiem liczby $p^n - 1$. Zatem dla każdego dzielnika $t \mid p^n - 1$ grupa $\mathbb{F}_{p^n}^*$ ma dokładnie jedną podgrupę rzędu t, która jest cykliczna. To też powoduje, że $\mathbb{F}_{p^n}^*$ jest cykliczna.

□

Przykłady 7.1.9.

(**1**) Ciało o 4 elementach $\mathbb{F}_2[X]/\langle X^2 + X + 1\rangle = \{0, 1, x, x+1\}$ zostało skonstruowane w przykładzie 4.5.8(1).

(**2**) Mamy ciało o trzech elementach $\mathbb{F}_3 = \{0, 1, 2\}$ z tabliczkami Cayleya (dodawania i mnożenia):

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Jeśli $h = X^2 + 1 \in \mathbb{F}_3[X]$, to $h(0) = 1 \neq 0,\ h(1) = 2 \neq 0,\ h(2) = 2 \neq 0$, czyli wielomian h jest nieprzywiedlny nad ciałem $\mathbb{F}_3$ oraz

$$\mathbb{F}_3[X]/\langle h\rangle = \{aX + b + \langle h\rangle \mid a, b \in \mathbb{F}_3\}$$

jest ciałem o 9 elementach. Jeśli wprowadzimy oznaczenia

$$\begin{array}{lcl} 0 & = & 0+\langle h\rangle,\\ 1 & = & 1+\langle h\rangle,\\ 2 & = & 2+\langle h\rangle,\\ \alpha & = & X+\langle h\rangle,\\ \alpha+1 & = & X+1+\langle h\rangle,\\ \alpha+2 & = & X+2+\langle h\rangle,\\ 2\alpha & = & 2X+\langle h\rangle,\\ 2\alpha+1 & = & 2X+1+\langle h\rangle,\\ 2\alpha+2 & = & 2X+2+\langle h\rangle, \end{array}$$

to ciało

$$\mathbb{F}_3[X]/\langle h\rangle = \{0,1,2,\alpha,\alpha+1,\alpha+2,2\alpha,2\alpha+1,2\alpha+2\}$$

ma takie tabliczki Cayleya (dodawania i mnożenia):

$+$	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
0	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
1	1	2	0	$\alpha+1$	$\alpha+2$	α	$2\alpha+1$	$2\alpha+2$	2α
2	2	0	1	$\alpha+2$	α	$\alpha+1$	$2\alpha+2$	2α	$2\alpha+1$
α	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$	0	1	2
$\alpha+1$	$\alpha+1$	$\alpha+2$	α	$2\alpha+1$	$2\alpha+2$	2α	1	2	0
$\alpha+2$	$\alpha+2$	α	$\alpha+1$	$2\alpha+2$	2α	$2\alpha+1$	2	0	1
2α	2α	$2\alpha+1$	$2\alpha+2$	0	1	2	α	$\alpha+1$	$\alpha+2$
$2\alpha+1$	$2\alpha+1$	$2\alpha+2$	2α	1	2	0	$\alpha+1$	$\alpha+2$	α
$2\alpha+2$	$2\alpha+2$	2α	$2\alpha+1$	2	0	1	$\alpha+2$	α	$\alpha+1$

$\cdot$	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
0	0	0	0	0	0	0	0	0	0
1	0	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$
2	0	2	1	2α	$2\alpha+2$	$2\alpha+1$	α	$\alpha+2$	$\alpha+1$
α	0	α	2α	2	$\alpha+2$	$2\alpha+2$	1	$\alpha+1$	$2\alpha+1$
$\alpha+1$	0	$\alpha+1$	$2\alpha+2$	$\alpha+2$	2α	1	$2\alpha+1$	2	α
$\alpha+2$	0	$\alpha+2$	$2\alpha+1$	$2\alpha+2$	1	α	$\alpha+1$	2α	2
2α	0	2α	α	1	$2\alpha+1$	$\alpha+1$	2	$2\alpha+2$	$\alpha+2$
$2\alpha+1$	0	$2\alpha+1$	$\alpha+2$	$\alpha+1$	2	2α	$2\alpha+2$	α	1
$2\alpha+2$	0	$2\alpha+2$	$\alpha+1$	$2\alpha+1$	α	2	$\alpha+2$	1	2α

(**3**) Jeśli p jest liczbą pierwszą, to ciało klas reszt $\mathbb{Z}_p = \{\overline{0},\ \overline{1},\ \ldots,\ \overline{p-1}\}$ ma cykliczną grupę multiplikatywną $\mathbb{Z}_p^* = \langle\overline{1}\rangle$ generowaną przez element $\overline{1} = \{1+pt \mid t\in\mathbb{Z}\}$ rzędu $p-1$.

Możemy teraz udowodnić twierdzenie Abela (patrz 6.5.8) w przypadku ciała skończonego.

Wniosek 7.1.10 (Abela)**.** *Niech $\mathbb{F}\leqslant\mathbb{E}$ będzie rozszerzeniem rozdzielczym ciał $\mathbb{F}$, $\mathbb{E}$ oraz $\alpha,\beta\in\mathbb{E}$ będą elementami algebraicznymi nad ciałem $\mathbb{F}$. Jeśli ciało $\mathbb{F}$ jest skończone, to istnieje element $\theta\in\mathbb{E}$ taki, że*

$$\mathbb{F}(\alpha,\beta) = \mathbb{F}(\theta)$$

(element θ jest nazywany *elementem prymitywnym* tego rozszerzenia).

Dowód. Ciało $\mathbb{F}(\alpha,\beta)$ jest skończone, a więc jego grupa multiplikatywna $\mathbb{F}(\alpha,\beta)^* = \langle\theta\rangle$ jest cykliczną generowaną przez pewien element $\theta \in \mathbb{F}(\alpha,\beta)$ i teza zachodzi. □

Konwencja. Wszędzie dalej ciało skończone o p^u elementach (gdzie $u > 0$ jest liczbą całkowitą) będziemy oznaczać przez $\mathbb{F}_{p^u}$ (lub przez $GF(p^u)$). Często zamiast p^u będziemy pisać q oraz zamiast $\mathbb{F}_{p^u}$ będziemy używać krótszego oznaczenia $\mathbb{F}_q$.

■ Ciało skończone $GF(q) = \mathbb{F}_q$ jest nazywane *ciałem Galois*. Ciało $\mathbb{F}_p$ jest nazywane *prostym*, a ciało $\mathbb{F}_{p^u}$ $(u \geqslant 2)$ *rozszerzonym*.

Ćwiczenia 7.1.11.

(**1**) Udowodnić, że suma elementów ciała skończonego $\mathbb{F}_q$ $(q \neq 2)$ jest równa 0.
(**2**) Udowodnić, że ciało $\mathbb{F}$ z cykliczną grupą multiplikatywną $\mathbb{F}^*$ jest skończone.
(**3**) Niech $\mathbb{F}_q$ będzie ciałem skończonym nieparzystej charakterystyki. Wtedy w ciele $\mathbb{F}_q$ istnieje pierwiastek kwadratowy z elementu $a \in \mathbb{F}_q$ w tym i tylko tym przypadku, gdy $a^{\frac{q-1}{2}} = 1$.
(**4**) Niech n będzie nieparzystą dodatnią liczbą całkowitą oraz $x, y \in \mathbb{F}_{2^n}$. Udowodnić, że z równości $x^2 + xy + y^2 = 0$ wynika, że $x = y = 0$.
(**5**) Udowodnić, że:
(a) przestrzeń $\mathbb{F}_2^n$ ma $\binom{n}{t}$ wektorów, które mają współrzędne 0 na dokładnie t pozycjach;
(b) przestrzeń $\mathbb{F}_q^n$ ma $\binom{n}{t}(q-1)^{n-t}$ wektorów, które mają współrzędne 0 na dokładnie t pozycjach.
(**6**) Mamy ciało $\mathbb{F}_q$, gdzie liczba q jest nieparzysta. Udowodnić, że iloczyn wszystkich niezerowych elementów z $\mathbb{F}_q$ jest równy -1.
(**7**) Udowodnić w ciele $\mathbb{F}$ charakterystyki $p > 0$ takie, że:
(a) jeśli $p = 2$, to każdy element ciała $\mathbb{F}$ jest kwadratem;
(b) jeśli p nie jest parzyste, to połowa niezerowych elementów ciała jest kwadratami.
(**8**) Zbudować ciało $\mathbb{F}_{16} = \mathbb{F}_2[X]/\langle X^4 + X^3 + X^2 + X + 1\rangle$.

Uwagi. Ciała proste $\mathbb{F}_p$ dla różnych liczb pierwszych p w różnym stopniu (i z różnych powodów) zostały przebadane przez P. de Fermata, J. Lagrange'a[(1)], L. Eulera, C. Gaussa, A. Legendre'a oraz innych. Osiągnięciem E. Galoisa było wprowadzenie rozszerzenia ciała prostego, czyli

(1) Arien-Marie Legendre (1752–1823)

ciała $\mathbb{F}_q$. E.H. Moore w 1893 r. na Międzynarodowym Kongresie Matematycznym zaprezentował dowód tego, że każde ciało skończone $\mathbb{F}_{p^n}$ jest izomorficzne z ciałem $\mathbb{F}_p[X]/\langle f\rangle$, gdzie $f \in \mathbb{F}_p[X]$ jest wielomianem nieprzywiedlnym stopnia n nad ciałem $\mathbb{F}_p$. Faktycznie E. Moore pierwszy użył angielskojęzycznego terminu „field". Twierdzenie 7.1.7 było udowodnione przez E.H. Moore'a w 1893 r.

7.2. Podciała i automorfizmy Frobeniusa ciała skończonego

Lemat 7.2.1. *Niech s, n będą dodatnimi liczbami całkowitymi. Wtedy nad każdym ciałem skończonym zachodzą własności:*

(1) $X^s - 1$ *dzieli* $X^n - 1 \quad \Leftrightarrow \quad s$ *dzieli* n,

(2) $\mathrm{NWD}(X^s - 1, X^n - 1) = X^{\mathrm{NWD}(s,n)} - 1$.

Dowód. Udowodnić samodzielnie. □

Twierdzenie 7.2.2 (o podciałach ciała skończonego)**.** *Niech n będzie dodatnią liczbą całkowitą. Wtedy zachodzą własności:*

(1) *jeśli $\mathbb{F}$ jest podciałem ciała $\mathbb{F}_{p^n}$, to istnieje taka dodatnia liczba całkowita m, że ciało $\mathbb{F}$ jest izomorficzne z ciałem $\mathbb{F}_{p^m}$ oraz m jest dzielnikiem liczby n,*

(2) *jeśli dodatnia liczba całkowita m jest dzielnikiem liczby n, to istnieje dokładnie jedno (z dokładnością do izomorfizmu ciał) ciało $\mathbb{F}$ o p^m elementach będące podciałem w ciele $\mathbb{F}_{p^n}$.*

Dowód. (1) Skoro $\mathbb{F}$ jest podciałem w $\mathbb{F}_{p^n}$, to $\operatorname{char}\mathbb{F} = p$ i na podstawie lematu 7.1.4 zachodzi $\mathbb{F} = \mathbb{F}_{p^m}$ dla pewnej dodatniej liczby całkowitej m. Wtedy $\mathbb{F}_{p^n}$ jest przestrzenią liniową nad ciałem $\mathbb{F}_{p^m}$, a więc istnieje dodatnia liczba całkowita s taka, że $p^n = (p^m)^s$, czyli m dzieli n.

(2) Jeśli m dzieli n, to $p^m - 1$ dzieli liczbę $p^n - 1$. Na mocy lematu 7.2.1 wielomian $X^{p^m} - X$ dzieli wielomian $X^{p^n} - X$. Każdy pierwiastek wielomianu $X^{p^m} - X$ jest pierwiastkiem dla $X^{p^n} - X$, a pierwiastki tego ostatniego wielomianu nad $\mathbb{F}_p$ tworzą ciało $\mathbb{F}_{p^n}$. Ponieważ pierwiastki wielomianu $X^{p^m} - X$ nad ciałem $\mathbb{F}_p$ tworzą ciało $\mathbb{F}_{p^m}$, to $\mathbb{F}_{p^m}$ jest jedynym podciałem ciała $\mathbb{F}_{p^n}$ o p^m elementach. □

Wniosek 7.2.3. *Niech $f \in \mathbb{F}_p[X]$ będzie unormowanym wielomianem nieprzywiedlnym nad $\mathbb{F}_p$. Wielomian f dzieli $X^{p^m-1} - 1$ wtedy i tylko wtedy, gdy* $\deg f$ *dzieli m.*

Dowód. Niech $k = \deg f$. Pierścień ilorazowy $\mathbb{F} = \mathbb{F}_p[X]/\langle f\rangle$ jest ciałem. Z twierdzenia o dzieleniu z resztą dla dowolnego $g \in \mathbb{F}_p[X]$ istnieją wie-

lomiany $q, r \in \mathbb{F}_p[X]$ takie, że $g = fq + r$ oraz $\deg r < \deg f$. Wtedy na podstawie kryterium równości warstw

$$g + \langle f \rangle = r + (fq + \langle f \rangle) = r + \langle f \rangle = a_0 X^0 + a_1 X^1 + \cdots + a_{k-1} X^{k-1} + \langle f \rangle$$

dla pewnych współczynników $a_0, a_1, \ldots, a_{k-1} \in \mathbb{F}_p$ oraz

$$\begin{gathered} \mathbb{F} = \{g + \langle f \rangle \mid g \in \mathbb{F}_p[X]\} = \\ = \{a_0\theta^0 + a_1\theta^1 + \cdots + a_{k-1}\theta^{k-1} \mid a_i \in \mathbb{F}_p \ (i = 0, 1, \ldots, k-1)\}, \end{gathered}$$

gdzie $\theta = X + \langle f \rangle$. Z tego wynika, że $\dim_{\mathbb{F}_p} \mathbb{F} = k$ oraz $\mathbb{F} = \mathbb{F}_{p^k}$.

($\Rightarrow$) Pierwiastki wielomianu $X^{p^m} - X \in \mathbb{F}_p[X]$ tworzą ciało $\mathbb{F}_{p^m}$. Zatem pierwiastki wielomianu f należą do ciała $\mathbb{F}_{p^m}$ oraz $\mathbb{F}_{p^k}$ jest podciałem w $\mathbb{F}_{p^m}$ i biorąc pod uwagę twierdzenie 7.2.2, wnosimy, że k dzieli m.

($\Leftarrow$) Skoro f jest nieprzywiedlny nad $\mathbb{F}_p$ oraz $f(\theta) = 0$ (patrz wyżej), to f jest wielomianem minimalnym elementu $\theta \in \mathbb{F}$ nad $\mathbb{F}_p$. Wiemy, że $\mathbb{F}$ jest ciałem o p^k elementach. Ponieważ k dzieli m, to $\mathbb{F}_{p^k}$ jest podciałem w $\mathbb{F}_{p^m}$. Pierwiastki wielomianu $X^{p^m} - X \in \mathbb{F}_p[X]$ tworzą ciało $\mathbb{F}_{p^m}$. Zatem θ jest pierwiastkiem wielomianu $X^{p^m} - X$ (a więc i wielomianu $X^{p^m-1} - 1$). Na mocy twierdzenia 6.2.1 wielomian f dzieli $X^{p^m-1} - 1$. □

Wniosek 7.2.4. *Wielomian $X^{p^m} - X \in \mathbb{F}_p[X]$ jest iloczynem nieprzywiedlnych nad $\mathbb{F}_p$ wielomianów, stopnie których są dzielnikami liczby całkowitej m.*

■ Element $\theta \in \mathbb{F}_{p^n}$ jest nazywany *pierwotnym* (lub *prymitywnym*), jeśli $\mathbb{F}_{p^n}^* = \langle \theta \rangle$, czyli θ jest generatorem grupy multiplikatywnej $\mathbb{F}_{p^n}^*$ ciała $\mathbb{F}_{p^n}$.

Przykład 7.2.5.

Niech $f = X^2 + X + 1 \in \mathbb{F}_2[X]$. Skoro $f(0) = 1 \neq 0$, $f(1) = 1 \neq 0$, to f jest nieprzywiedlny nad $\mathbb{F}_2$, a więc $\mathbb{F}_4 = \mathbb{F}_2[X]/\langle f \rangle = \{0, 1, x, x+1\}$ (patrz przykład 4.5.8(1)) jest ciałem. Ponieważ

$$\begin{array}{lcllcl} x^1 & = & x, & (x+1)^1 & = & x+1, \\ x^2 & = & x+1, & (x+1)^2 & = & x^2+1, \\ x^3 & = & x^2+x=1, & (x+1)^3 & = & x^2+x=1, \end{array}$$

to $\mathbb{F}_4^* = \langle x \rangle = \langle x+1 \rangle$. Zatem x (odpowiednio $x+1$) jest elementem prymitywnym ciała $\mathbb{F}_4$.

Lemat 7.2.6. *Ciało skończone $\mathbb{F}_{p^n}$ jest rozszerzeniem prostym ciała $\mathbb{F}_p$, czyli $\mathbb{F}_{p^n} = \mathbb{F}_p(\theta)$ dla każdego prymitywnego elementu $\theta \in \mathbb{F}_{p^n}$.*

Dowód. Skoro grupa $\mathbb{F}_{p^n}^*$ jest cykliczna na mocy twierdzenia 7.1.8, to $\mathbb{F}_{p^n}^* = \langle\theta\rangle$ dla pewnego $\theta \in \mathbb{F}_{p^n}$. Wtedy

$$\theta \in \mathrm{Lin}_{\mathbb{F}_p}\{\theta^0, \theta^1, \ldots, \theta^{p^n-1}\} \subseteq \mathbb{F}_p(\theta) \subseteq \mathbb{F}_{p^n},$$

a więc $\mathbb{F}_{p^n} = \mathbb{F}_p(\theta)$. □

Otrzymujemy takie bardzo ważne

Twierdzenie 7.2.7 (o istnieniu wielomianów nieprzywiedlnych)**.** *Dla każdej liczby pierwszej p oraz dodatniej liczby całkowitej n istnieje wielomian $f \in \mathbb{F}_p[X]$ stopnia $n \geqslant 1$ nieprzywiedlny nad ciałem $\mathbb{F}_p$.*

Dowód. Niech $\theta \in \mathbb{F}_{p^n}$ będzie elementem prymitywnym. Z lematu 7.2.6 ciało $\mathbb{F}_{p^n} = \mathbb{F}_p(\theta)$ jest rozszerzeniem ciała $\mathbb{F}_p$, a z twierdzenia 6.2.3 mamy $\mathbb{F}_p(\theta) \cong \mathbb{F}_p[X]/\langle m_\theta\rangle$, gdzie $m_\theta \in \mathbb{F}_p[X]$ jest wielomianem minimalnym elementu θ nad ciałem $\mathbb{F}_p$, oraz

$$\deg m_\theta = |\mathbb{F}_p(\theta) : \mathbb{F}_p| = n.$$

Zatem $f = m_\theta$ istnieje. □

Twierdzenie 7.2.8. *Jeśli $f \in \mathbb{F}_p[X]$ jest wielomianem stopnia $n \geqslant 1$ nieprzywiedlnym nad ciałem $\mathbb{F}_p$, to:*

(1) *każdy pierwiastek θ wielomianu f nad ciałem $\mathbb{F}_p$ zawiera się w ciele $\mathbb{F}_{p^n}$,*

(2) *ciało $\mathbb{F}_{p^n}$ zawiera n parami różnych pierwiastków wielomianu f postaci*

$$\theta, \theta^p, \theta^{p^2}, \ldots, \theta^{p^{n-1}}.$$

Dowód. Jeśli $f(\theta) = 0$, to stopień rozszerzenia $|\mathbb{F}_p(\theta) : \mathbb{F}_p| = \deg f = n$ na mocy twierdzenia 6.2.3, a zatem $\mathbb{F}_{p^n} = \mathbb{F}_p(\theta)$. Biorąc pod uwagę wniosek 7.1.6, otrzymujemy

$$0 = f(\theta)^{p^k} = f(\theta^{p^k})$$

dla dowolnej nieujemnej liczby całkowitej k, czyli θ^{p^k} jest pierwiastkiem wielomianu f. Skoro grupa multiplikatywna $\mathbb{F}_{p^n}^*$ ma rząd $p^n - 1$, to zachodzi równość $\theta^{p^n} = \theta$.

Jeśli założyć, że $\theta^{p^m} = \theta^{p^l}$ dla takich liczb całkowitych m, l, że $0 \leqslant m < l \leqslant n-1$, to otrzymujemy

$$\theta^{p^{m-l+n}} = (\theta^{p^m})^{p^{n-l}} = (\theta^{p^l})^{p^{n-l}} = \theta^{p^n} = \theta,$$

co implikuje, że wielomian f jest dzielnikiem wielomianu $X^{p^{m-l+n}} - X \in \mathbb{F}_p[X]$ i z wniosku 7.2.3 liczba n jest dzielnikiem liczby $m-l+n$. Lecz $0 < m-l+n < n$, a to prowadzi do sprzeczności, ponieważ elementy ciała $\mathbb{F}_{p^n}$ są pierwiastkami wielomianu $X^{p^n} - X \in \mathbb{F}_p[X]$.

□

Twierdzenie 7.2.9. *Niech p będzie liczbą pierwszą oraz n będzie dodatnią liczbą całkowitą. Wtedy zachodzą własności:*

(1) *każdy automorfizm σ_j ciała $\mathbb{F}_{p^n}$ ma postać*

$$\sigma_j : \mathbb{F}_{p^n} \ni a \mapsto a^{p^j} \in \mathbb{F}_{p^n} \qquad (0 \leqslant j \leqslant n-1)$$

(i jest nazywany *automorfizmem Frobeniusa* ciała $\mathbb{F}_{p^n}$),

(2) *podciało proste $\mathbb{F}_p$ jest zbiorem elementów stałych automorfizmu σ_1, czyli $\{a \in \mathbb{F}_{p^n} \mid \sigma_1(a) = a\} = \mathbb{F}_p$,*

(3) *rząd automorfizmu σ_1 jest równy n (czyli $\sigma_1^n = \mathrm{id}_{\mathbb{F}_{p^n}}$ oraz $\sigma_1^k \neq \mathrm{id}_{\mathbb{F}_{p^n}}$ dla wszystkich liczb całkowitych k takich, że $1 \leqslant k < n$),*

(4) *elementy ciała, które są sprzężone z elementem $\theta \in \mathbb{F}_{p^n}$, mają postać $\sigma_1^j(\theta) = \theta^{p^j}$ ($j = 0, 1, \ldots, n-1$).*

Dowód. (1) Rzeczywiście,

$$\begin{array}{lcl} \sigma_j(a+b) & = & (a+b)^{p^j} = a^{p^j} + b^{p^j} = \sigma_j(a) + \sigma_j(b), \\ \sigma_j(ab) & = & (ab)^{p^j} = a^{p^j} b^{p^j} = \sigma_j(a)\sigma_j(b), \\ \sigma_j(1) & = & 1 \end{array}$$

dla dowolnych $a, b \in \mathbb{F}_{p^n}$. Na mocy wniosku 7.1.6(2) odwzorowanie σ_j jest suriektywne. Jeśli założyć, że $\sigma_j(c) = \sigma_j(d)$ dla pewnych $c, d \in \mathbb{F}_{p^n}$, to $(c-d)^{p^j} = 0$ w ciele $\mathbb{F}_{p^n}$, a stąd $c - d = 0$. To znaczy, że σ_j jest iniektywne. Zatem σ_j jest automorfizmem ciała $\mathbb{F}_{p^n}$.

Skoro $\sigma_j(1) = \sigma_j(1^2) = \sigma_j(1)^2$, to $\sigma_j(1) = 1$, a więc $\sigma(a) = a$ dla każdego $a \in \mathbb{F}_p$. Niech θ będzie elementem prymitywnym ciała $\mathbb{F}_{p^n}$. Jeśli

$m_\theta = X^n + a_1X^{n-1} + \cdots + a_n \in \mathbb{F}_p[X]$ jest jego wielomianem minimalnym, to

$$0 = \sigma_j(m_\theta(\theta)) = \sigma_j(\theta)^n + a_1\sigma_j(\theta)^{n-1} + \cdots + a_n,$$

czyli $\sigma_j(\theta)$ jest pierwiastkiem wielomianu m_θ. Na podstawie twierdzenia 7.2.8(1) wnosimy, że $\sigma_j(\theta) = \theta^{p^j}$ dla pewnej liczby całkowitej j $(0 \leqslant j \leqslant n-1)$.

(2) Niech

$$S = \{a \in \mathbb{F}_{p^n} \mid \sigma_1(a) = a\}.$$

Wiadomo, że $\mathbb{F}_p \leqslant \mathbb{F}_{p^n}$. Jeśli $a \in \mathbb{F}_p$, to $a = 1 + \cdots + 1$ jest pewną sumą skończoną jedności 1 ciała $\mathbb{F}_{p^n}$ oraz

$$\sigma_1(a) = (1 + \cdots + 1)^p = 1^p + \cdots + 1^p = 1 + \cdots + 1 = a,$$

czyli $\mathbb{F}_p \subseteq S$. Jeśli $b \in S$, to $b^p = b$, co implikuje, że $b^{p-1} = 1$ oraz $\langle b\rangle$ jest podgrupą cykliczną rzędu $\leqslant p-1$. Ponieważ grupa cykliczna $\mathbb{F}^*_{p^n}$ ma rząd równy $p^n - 1$ oraz liczba $p-1$ dzieli $p^n - 1$ oraz na mocy twierdzenia 7.1.8 grupa $\mathbb{F}^*_{p^n}$ zawiera dokładnie jedną podgrupę cykliczną rzędu $p-1$ (a mianowicie $\mathbb{F}^*_p$), to $\langle b\rangle \leqslant \mathbb{F}^*_p$. Stąd wynika, że $b \in \mathbb{F}_p$ oraz w konsekwencji $S = \mathbb{F}_p$.

(3) Na mocy wniosku 7.1.6 mamy $\mathrm{id}_{\mathbb{F}_{p^n}}(b) = b = b^{p^n} = \sigma_1^n(b)$ dla każdego $b \in \mathbb{F}_{p^n}$, co oznacza, że $\sigma_1^n = \mathrm{id}_{\mathbb{F}_{p^n}}$ jest odwzorowaniem tożsamościowym. Twierdzenie 7.1.8 implikuje, że $\mathbb{F}^*_{p^n} = \langle a\rangle$ jest grupą cykliczną (z pewnym generatorem a). Jeśli k jest liczbą całkowitą taką, że $1 \leqslant k < n$ oraz $\sigma_1^k = \mathrm{id}_{\mathbb{F}_{p^n}}$, to $a^{p^k} = a$, skąd $a^{p^k-1} = 1$. Ponieważ

$$p^n - 1 = |\mathbb{F}^*_{p^n}| = |a| \leqslant p^k - 1,$$

to mamy sprzeczność. Zatem $\sigma_1^k \neq \mathrm{id}_{\mathbb{F}_{p^n}}$, gdzie $1 \leqslant k < n$.

(4) Niech $\theta \in \mathbb{F}_{p^n} \setminus \mathbb{F}_p$ oraz stopień rozszerzenia $|\mathbb{F}_p(\theta) : \mathbb{F}_p| = d$. Wtedy $\mathbb{F}_p(\theta) = \mathbb{F}_{p^d}$, czyli θ jest pierwiastkiem wielomianu $X^{p^d} - X \in \mathbb{F}_p[X]$. Skoro θ jest elementem prymitywnym ciała $\mathbb{F}_{p^d}$, to θ nie jest pierwiastkiem wielomianu $X^{p^k} - X \in \mathbb{F}_p[X]$ $(1 \leqslant k < d)$. Jak w dowodzie twierdzenia 7.2.8 otrzymujemy, że $m_\theta = m_{\theta^p} = m_{\sigma_1(\theta)}$, a stąd wynika teza. □

Wniosek 7.2.10. *Grupa automorfizmów* $\operatorname{Aut}\mathbb{F}_{p^n} = \langle\sigma_1\rangle$ *jest cykliczna z generatorem* σ_1 *rzędu* n.

Przykłady 7.2.11.

(**1**) W ciele $\mathbb{F}_3 = \{0, 1, 2\}$ automorfizm

$$\sigma : \mathbb{F}_3 \ni a \mapsto a^3 \in \mathbb{F}_3$$

jest taki, że $\sigma(0) = 0,\ \sigma(1) = 1,\ \sigma(2) = 2^3 = 2$, czyli $\sigma = \mathrm{id}_{\mathbb{F}_3}$ jest tożsamościowy. Zatem grupa automorfizmów $\operatorname{Aut}\mathbb{F}_3$ jest jednostkowa.

(**2**) Niech $\mathbb{F}_4 = \mathbb{F}_2[X]/\langle X^2+X+1\rangle = \{0, 1, \theta, \theta+1\}$, gdzie $\theta = X+\langle X^2+X+1\rangle$. Wtedy automorphism

$$\sigma : \mathbb{F}_4 \ni a \mapsto a^2 \in \mathbb{F}_4$$

spełnia warunki

$$\begin{array}{lcl} \sigma(0) & = & 0, \\ \sigma(1) & = & 1, \\ \sigma(\theta) & = & \theta^2 = \theta + 1, \\ \sigma(\theta+1) & = & \sigma(\theta) + 1 = \theta, \end{array}$$

a więc $\sigma \neq \mathrm{id}_{\mathbb{F}_4}$. Skoro

$$\begin{array}{lcl} \sigma^2(0) & = & 0, \\ \sigma^2(1) & = & 1, \\ \sigma^2(\theta) & = & \sigma(\theta+1) = \sigma(\theta) + 1 = \theta, \\ \sigma^2(\theta+1) & = & \sigma(\theta) = \theta + 1, \end{array}$$

to wnosimy, że $\sigma^2 = \mathrm{id}_{\mathbb{F}_4}$ oraz $\operatorname{Aut}\mathbb{F}_4$ jest grupą cykliczną rzędu 2.

Ćwiczenia 7.2.12.

(**1**) Znaleźć liczbę podciał w ciele $\mathbb{F}_{q^r}$, jeśli:
(a) $q^r = 3^6$;
(b) $q^r = 5^{18}$;
(c) $q^r = 3^{17}$;
(d) $q^r = 2^{14}$;
(e) $q^r = 2^{15}$;
(f) $q^r = 2^{12}$;
(g) $q^r = 2^{11}$;
(h) $q^r = 3^5$.
(**2**) Niech $\mathbb{F}$ będzie ciałem oraz $\varphi : \mathbb{F} \to \mathbb{F}$ będzie odwzorowaniem takim, że

$$\varphi(a) = \begin{cases} a^{-1}, & \text{gdy } a \neq 0, \\ 0, & \text{gdy } a = 0. \end{cases}$$

Udowodnić, że $\varphi \in \operatorname{Aut}\mathbb{F}$ wtedy i tylko wtedy, gdy $|\mathbb{F}| \leqslant 4$.
(**3**) Znaleźć elementy prymitywne ciała:
(a) $\mathbb{F}_3$;
(b) $\mathbb{F}_5$;

(*c*) $\mathbb{F}_7$;
(*d*) $\mathbb{F}_{11}$;
(*e*) $\mathbb{F}_{13}$.

Uwagi. W 1893 r. E.H. Moore jako pierwszy udowodnił, że rząd ciała skończonego jest potęgą liczby pierwszej oraz dla każdej liczby pierwszej p i dodatniej liczby całkowitej n istnieje jedyne (z dokładnością do izomorfizmu) ciało rzędu p^n.

7.3. Wielomiany prymitywne

Twierdzenie 7.3.1. *Jeśli $f \in \mathbb{F}_q[X]$ jest wielomianem stopnia $r \geqslant 1$ oraz $f(0) \neq 0$, to istnieje liczba całkowita e, taka, że $0 < e \leqslant q^r - 1$ oraz f dzieli wielomian $X^e - 1 \in \mathbb{F}_q[X]$.*

Dowód. Ponieważ $|\mathbb{F}_q[X]/\langle f\rangle| = q^r = |A|$, gdzie

$$A = \{X^i + \langle f\rangle \mid i = 0, 1, \ldots, q^r - 1\} \subseteq \mathbb{F}_q[X]/\langle f\rangle$$

oraz wszystkie elementy z A są niezerowe, to $X^k + \langle f\rangle = X^l + \langle f\rangle$ dla pewnych liczb całkowitych k, l takich, że $0 \leqslant k < l \leqslant q^r - 1$. Skoro f dzieli $X^l - X^k$ oraz $\mathrm{NWD}(X, f) = 1$, to f jest dzielnikiem wielomianu $X^{l-k} - 1 \in \mathbb{F}_q[X]$ i teza zachodzi. □

■ Jeśli $f \in \mathbb{F}_q[X]$ oraz $f(0) \neq 0$, to najmniejsza dodatnia liczba całkowita e, taka, że f dzieli $X^e - 1 \in \mathbb{F}_q[X]$, jest nazywana *rzędem* (=the order) wielomianu f (i oznaczana przez $\operatorname{ord} f$). Jeśli $f(0) = 0$, to $f = X^k g$ dla pewnej liczby całkowitej $k > 0$ i pewnego wielomianu $g \in \mathbb{F}_q[X]$ takiego, że $\mathrm{NWD}(X, g) = 1$. Wtedy przyjmujemy, że $\operatorname{ord} g = \operatorname{ord} f$.

Przykłady 7.3.2.

(**1**) Niech $f = X^2 + X + 1 \in \mathbb{F}_3[X]$ oraz $e = \operatorname{ord} f$. Wtedy:
- f nie dzieli $X - 1 \Rightarrow e > 1$,
- f nie dzieli $X^2 - 1 \Rightarrow e > 2$,
- $X^3 - 1 = f \cdot (X - 1)$, czyli f dzieli $X^3 - 1 \Rightarrow e \leqslant 3$.

Zatem rząd wielomianu $\operatorname{ord} f = 3$.

(**2**) Niech $f = X^3 + X + 1 \in \mathbb{F}_2[X]$ oraz $e = \operatorname{ord} f$. Wtedy:
- f nie dzieli $X - 1 \Rightarrow e > 1$,
- f nie dzieli $X^2 - 1 \Rightarrow e > 2$,
- f nie dzieli $X^3 - 1 \Rightarrow e > 3$,
- $X^4 - 1 = fX + (X^2 + X + 1)$, czyli f nie dzieli $X^4 - 1 \Rightarrow e > 4$,
- $X^5 - 1 = f(X^2 - 1) + (X^2 + X)$, czyli f nie dzieli $X^5 - 1 \Rightarrow e > 5$,
- $X^6 - 1 = f(X^3 + X + 1) + X^2$, a więc f nie dzieli $X^6 - 1 \Rightarrow e > 6$,
- $X^7 - 1 = f(X^4 + X^2 + X)$, a stąd wynika, że $e \leqslant 7$.

Wnosimy, że rząd wielomianu $\operatorname{ord} f = 7$.

W ogólnym przypadku ma miejsce takie twierdzenie (które zostawiamy bez dowodu).

Twierdzenie 7.3.3. *Niech $f, g_1, \ldots, g_l \in \mathbb{F}_q[X]$ będą wielomianami unormowanymi nad ciałem $\mathbb{F}_q$. Wtedy zachodzą następujące własności:*

(1) *jeśli f jest wielomianem nieprzywiedlnym nad ciałem $\mathbb{F}_q$, $f(0) \neq 0$ oraz $m \geqslant 1$ jest liczbą całkowitą, to*

$$\operatorname{ord}(f^m) = p^t \cdot \operatorname{ord} f,$$

gdzie t jest taką najmniejszą liczbą całkowitą, że $p^t \geqslant m$,

(2) *jeśli wielomiany $g_1, \ldots, g_l$ są parami względnie pierwsze, to*

$$\operatorname{ord}(g_1 \cdots g_l) = \operatorname{NWW}(\operatorname{ord} g_1, \ldots, \operatorname{ord} g_l),$$

(3) *jeśli wielomiany $g_1, \ldots, g_l$ są parami względnie pierwsze, $g_i(0) \neq 0$ oraz $0 \neq a \in \mathbb{F}_q$, oraz $m_1, \ldots, m_l$ są dodatnimi liczbami całkowitymi, to*

$$\operatorname{ord}(a g_1^{m_1} \cdots g_l^{m_l}) = p^t \cdot \operatorname{NWW}(\operatorname{ord} g_1, \ldots, \operatorname{ord} g_l),$$

gdzie t jest taką najmniejszą liczbą całkowitą, że $p^t \geqslant \max\{m_1, \ldots, m_l\}$.

Własności rzędu wielomianu opisuje następne

Twierdzenie 7.3.4. *Niech $f \in \mathbb{F}_q[X]$ będą wielomianem nieprzywiedlnym nad ciałem $\mathbb{F}_q$ stopnia $r \geqslant 1$. Wtedy zachodzą własności:*

(1) *wszystkie pierwiastki wielomianu f mają ten sam rząd w grupie $\mathbb{F}_{q^r}^*$,*

(2) $\operatorname{ord} f$ *jest równy rzędowi każdego pierwiastka wielomianu f w grupie $\mathbb{F}_{q^r}^*$,*

(3) $\operatorname{ord} f$ *dzieli $q^r - 1$.*

Dowód. (1) Wszystkie pierwiastki wielomianu f zawierają się w ciele $\mathbb{F}_{q^r}$, a więc ich rzędy są dzielnikami liczby $q^r - 1 = |\mathbb{F}_{q^r}^*|$. Jeśli $\theta \in \mathbb{F}_{q^r}$ jest pierwiastkiem wielomianu f oraz $\mathbb{F}_{q^r}$ ma generator a, to $\theta = a^k$ dla pewnej liczby całkowitej k. Wtedy z twierdzenia 3.1.8(1) wnosimy, że

$$|\theta^{q^i}| = |a^{kq^i}| = \frac{q^r - 1}{\operatorname{NWD}(kq^i, q^r - 1)} = \frac{q^r - 1}{\operatorname{NWD}(k, q^r - 1)} = |a^k| = |\theta|.$$

(2) Jeśli $\theta \in \mathbb{F}_{q^r}$ jest pierwiastkiem wielomianu f oraz $e = |\theta|$ jest jego rzędem w grupie $\mathbb{F}_{q^r}^*$, to θ jest pierwiastkiem wielomianu $X^e - 1 \in \mathbb{F}_q[X]$, a więc f dzieli $X^e - 1$ na podstawie twierdzenia 6.2.1(2). Z innej strony,

skoro $\theta^i \neq 1$ dla każdego i $(1 \leqslant i \leqslant e-1)$, to θ nie jest pierwiastkiem wielomianu $X^i - 1 \in \mathbb{F}_p[X]$, a więc f nie dzieli $X^i - 1$.

(3) Jeśli $\theta \in \mathbb{F}_{q^r}$ oraz $f(\theta) = 0$, to $\theta^{q^r-1} = 1$, a zatem $\operatorname{ord} f$ dzieli $q^r - 1$. □

Wniosek 7.3.5. *Jeśli k jest dodatnią liczbą całkowitą, $f \in \mathbb{F}_q[X]$ oraz $f(0) \neq 0$, to f dzieli $X^k - 1$ wtedy i tylko wtedy, gdy* $\operatorname{ord} f$ *dzieli k.*

Dowód. Niech $e = \operatorname{ord} f$.

($\Rightarrow$) Jeśli $f \mid (X^k - 1)$, to $e \leqslant k$. Z twierdzenia o dzieleniu z resztą $k = eq + r$ oraz $0 \leqslant r < |e| = e$ dla pewnych liczb całkowitych q oraz r. Skoro

$$X^k - 1 = (X^{eq} - 1)X^r + (X^r - 1),$$

to f dzieli $X^r - 1$, a więc $r = 0$. Zatem $k = eq$.

($\Leftarrow$) Jeśli $e \mid k$, to $X^e - 1$ dzieli $X^k - 1$. Wtedy $f \mid (X^k - 1)$, ponieważ $f \mid (X^e - 1)$. □

■ Wielomian $f \in \mathbb{F}_q[X]$ stopnia $r \geqslant 1$ jest nazywany *prymitywnym* (lub *pierwotnym*) nad ciałem $\mathbb{F}_q$, jeśli $f = m_\theta$ jest wielomianem minimalnym pewnego elementu prymitywnego θ ciała $\mathbb{F}_{q^r}$. Z tego wynika, że jednomian niezerowy $aX \in \mathbb{F}_p[X]$ nie jest wielomianem prymitywnym. Poza tym wielomian prymitywny zawsze jest unormowany.

Wniosek 7.3.6. *Wielomian prymitywny $f \in \mathbb{F}_q[X]$ jest nieprzywiedlny nad ciałem $\mathbb{F}_q$.*

Twierdzenie 7.3.7 (kryterium prymitywności wielomianu)**.** *Niech $f \in \mathbb{F}_q[X]$ będzie wielomianem unormowanym stopnia $r \geqslant 1$. Wielomian f jest prymitywny nad ciałem $\mathbb{F}_q$ wtedy i tylko wtedy, gdy $f(0) \neq 0$ oraz* $\operatorname{ord} f = q^r - 1$.

Dowód. ($\Rightarrow$) Załóżmy, że f jest prymitywny nad $\mathbb{F}_q$. Wtedy $f(0) \neq 0$ na mocy wniosku 7.3.6. Skoro f ma pierwiastek $\theta \in \mathbb{F}_{q^r}$ taki, że

$$\mathbb{F}_{q^r}^* = \langle \theta \rangle$$

oraz θ ma rząd $q^r - 1$, to $\operatorname{ord} f = q^r - 1$ na mocy twierdzenia 7.3.4(2).

($\Leftarrow$) Niech $e = \operatorname{ord} f$. Załóżmy, że $e = q^r - 1$. Jeśli założyć, że f nie jest nieprzywiedlny nad ciałem $\mathbb{F}_q$, to mamy takie możliwe przypadki:

• $f = h^s$ jest potęgą wielomianu $h \in \mathbb{F}_q[X]$ nieprzywiedlnego nad ciałem $\mathbb{F}_q$ ($s \geqslant 2$ jest dodatnią liczbą całkowitą) i wtedy z twierdzenia 7.3.3(1) liczba p dzieli $\operatorname{ord} f$. Natomiast $\mathrm{NWD}(p, q^r - 1) = 1$ i mamy sprzeczność.

• $f = h_1 \cdot h_2$ jest iloczynem względnie pierwszych wielomianów $h_1, h_2 \in \mathbb{F}_q[X]$. Jeśli $e_i = \operatorname{ord} h_i$ ($i = 1, 2$), to z twierdzenia 7.3.3(2) oraz (3) wnioskujemy, że $e \leqslant e_1 e_2$. Z lematu 7.3.1 wynika, że $e_i \leqslant q^{\deg h_i} - 1$, a więc

$$e \leqslant (q^{\deg h_1} - 1)(q^{\deg h_2} - 1) < q^r - 1$$

i mamy sprzeczność z założeniem. To znaczy, że wielomian f jest nieprzywiedlny nad ciałem $\mathbb{F}_q$. Wtedy każdy pierwiastek wielomianu f ma rząd e w grupie $(\mathbb{F}_q[X]/\langle f\rangle)^*$ na podstawie twierdzenia 7.3.4(2) i teza zachodzi. □

Wniosek 7.3.8. *Niech $f \in \mathbb{F}_q[X]$. Jeśli $\mathbb{F}_{q^r} = \mathbb{F}_q[X]/\langle f\rangle$ oraz rząd $\operatorname{ord} f = q^r - 1$, to $x = X + \langle f\rangle$ jest elementem prymitywnym ciała $\mathbb{F}_{q^r}$*
(i wtedy

$$\mathbb{F}_{q^r} = \{0, 1, x, x^2, \ldots, x^{q^r-2}\},$$

gdyż $x^{q^r-1} = 1$).

Ćwiczenia 7.3.9.

(**1**) Znaleźć rząd $\operatorname{ord} f$ wielomianu $f \in \mathbb{F}_q[X]$, jeśli:
(a) $q = 3$ oraz $f = X^8 + 1$;
(b) $q = 3$ oraz $f = X^4 + X^3 + X^2 + 2X + 2$;
(c) $q = 2$ oraz $f = X^6 + X + 1$;
(d) $q = 3$ oraz $f = X^6 + X + 1$;
(e) $q = 3$ oraz $f = X^7 + 2X^6 + X^4 + 2X^2 + X$;
(f) $q = 2$ oraz $f = X^3(X^3 + X^2 + 1)(X^2 + X + 1)^4$;
(g) $q = 2$ oraz $f = X^8 + X^7 + X^3 + X + 1$;
(h) $q = 3$ oraz $f = X^7 - X^6 + X^4 - X^2 + X$;
(i) $q = 2$ oraz $f = (X^2 + X + 1)^4(X^3 + X + 1)$.
(**2**) Znaleźć wszystkie wielomiany nieprzywiedlne stopnia n nad ciałem $\mathbb{F}_q$, jeśli:
(a) $q = 2$ oraz $n = 2$;
(b) $q = 2$ oraz $n = 3$;
(c) $q = 3$ oraz $n = 2$;
(d) $q = 3$ oraz $n = 3$.
(**3**) Znaleźć liczbę wielomianów $f \in \mathbb{F}_q[X]$ stopnia $n \geqslant 1$ prymitywnych nad ciałem $\mathbb{F}_q$, jeśli:

(a) $q = 5$ oraz $n = 2$;
(b) $q = 7$ oraz $n = 3$;
(c) $q = 11$ oraz $n = 4$;
(d) $q = 13$ oraz $n = 6$;
(e) $q = 2$ oraz $n = 6$;
(f) $q = 2$ oraz $n = 8$;
(g) $q = 2$ oraz $n = 10$;
(h) $q = 3$ oraz $n = 8$.
(**4**) Niech n będzie dodatnią liczbą całkowitą. Udowodnić, że $X^q - X + a$ jest dzielnikiem wielomianu $X^{q^n} - X + na$, gdzie $a \in \mathbb{F}_q$.
(**5**) Znaleźć elementy prymitywne w ciele:
(a) $\mathbb{F}_5$;
(b) $\mathbb{F}_{11}$;
(c) $\mathbb{F}_{16}$;
(d) $\mathbb{F}_3[X]/\langle x^2 - 2\rangle$;
(e) $\mathbb{F}_5[X]/\langle x^3 + 2\rangle$.
(**6**) Znaleźć wszystkie elementy prymitywne ciała:
(a) $\mathbb{F}_7$;
(b) $\mathbb{F}_9$;
(c) $\mathbb{F}_{11}$;
(d) $\mathbb{F}_{13}$;
(e) $\mathbb{F}_{17}$.
(**7**) Znaleźć rząd wszystkich wielomianów $f \in \mathbb{F}_q[X]$, jeśli:
(a) $q = 2$ oraz $\deg f = 2$;
(b) $q = 2$ oraz $\deg f = 2$;
(c) $q = 3$ oraz $\deg f = 3$.
(**8**) Znaleźć rząd $\operatorname{ord} f$ wielomianu $f \in \mathbb{F}_q[X]$, jeśli:
(a) $q = 2$ oraz $f = (X^2 + X + 1)^3(X^3 + X^2 + 1)$;
(b) $q = 3$ oraz $f = X^4 + X^3 + X^2 + 2X + 1$;
(c) $q = 2$ oraz $f = X^{31} + X^3 + 1$.

Uwagi. Główne wyniki twierdzeń 7.3.1 oraz 7.3.4 są udowodnione przez C. Gaussa.

7.4. Postać macierzowa elementów ciała skończonego

■ Niech f będzie wielomianem unormowanym w postaci „rosnących potęg"

$$f = a_0 + a_1X + \cdots + a_{k-1}X^{k-1} + X^k \in \mathbb{F}_q[X]. \tag{7.2}$$

Macierz

$$A_f = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix} \in M_k(\mathbb{F}_q) \tag{7.3}$$

jest nazywana *macierzą stowarzyszoną z wielomianem* f. Mówi się też, że postać (7.3) jest *postacią kanoniczną Frobeniusa.*

Przykład 7.4.1.

Jeśli wielomian $f = X^3 + 2X^2 + 3X + 4 \in \mathbb{F}_5[X]$, to stowarzyszona z nim macierz ma postać

$$A_f = \begin{bmatrix} 0 & 0 & -4 \\ 1 & 0 & -3 \\ 0 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix} \in M_3(\mathbb{F}_5).$$

Lemat 7.4.2. *Jeśli* $A_f \in M_k(\mathbb{F}_q)$ *jest macierzą postaci* (7.3)*, to jej wyznacznik charakterystyczny*

$$\det(A_f - X \cdot I_k) = (-1)^k f.$$

Dowód. Stosując rozwinięcie Laplace'a względem k-tego wiersza przekonujemy się, że wyznacznik charakterystyczny macierzy A_f jest równy

$$|A_f - X \cdot I_k| = \begin{vmatrix} -X & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & -X & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & -X & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -X & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} - X \end{vmatrix} = (-1)^k f.$$

□

Twierdzenie 7.4.3 (Hamiltona-Cayleya)**.** *Macierz $A \in M_k(\mathbb{F})$ nad ciałem $\mathbb{F}$ jest pierwiastkiem swojego wielomianu charakterystycznego*

$$\det(A - X \cdot I_k) \in \mathbb{F}[X].$$

Dowód. (Szkic) Niech

$$B = A - X \cdot I_k = [b_{ij}] \in M_k(\mathbb{F}[X])$$

oraz $B^D = [B_{ij}] \in M_k(\mathbb{F}[X])$ będzie macierzą dołączoną do macierzy B, czyli B_{ij} jest dopełnieniem algebraicznym do elementu b_{ij} macierzy B (tutaj i dalej $i, j = 1, \ldots, k$). Wtedy

$$B^D = B_0 X^{k-1} + B_1 X^{k-2} + \cdots + B_{k-2} X + B_{k-1},$$

gdzie współczynniki $B_i \in M_k(\mathbb{F})$ są macierzami kwadratowymi stopnia k. Z algebry liniowej wiadomo, że zachodzą równości

$$B^D B = (\det B) \cdot I_k = B B^D.$$

Skoro

$$\det(A - X \cdot I_k) = (-1)^k X^k + a_{k-1} X^{k-1} + a_{k-2} X^{k-2} + \cdots + a_1 X + a_0 \in \mathbb{F}[X]$$

jest wielomianem oraz, w szczególności,

$$(A - X \cdot I_k)(\sum_{i=0}^{k-1} B_i X^{k-1-i}) = B B^D = \det(A - X \cdot I_k) I_k,$$

to otrzymujemy, że odpowiednie współczynniki tych wielomianów

$$\begin{aligned}
-I_k B_0 &= (-1)^k I_k, \\
A B_0 - I_k B_1 &= a_{k-1} I_k, \\
A B_1 - I_k B_2 &= a_{k-2} I_k, \\
\vdots \quad &\vdots \quad \vdots \\
A B_{k-2} - I_k B_{k-1} &= a_1 I_k, \\
A B_{k-1} &= a_0 I_k
\end{aligned}$$

są równe. Stąd mnożąc równości odpowiednio przez $A^k, A^{k-1}, \ldots, A, I_k$ i sumując stronami, otrzymujemy, że ich suma jest równa

$$\begin{gathered}\mathcal{O}_k = A^k \cdot (-I_k B_0) + A^{k-1} \cdot (AB_0 - I_k B_1) + A^{k-2} \cdot (AB_1 - I_k B_2) + \cdots + \\ + A \cdot (AB_{k-2} - I_k B_{k-1}) + A^0 (AB_{k-1}) = \\ = (-1)^k A^k + a_{k-1} A^{k-1} + \cdots + a_1 A + a_0 I_k = \\ = \det(A - X \cdot I_k)_{|X=A},\end{gathered}$$

czyli macierz A jest pierwiastkiem wielomianu $\det(A - X \cdot I_k)$. □

Wniosek 7.4.4. *Jeśli wielomian $f \in \mathbb{F}_q[X]$ ma postać (7.2), to stowarzyszona z nim macierz A_f jest pierwiastkiem tego wielomianu.*

Dowód. Wynika z twierdzenia 7.4.3 oraz lematu 7.4.2. □

Lemat 7.4.5 (istnienie wielomianu minimalnego macierzy)**.** *Niech $A \in M_k(\mathbb{F}_q)$. Wtedy zachodzą następujące własności:*

(1) *istnieje dokładnie jeden wielomian $m_A \in \mathbb{F}_q[X]$ spełniający własności:*
   (a) *wielomian m_A jest unormowany,*
   (b) *$m_A(A) = \mathcal{O}_k \in M_k(\mathbb{F}_q)$ jest macierzą zerową,*
   (c) *wśród wszystkich wielomianów z pierścienia $\mathbb{F}_q[X]$, spełniających warunki (a) oraz (b), m_A jest najmniejszego możliwego stopnia*

   (taki wielomian m_A jest nazywany *wielomianem minimalnym* macierzy A),

(2) jeśli $h \in \mathbb{F}_q[X]$ oraz $h(A) = \mathcal{O}_k$, to m_A dzieli h.

Dowód. Oczywiście, że $M_k(\mathbb{F}_q)$ jest przestrzenią liniową wymiaru k^2 nad ciałem $\mathbb{F}_q$, a więc układ

$$\{I_k, A, A^2, \ldots, A^{k^2}\}$$

jest liniowo zależny nad ciałem $\mathbb{F}_q$. Zatem istnieją takie współczynniki $a_0, a_1, \ldots, a_i, \ldots, a_{k^2} \in \mathbb{F}_q$, że $a_i \neq 0$ dla pewnego i $(0 \leqslant i \leqslant k^2)$ oraz

$$a_0 I_k + a_1 A + a_2 A^2 + \cdots + a_i A^i + \cdots + a_{k^2} A^{k^2} = \mathcal{O}_k.$$

To znaczy, że mamy wielomian niezerowy

$$f = a_0 + a_1 X + a_2 X^2 + \cdots + a_i X^i + \cdots + a_{k^2} X^{k^2} \in \mathbb{F}_q[X]$$

taki, że $f(A) = \mathcal{O}_k$. Zbiór

$$S = \{\deg f \mid 0 \neq f \in \mathbb{F}_q[X] \text{ oraz } f(A) = \mathcal{O}_k\}$$

posiada element minimalny $n \in \mathbb{N}^*$, czyli istnieje wielomian niezerowy $g \in \mathbb{F}_q[X]$ stopnia n taki, że $g(A) = \mathcal{O}_k$. Jeśli $g = a_0X^n + \cdots$, to przyjmujemy

$$m_A = \frac{1}{a_0}g.$$

Wtedy $m_A \in \mathbb{F}_q[X]$ jest unormowany, $m_A(A) = \mathcal{O}_k$ oraz m_A ma najmniejszy możliwy stopień wśród wielomianów $f \in \mathbb{F}_q[X]$ o własności $f(A) = \mathcal{O}_k$.

(2) Załóżmy, że $h \in \mathbb{F}_q[X]$ oraz $h(A) = \mathcal{O}_k$. Ponieważ istnieją takie wielomiany $q, r \in \mathbb{F}_q[X]$, że $h = m_Aq + r$, $\deg r < \deg m_A$ oraz

$$\mathcal{O}_k = h(A) = m_A(A)q(A) + r(A) = r(A),$$

to $r = 0$ jest wielomianem zerowym. Zatem $m_A \mid h$. □

■ Jeśli $f \in \mathbb{F}_q[X]$ jest wielomianem stopnia k postaci (7.2) nieprzywiedlnym nad ciałem $\mathbb{F}_q$ oraz $A = A_f$, to $f = m_A$ oraz

$$\begin{gathered}\mathbb{F}_{q^k} = \mathbb{F}_q[X]/\langle f\rangle = \\ = \{a_0 + a_1X + \cdots + a_{k-1}X^{k-1} + \langle f\rangle \mid a_i \in \mathbb{F}_q \ (i = 0, 1, \ldots, k-1)\} \cong \\ \cong \{a_0I_k + a_1A + \cdots + a_{k-1}A^{k-1} \mid a_i \in \mathbb{F}_q \ (i = 0, 1, \ldots, k-1)\}\end{gathered}$$

jest ciałem o q^k elementach, czyli elementy ciała $\mathbb{F}_{q^k}$ mają prezentację macierzową.

Przykłady 7.4.6.

(**1**) Niech $f = X^2 + X + 1 \in \mathbb{F}_2[X]$. Skoro $f(0) = 1 \neq 0$ oraz $f(1) = 1 \neq 0$, to f jest nieprzywiedlny nad ciałem $\mathbb{F}_2$. Macierz stowarzyszona z wielomianem f ma postać

$$A = A_f = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{F}_2),$$

a więc

$$\mathbb{F}_4 = \{aI_2 + bA \mid a, b \in \mathbb{F}_2\} = \{\mathcal{O}_2, I_2, A, I_2 + A\}.$$

Łatwo przekonać się, że

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad A^3 = I_2,$$

a więc $\mathbb{F}_4^* = \langle A \rangle$ i dlatego A jest elementem prymitywnym ciała $\mathbb{F}_4$.

(**2**) Jeśli $f = X^2 + X - 1 \in \mathbb{F}_3[X]$, to

$$\begin{array}{rcl} f(0) & = & 2 \neq 0, \\ f(1) & = & 1 \neq 0, \\ f(2) & = & 2 \neq 0, \end{array}$$

a zatem f jest nieprzywiedlny nad cialem $\mathbb{F}_3$. Ponieważ

$$A = A_f = \begin{bmatrix} 0 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \in M_2(\mathbb{F}_3),$$

to ciało

$$\mathbb{F}_9 = \{aI_2 + bA \mid a, b \in \mathbb{F}_3\} = \{\mathcal{O}_2, I_2, A, I_2 + A, 2I_2, 2A, 2I_2 + A, I_2 + 2A, 2I_2 + 2A\}.$$

Obliczając kolejne potęgi macierzy A

$$\begin{array}{rcl} A^2 & = & \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} = I_2 + 2A, \\ A^3 & = & \begin{bmatrix} 2 & 2 \\ 2 & 0 \end{bmatrix} = 2I_2 + 2A, \\ A^4 & = & \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2I_2, \\ A^5 & = & \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix} = 2A, \\ A^6 & = & \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = 2I_2 + A, \\ A^7 & = & \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = I_2 + A, \\ A^8 & = & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2, \end{array}$$

przekonujemy się, że A ma rząd 8 w grupie $GL_2(\mathbb{F}_3)$ oraz grupa multiplikatywna $\mathbb{F}_9^* = \langle A \rangle$, czyli A jest elementem prymitywnym ciała $\mathbb{F}_9$.

(**3**) Wielomian $f = X^3 + X + 1 \in \mathbb{F}_2[X]$ jest nieprzywiedlny nad ciałem $\mathbb{F}_2$ oraz jego macierz stowarzyszona ma postać

$$A = A_f = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \in M_3(\mathbb{F}_2).$$

Zatem ciało

$$\begin{gathered} \mathbb{F}_8 = \{aI_3 + bA + cA^2 \mid a, b, c \in \mathbb{F}_2\} = \\ = \{\mathcal{O}_3, I_3, A, A^2, I_3 + A, I_3 + A^2, A + A^2, I_3 + A + A^2\} = \end{gathered}$$

$$= \left\{ \begin{array}{cccc} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, & \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, & \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, & \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} & & \end{array} \right\}.$$

Skoro

$$
\begin{aligned}
A^2 &= \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 0&1&0\\ 0&1&1\\ 1&0&1 \end{bmatrix} = A^2,\\
A^3 &= \begin{bmatrix} 0&1&0\\ 0&1&1\\ 1&0&1 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 1&0&1\\ 1&1&1\\ 0&1&1 \end{bmatrix} = I_3 + A,\\
A^4 &= \begin{bmatrix} 1&0&1\\ 1&1&1\\ 0&1&1 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 0&1&1\\ 1&1&0\\ 1&1&1 \end{bmatrix} = A + A^2,\\
A^5 &= \begin{bmatrix} 0&1&1\\ 1&1&0\\ 1&1&1 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 1&1&1\\ 1&0&0\\ 1&1&0 \end{bmatrix} = I_3 + A + A^2,\\
A^6 &= \begin{bmatrix} 1&1&1\\ 1&0&0\\ 1&1&0 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 1&1&0\\ 0&0&1\\ 1&0&0 \end{bmatrix} = I_3 + A^2,\\
A^7 &= \begin{bmatrix} 1&1&0\\ 0&0&1\\ 1&0&0 \end{bmatrix} \cdot \begin{bmatrix} 0&0&1\\ 1&0&1\\ 0&1&0 \end{bmatrix} = \begin{bmatrix} 1&0&0\\ 0&1&0\\ 0&0&1 \end{bmatrix} = I_3,
\end{aligned}
$$

to

$$\mathbb{F}_8^* = \langle A \rangle$$

oraz A jest elementem prymitywnym ciała $\mathbb{F}_8$. Macierz A ma rząd 7 w grupie liniowej $GL_3(\mathbb{F}_2)$.

Ćwiczenia 7.4.7.

(**1**) Znaleźć prezentację macierzową ciała $\mathbb{F}_{q^k} = \mathbb{F}_q[X]/\langle f \rangle$, jeśli:
(a) $f = X^2 + 2X + 2 \in \mathbb{F}_3[X]$;
(b) $f = X^2 + 1 \in \mathbb{F}_3[X]$;
(c) $f = X^3 + X^2 + 1 \in \mathbb{F}_2[X]$;
(d) $f = X^3 + X + 1 \in \mathbb{F}_2[X]$.
(**2**) Znaleźć wszystkie elementy prymitywne (w postaci macierzowej) z przykładu (**1**).

Uwagi. Pojęcie macierzy zdefiniował J. Sylvester[(2)]. Początkowe wyniki o własnościach macierzy są zawarte w publikacjach A. Cayleya i E. Laguerre'a[(3)].

Twierdzenie 7.4.3 zostało udowodnione przez W. Hamiltona w 1853 r., a później uogólnione przez A. Cayleya w 1858 r. W ogólnym przypadku to twierdzenie udowodnił F. Frobenius w 1878 r.

(2) James Joseph Sylvester (1814–1897)
(3) Edmond Nicolas Laguerre (1834–1886)

7.5. Liniowe ciągi rekurencyjne

■ Niech wielomian $f \in \mathbb{F}_q[X]$ ma postać (7.2). Wtedy równość

$$s_{j+k} = -a_{k-1}s_{j+k-1} - \cdots - a_1 s_{j+1} - a_0 s_j \tag{7.4}$$

jest nazywana (jednorodną) *zależnością rekurencyjną stowarzyszoną z wielomianem* f, gdzie $j = 0, 1, 2, \ldots$

■ Najpierw losowo wybieramy wektor

$$\mathbf{s}_0 = (s_0, s_1, \ldots, s_{k-1}) \in \mathbb{F}_q^k,$$

który jest nazywany *wektorem stanu początkowego*. Wtedy za pomocą zależności (7.4) możemy obliczyć inne wartości s_i ($i \in \mathbb{N}$) i w taki sposób zbudować *liniowy ciąg rekurencyjny*

$$\mathbf{S} = \{s_i\}_{i\in\mathbb{N}}$$

stowarzyszony z wielomianem f. Ciąg $\mathbf{S} = \{s_i\}_{i\in\mathbb{N}}$ będziemy często krótko oznaczać przez $\mathbf{S} = \{s_i\}$. Ciąg $\mathbf{S}$, spełniający warunek (7.4), jest także nazywany *liniowym ciągiem rekurencyjnym rzędu* k *nad ciałem* $\mathbb{F}_q$. Wtedy wektor

$$\mathbf{s}_n = (s_n, s_{n+1}, \ldots, s_{n+k-1}) \in \mathbb{F}_q^k \quad (n \in \mathbb{N})$$

jest nazywany *wektorem* n*-tego stanu*.

■ Ciąg $\mathbf{S} = \{s_i\}_{i\in\mathbb{N}}$ jest nazywany *okresowym*, jeśli istnieją takie dodatnie liczby całkowite t, N, że

$$s_{n+t} = s_n \tag{7.5}$$

zachodzi dla wszystkich $n \geqslant N$ (wtedy liczba t jest nazywana *okresem* ciągu $\mathbf{S}$). Jeśli ciąg $\mathbf{S}$ jest okresowy, to najmniejsza dodatnia liczba całkowita t, spełniająca warunek (7.5), jest nazywana *okresem minimalnym* ciągu $\mathbf{S}$ (wtedy zapisujemy, że $p(\mathbf{S}) = t$).

Lemat 7.5.1. *Minimalny okres* $p(\mathbf{S})$ *liniowego ciągu rekurencyjnego* $\mathbf{S} = \{s_i\}$*, stowarzyszonego z wielomianem postaci* (7.2)*, jest dzielnikiem każdego okresu tego ciągu.*

Dowód. Niech t będzie okresem ciągu $\mathbf{S}$. Wtedy istnieją takie liczby $N_1, N_2 \in \mathbb{N}$, że

$$s_{n+t} = s_n$$

dla $n \geqslant N_1$ oraz

$$s_{n+p(\mathbf{S})} = s_n$$

dla $n \geqslant N_2$. Z twierdzenia o dzieleniu z resztą otrzymujemy, że $t = p(\mathbf{S})q + r$ oraz $0 \leqslant r < p(\mathbf{S})$ dla pewnych liczb całkowitych q, r, a stąd

$$s_n = s_{n+t} = s_{n+p(\mathbf{S})q+r} = s_{n+(p(\mathbf{S})-1)q+r} = \cdots = s_{n+r},$$

czyli r jest okresem ciągu $\mathbf{S}$ oraz $r < p(\mathbf{S})$, a więc $r = 0$. Zatem $t = p(\mathbf{S})q$ i twierdzenie zostało udowodnione. □

Twierdzenie 7.5.2. *Niech k będzie dodatnią liczbą całkowitą. Wtedy każdy (jednorodny) liniowy ciąg rekurencyjny $\mathbf{S} = \{s_i\}$, stowarzyszony z wielomianem $f \in \mathbb{F}_q[X]$ stopnia $k \geqslant 1$ postaci (7.2), jest okresowy z okresem minimalnym $p(\mathbf{S})$ spełniającym warunek $p(\mathbf{S}) \leqslant q^k - 1$.*

Dowód. *a*) Niech

$$A_k = \{\mathbf{s}_l \in \mathbb{F}_q^k \mid 0 \leqslant l \leqslant q^k \text{ oraz } \mathbf{s}_l \text{ jest wektorem } l\text{-tego stanu}\}.$$

Skoro

$$|A_k| = q^k + 1 \quad \text{oraz} \quad |\mathbb{F}_q^k| = q^k,$$

to istnieją takie różne liczby całkowite i, j, że $\mathbf{s}_i = \mathbf{s}_j$ oraz $0 \leqslant i < j \leqslant q^k$. Indukcją względem indeksu m otrzymujemy

$$s_m = s_{m+j-i}$$

dla $m \geqslant i$, co implikuje, że ciąg $\mathbf{S}$ jest okresowy oraz $p(\mathbf{S}) \leqslant j - i$.

b) Jeśli istnieje $l \in \mathbb{N}$ takie, że $\mathbf{s}_l = \mathbf{0}$, to $\mathbf{s}_t = \mathbf{0}$ dla wszystkich liczb całkowitych $t \geqslant l$. Wtedy na podstawie lematu 7.5.1 wnosimy, że $p(\mathbf{S}) = 1 \leqslant q^k - 1$.

c) Załóżmy, że wektor żadnego stanu nie jest zerowy. Wtedy

$$|\mathbb{F}_q^k \setminus \{\mathbf{0}\}| = q^k - 1 \quad \text{oraz} \quad |A_{k-1}| = q^k.$$

Rozumując podobnie jak w części *a*), wnioskujemy, że okres $p(\mathbf{S}) \leqslant q^k - 1$.

Zatem teza zachodzi. □

Wniosek 7.5.3. *Niech* $\mathbf{S}$ *będzie liniowym ciągiem rekurencyjnym stowarzyszonym z wielomianem* $f \in \mathbb{F}_q[X]$ *stopnia* $k \geqslant 1$ *postaci* (7.2). *Jeśli* $f(0) \neq 0$, *to indeks* $N = 0$ (z definicji ciągu okresowego) *jest zerowy.*

Dowód. Skoro $f(0) \neq 0$, to wyraz wolny $a_0 \neq 0$. Z twierdzenia 7.5.2 wynika, że ciąg $\mathbf{S} = \{s_i\}$ jest okresowy z okresem minimalnym $p(\mathbf{S})$, czyli

$$s_{n+p(\mathbf{S})} = s_n$$

dla $n \geqslant N$. Załóżmy nie wprost, że $N \neq 0$. Wtedy biorąc $n = N+p(\mathbf{S})-1$ z zależności (7.4), otrzymujemy

$$\begin{aligned} s_{N+p(\mathbf{S})-1} &= a_0^{-1}(-s_{N+p(\mathbf{S})-1+k} + a_{k-1}s_{N+p(\mathbf{S})-2+k} + \cdots + a_1 s_{N+p(\mathbf{S})} = \\ &= a_0^{-1}(-s_{N+k-1} + a_{k-1}s_{N+k-2} + \cdots + a_1 s_N) = s_{N-1}, \end{aligned}$$

a to jest niemożliwe. Zatem $N = 0$. □

Za pomocą macierzy stowarzyszonej A_f z wielomianem f postaci (7.2) możemy opisać pewne własności liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z tym wielomianem.

Twierdzenie 7.5.4. *Niech* $\mathbf{S} = \{s_i\}$ *będzie liniowym ciągiem rekurencyjnym, stowarzyszonym z wielomianem* $f \in \mathbb{F}_q[X]$ *stopnia* $k \geqslant 1$ *postaci* (7.2), *a* $A_f \in M_k(\mathbb{F}_q)$ *będzie stowarzyszoną z nim macierzą. Wtedy zachodzą własności:*

(1) *wektor n-tego stanu*

$$\mathbf{s}_n = \mathbf{s}_0 A_f^n \quad (n \in \mathbb{N}), \tag{7.6}$$

(2) *jeśli* $f(0) \neq 0$, *to okres minimalny* $p(\mathbf{S})$ *jest dzielnikiem rzędu macierzy* A_f (*jako elementu grupy liniowej* $GL_k(\mathbb{F}_q)$),

(3) *jeśli wielomian* f *jest nieprzywiedlny nad ciałem* $\mathbb{F}_q$ *oraz* $f(0) \neq 0$, *to:*

(a) *rząd* $\operatorname{ord} f$ *wielomianu* f *jest równy rzędowi macierzy* A_f (*jako elementu grupy liniowej* $GL_k(\mathbb{F}_q)$),

(b)

$$s_{n+p(\mathbf{S})} = s_n$$

dla wszystkich $n \in \mathbb{N}$ *oraz* $p(\mathbf{S}) = \operatorname{ord} f$.

Dowód. (1) Ćwiczenie.

(2) Wyznacznik macierzy stowarzyszonej A_f jest równy

$$\det A_f = (-1)^k a_0 \neq 0,$$

a więc macierz $A_f \in GL_k(\mathbb{F}_q)$ jest odwracalna. Załóżmy, że l jest najmniejszą dodatnią liczbą całkowitą taką, że $A_f^l = I_k$. Wtedy biorąc pod uwagę część (1), dostajemy

$$\mathbf{s}_{n+l} = \mathbf{s}_0 A_f^{n+l} = \mathbf{s}_0 A_f^n = \mathbf{s}_n \quad (n \in \mathbb{N}).$$

Z lematu 7.5.1 wynika, że $p(\mathbf{S})$ jest dzielnikiem liczby l.

(3a) Na mocy wniosku 7.4.4 otrzymujemy, że $f = m_{A_f}$ jest wielomianem minimalnym macierzy A_f. Jeśli l jest rzędem macierzy A_f w grupie $GL_k(\mathbb{F}_q)$, to A_f jest pierwiastkiem wielomianu $X^l - 1 \in \mathbb{F}_q[X]$. Z lematu 7.4.5(2) wynika, że m_{A_f} dzieli $X^l - 1$, a stąd $\operatorname{ord} f \leqslant l$.

Jeśli założyć, że $e = \operatorname{ord} f < l$, to $f \mid (X^e - 1)$, a więc $A_f^e = I_k$, sprzeczność. Zatem rząd wielomianu $\operatorname{ord} f = l$.

(3b) Skoro wielomian f jest nieprzywiedlny nad ciałem $\mathbb{F}_q$, to $s_{n+p(\mathbf{S})} = s_n$ dla wszystkich $n \in \mathbb{N}$ na podstawie wniosku 7.5.3. Część (3a) implikuje, że $\operatorname{ord} f$ jest równy rzędowi l macierzy A_f (jako elementu grupy $GL_k(\mathbb{F}_q)$).

Natomiast ciąg $\mathbf{S}$ stowarzyszony z f ma dokładnie $p(\mathbf{S})$ parami różnych stanów $\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{p(\mathbf{S})-1}$, a więc $p(\mathbf{S}) = l$ w wyniku związku (7.6). Zatem $\operatorname{ord} f = p(\mathbf{S})$. □

Ćwiczenia 7.5.5.

(**1**) Sprawdzić, czy wielomian f jest prymitywny nad ciałem $\mathbb{F}_q$, gdzie:
(a) $f = X^5 - X + 1$ oraz $q = 5$;
(b) $f = X^6 + X^5 + X^2 + X + 1$ oraz $q = 2$;
(c) $f = X^8 + X^6 + X^5 + X + 1$ oraz $q = 2$.
(**2**) Zbudować liniowy ciąg rekurencyjny S stowarzyszony z wielomianem $f \in \mathbb{F}_q[X]$ i znaleźć jego okres, jeśli:
(a) $q = 2$ oraz $f = X^6 + X^4 + X^2 + X + 1$;
(b) $q = 2$ oraz $f = X^5 + X + 1$;
(c) $q = 2$ oraz $f = X^7 + X^4 + X^3 + X^2 + 1$;
(d) $q = 2$ oraz $f = X^4 + X^3 + X + 1$;
(e) $q = 2$ oraz $f = X^5 + X^3 + X^2 + 1$;
(f) $q = 2$ oraz $f = X^4 + X^2 + X$;
(g) $q = 2$ oraz $f = X^5 + X^4 + X^2$.

Uwagi. Liniowe ciągi rekurencyjne mają długą historię, co najmniej od 1202 r., kiedy został wprowadzony ciąg Fibonacciego[4] $\{F_j\}_{j\in\mathbb{N}}$, gdzie $F_0 = 0$, $F_1 = 1$ oraz

$$F_{j+2} = F_{j+1} + F_j.$$

S.W. Golomb[5] w 1955 r. pierwszy zaczął szczegółowo badać ciągi maksymalnej długości.

[4] Leonardo z Pizy (Leonardo Bonacci) zwany Fibonaccim (ok. 1170–1240-50)
[5] Solomon Wolf Golomb (1932–2016)

7.6. Postać wektorowa elementów ciała skończonego

■ Jeśli liniowy ciąg rekurencyjny $\mathbf{S} = \{s_i\}_{i\in\mathbb{N}}$, stowarzyszony z wielomianem $f \in \mathbb{F}_q[X]$ stopnia $k \geqslant 1$, ma maksymalny okres (czyli $p(\mathbf{S}) = q^k - 1$), to jest nazywany *ciągiem pseudolosowym* (lub *ciągiem maksymalnego okresu*).

Twierdzenie 7.6.1 (kryterium prymitywności wielomianu)**.** *Niech $f \in \mathbb{F}_q[X]$ będzie wielomianem stopnia $k \geqslant 1$. Wtedy f jest prymitywny nad ciałem $\mathbb{F}_q$ wtedy i tylko wtedy, gdy $f(0) \neq 0$ oraz liniowy ciąg rekurencyjny $\mathbf{S}$, stowarzyszony z f, jest pseudolosowy.*

Dowód. Niech $\mathbf{S}$ będzie liniowym ciągiem rekurencyjnym stowarzyszonym z wielomianem f.

($\Rightarrow$) Z twierdzenia 7.3.7 wynika, że $f(0) \neq 0$ oraz rząd $\operatorname{ord} f = q^k - 1$, a więc $\mathbf{S}$ jest pseudolosowy.

($\Leftarrow$) Załóżmy, że $\mathbf{S}$ jest pseudolosowy. Wtedy $p(\mathbf{S}) = q^k - 1$. Skoro $f(0) \neq 0$, to f jest prymitywny w wyniku twierdzeń 7.5.4 oraz 7.3.4. □

■ Istnieje wiele sposobów przedstawienia elementów ciała skończonego $\mathbb{F}_{p^n}$. Takie przedstawienia są użyteczne, ponieważ ułatwiają różne (w szczególności niezbędne w teorii kodowania) obliczenia w ciele skończonym $\mathbb{F}_{p^n}$.

Niech dalej $f \in \mathbb{F}_p[X]$ będzie wielomianem stopnia $n \geqslant 1$.

1°. Jeśli f jest nieprzywiedlny nad ciałem $\mathbb{F}_p$, to pierścień ilorazowy $\mathbb{F}_p[X]/\langle f\rangle$ jest ciałem (patrz wniosek 4.5.7) o p^n elementach. Jeśli θ jest pierwiastkiem wielomianu f oraz $\theta \in \mathbb{F}_{p^n}$, to (patrz początek dowodu wniosku 7.2.3)

$$\mathbb{F}_{p^n} = \{g(\theta) \mid g \in \mathbb{F}_p[X] \text{ oraz } \deg g < n\} = \mathbb{F}_p(\theta),$$

czyli elementy ciała skończonego $\mathbb{F}_{p^n}$ są przedstawiane jako wartości wszystkich wielomianów stopni $< n$ nad ciałem $\mathbb{F}_p$ w punkcie θ.

2°. Jeśli wielomian f jest prymitywny nad ciałem $\mathbb{F}_p$ oraz $\theta \in \mathbb{F}_{p^n}$ jest jego pierwiastkiem, to

$$\mathbb{F}_{p^n}^* = \langle\theta\rangle,$$

a więc $\theta^{p^n-1} = 1$ oraz

$$\mathbb{F}_{p^n} = \{0,\ 1,\ \theta^1,\ \theta^2,\ \ldots,\ \theta^{p^n-2}\}, \tag{7.7}$$

czyli elementy niezerowe ciała $\mathbb{F}_{p^n}$ są przedstawiane jako potęgi jego elementu prymitywnego θ.

3°. Jeśli f jest nieprzywiedlny nad ciałem $\mathbb{F}_p$ oraz $A_f \in M_n(\mathbb{F}_p)$ jest macierzą stowarzyszoną z nim, to A_f jest pierwiastkiem wielomianu f (patrz wniosek 7.4.4), a więc elementy ciała

$$\mathbb{F}_{p^n} = \{g(A_f) \mid g \in \mathbb{F}_p[X] \text{ oraz } \deg g < n\}$$

przedstawiamy jako wartości wszystkich wielomianów stopni $< n$ nad ciałem $\mathbb{F}_p$ w punkcie A_f, czyli jako pewne macierze z pierścienia $M_n(\mathbb{F}_p)$.

4°. Jeśli wielomian f jest prymitywny nad ciałem $\mathbb{F}_p$ oraz $A = A_f \in M_n(\mathbb{F}_p)$, to $A^{p^n-1} = I_n$ jest macierzą jednostkową stopnia n oraz

$$\mathbb{F}_{p^n} = \{\mathcal{O}_n, I_n, A^1, A^2, \ldots, A^{p^n-2}\}. \tag{7.8}$$

Wtedy, jak wiadomo z definicji wielomianu prymitywnego, elementu prymitywnego ciała, lematu 7.2.6 oraz wniosku 7.2.2, ciało

$$\mathbb{F}_{p^n} = \{B \in M_n(\mathbb{F}_p) \mid B^{p^n} = B\}$$

składa się ze wszystkich macierzy $B \in M_n(\mathbb{F}_p)$ będących pierwiastkami wielomianu $X^{p^n} - X \in \mathbb{F}_p[X]$ i, w szczególności, $\mathcal{O}_n, A \in \mathbb{F}_{p^n}$. Oprócz tego

$$\begin{array}{rcl} (A^i)^{p^n} & = & A^i, \\ (A^i + A^j)^{p^n} & = & A^i + A^j, \\ (\alpha A^i)^{p^n} & = & \alpha A^i \end{array}$$

dla wszystkich $\alpha \in \mathbb{F}_p$ oraz liczb całkowitych i, j. To daje

$$A^i + A^j = A^s$$

dla pewnej liczby całkowitej $s = s(i, j)$ zależnej od i oraz j.

Zatem, podsumowując, mamy następny

Lemat 7.6.2. *Jeśli $f \in \mathbb{F}_p[X]$ jest wielomianem stopnia $n \geqslant 1$ prymitywnym nad ciałem $\mathbb{F}_p$, to*

$$\mathbb{F} = \{A^i \mid i \in \mathbb{Z}\} \bigcup \{\mathcal{O}_n\} = \{\mathcal{O}_n, A^1, A^2, \ldots, A^{p^n-2}, A^{p^n-1} = I_n\}$$

jest ciałem o p^n elementach oraz $\mathbb{F}$ jest przestrzenią liniową nad ciałem $\mathbb{F}_p$.

Lemat 7.6.3. *Jeśli $f \in \mathbb{F}_p[X]$ jest wielomianem stopnia $n \geqslant 1$ prymitywnym nad ciałem $\mathbb{F}_p$ z macierzą stowarzyszoną $A_f \in M_n(\mathbb{F}_p)$, to:*

(1)
$$\varphi : \mathbb{F} \ni A_f^i \mapsto \mathbf{s}_0 A_f^i \in \mathbb{F}_p^n$$
jest homomorfizmem przestrzeni liniowych nad ciałem $\mathbb{F}_p$, gdzie $\mathbf{s}_0$ jest wektorem stanu początkowego liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z wielomianem f (tutaj $\mathbb{F}$ jest ciałem jak w lemacie 7.6.2),

(2) *obraz*
$$\operatorname{Im}\varphi = \{\mathbf{0}, \mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{n-2}\}$$
jest zbiorem składającym się z parami różnych stanów
$$\mathbf{0}, \mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{n-2}$$
liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z wielomianem f oraz wektora zerowego $\mathbf{0} \in \mathbb{F}_p^n$,

(3) *przestrzenie liniowe $\mathbb{F}$ oraz $\operatorname{Im}\varphi$ są izomorficzne (nad ciałem $\mathbb{F}_p$).*

Dowód. Ćwiczenie. □

■ Jeśli:

a) ciało $\mathbb{F}_{p^n}$ postaci (7.7) jest przedstawione jako zbiór składający się z elementu zerowego 0 oraz potęg pierwiastka θ wielomianu $f \in \mathbb{F}_p[X]$ prymitywnego nad ciałem $\mathbb{F}_p$
oraz

b) ciało $\mathbb{F}_{p^n}$ postaci (7.8) jest przedstawione jako zbiór składający się z macierzy zerowej $\mathcal{O}_n$ stopnia n oraz potęg macierzy $A_f \in M_n(\mathbb{F}_p)$ stowarzyszonej z wielomianem prymitywnym $f \in \mathbb{F}_p[X]$ stopnia $n \geqslant 1$ (a przypominamy, że A_f jest pierwiastkiem wielomianu f),

to mamy następujące izomorfizmy przestrzeni liniowych (nad ciałem $\mathbb{F}_p$):
i)

$$\varphi : \mathbb{F}_{p^n} \to \mathbb{F}_p^n$$

takie, że

$$\begin{cases} \varphi(0) &= \mathcal{O}_n \\ \varphi(\theta^i) &= A_f^i \ (i \in \mathbb{N}) \end{cases}$$

oraz
ii)

$$\psi : \mathbb{F}_{p^n} \to \{\mathbf{0}, \mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{n-2}\} \subseteq \mathbb{F}_p^n$$

takie, że

$$\begin{cases} \varphi(\mathcal{O}_n) &= \mathbf{0} \\ \varphi(A_f^i) &= \mathbf{s}_i \ (i \in \mathbb{N}), \end{cases}$$

gdzie

$$\mathbb{F}_{p^n} = \{\mathcal{O}_n, A_f^1, A_f^2, \ldots, A_f^{p^n-2}, A_f^{p^n-1} = I_n\}.$$

Zatem elementy ciała $\mathbb{F}_{p^n}$ (rozpatrywanego jako przestrzeń liniowa nad ciałem $\mathbb{F}_p$) możemy przedstawiać jako kolejne stany ciągu pseudolosowego $\mathbf{S}$ stowarzyszonego z wielomianem $f \in \mathbb{F}_p[X]$ prymitywnym nad ciałem $\mathbb{F}_p$ w następujący sposób:

$$\begin{cases} \mathbb{F}_{p^n} \ni 0 &= \mathbf{0} \in \mathbb{F}_p^n, \\ \mathbb{F}_{p^n} \ni \theta^i &= \mathbf{s}_i \in \mathbb{F}_p^n, \end{cases}$$

gdzie $i \in \mathbb{N}$.

Zilustrujmy to na takich przykładach.

Przykłady 7.6.4.

(**1**) Sprawdźmy, czy wielomian $f = X^4 + X + 1 \in \mathbb{F}_2[X]$ jest prymitywny nad ciałem $\mathbb{F}_2$. W tym celu budujemy zależność rekurencyjną stowarzyszoną z f:

- $X^4 = -X - 1$,
- $X^4 = X + 1$,
- $s_{j+4} = s_{j+1} + s_j$.

Losowo wybieramy wartości początkowe (ale różne) z ciała $\mathbb{F}_2$:

$$\begin{aligned} s_0 &= 1, \\ s_1 &= 0, \\ s_2 &= 0, \\ s_3 &= 1 \end{aligned}$$

i konstruujemy liniowy ciąg rekurencyjny $\mathbf{S} = \{s_i\}_{i\in\mathbb{N}}$. Ciąg $\mathbf{S}$ może mieć okres $p(\mathbf{S}) \leqslant 2^4 - 1 = 15$. Obliczając wartości s_i, otrzymujemy

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
1	0	0	1	1	0	1	0	1	1	1	1

s_{12}	s_{13}	s_{14}	s_{15}	s_{16}	s_{17}	s_{18}
0	0	0	1	0	0	1

Zatem $p(\mathbf{S}) = 15$, ciąg $\mathbf{S}$ jest pseudolosowy i wielomian f jest prymitywny nad ciałem $\mathbb{F}_2$. Jeśli $\mathbb{F}_{16}^* = \langle\theta\rangle$ ($\theta \in \mathbb{F}_{16}$ jest pierwiastkiem wielomianu f), to możemy przedstawić elementy ciała $\mathbb{F}_{16}$ w postaci wektorowej:

$$\begin{aligned}
0 &= 0000,\\
1 &= 1001,\\
\theta^1 &= 0011,\\
\theta^2 &= 0110,\\
\theta^3 &= 1101,\\
\theta^4 &= 1010,\\
\theta^5 &= 0101,\\
\theta^6 &= 1011,\\
\theta^7 &= 0111,\\
\theta^8 &= 1111,\\
\theta^9 &= 1110,\\
\theta^{10} &= 1100,\\
\theta^{11} &= 1000,\\
\theta^{12} &= 0001,\\
\theta^{13} &= 0010,\\
\theta^{14} &= 0100,\\
\theta^{15} &= 1001 = 1.
\end{aligned}$$

(**2**) Z wielomianem $f = X^2 + X + 2 \in \mathbb{F}_3[X]$ jest związana zależność rekurencyjna

$$s_{j+2} = 2s_{j+1} + s_j$$

generująca liniowy ciąg rekurencyjny $\mathbf{S} = \{s_i\}$, który może mieć okres minimalny co najwyżej $\leqslant 3^2 - 1 = 8$. Znajdźmy jego okres minimalny. Wybierzmy losowo dwie wartości początkowe $s_0 = 1$ oraz $s_1 = 0$ w ciele $\mathbb{F}_3$. W wyniku obliczeń wypełniamy taką tablicę:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
1	0	1	2	2	0	2	1	1	0	1

Zatem ciąg $\mathbf{S}$ jest pseudolosowy (o okresie $p(\mathbf{S}) = 8$) i możemy przedstawić elementy ciała $\mathbb{F}_9$ w postaci wektorowej w taki sposób:

$$\begin{aligned}
0 &= 00,\\
1 &= 10,\\
\theta^1 &= 01,\\
\theta^2 &= 12,\\
\theta^3 &= 22,\\
\theta^4 &= 20,\\
\theta^5 &= 02,\\
\theta^6 &= 21,\\
\theta^7 &= 11,\\
\theta^8 &= 10 = 1.
\end{aligned}$$

(**3**) Z wielomianem $g = X^4 + X^3 + X \in \mathbb{F}_2[X]$ jest związana zależność rekurencyjna

$$s_{j+4} = s_{j+3} + s_{j+1}.$$

Biorąc wartości początkowe $s_0 = 1, s_1 = 1, s_2 = 0, s_3 = 1$ (bo $\deg g = 4$) w ciele $\mathbb{F}_2$ (a ma ich być tyle, ile wynosi stopień $\deg f$), obliczamy kolejne wartości liniowego ciągu **S** stowarzyszonego z wielomianem g:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}
1	1	0	1	0	0	1	1	1	0	1	0

Zatem

$$p(\mathbf{S}) = 7 \neq 15 = 2^{\deg f} - 1,$$

a więc ciąg **S** nie jest pseudolosowy i wielomian f nie jest prymitywny.

* * *

■ **Logarytmy Jacobiego-Zecha.** Niech θ będzie elementem prymitywnym ciała $\mathbb{F}_q$, a więc $\theta^{q-1} = 1$. Jeśli przyjąć, że

$$\theta^{-\infty} = 0,$$

to każdy element ciała $\mathbb{F}_q$ będzie pewną potęgą elementu θ. Wtedy dla dodatnich liczb całkowitych n takich, że $\theta^n + 1 \neq 0$ możemy zdefiniować *logarytmy Jacobiego-Zecha* $L(n)$ w taki sposób:

$$\theta^{L(n)} = \theta^n + 1.$$

Ponadto

- jeśli $\operatorname{char} \mathbb{F}_q = 2$, to przyjmujemy, że $L(0) = -\infty$,
- jeśli $\operatorname{char} \mathbb{F}_q > 2$, to $L(\frac{q-1}{2}) = -\infty$.

Jeśli $n \equiv m \pmod{q-1}$, to $L(n) = L(m)$, a zatem dalej zamiast $L(n)$ będziemy pisać $L(n \pmod{q-1})$.

Twierdzenie 7.6.5. *Niech $q = p^m$, $m, k, n \in \mathbb{N}^*$, $i \in \mathbb{N}$, p będzie liczbą pierwszą oraz $\mathbb{F}_q^* = \langle\theta\rangle$. Wtedy logarytmy Jacobiego-Zecha spełniają warunki:*

(1) $\theta^n + \theta^k = \theta^{n+L(k-n)}$, *gdzie $k > n$;*

(2)

$$L((q-1-n)p^i \pmod{q-1}) = (L(n) - n)p^i \pmod{q-1}; \quad (7.9)$$

(3) $\theta^{L(n)p} = \theta^{L(np)}$;
(4)

$$L(np^i \pmod{q-1}) = L(n)p^i \pmod{q-1}. \tag{7.10}$$

Dowód. (1) Rzeczywiście zachodzą równości

$$\theta^n + \theta^k = \theta^n(1+\theta^{k-n}) = \theta^n\theta^{L(k-n)} = \theta^{n+L(k-n)}.$$

(2) Skoro

$$\theta^{L(n)-n} = (\theta^n+1)\theta^{-n} = 1+\theta^{-n}$$

oraz $\theta^{q-1} = 1$, to $\theta^{L(n)-n} = 1+\theta^{q-1-n}$ oraz

$$\left(\theta^{L(n)-n}\right)^{p^i} = (1+\theta^{q-1-n})^{p^i} = 1+\theta^{(q-1-n)p^i} = \theta^{L((q-1-n)p^i)},$$

a stąd wynika teza.
(3) Mamy

$$\theta^{L(n)p} = (1+\theta^n)^p = 1+\theta^{np} = \theta^{L(np)}.$$

(4) Wynika z własności (3). □

■ Za pomocą wzorów (7.9) oraz (7.10) nie zawsze możemy znaleźć wartości logarytmów Jacobiego-Zecha dla wszystkich elementów ciała skończonego. Czasem przed stosowaniem wzorów (7.9) oraz (7.10) wynika konieczność obliczenia wartości logarytmów Jacobiego-Zecha dla niektórych elementów ciała skończonego w inny sposób.

Przykłady 7.6.6.

(**1**) Znajdźmy wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_9$. Ponieważ ciało $\mathbb{F}_9$ jest generowane przez pierwiastek θ wielomianu prymitywnego f stopnia 2 nad ciałem $\mathbb{F}_3$ oraz takim wielomianem jest $f = X^2+X+2 \in \mathbb{F}_3[X]$, to $\theta^2+\theta+2=0$,

$$\begin{array}{rcl} \theta^2 & = & -\theta-2 = 2\theta+1, \\ \theta^4 & = & (2\theta+1)^2 = 2 = -1 \end{array}$$

oraz

$$\theta^{L(2)} = \theta^2+1 = 2(\theta+1) = -(\theta+1) = -\theta^{L(1)} = \theta^{4+L(1)},$$

czyli $L(2) = 4+L(1)$. Podobnie

$$\theta^{L(6)} = \theta^6+1 = (\theta^2+1)^3 = (-(\theta+1))^3 = -(\theta^3+1) = -\theta^{L(3)} = \theta^{4+L(3)},$$

czyli $L(6) = 4 + L(3)$. Oprócz tego $\frac{q-1}{2} = \frac{9-1}{2} = 4$, a zatem

$$L(4) = -\infty.$$

• Ze wzoru (7.10) dla $i = 1$ oraz $n = 2,\ q = 9$ otrzymujemy

$$4 + L(3) = L(6) = L(2) \cdot 3 = 12 + 3L(1) \pmod 8 = 4 + 3L(1),$$

na podstawie czego $L(3) = 3L(1) = 3(L(2) - 4) = 3L(2) - 4 \pmod 8$ oraz $L(6) = 3L(2)$.

• Z zależności (7.9) dla $i = 0$ oraz $n = 2,\ q = 9$ dostajemy $3L(2) = L(6) = L(2) - 2$, a zatem modulo 8 mamy

$$\begin{array}{rcl} L(2) & = & 3, \\ L(1) & = & 7, \\ L(6) & = & 1, \\ L(3) & = & 5. \end{array}$$

Ponieważ

$$\theta^4 = (\theta^2)^2 = (2\theta + 1)^2 = 4\theta^2 + 4\theta + 1 = \theta^2 + \theta + 1 = 2,$$

to $\theta^{L(5)} = \theta^5 + 1 = \theta^4 \cdot \theta + 1 = 2\theta + 1 = \theta^2$ i mamy

$$L(5) = 2.$$

Skoro $\theta^8 = (\theta^4)^2 = 2^2 = 1$, to $\theta^{L(8)} = 1 + \theta^8 = 2 = \theta^4$, w wyniku czego

$$L(8) = 4.$$

Z zależności (7.9) dla $n = 1$ oraz $i = 0$ także wnioskujemy, że $L(7) = L(9 - 1 - 1) = L(1) - 1 \pmod 8 = 6$, czyli

$$L(7) = 6.$$

(**2**) Znajdźmy wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_8$. Ciało $\mathbb{F}_8$ jest generowane przez pierwiastek θ wielomianu $f = X^3 + X + 1 \in \mathbb{F}_2[X]$ prymitywnego nad ciałem $\mathbb{F}_2$ oraz $\theta^3 = \theta + 1$. Zatem

$$L(1) = 3.$$

Poza tym we wzorach (7.9) i (7.10) mamy $q = 8$, $n = 1$ oraz $p = 2$ i wtedy:

• z wykorzystaniem wzoru (7.9) obliczamy:

$$\begin{array}{lllllll} i = 0 & \Rightarrow & L(6 \pmod 7) = L(1) - 1 \pmod 7 = 2 & & \Rightarrow & L(6) = 2, \\ i = 1 & \Rightarrow & L(6 \cdot 2^1 \pmod 7) = (L(1) - 1)2^1 \pmod 7 = 4 & & \Rightarrow & L(5) = 4, \\ i = 2 & \Rightarrow & L(6 \cdot 2^2 \pmod 7) = (L(1) - 1)2^2 \pmod 7 = 2 & & \Rightarrow & L(3) = 1; \end{array}$$

• z wykorzystaniem wzoru (7.10) obliczamy:

$$\begin{array}{lllllll} i = 1 & \Rightarrow & L(1 \cdot 2^1 \pmod 7) = L(1)2^1 \pmod 7 = 6 & & \Rightarrow & L(2) = 6, \\ i = 2 & \Rightarrow & L(1 \cdot 2^2 \pmod 7) = L(1)2^2 \pmod 7 = 5 & & \Rightarrow & L(4) = 5. \end{array}$$

Dla innych indeksów i wartości $L(6\cdot 2^i \pmod 7)$ powtarzają się. Zatem podsumowujemy wynik w takiej tabelce:

θ^i	i	$L(i)$
000	$-\infty$	0
100	0	$-\infty$
010	1	3
001	2	6
110	3	1
011	4	5
111	5	4
101	6	2

(**3**) Obliczmy logarytmy Jacobiego-Zecha dla elementów ciała $\mathbb{F}_{16} = \mathbb{F}_2[X]/\langle X^4 + X + 1\rangle$. Niech $\theta \in \mathbb{F}_{16}$ będzie pierwiastkiem wielomianu $X^4 + X + 1 \in \mathbb{F}_2[X]$, który jest prymitywny nad ciałem $\mathbb{F}_2$. Wtedy (patrz postacie wektorowe elementów ciała $\mathbb{F}_{16}$ obliczone w przykładzie 7.6.4(1)) otrzymujemy:

$$
\begin{array}{lllcll}
1+\theta = L(1) & \text{oraz } 1+\theta = 1001+0011 = 1010 & = & \theta^4 & \Rightarrow & L(1)=4,\\
1+\theta^2 = L(2) & \text{oraz } 1+\theta^2 = 1001+0110 = 1111 & = & \theta^8 & \Rightarrow & L(2)=8,\\
1+\theta^3 = L(3) & \text{oraz } 1+\theta^3 = 1001+1101 = 0100 & = & \theta^{14} & \Rightarrow & L(3)=14,\\
1+\theta^4 = L(4) & \text{oraz } 1+\theta^4 = 1001+1010 = 0011 & = & \theta & \Rightarrow & L(4)=1,\\
1+\theta^5 = L(5) & \text{oraz } 1+\theta^5 = 1001+0101 = 1100 & = & \theta^{10} & \Rightarrow & L(5)=10,\\
1+\theta^6 = L(6) & \text{oraz } 1+\theta^6 = 1001+1011 = 0010 & = & \theta^{13} & \Rightarrow & L(6)=13,\\
1+\theta^7 = L(7) & \text{oraz } 1+\theta^7 = 1001+0111 = 1110 & = & \theta^9 & \Rightarrow & L(7)=9,\\
1+\theta^8 = Z(8) & \text{oraz } 1+\theta^8 = 1001+1111 = 0110 & = & \theta^2 & \Rightarrow & L(8)=2,\\
1+\theta^9 = Z(9) & \text{oraz } 1+\theta^9 = 1001+1110 = 0111 & = & \theta^7 & \Rightarrow & L(9)=7,\\
1+\theta^{10} = L(10) & \text{oraz } 1+\theta^{10} = 1001+1100 = 0101 & = & \theta^5 & \Rightarrow & L(10)=5,\\
1+\theta^{11} = L(11) & \text{oraz } 1+\theta^{11} = 1001+1000 = 0001 & = & \theta^{12} & \Rightarrow & L(11)=12,\\
1+\theta^{12} = L(12) & \text{oraz } 1+\theta^{12} = 1001+0001 = 1000 & = & \theta^{11} & \Rightarrow & L(12)=11,\\
1+\theta^{13} = L(13) & \text{oraz } 1+\theta^{13} = 1001+0010 = 1011 & = & \theta^6 & \Rightarrow & L(13)=6,\\
1+\theta^{14} = L(14) & \text{oraz } 1+\theta^{14} = 1001+0100 = 1101 & = & \theta^3 & \Rightarrow & L(14)=3.
\end{array}
$$

Oprócz tego $L(-\infty) = 0$ oraz $L(0) = -\infty$.

■ Za pomocą logarytmów Jacobiego-Zecha można zręcznie obliczać sumy elementów w ciele skończonym $\mathbb{F}_q = \mathbb{F}_p(\theta)$, które jest generowane przez element prymitywny θ, stosując wzór $(k > n)$

$$\theta^k + \theta^n = \theta^{n+L(k-n)} \pmod{q-1}.$$

Na przykład w ciele $\mathbb{F}_9 = \mathbb{F}_3(\theta)$ obliczamy

$$\theta^6 + \theta^5 = \theta^{5+L(6-5)} = \theta^{12} \pmod 8 = \theta^4$$

oraz

$$\theta^2 + \theta^4 = \theta^{2+L(4-2)} = \theta^5.$$

* * *

■ **Ślad i norma elementów ciała skończonego.** Niech $\theta \in \mathbb{F}_{q^n}$. Wtedy jego *śladem* (z ciała $\mathbb{F}_{q^n}$ do ciała $\mathbb{F}_q$) jest nazywany element

$$\mathrm{tr}_{q^n/q}(\theta) = \sum_{i=0}^{n-1} \theta^{q^i}.$$

Jeśli q jest liczbą pierwszą, to wartość $\mathrm{tr}_{q^n/q}(\theta)$ jest nazywana *śladem bezwzględnym* elementu θ. Jeśli $m_\theta \in \mathbb{F}_q[X]$ jest wielomianem minimalnym elementu θ, to m_θ dzieli $X^{q^n} - X$ (a więc $\deg m_\theta$ dzieli n na podstawie wniosku 7.2.3). Wielomian

$$g = (m_\theta)^{\frac{n}{m}} \in \mathbb{F}_q[X] \text{ (tutaj } \deg m_\theta = m \geqslant 1)$$

jest nazywany wielomianem *charakterystycznym* elementu θ nad ciałem $\mathbb{F}_q$. Skoro pierwiastkami wielomianu m_θ są

$$\theta, \theta^q, \ldots, \theta^{q^{m-1}},$$

to pierwiastkami wielomianu g są elementy

$$\theta, \theta^q, \ldots, \theta^{q^{n-1}},$$

a więc

$$g = \prod_{i=0}^{n-1} (X - \theta^{q^i}) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{F}_q[X] \tag{7.11}$$

dla pewnych współczynników $a_{n-1}, \ldots, a_0 \in \mathbb{F}_q$.

Opiszmy podstawowe własności śladu.

Twierdzenie 7.6.7. *Niech $\theta \in \mathbb{F}_{q^n}$. Wtedy zachodzą następujące własności:*

(1) $\mathrm{tr}_{q^n/q}(\theta^q) = \mathrm{tr}_{q^n/q}(\theta)$;
(2) $\mathrm{tr}_{q^n/q}(\theta) = -a_{n-1} \in \mathbb{F}_q$;
(3) $\mathrm{tr}_{q^n/q} : \mathbb{F}_{q^n} \ni \theta \mapsto \mathrm{tr}_{q^n/q}(\theta) \in \mathbb{F}_q$ *jest $\mathbb{F}_q$-liniowym odwzorowaniem;*
(4) *odwzorowanie* $\mathrm{tr}_{q^n/q}$ *jest suriekcją;*
(5) $\mathrm{tr}_{q^n/q}(a) = na$ *dla każdego elementu* $a \in \mathbb{F}_q$;

(6) (przechodniość śladu) *dla każdego rozszerzenia ciał* $\mathbb{F}_q \leqslant \mathbb{F}_{q^n} \leqslant \mathbb{F}_{q^l}$ *oraz elementu* $\vartheta \in \mathbb{F}_{q^l}$ *zachodzi*

$$\mathrm{tr}_{q^l/q}(\vartheta) = \mathrm{tr}_{q^n/q}(\mathrm{tr}_{q^l/q^n}(\vartheta)).$$

Dowód. (1) Stosując wniosek 7.1.6, otrzymujemy

$$\mathrm{tr}_{q^n/q}(\theta^q) = \theta^q + \theta^{q^2} + \cdots + \theta^{q^n} = \mathrm{tr}_{q^n/q}(\theta).$$

(2) Porównując współczynniki w (7.11), otrzymujemy wynik.

(3) Przekonujemy się, że dla dowolnych elementów $\alpha, \beta \in \mathbb{F}_{q^n}$ oraz $a \in \mathbb{F}_q$ zachodzą związki:

- $$\mathrm{tr}_{q^n/q}(\alpha+\beta) = \sum_{i=0}^{n-1}(\alpha+\beta)^{q^i} = \sum_{i=0}^{n-1}(\alpha^{q^i}+\beta^{q^i}) =$$
$$= \sum_{i=0}^{n-1}\alpha^{q^i} + \sum_{i=0}^{n-1}\beta^{q^i} = \mathrm{tr}_{q^n/q}(\alpha) + \mathrm{tr}_{q^n/q}(\beta);$$
- $$\mathrm{tr}_{q^n/q}(a\cdot\alpha) = \sum_{i=0}^{n-1}(a\alpha)^{q^i} = \sum_{i=0}^{n-1}a^{q^i}\alpha^{q^i} =$$
$$= \sum_{i=0}^{n-1}a\alpha^{q^i} = a\cdot\mathrm{tr}_{q^n/q}(\alpha),$$

a więc teza zachodzi.

(4) Zauważamy, że $\mathrm{tr}_{q^n/q}(\mu) = 0$ wtedy i tylko wtedy, gdy $\mu \in \mathbb{F}_{q^n}$ jest pierwiastkiem wielomianu

$$f = X^{q^{n-1}} + \cdots + X^q + X \in \mathbb{F}_q[X].$$

Skoro $\deg f = q^{n-1}$, to ten wielomian ma co najwyżej q^{n-1} pierwiastków w ciele $\mathbb{F}_{q^n}$ (które składa się z q^n elementów). To implikuje, że znajdzie się element $\theta \in \mathbb{F}_{q^n}$ taki, że $f(\theta) \neq 0$, co zezwala wnioskować, że odwzorowanie $\mathrm{tr}_{q^n/q}$ jest suriekcją.

(5) Wynika na podstawie definicji i wniosku 7.1.6.

(6) W rzeczy samej, n dzieli l, a więc $\frac{l}{n} = m \in \mathbb{N}$ oraz

$$\mathrm{tr}_{q^n/q}(\mathrm{tr}_{q^l/q^n}(\vartheta)) = \sum_{i=0}^{n-1}\mathrm{tr}_{q^l/q^n}(\vartheta)^{q^i} = \sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1}\vartheta^{q^{nj}}\right)^{q^i} =$$
$$= \sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\vartheta^{q^{nj+i}} = \sum_{s=0}^{nm-1}\vartheta^{q^s} = \mathrm{tr}_{q^l/q}(\vartheta).$$

□

Za pomocą funkcji śladu możemy teraz opisać wszystkie funkcjonały postaci $\mathbb{F}_{q^n} \to \mathbb{F}_q$.

Stwierdzenie 7.6.8. *Niech $n \geqslant 1$ będzie liczbą całkowitą. Wtedy zachodzą następujące własności:*

(1) *dla każdego $\mathbb{F}_q$-liniowego odwzorowania $l_\nu : \mathbb{F}_{q^n} \to \mathbb{F}_q$ istnieje taki element $\nu \in \mathbb{F}_{q^n}$, że*

$$l_\nu : \mathbb{F}_{q^n} \ni e \mapsto \mathrm{tr}_{q^n/q}(e\nu) \in \mathbb{F}_q;$$

(2) *jeśli μ, ν są różnymi elementami ciała $\mathbb{F}_{q^n}$, to $l_\mu \neq l_\nu$.*

Dowód. (1) Istotnie znajdzie się element $\theta \in \mathbb{F}_{q^n}$ taki, że

$$l_\mu(\theta) - l_\nu(\theta) = \mathrm{tr}_{q^n/q}(\mu\theta) - \mathrm{tr}_{q^n/q}(\nu\theta) = \mathrm{tr}_{q^n/q}((\mu - \nu)\theta) \neq 0,$$

a więc $l_\mu \neq l_\nu$. Zatem moc

$$|\{l_\mu \mid \mu \in \mathbb{F}_{q^n}\}| = |\mathbb{F}_{q^n}| = q^n.$$

(2) Niech $l : \mathbb{F}_{q^n} \to \mathbb{F}_q$ będzie dowolnym $\mathbb{F}_q$-liniowym odwzorowaniem oraz $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ będzie bazą przestrzeni liniowej $\mathbb{F}_{q^n}$ nad ciałem $\mathbb{F}_q$. Skoro obraz $l(\mathbf{e}_i) \in \mathrm{Lin}_{\mathbb{F}_q}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, to istnieje q^n wszystkich $\mathbb{F}_q$-liniowych odwzorowań postaci $\mathbb{F}_{q^n} \to \mathbb{F}_q$. Zatem dla każdego $\mathbb{F}_q$-liniowego odwzorowania l istnieje $\nu \in \mathbb{F}_{q^n}$ takie, że $l = l_\nu$. □

■ Niech $\theta \in \mathbb{F}_{q^n}$. *Normą* elementu θ (z ciała $\mathbb{F}_{q^n}$ do ciała $\mathbb{F}_q$) jest nazywany element

$$N_{q^n/q}(\theta) = \prod_{i=0}^{n-1} \theta^{q^i} = \theta^{\frac{q^n-1}{q-1}}.$$

Twierdzenie 7.6.9. *Niech $\theta \in \mathbb{F}_{q^n}$. Wtedy zachodzą następujące własności:*

(1) *$N_{q^n/q}(\theta) = (-1)^n a_0 \in \mathbb{F}_q$, gdzie a_0 jest wyrazem wolnym wielomianu charakterystycznego g postaci* (7.11)*;*

(2) *$N_{q^n/q}(\theta^q) = N_{q^n/q}(\theta)$;*

(3) *$N_{q^n/q}(a) = a^n$ dla każdego $a \in \mathbb{F}_q$;*

(4) $N_{q^n/q}(\theta) = 0$ *w tym i tylko tym przypadku, gdy* $\theta = 0$*;*
(5) $N_{q^n/q} : \mathbb{F}_{q^n}^* \ni \theta \mapsto N_{q^n/q}(\theta) \in \mathbb{F}_q^*$ *jest epimorfizmem grup;*
(5′) $N_{q^n/q} : \mathbb{F}_{q^n} \to \mathbb{F}_q$ *jest suriekcją;*
(6) (przechodniość) *dla każdego rozszerzenia ciał* $\mathbb{F}_q \leqslant \mathbb{F}_{q^n} \leqslant \mathbb{F}_{q^l}$ *oraz elementu* $\theta \in \mathbb{F}_{q^l}$ *jego norma*

$$N_{q^l/q}(\theta) = N_{q^n/q}(N_{q^l/q^n}(\theta)).$$

Dowód. (1) Wynika z definicji normy i wielomianu charakterystycznego.
(2) Ponieważ $N_{q^n/q}(\theta) \in \mathbb{F}_q$, to

$$N_{q^n/q}(\theta^q) = \big(N_{q^n/q}(\theta)\big)^q = N_{q^n/q}(\theta).$$

(3) Skoro $a^q = a$ dla każdego $a \in \mathbb{F}_q$ na podstawie wniosku 7.1.6, to otrzymujemy tezę.
(4) Ćwiczenie.
(5) Z definicji normy otrzymujemy

$$N_{q^n/q}(\alpha \cdot \beta) = N_{q^n/q}(\alpha) \cdot N_{q^n/q}(\beta)$$

dla dowolnych $\alpha, \beta \in \mathbb{F}_{q^n}$, czyli $N_{q^n/q}$ jest homomorfizmem grup. Jądro tego homomorfizmu

$$\operatorname{Ker} N_{q^n/q} = \left\{\alpha \in \mathbb{F}_{q^n} \mid \alpha \text{ jest pierwiastkiem } X^{\frac{q^n-1}{q-1}} - 1 \in \mathbb{F}_q[X]\right\},$$

a więc jego rząd r (jako grupy) spełnia warunek

$$r \leqslant \frac{q^n - 1}{q - 1}.$$

Lecz na mocy twierdzenia Lagrange'a rząd obrazu

$$\left|\operatorname{Im} N_{q^n/q}\right| = \frac{q^n - 1}{r} \geqslant q - 1.$$

Na tej podstawie wnosimy, że $N_{q^n/q}$ jest suriekcją i teza zachodzi.
(5′) Wynika z części (5).
(6) Rzeczywiście sprawdzamy, że

$$N_{q^n/q}(N_{q^l/q^n}(\theta)) = N_{q^n/q}(\theta^{\frac{q^l-1}{q^n-1}}) = \left(\theta^{\frac{q^l-1}{q^n-1}}\right)^{\frac{q^n-1}{q-1}} = \theta^{\frac{q^l-1}{q-1}} = N_{q^l/q}(\theta).$$

□

Zaznaczmy też takie

Twierdzenie 7.6.10 (Hilberta 90)**.** *Niech $\theta \in \mathbb{F}_{q^n}$. Wtedy zachodzą następujące własności:*

(1) $\mathrm{tr}_{q^n/q}(\theta) = 0$ *w tym i tylko tym przypadku, gdy* $\theta = \alpha^q - \alpha$ *dla pewnego* $\alpha \in \mathbb{F}_{q^n}$*;*

(2) *jeśli* $0 \neq \beta \in \mathbb{F}_{q^n}$*, to*

$$N_{q^n/q}(\beta) = 1 \quad \Leftrightarrow \quad \beta = \frac{\theta^q}{\theta} \text{ dla pewnego } \theta \in \mathbb{F}_{q^n}^*.$$

Dowód. (1) ($\Leftarrow$) Wynika na podstawie twierdzenia 7.6.7(1).

($\Rightarrow$) Niech α będzie pierwiastkiem wielomianu $f = X^q - X - \theta \in \mathbb{F}_{q^n}[X]$ zawierającym się w pewnym rozszerzeniu ciała $\mathbb{F}_{q^n}$. Wtedy $\alpha^q - \alpha = \theta$ oraz

$$0 = \mathrm{tr}_{q^n/q}(\theta) = \sum_{i=0}^{n-1} \theta^{q^i} = \sum_{i=0}^{n-1} (\alpha^q - \alpha)^{q^i} = \sum_{i=0}^{n-1} (\alpha^{q^{i+1}} - \alpha^{q^i}) = \alpha^{q^n} - \alpha.$$

Zatem $\alpha \in \mathbb{F}_{q^n}$.

(2) Ćwiczenie. □

Przykłady 7.6.11.

(**1**) Obliczmy ślady i normy elementów x oraz $1 + x$ ciała $\mathbb{F}_4$ z przykładu 4.5.8(1) (tutaj $q = 2$ oraz $n = 2$):

- ślad

$$\mathrm{tr}_{4/2}(x) = x^{2^0} + x^{2^1} = x + x + 1 = 1$$

oraz

$$\mathrm{tr}_{4/2}(1 + x) = (1 + x)^{2^0} + (1 + x)^{2^1} = 1 + x + 1 + x^2 = 1;$$

- norma

$$N_{4/2}(x) = x^{\frac{4-1}{2-1}} = x^3 = 1$$

oraz

$$N_{4/2}(1 + x) = (1 + x)^{\frac{4-1}{2-1}} = (1 + x)^3 = 1.$$

(**2**) Podobnie dla elementu $\theta^4 \in \mathbb{F}_{16}$ ciała zbudowanego w przykładzie 7.6.4(1) (tutaj $q = 2$ oraz $n = 4$):

- ślad

$$\mathrm{tr}_{16/2}(\theta^4) = \theta^4 + (\theta^4)^2 + (\theta^4)^4 = \theta^8 + \theta + \theta^2 = 1010 + 1111 + 0011 = 0110 = \theta^2;$$

- norma

$$N_{16/2}(\theta^4) = \left(\theta^4\right)^{\frac{16-1}{2-1}} = (\theta^4)^{15} = 1.$$

Ćwiczenia 7.6.12.

(**1**) Zbudować liniowy ciąg rekurencyjny **S** stowarzyszony z wielomianem $f \in \mathbb{F}_p[X]$ i sprawdzić, czy ten ciąg jest pseudolosowy. Jeśli jest, to przedstawić elementy ciała $\mathbb{F}_p[X]/\langle f\rangle$ w postaci wektorowej, jeśli:
(a) $p = 2$ oraz $f = X^4 + X^3 + X^2 + X + 1$;
(b) $p = 2$ oraz $f = X^3 + X^2 + 1$;
(c) $p = 2$ oraz $f = X^5 + X^3 + X$;
(d) $p = 2$ oraz $f = X^3 + X + 1$;
(e) $p = 3$ oraz $f = X^3 + 2X + 1$;
(f) $p = 2$ oraz $f = X^6 + X^5 + X^4 + 1$;
(g) $p = 3$ oraz $f = X^2 + 2X + 2$;
(h) $p = 3$ oraz $f = X^2 + X + 2$;
(i) $p = 11$ oraz $f = X^2 + 1$;
(j) $p = 11$ oraz $f = X^2 + X + 4$.
(**2**) Przekonać się, że wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_{16}$ podane w takiej tablicy

i	$Z(i)$	i	$Z(i)$
∞	0	7	9
0	∞	8	2
1	4	9	7
2	8	19	5
3	14	11	12
4	1	12	11
5	10	13	6
6	13	14	3

są poprawne.
(**3**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_2[X]/\langle X^3 + X^2 + 1\rangle$.
(**4**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_2[X]/\langle X^4 + X^3 + 1\rangle$.
(**5**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_3[X]/\langle X^2 + X + 2\rangle$.
(**6**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_2[X]/\langle X^4 + X^3 + X^2 + 1\rangle$.
(**7**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_q$, jeśli:
(a) $q = 27$;
(b) $q = 25$;
(c) $q = 32$;
(d) $q = 64$;
(e) $q = 9$.
(**8**) Obliczyć wartości logarytmów Jacobiego-Zecha dla elementów ciała $\mathbb{F}_2[X]/\langle X^4+X^3+X^2+X+1\rangle$.
(**9**) Obliczyć ślady i normy wszystkich elementów ciała:
(a) $\mathbb{F}_8$;
(b) $\mathbb{F}_9$;
(c) $\mathbb{F}_{16}$;
(d) $\mathbb{F}_{25}$.
(**10**) Niech $\theta \in \mathbb{F}_{q^n}$. Udowodnić, że element $\mu = \theta^q - \theta = \beta^q - \beta$ dla pewnego elementu $\beta \in \mathbb{F}_{q^n}$ wtedy i tylko wtedy, gdy $\theta - \beta \in \mathbb{F}_q$.
(**11**) Niech θ będzie pierwiastkiem wielomianu $X^n - 1 \in \mathbb{F}_q[X]$. Udowodnić, że

$$1 + \theta + \theta^2 + \cdots + \theta^{n-1} = \begin{cases} 0, & \text{gdy } \theta \neq 1, \\ n, & \text{gdy } \theta = 1. \end{cases}$$

Uwagi. Logarytmy Zecha[6] (nazywane też logarytmami Jacobiego-Zecha) C. Jacobi[7] wykorzystywał w badaniach teorioliczbowych.

Twierdzenie 7.6.10 jest zawarte w książce D. Hilberta *Zahlbericht* (1897 r.) pod numerem 90.

[6] Julius August Christoph Zech (1821–1864)

[7] Carl Gustav Jacob Jacobi (1804–1851)

7.7. Liczby wielomianów nieprzywiedlnych i prymitywnych

Lemat 7.7.1. *Niech r oraz t będą dodatnimi liczbami całkowitymi. Wtedy zachodzą następujące własności:*

(1) *iloczyn wszystkich unormowanych wielomianów z pierścienia $\mathbb{F}_q[X]$ nieprzywiedlnych nad $\mathbb{F}_q$, których stopnie dzielą liczbę r, jest równy $X^{q^r} - X \in \mathbb{F}_q[X]$,*

(2) *jeśli $I_q(t)$ jest liczbą unormowanych wielomianów nieprzywiedlnych stopnia $t \geqslant 1$ z pierścienia $\mathbb{F}_q[X]$, to*

$$q^r = \sum_{t|r} t \cdot I_q(t).$$

Dowód. (1) Jak wiadomo, każdy unormowany wielomian (w szczególności $X^{q^r} - X$) jest iloczynem pewnych unormowanych wielomianów z pierścienia $\mathbb{F}_q[X]$ nieprzywiedlnych nad ciałem $\mathbb{F}_q$. Ale jeśli stopień wielomianu $f \in \mathbb{F}_q[X]$ dzieli liczbę r, to f dzieli $X^{q^r} - X$ w pierścieniu $\mathbb{F}_q[X]$ na mocy wniosku 7.2.3. Oprócz tego wielomian $X^{q^r} - X$ nie posiada pierwiastków krotnych w ciele $\mathbb{F}_{q^r}$. Zatem teza zachodzi.

(2) Wynika na podstawie części (1). □

* * *

■ Najpierw przypomnijmy własności funkcji Möbiusa stosowane w następnym twierdzeniu o wielomianach nieprzywiedlnych.

Funkcja

$$\mu : \mathbb{N}^* \ni n \mapsto \mu(n) \in \{-1, 0, 1\}$$

jest nazywana *funkcją Möbiusa*, jeśli

$$\mu(n) = \begin{cases} 1, & \text{gdy } n = 1, \\ (-1)^t, & \text{gdy } n \text{ jest iloczynem } t \text{ parami} \\ & \text{różnych liczb pierwszych,} \\ 0, & \text{gdy } n \text{ jest podzielne przez kwadrat} \\ & \text{pewnej liczby pierwszej.} \end{cases}$$

Zachodzi twierdzenie Frobeniusa o odwracaniu (patrz część (2)).

Twierdzenie 7.7.2. *Niech n będzie dodatnią liczbą całkowitą. Wtedy zachodzą następujące własności:*

(1)
$$\sum_{t|n} \mu(t) = \begin{cases} 1, & \text{jeśli } n = 1, \\ 0, & \text{jeśli } n > 1; \end{cases}$$

(2) (postać addytywna wzoru Möbiusa) *jeśli $(G, +)$ jest grupą abelową oraz $f, g : \mathbb{N}^* \to G$ są funkcjami, to*
$$f(n) = \sum_{t|n} g(t) \tag{7.12}$$
zachodzi dla wszystkich $n \in \mathbb{N}^*$ wtedy i tylko wtedy, gdy
$$g(n) = \sum_{t|n} \mu\left(\frac{n}{t}\right) f(t) = \sum_{t|n} \mu(t) f\left(\frac{n}{t}\right)$$
dla wszystkich $n \in \mathbb{N}^*$;

(3)
$$\varphi(n) = \sum_{t|n} t\mu\left(\frac{n}{t}\right),$$
gdzie $n \geqslant 1$ jest liczbą całkowitą.

Dowód. Jeśli $n = 1$, to teza zachodzi. Załóżmy więc, że $n > 1$. W sumie $\sum\limits_{t|n} \mu(t)$ niezerowymi będą składniki $\mu(t)$, dla których $t = 1$ lub t jest liczbą czynników w iloczynie parami różnych pewnych liczb pierwszych z listy $p_1, \ldots, p_k$. Wtedy
$$\sum_{t|n} \mu(t) = \mu(1) + \sum_{i=1}^{k} \mu(p_i) + \sum_{1 \leqslant i_1 < i_2 \leqslant k} \mu(p_1 p_2) + \cdots + \mu(p_1 \cdots p_k) = \\ = \sum_{i=0}^{t} \binom{k}{i} (-1)^i = (1 + (-1))^k = 0.$$

(2) ($\Rightarrow$) Załóżmy, że zachodzi (7.12) dla wszystkich $n \in \mathbb{N}^*$. Wtedy na podstawie części (1) otrzymujemy
$$\sum_{t|n} \mu\left(\tfrac{n}{t}\right) f(t) = \sum_{t|n} \mu(t) f\left(\tfrac{n}{t}\right) = \sum_{t|n} \mu(t) \sum_{c|\frac{n}{t}} g(c) = \\ = \sum_{c|n} \sum_{t|\frac{n}{c}} \mu(t) g(c) = \sum_{c|n} g(c) \sum_{t|\frac{n}{c}} \mu(t) = g(n).$$

dla wszystkich $n \in \mathbb{N}^*$.

($\Leftarrow$) Podobnie mamy

$$\begin{aligned}\sum_{t|n} g(t) &= \sum_{t|n} g\left(\tfrac{n}{t}\right) = \sum_{t|n}\sum_{c|\frac{n}{t}} \mu(c) f\left(\tfrac{n}{tc}\right) = \\ &= \sum_{t|n}\sum_{c|\frac{n}{t}} \mu\left(\tfrac{n}{tc}\right) f(c) = \sum_{tc|n} \mu\left(\tfrac{n}{tc}\right) f(c) = \\ &= \sum_{c|n} f(c) \sum_{t|\frac{n}{c}} \mu\left(\tfrac{n}{tc}\right) = f(n).\end{aligned}$$

(3) Z części (3) oraz (2) wynika, że

$$\varphi(n) = \sum_{t|n} \mu(t) \cdot \frac{n}{t} = \sum_{t|n} t \cdot \mu\left(\frac{n}{t}\right).$$

□

Twierdzenie 7.7.3 (o liczbie wielomianów nieprzywiedlnych)**.** *Liczba $I_q(r)$ wielomianów unormowanych stopnia $r \geqslant 1$ z pierścienia $\mathbb{F}_q[X]$, które są nieprzywiedlne nad ciałem $\mathbb{F}_q$, jest równa*

$$I_q(r) = \frac{1}{r} \sum_{t|r} \mu\left(\frac{r}{t}\right) q^t = \frac{1}{r} \sum_{t|r} \mu(t) q^{\frac{r}{t}}.$$

Dowód. Niech $f, g : \mathbb{N}^* \to \mathbb{Z}$ będą funkcjami odwzorowującymi zbiór dodatnich liczb całkowitych $\mathbb{N}^*$ w grupę addytywną liczb całkowitych $\mathbb{Z}$ określonymi w następujący sposób

$$f(r) = q^r \text{ oraz } g(r) = rI_q(r)$$

(tutaj i niżej $r \in \mathbb{N}^*$). Wtedy z lematu 7.7.1(2) otrzymujemy, że

$$\sum_{t|r} tI_q(t) = q^r = f(r),$$

a więc na podstawie twierdzenia 7.7.2(2) teza zachodzi. □

Przykład 7.7.4.

Liczba wielomianów unormowanych (stopnia $r \geqslant 1$ w pierścieniu $\mathbb{F}_q[X]$) nieprzywiedlnych nad ciałem $\mathbb{F}_q$ wynosi:

- $I_2(1) = 2$, jeśli $r = 1$ oraz $q = 2$,
- $I_2(2) = \frac{1}{2}\mu(2)2 + \frac{1}{2}\mu(1)2^2 = 1$, jeśli $r = 2$ oraz $q = 2$,
- $I_2(3) = \frac{1}{3}\mu(3)2 + \frac{1}{3}\mu(1)2^3 = 2$, jeśli $r = 3$ oraz $q = 2$,
- $I_2(4) = \frac{1}{4}\mu(4)2 + \frac{1}{4}\mu(2)2^2 + \frac{1}{4}\mu(1)2^4 = 3$, jeśli $r = 4$ oraz $q = 2$,
- $I_3(1) = 3$, jeśli $r = 1$ oraz $q = 3$,
- $I_3(2) = \frac{1}{2}\mu(2)3 + \frac{1}{2}\mu(1)3^2 = 3$, jeśli $r = 2$ oraz $q = 3$,
- $I_3(3) = \frac{1}{3}\mu(3)3 + \frac{1}{3}\mu(1)3^3 = 8$, jeśli $r = 3$ oraz $q = 3$,
- $I_3(6) = \frac{1}{6}\mu(6)3 + \frac{1}{6}\mu(3)3^2 + \frac{1}{6}\mu(2)3^3 + \frac{1}{6}\mu(1)3^6 = 116$, jeśli $r = 6$ oraz $q = 3$.

■ Oczywiście, że $I_q(r) \geqslant 1$ dla dowolnego niezerowego $r \in \mathbb{N}^*$.

■ Niżej są przytoczone listy wielomianów stopnia n (oraz rzędu e) nieprzywiedlnych nad ciałem $\mathbb{F}_q$:

	stopień	wielomian	rząd e
$q = 2$	$n = 1$	$X + 1$	1
$q = 2$	$n = 2$	$X^2 + X + 1$	3
$q = 2$	$n = 3$	$X^3 + X + 1$	7
		$X^3 + X^2 + 1$	7
$q = 2$	$n = 4$	$X^4 + X^2 + 1$	15
		$X^4 + X^3 + 1$	15
		$X^4 + X^3 + X^2 + X + 1$	5
$q = 2$	$n = 5$	$X^5 + X^2 + 1$	31
		$X^5 + X^3 + 1$	31
		$X^5 + X^3 + X^2 + X + 1$	31
		$X^5 + X^4 + X^2 + X + 1$	31
		$X^5 + X^4 + X^3 + X + 1$	31
		$X^5 + X^4 + X^3 + X^2 + 1$	31

	stopień	wielomian	rząd e
$q = 3$	$n = 1$	$X + 1$	2
		$X + 2$	1
$q = 3$	$n = 2$	$X^2 + 1$	4
		$X^2 + X + 2$	8
		$X^2 + 2X + 2$	9
$q = 3$	$n = 3$	$X^3 + 2X + 1$	26
		$X^3 + 2X + 2$	13
		$X^3 + X^2 + 2$	13
		$X^3 + X^2 + X + 2$	13
		$X^3 + X^2 + 2X + 1$	26
		$X^3 + 2X^2 + 1$	26
		$X^3 + 2X^2 + X + 1$	26
		$X^3 + 2X^2 + 2X + 2$	13

	stopień	wielomian	rząd e
$q=3$	$n=4$	X^4+X+2	80
		X^4+2X+2	80
		X^4+X^2+2	16
		X^4+X^2+X+1	40
		X^4+X^2+2X+1	40
		X^4+2X^2+2	16
		X^4+X^3+2	80
		X^4+X^3+2X+1	20
		$X^4+X^3+X^2+1$	40
		$X^4+X^3+X^2+X+1$	5
		$X^4+X^3+X^2+2X+2$	80
		$X^4+X^3+2X^2+2X+2$	80
		X^4+2X^3+2	80
		X^4+2X^3+X+1	20
		$X^4+2X^3+X^2+1$	40
		$X^4+2X^3+X^2+X+2$	80
		$X^4+2X^3+X^2+2X+1$	10
		$X^4+2X^3+2X^2+X+2$	80

	stopień	wielomian	rząd e
$q=5$	$n=1$	$X+1$	2
		$X+2$	4
		$X+3$	4
		$X+4$	1
$q=5$	$n=2$	X^2+2	8
		X^2+3	8
		X^2+X+1	3
		X^2+X+2	24
		X^2+2X+3	24
		X^2+2X+4	12
		$2X^2+3X+3$	24
		X^2+3X+4	12
		X^2+4X+1	6
		X^2+4X+2	24

Twierdzenie 7.7.5. *Liczba unormowanych wielomianów $f \in \mathbb{F}_q[X]$ stopnia $r \geqslant 1$ oraz rzędu e, które są nieprzywiedlne nad ciałem $\mathbb{F}_q$, jest rów-*

na[8]

$$\begin{cases} \frac{\varphi(e)}{r}, & \textit{gdy } e \geqslant 2 \textit{ oraz } r \textit{ jest rzędem multiplikatywnym} \\ & \textit{liczby } q \textit{ modulo } e, \\ 2, & \textit{gdy } r = e = 1, \\ 0 & \textit{w pozostałych przypadkach.} \end{cases}$$

Dowód. Niech $f \in \mathbb{F}_q[X]$. Jeśli $r = e = 1$, to $f = X \in \mathbb{F}_q[X]$. Załóżmy, że $f(0) \neq 0$. Wtedy $\operatorname{ord} f = e$ w tym i tylko tym przypadku, gdy każdy pierwiastek θ wielomianu f jest pierwiastkiem prymitywnym stopnia e z jedności 1 nad ciałem $\mathbb{F}_q$ (czyli $e \in \mathbb{N}^*$ jest najmniejsze z własnością $\theta^e = 1$). Ponieważ $X^e - 1$ jest dzielnikiem wielomianu $X^{q^r} - X$, to ciało $\mathbb{F}_{q^r}$ zawiera wszystkie pierwiastki stopnia e z jedności 1, a więc $\mathbb{F}_{q^r}$ zawiera $\varphi(e)$ pierwiastków prymitywnych stopnia e z 1. Zatem teza zachodzi. □

Przykład 7.7.6.

Skoro $2^k \not\equiv 1 \pmod{11}$ dla $k \in \mathbb{N}$ takich, że $1 \leqslant k \leqslant 9$ oraz $2^{10} \equiv 1 \pmod{11}$, to rząd (multiplikatywny) liczby 2 modulo 11 jest równy 10.

* * *

■ **Liczba wielomianów prymitywnych.** Zachodzi następny

Lemat 7.7.7. *Jeśli θ jest elementem prymitywnym ciała $\mathbb{F}_q$, to sprzężone z nim elementy*

$$\theta, \theta^p, \theta^{p^2}, \ldots$$

również są prymitywne.

Dowód. Nie wprost. Załóżmy, że element θ^p nie jest prymitywny. Wtedy jego rząd

$$o(\theta^p) = s < q - 1$$

dla pewnego $s \in \mathbb{N}^*$. Skoro $\sigma : \mathbb{F}_q \ni a \mapsto a^p \in \mathbb{F}_q$ jest automorfizmem ciała $\mathbb{F}_q$, to

$$1 = \sigma(\theta^s) = (\theta^s)^p \quad \Rightarrow \quad \theta^s = 1,$$

[8] Przypomnijmy, że najmniejsza liczba naturalna $n \in \mathbb{N}^*$, taka, że $a^k \equiv 1 \pmod{n}$, jest nazywana *rzędem multiplikatywnym* liczby $a \in \mathbb{Z}$ modulo n, gdzie $\mathrm{NWD}(a, n) = 1$.

co prowadzi do sprzeczności. □

Twierdzenie 7.7.8 (o liczbie wielomianów prymitywnych)**.** *Liczba wielomianów prymitywnych stopnia s nad ciałem $\mathbb{F}_p$ jest równa*

$$\frac{\varphi(p^s-1)}{s}. \tag{7.13}$$

Dowód. Na podstawie twierdzenia 3.1.10 istnieje $\varphi(p^s-1)$ elementów prymitywnych w ciele $\mathbb{F}_{p^s}$. Każdy wielomian prymitywny $f \in \mathbb{F}_p[X]$ stopnia s ma s pierwiastków w ciele $\mathbb{F}_{p^s}$, które są prymitywne. Zatem liczba takich wielomianów f jest równa liczbie (7.13). □

Przykłady 7.7.9.

(**1**) Nad ciałem $\mathbb{F}_2$ istnieje:

- $$\frac{\varphi(2^4-1)}{4} = \frac{8}{4} = 2$$

wielomianów prymitywnych stopnia 4;

- oraz 3 wielomianów nieprzywiedlnych stopnia 4 (patrz przykład 7.7.4).

(**2**) Znajdźmy wszystkie wielomiany f stopnia 2, które są nierozkładalne nad ciałem $\mathbb{F}_3$.

a) Skoro $f = aX^2+bX+c \in \mathbb{F}_3[X]$ oraz $\deg f = 2$, to $a \in \{1,2\}$, a więc mamy 18 wielomianów stopnia 2, a mianowicie:

$$\begin{array}{llllllll}
(b,c)=(0,0) & \leftrightarrow & f_1 & = & X^2, & f_{10} & = & 2X^2,\\
(b,c)=(1,0) & \leftrightarrow & f_2 & = & X^2+X, & f_{11} & = & 2X^2+X,\\
(b,c)=(2,0) & \leftrightarrow & f_3 & = & X^2+2X, & f_{12} & = & 2X^2+2X,\\
(b,c)=(1,1) & \leftrightarrow & f_4 & = & X^2+X+1, & f_{13} & = & 2X^2+X+1,\\
(b,c)=(1,2) & \leftrightarrow & f_5 & = & X^2+X+2, & f_{14} & = & 2X^2+X+2,\\
(b,c)=(0,1) & \leftrightarrow & f_6 & = & X^2+1, & f_{15} & = & 2X^2+1,\\
(b,c)=(0,2) & \leftrightarrow & f_7 & = & X^2+2, & f_{16} & = & 2X^2+2,\\
(b,c)=(2,2) & \leftrightarrow & f_8 & = & X^2+2X+2, & f_{17} & = & 2X^2+2X+2,\\
(b,c)=(2,1) & \leftrightarrow & f_9 & = & X^2+2X+1, & f_{18} & = & 2X^2+2X+1.
\end{array}$$

b) Teraz sprawdźmy, które z tych wielomianów są rozkładalne:

- $_1$ $$f_1 = X\cdot X \quad\Rightarrow\quad f_1 \text{ – rozkładalny;}$$
- $_2$ $$f_2 = X\cdot(X+1) \quad\Rightarrow\quad f_2 \text{ – rozkładalny;}$$
- $_3$ $$f_3 = X\cdot(X+2) \quad\Rightarrow\quad f_3 \text{ – rozkładalny;}$$
- $_4$ $$\left.\begin{array}{lll} f_4(0) & = & 1\neq 0\\ f_4(1) & = & 0\\ f_4(2) & = & 1\neq 0 \end{array}\right\} \quad\Rightarrow\quad f_4 \text{ – rozkładalny;}$$

•5
$$\left.\begin{array}{rcl} f_5(0) & = & 2 \neq 0 \\ f_5(1) & = & 1 \neq 0 \\ f_5(2) & = & 2 \neq 0 \end{array}\right\} \Rightarrow \quad f_5 \text{ – nierozkładalny;}$$

•6
$$\left.\begin{array}{rcl} f_6(0) & = & 1 \neq 0 \\ f_6(1) & = & 2 \neq 0 \\ f_6(2) & = & 2 \neq 0 \end{array}\right\} \Rightarrow \quad f_6 \text{ – nierozkładalny;}$$

•7
$$\left.\begin{array}{rcl} f_7(0) & = & 2 \neq 0 \\ f_7(1) & = & 0 \\ f_7(2) & = & 0 \end{array}\right\} \Rightarrow \quad f_7 \text{ – rozkładalny;}$$

•8
$$\left.\begin{array}{rcl} f_8(0) & = & 2 \neq 0 \\ f_8(1) & = & 2 \neq 0 \\ f_8(2) & = & 1 \neq 0 \end{array}\right\} \Rightarrow \quad f_8 \text{ – nierozkładalny;}$$

•9
$$\left.\begin{array}{rcl} f_9(0) & = & 1 \neq 0 \\ f_9(1) & = & 1 \neq 0 \\ f_9(2) & = & 0 \end{array}\right\} \Rightarrow \quad f_9 \text{ – rozkładalny;}$$

•10
$$f_{10} = 2X \cdot X \quad \Rightarrow \quad f_{10} \text{ – rozkładalny;}$$

•11
$$f_{11} = X \cdot (2X + 1) \quad \Rightarrow \quad f_{11} \text{ – rozkładalny;}$$

•12
$$f_{12} = 2X \cdot (X + 1) \quad \Rightarrow \quad f_{12} \text{ – rozkładalny;}$$

•13
$$\left.\begin{array}{rcl} f_{13}(0) & = & 1 \neq 0 \\ f_{13}(1) & = & 1 \neq 0 \\ f_{13}(2) & = & 2 \neq 0 \end{array}\right\} \Rightarrow \quad f_{13} \text{ – nierozkładalny;}$$

•14
$$\left.\begin{array}{rcl} f_{14}(0) & = & 2 \neq 0 \\ f_{14}(1) & = & 2 \neq 0 \\ f_{14}(2) & = & 0 \end{array}\right\} \Rightarrow \quad f_{14} \text{ – rozkładalny;}$$

•15
$$\left.\begin{array}{rcl} f_{15}(0) & = & 1 \neq 0 \\ f_{15}(1) & = & 0 \\ f_{15}(2) & = & 0 \end{array}\right\} \Rightarrow \quad f_{15} \text{ – rozkładalny;}$$

•16
$$\left.\begin{array}{rcl} f_{16}(0) & = & 2 \neq 0 \\ f_{16}(1) & = & 1 \neq 0 \\ f_{16}(2) & = & 1 \neq 0 \end{array}\right\} \Rightarrow \quad f_{16} \text{ – nierozkładalny;}$$

•17
$$\left.\begin{array}{rcl} f_{17}(0) & = & 2 \neq 0 \\ f_{17}(1) & = & 0 \\ f_{17}(2) & = & 2 \neq 0 \end{array}\right\} \Rightarrow \quad f_{17} \text{ – rozkładalny;}$$

•18
$$\left.\begin{array}{rcl} f_{18}(0) & = & 1 \neq 0 \\ f_{18}(1) & = & 2 \neq 0 \\ f_{18}(2) & = & 1 \neq 0 \end{array}\right\} \Rightarrow \quad f_{18} \text{ – nierozkładalny.}$$

Zatem istnieje 6 wielomianów stopnia 2 nierozkładalnych nad ciałem $\mathbb{F}_3$:

$$\begin{array}{rcl} f_5 & = & X^2 + X + 2; \\ f_6 & = & X^2 + 1; \\ f_8 & = & X^2 + 2X + 2; \\ f_{13} & = & 2X^2 + X + 1; \\ f_{16} & = & 2X^2 + 2; \\ f_{18} & = & 2X^2 + 2X + 1. \end{array}$$

(**3**) Znajdźmy wszystkie wielomiany f stopnia 2, które są prymitywne nad ciałem $\mathbb{F}_3$. Ponieważ każdy wielomian prymitywny jest nierozkładalny, to w tym celu skorzystamy z wyników poprzedniego przykładu (**2**). Zbadajmy, który z wielomianów $f_5,\ f_6,\ f_8,\ f_{13},\ f_{16},\ f_{18}$ jest prymitywny nad $\mathbb{F}_3$.

•1 Wielomianowi $f_5 = X^2 + X + 2$ odpowiada stowarzyszona liniowa zależność rekurencyjna

$$s_{j+2} = 2s_{j+1} + s_j.$$

Losowo wybierając dwie (bo $\deg f_5 = 2$) wartości początkowe $s_0 = 1$ oraz $s_1 = 0$, obliczamy wartości liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z wielomianem f_5:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
1	0	1	2	2	0	2	1	1	0	1

Ponieważ okres tego ciągu $p(\mathbf{S}) = 8 = 3^{\deg f_5} - 1$, to ciąg $\mathbf{S}$ jest pseudolosowy, a wielomian f_5 jest prymitywny.

•2 Wielomianowi $f_6 = X^2 + 1$ odpowiada stowarzyszona liniowa zależność rekurencyjna

$$s_{j+2} = 2s_j.$$

Losowo wybierając dwie (bo $\deg f_6 = 2$) wartości początkowe $s_0 = 2$ oraz $s_1 = 1$, obliczamy wartości liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z wielomianem f_6:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7
2	1	1	2	2	1	1	2

Ponieważ okres $p(\mathbf{S}) = 4 \neq 3^{\deg f_6} - 1$, to ciąg $\mathbf{S}$ nie jest pseudolosowy, a wielomian f_6 nie jest prymitywny.

•3 Wielomianowi $f_8 = X^2 + 2X + 2$ odpowiada stowarzyszona liniowa zależność rekurencyjna

$$s_{j+2} = s_{j+1} + s_j.$$

Losowo wybierając dwie (bo $\deg f_8 = 2$) wartości początkowe $s_0 = 2$ oraz $s_1 = 0$, obliczamy wartości liniowego ciągu rekurencyjnego $\mathbf{S}$ stowarzyszonego z wielomianem f_8:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}
2	0	2	2	1	0	1	1	2	0	2

Ponieważ okres $p(\mathbf{S}) = 8 = 3^{\deg f_8} - 1$, to ciąg $\mathbf{S}$ jest pseudolosowy, a wielomian f_8 jest prymitywny.

•4 Wielomiany $f_{13},\ f_{16},\ f_{18}$ nie są unormowane (a więc nie są prymitywne), bo $f_{13} = 2f_8,\ f_{16} = 2f_6,\ f_{18} = 2f_5$.

Wnosimy, że nad ciałem $\mathbb{F}_3$ dokładnie dwa wielomiany są prymitywne, a mianowicie:

$$\begin{array}{ccccc} X^2 & + & X & + & 2, \\ X^2 & + & 2X & + & 2. \end{array}$$

Zanotujmy, że na podstawie twierdzenia 7.7.8 liczba wielomianów stopnia 2 prymitywnych nad ciałem $\mathbb{F}_3$ jest równa

$$\frac{\varphi(3^2-1)}{2} = 2.$$

Ćwiczenia 7.7.10.

(**1**) Podobnie jak w przykładzie 7.3.2 zbudować ciała:
(a) $\mathbb{F}_4$;
(b) $\mathbb{F}_9$;
(c) $\mathbb{F}_8$;
(d) $\mathbb{F}_{16}$;
(e) $\mathbb{F}_{25}$.
(**2**) Znaleźć wszystkie wielomiany $f \in \mathbb{F}_2[X]$ stopnia n nieprzywiedlne nad ciałem $\mathbb{F}_2$, jeśli:
(a) $n = 2$;
(b) $n = 3$;
(c) $n = 4$.
(**3**) Znaleźć wszystkie wielomiany $f \in \mathbb{F}_3[X]$ stopnia n nieprzywiedlne nad ciałem $\mathbb{F}_3$, jeśli:
(a) $n = 3$;
(b) $n = 4$.
(**4**) Znaleźć wszystkie wielomiany $f \in \mathbb{F}_5[X]$ stopnia n nieprzywiedlne nad ciałem $\mathbb{F}_5$, jeśli:
(a) $n = 2$;
(b) $n = 3$.
(**5**) Znaleźć wielomian $f \in \mathbb{F}_m[X]$ stopnia n nieprzywiedlny nad ciałem $\mathbb{F}_m$, jeśli:
(a) $n = 3$ oraz $m = 4$;
(b) $n = 4$ oraz $m = 4$;
(c) $n = 3$ oraz $m = 11$.
(**6**) Udowodnić, że jeśli p jest liczbą pierwszą oraz $0 \neq a \in \mathbb{F}_p$, to wielomian $f = X^p - X - a \in \mathbb{F}_p[X]$ jest nierozkładalny nad ciałem $\mathbb{F}_p$.
(**7**) Udowodnić, że pierwiastki wielomianu $a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n \in \mathbb{F}[X]$ są odwrotne do pierwiastków wielomianu $a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in \mathbb{F}[X]$ nad ciałem $\mathbb{F}$.
(**8**) Znaleźć $I_2(6)$.

Uwagi. L. Carlitz[(9)] w 1964 r. udowodnił, że nad ciałem $\mathbb{F}_q$ istnieje jeden wielomian prymitywny stopnia n dokładnie wtedy, gdy $q = 2$, $n \leqslant 2$ lub $q = 3$, $n = 1$.

[(9)] Leonard Carlitz (1907–1999)

7.8. Faktoryzacja wielomianu $X^n - 1$

Niech $r \geqslant 1$ będzie liczbą całkowitą. Jak zwykle $q = p^u$ dla pewnej liczby pierwszej p i pewnej liczby dodatniej całkowitej u. W wielu zastosowaniach ważny jest opis rozłożenia wielomianu $X^n - 1$ na czynniki nieprzywiedlne nad ciałem $\mathbb{F}_q$.

Jeśli $n = q^i t$ dla pewnych dodatnich liczb całkowitych $i \geqslant 1$ oraz t, gdzie $\mathrm{NWD}(q, t) = 1$, to

$$(X^t - 1)^{q^i} = X^n - 1.$$

W związku z tym załóżmy, że

$$\mathrm{NWD}(q, n) = 1,$$

bo wtedy pochodna

$$(X^n - 1)' = nX^{n-1}$$

nie jest zerowa, a więc wszystkie pierwiastki wielomianu $X^n - 1$ są parami różne i ten wielomian nie posiada dzielników krotnych.

■ Jeśli $\theta \in \mathbb{F}_{q^r}$ jest pierwiastkiem wielomianu $f \in \mathbb{F}_q[X]$, to

$$f(\theta^{q^i}) = f(\theta)^{q^i} = 0$$

dla każdego $i \in \mathbb{N}$, czyli θ^{q^i} także jest pierwiastkiem wielomianu f. Zatem mamy taki

Lemat 7.8.1. *Jeśli θ jest elementem prymitywnym ciała $\mathbb{F}_{q^r}$, $m_i \in \mathbb{F}_q[X]$ jest wielomianem minimalnym elementu θ^i oraz stopień $\deg m_i = k$, to wielomian m_i ma w ciele $\mathbb{F}_{q^r}$ parami różnych k pierwiastków postaci*

$$\theta^i,\ \theta^{iq},\ \ldots, \theta^{iq^{k-1}} \pmod{q-1}, \tag{7.14}$$

gdzie $\theta^{iq^k} = \theta^i$.

■ Jeśli $\mathrm{NWD}(q, n) = 1$ oraz i jest taką liczbą całkowitą, że $0 \leqslant i \leqslant n-1$, to zbiór

$$C_i = \{i, iq, iq^2, \ldots, iq^{k-1} \pmod{n}\}$$

jest nazywany *warstwą cyklotomiczną* liczby całkowitej q modulo liczby n z *reprezentantem* i, gdzie elementy tego zbioru bierzemy modulo n, a k jest najmniejszą liczbą całkowitą dodatnią taką, że

$$iq^k \equiv i \pmod{n}.$$

Rząd liczby q modulo n jest równy *pojemności* (=liczbie elementów) warstwy cyklotomicznej C_1 z reprezentantem 1. Wtedy istnieje zbiór $I = \{i_1, i_2, \ldots, i_l\}$ liczb całkowitych i_j $(0 \leqslant i_j \leqslant n-1)$ taki, że warstwy cyklotomiczne $C_{i_1}, C_{i_2}, \ldots C_{i_l}$ są parami różne oraz

$$\bigcup_{j=1}^{l} C_{i_j} = \{0, 1, \ldots, n-1\}.$$

Taki zbiór I jest nazywany *pełnym zbiorem reprezentantów* warstw cyklotomicznych liczby q modulo n. Liczby, które należą do jednej warstwy cyklotomicznej, są nazywane *sprzężonymi* (modulo n).

Przykłady 7.8.2.

(**1**) Znajdźmy pełny zbiór reprezentantów warstw cyklotomicznych liczby 2 modulo 51. Mamy zbiór $A = \{0, 1, \ldots, 50\}$ wszystkich reszt modulo 51.
Najpierw obliczmy warstwy cyklotomiczne:

- $0 \in A$, a zatem $i = 0$ oraz

$$0 \cdot 2^s \equiv 0 \pmod{51}\ (s \in \mathbb{N}) \quad \Rightarrow \quad C_0 = \{0\};$$

- liczba $1 \in A \setminus C_0$, a więc $i = 1$ oraz

$$1 \cdot 2^0 \equiv 1,\ 1 \cdot 2^1 \equiv 2,\ 1 \cdot 2^2 \equiv 4,\ 1 \cdot 2^3 \equiv 8,\ 1 \cdot 2^4 \equiv 16,\ 1 \cdot 2^5 \equiv 32,\ 1 \cdot 2^6 \equiv 13,$$
$$1 \cdot 2^7 \equiv 26,\ 1 \cdot 2^8 \equiv 1 \pmod{51} \quad \Rightarrow \quad C_1 = \{1, 2, 4, 8, 16, 32, 13, 26\};$$

- liczba $3 \in A \setminus (C_0 \cup C_1)$, a więc $i = 3$ oraz

$$3 \cdot 2^0 \equiv 3,\ 3 \cdot 2^1 \equiv 6,\ 3 \cdot 2^2 \equiv 12,\ 3 \cdot 2^3 \equiv 24,\ 3 \cdot 2^4 \equiv 48,\ 3 \cdot 2^5 \equiv 45,\ 3 \cdot 2^6 \equiv 39,$$
$$3 \cdot 2^7 \equiv 27,\ 3 \cdot 2^8 \equiv 3 \pmod{51} \quad \Rightarrow \quad C_3 = \{3, 6, 12, 24, 48, 45, 39, 27\};$$

- liczba $5 \in A \setminus (C_0 \cup C_1 \cup C_3)$, a więc $i = 5$ oraz

$$5 \cdot 2^0 \equiv 5,\ 5 \cdot 2^1 \equiv 10,\ 5 \cdot 2^2 \equiv 20,\ 5 \cdot 2^3 \equiv 40,\ 5 \cdot 2^4 \equiv 29,\ 5 \cdot 2^5 \equiv 7,\ 5 \cdot 2^6 \equiv 14,$$
$$5 \cdot 2^7 \equiv 28,\ 5 \cdot 2^8 \equiv 5 \pmod{51} \quad \Rightarrow \quad C_5 = \{5, 10, 20, 40, 29, 7, 14, 28\};$$

- liczba $9 \in A \setminus (C_0 \cup C_1 \cup C_3 \cup C_5)$, a więc $i = 9$ oraz

$$9 \cdot 2^0 \equiv 9,\ 9 \cdot 2^1 \equiv 18,\ 9 \cdot 2^2 \equiv 36,\ 9 \cdot 2^3 \equiv 21,\ 9 \cdot 2^4 \equiv 42,\ 9 \cdot 2^5 \equiv 33,\ 9 \cdot 2^6 \equiv 15,$$
$$9 \cdot 2^7 \equiv 30,\ 5 \cdot 2^8 \equiv 9 \pmod{51} \quad \Rightarrow \quad C_9 = \{9, 18, 36, 21, 42, 33, 15, 30\};$$

- liczba $11 \in A \setminus (C_0 \cup C_1 \cup C_3 \cup C_5 \cup C_9)$, a więc $i = 11$ oraz

$$11 \cdot 2^0 \equiv 11,\ 11 \cdot 2^1 \equiv 22,\ 11 \cdot 2^2 \equiv 44,\ 11 \cdot 2^3 \equiv 37,\ 11 \cdot 2^4 \equiv 23,\ 11 \cdot 2^5 \equiv 46,\ 11 \cdot 2^6 \equiv 41,$$
$$11 \cdot 2^7 \equiv 31,\ 11 \cdot 2^8 \equiv 11 \pmod{51} \quad \Rightarrow \quad C_{11} = \{11, 22, 44, 37, 23, 46, 41, 31\};$$

- liczba $17 \in A \setminus (C_0 \cup C_1 \cup C_3 \cup C_5 \cup C_9 \cup C_{11})$, a więc $i = 17$ oraz

$$17 \cdot 2^0 \equiv 17,\ 17 \cdot 2^1 \equiv 34,\ 17 \cdot 2^2 \equiv 17 \pmod{51} \Rightarrow \quad C_{17} = \{17, 34\};$$

- liczba $19 \in A \setminus (C_0 \cup C_1 \cup C_3 \cup C_5 \cup C_9 \cup C_{11} \cup C_{17})$, a więc $i = 19$ oraz

$$19 \cdot 2^0 \equiv 19,\ 19 \cdot 2^1 \equiv 38,\ 19 \cdot 2^2 \equiv 25,\ 19 \cdot 2^3 \equiv 50,\ 19 \cdot 2^4 \equiv 49,\ 19 \cdot 2^5 \equiv 47,$$
$$19 \cdot 2^6 \equiv 43,\ 19 \cdot 2^7 \equiv 35,\ 19 \cdot 2^8 \equiv 19 \pmod{51} \quad \Rightarrow \quad C_{19} = \{19, 38, 25, 50, 49, 47, 43, 35\}.$$

Zatem

$$A = C_0 \bigcup C_1 \bigcup C_3 \bigcup C_5 \bigcup C_9 \bigcup C_{11} \bigcup C_{17} \bigcup C_{19},$$

rodzina zbiorów

$$\{C_0,\ C_1,\ C_3,\ C_5,\ C_9,\ C_{11},\ C_{17},\ C_{19}\}$$

tworzy rozbicie zbioru A oraz liczba $|C_1| = 8$ jest rzędem liczby 2 modulo 51.

Pełnym zbiorem reprezentantów warstw liczby 2 modulo 51 jest, na przykład,

$$\{0, 1, 3, 5, 9, 11, 17, 19\}.$$

(**2**) Obliczmy warstwy cyklotomiczne liczby 2 modulo 21:

- $[i = 0]$

$$0 \cdot 2^s \equiv 0 \pmod{21} \quad (s \in \mathbb{N}) \quad \Rightarrow \quad C_0 = \{0\};$$

- $[i = 1]$

$$1 \cdot 2^0 \equiv 1,\ 1 \cdot 2^1 \equiv 2,\ 1 \cdot 2^2 \equiv 4,\ 1 \cdot 2^3 \equiv 8,\ 1 \cdot 2^4 \equiv 16,\ 1 \cdot 2^5 \equiv 11,$$
$$1 \cdot 2^6 \equiv 1 \pmod{21} \quad \Rightarrow \quad C_1 = \{1, 2, 4, 8, 16, 11\};$$

- $[i = 3]$

$$3 \cdot 2^0 \equiv 3,\ 3 \cdot 2^1 \equiv 6,\ 3 \cdot 2^2 \equiv 12,\ 3 \cdot 2^3 \equiv 3 \pmod{21} \quad \Rightarrow \quad C_3 = \{3, 6, 12\};$$

- $[i = 5]$

$$5 \cdot 2^0 \equiv 5,\ 5 \cdot 2^1 \equiv 10,\ 5 \cdot 2^2 \equiv 20,\ 5 \cdot 2^3 \equiv 19,\ 5 \cdot 2^4 \equiv 17,\ 5 \cdot 2^5 \equiv 13,$$
$$5 \cdot 2^6 \equiv 5 \pmod{21} \quad \Rightarrow \quad C_5 = \{5, 10, 20, 19, 17, 13\};$$

- $[i = 7]$

$$7 \cdot 2^0 \equiv 7,\ 7 \cdot 2^1 \equiv 14,\ 7 \cdot 2^2 \equiv 7 \pmod{21} \quad \Rightarrow \quad C_7 = \{7, 14\};$$

- $[i = 9]$

$$9 \cdot 2^0 \equiv 9,\ 9 \cdot 2^1 \equiv 18,\ 9 \cdot 2^2 \equiv 15,\ 9 \cdot 2^3 \equiv 9 \pmod{21} \quad \Rightarrow \quad C_9 = \{9, 18, 15\}.$$

Zatem

$$\{0, 1, \ldots, 20\} = C_0 \bigcup C_1 \bigcup C_3 \bigcup C_5 \bigcup C_7 \bigcup C_9.$$

■ Mówimy też, że elementy ciągu (7.14) tworzą *warstwę cyklotomiczną* elementów ciała $\mathbb{F}_q$. Łatwo zauważyć, że elementy ciągu (7.14) mają ten sam rząd w grupie multiplikatywnej $\mathbb{F}_q^*$ (udowodnić samodzielnie). Zatem elementy każdego ciała skończonego możemy rozłożyć na parami nieprzecinające się warstwy cyklotomiczne.

■ Jeśli θ jest elementem prymitywnym ciała $\mathbb{F}_{q^r}$, to `wielomian minimalny` $m_i \in \mathbb{F}_q[X]$ elementu $\theta^i \in \mathbb{F}_{q^r}$ `możemy znaleźć w taki sposób`:

- najpierw obliczamy jego stopień $k = \deg m_i$ ze wzoru

$$iq^k \equiv i \pmod{q^r - 1},$$

- dalej obliczamy współczynniki wielomianu m_i, korzystając z reprezentacji wektorowej elementów ciała $\mathbb{F}_{q^r}$.

Twierdzenie 7.8.3. *Jeśli θ jest elementem prymitywnym ciała $\mathbb{F}_{q^r}$, to*

$$m_i = \prod_{j \in C_i} (X - \theta^j) \in \mathbb{F}_q[X]$$

jest wielomianem minimalnym elementu θ^i, gdzie C_i jest warstwą cyklotomiczną (liczby q modulo $q^r - 1$) zawierającą i $(0 \leqslant i \leqslant q^r - 2)$ (a więc $\deg m_i = |C_i|$).

Dowód. Jeśli

$$m_i = a_0 + a_1 X + \cdots + a_s X^s \in \mathbb{F}_{q^r}[X],$$

gdzie $|C_i| = s$, to

$$\begin{aligned} {a_0}^q + {a_1}^q X + \cdots + {a_s}^q X^s &= \prod_{j \in C_i} (X - \theta^{qj}) = \\ &= \prod_{j \in C_{qi}} (X - \theta^j) = \prod_{j \in C_i} (X - \theta^j) = m_i, \end{aligned}$$

bo $C_i = C_{qi}$. Wtedy $a_l = a_l{}^q$ $(0 \leqslant l \leqslant s)$ i w konsekwencji $a_l \in \mathbb{F}_q$. Zatem $m_i \in \mathbb{F}_q[X]$.

Ponieważ $\theta^k \neq \theta^j$ dla różnych $k, j \in C_i$, to wielomian m_i nie posiada krotnych pierwiastków. Załóżmy, że $f = b_0 + b_1 X + \cdots + b_v X^v \in \mathbb{F}_q[X]$ oraz $f(\theta^i) = 0$. Dla każdego $j \in C_i$ zachodzi warunek

$$j \equiv iq^l \pmod{q^r - 1}$$

dla pewnej liczby całkowitej nieujemnej l oraz

$$f(\theta^j) = f(\theta^{iq^l}) = f(\theta^i)^{q^l} = 0,$$

a to znaczy, że wielomian m_i dzieli f. Skoro θ^i jest pierwiastkiem wielomianu m_i, to z przytoczonych rozumowań wynika, że m_i jest jego wielomianem minimalnym. □

* * *

■ **Przypadek $\mathbf{n = q^r - 1}$.** Załóżmy, że $\theta \in \mathbb{F}_{q^r}$ jest elementem prymitywnym, czyli $\theta^{q^r-1} = 1$. Jak wiemy, wtedy

$$\mathbb{F}_{q^r} = \{0, 1, \theta, \ldots, \theta^{q^r-2}\}.$$

Niech i będzie taką liczbą całkowitą, że $0 \leqslant i \leqslant q^r - 2$. Elementy postaci (7.14) są parami różnymi pierwiastkami wielomianu minimalnego $m_i \in \mathbb{F}_q[X]$ elementu θ^i. Zatem wnioskujemy, że zachodzi takie

Twierdzenie 7.8.4. *Niech r będzie dodatnią liczbą całkowitą. Wtedy zachodzą następujące własności:*

(1) *wielomian*

$$X^{q^r-1} - 1 = g_1 g_2 \cdots g_l$$

jest iloczynem wielomianów $g_i \in \mathbb{F}_q[X]$ nieprzywiedlnych nad ciałem $\mathbb{F}_q$, przy czym stopnie $\deg g_i$ dzielą r $(i = 1, \ldots, l)$;

(2) *istnieje taki element $\alpha_i \in \mathbb{F}_{q^r}$, że g_i jest jego wielomianem minimalnym $(i = 1, 2, \ldots, l)$;*

(3) *jeśli stopień $\deg m_\theta = k$, to*

$$\theta, \theta^q, \ldots, \theta^{q^{k-1}} \in \mathbb{F}_{q^r}$$

są parami różnymi pierwiastkami wielomianu minimalnego $m_\theta \in \mathbb{F}_q[X]$ elementu $\theta \in \mathbb{F}_{q^r}$, a zatem:

(a)

$$m_\theta = m_{\theta^q} = \cdots = m_{\theta^{q^{k-1}}};$$

(b)

$$m_\theta = \prod_{i=0}^{k-1}(X - \theta^{p^i}).$$

Przykład 7.8.5.

(**1**) Załóżmy, że chcemy rozłożyć wielomian dwójkowy $X^7 - 1 \in \mathbb{F}_2[X]$. Mamy NWD$(7, 2) = 1$ oraz $7 = 2^3 - 1$. Niech θ będzie elementem prymitywnym ciała $\mathbb{F}_8$.
Obliczamy warstwy cyklotomiczne liczby 2 modulo 7:

- $$0 \cdot 2^s \equiv 0 \pmod 7 \quad (s \in \mathbb{N}) \quad \Rightarrow \quad C_0 = \{0\};$$
- $$1 \cdot 2^0 \equiv 1,\ 1 \cdot 2^1 \equiv 2,\ 1 \cdot 2^2 \equiv 4,\ 1 \cdot 2^3 \equiv 1 \pmod 7 \quad \Rightarrow \quad C_1 = \{1, 2, 4\};$$
- $$3 \cdot 2^0 \equiv 3,\ 3 \cdot 2^1 \equiv 6,\ 3 \cdot 2^2 \equiv 5,\ 3 \cdot 2^3 \equiv 3 \pmod 7 \quad \Rightarrow \quad C_3 = \{3, 6, 5\}.$$

Zatem

$$\{0, 1, 2, 3, 4, 5, 6\} = C_0 \bigcup C_1 \bigcup C_3$$

oraz

$$X^7 - 1 = m_0 m_1 m_3$$

jest iloczynem wielomianów minimalnych elementów $\theta^0,\ \theta^1$ oraz θ^3, gdzie

$$\begin{array}{lcl} m_0 & = & X - 1 \in \mathbb{F}_2[X], \\ m_1 = m_2 = m_4 = & = & (X - \theta^1)(X - \theta^2)(X - \theta^4) = X^3 + X + 1 \in \mathbb{F}_2[X], \\ m_3 = m_5 = m_6 = & = & (X - \theta^3)(X - \theta^5)(X - \theta^6) = X^3 + X^2 + 1 \in \mathbb{F}_2[X]. \end{array}$$

(**2**) Obliczmy wielomiany minimalne elementów ciała $\mathbb{F}_{16}$ nad ciałem $\mathbb{F}_2$. Skoro wielomian $f = X^4 + X + 1$ jest prymitywny nad ciałem $\mathbb{F}_2$, to istnieje pierwiastek $\theta \in \mathbb{F}_{16}$ taki, że $\mathbb{F}_{16}^* = \langle\theta\rangle$. Jeśli $\mathbf{S} = \{s_i\}$ jest liniowym ciągiem rekurencyjnym stowarzyszonym z f, to

$$s_{j+4} = s_{j+1} + s_j.$$

Biorąc wartości początkowe

$$s_0 = 1,\ s_1 = 0,\ s_2 = 0,\ s_3 = 1,$$

obliczamy kolejne wartości ciągu $\mathbf{S}$:

s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
1	0	0	1	1	0	1	0	1	1

s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	s_{16}	s_{17}	s_{18}	s_{19}
1	1	0	0	0	1	0	0	1	1

Okres ciągu $p(\mathbf{S}) = 15$, a więc ten ciąg jest pseudolosowy, a wielomian f jest prymitywny nad ciałem $\mathbb{F}_2$. Wtedy elementy ciała możemy w taki sposób przedstawić w postaci wektorowej:

$$\mathbb{F}_{16} = \left\{ \begin{array}{lcl} 0 & = & 0000, \\ 1 & = & 1001, \\ \theta^1 & = & 0011, \\ \theta^2 & = & 0110, \\ \theta^3 & = & 1101, \\ \theta^4 & = & 1010, \\ \theta^5 & = & 0101, \\ \theta^6 & = & 1011, \\ \theta^7 & = & 0111, \\ \theta^8 & = & 1111, \\ \theta^9 & = & 1110, \\ \theta^{10} & = & 1100, \\ \theta^{11} & = & 1000, \\ \theta^{12} & = & 0001, \\ \theta^{13} & = & 0010, \\ \theta^{14} & = & 0100 \end{array} \right\}.$$

Teraz obliczmy wielomiany minimalne elementów ciała $\mathbb{F}_{16}$.

- Wielomian minimalny $m_0 = (X - \theta^0) = X + 1 \in \mathbb{F}_2[X]$ elementu $\theta^0 = 1$.
- Ponieważ zachodzi implikacja

$$i = 1,\ k \geqslant 1,\ p = 2,\ 2^k \equiv 1 \pmod{15} \quad \Rightarrow \quad k = 4,$$

to wielomianu minimalnego m_1 elementu θ^1 poszukujemy w postaci $m_1 = X^4 + aX^3 + bX^2 + cX + d \in \mathbb{F}_2[X]$. Z zależności

$$0000 = 0 = m_1(\theta) = \theta^4 + a\theta^3 + b\theta^2 + c\theta^1 + d\theta^0 = 1010 + a \cdot 1101 + b \cdot 0110 + c \cdot 0011 + d \cdot 1001$$

otrzymujemy taki jednorodny układ równań liniowych

$$\begin{cases} 1 + a + + d = 0, \\ \ a + b = 0, \\ 1 + \ b + c = 0, \\ \ a + c + d = 0, \end{cases}$$

skąd obliczamy, że

$$\begin{cases} a = 0, \\ b = 0, \\ c = 1, \\ d = 1 \end{cases}$$

oraz

$$m_1 = X^4 + X + 1 \in \mathbb{F}_2[X].$$

Skoro mamy ciąg $\theta^1, \theta^2, (\theta^2)^2 = \theta^4, (\theta^4)^2 = \theta^8, (\theta^8)^2 = \theta^{16} = \theta^1$, to otrzymujemy warstwę cyklotomiczną

$$\{\theta^1,\ \theta^2,\ \theta^4,\ \theta^8\}$$

elementów ciała $\mathbb{F}_{16}$ oraz wielomiany minimalne elementów tej warstwy

$$m_1 = m_2 = m_4 = m_8 = (X-\theta^1)(X-\theta^2)(X-\theta^4)(X-\theta^8) \in \mathbb{F}_2[X]$$

są równe.

• Skoro

$$i = 3,\ k \geqslant 1,\ p = 2,\ 3 \cdot 2^k \equiv 3 \pmod{15} \quad \Rightarrow \quad k = 4,$$

to wielomianu minimalnego m_3 elementu θ^3 poszukujemy w postaci $m_3 = X^4 + aX^3 + bX^2 + cX + d \in \mathbb{F}_2[X]$. Z równości

$$0000 = 0 = m_3(\theta^3) = \theta^{12} + a\theta^9 + b\theta^6 + c\theta^3 + d\theta^0 = 0001 + a \cdot 1110 + b \cdot 1011 + c \cdot 1101 + d \cdot 1001$$

otrzymujemy jednorodny układ równań liniowych

$$\begin{cases} a+b+c+d &= 0, \\ a+\ \ +c &= 0, \\ a+b &= 0, \\ 1\quad +b+c+d &= 0, \end{cases}$$

i obliczamy, że

$$\begin{cases} a &= 1, \\ b &= 1, \\ c &= 1, \\ d &= 1 \end{cases}$$

oraz

$$m_3 = X^4 + X^3 + X^2 + X + 1 \in \mathbb{F}_2[X].$$

Skoro mamy ciąg $\theta^3, (\theta^3)^2 = \theta^6, (\theta^6)^2 = \theta^{12}, (\theta^{12})^2 = \theta^{24} = \theta^9, (\theta^9)^2 = \theta^{18} = \theta^3$, to otrzymujemy warstwę cyklotomiczną

$$\{\theta^3,\ \theta^6,\ \theta^{12},\ \theta^9\}$$

elementów ciała $\mathbb{F}_{16}$ oraz wielomiany minimalne elementów tej warstwy

$$m_3 = m_6 = m_{12} = m_9 = (X-\theta^3)(X-\theta^6)(X-\theta^9)(X-\theta^{12}) \in \mathbb{F}_2[X]$$

są równe.

• Ponieważ zachodzi implikacja

$$i = 5,\ k \geqslant 1,\ p = 2,\ 5 \cdot 2^k \equiv 5 \pmod{15} \quad \Rightarrow \quad k = 2,$$

to wielomianu minimalnego m_5 elementu θ^5 poszukujemy w postaci $m_5 = X^2 + aX + b \in \mathbb{F}_2[X]$. Skoro

$$0000 = 0 = m_5(\theta^5) = \theta^{10} + a\theta^5 + b\theta^0 = 1100 + a \cdot 0101 + b \cdot 1001,$$

dostajemy jednorodny układ równań liniowych

$$\begin{cases} 1\quad +b &= 0, \\ 1+a &= 0 \\ \quad a+b &= 0 \end{cases}$$

i obliczamy, że

$$\begin{cases} a &= 1, \\ b &= 1 \end{cases}$$

oraz

$$m_5 = X^2 + X + 1 \in \mathbb{F}_2[X].$$

Skoro mamy ciąg $\theta^5, (\theta^5)^2 = \theta^{10}, (\theta^{10})^2 = \theta^{20} = \theta^5$, to otrzymujemy warstwę cyklotomiczną

$$\{\theta^5,\ \theta^{10}\}$$

elementów ciała $\mathbb{F}_{16}$ oraz wielomiany minimalne elementów tej warstwy

$$m_5 = m_{10} = (X - \theta^5)(X - \theta^{10}) \in \mathbb{F}_2[X]$$

są równe.

- Z tego, że

$$i = 7,\ k \geqslant 1,\ p = 2,\ 7 \cdot 2^k \equiv 7 \pmod{15} \quad \Rightarrow \quad k = 4$$

wnosimy, że wielomianu minimalnego m_7 elementu θ^7 poszukujemy w postaci $m_7 = X^4 + aX^3 + bX^2 + cX + d \in \mathbb{F}_2[X]$. Ponieważ

$$\begin{gathered} 0000 = 0 = m_7(\theta^7) = \theta^{28} + a\theta^{21} + b\theta^{14} + c\theta^7 + d\theta^0 = \\ = \theta^{13} + a\theta^6 + b\theta^{14} + c\theta^7 + d\theta^0 = 0010 + a \cdot 1011 + b \cdot 0100 + c \cdot 0111 + d \cdot 1001, \end{gathered}$$

otrzymujemy taki jednorodny układ równań liniowych

$$\begin{cases} \quad a \quad\quad + d &= 0, \\ \quad\quad b + c \quad\quad &= 0, \\ 1 \quad + a \quad + c \quad\quad &= 0, \\ \quad\quad a \quad + c + d &= 0, \end{cases}$$

i obliczamy, że

$$\begin{cases} a &= 1, \\ b &= 0, \\ c &= 0, \\ d &= 1 \end{cases}$$

oraz

$$m_7 = X^4 + X^3 + 1 \in \mathbb{F}_2[X].$$

Skoro mamy ciąg $\theta^7, (\theta^7)^2 = \theta^{14}, (\theta^{14})^2 = \theta^{28} = \theta^{13}, (\theta^{13})^2 = \theta^{26} = \theta^{11}, (\theta^{11})^2 = \theta^{22} = \theta^7$, to otrzymujemy warstwę cyklotomiczną

$$\{\theta^7,\ \theta^{14}, \theta^{13}, \theta^{11}\}$$

elementów ciała $\mathbb{F}_{16}$ oraz wielomiany minimalne elementów tej warstwy

$$m_7 = m_{14} = m_{13} = m_{11} = (X - \theta^7)(X - \theta^{11})(X - \theta^{13})(X - \theta^{14}) \in \mathbb{F}_2[X]$$

są równe.

Następnie możemy bezpośrednio obliczyć warstwy cyklotomiczne liczby 2 modulo 15:

$$\begin{array}{rcr} C_0 & = & \{0\}, \\ C_1 & = & \{1,2,4,8\}, \\ C_3 & = & \{3,6,12,9\}, \\ C_5 & = & \{5,10\}, \\ C_7 & = & \{7,14,13,11\} \end{array}$$

oraz

$$\{0,1,\ldots,14\} = C_0 \bigcup C_1 \bigcup C_3 \bigcup C_5 \bigcup C_7.$$

Zatem mamy rozkład

$$X^{15} - 1 = m_0 m_1 m_3 m_5 m_7$$

nad ciałem dwójkowym $\mathbb{F}_2$. Zostawiamy Czytelnikowi do obliczenia współczynniki wszystkich wielomianów $m_i \in \mathbb{F}_2[X]$.

(**3**) Teraz rozłóżmy wielomian $X^{15} - 1$ na czynniki nieprzywiedlne nad ciałem $\mathbb{F}_4$, czyli znajdźmy wielomiany minimalne elementów ciała $\mathbb{F}_{16}$ nad ciałem $\mathbb{F}_4$. Niech $\theta \in \mathbb{F}_{16}$ będzie elementem prymitywnym. Skoro θ^5 oraz θ^{10} są pierwiastkami wielomianu $X^2 + X + 1 \in \mathbb{F}_2[X]$ oraz $\mathbb{F}_4 = \mathbb{F}_2[X]/\langle X^2 + X + 1\rangle$, to

$$\mathbb{F}_4 = \{0,\ 1,\ \theta^5,\ \theta^{10}\}.$$

Najpierw obliczamy warstwy cyklotomiczne liczby 4 modulo 15:

- $$0 \cdot 4^s \equiv 0 \pmod{15} \quad (s \in \mathbb{N}^*) \quad \Rightarrow \quad C_0 = \{0\};$$
- $$1 \cdot 4^0 \equiv 0,\ 1 \cdot 4^1 \equiv 4,\ 1 \cdot 4^2 \equiv 1 \pmod{15} \quad \Rightarrow \quad C_1 = \{1,4\};$$
- $$2 \cdot 4^0 \equiv 2,\ 2 \cdot 4^1 \equiv 8,\ 2 \cdot 4^2 \equiv 2 \pmod{15} \quad \Rightarrow \quad C_2 = \{2,8\};$$
- $$3 \cdot 4^0 \equiv 3,\ 3 \cdot 4^1 \equiv 12,\ 3 \cdot 4^2 \equiv 3 \pmod{15} \quad \Rightarrow \quad C_3 = \{3,12\};$$
- $$5 \cdot 4^0 \equiv 5,\ 5 \cdot 4^1 \equiv 5 \pmod{15} \quad \Rightarrow \quad C_5 = \{5\};$$
- $$6 \cdot 4^0 \equiv 6,\ 6 \cdot 4^1 \equiv 9,\ 6 \cdot 4^2 \equiv 6 \pmod{15} \quad \Rightarrow \quad C_6 = \{6,9\};$$
- $$7 \cdot 4^0 \equiv 7,\ 7 \cdot 4^1 \equiv 13,\ 7 \cdot 4^2 \equiv 7 \pmod{15} \quad \Rightarrow \quad C_7 = \{7,13\};$$
- $$10 \cdot 4^0 \equiv 10,\ 10 \cdot 4^1 \equiv 10 \pmod{15} \quad \Rightarrow \quad C_{10} = \{10\};$$
- $$11 \cdot 4^0 \equiv 11,\ 11 \cdot 4^1 \equiv 14,\ 11 \cdot 4^2 \equiv 11 \pmod{15} \quad \Rightarrow \quad C_{11} = \{11,14\}.$$

Zatem

$$\{0,1,\ldots,14\} = C_0 \bigcup C_1 \bigcup C_2 \bigcup C_3 \bigcup C_5 \bigcup C_6 \bigcup C_7 \bigcup C_{10} \bigcup C_{11}$$

oraz

$$X^{15} - 1 = m_0 m_1 m_2 m_3 m_5 m_6 m_7 m_{10} m_{11},$$

gdzie

$$\begin{array}{lcl}
m_0 & = & X-\theta^0 = X-1 \in \mathbb{F}_4[X],\\
m_1 = m_4 & = & (X-\theta^1)(X-\theta^4) = X^2+X+\theta^5 \in \mathbb{F}_4[X],\\
m_2 = m_8 & = & (X-\theta^2)(X-\theta^8) = X^2+X+\theta^{10} \in \mathbb{F}_4[X],\\
m_3 = m_{12} & = & (X-\theta^3)(X-\theta^{12}) = X^2+\theta^{10}X+1 \in \mathbb{F}_4[X],\\
m_5 & = & X-\theta^5 \in \mathbb{F}_4[X],\\
m_6 = m_9 & = & (X-\theta^6)(X-\theta^9) = X^2+\theta^5X+1 \in \mathbb{F}_4[X],\\
m_7 = m_{13} & = & (X-\theta^7)(X-\theta^{13}) = X^2+\theta^5X+\theta^5 \in \mathbb{F}_4[X],\\
m_{10} & = & X-\theta^{10} \in \mathbb{F}_4[X],\\
m_{11} = m_{14} & = & (X-\theta^{11})(X-\theta^{14}) = X^2+\theta^{10}X+\theta^{10} \in \mathbb{F}_4[X].
\end{array}$$

Rekomendujemy Czytelnikowi, aby porównał otrzymane wyniki z wynikami przykładu (**2**).

(**4**) Znajdźmy warstwy cyklotomiczne liczby 3 modulo 26. Mamy $26 = 3^3 - 1$. Obliczamy, że:

$$\begin{array}{lcl}
i=0 & \Rightarrow & 0\cdot 3^s \equiv 0 \pmod{26} \quad (s\in\mathbb{N}) \quad \Rightarrow\\
 & \Rightarrow & C_0 = \{0\},\\
i=1 & \Rightarrow & 1\cdot 3^0 = 1,\ 1\cdot 3^1 = 3,\ 1\cdot 3^2 = 9,\ 1\cdot 3^3 \equiv 1 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_1 = \{1,3,9\},\\
i=2 & \Rightarrow & 2\cdot 3^0 = 2,\ 2\cdot 3^1 = 6,\ 2\cdot 3^2 = 18,\ 2\cdot 3^3 \equiv 2 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_2 = \{2,6,18\},\\
i=4 & \Rightarrow & 4\cdot 3^0 = 4,\ 4\cdot 3^1 = 12,\ 4\cdot 3^2 \equiv 10,\ 4\cdot 3^3 \equiv 4 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_4 = \{4,12,10\},\\
i=5 & \Rightarrow & 5\cdot 3^0 = 5,\ 5\cdot 3^1 = 15,\ 5\cdot 3^2 \equiv 19,\ 5\cdot 3^3 \equiv 5 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_5 = \{5,15,19\},\\
i=7 & \Rightarrow & 7\cdot 3^0 = 7,\ 7\cdot 3^1 = 21,\ 7\cdot 3^2 \equiv 11,\ 7\cdot 3^3 \equiv 7 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_7 = \{7,21,11\},\\
i=8 & \Rightarrow & 8\cdot 3^0 = 8,\ 8\cdot 3^1 = 24,\ 8\cdot 3^2 \equiv 20,\ 8\cdot 3^3 \equiv 8 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_8 = \{8,24,20\},\\
i=13 & \Rightarrow & 13\cdot 3^0 = 13,\ 13\cdot 3^1 \equiv 13 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_{13} = \{13\},\\
i=14 & \Rightarrow & 14\cdot 3^0 = 14,\ 14\cdot 3^1 \equiv 16,\ 14\cdot 3^2 \equiv 22,\ 14\cdot 3^3 \equiv 14 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_{14} = \{14,16,22\},\\
i=17 & \Rightarrow & 17\cdot 3^0 = 17,\ 17\cdot 3^1 \equiv 25,\ 17\cdot 3^2 \equiv 23,\ 17\cdot 3^3 \equiv 17 \pmod{26} \quad \Rightarrow\\
 & \Rightarrow & C_{17} = \{17,25,23\}.
\end{array}$$

Zatem mamy takie wielomiany minimalne (nad ciałem $\mathbb{F}_3$) elementów ciała $\mathbb{F}_{27}$ (tutaj θ jest elementem prymitywnym ciała $\mathbb{F}_{27}$):

$$\begin{array}{lcllcll}
m_0 & = & (X-\theta^0) & = & (X-1) & = & (X+2) \in \mathbb{F}_3[X],\\
m_1 & = & (X-\theta)(X-\theta^3)(X-\theta^9) & = & m_3 & = & m_9 \in \mathbb{F}_3[X],\\
m_2 & = & (X-\theta^2)(X-\theta^6)(X-\theta^{18}) & = & m_6 & = & m_{18} \in \mathbb{F}_3[X],\\
m_4 & = & (X-\theta^4)(X-\theta^{12})(X-\theta^{10}) & = & m_{12} & = & m_{10} \in \mathbb{F}_3[X],\\
m_5 & = & (X-\theta^5)(X-\theta^{15})(X-\theta^{19}) & = & m_{15} & = & m_{19} \in \mathbb{F}_3[X],\\
m_7 & = & (X-\theta^7)(X-\theta^{21})(X-\theta^{11}) & = & m_{21} & = & m_{11} \in \mathbb{F}_3[X],\\
m_8 & = & (X-\theta^8)(X-\theta^{24})(X-\theta^{20}) & = & m_{24} & = & m_{20} \in \mathbb{F}_3[X],\\
m_{13} & = & (X-\theta^{13}) \in \mathbb{F}_3[X], & & & & \\
m_{14} & = & (X-\theta^{14})(X-\theta^{16})(X-\theta^{22}) & = & m_{16} & = & m_{22} \in \mathbb{F}_3[X],\\
m_{17} & = & (X-\theta^{17})(X-\theta^{25})(X-\theta^{23}) & = & m_{25} & = & m_{23} \in \mathbb{F}_3[X].
\end{array}$$

Możemy teraz obliczyć, na przykład, że $m_2 = X^3+X^2+X+2 \in \mathbb{F}_3[X]$ oraz $m_8 = X^3+2X^2+2X+2 \in \mathbb{F}_3[X]$. Proponujemy Czytelnikowi znaleźć inne wielomiany minimalne w takiej postaci.

* * *

■ **Przypadek $\mathbf{n \neq q^r - 1}$.** Ważne miejsce posiada następujące

Twierdzenie 7.8.6. *Niech* $\mathrm{NWD}(n, q) = 1$, *n dzieli $q^r - 1$ oraz $\theta \in \mathbb{F}_{q^r}$ będzie elementem prymitywnym. Jeśli $\{c_1, \ldots, c_l\}$ jest pełnym zbiorem reprezentantów warstw cyklotomicznych liczby q modulo n, to*

$$X^n - 1 = \prod_{i=1}^{l} m_{\frac{(q^r-1)c_i}{n}},$$

gdzie $m_j \in \mathbb{F}_q[X]$ jest wielomianem minimalnym elementu θ^j.

Dowód. Niech $a = \frac{q^r-1}{n}$. Skoro θ^a jest pierwiastkiem prymitywnym stopnia n z jedności 1, to zbiór

$$\theta^0, \theta^a, \theta^{2a}, \ldots, \theta^{(n-1)a}$$

składa się ze wszystkich pierwiastków wielomianu $X^n - 1 \in \mathbb{F}_q[X]$ oraz

$$m_{ia} \quad \text{dzieli} \quad X^n - 1 \quad (0 \leqslant i \leqslant n-1).$$

Wtedy

$$X^n - 1 = \mathrm{NWW}(m_0, m_a, , \ldots, m_{(n-1)a}).$$

W wyniku twierdzenia 7.8.3 wnosimy, że

$m_{ia} = m_{ja} \Leftrightarrow$ i oraz j leżą w tej samej warstwie cyklotomicznej liczby q modulo $q^r - 1$,

a zatem wielomiany $m_{c_1 a}, m_{c_2 a}, \ldots, m_{c_l a}$ są parami różne oraz

$$X^n - 1 = m_{c_1 a} m_{c_2 a} \cdots m_{c_l a}.$$

□

■ Biorąc pod uwagę twierdzenie 7.8.6, otrzymujemy taki `algorytm faktoryzacji wielomianu` $X^n - 1$ nad ciałem $\mathbb{F}_q$, gdzie $\mathrm{NWD}(n, q) = 1$.

1°. Znajdujemy najmniejszą liczbę całkowitą dodatnią r taką, że n dzieli $q^r - 1$.

2°. Obliczamy dokładnie warstwy cyklotomiczne liczby q modulo $q^r - 1$ takie, że ich reprezentanty są podzielne przez liczbę $\frac{q^r-1}{n}$. Niech C będzie pełnym zbiorem reprezentantów takich warstw.

3°. Niech θ będzie elementem prymitywnym ciała $\mathbb{F}_{q^r}$. Obliczamy wielomiany minimalne m_i elementów θ^i nad ciałem $\mathbb{F}_q$, gdzie $i \in C$.

4°. Otrzymujemy rozkład wielomianu $X^n - 1$:

$$X^n - 1 = \prod_{i \in C} m_i.$$

■ Przytoczmy taką tablicę rozkładów wielomianu dwójkowego $X^n - 1 \in \mathbb{F}_2[X]$ na czynniki nieprzywiedlne:

n	rozkład binarnego wielomianu $X^n - 1 \in \mathbb{F}_2[X]$
1	$X + 1$
2	$(X + 1)^2$
3	$(X + 1)(X^2 + X + 1)$
4	$(X + 1)^4$
5	$(X + 1)(X^4 + X^3 + X^2 + X + 1)$
6	$(X + 1)^2(X^2 + X + 1)^2$
7	$(X + 1)(X^3 + X + 1)(X^3 + X^2 + 1)$
8	$(X + 1)^8$
9	$(X + 1)(X^2 + X + 1)(X^6 + X^3 + 1)$
10	$(X + 1)^2(X^4 + X^3 + X^2 + X + 1)$
11	$(X + 1)(X^{10} + \cdots + X + 1)$
12	$(X + 1)^4(X^2 + X + 1)^4$
13	$(X + 1)(X^{12} + \cdots + X + 1)$
14	$(X + 1)^2(X^3 + X + 1)^2(X^3 + X^2 + 1)$
15	$(X + 1)(X^2 + X + 1)(X^4 + X^3 + X^2 + X + 1)(X^4 + X + 1)(X^4 + X^3 + 1)$
16	$(X + 1)^{16}$
17	$(X + 1)(X^8 + X^7 + X^6 + X^4 + X^2 + X + 1)(X^8 + X^5 + X^4 + X^3 + 1)$
18	$(X + 1)^2(X^2 + X + 1)^2(X^6 + X^3 + 1)^2$
19	$(X + 1)(X^{18} + \cdots + X + 1)$
20	$(X + 1)^4(X^4 + X^3 + X^2 + X + 1)^4$

Przykłady 7.8.7.

(**1**) Rozłóżmy wielomian $X^{51} - 1 \in \mathbb{F}_2[X]$ na czynniki nieprzywiedlne nad ciałem $\mathbb{F}_2$. Tutaj $q = 2$ oraz najmniejszą liczbą r, dla której 51 dzieli $2^r - 1$, jest $r = 8$. Niech θ będzie elementem prymitywnym ciała $\mathbb{F}_{256}$ oraz $m_i \in \mathbb{F}_2[X]$ będzie wielomianem minimalnym elementu θ^i. Wtedy

$$\frac{q^r - 1}{n} = \frac{2^8 - 1}{51} = 5$$

oraz mamy taką listę A elementów ze zbioru $\{0, 1, \ldots, 254\}$, które są podzielne przez 5:

$$\begin{array}{rrrrrrrrrrrrrr} 0, & 5, & 10, & 15, & 20, & 25, & 30, & 35, & 40, & 45, & 50, & 55, & 60, & 65, \\ 70, & 75, & 80, & 85, & 90, & 95, & 100, & 105, & 110, & 115, & 120, & 125, & 130, & 135, \\ 140, & 145, & 150, & 155, & 160, & 165, & 170, & 175, & 180, & 185, & 190, & 195, & 200, & 205, \\ 210, & 215, & 220, & 225, & 230, & 235, & 240, & 245, & 250. & & & & & \end{array}$$

Obliczamy warstwy liczby 2 modulo 255, reprezentanty których są podzielne przez 5:

- $\bullet_1$ liczba 5 dzieli $0 \in A$, a więc

$$i = 0 \;\Rightarrow\; 0 \cdot 2^s \equiv 0 \pmod{255}\ (s \in \mathbb{N}) \quad\Rightarrow\quad C_0 = \{0\};$$

- $\bullet_2$ liczba 5 dzieli $5 \in A \setminus C_0$, a zatem

$$i = 5 \;\Rightarrow\; 5 \cdot 2^0 \equiv 5,\ 5 \cdot 2^1 \equiv 10,\ 5 \cdot 2^2 \equiv 20,\ 5 \cdot 2^3 \equiv 40,\ 5 \cdot 2^4 \equiv 80,\ 5 \cdot 2^5 \equiv 160,$$
$$5 \cdot 2^6 \equiv 65,\ 5 \cdot 2^7 \equiv 130,\ 5 \cdot 2^8 \equiv 5 \pmod{255} \quad\Rightarrow\quad C_5 = \{5, 10, 20, 40, 80, 160, 65, 130\};$$

- $\bullet_3$ liczba 5 dzieli $15 \in A \setminus (C_0 \cup C_5)$, a więc

$$i = 15 \;\Rightarrow\; 15 \cdot 2^0 \equiv 15,\ 15 \cdot 2^1 \equiv 30,\ 15 \cdot 2^2 \equiv 60,\ 15 \cdot 2^3 \equiv 120,$$
$$15 \cdot 2^4 \equiv 240,\ 15 \cdot 2^5 \equiv 225,\ 15 \cdot 2^6 \equiv 195,$$
$$51 \cdot 2^7 \equiv 135,\ 15 \cdot 2^8 \equiv 15 \pmod{255} \quad\Rightarrow\quad C_{15} = \{15, 30, 60, 120, 240, 225, 195, 135\};$$

- $\bullet_4$ liczba 5 dzieli $25 \in A \setminus (C_0 \cup C_5 \cup C_{15})$, a zatem

$$i = 25 \;\Rightarrow\; 25 \cdot 2^0 \equiv 25,\ 25 \cdot 2^1 \equiv 50,\ 25 \cdot 2^2 \equiv 100,\ 25 \cdot 2^3 \equiv 200,$$
$$25 \cdot 2^4 \equiv 145,\ 25 \cdot 2^5 \equiv 35,\ 25 \cdot 2^6 \equiv 70,$$
$$25 \cdot 2^7 \equiv 140,\ 25 \cdot 2^8 \equiv 25 \pmod{255} \quad\Rightarrow\quad C_{25} = \{25, 50, 100, 200, 145, 35, 70, 140\};$$

- $\bullet_5$ liczba 5 dzieli $45 \in A \setminus (C_0 \cup C_5 \cup C_{15} \cup C_{25})$, a więc

$$i = 45 \;\Rightarrow\; 45 \cdot 2^0 \equiv 45,\ 45 \cdot 2^1 \equiv 90,\ 45 \cdot 2^2 \equiv 180,\ 45 \cdot 2^3 \equiv 105,$$
$$45 \cdot 2^4 \equiv 210,\ 45 \cdot 2^5 \equiv 165,\ 45 \cdot 2^6 \equiv 75,$$
$$45 \cdot 2^7 \equiv 150,\ 45 \cdot 2^8 \equiv 45 \pmod{255} \quad\Rightarrow\quad C_{45} = \{45, 90, 180, 105, 210, 165, 75, 150\};$$

- $\bullet_6$ liczba 5 dzieli $55 \in A \setminus (C_0 \cup C_5 \cup C_{15} \cup C_{25} \cup C_{45})$, a więc

$$i = 55 \;\Rightarrow\; 55 \cdot 2^0 \equiv 55,\ 55 \cdot 2^1 \equiv 110,\ 55 \cdot 2^2 \equiv 220,\ 55 \cdot 2^3 \equiv 185,$$
$$55 \cdot 2^4 \equiv 115,\ 55 \cdot 2^5 \equiv 230,\ 55 \cdot 2^6 \equiv 205,$$
$$55 \cdot 2^7 \equiv 155,\ 55 \cdot 2^8 \equiv 55 \pmod{255} \quad\Rightarrow\quad C_{55} = \{55, 110, 220, 185, 115, 230, 205, 155\};$$

- $\bullet_7$ liczba 5 dzieli $85 \in A \setminus (C_0 \cup C_5 \cup C_{15} \cup C_{25} \cup C_{45} \cup C_{55})$, a więc

$$i = 85 \;\Rightarrow\; 85 \cdot 2^0 \equiv 85,\ 85 \cdot 2^1 \equiv 170,\ 85 \cdot 2^2 \equiv 85 \quad\Rightarrow\quad C_{85} = \{85, 170\};$$

- $\bullet_8$ liczba 5 dzieli $95 \in A \setminus (C_0 \cup C_5 \cup C_{15} \cup C_{25} \cup C_{45} \cup C_{55} \cup C_{85})$, a więc

$$i = 95 \;\Rightarrow\; 95 \cdot 2^0 \equiv 95,\ 95 \cdot 2^1 \equiv 190,\ 95 \cdot 2^2 \equiv 125,\ 95 \cdot 2^3 \equiv 250,$$
$$95 \cdot 2^4 \equiv 245,\ 95 \cdot 2^5 \equiv 235,\ 95 \cdot 2^6 \equiv 215,$$
$$95 \cdot 2^7 \equiv 175,\ 95 \cdot 2^8 \equiv 95 \pmod{255} \quad\Rightarrow\quad C_{95} = \{95, 190, 125, 250, 245, 235, 215, 175\}.$$

Zatem

$$A = C_0 \bigcup C_5 \bigcup C_{15} \bigcup C_{25} \bigcup C_{45} \bigcup C_{55} \bigcup C_{85} \bigcup C_{95}$$

oraz

$$X^{51} - 1 = m_0 m_5 m_{15} m_{25} m_{45} m_{55} m_{85} m_{95},$$

gdzie

$$\begin{aligned}
m_0 &= \prod_{i \in C_0} (X - \theta^i) = (X-1) \in \mathbb{F}_2[X] && (\deg m_0 = |C_0| = 1),\\
m_5 &= \prod_{i \in C_5} (X - \theta^i) = (X-\theta^5)(X-\theta^{10})(X-\theta^{20})(X-\theta^{40})\cdot\\
&\quad \cdot(X-\theta^{65})(X-\theta^{80})(X-\theta^{130})(X-\theta^{160}) \in \mathbb{F}_2[X] && (\deg m_5 = |C_5| = 8),\\
m_{15} &= \prod_{i \in C_{15}} (X - \theta^i) = (X-\theta^{15})(X-\theta^{30})(X-\theta^{60})(X-\theta^{120})\cdot\\
&\quad \cdot(X-\theta^{135})(X-\theta^{195})(X-\theta^{225})(X-\theta^{240}) \in \mathbb{F}_2[X] && (\deg m_{15} = |C_{15}| = 8),\\
m_{25} &= \prod_{i \in C_{25}} (X - \theta^i) = (X-\theta^{25})(X-\theta^{35})(X-\theta^{50})(X-\theta^{70})\cdot\\
&\quad \cdot(X-\theta^{100})(X-\theta^{140})(X-\theta^{145})(X-\theta^{200}) \in \mathbb{F}_2[X] && (\deg m_{25} = |C_{25}| = 8),\\
m_{45} &= \prod_{i \in C_{45}} (X - \theta^i) = (X-\theta^{45})(X-\theta^{75})(X-\theta^{90})(X-\theta^{105})\cdot\\
&\quad \cdot(X-\theta^{150})(X-\theta^{165})(X-\theta^{180})(X-\theta^{210}) \in \mathbb{F}_2[X] && (\deg m_{45} = |C_{45}| = 8),\\
m_{55} &= \prod_{i \in C_{55}} (X - \theta^i) = (X-\theta^{55})(X-\theta^{110})(X-\theta^{115})(X-\theta^{155})\cdot\\
&\quad \cdot(X-\theta^{185})(X-\theta^{205})(X-\theta^{220})(X-\theta^{230}) \in \mathbb{F}_2[X] && (\deg m_{55} = |C_{55}| = 8),\\
m_{85} &= \prod_{i \in C_{85}} (X - \theta^i) = (X-\theta^{85})(X-\theta^{170}) \in \mathbb{F}_2[X] && (\deg m_{85} = |C_{85}| = 2),\\
m_{95} &= \prod_{i \in C_{95}} (X - \theta^i) = (X-\theta^{95})(X-\theta^{125})(X-\theta^{175})(X-\theta^{190})\cdot\\
&\quad \cdot(X-\theta^{215})(X-\theta^{235})(X-\theta^{245})(X-\theta^{250}) \in \mathbb{F}_2[X] && (\deg m_{95} = |C_{95}| = 8).
\end{aligned}$$

(**2**) Rozłóżmy wielomian $X^9 - 1 \in \mathbb{F}_2[X]$ na czynniki nieprzywiedlne (nad ciałem $\mathbb{F}_2$). Przekonujemy się, że najmniejsze $r \in \mathbb{N}^*$ takie, że 9 dzieli $2^r - 1$ jest liczbą $r = 6$. Niech $\mathbb{F}_{2^6} = \langle\theta\rangle$ oraz $m_i \in \mathbb{F}_2[X]$ będzie wielomianem minimalnym elementu θ^i. Skoro

$$\frac{q^r - 1}{n} = \frac{2^6 - 1}{9} = 7$$

oraz 7 dzieli $\frac{63 \cdot c_i}{9}$, to obliczamy warstwy cyklotomiczne liczby 2 modulo 63 takie, że ich reprezentanty dzielą liczbę 7. Wybieramy z listy

$$A = \{0, 1, \ldots, 62\}$$

liczby, które są podzielne przez 7 i których nie było w poprzednich warstwach:

- $\bullet_1$ liczba 7 dzieli $0 \in A$, a więc

$$i = 0 \;\Rightarrow\; 0 \cdot 2^s \equiv 0 \pmod{63} \quad (s \in \mathbb{N}) \quad \Rightarrow \quad C_0 = \{0\};$$

- $\bullet_2$ liczba 7 dzieli $7 \in A \setminus C_0$, a zatem

$$\begin{aligned}
&i = 7 \;\Rightarrow\; 7 \cdot 2^0 \equiv 7,\; 7 \cdot 2^1 \equiv 14,\; 7 \cdot 2^2 \equiv 28,\; 7 \cdot 2^3 \equiv 56,\; 7 \cdot 2^4 \equiv 49,\; 7 \cdot 2^5 \equiv 35,\\
&7 \cdot 2^6 \equiv 7 \pmod{63} \quad \Rightarrow \quad C_7 = \{7, 14, 28, 56, 49, 35\};
\end{aligned}$$

- $\bullet_3$ liczba 7 dzieli $21 \in A \setminus (C_0 \cup C_7)$, a więc

$$i = 21 \;\Rightarrow\; 21 \cdot 2^0 \equiv 21,\; 21 \cdot 2^1 \equiv 42,\; 21 \cdot 2^2 \equiv 21 \quad \Rightarrow \quad C_{21} = \{21, 42\}.$$

Otrzymaliśmy

$$A = C_0 \bigcup C_7 \bigcup C_{21},$$

a zatem

$$X^9 - 1 = m_0 m_7 m_{21},$$

gdzie

$$\begin{aligned}
m_0 &= \prod_{i \in C_0} (X - \theta^i) = (X-1) \in \mathbb{F}_2[X] && (\deg m_0 = |C_0| = 1),\\
m_7 &= \prod_{i \in C_7} (X - \theta^i) = (X-\theta^7)(X-\theta^{14})(X-\theta^{28})(X-\theta^{35})\cdot\\
&\quad \cdot(X-\theta^{49})(X-\theta^{56}) = X^6 + X + 1 \in \mathbb{F}_2[X] && (\deg m_7 = |C_7| = 6),\\
m_{21} &= \prod_{i \in C_{21}} (X - \theta^i) = (X-\theta^{21})(X-\theta^{42}) = X^2 + X + 1 \in \mathbb{F}_2[X] && (\deg m_{21} = |C_{21}| = 2).
\end{aligned}$$

(**3**) Znajdźmy rozkład wielomianu $X^{21} - 1$ nad ciałem dwójkowym $\mathbb{F}_2$. Liczba $r = 6$ jest najmniejszą taką, że 21 dzieli $2^6 - 1$. Niech $\theta \in \mathbb{F}_{64}$ będzie elementem prymitywnym oraz $m_i \in \mathbb{F}_2[X]$ będzie wielomianem minimalnym elementu θ^i. Ponieważ

$$\frac{q^r - 1}{n} = \frac{2^6 - 1}{21} = 3,$$

to 3 dzieli $\frac{63 \cdot c_i}{21}$ (patrz twierdzenie 7.8.6), a zatem obliczamy dokładnie warstwy cyklotomiczne C_i liczby 2 modulo 63 takie, że ich reprezentanty są podzielne przez 3. To znaczy, że ze zbioru $\{0, 1, \ldots, 62\}$ wybieramy liczby podzielne przez 3; otrzymujemy zbiór

$$A = \{0, 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60\}.$$

Obliczamy:

- $\bullet_1$ liczba 3 dzieli 0, a więc

$$i = 0 \;\Rightarrow\; 0 \cdot 2^s \equiv 0 \pmod{63} \quad (s \in \mathbb{N}) \quad \Rightarrow \quad C_0 = \{0\};$$

- $\bullet_2$ liczba 3 dzieli $3 \in A \setminus C_0$, a zatem

$$\begin{aligned}
&i = 3 \;\Rightarrow\; 3 \cdot 2^0 \equiv 3,\; 3 \cdot 2^1 \equiv 6,\; 3 \cdot 2^2 \equiv 12,\; 3 \cdot 2^3 \equiv 24,\; 3 \cdot 2^4 \equiv 48,\; 3 \cdot 2^5 \equiv 33,\\
&3 \cdot 3^6 \equiv 3 \pmod{63} \quad \Rightarrow \quad C_3 = \{3, 6, 12, 24, 48, 33\};
\end{aligned}$$

- $\bullet_3$ liczba 3 dzieli $9 \in A \setminus (C_0 \cup C_3)$, a więc

$$\begin{aligned}
&i = 9 \;\Rightarrow\; 9 \cdot 2^0 \equiv 9,\; 9 \cdot 2^1 \equiv 18,\; 9 \cdot 2^2 \equiv 36,\\
&9 \cdot 2^3 \equiv 9 \pmod{63} \quad \Rightarrow \quad C_9 = \{9, 18, 36\};
\end{aligned}$$

- $\bullet_4$ liczba 3 dzieli $15 \in A \setminus (C_0 \cup C_3 \cup C_9)$, a zatem

$$\begin{aligned}
&i = 15 \;\Rightarrow\; 15 \cdot 2^0 \equiv 15,\; 15 \cdot 2^1 \equiv 30,\; 15 \cdot 2^2 \equiv 60,\; 15 \cdot 2^3 \equiv 57, 15 \cdot 2^4 \equiv 51,\\
&15 \cdot 2^5 \equiv 39,\; 15 \cdot 2^6 \equiv 15 \pmod{63} \quad \Rightarrow \quad C_{15} = \{15, 30, 60, 57, 51, 39\};
\end{aligned}$$

- $\bullet_5$ liczba 3 dzieli $21 \in A \setminus (C_0 \cup C_3 \cup C_9 \cup C_{15})$, a więc

$$i = 21 \;\Rightarrow\; 21 \cdot 2^0 \equiv 21,\; 21 \cdot 2^1 \equiv 42,\; 21 \cdot 2^2 \equiv 21 \pmod{63} \quad \Rightarrow \quad C_{21} = \{21, 42\};$$

•6 liczba 3 dzieli $27 \in A \setminus (C_0 \cup C_3 \cup C_9 \cup C_{15} \cup C_{21})$, a zatem

$$\begin{array}{c} i = 27 \ \Rightarrow \ 27 \cdot 2^0 \equiv 27, \ 27 \cdot 2^1 \equiv 54, \ 27 \cdot 2^2 \equiv 45, \\ 27 \cdot 2^3 \equiv 27 \pmod{63} \quad \Rightarrow \quad C_{27} = \{27, 54, 45\}. \end{array}$$

Otrzymaliśmy

$$A = C_0 \bigcup C_3 \bigcup C_9 \bigcup C_{15} \bigcup C_{21} \bigcup C_{27},$$

a zatem

$$X^{21} - 1 = m_0 m_3 m_9 m_{15} m_{21} m_{27},$$

gdzie

$$\begin{array}{rcll} m_0 & = & \prod\limits_{i \in C_0} (X - \theta^i) = (X - 1) \in \mathbb{F}_2[X] & (\deg m_0 = |C_0| = 1), \\ m_3 & = & \prod\limits_{i \in C_3} (X - \theta^i) = (X - \theta^3)(X - \theta^6)(X - \theta^{12}) \cdot & \\ & \cdot & (X - \theta^{24})(X - \theta^{33})(X - \theta^{48}) = & \\ & = & X^6 + X^4 + X^2 + X + 1 \in \mathbb{F}_2[X] & (\deg m_3 = |C_3| = 6), \\ m_9 & = & \prod\limits_{i \in C_9} (X - \theta^i) = (X - \theta^9)(X - \theta^{18})(X - \theta^{36}) = & \\ & = & X^3 + X^2 + 1 \in \mathbb{F}_2[X] & (\deg m_9 = |C_9| = 3), \\ m_{15} & = & \prod\limits_{i \in C_{15}} (X - \theta^i) = (X - \theta^{15})(X - \theta^{30})(X - \theta^{39}) \cdot & \\ & \cdot & (X - \theta^{51})(X - \theta^{57})(X - \theta^{60}) = & \\ & = & X^6 + X^5 + X^4 + X^2 + 1 \in \mathbb{F}_2[X] & (\deg m_{15} = |C_{15}| = 6), \\ m_{21} & = & \prod\limits_{i \in C_{21}} (X - \theta^i) = (X - \theta^{21})(X - \theta^{42}) = & \\ & = & X^2 + X + 1 \in \mathbb{F}_2[X] & (\deg m_{21} = |C_{21}| = 2), \\ m_{27} & = & \prod\limits_{i \in C_{27}} (X - \theta^i) = (X - \theta^{27})(X - \theta^{45})(X - \theta^{54}) = & \\ & = & X^3 + X + 1 \in \mathbb{F}_2[X] & (\deg m_{27} = |C_{27}| = 3). \end{array}$$

Twierdzenie 7.8.8. *Niech* $\mathrm{NWD}(n, q) = 1$. *Istnieje bijekcja między unormowanymi wielomianami nieprzywiedlnymi* $f_i \in \mathbb{F}_q[X]$ *z rozkładu*

$$X^n - 1 = \prod_{i=0}^{l-1} f_i$$

a różnymi warstwami cyklotomicznymi

$$C_j = \{jq^i \pmod{n} \mid i = 0, 1, \ldots, n-1\}$$

liczby q *modulo* n, *przy tym stopień* $\deg f_i$ *jest liczbą elementów warstwy cyklotomicznej* C_j *odpowiadającej wielomianowi* f_i (*czyli liczba unormowanych czynników nieprzywiedlnych wielomianu* $X^n - 1 \in \mathbb{F}_q[X]$ *jest równa liczbie warstw cyklotomicznych liczby* q *modulo* n).

Dowód. Niech θ będzie pierwiastkiem wielomianu $X^n - 1 \in \mathbb{F}_q[X]$ (z pewnego rozszerzenia ciała $\mathbb{F}_q$). Wtedy $\theta^n = 1$, a więc $(\theta^m)^n = 1$ dla każdej liczby całkowitej m. To znaczy, że θ^m również jest pierwiastkiem wielomianu $X^n - 1$. Łatwo zauważyć, że

$$\theta^0, \theta^1, \ldots, \theta^{n-1}$$

są parami różnymi pierwiastkami wielomianu $X^n - 1$. Skoro pierwiastki wielomianu f_i są pierwiastkami wielomianu $X^n - 1$, to istnieje j $(0 \leqslant j \leqslant n-1)$ takie, że θ^j jest pierwiastkiem wielomianu f_i. Wtedy

$$\theta^{jq^s} \pmod{n}$$

także jest pierwiastkiem wielomianu f_i dla każdego $s \in \mathbb{N}$, a więc wielomianowi f_i bijektywnie odpowiada warstwa cyklotomiczna C_j.

Stopień $\deg f_i$ jest liczbą pierwiastków wielomianu f_i (w pewnym rozszerzeniu ciała $\mathbb{F}_q$), dla każdego reprezentanta $r \in C_j$ element θ^r jest pierwiastkiem wielomianu f_i oraz teza zachodzi. □

Ćwiczenia 7.8.9.

(**1**) Rozłożyć na czynniki nieprzywiedlne wielomian $X^n - 1 \in \mathbb{F}_p[X]$, jeśli:
(a) $n = 7$ oraz $p = 2$;
(b) $n = 7$ oraz $p = 3$;
(c) $n = 8$ oraz $p = 2$;
(d) $n = 8$ oraz $p = 3$;
(e) $n = 8$ oraz $p = 5$;
(f) $n = 6$ oraz $p = 2$;
(g) $n = 6$ oraz $p = 3$;
(h) $n = 5$ oraz $p = 2$;
(k) $n = 5$ oraz $p = 3$;
(l) $n = 3$ oraz $p = 5$;
(m) $n = 7$ oraz $p = 5$.
(**2**) Rozłożyć na czynniki nieprzywiedlne wielomian $X^n - 1 \in \mathbb{F}_p[X]$, jeśli:
(a) $n = 15$ oraz $p = 2$;
(b) $n = 23$ oraz $p = 2$;
(c) $n = 11$ oraz $p = 3$;
(d) $n = 4$ oraz $p = 2$;
(e) $n = 13$ oraz $p = 3$;
(f) $n = 21$ oraz $p = 2$;
(g) $n = 31$ oraz $p = 2$;
(h) $n = 12$ oraz $p = 5$;
(i) $n = 24$ oraz $p = 7$;
(j) $n = 9$ oraz $p = 2$.
(**3**) Znaleźć wielomiany minimalne wszystkich elementów ciała $\mathbb{F}_q$ i rozłożyć jego elementy na warstwy cyklotomiczne nad ciałem $\mathbb{F}_p$, jeśli:

(a) $q = 243$ oraz $p = 3$;
(b) $q = 16$ oraz $p = 2$;
(c) $q = 8$ oraz $p = 2$;
(d) $q = 25$ oraz $p = 5$;
(e) $q = 49$ oraz $p = 7$.
(**4**) Zbudować warstwy cyklotomiczne: (a) modulo 23 nad ciałem $\mathbb{F}_2$; (b) modulo 17 nad ciałem $\mathbb{F}_2$. Rozłożyć wielomiany: (c) $X^{23} - 1$; (d) $X^{17} - 1$ w iloczyn wielomianów minimalnych nad $\mathbb{F}_2$.
(**5**) Rozłożyć na czynniki nieprzywiedlne wielomian $X^n - 1 \in \mathbb{F}_q[X]$, jeśli:
(a) $n = 8$ oraz $q = 4$;
(b) $n = 10$ oraz $q = 2$;
(c) $n = 10$ oraz $q = 3$;
(d) $n = 13$ oraz $q = 2$;
(e) $n = 15$ oraz $q = 3$;
(f) $n = 24$ oraz $q = 2$.
(**6**) Rozłożyć na czynniki nieprzywiedlne wielomian $X^n - 1 \in \mathbb{F}_2[X]$, jeśli:
(a) $n = 9$;
(b) $n = 73$;
(c) $n = 85$;
(d) $n = 18$.
(**7**) Znaleźć liczbę elementów prymitywnych w ciele $\mathbb{F}_{2^n}$, jeśli:
(a) $n = 5$;
(b) $n = 6$;
(c) $n = 7$;
(d) $n = 9$.

Uwagi. Pierwszy algorytm faktoryzacji wielomianów opublikował T. von Schubert[(10)] w 1793 r. Epoka komputerowa z 1965 r. przyspieszyła poszukiwania algorytmów faktoryzacji nad różnymi ciałami.

[(10)] Teodor von Schubert (1758–1825)

Bibliografia

[1] Artemowicz O., Piękosz A., *Algebra*, Wyd. PK, Kraków 2010.

[2] Bagiński Cz., *Wstęp do teorii grup*, SCRIPT, Warszawa 2002.

[3] Białynicki-Birula A., *Algebra*, wyd. III, PWN, Warszawa 1987.

[4] Birkhoff G., Thomas C.B., *Współczesna algebra stosowana*, PWN, Warszawa 1983.

[5] Browkin J., *Teoria ciał* (Biblioteka Matematyczna 49), PWN, Warszawa 1978.

[6] Bryński M., Jurkiewicz J., *Zbiór zadań z algebry*, PWN, Warszawa 1982.

[7] Dummit D.S., Foote R.M., *Abstract algebra*, Prentice Hall, Upper Saddle River, NJ 1991.

[8] Gancarzewicz J., *Arytmetyka: podręcznik dla licencjatów*, Wyd. UJ, Kraków 2000.

[9] Gilbert W.J., Nicholson W.K., *Algebra współczesna z zastosowaniami*, WNT, Warszawa 2008.

[10] Gleichgewicht B., *Algebra*, GiS, Wrocław 2002.

[11] Klin M.Ch., Pöchel R., Rosenbaum K., *Algebra stosowana dla matematyków i informatyków*, WNT, Warszawa 1993.

[12] Koblitz N., *Algebraiczne aspekty kryptografii*, WNT, Warszawa 2000.

[13] Kostrikin A., *Wstęp do algebry*, (t. 1-3), PWN, Warszawa 2004–2005.

[14] Kostrikin A., *Zbiór zadań z algebry*, PWN, Warszawa 2005.

[15] Lang S., *Algebra*, Addison-Wesley, Mass. 1970 (przekład polski: Algebra, PWN, Warszawa 1995).

[16] Lidl R., Niederreiter H., *Finite fields*, Addison-Wesley, Reading 1983.

[17] Piękosz A., *Algebra liniowa*, Wyd. PK, Kraków 2009.

[18] Piękosz A., *Wstęp do matematyki*, Wyd. PK, Kraków 2010.

[19] Rutkowski J., *Algebra abstrakcyjna w zadaniach*, PWN, Warszawa 2000.

[20] Rutkowski J., *Teoria liczb w zadaniach*, PWN, Warszawa 2018.
[21] Więsław W., *Grupy, pierścienie, ciała*, UW, Wrocław 1979.

Spis oznaczeń

$\forall$	kwantyfikator ogólny (uniwersalny)
$\exists$	kwantyfikator szczególny (egzystencjalny)
$\Rightarrow$	symbol implikacji
$\Leftrightarrow$	symbol równoważności
$\cong$	symbol izomorficzności
$n!$	n silnia
$\binom{n}{k}$	współczynnik dwumianowy (symbol Newtona)
$\mathbb{N}$	zbiór liczb naturalnych $\{0, 1, 2, \ldots\}$
$\mathbb{N}^*$	zbiór liczb naturalnych dodatnich $\mathbb{N}\setminus\{0\} = \{1, 2, \ldots\}$
$\mathbb{Z}$	zbiór liczb całkowitych
$\mathbb{Q}$	zbiór liczb wymiernych
$\mathbb{R}$	zbiór liczb rzeczywistych
$\mathbb{C}$	zbiór liczb zespolonych
$\mathbb{Q}^*$	zbiór liczb wymiernych niezerowych
$\mathbb{R}^*$	zbiór liczb rzeczywistych niezerowych
$\mathbb{C}^*$	zbiór liczb zespolonych niezerowych
$\mathbb{R}_+$	zbiór liczb rzeczywistych dodatnich
$\mathbb{Z}_n$	zbiór klas reszt modulo n
$\mathbb{F}_q$ lub $GF(q)$	ciało skończone o q elementach
$n\mathbb{Z}$	zbiór liczb całkowitych podzielnych przez n
$\{a \mid \phi(a)\}$	zbiór składający się z elementów a spełniających warunek $\phi(a)$

$[a,b]$	przedział domknięty w zbiorze $\mathbb{R}$
$C_{[a,b]}$	zbiór (pierścień) funkcji rzeczywistych jednej zmiennej ciągłych na przedziale domkniętym $[a,b]$
$M_{m,n}(A)$	zbiór macierzy wymiaru $m \times n$ o współczynnikach z A
$M_n(A)$	zbiór (pierścień) macierzy (kwadratowych) stopnia n o współczynnikach z A
$\begin{pmatrix} 1 & 2 & \dots n \\ i_1 & i_2 & \dots i_n \end{pmatrix}$	permutacja
$(a_1, a_2, \dots, a_k)$	cykl długości k
$[a_1, a_2, \dots, a_k]$ lub $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix}$	wektor o n współrzędnych $a_1, a_2, \dots, a_k$
$L(A)$	zbiór przekształceń liniowych pierścienia A o niezerowym współczynniku wiodącym
$L_1(A)$	zbiór przekształceń liniowych pierścienia A o współczynniku wiodącym 1
$(a_n)_{n\in I},\ \{a_n\}_{n\in I}$	ciąg nieskończony o n-tym wyrazie a_n oraz dziedzinie I
$\varnothing$	zbiór pusty
$a \in A$	element a należy do zbioru A
$A = B$	zbiory A i B są równe
$A \subseteq B$ lub $B \supseteq A$	zbiór A zawiera się w zbiorze B
$A \subsetneq B$ lub $A \subset B$	A jest podzbiorem właściwym w zbiorze B
$A \cup B$	suma mnogościowa zbiorów A i B
$A \cap B$	przekrój (iloczyn mnogościowy) zbiorów A i B
$\bigcap \mathcal{A}$	przekrój rodziny $\mathcal{A}$

$A \setminus B$	różnica zbiorów A i B
$\mathcal{P}(A)$	zbiór wszystkich podzbiorów zbioru A (=zbiór potęgowy zbioru A)
$\|A\|$	moc zbioru A
B^A	zbiór wszystkich odwzorowań ze zbioru A w zbiór B
$\prod_{\lambda\in\Lambda} A_\lambda$	iloczyn kartezjański rodziny zbiorów A_λ
$A \times B$ lub $A_1 \times \cdots \times A_n$	skończony produkt kartezjański zbiorów
A^n	n-ta potęga kartezjańska zbioru A
$a \mid b$	a dzieli b
$\mathrm{NWD}(a, b)$	największy wspólny dzielnik a i b
$\mathrm{NWW}(a, b)$	najmniejsza wspólna wielokrotność a i b
(a, b)	para uporządkowana elementów a i b
$a \equiv b \pmod{n}$	a przystaje do b modulo n
p	liczba pierwsza
$\varphi(n)$	funkcja Eulera
$\overline{k} = [k]_n = k_n$	klasa reszt modulo $n \in \mathbb{N}^*$ z reprezentantem $k \in \mathbb{Z}$
$\|a\|$	wartość bezwzględna liczby rzeczywistej a
$\|z\|$	moduł liczby zespolonej z
$\|-\|$ lub $\|\|-\|\|$	wartość bezwzględna lub norma na ciele
$f : X \to Y$	funkcja (=odwzorowanie) f określona w zbiorze X o wartościach w zbiorze Y
$y = f(a)$	y jest wartością funkcji f w punkcie a
$g \circ f$	złożenie (superpozycja) funkcji f i g
$f\|_A$	zawężenie funkcji $f : X \to Y$ w dziedzinie do podzbioru $A \subseteq X$
id_X, id, e_X, 1_X, i_X	odwzorowanie identycznościowe (=tożsamościowe) na zbiorze X
$f^{-1}(B)$	przeciwobraz zbioru $B \subseteq Y$ względem funkcji $f : X \to Y$

$f^{-1}(y)$	przeciwobraz elementu $y \in Y$ względem funkcji $f : X \to Y$
$\operatorname{Im} f$	zbiór wartości funkcji f
$\operatorname{Ker} f$	jądro homomorfizmu f
π_α	α-ta projekcja
i_α	α-te zanurzenie (włożenie)
$(G, *)$	grupa względem działania $*$
$(G, \cdot)$	grupa multiplikatywna
$(G, +)$	grupa addytywna
$\mathbb{S}(X)$	grupa symetryczna zbioru X
$\mathbb{S}_n$	grupa symetryczna stopnia n
$\mathbb{C}_{p^\infty}$	kwazicykliczna p-grupa
$\operatorname{supp} \sigma$	nośnik permutacji σ
e_G, 0_R, e lub 1	element neutralny grupy G lub pierścienia R
0_G lub 0	element neutralny grupy addytywnej G lub zero (=element zerowy)
$\|G\|$	rząd grupy G
$\|g\|$ lub $o(g)$	rząd elementu g w grupie
g^{-1}	element odwrotny do elementu g w grupie
$H \leqslant G$	H jest podgrupą grupy G
$H < G$	H jest podgrupą właściwą grupy G
$H \triangleleft G$	H jest podgrupą normalną (=dzielnikiem normalnym) grupy G
G/H	grupa ilorazowa grupy G względem jej podgrupy normalnej H
gH, Hg	warstwa lewostronna, prawostronna w grupie względem podgrupy H elementu g
$\|G : H\|$	indeks podgrupy H w grupie G
$H + K$	suma podgrup H i K w grupie addytywnej
$H \oplus K$	suma prosta podgrup H i K w grupie addytywnej

$H \cdot K$	iloczyn podgrup H i K w grupie multiplikatywnej
$H \times K$	iloczyn prosty podgrup H i K w grupie multiplikatywnej
$G = H \rtimes K$	G jest iloczynem półprostym podgrup H oraz K, gdzie $H \triangleleft G$
$\langle g \rangle$, $\langle X \rangle$	grupa generowana przez element g, podzbiór X
$\mathbb{A}_n$	grupa alternująca stopnia n
Q_8	grupa kwaternionów
$\operatorname{Aut} G$	grupa automorfizmów grupy G
$\operatorname{Inn} G$	grupa automorfizmów wewnętrznych G
$\mathbb{C}_n$	grupa cykliczna pierwiastków zespolonych stopnia n z 1
$\mathbb{S}^1$	grupa liczb zespolonych o module 1
$H(\mathbb{F})$	grupa Heisenberga nad ciałem $\mathbb{F}$
G_x lub $\operatorname{St}(x)$	stabilizator elementu x
$G(x)$ lub $\operatorname{Orb}(x)$	orbita elementu x
$A^g = g \cdot A \cdot g^{-1}$	podgrupa sprzężona do podgrupy A wyznaczona przez element g
$Z(G)$	centrum grupy G
$C_G(x)$	centralizator elementu $x \in G$ w grupie G
$C_G(H)$	centralizator podgrupy $H \leqslant G$ w grupie G
$N_G(H)$	normalizator podgrupy $H \leqslant G$ w grupie G
$GL_n(\mathbb{F})$	ogólna (=pełna) grupa liniowa stopnia n nad ciałem $\mathbb{F}$
$SL_n(\mathbb{F})$	szczególna grupa liniowa stopnia n nad ciałem $\mathbb{F}$
$PGL_n(\mathbb{F})$	rzutowa (ogólna) grupa liniowa stopnia n nad ciałem $\mathbb{F}$
$PSL_n(\mathbb{F})$	rzutowa szczególna grupa liniowa stopnia n nad ciałem $\mathbb{F}$

$GL_n(q)$	ogólna (=pełna) grupa liniowa stopnia n nad ciałem $\mathbb{F}_q$
$SL_n(q)$	szczególna grupa liniowa stopnia n nad ciałem $\mathbb{F}_q$
$PGL_n(q)$	rzutowa (ogólna) grupa liniowa stopnia n nad ciałem $\mathbb{F}_q$
$PSL_n(q)$	rzutowa szczególna grupa liniowa stopnia n nad ciałem $\mathbb{F}_q$
$O(n, F)$ lub $O_n(F)$	n-ta ortogonalna grupa nad ciałem F
$O(n)$, $O_n(\mathbb{R})$ lub $O(n, \mathbb{R})$	n-ta grupa ortogonalna (nad ciałem liczb rzeczywistych $\mathbb{R}$)
$U(n)$ lub $U(n, \mathbb{C})$	n-ta grupa unitarna (nad ciałem liczb zespolonych $\mathbb{C}$)
$SO(n)$ lub $SO_n(\mathbb{R})$	n-ta grupa specjalna ortogonalna (nad ciałem liczb rzeczywistych $\mathbb{R}$)
$SU(n)$	n-ta specjalna grupa unitarna
$UT_n(F)$	grupa macierzy unitrójkątnych stopnia n nad ciałem F
$\det A$	wyznacznik macierzy A
I_n	macierz jednostkowa stopnia n
A^T	transpozycja macierzy A
$A[X]$	pierścień wielomianów zmiennej X o współczynnikach z A
$\mathbb{F}(X)$	ciało funkcji wymiernych jednej zmiennej X o współczynnikach z ciała $\mathbb{F}$
$A[[X]]$	pierścień formalnych szeregów potęgowych zmiennej X o współczynnikach z A
$F((X))$	ciało (formalnych) szeregów Laurenta zmiennej X o współczynnikach z F
$\mathbb{F}[X_1, \ldots, X_n]$	pierścień wielomianów zmiennych $X_1, \ldots, X_n$ o współczynnikach z $\mathbb{F}$

$\mathbb{F}(X_1, \ldots, X_n)$	ciało funkcji wymiernych zmiennych $X_1, \ldots, X_n$ o współczynnikach z $\mathbb{F}$
$F((X_1, \ldots, X_n))$	ciało ułamków pierścienia $F[[X_1, \ldots, X_n]]$
$\mathbb{Z}[\frac{1}{p}]$	podpierścień w $\mathbb{Q}$ generowany przez $\frac{1}{p}$
$\mathbb{Z}[i]$	pierścień liczb całkowitych Gaussa
$\operatorname{char} A$	charakterystyka pierścienia A
$\dim_F V$	wymiar przestrzeni V nad ciałem F
aA, Aa, $\langle a \rangle$	ideał główny prawostronny, lewostronny, obustronny pierścienia A z 1 generowany przez element $a \in A$
A/I	pierścień ilorazowy pierścienia A przez ideał obustronny I
$a + I$	warstwa ideału I z reprezentantem a
$I_1 \cdot I_2$	iloczyn algebraiczny zbiorów I_1 i I_2
$\pi : A \to A/I$	epimorfizm kanoniczny
$Z(A)$	centrum pierścienia A
A^* lub $U(A)$	grupa elementów odwracalnych w A
$\mathbf{s}_j$	wektor j-tego stanu ciągu $\mathbf{S}$
$p(\mathbf{S})$	okres ciągu $\mathbf{S}$
$\mathbb{F}(\alpha)$	ciało generowane przez ciało $\mathbb{F}$ i element α
$\mathbb{F}(\alpha_1, \ldots, \alpha_n)$	ciało generowane przez ciało $\mathbb{F}$ i elementy $\alpha_1, \ldots, \alpha_n$
m_α	wielomian minimalny elementu α (nad ciałem F)
m_i	wielomian minimalny elementu θ^i (nad ciałem skończonym)
$\lvert F : K \rvert$	stopień rozszerzenia $K \subseteq F$ ciał K i F
$\operatorname{ord} f$	rząd wielomianu f

Skorowidz

eISBN 978-83-67188-35-7

Politechnika Krakowska
im. Tadeusza Kościuszki