Konspekt do wykładu G.C. z Algebry 'C'

Część sem. I

(wersja robocza 1997/98)

Spis treści

1 Pojęcie ciała. Liczby zespolone

1.1 Krótki rys historyczny

1. Liczby zespolone zostały odkryte w XVI w. przy okazji zmagań z rozwiązywaniem równań algebraicznych trzeciego stopnia. Oto zarys tej ciekawej i burzliwej historii, która działa się we Włoszech:

   Pierwszym, który wymyślił metodę znajdowania pierwiastka równania 3. stopnia był zapewne Scipio del Ferro (1465..1526, Bolonia). Zdradził swój sekret na łożu śmierci swoim uczniom, którymi byli Hannibal della Nave i Antonio Mario Fior.

   A.M.Fior wygrał kilka turniejów matematycznych, przegrał dopiero z Tartaglią. Tartaglia ('jąkała', właściwe nazwisko Nicolo Fontana, 1500..57) był samoukiem, doszedł do stanowiska nauczyciela matematyki w szkole handlowej; wygrał także konkurs na najdalszy strzał — ustawiając armatę pod kątem 45°; został dzięki temu ekspertem artyleryjskim.

   Turniej Fontana – Tartaglia trwał aż 50 dni (24.12'1534 .. 12.2'1535); wszystkie zadania wymyślone przez Fiora były na równania 3. stopnia. Dnia 4 lutego 1535 (8 dni przed końcem terminu) Tartaglia odkrył metodę, dzięki której wygrał.

   Girolamo Cardano (1501..76) zaskarbił sobie zaufanie Tartaglii i w końcu wydobył odeń tajemnicę; opublikował metodę w ARS MAGNA (1545), dziele nt. algebry.

   Awantura i pretensje Tartaglii skończyły się turniejem (1548), na który Cardan nie raczył nawet osobiście przybyć, przysyłając w zamian swego ucznia, Lodovico Ferrariego, odkrywcę metody (też opisanej w ARS MAGNA, z nazwiskiem autora) rozwiązywania równań stopnia 4. Tartaglia tym razem przegrał turniej.

   Warto uświadomić sobie, że w owym czasie nie istniała jeszcze dzisiejsza notacja algebraiczna, a więc nie stosowano jeszcze np. znaków $+$, $-$, $\cdot$, $\sqrt{\ }$, itp. Co gorsza, każdy z uczestników tych wydarzeń używał własnych, zwykle bardzo nieprzystępnych oznaczeń lub wręcz opisywał słownie operacje arytmetyczne. Oto dla przykładu próbka notacji Bombellego z jego podręcznika ALGEBRA (1572):

   $\boxed{\text{0m.3a.R[0m.16]}}$ w notacji Bombellego dziś zapisujemy jako $\boxed{-3+4i}$;

   użyte tu litery a, m i R są skrótami łacińskich słów *addo* — dodać, *minno, meno* — odjąć, *radical* — pierwiastek; zatem R[0m.16] oznacza 'pierwiastek z −16'.

2. **Wzór Tartaglii, 1535**([1]): Równanie $x^3 = 3ax + 2b$ (współczynniki a, b są dane) ma pierwiastek

   $$x = \sqrt[3]{b + \sqrt{b^2 - a^3}} + \sqrt[3]{b - \sqrt{b^2 - a^3}}.$$

   Sprawdzenie: $x = u+v$, gdzie $\begin{cases} u^3 + v^3 = 2b, \\ uv = a, \end{cases}$ więc $x^3 = 3uvx + u^3 + v^3 = 3ax + 2b$.

3. Dla $x^3+3x+2 = 0$ dostajemy $x = \sqrt[3]{-1+\sqrt{2}} - \sqrt[3]{1+\sqrt{2}} \approx -0.596071$.
   Dla $x^3 - 3x + 2 = 0$ dostajemy: $x = \sqrt[3]{-1} + \sqrt[3]{-1} = -2$.
   Dla $x^3 - 3\sqrt[3]{2}x + 2 = 0$ (czyli $a = \sqrt[3]{2}$, $b = -1$) dostajemy:

[1] Zapewne równoważny wzór odkrył kilkanaście lat wcześniej Scipion del Ferro. Więcej o tych sprawach można znaleźć w HISTORII MATEMATYKI Marka Kordosa, WSiP'94.

$$x = \sqrt[3]{-1+\sqrt{-1}} + \sqrt{-1-\sqrt{-1}}.$$

Jeśli nie zrazimy się tym, że $\sqrt{-1}$ 'nie istnieje', to licząc formalnie otrzymamy równości $\begin{cases} (1+\sqrt{-1})^3 = \ldots = -2+2\sqrt{-1} \\ (1-\sqrt{-1})^3 = \ldots = -2-2\sqrt{-1} \end{cases}$, z których wynika $x = \frac{1+\sqrt{-1}}{\sqrt[3]{2}} + \frac{1-\sqrt{-1}}{\sqrt[3]{2}} = \frac{2}{\sqrt[3]{2}} = \sqrt[3]{4}$, co jest wynikiem poprawnym!

1.2 Pojęcie ciała. Ciało liczb zespolonych

4. **Definicja ciała**. *Ciałem* nazywa się zbiór $\mathbb{K}$ wyposażony w dwa działania 2-argumentowe '$+$' i '$\cdot$', spełniające następujące aksjomaty:

   1° *Łączność* $\begin{cases} \textit{dodawania:} & (a+b)+c = a+(b+c), \\ \textit{i mnożenia:} & (ab)c = a(bc). \end{cases}$

   2° *Przemienność dodawania i mnożenia:* $a+b = b+a$, $ab = ba$.

   3° *Istnienie elementów neutralnych:*
   $\begin{cases} \textit{dla dodawania:} & O \in \mathbb{K} \text{ ('zero'), } \textit{taki że } O+a = a \textit{ dla } a \in \mathbb{K}, \\ \textit{i dla mnożenia:} & I \in \mathbb{K} \text{ ('jedynka'), } \textit{taki że } Ia = a \textit{ dla } a \in \mathbb{K}. \end{cases}$

   4° *Istnienie elementów odwrotnych:*
   $\begin{cases} \textit{dla dodawania: dla } a \in \mathbb{K} \textit{ istnieje } b \in \mathbb{K}, \textit{ takie że } a+b = O, \\ \textit{i dla mnożenia: dla } a \in \mathbb{K} \setminus \{O\} \textit{ istnieje } b \in \mathbb{K}, \textit{ takie że } ab = I. \end{cases}$

   Oznaczenia: $\begin{cases} -a = b, \text{ jeśli } a+b = O \text{ (element przeciwny do } a), \\ a^{-1} = \frac{1}{a} = b, \text{ jeśli } ab = I \text{ (element odwrotny do } a). \end{cases}$

   5° *Rozdzielność mnożenia względem dodawania:* $a(b+c) = ab+ac$.

   6° *Nietrywialność:* $O \neq I$ (bez 6° byłoby $\mathbb{K} = \{O\}$).

5. **Przykłady ciał**. Ciała liczbowe (tzn. 'podciała' ciała $\mathbb{R}$): $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}(\sqrt{2})$.

   Ciała nieliczbowe (skończone!): $\mathrm{GF}_2 = \{O, I\}$ z działaniami

$+$	O	I
O	O	I
I	I	O

   ,

$\cdot$	O	I
O	O	O
I	O	I

   ; $\mathrm{GF}_4 = \{O, I, a, b\}$ z działaniami

$+$	O	I	a	b
O	O	I	a	b
I	I	O	b	a
a	a	b	O	I
b	b	a	I	O

   ,

$\cdot$	O	I	a	b
O	O	O	O	O
I	O	I	a	b
a	O	a	b	I
b	O	b	I	a

   Uwaga (dla znających pojęcie *macierzy*). GF_4, czyli tzw. 'ciało Galois', można zrealizować wychodząc z ciała $\mathrm{GF}_2 = \{O, I\}$ i biorąc cztery następujące *macierze wymiaru* 2×2: $\boldsymbol{O} := \begin{bmatrix} O & O \\ O & O \end{bmatrix}$, $\boldsymbol{I} := \begin{bmatrix} I & O \\ O & I \end{bmatrix}$, $\boldsymbol{a} := \begin{bmatrix} O & I \\ I & I \end{bmatrix}$, $\boldsymbol{b} := \begin{bmatrix} I & I \\ I & O \end{bmatrix} = \boldsymbol{a}^2$.

1.3 Definicja i własności liczb zespolonych

6. **Konstrukcja ciała** $\mathbb{C}$. W zbiorze $\mathbb{C} := \mathbb{R} \times \mathbb{R} = \{(a,b) : a, b \in \mathbb{R}\}$, którego elementy (tzn. pary liczb rzeczywistych) nazywać będziemy *liczbami zespolonymi*, wprowadźmy następujące działania:

dodawanie: $(a,b)+(c,d) := (a+c, b+d)$,
mnożenie: $(a,b)(c,d) := (ac-bd, ad+bc)$.

7. **Fakt**. Zbiór $\mathbb{C} = \mathbb{R}\times\mathbb{R}$ z powyższymi działaniami jest ciałem; w szczególności zerem jest $O = (0,0)$, jedynką $I = (1,0)$; elementem przeciwnym do (a,b) jest $(-a,-b)$, a odwrotnością — element
$$\Big(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\Big).$$

8. **Część rzeczywista i część urojona liczby zespolonej**. Dwa człony pary, jaką jest liczba zespolona $z\in\mathbb{C}$, nazywa się *częścią rzeczywistą* i *częścią urojoną* liczby z; oznacza się je symbolami $\operatorname{Re} z$ i $\operatorname{Im} z$ (są to skróty łacińskich słów *realis, imaginaris*):
$$z=(a,b) \;\Rightarrow\; \begin{cases} \operatorname{Re} z := a, \\ \operatorname{Im} z := b. \end{cases}$$
Wprost z definicji działań w $\mathbb{C}$ wynikają wzory
$$\begin{cases} \operatorname{Re}(z_1+z_2) = \operatorname{Re} z_1 + \operatorname{Re} z_2, \\ \operatorname{Im}(z_1+z_2) = \operatorname{Im} z_1 + \operatorname{Im} z_2, \end{cases} \quad \begin{cases} \operatorname{Re}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Re} z_2 - \operatorname{Im} z_1 \operatorname{Im} z_2, \\ \operatorname{Im}(z_1 z_2) = \operatorname{Re} z_1 \operatorname{Im} z_2 + \operatorname{Im} z_1 \operatorname{Re} z_2. \end{cases}$$

9. **Wygodne spostrzeżenia i oznaczenia**. Zauważmy, że:

   (a) podzbiór elementów $z\in\mathbb{C}$ postaci $(a,0)$, tzn. takich że $\operatorname{Im} z = 0$, jest zamknięty względem obu działań (czyli jest *podciałem*), przy czym
   $$(a,0)+(b,0) = (a+b,0),$$
   $$(a,0)(b,0) = (ab,0),$$
   a więc sensowne (nie prowadzące do kolizji) jest stosowanie 'stenografii', polegającej na pisaniu a zamiast $(a,0)$. W tym sensie będziemy odtąd traktować liczby rzeczywiste jako *liczby zespolone o zerowej części urojonej*, czyli traktować $\mathbb{R}$ jako podzbiór (a nawet pociało) $\{z : \operatorname{Im} z = 0\}$ ciała $\mathbb{C}$; w konsekwencji będziemy pisać 0 zamiast O, a 1 zamiast I;

   (b) jeśli oznaczymy $i := (0,1)$, to $i^2 = (0,1)(0,1) = (-1,0) = -1$, gdzie ostatnia równośc jest wynikiem konwencji (a);

   (c) każdą liczbę zespoloną $(a,b)\in\mathbb{C}$ można przedstawić w postaci $(a,b) = (a,0)+(0,1)(b,0)$, a więc — po zastosowaniu konwencji (a),
   $$(a,b) = a+ib, \quad a,b\in\mathbb{R},$$
   co można także zapisać w postaci
   $$z = \operatorname{Re} z + i \operatorname{Im} z.$$

10. Te spostrzeżenia i poręczna notacja uwalniają nas od stosowania skomplikowanych wzorów definiujących działania na liczbach zespolonych; zamiast nich w rachunkach na liczbach zespolonych $a+ib$ wystarczy pamiętać o własnościach działań (przemienność, łączność, rozdzielność, własności 0 i 1) oraz o regule $i^2=-1$, na przykład
$$(-3+12i)+(2-i)(3+4i) = -3+12i+6+8i-3i-4i^2 = 7+170i,$$
$$\frac{2+3i}{3-4i} = \frac{(2+3i)(3+4i)}{(3-4i)(3+4i)} = \frac{6+8i+9i+12i^2}{3^2+4^2} = \ldots = -\frac{6}{25}+i\frac{17}{25}.$$

1.4 Operacje sprzężenia i modułu

11. **Operacja sprzężenia**. *Sprzężeniem* liczby $z = (a, b) \in \mathbb{C}$ nazywa się liczbę $\overline{z} := (a, -b)$; w poręczniejszej notacji:

$$\overline{a + ib} := a - ib, \quad \text{jeśli } a, b \in \mathbb{R},$$

co można też zapisać w postaci

$$\overline{z} = \operatorname{Re} z - i \operatorname{Im} z.$$

12. **Fakt**. Operacja sprzężenia $\mathbb{C} \to \mathbb{C}$ ma następujące własności:

1° $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2$ (*addytywność*); 2° $\overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$ (*multiplikatywność*);
3° $\overline{z_1/z_2} = \overline{z}_1/\overline{z}_2$; 4° $\overline{z^n} = (\overline{z})^n$ $(n \in \mathbb{Z})$; 5° $\overline{\overline{z}} = z$ (*inwolutywność*);
6° $\overline{z}z = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2$, więc $\overline{z}z \in \mathbb{R}_+$; 7° $\overline{z} = z \iff z \in \mathbb{R}$;
8° $\overline{z} = -z \iff z \in i\mathbb{R} := \{iy : y \in \mathbb{R}\}$ (zbiór tzw. *liczb urojonych*).

Ad 2°: $z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2)$, więc $\overline{z_1 z_2} = (x_1 x_2 - y_1 y_2) - i(x_1 y_2 + y_1 x_2)$, zaś $\overline{z}_1 \overline{z}_2 = \ldots = (x_1 x_2 - y_1 y_2) - i(x_1 y_2 + y_1 x_2)$. *Ad* 3° $z := \frac{z_1}{z_2}$, wtedy $z_1 = z z_2$ zgodnie z 2° daje $\overline{z}_1 = \overline{z}\,\overline{z}_2$, czyli $\overline{z} = \frac{\overline{z}_1}{\overline{z}_2}$. *Ad* 4°: Dla $n \in \mathbb{N}$ z 2° przez indukcję dostajemy $\overline{z_1 \ldots z_n} = \overline{z}_1 \ldots \overline{z}_n$, w szczególności $\overline{z^n} = \overline{z}^n$; dla $n < 0$ mamy $\overline{z^n} = \overline{1/z^{-n}} \overset{3^\circ}{=} \frac{1}{\overline{z}^{-n}} = \overline{z}^{-n}$. *Ad* 6 : $(x + iy)(x - iy) = x^2 + y^2 \geqslant 0$.

13. Skoro $\begin{cases} z = \operatorname{Re} z + i \operatorname{Im} z, \\ \overline{z} = \operatorname{Re} z - i \operatorname{Im} z, \end{cases}$ to mamy też ważne wzory $\begin{cases} \operatorname{Re} z = \frac{1}{2}(z + \overline{z}), \\ \operatorname{Im} z = \frac{1}{2i}(z - \overline{z}). \end{cases}$

14. **Fakt**. *Moduł* $|z| := \sqrt{z\overline{z}}$ jest funkcją $\mathbb{C} \to \mathbb{R}_+ := \{r \in \mathbb{R} : r \geqslant 0\}$, mającą następujące własności:

$$1^\circ \ |z| = 0 \iff z = 0, \qquad 2^\circ \ |z_1 z_2| = |z_1| \cdot |z_2|,$$
$$3^\circ \ \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \qquad 4^\circ \ |z_1 + z_2|^2 = |z_1|^2 + 2\operatorname{Re}(z_1 \overline{z}_2) + |z_2|^2,$$
$$5^\circ \ |z_1 + z_2| \leqslant |z_1| + |z_2|, \qquad 6^\circ \ \big||z_1| - |z_2|\big| \leqslant |z_1 - z_2|.$$

2° i 5° mają oczywiste uogólnienia: $|z_1 z_2 \ldots z_n| = |z_1| \cdot |z_2| \cdot \ldots \cdot |z_n|$, $|z_1 + z_2 + \ldots z_n| \leqslant |z_1| + |z_2| + \ldots + |z_n|$. Ostatnia nierówność staje się równością $\iff$ $\exists z \in \mathbb{C} : \exists r_1, \ldots r_n \in \mathbb{R}_+ : z_1 = r_1 z, \ldots, z_n = r_n z$, tzn. gdy z_i leżą na półprostej przechodzącej przez $0 \in \mathbb{C}$ (ćwiczenie).

Zauważmy przede wszystkim, że definicja modułu ma sens, gdyż dla każdego $z = (a, b) = a + ib \in \mathbb{C}$ wielkość $z\overline{z} = (a + ib)(a - ib) = a^2 + b^2$ należy do $\mathbb{R}_+$.

Ad 2°: Jeśli $z = z_1 z_2$, to z własności sprzężenia $z\overline{z} = z_1\overline{z}_1 \cdot z_2\overline{z}_2$; przy tym liczby $z\overline{z}$, $z_1\overline{z}_1$ i $z_2\overline{z}_2$ należą do $\mathbb{R}_+$, więc można obustronnie spierwiastkować tę równość.

Ad 3°: Niech $z = z_1/z_2$, wtedy $z_1 = z z_2$, więc $|z_1| = |z| \cdot |z_2|$ na mocy 2°; stąd teza.
Ad 4°: $L = (z_1 + z_2)\overline{(z_1 + z_2)} = (z_1 + z_2)(\overline{z}_1 + \overline{z}_2) = z_1\overline{z}_1 + z_1\overline{z}_2 + z_2\overline{z}_1 + z_2\overline{z}_2 = |z_1|^2 + w + \overline{w} + |z_2|^2 = P$, gdzie $w = z_1\overline{z}_2$.

Ad 5°: Korzystamy z $\operatorname{Re} z \leqslant |z|$ oraz z 3° (lub z 2°) i $\operatorname{Re} \frac{z_1}{z_2 + z_3} + \operatorname{Re} \frac{z_2}{z_1 + z_3} = 1$.
Ad 6°: Dla $a, b \in \mathbb{R}$: $|b| \leqslant a \iff -a \leqslant b \leqslant a$, więc 5° oznacza nierówności $-|z_1 - z_2| \leqslant |z_1| - |z_2| \leqslant |z_1 - z_2|$; dostajemy je z 4°, biorąc sumy $\left\{ \begin{matrix} z_1 + (z_2 - z_1) \\ (z_1 - z_2) + z_2 \end{matrix} \right\}$.

1.5 O niemożności uporządkowania

15. *Ciało uporzadkowane*: Jest to takie ciało $\mathbb{K}$, w którym jest zadana relacja "$<$", zwana *uporządkowaniem*, spełniająca następujące aksjomaty:

    1° $\forall a,b\in\mathbb{K}:\ a<b$ albo $a>b$ (tzn. $b<a$) albo $a=b$ (*trichotomia*);
    2° $\forall a,b,c\in\mathbb{K}:\ (a<b,\ b<c)\Rightarrow a<c$ (*przechodniość*);
    3° $\forall a,b,c\in\mathbb{K}:\ a<b\Rightarrow a+c<b+c$ (*monotoniczność dodawania*);
    4° $\forall a,b,c\in\mathbb{K}:\ (a<b,\ c>0)\Rightarrow ac<bc$ (..... *mnożenia*).

    Odnotujmy, że z 3° wynika (dla $b:=0, c:=-a$) ważna własność $a<0\ \Rightarrow\ -a>0$.

16. **Fakt**. Ciała liczb zespolonych nie da się uporządkować, tzn. w zbiorze $\mathbb{C}$ nie istnieje relacja "$<$" o własnościach 1°...4°.

    Zauważmy, że z 4° wynika, że $\forall b,c\in\mathbb{K}:\ (b>0,\ c>0)\Rightarrow bc>0$, a więc też $(b<0,\ c<0)\Rightarrow bc=(-b)(-c)>0$; stąd i z 1° dostajemy $\forall a\in\mathbb{K}^*:\ a^2>0$; w szczególności $1=1^2>0$, a więc $-1<0$. Teraz już łatwo dowieść nie wprost naszą tezę: Dla $\mathbb{K}=\mathbb{C}$ oprócz $-1<0$ byłoby też $-1=i^2>0$, sprzeczność z 1°.

17. **Uwaga**. Dla $z\in\mathbb{C}$ zapis '$z>0$' zawsze będzie oznaczać, że z jest *rzeczywistą* liczbą dodatnią, tzn. $\left\{\begin{array}{l} z\in\mathbb{R}\\ z>0\end{array}\right\}$; innymi słowy $z=(a,b)>0\overset{\text{df}}{\iff}\left\{\begin{array}{l} a>0\\ b=0\end{array}\right\}$.
    Jeśli teraz określimy w zbiorze $\mathbb{C}$ relację '$<$' wzorem $z_1<z_2\overset{\text{df}}{\iff}z_2-z_1>0$, tzn. $\left\{\begin{array}{l}\operatorname{Re}z_1<\operatorname{Re}z_2\\ \operatorname{Im}z_1=\operatorname{Im}z_2\end{array}\right\}$, to oczywiście spełnione będą warunki $2^\circ,3^\circ,4^\circ$, ale nie 1°.

18. **Ćwiczenie**. Określmy dla liczb zespolonych inną 'namiastkę uporządkowania': $z_1\prec z_2\ \overset{\text{df}}{\iff}\left\{\begin{array}{l}\operatorname{Re}z_1<\operatorname{Re}z_2\\ \operatorname{Im}z_1<\operatorname{Im}z_2\end{array}\right\}$. Sprawdzić, że ma ona 3 spośród własności $1^\circ\ldots4^\circ$.

1.6 Postać biegunowa i trygonometryczna

19. **Fakt** (rozkład biegunowy). Każda liczba $z\in\mathbb{C}$ różna od zera ma jednoznaczny rozkład, zwany *postacią biegunową liczby* z, postaci

$$z=ru,\ \text{gdzie}\ \ r>0,\ \ \text{zaś}\ \ u\in U:=\{u\in\mathbb{C}:\ |u|=1\}. \qquad (\mathrm{R})$$

    Istotnie, dla $r:=|z|$ oraz $u:=\frac{z}{r}$ warunki (R) są spełnione, co dowodzi *istnienia*. *Jednoznaczność*: Z (R) wynika, że $|z|=|r|\cdot|u|=|r|=r$, więc $\left\{\begin{array}{l} r=|z|\\ u=\frac{z}{|z|}\end{array}\right\}$, Q.E.D.

20. **Oznaczenie** (*symbol Eulera*). $\boxed{e^{i\varphi}:=\cos\varphi+i\sin\varphi\ \text{dla}\ \varphi\in\mathbb{R}}$.

    Uwaga. Nie należy traktować $e^{i\varphi}$ jako "potęgi liczby e o urojonym wykładniku". Liczbę $\cos\varphi+i\sin\varphi\in\mathbb{C}$, zależącą od $\varphi\in\mathbb{R}$, moglibyśmy tu oznaczać jakimś 'neutralnym' symbolem, np. $u(\varphi)$ lub u_φ lub tp. Otóż zobaczymy zaraz, że funkcja $\mathbb{R}\ni\varphi\mapsto\cos\varphi+i\sin\varphi\in\mathbb{C}$ ma **własności** podobne do funkcji wykładniczej $\varphi\mapsto a^\varphi$, a więc symbol $e^{i\varphi}$ okaże się wygodny dzięki swojej **sugestywności**. Głębsze uzasadnienie symbolu $e^{i\varphi}$ przyniesie wykład z analizy, na którym wzór $e^{i\varphi}=\cos\varphi+i\sin\varphi$ pojawi się ponownie, lecz już jako **twierdzenie**, a nie definicja. Dla niecierpliwych: Na analizie definicją e^z będzie wzór $e^z:=\lim\limits_{n\to\infty}\sum\limits_{k=0}^{n}\frac{1}{k!}z^k$, $z\in\mathbb{C}$.

21. **Fakt** (*własności symbolu Eulera*).

(1) $|e^{i\varphi}| = 1$, tzn. $e^{i\varphi} \in U$;
(1') co więcej, zbiór $\{e^{i\varphi} : \varphi \in \mathbb{R}\}$ jest całym okręgiem jednostkowym U;
(2) $(e^{i\varphi})^{-1} = \overline{e^{i\varphi}} = e^{-i\varphi}$;
(3) $e^{i\varphi_1} \cdot e^{i\varphi_2} = e^{i(\varphi_1+\varphi_2)}$;
(4) $e^{i\varphi} = 1 \iff \varphi \in 2\pi\mathbb{Z}$, tzn. $\exists n \in \mathbb{Z} :\ \varphi = 2\pi n$, ogólniej:
(4') $e^{i\varphi_1} = e^{i\varphi_2} \iff \varphi_2 - \varphi_1 \in 2\pi\mathbb{Z}$, tzn. $\exists n \in \mathbb{Z} :\ \varphi_2 - \varphi_1 = 2\pi n$.

Sprawdzenie (1), (2), (3) jest prostym ćwiczeniem na własności modułu, sprzężenia oraz funkcji cos i sin. Własność (4') wynika z (4), gdyż $e^{i\varphi_2}\left[e^{i\varphi_1}\right]^{-1} = e^{i(\varphi_2-\varphi_1)}$ wskutek (2) i (3). Natomiast własności (1') i (4) wypowiadają następujący

22. **Fakt** (*znany ze szkoły*). Jeśli $c, s \in \mathbb{R}$ spełniają warunek $c^2 + s^2 = 1$, to $\exists\varphi \in \mathbb{R} : \begin{cases} c = \cos\varphi \\ s = \sin\varphi \end{cases}$. Ponadto $\begin{cases} \cos\psi = \cos\varphi \\ \sin\psi = \sin\varphi \end{cases} \Leftrightarrow \begin{cases} \exists n \in \mathbb{Z} : \\ \psi = \varphi + 2n\pi \end{cases}$.

23. **Wniosek** (*własności okręgu jednostkowego* $U = \{u \in \mathbb{C} :\ |u| = 1\}$).

(1) $u_1, u_2 \in U \Rightarrow u_1u_2 \in U$ (zamkniętość względem mnożenia);
(2) $u \in U \Rightarrow u^{-1} = \overline{u} \in U$ (..... odwrotności i sprzężenia);
(3) $U = \{e^{i\varphi} :\ \varphi \in \mathbb{R}\}$ (parametryzacja).

24. **Uwaga**. Zauważmy, że jeśli $u = e^{i\varphi}$, to $\begin{cases} \cos\varphi = \operatorname{Re} u = \frac{1}{2}(u + \overline{u}), \\ \sin\varphi = \operatorname{Im} u = \frac{1}{2i}(u - \overline{u}), \end{cases}$ więc wzory $\begin{cases} \operatorname{Re}(u_1u_2) = (\operatorname{Re} u_1)(\operatorname{Re} u_2) - (\operatorname{Im} u_1)(\operatorname{Im} u_2), \\ \operatorname{Im}(u_1u_2) = (\operatorname{Re} u_1)(\operatorname{Im} u_2) + (\operatorname{Im} u_1)(\operatorname{Re} u_2) \end{cases}$ (zobacz określenie mnożenia w $\mathbb{C}$) prowadzą wprost do tożsamości $\begin{cases} \cos(\varphi_1 + \varphi_2) = \cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2, \\ \sin(\varphi_1 + \varphi_2) = \sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2. \end{cases}$
Dzięki temu znając własności symbolu $e^{i\varphi}$ (oraz własności operacji na liczbach zespolonych) możemy łatwo wyprowadzić każdą tożsamość trygonometryczną. Np.

$$\cos^3\varphi = \left(\frac{u+\overline{u}}{2}\right)^3 = \frac{u^3 + 3u + 3\overline{u} + \overline{u}^3}{8} = \frac{1}{4}\cos 3\varphi + \frac{3}{4}\cos\varphi,$$

$$\sin^2\varphi\cos\psi = \left(\frac{u-\overline{u}}{2i}\right)^2 \frac{v+\overline{v}}{2} = -\frac{1}{8}(u^2 - 2 + \overline{u}^2)(v + \overline{v}) =$$
$$= \frac{v+\overline{v}}{4} - \frac{u^2v + \overline{u}^2\overline{v}}{8} - \frac{u^2\overline{v} + \overline{u}^2v}{8} = \frac{1}{2}\cos\psi - \frac{1}{4}\cos(2\varphi + \psi) - \frac{1}{4}\cos(2\varphi - \psi).$$

25. **Wniosek** (*postać trygonometryczna*). Każdą liczbę $0 \neq z \in \mathbb{C}$ można przedstawić w postaci $z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$, gdzie $\varphi \in \mathbb{R}$, $r > 0$.

Oczywiście $r = |z|$; liczbę φ nazywa się *argumentem* z i oznacza symbolem $\operatorname{Arg} z$; jest ona określona z dokładnością do krotności 2π. Liczbę $\varphi \in [0, 2\pi[$ taką, że $z = |z|e^{i\varphi}$, tzn. wartośc $\operatorname{Arg} z$ należącą do $[0, 2\pi[$, nazywa się *argumentem głównym* liczby z i oznacza symbolem $\arg z$.

Powyższa tradycyjna definicja $\operatorname{Arg} z$ ma, jak teraz zobaczymy, pewne wady, więc potraktujemy ją jako tymczasową i za chwilę zastąpimy lepszą formalnie definicją.

26. Gdzie tkwi błąd w wywodzie $\left.\begin{array}{l} -i = e^{\frac{3}{2}\pi i} \Rightarrow \operatorname{Arg}(-i) = \frac{3\pi}{2} \\ -i = e^{-\frac{1}{2}i} \Rightarrow \operatorname{Arg}(-i) = -\frac{\pi}{2} \end{array}\right\} \Rightarrow\ \frac{3\pi}{2} = -\frac{\pi}{2}$?

W tym, że kierując się naszymi nawykami (związanymi z symbolami postaci $\sin x$, $\operatorname{Re} z$ itd.) potraktowaliśmy $\operatorname{Arg} z$ jako wartość pewnej funkcji, przyporządkowującej liczbie z pewien kąt $\operatorname{Arg} z \in \mathbb{R}$. Tymczasem tak nie jest: zapis $\operatorname{Arg} z = \varphi$ oznacza $z = |z|e^{i\varphi}$, a ta ostatnia konstatacja wcale nie określa jednoznacznie wartości φ.

Są dwa (w gruncie rzeczy równoważne) sposoby nadania sensu zdaniu $\varphi = \operatorname{Arg} z$:

(1) Przyporządkowanie $\mathbb{C}^* := \mathbb{C} \setminus \{0\} \ni z \mapsto \operatorname{Arg} z \in \mathbb{R}$ jest *funkcją wieloznaczną*; taka "funkcja" nie jest prawdziwą funkcją, bo wyrażenie $\operatorname{Arg} z$ jest wieloznaczne, określa nie liczbę, ale całą klasę liczb (kątów równoważnych w tym sensie, że wszystkim odpowiada ten sam punkt $e^{i\varphi}$ na "okręgu trygonometrycznym" U);
(2) $\operatorname{Arg} z$ nie jest liczbą, lecz **zbiorem liczb**: $\operatorname{Arg} z := \{\varphi \in \mathbb{R} :\ z = |z|e^{i\varphi}\}$, zaś zapis $\varphi = \operatorname{Arg} z$ jest ('nonszalancką', 'niefrasobliwą', 'tradycyjną' albo 'staroświecką') formą zapisu zdania $\varphi \in \operatorname{Arg} z$. W takim samym sensie zapis $a = \pm 1$ oznacza $a \in \{1, -1\}$, zapis $\sqrt{-1} = \pm i$ oznacza $\sqrt{-1} = \{i, -i\}$, a zapis $z = \sqrt{-1}$ oznacza tyle, co $z = \pm i$, czyli (już w precyzyjnej postaci) $z \in \{i, -i\}$.

27. **Fakt**. Dla $z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ określmy $\boxed{\operatorname{Arg} z := \{\varphi \in \mathbb{R} :\ z = |z|e^{i\varphi}\}}$. Wtedy:

(1) Zbiór $\operatorname{Arg} z$ jest jedną z klas następującej relacji równoważności

$$\varphi_1 \cong \varphi_2 \overset{\text{def}}{\iff} \varphi_2 - \varphi_1 \in 2\pi\mathbb{Z}$$

w zbiorze $\mathbb{R}$; zatem Arg jest odwzorowaniem $\operatorname{Arg} : \mathbb{C}^* \to \mathbb{R}/2\pi\mathbb{Z}$, gdzie $\mathbb{R}/2\pi\mathbb{Z} = \mathbb{R}/\cong$ oznacza zbiór wszystkich klas relacji $\cong$.

(2) $\operatorname{Arg}(z_1 \cdot z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2$, gdzie po prawej stronie mamy tzw. *sumę algebraiczną podzbiorów* ciała $\mathbb{C}$:

$$A + B := \{a + b :\ a \in A,\ b \in B\}.$$

(3) $\operatorname{Arg}(z^{-1}) = \operatorname{Arg}\overline{z} = -\operatorname{Arg} z$, $\operatorname{Arg}\frac{z_1}{z_2} = \operatorname{Arg} z_1 - \operatorname{Arg} z_2$.

(4) Jeśli $z_1, z_2 \in \mathbb{C}^*$, to $z_1 = z_2 \iff \left\{\begin{array}{c} |z_1| = |z_2| \\ \operatorname{Arg} z_1 = \operatorname{Arg} z_2 \end{array}\right\}$.

Ad (1) Wynika to wprost stąd, że $e^{i\varphi_1} = e^{i\varphi_2} \Leftrightarrow \exists n : \varphi_2 - \varphi_1 = 2\pi n$, tzn. $\varphi_1 \cong \varphi_2$. Niech $[\varphi] = \varphi + 2\pi\mathbb{Z}$ oznacza $\cong$-klasę elementu $\varphi \in \mathbb{R}$; wtedy, wprost z definicji, $[\varphi_1+\varphi_2] = [\varphi_1]+[\varphi_2]$; otóż $r_1e^{i\varphi_1}\cdot r_2e^{i\varphi_2} = (r_1r_2)e^{i(\varphi_1+\varphi_2)}$ oraz $\operatorname{Arg}(re^{i\varphi}) = [\varphi]$ (dla $r, \varphi \in \mathbb{R}, r > 0$); stąd teza. *Ad* (3) Jeśli $z = re^{i\varphi}$, to $z^{-1} = \frac{1}{r}e^{-i\varphi}$ oraz $\overline{z} = re^{-i\varphi}$, skąd teza. *Ad* (4). Jeśli $\varphi \in \operatorname{Arg} z$, to $z = |z|e^{i\varphi}$, skąd wynika '$\Leftarrow$'.

28. Argument głowny $\arg z \in \operatorname{Arg} z$ jest wyróżniony arbitralnym warunkiem należenia do $[0, 2\pi[$; w odróżnieniu od Arg nie jest on addytywny: różnica $d := \arg z_1 + \arg z_2 - \arg(z_1z_2)$ jest równa albo 0, albo 2π. Gdybyśmy zamiast $[0, 2\pi[$ wzięli przedział $]-\pi, \pi]$, byłoby jeszcze gorzej: $d \in \{-2\pi, 0, 2\pi\}$. Czy jakiś inny sposób wyboru reprezentanta mógłby zapewnić addytywność? Okazuje się, że nie, mianowicie:

29. **Fakt**. Nie istnieje funkcja $\phi : \mathbb{C}^* \to \mathbb{R}$, spełniająca warunki

$$\left(\begin{array}{c} \phi(z) \in \operatorname{Arg} z \\ \text{tzn. } z = |z|e^{i\phi(z)} \end{array}\right) \quad \text{oraz} \quad \forall z_1, z_2 \in \mathbb{C} :\ \phi(z_1z_2) = \phi(z_1) + \phi(z_2).$$

$\phi(1) = \phi(1)+\phi(1)$ (gdyż $1 = 1\cdot 1$) oraz $\phi(1) = \phi(-1)+\phi(-1)$ (gdyż $1 = (-1)(-1)$); stąd $\phi(1) = 0$, $\phi(-1) = 0$, co jest sprzeczne z warunkiem $-1 = e^{i\phi(-1)}$.

30. **Wzór de Moivre'a**. Jeśli $n \in \mathbb{Z}$, $r \geqslant 0$ oraz $\varphi \in \mathbb{R}$, to

$$\Big(r(\cos\varphi + i\sin\varphi)\Big)^n = r^n(\cos n\varphi + i\sin n\varphi).$$

Wynika on wprost z własności $\left(e^{i\varphi}\right)^n = e^{in\varphi}$ symbolu Eulera i ma raczej historyczne niż praktyczne znaczenie: Stosując symbol «$e^{i\varphi}$» możemy tego wzoru nie używać.

1.7 Pierwiastkowanie liczb zespolonych

31. **Definicja**. Dla $z \in \mathbb{C}$ i $n \in \mathbb{N}$ określmy $\boxed{\sqrt[n]{z} := \{w \in \mathbb{C} : \ w^n = z\}}$. Elementy zbioru $\sqrt[n]{z}$ nazywa się *pierwiastkami stopnia n z liczby z*; często spotykany tradycyjny zapis '$w = \sqrt[n]{z}$' oznacza '$w \in \sqrt[n]{z}$', więc np. wzór $z = \frac{-b\pm\sqrt{\Delta}}{2a}$ oznacza to samo, co '$z \in \left\{\frac{-b+w}{2a} : \ w \in \sqrt{\Delta}\right\}$'.

32. **Przykład**. $w \in \sqrt[3]{1} \Leftrightarrow w^3 = 1 \Leftrightarrow 0 = w^3 - 1 = (w-1)(w^2+w+1)$; ponieważ pierwiastkami równania $w^2 + w + 1 = 0$ są $w = \frac{-1 \pm i\sqrt{3}}{2}$, to $\boxed{\sqrt[3]{1} = \{1, \omega, \overline{\omega}\}}$, gdzie liczba $\boxed{\omega := \frac{-1+i\sqrt{3}}{2} = e^{i\frac{2\pi}{3}}}$ ma własności

$$\boxed{\omega^2 + \omega + 1 = 0, \ \ \omega^3 = 1, \ \ \overline{\omega}\omega = 1, \ \ \omega^2 = \omega^{-1} = \overline{\omega} \ .}$$

33. **Fakt**. $\sqrt[n]{re^{i\varphi}} = \left\{ r^{\frac{1}{n}} e^{i\frac{\varphi+2k\pi}{n}} : \ k \in \overline{0, n-1} \right\}$ dla $r > 0$, $\varphi \in \mathbb{R}$.

Wniosek (*Interpretacja geometryczna*). $\sqrt[n]{z}$ jest zbiorem wierzchołków pewnego n-kąta foremnego wpisanego w okrąg $C(0; r^{\frac{1}{n}})$, $r := |z|$.

Istotnie, zapiszmy w w postaci trygonometrycznej: $w = \varrho e^{i\alpha}$, wtedy $w^n = \varrho^n e^{in\alpha}$, więc $w \in \sqrt[n]{z} \iff w^n = z \iff \varrho^n = r$ (tzn. $\varrho = r^{\frac{1}{n}}$) oraz $n\alpha \cong \varphi$, tzn. $n\alpha = \varphi + 2k\pi$, $k \in \mathbb{Z}$. Przy tym wielkość $e^{i\frac{\varphi+2k\pi}{n}}$ nie zmienia się, gdy do k dodamy krotność n, więc wystarczy się ograniczyć do $k \in \overline{0, n-1}$.

34. **Przykład**. Liczba -4 ma postać trygonometryczną $4e^{i\pi}$, a więc
$\sqrt[4]{-4} = \sqrt[4]{4e^{i\pi}} = \left\{\sqrt{2}e^{i\frac{\pi}{4}}, \sqrt{2}e^{i\frac{3\pi}{4}}, \sqrt{2}e^{i\frac{5\pi}{4}}, \sqrt{2}e^{i\frac{7\pi}{4}}\right\} = \{1+i, -1+i, -1-i, 1-i\}$.

35. **Fakt**. $\sqrt[n]{z_1 z_2} = \sqrt[n]{z_1}\sqrt[n]{z_2}$ dla $z_1, z_2 \in \mathbb{C}$,
gdzie po prawej stronie występuje 'iloczyn algebraiczny' podzbiorów $\mathbb{C}$.

Możemy założyć, że $z_1 \neq 0$. Jeśli $w \in \sqrt[n]{z_1 z_2}$ to dla dowolnego $w_1 \in \sqrt[n]{z_1}$ biorąc $w_2 := \frac{w}{w_1}$ mamy ${w_2}^n = \frac{w^n}{{w_1}^n} = \frac{z_1 z_2}{z_1} = z_2$, a zatem $w = w_1 w_2 \in \sqrt[n]{z_1}\sqrt[n]{z_2}$. Odwrotnie, gdy $w_k \in \sqrt[n]{z_k}$, wtedy $(w_1 w_2)^n = {w_1}^n {w_2}^n = z_1 z_2$, czyli $w_1 w_2 \in \sqrt[n]{z_1 z_2}$.

36. **Fakt**. Niech $n \in \mathbb{N}$, $n \geqslant 2$. Nie istnieje funkcja $p_n : \mathbb{C} \to \mathbb{C}$, taka że

$$p_n(z) \in \sqrt[n]{z} \ \text{ oraz } \ p_n(z_1 z_2) = p_n(z_1) p_n(z_2).$$

$p_n(z^n) = \left(p_n(z)\right)^n = z$, więc $1 = p_n(1) = p_n(\epsilon^n) = \epsilon$ dla $z = \epsilon = e^{\frac{2\pi i}{n}}$, sprzeczność.

1.8 Interpretacje geometryczne niektórych rzeczy

37. **Ćwiczenie**. Znaleźć interpretację geometryczną:

(a) Dodawania $z = z_1 + z_2$ liczb zespolonych ('reguła równoległoboku').
(b) Części rzeczywistej i urojonej oraz modułu i sprzężenia liczby z.

(c) Mnożenia $z = z_1 z_2$ (zastosować wzór $r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = (r_1 r_2) e^{i(\varphi_1+\varphi_2)}$; odnotujmy, że dla znalezienia długości $r_1 r_2$ musimy mieć oprócz punktów $0, r_1, r_2$ osi rzeczywistej także punkt $1 \in \mathbb{C}$, wtedy przedział $[0,1]$ zadaje jednostkę długości.

(d) Odwzorowania $\mathbb{C} \ni z \mapsto F(z) := az + b$, jeśli $a \in \mathbb{C}^*$, $b \in \mathbb{C}$ są dane *Odpowiedź.* F jest złożeniem $T_b \circ \mathcal{H}_k \circ \mathcal{O}_\varphi$ obrotu o kąt $\varphi := \arg a$, jednokładności (tj. homotetii) o skali $k := |a|$ oraz translacji o b.

(e) Inwersji, czyli odwzorowania $\mathbb{C}^* \ni z \mapsto \mathcal{J}(z) := z^{-1} \in \mathbb{C}^*$. *Odpowiedź.* $\mathcal{J}$ jest złożeniem (w dowolnej kolejności) operacji sprzężenia i operacji $z \mapsto w$, gdzie w jest punktem półosi $z\mathbb{R}_+$ (wychodzącej z punktu 0 i przechodzącej przez z) odległym od 0 o $1/r$, gdzie $r = |z|$.

1.9 Równanie trzeciego stopnia; metoda Cardana

38. **Problem**. Znaleźć parę liczb, mających daną sumę S i dany iloczyn I; innymi słowy: rozwiązać względem w_1, w_2 układ równań $\begin{cases} w_1 + w_2 = S, \\ w_1 w_2 = I. \end{cases}$

Rozwiązanie. Skorzystajmy ze szkolnych wzorów Viete'a: pierwiastki w_1, w_2 równania $\boxed{w^2 - Sw + I = 0}$ mają potrzebne nam własności!

39. **Redukcja równania**. Równanie $x^3+ax^2+bx+c = 0$ ($a, b, c \in \mathbb{C}$ -dane) można zawsze 'zredukować', podstawiając $x = z + C$ i dobierając stałą $C \in \mathbb{C}$ tak, by w otrzymanym równaniu na z nie wystąpił kwadrat z.

40. **Metoda Cardana**. Zajmijmy się równaniem zredukowanym postaci

$$\begin{gathered} z^3 + pz + q = 0 \\ (p, q \in \mathbb{C} \text{ — dane}, p \neq 0). \end{gathered} \tag{R}$$

Przedstawmy szukany pierwiastek (R) w postaci sumy $z = u+v$; wtedy $z^3 = u^3 + 3u^2 v + 3uv^2 + v^3 = 3uvz + (u^3 + v^3)$. Prawa strona będzie równa $-pz - q$, jeśli parę (u, v) dobierzemy tak, by

$$\begin{cases} u^3 + v^3 = -q \\ 3uv = -p \end{cases}; \tag{U}$$

zatem spełnienie tego układu jest warunkiem *wystarczającym* (ale wcale nie koniecznym!), na to, by liczba $z = u + v$ spełniała równanie (R). Konsekwencją warunków (U) jest układ typu 'dane są suma i iloczyn':

$$\begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -\dfrac{p^3}{27} \end{cases}. \tag{U'}$$

Jak wiemy, dla rozwiązania układu (U') względem wielkości u^3, v^3 trzeba rozwiązać równanie kwadratowe

$$w^2 + qw - \frac{p^3}{27} = 0; \tag{RR}$$

nosi ono nazwę *równania rozwiązującego* (albo *równania rezolwenty*) dla równania (R). Jeśli w_0, w_1 są pierwiastkami (RR), to dowolna para (u, v), taka że $u^3 = w_0$, $v^3 = w_1$ (tzn. $u \in \sqrt[3]{w_0}$, $v \in \sqrt[3]{w_1}$) będzie spełniać (U'); jednakże chcąc, by oprócz (U') spełniony był mocniejszy układ (U), musimy zadbać o właściwe skorelowanie wyboru wartości

$u \in \sqrt[3]{w_0}$ i $v \in \sqrt[3]{w_1}$. Dlatego też postąpmy w następujący sposób:

Znajdźmy tylko **jeden** z pierwiastków w_0 równania (RR), następnie zaś **jedną** z wartości $u_0 \in \sqrt[3]{w_0}$ oraz dobierzmy v_0 tak, by spełnione było drugie z równań układu (U), tzn. $v_0 := -\frac{p}{3u_0}$. Wtedy także pierwsze z równań (U) będzie spełnione, skoro bowiem $w_0 = {u_0}^3$ jest jednym z pierwiastków (RR), to drugim (ze wzoru Viete'a na $w_0 w_1$) jest $w_1 = -\frac{p^3}{27 w_0} = {v_0}^3$, a więc ${u_0}^3 + {v_0}^3 = -q$ ze wzoru Viete'a na $w_0 + w_1$.

Zauważmy teraz, że dwie inne pary (u_1, v_1), (u_2, v_2), otrzymane z pary (u_0, v_0) za pomocą wzorów

$$\begin{cases} u_1 := \omega u_0 \\ v_1 := \overline{\omega} v_0 \end{cases}, \qquad \begin{cases} u_2 := \overline{\omega} u_0 \\ v_2 := \omega v_0 \end{cases},$$

dają te same wartości $u^3 + v^3$ (gdyż $\omega^3 = 1$) oraz te same wartości uv (gdyż $\omega\overline{\omega} = 1$), a więc na równi z parą (u_0, v_0) spełniają układ (U).

Wobec tego dowiedliśmy, że każda z liczb $\begin{cases} z_0 = u_0 + v_0 \\ z_1 = u_1 + v_1 = \omega u_0 + \overline{\omega} v_0 \\ z_2 = u_2 + v_2 = \overline{\omega} u_0 + \omega v_0 \end{cases}$
(są to właśnie tzw. *wzory Cardana*) jest pierwiastkiem równania (R).

41. **Przykład**. $z^3 - 6z - 40 = 0$; $\begin{cases} u^3 + v^3 = 40, \\ uv = 2, \end{cases}$ $w^2 - 40w + 8 = 0$, $\Delta = 16(100 - 2) =$ $= 16 \cdot 49 \cdot 2$, $w_0 = 20 + 14\sqrt{2}$, $\begin{cases} u_0 = \sqrt[3]{20 + 14\sqrt{2}}, \\ v_0 = \sqrt[3]{20 - 14\sqrt{2}}. \end{cases}$ Skądinąd $z_0 = u_0 + v_0 = 4$, gdyż $W(z) = (z - 4)(z^2 + 4z + 10)$; wobec tego $\{z_0, z_1, z_2\} = \{4, -2 \pm i\sqrt{6}\}$. Co więcej, można zauważyć, że $\begin{cases} u_0 = 2 + \sqrt{2}, \\ v_0 = 2 - \sqrt{2}, \end{cases}$ a zatem $\begin{cases} z_1 = \ldots = -2 + i\sqrt{6}, \\ z_2 = \ldots = -2 - i\sqrt{6}. \end{cases}$

42. **Przykład**. $z^3 - 15z - 10 = 0$; $\begin{cases} u^3 + v^3 = 10, \\ uv = 5, \end{cases}$ $w^2 - 10w + 125 = 0$, $\Delta = -400$, $w_0 = 5 + 10i = \sqrt{125}\frac{1 + 2i}{\sqrt{5}} = \left(\sqrt{5}\right)^3 e^{i\varphi}$, gdzie $\left\{ \begin{array}{l} \cos\varphi = \frac{1}{\sqrt{5}} \\ \sin\varphi = \frac{2}{\sqrt{5}} \end{array} \right\}$, tzn. $\varphi = \operatorname{arc\,tg} 2$. Biorąc $\begin{cases} u_0 = \sqrt{5} e^{\frac{i\varphi}{3}} \\ v_0 = \sqrt{5} e^{-\frac{i\varphi}{3}} \end{cases}$ dostajemy $z_0 = 2\sqrt{5}\cos\frac{\varphi}{3}$, $z_{1,2} = 2\sqrt{5}\cos\frac{\varphi \pm 2\pi}{3}$. Numerycznie: $\varphi = 1.107149 = 63.4349°$, $z_0 = 4.17103$, $x_z = -3.48260$, $z_2 = -0.688429$.

43. **Fakt**. Jeśli $\begin{cases} u^3 + v^3 = -q \\ 3uv = -p \end{cases}$ oraz $z^3 + pz + q = 0$, to $z = u + v$ lub $z = \omega u + \overline{\omega} v$ lub $z = \overline{\omega} u + \omega v$. Zatem metoda Cardana daje **wszystkie** pierwiastki równania $z^3 + pz + q = 0$.

Łatwo sprawdzić tożsamość $(z - u - v)(z - \omega u - \overline{\omega} v)(z - \overline{\omega} u - \omega v) = z^3 - 3uvz - u^3 - v^3$ (wymnażając i porządkując względem potęg z), z której oczywiście wynika teza.

44. **Fakt**. Każdy wielomian $w(x) = x^3 + ax^2 + bx + c$ (gdzie $a, b, c \in \mathbb{C}$) można rozłożyć na czynniki stopnia pierwszego([2]):

[2]Jak zobaczymy w następnym podrozdziale, każdy wielomian *dowolnego stopnia* ma rozkład na czynniki stopnia pierwszego o współczynikach z $\mathbb{C}$; jednakże w przypadku ogólnym dowód wymaga zastosowania tzw. *podstawowego twierdzenia algebry*.

$$\exists x_0, x_1, x_2 \in \mathbb{C}: \; w(x) = (x - x_0)(x - x_1)(x - x_2);$$

przy tym $x_0 + x_1 + x_2 = -a$ (wzór Viete'a), więc równanie zredukowane (i tylko takie) ma pierwiastki o zerowej sumie.

Weźmy $C \in \mathbb{C}$ takie, że $W(z) := w(z + C)$ ma postać $z^3 + pz + q$, zob. 39; określmy z_0, z_1, z_2 wzorami Cardana, wtedy $W(z) = (z - z_0)(z - z_1)(z - z_2)$ (zob. tożsamość odnotowaną w 43), więc $w(x) = W(x - C)$ ma żądany rozkład, gdzie $x_k = C + z_k$. Ostatnia część tezy wynika z rozwinięcia $(x - x_0)(x - x_1)(x - x_2)$ wzgl. potęg x.

Przypadek, gdy współczynniki są rzeczywiste

Niech teraz $p, q \in \mathbb{R}$; wtedy $W(\overline{z}) = \overline{z}^3 + p\overline{z}^2 + q = \overline{z}^3 + \overline{p}\,\overline{z}^2 + \overline{q} = \overline{W(z)}$, więc jeśli $W(z) = 0$, to także $W(\overline{z}) = 0$; wobec tego zbiór $\{z_0, z_1, z_2\}$ pierwiastków $W(z)$ jest *niezmienniczy względem operacji sprzężenia* (tzn. jest symetryczny względem osi rzeczywistej). Co więcej, wobec $W(z) = (z - z_0)(z - z_1)(z - z_2)$ tożsamość $\overline{W(\overline{z})} = W(z)$ oznacza, że trójka $\overline{z}_0, \overline{z}_1, \overline{z}_2$ może się różnić tylko kolejnością od z_0, z_1, z_2; łatwo stąd wynika, że choć jeden z pierwiastków (powiedzmy z_0) należy do $\mathbb{R}$.

W takim razie mamy dwie następujące możliwości:

(1) $z_0, z_1, z_2 \in \mathbb{R}$, czyli wszystkie pierwiastki $W(z)$ są rzeczywiste;
(2) $z_0 \in \mathbb{R}$, $z_1 \in \mathbb{C} \setminus \mathbb{R}$, $z_2 = \overline{z}_1$, czyli $W(z)$ ma dwa nierzeczywiste pierwiastki wzajemnie sprzężone.

Przypomnijmy, że obu przypadkach $z_0 + z_1 + z_2 = 0$ (wzór Viete'a).

45. **Fakt**. Jeśli $W(z) = z^3 + pz + q$ ma rzeczywiste współczynniki: $p, q \in \mathbb{R}$, to $\Big(W(z) = 0$ ma tylko rzeczywiste pierwiastki$\Big) \iff \Delta = q^2 + \frac{4p^3}{27} \leqslant 0$.

$\boxed{\Leftarrow}$ Jeśli $W(a + ib) = 0$, gdzie $a, b \in \mathbb{R}$, $b \neq 0$, to $W(z)$ ma pierwiastki $a \pm ib$ oraz $-2a$ (wzór Viete'a), więc $W(z) = (z^2 - 2az + a^2 + b^2)(z + 2a)$, skąd $p = -3a^2 + b^2$, $q = 2a(a^2 + b^2)$; wtedy $\Delta = \ldots = \frac{4b^2}{27}(9a^2 + b^2)^2 > 0$. $\boxed{\Rightarrow}$. Jeśli $W(z)$ ma tylko rzeczywiste pierwiastki $a, b, c = -a - b$, to $W(z) = (z - a)(z - b)(z + a + b)$, więc $p = -(a^2 + ab + b^2)$, $q = ab(a + b)$, skąd $\Delta = -\frac{1}{27}(a - b)^2(a + 2b)^2(2a + b)^2 \leqslant 0$.

Alternatywny dowód:
$W(z)$ jest nieparzystego stopnia, więc ma choć jeden rzeczywisty pierwiastek; stąd $\exists a, b \in \mathbb{R}: W(z) = (z - a)(z^2 + az + b)$, a wobec tego $\left\{ \begin{array}{l} p = -a^2 + b \\ q = -ab \end{array} \right\}$. Łatwy rachunek daje teraz $\Delta = \frac{1}{27}(2a^2 + b)^2(4b - a^2)$, więc $\Delta \leqslant 0 \iff D := a^2 - 4b \geqslant 0$; stąd wynika teza, ponieważ D jest wyróżnikiem trójmianu $z^2 + az + b$.

1.10 Algebraiczna domkniętość ciała

46. Przypomnijmy najpierw kilka pojęć związanych z wielomianami([3]). *Wielomianem zmiennej z o współczynnikach z ciała $\mathbb{K}$* nazywamy wyrażenie postaci $W(z) = a_0 + a_1 z + \ldots + a_n z^n$, gdzie $a_0, \ldots, a_n \in \mathbb{K}$. Dwa wielomiany są równe $\overset{\text{def}}{\iff}$ mają jednakowe współczynniki; zatem wielomian $W(z) = a_0 + a_1 z + \ldots + a_n z^n$

[3]Bardziej gruntownie omówimy te sprawy w następnym rozdziale; tu wystarczy nam właściwie podstawowa szkolna wiedza.

jest *niezerowy*, $W(\cdot) \neq 0$, jeśli ma różny od zera choć jeden ze współczynników $a_0, \ldots, a_n \in \mathbb{K}$. Liczbę $\deg W(\cdot) := \max\{k : a_k \neq 0\}$, tzn. maksymalny wykładnik, z jakim zmienna z występuje faktycznie w $W(z)$, nazywamy *stopniem* niezerowego wielomianu $W(\cdot)$. Oczywiście stopień iloczynu wielomianów jest sumą ich stopni.

Zauważmy, że wielomian $W(z) := z + z^2$ jest niezerowy nawet dla ciała $\mathbb{K} = \mathrm{GF}_2$ (zob. punkt 5), chociaż wtedy $W(O) = W(I) = O$, tzn. $\forall z \in \mathbb{K} : W(z) = 0$ ([4]).

47. **Oznaczenie.** Symbolem $\mathbb{K}[\,\cdot\,]$ będziemy oznaczać zbiór wszystkich wielomianów (jednej zmiennej) o współczynnikach z ciała $\mathbb{K}$.

Często zamiast $W(\cdot) \in \mathbb{K}[\,\cdot\,]$ pisze się bardziej tradycyjnie $W(z) \in \mathbb{K}[z]$; symbol $\mathbb{K}[z]$ oznacza to samo co $\mathbb{K}[\,\cdot\,]$, lecz dodatkowo nakazuje zmienną oznaczać literą z.

48. **'Twierdzenie Bezouta'.** Jeśli $W(z) \in \mathbb{K}[z]$ oraz $z_0 \in \mathbb{K}$, to

$$\begin{pmatrix} W(z_0) = 0, \text{ tzn. } z_0 \text{ jest} \\ \text{pierwiastkiem } W(z) \end{pmatrix} \Leftrightarrow \begin{pmatrix} (z - z_0) \mid W(z), \text{ tzn. } \exists W_1(z) \in \mathbb{K}[z] : \\ W(z) = (z - z_0)W_1(z) \end{pmatrix}$$

Zauważmy, że wielomian $W(z) - W(z_0)$ jest podzielny przez wielomian $z - z_0$, gdyż $W(z)$ jest sumą składników postaci $a_n z^n$, a mamy znaną tożsamość $a_n(z_n - {z_0}^n) = a_n(z - z_0)(z^{n-1} + z^{n-2}z_0 + \ldots + {z_0}^{n-1})$. Zatem $W(z) = (z - z_0)Q(z) + W(z_0)$, gdzie $Q(z)$ jest wielomianem. Stąd natychmiast wynika teza.

Nieco inny wariant dowodu:
Dzieląc $W(z)$ przez $z - z_0$ dostajemy wielomiany $Q(z)$ oraz $R(z)$ (*iloraz i reszta*), takie że $W(z) = (z - z_0)Q(z) + R(z)$, przy czym $\deg R(z) < \deg(z - z_0) = 1$, czyli $R(z) = C$ jest stałą: $W(z) = (z - z_0)Q(z) + C$. Ta tożsamość dla $z = z_0$ daje $W(z_0) = C$, a więc $W(z_0)$ *jest resztą z dzielenia* $W(z)$ *przez* $z - z_0$; stąd teza.

49. **Wnioski z twierdzenia Bezouta.** Niech $\mathbb{K}$ — dowolne ciało; wtedy

(A) Liczba pierwiastków niezerowego wielomianu $W(\cdot)$ jest $\leqslant \deg W(\cdot)$.
(B) Jeśli $W(\cdot) \in \mathbb{K}[\,\cdot\,]$ jest niezerowy, a podzbiór $S \subset \mathbb{K}$ ma więcej niż $\deg W(\cdot)$ elementów, to $\exists z_0 \in S : W(z_0) \neq 0$.
(C) Jeśli $n \in \mathbb{N}$, podzbiór $S \subset \mathbb{K}$ ma więcej niż n elementów oraz $\forall z \in S : a_0 + a_1 z + \ldots + a_n z^n = 0$, to $a_0 = a_1 = \ldots = a_n = 0$. ([5])

(A) Jeśli $W(z_1) = 0, \ldots, W(z_r) = 0$, gdzie $z_1, \ldots, z_r \in \mathbb{K}$ są parami różne, to $W(z) = (z - z_1)\ldots(z - z_r)\tilde{W}(z)$, gdzie $\tilde{W}(\cdot) \neq 0$, więc $\deg W(\cdot) = r + \deg \tilde{W} \geqslant r$. Punkty (B) oraz (C) wynikają natychmiast z (A).

50. **Definicja.** Jeśli z_0 jest pierwiastkiem $W(\cdot)$, to największą liczbę $k \in \mathbb{N}$, dla której $(z - z_0)^k \mid W(z)$, nazywamy *krotnością* pierwiastka z_0.

51. **Fakt.** Następujące warunki charakteryzujące ciało $\mathbb{K}$ są równoważne:
1° Każdy wielomian $W(\cdot) \in \mathbb{K}[\,\cdot\,]$ stopnia $\geqslant 1$ ma pierwiastek w $\mathbb{K}$, tzn.

$$\exists z_0 \in \mathbb{K} : \; W(z_0) = 0.$$

2° Każdy wielomian $W(\cdot) \in \mathbb{K}[\,\cdot\,]$ stopnia $\geqslant 1$ ma rozkład na czynniki

[4] Zauważmy, że dla innego ciała $\mathbb{K} = \mathrm{GF}_4$ jest inaczej: $W(a) = W(b) = I \neq O$.
[5] Fakt (C) wysławia się zwykle krótko, mówiąc że dla $|S| > n$ jednomiany $z^0, z^1 \ldots, z^n$ są *liniowo niezależne* jako funkcje na zbiorze S.

liniowe (tzn. stopnia 1):

$$\exists c \neq 0, z_1, ..., z_n \in \mathbb{K}: \; W(z) = c(z - z_1) \dots (z - z_n).$$

3° Każdy niezerowy wielomian $W(\cdot) \in \mathbb{K}[\,\cdot\,]$ ma w $\mathbb{K}$ tyle pierwiastków (liczonych z uwzględnieniem ich krotności), ile wynosi jego stopień.

$\boxed{1° \Rightarrow 2°}$ Z tw. Bezouta: $W(z_1) = 0 \Rightarrow \exists W_1(\cdot) \in \mathbb{K}[\,\cdot\,]: \; W(z) = (z - z_1)W_1(z)$; jeśli $\deg W_1 > 1$, to $\exists z_2 \in \mathbb{K}: \; W_1(z_2) = 0$, czyli $W_1(z) = (z - z_2)W_2(z)$, itd. $\boxed{2° \Rightarrow 3°}$ Wynika wprost z definicji krotności pierwiastka. $\boxed{3° \Rightarrow 1°}$ Oczywiste.

52. **Definicja**. Ciało $\mathbb{K}$ nazywa się *algebraicznie domknięte*, jeśli spełnione są powyższe równoważne warunki $1°, \dots, 3°$.

53. **Przykłady**. Nie są algebraicznie domknięte ciała: $\mathbb{R}$, $\mathbb{Q}$ (vide $z^2 + 1$), $\mathbb{Q} + i\mathbb{Q}$ (vide $z^2 - 2$), $\mathbb{Q}(\sqrt{2})$ (vide $z^2 - 3$), GF_2 (vide $z^2 + z + 1$).

54. Ważny przykład ciała algebraicznie domkniętego stanowi tzw. *ciało liczb algebraicznych.*

Liczbę $\lambda \in \mathbb{C}$ nazywamy *liczbą algebraiczną*, jeśli λ jest pierwiastkiem jakiegoś $\neq 0$ wielomianu z $\mathbb{Q}[\,\cdot\,]$, tzn. jeśli istnieje $n \in \mathbb{N}$ oraz wymierne (a nawet całkowite) współczynniki $a_0, \dots, a_n$, $a_n \neq 0$, takie że $a_0 + a_1\lambda + \dots + a_n\lambda^n = 0$.

Przykładami liczb algebraicznych są: $\frac{7}{13}$ (oczywiste), $\frac{2+3i}{7}$ (gdyż $\left(\lambda - \frac{2}{7}\right)^2 + \frac{9}{49} = 0$), $\sqrt[3]{-1 + \sqrt{5}} + 2$ (gdyż $\left((\lambda - 2)^3 + 1\right)^2 - 5 = 0$), $\sin 10°$ (gdyż $4\lambda^3 - 3\lambda + \frac{1}{2} = 0$), $\cos\left(\frac{1}{3} \operatorname{arc\,tg} 2\right)$ (gdyż $4\lambda^3 - 3\lambda = \cos(\operatorname{arc\,tg} 2) = \frac{1}{\sqrt{5}}$, czyli $(4\lambda^3 - 3\lambda)^2 - \frac{1}{5} = 0$), $\frac{2}{7}\sqrt[5]{3 - \sqrt{2}} - \frac{4i}{3}\sqrt{1 + 3\sqrt[3]{4}}$ (tu bezpośrednie sprawdzenie nie jest już tak proste).

Liczby *niealgebraiczne* istnieję także: skoro zbiór $\mathbb{Q}[\,\cdot\,]$ jest, na równi ze zbiorem $\mathbb{Q}$, przeliczalny, zaś każdy wielomian ma jedynie skończoną liczbę pierwiastków, to *zbiór liczb algebraicznych jest przeliczalny*, a więc (wskutek nieprzeliczalności $\mathbb{C}$) nie jest on całym zbiorem $\mathbb{C}$. Znacznie trudniej jest dowodzić, że jakaś konkretna liczba (np. $\pi \approx 3.14159265$, $e \approx 2.7182818$, $2^{\sqrt{2}}$, itp.) jest niealgebraiczna.

Okazuje się, że

- zbiór liczb algebraicznych jest ciałem — podciałem ciała $\mathbb{C}$; oznacza to, że operacje algebraiczne (dodawanie, odejmowanie, mnożenie i dzielenie) wykonywane na liczbach algebraicznych zawsze dają w wyniku liczby algebraiczne;
- ciało liczb algebraicznych jest algebraicznie domknięte (innymi słowy: jeśli $\lambda \in \mathbb{C}$ jest pierwiastkiem jakiegoś wielomianu, którego współczynniki są liczbami algebraicznymi, to λ jest też pierwiastkiem jakiegoś wielomianu, na ogół wyższego stopnia, o współczynnikach całkowitych.

Dowody tych twierdzeń można znaleźć w bardziej zaawansowanych podręcznikach algebry, np. w *Teorii ciał* J.Browkina.

1.11 Podstawowe Twierdzenie Algebry

55. **Twierdzenie** (*'Podstawowe twierdzenie algebry'*):

Ciało liczb zespolonych jest algebraicznie domknięte.

Znanych jest obecnie wiele rozmaitych dowodów tego (rzeczywiście bardzo ważnego) twierdzenia; każdy z nich zawiera jakiś 'element niealgebraiczny', a jest tak mówiąc z grubsza dlatego, że nieodzowne jest odwołanie się do tych własności ciała $\mathbf{C}$, które odróżniają je od ciała $\mathbf{Q} + i\mathbf{Q}$ (wymiernych liczb zespolonych), które nie jest algebraicznie domknięte. Taką cechą jest np. 'zupełność' ciała $\mathbf{C}$ jako przestrzeni

metrycznej, wynikająca z zupełności ciała $\mathbf{R}$ (opisanej np. aksjomatem o istnieniu kresów każdego podzbioru $\mathbf{R}$).

Przedstawimy tu (oczywiście we współczesnej formie) pierwszy historycznie (znaleziony przez C.F.Gaussa) dowód PTA; skorzystamy w nim z paru pojęć topologicznych (ciągłość, spójność), odwołując się do ich intuicyjnego, potocznego sensu; formalne i zupełne zrozumienie dowodu stanie się możliwe wkrótce, gdy zapoznamy się z tymi pojęciami topologicznymi na zajęciach z analizy matematycznej.

56. **Fakt**. Jeśli $P = [a,b] \subset \mathbb{R}$ oraz $\nu : P \to U$ jest funkcją ciągłą, to istnieje funkcja ciągła $\phi : P \to \mathbb{R}$ taka, że $\forall s \in P : \nu(s) = \exp\big(i\phi(s)\big)$.

Skorzystamy z tożsamości $\boxed{\forall u \in U :\ \mathrm{Re}\, u > 0 \ \Rightarrow\ \exp\left(i \operatorname{arc\,tg} \frac{\mathrm{Im}\, u}{\mathrm{Re}\, u}\right) = u}$ (T); wynika ona natychmiast z dwóch faktów: (a) $\forall \varphi \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[:\ \operatorname{arc\,tg} \varphi = \varphi$ oraz (b) $\forall u \in U : \mathrm{Re}\, u > 0 \ \Rightarrow\ \exists \varphi \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[: u = \exp(i\varphi)$.

P jest zwarty, więc funkcja ciągła ν jest jednostajnie ciągła; stąd $\exists \varepsilon > 0 : \forall s, s' \in P :$ $|s - s'| < \varepsilon \ \Rightarrow\ |\nu(s) - \nu(s')| < 1$; mamy zatem $|s - s'| < \varepsilon \ \Rightarrow\ \left|\frac{\nu(s')}{\nu(s)} - 1\right| =$ $|\nu(s') - \nu(s)| < 1$. Zatem ustalając $n \in \mathbb{N}$ takie, że $n > \frac{b-a}{\varepsilon}$, dla $a_k := a + k\frac{b-a}{n}$ mamy $a_0 = a < a_1 < \ldots < a_n = b$ oraz $\forall s \in P_k := [a_k, a_{k+1}] : \left|\frac{\nu(s)}{\nu(a_k)} - 1\right| < 1$, więc $\mathrm{Re}\, \frac{\nu(s)}{\nu(a_k)} > 0$. Możemy więc kolejno, dla $k = 0, 1, \ldots, n-1$, określić funkcje $\phi_k : P_k \to \mathbb{R}$, wzorem $\phi_k(s) := \phi_{k-1}(a_k) + \operatorname{arc\,tg} \frac{\mathrm{Im}\, {\nu(s)}/{\nu(a_k)}}{\mathrm{Re}\, {\nu(s)}/{\nu(a_k)}}$, przy czym dla $k = 0$ jako $\phi_{k-1}(a_k)$ bierzemy $\mathrm{Arg}\, \varrho(a_0)$. Z tożsamości (T) wynika, że $\exp\big(i\Phi_k(s)\big) = \nu(s)$; ponadto $\phi_k(a_k) = \exp\big(i\phi_{k-1}(a_k)\big) = \phi_{k-1}(a_k)$, więc: (I) poprawne jest określenie funkcji $\phi : P = P_0 \cup \ldots \cup P_{n-1} \to \mathbb{R}$ wzorem $s \in P_k \ \Rightarrow\ \phi(s) := \phi_k(s)$ oraz (II) funkcja ϕ jest ciągła (dzięki ciągłości ϕ_k i ich 'zgodności na końcach dziedzin').

Definicja. Odwzorowanie $\gamma : [0, 2\pi] \to \Omega$ nazywamy *pętlą w zbiorze* $\Omega \subset \mathbb{C}$, jeżeli jest ciągłe oraz $\gamma(0) = \gamma(2\pi)$ (warunek *zamkniętości*).

Przykład. $s \mapsto \gamma(s) := Ce^{ins}$ jest pętlą $\iff e^{2\pi i n} = 1 \iff n \in \mathbb{Z}$.

57. **Fakt** (*liczba obiegów*). Jeśli γ jest pętlą w zbiorze $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, to:

(a) istnieje taka ciągła funkcja $\phi : [0, 2\pi] \to \mathbb{R}$, że $\gamma(s) = |\gamma(s)|e^{i\phi(s)}$;

(b) liczba $N(\gamma) := \dfrac{\phi(2\pi) - \phi(0)}{2\pi}$ jest całkowita oraz zależy tylko od γ, a nie od wyboru funkcji ϕ.

Ad (a) Wystarczy zastosować fakt 56, biorąc $\nu(s) := \dfrac{\gamma(s)}{|\gamma(s)|}$.

Ad (b) $\gamma(0) = \gamma(2\pi) \ \Rightarrow\ e^{i\phi(0)} = e^{i\phi(2\pi)} \ \Rightarrow\ \phi(2\pi) - \phi(0) \in 2\pi\mathbb{Z} \ \Rightarrow\ N(\gamma) \in \mathbb{Z}$. Jeśli funkcje $\phi, \tilde{\phi} : [0, 2\pi] \to \mathbb{R}$ są ciągłe oraz $\gamma(s) = |\gamma(s)|e^{i\phi(s)} = |\gamma(s)|e^{i\tilde{\phi}(s)}$, to $\tilde{\phi}(s) - \phi(s) = 2\pi n(s)$, gdzie $n(s) \in \mathbb{Z}$ oraz funkcja $[0, 2\pi] \ni s \mapsto n(s) \in \mathbb{Z}$ jest ciągła; stąd n-obraz jest spójnym podzbiorem $\mathbb{Z}$, tzn. $n(s) = \mathrm{const}$; zatem $\phi(2\pi) - \phi(0) = \tilde{\phi}(2\pi) - \tilde{\phi}(0)$.

Definicja. $N(\gamma)$ nazywa się *liczbą obiegów* (wokół punktu 0) pętli γ.

Przykład. Pętla $s \mapsto \gamma(s) = Ce^{ins}$ ma liczbę obiegów $N(\gamma) = n \in \mathbb{Z}$.

58. **Fakt**. Niech γ, δ będą pętlami w $\mathbb{C}^*$. Wtedy:

(a) $N(\gamma\delta) = N(\gamma) + N(\delta)$ (iloczyn 'punktowy': $(\gamma\delta)(s) := \gamma(s)\delta(s)$);

(b) $N(\gamma^{-1}) = -N(\gamma)$ (odwrotność 'punktowa': $(\gamma^{-1})(s) := \frac{1}{\gamma(s)}$).

Ad (a) $\Big(\gamma(s) = |\gamma(s)|e^{i\,\phi(s)},\ \delta(s) = |\delta(s)|e^{i\psi(s)}\Big) \Rightarrow \gamma(s)\delta(s) = |\gamma(s)\delta(s)|e^{i(\phi(s)+\psi(s))}$.
Ad (b) Wynika to wprost z (a): $N(\gamma) + N(\gamma^{-1}) = N(\gamma\gamma^{-1}) = N(\text{const}) = 0$.

59. **Fakt**. Jeśli γ jest pętlą w zbiorze $\mathcal{V}_+ = \{z \in \mathbb{C} : \operatorname{Re} z > 0\}$, to $N(\gamma) = 0$.

Z założenia $\operatorname{Re}\gamma(s) > 0$, więc ma sens $\phi(s) := \operatorname{arc\,tg}\frac{\operatorname{Im}\gamma(s)}{\operatorname{Re}\gamma(s)}$; tożsamość (T) z punktu 56 dla $u = \frac{\gamma(s)}{|\gamma(s)|}$ pokazuje, że $\gamma(s) = |\gamma(s)|e^{i\,\phi(s)}$, a przy tym $\phi(0) = \phi(2\pi)$.

60. **Fakt**. Jeśli 1-parametrowa rodzina $\{\gamma_r\}$ pętli w $\mathbb{C}^*$ zależy w sposób ciągły od parametru $r \in \mathbb{R}$ (tzn. funkcja $(r, s) \mapsto \gamma_r(s)$ jest ciągła), to liczba obiegów $N(\gamma_r)$ jest stała (nie zależy od r).

Ustalmy parametr r_0; pętla $\gamma_r{\gamma_{r_0}}^{-1}$ dla $r = r_0$ jest punktem $1 \in \mathcal{V}_+$, więc $\exists\varepsilon > 0$ takie, że $\gamma_r{\gamma_{r_0}}^{-1}$ dla $|t - t_0| < \varepsilon$ jest pętlą w $\mathcal{V}_+$, więc $N(\gamma_r{\gamma_{r_0}}^{-1}) = 0$ (zob. 59), skąd $N(\gamma_r) = N(\gamma_{r_0})$ (zob. 58).

61. **Fakt**. Jeśli $W(z) = a_0 + a_1 z + \ldots + a_n z^n$, gdzie $a_0, \ldots, a_n \in \mathbb{C}$, przy czym $a_n \neq 0$, to

$$\exists r_0 > 0 : \forall z \in \mathbb{C} :\ |z| > r_0 \Rightarrow \frac{W(z)}{a_n z^n} \in K(1;1) \subset \mathcal{V}_+.$$

Oznaczmy $A := \max\Big\{\Big|\frac{a_0}{a_n}\Big|, \ldots, \Big|\frac{a_{n-1}}{a_n}\Big|\Big\}$, wtedy dla $|z| > 1$ mamy:

$$\Big|\frac{W(z)}{a_n z^n} - 1\Big| = \Big|\frac{a_0}{a_n z^n} + \frac{a_1 z}{a_n z^n} + \ldots + \frac{a_{n-1}z^{n-1}}{a_n z^n}\Big| \leqslant \Big|\frac{a_0}{a_n z^n}\Big| + \ldots + \Big|\frac{a_{n-1}z^{n-1}}{a_n z^n}\Big| \leqslant$$
$$\leqslant \frac{A}{|z|^n}(1 + |z| + \ldots + |z|^{n-1}) = \frac{A}{|z|^n}\,\frac{|z|^n - 1}{|z| - 1} < \frac{A}{|z| - 1},$$

więc dla $|z| > 1 + A =: r_0$ spełniona jest nierówność $\Big|\frac{W(z)}{a_n z^n} - 1\Big| < 1$, QED.

Dowód Podstawowego Twierdzenia Algebry:

Niech $W(z) = a_0 + \ldots + a_n z^n$. Przypuśćmy, że $\forall z \in \mathbb{C} :\ W(z) \in \mathbb{C}^*$. Dla $r \geqslant 0$ określmy pętlę γ_r w $\mathbb{C}^*$ wzorem $\gamma_r(s) := W(re^{is})$. Wtedy γ_r jest iloczynem punktowym $\gamma_r = \tilde{\gamma}_r\delta_r$, gdzie $\tilde{\gamma}_r(s) := \frac{W(z)}{a_n z^n}\Big|_{z=re^{is}}$ jest dla $r > r_0$ pętlą w $\mathcal{V}_+$ (zob. 61), więc $N(\tilde{\gamma}_r) = 0$ (zob. 59), zaś $\delta_r(s) := a_n r^n e^{ins}$, jest pętlą taką, że $N(\delta_r) = n$. Stąd $N(\gamma_r) = n$ dla dostatecznie dużych r, wiemy zaś z 60, że $N(\gamma_r)$ jest stałe. Tymczasem $N(\gamma_0) = 0$, gdyż $\gamma_0 = a_0$ jest pętlą 'punktową'. Sprzeczność!

62. **Wnioski z Podstawowego Twierdzenia Algebry**.

(1) Każdy wielomian $W \in \mathbb{C}[\,\cdot\,]$ stopnia $\geqslant 1$ ma w $\mathbb{C}[\,\cdot\,]$ rozkład na czynniki stopnia 1:

$$\exists c \in \mathbb{C}^*, z_1, \ldots, z_n \in \mathbb{C} :\ W(z) = c(z - z_1)\ldots(z - z_n).$$

(2) Każdy wielomian $W \in \mathbb{R}[\,\cdot\,]$ stopnia $\geqslant 1$ ma w $\mathbb{R}[\,\cdot\,]$ rozkład na czynniki stopnia 1 i/lub czynniki stopnia 2 o ujemnych wyróżnikach:

$$\left(\begin{array}{c}\exists c \in \mathbb{R}^*, \exists x_1, \ldots, x_k \in \mathbb{R}\\ \exists p_1, q_1, \ldots, p_l, q_l \in \mathbb{R}\end{array}\right) :\ W(x) = c\prod_{i=1}^{k}(x - x_i)\cdot\prod_{j=1}^{l}(x^2 + p_j x + q_j),$$

przy czym $\Delta_j = p_j^2 - 4q_j < 0$ (może być $k = 0$ lub $l = 0$, gdy brak w rozkładzie $W(x)$ czynników stopnia 1 lub 2).

(3) Jeśli k jest liczbą rzeczywistych pierwiastków (z uwzględnieniem krotności) wielomianu $W \in \mathbb{R}[\,\cdot\,]$, to $\deg W - k$ jest liczbą parzystą. W szczególności każdy wielomian stopnia nieparzystego ma pierwiastek rzeczywisty, gdyż wtedy k jest nieparzyste, a tym bardziej $k \neq 0$.

Ad(1) Wynika wprost z 51 oraz Podstawowego Twierdzenia Algebry.

Ad(2) Zastosujemy indukcję względem stopnia W. Jeśli $\deg W = 1$, teza jest oczywista. Niech więc $\deg W > 1$; wiemy, że $W(z)$ ma pierwiastek w zbiorze $\mathbb{C}$, tj. $\exists z_0 = x_0 + iy_0 \in \mathbb{C} : W(z_0) = 0$. Rozważmy dwie możliwości:

1° Jeśli $y_0 = 0$, to $(x - x_0) \mid W(x)$, więc $\exists \tilde{W} \in \mathbb{R}[\,\cdot\,] :\ W(x) = (x - x_0)\tilde{W}(x)$.

2° Jeśli $y_0 \neq 0$, to $W(z_0) = 0$ oraz $W(\overline{z}_0) = 0$ (to ostatnie dostaje się, biorąc sprzężenie obu stron równości $W(z_0) = 0$ i korzystając z tego, że współczynniki W należą do $\mathbb{R}$) powodują, że $W(x)$ jest podzielny przez $(x - z_0)(x - \overline{z}_0) = x^2 + px + q$, przy czym $p = -(z_0 + \overline{z}_0) = -2x_0$, $q = |z_0|^2 = x_0^2 + y_0^2$, więc $p, q \in \mathbb{R}$; wobec tego $\exists \tilde{W} \in \mathbb{R}[\,\cdot\,] : W(x) = (x^2 + px + q)\tilde{W}(x)$, ponadto $\Delta = p^2 - 4q = -4y_0^2$.

Ponieważ w obu przypadkach $\deg \tilde{W} < \deg W$, więc z założenia indukcyjnego $\tilde{W}$ ma rozkład na czynniki stopnia $\leqslant 2$, stąd zaś natychmiast dostajemy rozkład W.

Ad(3) Wynika bezpośrednio z (2).

63. **Ćwiczenie**. Dowieść, że każde ciało algebraicznie domknięte jest nieskończone.

Wskazówka. Jeśli $\mathbb{K}$ jest skończone i składa się z elementów $x_1, \ldots, x_n$, to łatwo określić wielomian $W(\,\cdot\,) \in \mathbb{K}[\,\cdot\,]$, taki że $\deg W(\,\cdot\,) = n$ oraz $\forall x \in \mathbb{K} : W(x) = 1$.

2 Wielomiany, pierścienie, podzielność

2.1 Funkcje wielomianowe i wielomiany

Niech $\mathbb{K}$ będzie ustalonym ciałem.

64. **Definicja**. *Funkcją wielomianową* o współczynnikach z ciała $\mathbb{K}$ nazywa się funkcję postaci $D \ni z \mapsto W(z) = a_0 + a_1 z + \ldots + a_n z_n$.

Zbiór D, będący dziedziną tej funkcji, jest zwykle podzbiorem ciała $\mathbb{K}$; często jednak mamy do czynienia z sytuacją ogólniejszą, gdy D jest podzbiorem jakiegoś większego ciała $\tilde{\mathbb{K}} \supset \mathbb{K}$; na przykład dla $\mathbb{K} = \mathbb{Q}$ wielkość $W(z)$ ma sens dla $z \in \mathbb{R}$, a nawet dla $z \in \mathbb{C}$. Nierzadko jednak nawet to ogólniejsze założenie o D byłoby nazbyt krępujące, np. przydatne bywa rozważanie $W(z)$ (o współczynnikach $a_i \in \mathbb{K}$) w przypadku, gdy z jest 'macierzą' o wyrazach z $\mathbb{K}$, np. $z = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, albo 'operatorem', np. 'operatorem różniczowania' $z = \frac{d}{dt}$, albo czymś jeszcze innym. Dlatego też w definicji funkcji wielomianowej *o zbiorze D założymy jedynie, że jego elementy da się dodawać i mnożyć, a także że da się mnożyć elementy zbioru D przez elementy ciała $\mathbb{K}$.*

65. W 'uproszczonym' kursie algebry nie odróżnia się «wielomianu» od «funkcji wielomianowej»; w dodatku w takiej ułatwionej definicji albo pomija się kwestię dziedziny 'funkcji' W, albo przyjmuje się $D := \mathbb{K}$. Jednak na dłuższą metę jest to niewystarczające i niewygodne; z tego względu w 'porządnej' definicji *wielomian* będziemy rozumieć nie jako funkcję, ale jako «receptę na funkcje» albo — innymi słowy — jako całą «klasę funkcji wielomianowych» różniących się między sobą dziedzinami, lecz określonych wspólnym wzorem $W(z) := a_0 + a_1 z + \ldots + a_n z^n$ z danymi $a_i \in \mathbb{K}$. Oczywiście taka «recepta», czy też «pre-funkcja» jest w pełni określona ciągiem współczynników $a_0, a_1, \ldots$, a więc wygodnie będzie przyjąć, że *wielomian* to to samo, co ciąg jego współczynników.

Np. $z^4 - 3z^2 + 5z - 2$ możemy utożsamić z ciągiem $(-2, 5, -3, 0, 1, 0, 0, 0, \ldots)$ itp.:

66. **Definicja**. *Wielomianem* $W \in \mathbb{K}[\cdot]$ nazywa się ciąg nieskończony $(a_0, a_1, a_2, \ldots)$ elementów ciała $\mathbb{K}$, mający co najwyżej skończoną liczbę wyrazów różnych od zera, tzn. ciąg spełniający następujący warunek

$$\exists n \in \mathbb{Z}_+ : \forall k > 0 : a_k = 0.$$

W szczególności *wielomianem zerowym* jest ciąg $(0, 0, \ldots)$ złożony z samych zer. Jeśli wielomian $W = (a_0, a_1, \ldots)$ jest niezerowy, to jego *stopniem* nazywa się liczbę $\deg W := \max\{k : a_k \neq 0\}$; wygodnie jest dodatkowo umówić się, że wielomian zerowy ma stopień równy $-\infty$ ([6]).

Zgodnie z naszą intencją każdemu wielomianowi $W = (a_0, a_1, a_2, \ldots)$ i dowolnej dziedzinie D, dobranej stosownie do ciała $\mathbb{K}$, odpowiada

[6] Dzięki temu wzór $\deg(WV) = \deg W + \deg V$ będzie prawdziwy zawsze, nawet wtedy, gdy $W = 0$ lub $V = 0$.

funkcja wielomianowa określona wzorem $W(z) := a_0 + a_1 z + a_2 z^2 + \ldots$ (suma występująca w tym wzorze jest w istocie skończona).

67. W zbiorze $\mathbb{K}[\,\cdot\,]$ wszystkich wielomianów o współczynnikach z $\mathbb{K}$ określa się operacje dodawania, mnożenia oraz mnożenia przez elementy z $\mathbb{K}$:
Jeśli $W = (a_0, a_1, a_2, \ldots)$, $V = (b_0, b_1, b_2, \ldots)$, to

suma wielomianów: $W + V = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$;
iloczyn wielomianów: $W \cdot V = (c_0, c_1, c_2, \ldots)$, gdzie

$$c_0 := a_0 b_0,\ c_1 := a_0 b_1 + a_1 b_0,\ c_2 := a_0 b_2 + a_1 b_1 + a_2 b_0,$$

ogólnie,

$$c_m = a_0 b_m + a_1 b_{m-1} + \ldots + a_m b_0 = \sum_{(k,l):\, k+l=m} a_k b_l\,;$$

iloczyn W przez liczbę $\lambda \in \mathbb{K}$: $\lambda W = (\lambda a_0, \lambda a_1, \lambda a_2, \ldots)$.

Powyższe definicje działań są oczywiście dobrane w taki sposób, by sumie wielomianów odpowiadała suma ich funkcji wielomianowych, iloczynowi wielomianów — iloczyn funkcji wielomianowych, itd.:

$$(W + V)(z) = W(z) + V(z), \qquad (W \cdot V)(z) = W(z) \cdot V(z),$$
$$(\lambda W)(z) = \lambda W(z).$$

68. **Fakt**. (1) $\deg(\lambda W) = \deg W$ dla $\lambda \in \mathbb{K}^*$;
(2) $\deg(W + V) \leqslant \max(\deg W, \deg V)$,
przy czym jeśli $\deg W \neq \deg V$, to zamiast $\leqslant$ mamy równość;
(3) $\deg(W \cdot V) = \deg W + \deg V$.

Własności (1) i (2) są widoczne wprost z definicji.

Prawdziwość wzoru (3) jest oczywista, gdy W lub V jest zerem, załóżmy więc, że $W, V \neq 0$. Niech $W = (a_0, \ldots, a_n, 0, \ldots)$, $V = (b_0, \ldots, b_n, 0, \ldots)$, przy czym niech $a_m \neq 0$, $b_n \neq 0$, tzn. $\deg W = m$ oraz $\deg V = n$. W sumie $c_{m+n} = \sum_{k+l=m+n} a_k b_l$ jeden składnik, mianowicie $a_m b_n$, jest niezerowy, wszystkie pozostałe są zerami, gdyż $k > m \ \Rightarrow\ a_k = 0$ oraz $l > n \ \Rightarrow\ b_l = 0$; zatem $c_{m+n} \neq 0$. To samo sprawia, że dla $r > m + n$ wszystkie składniki sumy $c_r = \sum_{k+l=r} a_k b_l$ są $=0$.

69. Podane w 67 definicje operacji na wielomianach prowadzą do wzoru

$$W = (a_0, a_1, a_2, \ldots, a_n, 0, \ldots) = a_0\mu_0 + a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n,$$

gdzie $\mu_0 := (1, 0, \ldots)$, $\mu_1 := (0, 1, 0, \ldots)$, $\mu_2 := (0, 0, 1, 0, \ldots)$, itd. Z definicji mnożenia w $\mathbb{K}[\,\cdot\,]$ wynika łatwo, że $\mu_k \cdot \mu_l = \mu_{k+l}$, a zatem $\mu_k = \zeta^k$ jest potęgą elementu $\zeta := \mu_1$; pozwala to nadać bardziej swojski wygląd powyższemu rozkładowi wielomianu $W \in \mathbb{K}[\,\cdot\,]$:

$$W = a_0\zeta^0 + a_1\zeta + a_2\zeta^2 + \ldots + a_n\zeta^n$$

(zwykle $a_0\zeta^0$ skracamy w zapisie do a_0). Wielomian $\zeta = (0, 1, 0, 0, \ldots)$ nazywa się zazwyczaj *zmienną* (albo *generatorem pierścienia wielomianów*); odpowiadającą mu funkcją wielomianową jest oczywiście id_D; elementy $a_k\zeta^k \in \mathbb{K}[\,\cdot\,]$ noszą nazwę *jednomianów*.

2.2 Pojęcie pierścienia; przykłady

70. **Definicja**. *Pierścieniem* nazywa się zbiór R, wyposażony w dwa działania 2-argumentowe '$+$' i '$\cdot$', spełniające następujące aksjomaty:

1° *Łączność* $\begin{cases} \text{dodawania:} & (a+b)+c = a+(b+c), \\ \text{i mnożenia:} & (ab)c = a(bc). \end{cases}$

2° *Przemienność dodawania:* $a+b = b+a$.

3° *Istnienie zera:* $\exists 0 \in R : \forall a \in R : a+0 = a$.

4° *Istnienie elementu przeciwnego:* $\forall a \in R : \exists b \in R : a+b = 0$.
Oznaczenie: $-a := b$, jeśli $a+b=0$ (element *przeciwny* do a).

5° *Rozdzielność mnożenia względem dodawania:* $\begin{cases} a(b+c) = ab+ac, \\ (b+c)a = ba+ca. \end{cases}$

[Umowa, że 'mnożenie ma pierwszeństwo przed dodawaniem', pozwala zmniejszyć w zapisie liczbę nawiasów, a więc np. $ab+ac$ jest skróconą formą $(a\cdot b)+(a\cdot c)$].

71. **Definicja**. Pierścień R jest *przemienny*, gdy ma przemienne mnożenie:

6° $\forall a, b \in R : ab = ba$;

pierścień z jedynką jest to pierścień mający element $1 \in R$, taki że

7° $\forall a \in R : 1\cdot a = a,\ a\cdot 1 = a$.

Ćwiczenie. Pierścień ma tylko jedno *zero*, tzn. element opisany aksjomatem 3°; każdy $a \in R$ ma tylko jeden *element przeciwny*; pierścień z jedynką ma tylko jedną *jedynkę*, tj. element o własności 7°.

72. **Definicja**. *Dziedziną całkowitości* nazywa się pierścień (przemienny, z jedynką) spełniający następujący 'substytut aksjomatu o istnieniu odwrotności':

8° $\forall a, b \in R :\ ab = 0 \Rightarrow (a = 0 \text{ lub } b = 0)$.

Element $a \in R \setminus \{0\}$ nazywa się (*nietrywialnym*) *dzielnikiem zera*, jeżeli $\exists b \in R \setminus \{0\} : ab = 0$. Zatem warunek 8° oznacza, że *w pierścieniu R nie ma nietrywialnych dzielników zera.*

73. **Przykłady pierścieni**.

(A) Każde ciało (np. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$) jest dziedziną całkowitości.

(B) Zbiór $\mathbb{Z}$ wszystkich liczb całkowitych ze zwykłymi działaniami '$+$' i '$\cdot$' jest dziedziną całkowitości.

(C) Zbiór wielomianów $\mathbb{K}[\,\cdot\,]$ o współczynnikach z ustalonego ciała $\mathbb{K}$ (ogólniej: z danej dziedziny całkowitości R) jest dziedziną całkowitości.

(D) Ustalmy $n \in \mathbb{N}$; określmy w zbiorze $\overline{0, n-1} = \{0, 1, \ldots, n-1\}$ następujące działania $\underset{n}{\oplus}$, $\underset{n}{\odot}$, zwane *dodawaniem i mnożeniem modulo n*:

$$a \underset{n}{\oplus} b := \text{ reszta z dzielania } a+b \text{ przez } n,$$
$$a \underset{n}{\odot} b := \text{ reszta z dzielania } ab \text{ przez } n.$$

Okazuje się, że aksjomaty 1°...7° są wtedy spełnione, więc otrzymujemy w ten sposób pierścień (przemienny, z jedynką); oznacza się go symbolem $\mathbb{Z}_n$ i nazywa *pierścieniem reszt modulo n*.

Zauważmy, że jeśli n nie jest liczbą pierwszą, tzn. $n = kl$, gdzie $k, l \in \overline{1, n-1}$, to aksjomat 8° nie jest spełniony: $k \underset{n}{\odot} l = 0$. Łatwo też pokazać, że na odwrót, jeśli n jest liczbą pierwszą, to spełniony jest aksjomat 8°; zatem $\mathbb{Z}_n$ *jest dziedziną całkowitości* $\Leftrightarrow$ n *jest pierwsza.*

(E) Zbiór $\mathbb{Z}[i] := \{a + ib : \ a, b \in \mathbb{Z}\} = \{z \in \mathbb{C} : \mathrm{Re}\, z, \mathrm{Im}\, z \in \mathbb{Z}\}$, wyposażony w 'zwykłe' działania (tzn. traktowany jako *podpierścień* pierścienia $\mathbb{C}$), jest dziedziną całkowitości. Nazywa się go *pierścieniem Gaussa*. W taki sam sposób można zdefiniować np. pierścień (a nawet dziedzinę całkowitości) $\mathbb{Z}[\sqrt{2}] := \{a + b\sqrt{2} : \ a, b \in \mathbb{Z}\}$.

(F) Pora na przykłady pierścieni nieprzemiennych: Bardzo ważny i interesujący jest *pierścień kwaternionów*. Przypomnijmy, że ciało $\mathbb{C}$ można skonstruować jako zbiór wyrażeń postaci $x+iy$, gdzie $x, y \in \mathbb{R}$, na których operacje algebraiczne wykonujemy w 'zwykły sposób', przy czym obowiązuje umowa, że $\forall y \in \mathbb{R} : yi = iy$ oraz $i^2 = -1$. Analogicznie *kwaterniony* można zdefiniować jako *formalne wyrażenia* postaci $\boldsymbol{x} = x_0\mathbf{1} + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$, gdzie $x_0, \ldots, x_3 \in \mathbb{R}$ [nietrudno się domyślić, że chodzi tu o swoisty zapis czwórki $(x_0, x_1, x_2, x_3) \in \mathbb{R}^4$]; zwykle piszemy x_0 zamiast $x_0\mathbf{1}$. Zbiór $\mathbb{H}$ wszystkich tych kwaternionów staje się pierścieniem, jeśli operacje dodawania i mnożenia określimy 'zwyczajnie'([7]), dołączając: (1) reguły przemienności: $\forall x \in \mathbb{R} : \ x\boldsymbol{i} = \boldsymbol{i}x$, $x\boldsymbol{j} = \boldsymbol{j}x$, $x\boldsymbol{k} = \boldsymbol{k}x$, (2) własność $\mathbf{1}$ jako jedynki: $\forall \boldsymbol{x} \in \mathbb{H} : \boldsymbol{x}\mathbf{1} = \mathbf{1}\boldsymbol{x} = \boldsymbol{x}$, oraz (3) swoistą 'tabelkę mnożenia':

	$\boldsymbol{i}$	$\boldsymbol{j}$	$\boldsymbol{k}$
$\boldsymbol{i}$	$-\mathbf{1}$	$\boldsymbol{k}$	$-\boldsymbol{j}$
$\boldsymbol{j}$	$-\boldsymbol{k}$	$-\mathbf{1}$	$\boldsymbol{i}$
$\boldsymbol{k}$	$\boldsymbol{j}$	$-\boldsymbol{i}$	$-\mathbf{1}$

Przykład. Jeśli $\boldsymbol{x} = 3+\boldsymbol{i}-2\boldsymbol{j}+5\boldsymbol{k}$, $\boldsymbol{y} = 2-3\boldsymbol{i}+4\boldsymbol{j}-\boldsymbol{k}$, to $\boldsymbol{x}+\boldsymbol{y} = 5-2\boldsymbol{i}+2\boldsymbol{j}+4\boldsymbol{k}$,

$\boldsymbol{xy} = (6+3+8+5)+(-9+2+2-20)\boldsymbol{i}+(12+1-4-15)\boldsymbol{j}+(-3+4-6+10)\boldsymbol{k} =$
$= 22-25\boldsymbol{i}-6\boldsymbol{j}+5\boldsymbol{k}$,

$\boldsymbol{yx} = (6+3+8+5)+(2-9+20-2)\boldsymbol{i}+(-4+15+12-1)\boldsymbol{j}+(10+6-4-3)\boldsymbol{k} =$
$= 22+11\boldsymbol{i}+22\boldsymbol{j}+9\boldsymbol{k}$.

Pierścień kwaternionów ma pewną ważną i rzadką własność: każdy niezerowy element $\boldsymbol{x} \in \mathbb{H}$ ma odwrotność: $\exists \boldsymbol{y} \in \mathbb{H} : \ \boldsymbol{xy} = \boldsymbol{yx} = \mathbf{1}$. Istotnie, dla dowodu wystarczy sprawdzić, że jeśli *sprzężenie* kwaternionu $\boldsymbol{x} = x_0\mathbf{1} + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$ określimy jako $\overline{\boldsymbol{x}} := x_0\mathbf{1} - x_1\boldsymbol{i} - x_2\boldsymbol{j} - x_3\boldsymbol{k}$, to $\boldsymbol{x}\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}}\boldsymbol{x} = {x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2 > 0$, skąd widać, że element $\boldsymbol{y} := (\boldsymbol{x}\overline{\boldsymbol{x}})^{-1}\overline{\boldsymbol{x}} \in \mathbb{H}$ jest odwrotnością $\boldsymbol{x}$.

(G) *Pierścień macierzy* $\mathrm{M}_2(R)$ wymiaru 2×2 o wyrazach z danego pierścienia R. Jego elementami są 2×2-*macierze*, tj. tablice postaci $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, gdzie $a, b, c, d \in R$; działania na takich macierzach określamy następującymi wzorami

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} := \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}, \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} := \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix}.$$

Na przykład $\begin{bmatrix} 1 & 3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 14 \\ 10 & -3 \end{bmatrix}$, $\begin{bmatrix} -2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -17 & 14 \\ -8 & 15 \end{bmatrix}$.

Jeśli R ma jedynkę, to $\mathrm{M}_2(R)$ także, jest nią $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (sprawdzić!). Zauważmy, że pierścień $\mathrm{M}_2(R)$ ma nietrywialne dzielniki zera, na przykład:

[7] Co oznacza, że dla zapewnienia rozdzielności mnożenia względem dodawania oraz łączności mnożenia obowiązują dwie (często traktowane domyślnie) umowy: (a) iloczyn $\boldsymbol{x} = x_0\mathbf{1} + x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$ i $\boldsymbol{y} = y_0\mathbf{1} + y_1\boldsymbol{i} + y_2\boldsymbol{j} + y_3\boldsymbol{k}$ jest sumą 16 iloczynów $(x_0\mathbf{1})(y_0\mathbf{1})$, $(x_0\mathbf{1})(y_1\boldsymbol{i}), \ldots, (x_3\boldsymbol{k})(y_3\boldsymbol{k})$; oraz (b) $(x_1\boldsymbol{i})(\boldsymbol{j}y_2) = \big(x_1(\boldsymbol{ij})\big)x_2$ itp.

$$\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}, \qquad \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}.$$

74. Rozważmy pierścień wielomianów $P = R[\,\cdot\,]$, gdzie R jest pierścieniem przemiennym z jedynką. Sprawdzić, że każdy z następujących warunków:

(1) $\forall \lambda \in R \setminus \{0\} : W \in P : \deg(\lambda W) = \deg W$); (2) P jest dziedziną całkowitości; (3) $\forall W, V \in P : \deg(W \cdot V) = \deg W + \deg V$

jest równoważny temu, że R jest dziedziną całkowitości, tj. nie ma dzielników zera.

W dalszym ciągu tego rozdziału termin "pierścień" będzie oznaczać "pierścień przemienny z jedynką"; takie pierścienie scharakteryzowane są zestawem aksjomatów ciała, pomniejszonym o dwa aksjomaty: o istnieniu odwrotności elementu $\neq 0$, oraz o nietrywialności ($0 \neq 1$).

2.3 Relacja podzielności

Zgodnie z naszą umową zakładamy, że pierścień R jest przemienny i ma jedynkę.

75. **Definicja** (relacja podzielności w pierścieniu R). Jeśli $a, b \in R$, to $\boxed{a \mid b \overset{\text{def}}{\iff} \exists c \in R : b = ac}$; taki fakt wysławiamy, mówiąc że a jest *dzielnikiem* b, lub że b jest *podzielny* przez a, lub że b jest *krotnością* a.

76. **Fakt** (oczywiste własności relacji '|' w pierścieniu przemiennym, z 1):

$\forall a \in R : a \mid a$ (zwrotność);
$\forall a, b, c \in R : (a \mid b,\ b \mid c) \ \Rightarrow\ a \mid c$ (przechodniość);
$(a_1 \mid b_1,\ a_2 \mid b_2) \ \Rightarrow\ a_1 a_2 \mid b_1 b_2$ (multiplikatywność),
w szczególności: jeśli $a \mid b$, to $ac \mid bc$, więc tym bardziej $a \mid bc$;
$(a \mid b_1, a \mid b_2) \ \Rightarrow\ a \mid (b_1 + b_2)$ (suma krotności a jest krotnością a);
$\forall a \in R :\ 1 \mid a \mid 0$ (1 jest elementem minimalnym, 0 — maksymalnym);
$ac \mid bc \ \Rightarrow\ a \mid b$, jeśli R jest dziedziną całkowitości oraz $c \in R \setminus \{0\}$.

Oznaczmy $\mathrm{D}_R(b) := \{a \in R : a \mid b\}$; tam, gdzie to nie prowadzi do niejasności będziemy zamiast $\mathrm{D}_R(b)$ pisać $\mathrm{D}(b)$. Zatem $\mathrm{D}_R(0) = R$, $\mathrm{D}_{\mathbb{Z}}(6) = \{\pm 1, \pm 2, \pm 3, \pm 6\}$, $\mathrm{D}_{\mathbb{K}}(a) = \mathbb{K}^*$ jeśli $\mathbb{K}$ jest ciałem i $a \in \mathbb{K}^*$.

77. **Definicja** (relacja stowarzyszenia w pierścieniu R). Elementy $a, b \in R$ są *stowarzyszone w* R, $a \sim b$, jeśli $a \mid b$ oraz $b \mid a$; oczywiście jest to relacja równoważności w zbiorze R.

Z multiplikatywności relacji podzielności wynika multiplikatywność relacji $\sim$:
$$a_1 \sim b_1,\ a_2 \sim b_2 \ \Rightarrow\ a_1 a_2 \sim b_1 b_2.$$

78. **Fakt**. Niech R^* oznacza podzbiór elementów odwracalnych w R; wtedy

1° każdy $a \in R^*$ ma tylko jedną *odwrotność*, tj. $b \in R$, taki że $ab = 1$;
2° $a \sim 1 \iff a \mid 1 \iff a \in R^*$, czyli $R^* = [1]_\sim$ oraz $R^* = \mathrm{D}_R(1)$;
3° $(a \in R^*,\ b \mid a) \ \Rightarrow\ b \in R^*$;
4° Jeśli $a \sim a'$ oraz $b \sim b'$, to $a \mid b \iff a' \mid b'$;

5° $a \sim b \iff \exists c \in R^* : a = bc$, gdy R jest dziedziną całkowitości([8]).

Ad 1°: Jeśli $ab = ac = 1$, to $c = (ab)c = (ba)c = b(ac) = b$. *Ad* 2°: Wprost z definicji. *Ad* 3°: $b \mid a \mid 1$, więc z przechodniości $b \mid 1$, czyli $b \in R^*$. *Ad* 4°: $a \mid b \Rightarrow a' \mid a \mid b \mid b' \Rightarrow a' \mid b'$ wskutek przechodniości relacji '$\mid$'. *Ad* 5°: Jeśli $a \sim b$, to $\exists c, c' : a = bc,\ b = ac'$, a zatem $a = acc'$, tzn. $a(1 - cc') = 0$; przy $a \neq 0$ wynika stąd $1 - cc' = 0$ (brak nietrywialnych dzielników zera!), a więc $c \in R^*$; z kolei gdy $a = 0$, wtedy $a \mid b$ daje $b = 0$, więc $a = bc$ dla $c = 1$. Odwrotne wynikanie nie wymaga braku dzielników zera: $\big(a = bc,\ c \mid 1\big) \Rightarrow \big(b \mid a$ oraz $a = bc \mid b \cdot 1 = b\big)$.

79. **Uwaga terminologiczna**. Jeżeli $\preceq$ jest "relacją porządkującą" (tzn. relacją przechodnią i zwrotną) w zbiorze X, $A \subset X$ oraz $a_0 \in A$, to

$$a_0 \text{ jest } \textit{największy} \text{ w } A \overset{\text{def.}}{\iff} \forall a \in A : a \preceq a_0,$$
$$a_0 \text{ jest } \textit{maksymalny} \text{ w } A \overset{\text{def.}}{\iff} \forall a \in A : a_0 \preceq a \Rightarrow a \preceq a_0.$$

Podobnie definiuje się pojęcia elementu najmniejszego i minimalnego. Jasne, że: każdy element największy jest maksymalny, a najmniejszy — minimalny; jeśli dwa elementy są największe w A, to są stowarzyszone (natomiast mogą być niestowarzyszone elementy maksymalne).

Przykłady. Niech $x \preceq y \Leftrightarrow x \mid y$. Dla $A = \overline{1,5}$ brak elementu największego, maksymalne są 3, 4 i 5, a minimalny (i najmniejszy) jest 1. Dla $A = \{2, 3, 6\}$ elementem maksymalnym (i największym) jest 6, najmniejszego brak, a minimalne są 2 i 3.

80. **Definicja**. Niech R — pierścień przemienny z jedynką. Element $d \in R$ nazywa się *największym wspólnym dzielnikiem* elementów $a, b \in R$, jeśli

(1) $d \mid a$ oraz $d \mid b$ (tzn. d jest wspólnym dzielnikiem a i b);
(2) d jest największym (w sensie relacji podzielności) wśród elementów mających własność (1), tzn. spełniony jest warunek

$$\forall c \in R : (c \mid a,\ c \mid b) \Rightarrow c \mid d.$$

Ogólniej, *największym wspólnym dzielnikiem* elementów $a_1, \ldots, a_n \in R$ nazywa się element $d \in R$, spełniający następujące warunki:

(1) $d \mid a_1, \ldots, d \mid a_n$, oraz
(2) $\forall c \in R : (c \mid a_1, \ldots, c \mid a_n) \Rightarrow c \mid d$.

Warunki (1),(2) można zapisać krócej: $\boxed{d \in \mathrm{D}(a_1, \ldots, a_n) \subset \mathrm{D}(d)}$, jeśli

$$\mathrm{D}(a_1, \ldots, a_n) := \mathrm{D}(a_1) \cap \ldots \cap \mathrm{D}(a_n) = (\text{wspólne dzielniki } a_1, \ldots, a_n).$$

Oznaczmy

$$\mathrm{NWD}(a_1, \ldots, a_n) := \{d \in R : d \text{ spełnia warunki (1),(2)}\}.$$

81. **Uwagi**. (A) Może się zdarzyć, że zbiór $\mathrm{NWD}(a, b)$ jest pusty, co oznacza, że $\mathrm{D}(a, b)$ nie ma elementu największego (w sensie '$\mid$'). Zobaczymy jednak, że NWD zawsze istnieje w niektórych ważnych pierścieniach R.

(B) Jeśli $d \in \mathrm{NWD}(a_1, \ldots, a_n)$ oraz $d \sim \tilde{d}$, to $\tilde{d} \in \mathrm{NWD}(a_1, \ldots, a_n)$.

[8] Jest to założenie istotne dla '$\Rightarrow$'. Oto kontrprzykład: $R :=$ (ciągłe funkcje $\mathbb{R} \to \mathbb{R}$), $a(x) := \max\big(|x| - 1, 0\big)$, $b(x) := a(x)\,\mathrm{sgn}(x)$; wtedy $a \sim b$, lecz jeśli $a = bc$, to $c(\pm 2) = \pm 2$, więc z ciągłości c mamy $\exists x_0 \in \mathbb{R} : c(x_0) = 0$ (własność Darboux); zatem $c \notin R^*$.

Odwrotnie, jeśli $d, \tilde{d} \in \mathrm{NWD}(a_1, \ldots, a_n)$, to oczywiście $d \sim \tilde{d}$. Zatem $\mathrm{NWD}(a_1, \ldots, a_n)$ *jest klasą relacji stowarzyszenia w* R; mówiąc to samo bardziej tradycyjnym językiem: NWD danych elementów, jeśli istnieje, określony jest z dokładnością do relacji stowarzyszenia.

(C) Zamiast $d \in \mathrm{NWD}(a_1, \ldots, a_n)$ pisze się często $d = \mathrm{NWD}(a_1, \ldots, a_n)$.

Jest to reliktem czasów 'przedteoriomnogościowych', na równi z tradycyjną notacją typu $z = \sqrt[3]{2-2i}$ (zamiast $z \in \sqrt[3]{2-2i}$) czy $\varphi = \arg z$ (zamiast $\varphi \in \arg z$). Należy więc pamiętać, że zapis $d = \mathrm{NWD}(a, b)$ określa d tylko z dokładnością do relacji '$|$'.

(D) Dla niektórych ważnych pierścieni R wybór $d \in \mathrm{NWD}(a_1, \ldots, a_n)$ można ujednoznacznić nałożeniem na d dodatkowego warunku, np.

— dla $R = \mathbb{Z}$ warunku $d \geqslant 0$;
— dla $R = \mathbb{K}[\,\cdot\,]$ warunku 'unormowania': wielomian jest *unormowany*, jeśli jego wyraz kierunkowy (czyli współczynnik przy najwyższej potędze zmiennej) jest równy 1.

W takim przypadku zwykle pisze się $d = ((a_1, \ldots, a_n))$, np. dla $R = \mathbb{Z}$:

$$\mathrm{D}(12, 20) = \{\pm 1, \pm 2, \pm 4\},\ \mathrm{NWD}(12, 20) = \{-4, 4\},\ ((12, 20)) = 4.$$

82. W dalszym ciągu stosując symbol $((a_1, \ldots, a_n))$ będziemy zakładać, że dane jest pewne odwzorowanie (*funkcja wyboru*) $f : R/_\sim \to R$, które klasie relacji $\sim$ (stowarzyszenia) przyporządkowuje pewien element tej klasy. Np. dla $R = \mathbb{Z}$ przyjmiemy $f\{-n, n\} := n$ dla $n \in \mathbb{Z}_+$; dla $R = \mathbb{K}[\,\cdot\,]$ z kolei $f(K)$ będzie tym (jedynym) wielomianem z klasy K, którego współczynnik przy maksymalnej potędze zmiennej jest równy 1. Symbol $((a_1, \ldots, a_n))$ będzie oznaczać f-obraz zbioru $\mathrm{NWD}(a_1, \ldots, a_n)$.

83. **Fakt** (podstawowe własności NWD w pierścieniach przemienych z 1).

1° $a_1 \mid b_1, \ldots, a_n \mid b_n \Rightarrow ((a_1, \ldots, a_n)) \mid ((b_1, \ldots, b_n))$ (monotoniczność);
2° $((a_1, a_2, \ldots, a_n)) = ((a_{i_1}, a_{i_2}, \ldots, a_{i_n}))$ (przemienność);
3° $\big(c \,|a_1, \ldots, c \mid a_n\big) \iff c \mid ((a_1, \ldots, a_n))$;
4° jeśli $a = qb + r$, to $((a, b)) = ((b, r))$ ([9]);
5° $(((a_1, \ldots, a_m)), b_1, \ldots, b_n)) = ((a_1, \ldots, a_m, b_1, \ldots, b_m))$ (łączność);
zatem np. $((a_1, (((a_2, a_3)), ((a_4, a_5)), a_6)), a_7)) = ((a_1, \ldots, a_7))$.

*Ad*3° $\boxed{\Leftarrow}$ $d := ((a_1, \ldots, a_n))$, wtedy $d \mid a_k$, więc $c \mid d \mid a_k$; $\boxed{\Rightarrow}$ z definicji NWD.

*Ad*4° $((a, b))$ dzieli a i b, więc także $r = a - qb$; zatem z 3° dostajemy $((a, b)) \mid ((b, r))$. Tak samo widać, że $((b, r))$ dzieli $a = qb + r$, więc $((b, r)) \mid ((a, b))$.

*Ad*5° Niech $d = ((a_1, \ldots, a_m, b_1, \ldots, b_n))$, $\tilde{d} = (((a_1, \ldots, a_m)), b_1, \ldots, b_n))$; należy wykazać, że $d \sim \tilde{d}$. Otóż korzystając z własności 3° widzimy, że:
(1) $\tilde{d} \mid ((a_1, \ldots, a_m))$, a więc $\tilde{d} \mid a_1, \ldots, \tilde{d} \mid a_m$; ponadto $\tilde{d} \mid b_1, \ldots, \tilde{d} \mid b_n$, skąd
$\tilde{d} \mid ((a_1, \ldots, a_m, b_1, \ldots, b_n))$, czyli $\tilde{d} \mid d$.
(2) $d \mid ((a_1, \ldots, a_m))$, gdyż $d \mid a_1, \ldots, d \mid a_m$; ponadto $d \mid b_k$, a zatem $d \mid \tilde{d}$.

84. **Ćwiczenie**. Przeformułować powyższe własności, używając zamiast symbolu $((\ldots))$ bardziej naturalnego obiektu jakim jest $\mathrm{NWD}(\ldots)$.

[9] Czyli nie zmienimy NWD, jeśli dodamy do jednego z elementów krotność drugiego.

85. **Definicja**. Element $w \in R$ jest *najmniejszą wspólną krotnością* elementów $a_1, \ldots, a_n \in R$, jeśli

(1) $a_1 \mid w, \ldots, a_n \mid w$ (tzn. w jest wspólną krotnością $a_1, \ldots, a_n$);
(2) $\forall v \in R : (a_1 \mid v, \ldots, a_n \mid v) \Rightarrow w \mid v$ (czyli każda wspólna krotność a_i jest krotnością w).

Piszemy wtedy $w \in \mathrm{NWK}(a_1, \ldots, a_n)$; zbiór $\mathrm{NWK}(a_1, \ldots, a_n)$, tak jak $\mathrm{NWD}(a_1, \ldots, a_n)$, jest jedną z klas relacji stowarzyszenia (lub zbiorem pustym). Zauważmy, że warunki (1),(2) można zapisać nieco krócej:

$$w \in a_1R \cap \ldots \cap a_nR \subset wR,$$

przy czym $wR := \{wx : x \in R\}$ jest zbiorem wszystkich krotności w, a więc $a_1R \cap \ldots \cap a_nR$ oznacza zbiór wspólnych krotności $a_1, \ldots, a_n$.

W przypadku wspomnianym wyżej w punkcie 82 wyróżnioną NWK oznacza się niekiedy symbolem $w = [\![a_1, \ldots, a_n]\!]$. Zatem np. dla $R = \mathbb{Z}$ mamy: $\mathrm{NWK}(12, -20) = \{-60, 60\}$, $[\![12, -20]\!] = 60$.

Uwaga. Anglojęzycznymi odpowiednikami skrótów NWD i NWK są LCM (*Least Common Multiple*) i GCD (*Greatest Common Divisor*).

86. **Przykład** (w którym R jest dziedziną całkowitości, a $\mathrm{NWD}(a, b) = \emptyset$). Niech $R = \mathbb{Z}[i\sqrt{6}]$ będzie zbiorem liczb postaci $a + bi\sqrt{6}$, gdzie $a, b \in \mathbb{Z}$; zbiór ten, jak łatwo sprawdzić, jest zamknięty względem dodawania, odejmowania i mnożenia, więc (ze zwykłymi operacjami dodawania i mnożenia) jest pierścieniem — podpierścieniem pierścienia $\mathbb{C}$. Jasne, że R, tak jak $\mathbb{C}$, jest dziedziną całkowitości.

Znajdźmy wspólne dzielniki elementów $a = 6$, $b = 10i\sqrt{6}$: Niech $d = d_1 + d_2i\sqrt{6}$; jeśli $d \in \mathrm{D}_R(a)$, to $\exists c = c_1 + c_2i\sqrt{6} \in R : a = cd$, więc $|a|^2 = |c|^2|d|^2$, przy czym $|c|^2 = c_1^2 + 6c_2^2 \in \mathbb{Z}$; zatem $|d|^2 \in \mathrm{D}_{\mathbb{Z}}(|a|^2)$, czyli $d_1^2 + 6d_2^2 \in \mathrm{D}_{\mathbb{Z}}(36)$; w szczególności $|d_1| \in \overline{0,6}$, $|d_2| \in \overline{0,2}$. Mamy więc $13 \cdot 5 = 65$ 'elementów podejrzanych'; jednak skoro $\frac{a}{d} = \frac{6}{d} = \frac{6}{d_1^2 + 6d_2^2}(d_1 - d_2i\sqrt{6})$, to $d \in \mathrm{D}_R(a) \Leftrightarrow d_1^2 + 6d_2^2 \in \mathrm{D}_{\mathbb{Z}}(6d_1, 6d_2)$, więc wystarczy sprawdzać tylko d, mające $d_1, d_2 \geqslant 0$; jest ich $7 \cdot 3 = 21$. Dostajemy

$$\mathrm{D}_R(a) = \{\pm 1, \pm 2, \pm 3, \pm 6 \pm i\sqrt{6}\}.$$

Łatwo stąd sprawdzić, że spośród elementów $\mathrm{D}(a)$ dzielnikami $b = 10i\sqrt{6}$ są tylko ± 1, ± 2, $\pm i\sqrt{6}$, skoro zaś $2 \not| i\sqrt{6}$ oraz $i\sqrt{6} \not| 2$, to wynika stąd, że $\mathrm{NWD}(a, b) = \emptyset$.

A oto inna osobliwość: W podobny sposób można sprawdzić, że jeśli $\delta := i\sqrt{6}$, to

$$\mathrm{D}_R(10) = \{\pm 1, \pm 2, \pm 5, \pm 10, \pm(2+\delta), \pm(2-\delta)\}, \qquad \mathrm{D}_R(\delta) = \{\pm 1, \pm\delta\},$$

natomiast $\mathrm{D}_R(10\delta) = \{\ \pm 1,\ \pm 2,\ \ \pm 5,\ \ \pm 10,\ \ \pm(2+\delta),\ \ \pm(2-\delta),$
$\pm\delta,\ \pm 2\delta,\ \pm 5\delta,\ \pm 10\delta,\ \pm(2+\delta)\delta,\ \pm(2-\delta)\delta,$
$\pm\ 2(2+2\delta),\ \pm 2(2-2\delta),\ \pm(3+\delta),\ \pm(3-\delta)\}$;

zatem 10δ ma dzielniki nie będące postaci xy, gdzie $x \in \mathrm{D}_R(10)$, $y \in \mathrm{D}_R(\delta)$.

2.4 Podpierścienie i ideały

87. **Definicja**. Podzbiór $S \subset R$ jest *podpierścieniem* pierścienia R, jeśli dla każdej pary $a, b \in S$ elementy $a + b$, $a - b$, ab należą do S; wtedy S jest pierścieniem, w którym działania (dodawanie i mnożenie) są

'działaniami w R, obciętymi do S'.

Nie warto wprowadzać odrębnych symboli dodawania i mnożenia dla podpierścienia $S \subset R$: suma elementów $a, b \in S$ jest «w sensie S» taka sama, jak «w sensie R» itd.

Podzbiór $J \subset R$ nazywa się *ideałem* pierścienia R, jeśli

$$\forall a, b \in J : a \pm b \in J \quad \text{oraz} \quad \forall a \in J, r \in R : ar \in J;$$

oczywiście J jest wtedy podpierścieniem pierścienia R.([10])

88. **Przykłady**. 1 Pierścień $\mathbb{Z}$ jest podpierścieniem pierścienia $\mathbb{R}$ (a także $\mathbb{C}$).

2 Jeśli w R każdy różny od 0 element ma odwrotność (w szczególności jeśli $R = \mathbb{K}$ jest ciałem), to pierścień R zawiera tylko dwa ideały: $\{0\}$ i R (ćwiczenie).

3 Jeśli $a \in R$ jest ustalony, to podzbiór $aR := \{ar : r \in R\}$ jest, jak łatwo sprawdzić, ideałem (najmniejszym ideałem zawierającym element a, a więc *generowanym przez* a); ideały takiej postaci nazywa się *ideałami głównymi*.

W pierścieniu $R = \mathbb{Z}$ ideał $6\mathbb{Z} = \{0, \pm 6, \pm 12, \pm 18, \ldots\}$ składa się z krotności 6, tzn. z liczb całkowitych podzielnych przez 6; podobnie $10\mathbb{Z} = \{0, \pm 10, \pm 20, \pm 30, \ldots\}$; zauważmy, że $2n = 12n + (-10)n \in 6\mathbb{Z} + 10\mathbb{Z}$ dla $n \in \mathbb{Z}$, a zatem $6\mathbb{Z} + 10\mathbb{Z} = 2\mathbb{Z}$. Przykład ten pokazuje, że na ogół ideał $aR + bR$ jest większy od $(a + b)R$.

Zauważmy, że każdy podpierścień pierścienia $\mathbb{Z}$ jest jego ideałem (gdyż mnożenie przez liczbę całkowitą jest kilkakrotnym dodawaniem); z kolei można pokazać, że każdy ideał $J \subset \mathbb{Z}$ jest równy $n\mathbb{Z}$, gdzie n jest najmniejszym z elementów $J \cap \mathbb{Z}_+$.

4 Ogólniej, dla ustalonych $a_1, \ldots, a_n \in R$ podzbiór $J := a_1R + \ldots + a_nR \subset R$, określony jako $\{a_1r_1 + \ldots + a_nr_n : r_k \in R\}$, jest ideałem; jest to oczywiście najmniejszy z ideałów zawierających wszystkie elementy $a_1, \ldots, a_n$, dlatego nazywa się go *ideałem generowanym przez elementy* $a_1, \ldots, a_n$.

5 Dla ustalonego $u \in \mathbb{C}$ zbiór $\mathbb{Z}[u] := \{a_0 + a_1u + \ldots + a_nu^n : n \in \mathbb{Z}_+,\ a_i \in \mathbb{Z}\}$ jest najmniejszym spośród podpierścieni ciała $\mathbb{C}$, zawierających u oraz 1; oczywiście $\mathbb{Z}[u] = \big\{a(u) : a \in \mathbb{Z}[\cdot]\big\}$. Zauważmy też, że np. $\mathbb{Z}[\sqrt{5}] = \{\alpha + \beta\sqrt{5} : \alpha, \beta \in \mathbb{Z}\}$, $\mathbb{Z}[\sqrt[3]{2}] = \{\alpha + \beta\sqrt[3]{2} + \gamma\sqrt[3]{4} : \alpha, \beta, \gamma \in \mathbb{Z}\}$, $\mathbb{Z}[i] = \{\alpha + i\beta : \alpha, \beta \in \mathbb{Z}\}$. Łatwo też np. sprawdzić, że podzbiór $\{\alpha + \beta\sqrt{5} : \alpha, \beta \in \mathbb{Z},\ 5 \mid \alpha\}$ jest najmniejszym podpierścieniem pierścienia $\mathbb{Z}[\sqrt{5}]$ (a więc i ciała $\mathbb{C}$), zawierającym element $\sqrt{5}$.

6 Niech teraz liczba $u \in \mathbb{C}$ będzie pierwiastkiem ustalonego wielomianu $b \in \mathbb{Z}[\cdot]$; jeśli jest = 1 współczynnik przy najwyższej potędze: $b(x) = b_0 + \ldots + b_{n-1}x^{n-1} + x^n$, to dla $a \in \mathbb{Z}[\cdot]$ iloraz i reszta z dzielenia a przez b, mają nie tylko wymierne, ale nawet całkowite współczynniki: $\forall a : \exists q, r \in \mathbb{Z}[\cdot] :\ a = qb + r,\ \deg r < n = \deg b$; skoro przy tym $b(u) = 0$, to $a(u) = r(u)$; zatem

$$\mathbb{Z}[u] = \{r_0 + r_1u + \ldots + r_{n-1}u^{n-1} : r_0, \ldots, r_{n-1} \in \mathbb{Z}\}.$$

7 Łatwe ćwiczenie: $\mathbb{Z}\left[\frac{1}{\sqrt{2}}\right] = \left\{\frac{\alpha+\beta\sqrt{2}}{2^n} :\ \alpha, \beta \in \mathbb{Z},\ n \in \mathbb{Z}_+\right\} = \bigcup_{n=0}^{\infty} \frac{1}{2^n}\mathbb{Z}[\sqrt{2}]$.

8 Jeśli ułamek $\frac{L}{M} \in \mathbb{Q}$ jest nieskracalny, to $\frac{1}{M^n} \in \mathbb{Z}\left[\frac{L}{M}\right]$ dla $n \in \mathbb{N}$; istotnie, dzięki temu, że $1 \in \mathrm{NWD}(L^n, M^n)$ mamy $\exists x, y \in \mathbb{Z} : 1 = L^nx + M^ny$ (zob. dalej punkt 95), a więc $\frac{1}{M^n} = \left(\frac{L}{N}\right)^n x + y$. Wynika stąd łatwo, że

[10] Dla pierścieni nieprzemiennych rozróżnia się trzy rodzaje ideałów: ideały *prawostronne* (zdefiniowane warunkami $\forall a, b, r : a \pm b \in J,\ ar \in J$), *lewostronne* ($a \pm b \in J,\ ra \in J$), oraz *obustronne* ($a \pm b \in J,\ ar \in J,\ ra \in J$).

$$\mathbb{Z}\left[\tfrac{L}{M}\right] = \bigcup_{n=1}^{\infty} \tfrac{1}{M^n}\mathbb{Z}.$$

9 Niech liczba $v \in \mathbb{C}$ będzie pierwiastkiem ustalonego wielomianu $c(\,\cdot\,) \in \mathbb{Z}[\,\cdot\,]$. Dla $n := \deg c(\,\cdot\,)$ i $M := c_n \in \mathbb{Z}$ wielomian $b(x) := M^{n-1} \cdot c\big(\frac{1}{M}x\big)$ należy do $\mathbb{Z}[\,\cdot\,]$, jest unormowany ($b_n = 1$) i zeruje liczbę $u = Mv$: $b(u) = 0$; zgodnie z 6 mamy stąd $\mathbb{Z}[u] = \{r_0 + r_1 u + \ldots + r_{n-1}u^{n-1} : r_0, \ldots, r_{n-1} \in \mathbb{Z}\}$. Natomiast $\mathbb{Z}[v] = \mathbb{Z}[\frac{1}{M}u] \subset \bigcup_{n=1}^{\infty} \frac{1}{M^n}\mathbb{Z}[u]$; niekiedy, np. dla $v = \frac{1}{\sqrt{2}}$, zamiast '$\subset$' jest równość, lecz na ogół dokładny opis $\mathbb{Z}[v]$ może być zadaniem niełatwym (nawet dla $n = 2$).

89. **Ćwiczenie.** $\quad aR \subset bR \iff b \mid a, \qquad aR = bR \iff a \sim b.$

90. **Fakt** (*operacje na ideałach*). Przecięcie dowolnej (nawet nieskończonej) liczby ideałów jest ideałem. Suma algebraiczna ideałów

$$J_1 + J_2 := \{a_1 + a_2 : a_1 \in J_1, a_2 \in J_2\} = \{a \in R : \exists a_k \in J_k : a = a_1 + a_2\}$$

jest ideałem. Suma mnogościowa $\bigcup_{n=1}^{\infty} J_n$ rosnącego (tj. wstępującego) ciągu ideałów $J_1 \subset J_2 \subset J_3 \subset \ldots$ jest ideałem.

2.5 Ideały główne

91. **Definicja**. Ideał $J \subset R$, dający się przedstawić w postaci $J = aR$ dla pewnego $a \in R$ (tzn. dający się 'wygenerować' jednym elementem pierścienia), nazywa się *ideałem głównym* w R. Pierścień R, będący dziedziną całkowitości, w którym każdy ideał jest główny, nazywa się *dziedziną ideałów głównych* (w skrócie: d.i.g.).

92. **Przykłady.**

(1) Pierścienie $\mathbb{Z}$ i $\mathbb{K}[\,\cdot\,]$ są d.i.g. Przekonamy się o tym w następnym paragrafie.

(2) $R := \mathbb{Z}[x]$ nie jest d.i.g., gdyż np. ideał J, złożony z wielomianów w, dla których $2 \mid w(0)$, nie jest główny. Istotnie, gdyby $J = cR$, wtedy $c(x) \mid 2$, $c(x) \mid x$, więc $c(x) = \pm 1$, skąd $J = R$, sprzeczność. Zauważmy też, że $J = 2R + xR$ oraz że wielomianu $1 \in \mathrm{NWD}(2, x)$ nie da się przedstawić w postaci $2p(x) + xq(x)$.

(3) $R := \mathbb{K}[x, y]$, wtedy ideał $J := xR + yR$, tzn. $J = \{w : w(0,0) = 0\}$, nie jest główny, bo wspólnym dzielnikiem $x, y \in J$ jest 1, zaś $J \neq 1R = R$.

93. **Fakt**. Jeśli R jest pierścieniem (przemiennym, z 1) oraz $a, b, d \in R$, to

$$aR + bR = dR \Leftarrow\Rightarrow \left\{ \begin{array}{c} d \mid a,\ d \mid b \\ \exists x, y \in R : d = ax + by \end{array} \right\} \overset{*}{\Rightarrow} d \in \mathrm{NWD}(a, b);$$

$\boxed{\Rightarrow}$ — oczywiste, np. $d \mid a$, gdyż $aR \subset aR + bR = dR$. $\boxed{\Leftarrow}$ Skoro $aR, bR \subset dR$, więc $aR + bR \subset dR$; z kolei $d = ax + by \in aR + bR$, więc także $dR \subset aR + bR$. Udowodnienie wynikania $\boxed{\overset{*}{\Rightarrow}}$ jest bardzo prostym ćwiczeniem.

94. **Fakt**. Jeśli R jest pierścieniem oraz $a, b, d, w \in R$, to

$$\text{(a)} \quad d \in \mathrm{NWD}(a, b) \iff \left(\begin{array}{c} dR \text{ jest najmniejszym z ideałów głó-} \\ \text{wnych zawierających ideał } aR + bR \end{array} \right),$$

(b) $\quad w \in \mathrm{NWK}(a,b) \iff wR = aR \cap bR$ ([11]).

(a) $\Big(dR \subset cR\Big) \Leftrightarrow c \mid d$ oraz $aR + bR \subset dR \Leftrightarrow \left\{\begin{matrix} aR \subset dR \\ bR \subset dR \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} d \mid a \\ d \mid b \end{matrix}\right\}$, skąd teza.

(b) Warunek $wR \subset aR \cap bR$, równoważny $w \in aR \cap bR$, oznacza, że $\left\{\begin{matrix} a \mid w \\ b \mid w \end{matrix}\right\}$, z kolei zawieranie $aR \cap bR \subset wR$ oznacza, że $\forall v \in R : \left\{\begin{matrix} a \mid v \\ b \mid v \end{matrix}\right\} \Rightarrow w \mid v$; stąd teza.

Natychmiastowym wnioskiem z tego faktu jest następujące podstawowe

95. **Twierdzenie**. Jeśli R jest dziedziną ideałów głównych i $a, b \in R$, to:

$\boxed{1}$ $\mathrm{NWD}(a,b) \neq \emptyset$ oraz $d \in \mathrm{NWD}(a,b) \iff dR = aR + bR$; zatem:
$\boxed{2}$ $\forall d \in \mathrm{NWD}(a,b) : \exists x, y \in R : d = ax + by$;
$\boxed{3}$ $\mathrm{NWK}(a,b) \neq \emptyset$ oraz $k \in \mathrm{NWK}(a,b) \iff kR = aR \cap bR$.

Uogólnienie:

96. **Twierdzenie**. Jeśli pierścień R jest d.i.g. oraz $a_1, \ldots, a_n \in R$, to:

$\boxed{1}$ $\mathrm{NWD}(a_1, \ldots, a_n) \neq \emptyset$ oraz

$$d \in \mathrm{NWD}(a_1, \ldots, a_n) \iff dR = a_1R + \ldots + a_nR;$$

$\boxed{2}$ $\forall d \in \mathrm{NWD}(a_1, \ldots, a_n) : \exists x_1, \ldots, x_n \in R : d = a_1x_1 + \ldots + a_nx_n$;

$\boxed{3}$ $\mathrm{NWK}(a_1, \ldots, a_n) \neq \emptyset$ oraz

$$w \in \mathrm{NWK}(a_1, \ldots, a_n) \iff wR = a_1R \cap \ldots \cap a_rR.$$

Bezpośredni i Elementarny Dowód:
Ad $\boxed{1}$,$\boxed{2}$: Zbiór $J := a_1R + \ldots + a_nR$ (złożony z elementów postaci $a_1x_1 + \ldots + a_nx_n$, $x_i \in R$) jest ideałem R; skoro R jest d.i.g., to J jest postaci $J = dR$, gdzie $d \in R$. Zatem $\exists x_i \in R : d = a_1x_1 + \ldots + a_nx_n$. Skoro $a_i \in a_iR \subset J = dR$, to $d \mid a_i$, czyli d jest wspólnym dzielnikiem a_i. Sprawdzimy, że jest to dzielnik największy: jeśli $c \mid a_i$ dla $i \in \overline{1,n}$, to także $c \mid a_1x_1 + \ldots + a_nx_n$, czyli $c \mid d$. Zatem $d \in \mathrm{NWD}(a_1, \ldots, a_n)$.
Ad $\boxed{3}$ $wR \subset I := a_1R \cap \ldots \cap a_nR \iff \forall k : \big(wR \subset a_kR$, tzn. $a_k \mid w\big)$. Z kolei $I \subset wR \iff \forall v \in R : \big(\forall k : v \in a_kR\big) \Rightarrow v \in wR$, tzn. $\forall v \in R : \big(\forall k : a_k \mid v\big) \Rightarrow w \mid v$. Zatem $wR \subset I \Leftrightarrow (1)$, $I \subset wR \Leftrightarrow (2)$, gdzie (1),(2) oznaczają warunki z definicji 85.

Poniższe własności $6^\circ \ldots 10^\circ$ stanowią uzupełnienie listy z punktu 83:

97. **Fakt** (dodatkowe własności największego wspólnego dzielnika w d.i.g.).

6° $\big((a,b)\big) \sim 1 \iff \exists p, q \in R : ap + bq = 1$;
7° $\big((av, bv)\big) \sim \big((a,b)\big)v$;
8° $\big((a,b_1)\big) \sim \big((a,b_2)\big) \sim 1 \Rightarrow \big((a,b_1b_2)\big) \sim 1$;
9° $(a \mid bc$ oraz $\big((a,b)\big) \sim 1) \Rightarrow a \mid c$;
10° $(a \mid c,\ b \mid c,\ \big((a,b)\big) \sim 1) \Rightarrow ab \mid c$.

Ad 6°: $\boxed{\Rightarrow}$ już było. $\boxed{\Leftarrow}$: Jeśli $ap + bq = 1$ i $d = \big((a,b)\big)$, to $d \mid \big((ap + bq)\big)$, tzn. $d \mid 1$.
Ad 7°: $P \mid L$ (w każdej dz. całkowitości), gdyż wskutek $P \mid av$, $P \mid bv$ mamy też $P \mid \big((av, bv)\big) = L$. Odwrotnie, skoro $\exists p, q : \big((a,b)\big) = ap + bq$, to $L \mid avp + bvq = dv$, skąd $L \mid avp + bvq = \big((a,b)\big)v = P$.

[11] $wR = aR \cap bR \iff wR$ jest największym z ideałów głównych zawartych w $aR \cap bR$. Dowód. $\boxed{\Rightarrow}$ jest oczywiste. $\boxed{\Leftarrow}$ $wR \subset J := aR \cap bR$ wprost z założenia o J. Odwrotnie, jeśli $v \in J$, to vR jest i.g. zawartym w J, więc $vR \subset wR$, skąd $v \in wR$; zatem $J \subset wR$.

Ad 8°: $1 = 1 \cdot 1 = (ap + b_1 q)(ar + b_2 s) = a(apr + pb_2 s + rb_1 q) + b_1 b_2 qs$.
Ad 9°: $\exists q : bc = aq$ i $\exists r, s : ar + bs = 1$, więc $c = (ar + bs)c = a(rc + qr)$, tj. $a \mid c$.
*Ad*10°: $\exists q : c = aq$; warunek $b \mid c$ oznacza więc $b \mid aq$, skoro zaś $((a, b)) \sim 1$, to wskutek własności 9° mamy $b \mid q$, więc $\exists r : q = br$; stąd $c = aq = abr$.

2.6 Pierścienie euklidesowe

98. **Definicja**. Pierścień R (przemienny, z 1) nazywa się *pierścieniem euklidesowym*, jeśli istnieje funkcja $\delta : R \setminus \{0\} \to \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ taka, że spełniony jest następujący 'warunek Euklidesa':

(E) dla każdej pary $a, b \in R$, gdzie $b \neq 0$, istnieją $q, r \in R$ takie, że

$$a = qb + r, \text{ przy czym } \Big(\text{jeśli } r \neq 0, \text{ to } \delta(r) < \delta(b)\Big).$$

Pierścień euklidesowy będący zarazem dziedziną całkowitości nazywa się *dziedziną euklidesową*([12]).

99. **Fakt**. Każdy pierścień euklidesowy jest dziedziną ideałów głównych.

Niech $J \subset R$ będzie ideałem; możemy założyć, że $J \neq \{0\}$. Niech n będzie najmniejszym elementem zbioru $\{\delta(b) : b \in J, b \neq 0\}$. Weźmy dowolne ustalone $b \in J$, takie że $\delta(b) = n$. Pokażemy że $\forall a \in J : \exists q \in R : a = qb$, a zatem $J \subset bR$, czyli $J = bR$ (jako że $bR \subset J$, gdyż $b \in J$ i J jest ideałem). Otóż z (E) mamy $a = bq + r$, przy czym albo $r = 0$ (co daje tezę), albo $\delta(r) < \delta(b)$, lecz wtedy byłoby $r = a - qb \in J$, przy czym $r \neq 0$ oraz $\delta(r) < n$, wbrew minimalności liczby n.

100. **Przykłady pierścieni euklidesowych**.

P1. Dla $R = \mathbb{Z}$ weźmy $\delta(a) := |a|$; sprawdzimy, że $q := E(\frac{a}{b})$ jest dobre: $r = a - bq = b\Big(\frac{a}{b} - E(\frac{a}{b})\Big) = bc$, gdzie $c \in [0, 1[$, więc $|r| < |b|$, tzn. $\delta(r) < \delta(b)$. Warto sprawdzić, że dla $q := 1 + E(\frac{a}{b})$ i $\frac{a}{b} \notin \mathbb{Z}$ też jest OK!

P2. Dla $R = \mathbb{Z}$ można też δ określić inaczej: $\hat{\delta}(a) := [\log_2 |a|]$. Sprawdzimy warunek (E): Mając liczby $\mathbb{Z} \ni a, b \neq 0$ określmy q jako element $\mathbb{Z}$ leżący najbliżej ułamka $\frac{a}{b}$; wtedy $\left|\frac{a}{b} - q\right| \leqslant \frac{1}{2}$, więc dla $r := a - bq \in Z$ mamy: $|r| = \left|\frac{a}{b} - q\right| \cdot |b| \leqslant \frac{1}{2}|b|$, więc $\log_2 |r| \leqslant -1 + \log_2 |b|$, QED.

Zauważmy, że $\forall a : \hat{\delta}(a) < |a| = \delta(a)$; okazuje się, że $\hat{\delta}$ jest najmniejszą z funkcji $\delta : \mathbb{Z} \setminus \{0\} \to \mathbb{Z}_+$, spełniających aksjomat (E).

P3. Dla $R = \mathbb{K}[x]$, gdzie $\mathbb{K}$ jest ciałem, weźmy $\delta(a) := \deg a$.

Twierdzenie (*o dzieleniu wielomianów*). Jeśli $a, b \in \mathbb{K}[x]$, przy czym $b \neq 0$, to istnieje dokładnie jedna para wielomianów $q, r \in \mathbb{K}[x]$,

[12] W wielu podręcznikach oprócz warunku (E) postuluje się jeszcze następujący warunek:

(E') Funkcja δ jest rosnąca w tym sensie, że $(a, b \neq 0,\ a \mid b) \Rightarrow \delta(a) \leqslant \delta(b)$.

Okazuje się, że warunek (E') jest właściwie zbędny: Jeśli para (E, δ) spełnia warunek (E), to istnieje funkcja $\delta' : R \setminus \{0\} \to \mathbb{Z}_+$ taka, że para (E, δ') spełnia oba warunki (E),(E'). Jako δ' można wziąć np. *najmniejszą* z funkcji δ, dla których spełniony jest warunek (E).

nazywanych *ilorazem* i *resztą* z dzielenia a przez b, takich że

$$a = qb + r \quad \text{oraz} \quad \deg r < \deg b.$$

Istnienie. Ustalmy b, a więc także $m = \deg b \geqslant 0$. Zastosujemy indukcję względem liczby $\deg a$. Jeśli $\deg a < m$, to para $q := 0$, $r := a$ ma żadane własności. Załóżmy teraz, że $\deg a = n \geqslant m$ oraz że twierdzenie jest prawdziwe dla wielomianów stopnia $< n$. Niech $a(x) = a_0 + \ldots + a_n x^n$, $b(x) = b_0 + \ldots + b_m x^m$, $a_n b_m \neq 0$. Zauważmy, że składniki z x^n w wielomianie $\tilde{a}(x) := a(x) - \frac{a_n}{b_m} x^{n-m} b(x)$ znoszą się, więc $\deg \tilde{a} < n$; wobec tego z założenia indukcyjnego $\tilde{a}$ ma rozkład $\tilde{a}(x) = \tilde{q}(x)b(x) + r(x)$, gdzie $\deg r < m$. Stąd

$$a(x) = \tilde{a}(x) + \tfrac{a_n}{b_m} x^{n-m} b(x) = \big(\tilde{q}(x) + \tfrac{a_n}{b_m} x^{n-m}\big) b(x) + r(x),$$

a więc dowód kroku indukcyjnego zakończymy, przyjmując $q(x) := \tilde{q}(x) + \frac{a_n}{b_m} x^{n-m}$.

Jednoznaczność. Jeśli $a = qb + r$ oraz $a = q_1 b + r_1$, to $(q_1 - q)b = r - r_1$; gdyby przy tym było $q_1 - q \neq 0$, wtedy $\deg\big((q_1 - q)b\big) = \deg(q_1 - q) + \deg b \geqslant \deg b$, podczas gdy warunki $\deg r, \deg r_1 < \deg b$ implikują $\deg(r - r_1) < \deg b$, sprzeczność.

P4. Dla $R = \mathbb{Z}[i]$ (pierścień Gaussa) weźmy $\delta(a) := |a|^2$; wtedy faktycznie $\delta(a) \in \mathbb{Z}_+$, gdyż $\delta(a) = {a_1}^2 + {a_2}^2$, gdzie $a_1 := \operatorname{Re} a$, $a_2 = \operatorname{Im} a$.

Sprawdźmy warunek (E): Dla $a, b \in R \setminus \{0\}$ można dobrać $q \in R$ tak, by różnica $c := \frac{a}{b} - q$ leżała w kwadracie $\Big\{z : \operatorname{Re} z, \operatorname{Im} z \in \Big[-\frac{1}{2}, \frac{1}{2}\Big]\Big\}$. Wtedy $|c|^2 \leqslant \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$; zarazem $r = b\Big(\frac{a}{b} - q\Big) = bc$, a zatem

$$\delta(r) = |r|^2 = |bc|^2 = |b|^2 |c|^2 \leqslant \tfrac{1}{2}|b|^2 < |b|^2 = \delta(b).$$

101. **Uwaga**. Z aksjomatu (E) nie wynika *jednoznaczność* operacji dzielenia z resztą, np. dla $R = \mathbb{Z}$ mamy $30 = 2 \cdot 13 + 4 = 3 \cdot 13 + (-9)$, więc są dwa możliwe wyniki dzielenia 30 przez 13: $q = 2, r = 4$ oraz $q = 3, r = -9$. W pierścieniu Gaussa są możliwe nawet 4 różne wyniki:

$$1 = 0\cdot(1+i)+1 = 1\cdot(1+i)+(-i) = (-i)(1+i)+i = (1-i)(1+i)+(-1).$$

Natomiast dzielenie z resztą jest operacją jednoznaczną w $R = \mathbb{K}[\,\cdot\,]$.

102. **Fakt**. Niech pierścień R ma własności (E) i (E') oraz $a, b \in R \setminus \{0\}$.

(1) Jeśli $a \mid b$, lecz $a \not\sim b$ (tzn. $b \not| a$), to $\delta(a) < \delta(b)$
(w tym sensie możemy o funkcji δ mówić, że jest «*ściśle rosnąca*»);
(2) Jeśli $a \mid b$ oraz $\delta(a) = \delta(b)$, to $a \sim b$;
(3) $\delta(b) = \delta(1) \iff b \in R^*$:= (elementy odwracalne pierścienia R).

Ad(1) Zastosujmy (E): $a = bq + r$, przy czym $\delta(r) < \delta(b)$ ($r \neq 0$, gdyż $b \not| a$); lecz $r = a - bq$ wraz z $a \mid b$ daje $a \mid r$, więc z (E') wynika $\delta(a) \leqslant \delta(r) < \delta(b)$.
Ad(2) Jest to oczywiście tylko inne sformułowanie (przez 'kontrapozycję') faktu (1).
Ad(3) Implikację '$\Rightarrow$' dostajemy z (2), biorąc $a = 1$; jeśli zaś $b \in R^*$, to $1 \mid b \mid 1$, co wraz z (E') daje $\delta(1) \leqslant \delta(b) \leqslant \delta(1)$, tzn. $\delta(b) = \delta(1)$.

2.7 Algorytm Euklidesa

103. **Twierdzenie**. Jeśli R jest pierścieniem euklidesowym, to dla każdej

pary $a, b \in R$, takiej że $b \neq 0$, NWD(a,b) zawiera «ostatnią niezerową resztę», tzn. element $r_k \in R$, uzyskany w następującym procesie:

$$\begin{array}{lll}
a = q_1 b + r_1 & r_1 \neq 0 & \delta(r_1) < \delta(b) \\
b = q_2 r_1 + r_2 & r_2 \neq 0 & \delta(r_2) < \delta(r_1) \\
r_1 = q_3 r_2 + r_3 & r_3 \neq 0 & \delta(r_3) < \delta(r_2) \\
\ldots\ldots & \ldots\ldots\ldots & \ldots \\
r_{k-2} = q_k r_{k-1} + r_k & r_k \neq 0 & \delta(r_k) < \delta(r_{k-1}) \\
r_{k-1} = q_{k+1} r_k & r_{k+1} = 0. &
\end{array}$$

Procedura musi się zakończyć po skończonej liczbie kroków, gdyż ściśle malejący ciąg liczb $\delta(r_k) \in \mathbb{Z}_+$ musi mieć skończoną długość.

Oznaczmy dla wygody $r_{-1} := a$, $r_0 := b$, wtedy $r_{-1}, r_0, \ldots, r_{k+1}$ są elementami spełniającymi relacje $\boxed{r_{j-2} = q_j r_{j-1} + r_j}$, $j \in \overline{1, k+1}$. Korzystając z własności $((qb + r, b)) = ((b, r))$, zob. 83, dostajemy $((r_{j-2}, r_{j-1})) = ((q_j r_{j-1} + r_j, r_{j-1})) = ((r_{j-1}, r_j))$, a więc $((a,b)) = ((r_{-1}, r_0)) = ((r_0, r_1)) = \ldots = ((r_k, r_{k+1})) = ((r_k, 0)) \sim r_k$.

104. **Fakt**. W oznaczeniach p. 103 niech $s_{-1}, s_0, s_1, \ldots, s_k \in R$ będzie dowolnym ciągiem spełniającym te same relacje, co r_j: $\boxed{s_{j-2} = q_j s_{j-1} + s_j}$ dla $j \in \overline{1,k}$. Wtedy $D_0 = -D_1 = D_2 = \ldots = (-1)^k D_k$, gdzie

$$D_j := \begin{vmatrix} r_{j-1} & r_j \\ s_{j-1} & s_j \end{vmatrix} = r_{j-1}s_j - r_j s_{j-1} \quad \text{dla} \quad j \in \overline{1,k}.$$

Istotnie, mamy $\begin{vmatrix} qr' + r & r' \\ qs' + s & s' \end{vmatrix} = (qr' + r)s' - r'(qs' + s) = rs' - r's = -\begin{vmatrix} r' & r \\ s' & s \end{vmatrix}$.

Jeśli więc weźmiemy $\begin{pmatrix} s_k := 0 \\ s_{k-1} := 1 \end{pmatrix}$ i kolejno wyliczymy $s_{k-2}, \ldots, s_0, s_{-1}$, to wtedy $\begin{vmatrix} a & b \\ s_{-1} & s_0 \end{vmatrix} = D_0 = \pm D_k = \pm \begin{vmatrix} r_{k-1} & r_k \\ 1 & 0 \end{vmatrix}$, czyli $as_0 - bs_{-1} = \mp r_k$; dostajemy w ten sposób wygodny algorytm znajdowania pary $x, y \in R$, spełniającej warunek $ax + by = ((a,b))$; por. punkt 95.

105. **Uwaga**. Wygodnie jest stosując powyższy algorytm posługiwać się jednym z dwu następujących 'formularzy' (poziomym lub pionowym):

a	b	r_1	r_2	$\ldots$	r_{k-1}		r_k
	q_1	q_2	q_3	$\ldots$	$\ldots$	q_{k+1}	
s_{-1}	s_0	s_1	s_2	$\ldots$	s_{k-1}		s_k

a		s_{-1}
	q_1	
b		s_0
	q_2	
r_1		s_1
	q_3	
r_2		s_2
	...	
...		...
	...	
r_{k-1}		s_{k-1}
	q_{k+1}	
r_k		s_k

Przykład. Dla pary (169,64) dostajemy tabelkę

169	64	41	23	18	5	3	2	1
	2	1	1	1	3	1	1	2
66	25	16	9	7	2	1	1	0

z której wynika, że $((169, 64)) = 1 = 169 \cdot 25 - 64 \cdot 66$.

2.8 Rozkład na czynniki pierwsze

106. **Definicja**. Niech R będzie dziedziną całkowitości. Element $p \in R$, taki że $p \neq 0$ oraz $p \not\sim 1$ ([13]), nazywa się:

$$\begin{aligned} \textit{rozkładalny} &\overset{\text{def}}{\iff} p \text{ ma dzielnik niestowarzyszony ani z } p, \text{ ani z } 1; \\ \textit{nierozkładalny} &\overset{\text{def}}{\iff} \text{każdy dzielnik } p \text{ jest stowarzyszony z } 1 \text{ lub z } p; \\ \textit{pierwszy} &\overset{\text{def}}{\iff} \text{zachodzi implikacja } \ p \mid ab \Rightarrow (p \mid a \text{ lub } p \mid b). \end{aligned}$$

Oczywiście element $p \in R$ jest rozkładalny $\Leftrightarrow$ ma nietrywialny rozkład na czynniki:

$$p = ab, \ a \not\sim 1, \ b \not\sim 1.$$

107. **Fakt**. (a) Każdy element pierwszy jest nierozkładalny.
(b) Jeśli R jest d.i.g., to każdy element nierozkładalny jest pierwszy.

Ad(a). Niech p — pierwszy i niech $a \mid p$; wtedy $\exists b : p = ab$. Skoro $p \mid ab$, to są dwie możliwości: $p \mid a$, wtedy $a \sim p$; lub $p \mid b$, wtedy $ab \mid b$, skąd $a \mid 1$, czyli $a \sim 1$.

Ad(b). Niech p — nierozkładalny i $p \mid ab$. Skoro $d := ((a,p))$ jest dzielnikiem p, to są dwie możliwości: 1° $d \sim p$, wtedy $p \mid d \mid a$; lub

$$2^\circ \ \ d \sim 1, \text{ wtedy } \exists x, y : 1 = ax + py, \text{ skąd } p \mid (\overbrace{ab}^{\text{krotność } p})x + pyb = b.$$

108. **Przykłady**. W każdym z poniższych przykładów R jest dziedziną euklidesową, a więc też d.i.g.; zatem *elementy pierwsze* to to samo, co *elementy nierozkładalne*.

(!) W $R = \mathbb{Z}$: (zbiór elementów pierwszych)$=\{\pm 2, \pm 3, \pm 5, \pm 7, \pm 11, \pm 13, \pm 17, \ldots\}$. Rozkładem na czyniki pierwsze elementu $a = 60$ jest $a = 2 \cdot 2 \cdot 3 \cdot 5$, ale mamy także inne (równoważne) rozkłady: $a = (-5) \cdot (-2) \cdot 2 \cdot 3 = (-2) \cdot (-3) \cdot (-5) \cdot (-2) = \ldots$.

(!!) W $\mathbb{C}[\,x\,]$ elementy pierwsze są to wielomiany stopnia 1; w $\mathbb{R}[\,x\,]$ elementy pierwsze są to wielomiany stopnia 1 oraz wielomiany stopnia 2 o wyróżnikach $\leqslant 0$. Wynika to wprost z twierdzenia o rozkładzie wielomianów w $\mathbb{C}[\,x\,]$ i w $\mathbb{R}[\,x\,]$, zob. 62.

(!!!) W $R = \mathbb{Q}[\,x\,]$ są elementy pierwsze dowolnego stopnia, można np. pokazać, że wielomiany $p_n(x) := x^n - 2$ są pierwsze dla $n \in \mathbb{N}$; wymaga to skorzystania np. z tzw. 'kryterium Eisensteina' (zob. np. Kostrykin, str. 168), nie wystarczy tu sam fakt braku pierwiastków wymiernych!

(!!!!) W pierścieniu Gaussa $R = \mathbb{Z}[i]$ element $2 \in R$ nie jest pierwszy, gdyż ma rozkład $2 = (1+i)(1-i)$; z kolei $1+i$ jest pierwszy w R, bo jeśli $1+i = ab$, to $2 = |ab|^2 = |a|^2|b|^2$, przy czym $|a|^2 \in \mathbb{N}$, $|b|^2 \in \mathbb{N}$, więc $|a|^2 = 1$ lub $|b|^2 = 1$. W taki sam sposób można sprawdzić np., że pierwsze są elementy $3, 7, 11, 19, 23$, a także $2+i$, $3+2i$, $4+i$, natomiast rozkładalne są $5 = (2+i)(2-i)$, $3+i = (1+i)(2-i)$, $13 = (3+2i)(3-2i)$, $17 = (4+i)(4-i)$ itd.

Ćwiczenie. Dowieść, że jeśli $p \in \mathbb{N}$ jest liczbą pierwszą, to

$$\left(\begin{array}{c} p \text{ jest elementem rozkła-} \\ \text{dalnym w pierścieniu } \mathbb{Z}[i] \end{array}\right) \iff \left(\begin{array}{c} p \text{ da się rozłożyć na sumę kwadratów:} \\ p = a^2 + b^2, \text{ gdzie } a, b \in \mathbb{N} = \{1, 2, \ldots\} \end{array}\right).$$

Dla dowodu '$\Rightarrow$' można zauważyć, że jeśli $z = a + ib \in \mathbb{Z}[i]$ jest dzielnikiem p, to (1) $a^2 + b^2$ dzieli pa i pb; (2) $\exists x, y \in \mathbb{Z} : ax + by = 1$, a więc $p = pax + pby$.

109. **Twierdzenie**. Każda d.i.g. jest *dziedziną z jednoznacznością rozkładu*, czyli taką dziedziną całkowitości R, że:

[13] Łatwo zauważyć, że takie elementy istnieją wtedy i tylko wtedy, gdy R nie jest ciałem.

(1) Każdy element $a \in R \setminus (\{0\} \cup [1]_\sim)$ ma *rozkład na czynniki nierozkładalne*, tzn. $a = p_1 p_2 \ldots p_m$, gdzie $p_i \in R$ są nierozkładalne;

(2) Powyższy rozkład jest *jednoznaczny*([14]) w następującym sensie: Jeśli

$$p_1 p_2 \ldots p_m \sim q_1 q_2 \ldots q_n, \text{ gdzie } p_i, q_j \in R \text{ — nierozkładalne,}$$

to $m = n$ oraz istnieje permutacja $i_1, \ldots, i_m$ ciągu $1, \ldots, m$, taka że

$$p_1 \sim q_{i_1}, \ldots, p_m \sim q_{i_m}.$$

Skoro R jest d.i.g., to *elementy nierozkładalne* to to samo, co *elementy pierwsze*.

Ad (1) Nazwijmy element $a \in R$ *zwykłym*, jeśli a ma rozkład $a = p_1 p_2 \ldots p_m$ na czynniki pierwsze $p_i \in R$. Zatem każdy element pierwszy jest zwykły ($m = 1$) oraz R jest sumą rozłączną $R = \{0\} \cup [1]_\sim \cup R_{\text{zw}} \cup R_{\text{nzw}}$, gdzie R_{zw} oznacza zbiór elementów *zwykłych*, a R_{nzw} — *niezwykłych*. Pokażemy, że dla d.i.g. zawsze $R_{\text{nzw}} = \emptyset$. Zauważmy najpierw, że każdy $a \in R_{\text{nzw}}$ ma rozkład $a = a_1 b_1$ taki, że $\begin{pmatrix} a_1 \in R_{\text{nzw}} \\ a \not| a_1 \end{pmatrix}$. Istotnie, a jest rozkładalny (gdyż elementy pierwsze należą do R_{zw}), a więc $a = a_1 b_1$, gdzie $a_1 \not\sim a$ oraz $b_1 \not\sim a$; skoro zaś iloczyn dwóch elementów R_{zw} należy do R_{zw} (z definicji!), to choć jeden z czynników a_1, b_1, powiedzmy a_1, należy do R_{nzw}.

Przypuśćmy teraz, że $R_{\text{nzw}} \neq \emptyset$ oraz niech $a \in R_{\text{nzw}}$; wtedy dzięki powyższej uwadze można indukcyjnie zbudować nieskończony ciąg $a = a_0, a_1, a_2, a_3, \ldots$ elementów R_{nzw} takich, że $\begin{pmatrix} a_{k+1} \mid a_k \\ a_k \not| a_{k+1} \end{pmatrix}$. Ciąg ideałów $a_k R$ jest rosnący, więc $J := \bigcup_{k=1}^{\infty} a_k R$ też jest ideałem; skoro zaś R jest d.i.g., to $\exists d \in R : J = dR$. Oczywiście $\forall k : a_k R \subset dR$, czyli $d \mid a_k$. Zarazem $d \in dR = J$, a zatem $\exists j : d \in a_j R$, tzn. $a_j \mid d$. Mamy więc $a_j \mid d \mid a_{j+1}$, co przeczy własności $\forall k : a_k \not| a_{k+1}$; zatem zbiór R_{nzw} jest pusty.

Ad (2). Zastosujemy indukcję względem liczby $k = \min(m, n) \geqslant 0$. Przyjmijmy tu naturalną umowę, że $p_1 \ldots p_m$ oznacza 1 w przypadku, gdy $m = 0$.

$\boxed{T_0}$ Mamy pokazać, że dla $m = 0$ nie może być $n \geqslant 1$; otóż gdyby $1 \sim q_1 \ldots q_n$, wtedy $q_1 \mid q_1 \ldots q_n \mid 1$, skąd $q_1 \sim 1$ wbrew założeniu, że element q_1 jest pierwszy.

$\boxed{T_k \Rightarrow T_{k+1}}$ Niech $p_1 \ldots p_m \sim q_1 \ldots q_n$, gdzie $\min(m, n) = k + 1$, $k \geqslant 0$; możemy dla ustalenia uwagi założyć, że $m = k + 1 \leqslant n$.
Skoro p_m jest pierwszy i $p_m \mid q_1 \ldots q_n$, to p_m dzieli choć jeden z czynników q_i, tzn. $\exists i_m \in \overline{1, n} : p_m \mid q_{i_m}$. Zarazem q_{i_m} jest nierozkładalny, więc jego dzielnik p_m jest stowarzyszony z q_{i_m} lub z 1; ta druga możliwość odpada, gdyż p_m jest pierwszy, wobec tego $p_m \sim q_{i_m}$.
Zauważmy teraz, że jeśli $ap \sim bq$ oraz $0 \neq p \sim q$, to $a \sim b$. Istotnie, $ap \sim bq \sim bp$ daje $ap \sim bp$, skąd $a \sim b$. Zastosujmy tę uwagę biorąc $a = a_1 \ldots a_{m-1}$, $p = p_m$, $q = q_{i_m}$ oraz $b = q_1 \ldots \widehat{q_{i_m}} \ldots q_n$ (gdzie $\ldots \widehat{q} \ldots$ oznacza pominięcie czynnika q). Dostajemy $p_1 \ldots p_{m-1} \sim q_1 \ldots \widehat{q_{i_m}} \ldots q_n$; skoro $\min(m-1, n-1) = k$, z założenia indukcyjnego T_k wynika stąd, że $m - 1 = n - 1$ oraz $p_1 \sim q_{i_1}, \ldots, p_{n-1} \sim q_{i_{n-1}}$; zarazem $m = n$ i $p_m = q_{i_m}$, co kończy dowód indukcyjny.

110. Z tego, co już wiemy, dostajemy następujący obraz zawierań pomiędzy dziedzinami euklidesowymi (DE), dziedzinami ideałów głównych (DIG) i dziedzinami z jednoznacznością rozkładu (DZJR):

[14] W istocie wykażemy tu, że w dziedzinie całkowitości (niekoniecznie d.i.g.) rozkład na czynniki *pierwsze* (to więcej niż *nierozkładalne*!), o ile istnieje, jest jednoznaczny.

$$\mathrm{DE} \subset \mathrm{DIG} \subset \mathrm{DZJR}.$$

Przykładami pierścieni kategorii $\mathrm{DZJR} \setminus \mathrm{DIG}$, są $R = \mathbb{Z}[x]$ (wielomiany zmiennej x o współczynnikach z $\mathbb{Z}$) oraz $R = \mathbb{K}[x, y]$ (wielomiany dwóch zmiennych); zob. 92. Trudniej wskazać przykłady pierścieni z $\mathrm{DIG} \setminus \mathrm{DE}$ (jak badać 'nieeuklidesowość'?).

111. **Przykład**. W dziedzinie całkowitości $\mathbb{Z}[i\sqrt{6}]$ nie ma jednoznaczności rozkładu, gdyż np. $-6 = 2(-3) = \delta \cdot \delta$, gdzie $\delta := i\sqrt{6}$, przy czym każdy z elementów $2, 3, \delta$ jest nierozkładalny (gdyż $|2|^2 = 4, |3|^2 = 9, |\delta|^2 = 6$, zaś $|a + b\delta|^2 = a^2 + 6b^2$). Ponadto elementy 2 i 3 nie są pierwsze, gdyż są dzielnikami $\delta \cdot \delta$, nie dzielą zaś δ.

Załóżmy teraz, że R jest d.i.g. Ustalmy parę $a, b \in R$; niech $p_1, p_2, \ldots, p_r$ będą wszystkimi czynnikami pierwszymi (z dokładnością do relacji stowarzyszenia, a więc $p_i \not\sim p_j$ dla $i \neq j$), występującymi w rozkładach elementów a i b.

112. **Fakt**. Rozłóżmy $a, b \in R$ na czynniki pierwsze: $a = \tilde{a} {p_1}^{\alpha_1} \cdot \ldots \cdot {p_r}^{\alpha_r}$, $b = \tilde{b} {p_1}^{\beta_1} \cdot \ldots \cdot {p_r}^{\beta_r}$, gdzie $\tilde{a}, \tilde{b} \in R^*$, $r \in \mathbb{N}$, $p_1, \ldots, p_r$ — elementy pierwsze w R, przy czym $i \neq j \Rightarrow p_i \not\sim p_j$, natomiast $\alpha_i, \beta_i \in \mathbb{Z}_+$. Wówczas

1. $a \mid b \iff \alpha_1 \leqslant \beta_1, \ \ldots, \ \alpha_r \leqslant \beta_r$;
2. $((a, b)) \sim {p_1}^{\min(\alpha_1, \beta_1)} \cdot \ldots \cdot {p_r}^{\min(\alpha_r, \beta_r)}$;
3. $[\![a, b]\!] \sim {p_1}^{\max(\alpha_1, \beta_1)} \cdot \ldots \cdot {p_r}^{\max(\alpha_r, \beta_r)}$.

Ad 1. Jeśli $b = ac$, to każdy z czynników pierwszych c jest czynnikiem pierwszym b, więc c ma rozkład $c = \tilde{c} {p_1}^{\gamma_1} \cdot \ldots \cdot {p_r}^{\gamma_r}$, zatem $\beta_i = \alpha_i + \gamma_i \geqslant \alpha_i$. Odwrotne wynikanie '$\Leftarrow$' jest oczywiste: $c = \tilde{a}^{-1} \tilde{b} \prod_i {p_i}^{\beta_i - \alpha_i}$. *Ad 2.* Każdy wspólny dzielnik a i b ma w rozkładzie tylko takie czynniki pierwsze, które występują w a i b, więc ma postać $d = \tilde{d} {p_1}^{\delta_1} \cdot \ldots \cdot {p_r}^{\delta_r}$, gdzie (zgodnie z 1.) $\delta_i \leqslant \alpha_i$ i $\delta_i \leqslant \beta_i$, tzn. $\delta_i \leqslant \min(\alpha_i, \beta_i)$; stąd teza. *Ad 3.* Jeśli $a \mid w$ i $b \mid w$, to ${p_i}^{\alpha_i} \mid w$ i ${p_i}^{\beta_i} \mid w$, więc ${p_i}^{\max(\alpha_i, \beta_i)} \mid w$; stąd, skoro czynniki ${p_i}^{\max(\cdot)}$ są względnie pierwsze, to w musi być krotnością ich iloczynu.

113. **Przykład**. Jeśli $R = \mathbb{Z}$, $a = 4200$, $b = 495$, to rozkładając na czynniki pierwsze dostajemy $a = 2^3 \cdot 3 \cdot 5^2 \cdot 7 \cdot 11^0$, $b = 2^0 \cdot 3^2 \cdot 5^1 \cdot 7^0 \cdot 11$, więc $((a, b)) = 2^0 \cdot 3 \cdot 5 \cdot 7^0 \cdot 11^0 = 15$, $[\![a, b]\!] = 2^3 \cdot 3^2 \cdot 5^2 \cdot 7 \cdot 11 = 138\,600$.

2.9 Ciało ułamków pierścienia

114. Pierścień $\mathbb{Z}$ jest podpierścieniem ciała $\mathbb{Q}$; oczywiście $\mathbb{Z}$ jest również podpierścieniem wielu innych ciał, np. $\mathbb{R}$ lub $\mathbb{C}$ lub $\mathbb{Q} + i\mathbb{Q}$, lecz $\mathbb{Q}$ jest najmniejszym spośród tych ciał, jeśli bowiem $\mathbb{Z} \subset \mathbb{K}$ i $\mathbb{K}$ jest zamknięte względem dzielenia, to każdy ułamek $q = \frac{l}{m} \in \mathbb{Q}$ musi należeć do $\mathbb{K}$.

Pokażemy teraz, że podobna sytuacja ma miejsce dla wielu innych pierścieni: jeśli pierścień R jest 'dostatecznie dobry', to można utworzyć pewne ciało $Q(R)$, zawierające R jako podpierścień; co więcej, każdy element $Q(R)$ będzie ilorazem dwóch elementów z R, a zatem $Q(R)$ będzie *najmniejszym* ciałem zawierającym R.

Co to znaczy, że pierścień R jest 'dostatecznie dobry', tzn. że jest podpierścieniem jakiegoś ciała? Zauważmy, że każde ciało ma mnożenie

łączne i przemienne, ponadto zaś $(a \neq 0, b \neq 0) \Rightarrow (ab \neq 0)$; jasne jest, że te własności (łączność i przemienność mnożenia oraz brak nietrywialnych dzielników zera) są 'dziedziczone' przez każdy podpierścień ciała. Wobec tego o danym pierścieniu R musimy założyć, że jest dziedziną całkowitości, jest to bowiem (dla pierścienia z jedynką) warunek konieczny możliwości 'zanurzenia' R w jakieś ciało.

115. **Twierdzenie**. Niech R będzie pierścieniem z jedynką. Wówczas

$$\begin{pmatrix}\text{istnieje ciało } Q\text{, zawiera-}\\ \text{jące } R \text{ jako podpierścień}\end{pmatrix} \Leftrightarrow \begin{pmatrix}R \text{ jest dziedzi-}\\ \text{ną całkowitości}\end{pmatrix}.$$

Dowód '$\Rightarrow$' właśnie przedstawiliśmy; dowód wynikania '$\Leftarrow$' stanowi następny punkt:

116. **Konstrukcja ciała** $Q = Q(R)$. Niech R będzie dziedziną całkowitości. W zbiorze $Z := R \times (R \setminus \{0\}) = \{(l,m) : l, m \in R, m \neq 0\}$ wprowadźmy następującą relację: $(l_1, m_1) \sim (l_2, m_2) \overset{\text{def}}{\iff} l_1 m_2 = l_2 m_1$. Łatwo sprawdzić, że jest to relacja równoważności([15]). Zatem zbiór Z rozkłada się na klasy równoważności tej relacji; niech $Q := Z/_\sim$ będzie zbiorem wszystkich tych klas. Oznaczmy

$$\frac{l}{m} := \text{klasa równoważności, zawierająca element } (l,m); \qquad (*)$$

wobec tego z definicji

$$\frac{l_1}{m_1} = \frac{l_2}{m_2} \iff l_1 m_2 = l_2 m_1$$

oraz

$$Q = \{\frac{l}{m} : l, m \in R, m \neq 0\} = Z/_\sim.$$

Zdefiniujmy w zbiorze Q działania mnożenia i dodawania następująco:

$$q_1 = \frac{l_1}{m_1},\ q_2 = \frac{l_2}{m_2} \ \Rightarrow\ q_1 q_2 := \frac{l_1 l_2}{m_1 m_2},\ q_1 + q_2 := \frac{l_1 m_2 + m_1 l_2}{m_1 m_2};$$

wzory te są sensowne: $m_1 m_2 \neq 0$ wskutek tego, że R jest dziedziną całkowitości i $m_1, m_2 \neq 0$, co więcej zaś, prawe strony nie zależą od wyboru przedstawienia elementów q_1 i q_2 jako 'ilorazów' typu $\frac{l}{m}$ ([16]).

Jest rzeczą bardzo łatwą (lecz nudną) sprawdzenie, że zbiór Q z tak zdefiniowanymi działaniami jest ciałem; w szczególności zerem jest $\frac{0}{1}$ (przy czym $\frac{l}{m} = \frac{0}{1} \Leftrightarrow l = 0$), jedynką jest $\frac{1}{1}$, elementem przeciwnym do $\frac{l}{m}$ — element $\frac{-l}{m}$, a odwrotnością $\frac{l}{m}$ (dla $l \neq 0$) — element $\frac{m}{l}$.

Wzory $\frac{a}{1} + \frac{b}{1} = \frac{a+b}{1}$, $\frac{a}{1} \cdot \frac{b}{1} = \frac{ab}{1}$ oraz to, że $\frac{a}{1} = \frac{b}{1} \iff a = b$, sprawiają, że podzbiór $\tilde{R} := \left\{\frac{a}{1} : a \in R\right\}$ jest podpierścieniem ciała Q, «wyglądającym z punktu widzenia własności algebraicznych tak samo,

---

[15] Np. jeśli $(l_1, m_1) \sim (l_2, m_2)$ oraz $(l_2, m_2) \sim (l_3, m_3)$, tzn. $\left.\begin{matrix} l_1 m_2 = l_2 m_1 \\ l_2 m_3 = l_3 m_2 \end{matrix}\right\}$, to $l_1 m_2 m_3 = l_2 m_1 m_3 = m_1 l_3 m_2$, czyli $m_2(l_1 m_3 - l_3 m_1) = 0$, skąd wskutek $m_2 \neq 0$ i braku dzielników zera wynika $l_1 m_3 - l_3 m_1 = 0$, tzn. mamy przechodniość $(l_1, m_1) \sim (l_3, m_3)$.

[16] Jeśli bowiem weźmiemy inne przedstawinia $q_1 = \frac{L_1}{M_1}$, $q_2 = \frac{L_2}{M_2}$, wtedy $l_1 M_1 = L_1 m_1$, $l_2 M_2 = L_2 m_1$, stąd zaś wynika, że $l_1 l_2 M_1 M_2 = L_1 L_2 m_1 m_2$, a więc $\frac{l_1 l_2}{m_1 m_2} = \frac{L_1 L_2}{M_1 M_2}$, a także $(l_1 m_2 + m_1 l_2) M_1 M_2 = (L_1 M_2 + M_1 L_2) m_1 m_2$, więc $\frac{l_1 m_2 + m_1 l_2}{m_1 m_2} = \frac{L_1 M_2 + M_1 L_2}{M_1 M_2}$.

jak pierścień R». Chcąc nadać ścisły sens ostatniej konstatacji określmy odwzorowanie $j : R \to Q$ wzorem

$$j(x) := \frac{x}{1} = \text{klasa zawierająca parę } (x, 1).$$

Jest widoczne, że j jest injektywne (a więc jest bijekcją R na $\tilde{R} = j(R) \subset Q$) oraz że ma własności $\left\{ \begin{matrix} j(a+b) = j(a) + j(b) \\ j(ab) = j(a)j(b) \end{matrix} \right\}$, czyli że jest *homomorfizmem pierścieni*. Taką rzecz wyraża się zwykle, mówiąc że j jest *zanurzeniem* pierścienia R w ciało Q. W takim razie pokazaliśmy, że *każda dziedzina całkowitości daje się zanurzyć w pewne ciało*; co więcej, *istnieje takie ciało $Q = Q(R)$, tzw.* ciało ułamków pierścienia R, *które zawiera podpierścień $\tilde{R}$ izomorficzny z R i generujący całe ciało Q* (tzn. taki, że każdy element Q jest ilorazem elementów z $\tilde{R}$).

117. **Przykłady**. 1 Dla $R = \mathbb{Z}$, oczywiście, $Q(R) = \mathbb{Q} =$ (ciało liczb wymiernych).

2 Pokażemy, że dla $R = \mathbb{Z}[i]$ (pierścień Gaussa) ciało $\mathbb{Q}(R)$ można utożsamić z ciałem $\mathbb{Q} + i\mathbb{Q} = \{z \in \mathbb{C} : \mathrm{Re}\, z, \mathrm{Im}\, z \in \mathbb{Q}\}$ *wymiernych liczb zespolonych*.

Rezygnując z oznaczenia (*) oznaczmy teraz $\sim$-klasę pary $(l, m) \in Z$ symbolem $[(l, m)]$, podczas gdy $\frac{l}{m} := l \cdot m^{-1}$ oznaczać będzie, jak zwykle, iloraz w ciele $\mathbb{C}$. Skoro $[(l_1, m_1)] = [(l, m)] \;\Rightarrow\; l_1 m = m_1 l \;\Rightarrow\; \frac{l_1}{m_1} = \frac{l}{m}$, to sensowna jest definicja $F\big([(l, m)]\big) := \frac{l}{m}$ odwzorowania $F : Q = Q(R) \to \mathbb{C}$. Wystarczy jeszcze zauważyć, że — wprost z określenia działań w Q — F jest homomorfizmem: $F(q_1 + q_2) = F(q_1) + F(q_2)$, $F(q_1 q_2) = F(q_1)F(q_2)$; ponadto F jest injektywne: $F(q_1) = F(q_2) \;\Rightarrow\; F(q_1 - q_2) = 0 \;\Rightarrow\; q_1 - q_2 = 0$, oraz $F(Q) = \mathbb{Q} + i\mathbb{Q}$, gdyż $\frac{l}{m} = \frac{l\overline{m}}{|m|^2}$ jest dla $l, m \in R$ wymierną liczbą zespoloną, a zarazem każda liczba $z \in Q + i\mathbb{Q}$ jest wartością F: $z = F\big([(l_1, m_1)] + [(il_2, m_2)]\big)$, gdzie $l_1, l_2 \in \mathbb{Z}$ są licznikami, a $m_1, m_2 \in \mathbb{Z}$ — mianownikami liczb $\mathrm{Re}\, z$ i $\mathrm{Im}\, z$. Wobec tego f jest *izomorfizmem* (tzn. bijektywnym homomorfizmem) ciała Q na ciało $\mathbb{Q} + i\mathbb{Q}$.

3 Załóżmy, że R jest podpierścieniem jakiegoś ciała $\mathbb{K}$. Powtarzając wywody 2 bez trudu stwierdzamy, że w tym przypadku $Q(R)$ można utożsamiać z podciałem $\mathbb{K}_0 \subset \mathbb{K}$ *generowanym przez R*, tzn. najmniejszym spośród takich podciał $\mathbb{K}_0 \subset \mathbb{K}$, że $R \subset \mathbb{K}_0$; ponadto $\mathbb{K}_0 = \{lm^{-1} : l, m \in R,\ m \neq 0\}$ składa się z ilorazów (w sensie operacji w $\mathbb{K}$) o licznikach i mianownikach należących do R.

4 Dla $R = \mathbb{K}[\,\cdot\,]$ ciało $Q = Q(R)$ nosi nazwę *ciała funkcji wymiernych o współczynnikach z $\mathbb{K}$*; słowo 'funkcje' jest tu — podobnie jak przy definiowaniu wielomianów — niezbyt adekwatne: wprawdzie każda 'funkcja wymierna' $q = \frac{L}{M} \in Q(R)$ określa rzeczywiście pewną funkcję:

$$f : D_f \to \mathbb{K},\ \ D_f := \{x \in \mathbb{K} : M(x) \neq 0\},\ \ f(x) := \frac{L(x)}{M(x)} \text{ (iloraz w sensie } \mathbb{K}),$$

lecz niekiedy różnym elementom Q mogą odpowiadać jednakowe funkcje, a nawet może się zdarzyć (dla skończonego ciała $\mathbb{K}$), że f ma pustą dziedzinę: $D_f = \emptyset$! Są i inne niedogodności, np. jeśli $q_1 \mapsto f_1$, $q_2 \mapsto f_2$, $q_1 + q_2 \mapsto f$, to f pokrywa się z $f_1 + f_2$ na $D_{f_1} \cap D_{f_2}$, lecz na ogół D_f jest większe od $D_{f_1} \cap D_{f_2}$, itd. Z tych względów elementy ciała $Q(R)$ należy traktować nie jako funkcje, lecz jako 'prefunkcje', czyli 'recepty na funkcje'; różne funkcje odpowiadające jednemu elementowi $q \in Q$ są określone identycznym 'wzorem', różnią się natomiast dziedzinami, np. można D wybrać jako stosowny podzbiór jakiegoś większego ciała $\tilde{\mathbb{K}} \supset \mathbb{K}$.

Dla $\mathbb{K} = \mathbb{Z}_2 = \{0,1\}$ element $f(x) = \frac{1+x^3}{x+x^2}$ ma pustą dziedzinę D_f: zarówno dla $x = 0$, jak i dla $x = 1$, mianownik znika. Niemniej jednak można traktować f jako funkcję, np. na podzbiorze $D_f = \{a, b\}$ 4-elementowego ciała Galois $\mathrm{GF}_4 = \{0, I, a, b\}$: $f(a) = \frac{I+a^3}{a+a^2} = \frac{I+I}{a+b} = \frac{0}{I} = 0$, $f(b) = \frac{I+I}{b+a} = \frac{0}{I} = 0$. Z tego bynajmniej nie wynika, że $f = 0$, dla jeszcze większego ciała (np. GF_8) może być $\exists x \in D_f : f(x) \neq 0$. Co ważne, operacje na 'funkcjach wymiernych' wykonuje się całkiem formalnie, np. $\frac{1}{f(x)} = \frac{x+x^2}{1+x^3}$, $\frac{f(x)}{f(x)+1} = \frac{1+x^3}{1+x+x^2+x^3}$, itd.

2.10 Rozkład na ułamki proste

118. **Fakt**. Jeśli R jest dziedziną ideałów głównych, a $Q = Q(R)$ — ciałem ułamków pierścienia R, to każdy element $q \in Q$ można przedstawić jako sumę ułamków o mianownikach postaci p^n, gdzie $p \in R$ jest elementem pierwszym oraz $n \in \mathbb{N}$.

Istotnie, $q = \frac{l}{m}$, gdzie $l, m \in R$ oraz $m \neq 0$. Jak wiemy (zob. 109) m ma rozkład na czynniki pierwsze; zatem $m = {p_1}^{\alpha_1} \cdot \ldots \cdot {p_r}^{\alpha_r}$, gdzie $p_1, \ldots, p_r \in R$ są pierwsze i parami niestowarzyszone. Weźmy $m_1 := \frac{m}{p_1^{\alpha_1}} = {p_2}^{\alpha_2} \cdot \ldots \cdot {p_r}^{\alpha_r}$, $\ldots$, $m_r := \frac{m}{p_r^{\alpha_r}}$; są to elementy R takie, że $\mathrm{NWD}(m_1, \ldots, m_r) = [1]_\sim$, więc (zob. 96) $\exists u_1, \ldots, u_r \in R : \sum_{k=1}^{r} m_k u_k = 1$. Stąd $q = \frac{l}{m} = \frac{l(m_1 u_1 + \ldots + m_r u_r)}{m} = \frac{l u_1}{p_1^{\alpha_1}} + \ldots + \frac{l u_r}{p_r^{\alpha_r}}$, QED.

119. W dalszej części tego paragrafu będzie nam potrzebny warunek nieco mocniejszy niż euklidesowość pierścienia([17]). Założymy mianowicie, że para (R, δ) spełnia następujący *mocny warunek Euklidesa*([18]):

$$(\mathrm{E}^*) \quad \forall a, b \in R \setminus \{0\} : \exists r, q \in R : \begin{pmatrix} a = bq + r,\ \delta(r) < \delta(b) \text{ oraz} \\ \text{jeśli } \delta(b) > \delta(1), \text{ to } \delta(q) < \delta(a) \end{pmatrix}$$

120. **Fakt** (*o rozwijaniu przy podstawie p*). Niech p będzie elementem R, takim że $\delta(p) > \delta(1)$. Wówczas każdy $a \in R$ ma rozkład postaci

$$a = r_0 + r_1 p + \ldots + r_n p^n, \qquad (*)$$

gdzie $n \in \mathbb{Z}_+$ oraz $r_0, \ldots, r_n \in C_p := \{r \in R : \delta(r) < \delta(p)\}$ ('p-cyfry').

Ad absurdum. Spośród elementów $a \in R$ nie mających rozkładu postaci (*) wybierzmy taki, który ma najmniejszą możliwą wartość $\delta(a)$. Weźmy parę (r, q) opisaną warunkiem (E*) dla $b = p$; skoro $\delta(q) < \delta(a)$, to q ma p-rozkład, a więc $q = r_1 + r_2 p + \ldots + r_{m+1} p^m$ dla stosownych $m \in \mathbb{Z}_+$ oraz $r_1, \ldots, r_m \in C_p$. Zarazem $r_0 := r \in C_p$ oraz $a = r_0 + pq = r_0 + r_1 p + \ldots + r_{m+1} p^{m+1}$; jest to p-rozkład elementu a, co stanowi sprzeczność z wyborem a, QED.

121. **Uwaga**. Do znalezienia rozwinięcia (*) można posłużyć się następującym prostym algorytmem, w którym znak ':=' oznaczać będzie podstawienie:

[17] Być może tylko pozornie mocniejszy: autor przypuszcza, że każdy pierścień euklidesowy spełnia ten warunek. Nietrudno dowieść, że jest tak dla pierścieni $\mathbb{K}[\,\cdot\,]$, $\mathbb{Z}$ oraz $\mathbb{Z}[i]$.

[18] Wartość $\delta(0)$ określamy tak, by $\forall a \in R \setminus \{0\} : \delta(0) < \delta(a)$, np. przyjmując $\delta(0) := -\infty$. Dzięki temu w warunku (E) zamiast '$r \neq 0 \ \Rightarrow\ \delta(r) < \delta(b)$' wystarczy pisać '$\delta(r) < \delta(b)$'.

$n := 0;\ \ c := a;\ \ r_0 := c;$

$\texttt{until}\big(\delta(c) \geqslant \delta(p)\big)$

$$\left\{\texttt{znajdź } r,q \in R \texttt{ takie, że } \begin{pmatrix} c = pq + r \\ r \in C_p \\ \delta(q) < \delta(c) \end{pmatrix};\ \ r_n := r;\ \ c := q;\ \ n := n+1; \right\}$$

Algorytm zawsze zakończy się po skończonej liczbie kroków, gdyż kolejne wartości $\delta(c)$ tworzą malejący ciąg liczb całkowitych nieujemnych.

Przykład. Dla $R = \mathbb{Z}$, $a = 474$ oraz $p = 7$ kolejne ilorazy i reszty są następujące:

ilorazy	474	67	9	1	0
reszty	$r_0 = 5$	$r_1 = 4$	$r_2 = 2$	$r_3 = 1$,	

a zatem

$474 = 5 + 67 \cdot 7 = 5 + (4 + 9 \cdot 7) \cdot 7 = 5 + \big(4 + (2 + 1 \cdot 7)7\big)7 = 5 + 4 \cdot 7 + 2 \cdot 7^2 + 1 \cdot 7^3.$

Zauważmy, że w kolejnych dzieleniach wybieraliśmy tutaj takie ilorazy, przy których reszty (a więc 'cyfry rozwinięcia') są nieujemne. Nie przestrzegając tej zasady można otrzymać inne rozwinięcia, np.

ilorazy	474	68	10	1	0
reszty	$r_0 = -2$	$r_1 = -2$	$r_2 = 3$	$r_3 = 1$,	

czyli $474 = (-2) + (-2) \cdot 7 + 3 \cdot 7^2 + 1 \cdot 7^3$.

Ponieważ każdy pierścień euklidesowy jest d.i.g., to rozkładając zgodnie z faktem 118 dany $q \in Q$ na sumę ułamków $\frac{a}{p^n}$ oraz rozwijając (zgodnie z faktem 120) każdy z liczników dostajemy natychmiast następujący wniosek:

122. **Twierdzenie** (*o rozkładaniu na ułamki proste*). Jeśli (R, δ) jest dziedziną euklidesową, to każdy element q ciała ułamków $Q = Q(R)$ można przedstawić jako sumę pewnego elementu pierścienia R i pewnej liczby *ułamków prostych*, czyli ułamków postaci $\frac{a}{p^n}$, gdzie $a, p \in R$, p jest elementem pierwszym oraz $\delta(a) < \delta(p)$.

123. **Uwagi**. 1 Z dowodu 118 i 122 widać, że jeśli $m = {p_1}^{\alpha_1} \cdot \ldots \cdot {p_r}^{\alpha_r}$ jest rozkładem na czynniki pierwsze mianownika elementu $q \in Q(R)$, to rozkład q na ułamki proste zawiera jedynie ułamki proste postaci $\frac{b}{p^k}$, gdzie $p = p_j \in \{p_1, \ldots, p_r\}$ oraz $k \in \overline{1, \alpha_j}$.

2 Dla $R = \mathbb{K}[\,\cdot\,]$ ma miejsce jednoznaczność p-rozwinięcia elementu.

Pokażemy najpierw, że *jeśli* $r_0 + r_1 p + \ldots + r_n p^n = 0$ *dla pewnych* $p, r_j \in R = \mathbb{K}[\mathbb{Z}]$ *takich, że* $\forall j : \delta(r_j) < d := \deg(p)$, *to* $r_0 = \ldots = r_n = 0$. Otóż $\delta(p) := \deg p$ ma własności $\delta(p+q) \leqslant \max\{\delta(p), \delta(q)\}$, $\delta(p) < \delta(q) \ \Rightarrow\ \delta(p+q) = \delta(q)$ oraz $\delta(pq) = \delta(p) + \delta(q)$, więc jeśli $\exists j : r_j \neq 0$, to biorąc $h := \max\{j : r_j \neq 0\}$ mamy $\delta(r_h p^h) = \delta(r_h) + h\delta(p) \geqslant hd$, zaś $\forall j < h : \delta(r_j p^j) \leqslant \delta(r_j) + \delta(p^j) < d + jd = (j+1)d \leqslant hd$, a zatem $\delta(r_0 + \ldots + r_n p^n) \geqslant hd \geqslant 0$, czyli $r_0 + \ldots + r_n p^n \neq 0$ (bo $\delta(0) = -\infty$).

Mając to zauważmy, że jeśli $q_0 + q_1 p + \ldots + q_n p^n = \tilde{q}_0 + \tilde{q}_1 p + \ldots + \tilde{q}_n p^n$, to $r_0 + r_1 p + \ldots + r_n p^n = 0$ dla $r_j := \tilde{q}_j - q_j$; implikuje to $\forall j : r_j = 0$, tzn. $\tilde{q}_j = q_j$.

3 Dla $R = \mathbb{K}[\,\cdot\,]$ ma miejsce jednoznaczność rozkładu na ułamki proste.

Niech $q = r + \sum\limits_{i=1}^{r} \frac{l_i}{p_i^{\alpha_i}}$, gdzie $p_1, \ldots, p_r \in R$ są pierwsze, $i \neq j \ \Rightarrow\ p_i \not\sim p_j$ oraz $r \in R$. Pokażemy, że jeśli $q = 0$ oraz $\forall i : \delta(l_i) < \delta(p_i)$, to $r = l_1 = \ldots l_r = 0$. W tym celu

weźmy $m := \prod_{i=1}^{r} p_i^{\alpha_i}$, $m_j := \prod_{i\neq j} p_i^{\alpha_i}$, tzn. $m = m_j p^j$. Skoro $0 = qm = rm + \sum_{i=1}^{r} l_i m_i$, to $\forall j : p_j^{\alpha_j} \mid (l_j m_j)$ bo wszystkie inne niż $l_j m_j$ składniki sumy mają czynnik $p_j^{\alpha_j}$; stąd, wobec $((p_j^{\alpha_j}, m_j)) = 1$, mamy $p_j^{\alpha_j} \mid l_j$. Zarazem $\delta(l_j) < \delta(p_j)$, więc daje to $l_j = 0$. Stąd mamy dowodzoną tezę; z niej oraz z punktu $\boxed{2}$ łatwo wynika $\boxed{3}$.

$\boxed{4}$ Dla innych R na ogół nie ma jednoznaczności ani p-rozwinięcia, ani rozkładu na ułamki proste. Dla $R = \mathbb{Z}$ powszechnie wiadomo (i łatwo dowieść), że p-rozwinięcie staje się *jednoznaczne*, jeśli dorzucimy dodatkowy warunek, by cyfry były wszystkie $\geqslant 0$ lub wszystkie $\leqslant 0$.

Można też dowieść, że każdy ułamek $q \in \mathbb{Q} = Q(\mathbb{Z})$ ma jednoznaczny rozkład na takie ułamki proste, których liczniki są $\geqslant 0$; poniższy przykład pokazuje, jak taki rozkład znaleźć.

124. **Przykład**. Liczba $q = \frac{77}{360} \in \mathbb{Q}$ ma mianownik $m = 360 = 2^3 3^2 5$, więc weźmy $m_1 = \frac{m}{2^3} = 45$, $m_2 = \frac{m}{3^2} = 40$, $m_3 = \frac{m}{5} = 72$. Znajdźmy najpierw rozkład $1 = m_1 u_1 + m_2 u_2 + m_3 u_3$. Skoro $((45, 40)) = 5$, zaś $1 = 5x_1 + 72x_2$ dla $x_1 = 29$, $x_2 = -2$, więc $1 = 45x_1 - 40x_1 + 72x_2$, czyli dobre są $u_1 = 29$, $u_2 = -29$, $u_3 = -2$. Zatem

$$\forall v_i \,:\, 77 = 45\underbrace{(77u_1 - 8v_1)}_{=:a_1} + 40\underbrace{(77u_2 - 9v_2)}_{=:a_2} + 72\underbrace{(77u_3 - 5v_3)}_{=:a_3} + 360(v_1 + v_2 + v_3).$$

Otóż można dobrać $v_1, v_2, v_3 \in \mathbb{Z}$ tak, by $0 \leqslant a_1 < 8$, $0 \leqslant a_2 < 9$, $0 \leqslant a_3 < 5$; tak będzie dla $v_1 = \left[\frac{77\cdot 29}{8}\right] = 279$, $v_2 := \left[\frac{-77\cdot 29}{9}\right] = -249$, $v_3 = \left[\frac{-77\cdot 2}{5}\right] = -31$. Stąd $\frac{77}{360} = \frac{a_1}{8} + \frac{a_2}{9} + \frac{a_3}{5} + (v_1 + v_2 + v_3) = \frac{1}{8} + \frac{8}{9} + \frac{1}{5} - 1 = \frac{1}{8} + \frac{2 + 2\cdot 3}{9} + \frac{1}{5} - 1$, czyli otrzymaliśmy rozkład $\boxed{q = \frac{1}{8} + \frac{2}{9} + \frac{2}{3} + \frac{1}{5} - 1}$.

125. **Przykład**. Niech $q(x) := \frac{x^4}{(x-1)^2(x+1)}$; dokonując dzielenia z resztą dostajemy $q(x) = x+1+\tilde{q}(x)$, gdzie $\tilde{q}(x) = \frac{2x^2 - 1}{(x-1)^2(x+1)}$; zatem $l_0(x) = x+1$ i przewidujemy rozkład $\tilde{q}(x) = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+1}$. Mmnożąc obustronnie przez $(x-1)^2(x+1)$ i porównując współczynniki przy potęgach x dostajemy $A = \frac{7}{4}$, $B = \frac{1}{2}$, $C = \frac{1}{4}$.

Odpowiedź. $q(x) = \frac{7/4}{x-1} + \frac{1/2}{(x-1)^2} + \frac{1/4}{x+1} + (x+1)$.

126. **Przykład**. Dla $q(x) := \frac{l(x)}{x(x-1)(x+2)} = \frac{A}{x} + \frac{B}{x-1} + \frac{C}{x+2}$, gdzie $l(x)$ jest danym trójmianem kwadratowym, współczynniki A, B, C można znaleźć szybciej: Skoro $\frac{l(x)}{(x-1)(x+2)} = A + \frac{Bx}{x-1} + \frac{Cx}{x+2}$, to dla $x := 0$ dostajemy $\boxed{A = -\frac{l(0)}{2}}$. Tak samo skoro $\frac{l(x)}{x(x+2)} = \frac{A(x-1)}{x} + B + \frac{C(x-1)}{x+2}$, to $\boxed{B = \frac{l(1)}{3}}$, zaś skoro $\frac{l(x)}{x(x-1)} = \frac{A(x+2)}{x} + \frac{B(x+2)}{x-1} + C$, to $\boxed{C = \frac{l(-2)}{6}}$.

127. **Przykład**. Rozkład na ułamki proste $q(x) := \frac{x^3 - 2x}{(x^2 + 2x + 2)^2}$ jest inny dla $R = \mathbb{R}[\,\cdot\,]$ niż dla $R = \mathbb{C}[\,\cdot\,]$, podobnie jak rozkład mianownika na czynniki nierozkładalne:

$\boxed{1}$ Dla $R = \mathbb{C}[\,\cdot\,]$ mianownik ma rozkład $(x + 1 + \imath)^2(x + 1 - \imath)^2$, więc szukamy $q(x)$ w postaci $q(x) = \frac{A}{x+1+\imath} + \frac{B}{(x+1+\imath)^2} + \frac{C}{x+1-\imath} + \frac{D}{(x+1-\imath)^2}$. Wtedy $x^3 - 2x = A(x + 1 + \imath)(x + 1 - \imath)^2 + B(x + 1 - \imath)^2 + C(x + 1 + \imath)^2(x + 1 - \imath) + D(x+1+\imath)^2$; porównując współczynniki przy kolejnych potęgach x dostajemy układ

$$\begin{cases} 0 = (2-2i)A - 2iB + (2+2i)C + 2iD, \\ -2 = (4-2i)A + (2-2i)B + (4+2i)C + (2+2i)D, \\ 0 = (3-i)A + B + (3+i)C + D, \\ 1 = A + C, \end{cases} \text{ dający } \begin{cases} A = \frac{1}{2}(1-i), \\ B = -1, \\ C = \frac{1}{2}(1+i), \\ D = -1. \end{cases}$$

$\boxed{2}$ Dla $R = \mathbb{R}[\,\cdot\,]$ czynnik x^2+2x+2 jest nierozkładalny, więc szukamy rozkładu postaci $q(x) = \frac{ax+b}{x^2+2x+2} + \frac{cx+d}{(x^2+2x+2)^2}$; postępując jak w punkcie $\boxed{1}$ znajdujemy, że $a=1$, $b=-2$, $c=0$, $d=4$, czyli $\boxed{q(x) = \frac{x-2}{x^2+2x+2} + \frac{4}{(x^2+2x+2)^2}}$.

2.11 Appendix A: Różniczkowanie wielomianów

128. **Definicja**. *Pochodną* wielomianu $w(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n \in \mathbb{K}[z]$ nazywa się wielomian

$$w'(z) := a_1 + 2a_2 z + \ldots + n a_n z^{n-1} = \sum_k k a_k z^{k-1} \in \mathbb{K}[z].$$

Jest to definicja ściśle algebraiczna (nie ma w niej żadnego przejścia granicznego typu 'granica ilorazu różnicowego'), a więc ma sens dla każdego ciała $\mathbb{K}$.

Łatwym ćwiczeniem jest sprawdzenie, że taka pochodna ma spodziewane własności, np. że zachodzi 'reguła Leibniza':

$$\big(w(z)v(z)\big)' = w'(z)v(z) + w(z)v'(z).$$

Mając pochodną wielomianów możemy (też algebraicznie) zdefiniować *pochodną funkcji wymiernej* tak, by zachować regułę Leibniza:

$$\Big(\tfrac{l(z)}{m(z)}\Big)' := \tfrac{l'(z)m(z) - l(z)m'(z)}{[m(z)]^2}.$$

Sprawdzimy poprawność tej definicji: równość $\frac{l}{m} = \frac{L}{M}$ oznacza, z definicji funkcji wymiernych, że $lM = Lm$. Różniczkując to i mnożąc obustronnie przez mM dostajemy $l'mM^2 + \underbrace{lmM}_{=m^2L} M' = L'm^2M + \underbrace{mML}_{=lM^2} m'$, co oznacza, że $\frac{l'm - lm'}{m^2} = \frac{L'M - LM'}{M^2}$.

129. **Ćwiczenie**. Sprawdzić, że jeśli $w \in \mathbb{K}[\,\cdot\,]$, $z_0, k \in \mathbb{K}$, zaś $q(z_0,\,\cdot\,)$ jest (jednoznacznie określonym) wielomianem takim, że $w(z) - w(z_0) = (z - z_0)q(z_0, z)$ (przypomnijmy, że $(z-z_0) \mid [w(z) - w(z_0)]$ wskutek twierdzenia Bezouta), to:

(1) $(z-z_0)^2 \mid [w(z) - w(z_0) - k\cdot(z-z_0)] \iff k = w'(z_0)$; (2) $w'(z) = q(z,z)$.

2.12 Appendix B: Wielokrotne pierwiastki wielomianu

130. $x_0 \in \mathbb{K}$ nazywamy *pierwiastkiem k-krotnym* wielomianu $W \in \mathbb{K}[\,\cdot\,]$, jeżeli $k = \mathrm{kr}_{x_0}(W)$, gdzie $\mathrm{kr}_{x_0}(W) := \max\{k \in \mathbb{Z}_+ : (x-x_0)^k \mid W(x)\}$; z twierdzenia Bezouta wynika, że $k > 0 \iff W(x_0) = 0$, zatem «pierwiastek o krotności 0» nie jest pierwiastkiem *sensu stricto*. Wprost z definicji mamy

$$\mathrm{kr}_{x_0}(W) = k \Leftrightarrow \exists \check{W} \in \mathbb{K}[\,\cdot\,] :\ W(x) = (x-x_0)^k \check{W}(x) \text{ oraz } \check{W}(x_0) \neq 0.$$

131. **Fakt**. Jeżeli $W, W_1, W_2 \in \mathbb{K}[\,\cdot\,]$, $D \in \mathrm{NWD}(W_1, W_2)$ oraz $x_0 \in \mathbb{K}$, to

$$1^\circ \quad \mathrm{kr}_{x_0}(W_1 W_2) = \mathrm{kr}_{x_0}(W_1) + \mathrm{kr}_{x_0}(W_2);$$

2° jeśli $W_1 \mid W$, to $\mathrm{kr}_{x_0}(W_1) \leqslant \mathrm{kr}_{x_0}(W)$;
3° $\mathrm{kr}_{x_0}(D) = \min\{\mathrm{kr}_{x_0}(W_1), \mathrm{kr}_{x_0}(W_2)\}$;
4° pierwiastkami D są wspólne pierwiastki W_1 i W_2.

Ad 1°: Jeśli $W_1(x) = (x-x_0)^{k_1}\tilde{W}_1(x)$ oraz $W_2(x) = (x-x_0)^{k_2}\tilde{W}_2(x)$, przy czym $W_1(x_0) \neq 0, W_2(x_0) \neq 0$, to $W = W_1W_2$ ma rozkład $W(x) = (x-x_0)^{k_1+k_2}\tilde{W}(x)$, przy czym $\tilde{W}(x_0) = \tilde{W}_1(x_0)\tilde{W}_2(x_0) \neq 0$; stąd teza. *Ad* 2°: Wynika wprost z 1°. *Ad* 3°: Z poprzedniego punktu mamy $\mathrm{kr}_{x_0}(D) \leqslant \min\{\mathrm{kr}_{x_0}(W_1), \mathrm{kr}_{x_0}(W_2)\}$; zarazem $D = W_1V_1 + W_2V_2$ dla pewnych $V_1, V_2 \in \mathbb{K}[\,\cdot\,]$, więc jeśli $(x-x_0)^k$ dzieli $W_1(x)$ i $W_2(x)$, to dzieli i $D(x)$, co daje przeciwną nierówność. *Ad* 4° Wynika wprost z 3°.

132. **Fakt**. Jeśli $W \in \mathbb{K}[\,\cdot\,]$, $\mathrm{kr}_{x_0}(W) \geqslant 1$ oraz $D \in \mathrm{NWD}(W, W')$, to

$$\mathrm{kr}_{x_0}(D) = \mathrm{kr}_{x_0}(W') = \mathrm{kr}_{x_0}(W) - 1, \;\; \mathrm{kr}_{x_0}\left(\tfrac{W}{D}\right) = 1.$$

Zatem pierwiastkami D są pierwiastki wielokrotne W, zaś wielomian $\frac{W}{D}$ ma te same pierwiastki co W, lecz z krotnościami równymi 1.

Niech $k := \mathrm{kr}_{x_0}(W)$, wtedy $W(x) = (x-x_0)^k\tilde{W}(x)$, gdzie $\tilde{W}(x_0) \neq 0$, więc z reguły Leibniza $W'(x) = (x-x_0)^{k-1}\big(k\tilde{W}(x) + (x-x_0)\tilde{W}'(x)\big) = (x-x_0)^{k-1}V(x)$, przy czym $V(x_0) = k\tilde{W}(x_0) + 0 \neq 0$, a więc $\mathrm{kr}_{x_0}(W') = k-1$ oraz $\mathrm{kr}_{x_0}(D) = \min\{k, k-1\} = k-1$. To oraz równość $\mathrm{kr}_{x_0}(W) = \mathrm{kr}_{x_0}\left(\frac{W}{D}\right) + \mathrm{kr}_{x_0}(D)$ dają tezę.

133. **Przykład**. $W(x) := x^{10}+2x^9+9x^8+14x^7+31x^6+36x^5+50x^4+40x^3+36x^2+16x+8$; stosując algorytm Euklidesa dostajemy $D(x) = x^6+x^5+5x^4+4x^3+8x^2+4x+4$, zaś ilorazem jest $\dfrac{W(x)}{D(x)} = x^4+x^3+3x^2+2x+2$, co ma łatwy do odgadnięcia (lub znalezienia metodą Ferrariego) rozkład: $\dfrac{W(x)}{D(x)} = (x^2+2)(x^2+x+1)$. Zatem pierwiastkami $W(x)$ są $\pm i\sqrt{2}0$ oraz $\frac{1}{2}(-1 \pm i\sqrt{3}1)$; zarazem $D(x) = (x^2+2)\dfrac{W(x)}{D(x)}$, więc wszystkie pierwiastki $W(x)$ mają krotności $\geqslant 2$; dzieląc wielomiany dostajemy

$$\frac{W(x)}{(x^2+2)^2(x^2+x+1)^2} = x^2+2, \;\; \text{tzn.} \;\; W(x) = (x^2+2)^3(x^2+x+1)^2.$$

134. **Fakt**. Jeśli ciało $\mathbb{K}$ ma charakterystykę równą 0, tzn. jeśli $1+\ldots+1 \neq 0$ w ciele $\mathbb{K}$ dla każdej liczby dodawanych jedynek, to

$$\mathrm{kr}_{x_0}(W) = \min\{k \in \mathbb{Z}_+ : W^{(k)}(x_0) \neq 0\},$$

czyli $k = \mathrm{kr}_{x_0}(W)$ wtedy i tylko wtedy, gdy

$$W(x_0) = W'(x_0) = \ldots = W^{(k-1)}(x_0) = 0, \;\; W^{(k)}(x_0) \neq 0.$$

Istotnie, przedstawmy $W(x)$ w postaci $W(x) = \sum\limits_{n=0}^{m} a_n(x-x_0)^n$; mamy wówczas $\mathrm{kr}_{x_0}(W) = k \iff \big(a_0 = \ldots = a_{k-1} = 0, \, a_k \neq 0\big)$, przy czym $a_n = \frac{1}{n!}W^{(n)}(x_0)$.

3 Grupy i permutacje

3.1 Definicja grupy; przykłady

135. **Definicja**. Zbiór G, wyposażony w operację (działanie) $\Phi : G\times G \to G$, nazywamy *grupą*, jeśli spełnione są następujące *aksjomaty grupy*:

(1) $\forall a,b,c \in G : \Phi(\Phi(a,b),c) = \Phi(a,\Phi(b,c))$ (*łączność*);
(2) $\exists e \in G : \forall a \in G : \Phi(e,a) = a,\ \Phi(a,e) = a$ (*element neutralny*);
(3) $\forall a \in G : \exists b \in G : \Phi(a,b) = e,\ \Phi(b,a) = e$ (*istnienie odwrotności*).

Grupa G nazywa się *przemienna*, albo *abelowa*, jeśli oprócz tego

(4) $\forall a,b \in G : \Phi(a,b) = \Phi(b,a)$ (*warunek przemienności*).

136. **Fakt**. W grupie jest tylko jeden element neutralny, tzn. element $e \in G$, scharakteryzowany aksjomatem (2).

Jeśli bowiem $\forall a : \left.\begin{array}{l}\Phi(e,a) = a = \Phi(a,e)\\ \Phi(\tilde{e},a) = a = \Phi(a,\tilde{e})\end{array}\right\}$, to w szczególności $\tilde{e} = \Phi(\tilde{e},e) = e$.

Element neutralny odgrywa istotną rolę także w aksjomacie (3), więc dopiero teraz możemy się zająć kwestią jednoznaczności odwrotności.

137. **Fakt**. Każdy element $a \in G$ ma tylko jedną odwrotność.

Istotnie, $\left.\begin{array}{l}\Phi(a,b) = e\\ \Phi(\tilde{b},a) = e\end{array}\right\} \Rightarrow \tilde{b} = \Phi(\tilde{b},e) = \Phi(\tilde{b},\Phi(a,b)) = \Phi(\Phi(\tilde{b},a),b) = \Phi(e,b) = b.$

138. **Fakt**. W grupie G obowiązują następujące 'prawa skracania':

(i) $\Phi(a,b) = \Phi(a,c) \ \Rightarrow\ b = c$; (ii) $\Phi(a,b) = \Phi(c,b) \ \Rightarrow\ a = c$.

Ad(i): Niech $\tilde{a}$ będzie odwrotnością a, wtedy $\Phi(\tilde{a},a) = e$; skoro $\Phi\big(\tilde{a},\Phi(a,b)\big) = \Phi\big(\tilde{a},\Phi(a,c)\big)$, to z łączności $\Phi\big(\Phi(\tilde{a},a),b\big)\big) = \Phi\big(\Phi(\tilde{a},a),c\big)\big)$, czyli $\Phi(e,b) = \Phi(e,c)$; na mocy własności elementu e oznacza to, że $b = c$. Analogicznie dowodzimy (ii).

139. **Fakt**. Jeśli elementy $a,b \in G$ spełniają warunek $\Phi(a,b) = e$, to także $\Phi(b,a) = e$, czyli a i b są wzajemnie odwrotne.

Niech $\tilde{a}$ będzie odwrotnością a; wtedy $b = \Phi(e,b) = \Phi\big(\Phi(\tilde{a},a),b\big) = \Phi\big(\tilde{a},\Phi(a,b)\big) = \Phi(\tilde{a},e) = \tilde{a}$, skąd oczywiście wynika teza.

140. **Uwagi**

1. Formalnie rzecz biorąc *zbiór G, wyposażony w operację Φ*, to para (G,Φ), dlatego puryści definiują grupę jako *parę (G,Φ), złożoną ze zbioru G i odwzorowania $\Phi : G\times G \to G$, spełniających warunki ....*

2. Chcąc mieć wzory wyglądające bardziej swojsko używa się zamiast «$\Phi(a,b)$» symbolu typu «$a\Phi b$», przy czym zwykle zamiast litery «Φ» używa się czegoś «kropkopodobnego» lub «plusopodobnego»:

W notacji i terminologii *multiplikatywnej*: pisze się $\Phi(a,b) = ab$ (lub $a\odot b$ lub $a\cdot b$ lub $a\circ b$ lub $a\bullet b$ lub $a\star b$ lub tp.), element neutralny nazywa się «jedynką» lub «jednością» (i często oznacza jako 1 lub I

lub **1** lub tp.), a odwrotność elementu a oznacza jako a^{-1}.

W notacji i terminologii *addytywnej*: $\Phi(a,b)$ pisze się jako $a+b$ (lub $a \oplus b$ lub $a \mathbin{\hat{+}} b$ lub tp.), element neutralny nazywa się «zerem» (i zwykle oznacza jako 0 lub **0** lub O lub tp.), a odwrotność a oznacza jako $-a$ i nazywa *elementem przeciwnym do* a.

Ważna konwencja:

notacji addytywnej używa się z reguły tylko dla grup przemiennych!

Możemy teraz przepisać wzory występujące w aksjomatach grupy w 'standardowej' (tzn. multiplikatywnej) notacji([19]):

$$(ab)c = a(bc),\ \ ea = a,\ \ ae = a,\ \ aa^{-1} = e,\ \ a^{-1}a = e$$

(ewentualnie z warunkiem przemienności $ab = ba$), lub w addytywnej:

$$\begin{array}{l}(a+b)+c = a+(b+c), \quad 0+a = a, \quad a+0 = a,\\ a+(-a) = 0, \qquad (-a)+a = 0, \qquad a+b = b+a.\end{array}$$

141. **Fakt** (*własności odwrotności*). Dla $a,b,c \in G$ zachodzą wzory

$$(a^{-1})^{-1} = a, \qquad (ab)^{-1} = b^{-1}a^{-1}, \qquad ab = c \Leftrightarrow a = cb^{-1} \Leftrightarrow b = a^{-1}c.$$

Dla $b := a^{-1}$ mamy $ab = e = ba$, a więc $a = b^{-1}$, co daje pierwszy dowodzony wzór. Dalej, $c := b^{-1}a^{-1}$ jest odwrotnością ab (lecz na ogół nie ba!), gdyż $(ab)c = a\big(bc\big) = a\big(b(b^{-1}a^{-1})\big) = a\big((bb^{-1})a^{-1}\big) = a\big(ea^{-1}\big) = aa^{-1} = e$, a także $c(ab) = \ldots = e$. Mnożąc lewostronnie obie strony $ab = c$ przez a^{-1} dostajemy $b = a^{-1}c$, itd.

142. **Definicja**. Jeśli dla G stosujemy notację multiplikatywną, to *potęgami* elementu $a \in G$ nazywamy iloczyny postaci $a \cdot \ldots \cdot a$ lub $a^{-1} \cdot \ldots \cdot a^{-1}$, tzn. wyrazy ciągu $(a^k)_{k\in\mathbb{Z}}$, określonego warunkami

$$a^0 := e,\ \ a^{k+1} = a^k a,\ \ a^{k-1} := a^k a^{-1}.$$

Łatwo się przekonać, że operacja potęgowania ma 'zwykłe' własności

$$a^{k+l} = a^k a^l,\ \ (a^k)^l = a^{kl},\ \ \text{w szczególności}\ \ a^{-k} = (a^k)^{-1} = (a^{-1})^k;$$

natomiast $(ab)^k$ jest na ogół różne od $a^k b^k$.

W notacji addytywnej odpowiednikiem potęgi jest *krotność* elementu: dla $a \in G$ określamy ka jako $\underbrace{a+\ldots+a}_{k \text{ razy}}$ lub $\underbrace{(-a)+\ldots(-a)}_{-k \text{ razy}}$, zależnie od znaku współczynnika $k \in \mathbb{Z}$.

143. *Półgrupą* nazywamy zbiór P, wyposażony w operację $\Phi : P \times P \to P$, spełniającą warunek (1) (łączności), tzn. $\Phi(\Phi(a,b),c) = \Phi(a,\Phi(b,c))$. *Półgrupa z jedynką* jest półgrupą, spełniającą także warunek (2), tzn. $\exists e \in G : \forall a \in G : \Phi(e,a) = a,\ \Phi(a,e) = a$.

Zwykle używa się w tym kontekście notacji multiplikatywnej, czyli np.

$$\circ : P \times P \to P, \qquad (a \circ b) \circ c = a \circ (b \circ c), \qquad e \circ a = a = a \circ e.$$

144. **Garść przykładów**

[19] Takiej właśnie notacji, z 'niewidzialnym' symbolem operacji, używa się najczęściej w teoriogrupowych rozważaniach, zwłaszcza jeśli się nie zakłada przemienności grupy.

(A) Jeśli «+» oznacza zwykłe dodawanie, a «·» — zwykłe mnożenie liczb, to przykładami grup są:

$(\mathbb{Z},+)$; $(\mathbb{C},+)$; $(\mathbb{R},+)$; $(\mathbb{Q},+)$; $(\mathbb{C}^*,\cdot)$; $(\mathbb{R}^*,\cdot)$; $(\mathbb{R}^+,\cdot)$; $(\mathbb{Q}^*,\cdot)$; $(\mathbb{Q}^+,\cdot)$; $(U,\cdot)$, gdzie $U := \{u \in \mathbb{C} : |u| = 1\}$; $(\sqrt[n]{1},\cdot)$ dla $n \in \mathbb{C}$.

(B) Określmy grupę $(\mathbb{Z}_6, \underset{6}{+})$, nazywaną *grupą reszt modulo 6*, biorąc jako $\mathbb{Z}_6$ dowolny zbiór, którego elementy są numerowane liczbami $0,\ldots,5$, dajmy na to $\mathbb{Z}_6 := \{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4},\overline{5}\}$, oraz określając działanie $\underset{6}{+}$ wzorem

$$\overline{k} \underset{6}{+} \overline{l} := \overline{m}, \text{ gdzie } m \text{ jest resztą z dzielenia } k+l \text{ przez } 6.$$

Łatwo się przekonać, że $(\mathbb{Z}_6, \underset{6}{+})$ rzeczywiście jest grupą (przemienną), w szczególności elementem neutralnym (zerem) jest $\overline{0}$, a elementem przeciwnym do $\overline{k}$ jest

$$-\overline{k} = \overline{6-k}.$$

$\underset{6}{+}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$

$\overline{k}$	$-\overline{k}$
$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{5}$
$\overline{2}$	$\overline{4}$
$\overline{3}$	$\overline{3}$
$\overline{4}$	$\overline{2}$
$\overline{5}$	$\overline{1}$

Oczywiście w podobny sposób dla $n \in \mathbb{N}$ można utworzyć grupę $(\mathbb{Z}_n, \underset{n}{+})$, zwaną *grupą reszt modulo* n; jej elementy bywają oznaczane różnie:

$$\overline{0},\overline{1},\ \ldots\ ,\overline{n-1} \text{ albo } [0]_n,[1]_n,\ \ldots\ ,[n-1]_n,$$
$$\text{albo } 0+n\mathbb{Z},1+n\mathbb{Z},2+n\mathbb{Z},\ \ldots\ ,n-1+n\mathbb{Z},$$
$$\text{albo wręcz po prostu jako } 0,1,2,\ \ldots\ ,n-1,\ \text{itp.},$$

lecz w każdej notacji wygodnie jest przyjąć zasadę, że *elementy zbioru* $\mathbb{Z}_n$ *zależą okresowo* (o okresie n) od swojego numeru $k \in \mathbb{Z}$, a więc np.

$$[0]_n = [n]_n = [-n]_n = \ldots\ ,\ [1]_n = [n+1]_n = [-n+1]_n = \ldots \text{ itd};$$

wobec tego element oznaczony jako $\overline{k}$ (albo $[k]_n$, albo $k+n\mathbb{Z}$ albo $\ldots$) zależy jedynie od reszty modulo n ze swojego numeru $k \in \mathbb{Z}$. Prosta jest wtedy definicja *dodawania modulo* n, np. $[k]_n \underset{n}{+} [l]_n := [k+l]_n$ itd.

(C) Łatwo sprawdzić, że jeśli P jest *półgrupą z jedynką*, to podzbiór $P^* := \{a \in P : \exists b \in P : a \circ b = e = b \circ a\}$ (*elementy odwracalne półgrupy*) jest grupą względem operacji '$\circ$' obciętej do $P^* \times P^*$; m.in. trzeba tu sprawdzić, że iloczyny oraz odwrotności elementów z P^* należą do P^*. W szczególności *elementy odwracalne pierścienia łącznego z jedynką tworzą grupę* (multiplikatywną).

(D) *Permutacją* zbioru X nazywa się bijektywne (czyli odwracalne) odwzorowanie $X \to X$. Symbolem S_X oznaczmy zbiór wszystkich permutacji X, tzn. $S_X := \{\varrho : X \to X :\ \varrho \text{ jest bijekcją}\}$. Sprawdzimy, że zbiór S_X wraz z operacją '$\circ$' składania odwzorowań stanowi grupę([20]):

Zauważmy najpierw, że '$\circ$' odwzorowuje $S_X \times S_X$ w S_X, łatwo bowiem spostrzec, że złożenie surjekcji jest surjekcją, złożenie injekcji — injekcją, a zatem złożenie

[20] Zauważmy, że jest to szczególny przypadek sytuacji z punktu (C), jeśli P jest zbiorem wszystkich odwzorowań $P \to P$, zaś '$\circ$' jest operacją składania odwzorowań.

bijekcji $\varrho, \sigma : X \to X$ jest bijekcją $\varrho \circ \sigma : X \to X$. ([21]) Aksjomat (1) (*łączność*) jest spełniony, gdyż składanie wszelkich odwzorowań jest łączne: Jeśli $\omega_1 = \varrho \circ (\sigma \circ \tau)$, a $\omega_2 = (\varrho \circ \sigma) \circ \tau$, to biorąc $x' := \tau(x)$, $x'' := \sigma(x') = \sigma(\tau(x)) = (\sigma \circ \tau)(x)$ dostajemy

$$\omega_1(x) = \varrho(x'') = \varrho\big(\sigma(\tau(x))\big) = \varrho\big(\sigma(x')\big) = (\varrho \circ \sigma)(x') = \omega_2(x).$$

Aksjomat (2) jest spełniony, gdyż odwzorowanie identycznościowe $e := \mathrm{id}_X$ jest elementem neutralnym w S_X: $\sigma \circ e = \sigma = e \circ \sigma$. Dla sprawdzenia aksjomatu (3) zauważmy, że relacja 'ρ jest odwzorowaniem odwrotnym do σ' (tzn. $\sigma \circ \varrho = \mathrm{id}_X = \varrho \circ \sigma$) jest symetryczna, więc odwzorowanie odwrotne do bijekcji $\sigma : X \to X$ jest bijekcją.

145. Grupy pojawiają się w wielu innych, bardziej złożonych obiektach algebraicznych, np. w pierścieniach i ciałach. Analizując naszą dotychczasową definicję ciała zauważymy, że struktura ciała «mieści w sobie» aż dwie struktury grupowe. Wobec tego wychodząc z pojęcia grupy możemy teraz sformułować nową, krótszą definicję ciała:

Definicja. Trójka $(\mathbb{K}, +, \cdot)$ nazywa się *ciałem*, jeśli $\mathbb{K}$ jest zbiorem o $\geqslant 2$ elementach, zaś $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ są takimi *działaniami* w $\mathbb{K}$, że:

1° $(\mathbb{K}, +)$ jest grupą przemienną;
2° $(\mathbb{K} \setminus \{0\}, \cdot)$ jest grupą przemienną ($0 \in \mathbb{K}$ jest zerem działania '+');
3° 'mnożenie' jest rozdzielne względem 'dodawania'.

3.2 Homomorfizmy grup

146. **Definicja** (*homomorfizm grup*). Niech (G_1, Φ_1), (G_2, Φ_2) będą grupami. Odwzorowanie $f : G_1 \to G_2$ nazywa się *homomorfizmem grupy* G_1 *w grupę* G_2, jeśli jest zgodne z działaniami w obu grupach w tym sensie, że

$$\forall a, b \in G_1 : \quad f(\Phi_1(a, b)) = \Phi_2(f(a), f(b)).$$

Jasne, że ostatni warunek można zapisać na cztery sposoby, zależne od notacji (multiplikatywnej lub addytywnej) używanej dla grup G_1 i G_2:

$$\begin{aligned} f(ab) &= f(a)f(b), \quad & f(ab) &= f(a) + f(b), \\ f(a+b) &= f(a)f(b), \quad & f(a+b) &= f(a) + f(b). \end{aligned}$$

147. Jeśli $f : G_1 \to G_2$ oraz $g : G_2 \to G_3$ są homomorfizmami grup, to ich złożenie $g \circ f : G_1 \to G_3$, jak łatwo sprawdzić, też jest homomorfizmem.

Izomorfizmem nazywa się bijektywny homomorfizm; o grupach G_1 i G_2 mówimy, że są *izomorficzne*, jeśli istnieje jakiś izomorfizm $f : G_1 \to G_2$. Łatwo pokazać, że wówczas odwzorowanie $f^{-1} : G_2 \to G_1$ także jest izomorfizmem, a więc «relacja izomorficzności» jest nie tylko zwrotna (id_G jest izmomorfizmem) i przechodnia (złożenie dwu izomorfizmów), ale także symetryczna, jest więc «relacją równoważności dla grup»([22]).

[21] Inne uzasadnienie: jeśli $\varrho, \sigma : X \to X$ są odwracalne (tzn. istnieją odwzorowania $\tilde{\varrho}, \tilde{\sigma}$, takie że $\varrho \circ \tilde{\varrho}$, $\tilde{\varrho} \circ \varrho$, $\sigma \circ \tilde{\sigma}$ i $\tilde{\sigma} \circ \sigma$ są równe id_X), to $\varrho \circ \sigma$ też jest odwracalne, gdyż dla $\alpha := \tilde{\sigma} \circ \tilde{\varrho}$ mamy $(\varrho \circ \sigma) \circ \alpha = \varrho \circ (\sigma \circ \tilde{\sigma}) \circ \tilde{\varrho} = \varrho \circ \tilde{\varrho} = \mathrm{id}_X$ oraz $\alpha \circ (\varrho \circ \sigma) = \ldots = \mathrm{id}_X$.

[22] Cudzysłowy zostały tu użyte dlatego, że nie istnieje «zbiór wszystkich grup», bowiem coś takiego prowadziłoby do antynomii, w podobny sposób jak «zbiór wszystkich zbiorów».

148. **Przykłady**. Łatwo sprawdzić, że poniższe odwzorowania $f : G_1 \to G_2$ (ewentualnie zależące od stałych $p, q \in \mathbb{R}$) są homomorfizmami grup:

Nr	$G_1,\ \Phi_1$	$G_2,\ \Phi_2$	f
1.	$\mathbb{R}^+,\ \cdot$	$\mathbb{R}^+,\ \cdot$	$f(t) := \sqrt{t}$, ogólniej $f(t) := t^p$
2.	$\mathbb{C}^*,\ \cdot$	$\mathbb{C}^*,\ \cdot$	$f(z) := z^3$, ogólniej $f(z) := z^n$, $n \in \mathbb{Z}$
3.	$\mathbb{C}^*,\ \cdot$	$\mathbb{C}^*,\ \cdot$	$f(z) := \overline{z}$
4.	$\mathbb{R}^+,\ \cdot$	$\mathbb{R},\ +$	$f(t) := p \log t$
5.	$\mathbb{R},\ +$	$\mathbb{R}^+,\ \cdot$	$f(t) := e^{pt}$
6.	$\mathbb{R},\ +$	$U,\ \cdot$	$f(t) := e^{ipt}$
7.	$\mathbb{C},\ +$	$\mathbb{R},\ +$	$f(z) := \mathrm{Re}\, z$, ogólniej $f(z) := p\,\mathrm{Re}\, z + q\,\mathrm{Im}\, z$
8.	$\mathbb{C},\ +$	$\mathbb{C},\ +$	$f(z) := \overline{z}$, ogólniej $f(z) := pz + q\overline{z}$
9.	$\mathbb{C}^*,\ \cdot$	$\mathbb{R}^+,\ \cdot$	$f(z) := \lvert z \rvert$
10.	$\mathbb{R}^+,\ \cdot$	$\mathbb{C}^*,\ \cdot$	$f(t) := t^q\big(\cos(p \log t) + i \sin(p \log t)\big)$

gdzie $p, q \in \mathbb{R}$, $\mathbb{R}^+ := \{t \in \mathbb{R} : t > 0\}$ oraz $U := \{u \in \mathbb{C} : |u| = 1\}$.

3.3 Podgrupy

149. **Definicja**. Podzbiór $\emptyset \neq H \subset G$ nazywa się *podgrupą* grupy G, jeśli:

(1) H jest zamknięty względem operacji grupowej: $\forall a, b \in H : ab \in H$,
(2) H jest zamknięty względem brania odwrotności: $\forall a \in H : a^{-1} \in H$.

Warunki te można zapisać krócej: (1) $HH \subset H$, (2) $H^{-1} \subset H$, używając działań na podzbiorach grupy: dla dowolnych $A, B \subset G$ $AB := \{ab : a \in A,\ b \in B\}$, $A^{-1} := \{a^{-1} : a \in A\}$.

Oczywiście używając dla G notacji addytywnej należy powyższe warunki zapisać w innej ('addytywnej') formie:

(1) $H + H \subset H$, tzn. $\forall a, b \in H : a + b \in H$,
(2) $-H \subset H$, tzn. $\forall a \in H : -a \in H$.

150. **Fakt**. Jeśli (G, Φ) jest grupą, a $H \subset G$ — podgrupą, to zbiór H, wyposażony w operację Φ obciętą do $H \times H$, czyli para $(H, \Phi|_{H\times H})$, także jest grupą.
Ponadto element neutralny dla G jest elementem neutralnym dla H, zaś dla $h \in H$ odwrotność h w G jest zarazem odwrotnością h w H.

Z warunku (1) $\tilde{\Phi} := \Phi|_{H\times H}$ jest odwzorowaniem $H \times H \to H$, przy czym z łączności Φ wynika łączność $\tilde{\Phi}$, gdyż jeśli $\forall a, b, c \in G : ***$, to tym bardziej $\forall a, b, c \in H : ***$. Ustalmy $a \in H$, wtedy $b := a^{-1} \in H$ oraz $e = ab \in H$, a zatem element neutralny e grupy G należy do H i e jest tym bardziej elementem neutralnym dla H. Wobec tego dla $h \in H$ element h^{-1}, należący do H wskutek (2), jest odwrotnością elementu h nie tylko w G, ale również w H, QED.

151. **Przykłady**. Każda z grup przykładu 144(A) jest albo podgrupą grupy addytywnej $\mathbb{C}$, albo podgrupą grupy multiplikatywnej $\mathbb{C}^*$.
Jeśli $n \in \mathbb{Z}$, to podzbiór $n\mathbb{Z} := \{kn : k \in \mathbb{Z}\}$ jest podgrupą grupy $\mathbb{Z}$; co więcej, każda podgrupa $H \subset \mathbb{Z}$ jest tej postaci.

Rozważmy przypadek $H \neq \{0\}$. Niech $n :=$ element najmniejszy zbioru $H \cap \mathbb{N}$.

Wtedy $n \in H$, więc skoro H jest grupą, to także $n\mathbb{Z} \subset H$. Odwrotnie, jeśli $h \in H$, to $h = nq + r$, gdzie $r \in \overline{0, n-1}$, zarazem $r = h - nq \in H$, więc $r \neq 0$ dałoby sprzeczność z minimalnością n; zatem $r = 0$, $h = nq \in n\mathbb{Z}$, co dowodzi, że $H \subset nZ$.

Inny dowód: Podgrupa $H \subset \mathbb{Z}$ jest ideałem pierścienia $\mathbb{Z}$ (bo krotność elementu $h \in H$ też jest sumą elementów H), zaś pierścień $\mathbb{Z}$ jest dziedziną ideałów głównych.

152. **Fakt**. Jeśli $f : G_1 \to G_2$ jest homomorfizmem, to (używając notacji multiplikatywno-multiplikatywnej) mamy:

(a) $f(e_1) = e_2$, gdzie e_1, e_2 są elementami neutralnymi dla G_1 i G_2;
(b) $f(a^{-1}) = (f(a))^{-1}$ dla każdego $a \in G_1$.

(a) $f(e_1) = f(e_1 e_1) = f(e_1)f(e_1)$, skąd teza;
(b) $f(a)f(a^{-1}) = f(aa^{-1}) = f(e_1) = e_2$ oraz $f(a^{-1})f(a) = \ldots = e_2$.

153. **Definicja**. Jeśli $f : G_1 \to G_2$ jest homomorfizmem grup, to podzbiory

$$\ker f := f^{-1}\{e_2\} = \{a \in G_1 : f(a) = e_2\} \subset G_1\,,$$
$$\operatorname{im} f := f(G_1) = \{f(a) : a \in G_1\} \subset G_2\,,$$

nazywamy, odpowiednio, *jądrem* (ang. *kernel*) i *obrazem* (ang. *image*) homomorfizmu f. Zauważmy, że f jest surjektywny $\iff \operatorname{im} f = G_2$.

154. **Ćwiczenie**. Znaleźć jądro i obraz każdego z homomorfizmów opisanych w 148.

Odpowiedź. Oznaczmy $K = \ker f$, $L = \operatorname{im} f$. W przykładach 1,4,5,6 dla $p = 0$ mamy oczywiście $K = G_1$, $L = \{e_2\}$; dlatego też poniżej będziemy tam zakładać, że $p \neq 0$.

1. $K = \{1\}$, $L = \mathbb{R}^+$; dla $p = 0$: $K = \mathbb{R}^+$, $L = \{1\}$. **2.** Dla $n \neq 0$: $K = \sqrt[|n|]{1}$, $L = \mathbb{C}^*$. **3.** $K = \{1\}$, $L = \mathbb{C}^*$. **4.** $K = \{1\}$, $L = \mathbb{R}$. **5.** $K = \{0\}$, $L = \mathbb{R}^+$. **6.** $K = \frac{2\pi}{p}\mathbb{Z}$, $L = U$. **7.** $K = i\mathbb{R}$, $L = \mathbb{R}$, w ogólniejszym przypadku, dla $(p, q) \neq (0, 0)$, K jest prostą $px + qy = 0$, zaś $L = \mathbb{R}$. **8.** $K = \{0\}$, $L = \mathbb{C}$, w ogólniejszym przypadku dla $p \neq \pm q$ mamy $K = \{0\}$, $L = \mathbb{C}$. **9.** $K = U$, $L = \mathbb{R}^+$. **10.** Dla $p \neq 0 \neq q$: $K = \{1\}$, zaś $L \subset \mathbb{C}$ jest *spiralą logarytmiczną* $r = e^{a\varphi}$, $a = \frac{q}{p}$; dla $p \neq 0 = q$: $K = \{k^n : n \in \mathbb{Z}\}$, gdzie $k = e^{\frac{2\pi}{p}}$, a spirala degeneruje się do okręgu $L = U$; dla $p = 0 \neq q$: $K = \{1\}$, $L = \mathbb{R}^+$; dla $p = q = 0$: $K = \mathbb{R}^+$, $L = \{1\}$.

155. Zauważmy, że $K := \ker f$ jest jedną z poziomic f: tą, która przechodzi przez element neutralny $e_1 \in G_1$. Jak wyglądają inne poziomice f ?

Użyjmy dla ustalenia uwagi notacji 'multiplikatywno-multiplikatywnej':

$$(a, b \in G_1 \text{ leżą na jednej } f\text{-poziomicy}) \Leftrightarrow f(a) = f(b) \Leftrightarrow$$
$$\Leftrightarrow e_2 = (f(a))^{-1}f(b) = f(a^{-1}b) \Leftrightarrow a^{-1}b \in K \Leftrightarrow b \in aK.$$

Odpowiedź. f-poziomicą przechodzącą przez element $a \in G_1$ jest zbiór $aK := \{ak : k \in K\}$, zwany *K-warstwą*. W szczególności *homomorfizm f jest injektywny $\iff$ K składa się jedynie z elementu neutralnego, tzn.* $\ker f = \{e_1\}$. Zauważmy, że odwzorowanie $K \ni k \mapsto ak \in aK$ jest bijekcją, a więc każda K-warstwa jest równoliczna z $K = \ker f$.

156. **Definicja**. Podgrupa $H \subset G$ nazywa się *podgrupą normalną*([23]), jeśli spełniony jest następujący warunek:

[23] Używane są również inne nazwy: *podgrupa niezmiennicza* oraz *dzielnik normalny*.

$$\forall g \in G:\ gHg^{-1} = H, \quad \text{tzn.} \quad \forall g \in G : \forall h \in H : ghg^{-1} \in H.$$

Zauważmy, że jeśli grupa G jest przemienna, to każda jej podgrupa jest normalna, bowiem $ghg^{-1} = gg^{-1}h = h$; z tego też powodu nie musimy się martwić, jak zapisać 'addytywną' wersję warunku normalności.

157. **Fakt**. Jeśli $f : G_1 \to G_2$ jest homomorfizmem grup, to:

(a) $\ker f$ jest podgrupą normalną grupy G_1;
(b) $\operatorname{im} f$ jest podgrupą (niekoniecznie normalną) grupy G_2.

Posłużmy się znów (dla wygody) notacją multiplikatywną dla grup G_1 i G_2. *Ad*(a) Jeśli $a, b \in \ker f$, tzn. $f(a) = f(b) = e_2$, to $f(ab) = f(a)f(b) = e_2e_2 = e_2$, więc $ab \in \ker f$, co dowodzi (1). Jeśli $a \in \ker f$, to $f(a^{-1}) = (f(a))^{-1} = {e_2}^{-1} = e_2$, więc $a^{-1} \in \ker f$, co dowodzi (2). Sprawdźmy warunek normalności dla $\ker f$: jeśli $h \in \ker f$, tzn. $f(h) = e_2$, to $f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)e_2f(g^{-1}) = = f(g)f(g^{-1}) = f(gg^{-1}) = f(e_1) = e_2$, czyli $f(ghg^{-1} = e_2$, a więc $ghg^{-1} \in \ker f$.

Ad(b) Jeśli $\alpha, \beta \in \operatorname{im} f$, tzn. $\exists a, b \in G_1 : \alpha = f(a)$, $\beta = f(b)$, to $\alpha\beta = f(a)f(b) = = f(ab) \in \operatorname{im} f$, a więc $\operatorname{im} f$ ma własność (1). Jeśli $\alpha \in \operatorname{im} f$, tzn. $\exists a : \alpha = f(a)$, to $\alpha^{-1} = (f_{(a)})^{-1} = f(a^{-1}) \in \operatorname{im} f$, co dowodzi własności (2).

3.4 Permutacje: rozkład na cykle i znak

158. **Fakt**. Jeśli $\sigma : X \to X$ jest permutacją zbioru skończonego X, to:

(1) Dla każdego ustalonego $x \in X$ okresowy jest ciąg o wyrazach

$$x_0 := x,\ x_1 := \sigma(x),\ x_2 := \sigma^2(x),\ x_3 := \sigma^3(x),\ \ldots .$$

(2) $x \underset{\sigma}{\sim} y \overset{\text{def.}}{\iff} \exists k \in \mathbb{Z}_+ : y = \sigma^k(x)$ jest relacją równoważności w X.

Ad(1) Skoro zbiór X jest skończony, to ciąg (x_k) nie jest różnowartościowy, więc $\exists k \geqslant 0, l > 0 : x_{k+l} = x_k$; lecz $x_{k+l} = \sigma^k(x_l)$, zaś równość $\sigma^k(x_l) = \sigma^k(x_0)$ wraz z injektywnością σ (a więc i σ^k) implikuje $x_l = x_0$. Aplikując σ^n do obu stron tej równości otrzymujemy $\forall n \geqslant 0 : x_{l+n} = x_n$, co oznacza l-okresowość ciągu.

Ad(2) *Zwrotność* jest oczywista, gdyż $\sigma^0(x) = \mathrm{id}_X(x) = x$. *Symetria*: Jeśli $x \underset{\sigma}{\sim} y$, tzn. $\exists k : y = \sigma^k(x)$, to biorąc $q \in \mathbb{N}$ tak duże, by $r := lq - k \geqslant 0$, dzięki l-okresowości ciągu (x_k) dostajemy $x = x_0 = x_{lq} = x_{k+r} = \sigma^r(x_k) = \sigma^r(y)$, a więc $y \underset{\sigma}{\sim} x$. *Przechodniość*: Jeśli $x \underset{\sigma}{\sim} y \underset{\sigma}{\sim} z$, to $\exists k, l \in \mathbb{Z}_+ : y = \sigma^k(x)$, $z = \sigma^l(y)$, a wtedy $z = \sigma^l(\sigma^k(x)) = \sigma^{k+l}(x)$, więc $x \underset{\sigma}{\sim} z$.

159. Klasy relacji równoważności $\underset{\sigma}{\sim}$ nazywa się *orbitami permutacji* σ lub, w skrócie, σ-*orbitami*. Jeśli np. $X = \overline{0,9}$ i $\sigma = \begin{pmatrix} 0&1&2&3&4&5&6&7&8&9 \\ 5&6&9&1&8&4&3&7&0&2 \end{pmatrix}$, to działanie σ można przedstawić schematycznie strzałkami

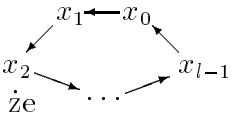
$$0 \mapsto 5 \mapsto 4 \mapsto 8 \mapsto 0, \quad 1 \mapsto 6 \mapsto 3 \mapsto 1, \quad 2 \mapsto 9 \mapsto 2, \quad 7 \mapsto 7,$$

a w takim razie σ-orbitami są w tym przykładzie podzbiory

$$X_1 = \{0,4,5,8\}, \quad X_2 = \{1,3,6\}, \quad X_3 = \{2,9\} \quad \text{i} \quad X_4 = \{7\}.$$

160. **Wnioski**

(1) Każde dwie σ-orbity są albo rozłączne, albo identyczne.
(2) Każda σ-orbita jest postaci $\{x_0, x_1, \ldots, x_{l-1}\}$, gdzie $l \in \mathbb{N}$ oraz
$$\sigma(x_{k-1}) = \begin{cases} x_k, \text{ dla } k = 1, 2, \ldots, l-1, \\ x_0, \text{ dla } k = l. \end{cases}$$
(3) Podzbiór $Y \subset X$ jest *σ-niezmienniczy* (tzn. spełnia warunek $\sigma(Y) = Y$) wtedy i tylko wtedy, gdy Y jest sumą mnogościową pewnej liczby σ-orbit; w szczególności σ-orbity są σ-niezmiennicze.

161. **Definicja**. Permutacja $\gamma \in S_X$ nazywa się *cyklem długości l*, jeśli istnieje l parami różnych elementów $x_0, x_1, \ldots, x_{l-1} \in X$, takich że
$$\gamma(x) = \begin{cases} x, & \text{dla } x \in X \setminus \{x_0, x_1, \ldots, x_{l-1}\}, \\ x_k, & \text{dla } x = x_{k-1},\ 1 \leqslant k < l, \\ x_0, & \text{dla } x = x_{l-1}. \end{cases}$$
Taki "cykl" γ wygodnie jest obrazować diagramem 

oraz oznaczać symbolem $(x_0\ x_1\ \ldots\ x_{l-1})$; zauważmy, że
$$(x_0\ x_1\ \ldots\ x_{l-1}) = (x_1\ x_2\ \ldots\ x_{l-1}\ x_0) = (x_2\ x_3\ \ldots\ x_{l-1} x_0\ x_1) = \ldots\ .$$
Warto również pamiętać, że symbol $(x_0\ x_1\ \ldots\ x_{l-1})$ nie określa całego zbioru X, na którym działa cykl, a więc np. $(1\ 2\ 5)$ równie dobrze może oznaczać permutację zbioru $\overline{1,5}$, jak i np. zbioru $\overline{1,10}$. Na szczęście ta nieprecyzyjność symbolu nie prowadzi do błędów rachunkowych.

162. Zbiór $\underline{\varrho} := \{x \in X : \varrho(x) \neq x\}$, tzn. dopełnienie zbioru punktów stałych ϱ, nazywamy *nośnikiem* permutacji $\varrho \in S_X$. Zauważmy, że zbiory $\underline{\varrho}$ oraz $X \setminus \underline{\varrho}$ są ϱ-niezmiennicze, co więcej, $\begin{cases} \varrho(\underline{\varrho}) = \underline{\varrho} \\ \varrho(X \setminus \underline{\varrho}) = X \setminus \underline{\varrho} \end{cases}$.

$\big(x \in \underline{\varrho}$, tzn. $\varrho(x) \neq x\big) \Leftrightarrow \varrho(\varrho(x)) \neq \varrho(x) \ \Leftrightarrow\ \varrho(x) \in \underline{\varrho}$; stąd też wynika drugi wzór.

163. **Fakt**. Jeśli dwie permutacje mają rozłączne nośniki, to są przemienne:
$$\underline{\varrho_1} \cap \underline{\varrho_2} = \emptyset \quad \Rightarrow \quad \varrho_1 \circ \varrho_2 = \varrho_2 \circ \varrho_1.$$
Niech $X_k := \underline{\varrho_k}$, $k \in \{1, 2\}$, wtedy $X = X_1 \cup X_2 \cup X_0$, gdzie X_0 jest dopełnieniem $X_1 \cup X_2$; skoro ϱ_1 na $X_2 \cup X_0$ jest identycznością, to każdy z tych trzech zbiorów jest ϱ_1-niezmienniczy, a także ϱ_2-niezmienniczy. Otóż dla $x \in X_1$ mamy $\varrho_2(\varrho_1(x)) = \varrho_1(x)$ (bo $\varrho_1(x) \in X_1$) oraz $\varrho_1(\varrho_2(x)) = \varrho_1(x)$ (bo $\varrho_2(x) = x$), skąd $\varrho_1(\varrho_2(x)) = \varrho_2(\varrho_1(x))$. Podobnie jest dla $x \in X_2$, zaś na X_0 obie permutacje są identycznościami.

164. **Ćwiczenie**. Sprawdzić, że $\boxed{\underline{\varrho \circ \sigma} \subset \underline{\varrho} \cup \underline{\sigma}}$, $\boxed{\underline{\varrho^{-1}} = \underline{\varrho}}$ i $\boxed{\underline{\varrho \circ \sigma \circ \varrho^{-1}} = \varrho(\underline{\sigma})}$, więc permutacje o skończonych nośnikach tworzą podgrupę (i to normalną) grupy S_X.

165. **Twierdzenie**. Jeśli X jest zbiorem skończonym, to każdą permutację $\sigma \in S_X$ można przedstawić jako złożenie cykli rozłącznych; przy tym taki rozkład jest jednoznaczny z dokładnością do kolejności cykli.

Istnienie rozkładu: Niech $X = X_1 \cup \ldots \cup X_r$ będzie rozkładem X na σ-orbity; określmy $\gamma_1, \ldots, \gamma_r$ wzorami
$$\gamma_i(x) = \begin{cases} \sigma(x), & x \in X_i \\ x, & x \in X \setminus X_i \end{cases}. \qquad (*)$$
Wtedy z 160. jest widoczne, że γ_i są rozłącznymi cyklami, oraz że $\sigma = \gamma_1 \circ \ldots \circ \gamma_r$.

Jednoznaczność rozkładu: Wystarczy oczywiście sprawdzić, że jeśli $\gamma_1, \ldots, \gamma_r$ są rozłącznymi cyklami, to znając $\sigma := \gamma_1 \circ \ldots \circ \gamma_r$ można poszczególne γ_i odtworzyć powyższymi wzorami (*), biorąc jako X_i σ-orbity. Istotnie, rozłączność cykli γ_i sprawia, że $\gamma_2, \ldots, \gamma_r$ nie ruszają punktów zbioru $X_1 := \underline{\gamma}_1$, więc σ pokrywa się z γ_1 na X_1, co dowodzi zarówno wzoru (*), jak również tego, że X_1 jest σ-orbitą.

166. **Fakt**. Każdą permutację $\sigma \in S_X$ można przedstawić w postaci złożenia pewnej liczby *transpozycji*, tzn. cykli długości 2.

Wobec 165. wystarczy pokazać, że każdy cykl $\gamma \in S_X$ można przedstawić w postaci złożenia pewnej liczby transpozycji; otóż łatwo w tym celu sprawdzić, że

$$(x_1\ x_2\ \ldots\ x_r) = (x_1\ x_2)(x_2\ x_3)\ \ldots\ (x_{r-1}\ x_r).$$

Oczywiście rozkład na transpozycje nie jest jednoznaczny, gdyż np.

$$(1\ 2)(2\ 3)(1\ 2) = (1\ 3).$$

167. **Twierdzenie**. Jeśli $\sigma = \tau_1\ \ldots\ \tau_r$ oraz $\sigma = \tau'_1\ \ldots\ \tau'_s$ są dwoma rozkładami permutacji σ na iloczyn transpozycji, to $(-1)^r = (-1)^s$, tzn. r i s mają jednakową parzystość. W konsekwencji sensowna jest

168. **Definicja**. *Znakiem* permutacji σ nazywa się liczbę $\operatorname{sgn}(\sigma) := (-1)^r$, gdzie $r \in \mathbb{N}$ *jest liczbą czynników dowolnego rozkładu σ na iloczyn transpozycji*. Wprost z takiego określenia liczby $\operatorname{sgn}(\sigma)$ wynika wzór

$$\operatorname{sgn}(\sigma \circ \varrho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\varrho),$$

a więc *odwzorowanie* $\operatorname{sgn} : S_X \to \{-1, +1\}$ *jest homomorfizmem grup.* Jego jądro oznacza się zwykle symbolem A_X:

$$A_X := \ker(\operatorname{sgn} : S_X \to \{\pm 1\}) = \{\sigma \in S_X : \operatorname{sgn}(\sigma) = 1\};$$

jak wiemy z 157., A_X jest podgrupą normalną grupy S_X. Elementy z A_X, tzn. takie, że $\operatorname{sgn}(\sigma) = 1$, nazywa się permutacjami *parzystymi*; pozostałe permutacje z S_X są *nieparzyste*. Zamiast $S_{\overline{1,n}}$ i $A_{\overline{1,n}}$ stosuje się krótsze symbole S_n i A_n; tradycyjnie S_n nosi nazwę *grupy symetrycznej stopnia n*, natomiast A_n — *grupy alternującej stopnia n*.

Poniższy dowód twierdzenia 167. oprzemy na następującym lemacie:

169. **Lemat**. Oznaczmy symbolem $N(\sigma)$ liczbę orbit permutacji $\sigma \in S_X$. Jeśli $\tau = (x\ \ y) \in S_X$ jest transpozycją, to

$$N(\tau\sigma) - N(\sigma) = \begin{cases} -1, & \text{gdy } x \overset{\sigma}{\not\sim} y, \\ +1, & \text{gdy } x \overset{\sigma}{\sim} y. \end{cases}$$

Niech $\sigma = \gamma_1 \ldots \gamma_N$ będzie rozkładem σ na cykle rozłączne, zawierającym również cykle długości 1 (równe id_X), odpowiadające punktom stałym σ; wtedy σ-orbitami są zbiory wyrazów poszczególnych cykli γ_i, a więc $N =$ (liczba czynników) $= N(\sigma)$.

1° Jeśli $x \not\sim y$, to x, y są wyrazami dwóch cykli rozkładu, powiedzmy γ_1 i γ_2; oznaczmy $\gamma_1 = (x_1\ \ldots\ x_r)$, $\gamma_2 = (y_1\ \ldots\ y_s)$ tak, aby $x_1 = x$, $y_1 = y$, wtedy

$$\tau\gamma_1\gamma_2 = (x_1\ y_1)(x_1\ \ldots\ x_r)(y_1\ \ldots\ y_s) = (x_1\ \ldots\ x_r\ y_1\ \ldots\ y_s) =: \gamma;$$

zatem rozkład $\tau\sigma$ ma $N-1$ czynników, gdyż powstaje z rozkładu σ przez 'sklejenie' dwóch cykli γ_1, γ_2 w jeden cykl γ. Permutacja $\tau\sigma$ ma więc $N-1$ orbit.

2° Jeśli $x \underset{\sigma}{\sim} y$, to $x,\ y$ są wyrazami jednego z cykli γ_i, powiedzmy γ_N; oznaczmy jego wyrazy tak, aby $\gamma_N = (x_1\ \ldots\ x_r\ y_1\ \ldots\ y_s)$ oraz $x_1 = x$ i $y_1 = y$, wtedy

$$\tau\gamma_N = (x_1\ y_1)(x_1\ \ldots\ x_r\ y_1\ \ldots\ y_s) = (x_1\ \ldots\ x_r)(y_1\ \ldots\ y_s) =: \gamma'_N\gamma'_{N+1},$$

przy czym cykle γ'_N i γ'_{N+1} są rozłączne i zawierają razem wszystkie wyrazy γ_N. Zatem rozkład $\tau\sigma$ ma $N+1$ cykli, czyli permutacja $\tau\sigma$ ma $N+1$ orbit.

170. Dowód twierdzenia 167.
Z lematu wynika istnienie takich liczb $\epsilon_1, \epsilon_2, \ldots, \epsilon_r$ ze zbioru $\{-1, +1\}$, że

$$N(\sigma) = N(\tau_1\tau_2\ \ldots\ \tau_r) = \epsilon_1 + N(\tau_2\tau_3\ \ldots\ \tau_r) = \\ = \epsilon_1 + \epsilon_2 + N(\tau_3\ \ldots\ \tau_r) = \ldots = \epsilon_1 + \ldots + \epsilon_r + N(\mathrm{id}_X);$$

przy czym oczywiście $N(\mathrm{id}_X) = |X|$ (gdyż orbity permutacji id_X są 1-elementowe). Ponieważ $(-1)^{\epsilon_i} = (-1)^{\pm 1} = -1$, z powyższego wzoru otrzymujemy

$$(-1)^{N(\sigma)} = (-1)\cdot\ldots\cdot(-1)(-1)^{|X|} = (-1)^r(-1)^{|X|}, \quad \text{tzn.} \quad (-1)^r = (-1)^{|X|-N(\sigma)}.$$

W taki sam sposób rozkładu $\sigma = \tau'_1 \ldots \tau'_s$ wynika $(-1)^s = (-1)^{|X|-N(\sigma)} = (-1)^r$.

171. **Wniosek.** Jeśli $\sigma = \gamma_1 \circ \ldots \circ \gamma_r$, gdzie γ_k jest cyklem długości l_k, to

$$\operatorname{sgn}\sigma = (-1)^{l_1-1} \cdot \ldots \cdot (-1)^{l_r-1}.$$

3.5 Liczba inwersji a znak permutacji

Niekiedy znalezienie rozkładu σ na cykle jest niełatwe lub nawet praktycznie niewykonalne; w takich przypadkach do obliczenia $\operatorname{sgn}(\sigma)$ można się posłużyć pojęciem tzw. *σ-inwersji*. Zbiór σ-inwersji zależy nie tylko od permutacji $\sigma : X \to X$, do jego zdefiniowania potrzebne jest jeszcze *liniowe uporządkowanie* zbioru X, tzn. przypisanie kolejnych numerów $1, \ldots, n$ jego elementom: $X = \{x_1, \ldots, x_n\}$.

Okazuje się, że od uporządkowania X zależy nie tylko zbiór σ-inwersji J_σ, ale nawet jego moc $|J_\sigma|$, zaś tym, co zależy wyłącznie od σ (i wiąże się ze znakiem σ) jest parzystość liczby $|J_\sigma|$.

Dla uproszczenia zapisu (aby pisać $1, 2, \ldots$ zamiast $x_1, x_2, \ldots$) przyjmiemy, że X jest przedziałem $\overline{1,n}$ z jego naturalnym uporządkowaniem.

172. Dla $\sigma \in S_n$ oznaczmy $J_\sigma := \{(i,j) : i,j \in \overline{1,n},\ i < j,\ \sigma(i) > \sigma(j)\}$; pary $(i,j) \in J_\sigma$ nazywać będziemy *σ-inwersjami*. Wprowadźmy też następujący zbiór transpozycji:

$$T_n := \{(j-1\ j) : j \in \overline{2,n},\ i < j\} \subset S_n.$$

173. **Fakt.** σ da się przedstawić w postaci złożenia $|J_\sigma|$ transpozycji z T_n, otrzymujemy więc następujący alternatywny wzór na znak permutacji:

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

Nazwijmy (chwilowo) 'defektem' ciągu liczbowego $s_1, \ldots, s_n$ taką parę (s_i, s_j), że $i < j$, lecz $s_i > s_j$. Wtedy oczywiste będą kolejno następujące konstatacje:

(1) Wskutek przestawienia dwóch **sąsiednich** wyrazów ciągu, tworzących defekt, np. wyrazów 8 i 3 w ciągu $2, 7, 8, 3, 1, 5, 4, 6$, liczba defektów zmniejsza się o 1;
(2) $(i,j) \in J_\sigma \iff$ para $(\sigma(i), \sigma(j))$ jest defektem ciągu $\sigma(1), \ldots, \sigma(n)$, więc $|J_\sigma|$ jest liczbą defektów tego ostatniego ciągu (kolejnych wartości σ). W takim razie,
(3) jeśli $(j-1, j) \in J_\sigma$, to dla $\tau := (j-1\ j)$ mamy $|J_{\sigma\tau}| = |J_\sigma| - 1$, gdyż ciąg

$\sigma\tau(1),\ldots,\sigma\tau(n)$ powstaje z ciągu $\sigma(1),\ldots,\sigma(n)$ przez przestawienie wyrazów $\sigma(j-1),\sigma(j)$, tworzących defekt.

(4) Jeśli $\mathcal{J}_\sigma \neq \emptyset$, to $\exists j \in \overline{2,n} : (j-1,j) \in J_\sigma$, w przeciwnym bowiem razie mielibyśmy $\sigma(1) < \sigma(2) < \ldots < \sigma(n)$, co jest możliwe tylko dla $\sigma = \mathrm{id}_X$, zaś $J_{\mathrm{id}_X} = \emptyset$. Możemy więc skorzystać $N := |J_\sigma|$ razy z (3), a zatem

(5) istnieją $\tau_1,\ldots\tau_N \in T_n$, takie że $|J_\sigma| = N$, $|J_{\sigma\tau_1}| = N-1$,...,$|J_{\sigma\tau_1\ldots\tau_r}| = N-r$ dla $r \in \overline{1,N}$. W szczególności $\sigma\tau_1\ldots\tau_N$ nie ma inwersji, więc $\sigma\tau_1\ldots\tau_N = \mathrm{id}_X$, a wobec tego, dzięki własności ${\tau_i}^{-1} = \tau_i$, otrzymujemy $\sigma = \tau_N\ldots\tau_1$, Q.E.D.

174. **Przykład**. Dla $n = 2m$ i $\sigma := \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & 2m \\ 2 & 4 & \ldots & 2m & 1 & 3 & \ldots & 2m-1 \end{pmatrix}$ σ-inwersjami są następujące pary:

$(1,m+1),(2,m+1),(2,m+2),(3,m+1),(3,m+2),(3,m+3),\ldots,$

tzn. $J_\sigma = \{(i,m+j) : i \in \overline{1,m},\ j \in \overline{1,i}\}$, skąd $|J_\sigma| = \sum\limits_{i=1}^{m} i = \frac{m(m+1)}{2}$, więc z powyższego wzoru dostajemy $\boxed{\mathrm{sgn}(\sigma) = (-1)^{\frac{1}{2}m(m+1)}}$.

Warto zaznaczyć, że znalezienie rozkładu σ na rozłączne cykle, choć całkiem łatwe dla konkretnych niedużych wartości $n = 2m$, jest zadaniem wręcz karkołomnym w przypadku ogólnym, gdy liczba m nie jest określona konkretnie.

175. **Uwaga**. Zauważmy, że dla $\varrho,\sigma \in S_n$ zachodzi nierówność ([24])

$$|J_{\varrho\sigma}| \leqslant |J_\varrho| + |J_\sigma|,$$

więc, przez indukcję, $|J_{\tau_1\ldots\tau_n}| \leqslant \sum_i |J_{\tau_i}|$, skąd $|J_{\tau_1\ldots\tau_n}| \leqslant r$ dla $\tau_i \in T_n$. Zatem fakt poprzedni można uzupełnić, dodając że

σ nie da się rozłożyć na mniej niż $|J_\sigma|$ transpozycji z T_n.

3.6 Rząd grupy i elementu grupy; warstwy

176. *Rzędem* grupy G nazywamy liczbę elementów (tzn. moc) zbioru G; oznaczamy ją — jak zwykle — symbolem $|G|$ albo $\#G$.

177. **Przykład**. Grupa S_n ma rząd równy $n\cdot(n-1)\cdot\ldots\cdot2\cdot1$, czyli $|S_n| = n!$.

Chcąc skonstruować permutację $\sigma \in S_n$ możemy wartość $\sigma(1) \in \overline{1,n}$ wybrać na n sposobów; mając $\sigma(1)$ możemy $\sigma(2) \in \overline{1,n} \setminus \{\sigma(1)\}$ wybrać na $n-1$ sposobów; mając $\sigma(1)$ i $\sigma(2)$ możemy $\sigma(3) \in \overline{1,n} \setminus \{\sigma(1),\sigma(2)\}$ wybrać na $n-2$ sposobów, itd.

178. **Rząd elementu grupy**. Mając ustalony element a grupy G określmy homomorfizm $f_a : \mathbb{Z} \to G$ wzorem $f_a(k) := a^k$. Jego obraz, tzn. zbiór $\mathrm{im}\, f_a = \{a^k : k \in \mathbb{Z}\} = \langle a\rangle$, jest podgrupą w G, generowaną przez a. Jądro f_a, tzn. zbiór $\ker f_a = \{k : a^k = 1\}$, jest podgrupą (normalną) grupy $\mathbb{Z}$. Każda podgrupa $\mathbb{Z}$ jest oczywiście ideałem w $\mathbb{Z}$, więc (skoro pierścień $\mathbb{Z}$ jest dziedziną ideałów głównych) ma postać $n\mathbb{Z}$ (krotności ustalonej liczby n), przy czym $n \in \mathbb{Z}_+$ jest określone jednoznacznie. Zatem $\exists n \in \mathbb{Z}_+ : \ker f_a = n\mathbb{Z}$; jeśli $n \neq 0$, to liczbę n nazywa się *rzędem elementu* a i oznacza symbolem $\mathrm{ord}\, a$. Wprost z definicji mamy:

[24] Albowiem, jak łatwo zauważyć, $(i,j) \in J_{\varrho\sigma} \Rightarrow (i,j) \in J_\sigma$ lub $(\sigma(i),\sigma(j)) \in J_\varrho$.

- $a^k = 1 \iff k \in n\mathbb{Z}$, tzn. $n \mid k$; ogólniej,
$$a^k = a^l \iff a^{l-k} = 1 \iff k \equiv l \bmod n;$$
- $n = \min\{k \in \mathbb{N} : a^k = 1\}$;
- $n = \#\langle a\rangle =$ rząd podgrupy $\langle a\rangle$,

gdyż $a^1, \ldots, a^n$ są parami różne, a ciąg $(a^k)_{k\in\mathbb{Z}}$ jest n-okresowy.

Gdy $n = 0$, tzn. gdy $\ker f_a = \{0\}$, ciąg $(a^k)_{k\in\mathbb{Z}}$ jest różnowartościowy; mówimy wtedy, że *a jest rzędu nieskończonego* i piszemy $\operatorname{ord} a = \infty$.

Przykład. Jeśli $\varrho = \gamma_1 \circ \ldots \gamma_r$ jest rozkładem permutacji $\varrho \in S_X$ na parami rozłączne cykle $\gamma_1, \ldots, \gamma_r \in S_X$ o długościach $l_1, \ldots, l_r$, to $\operatorname{ord}\varrho = \mathrm{NWK}(l_1, \ldots, l_r)$. Wynika to stąd, że $\varrho^n = \gamma_1^n \circ \ldots \gamma_r^n = \mathrm{id}_X \iff \gamma_1^n = \ldots = \gamma_r^n = \mathrm{id}_X$ (gdyż permutacje γ_i^n, tak jak γ_i, mają rozłączne nośniki) $\iff \forall i : l_i \mid n$.

179. **Indeks podgrupy; warstwy.** Mając podgrupę $H \subset G$ określmy w zbiorze G relację $\boxed{a \underset{H}{\sim} b \overset{\text{def}}{\iff} a^{-1}b \in H}$; jest ona zwrotna (bo $a^{-1}a = 1 \in H$), symetryczna (bo gdy $a^{-1}b \in H$, to $b^{-1}a = \left(a^{-1}b\right)^{-1} \in H$) oraz przechodnia (jeśli $a^{-1}b, b^{-1}c \in H$, to $a^{-1}c = a^{-1}b \cdot b^{-1}c \in H$). Relację $\underset{H}{\sim}$ nazywa się *(lewą) H-równoważnością*, a jej klasy abstrakcji $[a]_H = \{b \in G : a \underset{H}{\sim} b$, tzn. $b \in aH\} = aH$ — *warstwami lewostronnymi względem podgrupy H*. Zatem (z własności relacji równoważności) każde dwie warstwy aH i bH są albo równe (gdy $a \underset{H}{\sim} b$) albo rozłączne; ich sumą jest cały zbiór G. Liczbę (ogólniej: moc zbioru) warstw lewostronnych względem H nazywa się *indeksem podgrupy H w grupie G* i oznacza symbolem $(G : H)$. ([25])

Przykłady. Skoro $1H = H1 = H$, to zawsze jedną jedną z H-warstw (tą, która zawiera element neutralny) jest H. Jeśli grupa G jest przemienna, to oczywiście $aH = Ha$, czyli lewe i prawe H-warstwy są takie same.

- Dla $G = \mathbb{Z}$ (musimy tu 'przestawić się' na notację addytywną) oraz $H = n\mathbb{Z}$, gdzie $n \in \mathbb{N}$, dwie H-warstwy, $H + x$ i $H + y$, są równe $\Leftrightarrow\ y \in H + x \ \Leftrightarrow\ x \equiv y \bmod n$. Zatem każda z H-warstw jest postaci $H + x$ dla pewnego $x \in \overline{0, n-1}$, przy czym warstwy $H, H+1, \ldots, H+n-1$ są parami różne. Zatem H-warstw jest n:
$$H = \{\ldots, -n, 0, n, 2n, \ldots\}, \quad H + 1 = \{\ldots, -n+1, 1, n+1, 2n+1, \ldots\},$$
$$H + 2 = \{\ldots, -n+2, 2, n+2, 2n+2, \ldots\}, \ H + 3 = \{\ldots, -n+3, 3, n+3, \ldots\},$$
$$\ldots, H + (n-1) = \{\ldots, -2n-1, -n-1, -1, n-1, 2n-1, \ldots\}.$$
W takim razie $(\mathbb{Z} : n\mathbb{Z}) = n$, czyli $n\mathbb{Z}$ jest podgrupą indeksu n.

- Niech A_n będzie podzbiorem grupy S_n, złożonym z permutacji parzystych, tzn. $A_n = \{\sigma \in S_n : \operatorname{sgn}\sigma = 1\} = \ker\big(\operatorname{sgn} : S_n \to \{-1, 1\}\big)$; jest to podgrupa grupy S_n, tradycyjnie nazywana *grupą alternującą stopnia n*.

  Pokażemy, że $(S_n : A_n) = 2$, a więc $|A_n| = \frac{1}{2}|S_n| = \frac{1}{2}\,n!$.

  Niech τ będzie ustalonym elementem S_n takim, że $\operatorname{sgn}\tau = -1$. Wtedy $\sigma \mapsto \sigma \circ \tau$ jest bijekcją A_n na dopełnienie A_n w S_n, więc $|A_n| = |S_n \setminus A_n|$, skąd wynika teza.

  Wzór $(S_n : A_n) = 2$ można też bez trudu uzyskać, stosując następujący

[25] Oprócz 'lewej' równoważności i 'lewych' H-warstw czasami przydatne są ich prawe odpowiedniki: $a \overset{H}{\sim} b \overset{\text{def}}{\iff} ab^{-1} \in H$, $[b]^H = Hb$; jednak ze wzoru $(ah)^{-1} = h^{-1}a^{-1}$ łatwo wynika, że warstwa Ha^{-1} jest zbiorem odwrotności elementów warstwy aH, więc odwzorowanie $\kappa : G \to G$, $\kappa(x) := x^{-1}$, odwzorowuje bijektywnie lewe H-warstwy na prawe i *vice versa*; w szczególności wynika stąd, że prawych H-warstw jest tyleż, co lewych, więc nie ma potrzeby definiowania 'lewego i prawego indeksu'.

180. **Fakt**. Jeśli $\varphi : G \to G_1$ jest homomorfizmem grup, przy czym grupa $\operatorname{im} G$ jest skończona, to $\boxed{(G : \ker\varphi) = |\operatorname{im}\varphi|}$.

Niech $H := \ker\varphi$, wtedy $\varphi(a) = \varphi(b) \Leftrightarrow \varphi(a^{-1}b) = e \Leftrightarrow a \underset{H}{\sim} b \Leftrightarrow aH = bH$, więc $\varphi(a) \mapsto aH$ jest poprawnie zdefiniowaną bijekcją $\operatorname{im}\varphi$ na zbiór H-warstw.

3.7 Twierdzenie Lagrange'a i jego konsekwencje

181. **Fakt** (*twierdzenie Lagrange'a*). Jeśli G jest grupą skończoną, a H — jej podgrupą, to $|G| = (G : H) \cdot |H|$; zatem $|H|$ dzieli $|G|$, tzn.

rząd podgrupy jest dzielnikiem rzędu grupy.

Wynika to natychmiast stąd, że zbiór G jest sumą wszystkich H-warst; jest ich $(G : H)$, są one parami rozłączne i każda ma elementów tyleż, co H, gdyż przy ustalonym $a \in G$ odwzorowanie $H \ni h \mapsto ah \in aH$ jest bijekcją H na aH.

182. **Wniosek**. Jeśli $a \in G$, a G jest skończona, to ord a jest dzielnikiem $|G|$, tzn. $\boxed{a^{|G|} = 1}$; w notacji addytywnej: $\boxed{|G|a = (|G|\text{-krotność } a) = 0}$.

$n := \operatorname{ord} a$ jest rzędem podgrupy $H = \langle a\rangle$, więc dzieli $|G|$; zatem $|G| \in n\mathbb{Z} = \ker f_a$.

183. **Wniosek**. Grupa skończona, której rząd jest liczbą pierwszą, jest *grupą cykliczną*, tzn. generowaną przez jeden ze swoich elementów: $\exists\, a \in G : G = \langle a\rangle$. Zatem jeśli $|G| = p$ jest liczbą pierwszą, to $G \cong \mathbb{Z}_p$.

Jeśli $a \in G$ i $a \neq 1$, to rząd podgrupy $\langle a\rangle$ jest różnym od 1 dzielnikiem liczby pierwszej p, więc $\big|\langle a\rangle\big| = p = |G|$, tj. $\langle a\rangle = G$. Ponieważ $a^k = a^l \Leftrightarrow k \equiv l \bmod p \Leftrightarrow k + p\mathbb{Z} = l + p\mathbb{Z}$, więc wzór $f(k + p\mathbb{Z}) := a^k$ poprawnie definiuje odwzorowanie $f : \mathbb{Z}_p \to G$; bez trudu sprawdzamy, że f jest izomorfizmem grup.

184. **Wniosek**. Jeśli $\varphi : G \to G_1$ jest homomorfizmem grup, a grupa G jest skończona, to

$$\boxed{|\ker\varphi| \cdot |\operatorname{im}\varphi| = |G|}.$$

Jest to natychmiastowa konsekwencja twierdzenia Lagrange'a i faktu 180.

185. **Wniosek**. Każda skończona podgrupa grupy multiplikatywnej $\mathbb{C}^*$ jest postaci $\sqrt[n]{1}$ dla pewnego $n \in \mathbb{N}$.

Jeśli $H \subset \mathbb{C}^*$ — podgrupa rzędu n, to $\forall a \in H : a^n = 1$, tzn. $H \subset \sqrt[n]{1}$; zarazem oba te zbiory są n-elementowe.

186. **Wniosek** (*twierdzenie Eulera* z teorii liczb). Tzw. *funkcję Eulera* $\varphi : \mathbb{N} \to \mathbb{N}$ określa się wzorem $\varphi(n) := \#\{k \in \overline{1,n} : \mathrm{NWD}(k,n) = 1\}$.

$$\underline{\text{Teza}}\text{:}\quad \forall x \in \mathbb{Z} :\ \mathrm{NWD}(x,n) = 1 \quad \Rightarrow \quad \left(\begin{array}{c} x^{\varphi(n)} \equiv 1 \bmod n \\ \text{tzn. } n \mid \left(x^{\varphi(x)} - 1\right) \end{array}\right).$$

Zauważmy, że $\big(x+n\mathbb{Z}$ jest odwracalny w $\mathbb{Z}_n\big) \Leftrightarrow \big(\exists y \in \mathbb{Z} :\ (x+n\mathbb{Z})(y+n\mathbb{Z}) = 1+n\mathbb{Z}$, tzn. $xy \equiv 1 \bmod n\big) \Leftrightarrow \big(\exists y, z \in \mathbb{Z} : xy - nz = 1\big) \Leftrightarrow \mathrm{NWD}(x,n) = 1$; zatem grupa (multiplikatywna) elementów odwracalnych pierścienia $\mathbb{Z}_n$ ma postać

$\mathbb{Z}_n^* = \{x + n\mathbb{Z} : x \in \overline{1,n},\ \mathrm{NWD}(x,n) = 1\}$. Wobec tego $|G| = \varphi(n)$ dla $G = \mathbb{Z}_n^*$, zaś $a^r = x^r + n\mathbb{Z}$ dla $a = x + n\mathbb{Z} \in G$, więc teza wynika wprost z wniosku 182.

187. **Wniosek** (*małe twierdzenie Fermata*). Jeżeli $p \in \mathbb{N}$ jest liczbą pierwszą, zaś $x \in \mathbb{Z}$ — liczbą niepodzielną przez p, to $\boxed{x^{p-1} \equiv 1 \bmod p}$.

Wszystkie $k \in \overline{1, p-1}$ są niepodzielne przez p, więc $\varphi(p) = p - 1$, skąd teza.

Na przykład przez $p = 11$ podzielne są liczby $2^{10} - 1 = 1023$, $3^{10} - 1 = 59048$, $4^{10} - 1 = 1048575$, $5^{10} - 1 = 9765624$, $6^{10} - 1 = 60466175$. W istocie, mając dość cierpliwości można sprawdzić, że mają one rozkłady $3 \cdot 11 \cdot 31$, $2^3 \cdot 11^2 \cdot 61$, $3 \cdot 5^2 \cdot 11 \cdot 31 \cdot 41$, $2^3 \cdot 3 \cdot 11 \cdot 71 \cdot 521$, $5^2 \cdot 7 \cdot 11 \cdot 101 \cdot 311$.

3.8 Twierdzenie Cayleya

188. **Fakt**. Jeśli X i Y są zbiorami (skończonymi), to

$$(\text{grupy } S_X \text{ i } S_Y \text{ są izomorficzne}) \Leftrightarrow \begin{pmatrix}\text{zbiory } X \text{ i } Y \text{ są równoliczne,}\\ \text{tzn. } |X| = |Y|\end{pmatrix}.$$

W szczególności grupa S_X jest izomorficzna z grupą S_n dla $n := |X|$.

$\boxed{\Rightarrow}$ wynika z tego, że liczba wszystkich bijekcji $X \to X$, równa $n!$, jest rosnącą (a więc i injektywną) funkcją $n := |X|$. $\boxed{\Leftarrow}$ Ustalmy jakąś bijekcję $\varphi : X \to Y$. Zauważmy, że każdemu odwzorowaniu $\sigma : X \to X$ odpowiada (dokładnie jedno) odwzorowanie $\tilde\sigma : Y \to Y$, dla którego diagram $\begin{array}{ccc} X & \xrightarrow{\varphi} & Y \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\tilde\sigma} \\ X & \xrightarrow{\varphi} & Y \end{array}$ "jest przemienny", przez co rozumie się równość $\boxed{\varphi \circ \sigma = \tilde\sigma \circ \varphi}$; istotnie, skoro φ ma odwrotność, to ostatni warunek da się "rozwiązać względem $\tilde\sigma$", otrzymując $\boxed{\tilde\sigma = \varphi \circ \sigma \circ \varphi^{-1}}$. Wychodząc z takiej inspiracji określmy odwzorowanie $f : S_X \to (\text{funkcje } Y \to Y)$ wzorem $f(\sigma) := \tilde\sigma = \varphi \circ \sigma \circ \varphi^{-1}$. Zauważmy najpierw, że $\boxed{f(\sigma \circ \varrho) = f(\sigma) \circ f(\varrho)}$ (*), gdyż $\varphi \circ (\sigma \circ \varrho) \circ \varphi^{-1} = (\varphi \circ \sigma \circ \varphi^{-1}) \circ (\varphi \circ \varrho \circ \varphi^{-1})$; zarazem $f(\mathrm{id}_X) = \mathrm{id}_Y$ (gdyż ...), zatem $f(\sigma) \circ f(\sigma^{-1}) = \ldots = f(\mathrm{id}_X) = \mathrm{id}_Y$; tak samo $f(\sigma^{-1}) \circ f(\sigma) = \mathrm{id}_Y$, więc $f(\sigma)$ **jest bijekcją**! Tym samym sprawdziliśmy, że f jest odwzorowaniem $S_X \to S_Y$, spełniającym warunek (*), tzn. jest homomorfizmem grup. Pozostaje jeszcze zauważyć, że f jest bijekcją, bowiem ma odwrotność: $f^{-1}(\tilde\sigma) = \varphi^{-1} \circ \tilde\sigma \circ \tilde\sigma$.

189. **Fakt** (*twierdzenie Cayleya*). Każdą grupę skończoną G dla dostatecznie dużego $n \in \mathbb{N}$ można *zanurzyć homomorficznie* w grupę S_n. Oznacza to, że istnieje **injektywny** homomorfizm $\varphi : G \to S_n$, a w takim razie $\varphi : G \to \tilde G$ jest izomorfizmem G na podgrupę $\tilde G = \mathrm{im}\,\varphi$ grupy S_n.

Przyporządkujmy elementowi $a \in G$ odwzorowanie $L_a : G \to G$, określone wzorem $L_a(g) := ag$. Zauważmy, że $L_a \circ L_b = L_{ab}$, wobec tego L_a ma odwrotność: $L_a \circ L_{a^{-1}} = L_{a^{-1}} \circ L_a = \mathrm{id}_G$, czyli L_a jest bijekcją, tzn. permutacją zbioru G. Mamy więc odwzorowanie $f : G \to S_G$, $f(a) := L_a$, spełniające warunek $f(ab) = f(a) \circ f(b)$, tzn. będące homomorfizmem. Zauważmy, że f jest injektywne: jeśli $f(a) = L_a = \mathrm{id}_G$, to w szczególności $L_a(e) = e$, tzn. $ae = e$, więc $a = e$, zob. 155. Można teraz dla $n := |G|$ określić homomorficzne zanurzenie $\varphi : G \to S_n$, składając zanurzenie $f : G \to S_G$ z izomorfizmem $S_G \to S_n$, opisanym w poprzednim punkcie.

3.9 Appendix: Inny sposób definiowania znaku permutacji

190. Niech X — zbiór skończony; weźmy $X^2_* := \{(x,y) : x,y \in X, x \neq y\}$ oraz rodzinę $\mathcal{A}$ wszystkich funkcji $p : X^2_* \to \{-1,1\}$ *antysymetryczych*, tzn. takich, że $p(x,y) = -p(y,x)$. Mając $\sigma \in S_X$ określmy dla $p \in \mathcal{A}$ wzorem $p^\sigma(x,y) := p(\sigma(x),\sigma(y))$ funkcję $p^\sigma : X^2_* \to \{-1,1\}$; jasne jest, że $p^\sigma \in \mathcal{A}$ oraz $(p^\sigma)^\rho = p^{\sigma\rho}$, gdyż p^σ to w istocie "p złożone z σ". Zauważmy, że wielkość $\langle p|q\rangle := \prod_{r(x,y)=1} p(x,y)q(x,y)$ dla $p,q \in \mathcal{A}$ nie zależy od $r \in \mathcal{A}$, ponieważ $s = pq$ jest symetryczna: $s(x,y) = s(y,x)$.

191. **Fakt**. Dla $p,q,\tilde{p},\tilde{q} \in \mathcal{A}$ oraz $\sigma \in S_X$ zachodzą następujące tożsamości:

$$\langle p^\sigma|q^\sigma\rangle = \langle p|q\rangle, \quad \langle p|q\rangle \cdot \langle \tilde{p}|\tilde{q}\rangle = \langle p|\tilde{q}\rangle \cdot \langle \tilde{p}|q\rangle$$

L $= \prod_{r(\sigma(x),\sigma(y))=1} p(\sigma(x),\sigma(y))q(\sigma(x),\sigma(y)) = \prod_{r(\tilde{x},\tilde{y})=1} p(\tilde{x},\tilde{y})q(\tilde{x},\tilde{y}) =$ P; drugi wzór jest oczywisty.

Wobec tego $\langle p^\sigma|p\rangle \cdot \langle q^\sigma|q\rangle = \langle p^\sigma|q^\sigma\rangle \cdot \langle p|q\rangle = (\pm 1)^2 = 1$, a więc wielkość $\langle p^\sigma|p\rangle \in \{-1,1\}$ nie zależy od wyboru $p \in \mathcal{A}$, czyli ma sens definicja

$$\boxed{\operatorname{sgn}(\sigma) := \langle p^\sigma|p\rangle, \text{ gdzie } p \in \mathcal{A} \text{ — dowolny.}}$$

192. **Fakt**. (a) Określone w taki sposób odwzorowanie $\operatorname{sgn} : S_X \to \{-1,1\}$ jest homomorfizmem grup, tzn. $\operatorname{sgn}(\sigma\rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\varrho)$ dla $\varrho,\sigma \in S_X$.
(b) $\operatorname{sgn}(\tau) = -1$ dla dowolnej transpozycji $\tau \in S_X$.

(a) $\operatorname{sgn}(\sigma\rho) = \langle p^{\sigma\rho}|p\rangle = \langle q^\rho|p\rangle = \langle q^\rho|q\rangle \cdot \langle q|p\rangle = \operatorname{sgn}(\rho)\operatorname{sgn}(\sigma)$, gdzie $q := p^\sigma$.

(b) Zgodnie z definicją $\operatorname{sgn}(\tau) = \langle p^\tau|p\rangle = (-1)^k$, gdzie k jest połową liczby elementów zbioru $A_\tau := \{(x,y) \in X^2_* : p(\tau(x),\tau(y)) \neq p(x,y)\}$; jeśli w szczególności τ jest transpozycją elementów $a,b \in X$, to oznaczając elementy zbioru X jako $x_1,x_2,\ldots,x_n$ w taki sposób, by $x_1 = a$, $x_2 = b$ oraz określając $p \in \mathcal{A}$ wzorem $p(x_i,x_j) := \begin{cases} +1, & \text{gdy } i < j \\ -1, & \text{gdy } i > j \end{cases}$, otrzymujemy $A_\tau = \{(x_1,x_2),(x_2,x_1)\}$, czyli $k = 1$.

193. **Wniosek**. Jeśli permutacja $\sigma \in S_X$ jest złożeniem r cykli parzystej długości oraz s cykli nieparzystej długości, to $\operatorname{sgn}(\sigma) = (-1)^r$.

Jeśli $\gamma = (x_1\ x_2\ \ldots\ x_l)$ jest cyklem długości l, to $\gamma = (x_1\ x_2)(x_2\ x_3)\ldots(x_{l-1}\ x_l)$ jest złożeniem l transpozycji, a więc $\operatorname{sgn}(\gamma) = (-1)^l$; to wraz z 192.(a) dowodzi tezy.

4 Przestrzenie wektorowe

4.1 Definicja przestrzeni wektorowej. Przykłady

Niech $\mathbb{K}$ będzie ustalonym ciałem. Będziemy zakładać zwykle, że jest to ciało liczbowe, tzn. $\mathbb{K} \subseteq \mathbb{C}$; co więcej, w zdecydowanej większości zastosowań w zupełności wystarczy ograniczyć się do przypadków $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$.

194. **Definicja**. *Przestrzenią wektorową nad ciałem* $\mathbb{K}$ nazywa się trójkę $(V, +, \cdot)$, w której

- V jest zbiorem,
- $+$ jest operacją w V, taką, że $(V, +)$ jest grupą przemienną;
- $\cdot$ jest odwzorowaniem $\cdot : \mathbb{K} \times V \to V$, spełniającym cztery następujące aksjomaty([26]):

$$\left.\begin{array}{ll} (1) & (\lambda+\mu)\cdot v = \lambda\cdot v + \mu\cdot v; \\ (2) & \lambda\cdot(v+w) = \lambda\cdot v + \lambda\cdot w; \\ (3) & (\lambda\mu)\cdot v = \lambda\cdot(\mu\cdot v); \\ (4) & 1\cdot v = v; \end{array}\right\} \begin{array}{l} \forall v, w \in V, \\ \forall \lambda, \mu \in \mathbb{K}. \end{array}$$

Zwykle używa się terminu 'przestrzeń wektorowa' odnośnie do zbioru V, traktując oczywiste człony trójki, '$+$' i '$\cdot$', jako domyślne — sporadycznie się zdarza, że z jakiegoś powodu '$+$' i '$\cdot$' należy zastąpić innymi znaczkami. Elementy zbioru V nazywa się *wektorami*, elementy $\mathbb{K}$ — *skalarami*, zaś zbiór $\mathbb{K}$ — *ciałem skalarów* przestrzeni. Zamiast 'przestrzeń nad ciałem $\mathbb{R}$ (lub $\mathbb{C}$)' mówi się 'przestrzeń rzeczywista (lub zespolona)'. Odwzorowanie $\cdot$ nazywa się *operacją mnożenia wektorów przez skalary*; często znak '$\cdot$' pomiędzy liczbą a wektorem pomija się; wygodnie też umówić się, że '$v\lambda$' oznacza to samo, co 'λv', tzn. '$\lambda \cdot v$'.

195. **Fakt**. Jeśli V jest przestrzemią wektorową, $v \in V$, $\lambda \in \mathbb{K}$, to:

$$\text{(a)}\ \ 0\cdot v = \mathbf{0}; \qquad \text{(b)}\ \ \lambda\cdot\mathbf{0} = \mathbf{0}; \qquad \text{(c)}\ \ (-\lambda)\cdot v = -(\lambda\cdot v);$$
$$\text{(d)}\ \ \lambda\cdot v = \mathbf{0} \iff (\lambda = 0 \text{ lub } v = \mathbf{0}).$$

Powyżej $\mathbf{0}$ oznacza element neutralny (zero) grupy $(V,+)$; nazywa się go *wektorem zerowym* i zwykle oznacza tak, jak liczbę zero, symbolem 0 (z kontekstu na ogół jest jasne, czy chodzi o liczbę, czy o wektor).

Ad (a): $w := 0\cdot v$, wtedy $w + w = (0+0)\cdot v = 0\cdot v = w$ wskutek (1), więc dzięki prawu skracania w grupie dostajemy $w = 0$. *Ad* (b): $w := \lambda\cdot\mathbf{0}$, wtedy $w + w = \lambda\cdot(\mathbf{0}+\mathbf{0}) = \lambda\cdot\mathbf{0} = w$ wskutek (2), co jak poprzednio daje $w = 0$. *Ad* (c): Dzięki (1) i (a) mamy: $(-\lambda)\cdot v + \lambda\cdot v = \big((-\lambda)+\lambda\big)\cdot v = 0\cdot v = \mathbf{0}$, a więc $(-\lambda)\cdot v$ jest elementem przeciwnym do $\lambda\cdot v$. *Ad* (d): Jeśli $\lambda\cdot v = \mathbf{0}$ i ponadto $\lambda \neq 0$, to wskutek (b) mamy: $v = (\lambda^{-1}\lambda)\cdot v = \lambda^{-1}\cdot(\lambda\cdot v) = \lambda^{-1}\cdot\mathbf{0} = \mathbf{0}$.

196. **Przykłady**.

1. Dla $n \in \mathbb{N}$ niech $\mathbb{K}^n$ będzie iloczynem kartezjańskim $\mathbb{K} \times \ldots \times \mathbb{K}$ (n razy); elementami $\mathbb{K}^n$ są więc n-elementowe ciągi $\boldsymbol{x} = (x_1, \ldots, x_n)$

[26] Jak zwykle, dla uproszczenia zapisu umawiamy się, że 'mnożenie ma pierwszeństwo przed dodawaniem', co oznacza np. że $\lambda\cdot v + \mu\cdot w$ jest uproszczonym zapisem $(\lambda\cdot v)+(\mu\cdot w)$.

elementów z $\mathbb{K}$. Określmy dodawanie elementów $\mathbb{K}^n$ wzorem

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n),$$

natomiast mnożenie elementów $\mathbb{K}^n$ przez liczbę $\lambda \in \mathbb{K}$ — wzorem

$$\lambda(x_1, \ldots, x_n) := (\lambda x_1, \ldots, \lambda x_n).$$

Łatwo się przekonać — korzystając z aksjomatów ciała — że spełnione są wówczas wszystkie warunki definicji przestrzeni wektorowej, np. sprawdzenie (2) wygląda następująco:

$$\begin{aligned}\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda((x_1, \ldots, x_n) + (y_1, \ldots, y_n)) = \lambda(x_1 + y_1, \ldots, x_n + y_n) = \\ = (\lambda(x_1 + y_1), \ldots, \lambda(x_n + y_n)) = (\lambda x_1 + \lambda y_1, \ldots, \lambda x_n + \lambda y_n) = \\ = (\lambda x_1, \ldots, \lambda x_n) + (\lambda y_1, \ldots, \lambda y_n) = \\ = \lambda(x_1, \ldots, x_n) + \lambda(y_1, \ldots, y_n) = \lambda\boldsymbol{x} + \lambda\boldsymbol{y}.\end{aligned}$$

W tzw. *rachunku macierzowym* używane są dwa różne 'egzemplarze' przestrzeni $\mathbb{K} \times \ldots \times \mathbb{K}$, różniące się od siebie (tylko) sposobem zapisu elementów, i przez to z definicji rozłączne:

- przestrzeń ${\mathbb{K}^n}_1$ *wektorów kolumnowych*, postaci $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$,
- przestrzeń ${\mathbb{K}^1}_n$ *wektorów wierszowych*, postaci $\boldsymbol{f} = \begin{bmatrix} f_1 & \ldots & f_n \end{bmatrix}$.

2. Jeśli X jest dowolnym zbiorem, to symbolem $\mathbb{K}^X$ oznaczmy zbiór, złożony z wszystkich funkcji $v : X \to K$. Wprowadźmy w tym zbiorze działanie 'zwykłego' (często zwanego *punktowym*) dodawania funkcji:

$$\forall x \in X : \quad (v_1 + v_2)(x) := v_1(x) + v_2(x);$$

inaczej mówiąc, jeśli $v, v_1, v_2 \in \mathbb{K}^X$, to relacja $v = v_1 + v_2$ oznacza *z definicji*, że $\forall x \in X : v(x) = v_1(x) + v_2(x)$. Podobnie określmy iloczyn funkcji $v \in \mathbb{K}^X$ przez liczbę $\lambda \in \mathbb{K}$:

$$\forall x \in X : \quad (\lambda v)(x) := \lambda v(x).$$

Łatwo sprawdzić, że w ten sposób $\mathbb{K}^X$ staje się przestrzenią wektorową.

3. Zbiór $\mathbb{K}[\,\cdot\,]$ wielomianów o współczynnikach z $\mathbb{K}$ tworzy przestrzeń wektorową (dla zdefiniowanych w rozdziale o wielomianach działań dodawania i mnożenia wielomianów przez liczby).

197. **Definicja**. Podzbiór $V_0 \subset V$ nazywa się *podprzestrzenią* przestrzeni V, jeśli jest niepusty oraz 'zamknięty względem działań w V', tzn.

$\forall v_1, v_2 \in V_0 : \; v_1 + v_2 \in V_0$ (zamkniętość wzgl. dodawania);
$\forall \lambda \in \mathbb{K}, \; v \in V_0 : \; \lambda v \in V_0$ (zamkniętość wzgl. mnożenia przez liczby).

Każda podprzestrzeń zawiera wektor zerowy: $v_0 \in V_0 \Rightarrow \mathbf{0} = 0 \cdot v_0 \in V_0$. Zauważmy, że poprzestrzeń V_0 jest przestrzenią wektorową względem działań (dodawania i mnożenia przez liczby) 'takich samych, jak w V'.

Mówiąc ściśle, operacja dodawania w V_0 jest określona jako *obcięcie* odwzorowania 'dodawanie w V': $V \times V \to V$, do podzbioru $V_0 \times V_0 \subset V \times V$; zamkniętość V_0 względem dodawania sprawia, że obcięcie to ma wartości tylko w V_0, a więc określa odwzorowanie $V_0 \times V_0 \to V_0$. Podobnie rzecz ma się z mnożeniem przez liczby.

198. **Dalsze przykłady.**

4. Dla dowolnego zbioru X podzbiór

$$E(X) := \{v \in \mathbb{K}^X : \{x : v(x) \neq 0\} \text{ jest skończony}\}$$

jest przestrzenią wektorową, mianowicie podprzestrzenią $\mathbb{K}^X$. Aby to pokazać, wystarczy sprawdzić, że $E(X)$ jest podzbiorem zamkniętym

- względem dodawania: jeśli $v_1, v_2 \in \mathbb{K}^X$ oraz $v = v_1 + v_2$, to
$$\{x : v(x) \neq 0\} \subset \{x : v_1(x) \neq 0\} \cup \{x : v_2(x) \neq 0\},$$
a suma zbiorów skończonych — gdy $v_1, v_2 \in E(X)$ — jest skończona;
- względem mnożenia: otóż zbiór $\{x : \lambda v(x) \neq 0\}$ jest dla $\lambda \neq 0$ równy $\{x : v(x) \neq 0\}$ (więc dla $v \in E(X)$ skończony), a dla $\lambda = 0$ — pusty.

5. Elementy przestrzeni $\mathbb{K}^{\mathbb{N}}$, tzn. funkcje $\xi : \mathbb{N} \to \mathbb{K}$, wygodnie jest przedstawiać nieskończonymi ciągami $\boldsymbol{x} = (x_1, x_2, \ldots)$ ich wartości $x_k = \xi(k) \in \mathbb{K}$; działania w $\mathbb{K}^{\mathbb{N}}$ (dodawanie i mnożenie przez liczby) są wtedy zwykłymi działaniami na ciągach.

Opierając się na podstawowych własnościach ciągów, granic i szeregów łatwo jest sprawdzić, że dla $\mathbb{K} = \mathbb{R}$ (jak również dla $\mathbb{K} = \mathbb{C}$) podzbiór

$$V_0 := \{\boldsymbol{x} \in \mathbb{K}^{\mathbb{N}} : \mathcal{W}\}$$

jest podprzestrzenią przestrzeni $\mathbb{K}^{\mathbb{N}}$, jeśli $\mathcal{W}$ jest którymkolwiek z następujących (przykładowych) warunków ([27]):

- $\forall k > 100 : x_k = 0$;
- (x_k) jest ciągiem arytmetycznym;
- istnieje skończona granica $\lim\limits_{k\to\infty} x_k$;
- $\lim\limits_{k\to\infty} k^3 x_k = 0$;
- $\exists a, b \in \mathbb{K} : \lim\limits_{k\to\infty} (x_k - ak - b) = 0$;
- szereg $\sum\limits_{k=1}^{\infty} |x_k|$ jest zbieżny;
- $\forall k : x_{k+100} = x_k$ (okresowość);
- $\forall^* k : \ x_{k+1} = x_k$;
- $\forall k, l : \ x_{kl} = \frac{1}{2}(x_k + x_l)$;
- $\forall^* k : \ x_k = 0$;
- $\forall k : \ x_{k+1} = \frac{1}{2} x_k$;
- $\lim\limits_{k\to\infty} x_k = 0$;
- ciąg $(\frac{x_k}{k})$ jest ograniczony;
- szereg $\sum\limits_{k=1}^{\infty} x_k$ jest zbieżny;
- szereg $\sum\limits_{k=1}^{\infty} |x_k|^2$ jest zbieżny;
- $\exists p \in \mathbb{N} : \forall k : x_{k+p} = x_k$;
- $\forall k : x_{k+2} = x_k + x_{k+1}$;
- $\forall k : x_{k^2} = \frac{1}{3} x_k$.

Z kolei V_0 *nie jest* podprzestrzenią $\mathbb{K}^{\mathbb{N}}$ dla następujących warunków $\mathcal{W}$:

- $\exists k > 100 : x_k = 0$;
- $\forall k : x_k x_{k+2} = {x_{k+1}}^2$ (geometryczny);
- $\lim\limits_{k\to\infty} x_k = 4$;
- $\forall k : \ x_{k^2} = (x_k)^2$;
- zero jest punktem skupienia (x_k);
- $\forall^* k : \ x_k = \frac{1}{k}$;
- $\forall k : \ |x_k| \leqslant 12$;
- ciąg $(|x_k|)$ jest malejący;
- $\sum\limits_{k=1}^{\infty} x_k = \frac{\pi}{2}$;
- $\exists^* k : \ x_k = 0$.

199. **Definicja.** Podzbiór $A + B := \{a + b : \ a \in A, \ b \in B\}$ nazywa się *sumą*

[27] Dla uproszczenia notacji używamy poniżej symboli 'uogólnionych kwantyfikatorów' $\forall^* k$ (oznaczającego *dla prawie wszystkich* k) oraz $\exists^* k$ (*dla nieskończenie wielu* k).

algebraiczną podzbiorów $A, B \subset V$; podobnie definiuje się podzbiory λA, $\lambda A + \mu B$ itp., np zbiór $3A - 2B$ składa się z wektorów, dających się przedstawić w postaci $3a - 2b$ dla pewnych $a \in A$, $b \in B$.

Jasne, że $V_0 \subset V$ jest zamknięty względem dodawania $\Leftrightarrow$ $V_0 + V_0 \subset V_0$, a wzgl. mnożenia przez liczby $\Leftrightarrow$ $\mathbb{K} \cdot V_0 \subset V_0$, tzn. $\forall \lambda \in \mathbb{K} : \lambda V_0 \subset V_0$.

Zatem $\left(\begin{array}{c}\text{podzbiór } V_0 \subset V \text{ jest}\\ \text{podprzestrzenią } V\end{array}\right) \iff \left(\begin{array}{c}V_0 + V_0 \subset V_0\\ \text{oraz } \mathbb{K} \cdot V_0 \subset V_0\end{array}\right)$.

200. **Fakt**. Przecięcie dowolnej liczby podprzestrzeni jest podprzestrzenią.
Suma algebraiczna dwóch podprzestrzeni jest podprzestrzenią.

Dowód stanowi proste ćwiczenie.

4.2 Przestrzenie ilorazowe

201. A teraz dość zabawny, lecz pożyteczny przepis na zrobienie przestrzeni wektorowej. Potrzebne są dwa 'surowce': przestrzeń wektorowa V i jakaś jej podprzestrzeń V_0.

Określmy w zbiorze V relację $v \sim v' \overset{\text{def}}{\iff} v' - v \in V_0$; jest to oczywiście relacja równoważności, więc zadaje klasy: $[v] := \{u \in V : u \sim v\} \subset V$, $[v] = [v'] \iff v \sim v'$. Niech $\tilde{V} = V/_\sim = \{[v] : v \in V\}$ oznacza zbiór wszystkich tych klas. Uczynimy go przestrzenią wektorową, określając w nim:

dodawanie: $[v_1] + [v_2] := [v_1 + v_2]$,
oraz
mnożenie przez skalary: $\lambda[v] := [\lambda v]$ dla $\lambda \in \mathbb{K}$.

Oczywiście należy sprawdzić, co jest to bardzo łatwe, sensowność tych definicji, tzn. że $\left\{\begin{array}{l}[u_1] = [v_1]\\ [u_2] = [v_2]\end{array}\right\}$ implikuje $[u_1 + u_2] = [v_1 + v_2]$, zaś $[u] = [v]$ implikuje $[\lambda v] = [\lambda u]$.

Następnie równie łatwo przekonujemy się, że zbiór $\tilde{V}$ z tak określonymi działaniami jest przestrzenią wektorową, tzn. spełnia wszystkie aksjomaty definicji 194.

202. **Definicja** (*przestrzeń ilorazowa*). Skonstruowana w powyższy sposób przestrzeń wektorowa nazywa się *przestrzenią ilorazową* (lub krócej: *ilorazem*) przestrzeni V przez jej podprzestrzeń V_0.

A oto standardowe oznaczenia:

$\tilde{V}$	$v \sim v'$	$[v]$
V/V_0	$v \equiv v' \bmod V_0$	$v + V_0$

Dlaczego użyliśmy innych oznaczeń, czyżby standardowe były niedobre? Ależ wręcz przeciwnie, są one poniekąd aż za dobre...Kłopot polega na tym, że np. definicja dodawania w postaci $(v_1 + V_0) + (v_2 + V_0) = (v_1 + v_2) + V_0$ wygląda bardziej na tożsamość, niż na definicję; w istocie jest to tożsamość dzięki własności $V_0 + V_0 = V_0$.

Podobnie: klasa $[v]$ *jest* zbiorem $v + V_0 = \{v + v_0 : v_0 \in V_0\} = \{u : u - v \in V_0\} \subset V$.

203. **Przykłady**. (1) Całka nieoznaczona $\int f(x)dx$ z jakiejś funkcji $f \in V := C(]a, b[)$ w gruncie rzeczy nie jest funkcją, czyli elementem V, ale elementem przestrzeni ilorazowej V/V_0, gdzie $V_0 \subset V$ oznacza podzbiór funkcji stałych na $]a, b[$.

(2) Ogólniej: jeśli $X \subset \mathbb{R}$ jest podzbiorem otwartym, np. sumą mnogościową kilku

otwartych przedziałów, to całka nieoznaczona $\int f(x)dx$ z funkcji $f \in V := C(X)$ (ciągłe $X \to \mathbb{R}$) jest elementem przestrzeni V/V_0, gdzie tym razem $V_0 :=$ (funkcje lokalnie stałe na X) (czyli stałe na każdej ze składowych spójnych X).

(3) Całkę nieoznaczoną $\int e^x(1+x^2+y^2)^{-1}dx$ można traktować (uzasadnić!) jako element przestrzeni ilorazowej V/V_0, gdzie $V = C(\mathbb{R}^2)$, $V_0 = C(\mathbb{R})$.

4.3 Kombinacje liniowe. Powłoka liniowa podzbioru

204. **Definicja**. *Kombinacją liniową* wektorów $v_1, \ldots, v_s \in V$ nazywa się wektor $v \in V$, dający się przedstawić w postaci

$$v = \lambda_1 v_1 + \ldots + \lambda_s v_s,$$

gdzie $\lambda_1, \ldots, \lambda_s$ są jakimiś elementami $\mathbb{K}$, nazywanymi *współczynnikami kombinacji liniowej*. Symbol $\mathrm{Lin}(v_1, \ldots, v_s)$, lub krótszy $\langle v_1, \ldots, v_s \rangle$, oznacza zbiór wszystkich kombinacji liniowych wektorów $v_1, \ldots, v_s$, tzn.

$$\langle v_1, \ldots, v_s \rangle := \{v \in V : \exists \lambda_1, \ldots, \lambda_s \in \mathbb{K} : v = \lambda_1 v_1 + \ldots + \lambda_s v_s\} \subset V.$$

205. **Fakt**. Podzbiór $\langle v_1, \ldots, v_s \rangle$

1° jest podprzestrzenią przestrzeni V;
2° zawiera każdy z wektorów $v_1, \ldots, v_s$;
3° jest najmniejszą z podprzestrzeni zawierających $v_1, \ldots, v_s$.

*Ad*1°: Suma kombinacji liniowych $\lambda_1 v_1 + \ldots + \lambda_s v_s$ i $\mu_1 v_1 + \ldots + \mu_s v_s$ jest kombinacją liniową $(\lambda_1 + \mu_1)v_1 + \ldots + (\lambda_s + \mu_s)v_s$, więc $\langle v_1, \ldots, v_s \rangle$ jest zamknięty względem dodawania; podobnie krotność λv wektora $v = \lambda_1 v_1 + \ldots + \lambda_s v_s$ jest wektorem $(\lambda\lambda_1)v_1 + \ldots + (\lambda\lambda_s)v_s \in \langle v_1, \ldots, v_s \rangle$ (zamkniętość wzgl. mnożenia przez skalary).

*Ad*2° : $v_1 = 1 \cdot v_1 + 0 \cdot v_2 + \ldots 0 \cdot v_s$ jest kombinacją liniową v_i; podobnie $v_2 = \ldots$.

*Ad*3°: Jeśli W jest jakąkolwiek podprzestrzenią V, zawierającą każdy z wektorów v_i, to każda kombin. liniowa $v = \lambda_1 v_1 + \ldots + \lambda_s v_s$ należy do W, więc $\langle v_1, \ldots, v_s \rangle \subset W$.

206. Zbiór $\langle v_1, \ldots, v_s \rangle$ nazywa się podprzestrzenią *rozpiętą* (lub *generowaną*) *przez wektory* $v_1, \ldots, v_s$, albo też *powłoką liniową* tych wektorów.

207. **Przykład**. Gdy $v_1, v_2 \in \mathbb{R}^3$, wtedy $\langle v_1, v_2 \rangle$ jest:
(a) płaszczyzną przechodzącą przez $0, v_1$ i v_2, gdy punkty $0, v_1, v_2$ nie są współliniowe;
(b) prostą przechodzącą przez $0, v_1$ i v_2, gdy te punkty są współliniowe, przy czym choć jeden z wektorów v_1 lub v_2 jest niezerowy;
(c) zbiorem $\{0\}$, gdy $v_1 = v_2 = 0$.

208. **Uwaga**. Dla dowolnego (niekoniecznie skończonego) podzbioru $Z \subset V$ istnieje najmniejsza podprzestrzeń zawierająca Z; jest nią mianowicie przecięcie wszystkich podprzestrzeni zawierających Z; oznaczamy ją symbolem $\mathrm{Lin}(Z)$ lub $\langle Z \rangle$ i nazywamy *podprzestrzenią rozpiętą* (lub *generowaną*) *przez zbiór* Z lub *powłoką liniową zbioru* Z. Oczywiście $\langle Z \rangle = \langle v_1, \ldots, v_s \rangle$ dla $Z = \{v_1, \ldots, v_s\}$. W ogólnym przypadku $\langle Z \rangle$, tak jak poprzednio, jest zbiorem kombinacji liniowych wektorów z Z, dokładniej: kombinacji liniowych *skończonych* układów wektorów z Z:

$$v \in \langle Z\rangle \Leftrightarrow \exists s \in \mathbb{N} : \exists v_1, \ldots, v_s \in Z : \left\{ \begin{array}{c} v \in \langle v_1, \ldots, v_s\rangle, \text{ tzn.} \\ \exists \lambda_i : v = \lambda_1 v_1 + \ldots + \lambda_s v_s \end{array} \right\}.$$

Oczywiście $\langle \emptyset \rangle = \{0\}$; z tego powodu wygodnie jest przyjąć umowę, że wyrażenie «v jest kombinacją liniową wektorów z $\emptyset$» oznacza «$v = 0$».

209. **Fakt**. 1° $Z_1 \subset Z_2 \quad \Rightarrow \quad \langle Z_1\rangle \subset \langle Z_2\rangle$ (*monotoniczność* operacji $\langle\,\cdot\,\rangle$);

2° $Z_1 \subset Z_2 \subset \langle Z_1\rangle \quad \Rightarrow \quad \langle Z_2\rangle = \langle Z_1\rangle$; w szczególności:

3° $\forall i:\ v_i \in \langle u_1, \ldots, u_r\rangle \ \Rightarrow\ \langle u_1, \ldots, u_r, v_1, \ldots\rangle = \langle u_1, \ldots, u_r\rangle$.

Ad2° Skoro $Z_2 \subset \langle Z_1\rangle$, to $\langle Z_1\rangle$ jest jedną (zaś $\langle Z_2\rangle$ — najmniejszą) z podprzestrzeni zawierających Z_2; zatem $\langle Z_2\rangle \subset \langle Z_1\rangle$, a monotoniczność daje odwrotne zawieranie.
Ad3° Wystarczy zastosować 2°, biorąc $Z_1 = \{u_1, \ldots, u_r\}$, $Z_2 = Z_1 \cup \{v_1, \ldots, v_s\}$.

210. **Uwaga**. Łatwo jest zauważyć, że jeśli $V_1, V_2 \subset V$ są podprzestrzeniami V, to

$$V_1 + V_2 = \langle V_1 \cup V_2\rangle.$$

Zatem właściwym i naturalnym uogólnieniem operacji sumy algebraicznej na przypadek dowolnej rodziny podprzestrzeni $V_\alpha \subset V$, $\alpha \subset A$, jest następująca definicja

$$\sum_{\alpha \in A} V_\alpha := \langle \bigcup_{\alpha \in A} V_\alpha \rangle.$$

4.4 Liniowa niezależność

211. **Definicja**. Układ $u_1, \ldots, u_r$ wektorów z V nazywamy *liniowo zależnym*, jeśli pewna *nietrywialna* (tzn. mająca choć jeden współczynnik $\neq 0$) kombinacja liniowa tych wektorów jest wektorem zerowym:

$$\exists \lambda_1, \ldots, \lambda_r \in \mathbb{K} : (\lambda_1, \ldots, \lambda_r) \neq (0, \ldots, 0),\ \lambda_1 u_1 + \ldots + \lambda_r u_r = 0.$$

Układ wektorów nazywamy *liniowo niezależnym*([28]), jeśli nie jest liniowo
zależny, tzn. jeśli dla każdych $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ prawdziwa jest implikacja

$$\lambda_1 u_1 + \ldots + \lambda_r u_r = 0 \quad \Rightarrow \quad \lambda_1 = 0,\ \ldots,\ \lambda_r = 0;$$

można to także wyrazić warunkiem, że *każda nietrywialna kombinacja liniowa wektorów $u_1, \ldots, u_r$ jest wektorem różnym od wektora zerowego.*

212. **Przykłady**. Jeśli jeden z wektorów u_j jest zerowy, układ jest liniowo zależny, np. gdy $u_1 = 0$, wtedy biorąc $\lambda_1 = 1$, $\lambda_2 = \ldots = \lambda_r = 0$ dostaniemy nietrywialną kombinację przedstawiającą wektor 0.
Gdy dwa z wektorów układu są jednakowe, wtedy układ jest liniowo zależny, np. przy $u_1 = u_2$ nietrywialna komb. liniowa o współczynnikach $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_j = 0$ dla $j > 2$, przedstawia wektor 0.
Układ wektorów mający liniowo zależny podukład sam jest liniowo zależny, bowiem (nietrywialna) komb. liniowa wektorów podukładu jest

[28] Zamiast «liniowo (nie)zależny układ wektorów» zwykło się mówić także «liniowo (nie)zależne wektory»; należy jednak pamiętać, że *liniowa (nie)zależność* nie jest cechą poszczególnych wektorów, lecz cechą ich wzajemnych relacji, czyli cechą całego układu.

(nietrywialną) komb. liniową wektorów całego układu. Równoważny fakt: *Podukład układu liniowo niezależnego jest liniowo niezależny.*

213. Oczywiście liniowa (nie)zależność układu nie zależy od kolejności wektorów; dzięki temu możemy zdefiniować pojęcie liniowej (nie)zależności zbioru wektorów:

Definicja. Podzbiór $Z \subset V$ jest liniowo (nie)zależny, jeśli liniowo (nie)zależny jest układ $u_1, \ldots, u_r$ wszystkich elemntów Z, ustawionych w dowolnej kolejności bez powtórzeń. Inaczej mówiąc:

Jeśli wektory $u_1, \ldots, u_r$ są *parami różne*, to **zbiór** $\{u_1, \ldots, u_r\}$ nazywa się *liniowo (nie)zależny*, jeśli liniowo (nie)zależny jest **układ** $u_1, \ldots, u_r$.

Zwróćmy uwagę, że zbiór $Z = \{u_1, u_2, u_3\}$ może być l. niezależny nawet w przypadku, gdy układ u_1, u_2, u_3 jest liniowo zależny, np. gdy u_1, u_2 są l. niezależne, a $u_2 = u_3$ (wtedy $\#Z = 2$), albo gdy $u_1 = u_2 = u_3 \neq 0$ (wtedy $\#Z = 1$).

214. **Uogólnienie**. O zbiorze $Z \subset V$ (niekoniecznie skończonym) mówi się, że jest *liniowo niezależny*, jeśli każdy skończony układ $u_1, \ldots, u_r$ parami różnych elementów Z jest liniowo niezależny. Tak więc zbiór Z jest *liniowo zależny* $\iff$ ma skończony liniowo zależny podzbiór.

215. **Lemat**. Dla $u_1, \ldots, u_r \in V$ trzy następujące warunki są równoważne:
1° układ $u_1, \ldots, u_r$ jest liniowo zależny;
2° jeden z wektorów u_j jest kombin. liniową wcześniejszych wektorów (zgodnie z 208. gdy 'jeden z wektorów' to u_1, oznacza to, że $u_1 = 0$);
3° jeden z wektorów u_j jest kombinacją liniową pozostałych wektorów.

$\boxed{1° \Rightarrow 2°}$ Niech $\lambda_1 u_1 + \ldots + \lambda_r u_r = 0$ oraz niech λ_j będzie ostatnim (tzn. o maksym. indeksie) niezerowym współczynnikiem. Tak więc $\lambda_j \neq 0$ oraz $\lambda_1 u_1 + \ldots + \lambda_j u_j = 0$. Dla $j > 1$ wynika stąd, że $u_j = (-\frac{\lambda_1}{\lambda_j})u_1 + \ldots + (-\frac{\lambda_{j-1}}{\lambda_j})u_{j-1}$; dla $j = 1$ — że $u_1 = 0$.
$\boxed{2° \Rightarrow 3°}$ — oczywiste. $\boxed{3° \Rightarrow 1°}$ Jeśli wektor u_j (gdzie $j \in \overline{1,r}$ — ustalone) jest kombinacją liniową pozostałych wektorów, tzn. $u_j = \sum_{i \in \overline{1,r} \setminus \{j\}} \alpha_i u_i$, to dodając obustronnie $-u_j$ dostajemy $0 = \alpha_1 u_1 + \ldots + \alpha_j u_j + \ldots + \alpha_r u_r$, gdzie $\alpha_j := -1$.

216. **Wniosek**. Każdy skończony układ([29]) (lub zbiór) wektorów zawiera podukład (podzbiór) liniowo niezależny, rozpinający tę samą przestrzeń.

Przeprowadźmy 'lustrację' kolejnych (od lewej) wektorów układu, usuwając te z nich, które są liniową kombinacją swych poprzedników. Dzięki $\neg 2° \Rightarrow \neg 1°$ dostaniemy układ l. niezależny, a na mocy faktu 3°p.209 nie zmienimy przestrzeni rozpinanej.

217. **Wniosek**. Dla podzbioru $Z \subset V$ następujące warunki są równoważne:

(a) Z jest liniowo zależny;
(b) pewien element Z jest kombinacją liniową innych elementów Z;
(c) istnieje podzbiór $Z_0 \subsetneq Z$, taki że $\langle Z_0 \rangle = \langle Z \rangle$.

(a)$\Rightarrow$(b): Zbiór l. zależny zawiera skończony układ l. zależny $u_1, \ldots, u_r \in Z$, więc możemy skorzystać z implikacji $1° \Rightarrow 2°$. (b)$\Rightarrow$(c): Wystarczy wziąć $Z_0 := Z \setminus \{v\}$,

[29] Teza pozostaje prawdziwa bez założenia o skończoności, lecz dowód wymaga wtedy zastosowania aksjomatu wyboru, np. w formie 'lematu Kuratowskiego-Zorna'; szkic tego dowodu przedstawimy dalej w punkcie 243.

gdzie $v \in Z$ jest kombinacją liniową innych elementów Z, i zastosować 2°p.209. (c)⇒(a): Weźmy dowolny $v \in Z \setminus Z_0$; skoro $v \in \langle Z_0 \rangle$, to v jest komb. liniową pewnych parami różnych wektorów $u_1, \ldots, u_r \in Z_0$; wobec $3° \Rightarrow 1°$ układ $u_1, \ldots, u_r, v$ jest l. zależny, ma elementy parami różne i należące do Z; zatem Z jest l. zależny.

4.5 Baza

218. **Definicja.** Układ wektorów $v_1, \ldots, v_n$ nazywamy *bazą (uporządkowaną)* przestrzeni V, jeśli jest liniowo niezależny i rozpina całą przestrzeń V.

219. **Uwaga terminologiczna**. W żargonie matematycznym przymiotniki *minimalny* i *najmniejszy* mają odmienne znaczenia, nie zawsze zgodne z sensem potocznym: *Najmniejszy* oznacza 'mniejszy/równy od wszystkich innych', przy czym zamiast '$\leqslant$' może być relacja zawierania (*najmniejsza podprzestrzeń $\langle Z \rangle$ zawierająca zbiór Z*) lub podzielności (*najmniejszy wspólny dzielnik*) lub inna 'relacja porządkująca'. Z kolei *minimalny* to 'nie mający «mniejszych» od siebie', a więc — innymi słowy — 'nie dający się zmniejszyć w obrębie rozważanej klasy'.([30])
Warto zauważyć, że jeśli porównywanie dotyczy zawierania $\subset$, to zbiór *minimalny* w danej klasie nie musi mieć minimalnej (w tej klasie) liczby elementów. Np. jeśli przyjmiemy, że marchew zawiera witaminy A i C, kapusta — B i C, burak — C i D, a brukiew — B, C i D, to *minimalnym zestawem pełnowitaminowym* będzie zarówno {marchew, brukiew}, jak również {marchew, kapusta, burak} (2 lub 3 elementy).

220. **Ćwiczenie.** Następujące cztery warunki są równoważne:

1° $v_1, \ldots, v_n$ jest bazą przestrzeni V;
2° $v_1, \ldots, v_n$ jest maksymalnym układem liniowo niezależnym w V;
3° $v_1, \ldots, v_n$ jest minimalnym układem rozpinającym przestrzeń V;
4° Każdy wektor $v \in V$ ma dokładnie jedno przedstawienie postaci

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n, \quad \lambda_1, \ldots, \lambda_n \in \mathbb{K}.$$

Ze względów dydaktycznych sprawdzimy równoważność 1° z każdym spośród warunków 2°, 3° i 4° z osobna, nie ograniczając się do minimalnej liczby implikacji. $\boxed{1° \Rightarrow 2°}$ Układ $v_1, \ldots, v_n$ powiększony o dowolny wektor $v \in V$ jest l.zależny, gdyż v jest jest komb. liniową v_i. $\boxed{2° \Rightarrow 1°}$ Maksymalność sprawia, że dla $v \in V$ układ $v_1, \ldots, v_n, v$ jest już l.zależny, a więc dzięki 215. albo v, albo któryś v_i jest liniową komb. swoich poprzedników; tę drugą możliwość wyklucza l.niezależność $v_1, \ldots, v_n$. $\boxed{1° \Rightarrow 3°}$ Gdyby $V = \langle v_2, \ldots, v_n \rangle$, wtedy v_1 byłby komb. liniową pozostałych v_i, wbrew liniowej niezależności układu. Zatem po odrzuceniu v_1 układ nie rozpina V. $\boxed{3° \Rightarrow 1°}$ Gdyby układ był liniowo zależny, wtedy pewien v_i byłby komb. liniową

[30] Formalna definicja: Jeżeli $\preceq$ jest "relacją porządkującą"(tzn. zwrotną i przechodnią) w zbiorze X, to element $x_0 \in X$ jest *najmniejszy* w X, gdy $\forall x \in X : x_0 \preceq x$, zaś x_0 jest *minimalny* w X, gdy $\forall x \in X : x \preceq x_0 \Rightarrow x_0 \preceq x$; zamieniając $\preceq$ na $\succeq$ definiuje się pojęcia elementu *największego* i *maksymalnego*. Jasne, że: (1) element najmniejszy jest minimalny, a największy — maksymalny; (2) jeśli dwa elementy są najmniejsze (lub największe) w X, to są *stowarzyszone* (koniunkcja $\preceq$ i $\succeq$); elementy minimalne lub maksymalne nie muszą być stowarzyszone. **Przykłady**. Niech $x \preceq y \Leftrightarrow x \mid y$. Dla $X = \overline{1,5}$ najmniejszy (a więc i minimalny) jest 1, elementu największego brak, a maksymalne są 3, 4 i 5. Dla $X = \{2, 3, 6\}$ elementu najmniejszego brak, minimalne są 2 i 3, a największym (i maksymalnym) jest 6.

pozostałych wektorów; odrzucenie v_i nie psułoby rozpinania, wbrew minimalności. $\boxed{1^\circ \Rightarrow 4^\circ}$ $v \in V$ ma rozkład, gdyż $V = \langle v_1, \ldots, v_n \rangle$; jeśli $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ oraz $v = \tilde{\lambda}_1 v_1 + \ldots + \tilde{\lambda}_n v_n$, to kładąc $\mu_i := \tilde{\lambda}_i - \lambda_i$ dostajemy $\mu_1 v_1 + \ldots + \mu_n v_n = 0$, co wskutek l.niezależności układu daje $\mu_i = 0$, tzn. $\tilde{\lambda}_1 - \lambda_1 = 0, \ldots, \tilde{\lambda}_n - \lambda_n = 0$. $\boxed{4^\circ \Rightarrow 1^\circ}$ Istnienie rozkładu każdego v sprawia, że $v_1, \ldots, v_n$ rozpinają V. Jeśli $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$, to lewa strona jest rozkładem wektora 0; rozkład ten musi być identyczny z rozkładem $0v_1 + \ldots + 0v_n$, skąd wynikają równości $\lambda_1 = 0, \ldots, \lambda_n = 0$.

221. **Wniosek**. Równoważność $1^\circ \iff 4^\circ$ powoduje, że baza $v_1, \ldots, v_n$ zadaje w przestrzeni V *układ współrzędnych*, tzn. bijekcję $V \longrightarrow \mathbb{K}^n$. Chodzi tu oczywiście o odwzorowanie

$$V \ni v = \lambda_1 v_1 + \ldots + \lambda_n v_n \longmapsto (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n.$$

Zauważmy, że osie tego układu współrzędnych, tzn. podzbiory V, określone warunkiem znikania wszystkich współrzędnych, oprócz jednej ustalonej, są (jednowymiarowymi) podprzestrzeniami $\langle v_1 \rangle$, $\ldots$, $\langle v_n \rangle$.

222. Jeśli $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$, to liczby $\lambda_1, \ldots, \lambda_n$ nazywa się *współrzędnymi wektora v w bazie $v_1, \ldots, v_n$*, natomiast wektory $\lambda_1 v_1, \ldots, \lambda_n v_n \in V$ — *składowymi wektora v w tej bazie* (v jest sumą swoich składowych).

223. **Przykład 1**. Wektory "zero-jedynkowe" w przestrzeni $\mathbb{K}^n$, tzn. wektory

$$\boldsymbol{e}_1 := (1, 0, \ldots, 0),\ \boldsymbol{e}_2 := (0, 1, 0, \ldots, 0),\ \ldots,\ \boldsymbol{e}_n := (0, \ldots, 0, 1),$$

lub w innej, wygodniejszej czasem formie 'kolumnowej' ('wertykalnej')

$$\boldsymbol{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \boldsymbol{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \boldsymbol{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \boldsymbol{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

mają oczywistą własność

$$\forall\ \lambda_1, \ldots \lambda_n \in \mathbb{K} : \lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_n \boldsymbol{e}_n = (\lambda_1, \ldots \lambda_n);$$

korzystając z niej można bez trudu sprawdzić bezpośrednio każdy z powyższych warunków $1^\circ, \ldots, 4^\circ$. Wobec tego układ $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ jest bazą przestrzeni $\mathbb{K}^n$; będziemy jej często używać — zwykle stosując notację 'kolumnową' — nazywając *bazą standardową przestrzeni* $\mathbb{K}^n$.

224. **Przykład 2**. Jednomiany $\mu_0, \ldots, \mu_n$, gdzie $\mu_k(t) := t^k$, tworzą bazę ('standardową') przestrzeni $\mathbb{K}_n[\,\cdot\,]$ wielomianów stopnia $\leqslant n$ jednej zmiennej. Czasami wygodniej jest posłużyć się inną bazą przestrzeni $\mathbb{K}_n[\,\cdot\,]$, np. utworzoną z 'przesuniętych' jednomianów $\mu_k(t-a) = (t-a)^k$. Przykładem bazy przestrzeni $\mathbb{K}_4[x, y]$ (wielomianów stopnia $\leqslant 4$ dwóch zmiennych x, y, o współczynnikach z $\mathbb{K}$) jest układ 15 jednomianów

$$1,\ x,\ y,\ x^2,\ xy,\ y^2,\ x^3,\ x^2y,\ xy^2,\ x^3,\ x^4,\ x^3y,\ x^2y^2,\ xy^3,\ y^4.$$

225. **Definicja**. *Wymiarem* przestrzeni wektorowej V nazywamy najmniejszą możliwą moc zbioru rozpinającego V. Jeśli przestrzeń jest *skończe-*

nie wymiarowa, tzn. ma skończony podzbiór generujący, wtedy wymiar V jest liczbą całkowitą $\geqslant 0$, oznaczaną symbolem $\boxed{\dim V}$.

226. Z tego, co już wiemy jest jasne, że skończenie wymiarowa przestrzeń V ma bazę o $\dim V$ elementach (minimalnej mocy zbiór generujący!) oraz że każda baza V ma $\geqslant \dim V$ elementów. Naszym najbliższym celem będzie m.in. pokazanie, że *każda baza V ma $\dim V$ elementów.*

227. **Ćwiczenie (obligatoryjne).** Znaleźć błąd w poniższym 'dowodzie' faktu (skądinąd prawdziwego!), że każda baza V ma $\dim V$ elementów:

"Baza jest minimalnym układem rozpinającym V, a więc ma minimalną możliwą moc, jaką może mieć zbiór rozpinający V, czyli — z definicji wymiaru — ma moc równą $\dim V$, QUOD ERAT DEMONSTRANDUM".

228. **Fakt.** Przestrzeń skończenie wymiarowa ma bazę; co więcej, każdy układ liniowo niezależny w przestrzeni V da się powiększyć do bazy V.

Niech $u_1, \ldots, u_r$ będzie układem l. niezależnym. Skoro $\dim V < \infty$, istnieje układ $v_1, \ldots, v_s$, rozpinający V. Układ $u_1, \ldots, u_r, v_1, \ldots, v_s$ rozpina V, więc zgodnie z 216 można z niego uzyskać podukład l.niezależny i rozpinający V (czyli bazę), usuwając wektory będące kombinacjami liniowymi swych poprzedników. Stojące z lewej wektory u_i są l.niezależne, więc ta 'lustracja' nie wyeliminuje żadnego z nich.

229. **Lemat Steinitza.** Jeśli przestrzeń V zawiera układ r wektorów liniowo niezależnych oraz układ s wektorów rozpinających V, to $r \leqslant s$.

Zastosujemy «pomysł Steinitza o zastępowaniu». Niech $u_1, \ldots, u_r$ będzie układem liniowo niezależnym, a $v_1, \ldots, v_s$ — układem rozpinającym V, tzn.

$$V = \langle v_1, \ldots, v_s \rangle. \tag{0}$$

Skoro $u_1 \in \langle v_1, \ldots, v_s \rangle$, to układ $u_1, v_1, \ldots, v_s$ jest l. zależny; stąd i z 215. któryś v_i jest komb. liniową swoich poprzedników (u_1 jest $\neq 0$, więc nie jest komb. liniową pustego zbioru swoich poprzedników). Zatem dzięki p.3°209

$$\exists i \in \overline{1,s} : \ V = \langle u_1, v_1, \ldots, v_s \,(\text{bez } v_i) \rangle. \tag{1}$$

Skoro $u_2 \in \langle u_1, v_1, \ldots, v_s(\text{bez } v_i) \rangle$, to układ $u_2, u_1, v_1, \ldots, v_s(\text{bez } v_i)$ jest l.zależny; stąd i z 215. któryś v_j jest komb. liniową swoich poprzedników (u_1 ani u_2 nie może być komb. liniową swoich poprzedników, gdyż u_2, u_1 są l.niezależne). Zatem

$$\exists i \neq j \in \overline{1,s} : \ V = \langle u_2, u_1, v_1, \ldots, v_s \,(\text{bez } v_i, v_j) \rangle. \tag{2}$$

Przypuśćmy teraz, że $s < r$; wtedy stosując s-krotnie taką 'procedurę zastępowania' otrzymamy w końcu

$$V = \langle u_s, \ldots, u_2, u_1 \rangle, \tag{s}$$

a w takim razie $u_{s+1} \in \langle u_s, \ldots, u_1 \rangle$, co przeczy l.niezależności. Zatem $r \leqslant s$, Q.E.D.

230. **Wniosek 1.** Każde dwie bazy V mają jednakową liczbę elementów.

231. **Wniosek 2.** Podprzestrzeń przestrzeni skończenie wymiarowej V jest przestrzenią skończenie wymiarową wymiaru $\leqslant \dim V$.

Niech $U \subset V$ oraz $\dim V < \infty$. Z 229. wynika, że każdy układ l.niez. wektorów z V (tym bardziej z U) ma nie więcej niż $\dim V$ wektorów. Dotyczy to w szczególności

maksymalnego układu l.niezależnego w U. Taki układ $u_1, \ldots, u_r$ — jak wiemy z 220. — jest bazą U; wobec tego $\dim U = r \leqslant \dim V$.

Uwaga. Implikację $U \leqslant V \Rightarrow \dim U \leqslant \dim V$ nazywa się zwykle *monotonicznością wymiaru*; jej dowód przy *dodatkowym* założeniu $\dim U < \infty$ jest oczywisty: baza U jest układem l.niezależnym w V, więc ma nie więcej niż $\dim V$ elementów, Q.E.D.

232. **Wniosek 3**. Jeśli $V_0 \subset V$ jest podprzestrzenią oraz $\dim V_0 = \dim V$, to $V_0 = V$; można to streścić hasłem *'funkcja'* dim *jest ściśle rosnąca.*

Niech $v_1, \ldots, v_r$ — baza V_0; gdyby $\exists v \in V : v \notin \langle v_1, \ldots, v_r \rangle$, wtedy układ $v_1, \ldots, v_r, v$ byłby l.niezal., więc $\dim V \geqslant r+1 = \dim V_0 + 1$, sprzeczność; zatem $\langle v_1, \ldots, v_r \rangle = V$.

233. **Wniosek 4**. Niech $\dim V = n$ oraz $Z \subset V$, wtedy:

(a) Z jest l. niezależny $\Rightarrow$ $|Z| \leqslant n$; (a') $|Z| > n \Rightarrow$ Z l. zależny;
(b) $\langle Z \rangle = V \Rightarrow |Z| \geqslant n$; (b') $|Z| < n \Rightarrow \langle Z \rangle \subsetneq V$;
(c) w przypadku, gdy $|Z| = n$, zachodzą następujące równoważności:

$$Z \text{ jest l. niezależny} \iff Z \text{ jest bazą } V \iff \langle Z \rangle = V.$$

(a) i (b) wynikają wprost z 229. (c) Załóżmy, że $|Z| = \dim V$. Gdy Z jest l.niezależny, wtedy Z zawiera się w jakiejś bazie (zob. 228) i ma tyle samo elementów, co ta baza (gdyż $|Z| = \dim V$), więc jest całą bazą. Gdy $\langle Z \rangle = V$, wtedy Z zawiera jakąś bazę (zob 216) i ma tyleż elementów, co ta baza, więc ta baza jest całym zbiorem Z.

234. **Wniosek 5**. W skończenie wymiarowej przestrzeni V układ $v_1, \ldots, v_n$ jest bazą wtedy i tylko wtedy, gdy spełnione są choć dwa spośród następujących warunków:

(a) $\langle v_1, \ldots, v_n \rangle = V$; (b) układ $v_1, \ldots, v_n$ jest lin. niezależny; (c) $n = \dim V$.

4.6 O sumie i przecięciu

235. **Fakt**. Jeśli $V_1, V_2 \subset V$ są dwiema podprzestrzeniami skończonego wymiaru, to

$$\dim V_1 + \dim V_2 = \dim(V_1 \cap V_2) + \dim(V_1 + V_2). \qquad (*)$$

Uwaga. Pokażemy nawet więcej, że V_1, V_2 mają takie bazy (zbiory) B_1, B_2 , że $B_1 \cap B_2$ jest bazą $V_1 \cap V_2$, a $B_1 \cup B_2$ — bazą $V_1 + V_2$; jest to wyjaśnieniem związku wzoru $(*)$ ze znanym z teorii mnogości wzorem

$$|B_1| + |B_2| = |B_1 \cap B_2| + |B_1 \cup B_2|. \qquad (\star)$$

wzór $(*)$ bywa (żartobliwie) nazywany *kwantową wersją wzoru* $(\star)$.

$V_1 \cap V_2$ ma — jako podprzestrzeń V_1 — skończony wymiar; niech $e_1, \ldots, e_k$ będzie bazą $V_1 \cap V_2$. Z 228. wiemy, że da się ją powiększyć do bazy V_1, a także do bazy V_2; zatem istnieją wektory $a_i, b_j \in V$ takie, że

$$e_1, \ldots, e_k, a_1, \ldots, a_r \text{ — baza } V_1, \qquad e_1, \ldots, e_k, b_1, \ldots, b_s \text{ — baza } V_2.$$

Układ $k + r + s$ wektorów $e_1, \ldots, e_k, a_1, \ldots, a_r, b_1, \ldots, b_s$ rozpina $V_1 + V_2$; jeśli więc wykażemy, że jest on l.niezależny, to da nam to wzór $\dim(V_1 + V_2) = k + r + s$, co wraz z równościami $\dim(V_1 \cap V_2) = k$, $\dim V_1 = k + r$, $\dim V_2 = k + s$ zakończy dowód wzoru $(*)$.

Zauważmy najpierw, że warunki $a \in \langle a_1, \ldots, a_r \rangle$ oraz $a \in V_1 \cap V_2$ implikują $a = 0$; istotnie, wtedy $\exists \alpha_i \in \mathbb{K} : a = \alpha_1 a_1 + \ldots + \alpha_r a_r$ oraz $\exists \gamma_i \in \mathbb{K} : a = \gamma_1 e_1 + \ldots + \gamma_k e_k$, więc $0 = a - a = \alpha_1 a_1 + \ldots + \alpha_r a_r + (-\gamma_1) e_1 + \ldots + (-\gamma_k) e_k$, co dzięki l. niezależności $a_1, \ldots, a_r, e_1, \ldots, e_k$ implikuje znikanie wszystkich współczynników, tzn. $a = 0$.

Dla dowodu l. niez. $e_1, \ldots, e_k, a_1, \ldots, a_r, b_1, \ldots, b_s$ wystarczy oczywiście pokazać, że jeśli $a+b+e = 0$, gdzie $a \in \langle a_i \rangle$, $b \in \langle b_i \rangle$ oraz $e \in \langle e_i \rangle$, to $a = b = e = 0$. Otóż $a \in V_1$ oraz $a = -b - e \in V_2$, a więc $a \in V_1 \cap V_2$; to — dzięki powyższej uwadze — daje $a = 0$. Stąd także $b + e = 0$, co — wskutek l.niez. b_i i e_j — implikuje $b = e = 0$.

236. **Uwaga 1**. Jeśli zbiór B_1 jest bazą V_1, a B_2 — bazą V_2, to — wbrew opinii często spotykanej na egzaminach i kolokwiach — zbiór $B_1 \cap B_2$ wcale *nie musi* być bazą $V_1 \cap V_2$, a zbiór $B_1 \cup B_2$ — bazą $V_1 + V_2$.

Kontrprzykład: Niech $e_1, e_2, e_3 \in V$ — l. niezależne; wtedy $B_1 := \{e_1, e_1 + e_2\}$ jest bazą $V_1 := \langle e_1, e_2 \rangle$, $B_2 := \{e_2 + e_3, e_3\}$ — bazą $V_2 := \langle e_2, e_3 \rangle$, tymczasem zaś $B_1 \cap B_2 = \emptyset$, mimo że $V_1 \cap V_2 = \langle e_2 \rangle$; podobnie $B_1 \cup B_2$ ma 4 elementy, nie może więc być bazą $V_1 + V_2 = \langle e_1, e_2, e_3 \rangle$.

237. **Uwaga 2**. Analogie pomiędzy operacjami $\cup, \cap$ na zbiorach, a ich odpowiednikami 'kwantowymi' $+, \cap$ na podprzestrzeniach, mają swoje granice. Warto sobie m.in. uświadomić, że odpowiednik tożsamości $(B_1 \cup B_2) \cap B_3 = (B_1 \cap B_3) \cup (B_2 \cap B_3)$, tj. wzór $(V_1 + V_2) \cap V_3 = V_1 \cap V_3 + V_2 \cap V_3$, nie jest na ogół prawdziwy: kontrprzykład łatwo znaleźć (i narysować!), biorąc np. 1-wymiarowe podprzestrzenie płaszczyzny.

Jak dobrze wiemy, że wzór $(\star)$ daje się (przez indukcję) uogólnić na większą liczbę zbiorów, np. $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|$. Z powodu, który właśnie przedstawiliśmy, dla 'kwantowego' odpowiednika tego wzoru indukcja 'nie idzie'; w istocie, taki odpowiednik jest wzorem nieprawdziwym!

4.7 Ogólne twierdzenie o istnieniu bazy

Oprócz pojęcia 'bazy uporządkowanej' (czyli układu wektorów) istnieje także pojęcie 'bazy nieuporządkowanej', będącej zbiorem wektorów:

238. **Definicja**. Podzbiór $Z \subset V$ nazywa się *bazą (nieuporządkowaną)* przestrzeni V, jeśli jest liniowo niezależny i rozpina całą przestrzeń V.

239. **Przykłady**. Niech V będzie przestrzenią tych ciągów $x = (x_1, x_2, \ldots) \in \mathbb{K}^{\mathbb{N}}$, które mają jedynie skończoną liczbę niezerowych wyrazów: $\exists n_0 : \forall n \geqslant n_0 : x_n = 0$. Bazą V jest np. zbiór przeliczlny, złożony z elementów $\boldsymbol{e_1}, \boldsymbol{e_2}, \ldots$, gdzie $\boldsymbol{e_r}$ jest ciągiem, którego r-ty wyraz jest jedynką, a wszystkie pozostałe — zerami.

Przestrzeń $\mathbb{K}[\cdot]$ wielomianów o współczynnikach z $\mathbb{K}$ stanowi przykład bliźniaczo podobny: zbiór jednomianów jest oczywistym przykładem bazy tej przestrzeni.

Jeśli X niepustym zbiorem, to funkcje $\delta_x : X \to \mathbb{K}$, gdzie $\delta_x(\xi) := \begin{cases} 1, \xi = x, \\ 0, \xi \in X \setminus \{x\} \end{cases}$ (o nośnikach jednoelementowych), tworzą zbiór liniowo niezależny([31]); oczywiście generują one podprzestrzeń $V \subset \mathbb{K}^X$, złożoną z funkcji o nośnikach skończonych:

[31] Istotnie, jeśli $v = \lambda_1 \delta_{x_1} + \ldots + \lambda_n \delta_{x_n}$, gdzie $\lambda_i \in \mathbb{K}$, zaś $x_i \in X$ są parami różne, to $v(x_i) = \lambda_i$; zatem $v = 0$ implikuje $\lambda_1 = \ldots = \lambda_n = 0$.

$$V = \Big\{ v \in \mathbb{K}^X : \ \#(\operatorname{supp} v) < \infty \Big\}, \text{ gdzie } \operatorname{supp} v := \{x \in X : \ v(x) \neq 0\}.$$

W takim razie zbiór $\{\delta_x : \ x \in X\}$, którego elementy są bijektywnie parametryzowane elementami zbioru X, jest bazą przestrzeni V. Z tego powodu przestrzeń V bywa czasem nazywana *przestrzenią wektorową rozpiętą na zbiorze* X i oznaczana symbolem $\mathrm{E}(X)$ lub E_X (literka 'E' od fr. słowa *espace*).

240. **Ćwiczenie.** Podzbiór $Z \subset V$ jest bazą $V \iff$ każdy wektor $v \in V$ da się w jeden jedyny sposób przedstawić w postaci kombinacji liniowej jakiegoś *skończonego* (zob. 214.) układu wektorów z Z.

241. Jak widzieliśmy (zob. 228.), *istnienie* bazy jest faktem dość błachym dla przestrzeni skończenie wymiarowej. Czy założenie $\dim V < \infty$ można by tu opuścić, tzn. czy każda przestrzeń nieskończonego wymiaru też ma bazę? Okazuje się, że tak, lecz udowodnienie tego wymaga aksjomatu wyboru, a więc dowód istnienia bazy jest niekonstruktywny, w przeciwieństwie do przypadku $\dim V < \infty$. W istocie, nawet dla 'najprostszych' przestrzeni nieskończonego wymiaru (choćby dla $\mathbb{K}^X$, gdzie zbiór X jest nieskończony) nie umiemy z reguły([32]) wskazać żadnej konkretnej bazy.

242. Zastosujemy słynny *lemat Kuratowskiego-Zorna*; jest to dalece nietrywialny wniosek z *aksjomatu wyboru*, często stosowany w 'niekonstruktywnych dowodach istnienia' jako narzędzie poręczniejsze od *aksjomatu wyboru* i niejako potęgujące jego moc.

Lemat Kuratowskiego-Zorna. Jeśli A jest zbiorem częściowo uporządkowanym([33]), w którym każdy łańcuch([34]) ma ograniczenie górne([35]), to A ma element maksymalny([36]).

243. **Twierdzenie.** Niech $Z_1 \subset Z_2$ będą dwoma podzbiorami przestrzeni V, przy czym załóżmy, że Z_1 jest liniowo niezależny. Wtedy istnieje taki liniowo niezależny zbiór Z, że $Z_1 \subset Z \subset Z_2$ oraz $\langle Z \rangle = \langle Z_2 \rangle$.

Wynika stąd w szczególności (gdy $Z_2 = V$ albo $Z_1 = \emptyset$), że:
Każdy liniowo niezależny podzbiór zawarty jest w jakiejś bazie V.
Każdy podzbiór rozpinający V zawiera jakąś bazę V.

* Szkic Dowodu. Utwórzmy zbiór A, którego elementami są liniowo niezależne zbiory Z, spełniające warunek $Z_1 \subset Z \subset Z_2$ (np. $Z_1 \in A$). Każdy 'łańcuch' w A (tzn. podzbiór A, uporządkowany 'liniowo' relacją $\subset$) ma ograniczenie górne (mianowicie sumę mnogościową swoich elementów); z lematu Kuratowskiego-Zorna wynika istnienie elementu maksymalnego w A, tzn. elementu $Z \in A$, nie zawierającego się w żadnym większym elemencie zbioru A. Dla zakończenia dowodu pozostało sprawdzenie, że $\langle Z \rangle = \langle Z_2 \rangle$, do tego zaś, dzięki p.2°209, wystarczy pokazać, że $Z_2 \subset \langle Z \rangle$. Gdyby tak nie było, istniałby wektor $u_0 \in Z_2$, nie należący do $\langle Z \rangle$; powiększenie Z o element u_0 nie zepsułoby liniowej niezależności, więc $\tilde{Z} := Z \cup \{u_0\}$ byłby elementem rodziny A większym od Z, wbrew maksymalności Z; sprzeczność.

244. **Przykład.** Zbiór $\mathbb{R}$ można traktować jako przestrzeń wektorową nad ciałem $\mathbb{Q}$, biorąc zwykłe dodawanie w $\mathbb{R}$ i definiując iloczyn 'wektora' z $\mathbb{R}$ przez 'skalar' z $\mathbb{Q}$ jako zwykły iloczyn liczb rzeczywistych. Baza tej przestrzeni jest takim zbiorem $B \subset \mathbb{R}$, że każda liczba $x \in \mathbb{R}$ ma rozkład $x = q_1 b_1 + \ldots + q_n b_n$, przy czym $n \in \mathbb{N}$,

[32] Wyjątkiem są przestrzenie postaci $E(X) := \big\{ v \in \mathbb{K}^X : \{x : v(x) \neq 0\} \text{ jest skończony} \big\}$.

[33] Tzn. wyposażonym w dwuczłonową relację $\prec$, spełniającą warunki $a \prec a$ (zwrotność), $\big(a \prec b \ \& \ b \prec a\big) \Rightarrow a = b$ (antysymetria) oraz $a \prec b \prec c \ \Rightarrow \ a \prec c$ (przechodniość).

[34] Tzn. podzbiór $L \subset A$, 'uporządkowany liniowo': $\forall\, l_1, l_2 \in L : l_1 \prec l_2$ lub $l_2 \prec l_1$.

[35] Tzn. element $a \in A$, taki że $\forall\, l \in L : l \prec a$.

[36] Tzn. $a^* \in A$ taki, że jedynym elementem $a \in A$, dla którego $a^* \prec a$, jest $a = a^*$.

$q_i \in \mathbb{Q}^*$ oraz $b_i \in B$ są liczbami określonymi *jednoznacznie* 'wektorem' x.

Zauważmy teraz, że każda funkcja $f_0 : B \to \mathbb{R}$ daje się 'przedłużyć' do funkcji $f : \mathbb{R} \to \mathbb{R}$, określonej wzorem $f(q_1 b_1 + \ldots + q_n b_n) := q_1 f_0(b_1) + \ldots + q_n f_0(b_n)$, przy czym taka funkcja f ma własności

$$\forall x, y \in \mathbb{R} : f(x+y) = f(x) + f(y), \quad \forall q \in \mathbb{Q},\, x \in \mathbb{R} : f(qx) = q f(x).$$

W szczególności wynika stąd na przykład, że istnieją funkcje $f : \mathbb{R} \to \mathbb{R}$, które

- spełniają warunek $f(x+y) = f(x) + f(y)$, lecz nie mają postaci $f(x) = ax$;
- spełniają warunek $f\left(\frac{x+y}{2}\right) \leqslant \frac{f(x)+f(y)}{2}$, lecz nie są wypukłe.

4.8 Suma prosta podprzestrzeni

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$, a $V_1, \ldots, V_r$ — podprzestrzeniami V.

245. **Ćwiczenie**. Sprawdzić, że
$$\left\{ \begin{array}{c} V_1 + V_2 = V \\ V_1 \cap V_2 = \{0\} \end{array} \right\} \iff \left\{ \begin{array}{l} \text{każdy wektor } v \in V \text{ ma jednoznaczny} \\ \text{rozkład } v = v_1 + v_2,\ v_1 \in V_1,\ v_2 \in V_2 \end{array} \right\}.$$

246. **Fakt**. Następujące warunki (1)...(4) — albo (1)...(6), przy dodatkowym założeniu, że $\dim V_i < \infty$ — są wzajemnie równoważne:

(1) $\forall v_1 \in V_1 : \ldots : \forall v_r \in V_r : v_1 + \ldots + v_r = 0 \Rightarrow (v_1 = 0, \ldots, v_r = 0)$;
(2) Każdy $v \in V_1 + \ldots + V_r$ ma jednoznaczny 'rozkład na składowe', tzn. rozkład postaci $v = v_1 + \ldots + v_r,\ v_i \in V_i$;
(3) $\forall i \in \overline{1,r} : V_i \cap V_i' = \{0\}$, gdzie $V_i' := V_1 + \ldots + V_r$ (suma bez V_i);
(4) $\forall i \in \overline{1,r} : V_i \cap V_i'' = \{0\}$, gdzie $V_1'' := \{0\}$, $V_i'' := V_1 + \ldots + V_{i-1}$.

(5) Konkatenacja baz przestrzeni $V_1, \ldots, V_r$ jest bazą $V_1 + \ldots + V_r$;
(6) $\dim(V_1 + \ldots + V_r) = \dim V_1 + \ldots + \dim V_r$.

Uwaga. *Konkatenacja* (od łac. *con-* — 'współ-' oraz *catena* — 'łańcuch') jest to operacja, polegająca na połączeniu kilku ciągów ('łańcuchów') w jeden ciąg.

Oczywiste są implikacje (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4), (1) $\Rightarrow$ (5) $\Rightarrow$ (6) oraz (5) $\Rightarrow$ (1).

$\boxed{(4) \Rightarrow (1)}$ Niech $v_1 + \ldots + v_r = 0$ oraz v_i — ostatni niezerowy składnik; wtedy $i > 1$ oraz $v_i = -v_1 - \ldots - v_{i-1} \in V_i \cap V_i'' = \{0\}$ — sprzeczność; zatem $\forall i :\ v_i = 0$.
$\boxed{(6) \Rightarrow (5)}$ Konkatenacja baz V_i jest układem rozpinającym $V^* := V_1 + \ldots + V_r$ i ma $\dim V^*$ elementów, więc jest bazą V^*.

247. **Definicja**. Mówimy, że podprzestrzenie $V_1, \ldots, V_r \subset V$ tworzą układ *liniowo niezależny* (lub że są *transwersalne* lub że *tworzą sumę prostą*), jeśli spełniają powyższe równoważne warunki.

248. **Uwaga**. Zauważmy, że — odmiennie niż dla l.niez. wektorów — wśród podprzestrzeni l.niezal. mogą być podprzestrzenie zerowe, równe $\{0\}$.

249. **Ćwiczenie**. Jeśli każda z podprzestrzeni $V_1, \ldots, V_r$ jest $\neq \{0\}$, to (1) $\Leftrightarrow$ (1*), gdzie

(1*) Każdy układ $v_1, \ldots, v_r$ taki, że $\forall i :\ v_i \in V_i$ oraz $v_i \neq 0$, jest liniowo niezależny.

250. **Uwaga**. Jasne, że dla $r = 2$ warunki (1)...(6) są równoważne warunkowi
$$V_1 \cap V_2 = \{0\}.$$

Natomiast dla $r > 2$ warunek "(7) $\forall i < j:\ V_i \cap V_j = \{0\}$" jest *słabszy* od (1)...(6), o czym świadczy przykład $V := \mathbb{K}^2$, $V_i := \langle \boldsymbol{e}_1 + i\boldsymbol{e}_2 \rangle$.

251. **Definicja**. V jest *sumą prostą* swoich podprzestrzeni $V_1, \ldots, V_r$, jeśli każdy wektor $v \in V$ ma, i to jednoznaczny, rozkład na składowe z V_i:

$$\forall v \in V : \exists!\ v_1 \in V_1 : \ldots : \exists!\ v_r \in V_r :\ v = v_1 + \ldots + v_r\ ([37]).$$

Oczywiście (patrz warunek (2)) jest to równoważne temu, że

- V jest sumą algebraiczną V_i, $i \in \overline{1,n}$, tzn. $V = V_1 + \ldots + V_r$, oraz
- $V_1, \ldots, V_r$ są liniowo niezależne.

252. **Oznaczenie**: $\boxed{V = V_1 \dot{+} \ldots \dot{+} V_r \ \text{ albo, równoważnie, } \ V = \sum_{i=1}^{r\bullet} V_i}$

oznaczać będzie, że V jest sumą prostą podprzestrzeni $V_1, \ldots, V_r \subset V$.

253. Zwróćmy uwagę, że nie nadajemy żadnego sensu samej 'prawej stronie', tzn. wyrażeniu $V_1 \dot{+} \ldots \dot{+} V_r$, bądź $\sum^{\bullet} V_i$. Trzeba również podkreślić, że 'suma prosta' nie jest jakąś operacją (jak dodawanie czy mnożenie): przestrzeń V nie jest 'rezultatem operacji', lecz czymś, co mamy od początku. Jest natomiast 'suma prosta' pewną relacją — konstatacją pewnych własności układu (pod)przestrzeni $V, V_1, \ldots, V_r$.

Zdefiniowaną w powyższy sposób 'sumę prostą' niekiedy opatruje się przymiotnikiem «wewnętrzna»; jest on związany z tym, że «wszystko odbywa się tu wewnątrz jednej przestrzeni wektorowej V».

254. **Fakt**. Niech V_1, V_2 będą podprzestrzeniami V oraz $\dim V < \infty$; wtedy $V = V_1 \dot{+} V_2 \iff$ spełnione są (choć) dwa z następujących warunków:

(a) $V_1 + V_2 = V$; (b) $V_1 \cap V_2 = \{0\}$; (c) $\dim V_1 + \dim V_2 = \dim V$.

Jasne, że $V = V_1 \dot{+} V_2 \Rightarrow$(a)(b)(c), więc zajmiemy się dowodem implikacji '$\Leftarrow$'. Skoro (a)(c) $\Rightarrow\ \dim V_1 + \dim V_2 = \dim(V_1 + V_2) \ \Rightarrow$ (6), to (a)(c) $\Rightarrow\ V = V_1 \dot{+} V_2$. Ponieważ z (b) wynika (4), to (a)(b)$\Rightarrow\ V = V_1 \dot{+} V_2$, z kolei dzięki 235 oraz 232 mamy
(b)(c)$\Rightarrow \dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2) = \dim V \Rightarrow V_1 + V_2 = V$, więc mamy także implikację (b)(c) $\Rightarrow\ V = V_1 \dot{+} V_2$.

255. **Ćwiczenie**. Niech $V_1, \ldots, V_r$ będą podprzestrzeniami V, przy czym $\dim V < \infty$; wtedy $V = V_1 \dot{+} \ldots \dot{+} V_r \iff$ spełnione są (choć) dwa z następujących warunków:

(a) $V_1 + \ldots + V_r = V$; (b′) $\forall i:\ V_i \cap V_i' = \{0\}$; (c) $\dim V_1 + \ldots + \dim V_r = \dim V$,

gdzie $V_i' := \sum_{j:\, i \neq j} V_j$.

Warunek (b′) można też zastąpić warunkiem (b″), zamieniając V_i' na $V_i'' := \sum_{j:\, j<i} V_j$.

256. **Przykład**. Jeśli $e_1, \ldots, e_n$ jest bazą V, to $V = \langle e_1 \rangle \dot{+} \langle e_2 \rangle \dot{+} \ldots \dot{+} \langle e_n \rangle$; mamy wtedy także np. dla $1 < i < j < n$ rozkład

$$V = \langle e_1, \ldots, e_i \rangle \dot{+} \langle e_{i+1}, \ldots, e_j \rangle \dot{+} \langle e_{j+1}, \ldots, e_n \rangle.$$

[37] Przypominamy, że użyty tutaj symbol '$\exists!$' oznacza '*istnieje dokładnie jeden*'.

257. **Przykład**. Ustalmy $n \in \mathbb{N}$ oraz niezerowy wielomian $b \in V := \mathbb{K}_n[\,\cdot\,]$; niech $m := \deg b$. Określmy podprzestrzenie $V_0, V_1 \subset V$ następująco:

$$V_0 := b\,\mathbb{K}_{n-m}[\,\cdot\,] = (\text{elementy } V \text{ podzielne przez } b), \quad V_1 := \mathbb{K}_{m-1}[\,\cdot\,].$$

Sprawdzimy, że $V = V_0 \dot{+} V_1$; uzyskamy w ten sposób nowy, krótszy dowód twierdzenia o dzieleniu wielomianów: każdy $a \in V$ (a więc każdy $a \in \mathbb{K}[\,\cdot\,]$, bo n może tu być dowolnie duże) ma jednoznaczny rozkład $a = bq + r$, gdzie q i r są wielomianami oraz $\deg r < \deg b$.

Warunek (b) jest oczywisty; dla sprawdzenia (c) zauważmy, że jeśli $e_i \in \mathbb{K}[\,\cdot\,]$ są jednomianami, $e_i(t) := t^i$, to $(e_0, \ldots, e_n)$ jest bazą V, $(e_0, \ldots, e_{m-1})$ — bazą V_1, zaś $(be_0, \ldots, be_{n-m})$ — bazą V_0, więc $\dim V_0 + \dim V_1 = (n-m+1)+m = n+1 = \dim V$.

258. **Definicja** ('zewnętrzna' suma prosta). Niech $V_1, \ldots, V_r$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$; teraz już nie muszą być one podprzestrzeniami jakiejś jednej przestrzeni V. W iloczynie kartezjańskim $V_1 \times \ldots \times V_r = \{(v_1, \ldots, v_r) : \forall i \in \overline{1,r} : v_i \in V_i\}$ zbiorów V_i określmy dodawanie elementów i ich mnożenie przez liczby wzorami:

$$(v_1, \ldots, v_r) + (w_1, \ldots, w_r) := (v_1 + w_1, \ldots, v_r + w_r),$$
$$\lambda(v_1, \ldots, v_r) := (\lambda v_1, \ldots, \lambda v_r);$$

dostajemy w ten sposób przestrzeń wektorową, nazywaną *sumą prostą* (*zewnętrzną*) przestrzeni wektorowych $V_1, \ldots, V_r$.

Zamiast $V_1 \times \ldots \times V_r$ używa się często oznaczenia $V_1 \oplus \ldots \oplus V_r$, zaś element $(v_1, \ldots, v_r)$ oznacza się wtedy symbolem $v_1 \oplus \ldots \oplus v_r$; wobec tego mamy np. $(v_1 \oplus \ldots \oplus v_r) + (w_1 \oplus \ldots \oplus w_r) = (v_1 + w_1) \oplus \ldots \oplus (v_1 + w_1)$.

Warto zwrócić uwagę na to, że np. $v_1 \oplus v_2$ to zazwyczaj co innego, niż $v_2 \oplus v_1$: w przypadku, gdy $V_1 \neq V_2$, są to zresztą elementy dwóch różnych przestrzeni!

Czy 'wewnętrzna' i 'zewnętrzna' suma prosta mają za sobą coś wspólnego? Odpowiedź jest krótka i prosta, a jej uzasadnienie stanie się łatwym ćwiczeniem do następnego rozdziału:

259. **Fakt**. Niech $V_1, \ldots, V_r$ będą podprzestrzeniami przestrzeni V. Wtedy

$$V = V_1 \dot{+} \ldots \dot{+} V_r \Leftrightarrow \begin{pmatrix} \text{odwzorowanie liniowe } F : V_1 \oplus \ldots \oplus V_r \to V, \\ \text{dane wzorem } F(v_1 \oplus \ldots \oplus v_r) := v_1 + \ldots + v_r\,, \\ \text{jest izomorfizmem przestrzeni wektorowych.} \end{pmatrix}$$

260. **Uwaga**. Bliskie pokrewieństwo obu sum prostych (i pojęć, i nazw) jest powodem tego, że najczęściej zamiast '$\dot{+}$' używa się symbolu '$\oplus$'. Z kontekstu zwykle łatwo wynika, o którą sumę prostą ('wewnętrzną' czy wewnętrzną') chodzi, zaś kropka nad '+' łatwo może być przeoczona lub (przy gorszej jakości druku) niewidoczna.

5 Odwzorowania liniowe

5.1 Odwzorowania liniowe; przykłady

Niech V i W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$.

261. **Definicja**. Mówimy, że odwzorowanie $F : V \to W$ jest *liniowe*, jeżeli

- $\forall v_1, v_2 \in V : \ F(v_1 + v_2) = F(v_1) + F(v_2)$ (addytywność);
- $\forall \lambda \in \mathbb{K} : \forall v \in V : \ F(\lambda v) = \lambda F(v)$ (jednorodność).

262. **Fakt** (*oczywiste własności odwzorowania liniowego* $F : V \to W$):

(1) $F(0) = 0$;
(2) $F(\lambda_1 v_1 + \ldots + \lambda_r v_r) = \lambda_1 F(v_1) + \ldots + \lambda_r F(v_r)$,
czyli F-obrazem kombinacji liniowej jest kombinacja liniowa F-obrazów z tymi samymi współczynnikami;
(3) $F(\langle v_1, \ldots, v_r \rangle) := (F\text{-obraz zbioru } \langle v_1, \ldots \rangle) = \langle F(v_1), \ldots, F(v_r) \rangle$;
(4) F-obrazem i F-przeciwobrazem podprzestrzeni są podprzestrzenie:

$$V_0 \text{ jest podprzestrzenią } V \ \Rightarrow \ \begin{pmatrix} F(V_0) := \{F(v) : v \in V_0\} \\ \text{jest podprzestrzenią } W \end{pmatrix};$$

$$W_0 \text{ jest podprzestrzenią } W \Rightarrow \begin{pmatrix} F^{-1}(W_0) := \{v \in V : F(v) \in W_0\} \\ \text{jest podprzestrzenią } V \end{pmatrix}.$$

263. **Definicja** (*jądro, obraz i rząd odwzorowania liniowego* $F : V \to W$):

$$\begin{aligned} &\ker F := F^{-1}\{0\} = \{v \in V : F(v) = 0\} \subset V && \text{(jądro } F\text{)}; \\ &\operatorname{im} F := F(V) = \{F(v) : v \in V\} \subset W && \text{(obraz } F\text{)}; \\ &\operatorname{rk} F := \dim(\operatorname{im} F) \in \mathbb{Z}_+ && \text{(rząd } F\text{)}. \end{aligned}$$

Z powyższej własności (4) wynika, że $\ker F$ jest podprzestrzenią V, natomiast $\operatorname{im} F$ — podprzestrzenią przestrzeni W.

264. **Fakt**. $F(v_1) = F(v_2) \iff v_2 - v_1 \in \ker F$, a zatem

$$(F \text{ jest injektywne}) \iff \ker F = \{0\}).$$

265. **Ćwiczenie**. Udowodnić fakt 259, badając, kiedy $\ker F = \{0\}$, oraz kiedy $\operatorname{im} F = V$.

266. **Fakt** (*postać kanoniczna* odwzorowania liniowego). Jeśli przestrzenie V i W są skończenie wymiarowe, a $F : V \to W$ jest odwzorowaniem liniowym, to istnieją bazy: $e_1, \ldots, e_n$ dla V i $f_1 \ldots, f_m$ dla W, oraz liczba $r \in \overline{0, \min(m,n)}$, takie że

$$F(e_i) = \begin{cases} f_i\,, & \text{dla } 1 \leqslant i \leqslant r, \\ 0, & \text{dla } r < i \leqslant n. \end{cases} \qquad (\dagger)$$

267. **Ćwiczenie**. Sprawdzić, że jeśli zachodzi (†), to układ $e_{r+1}, \ldots, e_n$ jest bazą $\ker F$, układ $f_1, \ldots, f_r$ — bazą $\operatorname{im} F$, zaś $r = \operatorname{rk} F$.

268. **Wniosek**. $\boxed{\dim \ker F + \dim \operatorname{im} F = \dim V}$ ('bilans wymiarów')

gdyż $\dim \ker F = n - r$, $\dim \operatorname{im} F = r$.

Konstrukcja baz, w których F ma postać kanoniczną. Niech $f_1, \ldots, f_r$ — dowolna baza $\operatorname{im} F$; skoro $f_i \in \operatorname{im} F$, to $\exists\, e_1, \ldots, e_r \in V$ takie, że $F(e_1) = f_1, \ldots, F(e_r) = f_r$. Wykażemy, że *jeśli $\tilde{e}_1, \ldots, \tilde{e}_s$ jest bazą $\ker F$, to układ $e_1, \ldots, e_r, \tilde{e}_1, \ldots, \tilde{e}_s$ jest bazą V*, dowodząc istnienia i jednoznaczności rozkładu $\boxed{v = \sum_{i=1}^{r} \lambda_i e_i + \sum_{j=1}^{s} \mu_j \tilde{e}_j}$ (1) dla każdego $v \in V$, zob.220. Ponieważ $F(e_i) = f_i$, $F(\tilde{e}_j) = 0$ i F jest liniowe, więc konsekwencją (1) byłoby $\boxed{F(v) = \sum_{i=1}^{s} \lambda_i f_i}$ (2); otóż f_i są bazą $\operatorname{im} F$, a $F(v) \in \operatorname{im} F$, więc $\lambda_i \in \mathbb{K}$, spełniające (2), istnieją i są określone jednoznacznie. Zarazem z (2) wynika, że $\tilde{v} := v - \sum_{i=1}^{r} \lambda_i e_i \in \ker F$, zaś $\tilde{e}_j$ są bazą $\ker F$, więc istnieją jednoznacznie określone $\mu_j \in \mathbb{K}$ takie, że $\tilde{v} = \sum_{j=1}^{s} \mu_j \tilde{e}_j$, tzn. że zachodzi równość (1).

Mając dla V bazę $e_1, \ldots, e_r, e_{r+1} := \tilde{e}_1, \ldots, e_{r+s} := \tilde{e}_s$ zakończymy konstrukcję, dopełniając układ $f_1, \ldots, f_r$ do bazy $f_1, \ldots, f_r, f_{r+1}, \ldots, f_m$ całej przestrzeni W.

269. **Ćwiczenie.** Uzasadnić alternatywną konstrukcję: wybieramy najpierw bazę $\tilde{e}_1, \ldots, \tilde{e}_s$ podprzestrzeni $\ker F$; następnie dopełniamy ją do bazy $e_1, \ldots, e_r, \tilde{e}_1 =: e_{r+1}, \ldots, \tilde{e}_s =: e_n$ całej V, znajdując stosowne $e_1, \ldots, e_r$; wtedy układ $f_1, \ldots, f_r$, gdzie $f_i := F(e_i)$, jest bazą $\operatorname{im} F$ (udowodnić!), więc można go dopełnić do bazy $f_1, \ldots, f_r, \ldots, f_m$ całej przestrzeni W.

270. **Przykłady odwzorowań liniowych.**

Niech dla uproszczenia $\mathbb{K} = \mathbb{R}$, choć wszystkie poniższe przykłady można uogólnić na przypadek dowolnego ciała $\mathbb{K}$.

$F : \mathbb{K}[\,\cdot\,] \to \mathbb{K}$, $F(v) := v(t_0)$, gdzie $t_0 \in \mathbb{K}$ jest ustaloną liczbą;

Uwaga. Liniowość F oznacza, że $v(t_0)$ zależy liniowo od v (a niekoniecznie od t_0).

$G : \mathbb{K}[\,\cdot\,] \to \mathbb{K}$, $F(v) := v(t_1) - 3v'(t_2) + 4v''(t_3)$;
$D : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $D(v) := v'$, tzn. $(D(v))(t) := v'(t)$;
$I : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $I(v)(t) := \int_0^t v(s)ds$;

Warto tu zauważyć, że $D \circ I = \mathrm{id}_V$, lecz $I \circ D \neq \mathrm{id}_V$, gdyż $D \circ I(v) = v - v(0)$;

$J : \mathbb{K}[\,\cdot\,] \to \mathbb{K}$, $J(v) := \int_a^b g(s)v(s)dt$; ($g$ — ustalona funkcja ciągła);
$\Delta_+ : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $(\Delta_+(v))(t) := v(t+1) - v(t)$;
$\Delta_- : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $(\Delta_-(v))(t) := v(t) - v(t-1)$;
$J : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $(J(v))(t) := \int_a^t g(s)v(s)ds$;
$M_w : \mathbb{K}[\,\cdot\,] \to \mathbb{K}[\,\cdot\,]$, $M_w(v) := vw$ ($w \in \mathbb{K}[\,\cdot\,]$ — ustalony wielomian);

Analogiczna operacja $A_w(v) := v + w$ (dodawania ustalonego w) nie jest liniowa!

$L : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$, $L(x_1, x_2, x_3, \ldots) := (x_2, x_3, \ldots)$ ('lewy przesuw');
$R : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$, $R(x_1, x_2, x_3, \ldots) := (0, x_1, x_2, \ldots)$ ('prawy przesuw');

Znów $L \circ R = \mathrm{id}_V$, lecz $R \circ L \neq \mathrm{id}_V$, gdyż $R \circ L(x_1, x_2, x_3, \ldots) = (0, x_2, x_3, \ldots)$.

271. **Ćwiczenie.** Jeśli $e_1, \ldots, e_n$ jest dowolnym układem wektorów z V, to $L : \mathbb{K}^n \to V$, dane wzorem $L(\boldsymbol{x}) = L\left(\begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}\right) := e_1 x^1 + \ldots + e_n x^n$,

jest liniowe; $\operatorname{im} L = \langle e_1, \ldots, e_n \rangle$, $\ker L = \{\boldsymbol{x} \in \mathbb{K}^n : \sum_{j=1}^{n} e_j x^j = 0\}$, a zatem ($L$ jest injektywne) $\iff$ ($e_1, \ldots, e_n$ są liniowo niezależne).

272. **Oznaczenie.** $\boxed{\mathrm{L}(V;W) := \{F : F \in W^V,\ F \text{ — liniowe}\} \subset W^V}$.

Proponujemy jako ćwiczenie (nie wymagające ani mozołu, ani inwencji) samodzielne wykazanie dwóch poniższych faktów:

273. **Fakt.** Określmy w naturalny sposób *sumę* odwzorowań $F_1, F_2 \in W^V$:

$$(F_1 + F_2)(v) := F_1(v) + F_2(v),$$

oraz iloczyn odwzorowania $F \in W^V$ przez liczbę: $(\lambda F)(v) := \lambda F(v)$. Wtedy

(a) zbiór W^V jest przestrzenią wektorową nad ciałem K;
(b) zbiór $\mathrm{L}(V;W)$ jest podprzestrzenią wektorową przestrzeni W^V.

Zauważmy, że fakt (b) można wypowiedzieć inaczej:

Kombinacja liniowa (w szczególności suma) odwzorowań liniowych jest odwzorowaniem liniowym.

274. **Fakt.** Złożenie odwzorowań liniowych jest odwzorowaniem liniowym:

$$\text{Jeśli } G \in \mathrm{L}(U;V),\ F \in \mathrm{L}(V;W), \text{ to } F \circ G \in \mathrm{L}(U;W).$$

Ponadto operacja składania $\circ : \mathrm{L}(V;W) \times \mathrm{L}(U;V) \to \mathrm{L}(U;W)$ jest 'dwuliniowa', tzn. $F \circ G$ zależy liniowo od F (przy ustalonym G) oraz od G (przy ustalonym F).

275. **Uwaga**. Bardzo często mając do czynienia z odwzorowaniami liniowymi([38]) dla uproszczenie notacji pomija się:
(a) symbol składania '$\circ$', pisząc np. FG i AFA^{-1} zamiast $F \circ G$ i $A \circ F \circ A^{-1}$;
(b) nawiasy zamykające argument; można więc pisać Fv zamiast $F(v)$.

276. **Fakt.** (1) Jeśli $F_1, F_2 \in \mathrm{L}(V;W)$, to $\boxed{\operatorname{rk}(F_1 + F_2) \leqslant \operatorname{rk} F_1 + \operatorname{rk} F_2}$.

(2) Jeśli $F \in \mathrm{L}(V,W)$, to $\boxed{\operatorname{rk} F \leqslant \min\{\dim V, \dim W\}}$.

(3) Jeśli $F \in \mathrm{L}(V,W)$ oraz $G \in \mathrm{L}(U,V)$, to

$$\boxed{\operatorname{rk}(F \circ G) \leqslant \min\{\operatorname{rk} F, \operatorname{rk} G\}}.$$

(1) Jeśli $F = F_1 + F_2$, to $\forall v : F(v) = F_1(v) + F_2(v) \in \operatorname{im} F_1 + \operatorname{im} F_2$, więc $\boxed{\operatorname{im} F \subset \operatorname{im} F_1 + \operatorname{im} F_2}$. Stąd $\operatorname{rk}(F_1 + F_2) = \dim \operatorname{im} F \leqslant \dim(\operatorname{im} F_1 + \operatorname{im} F_2) = \dim \operatorname{im} F_1 + \dim \operatorname{im} F_2 - \dim(\operatorname{im} F_1 \cap \operatorname{im} F_2) \leqslant \operatorname{rk} F_1 + \operatorname{rk} F_2$.

(2) $\operatorname{im} F \subset W$ daje $\operatorname{rk} F \leqslant \dim W$, a zarazem $\operatorname{rk} F = \dim V - \dim \ker F \leqslant \dim V$.

(3) Stosując (2) do $\tilde{F} := F|_{\operatorname{im} G} \in \mathrm{L}(\operatorname{im} G, W)$ dostajemy $\operatorname{rk} \tilde{F} \leqslant \dim \operatorname{im} G = \operatorname{rk} G$, zaś $\operatorname{im} \tilde{F} = F(\operatorname{im} G) = F\big(G(U)\big) = \operatorname{im}(F \circ G)$, a więc $\operatorname{rk} \tilde{F} = \operatorname{rk}(F \circ G)$. Zarazem $G(U) \subset V$ daje $\operatorname{im}(F \circ G) = F\big(G(U)\big) \subset F(V) = \operatorname{im} F$, a więc $\operatorname{rk}(F \circ G) \leqslant \operatorname{rk} F$.

[38] Słowa 'odwzorowanie' i 'operator' w zestawieniu z przymiotnikami 'liniowe–liniowy' są zwykle używane wymiennie i traktowane jako synonimy.

277. **Fakt**. Jeśli $F \in \mathrm{L}(V;W)$ jest bijekcją, tzn. ma odwrotność, to odwzorowanie $F^{-1} : W \to V$ też jest liniowe, tzn. $F^{-1} \in \mathrm{L}(W;V)$.

Addytywność: Gdy dla zadanych $w_1, w_2 \in W$ weźmiemy $v_i := F^{-1}(w_i)$, wtedy $w_1 + w_2 = F(v_1) + F(v_2) = F(v_1 + v_2)$, więc

$$F^{-1}(w_1 + w_2) = F^{-1}\big(F(v_1 + v_2)\big) = v_1 + v_2 = F^{-1}(w_1) + F^{-1}(w_2).$$

Jednorodność. Jeśli dla $w \in W$ wziąć $v := F^{-1}(w)$, to $\lambda w = \lambda F(v) = F(\lambda v)$, więc

$$F^{-1}(\lambda w) = F^{-1}\big(F(\lambda v)\big) = \lambda v = \lambda F^{-1}(w).$$

278. **Terminologia**. Odwzorowanie, będące liniową bijekcją, nazywa się *izomorfizmem przestrzeni wektorowych*. Zapis $F : V \xrightarrow{\cong} W$ oznacza, że F jest izomorfizmem. Dwie przestrzenie V i W, dla których istnieje izomorfizm $F : V \to W$, nazywamy *izomorficznymi*, pisząc $V \cong W$.

279. **Ćwiczenie**. Jeśli $F \in \mathrm{L}(V;W)$, to $F : V \xrightarrow{\cong} W \iff \begin{pmatrix} \ker F = \{0\} \\ \mathrm{im}\, F = W \end{pmatrix}$.

280. **Fakt**. Niech $F \in \mathrm{L}(V;W)$, przy czym V, W są skończonego wymiaru. Wtedy F jest izomorfizmem $\iff$ spełnione są (choć) dwa z warunków

$$(1)\ \ker F = \{0\}; \qquad (2)\ \mathrm{im}\, F = W; \qquad (3)\ \dim V = \dim W.$$

Wiemy już, że (1)&(2) $\Rightarrow$ (F jest izomorfizmem), i że dla izomorfizmu warunki (1)..(3) są spełnione; do zakończenia dowodu wystarczy sprawdzić, że (1)&(3) $\Rightarrow$ (2) oraz (2)&(3) $\Rightarrow$ (1), a to jest oczywistą konsekwencją bilansu wymiarów (p.268).

281. **Ćwiczenie**. Znaleźć przykłady na to, że dla $F \in \mathrm{L}(V;V)$, $\dim V = \infty$, ani injektywność ($\ker F = \{0\}$), ani surjektywność ($\mathrm{im}\, F = V$) nie implikują bijektywności.

282. **Uwaga**. Jasne, że odwrotność izomorfizmu, a także złożenie izomorfizmów, jest izomorfizmem. Wynika stąd, że izomorficzność przestrzeni wektorowych jest 'relacją równoważności', tzn. jest zwrotna (zawsze $V \cong V$), symetryczna (jeśli $V \cong W$, to $W \cong V$) oraz przechodnia (jeżeli $U \cong V \cong W$, to także $U \cong W$).

Zwrotność: $\mathrm{id}_V : V \xrightarrow{\cong} V$. Symetria: jeśli $F : V \xrightarrow{\cong} W$, to $F^{-1} : W \xrightarrow{\cong} V$. Przechodniość: jeśli $G : U \xrightarrow{\cong} V$ & $F : V \xrightarrow{\cong} W$, to $F \circ G : U \xrightarrow{\cong} W$.

283. **Ćwiczenie**. Jeśli $e = (e_1, \ldots, e_n)$ jest układem wektorów z V, to

$$\begin{pmatrix} \text{odwzorowanie } L_e : \mathbb{K}^n \to V, \\ L_e(\boldsymbol{x}) := x^1 e_1 + \ldots x^n e_n, \text{ jest izomorfizmem} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{układ } e \\ \text{jest bazą } V \end{pmatrix}.$$

284. **Fakt**. $\begin{pmatrix} \text{Dwie przestrzenie} \\ \text{skończenie wymiarowe} \\ \text{są izomorficzne} \end{pmatrix} \iff \begin{pmatrix} \text{mają jednakowe} \\ \text{wymiary} \end{pmatrix}$.

$\boxed{\Leftarrow}$ Gdy $n = \dim V = \dim V$, e — baza V, a f — baza W, to mamy $L_e : \mathbb{K}^n \xrightarrow{\cong} V$ oraz $L_f : \mathbb{K}^n \xrightarrow{\cong} W$, a zatem $V \cong W$ dzięki symetrii i przechodniości 'relacji' $\cong$.
$\boxed{\Rightarrow}$ Odwrotnie, jeśli $F : V \xrightarrow{\cong} W$, to $\ker F = \{0\}$, $\mathrm{im}\, F = W$, więc z 'bilansu wymiarów' $\dim\ker F + \dim\mathrm{im}\, F = \dim V$ dostajemy równość $\dim W = \dim V$.

285. **Fakt**. Niech $F \in \mathrm{L}(V;W)$ oraz $G \in \mathrm{L}(W;V)$, gdzie przestrzenie V, W są skończenie wymiarowe. Załóżmy, że $\boxed{F \circ G = \mathrm{id}_W}$, oraz że spełniony jest choć jeden z następujących warunków:

$$(1)\ F \text{ jest bijekcją}; \qquad (2)\ G \text{ jest bijekcją}; \qquad (3)\ \ker F = \{0\};$$
$$(4)\ \mathrm{im}\, G = V; \qquad (5)\ \dim V = \dim W.$$

Wtedy operatory F i G są wzajemnie odwrotnymi izomorfizmami (a więc pozostałe warunki też są spełnione).

Z założenia $FG = \mathrm{id}_W$ wynika, że $\left\{ \begin{matrix} \mathrm{im}\, F = W \\ \ker G = \{0\} \end{matrix} \right\}$ (łatwe spostrzeżenie!), a teza sprowadza się do równości $GF = \mathrm{id}_V$.

(1) Jeśli F odwracalny, to $GF = F^{-1}(FG)F = F^{-1}\,\mathrm{id}_W\, F = F^{-1}F = \mathrm{id}_V$, QED.
(2) Jeśli G jest odwracalny, to $GF = G(FG)G^{-1} = G\,\mathrm{id}_W\, G^{-1} = \mathrm{id}_V$, QED.
(3) Założenie $\ker F = 0$ wraz z $\mathrm{im}\, F = W$ implikuje (1), a to już daje tezę, QED.
(4) Założenie $\mathrm{im}\, G = V$ wraz z $\ker G = \{0\}$ implikuje (2), skąd mamy tezę, QED.
(5) Bilans wymiarów daje $\dim \ker F = \dim V - \dim \mathrm{im}\, F = \dim V - \dim W = 0$, a zatem spełniony jest warunek (3), z którego już wywiedliśmy tezę, QED.

286. **Uwaga**. Założenie o skończoności wymiarów przestrzeni V i W jest w tym fakcie istotne, o czym świadczy przykład pary operatorów D i I (lub L i R) z punktu 270.

5.2 Macierz wektora i macierz operatora liniowego

Zbiór macierzy wymiaru $m \times n$ o wyrazach z $\mathbb{K}$ będziemy oznaczać symbolem ${\mathbb{K}^m}_n$; począwszy od tego miejsca symbole $\mathbb{K}^n$ i $\mathbb{K}_n$ będą uproszczonymi formami dla ${\mathbb{K}^n}_1$ i ${\mathbb{K}^1}_n$, oznaczając, odpowiednio, przestrzeń wektorów kolumnowych i wierszowych.

Ustalmy, jak poprzednio, przestrzenie skończonego wymiaru V i W nad ciałem $\mathbb{K}$.

287. **Definicja** (*macierz wektora w bazie*). Jeśli $e = (e_1, \ldots, e_n)$ jest bazą V, to dla $v \in V$ symbolem $[v]^e$ oznaczamy wektor kolumnowy utworzony ze współczynników rozkładu wektora v w bazie e, tzn.

$$[v]^e = \boldsymbol{x} \iff v = x^1 e_1 + \ldots + x^n e_n.$$

Tak więc odwzorowanie $[\,\cdot\,]^e : V \to \mathbb{K}^n$ jest izomorfizmem, mianowicie odwrotnością izomorfizmu $L_e : \mathbb{K}^n \xrightarrow{\cong} V$, $F(\boldsymbol{x}) = x^1 e_1 + \ldots x^n e_n$.

Wektor kolumnowy $[v]^e \in \mathbb{K}^n$ nazywamy *macierzą wektora v w bazie e*.

288. **Fakt**. Jeśli $e_1, \ldots, e_n$ jest bazą V, to dla dowolnego układu $w_1, \ldots, w_n$ wektorów przestrzeni W istnieje dokładnie jedno odwzorowanie liniowe $F \in \mathrm{L}(V;W)$, takie że

$$F(e_1) = w_1, \ \ldots, \ F(e_n) = w_n. \qquad (*)$$

Można to streścić konstatacją *odwzorowanie liniowe jest jednoznacznie określone przez swoje wartości na wektorach bazy*. Inny wariant: *jeśli podzbiór $B \subset V$ jest bazą V, to każde odwzorowanie $f : B \to W$ da się, i to jednoznacznie, rozszerzyć do odwzorowania liniowego $F : V \to W$.*

Jeśli $F \in \mathrm{L}(V;W)$ spełnia warunki (*), to wskutek liniowości mamy

$$F(e_1x^1+\ldots+e_nx^n)=w_1x^1+\ldots+w_nx^n; \qquad (**)$$

wynika stąd, że wartości $w_1,\ldots,w_n$ w pełni określają F, ponieważ każdy $v\in V$ **ma** rozkład postaci $v=e_1x^1+\ldots+e_nx^n$. Z kolei **jednoznaczność** takiego rozkładu v pozwala, dla danych dowolnie $w_i\in W$, *zdefiniować* $F:V\to W$ wzorem (**); widoczna jest wtedy liniowość F oraz to, że spełnione są warunki (*), gdyż np. $e_1=e_1x^1+\ldots+e_nx^n$, gdzie $x^1=1$, $x^2=\ldots=x^n=0$.

289. **Wniosek**. Baza $e_1,\ldots,e_n$ przestrzeni V określa bijekcję

$$\mathrm{L}(V;W)\longrightarrow W\times\ldots\times W, \qquad F\mapsto(F(e_1),\ldots,F(e_n)).$$

Jeśli mamy także bazę W, to wektory W możemy opisywać macierzami z $\mathbb{K}^m$, zaś n-elementowe układy wektorów W — macierzami z $\mathbb{K}^m{}_n$.

290. **Definicja**. Jeśli $e=(e_1,\ldots,e_n)$ jest bazą w V, zaś $f=(f_1,\ldots,f_m)$ — bazą w W, to *macierzą odwzorowania* $F\in\mathrm{L}(V;W)$ *względem obu tych baz* nazywamy macierz

$$[F]^f_{\ e}=[F^i{}_j]_{i\in\overline{1,m},\ j\in\overline{1,n}}=\begin{bmatrix}F^1{}_1 & \ldots & F^1{}_n\\ \vdots & \ldots & \vdots\\ F^m{}_1 & \ldots & F^m{}_n\end{bmatrix}\in\mathbb{K}^m{}_n\ ,$$

której kolejnymi kolumnami są wektory $[F(e_1)]^f,\ldots,[F(e_n)]^f\in\mathbb{K}^m$, tzn. macierze wektorów $F(e_j)\in W$ w bazie f. Oznacza to, że liczby $F^1{}_j,\ \ldots,\ F^m{}_j$ są współrzędnymi wektora $F(e_j)$ w bazie f, czyli że wyrazy $F^i{}_j$ macierzy $[F]^f_{\ e}$ są określone warunkami $\boxed{F(e_j)=\sum_{i=1}^m f_i\,F^i{}_j}$.

291. **Przypomnienie**. *Iloczyn* $\boldsymbol{Fx}$ wektora $\boldsymbol{x}\in\mathbb{K}^n$ przez macierz $\boldsymbol{F}\in\mathbb{K}^m{}_n$ definiujemy jako wektor $\boldsymbol{y}\in\mathbb{K}^m$ o współrzędnych danych wzorem $y^i:=\sum_{j=1}^n F^i{}_j\,x^j$, $i\in\overline{1,m}$, tzn.

$$\begin{bmatrix}F^1{}_1 & \ldots & F^1{}_n\\ \vdots & \ldots & \vdots\\ F^m{}_1 & \ldots & F^m{}_n\end{bmatrix}\begin{bmatrix}x^1\\ \vdots\\ x^n\end{bmatrix}:=\begin{bmatrix}F^1{}_1x^1+\ldots+F^1{}_nx^n\\ \vdots\\ F^m{}_1x^1+\ldots+F^m{}_nx^n\end{bmatrix}.$$

Zobaczymy teraz, że macierz operatora zawiera pełną informację o operatorze:

292. **Fakt**. Dla $v\in V$ zachodzi równość $\boxed{[F(v)]^f=[F]^f_{\ e}\,[v]^e}$. Oznacza to, że dla $\boldsymbol{x}=\begin{bmatrix}x^1\\ \vdots\\ x^n\end{bmatrix}\in\mathbb{K}^n$ oraz $\boldsymbol{y}=\begin{bmatrix}y^1\\ \vdots\\ y^m\end{bmatrix}\in\mathbb{K}^m$ mamy równoważność

$$F\Big(\sum_{j=1}^n e_j\,x^j\Big)=\sum_{i=1}^m f_i\,y^i \quad\Longleftrightarrow\quad \forall i:\ y^i=\sum_{j=1}^n F^i{}_j\,x^j,\ \text{tzn. } \boldsymbol{y}=[F]^f_{\ e}\,\boldsymbol{x}.$$

Istotnie, liniowość F sprawia, że $F\big(\sum_{j=1}^n e_j\,x^j\big)=\sum_{j=1}^n F(e_j)\,x^j=\sum_{j=1}^n\Big(\sum_{i=1}^m f_iF^i{}_j\Big)x^j$, co jest równe $\sum_{i=1}^m f_i\Big(\sum_{j=1}^n F^i{}_j\,x^j\Big)$ dzięki łączności '+' i rozdzielności '$\cdot$' względem '+'.

293. **Uwaga**. $[F]^f_{\ e}$ jest *jedyną* macierzą $\boldsymbol{A}\in\mathbb{K}^m{}_n$, spełniającą warunek $\forall v:\ [F(v)]^f_{\ e}=\boldsymbol{A}\,[v]^e$, a więc powyższy wzór $[F(v)]^f=[F]^f_{\ e}\,[v]^e$

można traktować jako ALTERNATYWNĄ DEFINICJĘ macierzy $[F]^f_e$.

Istotnie, każdy wektor $\boldsymbol{x} \in \mathbb{K}^n$ można przedstawić jako $[v]^e$, $v \in V$; zarazem macierz zerowa jest oczywiście jedyną macierzą z $\mathbb{K}^m{}_n$, zerującą wszystkie wektory $\boldsymbol{x} \in \mathbb{K}^n$.

294. **Przykład**. Wybierzmy dla $V := \mathbb{K}^n$, $W := \mathbb{K}^m$ *bazy standardowe*, tzn.

$$\boldsymbol{e}_j := \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{K}^n, \quad \text{('1' tylko na } j\text{-tym miejscu)} \qquad \boldsymbol{f}_i := \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{K}^m. \quad \text{('1' na } i\text{-tym miejscu)}$$

Wtedy oczywiście $\forall \boldsymbol{x} \in \mathbb{K}^n : [\boldsymbol{x}]^e = \boldsymbol{x}$ oraz $\forall \boldsymbol{y} \in \mathbb{K}^n : [\boldsymbol{y}]^f = \boldsymbol{y}$, a zatem z tożsamości $[F(v)]^f = [F]^f_e [v]^e$ wynika następujący prosty

295. **Fakt**. Dla $F \in \mathrm{L}(\mathbb{K}^n; \mathbb{K}^m)$ istnieje jednoznacznie określona macierz $\boldsymbol{F} \in \mathbb{K}^m{}_n$, taka że $\forall \boldsymbol{x} \in \mathbb{K}^n : \; F(\boldsymbol{x}) = \boldsymbol{F}\boldsymbol{x}$ (iloczyn macierzy i wektora kolumnowego); mianowicie $\boldsymbol{F}$ jest macierzą F wzgl. baz standardowych $\mathbb{K}^n$ i $\mathbb{K}^m$. Odwrotnie, dla $\boldsymbol{F} \in \mathbb{K}^m{}_n$ wzór $F(\boldsymbol{x}) := \boldsymbol{F}\boldsymbol{x}$ określa operator $F \in \mathrm{L}(\mathbb{K}^n; \mathbb{K}^m)$, którego macierzą wzgl. baz standardowych jest $\boldsymbol{F}$.

Tak więc, mamy parę wzajemnie odwrotnych izomorfizmów liniowych:

$$[\,\cdot\,]^f_e : \mathrm{L}(\mathbb{K}^n; \mathbb{K}^m) \xrightarrow{\cong} \mathbb{K}^m{}_n, \quad F \mapsto [F]^f_e,$$

oraz

$$\mathbb{K}^m{}_n \xrightarrow{\cong} \mathrm{L}(\mathbb{K}^n; \mathbb{K}^m), \; \boldsymbol{F} \mapsto \begin{pmatrix} \text{operator mnożenia przez macierz } \boldsymbol{F}, \\ \text{tzn. } \mathbb{K}^n \ni \boldsymbol{x} \mapsto F(\boldsymbol{x}) := \boldsymbol{F}\,\boldsymbol{x} \in \mathbb{K}^m \end{pmatrix}.$$

W ten sposób wyczerpująco opisaliśmy ogólną postać odwzorowań liniowych między 'przestrzeniami arytmetycznymi', tzn. przestrzeniami postaci $\mathbb{K}^n$.

W dalszym ciągu, dla uniknięcia konieczności oznaczania baz standardowych $\mathbb{K}^n, \mathbb{K}^m$ jakimiś literami, będziemy zamiast $[\,\cdot\,]^f_e$ pisać $[\,\cdot\,]^{\mathrm{st}}_{\mathrm{st}}$.

296. **Przypomnienie**. Jeśli $\boldsymbol{A} \in \mathbb{K}^m{}_n$, to $\ker \boldsymbol{A} := \{\boldsymbol{x} \in \mathbb{K}^n : \boldsymbol{A}\boldsymbol{x} = 0\} \subset \mathbb{K}^n$, $\operatorname{im} \boldsymbol{A} := \langle \boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \rangle \subset \mathbb{K}^m$ oraz $\operatorname{rk} \boldsymbol{A} := \dim \operatorname{im} \boldsymbol{A}$; są to *jądro, obraz* i *rząd* $\boldsymbol{A}$.

297. **Ćwiczenie**. Przy dotychczasowych oznaczeniach niech $\boldsymbol{A} := [F]^f_e$. Wtedy $\operatorname{rk} F = \operatorname{rk} \boldsymbol{A}$, a ponadto

$$v \in \ker F \iff [v]^e \in \ker \boldsymbol{A}, \quad \text{tzn.} \quad \ker F = \{\sum_{j=1}^{n} e_j x^j : \; \boldsymbol{x} \in \ker \boldsymbol{A}\};$$

$$w \in \operatorname{im} F \iff [w]^f \in \operatorname{im} \boldsymbol{A}, \quad \text{tzn.} \quad \operatorname{im} F = \{\sum_{i=1}^{m} f_i y^i : \; \boldsymbol{y} \in \operatorname{im} \boldsymbol{A}\}.$$

Ponieważ $\ker \boldsymbol{A}$, $\operatorname{im} \boldsymbol{A}$ oraz $\operatorname{rk} \boldsymbol{A}$ są jądrem, obrazem i rzędem odpowiadającego macierzy $\boldsymbol{A}$ operatora liniowego, więc z 'bilansu wymiarów' 268. dostajemy

298. **Fakt** (*bilans wymiarów*): $\boxed{\dim \ker \boldsymbol{A} + \dim \operatorname{im} \boldsymbol{A} = n}$ dla $\boldsymbol{A} \in \mathbb{K}^m_n$.

Posłużmy się teraz faktem 274 mówiącym o tym, że złożenie odwzorowań liniowych jest odwzorowaniem liniowym, oraz 295 o tym, że bijektywna jest odpowiedniość pomiędzy macierzami a odwzorowaniami liniowymi przestrzeni kartezjańskich.

299. **Definicja**. *Iloczyn* dwóch macierzy $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^n{}_p$ jest macierzą $\boldsymbol{AB} \in \mathbb{K}^m{}_p$, odpowiadającą złożeniu $A \circ B : \mathbb{K}^p \xrightarrow{B} \mathbb{K}^n \xrightarrow{A} \mathbb{K}^m$, operatorów A, B, odpowiadających $\boldsymbol{A}$ i $\boldsymbol{B}$. Ponieważ $A(\boldsymbol{x}) = \boldsymbol{Ax}$, więc

$$\boldsymbol{C} = \boldsymbol{AB} \overset{\text{def}}{\iff} C = AB \overset{\text{def}}{\iff} \forall \boldsymbol{u} \in \mathbb{K}^p : \boldsymbol{Cu} = \boldsymbol{A}(\boldsymbol{Bu}).$$

300. Mnożenie macierzy jest *łączne* (bo łączne jest składanie odwzorowań); mamy też wzór $\boxed{\boldsymbol{C} = \boldsymbol{AB} \iff \forall i \in \overline{1,m},\ k \in \overline{1,p} : C^i{}_k = \sum_{j=1}^{n} A^i{}_j B^j{}_k}$.

$\boldsymbol{x} = \boldsymbol{Bu} \iff x^j := \sum_k B^j{}_k u^k$, więc $\boldsymbol{y} = \boldsymbol{Cu} = \boldsymbol{A}(\boldsymbol{Bu}) = \boldsymbol{Ax}$ ma współrzędne $y^i = \sum_j A^i{}_j x^j = \sum_j A^i{}_j \sum_k B^j{}_k u^k = \sum_k \Big(\sum_j A^i{}_j B^j{}_k\Big) u^k$, zarazem zaś $y^i = \sum_k C^i{}_k u^k$.

Uwaga. W przypadku, gdy $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^\nu{}_p$ i $n \neq \nu$ (tzn. gdy wartości operatora B nie należą do dziedziny A) iloczyn $\boldsymbol{AB}$ nie ma sensu (nie jest zdefiniowany).

301. **Definicja**. Macierz $\boldsymbol{I}_n := \begin{bmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ 0 & 0 & \dots \\ \cdot & \cdot & \cdots \end{bmatrix} \in \mathbb{K}^n{}_n$ o wyrazach $\delta^i{}_j = \begin{cases} 0, & i \neq i \\ 1, & i = j \end{cases}$ (*delta Kroneckera*), nazywamy *macierzą jednostkową*; ma ona własność $\boldsymbol{AI}_n = \boldsymbol{A}$, $\boldsymbol{I}_n\boldsymbol{B} = \boldsymbol{B}$ dla każdych $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^n{}_m$. Jeśli $\boldsymbol{BA} = \boldsymbol{I}_n$, to $\boldsymbol{B}$ nazywamy *lewą odwrotnością* $\boldsymbol{A}$, a $\boldsymbol{A}$ — *prawą odwrotnością* $\boldsymbol{B}$.

Uwaga. Z równości $\boldsymbol{BA} = \boldsymbol{I}_n$ wynika $m \geqslant n$, bowiem wtedy kolumny $\boldsymbol{A}$ (których jest n, a należą do $\boldsymbol{K}^m$) są liniowo niezależne: $\boldsymbol{Ax} = 0 \Rightarrow (\boldsymbol{BA})\boldsymbol{x} = 0$, tzn. $\boldsymbol{x} = 0$.

Odwrotnością macierzy $\boldsymbol{A}$ nazywamy macierz $\boldsymbol{B}$, dla której każdy z iloczynów $\boldsymbol{AB}, \boldsymbol{BA}$ jest macierzą jednostkową; jest to możliwe tylko przy $m = n$, więc tylko macierz kwadratowa $\boldsymbol{A} \in \mathbb{K}^n{}_n$ może być *odwracalna*.

302. **Fakt**. Przyporządkowanie $F \mapsto [F]^f{}_e$ jest nie tylko liniowe, ale także *multiplikatywne*, tzn. składaniu odwzorowań odpowiada mnożenie ich macierzy. Dokładniej, jeśli d, e i f są bazami przestrzeni (odpowiednio) U, V i W oraz $U \xrightarrow{G} V \xrightarrow{F} W$ (tzn. $G \in \mathrm{L}(U;V)$, $F \in \mathrm{L}(V;W)$), to

$$\boxed{[F \circ G]^f{}_d = [F]^f{}_e\,[G]^e{}_d}.$$

Z definicji $[G(u)] = [G][u]$ oraz $[F(v)] = [F][v]$, więc biorąc $v = G(u)$ dostajemy $[(F \circ G)(u)] = [F(v)] = [F][v] = [F][G(u)] = [F][G][u]$, skąd dzięki 293. wynika teza.

303. 'Algebrą' nazywa się strukturę algebraiczną, będącą 'połączeniem dwóch innych struktur: pierścienia i przestrzeni wektorowej'. Dokładniej:

Przestrzeń wektorowa A nad ciałem $\mathbb{K}$, wyposażona w operację 'mnożenia' swoich elementów: $A \times A \ni (a_1, a_2) \mapsto a_1 a_2 \in A$, taką że iloczyn $a_1 a_2$ zależy liniowo od każdego z czynników (tym bardziej więc mnożenie to jest rozdzielne względem dodawania), nosi nazwę $\mathbb{K}$-*algebry*.

Przykładem K-algebry jest przestrzeń $M_n(\mathbb{K}) := \mathbb{K}^n{}_n$, wyposażona w zwykłe mnożenie macierzy; jak wiemy, jest to algebra łączna, nieprzemienna (gdy $n > 1$) oraz ma *jedynkę*, którą jest macierz jednostkowa $\boldsymbol{I}_n$. Innym przykładem $\mathbb{K}$-algebry

jest przestrzeń $\operatorname{End} V := \mathrm{L}(V;V)$ tzw. *endomorfizmów przestrzeni wektorowej* V, w której mnożeniem jest składanie odwzorowań; id_V jest oczywiście jedynką $\operatorname{End} V$.

304. **Wniosek**. Przy ustalonej bazie e odwzorowanie $[\,\cdot\,]^e{}_e : \operatorname{End} V \to \mathbb{K}^n{}_n$ jest izomorfizmem algebr.

305. Wprost z definicji $\boxed{[\mathrm{id}_V]^e{}_e = \boldsymbol{I}_n}$ dla dowolnej bazy $e = (e_1, \ldots, e_n)$ przestrzeni V. Jeśli $\tilde{e} = (\tilde{e}_1, \ldots, \tilde{e}_n)$ jest inną bazą V, to macierze $[\mathrm{id}_V]^e{}_{\tilde{e}}$ oraz $[\mathrm{id}_V]^{\tilde{e}}{}_e$ noszą nazwę *macierzy przejścia* (lub *transformacji*) między tymi bazami albo krócej: *macierzy zmiany bazy*. Zauważmy, że te dwie macierze przejścia są wzajemnie odwrotne:

$$[\mathrm{id}_V]^e{}_{\tilde{e}}[\mathrm{id}_V]^{\tilde{e}}{}_e = [\mathrm{id}_V]^e{}_e = \boldsymbol{I}_n, \qquad [\mathrm{id}_V]^{\tilde{e}}{}_e[\mathrm{id}_V]^e{}_{\tilde{e}} = [\mathrm{id}_V]^{\tilde{e}}{}_{\tilde{e}} = \boldsymbol{I}_n$$

Natychmiastową konsekwencją 302 jest

306. **Fakt** (*transformacje macierzy wektora i operatora przy zmianie bazy*):

$$\boxed{[v]^{\tilde{e}} = [\mathrm{id}_V]^{\tilde{e}}{}_e[v]^e}, \qquad \boxed{[F]^{\tilde{f}}{}_{\tilde{e}} = [\mathrm{id}_W]^{\tilde{f}}{}_f[F]^f{}_e[\mathrm{id}_V]^e{}_{\tilde{e}}}.$$

Pierwszy wzór to równość typu $[F(v)] = [F][v]$ dla $F = \mathrm{id}_V$. Do uzyskania drugiego wystarczy zastosować równość typu $[FG] = [F][G]$ dla złożeń $\mathrm{id}_W \circ F$ oraz $F \circ \mathrm{id}_V$.

307. **Fakt**. Ustalmy (*starą*) bazę $e = (e_1, \ldots, e_n)$ przestrzeni V;

(1) Jeśli $\boldsymbol{A} \in \mathbb{K}^n{}_n$ jest macierzą kwadratową, to układ wektorów

$$\tilde{e}_1 := \sum_{i=1}^n e_i A^i{}_1, \ \ldots, \ \boxed{\tilde{e}_j := \sum_{i=1}^n e_i A^i{}_j}, \ \ldots, \ \tilde{e}_n := \sum_{i=1}^n e_i A^i{}_n$$

jest (*nową*) bazą $V \iff \boldsymbol{A}$ jest macierzą odwracalną. (2) W tym przypadku wektor $v \in V$, mający w nowej bazie współrzędne $\tilde{x}^1, \ldots, \tilde{x}^n$, ma w starej bazie współrzędne $\boxed{x^i = \sum_{j=1}^n A^i{}_j \tilde{x}^j}$. (3) Odwrotne zależności mają postać $\boxed{e_i = \sum_{j=1}^n \tilde{e}_j B^j{}_i}$, $\boxed{\tilde{x}^j = \sum_{i=1}^n B^j{}_i x^i}$, gdzie $\boldsymbol{B} := \boldsymbol{A}^{-1}$.

(4) Powyższe wzory można zapisać w krótszej formie macierzowej:

$$\boxed{\boldsymbol{x} = \boldsymbol{A}\tilde{\boldsymbol{x}}}, \qquad \boxed{\tilde{\boldsymbol{x}} = \boldsymbol{B}\boldsymbol{x}}, \qquad \boxed{\tilde{e} = e\boldsymbol{A}}, \qquad \boxed{e = \tilde{e}\boldsymbol{B}}.$$

(1) Gdy układ $\tilde{e}_i$ jest bazą, wtedy $\boldsymbol{A} = [\mathrm{id}_V]^e{}_{\tilde{e}}$ ma odwrotność, mianowicie $[\mathrm{id}_V]^{\tilde{e}}{}_e$. Odwrotnie, $\boldsymbol{A}^{-1} =: \boldsymbol{B} \Rightarrow \sum_{j=1}^n \tilde{e}_j B^j{}_k = \sum_{i,j=1}^n e_i A^i{}_j B^j{}_k = \sum_{i=1}^n e_i (\sum_{j=1}^n A^i{}_j B^j{}_k) = \sum_{i=1}^n e_i \delta^i{}_k =$ e_k, więc $\tilde{e}_*$, tak jak e_*, rozpinają V i jest ich $\dim V$, więc tworzą bazę, zob. 233.(c).

---

308. **Przykład**. Niech $V = W = \mathbb{K}_n[\,\cdot\,]$ (a więc $\dim V = \dim W = n+1$, odmiennie niż dotychczas); określmy operator $F \in \mathrm{L}(V;V)$ wzorem

$$(F(v))(t) := v(t+a), \text{ gdzie } a \in \mathbb{K} \text{ jest ustaloną liczbą.}$$

Niech $e = (e_0, e_1, \ldots, e_n)$ będzie bazą V, złożoną z jednomianów, tzn. $e_0(t) := 1$, $e_1(t) := t$, $\ldots$, $e_n(t) := t^n$ (potęgi t). Wzór dwumienny $(t+a)^j = \sum_{i=0}^n \binom{j}{i} a^{j-i} t^i$ oznacza, że $F(e_j) = \sum_{i=0}^j \binom{j}{i} a^{j-i} e_i$, a zatem macierz

$[F]^e_{\ e} = [F^i{}_j]_{i,j\in\overline{0,n}}$ ma wyrazy([39]) $F^i{}_j = \begin{cases} \binom{j}{i}a^{j-i}, & \text{dla } 0 \leqslant i \leqslant j \leqslant n, \\ 0, & \text{dla } 0 \leqslant j < i \leqslant n. \end{cases}$

Np. dla $n = 3$ dostajemy macierz $[F]^e_{\ e} = \begin{bmatrix} 1 & a & a^2 & a^3 \\ 0 & 1 & 2a & 3a^2 \\ 0 & 0 & 1 & 3a \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Na marginesie tego przykładu zauważmy, że czasem warto odstąpić od reguły, że zbiór numerów (wektorów bazy, wierszy lub kolumn macierzy, współrzędnych wektora itd.) musi być przedziałem postaci $\overline{1, \text{liczba}}$. W tym przykładzie odrzucenie tej reguły nie odgrywało istotnej roli: zamiast $e_i(t) = t^i$ moglibyśmy wprowadzić jednomiany $\varepsilon_i(t) := t^{i-1}$ numerowane liczbami $1, 2, \ldots$; lecz już np. w przestrzeni stanów wewnętrznych cząstki o spinie $\frac{3}{2}$ 'numery' $-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ bywają wygodniejsze od numerów $1, 2, 3, 4$ choćby o tyle, że mają bezpośrednią interpretację fizyczną.

W ogólnym przypadku operowanie wzorami postaci "$F(e_j) = \sum\limits_{i\in I} f_i\, F^i{}_j$ *dla* $j \in J$" nie jest ani trochę trudniejsze, niż wzorami postaci "$F(e_j) = \sum\limits_{i=1}^{m} f_i\, F^i{}_j$ *dla* $j \in \overline{1,n}$".

309. **Przykład**. Jeśli dla $n = 3$ oprócz $e = (e_0, e_1, e_2, e_3)$ weźmiemy drugą bazę $f := (e_3, e_2, e_1, e_0)$ przestrzeni $V = W$ z poprzedniego przykładu, to

$$[F]^f_{\ e} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3a \\ 0 & 1 & 2a & 3a^2 \\ 1 & a & a^2 & a^3 \end{bmatrix},\ [F]^e_{\ f} = \begin{bmatrix} a^3 & a^2 & a & 1 \\ 3a^2 & 2a & 1 & 0 \\ 3a & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ [F]^f_{\ f} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3a & 1 & 0 & 0 \\ 3a^2 & 2a & 1 & 0 \\ a^3 & a^2 & a & 1 \end{bmatrix}.$$

310. **Przykład**. Układ $\tilde{e} = (\tilde{e}_0, \tilde{e}_1, \tilde{e}_2, \tilde{e}_3)$, gdzie $\tilde{e}_i(t) := (t + a)^i$ ('przesunięte jednomiany') jest oczywiście także bazą $V = \mathbb{K}_3[\,\cdot\,]$ oraz dla F z poprzedniego przykładu $[F]^{\tilde{e}}_{\ e} = \boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (macierz jednostkowa!).

311. **Przykład**. Określmy operator $F \in \mathrm{L}(V; \mathbb{K}^3)$, $V := \mathbb{K}_2[\,\cdot\,]$, wzorem $F(v) := \begin{bmatrix} v(a) \\ v(b) \\ v(c) \end{bmatrix}$, gdzie $a, b, c \in \mathbb{K}$ są ustalone. Jeśli $e := (e_0, e_1, e_2)$ jest, jak wyżej, bazą jednomianów w V, zaś f — bazą standardową $\mathbb{K}^3$, to $[F]^f_{\ e} = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$, mamy bowiem $F(e_j) = \begin{bmatrix} e_j(a) \\ e_j(b) \\ e_j(c) \end{bmatrix} = \begin{bmatrix} a^j \\ b^j \\ c^j \end{bmatrix}$.

W przypadku, gdy dane $a, b, c \in \mathbb{K}$ są parami różne, operator F ma odwrotność: dla $\alpha, \beta, \gamma \in \mathbb{K}$ istnieje (dokładnie jeden) wielomian $v \in V$ spełniający warunek

$$F(v) = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix},\ \text{tzn. taki, że } v(a) = \alpha,\ v(b) = \beta,\ v(c) = \gamma;$$

gdyż $v \in V$, określony wzorem $v(t) = \alpha\frac{(t-b)(t-c)}{(a-b)(a-c)} + \beta\frac{(t-a)(t-c)}{(b-a)(b-c)} + \gamma\frac{(t-a)(t-b)}{(c-a)(c-b)}$, ma

[39] Symbol dwumienny $\binom{x}{i} := \frac{x(x-1)\ldots(x-i+1)}{i!}$, $i \in \mathbb{Z}_+$, ma dobrze znaną własność $\binom{0}{i} = \ldots = \binom{i-1}{i} = 0$, a więc wzór $F^i{}_j = \binom{j}{i}a^{j-i}$ jest słuszny dla *wszystkich* par $i, j \in \overline{0,n}$.

żądane własności; rozkładając w bazie e ten wielomian v, tzn. $F^{-1}\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix}$, dostajemy $[F^{-1}]^e{}_f = \frac{1}{D}\begin{bmatrix} bc^2-b^2c & ca^2-c^2a & ab^2-a^2b \\ b^2-c^2 & c^2-a^2 & a^2-b^2 \\ c-b & a-c & b-a \end{bmatrix}$, gdzie $D := (b-a)(c-a)(c-b)$.

Z własności $[FG]=[F][G]$ widać, że dla odwracalnego operatora F macierz $[F^{-1}]^e{}_f$ jest odwrotnością $[F]^f{}_e$; zatem znaleźliśmy $\begin{bmatrix}1 & a & a^2\\ 1 & b & b^2\\ 1 & c & c^2\end{bmatrix}^{-1}$. W ten sam sposób można znaleźć odwrotność analogicznej «macierzy Vandermonde'a» $\boldsymbol{A}\in\mathbb{K}^n{}_n$ dla $n\in\mathbb{N}$.

5.3 Równanie liniowe niejednorodne Twierdzenie Kroneckera-Capelliego

312. **Ćwiczenie**. Wychodząc ze spostrzeżenia, że $\boldsymbol{Ax} = \boldsymbol{A}_1x^1+\ldots+\boldsymbol{A}_nx^n$ (kombinacja liniowa kolumn $\boldsymbol{A}$ ze współczynnikami $x^1,\ldots,x^n$) dowieść, że dla danych $\boldsymbol{A}\in\mathbb{K}^m{}_n$, $\boldsymbol{b}\in\mathbb{K}^m$ następujące warunki są równoważne:

(1) równanie $\boldsymbol{Ax}=\boldsymbol{b}$ ma rozwiązanie $\boldsymbol{x}\in\mathbb{K}^n$;
(2) $\boldsymbol{b}\in\operatorname{im}\boldsymbol{A}$;
(3) $\operatorname{rk}\boldsymbol{A} = \operatorname{rk}\widetilde{\boldsymbol{A}}$, gdzie $\widetilde{\boldsymbol{A}} = [\boldsymbol{A}\,|\,\boldsymbol{b}]\in\mathbb{K}^m{}_{n+1}$ jest *macierzą rozszerzoną* układu $\boldsymbol{Ax}=\boldsymbol{b}$, tzn. macierzą powstałą z zestawienia macierzy $\boldsymbol{A}$ i $\boldsymbol{b}$.

Jeśli równoważne warunki (1)..(3) nie są spełnione, to równanie $\boldsymbol{Ax}=\boldsymbol{b}$ nie ma rozwiązań. Zajmiemy się teraz zbadaniem przypadku, gdy są one spełnione:

313. **Spostrzeżenie**. Jeśli wektor $\hat{\boldsymbol{x}}\in\mathbb{K}^n$ spełnia równanie $\boldsymbol{A}\hat{\boldsymbol{x}}=\boldsymbol{b}$, to

$$\boldsymbol{Ax}=\boldsymbol{b} \quad\Longleftrightarrow\quad \boldsymbol{A}(\boldsymbol{x}-\hat{\boldsymbol{x}})=0,\ \text{tzn. } \boldsymbol{x}-\hat{\boldsymbol{x}}\in\ker\boldsymbol{A};$$

zatem każde rozwiązanie równania $\boldsymbol{Ax}=\boldsymbol{b}$ jest postaci $\boldsymbol{x}=\hat{\boldsymbol{x}}+\boldsymbol{u}$, gdzie $\boldsymbol{u}\in\ker\boldsymbol{A}$, a więc zależy od $\dim\ker\boldsymbol{A} = n-\operatorname{rk}\boldsymbol{A}$ parametrów. W szczególności

$$\begin{pmatrix}\text{równanie } \boldsymbol{Ax}=\boldsymbol{b} \text{ ma}\\ \text{dokładnie jedno rozwiązanie}\end{pmatrix} \Longleftrightarrow \quad \operatorname{rk}\boldsymbol{A} = \operatorname{rk}\widetilde{\boldsymbol{A}} = n.$$

Uwagi. (a) Ostatni warunek $\operatorname{rk}\boldsymbol{A}=\operatorname{rk}\widetilde{\boldsymbol{A}}=n$ oznacza oczywiście, że kolumny $\boldsymbol{A}$, których jest n, są liniowo niezależne, a wektor $\boldsymbol{b}$ jest ich kombinacją liniową.
(b) Warunek $\boldsymbol{Ax}=\boldsymbol{b}$ można równie dobrze nazywać *równaniem na wektor* $\boldsymbol{x}$, jak też *układem równań na niewiadome* $x^1,\ldots,x^n$.
(c) Powyższy rozkład $\boldsymbol{x}=\hat{\boldsymbol{x}}+\boldsymbol{u}$ często tradycyjnie zapisuje się w postaci

$$\text{RORN} = \text{RSRN} + \text{RORJ}, \qquad (*)$$

gdzie

RORN	oznacza	*rozwiązanie ogólne równania niejednorodnego,*
RSRN	—	*rozwiązanie szczególne równania niejednorodnego,*
a RORJ	—	*rozwiązanie ogólne równania jednorodnego.*

Zwykle w rozkładzie (*) przedstawia się RORJ w postaci $\boldsymbol{u}=C_1\boldsymbol{u}_1+\ldots C_s\boldsymbol{u}_s$, gdzie $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_s$ jest jakąś bazą przestrzeni $\ker\boldsymbol{A}=\{\boldsymbol{u}:\boldsymbol{Au}=0\}$, zaś $C_1,\ldots,C_s\in\mathbb{K}$ — dowolnymi parametrami. Można jednak formalny rozkład (*) rozumieć inaczej, jako równość dwóch zbiorów, traktując RORJ i RORJ jako zbiory $\{\boldsymbol{x}:\boldsymbol{Ax}=\boldsymbol{b}\}$

i $\ker \boldsymbol{A}$; wtedy (*) oznacza, że *zbiór* RORN *jest rezultatem przesunięcia o wektor* $\hat{\boldsymbol{x}}$ = RSRN *podprzestrzeni* RORJ. Przy tym $\hat{\boldsymbol{x}}$ można wybrać dowolnie z RORN.

(d) Obliczanie rzędów $\boldsymbol{A}$ i $\widetilde{\boldsymbol{A}}$ dla zbadania istnienia rozwiązania równania $\boldsymbol{Ax} = \boldsymbol{b}$ o zadanych *liczbowych* macierzach $\boldsymbol{A}, \boldsymbol{b}$ jest niepraktyczne: wymaga bowiem prawie tylu obliczeń, co całkowite rozwiązanie równania metodą operacji elementarnych, przy którym oba rzędy dostajemy gratis, jako 'produkty uboczne'. Natomiast obliczanie rzędów tych macierzy metodą wyznacznikową (propagowane w niektórych podręcznikach i poradnikach *szukanie niezerowych minorów maksymalnego stopnia*) jest już zupełnym absurdem: nakład obliczeń, zwłaszcza gdy wymiary m, n nie są zbyt małe, jest tu niepomiernie duży w stosunku do otrzymanego rezultatu!

Inaczej się ma sprawa w przypadku, gdy dane macierze $\boldsymbol{A}, \boldsymbol{b}$ nie są numeryczne, lecz np. zależą od jakichś parametrów: wtedy metoda operacji elementarnych (i inne metody numeryczne) stają się nieadekwatne i, chcąc niechcąc, musimy się uciekać do metod nienumerycznych, takich jak wyznaczniki, minory itp., czy choćby oba powyższe rezultaty, zwane tradycyjnie *twierdzeniem Kroneckera-Capelliego*.

5.4 Transpozycja macierzy

314. **Definicja**. Niech $\boldsymbol{A} \in \mathbb{K}^m{}_n$. Macierz $\tilde{\boldsymbol{A}} \in \mathbb{K}^n{}_m$, o wyrazach $\tilde{A}^j{}_i := A^i{}_j$, $i \in \overline{1,m}$, $j \in \overline{1,n}$, nazywamy *macierzą transponowaną względem* $\boldsymbol{A}$ lub, krócej, *transpozycją* macierzy $\boldsymbol{A}$; oznaczamy ją symbolem $\boldsymbol{A}^T$:

$$\boldsymbol{A} = \begin{bmatrix} A^1{}_1 & \cdots & A^1{}_n \\ \vdots & \dots & \vdots \\ A^m{}_1 & \cdots & A^m{}_n \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{A}^T = \begin{bmatrix} A^1{}_1 & \cdots & A^m{}_1 \\ \vdots & \dots & \vdots \\ A^1{}_n & \cdots & A^m{}_n \end{bmatrix}.$$

315. **Fakt**. Operacja transpozycji $\boldsymbol{A} \mapsto \boldsymbol{A}^T$ jest odwzorowaniem:
(1) liniowym; (2) inwolutywnym, tzn. $(\boldsymbol{A}^T)^T = \boldsymbol{A}$;
(3) 'antymultiplikatywnym', tzn. $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$.

Ad(3): Niech $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^n{}_p$, $\boldsymbol{C} := \boldsymbol{AB} \in \mathbb{K}^m{}_p$. Oznaczmy dla wygody $\tilde{\boldsymbol{A}} := \boldsymbol{A}^T$, $\tilde{\boldsymbol{B}} := \boldsymbol{B}^T$, $\tilde{\boldsymbol{C}} := \boldsymbol{C}^T$. Równość $\boldsymbol{C} = \boldsymbol{AB}$ oznacza, że $C^i{}_k = \sum\limits_j A^i{}_j B^j{}_k$ ($i \in \overline{1,m}$, $j \in \overline{1,n}$, $k \in \overline{1,p}$); wstawiając tu $A^i{}_j = \tilde{A}^j{}_i$, $B^k{}_j = \tilde{B}^k{}_j$, $C^i{}_k = \tilde{C}^k{}_i$ dostajemy równość $\tilde{C}^k{}_i = \sum\limits_j \tilde{A}^j{}_i \tilde{B}^k{}_j = \sum\limits_j \tilde{B}^k{}_j \tilde{A}^j{}_i$, oznaczającą, że $\tilde{\boldsymbol{C}} = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{A}}$.

Ćwiczenie. Jeśli $\boldsymbol{A} \in M_n(\mathbb{K}) = \mathbb{K}^n{}_n$ ma odwrotność, to ma ją również $\boldsymbol{A}^T$ oraz

$$\boxed{(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T}.$$

5.5 Rząd 'wierszowy' i 'kolumnowy' macierzy

316. Zdefiniujmy *rząd kolumnowy* i *rząd wierszowy* macierzy $\boldsymbol{A} \in \mathbb{K}^m{}_n$ jako wymiary podprzestrzeni rozpiętych przez jej kolumny bądź wiersze:

$$\operatorname{rk}_K \boldsymbol{A} := \dim \langle \boldsymbol{A}_1, \dots, \boldsymbol{A}_n \rangle, \qquad \operatorname{rk}_K \boldsymbol{A} := \dim \left\langle \boldsymbol{A}^1, \dots, \boldsymbol{A}^m \right\rangle.$$

Wprost z określenia $\mathrm{rk}_K \boldsymbol{A}, \mathrm{rk}_W \boldsymbol{A} \in \overline{0, \min\{m,n\}}$. Pokażemy wkrótce m.in., że dla każdej macierzy $\boldsymbol{A}$ zachodzi równość $\mathrm{rk}_K \boldsymbol{A} = \mathrm{rk}_W \boldsymbol{A}$.

317. Przypomnijmy najpierw, że jeśli $\boldsymbol{w} = [w_1 \ \ldots \ w_n]$ jest wektorem wierszowym, a $\boldsymbol{x} = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}$ — wektorem kolumnowym, to $\boldsymbol{wx} = \sum_{n=1}^{n} w_n x^n \in \mathbb{K}$ jest liczbą; zależy ona liniowo zarówno od $\boldsymbol{w}$, jak i od $\boldsymbol{x}$. W przypadku, gdy $\boldsymbol{wx} = 0$, mówi się, że $\boldsymbol{w}$ i $\boldsymbol{x}$ wzajemnie się *zerują* lub *anihilują* albo że są *prostopadłe*([40]), pisząc $\boldsymbol{w} \perp \boldsymbol{x}$.

Jeśli $W \subset \mathbb{K}_n$ jest podzbiorem przestrzeni $\mathbb{K}_n$ (wektorów wierszowych), to

$$W^0 := \{\boldsymbol{x} \in \mathbb{K}^n : x \perp W, \text{ tzn. } \forall w \in W : \boldsymbol{wx} = 0\} \subset \mathbb{K}^n$$

jest (oczywiście) podprzestrzenią, zwaną *anihilatorem*([41]) zbioru W. W taki sam sposób określa się anihilator $X^0 \subset \mathbb{K}_n$ podzbioru $X \subset \mathbb{K}^n$.

318. **Fakt** ('metamatematyczny'). Każde z trzech poniższych twierdzeń jest prostą konsekwencją dwóch pozostałych:

T.1 $\forall \boldsymbol{A} \in K^m{}_n : \dim\ker \boldsymbol{A} + \dim\mathrm{im}\, \boldsymbol{A} = n$ ('bilans wymiarów').
T.2 $\forall \boldsymbol{A} \in K^m{}_n : \mathrm{rk}_W \boldsymbol{A} = \mathrm{rk}_K \boldsymbol{A}$ ('o równości rk_K i rk_W').
T.3 $\forall\, W \subset \mathbb{K}_n : \dim W^0 + \dim W = n$ ('o wymiarze anihilatora').

Jeśli $\boldsymbol{A}$ i W są 'stowarzyszone' w tym sensie, że $W = \langle \boldsymbol{A}^1, \ldots, \boldsymbol{A}^m \rangle \subset \mathbb{K}_n$, to $\dim W = \mathrm{rk}_W \boldsymbol{A}$, $\ker \boldsymbol{A} = W^0$. Ponadto $\dim\mathrm{im}\, \boldsymbol{A} = \mathrm{rk}_K \boldsymbol{A}$, więc T.1, T.2, T.3 można przepisać w formie $n_0 + \mathrm{rk}_K \boldsymbol{A} = n$, $\mathrm{rk}_W \boldsymbol{A} = \mathrm{rk}_K \boldsymbol{A}$, $n_0 + \mathrm{rk}_W \boldsymbol{A} = n$, gdzie $n_0 := \dim\ker \boldsymbol{A}$, co powoduje, że teza staje się oczywista.

Przypomnijmy, że twierdzenie T.1. udowodniliśmy już wcześniej, zob. 298.

319. *Bezpośredni dowód* T.2. Niech $\boldsymbol{A}^{i_1}, \ldots, \boldsymbol{A}^{i_\mu}$ będzie maksymalnym układem l. niez. wierszy, natomiast $\boldsymbol{A}_{j_1}, \ldots, \boldsymbol{A}_{j_\nu}$ — maksymalnym układem l. niezal. kolumn $\boldsymbol{A}$. Niech $\tilde{\boldsymbol{A}} \in \mathbb{K}^\mu{}_n$ będzie podmacierzą $\boldsymbol{A}$, utworzoną z wierszy $\boldsymbol{A}^{i_1}, \ldots, \boldsymbol{A}^{i_\mu}$; mamy więc $\ker \tilde{\boldsymbol{A}} = \ker \boldsymbol{A}$. Pokażemy teraz, że kolumny $\tilde{\boldsymbol{A}}_{j_1}, \ldots, \tilde{\boldsymbol{A}}_{j_\nu}$ są l. niezależne:

Jeśli $\tilde{\boldsymbol{A}}_{j_1}\lambda^{j_1} + \ldots + \tilde{\boldsymbol{A}}_{j_\nu}\lambda^{j_\nu} = 0$ dla pewnych współczynników $\lambda^{j_1}, \ldots, \lambda^{j_\nu} \in \mathbb{K}$, to dookreślając $\lambda^i := 0$ dla $i \in \overline{1,n} \setminus \{j_1, \ldots, j_\nu\}$, otrzymamy $\tilde{\boldsymbol{A}}_1\lambda^1 + \ldots + \tilde{\boldsymbol{A}}_n\lambda^n = 0$, czyli $\begin{bmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{bmatrix} \in \ker \tilde{\boldsymbol{A}}$, skoro zaś $\ker \tilde{\boldsymbol{A}} = \ker \boldsymbol{A}$, to mamy stąd $\boldsymbol{A}_1\lambda^1 + \ldots + \boldsymbol{A}_n\lambda^n = 0$, tzn. $\boldsymbol{A}_{j_1}\lambda^{j_1} + \ldots + \boldsymbol{A}_{j_\nu}\lambda^{j_\nu} = 0$, a więc $\lambda^{j_1} = \ldots = \lambda^{j_\nu} = 0$ wskutek l. niezależności.

Z istnienia l. niezal. układu ν kolumn $\tilde{\boldsymbol{A}}_{j_*}$ w przestrzeni $\mathbb{K}^\mu$ wynika, że $\nu \leqslant \mu$, czyli $\mathrm{rk}_K \boldsymbol{A} \leqslant \mathrm{rk}_W \boldsymbol{A}$. Odwracając role kolumn i wierszy dostajemy $\mathrm{rk}_K \boldsymbol{A} \geqslant \mathrm{rk}_W \boldsymbol{A}$.

320. **Ćwiczenie**. Uzupełnić szczegóły następującego bezpośredniego dowodu twierdzenia T.3; inny jego dowód poznamy w 348.

Jak wiemy (z ćwiczeń), każda macierz jest wierszowo równoważna pewnej macierzy wierszowo zredukowanej. Zatem daną podprzestrzeń $W \subset \mathbb{K}_n$ można przedstawić w

[40] Słowa 'prostopadłość' czy 'ortogonalność' są tu niezbyt trafne: nie ma sensu pojęcie 'kąta' między $\boldsymbol{w}$ a $\boldsymbol{x}$, zwłaszcza że te dwa wektory należą do dwóch różnych przestrzeni!
[41] Używa się też nazw *polara* i *dopełnienie ortogonalne* oraz oznaczeń $\mathrm{An}(W)$, $W^\perp$, itp.

postaci $W = \langle \boldsymbol{A}^1, \ldots, \boldsymbol{A}^m \rangle$, gdzie $\boldsymbol{A} \in \mathbb{K}^m{}_n$ jest WZ-macierzą o niezerowych wierszach, $m = \dim W$. Niech wyrazy $A^1_{j_1}, \ldots, A^m_{j_m}$ będą samotnikami kolumnowymi $\boldsymbol{A}$; skoro $j_1, \ldots, j_m \in \overline{1,n}$ są parami różne, to $Z := \{j_1, \ldots, j_m\}$ ma m elementów. Mamy przy tym: $\boldsymbol{x} \in W^0 \Leftrightarrow \boldsymbol{A}\boldsymbol{x} = 0 \Leftrightarrow \forall\, i \in \overline{1,m} : x^{j_i} = -\frac{1}{A^i_{j_i}} \sum\limits_{k \in \overline{1,n} \setminus Z} A^i{}_k x^k$. Oznacza to, że $W^0 = \{\boldsymbol{x} \in \mathbb{K}^n : \forall j \in Z : x^j = \sum\limits_{k \in \overline{1,n} \setminus Z} B^j{}_k x^k\}$, gdzie $B^j{}_k$ są pewnymi współczynnikami; zatem podprzestrzeń W^0 jest parametryzowana przez $n-m$ niezależnych wielkości x^k, $k \in \overline{1,n} \setminus Z$, czyli $\dim W^0 = n - m$.

321. **Definicja**. Widząc już, że zawsze $\mathrm{rk}_K \boldsymbol{A} = \mathrm{rk}_W \boldsymbol{A}$, wspólną wartość obu tych rzędów nazwijmy *rzędem* $\boldsymbol{A}$ oraz oznaczmy symbolem $\boxed{\mathrm{rk}\,\boldsymbol{A}}$.

322. **Fakt**. Jeśli $\boldsymbol{A}, \boldsymbol{A}' \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^n{}_p$, to spełnione są nierówności

$$\boxed{\mathrm{rk}(\boldsymbol{A} + \boldsymbol{A}') \leqslant \mathrm{rk}\,\boldsymbol{A} + \mathrm{rk}\,\boldsymbol{A}'},$$

$$\boxed{\mathrm{rk}\,\boldsymbol{A} \leqslant \min\{m,n\}}, \qquad \boxed{\mathrm{rk}(\boldsymbol{A}\boldsymbol{B}) \leqslant \min\{\mathrm{rk}\,\boldsymbol{A},\, \mathrm{rk}\,\boldsymbol{B}\}}.$$

$\mathrm{rk}\,\boldsymbol{A} = \mathrm{rk}\,A$, jeśli $A \in \mathrm{L}(\mathbb{K}^n;\mathbb{K}^m)$, $A(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$; stąd i z faktu 276. wynika teza.

Oto trzy przykładowe wnioski z twierdzenia o równości $\mathrm{rk}_W = \mathrm{rk}_K$:

323. **Fakt**. Operacje elementarne zarówno na kolumnach, jak i na wierszach macierzy, nie zmieniają jej rzędu.

Istotnie, przestrzeń rozpięta przez kolumny, a więc i $\mathrm{rk}_K \boldsymbol{A}$, jest niezmiennikiem operacje elementarnych na kolumnach; z kolei operacje elementarne na wierszach zachowują przestrzeń rozpiętę przez wiersze, a więc i $\mathrm{rk}_W \boldsymbol{A}$.

324. **Fakt**. Transpozycja nie zmienia rzędu macierzy, tzn. $\mathrm{rk}\left(\boldsymbol{A}^T\right) = \mathrm{rk}\,\boldsymbol{A}$.

Kolumny $\boldsymbol{A}$ są 'takie same', jak wiersze $\boldsymbol{A}^T$, więc $\mathrm{rk}_K \boldsymbol{A} = \mathrm{rk}_W \boldsymbol{A}^T$, skąd teza.

325. **Fakt**. Układ równań liniowych ma rozwiązanie wtedy i tylko wtedy, gdy jest 'liniowo niesprzeczny'. Dokładniej: jeśli $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $b \in \mathbb{K}^n$, to dwa następujące warunki są równoważne:

(1) $\boldsymbol{b} \notin \mathrm{im}\,\boldsymbol{A}$, czyli[42] układ $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ nie ma rozwiązania $\boldsymbol{x} \in \mathbb{K}^n$;

(2) układ $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ jest 'liniowo sprzeczny' w tym sensie, że równanie $0x^1 + \ldots + 0x^n = 1$ jest kombinacją liniową równań układu:

$$\exists\, \lambda_1, \ldots, \lambda_m \in \mathbb{K} : \left(\begin{array}{c} \lambda_1 \boldsymbol{A}^1 + \ldots + \lambda_m \boldsymbol{A}^m = 0, \\ \text{lecz } \lambda_1 b^1 + \ldots + \lambda_m b^m \neq 0 \end{array} \right).$$

(1) $\Rightarrow$ (2): Niech $\tilde{\boldsymbol{A}} := [\boldsymbol{A}|\boldsymbol{b}]$ będzie *macierzą rozszerzoną układu*. Z (1) wynika, że przestrzeń $\mathrm{im}\,\tilde{\boldsymbol{A}}$ jest większa od $\mathrm{im}\,\boldsymbol{A}$, a więc $\mathrm{rk}_K \tilde{\boldsymbol{A}} > \mathrm{rk}_K \boldsymbol{A}$. Skoro $\mathrm{rk}_K = \mathrm{rk}_W$, mamy $\mathrm{rk}_W \tilde{\boldsymbol{A}} > \mathrm{rk}_W \boldsymbol{A}$, więc jeśli $\tilde{\boldsymbol{A}}^{i_1}, \ldots, \tilde{\boldsymbol{A}}^{i_\mu}$ jest maksymalnym l. niez. układem wierszy $\tilde{\boldsymbol{A}}$, to układ $\boldsymbol{A}^{i_1}, \ldots, \boldsymbol{A}^{i_\mu}$ jest l. zależny: $\exists\,(\lambda_i) : \sum \lambda_i \boldsymbol{A}^i = 0,\ \sum \lambda_i \tilde{\boldsymbol{A}}^i \neq 0$.

326. **Uwaga**. Negacja warunku (2) ma postać $\forall \boldsymbol{\lambda} \in \mathbb{K}_m : \boldsymbol{\lambda}\boldsymbol{A} = 0 \Rightarrow \lambda \boldsymbol{b} = 0$, tzn. $\boldsymbol{b} \in W^0$, gdzie $W := \{\boldsymbol{\lambda} : \lambda \boldsymbol{A} = 0\} = \ker G$, zaś $G \in \mathrm{L}(\mathbb{K}_m;\mathbb{K}_n)$, $G(\boldsymbol{\lambda}) := \boldsymbol{\lambda}\boldsymbol{A}$.

Zatem fakt 325 oznacza, że $\mathrm{im}\,\boldsymbol{A} = (\ker G)^0$; w następnym rozdziale zobaczymy, że

[42] $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}_1 x^1 + \ldots + \boldsymbol{A} x^n$ jest kombinacją liniową kolumn $\boldsymbol{A}$ ze współczynnikami x^i.

operator G jest *transpozycją* ($G = F^T$) operatora $F \in \mathrm{L}(\mathbb{K}^n;\mathbb{K}^m)$, $F(\boldsymbol{x}) := \boldsymbol{A}\boldsymbol{x}$, zaś uogólnieniem faktu 325 jest ważny wzór $\operatorname{im} F = \left(\ker F^T\right)^0$, zob. dalej punkt 349.

5.6 Przestrzeń sprzężona

327. **Definicja**. Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$. Elementy zbioru $V^* := \mathrm{L}(V;\mathbb{K})$, tzn. odwzorowania liniowe $V \to \mathbb{K}$, nazywa się *formami* (lub *funkcjonałami*) liniowymi na przestrzeni V. Zbiór V^* wszystkich form liniowych na V jest, jak wiemy, przestrzenią wektorową (nad $\mathbb{K}$) względem zwykłych działań na funkcjach (dodawania i mnożenia przez skalary); nazywamy ją *przestrzenią sprzężoną* (lub *dualną*, lub *dwoistą*) do przestrzeni V.

Elementy przestrzeni V^* niekiedy (np. w *geometrii różniczkowej*), dla podkreślenia ich związku z 'wektorami' (tzn. elementami V) są nazywane 'kowektorami' przestrzeni V; tak więc 'kowektory' są wektorami, lecz innej — dualnej — przestrzeni.

328. **Przykład**. Jeśli $e_1,\ldots,e_n$ jest bazą V, to funkcje $e^1,\ldots,e^n : V \to \mathbb{K}$, określone wzorem $\boxed{e^i(v) := (i\text{-ta współrzędna } v \text{ w bazie } e_1,\ldots,e_n)}$, są formami liniowymi, tzn. $e^1,\ldots,e^n \in V^*$.

Liniowość funkcji e^i wyraża oczywisty fakt, że działaniom na wektorach (dodawaniu i mnożeniu przez liczby) odpowiadają analogiczne działania na ich współrzędnych.

329. **Fakt**. $\boxed{e^i(e_j) = \delta^i{}_j}$, gdzie $\delta^i{}_j := \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$ jest *symbolem Kroneckera*; mamy ponadto następujące równości:

$$\boxed{\begin{array}{ll} (1) & \forall v \in V : \quad v = \sum_{i=1}^{n} e_i\, e^i(v) \\ (2) & \forall \phi \in V^* : \phi = \sum_{i=1}^{n} \phi(e_i) e^i \end{array}}$$

Ad (1): v ma w bazie rozkład $v = \sum e_i x^i$, przy czym $x^i = e^i(v)$ z definicji e^i, QED.
Ad (2): Liniowość ϕ daje $\phi(v) = \phi\left(\sum e_i x^i\right) = \sum \phi(e_i) x^i = \sum \phi(e_i) e^i(v)$, QED([43]).

330. **Wniosek**. Układ form $e^1,\ldots,e^n$ jest bazą przestrzeni V^*; wobec tego:

Jeśli przestrzeń V ma skończony wymiar, to $\boxed{\dim V^* = \dim V}$.

Skoro $e^i(e_j) = \delta^i{}_j$, to kombinacja liniowa $\sum_i \lambda_i e^i$ przyjmuje na e_j wartość λ_j, skąd wynika liniowa niezależność. Z kolei ze wzoru (2) widać, że $e^1,\ldots,e^n$ rozpinają V^*.

331. **Definicja**. Bazę $e^1,\ldots,e^n$ (przestrzeni V^*) nazywamy *bazą sprzężoną* (lub *dualną*, lub *dwoistą*) względem bazy $e_1,\ldots,e_n$ (przestrzeni V).

332. **Uwaga**. Wzory (1) i (2) pokazują, że pod względem algebraicznych własności relacje między przestrzeniami V i V^* oraz bazami $e_1,\ldots,e_n$ i $e^1,\ldots,e^n$ są symetryczne. Z tego powodu często używa się innej ('braketowej') notacji, oznaczając wartość $\phi(v)$ symbolem $\langle \phi, v\rangle$. W takiej notacji wzory (1),(2) przyjmują następującą postać:

[43] Inny sposób: Obie strony wzoru (2) są formami liniowymi i przyjmują na każdym z wektorów e_i tę samą wartość $\phi(e_i)$; zatem ich różnica jest formą zerową.

$$\boxed{\begin{array}{ll}(1) & \forall v \in V: \quad v = \sum_{i=1}^{n} e_i \langle e^i, v\rangle \\ (2) & \forall \phi \in V^*: \phi = \sum_{i=1}^{n} \langle \phi, e_i\rangle e^i\end{array}}$$

Oczywiście z kontekstu należy się domyślić, czy np. $\langle a, b\rangle$ jest wartością formy a na wektorze b, czy przestrzenią rozpiętą przez dwa wektory a, b; wątpliwości zdarzają równie rzadko, jak np. wątpliwości, czy 'z' ma być ostatnią literą alfabetu, czy $\frac{dz}{dt}$.

333. **Fakt** (*kanoniczny izomorfizm* $V \xrightarrow{\cong} V^{**}$). Określmy $\mathsf{K} \in \mathrm{L}(V; (V^*)^*)$ wzorem $\boxed{(\mathsf{K}(v))(\phi) := \phi(v)}$, czyli $\boxed{\mathsf{K}(v) =: \widehat{v}, \text{ gdzie } \widehat{v}(\phi) := \phi(v)}$. Wtedy K jest liniową injekcją, a gdy $\dim V < \infty$ — izomorfizmem.

$\ker \mathsf{K} = \{v: \forall \phi \in V^*: \phi(v) = 0\}$, więc injektywność K, tzn. $\ker \mathsf{K} = \{0\}$, oznacza że *formy liniowe rozdzielają elementy* V: gdy $0 \neq v \in V$, wtedy $\exists\, \phi \in V^*: \phi(v) \neq 0$. Dla sprawdzenia tego *warunku rozdzielania* w przypadku $\dim V < \infty$ wystarczy (mając ustaloną bazę V) dla zadanego $0 \neq v \in V$ wziąć wziąć $\phi := e^i$, gdzie numer $i \in \overline{1,n}$ dobieramy do v w taki sposób, by i-ta współrzędna v była $\neq 0$.

W przypadku $\dim V = \infty$ rozumowania pozostaje słuszne, a jedyną różnicą jest to, że teraz istnienie bazy jest nietrywialne i wymaga aksjomatu wyboru, zob. 243([44]).

Dla $\dim V < \infty$ oprócz $\ker \mathsf{K} = \{0\}$ mamy też $\dim V = \dim V^* = \dim (V^*)^*$, a stąd jak wiemy wynika bijektywność K, zob. 285, QED.

334. **Uwaga 1**. Nie tylko formy liniowe (tj. elementy V^*) rozdzielają punkty V, ale (co jest oczywiste!) także elementy V (ściślej: funkcje $\widehat{v}$, $v \in V$) rozdzielają punkty V^*.

Uwaga 2. Przymiotnik 'kanoniczny' (lub 'naturalny') używany jest w matematyce w odniesieniu do obiektów, które nie zależą od żadnych arbitralnych wyborów (np. od wyborów baz lub układów współrzędnych), tzn. do obiektów określonych jedynie przez obiekty i struktury występujące *explicite* w rozważanej sytuacji([45]).

Przykład. Mając bazę $e_1, \ldots, e_n$ przestrzeni V, określmy izomorfizm $F: V \xrightarrow{\cong} V^*$ wzorem 'najprostszym możliwym', $F(e_1 x^1 + \ldots + e_n x^n) := e^1 x^1 + \ldots + e^n x^n$. Zobaczmy, jakim wzorem wyraża się F, jeśli zamiast e weźmiemy inną bazę, np. $f_1 := 2e_1 + 3e_2$, $f_i := e_i$ dla $\in \overline{2,n}$; wtedy $f^1 = \frac{1}{2}e^1$, $f^2 = e^2 - \frac{3}{2}e^1$, $f^i = e^i$ dla $i \in \overline{3,n}$, więc $F(f_1) = 2e^1 + 3e^2 = 4f^1 + 3(f^2 + \frac{3}{2} \cdot 2f^1) = 13f^1 + 3f^2$, $F(f_2) = e^2 = f^2 + \frac{3}{2} \cdot 2f^1 = 3f^1 + f^2$, $F(f_i) = f^i$ dla $i \in \overline{3,n}$. Widać więc, że F zależy w istotny sposób od wyboru bazy, a więc jest izomorfizmem niekanonicznym.

335. **Definicja**. Dla $F \in \mathrm{L}(V; W)$ oraz $\psi \in W^*$ operator $\psi \circ F: V \to \mathbb{K}$ jest liniowy (złożenie operatorów liniowych), a więc $\psi \circ F \in V^*$; tak określoną formę $F^T(\psi) := \psi \circ F$ nazywamy F-*cofnięciem*([46]) formy ψ.

Co więcej, zależność $\psi \circ F$ od formy ψ jest liniowa, więc otrzymany

[44] Inny dowód warunku rozdzielania opiera się na fakcie, że każda podprzestrzeń $V_1 \subset V$ ma podprzestrzeń dopełniającą, tzn. $V_0 \subset V$, taką że $V = V_0 \dotplus V_1$. Biorąc $V_1 := \langle v_1\rangle$, gdzie $v_1 \in V$ jest danym wektorem niezerowym, możemy określić wartość $\phi(v)$ jako jedyną liczbę λ taką, że $v - \lambda v_1 \in V_0$; wtedy $\phi \in V^*$ oraz $\phi(v_1) = 1$.

[45] Ścisłe zdefiniowanie pojęcia 'kanoniczności' jest możliwe w języku tzw. teorii kategorii.

[46] Używany jest też angielskojęzyczny termin *pullback*, a także termin F-*transport formy*.

w ten sposób operator $F^T : W^* \to V^*$ jest liniowy; nazywamy go([47]) *transpozycją operatora* F lub *operatorem transponowanym względem* F.

Zatem z definicji mamy $\forall \psi \in W^* : \ \forall v \in V : \ (F^T(\psi))(v) = \psi(F(v))$, co oczywiście wygląda przejrzyściej w notacji 'braketowej':

$$\langle F^T(\psi), v \rangle = \langle \psi, F(v) \rangle. \qquad \text{(T)}$$

336. **Fakt**. Operacja transponowania operatora ma następujące własności:

(1) jest liniowa, tzn. $(\lambda_1 F_1 + \lambda_2 F_s)^T = \lambda_1 F_1^T + \lambda_2 F_2^T$;

(2) jest 'antymultiplikatywna': $(F \circ G)^T = G^T \circ F^T$;

zauważmy, że jeśli $U \xrightarrow{G} V \xrightarrow{F} W$, to $W^* \xrightarrow{F^T} V^* \xrightarrow{G^T} U^*$, więc tylko taka 'odwrócona' kolejność transpozycji wchodzi w grę.

(3) jest inwolutywna, tzn. $(F^T)^T = F$;

w tym wzorze, zgodnie z twierdzeniem 333, utożsamimy $(V^*)^*$ z V;

(4) Jeśli $F = w \otimes \phi$, gdzie $w \in W$, $\phi \in V^*$([48]), to $F^T = \phi \otimes w$.

(5) Macierz $\left[F^T\right]^{e^*}_{f^*}$ jest transpozycją macierzy $[F]^f_e$,

jeśli e^* jest bazą sprzężoną względem e, a f^* — bazą sprzężoną względem f.

(1) jest oczywista. (2) $(F \circ G)^T(\psi) = \psi \circ (F \circ G) = (\psi \circ F) \circ G = G^T\left(F^T(\psi)\right)$. (3) gdyż wzór (T) można przepisać w postaci $\langle F(v), \psi \rangle = \left\langle v, F^T(\psi) \right\rangle$, jeśli wektory v i $F(v)$ traktujemy zgodnie z 333 jako formy liniowe na przestrzeniach V^* i W^*. (4) gdyż $\left\langle F^T(\psi), v \right\rangle = \left\langle \psi, w \langle \phi, v \rangle \right\rangle = \langle \psi, w \rangle \cdot \langle \phi, v \rangle$, co oznacza, że $F^T(\psi) = \langle \psi, w \rangle \, \phi$ wskutek dowolności v. (5) gdyż jeśli $F = \sum\limits_{i,j} F^i{}_j \, f_i \otimes e^j$, to $F^T = \sum\limits_{i,j} F^i{}_j \, e^j \otimes f_i$.

Dalsze własności wprowadzonych tu pojęc warto formułować i rozważać w trochę ogólniejszej formie, w której dwie wzajemnie sprzeżone przestrzenie są od początku traktowane równoprawnie. Korzyścią takiego podejścia jest m.in. to, że uzyskane pojęcia, własności i fakty odnoszą się nie tylko do przestrzeni sprzężonych, ale i do 'przestrzeni z iloczynem skalarnym' lub np. 'przestrzeni symplektycznych'.

5.7 Pary dwoiste

337. **Definicja**. Niech U, V będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$, zaś $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{K}$, $(u, v) \mapsto \langle u, v \rangle \in \mathbb{K}$ — odwzorowaniem dwuliniowym (tzn. takim, że liczba $\langle u, v \rangle$ zależy liniowo i od u, i od v). Trójkę([49]) $(U, V, \langle \cdot, \cdot \rangle)$ nazywamy *parą dwoistą* (lub *dualną*), a odwzorowanie $\langle \cdot, \cdot \rangle$ — *dwoistością* lub *dualnością* między przestrzeniami U i V, jeśli $\langle \cdot, \cdot \rangle$ ma następujące 'własności rozdzielania':

[47] Nieco starsza, lecz wciąż spotykana jest nazwa *sprzężenie* F lub *operator sprzężony względem* F; jedynym jej mankamantem jest to, że termin *operator sprzężony* jest często używany w innym (choć pokrewnym) znaczeniu, w teorii przestrzeni z iloczynem skalarnym (np. unitarnych lub przestrzeni Hilberta). Termin 'sprzężenie' łączy się z symbolem F^*.

[48] Tzn. $F(v) = w \langle \phi, v \rangle$, a ϕ i w są tu ustalone; tej (i tylko tej) postaci F ma rząd $\leqslant 1$.

[49] Czyli parę przestrzeni wyposażoną w odwzorowanie dwuliniowe.

$$1^\circ \quad \forall\, 0 \neq u \in U : \exists\, v \in V : \langle u, v\rangle \neq 0;$$
$$2^\circ \quad \forall\, 0 \neq v \in V : \exists\, u \in U : \langle u, v\rangle \neq 0.$$

338. **Fakt**. Trójka $(V^*, V, \langle\,\cdot\,,\,\cdot\,\rangle)$, gdzie $\langle \phi, v\rangle := \phi(v)$, jest parą dwoistą. Odwrotnie, jeśli $(U, V, \langle\,\cdot\,,\,\cdot\,\rangle)$ jest parą dwoistą, a przestrzenie U, V są skończenie wymiarowe, to odwzorowanie

$$\boxed{U \ni u \mapsto \widehat{u} \in V^*, \qquad \widehat{u}(v) := \langle u, v\rangle,}$$

jest (kanonicznym) izomorfizmem $U \xrightarrow{\cong} V^*$, czyli «każda para dwoista wygląda tak samo, jak para V^*, V'», w szczególności $\dim U = \dim V$.

$u \mapsto \widehat{u}$ jest (wprost z własności 1°) liniową injekcją $U \to V^*$, jeśli więc stosując fakt 280 wystarczy pokazać, że $\dim U = \dim V^*$. Otóż z istnienia injekcji $U \to V^*$ mamy $\dim U \leqslant \dim V^*$, zaś dzięki symetrii sytuacji mamy też przeciwną nierówność.

Uwaga. Odwzorowania $U \ni u \mapsto \widehat{u} = \langle u, \cdot\,\rangle \in V^*$ oraz $V \ni v \mapsto \widehat{v} = \langle\,\cdot\,, v\rangle \in U^*$ w przypadku nieskończonego wymiaru są injektywne, lecz nie muszą (a nawet oba naraz nie mogą, jak pokazujemy w Appandiksie) być surjektywne, zob. poniżej (D).

339. **Przykłady**. (A) Dla przestrzeni $U = \mathbb{K}^m_{\ n}$, $V := \mathbb{K}^n_{\ m}$ dwuliniowe odwzorowanie $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \langle \boldsymbol{u}, \boldsymbol{v}\rangle := \mathrm{tr}(\boldsymbol{uv})$ jest dwoistością: własności rozdzielania 1°, 2° łatwo widać ze wzoru $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} u^i_{\ j}\, v^j_{\ i}$, gdzie $u^i_{\ j}$ i $v^j_{\ i}$ są wyrazami macierzy $\boldsymbol{u}, \boldsymbol{v}$.

W szczególności dla $m = 1$ dostajemy dwoistość między przestrzenią wektorów wierszowych $U = \mathbb{K}_n$ i wektorów kolumnowych $V = \mathbb{K}^n$; jest to właśnie dwoistość

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \mathrm{tr}(\boldsymbol{uv}) = \boldsymbol{uv} = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix} \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix} = u_1 v^1 + \ldots + u_n v^n,$$

której używamy, traktując wektory wierszowe jako formy liniowe na przestrzeni $\mathbb{K}^n$.

(B) Jeśli X, Y są przestrzeniami wektorowymi, $U = \mathrm{L}(X; Y)$, $V = \mathrm{L}(Y; X)$, to wzór $\langle u, v\rangle := \mathrm{tr}(u \circ v) = \mathrm{tr}(v \circ u)$ określa dwoistość pomiędzy U i V (zauważmy, że oba te złożenia są endomorfizmami, a więc oba ślady są dobrze określone). Warunek 1° można sprawdzić np. korzystając ze wzoru $\langle u, x \otimes \psi\rangle = \langle \psi, u(x)\rangle$([50]); można także skorzystać z tego, że $\mathrm{tr}(uv) = \mathrm{tr}\big([uv]^e_{\ e}\big) = \mathrm{tr}(\boldsymbol{uv})$, gdzie $\boldsymbol{u} = [u]^e_{\ f}$, $\boldsymbol{v} = [v]^f_{\ e}$.

(C) Niebawem poznamy bardzo ważny przykład, w którym $U = V$, $\dim V < \infty$, a $\langle u, v\rangle = b(u, v)$, gdzie $b : V \times V \to \mathbb{K}$ jest pewną formą dwuliniową *nieosobliwą* na V. Taka forma najczęściej jest symetryczna (wtedy nazywa się ją *iloczynem skalarnym na* V) lub antysymetryczna (wtedy jest to *forma symplektyczna na* V). Najprostszy przykład: $V = \mathbb{K}^n$, zaś $\langle \boldsymbol{x}, \boldsymbol{y}\rangle := \sum\limits_{i=1}^{n} x_i y_i = \begin{pmatrix} \text{standardowy iloczyn} \\ \text{skalarny na } \mathbb{K}^n \end{pmatrix}$.

(D) Niech $U = V = \{\boldsymbol{x} \in \mathbb{K}^{\mathbb{N}} : \ \forall^* i : \ x_i = 0\}$ (nieskończone ciągi $\boldsymbol{x} = (x_1, x_2, \ldots)$, mające *prawie wszystkie* wyrazy równe zero). Oczywiście wzór $\langle \boldsymbol{x}, \boldsymbol{y}\rangle := \sum\limits_{i=1}^{\infty} x_i y_i$ jest sensowny (liczba $\neq 0$ składników jest skończona), a określona nim forma jest dwoistością. W tym przykładzie odwzorowanie $U \ni u \mapsto \widehat{u} \in V^*$ jest niesurjektywne, np. forma liniowa $\phi \in V^*$, dana wzorem $\phi(\boldsymbol{x}) := \sum\limits_{i=1}^{\infty} x_i$ (sensowna definicja:

[50] Jeśli operator $u \in U$ jest $\neq 0$, to $\exists x \in X : \ y := u(x) \neq 0$, a więc $\exists \psi \in Y^* : \langle \psi, y\rangle \neq 0$.

jest tylko skończona liczba niezerowych składników) nie jest postaci $\phi(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.

(E) Niech $U = V = \mathbb{K}_n[\cdot]$ (wielomiany stopnia $\leqslant n$); dopuszczamy tutaj $n = \infty$. Łatwo zauważyć, że jest dwoistością wielkość $\langle u, v \rangle = b(u,v) := \sum_k a_k b_k$, gdzie a_k, b_k są współczynnikami u, v, tzn. $u(t) = \sum_k a_k t^k$, $v(t) = \sum_k b_k t^k$. Ciekawsza jest inna dwoistość: $\langle u, v \rangle = \hat{b}(u,v) := \left[u\left(\frac{d}{dt}\right) v(t) \right]\big|_{t=0} = \sum_k \frac{a_k b_k}{k!}$; w tym wzorze $u\left(\frac{d}{dt}\right)$ jest 'operatorem różniczkowym' $\sum_k a_k \frac{d^k}{dt^k}$, którym działamy na wielomian $v(t)$.

Od tego miejsca założymy, że przestrzenie U i V są skończenie wymiarowe.

340. **Bazy dwoiste**. Para baz: $e = (e_1, \ldots, e_n)$ dla V oraz $e^* = (e^1, \ldots, e^n)$ dla U, nazywamy (wzajemnie) *sprzężona* (lub *dualna* lub *dwoista*), jeśli

$$\boxed{\forall i,j \in \overline{1,n} : \ \langle e^i, e_j \rangle = \delta^i{}_j}; \quad \text{wtedy} \quad \boxed{\begin{array}{ll} \forall u \in U : & u = \sum_i \langle u, e_i \rangle \, e^i \\ \forall v \in V : & v = \sum_i e_i \, \langle e^i, v \rangle \end{array}}.$$

W tym sensie elementy bazy e^* określają liniowe współrzędne na przestrzeni V, zaś elementy bazy e — liniowe współrzędne na U.

341. **Fakt**. Relacja sprzężenia baz jest odpowiedniością bijektywną $e \leftrightarrow e^*$.

Określmy $\phi^i(v) :=$ (i-ta współrzędna v w bazie e), wtedy $\phi^i \in V^*$; skoro odwzorowanie $U \ni u \mapsto \widehat{u} \in V^*$ jest bijektywne, to istnieją $e^i \in U$, takie że $\widehat{e}^i = \phi^i$, tzn. $\langle e^i, v \rangle = \phi^i(v)$, w szczególności $\langle e^i, e_j \rangle = \phi^i(e_j) = \delta^i{}_j$. Zatem baza sprzężona względem e istnieje. Gdyby oprócz $\langle e^i, e_j \rangle = \delta^i{}_j$ było także $\langle f^i, e_j \rangle = \delta^i{}_j$, wtedy $\forall j : \ \langle f^i - e^i, e_j \rangle = 0$, skąd $\forall v : \ \langle f^i - e^i, v \rangle = 0$, a to z warunku rozdzielania oznacza, że $f^i - e^i = 0$; wobec tego baza sprzężona jest określona jednoznacznie.

342. **Ćwiczenie**. Czasem przydatne jest ogólniejsze pojęcie *układów wzajemnie sprzężonych*, tzn. układów wektorów $e^1, \ldots, e^r \in U$ i $e_1, \ldots, e_r \in V$, spełniających relacje $\langle e^i, e_j \rangle = \delta^i{}_j$. Wykazać, że: (1) wtedy oba układy są liniowo niezależne oraz (2) zachodzą wzory $\forall u \in U_0 : \ u = \sum_i \langle u, e_i \rangle \, e^i$, $\forall v \in V_0 : \ v = \sum_i e_i \, \langle e^i, v \rangle$, gdzie $U_0 \subset U$, $V_0 \subset V$ są podprzestrzeniami rozpiętymi przez oba układy; (3) każdy układ liniowo niezależny w V ma układ sprzężony; dla $r < \dim U = \dim V$ nie ma jednoznaczności: do e^i można dodać 'poprawki' u^i, takie że $\forall v \in V_0 : \ \langle u^i, v \rangle = 0$.

343. **Przykład**. Dla 'modelowej' pary dwoistej $U = \mathbb{K}_n$, $V = \mathbb{K}^n$, gdzie $\boxed{\langle \boldsymbol{u}, \boldsymbol{v} \rangle := \boldsymbol{u}\boldsymbol{v} = \sum_i u_i v^i}$, mamy następujące pożyteczne spostrzeżenie:

Ćwiczenie!

Jeśli $\boldsymbol{X} \in \mathbb{K}^n{}_n$ jest macierzą odwracalną, to bazę dwoistą względem bazy utworzonej z kolumn macierzy $\boldsymbol{X}$ tworzą wiersze macierzy $\boldsymbol{X}^{-1}$.

344. **Transformacje baz**: $\boxed{f_j = \sum_i e_i A^i{}_j \iff f^i = \sum_j B^i{}_j e^j}$,

gdzie obie 'macierze przejścia', $\boldsymbol{A} = [A^i{}_j]$ oraz $\boldsymbol{B} = [B^i{}_j] \in \mathbb{K}^n{}_n$, są wzajemnie odwrotne:

$$\delta^i{}_j = \langle f^i, f_j \rangle = \langle \sum_r B^i{}_r e^r, \sum_s e^s A^s{}_j \rangle = \sum_r \sum_s \ldots = \sum_r B^i{}_r A^r{}_j.$$

345. **Fakt**. Niech (U_1, V_1) oraz (U_2, V_2) będą dwiema parami dwoistymi, a $F \in \mathrm{L}(V_1; V_2)$. Wtedy istnieje dokładnie jeden operator $F^T \in \mathrm{L}(U_2; U_1)$,

zwany *transpozycją* F lub *operatorem transponowanym wzgl.* F, taki że

$$\boxed{\forall u_2 \in U_2 : \forall v_1 \in V_1 : \left\langle F^T(u_2), v_1 \right\rangle_1 = \langle u_2, F(v_1) \rangle_2}. \qquad (*)$$

Ćwiczenie. Dla dla $U_i = V_i^*$ warunek (*) oznacza, że $\boxed{\forall \phi \in V_2^* : F^T(\phi) = \phi \circ F}$.

346. **Fakt** (*podstawowe własności operacji transponowania operatora*):

(1) $F \mapsto F^T$ jest liniowe, (2) $(F \circ G)^T = G^T \circ F^T$, (3) $F^{TT} = F$;
(4) Macierz $[F^T]^{e^*}_{f^*}$ jest transpozycją macierzy $[F]^f_e$,

jeśli e, e^*, f i f^* są (stosownie sprzężonymi) bazami kolejno V_1, U_1, V_2 i U_2.

Sprawdzenie tych własności pozostawiamy jako proste ćwiczenie.

347. **Definicja** (*anihilator podzbioru*). Jeśli $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{K}$ jest dwoistością, to *anihilatorem* podzbioru $W \subset V$ nazywamy podzbiór $W^0 \subset U$, określony wzorem $W^0 := \{u \in U : \forall w \in w : \langle u, w \rangle = 0\}$. Oczywiście W^0 jest zawsze podprzestrzenią U. Analogicznie definiujemy anihilator podzbioru $Z \subset U$:

$$Z^0 := \{v \in V : \forall z \in Z : \langle z, v \rangle = 0\} \subset V.$$

348. **Fakt** (własności anihilatora podprzestrzeni). Niech (U, V) będzie parą dwoistą, $n := \dim U = \dim V$ oraz $W \subset V$ (lub $W \subset U$) — podprzestrzenią; wtedy

(1) $\dim W^0 = n - \dim W$; (2) $W^{00} = W$;
(3) $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$; (4) $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

Ad (1): Jeśli $e_1, \ldots, e_r$ jest bazą W, zaś $e_1, \ldots, e_r, \ldots, e_n$ — bazą V, to oczywiście $e^{r+1}, \ldots, e^n$ jest bazą W^0. *Ad* (2): Wprost z definicji $W \subset W^{00}$, zaś $\dim W = \dim W^{00}$. *Ad* (3): Oczywiste. *Ad* (4): Istnieją $\tilde{W}_i \subset U$, takie, że $W_i = \tilde{W}^0$; korzystając z (3),(2) dostajemy $(W_1 \cap W_2)^0 = (\tilde{W}_1^0 \cap \tilde{W}_2^0)^0 = (\tilde{W}_1 + \tilde{W}_2)^{00} = W_1^0 + W_2^0$.

349. **Fakt**. Przy założeniach i oznaczeniach z punktu 345 zachodzą równości:

$$(1) \ \boxed{\ker F^T = (\operatorname{im} F)^0}, \qquad (2) \ \boxed{\operatorname{im} F^T = (\ker F)^0}.$$

Ad (1): $u_2 \in \ker F^T \Leftrightarrow \forall v_1 : \left(\left\langle F^T(u_2), v_1 \right\rangle = 0, \text{ tzn. } \langle u_2, F(v_1) \rangle = 0\right) \Leftrightarrow$
$\Leftrightarrow u_2 \in \{F(v_1) : v_1 \in V_1\}^0 \Leftrightarrow u_2 \in (\operatorname{im} F)^0$.

Ad (2): $(\operatorname{im} F^T)^0 = \ker F^{TT} = \ker F$ dzięki (1) i własności sprzężenia; stąd teza.

Appendix 0: Przestrzeń sprzężona w przypadku nieskończonego wymiaru

Pokażemy teraz, że założenie $\dim V < \infty$ jest nie tylko wystarczające, lecz także *konieczne*, by przestrzenie V^* oraz V^{**} były izomorficzne z V. Wobec tego dla $\dim V = \infty$ kanoniczne odwzorowanie $V \to V^{**}$ jest injektywne, a niesurjektywne.

* **Fakt**. Jeśli przestrzeń V ma nieskończony wymiar, to $\dim V < \dim V^*$.

Niech $X \subset V$ będzie bazą V; oznaczmy $x := |X| = \dim V$, $k := |\mathbf{K}|$. Pokażemy najpierw, że $\boxed{|V| = kx, \ |V^*| = k^x.}$ Każdy $v \in V$ jest skończoną kombinacją liniową elementów $x \in X$, więc $V = V_1 \cup V_2 \cup V_3 \cup \ldots$, gdzie V_n jest zbiorem tych $v \in V$, które

są kombinacją liniową $\leqslant n$ elementów bazy. Dzięki surjekcji $S_n : \mathbf{K}^n \times X^n \to V_n$, $S_n(\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_n) := \sum_{i=1}^{n} \lambda_i x_i$, mamy $|V_n| \leqslant k^n x^n = kx$; skoro $kx \geqslant |\mathbf{N}|$, mamy stąd $|V| = |\bigcup_n V_n| \leqslant |\mathbf{N}| kx = kx$, zarazem $|V| \geqslant |V_1| = kx$, a więc $|V| = kx$. Z kolei ponieważ każda forma $\phi : V \to \mathbf{K}$ jest jednoznacznie określona przez swe wartości na bazie, tzn. przez obcięcie $\phi|_X : X \to \mathbf{K}$, to $V^* \cong \mathbf{K}^X$, więc $|V^*| = k^x$.

Jeśli $x \geqslant k$, to $kx = x < 2^x \leqslant k^x$, czyli $|V| < |V^*|$; tym bardziej $\dim V < \dim V^*$.

Jeśli $x < k$, to może być $kx = k^x$, a więc $|V| = |V^*|$, np. jeśli k jest postaci $k = a^x$, to $kx = k = a^x = a^{x^2} = \left(a^x\right)^x = k^x$. Pokażemy jednak, że w tym przypadku także $\dim V < \dim V^*$. Wybierzmy przeliczlny podzbiór $X_0 := \{x_0, x_1, x_2, \ldots\} \subset X$ i określmy dla $a \in \mathbf{K}$ element $\phi_a \in V^*$ wzorem $\phi_a(x) := \begin{cases} 0, & x \in X \setminus \in X_0, \\ a^j, & x = x_j. \end{cases}$ Sprawdzimy, że ϕ_a dla $a \in \mathbf{K}$ są liniowo niezależne, a więc $\dim V^* \geqslant k > x = \dim V$. Niech elementy $a_1, \ldots, a_n \in \mathbf{K}$ będą parami różne; pokażemy, że jeśli dla pewnych $\lambda_1, \ldots, \lambda_n \in \mathbf{K}$ zachodzi równość $\lambda_1 \phi_{a_1} + \ldots + \lambda_n \phi_{a_n} = 0$, to $\lambda_1 = \ldots = \lambda_n = 0$. Istnieje wielomian $L_1(t)$ taki, że $L_1(a_1) \neq 0$, lecz $L_1(a_2) = \ldots = L_1(a_n) = 0$; można np. wziąć $L_1(t) := (t - a_2) \ldots (t - a_n)$. Niech $L_1(t) = c_0 + c_1 t + \ldots + c_N t^N$; wtedy $0 = \sum_{j=0}^{N} c_j \Big(\sum_{i=1}^{n} \lambda_i \phi_{a_i}(x_j) \Big) = \sum_{j=0}^{N} c_j \Big(\sum_{i=1}^{n} \lambda_i {a_i}^j \Big) = \sum_{i=1}^{n} \lambda_i \Big(\sum_{j=0}^{N} c_j {a_i}^j \Big) = \sum_{i=1}^{n} \lambda_i L_1(a_i) =$ $= \lambda_1 L_1(a_1)$, czyli $\lambda_1 = 0$. Analogicznie pokazujemy, że $\lambda_2 = 0, \ldots, \lambda_n = 0$.

5.8 Ślad macierzy i endomorfizmu

350. *Śladem* macierzy kwadratowej $\boldsymbol{A} \in \mathbb{K}^n{}_n$ nazywamy sumę wszystkich jej wyrazów diagonalnych $A^1{}_1, \ldots, A^n{}_n$; oznaczamy go symbolem $\operatorname{tr} \boldsymbol{A}$ ([51]):

$$\operatorname{tr} \boldsymbol{A} = A^1{}_1 + \ldots + A^n{}_n.$$

351. **Fakt**. Ślad jest liniową funkcją $\operatorname{tr} : \mathbb{K}^n{}_n \to \mathbb{K}$ o następującej własności:

$$\boxed{\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})}\,([52]),$$

gdy $\boldsymbol{A} \in \mathbb{K}^m{}_n$, $\boldsymbol{B} \in \mathbb{K}^n{}_m$, tzn. gdy określone są iloczyny $\boldsymbol{AB}$ i $\boldsymbol{BA}$.

Macierze $\boldsymbol{AB} \in \mathbb{K}^m{}_m$ i $\boldsymbol{BA} \in \mathbb{K}^n{}_n$ są kwadratowe, a z definicji śladu i mnożenia macierzowego oraz z łączności i przemienności dodawania liczb mamy: $\operatorname{tr}(\boldsymbol{AB}) = \sum_i (\boldsymbol{AB})^i{}_i = \sum_i (\sum_j A^i{}_j B^j{}_i) = \sum_{i,j} A^i{}_j B^j{}_i = \sum_j (\sum_i B^j{}_i A^i{}_j) = \sum_j (\boldsymbol{BA})^j{}_j = \operatorname{tr}(\boldsymbol{BA})$.

352. **Wniosek**. Jeśli $F \in \operatorname{End} V := \mathrm{L}(V; V)$, to ślad macierzy $[F]^e{}_e$ nie zależy od wyboru bazy e w przestrzeni V, co powoduje, że poniższa definicja jest sensowna.

Jeśli f jest drugą bazą V, to ze wzoru 306 mamy $[F]^f{}_f = \boldsymbol{AB}$, gdzie $\boldsymbol{A} := [\mathrm{id}]^f{}_e [F]^e{}_e$, $\boldsymbol{B} := [\mathrm{id}]^e{}_f$, więc $\boldsymbol{BA} = \left([\mathrm{id}]^e{}_f [\mathrm{id}]^f{}_e \right) [F]^e{}_e = [F]^e{}_e$; stąd $\operatorname{tr}\left([F]^f{}_f\right) = \operatorname{tr}\left([F]^e{}_e\right)$.

353. **Definicja**. *Ślad* endomorfizmu $F \in \operatorname{End} V$ jest to liczba dana wzorem

[51] Z ang. *trace* 'ślad, trop'; spotyka się także symbole $\operatorname{Tr} \boldsymbol{A}$ oraz $\operatorname{Sp} \boldsymbol{A}$ (z niem. *Spuhr*).

[52] Wynika stąd od razu, że *cykliczne* przestawianie czynników nie zmienia śladu, lecz np. $\operatorname{tr}(\boldsymbol{ABC}) = \operatorname{tr}(\boldsymbol{BCA}) = \operatorname{tr}(\boldsymbol{CAB})$ na ogół jest $\neq \operatorname{tr}(\boldsymbol{ACB}) = \operatorname{tr}(\boldsymbol{BAC}) = \operatorname{tr}(\boldsymbol{CBA})$.

$$\operatorname{tr} F := \operatorname{tr}([F]^e_e) = \left(\begin{array}{c} \text{suma wyrazów diagonalnych macierzy } [F]^e_e \\ \text{gdzie } e \text{ jest dowolną bazą } V \end{array} \right).$$

Ponieważ i-ta współrzędna $w \in V$ w bazie e wyraża się wzorem $\langle e^i, w\rangle$, gdzie $e^* = (e^1, \ldots, e^n)$ jest bazą sprzężoną względem $e = (e_1, \ldots, e_n)$, to $[F]^e_e$ ma wyrazy $F^i_{\ j} = \langle e^i, F(e_j)\rangle$; wobec tego otrzymujemy wzór

$$\boxed{\operatorname{tr} F = \sum_{i=1}^{n} \langle e^i, F(e_i)\rangle}.$$

354. **Fakt**. (1) Ślad jest funkcjonałem liniowym na przestrzeni $\operatorname{End} V$.
(2) Dla dowolnych operatorów $F \in \mathrm{L}(V;W)$ i $G \in \mathrm{L}(W;V)$ mamy:

$$\operatorname{tr}(F \circ G) = \operatorname{tr}(G \circ F).$$

(3) Wobec tego jeśli $A \in \mathrm{L}(V;W)$ jest liniową bijekcją, a $F \in \operatorname{End} V$, to

$$\operatorname{tr}(A \circ F \circ A^{-1}) = \operatorname{tr} F.$$

Jest to oczywistą konsekwencją własności śladu macierzy i macierzy odwzorowania.

355. **Uwagi.**

1. Liczba $S := \operatorname{tr}([F]^f_e) = \sum_{i=1}^{n} \langle f^i, F(e_i)\rangle$, gdzie e i f są dwiema bazami V, zależy nie tylko od operatora F, ale także od obu baz, a więc ze śladem F nie ma nic wspólnego; co więcej, jedyną relacją między F a liczbą S jest implikacja $F = 0 \Rightarrow S = 0$.

2. Nie da się uogólnić pojęcia śladu na inne odwzorowania liniowe. Nietrudno też pokazać, że dla ustalonego $F \in \mathrm{L}(V;W)$ przez stosowny wybór baz można jako $[F]^f_e$ uzyskać każdą macierz (wymiaru $\dim W \times \dim V$), której rząd jest równy $\operatorname{rk} F$.

3. Oprócz endomorfizmów istnieją inne *obiekty geometryczne związane z* V, które w bazie V można scharakteryzować macierzą kwadratową; przykładem może być odwzorowanie dwuliniowych $\Phi : V \times V \to \mathbb{K}$. Otóż ślad macierzy takiego obiektu (np. macierzy $[\Phi(e_i, e_j)]$ dla formy dwuliniowej) zależy od wyboru bazy, a zatem nie stanowi żadnej istotnej charakterystyki takiego obiektu.

4. Ślad $\operatorname{tr} : \operatorname{End} V \to \mathbb{K}$ stanowi nowy przykład *kanonicznego* (lub *naturalnego*) odwzorowania liniowego, zob. punkt 334.

6 Odwzorowania wieloliniowe i (anty)symetryczne

6.1 Wieloliniowość, symetria i antysymetria

356. **Definicja**. Niech $r \in \mathbb{N}$, zaś $V_1, \ldots, V_r$ oraz W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$. Odwzorowanie $f : V_1 \times \ldots \times V_r \to W$ nazywamy *r-liniowym*, jeśli jest liniowe względem każdego ze swoich argumentów (przy każdym zestawie pozostałych argumentów), tzn. jeśli

$$f(v_1, \ldots, \lambda' v_j' + \lambda'' v_j'', \ldots, v_r) = \\ = \lambda' f(v_1, \ldots, v_j', \ldots, v_r) + \lambda'' f(v_1, \ldots, v_j'', \ldots, v_r)$$

dla każdych $j \in \overline{1,r}$, $\lambda', \lambda'' \in \mathbb{K}$, $v_1 \in V_1, \ldots, v_j', v_j'' \in V_j, \ldots, v_r \in V_r$.

357. **Przykłady**. Dla $\phi_1 \in V_1^*, \ldots, \phi_r \in V_r^*$ funkcja $f : V_1 \times \ldots \times V_r \to \mathbb{K}$,

$$f(v_1, \ldots, v_r) := \phi_1(v_1) \cdot \ldots \cdot \phi_r(v_r),$$

jest r-liniowa; oznacza się ją zwykle symbolem $\phi_1 \otimes \ldots \otimes \phi_r$.

Odwzorowanie 'ewaluacji' $f : \mathrm{L}(V;W) \times V \to W$, $f(F, v) := F(v)$, jest dwuliniowe: $F(v)$ zależy liniowo zarówno od v, jak i od F.

Odwzorowanie $f : V^* \times V \to \mathbb{K}$, $f(\phi, v) := \langle \phi, v \rangle$, jest 2-liniowe. Funkcja $g : W^* \times \mathrm{L}(V;W) \times V \to \mathbb{K}$, $g(\phi, F, v) := \langle \phi, F(v) \rangle$, jest trójliniowa: liczba $\langle \phi, F(v) \rangle$ zależy liniowo od ϕ, od F i od v.

Macierz $\boldsymbol{ABCD}$ oraz liczba $\mathrm{tr}(\boldsymbol{ABCD})$ są czteroliniowymi funkcjami macierzy $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$; mamy więc, na przykład, czteroliniowe odwzorowania $\mathbb{K}^3{}_5 \times \mathbb{K}^5{}_4 \times \mathbb{K}^4{}_7 \times \mathbb{K}^7{}_3$ o wartościach w $\mathbb{K}^3{}_3$ i w $\mathbb{K}$.

Operacja składania operatorów liniowych jest wieloliniowa, tzn. np.

$$\mathrm{L}(V_3;V_4) \times \mathrm{L}(V_2;V_3) \times \mathrm{L}(V_1;V_2) \to \mathrm{L}(V_1;V_4), \quad (F, G, H) \mapsto F \circ G \circ H,$$

jest odwzorowaniem trójliniowym.

358. **Definicja**. Symbolem $\mathrm{L}(V_1, \ldots, V_r; W)$ oznaczmy zbiór wszystkich r-liniowych odwzorowań $V_1 \times \ldots \times V_r \to W$; jest on przestrzenią wektorową względem 'zwykłych', tj. punktowych działań dodawania i mnożenia przez liczby. W dalszym ciągu zajmiemy się przede wszystkim przypadkiem $V_1 = \ldots = V_r$, więc wprowadźmy jeszcze oznaczenia

$$\mathrm{L}_r(V;W) := \mathrm{L}(\underbrace{V, \ldots, V}_{r}; W), \quad \mathrm{L}_r(V) := \mathrm{L}_r(V;\mathbb{K}) = \mathrm{L}(\underbrace{V, \ldots, V}_{r}; \mathbb{K});$$

elementy przestrzeni $\mathrm{L}_r(V)$, tzn. r-liniowe funkcje $V \times \ldots \times V \to \mathbb{K}$, nazywa się *formami r-liniowymi na V*.

359. Niech $1 \leqslant i < j \leqslant r$; mówimy, że odwzorowanie $f : V \times \ldots \times V \to W$ jest *antysymetryczne względem i-tego oraz j-tego argumentu*, jeśli

$$\forall\ v_1, \ldots, v_r \in V : f(\ldots, v_j, \ldots, v_i, \ldots) = -f(\ldots, v_i, \ldots, v_j, \ldots).$$

360. **Fakt**. Dla $f : V \times \ldots \times V \to W$ następujące warunki są równoważne:

1. Dla każdej permutacji $\sigma \in S_r$ zachodzi równość:

$$\forall\ v_1, \ldots, v_r \in V: \quad f(v_{\sigma(1)}, \ldots, v_{\sigma(r)}) = \operatorname{sgn}(\sigma) \cdot f(v_1, \ldots, v_r);\ (*)$$

2. $\forall\ 1 \leqslant i < j \leqslant r:$ f jest antysym. wzgl. argumentów o n.rach $i,\ j$;
3. $\forall\ 1 \leqslant i < r:$ f jest antysym. wzgl. argumentów o n.rach $i,\ i+1$.
4. $\forall\ 1 < j \leqslant r:$ f jest antysym. wzgl. argumentów o numerach $1,\ j$.

Biorąc $\sigma := (i\ \ j)$ dostajemy 1.⇒2.3.4. Dowód odwrotnych implikacji zaczniemy od sprawdzenia, że zbiór $\Sigma_f := \{\sigma \in S_r : (*)\}$ jest zamknięty względem składania: jeśli $\varrho, \sigma \in \Sigma_f$, to opócz (*) mamy $f(w_{\varrho(1)}, \ldots, w_{\varrho(r)}) = \operatorname{sgn}(\varrho) \cdot f(w_1, \ldots, w_r)$ dla $w_1, \ldots, w_r \in V$; biorąc $w_i := v_{\sigma(i)}$ dostajemy $f(v_{\sigma\circ\varrho(1)}, \ldots, v_{\sigma\circ\varrho(r)}) = \operatorname{sgn}(\varrho) \cdot \operatorname{sgn}(\sigma) \cdot f(v_1, \ldots, v_r)$, tzn. $\sigma \circ \varrho \in \Sigma_f$, gdyż sgn jest homomorfizmem grup. Implikacja 3.⇒1. wynika teraz z tego, że transpozycje $(1\ 2)$, $(2\ 3), \ldots, (r-1\ r)$ generują grupę S_r oraz że — na mocy 3. — każda z nich należy do Σ_f. Analogicznie 2.⇒1. i 4.⇒1.

361. **Definicja**. Odwzorowanie $f : V \times \ldots \times V \to W$ nazywa się *(całkowicie) antysymetryczne*, jeśli spełnia powyższe równoważne warunki.

362. **Ćwiczenie**. Sformułować i dowieść analogiczny fakt charakteryzujący odwzorowania *(całkowicie) symetryczne*.

363. **Fakt**. Jeśli odwzorowanie f jest r-liniowe, tzn. $f \in \mathrm{L}_r(V; W)$, to dla każdej (ustalonej) pary $1 \leqslant i < j \leqslant r$ równoważne są następ. warunki:

(a) f jest antysymetryczne względem i-tego oraz j-tego argumentu;
(b) $\forall\ v_1, \ldots, v_r \in V: \ \Big(v_i = v_j \ \Rightarrow\ f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_r) = 0\Big)$.

Pozwala to dla odwzorowań wieloliniowych przeformułować warunki antysymetrii z faktu 360., np. warunkowi 2. możemy nadać postać

$$f(v_1, \ldots, v_r) = 0, \textit{ ilekroć dwa spośród wektorów } v_1, \ldots, v_r \textit{ są równe.}$$

(a)⇒(b) jest oczywiste. (b)⇒(a): Dla uproszczenia zapisu weźmy $i = 1,\ j = r = 2$. Z 2-liniowości mamy $f(v_1+v_2, v_1+v_2) = f(v_1, v_1)+f(v_1, v_2)+f(v_2, v_1)+f(v_2, v_2)$, co po uwzględnieniu własności (b) daje równość $0 = f(v_1, v_2) + f(v_2, v_1)$.

364. **Twierdzenie**. Niech $f \in \mathrm{L}_r(V; W)$; wtedy

$$\Big(f \text{ jest antysymetryczne }\Big) \Leftrightarrow \left(\begin{array}{c} f \text{ zeruje się na każdym liniowo} \\ \text{zależnym układzie } r \text{ wektorów z } V \end{array}\right).$$

Zatem jeśli $f \in \mathrm{L}_r(V; W)$ jest antysymetryczne i $r > \dim V$, to $f = 0$.

$\boxed{\Rightarrow}$ Gdy $v_1, \ldots, v_r$ są liniowo zależne, wtedy jeden z wektorów układu, np. v_r, jest kombinacją liniową pozostałych, więc z liniowości względem r-tego argumentu

$$f(v_1, \ldots, v_r) = f(v_1, \ldots, v_{r-1}, \sum_{j=1}^{r-1} v_j \lambda^j) = \sum_{j=1}^{r-1} f(v_1, \ldots, v_j, \ldots, v_{r-1}, v_j)\lambda^j = 0.$$

$\boxed{\Leftarrow}$ Układ $v_1, v_1, v_2, \ldots, v_r$ jest liniowo zależny, więc $f(v_1, v_1, \ldots, v_r) = 0$, to zaś oznacza antysymetrię f wzgl. 1. i 2. argumentu itd.; stąd i z faktu 360 wynika teza.

365. **Wniosek**. Jeśli $f \in \mathrm{L}_n(V; W)$ jest odwzorowaniem antysymetrycznym, a wektory $u_i, v_i \in V$ są takie, że każdy z układów $u_1, u_2, \ldots, u_k, v_k$, gdzie $k \in \overline{1, n-1}$, jest liniowo zależny, np. $v_k \in \langle u_1, \ldots, u_k\rangle$, to

$$f(u_1, u_2, \ldots, u_n) = f(u_1, v_1 + u_2, v_2 + u_3, \ldots, v_{n-1} + u_n).$$

Istotnie, wartość prawej strony nie zmieni się, gdy pominiemy w niej składnik

- v_1, gdyż liniowa zależność $u_1, v_1, \ldots$ sprawia, że $f(u_1, v_1, v_2 + u_3, \ldots) = 0$;
- v_2, gdyż lin. zależność $u_1, u_2, v_2, \ldots$ sprawia, że $f(u_1, u_2, v_2, v_3 + u_4, \ldots) = 0$;
- v_3, gdyż lin. zależność $u_1, u_2, u_3, v_3, \ldots$ sprawia, że $\ldots$

366. **Definicja**. Niech $\bigwedge^r(V^*) := \{f \in \mathrm{L}_r(V) : \ f \text{ jest antysymetryczne}\}$; jest to podprzestrzeń $\mathrm{L}_r(V)$, złożona z antysymetrycznych form r-liniowych na V; jej elementy nazywają się *formami zewnętrznymi stopnia* r (albo, krócej, *r-formami*) *na* V; $\bigwedge^r(V^*)$ nosi nazwę *r-tej potęgi zewnętrznej przestrzeni* V^*; oczywiście $\bigwedge^1(V^*) = V^*$; z 364 wiemy, że

$$\boxed{\bigwedge^r(V^*) = \{0\} \ \text{ dla } \ r > \dim V}.$$

367. **Przykład**. Jeśli $\varphi_1, \varphi_2 \in V^*$, to wzór $f(v_1, v_2) := \varphi_1(v_1)\varphi_2(v_2) - \varphi_1(v_2)\varphi_2(v_1)$ określa pewną 2-formę $f \in \bigwedge^2(V^*)$ (oznaczaną zwykle symbolem $\varphi_1 \wedge \varphi_2$ i zwaną *iloczynem zewnętrznym* φ_1 *i* φ_2). Zauważmy, że jeśli φ_1, φ_2 są l. niezależne, to $f \neq 0$, bowiem V ma bazę $e_1, \ldots, e_n$, taką że $\varphi_1 = e^1$, $\varphi_2 = e^2$, gdzie $e^i(e_j) = \delta^i_j$ (baza dualna), mamy więc $f(e_1, e_2) = 1$. Odwrotnie, $f = 0$ dla l.zależnych form φ_1, φ_2.

368. **Definicja**. *Formą objętości* na przestrzeni wektorowej V nazywa się niezerowy element $f \in \bigwedge^n(V^*)$, gdzie $n := \dim V$.

W przypadku $\mathbb{K} = \mathbb{R}$ wartość $f(v_1, \ldots, v_n) \in \mathbb{R}$ nazywa się *objętością zorientowaną* (a jej moduł — *objętością niezorientowaną*) równoległościanu 'rozpostartego' na $v_1, \ldots, v_n$, tzn. zbioru $[0,1]v_1 + \ldots + [0,1]v_n = \{t_1v_1 + \ldots + t_nv_n : \ t_i \in [0,1]\} \subset V$.

369. **Fakt**. Niech $n = \dim V$ oraz $e_1, \ldots, e_n$ będzie bazą V. Wtedy każda n-forma $f \in \bigwedge^n(V^*)$ jest jednoznacznie określona przez swoją wartość $f(e_1, \ldots, e_n) \in \mathbb{K}$. Wobec tego: (a) przestrzeń $\bigwedge^n(V^*)$ jest 1-wymiarowa; (b) jeśli f i $\tilde{f}$ są dwiema formami objętości na V, to $\exists \lambda \in \mathbb{K}^* : \tilde{f} = \lambda \cdot f$.

Rozłóżmy każdy z wektorów $v_1, \ldots, v_n$ w bazie e: $v_j = \sum_{i=1}^{n} e_i x^i_j$, wtedy dzięki n-liniowości

$$f(v_1, \ldots, v_n) = f\Big(\sum_{i_1=1}^{n} e_{i_1} x_1^{i_1}, \ldots, \sum_{i_n=1}^{n} e_{i_n} x_n^{i_n}\Big) = \sum_{i_1, \ldots, i_n = 1}^{n} f(e_{i_1}, \ldots, e_{i_n}) \cdot x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}.$$

Wskutek antysymetrii liczba $f(e_{i_1}, \ldots, e_{i_n})$ jest $= 0$, gdy ciąg wskaźników $i_1, \ldots, i_n$ nie jest różnowartościowy, więc wkład do sumy mogą wnosić tylko ciągi postaci $\sigma(1), \ldots, \sigma(n)$, gdzie $\sigma \in S_n$; ponadto $f(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \mathrm{sgn}(\sigma) f(e_1, \ldots, e_n)$, więc

$$\boxed{f(v_1, \ldots, v_n) = \alpha \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x_1^{\sigma(1)} \cdot \ldots \cdot x_n^{\sigma(n)}}, \qquad (*)$$

gdzie $\alpha := f(e_1, \ldots, e_n) \in \mathbb{K}$. Dowiedliśmy tym samym, że istnieje co najwyżej jedna n-forma f o danej wartości $\alpha := f(e_1, \ldots, e_n) \in \mathbb{K}$; pozostaje teraz pokazać, funkcja f zdefiniowana wzorem (*) jest n-formą. Sprawdźmy liniowość względem (przykładowo) 1. argumentu: $v = \lambda' v'_1 + \lambda'' v''_1$ oznacza, że $x^i_1 = \lambda' x'^i_1 + \lambda'' x''^i_1$, więc

$$f(\lambda' v'_1 + \lambda'' v''_1, v_2, \ldots, v_n) = \alpha \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\big(\lambda' x'^{\sigma(1)}_1 + \lambda'' x''^{\sigma(1)}_1\big) \cdot \ldots \cdot x_n^{\sigma(n)} =$$

$$= \lambda' \alpha \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x'^{\sigma(1)}_1 \cdot \ldots \cdot x_n^{\sigma(n)} + \lambda'' \alpha \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x''^{\sigma(1)}_1 \cdot \ldots \cdot x_n^{\sigma(n)} =$$

$$= \lambda' f(v'_1, v_2, \ldots, v_n) + \lambda'' f(v''_1, v_2, \ldots, v_n).$$

Antysymetria: zastępując w (*) wektory v_j wektorami $w_j = v_{\varrho(j)} = \sum_{i=1}^{n} e_i x^i_{\varrho(j)}$, gdzie $\varrho \in S_n$, musimy w miejsce współrzędnych x^i_j wstawić $y^i_j = x^i_{\varrho(j)}$, zatem

$$f(v_{\varrho(1)},\ldots,v_{\varrho(n)}) = \alpha \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) x_{\varrho(1)}^{\sigma(1)} \cdot \ldots \cdot x_{\varrho(n)}^{\sigma(n)}.$$

Suma nie zmieni się, gdy 'wskaźnik' σ przemianujemy na $\sigma\circ\varrho$, co wraz z σ przebiega całą grupę S_n; otóż iloczyn $x_{\varrho(1)}^{\sigma\circ\varrho(1)} \cdot \ldots \cdot x_{\varrho(n)}^{\sigma\circ\varrho(n)}$ rożni się tylko kolejnością czynników od $x_1^{\sigma(1)} \cdot \ldots \cdot x_n^{\sigma(n)}$, zaś $\operatorname{sgn}(\sigma\circ\varrho) = \operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\varrho)$; tak więc faktycznie mamy

$$f(v_{\varrho(1)},\ldots,v_{\varrho(n)}) = \operatorname{sgn}(\varrho)\cdot f(v_1,\ldots,v_n).$$

370. **Uwaga**. W taki sam sposób można otrzymać ogólniejszy rezultat: każda r-forma $f \in \bigwedge^r(V^*)$, $r \in \overline{1,n}$, jest całkowicie określona([53]) przez swe wartości $f(e_{i_1},\ldots,e_{i_r})$ dla wszystkich możliwych rosnących ciągów $i_1,\ldots,i_r$ indeksów z $\overline{1,n}$. Ponieważ takich ciągów jest tyle, co podzbiorów ('kombinacji') r-elementowych zbioru n-elementowego, to

$$\boxed{\dim \textstyle\bigwedge^r(V^*) = \binom{n}{r},\ n := \dim V}.$$

6.2 Wyznacznik macierzy

371. Z ostatniego faktu wynika, że dla $n \in \mathbb{N}$ na przestrzeni $\mathbb{K}^n$ istnieje dokładnie jedna forma objętości, przyjmująca na bazie standardowej $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ przestrzeni $\mathbb{K}^n$ wartość równą 1; formę tę nazywa się *wyznacznikiem stopnia n*.

W takim razie wykazaliśmy następujące

Twierdzenie o Istnieniu i Jednoznaczności Wyznacznika:
Istnienie: Istnieje funkcja $\operatorname{Det}_n : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$, mająca trzy następujące własności:

(1) n-liniowość; (2) antysymetria;

(3) $\operatorname{Det}_n\left(\begin{bmatrix}1\\ \vdots\\ 0\end{bmatrix},\ldots,\begin{bmatrix}0\\ \vdots\\ 1\end{bmatrix}\right) = \operatorname{Det}_n(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n) = 1$ (unormowanie).

Jednoznaczność: Każde $f : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$, mające własności (1) i (2), ma postać $f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) = f(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)\cdot\operatorname{Det}_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$, czyli

$$f = \alpha\cdot\operatorname{Det}_n, \text{ gdzie } \alpha = f(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n).$$

372. **Definicja**. *Wyznacznikiem* macierzy kwadratowej $\boldsymbol{A} \in \mathbb{K}^n{}_n$ nazywamy liczbę $\det_n \boldsymbol{A} := \operatorname{Det}_n(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_n)$, gdzie $\boldsymbol{A}_i \in \mathbb{K}^n$ są kolumnami $\boldsymbol{A}$. Zatem $\det_n : \mathbb{K}^n{}_n \to \mathbb{K}$ jest jedyną funkcją, która:

(I) jest n-liniową antysymetryczną funkcją kolumn;
(II) przyjmuje wartość 1 na macierzy jednostkowej.

TIJW mówi, że taka funkcja $\det_n$ istnieje, oraz że

[53] Mamy mianowicie wzór $f(v_1,\ldots,v_r) = \sum_{(i_1,\ldots,i_r)\in I_r^n} f(e_{i_1},\ldots,e_{i_r}) X^{i_1,\ldots,i_r}$, stanowiący uogólnienie (*); I_r^n jest $\binom{n}{r}$-elementowym zbiorem wszystkich rosnących ciągów $(i_1,\ldots,i_r)$ o wyrazach z $\overline{1,n}$, natomiast $X^{i_1,\ldots,i_r} = \sum_{\sigma\in S_r} \operatorname{sgn}(\sigma) x_1^{i_{\sigma(1)}} \cdot \ldots \cdot x_r^{i_{\sigma(r)}} = \det\left[x_q^{i_p}\right]_{p,q\in\overline{1,r}}$.

każde odwzorowanie $f : \mathbb{K}^n{}_n \to \mathbb{K}$, *spełniające warunek* (I), *jest postaci*

$$\boxed{f(\boldsymbol{A}) = f(\boldsymbol{I}_n) \cdot \det \boldsymbol{A}}.$$

373. Odnotujmy, że zgodnie ze wzorem (*) z punktu 369 zachodzi równość([54])

$$\boxed{\det \boldsymbol{A} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, A^{\sigma(1)}{}_1 \cdot \ldots \cdot A^{\sigma(n)}{}_n}.$$

374. Wprawdzie wzór powyższy odgrywa doniosłą rolę, lecz rosnąca lawinowo wraz z n liczba $n!$ składników sumy sprawia, że bezpośrednie jego stosowanie do obliczania wyznaczników ma sens tylko dla $n \leqslant 3$. Biorąc $n = 2$ i $n = 3$ otrzymujemy z niego:

$$\det \begin{bmatrix} A^1{}_1 & A^1{}_2 \\ A^2{}_1 & A^2{}_2 \end{bmatrix} = A^1{}_1 A^2{}_2 - A^2{}_1 A^1{}_2;$$

$$\det \begin{bmatrix} A^1{}_1 & A^1{}_2 & A^1{}_3 \\ A^2{}_1 & A^2{}_2 & A^2{}_3 \\ A^3{}_1 & A^3{}_2 & A^3{}_3 \end{bmatrix} = A^1{}_1 A^2{}_2 A^3{}_3 + A^2{}_1 A^3{}_2 A^1{}_3 + A^3{}_1 A^1{}_2 A^2{}_3 + \\ - A^3{}_1 A^2{}_2 A^1{}_3 - A^1{}_1 A^3{}_2 A^2{}_3 - A^2{}_1 A^1{}_2 A^3{}_3.$$

Stosowanie ostatniego wzoru ułatwia mnemotechniczna *reguła Sarrusa*; polega ona na zapamiętaniu tych zestawów wyrazów macierzy, których iloczyny opatrujemy znakiem $+$, oraz tych zestawów, których iloczyny opatrujemy znakiem $-$:

$$\underbrace{\begin{bmatrix} \bullet & & \\ & \bullet & \\ & & \bullet \end{bmatrix}, \begin{bmatrix} & \bullet & \\ \bullet & & \\ & \bullet & \end{bmatrix}, \begin{bmatrix} & & \bullet \\ & \bullet & \\ \bullet & & \end{bmatrix}}_{\text{ze znakiem } \boxed{+}}, \underbrace{\begin{bmatrix} & \bullet & \\ \bullet & & \\ & & \bullet \end{bmatrix}, \begin{bmatrix} \bullet & & \\ & & \bullet \\ & \bullet & \end{bmatrix}, \begin{bmatrix} & & \bullet \\ \bullet & & \\ & \bullet & \end{bmatrix}}_{\text{ze znakiem } \boxed{-}}.$$

Łatwo to zapamiętać, zauważając że znaki '$+$' przypisujemy iloczynom trójek, mającym 'bok' równoległy do $\diagdown$, a znaki '$-$' — mającym 'bok' równoległy do $\diagup$.

375. **Uwaga**. Liczbę $\det \boldsymbol{A}$ oznacza się często symbolem $\begin{vmatrix} A^1{}_1 & A^1{}_2 & \cdots & A^1{}_n \\ \vdots & \vdots & \ldots & \vdots \\ A^n{}_1 & A^n{}_2 & \cdots & A^n{}_n \end{vmatrix}$. Pamiętajmy więc, że np. $\begin{bmatrix} 2 & 3 \\ -7 & 6 \end{bmatrix}$ jest macierzą, a więc 'zestawem liczb' (niektórzy piszą to jako $\begin{pmatrix} 2 & 3 \\ -7 & 6 \end{pmatrix}$), zaś $\begin{vmatrix} 2 & 3 \\ -7 & 6 \end{vmatrix}$ — wyznacznikiem tej macierzy, a więc liczbą.

6.3 Podstawowe własności wyznacznika

376. W definicji wyznacznika traktowaliśmy macierz jako zestaw kolumn; można jednak też postrzegać macierz jako zestaw wierszy, a w konsekwencji oprócz '*wyznacznika kolumnowego*' zdefiniować analogiczny '*wyznacznik wierszowy*' macierzy. Okazuje się jednak, że nie ma takiej potrzeby: oba wyznaczniki, 'wierszowy' i 'kolumnowy', są równe. Aby się o tym przekonać wystarczy pokazać, że nasz 'kolumnowy' wyznacznik jest antysymetryczną funkcją wieloliniową nie tylko kolumn, ale i wierszy; to zaś wynika w oczywisty sposób z następującego faktu:

377. **Fakt**. Transpozycja nie zmienia wyznacznika, tzn. $\boxed{\det(\boldsymbol{A}^T) = \det \boldsymbol{A}}$.

[54] Od tego miejsca będziemy zazwyczaj pisać krótko '$\det \boldsymbol{A}$' zamiast '$\det_n \boldsymbol{A}$'.

Zamiana $\boldsymbol{A}$ na $\boldsymbol{A}^T$ we wzorze 373 daje $\det(\boldsymbol{A}^T) = \sum\limits_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, A^1{}_{\sigma(1)} \cdot \ldots \cdot A^n{}_{\sigma(n)}$. Ustawiając czynniki w kolejności numerów kolumn dostajemy $A^1{}_{\sigma(1)} \cdot \ldots \cdot A^n{}_{\sigma(n)} = A^{\sigma^{-1}(1)}{}_1 \cdot \ldots \cdot A^{\sigma^{-1}(n)}{}_n$; zarazem $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, więc po przemianowaniu σ^{-1} na ϱ, a $\sum\limits_{\sigma}$ na $\sum\limits_{\varrho}$ (co możemy zrobić dzięki temu, że przyporządkowanie $\sigma \mapsto \sigma^{-1}$ jest bijekcją $S_n \xrightarrow{\varrho} S_n$; mówiąc obrazowo: wraz z σ także σ^{-1} przebiega jednokrotnie cały zbiór S_n) dostajemy $\det(\boldsymbol{A}^T) = \sum\limits_{\varrho\in S_n} \operatorname{sgn}(\varrho)\, A^{\varrho(1)}{}_1 \cdot \ldots \cdot A^{\varrho(n)}{}_n = \det \boldsymbol{A}$.

378. **Wniosek**([55]). Dla $f : \mathbb{K}^n{}_n \to \mathbb{K}$ następujące warunki są równoważne:

- f jest antysymetryczną funkcją wieloliniową kolumn macierzy;
- f jest antysymetryczną funkcją wieloliniową wierszy macierzy;
- f jest postaci $f(\boldsymbol{A}) = c \cdot \det \boldsymbol{A}$, gdzie (oczywiście) $c = f(\boldsymbol{I}_n)$.

Pojęcie wyznacznika jest ogromnie ważne głównie dlatego, że zachodzi następujący

379. **Fakt** (*wzór Cauchy'ego*). Wyznacznik $\det : \mathbb{K}^n{}_n \to \mathbb{K}$ jest funkcją multiplikatywną, tzn. $\boxed{\det(\boldsymbol{AB}) = \det \boldsymbol{A} \cdot \det \boldsymbol{B}}$ dla $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{K}^n{}_n$.

Kolumnami macierzy $\boldsymbol{AB}$ są wektory $\boldsymbol{AB}_1, \ldots, \boldsymbol{AB}_n$, więc $f_{\boldsymbol{A}}(\boldsymbol{B}) := \det(\boldsymbol{AB})$ przy ustalonej macierzy $\boldsymbol{A}$ jest antysymetryczną funkcją n-liniową kolumn $\boldsymbol{B}$; stąd

$$f_{\boldsymbol{A}}(\boldsymbol{B}) = f_{\boldsymbol{A}}(\boldsymbol{I}_n) \cdot \det \boldsymbol{B} = \det \boldsymbol{A} \cdot \det \boldsymbol{B}.$$

380. **Fakt**. Wyznacznik macierzy górnotrójkątnej (lub dolnotrójkątnej) jest iloczynem jej wyrazów diagonalnych([56]).

Zastosujmy wzór 373. Z założenia spełniony jest warunek $A^i{}_j \neq 0 \Rightarrow i \leqslant j$, więc $A^{\sigma(1)}{}_1 \ldots A^{\sigma(n)}{}_n \neq 0$ implikuje $\sigma(1) \leqslant 1, \sigma(2) \leqslant 2, \ldots, \sigma(n) \leqslant n$, co dla permutacji oznacza, że $\sigma(1) = 1, \ldots, \sigma(n) = n$, tj. $\sigma = \mathrm{id}$; zatem suma, jaką jest $\det \boldsymbol{A}$, ma co najwyżej jeden niezerowy składnik, mianowicie $A^1{}_1 \ldots A^n{}_n$, QED.

Inny dowód (bardziej bezpośredni, nie wymagający wzoru 373). Zastosujemy 365. Macierz górnotrójkątna ma kolumny postaci $\boldsymbol{A}_k = \boldsymbol{v}_{k-1} + \boldsymbol{u}_k$, gdzie $\boldsymbol{u}_k := A^k{}_k \boldsymbol{e}_k$, $\boldsymbol{v}_0 := 0$, zaś $\boldsymbol{v}_j \in V_j := \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_j \rangle$; zatem każdy z układów $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_k$ jest liniowo zależny ($k+1$ wektorów w k-wymiarowej przestrzeni V_k). Stąd teza.

381. **Fakt**. $$\det \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{0} & \boldsymbol{D} \end{array}\right] = \det \boldsymbol{A} \cdot \det \boldsymbol{D} = \det \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{0} \\ \hline \boldsymbol{C} & \boldsymbol{D} \end{array}\right]$$

dla $\boldsymbol{A} \in \mathbb{K}^p{}_p$, $\boldsymbol{B} \in \mathbb{K}^p{}_q$, $\boldsymbol{C} \in \mathbb{K}^q{}_p$, $\boldsymbol{D} \in \mathbb{K}^q{}_q$, czyli *wyznacznik macierzy «blokowo trójkątnej» jest iloczynem wyznaczników bloków diagonalnych.* Oczywiście przez indukcję możemy ten rezultat uogólnić na dowolną liczbę bloków, np. $\det \left[\begin{array}{c|c|c} \boldsymbol{A} & \boldsymbol{B} & \boldsymbol{C} \\ \hline \boldsymbol{0} & \boldsymbol{D} & \boldsymbol{E} \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{F} \end{array}\right] = \det \boldsymbol{A} \cdot \det \boldsymbol{D} \cdot \det \boldsymbol{F}$.

Z własności wyznacznika jest jasne, że $f(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}) := \det \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{0} & \boldsymbol{D} \end{array}\right]$ jest antysymetryczną p-liniową funkcją kolumn $\boldsymbol{A}$ (przy ustalonych $\boldsymbol{B}, \boldsymbol{D}$), więc z 378. mamy $f(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}) = f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{D}) \cdot \det \boldsymbol{A}$; tak samo $f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{D})$ jest antysymetryczną

[55] Szalenie ważny.

[56] Macierz kwadratowa $\boldsymbol{A} = [\, A^i{}_j \,] \in \mathbb{K}^n{}_n$ nazywa się *górnotrójkątna*, jeśli są zerami wszystkie jej wyrazy leżące pod główną przekątną, tzn. jeśli $A^i{}_j = 0$ dla $i > j$; $\boldsymbol{A}$ jest *ściśle trójkąna*, gdy ponadto $\forall i : A^i{}_i = 0$. Analogicznie definiujemy *macierze dolnotrójkątne*.

q-liniową funkcją wierszy $\boldsymbol{D}$, więc $f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{D}) = f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{I}_q) \cdot \det \boldsymbol{D}$. Wreszcie $f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{I}_q) =$ (wyznacznik macierzy trójkątnej o jedynkach na diagonali) $= 1$. ([57])

382. **Definicja**. Dla $\boldsymbol{A} = [A^i{}_j] \in \mathbb{K}^n{}_n$ niech $\boldsymbol{A}^{\neq i}{}_{\neq j}$ oznacza podmacierz $\boldsymbol{A}$, uzyskaną przez wykreślenie kolumny i wiersza zawierających wyraz $A^i{}_j$. Liczbę $\boxed{(-1)^{i+j} \det \boldsymbol{A}^{\neq i}{}_{\neq j}}$ nazywa się *dopełnieniem algebraicznym* wyrazu $A^i{}_j$; będziemy ją oznaczać symbolem $\tilde{A}^j{}_i$. Zwróćmy uwagę na notację: $\tilde{A}^1{}_2$ jest dopełnieniem algebraicznym wyrazu $A^2{}_1$, a nie $A^1{}_2$.

Niech $\tilde{\boldsymbol{A}} = [\tilde{A}^j{}_i]$ będzie macierzą dopełnień algebraicznych.

383. **Przykłady.** $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \tilde{\boldsymbol{A}} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$,

$$\boldsymbol{A} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \Rightarrow \tilde{\boldsymbol{A}} = \begin{bmatrix} b_2c_3 - c_2b_3 & c_2a_3 - a_2c_3 & a_2b_3 - b_2a_3 \\ b_3c_1 - c_3b_1 & c_3a_1 - a_3c_1 & a_3b_1 - b_3a_1 \\ b_1c_2 - c_1b_2 & c_1a_2 - a_1c_2 & a_1b_2 - b_1a_2 \end{bmatrix}.$$

384. **Twierdzenie**. Zachodzi wzór $\boxed{\boldsymbol{A}\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}\boldsymbol{A} = (\det \boldsymbol{A})\boldsymbol{I}_n}$, co oznacza, że

$$\forall i, k \in \overline{1,n}: \ \sum_{j=1}^{n} A^i{}_j \tilde{A}^j{}_k = (\det \boldsymbol{A})\delta^i{}_k, \quad \sum_{j=1}^{n} \tilde{A}^i{}_j A^j{}_k = (\det \boldsymbol{A})\delta^i{}_k.$$

Dla ustalonej pary i, k oznaczmy $f(\boldsymbol{A}) := \sum_{j=1}^{n} A^i{}_j \tilde{A}^j{}_k$. Skoro wielkość $A^i{}_j$ jest liniową funkcją kolumny $\boldsymbol{A}_j$, zaś wielkość $\tilde{A}^j{}_k$ — wieloliniową funkcją pozostałych (wszystkich, prócz j-tej) kolumn $\boldsymbol{A}$, to składnik $A^i{}_j \tilde{A}^j_k$, a więc także suma $f(\boldsymbol{A})$, jest wieloliniową funkcją kolumn $\boldsymbol{A}$. Sprawdzimy, że $f(\boldsymbol{A})$ jest antysymetryczną funkcją kolumn: jeśli $\exists j : \boldsymbol{A}_j = \boldsymbol{A}_{j+1}$, to $\tilde{A}^j{}_k = -\tilde{A}^{j+1}_k$ (gdyż $\boldsymbol{A}^{\neq k}_{\neq j} = \boldsymbol{A}^{\neq k}_{\neq j+1}$, zaś $(-1)^{j+k} = -(-1)^{j+1+k}$), natomiast $\tilde{A}^l{}_k = 0$ dla $l \in \overline{1,n} \setminus \overline{j, j+1}$ (bo macierz $\boldsymbol{A}^{\neq k}_{\neq l}$ ma dwie jednakowe kolumny); zatem $(\exists j : \boldsymbol{A}_j = \boldsymbol{A}_{j+1}) \Rightarrow f(\boldsymbol{A}) = 0$, co, jak wiemy, jest równoważne antysymetrii $f(\boldsymbol{A}) = f(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n)$.

Wobec tego $f(\boldsymbol{A}) = f(\boldsymbol{I}_n) \cdot \det \boldsymbol{A}$; pozostaje zauważyć, że $f(\boldsymbol{I}_n) = \delta^i{}_k$, gdyż dla $\boldsymbol{A} = \boldsymbol{I}_n$ mamy $A^i{}_j = \delta^i{}_j$ oraz $\tilde{A}^j{}_k = \delta^j{}_k$. Drugi wzór wyprowadza się analogicznie.

385. **Wniosek**. $\forall i \leqslant n: \boxed{\sum_{j=1}^{n} A^i{}_j \tilde{A}^j{}_i = \det \boldsymbol{A}}$, $\quad \forall j \leqslant n: \boxed{\sum_{i=1}^{n} A^i{}_j \tilde{A}^j{}_i = \det \boldsymbol{A}}$.

Wzory te nazywa się *rozwinięciami Laplace'a*; dają one rozwinięcia wyznacznika względem i-tego wiersza (1. wzór) oraz względem j-tej kolumny (2. wzór).

386. **Przykład**. Dla macierzy $\boldsymbol{A} = \begin{bmatrix} 2 & 11 & 4 & 9 \\ 16 & 7 & 5 & 14 \\ 3 & 12 & 10 & 1 \\ 8 & 6 & 13 & 15 \end{bmatrix}$ rozwinięcia $\det \boldsymbol{A}$ względem drugiej kolumny oraz względem trzeciego wiersza wyglądają następująco:

$$\det \boldsymbol{A} = (-11)\begin{vmatrix} 16 & 5 & 14 \\ 3 & 10 & 1 \\ 8 & 13 & 15 \end{vmatrix} + 7\begin{vmatrix} 2 & 4 & 9 \\ 3 & 10 & 1 \\ 8 & 13 & 15 \end{vmatrix} - 12\begin{vmatrix} 2 & 4 & 9 \\ 16 & 5 & 14 \\ 8 & 13 & 15 \end{vmatrix} + 6\begin{vmatrix} 2 & 4 & 9 \\ 16 & 5 & 14 \\ 3 & 10 & 1 \end{vmatrix};$$

[57] Wzór $f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{I}_q) = 1$ można wykazać bez tw. o wyznaczniku macierzy trójkątnej: $\forall \lambda : \lambda^p f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{I}_q) \underset{1.}{=} f(\lambda \boldsymbol{I}_p, \lambda \boldsymbol{B}, \boldsymbol{I}_q) \underset{2.}{=} \lambda^p f(\boldsymbol{I}_p, \lambda \boldsymbol{B}, \boldsymbol{I}_q)$ dzięki jednorodności wyznacznika względem wierszy (1.) i kolumn (2.). Stąd tożsamość $f(\boldsymbol{I}_p, \boldsymbol{B}, \boldsymbol{I}_q) = f(\boldsymbol{I}_p, \lambda \boldsymbol{B}, \boldsymbol{I}_q)$; musi ona zachodzić nawet dla $\lambda = 0$, gdyż zależność $f(\boldsymbol{I}_p, \lambda \boldsymbol{B}, \boldsymbol{I}_q)$ od λ jest wielomianowa.

$$\det \boldsymbol{A} = 3\begin{vmatrix} 11 & 4 & 9 \\ 7 & 5 & 14 \\ 6 & 13 & 15 \end{vmatrix} - 12\begin{vmatrix} 2 & 4 & 9 \\ 16 & 5 & 14 \\ 8 & 13 & 15 \end{vmatrix} + 10\begin{vmatrix} 2 & 11 & 9 \\ 16 & 7 & 14 \\ 8 & 6 & 15 \end{vmatrix} - 1\begin{vmatrix} 2 & 11 & 4 \\ 16 & 7 & 5 \\ 8 & 6 & 13 \end{vmatrix}.$$

387. **Przykład**. Mając dane $n \geqslant 2$ liczb $x_1, \ldots, x_n \in \mathbb{K}$ utwórzmy następujący wyznacznik stopnia n, nazywany *wyznacznikiem Vandermonde'a*:

$$V(x_1, \ldots, x_n) := \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ {x_1}^2 & {x_2}^2 & \ldots & {x_n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ {x_1}^{n-1} & {x_2}^{n-1} & \ldots & {x_n}^{n-1} \end{vmatrix}.$$

Na przykład dla $n = 2$ i $n = 3$ mamy: $V(a,b) = \begin{vmatrix} 1 & 1 \\ a & b \end{vmatrix} = b - a$,

$$V(a,b,c) = \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = bc^2 + ca^2 + ab^2 - b^2c - c^2a - a^2b = (b-a)(c-a)(c-b).$$

Fakt. Wyznacznik Vandermonde'a wyraża się następującym wzorem:

$$V(x_1, \ldots, x_n) = \prod_{i<j} (x_j - x_i).$$

W konsekwencji $V(x_1, \ldots, x_n) \neq 0 \iff x_1, \ldots, x_n$ są parami różne([58]).

Niech $f(x) := V(x_1, \ldots, x_{n-1}, x)$; traktując $x_1, \ldots, x_{n-1}$ jako parametry rozwińmy wyznacznik $f(x)$ względem ostatniej kolumny; widać wtedy, że $f(x)$ jest wielomianem stopnia $\leqslant n-1$, przy czym współczynnikiem przy x^{n-1} jest $V(x_1, \ldots, x_{n-1})$. Ponadto $f(x_1) = 0, \ldots, f(x_{n-1}) = 0$ (wyznacznik macierzy o dwóch jednakowych kolumnach), więc $f(x) = C(x - x_1)\ldots(x - x_{n-1})$, gdzie $C = V(x_1, \ldots, x_{n-1})$. Dla zakończenia dowodu wystarczy teraz zastosować prostą indukcję względem n.

Poznamy teraz dalsze przykłady pożytków z twierdzenia 384:

6.4 Macierze nieosobliwe, wzory Cramera

388. **Twierdzenie**. Jeśli $\boldsymbol{A} \in \mathbb{K}^n{}_n$, to następujące warunki są równoważne:

(1) $\det \boldsymbol{A} \neq 0$;
(2) $\boldsymbol{A}$ jest odwracalna (czyli $\exists \boldsymbol{B} : \boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n$);
(3) $\boldsymbol{A}$ ma lewą lub prawą odwrotność: $\exists \boldsymbol{B} : \boldsymbol{AB} = \boldsymbol{I}_n$ lub $\boldsymbol{BA} = \boldsymbol{I}_n$;
(4) $\forall \boldsymbol{b} \in \mathbb{K}^n$: układ $\boldsymbol{Ax} = \boldsymbol{b}$ ma dokładnie jedno rozwiązanie $\boldsymbol{x}$;
(5) $\exists \boldsymbol{b} \in \mathbb{K}^n$: układ $\boldsymbol{Ax} = \boldsymbol{b}$ ma dokładnie jedno rozwiązanie $\boldsymbol{x}$;
(6) $\ker \boldsymbol{A} = 0$ (czyli dla $\boldsymbol{b} = 0$ układ ma tylko jedno rozwiązanie);
(7) $\operatorname{rk} \boldsymbol{A} = n$ (czyli dla każdego $\boldsymbol{b} \in \mathbb{K}^n$ układ ma rozwiązanie).

(1) $\Rightarrow$ (2), gdyż na mocy twierdzenia 384 przy $D = \det \boldsymbol{A} \neq 0$ macierz $\boldsymbol{B} := \frac{1}{D}\tilde{\boldsymbol{A}}$ ma własności $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n$, tzn. jest odwrotnością $\boldsymbol{A}$. (2) $\Rightarrow$ (3) oczywiste. (3) $\Rightarrow$ (1), gdyż $\boldsymbol{AB} = \boldsymbol{I}_n \Rightarrow \det \boldsymbol{A} \det \boldsymbol{B} = 1$ (ze wzoru Cauchy'ego) $\Rightarrow \det \boldsymbol{A} \neq 0$. (4) $\Rightarrow$ (5) oczywiste. (5) $\Rightarrow$ (6), gdyż jeśli $\boldsymbol{u} \in \ker \boldsymbol{A}$ i $\boldsymbol{Ax} = \boldsymbol{b}$, to $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{u}) = \boldsymbol{b}$,

[58] Wynikanie '$\Rightarrow$' jest widoczne wprost z definicji: jeśli $x_i = x_j$, $i \neq j$, to $V(x_1, \ldots, x_n)$ jest wyznacznikiem macierzy, która ma dwie kolumny równe (i-tą i j-tą), więc jest zerem.

skąd (jednoznaczność) $\boldsymbol{x}+\boldsymbol{u}=\boldsymbol{x}$, tj. $\boldsymbol{u}=0$. (6) $\Leftrightarrow$ (7), gdyż $\dim\ker\boldsymbol{A}+\operatorname{rk}\boldsymbol{A}=n$. (7) $\Leftrightarrow$ (4): $\operatorname{rk}\boldsymbol{A}=n \iff \dim\langle\boldsymbol{A}_1,\ldots,\boldsymbol{A}_n\rangle=n \iff \langle\boldsymbol{A}_1,\ldots,\boldsymbol{A}_n\rangle=\mathbb{K}^n \iff$ kolumny $\boldsymbol{A}$ tworzą bazę $\mathbb{K}^n$, zob. 233(c); z kolei wobec wzoru $\boldsymbol{Ax}=\sum\boldsymbol{A}_i x^i$ mamy $\boldsymbol{Ax}=\boldsymbol{b} \iff \boldsymbol{b}$ jest kombinacją liniową kolumn $\boldsymbol{A}$ o współczynnikach x^i.

(4) $\Rightarrow$ (2) Niech $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ będzie standardową (zerojedynkową) bazą $\mathbb{K}^n$, zaś dla $i\in\overline{1,n}$ niech $\boldsymbol{x}_i$ będzie rozwiązaniem układu $\boldsymbol{Ax}_i=\boldsymbol{e}_i$. Wtedy $\boldsymbol{A}[\boldsymbol{x}_1|\ldots|\boldsymbol{x}_n]=[\boldsymbol{e}_1|\ldots|\boldsymbol{e}_n]=\boldsymbol{I}_n$, tzn. mamy $\boldsymbol{AX}=\boldsymbol{I}_n$, skąd $\det\boldsymbol{A}\,\det\boldsymbol{X}=1$, a więc $\det\boldsymbol{A}\neq 0$. (2) $\Rightarrow$ (4), gdyż wtedy $\boldsymbol{Ax}=\boldsymbol{b} \iff \boldsymbol{A}^{-1}(\boldsymbol{Ax})=\boldsymbol{A}^{-1}\boldsymbol{b} \iff \boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$.

389. *Ćwiczenie.* Dowieść, że jeśli $\boldsymbol{A},\boldsymbol{B}\in\mathbb{K}^n{}_n$ są takie, że $\boldsymbol{AB}=\boldsymbol{I}_n$, to także $\boldsymbol{BA}=\boldsymbol{I}_n$.

390. Macierzą *nieosobliwą* nazywamy macierz kwadratową, spełniającą równoważne warunki (1)..(7); jak widzieliśmy w '(1) $\Rightarrow$ (2)' mamy wzór

$$\boldsymbol{A}^{-1}=\frac{1}{\det\boldsymbol{A}}\tilde{\boldsymbol{A}}\,.$$

Jest to ważny i pożyteczny wzór, lecz jeśli chodzi o odwracanie macierzy *numerycznych*, jego przydatność ogranicz się do $n\leqslant 3$; warto np. stosować wzór

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}^{-1}=\tfrac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}.$$

391. **Definicja**. Układ równań liniowych $\begin{cases}A^1_1x^1+\ldots+A^1{}_nx^n=b^1\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ A^m_1x^1+\ldots+A^m_nx^n=b^m\end{cases}$ tzn. $\boldsymbol{Ax}=\boldsymbol{b}$, nazywa się *cramerowski*, jeśli $m=n$ (tzn. niewiadomych jest tyle, co równań) oraz macierz $\boldsymbol{A}$ jest nieosobliwa, tzn. $\det\boldsymbol{A}\neq 0$.

392. **Fakt**. Jeśli układ równań $\boldsymbol{Ax}=\boldsymbol{b}$ (gdzie $\boldsymbol{A}\in\mathbb{K}^n{}_n$ oraz $\boldsymbol{b}\in\mathbb{K}^n$) jest cramerowski, to: (a) ma dokładnie jedno rozwiązanie $\boldsymbol{x}=\begin{bmatrix}x^1\\\cdots\\x^n\end{bmatrix}$;

(b) rozwiązanie tego układu można otrzymać z tzw. *wzorów Cramera*:

$$\boxed{\begin{aligned}x^j&=\frac{D_j}{D},\quad \text{gdzie}\quad D:=\det\boldsymbol{A},\\ D_j&=\begin{pmatrix}\text{wyznacznik macierzy, otrzymanej z }\boldsymbol{A}\\ \text{przez zastąpienie kolumny }\boldsymbol{A}_j\text{ wektorem }\boldsymbol{b}\end{pmatrix}.\end{aligned}}$$

(a) Ponieważ $\boldsymbol{A}^{-1}$ istnieje, to $\boldsymbol{Ax}=\boldsymbol{b} \iff \boldsymbol{A}^{-1}\boldsymbol{Ax}=\boldsymbol{A}^{-1}\boldsymbol{b} \iff \boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$, co oznacza istnienie i jednoznaczność rozwiązania.

(b) Pokażemy, że $\boxed{\text{jeśli } \boldsymbol{Ax}=\boldsymbol{b}, \text{ to } D_j=Dx^j}$, i to nawet *bez założenia, że* $D\neq 0$:

Sposób 1. $\boldsymbol{Ax}=\boldsymbol{b} \iff \boldsymbol{b}=\sum\limits_i\boldsymbol{A}_ix^i$, zatem z liniowości wyznacznika wzgl. j-tej kolumny dostajemy $D_j=\det[\boldsymbol{A}_1,\ldots,\boldsymbol{b},\ldots,\boldsymbol{A}_n]=\det[\boldsymbol{A}_1,\ldots,\sum\limits_i\boldsymbol{A}_ix^i,\ldots,\boldsymbol{A}_n]=$
$$=\sum_i x^i\det[\boldsymbol{A}_1,\ldots,\boldsymbol{A}_i,\ldots,\boldsymbol{A}_n]=x^j\det\boldsymbol{A}=x^j\cdot D.$$

Sposób 2. $\tilde{\boldsymbol{A}}\boldsymbol{A}=D\,\boldsymbol{I}_n$, więc konsekwencją $\boldsymbol{Ax}=b$ jest $D\boldsymbol{x}=\tilde{\boldsymbol{A}}\boldsymbol{A}x=\tilde{\boldsymbol{A}}\boldsymbol{b}$; stąd
$$Dx^j=\sum_i\tilde{A}^j{}_ib^i=\big(\text{rozwinięcie }\det[\boldsymbol{A}_1,\ldots,\boldsymbol{b},\ldots,\boldsymbol{A}_n]\text{ wzgl. kolumny }\boldsymbol{b}\big)=D_j\,.$$

393. **Przykład**. $\begin{bmatrix}p-4&p-3&p-2\\p+1&p-1&p+2\\p+3&p&p+4\end{bmatrix}\begin{bmatrix}x^1\\x^2\\x^3\end{bmatrix}=\begin{bmatrix}3\\p\\5\end{bmatrix}$, $p\in\mathbb{K}$ — parametr. Stosując 'sche-

mat Sarrusa' dostajemy $D = 4-p$, $D_1 = -3p^2+25p-52 = (4-p)(3p-13)$, $D_2 = -p^2-5p+36 = (4-p)(p+9)$, $D_3 = 4p^2-27p+44 = (p-4)(4p-11)$. Zatem układ jest cramerowski $\iff p \neq 4$ i ma wtedy dokładnie jedno rozwiązanie $\boldsymbol{x} = \begin{bmatrix} 3p-13 \\ p+9 \\ -4p+11 \end{bmatrix}$. Z kolei w przypadku $p=4$ możemy zastosować metodę operacji elementarnych, dostając rozwiązanie ogólne $\boldsymbol{x} = \begin{bmatrix} -1 \\ 13 \\ -5 \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$.

394. **Uwaga**. Jak widzieliśmy, relacje $D_j = Dx^j$ są konsekwancją równania $\boldsymbol{Ax} = \boldsymbol{b}$ nawet wtedy, gdy $D = \det \boldsymbol{A} = 0$. Zatem *jeśli układ niecramerowski* (tzn. taki, że $D=0$) *ma rozwiązanie, to* $D_1 = \ldots = D_n = 0$. To samo można wyrazić w innej formie: *jeśli* $D=0$, *przy czym choć jeden z wyznaczników* D_j *jest* $\neq 0$, *to układ jest sprzeczny.*

Jednakże([59]) odwrotne wynikanie jest fałszywe. Aby to sobie wyraźnie uświadomić, wystarczy rozważyć i zapamiętać dwa bardzo proste, lecz pouczające przykłady, w których $D = D_1 = D_2 = 0$:

$$\text{(I)} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \text{(II)} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

6.5 Wyznacznik endomorfizmu

395. **Definicja**. *Wyznacznikiem endomorfizmu* $F \in \operatorname{End} V$ nazywamy liczbę $\boxed{\det F := \det([F]^e_e)}$; nie zależy ona od wyboru bazy e przestrzeni V.

$$\det([F]^f_f) = \det(P[F]^e_e P^{-1}) = \cancel{\det P}\cdot\det([F]^e_e)\cdot\cancel{\det(P^{-1})}, \text{ gdzie } P := [\mathrm{id}_V]^f_e.$$

396. Multiplikatywność $\boldsymbol{A} \mapsto \det \boldsymbol{A}$ daje następujące własności funkcji $\det : \operatorname{End} V \to \mathbb{K}$

$\boxed{\det(F_1F_2) = \det F_1 \cdot \det F_2}$ dla $F_1, F_2 \in \operatorname{End} V$ (*wzór Cauchy'ego*);
$\boxed{\det(FG) = \det(GF)}$, gdy $V \xrightarrow{F} W$, $W \xrightarrow{T} V$, $\dim V = \dim W$([60]);
wobec tego $\boxed{\det(AFA^{-1}) = \det F}$, gdy $F \in \operatorname{End} V$ oraz $A : V \xrightarrow{\cong} W$;
$\det F = 0 \iff$ operator F jest *osobliwy*, tzn. $\ker F \neq \{0\}$.

397. **Uwaga**. Nie da się sensownie zdefiniować wyznacznika operatora $F \in \mathrm{L}(V;W)$ dla dowolnej pary przestrzeni V, W; co więcej, nie da się tego zrobić nawet wtedy, gdy ograniczymy się do par przestrzeni V, W takich, że $\dim V = \dim W$.

Nie da się także sensownie zdefiniować wyznacznika np. *formy kwadratowej*; to samo dotyczy wielu innych 'liniowych obiektów geometrycznych nad przestrzenią V', opisujących się macierzą kwadratową zależną od wyboru bazy e.

[59] Wbrew zadziwiająco popularnej opinii, od lat spotykanej często na kolokwiach i egzaminach.

[60] Jest to istotne założenie, gdyż np. $[1\ 0]\begin{bmatrix} 1 \\ 0 \end{bmatrix} = [1] \neq 0$, $\det(\begin{bmatrix} 1 \\ 0 \end{bmatrix}[1\ 0]) = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = 0$.

7 Formy kwadratowe i przestrzenie euklidesowe

7.1 Formy dwuliniowe

398. Niech V będzie skończenie wymiarową przestrzenią wektorową nad ciałem $\mathbb{K}$. Przypomnijmy, że $\mathrm{L}_2(V) = \mathrm{L}(V,V;\mathbb{K}) = \{\text{dwuliniowe odwzorowania } b : V \times V \to \mathbb{K}\}$, a elementy przestrzeni $\mathrm{L}_2(V)$ nazywa się *formami dwuliniowymi na* V([61]).

Ważny Przykład formy dwuliniowej stanowi funkcja $b = \phi \otimes \psi$, określona dla danych $\phi, \psi \in V^*$ wzorem $b(u,v) := \phi(u)\psi(v)$. Zobaczymy wkrótce, że każdy $b \in \mathrm{L}_2(V)$ jest sumą elementów tej postaci; co więcej, jeśli układ $e^1, \ldots, e^n$ jest bazą V, to n^2 form $e^i \otimes e^j$ tworzy bazę przestrzeni $\mathrm{L}_2(V)$ ([62]).

399. **Fakt** (*kanoniczny izomorfizm* $\boxed{\mathrm{L}_2(V) \cong \mathrm{L}(V;V^*)}$).

(1) Jeśli $b \in \mathrm{L}_2(V)$, to $F(v) := b(\,\cdot\,, v) \in V^*$ dla $v \in V$, przy czym tak określone odwzorowanie F jest liniowe: $F \in \mathrm{L}(V;V^*)$.
(2) Jeśli $F \in \mathrm{L}(V;V^*)$, to wzór $b(u,v) := \langle F(v), u\rangle$ określa pewną formę dwuliniową $b \in \mathrm{L}_2(V)$.
(3) Opisane w (1) i (2) odwzorowania $b \mapsto F$ i $F \mapsto b$ są wzajemnie odwrotnymi izomorfizmami przestrzeni $\mathrm{L}_2(V)$ i $\mathrm{L}(V;V^*)$.

Jak łatwo się domyślić, zapis typu $\phi = b(\,\cdot\,, v)$ oznacza, że $\forall u : \phi(u) = b(u,v)$. Sprawdzenie (1)..(3) stanowi łatwe ćwiczenie. Zauważmy, że 'teoriomnogościowym aspektem' (na poziomie samych zbiorów i odwzorowań, bez struktur wektorowych i liniowości) jest spostrzeżenie, że dla dowolnych zbiorów K, U, V mamy kanoniczną bijekcję $K^{U\times V} \cong \left(K^U\right)^V$, daną przez $b \mapsto \hat{b}$, gdzie $\hat{b}(v) := b(\,\cdot\,, v)$, tzn. $(\hat{b}(v))(u) = b(u,v)$ (dla przypomnienia: K^V oznacza zbiór wszystkich odwzorowań $V \to K$).

W dalszym ciągu chcąc zaznaczyć, że b i F są związane powyższymi relacjami, tzn. że $F(v) = b(\,\cdot\,, v)$ oraz $b(u,v) = \langle F(v), u\rangle$, będziemy pisać $F = F_b$ lub $b = b_F$.

400. **Definicja**. *Rzędem* formy dwuliniowej b nazywamy rząd operatora F_b, tzn. liczbę $\operatorname{rk}(F_b) \in \overline{0, \dim V}$; oznaczamy ją symbolem $\operatorname{rk} b$. Mówimy, że forma b jest *niezdegenerowana*, jeśli $\operatorname{rk} b = \dim V$.

401. **Oznaczenie**. Dla $b \in \mathrm{L}_2(V)$ oznaczmy $b^T(u,v) := b(v,u)$, wtedy oczywiście $b^T \in \mathrm{L}_2(V)$; operację $b \mapsto b^T$ nazywa się *transpozycją* lub *przestawieniem* na przestrzeni $\mathrm{L}_2(V)$. Jej oczywistymi własnościami są:

$$\text{(1) liniowość;} \qquad \text{(2) inwolutywność: } \left(b^T\right)^T = b.$$

Forma b jest *symetryczna*, jeśli $b(u,v) = b(v,u)$, zaś *antysymetryczna*, jeśli $b(u,v) = -b(v,u)$ dla każdych $u, v \in V$. Oczywiście można te warunki zapisać w krótszej postaci:

[61] Litera 'b' nawiązuje tu do angielskiego słowa *bilinear* ('dwuliniowy'); spotykane czasem w żargonie matematycznym słowo 'biliniowy' wydaje się stylistycznie niefortunne: brzmi równie dziwacznie (i pretensjonalnie), jak choćby słowa 'bikrotny', 'bistronny' itp.

[62] Posługując się pojęciem tzw. *iloczynu tensorowego* mówi się w takiej sytuacji, że przestrzeń wektorowa $\mathrm{L}_2(V)$ może być utożsamiana z *iloczynem tensorowym* $V^* \otimes V^*$.

$$b^T = b \text{ (symetria)} \qquad \text{bądź} \qquad b^T = -b \text{ (antysymetria)}.$$

402. **Fakt**. $\mathrm{L}_2(V)$ jest sumą prostą $\mathrm{L}_2(V) = \mathrm{L}_2^+(V) \dot{+} \mathrm{L}_2^-(V)$ podprzestrzeni $\mathrm{L}_2^+(V) =$ (formy symetryczne) i $\mathrm{L}_2^-(V) =$ (formy antysymetryczne). Inaczej mówiąc: każda forma dwuliniowa $b \in \mathrm{L}_2(V)$ ma jednoznaczny rozkład $b = b_+ + b_-$ na sumę formy symetrycznej b_+ i antysymetrycznej b^-.

Jeśli b ma rozkład $b = b_+ + b_-$, gdzie $b_\pm^T = \pm b_\pm$, to $b^T = b_+ - b_-$; z tych dwóch równań wynikają wzory $\boxed{b_+ = \frac{1}{2}(b + b^T),\ \ b_- = \frac{1}{2}(b - b^T)}$, a zatem rozkład może być **co najwyżej jeden**. Gdy z kolei ostatnie wzory przyjmiemy za definicję $b_\pm$, wtedy oczywiście będzie $b_\pm \in \mathrm{L}_2^\pm(V)$ oraz $b = b_+ + b_-$, co dowodzi **istnienia** rozkładu.

Ćwiczenie. Sprawdzić, że

(1) $b = \phi \otimes \psi \ \Rightarrow\ b^T = \psi \otimes \phi$. (2) $\operatorname{rk} b = 1 \iff \exists \phi, \psi \in V^* :\ b = \phi \otimes \psi$.

(3) $F_{b^T} = (F_b)^T$, co oznacza, że jeśli formie b odpowiada operator $F \in \mathrm{L}(V; V^*)$, to b^T odpowiada operator $F^T \in \mathrm{L}(V^{**}; V^*) = \mathrm{L}(V; V^*)$, transponowany względem F.

(4) $\ker F_b = \{v \in V :\ b(\,\cdot\,, v) = 0\}$ (tzw. 'lewe jądro b') jest podprzestrzenią V wymiaru $\dim V - \operatorname{rk} b$; 'prawe jądro b', tzn. podprzestrzeń $\{u \in V :\ b(u, \cdot\,) = 0\}$, ma taki sam wymiar, gdyż jest jądrem operatora ${F_b}^T$, a jak wiemy $\operatorname{rk} F^T = \operatorname{rk} F$.

(5) Jeśli $b \in \mathrm{L}_2^+(V)$ lub $b \in \mathrm{L}_2^-(V)$, to 'prawe jądro' pokrywa się z 'lewym'.

(6) b jest niezdegenerowana $\iff$ jest dwoistością, tzn. $(V, V; b)$ jest *parą dwoistą*.

403. **Definicja**. Jeśli $b \in \mathrm{L}_2(V)$, a e jest bazą $e_1, \ldots, e_n$ przestrzeni V, to macierz kwadratowa $[b_{ij}] \in \mathbb{K}^n{}_n$ o wyrazach $b_{ij} := b(e_i, e_j)$ nazywa się *macierzą formy dwuliniowej b w bazie e* i oznacza symbolem $[b]_e$. Zauważmy, że:

i. Macierz $[b]_e$ w pełni określa formę b, gdyż rozpisując $u, v \in V$ w bazie e: $u = \sum\limits_i e_i x^i$, $v = \sum\limits_j e_j y^j$, wskutek dwuliniowości b dostajemy:
$$b(u, v) = \sum_i \sum_j b(e_i x^i, e_j y^j) = \sum_{i,j=1}^n b(e_i, e_j) x^i y^j.$$

ii. Jeśli $F = F_b$, to $[b]_e = [F]^{e^*}{}_e$, gdzie e^* jest bazą V^* sprzężoną względem e. Istotnie, i-tą współrzędną w bazie e^* formy $\phi = F(e_j)$ jest $\phi(e_i) = \langle F(e_j), e_i \rangle = b(e_i, e_j)$, a więc $F^i{}_j = b_{ij}$.

iii. $[b^T]_e = ([b]_e)^T$, czyli transponowaniu formy odpowiada transpozycja jej macierzy; stąd (b jest (anty-)symetryczna) $\iff$ (macierz $[b]_e$ jest (anty-)symetryczna, zaś macierzami części b_+ i b_- są część symetryczna i antysymetryczna $\boldsymbol{B}_\pm = \frac{1}{2}(\boldsymbol{B} \pm \boldsymbol{B}^T)$ macierzy $\boldsymbol{B} = [b]_e$.

404. **Uwaga**. W notacji macierzowej wzór i. ma postać $\boxed{b(u, v) = [u]^T [b] [v]}$, gdzie $[u] = [u]^e$ i $[v] = [v]^e$ są macierzami wektorów u, v, a $[b] = [b]_e$ jest macierzą formy b, przy czym oczywiście wszystkie te macierze muszą się odnosić do tej samej bazy.

Wniosek. Dla $V = \mathbb{K}^n$ każda forma dwuliniowa $b \in \mathrm{L}_2(V)$ jest postaci
$$b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T \boldsymbol{B} \boldsymbol{v},$$

(iloczyn macierzy z $\mathbb{K}^1{}_n$, $\mathbb{K}^n{}_n$, $\mathbb{K}^n{}_1$ jest liczbą!), gdzie $\boldsymbol{B} \in \mathbb{K}^n{}_n$ jest pewną (określoną jednoznacznie) macierzą, mianowicie $\boldsymbol{B} = [b]_{\mathrm{st}}$.

405. **Fakt**. Jeśli $e^1, \ldots, e^n$ jest bazą V^*, sprzężoną względem bazy $e_1, \ldots, e_n$, to układ n^2 form dwuliniowych $e^i \otimes e^j$, gdzie $i, j \in \overline{1,n}$, tworzy bazę przestrzeni $\mathrm{L}_2(V)$; w konsekwencji $\dim \mathrm{L}_2(V) = n^2$.

W oznaczeniach 403 mamy $(e^i \otimes e^j)(u,v) = e^i(u)e^j(v) = x^i y^j$, a więc wzór i. oznacza, że $b = b_{ij} e^i \otimes e^j$; zatem układ form $e^i \otimes e^j$ rozpina $\mathrm{L}_2(V)$. Ponadto układ ten jest l. niezależny, gdyż wartością formy $\sum\limits_{i,j} \lambda_{i,j} e^i \otimes e^j$ na parze $(u,v) = (e_k, e_l)$ jest $\lambda_{k,l}$.

406. **Wniosek**. Formy $e^i \otimes e^j + e^j \otimes e^i$, gdzie $1 \leqslant i \leqslant j \leqslant n$, tworzą bazę $L_2^+(V)$; formy $e^i \otimes e^j - e^j \otimes e^i$, gdzie $1 \leqslant i < j \leqslant n$, tworzą bazę $L_2^-(V)$. Zatem

$$\dim L_2^+(V) = \tfrac{1}{2}n(n+1), \quad \dim L_2^-(V) = \tfrac{1}{2}n(n-1).$$

407. **Zmiana bazy**. Jeśli oprócz e mamy inną bazę f, zaś $\boldsymbol{A} = [\mathrm{id}_V]^e{}_f$, tzn. $f_k = \sum\limits_i e_i A^i{}_k$, to korzystając z dwuliniowości dostajemy

$$b(f_k, f_l) = b(\sum_i e_i A^i{}_k, \sum_j e_j A^j{}_l) = \sum_{i,j} A^i{}_k A^j{}_l b(e_i, e_j).$$

W zapisie macierzowym oznacza to, że $\boxed{[b]_f = A^T [b]_e A, \;\; A = [\mathrm{id}_V]^e{}_f}$.

408. **Uwaga**. Wyznacznik $\det[\ b\ ]_e$ zależy od bazy, a więc nie jest niezmiennikiem b; wobec tego nie ma pojęcia 'wyznacznika formy dwuliniowej', tym bardziej nie mają sensu takie obiekty, jak 'wielomian charakterystyczny' czy 'wartości i wektory własne' formy dwuliniowej; często jednak spotyka się z ignorancją w tych sprawach. Natomiast jest niezmiennikiem (tzn. nie zależy od bazy) znak wyznacznika $\det[\ b\ ]_e$.

7.2 Formy kwadratowe

409. **Definicja**. Dla $b \in \mathrm{L}_2(V)$ funkcję $Q_b : V \to \mathbb{K}$, określoną wzorem $Q_b(v) := b(v,v)$, nazywamy *formą kwadratową, określoną przez formę* b lub *stowarzyszoną z* b([63]). Funkcję $Q : V \to \mathbb{K}$ nazywamy *formą kwadratową* na przestrzeni V, jeśli istnieje forma $b \in \mathrm{L}_2(V)$, taka że $Q = Q_b$.

410. **Uwaga**. Jeśli $b = \phi \otimes \psi$, to $Q_b = \phi\psi =$ (zwykły iloczyn funkcji $\phi, \psi : V \to \mathbb{K}$) jest *funkcją wielomianową jednorodną stopnia 2.* Dla $\dim V < \infty$ jest to też prawdą dla każdej formy kwadratowej, gdyż $b \in \mathrm{L}_2(V)$ jest sumą składników postaci $b_{ij} e^i \otimes e^j$.

411. **Fakt** (*własności form kwadratowych*): Dla każdych $u, v \in V$ oraz $\lambda \in \mathbb{K}$

$$1^\circ \;\; Q(\lambda v) = \lambda^2 Q(v); \quad 2^\circ \;\; Q(u+v) + Q(u-v) = 2(Q(u) + Q(v)).$$

[63] Q_b jest 'częścią diagonalną funkcji b', tj. jej obcięciem do 'diagonali' $\{(v,v) : v \in V\}$ 'kwadratu kartezjańskiego' $V \times V = \{(u,v) : u, v \in V\}$.
Nietrudno się domyślić, że po formach liniowych i kwadratowych można by rozważać *formy sześcienne na przestrzeni* V, czyli funkcje $V \to \mathbb{K}$ postaci $v \mapsto t(v,v,v)$, gdzie $t \in \mathrm{L}_3(V)$.

Wzór 2° nazywa się *regułą równoległoboku*([64]); wynika on natychmiast z równości

$$Q(u \pm v) = b(u \pm v, u \pm v) = b(u,u) \pm b(u,v) \pm b(v,u) + b(v,v).$$

412. Jak sprawdzić, czy dana funkcja $Q : V \to \mathbb{K}$ jest formą kwadratową? Okazuje się, że w gruncie rzeczy tym pytaniem nie warto sobie zawracać głowy, podobnie np. jak pytaniem, w jaki sposób rozpoznać, czy coś jest wielomianem([65]). Jeśli Q jest formą kwadratową, to jest to po prostu widoczne wprost z jej określenia([66]), i nawet gdy w tym określeniu b nie pojawia wprost, to zawsze łatwo odgadnąć, jak wygląda:

Przykład. Dla $Q : \mathbb{K}^3 \to \mathbb{K}$, $Q(\boldsymbol{x}) = x_1^2 + 3x_1x_2 + 4x_2x_3$ równość $Q(\boldsymbol{x}) = b(\boldsymbol{x},\boldsymbol{y})$ dostaniemy, biorąc $b(\boldsymbol{x},\boldsymbol{y}) = x_1y_1 + 3x_1y_2 + 4x_2y_3$; równie dobrym $b \in \mathrm{L}_2(\mathbb{K}^3)$ jest $b(\boldsymbol{x},\boldsymbol{y}) = x_1y_1 + 3x_2y_1 + 4x_3y_2$ lub $b(\boldsymbol{x},\boldsymbol{y}) = x_1y_1 + x_1y_2 + 2x_2y_1 + 3x_2y_3 + x_3y_2$. Możmy także wybrać $b(\boldsymbol{x},\boldsymbol{y}) = x_1y_1 + \frac{3}{2}x_1y_2 + \frac{3}{2}x_2y_1 + 2x_2y_3 + 2x_3y_2$, a wtedy będzie spełniony warunek symetrii $b(\boldsymbol{x},\boldsymbol{y}) = b(\boldsymbol{y},\boldsymbol{x})$, czyli $b \in \mathrm{L}_2^+(\mathbb{K}^3)$.

Jeśli jednak koniecznie chcemy mieć ogólny sposób wykazywania, że jakieś zadane Q *nie jest* formą kwadratową, albo jeśli chcemy rozwiać wątpliwości, czy napotkana w jakimś podręczniku 'dziwna' definicja formy kwadratowej jest równoważna z naszą, to warto w charakterze bardzo łatwego ćwiczenia udowodnić następujący

Fakt. Mając daną funkcję $Q : V \to \mathbb{K}$ określmy funkcję $b : V \times V \to \mathbb{K}$ wzorem $b(u,v) := \frac{1}{2}\big[Q(u+v) - Q(u) - Q(v)\big]$; wtedy (0) $\Rightarrow$ (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (0), gdzie:

(0) Q jest formą kwadratową;
(1) $b \in \mathrm{L}_2(V)$ oraz $\forall \lambda,\, v:\ Q(\lambda v) = \lambda^2 Q(v)$;
(2) $b \in \mathrm{L}_2(V)$ oraz $\forall v:\ Q(2v) = 4Q(v)$;
(3) $b \in \mathrm{L}_2(V)$ oraz $Q = Q_b$.

Zatem każdy z warunków (1), (2), (3) może być definicją formy kwadratowej([67]).

413. **Dygresja**. Jeśli $Q : V \to \mathbb{K}$ jest funkcją, a $b(u,v) := \frac{1}{2}\big[Q(u+v) - Q(u) - Q(v)\big]$, to nietrudno pokazać, że warunek równoległoboku 2° jest równoważny temu, że b jest formą $\mathbb{Q}$-dwuliniową. Wynika stąd, że dla $\mathbb{Q} \subsetneq \mathbb{K}$ warunki 1° i 2° są niezależne, a także, że istnieją Q spełniające 1° i 2°, lecz nie będące formą kwadratową; natomiast dla $\mathbb{K} = \mathbb{Q}$ warunek 2° implikuje, że $Q = Q_b$ jest formą kwadratową; zatem w tym przypadku warunki 1° i 2° w pełni charakteryzują formy kwadratowe.

Znacznie trudniej pokazać, że dla $\mathbb{K} = \mathbb{R}$ lub $\mathbb{C}$ (ogólniej: dla ciała $\mathbb{K}$ charakterystki $\neq 2$, mającego elementy niealgebraiczne) oraz $\dim V \geqslant 2$ istnieją funkcje $Q : V \to \mathbb{K}$ o własnościach 1°, 2°, nie będące formami kwadratowymi.

414. «**Wzór polaryzacyjny**»: $\boxed{b_+(u,v) = \frac{1}{2}[Q_b(u+v) - Q_b(u) - Q_b(v)]}$ wyraża b_+, tzn. symetryczną część formy $b \in \mathrm{L}_2(V)$, przez Q_b; zatem

forma kwadratowa Q_b zawiera pełną informację o symetrycznej części b.

Mamy $Q_b(u+v) = b(u+v,u+v) = \underbrace{b(u,u)}_{=Q_b(u)} + \underbrace{b(v,v)}_{=Q_b(v)} + \underbrace{b(u,v)+b(v,u)}_{=2b_+(u,v)}$ dzięki 2-liniowości formy b i wprost z definicji formy kwadratowej Q_b, co kończy dowód([68]).

---

[64] Dla $V = \mathbb{R}^2$ oraz $Q(v) :=$ (kwadrat długości v) wzór 2° mówi, że jeśli liczby a, b są długościami boków, a d_1, d_2 — przekątnych równoległoboku, to $d_1^2 + d_2^2 = 2(a^2 + b^2)$.

[65] *'Koń jaki jest, każdy widzi'* (Benedykt Chmielowski, 1700-1763, *Nowe Ateny*).

[66] Chyba, że mamy do czynienia z dziwacznym i pozbawionym głębszego sensu zadaniem.

[67] Niech $0 \neq \phi \in V^*$ i $0 \neq v_0 \in V$; wtedy funkcje $Q_1 := \phi$ oraz $Q_2(v) := \begin{cases} \alpha^2, & v = \alpha v_0, \\ 0, & v \notin \langle v_0 \rangle, \end{cases}$ pokazują, że może być spełniony tylko jeden z dwu warunków $b \in \mathrm{L}_2(V)$ i $Q(\lambda v) = \lambda^2 Q(v)$.

[68] Równie łatwo sprawdzić inny wzór polaryzacyjny $\boxed{b_+(u,v) = Q_b\left(\frac{u+v}{2}\right) - Q_b\left(\frac{u-v}{2}\right)}$;

415. **Fakt**. (1) Zależność $b \mapsto Q_b$ jest liniowa; (2) jej jądrem jest podprzestrzeń $\mathrm{L}_2^-(V)$ form antysymetrycznych; (3) $Q_b = Q_{b'} \iff [b_+ = b'_+$, tzn. b i b' mają jednakowe części symetryczne]. Zatem odwzorowanie $b \mapsto Q_b$ jest izomorfizmem $\mathrm{L}_2^+(V) \xrightarrow{\cong}$ (formy kwadratowe na V).

Ad(2): Jeśli $b \in \mathrm{L}_2^-$, to $\forall v : b(v,v) = 0$, tzn. $Q_b = 0$, więc $\mathrm{L}_2^- \subset K := \{b : Q_b = 0\}$; odwrotnie, dla $b \in K$, tj. $Q_b = 0$, wzór polaryzacyjny daje $b_+ = 0$, tj. $b = b_- \in \mathrm{L}_2^-$.
Ad(3): $Q_b = Q_{b'} \Leftrightarrow Q_{b'-b} = 0 \Leftrightarrow b'-b \in K \Leftrightarrow b'-b \in \mathrm{L}_2^- \Leftrightarrow (b'-b)_+ = 0$.

416. **Definicja**. Niech $b \in \mathrm{L}_2^+(V)$; baza e przestrzeni V, taka że macierz $[b]_e$ jest diagonalna, tzn.

$$\forall i \neq j : \; b_{ij} = b(e_i, e_j) = 0,$$

nazywa się bazą *diagonalizującą formę* b lub *ortogonalną względem* b.

Macierz formy kwadratowej Q zdefiniujmy wzorem $[Q]_e := [b]_e$, gdzie $b \in \mathrm{L}_2^+(V)$ jest określona (jednoznacznie wobec 415) warunkiem $Q = Q_b$. Zatem macierz $[Q]_e = \boldsymbol{B} = [b_{ij}] \in \mathbb{K}^n{}_n$ jest symetryczna ($b_{ij} = b_{ji}$) oraz

$$Q(v) = \sum_{i,j=1}^{n} b_{ij} x^i x^j \quad \text{dla } v = \sum_i e_i x^i.$$

Można ostatni warunek zapisać inaczej: $Q(v) = \boldsymbol{x}^T \boldsymbol{B} \boldsymbol{x}$, gdzie $\boldsymbol{x} = [v]^e$.

Ponieważ $[Q]_e = [b]_e$, to bazę diagonalizującą b nazywa się także *bazą diagonalizującą formę kwadratową* $Q = Q_b$. Jeśli zestaw $\phi_1, \ldots, \phi_n \in V^*$ jest bazą sprzężoną względem e, tzn. funkcje $\phi_i : V \to \mathbb{K}$ tworzą odpowiadający tej bazie układ współrzędnych na V, to

$$b = \sum_{i,j} b_{ij} \phi_i \otimes \phi_j, \quad \text{zaś} \quad Q = \sum_{i,j} b_{ij} \phi_i \phi_j \,.\, ({}^{69})$$

Zatem $\begin{pmatrix}\text{współrzędne } \phi_1, \ldots, \phi_n \text{ odpowiadają bazie diagonalizującej } b\end{pmatrix} \Leftrightarrow \begin{pmatrix} Q = c_1 {\phi_1}^2 + \ldots + c_n {\phi_n}^2 \\ \text{gdzie } c_i = b_{ii} \in \mathbb{K}\end{pmatrix}$; z tego względu o takich *współrzędnych* (tzn. bazie V^*) mówimy, że *diagonalizują formę* Q.

417. **Twierdzenie** (*Lagrange'a, o diagonalizacji*). Każda forma kwadratowa (więc i każda symetryczna forma dwuliniowa) ma bazę diagonalizującą.

Dowód indukcyjny wzgl. $\dim V$. Dla $\dim V = 1$ teza jest oczywista. Dla $\dim V > 1$ oraz $b \neq 0$ ustalmy $v_0 \in V$ taki, że $Q(v_0) \neq 0$, tzn. $b(v_0, v_0) \neq 0$; następnie określmy $U := \{u \in V : \; b(u, v_0) = 0\} = \ker b(\,\cdot\,, v_0) = \ker \phi_0$, $0 \neq \phi_0 = F(v_0) \in V^*$. Zatem $\dim U + 1 = \dim V$ (bilans wymiarów dla $\phi_0 : V \to \mathbb{K}$) oraz $U \cap \langle v_0 \rangle = \{0\}$ (gdyż $v_0 \notin U$), skąd $V = U \dotplus \langle v_0 \rangle$. Ponadto, z określenia U, $v_0 \perp U$ w sensie b, tzn. $\forall u \in U : \; b(u, v_0) = 0$; zatem biorąc bazę $e_1, \ldots, e_{n-1}$ diagonalizującą $b|_{U \times U}$ (istnieje z założenia indukcyjnego) oraz $e_n := v_0$ dostajemy bazę diagonalizującą b.

Metoda Lagrange'a

Opiszemy konstrukcję *współrzędnych* diagonalizujących, zwaną *metodą Lagrange'a* lub *metodą uzupełniania do kwadratu*. Wychodząc z dowolnych 'początkowych' współrzędnych $\phi_1, \ldots, \phi_n$ na V (czyli bazy V^*), będziemy krok po kroku poprawiać te współrzędne, zmniejszając liczbę pozadiagonalnych wyrazów macierzy Q.

---

można wymyślać i inne wzory, wyrażające $b_+(u,v)$ przez wartości formy kwadratowej Q_b.
[69] Dla przypomnienia: składniki $\phi_i \phi_j$ są zwykłymi iloczynami funkcji $\phi_i, \phi_j : V \to \mathbb{K}$.

A oto pojedynczy krok, który wykonujemy, dopóki $\phi_1, \ldots, \phi_n$ nie diagonalizują b.

1° Jeśli $\exists i \in \overline{1,n}: \ b_{ii} \neq 0$, przejdźmy od razu do 2°; w przeciwnym razie weźmy ustalone $i < j$ takie, że $b_{ij} \neq 0$ i oznaczmy $\tilde{\phi}_i = \frac{1}{2}(\phi_i + \phi_j)$, $\tilde{\phi}_j = \frac{1}{2}(\phi_i - \phi_j)$, wtedy $2b_{ij}\phi_i\phi_j = 2b_{ij}\left(\tilde{\phi}_i^2 - \tilde{\phi}_j^2\right)$, więc b wyrażona w 'poprawionych' współrzędnych $\tilde{\phi}_1, \ldots, \tilde{\phi}_n$, gdzie $\tilde{\phi}_k := \phi_k$ dla $k \in \overline{1,n} \setminus \{i,j\}$, ma już $\neq 0$ wyraz diagonalny.

[W dalszym ciągu kolejne — 'coraz lepsze' — układy współrzędnych będziemy oznaczać tu tak, jak wyjściowe, opuszczając falki $\tilde{\cdot}$ dla uproszczenia zapisu].

2° Weźmy ustalone $i \in \overline{1,n}$, takie że $b_{ii} \neq 0$; dla ułatwienia zapisu przyjmijmy, że jest to $i = 1$. Zauważmy, że tę część Q, która zawiera *wszystkie* składniki z ϕ_1, tzn. $\boxed{b_{11}{\phi_1}^2 + 2\sum\limits_{j\geqslant 2} b_{1j}\phi_1\phi_j}$ można zapisać w postaci $\boxed{\frac{1}{b_{11}}\tilde{\varphi}_1^2 - \frac{1}{b_{11}}\Big(\sum\limits_{j\geqslant 2} b_{1j}\phi_j\Big)^2}$, gdzie $\boxed{\tilde{\phi}_1 := b_{11}x_1 + \ldots + b_{1n}\phi_n}$. Stąd wyrażając Q we współrzędnych $\tilde{\phi}_1, \ldots, \tilde{\phi}_n$, gdzie $\tilde{\phi}_j := \phi_j$ dla $j \geqslant 2$, otrzymamy wyrażenie nie zawierające $\tilde{\phi}_1\tilde{\phi}_j$ dla $j \geqslant 2$.

418. **Wniosek** (*postać kanoniczna*([70]) *formy kwadratowej dla* $\mathbb{K} = \mathbb{R}$ *lub* $\mathbb{C}$). Istnieje baza $\phi_1, \ldots, \phi_n$ przestrzeni V^* (tj. współrzędne na V), taka że

$$\boxed{\text{dla } \mathbb{K} = \mathbb{R}}: \qquad Q = \sum_{i=1}^{p} {\phi_i}^2 - \sum_{j=p+1}^{p+q} {\phi_j}^2, \tag{$\dagger$}$$

$$\boxed{\text{dla } \mathbb{K} = \mathbb{C}}: \qquad Q = \sum_{i=1}^{r} {\phi_i}^2. \tag{$\ddagger$}$$

Istotnie, weźmy najpierw dowolne współrzędne diagonalizujące: $Q = \sum\limits_{i=1}^{n} b_{ii}{\phi_i}^2$; następnie je poprzestawiajmy i przenumerujmy w taki sposób, by:

dla $\mathbb{K} = \mathbb{R}$: najpierw stały dodatnie b_{ii}, potem ujemne, a na końcu zerowe;
dla $\mathbb{K} = \mathbb{C}$: niezerowe b_{ii} poprzedzały zerowe.

Ostatni krok polega na zastąpieniu każdej z form ϕ_i, dla $b_{ii} \neq 0$, formą $\lambda_i\phi_i$, gdzie

$$\text{dla } \mathbb{K} = \mathbb{R}: \ \lambda_i := \sqrt{|b_{ii}|}; \qquad \text{dla } \mathbb{K} = \mathbb{C}: \ \lambda_i \in \sqrt{b_{ii}} \text{ (dowolna z dwu wartości).}$$

Uwaga. Oczywiście baza V (sprzężona z $\phi_1, \ldots, \phi_n$), w której Q ma postać kanoniczną, jest *bazą ortonormalną* dla formy dwuliniowej $b \sim Q$, co oznacza, że oprócz *ortogonalności* ($i \neq j \Rightarrow b(e_i, e_j) = b_{ij} = 0$) mamy jeszcze warunek *unormowania*:

$$\forall i \in \overline{1,n}: \ b(e_i, e_i) = b_{ii} \in \begin{cases} \{0, -1, 1\}, & \text{gdy } \mathbb{K} = \mathbb{R}, \\ \{0, 1\}, & \text{gdy } \mathbb{K} = \mathbb{C}. \end{cases}$$

Wprost z definicji: jeśli baza e jest ortonormalna i stosownie uporządkowana, to macierz $\big[\, Q \,\big]_e = \big[\, b \,\big]_e = \big[\, b(e_i, e_j) \,\big]$ jest postaci

$$\text{dla } \mathbb{K} = \mathbb{R}: \ \left[\begin{array}{c|c|c} \boldsymbol{I}_p & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & -\boldsymbol{I}_q & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right], \qquad \text{dla } \mathbb{K} = \mathbb{C}: \ \left[\begin{array}{c|c} \boldsymbol{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right].$$

Z bazy ortogonalnej można zawsze zrobić bazę ortonormalną, *normując* jej wektory, tzn. biorąc

$$\hat{e}_i := \begin{cases} \frac{1}{\lambda_i}e_i, \text{ gdy } b(e_i, e_i) \neq 0, \\ e_i, \quad \text{gdy } b(e_i, e_i) = 0, \end{cases} \text{ gdzie } \lambda_i \in \mathbb{K}, \ \lambda_i^2 = \begin{cases} |b(e_i, e_i)|, \text{ gdy } \mathbb{K} = \mathbb{R}, \\ b_{ii}, \quad \text{gdy } \mathbb{K} = \mathbb{C}. \end{cases}$$

Przypomnijmy, że liczba $\#\{i : b_{ii} \neq 0\}$, równa $p+q$ lub r dla Q danej wzorem wzorem ($\dagger$) lub ($\ddagger$), jest rzędem formy Q (a, z definicji, $\operatorname{rk} Q_b = \operatorname{rk} b$), a więc nie

[70] Zwróćmy uwagę, że często spotykane w tym kontekście wyrażenia 'baza kanoniczna' i 'współrzędne kanoniczne' są bardzo nietrafne merytorycznie: jest wiele baz (i współrzędnych) diagonalizujących daną formę, więc o ich 'kanoniczności' mowy być nie może.

zależy od wyboru bazy diagonalizującej (*jest niezmiennikiem* Q). Zobaczymy teraz, że w przypadku $\mathbb{K} = \mathbb{R}$ także liczby $p, q \in \mathbb{Z}_+$ są *niezmiennikami* formy Q; tradycyjne słowo 'bezwładność' w poniższej nazwie nawiązuje właśnie do tego: zmiana współrzędnych diagonalizujących nie powoduje zmiany liczb p, q.

419. **Twierdzenie** (*Sylvestera, 'o bezwładności'*). Niech $\mathbb{K} = \mathbb{R}$, tzn. V jest przestrzenią rzeczywistą. Jeśli $\phi_1, \ldots, \phi_n$ oraz $\psi_1, \ldots, \psi_n$ są dwiema bazami przestrzeni V^* oraz

$$Q := \sum_{i=1}^{p} {\phi_i}^2 - \sum_{j=p+1}^{p+q} {\phi_j}^2 = \sum_{i=1}^{r} \psi_i^2 - \sum_{j=r+1}^{r+s} \psi_j^2,$$

dla pewnych $p, q, r, s \in \mathbb{Z}_+$ $(p+q \leqslant n,\ r+s \leqslant n)$, to $p = r$ oraz $q = s$.

Wiemy już, że $p+q = r+s =$ (rząd dwuliniowej formy $b = b^T$, odpowiadającej Q). Przypuśćmy, że $p < r$, wtedy układ równań liniowych

$$\phi_1(v) = 0, \ldots, \phi_p(v) = 0, \psi_{r+1}(v) = 0, \ldots, \psi_n(v) = 0$$

ma niezerowe rozwiązanie $v \in V$, gdyż równań jest mniej niż $\dim V$: $p + n - r < n$. Lecz równość $Q(v) = -\sum_{j=p+1}^{p+q} \phi_j(v)^2 = \sum_{i=1}^{r} \psi_i(v)^2$ daje $\sum_i \phi_j(v)^2 + \sum_i \psi_i(v)^2 = 0$, więc $\forall i \in \overline{1,r}\,: \psi_i(v) = 0$; zatem $\psi_1(v) = \ldots = \psi_n(v) = 0$, wbrew temu, że $v \neq 0$.

420. **Definicja**. Niech $Q : V \to \mathbb{R}$ będzie formą kwadratową na rzeczywistej przestrzeni V $(\mathbb{K} = \mathbb{R})$; *sygnaturą formy* Q nazywamy parę (p, q) liczb z $\mathbb{Z}_+$, taką że Q ma w jakichś liniowych współrzędnych (tj. bazie V^*) przedstawienie postaci (†). Piszemy wtedy $\operatorname{sgn} Q = (p, q)$. Taką samą sygnaturę przypisuje się także formie $b \in \mathrm{L}_2^+(V)$, stowarzyszonej z Q:

$$\operatorname{sgn} b = \operatorname{sgn} Q_b.$$

Często zamiast $\operatorname{sgn} Q = (p, q)$ pisze się też $\operatorname{sgn} Q = (\underbrace{+, \ldots, +}_{p}, \underbrace{-, \ldots, -}_{q})$, a więc np. $(+,+,+,-)$ zamiast (3,1), a $(+,-,-)$ zamiast (1,2).

421. **Definicja**([71]). Niech $\mathbb{K} = \mathbb{R}$ oraz $Q : V \to \mathbb{R}$ będzie formą kwadratową; wtedy

$$Q \geqslant 0 \overset{\text{def}}{\iff} \forall v \in V : Q(v) \geqslant 0 \qquad (Q \text{ jest } \textit{dodatnio półokreślona});$$
$$Q > 0 \overset{\text{def}}{\iff} \forall v \in V \setminus \{0\} : Q(v) > 0 \quad (Q \text{ jest } \textit{dodatnio określona});$$

analogicznie nadajemy sens i nazwy własnościom '$Q \leqslant 0$' i '$Q < 0$'.

Dla $b \in \mathrm{L}_2^+(V)$ własności $b \geqslant 0$, $b > 0$, $b \leqslant 0$ i $b < 0$ definiujemy jako równoważne, odpowiednio, z $Q_b \geqslant 0$, $Q_b > 0$, $Q_b \leqslant 0$ i $Q_b < 0$; np.

$$b > 0 \iff \forall v \in V \setminus \{0\} : b(v, v) > 0.$$

422. **Uwaga**. Warto sobie uświadomić, że istnieją formy kwadratowe *nieokreślone*, tzn. nie spełniające ani relacji $Q \geqslant 0$, ani $Q \leqslant 0$; przykładem są formy kwadratowe $Q(\begin{bmatrix} x \\ y \end{bmatrix}) := xy$ oraz $\tilde{Q}(\begin{bmatrix} x \\ y \end{bmatrix}) := x^2 - y^2$, przyjmujące i dodatnie, i ujemne wartości.

Przykład $Q(\begin{bmatrix} x \\ y \end{bmatrix}) := x^2$ dowodzi fałszywości implikacji $\big(Q \geqslant 0,\ Q \neq 0\big) \ \Rightarrow\ Q > 0$.

[71] W tej definicji nie jest potrzebny skończony wymiar V. Wspomnijmy też o innej terminologii: własność '$Q \geqslant 0$' bywa nazywana *dodatniością*, a '$Q > 0$' — *ścisłą dodatniością*.

Natomiast pod pozostałymi względami relacje $\geqslant$, $\leqslant$, $>$, $<$ dla form przypominają uporządkowanie zbioru $\mathbb{R}$: Jeśli oznaczymy $Q_1 \leqslant Q_2 \overset{\text{def}}{\iff} Q_2 - Q_1 \leqslant 0$, to mamy

1° $Q \leqslant Q$ (zwrotność), $\qquad 2^\circ$ $Q_1 \leqslant Q \leqslant Q_2 \;\Rightarrow\; Q_1 \leqslant Q_2$ (przechodniość),

3° $Q_1 \leqslant Q_2 \leqslant Q_1 \;\Rightarrow\; Q_1 = Q_2$ (przeciwsymetria),

4° $\left.\begin{matrix} Q_1 \leqslant \tilde{Q}_1 \\ Q_2 \leqslant \tilde{Q}_2 \end{matrix}\right\} \Rightarrow\; Q_1 + Q_2 \leqslant \tilde{Q}_1 + \tilde{Q}_2$, $\qquad 5^\circ$ $\left.\begin{matrix} Q \geqslant 0 \\ \lambda \in \mathbb{R}_+ \end{matrix}\right\} \;\Rightarrow\; \lambda Q \geqslant 0$;

oczywiście $1^\circ, 3^\circ, 4^\circ, 5^\circ$ oznaczają, że zbiór S wszystkich form $\geqslant 0$ form na V jest *stożkiem wypukłym*, tzn. $S+S \subset S$ i $\mathbb{R}_+ S \subset S$, natomiast 2° — że $S \cap (-S) = \{0\}$.

423. **Fakt** (*nierówność Schwarza*). Jeśli forma $b \in \mathrm{L}_2^+(V)$ jest półokreślona dodatnio, tzn. $b \geqslant 0$, to

$$\forall u, v \in V : \; \Big[\, b(u,v) \Big]^2 \leqslant b(u,u)\, b(v,v).$$

Ustalmy $u, v \in V$ i określmy $f : \mathbb{R} \to \mathbb{R}$ wzorem $f(t) := b(tu+v, tu+v) = b(u,u)t^2 + 2b(u,v)t + b(v,v)$; jest to wielomian stopnia $\leqslant 2$ oraz $\forall t : f(t) \geqslant 0$. Stąd $\Delta \leqslant 0$ (nawet, gdy $b(u,u) = 0$), zaś Δ = wyróżnik $= 4\big(b(u,v)^2 - b(u,u)b(v,v)\big)$.

424. **Fakt**. Niech $\mathbb{K} = \mathbb{R}$, $n := \dim V$, $Q : V \to \mathbb{R}$ — forma kwadratowa, $\operatorname{sgn} Q =: (p,q)$. Wtedy oczywiście

$$\begin{array}{llll} Q \geqslant 0 \iff q = 0, & & Q > 0 \iff (p,q) = (n,0), \\ Q \leqslant 0 \iff p = 0, & & Q < 0 \iff (p,q) = (0,n). \end{array}$$

425. **Twierdzenie** (*Sylvestera-Jacobiego wyznacznikowy wzór na sygnaturę*). Niech Q będzie formą kwadratową na rzeczywistej ($\mathbb{K} = \mathbb{R}$) przestrzeni V, $b \in \mathrm{L}_2^+(V)$ — odpowiadającą Q formą dwuliniową, zaś $b_{ij} := b(e_i, e_j)$, gdzie $e_1, \ldots, e_n$ jest pewną bazą V. Załóżmy, że różne od 0 są wszystkie wyznaczniki $D_1, \ldots, D_n$, gdzie $D_k := \det[b_{i,j}]_{i,j \in \overline{1,k}}$, w szczególności $D_1 = b_{1,1}$, $D_n = \det[\ b\ ]_e$. Wówczas Q ma sygnaturę (p,q), gdzie p jest liczbą dodatnich, a q — ujemnych wyrazów ciągu

$$\frac{D_1}{1},\ \frac{D_2}{D_1},\ \frac{D_3}{D_2},\ \ldots,\ \frac{D_n}{D_{n-1}}.$$

Dowód przeprowadzimy w dwóch krokach:

(1) Pokażemy, że istnieje baza $f_1, \ldots, f_n$ diagonalizująca b i taka, że macierz $[\mathrm{id}_V]^e_f$ jest postaci $\begin{bmatrix} 1 & * & * & \ldots \\ 0 & 1 & * & \ldots \\ 0 & 0 & 1 & \ldots \\ \cdot & \cdot & \cdot & \ldots \end{bmatrix}$, tzn. że $f_j - e_j \in \langle e_1, \ldots, e_{j-1} \rangle$ dla $j \in \overline{1,n}$. W tym celu weźmy $f_1 := e_1$, a dla $j \in \overline{2,n}$ połóżmy $f_j = x_1 e_1 + \ldots + x_{j-1} e_{j-1} + e_j$ i spróbujmy (dla każdego $j \geqslant 2$ z osobna) tak dobrać $x_1, \ldots, x_{j-1}$, by $\boxed{b(e_i, f_j) = 0 \text{ dla } i < j}$. Oznacza to, że $b_{i,1}x_1 + \ldots + b_{i,j-1}x_{j-1} = -b_{i,j}$ dla $i \in \overline{1, j-1}$; ten układ $j-1$ równań na $x_1, \ldots, x_{j-1}$ jest **cramerowski**, gdyż ma wyznacznik główny równy $D_{j-1} \neq 0$. Zatem wektor f_j o postulowanych własnościach istnieje (i to dokładnie jeden). Łatwo sprawdzić, że $f_1, \ldots, f_j$ rozpinają $V_j := \langle e_1, \ldots, e_j \rangle$, tzn. tworzą bazę V_j ([72]).

Otrzymana baza $f_1, \ldots, f_n$ jest ortogonalna względem b, gdyż wskutek warunków $f_i \in \langle e_1, \ldots, e_i \rangle$ i $b(e_1, f_j) = \ldots = b(e_i, f_j) = 0$ dla $i < j$ dostajemy $b(f_i, f_j) = 0$.

(2) Niech $c_{i,j} := b(f_i, f_j)$. Macierze $\boldsymbol{C}_k := [c_{i,j}]_{i,j \in \overline{1,k}}$ oraz $\boldsymbol{B}_k := [b_{i,j}]_{i,j \in \overline{1,k}}$ dla

[72] Mając $v \in V_j$ dobierzmy λ_j tak, by $v - \lambda_j e_j \in V_{j-1}$, wtedy $v_{j-1} := v - \lambda_j f_j \in V_{j-1}$; tak samo można dobrać λ_{j-1} tak, by $v - \lambda_{j-1} f_{j-1} - \lambda_j f_j = v_{j-1} - \lambda_{j-1} f_{j-1} \in V_{j-2}$ itd.

$k \in \overline{1,n}$ są macierzami formy $Q|_{V_k}$ względem baz $f_1, \ldots, f_k$ oraz $e_1, \ldots, e_k$. Zatem ze wzoru na transformację macierzy formy $\boldsymbol{C}_k = \boldsymbol{A}_k^T \boldsymbol{B}_k \boldsymbol{A}_k$, gdzie $\boldsymbol{A}_k = [\mathrm{id}_{V_k}]^{e_k}_{f_k}$. Stąd $\det \boldsymbol{C}_k = (\det \boldsymbol{A}_k)^2 \cdot \det \boldsymbol{B}_k = D_k$ z definicji D_k i stąd, że $\boldsymbol{A}_k$ jest górnotrójkątna i ma jedynki na diagonali. Zarazem macierz $\boldsymbol{C}_k$ jest diagonalna (ortogonalność bazy f), więc (1) $D_k = \det \boldsymbol{C}_k = c_{1,1} \cdot \ldots \cdot c_{k,k}$, skąd $c_{k,k} = \frac{D_k}{D_{k-1}}$, oraz (2) $\operatorname{sgn} Q = (p,q)$, gdzie $p = \#\{k : c_{k,k} > 0\}$, $q = \#\{k : c_{k,k} < 0\}$. To kończy dowód.

426. **Wniosek 1** (*kryterium wyznacznikowe dodatniej określoności formy*). W oznaczeniach twierdzenia Sylvestera-Jacobiego mamy:

$$Q > 0 \iff D_1 > 0, \ldots, D_n > 0.$$

$\boxed{\Leftarrow}$ Wprost z twierdzenia. $\boxed{\Rightarrow}$ Wystarczy oczywiście pokazać, że jeśli forma b jest dodatnio określona, to dla każdej bazy $e_1, \ldots, e_n$ mamy $D_1, \ldots, D_n \neq 0$, gdzie

$$D_k := \det \left[\, b(e_i, e_j) \,\right]_{i,j \in \overline{1,k}} \text{ dla } k \in \overline{1,n}.$$

Zauważmy w tym celu, że warunek $b > 0$ implikuje *niezdegenerowanie* formy b:

$$\Big(b(\,\cdot\,, v) = 0, \text{ tzn. } \forall u \in V : b(u,v) = 0\Big) \;\Rightarrow\; b(v,v) = 0 \;\Rightarrow\; v = 0.$$

Zatem macierz $\left[\, b_{ij} \,\right]$ jest niezdegenerowana, a więc ma wyznacznik $\neq 0$, tj. $D_n \neq 0$. Ostatnim krokiem jest spostrzeżenie, że dodatniość $b : V \times V \to \mathbb{R}$ implikuje dodatniość $b|_{W \times W}$ (a więc i niezdegenerowanie!) dla każdej podprzestrzeni $W \subset V$([73]). Biorąc w szczególności $W = \langle e_1, \ldots, e_k \rangle$ dostajemy stąd $D_k \neq 0$.

427. **Uwaga**. Nie jest prawdą, że jeśli $D_1, \ldots, D_n \geqslant 0$, to $Q \geqslant 0$. Kontrprzykładem może być forma $Q : \mathbb{R}^3 \to \mathbb{R}$, dana wzorem $Q(\boldsymbol{x}) := x_1 x_3 - {x_2}^2$; oczywiście $\operatorname{sgn} Q = (1,2)$, gdyż $x_1 x_3 = \left(\frac{x_1+x_3}{2}\right)^2 - \left(\frac{x_1-x_3}{2}\right)^2$, natomiast $\left[\, Q \,\right]_{\mathrm{st}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, mamy więc $D_1 = D_2 = 0$, $D_3 \geqslant 0$.

Ćwiczenie. Dowieść, że $Q \geqslant 0 \;\Rightarrow\; D_1, \ldots, D_n \geqslant 0$.

428. **Wniosek 2**. W oznaczeniach twierdzenia Sylvestera-Jacobiego mamy

$$Q < 0 \iff -Q > 0 \iff \forall k \in \overline{1,n} : \; (-1)^k D_k > 0.$$

Wynika to wprost z wniosku 1 oraz tego, że $\det(-\boldsymbol{B}) = (-1)^n \det \boldsymbol{B}$ dla $\boldsymbol{B} \in \mathbb{K}^n_n$.

7.3 Przestrzenie euklidesowe

429. **Definicja**. Przestrzenią *euklidesową* nazywamy przestrzeń wektorową nad ciałem $\mathbb{R}$, wyposażoną w dodatnio określoną symetryczną formę dwuliniową b, zwaną (euklidesowym) *iloczynem skalarnym* lub *metryką euklidesową* w tej przestrzeni.

Sztandarowym przykładem jest $V = \mathbb{R}^n$ ze 'standardowym' iloczynem skalarnym

$$(\boldsymbol{x}|\boldsymbol{y}) := \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^T \boldsymbol{y}.$$

Zwykle 'uprawiając geometrię euklidesową' formę $b \in \mathrm{L}_2^+(V)$ oznacza się symbolem

[73] Dygresja: W przeciwieństwie do *dodatniości* własność *niezdegenerowania* dla form nie jest 'dziedziczna': obcięcie do $W \subset V$ formy niezdegenerowanej na V może być formą zdegenerowaną; np. forma $Q\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) := xy$ jest niezdegenerowana na $\mathbb{R}^2$, a znika na $\left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle$.

$(\,\cdot\,|\,\cdot\,)$, pisząc $(u|v)$ zamiast $b(u|v)$; podobnie $b(v,v)$ oznacza się zwykle jako $\|v\|^2$, przy czym liczbę $\|v\| := \sqrt{(v|v)}$ nazywa się (euklidesową) *normą* wektora $v \in V$. Oczywiste własności

$$\|v\| \geqslant 0,\ \ \|v\| = 0 \iff v = 0,\ \ \|\lambda v\| = |\lambda|\,\|v\|,\ \ \|u+v\| \leqslant \|u\| + \|v\|$$

(z których ostatnia jest konsekwencją nierówności Schwarza) powodują, że wielkość $d(u,v) := \|u - v\|$ jest *metryką* na przestrzeni V; jest to metryka *niezmiennicza względem przesunięć*: $d(u,v) = d(u+w, v+w)$, oraz *współzmiennicza* z jednokładnościami: $d(\lambda u, \lambda v) = |\lambda|\, d(u,v)$. Dalszym obiektem jest np. *kąt między parą wektorów*:

$$\angle(u,v) := \arccos \frac{(u|v)}{\|u\|\,\|v\|} \in [-\pi, \pi]$$

(nierówność Schwarza upewnia nas, że takiew określenie jest poprawne). Można teraz łatwo zauważyć, że wiele pojęć, wzorów i faktów szkolnej geometrii przenosi się 'na teren geometrii euklidesowej', przykładowo

$$\cos\angle(u,v) = \frac{a^2+b^2-c^2}{2ac}, \text{ jeśli } a = \|u\|,\ b = \|v\|,\ c = \|u-v\|.$$

Twierdzenie o istnieniu bazy diagonalizującej mówi, że każda przestrzeń euklidesowa V ma bazę ortonormalną, tj. bazę $e_1, \ldots, e_n$ taką, że $(u|v) = \sum\limits_{i=1}^{n} x_i y_i$ dla $x = \sum\limits_i x_i e_i$, $y = \sum\limits_i y_i e_i$. Zatem odwzorowanie $F : \mathbb{R}^n \to V$, $n = \dim V$, określone wzorem $F\Big(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\Big) := x_1 e_1 + \ldots + x_n e_n$, jest izomorfizmem zgodnym z iloczynami skalarnymi obu przestrzeni: $\big(F(\boldsymbol{x})\big|F(\boldsymbol{y})\big) = (\boldsymbol{x}|\boldsymbol{y})$. Można to wysłowić, mówiąc że *każda n-wymiarowa przestrzeń euklidesowa wygląda 'tak samo', jak modelowa przestrzń euklidesowa, czyli $\mathbb{R}^n$ ze standardowym iloczynem skalarnym $(\boldsymbol{x}|\boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}$.*

8 Struktura endomorfizmu

8.1 Wielomiany od macierzy i operatorów

430. **Definicja**. Jeśli $w(\lambda) = c_0 + c_1\lambda + \ldots + c_m\lambda^m \in \mathbb{K}[\,\cdot\,]$, to *wartością wielomianu w na macierzy kwadratowej* $\boldsymbol{A} \in \mathbb{K}^n_{\ n}$ nazywamy macierz

$$w(\boldsymbol{A}) := c_0\boldsymbol{I}_n + c_1\boldsymbol{A} + c_2\boldsymbol{A}^2 + \ldots + c_m\boldsymbol{A}^m,$$

zaś *wartością wielomianu w na operatorze* $F \in \operatorname{End} V$ — operator([74])

$$w(F) := c_0\,\mathrm{id}_V + c_1F + c_2F^2 + \ldots + c_mF^m.$$

431. Przyporządkowania $w \mapsto w(\boldsymbol{A})$ i $w \mapsto w(F)$ (zwane *ewaluacjami* na $\boldsymbol{A}$ i na F) są nie tylko *liniowe*, ale i *multiplikatywne*([75]): jeśli $w(\lambda)$ jest iloczynem $w_1(\lambda)w_2(\lambda)$ dwóch wielomianów, to $w(\boldsymbol{A}) = w_1(\boldsymbol{A})w_2(\boldsymbol{A})$ (iloczyn macierzy) oraz $w(F) = w_1(F) \circ w_2(F)$ (złożenie operatorów). Fakt ten wyrażamy, mówiąc że *odwzorowania $w \mapsto w(\boldsymbol{A})$ i $w \mapsto w(F)$ są homomorfizmami algebry $\mathbb{K}[\,\cdot\,]$ w algebrę $\mathbb{K}^n_{\ n}$ i w algebrę* $\operatorname{End} V$.

Warto pamiętać, że wskutek przemienności mnożenia w $\mathbb{K}[\,\cdot\,]$ macierze $\boldsymbol{B}_1 = w_1(\boldsymbol{A})$, $\boldsymbol{B}_2 = w_2(\boldsymbol{A})$ są zawsze przemienne: $\boldsymbol{B}_1\boldsymbol{B}_2 = \boldsymbol{B}_2\boldsymbol{B}_1$; tak samo $G_1G_2 = G_2G_1$ dla operatorów $G_1 = w_1(F)$, $G_2 = w_2(F)$. W szczególności $\boldsymbol{A}w(\boldsymbol{A}) = w(\boldsymbol{A})\boldsymbol{A}$ oraz $Fw(F) = w(F)F$.

432. Każda z algebr $\mathbb{K}[\,\cdot\,]$, $\mathbb{K}^n_{\ n}$ i $\operatorname{End} V$ ma jedynkę (odpowiednio: wielomian stały równy 1, $\boldsymbol{I}_n$ oraz id_V), a każdy z powyższych homomorfizmów algebr odzorowuje jedynkę na jedynkę.

433. Zauważmy, że jeśli $w_1 \mid w$, tzn. $\exists w_2 : w = w_1w_2$, to mamy implikacje

$$w_1(\boldsymbol{A}) = 0 \;\Rightarrow\; w(\boldsymbol{A}) = 0 \quad \text{oraz} \quad w_1(F) = 0 \;\Rightarrow\; w(F) = 0.$$

434. Ustalmy bazę e przestrzeni V i piszmy krócej $[\,\cdot\,]$ zamiast $[\,\cdot\,]^e_{\ e}$; ponieważ przyporządkowanie $F \mapsto [F]$ jest *homomorfizmem algebr* $\operatorname{End} V$ i $\mathbb{K}^n_{\ n}$ (tzn. jest 'zgodne z działaniami algebraicznymi': dodawaniu i mnożeniu operatorów odpowiadają takie same operacje na ich macierzach), to

$$[w(F)] = w([F]), \quad \text{tzn.} \quad [w(F)] = w(\boldsymbol{A}), \;\text{ gdzie } \boldsymbol{A} := [F].$$

8.2 Wektory i wartości własne

435. **Definicja**. Wektor $v \in V$ nazywamy *wektorem własnym operatora* $F \in \operatorname{End} V$, jeśli $F(v) \in \langle v\rangle$, tzn. jeśli $F(v) = \lambda v$ dla pewnej liczby $\lambda \in \mathbb{K}$. Zauważmy, że niezerowy wektor własny w pełni determinuje tę liczbę λ; nazywamy ją *wartością własną operatora F, odpowiadającą wektorowi własnemu v* ([76]). Tak więc

[74] W tym rozdziale słowa 'operator' i 'endomorfizm' będą synonimami; $F^n := F \circ \ldots \circ F$ jest n-tą potęgą operatora, a symbol składania '$\circ$' będzie często pomijany, np. $FG = F \circ G$.

[75] Jest to oczywiste, gdy oba czynniki są jednomianami; stąd i z rozdzielności mnożenia macierzy (lub składania operatorów) względem dodawania wynika ogólny przypadek.

[76] W tekstach anglojęzycznych używane są terminy *eigenvectors* i *eigenvalues*.

$$\left(\begin{array}{c}\lambda\in\mathbb{K}\text{ jest wartością}\\ \text{własną operatora } F\end{array}\right)\overset{\text{def}}{\iff}\left(\begin{array}{c}F(v)=\lambda v,\text{ tzn. } v\in\ker(F-\lambda\,\mathrm{id}_V)\\ \text{dla pewnego niezerowego } v\in V\end{array}\right).$$

436. Warto uzmysłowić sobie, że jeśli $F(v)=\lambda v$, to także $F^n(v)=\lambda^n v$ dla $n\in\mathbb{Z}_+$ (łatwa indukcja), a w takim razie $\boxed{(w(F))(v)=w(\lambda)v}$ dla dowolnego wielomianu $w\in\mathbb{K}[\,\cdot\,]$ (bo $w(\lambda)$ jest kombinacją liniową λ^n).

Wobec tego łatwe jest obliczenie wartości operatora $G=w(F)$ na wektorze, który jest sumą([77]) kilku wektorów własnych F:

$$\left(\begin{array}{c}v=v_1+\ldots+v_r,\text{ gdzie}\\ F(v_j)=\lambda_j v_j,\ j\in\overline{1,r}\end{array}\right)\quad\Rightarrow\quad G(v)=w(\lambda_1)v_1+\ldots+w(\lambda_r)v_r.$$

437. **Fakt**. Niech liczby $\lambda_1,\ldots,\lambda_r\in\mathbb{K}$ będą parami różne; mamy wtedy([78])

$$\left(\begin{array}{c}F(v_j)=\lambda_j v_j\text{ dla } j\in\overline{1,r}\\ v_1+\ldots+v_r=0\end{array}\right)\quad\Rightarrow\quad v_1=\ldots=v_r=0.$$

Zatem *każdy układ niezerowych wektorów własnych F, odpowiadających parami różnym wartościom własnym, jest liniowo niezależny.*

Weźmy $w_1(\lambda):=(\lambda-\lambda_2)\ldots(\lambda-\lambda_r)$, wtedy $w_1(\lambda_1)\neq 0, w_1(\lambda_2)=\ldots=w_1(\lambda_r)=0$, więc dla $G=w_1(F)$ mamy $0=G(v_1+\ldots+v_r)=v_1$. Analogicznie $v_2=0$, itd([79]).

438. **Ćwiczenie**. Jeśli F ma $n=\dim V$ różnych wartości własnych, to V ma bazę złożoną z wektorów własnych F. Z kolei jeśli V ma bazę złożoną z wektorów własnych F, tzn. $F(e_i)=\lambda_i e_i$ dla $i\in\overline{1,n}$, to $\left[\,F\,\right]^e_e$ jest macierzą diagonalną $\mathrm{diag}\left[\,\lambda_1,\ldots,\lambda_n\,\right]$; wtedy także $\left[\,w(F)\,\right]^e_e=\mathrm{diag}\left[\,w(\lambda_1),\ldots,w(\lambda_n)\,\right]$ dla dowolnego wielomianu w.

439. Wektorami i wartościami własnymi macierzy kwadratowej $\boldsymbol{A}$ nazywamy wektory i wartości własne operatora mnożenia przez macierz $\boldsymbol{A}$, tzn. operatora $A\in\mathrm{End}\,\mathbb{K}^n$, określonego wzorem $A(\boldsymbol{x}):=\boldsymbol{A}\boldsymbol{x}$; jasne, że są one związane z równaniem $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$.

Ponieważ $F(v)=\lambda v\iff\left[\,F\,\right]\left[\,v\,\right]=\lambda\left[\,v\,\right]$ (w ustalonej bazie e przestrzeni V), więc w szczególności F *ma wartości własne takie same, jak macierz* $\boldsymbol{A}:=\left[\,F\,\right]^e_e$.

440. **Ćwiczenie**. Sprawdzić, że: (1) Operator $F\begin{bmatrix}x_1\\x_2\end{bmatrix}:=\begin{bmatrix}-x_2\\x_1\end{bmatrix}$ w $V:=\mathbb{R}^2$ jest obrotem o $90°$; z tej interpretacji widać, że F nie ma $\neq 0$ wektora własnego. (2) Określony tym samym wzorem operator $F\in\mathrm{End}(\mathbb{C}^2)$ ma dwa (tworzące bazę $\mathbb{C}^2$) wektory własne $\begin{bmatrix}1\\i\end{bmatrix}$, $\begin{bmatrix}1\\-i\end{bmatrix}$, odpowiadające wartościom własnym $\lambda_1=-i$, $\lambda_2=i$.

(3) Niech $F\in\mathrm{End}(\mathbb{R}^2)$, $F\begin{bmatrix}x_1\\x_2\end{bmatrix}:=\begin{bmatrix}x_1\\x_1+x_2\end{bmatrix}$; wtedy $\left(\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}\text{ jest wektorem własnym } F\right)\iff x_1=0\iff F\boldsymbol{x}=\boldsymbol{x}$; zatem F ma tylko jedną wartość własną $\lambda=1$; ponadto dla $x_1\neq 0$ wektor $\boldsymbol{x}$ nie jest sumą wektorów własnych F.

---

[77] Skoro krotność wektora własnego jest także wektorem własnym, to kombinacja liniowa wektorów własnych jest sumą wektorów własnych.

[78] Jest to równoważne liniowej niezależności *przestrzeni własnych* $\ker(F-\lambda_i\,\mathrm{id}_V)$, $i\in\overline{1,r}$. Wkrótce poznamy mocniejszy fakt o liniowej niezależności przestrzeni pierwiastkowych.

[79] Inny dowód: $F^k(v_i)={\lambda_i}^k v_i$, więc mamy $\sum_{i=1}^{r} v_i {A^i}_j=0$ dla $j\in\overline{1,r}$, gdzie ${A^i}_j={\lambda_i}^{j-1}$ są wyrazami macierzy nieosobliwej: $\det\boldsymbol{A}=V(\lambda_1,\ldots,\lambda_r)\neq 0$ (wyznacznik Vandermonde'a).

441. **Oznaczenia**. Dla $F \in \operatorname{End} V$, $\lambda \in \mathbb{K}$ oraz $k \in \mathbb{Z}_+ = \{0,1,2,\ldots\}$ zdefiniujmy następujące podprzestrzenie przestrzeni V:

$$V_k(\lambda) =: V_k(\lambda, F) := \ker {F_\lambda}^k = \{v : (F - \lambda \operatorname{id}_V)^k v = 0\},$$

gdzie dla uproszczenia zapisu $F_\lambda := F - \lambda \operatorname{id}_V$. Wskutek oczywistych zawierań $\{0\} = V_0(\lambda) \subset V_1(\lambda) \subset V_2(\lambda) \subset \ldots$ suma mnogościowa

$$V(\lambda) =: V(\lambda, F) := \bigcup_{k=0}^{\infty} V_k(\lambda)$$

także jest podprzestrzenią; co więcej, ciąg liczb całkowitych $\dim V_k(\lambda)$ jest (słabo) rosnący, a także — przy założeniu $\dim V < \infty$ — ograniczony z góry, więc stały od pewnego miejsca; zatem przy ustalonym λ

$$\exists h \geqslant 1 : \ \forall k \geqslant h : \ V_h(\lambda) = V_k(\lambda), \ \text{ czyli } \ V(\lambda) = V_h(\lambda).$$

Zakładać będziemy odtąd, że V ma skończony wymiar oraz $n = \dim V$.

442. **Definicja**. Oznaczmy $\boxed{w_F(\lambda) := \det(F_\lambda) = \det(F - \lambda \operatorname{id}_V)}$; ponieważ $\det G = \det[G]^e_e$ i $[F_\lambda]^e_e = [F]^e_e - \lambda \boldsymbol{I}_n$, więc biorąc $\boldsymbol{A} := [F]^e_e$ mamy

$$w_F(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}_n) = \begin{vmatrix} A^1{}_1 - \lambda & A^1{}_2 & \ldots & A^1{}_n \\ A^2{}_1 & A^2{}_2 - \lambda & \ldots & A^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ A^n{}_1 & A^n{}_2 & \ldots & A^n{}_n - \lambda \end{vmatrix}. \qquad (*)$$

Z rozdziału o wyznaczniku endomorfizmu wiemy już, że ostatni wyznacznik, w odróżnieniu od macierzy $[F]^e_e$, nie zależy od wyboru bazy e przestrzeni V. Ponadto ze wzoru (*) łatwo wywnioskować, że $w_F(\lambda)$ jest wielomianem stopnia n względem λ; nazywa się on *wielomianem charakterystycznym operatora* (sciślej: endomorfizmu) F.

443. **Fakt**. Dla $F \in \operatorname{End} V$ oraz $\lambda \in \mathbb{K}$ następujące warunki są równoważne:

(1) $w_F(\lambda) = 0$;
(2) $V_1(\lambda) \neq \{0\}$, czyli λ jest wartością własną F;
(3) $V(\lambda) \neq \{0\}$, czyli $\exists k \geqslant 1 : \ V_k(\lambda) \neq \{0\}$;
(4) $V_k(\lambda) \neq \{0\}$ dla wszystkich $k \geqslant 1$.

Jak wiemy, $\ker G \neq \{0\} \Leftrightarrow \det G = 0$ dla $G \in \operatorname{End} V$; daje to (1) $\Leftrightarrow$ (2) przy $G = F_\lambda$. Biorąc $G = {F_\lambda}^k$ dostajemy $\det G = \big(w_F(\lambda)\big)^k$, skąd (1) $\Leftrightarrow$ (3) i (1) $\Leftrightarrow$ (4).

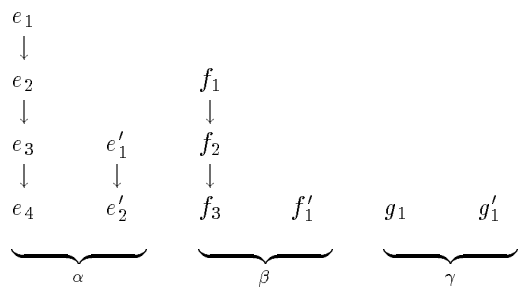
444. **Definicja**. Zbiór $\boxed{\operatorname{Sp} F := \{\lambda \in \mathbb{K} : \ w_F(\lambda) = 0\}}$, czyli zbiór wartości własnych F, nazywamy *spektrum* lub *widmem* operatora F.
Dla $\lambda \in \operatorname{Sp} F$ przestrzenie $V_1(\lambda)$ (niezerowe!) nazywa się *przestrzeniami własnymi operatora* F, zaś $V_k(\lambda) \supseteq V_1(\lambda)$ — *uogólnionymi przestrzeniami własnymi* F. Przestrzenie $V(\lambda)$, tzn. maksymalne spośród przestrzeni $V_k(\lambda)$, nazywamy *przestrzeniami pierwiastkowymi operatora* F.

445. **Przykład**. Niech przestrzeń V ma bazę $e_1, e_2, e_3, e_4, e'_1, e'_2, f_1, f_2, f_3, f'_1, g_1, g'_1$ (a zatem $\dim V = 12$), w której operator F wyraża się wzorami

$$\begin{aligned} F(e_i) &= \alpha e_i + e_{i+1}, & F(e'_i) &= \alpha e'_i + e'_{i+1}, \\ F(f_i) &= \beta f_i + f_{i+1}, & F(f'_i) &= \beta f'_i + f'_{i+1}, \\ F(g_1) &= \gamma g_1, & F(g'_1) &= \gamma g'_1, \end{aligned}$$

gdzie $\alpha, \beta, \gamma \in \mathbb{K}$ są parami różne; przy tym dla skrócenia zapisu umawiamy się,

że $e_5 := 0$, $e'_3 := 0$, $f_4 := 0$, $f'_2 := 0$. Powyższą sytuację wygodnie jest obrazować następującym diagramem:



Nietrudne (a warte zrobienia i przemyślenia) ćwiczenie polega na sprawdzeniu, że mamy wtedy $\operatorname{Sp} F = \{\alpha, \beta, \gamma\}$, $V_1(\alpha) = \langle e_4, e'_2\rangle$, $V_2(\alpha) = \langle e_3, e_4, e'_1, e'_2\rangle$, $V_3(\alpha) = \langle e_2, e_3, e_4, e'_1, e'_2\rangle$, $V_4(\alpha) = \langle e_1, e_2, e_3, e_4, e'_1, e'_2\rangle$, $V(\alpha) = V_4(\alpha)$; $V_1(\beta) = \langle f_3, f'_1\rangle$, $V_2(\beta) = \langle f_2, f_3, f'_1\rangle$, $V_3(\beta) = \langle f_1, f_2, f_3, f'_1\rangle$, $V(\beta) = V_3(\beta)$; $V_1(\gamma) = V(\gamma) = \langle g_1, g'_1\rangle$.

8.3 Podprzestrzenie niezmiennicze wzgl. operatora

446. **Definicja**. Podprzestrzeń $W \subset V$ nazywamy *niezmienniczą względem operatora* F (krócej: *F-niezmienniczą*), jeśli $\boxed{F(W) \subset W}$, tzn. jeśli

$$\forall w \in W : F(w) \in W.$$

Oznacza to, że obcięcie $F|_W : W \to V$ przyjmuje wartości jedynie z W, a więc może być traktowane jako operator należący do $\operatorname{End} W$.

447. **Przykład**. Sprawdzimy, że jeśli operatory $F, G \in \operatorname{End} V$ *komutują*, tzn. $FG = GF$, to podprzestrzenie $\ker G$ i $\operatorname{im} G$ są F-niezmiennicze:

$$v \in \ker G \;\Leftrightarrow\; G(v) = 0 \;\Rightarrow\; G\big(F(v)\big) = F\big(G(v)\big) = F(0) = 0 \;\Rightarrow\; F(v) \in \ker G;$$
$$v \in \operatorname{im} G \;\Leftrightarrow\; \exists u \in V :\; v = G(u) \;\Rightarrow\; F(v) = F\big(G(u)\big) = G\big(F(u)\big) \in \operatorname{im} G.$$

448. **Fakt**. Każda z podprzestrzeni $V_k(\lambda)$ oraz $V(\lambda)$ jest F-niezmiennicza.

$V_k(\lambda) = \ker G$, gdzie $G = (F - \lambda \operatorname{id}_V)^k$ jest wielomianem od F, więc $FG = GF$; zatem $V_k(\lambda)$ jest F-niezmiennicz. Zarazem $V(\lambda) = V_k(\lambda)$ dla dużych k, skąd teza.

449. Jeśli $e_1, \ldots, e_n$ jest bazą V, zaś $F^i{}_j$ — wyrazami macierzy $[F]^e{}_e$, tzn. $F(e_j) = \sum\limits_i e_i F^i{}_j$, to wprost z określenia wynika, że

   i. $\langle e_1, \ldots, e_k\rangle$ jest F-niezmiennicza wtedy i tylko wtedy, gdy $\forall j \in \overline{1,k} : F(e_j) \in \langle e_1, \ldots, e_k\rangle$, tzn. $F^{k+1}{}_j = \ldots = F^n_j = 0$; oznacza to, że macierz $[F]^e{}_e$ podzielona na bloki jest postaci

   $$[F]^e{}_e = \left[\begin{array}{c|c} * & * \\ \hline \mathbf{0} & * \end{array}\right],$$

   gdzie $\mathbf{0}$ jest macierzą zerową z $\mathbb{K}^{n-k}{}_k$, zaś $*$ oznaczają dowolne macierze ('bloki') stosownych wymiarów.

   ii. Analogicznie, F-niezmienniczość $\langle e_{k+1}, \ldots, e_n\rangle$ oznacza, że

   $$[F]^e{}_e = \left[\begin{array}{c|c} * & \mathbf{0} \\ \hline * & * \end{array}\right],$$

   gdzie $\mathbf{0}$ jest macierzą zerową z $\mathbb{K}^k{}_{n-k}$.

iii. Zatem F-niezmienniczość obu podprzestrzeni: $\langle e_1, \ldots, e_k\rangle$ oraz $\langle e_{k+1}, \ldots, e_n\rangle$, jest równoważna postaci $[F]^e_e = \left[\begin{array}{c|c} * & \mathbf{0} \\ \hline \mathbf{0} & * \end{array}\right]$, gdzie macierze zerowe są, odpowiednio, z przestrzeni $\mathbb{K}^k{}_{n-k}$ i $\mathbb{K}^{n-k}{}_k$.

iv. Ogólniej, gdy $V = V_1 \dot{+} \ldots \dot{+} V_r$ jest sumą prostą swoich podprzestrzeni $V_1, \ldots, V_r$, wtedy F-niezmienniczość wszystkich V_i jest równoważna temu, że w bazie V zgodnej z tym rozkładem F wyraża się macierzą 'blokowo diagonalną', postaci

$$[F]^e_e = \left[\begin{array}{c|c|c|c} * & \mathbf{0} & \ldots & \mathbf{0} \\ \hline \mathbf{0} & * & \ldots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \ldots & * \end{array}\right].$$

450. **Fakt**. Jeśli V jest sumą prostą swoich podprzestrzeni $V_1, \ldots, V_k$, przy czym są one F-niezmiennicze, to

$$\det F = \det F_1 \ldots \det F_r,$$
$$w_F(\lambda) = w_{F_1}(\lambda) \ldots w_{F_r}(\lambda),$$

gdzie $F_i \in \operatorname{End} V_i$ jest obcięciem $F|_{V_i}$ operatora F do podprzestrzeni V_i.

W bazie V, będącej konkatenacją baz poszczególnych podprzestrzeni V_i, macierz $[F]$ jest blokowo-diagonalna, więc $\det F = \det[F]$ jest iloczynem wyznaczników bloków diagonalnych $[F_i]$, zaś $\det[F_i] = \det F_i$. Drugi wzór dostajemy z pierwszego, zastępując F operatorem $F_\lambda = F - \lambda \operatorname{id}_V$.

8.4 Wielomian charakterystyczny i niezmienniki endomorfizmu

451. **Definicja**. Jeśli $\boldsymbol{A} \in \mathbb{K}^n{}_n$, to $w_{\boldsymbol{A}}(\lambda) := \det(\boldsymbol{A} - \lambda \boldsymbol{I}_n)$ jest oczywiście wielomianem stopnia n od λ:

$$w_{\boldsymbol{A}}(\lambda) = \tau_n(\boldsymbol{A}) - \tau_{n-1}(\boldsymbol{A})\lambda + \ldots + \tau_1(\boldsymbol{A})(-\lambda)^{n-1} + (-\lambda)^n;$$

nazywamy go *wielomianem charakterystycznym macierzy* $\boldsymbol{A}$, a jego współczynniki $\tau_k(\boldsymbol{A})$ — *niezmiennikami podstawowymi* $\boldsymbol{A}$.

Jeśli $F \in \operatorname{End} V$, to $w_F(\lambda) := \det(F - \lambda \operatorname{id}_V)$ nazywa się *wielomianem charakterystycznym operatora* F:

$$w_F(\lambda) = \tau_n(F) - \tau_{n-1}(F)\lambda + \ldots + \tau_1(F)(-\lambda)^{n-1} + (-\lambda)^n;$$

współczynniki $\tau_k(F)$ nazywają się *niezmiennikami podstawowymi* operatora F. Skoro $[F - \lambda \operatorname{id}_V]^e_e = [F]^e_e - \lambda \boldsymbol{I}_n$, to $\boxed{w_F(\lambda) = w_{\boldsymbol{A}}(\lambda)}$, a więc również $\boxed{\tau_k(F) = \tau_k(\boldsymbol{A})}$, jeśli $\boldsymbol{A} = [F]^e_e$, gdzie e jest *dowolną* bazą V.

Biorąc $\lambda = 0$ widzimy natychmiast, że $\tau_n(\boldsymbol{A}) = \det \boldsymbol{A}$, $\tau_n(F) = \det F$.

452. **Uwaga**. Załóżmy, że $\mathbb{K} = \mathbb{C}$ lub, ogólniej, że wielomian charakterystyczny jest rozkładalny na czynniki stopnia 1. Wzory Viete'a (wyrażające współczynniki wielomianu przez jego pierwiastki) dają wtedy dla $\tau_k = \tau_k(F)$ lub $\tau_k = \tau_k(\boldsymbol{A})$ relacje

$$\tau_1 = \lambda_1 + \ldots + \lambda_n, \qquad \tau_n = \lambda_1 \cdot \ldots \cdot \lambda_n,$$
$$\tau_2 = \lambda_1\lambda_2 + \ldots + \lambda_1\lambda_n + \lambda_2\lambda_3 + \ldots + \lambda_2\lambda_n + \ldots = \sum_{i<j} \lambda_i\lambda_j,$$

gdzie $\lambda_1, \ldots, \lambda_n$ są wszystkimi pierwiastkami wielomianu $w_F(\lambda)$ lub $w_{\boldsymbol{A}}(\lambda)$, tzn. $\{\lambda_1, \ldots, \lambda_n\} = \operatorname{Sp} F$ lub $\operatorname{Sp} \boldsymbol{A}$; przyjmujemy oczywiście konwencję, że ciąg $\lambda_1, \ldots, \lambda_n$ uwzględnia krotności pierwiastków.

Ogólnie, $\tau_k = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_k}$ jest dla $k \in \overline{1,n}$ tzw. *k-tym symetrycznym wielomianem podstawowym* od wartości własnych operatora F lub macierzy $\boldsymbol{A}$.

453. **Fakt** (*niezmienniczość*([80])). Jeśli $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{K}^n{}_n$, to $\tau_k(\boldsymbol{AB}) = \tau_k(\boldsymbol{BA})$, więc tym bardziej $\tau_k(\boldsymbol{ABA^{-1}}) = \tau_k(\boldsymbol{B})$.

Analogiczne wzory zachodzą dla endomorfizmów: jeśli $F, G \in \operatorname{End} V$, to $\boxed{\tau_k(FG) = \tau_k(GF)}$, a także $\boxed{\tau_k(AFA^{-1}) = \tau_k(F)}$ dla $A : V \xrightarrow{\cong} W$.

$\boldsymbol{ABA}^{-1} - \lambda \boldsymbol{I}_n = \boldsymbol{A}(\boldsymbol{B} - \lambda \boldsymbol{I}_n)\boldsymbol{A}^{-1}$, więc licząc wyznacznik i opuszczając znoszące się czynniki $\det \boldsymbol{A} \cdot \det \boldsymbol{A}^{-1}$ widzimy, że wielomiany charakterystyczne $\boldsymbol{ABA}^{-1}$ i $\boldsymbol{B}$ są jednakowe, więc $\tau_k(\boldsymbol{ABA}^{-1}) = \tau_k(\boldsymbol{B})$. Zamieniając tu $\boldsymbol{B}$ na $\boldsymbol{BA}$ dostajemy wzór $\tau_k(\boldsymbol{AB}) = \tau_k(\boldsymbol{BA})$, lecz przy dodatkowym założeniu $\det \boldsymbol{A} \neq 0$; otóż założenie to można opuścić, bo obie strony są wielomianowymi funkcjami wyrazów macierzy $\boldsymbol{A}$ (dwa wielomiany, przyjmujące jednakowe wartości na 'dużym' zbiorze, są równe, a zbiór $\{\boldsymbol{A} : \det \boldsymbol{A} \neq 0\}$ jest 'duży')([81]). Z kolei wzory dla operatorów wynikają z analogicznych wzorów dla macierzy oraz z relacji $\tau_k(F) = \tau_k\big([F]^e_e\big)$.

454. **Uwaga**. Niezmienniki $\tau_k(\,\cdot\,)$ są nazywane *podstawowymi* (lub *fundamentalnymi*) z powodu następującego faktu: *każdy wielomianowy 'niezmiennik', tzn. funkcja $\tau : \operatorname{End} V \to \mathbb{K}$ o własności $\tau(AFA^{-1}) = \tau(F)$, jest wielomianem od $\tau_1, \ldots, \tau_n$*([82]).

Można np. sprawdzić (zakładając najpierw diagonalizowalność operatora F), że

$$\operatorname{tr}(F^2) = \big(\tau_1(F)\big)^2 - \tau_2(F).$$

455. **Fakt**. $\tau_k(\boldsymbol{A})$ jest sumą wszystkich $\binom{n}{k}$ minorów głównych([83]) stopnia k macierzy $\boldsymbol{A}$; w szczególności

$$\tau_1(\boldsymbol{A}) = \sum_{i=1}^{n} A^i{}_i = \operatorname{tr} \boldsymbol{A},\ \tau_2(\boldsymbol{A}) = \sum_{i<j} \begin{vmatrix} A^1_1 & A^1_2 \\ A^2_1 & A^2_2 \end{vmatrix},\ \ldots,\ \ \tau_n(\boldsymbol{A}) = \det \boldsymbol{A}.$$

$\tau_k(\boldsymbol{A})$ jest współczynnikiem przy λ^{n-k} w $w_{\boldsymbol{A}}(-\lambda)$. Niech $S^p_j := \left\{ \begin{array}{l} \boldsymbol{A}_j,\ p = 0 \\ \boldsymbol{e}_j,\ p = 1 \end{array} \right\}$, wtedy

---

[80] Przymiotnik ten (niezbyt trafny, lecz uświęcony tradycją) oznacza tu *stałość* funkcji $\boldsymbol{A} \mapsto \tau_k(\boldsymbol{A})$ i $F \mapsto \tau_k(F)$ na każdej z klas relacji '$\sim$' (zwanej czasem *podobieństwem*), określonej w zbiorach $\mathbb{K}^n{}_n$ i $\operatorname{End} V$ wzorami $\boldsymbol{B}' \sim \boldsymbol{B} \overset{\text{def}}{\iff} \exists \boldsymbol{A} :\ \boldsymbol{B}' = \boldsymbol{ABA}^{-1}$ oraz $F' \sim F \overset{\text{def}}{\iff} \exists A :\ F' = AFA^{-1}$.

[81] Jest to prosta i wręcz modelowa ilustracja zastosowania tzw. *Zasady Usuwalności Nierówności Algebraicznych*, będącej wygodnym i skutecznym narzędziem w wielu dowodach.

[82] Dowód opiera się na tym, że: (1) operatory diagonalizowalne tworzą 'duży' podzbiór $\operatorname{End} V$, oraz (2) każdy wielomian *symetryczny* $f(\lambda_1, \ldots, \lambda_n)$, tzn. niezmienniczy względem przestawiania zmiennych λ_i, jest wielomianem od tzw. *wielomianów symetrycznych podstawowych* $\sum_{i=1}^{n} \lambda_i,\ \sum_{i<j} \lambda_i\lambda_j,\ \ldots,\ \lambda_1\lambda_2 \ldots \lambda_n$, tzn. wzpółczynników $W(\lambda) = \prod_{i=1}^{n} (\lambda + \lambda_i)$.

[83] *Minorem* stopnia k macierzy $\boldsymbol{A}$ nazywamy wyznacznik jej podmacierzy wymiaru $k \times k$; minor jest *główny*, jeśli utworzony jest z wierszy i kolumn o tych samych numerach.

$$w_{\boldsymbol{A}}(-\lambda) = \det(\boldsymbol{A}+\lambda\boldsymbol{I}_n) = \det[\boldsymbol{A}_1+\lambda\boldsymbol{e}_1,\ldots,\boldsymbol{A}_n+\lambda\boldsymbol{e}_n] =$$
$$= \sum \det[\boldsymbol{A}_1 \text{ lub } \lambda\boldsymbol{e}_1,\ldots,\boldsymbol{A}_n \text{ lub } \lambda\boldsymbol{e}_n] = \sum_{p_1,\ldots,p_n} \lambda^{p_1+\ldots+p_n}\det[S_1^{p_1},\ldots,S_n^{p_n}],$$

gdzie indeksy p_j przebiegają zbiór $\{0,1\}$, a więc cała suma ma 2^n składników. Zatem współczynnik przy λ^{n-k} możemy wyrazić sumą $\sum_{p_1+\ldots+p_n=n-k} \det(S_1^{p_1},\ldots,S_n^{p_n})$ mającą $\binom{n}{k}$ składników. Składnik z $p_1=\ldots=p_k=0, p_{k+1}=\ldots=p_n=1$ jest na mocy tw. o wyznaczniku macierzy blokowo-trójkątnej równy $\det[A_1,\ldots,A_k,e_{k+1},\ldots,e_n] = \begin{vmatrix} A^1{}_1 & \ldots & A^1{}_k \\ \ldots & \ldots & \ldots \\ A^k{}_1 & \ldots & A^k{}_k \end{vmatrix}$; pozostałe składniki $\tau_k(\boldsymbol{A})$ są także minorami głownymi stopnia k.

Przykład. Dla $\boldsymbol{A} = \begin{bmatrix} 3 & 5 \\ 9 & 7 \end{bmatrix}$: $w_{\boldsymbol{A}}(\lambda) = \det\boldsymbol{A} - (\operatorname{tr}\boldsymbol{A})\lambda + \lambda^2 = -24 - 10\lambda + \lambda^2 = (\lambda+2)(\lambda-12)$. Dla $\boldsymbol{A} = \begin{bmatrix} 8 & 2 & 5 \\ 3 & 9 & 4 \\ 1 & 7 & 6 \end{bmatrix}$ z kolei $\tau_1(\boldsymbol{A}) = \operatorname{tr}\boldsymbol{A} = 8+9+6 = 23$, $\tau_2(\boldsymbol{A}) = \begin{bmatrix} 8 & 2 \\ 3 & 9 \end{bmatrix} + \begin{bmatrix} 8 & 5 \\ 1 & 6 \end{bmatrix} + \begin{bmatrix} 9 & 4 \\ 7 & 6 \end{bmatrix} = 66+43+26 = 135$, $\tau_3(\boldsymbol{A}) = \det\boldsymbol{A} = 240$. Zatem $w_{\boldsymbol{A}}(\lambda) = 240 - 135\lambda + 23\lambda^2 - \lambda^3$.

8.5 Wielomiany zerujące operator

456. **Twierdzenie** (Cayleya-Hamiltona)

(a) Każda macierz kwadratowa jest zerowana przez swój wielomian charakterystyczny: $w_{\boldsymbol{A}}(\boldsymbol{A}) = 0$.

(b) Każdy endomorfizm $F \in \operatorname{End} V$ jest zerowany przez swój wielomian charakterystyczny: $w_F(F) = 0$.

(a) Niech $\boldsymbol{B}(\lambda) := \widetilde{\boldsymbol{A}-\lambda\boldsymbol{I}_n}$ będzie macierzą dopełnień alg. macierzy $\boldsymbol{A}-\lambda\boldsymbol{I}_n$; zamieniając $\boldsymbol{A}$ na $\boldsymbol{A}-\lambda\boldsymbol{I}_n$ we wzorze $(\det\boldsymbol{A})\boldsymbol{I}_n = \boldsymbol{A}\widetilde{\boldsymbol{A}}$ dostajemy tożsamość $w_{\boldsymbol{A}}(\lambda)\boldsymbol{I}_n = (\boldsymbol{A}-\lambda\boldsymbol{I}_n)\boldsymbol{B}(\lambda)$. Wyrazy macierzy $\boldsymbol{B}(\lambda)$ są oczywiście wielomianami stopnia $\leqslant n-1$, a zatem $\boldsymbol{B}(\lambda) = \boldsymbol{B}_0+\boldsymbol{B}_1\lambda+\ldots+\boldsymbol{B}_{n-1}\lambda^{n-1}$ dla pewnych $\boldsymbol{B}_k \in \mathbb{K}^n{}_n$; oznaczmy ponadto $w_{\boldsymbol{A}}(\lambda) = c_0+c_1\lambda+\ldots+c_n\lambda^n$. Dostajemy wtedy tożsamość

$$(c_0+c_1\lambda+\ldots+c_n\lambda^n)\boldsymbol{I}_n = (\boldsymbol{A}-\lambda\boldsymbol{I}_n)(\boldsymbol{B}_0+\boldsymbol{B}_1\lambda+\ldots+\boldsymbol{B}_{n-1}\lambda^{n-1}),$$

z której wynikają równości współczynników przy kolejnych potęgach zmiennej λ:

$$c_0\boldsymbol{I}_n = \boldsymbol{A}\boldsymbol{B}_0,\ c_1\boldsymbol{I}_n = \boldsymbol{A}\boldsymbol{B}_1-\boldsymbol{B}_0,\ \ldots,\ c_{n-1}\boldsymbol{I}_n = \boldsymbol{A}\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2},\ c_n\boldsymbol{I}_n = -\boldsymbol{B}_{n-1}.$$

Mnożąc te równości lewostronnie przez $\boldsymbol{I}, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1}, \boldsymbol{A}^n$ oraz dodając dostajemy:

$$w_{\boldsymbol{A}}(\boldsymbol{A}) = c_0\boldsymbol{I}_n + c_1\boldsymbol{A} + \ldots + c_n\boldsymbol{A}^n =$$
$$= \boldsymbol{A}\boldsymbol{B}_0 + \boldsymbol{A}(\boldsymbol{A}\boldsymbol{B}_1-\boldsymbol{B}_0) + \ldots + \boldsymbol{A}^{n-1}(\boldsymbol{A}\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}) + \boldsymbol{A}^n(-\boldsymbol{B}_{n-1}) = 0.$$

(b) Ustalmy bazę e w V i piszmy $[\,\cdot\,]$ zamiast $[\,\cdot\,]^e_e$; wtedy $[w(F)] = w\big([F]\big) = w(\boldsymbol{A})$ dla każdego wielomianu $w(\lambda)$ (gdyż $F \mapsto [F]$ jest homomorfizmem algebr), więc dla $w(\lambda) := w_F(\lambda) = w_{\boldsymbol{A}}(\lambda)$, $\boldsymbol{A} := [F]$, dostajemy $[w_F(F)] = [w_{\boldsymbol{A}}(F)] = w_{\boldsymbol{A}}(\boldsymbol{A}) = 0$.

457. **Wniosek**. Każdy wielomian od macierzy $\boldsymbol{A} \in \mathbb{K}^n{}_n$, w szczególności każdą jej potęgę, można przedstawić jako wielomian stopnia $< n$ od $\boldsymbol{A}$, tzn. jako kombinację liniową $\boldsymbol{I}, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1}$. Analogicznie: każdy wielomian od operatora $F \in \operatorname{End} V$ można przedstawić jako kombinację

liniową operatorów $\mathrm{id}_V, F, \ldots, F^{n-1}$, gdzie $n = \dim V$.

Istotnie, dzieląc $w(\lambda)$ przez $w_{\boldsymbol{A}}(\lambda)$ dostajemy $w(\lambda) = q(\lambda)w_{\boldsymbol{A}}(\lambda)+\varrho(\lambda)$, gdzie $\varrho(\lambda)$ jest wielomianem stopnia $< n$. Skoro $w_{\boldsymbol{A}}(\boldsymbol{A}) = 0$, to $w(\boldsymbol{A}) = \varrho(\boldsymbol{A})$.
Jest też możliwy inny dowód: to, że $\boldsymbol{A}^k \in \langle \boldsymbol{I}, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1} \rangle$, dowodzimy indukcyjnie względem $k \geqslant n$, korzystając przy tym z relacji $0 = w_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{A}^n - \tau_1 \boldsymbol{A}^{n-1} + \ldots$.

Uwaga. Współczynniki wielomianu lub kombinacji liniowej zależą oczywiście zarówno od wielomianu $w(\lambda)$ czy wykładnika, jak i od $\boldsymbol{A}$ lub F.

458. **Fakt**. Dla $F \in \mathrm{End}\, V$ istnieje w $\mathbb{K}[\,\cdot\,]$ dokładnie jeden wielomian, który

- zeruje F;
- jest dzielnikiem każdego wielomianu zerującego F;
- jest unormowany (ma współczynnik 1 przy najwyższej potędze λ).

Nazywamy go *wielomianem minimalnym* operatora F i oznaczamy symbolem $\widehat{w}_F(\lambda)$. Skoro $w_F(F) = 0$ (tw. Cayleya-Hamiltona), to $\widehat{w}_F \mid w_F$.

$\mathcal{J}_F := \{w \in \mathbb{K}[\,\cdot\,] : w(F) = 0\}$ jest oczywiście ideałem w pierścieniu $\mathbb{K}[\,\cdot\,]$; ideał ten jest główny, bo $\mathbb{K}[\,\cdot\,]$ jest dziedziną ideałów głównych, oraz $\neq \{0\}$, gdyż $w_F \in \mathcal{J}_F$.

459. **Ćwiczenie**. Dowieść, że wielomiany $w_F(\lambda)$ i $\widehat{w}_F(\lambda)$ mają jednakowe pierwiastki.

Rozwiązanie. Oczywiście $\widehat{w}_F^{-1}\{0\} \subset {w_F}^{-1}\{0\}$, gdyż $\widehat{w}_F \mid w_F$. Odwrotne zawieranie: jeśli $w_F(\lambda) = 0$, to $\exists v \neq 0 : \; Fv = \lambda v$, więc $w(F)v = w(\lambda)v$ dla dowolnego wielomianu w; biorąc $w = \widehat{w}_F$ dostajemy $0 = \widehat{w}_F(F)v = \widehat{w}_F(\lambda)v$, tzn. $\widehat{w}_F(\lambda) = 0$.

460. Zatem jeśli $w_F(\lambda) = (\lambda_1 - \lambda)^{k_1} \ldots (\lambda_r - \lambda)^{k_r}$, to wielomian minimalny ma postać

$$\widehat{w}_F(\lambda) = (\lambda - \lambda_1)^{h_1} \ldots (\lambda - \lambda_r)^{h_r}, \qquad \text{gdzie } 1 \leqslant h_i \leqslant k_i.$$

Wykładniki k_i są *krotnościami* wartości własnych λ_i; jak zobaczymy niebawem (zob. 488), wykładniki h_i są tzw. *wysokościami* przestrzeni pierwiastkowych $V(\lambda_i)$.

8.6 Funkcje od operatora

Wiemy już, co to jest $\varphi(F)$, gdy $F \in \mathrm{End}\, V$, a $\varphi(\lambda)$ jest wielomianem. Zobaczymy teraz, jak można uogólnić definicję $\varphi(F)$ na większą od wielomianów klasę funkcji. Założymy tu, że $\mathbb{K} = \mathbb{C}$ albo $\mathbb{K} = \mathbb{R}$, aby można się było zajmować funkcjami analitycznymi, tzn. rozwijalnymi w szereg potęgowy. Jeśli $\mathbb{K} = \mathbb{R}$, to o takiej funkcji $\varphi : \mathcal{O} \to \mathbb{C}$ założymy, że jest *rzeczywista*([84]): $\forall \lambda \in \mathcal{O} : \; (\overline{\lambda} \in \mathcal{O}$ oraz $\varphi(\overline{\lambda}) = \overline{\varphi(\lambda)})$.

461. **Lemat 1** (*kryterium podzielności funkcji analitycznej przez wielomian*). Niech $w(\lambda) = (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_r)^{k_r}$, gdzie $k_i \in \mathbb{N}$, a $\lambda_i \in \mathbb{C}$ są parami różne; niech ponadto $\varphi : \mathcal{O} \to \mathbb{C}$ będzie funkcją analityczną na otwartym podzbiorze $\mathcal{O} \subset \mathbb{C}$, zawierającym $w^{-1}\{0\} = \{\lambda_1, \ldots, \lambda_r\}$. Wtedy następujące warunki są równoważne:

(1) Istnieje funkcja analityczna $\sigma : \mathcal{O} \to \mathbb{C}$, taka że $\varphi = \sigma w$;
(2) $\forall i \in \overline{1, r} : \forall k \in \overline{0, k_i - 1} : \; \varphi^{(k)}(\lambda_i) = 0$. ([85])

[84] Warto sprawdzić, że dla wielomianu taki warunek faktycznie oznacza, że $\varphi \in \mathbb{R}[\,\cdot\,]$.
[85] Czyli każde zero dla w jest zerem dla φ, z krotnością co najmniej taką, jak dla w.

$\boxed{(1)\Rightarrow(2)}$ b. proste: skoro $w^{(j)}(\lambda_i)=0$ dla $j<k_i$, to $\left[\sigma w\right]^{(k)}=\sum\limits_{j\leqslant k}\binom{k}{j}\sigma^{(k-j)}w^{(j)}$ znika w $\lambda=\lambda_i$ dla $k<k_i$. $\boxed{(2)\Rightarrow(1)}$: Na otoczeniu λ_1 funkcja φ ma rozwinięcie $\varphi(\lambda)=\sum\limits_{k=0}^{\infty}c_k(\lambda-\lambda_1)^k$, przy czym $c_0=\ldots=c_{k_1-1}=0$; wobec tego funkcja $\frac{\varphi(\lambda)}{w(\lambda)}=\frac{1}{(\lambda-\lambda_2)^{k_2}\ldots(\lambda-\lambda_r)^{k_r}}\sum\limits_{k=0}^{\infty}c_{k+k_1}(\lambda-\lambda_1)^k$ na otoczeniu λ_1 jest analityczna([86]).

Przy spełnieniu (1),(2) mówimy, że φ *jest podzielna przez* w, pisząc

$$w\mid\phi.$$

462. **Lemat 2** (*reszta z dzielenia funkcji przez wielomian*). Jeśli $w\in\mathbb{C}[\,\cdot\,]$ jest niezerowy, zaś $\varphi:\mathcal{O}\to\mathbb{C}$ jest funkcją analityczną na otoczeniu zbioru $w^{-1}\{0\}$, to istnieje dokładnie jeden wielomian $\varrho\in\mathbb{C}[\,\cdot\,]$, taki że

$$w\mid(\varphi-\varrho)\ \text{ oraz }\ \deg\varrho<d:=\deg w;$$

zatem istnieje funkcja analityczna $\sigma:\mathcal{O}\to\mathbb{C}$ taka, że $\varphi=\sigma w+\varrho$. Jeśli oprócz tego φ jest rzeczywista, a $w\in\mathbb{R}[\,\cdot\,]$, to również $\varrho\in\mathbb{R}[\,\cdot\,]$.

Niech $w(\lambda)=(\lambda-\lambda_1)^{k_1}\ldots(\lambda-\lambda_r)^{k_r}$, gdzie λ_i parami różne, $k_i\in\mathbb{N}$; wtedy $w^{-1}\{0\}=\{\lambda_1,\ldots,\lambda_r\}$. Można sprawdzić, że formy liniowe $v\mapsto v^{(k)}(\lambda_i)$ na przestrzeni $W=\mathbb{K}_{d-1}[\,\cdot\,]$, dla $i\in\overline{1,r}$, $k\in\overline{0,k_i-1}$, są l. niezależne; skoro jest ich $\sum k_i=d=\dim W$, więc tworzą bazę W^*. Zatem istnieje jedyny $\varrho\in W$, przyjmujący zadane dowolnie wartości $\varrho^{(k)}(\lambda_i)=C_i^k$; dla $C_i^k:=\varphi^{(k)}(\lambda_i)$ mamy $w\mid(\varphi-\varrho)$.

Funkcja $\varphi^*(\lambda):=\overline{\varphi(\overline{\lambda})}$ też jest analityczna, a rzeczywistość φ i w oznacza $\varphi^*=\varphi$, $w^*=w$; stąd $\varphi=(w\sigma+\varrho)^*=w\sigma^*+\varrho^*$, więc $\varrho^*=\varrho$ dzięki jednoznaczności reszty.

463. Wielomian, o którym mowa w lemacie, nazywa się *resztą z dzielenia* φ *przez* w albo *wielomianem interpolacyjnym dla* φ *względem* w; będziemy go zwykle oznaczać symbolem $\tilde{\varphi}$. Jak widać z dowodu lematu, jest on niezależny od wyboru otoczenia $\mathcal{O}$ i określony jednoznacznie warunkami

$$\forall i\in\overline{1,r}:\forall k\in\overline{0,k_i-1}:\ \tilde{\varphi}^{(k)}(\lambda_i)=\varphi^{(k)}(\lambda_i)\ \text{ oraz }\ \deg\tilde{\varphi}<\deg w.$$

464. **Definicja** (*funkcja od operatora* $F\in\operatorname{End}V$). Niech $\varphi:\mathcal{O}\to\mathbb{C}$ będzie funkcją analityczną na pewnym otoczeniu $\mathcal{O}$ zbioru

$$\operatorname{Sp}_{\mathbb{C}}F:=(w_F)^{-1}\{0\}=\{\lambda\in\mathbb{C}:\ w_F(\lambda)=0\};$$

dla $\mathbb{K}=\mathbb{R}$ zakładamy przy tym, że φ jest rzeczywista: $\varphi(\overline{\lambda})=\overline{\varphi(\lambda)}$. Wówczas

$$\boxed{\varphi(F):=\tilde{\varphi}(F)},\qquad(\star)$$

gdzie $\tilde{\varphi}\in\mathbb{K}[\,\cdot\,]$ jest resztą z dzielenia φ przez dowolny wielomian w, zerujący operator F: $\boxed{w(F)=0}$, co oznacza (zob. 458), że $\boxed{\widehat{w}_F\mid w}$.

Są dwa *praktyczne* powody, dla których będziemy dodatkowo wymagać, by $\boxed{w\mid w_F}$:
(1) stopień $d=\deg w$ jest liczbą niewiadomych (tj. współczynników wielomianu $\tilde{\varphi}$), a także liczbą równań (tj. warunków $\tilde{\varphi}^{(k)}(\lambda_i)=\ldots$), które musimy rozwiązać,

[86] Korzystamy tu z następujących rezultatów kursu analizy: (1) analityczność jest własnością lokalną; (2) iloczyn funkcji analitycznych jest funkcją analityczną; (3) odwrotność funkcji analitycznej o wartościach $\neq 0$ jest funkcją analityczną. Wynika z tego także i to, że iloraz $\frac{\varphi}{w}$ jest f. analityczną na $\mathcal{O}\setminus\{\lambda_1,\ldots,\lambda_r\}$, więc tylko punkty λ_i stanowią problem.

szukając wielomianu interpolacyjnego $\tilde\varphi$; zatem nakład pracy rośnie wraz z d.

(2) w definicji reszty i ilorazu z dzielenia φ przez w przyjęliśmy (dla uproszczenia), że zbiór $w^{-1}\{0\}$ musi się zawierać w dziedzinie funkcji analitycznej φ.

465. **Fakt**. Definicja $(\star)$ jest sensowna, tzn. rezultat nie zależy od wyboru w.

Niech $\tilde\varphi$ i $\tilde\varphi_1$ będą resztami z dzielenia φ przez wielomiany w i w_1, takie że $w(F) = w_1(F) = 0$; wtedy $\tilde\varphi_1 - \tilde\varphi = (\varphi - \tilde\varphi) - (\varphi - \tilde\varphi_1) = \sigma w - \sigma_1 w_1$, przy czym w i w_1 są podzielne przez $\widehat{w}_F$; zatem $\widehat{w}_F \mid (\tilde\varphi_1 - \tilde\varphi)$, a więc wielomian $\tilde\varphi_1 - \tilde\varphi$ zeruje F, QED.

466. **Uwaga**. Zauważmy, że jeśli $k_1 = \ldots = k_r = 1$, tzn. wielomian $\widehat{w}_F$ nie ma pierwiastków wielokrotnych, to $\varphi(F)$ zależy jedynie od wartości (a nie pochodnych) φ w punktach zbioru $\operatorname{Sp}_{\mathbb{C}} F$; w tym przypadku mamy więc implikację

$$\Big(\psi - \varphi \text{ zeruje się na } \operatorname{Sp}_{\mathbb{C}} F\Big) \;\Rightarrow\; \psi(F) = \varphi(F);$$

dzięki temu można operator $\varphi(F)$ zdefiniować nie tylko dla analitycznej, ale wręcz dla *każdej* (rzeczywistej dla $\mathbb{K} = \mathbb{R}$) funkcji $\varphi : \operatorname{Sp}_{\mathbb{C}} F \to \mathbb{C}$.

Kiedy ten pożyteczny warunek $k_1 = \ldots = k_r = 1$ jest spełniony? Na przykład, gdy F jest *diagonalizowalny*, tzn. V ma bazę złożoną z wektorów własnych F (jest tak m.in. wtedy, gdy wielomian w_F ma $n = \dim V$ pierwiastków w $\mathbb{K}$, zob. 438)([87])

467. **Własności funkcji od operatora.**

(1) Jeśli $\varphi \in \mathbb{K}[\,\cdot\,]$, to *tak zdefiniowane $\varphi(F)$ jest 'zwykłą' ewaluacją wielomianu na operatorze*, gdyż dla wielomianów $\big[\varphi = \sigma w + \tilde\varphi,\ w(F) = 0\big]$ implikuje $\varphi(F) = \tilde\varphi(F)$.

(2) Ogólniej, *operacjom algebraicznym na funkcjach odpowiadają operacje na operatorach:* $\varphi = \varphi_1 + \varphi_2 \;\Rightarrow\; \varphi(F) = \varphi_1(F) + \varphi_2(F)$ (oczywista liniowość $\varphi \mapsto \tilde\varphi$), $\varphi = \varphi_1 \cdot \varphi_2 \;\Rightarrow\; \varphi(F) = \varphi_1(F) \circ \varphi_2(F)$; istotnie, jest tak, jak wiemy, dla wielomianów, więc skoro $\tilde\varphi_1\tilde\varphi_2 - \varphi = (\tilde\varphi_1 - \varphi_1)\varphi_2 + \varphi_1(\tilde\varphi_2 - \varphi_2)$ jest podzielna przez wielomian w, zerujący F, więc $\tilde\varphi_1(F)\tilde\varphi_1(F) - \varphi(F) = 0$, zaś $\tilde\varphi_i(F) = \varphi_i(F)$.

(3) *Jeśli podprzestrzeń $U \subset V$ jest F-niezmiennicza: $FU \subset U$, to także $\varphi(F)U \subset U$.* Istotnie, zbiór $\mathcal{A}_U = \{G \in \operatorname{End} V : GU \subset U\}$ jest zamkniety względem operacji algebraicznych, więc $\varphi(F)$, będąc pewnym wielomianem od $F \in \mathcal{A}_U$, należy do $\mathcal{A}_U$.

(4) *Jeśli $\varphi(\lambda) = \frac{L(\lambda)}{M(\lambda)}$, gdzie funkcje L, M są analityczne na otoczeniu $\operatorname{Sp} F$ (np. są wielomianami), a M nie zeruje się na $\operatorname{Sp} F$, to $M(F)$ jest odwracalny oraz*

$$\varphi(F) = L(F)\big(M(F)\big)^{-1} = \big(M(F)\big)^{-1}L(F).$$

Istotnie, z $\varphi M = L$ dzięki (2) wynika $M(F)\varphi(F) = L(F) = \varphi(F)M(F)$, więc pozostaje wykazać istnienie $M(F)^{-1}$. Otóż $M(F) = \tilde M(F)$, gdzie wielomian $\tilde M$, interpolujący M, pokrywa się na $\operatorname{Sp} F$ z M, więc nie znika. Zatem wielomiany w_F i $\tilde M$ są względnie pierwsze, skąd $\exists p, q \in \mathbb{K}[\,\cdot\,] : 1 = p\tilde M + q w_F$, co wraz z tw. Cayleya-Hamiltona daje $\mathrm{id}_V = p(F)M(F)$, tzn. odwracalność $M(L)$.

(5) *Jeśli $Fv = \mu v$ dla pewnego wektora $v \in V$ i liczby $\mu \in \mathbb{K}$, to $\varphi(F)v = \varphi(\mu)v$.* Istotnie, jest tak, jak wiemy, dla wielomianów, więc $\varphi(F)v = \tilde\varphi(F)v = \tilde\varphi(\mu)v$, a ponadto przy $v \neq 0$ mamy $\mu \in \operatorname{Sp} F$, skąd $\tilde\varphi(\mu) = \varphi(\mu)$, QED. W konsekwencji

[87] Okazuje się, że $(\widehat{w}_F$ nie ma wielokrotnych pierwiastków $) \Leftrightarrow F$ jest diagonalizowalny lub — w przypadku $\mathbb{K} = \mathbb{R}$ — diagonalizowalna jest tzw. *kompleksyfikacja* operatora F.

(5') *Gdy V ma bazę wektorów własnych F: $Fe_i = \lambda_i e_i$, tj.* $[F]^e_e = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, *wtedy*

$$[\varphi(F)]^e_e = \operatorname{diag}\big(\varphi(\lambda_1), \ldots, \varphi(\lambda_n)\big).$$

(6) Uogólnienie: *Jeśli $v \in V(\mu, F)$, a więc $\exists h \in \mathbb{N}: (F_\mu)^h v = 0$* (zob. 441), *to wtedy*

$$\varphi(F)v = \sum_{0 \leqslant k < h} \frac{\varphi^{(k)}(\mu)}{k!} (F_\mu)^k v.$$

Dowód: Skoro $\tilde\varphi(\lambda) = \sum_{k=0}^{N} \frac{\tilde\varphi^{(k)}(\mu)}{k!}(\lambda-\mu)^k$, to $\varphi(F)v = \tilde\varphi(F)v = \sum_{k=0}^{N} \frac{\tilde\varphi^{(k)}(\mu)}{k!}(F_\mu)^k v$, gdzie $N = \deg\tilde\varphi$; stąd teza, gdyż $\forall k \geqslant 0 : \big[\, \tilde\varphi^{(k)}(\mu) = \varphi^{(k)}(\mu)$ lub $(F_\mu)^k v = 0 \,\big]$.

(7) *Jeśli $\varphi(\lambda) = \sum_{k=-\infty}^{\infty} c_k(\lambda - \lambda_0)^k$, przy czym szereg ten jest zbieżny na* $\operatorname{Sp} F$, *to szereg $\sum_{k=-\infty}^{\infty} c_k (F_{\lambda_0})^k$ jest zbieżny punktowo* (tzn. na każdym $v \in V$) *do $\varphi(F)$.*

Niech $\varphi_m(\lambda) := \sum_{k=-m}^{m} c_k(\lambda - \lambda_0)^k$. Jeśli v jest wektorem własnym, $Fv = \mu v$, $\mu \in \operatorname{Sp} F$, to z (5) mamy $\varphi_m(F)v = \varphi_m(\mu)v \overset{m\to\infty}{\longrightarrow} \varphi(\mu)v = \varphi(F)v$. Podobnie jest w sytuacji opisanej w (6). W ogólnym przypadku v jest sumą takich uogólnionych wektorów własnych (patrz dalej tw. o bazie jordanowskiej), skąd teza.

468. **Przykład.** Jeśli $\varphi(\lambda) = e^{t\lambda}$, gdzie $t \in \mathbb{R}$ jest ustalone, to $\varphi(F) = e^{tF}$, przy czym $e^{tF}v = \sum_{k=0}^{\infty} \frac{t^k}{k!} F^k v$, więc $\frac{d}{dt} e^{tF} = Fe^{tF} = e^{tF}F$. Podobnie dowodzimy, że $e^{F_1+F_2} = e^{F_1}e^{F_2}$, jeśli $F_1F_2 = F_2F_1$, więc w szczególności

$$e^{(t+s)F} = e^{tF}e^{sF}, \quad \left(e^{tF}\right)^{-1} = e^{-tF}.$$

469. **Ćwiczenie.** Sprawdzić, że: (1) $\varphi(AFA^{-1}) = A\,\varphi(F)\,A^{-1}$ dla $A : V \overset{\cong}{\longrightarrow} W$; (2) $\varphi(F^T) = \big(\varphi(F)\big)^T$.

8.7 Operatory rzutowe

W tym paragrafie nie musimy zakładać, że przestrzeń V ma skończony wymiar.

470. **Definicja.** Operator $P \in \operatorname{End} V$ nazywa się *rzutem* (albo *operatorem rzutowym*, albo *projektorem*), jeśli spełnia warunek $P^2 := P \circ P = P$.

471. **Definicja.** Niech $V = V_0 \dot+ V_1$. Jeśli $v = v_0 + v_1$ jest rozkładem wektora $v \in V$ na składowe $v_i \in V_i$, to v_1 nazywamy *rzutem v na V_1 wzdłuż* (albo *równolegle do*) V_0 i oznaczamy symbolem $P^{V_0}_{V_1}(v)$. Zauważmy, że

$$P^{V_0}_{V_1}(v) = \begin{pmatrix} \text{taki (określony jednoznacznie!)} \\ \text{wektor } v_1 \in V_1, \text{ że } v - v_1 \in V_0 \end{pmatrix}.$$

Pokażemy teraz, że operator jest rzutem $\iff$ jest 'rzutem na coś wzdłuż czegoś':

472. **Fakt.** (a) Jeśli $V = V_0 \dot+ V_1$, to $P^{V_0}_{V_1}$ należy do $\operatorname{End} V$ i jest rzutem.

(b) Jeśli $P \in \operatorname{End} V$ jest rzutem, to $V = \ker P \dot+ \operatorname{im} P$ oraz $P = P^{\ker P}_{\operatorname{im} P}$.

(a) Oznaczmy dla wygody $P := P^{V_0}_{V_1}$.
Liniowość P: dla sprawdzenia, że $P(v' + v'') = P(v') + P(v'')$ należy pokazać, że

$$(1)\ P(v')+P(v'')\in V_1 \quad\text{oraz}\quad (2)\ v_0:=(v'+v'')-\big(P(v')+P(v'')\big)\in V_0;$$

otóż (1) wynika stąd, że $P(v'),P(v'')\in V_1$, a V_1 jest podprzestrzenią, zaś (2) stąd, że v_0 jest sumą wektorów $v'-P(v')$ i $v''-P(v'')$ należących do podprzestrzeni V_0. Podobnie równość $P(\lambda v)=\lambda P(v)$ wynika stąd, że $\lambda P(v)\in V_1$ oraz $\lambda v-\lambda P(v)\in V_0$. Dla dowodu równości $P^2=P$ zauważmy, że jeśli $v\in V_1$, to warunki $\left\{\begin{array}{c} v_1\in V_1 \\ v-v_1\in V_0\end{array}\right\}$ są spełnione dla $v_1:=v$, a więc $\boxed{\forall v\in V_1: P(v)=v}$. Stąd dla każdego $v\in V$, dzięki temu że $P(v)\in V_1$, mamy $P\big(P(v)\big)=P(v)$, a zatem $P=P^{V_0}_{V_1}$ jest rzutem.

(b) Zauważmy, że $v_1\in V_1\Rightarrow\exists v: v_1=P(v)\Rightarrow P(v_1)=P^2(v)=P(v)=v_1$, więc P jest identycznością na $V_1:=\operatorname{im}P$. Stąd jeśli $v\in V$ ma rozkład $v=v_0+v_1$, gdzie $v_0\in V_0:=\ker P$, $v_1\in V_1$, to $P(v)=P(v_0)+P(v_1)=0+v_1$, czyli $\left\{\begin{array}{c} v_1=P(v) \\ v_0=v-P(v)\end{array}\right\}$, więc rozkład może być co najwyżej jeden. Z drugiej strony $v_1:=P(v)\in V_1$, zaś $v_0:=v-P(v)\in V_0$, gdyż $P(v_0)=P(v)-P^2(v)=0$, więc każdy $v\in V$ ma taki rozkład, co dowodzi że $V=V_0\dot{+}V_1$. Ponadto $P^{V_0}_{V_1}(v)=v_1=P(v)$, więc $P=P^{V_0}_{V_1}$.

Najważniejsze własności operatora rzutowego zestawia następujący

473. **Fakt**. Niech $P\in\operatorname{End}V$ będzie rzutem, $V_0:=\ker P$, $V_1:=\operatorname{im}P$. Wtedy

$$P\big|_{V_0}=0;\qquad P\big|_{V_1}=\mathrm{id};\qquad \ker(\mathrm{id}_V-P)=V_1;\qquad \operatorname{im}(\mathrm{id}_V-P)=V_0.$$

Jeśli oprócz powyższych założeń $\dim V<\infty$, to

$$\exists e\ (\text{baza}):\ [P]^e_e=\left[\begin{array}{c|c} \boldsymbol{I}_r & 0 \\ \hline 0 & 0\end{array}\right],\qquad \operatorname{tr}P=r:=\operatorname{rk}P\in\mathbb{Z}_+.$$

Dwie pierwsze własności już znamy; zatem $v\in\ker(\mathrm{id}-P)\ \Leftrightarrow\ v=P(v)\ \Leftrightarrow\ v\in V_1$. Skoro $P(\mathrm{id}-P)=P-P^2=0$, to $\operatorname{im}(\mathrm{id}-P)\subset\ker P=V_0$; odwrotnie, $v\in V_0$, tzn. $P(v)=0$, implikuje $v=(\mathrm{id}-P)(v)\in\operatorname{im}(\mathrm{id}-P)$. Bazą o własności $[P]^e_e=\left[\begin{array}{c|c} \boldsymbol{I}_r & 0 \\ \hline 0 & 0\end{array}\right]$ jest konkatenacja dowolnych baz dla V_1 i V_0; stąd także wynika ostatnia własność.

474. Zauważmy, że jeśli $V=V_0\dot{+}V_1$, to rzuty $P_0:=P^{V_1}_{V_0}$ i $P_1:=P^{V_0}_{V_1}$ spełniają następujące warunki:

$$P_0^2=P_0,\ \ P_1^2=P_1,\ \ P_0P_1=P_1P_0=0,\ \ P_0+P_1=\mathrm{id}_V;$$

uogólnimy to teraz na przypadek sumy prostej $r\geqslant 2$ podprzestrzeni.

475. **Fakt**. (a) Niech $V=V_1\dot{+}\ldots\dot{+}V_r$ i niech $P_1(v)\in V_1,\ldots,P_r(v)\in V_r$ będą składowymi wektora $v\in V$. Wtedy $P_1,\ldots,P_r$ są operatorami z $\operatorname{End}V$, spełniającymi następujące warunki:

$$\forall i: P_i^2=P_i,\qquad \forall i\neq j: P_j\circ P_i=0,\qquad P_1+\ldots+P_r=\mathrm{id}_V. \qquad (*)$$

(b) Odwrotnie, jeśli dane są operatory $P_1,\ldots,P_r\in\operatorname{End}V$ spełniające warunki $(*)$, to mamy rozkład $V=V_1\dot{+}\ldots\dot{+}V_r$, gdzie $V_i:=\operatorname{im}P_i$; nadto $P_i(v)$ są składowymi $v\in V$ odpowiadającymi temu rozkładowi.

(a) Ustalmy chwilowo $i\in\overline{1,n}$; jeśli $v\in V_i$, to rozkładem v jest $0+\ldots+v+\ldots+0$, a więc $P_i(v)=v$ oraz $P_j(v)=0$ dla $j\neq i$. Stąd $P_i\big|_{V_i}=\mathrm{id}$, $P_j\big|_{V_i}=0$, więc skoro $\forall v\in V: P_i(v)\in V_i$, to $P_i\circ P_i(v)=P_i(v)$, $P_j\circ P_i(v)=0$. Ponadto każdy v jest sumą swych składowych $P_i(v)$, więc $P_1(v)+\ldots+P_r(v)=v$.

(b) Skoro $v=(P_1+\ldots+P_r)(v)=P_1(v)+\ldots+P_r(v)$ oraz $P_i(v)\in\operatorname{im}P_i=V_i$,

to każdy wektor **ma** rozkład. Sprawdzimy, że jest to **jedyny** rozkład: Jeśli $v = v_1 + \ldots + v_r$, gdzie $v_i \in V_i$, to $P_i(v_i) = v_i$ (gdyż $\exists w : v_i = P_i(w)$, zaś $P_i^2 = P_i$) oraz $P_j(v_i) = 0$ dla $i \neq j$ (gdyż $P_j\big(P_i(w)\big) = 0$), więc $P_1(v) = P_1(v_1 + v_2 + \ldots) = = P_1(v_1) + P_1(v_2) + \ldots = v_1 + 0 + \ldots = v_1$, itd., co kończy dowód.

476. Oczywiście w powyższej sytuacji P_i jest rzutem na V_i wzdłuż podprzestrzeni $V^i := V_1 + \ldots(\text{bez } V_i)\ldots + V_r$, tzn. $P_i = P_{V_i}^{V^i}$. W takim razie np. P_1 zależy nie tylko od przestrzeni V_1, na którą się rzutuje, ale i od pozostałych przestrzeni $V_2, \ldots, V_r$, poprzez ich sumę algebraiczną V^1, określającą 'kierunek' rzutowania.

477. **Fakt**. Jeśli spełnione są powyższe warunki (a),(b) oraz $F \in \operatorname{End} V$, to

$$\left(\begin{array}{c} V_i \text{ są } F\text{-niezmiennicze,} \\ \text{tzn. } \forall i : F(V_i) \subset V_i \end{array}\right) \Leftrightarrow \left(\begin{array}{c} F \text{ komutuje z } P_1, \ldots, P_r, \\ \text{tzn. } \forall i : FP_i = P_iF \end{array}\right).$$

$\boxed{\Rightarrow}$ Dla $v \in V$ mamy $F(v) = \sum\limits_j P_jF(v)$, a także $F(v) = F\big(\sum\limits_j P(v)\big) = \sum\limits_j FP_j(v)$. Przy tym zarówno $P_jF(v) \in V_j$, jak również $FP_j(v) \in V_j$ (gdyż $F(V_j) \subset V_j$), więc z jednoznaczności rozkładu na składowe wynika $P_jF(v) = FP_j(v)$. *Krótszy sposób*:

$\boxed{\Rightarrow}$ Skoro $w_j := FP_j(v)$ należy do $F(V_j) \subset V_j$, to $P_j(w_j) = w_j$, więc z dowolności $v \in V$ mamy $P_jFP_j = FP_j$. Stąd $P_iF = P_i\sum\limits_j FP_j = P_i\sum\limits_j P_jFP_j = P_iFP_i = FP_i$.

$\boxed{\Leftarrow}$ Jeśli $v \in V_i$, to $v = P_i(v)$, więc $F(v) = FP_i(v) = P_iF(v) \in V_i$; stąd $F(V_i) \subset V_i$.

478. **Uwaga**. W najprostszym przypadku $r = 2$ wynik ten oznacza, że jeśli $F, P \in \operatorname{End} V$ oraz P jest rzutem, to $FP = PF \iff \left\{\begin{array}{l} F(\operatorname{im} P) \subset \operatorname{im} P \\ F(\ker P) \subset \ker P \end{array}\right\}$; ostatni warunek jest oczywiście **mocniejszy**, niż samo zawieranie $F(\operatorname{im} P) \subset \operatorname{im} P$.

8.8 Operatory nilpotentne

479. **Definicja**. Operator $N \in \operatorname{End} V$ jest *nilpotentny*, jeśli istnieje liczba naturalna k taka, że $N^k = 0$.

Czy operator $N \in \operatorname{End}\mathbb{R}^2$, określony macierzą $\boldsymbol{N} := \begin{bmatrix} 2 & 5 \\ -1 & -3 \end{bmatrix}$ jest nilpotentny? Kolejnymi potęgami macierzy $\boldsymbol{N}$ są $\begin{bmatrix} -1 & -5 \\ 1 & 4 \end{bmatrix}$, $\begin{bmatrix} 3 & 10 \\ -2 & -7 \end{bmatrix}$, $\begin{bmatrix} -4 & -15 \\ 3 & 11 \end{bmatrix}$, $\begin{bmatrix} 7 & 25 \\ -5 & -18 \end{bmatrix}$, $\begin{bmatrix} -11 & -40 \\ 8 & 29 \end{bmatrix}$, $\begin{bmatrix} 18 & 65 \\ -13 & -47 \end{bmatrix}$, $\begin{bmatrix} -29 & -105 \\ 21 & 76 \end{bmatrix}$, $\begin{bmatrix} 47 & 170 \\ -34 & -123 \end{bmatrix}$, $\begin{bmatrix} -76 & -275 \\ 55 & 199 \end{bmatrix}$; nie da się na podstawie tych wyliczeń wykluczyć, że któraś z następnych potęg okaże się zerem.

480. **Fakt**. Jeśli $\dim V < \infty$ oraz $N \in \operatorname{End} V$, to następujące warunki są równoważne:

1. N jest operatorem nilpotentnym;
2. Istnieje ciąg podprzestrzeni $\{0\} = V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_n = V$, taki że $\dim V_k = k$ oraz $N(V_k) \subset V_{k-1}$ dla $k \in \overline{1,n}$;
3. V ma bazę, w której N wyraża się macierzą ściśle (górno-)trójkątną;
4. $w_N(\lambda) = (-\lambda)^n$, $n := \dim V$;
5. $\forall k : \tau_k(N) = 0$ (tzn. wszystkie niezm. podstawowe N znikają);
6. $N^n = 0$, gdzie $n := \dim V$.

$\boxed{1. \Rightarrow 2.}$ Niech $V_n := V$. Gdyby N był surjektywny, wtedy każda potęga N też byłaby surjektywna, co jest sprzeczne z nilpotentnością ($\exists\ k : N^k = 0$); zatem

$N(V_n) = \operatorname{im} N$ ma wymiar $\leqslant n-1$, więc istnieje $(n-1)$-wymiarowa podprzestrzeń $V_{n-1} \subset V_n$ taka, że $N(V_n) \subset V_{n-1}$. Oczywiście V_{n-1} jest N-niezmiennicza: $N(V_{n-1}) \subset N(V_n) \subset V_{n-1}$, więc możemy powtórzyć nasz zabieg, zastępując (V_n, N) parą $(V_{n-1}, N|_{V_{n-1}})$, itd. W ten sposób znajdziemy ciąg V_k o żądanych własnościach. $\boxed{2. \Rightarrow 3.}$ Dla $k \in \overline{1,n}$ wybierzmy dowolny wektor $e_k \in V_k \setminus V_{k-1}$; dostajemy w ten sposób bazę V([88]), przy czym $N(e_k) \in \langle e_1, \ldots, e_{k-1}\rangle$, co oznacza, że macierz $[N]^e_e$ jest ściśle górno-trójkątna. $\boxed{3. \Rightarrow 4.}$ Z tw. o wyznaczniku macierzy trójkątnej. $\boxed{4. \Leftrightarrow 5.}$ Wprost z określenia liczb $\tau_k(N)$ jako współczynników wielomianu $w_N(\lambda)$. $\boxed{4. \Rightarrow 6.}$ Wprost z tw. Cayleya-Hamiltona. $\boxed{6. \Rightarrow 1.}$ Z definicji nilpotentności.

Uwaga. Jeśli ciało $\mathbb{K}$ jest algebraicznie domknięte, np. $\mathbb{K} = \mathbb{C}$, to 4. $\Leftrightarrow$ $\overline{4}$., gdzie

$\overline{4}$. $\operatorname{Sp} N = \{0\}$;

istotnie, wielomian $w_N(\lambda)$ ma wtedy rozkład postaci $\prod_{i=1}^{n} (\lambda_i - \lambda)$ w pierścieniu $\mathbb{K}[\,\cdot\,]$. Bez algebraicznej domkniętości, np. dla $\mathbb{K} = \mathbb{R}$, mamy wciąż wynikanie 4. $\Rightarrow$ $\overline{4}$., lecz $\overline{4}$. $\Rightarrow$ 4., a więc i $\overline{4}$. $\Rightarrow$ 1., może być fałszywe, czego dowodzi kontrprzykład

$$V := \mathbb{R}^3, \quad N \in \operatorname{End}\mathbb{R}^3, \quad N\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) := \begin{bmatrix} -x_2 \\ x_1 \\ 0 \end{bmatrix};$$

mamy tutaj $\operatorname{Sp} N = \{0\}$, lecz $0 \neq N = N^5 = N^9 = \ldots$ oraz $w_N(\lambda) = -\lambda(\lambda^2+1)$.

Ćwiczenie. Znaleźć bezpośrednie dowody implikacji 1. $\Rightarrow$ $\overline{4}$. oraz 1. $\Rightarrow$ 3. $\Rightarrow$ 2. $\Rightarrow$ 6.

Rozwiązanie.

$\boxed{1. \Rightarrow \overline{4}.}$ Jeśli $N^k = 0$ oraz $\exists v \neq 0 : N(v) = \lambda v$, to $0 = N^k(v) = \lambda^k v$, skąd $\lambda = 0$. $\boxed{1. \Rightarrow 3.}$ Przypuśćmy, że znaleźliśmy $e_1, \ldots, e_k \in V$, takie że $V_k := \langle e_1, \ldots, e_k\rangle$ jest k-wymiarowa oraz $N(V_k) \subset V_{k-1}$. Skoro $\forall^* p : \operatorname{im}(N^p) \subset V_k$, to ma sens określenie $q := \min\{p \geqslant 1 : \operatorname{im}(N^p) \subset V_k\}$; wtedy $\exists v \in V : N^{q-1}(v) \notin V_k$, wobec tego $e_{k+1} := N^{q-1}(v)$ ma własności $e_{k+1} \notin V_k$ i $N(e_{k+1}) \in V_k$, więc $N(V_{k+1}) \subset V_k$. Oczywiście konstrukcję rozpocząć musimy od znalezienia e_1, przyjmując $V_0 := \{0\}$. $\boxed{3. \Rightarrow 2.}$ Jeśli $[N]^e_e$ jest ściśle trójkątna, to $V_k := \langle e_1, \ldots, e_k\rangle$ ma żądane własności. $\boxed{2. \Rightarrow 6.}$ Skoro $N(V_k) \subset V_{k-1}$, to $N^2(V_k) \subset V_{k-2}$, itd.; ogólnie $N^p(V_k) = V_{k-p}$, jeśli oznaczymy $V_j := \{0\}$ dla $j \leqslant 0$. Zatem $N^n(V_n) \subset V_0$, tzn. $N^n = 0$, QED.

481. **Definicja**. Jeśli $N \in \operatorname{End} V$ jest operatorem nilpotentnym, to *N-serią długości $k \geqslant 1$* nazywamy ciąg niezerowych wektorów z V postaci

$$v_1,\ v_2 = N(v_1),\ \ldots,\ v_k = N(v_{k-1}),$$

taki że $N(v_k) = 0$, tzn. nie dający się 'przedłużyć w prawo' (lub 'w dół', gdy na diagramie umieszczamy v_2 pod v_1, v_3 pod v_2 itd.); jasne, że ten warunek gwarantuje N-niezmienniczość podprzestrzeni $\langle v_1, \ldots, v_k\rangle$.

Ćwiczenie. Dowieść, że N-seria jest układem liniowo niezależnym, a więc bazą $W := \langle v_1, \ldots, v_k\rangle$. Napisać macierz $\left[N|_W\right]^e_e$, gdy bazą e jest $v_1, \ldots, v_k$ lub $v_k, \ldots, v_1$.

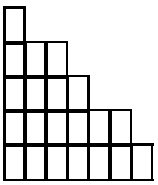
482. **Twierdzenie**. Jeśli operator $N \in \operatorname{End} V$ jest nilpotentny, to przestrzeń V ma bazę będącą zestawem (tj. konkatenacją) pewnej liczby N-serii.

[88] Istotnie, niech $\lambda_1 e_1 + \ldots \lambda_n e_n = 0$; przy $\lambda_n \neq 0$ wynika stąd, że $e_n \in \langle e_1, \ldots, e_{n-1}\rangle$, czyli $e_n \in V_{n-1}$, wbrew założeniu; zatem $\lambda_n = 0$ oraz $\lambda_1 e_1 + \ldots \lambda_{n-1} e_{n-1} = 0$, więc możemy powtórzyć argumentację, dowodząc że $\lambda_{n-1} = 0$, i tak dalej.

Dowód (niezbyt trudny, będący konstrukcją takiej bazy) podamy w Appendiksie A.

Bazę zestawioną z N-serii nazywamy *bazą jordanowską* operatora N.

483. **Fakt**. Zbiór długości (a więc i liczba) wszystkich N-serii, z jakich składa się dana baza jordanowską N, jest dla wszystkich baz jordanowskich jednakowy, a więc zależy tylko od N (jest *niezmiennikiem operatora*).



Wygodnie jest bazę jordanowską obrazować jej *diagramem*: pola diagramu odpowiadają wektorom bazy, kolumny — N-seriom tworzącym tę bazę, przy czym obowiązują tu dwie konwencje:

(1) (pole wektora v leży nad polem wektora v') $\iff N(v) = v'$;
(2) pola odpowiadające końcom serii leżą w jednym (najniższym) wierszu; będziemy ponadto ustawiać kolumny wg ich długości.

Przedstawiony na rysunku diagram odpowiada bazie, zestawionej z serii o długościach 5, 4, 4, 3, 2, 2, 1. Zauważmy, że wektory takiej bazy, odpowiadające najniższemu wierszowi, rozpinają $\ker N$; z kolei $\ker N^2$ rozpinają wektory z dwóch najniższych wierszy, $\ker N^3$ — wektory z trzech najniższych, itd. Zatem wymiary $d_k = \dim\ker N^k$ są tu równe $d_1 = 7$, $d_2 = 13$, $d_3 = 17$, $d_4 = 20$, $d_5 = d_6 = \ldots = 21$.

Jeśli teraz dowodzoną tezę sformułujemy w postaci '*liczby* $d_k = \dim\ker N^k$ *całkowicie determinują długości* $k_1 \geqslant k_2 \geqslant k_3 \ldots$ *poszczególnych kolumn diagramu*', to staje się ona oczywista: diagram możemy zbudować, stawiając jego kolejne wiersze, których długości w_k znajdujemy, znając wszystkie sumy $w_1 + \ldots + w_k = d_k$. ([89])

21							$w_1 + w_2 + w_3 + w_4 + w_5 = 21$
18	19	20					$w_1 + w_2 + w_3 + w_4 = 20$
14	15	16	17				$w_1 + w_2 + w_3 = 17$
8	9	10	11	12	13		$w_1 + w_2 = 13$
1	2	3	4	5	6	7	$w_1 = 7$

Pojęcie N-serii można i warto uogólnić na dowolny operator w następujący sposób:

484. **Definicja**. Jeśli $F \in \operatorname{End} V$, to *$F$-serią długości $k \geqslant 1$* nazywamy ciąg

$$v_1,\ v_2 = F_\lambda(v_1),\ \ldots,\ v_k = F_\lambda(v_{k-1})$$

(gdzie λ jest pewną liczbą), taki że $F_\lambda(v_k) = 0$, lecz $v_k \neq 0$. Wtedy $\langle v_1, \ldots, v_k\rangle$ jest podprzestrzenią F-niezmienniczą; ponadto $\lambda \in \mathbb{K}$ jest wartością własną F (gdyż v_k jest $\neq 0$ wektorem własnym); mówimy, że dana seria *odpowiada* wartości własnej λ lub że *jest związana* z λ.

Prostym i pouczającym ćwiczeniem jest sprawdzenie liniowej niezależności F-serii oraz zbadanie, jak wyglądają macierze (tzw. *klatki Jordana 1. i 2. rodzaju*) operatora $F\big|_W$ w każdej z dwu baz $v_1, \ldots, v_k$ oraz $v_k, \ldots, v_1$ podprzestrzeni $W := \langle v_1, \ldots, v_k\rangle$.

8.9 Pewien Ważny Lemat i jego konsekwencje

W poniższym lemacie nie zakłada się, że przestrzeń V ma skończony wymiar!

485. **Lemat**. Jeśli wielomian $w \in \mathbb{K}[\,\cdot\,]$ zeruje operator $F \in \operatorname{End} V$ i ma rozkład $w = u_1 \cdot \ldots \cdot u_r$ na parami względnie pierwsze czynniki $u_i \in \mathbb{K}[\,\cdot\,]$, to

$$V = V_1 \dot{+} \ldots \dot{+} V_r, \quad \text{gdzie} \quad V_i = \ker u_i(F);$$

[89] Oczywiście mamy $w_k = d_k - d_{k-1}$, gdzie $d_0 := 0$. Nietrudno też uzasadnić wzory $k_i = \max\{k :\ w_k \geqslant i\}$; odwrotne zależności mają podobną postać: $w_i = \max\{j :\ k_j \geqslant i\}$.

oczywiście podprzestrzenie V_i są F-niezmiennicze, a każdy z operatorów $F_{(i)} := F|_{V_i} \in \operatorname{End} V_i$ jest zerowany przez wielomian u_i, tzn. $u_i(F_{(i)}) = 0$.

Wielomiany $w_i := u_1 \cdot \ldots \widehat{u_i} \ldots \cdot u_r$ są względnie pierwsze, $\mathrm{NWD}(w_1, \ldots, w_n) = 1$, więc $\exists \varphi_1, \ldots, \varphi_r \in \mathbb{K}[\,\cdot\,] : \varphi_1(\lambda)w_1(\lambda) + \ldots + \varphi_r(\lambda)w_r(\lambda) = 1$. Weźmy $p_i := \varphi_i w_i$ oraz $P_i := p_i(F) = \varphi_i(F)w_i(F)$. Wtedy

- $P_1 + \ldots + P_r = \mathrm{id}_V$, gdyż $p_1(\lambda) + \ldots + p_r(\lambda) = 1$;
- $i \neq j \;\Rightarrow\; P_i P_j = 0$, gdyż wtedy $w(\lambda) \mid p_i(\lambda)p_j(\lambda)$;
- $P_i^2 = P_i$, gdyż $P_i^2 = P_i(\mathrm{id}_V - \sum_{j \neq i} P_j) = P_i$ dzięki poprzedniej własności;
- zatem stosując fakt 475. dostajemy rozkład $V = V_1 \dot{+} \ldots \dot{+} V_r$, gdzie $V_i := \operatorname{im} P_i$;
- $V_i \subset \ker u_i(F)$, gdyż $w(\lambda) = u_i(\lambda)w_i(\lambda) \mid u_i(\lambda)p_i(\lambda)$, skąd $0 = u_i(F)p_i(F) = u_i(F)P_i$;
- odwrotnie, $V_i \supset \ker u_i(F)$; istotnie, jeśli $v \in \ker u_i(F)$, tzn. $u_i(F)v = 0$, to $P_j(v) = 0$ dla $j \neq i$, gdyż $u_i(\lambda) \mid w_j(\lambda) \mid p_j(\lambda)$; stąd $v = \sum\limits_j P_j(v) = P_i(v) \in V_i$.

486. **Uwaga**. Najważniejszy jest przypadek, gdy $u_i(\lambda) = (\lambda - \lambda_i)^{k_i}$, gdzie $\lambda_1, \ldots, \lambda_i \in \mathbb{K}$ są parami różne, zaś $k_1, \ldots, k_r \in \mathbb{N}$; dostajemy wówczas

$$V = V_1 \dot{+} \ldots \dot{+} V_r, \quad \text{gdzie} \quad V_i = V_{k_i}(\lambda_i, F) := \ker(F_{\lambda_i})^{k_i};$$

ponadto $V_i = V_k(\lambda_i, F)$ dla $k \geqslant k_i$, gdyż zwiększenie wykładnika k_i nie narusza równania $w(F) = 0$; zatem każda z przestrzeni V_i jest teraz albo przestrzenią pierwiastkową $V(\lambda_i)$ — gdy $\lambda_i \in \operatorname{Sp} F$ — albo zerową.

Najważniejszym chyba wnioskiem z powyższego lematu jest następujące

487. **Twierdzenie** (*rozkład V na przestrzenie pierwiastkowe*). Jeśli wielomian charakterystyczny operatora $F \in \operatorname{End} V$ ma rozkład na czynniki stopnia 1 w $\mathbb{K}[\lambda]$ (a tak jest zawsze np. dla $\mathbb{K} = \mathbb{C}$), to V jest sumą prostą wszystkich przestrzeni pierwiastkowych $V(\lambda^*)$, $\lambda^* \in \operatorname{Sp} F$.

Wystarczy w lemacie wziąć $w(\lambda) = \pm w_F(\lambda)$ i skorzystać z tw. Cayleya-Hamiltona.

488. **Fakt**. Wielomian minimalny operatora $F \in \operatorname{End} V$ wyraża się wzorem

$$\widehat{w}_F(\lambda) = (\lambda - \lambda_1)^{h_1} \cdots (\lambda - \lambda_r)^{h_r}, \quad \text{gdzie} \quad \{\lambda_1, \ldots, \lambda_r\} = \operatorname{Sp} F,$$

w którym $h_i := h(\lambda_i, F)$ jest *wysokością przestrzeni* $V(\lambda_i, F)$, określoną jako najmniejszy numer $h \geqslant 1$, taki że $V(\lambda_i, F) = V_h(\lambda_i, F)$.

Z lematu wiemy, że $w(F) = 0$ implikuje $V_{k_i}(\lambda_i) = V(\lambda_i)$, a więc każdy wielomian zerujący F ma wykładniki $k_i \geqslant h_i$. Z drugiej strony dla $w(\lambda) := \prod_i (\lambda - \lambda_i)^{h_i}$ mamy oczywiście $w(F) = 0$, gdyż na $V(\lambda_i)$ zeruje się $(F_{\lambda_i})^{h_i}$, zaś $(\lambda - \lambda_i)^{h_i} \mid w(\lambda)$.

489. **Fakt**. $\dim V(\lambda_i, F) = k(\lambda_i, F) :=$ krotność λ_i jako pierwiastka $w_F(\lambda)$. W konsekwencji *wysokość* $h_i = h(\lambda_i, F)$, zdefiniowana 'niekonstruktywnie' jako najmniejsze $h \in \mathbb{N}$, dla którego $V_h(\lambda_i) = V(\lambda_i)$, jest zarazem najmniejszym numerem h, takim że $\dim V_h(\lambda_i) = k_i = k(\lambda_i, F)$.

Z 450. wiemy, że wielomian $w_F(\lambda)$ jest iloczynem $w_{F_i}(\lambda)$, gdzie $F_i \in \operatorname{End} V_i$ są składowymi F odpowiadającymi rozkładowi V na sumę prostą $V_i := V(\lambda_i)$. Lecz skoro operator $N_i = F_i - \lambda_i \operatorname{id}$ jest nilpotentny, to $w_{F_i} = (\lambda_i - \lambda)^{k_i}$, gdzie $k_i = \dim V_i$.

490. **Twierdzenie**. Jeśli w_F jest rozkładalny na czynniki stopnia 1 w $\mathbb{K}[\,\cdot\,]$, to przestrzeń V ma bazę, w której macierz F jest górnotrójkątna.

Niech $\operatorname{Sp} F = \{\lambda_1, \ldots, \lambda_r\}$; przestrzeni $V(\lambda_i)$ jest niezmiennicza względem F i — wprost z określenia $V(\lambda_i)$ — operator $N_i = F_{\lambda_i}\big|_{V(\lambda_i)}$ jest nilpotentny; zatem $V(\lambda_i)$ ma bazę, w której N_i ma macierz ściśle górno-trójkątną, a więc $F\big|_{V(\lambda_i)}$ — macierz trójkątną, o wyrazach λ_i na diagonali. Skoro $V = \sum^{\cdot} V(\lambda_i)$, to baza V, będąca konkatenacją takich baz dla $V(\lambda_i)$, gdzie $i \in \overline{1,r}$, ma oczywiście żądaną własność.

Inny (nieco dłuższy, ale bardziej elementarny) dowód tego twierdzenia:

Indukcja względem $n = \dim V$: Gdy $\dim V = 1$ teza jest oczywista. Krok indukcyjny: weźmy dowolne $\lambda_n \in \operatorname{Sp} F$; operator $F_{\lambda_n} \in \operatorname{End} V$ jest osobliwy, więc osobliwy jest także $F^*_{\lambda_n} \in \operatorname{End} V^*$ tzn. $\exists\, 0 \neq \phi \in V^* : F^*_{\lambda_n}(\phi) = 0$. Skoro $F^*(\phi) = \phi \circ F$, oznacza to, że $\phi \circ F_{\lambda_n} = 0$. Stąd $\forall v : 0 = \big(\phi \circ F_{\lambda_n}\big)(v) = \phi\big(F(v) - \lambda_n v\big)$, tzn. $\forall v : F(v) - \lambda_n v \in V_{n-1} := \ker \phi$. W szczególności podprzestrzeń V_{n-1} jest F-niezmiennicza, a biorąc dowolny $e_n \in V \setminus V_{n-1}$ mamy $F(e_n) - \lambda_n e_n \in V_{n-1}$. Wynika z tego, że $w_F(\lambda) = w_{\tilde F}(\lambda)(\lambda_n - \lambda)$ (gdzie $\tilde F \in \operatorname{End} V_{n-1}$ jest obcięciem F do V_{n-1}), więc możemy powtórzyć nasz zabieg, biorąc V_{n-1} zamiast V, itd. Dostaniemy w ten sposób łańcuch podprzestrzeni $V = V_n \supset V_{n_1} \supset \ldots \supset V_2 \supset V_1$ oraz bazę $e = (e_1, \ldots, e_n)$, takie że $V_k = \langle e_1, \ldots, e_k\rangle$ oraz $F(V_k) \subset V_{k-1}$.

Uwaga. Zamiast $\ker \phi$ można użyć innej konstrukcji V_{n-1}: Dla $\lambda_n \in \operatorname{Sp} F$ operator F_{λ_n} jest niesurjektywny, więc $\dim \operatorname{im} F_{\lambda_n} < n$; jasne jest też, że każda $(n-1)$-wymiarowa podprzestrzeń V_{n-1}, zawierająca $\operatorname{im} F_{\lambda_n}$, ma własność $F_{\lambda_n} V \subset V_{n-1}$.

491. **Ćwiczenie**. Przy założeniach powyższego twierdzenia dowieść, że

(a) $\det(e^F) = e^{\operatorname{tr} F}$;

(b) $w_F(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda) \quad \Rightarrow \quad w_{\varphi(F)}(\lambda) = \prod_{i=1}^{n} (\varphi(\lambda_i) - \lambda)$

dla dowolnego wielomianu $\phi(\lambda)$, a więc także dla 'dowolnej' funkcji ϕ, dla której operator $\phi(F)$ jest określony.

Rozwiązanie. Łatwo sprawdzić, że jeśli $\mathrm{GT}(\mathbb{K}^n)$ oznacza podzbiór ${\mathbb{K}^n}_n$, złożony z macierzy górnotrójkątnych, zaś $\delta_1, \ldots, \delta_n$ — funkcje na $\mathrm{GT}(\mathbb{K}^n)$, dane wzorem $\delta_i(\boldsymbol{A}) := {A^i}_i$ = i-ty wyraz diagonalny $\boldsymbol{A}$, to: (1) $\mathrm{GT}(\mathbb{K}^n)$ jest podprzestrzenią przestrzeni ${\mathbb{K}^n}_n$, zamkniętą względem mnożenia macierzy (*podalgebrą* algebry ${\mathbb{K}^n}_n$); (2) każde δ_i jest homomorfizmem algebry $\mathrm{GT}(\mathbb{K}^n)$ w $\mathbb{K}$, tzn. że δ_i jest liniowe oraz multiplikatywne: $\delta_i(\boldsymbol{AB}) = \delta_i(\boldsymbol{A})\delta_i(\boldsymbol{B})$. Wobec tego (3) $\delta_i(w(\boldsymbol{A})) = w(\delta_i(\boldsymbol{A}))$ dla każdego wielomianu $w(\lambda)$, więc i dla 'dowolnej' funkcji, dla której $w(\boldsymbol{A})$ ma sens.

Dzięki twierdzeniu dla pewnej bazy e macierz $\boldsymbol{A} = [F]^e_e$ należy do $\mathrm{GT}(\mathbb{K}^n)$, mamy więc: $\det e^{\boldsymbol{A}} = \prod \delta_i(e^{\boldsymbol{A}}) = \prod e^{\delta_i \boldsymbol{A}} = e^{\sum \delta_i \boldsymbol{A}} = e^{\operatorname{tr} \boldsymbol{A}}$, co daje (a), oraz $w_{\boldsymbol{A}}(\lambda) = \prod(\lambda - \delta_i \boldsymbol{A})$, $w_{\varphi(\boldsymbol{A})}(\lambda) = \prod\big(\lambda - \delta_i(\varphi(\boldsymbol{A}))\big) = \prod\big(\lambda - \varphi(\delta_i \boldsymbol{A})\big)$, co daje (b).

Uwaga. We wzorach (a),(b) założenie o rozkładalności $w_F(\lambda)$ nie jest istotne; można się go pozbyć, stosując tzw. *kompleksyfikację* przestrzeni V i operatora F.

A oto najważniejszy (a przynajmniej najbardziej znany) wynik tego rozdziału:

492. **Twierdzenie** (*baza jordanowska operatora*). Jeśli wilomian $w_F(\lambda)$ ma w $\mathbb{K}[\lambda]$ rozkład na czynniki stopnia 1, to V ma bazę, złożoną z F-serii.

W każdej z przestrzeni pierwiastkowych $V_i = V(\lambda_i)$ weźmy bazę, złożoną z N_i-serii operatora nilpotentnego $N_i = F_{\lambda_i}\big|_{V_i} = F\big|_{V_i} - \lambda_i \operatorname{id}_{V_i}$; konkatenacja wszystkich tych baz jest oczywiście bazą V złożoną z F-serii (tzw. *bazą jordanowską* operatora F).

493. **Uwaga**. Tak, jak dla operatora nilpotentnego, zbiór długości F-serii związanych z dowolną ustaloną wartością wł. nie zależy od wyboru bazy jordanowskiej, por. 483.

* **Przykład 1**. Znajdziemy najpierw opis przestrzeni rozwiązań równania różniczkowego zwyczajnego liniowego o stałych współczynnikach, mianowicie $w(\frac{d}{dt})v(t) = 0$. Wielomian określający równanie $w(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - \ldots - a_0 \in \mathbb{C}[\lambda]$ ma rozkład na czynniki $w(\lambda) = (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_r)^{k_r}$, $\lambda_i \in \mathbb{C}$, $k_i \in \mathbb{N}$, $k_1 + \ldots + k_r = n$. Mając zadany przedział otwarty $J \subset \mathbb{R}$ określmy zaspoloną przestrzeń wektorową

$$V := \big\{v : J \to \mathbb{C} : \ v \text{ jest } n\text{-krotnie różniczkowalna oraz } w(\tfrac{d}{dt})v(t) = 0\big\}.$$

Zauważmy, że $V \subset C^\infty(J;\mathbb{C})$ ([90]), a ponadto $v \in V \ \Rightarrow \ \frac{d}{dt} \in V$ dzięki oczywistej przemienności operacji $\frac{d}{dt}$ i $w(\frac{d}{dt})$. Możemy więc określić operator $F := \frac{d}{dt}\big|_V \in \operatorname{End} V$ i spostrzec, że $w(F) = 0$ wprost z określenia V. Z lematu wynika więc rozkład

$$V = V_1 \dot{+} \ldots \dot{+} V_r, \text{ gdzie } V_i = \ker(F_{\lambda_i})^{k_i} = \big\{v \in C^\infty(J;\mathbb{C}) : \ (\tfrac{d}{dt} - \lambda_i)^{k_i} v(t) = 0\big\}.$$

Lecz przez indukcję wzgl. k dostajemy tożsamość $(\frac{d}{dt} - \lambda)^k \left[e^{\lambda t}u(t)\right] = e^{\lambda t}u^{(k)}(t)$, z której natychmiast wynika postać poszczególnych podprzestrzeni V_i:

$$V_i = \big\{v : \ v(t) = e^{\lambda_i t}(\text{wielomian stopnia} < k_i \text{ względem } t)\big\}$$

W ten sposób rozwiązaliśmy równanie różniczkowe ściśle algebraicznymi środkami!

* **Przykład 2**. Opiszemy teraz przestrzeń rozwiązań rekurencji liniowej stopnia n:

$$\forall\, k \geqslant 0 : \ x_{k+n} = a_0 x_k + a_1 x_{k+1} + \ldots + a_{n-1} x_{k+n-1}, \qquad (*)$$

gdzie współczynniki $a_0, \ldots, a_{n-1} \in \mathbb{K}$ są dane. Zauważmy, że jeśli w przestrzeni $\mathbb{K}^{\mathbb{Z}_+} = \{\boldsymbol{x} = (x_0, x_1, x_2, \ldots) : x_k \in \mathbb{K}\}$ określimy *operator przesunięcia* ('shift') wzorem $(S\boldsymbol{x})_k := x_{k+1}$, tzn. $S(x_1, x_1, x_2, \ldots) := (x_1, x_2, x_3, \ldots)$, zaś wielomian $w(\lambda)$ ma postać $w(\lambda) := \lambda^n - a_{n-1}\lambda^{n-1} - \ldots - a_1\lambda - a_0$, to $(*) \iff \boldsymbol{x} \in \ker w(F)$, tzn. przestrzenią rozwiązań rekurencji $(*)$ jest $V := \ker w(F)$. Zauważmy przy tym, że

(1) $\dim V = n$, gdyż rozwiązanie $(*)$ jest w pełni określone przez $x_0, \ldots, x_{n-1}$;
(2) przestrzeń V jest S-niezmiennicza, gdyż $w(S)\boldsymbol{x} = 0 \ \Rightarrow \ w(S)S\boldsymbol{x} = Sw(S)\boldsymbol{x} = 0$;
(3) operator $F := S\big|_V \in \operatorname{End} V$ spełnia równanie $w(F) = 0$ (wprost z definicji V).

Wobec tego, jeśli $w(\lambda)$ ma rozkład $w(\lambda) = (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_r)^{k_r}$ (a jest tak zawsze dla $\mathbb{C} = \mathbb{K}$), to z lematu dostajemy $V = V_1 \dot{+} \ldots \dot{+} V_r$, gdzie $V_i = \ker(S - \lambda_i)^{k_i}$. Otóż jeśli $\boldsymbol{x} = \boldsymbol{g}(\lambda)\boldsymbol{u}$ oznacza zwykły (punktowy) iloczyn ciągu geometrycznego $\boldsymbol{g}(\lambda) = (1, \lambda, \lambda^2, \ldots)$ i ciągu $\boldsymbol{u} = (u_0, u_1, u_2, \ldots)$, tzn. $x_j = \lambda^j u_j$, to przez indukcję względem $k \geqslant 0$ dostajemy tożsamość $\boxed{(S - \lambda)^k[\boldsymbol{g}(\lambda)\boldsymbol{u}] = \lambda^k \boldsymbol{g}(\lambda)(S - 1)^k \boldsymbol{u}}$; z kolei

$$(S-1)^k \boldsymbol{u} = 0 \iff \begin{pmatrix} \text{ciąg } \boldsymbol{u} \text{ jest wielomianowy stopnia} < k, \text{ tzn. istnieje} \\ \text{wielomian } U(\lambda) \text{ stopnia} < k, \text{ taki że } \forall j : u_j = U(j) \end{pmatrix}$$

['$\Leftarrow$', bo $\deg\big[U(\lambda+1) - U(\lambda)\big] = \deg U(\lambda) - 1$; '$\Leftrightarrow$', bo $\ker(S-1)^k$ ma wymiar k i zawiera, wobec '$\Leftarrow$', k-wymiarową przestrzeń ciągów wielomianowych stopnia $< k$].

Zatem *każde rozwiązanie* $\boldsymbol{x}$ *rekurencji* $(*)$ *ma postać* $\boldsymbol{x} = \boldsymbol{g}(\lambda_1)\boldsymbol{u}_{(1)} + \ldots + \boldsymbol{g}(\lambda_r)\boldsymbol{u}_{(r)}$, *gdzie* $\boldsymbol{u}_{(i)}$ *jest* (jednoznacznie określonym) *ciągiem wielomianowym stopnia* $< k_i$.

[90] Stosując indukcję względem $m \geqslant 1$ bez trudu otrzymujemy, że

$$\forall\, m \geqslant 1 : \exists\, c_{m,0}, c_{m,1}, \ldots, c_{m,n-1} \in \mathbb{C} : \ \forall v \in V : \ v^{(n-1+m)} = \sum_{k=0}^{n-1} c_{m,k} v^{(k)}.$$

8.10 Operatory diagonalizowalne

494. **Fakt**. Dla operatora $D \in \operatorname{End} V$ następujące warunki są równoważne:

1. $\exists\ e$ (baza V): macierz $[D]^e_e$ jest diagonalna;
2. V ma bazę złożoną z wektorów własnych operatora D;
3. $\forall \alpha \in \operatorname{Sp} D : V(\alpha) = V_1(\alpha)$ [co można zapisać w postaci $h(\alpha, D) = 1$ (wysokość) lub $\dim V_1(\alpha) = k(\alpha, D)$ (krotność)], ponadto zaś $w_D(\lambda)$ jest rozkładalny w $\mathbb{K}[\lambda]$ na czynniki stopnia 1.
4. D ma *rozkład spektralny*: $D = \lambda_1 P_1 + \ldots + \lambda_r P_r$, gdzie zestaw $P_1, \ldots, P_r$ jest *rzutowym rozkładem jedynki* oraz $\lambda_i \in \mathbb{K}$;
5. Istnieje wielomian $w(\lambda)$ o pierwiastkach krotności 1, zerujący D i rozkładalny w $\mathbb{K}[\lambda]$ na czynniki stopnia 1.

Operator mający powyższe własności nazywamy *diagonalizowalnym*.

$\boxed{4. \Rightarrow 1.}$ Konkatenacja baz poszczególnych $V_i := \operatorname{im} P_i$ jest bazą diagonalizującą D. $\boxed{1. \Rightarrow 4.}$ Niech $P_1, \ldots, P_n$ — rzuty odpowiadające rozkładowi $V = \langle e_1 \rangle \dot{+} \ldots \dot{+} \langle e_n \rangle$ na 1-wymiarowe przestrzenie rozpinane przez wektory bazy diagonalizującej. Wtedy $D = \sum_{i=1}^{n} \delta_i P_i$. $\boxed{4. \Rightarrow 5.}$ Zauważmy, że $(\sum_i \alpha_i P_i)(\sum_j \beta_j P_j) = \sum_k (\alpha_i \beta_j) P_i$, skąd widać, że $w(\sum_i \lambda_i P_i) = \sum_i w(\lambda_i) P_i$ dla dowolnego wielomianu $w(\lambda)$; wystarczy więc jako $w(\lambda)$ wziąć dowolny wielomian o pojedynczych pierwiastkach, zerujący się na wszystkich λ_i. $\boxed{5. \Rightarrow 3.}$ Wprost z lematu 485. $\boxed{3. \Rightarrow 2.}$ Jasne, że konkatenacja baz poszczególnych przestrzeni $V_1(\lambda_i)$ jest bazą diagonalizującą D. $\boxed{1. \Leftrightarrow 2.}$ oczywiste.

495. **Uwaga**. Dany rozkład spektralny $D = \sum_i \lambda_i P_i$ można zawsze 'zredukować', odrzucając ewentualne zerowe rzuty P_i oraz zastępując P_i, odpowiadające jednakowym współczynnikom λ_i, sumą tych P_i (suma kilku P_i też jest rzutem!). Po takiej redukcji współczynniki λ_i rozkładu $D = \sum_i \lambda_i P_i$ będą już parami różne, $\{\lambda_i : i \in \overline{1,r}\}$ będzie zbiorem wszystkich wartości własnych D (może być jedno λ_i równe 0), a $V_i := \operatorname{im} P_i = V_1(\lambda_i, D)$ będą przestrzeniami własnymi D. Co więcej, *zredukowany rozkład spektralny D jest jednoznaczny z dokładnością do kolejności składników.*

496. **Wniosek** (z $3° \Rightarrow 1°$, oczywisty, lecz często przywoływany): Jeśli wielomian $w_F(\lambda)$ ma $n = \dim V$ pierwiastków parami różnych w ciele $\mathbb{K}$, to F jest diagonalizowalny([91]).

497. **Ćwiczenie**. Diagonalizowalność macierzy $\boldsymbol{A} \in \mathbb{K}^n_n$ (rozumiana oczywiście jako diagonalizowalność operatora $A \in \operatorname{End} \mathbb{K}^n$, $A(\boldsymbol{x}) := \boldsymbol{A}\boldsymbol{x}$) jest równoważna warunkowi

$$\exists \boldsymbol{B} \in \mathbb{K}^n_n : \boldsymbol{B}^{-1} \text{ istnieje oraz macierz } \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B} \text{ jest diagonalna.}$$

Wskaz. Zauważmy, że jeśli $\boldsymbol{B} \in \mathbb{K}^n_m$, to $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B} \operatorname{diag}(\lambda_1, \ldots, \lambda_m) \Leftrightarrow \forall j \in \overline{1,m}$: $\boldsymbol{A}\boldsymbol{B}_j = \lambda_j \boldsymbol{B}_j$; zatem można $\boldsymbol{B}$ zestawić z (kolumnowych) wektorów własnych $\boldsymbol{A}$.

498. **Fakt**. Niech $F, D \in \operatorname{End} V$, przy czym załóżmy, że operator D jest diagonalizowalny. Wtedy

$$\begin{pmatrix} F \text{ i } D \text{ komutują,} \\ \text{tzn. } FD = DF \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{przestrzenie własne } D \text{ są } F\text{-niezmiennicze,} \\ \text{tzn. } \forall \mu \in \operatorname{Sp} D : F(V(\mu, D)) \subset V(\mu, D) \end{pmatrix}.$$

$\boxed{\Rightarrow}$ Jeśli $v \in V(\mu, D)$, tzn. $D(v) = \mu v$, to $D\big(F(v)\big) = F\big(D(v)\big) = F\big(\mu v\big) = \mu F(v)$,

[91] Można to uzyskać w sposób bardziej elementarny, zob. 438.

tzn. $F(v) \in V(\mu, D)$. $\boxed{\Leftarrow}$ Dowolny $v \in V$ ma rozkład $v = \sum_i v_i$ na składowe $v_i \in V(\mu_i, D) = V_1(\mu_i, D)$. Z założenia $F(v_i) \in V(\mu_i, D)$, więc $D\big(F(v_i)\big) = \mu_i F(v_i)$, skąd $D\big(F(v)\big) = D\big(\sum_i F(v_i)\big) = \sum_i \ldots = \sum_i \mu_i F(v_i) = F\big(\sum_i \mu_i v_i\big) = F\big(D(v)\big)$.

8.11 Rozkład na część diagonalizowalną i nilpotentną

Załóżmy o operatorze $F \in \operatorname{End} V$, że jego wielomian charakterystyczny $w_F(\lambda)$ jest rozkładalny na czynniki stopnia pierwszego w $\mathbb{K}[\cdot]$.

499. **Twierdzenie.** Operator F ma jednoznaczny rozkład postaci

$$F = D + N, \text{ gdzie } \left\{ \begin{array}{c} D, N \in \operatorname{End} V, \\ D \text{ — diagonalizowalny, } N \text{ — nilpotentny,} \\ D \text{ i } N \text{ komutują, tzn. } DN = ND \end{array} \right\}.$$

Ponadto D i N można wyrazić jako wielomiany od operatora F.

Istnienie.

Niech $\operatorname{Sp} F = \{\lambda_1, \ldots, \lambda_r\}$, gdzie λ_i są parami różne, oraz niech P_i będzie układem rzutów odowiadających rozkładowi $V = \sum_{i=1}^{r} V(\lambda_i, F)$. Wtedy operator $D := \sum_{i=1}^{r} \lambda_i P_i$ jest diagonalizowalny (bo ma rozkład spektralny), $DF = FD$ (gdyż $V(\lambda_i, F)$ są F-niezmiennicze, co daje $FP_i = P_i F$), a więc także $ND = DN$ dla $N := F - D$. Ponadto N jest nilpotentny, gdyż N pokrywa się z F_{λ_i} na podprzestrzeni $V(\lambda_i, F)$.

Jednoznaczność.

(1) Przestrzeń $V(\lambda_i, D)$ jest niezmiennicza względem F (a więc i $N = F - D$), bowiem $v \in V(\lambda_i, D) \;\Rightarrow\; D(v) = \lambda_i v \;\Rightarrow\; D(F(v)) = F(D(v)) = F(\lambda_i v) = \lambda_i F(v)$.

(2) $V(\lambda_i, D) \subset V(\lambda_i, F)$. Istotnie, niech $v \in V(\lambda_i, D)$; mamy więc $D(v) = \lambda_i v$, skąd $F(v) = \lambda_i v + N(v)$, czyli $\boxed{F_{\lambda_i}(v) = N(v)}$. Stąd $\big(F_{\lambda_i}\big)^h(v) = N^h(v)$ (dzięki (1), przez indukcję), więc $v \in \ker\big(F_{\lambda_i}\big)^h$ (gdy $N^h = 0$, dla dużych h), tzn. $v \in V(\lambda_i, F)$.

(3) $V(\lambda_i, D) = V(\lambda_i, F)$, wskutek (2) i tego, że V jest sumą prostą zarówno $V(\lambda_i, D)$, jak też $V(\lambda_i, F)$. Stąd $D = \sum_i \lambda_i P_i$, gdzie λ_i są wartościami własnymi F, zaś P_i — rzutami odpowiadającymi rozkładowi V na $V(\lambda_i, F)$; to kończy dowód.

Wyrażenie D i N w postaci wielomianów od F.

Wielomiany $p_i(\lambda)$ skonstruowane w 485. dają rzuty $p_i(F) = P_i$, więc $D = \varphi(F)$, $N = \psi(F)$, gdzie $\varphi(\lambda) := \sum_{i=1}^{r} \lambda_i p_i(\lambda)$, $\psi(\lambda) := 1 - \varphi(\lambda)$.

500. **Fakt.** Dla ustalonego $\mu \in \operatorname{Sp} F$ niech $h \geqslant 1$ będzie na tyle duże, że $V(\mu) = V_h(\mu) = \ker(F_\mu)^h$. Wówczas $\sum_{\lambda \neq \mu} V(\lambda) = \operatorname{im}(F_\mu)^h$.

Pozwala to m.in. znaleźć rzut P_μ na $V(\mu)$ bez znajdowania pozostałych wartości własnych $\lambda \in \operatorname{Sp} F$, a tym bardziej przestrzeni $V(\lambda)$.

Skoro $G := \big(F_\mu\big)^h$ jest wielomianem od F, to każda z przestrzeni $V(\lambda)$, będąc F-niezmiennicza, jest także G-niezmiennicza. Przy tym G znika na $V(\mu)$, a więc $\operatorname{im} G \subset \widehat{V}(\mu) := \sum_{\lambda \neq \mu} V(\lambda)$. Z drugiej strony mamy $\dim \widehat{V}(\mu) = \dim V - \dim V(\mu) =$ $= \dim V - \dim \ker G = \dim \operatorname{im} G$, więc zawieranie to jest w istocie równością.

Alternatywny dowód twierdzenia Cayleya-Hamiltona

501. Ustalmy $0 \neq v \in V$; pokażemy, że $w_F(F)v = 0$. Niech $m \in \mathbb{N}$ będzie najmniejszą liczbą taką, że układ $v, Fv, \ldots, F^m v$ jest l.zależny. Wtedy układ $e_1 = v,\ e_2 = Fv,\ \ldots,\ e_m = F^{m-1}v$ jest l.niezależny oraz $F^m v = a_1 e_1 + \ldots + a_m e_m$ dla stosownych $a_i \in \mathbb{K}$. Wybierzmy $e_{m+1}, \ldots, e_n$ tak, by układ $e_1, \ldots, e_n$ był bazą V. W tej bazie $Fe_i = e_{i+1}$ dla $i \in \overline{1, m-1}$, $Fe_m = a_1 e_1 + \ldots + a_m e_m$, więc $[F]^e_e$ jest blokowo górno-trójkątna: $j \leqslant m < i \ \Rightarrow\ F^i_j = 0$. Stąd i ze wzoru $\begin{vmatrix} \lambda & 0 & \ldots & 0 & a_1 \\ -1 & \lambda & \ldots & 0 & a_2 \\ . & . & \ldots & . & . \\ 0 & 0 & \ldots & \lambda & a_{m-1} \\ 0 & 0 & \ldots & -1 & a_m \end{vmatrix} =$ $a_1 + a_2\lambda + \ldots + a_m\lambda^{m-1}$ (gdyż $D_m = D_{m-1} + a_m\lambda^{m-1}$) wielomian charakterystyczny F ma postać $w_F(\lambda) = g(\lambda)(c_1 + c_2\lambda + \ldots + c_m\lambda^{m-1} - \lambda^m)g(\lambda)$. Zatem $w_F(F) = g(F)(c_1 \mathrm{id}_V + c_2 F + \ldots + c_m F^{m-1} - F^m)$, skąd $w_F(F)v = g(F)(c_1 e_1 + \ldots + c_m e_m - F^m v) = 0$.

9 Przestrzenie unitarne

9.1 Hermitowski iloczyn skalarny

502. Niech V będzie przestrzenią unitarną lub euklidesową, tzn. przestrzenią wektorową nad ciałem $\mathbb{C}$ (odpowiednio: nad $\mathbb{R}$), wyposażoną w (dodatni) iloczyn skalarny (półtoraliniowy hermitowski dla $\mathbb{K} = \mathbb{C}$, dwuliniowy symetryczny dla $\mathbb{K} = \mathbb{R}$); będziemy go oznaczać symbolem $\left\langle \cdot \middle| \cdot \right\rangle$ lub (gdy mamy kilka przestrzeni) $\left\langle \cdot \middle| \cdot \right\rangle_V$.

O wszystkich rozważanych tu przestrzeniach zakładamy, że mają skończony wymiar.

503. **Antyizomorfizm Frecheta-Riesza:** $\boxed{V \ni v \mapsto v^\dagger := \left\langle v \middle| \cdot \right\rangle \in V^*}$.

Jest to wyróżniony przez iloczyn skalarny antyliniowy izomorfizm $V \cong V^*$; pozwala on w różnych sytuacjach operować przestrzenią V zamiast V^*; taka zamiana V^* na V prowadzi m.in. do takich obiektów, jak $W^\perp \subset V$ (odpowiednika $W^0 \subset V^*$ dla $W \subset V$) oraz $F^\dagger \in \mathrm{L}(W;V)$ (odpowiednika $F^* \in \mathrm{L}(W^*;V^*)$ dla $F \in \mathrm{L}(V;W)$).

Odwrotność (anty)izomorfizmu $(\cdot)^\dagger : V \to V^*$ oznaczmy tym samym symbolem $(\cdot)^\dagger$:

$$(\cdot)^\dagger : V^* \to V,\ \phi^\dagger = v \iff v^\dagger = \phi \iff \forall w \in V : \left\langle v \middle| w \right\rangle = \phi(v).$$

Ta 'nieroztropność' nie prowadzi do kolizji; okaże się wkrótce, że oba odwzorowania $(\cdot)^\dagger$ są tzw. sprzężeniami hermitowskimi operatorów: $\mathbb{K} \to V$, mianowicie $\lambda \mapsto \lambda v$ (naturalne i 'bezpieczne' jest oznaczanie tego operatora symbolem v) i $\phi : V \to \mathbb{K}$.

504. **Ortogonalność i dopełnienie ortogonalne podprzestrzeni:**

$v \perp w \overset{\text{def}}{\iff} \left\langle v \middle| w \right\rangle = 0$; mówimy wtedy, że wektory v, w przestrzeni unitarnej V są *prostopadłe* (lub *ortogonalne*); jeśli $V_1, V_2, W \subset V$, to

$$V_1 \perp V_2 \overset{\text{def}}{\iff} \forall v_1 \in V_1, v_2 \in V_2 :\ v_1 \perp v_2;$$
$$W^\perp := \{v \in V :\ \forall w \in w :\ v \perp w\};$$

oczywiście $W^\perp$ jest największym podzbiorem V ortogonalnym do W.

505. **Fakt** (*własności* $(\cdot)^\perp$ *w przestrzeni unitarnej lub euklidesowej* V):

(1) $W^\perp$ jest podprzestrzenią V dla każdego podzbioru $W \subset V$;
w (2)..(5) założymy, że $W, W_1, W_2 \subset V$ są podprzestrzeniami:
(2) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$;
(3) $V = W \dotplus W^\perp$, w szczególności $\dim(W^\perp) = \dim V - \dim W$;
(4) $(W^\perp)^\perp = W$;
(5) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

Są to własności analogiczne i związane z własnościami anihilatora, a więc też ich sprawdzenie jest podobne: (1) oczywiste; (2) proste sprawdzenie; (3) $W \cap W^\perp = \{0\}$, bo $w \perp w \Rightarrow \|w\| = 0 \Rightarrow w = 0$ wskutek dodatniości $\left\langle \cdot \middle| \cdot \right\rangle$; zarazem $\dim W^\perp = \dim W^0 = \dim V - \dim W$; (4) '$\supset$' jest oczywiste, a zarazem $\dim W^{\perp\perp} = \dim W$; (5) wystarczy wziąć $(\cdot)^\perp$ obu stron (2) oraz zamienić W_i na $W_i^\perp$.

9.2 Sprzężenie hermitowskie operatora

506. **Definicja** (*sprzężenie hermitowskie*). Jeśli $F \in \mathrm{L}(V;W)$, gdzie obie przestrzenie V i W są unitarne lub euklidesowe, to wzór

$$\boxed{\left\langle v \middle| F^\dagger w \right\rangle_V = \left\langle Fv \middle| w \right\rangle_W}$$

jest poprawną definicją pewnego operatora $F^\dagger \in \mathrm{L}(W;V)$, zwanego *hermitowskim sprzężeniem* F. Z operatorem $F^* \in \mathrm{L}(W^*;V^*)$ [takim, że $F^*\phi := \phi \circ F$] jest on związany przemiennością następ. diagramu:

$$\begin{array}{ccc} W & \xrightarrow{F^\dagger} & V \\ {\scriptstyle\dagger}\downarrow & & \downarrow{\scriptstyle\dagger} \\ W^* & \xrightarrow{F^*} & V^* \end{array}\,, \qquad \begin{array}{c}\text{co oznacza, że } F^\dagger \text{ jest złożeniem} \\ W \xrightarrow{\dagger} W^* \xrightarrow{F^*} V^* \xrightarrow{\dagger} V.\end{array}$$

Uwaga. $F^\dagger$ zależy oczywiście zarówno od F, jak i iloczynów skalarnych w V i W.

Własności: (1) $F \mapsto F^\dagger$ jest antyliniowe; (2) inwolutywność: $(F^\dagger)^\dagger = F$; (3) $(F \circ G)^\dagger = G^\dagger \circ F^\dagger$ (odwraca kolejność składania); (4) $\mathrm{id}_V^\dagger = \mathrm{id}_V$; stąd oraz z (3) wynika łatwo (5) jeśli F^{-1} istnieje, to $(F^{-1})^\dagger = (F^\dagger)^{-1}$.

507. **Fakt**. Jeśli $F \in \mathrm{L}(V;W)$, gdzie obie przestrzenie V, W są unitarne (lub euklidesowe), to zachodzą wzory

$$\text{(a) } \ker F^\dagger = (\operatorname{im} F)^\perp; \qquad \text{(b) } \operatorname{im} F^\dagger = (\ker F)^\perp.$$

(a) $w \in \ker F^\dagger \;\Leftrightarrow\; F^\dagger w = 0 \;\Leftrightarrow\; \forall v : \big(0 = \langle F^\dagger w | v\rangle, \text{tzn. } 0 = \langle w | Fv\rangle\big) \;\Leftrightarrow\; w \perp \operatorname{im} F$.
(b) Stosując (a) do $F^\dagger$ widzimy, że dopełnienia ortogonalne obu stron są równe; stąd teza wskutek inwolutywności operacji $(\,\cdot\,)^\perp$.

508. **Uwaga**. Bezpośrednio widoczne jest zawieranie $\operatorname{im} F^\dagger \subset (\ker F)^\perp$:

jeśli $v \in \operatorname{im} F^\dagger$ (tzn. $\exists w : v = F^\dagger w$), to $v \perp \ker F$, gdyż $\langle v | \cdot \rangle = \langle w | F(\cdot)\rangle$.

Okazuje się, że dla *nieskończenie wymiarowych przestrzeni Hilberta* inkluzja '$\subset$' nie zawsze jest równością; ważne jest tzw. *twierdzenie Banacha o obrazie domkniętym*, mówiące że równość (b) jest równoważna domkniętości (w W) podprzestrzeni $\operatorname{im} F$.

509. **Definicja**. Operator $F \in \operatorname{End} V$ w przestrzeni unitarnej nazywa się:

(N) *normalny*, jeśli $F^\dagger F = F F^\dagger$;
(H) *hermitowski* (lub *samosprzężony* lub *symetryczny*), jeśli $F^\dagger = F$;
(A) *antyhermitowski* (a *antysymetryczny* dla $\mathbb{K} = \mathbb{R}$), jeśli $F^\dagger = -F$;
(U) *unitarny* (lub *ortogonalny* — gdy V jest euklidesowa), jeśli

$$F^\dagger F = \mathrm{id}_V, \;\text{ tzn. }\; F^{-1} \text{ istnieje i } \; F^\dagger = F^{-1}.$$

Zauważmy, że operatory normalne tworzą najobszerniejszą z tych klas: zawiara ona wszystkie pozostałe wymienione tu klasy. Łatwo też sprawdzić, że podzbiór operatorów hermitowskich jest zamknięty względem dodawania i mnożenia przez liczby rzeczywiste, a więc jest $\mathbb{R}$-podprzestrzenią przestrzeni $\operatorname{End} V$ (ale nie podprzestrzenią sensu stricte, gdy $\mathbb{K} = \mathbb{C}$). To samo dotyczy zbioru operatorów antyhermitowskich.

Uwaga. Słowo 'normalny' jest tu niefortunne o tyle, że operatory normalne stanowią jedynie 'znikomą część' zbioru $\operatorname{End} V$. Ponadto podzbiór operatorów normalnych w $\operatorname{End} V$ jest trudny do sparametryzowania i nie ma dobrych własności algebraicznych, np. na ogół 'normalność' nie

przenosi się z 'półproduktów' na produkt końcowy przy operacjach algebraicznych, np. dodawaniu i składaniu. Niemniej jednak w zastosowaniach 'normalność' jest bardzo ważna i często spotykana.

510. Złożenie dwu operatorów hermitowskich na ogół nie jest operatorem hermitowskim; wzór $(F_1 \circ F_2)^\dagger = F_2^\dagger \circ F_1^\dagger = F_2 \circ F_1$ pokazuje natomiast, że operatory hermitowskie tworzą zbiór zamknięty względem *antykomutatora* $[F_1, F_2]_+ := F_1 \circ F_2 + F_2 \circ F_1$, zaś operatory hermitowskie — względem *komutatora* $[F_1, F_2] := F_1 \circ F_2 - F_2 \circ F_1$.

Operatory hermitowskie i antyhermitowskie w przestrzeni unitarnej (gdy $\mathbf{K} = \mathbf{C}$) są ściśle ze sobą związane: $\big(H \in \operatorname{End} V \text{ jest hermitowski}\big) \iff \big(A := iH \text{ jest antyhermitowski}\big)$, gdyż $F \mapsto F^\dagger$ jest antyliniowe; z tego powodu zajmowanie się operatorami antyhermitowskimi jest właściwie niepotrzebne. Zupełnie inaczej jest w przestrzeniach euklidesowych, gdy $\mathbf{K} = \mathbf{R}$: tu operatory symetryczne i antysymetryczne stanowią niejeko dwa odrębne i zupełnie niepodobne światy.

511. Zbiór operatorów unitarnych z $\operatorname{End} V$ zawiera id_V, jest zamknięty wzgl. składania:

$$\big(F = F_1 F_2,\ F_i^\dagger F_i = \mathrm{id}_V\big) \ \Rightarrow\ F^\dagger F = F_2^\dagger (F_1^\dagger F_1) F_2 = F_2^\dagger F_2 = \mathrm{id}_V,$$

a także względem brania odwrotności: $F^\dagger F = \mathrm{id}_V \ \Rightarrow\ F F^\dagger = \mathrm{id}_V$. Oznacza to, że jest podgrupą (nazywaną *grupą unitarną*, a w przypadku $\mathbb{K} = \mathbb{R}$ — *ortogonalną*) grupy wszystkich odwracalnych operatorów z $\operatorname{End} V$ ([92]).

Fakt. Warunek $F^\dagger F = \mathrm{id}_V$ na operator $F \in \operatorname{End} V$ jest równoważny F-niezmienniczości (tzn. 'zachowywaniu przez F') iloczynu skalarnego:

$$\forall v, w \in V : \big\langle Fv \big| Fw \big\rangle = \big\langle v \big| w \big\rangle.$$

W konsekwencji każdy operator unitarny lub ortogonalny jest izometrią (tzn. zachowuje odległości) w przestrzeni metrycznej $(V, d = \|\cdot - \cdot\|)$:

$$d(v, w) := \|v - w\| = d(F(v), F(w)).$$

Istotnie, $\big\langle Fv \big| Fw \big\rangle - \big\langle v \big| w \big\rangle = \big\langle v \big| (F^\dagger F - \mathrm{id}_V) w \big\rangle$ z definicji sprzężenia operatora.

512. **Fakt**. Każdy operator $F \in \operatorname{End} V$ ma jednoznaczny rozkład na część symetryczną i antysymetryczną:

$$F = H + A, \ \text{ gdzie } \ H^\dagger = H,\ A^\dagger = -A,$$
$$\text{mianowicie } \ H = \tfrac{1}{2}(F + F^\dagger), \ \ A = \tfrac{1}{2}(F - F^\dagger).$$

Przy tym F *jest normalny* $\iff$ *jego obie części komutują:* $HA = AH$. Gdy $\mathbb{K} = \mathbb{C}$, tzn. w przestrzeni unitarnej, F ma jednoznaczny rozkład

$$F = H_1 + iH_2, \text{ gdzie } H_1, H_2 \text{ są hermitowskie}$$
$$\text{mianowicie } \ H_1 = \tfrac{1}{2}(F + F^\dagger), \ \ H_2 = \tfrac{1}{2i}(F - F^\dagger),$$

zwane *częścią rzeczywistą i urojoną* operatora F (m.in. dlatego, że wzory na nie przypominają wzory $\operatorname{Re} z = \frac{z+\overline{z}}{2}$, $\operatorname{Im} z = \frac{z-\overline{z}}{2i}$). Przy tym

$$F \textit{ jest normalny} \iff \textit{obie te części komutują: } H_1 H_2 = H_2 H_1.$$

Posprawdzanie tego wszystkiego stanowi bardzo proste (obligatoryjne!) ćwiczenie.

513. **Fakt**. $F \in \operatorname{End} V$ jest normalny $\iff\ \forall v \in V :\ \|F^\dagger v\| = \|Fv\|$. W konsekwencji jeśli operator F jest normalny, to:

[92] Ta ostatnia grupa, nazywana jest *pełną grupą liniową* albo *grupą automorfizmów przestrzeni* V i oznaczana bywa symbolem $\mathrm{GL}(V)$ (od ang. skrótu *general linear*) lub $\mathrm{Aut}(V)$.

(1) $\ker F^\dagger = \ker F$;
(2) $Fv = \lambda v \iff F^\dagger v = \overline{\lambda} v$;
(3) $\ker F = (\operatorname{im} F)^\perp$, ogólniej: $\forall \lambda \in \mathbb{K}\colon \ker F_\lambda = (\operatorname{im} F_\lambda)^\perp$.

Oznaczmy $H := F^\dagger F - F F^\dagger$. $\boxed{\Rightarrow}$ Mamy $H = 0$, więc $\|Fv\|^2 = \langle Fv | Fv \rangle = \langle F^\dagger F v | v \rangle = \langle F F^\dagger v | v \rangle = \langle F^\dagger v | F^\dagger v \rangle = \|F^\dagger v\|^2$. $\boxed{\Leftarrow}$ Określmy $\phi(v,w) := \langle v | Hw \rangle$; skoro $H^\dagger = H$, to forma ϕ jest hermitowska (dla $\mathbb{K} = \mathbb{C}$) lub symetryczna (dla $\mathbb{K} = \mathbb{R}$), zaś poprzedni rachunek daje $\forall v : \phi(v,v) = 0$; stąd $\phi = 0$ dzięki wzorowi polaryzacyjnemu, a to oznacza $H = 0$, tzn. normalność F.

(1) $v \in \ker F \iff \|Fv\| = 0 \iff \|F^\dagger v\| = 0 \iff v \in \ker F^\dagger$. (2) Skoro $(F_\lambda)^\dagger = (F - \lambda \operatorname{id}_V)^\dagger = F^\dagger - \overline{\lambda} \operatorname{id}_V = (F^\dagger)_{\overline{\lambda}}$, to z normalności F wynika normalność F_λ, więc można zastosować (1) do operatora F_λ. (3) Gdyż zawsze $\ker F^\dagger = (\operatorname{im} F)^\perp$.

514. Iloczyn skalarny pozwala wyróżnić wśród wszystkich rzutów $P = P^U_W \in \operatorname{End} V$ tzw. *rzuty prostopadłe (ortogonalne)*, tzn. takie, że $U = W^\perp$. Niech $P_W := P^{W^\perp}_W$ oznacza rzut prostopadły na podprzestrzeń $W \subset V$.

515. **Fakt**. Jeśli V jest unitarna lub euklidesowa, a $P \in \operatorname{End} V$ jest rzutem, to

$$(\ P \text{ jest ortogonalny, tzn. } \operatorname{im} P \perp \ker P) \iff P^\dagger = P.$$

$\boxed{\Rightarrow}$ Skoro $v - Pv \in \ker P$, $Pv' \in \operatorname{im} P$, to $0 = \langle v - Pv | Pv' \rangle = \langle v | (\operatorname{id} - P)^\dagger P v' \rangle$, więc z dowolności $v, v' \in V$ wynika, że $0 = (\operatorname{id} - P)^\dagger P$, tzn. $P = P^\dagger P$; stąd $P^\dagger = (P^\dagger P)^\dagger = P^\dagger P^{\dagger\dagger} = P$. $\boxed{\Leftarrow}$ $P = P^\dagger$ implikuje równość $(\operatorname{id} - P)^\dagger P = 0$, oznaczającą (jak przed chwilą widzieliśmy) prostopadłość $\operatorname{im}(\operatorname{id} - P)$ (czyli $\ker P$) do $\operatorname{im} P$.

Inny sposób. $P = P^\dagger \iff \forall v, v' \in V : \langle Pv | v' \rangle = \langle v | Pv' \rangle$; kładąc tu $P = P^U_W$, gdzie $V = W \dot{+} U$, dostajemy warunek $\forall (w, w' \in W,\ u, u' \in U) : \langle w | w' + u' \rangle = \langle w + u | w' \rangle$, tzn. $\langle w | u' \rangle = \langle u | w' \rangle$; ponieważ lewa i prawa strona zależą tu od innych niezależnych składowych, to jest on równoważny tożsamościowemu zerowaniu się obu stron, tzn. $W \perp U$, co przy $V = W \dot{+} U$ oznacza $U = W^\perp$.

Ćwiczenie. Dowieść że *sprzężeniem rzutu P^U_W jest rzut $P^{W^\perp}_{U^\perp}$*; jest to łatwe w oparciu o wzór $\ker P^\dagger = (\operatorname{im} P)^\perp$, a stanowi oczywiście uogólnienie poprzedniego faktu.

9.3 Twierdzenie spektralne

516. **Fakt**. Spektrum operatora hermitowskiego jest rzeczywiste, antyhermitowskiego — urojone, a unitarnego — zawarte w okręgu $\mathbf{U} = \mathbf{C}(\mathbf{1};\mathbf{1})$. Inaczej mówiąc, jeśli F jest: (1) hermitowski, to $\operatorname{Sp} F \subset \mathbb{R}$; (2) antyhermitowski, to $\operatorname{Sp} F \subset i\mathbb{R}$; (3) unitarny, to $\operatorname{Sp} F \subset \mathbf{U} = \{\lambda \in \mathbb{C}\colon |\lambda| = \mathbf{1}\}$.

Niech $\lambda \in \operatorname{Sp} F$ oraz $0 \neq v \in \ker F_\lambda$, tzn. $Fv = \lambda v$. (1) $\lambda \|v\|^2 = \langle v | Fv \rangle = \langle Fv | v \rangle = \overline{\lambda} \|v\|^2$, czyli $\|v\|^2(\lambda - \overline{\lambda}) = 0$, skąd $\overline{\lambda} = \lambda$. (2) $\lambda \|v\|^2 = \langle v | Fv \rangle = \langle -Fv | v \rangle = -\overline{\lambda} \|v\|^2$, czyli $\|v\|^2(\lambda + \overline{\lambda}) = 0$, skąd $\overline{\lambda} = -\lambda$. (3) F-niezmienniczość iloczynu skalarnego sprawia, że $\|v\| = \|Fv\| = \|\lambda v\| = |\lambda| \cdot \|v\|$, czyli $\|v\|(1 - |\lambda|) = 0$.

517. Dla operatora normalnego implikacje odwrotne do (1), (2) i (3) też są prawdziwe, lecz by tego dowieść, skorzystamy z *twierdzenia spektralnego*, które wykażemy dalej: Jeśli $F = \sum_{i=1}^{r} \lambda_i P_i$, gdzie $P_1, \ldots, P_r \in \operatorname{End} V$ tworzą 'rzutowy rozkład jedynki', przy czym $P_i^\dagger = P_i$, to $F^\dagger = \sum_{i=1}^{r} \overline{\lambda}_i P_i$, więc F jest normalny; otóż tw. spektralne mówi, że każdy operator normalny ma taki rozkład spektralny. Możemy ponadto założyć,

że wszystkie rzuty P_i są $\neq 0$, wtedy $\operatorname{Sp} F = \{\lambda_1, \ldots, \lambda_r\}$, a zatem

(1) $F^\dagger = F \iff \forall i : \overline{\lambda}_i = \lambda_i \iff \operatorname{Sp} F \subset \mathbb{R}$; (2) $F^\dagger = -F \iff \operatorname{Sp} F \subset \imath\mathbb{R}$ analogicznie; (3) $F^\dagger F = \sum\limits_i (\overline{\lambda}_i \lambda_i) P_i = \mathrm{id}_V \iff \forall i : \overline{\lambda}_i \lambda_i = 1 \iff \operatorname{Sp} F \subset \mathbf{U}$.

518. **Fakt**. Jeśli operator $F \in \operatorname{End} V$ (w przestrzeni unitarnej lub euklidesowej) jest normalny, to $V(\lambda^*) = V_1(\lambda^*)$ dla każdej wartości własnej λ^*.

Wykażemy najpierw, że $\boxed{F^k(v) = 0 \;\Rightarrow\; F(v) = 0}\,(Z_k)$ dla każdego $k \geqslant 2$. Dla $k = 2$ mamy $\|F^2(v)\| = \|F(w)\| = \|F^\dagger(w)\| = \|F^\dagger F(v)\|$, gdzie $w = F(v)$, więc $F^2(v) = 0 \;\Rightarrow\; 0 = \langle v | F^\dagger F(v)\rangle = \langle F(v) | F(v)\rangle = \|F(v)\|^2$, co dowodzi (Z_2). Stąd dla $k \geqslant 2$, kładąc $w = F^{k-2}(v)$, dostajemy $F^k(v) = 0 \;\Rightarrow\; F^2(w) = 0 \;\Rightarrow\; F(w) = 0 \;\Rightarrow\; F^{k-1}(w) = 0$, co indukcyjnie dowodzi implikacji (Z_k).

Tezę $V(\lambda^*) = V_1(\lambda^*)$ dostajemy natychmiast, stosując (Z_k) do $F_{\lambda^*} = F - \lambda^* \mathrm{id}_V$.

519. **Twierdzenie** (*'spektralne' dla operatora normalnego*). Jeśli F jest operatorem normalnym w przestrzeni unitarnej V, to V jest *ortogonalną* sumą prostą przestrzeni własnych F, a więc F jest diagonalizowalny; ponadto V ma ortonormalną bazę wektorów własnych F, zaś rzuty P_i w rozkładzie spektralnym $F = \sum\limits_{i=1}^{r} \lambda_i P_i$ są hermitowskie: $P_i^\dagger = P_i$.

V jest sumą prostą przestrzeni pierwiastkowych, a więc — wobec powyższego faktu — przestrzeni własnych F. Ortogonalność $V(\lambda^*)$: jeśli $v_1 \in V(\lambda_1)$, $v_2 \in V(\lambda_2)$, tzn. $F(v_i) = \lambda_i v_i$, to także $F^\dagger v_i = \overline{\lambda}_i v_i$, więc $\lambda_2 \langle v_1 | v_2\rangle = \langle v_1 | F v_2\rangle = \langle F^\dagger v_1 | v_2\rangle = \langle \overline{\lambda}_1 v_1 | v_2\rangle = \lambda_1 \langle v_1 | v_2\rangle$, czyli $(\lambda_2 - \lambda_1)\langle v_1 | v_2\rangle = 0$, tzn. $v_1 \perp v_2$ przy $\lambda_1 \neq \lambda_2$.

Jasne stąd, że ortogonalną bazę złożoną z wektorów własnych F dostaniemy jako konkatenację ortonormalnych baz wszystkich przestrzeni własnych $V(\lambda_i)$, $i \in \overline{1,r}$.

Hermitowskość rzutów P_i wynika wprost z ortogonalności przestrzeni $V(\lambda_i)$.

520. **Uwaga**. Operator normalny w przestrzeni euklidesowej (nad $\mathbb{K} = \mathbb{R}$) nie musi być diagonalizowalny (np. niediagonalizowalne są nietrywialne obroty w $\mathbb{R}^2$ i w $\mathbb{R}^3$); jednakże także tu operatory normalne są 'półproste':

521. **Fakt**. Jeśli operator $F \in \operatorname{End} V$ (w przestrzeni unitarnej lub euklidesowej) jest normalny, a podprzestrzeń $W \subset V$ jest F-niezmiennicza, to $W^\perp$ jest także F-niezmiennicza, zaś operator $F|_W$ jest normalny.

Gdy V jest sumą prostą podprzestrzeni W, U, wtedy każdy operator $F \in \operatorname{End} V$ można 'rozłożyć na bloki' $A \in \operatorname{End} W$, $B \in \mathrm{L}(U;W)$, $C \in \mathrm{L}(W;U)$, $D \in \operatorname{End} U$, pisząc $F = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, przez co należy rozumieć, że

$$F(w+u) = \underbrace{A(w) + B(u)}_{\in W} + \underbrace{C(w) + D(u)}_{\in U}$$

dla każdej pary $w \in W$, $u \in U$. Łatwo się przekonać, że na takich 'rozkładach blokowych' można operować (dodawać i 'mnożyć') tak samo, jak na liczbowych 2×2-macierzach, dbając jednakże o właściwą kolejność przy składaniu operatorów.

Weźmy teraz $U = W^\perp$; wtedy zachodzi wzór $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^\dagger = \begin{bmatrix} A^\dagger & C^\dagger \\ B^\dagger & D^\dagger \end{bmatrix}$, wynikający bez trudu z tożsamości $\langle w+u | w'+u'\rangle = \langle w | w'\rangle + \langle u | u'\rangle$.

Załóżmy wreszcie, że W jest F-niezmiennicza, $F(W) \subset W$, co oczywiście oznacza,

że $C = 0$. Pisząc warunek $FF^\dagger = F^\dagger F$ i przyrównując lewe-górne bloki dostajemy warunek $AA^\dagger + BB^\dagger = A^\dagger A$. Obie strony są tu operatorami w W, możemy więc obliczyć ich ślady; korzystając z własności $\mathrm{tr}(AH) = \mathrm{tr}(HA)$ dostajemy stąd $\mathrm{tr}(BB^\dagger) = 0$. Lecz $\mathrm{tr}\, H = \sum_{i=1}^{k} \langle e_i | He_i \rangle$ dla dowolnej ortonormalnej bazy $e_1, \ldots, e_k$ w W i operatora $H \in \mathrm{End}\, W$ (gdyż $\langle e_i | \cdot \rangle \in W^*$ dają bazę sprzężoną względem e_i); dla $H = BB^\dagger$ mamy przy tym $\langle e_i | He_i \rangle = \|B^\dagger e_i\| \geqslant 0$, więc $\mathrm{tr}(BB^\dagger) = 0$ implikuje $\forall i : B^\dagger(e_i) = 0$, tzn. $B = 0$, co oznacza F-niezmienniczość $W^\perp$.

W konsekwencji dostajemy również $AA^\dagger = A^\dagger A$, czyli $A = F|_W$ jest normalny.

522. **Uwaga**. Dla każdego z poniższych dodatkowych założeń (a),(b),(c) można łatwo napisać znacznie krótszy dowód, nie korzystający z rozkładu blokowego ani śladu:

(a) $\exists \lambda : W \subset \ker F_\lambda$; (b) $F^\dagger = F$; (c) $F^\dagger = F^{-1}$.

Ad(a). $(F_\lambda)^\dagger = (F - \lambda\,\mathrm{id})^\dagger = F^\dagger - \overline{\lambda}\,\mathrm{id} = (F^\dagger)_{\overline{\lambda}}$, więc $\big(F \text{ normalny} \Rightarrow F_\lambda \text{ normalny}\big)$ oraz $F(w) = \lambda w \;\Rightarrow\; F^\dagger(w) = \overline{\lambda} w$; mamy stąd $\big(v \in W^\perp,\, w \in W\big) \Rightarrow \langle Fv | w \rangle = \langle v | F^\dagger w \rangle = \overline{\lambda}\langle v | w \rangle = 0$, czyli $F(W^\perp) \subset W^\perp$. *Ad*(b). Jeśli $v \in W^\perp$, $w \in W$, to $\langle Fv | w \rangle = \langle v | Fw \rangle = 0$, bo $Fw \in W$; stąd teza. Inny sposób: Zauważmy, że $\boxed{v \in (FW)^\perp \;\Rightarrow\; F^\dagger v \in W^\perp}$, więc $\boxed{F(W) \subset W \Rightarrow\; F^\dagger(W^\perp) \subset W^\perp}$ dla *każdych* $W \subset V$, $F \in \mathrm{End}\, V$. *Ad*(c). Dla F unitarnego $F(W) \perp F(W^\perp)$ (bo $\langle F(w) | F(v) \rangle = \langle w | v \rangle$) oraz $\dim F(W) = \dim W$, $\dim F(W^\perp) = \dim W^\perp$; stąd $F(W^\perp) = F(W)^\perp$.

10 Uzupełnienia

10.1 Appendix A: Baza V złożona z N-serii

Opiszemy teraz konstrukcję bazy złożonej z N-serii dla danego operatora nilpotentnego N. Zastępując w tej konstrukcji V przestrzenią pierwiastkową $V(\mu, F)$, a N — operatorem F_μ (obciętym do $V(\mu, F)$) dla μ przebiegających całe $\operatorname{Sp} F$, dostaje się konstrukcję tzw. 'bazy jordanowskiej operatora F'.

523. **Definicja** (*baza przestrzeni względem podprzestrzeni*). Niech V_0 będzie podprzestrzenią V. Będziemy mówić, że układ wektorów $u_1, \ldots, u_r \in V$

I *rozpina V względem V_0*, jeśli **każdy** wektor $v \in V$ ma rozkład

$$v = \sum_{i=1}^{r} u_i \lambda^i + u_0, \quad \lambda^i \in \mathbb{K},\ u_0 \in V_0; \tag{*}$$

II jest *liniowo niezależny względem V_0*, jeśli wektor zerowy $v = 0$ ma **jednoznaczny** (tzn. jedynie trywialny: $\lambda^i = 0$) rozkład (*);

III jest *bazą V względem V_0*, jeśli spełnione są oba warunki I i II, tzn. jeśli **każdy** $v \in V$ ma **jednoznaczny** rozkład (*). ([93])

Łatwo sprawdzić, że pojęcia te mają oczekiwane własności, np. układ lin. niezależny względem V_0 można dopełnić do bazy V względem V_0.

524. **Konstrukcja bazy złożonej z N-serii.** Dany operator $N \in \operatorname{End} V$ określa łańcuch $V_0 \subset V_1 \subset V_2, \ldots$, gdzie $V_0 := \{0\}$, $V_k := \ker N^k$, $k \in \mathbb{N}$. Zauważmy najpierw, że jeśli $k \geqslant 1$ i dany układ $u_1, \ldots, u_r$ jest liniowo niezależny względem V_k, to jego N-obraz, tj. układ $N(u_1), \ldots, N(u_r)$, jest liniowo niezależny względem V_{k-1}.

Jeśli $\sum_i N(u_i)\lambda^i \in V_{k-1}$, to działając na to N^{k-1} dostajemy $N^k\left(\sum_i u_i\lambda^i\right) = 0$, więc $\sum_i u_i\lambda^i \in V_k$, to zaś — z założenia o układzie $u_1, \ldots, u_r$ — implikuje $\lambda^i = 0$.

Jeśli na dodatek $u_i \in V_{k+1}$, to oczywiście $N(u_i) \in V_k$. Wybierzmy teraz (dla ustalonego $k \geqslant 1$) dowolną (nieuporządkowaną) bazę — oznaczmy ją $\mathcal{B}_{(k+1)}$ — przestrzeni V_{k+1} wzgl. V_k; jej N-obraz jest układem w V_k, lin. niezależnym wzgl. V_{k-1}, można więc dopełnić go do pewnej bazy $\mathcal{B}_{(k)}$ dla V_k wzgl. V_{k-1}; postępując w taki sposób otrzymamy po k krokach bazę $\mathcal{B}_{(1)}$ dla V_1 wzgl. $V_0 = \{0\}$, tzn. 'zwykłą' bazę V_1.

Zbiór $\mathcal{B} := \mathcal{B}_{(1)} \cup \mathcal{B}_{(2)} \cup \ldots \cup \mathcal{B}_{(k+1)}$ składa się oczywiście z N-serii, a ponadto rozpina V_{k+1}; pokażemy teraz, że jest on liniowo niezależny:

Niech $\sum_{v \in \mathcal{B}} \lambda(v)v = 0$ dla pewnego zestawu współczynników $\lambda(\,\cdot\,) : \mathcal{B} \to \mathbb{K}$. Skoro N^k zeruje się na $\mathcal{B}_{(j)}$ dla $j \in \overline{1,k}$, więc działając operatorem N^k otrzymamy równość $\sum_{v \in \mathcal{B}_{(k+1)}} \lambda(v)N^k(v) = 0$. Otóż liniowa niezależność względem V_k zbioru $\mathcal{B}_{(k+1)}$ implikuje liniową niezależność względem $V_0 = \{0\}$ (tzn. zwykłą) zbioru $N^k(\mathcal{B}_{(k+1)})$. Zatem z ostatniej równości wynika znikanie współczynników $\lambda(v)$ dla $v \in \mathcal{B}_{(k+1)}$. Powtarzając k-krotnie ten wywód dostajemy tezę. Wobec tego dowiedliśmy

[93] Równoważne określenie: układ $u_1, \ldots, u_r \in V$ jest 'taki, siaki lub owaki' *względem V_0*, jeśli układ $u_1 + V_0, \ldots, u_r + V_0$ (wektorów przestrzeni ilorazowej V/V_0) jest ... w V/V_0.

Twierdzenie. Jeśli operator N jest nilpotentny, zaś k jest (najlepiej minimalną) liczbą taką, że $N^{k+1}=0$, tj. $V_{k+1}=V$, to powyżej skonstruowany zbiór $\mathcal{B}=\mathcal{B}_{(1)}\cup\ldots\cup\mathcal{B}_{(k+1)}$ jest bazą V, złożoną z N-serii.

525. Oczywiste uogólnienie: jeśli zamiast nilpotentności N założymy jedynie, że $\ker N\neq\{0\}$, to biorąc minimalne k takie, że $V_{k+1}=V(0,N)$, dostaniemy złożoną z N-serii bazę przestrzeni pierwiastkowej $V(0,N)$.

526. **Ćwiczenie**. Niech $V'\subset V$ ma opis $V'=\{v:\phi_1(v)=\ldots=\phi_s(v)=0\}$, zaś $V''\subset V'$ — opis $V''=\{v\in V':\psi_1(v)=\ldots=\psi_r(v)=0\}$, gdzie dane równania, tzn. elementy $\phi_1,\ldots,\phi_s,\psi_1,\ldots,\psi_r\in V^*$, są liniowo niezależne. Sprawdzić, że wówczas wektory $v_1,\ldots,v_r\in V'$ tworzą bazę V' względem V'' wtedy i tylko wtedy, gdy macierz kwadratowa $\big[\psi_i(v_j)\big]$ jest nieosobliwa. [Kryterium to jest użyteczne przy konstruowaniu 'bazy jordanowskiej' danego operatora w przestrzeni $V=\mathbb{K}^n$].

10.2 Appendix B: Rekurencje liniowe jednorodne stopnia d o stałych współczynnikach

527. Dla danych $d\in\mathbb{N}$ oraz $a_0,\ldots,a_{d-1}\in\mathbb{K}$ rozważymy następujące równanie rekurencyjne na ciąg $\boldsymbol{x}=(x_0,x_1,x_2,\ldots)\in\mathbb{K}^{\mathbb{Z}_+}$:

$$\forall n\in\mathbb{Z}_+:\ x_{n+d}=a_0x_n+a_1x_{n+1}+\ldots+a_{d-1}x_{n+d-1};\qquad\text{(R)}$$

jest widoczne, że dowolnie zadane wartości $x_0,\ldots,x_{d-1}\in\mathbb{K}$ całkowicie określają ciąg $\boldsymbol{x}$ spełniający rekurencje (R), więc *przestrzeń rozwiązań*

$$W:=\{\boldsymbol{x}\in\mathbb{K}^{\mathbb{Z}_+}:\text{spełnione są równania (R)}\}$$

stanowi d-wymiarową podprzestrzeń przestrzeni $\mathbb{K}^{\mathbb{Z}_+}$.

Określmy operator $S\in\operatorname{End}\mathbb{K}^{\mathbb{Z}_+}$ wzorem $S(x_0,x_1,\ldots):=(x_1,x_2,\ldots)$, tzn. $(S\boldsymbol{x})_n=x_{n+1}$, wtedy (R) $\iff S^d\boldsymbol{x}=a_0\boldsymbol{x}+a_1S\boldsymbol{x}+\ldots+a_{d-1}S^{d-1}\boldsymbol{x}$ $\iff w(S)\boldsymbol{x}=0$, gdzie $\boxed{w(\lambda):=\lambda^d-a_0-a_1\lambda-\ldots-a_{d-1}\lambda^{d-1}}$; zatem $\boxed{W=\ker w(S)}$; z tego względu wielomian $w(\lambda)$ nazywamy *wielomianem charakterystycznym* rekurencji (R). Założymy tutaj, że $w(\lambda)$ *ma rozkład na czynniki stopnia 1 w* $\mathbb{K}[\,\cdot\,]$; jest tak zawsze, gdy $\mathbb{K}=\mathbb{C}$.

528. Znajdziemy pewną bazę, a więc także parametryzację przestrzeni W. Niech $\boldsymbol{g}(\lambda):=(1,\lambda,\lambda^2,\ldots)\in\mathbb{K}^{\mathbb{Z}_+}$; jest to ciąg geometryczny o ilorazie $\lambda\in\mathbb{K}$. Zauważmy, że $S\boldsymbol{g}(\lambda)=\lambda\boldsymbol{g}(\lambda)$, więc także $S^k\boldsymbol{g}(\lambda)=\lambda^k\boldsymbol{g}(\lambda)$ dla $k\in\mathbb{N}$ (przez indukcję); zatem $\boxed{\phi(S)\boldsymbol{g}(\lambda)=\phi(\lambda)\boldsymbol{g}(\lambda)}$ dla dowolnego wielomianu ϕ. Niech $\boldsymbol{g}'(\lambda)=(0,1,2\lambda,3\lambda^2,\ldots)$, $\boldsymbol{g}''(\lambda)=(0,0,2,6\lambda,\ldots)$, ogólnie $\boldsymbol{g}^{(s)}(\lambda)$, oznaczają pochodne odwzorowania $\lambda\mapsto\boldsymbol{g}(\lambda)$.

529. **Fakt**. Zbiór $\big\{\boldsymbol{g}^{(s)}(\lambda):\ s\in\mathbb{Z}_+,\ \lambda\in\mathbb{K}\big\}\subset\mathbb{K}^{\mathbb{Z}_+}$ jest liniowo niezależny.

Załóżmy, że $\sum_{j=1}^{r}\sum_{s=0}^{m}x_{j,s}\boldsymbol{g}^{(s)}(\lambda_j)=0$ dla pewnych parami różnych $\lambda_1,\ldots,\lambda_r\in\mathbb{K}$ oraz pewnego zestawu współczynników $\big(x_{j,s}\big)_{(j,s)\in\overline{1,r}\times\overline{0,m}}$. Z określenia $\boldsymbol{g}(\lambda)$ oznacza to, że $\forall n\geqslant 0:\ \sum_{j,s}x_{j,s}\frac{d^s}{d\lambda^s}\big|_{\lambda=\lambda_j}\lambda^n=0$; wynika stąd, że dla dowolnego wielomianu

$\phi(\lambda)$, tzn. kombinacji liniowej λ^n, mamy $\sum_{j,s} x_{j,s}\phi^{(s)}(\lambda_j) = 0$. Weźmy teraz $\phi(\lambda) = (\lambda-\lambda_1)^k(\lambda-\lambda_2)^{m+1}\ldots(\lambda-\lambda_r)^{m+1}$; wtedy to, że $\frac{d^k}{d\lambda^k}\big|_{\lambda=a}(\lambda-a)^n \neq 0 \;\Leftrightarrow\; k=n$, oraz wzór Leibniza $\big[\phi_1\phi_2\big]^{(k)} = \sum \binom{k}{l}\phi_1^{(l)}\phi_2^{(k-l)}$, dają $\phi^{(k)}(\lambda_1) \neq 0$, $\phi^{(s)}(\lambda_1) = 0$ dla $s \in \overline{0,k-1}$, a także $\phi^{(s)}(\lambda_j) = 0$ dla $(j,s) \in \overline{2,r} \times \overline{0,m}$; wobec tego biorąc kolejno $k = m, m-1, \ldots, 0$ dostaniemy $x_{1,m} = x_{1,m-1} = \ldots = x_{1,0} = 0$. W taki sam sposób pokazujemy, że $\forall s \in \overline{0,m}:\; x_{2,s} = \ldots = x_{r,s} = 0$, QED.[94]

530. **Wniosek**. Jeśli $\lambda_1, \ldots, \lambda_r$ jest układem wszystkich (bez powtórzeń) pierwiastków wielomianu $w(\lambda)$, zaś $k_1, \ldots, k_r \in \mathbb{N}$ — ich krotnościami, to wektory (tj. ciągi)

$$\underbrace{\boldsymbol{g}(\lambda_1), \boldsymbol{g}'(\lambda_1), \ldots}_{k_1},\; \ldots,\; \underbrace{\boldsymbol{g}(\lambda_r), \boldsymbol{g}'(\lambda_r), \ldots}_{k_r} \qquad (*)$$

tworzą bazę przestrzeni $W = \ker w(S)$.

Ponieważ wiemy już, że są one liniowo niezależne, a ich liczba $k_1 + \ldots + k_r = d$ jest równa $\dim W$, więc wystarczy sprawdzić, że wektory (*) należą do W. Zrobimy to, nie używając wprost wzoru na $\boldsymbol{g}(\lambda)$, a jedynie własności $\boldsymbol{g}(\lambda) \in \ker(S-\lambda)$:

Różniczkowanie k-krotne $w(S)\boldsymbol{g}(\lambda) = w(\lambda)\boldsymbol{g}(\lambda)$ daje $w(S)\boldsymbol{g}^{(k)}(\lambda) = \sum_{j=0}^{k}\binom{k}{j}w^{(j)}(\lambda)\cdot \boldsymbol{g}^{(k-j)}(\lambda)$, więc $\left(\begin{matrix}\forall j \in \overline{0,k}:\; w^{(j)}(\lambda_1) = 0, \text{ czyli } \lambda_1 \text{ jest} \\ \text{pierwiastkiem } w(\lambda) \text{ krotności} \geqslant k+1\end{matrix}\right) \Rightarrow\; \boldsymbol{g}^{(k)}(\lambda_1) \in \ker w(S)$.

Inny sposób. Przez indukcję wykażemy zaraz, że $\boxed{\boldsymbol{g}^{(k-1)}(\lambda) \in \ker(S-\lambda)^k}\,(Z_k)$. Z tego wzoru wynika, że jeśli λ_1 jest pierwiastkiem krotności $\geqslant k$ wielomianu $w(\lambda)$, to $\boldsymbol{g}^{(k-1)}(\lambda_1) \in W = \ker w(S)$, gdyż wtedy $w(S) = \tilde{w}(S)(S-\lambda_1)^k$.

Dowód '$(Z_k) \Rightarrow (Z_{k+1})$': Jeśli $(S-\lambda)^k\boldsymbol{g}^{(k-1)}(\lambda) = 0$, to $(S-\lambda)^{k+1}\boldsymbol{g}^{(k-1)}(\lambda) = 0$; różniczkując to dostajemy $(k+1)\cdot 0 + (S-\lambda)^{k+1}\boldsymbol{g}^{(k)}(\lambda) = 0$, czyli mamy (Z_{k+1}).

531. Zatem każde rozwiązanie rekurencji (R) można przedstawić w postaci

$$x_n = C_{1,1}\lambda_1^n + C_{1,1}n\lambda^{n-1} + \ldots + C_{r,0}\lambda_r^n + C_{r,1}n\lambda_r^{n-1} + \ldots,$$

przy czym wartości d stałych $C_{j,s}$, $j \in \overline{1,r}$, $s \in \overline{0,k_j-1}$, da się jednoznacznie określić, znając np. wartości początkowe $x_0, \ldots, x_{d-1}$.

532. **Przykład**. Dla rekurencji $x_{n+3} = 5x_{n+2} - 8x_{n+1} + 4x_n$ wielomian charakterystyczny $w(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda-1)(\lambda-2)^2$ ma pierwiastki $\lambda_1 = 1$, $\lambda_2 = 2$ z krotnościami $k_1 = 1$, $k_2 = 2$; zatem rozwiązanie ogólne ma postać $x_n = a + b\cdot 2^n + c\cdot n2^{n-1}$. Dla określenia wartości współczynników przez wartości początkowe x_0, x_1, x_2 należy rozwiązać układ $\left\{\begin{matrix} a+b &= x_0 \\ a+2b+c &= x_1 \\ a+4b+4c &= x_2\end{matrix}\right.$; dostajemy wtedy $\left\{\begin{matrix} a &= 4x_0 - 4x_1 + x_2 \\ b &= -3x_0 + 4x_1 - x_2 \\ c &= 2x_0 - 3x_1 + x_2\end{matrix}\right.$,

[94] Inny dowód opiera się na pojęciu i własności przestrzeni pierwiastkowych operatora. Dla ustalonych $\lambda_1, \ldots, \lambda_r$ oraz $m \in \mathbb{N}$ weźmy $V = \ker v(S)$, gdzie $v(\lambda) := \prod_{j=1}^{r}(\lambda-\lambda_j)^{m+1}$; jest to skończenie wymiarowa przestrzeń niezmiennicza względem S, więc możemy określić $F := S\big|_W \in \operatorname{End} W$. Otóż rachunek przeprowadzony w poniższym dowodzie pokazuje, że $\boldsymbol{g}^{(s)}(\lambda_j)$ dla $s \in \overline{0,m}$ należą do $\ker(F - \lambda_j \operatorname{id}_V)^{m+1}$, tzn. do przestrzeni pierwiastkowej $V(\lambda_j, F)$. Stąd, wobec liniowej niezależności przestrzeni $V(\lambda_j, F)$, wystarczy sprawdzić liniową niezależność $\boldsymbol{g}^{(s)}(\lambda_j)$ dla ustalonego $\lambda = \lambda_j$. Jest to bardzo proste: jeśli $\lambda \neq 0$, to $\frac{1}{\lambda^n}\boldsymbol{g}^{(s)}(\lambda)_n = \frac{n(n-1)\ldots(n-s+1)}{\lambda^s} =$ (wielomian stopnia s względem n), zaś $\boldsymbol{g}^{(s)}(0)_n = n!\delta_n^s$.

więc rozwiązaniem rekurencji jest $x_n = 2^{n-1}\big[n(2x_0-3x_1+x_2)-6x_0+8x_1-2x_2\big]+$ $+(4x_0-4x_1+x_2) = [4+2^n(n-3)]x_0+[-4+2^{n-1}(8-3n)]x_1+[1+2^{n-1}(n-2)]x_2$.

533. **Uwaga 1**. Jeśli oznaczymy $W(\lambda) := \big\langle \boldsymbol{g}(\lambda), \boldsymbol{g}'(\lambda), \ldots, \boldsymbol{g}^{(k)}(\lambda), \ldots \big\rangle \subset \mathbb{K}^{\mathbb{Z}_+}$, to

$$W(\lambda) = \big\{(\varphi(0), \lambda\varphi(1), \lambda^2\varphi(2), \lambda^3\varphi(3), \ldots) : \ \varphi \in \mathbb{K}[\,\cdot\,]\big\} \ \text{ dla } \ \lambda \neq 0,$$

gdyż $\frac{1}{\lambda^n}\big(\boldsymbol{g}^{(k)}(\lambda)\big)_n = \frac{1}{\lambda^n}\frac{d^k}{d\lambda^k}\lambda^n = \frac{k!}{\lambda^k}\binom{n}{k} =$ (funkcja wielomianowa względem n); natomiast oczywiście $W(0) = \{\boldsymbol{x} : \ x_n = 0 \text{ dla p.w. } n \in \mathbb{N}\}$.

Uwaga 2. Gdy $w(0) = 0$, wtedy w ma rozkład $w(\lambda) = \tilde{w}(\lambda)\cdot\lambda^k$, gdzie $k \in \mathbb{N}$, $\tilde{w}(0) \neq 0$, więc $\ker w(S) = \{\boldsymbol{x} : \ S^k\boldsymbol{x} \in \ker\tilde{w}(S)\}$. Zatem opis $W = \ker w(S)$ łatwo uzyskać z opisu $\tilde{W} = \ker\tilde{w}(S)$, mianowicie: *$W$ składa się z tych ciągów, które po opuszczeniu k początkowych wyrazów dają element z $\tilde{W}$.* ([95])

Zatem bez istotnej straty ogólności można założyć, że $w(0) \neq 0$, tzn. że $\forall j : \ \lambda_j \neq 0$; zauważmy w związku z tym, że jeśli $\lambda \neq 0$, to $\big($ciąg $\boldsymbol{x}$ jest kombinacją liniową $\boldsymbol{g}(\lambda), \ldots, \boldsymbol{g}^{(k-1)}(\lambda)\big) \iff \big($ma postać $x_n = \lambda^n\varphi(n)$, gdzie $\varphi(\,\cdot\,) \in \mathbb{K}_{k-1}[\,\cdot\,]\big)$.

[95] Inne uzasadnienie: $a_0 = \ldots = a_{k-1} = 0$, więc w rekurencji (R) nie ma $x_0, \ldots, x_{k-1}$.