Rozdział 1. Liczby zespolone

1.1. Definicja i podstawowe własności.

Oznaczmy przez $\mathbb{R}^2$ iloczyn (produkt) kartezjański zbioru liczb rzeczywistych $\mathbb{R}$ przez siebie, t.j., zbiór par uporządkowanych (a, b), gdzie $a, b \in \mathbb{R}$. W $\mathbb{R}^2$ wprowadzamy działania dodawania i mnożenia:

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$
$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 \cdot a_2 - b_1 \cdot b_2, a_1 \cdot b_2 + b_1 \cdot a_2).$$

Zbiór $\mathbb{R}^2$ z tak określonymi działaniami będziemy oznaczać $\mathbb{C}$, a jego elementy będziemy nazywać *liczbami zespolonymi.*

DEFINICJA 1.1. *Niech para $z = (a, b)$ będzie liczbą zespoloną. Liczbę (rzeczywistą!)* $\operatorname{Re} z = a$ *nazywamy* częścią rzeczywistą *liczby zespolonej z, a liczbę (rzeczywistą!)* $\operatorname{Im} z = b$ częścią urojoną *liczby zespolonej z.*

Tak określone działania mają szereg własności, analogicznych do własności działań dodawania i mnożenia w zbiorze liczb rzeczywistych. Mamy więc dla z, z_1, z_2 i $z_3 \in \mathbb{C}$

(1) przemienność dodawania i przemienność mnożenia

$$z_1 + z_2 = z_2 + z_1, \ z_1 \cdot z_2 = z_2 \cdot z_1,$$

(2) łączność dodawania i łączność mnożenia

$$(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3), \ (z_1 + z_2) + z_3 = z_1 + (z_2 + z_3),$$

(3) rozdzielność mnożenia względem dodawania

$$(z_1 + z_2) \cdot z = z_1 \cdot z + z_2 \cdot z,$$

(4) istnienie elementów neutralnych ze względu na dodawanie i mnożenie

$$z + (0, 0) = z, \ (1, 0) \cdot z = z,$$

(5) dla każdego z istnieje element $-z$, odwrotny ze względu na dodawanie, tzn., $z + (-z) = (0, 0)$,

$$-z = (-a, -b) = (-1, 0) \cdot (a, b),$$

(6) dla każdego $z \neq (0, 0)$ istnieje element z^{-1}, odwrotny ze względu na mnożenie, tzn., $z \cdot (z^{-1}) = (1, 0)$

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

Zauważamy natychmiast, że

$$z = (a, b) = (a, 0) + (b, 0) \cdot (0, 1).$$

Zdefiniujmy odwzorowanie $\mathbb{R} \to \mathbb{C}\colon a \mapsto (a, 0)$. Pozwala ono utożsamiać liczby rzeczywiste z liczbami zespolonymi o zerowej części urojonej: liczbę rzeczywistą a utożsamiamy z liczbą zespoloną $(a, 0)$. Utożsamienie to zachowuje działania dodawania i mnożenia

$$\begin{aligned} a + b &\mapsto (a + b, 0) = (a, 0) + (b, 0), \\ a \cdot b &\mapsto (ab, 0) = (a, 0) \cdot (b, 0) \end{aligned}$$

Korzystając z tak wprowadzonego utożsamienia możemy napisać

$$z = (a, b) = (a, 0) + (0, b) = a + b \cdot \iota, \text{gdzie } \iota = (0, 1).$$

Zauważmy, że $\iota^2 = (0, 1) \cdot (0, 1) = -1$.

DEFINICJA 1.2. Liczbą sprzężoną *do liczby zespolonej* $z = a + b\iota$, $a, b \in \mathbb{R}$, *nazywamy liczbę zespoloną* $\bar{z} = a - b\iota$, *tzn.*, $\operatorname{Re} \bar{z} = \operatorname{Re} z$ *oraz* $\operatorname{Im} \bar{z} = -\operatorname{Im} z$.

Oczywiste, że $\bar{\bar{z}} = z$.

STWIERDZENIE 1.3. *Zachodzą związki*

$$\begin{aligned} \overline{z_1 + z_2} &= \bar{z}_1 + \bar{z}_2, \\ \overline{z_1 z_2} &= \bar{z}_1 \bar{z}_2, \\ \frac{1}{\bar{z}} &= \overline{\left(\frac{1}{z}\right)}. \end{aligned}$$

DOWÓD: Sprawdzamy prostym rachunkiem

$$\begin{aligned} \overline{z_1 + z_2} &= (a_1 + a_2) - \iota(b_1 + b_2) = (a_1 - \iota b_1) + (a_2 - \iota b_2) = \bar{z}_1 + \bar{z}_2, \\ \overline{z_1 z_2} &= a_1 a_2 - b_1 b_2 - \iota(a_1 b_2 + a_2 b_1) = (a_1 - \iota b_1) \cdot (a_2 - \iota b_2) = \bar{z}_1 \bar{z}_2, \\ z\left(\frac{1}{z}\right) &= 1, \text{ czyli, z poprzedniego, } \bar{z}\overline{\left(\frac{1}{z}\right)} = 1, \end{aligned}$$

co oznacza $\frac{1}{\bar{z}} = \overline{\left(\frac{1}{z}\right)}$. ■

Wnioski:

$$\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2 \quad \text{oraz} \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2},$$

$$\operatorname{Re} z = \frac{z + \bar{z}}{2} \quad \text{oraz} \quad \operatorname{Im} z = \frac{z - \bar{z}}{2\iota}.$$

Jeśli $z = -\bar{z}$, to z jest liczbą czysto urojoną. Jeśli $z = \bar{z}$, to z jest liczbą rzeczywistą.

1.2. Interpretacja geometryczna.

$\mathbb{R}^2$ możemy uważać za płaszczyznę z wyróżnionymi osiami współrzędnych. Dodawanie liczb zespolonych ma prostą interpretację dodawania wg reguły równoległoboku:

Interpretację geometryczną mnożenia można łatwo otrzymać z punktów 1 i 4 poniższego stwierdzenia 1.5.

Zacznijmy od tego, że punkty na płaszczyźnie można opisywać współrzędnymi biegunowymi (r, φ). Para liczb rzeczywistych (r, φ), $r \geqslant 0$, opisuje liczbę $z = a + b\iota \in \mathbb{C}$, jeśli $a = r\cos\varphi,\ b = r\sin\varphi$. Zauważmy, że wówczas (o ile $z \neq 0$)

$$r = \sqrt{a^2 + b^2},\ \cos\varphi = \frac{a}{\sqrt{a^2+b^2}},\ \sin\varphi = \frac{a}{\sqrt{a^2+b^2}}. \tag{1.1}$$

DEFINICJA 1.4. *Jeśli $z = a + b\iota$, gdzie $a, b \in \mathbb{R}$, to liczbę rzeczywistą $|z| = \sqrt{a^2 + b^2}$ nazywamy* modułem *liczby z; jeśli ponadto $z \neq 0$ lub, równoważnie, $|z| \neq 0$, to $\varphi \in \mathbb{R}$ takie, że*

$$\cos\varphi = \frac{a}{\sqrt{a^2+b^2}},\ \sin\varphi = \frac{b}{\sqrt{a^2+b^2}}$$

nazywamy argumentem *liczby z i oznaczamy* $\arg z$.

Oczywiste, że argument liczby zespolonej jest wyznaczony z dokładnością do wielokrotności 2π. Jeśli argument φ liczby z spełnia nierówność $0 \leq \varphi < 2\pi$, to nazywamy go *argumentem głównym* i oznaczamy $\operatorname{Arg} z$. argument główny liczby zespolonej Z wzorów (1.1) wynikają równości

$$a = r\cos\varphi, \quad b = r\sin\varphi$$

i, w konsekwencji,

$$z = r(\cos\varphi + \iota\sin\varphi). \tag{1.2}$$

Powyższe przedstawienie liczby zespolonej z nazywamy *reprezentacją trygonometryczną* lub *rozkładem biegunowym* liczby zespolonej z.

STWIERDZENIE 1.5. *Zachodzą związki:*

(1) $|z|^2 = z\bar{z}$,
(2) $|z_1 z_2| = |z_1| \cdot |z_2|$,
(3) $|\bar{z}| = |z|$,
(4) $\dfrac{|z_1|}{|z_2|} = \left|\dfrac{z_1}{z_2}\right|$ *dla* $z_2 \neq 0$,
(5) $\arg(z_1 z_2) = \arg z_1 + \arg z_2$,
(6) $\arg \bar{z} = -\arg z$.

DOWÓD:

(1) Niech $z = a + \imath b$, gdzie $a, b \in \mathbb{R}$. Wówczas

$$z\bar{z} = (a + \imath b)(a - \imath b) = a^2 + b^2.$$

(2) Mamy z poprzedniego $|z|^2 = z\bar{z}$, więc

$$|z_1 z_2|^2 = z_1 \bar{z}_1 z_2 \bar{z}_2 = |z_1|^2 |z_2|^2 : .$$

(3) Oczywiste.
(4) Mamy $1 = z\frac{1}{z}$. Z (2) wynika $\left|z\frac{1}{z}\right| = 1 = |z| \left|\frac{1}{z}\right|$. Zatem $\left|\frac{1}{z}\right| = \frac{1}{|z|}$. Stąd i z (2) wynika żądana równość.
(5) Niech $z_i = r_i(\cos\varphi_i + \imath \sin\varphi_i)$, $i = 1, 2$. Mamy

$$\begin{aligned} z_1 z_2 &= (r_1 \cos\varphi_1 + \imath r_1 \sin\varphi_1)(r_2 \cos\varphi_2 + \imath r_2 \sin\varphi_2) \\ &= r_1 r_2 (\cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2 + \imath \cos\varphi_1 \sin\varphi_2 + \imath \sin\varphi_1 \cos\varphi_2) \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + \imath \sin(\varphi_1 + \varphi_2)). \end{aligned}$$

Stąd żądana równość.
(6) W reprezentacji biegunowej

$$\bar{z} = r(\cos\varphi - \imath \sin\varphi) = r\cos(-\varphi) + r\imath \sin(-\varphi).$$

■

Wniosek (wzór de Moivre'a): Jeżeli $z = r(\cos\varphi + \imath \sin\varphi)$, to dla n całkowitego

$$z^n = r^n(\cos(n\varphi) + \imath \sin(n\varphi)).$$

Czytelnikowi pozostawiam znalezienie interpretacji geometrycznej następującego stwierdzenia:

STWIERDZENIE 1.6. *Dla dowolnych* $z_1, z_2 \in \mathbb{C}$ *zachodzą nierówności:*

(1) $|z_1 + z_2| \leqslant |z_1| + |z_2|$
(2) $||z_1| - |z_2|| \leqslant |z_1 - z_2|$

Dowód:

(1) Ze Stwierdzenia 1.5 wynika, że

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) \\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + \bar{z}_1 z_2 + z_1\bar{z}_1 = |z_1|^2 + z_1\bar{z}_2 + \overline{z_1\bar{z}_2} + |z_2|^2 \\ &= |z_1|^2 + 2\operatorname{Re}(z_1\bar{z}_2) + |z_2|^2. \end{aligned} \tag{1.3}$$

Z kolei, dla dowolnej liczby zespolonej z, mamy $|\operatorname{Re} z| \leqslant |z|$, więc

$$|\operatorname{Re}(z_1\bar{z}_2)| \leqslant |z_1\bar{z}_2| = |z_1||z_2|.$$

Z równości (1.3) dostajemy więc nierówność

$$|z_1 + z_2|^2 \leqslant |z_1|^2 + 2|z_1||z_2| + |z_2|^2 = (|z_1| + |z_2|)^2.$$

(2) Z poprzedniego punktu wynikają dwie nierówności

$$\begin{aligned} |z_1| &= |z_1 - z_2 + z_2| \leqslant |z_1 - z_2| + |z_2|, \\ |z_2| &= |z_2 - z_1 + z_1| \leqslant |z_1 - z_2| + |z_1|, \end{aligned}$$

a stąd

$$\begin{aligned} |z_1| - |z_2| &\leqslant |z_1 - z_2|, \\ |z_2| - |z_1| &\leqslant |z_1 - z_2| \end{aligned}$$

i teza.

■

1.3. Pierwiastkowanie liczb zespolonych.

DEFINICJA 1.7. *Niech* $w \in \mathbb{C}$. Pierwiastkiem stopnia n z w *nazywać będziemy taką liczbę zespoloną* z, *że* $z^n = w$.

Symbolem $\{\sqrt[n]{w}\}$ oznaczać będziemy zbiór pierwiastków stopnia n z w. Ustalony pierwiastek ozaczać będziemy (jeżeli nie prowadzi to do nieporozumień) $\sqrt[n]{w}$.

TWIERDZENIE 1.8. *Dla $w \in \mathbb{C}$, $w \neq 0$, oraz dla $n \in \mathbb{N}$, istnieje dokładnie n różnych pierwiastków stopnia n z w. Pierwiastki te dane są wzorami:*

$$z_k = \sqrt[n]{|w|}\left(\cos\frac{\varphi + 2k\pi}{n} + + \imath\sin\frac{\varphi + 2k\pi}{n}\right) \quad k = 0, 1, 2, .., n - 1,$$

gdzie $\sqrt[n]{|w|}$ jest pierwiastkiem arytmetycznym i $\varphi = \arg z$.

DOWÓD: Niech $w = r(\cos\varphi + \imath\sin\varphi)$ i niech $z = |z|(\cos\alpha + \imath\sin\alpha)$. Z wzorów de Moivre'a

$$|z|^n(\cos n\alpha + \imath\sin n\alpha) = r(\cos\varphi + \imath\sin\varphi),$$

czyli $|z|^n = r$ ($|z| = \sqrt[n]{r}$) i $n\alpha = \varphi + 2k\pi$ (k jest liczbą całkowitą). Ile różnych liczb dostaniemy, gdy k przebiega cały zbiór $\mathbb{Z}$?

Niech $\alpha = \dfrac{\varphi + 2k\pi}{n}$ i $\alpha' = \dfrac{\varphi + 2k'\pi}{n}$. Argumenty te są równoważne (tzn. dają to samo z), jeśli $\alpha - \alpha' = 2l\pi$, to znaczy, jeśli

$$\frac{\varphi + 2k\pi}{n} - \frac{\varphi + 2k'\pi}{n} = 2l\pi,$$

czyli $k - k' = nl$ dla pewnego całkowitego l. Zatem, aby otrzymać <u>wszystkie różne</u> pierwiastki wystarczy podstawić $k = 0, 1, 2, \dots, n - 1$. ∎

Przykłady:

(1) Znajdźmy $\{\sqrt{\imath}\}$ metodą algebraiczną. Równość $(x+\imath y)^2 = \imath$ jest równoważna równościom $x^2 - y^2 = 0, 2xy = 1$. Stąd $x = y$ i $2x^2 = 1$ lub $x = -y$ i $-2x^2 = 1$. Druga para równości nie daje rozwiązania, bo x jest liczbą rzeczywistą. Z pierwsze pary dostajemy $x = \pm\frac{1}{\sqrt{2}}$. Zatem

$$\{\sqrt{\imath}\} = \left\{\frac{1+\imath}{\sqrt{2}}, -\frac{1+\imath}{\sqrt{2}}\right\}.$$

(2) Znajdźmy $\{\sqrt[3]{1}\}$. Mamy $1 = \cos 0 + \imath\sin 0$, zatem $z_1 = 1$, $z_2 = \cos\frac{2\pi}{3} + \imath\sin\frac{2\pi}{3} = -\frac{1}{2} + \imath\frac{\sqrt{3}}{2}$, $z_3 = \cos\frac{4\pi}{3} + \imath\sin\frac{4\pi}{3} = -\frac{1}{2} - \imath\frac{\sqrt{3}}{2}$. Pierwiastki te tworzą na płaszczyźnie zespolonej wierzchołki trójkąta równobocznego, wpisanego w okrąg jednostkowy.

1.4. Pierwiastki z jedności.

Ponieważ $|1| = 1$, $\arg 1 = 0$, więc, na mocy Twierdzenia 1.8, pierwiastki n-tego stopnia z jedności dane są wzorami:

$$\varepsilon_k = \cos\frac{2k\pi}{n} + \imath\sin\frac{2k\pi}{n}, \quad k = 0, 1, 2, \dots, n - 1.$$

Ze wzorów de Moivre'a mamy $\varepsilon_k = \varepsilon_1^k$. Pierwiastki są więc równe:

$$1, \varepsilon_1, \varepsilon_1^2, \dots, \varepsilon_1^{n-1}.$$

Dowolny pierwiastek ε n-tego stopnia z 1 taki, że ciąg: $1, \varepsilon, \dots, \varepsilon^{n-1}$ daje wszystkie pierwiastki n-tego stopnia z jedności nazywa się *pierwiastkiem pierwotnym.* Jeżeli n jest liczbą pierwszą, to każdy, różny od 1, pierwiastek jest pierwotny.

Uwagi:

(1) Pierwiastki n-tego stopnia z 1 tworzą na płaszczyźnie wierzchołki n-kąta równobocznego wpisanego w okrąg o środku w zerze i promieniu jeden.
(2) Niech $0 \neq w \in \mathbb{C}$ i niech $z^n = w$. Jeżeli ε jest pierwiastkiem n-tego stopnia z 1, to εz jest pierwiastkiem n-tego stopnia z w. W ten sposób z jednego $z \in \{\sqrt[n]{w}\}$ można otrzymać wszystkie pierwiastki stopnia n z w.

1.5. Równania trzeciego stopnia.

1.5.1. Trochę historii.

Leonardo z Pizy (Pisano) zwany Fibonacci, autor „Liber abaci" (1202), zauważył, że jeśli: $x = \sqrt[3]{32} + \sqrt[3]{4}$ to $x^3 = 108$. Korzystał przy tym z tożsamości: $x^3 = u^3 + v^3 + 3u^2v + 3uv^2$, $x = u + v$. Tożsamość ta, jak zobaczymy, mogła go doprowadzić do rozwiązania równań trzeciego stopnia. Scipione del Ferro (1465-1526), profesor w Bolonii, umiał rozwiązywać równania postaci $x^3 + px = q$. Sposób rozwiązania otrzymał „w spadku" Antonio Maria del Fiore. Niccolo Tartaglia (1500-1557) wyzwany na pojedynek matematyczny przez del Fiore znalazł metodę rozwiązywania równań w wigilię zawodów, 12-tego lutego 1535.

Hieronimo Cardano (1501-1576) dostał od Tartagli wskazówki:

- Jeśli $x^3+px+q = 0$, to należy szukać takich α i β, że: $\alpha-\beta = q$ i $\alpha\beta = \left(\frac{p}{3}\right)^3$. Wówczas pierwiastkiem jest $x = \sqrt[3]{\beta} - \sqrt[3]{\alpha}$.
- Jeśli $x^3 = px + q$, to należy szukać takich α i β, że: $\alpha + \beta = q$ i $\alpha\beta = \left(\frac{p}{3}\right)^3$. Wówczas pierwiastkiem jest $x = \sqrt[3]{\alpha} + \sqrt[3]{\beta}$.
- Jeżeli $x^3 + q = px$, to zamiana $x \to -x$ sprowadza problem do już rozwiązanego.

Przykład: Spróbujmy rozwiązać według tych wskazówek równanie $x^3 = x$. Szukamy więc takich α, β, że $\alpha + \beta = 0$, $\alpha\beta = \frac{1}{27}$. Stąd $\alpha^2 = -\frac{1}{27}$. I co dalej?

Cardano w 1545r w „Artis magnæ sive de rebus algebraicis liber unus" dał pewne recepty na posługiwanie się urojonymi (bo nie istniejącymi) pierwiastkami z liczb ujemnych, np:

$$(5 + \sqrt{-15})(5 - \sqrt{-15}) = 25 - (-15) = 40.$$

1.5.2. Wzory Cardano.

Zauważmy, że każde równanie trzeciego stopnia da się sprowadzić do postaci $x^3 + px + q = 0$. Rozwiążmy je metodą Tartaglii: podstawmy $x = v + u$. Mamy

$$x^3 = v^3 + u^3 + 3vu(v + u),$$

czyli

$$x^3 - 3vux - (v^3 + u^3) = 0.$$

x ma być rozwiązaniem równania, więc żądamy, by $-3vu = p$ i $-(v^3+u^3) = q$. Stąd $-27v^3u^3 = p^3$ i spostrzeżenie, że v^3 i u^3 są parą rozwiązań równania kwadratowego

$$y^2 + qy - \left(\frac{p}{3}\right)^3 = 0.$$

Rozwiązujemy je:

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

Mamy ostatecznie

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

Jeśli $1, \varepsilon, \varepsilon^2$ są pierwiastkami trzeciego stopnia z jedności i $v+u$ jest rozwiązaniem równania, to: $\varepsilon v + \varepsilon^2 u$ oraz $\varepsilon^2 v + \varepsilon u$ są pozostałymi rozwiązaniami równania.

Przykład (dokończenie rozwiązania równania $x^3 = x$):

Mamy $v^3 + u^3 = 0$, $v^3u^3 = \dfrac{1}{27}$ i stąd $v^3 = \dfrac{\iota}{3\sqrt{3}}$. Zatem

$$v = \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right) \text{ i } u = \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} - \frac{i}{2}\right).$$

Liczby

$$\begin{aligned}
x_1 &= v + u = 1,\\
x_2 &= \varepsilon v + \varepsilon^2 u\\
&= \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} + \frac{\iota}{2}\right)\left(-\frac{1}{2} + \iota\frac{\sqrt{3}}{2}\right) + \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} - \frac{\iota}{2}\right)\left(-\frac{1}{2} - \iota\frac{\sqrt{3}}{2}\right)\\
&= -\frac{1}{4} - \frac{1}{4} - \frac{1}{4} - \frac{1}{4} = -1,\\
x_3 &= \varepsilon^2 v + \varepsilon u\\
&= \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} + \frac{\iota}{2}\right)\left(-\frac{1}{2} - \iota\frac{\sqrt{3}}{2}\right) + \frac{1}{\sqrt{3}}\left(\frac{\sqrt{3}}{2} - \frac{\iota}{2}\right)\left(-\frac{1}{2} + \iota\frac{\sqrt{3}}{2}\right)\\
&= -\frac{1}{4} + \frac{1}{4} - \frac{1}{4} + \frac{1}{4} = 0
\end{aligned}$$

są więc rozwiązaniami równania $x^3 = x$.

Rozdział 2. Trochę o wielomianach

2.1. Pojęcie wielomianu.

Wielomianem stopnia $\leqslant k$, gdzie $k \geqslant 0$, o współczynnikach zespolonych (lub rzeczywistych, wymiernych, całkowitych) nazywamy formalne wyrażenie

$$a_0 + a_1 x + \cdots + a_k x^k, \tag{2.1}$$

gdzie $a_0, a_1, \ldots, a_k \in \mathbb{C}$ (lub $\mathbb{R}, \mathbb{Q}, \mathbb{Z}$). Wielomiany można dodawać i mnożyć:

$$(a_0 + \cdots + a_k x^k) + (b_0 + \cdots + b_k x^k) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_k + b_k)x^k,$$

$$(a_0 + \cdots + a_k x^k)(b_0 + \cdots + b_l x^l) = c_0 + \cdots + c_{k+l} x^{k+l},$$

gdzie $c_i = \sum_{m+n=i} a_m b_n$. Przestrzeń wielomianów stopnia $\leqslant n$ o współczynnikach w $\mathbb{K}$ oznaczać będziemy $\mathbb{K}[n]$. *Stopniem wielomianu* w nazywamy najmniejszą liczbę naturalną (lub zero) n taką, że $w \in \mathbb{K}[n]$. Stopień wielomianu w oznaczamy $\deg w$.

Wielomianom odpowiadają *funkcje wielomianowe*. Otrzymujemy je zastępując w wyrażeniu (2.1) element x liczbami zespolonymi (rzeczywistymi, wymiernymi, całkowitymi), lub innymi obiektami (z jakiegoś zbioru A), które można mnożyć, dodawać i mnożyć przez liczbę (współczynnik wielomianu). Mówimy o funkcji wielomianowej zmiennej $a \in A$, o współczynnikach liczbowych (zespolonych, rzeczywistych, wymiernych, całkowitych). Jeżeli jako A weźmiemy $\mathbb{C}$ ($\mathbb{R}, \mathbb{Q}, \mathbb{Z}$), to, jak łatwo zauważyć, istnieje wzajemnie jednoznaczna odpowiedniość między wielomianami, a funkcjami wielomianowymi. Dlatego, jeżeli nie prowadzi to do nieporozumień, funkcję wielomianową będziemy nazywać po prostu wielomianem i z nim utożsamiać.

2.2. Podzielność wielomianów.

STWIERDZENIE 2.1. *Niech v, w będą wielomianami stopni k i l o współczynnikach w $\mathbb{K}$ ($\mathbb{K} = \mathbb{C}, \mathbb{R}, \mathbb{Q}$). Istnieją wielomiany f i r takie, że stopień wielomianu r jest mniejszy od l, oraz $v = fw + r$. Wielomiany te są wyznaczone jednoznacznie, gdy $w \neq 0$.*

DOWÓD: Dowód istnienia przeprowadzimy indukcyjnie ze względu na na k.

Niech $v = a_0 + \cdots + a_k x^k$, $a_k \neq 0$ i $b_0 + \cdots + b_l x^l$, $b_l \neq 0$. Jeżeli $k < l$, to kładziemy $f = 0, r = v$. Niech teraz $k \geqslant 0$. Wielomian

$$v - \left(\frac{a_k}{b_k} x^{k-l} \right) w$$

jest stopnia $< k$. Z założenia indukcyjnego istnieją wielomiany f_1, r takie, że

$$v - \left(\frac{a_k}{b_k} x^{k-l} \right) w = f_1 w + r$$

i, w konsekwencji,

$$v = \left(\frac{a_k}{b_k}x^{k-l} + f_1\right) w + r.$$

Jednoznaczność. Niech $v = fw + r = f_1 w + r_1$, czyli $(f - f_1)w = r_1 - r$. Ponieważ stopień wielomianu w jest większy od stopnia wielomianu $r_1 - r$, to musimy mieć $f - f_1 = 0$ i, w konsekwencji, $r_1 - r = 0$. ∎

Jeśli w równości $v = fw + r$ wielomian r jest wielomianem zerowym ($r = 0$), to mówimy, że w dzieli v (w jest *dzielnikiem* v). **Spostrzeżenia:**

(1) Jeżeli w dzieli v_1 i v_2, to w dzieli $v_1 + v_2$.
(2) Jeżeli w dzieli v i v dzieli f, to w dzieli f.
(3) Jeżeli w dzieli v i c jest liczbą różną od zera, to cw dzieli v.

Jeżeli w dzieli wielomiany v_1 i v_2, to w nazywamy *wspólnym dzielnikiem* v_1 i v_2. Najwyższym wspólnym dzielnikiem(NWD) wielomianów v_1 i v_2 nazywamy wspólny dzielnik najwyższego stopnia.

2.2.1. Algorytm Euklidesa. Niech n_1 będzie stopniem v_1 a n_2 stopniem v_2. Przyjmijmy, że $n_1 \geqslant n_2$ i że $v_1 \neq 0$. Ze Stwierdzenia 2.1 $v_1 = f_1 v_2 + v_3$, gdzie stopień v_3 jest mniejszy od n_2. Jeżeli w jest wspólnym dzielnikiem v_1, v_2, to w jest również dzielnikiem v_3 (bo $v_3 = v_1 - f_1 v_2$). Zatem w -wspólny dzielnik v_1, v_2, jest wspólnym dzielnikiem v_2, v_3. Jeżeli $v_3 = 0$, to kończymy procedurę.

Niech teraz $v_3 \neq 0$ i $v_2 = f_2 v_3 + v_4$; stąd w jest wspólnym dzielnikiem v_3 i v_4. Jeżeli $v_4 = 0$, to kończymy procedurę. Jeżeli $v_4 \neq 0$, to.... itd,
aż dojdziemy do $v_p = f_p v_{p+1}$, to znaczy $v_{p+2} = 0$. Oznacza to, że $v_{p+1} \neq 0$ jest dzielnikiem v_p i w jest dzielnikiem v_{p+1}. Stąd, ponieważ $v_{p-1} = f_{p-1} v_p + v_{p+1} = f_p f_{p-1} v_{p+1} + v_{p+1}$, v_{p+1} jest wspólnym dzielnikiem v_{p-1} i v_p. W końcu dostajemy, że v_{p+1} jest wspólnym dzielnikiem v_1 i v_2. Ponieważ każdy wspólny dzielnik v_1 i v_2 jest również dzielnikiem v_{p+1}, v_{p+1} jest NWD v_1 i v_2.

Z tej konstrukcji wynika, że NWD wielomianów v_1, v_2 jest wyznaczony jednoznacznie z dokładnością do współczynnika liczbowego (różnego od zera). Oznaczać go będziemy $\mathsf{NWD}(v_1, v_2)$

STWIERDZENIE 2.2. *Jeśli w jest* NWD *wielomianów v_1 i v_2, to istnieją wielomiany h_1 i h_2 takie, że $w = h_1 v_1 + h_2 v_2$.*

DOWÓD: Możemy przyjąć, że $w = v_{p+1}$, jak w algorytmie Euklidesa. Zatem

$$w = v_{p-1} - f_{p-1} v_p = -f_{p-1} v_{p-2} + (1 + f_{p-1} f_{p-2}) v_{p-1} = \dots .$$

∎

Wniosek 2.3. *Jeżeli wielomiany v, w są wzajemnie proste, tzn. $\mathsf{NWD}(v, w) = 1$, to dla dowolnego wielomianu f stopnia $< \deg v + \deg w$ istnieją wielomiany r, s takie, że*

$$f = rv + sw$$

i $\deg r < \deg w, \ \deg s < \deg v$.

DOWÓD: Ze Stwierdzenia wynika istnienie wielomianów r_1, s_1 takich że $1 = r_1 v + s_1 w$. Stąd $f = (fr_1)v + (fs_1)w$. Niech $fr_1 = pw + r$ będzie rozkładem, o którym mówi Stwierdzenie 2.1. Mamy

$$f = rv + sw,$$

gdzie $s = fs_1 + p$ i $\deg r < \deg w$. Ponieważ $\deg f, \deg(rv) < \deg v + \deg w$, to również $\deg(sw) < \deg v + \deg w$, a stąd $\deg s < \deg v$. ■

Podobnie jak dla dwóch wielomianów, definiujemy NWD dla rodziny $v_1, \ldots, v_k$ wielomianów i oznaczamy go $\mathsf{NWD}(v_1, \ldots, v_k)$. Pozostawiam jako ćwiczenie dowód dość oczywistej formuły indukcyjnej

$$\mathsf{NWD}(v_1, \ldots, v_k) = \mathsf{NWD}(v_k, \mathsf{NWD}(v_1, \ldots, v_{k-1})), \tag{2.2}$$

z której wynika istnienie $\mathsf{NWD}(v_1, \ldots, v_k)$.

Przez indukcję łatwo udowodnić odpowiednik twierdzenia 2.2 dla dowolnej liczby wielomanów:

Ze wzoru (2.2) i twierdzenia 2.2 wynika istnienie wielomianów h_k, f, że

$$\mathsf{NWD}(v_1, \ldots, v_k) = h_k v_k + f\mathsf{NWD}(v_1, \ldots, v_{k-1}).$$

Z założenia indukcyjnego

$$\mathsf{NWD}(v_1, \ldots, v_{k-1}) = g_1 v_1 + \cdots + g_{k-1} v_{k-1},$$

więc, kładąc $h_l = fg_l$ dla $l < k$, dostajemy

$$\mathsf{NWD}(v_1, \ldots, v_k) = h_1 v_1 + \cdots + h_k v_k. \tag{2.3}$$

Na zakończenie trzy proste, ale użyteczne stwierdzenia.

STWIERDZENIE 2.4. *Niech wielomiany $v_1, \ldots, v_k$ będą parami wzajemnie proste, tzn.* $\mathsf{NWD}(v_i, v_j) = 1$ *dla* $i \neq j$. *Wówczas* $\mathsf{NWD}(w_1, w_2, \ldots, w_k) = 1$, *gdzie* $w_i = v_1 \cdots v_{i-1} v_{i+i} \cdots v_k$.

DOWÓD: Z założeń wynika, że $\mathsf{NWD}(w_{k-1}, w_k) = v_1 \cdots v_{k-2}$. Z wzoru (2.2)

$$\mathsf{NWD}(w_{k-2}, w_{k-1}, w_k) = \mathsf{NWD}(w_{k-2}, \mathsf{NWD}(w_{k-1}, w_k)) = \\ \mathsf{NWD}(w_{k-1}, v_1 \cdots v_{k-2}) = v_1 \cdots v_{k-3}, \quad \text{i.t.d.}$$

■

Poniższe stwierdzenie jest uogólnieniem wniosku 2.3.

STWIERDZENIE 2.5. *Niech wielomiany* $v_1, \dots, v_k$ *będą parami wzajemnie proste i niech* f *będzie wielomianem stopnia* $< \deg v_1 + \cdots + \deg v_k$. *Wówczas istnieją wielomiany* $r_1, \dots, r_k$ *takie, że* $\deg r_i < \deg v_i$ *oraz*

$$f = r_1 w_1 + \cdots + r_k w_k,$$

gdzie w_i *są jak w poprzednim stwierdzeniu.*

DOWÓD: Z poprzedniego stwierdzenia

$$\mathsf{NWD}(w_1, w_2, \dots, w_k) = 1.$$

Ze wzoru (2.3) wynika istnienie wielomianów $f_1, \dots, f_k$ takich, że

$$f = f_1 w_1 + f_2 w_2 + \cdots + f_k w_k.$$

Niech, z kolei, $f_i = g_i v_i + r_i$, gdzie $\deg r_i < \deg v_i$. Dostajemy

$$f = r_1 w_1 + \cdots + r_k w_k + r v_1 \cdots v_k,$$

a ponieważ $\deg f, \deg(r_i w_i) < \deg v_1 + \cdots + \deg v_k$, to $r = 0$. ∎

STWIERDZENIE 2.6. *Niech* v, f *będą wielomianami i niech* $\deg f < (\deg v)^l$. *Wówczas istnieją wielomiany* s_i, $\deg s_i < \deg v$, *że*

$$f = s_l + s_{l-1} v + \cdots + s_1 v^{l-1}.$$

DOWÓD: Mamy $f = s_1 v^{l-1} + f_1$, gdzie $\deg f_1 < \deg v^{l-1}$. Wynika stąd, że $\deg s_1 < \deg v$. Postępując podobnie dostajemy $f_1 = s_2 v^{l-2} + f_2$, gdzie $\deg f_2 < \deg v^{l-2}$ i $\deg s_2 < \deg v$, i.t.d. Stąd dostajemy (indukcyjnie) żądany rozkład. ∎

2.3. Podstawowe twierdzenie algebry.

Niech w będzie wielomianem stopnia $\geqslant 1$ i niech $a \in \mathbb{C}$. Mamy ze Stwierdzenia 2.1 $w(z) = w_1(z)(z-a) + r$, gdzie r jest wielomianem stopnia zerowego, więc liczbą, $r \in \mathbb{C}$. Stąd,

$$(z-a) \text{ dzieli } w \Leftrightarrow r = 0 \Leftrightarrow w(a) = 0.$$

Mówimy, że a jest pierwiastkiem wielomianu w.

TWIERDZENIE 2.7. *Wielomian* w *stopnia* n *ma dokładnie* n *pierwiastków liczonych z uwzględnieniem krotności.*

Uwaga. Gdy a jest pierwiastkiem w, tzn., gdy $z - a$ dzieli $w(z)$, to *krotnością* pierwiastka a nazywamy największą liczbę $k \in \mathbb{N}$ taką, że $(z-a)^k$ dzieli w. Aby udowodnić powyższe twierdzenie wystarczy wykazać istnienie jednego pierwiastka, dzieląc bowiem przez $(z-a)$ dostajemy wielomian stopnia $n-1$ i, przez indukcję, dostajemy tezę. Dowód istnienia pierwiastka jest trudny i wykracza poza ramy naszego wykładu.

2.4. Rozkład funkcji wymiernej na ułamki proste.

Jako przykład zastosowania reprezentacji (2.3) podamy dowód rozkładu rzeczywistej funkcji wymiernej, intensywnie wykorzystywanego w praktyce całkowania. *Funkcją wymierną* nazywamy funkcję będącą ilorazem funkcji wielomianowych. Jak już było powiedziane, zajmujemy się tylko przypadkiem $\mathbb{K} = \mathbb{R}$. Niech P, Q będą funkcjami wielomianowymi stopni, odpowiednio, n, m i takimi, że $n < m$ oraz $\mathsf{NWD}(P, Q) = 1$. Z Twierdzenia Podstawowego Algebry wynika następujący rozkład wielomianu Q na czynniki:

$$Q(x) = c(x-\alpha_1)^{k_1}\cdots(x-\alpha_l)^{k_l}(x^2+\beta_1 x+\gamma_1)^{m_1}\cdots(x^2+\beta_r x+\gamma_r)^{m_r}, \quad (2.4)$$

gdzie $c, \alpha_i, \beta_i, \gamma_i$ sa liczbami rzeczywistymi, a czynniki rozkładu są parami wzajemnie proste. Ze Stwierdzenia 2.5 wynika, że

$$\frac{Q(x)}{P(x)} = \frac{Q_1(x)}{(x-\alpha_1)^{k_1}}+\cdots+\frac{Q_l(x)}{(x-\alpha_l)^{k_l}}+\frac{\bar{Q}_1(x)}{(x^2+\beta_1 x+\gamma_1)^{m_1}}+\cdots+\frac{\bar{Q}_r(x)}{(x^2+\beta_r x+\gamma_r)^{m_r}},$$

gdzie $\deg Q_i < k_i$ i $\deg \bar{Q}_i < 2m_i$. Z kolei, ze Stwierdzenia 2.6 wynika, że

$$\frac{Q_i(x)}{(x-\alpha_1)^{k_i}} = \frac{a_{i1}}{(x-\alpha_i)}+\frac{a_{i2}}{(x-\alpha_i)^2}+\cdots+\frac{a_{ik_i}}{(x-\alpha_i)^{k_i}}$$

oraz

$$\frac{\bar{Q}_i(x)}{(x^2+\beta_i x+\gamma_i)^{m_i}} = \frac{b_{i1}x+c_{i1}}{(x^2+\beta_i x+\gamma_i)}+\cdots+\frac{b_{im_i}x+c_{im_i}}{(x^2+\beta_i x+\gamma_i)^{m_i}},$$

gdzie a_{ij}, b_{ij}, c_{ij} są liczbami rzeczywistymi.

Ostatecznie,

$$\begin{aligned}\frac{P(x)}{Q(x)} = {} & \frac{a_{11}}{(x-\alpha_1)}+\frac{a_{12}}{(x-\alpha_1)^2}+\cdots+\frac{a_{1k_1}}{(x-\alpha_1)^{k_1}}+\cdots \\ & +\frac{a_{l1}}{(x-\alpha_l)}+\frac{a_{l2}}{(x-\alpha_l)^2}+\cdots+\frac{a_{lk_l}}{(x-\alpha_l)^{k_l}}+ \\ & +\frac{b_{11}x+c_{11}}{(x^2+\beta_1 x+\gamma_1)}+\cdots+\frac{b_{1m_1}x+c_{1m_1}}{(x^2+\beta_1 x+\gamma_1)^{m_1}}+\cdots \\ & +\frac{b_{r1}x+c_{r1}}{(x^2+\beta_r x+\gamma_r)}+\cdots+\frac{b_{rm_r}x+c_{rm_r}}{(x^2+\beta_r x+\gamma_r)^{m_r}}. \quad (2.5)\end{aligned}$$

Rozdział 3. Grupy. Ciała

3.1. Grupy.

DEFINICJA 3.1. *Grupą nazywamy niepusty zbiór G z działaniem dwuargumentowym, to jest odwzorowaniem*

$$*\colon G \times G \to G\colon (a,b) \mapsto *(a,b)$$

(dla wygody będziemy oznaczać $*(a,b)$ przez $a*b$) takim, że:

(1) *$\forall a,b,c \in G \quad *(a,*(b,c)) = *(*(a,b),c)$, czyli $(a*(b*c) = (a*b)*c$ - łączność,*
(2) *istnieje $e \in G$ takie, że $\forall a \in G$ mamy $e*a = a = a*e$ - istnienie jedności (elementu neutralnego) grupy,*
(3) *$\forall a \in G\ \exists x \in G$ taki, że $a*x = x*a = e$ - istnienie elementu odwrotnego.*

Jeżeli ponadto

(4) $\forall a,b \in G \quad a*b = b*a$,

to mówimy, że grupa jest *przemienna* (*abelowa*).

W dalszym ciągu będziemy pisać $a*b*c$ zamiast $a*(b*c) = (a*b)*c$.
Przykłady:

(1) $(\mathbb{Z},+), (\mathbb{C},+), (\mathbb{R},+), (\mathbb{R}_+,\cdot)$, gdzie $\mathbb{R}_+ = \{r \in \mathbb{R}\colon r > 0\}$, są grupami przemiennymi.
(2) Okrąg jednostkowy S^1 na płaszczyźnie zespolonej z działaniem mnożenia.
(3) Pierwiastki z jedynki wszystkich możliwych stopni, też z działaniem mnożenia.
(4) Pierwiastki ustalonego stopnia z jedynki.
(5) Bijekcje ustalonego zbioru A z działaniem złożenia (superpozycji) tworzą grupę nieprzemienną. Jednością jest tu odwzorowanie tożsamościowe. Elementem odwrotnym jest odwzorowanie odwrotne.

Uwaga. Przedstawiony zespół aksjomatów grupy nie jest minimalny. Warunki (2) i (3) można zastąpić słabszymi warunkami o istnieniu lewej jedności ($e*a = a$) i lewego elementu odwrotnego ($x*a = e$). Dla przykładu pokażemy, że z tych warunków wynika $a*x = e$ (lewy element odwrotny jest też prawym elementem odwrotnym). Jeżeli bowiem $x*a = e$, to $x*a*x = (x*a)*x = e*x = x$, a stąd, dla b takiego, że $b*x = e$, mamy

$$a*x = (b*x)*(a*x) = b*x*a*x = b*(x*a*x) = b*x = e.$$

Jako ćwiczenie pozostawiam dowód, że lewa jedność jest też prawą jednością.

STWIERDZENIE 3.2. *Grupa* $(G, *)$ *ma dokładnie jedną jedność i dla każdego* $a \in G$ *istnieje dokładnie jeden element odwrotny (oznaczany* a^{-1}*).*

Dowód: Niech e i e' będą jednościami grupy $(G, *)$. Wówczas $e = *(e, e') = e'$. Niech teraz x i x' będą elementami odwrotnymi do $a \in G$. Mamy $x * a = e = x' * a$. Stąd

$$x = x * e = x * (x' * a) = ((x * a) * x') = x'.$$

■

Wnioski:

(1) $(a^{-1})^{-1} = a$,
(2) $(a * b)^{-1} = b^{-1} * a^{-1}$,
(3) jeśli $a * x = b$ to $x = a^{-1} * b$,
(4) jeśli $x * a = b$ to $x = b * a^{-1}$,
(5) jeśli dla pewnego a mamy $a * x = a * y$ to $x = y$ i, podobnie, jeśli $x * a = y * a$, to $x = y$.

Podzbiór $H \subset G$ jest *podgrupą* jeżeli $(H, *|_H)$ jest grupą. Aby stwierdzić, że H jest podgrupą wystarczy sprawdzić, czy dla $a, b \in H$ również $a * b, a^{-1} \in H$ lub, równoważnie, czy $a * b^{-1} \in H$.

3.2. Grupa permutacji.

Niech będzie $I_n = \{1, 2, ..., n\}$. Przez $\mathsf{S}(n)$ oznaczać będziemy grupę bijekcji zbioru I_n. Elementy tej grupy nazywać będziemy permutacjami elementów zbioru I_n.

Bijekcję $\mathsf{S}(n) \ni s$ można zapisać w postaci:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ s(1) & s(2) & \cdots & s(n) \end{pmatrix}$$

lub, ogólnie,

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ s(i_1) & s(i_2) & \cdots & s(i_n) \end{pmatrix}$$

Na przykład

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

i

$$\begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

oznaczają tą samą permutację.

Moc (liczba elementów) grupy permutacji $\mathsf{S}(n)$ wynosi, jak łatwo policzyć, $n!$

DEFINICJA 3.3. *Permutację* $s \in \mathsf{S}(n)$ *nazywamy* cyklem o długości $k > 1$, *jeśli istnieje* k *różnych liczb* $i_1, \ldots, i_k$ *ze zbioru* $\{1, \ldots, n\}$ *takich, że*

$$s(i_1) = i_2, s(i_2) = i_3, \cdots, s(i_{k-1}) = i_k, s(i_k) = i_1$$

oraz $s(j) = j$ *dla pozostałych liczb* $j \in \{1, \ldots, n\}$

Przykład: Permutacja

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 1 & 4 & 3 \end{pmatrix}$$

jest cyklem ($i_1 = 1, i_2 = 5, i_3 = 3$).

Cykl taki jak w definicji oznaczać będziemy $(i_1, ..., i_k)$. Permutację z przykładu oznaczamy więc (1,5,3) lub (5,3,1) lub (3,1,5). Uwaga : z zapisu nie wynika, do jakiej grupy permutacji należy cykl, tzn., czemu jest równe n.

DEFINICJA 3.4. *Cykle* $s = (i_1, \ldots, i_k)$ *i* $t = (j_1, \ldots, j_l)$ *nazywamy* rozłącznymi *jeśli*

$$\{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_l\} = \emptyset.$$

TWIERDZENIE 3.5.

(1) *Jeżeli* s, t *są cyklami rozłącznymi, to* $s \circ t = t \circ s$.
(2) *Każda permutacja, różna od tożsamości, jest superpozycją cykli rozłącznych.*

Dowód:

(1) Oczywiste.
(2) Niech $s \in \mathsf{S}(n)$. Weźmy pierwszą liczbę i_1 w ciągu $(1, \ldots, n)$ taką, że $s(i_1) \neq i_1$. Definiujemy $i_2 = s(i_1)$ i obserwujemy, że $s(i_2) \neq i_2$ (s jest bijekcją).

   Jeżeli $s(i_2) = i_1$ to mamy cykl, a jeżeli nie, to kładziemy $s(i_2) = i_3$. Ponieważ s jest bijekcją mamy $s(i_3) \neq i_3, i_2$. Jeżeli $i_3 = i_1$ to dostajemy cykl a jeżeli nie, to kładziemy $i_4 = s(i_3)$ i.t.d. Ponieważ n jest liczbą skończoną, dostajemy w końcu $s(i_k) = i_1$, a więc cykl.

   Niech teraz l_1 będzie pierwszą liczbą różną od $i_1, \ldots, i_k$, i taką, że $s(l_1) \neq l_1$. Kładziemy $l_2 = s(l_2)$ i.t.d. Otrzymujemy znowu cykl $(l_1, \ldots, l_p)$.

   Po skończonej liczbie kroków dostajemy ciąg cykli

$$(i_1, \ldots, i_k), (l_1, \ldots, l_p), \ldots, (m_1, \ldots, m_q).$$

   Z konstrukcji tych cykli widać natychmiast, że:

$$s = (i_1, \ldots, i_k)(l_1, \ldots, l_p) \cdots (m_1, \ldots, m_q).$$

■

Przykład:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 3 & 4 & 5 & 1 & 2 \end{pmatrix} = (1,8,2,7)(3,6,5,4)$$

DEFINICJA 3.6. *Cykl o długości 2 nazywamy* transpozycją.

STWIERDZENIE 3.7. *Cykl o długości k można przedstawić jako złożenie $k-1$ transpozycji.*

Dowód: Łatwo sprawdzić następującą równość

$$(i_1, \dots, i_k) = (i_1, i_k)(i_1, i_{k-1}) \cdots (i_1, i_2).$$

■

Wniosek: każda permutacja $s \in \mathsf{S}(n)$ da się przedstawić jako złożenie co najwyżej $n-1$ transpozycji.

TWIERDZENIE 3.8. *Niech $s = t_k \cdots t_1$ gdzie t_i są transpozycjami. Parzystość liczby k zależy tylko od s.*

Dowód: Dla $q \in \mathsf{S}(n)$ oznaczmy przez $m(q)$ liczbę czynników występujących w rozkładzie q na cykle rozłączne plus liczbę elementów stałych w q (tzn. liczb j takich, że $q(j) = j$).

Pokażemy, że jeśli t jest transpozycją, to $m(tq) = m(q) \pm 1$. Niech $t = (i,j)$, wówczas mamy następujące możliwości:

(1) Jeżeli i, j nie wystąpują w żadnym cyklu, to $m(tq) = m(q) - 1$ (dwa punkty stałe mniej, jedna transpozycja więcej).

(2) Jeżeli i występuje w cyklu $(i, i_2, \dots, i_r)$, a j jest punktem stałym, to w rozkładzie permutacji tq na cykle rozłączne wystąpi

$$(i,j)(i, i_2, \dots, i_r) = (i, i_2, \dots, i_r, j),$$

czyli $m(tq) = m(q) - 1$ (jeden punkt stały mniej).

(3) Jeżeli i, j występują w różnych cyklach $(i, i_2, \dots, i_r)$ i $(j, j_2, \dots, j_l)$, to

$$(i,j)(i, i_2, \dots, i_r)(j, j_2, \dots, j_l) = (j, j_2, \dots, j_l, i, i_2, \dots, i_r).$$

Zatem $m(tq) = m(q) - 1$ (jeden cykl mniej).

(4) Jeżeli i, j występują w jednym cyklu $(i, i_2, \dots, i_{p-1}, j, i_{p+1}, \dots, i_r)$, to

$$(i,j)(i, i_2, \dots, i_{p-1}, j, i_{p+1}, \dots, i_r) = (i, i_2, \dots, i_{p-1})(j, i_{p+1}, \dots, i_r),$$

czyli $m(tq) = m(q) + 1$ (jeden cykl więcej).

Przypuśćmy teraz, że $s = t_{2r} \cdots t_1 = s_{2p+1} \cdots s_1$, gdzie t_i, s_j są transpozycjami. Mamy $m(t_1) = n - 1 = m(s_1)$, więc

$$m(s) = (n-1) \underbrace{\pm 1 \pm 1, \dots, \pm 1}_{2r-1 \ razy}$$

i, z drugiej strony,

$$m(s) = (n-1) \underbrace{\pm 1 \pm 1, \dots, \pm 1}_{2p \ razy}$$

- sprzeczność! (parzystość liczby $m(s)$ się nie zgadza). ■

Wniosek: dla permutacji s liczba $(-1)^{liczba\ transpozycji\ w\ rozkladzie\ s}$ jest dobrze określona. Nazywamy ją *znakiem permutacji* s i oznaczamy $\operatorname{sgn}(s)$.

STWIERDZENIE 3.9. $\operatorname{sgn}(s_1 \circ s_2) = \operatorname{sgn}(s_1) \operatorname{sgn}(s_2)$.

STWIERDZENIE 3.10. *Jeśli p jest liczbą cykli parzystej długości występujących w rozkładzie permutacji $s \in \mathsf{S}(n)$, to* $\operatorname{sgn} s = (-1)^p$.

Dowód: Wynika natychmiast ze Stwierdzenia 3.7. ■

Permutację s nazywamy parzystą, jeżeli $\operatorname{sgn}(s) = 1$ i nieparzystą, jeżeli $\operatorname{sgn}(s) = -1$.

Alternatywny sposób wprowadzania znaku permutacji.

DEFINICJA 3.11. *Dla $s \in \mathsf{S}(n)$* s-inwersją *nazywamy parę (i, j) taką, że $1 \leqslant i < j \leqslant n$ oraz $s(i) > s(j)$.*

STWIERDZENIE 3.12. *Jeśli $s, t \in \mathsf{S}(n)$, przy czym t jest transpozycją, to suma*
(liczba s-inwersji) + (liczba $s \circ t$-inwersji)
jest liczbą nieparzystą).

Dowód: Niech $t = (i, j)$, gdzie $i < j$; wtedy

$$s \circ t = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ s(1) & \dots & s(j) & \dots & s(i) & \dots & s(n) \end{pmatrix}.$$

Stąd

(a) dla $i < k < j$:

(i, j) jest s-inwersją albo $s \circ t$-inwersją,
(i, k) jest s-inwersją albo $s \circ t$-inwersją,
(k, j) jest s-inwersją albo $s \circ t$-inwersją.

Tych par jest nieparzysta liczba!

(b) dla wszystkich pozostałych par (k, l), $1 \leq k < l \leq n$, mamy:
(k, l) jest s-inwersją $\Leftrightarrow$ (k, l) jest $s \circ t$-inwersją.

Stąd teza. ■

Wnioski:

(1) Jeżeli zdefiniujemy $\operatorname{sgn} s = (-1)^{liczba\ s-inwersji}$, to $\operatorname{sgn}(s \circ t) = -\operatorname{sgn} s$ gdy t jest transpozycją.
(2) Jeżeli s da się przedstawić jako złożenie k transpozycji, to $\operatorname{sgn} s = (-1)^k$.

3.3. Ciała.

Ciałem K nazywamy zbiór K z działaniami dwuargumentowymi „+" i „·" takimi, że

(1) $(K, +)$ jest grupą abelową,
(2) $(K \setminus 0, \cdot)$ jest grupą abelową, gdzie 0 jest elementem neutralnym ze względu na „+",
(3) $\forall a, b, c, \in K$ mamy $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (rozdzielność mnożenia względem dodawania).

Przykłady:

(1) Liczby zespolone z działaniami dodawania i mnożenia.
(2) Liczby rzeczywiste z działaniami dodawania i mnożenia.
(3) Liczby wymierne z działaniami dodawania i mnożenia.
(4) Zbiór dwuelementowy $\{0, 1\}$ z działaniami dodawania modulo 2 i mnożenia.
(5) Zbiór $Z_n = \{1, 2, \ldots, n\}$ z działaniami dodawania i mnożenia modulo n, gdzie n jest liczbą pierwszą. (Udowodnić!)

Rozdział 4. Przestrzenie wektorowe

4.1. Definicja i przykłady.

DEFINICJA 4.1. Przestrzenią wektorową *nad ciałem* $\mathbb{K}$ *nazywamy grupę abelową* $(V,+)$ *z odwzorowaniem*

$$\mathbb{K} \times V \to V : (\lambda, v) \longmapsto \lambda \cdot v$$

takim, że dla wszystkich $\lambda, \mu \in \mathbb{K}$, $v, w \in V$*:*

(1) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,
(2) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$,
(3) $1 \cdot v = v$,
(4) $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$.

Elementy przestrzeni wektorowej nazywać będziemy *wektorami*, a elementy ciała $\mathbb{K}$ *skalarami*. W dalszym ciągu zajmować się będziemy jedynie przestrzeniami wektorowym nad ciałami $\mathbb{K} = \mathbb{R}$ i $\mathbb{K} = \mathbb{C}$. Element nautralny grupy $(V,+)$ (wektor zerowy) oznaczać będziemy, dla odróżnienia od zera liczbowego, przez $\mathbf{0}$. Będziemy też pisać po prostu λv zamiast $\lambda \cdot v$.

STWIERDZENIE 4.2. *Dla każdego wektora* $v \in V$ *i każdej liczby* $\lambda \in \mathbb{K}$

(1) $0v = \mathbf{0}$,
(2) $(-1)v = -v$, *to znaczy*, $v + (-1)v = \mathbf{0}$,
(3) $\lambda\mathbf{0} = \mathbf{0}$,
(4) *jeżeli* $\lambda v = \mathbf{0}$ *to* $\lambda = 0$ *lub* $v = \mathbf{0}$.

DOWÓD: Niech $v \in V$ i $\lambda \in \mathbb{K}$.

(1) Mamy $v = (1 + 0)v = 1v + 0v = v + 0v$ i stąd $\mathbf{0} = 0v$.
(2) Z powyższego i z punktu pierwszego definicji $v + (-1)v = (1 + (-1))v = 0v = \mathbf{0}$, czyli $-v = (-1) \cdot v$
(3) Z punktu trzeciego definicji $\lambda v = \lambda(v + \mathbf{0}) = \lambda v + \lambda\mathbf{0}$ i stąd $\lambda\mathbf{0} = \mathbf{0}$.
(4) Jeżeli $\lambda v = \mathbf{0}$ i $\lambda \neq 0$, to $v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \mathbf{0}$.

■

4.1.1. Przykłady.

(A) Zbiór $\mathbb{K}^n$ z dodawaniem

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$$

i mnożeniem

$$\lambda(x_1, x_2, \cdots, x_n) = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n)$$

jest przestrzenią wektorową nad $\mathbb{K}$.

(B) Niech A będzie dowolnym zbiorem. Symbolem $\mathrm{Map}(A, B)$ oznaczamy zbiór wszystkich odwzorowań ze zbioru A w zbiór B. Weźmy $V = \mathrm{Map}(A, K)$ z działaniami:

$$(f + g)(a) = f(a) + g(a)$$

oraz

$$(\lambda f)(a) = \lambda f(a).$$

Jest to przestrzeń wektorowa nad $\mathbb{K}$. W szczególnosci, biorąc $A = I_n = \{1, 2, \dots, n\}$, dostaniemy poprzedni przykład.

(C) Zbiór wielomianów $\mathbb{K}[n]$ stopnia $\leqslant n$ z działaniami określonymi w rozdziale drugim jest przestrzenią wektorową nad $\mathbb{K}$.

DEFINICJA 4.3. *Niepusty podzbiór S przestrzeni wektorowej V nazywamy* podprzestrzenią wektorową *przestrzeni V, jeżeli S z działaniami indukowanymi z V jest przestrzenią wektorową.*

STWIERDZENIE 4.4. *S jest podprzestrzenią wektorową wtedy i tylko wtedy, gdy*

$$\forall \lambda_1, \lambda_2 \in \mathbb{K}, \forall v_1, v_2, \in S$$

mamy

$$\lambda_1 v_1 + \lambda_2 v_2 \in S$$

DOWÓD: Jedyną rzeczą do sprawdzenia jest (oczywista) wykonalność działań dodawania wektorów i mnożenia ich przez liczbę. Pozostałe własności działań spełnione są automatycznie. ■

Dalszy ciąg przykładów.

(D) Funkcje wielomianowe na $\mathbb{R}$ o współczynnikach w $\mathbb{K}$ tworzą podprzestrzeń wektorową przestrzeni wszystkich funkcji na $\mathbb{R}$ o wartościach w $\mathbb{K}$.

(E) Niech A będzie dowolnym zbiorem i W - dowolną przestrzenią wektorową nad $\mathbb{K}$. $V = \mathrm{Map}(A, W)$ z działaniami określonymi wzorami jak w przykładzie (B).

(F) Inne podprzestrzenie przestrzeni $\mathrm{Map}(\mathbb{R}, \mathbb{K})$: wielomianów parzystych, funkcji ciągłych, funkcji różniczkowalnych, etc.

(G) Bardzo ważny przykład-wytrych: zbiór rozwiązań układu

$$(*)\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_2 & = & 0 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0 \end{cases}$$

tworzy podprzestrzeń wektorową przestrzeni $\mathbb{K}^n$, co jest oczywiste. Pokażemy, że gdy liczba niewiadomych jest większa od liczby równań, to podprzestrzeń ta zawiera wektory różne od zera.

TWIERDZENIE 4.5. *Jeżeli $n > m$ to istnieje niezerowe rozwiązanie układu $(*)$.*

DOWÓD: (indukcyjny ze względu na liczbę równań): Dla $m = 1$ i $n > 1$ twierdzenie jest w oczywisty sposób prawdziwe. Załóżmy, że jest również prawdziwe dla $m - 1$.

Jeżeli wszystkie współczynniki w $(*)$ są równe zeru, to zbiór rozwiązań jest równy całemu $\mathbb{K}^n$. Przyjmijmy teraz, że choć jeden ze współczynników w $(*)$ jest różny od zera, na przykład a_{11}. Układ $(*)$ zastępujemy równoważnym (to znaczy mającym te same rozwiązania)

$$(**)\left\{\begin{array}{lcccccc} a_{11}x_1 + a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ (a_{22} - \frac{a_{21}}{a_{11}}a_{12})x_2 & + & \cdots & + & (a_{2n} - \frac{a_{21}}{a_{11}}a_{1n})x_2 & = & 0 \\ & & \vdots & & & & \\ (a_{m2} - \frac{a_{m1}}{a_{11}}a_{12})x_2 & + & \cdots & + & (a_{mn} - \frac{a_{m1}}{a_{11}}a_{1n})x_n & = & 0 \end{array}\right.$$

Układ powstały z $(**)$ przez usunięcie pierwszego równania ma $m - 1$ równań i $n - 1$ niewiadomych. Z założenia indukcyjnego ma on niezerowe rozwiązanie. Rozwiązanie niezerowe układu $(**)$ dostaniemy wyznaczając x_1 z pierwszego równania. ■

DEFINICJA 4.6. *Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$ i niech będzie dany ciąg wektorów $v_1, v_2, \dots, v_n \in V$. Wektor przestrzeni V postaci*

$$\lambda^1 v_1 + \lambda^2 v_2 + \cdots + \lambda^n v_n,$$

gdzie $\lambda^i \in \mathbb{K}$, nazywamy kombinacją liniową *wektorów $v_1, \dots, v_2$.*

Niech teraz S będzie dowolnym podzbiorem przestrzeni V. Zbiór kombinacji liniowych wektorów z S oznaczać będziemy $\langle S\rangle$. Jeżeli $S = \emptyset$, to przyjmujemy $\langle S\rangle = \{\mathbf{0}\}$. Inaczej mówiąc, kombinacja liniowa pustego ciągu wektorów jest wektorem zerowym.

STWIERDZENIE 4.7. *$\langle S\rangle$ jest podprzestrzenią wektorową przestrzeni V.*

DOWÓD: Wystarczy rozpatrzyć przypadek niepustego S. Niech $v, w \in \langle S\rangle$, tzn. $v = \lambda^1 v_1 + ... + \lambda^n v_n$ i $w = \mu^1 w_1 + .. + \mu^n w_n$ gdzie v_i, $w_i \in S$ i $\lambda^i, \mu^i \in \mathbb{K}$. Dla dowolnych $\lambda, \mu \in \mathbb{K}$ mamy

$$\lambda v + \mu w = (\lambda\lambda^1)v_1 + \cdots + (\lambda\lambda^n)v_n + (\mu\mu_1)w_1 + \cdots + (\mu\mu^m)w_m \in S$$

■

Uwagi:
a) Jeżeli $V \supset W \supset S$ i W jest podprzestrzenią wektorową to $\langle S\rangle \subset W$.
b) $\langle S\rangle$ jest najmniejszą podprzestrzenią wektorową zawierającą S.

Przykład: $S = \{1, x, x + x^2, x\}$. $\langle S\rangle = \mathbb{K}[2]$.
Inne przykłady będą podane później.

4.2. Liniowa niezależność. Baza.

DEFINICJA 4.8. *Przestrzeń wektorową V nazywamy* skończenie wymiarową, *jeżeli istnieje skończony zbiór wektorów $S = \{v_1, v_2, \dots, v_k\} \subset V$ taki, że $\langle S \rangle = V$.*

Przykłady:

(1) $V = \mathbb{K}^n$ i $S = \{e_1, \dots, e_n\}$ gdzie $e_i = (\delta_{1i}, \dots, \delta_{ni})$.
(2) Przestrzeń wielomianów stopnia $\leqslant 2$ i $S = \{1, x, x^2\}$
(3) Przestrzeń odwzorowań $\mathrm{Map}(\mathbb{N}, \mathbb{K})$ nie jest skończenie wymiarowa (jest nieskończenie wymiarowa).
(4) Przestrzeń $\mathbb{R}$ jako przestrzeń wektorowa nad ciałem liczb wymiernych nie jest skończenie wymiarowa.

DEFINICJA 4.9. *Układ wektorów (ciąg wektorów - jeśli uporządkowany)*

$$\{v_1, v_2, \dots, v_k\}, v_i \in V,$$

nazywamy linowo niezależnym, *jeżeli zachodzi implikacja*

$$\left(\lambda^1 v_1 + \dots + \lambda^k v_k = \mathbf{0}\right) \Longrightarrow \left(\lambda^1 = \lambda^2 = \dots = \lambda^k = 0\right).$$

Jeżeli układ wektorów nie jest liniowo niezależny, to mówimy, że jest liniowo zależny.

Przykłady:

(1) Wielomiany $\{1, t, t^3\}$ sa liniowo niezależne.
(2) $V = \mathbb{C}^1$ jako przestrzeń wektorowa nad ciałem liczb rzeczywistych. 1, i są liniowo niezależne.
(3) $V = \mathbb{C}^1$ jako przestrzeń wektorowa nad ciałem liczb zespolonych. 1, i nie są liniowo niezależne.
(4) Dowolny układ zawierający wektor zerowy jest liniowo zależny.
(5) Jeżeli $v \neq \mathbf{0}$ to układ $\{v\}$ składający się z jednego wektora jest liniowo niezależny.

DEFINICJA 4.10. *Mówimy, że wektor v jest* liniowo zależny *od układu wektorów $v_1, v_2, \dots, v_k$, jeżeli istnieją liczby $\lambda^1, \dots, \lambda^k$ takie, że*

$$v = \lambda^1 v_1 + \dots + \lambda^k v_k$$

lub, równoważnie,

$$v \in \langle\{v_1, v_2, \dots, v_k\}\rangle,$$

lub, równoważnie,

$$\langle\{v_1, v_2, \dots, v_k\}\rangle = \langle\{v_1, v_2, \dots, v_k, v\}\rangle.$$

Poniższe stwierdzenie nie wymaga dowodu.

STWIERDZENIE 4.11. *Niech $S = \{v_1, \ldots, v_k\}$ będzie skończonym układem wektorów z przestrzeni wektorowej V. Wówczas*

(1) *Jeśli $S_0 \subset S$ i S_0 jest liniowo zależny, to S też jest liniowo zależny.*
(2) *Jeśli $S_0 \subset S$ i S jest liniowo niezależny, to S_0 też jest liniowo niezależny.*
(3) *Jeśli $\mathbf{0} \in S$, to S jest liniowo zależny*
(4) *S jest liniowo zależny wtedy i tylko wtedy, gdy dla pewnego i wektor v_i jest kombinacją liniową pozostałych wektorów z S.*

DEFINICJA 4.12. *Ciąg $(v_1, \ldots, v_k)$ wektorów z V nazywamy bazą, jezeli jest liniowo niezależny i $\langle\{v_1, \ldots, v_k\}\rangle = V$*

Przykład:
Niech $(\delta_{1i}, \ldots, \delta_{ni}) = e_i \in \mathbb{K}^n$. Ciąg $(e_1, e_2, \cdots, e_n)$ jest bazą w $\mathbb{K}^n$.
Bazą przestrzeni jednopunktowej $V = \{\mathbf{0}\}$ jest zbiór pusty.

STWIERDZENIE 4.13. *Jeżeli przestrzeń wektorowa V ma bazę, to każdy jej wektor da się przedstawić jednoznacznie jako kombinacja liniowa wektorów bazy.*

DOWÓD: Niech $(v_1, v_2, \cdots, v_n)$ bedzie bazą przestrzeni V. Wektory bazy rozpinają całą przestrzeń, więc każdy wektor jest kombinacją liniową wektorów bazy. Niech teraz

$$v = \lambda^1 v_1 + \cdots + \lambda^n v_n = \mu^1 v_1 + \cdots + \mu^n v_n,$$

a stąd

$$(\lambda^1 - \mu^1)v_1 + \cdots + (\lambda^n - \mu^n)v_k = \mathbf{0}.$$

Z liniowej niezależności wektorów bazy współczynniki kombinacji liniowej równej zero są zerowe:

$$\lambda^1 = \mu^1, \ldots, \lambda^n = \mu^n.$$

■

TWIERDZENIE 4.14. *Jeśli przestrzeń wektorowa V posiada bazę n-elementową i $S = \{w_1, \ldots, w_k\}$, przy czym $k > n$, to układ wektorów S jest liniowo zależny.*

DOWÓD: Niech $(v_1, \ldots, v_k)$ będzie bazą w V. Wówczas $w_i = \sum_j \lambda_i^j v_j$. Stąd równość

$$\mu^1 w_1 + \cdots + \mu^k w_k = \mathbf{0}$$

jest równoważna równości,

$$(\sum_i \mu_i \lambda_i^1)v_1 + \cdots + (\sum_i \mu^i \lambda_i^n)v_n = \mathbf{0},$$

a ta, ponieważ $(v_1, \cdots, v_k)$ jest bazą, równoważna jest układowi równań

$$\begin{cases} \mu^1 \lambda_1^1 + \mu^2 \lambda_2^1 + \cdots + \mu^k \lambda_k^1 = \mathbf{0} \\ \qquad \cdots \\ \mu^1 \lambda_1^n + \mu^2 \lambda_2^n + \cdots + \mu^k \lambda_k^n = \mathbf{0} \end{cases}$$

Ponieważ $k > n$, więc istnieje (patrz przykład-wytrych, Twierdzenie 4.5) niezerowe rozwiązanie $(\mu^1, \ldots, \mu^k)$ tego układu i, w konsekwencji,

$$\mu^1 w_1 + \cdots + \mu^k w_k = \mathbf{0}.$$

■

Wnioski:

(1) Jeżeli $(v_1, \ldots, v_n)$ jest bazą i układ wektorów $\{w_1, \ldots, w_k\}$ jest liniowo niezależny, to $k \leqslant n$.
(2) Jeżeli $(v_1, \ldots, v_n)$ i $(w_1, \ldots, w_m)$ są bazami w V, to $m = n$.

TWIERDZENIE 4.15. *Jeśli V jest przestrzenią skończenie wymiarową i $V \neq \{\mathbf{0}\}$, to V posiada niepustą bazę.*

Dowód: Niech $\langle\{v_1, \ldots, v_k\}\rangle = V$. Mamy dwie możliwości:

a) układ wektorów $(v_1, \ldots, v_k)$ jest liniowo niezależny, więc jest bazą,
b) układ $(v_1, \cdots, v_k)$ jest liniowo zależny, więc istnieje wektor v_i będący kombinacją liniową pozostałych.

W drugim przypadku

$$\langle\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}\rangle = \langle\{v_1, \ldots, v_k\}\rangle = V.$$

Znów mamy dwie możliwości

a) układ $\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}$ jest bazą, lub
b) jeden z wektorów $\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}$ jest liniową kombinacją pozostałych. W tym przypadku postępujemy jak poprzednio.

i.t.d.

W efekcie końcowym dostajemy bazę lub pojedyńczy wektor v_j taki, że $\langle\{v_j\}\rangle = V$. Ponieważ V zawiera element różny od zera, więc $v_j \neq \mathbf{0}$, a stąd układ $\{v_j\}$ jest liniowo niezależny, więc jest bazą. ■

Wniosek: Dla przestrzeni wektorowych wymiaru skończonego liczba elementów bazy jest dobrze określoną funkcją przestrzeni. Oznaczamy ją dim i nazywamy *wymiarem.* dim V jest liczbą elementów dowolnej bazy przestrzeni V. Dla podkreślenia z jakim ciałem mamy do czynienia, piszemy $\dim_{\mathbb{K}}$.

Przykłady:

(1) $\dim_{\mathbb{K}} \mathbb{K}^n = n$. Jako bazę możemy wybrać układ $(e_1, e_2, \ldots, e_n)$ (przykład po Definicji 4.12).
(2) $\dim_{\mathbb{C}} \mathbb{C}^1 = 1$.
(3) $\dim_{\mathbb{R}} \mathbb{C}^1 = 2$. Jako bazę możemy wybrać parę $(1, \imath)$.

TWIERDZENIE 4.16. *Niech V będzie przestrzenią skończenie wymiarową nad ciałem $\mathbb{K}$. Dowolny ciąg wektorów liniowo niezależnych da się uzupełnić do bazy.*

DOWÓD: Jeśli układ $(v_i, \cdots, v_k)$ jest liniowo niezależny i nie jest bazą, to istnieje wektor $v_{k+1} \in V$ taki, że v_{k+1} nie należy do $\langle(v_1, \cdots, v_k)\rangle$. Układ $(v_1, \cdots, v_k)$ jest liniowo niezależny, więc jeśli $\lambda^1 v_1 + \cdots + \lambda^k v_k + \lambda^{k+1} v_{k+1} = \mathbf{0}$ to mamy dwie możliwości:

a) $\lambda^{k+1} \neq 0$ oznacza, że v_{k+1} jest zależny od $(v_i, \cdots, v_k)$. Sprzeczność.
b) $\lambda^{k+1} = 0$ czyli $\lambda^1 v_1 + \cdots + \lambda^k v_k = \mathbf{0}$ i stąd $\lambda^1 = \cdots = \lambda^k = 0$.

Wynika stąd, że układ $(v_1, \cdots, v_k, v_{k+1})$ jest liniowo niezależny. Znów mamy dwie możliwości:
$k + 1 = \dim V$, czyli układ $(v_1, \cdots, v_k, v_{k+1})$ jest bazą, lub $\langle(v_1, \cdots, v_{k+1})\rangle \neq V$. W drugim przypadku powtarzamy całą procedurę.

Po $n - k$ krokach otrzymujemy układ n liniowo niezależnych wektorów, które tworzą bazę. ■

4.2.1. Dalsze przykłady przestrzeni wektorowych.

(G) Niech V i W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$. Iloczyn kartezjański $V \times W$ z działaniami:
a) $(v, w) + (v', w') = (v + v', w + w')$
b) $\lambda(v, w) = (\lambda v, \lambda w)$
jest też przestrzenią wektorową nad ciałem $\mathbb{K}$. Nazywamy ją *iloczynem kartezjańskim przestrzeni wektorowych.*

Jeśli układ $(v_1, \cdots, v_n)$ jest bazą V i układ $(w_1, \cdots, w_m)$ jest bazą W, to układ $n + m$ wektorów

$$((v_1, \mathbf{0}), \cdots, (v_n, \mathbf{0}), (\mathbf{0}, w_1), \cdots, (\mathbf{0}, w_m))$$

tworzy bazę $V \times W$.

Istotnie, jeśli

$$\lambda^1(v_1, \mathbf{0}) + \cdots + \lambda^n(v_n, \mathbf{0}) + \mu^1(\mathbf{0}, w_1) + \cdots + \mu^m(\mathbf{0}, w_m) \\ = (\lambda^1 v_1 + \cdots + \lambda^n v_n, \mu^1 w_1 + \cdots + \mu^m w_m) = \mathbf{0},$$

to

$$\lambda^1 = \cdots = \lambda^n = 0, : \mu^1 = \cdots = \mu^m = 0.$$

Stąd mamy

STWIERDZENIE 4.17. $\dim(V \times W) = \dim V + \dim W$

Niech V będzie przestrzenią wektorową i niech $W_1, W_2 \subset V$ będą jej podprzestrzeniami. Wówczas

(H) $W_1 \cap W_2$ jest podprzestrzenią wektorową

(I) Zbiór $W_1 \cup W_2$ nie jest w ogólności przestrzenią wektorową. (Jeżeli jest, to $W_1 \subset W_2$ lub $W_2 \subset W_1$.) *Sumą algebraiczną* podprzestrzeni W_1 i W_2 nazywamy podprzestrzeń $\langle W_1 \cup W_2 \rangle$ i oznaczamy ją $W_1 + W_2$. Jest to najmniejsza podprzestrzeń zawierająca W_1 i W_2.

Uwaga. Reprezentacja wektora $v \in W_1 + W_2$ jako sumy $v = w_1 + w_2$, gdzie $w_1 \in W_1$ a $w_2 \in W_2$, nie jest na ogół jednoznaczna np. dla $W_1 = W_2 = W$ mamy $W_1 + W_2 = W$ i $w = \mathbf{0} + w = w + \mathbf{0}$.

TWIERDZENIE 4.18.
Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$, a W_1 i W_2 jej podprzestrzeniami. Poniższe warunki są równoważne:

a) $W_1 \cap W_2 = \{\mathbf{0}\}$,

b) *dla każdego $v \in W = W_1 + W_2$ istnieją jednoznacznie określone wektory $w_1 \in W_1$, $w_2 \in W_2$ takie, że $v = w_1 + w_2$,*

c) *zachodzi wynikanie:*
jeśli $w_1 + w_2 = \mathbf{0}$ gdzie $w_1 \in W_1$ i $w_2 \in W_2$, to $w_1 = w_2 = \mathbf{0}$.

DOWÓD:
$a \Rightarrow b$ Niech $w_1 + w_2 = w_1' + w_2'$. Stąd $(w_1 - w_1') = (w_2' - w_2) = \mathbf{0}$, czyli $w_1 = w_1'$ i $w_2 = w_2'$, gdzie $(w_1 - w_1') \in W_1$ a $(w_2') \in W_2$.

$b \Rightarrow c$ Niech $\mathbf{0} + \mathbf{0} = \mathbf{0} = w_1 + w_2$. Stąd $w_1 = \mathbf{0}$ i $w_2 = \mathbf{0}$.

$c \Rightarrow a$ Niech $w \in W_1 \cap W_2$. Kładć $w_1 = w \in W_1$ i $w_2 = -w \in W_2$ dostajemy $w_1 + w_2 = \mathbf{0}$. Z jednoznaczności rozkładu $w_1 = w_2 = \mathbf{0}$, czyli $w = \mathbf{0}$. ■

Jeżeli spełnione są warunki o których mówi twierdzenie, wprowadzamy oznaczenie $W_1 + W_2 = W_1 \oplus W_2$ i mówimy, że mamy *sumę prostą* podprzestrzeni W_1 i W_2.

Ogólnie: jeżeli mamy ciąg podprzestrzeni $W_1, \cdots, W_k$, to

$$W_1 + \cdots + W_k = \langle W_1 \cup \cdots \cup W_k \rangle.$$

Jeżeli reprezentacja $w = w_1 + \cdots + w_k$, $w_i \in W_i$, wektora w jest jednoznaczna, to piszemy

$$W_1 + \cdots + W_k = W_1 \oplus \cdots \oplus W_k = \oplus_{i=1}^k W_i.$$

TWIERDZENIE 4.19. $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$

DOWÓD: Niech $(w_1, \dots, w_k)$ będzie bazą $W_1 \cap W_2$. Można ją uzupełnić do bazy

$$(w_1, \dots, w_k, v_1, \dots, v_m)$$

w W_1 i do bazy

$$(w_1, \dots, w_k, u_1, \dots, u_n)$$

w W_2. Twierdzę, że

$$(w_1, \dots, w_k, v_1, \dots, v_m, u_1, \dots, u_n)$$

jest bazą $W_1 + W_2$. Sprawdzenie: oczywiste, że te wektory rozpinają całą podprzestrzeń $W_1 + W_2$. Czy układ jest liniowo niezależny? Jeżeli

$$\lambda^1 w_1 + \cdots + \lambda^k w_k + \mu^1 v_1 + \cdots + \mu^m v_m + \nu^1 u_1 + \cdots + \nu^n u_n = \mathbf{0},$$

to $\lambda^1 w_1 + \cdots + \mu^m v_m \in W_1 \cap W_2$, czyli również $\mu^1 v_1 + \cdots + \mu^m v_m \in W_1 \cap W_2$, a to oznacza, że jest kombinacją liniową wektorów $w_1, \dots, w_k$. Ponieważ układ wektorów $(w_1, \dots, w_k, v_1, \dots, v_m)$ jest liniowo niezależny, dostajemy $\mu_1 = \cdots = \mu_m = 0$. Podobnie $\nu_1 = \cdots = \nu_n = 0$ i, w konsekwencji, również $\lambda_1 = \cdots = \lambda_k = 0$. ■

(J) Niech V będzie przestrzenią wektorową a W jej podprzestrzenią wektorową. W V wprowadzamy relację równoważności zdefiniowaną następująco:

$$v_1 \sim v_2 \equiv v_1 - v_2 \in W\ .$$

Zbiór klas równoważności oznaczamy V/W. W zbiorze tym określamy działanie dodawania

$$[v_1] + [v_2] = [v_1 + v_2]$$

i mnożenia przez liczbę

$$\lambda[v] = [\lambda v],$$

gdzie $[v]$ jest klasą równoważności wektora v. Łatwo sprawdzić, że wyniki działań nie zależą od wyboru reprezentanta i że określają w V/W strukturę przestrzeni wektorowej. Nazywamy ją *przestrzenią ilorazową* przestrzeni V przez podprzestrzeń W.

Przykład: całka nieoznaczona może być uważana jako elementem przestrzeni ilorazowej przestrzeni funkcji ciągłych przez podprzestrzeń funkcji stałych.

STWIERDZENIE 4.20. $\dim V/W = \dim V - \dim W$

DOWÓD: Niech $(w_1, w_2, \cdots, w_n)$ będzie bazą W. Uzupełniamy ją do bazy

$$(w_1, \ldots, w_n, v_1, \ldots, v_m)$$

w V. Pokażemy, że $([v_1], \ldots, [v_m])$ jest bazą w V/W. Istotnie, niech $v \in V$, wówczas

$$v = \lambda^1 w_1 + \cdots + \lambda^n w_n + \mu^1 v_1 + \cdots + \mu^m v_m$$

i stąd

$$\begin{aligned}[v] &= [\mu^1 v_1 + \cdots + \mu^m v_m] \\ &= \mu^1 [v_1] + \cdots + \mu^m [v_m].\end{aligned}$$

Ponadto, jeśli

$$\mu^1 [v_1] + \cdots + \mu^m [v_m] = [\mu^1 v_1 + \cdots + \mu^m v_m] = \mathbf{0},$$

to $\mu^1 v_1 + \cdots + \mu^m v_m \in W$, czyli $\mu^1 v_1 + \cdots + \mu^m v_m = 0$ i stąd $\mu_1 = \cdots = \mu_m = 0$. Ciąg $([v_1], \ldots, [v_m])$ rozpina więc W/V i jest liniowo niezależny. ■

Rozdział 5. Odwzorowania liniowe

5.1. Definicja i postawowe własności.

DEFINICJA 5.1. *Niech V, W będą przestrzeniami wektorowymi nad ciałem $\mathbb{K}$. Odwzorowanie $F: V \to W$ nazywamy* liniowym, *jeżeli $\forall v_1, v_2 \in V$ i $\forall \lambda_1, \lambda_2 \in \mathbb{K}$,*

$$F(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 F(v_1) + \lambda_2 F(v_2).$$

Równoważnie, odwzorowanie jest liniowe, jeżeli spełnione są dwa warunki:

$$F(v_1 + v_2) = F(v_1) + F(v_2) \quad \text{i} \quad F(\lambda v) = \lambda F(v).$$

Z definicji wynika natychmiast, że

$$F(\mathbf{0}) = \mathbf{0}.$$

Istotnie, $F(\mathbf{0}) = F(0 \cdot \mathbf{0}) = 0 \cdot F(\mathbf{0}) = \mathbf{0}$.

Przykłady.

(1) $V = \mathscr{C}([-1, 1])$ - przestrzeń funkcji ciągłych na odcinku $[-1, 1]$ i $W = \mathbb{R}^1$. Definiujemy odwzorowanie $F: V \to W$ wzorem $F(f) = f(0)$. Liniowość F jest oczywista.
(2) $V = \mathscr{C}^1(\mathbb{R}^1)$, $W = \mathscr{C}(\mathbb{R}^1)$ i $F(f) = f'$ (pochodna funkcji f).
(3) $V = W = \mathbb{R}^1$. Które z odwzorowań:

$$F_1(x) = x^2, \;\; F_2(x) = x + 1, \;\; F_3(x) = 4x$$

jest liniowe?

Odwzorowania liniowe z V do W są elementami przestrzeni wektorowej $\mathrm{Map}(V, W)$ wszystkich odwzorowań z V do W (patrz przykład (E) z poprzedniego rozdziału). Pokażemy, że tworzą one podprzestrzeń wektorową.

STWIERDZENIE 5.2. *Niech F, $G: V \to W$ będą odwzorowaniami liniowymi i niech $\lambda \in K$. Wówczas*

(1) *$F + G$ jest odwzorowaniem liniowym,*
(2) *λF jest odwzorowaniem liniowym.*

DOWÓD: Zgodnie z definicją działań w $\mathrm{Map}(V, W)$

$$\begin{aligned}(F + G)(v_1 + v_2) &= F(v_1 + v_2) + G(v_1 + v_2)\\ &= F(v_1) + F(v_2) + G(v_1) + G(v_2)\\ &= (F + G)(v_1) + (F + G)(v_2).\end{aligned}$$

Podobnie

$$(F+G)(\mu v) = F(\mu v) + G(\mu v) = \mu F(v) + \mu G(v) = \mu(F+G)(v).$$

Zatem $F+G$ jest odwzorowaniem liniowym. Tak samo pokazujemy, że λF jest liniowe. ■

Wniosek: Wszystkie odwzorowania liniowe z V do W tworzą przestrzeń wektorową; oznaczamy ją $\mathsf{L}(V,W)$. Ponadto używać będziemy oznaczeń:

$$\mathsf{L}(V,V) = \mathrm{End}(V) \ \text{ i } \mathsf{L}(V,\mathbb{K}) = V^*.$$

STWIERDZENIE 5.3.
Niech V, W, U będą przestrzeniami wektorowymi nad $\mathbb{K}$. Jeżeli $F\colon V\to W$ oraz $G\colon W\to U$ są odwzorowaniami liniowymi, to złożenie $G\circ F\colon V\to U$ jest też odwzorowaniem liniowym.

DOWÓD: Mamy

$$\begin{aligned} G\circ F(\lambda_1 v_1 + \lambda_2 v_2) &= G(F(\lambda_1 v_1 + \lambda_2 v_2)) \\ &= G(\lambda_1 F(v_1) + \lambda_2 F(v_2)) \\ &= \lambda_1 G(F(v_1)) + \lambda_2 G(F(v_2)) \\ &= \lambda_1 G\circ F(v_1) + \lambda_2 G\circ F(v_2) \end{aligned}$$

■

Uwaga: Niech $F\colon V\to\mathbb{K}^n$ będzie jakimś odwzorowaniem. Ponieważ odwzorowania

$$\pi^i\colon \mathbb{K}^n \to \mathbb{K}^1\colon (x_1, x_2, \cdots, x_n) \mapsto x_i$$

są liniowe to, jak łatwo zauważyć, F jest liniowe wtedy i tylko wtedy, gdy dla każdego i złożenie $\pi^i\circ F$ jest odwzorowaniem liniowym.

STWIERDZENIE 5.4. *Odwzorowanie liniowe F jest wyznaczone jednoznacznie przez jego wartości na wektorach bazy.*

DOWÓD: Niech $(e_1,\dots,e_n)$ będzie bazą V i niech $v\in V$. Wówczas $v=\lambda^1 e_1 + \cdots + \lambda^n e_n$ i, z liniowości F, mamy $F(v) = \lambda^1 F(e_1) + \cdots + \lambda^n F(e_n)$. ■

STWIERDZENIE 5.5. *Jeżeli $F\colon V\to W$ i $V_1\subset V$ jest podprzestrzenią wektorową, to $F(V_1)$ jest podprzestrzenią wektorową przestrzeni W i $\dim F(V_1)\leqslant \dim V_1$.*

DOWÓD: To że $F(V_1)$ jest podprzestrzenią wektorową wynika natychmiast z liniowości F. Jeżeli $(e_1,\dots,e_{n_1})$ jest bazą V_1, podprzestrzeń $F(V_1)$ jest rozpięta na wektorach $F(e_1),\dots,F(e_{n_1})$. ■

STWIERDZENIE 5.6. *Jeżeli $F \in \mathsf{L}(V, W)$ jest bijekcją (tzn. F^{-1} istnieje), to $F^{-1} \in \mathsf{L}(W, V)$.*

DOWÓD: Niech $w_1,\ w_2 \in W$. Istnieją $v_1,\ v_2$ takie, że $F(v_1) = w_1$ i $F(v_2) = w_2$. Wówczas

$$\begin{aligned} F^{-1}(\lambda_1 w_1 + \lambda_2 w_2) &= F^{-1}(\lambda_1 F(v_1) + \lambda_2 F(v_2)) \\ &= F^{-1}(F(\lambda_1 v_1 + \lambda_2 v_2)) \\ &= \lambda_1 v_1 + \lambda_2 v_2 \\ &= \lambda_1 F^{-1}(w_1) + \lambda_2 F^{-1}(w_2) \end{aligned}$$

■

Oznaczenia: $\mathsf{Aut}(V) = \mathsf{Gl}(V) = \{F \in \mathsf{End}(V) \colon F^{-1}$ jest odwzorowaniem$\}$.
Jeżeli $F \in \mathsf{L}(V, W)$ jest takie, że F^{-1} istnieje, to mówimy, że F jest *izomorfizmem przestrzeni wektorowych*.

Przykłady:

(1) Jako V weźmy przestrzeń $\mathbb{K}[n]$ wielomianów stopnia $\leqslant n$ i współczynnikach z $\mathbb{K}$. Odwzorowanie liniowe

$$F \colon \mathbb{K}[n] \to \mathbb{K}^{n+1} \colon a_0 + a_1 x^1 + \cdots + a_n x^n \mapsto (a_0, a_1, \dots, a_n) \in \mathbb{K}^{n+1} \tag{5.1}$$

jest izomorfizmem.

(2) Przy ustalonej permutacji $\sigma \in \mathsf{S}(n)$ definiujemy odwzorowanie

$$\mathbb{K}^n \ni (x_1, \dots, x_n) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \in \mathbb{K}^n. \tag{5.2}$$

Jest to izomorfizm przestrzeni $\mathbb{K}^n$ w siebie (tzn. automorfizm).

5.2. Obraz i jądro odwzorowania liniowego.

STWIERDZENIE 5.7. *Jeżeli odwzorowanie $F \colon V \to W$ jest liniowe, to zbiory $F(V) \subset W$ i $F^{-1}(\mathbf{0}) \subset V$ są podprzestrzeniami wektorowymi.*

DOWÓD:

(1) Jeżeli $w_1, w_2 \in F(V)$ to istnieją wektory $v_1, v_2 \in V$ takie,że $w_1 = F(v_1)$ i $w_2 = F(v_2)$. Stąd $\lambda^1 w_1 + \lambda^2 w_2 = \lambda^1 F(v_1) + \lambda^2 F(v_2) = F(\lambda^1 v_1 + \lambda^2 v_2)$, więc $\lambda^1 w_1 + \lambda^2 w_2 \in F(V)$.

(2) Jeżeli $F(v_1) = \mathbf{0}$ i $F(v_2) = \mathbf{0}$ to $F(\lambda^1 v_1 + \lambda^2 v_2) = \lambda^1 F(v_1) + \lambda^2 F(v_2) = \mathbf{0}$.

■

Wniosek: Jeżeli $U \subset V$ jest podprzestrzenią wektorową i $F \colon V \to W$ jest liniowe, to $F(U) \subset W$ też jest podprzestrzenią wektorową.

Terminologia i oznaczenia:

Podprzestrzeń wektorową $F(V)$ przestrzeni W nazywamy *obrazem* odwzorowania liniowego F i oznaczamy $\operatorname{im} F$. Podprzestrzeń wektorową $F^{-1}(\mathbf{0})$ przestrzeni V nazywamy *jądrem* odwzorowania liniowego F i oznaczamy $\ker F$.

STWIERDZENIE 5.8. *Niech $F: V \to W$ będzie odwzorowaniem liniowym. Wówczas*

$$F(v_1) = F(v_2) \iff v_1 - v_2 \in \ker F.$$

Wnioski:

(1) F jest injekcją wtedy i tylko wtedy, gdy $\ker F = \{\mathbf{0}\}$,
(2) F jest izomorfizmem wtedy i tylko wtedy, gdy $\operatorname{im} F = W$ i $\ker F = \{\mathbf{0}\}$

TWIERDZENIE 5.9. *Niech V będzie skończenie wymiarową przestrzenią wektorową. Jeżeli $F: V \to W$ jest izomorfizmem, to przestrzeń W jest wymiaru skończonego i $\dim V = \dim W$.*

DOWÓD: Niech $(e_1, \dots, e_n)$ będzie bazą V, wówczas $(F(e_1), \dots, F(e_n))$ rozpina $\operatorname{im} F = W$. W jest więc wymiaru skończonego. Jeżeli z kolei

$$\mathbf{0} = \lambda^1 F(e_1) + \cdots + \lambda^n F(e_n) = F(\lambda^1 e_1 + \cdots + \lambda^n e_n),$$

to

$$\lambda^1 e_1 + \cdots + \lambda^n e_n = \mathbf{0},$$

czyli $\lambda^1 = \cdots = \lambda^n = 0$. Stąd wynika, że układ $(F(e_1), \dots, F(e_n))$ jest liniowo niezależny, więc jest bazą przestrzeni W. ■

Uwaga: Jeżeli $\dim V = \dim W$, to zawsze istnieje izomorfizm $F: V \to W$. Na przykład, jeżeli $(e_1, \dots, e_n)$ jest bazą V i $(f_1, \dots, f_n)$ jest bazą W, to kładąc $F(e_i) = f_i$ dostajemy izomorfizm.

TWIERDZENIE 5.10. *Jeżeli $F \in \mathsf{L}(V, W)$ to*

$$\dim V = \dim(\ker F) + \dim(\operatorname{im} F). \tag{5.3}$$

DOWÓD: Niech $V_1 = \ker F$. Z Twierdzenia 4.16 wynika, że istnieje podprzestrzeń $V_2 \subset V$ taka, że $V = V_1 \oplus V_2$. Oznaczmy przez F_2 odwzorowanie F obcięte do V_2. Mamy więc

$$\operatorname{im} F = \operatorname{im} F_2 \overset{\text{ozn}}{=} W_1.$$

Odwzorowanie $F_2: V_2 \to W_1$ jest bijekcją, więc izomorfizmem. Stąd

$$\dim \operatorname{im} F = \dim W_1 = \dim V_2.$$

Ponieważ $\dim V = \dim V_2 + \dim V_1$, mamy $\dim V = \dim \operatorname{im} F + \dim \ker F$. ■

Wnioski:

(1) $F \in \mathsf{L}(V, W)$ i F jest surjekcją, to $\dim V \geqslant \dim W$,
(2) $F \in \mathsf{L}(V, W)$ i F jest injekcją, to $\dim V \leqslant \dim W$,
(3) $\dim V > \dim W$, to $\ker F \neq \{\mathbf{0}\}$

Czytelnikowi pozostawiam do udowodnienia dwa proste, ale ważne fakty:

a) Niech $W \subset V$ będzie podprzestrzenią wektorową o wymiarze m. Istnieje odwzorowanie liniowe $F\colon \mathbb{K}^m \to V$ takie, że $W = \operatorname{im} F$. Mówimy, że F jest opisem parametrycznym podprzestrzeni W.
b) Niech $W \subset V$ będzie jak wyżej i niech $\dim V = n$. Istnieje odwzorowanie liniowe $F\colon V \to \mathbb{K}^{n-m}$ takie, że $W = \ker F$. Mówimy w tym przypadku, że podprzestrzeń jest opisana układem równań (patrz punkt następny).

Uwagi końcowe:

(1) $\mathsf{End}(V)$ jest przestrzenią wektorową.
(2) $\mathsf{Aut}(V)$ nie jest przestrzenią wektorową, ale jest grupą ze względu na składanie odwzorowań.
(3) Odwzorowanie $F \in \mathsf{L}(V, W)$ indukuje odwzorowanie liniowe

$$\tilde{F}\colon V/\ker F \to W\colon [v] \mapsto F(v).$$

Odwzorowanie to jest injekcją.

5.3. Równania liniowe (teoria ogólna).

Równaniem liniowym nazywamy równanie postaci $F(x) = b$ gdzie $F \in \mathsf{L}(V, W)$, $b \in W$. Inaczej mówiąc, szukamy $x \in V$ takich, że $Fx = b$. Jeśli $b = \mathbf{0}$ to równanie nazywamy *jednorodnym* a jeśli $b \neq \mathbf{0}$ to równanie nazywamy *niejednorodnym*.

Fakty oczywiste:

(1) Aby zbiór rozwiązań równania $Fx = b$ był niepusty (inaczej mówiąc – aby istniało rozwiązanie równania $Fx = b$) potrzeba i wystarcza, by $b \in \operatorname{im} F$.
(2) Jeśli $b = \mathbf{0}$, to zbiór rozwiązań jest niepusty ($F\mathbf{0} = \mathbf{0}$).
(3) Jeśli $b = \mathbf{0}$, to zbiorem rozwiązań jest $\ker F$. W tym przypadku zbiór rozwiązań jest podprzestrzenią wektorową (dla $b \neq 0$, jak łatwo sprawdzić, nie jest).
(4) Jeśli x_1, x_2 są rozwiązaniami równania $Fx = b$, to $x_1 - x_2 \in \ker F$ czyli $x_1 - x_2$ jest rozwiązaniem równania jednorodnego $Fx = \mathbf{0}$.
(5) Jeśli x_1 jest rozwiązaniem równania $Fx = b$ i $x_0 \in \ker F$, to $x_1 + x_0$ jest też rozwiązaniem równania $Fx = b$.
(6) Jeżeli F jest izomorfizmem, to dla każdego b istnieje dokładnie jedno rozwiązanie równania $Fx = b$. Równanie takie nazywa się układem Cramera.

Jeżeli w V mamy bazę $(e_1, e_2, \ldots, e_n)$, to punkt 1 równoważny jest

(1') $b \in \langle F(e_1), \ldots, F(e_n)\rangle$, co z kolei jest równoważne
(1")

$$\langle F(e_1), \ldots, F(e_n)\rangle = \langle F(e_1), \ldots, F(e_n), b\rangle. \tag{5.4}$$

Jak opisać zbiór rozwiązań równania $Fx = b$?

Jeżeli $b = 0$ to wystarczy podać bazę podprzestrzeni $\ker F$. Nazywamy ją *fundamentalnym układem rozwiązań*. Jeżeli $b \neq 0$ to, jak wynika z punktu 5, należy podać jedno rozwiązanie (szczególne) równania $Fx = b$ i fundamentalny układ rozwiazań równania jednorodnego $Fx = \mathbf{0}$.

Innym sposobem opisu jest podanie bijekcji jakiegoś zbioru parametrów (najczęściej jest to odwzorowanie liniowe z przestrzeni $\mathbb{K}^l$) w zbiór rozwiązań.

Przykład. Niech $F\colon \mathbb{R}^2 \mapsto \mathbb{R}^2\colon (x, y) \longmapsto (x + y, 2x + 2y)$ i niech $b = (2, 4)$
Rozwiązania można sparametryzować następująco: $\mathbb{R}^1 \ni \lambda \mapsto (\lambda + 1, 1 - \lambda)$.

Rozdział 6. Przestrzeń macierzy. Macierze odwzorowań liniowych

6.1. Definicja i podstawowe operacje.

DEFINICJA 6.1. Macierzą *o m wierszach, n kolumnach i o elementach ze zbioru X nazywamy odwzorowanie* $\{1, \cdots, m\} \times \{1, \cdots, n\} \to X$.

Na macierz możemy patrzeć jak na „tabliczkę" o m wierszach i n kolumnach, złożoną z elementów ze zbioru X. Będziemy pisać

$$\begin{bmatrix} a^1{}_1 & a^1{}_2 & \cdots & a^1{}_n \\ \vdots & \vdots & & \vdots \\ a^m{}_1 & a^m{}_2 & \cdots & a^m{}_n \end{bmatrix} = [a^i{}_j]$$

Zbiór macierzy o m wierszach, n kolumnach i o elementach z X oznaczamy $\mathbf{M}^m{}_n(X)$.

W dalszym ciągu będziemy się zajmować macierzami, dla których $a^i{}_j \in \mathbb{K}$. Nazywać je będziemy macierzami liczbowymi. W zbiorze $\mathbf{M}^m{}_n(\mathbb{K})$ określamy dodawanie i mnożenie przez liczbę:

$$[a^i{}_j] + [b^i{}_j] = [a^i{}_j + b^i{}_j]$$
$$\lambda[a^i{}_j] = [\lambda a^i{}_j]$$

Z tymi działaniami $\mathbf{M}^m{}_n(\mathbb{K})$ tworzy, co łatwo sprawdzić, przestrzeń wektorową nad $\mathbb{K}$. Układ macierzy $E^k_l = [\delta_{ik}\delta_{jl}]$ (tzn. tylko jeden element macierzy jest różny od zera - leżący w k-tym wierszu i l-tej kolumnie) rozpina przestrzeń $\mathbf{M}^m{}_n(\mathbb{K})$ i jest liniowo niezależny (po ułożeniu w ciąg stanowi więc bazę). Wynika stąd, że $\dim_{\mathbb{K}} \mathbf{M}^m{}_n(\mathbb{K}) = n \cdot m$.

Wprowadzimy operację na macierzach zwaną *transpozycją*:

$$^{\mathsf{T}} \colon \mathbf{M}^m{}_n(\mathbb{K}) \to \mathbf{M}^n{}_m(\mathbb{K}) \colon A \mapsto A^{\mathsf{T}}$$

zdefiniowaną następująco: jeśli $A = [a^i{}_j]$, to $A^{\mathsf{T}} = [a^{\mathsf{T}i}{}_j]$ gdzie $a^{\mathsf{T}i}{}_j = a^j{}_i$. macierz, transponowana Transpozycja jest izomorfizmem przestrzeni wektorowych.

Przez $\bar{a}^i \in \mathbf{M}^1{}_n(\mathbb{K})$ oznaczać będziemy i-ty wiersz, a przez $\bar{a}_j \in \mathbf{M}^m{}_1(\mathbb{K})$ j-tą kolumnę macierzy $[a^i{}_j]$.

Oczywiste, że odwzorowania

$$A = [a^i{}_j] \mapsto (\bar{a}^1, \dots, \bar{a}^m)$$
$$A = [a^i{}_j] \mapsto (\bar{a}_1, \dots, \bar{a}_n)$$

zadają izomorfizmy

$$\mathbf{M}^m{}_n(\mathbb{K}) \simeq \underbrace{\mathbf{M}^m{}_1(\mathbb{K}) \times \cdots \times \ \mathbf{M}^m{}_1(\mathbb{K})}_{n \ razy}$$

$$\mathbf{M}^m{}_n(\mathbb{K}) \simeq \underbrace{\mathbf{M}^1{}_n(\mathbb{K}) \times \cdots \times \ \mathbf{M}^1{}_n(\mathbb{K})}_{m \ razy}$$

W dalszym ciągu będziemy (czasami) oznaczać macierz A jako wiersz kolumn

$$A = [\bar{a}_1, \ldots, \bar{a}_n]$$

lub jako kolumnę wierszy

$$A = \begin{bmatrix} \bar{a}^1 \\ \vdots \\ \bar{a}^m \end{bmatrix}.$$

DEFINICJA 6.2.
Rzędem wierszowym *macierzy* $A = [a^i{}_j]$ *nazywamy liczbę* $\dim\langle\{\bar{a}^1, \ldots, \bar{a}^m\}\rangle$.
Rzędem kolumnowym *macierzy* $A = [a^i{}_j]$ *nazywamy* $\dim\langle\{\bar{a}_1, \ldots, \bar{a}_n\}\rangle$.

TWIERDZENIE 6.3. *Rząd wierszowy jest równy rzędowi kolumnowemu.*

Dowód: Niech r będzie rzędem wierszowym, a s rzędem kolumnowym macierzy A. Niech układ $(\bar{a}^{i_1} \cdots \bar{a}^{i_r})$ będzie bazą $\langle\{\bar{a}^1, \ldots, \bar{a}^m\}\rangle$, a układ $(\bar{a}_{j_1} \cdots \bar{a}_{i_s})$ bazą $\langle\{\bar{a}_1, \ldots, \bar{a}_n\}\rangle$.

Oznaczmy przez $B = [b^k{}_l]$ macierz otrzymaną z A przez wykreślenie wierszy i kolumn nie należących do bazy, tzn. $b^k{}_l = a^{i_k}{}_{j_l}$ gdzie $k = 1, \ldots, r$ i $l = 1, \ldots, s$

Pokażemy najpierw, że jeżeli $\bar{b}_1, \ldots, \bar{b}_s$ są liniowo zależne, to $\bar{a}_{j_1}, \ldots, \bar{a}_{j_s}$ też są liniowo zależne. Istotnie, jeśli $\lambda^1\bar{b}_1 + \cdots + \lambda^s\bar{b}_s = \mathbf{0}$ to, ponieważ dla $i \neq i_1, \ldots, i_r$ mamy $\bar{a}^i = \alpha_1\bar{a}^{i_1} + \cdots + \alpha_r\bar{a}^{i_r}$,

$$\begin{aligned} \lambda^1 a^i{}_{j_1} + \cdots + \lambda^s a^i{}_{j_s} &= \\ \lambda^1(\alpha_1 a^{i_1}{}_{j_r} + \cdots + \alpha_r a^{i_r}{}_{j_1}) + \cdots + \lambda^s(\alpha_1 a^{i_1}{}_{j_s} + \cdots + \alpha_r a^{i_r}{}_{j_s}) &= \\ \alpha_1(\lambda^1 a^{i_1}{}_{j_1} + \cdots + \lambda^s a^{i_1}{}_{j_s}) + \cdots + \alpha_r(\lambda^1 a^{i_r}{}_{j_1} + \cdots + \lambda^s a^{i_r}{}_{j_s}) &= 0 \end{aligned}$$

Wynika stąd, że kolumny $\{\bar{b}_1, \ldots, \bar{b}_s\}$ są liniowo niezależne, czyli układ równań:

$$\begin{array}{ccccccc} x^1 b^1{}_1 & + & \cdots & + & x^s b^1{}_s & = & 0 \\ \vdots & & \vdots & & \vdots & & \\ x^1 b^r{}_1 & + & \cdots & + & x^s b^r{}_s & = & 0 \end{array}$$

nie ma niezerowych rozwiązań. Zatem na mocy przykładu-wytrycha (Twierdzenie 4.5) dostajemy $s \leqslant r$. Symetrycznie, stosując powyższe rozumowanie do macierzy A^T dostajemy nierówność $r \leqslant s$. W konsekwencji $r = s$. ∎

DEFINICJA 6.4. *Rząd wierszowy (lub kolumnowy) macierzy A nazywamy* rzędem *macierzy A i oznaczamy* $\operatorname{rz} A$.

STWIERDZENIE 6.5.

(1) $\operatorname{rz} A = \operatorname{rz} A^{\mathsf{T}}$.
(2) *Jeżeli macierz B otrzymaliśmy z macierzy A przez dodanie do wiersza $\bar{a}^i$ kombinacji liniowej wierszy*

$$\bar{a}^1, \cdots, \bar{a}^{i-1}, \bar{a}^{i+1}, \cdots \bar{a}^m,$$

to $\operatorname{rz} B = \operatorname{rz} A$.
(3) *Jeżeli B otrzymaliśmy przez dodanie do ustalonej kolumny kombinacji liniowej pozostałych, to* $\operatorname{rz} B = \operatorname{rz} A$.
(4) *Jeżeli B otrzymaliśmy z A przez permutację kolumn (wierszy), to* $\operatorname{rz} A = \operatorname{rz} B$.

DOWÓD: Rząd wierszowy macierzy A jest równy rzędowi kolumnowemu macierzy transponowanej A^{T}. Stąd pierwsza równość.

Pozostałe równości wynikają z oczywistej uwagi, że przestrzenie rozpięte na wierszach (lub kolumnach) macierzy A i B są równe. ■

Zdefiniujemy teraz mnożenie macierzy. Dla każdych m, n, p jest to odwzorowanie $\mathbf{M}^n{}_m(\mathbb{K}) \times \mathbf{M}^m{}_p(\mathbb{K}) \to \mathbf{M}^n{}_p(\mathbb{K})$ zdefiniowane przez

$$(A, B) = ([a^i{}_j], [b^i{}_j]) \longmapsto AB = [c^i{}_j], \quad c^i{}_j = \sum_{k=1}^{m} a^i{}_k b^k{}_j$$

Uwagi:

(1) Mnożenie macierzy jest nieprzemienne, tzn., na ogół $AB \neq BA$. Znalezienie przykładu dla $m = n = 2$ zostawiamy jako ćwiczenie.
(2) Mnożenie macierzy jest łączne i rozdzielne względem dodawania.
(3) Jeżeli $A, B \in \mathbf{M}^n{}_n(\mathbb{K})$, to $AB \in \mathbf{M}^n{}_n(\mathbb{K})$. Z tym mnożeniem $\mathbf{M}^n{}_n(\mathbb{K})$ jest *algebrą*, to znaczy, przestrzenią wektorową wraz z działaniem (mnożeniem), które jest łączne i rozdzielne względem dodawania.
(4) Algebra ta posiada element neutralny ze względu na mnożenie (jedynkę). Jedynką jest macierz $I = [\delta^i{}_j]$, gdzie $\delta^i{}_j = 0$ dla $i \neq j$ i $\delta^i{}_i = 1$.
(5) Innym przykładem algebry jest przestrzeń endomorizmów $\mathsf{End}(V)$ przestrzeni wektorowej V z działaniem składania.
(6) Jeżeli $A \in \mathbf{M}^n{}_n(\mathbb{K})$, to macierz $B \in \mathbf{M}^n{}_n(\mathbb{K})$ taką, że $BA = I$ nazywamy *macierzą odwrotną* do macierzy A i oznaczamy A^{-1}. Łatwo zauważyć (ćwiczenie!), że nie każda macierz (nawet różna od zera ma do niej odwrotną).

Te i inne własności mnożenia macierzy wynikają natychmiast z interpretacji macierzy jako macierzy odwzorowań, o czym będzie mowa w następnej części.

6.2. Macierze odwzorowań.

Niech $e = (e_1, \dots, e_n)$ będzie bazą przestrzeni wektorowej V. Każdy wektor $v \in V$ ma jednoznaczną reprezentację $v = \lambda^1 e_1 + \dots + \lambda^n e_n$. Odwzorowanie

$$V \ni v \mapsto \begin{bmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{bmatrix} = [v]^e \in \mathbf{M}^n{}_1$$

jest izomorfizmem przestrzeni wektorowych.

Niech $f = (f_1, \dots, f_m)$ będzie bazą przestrzeni W i niech $F : V \to W$ będzie odwzorowaniem liniowym. Mamy

$$F(v) = \lambda^1 F(e_1) + \dots + \lambda^n F(e_n)$$

i

$$[F(v)]^f = \lambda^1 [F(e_1)]^f + \dots + \lambda^n [F(e_n)]^f = B \begin{bmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{bmatrix},$$

gdzie $B = [b^i{}_j]$ i $\bar{b}_j = [F(e_j)]^f$. Wprowadzoną tak macierz B oznaczać będziemy $[F]^f{}_e$. Nazywamy ją *macierzą odwzorowania* liniowego F w bazach e i f. Ponieważ

$$\begin{bmatrix} \lambda^1 \\ \vdots \\ \lambda^n \end{bmatrix} = [v]^e,$$

mamy

$$[F(v)]^f = [F]^f{}_e [v]^e. \tag{6.1}$$

STWIERDZENIE 6.6.

(1) $[F + G]^f{}_e = [F]^f{}_e + [G]^f{}_e$.
(2) $[\lambda F]^f{}_e = \lambda [F]^f{}_e$.
(3) *Odwzorowanie* $\mathsf{L}(V, W) \to \mathbf{M}^m{}_n : F \mapsto [F]^f{}_e$ *jest izomorfizmem przestrzeni wektorowych.*

DOWÓD: i-ta kolumna macierzy $[F + G]^f{}_e$ jest równa $[(F + G)(e_i)]^f$, ale

$$[(F + G)(e_i)]^f = [F(e_i) + G(e_i)]^f = [F(e_i)]^f + [G(e_i)]^f$$

i stąd pierwsza równość. Podobnie dostajemy drugą.

Wystarczy teraz pokazać, że odwzorowanie $F \mapsto [F]^f{}_e$ jest surjekcją (Stwierdzenie 5.6) (injektywność jest oczywista). Niech $A \in \mathbf{M}^m{}_n(\mathbb{K})$. Ze Stwierdzenia 5.3. wynika, że przyporządkowanie

$$e_i \mapsto \sum_j a^j{}_i f_j$$

zadaje odwzorowanie liniowe $F: V \to W$ takie, że $[F]^f{}_e = A$. ■

Z trzeciego punktu tego stwierdzenia dostajemy

$$\dim \mathsf{L}(V, W) = \dim \mathbf{M}^m{}_n(\mathbb{K}) = m \cdot n = (\dim V) \cdot (\dim W).$$

STWIERDZENIE 6.7. *Jeżeli e jest bazą w V, f bazą w W, g bazą w U i jeśli $F \in \mathsf{L}(V, W)$, $G \in \mathsf{L}(W, U)$, to $[G \circ F]^g{}_f = [G]^g{}_f [F]^f{}_e$.*

DOWÓD: Mamy dla każdego wektora $v \in V$

$$\begin{aligned}[G \circ F(v)]^g = [G(F(v))]^g &= [G]^g{}_f [F(v)]^f \\ &= [G]^g{}_f ([F]^f{}_e [v]^e) = ([G]^g{}_f [F]^f{}_e)[v]^e.\end{aligned}$$

■

Wnioski:

(1) Ponieważ składanie odwzorowań jest łączne, więc również mnożenie macierzy jest łączne.
(2) Jeżeli $F \in \mathsf{L}(V, W)$ jest izomorfizmem, to $[F^{-1}]^e{}_f = [F]^f{}_e{}^{-1}$.
Istotnie,
$$I = [Id]^e{}_e = [F^{-1}F]^e{}_e = [F^{-1}]^e{}_f [F]^f{}_e.$$
(3) Ponieważ $(F^{-1})^{-1} = F$, więc również dla macierzy zachodzi $(A^{-1})^{-1} = A$.
(4) Ponieważ dla odwzorowań $(F \circ G)^{-1} = G^{-1} \circ F^{-1}$, więc i dla macierzy mamy podobnie: $(AB)^{-1} = B^{-1}A^{-1}$.

Spostrzeżenie: $\operatorname{rz} [F]^f{}_e = \dim \operatorname{im} F$

STWIERDZENIE 6.8. *Zachodzi nierówność* $\operatorname{rz} BA \leqslant \min(\operatorname{rz} A, \operatorname{rz} B)$.

DOWÓD: Niech $A = [F]^f{}_e$, $B = [G]^g{}_f$. Wówczas ze Stwierdzenia 5.5

$$\begin{aligned}\operatorname{rz} BA &= \dim \operatorname{im} G \circ F \leqslant \dim \operatorname{im} G = \operatorname{rz} B \\ \operatorname{rz} BA &= \dim \operatorname{im} G \circ F \leqslant \dim \operatorname{im} F = \operatorname{rz} A.\end{aligned}$$

■

6.2.1. Zamiana bazy.

Niech e, e' będą bazami w V, f, f' bazami w W, i niech $F \in \mathsf{L}(V, W)$. Z poprzednich stwierdzeń

$$\begin{aligned} [F]^f{}_e &= [F \circ Id_V]^f{}_e = [F]^f{}_{e'}[Id_V]^{e'}{}_e \\ &= [Id_W]^f{}_{f'}[F]^{f'}{}_{e'}[Id_V]^{e'}{}_e. \end{aligned} \tag{6.2}$$

Przez Id_W, Id_V oznaczyliśmy odwzorowania tożsamościowe w W i V.

Podobnie

$$[v]^{e'} = [Id_V v]^{e'} = [Id_V v]^{e'}{}_e[v]^e.$$

6.3. Równania liniowe.

Niech $F\colon V \to W$ będzie odwzorowaniem liniowym i niech $b \in W$. Jeżeli e, f są bazami odpowiednio przestrzeni V, W, to równanie liniowe $Fx = b$ możemy zapisać równoważnie:

$$[F]^f{}_e[x]^e = [b]^f.$$

Abstrahując od odwzorowania, mamy równanie macierzowe $Ax = b$, gdzie szukamy kolumny $x \in \mathbf{M}^n{}_1(\mathbb{K})$, przy zadanych $A \in \mathbf{M}^m{}_n(\mathbb{K}), b \in \mathbf{M}^m{}_1(\mathbb{K})$.

Przetłumaczmy na język macierzy uwagi na temat równań wypowiadane wcześniej.

(1) Aby istniało rozwiązanie potrzeba i wystarcza, by przestrzenie rozpięte na kolumnach macierzy $A = [\bar{a}_1, \ldots, \bar{a}_n]$ i $[A, b] = [\bar{a}_1, \ldots, \bar{a}_n, b]$ były równe. Do tego potrzeba i wystarcza, by ich wymiary były równe czyli, by $\operatorname{rz} A = \operatorname{rz}[A, b]$ (tw.Kroneckera-Capelliego).
(2) Jeśli $m = n$, to równanie $Ax = b$ ma dla każdego b dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy istnieje A^{-1}. Wówczas $x = A^{-1}b$.
(3) Dodając do równania kombinację liniową pozostałych dostajemy układ równoważny, tzn., mający te same rozwiązania. Operacja ta odpowiada przejściu do innej bazy w przestrzeni W. Można zmieniać bazę również w przestrzeni V, ale ze względów praktycznych tego się nie robi.

Przykład: Rozwiążmy układ równań

$$\begin{cases} 5x_1 + 3x_2 + 5x_3 + 12x_4 = 10 \\ 2x_1 + 2x_2 + 3x_3 + 5x_4 = 4 \\ x_1 + 7x_2 + 9x_3 + 4x_4 = 2 \end{cases}$$

Szukamy możliwie prostego układu równoważnego. Macierz układu A jest równa

$$A = \begin{bmatrix} 5 & 3 & 5 & 12 \\ 2 & 2 & 3 & 5 \\ 1 & 7 & 9 & 4 \end{bmatrix}$$

Przez $\sim$ oznaczę, że macierze dają układy równoważne. Mamy więc

$$\begin{bmatrix} 5 & 3 & 5 & 12 & 10 \\ 2 & 2 & 3 & 5 & 4 \\ 1 & 7 & 9 & 4 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 7 & 9 & 4 & 2 \\ 0 & -12 & -15 & -3 & 0 \\ 0 & -32 & -40 & -8 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 7 & 9 & 4 & 2 \\ 0 & 4 & 5 & 1 & 0 \\ 0 & 4 & 5 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 3 & 2 \\ 0 & 4 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 4 & 0 & -1 & 9 & 8 \\ 0 & 4 & 5 & 1 & 0 \end{bmatrix}$$

Otrzymaliśmy

$$\begin{aligned} x_1 &= \frac{1}{4}(8 + x_3 - 9x_4) \\ x_2 &= \frac{1}{4}(-5x_3 - x_4). \end{aligned}$$

Stąd

$$x = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 4 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ -5 \\ 4 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -9 \\ -1 \\ 0 \\ 4 \end{bmatrix}.$$

Rozdział 7. Wyznaczniki

7.1. Definicja i istnienie.

W tej części korzystać będziemy intensywnie z kanonicznego izomorfizmu między $\mathbf{M}^m{}_n(\mathbb{K})$ i $\mathbf{M}^m{}_1(\mathbb{K}) \times \cdots \times \mathbf{M}^m{}_1(\mathbb{K})$ (n razy).

DEFINICJA 7.1. *Odwzorowanie* $\det \colon \mathbf{M}^n{}_n(\mathbb{K}) \to \mathbb{K}$ *nazywamy* wyznacznikiem macierzy, *jeżeli spełnione są następujące warunki:*

(1) $\det[\bar{a}_1, \dots, \alpha\bar{a}_i + \beta\bar{b}, \dots, \bar{a}_n] =$
$= \quad \alpha \det[\bar{a}_1, \dots, \bar{a}_i, \dots, \bar{a}_n] + \beta \det[\bar{a}_1, \dots, \bar{b}, \dots, \bar{a}_n]$
dla $i = 1, \dots, n$,
(2) $\det[\bar{a}_1, \dots, \bar{a}_i, \dots, \bar{a}_j, \dots, \bar{a}_n] = -\det[\bar{a}_1, \dots, \bar{a}_j, \dots, \bar{a}_i, \dots, \bar{a}_n]$
dla każdej pary $i \neq j$,
(3) $\det I = 1$, *gdzie*

$$I = [\delta^i{}_j], \quad \delta^i{}_j = \begin{cases} 0 & i \neq j \\ 1 & i = 1 \end{cases}.$$

Liczbę $\det A$ nazywamy wyznacznikiem macierzy A.

STWIERDZENIE 7.2. *Jeżeli funkcja* $\det$ *jest wyznacznikiem, to*

(1) $\det 0 = 0$,
(2) *jeżeli dla pewnych* $i \neq j$ $\bar{a}_i = \bar{a}_j$, *to* $\det A = 0$,
(3) $\det[\bar{a}_1, \dots, \bar{b}_i, \dots, \bar{a}_n] = \det A$, *jeżeli* $\bar{b}_i = \bar{a}_i + \lambda^1\bar{a}_1 + \cdots + \lambda^{i-1}\bar{a}_{i-1} + \lambda^{i+1}\bar{a}_{i+1} + \cdots + \lambda^n\bar{a}_n$.

DOWÓD: Oczywiste (punkty (1) i (3) definicji). ■

Uwaga! W dalszym ciągu będziemy, dla przejrzystości zapisu, używać symbolu a^i_j (zamiast $a^i{}_j$) dla oznaczenia elementu macierzowego.

STWIERDZENIE 7.3. *Jeżeli* $\det$ *jest wyznacznikiem, to*

$$\det AB = \det A \sum_{\sigma \in \mathsf{S}(n)} (\operatorname{sgn}\sigma)\, b_1^{\sigma(1)} \cdots b_n^{\sigma(n)}.$$

DOWÓD: Oznaczmy $\bar{e}_j = \left[\delta^i_j\right]$ i $\bar{a}^i\bar{b}_j = \sum_k a^i_k b^k_j$. Zatem $\bar{b}_i = \sum_j \bar{b}^j_i \bar{e}_j$ i

$$AB = \begin{bmatrix} \bar{a}^1\bar{b}_1 & \bar{a}^1\bar{b}_2 & \cdots & \bar{a}^1\bar{b}_n \\ \vdots & \vdots & \vdots & \vdots \\ \bar{a}^n\bar{b}_1 & \bar{a}^n\bar{b}_2 & \cdots & \bar{a}^n\bar{b}_n \end{bmatrix},$$

a stąd

$$\det AB = \sum_j b_1^j \det \begin{bmatrix} \bar{a}^1\bar{e}_j & \bar{a}^1\bar{b}_2 & \cdots \\ \vdots & \vdots & \vdots \\ \bar{a}^n\bar{e}_j & \bar{a}^n\bar{b}_n & \cdots \end{bmatrix}$$

$$= \sum_{j,l} b_1^j b_2^l \det \begin{bmatrix} \bar{a}^1\bar{e}_j & \bar{a}^1\bar{e}_l & \cdots \\ \vdots & \vdots & \vdots \\ \bar{a}^n\bar{e}_j & \bar{a}^n\bar{e}_l & \cdots \end{bmatrix}$$

$$= \sum_{\sigma\in \mathsf{S}(2)} b_1^{\sigma(1)} b_2^{\sigma(2)} \det \begin{bmatrix} \bar{a}^1\bar{e}_{\sigma(1)} & \bar{a}^1\bar{e}_{\sigma(2)} & \cdots \\ \vdots & \vdots & \vdots \\ \bar{a}^n\bar{e}_{\sigma(1)} & \bar{a}^n\bar{e}_{\sigma(2)} & \cdots \end{bmatrix} = \cdots$$

$$= \sum_{\sigma\in \mathsf{S}(n)} b_1^{\sigma(1)} b_2^{\sigma(2)} \cdots b_n^{\sigma(n)} \det \begin{bmatrix} \bar{a}^1\bar{e}_{\sigma(1)} & \bar{a}^1\bar{e}_{\sigma(2)} & \cdots \\ \vdots & \vdots & \vdots \\ \bar{a}^n\bar{e}_{\sigma(1)} & \bar{a}^n\bar{e}_{\sigma(2)} & \cdots \end{bmatrix}$$

$$= \sum_{\sigma\in \mathsf{S}(n)} b_1^{\sigma(1)} b_2^{\sigma(2)} \cdots b_n^{\sigma(n)} \det[\bar{a}_{\sigma(1)}, \bar{a}_{\sigma(2)}, \dots, \bar{a}_{\sigma(n)}]$$

$$= \det A \sum_{\sigma\in \mathsf{S}(n)} \operatorname{sgn}\sigma b_1^{\sigma(1)} b_2^{\sigma(2)} \cdots b_n^{\sigma(n)}$$

■

TWIERDZENIE 7.4. *Istnieje dokładnie jeden wyznacznik* $\det : \mathbf{M}^n{}_n(\mathbb{K}) \to \mathbb{K}$. *Dany jest on wzorem*

$$\det A = \sum_{\sigma\in \mathsf{S}(n)} \operatorname{sgn}\sigma(a_1^{\sigma(1)} \dots a_n^{\sigma(n)}).$$

DOWÓD: Z poprzedniego stwierdzenia wynika, że jeżeli wyznacznik istnieje, to

$$\det A = \det(IA) = \det I \sum_{\sigma\in \mathsf{S}(n)} \operatorname{sgn}\sigma(a_1^{\sigma(1)} \dots a_n^{\sigma(n)})$$
$$= \sum_{\sigma\in \mathsf{S}(n)} \operatorname{sgn}\sigma(a_1^{\sigma(1)} \dots a_n^{\sigma(n)}).$$

Wystarczy więc sprawdzić, że powyższy wzór określa wyznacznik, tzn., spełnia trzy warunki z definicji wyznacznika:

(1) Oczywiste.

(2) Niech $\sigma' = \sigma \circ (i,j)$. Mamy

$$\det[\bar{a}_1, \dots, \bar{a}_j, \dots, \bar{a}_i, \dots, \bar{a}_n] = \sum_{\sigma \in \mathsf{S}(n)} \operatorname{sgn} \sigma (a_1^{\sigma'(1)} \dots a_n^{\sigma'(n)})$$

$$= \sum_{\sigma' \in \mathsf{S}(n)} \operatorname{sgn}(\sigma' \circ (i,j))(a_1^{\sigma'(1)} \dots a_n^{\sigma'(n)}) = -\det[\bar{a}_1, \dots, \bar{a}_i, \dots, \bar{a}_j, \dots, \bar{a}_n].$$

(3) Oczywiste.

■

Wnioski:

(1) Z powyższego twierdzenia i z poprzedniego stwierdzenia wynika Twierdzenie Cauchy'ego:

$$\det AB = \det A \det B,$$

(2) $\det A = \det A^{\mathsf{T}}$. Istotnie,

$$\det A = \sum_{\sigma} \operatorname{sgn} \sigma (a_1^{\sigma(1)}, \dots, a_n^{\sigma(n)}) = \sum_{\sigma} \operatorname{sgn} \sigma (a^1_{\sigma^{-1}(1)}, \dots, a^n_{\sigma^{-1}(n)})$$

Ale $\operatorname{sgn} \sigma = \operatorname{sgn} \sigma^{-1}$, więc

$$\det A = \sum_{\sigma} \operatorname{sgn} \sigma (a_1^{\sigma(1)}, \dots, a_n^{\sigma(n)}) = \sum_{\sigma} \operatorname{sgn} \sigma (a_1^{\mathsf{T}\sigma(1)}, \dots, a_n^{\mathsf{T}\sigma(n)}) = \det A^{\mathsf{T}}.$$

STWIERDZENIE 7.5. *Niech $A \in \mathbf{M}^n{}_n(\mathbb{K})$.* $\det A = 0$ *wtedy i tylko wtedy, gdy* $\operatorname{rz} A < n$

Dowód: Jeżeli $\operatorname{rz} A < n$, to jedna z kolumn jest liniową kombinacją pozostałych i $\det A = 0$ na podstawie Stwierdzenia 7.2.

Odwrotnie, załóżmy, że kolumny są liniowo niezależne, tzn. tworzą bazę w $\mathbf{M}^n{}_1$. Stąd dla $\bar{e}_k = [\delta^i_k]$ mamy $\bar{e}_k = \sum_{j=1}^n b^j_k \bar{a}_j$ i, w efekcie, $I = AB$, gdzie $B = [b^j_k]$. Stąd $1 = \det(AB) = \det A \det B$ i $\det A \neq 0$. ■

Wniosek: A^{-1} istnieje wtedy i tylko wtedy, gdy $\det A \neq 0$. Ponadto

$$\det A^{-1} = (\det A)^{-1}.$$

7.2. Dopełnienie algebraiczne. Rozwinięcie Laplace'a.

DEFINICJA 7.6. Dopełnieniem algebraicznym *elementu a^i_j macierzy A nazywamy wyznacznik macierzy otrzymanej przez skreślenie i-tego wiersza i j-tej kolumny, pomnożony przez $(-1)^{j+i}$. Oznaczamy je A^j_i*

TWIERDZENIE 7.7 (ROZWINIĘCIE LAPLACE'A).

$$\det A = \sum_{i=1}^{n} a_i^k A_k^i = \sum_{i=1}^{n} a_k^i A_i^k \tag{7.1}$$

DOWÓD: Sprawdzimy, że odwzorowanie $F\colon \mathbf{M}^m{}_n(\mathbb{K}) \to \mathbb{K}$ dane wzorem

$$A \mapsto F(A) = \sum_{i=1}^{n} a_i^1 A_1^i$$

jest wyznacznikiem:

(1) Niech $B = [\bar{a}_1, \dots, \bar{a}_i + \bar{a}'_i, \dots, \bar{a}_n]$ i $A' = [\bar{a}_i, \dots, \bar{a}'_i, \dots, \bar{a}_n]$. Mamy

$$F(B) = a_1^1(A_1^1 + {A'}_1^1) + \dots + (a_i^1 + {a'}_i^1)A_1^i + \dots + a_n^1(A_1^n + {A'}_1^n) = F(A) + F(A').$$

(2) Jeżeli $\bar{a}_i = \bar{a}_j$, to $A_1^j = -A_1^i$ i $A_1^k = 0$ dla $k \neq i, j$.
Stąd $F(A) = a_j^i A_1^i + a_j^1 A_1^j = 0$.

(3) $F(I) = 1$.

Dowodzi to, że $\det A = \sum_{i=1}^{n} a_i^1 A_1^i$. Pozostałe przypadki dowodzi się analogicznie. ■

7.3. Przykłady i zastosowania.

Przykłady:

(1) Schemat Sarrusa obliczania wyznaczników 3×3.

$$\begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = \begin{array}{|ccc|cc} & & & - \quad - & \quad - \\ a & b & c & a & b \\ a_1 & b_1 & c_1 & a_1 & b_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ & & & + \quad + & \quad + \end{array} = \tag{7.2}$$

$= ab_1c_2 + bc_1a_2 + ca_1b_2 - a_2b_1c - b_2c_1a - c_2a_1b$

(2)

$$\begin{vmatrix} 1 & 3 & 4 & 2 \\ 3 & 0 & 2 & -1 \\ 2 & 1 & 0 & 3 \\ 0 & 0 & 5 & 2 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 0 & 5 & 8 & 1 \\ 3 & 0 & 2 & -1 \\ 2 & 1 & 0 & 3 \\ 0 & 0 & 5 & 2 \end{vmatrix}$$

$$= \frac{1}{2}\left(-3\begin{vmatrix} 5 & 8 & 1 \\ 1 & 0 & 3 \\ 0 & 5 & 2 \end{vmatrix} + 2\begin{vmatrix} 5 & 8 & 1 \\ 0 & 2 & -1 \\ 0 & 5 & 2 \end{vmatrix}\right)$$

$$= \frac{1}{2}(-3(5-75-16) + 2(20+25)) = (3\cdot 43 + 45) = 174.$$

(3) Niech A i B będą macierzami kwadratowymi.

$$\det\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \det\left(\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & B \end{bmatrix}\right) = \det A \det B,$$

bo, korzystając z rozwinięcia Laplace'a względem ostatnich kolumn (wierszy), mamy

$$\det\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = \det A.$$

(4) Wyznacznik Vandermonde'a. Niech $x_1, \dots, x_n$ będą liczbami. Oznaczmy

$$V_n = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix}.$$

V_n traktowane jako funkcja od x_n jest wielomianem stopnia $n-1$. Oczywiste, że pierwiastkami tego wielomianu są $x_1, \dots, x_{n-1}$, i że współczynnik przy najwyższej potędze jest równy V_{n-1}. Mamy więc wzór rekurencyjny $V_n = (x_n - x_1)\cdots(x_n - x_{n-1})V_{n-1}$. Stąd

$$V_n = \prod_{k>l}(x_k - x_l).$$

Pewne zastosowania wyznaczników:

(A) Wzory Cramera. Rozpatrzmy równanie $Ax = b$, gdzie $A \in \mathbf{M}^n{}_n(\mathbb{K})$ i $\det A \neq 0$. Pisząc

$$x = \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix},$$

dostajemy to równanie w postaci $\bar{a}_1 x^1 + \cdots + \bar{a}_n x^n = b$ lub, równoważnie, $(\bar{a}_1 x^1 - b) + \bar{a}_2 x^2 + \cdots + \bar{a}_n x^n = 0$, czyli

$$\det[\bar{a}_1 x^1 - b, \bar{a}_2, \dots, \bar{a}_n] = 0.$$

Stąd

$$x^1 \det[\bar{a}_1, \dots, \bar{a}_n] = \det[b, \bar{a}_2, \dots, \bar{a}_n],$$

czyli

$$x^1 = \frac{\det[b, \bar{a}_2, \dots, \bar{a}_n]}{\det[\bar{a}_1, \dots, \bar{a}_n]}$$

i, ogólnie,

$$x^i = \frac{\det[\bar{a}_1, \dots, b, \dots, \bar{a}_n]}{\det[\bar{a}_1, \dots, \bar{a}_n]} \tag{7.3}$$

Są to *wzory Cramera.*

(B) Jeżeli $A \in \mathbf{M}^n{}_n(\mathbb{K})$ i $\det A \neq 0$ to, jak wiemy, istnieje A^{-1}. Pokażemy, że elementy macierzy odwrotnej zadane są wzorem

$$b^i_j = A^i_j (\det A)^{-1},$$

gdzie A^i_j jest dopełnieniem algebraicznym elementu a^j_i macierzy A. Istotnie, niech B będzie macierzą o elementach macierzowych $b^i_j = A^i_j(\det A)^{-1}$. Mamy z rozwinięcia Laplace'a (7.1)

$$\sum_k b^i_k a^k_j = \frac{1}{\det A} \sum A^i_k a^k_j = \frac{1}{\det A} \det[\bar{a}_1, \dots, \bar{a}_{i-1}, \bar{a}_j, \bar{a}_{i+1} \dots, \bar{a}_n] = \delta^i_j.$$

Zatem $BA = I$, czyli $B = A^{-1}$.

(C) Jeżeli A^{D} jest macierzą dopełnień algebraicznych macierzy A, to z poprzedniego punktu mamy

$$AA^{\mathsf{D}} = A^{\mathsf{D}} A = (\det A) I. \tag{7.4}$$

Rozdział 8. Struktura endomorfizmu liniowego

8.1. Operatory rzutowe.

Niech V będzie przestrzenią wektorową nad $\mathbb{K}$ i niech $V = V_1 \oplus V_2$. Każdy wektor $v \in V$ da się więc przedstawić jednoznacznie jako suma $v = v_1 + v_2$, gdzie $v_i \in V_i$. Odwzorowanie

$$P\colon V \to V\colon v \mapsto v_1$$

nazywamy *rzutem* na podprzestrzeń V_1 wzdłuż podprzestrzeni V_2. Zauważmy, że P jest odwzorowaniem liniowym i $P^2 = P$. Na odwrót, niech $F \in \mathsf{End}(V)$ i niech $F^2 = F$. Połóżmy $W_1 = \operatorname{im} F$ i $W_2 = \ker F$. Jeżeli $v \in W_1 \cap W_2$, to istnieje w takie, że $v = F(w)$. Stąd

$$v = F(w) = F(F(w)) = F(v) = \mathbf{0}.$$

Zatem $W_1 \cap W_2 = \{\mathbf{0}\}$. Mamy ponadto

$$\dim V = \dim \operatorname{im} F + \dim \ker F = \dim W_1 + \dim W_2$$

i

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

Ponieważ $\dim(W_1 \cap W_2) = 0$, mamy stąd $V = W_1 \oplus W_2$. Jeżeli $w = w_1 + w_2$ jest rozkładem wektora w, to $F(w_2) = 0$ i $w_1 = F(v)$ dla pewnego $v \in V$. Stąd

$$F(w) = F(w_1) = F(F(v)) = F(v) = w_1.$$

F jest więc rzutem na $\operatorname{im} F$ wzdłuż $\ker F$. Mamy zatem

STWIERDZENIE 8.1. *Endomorfizm (operator) F jest rzutem (jest rzutowy) wtedy i tylko wtedy, gdy $F = F^2$.*

8.2. Wektory i wartości własne.

Niech V będzie przestrzenią wektorową nad $\mathbb{K}$, $F \in \mathsf{L}(V, V)$ i niech f, e będą bazami w V. Mamy

$$[F]^e{}_e = [Id]^e{}_f [F]^f{}_f [Id]^f{}_e\ ,$$

ale $[Id]^e{}_f = ([Id]^f{}_e)^{-1}$, czyli $\det [Id]^e{}_f = (\det([Id]^f{}_e))^{-1}$ i, w konsekwencji,

$$\det([F]^e{}_e) = \det([F]^f{}_f)\ .$$

DEFINICJA 8.2. *Wyznacznik*

$$\det([F]^e{}_e)\ .$$

macierzy endomorfizmu F nazywamy wyznacznikiem endomorfizmu F.

Wyznacznik endomorfizmu oznaczamy $\det F$. Jak wiadomo, F jest izomorfizmem (tzn. istnieje F^{-1}) wtedy i tylko wtedy, gdy macierz odwzorowania $[F]^e{}_e$ jest odwracalna. Z kolei, macierz jest odracalna wtedy i tylko wtedy, gdy jej wyznacznik jest różny od zera. Zatem F jest izomorfizmem wtedy i tylko wtedy, gdy $\det F \neq 0$.

DEFINICJA 8.3. *Wielomian w zmiennej λ określony wzorem*

$$w(\lambda) = \det(F - \lambda Id_V) \tag{8.1}$$

nazywamy wielomianem charakterystycznym *endomorfizmu $F \in \mathsf{End}(V)$ i oznaczamy ω_F.*

Interesować nas będą pierwiastki wielomianu charakterystycznego.

DEFINICJA 8.4. Wartością własną *endomorfizmu (operatora) F nazywamy pierwiastek jego wielomianu charakterystycznego.*

Jeżeli λ jest wartością własną, to $\det(F - \lambda Id_V) = 0$, czyli jądro endomorfizmu $F - \lambda Id_V$ nie jest trywialne, tzn., istnieje $v \neq 0$ takie, że $Fv = \lambda v$. Na odwrót, jeżeli dla pewnego λ istnieje niezerowy v taki, że $Fv = \lambda v$, to $F - \lambda Id_V$ ma nietrywialne jądro, wiec nie jest izomorfizmem. Stąd $\det(F - \lambda Id_V) = 0$, czyli λ jest wartością własną.

DEFINICJA 8.5. *Niech λ będzie wartością własną F. Wektor $v \neq \mathbf{0}$ taki, że $Fv = \lambda v$ nazywamy* wektorem własnym *operatora (endomorfizmu) F odpowiadającym wartości własnej λ.*

Przykłady.

(a) Niech $V = \mathbb{R}^2$ i niech F będzie odbiciem względem osi x: $F((x, y)) = (x, -y)$. Warunek $F((x, y)) = \lambda(x, y)$ może być spełniony dla $\lambda = 1$ lub $\lambda = -1$. Są to wartości własne. Wektorami własnymi wartości własnej $\lambda = 1$ są wektory postaci $(x, 0)$. Wektorami własnymi wartości własnej $\lambda = -1$ są wektory postaci $(0, y)$.

(b) Niech $V = \mathbb{R}^2$ i niech F będzie obrotem wokół punktu $(0, 0)$ o kąt $\pi/2$. F nie ma wartości i wektorów własnych ($\mathbb{K} = \mathbb{R}$!).

(c) Niech $V = \mathbb{C}^1$ i niech F będzie jak w poprzednim przykładzie. Wartością własną jest teraz liczba $\lambda = \iota$, a wektorami własnymi wszystkie wektory niezerowe ($\mathbb{K} = \mathbb{C}$!).

DEFINICJA 8.6. *Podprzestrzeń wektorową W przestrzeni V nazywamy podprzestrzenią* niezmienniczą *operatora $F \in \mathsf{End}(V)$, jeżeli $FW \subset W$.*

Przykład: Podprzestrzeń wektorów własnych ustalonej wartości własnej, uzupełnionych zerem, jest podprzestrzenią niezmienniczą.

Endomorfizmy przestrzeni wektorowej (jak również macierze kwadratowe) można dodawać, mnożyć i mnożyć przez liczbę. Jest zatem sens mówić o fukcjach wielomianowych zmiennej F (zmiennej macierzowej) dla wielomianów z $\mathbb{K}[\cdot]$

TWIERDZENIE 8.7 (CAYLEYA-HAMILTONA). *Niech* $F \in \mathsf{L}(V, V)$ *i niech* ω_F *będzie wielomianem charakterystycznym* F*; wówczas* $w(F) = 0$.

DOWÓD: Oznaczmy $A = [F]^e{}_e$. Mamy, zgodnie z (7.4),

$$(A - \lambda I)(A - \lambda I)^{\mathsf{D}} = \det(A - \lambda I) I \ . \qquad (*)$$

Z definicji dopełnienia algebraicznego wyrazy macierzy $(A - \lambda I)^{\mathsf{D}}$ są wielomianami zmiennej λ stopnia co najwyżej $n - 1$, czyli

$$(A - \lambda I)^{\mathsf{D}} = B_0 + \lambda B_1 + \cdots + \lambda^{n-1} B_{n-1} \ .$$

Jeżeli $\omega_F(\lambda) = b_0 + \lambda b_1 + \cdots + \lambda^n b_n$, to z równości $(*)$ wynika, że

$$\begin{array}{rrcl} AB_0 & & = & b_0 I \\ AB_1 - & B_0 & = & b_1 I \\ AB_2 - & B_1 & = & b_2 I \\ & \cdots & & \\ AB_{n-1} - & B_{n-2} & = & b_{n-1} I \\ - & B_{n-1} & = & b_n I \ . \end{array}$$

Mnożąc te równania lewostronnie przez stosowne potęgi A i dodając stronami dostajemy

$$b_0 I + b_1 A + \cdots + b_n A^n = 0.$$

Ponieważ $[\omega_F(F)]^e{}_e = \omega_F(A)$, mamy tezę. ■

8.3. Rozkład na podprzestrzenie pierwiastkowe.

Uwaga! W tym podrozdziale (aż do odwołania) $\mathbb{K} = \mathbb{C}$, czyli rozpatrujemy przestrzenie **nad ciałem liczb zespolonych**.

Niech $\lambda_1, \ldots, \lambda_r$ będą pierwiastkami ω_F, a $n_1, \ldots, n_r$ - ich krotnościami. Zatem

$$w(\lambda) = \prod_{k=1}^{r} (\lambda_k - \lambda)^{n_k}.$$

Niech teraz λ_i będzie pierwiastkiem w o krotności n_i. Zdefiniujemy wielomian ω_i :

$$\omega_i(\lambda) = \prod_{k \neq i} (\lambda_k - \lambda)^{n_k} = \frac{w(\lambda)}{(\lambda_i - \lambda)^{n_i}} \ .$$

Spostrzeżenie: $\mathsf{NWD}(\omega_1, \ldots, \omega_r)$ jest, jak łatwo widać, równy 1. Wynika stąd (wzór (2.3)), że istnieją wielomiany u_i takie, że

$$1 = \sum_{i=1}^{r} u_i \omega_i .$$

Zdefiniujmy endomorfizmy P_i przestrzeni V wzorami:

$$P_i = u_i(F)\omega_i(F) \ .$$

Mają one następujące własności:

(1) $\sum_i u_i\omega_i) = 1$, więc $\sum_i P_i = Id_V$

(2) $P_iP_j = 0$ dla $i \neq j$ bo $u_i\omega_i u_j\omega_j = (coś)\omega_F$ jeśli $i \neq j$, a $\omega(F) = 0$ (Twierdzenie Cayleya-Hamiltona).

(3) $P_iP_i = P_i$ bo $Id_V = \sum_j P_j$ i, z poprzedniego, $P_i = \sum_j P_iP_j = P_i^2$. P_i jest więc operatorem rzutowym.

(4) Oznaczmy $V_i = \operatorname{im} P_i$. Pokażemy, że $V_i = \ker(F - \lambda_i Id_V)^{n_i}$.
Istotnie, jeżeli $x \in \ker(F - \lambda_i Id_V)^{n_i}$, to $P_j x = \mathbf{0}$ dla $j \neq i$ i, w konsekwencji, $x = Id_V \ x = \sum_k P_k x = P_i x \in V_i$.
Na odwrót, jeżeli $x \in \operatorname{im} P_i$, to istnieje v takie, że $x = P_i v$, czyli

$$(F - \lambda_i Id_V)^{n_i} x = (F - \lambda_i Id_V)^{n_i} P_i v = (-1)^{n_i} \omega_F(F) v = \mathbf{0} \ .$$

(5) V_i jest podprzestrzenią niezmienniczą dla F.
Istotnie, $FP_i = P_iF$, więc dla $x = P_i v$ mamy $Fx = F(P_i v) = P_i(Fv) \in V_i$.

(6) Jeżeli $i \neq j$, to $V_i \cap V_j = \{\mathbf{0}\}$ bo dla $x = P_i v_i = P_j v_j$ mamy z (2)

$$x = P_i v_i = P_i(P_i v_i) = P_i x = P_i P_j v_j = \mathbf{0} \ .$$

(7) $V = \oplus_i V_i$, bo $x = (\sum_i P_i)x = \sum_i P_i x$, a $V_i = \operatorname{im} P_i$. Z (2) wynika, że rozkład ten jest jednoznaczny: jeżeli $x = \sum_i y_i, \ y_i \in V_i$, to $P_i x = P_i y_i = y_i$.

(8) Oznaczamy $d_i = \dim V_i$. Niech $(e_{i,1}, \ldots, e_{i,d_i})$ będzie bazą V_i. Z poprzedniego $(e_{1,1}, \ldots, e_{r,d_r})$ jest bazą V. Oznaczmy ją e. Na mocy (4) macierz ${[F]^e}_e$ jest postaci

$$\begin{bmatrix} A_1 & & & & \\ & \ddots & & 0 & \\ & & A_i & & \\ & 0 & & \ddots & \\ & & & & A_r \end{bmatrix}$$

gdzie i-ta klatka A_i ma rozmiar $d_i \times d_i$. Stąd

$$\det(F - \lambda Id_V) = \prod_i \det(A_i - \lambda I) \ .$$

Zatem pierwiastek równania $\det(A_i - \lambda I) = 0$ jest wartością własną F. Niech będzie to λ_j. Istnieje więc $\mathbf{0} \neq v \in V_i$, że $Fv = \lambda_j v$. Zatem $v \in V_i \cap V_j$, więc z (6) wynika, że $j = i$. Zatem $\det(A_i - \lambda I) = (\lambda_i - \lambda)^{d_i}$ i, w konsekwencji,

$$\omega_F(\lambda) = \prod_i (\lambda_i - \lambda)^{n_i} = \det(F - \lambda Id_V) = \prod_i \det(A_i - \lambda) = \prod_i (\lambda_i - \lambda)^{d_i} \ .$$

Stąd $d_i = n_i$, tzn., $\dim V_i = n_i$.

Podprzestrzeń V_i nazywamy *podprzestrzenią pierwiastkową* wartości własnej λ_i. Powyższe możemy zebrać w następujące twierdzenie

TWIERDZENIE 8.8 (O ROZKŁADZIE NA PODPRZESTRZENIE PIERWIASTKOWE). *Niech $F \in \mathsf{L}(V, V)$. Wówczas $V = \oplus_i V_i$, gdzie $V_i = \ker(F - \lambda Id_V)^{n_i}$ są podprzestrzeniami niezmienniczymi dla F. Ponadto $\dim V_i = n_i$, gdzie n_i jest krotnością wartości własnej λ_i.*

Każda podprzestrzeń pierwiastkowa V_i zawiera wektor własny odpowiadający wartości własnej λ_i. Może się zdażyć, że wszystkie niezerowe wektory przestrzeni pierwiastkowych są wektorami własnymi i, w konsekwencji, istnieje baza przestrzeni V złożona z wektorów własnych. W takim przypadku operator nazywa się *diagonalizowalnym*. Nazwa bierze się stąd, że jeżeli e jest bazą wektorów własnych, to macierz $[F]^e{}_e$ jest diagonalna.

8.4. Funkcje od operatora.

Argumentami funkcji wielomianowych mogą być (patrz podrozdział 2.1.) dowolone obiekty, które można dodawać, mnożyć przez siebie i przez współczynniki wielomianu. W szczególności, jako argumenty funkcji wielomianowej o współczynnikach w $\mathbb{K}$ mogą służyć endomorfizmy (operatory) przestrzeni wektorowej nad $\mathbb{K}$. Działaniem mnożenia jest tu składanie operatorów. Jeżeli więc mamy wielomian

$$p(\lambda) = a_0 + \cdots + a_k x^k,$$

to odpowiednia funkcja wielomianowa o argumencie operatorowym wygląda tak:

$$p(F) = a_0 + a_1 F + \cdots + a_k F^k,$$

gdzie $F \in \mathsf{End}(V)$ i $F^k = F \circ F \circ \cdots \circ F$ (k razy).

W sposób oczywisty możemy rozszerzyć powyższe postępowanie do funkcji zadawanych szeregami potęgowymi. Jeżeli

$$f(\lambda) = \sum_{0}^{\infty} a_i(\lambda - \lambda_0)^i,$$

to

$$f(F) = \sum_{0}^{\infty} a_i(F - \lambda_0 Id_V)^i.$$

Powstaje tu, oczywiście, problem zbieżności szeregu, omawiany w kursie analizy.

8.4.1. Praktyczne sposoby. Praktyczne sposoby znajdowanie wartości funkcji operatorowej są różne dla przypadku wielomianu i dla przypadku szeregu. Zacznijmy od wielomianu.

Niech q będzie wielomianem i niech $F \in \mathsf{End}(V)$. Załóżmy, że znamy wielomian ω dla którego $\omega(F) = \mathbf{0}$. Wielomian o tej własności zawsze istnieje. Jest nim, na przykład, wielomian charakterystyczny ω_F, stopnia równego $\dim V$. Dla operatora diagonalizowalnego możemy wziąć wielomian $(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_r)$, gdzie $\lambda_1, \ldots, \lambda_r$ są wszystkimi wartościami własnymi F.

Ze Stwierdzenia 2.1 o dzieleniu wielomianów istnieją wielomiany q, r takie, że $p = q\omega + r$ i $\deg r < \deg \omega$. Stąd

$$f(F) = g(F) \circ \omega(F) + r(F) = r(F).$$

Zamiast liczyć $f(F)$ wystarcza obliczyć $r(F)$.

Załóżmy teraz, że $\mathbb{K} = \mathbb{C}$.

Niech funkcja $f\colon \mathbb{C} \to \mathbb{C}$ będzie zadana szeregiem w otoczeniu każdej wartości własnej λ_j operatora F, to znaczy,

$$f(\lambda) = \sum_{i=0}^{\infty} a_{j,i}\lambda^i$$

dla λ będącego w otoczeniu λ_j. (Z analizy wiadomo, że $a_{j,i} = \dfrac{1}{i!}\dfrac{\mathrm{d}^i f}{\mathrm{d}\lambda^i}(\lambda_j)$.) Oznaczmy przez V_{λ_j} podprzestrzeń pierwiastkową odpowiadającą wartości własnej λ_j, a przez P_j rzut na nią. Ponieważ V_j jest niezmiennicza dla operatora F, to

$$(F \circ P_j)^i = F^i \circ P_j$$

i, w konsekwencji,

$$f(F) \circ P_j = f(F \circ P_j) = \sum_{i=0}^{\infty} a_{j,i}(F \circ P_j - \lambda_j P_j)^i.$$

Ale $(F \circ P_j - \lambda_j P_j)^i = \mathbf{0}$ dla $i \geqslant n_j$, więc dostajemy

$$f(F) \circ P_j = \sum_{i=0}^{\infty} a_{j,i}(F \circ P_j - \lambda_j P_j)^i = \sum_{i=0}^{n_j-1} a_{j,i}(F - \lambda_j)^i \circ P_j.$$

Ponieważ zaś $\sum_{j=1}^{r} P_j = Id_V$, to dostajemy ostatecznie

$$f(F) = \sum_{j=1}^{r} \sum_{i=0}^{\infty} a_{j,i}(F - \lambda_j)^i \circ P_j. \tag{8.2}$$

Jest jeszcze jeden sposób znajdowania funkcji od operatora – przez znalezienie wygodnej reprezentacji macierzowej operatora, o czym poniżej.

8.4.2. Reprezentacja macierzowa. Znajdowanie operatora jest równoważne znajdowaniu jego macierzy w jakiejś bazie. Spróbujmy więc znaleźć $[f(F)]^e{}_e$ zakładając znajomość macierzy $[F]^e{}_e$.

Zauważmy najpierw, że, podobnie jak w przypadku operatorów, jest sens mówić o wielomianach i szeregach zmiennej macierzowej (ale tylko w przypadku macierzy kwadratowej).

Ponieważ $[F \circ F]^e{}_e = [F]^e{}_e [F]^e{}_e$, to, dla funkcji f zadanej szeregiem potęgowym, mamy

$$[f(F)]^e{}_e = f([F]^e{}_e).$$

Problem nasz sprowadza się więc do znalezienia wartości funkcji f dla pewnej macierzy A. Jeżeli macierz A jest diagonalna, $A = [a^i{}_j]$, gdzie $a^i{}_j = \lambda_j \delta^i{}_j$, to

$$A^k = [(\lambda_j)^k \delta^i{}_j]$$

i, w konsekwencji,

$$f(A) = [f(\lambda_j)\delta^i{}_j]. \tag{8.3}$$

Formułę (8.3) można przyjąć jako definicję $f(A)$ dla macierzy diagonalnej (więc i $f(F)$ dla operatora diagonalizowalnego) i dla dowolnej funkcji f (nie koniecznie zadanej szeregiem).

Dla dowolnej macierzy A procedura znajdowania $f(A)$ jest analogiczna do opisanej w części 8.4.1., jeżeli uwzględnimy, że każda macierz kwadratowa $n \times n$ odpowiada endomorfizmowi $V = \mathbb{K}^n$ lub $V = \mathbf{M}^n{}_1(\mathbb{K})$ (zapisanemu jako macierz w bazie kanonicznej). W tym kontekście mówić będziemy mówić o wartościach własnych i wektorach własnych macierzy.

8.5. Przykłady.

Przykład 1.

Obliczmy $f(A) = A^{50}$, gdzie

$$A = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 3 & 5 \\ -1 & -2 & -4 \end{bmatrix}.$$

Wielomian charakterystyczny ω_A jest równy

$$\omega_A(\lambda) = \det(A - \lambda I) = (1-\lambda)(\lambda+1)^2.$$

Rozkładamy teraz wielomian $p(\lambda) = \lambda^{50}$:

$$p(\lambda) = q(\lambda)\omega_A(\lambda) + r(\lambda), \tag{8.4}$$

gdzie $\deg r < 3$. Resztę r przewidujemy zatem w postaci $r(\lambda) = a\lambda^2 + b\lambda + c$. Współczynniki tego trójmianu znajdziemy porównując wartości lewej i prawej strony (8.4), oraz ich pochodnych do rzędu $n_i - 1$, dla pierwiastków wielomianu charakterystycznego ω_A:

$$\left\{ \begin{array}{c} p(1) = r(1) \\ p(-1) = r(-1) \\ p'(-1) = r'(-1) \end{array} \right\} \Longrightarrow \left\{ \begin{array}{r} 1 = a + b + c \\ 1 = a - b + c \\ -50 = -2a + b \end{array} \right\}.$$

Stąd $a = 25,\ b = 0,\ c = -24$ i $r(\lambda) = 25\lambda^2 - 24$. Ostatecznie

$$A^{50} = r(A) = 25 \begin{bmatrix} -1 & 0 & -2 \\ -2 & 1 & -2 \\ 2 & 0 & 3 \end{bmatrix} - 24 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -49 & 0 & -50 \\ -50 & 1 & -50 \\ 50 & 0 & 51 \end{bmatrix}.$$

Przykład 2.

Obliczmy $\exp(A)$, A jak w poprzednim przykładzie.
Mamy już znane:

$$\begin{array}{lll} \lambda_1 = 1 & n_1 = 1 & \dim V_1 = 1 \\ \lambda_2 = 2 & n_2 = 2 & \dim V_{-1} = 2. \end{array}$$

Rozwiązując równania $(A - I)x = \mathbf{0}$ i $(A + I)^2 x = \mathbf{0}$ dostajemy, że przestrzeń V_1 jest rozpięta na kolumnie $\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$, a V_{-1} jest rozpięta na kolumnach $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ i $\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$.
Rozkładamy teraz wektory bazy standardowej na sumę elementów z V_1 i V_{-1}.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in V_{-1}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}.$$

Stąd

$$P_1 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & -1 & -1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 1 & -1 & -1 \\ 0 & -1 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

i, jak w (8.2),

$$\exp(A) = e^1 P_1 + e^{-1}(I + (A+I))P_2 = \begin{bmatrix} 2e & e + e^{-1} & e \\ e^{-1} & 2e - e^{-1} & 2e - e^{-1} \\ -2e^{-1} & -e + e^{-1} & -e + e^{-1} \end{bmatrix}.$$

Przykład 3.
Niech (u_n) będzie ciągiem takim że, $u_n = u_{n-1} + u_{n-2}$, $u_1 = u_2 = 1$ (ciąg Fibonacciego). Szukamy u_n. Weźmy $V = \mathbb{C}^2$ i $F\colon V \to V$ zadane w bazie kanonicznej macierzą

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Zauważmy, że wyrazy ciągu Fibonacciego spełniają związek $F(u_{n-1}, u_{n-2}) = (u_n, u_{n-1})$. Stąd $(u_n, u_{n-1}) = F^{n-2}(u_2, u_1) = F^{n-2}(1,1)$. Z twierdzenia o rozkładzie na podprzestrzenie niezmiennicze $V = V_1 \oplus V_2$ i $(1,1) = v_1 + v_2$ gdzie $v_i \in V_i$. Stąd $F^{n-2}(1,1) = F^{n-2}v_1 + F^{n-2}v_2 = \lambda_1^{n-2}v_1 + \lambda_2^{n-2}v_2$. Liczymy wartości i wektory własne endomorfizmu F. Wyznacznik $\det(F - \lambda Id_V) = \lambda^2 - \lambda - 1$. Wartości własne $\lambda_{1,2} = (1 \pm \sqrt{5})/2$. Prosty rachunek daje, że

$$v_1 = \left(\left(\frac{1+\sqrt{5}}{2}\right)^2 \frac{1}{\sqrt{5}}, \left(\frac{1+\sqrt{5}}{2}\right) \frac{1}{\sqrt{5}} \right),$$

$$v_2 = -\left(\left(\frac{1-\sqrt{5}}{2}\right)^2 \frac{1}{\sqrt{5}}, \left(\frac{1-\sqrt{5}}{2}\right) \frac{1}{\sqrt{5}} \right)$$

i

$$u_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n \frac{1}{\sqrt{5}} - \left(\frac{1-\sqrt{5}}{2}\right)^n \frac{1}{\sqrt{5}} \right) .$$

8.6. Kompleksyfikacja.

Co zrobić, jeżeli V jest przestrzenią wektorową nad $\mathbb{R}$?

Niech V będzie przestrzenią wektorową nad $\mathbb{R}$. W iloczynie kartezjańskim $V \times V$ wprowadzamy strukturę przestrzeni wektorowej nad $\mathbb{C}$ kładąc $\iota(v,w) = (-w, v)$. Będziemy pisać $(v,w) = v + \iota w$. Tak wprowadzoną przestrzeń wektorową nad $\mathbb{C}$ oznaczać będziemy $V_{\mathbb{C}}$ i nazwiemy ją *kompleksyfikacją* przestrzeni V. Odwzorowanie liniowe $F \in \mathsf{L}(V, W)$ jednoznacznie przedłuża się do odwzorowania liniowego

$F_{\mathbb{C}}: V_{\mathbb{C}} \to W_{\mathbb{C}}$. Mamy $F_{\mathbb{C}}(v + \iota w) = F(v) + \iota F(w)$. Niech teraz F będzie endomorfizmem V i niech λ będzie wartością własną $F_{\mathbb{C}}$. Rozpatrzmy dwa przypadki:

(1) $\lambda \in \mathbb{R}$. Łatwo zauważyć, że w tym przypadku podprzestrzeń niezmiennicza jest postaci $V_{\lambda\mathbb{C}}$, tzn., jest kompleksyfikacją podprzestrzeni $V_\lambda = \{v \in V : (F - \lambda)^{n_\lambda} v = \mathbf{0}\}$, gdzie n_λ jest krotnością λ.

(2) $\lambda \notin \mathbb{R}$. W tym przypadku również $\overline{\lambda}$ jest wartością własną z tą samą krotnością co λ. Niech V_λ, $V_{\overline{\lambda}}$ będą odpowiednimi podprzestrzeniami niezmienniczymi.

STWIERDZENIE 8.9. $v + \iota w \in V_\lambda$ *wtedy i tylko wtedy, gdy* $v - \iota w \in V_{\overline{\lambda}}$.

DOWÓD: Zauważmy najpierw, że jeżeli $\mu F_{\mathbb{C}}(v + \iota w) = x + \iota y$, to $\bar{\mu} F_{\mathbb{C}}(v - \iota w) = x - \iota y$. Istotnie, mamy

$$\mu F_{\mathbb{C}}(v + \iota w) = (\operatorname{Re} \mu F(v) - \operatorname{Im} \mu F(w)) + \iota (\operatorname{Re} \mu F(w) + \operatorname{Im} \mu F(v)) .$$

Teza wynika teraz z faktu, że $V_\lambda = \ker(F_{\mathbb{C}} - \lambda Id_V)^{n_\lambda}$ i $V_{\overline{\lambda}} = \ker(F_{\mathbb{C}} - \overline{\lambda} Id_V)^{n_\lambda}$. ∎

Zajmijmy się podprzestrzenią $V_{\mathbb{C}} \supset W = V_\lambda + V_{\overline{\lambda}}$. Oznaczmy $V_{(\lambda,\overline{\lambda})}$ podprzestrzeń V określoną przez

$$V_{(\lambda,\overline{\lambda})} \times \{\mathbf{0}\} = (V_\lambda + V_{\overline{\lambda}}) \cap (V \times \{\mathbf{0}\}).$$

Pokażemy, że $W = V_{(\lambda,\overline{\lambda})\mathbb{C}}$. Istotnie, jeżeli $v + \iota w \in W$, to $v = v_1 + v_2$ i $w = w_1 + w_2$, gdzie $v_1 + \iota w_1 \in V_\lambda$ i $v_2 + \iota w_2 \in V_{\overline{\lambda}}$. Ze stwierdzenia wynika, że również $v - \iota w \in W$. Zatem $v \in V_{(\lambda,\overline{\lambda})}$ i, analogicznie, $w \in V_{(\lambda,\overline{\lambda})}$. Na odwrót, jeżeli $v, w \in V_{(\lambda,\overline{\lambda})}$, to $v + \iota w \in W$. Mamy więc żądaną równość.

Jest oczywistym, że $V_{(\lambda,\overline{\lambda})}$ jest podprzestrzenią niezmienniczą F.

Podsumowując: V rozkłada się na sumę prostą podprzestrzeni niezmienniczych V_λ (λ rzeczywiste) i $V_{(\lambda,\overline{\lambda})}$ (λ nierzeczywiste), przy czym $\dim_{\mathbb{R}} V_{(\lambda,\overline{\lambda})} = 2n_\lambda$.

Rozdział 9. Przestrzenie dualne. Odwzorowania sprzężone

9.1. Przestrzenie dualne.

DEFINICJA 9.1. *Niech V będzie przestrzenią wektorową nad $\mathbb{K}$. Przestrzeń $V^* = \mathsf{L}(V, \mathbb{K})$ nazywamy przestrzenią* sprzężoną (dualną) *do przestrzeni V. Elementy przestrzeni dualnej nazywamy formami liniowymi (funkcjonałami).*

Uwaga. $\dim V^* = \dim \mathsf{L}(V, K) = \dim V$. Ponieważ V^* jest przestrzenią wektorową, jest sens mówić o przestrzeni sprzężonej do V^* (mówimy–drugiej sprzężonej).
Niech $v \in V$ i niech $f \in V^*$. Wzór

$$f \longmapsto \varphi_v(f) = f(v)$$

określa odwzorowanie liniowe $\varphi_V : V \to (V^*)^*$.

STWIERDZENIE 9.2. *Odwzorowanie φ_V jest injekcją.*

DOWÓD: Jeżeli $\varphi_v = 0$, to dla wszystkich $f \in V^*$ mamy $f(v) = 0$. Wynika stąd, że $v = 0$. Wniosek ten, oczywisty dla przestrzeni wymiaru skończonego, wymaga użycia pewnika wyboru w przypadku przestrzeni wymiaru nieskończonego. ∎

STWIERDZENIE 9.3. *Jeśli* $\dim V$ *jest skończony, to odwzorowanie*

$$V \to (V^*)^* : v \mapsto \varphi_v,$$

gdzie $\varphi_v(f) = f(v)$, określa izomorfizm

$$V \simeq (V^*)^* .$$

DOWÓD: Odwzorowanie $v \mapsto \varphi_v$ jest liniowe i jest injekcją. Zatem $\dim(\operatorname{im} \varphi_V) = \dim V$. Ponadto $\dim(V^*)^* = \dim V^* = \dim V$, więc φ_V jest surjekcją, więc izomorfizmem. ∎

Niech $B \subset V$, $A \subset V^*$ będą dowolnymi podzbiorami.

DEFINICJA 9.4. *Zbiór*

$$B^\circ := \{f \in V^* : \forall v \in B \ f(v) = 0\}$$

nazywamy anihilatorem *zbioru B. Zbiór*

$$A^\circ := \{v \in V : \forall f \in A \ f(v) = 0\}$$

nazywamy anihilatorem *zbioru A.*

STWIERDZENIE 9.5.

(1) *Jeżeli* $B_1 \supset B$ *to* $B_1^\circ \subset B^\circ$. *Jeżeli* $A_1 \supset A$ *to* $A_1^\circ \subset A^\circ$.
(2) A°, B° *są podprzestrzeniami wektorowymi.*
(3) $B^\circ = \langle B \rangle^\circ$ *oraz* $A = \langle A \rangle^\circ$ *(dla przypomienia:* $\langle C \rangle$ *oznacza najmniejszą podprzestrzeń zawierającą zbiór* C*).*
(4)

$$(B_1 \cup B)^\circ = B_1^\circ \cap B^\circ$$
$$(A_1 \cup A)^\circ = A_1^\circ \cap A^\circ$$

W konsekwencji, gdy B, B_1, A, A_1 *są podprzestrzeniami, mamy*

$$(B_1 + B)^\circ = B_1^\circ \cap B^\circ$$
$$(A_1 + A)^\circ = A_1^\circ \cap A^\circ$$

DOWÓD: Przez proste sprawdzenie. ∎

Uwaga: na ogół $(B \cap B_1)^\circ \neq B^\circ + B_1^\circ$.

Przykład: Niech $V = \mathbb{R}$ i $B = \{0, 1\}, B_1 = \{-1, 0\}$. Mamy $B \cap B_1 = \{0\}$ i $B^\circ = \{0\} = B_1^\circ$ podczas gdy $(B_1 \cap B)^\circ = V^*$.

STWIERDZENIE 9.6. *Jeżeli* $\dim V = n$ *i* $\dim\langle B \rangle = k$, *to* $\dim B^\circ = n - k$.

DOWÓD: Niech $(e_1, \dots, e_k)$ będzie bazą $\langle B \rangle$, a $e = (e_1, \dots, e_n)$ jej uzupełnieniem do bazy w V. Mamy

$$f \in B^\circ = \langle B \rangle^\circ \iff f(e_1) = \cdots = f(e_k) = 0 \ .$$

Zdefiniujmy

$$f_{k+1}(e_i) = \begin{cases} 0 & i \neq k+1 \\ 1 & i = k+1 \end{cases}$$

i, podobnie, $f_{k+2}, \dots, f_n$. Oczywiste, że jeżeli $f \in B^\circ$, to

$$f = f(e_{k+1}) f_{k+1} + \cdots + f(e_n) f_n.$$

Funkcjonały $(f_{k+1}, \dots, f_n)$ są liniowo niezależne, bo jeżeli $f = \lambda^{k+1} f_{k+1} + \cdots + \mu^n f_n = 0$, to $\lambda^i = f(e_i) = 0$. Zatem układ $(f_{k+1}, \dots, f_n)$ tworzy bazę B°. ∎

Wnioski:

(1) Dla $A \subset V^*$ mamy $\dim A^\circ = n - \dim\langle A \rangle$.
(2) Jeżeli B_1, B_2 są podprzestrzeniami, to $(B_1 \cap B_2)^\circ = B_1^\circ + B_2^\circ$. Istotnie, mamy oczywiste zawieranie $(B_1 \cap B_2)^\circ \supset B_1^\circ + B_2^\circ$, z drugiej zaś strony

$$\dim(B_1 \cap B_2)^\circ = n - \dim(B_1 \cap B_2),$$

a więc na mocy Twierdzenia 4.19 i Stwierdzenia 9.5

$$\begin{aligned}\dim(B_1^\circ + B_2^\circ) &= \dim B_1^\circ + \dim B_2^\circ - \dim(B_1^\circ \cap B_2^\circ)\\ &= \dim B_1^\circ + \dim B_2^\circ - \dim(B_1 + B_2)^\circ\\ &= n - \dim B_1 + n - \dim B_2 - (n - (\dim B_1 + \dim B_2 - \dim(B_1 \cap B_2)))\\ &= n - \dim(B_1 \cap B_2) = \dim(B_1 \cap B_2)^\circ \ .\end{aligned}$$

(3) Przy utożsamieniu $(V^*)^*$ z V mamy $(B^\circ)^\circ = \langle B \rangle$, bo oczywiście $\langle B \rangle \subset (B^\circ)^\circ$, a z poprzednich rozważań $\dim(B^\circ)^\circ = \dim\langle B \rangle$.

Ze względu na symetrię między V i V^* będziemy pisać $\langle v, f \rangle$ zamiast $f(v)$ i $\varphi_v(f)$.

DEFINICJA 9.7. *Niech* $e = (e_1, \ldots, e_n)$ *będzie bazą w* V*. Bazę* $e^* = (e_1^*, \ldots, e_n^*)$ *w* V^* *zdefiniowaną przez warunki*

$$\langle e_i, e_j^* \rangle = \delta_{ij}$$

nazywamy bazą dualną *do* e*.*

Zwróćmy uwagę, że $e_i^* \in V^*$ zależy nie tylko od e_i, ale i od pozostałych wektorów bazy e. Mając bazę e w przestrzeni V możemy reprezentować $f \in V^*$ poprzez macierz jednowierszową $[f]^1{}_e \in \mathbf{M}^1{}_n$. Traktujemy tu $\mathbb{K}$ jako przestrzeń wektorową $\mathbb{K}^1$ z jedynką jako wektorem bazowym. Mając bazę e^* w V^* możemy reprezentować $f \in V^*$ również jako macierz jednokolumnową $[f]^{e^*} \in \mathbf{M}^n{}_1$.

STWIERDZENIE 9.8.

$$[f]^{e^*} = ([f]^1{}_e)^\mathsf{T} \ .$$

DOWÓD: i-tym wierszem macierzy $[f]^{e^*}$ jest i-ty współczynnik w rozwinięciu

$$f = \lambda^1 e_1^* + \cdots + \lambda^n e_n^* \ .$$

i-tą kolumną macierzy $[f]_e^1$ jest $\langle e_i, f \rangle$. Ale $\langle e_i, f \rangle = \lambda^i$. ■

Oznaczając i-tą kolumnę macierzy $[f]^1{}_e$ przez μ_i dostajemy

$$\langle v, f \rangle = \sum_i \mu_i \lambda_i,$$

gdzie $v = \sum_i \lambda_i e_i$ lub, równoważnie,

$$\langle v, f \rangle = ([f]^{e^*})^\mathsf{T} [v]^e.$$

Przykład bazy dualnej: $V = \mathbb{K}[n]$ i $e = (1, t, \ldots, t^n)$. e_i^* możemy zapisać jako $\left.\frac{1}{i!}\frac{\mathrm{d}^i}{\mathrm{d}t^i}\right|_{t=0}$.

9.2. Odwzorowania sprzężone.

Niech V, W będą przestrzeniami wektorowymi nad $\mathbb{K}$ i niech $F \in \mathsf{L}(V,W)$.

DEFINICJA 9.9. *Odwzorowanie*

$$W^* \to V^* \colon f \mapsto f \circ F$$

nazywamy odwzorowaniem sprzężonym *do F i oznaczamy F^*.*

Mamy więc $\langle v, F^* f\rangle = \langle Fv, f\rangle$.

STWIERDZENIE 9.10.

(1) $(\lambda F)^* = \lambda F^*$
(2) $(F_1 + F_2)^* = F_1^* + F_2^*$
(3) $F \in \mathsf{L}(V,W)$, $G \in \mathsf{L}(W,U)$, *to* $(G \circ F)^* = F^* \circ G^*$.

DOWÓD: Powyższe fakty są łatwe do sprawdzenia i wynikają z łączności składania odwzorowań oraz z rozdzielności składania odwzorowań względem ich dodawania. Dla przykładu,

$$(G \circ F)^*(h) = h \circ (G \circ F) = (h \circ G) \circ F = F^*(G^*(h)).$$

■

STWIERDZENIE 9.11. *Przy utożsamieniu $(V^*)^*$ z V oraz $(W^*)^*$ z W mamy*

$$F^{**} = F.$$

DOWÓD: Niech $v \in V$ i niech odpowiada mu $\varphi_v \in V^{**}$ w kanonicznej identyfikacji V i V^{**}. Mamy z definicji

$$F^{**}(\varphi_v) = \varphi_v \circ F^* \ .$$

Zatem, dla $g \in W^*$,

$$\langle \varphi_v \circ F^*, g\rangle = \langle v, F^* g\rangle = \langle Fv, g\rangle = \langle \varphi_{Fv}, g\rangle,$$

czyli $F^{**}(\varphi_v) = \varphi_{Fv}$ więc $F^{**}(v) = F(v)$ jeżeli utożsamimy v z φ_v i φ_{Fv} z Fv. ■

Wniosek: Odwzorowanie $\mathsf{L}(V,W) \ni F \mapsto F^* \in \mathsf{L}(W^*,V^*)$ jest izomorfizmem przestrzeni wektorowych: ze Stwierdzenia 9.10 wynika liniowość, a ze Stwierdzenia 9.11 – że odwzorowania sprzężenia $\mathsf{L}(V,W) \to \mathsf{L}(W^*,V^*)$ i $\mathsf{L}(W^*,V^*) \to \mathsf{L}(V,W)$ są wzajemnie odwrotne.

TWIERDZENIE 9.12 (FREDHOLMA). *Jeżeli V i W są skończenie wymiarowe, to*

$$\operatorname{im} F = (\ker F^*)^\circ, \quad \operatorname{im} F^* = (\ker F)^\circ$$

i wymiary tych przestrzeni (tzn. rzędy odwzorowań F i F^*) są równe.

DOWÓD: Niech $w \in \operatorname{im} F$ i $g \in \ker F^*$. Istnieje zatem $v \in V$ takie, że $w = Fv$. Mamy

$$\langle w, g\rangle = \langle Fv, g\rangle = \langle v, F^* g\rangle = 0,$$

czyli $w \in (\ker F^*)^\circ$, to znaczy, $\operatorname{im} F \subset (\ker F^*)^\circ$. Podobnie $\operatorname{im} F^* \subset (\ker F)^\circ$. Oznaczmy $k = \dim \operatorname{im} F$ i $k' = \dim \operatorname{im} F^*$. Powyższe zawierania oznaczają

$$k \leqslant n - \dim \ker F^* = n - (n - k') = k'$$
$$k' \leqslant n - \dim \ker F = n - (n - k) = k.$$

Stąd $k = k'$, a więc $\dim \operatorname{im} F = \dim(\ker F^*)^\circ$. Drugą równość dowodzimy analogicznie. ■

9.2.1. Macierz odwzorowania sprzężonego. Niech e będzie bazą V i $\widetilde{e}$ bazą W. Niech $F \in \mathsf{L}(V, W)$. Szukamy macierzy $[F^*]^{e^*}{}_{\tilde{e}^*}$. Macierz ta zdefiniowana jest przez równość

$$[F^* g]^{e^*} = [F^*]^{e^*}{}_{\tilde{e}^*}[g]^{\tilde{e}^*} .$$

Ale $[g]^{e^*} = ([g]^1{}_e)^\mathsf{T}$ oraz

$$[F^* g]^1{}_e = [g \circ F]^1{}_e = [g]^1{}_{\tilde{e}}[F]^{\tilde{e}}{}_e .$$

Stąd $[F^* g]^{e^*} = ([F]^{\tilde{e}}{}_e)^\mathsf{T}[g]^{\tilde{e}^*}$, a zatem

$$[F^*]^{e^*}{}_{\tilde{e}^*} = ([F]^{\tilde{e}}{}_e)^\mathsf{T} . \tag{9.1}$$

Spostrzeżenie: Z powyższego, łącznie z twierdzeniem Fredholma, wynika, że rząd wierszowy macierzy jest równy rzędowi kolumnowemu.

9.2.2. Odwzorowania samosprzężone. Niech $F\colon V \to V^*$, zatem $F^*\colon (V^*)^* = V \to V^*$

DEFINICJA 9.13. *Odwzorowanie liniowe $F\colon V \to V^*$ nazywamy* samosprzężonym, *jeżeli $F = F^*$.*

Dla odwzorowania samosprzężonego mamy $\langle w, Fv\rangle = \langle v, Fw\rangle$. Ponieważ

$$[F^*]^{e^*}{}_e = ([F]^*{}_e)^\mathsf{T}$$

F jest samosprzężone wtedy i tylko wtedy, gdy

$$[F]^{e^*}{}_e = ([F]^{e^*}{}_e)^\mathsf{T} . \tag{9.2}$$

Dla odwzorowań samosprzężonych sformułowanie twierdzenie Fredholma nieco się upraszcza.

TWIERDZENIE 9.14 (FREDHOLMA DLA ODWZOROWAŃ SAMOSPRZĘŻONYCH). *Jeżeli $F\colon V \to V^*$ jest samosprzężone, to* $\operatorname{im} F = (\ker F)^\circ$.

Rozdział 10. Formy dwuliniowe. Formy kwadratowe

10.1. Formy dwuliniowe.

DEFINICJA 10.1. Formą dwuliniową *(biliniową) na przestrzeni wektorowej V nazywamy odwzorowanie $\Phi: V \times V \to \mathbb{K}$ takie, że dla każdego $v \in V$ liniowe są odwzorowania*

$$V \ni w \mapsto \Phi(v, w) \in \mathbb{K}$$

i

$$V \ni w \mapsto \Phi(w, v) \in \mathbb{K}.$$

Jak łatwo zauważyć, zbiór form dwuliniowych na przestrzeni V tworzy przestrzeń wektorową. Oznaczamy ją $\mathsf{L}(V, V; \mathbb{K})$. Z każdą formą dwuliniową Φ stowarzyszamy odwzorowanie liniowe $F_\Phi: V \to V^*$, kładąc

$$\langle w, F_\Phi v\rangle = \Phi(w, v)\ .$$

STWIERDZENIE 10.2. *Przyporządkowanie $\mathsf{L}(V, V; \mathbb{K}) \ni \Phi \mapsto F_\Phi \in \mathsf{L}(V, V^*)$ jest liniowym izomorfizmem.*

DOWÓD: Dla form dwuliniowych Φ i Ψ mamy

$$\begin{aligned}\langle w, (F_\Phi + F_\Psi)v\rangle &= \Phi(w, v) + \Psi(w, v)\\ &= (\Phi + \Psi)(w, v)\\ &= \langle w, F_{\Phi+\Psi} v\rangle,\end{aligned}$$

więc przyporządkowanie o którym mowa jest addytywne. Podobnie dowodzimy jednorodności. Mamy zatem liniowość. Jeżeli $F_\Phi = 0$ to, dla wszystkich $v, w \in V$, mamy $\langle w, F_\Phi v\rangle = 0 = \Phi(w, v)$. Przyporządkowanie jest więc injektywne.

Jeśli z kolei $F \in \mathsf{L}(V, V^*)$, to $F = F_\Phi$, gdzie $\Phi(w, v) = \langle w, Fv\rangle$. Dowodzi to surjektywności. ■

Rząd formy dwuliniowej Φ definiujemy jako rząd (t. j. wymiar obrazu) odwzorowania F_Φ.

10.1.1. Macierz formy dwuliniowej. Macierz F_Φ w bazach e, e^* wygląda następująco: $[F_\Phi]^{e^*}{}_e = [a^i_j]$, gdzie

$$\bar{a}_j = [F_\Phi(e_j)]^{e^*} = ([F_\Phi(e_j)]^1{}_e)^{\mathsf{T}},$$

zatem $a^i_j = \langle e^i, F_\Phi(e_j)\rangle = \Phi(e_i, e_j)$. Macierz tę nazywamy *macierzą formy dwuliniowej Φ w bazie e*. Oznaczać ją będziemy $[\Phi]_e$. Oczywisty jest wzór

$$\Phi(w, v) = ([F_\Phi(v)]^1{}_e[w]^e)^{\mathsf{T}} = [w]^e[F_\Phi]^{e^*}_e[v]^e = [w]^e[\Phi]_e[v]^e. \qquad (10.1)$$

Weźmy teraz inną bazę $\tilde{e}$ w V. Zgodnie z ogólnym wzorem (6.2) mamy

$$[F_\Phi]^{\tilde{e}^*}{}_e = [Id_{V^*}]^{\tilde{e}^*}{}_{e^*}[F_\Phi]^{e^*}{}_e[Id_V]^e{}_{\tilde{e}},$$

z drugiej zaś strony, na mocy (9.1),

$$[Id_{V^*}]^{\tilde{e}^*}{}_{e^*} = \left([Id_V]^e{}_{\tilde{e}}\right)^{\mathsf{T}},$$

więc

$$[\Phi]_{\tilde{e}} = \left([Id_V]^e{}_{\tilde{e}}\right)^{\mathsf{T}}[\Phi]_e[Id_V]^e{}_{\tilde{e}}. \tag{10.2}$$

Przykład: V - przestrzeń wielomianów stopnia $\leqslant k$. $\Phi(w,v) = w'(0)v(a)$, gdzie $a \in \mathbb{R}$ jest ustaloną liczbą. Mamy

$$F_\Phi(v) = v(a)D^1, \quad \text{gdzie} \quad \langle w, D^i\rangle \overset{\text{def}}{=} \frac{\mathrm{d}^i w}{\mathrm{d}t^i}(0).$$

DEFINICJA 10.3. *Formę dwuliniową Φ na V nazywamy:*

(1) Symetryczną, *jeżeli* $\Phi(w,v) = \Phi(v,w)$.
(2) Antysymetryczną, *jeżeli* $\Phi(w,v) = -\Phi(w,v)$.
(3) Niezdegenerowaną, *jeżeli F_Φ jest izomorfizmem. Gdy* $\dim V < \infty$ *jest to równoważne warunkowi:* $\forall v \neq \mathbf{0}\ \exists w$ *takie, że* $\Phi(v,w) \neq 0$.

Bezpośrednio z definicji wynika następujące stwierdzenie.

STWIERDZENIE 10.4. *Forma dwuliniowa Φ jest symetryczna (antysymetryczna, niezdegenerowana) wtedy i tylko wtedy, gdy odwzorowanie F_Φ jest samosprzężone (antysamosprzężone, bijektywne), a także wtedy i tylko wtedy, gdy macierz $[\Phi]_e$ jest symetryczna (antysymetryczna, nieosobliwa).*

STWIERDZENIE 10.5. *Każdą formę dwuliniową Φ można jednoznacznie rozłożyć na sumę $\Phi = \Phi_a + \Phi_s$, gdzie Φ_s jest symetryczną, a Φ_a antysymetryczną formą dwuliniową.*

Dowód: Istnienie:
Kładziemy

$$\Phi_s(v,w) = \frac{1}{2}\{\Phi(v,w) + \Phi(w,v)\}$$

oraz

$$\Phi_a(v,w) = \frac{1}{2}\{\Phi(v,w) - \Phi(w,v)\}.$$

Jednoznaczność:
Jeżeli mamy rozkład $\Phi = \Phi'_s + \Phi'_a$ na część symetryczną i antysymetryczną, to $\Phi(w,v) + \Phi(v,w) = 2\Phi'_s(w,v)$ i $\Phi(w,v) - \Phi(v,w) = 2\Phi'_a(w,v)$. ■

Przykłady:

(1) Forma $\Phi(w,v) = w'(0)v(a)$ (jak w poprzednim przykładzie) jest formą niesymetryczną (t. j. nie jest symetryczna) i zdegenerowaną.
(2) $V = W \times W^*$. Forma

$$\Omega((w,f),(w_1,f_1)) = \langle w, f_1\rangle - \langle w_1, f\rangle,$$

zwana formą kanoniczną, jest antysymetryczna i niezdegenerowana.

Mając ustaloną formę Φ, dla każdego $A \subset V$ zdefiniujmy

$$A^{\&} = F_\Phi^{-1}(A^\circ) = \{V \ni v : \Phi(w,v) = 0 \;\text{ dla każdego } w \in A\}\ .$$

Przykład: $V = W \times W^*$ i $\Phi = \Omega$ – forma kanoniczna z poprzedniego przykładu. Niech $F\colon W \to W^*$ oraz A - wykres F tzn $A = \{(w,f)\colon Fw = f\}$. Mamy

$$\begin{aligned} A^{\&} &= \{(w,f)\colon \forall w_1\ \Omega((w_1,Fw_1),(w,f)) = \langle w_1, f\rangle - \langle w, Fw_1\rangle = 0\} \\ &= \{(w,f)\colon f = F^* w\}, \end{aligned}$$

zatem $A^{\&}$ jest wykresem F^*. Stąd F jest samosprzężony wtedy i tylko wtedy, gdy $A = A^{\&}$.

10.2. Formy kwadratowe.

Mając formę dwuliniową $\Phi\colon V \times V \to \mathbb{K}$ możemy zdefiniować funkcję φ na V wzorem

$$\varphi\colon V \to \mathbb{K}\colon v \mapsto \varphi(v) = \Phi(v,v).$$

Mamy

$$\varphi(v) = \Phi(v,v) = \Phi_s(v,v) + \Phi_a(v,v) = \Phi_s(v,v). \tag{10.3}$$

Zauważmy, że

$$\varphi(v+w) = \Phi_s(v,v) + \Phi_s(w,w) + 2\Phi_s(v,w),$$

czyli

$$\Phi_s(v,w) = \frac{1}{2}(\varphi(v+w) - \varphi(v) - \varphi(w)). \tag{10.4}$$

Wzór (10.4) nazywamy *formułą polaryzacyjną*. Funkcję φ nazywamy *formą kwadratową* odpowiadającą formie dwuliniowej Φ. Z wzorów (10.3) (10.4) wynika, że odpowiedność miedzy formami kwadratowymi a symetrycznymi formami dwuliniowymi jest bijektywna.

STWIERDZENIE 10.6. *Funkcja* $\varphi: V \to \mathbb{K}$ *jest formą kwadratową wtedy i tylko wtedy, gdy spełnione są dwa warunki:*

(1) *Funkcja*

$$V \times V \ni (v, w) \mapsto \frac{1}{2}(\varphi(v+w) - \varphi(v) - \varphi(w))$$

jest formą dwuliniową. Oznaczamy ją Φ.

(2) $\Phi(v, v) = \varphi(v)$,

a także wtedy i tylko wtedy, gdy

(1') $\forall \lambda \in \mathbb{K},\ \forall v \in V\ \varphi(\lambda v) = \lambda^2 \varphi(v)$

(2') $\forall v, w \in V$ *spełniony jest warunek równoległoboku*

$$\varphi(v+w) + \varphi(v-w) = 2\varphi(v) + 2\varphi(w).$$

Mamy więc ciąg bijektywnych odpowiedniości:
Formy kwadratowe $\Leftrightarrow$ formy dwuliniowe symetryczne $\Leftrightarrow$ odwzorowania samosprzężone.

DEFINICJA 10.7. Macierzą formy kwadratowej *w bazie* e *nazywamy macierz w bazie* e *odpowiadającej jej symetrycznej formy dwuliniowej.*

Spostrzeżenie: jeżeli $e = (e_1, \dots, e_n)$ jest bazą w V i $e^* = (\phi_1, \dots, \phi_n)$ jest bazą do niej dualną, to dla każdej formy kwadratowej φ mamy

$$\varphi = \sum_{i,j} \Phi(e_i, e_j)\phi_i\phi_j\ .$$

DEFINICJA 10.8. Rzędem *formy kwadratowej nazywamy rząd stowarzyszonej z nią formy dwuliniowej.*

Rząd formy kwadratowej jest więc równy rzędowi jej macierzy, jak też wymiarowi obrazu odpowiadajacego jej odwzorowania samosprzężonego (macierz formy kwadratowej jest równa macierzy odpowiadającej jej formy dwuliniowej i macierzy odpowiadajacego tej formie odwzorowania samosprzężonego).

DEFINICJA 10.9. *Jeśli forma* φ *ma w bazie* e *macierz diagonalną, to mówimy, że* e *jest* bazą diagonalizującą *formy* φ.

TWIERDZENIE 10.10 (LAGRANGE). *Każda forma kwadratowa ma bazę diagonalizującą.*

SZKIC DOWODU: Niech, jak w spostrzeżeniu, $\varphi = \sum_{i,j} a^{ij}\phi_i\phi_j$, gdzie $a^{ij} = a^{ji}$. Załóżmy, że istnieje i takie, że $a^{ii} \neq 0$, na przykład $i = 1$. Kładziemy wówczas

$$\psi_1 = \phi_1 + \frac{1}{a^{11}} \sum_{i \neq 1} a^{1i}\phi_i,$$

a wtedy w $\varphi - a^{11}\psi_1^2$ redukują się składniki zawierające ϕ_1, więc

$$\varphi = a^{11}\psi_1^2 + \sum_{i,j=2}^{n} b^{ij}\phi_i\phi_j.$$

Zauważmy, że $(\psi_1, \phi_2, \phi_3, \dots, \phi_n)$ jest bazą w V^*.
Jeżeli jednak $a^{ii} = 0$ dla każdego i, lecz np. $a^{12} \neq 0$, to kładziemy $\psi_1 = \phi_1 + \phi_2$ i $\psi_2 = \phi_1 - \phi_2$ i wtedy

$$\phi_1\phi_2 = \frac{1}{4}(\psi_1^2 - \psi_2^2),$$

mamy więc

$$\varphi = \frac{a^{12}}{4}\psi_1^2 + \cdots .$$

Ponadto widzimy, że $(\psi_1, \psi_2, \phi_3, \dots, \phi_n)$ jest bazą w V^*. Sprowadziliśmy w ten sposób problem do przypadku już rozpatrywanego: współczynnik przy ψ_1^2 jest $\neq 0$. Indukcyjnie dostaniemy w końcu bazę diagonalizującą formy φ. ■

Rozpatrzmy dwa przypadki:

(1) $\mathbb{K} = \mathbb{C}$. W bazie diagonalizującej mamy $\varphi = \sum_i a^i\phi_i^2$. Kładąc dla $a^i \neq 0$ $\psi_i = b_i\phi_i$, gdzie $b_i^2 = a^i$ dostajemy $\varphi = \sum_i \psi_i^2$, czyli współczynniki są równe 0 lub 1.

(2) $\mathbb{K} = \mathbb{R}$ Przyjmijmy $\psi_i = b_i\phi_i$, gdzie $b_i = \sqrt{|a_i|}$, jeśli $a_i \neq 0$, oraz $\psi_i = \phi_i$, jeśli $a_i = 0$. W rezultacie dostajemy postać diagonalną formy kwadratowej

$$\varphi = \sum_{i=1}^{p} \psi_i^2 - \sum_{i=1}^{q} \psi_{p+i},$$

gdzie $p + q = k =$ rząd formy. Tutaj współczynniki są równe $0, 1$ lub -1.

TWIERDZENIE 10.11 (SYLVESTERA O „BEZWŁADNOŚCI"). *Niech* $\mathbb{K} = \mathbb{R}$ *i niech*

$$(\phi_1, \dots, \phi_n),\ (\psi_1, \dots, \psi_n)$$

będą dwiema bazami diagonalizującymi formy φ*:*

$$\varphi = \sum_{i=1}^{p} \phi_i^2 - \sum_{i=1}^{q} \phi_{p+i}^2 = \sum_{i=1}^{r} \psi_i^2 - \sum_{i=1}^{s} \psi_{r+i}^2 .$$

Wówczas $p = r,\ q = s$.

DOWÓD: Oczywiste, że sumy $p+q$, $r+s$ są równe rzędowi φ (są więc sobie równe). Przypuśćmy, że $r > p$, czyli $s < q$. Zatem układ form: $(\phi_1, \ldots, \phi_p, \psi_{r+1}, \ldots, \psi_n)$ ma $p+n-r < n$ elementów. Rozpatrzmy układ równań:

$$\begin{cases} \langle v, \phi_i \rangle = 0, & i = 1, \ldots, p \\ \langle v, \psi_i \rangle = 0, & i = r+1, \ldots, n. \end{cases}$$

Jest to układ $p+n-r$ równań liniowych na wektor w przestrzeni n wymiarowej. Istnieje zatem rozwiązanie niezerowe. Niech $v_0 \neq 0$ będzie takim rozwiązaniem, tzn.,

$$\varphi(v_0) = -\sum_{i=1}^{q} (\phi_{p+i}(v_0))^2 \leqslant 0$$

i

$$\varphi(v_0) = \sum_{i=1}^{r} (\psi_i(v_0))^2 \geqslant 0.$$

W konsekwencji, $\psi_1(v_0) = \cdots = \psi_r(v_0) = 0$. Zatem $\langle v_0, \psi_i \rangle = 0$ dla $i = 1, \ldots, n$, czyli $v_0 = \mathbf{0}$. Sprzeczność. Mamy więc $r \leqslant p$. Analogicznie $p \leqslant r$, czyli $p = r$. ■

DEFINICJA 10.12. *Niech φ będzie formą kwadratową na przestrzeni wektorowej V nad ciałem $\mathbb{K} = \mathbb{R}$. Niech w pewnej bazie $\varphi = \sum_{i=1}^{p} \phi_i^2 - \sum_{i=1}^{q} \phi_{p+i}^2$. Parę (p, q) nazywamy* sygnaturą *formy φ. Liczbę $p - q$ nazywamy* indeksem *formy φ .*

10.3. Znajdowanie bazy diagonalizującej metodą Grama - Schmidta.

Niech $e = (e_1, \ldots, e_n)$ będzie bazą w V i niech φ będzie formą kwadratową na V. Przez Φ oznaczmy odpowiadającą jej symetryczną formę dwuliniową. Wprowadźmy podprzestrzenie $L_i = \langle \{e_1, \ldots, e_i\} \rangle$. Oczywiście $L_n = V$. Załóżmy, że $\varphi|_{L_i}$ jest niezdegenerowana dla każdego i. Definiujemy indukcyjnie układ wektorów $(\widetilde{e_1}, \ldots \widetilde{e_n})$:
$\widetilde{e}_1 = e_1$, a z niezdegenerowania na L_1 mamy $\Phi(\widetilde{e}_1, \widetilde{e}_1) = \varphi(e_1) \neq 0$;
$\widetilde{e}_2 = e_2 - \dfrac{1}{\varphi(\widetilde{e}_1)} \Phi(e_2, \widetilde{e}_1)\widetilde{e}_1$. Mamy $\Phi(\widetilde{e}_1, \widetilde{e}_2) = 0$, więc z niezdegenerowania Φ na L_2 wynika, że $\Phi(\widetilde{e}_2, \widetilde{e}_2) = \varphi(\widetilde{e}_2) \neq 0$. i.t.d.;
$\widetilde{e}_k = e_k - \displaystyle\sum_{i=1}^{k-1} \frac{1}{\varphi(\widetilde{e}_i)} \Phi(e_k, \widetilde{e}_i)\widetilde{e}_i$. Mamy $\Phi(\widetilde{e}_i, \widetilde{e}_k) = 0$ dla $i < k$, więc z niezdegenerowania Φ na L_k dostajemy $\Phi(\widetilde{e}_k, \widetilde{e}_k) = \varphi(\widetilde{e}_k) \neq 0$. Oczywiście $\langle \{\widetilde{e}_1, \ldots, \widetilde{e}_k\} \rangle = \langle \{e_1, \ldots, e_k\} \rangle = L_k$. Baza $(\widetilde{e}_1, \ldots, \widetilde{e}_k)$ jest bazą diagonalizującą formy φ.

10.4. Znajdowanie sygnatury nieosobliwej formy metodą Jacobiego.

Niech φ będzie formą kwadratową na V i niech $e = (e_1, \ldots, e_n)$ będzie bazą w V taką, że na $L_i = \langle \{e_1, \ldots, e_i\} \rangle$ forma φ jest niezdegenerowana. Zastosujmy metodę diagonalizacji Grama-Schmidta. W jej wyniku dostajemy bazę diagonalizującą $\widetilde{e} = (\widetilde{e}_1, \ldots, \widetilde{e}_n)$. Mamy, jak łatwo zauważyć,

$$[Id]^e_{\widetilde{e}} = \begin{bmatrix} 1 & * & * & \cdots & * \\ 0 & 1 & * & \cdots & * \\ 0 & 0 & 1 & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \overset{\text{ozn}}{=} A\,.$$

oraz, na mocy (10.2),

$$[\varphi]_{\widetilde{e}} = ([Id]^e_{\widetilde{e}})^{\mathsf{T}} [\varphi]_e [Id]^e_{\widetilde{e}} = A^{\mathsf{T}} [\varphi]_e A\,.$$

Oznaczmy $G = [g^i_j] = [\varphi]_e$. Ponieważ

$$[\varphi]_{\widetilde{e}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$

jest macierzą diagonalną i A jest macierzą trójkątną, mamy

$$\begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \alpha_i \end{bmatrix} = A_i^{\mathsf{T}} G_i A_i,$$

gdzie A_i, G_i są macierzami złożonymi z pierwszych i wierszy oraz pierwszych i kolumn macierzy A, G. Zatem

$$\alpha_1 \cdots \alpha_i = (\det A_i)^2 \det G_i$$

i $\operatorname{sgn}(\alpha_1 \cdots \alpha_i) = \operatorname{sgn}(\det G_i)$. Stąd,

STWIERDZENIE 10.13. *Jeżeli φ jest niezdegenerowana na każdej przestrzeni L_i, tzn. wszystkie $\det G_i$ są różne od 0, to sygnatura formy φ równa jest (p, q), gdzie p jest liczbą dodatnich, a q liczbą ujemnych wyrazów ciągu*

$$\det G_1, \frac{\det G_2}{\det G_1}, \ldots, \frac{\det G_n}{\det G_{n-1}}\,.$$

Przykład ogólny:

DEFINICJA 10.14. *Forma kwadratowa* φ *jest* dodatnio określona, *jeżeli*

$$\forall v \neq \mathbf{0}\ \varphi(v) > 0.$$

Forma φ *jest* ujemnie określona *jeżeli* $\forall v \neq \mathbf{0}\ \varphi(v) < 0$.

Oczywiste, że:

$$\varphi \text{ jest dodatnio określona (d.o.)} \iff \text{sygnatura } \varphi = (n, 0)$$

i

$$\varphi \text{ jest ujemnie określona (u.o.)} \iff \text{sygnatura } \varphi = (0, n).$$

STWIERDZENIE 10.15 (KRYTERIUM SYLVESTERA). φ *jest d. o.* $\iff \det G_i > 0$ *dla* $i = 1, \dots, n$. φ *jest u. o.* $\Leftrightarrow (-1)^i \det G_i > 0$ *dla* $i = 1, \dots, n$.

DOWÓD: Jeżeli φ jest d.o., to dla każdej bazy e φ obcięta do $L_i = \langle\{e_1, \dots, e_i\}\rangle$ jest dodatnio określona, a więc – tym bardziej – niezdegenerowana. Zatem

$$\det G_1, \frac{\det G_2}{\det G_1}, \dots, \frac{\det G_n}{\det G_{n-1}}$$

są dodatnie, czyli $\det G_1, \dots, \det G_n > 0$. Druga część wynika z pierwszej oraz z tego, że φ jest d.o. wtedy i tylko wtedy, gdy $(-\varphi)$ jest ujemnie określona. ■

Przykład: Dla jakich $\lambda \in \mathbb{R}$ forma kwadratowa φ na $\mathbb{R}^3$, gdzie $\varphi((x_1, x_2, x_3)) = 5x_1^2 + x_2^2 + \lambda x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$, jest dodatnio określona? W bazie kanonicznej e mamy

$$[\varphi]_e = G = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & \lambda \end{bmatrix}.$$

Stąd $\det G_1 = 5,\ \det G_2 = 1$, i $\det G_3 = \det G = \lambda - 2 > 0$. Odpowiedź: dla $\lambda > 2$.

Rozdział 11. Przestrzenie euklidesowe

11.1. Iloczyn skalarny.

W tym rozdziale $\mathbb{K} = \mathbb{R}$!

DEFINICJA 11.1. Iloczynem skalarnym *w przestrzeni wektorowej V (nad $\mathbb{R}$) nazywamy symetryczną i dodatnio określoną formę dwuliniową. To znaczy, funkcja $g: V \times V \to \mathbb{R}$ jest iloczynem skalarnym, jeśli*

(1) $g(v, v) > 0$ *dla* $v \neq 0$,
(2) $g(v, w) = g(w, v)$,
(3) g *jest formą dwuliniową.*

Przestrzeń wektorową nad ciałem liczb rzeczywistych z ustalonym iloczynem skalarnym nazywamy *przestrzenią euklidesową.*

Ponieważ mamy wzajemnie jednoznaczną odpowiedniość między formami kwadratowymi i symetrycznymi formami dwuliniowymi, g jest wyznaczone jednoznacznie przez formę kwadratową: $v \mapsto g(v, v)$.

Oznaczenia:

(1) $g(v, w)$ oznaczać będziemy $(v|w)$.
(2) $(g(v, v))^{\frac{1}{2}}$, oznaczać będziemy $\|v\|$ i nazywać będziemy normą (długością) wektora.

11.1.1. Podstawowe własności iloczynu skalarnego:.

STWIERDZENIE 11.2 (Tożsamość równoległoboku). $\|v+w\|^2+\|v-w\|^2 = 2(\|v\|^2 + \|w\|^2)$

Dowód: $(v+w|v+w)+(v-w|v-w) = 2(v|v)+2(w|w)$ z dwuliniowości iloczynu skalarnego. ■

TWIERDZENIE 11.3 (Nierówność Schwarza). *Jeśli V jest przestrzenią wektorową z iloczynem skalarnym, to*

$$|(v|w)| \leqslant \|v\| \, \|w\|.$$

Równość zachodzi wtedy i tylko wtedy, gdy v i w są liniowo zależne.

Dowód: Jeśli $v = \mathbf{0}$, to twierdzenie jest trywialne.
Jeśli $v \neq \mathbf{0}$, to rozpatrzmy funkcję $\alpha: \mathbb{R} \ni t \mapsto \|tv + w\|^2 \in \mathbb{R}$. Mamy $\alpha(t) = t^2(v|v) + 2t(v|w) + (v|w)$. Oczywiście $\alpha(t) \geqslant 0$, zatem wyróżnik tego trójmianu jest niedodatni, tzn:

$$(v|w)^2 - (\|v\| \, \|w\|)^2 \leqslant 0\ .$$

Jeżeli $w = \lambda v$, to

$$|(v|w)| = |\lambda| \|v\|^2 = \|\lambda v\| \cdot \|v\| = \|v\| \cdot \|w\|.$$

Niech teraz $|(v|w)| = \|v\| \cdot \|w\|$ i $|(v|w)| = \epsilon(v|w)$. Rozważmy funkcję

$$\beta: t \mapsto \beta(t) = \|\epsilon t v + w\|^2 = t^2\|v\|^2 + 2t|(v|w)| + \|w\|^2 =$$

$$= t^2\|v\|^2 + 2t\|v\|\|w\| + \|w\|^2 = (t\|v\| + \|w\|)^2.$$

β jest więc równe zero dla $t_0 = -\frac{\|w\|}{\|v\|}$, czyli $0 = -\epsilon\frac{\| w \|}{\| v \|}v + w$ i $w = \epsilon\frac{\|w\|}{\|v\|}v$. ■

STWIERDZENIE 11.4 (NIERÓWNOŚĆ TRÓJKĄTA).

$$\|v + w\| \leqslant \|v\| + \|w\|.$$

Równość zachodzi wtedy i tylko wtedy, gdy $(v|w) = \|v\|\|w\|$ *lub, równoważnie, gdy* v *i* w *są liniowo zależne.*

DOWÓD:

$$\|v + w\|^2 = (v + w|v + w) = \|v\|^2 + 2(v|w) + \|w\|^2 \leqslant$$

$$\leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

Pozostała część stwierdzenia wynika bezpośrednio z tego rachunku i z poprzedniego stwierdzenia. ■

TWIERDZENIE 11.5 (O REPREZENTACJI FUNKCJONAŁU LINIOWEGO). *Odwzorowanie* F_g *stowarzyszone z iloczynem skalarnym* g, $F_g: V \to V^*$,

$$\langle v, F_g w\rangle = (v|w) = g(v, w)$$

jest liniowym izomorfizmem (patrz 2.1). W konsekwencji, dla każdego fukcjonału $f \in V^*$ *istnieje dokładnie jeden wektor* $w_f \in V$ *taki, że dla każdego wektora* $v \in V$ *zachodzi równość*

$$\langle v, f\rangle = (v|w_f) = (w_f|v).$$

DOWÓD: Jeżeli $v \in \ker F_g$, to $F_g(v) = 0$, a więc

$$0 = \langle v, F_g(v)\rangle = g(v, v) = \|v\|^2.$$

Stąd $v = \mathbf{0}$, czyli $\ker F_g = \{\mathbf{0}\}$. Ponieważ wymiary przestrzeni V i V^* są równe oznacza to, że F_g jest izomorfizmem. ■

11.2. Prostopadłość. Rzut prostopadły.

DEFINICJA 11.6. *Niech $v, w \in V$. Mówimy, że wektor v jest* prostopadły *do w (piszemy $v \perp w$) jeżeli $(v|w) = 0$.*

STWIERDZENIE 11.7 „PITAGORASA”. *Jeżeli $(v|w) = 0$ to*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Niech $A \subset V$ będzie dowolnym podzbiorem. Zdefiniujemy *dopełnienie ortogonalne* $A^\perp$ przestrzeni A wzorem

$$A^\perp = \{v \in V : (v|w) = 0 \ \ \forall w \in A\} = F_g^{-1}(A^\circ).$$

(Porównaj z rozdziałem 10.1. strona 66 i zdefiniowanym tam zbiorem $A^\&$).

TWIERDZENIE 11.8. *Niech V będzie przestrzenią wektorową nad $\mathbb{R}$ z iloczynem skalarnym g. Niech $W \subset V$ będzie podprzestrzenią wektorową. Wówczas*

$$V = W \oplus W^\perp.$$

DOWÓD: Niech $v \in W \cap W^\perp$. Wtedy $(v|v) = \|\, v \,\|^2 = 0$, czyli $v = \mathbf{0}$; zatem $W \cap W^\perp = 0$.
Czy $V = W + W^\perp$?
Wystarczy policzyć wymiary. Jeżeli $\dim V = n$ i $\dim W = k$, to (patrz rozdział 9.1.) $\dim W^\circ = n - k$. Ale $W^\perp = F_g^{-1}(W^\circ)$, gdzie $F_g : V \to V^*$ jest samosprzężonym izomorfizmem stowarzyszonym z g. Zatem $\dim W^\perp = n - k$ i $\dim(W + W^\perp) = n$.

■

DEFINICJA 11.9. *Niech (przy założeniach powyższego twierdzenia) $v = w + w'$, gdzie $v \in V$, , $w \in W$, $w' \in W^\perp$. Wektor $P_W v \stackrel{\text{def}}{=} w$ nazywamy* rzutem ortogonalnym *wektora v na podprzestrzeń W.*

DEFINICJA 11.10. Odstępem *wektora v od podprzestrzeni W nazywamy liczbę*

$$\mathsf{d}(v, W) = \inf_{y \in W} \|v - y\|.$$

STWIERDZENIE 11.11.

$$\mathsf{d}(v, W) = \|v - P_W v\|.$$

DOWÓD: Dla $y \in W$ mamy $P_W v - y \in W$ a ponieważ $v - P_W v \in W^\perp$, więc, z twierdzenia Pitagorasa,

$$\|v - y\|^2 = \|(v - P_W v) + (P_W v - y)\|^2 = \|v - P_W v\|^2 + \|P_W v - y\|^2 \geqslant \|v - P_W v\|^2.$$

■

11.3. Przekształcenia ortogonalne.

DEFINICJA 11.12. *Odwzorowanie $F\colon V \to V$ nazywamy przekształceniem (odwzorowaniem, operatorem)* ortogonalnym*, jeżeli $(Fx|Fy) = (x|y)$ dla wszystkich $x, y \in V$.*

Uwagi:

(1) Operator ortogonalny jest nieosobliwy (ma trywialne jądro). Istotnie, mamy $\|Fx\| = \|x\|$, jeśli więc $Fx = \mathbf{0}$, to $x = 0$.
(2) Jeżeli operatory F i G są ortogonalne, to F^{-1}, $F \circ G$ są też ortogonalne. Nie są natomiast, na ogół ortogonalne odwzorowania $F + G$, λG. Przekształcenia ortogonalne (z działaniem składania) tworzą grupę (ortogonalną), którą oznaczamy $\mathsf{O}(V, g)$.

11.3.1. Reprezentacja macierzowa odwzorowań ortogonalnych. Niech $e = (e_1, ..., e_n)$ będzie bazą w V. Mamy

$$(v|w) = ([v]^e)^{\mathsf{T}}[g]_e[w]^e$$

oraz

$$(Fv|Fv) = ([F]^e{}_e[v]^e)^{\mathsf{T}}[g]_e([F]^e{}_e[w]^e) = ([v]^e)^{\mathsf{T}}(([F]^e{}_e)^{\mathsf{T}}[g]_e[F]^e{}_e)[w]^e.$$

Stąd F jest ortogonalny wtedy i tylko wtedy, gdy

$$[g]_e = ([F]^e{}_e)^{\mathsf{T}}[g]_e[F]^e{}_e. \tag{11.1}$$

DEFINICJA 11.13. *Baza e w której macierz iloczynu skalarnego $[g]_e$ jest diagonalna nazywa się bazą* ortogonalną.
Baza e w której macierz iloczynu skalarnego $[g]_e$ jest równa macierzy jednostkowej I nazywa się bazą ortonormalną.
Innymi słowy – e jest bazą ortonormalną jeżeli $(e_i|e_j) = \delta_{ij}$.

Zauważmy, że jeżeli e jest bazą ortonormalną, $v = \sum \lambda^i e_i$, to $\lambda^i = (v|e_i)$. Istnienie bazy ortogonalnej i, w konsekwencji, ortonormalnej, wynika z konstrukcji bazy diagonalizującej metodą Grama–Schmidta.

Odwzorowanie F jest więc ortogonalne wtedy i tylko wtedy, gdy w (dowolnej) bazie ortonormalnej e

$$[F]^e{}_e{}^{\mathsf{T}}[F]^e{}_e = I.$$

DEFINICJA 11.14. *Kwadratową macierz A taką, że $A^{\mathsf{T}}A = I$ nazywamy* macierzą ortogonalną.

W bazie ortonormalnej macierz przekształcenia ortogonalnego jest więc macierzą ortogonalną. Oczywiście, macierz A jest macierzą ortogonalną wtedy i tylko wtedy, gdy

$$\bar{a}_i^{\mathsf{T}}\bar{a}_j = \delta_{ij}, \tag{11.2}$$

gdyż (i,j)-tym wyrazem $A^{\mathsf{T}}A$ jest $\bar{a}_i^{\mathsf{T}}\bar{a}_j$.

STWIERDZENIE 11.15. *Niech F będzie operatorem ortogonalnym a e - bazą ortonormalną. Wtedy Fe jest też bazą ortonormalną.*

Dowód:

$$(Fe_i \mid Fe_j) = (e_i \mid e_j)\delta_{ij}.$$

■

Twierdzenie odwrotne jest też prawdziwe: jeżeli dla pewnej bazy ortonormalnej $(e_1, \ldots, e_n)$ ciąg $(Fe_1, \ldots, Fe_n)$ jest też bazą ortonormalną, to F jest ortogonalny. Wynika to z prostego rachunku:

$$\begin{aligned}(Fv, Fw) &= (F(\lambda^1 e_1 + \cdots + \lambda^n e_n) \mid F(\mu^1 e_1 + \cdots + \mu^n e_n)) \\ &= \sum_{i,j} \lambda^i\mu^j(F(e_i) \mid F(e_j)) = \sum_i \lambda^i\mu^i = (v \mid w).\end{aligned}$$

11.4. Przekształcenia (operatory, odwzorowania) symetryczne.

DEFINICJA 11.16. *Operator $F\colon V \to V$ nazywamy* symetrycznym, *jeżeli dla $v, w \in V$ zachodzi równość*

$$(v \mid Fw) = (Fv \mid w).$$

Z definicji wynika natychmiast, że F jest symetryczny wtedy i tylko wtedy, gdy

$$\langle v, (F_g \circ F)w\rangle = (v \mid Fw) = (Fv \mid w) = \langle w, (F_g \circ F)v\rangle,$$

czyli wtedy i tylko wtedy, gdy odwzorowanie

$$(F_g \circ F)\colon V \longrightarrow V^*$$

jest samosprzężone. Z tego powodu operatory symetryczne nazywa się też samosprzężonymi. W przeciwieństwie do operatorów ortogonalnych, kombinacja liniowa operatorów symetrycznych jest operatorem symetrycznym. Tworzą one przestrzeń wektorową. Z kolei złożenie operatorów symetrycznych nie jest, na ogół, symetryczne.

11.4.1. Reprezentacja macierzowa operatorów symetrycznych. Ponieważ dla symetrycznego F odwzorowanie $F_g \circ F$ jest samosprzężone, więc z (9.2)

$$[F_g \circ F]^{e^*}{}_e = ([F_g \circ F]^{e^*}{}_e)^{\mathsf{T}}$$

i

$$[F_g]_e [F]^e{}_e = ([F]^e{}_e)^{\mathsf{T}} [F_g]_e$$

(macierz $[F_g]_e$ jest symetryczna). Jeżeli e jest bazą ortonormalną, to ostatnia równość przyjmuje postać

$$[F]^e{}_e = ([F]^e{}_e)^{\mathsf{T}}.$$

11.5. Operatory sprzężone.

Z poprzednich podrozdziałów widać, że użyteczne jest przyporządkowanie każdemu operatorowi $F\colon V \to V$ operatora $F^{\#}\colon V \to V$

$$F^{\#} = (F_g)^{-1} \circ F^* \circ F_g \tag{11.3}$$

lub, równoważnie,

$$(v \mid Fw) = \langle v, (F_g \circ F)w \rangle = \langle w, (F^* \circ F_g)v \rangle = (F^{\#}v \mid w), \tag{11.4}$$

bo F_g jest odwzorowaniem samosprzężonym. Operator $F^{\#} \in \mathsf{L}(V, V)$ nazywamy operatorem sprzężonym do operatora F. Należy tu odróżniać operator sprzężony od odwzorowania sprzężonego $F^* \in \mathsf{L}(V^*, V^*)$. Z (11.4) wynika, że
F jest ortogonalny wtedy i tylko wtedy, gdy $F^{\#} = F^{-1}$
oraz, że
F jest symetryczny wtedy i tylko wtedy, gdy $F^{\#} = F$.

Z kolei z (11.3) wynika, że przyporządkowanie

$$\mathsf{End}\ (V) \ni F \to F^{\#} \in \mathsf{End}\ (V)$$

jest izomorfizmem przestrzeni wektorowych. Mamy też reprezentację macierzową warunku (11.3):

$$[F^{\#}]^e{}_e = ([F_g]^{e^*}{}_e)^{-1} [F^*]^{e^*}{}_{e^*} [F_g]^{e^*}{}_e = ([g]_e)^{-1} ([F]^e{}_e)^{\mathsf{T}} [g]_e.$$

Jeżeli e jest bazą ortonormalną, czyli $[g]_e = I$, to dostajemy

$$[F^{\#}]^e{}_e = ([F]^e{}_e)^{\mathsf{T}}.$$

Rozdział 12. Miara układu wektorów

12.1. Macierz Grama układu wektorów.

DEFINICJA 12.1. Macierzą Grama *ciągu wektorów* $(v_1, \dots, v_k)$ *nazywamy macierz*

$$G(v_1, \dots, v_k) = [a^i{}_j],$$

gdzie $a^i{}_j = g(v_i, v_j) = (v_i \mid v_j)$

Spostrzeżenie:

$$\det G(v_1, \dots, v_k) \neq 0 \iff (v_1, \dots, v_k) \text{ jest liniowo niezależny}$$

Wzór na $\mathsf{d}(v, L)$. Niech L będzie podprzestrzenią rozpiętą przez $(v_1, \dots, v_k)$ oraz niech P będzie rzutem ortogonalnym na L. Dla $v \in V$ mamy

$$\mathsf{d}(v, L) = \|v - Pv\|, \ Pv = x^1 v_1 + \ldots + x^k v_k, \quad v - Pv \in L^\perp$$

to znaczy $(v_i|v) = (v_i|Pv)$ dla każdego i, czyli

$$\begin{aligned}(v_1 \mid v_1)x^1 + \cdots + (v_1 \mid v_k)x^k &= (v_1 \mid v)\\ &\vdots\\ (v_k \mid v_1)x^1 + \cdots + (v_k \mid v_k)x^k &= (v_k \mid v).\end{aligned}$$

Ponadto, jeżeli przez δ oznaczymy $\mathsf{d}(v, L)$, to

$$\delta^2 = (v - Pv \mid v - Pv) = (v \mid v - Pv) = (v \mid v) - (v \mid Pv) = (v \mid v) - \sum_i (v \mid v_i)x^i.$$

Rozpatrzmy teraz układ $k+1$ równań z $k+1$ niewiadomymi:

$$\begin{aligned}(v_1 \mid v_1)x^1 + \cdots + (v_1 \mid v_k)x^k + \qquad (v_1 \mid v)x^{k+1} &= 0\\ &\vdots\\ (v_k \mid v_1)x^1 + \cdots + (v_k \mid v_k)x^k + \qquad (v_k \mid v)x^{k+1} &= 0\\ (v \mid v_1)x^1 + \cdots + (v \mid v_k)x^k + ((v \mid v) - \delta^2)x^{k+1} &= 0\end{aligned}$$

Ma on niezerowe rozwiazanie $(x^{k+1} = -1)$ co oznacza, że

$$0 = \det \begin{bmatrix} & & (v \mid v_1) \\ & G(v_1, ..., v_k) & \vdots \\ (v \mid v_1) & \cdots & (v \mid v) - \delta^2 \end{bmatrix} =$$

$$= \det G(v_1, \dots, v_k, v) - \delta^2 \det G(v_1, \dots, v_k),$$

a zatem w przypadku, gdy $\det G(v_1, \dots, v_k) \neq 0$, to znaczy gdy $(v_1, \dots, v_k)$ tworzą bazę L,

$$\mathsf{d}(v, L) = \delta = \left(\frac{\det G(v_1, \dots, v_k, v)}{\det G(v_1, \dots, v_k)} \right)^{\frac{1}{2}}. \tag{12.1}$$

12.2. Miara (objętość) układu wektorów.

Miarę ciągu wektorów definiujemy indukcyjnie

DEFINICJA 12.2.

(1) miarą *wektora nazywamy jego długość:* $\mathsf{vol}(v) = \|v\|$.
(2) miarą (objętością) *ciągu wektorów* $(v_1, \dots, v_k, v)$ *nazywamy iloczyn:*

$$\mathsf{vol}(v_1 \dots, v_k, v) \stackrel{\text{def}}{=} \mathsf{d}(v, \langle \{v_1 \dots, v_k\} \rangle) \mathsf{vol}(v_1 \dots, v_k).$$

TWIERDZENIE 12.3. *Dla każdego ciągu wektorów* $(v_1, \dots, v_k)$ *prawdziwa jest równość:*

$$\mathsf{vol}(v_1, \dots, v_k) = (\det G(v_1 \dots, v_k))^{\frac{1}{2}}$$

Dowód indukcyjny: Dla $k = 1$ równość oczywista.
Z definicji i założenia indukcyjnego mamy że:

$$\begin{aligned} \mathsf{vol}(v_1 \dots, v_k) &= \mathsf{d}(v, \langle \{v_1 \dots, v_{k-1}\} \rangle) \mathsf{vol}(v_1 \dots, v_{k-1}) \\ &= \mathsf{d}(v, \langle \{v_1, \dots, v_{k-1}\} \rangle)(\det G(v_1, \dots, v_{k-1}))^{\frac{1}{2}}. \end{aligned}$$

Rozpatrzmy dwa przypadki

a) Jeżeli $\det G(v_1, \dots, v_{k-1}) = 0$ (więc $\mathsf{vol}(v_1, \dots, v_k) = 0$), to $(v_1, \dots, v_{k-1})$ jest układem liniowo zależnym, więc także $(v_1, \dots, v_{k-1}, v_k)$ jest liniowo zależny i $\det G(v_1, \dots, v_k) = 0$. Zatem

$$\mathsf{vol}(v_1 \dots, v_k) = 0 = (\det G(v_1 \dots, v_k))^{\frac{1}{2}}.$$

b) Jeśli $\det G(v_1, \dots, v_{k-1}) \neq 0$, to

$$\mathsf{d}(v_k, \langle \{v_1, \dots, v_{k-1}\} \rangle) = \left(\frac{\det G(v_1, \dots, v_{k-1}, v_k)}{\det G(v_1, \dots, v_{k-1})} \right)^{\frac{1}{2}}$$

i zatem

$$\mathsf{vol}(v_1, \dots, v_k) = (\det G(v_1, \dots, v_k))^{\frac{1}{2}}$$

■

Wnioski:

(1) Objętość układu $(v_1, \dots, v_k)$ nie zależy od uporządkowania wektorów.
(2) $\mathsf{vol}(v_1, \dots, v_k) = 0 \Longleftrightarrow$ układ $(v_1, \dots, v_k)$ jest liniowo zależny.

Spostrzeżenie: Niech $e = (e_1, \dots, e_n)$ będzie bazą w V. Macierz iloczynu skalarnego g w bazie e jest równa $G(e_1, \dots, e_n)$. Niech $(v_1, \dots, v_n)$ będzie układem wektorów, zaś macierz $A = [a^i{}_j]$ macierzą ich współrzędnych:

$$\overline{a}_j = [v_j]^e.$$

Mamy w tym przypadku

$$G(v_1, \dots, v_n) = [b^i{}_j],$$

gdzie

$$b^i{}_j = (v_i \mid v_j) = \sum_{k,l} a^k{}_i a^l{}_j (e_k \mid e_l) = \overline{a}_i{}^{\mathsf{T}} G(e_1, \dots, e_n) \overline{a}_j.$$

W konsekwencji otrzymujemy:

$$G(v_1, \dots, v_n) = A^{\mathsf{T}} G(e_1, \dots, e_n) A,$$

$$\det G(v_1, \dots, v_n) = \det A^2 \det G(e_1, \dots, e_n).$$

Niech w szczególności baza e będzie ortonormalna, wówczas

$$\det G(v_1, \dots, v_n) = A^2,$$

$$\mathsf{vol}(v_1, \dots, v_n) = |\det A| = |\det[[v_1]^e, \dots, [v_n]^e]|. \tag{12.2}$$

Przykład - pole trójkąta: Obliczamy pole trójkąta w $\mathbb{R}^2$ o wierzchołkach: $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2)$, $v_3 = (x_3, y_3)$.Mamy

$$|\triangle| = \frac{1}{2}\mathsf{vol}(v_2 - v_1, v_3 - v_1) = \left| \frac{1}{2} \det \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix} \right|$$

$$= \frac{1}{2} \left| \det \begin{bmatrix} 1 & 0 & 0 \\ 1 & x_2 - x_1 & y_2 - y_1 \\ 1 & x_3 - x_1 & y_3 - y_1 \end{bmatrix} \right| = \frac{1}{2} \left| \det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \right|.$$

Wzór (12.2) uzasadnia sens wprowadzenia następującego pojęcia.

DEFINICJA 12.4. Formą objętości zorientowanej *(objętością zorientowaną) na przestrzeni wektorowej* V, $\dim V = n$, *nazywamy odwzorowanie*

$$\omega\colon \underbrace{V \times \cdots \times V}_{n\ razy} \mapsto \mathbb{R}$$

takie, że:

(1) ω *jest liniowe w każdej zmiennej,*
(2) *jest antysymetryczne.*
Jeżeli ponadto
(3) *dla dowolnej bazy ortonormalnej e mamy* $\omega(e_1, \ldots, e_n) = \pm 1$, *to mówimy, że* ω *jest formą objętości zorientowanej* zgodną z iloczynem skalarnym.

Liczbę $\omega(v_1, \ldots, v_n)$ nazywamy *objętością zorientowaną* układu $(v_1, \ldots, v_n)$.

Spostrzeżenie: Przy ustalonym iloczynie skalarnym mamy dwie objętości zorientowane zgodne z tym iloczynem:

$$\omega_1(v_1, \ldots, v_n) = \det A$$

oraz

$$\omega_1(v_1, \ldots, v_n) = -\det A,$$

gdzie A jest, jak poprzednio, macierzą współrzędnych wektorów $v_1, \ldots, v_n$ w bazie ortonormalnej.

12.3. Iloczyn wektorowy.

Niech będą dane: V- przestrzeń wektorowa nad R, $g = (\ |\ )$- iloczyn skalarny w V i ω - forma objętości zorientowanej w V. Kanoniczne odwzorowanie

$$F_g\colon V \to V^*$$

jest izomorfizmem, więc jeżeli $(v_1, \ldots, v_{n-1})$ jest ciągiem wektorów, to element z przestrzeni V^* zdefiniowany przez

$$v \longmapsto \omega(v_1, \ldots, v_{n-1}, v) \in \mathbb{R}$$

jest postaci $F_g(w)$ dla dokładnie jednego wektora $w \in V$. Mamy zatem dla każdego $v \in V$

$$\omega(v_1, \ldots, v_{n-1}, v) = (v \mid w).$$

DEFINICJA 12.5. *Zdefiniowany powyżej wektor w nazywamy* iloczynem wektorowym *ciągu wektorów* $(v_1, \ldots, v_{n-1})$ *i oznaczamy*

$$v_1 \times v_2 \times \cdots \times v_{n-1}.$$

Spostrzeżenia:

(1) Iloczyn wektorowy jest skośnie symetryczny, tzn.,

$$v_1 \times \cdots \times v_i \times \cdots \times v_j \times \cdots \times v_{n-1} = -v_1 \times \cdots \times v_j \times \cdots \times v_i \times \cdots \times v_{n-1},$$
$$v_1 \times v_1 \times \cdots = 0.$$

(2) Współrzędne iloczynu wektorowego w bazie ortonormalnej $e = (e_1, \dots, e_n)$ są równe

$$w^i = (e_i \mid w) = \omega(v_1, \dots, v_{n-1}, e_i).$$

(3) Dla każdego i mamy

$$(v_i|w) = \omega(v_1, \dots, v_{n-1}, v_i) = 0,$$

więc

$$(v_1 \times \cdots \times v_{n-1}) \perp v_i.$$

(4) Z poprzedniego wynika, że rzut ortogonalny $v_1 \times \cdots \times v_{n-1}$ na podprzestrzeń $\langle\{v_1, \dots, v_{n-1}\}\rangle$ jest zerem, więc

$$\mathsf{d}(v_1 \times \ldots \times v_{n-1}, \langle\{v_1, \dots, v_{n-1}\}\rangle) = \|v_1 \times \cdots \times v_{n-1}\|.$$

Jeżeli zatem forma objętości ω jest zgodna z iloczynem skalarnym, to

$$\begin{aligned}\|v_1 \times \cdots \times v_{n-1}\|^2 &= (v_1 \times \cdots \times v_{n-1}|v_1 \times \cdots \times v_{n-1}) = \\ &= \omega(v_1, \dots, v_{n-1}, v_1 \times \cdots \times v_{n-1}) = \mathsf{vol}(v_1, \dots, v_{n-1}, v_1 \times \cdots \times v_{n-1}) = \\ &= \mathsf{d}(v_1 \times \cdots \times v_{n-1}, \langle\{v_1, \dots, v_{n-1}\}\rangle)\mathsf{vol}(v_1, \dots, v_{n-1}) = \\ &= \|v_1 \times \cdots \times v_{n-1}\|\mathsf{vol}(v_1, \dots, v_{n-1}).\end{aligned}$$

Stąd

$$\|v_1 \times \cdots \times v_{n-1}\| = \mathsf{vol}(v_1, \dots, v_{n-1}).$$

Przykład trójwymiarowy ($\dim V = 3$):
Niech (e_1, e_2, e_3) będzie bazą ortonormalną w V i niech $\omega(e_1, e_2, e_3) = 1$.

Weźmy $w = v_1 \times v_2$. Mamy, dla $v_i = v_i^1 e_1 + v_i^2 e_2 + v_i^3 e_3$,

$$w^1 = \omega(v_1, v_2, e_1) = \det \begin{bmatrix} 1 & 0 & 0 \\ v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \end{bmatrix}.$$

$$w^2 = \omega(v_1, v_2, e_2) = \det \begin{bmatrix} 0 & 1 & 0 \\ v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \end{bmatrix}.$$

$$w^3 = \omega(v_1, v_2, e_3) = \det \begin{bmatrix} 0 & 0 & 1 \\ v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \end{bmatrix}.$$

Używa się również skróconego zapisu

$$w = \begin{vmatrix} e_1 & e_2 & e_3 \\ {v^1}_1 & {v^2}_1 & {v^3}_1 \\ {v^1}_2 & {v^2}_2 & {v^3}_2 \end{vmatrix} .$$

Rozdział 13. Elementy geometrii afinicznej

13.1. Podprzestrzenie afiniczne.

W tym rozdziale V jest przestrzenią wektorową nad ciałem $\mathbb{K} = \mathbb{R}$ z iloczynem skalarnym g.

DEFINICJA 13.1. *Zbiór $M \subset V$ nazywamy* podprzestrzenią afiniczną, *jeżeli jest on postaci $M = a + M_0$, gdzie M_0 jest podprzestrzenią wektorową przestrzeni V, zaś a wektorem z V.*

STWIERDZENIE 13.2. *Zbiór $M \subset V$ jest podprzestrzenią afiniczną wtedy i tylko wtedy, gdy dla pewnego (i, w konsekwencji, dowolnego) $y \in M$ zbiór*

$$M_0 = \{w \in V \colon w = v - y,\ v \in M\}$$

jest podprzestrzenią wektorową.

DOWÓD: Jeżeli M jest podprzestrzenią afiniczną, to znaczy jest postaci $M = a + M_0$, gdzie M_0 jest podprzestrzenią wektorową, to dla dowolnego $M \ni y = x+a,\ x \in M_0$ mamy

$$\begin{aligned} M_0 &= \{w \in V \colon w = v - a,\ v \in M\} \\ &= \{w \in V \colon w + x = v - a,\ v \in M \\ &= \{w \in V \colon w = v - y = (v - a) - x,\ v \in V\}. \end{aligned}$$

Na odwrót, jeżeli $M \subset V$ i dla pewnego $y \in M$

$$M_0 = \{w \in V \colon w = v - y \quad v \in M\}$$

jest podprzestrzenią wektorową to, ponieważ $M = y + M_0$, M jest podprzestrzenią afiniczną. ■

Wniosek: Wektor a w definicji podprzestrzeni afinicznej M może być dowolnym elementem z M, M_0 jest takie samo dla wszystkich reprezentacji podprzestrzeni.

13.2. Sposoby opisu podprzestrzeni afinicznej.

(1) Opis parametryczny:
Tworzymy odwzorowanie $\Phi \colon \mathbb{R}^k \mapsto V$ takie, że Φ jest injekcją i że $\mathrm{im}\Phi = M$. Możemy się ograniczyć do Φ postaci $\Phi(x) = \Phi(0) + \Phi_0(x)$, gdzie $\Phi_0 \colon \mathbb{R}^k \mapsto V$ jest odwzorowaniem liniowym. W tym przypadku mówimy, że odwzorowanie Φ jest afiniczne. Powstaje pytanie: jak dla danej przestrzeni M znaleźć taki opis? Niech $M = y + M_0$ i niech $e_1, ..., e_k$ będzie bazą w przestrzeni M_0. Kładziemy po prostu $\Phi(x) = y + \sum x^i e_i$.

(2) Opis poprzez układ równań:
Szukamy odwzorowania liniowego $F\colon V \mapsto W$ takiego, by dla pewnego $b \in W$

$$W = \{x \in V\colon Fx = b\},$$

to znaczy, W jest zbiorem rozwiązań równania liniowego $Fx = b$.
Jak znaleźć odpowiednie F?
Niech $(f_1, \dots, f_{n-k})$ będzie bazą ${M_0}^{\perp}$. Mamy

$$\begin{aligned} M &= \{v \in V\colon v - a \in M_0\} = \{v \in V\colon (v - a \mid f_i) = 0\} \\ &= \{v \in V\colon (v \mid f_i) = (a \mid f_i)\}. \end{aligned}$$

Możemy więc położyć

$$F\colon V \mapsto \mathbb{R}^{n-k}\colon v \mapsto ((v \mid f_1), \dots, (v \mid f_{n-k}))$$

i

$$b = ((a \mid f_1), ..., (a \mid f_{n-k})).$$

Przykład: równanie prostej L w $\mathbb{R}^3$. Niech

$$L = \{(x, y, z)\colon (x, y, z) = (a, b, c) + t(m, n, l), t \in \mathbb{R}\},$$

czyli

$$x = a + tm, \;\; y = b + tn, \;\; z = c + tl.$$

Stąd, obliczając t z każdego z tych równań dostajemy, przy $m, n, l \neq 0$,

$$\frac{x-a}{m} = \frac{y-b}{n} = \frac{z-c}{l}.$$

Jeżeli np. $l = 0$, to drugim równaniem jest $z - c = 0$.

DEFINICJA 13.3. *Mówimy, że dwie podprzestrzenie afiniczne M, M' są* równoległe, *jeżeli $M_0 \subset M_0'$ lub $M_0' \subset M_0$.*

Równoległość podprzestrzeni oznaczamy $M \parallel M'$.

DEFINICJA 13.4. *Mówimy, że dwie podprzestrzenie afiniczne M, M' są* prostopadłe, *jeżeli $M_0' \subset {M_0}^{\perp}$.*

Prostopadłość podprzestrzeni oznaczamy $M \perp M'$.

13.3. Odstęp podprzestrzeni afinicznych.

DEFINICJA 13.5. Odstępem $\mathsf{d}(M, M')$ *między podprzestrzeniami afinicznymi M i M' nazywamy liczbę*

$$\mathsf{d}(M, M') = \inf_{\substack{m \in M \\ m' \in M'}} \|m - m'\| = \inf_{m \in M} \mathsf{d}(m, M')$$

STWIERDZENIE 13.6. *Jeżeli $M \parallel M'$, przy czym $M_0 \subset M_0'$, to $\mathsf{d}(m, M')$ jest takie samo dla wszystkich $m \in M$.*

DOWÓD: Niech $m \in M$, $m' \in M'$. Zatem $m = a + v'$ i $m' = a' + v'$ gdzie $v \in M_0$ oraz $v' \in M_0'$. Mamy więc

$$\mathsf{d}(m, m') = \|m - m'\| = \|(a - a') + (v - v')\|$$

i

$$\inf_{m' \in M'} \mathsf{d}(m, m') = \inf_{v' \in M_0'} \|(a - a') + (v - v')\| = \inf_{v' \in M_0'} \|(a - a') + v'\|,$$

co nie zależy od v, a więc nie zależy od m. ■

Przykłady:

(1) Objętość równoległościanu, którego krawędzie mają długości $1, 1, 2$, a kąty między krawędziami wynoszą $\frac{2}{3}\pi$, $\frac{5}{6}\pi$, $\frac{1}{3}\pi$. Możemy przyjąć, że równoległościan jest rozpięty na trzech wektorach:

$$v_1 = (1, 0, 0)$$
$$v_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$$
$$v_3 = (x, y, z).$$

Obliczamy x, y, z:

$$\cos(\frac{2}{3}\pi) = -\frac{1}{2} = \frac{(v_3 \mid v_1)}{\|v_3\|\|v_1\|} = \frac{1}{2}x$$

$$\cos(\frac{5}{6}\pi) = -\frac{\sqrt{3}}{2} = \frac{(v_2 \mid v_1)}{\|v_2\|\|v_1\|} = \frac{1}{2}(\frac{1}{2}x + \frac{\sqrt{3}}{2}y)$$

stąd:

$$x = -1$$
$$y = -2 + \frac{1}{\sqrt{3}}$$
$$x^2 + y^2 + z^2 = 1 + 4 + \frac{1}{3} - \frac{4}{\sqrt{3}} + z^2$$
$$z^2 = \frac{4}{3}(\sqrt{3} - 1), \quad z = \pm\sqrt{\frac{4}{3}(\sqrt{3} - 1)}$$

Szukana objętość jest równa

$$\left|\det\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -1 & -2+\frac{1}{\sqrt{3}} & \sqrt{\frac{4}{3}(\sqrt{3}-1)} \end{bmatrix}\right| = \sqrt{\sqrt{3}-1}.$$

(2) Liczymy odległość prostych skośnych, zadanych równaniami:

$$L_1 = \left\{\begin{array}{r} x+y-z-1=0 \\ 2x+y-z-2=0 \end{array}\right\}$$

i

$$L_2 = \left\{\begin{array}{r} x+2y-z-2=0 \\ x+2y+2z+4=0 \end{array}\right\}$$

Najpierw znajdujemy opis parametryczny prostych:

$$(1,1,-1)\times(2,1,-1) = \begin{vmatrix} i & j & k \\ 1 & 1 & -1 \\ 2 & 1 & -1 \end{vmatrix} = (0,-1,-1),$$

$$L_1 = (1,0,0)+t(0,1,1).$$

Podobnie $(1,2,-1)\times(1,2,2) = (6,-3,0)$ i

$$L_2 = (0,0,-2)+t(2,-1,0).$$

Płaszczyzna P równoległa do L_1 i L_2 jest opisywana przez równanie:

$$x+2y-2z+c=0.$$

Jeśli zażądamy, by $L_2 \subset P$, to dostaniemy $c=-4$. Odległość $\mathsf{d}(L_1,L_2)$ między prostymi jest równa odległości dowolnego punktu na prostej L_1 od płaszczyzny P. Wybierzmy jakiś punkt a na prostej L_1 i punkt m na P. Na przykład, niech $a=(1,0,0)$ $m=(0,0,-2)$. Odległość $\mathsf{d}(a,P)$ jest równa odległości $a-m=(1,0,2)$ od podprzestrzeni $P_0=\langle(0,1,1),(2,-1,0)\rangle$.

$$\mathsf{d}((a-m),P_0)^2 = \frac{G((1,0,2),(0,1,1),(2,-1,0))}{G(0,1,1),(2,-1,0))} = \frac{\begin{vmatrix} 0 & 1 & 1 \\ 2 & -1 & 0 \\ 1 & 0 & 2 \end{vmatrix}^2}{\begin{vmatrix} 2 & -1 \\ -1 & 5 \end{vmatrix}} = \frac{9}{9} = 1.$$

13.4. Podrozmaitości stopnia drugiego - kwadryki.

Funkcję $f: V \mapsto \mathbb{R}$ nazywamy funkcją stopnia drugiego, jeżeli jest ona postaci

$$f = f_0 + f_1 + f_2,$$

gdzie f_2 jest formą kwadratową, f_1 jest formą liniową, a f_0 jest funkcją stałą. Jeżeli $e = (e_1, ..., e_n)$ jest bazą w V i $e^* = (\phi_1, ..., \phi_n)$ jest bazą dualną w V^*, to funkcja stopnia drugiego zapisuje się w postaci

$$f(v) = \sum a_{ij}\phi_i(v)\phi_j(v) + \sum_i b_i\phi_i(v) + c.$$

W dalszym ciągu będziemy zakładać, że e jest bazą ortonormalną.

DEFINICJA 13.7. Podrozmaitość stopnia drugiego (kwadryka) *w V jest to podzbiór V postaci $f^{-1}(0)$ gdzie f jest funkcją stopnia drugiego na V.*

Dwie podrozmaitości stopnia drugiego będziemy uważać za równoważne, jeżeli jedną można otrzymać z drugiej przez przesunięcia, transformacje ortogonalne i ich złożenia, czyli *ruchy sztywne.* Niech C będzie podrozmaitością stopnia drugiego w V

$$C = f^{-1}(0).$$

Niech $F: V \to V$ będzie dowolną bijekcją, wówczas

$$\begin{aligned} F(C) &= \{V \ni v: v = Fw, w \in C\} = \{V \ni v: F^{-1}v \in C\} \\ &= \{V \ni v: f(F^{-1}v) = 0\} = \tilde{f}^{-1}(0) \end{aligned}$$

gdzie $\tilde{f} = f \circ F^{-1}\}$. Widać stąd, że jeśli F jest odwzorowaniem liniowym lub przesunięciem, to $F(C)$ jest też podrozmaitością stopnia drugiego. Będziemy starali się znaleźć wśród równoważnych podrozmaitości stopnia drugiego podrozmaitość charakteryzującą całą klasę.

I Krok - transformacja ortogonalna.

Niech $e' = (e'_1, ..., e'_n)$ będzie bazą ortonormalną w V, diagonalizującą formę $f^{(2)}$. Istnienie takiej bazy jest prostym wnioskiem z twierdzenia spektralnego dla operatorów symatrycznych, o którym będzie mowa w następnym rozdziale. Ponieważ baza e jest ortonormalna, odwzorowanie $F: V \to V$ zdefiniowane przez $F: e'_i \mapsto e_i$ jest ortogonalne. Oznaczmy $\tilde{f} = f \circ F^{-1}$. Mamy

$$\begin{aligned} \tilde{f}(v) &= f \circ F^{-1}(v) = \sum_{i=1} n\lambda_i(\phi'_i)^2(F^{-1}v) + \cdots \\ &= \sum_{i=1}^{n} \lambda_i{\phi_i}^2 + \sum_{i=1}^{n} \tilde{b}_i\phi_i + c \end{aligned}$$

i stąd każda podrozmaitość drugiego stopnia jest równoważna podrozmaitości opisywanej funkcją drugiego stopnia o diagonalnej składowej kwadratowej.

II Krok - przesunięcia.

Niech $\tau\colon V \to V$ będzie przesunięciem o wektor a:

$$\tau(v) = v + a \qquad \tau^{-1}(v) = v - a$$

i niech f ma diagonalną część kwadratową. Mamy

$$\tilde{f}(v) = f \circ \tau^{-1}(v) = \sum_{i=1}^{n} \lambda_i(\phi_i(v-a))^2 + \sum b_i\phi_i(v-a) + c$$

$$= \sum_{i=1}^{n} \lambda_i{\phi_i}^2(v) - \sum_{i=1}^{n} \lambda_i 2\phi_i(v)\phi(a) + \sum_{i=1}^{n} \lambda_i{\phi_i}^2(a) + \sum b_i\phi_i(v) - \sum b_i\phi_i(a) + c.$$

Permutując bazę możemy dostać, że $\lambda_i \neq 0$ dla $i = 1, \dots, r$ i $\lambda_i = 0$ dla $i = r+1, \dots, n$. Wybieramy a takie, że: $-2\lambda_i\phi_i(a) + b_i = 0$ dla $i = 1, \dots, r$. Jest to możliwe dzięki $\lambda_i \neq 0$. W efekcie w klasie równoważnych podrozmaitości stopnia drugiego mamy podrozmaitość opisaną funkcją postaci

$$f = \sum_{i=1}^{r} \lambda_i{\phi_i}^2 + \sum_{i=r+1}^{n} b_i\phi_i + c.$$

III krok - znów odwzorowanie ortogonalne, ale nie ruszające $\phi_1, \dots, \phi_r$

$$f \circ F^{-1} = \sum_{i=1}^{r} \lambda_i{\phi_i}^2 + f_1 \circ F^{-1} + c$$

Można znaleźć takie F, że $f_1 \circ F^{-1} = \mu\phi_{r+1}$. Zatem w klasie równoważnych podrozmaitości stopnia drugiego mamy podrozmaitość opisaną funkcją postaci

$$f = \sum_{i=1}^{r} \lambda_i{\phi_i}^2 + \mu\phi_{r+1} + c.$$

Ponieważ dla $c \neq 0$ $cf = 0 \Leftrightarrow f = 0$, są dwie możliwości:

(1) $\mu = 0$ wtedy

$$f = \sum_{i=1}^{r} \lambda_i{\phi_i}^2 + 1$$

lub

$$f = \sum_{i=1}^{r} \lambda_i{\phi_i}^2.$$

(2) $\mu \neq 0$ wtedy

$$f = \sum_{i=1}^{r} \lambda_i {\phi_i}^2 + \phi_{r+1}$$

(c wyeliminowaliśmy wykonując jeszcze jedno przesunięcie).

13.4.1. Klasyfikacja kwadryk. Z poprzedniej części wynika, że w klasie równoważnych podrozmaitości stopnia drugiego mamy podrozmaitość opisaną funkcją postaci

$$f = \sum_{i=1}^{r} \lambda_i {\phi_i}^2 + 1,$$

lub

$$f = \sum_{i=1}^{r} \lambda_i {\phi_i}^2,$$

lub

$$f = \sum_{i=1}^{r} \lambda_i {\phi_i}^2 + \phi_{r+1}.$$

Klasa równoważnych podrozmaitości jest więc scharakteryzowana przez zespół liczb

$$\{\{\lambda_1, \ldots, \lambda_r\}; 0, 1\}$$

lub

$$\{\{\lambda_1, \ldots, \lambda_r\}; \mu, 0\}$$

gdzie $\mu = 0, 1$. Jest jeszcze (niewielka) niejednoznaczność w tej charakterystyce. Na przykład $\{\{\lambda_1, \lambda_2\}; 1, 0\}$ jest równoważne $\{\{-\lambda_1, -\lambda_2\}; 1, 0\}$. Poniżej podane są przykłady powierzchni z najważniejszych klas w przypadku $n = 2, 3$. Zespół liczb charakteryzujący klasę zawiera jedynie znaki λ_i.

n=2

$f(x,y) = (\frac{x}{2})^2 + y^2 - 1;$ typ $\{\{-,-\};0,1\}$ – Elipsa

$f(x,y) = x^2 - y^2;$ typ $\{\{+,-\};0,0\}$ – Przecinające się proste

$$f(x,y) = (\frac{10}{7}y)^2 - x^2 - 1;\ \text{typ}\ \{\{+,-\};0,1\}\ \text{– Hiperbola}$$

$$f(x,y) = (2y)^2 - 1;\ \ \text{typ}\ \{\{-\};0,1\}\ \text{– Para prostych równoległych}$$

$f(x,y) = x^2 - y$; typ $\{\{-\};1,0\}$ – Parabola n=3

Typ $\{\{+,+,-\};0,0\}$ – Stożek

Typ $\{\{-,-,-\};0,1\}$ – Elipsoida

Typ $\{\{+,+,-\};0,1\}$ – Hiperboloida dwupowłokowa

Typ $\{\{+,-,-\};0,1\}$ – Hiperboloida jednopowłokowa

Typ $\{\{-,-\};0,1\}$ – Walec eliptyczny

Typ $\{\{+,+\};1,0\}$ – Paraboloida eliptyczna

Typ $\{\{+,-\};1,0\}$ – Paraboloida hiperboliczna

Rozdział 14. Przestrzenie unitarne. Twierdzenia spektralne

14.1. Odwzorowania antyliniowe.

W tym rozdziale K=C!!

Niech V, W będą przestrzeniami wektorowymi nad $\mathbb{C}$. Odwzorowanie $F: V \to W$ nazywamy *antyliniowym*, jeżeli

$$F(\lambda v + \mu w) = \overline{\lambda} F(v) + \overline{\mu} F(w).$$

STWIERDZENIE 14.1.

(1) *Jeżeli* $F: V \to W$, $G: W \to U$ *są antyliniowe, to* $G \circ F$ *jest liniowe.*
(2) *Jeżeli jedno z odwzorowań* $F: V \to W$, $G: W \to U$ *jest liniowe, a drugie jest antyliniowe, to* $G \circ F$ *jest antyliniowe.*

Zbiór odwzorowań antyliniowych z V do W tworzy przestrzeń wektorową, którą oznaczać będziemy $\mathsf{AL}(V, W)$. Przestrzeń $\mathsf{AL}(V, \mathbb{C})$ fukcjonałów antyliniowych oznaczać będziemy $V^{\#}$. Niech $F: V \to W$ będzie odwzorowaniem liniowym. Odwzorowanie

$$W^{\#} \to V^{\#}: f \mapsto f \circ F$$

jest, jak łatwo zauważyć, liniowe. Nazywać je będziemy odwzorowaniem *antydualnym* lub *hermitowsko sprzężonym* do F i oznaczać będziemy $F^{\#}$.

DEFINICJA 14.2. Sprzężeniem zespolonym *funkcjonału liniowego (antyliniowego)* f *na* V *nazywamy funkcjonał antyliniowy (liniowy)* $\overline{f}$ *na* V, *zdefiniowany przez*

$$\overline{f}(v) = \overline{f(v)}.$$

Sprzężenie zespolone daje kanoniczny antyizomorfizm V^* i $V^{\#}$. W konsekwencji daje antyizomorfizm (t. j. antyliniowy izomorfizm) przestrzeni $\mathsf{L}(V^*, W^*)$ i $\mathsf{L}(V^{\#}, W^{\#})$, zdefiniowany wzorem

$$\overline{F}(f) = \overline{F(\overline{f})}.$$

Niech $F \in \mathsf{L}(V, W)$. Oczywistym jest związek

$$F^{\#} = \overline{F}^*.$$

STWIERDZENIE 14.3.

(1) *Istnieje kanoniczny izomorfizm przestrzeni* $(V^{\#})^*$ *i* $(V^*)^{\#}$.
(2) *Istnieje kanoniczny antyizomorfizm przestrzeni* $(V^*)^{\#}$ *i* V.
(3) *Istnieje kanoniczny antyizomorfizm przestrzeni* $(V^{\#})^*$ *i* V.
(4) *Istnieje kanoniczny izomorfizm przestrzeni* $(V^{\#})^{\#}$ *i* V.

DOWÓD:

(1) Niech $\varphi \in (V^*)^\#$. Wzór
$$V^\# \ni f \mapsto \overline{\varphi}(\overline{f}) \in \mathbb{C}$$
określa, jak łatwo sprawdzić, odwzorowanie liniowe, więc element z $(V^\#)^*$. Oznaczmy go przez $\widetilde{\varphi}$. Przyporządkowanie $\varphi \mapsto \widetilde{\varphi}$ jest izomorfizmem przestrzeni wektorowych. (Co się zmieni, jeżeli w powyższym wzorze $\overline{\varphi}$ zastąpimy przez φ?)

(2) Sprzężenie zespolone daje antyizomorfizm $(V^*)^*$ i $(V^*)^\#$. Sprzężenie zespolone złożone z kanonicznym izomorfizmem V i $(V^*)^*$ daje żądany antyizomorfizm.

(3) Otrzymujemy przez złożenie antyizomorfizmu z punktu drugiego z izomorfizmem z punktu pierwszego.

(4) Otrzymujemy przez złożenie antyizomorfizmu z punktu poprzedniego ze sprzężeniem zespolonym. ■

Jako proste ćwiczenie zostawiamy dowód następującego stwiedzenia.

STWIERDZENIE 14.4.

(1) $(F+G)^\# = F^\# + G^\#$,
(2) $(\lambda F)^\# = \overline{\lambda} F^\#$,
(3) $(G \circ F)^\# = F^\# \circ G^\#$,
(4) $(F^\#)^\# = F$.

Niech $F\colon V \to V^\#$ będzie odwzorowaniem liniowym. Zatem $F^\# = \overline{F}^*\colon V = (V^\#)^\# \to V^\#$. Odwzorowanie F nazywamy *hermitowsko samosprzężone* (h.s.s.), jeżeli $F^\# = F$ (przy utożsamieniu $(V^\#)^\# = V$).

14.1.1. Reprezentacja macierzowa. Niech e będzie bazą w V i f bazą w W. Odwzorowania antyliniowe $F\colon V \to W$ można, podobnie jak odwzorowania liniowe, reprezentować macierzami z tym, że, ponieważ $F(\lambda^1 e_1 + \dots + \lambda^n e_n) = \bar{\lambda}_1 F(e_1) + \dots + \bar{\lambda}_n F(e_n)$, mamy
$$[Fv]^f = [F]^f{}_e \overline{[v]^e}, \tag{14.1}$$
$\overline{[\cdot]}$ oznacza tu sprzężenie zespolone elementów macierzowych.

Niech $e^* = (e_1^*, \dots, e_n^*))$ będzie bazą w V^*, dualną do bazy e. Oczywistym jest, że ciąg $e^\# = (\overline{e_1^*}, \dots, \overline{e_n^*})$ tworzy bazę (dualną) w $V^\#$. Mamy dla $\varphi \in V^*$
$$\overline{[\varphi]^{e^*}} = [\overline{\varphi}]^{e^\#}.$$
Niech $F\colon V \to W$ będzie odwzorowaniem liniowym i niech f będzie bazą w W. Łatwo dostajemy, że
$$[F^\#]^{e^\#}{}_{f^\#} = \overline{([F]^f{}_e)^\mathsf{T}}. \tag{14.2}$$

DEFINICJA 14.5. Sprzężeniem hermitowskim *macierzy* A *nazywamy macierz* $A^\dagger = \overline{A}^{\mathsf{T}}$. *Macierz* A *nazywamy* hermitowską, *jeżeli* $A^\dagger = A$.

Niech będzie dane odwzorowanie liniowe $F\colon V \to \overline{V}^* = V^\#$. F jest hermitowsko samosprzężone wtedy i tylko wtedy, gdy macierz $[F]^{e^\#}{}_e$ jest hermitowska.

14.2. Przestrzenie unitarne.

DEFINICJA 14.6. Iloczynem skalarnym *w przestrzeni wektorowej* V *nad* $\mathbb{C}$ *nazywamy odwzorowanie:*

$$h\colon V \times V \to \mathbb{C}$$

takie, że dla $v, w \in V$ *i* $\alpha \in \mathbb{C}$ *mamy*

(1) $h(v, \alpha w) = \alpha h(v, w)$,
(2) $h(v, w + w') = h(v, w) + h(v, w')$,
(3) $h(v, w) = \overline{h(w, v)}$,
(4) $h(v, v) > 0$ *dla* $v \neq 0$ $(h(v, v) \in \mathbb{R})$.

Mówimy, że h *jest* formą półtoraliniową. *Przestrzeń z iloczynem skalarnym nazywamy* przestrzenią unitarną.

Przypomnijmy, że odwzorowanie $F\colon V \to W$ takie, że $F(v + w) = F(v) + F(w)$ i $F(\lambda v) = \bar{\lambda} F(v)$ nazywamy antyliniowym. Forma h jest zatem liniowa ze względu na drugi argument i antyliniowa ze względu na pierwszy argument. Czy możemy odtworzyć h mając odpowiednią formę kwadratową $v \mapsto h(v, v)$? Mamy

$$h(v + w, v + w) = h(v, v) + h(w, w) + h(v, w) + h(v, w),$$

czyli

$$\frac{1}{2}(h(v + w, v + w) - h(v, v) - h(w, w)) = \operatorname{Re} h(v, w)$$

i, z drugiej strony,

$$\begin{aligned} \operatorname{Im} h(v, w) &= -\operatorname{Re}(\imath h(v, w)) = \operatorname{Re} h(\imath v, w) \\ &= \frac{1}{2}(h(\imath v + w, \imath v + w) - h(\imath v, \imath v) - h(w, w)). \end{aligned}$$

Stąd dostajemy odpowiednio zmodyfikowaną formułę polaryzacyjną:

$$h(v, w) = \frac{1}{2}\{h(v+w, v+w) + \imath h(\imath v+w, \imath v+w) - (1+\imath)(h(v, v) + h(w, w))\}. \quad (14.3)$$

Spostrzeżenie: $\operatorname{Im} h$ jest formą R-dwuliniową, antysymetryczną i niezdegenerowaną. $\operatorname{Re} h$ jest formą R-dwuliniową, symetryczną i niezdegenerowaną, więc zdaje na V strukturę przestrzeni euklidesowej.

Tak jak w przypadku rzeczywistym

(1) $h(v, w)$ oznaczamy $(v \mid w)$
(2) $(h(v, v))^{\frac{1}{2}}$ będziemy oznaczać $\|v\|$ i nazywać *normą* (*długością*) wektora v.

Przykłady:

(1) $\mathbb{C}^n$ z iloczynem skalarnym

$$(x \mid y) = \overline{x_1} y_1 + \cdots + \overline{x_n} y_n.$$

(2) Przestrzeń wielomianów $\mathbb{C}[n]$ z iloczynem skalarnym

$$(f \mid g) = \int_a^b \bar{f} g \ .$$

Podstawowe własności iloczynu skalarnego w przestrzeniach euklidesowych przenoszą się na przestrzenie unitarne. Dla zilustrowania komplikacji wynikających z przejścia od liczb rzeczywistych do zespolonych podamy dowód nierówności Schwarza

TWIERDZENIE 14.7 (Nierówność Schwarza). *Jeśli V jest przestrzenią wektorową z iloczynem skalarnym, to:*

$$| (v \mid w) | \leqslant \|v\| \, \|w\|. \tag{14.4}$$

Równość zachodzi wtedy i tylko wtedy, gdy v i w są liniowo zależne.

Dowód: Jeśli $v = \mathbf{0}$, to twierdzenie jest trywialne.
Jeśli $v \neq \mathbf{0}$, to rozpatrzmy funkcję $\alpha\colon \mathbb{R} \ni t \mapsto \|tv + w\|^2 \in \mathbb{R}$, czyli

$$\alpha(t) = t^2 (v \mid v) + 2t \operatorname{Re}(v \mid w) + (v \mid w).$$

Oczywiście $\alpha(t) \geqslant 0$, zatem wyróżnik tego trójmianu jest niedodatni, tzn.,

$$(\operatorname{Re}(v \mid w))^2 - (\|v\| \, \|w\|)^2 \leqslant 0. \tag{14.5}$$

Dla pewnego rzeczywistego φ mamy

$$(v \mid w) = e^{\imath\varphi} |(v \mid w)|,$$

zatem

$$| (v \mid w) | = e^{-\imath\varphi} (v \mid w) = (e^{\imath\varphi} v \mid w) = \operatorname{Re}(e^{\imath\varphi} v \mid w).$$

Stąd i z (14.5) dostajemy

$$| (v \mid w) |^2 \leqslant \|e^{i\varphi} v\|^2 \, \|w\|^2 = \|v\|^2 \|w\|^2.$$

Jeżeli $w = \lambda v$, to

$$|(v \mid w)| = |\lambda| \|v\|^2 = \|\lambda v\| \, \|v\| = \|v\| \, \|w\|.$$

Niech teraz $|(v \mid w)| = \|v\|\|w\|$ i $|(v \mid w)| = e^{-\iota\varphi}(v \mid w)$. Rozpatrzmy funkcję

$$\begin{aligned}\beta\colon t \mapsto \beta(t) = \|e^{\iota\varphi} t v + w\|^2 = t^2\|v\|^2 + 2t \mid (v \mid w) \mid + \|w\|^2 = \\ = t^2\|v\|^2 + 2t\|v\|\|w\| + \|w\|^2 = (t\|v\| + \|w\|)^2.\end{aligned}$$

β jest równe zero dla $t_0 = -\frac{\|w\|}{\|v\|}$, czyli $w = (e^{\iota\varphi}\frac{\|w\|}{\|v\|})v$. ■

Poniższa wersja zespolona twierdzenia o rzucie ortogonalnym zostanie udowodniona w sposób istotnie różny od wersji rzeczywistej. Przedstawiony dowód nie korzysta z faktu, że przestrzeń jest skończenie wymiarowa.

TWIERDZENIE 14.8 (O RZUCIE PROSTOPADŁYM). *Niech V będzie przestrzenią wektorową nad $\mathbb{C}$ z iloczynem skalarnym h i niech $W \subset V$ będzie podprzestrzenią wektorową. Wówczas*

$$V = W \oplus W^{\perp}.$$

DOWÓD: Jeżeli $v \in W \cap W^{\perp}$, to $(v \mid v) = 0$, czyli $v = 0$.
Czy $V = W + W^{\perp}$?
Rozpatrzmy funkcję

$$V \ni v \mapsto \alpha(v) = \inf_{y \in W} \|v - y\|^2.$$

Z definicji inf istnieje ciąg $(y_n \in W)$ taki, że $\alpha(v) = \lim_{n\to\infty} \|v - y_n\|^2$. Zatem możemy przyjąć, że

$$\alpha(v) \leqslant \|v - y_n\|^2 \leqslant \alpha(v) + \frac{1}{n}.$$

Stąd

$$\begin{aligned}\|y_n - y_m\|^2 &= \|(v - y_m) - (v - y_n)\|^2 = \\ &= 2(\|v - y_m\|^2 + \|v - y_n\|^2) - \|2v - (y_m + y_n)\|^2 \\ &\leqslant 4\alpha + 2(\frac{1}{n} + \frac{1}{m}) - 4\alpha = 2(\frac{1}{n} + \frac{1}{m}) \to 0\end{aligned}$$

Zatem ciąg y_n jest ciągiem Cauchy. Ponieważ podprzestrzeń W jest domknięta[1] jego granica $P(v)$ należy do W i $\alpha(v) = \|v - P(v)\|$.

Pokażemy, że $v - P(v) \in W^{\perp}$.
Niech $y \in W$, wówczas funkcja

$$f(t) \mapsto \|(v - P(v)) + ty\|^2$$

[1] W przypadku nieskończonego wymiaru jest to dodatkowe założenie obok założenia zupełności przestrzeni V.

ma minimum w $t = 0$. Z drugiej strony

$$f(t) = \|v - P(v)\|^2 + 2t\ \mathrm{Re}(y \mid v - P(v)) + t^2\|y\|^2,$$

czyli

$$0 = f'\ (0) = 2\,\mathrm{Re}(y \mid v - P(v)).$$

Ponadto

$$\mathrm{Im}(y \mid v - P(v)) = \mathrm{Re}(\imath y \mid v - P(v)) = 0,$$

bo $\imath y \in W$. Oznacza to, że $v - P(v) \in W^{\perp}$. ■

Podobnie jak w przypadku przestrzeni euklidesowych sformułujemy twierdzenie o reprezentacji funkcjonałów liniowych. Ponieważ iloczyn skalarny h jest formą półtoraliniową, odpowiadają mu dwa odwzorowania $F_h\colon V \to V^*$ i ${}_hF\colon V \to V^{\#}$ zdefiniowane przez

$$\begin{aligned} \langle w, F_h(v)\rangle &= (v \mid w) \\ \langle w, {}_hF(v)\rangle &= (w \mid v). \end{aligned} \tag{14.6}$$

F_h jest odwzorowaniem antyliniowym i, w konsekwencji, $\mathbb{R}$–liniowym (to znaczy liniowym, jeżeli przestrzenie argumentów i wartości traktowane są jako przestrzenie wektorowe nad $\mathbb{R}$). Natomiast odwzorowanie ${}_hF$ jest liniowe i, jak łatwo zauważyć, hermitowsko samosprzężone. Jeżeli $v \in \ker F_h$, to $F_h(v) = 0$ i

$$0 = \langle v, F_h(v)\rangle = h(v, v) = \|v\|^2.$$

Stąd $v = \mathbf{0}$. Ponieważ wymiary przestrzeni V i V^* nad ciałem liczb rzeczywistych są równe oznacza to, że F_h jest bijekcją. Mamy więc

TWIERDZENIE 14.9 (O REPREZENTACJI FUNKCJONAŁU LINIOWEGO). *Dla każdego funkcjonału $f \in V^*$ istnieje dokładnie jeden wektor $w_f \in V$ taki, że dla każdego wektora $v \in V$ zachodzi równość*

$$\langle v, f\rangle = (w_f \mid v).$$

Uwaga: w odróżnieniu od przypadku euklidesowego istotny jest porządek wektorów w iloczynie skalarnym.

14.3. Operatory w przestrzeni unitarnej.

14.3.1. Sprzężenie hermitowskie. Niech (V, g) będzie przestrzenią unitarną i niech $F \in \mathsf{End}(V)$. Dla każdego $w \in V$ odwzorowanie

$$V \ni v \mapsto (w|Fv) \in \mathbb{C}$$

jest liniowe, zatem na mocy twierdzenia o reprezentacji funkcjonału liniowego istnieje $\widetilde{w} \in V$ takie, że $(w|Fv) = (\widetilde{w}|v)$ dla każdego v. Oznaczmy $\widetilde{w} = F^{\dagger}w$. Jak łatwo zauważyć, odwzorowanie $F^{\dagger}$ jest liniowe. Nazywamy je *sprzężeniem hermitowskim* operatora F.

Uwaga. Sprzężenie hermitowskie operatora F możemy zdefiniować używając odwzorowania[2] $F_h\colon V \to V^*$. Mianowicie, $F^{\dagger} = F_h^{-1} \circ F^* \circ F_h$.

[2] Lub ${}_hF\colon V \to V^{\#}$, $F^{\dagger} = {}_hF^{-1} \circ F^{\#} \circ {}_hF$.

STWIERDZENIE 14.10. *Dla* $F, G \in \mathsf{End}(V)$, $\lambda \in \mathbb{C}$ *mamy*

(1) $(F^\dagger)^\dagger = F$,
(2) $(F + G)^\dagger = F^\dagger + G^\dagger$,
(3) $(\lambda F)^\dagger = \overline{\lambda} F^\dagger$,
(4) $(F \circ G)^\dagger = G^\dagger \circ F^\dagger$.

Dowód:

(1) Dla $v, w \in V$ mamy

$$(w|Fv) = (F^\dagger w|v) = \overline{(v|F^\dagger w)} = \overline{((F^\dagger)^\dagger v|w)} = (w|(F^\dagger)^\dagger v).$$

Stąd $Fv = (F^\dagger)^\dagger v$ i $F = (F^\dagger)^\dagger$.
(2) Oczywiste.
(3) Mamy

$$((\lambda F)^\dagger w|v) = (w|(\lambda F)v) = (w|F(\lambda v)) = (F^\dagger w|\lambda v) = (\overline{\lambda} F^\dagger w|v).$$

Stąd $(\lambda F)^\dagger = \overline{\lambda} F^\dagger$.
(4) $((F \circ G)^\dagger w|v) = (w|F \circ Gv) = (F^\dagger w|Gv) = (G^\dagger \circ F^\dagger w|v)$. ∎

Wniosek. Przyporządkowanie $F \mapsto F^\dagger$ jest antyliniowym izomorfizmem w przestrzeni $\mathsf{End}(V)$.

DEFINICJA 14.11. *Operator* $F\colon V \to V$ *nazywamy* hermitowskim, *jeżeli* $F = F^\dagger$, *to znaczy dla wszystkich* $v, w \in V$ $(w|Fv) = (Fw|v)$.

Łatwo zauważyć, że jeżeli operatory F, G są hermitowskie, to $F + G$ też jest hermitowski, a $F \circ G$ na ogół nie jest hermitowski. Z kolei λF jest hermitowski wtedy i tylko wtedy, gdy $\lambda \in \mathbb{R}$.

DEFINICJA 14.12. *Operator* $F\colon V \to V$ *nazywamy* unitarnym, *jeżeli dla wszystkich* $v, w \in V$ *mamy* $(Fv|Fw) = (v|w)$.

STWIERDZENIE 14.13.

(1) *F jest unitarny wtedy i tylko wtedy, gdy dla wszystkich $v \in V$ mamy $(Fv|Fv) = (v|v)$.*
(2) *F jest unitarny wtedy i tylko wtedy, gdy $F^\dagger F = Id$.*
(3) *Jeżeli F, G są unitarne, to $F \circ G$ też jest unitarny.*

Dowód:

(1) Wynika z formuły polaryzacyjnej (14.3).
(2) Mamy $(v|w) = (Fv|Fw) = (v|F^\dagger Fw)$. Stąd $F^\dagger F = Id$.
(3) $(v|w) = (Gv|Gw) = (F(Gv)|F(Gw))$. ∎

Łatwo zauważyć, że suma operatorów unitarnych na ogół nie jest operatorem unitarnym. Z kolei dla F unitarnego λF jest unitarny wtedy i tylko wtedy, gdy $|\lambda| = 1$.

Teraz kilka słów o operatorach rzutowych. Operator rzutowy P (tzn. taki, że $P^2 = P$) jest *operatorem rzutu ortogonalnego*, jeżeli $\ker P = (\operatorname{im} P)^{\perp}$, to znaczy, przestrzeń wzdłuż której się rzutuje jest prostopadła do przestrzeni na którą się rzutuje.

STWIERDZENIE 14.14. *$P\colon V \to V$ jest operatorem rzutu ortogonalnego wtedy i tylko wtedy, gdy $P^2 = P$ i $P = P^{\dagger}$.*

Dowód: Niech $P^2 = P$ i $P^{\dagger} = P$. Równość $P^2 = P$ oznacza, że P jest operatorem rzutowym. Mamy pokazać, że $(\operatorname{im} P)^{\perp} = \ker P$. Wymiary obu przestrzeni są równe, więc wystarczy udowodnić zawieranie w jedną stronę. Niech $v \in \ker P$ i niech $w \in \operatorname{im} P$, tzn., $w = Pw'$. Wówczas, ponieważ $P = P^{\dagger}$,

$$(v|w) = (v|Pw') = (P^{\dagger}v|w') = (Pv|w) = 0.$$

Załóżmy teraz, że P jest operatorem rzutu ortogonalnego na W. Niech $v, v_1 \in V$ i niech $v = w + w', v_1 = w_1 + w_1'$ gdzie $w, w_1 \in W$ a $w', w_1' \in W^{\perp}$. Mamy wówczas

$$(v|Pv_1) = (v|w_1) = (w|w_1) = (w|w_1 + w_1') = (Pv|v_1).$$

■

Na koniec ciekawostka:

STWIERDZENIE 14.15. *Jeżeli $F\colon V \to V$ jest jednocześnie unitarny i hermitowski, to $\frac{1}{2}(Id - F)$ jest operatorem rzutu ortogonalnego.*

Dowód: Zauważmy, że $\frac{1}{2}(Id - F)$ jest hermitowski. Ponadto

$$(\frac{1}{2}(Id - F))^2 = \frac{1}{4}(Id - 2F + F^2) = \frac{1}{4}(Id - 2F + F^{\dagger}F) = \frac{1}{2}(Id - F).$$

Na mocy poprzedniego stwierdzenia dostajemy tezę. ■

14.3.2. Reprezentacja macierzowa. Niech e będzie bazą w V. Macierz iloczynu skalarnego definiujemy jako macierz $[h]_e = [{}_hF]^{e^{\#}}{}_e$ lub, równoważnie, jako macierz $\overline{[F_h]^{e^*}{}_e}$. Elementy macierzowe daną są wzorem $a^i{}_j = (e_j \mid e_i)$. Iloczyn skalarny zapisuje się więc

$$(v|w) = ([v]^e)^{\dagger}[h]_e[w]^e.$$

Oczywistym jest, że macierz $[h]_e$ jest hermitowska. Niech teraz $F \in \mathsf{End}(V)$. Mamy

$$(v|Fw) = ([v]^e)^{\dagger}[h]_e[F]^e{}_e[w]^e$$

i

$$(F^\dagger v|w) = ([F^\dagger]^e{}_e[v]^e)^\dagger [h]_e[w]^e = ([v]^e)^\dagger ([F^\dagger]^e{}_e)^\dagger [h]_e[w]^e$$

co, po porównaniu, daje (uwzględniając hermitowskość macierzy $[h]_e$)

$$[F^\dagger]^e{}_e = [h]_e^{-1}([F]^e{}_e)^\dagger [h]_e.$$

Jeżeli baza e jest ortonormalna (definicja jak w przypadku euklidesowym, dowód istnienia nieco zmodyfikowaną metodą Grama-Schmidta), to $[h]_e = I$ i

$$[F^\dagger]^e{}_e = ([F]^e{}_e)^\dagger.$$

DEFINICJA 14.16. *Macierz* $A \in \mathbf{M}^n{}_n(\mathbb{C})$ *nazywamy* unitarną *jeżeli* $A^\dagger A = I$, *tzn* $A^\dagger = A^{-1}$.

Macierz odwzorowania unitarnego w bazie ortonormalnej jest więc unitarna.

14.4. Twierdzenie spektralne dla operatorów normalnych.

DEFINICJA 14.17. $F\colon V \longmapsto V$ *nazywamy* normalnym *jeżeli* $FF^\dagger = F^\dagger F$.

Przykłady: operatory hermitowskie, unitarne są operatorami normalnymi.

STWIERDZENIE 14.18. *Operator F jest normalny wtedy i tylko wtedy, gdy* $\|Fv\| = \|F^\dagger v\|$ *dla każdego* $v \in V$.

DOWÓD: Jeżeli operator F jest normalny, to

$$(Fv \mid Fv) = (v \mid F^\dagger F v) = (v \mid FF^\dagger v) = (F^\dagger v \mid F^\dagger v).$$

Jeśli $(Fv \mid Fv) = (F^\dagger v \mid F^\dagger v)$, to z formuły polaryzacyjnej (14.3)

$$(Fv \mid Fw) = (F^\dagger v \mid F^\dagger w)$$

i

$$(v \mid F^\dagger F w) = (v \mid FF^\dagger w).$$

■

STWIERDZENIE 14.19 (PODSTAWOWE). *Jeżeli F jest operatorem normalnym, to*

$$\ker F = \ker F^\dagger.$$

DOWÓD: $v \in \ker F$ wtedy i tylko wtedy, gdy $0 = (Fv \mid Fv) = (F^\dagger v \mid F^\dagger v)$ tzn., gdy $v \in \ker F^\dagger$. ■

Z twierdzenia 8.8 o rozkładzie na podprzestrzenie pierwiastkowe mamy $V = \oplus V_i$, gdzie V_i jest niezmienniczą podprzestrzenią pierwiastkową odpowiadającą wartosci własnej λ_i. Niech $L_i \subset V_i$ będą podprzestrzeniami wektorów własnych (podprzestrzenie te są też niezmiennicze).

LEMAT 14.20. *Jeżeli operator F jest normalny i $w \in L_i$, to w jest wektorem własnym operatora $F^\dagger$ dla wartości własnej $\overline{\lambda_i}$.*

DOWÓD: Niech będzie $w \in \ker(F - \lambda_i I)$. Operator $(F - \lambda_i I)$ jest też normalny, więc

$$w \in \ker(F - \lambda_i I)^\dagger = \ker(F^\dagger - \overline{\lambda_i} I).$$

∎

LEMAT 14.21. *Jeżeli F jest operatorem normalnym, to $L_i^\perp$ jest podprzestrzenią F-niezmienniczą.*

DOWÓD: Niech $v \in L_i^\perp$, wówczas dla $w \in L_i$ mamy

$$(w \mid Fv) = (F^\dagger w \mid v) = \lambda_i (w \mid v) = 0.$$

Zatem $Fv \in L_i^\perp$. ∎

LEMAT 14.22. *Jeżeli operator F jest normalny oraz $i \neq j$, to $L_i \perp L_j$.*

DOWÓD: Niech $v \in L_i$ oraz $w \in L_j$, $i \neq j$, wówczas z Lematu 14.20

$$\lambda_j (v \mid w) = (v \mid Fw) = (F^\dagger v \mid w) = (\overline{\lambda_i} v \mid w) = \lambda_i (v \mid w).$$

Stąd

$$(\lambda_j - \lambda_i)(v \mid w) = 0,$$

czyli $(v \mid w) = 0$.

TWIERDZENIE 14.23 (SPEKTRALNE - WERSJA PIERWSZA). *Jeżeli operator F jest normalny, to V posiada bazę ortonormalną wektorów własnych F.*

DOWÓD: Wybierzmy w każdej z przestrzeni L_i bazę ortonormalną. Wektory tych baz, razem wzięte, tworzą, na mocy Lematu 3, zbiór wektorów ortonormalnych. Wystarczy więc pokazać, że $L_i = V_i$.

Oznaczmy $L = \oplus_i L_i$. L jest podprzestrzenią rozpiętą przez wszystkie wektory własne. Mamy $L^\perp = \bigcap_i {L_i}^\perp$, więc, ponieważ ${L_i}^\perp$ są F-niezmiennicze, także $L^\perp$ jest niezmiennicze względem F. Mamy zatem odwzorowanie $F \colon L^\perp \mapsto L^\perp$. W $L^\perp$ istnieje więc wektor własny operatora F, albo $L^\perp = \{0\}$. Ponieważ wszystkie wektory własne F są w L mamy $L^\perp = \{0\}$. Stąd $V = L$, czyli $L_i = V_i$. ∎

Twierdzenie odwrotne do powyższego jest też prawdziwe (udowodnić!):
Jeśli F posiada ortonormalną bazę wektorów własnych to F jest normalny.

TWIERDZENIE 14.24 (SPEKTRALNE - WERSJA DRUGA). *Operator F jest normalny wtedy i tylko wtedy, gdy F jest postaci $F = \sum_i \alpha_i P_i$, gdzie P_i sa operatorami rzutu ortogonalnego takimi, że $P_i P_j = 0$ dla $i \neq j$ i $\sum P_i = Id_V$.*

Dowód: Niech P_i będzie rzutem ortogonalnym na V_i. Ponieważ $v \in V_i$ jest wektorem własnym mamy $Fv = \lambda_i v$, gdzie λ_i jest odpowiednią wartością własną. Stąd dla dowolnego $v \in V$ mamy $v = \sum_i P_i v$, zatem

$$Fv = F \sum_i P_i v = \lambda_i P_i v.$$

Na odwrót, jeśli $F = \sum \alpha_i P_i$, gdzie $P_i P_j = \delta_{ij} P_i$, $P_i^\dagger = P_i$ i $\sum P_i = Id_V$, to

$$F^\dagger F = \sum |\alpha_i|^2 P_i = FF^\dagger.$$

Operator F jest więc normalny. ■

Reprezentacja $F = \sum \lambda_i P_i$, gdzie $\lambda_i \neq \lambda_j$ dla $i \neq j$, $P_i P_j = \delta_{ij} P_i$, $P_i^\dagger = P_i$ i $\sum P_i = Id_V$ nazywa się *rozkładem spektralnym* operatora F. Jest on jednoznaczny.

Czy można rozpoznać operatory hermitowskie i unitarne po ich własnościach spektralnych?

STWIERDZENIE 14.25. *Operator F jest hermitowski wtedy i tylko wtedy, gdy*

$$F = \sum \lambda_i P_i, \quad \lambda_i \in \mathbb{R},$$

gdzie P_i są jak w drugiej wersji twierdzenia spektralnego.

Dowód: Niech F będzie operatorem hermitowskim. Ponieważ operator hermitowski jest też normalny, to $F = \sum \lambda_i P_i$. Ale dla dowolnego v

$$\overline{\lambda_i}(P_i v \mid P_i v) = (FP_i v \mid P_i v) = (P_i v \mid FP_i v) = \lambda_i (P_i v \mid P_i v),$$

skąd $\lambda_i = \overline{\lambda_i}$.

Niech teraz $F = \sum \lambda_i P_i$ oraz $\lambda_i \in \mathbb{R}$. Mamy

$$F^\dagger = \sum \overline{\lambda_i} P_i = \sum \lambda_i P_i = F.$$

■

STWIERDZENIE 14.26. *F jest operatorem unitarnym wtedy i tylko wtedy, gdy w rozkładzie spektralnym $F = \sum \lambda_i P_i$ mamy $|\lambda_i| = 1$.*

Dowód: Jeżeli F jest unitarny, to $F^\dagger F = \sum |\lambda_i|^2 P_i = Id_V$. Ale $Id_V = \sum_i P_i$ i stąd $P_i = |\lambda_i|^2 P_i$ i $|\lambda_i| = 1$.

Na odwrót, niech $|\lambda_i| = 1$, wtedy

$$F^\dagger F = \sum_i |\lambda_i|^2 P_i = \sum P_i = Id_V.$$ ■

14.4.1. Twierdzenie spektralne dla operatora symetrycznego.
Niech V będzie przestrzenią wektorową nad $\mathbb{R}$ z metryką g. Formę dwuliniową g można jednoznacznie przedłużyć do formy półtoraliniowej h na kompleksyfikacji $V_{\mathbb{C}}$ przestrzeni V.

$$h(v + \iota w, v' + \iota w') = g(v, v') + g(w, w') + \iota(g(v, w') - g(w, v')).$$

Oczywiste, że h jest iloczynem skalarnym. Niech F będzie operatorem symetrycznym na V. Jego kompleksyfikacja $F_{\mathbb{C}}$ jest operatorem hermitowskim:

$$\begin{aligned}(v + \iota w \mid F_{\mathbb{C}}(v' + \iota w')) &= g(v, Fv') + g(w, Fw') + \iota(g(v, Fw') - g(w, Fv')) = \\ &= g(Fv, v) + g(Fw, w') + \iota(g(Fv, w') - g(Fw, v')) = (F_{\mathbb{C}}(v + \iota w) \mid v' + \iota w').\end{aligned}$$

Operator hermitowski ma rzeczywiste wartości własne, więc z podrozdziału 8.6. dostajemy, że podprzestrzeń pierwiastkowa operatora $F_{\mathbb{C}}$ jest kompleksyfikacją odpowiedniej podprzestrzeni pierwiastkowej operatora F. Dla operatora hermitowskiego wektor podprzestrzeni pierwiastkowej jest wektorm własnym, więc i wektor podprzestrzeni pierwiastkowej operatora F jest wektorem własnym. Podprzestrzenie pierwiastkowe różnych wektorów własnych są do siebie ortogonalne, więc dostajemy rozkład spektralny operatora symetrycznego

$$F = \sum_i \lambda_i P_i,$$

gdzie $\lambda_i \in \mathbb{R}$, a P_i są operatorami rzutu ortogonalnego.

Uwaga. Podobnie jak i w przypadku unitarnym, operator P jest operatorem rzutu ortogonalnego wtedy i tylko wtedy, gdy $P^2 = P$ i P jest operatorem symetrycznym.

Jako przykład rozpatrzmy operator F_φ odpowiadający formie kwadratowej φ na przestrzeni V. Operator F_φ jest symetryczny, więc posiada bazę ortonormalną wektorów własnych. Baza ta jest bazą diagonalizującą formę φ. Z istnienia takiej bazy korzystaliśmy przy klasyfikacji kwadryk.

14.4.2. Twierdzenie Eulera. Jako przykład twierdzenia opartego na analizie własności rozkładu spektralnego posłuży

TWIERDZENIE 14.27 (EULERA). *W $\mathbb{R}^3$ dowolne odwzorowanie ortogonalne nie zmieniające orientacji jest obrotem.*

Odwzorowanie $F: \mathbb{R}^3 \mapsto \mathbb{R}^3$ rozszerzamy do odwzorowania $\mathbb{C}$-liniowego $\widetilde{F}: \mathbb{C}^3 \mapsto \mathbb{C}^3$. Jezeli F jest ortogonalne, to $\widetilde{F}$ jest unitarne. $\widetilde{F}^\dagger = \widetilde{F}^{-1}$, więc $\det \widetilde{F} = \det F = \pm 1$. F nie zmienia orientacji, więc $\det \widetilde{F} = 1$. Niech λ_1, λ_2, λ_3 będą wartościami własnymi $\widetilde{F}$. Ponieważ wielomian charakterystyczny $\widetilde{F}$ ma współczynniki rzeczywiste, co najmniej jedna z wartości własnych jest rzeczywista. Niech będzie to na przykład λ_1. Dla λ_2, λ_3 mamy teraz następujące możliwości:

(1) λ_2 nie jest rzeczywiste, wtedy $\lambda_3 = \overline{\lambda_2}$ i $\det F = \lambda_1 \lambda_2 \overline{\lambda_2} = \lambda_1 = 1$

(2) λ_2 jest rzeczywiste, więc λ_3 też jest rzeczywiste. Zatem λ_2 i λ_3 są równe ± 1. Ponieważ $\lambda_1\lambda_2\lambda_3 = 1$ jeden z pierwiastków jest równy 1 (np. λ_1) a pozostałe dwa są sobie równe. Zatem $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = \pm 1$.

Niech L_1 będzie podprzestrzenią w $\mathbb{C}^3$ złożoną z wektorów własnych wartości własnej $\lambda_1 = 1$. Z powyższych rozważań $\dim L_1 = 1$ lub 3. Jeśli $\dim L_1 = 3$, to $F = Id$. Jeśli $\dim L_1 = 1$, to podprzestrzeń ${L_1}^{\perp}$ jest niezmiennicza, 2-wymiarowa. Indukowane odwzorowanie $\widetilde{F}\colon L_1^{\perp} \mapsto L_1^{\perp}$ jest unitarne. Mamy oczywiście, $L_1^{\perp} = L + iL$, $L \subset \mathbb{R}^3$, $\dim L = 2$ i $\widetilde{F}\mid_L = F\mid_L \colon L \to L$ jest ortogonalne.

W bazie ortonormalnej $F|_L$ jest reprezentowane przez macierz postaci

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Jest ona ortogonalna, więc

$$ab + cd = 0$$

$$a^2 + c^2 = b^2 + d^2 = 1.$$

Stąd istnieją kąty ϕ i ψ takie, że $a = \cos\phi$, $b = \sin\phi$, $c = -\sin\psi$, $d = \cos\psi$. Uzupełnijmy bazę w $L_1^{\perp}$ o wektor z L_1 do bazy e. Mamy

$$[F]^e{}_e = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{bmatrix}$$

Wyznacznik $[F]^e{}_e$ jest równy 1, więc

$$ad - bc = 1$$

i dostajemy $\sin(\phi - \psi) = 0$, $\cos(\phi - \psi) = 1$ stąd $\phi = \psi$. Ostatecznie mamy

$$[F]^e{}_e = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix}$$

F jest więc obrotem wokół osi L_1.

14.5. Rozkład biegunowy.

Niech $F\colon V \to V$ będzie operatorem hermitowskim. Wtedy, dla dowolnego $v \in V$,

$$(v \mid Fv) = (Fv \mid v) = \overline{(v \mid Fv)} \in \mathbb{R}.$$

Wprowadźmy nastepujące pojęcia:

DEFINITION 14.28. *F nazywamy operatorem* dodatnim, *jeżeli dla każdego wektora v spełniona jest nierówność:*

$$(v \mid Fv) \geqslant 0,$$

F nazywamy ściśle dodatnim, *jeśli dla każdego wektora $v \neq \mathbf{0}$ zachodzi:*

$$(v \mid Fv) > 0.$$

STWIERDZENIE 14.29. *Operator F jest dodatni, jeżeli jego wartości własne są nieujemne.*

Operator F jest ściśle dodatni, jeżeli jego wartości własne są dodatnie.

DOWÓD: Niech $F = \sum \lambda_i P_i$ będzie rozkładem spektralnym operatora F. Mamy wówczas
$$(v \mid Fv) = \sum_i \lambda_i (P_i v \mid P_i v) = \sum_i \lambda_i \|P_i v\|^2.$$

Stąd natychmiast wynika teza. ■

STWIERDZENIE 14.30. *Jeżeli operator F jest dodatni, to istnieje dokładnie jeden hermitowski, dodatni operator G taki, że $G^2 = F$.*

DOWÓD: Niech rozkład $G = \sum \alpha_i R_i$, $\alpha_i > 0$, będzie rozkładem spektralnym operatora G. Stąd $G^2 = \sum {\alpha_i}^2 R_i$, ale jeśli $\alpha_i \neq \alpha_j$, to ${\alpha_i}^2 \neq {\alpha_j}^2$, zatem rozkład ten jest rozkładem spektralnym G^2.

Niech $F = \sum \lambda_j P_j$ będzie rozkładem spektralnym operatora F. Ponieważ chcemy mieć $F = G^2$, z jednoznaczności rozkładu spektralnego wynika, że dla każdego i $P_i = R_i$. Wystarczy zatem wziąć $\alpha_i = \sqrt{\lambda_i}$. ■

Operator G z powyższego stwierdzenia oznaczamy $F^{\frac{1}{2}}$.

TWIERDZENIE 14.31 (O ROZKŁADZIE BIEGUNOWYM). *Każdy operator $F \colon V \to V$ da się przedstawić w postaci $F = UR$, gdzie U jest operatorem unitarnym, zaś R dodatnim i hermitowskim. Jeżeli F jest odwracalny, to rozkład jest jednoznaczny.*

DOWÓD: Jeśli $F = UR$ jest rozkładem, o którym mówi twierdzenie, to $F^\dagger = R^\dagger U^\dagger = RU^{-1}$. Wynika stąd, że $F^\dagger F = R^2$. Operator $F^\dagger F$ jest dodatni i hermitowski ponieważ

$$(v \mid F^\dagger F v) = (F^\dagger F v \mid v) = (Fv \mid Fv) \geqslant 0.$$

Istnienie R wynika więc z poprzedniego stwierdzenia:

$$R = (F^\dagger F)^{\frac{1}{2}}.$$

Dostajemy teraz równanie na U:

$$F = UR.$$

Zauważmy najpierw, że, ponieważ $(Rv \mid Rv) = (Fv \mid Fv)$, to $\ker R = \ker F$. R jest hermitowski, zatem $\operatorname{im} R = (\ker R)^{\perp} = (\ker F)^{\perp}$. Oznaczmy przez $\widetilde{R}$ indukowane przez R odwzorowanie $\widetilde{R}\colon (\ker F)^{\perp} \to (\ker F)^{\perp}$ a przez $\widetilde{F}$ indukowane przez F odwzorowanie $\widetilde{F}\colon (\ker F)^{\perp} \to \operatorname{im} F$. Zarówno $\widetilde{F}$ jak i $\widetilde{R}$ są izomorfizmami, więc odwzorowanie

$$\widetilde{U} = \widetilde{F}\widetilde{R}^{-1}\colon (\ker F)^{\perp} \to \operatorname{im} F$$

jest dobrze określone. Zachowuje ono iloczyn skalarny: dla $v, w \in (\ker F)^{\perp}$ mamy

$$(\widetilde{F}\widetilde{R}^{-1}v \mid \widetilde{F}\widetilde{R}^{-1}w) = (F\widetilde{R}^{-1}v \mid F\widetilde{R}^{-1}w) = (F^{\dagger}F\widetilde{R}^{-1}v \mid \widetilde{R}^{-1}w)$$

$$= (Rv \mid \widetilde{R}^{-1}w) = (v \mid R\widetilde{R}^{-1}w) = (v \mid w).$$

Ponieważ wymiary przestrzeni $\ker F$ i $\operatorname{im} F^{\perp}$ są równe możemy uzupełnić $\widetilde{U}$ do odwzorowania unitarnego

$$U\colon V \to V.$$

Jeśli F jest operatorem odwracalnym, to $U = \widetilde{U} = F(F^{\dagger}F)^{-\frac{1}{2}}$. Stąd jednoznaczność rozkładu. ■

W analogiczny sposób dostajemy dwoisty rozkład biegunowy

$$F = R'U'.$$

Rozdział 15. Operacje elementarne

Rozdział ten zawdzięczamy uprzejmości Grzegorza Cieciury.

15.1. Definicja operacji elementarnych.

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$. W zbiorze $V^n = V\times\cdots\times V$ n–elementowych układów $(v_1, \ldots, v_n)$ rozważać będziemy następujące trzy rodzaje operacji (zwane *operacjami elementarnymi*):

1^0 Przestawianie danych wektorów:

$$(v_1, \ldots, v_n) \mapsto (v_{\sigma(1)}, \ldots, v_{\sigma(n)}), \quad \sigma \in \mathsf{S}(n).$$

2^0 Pomnożenie poszczególnych wektorów przez dowolne $\neq 0$ liczby $\lambda_i \in \mathbb{K}$:

$$(v_1, \ldots, v_n) \mapsto (\lambda^1 v_1, \ldots, \lambda^n v_n).$$

3^0 Dodanie krotności jednego z wektorów układu (np. wektora v_i) do pozostałych wektorów układu:

$$(v_1, \ldots, v_i, \ldots, v_n) \mapsto (v_1 + \lambda^1 v_i, \ldots, v_i, \ldots, v_n + \lambda^n v_i)$$

(liczby $\lambda^1, \ldots, \lambda^{i-1}, \lambda^{i+1}, \ldots, \lambda^n) \in \mathbb{K}$ są tutaj dowolne.

FAKT. *Każda operacja elementarna jest odwracalna, przy czym jej odwrotność jest operacją elementarną tego samego typu ($1^0, 2^0$ lub 3^0).*

Dowód: Gdy $(w_1, \ldots, w_n) = (v_{\sigma(1)}, \ldots, v_{\sigma(n)})$, wtedy

$$(v_1, \ldots, v_n) = (w_{\sigma^{-1}(1)}, \ldots, w_{\sigma^{-1}(n)}).$$

Gdy $(w_1, \ldots, w_n) = (\lambda^1 v_1, \ldots, \lambda^n v_n)$, wtedy

$$(v_1, \ldots, v_n) = ((\lambda^1)^{-1} w_1, \ldots, (\lambda^n)^{-1} w_n).$$

Gdy wreszcie $(w_1, \ldots, w_n) = (v_1 + \lambda^1 v_i, \ldots, v_i, \ldots, v_n + \lambda^n v_i)$, wtedy

$$(v_1, \ldots, v_n) = (w_1 - \lambda^1 w_i, \ldots, w_i, \ldots, w_n - \lambda^n w_i).$$

■

DEFINICJA. *Dwa układy $(v_1, \ldots, v_n)$ i $(w_1, \ldots, w_n)$ nazywamy* równoważnymi, *pisząc*

$$(v_1, \ldots, v_n) \sim (w_1, \ldots, w_n),$$

jeżeli układ $(w_1, \ldots, w_n)$ można otrzymać z $(v_1, \ldots, v_n)$ stosując pewną liczbę operacji elementarnych.

Przykład (n=3): $(v_1, v_2, v_3) \sim (9v_2 + v_3, 4v_3 + 3v_1, -2v_1 + 3v_2)$ gdyż

$$\begin{aligned}(v_1, v_2, v_3) &\xmapsto{3^0} (v_1, v_2 - \frac{2}{3}v_1, v_3 + \frac{3}{4}v_1)\\ &\xmapsto{3^0} (v_1 + 2v_2 - \frac{4}{3}v_1, v_2 - \frac{2}{3}v_1, v_3 + \frac{3}{4}v_1)\\ &\xmapsto{3^0} (-\frac{1}{3}v_1 + 2v_2 + \frac{4}{9}(v_3 + \frac{3}{4}v_1), v_2 - \frac{2}{3}v_1, v_3 + \frac{3}{4}v_1)\\ &\xmapsto{2^0} (9v_2 + 2v_3, 3v_2 - 2v_1, 4v_3 + 3v_1)\\ &\xmapsto{1^0} (9v_2 + 2v_3, 4v_3 + 3v_1, -2v_1 + 3v_2).\end{aligned}$$

FAKT. *Określona powyżej relacja $\sim$ jest relacją równoważności w V^n.*

DOWÓD: Zwrotność i przechodniość jest oczywista, natomiast symetria wynika z poprzedniego Faktu: Odwrotnością złożenia $E_1 \circ E_2 \circ \cdots \circ E_k$ operacji elementarnych jest złożenie $E_1^{-1} \circ E_2^{-1} \circ \cdots \circ E_k^{-1}$ również operacji elementarnych. ■

FAKT. *Jeśli $(v_1, \ldots, v_n) \sim (w_1, \ldots, w_n)$ to*

a) *$(v_1, \ldots, v_n$ są liniowo niezależne$) \iff (w_1, \ldots, w_n$ są liniowo niezależne,$)$*
b) *$< v_1, \ldots, v_n >=< w_1, \ldots, w_n >$.*

DOWÓD: Wystarczy sprawdzić a) i b) w przypadku, gdy $(w_1, \ldots, w_n)$ powstaje z $(v_1, \ldots, v_n)$ przez zastosowanie jednej operacji elementarnej, a jest to banalnie prostym ćwiczeniem. ■

Uwaga. Można także wykazać (lecz jest to znacznie trudniejsze i nie będziemy z tego korzystać), że

$$(v_1, \ldots, v_n) \sim (w_1, \ldots, w_n) \iff < v_1, \ldots, v_n >=< w_1, \ldots, w_n > .$$

15.2. Operacje elementarne na kolumnach i wierszach macierzy.

Będziemy odtąd zajmować się przypadkiem, gdy $V = \mathbf{M}^m{}_1(\mathbb{K})$ (przestrzeń macierzy jednokolumnowych) lub $V = \mathbf{M}^1{}_n(\mathbb{K})$ (przestrzeń macierzy jednowierszowych. W takim przypadku wygodnie jest układ $(v_1, \ldots, v_n)$ wektorów utożsamić z macierzą A, gdzie $A \in \mathbf{M}^m{}_n(\mathbb{K})$ gdy $v_i \in \mathbf{M}^m{}_1(\mathbb{K})$. Kolumny macierzy A zdefiniowane są następująco:

$$\bar{a}_1 = v_1, \ldots, \bar{a}_n = v_n.$$

Gdy $v_i \in \mathbf{M}^1{}_m(\mathbb{K})$ wiersze macierzy $A \in \mathbf{M}^n{}_m(\mathbb{K})$ definiujemy analogicznie:

$$\bar{a}^1 = v_1, \ldots, \bar{a}^n = v_n.$$

DEFINICJA.

a *Dwie macierze* $A, B \in \mathbf{M}^m{}_n(\mathbb{K})$ *są* kolumnowo równoważne, $A \underset{k}{\sim} B$, *jeżeli*

$$(\bar{a}_1, \ldots, \bar{a}_n) \sim (\bar{b}_1, \ldots, \bar{b}_n).$$

b *Dwie macierze* $A, B \in \mathbf{M}^n{}_m(\mathbb{K})$ *są* wierszowo równoważne, $A \underset{k}{\sim} B$, *jeżeli*

$$(\bar{a}^1, \ldots, \bar{a}^n) \sim (\bar{b}^1, \ldots, \bar{b}^n).$$

Na przykład:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 1 & 2-2 \\ 3 & 4-6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & -2 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 1-3\cdot 0 & 0 \\ 3-3\cdot 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \underset{w}{\sim} \begin{bmatrix} 2 & 3 \\ 4-2\cdot 2 & 5-2\cdot 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix} \underset{w}{\sim} \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix},$$

a więc

$$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \underset{w}{\sim} \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

DEFINICJA. *Jeśli* $A \in \mathbf{M}^n{}_m(\mathbb{K})$, *to*

$$\operatorname{im} A := \langle \{\bar{a}_1, \ldots, \bar{a}_m\} \rangle \subset \mathbf{M}^n{}_1(\mathbb{K})$$

(podprzestrzeń rozpięta przez kolumny macierzy A) nazywamy obrazem macierzy,

$$\ker A := \{x \in \mathbf{M}^m{}_1(\mathbb{K}) \colon Ax = 0\}$$

(przestrzeń rozwiązań równania $Ax = 0$*) nazywamy* jądrem macirzy.

FAKT.

$$A \underset{k}{\sim} B \Longrightarrow \operatorname{im} A = \operatorname{im} B$$

$$A \underset{w}{\sim} B \Longrightarrow \ker A = \ker B$$

DOWÓD: Wynika natychmiast z punktu (b) poprzedniego Faktu. ■

Okazuje się, że w wielu sytuacjach szczególnie wygodne są tzw. *macierze zredukowane* (kolumnowo i wierszowo) i, co więcej, że z każdej macierzy A można, stosując operacje elementarne na jej kolumnach (wierszach) otrzymać równoważną jej macierz kolumnowo (wierszowo) zredukowaną.

15.3. Definicje i terminologia.

Macierz kwadratową $M \in \mathbf{M}^r{}_r(\mathbb{K})$ nazywamy *macierzą diagonalną* lub D-macierzą, jeśli ma zerowe wyrazy poza główną przekątną, tzn. jeśli $m^i{}_j = 0$ dla $i \neq j$. M nazywamy D^*-macierzą, jeżeli jest diagonalna i ma różne od zera wyrazy na diagonali, tzn. jeśli $m^i{}_j = 0$ dla $i \neq j$ oraz $m^i{}_i \neq 0$

M nazywamy PD^*-macierzą, jeśli przez stosowne przestawienie jej kolumn (lub– równoważnie–wierszy) można z niej otrzymać pewną D^*-macierz, tzn. jeśli $\exists \sigma \in \mathsf{S}(r)$ takie, że $m^i{}_j \neq 0 \Leftrightarrow j = \sigma(i)$. Na przykład macierze

$$\begin{bmatrix} 0 & 5 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 2 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & -5 \\ 0 & 3 & 0 \end{bmatrix}$$

są PD^*- macierzami.

Jeżeli $A \in \mathbf{M}^n{}_m(\mathbb{K})$, to *podmacierzą* macierzy A nazywamy taką macierz, która jest utworzona z wyrazów macierzy A, położonych w pewnych (dowolnie) wybranych wierszach macierzy A. Na przykład macierz $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ ma następujące podmacierze wymiaru 2×2:

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

DEFINICJA. *$A \in \mathbf{M}^n{}_m(\mathbb{K})$ nazywa się kolumnowo zredukowaną (krócej - KZ–macierzą), jeżeli A zawiera pewną podmacierz typu PD^* wymiaru $r \times r$, gdzie $r =$ (liczba $\neq 0$ kolumn macierzy} A. Analogicznie definiujemy macierze wierszowo zredukowane (WZ-macierze) (tym razem $r =$ liczbie różnych od zera wierszy macierzy A).*

Na przykład

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 0 & 5 \\ 2 & 0 & 6 \\ 3 & 0 & 0 \end{bmatrix} \text{ jest KZ-macierzą}$$

gdyż ma podmacierz $\begin{bmatrix} 0 & 5 \\ 3 & 0 \end{bmatrix}$ typu PD^* ,

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix} \text{ jest WZ-macierzą}$$

gdyż ma podmacierz $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ typu PD^*.

FAKT.

> *Niezerowe kolumny KZ–macierzy są liniowo niezależne.*
> *Niezerowe wiersze WZ–macierzy są liniowo niezależne.*

DOWÓD: PD^*–macierz ma liniowo niezależne kolumny i liniowo niezależne wiersze. ■

TWIERDZENIE. *Z dowolnej macierzy A można otrzymać, stosując operacje elementarne typu 3^0 na jej kolumnach, równoważną jej kolumnowo KZ–macierz. Analogicznie, każda macierz jest wierszowo równoważna pewnej WZ–macierzy.*

15.4. Algorytm redukcji kolumnowych.

(zamiast dowodu formalnego) Jeśli $a^i{}_j \neq 0$ jest ustalonym wyrazem macierzy A (nazywać go będziemy *wyrazem bazowym*), to *redukcja kolumnowa* względem $a^i{}_j$ jest operacją elementarną typu 3^0, wyspecyfikowaną następująco:

do poszczególnych kolumn $\bar{a}_k$ dla $k \neq j$ dodajemy takie krotności kolumny $\bar{a}_j$, które dają zerowe i-te współrzędne tych nowych kolumn

Na przykład dla $\begin{bmatrix} 2 & 3 & 4 \\ \boxed{1} & 2 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ redukcja kolumnowa względem wyrazu bazowego $a^2{}_1 = 1$ produkuje z macierzy A macierz

$$\begin{bmatrix} 2 & 3-2\cdot 2 & 4-5\cdot 2 \\ 1 & 2-2\cdot 1 & 5-5\cdot 1 \\ 3 & 5-2\cdot 3 & 6-5\cdot 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \\ 3 & -1 & -9 \end{bmatrix}.$$

A oto cały ALGORYTM:

Przeprowadzamy redukcje kolumnowe względem kolejno wybieranych wyrazów bazowych,kierując się przy ich wyborze następującymi zasadami:

1^0 nie mogą powtarzać się numery kolumn wyrazów bazowych.
2^0 nie mogą powtarzać się numery wierszy wyrazów bazowych.

Dodatkową (już nie obligatoryjną) zasadą jest wybór takich wyrazów bazowych przez które łatwo jest dzielić, gdyż jak łatwo zauważyć, redukcja odbywa się według schematu

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \boxed{a} & \cdots & b & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & c & \cdots & d & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \Longrightarrow \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & a & \cdots & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & c & \cdots & d-\frac{bc}{a} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

(ramką $\Box$ zaznaczamy zawsze wyraz bazowy). Procedura kończy się w chwili, gdy nie jest możliwy wybór nowego wyrazu bazowego, zgodny z zasadami 1^0 i 2^0 – wtedy otrzymana macierz jest już KZ–macierzą.

Przykład.

$$\begin{bmatrix} 2 & 3 & 3 & 4 \\ 3 & 5 & 5 & 6 \\ \boxed{1} & 2 & 1 & 1 \\ 2 & 1 & -5 & -2 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 2 & -1 & \boxed{1} & 2 \\ 3 & -1 & 2 & 3 \\ 1 & 0 & 0 & 0 \\ 2 & -3 & -7 & -4 \end{bmatrix} \underset{k}{\sim}$$

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & \boxed{1} & 2 & -1 \\ 1 & 0 & 0 & 0 \\ 16 & -10 & -7 & 10 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 6 & -10 & 13 & 0 \end{bmatrix} \text{ i już!}$$

Dwa sposoby opisu podprzestrzeni $V \subset \mathbb{K}^m$. Opis typu (W): Przez podanie wektorów rozpinających przestrzeni V, np.

$$V = \left\langle \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \right\rangle = \left\{ \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} u_1 + \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} u_2 : u_1, u_2 \in \mathbb{K} \right\} =$$

$$= \left\{ x = \begin{bmatrix} 2 & 3 \\ 3 & -2 \\ 1 & 5 \end{bmatrix} u : u \in \mathbb{K}^2 \right\} = \operatorname{im} \begin{bmatrix} 2 & 3 \\ 3 & -2 \\ 1 & 5 \end{bmatrix}$$

Opis typu (R): Poprzez układ równań określających przestrzeń V, np.

$$V = \left\{ x \in \mathbb{K}^3 : \begin{matrix} 2x_1 + 3x_2 - 4x_3 = 0 \\ x_1 - 4x_2 + 3x_3 = 0 \end{matrix} \right\} =$$

$$= \left\{ x \in \mathbb{K}^3 : \begin{bmatrix} 2 & 3 & -4 \\ 1 & -4 & 3 \end{bmatrix} x = 0 \right\} = \ker \begin{bmatrix} 2 & 3 & 4 \\ 1 & -4 & 3 \end{bmatrix}$$

15.5. Podstawowe Zadania Wyliczeniowe Algebry Wektorowej.

(1) Uproszczenie opisu typu (W), w szczególności badanie liniowej niezależności wektorów.
(2) Przejście: opis(W) $\rightarrow$opis(R) (t. j. wyznaczenie zupełnego układu równań spełnianych przez dane wektory).
1^* Uproszczenie opisu typu (R), w szczególności badanie liniowej niezależności równań.
2^* Przejście: opis (R) $\rightarrow$ opis (W) (t. j. rozwiązywanie układu równań liniowych jednorodnych).
(3) Rozwiązywanie układu równań liniowych niejednorodnych.

(4) Odwracanie macierzy kwadratowej (w szczególności badanie nieosobliwości macierzy).
(5) Wyznaczanie przecięcia danych podprzestrzeni w $\mathbb{K}^m$.
(6) Wyznaczanie sumy algebraicznej danych podprzestrzeni w $\mathbb{K}^m$.

15.5.1. Rozwiązywanie tych zadań metodą operacji elementarnych.
Ad(1)
Mając $V := \operatorname{im} A$ stosujemy do A algorytm redukcji kolumnowych, n. p.

$$V = \left\langle \begin{bmatrix} 7 \\ 4 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 10 \\ -5 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -7 \\ 4 \end{bmatrix} \right\rangle;$$

$$\begin{bmatrix} 7 & 3 & \boxed{1} & 0 \\ 4 & 2 & 2 & -1 \\ 0 & 2 & 10 & -7 \\ 5 & 1 & -5 & 4 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 0 & 0 & 1 & 0 \\ -10 & -4 & 2 & \boxed{-1} \\ -70 & -28 & 10 & -7 \\ 40 & 16 & -5 & 4 \end{bmatrix} \underset{k}{\sim} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -4 & -7 \\ 0 & 0 & 3 & 4 \end{bmatrix},$$

co daje

$$v = \left\langle \begin{bmatrix} 1 \\ 0 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -7 \\ 4 \end{bmatrix} \right\rangle = \operatorname{im} \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -4 & -7 \\ 3 & 4 \end{bmatrix}.$$

Ad(2)
Podanie opisu typu (R) dla $V = \operatorname{im} A$, gdzie A jest KZ-macierzą (co uzyskujemy stosując (1)) jest bardzo proste, gdyż w rozkładzie $x = \sum \lambda^i \bar{a}_i$ współczynniki λ^i przy $A_i \neq 0$ są proporcjonalne do odpowiednich x_j, n. p. (kontynuacja poprzedniego przykładu):

$$\exists \lambda^1, \lambda^2 : x = \lambda^1 \begin{bmatrix} 1 \\ 0 \\ -4 \\ 3 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 \\ -1 \\ -7 \\ 4 \end{bmatrix}$$

co jest równoważne

$$\begin{cases} x_3 = -4x_1 + 7x_2 \\ x_4 = 3x_1 - 4x_2 \end{cases},$$

a zatem

$$V = \left\{ x \in \mathbb{K}^4 : \begin{matrix} x_3 = -4x_1 + 7x_2 \\ x_4 = 3x_1 - 4x_2 \end{matrix} \right\} = \ker \begin{bmatrix} 4 & -7 & 10 & \\ 3 & -4 & 0 & -1 \end{bmatrix}$$

Warto zauważyć, że uzyskany w ten sposób opis $V = \ker B$ ma postać możliwie najwygodniejszą: B jest WZ-macierzą! Podobnie jest dla pozostałych omawianych tu procedur.

Ad(1*)
Mając $V = \ker B$ stosujemy do B algorytm redukcji wierszowych, n. p. jeśli

$$V = \left\{ x \in \mathbb{K}^4 : \begin{array}{rcrcrcrcl} 7x_1 & + & 3x_2 & + & x_3 & & & = & 0 \\ 4x_1 & + & 2x_2 & + & 2x_3 & - & x_4 & = & 0 \\ & & 2x_2 & + & 10x_3 & - & 7x_4 & = & 0 \\ 5x_1 & + & x_2 & - & 5x_3 & + & 4x_4 & = & 0 \end{array} \right\} = \ker B,$$

to

$$B = \begin{bmatrix} 7 & 3 & 1 & 0 \\ 4 & 2 & 2 & -1 \\ 0 & 2 & 10 & -7 \\ 5 & \boxed{1} & -5 & 4 \end{bmatrix} \underset{w}{\sim} \begin{bmatrix} -8 & 0 & 16 & -12 \\ \boxed{-6} & 0 & 12 & -9 \\ -10 & 0 & 20 & -15 \\ 5 & 1 & -5 & 4 \end{bmatrix}$$
$$\underset{w}{\sim} \begin{bmatrix} 0 & 0 & 0 & 0 \\ -6 & 0 & 12 & -9 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 5 & -\frac{7}{2} \end{bmatrix} \underset{w}{\sim} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & -4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 10 & -7 \end{bmatrix}$$

a więc

$$V = \left\{ x \in \mathbb{K}^4 : \begin{array}{r} 2x_1 - 4x_3 + 3x_4 = 0 \\ 2x_2 + 10x_3 - 7x_4 = 0 \end{array} \right\} = \ker \begin{bmatrix} 2 & 0 & -4 & 3 \\ 0 & 2 & 10 & -7 \end{bmatrix}$$

Ad (2*)
Podanie opisu typu (W) dla $V = \ker B$, gdzie B jest WZ-macierzą (co uzyskujemy stosując (1*)) jest bardzo proste: te współrzędne x_i wektora $x \in \ker B$, dla których i jest numerem jednej z kolumn PD*-podmacierzy macierzy B, można wyliczyć jednoznacznie przy dowolnie zadanych wartościach x_j dla j spoza numerów podmacierzy. N. p. (kontynuacja poprzednich przykładów)

$$V = \left\{ x \in \mathbb{K}^4 : \begin{array}{l} x_1 = 2x_3 - \dfrac{3}{2}x_4 \\ x_2 = -5x_3 + \dfrac{7}{2}x_4 \end{array} \right\} = \left\{ x = \begin{bmatrix} 2x_3 - \frac{3}{2}x_4 \\ -5x_3 + \frac{7}{2}x_4 \\ x_3 \\ x_4 \end{bmatrix} : x_3, x_4 \in \mathbb{K} \right\} =$$

$$= \left\{ x = \begin{bmatrix} 2 \\ -5 \\ 1 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} -3 \\ 7 \\ 0 \\ 2 \end{bmatrix} \frac{x_4}{2} : x_3, x_4 \in \mathbb{K} \right\} =$$

$$\left\langle \begin{bmatrix} 2 \\ -5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 0 \\ 2 \end{bmatrix} \right\rangle = \operatorname{im} \begin{bmatrix} 2 & -3 \\ -5 & 7 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Ad(3)
Stosujemy algorytm redukcji wierszowych do macierzy rozszerzonej układu, pamiętając przy tym, by bazowe wyrazy wybierać jedynie z części głównej tej macierzy.
Trzy przykłady:

$$\begin{bmatrix} 0 & 2 & 1 & 3 \\ 1 & 1 & 1 & 3 \\ 5 & 10 & 6 & 15 \\ 1 & 2 & 1 & 2 \end{bmatrix} x = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 1 \end{bmatrix}; \tag{a}$$

$$\begin{bmatrix} 3 & -4 & 1 \\ -5 & 2 & 3 \\ 2 & 1 & -3 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \tag{b}$$

$$\begin{bmatrix} 3 & -4 & 1 \\ -5 & 2 & 3 \\ 2 & 1 & -3 \end{bmatrix} x = \begin{bmatrix} 2 \\ -8 \\ 5 \end{bmatrix}. \tag{c}$$

Rozwiązanie:

$$\left[\begin{array}{cccc|c} 0 & 2 & \boxed{1} & 3 & 2 \\ 1 & 1 & 1 & 3 & 2 \\ 5 & 10 & 6 & 15 & 1 \\ 1 & 2 & 1 & 2 & 1 \end{array}\right] \underset{w}{\sim} \left[\begin{array}{cccc|c} 0 & 2 & 1 & 3 & 2 \\ \boxed{1} & -1 & 0 & 0 & -4 \\ 5 & -2 & 0 & -3 & -11 \\ 1 & 0 & 0 & -1 & 3 \end{array}\right]$$

$$\underset{w}{\sim} \left[\begin{array}{cccc|c} 0 & 2 & 1 & 3 & 2 \\ 1 & -1 & 0 & 0 & -4 \\ 0 & 3 & 0 & -3 & 9 \\ 0 & \boxed{1} & 0 & -1 & 3 \end{array}\right] \underset{w}{\sim} \left[\begin{array}{cccc|c} 0 & 0 & 1 & 5 & -4 \\ 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 3 \end{array}\right] \Rightarrow$$

$$\Rightarrow \begin{cases} x_1 &= & -1 + x_4 \\ x_2 &= & 3 + x_4 \\ x_3 &= & -4 - 5x_4 \end{cases}$$

$$\left[\begin{array}{ccc|cc} 3 & -4 & 1 & 1 & 2 \\ -5 & 2 & 3 & 2 & -8 \\ 2 & \boxed{1} & -3 & 3 & 5 \end{array}\right] \underset{w}{\sim} \left[\begin{array}{ccc|cc} 11 & 0 & -11 & 13 & 22 \\ -9 & 0 & \boxed{9} & -4 & -18 \\ 2 & 1 & -3 & 3 & 5 \end{array}\right] \underset{w}{\sim}$$

$$\underset{w}{\sim} \left[\begin{array}{ccc|cc} 0 & 0 & 0 & \frac{73}{9} & 0 \\ -9 & 0 & -9 & -4 & -18 \\ -1 & 1 & 0 & \frac{5}{3} & -1 \end{array}\right].$$

Odpowiedzi:

$$x = \begin{bmatrix} -1 \\ 3 \\ -4 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \\ -5 \\ 1 \end{bmatrix}, \lambda \in \mathbb{K} \tag{a}$$

$$\text{układ sprzeczny} \tag{b}$$

$$\begin{cases} x_2 &= -1 + x_1 \\ x_3 &= -2 + x_1 \end{cases} \quad \text{tzn.} \quad x = \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \lambda \in \mathbb{K} \tag{c}$$

Ad(4)

Ze wzoru $AB = [A\bar{b}_1, \dots, A\bar{b}_n]$ (rozkład na kolumny; $A, B \in \mathbf{M}^n{}_n(\mathbb{K})$) wynika, że kolumna $\bar{b}_k$ macierzy $B = A^{-1}$ jest rozwiązaniem układu równań

$$Ax = \bar{e}_k = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}.$$

To spostrzeżenie wraz z (3) daje następujący sposób znajdowania B^{-1} (przykłady):

$$\left[\begin{array}{ccc|ccc} 3 & \boxed{1} & 4 & 1 & 0 & 0 \\ 5 & 2 & 1 & 0 & 1 & 0 \\ 6 & 2 & 7 & 0 & 0 & 1 \end{array}\right] \underset{w}{\sim} \left[\begin{array}{ccc|ccc} 3 & 1 & 4 & 1 & 0 & 0 \\ \boxed{-1} & 0 & -7 & -2 & 1 & 0 \\ 0 & 0 & -1 & -2 & 0 & 1 \end{array}\right] \underset{w}{\sim}$$

$$\underset{w}{\sim} \left[\begin{array}{ccc|ccc} 0 & 1 & -17 & -5 & 3 & 0 \\ -1 & 0 & -7 & -2 & 1 & 0 \\ 0 & 0 & \boxed{-1} & -2 & 0 & 1 \end{array}\right] \underset{w}{\sim} \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 29 & 3 & -17 \\ -1 & 0 & 0 & 12 & 1 & -7 \\ 0 & 0 & -1 & -2 & 0 & 1 \end{array}\right] \underset{w}{\sim}$$

$$\underset{w}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -12 & -1 & 7 \\ 0 & 1 & 0 & 29 & 3 & -17 \\ 0 & 0 & 1 & 2 & 0 & -1 \end{array}\right].$$

Zatem

$$\begin{bmatrix} 3 & 1 & 4 \\ 5 & 2 & 1 \\ 6 & 2 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} -12 & -1 & 7 \\ 29 & 3 & -17 \\ 2 & 0 & -1 \end{bmatrix}.$$

W sposób alternatywny obliczymy teraz, że

$$\begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 6 \\ 1 & 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 7 & -5 & 2 \\ -3 & 2 & 0 \\ -1 & 1 & -1 \end{bmatrix}:$$

$$\left[\begin{array}{ccc} 2 & 3 & 4 \\ 3 & 5 & 6 \\ \boxed{1} & 2 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \underset{k}{\sim} \left[\begin{array}{ccc} 2 & \boxed{-1} & 2 \\ 3 & -1 & 3 \\ 1 & 0 & 0 \\ \hline 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \underset{k}{\sim} \left[\begin{array}{ccc} 0 & -1 & 0 \\ 1 & -1 & \boxed{1} \\ 1 & 0 & 0 \\ \hline -1 & -2 & -5 \\ 2 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right] \underset{k}{\sim}$$

$$\underset{k}{\sim} \left[\begin{array}{ccc} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ \hline 2 & -7 & -5 \\ 0 & 3 & 2 \\ -1 & 1 & 1 \end{array}\right] \underset{k}{\sim} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 7 & -5 & 2 \\ -3 & 2 & 0 \\ -1 & 1 & -1 \end{array}\right].$$

Ad(5)

Gdy dla V_1 znajdziemy opis typu (W): $V = \langle v_1, \dots, v_n \rangle$, a dla V_2–typu (R), wówczas przecięcie $V_1 \cap V_2$ można przedstawić jako zbiór tych $v = \sum \lambda^i v_i$, których współczynniki λ^i spełniają równanie indukowane z równań opisujących V_2.

Przykład.

Dla $V_1 = \ker \begin{bmatrix} 5 & 1 & -8 & 7 & 0 \\ 3 & 0 & -5 & 5 & 1 \end{bmatrix}$, $V_2 = \ker \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 4 & -1 & -3 & 0 & -1 \end{bmatrix}$ mamy (patrz p.(2*):

$$V_1 = \operatorname{im} \begin{bmatrix} 1 & 0 & 0 \\ -5 & 8 & -7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & 5 & -5 \end{bmatrix} = \operatorname{im} A.$$

Przy tym wektor

$$v = A \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda^1 \begin{bmatrix} 1 \\ -5 \\ 0 \\ 0 \\ -3 \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 \\ 8 \\ 1 \\ 0 \\ 5 \end{bmatrix} + \lambda^3 \begin{bmatrix} 0 \\ -7 \\ 01 \\ -5 \end{bmatrix} \in V_1$$

spełnia równanie opisujące V_2 wtedy i tylko wtedy, gdy

$$0 = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 4 & -1 & -3 & 0-1 & \end{bmatrix} A \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \begin{bmatrix} 6 & -8 & 6 \\ 12 & -16 & 12 \end{bmatrix} \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{bmatrix},$$

a stąd $0 = 3\lambda^1 - 4\lambda^2 + 3\lambda^3$, czyli $\lambda^3 = -\lambda^1 + \frac{4}{3}\lambda^2$, tzn.

$$\begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} \lambda^1 & \dfrac{\lambda^2}{3} \end{bmatrix},$$

gdzie $\lambda^1, \lambda^2 \in \mathbb{K}$ są dowolne.
Odpowiedź:

$$V_1 \cap V_2 = \operatorname{im} A \begin{bmatrix} 1 & 0 \\ 0 & 3 \\ -1 & 4 \end{bmatrix} = \operatorname{im} \begin{bmatrix} 1 & 0 \\ 2 & -4 \\ 0 & 3 \\ -1 & 4 \\ 2 & -5 \end{bmatrix}.$$

Ad(6)
Sposób jest „dwoisty” do poprzedniego: Tak jak w (5) znajdujemy dla jednej z podprzestrzeni opis typu (W), a dla drugiej–typu (R); następnie badamy, jakie kombinacje liniowe równań opisujących V_2 będą zerować generatory V_1.

Przykład. Dla V_1 i V_2 z poprzedniego przykładu kombinacja liniowa (ze współczynnikami λ^1 i λ^2) równań opisujących V_1 ma postać

$$[\,\lambda^1 \quad \lambda^2\,] \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 4 & -1 & -3 & 0 & -1 \end{bmatrix}.$$

Będzie ona zerować generatory V_1 wtedy i tylko wtedy, gdy

$$0 = [\,\lambda^1 \quad \lambda^2\,] \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 4 & -1 & -3 & 0 & -1 \end{bmatrix} A =$$

$$= [\,\lambda^1 \quad \lambda^2\,] \begin{bmatrix} 6 & -8 & 6 \\ 12 & -16 & 6 \end{bmatrix} = [\,6\lambda^1 + 12\lambda^2 \quad -8\lambda^1 - 16\lambda^2 \quad 6\lambda^1 + 12\lambda^2\,]$$

to znaczy,

$$\lambda^1 = -2\lambda^2 \Leftrightarrow [\,\lambda^1 \quad \lambda^2\,] = \lambda\,[\,-2 \quad 1\,], \lambda \in \mathbb{K} \text{ dowolne.}$$

Odpowiedź:

$$V_1 + V_2 = \ker\,[\,-2 \quad 1\,] \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 4 & -1 & -3 & 0 & -1 \end{bmatrix} = \ker\,[\,2 \quad 1 \quad -3 \quad 2 \quad -1\,]$$

UWAGA
Alternatywne sposoby dla (5) i (6), polegające na „połączeniu” obu zestawów równań (dla $V_1 \cap V_2$) lub generatorów (dla $V_1 + V_2$) opisujących V_1 i V_2 , a następnie dokonaniu ich redukcji, są na ogół bardziej pracochłonne!