VII FORMY KWADRATOWE, DWULINIOWE, WIELOLINIOWE.

W tym rozdziale zakładamy, że

a) rozważane przestrzenie wektorowe są skończonego wymiaru, oraz

b) w ciele skalarów $\mathbb{F}$ element $2_{\mathbb{F}} := 1_{\mathbb{F}} + 1_{\mathbb{F}}$ **nie** jest równy $0_{\mathbb{F}}$.

Ostatnie założenie wyrażane jest następująco: „charakterystyka ciała $\mathbb{F}$ jest różna od 2". Umożliwia ono wykonywanie w $\mathbb{F}$ „dzielenia przez 2": gdy przez $\frac{1}{2} \in \mathbb{F}$ oznaczyć odwrotność elementu $2_{\mathbb{F}}$, to $\frac{1}{2}a + \frac{1}{2}a = a$ dla $a \in \mathbb{F}$. (Przykładowo, jeśli $\mathbb{F} = \mathbb{Z}_5$ i $a = 3$, to $\frac{1}{2}a = 4$.) Najważniejsze dla nas przypadki, to gdy $\mathbb{F}$ jest ciałem liczbowym; wtedy wykonalność dzielenia przez 2 nie wymaga żadnych uzasadnień.

Wstęp.* Po afinicznych, najprostszymi są funkcje kwadratowe kilku zmiennych. Występują one w wielu zagadnieniach geometrii i analizy. W tym rozdziale pokażemy, jak do badania tych funkcji wykorzystać można własności macierzy, a zwłaszcza macierzy symetrycznych. Operacje elementarne na takich macierzach umożliwią uproszczenie wielomianu kwadratowego liniową zamianą zmiennych. Wyniki algebry liniowej są pomocne w ustaleniu, kiedy wielomian kwadratowy rzeczywisty przyjmuje tylko nieujemne (bądź tylko dodatnie) wartości. Umożliwią one również dowód ważnego „twierdzenia o bezwładności", sformułowanego w §2.

Badamy tu też funkcje dwuliniowe $V \times V \to \mathbb{F}$, gdzie V jest przestrzenią wektorową, a $\mathbb{F}$ jej ciałem skalarów. Oto pewne powody znaczenia tych funkcji:

1) Funkcje dwuliniowe przekazują pełną informację o endomorfizmach liniowych. Istotnie, gdy $L \in \mathcal{L}(\mathbb{F}^k)$ jest operatorem na przestrzeni $\mathbb{F}^k$, to funkcja $(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} \cdot L(\mathbf{v})$ jest dwuliniowa i nietrudno jest zauważyć, że wyznacza ona L. Pozwala to badanie endomorfizmów sprowadzić, przynajmniej formalnie, do badania funkcji dwuliniowych.

2) Funkcja dwuliniowa $V \times V \to \mathbb{F}$ wyznacza w przestrzeni V namiastkę geometrii euklidesowej: umożliwia zdefiniowanie ortogonalności wektorów, rzutu ortogonalnego, przekształcenia sprzężonego, izometrii. (Opiszemy to dokładniej w §§3 i 4.) Tym samym pewne intuicje i wyobrażenia, które wiążemy z przestrzeniami euklidesowymi, mogą być choć w części przeniesione na przestrzenie z wyróżnioną funkcją dwuliniową.

3) Każda jednorodna funkcja kwadratowa $V \to \mathbb{F}$ wyznacza funkcję dwuliniową $V \times V \to \mathbb{F}$. Ta prosta, lecz podstawowa obserwacja poczyniona w §3 umożliwia użycie opisanych wyżej pojęć geometrycznych do badania funkcji kwadratowych.

Rozwinięcie powyższych punktów 1)–3) wykracza jednak poza zakres wykładu i dotkniemy ich tylko w końcowych zadaniach uzupełniających. W materiale uzupełniającym w §6 omówimy też podstawowe pojęcia rachunku tensorowego i jego związku z badaniem funkcji n–liniowych dla $n>2$.

§ 1. Wielomiany i funkcje wielomianowe stopnia ≤ 2

1. Wielomiany stopnia ≤ 2 i wyznaczone przez nie funkcje wielomianowe $\mathbb{F}^k \to \mathbb{F}$.

Definicja. a) **Wielomian stopnia** ≤ 2, przemiennych zmiennych $x_1, ..., x_k$ i współczynnikach w ciele $\mathbb{F}$ (czyli takich, które są „nad $\mathbb{F}$ ”), to wyrażenie

$$p = \sum_{1 \leq i \leq j \leq k} b_{ij} x_i x_j + \sum_{i=1}^{k} b_i x_i + c, \tag{1}$$

gdzie wszystkie współczynniki b_{ij}, b_i, c są elementami $\mathbb{F}$.

b) W zbiorze wszystkich takich wielomianów, oznaczanym tu przez $\mathbb{F}_{\leq 2}[x_1, ..., x_k]$, wprowadzamy w naturalny sposób dodawanie i mnożenie przez skalar.

c) **Wartością** wielomianu (1) **w punkcie** $\mathbf{u} = (u_1, ..., u_k) \in \mathbb{F}^k$ nazywamy skalar

$$p(\mathbf{u}) := \sum_{1 \leq i \leq j \leq k} b_{ij} u_i u_j + \sum_{i=1}^{k} b_i u_i + c.$$

Funkcję $\mathbb{F}^k \to \mathbb{F}$ zadaną przez $\mathbf{u} \mapsto p(\mathbf{u})$ nazywamy **funkcją wielomianową wyznaczoną przez wielomian** p. Gdy $b_{ij} \neq 0$ dla pewnych i, j, to zarówno o wielomianie p, jak i o wyznaczonej przez niego funkcji mówimy, że są **stopnia 2** lub **kwadratowe**. W przeciwnym razie mówimy, że są one: **stopnia** 1, gdy $b_i \neq 0$ dla pewnego i, **stopnia 0** gdy $c \neq 0$ i $b_1 = \cdots = b_k = 0$, zaś **stopnia** $-\infty$ w pozostałym razie, tzn. gdy $p = 0$.

Poprawność tej definicji w odniesieniu do funkcji zapewnia następujące

Twierdzenie 1. *Gdy $2_{\mathbb{F}} \neq 0_{\mathbb{F}}$, to funkcja, wyznaczona przez wielomian $p \in \mathbb{F}_{\leq 2}[x_1, ..., x_k]$, określa go (czy równoważnie: jego współczynniki) jednoznacznie.*

Dowód. Przy oznaczeniach (1) mamy

$$c = p(\mathbf{0}) \tag{2}$$

Badana funkcja wyznacza więc współczynnik c wielomianu p oraz następujące funkcje $f_1, f_2 : \mathbb{F}^k \to \mathbb{F}$

$$f_1(\mathbf{u}) := \frac{1}{2}(p(\mathbf{u}) - p(-\mathbf{u})), \quad f_2(\mathbf{u}) := p(\mathbf{u}) - f_1(\mathbf{u}) - c = \frac{1}{2}(p(\mathbf{u}) + p(-\mathbf{u})) - c. \tag{3}$$

Latwe rachunki pokazują, że przy tych definicjach,

$$f_1(\mathbf{u}) = \sum_{i=i}^{k} b_i u_i \quad \text{oraz} \quad f_2(\mathbf{u}) = \sum_{1 \leq i \leq j \leq k} b_{ij} u_i u_j \text{ dla } \mathbf{u} \in \mathbb{F}^k \tag{4}$$

skąd $b_i = f_1(\mathbf{e}_i)$, $b_{ii} = f_2(\mathbf{e}_i)$ i $b_{ij} = f_2(\mathbf{e}_i + \mathbf{e}_j) - f_2(\mathbf{e}_i) - f_2(\mathbf{e}_j)$ dla $1 \leq i < j \leq k$. Wraz ze wzorami (2) i (3) wyznacza to współczynniki wielomianu p. $\square$

Uwaga 1. a) Oczywiście, podobnie do wielomianów k–zmiennych stopnia ≤ 2 można definiować wielomiany wyższych stopni. Zbiór wszystkich wielomianów nad $\mathbb{F}$, zmiennych $x_1, ..., x_k$, oznaczamy przez $\mathbb{F}[x_1, ..., x_k]$; jest on w naturalny sposób pierścieniem przemiennym. Odpowiednie uogólnienie twierdzenia 1 wymaga wymaga jednak dodatkowych założeń o ciele $\mathbb{F}$. (Np., gdy $\mathbb{F} = \mathbb{Z}_3$, to wielomian $x(x-1)(x-2)$ jest różny od 0, lecz wyznacza funkcję zerową.) Wystarczające jest założenie, by ciało $\mathbb{F}$ było nieskończone; dowód będzie podany na wykładzie Algebry.

b) Ze względu na twierdzenie 1, będziemy niekiedy utożamiać funkcję wielomianową stopnia ≤ 2 z wyznaczającym ją wielomianem i np. mówić o jej **współczynnikach**.

Przyjmijmy $p_2 := \sum_{1 \leq i \leq j \leq k} b_{ij} x_i x_j$ i $p_1 := \sum_{i=1}^{k} b_i x_i$. Wielomiany p_2 i $p_1 + c$ nazywamy, odpowiednio, **częścią główną** i **częścią liniową** wielomianu p. Podobnie funkcje f_2 i $f_1 + c$ nazywamy częścią główną i liniową [1] wyznaczonej przez p funkcji. Wielomian p, a także odpowiadającą mu funkcję, nazwiemy **formą kwadratową**, gdy $p = p_2$, zaś **formą liniową**, gdy $p = p_1$.

Uwaga 2. Forma kwadratowa może być stopnia $-\infty$ (tzn. być zerowa), podczas gdy kwadratowa funkcja czy wielomian są, z przyjętej definicji, zawsze stopnia 2.

Uwaga 3. Przez x_i oznaczamy na ogół i–tą z rozważnych zmiennych (wówczas jest to pewne wyrażenie algebraiczne), lecz niekiedy może tak być oznaczony i skalar. Ciąg zmiennych $x_1, ..., x_k$ będziemy oznaczać przez $\mathbf{x}$, i tak samo może być oznaczony wektor w $\mathbb{F}^k$ (którego współrzędne $x_1, ..., x_k$ są skalarami). Nie prowadzi to do nieporozumień, bo omawiamy na ogół lub jest skądinąd jasne, czy $\mathbf{x}$ jest ciągiem zmiennych, czy skalarów.

Zadanie 1. Gdy $f : \mathbb{F}^k \to \mathbb{F}$ jest funkcją wielomianową stopnia ≤ 2, to równoważne są warunki:

a) f jest formą kwadratową;

b) $f(\mathbf{0}) = 0$ i f jest funkcją parzystą, tzn. $f(\mathbf{u}) = f(-\mathbf{u})$ dla $\mathbf{u} \in \mathbb{F}^k$;

c) f jest funkcją 2–jednorodną, tzn. $f(t\mathbf{u}) = t^2 f(\mathbf{u})$ dla $t \in \mathbb{F}$ i $\mathbf{u} \in \mathbb{F}^k$.

<u>Zadanie uzupełniające 1.</u> * Niech funkcja $f : \mathbb{R}^2 \to \mathbb{R}$ ma tę własność, że dla każdych $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, wzór $t \mapsto f(\mathbf{u} + t\mathbf{v})$ zadaje funcję wielomianową stopnia ≤ 2. Czy f jest funkcją wielomianową stopnia ≤ 2?

[1] Nazwa ta odzwierciedla istniejącą niestety w nazewnictwie matematycznym niekonsekwencję: jeśli $h : \mathbb{F}^k \to \mathbb{F}^l$ jest postaci $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$, gdzie $\mathbf{A} \in \mathcal{M}_{l,k}$ i $\mathbf{b} \in \mathbb{F}^l$, to gdy $l > 1$ mówi się o h, że jest „przekształceniem afinicznym", a gdy $l = 1$ –że jest „funkcją liniową"; nazwa „funkcjonał liniowy" zaś oznacza, że $l = 1$ i $\mathbf{b} = \mathbf{0}$.

2. Formy kwadratowe a macierze.

Forma kwadratowa
$$q = \sum_{1\le i\le j\le k} b_{ij}x_ix_j \tag{5}$$
jest wyznaczona przez swe współczynniki b_{ij}. Ponieważ zakładamy w (5), że $i \le j$, to współczynniki te tworzą tylko „górną połówke" macierzy rozmiaru $k \times k$. By otrzymać pełną macierz, możemy dopisać w niej zera poniżej przekątnej; odpowiada to rozszerzeniu w (1) sumowania na wszystkie pary (i,j) takie, że $i,j \in \{1,...,k\}$, przy czym przyjmujemy $b_{ij} = 0$ gdy $i > j$. Jednak przy takiej zmianie zakresu wskaźników istnieją inne jeszcze możliwości wyboru $k \times k$–macierzy współczynników.

Definicja. Formą kwadratową **wyznaczoną** przez macierz $\mathbf{A} \in \mathcal{M}_k$ nazywamy zarówno wielomian
$$q_{\mathbf{A}} := \sum_{i,j=1}^{k} a_{ij}x_ix_j \tag{6}$$
jak i odpowiadająca mu funkcję $\mathbb{F}^k \to \mathbb{F}$, którą oznaczymy $f_{\mathbf{A}}$. By $q_{\mathbf{A}}$ zapisać w postaci (5), należy dokonać redukcji wyrazów x_ix_j oraz x_jx_i $(i < j)$. Oczywisty jest

Lemat 1. *Dla form q i $q_{\mathbf{A}}$ zadanych wzorami (5) i (6), równość $q = q_{\mathbf{A}}$ ma miejsce wtedy i tylko wtedy, gdy* $a_{ii} = b_{ii}$ oraz $a_{ij} + a_{ji} = b_{ij}$ dla $i < j$ $(i,j = 1,...,k)$.

Oznaczenia. Jak wcześniej, we wzorach wykorzystujących mnożenie przez macierz, skończone ciągi skalarów traktujemy jako macierze jednokolumnowe. Umowę tę rozszerzamy obecnie i na ciągi wielomianów. Tak więc dla $\mathbf{A} \in \mathcal{M}_k$ i ciągu zmiennych $\mathbf{x} = (x_1,...,x_k)$, przex $\mathbf{Ax}$ oznaczamy ciąg wielomianów $p_i = \sum_{j=1}^k a_{ij}x_j$ $(i = 1,...,k)$, a równość (6) zapisujemy tak:
$$q_{\mathbf{A}} = \mathbf{x}^t\mathbf{A}\mathbf{x} \tag{7}$$
Jest to oczywiście spowodowane tym, że $q_{\mathbf{A}}(\mathbf{v}) = \sum_i v_ip_i(\mathbf{v}) = \mathbf{v}^t\mathbf{A}\mathbf{v}$ dla $\mathbf{v} \in \mathbb{F}^k$.

Uwaga i definicja. a) Okazuje się (jeden z powodów wskażemy w p.3), że spośród macierzy wyznaczających formę kwadratową q, daną wzorem (5), najdogodniej jest wybrać symetryczną. Z lematu wynika, że taka symetryczna macierz $\mathbf{A}$ istnieje i jest wyznaczona jednoznacznie; nazywać ją będziemy **macierzą** (Gaussa) **formy** q.

b) Macierz ta jest wyznaczona tym, że $a_{ii} = b_{ii}$ i $a_{ij} = a_{ji} = \frac{1}{2}b_{ij}$ dla $1 \le i < j \le k$, a także – na podstawie twierdzenia 1 w p.1 – tym, że
$$\mathbf{A} = \mathbf{A}^t \text{ i } q(\mathbf{v}) = \mathbf{v}^t\mathbf{A}\mathbf{v} \text{ dla wszystkich } \mathbf{v} \in \mathbb{F}^k.$$

Ćwiczenie. Przestrzeń $\mathbb{R}^k$ wyposażamy w standardowy iloczyn skalarny. Dla macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ dowieść, że:

a) Jeśli $\mathbf{A}\mathbf{v}\perp\mathbf{v}$ $\forall\mathbf{v}\in\mathbb{R}^k$, to $\mathbf{A}^t=-\mathbf{A}$. (Por. zad. uz. 2 w §VI.3.1.)
b) Jeśli funkcja $\mathbf{v}\mapsto\langle\mathbf{A}\mathbf{v},\mathbf{v}\rangle$ jest na sferze $\|\mathbf{v}\|=1$ stale równa c, to $\mathbf{A}+\mathbf{A}^t=2c\mathbf{I}$.

* Wielomianowi kwadratowemu nie będącemu formą też można przyporządkować macierz symetryczną. (Nie zostanie to jednak tu wykorzystane i dlatego pozostała część tego punktu to materiał uzupełniający). W tym celu dla wielomianu p zadanego wzorem (1) przyjmijmy

$$a_{ij}=a_{ji}=b_{ij}/2 \text{ gdy } i<j,\ a_{ii}=b_{ii},\ a_{0i}=a_{i0}=b_i/2 \ \text{ oraz } \ a_{00}=c$$

Macierz symetryczną $(a_{ij})_{i,j=0,1,\dots,k}$ oznaczmy $\widetilde{\mathbf{A}}$ i nazwijmy **rozszerzoną macierzą** wielomianu p, a także wyznaczonej przez niego funkcji. Przy $\widetilde{p}:=\sum_{i,j=0}^k a_{ij}x_ix_j$ mamy $p(\mathbf{x})=\widetilde{p}(1,x_1,...,x_k)$, skąd

$$p=\widetilde{\mathbf{x}}^t\widetilde{\mathbf{A}}\widetilde{\mathbf{x}},\ \text{gdzie}\ \ \widetilde{\mathbf{x}}:=(1,x_1,...,x_k). \tag{8}$$

3. Upraszczanie formy podstawieniem liniowym. Kongruentność macierzy.

Definicja. Niech $p=\sum_{i,j=1}^k b_{ij}x_ix_j+\sum_{i=1}^k b_ix_i+c$ i niech $\mathbf{C}\in\mathcal{M}_k$ będzie macierzą nieosobliwą. Zastąpmy każdą ze zmiennych x_i wielomianem $\sum_{j=1}^k c_{ij}y_j$. Powiemy, że **zamiana zmiennych** (lub: **podstawienie**) $\mathbf{x}=\mathbf{C}\mathbf{y}$ **przeprowadza** p w otrzymany wielomian p' zmiennych $y_1,...,y_k$.

Podstawienie $\mathbf{x}=\mathbf{C}\mathbf{y}$ nazywamy **liniowym**. Zakładamy w nim zawsze nieosobliwość macierzy $\mathbf{C}$. Umożliwia to wyrażenie $\mathbf{y}$ poprzez $\mathbf{x}$ wzorem $\mathbf{y}=\mathbf{C}^{-1}\mathbf{x}$, analogicznym do $\mathbf{x}=\mathbf{C}\mathbf{y}$. Latwo widzieć, że stopień wielomianu p' nie przewyższa stopnia wielomianu p. Wobec powyższej symetrii między p i p', są więc one tego samego stopnia.

Oczywiście, gdy podstawienia $\mathbf{x}=\mathbf{C}\mathbf{y}$ i $\mathbf{y}=\mathbf{S}\mathbf{z}$ przeprowadzają p w p' i p' w p'', odpowiednio, to podstawienie $\mathbf{x}=\mathbf{C}\mathbf{S}\mathbf{z}$ przeprowadza p w p''.

Twierdzenie 1. *Gdy $q\in\mathbb{F}[x_1,...,x_k]$ jest formą kwadratową o macierzy $\mathbf{A}$, to podstawienie $\mathbf{x}=\mathbf{C}\mathbf{y}$ przeprowadza ją w formę q' o macierzy $\mathbf{C}^t\mathbf{A}\mathbf{C}$.*

Dowód. Dla $\mathbf{y}\in\mathbb{F}^k$ zachodzi

$$q'(\mathbf{y})=q(\mathbf{C}\mathbf{y})=(\mathbf{C}\mathbf{y})^t\mathbf{A}(\mathbf{C}\mathbf{y})=\mathbf{y}^t(\mathbf{C}^t\mathbf{A}\mathbf{C})\mathbf{y}$$

Ponadto, $\mathbf{C}^t\mathbf{A}\mathbf{C}$ jest macierzą symetryczną, w ślad za $\mathbf{A}$. (Patrz poniższe zadanie). Stąd i z części b) definicji-uwagi z p.2 wynika, że macierzą formy q' jest $\mathbf{C}^t\mathbf{A}\mathbf{C}$. □

Definicja. Macierze $\mathbf{A},\mathbf{B}\in\mathcal{M}_k(\mathbb{F})$ nazwiemy **kongruentnymi**, jeśli istnieje macierz nieosobliwa $\mathbf{C}\in\mathcal{M}_k(\mathbb{F})$ taka, że $\mathbf{B}=\mathbf{C}^t\mathbf{A}\mathbf{C}$.

Zadanie 1. a) Kongruentność jest relacją równoważności w zbiorze $\mathcal{M}_k(\mathbb{F})$.

b) Gdy jedna z kongruentnych macierzy jest symetryczna (odp. antysymetryczna), to druga też.
c) Gdy $\mathbf{A}_i \sim \mathbf{B}_i$ dla $i = 1, 2$, to $\mathrm{diag}(\mathbf{A}_1, \mathbf{A}_2) \sim \mathrm{diag}(\mathbf{B}_1, \mathbf{B}_2)$, gdzie $\sim$ to kongruentność.
d) Macierze kongruentne mają ten sam rząd.

Uwaga 1. Część b) zadania uwidacznia korzyść, jaką niesie wybór macierzy symetrycznej spośród wszystkich, zadających daną formę: zbiór macierzy symetrycznych jest zamknięty względem odpowiadającej zamianie zmiennych relacji kongruencji, podczas gdy np. narzucający się wybór macierzy górnie trójkątnej nie prowadzi do takiego zbioru.

Wykorzystamy teraz podstawienia liniowe do upraszczania macierzy formy.

Twierdzenie 2 (Lagrange'a o diagonalizacji form kwadratowych, 2 wersje)**.** *Każda macierz symetryczna jest kongruentna z pewną macierzą diagonalną.*

Równoważne sformułowanie: Każdą formę kwadratową $q \in \mathbb{F}[x_1, ..., x_k]$ można podstawieniem liniowym przeprowadzić w formę $\sum_{i=1}^{k} \lambda_i y_i^2$, dla pewnych $\lambda_1, ..., \lambda_k \in \mathbb{F}$.

By dostrzec równoważność obu wersji wystarcza zapisać q w postaci $q_{\mathbf{A}}$, dla odpowiedniej macierzy symetrycznej $\mathbf{A}$, i skorzystać z twierdzenia 1.

Twierdzenie 2 udowodnimy opisując sposób wyznaczenia macierzy $\mathbf{C}$ i skalarów $\lambda_1, ..., \lambda_k$ takich, że $\mathbf{C}^t\mathbf{A}\mathbf{C} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$. Różni się on tym od opisanego w §II.3.2 sposobu doprowadzenia macierzy do postaci schodkowej, że operacje wierszowe replikowane są jako kolumnowe.

Sposób diagonalizacji macierzy symetrycznej $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ przez kongruencję. Wykonujemy kolejno k kroków opisanych niżej.

<u>Krok s–ty ($s = 1, ..., k$).</u> Niech $\mathbf{B}$ oznacza macierz symetryczną, otrzymaną w wyniku wykonania poprzedzających kroków i mającą wyrazy różne od 0 tylko na przekątnej i w miejscach ij dla $i, j \geq s$. (Gdy $s = 1$, przyjmujemy $\mathbf{B} = \mathbf{A}$.) Wyróżnimy dwie części tego kroku:

<u>Część 1.</u> Wykonujemy ją tylko, gdy $b_{ss} = 0$ i $b_{ts} \neq 0$ dla pewnego $t > s$. Wtedy do wiersza s dodajemy c–krotność wiersza t taką, że $c \neq 0$ i $d := 2b_{st} + c \cdot b_{tt} \neq 0$. (Można n.p. obrać jeśli nie $c = 1$, to $c = -1$.) Następnie, powtarzamy tę operację na <u>kolumnach</u> otrzymanej macierzy (dodajemy tę samą krotność t–tej kolumny do s–tej). Końcową macierz oznaczamy nadal przez $\mathbf{B}$; zauważmy, że $b_{ss} = c \cdot d \neq 0$.

<u>Część 2.</u> Od wierszy $s{+}1, s{+}2, ..., k$ macierzy $\mathbf{B}$ odejmujemy takie krotności wiersza p, by stojące poniżej przekątnej wyrazy kolumny s uczynić zerami. Następnie, zamieniamy zerami $(s+1)$-szy i dalsze wyrazy wiersza s otrzymanej macierzy.

To kończy opis obu części kroku s. Macierz $\mathbf{B}$, otrzymana w wyniku wykonania wszystkich k kroków, jest diagonalna (uzasadnienie poniżej). Dla otrzymania macierzy $\mathbf{C}$ takiej, że $\mathbf{C}^t\mathbf{A}\mathbf{C} = \mathbf{B}$, należy powyższą konstrukcję rozszerzyć, dopisując do $\mathbf{B}$ klatki kwadratowe, z których pierwsza (tj. przy $s = 1$) jest równa $\mathbf{I}_k$. W obu częściach każdego z kroków, klatkę dopisaną zmieniamy tylko wtedy, gdy na macierzy $\mathbf{B}$ wykonano operację wierszową, i wtedy powtarzamy ją na klatce dopisanej. Końcową klatkę dopisaną przyjmujemy za $\mathbf{C}^t$; po transpozycji, da ona szukaną macierz $\mathbf{C}$.

Wykazanie poprawności tego sposobu poprzedzimy przykładem.

Przykład 1. Niech

$$q = x_1^2 + 3x_2^2 + 19x_4^2 - 2x_1x_2 + 4x_1x_3 + 2x_1x_4 - 6x_2x_4 - 12x_3x_4.$$

Macierzą tej formy jest

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 & 1 \\ -1 & 3 & 0 & -3 \\ 2 & 0 & 0 & -6 \\ 1 & -3 & -6 & 19 \end{pmatrix} \in \mathcal{M}_4(\mathbb{R})$$

Wykonujemy kolejno opisane wcześniej kroki (nad strzałkami zaznaczono, czy wykonano część 1, czy 2 odpowiedniego kroku, oraz czy operacje były wierszowe, czy kolumnowe):

$$(\mathbf{A}, \mathbf{I}) \xrightarrow{2w} \left(\begin{array}{cccc|cccc} 1 & -1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 2 & -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & -4 & -8 & -2 & 0 & 1 & 0 \\ 0 & -2 & -8 & 18 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{2k} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \mathbf{2} & 2 & -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & -4 & -8 & -2 & 0 & 1 & 0 \\ 0 & -2 & -8 & 18 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{2w} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 2 & -2 & 1 & 1 & 0 & 0 \\ 0 & 0 & -6 & -6 & -3 & -1 & 1 & 0 \\ 0 & 0 & -6 & 16 & 0 & 1 & 0 & 1 \end{array}\right) \xrightarrow{2k} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & \mathbf{-6} & -6 & -3 & -1 & 1 & 0 \\ 0 & 0 & -6 & 16 & 0 & 1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{2w} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -6 & -6 & -3 & -1 & 1 & 0 \\ 0 & 0 & 0 & 22 & 3 & 2 & -1 & 1 \end{array}\right) \xrightarrow{2k} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -6 & 0 & -3 & -1 & 1 & 0 \\ 0 & 0 & 0 & 22 & 3 & 2 & -1 & 1 \end{array}\right)$$

Zatem przy $\mathbf{C} = \begin{pmatrix} 1 & 1 & -3 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ otrzymamy $\mathbf{C}^t\mathbf{A}\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -6 & 0 \\ 0 & 0 & 0 & 22 \end{pmatrix}$. Inaczej mówiąc, podstawienie $\mathbf{x} = \mathbf{C}\mathbf{y}$ przeprowadza q w formę $y_1^2 + 2y_2^2 - 6y_3^2 + 22y_4^2$.

Dowód twierdzenia 2. Wykażemy, że końcowe macierze $\mathbf{B}$ i $\mathbf{C}$ mają żądane własności.

Weźmy część 2 kroku s. Wykonujemy w niej ciąg operacji wierszowych, który wobec twierdzenia 1 z §I.4.3 powoduje zastąpienie $\mathbf{B}$ macierzą $\mathbf{EB}$, dla pewnej nieosobliwej macierzy $\mathbf{E}$. Ponieważ macierz $\mathbf{B}$ jest symetryczna, a s-ta kolumna macierzy $\mathbf{EB}$ ma tylko jeden (s-ty) wyraz niezerowy, to wykonanie na macierzy $\mathbf{EB}$ ciągu operacji kolumnowych, odpowiadających wykonanym poprzednio operacjom wierszowym, skutkuje jedynie zastąpieniem zerami wyrazów $s+1, ..., k$ wiersza s macierzy $\mathbf{EB}$. Wykorzystując ponownie przywołane twierdzenie stwierdzamy, że wykonanie całej części 2 powoduje zastąpienie macierzy $\mathbf{B}$ przez $\mathbf{EBE}^t$, dla pewnej nieosobliwej macierzy $\mathbf{E}$.

Tak samo zmienia się macierz $\mathbf{B}$ w części 1 kroku s. Po tym kroku pozostanie więc ona symetryczna i spełni warunek $b_{ij} = 0$ gdy $i \neq j$ i $\min(i,j) \leq s$. (Wynika to z opisu tego kroku.) Końcowa macierz $\mathbf{B}$ jest więc zarówno diagonalna, jak i równa $\mathbf{E}_n(...(\mathbf{E}_1\mathbf{A}\mathbf{E}_1^t)...)\mathbf{E}_n^t$, gdzie $\mathbf{E}_1, ..., \mathbf{E}_n$ to macierze nieosobliwe, odpowiadające wykonywanym częściom kolejnych kroków. Stąd $\mathbf{B} = \mathbf{SAS}^t$ dla $\mathbf{S} = \mathbf{E}_n ... \mathbf{E}_1$ – czyli przy $\mathbf{C} := \mathbf{S}^t$ zachodzi $\mathbf{B} = \mathbf{C}^t\mathbf{AC}$, a macierz $\mathbf{C}^t = \mathbf{S}$ otrzymujemy z klatki $\mathbf{I}_k$ przez wykonanie kolejno wszystkich opisanych operacji wierszowych. (Korzystamy z tego, że i–ta z tych operacji polega na mnożeniu macierzy z lewej strony przez $\mathbf{E}_i$.) $\square$

Uwaga 2. Wykonanie części 2 jakiegokolwiek kroku nie zmienia wartości wyznacznika klatki wyznaczonej przez pierwszych s wierszy i kolumn macierzy $\mathbf{B}$. (Tu s jest dowolną liczbę niewiększą od stopnia macierzy.) Istotnie, żadna z operacji wykonywanych w części 2 nie zmienia tego wyznacznika.

Uwaga 3. * Jeśli, jak w przykładzie 1, dla każdego s wykonanie części 1 jest zbędne, to otrzymana macierz $\mathbf{C}$ jest górnie trójkątna i ma tylko jedynki na przekątnej. (Istotnie, „dopisana klatka" $\mathbf{C}^t$ jest wtedy dolnie trójkątna i ma wyłącznie jedynki na przekątnej.)

Zadanie 2. a) Dowieść, że formę kwadratową $q \in \mathbb{F}[x_1, ..., x_k]$ można podstawieniem liniowym przeprowadzić w formę postaci $z_1^2 + ... + z_s^2 - z_{s+1}^2 - ... - z_r^2$ gdy $\mathbb{F} = \mathbb{R}$, zaś postaci $z_1^2 + ... + z_r^2$ gdy $\mathbb{F} = \mathbb{C}$; tu $0 \leq r \leq k$. (Wskazówka: wyjść od twierdzenia 2 i dokonać dodatkowych podstawień, w tym zmiany kolejności zmiennych.)

b) Wywnioskować, że w $\mathcal{M}_k(\mathbb{F})$ każda macierz symetryczna jest kongruenta z macierzą postaci $\mathrm{diag}(1, ..., 1, -1, ..., -1, 0, ..., 0)$ gdy $\mathbb{F} = \mathbb{R}$, zaś $\mathrm{diag}(1, ..., 1, 0, ..., 0)$ gdy $\mathbb{F} = \mathbb{C}$. (Liczby jedynek, minus jedynek i zer mogą być zerowe i różne.)

Zadania uzupełniające.

1. Zauważyć, że twierdzenie 1 pozostanie słuszne, gdy dopuścić podstawienia o osobliwej macierzy $\mathbf{C}$. Wykorzystać to do obliczenia wyznacznika macierzy formy $(a_1x_1 + a_2x_2 + a_2x_3)^2 + (b_1x_1 + b_2x_2 + b_3x_3)^2 + (c_1x_1 + c_2x_2 + c_3x_3)^2$.

2. a) Dowieść kongruentności macierzy $\mathrm{diag}(a, b)$ i $\mathrm{diag}((a+b)ab, a+b)$.

b) Dowieść, że gdy macierz symetryczna $\mathbf{A}$ jest nieosobliwa, to macierz $\mathrm{diag}(\mathbf{A}, -\mathbf{A})$ jest kongruentna z macierzą $\mathrm{diag}(\mathbf{I}, -\mathbf{I})$.

3. a) Dowieść, że gdy macierz $\mathbf{A}$ jest nieosobliwa i symetryczna, to $\mathbf{A}^{-1} = \mathbf{CB}^{-1}\mathbf{C}^t$, gdzie $\mathbf{B}$ i $\mathbf{C}$ otrzymano w sposób opisany w tym punkcie.

b) Ponieważ macierz diagonalną $\mathbf{B}$ łatwo jest „odwrócić”, więc daje to pewien sposób obliczania macierzy $\mathbf{A}^{-1}$. Można go użyc do wyznaczenia odwrotności dowolnej macierzy nieosobliwej $\mathbf{X}$ dzięki tożsamości $\mathbf{X}^{-1} = \mathbf{A}^{-1}\mathbf{X}^t$, gdzie macierz $\mathbf{A} := \mathbf{X}^t\mathbf{X}$ jest symetryczna (a dla $\mathbb{F} = \mathbb{R}$ dodatnio określona, patrz §....).

4. (**twierdzenie Kroneckera**) Niech $\mathbf{A} \in \mathcal{M}_k$ będzie macierzą symetryczną, której klatka wyznaczona przez pierwszych r wierszy i kolumn jest nieosobliwa. Dowieść, że formę $q_{\mathbf{A}}$ można podstawieniem liniowym przeprowadzić w formę postaci $\sum_{i,j=1}^{r} a_{ij}y_iy_j + \sum_{i,j=r+1}^{k} b_{ij}y_iy_j$, dla pewnych współczynników b_{ij}. Ponadto, można uzyskać, by podstawienie nie zmieniało zmiennych $x_{r+1}, \ldots, x_k$.
b) Dowieść, że gdy $r = \mathrm{rk}(\mathbf{A})$, to wszystkie współczynniki b_{ij} są równe 0.

5. Niech $\mathbf{A} \in \mathcal{M}_k \setminus \{\mathbf{0}\}$ będzie macierzą symetryczną i niech $r \in \mathbb{N}$. Dowieść, że:
a) Istnieje niezerowy minor główny[2] stopnia 1 lub 2.
b) Jeśli $r \leq k - 2$ i istnieje niezerowy minor główny stopnia r, taki, że wszystkie obejmujące go [2] minory główne stopni $r+1$ i $r+2$ są zerowe, to $\mathrm{rk}(\mathbf{A}) = r$. (Wskazówka: założyć, że minor wyznaczony jest przez początkowych r wierszy i kolumn, po czym wyzerować wszystkie wyrazy pod odpowiadającą mu klatką i obok niej.)
c) Sformułować podobną tezę gdy $r = k - 1$ i dowieść jej i tego, że istnieje niezerowy minor główny stopnia $\mathrm{rk}(\mathbf{A})$.

Zadania ze zbioru Kostrykina: §II.2.2.17.

4. Funkcje wielomianowe stopnia ≤ 2 na przestrzeni wektorowej.

Niech $f : V \to \mathbb{F}$ będzie funkcją na k–wymiarowej przestrzeni wektorowej V nad $\mathbb{F}$.

Definicja. Powiemy, że funkcji tej **w bazie** $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ **odpowiada wielomian** $p \in \mathbb{F}[x_1, ..., x_k]$ (lub: że funkcja f jest w bazie $\mathcal{V}$ **zadana wielomianem** p), jeśli $f(\mathbf{v}) = p([\mathbf{v}]_{\mathcal{V}})$ dla wszystkich $\mathbf{v} \in V$, lub równoważnie:

$$f(x_1\mathbf{v}_1 + ... + x_k\mathbf{v}_k) = p(x_1, ..., x_k) \text{ dla } x_1, ..., x_k \in \mathbb{F} \tag{9}$$

Uwaga 1. a) Gdy, w bazie $\mathcal{V}$, funkcji $f_i : V \to \mathbb{F}$ odpowiada wielomian p_i $(i = 1, 2)$, to sumie $f_1 + f_2$ i iloczynowi f_1f_2 odpowiadają wielomiany $p_1 + p_2$ i p_1p_2, odp.
b) W danej bazie funkcji f odpowiadać może tylko jeden, lub żaden, wielomian stopnia ≤ 2. (Wynika to z twierdzenia 1 w p.1.)
c) Latwo o przykład funkcji, której w żadnej bazie nie w odpowiada wielomian. Gdy $V = \mathbb{R} = \mathbb{F}$, jest nią np. funkcja sin – bo jest niezrowa, lecz ma nieskończenie wiele zer.

Stwierdzenie 1. *Jeśli funkcji f odpowiada w bazie $\mathcal{V}$ wielomian p, to w bazie $\mathcal{V}'$ odpowiada jej wielomian p' powstały z p przez podstawienie $\mathbf{x} = \mathbf{C}\mathbf{y}$, gdzie $\mathbf{C} = [I]_{\mathcal{V}}^{\mathcal{V}'}$.*

[2] patrz dalej zad. uz. 2 w §2.4

Dowód. Dla $\mathbf{v} \in V$ jest $f(\mathbf{v}) = p([\mathbf{v}]_{\mathcal{V}})$ oraz $[\mathbf{v}]_{\mathcal{V}} = \mathbf{C}[\mathbf{v}]_{\mathcal{V}'}$, więc także $f(\mathbf{v}) = p(\mathbf{C}[\mathbf{v}]_{\mathcal{V}'})$. Wraz z definicją p' daje to żądaną tezę. □

Definicja. Powiemy, że funkcja $f : V \to \mathbb{F}$ na przestrzeni wektorowej V jest **kwadratowa** (odpowiednio: jest **wielomianowa** danego stopnia $i \leq 2$, jest **formą kwadratową**), jeśli w pewnej bazie $\mathcal{V}$ odpowiada jej wielomian o tej własności[3]. Ze stwierdzenia 1 i wiadomości z p.1 wynika, że wybór bazy $\mathcal{V}$ nie jest istotny.

b) **Macierzą formy kwadratowej f w bazie $\mathcal{V}$** przestrzeni V nazywamy macierz formy $q \in \mathbb{F}_{\leq 2}[x_1, ..., x_k]$, odpowiadającej f w bazie $\mathcal{V}$. Oznaczamy ją $[f]_{\mathcal{V}}$. Zatem:

$$\mathbf{A} := [f]_{\mathcal{V}} \text{ jest macierzą taką, że } \mathbf{A}^t = \mathbf{A} \text{ i } f(\mathbf{v}) = ([\mathbf{v}]_{\mathcal{V}})^t \mathbf{A} [\mathbf{v}]_{\mathcal{V}} \text{ dla } \mathbf{v} \in V. \tag{10}$$

Uwaga 2. Jak w (9), można to wyrazić i tak: $\mathbf{A}^t = \mathbf{A}$ i $f(\sum_i x_i \mathbf{v}_i) = \mathbf{x}^t \mathbf{A} \mathbf{x} = \sum_{i,j=1}^k a_{ij} x_i x_j$ dla $\mathbf{x} = (x_1, ..., x_k) \in \mathbb{F}^k$. (Tu, $\mathbf{v}_1, ..., \mathbf{v}_k$ to kolejne wektory bazy $\mathcal{V}$.)

Wniosek 1. *Gdy* $\mathbf{A}$ *i* $\mathbf{B}$ *są macierzami formy kwadratowej* f *w bazach* $\mathcal{V}$ *i* $\mathcal{W}$*, odpowiednio, to* $\mathbf{B} = \mathbf{C}^t \mathbf{A} \mathbf{C}$*, gdzie* $\mathbf{C} := [I]_{\mathcal{V}}^{\mathcal{W}}$.

Dowód. Wynika to ze stwierdzenia 1 i twierdzenia 1 z p.2. □

Nadamy teraz twierdzeniu o diagonalizacji form kwadratowych z p.3 nową (lecz równoważną) postać. Potrzebna jest

Definicja. **Rzędem formy kwadratowej** $f : V \to \mathbb{F}$, oznaczanym przez $\mathrm{rk}(f)$, nazywamy rząd jej macierzy w dowolnej bazie przestrzeni V. (Poprawność definicji wynika z zadania 1 d) w p.3 i wniosku 1.) Formę nazywamy **niesosobliwą** lub **niezdegenerowaną**, gdy $\mathrm{rk}(f) = \dim V$, a **osobliwą** lub **zdegenerowaną** w przeciwnym razie.

Twierdzenie 1 (o diagonalizacji formy kwadratowej, wersja dla funkcji)**.** *Niech* $f : V \to \mathbb{F}$ *będzie formą kwadratową na przestrzeni wektorowej* V*. Wówczas istnieje baza tej przestrzeni, w której formie* f *odpowiada wielomian postaci* $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$*, dla pewnych skalarów* $\lambda_1, .., \lambda_k$*.*
Dodatek: *Przy tych oznaczeniach, liczba niezerowych skalarów* λ_i *jest równa* $\mathrm{rk}(f)$*.*

Dowód. Niech $\mathbf{A} \in \mathcal{M}_k$ będzie macierzą formy f w bazie $\mathcal{V} = (\mathbf{v}_i)_{i=1}^k$ przestrzeni V. Na mocy twierdzenia 2 w p.3, istnieje macierz nieosobliwa $\mathbf{C} \in \mathcal{M}_k$, dla której $\mathbf{B} := \mathbf{C}^t \mathbf{A} \mathbf{C}$ jest macierzą diagonalną. Obierzmy bazę $\mathcal{W}$ przestrzeni V tak, by $[I]_{\mathcal{V}}^{\mathcal{W}} = \mathbf{C}$. Z wniosku 1 wynika, że macierz formy f w bazie $\mathcal{W}$ jest równa $\mathbf{B}$, a zatem jest diagonalna. Oznacza to, że w bazie $\mathcal{W}$ formie f odpowiada wielomian postaci $\sum_{i=1}^k \lambda_i x_i^2$, przy czym rząd macierzy $\mathbf{B}$ jest równy liczbie niezerowych wyrazów jej przekątnej $(\lambda_1, ..., \lambda_k)$. □

[3] Nazwa „forma kwadratowa" będzie więc używana zarówno w odniesieniu do wielomianów kilku zmiennych, jak i do funkcji skalarnych na przestrzeniach wektorowych. W wielu podręcznikach unika się tej dwuznaczności tak, że funkcje będące formą kwadratową nazywa się „funkcjonałami kwadratowymi jednorodnymi"; por zadanie 1 w p.1.

Definicja. O bazie $\mathcal{W}$ powiemy, że **diagonalizuje formę kwadratową** f, jeśli macierz $[f]_{\mathcal{W}}$ jest diagonalna. Odnotujmy, że wyżej baza diagonalizująca $\mathcal{W} = (\mathbf{w}_i)_{i=1}^k$ określona była wzorem $\mathbf{w}_j = \sum_{i=1}^k c_{ij}\mathbf{v}_i$ dla $j = 1, ..., k$. (Wynika to z definicji macierzy $[I]_{\mathcal{V}}^{\mathcal{W}}$.)

Uwaga 3. Oba twierdzenia diagonalizacyjne (powyższe i z p.3) nazywane są **twierdzeniem Lagrange'a o diagonalizacji form kwadratowych**.

Ćwiczenie. Dla $i = 1, 2$, niech $\mathbf{A}_i$ będzie macierzą formy kwadratowej f_i w bazie $\mathcal{V}$, zaś $\mathbf{B}_i$ – macierzą tej formy w bazie $\mathcal{W}$. Dowieść, że jeśli macierze te są nieosobliwe, to $\mathrm{tr}(\mathbf{A}_1^{-1}\mathbf{A}_2) = \mathrm{tr}(\mathbf{B}_1^{-1}\mathbf{B}_2)$.

Zadania uzupełniające.

1. Dowieść, że gdy V i W są przestrzeniami wektorowymi, z bazami $\mathcal{V}$ i $\mathcal{W}$, odpowiednio, to złożenie $f \circ L$ formy kwadratowej $f : W \to \mathbb{F}$ z operatorem $L \in \mathcal{L}(V, W)$ jest formą kwadratową, przy czym $[f \circ L]_{\mathcal{V}} = \mathbf{C}^t[f]_{\mathcal{W}}\mathbf{C}$, gdzie $\mathbf{C} := [L]_{\mathcal{W}}^{\mathcal{V}}$.

2. Dowieść, że wyraz a_{ij} macierzy formy kwadratowej f w bazie $(\mathbf{v}_1, ..., \mathbf{v}_k)$ jest równy $\frac{1}{2}(f(\mathbf{v}_i + \mathbf{v}_j) - f(\mathbf{v}_i) - f(\mathbf{v}_j))$. (Wskazówka: wzór na b_{ij} w dowodzie tw. 1 w p.1.)

Zadania ze zbioru Kostrykina: II.2.2.32

§ 2. Przypadek rzeczywistego ciała skalarów i kilka słów o zespolonym.

Poza wnioskiem 1 w p.2, gdzie rozpatrujemy również przypadek zespolony, w paragrafie tym zakładamy, że $\mathbb{F} = \mathbb{R}$. Znaczenie rzeczywistych funkcji kwadratowych w Analizie bierze się m.in. stąd, że gładkie funkcje $\mathbb{R}^k \to \mathbb{R}$ można aproksymować ich rozwinięciami Taylora drugiego stopnia, a te są funkcjami kwadratowymi. Wykorzystując własności ciała $\mathbb{R}$, w tym jego uporządkowanie relacją $<$, możemy też dla $\mathbb{F} = \mathbb{R}$ uzyskać o funkcjach kwadratowych więcej informacji, niż w ogólnym przypadku.

1. Określoność form kwadratowych i macierzy symetrycznych ($\mathbb{F} = \mathbb{R}$).

Niech f będzie formą kwadratową na rzeczywistej przestrzeni liniowej V.

Definicja. Powiemy, że forma f jest **dodatnio określona**, jeśli $f(\mathbf{v}) > 0$ dla wszystkich $\mathbf{v} \in V \setminus \{\mathbf{0}\}$, a jest **ujemnie określona**, jeśli $f(\mathbf{v}) < 0$ dla wszystkich $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. O macierzy symetrycznej $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ powiemy, że jest dodatnio (odp. ujemnie) określona, gdy forma $f_{\mathbf{A}} : \mathbb{R}^k \to \mathbb{R}$ ma tę własność. Gdy któryś z tych warunków jest spełniony przy ostrym znaku nierówności zastąpionym przez tępy, to mówimy, że forma lub macierz jest **dodatnio** (odp. **ujemnie) półokreślona**. W pozostałym przypadku formę czy macierz nazywamy **nieokreśloną.**

Przykład 1. Forma $f(x_1, x_2, x_3) = (x_1 - x_2)^2 + (x_1 - 2x_3)^2$ jest dodatnio półokreślona, a w ślad za nią taka jest jej macierz

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & -2 \\ -1 & 1 & 0 \\ -2 & 0 & 4 \end{pmatrix}$$

Nie są one dodatnio określone, bo $f(2, 2, 1) = 0$.

Uwaga 1. a) Jeśli $\mathbf{A}$ jest macierzą formy f w pewnej bazie $\mathcal{V}$ przestrzeni V, to f ma którąś ze zdefiniowanych wyżej własności wtedy i tylko wtedy, gdy ma ją $\mathbf{A}$. (Korzystamy z tego, że $f(\mathbf{v}) = [\mathbf{v}]^t_{\mathcal{V}}\mathbf{A}[\mathbf{v}]_{\mathcal{V}} = f_{\mathbf{A}}([\mathbf{v}]_{\mathcal{V}})$ dla $\mathbf{v} \in V$.)

b) Wynika stąd, że gdy macierze symetryczne $\mathbf{A}, \mathbf{B}$ są kongruentne i $\mathbf{A}$ ma którąś z tych własności, to i $\mathbf{B}$ ją ma (bo jest macierzą formy $f_{\mathbf{A}}$ w pewnej bazie przestrzeni $\mathbb{R}^k$).

c) Jeśli więc badamy macierz symetryczną ze względu na określoność czy półokreśloność, to wolno nam zastąpić ją przez dowolną kongruentną z nią macierz diagonalną. Ta zaś jest określona dodatnio (odp. ujemnie) wtedy i tylko wtedy, gdy wszystkie wyrazy jej przekątnej są dodatnie (odp. ujemne); podobnie jest dla półokreśloności (nierówności są wtedy tępe).

d) Gdy macierze symetryczne $\mathbf{A}$ i $\mathbf{B}$ są dodatnio określone, to $\operatorname{diag}(\mathbf{A}, \mathbf{B})$ też, i vice versa. Tak samo dla określoności ujemnej i obu półokreśloności. ⊟

Przykład 2. Macierz z przykładu 1 w §1.3 nie jest półokreślona, bo kongruentna z nią macierz diagonalna $\mathbf{B}$ ma na przekątnej wyrazy różnych znaków.

Uwaga 2. Z części d) uwagi 1 wynika, że gdy macierz symetryczna jest półokreślona i nieosobliwa, to jest określona, i odwrotnie. (Bo jest tak dla macierzy diagonalnych).

Dla małych k, a także w zastosowaniach teoretycznych, użyteczne może być wyznacznikowe kryterium określoności formy. By je sformułować umówmy się nazywać minor macierzy **początkowym**, gdy jest wyznaczony przez pierwszych jej s wierszy i kolumn, dla pewnej liczby s.

Twierdzenie 1. *Rzeczywista macierz symetryczna wtedy i tylko wtedy jest dodatnio określona, gdy wszystkie jej minory początkowe są dodatnie.*

Dowód. Oznaczmy macierz przez $\mathbf{A}$. Ponieważ $a_{11} = q_{\mathbf{A}}(\mathbf{e}_1)$, więc każdy z rozważanych warunków implikuje $a_{11} > 0$. Zakładamy więc niżej, że $a_{11} > 0$. Wówczas krok 1 algorytmu z §1.3 przeprowadza macierz $\mathbf{A}$ w macierz $\mathbf{B}$ postaci $\begin{pmatrix} a_{11} & 0 \\ 0 & \mathbf{K} \end{pmatrix}$, kongruentną z $\mathbf{A}$ i mającą te same co ona minory początkowe. (Patrz uwaga 1 w §1.3). Wobec tego:

1) Minory początkowe macierzy $\mathbf{A}$ wtedy i tylko wtedy wszystkie są dodatnie, gdy jest to prawdą dla $\mathbf{K}$. (Korzystamy z tego, że i–ty minor początkowy macierzy $\mathbf{B}$ jest iloczynem $i - 1$–szego minora początkowego macierzy $\mathbf{K}$ i liczby dodatniej a_{11}.)

Teza twierdzenia, oczywista dla $k = 1$, wynika więc przez indukcję względem k, bo:

2) Na mocy części b) i d) uwagi 1, macierz $\mathbf{A}$ wtedy i tylko wtedy jest dodatnio określona, gdy macierz $\mathbf{K}$ ma tę własność. (Gra rolę to, że $a_{11} > 0$.) $\square$

Z równości $\det(-\mathbf{B}) = (-1)^s \det(\mathbf{B})$ dla $\mathbf{B} \in \mathcal{M}_s$ i twierdzenia 1, zastosowanego do macierzy $-\mathbf{A}$, otrzymujemy

Wniosek 1. *Rzeczywista macierz symetryczna jest ujemnie określona wtedy i tylko wtedy, gdy jej minory początkowe stopnia nieparzystego są ujemne, a parzystego dodatnie.*

Twierdzenie 1 i wniosek 1 noszą nazwę **kryterium Sylvestera–Jacobiego**.[4] Zapiszmy je dla

$$\mathbf{A} = \begin{pmatrix} \underline{a|} & b| & d| \\ \underline{b} & \underline{c|} & e| \\ \underline{d} & \underline{e} & \underline{f|} \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}) \quad \text{(zaznaczono klatki początkowe)}$$

a) macierz $\mathbf{A}$ jest określona dodatnio$\Leftrightarrow$($a > 0$ i $\det \begin{pmatrix} a & b \\ b & c \end{pmatrix} > 0$ i $\det(\mathbf{A}) > 0$);

b) macierz $\mathbf{A}$ jest określona ujemnie$\Leftrightarrow$($a < 0$ i $\det \begin{pmatrix} a & b \\ b & c \end{pmatrix} > 0$ i $\det(\mathbf{A}) < 0$).

Uwaga 3. Już dla 4×4–macierzy kryterium Sylvestera–Jacobiego jest znacznie trudniej stosować, niż kryterium z uwagi 1c). Nie wolno też ulec pokusie zamiany w założeniu i tezie słów „dodatnio" i „dodatnie" przez „nieujemnie" i „nieujemne", odpowiednio, o czym zaświadcza macierz $\mathrm{diag}(0, 1, -1)$.

Zadania uzupełniające.

1. Dowieść, że odwrotność macierzy dodatnio określonej jest dodatnio określona.

2. Niech $\mathbf{A}$ będzie macierzą symetryczną nad $\mathbb{F}$ (tu niekoniecznie $\mathbb{F} = \mathbb{R}$) i niech r oznacza jej rząd, zaś a_i jej i-ty minor początkowy. Dowieść, że jeśli $a_i \neq 0$ dla $i = 1, ..., r$, to $\mathbf{A} = \mathbf{C}^t\mathbf{D}\mathbf{C}$, gdzie $\mathbf{D} = \mathrm{diag}(a_1, a_2/a_1, ..., a_r/a_{r-1}, 0, ..., 0)$, a $\mathbf{C}$ jest pewną macierzą górnie trójkątną, z jedynkami na przekątnej. (Wskazówka: dowód tw. 1.)

Uwaga 4. Rezultat ten nosi nazwę **twierdzenia Jacobiego.** Wynika z niego, że jeśli macierz symetryczna $\mathbf{A}$ ma wszystkie minory początkowe $a_1, \ldots, a_r$ ($r = \mathrm{rk}(\mathbf{A})$) niezerowe, to formę $q_{\mathbf{A}}$ można liniową zamianą zmiennych przeprowadzić w formę $\sum_i (a_i/a_{i-1})x_i^2$, gdzie przyjmujemy $a_0 := 1$.

[4]Na ogół kryterium przypisuje się (tylko) Sylvesterowi, choć jest ku temu chyba równie mało powodów, jak przypisywanie go Jacobiemu. Por. http://math.sfsu.edu/smith/Math880/General/Epilog.pdf i uwaga 4.

3. Dla symetrycznej macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ dowieść implikacji $a) \Rightarrow b) \Rightarrow c) \Rightarrow d)$ oraz b)$\Rightarrow$a) i b)$\Rightarrow$e), dotyczących poniższych warunków. (Wskazówka: w dowodach, że b)$\Rightarrow$ e) i b)$\Rightarrow$ a), wykorzystać d) i e), odpowiednio.)

a) Macierz $\mathbf{A}$ jest dodatnio półokreślona;

b) Każdy **minor główny** macierzy $\mathbf{A}$ (tzn. minor wyznaczony przez jej wiersze i kolumny należące do tego samego podzbioru zbioru $\{1, ..., k\}$) jest nieujemny.

c) $a_{ii} \geq 0$ i $|a_{ij}| \leq \sqrt{a_{ii}a_{jj}} \leq (a_{ii} + a_{jj})/2$ dla $i, j = 1, ..., k$.

d) Jeśli $a_{ii} = 0$, to $a_{ij} = a_{ji} = 0$ $(i, j = 1, ..., k)$.

e) W algorytmie z §1.3, zastosowanym do macierzy $\mathbf{A}$, część 1 każdego kroku jest pomijana.

4. a) Dla każdej macierzy $\mathbf{B} \in \mathcal{M}_k(\mathbb{R})$, macierz $\mathbf{B}^t\mathbf{B}$ jest symetryczna i dodatnio półokreślona.

b) Gdy macierz symetryczna $\mathbf{A}$ jest dodatnio półokreślona, to $\mathbf{A} = \mathbf{B}^t\mathbf{B}$ dla pewnej macierzy górnie trójkątnej $\mathbf{B}$ o nieujemnych wyrazach na przekątnej. (Por. zadanie uzupełniające 1 w §V.3.4.)

Zadania ze zbioru Kostrykina: 9, 11, 13, 31* w §II.2.2.

2. Twierdzenie o bezwładności.

Danej formie kwadratowej $f : V \to \mathbb{R}$ na rzeczywistej przestrzeni wektorowej V odpowiadają w różnych bazach diagonalizujących różne wielomiany postaci $\sum_i \lambda_i x_i^2$. Pokażemy jednak, że liczby dodatnich i ujemnych współczynników λ_i są jednoznacznie przez f wyznaczone.

Twierdzenie 1 (J.J.Sylvestera o bezwładności, trzy wersje)**.** *a) Gdy formie kwadratowej $f : V \to \mathbb{R}$ na rzeczywistej przestrzeni wektorowej V odpowiada w pewnej bazie $\mathcal{V} = (\mathbf{v}_i)_{i=1}^k$ wielomian $\sum_{i=1}^k \lambda_i x_i^2$, to liczba s dodatnich współczynników λ_i jest od bazy $\mathcal{V}$ niezależna, i tak samo jest dla współczynników ujemnych i dla równych 0.*

b) Gdy rzeczywiste macierze diagonalne sa kongruentne, to mają tę samą liczbę wyrazów dodatnich, i tak samo dla wyrazów ujemnych i dla równych 0.

c) Gdy podstawienia liniowe przeprowadzają pewną formę $q \in \mathbb{R}_{\leq 2}[x_1, ..., x_k]$ w formy $\lambda_1 y_1^2 + ... + \lambda_k y_k^2$ i $\lambda_1' y_1^2 + ... + \lambda_k' y_k^2$, odpowiednio, to w ciągu $\lambda_1, ..., \lambda_k$ jest tyle wyrazów dodatnich, co w $\lambda_1', ..., \lambda_k'$. Tak samo jest też z wyrazami ujemnymi i z równymi 0.

Dowód. Wersja c) wynika z b), bo macierze $\text{diag}(\lambda_1, ..., \lambda_k)$ i $\text{diag}(\lambda_1', ..., \lambda_k')$ w c) są kongruentne. Z kolei, b) wynika z a), bo kongruentne macierze są macierzami, w różnych bazach, pewnej wspólnej formy kwadratowej $\mathbb{R}^k \to \mathbb{R}$. (Korzystamy z wniosku 1 w §1.4.) Pozostaje dowieść a), i to tylko w odniesieniu do liczby s, by potem jej niezależność od bazy zastosować do formy $-f$. Teza wynika więc z poniższego lematu, wyrażającego s w sposób niezależny od bazy:

Lemat 1. *Przy oznaczeniach części a) twierdzenia, s jest maksymalnym wymiarem liniowych podprzestrzeni przestrzeni V, na których forma f jest dodatnio określona.*

Dowód. Niech dodatnimi współczynnikami będą $\lambda_1, ..., \lambda_s$. Gdy $\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_s\mathbf{v}_s$, gdzie $c_i \neq 0$ dla pewnego i, to $f(\mathbf{u}) = \lambda_1 c_1^2 + \cdots + \lambda_s c_s^2 > 0$; patrz (9). Zatem:

$$\text{przy } U := \mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_s), \text{ forma } f_{|U} \text{ jest dodatnio określona i } \dim(U) = s. \tag{11}$$

Z drugiej strony, gdy podprzestrzeń $W \subset V$ jest wymiaru większego niż s, to na postawie wniosku 1 w §III.6.1 zawiera niezerowy wektor $\mathbf{w} \in \mathrm{lin}\{\mathbf{v}_{s+1}, ..., \mathbf{v}_k\}$. Forma $f_{|W}$ nie jest więc wtedy dodatnio określona, bo $f(\mathbf{w}) \leq 0$ (uzasadnienie takie, jak dodatniości $f(\mathbf{u})$). To kończy dowód lematu. $\square$

Uwaga 1. Podobne rozumowania pozwalają też wyznaczyć maksymalny wymiar podprzestrzeni, na których forma f jest dodatnio półokreślona czy zerowa; patrz niżej zadanie uzupełniające 3.

Definicja. a) **Dodatnim** (odp.: **ujemnym**) **indeksem bezwładności** formy kwadratowej $f : V \to \mathbb{R}$ nazywamy liczbę dodatnich (odp. ujemnych) wyrazów macierzy tej formy w dowolnej bazie diagonalizującej. (Poprawność definicji wynika z wersji a) twierdzenia.) Oznaczamy je przez $\sigma_+(f)$ i przez $\sigma_-(f)$, odpowiednio. Parę $(\sigma_+(f), \sigma_-(f))$ oznaczamy przez $\sigma(f)$ i nazywamy **sygnaturą formy** f.[5]

b) Podobnie definiujemy i oznaczamy indeksy bezwładności i sygnaturę macierzy $\mathbf{A}$, która jest rzeczywista i symetryczna: $\sigma_+(\mathbf{A})$ (odp. $\sigma_-(\mathbf{A})$) jest liczbą dodatnich (odp. ujemnych) wyrazów na przekątnej dowolnej macierzy diagonalnej, kongruentnej z $\mathbf{A}$, zaś $\sigma(\mathbf{A}) := (\sigma_+(\mathbf{A}), \sigma_-(\mathbf{A}))$.

Przykład 1. Macierz z przykładu 1 w §1.3 ma sygnaturę $(3, 1)$, ponieważ jest kongruentna z macierzą diagonalną o 3 wyrazach dodatnich i 1 ujemnym.1 Wyznaczyć sygnaturę formy $\det : \mathcal{M}_2(\mathbb{R}) \to \mathbb{R}$ oraz jej obcięcia do $\{\mathbf{A} \in \mathcal{M}_2 : \mathrm{tr}(\mathbf{A}) = 0\}$.

Uwaga 2. a) Suma dodatniego i ujemnego indeksu bezwładności jest równa rzędowi (formy czy macierzy). Patrz „Dodatek" w twierdzeniu 1 z §1.4.

b) Jeśli rzeczywiste macierze symetryczne są kongruentne, to mają tę samą sygnaturę. (Wynika to z definicji sygnatury i przechodniości kongruentności.)

c) Gdy $\mathbf{A}$ jest macierzą formy f w pewnej bazie, to $\sigma(f) = \sigma(\mathbf{A})$. Istotnie, dla baz diagonalizujących wynika to z definicji, a dla innych – z b), bo macierze formy f w różnych bazach są kongruentne.

Wniosek 1. *Warunkiem koniecznym i dostatecznym, by macierze symetryczne $\mathbf{A}, \mathbf{B}$ były kongruentne w $\mathcal{M}_k(\mathbb{F})$, jest równość $\sigma(\mathbf{A}) = \sigma(\mathbf{B})$ gdy $\mathbb{F} = \mathbb{R}$, zaś równość $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{B})$ gdy $\mathbb{F} = \mathbb{C}$.*

[5] Często sygnaturą formy f nazywana bywa liczba $\sigma_+(f) - \sigma_-(f)$, która wraz z $\mathrm{rk}(f)$ wyznacza parę $\sigma(f)$; por uwaga 2.

Dowód. Konieczność tych warunków odnotowano już w zadaniu 1d) z §1.3 i uwadze 2b).

Dowiedziemy dostateczności, gdy $\mathbb{F} = \mathbb{C}$. Macierz $\mathbf{A}$ jest kongruentna z macierzą diagonalną, na przekątnej której stoją wpierw jedynki, a potem zera; przy tym liczba jedynek jest równa $\mathrm{rk}(\mathbf{A})$. (Korzystamy tu z zadań 2 i 1d) z §1.3.) Analogicznie jest dla macierzy $\mathbf{B}$. Jeśli więc $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{B})$, to $\mathbf{A}$ i $\mathbf{B}$ są kongruentne z tą samą macierzą diagonalną, a więc i jedna z drugą.

Gdy $\mathbb{F} = \mathbb{R}$ rozumujemy tak samo, wykorzystując część b) twierdzenia 1. ⊟

Uwaga 3. W przypadku dowolnego ciała $\mathbb{F}$ na ogół trudno jest ustalić, czy dane dwie macierze symetryczne $\mathbf{A}, \mathbf{B}$ są w $\mathcal{M}_k(\mathbb{F})$ kongruentne.(Dla ciała $\mathbb{Q}$ liczb wymiernych dyskusji tych zagadnień poświęcona jest książka.) Warunkiem koniecznym jest oczywiście, by $\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{B})$ i iloczyn $\det(\mathbf{A}) \cdot \det(\mathbf{B})$ był był kwadratem w $\mathbb{F}$. (Dlaczego?) Nie jest to jednak warunek wystarczający.

Zadania uzupełniające. (Poza ostatnim zadaniem, ciałem skalarów jest $\mathbb{R}$.)

1. a) Wyrazić znak liczby $\det(\mathbf{A})$ przez $\sigma(\mathbf{A})$.

b) Gdy macierz $\mathbf{A}$ jest nieosobliwa, wyrazić $\sigma(-\mathbf{A}^{-1})$ przez $\sigma(\mathbf{A})$.

2. Niech $p = \sum_{i=1}^{r} \lambda_i x_i^2 + \sum_{i=1}^{k} c_k x_k$, przy czym współczynniki $\lambda_1, ..., \lambda_s$ są dodatnie, dla pewnego $s \geq 1$. Dowieść, że jeśli $r < k$ i $c_k \neq 0$, to dla każdej podprzestrzeni liniowej $V_0 \subset \mathbb{R}^k$ takiej, że $\dim(V_0) > k - s - 1$, zachodzi $p(V_0) \supset [0, \infty)$.

Ponieważ zadanie to będzie wykorzystane w rozdziale VIII, więc daję wskazówke: gdy $\mathbf{v} \in \mathrm{lin}(\mathbf{e}_1, ..., \mathbf{e}_s, \mathbf{e}_k)$ to funkcja $\mathbb{R} \ni t \mapsto p(t\mathbf{v})$ przyjmuje wszystkie wartości nieujemne.

3. Niech $f : V \to \mathbb{R}$ będzie formą kwadratową i niech $\dim(V) = k$ i $\sigma(f) = (s, t)$. Dowieść, że:

a) Maksimum wymiarów podprzestrzeni, na których forma f jest dodatnio półokreślona, jest równe $k - t$.

b) Maksimum wymiarów podprzestrzeni, na których forma f jest zerowa, jest równe $k - \max(s, t)$.

4. a) Niech $V = U \oplus W$, a forma kwadratowa $f : V \to \mathbb{R}$ będzie dodatnio określona na U i ujemnie półokreślona na W. Udowodnić, że $\sigma_+(f) = \dim(U)$.

b) Wyznaczyć rząd i sygnaturę formy $f : \mathcal{M}_k(\mathbb{R}) \to \mathbb{R}$, danej wzorem $f(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^2)$.

5. Niech W będzie podprzestrzenią przestrzeni V, a $f : V \to \mathbb{R}$ formą kwadratową. Dowieść, że:

a) $f_{|W} : W \to \mathbb{R}$ jest formą kwadratową i $\sigma_+(f_{|W}) \leq \sigma_+(f)$, $\sigma_-(f_{|W}) \leq \sigma_-(f)$.

b) $\sigma_+(f) - \sigma_+(f_{|W}) \leq \dim(V) - \dim(W)$, i tak samo dla σ_-.

6. Niech $p \in \mathbb{R}_{\leq 2}[x_1, ..., x_k]$ będzie wielomianem o części głównej q.

a) Dowieść, że gdy $\mathbf{v}, \mathbf{w} \in \mathbb{R}^k$ i $q(\mathbf{v}) > 0$, to $\sup_{t \in \mathbb{R}} p(t\mathbf{v} + \mathbf{w}) = \infty$.

b) Wywnioskować, że dla każdego wektora $\mathbf{w} \in \mathbb{R}^k$ i każdej podprzestrzeni liniowej $V_0 \subset V$ takiej, że $\dim(V_0) > k - \sigma_+(q)$, zachodzi $\sup_{\mathbf{v}\in V_0} p(\mathbf{v}+\mathbf{w}) = \infty$.

7. Niech $q \in \mathbb{R}[x_1, ..., x_k]$ będzie formą kwadratową. Dowieść, że warunek $\sigma(q) = (s, t)$ jest równoważny temu, by $q = \ell_1^2 + ... + \ell_s^2 - \ell_{s+1}^2 - ... - \ell_{s+t}^2$ dla pewnych liniowo niezależnych form liniowych $\ell_1, ..., \ell_{s+t} \in \mathbb{R}_{\le 1}[x_1, ..., x_k]$. (Wielomian liniowy ℓ nazywamy formą, gdy $\ell(\mathbf{0}) = 0$.)

8. Rozważmy następujące własności niezerowej formy kwadratowej $q \in \mathbb{F}[x_1, ..., x_k]$:
a) q jest kwadratem wielomianu stopnia 1;
b) q jest iloczynem dwóch wielomianów stopnia 1.
Dowieść, że dla $\mathbb{F} = \mathbb{C}$ własność a) jest równoważna temu, by $\mathrm{rk}(q) = 1$, a b) temu, by $\mathrm{rk}(q) \in \{1, 2\}$; zaś dla $\mathbb{F} = \mathbb{R}$ własność a) jest równoważna temu, by $\sigma(q) = (1, 0)$, a b) temu, by $\sigma(q) \in \{(1, 0), (0, 1), (1, 1)\}$.

Zadania ze zbioru Kostrykina: §II.2.1.10 oraz 1, 2, 7, 16, 18, 20, 21 i 27 * w §II.2.2. (Stosować dowolną metodę ustalania równoważności.)

3. Ortogonalna diagonalizacja rzeczywistych form kwadratowych.

Dla macierzy ortogonalnej $\mathbf{C}$ mamy $\mathbf{C}^{-1} = \mathbf{C}^t$. Ortogonalnie podobne macierze są więc kongruentne, co umożliwia wyrażenie w języku form twierdzenia o ortogonalnej diagonalizowalności macierzy symetrycznych. Ta nowa interpretacja wyników z rozdziału VI pozwala na uzyskanie wielu dodatkowych informacji o formie czy macierzy symetrycznej, patrz m.in. poniższy wniosek 1, zadania z p.4, a także tw.2 z §3.2.

Definicja. Gdy $\mathbf{C} \in \mathcal{M}_k(\mathbb{R})$ jest macierzą ortogonalną i podstawienie $\mathbf{x} = \mathbf{C}\mathbf{y}$ przeprowadza formę $q \in \mathbb{R}_{\le 2}[x_1, ..., x_k]$ w formę q' to mówimy, że q przeprowadzono w q' **podstawieniem ortogonalnym** (lub: **ortogonalną zamianą zmiennych**).

Twierdzenie 1 (o ortogonalnej diagonalizacji form, dwie wersje). *a) Daną formę $q \in \mathbb{R}_{\le 2}[x_1, .., x_k]$ można podstawieniem ortogonalnym przeprowadzić w formę postaci $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$. Ciąg liczb $\lambda_1, ..., \lambda_k \in \mathbb{R}$ jest z dokładnością do kolejności wyznaczony jednoznacznie: są w nim wszystkie wartości własne macierzy formy q, każda powtórzona tylekroć, ile wynosi jej krotność jako pierwiastka wielomianu charakterystycznego.*

b) Danej formie kwadratowej na przestrzeni euklidesowej E odpowiada w pewnej bazie ortonormalnej wielomian postaci $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$, gdzie $k = \dim(E)$. Współczynniki $\lambda_1, ..., \lambda_k \in \mathbb{R}$ są wartościami własnymi macierzy formy w dowolnej bazie ortonormalnej, powtarzanymi zgodnie z ich krotnościami, jak wyżej.

Dowód. Ad a). Ponieważ macierz $\mathbf{A}$ formy q jest symetryczna i rzeczywista, więc istnieje macierz ortogonalna $\mathbf{C}$ taka, że $\mathbf{D} := \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ jest macierzą diagonalną. Wówczas

$\mathbf{C}^t\mathbf{A}\mathbf{C} = \mathbf{D}$, skąd $\mathbf{x} = \mathbf{C}\mathbf{y}$ jest żądanym podstawieniem. Dalsza część a) wynika z podobieństwa macierzy $\mathbf{A}$ i $\mathbf{D} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, por. uwaga 5 w §VI.2.3.

Ad b). Wystarcza powtórzyć uzasadnienie twierdzenia z §1.4. (Tym razem zaczynamy od bazy ortonormalnej, a za $\mathbf{C}$ obieramy macierz ortogonalną). ⊟

Wniosek 1. *a) Sygnatura rzeczywistej macierzy symetrycznej wynosi (s, t), gdzie s jest liczbą dodatnich, a t liczbą ujemnych wartości własnych tej macierzy. (Tu i niżej uwzględniamy krotności algebraiczne wartości własnych.)*

b) Sygnatura formy kwadratowej na przestrzeni rzeczywistej wynosi (s, t), gdzie s to liczba dodatnich, a t –ujemnych wartości własnych macierzy formy w dowolnej bazie przestrzeni.

Dowód. a) wynika z części a) twierdzenia, zaś b) –z a) i uwagi 2c) w p.2. □

Zadania uzupełniające. Dla rzeczywistych macierzy udowodnić, że:

1. Podobne macierze symetryczne są kongruentne.

2. a) Gdy macierz $\mathbf{S}$ jest symetryczna i nieosobliwa, a $\mathbf{B}$ – dowolna, to macierz $\mathbf{S}^2\mathbf{B}$ jest podobna do macierzy $\mathbf{SBS}$, kongruentnej z $\mathbf{B}$.

b) Gdy $\mathbf{A}$ i $\mathbf{B}$ są symetryczne, przy czym $\mathbf{A}$ dodatnio określona, to $\mathbf{AB}$ ma tylko rzeczywiste wartości własne, w tym tyle samo dodatnich (odp. ujemnych), co $\mathbf{B}$.

Zadania ze zbioru Kostrykina: §II.4.3: 18 i 19. („Sprowadzić formę na osie główne" to znaleźć podstawienie ortogonalne, diagonalizujące tę formę.)

4. * Wartości własne rzeczywistych macierzy symetrycznych (zadania uzupełniające).

1. Niech $f : \mathbb{R}^k \to \mathbb{R}$ będzie formą kwadratową, a $(\lambda_1, \ldots, \lambda_k)$ ciągiem wszystkich wartości własnych (z powtórzeniami) jej macierzy w ortonormalnej bazie przestrzeni $\mathbb{R}^k$. Niech też $S := \{\mathbf{v} \in \mathbb{R}^k : \|\mathbf{v}\| = 1\}$ oznacza sferę jednostkową. Dowieść, że:

a) $\sup f(S) = \sup_i \lambda_i$ i $\inf f(S) = \inf_i \lambda_i$.

b) Dla $\lambda \in \mathbb{R}$, liczba $\#\{i : \lambda_i \geq \lambda\}$ jest równa maksimum wymiarów podprzestrzeni W takich, że $f_{|W \cap S} \geq \lambda$.

c) Podobnie, liczba $\#\{i : \lambda_i \leq \lambda\}$ jest równa maksimum wymiarów podprzestrzeni W takich, że $f_{|W \cap S} \leq \lambda$.

d) Jeśli $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_k$, to dla $1 \leq i \leq k$ mają miejsce następujące **równości Couranta–Fischera**:

$$\lambda_i = \inf_W C_W \quad \text{i} \quad \lambda_{k-i+1} = \sup_W c_W,$$

gdzie W przebiega i–wymiarowe podprzestrzenie przestrzeni $\mathbb{R}^k$, zaś c_W i C_W oznaczają dla każdej takiej podprzestrzeni kresy zbioru $f(S \cap W)$, dolny i górny, odpowiednio.

2. Niech dalej $f' := f_{|V'}$ oznacza obcięcie formy f do pewnej podprzestrzeni $V' \subset V$, i niech $\lambda_1 \leq \lambda_2 ... \leq \lambda_k$ i $\lambda'_1 \leq ... \leq \lambda'_l$ będą wartościami własnymi macierzy form f i f' w pewnych bazach ortonormalnych przestrzeni V i V', odpowiednio. Dowieść, że $\lambda_i \leq \lambda'_i \leq \lambda_{i+k-l}$ dla $i = 1, ..., l$. (Jest to **twierdzenie Cauchy'ego o przeplataniu**; gdy $l = k - 1$ mówi ono, że $\lambda'_i \in [\lambda_i, \lambda_{i+1}]$ dla $i = 1, ..., k - 1$.)

3. Niech macierze symetryczne $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{R})$ mają tę własność, że $f_{\mathbf{A}} \leq f_{\mathbf{B}}$ (tzn. $\mathbf{v}^t\mathbf{A}\mathbf{v} \leq \mathbf{v}^t\mathbf{B}\mathbf{v}$, $\forall \mathbf{v} \in \mathbb{R}^k$). Dowieść, że:

a) Liczba dodatnich wartości własnych macierzy $\mathbf{B}$ jest niemniejsza, niż liczba dodatnich wartości własnych $\mathbf{A}$, i tak samo dla wartości nieujemnych. (Uwzględniamy krotności wartości własnych.)

b) Jeśli $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$ i $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k$ są wszystkimi pierwiastkami wielomianów $\chi_{\mathbf{A}}$ i $\chi_{\mathbf{B}}$, odpowiednio, to $\lambda_i \leq \mu_i$ dla $i = 1, ..., k$.

§ 3. Formy (funkcje) dwuliniowe.

1. Funkcje dwuliniowe i ich macierze.

Definicja. Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Funkcję $g : V \times V \to \mathbb{F}$ nazywamy **dwuliniową**, gdy

$$\text{dla każdego } \mathbf{v} \in V, \text{ funkcje } \mathbf{u} \mapsto g(\mathbf{u}, \mathbf{v}) \text{ i } \mathbf{u} \mapsto g(\mathbf{v}, \mathbf{u}) \text{ są liniowe.} \tag{12}$$

Funkcja dwuliniowa $V \times V \to \mathbb{F}$ nazywana jest też **formą dwuliniową**, przy czym mówi się o dwuliniowej funkcji czy formie **na** przestrzeni V. Jest to ogólnie przyjęte, choć nieco mylące: dziedziną nie jest tu bowiem przestrzeń V, lecz $V \times V$.

Zadanie 1. Dla takiej funkcji g ma miejsce tożsamość

$$g(\sum_{i \in I} x_i \mathbf{u}_i, \sum_{j \in J} y_j \mathbf{w}_j) = \sum_{i \in I, j \in J} x_i y_j g(\mathbf{u}_i, \mathbf{w}_j) \tag{13}$$

gdy $(x_i)_{i \in I}$ i $(y_j)_{j \in J}$ są skończonymi układami skalarów, a $(\mathbf{u}_i)_{i \in I}$ i $(\mathbf{w}_j)_{j \in J}$ – układami wektorów przestrzeni V.

Definicja. **Macierzą funkcji dwuliniowej** $g : V \times V \to \mathbb{F}$ **w bazie** $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ przestrzeni V nazywamy macierz $(g(\mathbf{v}_i, \mathbf{v}_j))_{i,j=1}^k$. Gdy wygodnie, oznaczamy ją $[g]_{\mathcal{V}}$.

Stwierdzenie 1. *Niech* $\mathbf{A}$ *i* $\mathbf{B}$ *będą macierzami funkcji dwuliniowej* $g : V \times V \to \mathbb{F}$ *w bazach* $\mathcal{V}$ *i* $\mathcal{W}$*, odpowiednio. Wówczas:*

$$g(\mathbf{u}, \mathbf{w}) = \sum_{i,j=1}^{k} a_{ij} x_i y_j = \mathbf{x}^t \mathbf{A} \mathbf{y} \quad \textit{dla } \mathbf{u}, \mathbf{w} \in V, \ \textit{gdzie } \mathbf{x} := [\mathbf{u}]_{\mathcal{V}}, \ \mathbf{y} := [\mathbf{w}]_{\mathcal{V}} \tag{14}$$

$$\mathbf{B} = \mathbf{C}^t \mathbf{A} \mathbf{C}, \ \textit{gdzie } \mathbf{C} = [I]_{\mathcal{V}}^{\mathcal{W}} \textit{ jest macierzą zmiany baz.} \tag{15}$$

Dowód. Równość (14) wynika z zadania 1, bo $\mathbf{u} = \sum_{i=1}^k x_i \mathbf{v}_i$ i $\mathbf{w} = \sum_{j=1}^k y_j \mathbf{v}_j$.

Ad (15). Dla $\mathbf{u}, \mathbf{w} \in V$ mamy więc $([\mathbf{u}]_\mathcal{V})^t \mathbf{A} [\mathbf{v}]_\mathcal{V} = g(\mathbf{u}, \mathbf{w}) = ([\mathbf{u}]_\mathcal{W})^t \mathbf{B} [\mathbf{w}]_\mathcal{W}$ oraz $[\mathbf{u}]_\mathcal{V} = \mathbf{C}[\mathbf{u}]_\mathcal{W}$ i $[\mathbf{w}]_\mathcal{V} = \mathbf{C}[\mathbf{w}]_\mathcal{W}$. To daje $\mathbf{x}^t \mathbf{B} \mathbf{y} = \mathbf{x}^t \mathbf{C}^t \mathbf{A} \mathbf{C} \mathbf{y}$ dla $\mathbf{x}, \mathbf{y} \in \mathbb{F}^k$, wobec czego $\mathbf{B} = \mathbf{C}^t \mathbf{A} \mathbf{C}$. (Por. §1.4.) $\boxminus$

Zależność (14) wyrażamy mówiąc, że w bazie $\mathcal{V}$, funkcja g jest **zadana** wielomianem $\mathbf{x}^t \mathbf{A} \mathbf{y} \in \mathbb{F}_{\leq 2}[x_1, ... x_k, y_1, ..., y_k]$.

Definicja. Funkcję $g : V \times V \to \mathbb{F}$ nazywamy

symetryczną, gdy $g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u})$ dla wszystkich $\mathbf{u}, \mathbf{v} \in V$,

antysymetryczną, gdy $g(\mathbf{u}, \mathbf{v}) = -g(\mathbf{v}, \mathbf{u})$ dla wszystkich $\mathbf{u}, \mathbf{v} \in V$.

Ćwiczenie. Niech $g = \langle \cdot, \cdot \rangle$ będzie iloczynem skalarnym na przestrzeni zespolonej V. Dowieść, że gdy V traktować jako przestrzeń rzeczywistą, to funkcje $\mathrm{Re}(g), \mathrm{Im}(g) : V \times V \to \mathbb{R}$ są dwuliniowe, przy czym pierwsza jest symetryczna, a druga antysymetryczna.

Wniosek 1. *Funkcja dwuliniowa $g : V \times V \to \mathbb{F}$ wtedy i tylko wtedy jest symetryczna (odp.: antysymetryczna), gdy jej macierz w zadanej bazie przestrzeni V jest taka.*

Dowód. Z (14) wynika, że gdy macierz $\mathbf{A}$ jest symetryczna (odp. antysymetryczna), to funkcja g też. Przeciwna implikacja wynika z definicji macierzy funkcji g. $\square$

Uwaga 1. (i definicja). Macierze funkcji dwuliniowej $g : V \times V \to \mathbb{F}$, rozpatrywane względem różnych baz, są kongruentne (patrz stwierdzenie 1). Niezmienników kongruentności macierzy użyć więc można do zdefiniowania własności takiej funkcji g. W szczególności, **rzędem funkcji dwuliniowej** g nazywamy rząd jej macierzy w dowolnej bazie przestrzeni V. Funkcję tę nazwiemy **nieosobliwą** lub **niezdegenerowaną**, gdy $\mathrm{rk}(g) = \dim(V)$, tzn. gdy jej macierz w dowolnej bazie jest nieosobliwa. W przeciwnym nazwiemy ją **osobliwą** lub **zdegenerowaną**. Możemy też użyć niezmienników rzeczywistych macierzy symetrycznych, by zdefiniować **sygnaturę** czy **określoność** wzgl. **półokreśloność** (dodatnią czy ujemną) symetrycznej funkcji dwuliniowej g na rzeczywistej przestrzeni wektorowej. (Są one takie, jak macierzy tej funkcji w dowolnej bazie przestrzeni.)

Zadania.

2. Gdy funkcja $g : V \times V \to \mathbb{F}$ jest dwuliniowa, to dla każdych liniowo zależnych wektorów $\mathbf{v}_1, ..., \mathbf{v}_k \in V$ macierz $(g(\mathbf{v}_i, \mathbf{v}_j))_{1 \leq i,j \leq k}$ jest osobliwa.

3. Niech $g : V \times V \to \mathbb{F}$ będzie funkcją dwuliniową i niech $g^t(\mathbf{u}, \mathbf{w}) := g(\mathbf{w}, \mathbf{u})$ dla $\mathbf{u}, \mathbf{w} \in V$. Zbadać zależność pomiędzy macierzami funkcji g i g^t w danej bazie przestrzeni V i dowieść równości $\mathrm{rk}(g) = \mathrm{rk}(g^t)$.

4. Niech funkcja $g : V \times V \to \mathbb{F}$ będzie dwuliniowa.

a) Gdy funkcja g jest **alternująca** (tzn. $g(\mathbf{v}, \mathbf{v}) = 0$ dla każdego wektora $\mathbf{v}$), to jest antysymetryczna. (Wskazówka: $g(\mathbf{u}+\mathbf{w}, \mathbf{u}+\mathbf{w}) = 0$.)

b) Implikacja przeciwna jest prawdziwa gdy $2_{\mathbb{F}} \neq 0_{\mathbb{F}}$, i tylko wtedy.

Zadania uzupełniające.

1. Niech SYM (odp. ANT) oznacza zbiór wszystkich funkcji symetrycznych (odpowiednio: antysymetrycznych) $V \times V \to \mathbb{F}$.

a) Dowieść, że SYM i ANT są podprzestrzeniami liniowymi przestrzeni FUN wszystkich funkcji $V \times V \to \mathbb{F}$, oraz $FUN = SYM \oplus ANT$;

b) opisać wzorem rzut liniowy P przestrzeni FUN na SYM wzdłuż ANT i zbadać, czy P przeprowadza funkcje dwuliniowe w dwuliniowe.

2. Niech $g : V \times V \to \mathbb{F}$ będzie funkcją dwuliniową i niech $f(\mathbf{v}) = g(\mathbf{v}, \mathbf{v})$ dla $\mathbf{v} \in V$. Udowodnić następującą **tożsamość Cauchy'ego**: $f(\mathbf{u})\big(f(\mathbf{u})f(\mathbf{w}) - g(\mathbf{u}, \mathbf{w})g(\mathbf{w}, \mathbf{u})\big) = f\big(f(\mathbf{u})\mathbf{w} - g(\mathbf{u}, \mathbf{w})\mathbf{u}\big)$ dla $\mathbf{u}, \mathbf{w} \in V$

3. Niech V będzie dwuwymiarową rzeczywistą przestrzenią wektorową, niech $g : V \times V \to \mathbb{R}$ będzie symetryczną funkcją dwuliniową i niech wektory $\mathbf{u}, \mathbf{w} \in V$ będą liniowo niezależne. Dowieść, że:

i) funkcja g jest dodatnio lub ujemnie określona $\Leftrightarrow (g(\mathbf{u}, \mathbf{w}))^2 < g(\mathbf{u}, \mathbf{u})g(\mathbf{w}, \mathbf{w})$;

ii) funkcja g jest osobliwa$\Leftrightarrow (g(\mathbf{u}, \mathbf{w}))^2 = g(\mathbf{u}, \mathbf{u})g(\mathbf{w}, \mathbf{w})$.

iii) funkcja g jest nieokreślona i nieosobliwa $\Leftrightarrow (g(\mathbf{u}, \mathbf{w}))^2 > g(\mathbf{u}, \mathbf{u})g(\mathbf{w}, \mathbf{w})$.

Wywnioskować, że znak liczby $(g(\mathbf{u}, \mathbf{w}))^2 - g(\mathbf{u}, \mathbf{u})g(\mathbf{w}, \mathbf{w})$ (dodatni, zerowy lub ujemny) nie zależy od wyboru liniowo niezależnych wektorów $\mathbf{u}, \mathbf{w} \in V$.

4. Funkcja dwuliniowa g na przestrzeni wektorowej wtedy i tylko wtedy jest iloczynem dwóch funkcji liniowych, gdy $\mathrm{rk}(g) \leq 1$.

5. (przygotowawcze; powinno się znaleźć w rozdz. II): Niech operatory $K, L \in \mathcal{L}(V, W)$ mają tę własność, że dla każdego $\mathbf{v} \in V$ wektor $L(\mathbf{v})$ jest proporcjonalny do $K(\mathbf{v})$. Dowieść, że $L = \lambda K$ dla pewnego skalara λ.

6. Niech $g, h : V \times V \to \mathbb{F}$ będą funkcjami dwuliniowymi. Dowieść, że

a) Jeśli $h(\mathbf{u}, \mathbf{w}) = 0$ dla wszystkich $\mathbf{u}, \mathbf{w} \in V$ takich, że $g(\mathbf{u}, \mathbf{w}) = 0$, to $h = \lambda g$ dla pewnego skalara λ. (Wskazówka: poprzednie zadanie, zastosowane do operatorów $V \to V^*$ zadanych wzorami $\mathbf{v} \mapsto g(\mathbf{v}, \cdot)$ i $\mathbf{v} \mapsto h(\mathbf{v}, \cdot)$, wraz z zad. uz. 3a) w §III.4.3.)

b) Jeśli $g(\mathbf{u}, \mathbf{w}) = 0 \Rightarrow g(\mathbf{w}, \mathbf{u}) = 0$ $(\mathbf{u}, \mathbf{w} \in V)$, to g jest funkcją symetryczną lub antysymetryczną.

Zadania ze zbioru Kostrykina: §II.2.1: 1 do 5, 9,10,11,12,16 i 19a); §II.2.2: 1,2,3,6,19*,30.

2. Funkcje dwuliniowe a formy kwadratowe.

Twierdzenie 1. *a) Gdy funkcja $g : V \times V \to \mathbb{F}$ jest dwuliniowa, to poniższy wzór definiuje formę kwadratową:*

$$f(\mathbf{v}) = g(\mathbf{v}, \mathbf{v}) \ \text{dla} \ \ \mathbf{v} \in V. \tag{16}$$

Macierz $\mathbf{B}$ tej formy w zadanej bazie przestrzeni V jest równa $\frac{1}{2}(\mathbf{A} + \mathbf{A}^t)$, gdzie $\mathbf{A}$ to macierz funkcji g w tejże bazie. (W szczególności, $\mathbf{B} = \mathbf{A}$ gdy g jest symetryczna.)

b) Odwrotnie, gdy $f : V \to \mathbb{F}$ jest formą kwadratową, to istnieje dokładnie jedna symetryczna funkcja dwuliniowa $g : V \times V \to \mathbb{F}$ spełniająca równość (16). Ponadto,

$$g(\mathbf{u}, \mathbf{v}) = \frac{1}{2}(f(\mathbf{u} + \mathbf{v}) - f(\mathbf{u}) - f(\mathbf{v})) = \frac{1}{4}(f(\mathbf{u} + \mathbf{v}) - f(\mathbf{u} - \mathbf{v})) \ \ \text{dla} \ \mathbf{u}, \mathbf{v} \in V. \tag{17}$$

Dowód. a) Niech $\mathbf{A}$ będzie macierzą funkcji g w pewnej bazie $\mathcal{V}$. Na podstawie stwierdzenia 1a) z p.1, funkcja f jest w tej bazie zadana wielomianem $\mathbf{x}^t\mathbf{A}\mathbf{x}$. Jest to więc forma kwadratowa, której macierz w bazie $\mathcal{V}$ jest równa $\frac{1}{2}(\mathbf{A} + \mathbf{A}^t)$; patrz §1.2.

b) Niech w pewnej bazie przestrzeni V forma f zadana będzie wielomianem $\mathbf{x}^t\mathbf{B}\mathbf{x}$, gdzie $\mathbf{B} = \mathbf{B}^t$. Funkcję $g : V \times V \to \mathbb{F}$ zadajemy w tej bazie wielomianem $\mathbf{x}^t\mathbf{B}\mathbf{y}$. Równość (16) i dwuliniowość g są oczywiste, a symetria wynika z wniosku 1 w p.1. Z tych własności wynikają łatwo tożsamości (17), a z każdej z nich – jedyność g. ⊟

Uwaga 1. (i definicja) a) Gdy symetryczna funkcja dwuliniowa g i forma kwadratowa f pozostają w zależności (16), to o każdej z nich mówimy, że jest **wyznaczona** przez pozostałą. Inne stosowane określenie to: g jest **formą** (czy **funkcją**) **biegunową** formy kwadratowej f. Formułę, wyrażającą g przez f, nazywamy **polaryzacyjną**. Dwóch przykładów takich formuł dostarcza tożsamość (17).

b) Z twierdzenia wynika, że w dowolnej bazie przestrzeni V, macierz formy kwadratowej f jest równa macierzy jej funkcji biegunowej g.

Przykład 1. Jeśli więc forma f w bazie $\mathcal{V}$ jest zadana wielomianem $\sum_{1 \le i \le j \le k} a_{ij}x_ix_j$, to jej funkcja biegunowa g jest w tej bazie zadana wielomianem $\frac{1}{2}\sum_{1 \le i \le j \le k} a_{ij}(x_iy_j + x_jy_i)$. Dla przykładu, niech $V = \mathcal{M}_2(\mathbb{F})$. Funkcja $\det : V \to \mathbb{F}$ jest formą kwadratową, bo w bazie

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

zadana jest wielomianem $x_1x_4 - x_2x_3$. Jej funkcja biegunowa $g : V \times V \to \mathbb{F}$ jest więc w tej bazie zadana wielomianem $\frac{1}{2}x_1y_4 + \frac{1}{2}x_4y_1 - \frac{1}{2}x_2y_3 - \frac{1}{2}x_3y_2$. Wynika stąd, że

$$g(\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}, \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix}) = \frac{1}{2}\det\begin{pmatrix} x_1 & x_2 \\ y_3 & y_4 \end{pmatrix} + \frac{1}{2}\det\begin{pmatrix} y_1 & y_2 \\ x_3 & x_4 \end{pmatrix},$$

co można zgadnąć bezpośrednio: prawa strona jest symetryczną funkcją dwuliniową i dla $\mathbf{Y} = \mathbf{X}$ przyjmuje wartość $\det(\mathbf{X})$. (A co dałyby wzory polaryzacyjne (17)?)

Przykład 2. Niech $(V, \langle\cdot,\cdot\rangle)$ będzie przestrzenią unitarną nad $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ i $f(\mathbf{v}) = \|\mathbf{v}\|^2$ dla $\mathbf{v} \in V$. Gdy $\mathbb{F} = \mathbb{R}$, to f jest formą kwadratową, o funkcji biegunowej $g(\mathbf{u}, \mathbf{v}) = \langle\mathbf{u}, \mathbf{v}\rangle$. (Wynika to z definicji normy.) Macierz formy f w bazie $(\mathbf{v}_1, ..., \mathbf{v}_k)$ przestrzeni V jest równa macierzy Grama $(\langle\mathbf{v}_i, \mathbf{v}_j\rangle)_{i,j=1}^k$, bo ta jest macierzą funkcji g. Natomiast gdy $\mathbb{F} = \mathbb{C}$, to funkcja f nie jest kwadratowa (dlaczego?).

Dowód poniższego twierdzenia ilustruje możliwość wykorzystania funkcji biegunowej.

Twierdzenie 2. *Gdy f i f' są formami kwadratowymi na rzeczywistej przestrzeni wektorowej i forma f jest dodatnio określona, to istnieje baza przestrzeni, diagonalizująca każdą z form f, f'. (Macierz formy f w tej bazie jest nawet jednostkowa.)*

Dowód. Niech $g : V \times V \to \mathbb{F}$ oznacza funkcję biegunową formy f. Ze względu na założoną dodatnią określoność, para (V, g) jest przestrzenią euklidesową (tzn. wzór $\langle\mathbf{v}_1, \mathbf{v}_2\rangle := g(\mathbf{v}_1, \mathbf{v}_2)$ zadaje iloczyn skalarny na V w sensie definicji z rodziału VI). Wobec twierdzenia z §2.4 istnieje więc g–ortonormalna baza $(\mathbf{v}_i)_{i=1}^k$ przestrzeni V, diagonalizująca formę f'. Ortonormalność oznacza, że $g(\mathbf{v}_i, \mathbf{v}_j) = 0$ gdy $i \neq j$ oraz $g(\mathbf{v}_i, \mathbf{v}_i) = 1$ dla $i, j = 1, ..., k$. Macierz funkcji g w tej bazie jest więc jednostkowa, a tym samym i macierz formy f jest taka (patrz twierdzenie 1). $\square$

Uwaga 2. * Niech macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{R})$ będą symetryczne, przy czym $\mathbf{A}$ dodatnio określona. Z twierdzenia 2 wynika istnienie macierzy nieosobliwej $\mathbf{C}$ takiej, że $\mathbf{C}^t\mathbf{A}\mathbf{C} = \mathbf{I}$ i $\mathbf{C}^t\mathbf{B}\mathbf{C} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, dla pewnych $\lambda_1, ..., \lambda_k \in \mathbb{R}$. Liczby λ_i można jawnie wyznaczyć: są to zera wielomianu $\det(\mathbf{B} - x\mathbf{A})$ (z krotnościami), bo ten jest proporcjonalny do wielomianu $\det(\mathbf{C}^t\mathbf{B}\mathbf{C} - x\mathbf{I}) = \prod_i(\lambda_i - x)$.

Zadania uzupełniające.

1. (Wskazówka: twierdzenie 2.) Przy założeniach uwagi 2 dowieść, że
   a) $\det(\mathbf{A} + i\mathbf{B}) \neq 0$.
   b) Gdy $\mathbf{B}$ jest dodatnio półokreślona, to $\det(\mathbf{A}+\mathbf{B}) \geq \det(\mathbf{A})$ i nierówność jest ostra dla $\mathbf{B} \neq \mathbf{0}$.

2. Gdy $\mathbf{A}$ i $\mathbf{B}$ są jak w uwadze 2, znajdźmy dla każdego pierwiastka λ wielomianu $\det(\mathbf{B} - x\mathbf{A})$ układ fundamentalny rozwiązań równania $(\mathbf{B} - \lambda\mathbf{A})\mathbf{x} = \mathbf{0}$ i poddajmy go ortonormalizacji względem iloczynu skalarnego $\mathbf{u}^t\mathbf{A}\mathbf{v}$. Dowieść, że otrzymamy łącznie k wektorów, a macierz $\mathbf{C}$, mająca je jako kolumny, spełnia warunki uwagi 2.

3. Niech $\widetilde{V} = V \oplus iV$ oznacza kompleksyfikację rzeczywistej przestrzeni wektorowej V i niech $f : V \to \mathbb{R}$ będzie formą kwadratową. Dowieść, że:

a) Istnieje dokładnie jedna forma kwadratowa $\widetilde{f} : \widetilde{V} \to \mathbb{C}$ taka, że $\widetilde{f}_{|V} = f$. Wyrazić też $\widetilde{f}(\mathbf{u} + i\mathbf{v})$ przez $f(\mathbf{u}), f(\mathbf{v})$ i $f(\mathbf{u} + \mathbf{v})$.

b) $\mathrm{rk}(\widetilde{f}) = \mathrm{rk}(f)$.

Zadania ze zbioru Kostrykina: 8, 14, 15 w §II.2.2.

3. Formy dwuliniowe a geometria (informacje wstępne).

Definicja. Niech w przestrzeniach V i V' wyróżnione będą formy dwuliniowe g i g', odpowiednio. Przekształcenie $L \in \mathcal{L}(V, V')$ nazywamy **zanurzeniem izometrycznym**, gdy $\ker(L) = \{\mathbf{0}\}$ i $g'(L(\mathbf{v}_1), L(\mathbf{v}_2)) = g(\mathbf{v}_1, \mathbf{v}_2)$ dla wszystkich $\mathbf{v}_1, \mathbf{v}_2 \in V$. Zanurzenie izometryczne nazywamy **izometrią**, gdy jest „na". Jeśli izometria taka istnieje, przestrzenie (V, g) i (V', g') nazywamy **izometrycznymi**, a formy g i g' – **równoważnymi**.

Zadanie 1. Przy oznaczeniach definicji, niech $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ i $\mathcal{V}'$ będą bazami w V i V', odpowiednio. Dla monomorfizmu $L \in \mathcal{L}(V, V')$ dowieść równoważności warunków:

a) L jest zanurzeniem izometrycznym;

b) $g'(L(\mathbf{v}_i), L(\mathbf{v}_j)) = g(\mathbf{v}_i, \mathbf{v}_j)$ dla $i, j = 1, ..., k$;

c) $\mathbf{A} = \mathbf{C}^t\mathbf{B}\mathbf{C}$, gdzie $\mathbf{C} := [L]^{\mathcal{V}}_{\mathcal{V}'}$, zaś $\mathbf{A}$ i $\mathbf{B}$ to macierze form g i g' w bazie $\mathcal{V}$ i $\mathcal{V}'$, odpowiednio.

Uwaga 1. * Ważny jest przypadek, gdy $V = V$ i $g = g'$. Wartościowych informacji o równaniu $\mathbf{A} = \mathbf{C}^t\mathbf{A}\mathbf{C}$ dostarcza zad. uz. 5 w §II.5.2.

Zadanie 2. a) Wywnioskować, że gdy forma g jest nieosobliwa i $L : (V, g) \to (V, g)$ jest izometrią, to $\det(L) = \pm 1$.

b)* Korzystając z zadania uz. 2 w §VI.1.5, uzyskać też wyżej zależności między współczynnikami wielomianu χ_L i dowieść, że zbiór jego pierwiastków jest zamknięty względem brania odwrotności (z uwzględnieniem krotności).

Funkcji dwuliniowej $g : V \times V \to \mathbb{F}$ użyć można do wprowadzenia w przestrzeni V pojęcia ortogonalności w następujący sposób: powiemy, że wektory $\mathbf{u}, \mathbf{v} \in V$ są g–**ortogonalne** i piszemy $\mathbf{u} \perp_g \mathbf{v}$, jeśli $g(\mathbf{u}, \mathbf{v}) = 0$. Na ogół, tak zdefiniowana relacja jest niesymetryczna (kolejność wektorów $\mathbf{u}, \mathbf{v}$ jest istotna). Z części b) zadania uzupełniającego 6 w p.1 wynika

Twierdzenie 1. *Funkcja dwuliniowa, zadająca symetryczną relację ortogonalności, jest symetryczna lub jest antysymetryczna.* □

Funkcje dwuliniowe, które są antysymetryczne (równoważnie: **alternujące**, por. zadanie 4 w p.1) lub symetryczne, są więc geometrycznie wyróżnione; nazwiemy je **formami metrycznymi**.

Uwaga 2. Nazwa „forma metryczna" jest nieco myląca: forma taka na ogół nie wyznacza na przestrzeni metryki w sensie znanym z wykładów Analizy czy Topologii. Jednak spotykanej też nazwie „iloczyn skalarny" nadaliśmy już w rozdziale V inne (choć pokrewne) znaczenie. W literaturze, obie nazwy używane bywają w różnoraki sposób.

Gdy wybór formy metrycznej g na przestrzeni V nie budzi wątpliwości, to zamiast o g–ortogonalności wektorów $\mathbf{u}, \mathbf{v} \in V$ mówimy o ich ortogonalności, oznaczając ją $\mathbf{u} \perp \mathbf{v}$. W miejsce $g(\mathbf{u}, \mathbf{v})$ używane też bywa oznaczenie $\langle \mathbf{u}, \mathbf{v} \rangle$. W odróżnieniu od przypadku euklidesowego, istnieć mogą różne od $\mathbf{0}$ wektory g–ortogonalne do każdego innego, które nazwiemy **anihilującymi**, lub g–ortogonalne do siebie, które nazwiemy **izotropowymi**. Ogólniej, podprzestrzeń $U \neq \{\mathbf{0}\}$ nazwiemy **izotropową** (lub: **całkowicie osobliwą**), gdy $g_{|U \times U} = 0$. [6] Natomiast **przestrzeń anizotropowa** lub **określona** to taka, w której nie ma wektorów izotropowych.

Uwaga 3. Przestrzeń z wyróżnioną formą symetryczną jest całkowicie osobliwa wtedy i tylko wtedy, gdy każdy jej wektor jest izotropowy. (Wynika to z formuły polaryzacyjnej (16) lub zadania 4 w p.1.)

Na przestrzenie z wyróżnioną formą metryczną przenieść można pojęcia rzutu ortogonalnego, symetrii ortogonalnej, przekształcenia sprzężonego. Poniżej i w §4 naszkicujemy tę część zarysowującej się teorii, którą otrzymać można nieznacznie modyfikując rozumowania przedstawione w rozdziale V. Modyfikacje te wymagają pewnej ostrożności: intuicja może zawodzić, bo w sformułowaniach lub dowodach uwzględniać trzeba istnienie wektorów izotropowych i to, że zdefiniowany jest odpowiednik iloczynu skalarnego wektorów, lecz nie ich długości. Głębsze wyniki, w tym kluczowe twierdzenia Witta i Clifforda, znaleźć można w książkach Langa „Algebra" oraz Kostrykina i Manina „Algebra liniowa i geometria".

Twierdzenie 2. *Przestrzeń, w której ortogonalność zadana jest formą symetryczną, ma bazę ortogonalną (tzn., istnieje baza $(\mathbf{w}_i)_{i=1}^k$ taka, że $\mathbf{w}_i \perp \mathbf{w}_j$ dla $i \neq j$).*

Równoważne sformułowanie: W odpowiedniej bazie, macierz symetrycznej formy dwuliniowej jest diagonalna.

Dowód. Niech $\mathbf{A} := (g(\mathbf{v}_i, \mathbf{v}_j))_{i,j=1}^k$ będzie macierzą rozważanej formy g w dowolnie obranej bazie $\mathcal{V} = (\mathbf{v}_i)_{i=1}^k$. Ponieważ $\mathbf{A} = \mathbf{A}^t$, więc istnieje macierz nieosobliwa $\mathbf{C}$ taka, że $\mathbf{C}^t\mathbf{A}\mathbf{C}$ jest macierzą diagonalną. Baza $\mathcal{W}$, dla której $[I]_{\mathcal{V}}^{\mathcal{W}} = \mathbf{C}$, ma żądaną własność. (Wynika to z (15).) $\square$

Uwaga 4. Twierdzenie jest więc kolejną interpretacją twierdzenia Lagrange'a z §1.3, a kolumny powyższej macierzy $\mathbf{C}$ są ciągami współrzędnych, w bazie $\mathcal{V}$, wektorów szukanej bazy g–ortogonalnej.

[6] Jest to terminologia np. J. P. Serre'a. U nowszych autorów, „(pod)przestrzeń izotropowa" to taka, której pewien wektor jest izotropowy – co nie odpowiada pochodzeniu słowa „izotropowy" („izo"=jednakowy", „tropos"=kierunek).

Uwaga 5. W g–ortogonalnej bazie $(\mathbf{w}_i)_{i=1}^k$, forma g jest zadana wielomianem $\sum \lambda_i x_i y_i$, gdzie $\lambda_i := g(\mathbf{w}_i, \mathbf{w}_i)$. Liczba nieizotropowych wektorów $\mathbf{w}_i$ jest równa $\mathrm{rk}(g)$, zaś gdy V jest przestrzenią rzeczywistą, to $\sigma_+(g) = \#\{i : \lambda_i > 0\}$ i $\sigma_-(g) = \#\{i : \lambda_i < 0\}$.

Inny dowód twierdzenia 2 (interesujący, bo „geometryczny") wskaże zadanie uz. 3 w §4.1.

Zadanie 3. Symetryczne formy dwuliniowe na przestrzeniach rzeczywistych są równoważne wtedy i tylko wtedy, gdy mają tę samą sygnaturę.

Pomiędzy ortogonalnością zadaną formą symetryczną a zadaną formą alternującą zachodzi zasadnicza różnica: w przypadku alternującym forma kwadratowa $\mathbf{v} \mapsto g(\mathbf{v}, \mathbf{v})$ jest zerowa, podczas gdy w przypadku symetrycznym daje ona pełną informację o formie g, zgodnie z twierdzeniem z p.2. W przypadku symetrycznym często mówimy, że rozważana ortogonalność zadana jest przez formę kwadratową (zamiast, że jest zadana przez jej dwuliniową funkcję biegunową). Przenosimy też na formy kwadratowe powyższe pojęcia i mówimy np. o ortogonalności i izotropowości względem takich form, czy o izometryczności przestrzeni, na których wyróżniono formy kwadratowe.

Ćwiczenie. Czy przestrzenie $(\mathbb{R}^4, x_1^2 + x_2^2 - x_3^2 - x_4^2)$ i $(\mathcal{M}_2(\mathbb{R}), \det)$ są izometryczne?

Przykład 1. $\mathbb{R}^4$ z formą kwadratową $x^2 + y^2 + z^2 - t^2$ (równoważnie: z formą dwuliniową $xx' + yy' + zz' - tt'$) nazywana jest przestrzenią Minkowskiego. Ogólniej, o **przestrzeni Minkowskiego** mówimy w odniesieniu do pary (V, f), gdzie V jest rzeczywistą przestrzenią wektorową wymiaru 4, a f formą kwadratową o sygnaturze $(3, 1)$ lub $(1, 3)$. (Oba przypadki prowadzą do „takiej samej" geometrii.) Interesująca podprzestrzeń przestrzeni Minkowskiego to **płaszczyzna Minkowskiego**: $\mathbb{R}^2$ z formą $x^2 - t^2$ czy, gdy tak woleć, dwuwymiarowa przestrzeń rzeczywista z formą o sygnaturze $(1, 1)$.

Zadanie 4. Niech na przestrzeniach V i V' wyróżnione będą formy kwadratowe f i f', odpowiednio, z funkcjami biegunowymi g i g'. Dla monomorfizmu $L \in \mathcal{L}(V, V')$ dowieść równoważności warunków:
a) L jest zanurzeniem izometrycznym przestrzeni (V, g) w (V', g');
b) $f' \circ L = f$.

Zadania ze zbioru Kostrykina: zadania 20 w §II.2.1 oraz 2 i 7 w §II.2.2.

§ 4. Pojęcia geometryczne wyznaczone przez formę metryczną, c.d.

1. Dopełnienia ortogonalne i sumy ortogonalne (zadania uzupełniające).

Niech V będzie przestrzenią z wyróżnioną formą metryczną g.

Definicja. Ortogonalność $\perp_g$ oznaczamy dalej przez $\perp$ i dla $A, B \subset V$ piszemy:
$A^\perp = \{\mathbf{v} \in V : \mathbf{a} \perp \mathbf{v}$ dla każdego $\mathbf{a} \in A\}$;
$A \perp B$ gdy $B \subset A^\perp$ (mówimy wtedy, że **zbiory** A, B **są ortogonalne**);

$V = A \textcircled{\perp} B$, gdy A i B są ortogonalnymi podprzestrzeniami liniowymi i $V = A \oplus B$.

1. a) Udowodnić, że $A^\perp$ jest podprzestrzenią liniową i $(\mathrm{lin}(A))^\perp = A^\perp = \bigcap_{\mathbf{a} \in A} \mathbf{a}^\perp$.

b) Udowodnić, że jeśli $\mathbf{0} \in A \cap B$, to $(A+B)^\perp = (A \cup B)^\perp = A^\perp \cap B^\perp$.

c) Dowieść, że $\dim(A^\perp) \geq \dim(V) - p$ gdy zbiór A liczy p elementów. Wywnioskować, że $\dim(V_0^\perp) \geq \dim(V) - \dim(V_0)$ dla każdej podprzestrzeni liniowej V_0 przestrzeni V.

d)* Dowieść, że gdy forma g jest nieosobliwa, to $\dim(V_0^\perp) = \dim(V) - \dim(V_0)$ dla każdej podprzestrzeni liniowej $V_0 \subset V$.

2. a) Dowieść, że $\dim(V^\perp) = \dim(V) - \mathrm{rk}(g)$.

b) Wywnioskować, że forma g wtedy i tylko wtedy jest nieosobliwa, gdy $V^\perp = \{\mathbf{0}\}$, tzn. gdy $\mathbf{0}$ jest jedynym wektorem g–ortogonalnym do każdego innego.

3. Dowieść twierdzenia 2 z §3.3 przez indukcję względem $\dim(V)$, jak następuje. Gdy wyróżniona forma symetryczna jest zerowa, to nie ma czego dowodzić, a w przeciwnym razie istnieje nieizotropowy wektor $\mathbf{v}_1 \in V$; patrz uwaga 2 w §3.3. Przy $V' := \mathbf{v}_1^\perp$ zauważyć, że $\dim(V') \geq \dim(V) - 1$ i $\mathbf{v}_1 \notin V'$. Stąd $V = \mathbb{F}\mathbf{v}_1 \textcircled{\perp} V'$ i $\dim V' < \dim V$, co pozwala wykorzystać założenie indukcyjne.

4. Niech teraz forma g będzie alternująca.

a) Gdy $g \neq 0$, to $g(\mathbf{v}, \mathbf{w}) \neq 0$ dla pewych $\mathbf{v}, \mathbf{w} \in V$. Zastępując $\mathbf{w}$ przez pewną jego krotność uzyskać $g(\mathbf{v}, \mathbf{w}) = 1$. Dowieść, że przy $V_1 := \mathrm{lin}(\mathbf{v}, \mathbf{w})$ forma $g_{|V_1 \times V_1}$ jest nieosobliwa, a jej macierz $\mathbf{A}$ w bazie $\mathbf{v}, \mathbf{w}$ ma wiersze $(0, 1)$ i $(-1, 0)$; ponadto $V_1 \cap V_1^\perp = \{\mathbf{0}\}$.

b) Stąd i z zadania 1c) wywnioskować, że $V = V_1 \textcircled{\perp} V_1^\perp$, a następnie wykorzystać indukcję względem $\dim V$ do dowodu, że w pewnej bazie macierz formy g ma postać $\mathrm{diag}(\mathbf{A}, ..., \mathbf{A}, 0, ..., 0)$, gdzie $\mathbf{A}$ jest 2×2–klatką opisaną w a).

c) Wywnioskować dalej, że każda macierz antysymetryczna $\mathbf{X}$ jest kongruentna z macierzą postaci $\mathrm{diag}(\mathbf{A}, ..., \mathbf{A}, 0, ..., 0)$, gdzie $\mathbf{A}$ jest jak wyżej; jej rząd $\mathrm{rk}(\mathbf{X})$ jest liczbą parzystą, a wyznacznik jest kwadratem w ciele $\mathbb{F}$.

Uwaga 1. i) Bazę, o której mowa w b), nazywamy **kanoniczną** dla przestrzeni z formą alternującą g. W bazie tej, formie g odpowiada wielomian $\sum_{i=1}^{s}(x_i y_{i+1} - x_{i+1} y_i)$, przy czym $2s = \mathrm{rk}(g)$.

ii) Z części tezy c) i własności wielomianów kilku zmiennych można wywnioskować, że dla każdego $k \in \mathbb{N}$ istnieje wielomian Pf_k zmiennych a_{ij}, gdzie $1 \leq i < j \leq k$, mający współczynniki całkowite i taki, że $\det(\mathbf{A}) = (\mathrm{Pf}_k((a_{ij})_{1 \leq i < j \leq k}))^2$ dla każdej macierzy antysymetrycznej $\mathbf{A} = (a_{ij}) \in \mathcal{M}_k(\mathbb{C})$, i podobnie jest dla innych ciał. (Oczywiście $\mathrm{Pf}_k = 0$ dla $k \in 2\mathbb{N}$. Więcej wiadomości o *Pfaffianie* patrz np. [Ko-Ma], str. 190-191.)

Zadania ze zbioru Kostrykina: §II.2.2.5.

2. Podprzestrzenie nieosobliwe i rzut ortogonalny (zadania uzupełniające).

Przypomnijmy, że przestrzeń V_0 z wyróżnioną formą metryczną g_0 nazywamy **podprzestrzenią przestrzeni** (V, g), gdy V_0 jest podprzestrzenią liniową przestrzeni V, zaś g_0 jest obcięciem g, tzn. $g_0(\mathbf{u}, \mathbf{w}) = g(\mathbf{u}, \mathbf{w})$ dla $\mathbf{u}, \mathbf{w} \in V_0$. Na ogół pomijamy oznaczanie formy g_0 zakładając, że jest nią obcięcie formy g.

Dla takiej podprzestrzeni przyjmijmy $\mathrm{rk}(V_0) := \mathrm{rk}(g_0)$. Podprzestrzeń V_0 nazywamy **nieosobliwą** (lub: **niezdegenerowana**), jeśli $\mathrm{rk}(V_0) = \dim(V_0)$; w przeciwnym razie nazywamy ją **osobliwą** lub **zdegenerowaną**.[7]

Uwaga 1. Na podstawie zadania 2 w p.1, podprzestrzeń V_0 wtedy i tylko wtedy jest nieosobliwa, gdy $V_0 \cap V_0^{\perp} = \{\mathbf{0}\}$.

1. Niech $V = V_0 \bigoplus V_1$.
   a) Udowodnić, że $\mathrm{rk}(V) = \mathrm{rk}(V_0) + \mathrm{rk}(V_1)$.
   b) Wywnioskować, że z nieosobliwości V wynika nieosobliwość V_0 i V_1, i odwrotnie.

Definicja. Przekształcenie $P \in \mathcal{L}(V)$ nazywamy **rzutem ortogonalnym**, gdy $P^2 = P$ i $\ker(P) \perp \mathrm{im}(P)$.

2. Niech podprzestrzeń W przestrzeni V będzie nieosobliwa. Udowodnić, że
   a) $V = W \oplus W^{\perp}$, wobec czego rzutowanie ortogonalne na W istnieje i jest jedyne.
   b) Jeśli przestrzeń V jest nieosobliwa, to $W^{\perp}$ też i $(W^{\perp})^{\perp} = W$.

3. Dowieść, że gdy forma g jest symetryczna i $\mathbf{v}_1, ..., \mathbf{v}_p$ jest ortogonalnym układem wektorów nieizotropowych, to wzór $\mathbf{v} \mapsto \sum_{i=1}^{p} \frac{g(\mathbf{v},\mathbf{v}_i)}{g(\mathbf{v}_i,\mathbf{v}_i)} \mathbf{v}_i$ zadaje rzutowanie ortogonalne przestrzeni V na podprzestrzeń $\mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_p)$.

4. * Dla podprzestrzeni W przestrzeni V dowieść równoważności warunków:
   a) W jest maksymalną podprzestrzenią nieosobliwą (tzn. jest ona nieosobliwa, lecz każda podprzestrzeń $W' \supsetneq W$ jest osobliwa);
   b) $V = V^{\perp} \oplus W$, przy czym zamiast $\oplus$ można użyć $\bigoplus$;
   c) podprzestrzeń W jest nieosobliwa i $\dim W = \mathrm{rk}(g)$.

5. * a) Dowieść, że gdy W jest maksymalną podprzestrzenią nieosobliwą, to rzut P przestrzeni V na W wzdłuż $V^{\perp}$ zachowuje wyróżnioną formę g (tzn. $g(P(\mathbf{v}_1), P(\mathbf{v}_2)) = g(\mathbf{v}_1, \mathbf{v}_2)$ dla $\mathbf{v}_1, \mathbf{v}_2 \in V$).
   b) Wywnioskować, że gdy W i W' są maksymalnymi podprzestrzeniami nieosobliwymi przestrzeni V, to $W' = L(W)$ dla pewnej izometrii $L : V \to V$.

6. * Niech $V = U \bigoplus W$, przy czym podprzestrzeń U jest całkowicie osobliwa. Dowieść, że równość $U = V^{\perp}$ zachodzi wtedy i tylko wtedy, gdy podprzestrzeń W jest nieosobliwa.

Zadania ze zbioru Kostrykina: §II.2.1: 13,19b); §II.2.2: 4,22,23*,24,32*.

[7] W niektórych podręcznikach, n.p. u [Kostr], terminologia jest inna.

3. Sprzężenie przekształcenia między przestrzeniami z formą dwuliniową (zadania uz.).

Twierdzenie 1. *Niech (V, g) i (V', g') będą przestrzeniami z funkcją dwuliniową i niech $K \in \mathcal{L}(V, V')$. Gdy funkcja g jest nieosobliwa, to istnieje jedyne przekształcenie $K^h \in \mathcal{L}(V', V)$ takie, że*

$$g(\mathbf{v}, K^h(\mathbf{w})) = g'(K(\mathbf{v}), \mathbf{w}) \quad \textit{dla wszystkich } \mathbf{v} \in V,\ \mathbf{w} \in V'. \tag{18}$$

Mówimy, że K^h jest **sprzężeniem** przekształcenia K (działającego pomiędzy przestrzeniami z funkcją dwuliniową). Dwa dowody twierdzenia dają zadania 1 i 7.

1. a) Przy oznaczeniach twierdzenia obierzmy bazy $\mathcal{V}$ przestrzeni V i $\mathcal{V}'$ przestrzeni V', i niech $\mathbf{G}$ będzie macierzą formy g w bazie $\mathcal{V}$, a $\mathbf{G}'$ – macierzą formy g' w bazie $\mathcal{V}'$. Niech dalej $L \in \mathcal{L}(V', V)$ i oznaczmy $\mathbf{K} := [K]^{\mathcal{V}}_{\mathcal{V}'}$, $\mathbf{L} := [L]^{\mathcal{V}'}_{\mathcal{V}}$. Dowieść, że tożsamość

$$g(\mathbf{v}, L(\mathbf{w})) = g'(K(\mathbf{v}), \mathbf{w}) \text{ dla wszystkich } \mathbf{v} \in V,\ \mathbf{w} \in V'$$

zachodzi wtedy i tylko wtedy, gdy $\mathbf{K}^t\mathbf{G}' = \mathbf{G}\mathbf{L}$.

b) Udowodnić twierdzenie 1.

c) Dowieść, że gdy obie formy g i g' są nieosobliwe i symetryczne, to $(K^h)^h = K$.

2. Dowieść, że gdy V jest przestrzenią z nieosobliwą formą dwuliniową, to $\det(K^h) = \det(K)$ dla każdego operatora $K \in \mathcal{L}(V)$.

3. Przy oznaczeniach twierdzenia załóżmy, że formy g i g' są symetryczne. Wtedy:

a) Przekształcenie $K^h \circ K$ (z przestrzeni (V, g) w nią samą) jest samosprzężone.

b) K jest zanurzeniem izometrycznym $\Leftrightarrow$ $K^h \circ K = I_V$.

Używane niżej pojęcia *bazy dualnej* i *przestrzeni sprzężonej* omówione są w §III.3.4. Dla funkcji dwuliniowej $g : V \times V \to \mathbb{F}$ i wektora $\mathbf{v} \in V$ definiujemy funkcjonał $J_g(\mathbf{v}) \in V^*$ wzorem

$$(J_g(\mathbf{v}))(\mathbf{x}) = g(\mathbf{x}, \mathbf{v}) \text{ dla } \mathbf{x} \in V$$

4. a) Dowieść, że gdy funkcja g jest nieosobliwa, to $J_g : V \to V^*$ jest izomorfizmem liniowym.

b) Odwrotnie, każdemu izomorfizmowi $J : V \to V^*$ odpowiada nieosobliwa funkcja dwuliniowa $g : V \times V \to \mathbb{F}$, dla której $J_g = J$.

5. a) Dla izomorfizmu $J : V \to V^*$, przeprowadzającego daną bazę $\mathcal{B}$ przestrzeni V na dualną do niej bazę $\mathcal{B}^*$, jaka jest macierz powyższej funkcji g w bazie $\mathcal{B}$?

b)* Gdy $\mathbb{F} = \mathbb{R}$, czy dla każdego izomorfizmu $J : V \to V^*$ istnieje baza $\mathcal{B} = (\mathbf{b}_i)_{i=1}^k$ przestrzeni V taka, że $(J(\mathbf{b}_i))_{i=1}^k$ jest bazą dualną do $\mathcal{B}$? A gdy $\mathbb{F} = \mathbb{C}$?

6. Niech $L \in \mathcal{L}(V_1, V_2)$ i niech g_1 i g_2 będą nieosobliwymi formami dwuliniowymi na przestrzeniach V_1 i V_2, odpowiednio. Zdefiniujmy $L^* : V_2^* \to V_1^*$ wzorem $L^*(\varphi) = \varphi \circ L$

dla $\varphi \in V_2^*$. Dowieść, że $L^* \circ J_{g_2} = J_{g_1} \circ L^h$. (Ze względu na to, przekształcenia L^* i L^h są utożsamiane i oznaczane wspólnie przez L^*.)

7. Dowieść twierdzenia 1 następująco. Dla danego wektora $\mathbf{w} \in W$, wzór $\mathbf{v} \mapsto g'(K(\mathbf{v}), \mathbf{w})$ określa funkcjonał liniowy na przestrzeni V, który oznaczmy $\varphi_{\mathbf{w}}$. Zadanie 4a) umożliwia więc określenie wektora $(J_g)^{-1}(\varphi_{\mathbf{w}})$, który *przyjmiemy* za wartość $K^h(\mathbf{w})$. (Pozostaje uzasadnić wzór (18) i liniowość przekształcenia $\mathbf{w} \mapsto K^h(\mathbf{w})$.)

4. Wokół twierdzenia o bezwładności (zadania uzupełniające).

1. Niech g będzie symetryczną formą dwuliniową na rzeczywistej przestrzeni wektorowej V. Dowieść, że gdy V jest g–ortogonalną sumą prostą podprzestrzeni V_1 i V_2, to $\sigma(g) = \sigma(g_1) + \sigma(g_2)$, gdzie σ oznacza sygnaturę oraz $g_i := g|_{V_i \times V_i}$ dla $i = 1, 2$.

2. Niech V i g będą jak w poprzednim zadaniu. Dowieść, że:
a) Istnieją podprzestrzenie V_+, V_- i V_{nul} takie, że $V = V_{nul} \bigoplus V_+ \bigoplus V_-$ i forma g jest na V_+ określona dodatnio, na V_- ujemnie, a na V_{nul} jest zerowa.
b) Gdy V_+, V_- i V_{nul} są takimi podprzestrzeniami, to $\sigma(g) = (\dim(V_+), \dim(V_-))$.

Uwaga 1. Zadanie to daje jeszcze jedną (ważną) wersję twierdzenia o bezwładności.

3. * Przy poprzednich założeniach, niech W i W' będą nieosobliwymi podprzestrzeniami przestrzeni V. Dowieść, że każdą izometrię $W \to W'$ (gdy taka istnieje) można przedłużyć do izometrii $V \to V$.

Uwaga 2. Jest to szczególny przypadek twierdzenia Witta, które dotyczy dowolnego ciała skalarów.

Zadania ze zbioru Kostrykina: §II.2.2.25.

5. * Formy hermitowskie na przestrzeniach zespolonych (zadania uzupełniające).

Niech V będzie zespoloną przestrzenią wektorową. Funkcję $g : V \times V \to \mathbb{C}$ nazywamy **formą hermitowską** lub **półtoraliniową funkcją hermitowsko–symetryczną**, gdy
a) dla każdego $\mathbf{v} \in V$, funkcja $V \ni \mathbf{u} \mapsto g(\mathbf{u}, \mathbf{v}) \in \mathbb{C}$ jest liniowa, oraz
b) $g(\mathbf{u}, \mathbf{v}) = \overline{g(\mathbf{v}, \mathbf{u})}$ dla każdych $\mathbf{u}, \mathbf{v} \in V$.
Macierzą tej formy w bazie $(\mathbf{v}_i)_{i=1}^k$ nazywamy macierz samosprzężoną $(g(\mathbf{v}_i, \mathbf{v}_j))_{1 \le i,j \le k}$.

1. * Dowieść, że część rzeczywista (odp.: część urojona) takiej formy jest formą dwuliniową symetryczną (odp.: antysymetryczną), gdy V traktować jako przestrzeń nad $\mathbb{R}$.

2. * Dowieść, że gdy $\mathbf{A}$ (odp. $\mathbf{B}$) jest macierzą g w bazie $\mathcal{V}$ (odp.: w bazie $\mathcal{W}$), to $\mathbf{B} = \mathbf{C}^h \mathbf{A} \mathbf{C}$, gdzie $\mathbf{C} = [I]_{\mathcal{V}}^{\mathcal{W}}$.

Wiele definicji i wyników dotyczących form dwuliniowych symetrycznych przenieść można na formy hermitowskie (w tym na iloczyny skalarne na przestrzeniach nad $\mathbb{C}$).

3. * Wzorując się na materiale z §1.3 obmyśleć sposób pozwalający dla danej macierzy samosprzężonej $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ wskazać macierz nieosobliwą $\mathbf{C} \in \mathcal{M}_k(\mathbb{C})$ taką, że $\mathbf{C}^h\mathbf{A}\mathbf{C}$ jest macierzą diagonalną.

4. * a) Obmyśleć definicję sygnatury formy hermitowskiej (odp. macierzy samosprzężonej) i dowieść jej poprawności.

b) Dowieść, że przy tej definicji kryterium Jacobiego–Sylvestera pozostanie prawdziwe dla macierzy samosprzężonych. Tak samo jest z zadaniami uzupełniającymi 1–3 z §2.4, po oczywistych modyfikacjach.

Zadania ze zbioru Kostrykina: 28 i 29 w §II.2.1.

6. * Miscelania (zadania).

1. Rozpatrzmy płaszczyznę Minkowskiego z przykładu 1 w §4.3: $\mathbb{R}^2$ z formą kwadratową $q = x^2 - t^2$.

a) Dowieść, że operator $L : (\mathbb{R}^2, q) \to (\mathbb{R}^2, q)$ wtedy i tylko wtedy jest izometrią liniową, gdy jego macierz w bazie standardowej jest postaci $\begin{pmatrix} a & \varepsilon b \\ b & \varepsilon a \end{pmatrix}$ dla $\varepsilon = \pm 1$ oraz $a, b \in \mathbb{R}$ takich, że $a^2 - b^2 = 1$. (Traktujemy x jako pierwszą zmienną, a t jako drugą.) Ponadto, $\varepsilon = 1$ gdy L zachowuje orientację i $\varepsilon = -1$ w przeciwnym razie.

b) Naszkicować zbiory $\{\mathbf{v} : q(\mathbf{v}) = c\}$, dla $c = 0, \pm 1, \pm 2$, będące odpowiednikami pewnych sfer o środku w $\mathbf{0}$ w geometrii euklidesowej (dlaczego?). Dowieść, że są one niezmiennicze względem każdej izometrii liniowej płaszczyzny Minkowskiego na nią samą i naszkicować analogiczne zbiory dla przestrzeni $\mathbb{R}^3$ z formą $x^2 + y^2 - t^2$.

c) Rozpatrzmy bazę płaszczyzny Minkowskiego utworzoną przez wektory izotropowe $\mathbf{w}_1 = (1, 1)$ i $\mathbf{w}_2 = (1, -1)$; współrzędne wektora $\mathbf{v}$ w tej bazie oznaczmy przez (c, d). Dowieść, że $q(\mathbf{v}) = 4cd$ i każda zachowująca orientację izometria płaszczyzny $(\mathbb{R}^2, q)$ jest zadana macierzą postaci $\pm\mathrm{diag}(e^\alpha, e^{-\alpha})$, dla pewnego $\alpha \in \mathbb{R}$.

Operator L_α, którego macierz w bazie $\mathbf{w}_1, \mathbf{w}_2$ jest równa $\mathrm{diag}(e^\alpha, e^{-\alpha})$, nazwiemy **obrotem hiperbolicznym płaszczyzny Minkowskiego** o α jednostek; oczywiście $L_{\alpha+\beta} = L_\alpha L_\beta$ dla $\alpha, \beta \in \mathbb{R}$.

2. * (Kontynuacja poprzedniego.) a) Niech $C := \{\mathbf{u} \in \mathbb{R}^2 : q(\mathbf{u}) < 0 \text{ i } u_2 > 0\}$. Dowieść, że gdy $\mathbf{u}, \mathbf{v} \in C$, to istnieje dokładnie jedna liczba $\alpha \in \mathbb{R}^+$ taka, że $L_\alpha(\mathbb{R}^+\mathbf{u}) = \mathbb{R}^+\mathbf{v}$. Liczbę tę nazwijmy **miarą Minkowskiego kąta zorientowanego** między $\mathbf{u}$ i $\mathbf{v}$ i oznaczmy $\angle_m(\mathbf{u}, \mathbf{v})$. Dowieść, że gdy $\mathbf{u}, \mathbf{v}, \mathbf{w} \in C$, to $\angle_m(\mathbf{u}, \mathbf{v}) + \angle_m(\mathbf{v}, \mathbf{w}) = \angle_m(\mathbf{u}, \mathbf{w})$.

b) Znaleźć macierz $\mathbf{A}_\alpha$ operatora L_α w standardowej bazie. (Wskazówka: jej wyrazy okazują się równe **cosinusowi hiperbolicznemu** $\cosh(\alpha) := (e^{\alpha}+e^{-\alpha})/2$ lub **sinusowi hiperbolicznemu** $\sinh(\alpha) := (e^{\alpha}-e^{-\alpha})/2$.) Wyznaczyć wzory na $\cosh(\alpha+\beta)$ i $\sinh(\alpha+\beta)$, odpowiadające tożsamości $L_{\alpha+\beta} = L_\alpha L_\beta$.

c) Dowieść, że gdy $q(\mathbf{u}) = q(\mathbf{v}) = -1$, to miara Minkowskiego α kąta pomiędzy $\mathbf{u}$ i $\mathbf{v}$ jest wyznaczona równością $\cosh(\alpha) = g(\mathbf{u}, \mathbf{v})$, gdzie g jest funkcją biegunową formy q.

3. * Niech (V, g) i (V', g') będą przestrzeniami z wyróżnionymi formami metrycznymi. O przekształceniu liniowym $L : V \to V'$ powiemy, że **zachowuje ortogonalność**, gdy $\mathbf{v}_1 \perp_g \mathbf{v}_2 \Rightarrow L(\mathbf{v}_1) \perp_{g'} L(\mathbf{v}_2)$ dla $\mathbf{v}_1, \mathbf{v}_2 \in V$.

a) Dowieść, że jeśli $L : (V, g) \to (V', g')$ zachowuje ortogonalność, to istnieje skalar λ taki, że $g'(L(\mathbf{v}_1), L(\mathbf{v}_2)) = \lambda g(\mathbf{v}_1, \mathbf{v}_2)$ dla $\mathbf{v}_1, \mathbf{v}_2 \in V$.

b) Wskazać przykład automorfizmu płaszczyzny Minkowskiego, który zachowuje ortogonalność, lecz nie jest proporcjonalny do izometrii.

4. Niech $L : V \to V$ będzie izometrią przestrzeni z wyróżnioną formą metryczną g, a $\mathbf{v}_i$ wektorem własnym dla L, odpowiadającymi wartości własnej λ_i $(i = 1, 2)$. Udowodnić, że jeśli $\lambda_1 \lambda_2 \neq 1$, to $\mathbf{v}_1 \perp_g \mathbf{v}_2$.

5. * Dowieść, że gdy w 2-wymiarowej przestrzeni z nieosobliwą formą symetryczną istnieje wektor izotropowy, to istnieje też baza, złożona z takich wektorów.

Zadania ze zbioru Kostrykina: 2,7,22,23,24 w §II.2.2.

§ 5. Przewodnik

skopiowany z wywieszki dla potoku „zwyczajnego”.

Hasła, dotyczące omawianych tematów (nie zachowuję kolejności z wykładu):

1. Wielomiany stopnia ≤ 2 kilku zmiennych i wyznaczone przez nie funkcje $\mathbb{F}^k \to \mathbb{F}$. Jedyność wielomianu, wyznaczającego funkcję. (§1.1)

2. Funkcje wielomianowe stopnia ≤ 2 na przestrzeni wektorowej (§1.4). Formy kwadratowe jako wielomiany (§1.2) i jako funkcje na przestrzeni wektorowej (§1.4). Macierz formy kwadratowej w obu przypadkach. Zmiana macierzy formy przy zmianie bazy czy podstawieniu liniowym; kongruentność macierzy i jej podstawowe własności (§1.3 i §1.4).

3. Twierdzenie Lagrange'a w 3 wersjach. Rząd formy kwadratowej (§1.3 i §1.4).

4. Określoność rzeczywistych form kwadratowych i macierzy; kryterium Sylvestera – Jacobiego.

5 (§2.1). Twierdzenie Sylvestera o bezwładności i jego związek z badaniem funkcji kwadratowych (lemat 1 w §2.1 i zadania, omawiane na ćwiczeniach); klasyfikacja zespolonych i rzeczywistych form kwadratowych przy pomocy rzędu i sygnatury, odpowiednio (wniosek 1 i zadanie 1 w §2.1).

6 (§2.2). Ortogonalna diagonalizacja rzeczywistych form kwadratowych; wyrażenie sygnatury rzeczywistej macierzy symetrycznej przez jej wartości własne.

7. Funkcje dwuliniowe: ich macierz w bazie i zmiana tej macierzy przy zmianie bazy, przeniesienie na funkcje dwuliniowe pojęcia rzędu i sygnatury (gdy $\mathbb{F} = \mathbb{R}$). Odpowiedniość między symetrycznymi funkcjami dwuliniowymi a formami kwadratowymi (§3.2).

8. (§3.3 i §4 w zakresie objętym ćwiczeniami.) Pojęcia geometryczne, wyznaczone przez formę metryczną (ortogonalność, izometryczność, rzut ortogonalny, izotropowość, sprzężenie operatora).

§ 6. Iloczyny tensorowe i zewnętrzne.

1. Iloczyny tensorowe i zewnętrzne funkcji.

W tym punkcie oznaczamy przez X, Y, Z, X_i dowolne zbiory, a przez $\mathbb{F}$ –ustalone ciało.

Definicja. **Iloczyn tensorowy** $f \otimes g$ funkcji $f : X \to \mathbb{F}$ i $g : Y \to \mathbb{F}$, to funkcja z $X \times Y$ do $\mathbb{F}$, określona wzorem

$$(f \otimes g)(x, y) := f(x) \cdot g(y) \text{ dla } (x, y) \in X \times Y$$

Tak samo definiujemy iloczyn tensorowy większej liczby funkcji: gdy dana jest jeszcze funkcja $h : Z \to \mathbb{F}$, to $(f \otimes g \otimes h)(x, y, z) := f(x)g(y)h(z)$ dla $(x, y, z) \in X \times Y \times Z$, itd. Możmy więc zdefiniować $\otimes_{i=1}^n f_i : \prod_{i=1}^n X_i \to \mathbb{F}$ dla skończonego układu funkcji $f_i : X_i \to \mathbb{F}$. Przyjmujemy konwencję, zgodnie z którą iloczyny kartezjańskie $(X\times Y)\times Z$ i $X\times(Y\times Z)$ są w naturalny sposób identyfikowane z $X\times Y\times Z$, i podobnie dla iloczynów większej liczby przestrzeni. (W szczególności $X^k\times X^l$ utożsamiamy z X^{k+l}.) Mamy więc

$$(f \otimes g) \otimes h = f \otimes g \otimes h = f \otimes (g \otimes h), \quad \text{i ogólniej } (\otimes_{i=1}^k f_i) \otimes (\otimes_{i=k+1}^{k+l} f_i) = \otimes_{i=1}^{k+l} f_i \quad (19)$$

Uwaga 1. Nawet gdy $X = Y$, funkcje $f\otimes g$ i $g\otimes f$ są na ogół różne, a iloczyn $f\otimes g$ jest inny niż iloczyn „zwykły" $f \cdot g$. Pierwszy jest bowiem funkcją z $X \times X$ do $\mathbb{F}$, określoną wzorem $(x_1, x_2) \mapsto f(x_1)g(x_2)$, a drugi funkcją $x \mapsto f(x)g(x)$, z X do $\mathbb{F}$.

Iloczyn tensorowy funkcji antysymetrycznych (a będą nas one interesować) nie musi być funkcją antysymetryczną. By taką otrzymać, należy go poddać antysymetryzacji.

Definicja. Niech $f : X^k \to \mathbb{F}$ będzie dowolną funkcją, przy czym $k \geq 2$ i o ciele $\mathbb{F}$ zakładamy, że $n_{\mathbb{F}} \neq 0_{\mathbb{F}}$ dla $n = 2, 3..., k$. Pozwala to mówić o odwrotnościach wymienionych elementów $n_{\mathbb{F}}$ i ich iloczynu, który oznaczamy $k!$. **Antysymetryzację** $Af : X^k \to \mathbb{F}$ funkcji f definiujemy wzorem:

$$(Af)(x_1, ..., x_k) = \frac{1}{k!} \sum_{\pi\in\mathbf{S}_k} \mathrm{Sgn}(\pi) f(x_{\pi(1)}, ..., x_{\pi(k)}) \text{ dla } x_1, ..., x_k \in X \quad (20)$$

Twierdzenie 1. *Operator antysymetryzacji $f \mapsto Af$ jest rzutem liniowym przestrzeni wszystkich funkcji $f : X^k \to \mathbb{F}$, na podprzestrzeń wszystkich funkcji antysymetrycznych.*

Dowód. Liniowość A, tzn. równość $A(f_1 + cf_2) = Af_1 + c \cdot Af_2$ dla $f_1, f_2 : X^k \to \mathbb{F}$ i $c \in \mathbb{F}$, jest oczywista. W §IV.2.2 udowodniono, że każda funkcja Af jest antysymetryczna. Ponadto gdy już funkcja f jest taka, to w (20) każdy składnik sumy jest równy $f(x_1, ..., x_k)$, skąd $Af = f$. □

Definicja. **Iloczynem zewnętrznym** uporządkowanego układu funkcji $f_i : X^{k_i} \to \mathbb{F}$ $(i = 1, ..., n; k_i \geq 1)$ nazywamy funkcję $\wedge_{i=1}^n f_i : \prod_i X^{k_i} \to \mathbb{F}$, określoną wzorem

$$\wedge_{i=1}^n f_i := C(k_1, ..., k_n) \cdot A(\otimes_{i=1}^n f_i) \quad (21)$$

gdzie $C(k_1,...,k_n) := (k_1+...+k_n)!/\prod_i k_i!$. Trzeba tu powiedzieć, że równie usprawiedliwione jest wzięcie $C(k_1,...,k_n)=1$, co ma miejsce np. w [Ko] i [Ko-Ma]; jednak poprzedni wybór, przyjęty w większości podręczników, czyni prostszym związek iloczynu zewnętrznego z wyznacznikiem we wzorze (22) i w jego konsekwencjach. Gdzie można, nie będziemy przywoływać wartości stałych $C(k_1,...,k_n)$, starając się uwypuklić zależności, które powinny one spełniać.

Przykład 1. Dla funkcji $f_1,...,f_n : X \to \mathbb{F}$ i dla $x_1,...,x_n \in X$ zachodzi równość

$$(\bigwedge_{i=1}^{n} f_i)(x_1,...,x_n) = \det(f_i(x_j))_{i,j=1}^{n}. \tag{22}$$

Istotnie, skoro funkcja $f := \otimes_{i=1}^{n} f_i$ zdefiniowana jest wzorem $f(x_1,...,x_n)=\prod_{i=1}^{n} f_i(x_i)$, to (z definicji) $n!Af(x_1,...,x_n) = \sum_\pi \mathrm{Sgn}(\pi)\prod_{i=1}^{n} f_i(x_{\pi(i)})$. Prawa strona jest pełnym rozwinięciem wyznacznika macierzy $(f_i(x_j))_{i,j=1}^{n}$, zaś $n! = C(1,...,1)$ (jedynek jest n). To daje żądaną zależność.

Wniosek 1. *Funkcje $f_1,...,f_n : X \to \mathbb{F}$ są liniowo niezależne wtedy i tylko wtedy, gdy funkcja $\bigwedge_i f_i : X^n \to \mathbb{F}$ jest niezerowa.*

Dowód. Z (22) wynika, że jeśli układ jest liniowo zależny, to $\bigwedge_i f_i = 0$ (bo wiersze macierzy $(f_i(x_j))_{i,j=1}^{n}$ są liniowo zależne, dla każdych $x_1,...,x_n \in X$). Przeciwnie, gdy układ jest liniowo niezależny, to – jak dowiedziono w semestrze I – wiersze tej macierzy sa liniowo niezależne dla pewnych $x_1,...,x_n \in X$, wobec czego $(\bigwedge_i f_i)(x_1,...,x_n) \neq 0$. □

Wniosek 2. *Niech $f_1,...,f_n : X \to \mathbb{F}$ i niech $g_i := \sum_{j=1}^{n} c_{ij} f_j$ dla $i=1,...,n$. (Tu, $(c_{ij})_{i,j=1}^{n}$ to macierz skalarów.) Wówczas $g_1 \wedge ... \wedge g_n = \det(c_{ij}) f_1 \wedge ... \wedge f_n$.*

Dowód. Funkcje $\bigwedge_{i=1}^{n} f_i$ i $\bigwedge_{i=1}^{n} g_i$ przyjmują na dowolnym ciągu $(x_s)_{s=1}^{n}$ wartość $\det(f_i(x_s))$ i $\det(g_i(x_s))$, odpowiednio. Teza wynika więc z twierdzenia Cauchy'ego o wyznaczniku iloczynu macierzy, bo $(g_i(x_s)) = (c_{ij}) \cdot (f_j(x_s))$. □

Ćwiczenie. Traktujmy wektory $\mathbf{x} \in \mathbb{R}^d$ jako funkcje z $S := \{1,...,d\}$ do $\mathbb{R}$. Dla $d=3$ opisać wzorami iloczyny $\mathbf{x}\otimes\mathbf{y}$ i $\mathbf{x}\wedge\mathbf{y}$. (Są to funkcje z $S\times S$ do $\mathbb{R}$, czyli 3×3–macierze.)

Stwierdzenie 1. *Mnożenie tensorowe funkcji jest rozdzielne względem dodawania, oraz przemienne z mnożeniem przez skalar: gdy $f,f' : X \to \mathbb{F}$, $g,g' : Y \to \mathbb{F}$ i $\lambda \in \mathbb{F}$, to*

$$f\otimes(g+g') = f\otimes g + f\otimes g',\ \ (f+f')\otimes g = f\otimes g + f'\otimes g \ \ i \ \ (\lambda f)\otimes g = \lambda(f\otimes g) = f\otimes(\lambda g).$$

Tak samo jest z mnożeniem zewnętrznym, dla $f,f' : X^k \to \mathbb{F}$, $g,g' : X^l \to \mathbb{F}$.

Dowód. Dla mnożenia tensorowego wynika to z definicji, a dla zewnętrznego – też, gdy uwzględnić liniowość antysymetryzacji.

Twierdzenie 2. *Dla funkcji skalarnych f, g, h, określonych na X^k, X^l i X^m, odpowiednio, mają miejsce równości:*
a) $f \wedge g = (-1)^{kl} g \wedge f$ (skośna przemienność z gradacją mnożenia zewnętrznego);
b) gdy wyżej k jest liczbą nieparzystą, to $f \wedge f = 0$.
c) $(f \wedge g) \wedge h = f \wedge (g \wedge h) = f \wedge g \wedge h$ (łączność mnożenia zewnętrznego).

Do dowodu potrzebny bedzie lemat, a także użycie oznaczeń z §IV.2.2.

Oznaczenia. Dla ustalonego zbioru X i permutacji $\pi \in \mathbf{S}_k$, piszemy ε_π w miejsce $\mathrm{Sgn}(\pi)$ i zadajemy bijekcję $\widetilde{\pi} : X^k \to X^k$ wzorem $\widetilde{\pi}(x_1, ..., x_k) = (x_{\pi(1)}, ..., x_{\pi(k)})$. Nieco ogólniej, gdy $J \subset \{1, ..., k\}$ i $\pi \in \mathbf{S}_J$, to przez $\widetilde{\pi} : X^k \to X^k$ oznaczamy bijekcję, przyporządkowującą każdemu układowi $(x_i)_{i=1}^k$ układ, którego i-ta współrzędna jest równa $x_{\pi(i)}$ gdy $i \in J$, zaś jest równa x_i w przeciwnym razie.
Odnotujmy, że $\widetilde{\pi\sigma} = \widetilde{\pi} \circ \widetilde{\sigma}$ dla $\pi, \sigma \in \mathbf{S}_J$.

Definicja. a) Antysymetryzacją funkcji $f : X^k \to \mathbb{F}$ **ze względu na wskaźniki ze zbioru** $J \subset \{1, ..., k\}$, nazywamy funkcję $A_J f := \frac{1}{j!} \sum_{\pi \in \mathbf{S}_J} \varepsilon_\pi f \circ \widetilde{\pi}$, gdzie j to liczebność zbioru J.
b) Funkcję $f : X^k \to \mathbb{F}$ nazywamy **antysymetryczną we wskaźnikach ze zbioru** J, jeśli $f \circ \widetilde{\pi} = \varepsilon_\pi f$ dla wszystkich $\pi \in \mathbf{S}_J$.

Lemat 1. *Dla funkcji $f : X^k \to \mathbb{F}$ i permutacji $\pi \in \mathbf{S}_k$ mają miejsce równości $A(f \circ \widetilde{\pi}) = \varepsilon_\pi A f$ oraz $A(A_J f) = Af$, gdzie A to antysymetryzacja pełna (we wszystkich wskaźnikach).*

Dowód. Pierwsza równość wynika stąd, że

$$k!A(f\circ\widetilde{\pi}) = \sum_{\sigma\in\mathbf{S}_k} \varepsilon_\sigma (f\circ\widetilde{\pi})\circ\widetilde{\sigma} = \sum_{\sigma\in\mathbf{S}_k} \varepsilon_\sigma f\circ\widetilde{\pi\sigma} = \sum_{\sigma\in\mathbf{S}_k} \varepsilon_\pi\varepsilon_{\pi\sigma} f\circ\widetilde{\pi\sigma} = \sum_{\tau\in\mathbf{S}_k} \varepsilon_\pi\varepsilon_\tau f\circ\widetilde{\tau} = \varepsilon_\pi k! Af.$$

Dla dowodu drugiej odnotujmy, że $j!A(A_J f) = A(\sum_{\pi\in S_J} \varepsilon_\pi f \circ \widetilde{\pi}) = \sum_{\pi\in S_J} \varepsilon_\pi A(f \circ \widetilde{\pi})$, co wobec pierwszej równości tezy daje $j!A(A_J f) = \sum_{\pi\in S_J} \varepsilon_\pi^2 Af = j!Af$. □

Dowód twierdzenia 2. a) $(f \otimes g)(x_1, ..., x_{k+l}) = f(x_1, ..., x_k) g(x_{k+1}, ..., x_{k+l})$ i $(g \otimes f)(x_1, ..., x_{k+l}) = g(x_1, ..., x_l) f(x_{l+1}, ..., x_{l+k})$, skąd $g \otimes f = (f \otimes g) \circ \widetilde{\pi}$ dla permutacji π, przeprowadzającej liczby $1, ..., k$ odpowiednio na $l+1, ..., l+k$, a $k+1, ..., k+l$ na $1, ..., l$. Z pierwszej tożsamości z lematu otrzymujemy więc, przy $c := C(k, l)$:

$$g \wedge f = cA(g \otimes f) = cA((f \otimes g) \circ \widetilde{\pi}) = c\varepsilon_\pi A(f \otimes g) = \varepsilon_\pi f \wedge g.$$

Teza a) wynika teraz z równości $\varepsilon_\pi = (-1)^{jk}$, pozostawionej jako zadanie.
b) Jak wiemy z a), $f \wedge f = -f \wedge f$ gdy k jest liczbą nieparzystą.
c) Niech $J = \{1, ..., k+l\} \subset \{1, ..., k+l+m\}$. Na podstawie definicji,

$$(f \wedge g) \wedge h = cA_J(f \otimes g) \wedge h = c'A(cA_J(f \otimes g) \otimes h) = c'cAA_J(f \otimes g \otimes h),$$

gdzie $c' := C(k+l,m)$ i nadal $c = C(k,l)$. Ponieważ $c'c = C(k,l,m)$ i $AA_J = A$ (co wiemy z lematu), więc $(f \wedge g) \wedge h = f \wedge g \wedge h$. Tak samo, $f \wedge (g \wedge h) = f \wedge g \wedge h$. $\square$

Uwaga 2. a) Ze względu na potrzeby rachunku tensorowego, warto powyższe oznaczenia i wyniki rozszerzyć na przypadek, gdy zamiast X^k mamy przestrzeń $\prod_{i=1}^k X_i$ nieco bardziej ogólną. Niech np. $X_i = Y$ dla $i = 1, ..., s$ oraz $X_i = Z$ dla $i = s+1, ..., k$, i niech $J \subset \{1, ..., s\}$ lub $J \subset \{s+1, ..., k\}$. Wówczas definicja bijekcji $\widetilde{\pi}$ nadal ma sens dla $\pi \in \mathbf{S}_J$, co pozwala jak poprzednio określić, kiedy funkcja $f : \prod_i X_i \to \mathbb{F}$ jest antysymetryczna we wskaźnikach ze zbioru J, a także zdefiniować jej antysymetryzację $A_J f$ ze względu na te wskaźniki.

b) Twierdzenie 1, wraz z dowodem, pozostaje w tej sytuacji słuszne: A_J nadal jest rzutem liniowym przestrzeni wszystkich funkcji $\prod_i X_i \to \mathbb{F}$ na podprzestrzeń, złożoną ze wszystkich funkcji, antysymetrycznych we wskaźnikach ze zbioru J.

Udowodnijmy na koniec dogodną własność iloczynu tensorowego funkcji:

Stwierdzenie 2. *Jeśli funkcje $f_1, ..., f_k : X \to \mathbb{F}$ są liniowo niezależne, a funkcje $g_1, ..., g_k : Y \to \mathbb{F}$ są niezerowe, to funkcje $f_1 \otimes g_1, ..., f_k \otimes g_k$ są liniowo niezależne. Tak samo jest, jeśli funkcje $g_1, ..., g_k$ są liniowo niezależne, a funkcje $f_1, ..., f_k$ –niezerowe.*

Dowód. Niech skalary $\lambda_1, ..., \lambda_k$ spelniają warunek $\sum_i \lambda_i f_i \otimes g_i = 0$. Obierzmy $y \in Y$ tak, by $g_1(y) \neq 0$. Skoro $\sum_i \lambda_i g_i(y) f_i(x) = 0 \ \forall x \in X$, to z liniowej niezależności funkcji $f_1, ..., f_k$ wynika, że $\lambda_1 = 0$. Podobnie, $\lambda_i = 0$ dla pozostałych i. $\square$

Zadanie 1. Dla funkcji $f_1, ..., f_n : X \to \mathbb{F}$ i permutacji $\pi \in \mathbf{S}_n$ zachodzi równość $(f_1 \otimes ... \otimes f_n) \circ \widetilde{\pi} = f_{\pi^{-1}(1)} \otimes ... \otimes f_{\pi^{-1}(n)}$.

Zadanie 2. Dla antysymetrycznych funkcji $f : X^k \to \mathbb{F}, g : X^l \to \mathbb{F}$ i $\mathbf{x} \in \mathbb{F}^{k+l}$ zachodzi równość $(f \wedge g)(\mathbf{x}) = \sum \mathrm{Sgn}(\pi) f(x_{\pi(1)}, ..., x_{\pi(k)}) g(x_{\pi(k+1)}, ..., x_{\pi(k+l)})$, gdzie sumowanie jest po tych permutacjach $\pi \in \mathbf{S}_{k+l}$, które są rosnące na $\{1, ..., k\}$ i na $\{k+1, ..., k+l\}$. Analogicznie, gdy funkcji antysymetrycznych jest więcej. (Pozwala to określić iloczyn zewnętrzny funkcji antysymetrycznych dla każdej charakterystyki ciała $\mathbb{F}$.)

Zadanie 3. Dla funkcji $f : X^k \to \mathbb{F}$ zdefiniujmy jej **symetryzację** $Sf : X^k \to \mathbb{F}$ wzorem $Sf := \frac{1}{k!} \sum_{\pi \in \mathbf{S}_k} f \circ \widetilde{\pi}$. Dowieść, że operator symetryzacji $f \mapsto Sf$ jest rzutem liniowym na przestrzeń tych funkcji $g : X^k \to \mathbb{F}$, które są symetryczne, tzn. spełniają warunek $g \circ \widetilde{\sigma} = g$ dla $\sigma \in \mathbf{S}_k$.

b) Dowieść równości $SA = AS = 0$.

2. Iloczyny tensorowy i zewnętrzny przestrzeni funkcyjnych

Definicja. Niech F, G i H będą podprzestrzeniami liniowymi przestrzeni funkcji skalarnych, określonych na X, Y i Z, odpowiednio. **Iloczyn tensorowy** $F \otimes G$ **przestrzeni**

funkcyjnych definiujemy wówczas wzorem $F \otimes G := \mathrm{lin}\{f \otimes g : f \in F, g \in G\}$; jest to podprzestrzeń liniowa przestrzeni wszystkich funkcji z $X \times Y$ do $\mathbb{F}$. Równoważnie (patrz wniosek 1 w p.1)

$$F \otimes G = \{f_1 \otimes g_1 + ... + f_s \otimes g_s : s \in \mathbb{N}, f_1, ..., f_s \in F, g_1, ..., g_s \in G\} \tag{23}$$

Tak samo definiujemy iloczyn tensorowy większej liczby przestrzeni funkcyjnych.

Z (19) wynika ówność:

$$(F \otimes G) \otimes H = F \otimes G \otimes H = F \otimes (G \otimes H) \tag{24}$$

Gdy zaś $F_1, F_2 \subset F$ i $G_1, G_2 \subset G$ są podprzestrzeniami liniowymi, to, wobec wniosku 1 w p.1,

$$(F_1 + F_2) \otimes G = F_1 \otimes G + F_2 \otimes G \ \text{ i } \ F \otimes (G_1 + G_2) = F \otimes G_1 + F \otimes G_2 \tag{25}$$

Uwaga 1. Istnieje „wzorcowa” bijekcja zbioru $X \times Y$ na $Y \times X$, zadana wzorem $(x, y) \mapsto (y, x)$. Wyznacza ona izomorfizm liniowy $F \otimes G \to G \otimes F$, przeprowadzający każdą sumę $\sum_{i=1}^{s} f_i \otimes g_i$ na $\sum_{i=1}^{s} g_i \otimes f_i$.

Analogiczne definicje wprowadzimy dla iloczynu zewnętrznego, przy założeniu, że F, G i H są przestrzeniami liniowymi (pewnych) funkcji skalarnych, określonych na skończonych iloczynach kartezjańskich tego samego zbioru. Przyjmujemy

$$F \wedge G = \{f_1 \wedge g_1 + ... + f_s \wedge g_s : s \in \mathbb{N}, f_1, ..., f_s \in F, g_1, ..., g_s \in G\} \tag{26}$$

Ze względu na liniowość antysymetryzacji i definicję iloczynu zewnętrznego,

$$F \wedge G = A(F \otimes G), \ \text{ gdzie } A \text{ jest antysymetryzacją funkcji } k + l \text{ zmiennych.} \tag{27}$$

Uwaga 2. Pozostają słuszne odpowiedniki równości (24) i (25), tzn. $(F \wedge G) \wedge H = F \wedge (G \wedge H)$, a jeśli $F_1, F_2 \subset F$ i $G_1, G_2 \subset G$ są podprzestrzeniami liniowymi, to $(F_1 + F_2) \wedge G = F_1 \wedge G + F_2 \wedge G$ i $F \wedge (G_1 + G_2) = F \wedge G_1 + F \wedge G_2$.

Dzięki łączności mnożenia, wyrażonej w (24) i powyżej, możemy w iloczynie tensorowym wielu czynników pomijać nawiasy, i tak samo w iloczynie zewnętrznym.

Definicja. Niech F będzie przestrzenią liniową (pewnych) funkcji, określonych na zbiorze X. Jej potęgi tensorowe i zewnętrzne definiujemy indukcyjnie, przyjmując $\otimes^1 F := \bigwedge^1 F := F$ i

$$\otimes^n F := (\otimes^{n-1} F) \otimes F \ \text{ i } \ \bigwedge^n F := (\bigwedge^{n-1} F) \wedge F \text{ dla } n > 1. \tag{28}$$

Przyjmujemy też, że $\otimes^0 F$ i $\bigwedge^0 F$ oznacza przestrzeń wszystkich funkcji skalarnych na przestrzeni jednopunktowej, w oczywisty sposób izomorficzną z $\mathbb{F}$. Ze względu na (27),

$$\bigwedge{}^n F = A(\otimes^n F) \quad \text{dla } n > 1. \tag{29}$$

Powyżej, A traktujemy jako endomorfizm przestrzeni wszystkich funkcji skalarnych, określonych na X^n.

Mimo podobieństw i związków mnożenia tensorowego i zewnętrznego, są i istotne różnice między nimi. Poniższa uwaga nie ma odpowiednika dla mnożenia zewnętrznego.

Uwaga 3. Jeśli w (25) po którejś stronie równości suma jest prosta, to i po drugiej jest ona taka. Istotnie, jeśli np. $F_1 \cap F_2 = \{\mathbf{0}\}$, to i $F_1 \otimes G \cap F_2 \otimes G = \{\mathbf{0}\}$, na podstawie stwierdzenia 2 z p.1.

Wyznaczymy teraz bazy i wymiar interesujących nas iloczynów i potęg.

Twierdzenie 1. *Gdy $\{f_i\}_{i=1}^k$ jest bazą przestrzeni F, a $\{g_j\}_{j=1}^l$– bazą przestrzeni G, to $\{f_i \otimes g_j : i = 1, ..., k, j = 1, ..., l\}$ jest bazą przestrzeni $F \otimes G$; w szczególności,* $\dim(F \otimes G) = (\dim F) \cdot (\dim G)$. *Podobnie jest dla iloczynu tensorowego większej liczby przestrzeni funkcyjnych.*

Dowód. Ponieważ $F = \oplus_{i=1}^k \mathbb{F} f_i$ i $G = \oplus_{j=1}^l \mathbb{F} g_j$, to z uwagi 3 wynika równość $F \otimes G = \oplus_{i=1}^k \oplus_{j=1}^l Z_{ij}$, gdzie Z_{ij} oznacza podprzestrzeń $(\mathbb{F} f_i) \otimes (\mathbb{F} g_j)$, w oczywisty sposób równą $\mathbb{F} \cdot (f_i \otimes g_j)$. A skoro $F \otimes G = \oplus_{i,j} \mathbb{F} \cdot (f_i \otimes g_j)$, to układ $(f_i \otimes g_j)_{i,j}$ jest bazą. □

Twierdzenie 2. *Jeśli $f_1, ..., f_d$ jest bazą przestrzeni F, to bazę przestrzeni $\bigwedge^n F$ tworzy układ funkcji $\{f_{i_1} \wedge ... \wedge f_{i_n} : \quad 1 \le i_1 < ... < i_n \le d\}$.*

Dowód. Z twierdzenia 1 wynika, że przestrzeń $\bigwedge^n F = A(\otimes^n F)$ jest rozpinana przez iloczyny $f_{i_1} \wedge ... \wedge f_{i_n}$, gdzie $i_j \in \{1, ..., d\}$ dla $j = 1, ..., n$. Na podstawie twierdzenia 2 z p.1, zmiana kolejności wskaźników $i_1, ..., i_n$ wpływa tylko na znak iloczynu zewnętrznego; możemy więc ograniczyć się do niemalejących układów $1 \le i_1 \le ... \le i_n \le d$. Tak samo, jeśli $i_j = i_{j+1}$ dla pewnego j, to $f_{i_1} \wedge ... \wedge f_{i_n} = 0$, bo $f_{i_j} \wedge f_{i_{j+1}} = 0$. Przestrzeń $\bigwedge^n F$ jest więc rozpinana przez żądany układ $\{f_{\mathbf{i}} : \mathbf{i} \in \mathcal{I}\}$, gdzie $\mathcal{I}$ oznacza zbiór wszystkich rozpatrywanych w (30) wielowskaźników $\mathbf{i} = (i_1, ..., i_n)$ i $f_{\mathbf{i}} := f_{i_1} \wedge ... \wedge f_{i_n}$ dla $\mathbf{i} \in \mathcal{I}$. W szczególności, $\bigwedge^n F = \{\mathbf{0}\}$ gdy $n > d$.

Pozostaje dowieść liniowej niezależności układu $(f_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}}$ gdy $n \le d$. Niech więc $\sum_{\mathbf{i} \in \mathcal{I}} c_{\mathbf{i}} f_{\mathbf{i}} = 0$; dowiedziemy, że wszystkie skalary $c_{\mathbf{i}}$ są zerowe. Dla uproszczenia oznaczeń zbadamy $c_{\mathbf{i}_0}$ dla $\mathbf{i}_0 = (1, ..., n)$. Mnożąc równość $\sum_{\mathbf{i}} c_{\mathbf{i}} f_{\mathbf{i}} = 0$ zewnętrznie przez $\bigwedge_{j=n+1}^d f_j$ stwierdzamy, że $\sum_{\mathbf{i}} c_{\mathbf{i}} f_{\mathbf{i}} \wedge \bigwedge_{j=n+1}^d f_j = 0$. Jak już wiemy, iloczyn $f_{\mathbf{i}} \wedge \bigwedge_{j=n+1}^d f_j$ jest zerowy gdy $\mathbf{i} \ne \mathbf{i}_0$, ze względu na powtórzenia w ciągu $i_1, ..., i_n, n+1, ..., d$; natomiast dla $\mathbf{i} = \mathbf{i}_0$ jest on niezerowy, na podstawie wniosku 1 w p.1. Zatem $c_{\mathbf{i}_0} = 0$, co kończy dowód. (Rozumowanie dla innych współczynników $c_{\mathbf{i}}$ jest analogiczne.) □

Wniosek 1. *Wymiar przestrzeni* $\bigwedge^n F$ *jest równy* $\binom{d}{n}$, *gdzie* $d := \dim F$.

Dowód. Podciągów $(i_1, ..., i_n)$ ciągu $(1, ..., d)$ jest tyle właśnie. □

Wniosek 2. *Gdy* $t \in \bigwedge^n F \setminus \{0\}$ *i* $\{f_i\}_{i=1}^d$ *jest bazą przestrzeni* F, *to* $t \wedge f_{j_1} \wedge ... \wedge f_{j_{d-n}} \neq 0$ *dla pewnych* $j_1, ..., j_{d-n} \in \{1, ..., d\}$.

Dowód. Wynika to z ostatniego akapitu dowodu twierdzenia 2. ⊟

Wniosek 3. *Gdy* $f_i \in F$ $(i = 1, ..., d)$ *i* $g_j := \sum_{i=1}^n c_{ij} f_i$ *dla* $j = 1, ..., n$, *to* $\bigwedge_{j=1}^n g_j = \sum_{\mathbf{i}} c_{\mathbf{i}} \cdot \bigwedge_{s=1}^n f_{i_s}$, *gdzie sumujemy po wszystkich* n*–elementowych podciągach* $\mathbf{i} = (i_s)_{s=1}^k$ *ciągu* $(1, ..., d)$, *zaś* $c_{\mathbf{i}} := \det(c_{i_s j})_{s,j=1}^k$. *(Gdy podciągów takich nie ma, to* $\bigwedge_j g_j = 0$.*)*

Dowód. Rozpatrzymy tylko (patrz jednak zadanie uzupełniające) najważniejszy przypadek, gdy układ (f_i) jest liniowo niezależny. Możemy założyć, że jest on bazą przestrzeni F (gdy nie jest, zmniejszymy F). Niech $t := \bigwedge_{j=1}^n g_j - \sum_{\mathbf{i}} c_{\mathbf{i}} \cdot \bigwedge_{s=1}^n f_{i_s}$; zgodnie z wnioskiem 2 należy dowieść, że $t \wedge \bigwedge_{s=1}^{d-n} f_{j_s} = 0$ gdy $1 \leq j_1 < ... < j_{d-n} \leq d$. Dla uproszczenia zapisu przyjmijmy $(j_1, ..., j_{d-n}) = (n+1, ..., d)$; otrzymamy (por. dowód twierdzenia 2)

$$t \wedge \bigwedge_{j=n+1}^d f_j = \bigwedge_{j=1}^n g_j \wedge \bigwedge_{j=n+1}^d f_j - c_{(1,...,n)} \bigwedge_{j=1}^d f_j.$$

Prawa strona jest równa $s \bigwedge_{i=1}^d f_i$ dla $s := \det(c_{ij})_{i,j=1}^n - c_{(1,...,n)} = 0$, co bez trudu wynika z wniosku 2 w p.1. Tak samo jest dla innych wielowskaźników $\mathbf{j}$ w miejsce $(1, ..., n)$. ⊟

Zadanie uzupełniające 1. a) Dowieść wniosku 3 w pełnej ogólności, w następujący sposób. Przyjąć $\widetilde{X} := X \cup \{1, ..., d\}$ i $\widetilde{f_i}(x) = f_i(x)$ dla $x \in X$, $\widetilde{f_i}(i) = 1$ i $\widetilde{f_i}(j) = 0$ gdy $i \neq j \in \{1, ..., d\}$. Wówczas funkcje $\widetilde{f_i}$ są już liniowo niezależne. Zastosować do nich i odpowiednich funkcji $\widetilde{g_j} : \widetilde{X} \to \mathbb{F}$ udowodniony przypadek szczególny.

b) Wyprowadzić z wniosku 3 twierdzenie Bineta–Cauchy'ego z §IV.3.3, i odwrotnie.

3. Iloczyny tensorowe i potęgi zewnętrzne a funkcje wieloliniowe.

Iloczyny tensorowy i zewnętrzny stosuje się w Algebrze Liniowej, gdy dziedziny X, Y, X_i itd. rozpatrywanych dotąd przestrzeni funkcyjnych są przestrzeniami liniowymi, a same przestrzenie funkcyjne składają się z funkcji wieloliniowych. Zamiast $X, Y, X_1, X_2...$ piszemy więc wtedy $V, W, V_1, V_2...$ i zakładamy, że są to skończenie–wymiarowe przestrzenie liniowe nad ustalonym ciałem $\mathbb{F}$. Ważnym przykładem liniowej przestrzeni funkcyjnej, z dziedziną V, jest przestrzeń V^* wszystkich liniowych funkcji $\varphi : V \to \mathbb{F}$.

Oznaczenia. Przypomnijmy, że przekształcenia wieloliniowe z $V_1 \times ... \times V_n$ do W to takie, które są liniowe ze względu na każdy argument $\mathbf{v}_i \in V_i$, przy ustalonych pozostałych. Tworzą one przestrzeń liniową, którą oznaczamy przez $\mathcal{L}(V_1, ..., V_n; W)$, zaś przez $\mathcal{L}_n(V; W)$ gdy $V_1 = ... = V_n = W$. Wśród przekształceń wieloliniowych można wyróżnić alternujące (=antysymetryczne, gdy $2_{\mathbb{F}} \neq 0_F$) i symetryczne. Złożone z nich podprzestrzenie przestrzeni $\mathcal{L}_n(V;W)$ oznaczamy przez $\mathcal{L}_n^a(V;W)$ i $\mathcal{L}_n^s(V;W)$, odpowiednio.

Jeśli $W = \mathbb{F}$, zamiast „przekształcenie" mówimy „funkcja" lub „forma".

Uwaga 1. a) Dla $k = 1$, każde przekształcenie liniowe z V do W jest alternujące, a zarazem jest symetryczne.

b) Przekształcenie wieliniowe na ogół nie jest liniowe, jeśli $\prod_i V_i$ rozpatrywać z naturalną strukturą przestrzeni liniowej!

Uwaga 2. a) Złożenie wieloliniowego przekształcenia $L : \prod_{i=1}^n V_i \to W$ z przekształceniem liniowym $W \to W'$ jest przekształceniem wieloliniowym. Gdy ponadto $V_1 = ... = V_n$ i przekształcenie L jest alternujące (odp. symetryczne), to i otrzymane złożenie jest takie.

b) Gdy funkcje $f : V_1 \times ... \times V_k \to \mathbb{F}$ i $g : V_{k+1} \times ... \times V_l \to \mathbb{F}$ są wieloliniowe, to wieloliniowy jest ich iloczyn tensorowy $f \otimes g : \prod_{i=1}^l V_i \to \mathbb{F}$.

c) Gdy funkcja $f : V^n \to \mathbb{F}$ jest wieloliniowa, to jej antysymetryzacja $Af : V^n \to \mathbb{F}$ też jest taka, i podobnie dla symetryzacj $Sf : V^n \to \mathbb{F}$.

Twierdzenie 1. *Niech $V, V_1, ..., V_n$ będą skończenie–wymiarowymi przestrzeniami liniowymi nad $\mathbb{F}$. Wówczas:*

a) iloczyn tensorowy przestrzeni funkcyjnych $V_1^*, ..., V_n^*$ jest zbiorem wszystkich wieloliniowych funkcji z $\prod_i V_i$ do $\mathbb{F}$, tzn. $V_1^* \otimes ... \otimes V_n^* = \mathcal{L}(V_1, ..., V_n; \mathbb{F})$;

b) n–ta potęga zewnętrzna przestrzeni funkcyjnej $\bigwedge^n V^*$ jest zbiorem wszystkich wieloliniowych funkcji alternujących z V^n do $\mathbb{F}$, tzn. $\bigwedge^n V^* = \mathcal{L}_n^a(V;\mathbb{F})$;

c) antysymetryzacja jest rzutem liniowym przestrzeni $\otimes^n V^* = \mathcal{L}_n(V;\mathbb{F})$ na $\bigwedge^n V^* = \mathcal{L}_n^a(V;\mathbb{F})$.

Dowód. By dowieść a), zastosujemy indukcję względem n. Jak zauważono, dla wieloliniowej funkcji $f : \prod_{i=1}^{n-1} V_i \to \mathbb{F}$ i funkcji liniowej $\varphi : V_n \to \mathbb{F}$, iloczyn $f \otimes \varphi$ jest wieloliniową funkcją na $\prod_{i=1}^n V_i$. Wobec łączności iloczynu tensorowego i równości (25) pozostaje dowieść, że odwrotnie, gdy h jest taką funkcją, to $h = \sum_{i=1}^s f_i \otimes \varphi_i$, dla pewnych $\varphi_1, ..., \varphi_s \in V_n^*$ i funkcji wieloliniowych $f_1, ..., f_s : \prod_{i=1}^{n-1} V_i \to \mathbb{F}$. W tym celu oznaczmy przez φ_i funkcjonał, który każdemu wektorowi $\mathbf{v} \in V$ przyporządkowuje jego i–tą współrzędną w ustalonej bazie $\mathbf{w}_1, ..., \mathbf{w}_s$. Wówczas $h(\mathbf{v}_1, ..., \mathbf{v}_{n-1}, \mathbf{v}_n) = h(\mathbf{v}_1, ..., \mathbf{v}_{n-1}, \sum_{i=1}^s \varphi_i(\mathbf{v}_n)\mathbf{w}_i) = \sum_{i=1}^s \varphi_i(\mathbf{v}_n) h(\mathbf{v}_1, ..., \mathbf{v}_{n-1}, \mathbf{w}_i)$. Przy $f_i(\mathbf{v}_1, ...\mathbf{v}_{n-1}) := f(\mathbf{v}_1, ..., \mathbf{v}_{n-1}, \mathbf{w}_i)$ zachodzi więc żądana równość $h = \sum_i f_i \otimes \varphi_i$. $\square$

Części b) i c) wynikają z a) i uwagi 2c), bo $\bigwedge^n V^* = A(\otimes^n V^*)$ i $A(\mathcal{L}_n(V;\mathbb{F})) = \mathcal{L}_n^a(V;\mathbb{F})$. $\square$

4. Tensory nad przestrzenią V.

Występujące tu przestrzenie V, W są skończonego wymiaru i nad ustalonym ciałem $\mathbb{F}$.

Jest dla nas ważne, by w parze (V, V^*) symetrycznie traktować obie przestrzenie. By to osiągnąć przyjmimy $W := V^*$, a wartość $\varphi(\mathbf{v})$ funkcjonału $\varphi \in W$ na wektorze $\mathbf{v}$ oznaczajmy też $\mathbf{v}(\varphi)$ lub $\langle \mathbf{v}, \varphi \rangle$. (Tu $\langle \cdot, \cdot \rangle$ nie jest iloczynem skalarnym.) Wówczas:

i) $\dim V < \infty$ i $\dim W < \infty$;

ii) funkcja $V \times W \ni (\mathbf{v}, \mathbf{w}) \mapsto \langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{F}$ jest dwuliniowa, oraz

iii) gdy $\mathbf{v}_0 \in V \setminus \{0\}$, to $\langle \mathbf{v}_0, \mathbf{w} \rangle \neq 0$ dla pewnego $\mathbf{w} \in W$, a gdy $\mathbf{w}_0 \in W \setminus \{0\}$, to $\langle \mathbf{v}, \mathbf{w}_0 \rangle \neq 0$ dla pewnego $\mathbf{v} \in V$.

Uwaga 1. a) Dzięki ii), każdemu wektorowi $\mathbf{v} \in V$ odpowiada funkcjonał $\varphi_{\mathbf{v}} \in W^*$, zadany wzorem $\varphi_{\mathbf{v}}(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$ dla $\mathbf{w} \in W$; przy tym z iii) wynika, że $\varphi_{\mathbf{v}} \neq 0$ gdy $\mathbf{v} \neq \mathbf{0}$. Przyporządkowanie $V \ni \mathbf{v} \mapsto \varphi_{\mathbf{v}} \in W^*$ jest więc monomorfizmem liniowym (korzystamy z ii)), co daje $\dim V \leq \dim W^* = \dim W$. Symetrycznie, $\dim W \leq \dim V$, skąd $\dim V = \dim W = \dim W^*$, a przyporządkowanie $\mathbf{v} \mapsto \varphi_{\mathbf{v}}$ jest izomorfizmem V na W^*.

b) Gdy więc warunki od i) do iii) są spełnione (i już niekoniecznie $W := V^*$), to uzasadnione jest nazwanie pary (V, W) **parą dualną**, a funkcji $\langle \cdot, \cdot \rangle$ z warunku ii) – **funkcją dualności**. Możemy bowiem wtedy zarówno utożsamić W z V^*, jak i V z W^*. (Prowadzi do tego utożsamienie każdego wektora $\mathbf{v} \in V$ z odpowiadającym mu funkcjonałem $\varphi_{\mathbf{v}} \in W^*$, zaś każdego wektora z W – z odpowiadającym mu funkcjonałem $\mathbf{v} \mapsto \langle \mathbf{v}, \mathbf{w} \rangle$ z V^*.) Z drugiej strony, każdy izmorfizm $L : V \to W^*$ wyznacza wzorem $\langle \mathbf{v}, \mathbf{w} \rangle := (L(\mathbf{v}))(\mathbf{w})$ funkcję $\langle \cdot, \cdot \rangle$ na $V \times W$, spełniającą warunki ii) oraz iii).

Przykład 1. Gdy $g : V \times V \to \mathbb{F}$ jest nieosobliwą formą dwuliniową, to warunki od i) do iii) są spełnione przy $W := V$ i $\langle \cdot, \cdot \rangle := g$.

Uwaga 2. a) Dla danej pary dualnej (V, W) będziemy więc elementy przestrzeni V traktować jako (liniowe) funkcjonały na W, a przestrzeni W – jako funkcjonały na V. Stosują się teraz definicje i wyniki z pp. 1–3, dotyczące iloczynów tensorowych przestrzeni funkcyjnych: dla każdych liczb $k, l \in \mathbb{N}$ określony jest iloczyn $V \otimes ... \otimes V \otimes W \otimes ... \otimes W$, gdzie czynników V i W jest k i l, odpowiednio. Iloczyn ten oznaczamy przez $T^k_l(V)$. Na podstawie twierdzenia 1 z p.3, możemy (i będziemy) utożsamiać $T^k_l(V)$ ze zbiorem wszystkich funkcji wieloliniowych z $W^k \times V^l$ do $\mathbb{F}$.

b) Każdą taką funkcję można poddać antysymetryzacji (lub symetryzacji) względem wybranych wskaźników $1, ..., k$, odpowiadających czynnikom W iloczynu $W^k \times V^l$, a także względem wybranych wskaźników $1, ..., l$, odpowiadających czynnikom V tego iloczynu. Wynikiem będzie nowy element przestrzeni $T^k_l(V)$.

c) W szczególności, $T^k(V)$ zawiera zbiór $\mathcal{L}^a_k(W; \mathbb{F})$ wszystkich wieloliniowych antysymetrycznych funkcji $t : W^k \to \mathbb{F}$, który oznaczymy $\bigwedge^k V$.

Umowa. Przyjmujemy konwencję, zgodnie z którą przestrzeń $(W^k \times V^l) \times (W^p \times V^q)$ utożsamiamy w naturalny sposób z $W^{k+p} \times V^{l+q}$. Pozwala to iloczyny $t \otimes t'$, dla $t \in T^k_l(V), t' \in T^p_q(V)$, traktować jako elementy $T^{k+p}_{l+q}(V)$.

Uwaga 3. Elementy przestrzeni $T^k_l(V)$ nazywamy **tensorami nad V, typu** (k,l). Celowe okaże się określenie ich *współrzędnych* względem bazy przestrzeni V czy przestrzeni W. Tu następuje zerwanie symetrii między V i W: musimy zdecydować, w której przestrzeni wybieramy bazy, i macierze zmiany baz od tego uzależniać.

Mówiąc o tensorach nad V, a nie nad W, i pisząc $T^k_l(V)$, a nie $T^k_l(W)$ czy podobnie, stwierdzamy, że bazy wybierane są w V; bazę w W zawsze obieramy jako dualną do wybranej w V. Umówiono się też wektory z W nazywać wtedy kowariantnymi, a z V – kontrawariantnymi. Latwy sposób zapamiętania tej umowy został wskazany przez Spivaka: jest ona „just opposite to what it logically ought to be". Dlatego nie będziemy jej stosować. Wektory z V od wektorów z W będziemy jednak odróżniać oznaczeniami.

Oznaczenia. Wektory z V będziemy oznaczali tłustymi literami $\mathbf{u}, \mathbf{v}, \mathbf{w}$ itp, a z W – greckimi φ, ψ itp. W przypadku numeracji wielu wektorów, wskaźniki będziemy odtąd umieszczać u góry przy wektorach z W, a u dołu przy wektorach z V; wskaźniki przy skalarach dobrze jest umieszczać tak, by sumowania iloczynów przebiegały po wskaźnikach, które powtarzają się w różnych ich czynnikach i występują w jednych u góry, a w innych u dołu.[8] (Np. indeksy skalarów w każdej ze sum $\sum_{i,j} a^{ij}\mathbf{v}_i \otimes \mathbf{v}_j$, $\sum_{i,j} a^j_i \varphi^i \otimes \mathbf{v}_j$, $\sum_{i,j} a_{ij}\varphi^i \otimes \psi^j$ zostały do tego dostosowane.) Z obu przestrzeni V i W, to W nazywać będziemy sprzężoną i oznaczać V^* (gdy mowa o tensorach nad V).

Definicja. a) Gdy $\mathbf{v}_1, ..., \mathbf{v}_r \in V$ i $\varphi^1, ..., \varphi^s \in V^*$ są ustalone, to dla wielowskaźników $\mathbf{i} = (i_1, ..., i_k) \in \{1, ..., r\}^k$ i $\mathbf{j} = (j_1, ..., j_l) \in \{1, ..., s\}^l$ piszemy

$$\mathbf{v}_\mathbf{i} := \mathbf{v}_{i_1} \otimes ... \otimes \mathbf{v}_{i_k} \in T^k_0(V),$$
$$\varphi^\mathbf{j} := \varphi^{j_1} \otimes ... \otimes \varphi^{j_l} \in T^0_l(V).$$

b) Niech $\mathcal{V} = (\mathbf{v}_i)^d_{i=1}$ będzie bazą przestrzeni V, a $\mathcal{V}^* = (\varphi^i)^d_{i=1}$ dualną z nią bazą przestrzeni V^*. Dla danego tensora $t \in T^k_l(V)$ istnieje na podstawie twierdzenia 1 w p.3 jedyny układ skalarów $t^\mathbf{i}_\mathbf{j}$ takich, że $t = \sum_{\mathbf{i},\mathbf{j}} t^\mathbf{i}_\mathbf{j} \mathbf{v}_\mathbf{i} \otimes \varphi^\mathbf{j}$, gdzie sumowanie przebiega po wszystkich $\mathbf{i} \in \{1, ..., d\}^k$ i $\mathbf{j} \in \{1, ..., d\}^l$. Skalary te nazywamy **współrzędnymi tensora t w bazie $\mathcal{V}$.**

Zadania uzupełniające.

1. Obmyśleć lub przeczytać w [Ko], jak zmienią się współrzędne tensora, gdy zmienimy bazę przestrzeni V.

2. Utwórzmy sumy proste: $T(V) = \oplus_{k,l} T^k_l(V)$ i $\oplus_k \bigwedge^k V$. W każdej z nich wskazać naturalnie określone działania, względem których jest ona algebrą łączną. (Są to **algebra tensorowa** i **zewnętrzna** przestrzeni V, odpowiednio.) Wyznaczyć ich wymiar.

[8] W związku z tym, indeksy górne często nie oznaczają już potęgowania.

5. Iloczyny tensorowe i zewnętrzne przestrzeni liniowych i ich własności uniwersalne.

Każdy wektory $\mathbf{v}$ przestrzeni liniowej V możemy interpretować jako funkcję na przestrzeni dualnej V^*, co wykorzystano w p.4. Mianowicie, wartość $\mathbf{v}(\varphi)$ tej funkcji na funkcjonale φ zdefiniowana jest jako wartość funkcjonału φ na wektorze $\mathbf{v}$, tzn. $\mathbf{v}(\varphi) := \varphi(\mathbf{v})$ dla $\varphi \in V^*$. Wzór ten zresztą nie odgrywa dalej większej roli; ważne jest to, że:

* Sumie wektorów odpowiada suma funkcji: $(\mathbf{v}_1 + \mathbf{v}_2)(\varphi) = \mathbf{v}_1(\varphi) + \mathbf{v}_2(\varphi)$ dla $\mathbf{v}_1, \mathbf{v}_2 \in V$ i $\varphi \in V^*$, i podobnie dla mnożenia przez skalar; przy tym $\mathbf{v} = \mathbf{0} \Leftrightarrow \mathbf{v}(\cdot) = 0$.

* Tak więc przestrzeń V jest izomorficzna z otrzymaną przestrzenią funkcji.

Powyższa interpretacja każdej przestrzeni liniowej jako przestrzeni funkcji pozwala dla dowolnej rodziny $V_1, ..., V_n$ przestrzeni skończonego wymiaru określić jednoznacznie iloczyn tensorowy $\otimes_{i=1}^n V_i$ otrzymanych przestrzeni funkcyjnych, który nazwiemy „funkcyjnym iloczynem tensorowym” tej rodziny. Analogicznie, określona jest „funkcyjna potęga zewnętrzna” $\bigwedge^n V$ przestrzeni V skończonego wymiaru. (Nazwy te są prowizoryczne.)

Systematyczne utożsamianie każdej przestrzeni V z przestrzenią funkcjonałów na V^* i używanie wyłącznie „funkcyjnych” iloczynów tensorowych i potęg zewnętrznych miałoby jednak istotne niedostatki. Rozpatrując np. iloczyny $V_1 \otimes V_2 \otimes V_2$ i $V_1 \otimes (V_2 \otimes V_3)$ należałoby je traktować jako przestrzenie (pewnych) funkcji na $V_1^* \times V_2^* \times V_3^*$ i na $V_1^* \times (V_2 \otimes V_3)^*$, odpowiednio, gdzie iloczyn $V_2 \otimes V_3$ też jest widziany jako zbiór (pewnych) funkcji na $V_2^* \times V_3^*$. Jest jasne, że w najlepszym razie dowiedziemy istnienia pewnego izomorfizmu (a nie równości) między $V_1 \otimes V_2 \otimes V_2$ a $V_1 \otimes (V_2 \otimes V_3)$, być może „kanonicznego”. Należy wyjaśnić, co ta „kanoniczność” oznacza i zrezygnować z jednoznaczności iloczynu tensorowego (jako zbioru), by dopuścić jego „kanonicznie izomorficzne” kopie i unikać opisów tak zawiłych, jak powyższy dla $V_1 \otimes (V_2 \otimes V_3)$.

Decydujemy się przyjąć następujące definicje, motywowane wynikami z p. 2.:

Definicja. a) Niech $V_1, ..., V_n$ będą skończenie–wymiarowymi przestrzeniami liniowymi. Ich (ogólnym) **iloczynem tensorowym** nazywamy każdą parę, złożoną z przestrzeni wektorowej Z i wieloliniowego przekształcenia $\tau : V_1 \times ... \times V_n \to Z$, taką, że

$$\dim Z = \prod_{i=1}^{n} \dim V_i \quad \text{i} \quad \mathrm{lin}(\mathrm{im}(\tau)) = Z \tag{30}$$

b) Niech V będzie przestrzenią liniową wymiaru d. Jej (ogólną) n–**tą potęgą zewnętrzną** nazywamy każdą parę, złożoną z przestrzeni liniowej Z i wieloliniowego, alternującego przekształcenia liniowego $\tau : V^n \to Z$, taką, że

$$\dim Z = \binom{d}{n} \quad \text{i} \quad \mathrm{lin}(\mathrm{im}(\tau)) = Z. \tag{31}$$

Oznaczenia (dotyczące obu definicji). Zamiast $\tau(\mathbf{v}_1, ..., \mathbf{v}_n)$ będziemy na ogół pisać $\otimes_{i=1}^n \mathbf{v}_i$ w przypadku iloczynu tensorowego, a $\bigwedge_{i=1}^n \mathbf{v}_i$ w przypadku potegi zewnętrznej; przestrzeń Z oznaczamy przez $\otimes_{i=1}^n V_i$ i $\bigwedge^n V$, odpowiednio. (Nie wymagamy już jednak, by Z, V_i czy V interpretować jako przestrzenie funkcji.)

Przykład 1. i) „Funkcyjne" iloczyny tensorowe i potęgi zewnętrzne spełniają warunki wymienione w a) i b), odpowiednio.

ii) Niech $V_1 := \mathbb{F}_{\leq k}[x]$ i $V_2 := \mathbb{F}_{\leq l}[y]$. Warunki części a) definicji są spełnione, gdy za Z wziąć zbiór tych wielomianów z $\mathbb{F}[x, y]$, ktorych stopień względem poszczególnych zmiennych x i y jest niewiększy niż k i l, odpowiednio, zaś $\tau(p_1, p_2)$ określić jako iloczyn p_1p_2, dla $(p_1, p_2) \in V_1 \times V_2$. ⊟

Twierdzenie 1. *W obu rozważanych w Definicji przypadkach, żądane pary (Z, τ) istnieją i mają następującą* **własność uniwersalności***:*

a) Własność uniwersalności w przypadku iloczynu tensorowego: dla każdej przestrzeni liniowej W i wieloliniowego przekształcenia $L : V_1 \times ... \times V_n \to W$, istnieje jedyne przekształcenie liniowe $\tilde{L} : Z \to W$ takie, że $L = \tilde{L} \circ \tau$.

b) Własność uniwersalności w przypadku potęgi zewnętrznej: dla każdej przestrzeni liniowej W i wieloliniowego przekształcenia alternującego $L : V^n \to W$, istnieje jedyne przekształcenie liniowe $\tilde{L} : Z \to W$ takie, że $L = \tilde{L} \circ \tau$.

Dowód twierdzenia podamy później, a teraz odniesiemy się teraz do „kanoniczności" przekształceń, powiedzmy z Y do Y'. Słowo to oznacza, że przekształcenie uznajemy za wzorcowe. Za takie jednak uznamy tylko przekształcenie niezależne od przypadkowych wyborów, np. od wyboru baz w Y i Y', gdy mowa o przekształceniach liniowych. Gdy ponadto Y czy/i Y' mają bogatszą strukturę, którą badamy, za kanoniczne możemy uznać i takie przekształcenia, które tę strukturę (ale tylko ją) wykorzystują. W odniesieniu do iloczynów tensorowych, kanoniczność przekształcenia $L : Y \to Y'$ oznacza więc, że zależy ono tylko od struktury liniowej obu przestrzeni Y, Y' i od przekształcenia $\tau : \prod_i V_i \to Y$, gdy (Y, τ) jest iloczynem tensorowym rodziny $\{V_i\}_{i=1}^n$. Gdy iloczynem tensorowym jest (Y', τ'), przekształcenie L może zależeć od τ', a gdy iloczynami są tak Y, jak i Y', może ono zależeć od τ i od τ'. Analogicznie rozumiemy kanoniczność, gdy Y i/lub Y' jest potęgą zewnętrzną, a także, gdy jedna z przestrzeni Y, Y' jest iloczynem tensorowym, a druga potęgą zewnętrzną. Nazwanie przekształcenia „kanonicznym" wyraża więc, prócz wyróżnienia, istotną treść matematyczną, zrozumiałą z kontekstu.

Wniosek 1. *Iloczyn tensorowy jest jednoznaczny w następującym sensie: gdy (Z, τ) i (Z', τ') są iloczynami tensorowymi rodziny $\{V_i\}_{i=1}^n$, to istnieje jedyny izomorfizm liniowy $K : Z \to Z'$ taki, że $K \circ \tau = \tau'$. Tak samo jest z n–tą potęgą zewnętrzną przestrzeni V.*

Dowód. Z własności uniwersalności pary (τ, Z), zastosowanej przy $W' := Z'$ i $L = \tau'$, wynika istnienie jedynego przekształcenia $\tilde{L} \in \mathcal{L}(Z, Z')$, takiego, że $\tilde{L} \circ \tau = \tau'$. Ponieważ $\mathrm{im}(\tau')$ rozpina Z', więc $S(Z) = Z'$, co wraz z równością $\dim Z = \dim Z'$ powoduje, że $\tilde{L}$ jest izomorfizmem. Możemy więc przyjąć $K := \tilde{L}$. □

Uwaga 1. Z jedyności i równości $K \circ \tau = \tau'$ wynika, że izomorfizm K jest kanoniczny.

W dowodzie twierdzenia wykorzystamy następujący

Lemat 1. *Dla $i = 1, ..., n$ niech B_i będzie bazą przestrzeni V_i i niech $B := \prod_i B_i$. Wówczas dla dowolnej przestrzeni liniowej W i układu jej wektorów $\{\mathbf{w}_b : b \in B\}$, istnieje jedyne przekształcenie $K \in \mathcal{L}(V_1, ..., V_n; W)$ takie, że $K(b) = \mathbf{w}_b$ dla $b \in B$.*

Dowód. Odnotujmy, że gdy $K : \prod_i V_i \to W$ jest przekształceniem wieloliniowym (niekoniecznie tym, o którym mowa w tezie) i wektory $\mathbf{v}_i \in V_i$ przedstawimy jako kombinacje wektorów bazowych: $\mathbf{v}_i = \sum_{\mathbf{b} \in B_i} c_i^{\mathbf{b}} \mathbf{b}$, to z wieloliniości uzyskujemy równość

$$K(\mathbf{v}_1, ..., \mathbf{v}_n) = K(\sum_{\mathbf{b} \in B_1} c_1^{\mathbf{b}} \mathbf{b}, ..., \sum_{\mathbf{b} \in B_n} c_n^{\mathbf{b}} \mathbf{b}) = \sum_{\mathbf{b}_1 \in B_1} ... \sum_{\mathbf{b}_n \in B_n} (\prod_{i=1}^{n} c_i^{\mathbf{b}_i}) K(\mathbf{b}_1, ..., \mathbf{b}_n). \quad (32)$$

Układ $(K(b))_{b \in B}$ wyznacza więc K jednoznacznie. Ponadto, dla danych $\mathbf{w}_b \in W$ $(b \in B)$ istnieje przekształcenie $K \in \mathcal{L}(V_1, ..., V_n; W)$ takie, że $K(b) = \mathbf{w}_b$ $\forall b \in B$. (Należy (32) wziąć za definicję, a wieloliniowość uzyskać ze wzoru.) Stąd wynika teza. ⊟

Dowód twierdzenia 1. Istnienie wynika w obu przypadkach z przykładu 1i). Pozostaje dowieść własności uniwersalności.

a). Wpierw rozpatrzmy przypadek iloczyniu tensorowego. Dla $i = 1, .., n$ ustalmy bazę B_i przestrzeni V_i i przyjmijmy $B := \prod_i B_i$. Dla danego przekształcenia $L \in \mathcal{L}(V_1, ..., V_n; W)$ przyjmijmy $\mathbf{w}_b := L(b)$ dla $b \in B$. Wiemy, że zbiór $\tau(B)$ generuje przestrzeń $\mathrm{lin}(\mathrm{im}(\tau))$, patrz wzór (32) przy $K = \tau$; a że $\mathrm{lin}(\mathrm{im}(\tau)) = Z$, to z nierówności $\#\tau(B) \leq \#B = \dim Z$ wynika, że $\tau(B)$ jest bazą przestrzeni Z. Istnieje więc jedyne przekształcenie $\tilde{L} \in \mathcal{L}(Z; W)$ takie, że $\tilde{L}(\tau(b)) = L(b)$ $\forall b \in B$. Na podstawie lematu, wieloliniowe przekształcenia $\tilde{L} \circ \tau$ i L są równe.

b). Teraz zajmijmy się przypadkiem potęgi zewnętrznej. Obierzmy w V dobrze uporządkowaną bazę $\mathcal{B}_0$ i niech $B := \mathcal{B}_0 \times ... \times \mathcal{B}_0$ (czynników jest n).

Dla każdego różnowartościowego (tzn. bez powtórzeń) ciągu $b = (\mathbf{b}_1, ..., \mathbf{b}_n) \in B$ oznaczmy przez $[b]$ jedyny ciąg z B, który jest rosnący (w porządku bazy $\mathcal{B}_0$) i różni się od b tylko kolejnością wyrazów, a przez $\mathrm{Sgn}(b)$ oznaczmy znak permutacji, która kolejnym wyrazom ciągu $[b]$ przyporządkowuje kolejne wyrazy ciągu b. Zbiór rosnących ciągów $b \in B$ oznaczmy przez C. Ze względu na antysymetrię przekształcenia τ zachodzi $\tau(C) \cup \{\mathbf{0}\} = \tau(B)$, przy czym $\mathrm{lin}(\tau(B)) = Z$, jak w a).

Tak więc $Z = \mathrm{lin}(\tau(C))$ i $\#\tau(C) \leq \#C = \dim(Z)$, skąd $\tau(C)$ jest bazą w Z. Dla danego przekształcenia $L : V^n \to W$ istnieje więc jedyne przekształcenie liniowe $\tilde{L} : Z \to W$ takie, że $\tilde{L}(\tau(b)) = L(b)$ $\forall b \in C$. Przekształcenie $\tilde{L} \circ \tau$ jest wieloliniowe i alternujące (w ślad za τ), więc gdy L też jest takie, to uzyskujemy $\tilde{L} \circ \tau(b) = \mathrm{Sgn}(b) L([b]) = L(b)$ dla różnowartościowych ciągów $b \in B$, a także $\tilde{L} \circ \tau(b) = \mathbf{0} = L(b)$ dla pozostałych $b \in B$. Tym samym z lematu 1 wynika, że $\tilde{L} \circ \tau = L$. □

Lemat 2. *a) Gdy zbiory $B_i \subset V_i$ są liniowo niezależne, to i zbiór $\widetilde{B} := \{\otimes_i \mathbf{b}_i : \mathbf{b}_i \in B_i$ dla $i = 1, ..., n\}$ jest liniowo niezależny w $\otimes_{i=1}^n V_i$. Podobnie, gdy $\mathcal{B}$ jest dobrze uporządkowanym układem liniowo niezależnym w V, to zbiór $\widetilde{C} := \{\bigwedge_{i=1}^n \mathbf{b}_i :$ ciąg $\mathbf{b}_1, ..., \mathbf{b}_n$ jest rosnący w $\mathcal{B}\}$ jest liniowo niezależny w $\bigwedge^n V$.*

b) Gdy zaś każdy z powyższych zbiorów B_i jest bazą w V_i, a układ $\mathcal{B}$ jest bazą w V, to $\widetilde{B}$ jest bazą w $\otimes_i V_i$, a $\widetilde{C}$ – bazą w $\bigwedge^n V$.

Dowód. Wystarczy dowieść tezy b), bo układ liniowo niezależny możemy rozszerzyć do bazy, a podzbiór bazy jest liniowo niezależny. Ale tezy tej dowiedliśmy już wyżej, bo przy poprzednich oznaczeniach mamy $\widetilde{B} = \tau(B)$ i $\widetilde{C} = \tau(C)$, odpowiednio. □

Tensorowo możemy mnożyć nie tylko przestrzenie, ale i przekształcenia:

Definicja. a) **Iloczynem tensorowym przekształceń liniowych** $L_i : V_i \to V_i'$ ($i = 1, ..., n$) nazywamy przekształcenie liniowe z $\otimes_{i=1}^n V_i$ do $\otimes_{i=1}^n V_i'$, przyjmujące na każdym iloczynie $\otimes_{i=1}^n \mathbf{v}_i$ wartość $\otimes_{i=1}^n L_i(\mathbf{v}_i)$. Oznaczamy je $\otimes_{i=1}^n L_i$.

b) n–**tą potęgą zewnętrzną przekształcenia** $L \in \mathcal{L}(V, V')$ nazywamy przekształcenie liniowe z $\bigwedge^n V$ do $\bigwedge^n V'$, przyjmujące na każdym iloczynie $\bigwedge_{i=1}^n \mathbf{v}_i$ wartość $\bigwedge_{i=1}^n L(\mathbf{v}_i)$. Oznaczamy [9] je $\bigwedge^n L$.

Jedyność żądanych przekształceń w obu przypadkach wynika z lematu 1, a istnienie – z twierdzenia. Np. w odniesieniu do $\bigwedge^n L$ należy zauważyć, że przekształcenie $V^n \ni (\mathbf{v}_i)_{i=1}^n \mapsto \bigwedge_{i=1}^n L(\mathbf{v}_i) \in \bigwedge^n V'$ jest wieloliniowe i alternujące, więc wyznacza przekształcenie z $\bigwedge^n V$ do $\bigwedge^n V'$, jak opisano w twierdzeniu 1b) – a to ma żądaną własność. Podkreślmy, że przy zadanych operatorach L_i czy L, iloczyny tensorowe $\otimes_{i=1}^n V_i$ i $\otimes_{i=1}^n V_i'$, czy zewnętrzne $\bigwedge^n V$ i $\bigwedge^n V'$, mogą wyżej być obrane dowolnie.

Z definicji i jednoznaczności wynika

Uwaga 2. a) Gdy $K_i \in \mathcal{L}(V_i; V_i')$ i $L_i \in \mathcal{L}(V_i', V_i'')$, to $\otimes_{i=1}^n L_i K_i = (\otimes_{i=1}^n L_i) \circ (\otimes_{i=1}^n K_i)$.

b) Podobnie, $\bigwedge^n LK = (\bigwedge^n L) \circ (\bigwedge^n K)$ dla $K \in \mathcal{L}(V; V')$ i $L \in \mathcal{L}(V', V'')$.

c) Iloczyn tensorowy przekształceń identycznościowych (odp. epimorfizmów, monomorfizmów) jest identycznością (odp. epimorfizmem, monomorfizmem), i tak samo dla potęgi zewnętrznej. (Gdy chodzi o monomorfizmy, wynika to z lematu 2a) – jak?) ⊟

Zadania uzupełniające.

1. Niech $P, Q \in \mathcal{L}(V)$ będą izomorfizmami. Przyjmijmy $P_l^k := \otimes^k P \otimes \otimes^l (P^*)^{-1}$ i analogicznie dla Q. Dowieść, że P_l^k jest izomorfizmem i $(P \circ Q)_l^k = P_l^k \circ Q_l^k$.

2. a) Dane są przekształcenia $K \in \mathcal{L}(V, V')$ i $L \in \mathcal{L}(W, W')$ oraz bazy $\mathcal{V}, \mathcal{V}', \mathcal{W}, \mathcal{W}'$ przestrzeni V, V', W, W', odpowiednio. Wyznaczyć, w zależności od $[K]_{\mathcal{V}'}^{\mathcal{V}}$ i $[L]_{\mathcal{W}'}^{\mathcal{W}}$, ma-

[9] Oznaczenia te nie prowadzą do nieporozumień, bo $\otimes_i L_i$ jest też (pewnym) iloczynem tensorowym wektorów L_i przestrzeni liniowych $\mathcal{L}(V_i, V_i')$, i podobnie dla $\bigwedge^n L$. Patrz „Dodatek" do twierdzenia 2 w następnym punkcie.

cierz przekształcenia $K \otimes L$ w uporządkowanych leksykograficznie bazach $\mathcal{V} \otimes \mathcal{W}$ i $\mathcal{V}' \otimes \mathcal{W}'$ (przestrzeni $V \otimes W$ i $V' \otimes W'$, odpowiednio).

b) Gdy $V = V'$ i $W = W'$, wyznaczyć $\mathrm{tr}(K \otimes L)$ i $\det(K \otimes L)$. (Wskazówka gdy chodzi o $\det(K \otimes L)$: zauważyć, że $K \otimes L = (K \otimes I_W)(I_V \otimes L)$, patrz uwaga 2.)

3. a) Obliczyć $\det(L \wedge L)$ dla operatora $L = L_{\mathbf{A}}$, gdzie $\mathbf{A} \in \mathcal{M}_3$.

b) Dowieść, że gdy $L \in \mathcal{L}(V)$ i $\dim(V) = d$, to $\bigwedge^d L : \bigwedge^d V \to \bigwedge^d V$ jest homotetią o skali $\det(L)$. (Wskazówka: wyrazić L w pewnej bazie; skorzystać z wniosku 2 w p.1.)

c) Wywnioskować, że wyżej $\det(L) = \mathrm{tr}(\bigwedge^d L)$ i $\det(L + xI) = \sum_{i=0}^{d} x^{d-i} \mathrm{tr}(\bigwedge^i L)$.