VII FORMY KWADRATOWE I DWULINIOWE

W tym rozdziale zakładamy, że

a) rozważane przestrzenie wektorowe są skończonego wymiaru, oraz

b) w ciele skalarów $\mathbb{F}$ element $1_\mathbb{F} + 1_\mathbb{F}$ **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 $1_\mathbb{F} + 1_\mathbb{F}$, 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. Przekształcenia takich macierzy 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 operatorach liniowych. Tu odnotujemy tylko, że 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 operator L jednoznacznie. W podobny sposób badanie operatorów i na innych przestrzeniach sprowadza się, 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.

Niestety, rozwinięcie powyższych punktów 1)–3) wykracza poza zakres wykładu i dotkniemy ich tylko w końcowych zadaniach uzupełniających.

§ 1. Wielomiany i funkcje wielomianowe stopnia ≤ 2

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

Definicja. Wielomianem stopnia ≤ 2, zmiennych $x_1, ..., x_k$ i o współczynnikach w $\mathbb{F}$, nazywamy 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 $\mathbb{F}$. Zbiór wszystkich takich wielomianów oznaczać będziemy przez $\mathbb{F}_2[x_1, ..., x_k]$. **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 powiemy, że są one **stopnia** 1 (gdy $b_i \neq 0$ dla pewnego i), bądź **stopnia** 0 (gdy $c \neq 0$ i $b_1 = \cdots = b_k = 0$), bądź też stopnia $-\infty$ (w pozostałym przypadku, tzn. gdy $p = 0$). Poprawność tej definicji w odniesieniu do funkcji wynika z następującego twierdzenia:

Twierdzenie 1. *Gdy $1_\mathbb{F} + 1_\mathbb{F} \neq 0_\mathbb{F}$, to funkcja, wyznaczona przez wielomian $p \in \mathbb{F}_2[x_1, ..., x_k]$, określa go jednoznacznie.*

Dowód. Przy oznaczeniach (1) mamy

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

Rozważana funkcja wyznacza więc współczynnik c wielomianu p, a także 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. 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.) Warto odnotować, że 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 **współczynnikach** takiej **funkcji**.

Funkcję f_2 zadaną wzorem (4) nazywamy **częścią główną**, a $f_1 + c$ –**częścią liniową** [1] rozważanej funkcji wielomanowej stopnia 2. Podobnie określamy i oznaczamy części główną i liniową wielomianu p postaci (1):

$$p_2 = \sum_{1\leq i\leq j\leq k} b_{ij}x_ix_j, \quad p_1 = \sum_{i=1}^{k} b_ix_i, \qquad \text{część liniowa} = p_1 + c.$$

Wielomian p, a także odpowiadającą mu funkcję wielomianową, 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 następujące warunki są równoważne:

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 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ funkcja $\mathbb{R} \to \mathbb{R}$ zadana przez $t \mapsto f(\mathbf{u} + t\mathbf{v})$ jest wielomianowa 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”, zaś gdy $l = 1$ –że jest „funkcją liniową”; nazwa „funkcjonał liniowy” oznacza natomiast, że $l = 1$ i $\mathbf{b} = \mathbf{0}$.

2. Formy kwadratowe a macierze.

Forma kwadratowa

$$q = \sum_{1 \le i \le j \le k} b_{ij} x_i x_j \in \mathbb{F}_2[x_1, ..., x_k] \tag{5}$$

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

$$q_{\mathbf{A}} := \sum_{i,j=1}^{k} a_{ij} x_i x_j \tag{6}$$

jak i odpowiadająca mu funkcję $\mathbb{F}^k \to \mathbb{F}$, którą oznaczymy $f_{\mathbf{A}}$. By $q_{\mathbf{A}}$ zapisać w postaci (1), należy dokonać redukcji wyrazów $x_i x_j$ oraz $x_j x_i$ $(i < j)$. Stąd i twierdzenia 1 w p.1 wynika:

Lemat 1. *Niech q i $q_{\mathbf{A}}$ zadana będą wzorami (5) i (6), odpowiednio. Równość $q = q_{\mathbf{A}}$ ma miejsce wtedy i tylko wtedy, gdy*

$$a_{ii} = b_{ii} \text{ oraz } a_{ij} + a_{ji} = b_{ij} \text{ dla } i < j \quad (i, j = 1, ..., k). \tag{7}$$

Uwaga i definicja. a) Okazuje się (jeden z powodów wskażemy w p.3), że spośród macierzy wyznaczających formę kwadratową q 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 więc wyznaczona tym, że

$$\mathbf{A} = \mathbf{A}^t \text{ i } q(\mathbf{v}) = \mathbf{v}^t \mathbf{A} \mathbf{v} \text{ dla wszystkich } \mathbf{v} \in \mathbb{F}^k$$

c) Jeśli zachodzi (5), to $a_{ii} = b_{ii}$ i $a_{ij} = a_{ji} = \frac{1}{2} b_{ij}$ dla $1 \le i < j \le k$.

Oznaczenia. Jak wcześniej, we wzorach wykorzystujących mnożenie przez macierz traktujemy skończone ciągi skalarów 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:

$$q_{\mathbf{A}} = \mathbf{x}^t \mathbf{A} \mathbf{x} \tag{8}$$

Jest to oczywiście spowodowane tym, że $q_{\mathbf{A}}(\mathbf{v}) = \sum_i v_i p_i(\mathbf{v}) = \mathbf{v}^t \mathbf{A} \mathbf{v}$ dla $\mathbf{v} \in \mathbb{F}^k$.

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

$$a_{ij} = a_{ji} = b_{ij}/2 \text{ gdy } i < j, \ a_{0i} = a_{i0} = b_i/2 \ \text{ oraz } \ a_{00} = c \ (i, j = 1, ..., k).$$

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

$$p = \widetilde{\mathbf{x}}^t \widetilde{\mathbf{A}} \widetilde{\mathbf{x}}, \text{ gdzie } \ \widetilde{\mathbf{x}} := (1, x_1, ..., x_k). \tag{9}$$

3. Upraszczanie formy podstawieniem liniowym. Kongruentność macierzy.

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

Twierdzenie 1. *Gdy* $q_{\mathbf{A}} \in \mathbb{F}_2[x_1, ..., x_k]$ *jest formą 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

$$q(\mathbf{y}) = q_{\mathbf{A}}(\mathbf{C}\mathbf{y}) = (\mathbf{C}\mathbf{y})^t \mathbf{A} (\mathbf{C}\mathbf{y}) = \mathbf{y}^t (\mathbf{C}^t \mathbf{A} \mathbf{C}) \mathbf{y}$$

Ponadto, $\mathbf{C}^t \mathbf{A} \mathbf{C}$ jest macierzą symetryczną (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 rozważaną formę: zbiór macierzy symetrycznych jest zamknięty względem odpowiadającej zamianie zmiennych relacji kongruencji, podczas gdy np. narzucający się wybór macierzy górnie trójkątnej nie prowadzi do zbioru o tej własności.

Zajmiemy się teraz możliwością wykorzystania podstawień liniowych 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 $q \in \mathbb{F}_2[x_1, ..., x_k]$ 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 dostrzeć 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.

Krok s–ty ($s = 1, ..., k$). Niech $\mathbf{B}$ oznacza macierz symetryczną, otrzymaną w wyniku wykonania poprzedzających kroków, której wyrazy różne od 0 występują 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:

Część 1. 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 kolumnach otrzymanej macierzy (dodajemy tę samą krotność t–tej kolumny do s–tej). Końcową macierz oznaczamy nadal przez $\mathbf{B}$; zauważmy, że $b_{ss} = c \cdot d \neq 0$.

Część 2. Od wierszy $s+1, s+2, ..., k$ macierzy $\mathbf{B}$ odejmujemy takie wielokrotności wiersza p, by stojące poniżej przekątnej wyrazy kolumny s uczynić zerami. Następnie, zamieniamy zerami $(s+1)$-szy i dalsze wyrazy wiersza s otrzymanej macierzy (co można też uzyskać w wyniku pewnych operacji kolumnowych).

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.

<u>Przykład 1.</u> 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):

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

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

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

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 $y_1^2 + 2y_2^2 - 6y_3^2 + 22y_4^2$.

<u>Dowód twierdzenia 2.</u> Wykażemy, że opisany sposób prowadzi do macierzy $\mathbf{C}$ o żądanych własnościach.

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 następnie 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 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 $\max(i, j) \leq s$. (Wynika to z opisu tego kroku.) Końcowa macierz $\mathbf{B}$ jest więc zarówno diagonalna, jak i równa $\mathbf{E}_n(...(\mathbf{E}_1\mathbf{AE}_1^t)...)\mathbf{E}_n^t$, 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 można przyjąć $\mathbf{C} = \mathbf{S}^t$, a $\mathbf{S} = \mathbf{C}^t$ otrzymać z klatki $\mathbf{I}_k$, wykonując kolejno wierszowe operacje algorytmu. (Korzystamy z tego, że i–ta z tych operacji jest równoważna 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 oznacza dowolną liczbę niewiększą od stopnia macierzy. Uwaga wynika stąd, że żadna z operacji wykonywanych w części 2 nie zmienia takiego 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.)

Zadania uzupełniające.

1. 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 $\mathrm{diag}(\mathbf{I}, -\mathbf{I})$.

2. 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 §....).

3. (**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.

4. 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 $\mathrm{rk}(\mathbf{A})$.

5. Udowodnić, że rzeczywista macierz symetryczna jest kongruentna z macierzą o zerowej przekątnej.

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 ciałem $\mathbb{F}$.

Definicja. Powiemy, że funkcji tej **w bazie** $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ **odpowiada wielomian** $p \in \mathbb{F}_2[x_1, ..., x_k]$ (lub: że funkcja f jest w bazie $\mathcal{V}$ **zadana wielomianem** p), jeśli zachodzi tożsamość

$$f(\lambda_1\mathbf{v}_1 + ... + \lambda_k\mathbf{v}_k) = p(\lambda_1, ..., \lambda_k) \text{ dla } \lambda_1, ..., \lambda_k \in \mathbb{F} \tag{10}$$

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 wielomian stopnia ≤ 2. (Wynika to z twierdzenia 1 w p.1.)

c) Latwo o przykład funkcji, której w żadnej bazie nie w odpowiada jakikolwiek wielomian. Gdy $V = \mathbb{R} = \mathbb{F}$, to jest nią np. funkcja sin – bo ma nieskończenie wiele zer, a wielomian $p \in \mathbb{R}[x]$ ma ich skończenie wiele.

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

[2] patrz dalej zad. uz. 2 w §2.4

Dowód. Tożsamość (11) oznacza, że $f(\mathbf{v}) = p([\mathbf{v}]_{\mathcal{V}})$ dla $\mathbf{v} \in V$. Ponieważ $[\mathbf{v}]_{\mathcal{V}} = \mathbf{C}[\mathbf{v}]_{\mathcal{W}}$, więc $f(\mathbf{v}) = p(\mathbf{C}[\mathbf{v}]_{\mathcal{W}})$ dla $\mathbf{v} \in V$ – co wraz z definicją wielomianu p' daje żądaną tezę. □

Definicja. a) Powiemy, że funkcja $f : V \to \mathbb{F}$ 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** $f : V \to \mathbb{F}$ **w bazie** $\mathcal{V}$ przestrzeni V nazywamy macierz odpowiadającej jej formy $q \in \mathbb{F}_2[x_1, ..., x_k]$. Jest to więc macierz symetryczna $\mathbf{A}$ taka, że dla wszystkich $\mathbf{v} \in V$ zachodzi $f(\mathbf{v}) = \mathbf{x}^t\mathbf{A}\mathbf{x}$, gdzie $\mathbf{x} = [\mathbf{v}]_{\mathcal{V}}$.

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 pewnej bazie $\mathcal{V} = (\mathbf{v}_1, .., \mathbf{v}_k)$ 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 tej 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 $\lambda_1, ..., \lambda_k$. □

Definicja. O bazie $\mathcal{W}$ powiemy, że **diagonalizuje formę kwadratową** $f : V \to \mathbb{F}$, jeśli macierz formy f w bazie $\mathcal{W}$ jest diagonalna. Odnotujmy, że w powyższym do-

[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.

wodzie baza diagonalizująca $\mathcal{W} = (\mathbf{w}_1, .., \mathbf{w}_k)$ określona jest 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 2. Oba twierdzenia diagonalizacyjne (powyższe i z p.3) nazywane są **twierdzeniem Lagrange'a o diagonalizacji form kwadratowych**.

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

Zadanie 1. Dowieść, że gdy V i W są przestrzeniami wektorowymi nad ciałem $\mathbb{F}$, to złożenie $f \circ L$ formy kwadratowej $f : W \to \mathbb{F}$ z operatorem $L \in \mathcal{L}(V, W)$ jest formą kwadratową. Wyrazić macierz tej formy przez macierz formy f w bazie $\mathcal{W}$ i macierz $[L]_{\mathcal{W}}^{\mathcal{V}}$, dla danych baz $\mathcal{V}$ i $\mathcal{W}$ przestrzeni V i W, odpowiednio.

<u>Zadanie uzupełniające 1.</u> 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 twierdzenia 1 w p.1.)

<u>Zadania ze zbioru Kostrykina:</u> II.2.2.32

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

Poza fragmentami punktu 1, 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. Twierdzenie o bezwładności.

Danej formie kwadratowej $f : V \to \mathbb{R}$ na rzeczywistej przestrzeni wektorowej V odpowiadają w różnych bazach diagonalizujących różne wielomiany postaci $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$. 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 $\mathcal{V}$ wielomian $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$ $(k = \dim V)$, to liczba s dodatnich współczynników λ_i jest od bazy diagonalizującej $\mathcal{V}$ niezależna, i tak samo jest dla współczynników ujemnych.*

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

c) Gdy podstawienia liniowe przeprowadzają pewną formę $q \in \mathbb{R}_2[x_1, ..., x_k]$ 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 $\mathrm{diag}(\lambda_1, ..., \lambda_k)$ i $\mathrm{diag}(\lambda_1', ..., \lambda_k')$ w c) są kongruentne. Z kolei, b) wynika z a), bo kongruentne macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ są macierzami, w różnych bazach, pewnej wspólnej formy kwadratowej $\mathbb{F}^k \to \mathbb{F}$. (Korzystamy z wniosku 1 w §1.4.) Pozostaje dowieść a), i to tylko w odniesieniu do liczby s – bo gdy jej niezależność od bazy zastosować do formy $-f$, to otrzymamy niezależność od bazy liczby t. Teza wynika więc z poniższego lematu, wyrażającego s w sposób niezależny od bazy:

Lemat 1. *Przy oznaczeniach części a) twierdzenia, $k-s$ jest maksymalnym wymiarem liniowych podprzestrzeni przestrzeni V, zawartych w $f^{-1}((-\infty, 0])$.*

Dowód. Niech dodatnimi współczynnikami będą $\lambda_1, ..., \lambda_s$ i niech $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$. Twierdzimy, że

$$\text{Przy } W_+ := \mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_s) \text{ zachodzi } f(\mathbf{w}) > 0 \text{ dla } \mathbf{w} \in W_+ \setminus \{\mathbf{0}\}. \tag{11}$$

Istotnie, gdy $\mathbf{w} = \mu_1 \mathbf{v}_1 + \cdots + \mu_s \mathbf{v}_s$, gdzie $\mu_i \neq 0$ dla pewnego i, to $f(\mathbf{v}) = \lambda_1 \mu_1^2 + \cdots + \lambda_s \mu_s^2 > 0$.

Dla dowolnej podprzestrzeni W przestrzeni V, jeśli więc $f(W) \subset (-\infty, 0]$, to $W \cap W_+ \subset \{\mathbf{0}\}$, skąd $\dim W \leq k - \dim W_+ = k - s$; patrz wniosek 1 w §III.6.1. Z drugiej strony, podprzestrzeń $\mathrm{lin}(\mathbf{v}_{s+1}, ..., \mathbf{v}_k)$ jest wymiaru $k - s$ i f przyjmuje na niej tylko nieujemne wartości. (Uzasadnienie jak dla (11).) Dowodzi to tezy lematu. □

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.

b) Podobnie definiujemy i oznaczamy indeksy bezwładności i sygnaturę macierzy $\mathbf{A}$, która jest rzeczywista i symetryczna: są one równe indeksom bezwładności i sygnaturze wyznaczonej przez $\mathbf{A}$ formy $f_{\mathbf{A}} : \mathbb{R}^k \to \mathbb{R}$.

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. □

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

b) Kongruentne macierze mają tę samą sygnaturę, i odwrotnie. (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 bazy diagonalizującej formę wynika to z definicji, a dla innej – z b), bo macierze formy f w różnych bazach są kongruentne.

Ćwiczenie. Wyznaczyć sygnaturę formy $\det : \mathcal{M}_2(\mathbb{R}) \to \mathbb{R}$ oraz jej obcięcia do $\{\mathbf{A} \in \mathcal{M}_2 : \mathrm{tr}(\mathbf{A}) = 0\}$.

Wniosek 1. *Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, zaś $f : V \to \mathbb{F}$ niech będzie niezerową formą kwadratową.*

a) Jeśli $\mathbb{F} = \mathbb{C}$, to w pewnej bazie przestrzeni V formie f odpowiada wielomian postaci $x_1^2 + ... + x_r^2$, dla pewnego $r \leq \dim(V)$.

b) Jeśli $\mathbb{F} = \mathbb{R}$, to w pewnej bazie przestrzeni V formie f odpowiada wielomian $x_1^2 + ... + x_s^2 - x_{s+1}^2 - ... - x_{s+t}^2$, gdzie $s, t \geq 0$ i $s + t \leq \dim(V)$.

c) Liczby s, t w części b) oraz liczba r w części a) są przez formę f jednoznacznie wyznaczone warunkami $(s, t) = \sigma(f)$ oraz $r = \mathrm{rk}(f)$.

Dowód. Ad a). Na podstawie twierdzenia Lagrange'a, formie f odpowiada w pewnej bazie $\mathbf{v}_1, ..., \mathbf{v}_k$ wielomian postaci $\sum_{i=1}^r \lambda_i x_i^2$, gdzie $\lambda_i \neq 0$ dla $i = 1, ..., r$. Latwo widzieć, że w bazie $\frac{1}{\sqrt{\lambda_1}}\mathbf{v}_1, ..., \frac{1}{\sqrt{\lambda_r}}\mathbf{v}_r, \mathbf{v}_{r+1}, ..., \mathbf{v}_k$ odpowiada jej wielomian $\sum_{i=1}^r x_i^2$.

Dowód pozostałych części pozostawiony jest jako ćwiczenie.⊟

Uwaga 2. Oto równoważne sformułowanie wniosku: dany jednorodny wielomian kwadratowy nad $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ można podstawieniem liniowym przeprowadzić w dokładnie jeden wielomian opisanej w a) i b) postaci.

Definicja. Nazwijmy funkcje $f : V \to \mathbb{F}$ i $f' : W \to \mathbb{F}$ **(liniowo) równoważnymi**, gdy istnieje izomorfizm liniowy $T : V \to W$ taki, że $f \circ T = f'$. (Zakładamy, że V i W są przestrzeniami wektorowymi nad $\mathbb{F}$; ograniczenie $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ nie jest tu istotne.)

Zadanie 1. Niech f i f' będą formami kwadratowymi na przestrzeni wektorowej V nad $\mathbb{F}$.

a) Gdy $\mathbb{F} = \mathbb{C}$, to dla równoważności f i f' potrzeba i wystarcza, by $\mathrm{rk}(f) = \mathrm{rk}(f')$.

b) Gdy $\mathbb{F} = \mathbb{R}$, to dla równoważności f i f' potrzeba i wystarcza, by $\sigma(f) = \sigma(f')$.

Uwaga 3. Wyniki dotyczące ciała $\mathbb{C}$ pozostają w mocy, gdy tylko w ciele skalarów każdy element ma pierwiastek kwadratowy. W przypadku dowolnego ciała $\mathbb{F}$, warunkiem koniecznym równoważności form kwadratowych $f, f' : V \to \mathbb{F}$ jest to, by macierze $\mathbf{A}$ i $\mathbf{A}'$ tych form w dowolnej bazie przestrzeni V spełniały warunek $\det(\mathbf{A}) = c^2 \det(\mathbf{A}')$, dla pewnego skalara $c \neq 0$. (Dlaczego?) Warunek ten nie jest jednak wystarczający i na ogół zagadnienie, czy równoważne są dane dwie formy kwadratowe nad ciałem $\mathbb{F} \notin \{\mathbb{R}, \mathbb{C}\}$, jest trudne. (Dla ciała $\mathbb{Q}$ liczb wymiernych jego dyskusji poświęcona jest książka.)

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

1. 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 $\mathbb{R} \ni x \mapsto p(x\mathbf{v})$ przyjmuje wszystkie wartości nieujemne.

2. Niech $f : V \to \mathbb{R}$ będzie formą kwadratową i niech $\dim(V) = k$.

a) Dowieść, że maksimum wymiarów podprzestrzeni W przestrzeni V, zawartych w $f^{-1}(0)$, jest równe $k - \max(s, t)$, gdzie $(s, t) := \sigma(f)$.

b) Wyznaczyć analogiczne maksimum przy $f^{-1}(0)$ zastąpionym przez $\{\mathbf{0}\} \cup f^{-1}((0, \infty))$.

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

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

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

4. Niech $p \in \mathbb{R}_2[x_1, ..., x_k]$ 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 $\sup_{\mathbf{v} \in V_0} p(\mathbf{v} + \mathbf{w}) = \infty$.

5. Niech $q \in \mathbb{R}_2[x_1, ..., x_k]$ 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 $\ell_1, ..., \ell_{s+t} \in \mathbb{R}_1[x_1, ..., x_k]$. (Wielomian liniowy ℓ nazywamy formą, gdy $\ell(\mathbf{0}) = 0$.)

6. Niech $\mathbf{A}$ będzie macierzą formy kwadratowej $f : V \to \mathbb{R}$ w pewnej bazie $(\mathbf{v}_1, ..., \mathbf{v}_k)$ przestrzeni V.

a) Wyrazić znak wyznacznika macierzy $\mathbf{A}$ przez sygnaturę formy f.

b) Dowieść, że liczba $\sum_i f(\mathbf{v}_i)$ jest równa sumie wartości własnych macierzy $\mathbf{A}$ (powtarzanych zgodnie z ich krotnościami algebraicznymi).

7. Rozważmy następujące własności niezerowej formy kwadratowej $q \in \mathbb{F}_2[x_1, ..., x_k]$:

a) q jest kwadratem wielomianu stopnia 1;

b) q jest iloczynem dwóch takich wielomianów.

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.)

2. Ortogonalna diagonalizacja 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.3, a także tw.2 z §3.2.

Definicja. Gdy $\mathbf{C} \in \mathcal{M}_k(\mathbb{R})$ jest macierzą ortogonalną i podstawienie $\mathbf{x} = \mathbf{Cy}$ przeprowadza formę $q \in \mathbb{R}_2[x_1, ..., x_k]$ 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 $q \in \mathbb{R}_2[x_1, .., x_k]$ 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 odpowiada w pewnej bazie ortonormalnej wielomian postaci $\lambda_1 x_1^2 + ... + \lambda_k x_k^2$, gdzie $k = \dim(V)$. Współczynniki $\lambda_1, ..., \lambda_k \in \mathbb{R}$ są wartościami własnymi macierzy formy w dowolnej bazie ortonormalnej, każda wartość powtórzona zgodnie z jej krotnością algebraiczną.

Dowód. Ad a). Niech $\mathbf{A}$ będzie macierzą formy q. Ponieważ $\mathbf{A} = \mathbf{A}^t$, więc istnieje macierz ortogonalna $\mathbf{C}$ taka, że $\mathbf{D} := \mathbf{C}^{-1}\mathbf{AC}$ jest macierzą diagonalną. Wówczas $\mathbf{C}^t\mathbf{AC} = \mathbf{D}$, skąd $\mathbf{x} = \mathbf{Cy}$ jest żądanym podstawieniem. Dalsza część a) wynika z podobieństwa $\mathbf{A}$ do macierzy $\mathbf{D} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, 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. (Uwzględniamy krotności algebraiczne wartości własnych.)*

b) Sygnatura formy kwadratowej na rzeczywistej przestrzeni wektorowej V wynosi (s, t), gdzie s jest liczbą dodatnich, a t liczbą ujemnych wartości własnych macierzy formy w dowolnej bazie przestrzeni V.

Dowód. a) wynika z części a) twierdzenia, zaś b) –z a) i uwagi 1b) w p.1. (Należy V traktować jako przestrzeń euklidesową, z dowolnie obranym iloczynem skalarnym.) □

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

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

1. Niech $f : V \to \mathbb{R}$ będzie formą kwadratową na przestrzeni euklidesowej V, a $(\lambda_1, \ldots, \lambda_k)$ ciągiem wszystkich wartości własnych (z powtórzeniami) jej macierzy w ortonormalnej bazie przestrzeni V. Dowieść, że:

a) Dla $\lambda \in \mathbb{R}$, liczba $\#\{i : \lambda_i \geq \lambda\}$ jest równa maksimum wymiarów podprzestrzeni W takich, że $f_{|W \cap S} \geq \lambda$, gdzie $S := \{\mathbf{v} \in \mathbb{R}^k : ||\mathbf{v}|| = 1\}$ oznacza sferę jednostkową.

b) Ma miejsce następująca **równość Couranta–Fischera**:

$$\lambda_i = \inf\{\max(f_{|W \cap S}) : W \subset \mathbb{R}^k \text{ jest podprzestrzenią wymiaru } k - i + 1\}.$$

2. Niech dalej $f' := f_{|V'}$ oznacza obcięcie formy f do pewnej podprzestrzeni $V' \subset V$, i niech $\lambda_1 \geq \lambda_2 ... \geq \lambda_k$ i $\mu_1 \geq ... \geq \mu_l$ będą wartościami własnymi macierzy formy f' w pewnych bazach ortonormalnych przestrzeni V i V', odpowiednio. Dowieść, że:

a) $\lambda_i \geq \mu_i \geq \lambda_{k-l+i}$ dla $i = 1, ..., l$.

b) $\lambda_1 + ... + \lambda_s = \sup\{\sum_{i=1}^s f(\mathbf{v}_i) :$ układ $\mathbf{v}_1, ..., \mathbf{v}_s$ jest ortonormalny$\}$, $\forall s \leq k$.

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 $\mathbf{A}$. (Uwzględniamy krotności wartości własnych.)

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

4. Określoność rzeczywistych form kwadratowych.

Niech f będzie formą kwadratową na rzeczywistej przestrzeni liniowej V.

<u>Definicja.</u> 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 $\mathbf{A} \in \mathcal{M}_k$ 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**, lub że jest **określona nieujemnie** (odp. **niedodatnio**). W pozostałym przypadku formę czy macierz nazywamy **nieokreśloną.**

<u>Przykład 1.</u> Forma $f(x_1, x_2, x_3) = (x_1 - x_2)^2 + (x_1 - 2x_3)^2$ jest nieujemnie określona, 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 $\mathbf{A}$. (Wynika to stąd, że $f(\mathbf{v}) = [\mathbf{v}]_{\mathcal{V}}^t \mathbf{A}[\mathbf{v}]_{\mathcal{V}} = f_{\mathbf{A}}([\mathbf{v}]_{\mathcal{V}})$ dla $\mathbf{v} \in V$.)

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

c) Macierz diagonalna jest określona dodatnio (odp. ujemnie etc.) wtedy i tylko wtedy, gdy wszystkie wyrazy jej przekątnej są dodatnie (odp. ujemne etc.). Określoność można więc badać diagonalizując przez kongruencję macierz formy.

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

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 on wyznaczony przez pierwszych jej s wierszy i kolumn, dla pewnej liczby s.

Twierdzenie 1. *Macierz symetryczna $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ wtedy i tylko wtedy jest dodatnio określona, gdy dodatnie są wszystkie jej minory początkowe.*

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

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

b) Na mocy uwagi 1a) i zadania 1, macierz $\mathbf{A}$ wtedy i tylko wtedy jest dodatnio określona, gdy macierz $\mathbf{Q}$ ma tę własność.

Teza twierdzenia (oczywista dla $k = 1$) wynika więc przez indukcję względem k. □

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. *Macierz symetryczna $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ 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 Jacobiego-Sylvestera**. Zapiszmy je przy $k = 3$:

$$\mathbf{A} = \begin{pmatrix} \underline{a_{11}}| & a_{12}| & a_{13}| \\ \underline{a_{21}} & \underline{a_{22}}| & a_{23}| \\ \underline{a_{31}} & \underline{a_{32}} & \underline{a_{33}}| \end{pmatrix} \qquad \text{(zaznaczono klatki początkowe)}$$

a) $\mathbf{A}$ jest macierzą dodatnio określoną $\Leftrightarrow (a_{11} > 0$ i $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0$ i $\det(\mathbf{A}) > 0)$;

b) $\mathbf{A}$ jest macierzą ujemnie określoną$\Leftrightarrow (a_{11} < 0$ i $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0$ i $\det(\mathbf{A}) < 0)$.

Uwaga 2. Macierz $\operatorname{diag}(0, 1, -1)$ ma zerowe minory początkowe i jest nieokreślona.

Zadania uzupełniające.

1. Niech $\mathbf{A}$ będzie macierzą symetryczną nad $\mathbb{F}$ (ograniczenie $\mathbb{F} = \mathbb{R}$ jest tu zbędne) 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 $\mathbf{D} = \operatorname{diag}(a_1, a_2/a_1, ..., a_r/a_{r-1}, 0, ..., 0)$, a $\mathbf{C}$ jest macierzą górnie trójkątna, z jedynkami na przekątnej.

Uwaga 3. Rezultat ten nosi nazwę **twierdzenia Jacobiego.** Wynika z niego, że gdy wszystkie minory początkowe $a_1, \ldots, a_r$ macierzy symetrycznej $\mathbf{A}$ są 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 $a_0 := 1$.

2. Dla macierzy $\mathbf{A} = \mathbf{A}^t \in \mathcal{M}_k(\mathbb{R})$ dowieść implikacji $a) \Rightarrow b) \Rightarrow c)$ 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 nieujemnie określona;

   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.

3. Gdy macierz symetryczna $\mathbf{A}$ jest nieujemnie określona, to $\mathbf{A} = \mathbf{B}^t\mathbf{B}$ dla pewnej macierzy górnie trójkątnej $\mathbf{B}$ o nieujemnych wyrazach na przekątnej. (Jest to tzw. **rozkład Cholesky'ego** macierzy $\mathbf{A}$, istotny dla metod numerycznych algebry liniowej.)

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

§ 3. Formy (funkcje) dwuliniowe.

1. Funkcje dwuliniowe i ich macierze.

Definicja. Niech V będzie przestrzenią wektorową. Funkcję $g : V \times V \to \mathbb{F}$ (gdzie $\mathbb{F}$ to ciało skalarów przestrzeni V) nazywamy **dwuliniową**, gdy

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

Funkcję dwuliniową $V \times V \to \mathbb{F}$ 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

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

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

Definicja. **Macierzą funkcji dwuliniowej** $g : V \times V \to \mathbb{F}$ **w bazie** $\mathbf{v}_1, ..., \mathbf{v}_k$ przestrzeni V nazywamy macierz, której (i, j)–ty wyraz jest równy $g(\mathbf{v}_i, \mathbf{v}_j)$, dla $i, j = 1, ..., k$.

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:*

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

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

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

b) Dla $\mathbf{u}, \mathbf{w} \in V$ mamy $([\mathbf{u}]_{\mathcal{V}})^t \mathbf{A} [\mathbf{v}]_{\mathcal{V}} = g(\mathbf{u}, \mathbf{w}) = ([\mathbf{u}]_{\mathcal{W}})^t \mathbf{B} [\mathbf{w}]_{\mathcal{W}}$ oraz $[\mathbf{u}]_{\mathcal{V}} = \mathbf{C}[\mathbf{u}]_{\mathcal{W}}$ i $[\mathbf{w}]_{\mathcal{V}} = \mathbf{C}[\mathbf{w}]_{\mathcal{W}}$. Stąd wynika, że $\mathbf{x}^t \mathbf{B} \mathbf{y} = \mathbf{x}^t \mathbf{C}^t \mathbf{A} \mathbf{C} \mathbf{y}$ dla $\mathbf{x}, \mathbf{y} \in \mathbb{F}^k$, wobec czego $\mathbf{B} = \mathbf{C}^t \mathbf{A} \mathbf{C}$. (Por. §1.4.)⊟

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

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 $\mathbf{u}, \mathbf{v} \in V$.

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 (15) 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 więc kongruentne. Dlatego niezmienników kongruentności macierzy użyć można do zdefiniowania własności funkcji g. W szczególności, **rzędem funkcji** g nazywamy rząd jej macierzy w dowolnej bazie przestrzeni V. Funkcję g 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 funkcji g w dowolnej bazie przestrzeni.)

Macierz symetrycznej formy dwuliniowej można diagonalizować w oparciu o wcześniejsze wyniki:

Stwierdzenie 2. *W odpowiednio dobranej bazie, macierz symetrycznej funkcji dwuliniowej jest diagonalna.*

Dowód. Niech $\mathbf{A}$ będzie macierzą rozważanej funkcji w dowolnie obranej bazie $\mathcal{V}$. 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).)

Uwaga 2. Dowód powyższy daje odmienną, niż dyskutowana w rozdziale IV, metodę konstruowania baz ortogonalnych przestrzeni euklidesowych.

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 $(g(\mathbf{v}_i, \mathbf{v}_j))_{1 \le i,j \le k}$ 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 $1_{\mathbb{F}} + 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$ (zaś jest nieprawdziwa, co nieco trudniejsze, gdy $1_{\mathbb{F}} + 1_{\mathbb{F}} = 0_{\mathbb{F}}$). Patrz też ......

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 $\mathbf{u}, \mathbf{w} \in V$ i uzasadnić przy jej pomocy nierówność CBS dla standardowego iloczynu skalarnego w $\mathbb{R}^n$.

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 macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{l,k}$ mają tę własność, że dla każdego $\mathbf{v} \in \mathbb{F}^k$ wektor $\mathbf{Bv}$ jest proporcjonalny do $\mathbf{Av}$. Dowieść, że $\mathbf{B} = \lambda\mathbf{A}$ dla pewnego skalara λ.

6. Niech g, h będą funkcjami dwuliniowymi na przestrzeni V. Dowieść, że

a) Jeśli $g(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow h(\mathbf{u}, \mathbf{v}) = 0$, to $h = \lambda g$ dla pewnego skalara λ. (Wskazówka: przy $V = \mathbb{F}^k$ użyteczne może być powyższe zadanie.)

b) Jeśli $g(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow g(\mathbf{v}, \mathbf{u}) = 0$ $(\mathbf{u}, \mathbf{v} \in V)$, to funkcja g jest symetryczna lub jest antysymetryczna.

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 $g : V \times V \to \mathbb{F}$ jest funkcją dwuliniową, 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 funkcja* 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,*

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

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 wówczas oczywiste. Z nich wynika łatwo tożsamość (17), a z tej jedyność g. ⊟

Uwaga 1. (i definicja) Gdy symetryczna funkcja dwuliniowa g oraz forma kwadratowa f pozostają w zależności (16), to o każdej z nich mówimy, że jest **wyznaczona** przez pozostałą. Inne stosowane nazwy to: g jest **formą** (lub: **funkcją**) **biegunową** formy kwadratowej f. Z twierdzenia wynika, że w dowolnej bazie przestrzeni V, macierz formy kwadratowej f jest równa macierzy jej funkcji biegunowej g.

Formułę, pozwalającą jawnie wyrazić funkcję biegunową g przez formę f, nazywamy **polaryzacyjną**. Dwóch przykładów takich formuł dostarcza tożsamość (17).

Przykład 1. Niech $V = \mathcal{M}_2(\mathbb{F})$. Funkcja $\det : V \to \mathbb{F}$ jest formą kwadratową: 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 ona 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 też zgadnąć bezpośrednio: prawa strona jest symetryczną funkcją dwuliniową, zaś dla $\mathbf{Y}=\mathbf{X}$ przyjmuje wartość $\det(\mathbf{X})$.

Przykład 2. Niech $(V, \langle\cdot,\cdot\rangle)$ będzie przestrzenią unitarną nad $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$ i niech $f(\mathbf{v}) = ||\mathbf{v}||^2$ 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 następującego 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. Wobec twierdzenia z §1.4 istnieje więc jej ortonormalna baza $(\mathbf{v}_i)_{i=1}^k$, 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).□

Uwaga 2. * Niech macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{R})$ będą symetryczne, w tym $\mathbf{A}$ dodatnio określona. Z twierdzenia 2 wynika istnienie macierzy nieosobliwej $\mathbf{C}$ takiej, że $\mathbf{C}^t\mathbf{AC} = \mathbf{I}$ i $\mathbf{C}^t\mathbf{BC} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, dla pewnych $\lambda_1, ..., \lambda_k \in \mathbb{R}$. Liczby λ_i można jawnie wyznaczyć: są one pierwiastkami wielomianu $\chi := \det(\mathbf{B} - x\mathbf{A})$ (z krotnościami), bo są pierwiastkami wielomianu $\det(\mathbf{C}^t\mathbf{BC} - x\mathbf{I})$, proporcjonalnego do χ. ⊟

Zadania uzupełniające.

1. 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, zaś macierz $\mathbf{C}$, mająca je jako kolumny, spełnia warunki uwagi 2.

2. (Wskazówka: twierdzenie 2.) Niech $\mathbf{A}$ będzie macierzą dodatnio określoną.
   a) Dowieść, że gdy macierz $\mathbf{B}$ jest symetryczna, to $\det(\mathbf{A} + i\mathbf{B}) \neq 0$.
   b) Dowieść, że gdy ponadto $\mathbf{B}$ jest nieujemnie określona, to $\det(\mathbf{A} + \mathbf{B}) \geq \det(\mathbf{A})$ i nierówność jest ostra gdy $\mathbf{B} \neq \mathbf{0}$.

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).

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 ortogonalności 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 2 w p.1) lub symetryczne, są więc geometrycznie wyróżnione; nazwiemy je **formami metrycznymi**.

Uwaga 1. Należy podkreślić, że nazwa „forma metryczna" jest umowna i myląca: forma taka na ogół nie wyznacza na przestrzeni żadnej metryki w sensie znanym z wykładów Analizy czy Topologii. Jednak spotykana też nazwa „iloczyn skalarny" była już użyta w rozdziale V w innym (choć pokrewnym) znaczeniu.

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 $\mathbf{u} \perp \mathbf{v}$. Zamiast $g(\mathbf{u}, \mathbf{v})$ 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, i te nazwiemy **osobliwymi**. Istnieć też mogą wektory $\mathbf{v} \neq \mathbf{0}$ takie, że $\mathbf{v} \perp \mathbf{v}$; nazywamy je **izotropowymi**. Ogólniej, podprzestrzeń $U \neq \{\mathbf{0}\}$ nazwiemy **izotropową** (lub: **całkowicie osobliwą**), gdy $g_{|U \times U} = 0$. [4] Natomiast **przestrzeń anizotropowa** lub **określona** to taka, w której nie ma wektorów izotropowych.

Uwaga 2. Przestrzeń z wyróżnioną formą symetryczną jest izotropowa wtedy i tylko wtedy, gdy każdy jej wektor jest izotropowy. (Wynika to z formuły polaryzacyjnej (17) 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 IV. Modyfikacje te wymagają pewnej ostrożności: intuicja może zawodzić, gdyż trzeba w sformułowaniach lub dowodach uwzględniać 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".

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 infor-

[4] Jest to terminologia np. J. P. Serre'a. U nowszych autorów, „(pod)przestrzeń izotropowa" to taka, której pewien wektor jest izotropowy – co nie odpowiada znaczeniu słowa „izotropowy" (jednorodny we wszystkich kierunkach).

mację o formie metrycznej g (wynika to z twierdzenia z p.2). W przypadku symetrycznym często mówimy, że rozważana ortogonalność zadana jest przy pomocy formy kwadratowej, zamiast, że jest zadana przez dwuliniową funkcję biegunową tej ostatniej.

Przykład 1. Forma kwadratowa $x^2 + y^2 + z^2 - t^2$ zadaje na $\mathbb{R}^4$ tzw. **geometrię Minkowskiego**, zaś odpowiadająca jej forma dwuliniowa to $xx' + yy' + zz' - tt'$. Ogólniej, o geometrii Minkowskiego można mówić, gdy rozważamy parę (V, f), gdzie V jest rzeczywistą przestrzenią wektorową wymiaru 4, zaś f jest 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 $(1, 1)$. □

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

Dowód. Jest to przeformułowanie stwierdzenia 2 w p.1. □

Uwaga 3. a) Inny (interesujący, bo geometryczny) dowód twierdzenia 2 wskazuje zadanie uz. 2 w §4.1.

b) Forma symetryczna g jest w g–ortogonalnej bazie $(\mathbf{v}_i)_{i=1}^k$ zadana wielomianem $\sum \lambda_i x_i y_i$, gdzie $\lambda_i = g(\mathbf{v}_i, \mathbf{v}_i)$. Liczba nieizotropowych wektorów $\mathbf{v}_i$ jest równa $\mathrm{rk}(g)$, zaś gdy V jest przestrzenią rzeczywistą, to $\sigma_+(g) = \#\{i : g(\mathbf{v}_i, \mathbf{v}_i) > 0\}$ oraz $\sigma_-(g) = \#\{i : g(\mathbf{v}_i, \mathbf{v}_i) < 0\}$. (Ćwiczenie: uzasadnić te stwierdzenia.)

Definicja. Niech w przestrzeniach V i W wyróżnione będą formy metryczne g i h, odpowiednio. Przekształcenie $L \in \mathcal{L}(V, W)$ nazywamy **zanurzeniem izometrycznym**, gdy jest ono różnowartościowe i $h(L(\mathbf{v}_1), L(\mathbf{v}_2)) = g(\mathbf{v}_1, \mathbf{v}_2)$ dla wszystkich $\mathbf{v}_1, \mathbf{v}_2 \in V$. Zanurzenie izometryczne, będące izomorfizmem, nazywamy **izometrią**. Gdy izometria taka istnieje, przestrzenie (V, g) i (W, h) nazywamy **izometrycznymi**, zaś formy g i h nazywamy **równoważnymi**.

Zadanie 1. Przy oznaczeniach definicji, niech $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ i $\mathcal{W}$ będą bazami w V i W, odpowiednio. Dowieść, że gdy przekształcenie L jest różnowartościowe, to równoważne są warunki:

a) L jest zanurzeniem izometrycznym;

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

c) $\mathbf{A} = \mathbf{C}^t\mathbf{B}\mathbf{C}$, gdzie $\mathbf{A}$ to macierz formy g w bazie $\mathcal{V}$, $\mathbf{B}$ to macierz formy h w bazie $\mathcal{W}$ i $\mathbf{C} = [L]_{\mathcal{W}}^{\mathcal{V}}$.

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

Zadanie 3. Wywnioskować też, że dwie symetryczne formy dwuliniowe na przestrzeni $\mathbb{R}^k$ są równoważne wtedy i tylko wtedy, gdy mają tę samą sygnaturę.

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?

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).

Niech V będzie przestrzenią z wyróżnioną formą metryczną g.

Definicja. 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**);

$V = A \bigoplus B$, gdy A i B są ortogonalnymi podprzestrzeniami liniowymi i $V = A \oplus B$.

1. a) Udowodnić, że $A^\perp$ jest podprzestrzenią liniową i $(A \cup B)^\perp = A^\perp \cap B^\perp$.

b) Udowodnić, że $(A^\perp)^\perp \supseteq A$, $(\mathrm{lin}(A))^\perp = A^\perp = \bigcap_{a \in A} a^\perp$ oraz gdy $\mathbf{0} \in A \cap B$, to $(A + B)^\perp = (A \cup B)^\perp$.

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

2. 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 $\mathbf{v}_1 \in V$. Przy $V_1 := \{\mathbf{v} \in V : \mathbf{v}_1 \perp \mathbf{v}\}$ zauważyć, że $\dim(V_1) \geq \dim(V) - 1$ i $\mathbf{v}_1 \notin V_1$. Stąd $V = V_1 \bigoplus \mathbb{F}\mathbf{v}_1$, co pozwala wykorzystać założenie indukcyjne.

3. 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.

Zadania ze zbioru Kostrykina: §II.2.2.5.

2. Podprzestrzenie nieosobliwe i rzut ortogonalny (zadania).

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 $\mathbf{u}, \mathbf{w} \in V_0$. Mówiąc o podprzestrzeni V_0 przestrzeni (V, g) pomijamy na ogół oznaczanie formy g_0 zakładając, że jest nią obcięcie formy g.

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

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

1. Dowieść, że gdy podprzestrzeń W przestrzeni V jest nieosobliwa, to
 a) $V = W \oplus W^{\perp}$.
 b) rzutowanie ortogonalne V na W istnieje i jest wyznaczone jednoznacznie.

2. Niech W będzie podprzestrzenią nieosobliwej przestrzeni V. Dowieść, że:
 a) Jeśli istnieje rzut ortogonalny V na W, to podprzestrzeń W jest nieosobliwa.
 b) Gdy podprzestrzeń W 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 $\operatorname{lin}(\mathbf{v}_1, ..., \mathbf{v}_p)$.

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 $\bigoplus\!\!\!\!\perp$;
 c) podprzestrzeń W jest nieosobliwa i $\dim W = \operatorname{rk}(g)$.

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 $V = U \bigoplus\!\!\!\!\perp W$, przy czym podprzestrzeń U jest izotropowa. 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 *.

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

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

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

Mówimy, że K^* jest **sprzężeniem** przekształcenia K (działającego pomiędzy przestrzeniami z funkcją dwuliniową). Dowód twierdzenia szkicujemy niżej w zadaniu 1.

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

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

zachodzi wtedy i tylko wtedy, gdy $\mathbf{K}^t\mathbf{H} = \mathbf{G}\mathbf{L}$.

b) Udowodnić twierdzenie 1.

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

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

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

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

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

Używane niżej pojęcia *bazy sprzężonej* 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 $J_g(\mathbf{v}) \in V^*$ 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. * Dla izomorfizmu $J : V \to V^*$, przeprowadzającego daną bazę $\mathcal{B}$ przestrzeni V na bazę dualną $\mathcal{B}^*$, jaka jest macierz powyższej funkcji g w bazie $\mathcal{B}$?

6. * Niech $L \in \mathcal{L}(V, W)$ i niech g i h będą nieosobliwymi formami dwuliniowymi na przestrzeniach V i W, odpowiednio.

a) Dowieść, że $J_h(\mathbf{w}) \circ L = J_g(L^*(\mathbf{w}))$ dla $\mathbf{w} \in W$.

b) Zdefiniujmy $L^\# : W^* \to \mathcal{V}^*$ wzorem $L^\#(\varphi) = \varphi \circ L$ dla $\varphi \in W^*$. Dowieść, że $L^\# \circ J_h = J_g \circ L^*$. (Ze względu na to, przekształcenia $L^\#$ i L^* są utożsamiane i oznaczane wspólnie L^*.)

4. Wokół twierdzenia o bezwładności (zadania).

1. Niech g będzie <u>symetryczną</u> formą dwuliniową na <u>rzeczywistej</u> 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_+ (odp. na V_-) dodatnio (odp. ujemnie) określona, 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. (Patrz §5.)

Zadania ze zbioru Kostrykina: §II.2.2.25.

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

Niech V będzie zespoloną przestrzenią wektorową. Funkcję $g : V \times V \to \mathbb{C}$ nazywamy **formą hermitowską** lub **półtoraliniową funkcją hermitowsko–symetryczną**, jeśli

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) formy hermitowskiej jest formą dwuliniową symetryczną (odp.: antysymetryczną).

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

Większość definicji i wyników dotyczących form dwuliniowych symetrycznych przenieść można na formy hermitowskie (a tym samym na iloczyny skalarne na przestrzeniach zespolonych).

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 $\mathbf{C}^*\mathbf{A}\mathbf{C}$ jest macierzą diagonalną.

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

b) Dowieść, że kryterium Jacobiego–Sylvestera pozostają prawdziwe dla macierzy samosprzężonych, gdy określoność rozumieć jak wzmiankowano wyżej. To 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: $\mathbb{R}^2$ z formą kwadratową $q = x^2 - t^2$. (Patrz przykład 1 z §4.3.)

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 postaci $\{\mathbf{v} : q(\mathbf{v}) = c\}$, 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 w płaszczyźnie Minkowskiego bazę 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 nazywamy miarą Minkowskiego kąta zorientowanego pomiędzy $\mathbf{u}$ i $\mathbf{v}$ i oznaczamy $\angle_m(\mathbf{u}, \mathbf{v})$. 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 (W, h) będą przestrzeniami z wyróżnionymi formami metrycznymi. O przekształceniu liniowym $L : V \to W$ powiemy, że **zachowuje ortogonalność**,

gdy $\mathbf{v}_1 \perp_g \mathbf{v}_2 \Rightarrow L(\mathbf{v}_1) \perp_h L(\mathbf{v}_2)$ dla $\mathbf{v}_1, \mathbf{v}_2 \in V$.

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

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

4. * Dowieść, że gdy w przestrzeni z 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

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 i przy 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 (§2.1). Twierdzenie 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).

5 (§2.2). Ortogonalna diagonalizacja form kwadratowych; wyrażenie sygnatury rzeczywistej macierzy symetrycznej przez jej wartości własne.

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

7. Funkcje dwuliniowe: 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 symetryczną lub antysymetryczną formę dwuliniową (ortogonalność, izometryczność, rzut ortogonalny, izotropowość).

1. Niech $\mathbf{A} = (g(\mathbf{v}_i, \mathbf{v}_j))_{i,j=1}^k$, gdzie g jest symetryczną formą dwuliniową na przestrzeni V, zaś $\mathbf{v}_1, ..., \mathbf{v}_k \in V$. Macierz nieosobliwą $\mathbf{C}$ taką, że macierz $\mathbf{C}^t\mathbf{AC}$ jest diagonalna, potraktujmy jako macierz przejścia od układu $\mathbf{v}_1, ..., \mathbf{v}_k$ do pewnego układu $\mathbf{w}_1, ..., \mathbf{w}_k$. (Oznacza to, że $\mathbf{w}_j = \sum_i c_{ij}\mathbf{v}_j$ dla $j = 1, ..., k$.) Dowieść, że:

a) Układ $\mathbf{w}_1, ..., \mathbf{w}_k$ jest g–ortogonalny i rozpina $\mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_k)$.

b) Gdy $\mathbf{C}$ otrzymano korzystając z algorytmu z §1.3, to $\mathbf{w}_1, ..., \mathbf{w}_k$ otrzymamy wykonując na układzie $\mathbf{v}_1, ..., \mathbf{v}_k$ kolejne operacje, wykonywane w algorytmie na wierszach macierzy $\mathbf{A}$.

c) Gdy ponadto nie wykonano części 1 żadnego z kroków algorytmu, to $\mathbf{w}_j \in \mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_j)$ dla $j = 1, ..., k$. (Ma to miejsce, gdy $\mathbb{F} = \mathbb{R}$ i forma g jest półokreślona, patrz zad. uz. 1 w §2.5.)

§ 6. ** Twierdzenia Witta i Clifforda o formach kwadratowych.

0. Wstęp. W tym paragrafie powaracamy do badania formy kwadratowej $q : V \to \mathbb{F}$. (Zakładamy nadal, że $\underline{1_{\mathbb{F}} + 1_{\mathbb{F}} \neq 0_{\mathbb{F}}}$.) Celem naszym jest udowodnienie poniższych dwóch ważnych twierdzeń:

Twierdzenie 1 (Witta o strukturze form kwadratowych). *Istnieje baza przestrzeni* V*, w której macierz formy* q *ma postać* $\mathrm{diag}(0, ..., 0, 1, -1, ..., 1, -1, c_1, ..., c_s)$*, gdzie* $\sum_{i=1}^{s} c_i v_i^2 \neq 0$ *dla* $(v_1, ..., v_s) \in \mathbb{F}^s \setminus \{\mathbf{0}\}$*. Powyżej, liczba zer jest równa* $\dim V - \mathrm{rk}(q)$*, liczba jedynek (równa liczbie wyrazów* -1*) jest przez formę* q *wyznaczona jednoznacznie, a macierz* $\mathrm{diag}(c_1, ..., c_s)$ *jest wyznaczona jednoznacznie z dokładnością do kongruentności.*

Dowód twierdzenia i inne jego sformułowanie podamy w p.1 i 2. Jest ono daleko idącym uogólnieniem, na przypadek dowolnego ciała o charakterystyce różnej od 2, twierdzenia Sylvestera o bezwładności. Istotnie, dla $\mathbb{F} = \mathbb{R}$ macierz formy jest w pewnej bazie postaci $\mathrm{diag}(0, ..., 0, 1, ..., 1, -1, ..., -1)$, więc pozostaje zmienić uporządkowanie bazy, by po zerach pojawiły się pary wyrazów 1 i -1 (liczba par wyniesie $\min(\sigma_+(q), \sigma_-(q))$, a potem same jedynki (gdy $\sigma_+(q) > \sigma_-(q)$) lub same wyrazy -1 (gdy $\sigma_+(q) < \sigma_-(q)$). Latwo dostrzec, że druga część tezy, dotycząca jednoznaczności, jest dla $\mathbb{F} = \mathbb{R}$ równoważna twierdzeniu o bezwładności.

Drugie twierdzenie, udowodnione w p.4, ma zupełnie inny charakter.

Twierdzenie 2 (Clifforda). *Przestrzeń* V *jest podprzestrzenią pewnej przestrzeni wektorowej* W_q *wymiaru* $2^{\dim(V)}$*, w której określone jest łączne mnożenie posiadające jedynkę* $\mathbf{e}$*, takie, że* $\mathbf{v}^2 = q(\mathbf{v})\mathbf{e}$ *dla każdego* $\mathbf{v} \in V$*.*

Ponadto, przestrzeń W_q *jest przez podprzestrzeń* V *generowana algebraicznie (wyjaśnimy to w p.3), i wymienione własności wyznaczają ją jednoznacznie z dokładnością do zachowującego mnożenie izomorfizmu przestrzeni liniowych.*

Twierdzenia te wyrażone są wyłącznie w terminach formy kwadratowej q. Charakterystyczne jest to, że oba są dowodzone w oparciu o własności wyznaczonej przez q formy dwuliniowej. Dowód jednoznaczności w twierdzeniu Witta wykorzystuje nawet piękne uogólnienia twierdzeń o własnościach izometrii przestrzeni euklidesowych, znanych nam z rozdziału IV. Te uogólnienia, mające niezależne znaczenie, przedstawimy w punktach 1 i 3; pochodzą one również od Witta.

O ile więc w §4 algebry użyto, by w dowolnej przestrzeni wektorowej wprowadzić namiastkę geometrii euklidesowej, o tyle teraz powstałe dzięki temu pojęcia geometryczne uczynią możliwym udowodnienie ważnych rezultatów algebraicznych.

Pewnym zaskoczeniem jest to, że struktura alternujących form dwuliniowych jest znacznie prostsza, niż symetrycznych. Wykazane to będzie w p.1.

1. Podprzestrzenie Artina i rozkład Witta.

W tym podpunkcie V oznacza przestrzeń z wyróżnioną formą metryczną g. (Dopuszczamy przypadek, gdy forma ta jest alternująca.)

Definicja. Układ izotropowych wektorów $\mathbf{u}_1, \mathbf{w}_1, ..., \mathbf{u}_n, \mathbf{w}_n$ nazywamy **artinowskim**, gdy dla $i, j = 1, \ldots, n$ zachodzi

$$\{\mathbf{u}_i, \mathbf{w}_i\} \perp \{\mathbf{u}_j, \mathbf{w}_j\} \text{ dla } i \neq j \text{ oraz } g(\mathbf{u}_i, \mathbf{w}_i) = 1 \tag{19}$$

Gdy $n = 1$ to mówimy też, że $\mathbf{u}_1, \mathbf{w}_1$ jest **parą Artina**. Podprzestrzeń nazwiemy artinowską, jeśli jest równa $\{\mathbf{0}\}$ lub pewna jej baza jest układem artinowskim.

Uwaga 1. W bazie artinowskiej, macierz wyróżnionej formy jest sumą zwnętrzną 2×2–klatek, mających na przekątnej zera, a poza nią parę $1, 1$ (gdy forma jest symetryczna) lub $1, -1$. W szczególności, podprzestrzeń artinowska jest nieosobliwa.

Zadanie 1. Gdy (V, g) jest przestrzenią Artina z symetryczną formą g, to:
a) w pewnej bazie, macierz formy g jest postaci $\mathrm{diag}(1, -1, \ldots, 1, -1)$;
b) w pewnej bazie, macierz formy g jest postaci $\begin{pmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_n & \mathbf{0} \end{pmatrix}$

Lemat 1. *Jeśli wektor izotropowy* $\mathbf{u}$ *nie jest osobliwy, to istnieje wektor* $\mathbf{w}$ *taki, że* $\mathbf{u}, \mathbf{w}$ *jest parą Artina.*

Dowód. Z założenia, istnieje wektor $\mathbf{v}$ taki, że $c := g(\mathbf{u}, \mathbf{v}) \neq 0$. Możemy założyć, że $c = 1$ (gdy $c \neq 1$, mnożymy $\mathbf{v}$ przez $1/c$). Pozostaje przyjąć $\mathbf{w} := \mathbf{v} + d\mathbf{u}$, gdzie $d = -\frac{1}{2} g(\mathbf{v}, \mathbf{v})$.$\square$

Twierdzenie 1 (strukturalne Witta)**.** *Istnieje rozkład ortogonalny* $V = V_{nul} \oplus\!\!\!\!\!\perp V_{art} \oplus\!\!\!\!\!\perp V_{def}$, *taki, że:*
i) podprzestrzeń V_{nul} *jest izotropowa;*
ii) podprzestrzeń V_{art} *jest artinowska;*
iii) podprzestrzeń V_{def} *jest* **określona**, *tzn. nie ma wektorów izotropowych.*
W tym tzw. **rozkładzie Witta** *podprzestrzeń* V_{nul} *jest wyznaczona jednoznacznie i równa przestrzeni* $V^{\perp}$ *wektorów osobliwych, a podprzestrzenie* V_{art} *i* V_{def} *są wyznaczone jednoznacznie z dokładnością do izometrii.*

Dowód. By szukany rozkład wskazać, obieramy za V_{art} dowolną podprzestrzeń artinowską, która jest maksymalna (tzn. nie jest zawarta w żadnej innej takiej podprzestrzeni). Ponieważ jest ona nieosobliwa, więc $V_{art} \cap V^{\perp} = \{\mathbf{0}\}$ i istnieje podprzestrzeń W taka, że $V_{art} \subset W$ oraz $V^{\perp} \oplus W = V$. Oczywiście, $W \perp V^{\perp}$ i podprzestrzeń W jest nieosobliwa. (Patrz zadanie 6 w §4.2).

Oznaczmy przez V_{def} dopełnienie ortogonalne (nieosobliwej) podprzestrzeni V_{art} w przestrzeni nieosobliwej W. Zauważmy, że żaden wektor podprzestrzeni V_{def} nie jest w niej osobliwy, bo będąc ortogonalnym do $V^\perp$ i do V_{art}, byłby ortogonalny do V – co nie ma miejsca, bo $V_{def} \cap V^\perp = \{\mathbf{0}\}$.

Jeśliby więc pewien wektor $\mathbf{u} \in V_{def}$ był izotropowy, to na podstawie lematu 1 istniałaby płaszczyzna Artina $A \subset V_{def}$. Jest to niemożliwe, bo przestrzeń $V_{art} \bigoplus A$ byłaby artinowska, wbrew maksymalności V_{art}. Przestrzeń V_{def} jest więc określona i $V^\perp \bigoplus V_{art} \bigoplus V_{def}$ jest szukanym rozkładem.

Jednoznaczność V_{nul} wynika z zadania 6 z §4.2, bo podprzestrzeń $V_{art} \bigoplus V_{def}$ jest nieosobliwa. Gdy wyróżniona forma g jest alternująca, to $V_{def} = \{\mathbf{0}\}$, skąd V_{art} jest przestrzenią Artina wymiaru $\mathrm{rk}(g)$ – co wyznacza ją jednoznacznie z dokładnością do izometrii. Zaś w najciekawszym przypadku form symetrycznych, jednoznaczności V_{def} i V_{art} dowiedziemy w następnym punkcie. (Wcześniej skorzystamy z niej tylko w zadaniach i uwadze 2.)□

Zadanie 2. W oparciu o twierdzenie 1 i zadanie 1a), uzyskać twierdzenie 1 wstępu.

Uwaga 2. a) Z dowodu wynika, że za V_{art} można w twierdzeniu przyjąć dowolną maksymalną podprzestrzeń artinowską.

b) Przeciwnie, gdy warunki tezy twierdzenia są spełnione, to podprzestrzeń V_{art} jest maksymalna (wśród artinowskich), bo maksymalna podprzestrzeń Artina $V'_{art} \supset V_{art}$ też jest składnikiem pewnego rozkładu Witta, skąd $V'_{art} = V_{art}$ na podstawie izomorficzności V_{art} i V'_{art}. (Korzystamy z nieudowodnionej jeszcze części twierdzenia.)

c) W szczególności, dwie maksymalne podprzestrzenie Artina są równego wymiaru.□

Dla form antysymetrycznych, z twierdzenia wynika ważny

Wniosek 1. *Antysymetryczna forma dwuliniowa g na przestrzeni V jest w pewnej bazie zadana wielomianem $\sum_{i=1}^n (x_{2i-1}y_{2i} - x_{2i}y_{2i-1})$. (Oczywiście, wtedy $\mathrm{rk}(g) = 2n \leq \dim V$.)*

<u>Dowód.</u> Ponieważ $V_{def} = \{\mathbf{0}\}$, więc wystarcza szukaną bazę „skleić” z bazy Artina przestrzeni V_{art} i bazy przestrzeni $V^\perp$. □

<u>Definicja.</u> Bazę, której istnienie stwierdza wniosek 1, nazywamy **symplektyczną** (względem g), a wielomian u nazywamy **postacią kanoniczną** formy g.

Wniosek 2. *Gdy dwuliniowe formy alternujące $g : V \times V \to \mathbb{F}$ i $g' : V' \times V' \to \mathbb{F}$ mają ten sam rząd i $\dim V = \dim V'$, to przestrzenie (V, g) i (V', g') są izometryczne.*

<u>Dowód.</u> W bazach symplektycznych, macierze obu form są równe.□

Na koniec udowodnimy lemat wykorzystywany w p.2 i poniższych zadaniach.

Lemat 2. *Niech przestrzeń (V, g) będzie nieosobliwa, a liniowo niezależne wektory $\mathbf{u}_1, ..., \mathbf{u}_n \in V$ będą izotropowe i wzajemnie ortogonalne. Wówczas istnieją wektory $\mathbf{w}_1, ..., \mathbf{w}_n \in V$ takie, że układ $\mathbf{u}_1, \mathbf{w}_1, ..., \mathbf{u}_n, \mathbf{w}_n$ jest artinowski.*

Dowód. Zastosujemy indukcję względem n. Niech $n \geq 1$ i niech $W_0 := \{\mathbf{u}_1, ..., \mathbf{u}_n\}^\perp$ i $W := \{\mathbf{u}_1, ..., \mathbf{u}_{n-1}\}^\perp$ (dla $n = 1$ przyjmujemy $W := V$). Ponieważ $\dim W > \dim W_0$ (patrz zadanie .... 3 w §4.3), więc istnieje wektor $\mathbf{w} \in W \setminus W_0$. Wektor $\mathbf{u}_n$ jest nieosobliwy w płaszczyźnie $A := \mathrm{lin}(\mathbf{u}_n, \mathbf{w})$, bo $g(\mathbf{u}_n, \mathbf{w}) \neq 0$, i z lematu 1 wynika istnienie wektora $\mathbf{w}_n \in A$ takiego, że $(\mathbf{u}_n, \mathbf{w}_n)$ jest parą Artina. Pozostaje zastosować założenie indukcyjne do nieosobliwej przestrzeni $A^\perp$ i wektorów $\mathbf{u}_1, ..., \mathbf{u}_{n-1} \in A^\perp$, by otrzymać pozostałe szukane wektory $\mathbf{w}_1, ..., \mathbf{w}_{n-1}$.$\square$

Zadania uzupełniające.

1. Z wniosku 1 wynika, że dowolna macierz antysymetryczna $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ jest nad $\mathbb{F}$ kongruentna z macierzą $\mathbf{B}$, będącą sumą zewnętrzną klatki zerowej i 2×2–klatek o wierszach (0,1) i (-1,0). Podać modyfikację algorytmu z § 1.3, umożliwiającą otrzymanie macierzy nieosobliwej $\mathbf{C}$, dla której $\mathbf{C}^t\mathbf{AC} = \mathbf{B}$.

2. Dowieść, że wyznacznik macierzy antysymetrycznej o wyrazach całkowitych jest kwadratem liczby całkowitej. (Wskazówka: macierz traktować jako element $\mathcal{M}_k(\mathbb{Q})$.)

3. Niech w ciele skalarów $\mathbb{F}$ przestrzeni V każdy element posiada pierwiastek kwadratowy. Dowieść, że $\dim(V_{def}) \leq 1$; w szczególności, gdy przestrzeń V jest nieosobliwa i $\dim(V)$ jest liczbą parzystą, to V jest przestrzenią Artina.

4. Dla przestrzeni nieosobliwej V dowieść równoważności warunków:
   a) V zawiera podprzestrzeń izotropową wymiaru $\geq \frac{1}{2}\dim V$.
   b) V jest przestrzenią Artina.

5. a) Dowieść, że gdy symetryczna forma dwuliniowa g jest nieosobliwa, to $W := (V, g) \bigoplus\!\!\!\!\!\perp (V, -g)$ jest przestrzenią Artina. Powiązać to z zadaniem 1 b) w §1.3.
   b) Dowieść, że każda (nawet i osobliwa) przestrzeń wymiaru k jest izometryczna z podprzestrzenią $2k$–wymiarowej przestrzeni Artina.

6. Dowieść, że gdy U jest maksymalną podprzestrzenią izotropową przestrzeni V, to:
   a) $U = V^\perp \oplus W$ dla pewnej podprzestrzeni W, zawartej w podprzestrzeni nieosobliwej.
   b) $\dim U = \dim V - \mathrm{rk}(g) + ind$, gdzie ind oznacza **indeks przestrzeni** V, czyli połowę wymiaru jej (dowolnej) maksymalnej podprzestrzeni artinowskiej.

2. Izometrie i symetrie zwierciadlane przestrzeni z symetryczną formą metryczną. Jednoznaczność rozkładu Witta.

Aż do końca rozdziału zakładamy, że V jest przestrzenią z wyróżnioną symetryczną formą dwuliniową g. Przez q oznaczamy wyznaczoną przez g formę kwadratową.

Przypomnijmy, że gdy V_0 jest podprzestrzenią nieosobliwą, to $V = V_0 \bigoplus V_0^\perp$. (Patrz §4.2.) Rzutowanie P przestrzeni V na V_0 wzdłuż $V_0^\perp$ nazywamy ortogonalnym, bo $\mathbf{v} - P(\mathbf{v}) \perp V_0$ dla każdego $\mathbf{v} \in V$. Przekształcenie $I_V - 2P$ nazywamy **symetrią ortogonalną** względem V_0, a gdy $\dim V_0 = \dim V - 1$ mówimy o **symetrii zwierciadlanej** lub **lustrzanej**. Oczywiście, $S(\mathbf{v}_0 + \mathbf{v}_1) = \mathbf{v}_0 - \mathbf{v}_1$ dla $\mathbf{v}_0 \in V_0$ i $\mathbf{v}_1 \in V_0^\perp$.

Definicje te są w pełni analogiczne do podanych w §IV, lecz mają sens tylko dla nieosobliwej podprzestrzeni V_0 (bo rzutowanie V na V_0 zdłuż $V_0^\perp$ nie istnieje gdy $V_0 \cap V_0^\perp \neq \{\mathbf{0}\}$).

Lemat 1. *Gdy $\mathbf{u}, \mathbf{w} \in V$ i $q(\mathbf{u}) = q(\mathbf{w}) \neq 0$, to wektor $\mathbf{v}_1 := \mathbf{u}+\mathbf{w}$ jest nieizotropowy i symetria lustrzana względem $\mathbf{v}_1^\perp$ przeprowadza $\mathbf{u}$ na $-\mathbf{w}$, lub też nieizotropowy jest wektor $\mathbf{v}_2 := \mathbf{u} - \mathbf{w}$ i symetria lustrzana względem $\mathbf{v}_2^\perp$ przeprowadza $\mathbf{u}$ na $\mathbf{w}$.*

Dowód. Wektory $\mathbf{v}_1$ i $\mathbf{v}_2$ są ortogonalne, bo $g(\mathbf{v}_1, \mathbf{v}_2) = g(\mathbf{u}, \mathbf{u}) + g(\mathbf{u}, \mathbf{w}) - g(\mathbf{w}, \mathbf{u}) - g(\mathbf{w}, \mathbf{w}) = 0$. Ze wzoru Pitagorasa $q(\mathbf{v}_1) + q(\mathbf{v}_2) = q(\mathbf{v}_1 + \mathbf{v}_2) = 4q(\mathbf{u}) \neq 0$ wynika więc, że pewien z wektorów $\mathbf{v}_1, \mathbf{v}_2$ jest nieizotropowy. Gdy jest nim $\mathbf{v}_1$, symetria lustrzana S_1 względem $\mathbf{v}_1^\perp$ jest zdefiniowana i $S_1(2\mathbf{u}) = S_1(\mathbf{v}_1 + \mathbf{v}_2) = -\mathbf{v}_1 + \mathbf{v}_2 = -2\mathbf{w}$, wobec czgo $S_1(\mathbf{u}) = -\mathbf{w}$. Gdy nieizotropowy jest wektor $\mathbf{v}_2$, to zdefiniowana jest symetria lustrzana S_2 względem $\mathbf{v}_2^\perp$ i $S_2(2\mathbf{u}) = \mathbf{v}_1 - \mathbf{v}_2 = 2\mathbf{w}$. □

Wniosek 1. *Przy założeniach lematu, wektor $\mathbf{u}$ jest przeprowadzany na $\mathbf{w}$ przez pewną izometrię, będącą symetrią lustrzaną lub złożeniem dwóch takich symetrii.*

Dowód. Gdy wektor $\mathbf{u} + \mathbf{w}$ jest nieizotropowy, to symetrię S_1 pozostaje złożyć z symetrią względem $\mathbf{w}^\perp$, przeprowadzającą $-\mathbf{w}$ na $\mathbf{w}$. □

Twierdzenie 1 (Witta o przedłużaniu)**.** *Niech $L_0 : V_0 \to V$ będzie zanurzeniem izometrycznym podprzestrzeni V_0 przestrzeni V. Gdy V lub V_0 jest przestrzenią nieosobliwą, to L_0 można przedłużyć do izometrii $L : V \to V$.*

Dowód. 1) Rozważymy wpierw przypadek, gdy podprzestrzeń V_0 jest nieosobliwa, stosując indukcję względem $n = \dim(V_0)$. (Gdy $n = 0$, teza jest oczywista.)

Jeśli istnieje nieizotropowy wektor $\mathbf{u} \in V_0$ taki, że $L_0(\mathbf{u}) = \mathbf{u}$, to rozpatrzmy obcięcie L_0 do podprzestrzeni $V_0' := \mathbf{u}^\perp \cap V_0$. Wyznacza ono zanurzenie izometryczne $V_0' \to \mathbf{u}^\perp$, które z założenia indukcyjnego można przedłużyć do izometrii $L' : \mathbf{u}^\perp \to \mathbf{u}^\perp$. (Wykorzystaliśmy nieosobliwość podprzestrzeni V_0', będącej dopełnieniem ortogonalnym, w nieosobliwej przestrzeni V_0, wektora nieizotropowego $\mathbf{u}$.) Niech $L := I \oplus L'$, gdzie I oznacza identyczność na prostej $\mathbb{F}\mathbf{u}$. Ponieważ $V = \mathbb{F}\mathbf{u} \bigoplus \mathbf{u}^\perp$, więc L jest izometrią; ponadto równość $L(\mathbf{v}) = L_0(\mathbf{v})$ zachodzi dla wszystkich $\mathbf{v} \in V_0$, gdyż zachodzi dla $\mathbf{v} \in \mathbb{F}\mathbf{u}$ i dla $\mathbf{v} \in \mathbf{u}^\perp \cap V_0$.

W ogólnym przypadku obieramy dowolny wektor nieizotropowy $\mathbf{u} \in V_0$ (istnieje na mocy uwagi 2 w §3.3, odniesionej do V_0), i korzystając z wniosku 1 znajdujemy izometrię $K : V \to V$ taką, że $K(L_0(\mathbf{u})) = \mathbf{u}$. Jak już wiemy, zanurzenie $KL_0 : V_0 \to V$ można przedłużyć do izometrii $M : V \to V$. Pozostaje przyjąć $L = K^{-1}M$.

2) Niech teraz przestrzeń V będzie nieosobliwa. (Przypadek ten nie jest dalej wykorzystany.)

Załóżmy wpierw dodatkowo, że V_0 jest przestrzenią izotropową. Wybierzmy jej bazę $\mathbf{u}_1, ..., \mathbf{u}_n$, przyjmijmy $\mathbf{u}'_i := L_0(\mathbf{u}_i)$ dla $i = 1, ..., n$ i utwórzmy artinowskie układy $\mathbf{u}_1, \mathbf{w}_1, ..., \mathbf{u}_n, \mathbf{w}_n$ i $\mathbf{u}'_1, \mathbf{w}'_1, ..., \mathbf{u}'_n, \mathbf{w}'_n$ i ich powłoki liniowe A i A'. Przekształcenie liniowe $L_1 : A \to A'$ takie, że $L_1(\mathbf{u}_i) = \mathbf{u}'_i$ oraz $L_1(\mathbf{w}_i) = \mathbf{w}'_i$ dla $i = 1, ..., n$, jest izometrią. (Patrz zadanie 1 w §3.3.). Z dowiedzionej części twierdzenia wynika, że L_1 można przedłużyć do izometrii $L : V \to V$. Ponieważ $L_1|_{V_0} = L_0$, to cel nasz został osiągnięty.

W ogólnym przypadku powtórzymy rozumowanie z 1), indukcyjne względem $\dim(V_0)$. Obierzmy wektor nieizotropowy $\mathbf{v} \in V_0$ (gdy takiego nie ma, tezy dowiedliśmy wyżej) i oznaczmy przez K izometrię $V \to V$ taką, że $K(\mathbf{v}) = L_0(\mathbf{v})$. Przyjmijmy $V'_0 := K(\mathbf{v}^\perp \cap V_0)$ i $\mathbf{w} := L_0(\mathbf{v}_0)$; oczywiście $V'_0 \subset \mathbf{w}^\perp$ i $L_0(V'_0) \subset \mathbf{w}^\perp$. Zatem $L_0K^{-1}|_{V'_0}$ jest zanurzeniem izometrycznym V'_0 w $\mathbf{w}^\perp$, które z założenia indukcyjnego można przedłużyć do izometrii $L' : \mathbf{w}^\perp \to \mathbf{w}^\perp$. Przyjmujemy $L := (I \oplus L')K$, gdzie I oznacza identyczność na $\mathbb{F}\mathbf{w}$. □

Wniosek 2 (Witta o dopełnieniu ortogonalnym)**.** *Niech przestrzenie W i W' będą izometryczne, a W_0 i W'_0 niech będą ich izometrycznymi podprzestrzeniami. Jeśli W lub W_0 jest przestrzenią nieosobliwą, to podprzestrzenie $W_0^\perp$ i $W_0'^\perp$ są izometryczne.*

Dowód. Niech $K_0 : W_0 \to W'_0$ i $K : W \to W'$ będą izometriami. Z twierdzenia wynika istnienie izometrii $L : W \to W$ takiej, że $L(\mathbf{w}) = K^{-1}L_0(\mathbf{w})$ dla $\mathbf{v} \in W_0$. Oczywiście, $KL : W \to W'$ jest izometrią przeprowadzającą W_0 na W'_0, a więc i $W_0^\perp$ na $W_0'^\perp$.□

Dowód jednoznaczności rozkładu $V = V_{nul} \bigoplus V_{art} \bigoplus V_{def}$ z p.1. Jak już zauważyliśmy, $V_{nul} = V^\perp$. Gdy oprócz rozważanego mamy inny rozkład Witta $V = V^\perp \bigoplus V'_{art} \bigoplus V'_{def}$, to przyjmijmy $W = V_{art} \bigoplus V_{def}$ i $W' = V'_{art} \bigoplus V'_{def}$. Ponieważ $V = V^\perp \bigoplus W = V^\perp \bigoplus W'$, wiec rzut $V \to W'$ wzdłuż $V^\perp$ wyznacza izometrię $W \to W'$. Niech $\dim V_{art} \leq \dim V'_{art}$. Wówczas $V'_{art} = A_0 \bigoplus A_1$, gdzie A_0 i A_1 są podprzestrzeniami Artina, przy czym $\dim A_0 = \dim V_{art}$. Na podstawie wniosku 2, zastosowanego do par (W, V_{art}) i (W', A_0), podprzestrzeń $A_1 \bigoplus V'_{def}$ jest izometryczna z podprzestrzenią V_{def}. Dowodzi to, że $A_1 = \{\mathbf{0}\}$, bo inaczej w A_1 istniałby wektor izotropowy, a w V_{def} takiego nie ma. Tym samym $A_1 \bigoplus V'_{def} = V'_{def}$ i $A_0 = V'_{art}$, tzn. podprzestrzenie V_{def} i V_{art} są izometryczne z V'_{def} i V'_{art}, odpowiednio. □

Oto szczególny przypadek wniosku 2, sformułowany w postaci pozornie nie mającej nic wspólnego z izometriami. (Można się nawet łudzić, że to łatwe zadanie dotyczące kongruentności macierzy.)

Wniosek 3 (Witta o skracaniu)**.** *Niech symetryczne macierze* $\operatorname{diag}(\mathbf{A}, \mathbf{B})$ *i* $\operatorname{diag}(\mathbf{A}', \mathbf{B}')$ *będą kongruentne. Jeśli macierze* $\mathbf{A}$ *i* $\mathbf{A}'$ *są nieosobliwe i kongruentne, to kongruentne są macierze* $\mathbf{B}$ *i* $\mathbf{B}'$*.*

Dowód. Oznaczmy stopień macierzy $\operatorname{diag}(\mathbf{A}, \mathbf{B})$ przez k, a przestrzeń $\mathbb{F}^k$ przez V lub przez V', zależnie od tego, czy wyróżniamy na niej formę g o macierzy $\operatorname{diag}(\mathbf{A}, \mathbf{B})$, czy formę g' o macierzy $\operatorname{diag}(\mathbf{A}', \mathbf{B}')$. Wówczas $V = U \bigoplus W$, gdzie w pewnej bazie przestrzeni U macierz formy $g|_{U \times U}$ jest równa $\mathbf{A}$, a w pewnej bazie przestrzeni W macierz formy $g|_{W \times W}$ jest równa $\mathbf{B}$. Tak samo, $V' = U' \bigoplus W'$, gdzie własności U' i W' są „primowane" w stosunku do powyższych (należy dopisać „primy" przy $\mathbf{A}, \mathbf{B}$ i g).

Przestrzenie V i V' są izometryczne, bo macierze form g i g' są kongruentne; podobnie, izometryczne są przestrzenie U i U'. (Patrz zadanie 1 w §3.3.) Z wniosku 2 wynika więc izometryczność przestrzeni W i W', a zatem kongruentność $\mathbf{B}$ i $\mathbf{B}'$.$\square$

Problem 1. Gdy nieosobliwe macierze symetryczne $\operatorname{diag}(\mathbf{A}, \mathbf{A})$ i $\operatorname{diag}(\mathbf{B}, \mathbf{B})$ są kongruentne, czy kongruentne są macierze $\mathbf{A}$ i $\mathbf{B}$?

Zadania uzupełniające. W 1 i 2 zakładamy nieosobliwość przestrzeni (V, q).

1. Udowodnić następujące twierdzenie E. Cartana: każda izometria przestrzeni (V, q) jest złożeniem l symetrii lustrzanych, dla pewnego $l \leq 2 \dim V$. (Można też, co trudniejsze, zastąpić $2 \dim V$ przez $\dim V$; patrz [Artin] czy [Berger].)

2. „Półpełnym obrotem" przestrzeni V nazywamy przekształcenie $-I_U \oplus I_{U^\perp}$, gdzie U jest podprzestrzenią nieosobliwą wymiaru 2. Dowieść, że każda izometria przestrzeni (V, q), mająca wyznacznik 1, jest złożeniem skończenie wielu półpełnych obrotów.

3. Dowieść, że gdy zanurzenie izometryczne $L_0 : V_0 \to V$ spełnia warunek $L_0^{-1}(V^\perp) = V^\perp \cap V_0$, to można je przedłużyć do izometrii $V \to V$. (Nie zakładamy nieosobliwości podprzestrzeni V_0 lub przestrzeni V.)

Zadania ze zbioru Kostrykina: 21, 22, 24 *, 25 i 27 w §II.2.1. (Podprzestrzeń „osobliwa" w [Ko] oznacza tyle, co „izotropowa".)

3. Algebra Clifforda formy kwadratowej.

Definicja. **Algebrą** nazywamy przestrzeń wektorową W z wyróżnioną operacją mnożenia $*$, taką, że:

1) $\mathbf{u} * \mathbf{v} \in W$ i $(c\mathbf{u}) * \mathbf{v} = \mathbf{u} * (c\mathbf{v}) = c(\mathbf{u} * \mathbf{v})$ dla $\mathbf{u}, \mathbf{v} \in W$ i $c \in \mathbb{F}$;

2) mnożenie jest obustronnie rozdzielne względem dodawania:

$(\mathbf{u} + \mathbf{v}) * \mathbf{w} = \mathbf{u} * \mathbf{w} + \mathbf{v} * \mathbf{w}$ i $\mathbf{w} * (\mathbf{u} + \mathbf{v}) = \mathbf{w} * \mathbf{u} + \mathbf{w} * \mathbf{v}$ dla $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W$.

O wszystkich rozważanych tu algebrach zakładamy, że są łączne, tzn. że

3) $(\mathbf{u} * \mathbf{v}) * \mathbf{w} = \mathbf{u} * (\mathbf{v} * \mathbf{w})$ dla $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W$.

Na ogół, iloczyny $\mathbf{u} * \mathbf{v}$ i $\mathbf{v} * \mathbf{v}$ oznaczane będą przez $\mathbf{uv}$ i $\mathbf{v}^2$, odpowiednio.

Przykład 1. Z wymienionym działaniem mnożenia, algebrą jest każda z poniższych przestrzeni wektorowych:

a) $\mathcal{M}_k(\mathbb{F})$, z naturalnym działaniem mnożenia macierzy;

b) przestrzeń wszystkich operatorów liniowych na przestrzeni wektorowej V, z mnożeniem będącym składaniem operatorów;

c) przestrzeń $\mathbb{R}^X$ wszystkich funkcji rzeczywistych na zadanym zbiorze X, z naturalnym mnożeniem funkcji.

Definicja. a) Element $\mathbf{e}$ algebry W nazywamy jej **jedynką**, gdy $\mathbf{ew} = \mathbf{w} = \mathbf{we}$ dla $\mathbf{w} \in W$. (Nie każda algebra ma jedynkę, lecz tu rozpatrujemy tylko algebry z jedynką.)

b) Zbiór $S \subset W$ **generuje algebraicznie** algebrę W, gdy ta jest najmniejszą swą podprzestrzenią liniową, zawierającą S i zamkniętą ze względu na mnożenie (tzn. gdy każdy element algebry jest kombinacją liniową skończonych iloczynów elementów zbioru S).

Badać będziemy przestrzeń wektorową V z wyróżnioną formą kwadratową q. Interesuje nas to, jak konstruować przekształcenia $L : V \to W$, spełniające następujące **warunki Clifforda**:

(C1) W jest algebrą z jedynką $\mathbf{e}$ i $(L(\mathbf{v}))^2 = q(\mathbf{v})\mathbf{e}$ dla wszystkich $\mathbf{v} \in V$. (Po lewej jest kwadrat elementu $L(\mathbf{v})$ algebry, a po prawej $q(\mathbf{v})$–krotność wektora $\mathbf{e}$.)

(C2) L jest monomorfizmem liniowym, którego obraz $L(V)$ generuje W algebraicznie.

Lemat 1. *Gdy $L : V \to W$ jest przekształceniem liniowym w algebrę z jedynką $\mathbf{e}$, to warunek* (C1) *jest równoważny każdemu z następujących:*

(C1′) $L(\mathbf{u})L(\mathbf{v}) + L(\mathbf{v})L(\mathbf{u}) = 2g(\mathbf{u}, \mathbf{v})\mathbf{e}$ *dla* $\mathbf{u}, \mathbf{v} \in V$, *gdzie* $g : V \times V \to \mathbb{F}$ *jest wyznaczoną przez q formą dwuliniową;*

(C1″) *istnieje baza ortogonalna $\mathbf{v}_1, ..., \mathbf{v}_n$ przestrzeni V taka, że przy $\mathbf{e}_i := L(\mathbf{v}_i)$ mają miejsce równości $\mathbf{e}_i\mathbf{e}_j = -\mathbf{e}_j\mathbf{e}_i$ oraz $\mathbf{e}_i^2 = q(\mathbf{v}_i)\mathbf{e}$ dla $i, j = 1, ..., n,\ i \neq j$.*

Dowód. Ponieważ $2g(\mathbf{u}, \mathbf{v}) = q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})$ oraz $(L(\mathbf{u}) + L(\mathbf{v}))^2 - L(\mathbf{u})^2 -$

$L(\mathbf{v})^2 = L(\mathbf{u})L(\mathbf{v}) + L(\mathbf{v})L(\mathbf{u})$, więc z (C1) wynika (C1′). Oczywiście, z (C1′) wynika (C1″) (obieramy dowolną bazę ortogonalną), zaś jeśli $\mathbf{v}_1, ..., \mathbf{v}_n$ jest bazą zaświadczającą o prawdziwości (C1″), to dla $\mathbf{v} = \sum_{i=1}^n \lambda_i \mathbf{v}_i$ zachodzi

$(L(\mathbf{v}))^2 = \sum_i \lambda_i^2 \mathbf{e}_i^2 + \sum_{i<j}(\lambda_i\lambda_j \mathbf{e}_i\mathbf{e}_j + \lambda_i\lambda_j \mathbf{e}_j\mathbf{e}_i) = (\sum_i \lambda_i^2 q(\mathbf{v}_i))\mathbf{e} = q(\mathbf{v})\mathbf{e}$.

(Pierwsza równość wynika z liniowości L, a ostatnia z równości Pitagorasa.)□

Definicja. Układ $(\mathbf{e}_1, ..., \mathbf{e}_n)$ elementów algebry W z jedynką $\mathbf{e}$ nazywamy **antykomutującym typu** $(\lambda_1, ..., \lambda_n)$, gdy $\mathbf{e}_i\mathbf{e}_j = -\mathbf{e}_j\mathbf{e}_i$ oraz $\mathbf{e}_i^2 = \lambda_i \mathbf{e}$ dla $i, j = 1, ..., n,\ i \neq j$.

Przykład 2. a) W ciele $\mathbb{C}$ liczb zespolonych, traktowanym jako algebra nad $\mathbb{R}$, (i) jest układem ortogonalnym typu (-1), rozpinającym $\mathbb{C}$ algebraicznie.

b) Niech nieprzemienne ciało kwaternionów $\mathbb{H}$ traktowane będzie jako algebra nad $\mathbb{R}$. Układ (i, j, k) jest ortogonalny typu $(-1, -1, -1)$ i każde jego dwa elementy rozpinają $\mathbb{H}$ algebraicznie.

c) W algebrze $\mathcal{M}_2(\mathbb{F})$, układ $\begin{pmatrix} 0 & c \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & -c \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ jest ortogonalny typu $(c, -c, 1)$, i dla $c \neq 0$ każde dwa jego elementy rozpinają $\mathcal{M}_2(\mathbb{F})$ algebraicznie.

d) Układ $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ rozpina algebraiczne przestrzeń $\mathcal{M}_2(\mathbb{C})$, traktowaną jako algebra nad $\mathbb{R}$, i jest ortogonalny typu $(1, 1, 1)$.⊟

Wniosek 1. *W każdym z poniższych przypadków istnieje monomorfizm $L : V \to W$, spełniający warunki Clifforda:*

a) $(V, q) = (\mathbb{R}, -x^2)$ *i* $W = \mathbb{C}$;

b) $(V, q) = (\mathbb{R}^n, -x_1^1 - ... - x_n^2)$, *dla* $n = 2, 3$, *i* $W = \mathbb{H}$ *(jako algebra nad* $\mathbb{R}$*)*;

c1) $W = \mathcal{M}_2(\mathbb{F})$, $V = \mathbb{F}^2$ *i* $q = x_1^2 + x_2^2$;

c2) $W = \mathcal{M}_2(\mathbb{F})$, $V = \mathbb{F}^2$ *i* $q = x_1^2 - x_2^2$;

c3) $W = \mathcal{M}_2(\mathbb{F})$, $V = \mathbb{F}^3$ *i* $q = x_1^2 + x_2^2 - x_3^2$.

d) $(V, q) = (\mathbb{R}^3, x_1^2 + x_2^2 + x_3^2)$ *i* $W = \mathcal{M}_2(\mathbb{C})$ *(jako algebra nad* $\mathbb{R}$*)*.

Dowód. Za każdym razem definiujemy $L \in \mathcal{L}(V, W)$ tak, by standardowa baza $(\mathbf{v}_i)$ przestrzeni V przeprowadzana była na antykomutujący układ odpowiedniego typu w W. Układy z przykładu 1 są liniowo niezależne, więc L jest monomorfizmem. □

By wyjść poza rozpatrywanie izolowanych (choć ważnych) przykładów, potrzebne są dwa stwierdzenia. Pierwsze uogólnia część c) przykładu, a drugie dotyczy iloczynów algebraicznie ortogonalnych elementów algebry.

Stwierdzenie 1. *W algebrze macierzy $\mathcal{M}_{2k}(\mathbb{F})$ rozpatrzmy następujące macierze:*

$$\mathbf{A}'_i = \operatorname{diag}(\mathbf{A}_i, -\mathbf{A}_i) \ \textit{dla}\ i \leq j, \quad \mathbf{A}'_{j+1} = \begin{pmatrix} \mathbf{0} & c\mathbf{I}_k \\ \mathbf{I}_k & \mathbf{0} \end{pmatrix}, \quad \mathbf{A}'_{j+2} = \begin{pmatrix} \mathbf{0} & -c\mathbf{I}_k \\ \mathbf{I}_k & \mathbf{0} \end{pmatrix}$$

Gdy $c \neq 0$ i macierze $\mathbf{A}_1, ..., \mathbf{A}_j \in \mathcal{M}_k(\mathbb{F})$ rozpinają algebrę $\mathcal{M}_k$ algebraicznie, to macierze $\mathbf{A}'_1, ..., \mathbf{A}'_{j+2}$ algebraicznie rozpinają $\mathcal{M}_{2k}$; gdy zaś macierze $\mathbf{A}_1, ..., \mathbf{A}_j$ tworzą antykomutujący układ typu $(c_1, ..., c_j)$, to macierze $\mathbf{A}'_1, ..., \mathbf{A}'_{j+2}$ tworzą taki układ typu $(c_1, ..., c_j, c, -c)$.

Dowód. Ostatnią własność sprawdzić jest bardzo łatwo; por. zadanie 1 w §II.2.4.

Oznaczmy przez W podprzestrzeń liniową algebry $\mathcal{M}_{2k}$, generowaną przez macierze $\mathbf{A}'_1, ..., \mathbf{A}'_{j+2}$ algebraicznie. Gdy macierze $\mathbf{A}_1, ..., \mathbf{A}_j$ rozpinają $\mathcal{M}_k$, to dla każdej klatki $\mathbf{A} \in \mathcal{M}_k$ istnieje klatka $\mathbf{B} \in \mathcal{M}_k$ taka, że $\operatorname{diag}(\mathbf{A}, \mathbf{B}) \in W$. Mnożąc $\operatorname{diag}(\mathbf{A}, \mathbf{B})$ przez jedną z macierzy $\mathbf{A}'_{j+1} \pm \mathbf{A}'_{j+2}$ lub przez ich iloczyn stwierdzamy, że do W należy każda macierz, której 3 z 4 klatek są równe $\mathbf{0}_k$, a pozostała jest zadana. Zatem $W = \mathcal{M}_{2k}$ (bo każda macierz z $\mathcal{M}_{2k}$ jest sumą czterech tej postaci).□

Oznaczenie. Dla $\mathbf{w}_1, ..., \mathbf{w}_n \in W$ i $\alpha \subset \{1, ..., n\}$ przyjmijmy $\mathbf{w}_\alpha = \mathbf{e}$ gdy $\alpha = \emptyset$ oraz, gdy (wszystkimi) elementami zbioru α są $\alpha(1) < \cdots < \alpha(s)$, to

$$\mathbf{w}_\alpha := \mathbf{w}_{\alpha(1)} \mathbf{w}_{\alpha(2)} ... \mathbf{w}_{\alpha(s)} \tag{20}$$

Stwierdzenie 2. *Niech antykomutujący układ $(\mathbf{e}_1, ..., \mathbf{e}_n)$ elementów algebry W będzie typu $(c_1, ..., c_n)$. Wówczas dla $\alpha, \beta \subset \{1, ..., n\}$ zachodzi $\mathbf{e}_\alpha \mathbf{e}_\beta = c\mathbf{e}_\gamma$, gdzie skalar c i zbiór $\gamma \subset \alpha \cup \beta$ zależą tylko od ciągu $(\alpha, \beta, c_1, ..., c_n)$ (a nie od algebry W).*

Dowód. Jeśli liczba $\#\alpha + \#\beta$ (gdzie $\#$ to moc zbioru) jest równa 1, to teza jest oczywista. Stosując indukcję względem tej liczby rozważamy dwa przypadki:

a) $\#\alpha > 1$. Niech $p = \inf \alpha$ i $\alpha' = \alpha \setminus \{p\}$. Wtedy $\mathbf{e}_\alpha \mathbf{e}_\beta = \mathbf{e}_p(\mathbf{e}_{\alpha'} \mathbf{e}_\beta) = c' \mathbf{e}_p \mathbf{e}_{\gamma'}$, gdzie skalar c' i zbiór γ' spełniają warunki tezy dla pary (α', β). Obierając zaś c'' i γ spełniające te warunki dla pary $(\{p\}, \gamma')$ stwierdzamy, że możemy przyjąć $c = c'c''$.

b) $\alpha = \{p\}$ dla pewnego p. Z definicji układu ortogonalnego zachodzi wówczas $\mathbf{e}_p \mathbf{e}_\beta = (-1)^l \mathbf{e}_\beta \mathbf{e}_p$, gdzie $l = \#(\beta \setminus \{p\})$. To pozwala zamienić α z β i sprowadzić ten przypadek do poprzedniego. (Patrz też zadanie uzupełniające 1.)□

Wniosek 2. *Niech $(\mathbf{e}_1, ..., \mathbf{e}_n)$ będzie antykomutującym układem elementów algebry W i niech $B := \{\mathbf{e}_\alpha : \alpha \subset \{1, ..., n\}\}$. Wówczas*

a) podprzestrzeń $\operatorname{lin} B$ jest zamknięta ze względu na mnożenie,

b) jeśli układ $(\mathbf{e}_i)$ generuje algebraicznie przestrzeń W, to zbiór B rozpina ją liniowo; zaś jest on jej bazą, jeśli ponadto $dim(W) \geq 2^n$. (W ostatnim przypadku układ $(\mathbf{e}_i)$ jest więc liniowo niezależny.)

Dowód. a) wynika ze stwierdzenia 2, a b) z a), bo $\#B = 2^n \leq \dim W$.□

Twierdzenie 1 (o istnieniu algebry Clifforda). *Istnieje spełniające warunki Clifforda przekształcenie $L : V \to W_q$ w pewną algebrę W_q, która jako przestrzeń wektorowa jest wymiaru $2^{\dim(V)}$.*

Dowód. a) Rozpatrzmy wpierw przypadek, gdy macierz formy q w pewnej bazie jest nieosobliwa i postaci $\mathrm{diag}(c_1, -c_1, ..., c_j, -c_j)$. Ponieważ $\dim \mathcal{M}_{2^j} = (2^j)^2 = 2^{\dim V}$, więc możemy wówczas przyjąć $W_q = \mathcal{M}_{2^j}$ i skonstruować L tak, by wektory bazy przesyłane były na elementy antykomutującego układu w $\mathcal{M}_{2^j}$, typu $(c_1, -c_1, ..., c_j, -c_j)$. Istnienie takiego układu wynika indukcyjnie z przykładu 2 c) i stwierdzenia 1, a różnowartościowość L – z liniowej niezależności tego układu, patrz wniosek 2.

b) W ogólnym przypadku przestrzeń V jest izometryczna z podprzestrzenią przestrzeni $\widetilde{V}$, spełniającej założenie wymienione w a). Wynika to z zadania uzupełniającego 5 w p.1 lub następującego rozumowania. V jest sumą ortogonalną przestrzeni nieosobliwej i przestrzeni izotropowej $V^\perp$, która dla $p := \dim V^\perp$ jest izometryczna z podprzestrzenią $\{\mathbf{x} : x_1 = x_2, x_3 = x_4, ...\}$ nieosobliwej przestrzeni $\mathbb{F}^{2p}$, rozpatrywanej z formą $\sum_{i=1}^p x_{2i-1}^2 - x_{2i}^2$. Wobec tego przestrzeń V jest izometryczna z podprzestrzenią pewnej przestrzeni nieosobliwej V', a ta jest izometryczna z podprzestrzenią żądanej przestrzeni $\widetilde{V}$. Można bowiem przyjąć $\widetilde{V} := V' \times V'$, przy czym gdy macierz wyróżnionej formy przestrzeni V' w pewnej bazie ortogonalnej $\mathbf{v}_1, ..., \mathbf{v}_j$ jest równa $\mathrm{diag}(c_1, ..., c_j)$, to wyróżnioną formę przestrzeni $\widetilde{V}$ definiujemy tak, by w bazie $(\mathbf{v}_1, \mathbf{0}), (\mathbf{0}, \mathbf{v}_1), ..., (\mathbf{v}_j, \mathbf{0}), (\mathbf{0}, \mathbf{v}_j)$ jej macierz była równa $\mathrm{diag}(c_1, -c_1, ..., c_j, -c_j)$.

c) Pozostaje dowieść, że gdy V jest podprzestrzenią przestrzeni $\widetilde{V}$ i teza jest prawdziwa dla $\widetilde{V}$, to jest prawdziwa i dla V. (Zakładamy oczywiście, że forma $q : V \to \mathbb{F}$ jest obcięciem wyróżnionej formy kwadratowej na $\widetilde{V}$.) W tym celu obierzmy spełniające warunki Clifforda przekształcenie $\widetilde{L} : \widetilde{V} \to \widetilde{W}$ w pewną algebrę $\widetilde{W}$ wymiaru $2^{\dim(\widetilde{V})}$. Gdy bazę ortogonalną $\mathbf{v}_1, ..., \mathbf{v}_k$ przestrzeni V rozszerzyć do bazy ortogonalnej $\mathbf{v}_1, ..., \mathbf{v}_l$ przestrzeni $\widetilde{V}$ i przyjąć $\mathbf{e}_i = L(\mathbf{v}_i)$, to zbiór $\{\mathbf{e}_\alpha : \alpha \subset \{1, ..., l\}\}$ jest bazą w $\widetilde{W}$. Jego podzbiór $\{\mathbf{e}_\alpha : \alpha \subset \{1, ..., k\}\}$ rozpina więc liniowo przestrzeń W_q wymiaru 2^k, zamkniętą ze względu na mnożenie. (Kilkukrotnie skorzystaliśmy z wniosku 2.) Indukowane przez $\widetilde{L}$ przekształcenie $L : V \to W_q$ spełnia warunki Clifforda. □

Twierdzenie 2 (o uniwersalności algebry Clifforda)**.** *Niech algebry W i W' i przekształcenia liniowe $L : V \to W$ i $L' : V \to W'$ spełniają warunki Clifforda. Jeśli* $\dim W = 2^{\dim(V)}$*, to istnieje jedyne przekształcenie liniowe $S : W \to W'$ takie, że $S \circ L = L'$ oraz $S(\mathbf{w}_1\mathbf{w}_2) = S(\mathbf{w}_1)S(\mathbf{w}_2)$ dla $\mathbf{w}_1, \mathbf{w}_2 \in W$. Przekształcenie to jest surjektywne, a gdy ponadto* $\dim(W') \geq 2^{\dim(V)}$*, to jest ono izomorfizmem.*

Dowód. Niech $\mathbf{v}_1, ..., \mathbf{v}_n$ będzie bazą ortogonalną przestrzeni V i niech $\mathbf{e}_i := L(\mathbf{v}_i)$ i $\mathbf{e}'_i = L'(\mathbf{v}_i)$ dla $i = 1, ..., n$. Zdefiniujmy $S : W \to W'$ wzorem $S(\mathbf{v}) := \sum_\alpha x_\alpha \mathbf{e}'_\alpha$ dla $\mathbf{v} = \sum_\alpha x_\alpha \mathbf{e}_\alpha \in W$; definicja jest poprawna, bo układ $\{\mathbf{e}_\alpha\}_\alpha$ jest na mocy wniosku 2 bazą przestrzeni W. Ze stwierdzenia 1 wynika, że $S(\mathbf{e}_\alpha \mathbf{e}_\beta) = \mathbf{e}'_\alpha \mathbf{e}'_\beta$, wobec czego S zachowuje mnożenie (bo jest ono rozdzielne względem dodawania). Stąd $S(W)$ jest podprzestrzenią liniową zamkniętą ze względu na mnożenie i zawierającą zbiór $L'(V) = SL(V)$, a więc równą W'. Jeśli więc $\dim W \leq \dim W'$, to S jest izomorfizmem. □

Uwaga 1. Twierdzenie to pozostaje prawdziwe, jeśli w jego sformułowaniu warunek $S(\mathbf{w}_1\mathbf{w}_2) = S(\mathbf{w}_1)S(\mathbf{w}_2)$ zastąpić przez $S(\mathbf{w}_1\mathbf{w}_2) = S(\mathbf{w}_2)S(\mathbf{w}_1)$ (zmieniona kolejność mnożenia po prawej). By się o tym przekonać, wystarczy w algebrze W' zmienić mnożenie na „przeciwne" (tzn. przyjąć $\mathbf{w}_1 * \mathbf{w}_2 := \mathbf{w}_2\mathbf{w}_1$ dla $\mathbf{w}_1, \mathbf{w}_2 \in W'$) i zastosować twierdzenie w poprzedniej formie.

Uwaga 2. a) Para (W_q, L) spełniająca tezę twierdzenia 1 nazywana jest **algebrą Clifforda** formy q. Jest ona jedyna w następującym sensie: gdy (W_q, L) i (W'_q, L') są takimi parami, to istnieje izomorfizm $S : W_q \to W'_q$, opisany w twierdzeniu 2.

b) Z dowodu twierdzenia 1 wynika, że gdy (V, q) jest $2j$–wymiarową przestrzenią Artina, to za W_q można obrać algebrę $\mathcal{M}_{2^j}$.

c) Przy zanurzeniu L spełniającym warunki Clifforda utożsamia się niejednokrotnie każdy wektor $\mathbf{v} \in V$ z jego obrazem $L(\mathbf{v})$. Poniżej dokonujemy wszędzie takich utożsamień. Gdy $(\mathbf{e}_1, ..., \mathbf{e}_n)$ jest ortogonalną bazą przestrzeni V, to bazą W_q (jako przestrzeni wektorowej) jest zbiór $\{\mathbf{e}_\alpha : \alpha \subset \{1, ..., n\}\}$, a mnożenie elementów bazy jest wyznaczone stwierdzeniem 1 (lub zadanie uzpełniającym 1).

d) Algebra Clifforda formy zerowej nazywana jest **algebrą zewnętrzną** (lub: **Grassmanna**) przestrzeni V. Algebrę tę oznaczamy przez $\bigwedge V$, a iloczyn dwóch jej elementów $\mathbf{v}, \mathbf{w}$ – przez $\mathbf{v} \wedge \mathbf{w}$. Algebra $\bigwedge V$ jest wyznaczona jednoznacznie następującymi własnościami: jej wymiar (jako przestrzeni wektorowej) jest równy $2^{\dim V}$, jest ona rozpinana przez podprzestrzeń V algebraicznie, i $\mathbf{v} \wedge \mathbf{v} = 0$ dla wszystkich $\mathbf{v} \in V$.

Uwaga 3. Niech $W_q \supset V$ będzie algebrą Clifforda formy q.

a) Każda izometria przestrzeni (V, q), traktowana jako zanurzenie w W_q, spełnia warunki Clifforda, więc na podstawie twierdzenia 2 i uwagi 1 przedłuża się do jedynego automorfizmu algebry W_q, a także do jedynego antyautomorfizmu tej algebry. (Izomorfizm liniowy $S : W_q \to W_q$ nazywamy **automorfizmem algebry** W_q, jeśli $S(\mathbf{w}_1\mathbf{w}_2) = S(\mathbf{w}_1)S(\mathbf{w}_2)$ dla $\mathbf{w}_1, \mathbf{w}_2 \in W_q$, zaś jej **antyautomorfizmem**, jeśli $S(\mathbf{w}_1\mathbf{w}_2) = S(\mathbf{w}_2)S(\mathbf{w}_1)$ dla $\mathbf{w}_1, \mathbf{w}_2 \in W_q$. Patrz uwaga 1.)

b) W szczególności, istnieje **antyautomorfizm główny** $T : W_q \to W_q$ –jedyny, który na V jest identycznością I_V. (Stosujemy a) do izometrii I_V.) Podobnie, istnieje **automorfizm główny** $S : W_q \to W_q$ – jedyny, który na V jest równy $-I_V$. Jeśli $(\mathbf{e}_i)_{i=1}^n$ jest bazą ortogonalną przestrzeni V, to działania S i T na elementach bazy $\{\mathbf{e}_\alpha : \alpha \subset \{1, \ldots, n\}\}$ są takie: przy $1 \leq \alpha(1) < \cdots < \alpha(j) \leq n$ zachodzi $S(\mathbf{e}_\alpha) = (-1)^j\mathbf{e}_\alpha$ i $T(\mathbf{e}_\alpha) = \mathbf{e}_{\alpha(j)} \ldots \mathbf{e}_{\alpha(1)} = (-1)^{j(j-1)/2}\mathbf{e}_a$, skąd m.in. $ST(\mathbf{e}_\alpha) = TS(\mathbf{e}_\alpha)$.

c) Każde z przekształceń $L \in \{S, T, ST = TS\}$ jest więc inwolutywne (tzn. spenia warunek $L^2 = I_W$) i gdy dla $\varepsilon = \pm$ przyjąć $W^\varepsilon_{q,L} := \{\mathbf{w} \in W_q : L(\mathbf{w}) = \varepsilon\mathbf{w}\}$, to L jest symetrią przestrzeni W_q wzdłuż $W^-_{q,L}$, względem podprzestrzeni $W^+_{q,L}$. (Ta jest i podalgebrą algebry W_q).

W miejsce $W^{\varepsilon}_{q,S}$ piszemy dalej W^{ε}_{q}. Z b) wynika, że W^{+}_{q} (odp. W^{-}_{q}) jest powłoką liniową elementów bazowych $\mathbf{e}_\alpha$, dla $\#\alpha$ będących liczbą parzystą (odp. nieparzystą). W szczególności, $\dim W^{+}_{q} = \dim W^{-}_{q} = 2^{\dim V - 1}$.

Uwaga 4. a) Oznaczmy przez Γ_q lub Γ zbiór wszystkich elementów algebry W_q, będących iloczynami nieizotropowych wektorów z V. Latwo sprawdzić, że elementy zbioru Γ posiadają w Γ odwrotność (wynika to stąd, że $\frac{1}{q(\mathbf{u})}\mathbf{u}$ jest odwrotnością wektora nieizotropowego $\mathbf{u} \in V$). Dla $\mathbf{u} \in \Gamma$ i $\mathbf{w} \in W_q$ przyjmijmy $F_{\mathbf{u}}(\mathbf{w}) := S(\mathbf{u})\mathbf{w}\mathbf{u}^{-1}$.

b) Gdy $\mathbf{u}, \mathbf{v} \in V$ i $\mathbf{u}$ jest wektorem nieizotropowym, to $F_{\mathbf{u}}(\mathbf{v})$ jest obrazem $\mathbf{v}$ przy symetrii ortogonalnej $V \to V$ względem podprzestrzeni $\mathbf{u}^{\perp}$. (Dowód: jeśli $\mathbf{v} = c\mathbf{u} + \mathbf{v}_1$, gdzie $\mathbf{v}_1 \perp \mathbf{u}$, to $F_{\mathbf{u}}(\mathbf{v}) = (-\mathbf{u})(c\mathbf{u} + \mathbf{v}_1)\mathbf{u}^{-1} = -c\mathbf{u} + \mathbf{v}_1$, bo $\mathbf{u}\mathbf{v}_1 = -\mathbf{v}_1\mathbf{u}$ na mocy $(C1')$.) Ponieważ $F_{\mathbf{u}_1}(F_{\mathbf{u}_2}(\mathbf{v})) = F_{\mathbf{u}_1\mathbf{u}_2}(\mathbf{v})$, więc wynika stąd, że dla każdego elementu $\mathbf{u} \in \Gamma$ wzór $V \ni \mathbf{v} \mapsto F_{\mathbf{u}}(\mathbf{v})$ zadaje przekształcenie $V \to V$, które oznaczymy przez $F_{\mathbf{u}}$ i które jest złożeniem skończenie wielu symetrii zwierciadlanych.

c) Dowiedliśmy więcej: $\{F_{\mathbf{u}} : \mathbf{u} \in \Gamma\}$ jest dokładnie zbiorem opisanych złożeń. Gdy przestrzeń V jest nieosobliwa, jest to zbiór wszystkich izometrii $V \to V$; patrz twierdzenie E. Cartana w p.1. Natomiast $\{F_{\mathbf{u}} : \mathbf{u} \in \Gamma \cap W^{+}_{q}\}$ jest wówczas zbiorem izometrii o wyznaczniku równym 1. (Jest to zadanie ....). Oczywiście, dla $\mathbf{u} \in \Gamma \cap W^{+}_{q}$ izometria $F_{\mathbf{u}}$ zadana jest prostszym wzorem $F_{\mathbf{u}}(\mathbf{v}) = \mathbf{u}\mathbf{v}\mathbf{u}^{-1}$. Można też udowodnić, że gdy $F_{\mathbf{u}} = F_{\mathbf{u}'}$ $(\mathbf{u}, \mathbf{u}' \in \Gamma)$, to $\mathbf{u}' = \lambda\mathbf{u}$ dla pewnego skalara λ.

Uwaga powyższa przesądza o znaczeniu algebry Clifforda W_q i zbioru Γ, zwanego **grupą Clifforda** formy q. Dalsze informacje o algebrach Clifforda znaleźć można w [Artin], [Komorowski] i [Porteous]; patrz też niżej.

Zadania uzupełniające.

1. Dowieść, że przy oznaczeniach stwierdzenia 1 i jego dowodu zachodzi
   a) $s = (-1)^t$, gdzie $t = (\#\alpha)(\#\beta) - \#(\alpha \cap \beta)$,
   b)* $c = ...$

2. Dowieść, że dla $c = 1$ stwierdzenie 2 pozostanie prawdziwe, gdy przyjąć

$$\mathbf{A}'_i = \begin{pmatrix} \mathbf{0} & \mathbf{A}_i \\ \mathbf{A}_i & \mathbf{0} \end{pmatrix} \text{ dla } i \leq j, \quad \mathbf{A}'_{j+1} = \mathrm{diag}(\mathbf{I}_k, -\mathbf{I}_k), \quad \mathbf{A}'_{j+2} = \begin{pmatrix} \mathbf{0} & -\mathbf{I}_k \\ \mathbf{I}_k & \mathbf{0} \end{pmatrix}$$

3. Niech $V = \mathbb{R}^n$ i $q(\mathbf{v}) = -||\mathbf{v}||^2$ dla $\mathbf{v} \in V$. Z wniosku 1 wynika, że można przyjąć $W_q = \mathbb{C}$ gdy $n = 1$ oraz $W_q = \mathbb{H}$ gdy $n = 2$.
   a) W obu tych przypadkach opisać wzorem włożenia Clifforda $L : V \to W_q$.
   b) Dowieść, że gdy $n = 3$, to można przyjąć $W_q = \mathbb{H} \times \mathbb{H}$, przy czym działania w W_q zadane są naturalnymi wzorami („po współrzędnych"), a włożenie $V \to W_q$ zadane jest tak: $L(x, y, z) = x(i, -i) + y(j, -j) + z(k, -k)$.

c) Dowieść, że gdy $n = 4$, to można przyjąć $W_q = \mathcal{M}_2(\mathbb{H})$ oraz $L(x, y, z, t) = x\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + y\begin{pmatrix} j & 0 \\ 0 & -j \end{pmatrix} + z\begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix} + t\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ dla $x, y, z, t \in \mathbb{R}$.

4. Korzystając ze stwierdzenia 1 lub zadania 1 dowieść, że:
a) Gdy $V = \mathbb{R}^3$ i $q = -x_1^2 - x_2^2 + x_3^2$, to za W_q przyjąć można algebrę $\mathcal{M}_2(\mathbb{C})$ (nad $\mathbb{R}$). Zdefiniować też jawnym wzorem włożenie Clifforda $L : V \to W_q$.
b) Jak wyżej, ale przy $V = \mathbb{R}^4$, $q = -x_1^2 - x_2^2 - x_3^2 + x_4^2$ i $W_q = \mathcal{M}_2(\mathbb{H})$.
c)* Jak w a), ale przy $V = \mathbb{R}^6$, $q = -x_1^2 - ... - x_5^2 + x_6^2$ i $W_q = \mathcal{M}_4(\mathbb{H})$.

5. a) Dowieść, że przekształcenie $\begin{pmatrix} a & c \\ b & d \end{pmatrix} \mapsto \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$ jest inwolutywnym antyizomorfizmem algebry $\mathcal{M}_2(\mathbb{F})$, przy czym gdy $\mathbf{A}$ jest jedną z macierzy $\mathrm{diag}(a, -a)$, $\begin{pmatrix} 0 & c \\ b & 0 \end{pmatrix}$, to jej obrazem jest $-\mathbf{A}$.
b) Opisać antyizomorfizm główny T, gdy i) $q = x_1^2 + x_2^2$, lub ii) $q = x_1^2 - x_2^2$, a algebrę W_q i włożenie $\mathbb{F}^2 \to W_q$ interpretujemy jak w częściach c1) i c2) wniosku 1.
c) To samo lecz z $\mathbb{F} = \mathbb{R}$, $q = -x_1^2 - x_2^2 + x_3^2$ i $W_q = \mathcal{M}_2(\mathbb{C})$, patrz zadanie 3 a).

6. Niech $q(\mathbf{v}) = -||\mathbf{v}||^2$ dla $\mathbf{v} \in \mathbb{R}^n$. Dla $n = 1$ i dla $n = 2$ zbadać, czy algebry W_{-q} i W_q są izomorficzne.

7. Dowieść, że gdy $\mathbf{v}_1, ..., \mathbf{v}_n$ jest ortogonalną bazą przestrzeni V, typu $(c_1, \ldots, c_n)$, to:
a) układ $(\mathbf{v}_1\mathbf{v}_n, ..., \mathbf{v}_{n-1}\mathbf{v}_n, \mathbf{v}_n)$ jest ortogonalny typu $(-c_1c_n, ..., -c_{n-1}c_n, c_n)$.
b) układ $(\mathbf{v}_1\mathbf{v}_n, ..., \mathbf{v}_{n-1}\mathbf{v}_n)$ rozpina algebrę W_q^+.

8. Dowieść, że gdy $\mathbf{u} \in V$ jest wektorem nieizotropowym, a formę kwadratową q' na przestrzeni V zdefiniować tak, by $q'(\mathbf{u}) = q(\mathbf{u})$ i $q'(\mathbf{v}) = -q(\mathbf{u})q(\mathbf{v})$ dla $\mathbf{v} \in \mathbf{u}^\perp$, to
a) przy odpowiednim włożeniu $V \to W_q$ można przyjąć $W_{q'} = W_q$,
b) przy odpowiednim włożeniu $\mathbf{u}^\perp \to W_q^+$ można za algebrę Clifforda formy $q'_{|\mathbf{u}^\perp}$ przyjąć W_q^+.

9. Gdy inkluzja przestrzeni V w algebrę W spełnia warunek (C1) i wektor $\mathbf{u} \in V$ jest nieizotropowy, to symetria ortogonalna przestrzeni V względem $\mathbf{u}^\perp$ jest zadana wzorem $\mathbf{v} \mapsto -\mathbf{u}^{-1}\mathbf{v}\mathbf{u}$. (Nie zakładamy warunku (C2) ani równości $\dim W = 2^{\dim V}$.) Powiązać to z .....

10. a) Dowieść, że iloczyn parzystej liczby wektorów $\mathbf{v}_i \in V$ należy do W_q^+, a nieparzystej – do W_q^-.
b) W uwadze 4c) dać brakujące uzasadnienie.

11. Niech wektor $\mathbf{u} \in V$ spełnia warunek $q(\mathbf{u}) = -1$ i niech przekształcenie $\mathbf{u}^\perp \to W$ spełnia warunki Clifforda. Przedłużmy je do przekształcenia liniowego $K : V \to W$, przyjmując $K(\mathbf{u}) = -\mathbf{e}$; zdefiniujmy też przekształcenie $L : V \to W_q^+$ wzorem $L(\mathbf{v}) =$

$\mathbf{vu}$, $\mathbf{v} \in V$. Przy założeniu, że $\dim W = 2^{\dim V - 1}$ dowieść, że

a) Istnieje izomorfizm algebr $J : W \to W_q^+$ taki, że $JK = L$.

b) Gdy na $K(V)$ i $L(V)$ rozpatrywać formy, odpowiadające q przy przekształceniach K i L, odpowiednio, to każda izometria z $Iso^+(K(V))$ jest postaci $\mathbf{z} \mapsto \mathbf{wzu}^{-1}\mathbf{w}^{-1}\mathbf{u}$, dla pewnego $\mathbf{w} \in \Gamma_q \cap W_q^+$, zaś każda izometria z $Iso^+(L(V))$ – postaci $\mathbf{y} \mapsto J^{-1}\big(J(\mathbf{w})J(\mathbf{y})\mathbf{u}^{-1}J(\mathbf{w}^{-1})\mathbf{u}\big)$ dla pewnego $\mathbf{w} \in \Gamma_W$, gdzie Γ_W oznacza zbiór iloczynów (w algebrze W) nieizotropowych wektorów z $K(V)$.

c) Zapisać ostatni wzór w postaci $\mathbf{y} \mapsto \mathbf{wy}R(\mathbf{w})$ (mnożenie w W), gdzie R jest antyautomorfizmem grupy Γ_W, danym wzorem $R(\mathbf{w}) = J^{-1}\big(\mathbf{u}^{-1}J(\mathbf{w}^{-1})\mathbf{u}\big)$ (mnożenie w W_q^+).

d) Dowieść, że $R(\mathbf{w}) = -\frac{1}{q(\mathbf{v})}\mathbf{w}$ gdy $\mathbf{w} = K(\mathbf{v})$ i $\mathbf{v} \in V$.

12. Na przestrzeni $\mathbb{R}^4$ rozpatrzmy formę $x^2 + y^2 + z^2 - t^2$, przyjmijmy $W = \mathcal{M}_2(\mathbb{C})$ (jako algebra nad $\mathbb{R}$) i zadajmy $K : \mathbb{R}^4 \to W$ wzorem $K(x, y, z, t) = \begin{pmatrix} x-t & y-zi \\ y+zi & -x-t \end{pmatrix}$. Dowieść, że:

a) Obrazem przekształcenia K jest zbiór $\mathcal{H}$ macierzy hermitowskich, a formie q odpowiada przy K forma $\mathcal{H} \ni \mathbf{A} \mapsto -\det(\mathbf{A})$. Ponadto, przekształcenie obcięte $K_{|\mathbf{u}^\perp}$, gdzie $\mathbf{u} = (0, 0, 0, 1)$, spełnia warunki Clifforda. (Por. przykład 2d).)

b) Przy oznaczeniach poprzedniego zadania zachodzi $\det(\mathbf{W})R(\mathbf{W}) = \mathbf{W}^*$ dla $\mathbf{W} \in \mathcal{H}$, a tym samym i dla wszystkich $\mathbf{W} \in \Gamma_W$.

c) Wywnioskować, że każda izometria z $Iso^+(\mathcal{H}, -\det)$ jest postaci $\mathbf{A} \mapsto \frac{1}{\det(\mathbf{W})}\mathbf{WAW}^*$, dla pewnej (zależnej od izometrii) macierzy $\mathbf{W} \in \mathcal{M}_2(\mathbb{C})$ takiej, że $\det(\mathbf{W}) \in \mathbb{R}$ lub, równoważnie, dla takiej, że $\det(\mathbf{W}) = \pm 1$.

???????????? Poprawić i uzupełnić

13. Dla układu $\mathcal{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ elementów algebry W przyjmijmy $B_\mathcal{V} = \{\mathbf{v}_\alpha : \alpha \subset \{1, \ldots, n\}\}$.

a) Gdy $\mathrm{lin}\mathcal{V}' \subset \mathrm{lin}\mathcal{V}$, to $\mathrm{lin}B_{\mathbf{v}'} \subset \mathrm{lin}B_\mathcal{V}$. W szczególności, $\mathrm{lin}B_{\mathcal{V}'} = \mathrm{lin}B_\mathcal{V}$ gdy $\mathrm{lin}\mathcal{V}' = \mathrm{lin}\mathcal{V}$.

b) Gdy W jest algebrą Clifforda przestrzeni V i $\mathcal{V}$ jest bazą przestrzeni V, to zbiór $B_\mathcal{V}$ jest bazą przestrzeni wektorowej W.

14. Dowieść, że gdy $\mathbf{v}_1, ..., \mathbf{v}_n$ jest ortogonalną bazą przestrzeni V i $\mathbf{w} \in W_q$, to:

a) $\mathbf{w} \in \mathbb{F}\mathbf{e}$ lub $\mathbf{wv} \neq \mathbf{vw}$ dla pewnego $\mathbf{v} \in V$. W szczególności $\mathbf{v}_1\mathbf{v}_2...\mathbf{v}_n \in \mathbb{F}\mathbf{e}$ jeśli $n \in 2\mathbb{Z}$.

b) n jest liczbą nieparzystą i $\mathbf{w} \in \mathbb{F}\mathbf{v}_1\mathbf{v}_2...\mathbf{v}_n$ lub $\mathbf{wv} \neq -\mathbf{wv}$ dla pewnego $\mathbf{v} \in V$.

§ 7. * Dodatek: czasoprzestrzeń Minkowskiego

1. Podstawowe własności metryczne

.

Przestrzeń Minkowskiego= przestrzeń rzeczywista wymiaru 4 z symetryczną formą metryczną o sygnaturze (3,1). Czasoprzestrzeń= $\mathbb{R}^4$ z iloczynem skalarnym

$$\langle \mathbf{u}, \mathbf{v} \rangle := -u_0v_0 + u_1v_1 + u_2v_2 + u_3v_3. \tag{1}$$

Jeśli nie powiedziano inaczej zakładamy, że rozważane iloczyny skalarne mają sygnaturę $(3, 1)$; oznaczamy je na ogól przez $\langle ., . \rangle$ lub literami μ, ν.

Wektory $\mathbf{u}$ przestrzeni Minkowskiego spełniające warunek $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ nazywamy **przestrzennymi**, spełniające warunek $\langle \mathbf{u}, \mathbf{u} \rangle < 0$ -**czasowymi**, zaś te, dla których $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ –**świetlnymi** (inna nazwa tych ostatnich to wektory **zerowe** lub **izotropowe**). Zbiór wektorów przestrzeni Minkowskiego, spełniających warunek $\langle \mathbf{u}, \mathbf{u} \rangle = c$, nazywamy **sferą o środku 0 i kwadracie promienia równym** c. (Gdy $c < 0$, promień sfery byłby urojony, więc go nie wprowadzamy.)

Gdy zamiast $\mathbb{R}^4$ z formą (1) rozpatrzymy $\mathbb{R}^3$ z formą $u_0v_0 + u_1v_1 + u_2v_2$, to zbiory wektorów przestrzennych, czasowych i świetlnych łatwo możemy sobie wyobrazić. Istotnie, zbiór wektorów świetlnych, zadany równaniem $-x_0^2+x_1^2+x_2^2 = 0$, tworzy powierzchnię boczną stożka obrotowego o osi $\mathbb{R}(1, 0, 0)$ i tworzącej $\mathbb{R}(1, 1, 0)$, zaś zbiór wektorów czasowych, zadany nierównością $-x_0^2 + x_1^2 + x_2^2 < 0$, tworzy wnętrze tego stożka; zewnętrze stożka to wektory przestrzenne. (Z tego względu i dla $(\mathbb{R}^4, \langle ., . \rangle$ mówimy o "stożku" wektorów czasowych.) Warto też naszkicować sferę o środku w 0 i promieniu 1 w $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ (jest nią hiperboloida jednopowłokowa $-x_0^2 + x_1^2 + x_2^2 = 1$) oraz sferę o tymże środku i kwadracie promienia równym -1 (jest nią hiperboloida dwupowłokowa $x_0^2 - x_1^2 - x_2^2 = 1$).

Rysunsk Wektory czasowe, przestrzenne i świetlne oraz sfery o kwadratach promieni równych ± 1 w przestrzeni $\mathbb{R}^3$ z metryką o sygnaturze $(2, 1)$.

Uwaga 1. Jeśli $\mathbf{u}$ jest wektorem czasowym, to na $\mathbf{u}^\perp$ sygnatura iloczynu skalarnego jest równa $(3, 0)$, skąd żaden wektor ortogonalny do $\mathbf{u}$ nie jest czasowy. Równoważnie: iloczyn skalarny dwóch wektorów czasowych jest różny od 0.

W przestrzeni Minkowskiego możemy wykorzystywać własności jej jednorodności, wynikające z rezultatów dowiedzionych w ....

Przykład 1. Dowiedziemy, że w przestrzeni Minkowskiego ma miejsce nierówność:

$$\text{gdy pewien z wektorów } \mathbf{u}, \mathbf{v} \text{ jest czasowy, to } \langle \mathbf{u}, \mathbf{v} \rangle^2 \geq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \tag{21}$$

W tym celu zauważmy, że istnieje izometria rozważanej przestrzeni Minkowskiego na czasoprzestrzeń, przeprowadzająca wektor czasowy $\mathbf{u}$ na proporcjonalny do $(1, 0, 0, 0)$.

(Patrz ....) Zastępując $\mathbf{u}$ i $\mathbf{v}$ przez ich obrazy możemy więc zakładać, że leżą one w czasoprzestrzeni i $\mathbf{u} = (t, 0, 0, 0)$. W tym jednak przypadku (.) sprowadza się do oczywistej nierówności $t^2 v_0^2 \geq -t^2(-v_0^2 + v_1^2 + v_2^2 + v_3^2)$. □

Dla wektora u przestrzeni Minkowskiego przyjmijmy

$$|\mathbf{u}| := \varepsilon\sqrt{|\langle \mathbf{u}, \mathbf{u}\rangle|}, \text{ gdzie } \varepsilon := \mathrm{Sgn}(\langle \mathbf{u}, \mathbf{u}\rangle)$$

Wniosek 1. *Jeśli wektory* $\mathbf{u}, \mathbf{v}$ *są czasowe i* $\langle \mathbf{u}, \mathbf{v}\rangle < 0$, *to* $\mathbf{u} + \mathbf{v}$ *jest wektorem czasowym i* $|\mathbf{u} + \mathbf{v}| < |\mathbf{u}| + |\mathbf{v}|$.

Dowód. Ponieważ wszystkie czynniki iloczynów występujących w nierówności (1.5) są ujemne, więc $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle = -|\mathbf{u}|^2 - |\mathbf{v}|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle < -|\mathbf{u}|^2 - |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}| = -(|\mathbf{u}| + |\mathbf{v}|)^2$. Stąd teza.□

Zadanie 1. Wektory czasowe $\mathbf{u}, \mathbf{v}$ nazwijmy **zgodnymi**, jeśli $\langle \mathbf{u}, \mathbf{v}\rangle < 0$. Dowieść, że:

a) Gdy wektory czasowe $\mathbf{u}, \mathbf{v}$ są zgodne, to $|t\mathbf{u} + (1 - t)\mathbf{v}| \leq t|\mathbf{u}| + (1 - t)|\mathbf{v}|$ dla $t \in [0, 1]$, skąd odcinek $\{t\mathbf{u} + (1 - t)\mathbf{v} : t \in [0, 1]\}$ składa się tylko z wektorów czasowych.

b) Wektory czasowe $\mathbf{u}$ i $\mathbf{v}$ wtedy i tylko wtedy są zgodne, gdy można je połączyć drogą wektorów czasowych. Tu **drogą wektorów czasowych** nazywamy każdą funkcję ciągłą $[0, 1] \to \mathbb{R}^4$, przyjmującą wartości w wektorach czasowych.

c) Zgodność wektorów czasowych jest relacją równoważności i każdy wektor czasowy jest zgodny badź z zadanym wektorem czasowym $\mathbf{u}$, bądź z $-\mathbf{u}$.

Wybór jednej z dwóch klas abstrakcji relacji zgodności wektorów czasowych nazywamy **orientacją czasu** rozważanej przestrzeni Minkowskiego.

Problem 1. Niech $S_+ := \{\mathbf{u} \in \mathbb{R}^4 : \langle \mathbf{u}, \mathbf{u}\rangle = -1 \text{ oraz } u_0 > 0\}$. Dowieść, że wzór $d(\mathbf{u}, \mathbf{v}) := \operatorname{arc\,cosh}\langle \mathbf{u}, \mathbf{v}\rangle$ zadaje metrykę na S_+.

2. Kwaternionowy i macierzowe modele czasoprzestrzeni.

Przestrzeń $\mathbb{R}^4$ utożsamić możemy ze zbiorem kwaternionów $\mathbb{H}$ w sposób opisany w §III.... . Prowadzi to do wyrażenia formy Minkowskiego przy pomocy operacji algebraicznych w $\mathbb{H}$ (patrz wzór .... w §III ....):

$$\mu(u, v) = -\mathrm{Re}(uv) = -\frac{1}{2}(uv + v^*u^*) \quad \text{dla} \quad u, v \in \mathbb{H}.$$

Zamiast ciała kwaternionów $\mathbb{H}$ rozpatrywać też można jego macierzowy model. Przypomnijmy, że każdy kwaternion $u = t + xi + yj + zk$ utożsamiony jest w nim z macierzą

$$\Phi(u) := \begin{pmatrix} t + zi & -x + yi \\ x + yi & t - zi \end{pmatrix}.$$

Wynika stąd, że $\mathrm{Re}(u) = \frac{1}{2}\mathrm{tr}(\Phi(u))$ dla $u \in \mathbb{H}$. Ponieważ ponadto Φ zachowuje mnożenie, więc formie μ odpowiada przy Φ forma μ' na $\mathbb{H}'$ zadana wzorem

$$\mu'(\mathbf{A},\mathbf{B}) = -\frac{1}{2}\mathrm{tr}(\mathbf{AB}) \ \text{ dla } \mathbf{A},\mathbf{B} \in \mathbb{H}'.$$

$(\mathbb{H},\mu)$ oraz $(\mathbb{H}',\mu')$ są przestrzeniami Minkowskiego, a ich bogata struktura algebraiczna może być użyteczna.

Opis odmiennego macierzowego modelu czasoprzestrzeni poprzedzimy następującą obserwacją. Obrazem przy Φ zbioru kwaternionów czystych jest zbiór wszystkich 2×2-macierzy antyhermitowskich o śladzie 0 (patrz §V.... i §VI.....). Nie zmieniając go na kwaternionach czystych, można odwzorowanie Φ zmodyfikować tak, by zbiór wszystkich kwaternionów przeprowadzany był na zbiór $\mathcal{A}$ macierzy antyhermitowskich (o dowolnym śladzie). Oto wzór na tak zmienione odwzorowanie:

$$\Psi(u) := \begin{pmatrix} (t+z)i & -x+yi \\ x+yi & (t-z)i \end{pmatrix}. \tag{1.75}$$

Ψ jest liniowym monomorfizmem nad $\mathbb{R}$, lecz <u>nie</u> zachowuje mnożenia ani sprzężenia. Bezpośredni rachunek wykazuje ważną zależność:

$$\det(\Psi(u)) = \mu(u,u) \text{ dla kadego } u, \ \text{ gdzie } \mu \text{ jest form Minkowskiego.} \tag{2}$$

Szukanym nowym modelem czasoprzestrzeni jest więc przestrzeń $\mathcal{A}$, wyposażona w formę kwadratową $\mathcal{A} \ni \mathbf{A} \mapsto \det(\mathbf{A})$. Równość .... pokazuje, że Ψ jest izometrią czasoprzestrzeni na $\mathcal{A}$.

Oznaczmy na koniec przez $\mathcal{H}$ zbiór 2×2–macierzy samosprzężonych. Mnożenie przez dowolną liczbę czysto urojoną c zmienia znak wyznacznika i ustala liniową bijekcję $\mathcal{A} \to \mathcal{H}$. W związku z tym kolejnym modelem czasoprzestrzeni jest $(\mathcal{H}, -\det)$. Model ten umożliwia wykorzystanie znanych własności macierzy samosprzężonych (np. tego, że mają rzeczywiste wartości własne, że można mówić o ich określoności itp). Pod tym względem jest on dogodniejszy, niż $(\mathcal{A}, \det)$; własności jego działań algebraicznych są jednak jeszcze bardziej odległe od działań w $\mathbb{H}$, niż to jest w przypadku $(\mathcal{A}, \nu)$. Jeśli chcieć, by $c\Psi(I) = \mathbf{I}$ należy przyjąć $c = -i$; wówczas $c\Psi$ jest zadane wzorem

$$(t,x,y,z) \leftrightarrow \begin{pmatrix} t+z & xi+y \\ -xi+y & t-z \end{pmatrix}. \tag{3}$$

3. Grupa Lorentza

Zbiór wszystkich automorfizmów zadanej przestrzeni Minkowskiego M jest grupą przekształceń; będziemy ją oznaczać przez $\mathrm{Lo}(M)$ i nazywać **grupą Lorentza** tej przestrzeni. Gdy nie jest zaznaczone, o jakiej przestrzeni Minkowskiego M mowa, zakładamy, że jest to czasoprzestrzeń ($\mathbb{R}^4$ z formą zadaną wzorem (....)) i piszemy Lo zamiast $\mathrm{Lo}(M)$. Oznaczając przez T dowolny izomorfizm rozważanej przestrzeni Minkowskiego M na czasoprzestrzeń otrzymujemy izomorfizm $L \mapsto TLT^{-1}$ grupy $\mathrm{Lo}(M)$ na Lo. Elementy grupy Lo nazywamy **przekształceniami Lorentza**.

Lemat 1. *Niech $T \in \mathrm{Lo}(M)$ i niech* $\mathbf{u}$ *oraz* $\mathbf{e}$ *będą wektorami czasowymi w M. Jeśli wektor* $\mathbf{e}$ *jest zgodny z $T(\mathbf{e})$, to wektor* $\mathbf{u}$ *jest zgodny z $T(\mathbf{u})$.*

Dowód. Możemy zakładać, że wektory $\mathbf{e}$ i $\mathbf{u}$ są zgodne (jeśli nie, zastąpimy $\mathbf{e}$ przez $-\mathbf{e}$). Wykorzystamy wielokrotnie to, że wektory czasowe wtedy i tylko twedy są zgodne, gdy można je połączyć drogą wektorów czasowych. Możemy więc połączyć $\mathbf{e}$ z $\mathbf{u}$ taką drogą. Jej obraz przy T jest drogą wektorów czasowych łączących $T(\mathbf{e})$ z $T(\mathbf{u})$. Jeśli więc $\mathbf{e}$ można połaczyć drogą wektorów czasowych z $T(\mathbf{e})$, to można taką drogą połączyć i $\mathbf{u}$ z $T(\mathbf{u})$: wystarcza wpierw iść po drodze z $\mathbf{u}$ do $\mathbf{e}$, następnie z $\mathbf{e}$ do $T(\mathbf{e})$ i na koniec z $T(\mathbf{e})$ do $T(\mathbf{u})$. $\square$

Gdy $T \in \mathrm{Lo}$ przeprowadza pewien (a więc i każdy) wektor czasowy na wektor z nim zgodny, to powiemy, że T **zachowuje orientację czasu**. Mamy więc rozbicie grupy Lorentza następujące: $\mathrm{Lo} = \mathrm{Lo}_+ \cup \mathrm{Lo}_- = \mathrm{Lo}_{++} \cup \mathrm{Lo}_{+-} \cup \mathrm{Lo}_{-+} \cup \mathrm{Lo}_{--}$, gdzie

$$\mathrm{Lo}_+ = \{T \in \mathrm{Lo} : \det(T) > 0\} \text{ oraz } \mathrm{Lo}_- = \{T \in \mathrm{Lo} : \det(T) < 0\},$$

$$\mathrm{Lo}_{\varepsilon+} = \{T \in \mathrm{Lo}_\varepsilon : T \text{ zachowuje orientację czasu}\} \text{ i } \mathrm{Lo}_{\varepsilon-} = \mathrm{Lo}_\varepsilon \setminus \mathrm{Lo}_{\varepsilon+} \quad (\varepsilon = \pm).$$

Zadanie 1. a) Gdy oba lub żadne z przekształceń Lorentza T_1, T_2 zachowuje orientację czasu, to zachowuje ją złożenie T_1T_2. Implikacja przeciwna też ma miejsce.

b) Przekształcenie $(t, x, y, z) \mapsto (-t, -x, y, z)$ zmienia orientację czasu czasoprzestrzeni, lecz ma dodatni wyznacznik.

Oprócz dróg w przestrzeni Minkowskiego M wygodnie jest wprowadzić **drogi w grupie Lorentza** $\mathrm{Lo}(M)$. Przyjmując dla prostoty za M czasoprzestrzeń, przez drogę taką rozumiemy każdą funkcję $[0, 1] \ni t \mapsto S_t \in \mathrm{Lo}$ o tej własności, że dla każdego wektora $\mathbf{v} \in \mathbb{R}^4$ funkcja $t \mapsto S_t(\mathbf{v})$ (z odcinka $[0, 1]$ w $\mathbb{R}^4$) jest ciągła.

Zadanie 2. Gdy (S_t) jest drogą w Lo, to $S_0 \in \mathrm{Lo}_+ \Rightarrow S_1 \in \mathrm{Lo}_+$, a także $S_0 \in \mathrm{Lo}_{++} \Rightarrow S_1 \in \mathrm{Lo}_{++}$.

Opiszemy teraz jeden ze sposobów badania grupy Lorentza, polegający na powiązaniu jej elementów z 2×2-macierzami zespolonymi. Opiera się on na obserwacji, że każda macierz nieosobliwa $\mathbf{V} \in \mathcal{M}_2(\mathbb{C})$ wyznacza operator liniowy $T_{\mathbf{V}} : \mathcal{H} \to \mathcal{H}$ zadany wzorem

$$T_{\mathbf{V}}(\mathbf{A}) := \mathbf{V}\mathbf{A}\mathbf{V}^* \text{ dla } \mathbf{A} \in \mathcal{H}.$$

(Dlaczego $T_{\mathbf{V}}(\mathcal{H}) \subset \mathcal{H}$?) Gdy $\det(\mathbf{V}) = \pm 1$, to $\det(\mathbf{VAV}^*) = \det(\mathbf{A})$ i wobec tego $T_{\mathbf{V}} \in \mathrm{Lo}$; ponadto $T_{\mathbf{U}}T_{\mathbf{V}} = T_{\mathbf{UV}}$. Wzór $H(\mathbf{V}) = T_{\mathbf{V}}$ zdaje więc homomorfizm grupy $G = \{\mathbf{V} \in \mathcal{M}_2(\mathbb{C}) : \det(\mathbf{V}) = \pm 1\}$ w grupę Lorentza $\mathrm{Lo}(\mathcal{H})$. Wyznaczymy obrazy przy tym homomorfiźmie zarówno grupy G, jak i grupy $\mathrm{SL}_2(\mathbb{C}) = \{\mathbf{V} \in \mathcal{M}_2(\mathbb{C}) : \det(\mathbf{V}) = 1\}$.

Twierdzenie 1. *Niech $S = T_{\mathbf{V}}$ dla pewnego $\mathbf{V} \in G$. Wówczas istnieje droga w $H(G)$, łącząca S z identycznością gdy* $\det(\mathbf{V}) = 1$, *zaś z przekształceniem ..... gdy* $\det(\mathbf{V}) = -1$.

Dowód. Rozpatrzmy wpierw przypadek, gdy $\det(\mathbf{V}) = 1$. Wystarcza wtedy przyjąć $S_t := H(\mathbf{V}_t)$, gdzie $(\mathbf{V}_t)$ jest drogą w $\mathrm{SL}_2(\mathbb{C})$, łączącą $\mathbf{I}$ z $\mathbf{V}$. (Istnienie drogi $(\mathbf{V}_t)$ wynika bezpośrednio z zadania uzupełniającego 2 w §V.4.3.)

Natomiast gdy $\det(\mathbf{V}) = -1$, to $\mathbf{V} = \mathbf{DU}$, gdzie $\mathbf{D} = \mathrm{diag}(1, -1)$ i $\det(\mathbf{U}) = 1$. Ponieważ $T_{\mathbf{D}} = L$ i istnieje droga (S_t), łącząca $T_{\mathbf{U}}$ z identycznością, więc (LS_t) jest drogą, łączącą $S = LT_{\mathbf{U}}$ z L.$\square$

Twierdzenie 2. *Obrazem* $\mathrm{im}(H)$ *homomorfizmu H jest zbiór tych transformacji Lorentza, których wyznacznik jest dodatni; natomiast obrazem przy H grupy* $\mathrm{SL}_2(\mathbb{C})$ *jest zbiór tych transformacji Lorentza, których wyznacznik jest dodatni i które zachowują orientację czasu. Równoważnie:*

$$\mathrm{im}(H) = \mathrm{Lo}_+(\mathcal{H}) \quad \text{i} \quad H(\mathrm{SL}_2(\mathbb{C})) = \mathrm{Lo}_{++}(\mathcal{H}).$$

Dowód. Teza wynika z następujących dwóch stwierdzeń, z których pierwsze zależy od wiadomości o algebrach Clifforda z poprzedniego paragrafu. (Dowód, zastępujący je wynikami o obrotach euklidesowej przestrzeni $\mathbb{R}^3$, relegowany jest do zadań.... i .....)

i) Z zadania .... wynika, że $\mathrm{Lo}_+(\mathcal{H}) \subset H(G)$.

ii) Niech $S = T_{\mathbf{V}}$, gdzie $\det(\mathbf{V}) = \pm 1$. Ponieważ przy oznaczeniach twierdzenia 1 zachodzi $L \in \mathrm{Lo}_{+-}$ (patrz zadanie 1), więc z twierdzenia 1 i zadania 2 wynika, że $S \in \mathrm{Lo}_+(\mathcal{H})$, przy czym $S \in \mathrm{Lo}_{++} \Leftrightarrow \det(\mathbf{V}) = 1$.$\square$

Uwaga 1. * a) Należy odnotować, że $\ker(H) = \{\mathbf{I}, -\mathbf{I}\}$, tzn. przekształcenie $T_{\mathbf{V}}$ jest identycznością tylko gdy $\mathbf{V} = \pm\mathbf{I}$. Istotnie, jeśli $\mathbf{VAV}^* = \mathbf{A}$ dla wszystkich $\mathbf{A} \in \mathcal{H}$, to kładąc $\mathbf{A} = \mathbf{I}$ stwierdzamy, że $\mathbf{V}^* = \mathbf{V}^{-1}$. Zatem $\mathbf{VA} = \mathbf{AV}$ m.in. dla każdej macierzy diagonalnej, wobec czego $\mathbf{V} = \pm\mathbf{I}$.

b) Stąd wynika już łatwo, że dane przekształcenie Lorentza o dodatnim wyznaczniku jest postaci $T_{\mathbf{V}}$ dla dokładne dwóch, różniacych się znakiem, macierzy $\mathbf{V}$ (o wyznaczniku ± 1).

Uwaga 2. * $\mathrm{Lo}_+ = \mathrm{Lo}_{++} \cup \mathrm{Lo}_{+-}$ jest rozbiciem Lo_+ na warstwy względem Lo_{++}, a $\mathrm{Lo} = \mathrm{Lo}_{++} \cup \mathrm{Lo}_{+-} \cup \mathrm{Lo}_{-+} \cup \mathrm{Lo}_{--}$ rozbiciem Lo na warstwy względem Lo_{++}. (Wynika to łatwo zadania 1.) Wobec twierdzeń 1 i 2, zbiór Lo_{++} jest łukowo spójny, a w ślad za nim spójna łukowo jest każda z wymienionych warstw. Ponieważ są one rozłączne, więc grupa Lo_+ ma dwie składowe łukowe (są nimi Lo_{++} i Lo_{+-}), a grupa Lo ma ich 4. (Inaczej, informacje te wypowiedzieć można tak: dwa przekształcenia Lorentza wtedy i tylko wtedy połączyć można drogą takich przekształceń, gdy oba należą do tej samej z wymienionych warstw.)

Zadania uzupełniające.

1. Dowieść, że gdy $T \in \mathrm{Lo}_{++}$ i $S \in \mathrm{Lo}$, to $S^{-1}TS \in L_{++}$.

2. Gdy $\mathbf{A} \in \mathcal{H}$ i $\det(\mathbf{A}) = 1$, to $T_{\mathbf{V}}(\mathbf{A}) \in \{\mathbf{I}, -\mathbf{I}\}$ dla pewnej macierzy $\mathbf{V}$, będącej iloczynem macierzy unitarnej i rzeczywistej macierzy diagonalnej (obie o wyznaczniku 1).

Definicja. Przekształcenie $T \in \mathrm{Lo}_+(M)$ przestrzeni Minkowskiego M nazywamy

a) **obrotem hiperbolicznym**, jeśli istnieje istnieje płaszczyzna $P \subset M$, nie zawierająca wektora czasowego i taka, że T jest identycznością na P;

b) **obrotem euklidesowym**, jeśli $T(\mathbf{v}_0) = \mathbf{v}_0$ dla pewnego wektora czasowego $\mathbf{v}_0$. Prostą $\mathbb{R}\mathbf{v}_0$ nazywamy **osią czasową** tego obrotu.

3. a) Wywnioskować z ...., że każdy obrót euklidesowy przestrzeni Minkowskiego $\mathcal{H}$, z osią czasową $\mathbb{R}\mathbf{I}$, jest postaci $T_{\mathbf{V}}$ dla pewnej macierzy $\mathbf{V} \in \mathrm{SU}_2(\mathbb{C})$.

b) W oparciu o a) i zadanie ... udowodnić krok i) dowodu twierdzenia 2.

4. a) Dowieść, że gdy $\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle$ i płaszczyzna $\mathrm{lin}(\mathbf{u}, \mathbf{v})$ nie zawiera wektora czasowego, to $T(\mathbf{u}) = \mathbf{v}$ dla pewnego obrotu euklidesowego T.

b) Dowieść, że gdy $\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle < 0$, to $T(\mathbf{u}) = \pm\mathbf{v}$ dla pewnego obrotu hiperbolicznego T. (Wskazówka: sprowadzić do przypadku czasoprzestrzeni i $\mathbf{u} = (1, 0, 0, 0)$ oraz $v_2 = v_3 = 0$.)

5. Dowieść, że każdy operator $T \in \mathrm{Lo}_{++}$ jest złożeniem obrotu hiperbolicznego z euklidesowym.

6. Dowieść, że jeśli $\mathbf{V}$ jest macierzą unitarną (odp. samosprzężoną) i $\det(\mathbf{V}) = 1$, to $H(\mathbf{V})$ jest obrotem euklidesowym (odp. hiperbolicznym). Czy prawdziwe są implikacje przeciwne?

4. Interpretacja fizyczna.

Podać krótką interpr., ale z uwzględnieniem Trautmana: Jeśli obserwator 1 umieszczony w punkcie 0 czasoprzestrzeni robi zdjęcie wszechświata, to rztutuje wektory świetlne o ujemnej współrzędnej czasowej na sferyczną euklidesową „kliszę” $S = \{(1, x, y, z) : x^2 + y^2 + z^2 = 1\}$; rzutowanie zadane jest wzorem

$$P_0(t, x, y, z) = (1, x/t, y/t, z/t). \tag{6}$$

(Wyjaśnienie: S identyfikujemy ze zbiorem punktów oddalonych o jednostkę świetlną od $\mathbf{0}$ i dla $t < 0$ punktowi $p = (t, x, y, z)$ przyporządkowujemy punkt $(1, q)$ kliszy, jeśli promień świetlny $\mathbb{R}p$ wysłany z p przejdzie przez $\mathbf{0}$ i dotrze do q w chwili 1.) Analogicznie obserwator 2 rzutuje wektory świetlne na „swą" kliszę S. Wektory te zostały poddane pewnej transformacji Lorentza $L \in \mathrm{Lo}_{++}$. Wynika stąd łatwo, że obraz na kliszy drugiego obserwatora otrzymany jest z takiegoż obrazu pierwszego, po poddaniu go wpierw transformacji L, a następnie rzutowaniu P_0.

Najbardziej czytelny opis otrzymanego przekształcenia klisz $\widetilde{L} = P_0 L : S \to S$ otrzymuje się, gdy w naturalny sposób utożsamimy z $\mathbb{C}$ plaszczyznę dwóch ostatnich osi w $\mathbb{R}^4$; umożliwia to traktowanie $\mathbb{R}^4$ jako $\mathbb{R} \times \mathbb{R} \times \mathbb{C}$. Odpowiedniość (3) pomiędzy kwaternionami a macierzami samosprzężonymi jest teraz zastąpiona przez

$$(3') \qquad \mathbb{R} \times \mathbb{R} \times \mathbb{C} \ni (t, x, v) \leftrightarrow \begin{pmatrix} t + x & v \\ \overline{v} & t - x \end{pmatrix} \in \mathcal{H}.$$

(Odpowiedniość (3) otrzymujemy więc używając utożsamienia $v = z - yi$, zaś $v = y + iz$ prowadzi do odpowiedniości Θ z zadania ....6 w p.3.) Jak uprzednio, $H(\mathbf{V}) \in \mathrm{Lo}_{++}$ oznacza dla $\mathbf{V} \in \mathrm{SL}_2(\mathbb{C})$ przekształcenie odpowiadające przy (3') przekształceniu $\mathcal{H} \to \mathcal{H}$ zadanemu przez $\mathbf{A} \mapsto \mathbf{VAV}^*$.

Odnotujmy też, że sfera S jest obecnie interpretowana jako $\{(1, x, v) \in \mathbb{R} \times \mathbb{R} \times \mathbb{C} : x^2 + |v|^2 = 1\}$.

Twierdzenie 1. *Gdy* $\mathbf{V} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$ *i* $L = H(\mathbf{V})$, *to omawiana transformacja klisz* $\tilde{L} : S \to S$ *jest równa* $P^{-1}QP$, *gdzie:*

P *oznacza rzut stereograficzny z bieguna* $(1, 1, 0) \in \mathbb{R} \times \mathbb{R} \times \mathbb{C}$ *sfery* S *na* $\tilde{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ *(zadany jest on wzorem* $P(1, 1, 0) = \infty$ *oraz* $P(1, x, v) = v/(1 - x)$ *gdy* $x \neq 1$*; patrz .....),*

$Q : \tilde{\mathbb{C}} \to \tilde{\mathbb{C}}$ *jest przekształceniem homograficznym* $v \mapsto \frac{av+b}{cv+d}$.

Inaczej mówiąc, przy rzucie stereograficznym $P : S \to \tilde{\mathbb{C}}$ *przekształceniu klisz* $\tilde{L} : S \to S$ *odpowiada homografia* $\tilde{\mathbb{C}} \to \tilde{\mathbb{C}}$, *wyznaczona przez macierz* $\mathbf{V}$.

Wniosek 1. *Ponieważ zarówno* P, *jak i* Q *zachowują kąty,* [5] *więc i przekształcenie klisz ma tę własność. Podobnie, zbiór zaobserwowany jako okrąg przez pierwszego obserwatora jest okręgiem i dla drugiego.*

Innym ważnym wnioskiem jest to, że rozważane przekształcenia klisz $\tilde{L}$ zachowują pełną informację o indukującym je przekształceniu Lorentza L. Ma też miejsce naturalna $1 - 1$ odpowiedniość pomiędzy elementami Lo_{++} a homografiami: transformacji $H(\mathbf{V}) \in \mathrm{Lo}_{++}$ odpowiada homografia, wyznaczona przez macierz $\mathbf{V}$. (Korzystamy z

[5] patrz ....

tego, że zarówno surjekcja $H : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{Lo}_{++}$, jak i przyporządkowanie macierzy $\mathbf{V} \in \mathrm{SL}_2(\mathbb{C})$ wyznaczonej przez nią homografii, „zlepia" każdą macierz $\mathbf{V}$ z $-\mathbf{V}$ i nie wprowadza innych utożsamień.)

Dowód twierdzenia wykorzystuje następujący

Lemat 1. *Niech* $\mathbf{w} \in \mathbb{R} \times \mathbb{R} \times \mathbb{C}$ *będzie wektorem świetlnym, zaś* $\mathbf{A} \in \mathcal{H}$ *macierzą odpowiadającą mu przy (3'). Wówczas* $a_{12}/a_{22} = PP_0(\mathbf{w})$.

Dowód. Gdy $\mathbf{w} = (t, x, v)$, to $P_0(\mathbf{w}) = (1, x/t, v/t)$ i $PP_0(\mathbf{w}) = \frac{v/t}{1-x/t} = v/(t-x) = a_{12}/_{22}$. □

Dowód twierdzenia. Ustalmy wektor $\mathbf{w} = (1, x, v) \in S,\ \mathbf{w} \neq (1, 1, 0)$. Rozumowanie podzielimy na 3 części.

i) Macierz $\mathbf{A} \in \mathcal{H}$ odpowiadająca wektorowi $\mathbf{w}$ przy (3') ma wyraz a_{12} równy v, zaś wyraz a_{22} równy $1 - x \geq 0$; ponadto, jej wyznacznik jest równy $\langle \mathbf{w}, \mathbf{w} \rangle = 0$. Stąd $a_{11} = |v|^2/a_{22}$ oraz

$$\text{dla} \ \ s := 1 - x \ \ \text{mamy} \quad \mathbf{A} = \begin{pmatrix} |v|^2/s & v \\ \overline{v} & s \end{pmatrix} = \begin{pmatrix} v/\sqrt{s} \\ \sqrt{s} \end{pmatrix} \begin{pmatrix} v/\sqrt{s} \\ \sqrt{s} \end{pmatrix}^*.$$

ii) Z definicji, przy (3') wektorowi $L(\mathbf{w})$ odpowiada w $\mathcal{H}$ macierz

$$\mathbf{A}' := \mathbf{VAV}^* = \left(\mathbf{V} \begin{pmatrix} v/\sqrt{s} \\ \sqrt{s} \end{pmatrix}\right) \left(\mathbf{V} \begin{pmatrix} v/\sqrt{s} \\ \sqrt{s} \end{pmatrix}\right)^* = \begin{pmatrix} u \\ r \end{pmatrix} \begin{pmatrix} u \\ r \end{pmatrix}^*$$

gdzie $u = a\frac{v}{\sqrt{s}} + b\sqrt{s},\ r = c\frac{v}{\sqrt{s}} + d\sqrt{s}$.

iii) Stosując lemat otrzymujemy więc:

$$PP_0(L(\mathbf{w})) = \frac{a'_{12}}{a'_{22}} = \frac{u\overline{r}}{|r|^2} = \frac{u}{r} = \frac{a\frac{v}{s} + b}{c\frac{v}{s} + d} = Q(v/s) = Q(v/(1-x))$$

Ponieważ $\frac{v}{1-x} = P(\mathbf{w})$, więc dla $\mathbf{w} \neq (1, 1, 0)$ dowiedliśmy żądanej równości: $P(\tilde{L}(\mathbf{w})) = QP(\mathbf{w})$. Sprawdzenie jej gdy $\mathbf{w} = (1, 1, 0)$ pozostawione jest jako ćwiczenie. (Wskazówka: wówczas $\mathbf{A}' = \begin{pmatrix} u \\ r \end{pmatrix} \begin{pmatrix} u \\ r \end{pmatrix}^*$, gdzie $u = a, r = c$, co pozwala powtórzyć dowód.) □

1. Niech $\mathbf{A} = (g(\mathbf{v}_i, \mathbf{v}_j))_{i,j=1}^k$, gdzie g jest symetryczną formą dwuliniową na przestrzeni V, zaś $\mathbf{v}_1, ..., \mathbf{v}_k \in V$. Macierz nieosobliwą $\mathbf{C}$ taką, że macierz $\mathbf{C}^t\mathbf{AC}$ jest diagonalna, potraktujmy jako macierz przejścia od układu $\mathbf{v}_1, ..., \mathbf{v}_k$ do pewnego układu $\mathbf{w}_1, ..., \mathbf{w}_k$. (Oznacza to, że $\mathbf{w}_j = \sum_i c_{ij}\mathbf{v}_j$ dla $j = 1, ..., k$.) Dowieść, że:

a) Układ $\mathbf{w}_1, ..., \mathbf{w}_k$ jest g–ortogonalny i rozpina $\mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_k)$.

b) Gdy $\mathbf{C}$ otrzymano korzystając z algorytmu z §1.3, to $\mathbf{w}_1, ..., \mathbf{w}_k$ otrzymamy wykonując na układzie $\mathbf{v}_1, ..., \mathbf{v}_k$ kolejne operacje, wykonywane w algorytmie na wierszach macierzy $\mathbf{A}$.

c) Gdy ponadto nie wykonano części 1 żadnego z kroków algorytmu, to $\mathbf{w}_j \in \mathrm{lin}(\mathbf{v}_1, ..., \mathbf{v}_j)$ dla $j = 1, ..., k$. (Ma to miejsce, gdy $\mathbb{F} = \mathbb{R}$ i forma g jest półokreślona, patrz zad. uz. 1 w §2.5.)

2. Niech macierz $\mathbf{A} \in \mathcal{M}_{l,k}$ ma tę własność, że liczba jej minorów początkowych różnych od zera jest równa $\mathrm{rk}(\mathbf{A})$. Wówczas istnieje **rozkład Gaussa** tej macierzy: $\mathbf{A} = \mathbf{LDU}$, gdzie macierz $\mathbf{D} \in \mathcal{M}_{l,k}$ jest diagonalna, a $\mathbf{L}$ i $\mathbf{U}$ są macierzami kwadratowymi, na przekątnych których stoją same jedynki, przy czym macierz $\mathbf{L}$ jest trójkątna dolnie, a $\mathbf{U}$ górnie. (Wskazówka: dowód istnienia rozkładu Cholesky'ego.)

3. Zbadać jednoznaczność rozkładów Gaussa i Cholesky'ego.

Uwaga 1. (i definicja) Dzieki twierdzeniu 1 możemy w odniesieniu do symetrycznych funkcji dwuliniowych używać nazw przyjętych dla form kwadratowych: powiemy, taka funkcja, g, ma daną własność, gdy ma ją forma kwadratowa wyznaczona przez g. W szczególności, w przypadku rzeczywistego ciała skalarów możemy mówić o **określoności** czy **półokreśloności**, a także o **sygnaturze** funkcji g. Na podstawie ..... i ...., są takie, dla jak dla macierzy $(g(\mathbf{v}_i\mathbf{v}_j))_{i,j=1}^k$ funkcji g w dowolnej bazie $(\mathbf{v}_i)_{i=1}^k$.

§ 8. Kwadryki w $\mathbb{F}^k$ i ich afiniczna klasyfikacja.

1. Równania kwadratowe i ich zbiory rozwiązań.

Nazwijmy **równaniem kwadratowym** w $\mathbb{F}^k$ równanie postaci $q(x) = c$, gdzie $c \in \mathbb{F}$ i q jest wielomianem kwadratowym k zmiennych. Interesuje nas zbiór rozwiązań tego równania, tzn. zbiór

$$f^{-1}(c) := \{\mathbf{v} \in \mathbb{F}^k : \ q(\underline{\mathbf{v}}) = c\},$$

o którym powiemy, że jest **zadany** rozważanym równaniem $q(x) = c$. Odnotujmy, że w wielomianie q pewne ze zmiennych mogą nie występować, i dlatego zbiór $f^{-1}(c)$ zależy od tego, w której z przestrzeni $\mathbb{F}^k$ równanie $q(x) = c$ jest ono rozpatrywane (patrz końcowe zdania w §1.1).

Zbiór zadany równaniem kwadratowym w $\mathbb{F}^k$ nazywamy [6] **kwadryką** w $\mathbb{F}^k$. Łatwo jest opisać kwadryki w $\mathbb{F}$: są nimi podzbiory $\mathbb{F}$ liczące nie więcej, niż 2 elementy.

[6] Nasza terminologia odbiega tu od powszechnie przyjętej, zgodnie z którą zbiór $X \subset \mathbb{F}^k$ nazywany jest kwadryką tylko gdy $k = 3$ i X posiada pewne dodatkowe (w stosunku do rozważanych przez nas) własności. Wygodnie jest jednak dopuścić wszystkie wartości k, zaś dyskusję własności kwadryk oddzielić; por. początkowe zadania uzupełniające w końcowym punkcie tego rozdziału.

(Ćwiczenie: podać uzasadnienie!) Gdy $k > 1$, kwadryki w $\mathbb{F}^k$ są zbiorami bardziej skomplikowanymi: jak się przekonamy, dla $\mathbb{F} = \mathbb{R}$ i $k = 2$ należy je sobie na ogół wyobrażać jako pewne krzywe na płaszczyźnie $\mathbb{R}^2$, zaś dla $\mathbb{F} = \mathbb{R}$ i $k = 3$ jako pewne powierzchnie w przestrzeni $\mathbb{R}^3$. Podział kwadryk w $\mathbb{F}^k$ na klasy kwadryk w dużej mierze do siebie podobnych, ułatwiający opis dowolnej kwadryki w $\mathbb{F}^k$ –to zadanie, którym się teraz zajmiemy, uzyskując (na drodze algebraicznej) całkowite jego rozwiązanie dla $\mathbb{F} = \mathbb{R}$. Okaże się też, że nie poznaliśmy jeszcze pojęć pozwalających na pełną geometryczną interpretację otrzymanych rezultatów; stąd pewną wstępną taką interpretację podamy tylko dla $k = 2$.

Ustalmy równanie kwadratowe $q(x) = c$. By zbadać zadaną nim kwadrykę, naturalne jest starać się je uprościć stosując pewną afiniczną zamianę zmiennych. Sprawdzimy więc, czym jest zbiór rozwiązań w ten sposób zmienionego równania. Poniżej rozważamy afiniczną zamianę k zmiennych

$$(0) \qquad x = \mathbf{C}y + \mathbf{v}$$

i oznaczamy ją dla krótkości przez $x = a(y)$; literą a oznaczamy też przekształcenie $\mathbb{F}^k \to \mathbb{F}^k$ zadane przez $u \mapsto \mathbf{C}u + \mathbf{v}$.

Stwierdzenie 1. *Niech zbiór K będzie zadany równaniem $q(x) = c$ i niech afiniczna zamiana zmiennych (0) przeprowadza q w pewien wielomian g. Wówczas dla zbioru L zadanego równaniem $g(y) = c$ mamy*

$$K = a(L) \text{ oraz } L = a^{-1}(K),$$

gdzie $a^{-1}(x) = \mathbf{C}^{-1}x - \mathbf{C}^{-1}\mathbf{v}$ dla $x \in \mathbb{F}^k$.

Dowód. Z implikacji $(x = a(y)) \Rightarrow (q(x) = g(y))$ dla $x, y \in \mathbb{F}^k$ wynika, że $y \in K \Leftrightarrow a(y) \in L$. Wyprowadzenie wzoru na a^{-1} pozostawione jest jako proste ćwiczenie. $\boxminus$

Zwróćmy uwagę na charakterystyczne „odwrócenie kierunku”: jeśli zamiana zmiennych $x = a(y)$ przeprowadza jedno równanie kwadratowe w drugie, to zbiór zadany drugim równaniem jest przez odwzorowanie $a : \mathbb{F}^k \to \mathbb{F}^k$ przekształcany na zbiór zadany równaniem pierwotnym.

Stwierdzenie 1 uzasadnia wprowadzenie następujących pojęć:

Definicja. Przekształcenie $a : \mathbb{F}^k \to \mathbb{F}^k$ nazwiemy **afinicznym**, jeśli jest zadane wzorem

$$a(\mathbf{u}) = \mathbf{C}u + \mathbf{v} \text{ dla } u \in \mathbb{F}^k,$$

gdzie $\mathbf{C} \in \mathcal{M}_k$ jest ustaloną macierzą, zaś $\mathbf{v} \in \mathbb{F}^k$ ustalonym wektorem. Podzbiory K, L przestrzeni $\mathbb{F}^k$ nazwiemy **afinicznie równoważnymi**, jeśli istnieje różnowartościowe przekształcenie afiniczne $a : \mathbb{F}^k \to \mathbb{F}^k$ takie, że $L = a(K)$.

Zadanie 1. a) Przekształcenie afiniczne $a(\mathbf{u}) = \mathbf{C}u + \mathbf{v}$ wtedy i tylko wtedy jest różnowartościowe, gdy $\mathbf{C}$ jest macierzą nieosobliwą. (Z tego względu różnowartościowe przekształcenia afiniczne $\mathbb{F}^k \to \mathbb{F}^k$ nazywane są też **nieosobliwymi**.)

b) Nieosobliwe przekształcenia afiniczne $\mathbb{F}^k \to \mathbb{F}^k$ tworzą grupę przekształceń.

c) Afiniczna równoważność jest relacją równoważności w zbiorze podzbiorów przestrzeni $\mathbb{F}^k$.

d) Jeśli zbiory K i L w $\mathbb{F}^k$ są afinicznie równoważne i jeden z nich jest kwadryką, to drugi też nią jest.

e) Jeśli $a : \mathbb{F}^k \to \mathbb{F}^k$ jest przekształceniem afinicznym, to $a(cu + (1-c)\underline{\mathbf{v}}) = ca(\mathbf{u}) + (1-c)a(\underline{\mathbf{v}})$ dla wszelkich $u, \mathbf{v} \in \mathbb{F}^k$ i $c \in \mathbb{F}$.

Ze stwierdzenia 1 wynika natychmiast natępujący

Wniosek 1. *Jeśli $f, g : \mathbb{F}^k \to \mathbb{F}$ są afinicznie równoważnymi wielomianami kwadratowymi, to dla każdego $c \in \mathbb{F}$ kwadryka zadana równaniem $q(x) = c$ jest afinicznie równoważna kwadryce zadanej równaniem $g(x) = c$.*

W dalszej części interesować nas będzie przede wszystkim przypadek, gdy $\mathbb{F} = \mathbb{R}$. Wykorzystamy następujący

Lemat 1. *Dowolny wielomian $f \in \mathbb{R}_2[x_1, ..., x_k]$ jest afinicznie równoważny wielomianowi proporcjonalnemu do wielomianu f' postaci*

$$\text{(1)}\quad \begin{array}{l} \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 + \varepsilon x_k - c, \text{ gdzie } \ 0 \le p \le p+q < k \text{ oraz } \varepsilon, c \in \{0,1\}, \\ \text{a ponadto}: \ \varepsilon = 1 \Rightarrow (c = 0 \ \ i \ \ p+q < k), \ c = 0 \Rightarrow p \ge q. \end{array}$$

Ponadto, wielomian f' jest przez q jednoznacznie wyznaczony.

<u>Dowód.</u> Dowiedziemy wpierw istnienia f'. Możemy zakładać, że rozważamy wielomian q postaci (1) z wniosku 3 w §8.3. Jeśli $c = 0$, to żadnej zamiany zmiennych nie musimy już stosować i w przypadku $p < q$ piszemy $q = (-1)(-q))$. Natomiast dla $c \neq 0$ stosujemy zamianę zmiennych $\varepsilon x_k / = -x'_k$ i $x_i/\sqrt{|c|} = x'_i$ dla $i \le p+q$, przeprowadzającą q w wielomian f' spełniający wymagane warunki i taki, że $q = (-c)f'$.

Dla dowodu jednoznaczności ustalmy q i przyjmijmy $C = q(p)$ dla dowolnego środka symetrii wielomianu q, jeśli taki środek istnieje, oraz $C = 0$ jeśli nie istnieje. Jeśli wielomiany q i $\mu f'$ są afinicznie równoważne, gdzie $\mu \neq 0$ i f' jest postaci 1), to z lematu 1 w §7.5 wynika, iż:

a) $\varepsilon = 0 \Leftrightarrow f$ ma środki symetrii i $c = 1 \Leftrightarrow C \neq 0$, oraz

b) Jeśli $c = 1$, to μ jest przeciwnego znaku, co C.

Ponadto, przy $(p_0, q_0) := \sigma(q)$ wnosimy z b) i twierdzenia o bezwładności, że $(p, q) = (p_0, q_0)$ gdy ($C < 0$ lub $C = 0$ i $p_0 \ge q_0$), oraz $(p, q) = (q_0, p_0)$ w przeciwnym razie. Wraz z a) wyznacza to współczynniki wielomianu p.□

Twierdzenie 1. *Każda niepusta kwadryka w $\mathbb{R}^k$ jest afinicznie równoważna kwadryce zadanej równaniem $g(x) = 0$ dla wielomianu g będącego jednej z trzech następujących postaci:*

(2a)	$x_1^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_q^2 - 1$	*gdzie* $1 \leq p \leq p + q \leq k$,
(2b)	$x_1^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_q^2$	*gdzie* $1 \leq p \leq p + q \leq k$ *oraz* $p \geq q$,
(2c)	$x_1^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_q^2 + x_k$	*gdzie* $1 \leq p \leq p + q < k$ *oraz* $p \geq q$.

Dowód. Niech kwadryka, której afiniczną równoważność z pewną kwadryką opisanej postaci chcemy wykazać, będzie zadana równaniem $q(x) = 0$, dla pewnego wielomianu $f \in \mathbb{R}_2[x_1, ..., x_k]$. Wielomian q jest afinicznie równoważny wielomianowi μg, gdzie $\mu \in \mathbb{R} \setminus \{0\}$ oraz g jest postaci (1), dla pewnych p, q, ε, c. Zauważmy, że jeśli $p = 0$, to dla $c = 0$ wielomian g jest stopnia 0, zaś dla $c = 1$ mamy $\varepsilon = 0$ i zbiór zadany równaniem $g(x) = 0$ jest pusty. Ponieważ $f^{-1}(0) \neq \emptyset$ i $\deg(q) = 2$, więc z .... wynika, iż $p \geq 1$. Stąd postaci (2a)-(2c) odpowiadają temu, czy $c = 1$ (przypadek (2a)), czy $c = 0 = \varepsilon$ (przypadek (2b)), czy też wreszcie $\varepsilon = 1$ (przypadek (2c)).□

Do **kanonu afinicznego równań kwadryk** w $\mathbb{R}^k$ zaliczymy równanie $0 = 1$ oraz każde równanie $g(x) = 0$, gdzie g jest jednej z postaci (2a), (2b) lub (2c). Twierdzenie 1 orzeka, że każda kwadryka jest afinicznie równoważna pewnej kwadryce L wyznaczonej przez równanie należące do kanonu. Okazuje się, że taka kwadryka L oraz opisujące ją równanie kanoniczne są jednoznacznie wyznaczone:

Twierdzenie 2. *Różne równania kanonu afinicznego opisują w $\mathbb{R}^k$ afinicznie nierównoważne kwadryki.*

Dowód tego twierdzenia podany będzie w punkcie 3.

Twierdzenia 1 i 2 ustalają pewną listę kwadryk „afinicznie kanonicznych” w $\mathbb{R}^k$, taką, że każda kwadryka w $\mathbb{R}^k$ jest afinicznie równoważna dokładnie z jedną umieszczoną na liście. Dla $k = 2, 3$ listy te omówione są poniżej.

Problem 1. (kombinatoryczny) Przy zadanym k, ile jest kwadryk na liście?

Zadania uzupełniające.

1. Każda kwadryka w $\mathbb{C}^k$ jest afinicznie równoważna pewnej kwadryce zadanej równaniem

$$\sum_{i=1}^{r} x_i^2 + \varepsilon x_k = c, \text{ gdzie } 1 \leq r \leq k, \ \varepsilon, c \in \{0, 1\} \text{ oraz } (\varepsilon = 1) \Rightarrow (r < k \text{ i } c = 0). \tag{3}$$

2. Kwadryki K i L w $\mathbb{C}^k$ nazwiemy $\mathbb{R}$-afinicznie równoważnymi, jeśli $(x \in K) \Leftrightarrow (\mathbf{A}x + \mathbf{v} \in L)$ dla pewnego wektora $\mathbf{v}$ oraz nieosobliwej macierzy $\mathbf{A}$ mających wyrazy rzeczywiste.

a) Udowodnić, że każda kwadryka w $\mathbb{C}^k$ zadana równaniem kwadratowym o rzeczywistych współczynnikach jest $\mathbb{R}$-afinicznie równoważna kwadryce o równaniu postaci (1) (lub, co to samo oznacza, jednej z postaci (2a), (2b) lub (2c), lecz przy warunku $p \geq 1$ zmienionym na $p \geq 0$).

b) Wywnioskować, że jeśli K i L są kwadrykami w $\mathbb{C}^k$ zadanymi równaniami o współczynnikach rzeczywistych oraz kwadryki $K \cap \mathbb{R}^k$ i $L \cap \mathbb{R}^k$ są niepuste i afinicznie równoważne jako kwadryki w $\mathbb{R}^k$, to K i L są w $\mathbb{C}^k$ $\mathbb{R}$-afinicznie równoważne.

2. Kwadryki afinicznie kanoniczne w $\mathbb{R}^2$ i w $\mathbb{R}^3$.

Dla $k = 2, 3$ podamy teraz zapowiedziane listy afinicznie kanonicznych kwadryk i ich równań. (Dla skrócenia zapisu, niektóre równania nieznacznie zmieniamy, dodając 1 do obu stron.) Każda kwadryka w $\mathbb{R}^2$ lub w $\mathbb{R}^3$ jest, w myśl twierdzeń 1 i 2 z p.1, afinicznie równoważna z dokładnie jedną z kwadryk umieszczonych na listach. Gdy $k = 2$, zgodnie z tradycją piszemy dalej "**stożkowa**" zamiast „kwadryka."

Kanon afiniczny stożkowych:

Lp.	Równanie kanoniczne	Nazwa kanonicznej stożkowej
1	$x_1^2 + x_2^2 = 1$	okrąg
2	$x_1^2 + x_2^2 = 0$	punkt
3	$x_1^2 - x_2^2 = 1$	hiperbola
4	$x_1^2 - x_2^2 = 0$	para przecinających się prostych
5	$x_1^2 = 1$	para prostych równoległych
6	$x_1^2 = 0$	prosta
7	$x_1^2 + x_2 = 0$	parabola
8	$0 = 1$	zbiór pusty

Stożkową afinicznie równoważną okręgowi nazywamy **elipsą**; każdą inną – tak jak afinicznie jej równoważną stożkową kanoniczną. (Wobec twierdzenia 2 z p.1 nie prowadzi to do niejasności, np. żadna kwadryka nie jest i elipsą, i parabolą). W niektórych poręcznikach prosta i zbiór pusty nie są zaliczane do stożkowych, gdyż są zbiorami stopnia niższego niż 2; por. przypis na początku p.1. (Podobna uwaga dotyczy poniższej tabeli: płaszczyzna i zbiór pusty nie zawsze są traktowane jako kwadryki w $\mathbb{R}^3$, dla tej samej przyczyny.)

Ćwiczenie. Naszkicować w prostokątnym układzie współrzędnych $0x_1x_2$ stożkowe kanoniczne.

Zadanie 1. Ustalmy stałe dodatnie e, c i niech K będzie zbiorem tych punktów

$(r\cos\phi, r\sin\phi)$ płaszczyzny $\mathbb{R}^2$, dla których

$$|r(1-e\cos\phi)| = c.$$

Dowieść, że K jest elipsą gdy $e \in [0,1)$, parabolą gdy $e = 1$ i hiperbolą gdy $e > 1$. Ograniczjąc się do $r > 0$, przy $e \leq 1$ równanie $r(1-e\cos\phi) = c$ opisuje zbiór K, zaś przy $e > 1$ –jedno z dwóch ramion hiperboli.

Analogiczną listę kwadryk można podać przy $k = 3$. Do zagadnienia tego wrócimy jeszcze w jednym z następnych rozdziałów, w którym zbadamy pewne własności geometryczne stożkowych i kwadryk. Najprostszych z nich, wyjaśniających pochodzenie „wytłuszczonych" poniżej nazw, dotyczą dalsze zadania.

Kanon afiniczny kwadryk w $\mathbb{R}^3$

Lp.	Równanie kanoniczne	Nazwa kanonicznej kwadryki
1	$x_1^2 + x_2^2 + x_3^2 = 1$	sfera
2	$x_1^2 + x_2^2 + x_3^2 = 0$	punkt
3	$x_1^2 + x_2^2 - x_3^2 = 1$	**hiperboloida jednopowłokowa**
4	$x_1^2 + x_2^2 - x_3^2 = 0$	stożek obrotowy
5	$x_1^2 - x_2^2 - x_3^2 = 1$	**hiperboloida dwupowłokowa**
6	$x_1^2 + x_2^2 = 1$	walec nad okręgiem (obrotowy)
7	$x_1^2 + x_2^2 = 0$	prosta
8	$x_1^2 + x_2^2 + x_3 = 0$	**paraboloida obrotowa**
9	$x_1^2 - x_2^2 = 1$	walec nad hiperbolą (hiperboliczny)
10	$x_1^2 - x_2^2 = 0$	para płaszczyzn przecinających się
11	$x_1^2 - x_2^2 + x_3 = 0$	**paraboloida hiperboliczna**
12	$x_1^2 = 1$	para płaszczyzn równoległych
13	$x_1^2 = 0$	płaszczyzna
14	$x_1^2 + x_3 = 0$	walec nad parabolą (paraboliczny)
15	$0 = 1$	zbiór pusty

Kwadrykę afinicznie równoważną sferze nazywamy **elipsoidą**, stożkowi obrotowemu –**stożkiem eliptycznym**, walcowi obrotowemu –**walcem eliptycznym**, zaś paraboloidzie obrotowej –**paraboloidą eliptyczną**; każdą inną kwadryke w $\mathbb{R}^3$ nazywamy tak, jak afinicznie jej równoważną kwadrykę afinicznie kanoniczną.

Zadanie 2. Niech $K := \{(x,y,z) \in \mathbb{R}^3 : \; z = q(\sqrt{x^2+y^2})\}$. Udowodnić, że dla każdego $s \in \mathbb{R}$ przecięcie K z płaszczyzną $\Pi_s := \{(x,y,z) \in \mathbb{R}^3 : \; z = s\}$ jest sumą wszystkich okręgów w Π_s zawierających punkty krzywej $\{(x,0,q(x)) : \; x \geq 0\}$.

Zadanie to uzasadnia nazwanie K **powierzchnią obrotową** otrzymaną przez obrót krzywej $z = q(x)$ wokół osi z. (Analogicznie otrzymujemy równania powierzchni obrotowych otrzymanych przez obroty krzywych wokół innych osi.)

Zadanie 3. Kwadryki o numerach 1-6 oraz 8 są powierzchniami obrotowymi. Dla każdej z nich, rozpoznać obracaną krzywą i oś obrotu.

Zadanie 4. Naszkicować każdą z (niepustych) kwadryk kanonicznych. (W przypadku paraboloidy hiperbolicznej zacząć od narysowania jej przecięć z płaszczyznami $x_i = \varepsilon$, gdzie $i = 1, 2, 3$ oraz $\varepsilon = -1, 0, 1$.)

3. * Dalsze własności kwadryk i funkcji kwadratowych, pewne użyteczne stwierdzenie algebraiczne i dowód dowód twierdzenia o afinicznej klasyfikacji kwadryk.

Udowodnimy tu twierdzenie 2 z p.1 i w ten sposób zakończymy dowód twierdzenia o afinicznej klasyfikacji kwadryk w $\mathbb{R}^k$. (Tak na ogół nazywane są łącznie oba twierdzenia z p.1) Pomocne nam będzie pojęcie środka symetrii funkcji, zdefiniowane w §7.5 , oraz pojęcie prostej w $\mathbb{F}^k$.

Definicja. **Prostą** $\mathbb{F}^k$ przechodzącą przez punkty $a, b \in \mathbb{F}^k,\ a \neq b$ nazywamy zbiór $L_{ab} := \{ta + (1-t)b : a, b \in \mathbb{F}\}$.

Zadanie 1. Jeśli $a', b' \in L_{ab}$, to $a, b \in L_{a'b'}$ $(a \neq b, a' \neq b')$.

Uwaga 1. Niech niepusta kwadryka $K \subset \mathbb{R}^k$ będzie zadana równaniem $\sum_i \lambda_i x_i^2 + \varepsilon x_k = c$, gdzie $\lambda_1 \neq 0$ i $\varepsilon\lambda_k = 0$. Jeśli nie istnieje prosta przecinająca K w dokładnie 2 punktach, to $\varepsilon = 0$ i $\lambda_1\lambda_i \geq 0$ dla każdego i. (Istotnie, jeśi $\varepsilon = 0$ i $\lambda_1\lambda_i < 0$ dla pewnego i, to K jest przecięte w dwóch punktach przez pewną prosta leżącą w płaszczyźnie x_1x_i, np. przez prostą zadana równaniami $x_i = 1$ oraz $x_j = 0$ dla $j \notin \{1, i\}$), zaś gdy $\varepsilon \neq 0$, to przez prostą leżącą w płaszczyźnie x_1x_k).□

Zadanie 2. $0 = 1$ jest jedynym równaniem kanonu afinicznego, opisującym zbiór pusty.

Zasadniczą rolę w dalszym rozumowaniu odgrywa następujące

Stwierdzenie 1. *Niech $f, g \in \mathbb{R}_2[x_1, ..., x_k]$ mają tę własność, że $f^{-1}(0) \subset g^{-1}(0)$, przy czym istnieje prosta przecinająca zbiór $f^{-1}(0)$ w dokładnie dwóch punktach. Wówczas $g = \lambda f$ dla pewnego $\lambda \in \mathbb{R}$.*

Podany niżej błyskotliwy dowód oparty jest na podanym przez N. Kuipera. Udowodnimy wpierw

Lemat 1. *Niech K będzie kwadryką, zaś L prostą w $\mathbb{R}^k$ przechodzącą przez punkty p i q, $p \neq q$.*

a) Jeśli $L \cap K$ liczy więcej, niż 2 punkty, to $L \subset K$.

b) Jeśli $\#(L \cap K) = 2$ to istnieją otocznia N_p punktu p i N_q punktu q takie, że dla każdej prostej L' przecinającej zarówno N_p jak i N_q mamy $\#(L' \cap K) = 2$.

Dowód. Niech $f \in \mathbb{R}_2[x_1, ..., x_k]$ będzie wielomianem takim, że $K = f^{-1}(0)$. Warunek $tp+(1-t)q \in K$ jest spełniony, gdy t jest pierwiastkiem równania $q(tp+(1-t)q) = 0$ względem t. Jeśli równanie to ma więcej niż 2 pierwiastki, to funkcja kwadratowa $t \mapsto q(tp + (1-t)q)$ jest stała i równa zeru, skąd $L \subset K$.

Ad b) Omawiane równanie ma dwa pierwiastki jeśli jego wyróżnik $\Delta_{p,q}$ jest dodatni. Wyróżnik ten w sposób ciągły zależy od pary (p, q) (można go łatwo określić wzorem zależnym od współczynników funkcji q i od współrzędnych punktów p i q) i wobec tego jeśli $\Delta_{p,q} > 0$, to $\Delta_{p',q'} > 0$ dla (p', q') dostatecznie bliskich (p, q).□

Wniosek 1. *Jeśli q jest funkcją kwadratową i $q(x) = 0$ dla wszystkich x z pewnego zbioru otwartego U, to $q = 0$.*

Dowód. Obierzmy $p \in U$ i niech $q \in \mathbb{F}^k \setminus \{p\}$ będzie dowolnym punktem. Prosta L przechodząca przez p i przez q przecina U, a zatem i $f^{-1}(0)$, w nieskończenie wielu punktach. Mamy więc $q(x) = 0$ dla $x \in L$ i wobec tego $q(q) = 0$. □

Podamy teraz dowód stwierdzenia. Niech punkty p i q wyznaczają prostą L przecinającą zbiór $K := f^{-1}(0)$ w dokładnie 2 punktach; zmieniając p możemy zakładać, że $p \in L \setminus K$. Obierzmy otoczenia N_p i N_q jak w lemacie i przyjmijmy $h(x) := q(p)g(x) - g(p)q(x)$. Wówczas $h^{-1}(0) \supset f^{-1}(0) \cup \{p\}$, przy czym h jest funkcją kwadratową. Z konstrukcji wynika, że każda prosta przechodząca przez p i dowolny punkt $q' \in N_q$ przecina $f^{-1}(0)$ w dokładnie 2 punktach, a zatem $h^{-1}(0)$ w co najmniej trzech (dochodzi punkt p). Stąd każda taka prosta jest zawarta w $h^{-1}(0)$ i wobec tego $N_q \subset h^{-1}(0)$. Na mocy wniosku mamy więc $h = 0$, tzn. $q = \frac{g(p)}{q(p)} g$.□

Dowód twierdzenia 2 z p.1. Niech $q(x) = 0$ i $f'(x) = 0$ będą równaniami kanonu afinicznego kwadryk, takimi, że kwadryki $f^{-1}(0)$ i $(q')^{-1}(0)$ są afinicznie równoważne; dowiedziemy, że $q = f'$. Ze na względu zadanie 2 możemy założyć, że obie kwadryki są niepuste.

Oznaczmy przez a nieosobliwe przekształcenie afiniczne $\mathbb{R}^k \to \mathbb{R}^k$ takie, że $a(K) = K'$; wówczas funkcja $x \mapsto f'(a(x))$ zeruje się na K. Rozpatrzymy dwa przypadki.

i) Istnieje prosta przecinająca zbiór $f^{-1}(0)$ w dokładnie dwóch punktach. Ze stwierdzenia 1 wnosimy, że $f' \circ a = \lambda f$ dla pewnego $\lambda \in \mathbb{R} \setminus \{0\}$. Dla tego λ, wielomiany f' i λf są afinicznie równoważne. Inaczej mówiąc, wielomian q jest afinicznie równoważny z wielomianem proporcjonalnym do f', i oczywiście też z wielomianem proporcjonalnym do q. Ponieważ zarówno q, jak i f' są postaci (1) (opisanej w lemacie 1 z p.1), więc z jednoznaczności f' ustalonej w tym lemacie wynika, że $q = f'$

To kończy dyskusję i). Ze względu na uwagę 1 i symetrię pomiędzy q i f' pozostaje do rozpatrzenia przypadek, gdy

ii) $q = x_1^2 + ... + x_p^2$ i $f' = x_1^2 + ... + x_{p'}^2$ dla pewnych $p, p' \geq 1$. Wtedy $K = f^{-1}(0)$

i $K' = (q')^{-1}(0)$ są zbiorami rozwiązań bardzo prostych układów równań:

(4) $$K = \{(x_1, ..., x_k) \in \mathbb{R}^k : x_1 = ... = x_p = 0\} \text{ i } K' = \{(x_1, ..., x_k) \in \mathbb{R}^k : x_1 = ... = x_{p'}$$

Żądaną równość $p = p'$ najszybciej „zobaczyć" można używając pojęć, które wprowadzimy dopiero w następnym rozdziale: $k - p$ jest wymiarem pierwszego, zaś $k - p'$ wymiarem drugiego z tych zbiorów rozwiązań, i z istnienia niesosobliwego przekształcenia afinicznego K na K' wynika, że wymiary te są równe. Czytelnik chcący już teraz uporać się z przypadkiem ii) powinien rozwiązać poniższe zadanie. ⊟

Zadanie 3. Niech $a(x) = \mathbf{C}x + \mathbf{v}$ będzie nieosobliwym przekształceniem afinicznym i niech K i K' będą określone jak w (4).

a) Równość $a(K) = K'$ pozostanie niezmieniona przy a zastąpionym przez przekształcenie $x \mapsto \mathbf{C}x$.

b) Jeśli $k - p' < k - p$ i $a(K) = K'$, to oznaczając przez $\mathbf{B}$ klatkę wyznaczoną przez ostanich $k - p'$ wierszy i ostatnich $k - p$ kolumn macierzy $\mathbf{C}$ zauważyć, że $\mathbf{B}x = 0$ dla pewnego ciągu $x = (x_{p+1}, ..., x_k) \neq 0$ i że dla $\tilde{x} := (0, ..., 0, x_{p+1}, ..., x_k)$ mamy $\mathbf{C}\tilde{x} = 0$, wbrew nieosobliwości $\mathbf{C}$. (Wskazówka: $\mathbf{C}\tilde{x} \in K'$.)

Na koniec kilka wiadomości o środkach kwadryk.

Definicja. Powiemy, że punkt $p \in \mathbb{F}^k$ jest **środkiem symetrii zbioru** (lub: **środkiem zbioru**) $K \subset \mathbb{F}^k$ jeśli dla wszystkich $\mathbf{v} \in \mathbb{F}^k$ mamy $p + \mathbf{v} \in K \Leftrightarrow p - \mathbf{v} \in K$.

Przykład 1. a) Gdy $K = \emptyset$ lub $K = \mathbb{F}^k$, to każdy punkt $p \in \mathbb{F}^k$ jest środkiem zbioru $\mathbb{F}^k$;

b) Punkt p wtedy i tylko wtedy jest środkiem zbioru $K = \{x \in \mathbb{R}^k : x_1 = ... = x_r = 0\}$, gdy $p \in K$. ⊟

Lemat 2. *a) Punkt p wtedy i tylko wtedy jest środkiem zbioru $K \subset \mathbb{F}^k$, gdy dla wszystkich $w \in K$ mamy $2p - w \in K$.*

b) Jeśli p jest środkiem zbioru $K \subset \mathbb{R}^k$ i $a : \mathbb{F}^k \to \mathbb{F}^k$ jest nieosobliwym przekształceniem afinicznym, to $a(p)$ jest środkiem zbioru $a(K)$.

Dowód pozostawiamy jako ćwiczenie (w b) użyteczne jest zadanie 1 e) z p.1).

Stwierdzenie 2. *Wielomian kwadratowy $f \in \mathbb{R}_2[x_1, ..., x_k]$ i kwadryka $f^{-1}(0)$ mają ten sam zbiór środków symetrii.*

Dowód. Ponieważ jedna z inkluzji jest oczywista, więc pozostaje dowieść, że jeśli punkt p jest środkiem kwadryki $f^{-1}(0)$, to jest i środkiem symetrii wielomianu q. W tym celu przypuśćmy wpierw, że q jest wielomianem postaci $\sum_i^r \lambda_{i=1} x_i^2 + \varepsilon x_k + c$, gdzie $\varepsilon \neq 0 \Rightarrow (r < k \text{ i } c = 0)$. Jeśli nie istnieje prosta przecinająca K w dokładnie 2 punktach, to na mocy uwagi 1 mamy $q = \sum_i^r \lambda_{i=1} x_i^2$, gdzie wszystkie współczynniki $_1, ..., \lambda_r$ są różne od zera i wspólnego znaku. Stąd, jak stwierdzono w części b) przykładu, $p_{r+1} = ... = p_k = 0$ i p jest środkiem symetrii q (łatwy rachunek pomijamy). Jeśli natomiast omawiana prosta istnieje, to przyjmijmy $g(x) := q(2p - x)$; wówczas $g(x) = 0$ dla $x \in K$ i na mocy stwierdzenia 1 istnieje $\neq 0$ takie, że $q(2p - x) = \lambda q(x)$ da każdego x. Równoważnie, mamy wtedy $q(p - x) = \lambda q(p + x)$ dla każdego x i porównując współczynniki jak w dowodzie lematu 2 w §7.5 wnosimy, że $\varepsilon = 0$ i $p_i \lambda_i = 0$ dla każdego i, skąd p jest środkiem symetrii wielomianu q.

To kończy dowód lematu gdy q jest rozważanej wyżej postaci. W ogólnym przypadku oznaczamy przez $y = a(x)$ afiniczną zamianę zmiennych przeprowadzającą q w wielomian f' powyższej postaci. (Zamiana taka istnieje na mocy twierdzenia Lagrange'a). Przy $S(g)$ oznaczającym zbiór środków symetrii wielomianu g, zaś $T(g)$ –zbiór środków kwadryki $g^{-1}(0)$, mamy więc

$a(S(q')) = S(q)$ na podstawie lematu 2 a) w §7.5, oraz

$a(T(q')) = T(q)$ na podstawie stwierdzenia 1 w p.1 i lematu 2 a).

Ponieważ $S(q') = T(q')$, więc wynika stąd, że $S(q) = T(q)$.$\square$

Uwaga 2. Ze stwierdzenia 2 i zadania uzupełniającego 2 w §7.5 wynika więc, że punkt p wtedy i tylko wtedy jest środkiem kwadryki $f^{-1}(0) \subset \mathbb{R}^k$, gdy $\frac{\partial q}{\partial x_i}(p) = 0$ dla $i = 1, ..., k$.

<u>Ćwiczenie.</u> Gdy $\deg(q) = 1$, to q nie ma środków symetrii, lecz zbiór $f^{-1}(0)$ je ma.

4. * Stopień hiperpowierzchni; kwadryki w $\mathbb{C}^k$ (zadania uzupełniające).

Poniżej proponujemy Czytelnikowi zadania, rozwiązując które może łatwo wykorzystać uzyskane rezultaty (lub ich dowody) do ustalenia dalszych własności kwadryk i funkcji kwadratowych.

1. Niech K będzie zbiorem rozwiązań układu równań liniowych w $\mathbb{R}^k$. Wówczas

a) K jest kwadryką.

b) Gdy $K \neq \emptyset$, to zbiór środków K jest równy K. (Jest tak i przy $\mathbb{R}$ zastąpionym dowolnym ciałem $\mathbb{F}$.)

c) Jeśli $K \neq \emptyset$, to K jest zbiorem afinicznie równoważnym kwadryce opisanej równaniem postaci $x_1^2 + ... + x_p^2 = 0$. (Wskazówka: dla zbioru opisanego innym równaniem kanonicznym teza b) jest fałszywa.)

Nazwijmy zbiór $K \subset \mathbb{F}^k$ **hiperpowierzchnią** w $\mathbb{F}^k$, jeśli $K = f^{-1}(0)$ dla pewnego wielomianu q zmiennych $x_1, ..., x_k$. **Stopniem** hiperpowierzchni K nazywamy liczbę

$$\mathrm{nul}(K) := \min\{\deg(q) : K = f^{-1}(0) \text{ i } f \in \mathbb{F}[x_1, ..., x_k]\}.$$

W szczególności, $\mathbb{F}^k$ jest hiperpowierzchnią w $\mathbb{F}^k$ stopnia -1 (bo $\mathbb{F}^k = f^{-1}(0)$ dla $q = 0$), zaś zbiór pusty hiperpowierzchnią stopnia 0 ((bo $\emptyset = f^{-1}(0)$ dla $q = 1$); są to jedyne hiperpowierzchnie w $\mathbb{F}^k$ stopni -1 i 0. Z definicji wynika więc, że niepuste kwadryki w $\mathbb{F}^k$ są hiperpowierzchniami stopni 1 lub 2.

2. a) Każda hiperpowierzchnia stopnia 1 w $\mathbb{R}^k$ jest afinicznie równoważna zbiorowi $K_0 := \{(u_1, ..., u_k) \in \mathbb{R}^k : u_1 = 0\}$. (Wskazówka: jeśli $K = g^{-1}(0)$ i $g = b_1x_1 + ... + b_kx_k + c$, gdzie $b_1 \neq 0$, to zastosować przekształcenie $u \mapsto (g(u), u_2, ..., u_k)$.)

b) Powyższy zbiór K_0, zadany równaniem $x_1^2 = 0$, jest jedyną kwadryką kanoniczną stopnia 1. (Wskazówka: część a) i twierdzenie klasyfikacyjne dla kwadryk.)

3. (Por. problem w p.1). Wyznaczyć liczbę wielomianów każdej z postaci (2a)-(2c).

4. Jeśli kwadryka $K \subset \mathbb{C}^k$ spełnia założenia uwagi 1, z $\mathbb{R}^k$ zastąpionym przez $\mathbb{C}^k$, to jest ona zadana równaniem $x_1^2 = 0$.

5. a) Niech $f^{-1}(0) \subset g^{-1}(0)$, gdzie $f, g : \mathbb{R}^k \to \mathbb{R}$ są funkcjami kwadratowymi. Wówczas albo $q = \lambda g$ dla pewnego $\lambda \in \mathbb{R}$, albo $f^{-1}(0)$ jest zbiorem rozwiązań pewnego układu równań liniowych.

b) Sformułować i udowodnić odpowiednik powyższej tezy przy $\mathbb{R}^k$ zastąpionym przez $\mathbb{C}^k$. (Wskazówka: poprzedzające zadanie.)

6. Stwierdzenie 2 pozostaje prawdziwe przy $\mathbb{R}$ zastąpionym przez $\mathbb{C}$.

7. (Por. zadanie uzupełniające 1 w p.1.) Dwa różne równania postaci (3) zadają w $\mathbb{C}^k$ kwadryki, które nie są afinicznie równoważne (nad $\mathbb{C}$).

8. (Por. zadanie uzupełniające 2 w p.1.) Dwa różne równania postaci (1) zadają w $\mathbb{C}^k$ kwadryki, które nie są $\mathbb{R}$-afinicznie równoważne.

5. Nieco heurezy (czyli rozważań przedwczesnych).

Przypuśćmy, że rozważane dwie kwadryki K i L zadane są równaniami kwadratowymi $q(x) = c$ oraz $g(x) = c$, odpowiednio, i że afiniczna zamiana zmiennych $x = \mathbf{C}y + \mathbf{v}$ przeprowadza funkcję q w g. Jaką interpretację geometryczną możemy temu nadać? Zajmiemy się tu przede wszystkim przypadkami, gdy $\mathbb{F} = \mathbb{R}$ oraz $k = 2$ lub $k = 3$. Wyobraźnia geometryczna, której możemy wówczas użyć, pomoże nam wskazać, jak w dalszej części należy w podobny sposób rozważać kwadryki w przestrzeniach bardziej ogólnych.