Rozdział 3

PRZESTRZENIE LINIOWE Z ILOCZYNEM SKALARNYM

Teorię form dwuliniowych rozwinęliśmy w rozdziale 1 częściowo w tym celu, by móc przejść od ogólnych przestrzeni liniowych do struktur bardziej konkretnych, a także bliższych naszemu doświadczeniu — przestrzeni liniowych wyposażonych w metrykę. Przypomnijmy, że bogactwo geometrii trójwymiarowej wiąże się w dużym stopniu z dwoma dodatkowymi pojęciami algebry wektorów — długością wektora i kątem między wektorami. Przejście od czysto jakościowego pojęcia liniowości, dla którego ciało skalarów jest w dużym stopniu obojętne, do związków o charakterze ilościowym wymusza na nas ograniczenie się w praktyce do dwóch ciał: $\mathbb{R}$ i $\mathbb{C}$. Geometria przestrzeni zespolonych zasługuje na szczególne potraktowanie zarówno ze względu na swoją ważność, jak i na konieczność wprowadzenia form nowego typu.

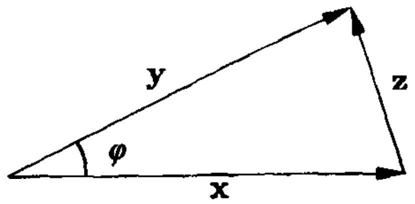
§ 1. PRZESTRZENIE EUKLIDESOWE

1. Rozważania heurystyczne i definicje. W geometrii analitycznej przestrzeni $\mathbb{R}^2$ i $\mathbb{R}^3$ ważną rolę odgrywa pojęcie iloczynu skalarnego wektorów, zdefiniowanego jako iloczyn ich długości i kosinusa kąta między nimi. Długość $\|\mathbf{x}\|$ wektora $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ w prostokątnym układzie współrzędnych jest określona wzorem

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2},$$

w szczególności $\|\mathbf{x}\| > 0$, jeśli $\mathbf{x} \neq \mathbf{0}$. Kwadrat długości wektora $\|\mathbf{x}\|^2 = x_1^2+x_2^2+x_3^2$ można interpretować jako wartość dodatnio określonej formy kwadratowej. Odpowiadająca jej forma dwuliniowa $(* \mid *)$ przypisuje wektorom $\mathbf{x} = x_1\mathbf{e}_1+x_2\mathbf{e}_2+x_3\mathbf{e}_3$ i $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3$ liczbę $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3$. Niech $\varphi = \sphericalangle(\mathbf{x}, \mathbf{y})$ będzie kątem między wektorami $\mathbf{x}$ i $\mathbf{y}$, a $\mathbf{z} = \mathbf{y} - \mathbf{x}$ ich różnicą (rys. 3). Wzór kosinusów z geometrii elementarnej, zastosowany do trójkąta o bokach $\mathbf{x}, \mathbf{y}, \mathbf{z}$, stwierdza, że

$$\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}^2\| - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\varphi.$$



Rys. 3

Z drugiej strony, wykorzystując dwuliniowość i symetryczność formy $(*\,|\,*)$, otrzymujemy

$$\|\mathbf{z}\|^2 = \|\mathbf{y}-\mathbf{x}\|^2 = (\mathbf{y}-\mathbf{x}\,|\,\mathbf{y}-\mathbf{x}) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2(\mathbf{x}\,|\,\mathbf{y}).$$

Porównanie dwóch wyrażeń na $\|\mathbf{z}\|^2$ daje

$$(\mathbf{x}\,|\,\mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\varphi,$$

tj. liczba $(\mathbf{x}\,|\,\mathbf{y})$ jest zwykłym iloczynem skalarnym wektorów $\mathbf{x}$ i $\mathbf{y}$. Ta okoliczność sugeruje przyjęcie następującej definicji iloczynu skalarnego wektorów w $\mathbb{R}^n$:

$$(\mathbf{x}\,|\,\mathbf{y}) = \sum_{i=1}^{n} x_i y_i. \tag{1}$$

Wadą tej definicji jest jej uzależnienie od wyboru układu współrzędnych. Aby pozbyć się tego defektu, przyjmujemy następującą ogólną definicję:

DEFINICJA 1. *Przestrzenią euklidesową* nazywamy rzeczywistą przestrzeń liniową V z wyróżnioną symetryczną formą dwuliniową $(\mathbf{x},\mathbf{y}) \mapsto (\mathbf{x}\,|\,\mathbf{y})$ o tej własności, że odpowiadająca jej forma kwadratowa $\mathbf{x} \mapsto (\mathbf{x}\,|\,\mathbf{x})$ jest dodatnio określona. Wartość $(\mathbf{x}\,|\,\mathbf{y})$ nazywamy *iloczynem skalarnym* wektorów $\mathbf{x}$ i $\mathbf{y}$.

Oznaczenie $(\mathbf{x}\,|\,\mathbf{y})$ (zamiast poprzednio używanego $f(\mathbf{x},\mathbf{y})$) wprowadziliśmy po to, by podkreślić, że z nieskończonego zbioru form symetrycznych wybraliśmy jedną i na niej budujemy strukturę danej przestrzeni euklidesowej. Używając iloczynu skalarnego $(\mathbf{x}\,|\,\mathbf{y})$, zamierzamy wprowadzić pojęcia długości wektora i kąta między wektorami; wynikającą stąd niejednoznaczność tych pojęć można porównać do dowolności związanej z wyborem skali przy mierzeniu odcinków na prostej. Iloczyn skalarny oznacza się często przez $(\mathbf{x},\mathbf{y})$ lub $\langle\mathbf{x},\mathbf{y}\rangle$, u nas jednak oba te oznaczenia są już zajęte: $(\mathbf{x},\mathbf{y})$ to para wektorów (element iloczynu kartezjańskiego $V \times V$), a $\langle\mathbf{x},\mathbf{y}\rangle$ jest podprzestrzenią rozpiętą na wektorach $\mathbf{x}$ i $\mathbf{y}$.

Powtórzmy zatem: przestrzeń euklidesowa to para $(V,(*\,|\,*))$, gdzie V jest przestrzenią liniową nad $\mathbb{R}$, a $(*\,|\,*)$ — ustaloną symetryczną dodatnio określoną formą dwuliniową na V, zwaną iloczynem skalarnym. Oto główne własności iloczynu skalarnego:

(i) $(\mathbf{x}\,|\,\mathbf{y}) = (\mathbf{y}\,|\,\mathbf{x})\ \forall \mathbf{x},\mathbf{y} \in V$;
(ii) $(\alpha\mathbf{x}+\beta\mathbf{y}\,|\,\mathbf{z}) = \alpha(\mathbf{x}\,|\,\mathbf{z}) + \beta(\mathbf{y}\,|\,\mathbf{z})\ \forall \alpha,\beta \in \mathbb{R}$;
(iii) $(\mathbf{x}\,|\,\mathbf{x}) > 0\ \forall \mathbf{x} \neq \mathbf{0}\ ((\mathbf{0}\,|\,\mathbf{x}) = 0)$.

Iloczyn skalarny w $\mathbb{R}^n$ określony wzorem (1) (zwany *standardowym* iloczynem skalarnym na $\mathbb{R}^n$) spełnia oczywiście powyższe warunki — w przeciwnym razie nasza ogólna definicja nie miałaby sensu.

Przykład 1. Niech $V = P_n$ będzie przestrzenią liniową wielomianów rzeczywistych stopni $\leqslant n-1$, a $[a, b] \subset \mathbb{R}$ — ustalonym przedziałem domkniętym. Z własności całek wynika, że przypisując dwóm wektorom (wielomianom) $f, g \in P_n$ liczbę

$$(f \mid g) = \int_a^b f(t)g(t)\,dt, \tag{2}$$

określamy iloczyn skalarny na V. Wyrażanie tego iloczynu skalarnego w terminach naturalnej bazy $(1, t, \ldots, t^{n-1})$ byłoby bardzo niewygodne. Warto zauważyć, że ten sam wzór (2) określa iloczyn skalarny w nieskończenie wymiarowej przestrzeni $C[a, b]$ funkcji ciągłych na $[a, b]$ o wartościach rzeczywistych. Otrzymaną w ten sposób nieskończenie wymiarową przestrzeń euklidesową będziemy oznaczać przez $CL^2[a, b]$.

2. Podstawowe pojęcia metryczne. Niech V będzie przestrzenią euklidesową z iloczynem skalarnym $(* \mid *)$.

DEFINICJA 2. *Długością* lub *normą* $\|\mathbf{v}\|$ wektora $\mathbf{v} \in V$ nazywamy liczbę rzeczywistą nieujemną

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v} \mid \mathbf{v})}. \tag{3}$$

Ponieważ $(\mathbf{v} \mid \mathbf{v}) \geqslant 0$, więc długość dowolnego wektora jest określona, przy czym $\mathbf{v} \neq \mathbf{0} \Rightarrow \|\mathbf{v}\| > 0$. Jeśli $\lambda \in \mathbb{R}$, to $\|\lambda\mathbf{v}\| = \sqrt{(\lambda\mathbf{v} \mid \lambda\mathbf{v})} = |\lambda| \cdot \|\mathbf{v}\|$.

Zauważmy w tym miejscu, że dowolna podprzestrzeń liniowa U przestrzeni euklidesowej V jest również przestrzenią euklidesową, gdyż obcięcie formy $(* \mid *)$ do U wyznacza formę dwuliniową $U \times U \to \mathbb{R}$, która jest oczywiście nadal symetryczna i dodatnio określona. W szczególności samo ciało $\mathbb{R}$ jest jednowymiarową przestrzenią euklidesową, w której długością wektora jest wartość bezwzględna liczby rzeczywistej. W ogólnym przypadku symbole $| * |$ i $\| * \|$ będą przez nas używane w różnych znaczeniach.

Wektor o długości 1 nazywamy *unormowanym*. Każdy wektor $\mathbf{x} \neq \mathbf{0}$ można unormować, mnożąc go przez odpowiedni skalar: mianowicie dla wektora $\mathbf{x}' = \frac{1}{\|\mathbf{x}\|}\mathbf{x}$ mamy

$$\|\mathbf{x}'\| = \left\| \frac{1}{\|\mathbf{x}\|}\mathbf{x} \right\| = \frac{1}{\|\mathbf{x}\|}\|\mathbf{x}\| = 1.$$

Zanim zdefiniujemy kąt między wektorami, wrócimy jeszcze raz do własności (iii) iloczynu skalarnego.

TWIERDZENIE 1 (nierówność Schwarza). *Dla dowolnych wektorów* $\mathbf{x}$ *i* $\mathbf{y}$ *przestrzeni euklidesowej* V *zachodzi nierówność*

$$|(\mathbf{x}\,|\,\mathbf{y})| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|. \tag{4}$$

Dowód. Z dodatniej określoności iloczynu skalarnego (własność (iii)) wynika, że dla dowolnej liczby rzeczywistej λ zachodzi nierówność

$$\lambda^2(\mathbf{x}\,|\,\mathbf{x}) - 2\lambda(\mathbf{x}\,|\,\mathbf{y}) + (\mathbf{y}\,|\,\mathbf{y}) = (\lambda\mathbf{x} - \mathbf{y}\,|\,\lambda\mathbf{x} - \mathbf{y}) \geqslant 0. \tag{5}$$

Dla ustalonych $\mathbf{x}, \mathbf{y} \in V$ lewa strona nierówności (5) jest trójmianem kwadratowym $f(\lambda)$. Ponieważ $f(\lambda) \geqslant 0$ dla każdego $\lambda \in \mathbb{R}$, więc wyróżnik trójmianu jest niedodatni: $\mathrm{D}(f) = (2(\mathbf{x}\,|\,\mathbf{y}))^2 - 4(\mathbf{x}\,|\,\mathbf{x}) \cdot (\mathbf{y}\,|\,\mathbf{y}) \leqslant 0$, skąd

$$(\mathbf{x}\,|\,\mathbf{y})^2 \leqslant (\mathbf{x}\,|\,\mathbf{x}) \cdot (\mathbf{y}\,|\,\mathbf{y}). \tag{6}$$

Wyciągając dodatni pierwiastek z obu stron i wykorzystując definicję (3) długości wektora, otrzymujemy (4). ■

Uwaga. Równość $|(\mathbf{x}\,|\,\mathbf{y})| = \|\mathbf{x}\|\,\|\mathbf{y}\|$ oznacza, że $\mathrm{D}(f) = 0$, czyli trójmian f ma (jeden) pierwiastek rzeczywisty λ_0. Wtedy z (5) wynika, że $(\lambda_0\mathbf{x} - \mathbf{y}\,|\,\lambda_0\mathbf{x} - \mathbf{y}) = 0$, czyli $\mathbf{y} = \lambda_0\mathbf{x}$. Tak więc iloczyn skalarny wektorów jest równy (co do wartości bezwzględnej) iloczynowi ich norm dokładnie wtedy, gdy wektory te są współliniowe (proporcjonalne).

Przykład 2. Zapiszmy nierówność (4) dla standardowego iloczynu skalarnego w $\mathbb{R}^n$:

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2} \tag{7}$$

oraz dla iloczynu skalarnego (2) w przestrzeni $CL^2[a,b]$ (rzeczywistych) funkcji ciągłych:

$$\left|\int_a^b f(t)g(t)\,dt\right| \leqslant \sqrt{\int_a^b f(t)^2\,dt} \cdot \sqrt{\int_a^b g(t)^2\,dt}. \tag{8}$$

Nierówność (8) odgrywa ważną rolę w analizie.

Nierówność Schwarza oznacza, że jeśli $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$, to

$$-1 \leqslant \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \leqslant 1.$$

Wynika stąd, że iloraz $(\mathbf{x}\,|\,\mathbf{y})/(\|\mathbf{x}\| \cdot \|\mathbf{y}\|)$ jest kosinusem ściśle określonego kąta φ:

$$\cos\varphi = \frac{(\mathbf{x}\,|\,\mathbf{y})}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}, \qquad 0 \leqslant \varphi \leqslant \pi. \tag{9}$$

Ten kąt właśnie, na mocy definicji, uznajemy za *kąt między wektorami* $\mathbf{x}$ i $\mathbf{y}$.

DEFINICJA 3. Wektory $\mathbf{x}$ i $\mathbf{y}$ nazywamy *ortogonalnymi* lub *prostopadłymi* (oznaczenie: $\mathbf{x} \perp \mathbf{y}$), jeśli kąt między nimi wynosi $\pi/2$, tj. $(\mathbf{x} \mid \mathbf{y}) = 0$.

Wektor zerowy jest ortogonalny do każdego wektora. Zauważmy ponadto, że

$$\mathbf{x} \perp \mathbf{y} \;\Rightarrow\; \|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

(*twierdzenie Pitagorasa*), przy czym u nas ten elementarny fakt geometryczny jest wnioskiem z formalnych własności iloczynu tensorowego. Nieco bardziej ogólne jest analogiczne stwierdzenie dla większej liczby wektorów: jeśli wektory $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \ldots$ są parami ortogonalne, to

$$\|\mathbf{x}+\mathbf{y}+\mathbf{z}+\mathbf{u}+\ldots\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \|\mathbf{z}\|^2 + \|\mathbf{u}\|^2 + \ldots$$

Jako ćwiczenie Czytelnik zechce sprawdzić, że

$$\|\mathbf{x}\| = \|\mathbf{y}\| \;\Rightarrow\; \mathbf{x}+\mathbf{y} \perp \mathbf{x}-\mathbf{y}$$

(przekątne rombu są prostopadłe).

Z twierdzenia 1 wynika

WNIOSEK (nierówność trójkąta). *Długości wektorów* $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ *i* $\|\mathbf{x} \pm \mathbf{y}\|$ *spełniają nierówność*

$$\|\mathbf{x} \pm \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \tag{10}$$

Dowód. Istotnie, z nierówności Schwarza wynika, że

$$\begin{aligned}\|\mathbf{x} \pm \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \pm 2(\mathbf{x} \mid \mathbf{y}) \leqslant \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2|(\mathbf{x} \mid \mathbf{y})| \\ &\leqslant \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \; \blacksquare\end{aligned}$$

Przykład 3. W przestrzeni funkcyjnej $CL^2[a,b]$ nierówność trójkąta przyjmuje postać

$$\sqrt{\int_a^b (f(t) \pm g(t))^2\,dt} \leqslant \sqrt{\int_a^b f(t)^2\,dt} + \sqrt{\int_a^b g(t)^2\,dt}$$

(*nierówność Minkowskiego*).

3. Ortogonalizacja. W przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym (1) wektory $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)$ (z jedynką na i-tym miejscu) są parami ortogonalne i stanowią bazę. Mamy prawo spodziewać się, że w każdej przestrzeni euklidesowej można znaleźć bazę o podobnych własnościach. Sformułujmy dokładnie nasze oczekiwania wobec takiej bazy.

DEFINICJA 4. Bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni euklidesowej V nazywamy *ortogonalną*, jeśli $(\mathbf{e}_i \,|\, \mathbf{e}_j) = 0$ dla dowolnych $i \neq j$ $(i, j = 1, \ldots, n)$. Jeśli ponadto $(\mathbf{e}_i \,|\, \mathbf{e}_i) = 1$ dla każdego i, to bazę nazywamy *ortonormalną*.

Inaczej mówiąc, wszystkie wektory bazy ortonormalnej mają długość 1. Z każdej bazy ortogonalnej można otrzymać bazę ortonormalną, normując wektory bazowe.

Następujący fakt jest niemal oczywisty:

TWIERDZENIE 2. *Jeśli niezerowe wektory* $\mathbf{e}_1, \ldots, \mathbf{e}_m \in V$ *są parami ortogonalne, to są liniowo niezależne. Jeśli więc* $m = \dim V$, *to wektory takie stanowią bazę ortogonalną w* V.

Dowód. Drugie stwierdzenie wynika z definicji bazy. Aby udowodnić pierwsze, przypuśćmy, że

$$\alpha_1 \mathbf{e}_1 + \ldots + \alpha_m \mathbf{e}_m = \mathbf{0}$$

jest nietrywialną zależnością liniową; niech np. $\alpha_k \neq 0$. Mnożąc skalarnie obie strony powyższej równości przez $\mathbf{e}_k$, otrzymujemy

$$\begin{aligned} 0 &= (\mathbf{0} \,|\, \mathbf{e}_k) = (\alpha_1 \mathbf{e}_1 + \ldots + \alpha_m \mathbf{e}_m \,|\, \mathbf{e}_k) \\ &= \alpha_1 (\mathbf{e}_1 \,|\, \mathbf{e}_k) + \ldots + \alpha_k (\mathbf{e}_k \,|\, \mathbf{e}_k) + \ldots + \alpha_m (\mathbf{e}_m \,|\, \mathbf{e}_k) = \alpha_k (\mathbf{e}_k \,|\, \mathbf{e}_k), \end{aligned}$$

gdyż z założenia $(\mathbf{e}_i \,|\, \mathbf{e}_k) = 0$ dla $i \neq k$. Z drugiej strony, $(\mathbf{e}_k \,|\, \mathbf{e}_k) \neq 0$, a więc $\alpha_k = 0$; sprzeczność ta kończy dowód. ■

Twierdzenie 2 będzie wykorzystywane przy konstrukcji bazy ortogonalnej, ale samo jej istnienie można otrzymać bezpośrednio:

TWIERDZENIE 3. *W każdej n-wymiarowej przestrzeni euklidesowej V istnieje baza ortonormalna.*

Dowód. Forma kwadratowa $q(\mathbf{x}) = (\mathbf{x} \,|\, \mathbf{x}) = \|\mathbf{x}\|^2$ na V jest dodatnio określona. Istnieje więc baza $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ w V, w której forma ta ma postać kanoniczną

$$q(\mathbf{x}) = x_1^2 + \ldots + x_n^2$$

dla $\mathbf{x} = x_1 \mathbf{e}_1 + \ldots + x_n \mathbf{e}_n$ (rozdz. 1, § 4, p. 8). Iloczynem skalarnym odpowiadającym tej formie kwadratowej jest (1). Oznacza to jednak, że $(\mathbf{e}_i \,|\, \mathbf{e}_j) = \delta_{ij}$, czyli baza $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest ortonormalna. ■

Zauważmy, że w bazie ortonormalnej współrzędne wektora $\mathbf{x}$ są równe iloczynom skalarnym tego wektora przez wektory bazy:

$$(\mathbf{x} \,|\, \mathbf{e}_i) = x_i.$$

DEFINICJA 5. Jeśli wektor $\mathbf{e}$ ma długość 1, to iloczyn skalarny $(\mathbf{x} \mid \mathbf{e})$ nazywamy *współrzędną wektora* $\mathbf{x}$ *na osi* $\langle \mathbf{e} \rangle_{\mathbb{R}}$.

Tak więc współrzędne wektora $\mathbf{x}$ w bazie ortonormalnej $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ są jego współrzędnymi na osiach $\langle \mathbf{e}_i \rangle_{\mathbb{R}}$.

Efektywna konstrukcja bazy ortonormalnej odbywa się za pomocą tzw. *ortogonalizacji Grama–Schmidta*, która to procedura występuje w rozmaitych działach analizy i geometrii. Zauważmy najpierw, że zbiór wszystkich wektorów ortogonalnych do danego wektora $\mathbf{v}$ stanowi podprzestrzeń liniową, zwaną *dopełnieniem ortogonalnym* wektora $\mathbf{v}$. Istotnie, jeśli $\mathbf{x} \perp \mathbf{v}$ i $\mathbf{y} \perp \mathbf{v}$, tj. $(\mathbf{x} \mid \mathbf{v}) = (\mathbf{y} \mid \mathbf{v}) = 0$, to również

$$(\alpha \mathbf{x} + \beta \mathbf{y} \mid \mathbf{v}) = \alpha(\mathbf{x} \mid \mathbf{v}) + \beta(\mathbf{y} \mid \mathbf{v}) = 0 \qquad \forall \alpha, \beta \in \mathbb{R}.$$

Mówimy też, że wektor $\mathbf{v}$ jest *ortogonalny do podprzestrzeni* $U \subset V$, jeśli $\mathbf{v} \perp \mathbf{u}$ dla każdego $\mathbf{u} \in U$. Jeśli $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ jest bazą w U, to oczywiście $\mathbf{v} \perp U \Leftrightarrow \mathbf{v} \perp \mathbf{e}_i$ dla $i = 1, \ldots, m$. Zbiór wszystkich wektorów ortogonalnych do podprzestrzeni U stanowi również podprzestrzeń liniową.

DEFINICJA 6. Podprzestrzeń liniową złożoną z wszystkich wektorów ortogonalnych do podprzestrzeni liniowej $U \subset V$ oznaczamy przez $U^{\perp}$ i nazywamy *dopełnieniem ortogonalnym* podprzestrzeni U w V.

TWIERDZENIE 4 (o ortogonalizacji Grama–Schmidta). *Niech* $\mathbf{e}_1, \ldots, \mathbf{e}_m$ *będzie układem wektorów liniowo niezależnych w przestrzeni euklidesowej* V. *Wtedy istnieje układ ortonormalny* $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ *taki, że* $\langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle = \langle \mathbf{e}'_1, \ldots, \mathbf{e}'_i \rangle$ *dla każdego* $i \leqslant m$.

Dowód. Dla $i = 1$ weźmy $\mathbf{e}'_1 = \lambda \mathbf{e}_1$, gdzie $\lambda = \|\mathbf{e}_1\|^{-1}$. Wtedy $\langle \mathbf{e}_1 \rangle = \langle \mathbf{e}'_1 \rangle$, czyli teza zachodzi dla $i = 1$. Przypuśćmy, że dla pewnego k $(1 \leqslant k < m)$ skonstruowaliśmy już układ ortonormalny $\mathbf{e}'_1, \ldots, \mathbf{e}'_k$ w taki sposób, że $L_i = L'_i$ dla $i = 1, \ldots, k$, gdzie $L_i := \langle \mathbf{e}_1, \ldots, \mathbf{e}_i \rangle$ i $L'_i := \langle \mathbf{e}'_1, \ldots, \mathbf{e}'_i \rangle$. Pokażemy, jak znaleźć wektor $\mathbf{e}'_{k+1}$.

Rozważmy wektor

$$\mathbf{v} = \mathbf{e}_{k+1} - \sum_{i=1}^{k} \lambda_i \mathbf{e}'_i,$$

gdzie λ_i są (na razie dowolnymi) skalarami. Wektor ten jest zawsze niezerowy, ponieważ $\mathbf{e}_{k+1} \notin L'_k = L_k$. Ponadto dla dowolnych λ_i zachodzi równość

$$\langle \mathbf{e}'_1, \ldots, \mathbf{e}'_k, \mathbf{v} \rangle = \langle \mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1} \rangle$$

(gdyż każdy wektor wypisany po lewej stronie równości należy do przestrzeni po prawej stronie i na odwrót). Dobierzemy liczby λ_i tak, by wektor $\mathbf{v}$ był ortogonalny do L'_k. Jak wiemy, wystarczy zapewnić ortogonalność do wektorów bazowych:

$$0 = (\mathbf{v} \,|\, \mathbf{e}'_j) = (\mathbf{e}_{k+1} \,|\, \mathbf{e}'_j) - \Big(\sum_{i=1}^{k} \lambda_i \mathbf{e}'_i \,\Big|\, \mathbf{e}'_j \Big)$$

$$= (\mathbf{e}_{k+1} \,|\, \mathbf{e}'_j) - \sum_{i=1}^{k} \lambda_i (\mathbf{e}'_i \,|\, \mathbf{e}'_j) = (\mathbf{e}_{k+1} \,|\, \mathbf{e}'_j) - \lambda_j, \qquad j = 1, \ldots, k.$$

Widać, że jeśli przyjmiemy $\lambda_j = (\mathbf{e}_{k+1} \,|\, \mathbf{e}'_j)$, to wektor $\mathbf{v} \neq \mathbf{0}$ będzie ortogonalny do L'_k. Definiujemy $\mathbf{e}'_{k+1} = \mu \mathbf{v}$, gdzie $\mu = \|\mathbf{v}\|^{-1}$; wtedy wektory $\mathbf{e}'_1, \ldots, \mathbf{e}'_{k+1}$ są ortonormalne i spełniają warunek $L'_{k+1} = L_{k+1}$, co kończy indukcyjny dowód twierdzenia. ∎

Wniosek. *Każdy układ ortonormalny w skończenie wymiarowej przestrzeni euklidesowej V można uzupełnić do bazy ortonormalnej tej przestrzeni.*

Dowód. Mając układ ortonormalny $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ w V, możemy go uzupełnić do bazy $(\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n)$ w V (rozdz. 1, § 2, twierdzenie 3). Wystarczy teraz przeprowadzić dla tej bazy procedurę ortogonalizacji Grama–Schmidta, zaczynając od wektora $\mathbf{e}_{m+1}$ (tzn. nie zmieniając wektorów $\mathbf{e}_1, \ldots, \mathbf{e}_m$; zob. poprzedni dowód). ∎

Rozumowanie podobne do procedury ortogonalizacji Grama–Schmidta prowadzi do dowodu kolejnego twierdzenia:

Twierdzenie 5. *Niech L będzie skończenie wymiarową podprzestrzenią liniową przestrzeni euklidesowej V, a $L^{\perp}$ – jej dopełnieniem ortogonalnym. Wtedy*

$$V = L \oplus L^{\perp}, \qquad L^{\perp\perp} = L. \tag{11}$$

Dowód (1). Niech $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ będzie bazą ortonormalną w L. Dla dowolnego wektora $\mathbf{v} \in V$ definiujemy

$$\mathbf{u} = \sum_{i=1}^{m} (\mathbf{v} \,|\, \mathbf{e}_i) \mathbf{e}_i, \qquad \mathbf{w} = \mathbf{u} - \mathbf{v}.$$

Wtedy oczywiście $\mathbf{v} = \mathbf{u} + \mathbf{w}$, przy czym $\mathbf{u} \in L$. Pokażemy, że $\mathbf{w} \in L^{\perp}$. Istotnie,

$$(\mathbf{w} \,|\, \mathbf{e}_j) = \sum_{i=1}^{m} (\mathbf{v} \,|\, \mathbf{e}_i)(\mathbf{e}_i \,|\, \mathbf{e}_j) - (\mathbf{v} \,|\, \mathbf{e}_j) = (\mathbf{v} \,|\, \mathbf{e}_j) \cdot 1 - (\mathbf{v} \,|\, \mathbf{e}_j) = 0, \qquad j = 1, \ldots, m.$$

Tak więc $V = L + L^{\perp}$.

---

(1) Sformułowanie twierdzenia i jego dowód zostały nieznacznie przeredagowane. W szczególności nie zakładamy, że przestrzeń V jest skończenie wymiarowa; będzie to wykorzystane w § 5 (*przyp. tłum.*).

Niech $\mathbf{x} \in L \cap L^{\perp}$. Ponieważ $\mathbf{x} \in L$, więc $\mathbf{x} \perp L^{\perp}$; ale również $\mathbf{x} \in L^{\perp}$, więc $\mathbf{x} \perp \mathbf{x}$, czyli $\mathbf{x} = \mathbf{0}$. Zatem suma $L + L^{\perp}$ jest prosta: $V = L \oplus L^{\perp}$.

Z rozkładu $\mathbf{v} = \mathbf{u} + \mathbf{w}$ ($\mathbf{u} \in L$, $\mathbf{w} \in L^{\perp}$) wynika, że jeśli $\mathbf{v} \in L^{\perp\perp} = (L^{\perp})^{\perp}$, to $0 = (\mathbf{v} \,|\, \mathbf{w}) = (\mathbf{u} + \mathbf{w} \,|\, \mathbf{w}) = (\mathbf{w} \,|\, \mathbf{w})$, czyli $\mathbf{w} = \mathbf{0}$, a więc $\mathbf{v} = \mathbf{u} \in L$. Wobec tego $L^{\perp\perp} \subset L$. Zawieranie w drugą stronę jest oczywiste: $L \perp L^{\perp}$, zatem $L \subset (L^{\perp})^{\perp}$ ([1]). ∎

4. Izomorfizmy przestrzeni euklidesowych. Widzieliśmy, że odpowiedni wybór bazy ortonormalnej w skończenie wymiarowej przestrzeni euklidesowej V pozwala na zapisanie iloczynu skalarnego w standardowej postaci (1). Fakt ten oznacza w istocie, że ze względu na swoje własności metryczne przestrzenie V i $\mathbb{R}^n$ są nierozróżnialne. Dokładniej wyraża to

TWIERDZENIE 6. *Dowolne dwie przestrzenie euklidesowe V i V' tego samego skończonego wymiaru są izomorficzne jako przestrzenie euklidesowe, tzn. istnieje izomorfizm $f : V \to V'$ przestrzeni liniowych* (zob. definicję w rozdziale 1, § 2, p. 3), *zachowujący iloczyn skalarny:*

$$(\mathbf{x} \,|\, \mathbf{y}) = (f(\mathbf{x}) \,|\, f(\mathbf{y}))', \tag{12}$$

gdzie $(* \,|\, *)'$ oznacza iloczyn skalarny w V'.

Dowód. Rozpatrzmy (dowolne) bazy ortonormalne $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ w V oraz $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ w V'. Odpowiedniość

$$f : \mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n \mapsto \mathbf{x}' = x_1\mathbf{e}'_1 + \ldots + x_n\mathbf{e}'_n$$

jest oczywiście izomorfizmem przestrzeni liniowych. Ponieważ w V i V' iloczyny skalarne $(\mathbf{x} \,|\, \mathbf{y})$ i $(\mathbf{x}' \,|\, \mathbf{y}')'$ oblicza się według tego samego wzoru (1) (na mocy ortonormalności obu baz), warunek (12) jest również spełniony. ∎

Twierdzenie to pozwala przetłumaczyć na język geometrii elementarnej każde stwierdzenie sformułowane w języku działań na wektorach z V i iloczynu skalarnego. Na odwrót: każde twierdzenie o charakterze metrycznym, odnoszące się do obiektów przestrzeni $\mathbb{R}^3$ lub $\mathbb{R}^2$, pozostaje prawdziwe w dowolnej przestrzeni euklidesowej V tego samego wymiaru.

Skoro już mówimy o izomorfizmach, rozważmy przestrzeń dualną (sprzężoną) V^* do przestrzeni euklidesowej V w sensie definicji z rozdziału 1 (§ 3). Dla każdego ustalonego wektora $\mathbf{v} \in V$ odwzorowanie $\mathbf{x} \mapsto (\mathbf{x} \,|\, \mathbf{v})$ określa oczywiście formę

([1]) Warto zwrócić uwagę, iż z twierdzenia 5 wynika, że rozkład wektora $\mathbf{v} \in L$ na sumę $\mathbf{u} + \mathbf{w}$, gdzie $\mathbf{u} \in L$ i $\mathbf{w} \in L^{\perp}$, jest jednoznaczny, w szczególności wektor $\mathbf{u}$ nie zależy od wyboru bazy ortonormalnej w L. Odwzorowanie $\mathcal{P} : V \to V$, $\mathbf{v} \mapsto \mathbf{u}$, nazywamy *rzutem ortogonalnym* przestrzeni V na podprzestrzeń L; inaczej mówiąc, jest to operator rzutowania na L równolegle do $L^{\perp}$ (*przyp. tłum.*).

liniową

$$\Phi_{\mathbf{v}} = (* \,|\, \mathbf{v}) : V \to \mathbb{R},$$

tj. $(* \,|\, \mathbf{v}) \in V^*$.

TWIERDZENIE 7. *Odwzorowanie $\Phi : \mathbf{v} \mapsto (* \,|\, \mathbf{v}) = \Phi_{\mathbf{v}}$ jest monomorfizmem przestrzeni liniowej V w jej przestrzeń dualną V^*. W szczególności, jeśli przestrzeń V jest skończenie wymiarowa, to Φ jest izomorfizmem. Przy tym izomorfizmie każda baza ortonormalna $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni V przechodzi na dualną do niej bazę $(e^1, \ldots, e^n)$ przestrzeni V^*.*

Dowód. Ponieważ iloczyn skalarny $(\mathbf{x} \,|\, \mathbf{v})$ jest liniowy względem $\mathbf{v}$, więc odwzorowanie Φ jest liniowe:

$$\Phi_{\alpha\mathbf{u}+\beta\mathbf{v}} = (* \,|\, \alpha\mathbf{u} + \beta\mathbf{v}) = \alpha(* \,|\, \mathbf{u}) + \beta(* \,|\, \mathbf{v}) = \alpha\Phi_{\mathbf{u}} + \beta\Phi_{\mathbf{v}}.$$

Ponadto $\operatorname{Ker}\Phi = \{\mathbf{0}\}$, ponieważ jeśli $\mathbf{v} \in \operatorname{Ker}\Phi$, to $(\mathbf{x} \,|\, \mathbf{v}) = 0$ dla każdego $\mathbf{x} \in V$ i w szczególności $(\mathbf{v} \,|\, \mathbf{v}) = 0$, czyli $\mathbf{v} = \mathbf{0}$.

Jeśli przestrzeń V jest skończenie wymiarowa i $(\mathbf{e}_i)$ jest dowolną jej bazą ortonormalną, to podobnie jak każdy element przestrzeni V^*, forma liniowa $(* \,|\, \mathbf{v})$ wyraża się liniowo przez wektory $e^1, \ldots, e^n \in V^*$ bazy dualnej do $(\mathbf{e}_i)$. W szczególności

$$\Phi_{\mathbf{e}_j} = (* \,|\, \mathbf{e}_j) = \sum_{k=1}^{n} a_{kj} e^k, \qquad j = 1, \ldots, n.$$

Baza $(\mathbf{e}_i)$ jest jednak ortonormalna, więc

$$\delta_{ij} = (\mathbf{e}_i \,|\, \mathbf{e}_j) = \sum_{k=1}^{n} a_{kj} e^k(\mathbf{e}_i) = \sum_{k=1}^{n} a_{kj}\delta_{ki} = a_{ij},$$

a stąd

$$(* \,|\, \mathbf{e}_j) = e^j, \qquad j = 1, \ldots, n. \; \blacksquare \tag{13}$$

Tak więc w skończenie wymiarowej przestrzeni euklidesowej V każdy wektor $\mathbf{v} \in V$ można uważać jednocześnie za formę liniową $\mathbf{v} : V \to \mathbb{R}$. Przy tym utożsamieniu każda baza ortonormalna w V jest dualna sama do siebie. Będziemy wykorzystywać ten naturalny izomorfizm (porównaj go z izomorfizmem $V \cong V^{**}$ dla zwykłych przestrzeni skończenie wymiarowych) przy badaniu operatorów liniowych. Izomorfizm Φ jest izomorfizmem przestrzeni euklidesowych w sensie twierdzenia 6, jeśli określimy w V^* iloczyn skalarny wzorem

$$((* \,|\, \mathbf{u}) \,|\, (* \,|\, \mathbf{v}))^* := (\mathbf{u} \,|\, \mathbf{v})$$

(Czytelnik zechce sprawdzić, że wszystkie aksjomaty iloczynu skalarnego są spełnione).

5. Bazy ortonormalne i macierze ortogonalne. Ponieważ bazy ortonormalne w przestrzeni euklidesowej pełnią szczególnie ważną rolę, warto się przyjrzeć regułom transformacji przy przejściu od bazy ortonormalnej $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ do innej takiej bazy $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$. Jak zwykle pisząc

$$\mathbf{e}'_j = a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \ldots + a_{nj}\mathbf{e}_n, \qquad j = 1, \ldots, n, \tag{14}$$

otrzymamy macierz przejścia $A = (a_{ij})$, której k-ta kolumna składa się ze współrzędnych wektora $\mathbf{e}'_k$ w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Na razie przepisaliśmy tylko wzory z rozdziału 1 (§ 2, (3)) i jedyne ograniczenie na A to $\det A \neq 0$. Wykorzystamy teraz ortonormalność baz:

$$\delta_{ij} = (\mathbf{e}'_i \mid \mathbf{e}'_j) = \Big(\sum_k a_{ki}\mathbf{e}_k \,\Big|\, \sum_l a_{lj}\mathbf{e}_l\Big) = \sum_{k,l} a_{ki}a_{lj}(\mathbf{e}_k \mid \mathbf{e}_l) = \sum_k a_{ki}a_{kj}.$$

Zatem

$$a_{1i}a_{1j} + a_{2i}a_{2j} + \ldots + a_{ni}a_{nj} = \begin{cases} 0 & \text{dla } i \neq j, \\ 1 & \text{dla } i = j. \end{cases} \tag{15}$$

Wprowadzając macierz transponowaną, można krótko zapisać (14) (lub (15)) w postaci

$${}^tA \cdot A = E, \tag{16}$$

czyli $A^{-1} = {}^tA$. Ale $A^{-1}A = E \Rightarrow AA^{-1} = E$, więc również

$$A \cdot {}^tA = E, \tag{16'}$$

co prowadzi do zależności

$$a_{i1}a_{j1} + a_{i2}a_{j2} + \ldots + a_{in}a_{jn} = \begin{cases} 0 & \text{dla } i \neq j, \\ 1 & \text{dla } i = j. \end{cases} \tag{15'}$$

DEFINICJA 7. Rzeczywistą macierz kwadratową $A = (a_{ij})$ spełniającą jeden z równoważnych warunków (15), (15′), (16) lub (16′) nazywamy *macierzą ortogonalną* ([1]). Zbiór wszystkich macierzy ortogonalnych stopnia n oznaczamy przez $\mathrm{O}(n)$.

Bezpośrednio można sprawdzić (wrócimy jeszcze do tej kwestii), że $\mathrm{O}(n)$ jest grupą. Nazywamy ją *grupą ortogonalną*. Jeśli A jest macierzą ortogonalną i $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest bazą ortonormalną, to układ $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ określony wzorami (14) jest też bazą ortonormalną.

Doszliśmy zatem do następującego wniosku:

TWIERDZENIE 8. *Macierz przejścia od jednej bazy ortonormalnej do drugiej jest ortogonalna; na odwrót, każda macierz ortogonalna jest taką macierzą przejścia.* ∎

([1]) Warto zwrócić uwagę na interpretację geometryczną warunków (15) i (15′): np. (15) oznacza, że kolumny macierzy A są ortonormalne jako wektory przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym (*przyp. tłum.*).

Zauważmy, że ze wzorów (14) i z ortonormalności wektorów bazowych można otrzymać interpretację geometryczną elementów a_{ij} macierzy przejścia A. Mianowicie

$$a_{ij} = (\mathbf{e}_i \,|\, \mathbf{e}'_j) = \cos\varphi_{ij}, \tag{17}$$

gdzie φ_{ij} jest kątem między starym wektorem bazowym $\mathbf{e}_i$ a nowym wektorem bazowym $\mathbf{e}'_j$.

Przechodząc do wyznaczników w równości (16), otrzymujemy $(\det A)^2 = 1$, czyli wyznacznik macierzy ortogonalnej jest równy 1 lub -1.

Z rozdziału 1 (§ 2) znamy regułę transformacji współrzędnych wektora przy zmianie bazy: jeśli $\mathbf{x} = \sum_i x_i \mathbf{e}_i = \sum_i x'_i \mathbf{e}'_i$, to

$$x_i = \sum_j a_{ij} x'_j, \qquad i = 1, \ldots, n.$$

Teraz wiemy jednak ponadto, że $A^{-1} = {}^tA$, więc

$$x'_i = \sum_j a_{ji} x_j, \qquad i = 1, \ldots, n.$$

6. Przestrzenie symplektyczne. Pojęcie iloczynu skalarnego ma wiele wariantów. Możemy np. rozpatrzyć (jak w § 4 rozdziału 1) niezdegenerowaną antysymetryczną formę dwuliniową na skończenie wymiarowej przestrzeni V. Formę taką, często oznaczaną przez $[\mathbf{x} \,|\, \mathbf{y}]$, nazywa się *skośnym iloczynem skalarnym* lub *formą symplektyczną*. Parę $(V, [* \,|\, *])$ nazywamy *przestrzenią symplektyczną*. Jak wiemy (rozdz. 1, § 4, twierdzenie 9), wymiar przestrzeni V jest wtedy parzysty, $\dim V = 2m$, a w odpowiedniej bazie (też zwanej *symplektyczną*) macierz formy dwuliniowej $[* \,|\, *]$ ma standardową postać J (lub J_0). Twierdzenie to można przeformułować następująco: *wszystkie przestrzenie symplektyczne tego samego wymiaru są izomorficzne.*

Podobnie jak w przypadku przestrzeni euklidesowych [1], naturalne jest rozpatrzenie grupy operatorów liniowych na V, zachowujących formę symplektyczną:

Definicja 8. Jeśli V jest przestrzenią symplektyczną, to operator liniowy $\mathcal{A} : V \to V$ nazywamy *symplektycznym*, jeśli

$$[\mathcal{A}\mathbf{x} \,|\, \mathcal{A}\mathbf{y}] = [\mathbf{x} \,|\, \mathbf{y}] \qquad \forall \mathbf{x}, \mathbf{y} \in V = \mathbb{R}^{2m}.$$

Zbiór wszystkich takich operatorów jest grupą; nazywamy ją *grupą symplektyczną* przestrzeni V i oznaczamy przez $\mathrm{Sp}(V)$.

Macierz A operatora symplektycznego $\mathcal{A}$ w bazie symplektycznej spełnia warunek

$${}^tA \cdot J_0 \cdot A = J_0 = \begin{Vmatrix} 0 & -E_m \\ E_m & 0 \end{Vmatrix}.$$

---

[1] Zob. § 3, p. 2 (*przyp. tłum.*).

Macierze takie, również zwane symplektycznymi, tworzą grupę symplektyczną $\mathsf{Sp}(2m)$, izomorficzną z grupą $\mathrm{Sp}(V)$. Z powyższego wzoru wynika, że $\det A = \pm 1$; w rzeczywistości $\det A = 1$, co wynika bezpośrednio z twierdzenia 10 w rozdziale 1 (§ 4, p. 10):

$$1 = \mathrm{Pf}(J_0) = \mathrm{Pf}({}^tA \cdot J_0 \cdot A) = (\det A)\,\mathrm{Pf}(J_0) = \det A.$$

O tym, że zbiór $\mathrm{Sp}(2m)$ jest rzeczywiście grupą, czyli jest zamknięty względem zwykłych działań $(A, B) \mapsto AB$ i $A \mapsto A^{-1}$, można się też łatwo przekonać bezpośrednio.

Dla $m = 1$ mamy izomorfizm $\mathrm{Sp}(2) \cong \mathrm{SL}_2(\mathbb{R})$ (macierze o wyznaczniku 1). Istotnie, dla $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ równość

$${}^t\begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix} \cdot \begin{Vmatrix} 0 & -1 \\ 1 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix} = \begin{Vmatrix} 0 & -(\alpha\delta - \beta\gamma) \\ \alpha\delta - \beta\gamma & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & -1 \\ 1 & 0 \end{Vmatrix}$$

jest równoważna warunkowi $\alpha\delta - \beta\gamma = 1$. ∎

W punkcie 5 widzieliśmy, że macierze ortogonalne to dokładnie macierze przejścia od jednej bazy ortonormalnej do drugiej. Analogicznie macierz przejścia od jednej bazy symplektycznej do drugiej jest macierzą symplektyczną i każda macierz symplektyczna jest taką macierzą przejścia.

Widmo macierzy symplektycznej (lub operatora symplektycznego) ma szereg interesujących własności:

TWIERDZENIE 9. *Zachodzi implikacja*

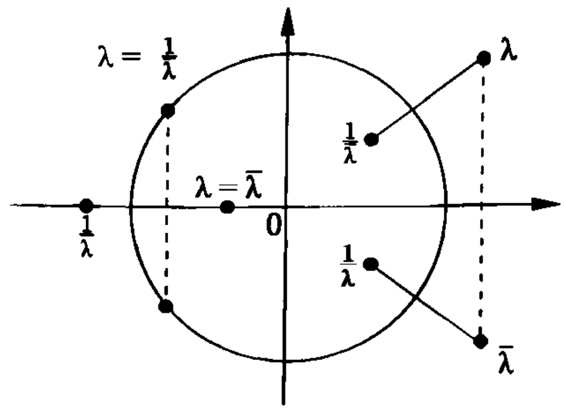
$$(\mathcal{A} \in \mathrm{Sp}(V) \text{ i } \dim V = 2m) \Rightarrow \chi_{\mathcal{A}}(t) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right).$$

Innymi słowy, $\chi_{\mathcal{A}}(t) = \sum_{i=0}^{2m} a_i t^i$ *jest wielomianem zwrotnym:* $a_i = a_{2m-i} \in \mathbb{R}$ *dla* $i = 1, 2, \ldots$ *W szczególności jeśli* $\lambda \in \mathbb{C}$ *jest pierwiastkiem charakterystycznym operatora symplektycznego, to są nimi również liczby* $1/\lambda$, $\overline{\lambda}$ *i* $1/\overline{\lambda}$.

Dowód. Wiemy, że każda macierz ma ten sam wielomian charakterystyczny co jej macierz transponowana oraz że wielomiany charakterystyczne macierzy podobnych są równe. Z równości ${}^tAJ_0A = J_0$ wynika, że $A^{-1} = J_0^{-1}\,{}^tAJ_0$. Zatem $\chi_{A^{-1}}(t) = \chi_A(t)$. Wystarczy teraz zauważyć, że ponieważ $\det A = 1$, więc

$$\begin{aligned} \chi_{A^{-1}}(t) &= \det(tE - A^{-1}) \\ &= \det(tE - A^{-1})\det A = \det(tA - E) = \det(E - tA) \\ &= t^{2m}\det\left(\frac{1}{t}E - A\right) = t^{2m}\chi_A\left(\frac{1}{t}\right). \text{ ∎} \end{aligned}$$

Widzimy, że zbiór pierwiastków charakterystycznych operatora symplektycznego jest symetryczny nie tylko względem osi x (jako zbiór pierwiastków wielo-



Rys. 4

mianu rzeczywistego), ale również względem okręgu $|z| = 1$, tzn. jest niezmienniczy względem inwersji $\lambda \mapsto 1/\overline{\lambda}$ (rys. 4).

Niekiedy w przestrzeni symplektycznej V ze standardową strukturą J_0 (w wybranej bazie):

$$[\mathbf{x}\,|\,\mathbf{y}] = \sum_{i=1}^{m}(x_{i+m}y_i - x_i y_{i+m}),$$

wprowadza się stowarzyszoną (choć nie kanonicznie, tzn. w sposób zależny od wyboru bazy) strukturę euklidesową

$$(\mathbf{x}\,|\,\mathbf{y}) = \sum_{k=1}^{2m} x_k y_k.$$

Dla formy $[\mathbf{x}\,|\,\mathbf{y}]$, podobnie jak dla każdej formy dwuliniowej (zob. § 3), istnieje operator liniowy $\mathcal{J}$ taki, że

$$[\mathbf{x}\,|\,\mathbf{y}] = (\mathbf{x}\,|\,\mathcal{J}\mathbf{y}).$$

Antysymetryczność formy $[\mathbf{x}\,|\,\mathbf{y}]$ pociąga za sobą antysymetryczność operatora $\mathcal{J}$ (tzn. $(\mathcal{J}\mathbf{x}\,|\,\mathbf{y}) = -(\mathbf{x}\,|\,\mathcal{J}\mathbf{y})$). Macierzą operatora $\mathcal{J}$ w wybranej bazie jest właśnie J_0. Mamy oczywiście $\mathcal{J}^2 = -\mathcal{E}$ i $\mathcal{J} \in \mathrm{Sp}(V)$. Jeśli przypomnimy sobie, że przestrzeń V rozkłada się na sumę prostą płaszczyzn hiperbolicznych (rozdz. 1, § 4, twierdzenie 9), to operator $\mathcal{J}$ można interpretować jako obrót o kąt $\pi/2$ na każdej z tych płaszczyzn.

Podprzestrzeń $\Pi \subset V$ nazywamy *osobliwą*, jeśli $[\Pi\,|\,\Pi] = 0$, tj. $[\mathbf{x}\,|\,\mathbf{y}] = 0$ dla dowolnych $\mathbf{x}, \mathbf{y} \in V$ (ortogonalność względem formy $[*\,|\,*]$). Z definicji operatora $\mathcal{J}$ wynika, że podprzestrzeń Π jest osobliwa dokładnie wtedy, gdy podprzestrzenie Π i $\mathcal{J}(\Pi)$ są ortogonalne w sensie euklidesowym, tzn. względem formy $(*\,|\,*)$. Ponieważ operator $\mathcal{J}$ jest nieosobliwy, więc $\dim \Pi = \dim \mathcal{J}(\Pi)$, w szczególności wymiar podprzestrzeni osobliwej Π w $V = \mathbb{R}^{2m}$ nie przekracza m.

Operator $\mathcal{J}$, spełniający warunek $\mathcal{J}^2 = -\mathcal{E}$, pozwala ponadto wprowadzić w przestrzeni symplektycznej V strukturę zespoloną. Co należy przez to rozumieć, wyjaśnimy w § 4.

ĆWICZENIA

1. W przestrzeni P_3 wielomianów rzeczywistych stopni < 3 wektory 1 i t są ortogonalne względem iloczynu skalarnego określonego wzorem $(f \mid g) = \int_{-1}^{1} f(t)g(t)\,dt$ (szczególny przypadek wzoru (2)). Znaleźć:

(a) podprzestrzeń $\langle 1, t\rangle^{\perp}$;
(b) bazę ortonormalną w P_3.

2. Niech $(V, (* \mid *))$ będzie trójwymiarową przestrzenią euklidesową, w której $\|\mathbf{x}\|^2 = (\mathbf{x} \mid \mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2 - 2x_1x_3 + 2x_2x_3$ (sprawdzić, że forma ta jest dodatnio określona). Znaleźć:

(a) kąt α między wektorami $\mathbf{x} = [1,1,1]$ i $\mathbf{y} = [2,2,1]$;
(b) wszystkie wektory ortogonalne do $\mathbf{x} = [1,1,1]$.

3. Wykorzystując algorytm ortogonalizacji Grama-Schmidta, udowodnić, że każdą macierz nieosobliwą $A = (a_{ij}) \in M_n(\mathbb{R})$ można zapisać w postaci iloczynu $A = BC$ macierzy ortogonalnej B i macierzy górnotrójkątnej C (w szczególności $\det C = \pm \det A$).

4. Zbiór $M_n(\mathbb{R})$, rozpatrywany jako n^2-wymiarowa przestrzeń euklidesowa ze standardowym iloczynem skalarnym, zawiera grupę $\mathrm{O}(n)$ macierzy ortogonalnych, określonych przez $n(n+1)/2$ równości (15) lub (15′). Tak więc $\mathrm{O}(n)$ można uważać za „rozmaitość algebraiczną" wymiaru $n^2 - n(n+1)/2 = n(n-1)/2 = \dim \mathfrak{o}(n)$, gdzie $\mathfrak{o}(n)$ jest przestrzenią liniową macierzy antysymetrycznych stopnia n. Naturalne jest oczekiwanie, że między zbiorami $\mathrm{O}(n)$ i $\mathfrak{o}(n)$ zachodzi jakiś naturalny związek. Przykład takiego związku otrzymuje się za pomocą *transformacji Cayleya*

$$K = (E - A)^{-1}(E + A), \qquad A = (E - K)^{-1}(E + K). \tag{18}$$

Należy wykazać, że jeśli $A \in \mathrm{O}(n)$ i $\det(E - A) \neq 0$, to $K \in \mathfrak{o}(n)$. Jest również na odwrót: każdej macierzy $K \in \mathfrak{o}(n)$ odpowiada $A \in \mathrm{O}(n)$ z $1 \notin \operatorname{Spec} A$. Przy tej odpowiedniości należy zatem z „rozmaitości algebraicznej" $\mathrm{O}(n)$ usunąć „hiperpowierzchnię" określoną równaniem $\det(E - A) = 0$.

Można rozpatrzyć inną transformację Cayleya:

$$K = (E + A)^{-1}(E - A), \qquad A = (E + K)^{-1}(E - K); \tag{19}$$

wtedy z $\mathrm{O}(n)$ trzeba usunąć „hiperpowierzchnię" o równaniu $\det(E + A) = 0$, tj. macierze ortogonalne, dla których $-1 \notin \operatorname{Spec} A$.

Należy wykazać, że przekształcenie (19) realizuje opisaną odpowiedniość.

5. Sprawdzić, że macierze ortogonalne odpowiadające macierzom antysymetrycznym przy transformacji Cayleya ((18) lub (19)) mają wyznacznik 1 (zbiór takich macierzy ortogonalnych oznaczamy przez $\mathrm{SO}(n)$).

6. Udowodnić, że wielomian charakterystyczny $\chi_A(t)$ macierzy $A \in \mathrm{O}(n)$ spełnia warunek

$$t^n \chi_A(1/t) = \pm \chi_A(t).$$

7. Niech $A = [A_{(1)}, \ldots, A_{(n)}]$ będzie dowolną macierzą, której wiersze są parami ortogonalne. Udowodnić, że

$$|\det A| = \|A_{(1)}\| \ldots \|A_{(n)}\|,$$

gdzie $\| * \|$ to standardowa norma wektorów w $\mathbb{R}^n$.

8. Niech $X = [X_{(1)}, \ldots, X_{(n)}] \in M_n(\mathbb{R})$. Udowodnić, że

$$|\det X| \leqslant \|X_{(1)}\| \ldots \|X_{(n)}\|$$

(*nierówność Hadamarda*).

§ 2. PRZESTRZENIE UNITARNE

1. Formy hermitowskie. Wiele zagadnień matematycznych sprowadza się do problemów dotyczących operatorów liniowych działających w przestrzeniach zespolonych, dlatego warto poświęcić im osobną uwagę. Bogactwo relacji metrycznych w przestrzeniach euklidesowych nad $\mathbb{R}$ sugeruje, by wprowadzić pojęcie iloczynu skalarnego również w przypadku zespolonym. Jednak standardowa forma dwuliniowa $s(\mathbf{x}, \mathbf{y}) = x_1y_1 + \ldots + x_ny_n$ z $x_j, y_j \in \mathbb{C}$ nie może już być teraz punktem wyjścia, ponieważ odpowiadająca jej forma kwadratowa $\|\mathbf{x}\|^2 = s(\mathbf{x}, \mathbf{x})$ miałaby nieprzyjemną własność

$$\|i\mathbf{x}\|^2 = s(i\mathbf{x}, i\mathbf{x}) = i^2 s(\mathbf{x}, \mathbf{x}) = -\|\mathbf{x}\|^2.$$

Jeśli więc $\mathbf{x} \neq \mathbf{0}$ i $\|\mathbf{x}\| > 0$, to mielibyśmy $\|i\mathbf{x}\|^2 < 0$, co jest sprzeczne z intuicyjnym rozumieniem długości wektora.

Odpowiednikiem przestrzeni euklidesowej jest pojęcie przestrzeni unitarnej. Wprowadźmy najpierw następującą definicję:

DEFINICJA 1. Odwzorowanie $f : V \times V \to \mathbb{C}$ nazywamy *formą półtoraliniową* na przestrzeni zespolonej V, jeśli:

(i) $f(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha f(\mathbf{x}, \mathbf{z}) + \beta f(\mathbf{y}, \mathbf{z})$ dla dowolnych $\alpha, \beta \in \mathbb{C}$ i $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, tj. forma f jest liniowa względem pierwszego argumentu, gdy drugi jest ustalony;

(ii) $f(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \overline{\alpha} f(\mathbf{x}, \mathbf{y}) + \overline{\beta} f(\mathbf{x}, \mathbf{y})$, gdzie kreski nad α i β oznaczają sprzężenie zespolone (*półliniowość* względem drugiego argumentu, gdy pierwszy jest ustalony).

Formę półtoraliniową nazywamy *hermitowską*, jeśli

$$f(\mathbf{y}, \mathbf{x}) = \overline{f(\mathbf{x}, \mathbf{y})}. \tag{1}$$

Załóżmy, że przestrzeń V jest skończenie wymiarowa z bazą $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Jeśli $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$ i $\mathbf{y} = \sum_{j=1}^n y_j\mathbf{e}_j$ oraz $f : V \times V \to \mathbb{C}$ jest formą półtoraliniową, to

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} f_{ij} x_i \overline{y}_j, \qquad f_{ij} = f(\mathbf{e}_i, \mathbf{e}_j).$$

Hermitowskość formy f oznacza, że elementy macierzy $F = (f_{ij})$ spełniają warunek $f_{ij} = \bar{f}_{ji}$. Inaczej mówiąc,

$$F^* = F, \tag{1'}$$

gdzie $F^* = {}^t\overline{F}$ (jeśli $A = (a_{ij})$, to $\overline{A} := (\bar{a}_{ij})$). Macierz F spełniającą warunek (1′) również nazywamy *hermitowską*.

Jeśli F' jest macierzą formy hermitowskiej f w bazie $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ otrzymanej z $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ za pomocą macierzy przejścia A, to

$$F' = {}^tA \cdot F \cdot \overline{A} \tag{2}$$

(por. rozdz. 1, § 4, wzór (5)). Wiemy, że macierz F' jest też hermitowska; wynika to również bezpośrednio z (1′) i (2):

$$(F')^* = {}^t({}^t\overline{A} \cdot \overline{F} \cdot A) = {}^tA \cdot {}^t\overline{F} \cdot \overline{A} = {}^tA \cdot F^* \cdot \overline{A} = {}^tA \cdot F \cdot \overline{A} = F'.$$

Formę kwadratową $f(\mathbf{x}, \mathbf{x})$ odpowiadającą formie hermitowskiej $f(\mathbf{x}, \mathbf{y})$ również nazywamy hermitowską. Ponieważ

$$f(\mathbf{x}, \mathbf{x}) = \overline{f(\mathbf{x}, \mathbf{x})},$$

więc forma kwadratowa hermitowska przyjmuje tylko wartości rzeczywiste. Jeśli przy tym $f(\mathbf{x}, \mathbf{x}) \geqslant 0$ oraz $f(\mathbf{x}, \mathbf{x}) = 0 \Rightarrow \mathbf{x} = \mathbf{0}$, to formę f nazywamy *dodatnio określoną*. Jeśli zapiszemy formę hermitowską f w postaci

$$f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y}) + ih(\mathbf{x}, \mathbf{y}),$$

gdzie funkcje g i h są rzeczywiste, to wykorzystując (1), łatwo się przekonać, że g i h są dwuliniowymi formami rzeczywistymi na V, przy czym forma g jest symetryczna, a h — antysymetryczna. Ponadto dodatnia określoność formy f jest równoważna dodatniej określoności g.

DEFINICJA 2. Przestrzeń liniową V nad $\mathbb{C}$, wyposażoną w dodatnio określoną formę hermitowską $(\mathbf{x} \mid \mathbf{y}) = f(\mathbf{x}, \mathbf{y})$, nazywamy *przestrzenią unitarną*. Liczbę zespoloną $(\mathbf{x} \mid \mathbf{y})$ nazywamy *iloczynem skalarnym* (lub *hermitowskim*) wektorów $\mathbf{x}, \mathbf{y} \in V$.

W nowych oznaczeniach mamy zatem

$$(\mathbf{x} \mid \mathbf{y}) = \overline{(\mathbf{y} \mid \mathbf{x})},$$
$$(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z}) = \alpha(\mathbf{x} \mid \mathbf{z}) + \beta(\mathbf{y} \mid \mathbf{z}),$$
$$(\mathbf{x} \mid \mathbf{x}) \geqslant 0, \quad (\mathbf{x} \mid \mathbf{x}) = 0 \text{ tylko dla } \mathbf{x} = \mathbf{0}.$$

Przykład 1. Przyjmując

$$(\mathbf{x} \mid \mathbf{y}) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n, \tag{3}$$

otrzymujemy dodatnio określoną formę hermitowską na $\mathbb{C}^n$. W ten sposób przestrzeń kartezjańska $\mathbb{C}^n$ stała się przestrzenią unitarną. Macierzą formy (3) w bazie standardowej jest macierz jednostkowa $F = E$. Jeśli przejdziemy do innej bazy w $\mathbb{C}^n$ za pomocą macierzy przejścia A, to zgodnie z (2) macierzą formy (3) w nowej bazie będzie macierz hermitowska $F' = {}^tA \cdot \overline{A}$.

Przestrzeń unitarna jest naturalnym odpowiednikiem przestrzeni euklidesowej. Podobnie jak w przestrzeni euklidesowej, długość $\|\mathbf{v}\|$ wektora $\mathbf{v} \in V$ określamy wzorem

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v} \mid \mathbf{v})}.$$

2. Związki metryczne. Łatwy do sprawdzenia związek

$$2(\mathbf{u} \mid \mathbf{v}) = \|\mathbf{u} + \mathbf{v}\|^2 + i\|\mathbf{u} + i\mathbf{v}\|^2 - (1 + i)\{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\}$$

pokazuje, że iloczyn skalarny można wyrazić za pomocą długości wektorów (polaryzacja). Z oczywistych równości

$$\|\lambda\mathbf{x}\| = \sqrt{(\lambda\mathbf{x} \mid \lambda\mathbf{x})} = \sqrt{|\lambda|^2(\mathbf{x} \mid \mathbf{x})} = |\lambda|\sqrt{(\mathbf{x} \mid \mathbf{x})}$$

wynika znana w przypadku rzeczywistym własność

$$\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|.$$

Ta zgodność z przypadkiem rzeczywistym dotyczy wielu innych stwierdzeń, w tym *nierówności Schwarza*

$$|(\mathbf{x} \mid \mathbf{y})| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\|, \tag{4}$$

przy czym znowu równość zachodzi dokładnie wtedy, gdy wektory $\mathbf{x}$ i $\mathbf{y}$ są proporcjonalne.

Dowód. Jeśli zapiszemy liczbę zespoloną $(\mathbf{x} \mid \mathbf{y})$ w postaci trygonometrycznej: $(\mathbf{x} \mid \mathbf{y}) = |(\mathbf{x} \mid \mathbf{y})|e^{i\varphi}$, gdzie $\varphi \in \mathbb{R}$, to widać, że dla każdego $t \in \mathbb{R}$ zachodzi nierówność

$$\|\mathbf{x}\|^2t^2 + ((\mathbf{x} \mid \mathbf{y})e^{-i\varphi} + \overline{(\mathbf{x} \mid \mathbf{y})}e^{i\varphi})t + \|\mathbf{y}\|^2 = (\mathbf{x}t + \mathbf{y}e^{i\varphi} \mid \mathbf{x}t + \mathbf{y}e^{i\varphi}) \geqslant 0.$$

Ponieważ $(\mathbf{x} \mid \mathbf{y})e^{-i\varphi} = |(\mathbf{x} \mid \mathbf{y})| = \overline{(\mathbf{x} \mid \mathbf{y})}e^{i\varphi}$, więc powyższą nierówność można przepisać w postaci

$$\|\mathbf{x}\|^2t^2 + 2|(\mathbf{x} \mid \mathbf{y})|t + \|\mathbf{y}\|^2 \geqslant 0.$$

Odpowiedni warunek na wyróżnik trójmianu prowadzi do nierówności (4). Nierówność ta staje się równością wtedy, gdy trójmian ma pierwiastek, tj. $\mathbf{x}t_0 + \mathbf{y}e^{i\varphi} = 0$ dla pewnego $t_0 \in \mathbb{R}$, czyli gdy wektory $\mathbf{x}$ i $\mathbf{y}$ są proporcjonalne. ∎

Z nierówności (4) wynika bezpośrednio *nierówność trójkąta*

$$\|\mathbf{x} \pm \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \tag{5}$$

lub w innej, równoważnej wersji

$$\|\mathbf{x} - \mathbf{z}\| \leqslant \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|.$$

Przykład 2. Przestrzenie $CL^2[a,b]$ i P_n nad $\mathbb{C}$, obie z iloczynem skalarnym

$$(f\,|\,g) = \int_a^b f(t)\overline{g(t)}\,dt,$$

są oczywiście unitarne. Nierówność (5) przyjmuje dla nich postać

$$\sqrt{\int_a^b |f(t) \pm g(t)|^2\,dt} \leqslant \sqrt{\int_a^b |f(t)|^2\,dt} + \sqrt{\int_a^b |g(t)|^2\,dt}$$

(porównać z nierównością Minkowskiego w § 1, przykład 3). W przestrzeni unitarnej $\mathbb{C}^n$ ze standardowym iloczynem skalarnym (3) zachodzi nierówność

$$\sqrt{\sum_{i=1}^n |x_i \pm y_i|^2} \leqslant \sqrt{\sum_{i=1}^n |x_i|^2} + \sqrt{\sum_{i=1}^n |y_i|^2}.$$

Z nierówności Schwarza (4) wynika, że jeśli $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$, to istnieje (jednoznacznie wyznaczony) kąt φ $(0 \leqslant \varphi \leqslant \pi/2)$, dla którego

$$\cos\varphi = \frac{|(\mathbf{x}\,|\,\mathbf{y})|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}.$$

Interpretację wielkości $\cos^2\varphi$ w mechanice kwantowej omówiono w podręczniku [2].

3. Ortogonalność. Podobnie jak w przypadku rzeczywistym, układ wektorów $\mathbf{e}_1, \ldots, \mathbf{e}_m$ przestrzeni unitarnej $(V, (*\,|\,*))$ nazywamy *ortonormalnym*, jeśli $(\mathbf{e}_i\,|\,\mathbf{e}_j) = \delta_{ij}$. Każdy taki układ jest liniowo niezależny i – jeśli przestrzeń V jest skończenie wymiarowa – można go uzupełnić do bazy ortonormalnej w V. Aby się o tym przekonać, należy znowu zastosować ortogonalizację Grama–Schmidta (§ 1, p. 3): jeśli wektory $\mathbf{e}_1, \ldots, \mathbf{e}_m$ są ortonormalne i $\mathbf{u} \notin \langle \mathbf{e}_1, \ldots, \mathbf{e}_m \rangle$, to wektor $\mathbf{v} = \mathbf{u} - \sum_{i=1}^m (\mathbf{u}\,|\,\mathbf{e}_i)\mathbf{e}_i$ jest niezerowy oraz $\mathbf{v} \in \langle \mathbf{e}_1, \ldots, \mathbf{e}_m \rangle^\perp$. Wektor $\mathbf{v}$ można teraz unormować i kontynuować proces. Ponadto

$$V = W \oplus W^\perp, \qquad W^{\perp\perp} = W, \tag{6}$$

dla dowolnej podprzestrzeni $W \subset V$.

Jako ćwiczenie proponujemy Czytelnikowi dowód następującego twierdzenia:

TWIERDZENIE 1. *Niech* $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ *będzie bazą ortonormalną przestrzeni unitarnej (lub euklidesowej)* V. *Wtedy:*

(i) $\mathbf{x} = \sum_i (\mathbf{x}\,|\,\mathbf{e}_i)\mathbf{e}_i$ *dla każdego* $\mathbf{x} \in V$;

(ii) $(\mathbf{x}\,|\,\mathbf{y}) = \sum_i (\mathbf{x}\,|\,\mathbf{e}_i)(\mathbf{e}_i\,|\,\mathbf{y})$ *dla dowolnych* $\mathbf{x}, \mathbf{y} \in V$ (*tożsamość Parsevala*);

(iii) $\|\mathbf{x}\|^2 = \sum_i |(\mathbf{x}\,|\,\mathbf{e}_i)|^2$ *dla każdego* $\mathbf{x} \in V$. ∎

Niech $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ będzie bazą ortonormalną przestrzeni unitarnej V. W twierdzeniu 1 wykorzystujemy następującą własność. Ponieważ iloczyn skalarny jest liniowy względem pierwszego argumentu, więc dla każdego wektora $\mathbf{x} = \sum_i x_i \mathbf{e}_i$ mamy równość

$$(\mathbf{x} \,|\, \mathbf{e}_j) = \Big(\sum_i x_i \mathbf{e}_i \,\Big|\, \mathbf{e}_j\Big) = \sum_i x_i (\mathbf{e}_i \,|\, \mathbf{e}_j) = x_j.$$

Tak więc $f_j = (* \,|\, \mathbf{e}_j) : V \to \mathbb{C}$ jest formą liniową przypisującą każdemu wektorowi jego j-tą współrzędną w bazie $(\mathbf{e}_i)$. Jeśli $\mathbf{y} = \sum_j y_j \mathbf{e}_j$ jest innym wektorem przestrzeni V, to

$$(\mathbf{x} \,|\, \mathbf{y}) = \sum_{ij} x_i \overline{y}_j (\mathbf{e}_i \,|\, \mathbf{e}_j) = x_1 \overline{y}_1 + \ldots + x_n \overline{y}_n,$$

tj. jeśli w V wybierzemy bazę ortonormalną, to obliczanie iloczynu skalarnego odbywa się według wzoru (3) dla standardowego iloczynu skalarnego w $\mathbb{C}^n$. Wykazaliśmy tym samym izomorfizm przestrzeni unitarnych $\mathbb{C}^n \cong V$: odwzorowanie $(x_1, \ldots, x_n) \mapsto \sum_i x_i \mathbf{e}_i$ jest liniową bijekcją zachowującą iloczyn skalarny.

W odróżnieniu od przestrzeni euklidesowych, przestrzeni unitarnej V nie można utożsamić z jej przestrzenią dualną, czyli przestrzenią form liniowych na V. Razem z formami liniowymi trzeba też rozpatrywać formy półliniowe, w sensie następującej definicji:

DEFINICJA 3. Niech f będzie formą liniową na przestrzeni zespolonej V. Wtedy forma do niej *sprzężona* $\bar{f} : V \to \mathbb{C}$, określona wzorem $\bar{f}(\mathbf{x}) = \overline{f(\mathbf{x})}$, jest *formą półliniową* na V, tzn. spełnia warunki

$$\bar{f}(\mathbf{x} + \mathbf{y}) = \bar{f}(\mathbf{x}) + \bar{f}(\mathbf{y}), \qquad \bar{f}(\lambda \mathbf{x}) = \overline{\lambda}\, \bar{f}(\mathbf{x}).$$

Oczywiście każda forma półliniowa na V jest postaci $\bar{f}$ dla pewnej formy liniowej f.

Każdą formę liniową f na przestrzeni unitarnej $(V, (* \,|\, *))$ można przedstawić w postaci $f(\mathbf{x}) = (\mathbf{x} \,|\, \mathbf{a})$ dla dokładnie jednego wektora $\mathbf{a}$ (por. § 1, twierdzenie 9), ale odpowiedniość między f i $\mathbf{a}$ nie jest już liniowa. Natomiast każda forma półliniowa $\bar{f}$ na V jest postaci $\bar{f}(\mathbf{x}) = (\mathbf{a} \,|\, \mathbf{x})$, przy czym odpowiedniość $\bar{f} \leftrightarrow \mathbf{a}$ jest liniowa. Istotnie, jeśli $(\mathbf{e}_i)$ jest bazą ortonormalną w V, to definiujemy $\mathbf{a} = \sum_i \bar{f}(\mathbf{e}_i)\mathbf{e}_i$; wtedy dla dowolnego wektora $\mathbf{x} = \sum_j x_j \mathbf{e}_j$ mamy

$$(\mathbf{a} \,|\, \mathbf{x}) = \sum_i \bar{f}(\mathbf{e}_i)\Big(\mathbf{e}_i \,\Big|\, \sum_j x_j \mathbf{e}_j\Big) = \sum_i \bar{f}(\mathbf{e}_i)\overline{x}_i = \bar{f}(\mathbf{x}).$$

Jednoznaczność wektora $\mathbf{a}$ wynika oczywiście z dodatniej określoności formy $(* \,|\, *)$.

4. Macierze unitarne. W przestrzeni euklidesowej przejście od jednej bazy ortonormalnej do drugiej odbywa się za pomocą macierzy ortogonalnej (§ 1, twierdzenie 8). Podobnie ma się rzecz w przypadku przestrzeni unitarnych. Niech $(\mathbf{e}_i)$

i $(\mathbf{e}'_j)$ będą dwiema bazami ortonormalnymi przestrzeni unitarnej $(V, (*\,|\,*))$, związanymi macierzą przejścia $A = (a_{ij})$, tj. $\mathbf{e}'_j = \sum_i a_{ij}\mathbf{e}_i$. Wtedy

$$\delta_{jk} = (\mathbf{e}'_j \,|\, \mathbf{e}'_k) = \sum_{i,s} a_{ij}\bar{a}_{sk}(\mathbf{e}_i \,|\, \mathbf{e}_s) = \sum_i a_{ij}\bar{a}_{ik}.$$

Inaczej mówiąc,

$$E = A^* \cdot A = A \cdot A^*, \tag{7}$$

gdzie $A^* = {}^t\bar{A}$ jest macierzą *hermitowsko sprzężoną* do A (przypominamy, że $\bar{A} = (\bar{a}_{ij})$ oraz że druga równość w (7) wynika z pierwszej).

DEFINICJA 4. Macierz A spełniającą warunek (7) nazywamy *unitarną*.

Widać, że macierz unitarna o wyrazach rzeczywistych jest ortogonalna. Ponadto ponieważ $\det \bar{A} = \overline{\det A}$, więc również $\det A^* = \overline{\det A}$ i z (7) wynika, że $|\det A| = 1$, czyli wyznacznik każdej macierzy unitarnej jest postaci $e^{i\varphi}$. W szczególności macierze unitarne są nieosobliwe.

Wprost z definicji A^* wynika, że

$$(A \cdot B)^* = B^* \cdot A^*. \tag{8}$$

Wobec tego iloczyn macierzy unitarnych A i B jest też unitarny:

$$(AB)(AB)^* = A(BB^*)A^* = AEA^* = AA^* = E.$$

Ponadto (7) oznacza, że $A^{-1} = A^*$, a ponieważ $A^{**} = A$, więc $A^{-1}(A^{-1})^* = A^*A^{**} = A^*A = E$ i podobnie $(A^{-1})^*A^{-1} = E$, czyli macierz odwrotna do macierzy unitarnej jest unitarna. Analogiczne własności, ma się rozumieć, mają macierze ortogonalne.

Z definicji grupy wynika więc

TWIERDZENIE 2.

(i) *Macierze unitarne stopnia n tworzą grupę* $\mathrm{U}(n)$, *zwaną grupą unitarną.*

(ii) *Grupa unitarna* $\mathrm{U}(n)$ *zawiera podgrupę* $\mathrm{O}(n)$ *rzeczywistych macierzy ortogonalnych, zwaną grupą ortogonalną.*

(iii) *Macierze ortogonalne (odpowiednio unitarne) o wyznaczniku* 1 *tworzą grupę, zwaną specjalną grupą ortogonalną* $\mathrm{SO}(n)$ *(odpowiednio specjalną grupą unitarną* $\mathrm{SU}(n)$*).*

Zatem

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}_n(\mathbb{R}) \subset \mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}_n(\mathbb{C}). \ \blacksquare$$

Ogólnie rzecz biorąc, bez większego trudu można by określić grupę ortogonalną $\mathrm{O}(n, \mathfrak{K})$ nad dowolnym ciałem $\mathfrak{K}$, a także grupę unitarną $\mathrm{U}(n, \mathfrak{K})$ nad każdym ciałem $\mathfrak{K}$, w którym istnieje automorfizm rzędu 2, będący odpowiednikiem sprzężenia zespolonego $\alpha \mapsto \bar{\alpha}$.

5. Przestrzenie unormowane. Nierówność (10) z § 1 (długość boku trójkąta nie przekracza sumy dwóch długości pozostałych boków) i jej odpowiednik unitarny (5) pozwalają traktować przestrzenie z iloczynem skalarnym jako przestrzenie metryczne w sensie następującej ogólnej definicji:

DEFINICJA 5. Niech E będzie dowolnym zbiorem, a $d : E \times E \to \mathbb{R}$ odwzorowaniem, które każdej parze elementów $u, v \in E$ przypisuje liczbę rzeczywistą nieujemną $d(u, v)$ (odległość między u i v), spełniającą następujące warunki:

(i) $d(u, v) = d(v, u)$ (symetria);
(ii) $d(u, v) = 0 \Leftrightarrow u = v$;
(iii) $d(u, w) \leqslant d(u, v) + d(v, w)$ (nierówność trójkąta).

Każdą funkcję d o tych własnościach nazywamy *metryką*, a parę (E, d) – *przestrzenią metryczną*.

Przykład 3. W przestrzeni liniowej V z iloczynem skalarnym, a tym samym z normą $\|\mathbf{x}\|$, za odległość między wektorami $\mathbf{x}$ i $\mathbf{y}$ przyjmujemy $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Na przykład dla $V = CL^2[a, b]$ metryką jest

$$d(f, g) = \sqrt{\int_a^b |f(t) - g(t)|^2 \, dt}$$

(nazywana w analizie *metryką* L^2). Dla $E = C[a, b]$ warunki (i)–(iii) definicji 5 są też jednak spełnione np. przez funkcje

$$d'(f, g) = \max_{a \leqslant t \leqslant b} |f(t) - g(t)|, \qquad d''(f, g) = \int_a^b |f(t) - g(t)| \, dt,$$

co łatwo sprawdzić bezpośrednim rachunkiem.

Istnienie metryki w danym zbiorze pozwala na wprowadzenie najprostszych pojęć topologii i analizy, w tym pojęcia granicy. Podzbiory

$$B(a_0, r) = \{x \in E \mid d(a_0, x) < r\},$$
$$\overline{B}(a_0, r) = \{x \in E \mid d(a_0, x) \leqslant r\},$$
$$S(a_0, r) = \{x \in E \mid d(a_0, x) = r\}$$

przestrzeni metrycznej (E, d) nazywamy odpowiednio *kulą otwartą*, *kulą domkniętą* i *sferą* o środku a_0 i promieniu r.

Podzbiór $F \subset E$ jest *ograniczony*, jeśli zawiera się w pewnej kuli.

Ciąg $(e_n)_{n=1}^{\infty}$ w (E, d) jest *zbieżny* do punktu $e \in E$, jeśli $\lim_{n\to\infty} d(e_n, e) = 0$. Ciąg (e_n) jest *ciągiem Cauchy'ego*, jeśli dla każdego $\varepsilon > 0$ istnieje takie $N = N(\varepsilon)$, że $d(e_n, e_m) < \varepsilon$ dla $n, m > N$. Przestrzeń metryczną E nazywamy *zupełną*, jeśli każdy ciąg Cauchy'ego w E jest zbieżny. Z zupełności przestrzeni $\mathbb{R}$ i $\mathbb{C}$, której dowodzi się w analizie, wynika, że przestrzenie $\mathbb{R}^n$ i $\mathbb{C}^n$ z każdą z metryk

$$d(\mathbf{x}, \mathbf{y}) = \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{1/2},$$

$$d_1(\mathbf{x}, \mathbf{y}) = \max_{i=1,\ldots,n} |x_i - y_i|, \qquad d_2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

są zupełne (Czytelnik powinien sprawdzić, że d_1 i d_2 to w istocie metryki; dla d stwierdziliśmy to już w przykładzie 3).

Niech teraz V będzie rzeczywistą lub zespoloną przestrzenią liniową z metryką d. Szczególnie ważne są przypadki, gdy metryka ma jeszcze dwie dodatkowe własności:

(a) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z})$ dla dowolnych $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ (niezmienniczość względem przesunięć);

(b) $d(\lambda\mathbf{x}, \lambda\mathbf{y}) = |\lambda| d(\mathbf{x}, \mathbf{y})$ (pomnożenie wektorów przez skalar λ zwiększa ich odległość $|\lambda|$ razy).

DEFINICJA 6. Jeśli metryka d w przestrzeni liniowej V spełnia warunki (a) i (b), to liczbę $d(\mathbf{x}, \mathbf{0})$ nazywamy *normą* wektora $\mathbf{x} \in V$ i oznaczamy przez $\|\mathbf{x}\|$.

Metryka, którą wprowadziliśmy w przestrzeni z iloczynem skalarnym (przykład 3), ma oczywiście własności (a) i (b); stara i nowa definicja normy są przy tym zgodne, co usprawiedliwia przyjęcie tego samego oznaczenia. Wracając do przypadku ogólnego, wykażemy, że norma określona w definicji 6 ma te same własności, co norma pochodząca od iloczynu skalarnego:

$$\|\mathbf{0}\| = 0; \qquad \|\mathbf{x}\| > 0, \quad \text{jeśli } \mathbf{x} \neq \mathbf{0};$$

$$\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| \quad \text{dla dowolnych } \lambda \in \mathbb{C},\ \mathbf{x} \in V;$$

$$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \quad \text{dla dowolnych } \mathbf{x}, \mathbf{y} \in V.$$

Pierwsze dwie własności wynikają bezpośrednio z aksjomatów odległości oraz z warunków (a) i (b), trzecią zaś udowadnia się tak:

$$\|\mathbf{x} + \mathbf{y}\| = d(\mathbf{x} + \mathbf{y}, \mathbf{0}) = d(\mathbf{x}, -\mathbf{y}) \leqslant d(\mathbf{x}, \mathbf{0}) + d(\mathbf{0}, -\mathbf{y}) = \|\mathbf{x}\| + \|\mathbf{y}\|.$$

DEFINICJA 7. Przestrzeń liniową V (nad $\mathbb{R}$ lub $\mathbb{C}$) wyposażoną w funkcję $\| * \| : V \to \mathbb{R}$ o powyższych własnościach, zwaną *normą*, nazywamy *przestrzenią unormowaną*. Mając normę, można w V określić metrykę wzorem $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$. Jeśli przestrzeń unormowana V z tą metryką jest zupełna, to nazywamy ją *przestrzenią Banacha*.

Przestrzenie $\mathbb{R}^n$ i $\mathbb{C}^n$ z normami odpowiadającymi każdej z metryk wymienionych powyżej są przestrzeniami Banacha.

Zbieżność ciągów, którą określiliśmy w dowolnej przestrzeni metrycznej, nosi w przypadku przestrzeni unormowanej nazwę *zbieżności w danej normie.*

Zachodzi następujące proste

TWIERDZENIE 3. *Niech V będzie skończenie wymiarową przestrzenią liniową nad $\mathbb{R}$ lub $\mathbb{C}$ z iloczynem skalarnym. Niech $\mathbf{x}_k \in V$ dla $k = 1, 2, \ldots$ oraz $\mathbf{x} \in V$. Wtedy następujące warunki są równoważne:*

(i) *ciąg $(\mathbf{x}_k)$ dąży do $\mathbf{x}$ w normie przestrzeni V, tj. $\|\mathbf{x}_k - \mathbf{x}\| \to 0$ przy $k \to \infty$;*
(ii) *$(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{y}) \to 0$ przy $k \to \infty$ dla każdego ustalonego $\mathbf{y} \in V$.*

Dowód. Mamy (i)$\Rightarrow$(ii), ponieważ z nierówności Schwarza wynika, że

$$|(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{y})| \leqslant \|\mathbf{x}_k - \mathbf{x}\| \cdot \|\mathbf{y}\| \to 0.$$

(ii)$\Rightarrow$(i). Aby to wykazać, rozpatrzmy w V dowolną bazę ortonormalną $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Jeśli zachodzi (ii), to $(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{e}_i) \to 0$ dla każdego $i = 1, \ldots, n$. Ale

$$\|\mathbf{x}_k - \mathbf{x}\|^2 = \sum_{i=1}^{n} |(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{e}_i)|^2$$

(twierdzenie 1(iii)), a zatem $\|\mathbf{x}_k - \mathbf{x}\| \to 0$. ■

Dla szeregów w przestrzeni unormowanej (podobnie jak dla szeregów liczbowych) można wprowadzić silniejsze pojęcie zbieżności szeregu niż zbieżność sum częściowych w sensie normy. Mówimy mianowicie, że szereg $\sum_{i=1}^{\infty} \mathbf{x}_i$ jest *bezwzględnie zbieżny*, jeśli zbieżny jest szereg liczbowy $\sum_{i=1}^{\infty} \|\mathbf{x}_i\|$.

ĆWICZENIA

1. Wzór $\mathrm{O}(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) \mid {}^t\!AA = E\}$ określa zespoloną grupę ortogonalną. Jasne jest, że $\mathrm{O}(n) \subset \mathrm{O}(n, \mathbb{C})$ i $\mathrm{SO}(n) \subset \mathrm{SO}(n, \mathbb{C})$ (podgrupy elementów o wyznaczniku 1). Czy, podobnie jak w twierdzeniu 2, zachodzą zawierania $\mathrm{O}(n, \mathbb{C}) \subset \mathrm{U}(n)$ i $\mathrm{SO}(n, \mathbb{C}) \subset \mathrm{U}(n)$?

2. Wykazać, że metryki d_1 i d_2 nie są indukowane przez żaden iloczyn skalarny w $\mathbb{R}^n$ (ani w $\mathbb{C}^n$).

3. Stosując metody analityczne, udowodnić, że dla każdego $p \geqslant 1$ wzór

$$\|\mathbf{x}\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}$$

określa normę w przestrzeniach $\mathbb{R}^n$ i $\mathbb{C}^n$ (tzw. *normę l^p*) [1]. Wykazać, że

$$\lim_{p\to\infty} \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} = \max_{1\leqslant i\leqslant n} |x_i|.$$

---

[1] Zob. np. G. Fichtenholz, *Rachunek różniczkowy i całkowy*, tom I, PWN, Warszawa 1966, s. 226 (*przyp. tłum.*).

Uzasadnia to przyjęcie oznaczenia $d_1(\mathbf{x},\mathbf{y}) =: \|\mathbf{x}-\mathbf{y}\|_\infty$ (mamy oczywiście $d_2(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_1$ i $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2$).

Istnieją też odpowiedniki tych norm na przestrzeni $C[a,b]$ funkcji ciągłych $f : [a,b] \to \mathbb{R}$ (lub $f : [a,b] \to \mathbb{C}$):

$$\|f\|_\infty = \max_{a \leqslant t \leqslant b} |f(t)|, \qquad \|f\|_1 = \int_a^b |f(t)|\,dt$$

i ogólnie

$$\|f\|_p = \left(\int_a^b |f(t)|^p\,dt\right)^{1/p};$$

są to tzw. *normy* L^p (zadanie to ma charakter nieobowiązkowy).

§ 3. OPERATORY LINIOWE NA PRZESTRZENIACH Z ILOCZYNEM SKALARNYM

1. Związki operatorów liniowych i form θ-liniowych. Przez *formę θ-liniową* na przestrzeni liniowej V rozumiemy formę dwuliniową ($\theta = 2$), gdy V jest przestrzenią rzeczywistą, a formę półtoraliniową ($\theta = 3/2$), gdy V jest przestrzenią zespoloną. Formy takie tworzą oczywiście przestrzeń liniową $\mathcal{L}_\theta(V,\mathfrak{K})$ nad ciałem $\mathfrak{K}$ ($\mathfrak{K} = \mathbb{R}$ lub $\mathbb{C}$).

W tym paragrafie V będzie skończenie wymiarową przestrzenią euklidesową lub unitarną z iloczynem skalarnym $(*\,|\,*)$. Niech $\mathcal{A}$ będzie dowolnym operatorem liniowym na V. W rozdziale 2 (§ 3, p. 6) określiliśmy operator $\mathcal{A}^*$ sprzężony do $\mathcal{A}$, działający w V^*. Dla przestrzeni z iloczynem skalarnym istnieje daleko idąca odpowiedniość pomiędzy operatorami i formami θ-liniowymi, co prowadzi (przynajmniej w przypadku rzeczywistym) do naturalnego izomorfizmu między V i V^*, dzięki któremu możemy uważać, że $\mathcal{A}^*$ działa bezpośrednio na V.

Zajmiemy się teraz tą kwestią bardziej szczegółowo. Rozpatrzmy odwzorowanie

$$f_\mathcal{A} : V \times V \to \mathfrak{K} \qquad (\mathfrak{K} = \mathbb{R} \text{ lub } \mathfrak{K} = \mathbb{C})$$

określone wzorem

$$f_\mathcal{A}(\mathbf{x},\mathbf{y}) := (\mathcal{A}\mathbf{x}\,|\,\mathbf{y}). \tag{1}$$

Z własności iloczynu skalarnego wynika natychmiast, że $f_\mathcal{A}$ jest formą θ-liniow na V, tzn. formą dwuliniową w przypadku rzeczywistym, a półtoraliniową w przy padku zespolonym.

Odwzorowanie $\mathcal{A} \mapsto f_\mathcal{A}$ z $\mathcal{L}(V)$ do $\mathcal{L}_\theta(V,\mathfrak{K})$ jest oczywiście liniowe, a tak różnowartościowe. Istotnie, jeśli $f_\mathcal{A} = 0$, to $(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = 0$ dla dowolnych $\mathbf{x} \in V$, w szczególności $(\mathcal{A}\mathbf{x}\,|\,\mathcal{A}\mathbf{x}) = 0$ dla każdego $\mathbf{x} \in V$, tj. $\mathcal{A} = \mathcal{O}$. Poniew $\dim\mathcal{L}(V) = \dim\mathcal{L}_\theta(V,\mathfrak{K})$, wynika stąd, że odwzorowanie $\mathcal{A} \mapsto f_\mathcal{A}$ jest bijekcją

Zbieżność ciągów, którą określiliśmy w dowolnej przestrzeni metrycznej, nosi przypadku przestrzeni unormowanej nazwę *zbieżności w danej normie.*

Zachodzi następujące proste

WIERDZENIE 3. *Niech V będzie skończenie wymiarową przestrzenią liniową d $\mathbb{R}$ lub $\mathbb{C}$ z iloczynem skalarnym. Niech $\mathbf{x}_k \in V$ dla $k = 1, 2, \ldots$ oraz $\mathbf{x} \in V$. tedy następujące warunki są równoważne:*

) *ciąg $(\mathbf{x}_k)$ dąży do $\mathbf{x}$ w normie przestrzeni V, tj. $\|\mathbf{x}_k - \mathbf{x}\| \to 0$ przy $k \to \infty$;*

) *$(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{y}) \to 0$ przy $k \to \infty$ dla każdego ustalonego $\mathbf{y} \in V$.*

owód. Mamy (i)$\Rightarrow$(ii), ponieważ z nierówności Schwarza wynika, że

$$|(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{y})| \leqslant \|\mathbf{x}_k - \mathbf{x}\| \cdot \|\mathbf{y}\| \to 0.$$

(ii)$\Rightarrow$(i). Aby to wykazać, rozpatrzmy w V dowolną bazę ortonormalną $_1, \ldots, \mathbf{e}_n)$. Jeśli zachodzi (ii), to $(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{e}_i) \to 0$ dla każdego $i = 1, \ldots, n$. le

$$\|\mathbf{x}_k - \mathbf{x}\|^2 = \sum_{i=1}^{n} |(\mathbf{x}_k - \mathbf{x} \,|\, \mathbf{e}_i)|^2$$

wierdzenie 1(iii)), a zatem $\|\mathbf{x}_k - \mathbf{x}\| \to 0$. ∎

Dla szeregów w przestrzeni unormowanej (podobnie jak dla szeregów liczbowych) można wprowadzić silniejsze pojęcie zbieżności szeregu niż zbieżność sum ęściowych w sensie normy. Mówimy mianowicie, że szereg $\sum_{i=1}^{\infty} \mathbf{x}_i$ jest *bezwzględnie zbieżny*, jeśli zbieżny jest szereg liczbowy $\sum_{i=1}^{\infty} \|\mathbf{x}_i\|$.

WICZENIA

Wzór $\mathrm{O}(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) \mid {}^tAA = E\}$ określa zespoloną grupę ortogonalną. Jasne jest, że $\mathrm{O}(n) \subset \mathrm{O}(n, \mathbb{C})$ i $\mathrm{SO}(n) \subset \mathrm{SO}(n, \mathbb{C})$ (podgrupy elementów o wyznaczniku 1). Czy, podobnie jak w twierdzeniu 2, zachodzą zawierania $\mathrm{O}(n, \mathbb{C}) \subset \mathrm{U}(n)$ i $\mathrm{SO}(n, \mathbb{C}) \subset \mathrm{U}(n)$?

Wykazać, że metryki d_1 i d_2 nie są indukowane przez żaden iloczyn skalarny w $\mathbb{R}^n$ (ani w $\mathbb{C}^n$).

Stosując metody analityczne, udowodnić, że dla każdego $p \geqslant 1$ wzór

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

określa normę w przestrzeniach $\mathbb{R}^n$ i $\mathbb{C}^n$ (tzw. *normę l^p*) ([1]). Wykazać, że

$$\lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} = \max_{1 \leqslant i \leqslant n} |x_i|.$$

([1]) Zob. np. G. Fichtenholz, *Rachunek różniczkowy i całkowy*, tom I, PWN, Warszawa 1966. 226 (*przyp. tłum.*).

Uzasadnia to przyjęcie oznaczenia $d_1(\mathbf{x},\mathbf{y}) =: \|\mathbf{x}-\mathbf{y}\|_\infty$ (mamy oczywiście $d_2(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_1$ i $d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2$).

Istnieją też odpowiedniki tych norm na przestrzeni $C[a,b]$ funkcji ciągłych $f : [a,b] \to \mathbb{R}$ (lub $f : [a,b] \to \mathbb{C}$):

$$\|f\|_\infty = \max_{a\leqslant t\leqslant b} |f(t)|, \qquad \|f\|_1 = \int_a^b |f(t)|\, dt$$

i ogólnie

$$\|f\|_p = \left(\int_a^b |f(t)|^p\, dt\right)^{1/p};$$

są to tzw. *normy* L^p (zadanie to ma charakter nieobowiązkowy).

§ 3. OPERATORY LINIOWE NA PRZESTRZENIACH Z ILOCZYNEM SKALARNYM

1. Związki operatorów liniowych i form θ-liniowych. Przez *formę θ-liniową* na przestrzeni liniowej V rozumiemy formę dwuliniową ($\theta = 2$), gdy V jest przestrzenią rzeczywistą, a formę półtoraliniową ($\theta = 3/2$), gdy V jest przestrzenią zespoloną. Formy takie tworzą oczywiście przestrzeń liniową $\mathcal{L}_\theta(V,\mathfrak{K})$ nad ciałem $\mathfrak{K}$ ($\mathfrak{K} = \mathbb{R}$ lub $\mathbb{C}$).

W tym paragrafie V będzie skończenie wymiarową przestrzenią euklidesową lub unitarną z iloczynem skalarnym $(*\,|\,*)$. Niech $\mathcal{A}$ będzie dowolnym operatorem liniowym na V. W rozdziale 2 (§ 3, p. 6) określiliśmy operator $\mathcal{A}^*$ sprzężony do $\mathcal{A}$, działający w V^*. Dla przestrzeni z iloczynem skalarnym istnieje daleko idąca odpowiedniość pomiędzy operatorami i formami θ-liniowymi, co prowadzi (przynajmniej w przypadku rzeczywistym) do naturalnego izomorfizmu między V i V^*, dzięki któremu możemy uważać, że $\mathcal{A}^*$ działa bezpośrednio na V.

Zajmiemy się teraz tą kwestią bardziej szczegółowo. Rozpatrzmy odwzorowanie

$$f_\mathcal{A} : V \times V \to \mathfrak{K} \qquad (\mathfrak{K} = \mathbb{R} \text{ lub } \mathfrak{K} = \mathbb{C})$$

określone wzorem

$$f_\mathcal{A}(\mathbf{x},\mathbf{y}) := (\mathcal{A}\mathbf{x}\,|\,\mathbf{y}). \tag{1}$$

Z własności iloczynu skalarnego wynika natychmiast, że $f_\mathcal{A}$ jest formą θ-liniową na V, tzn. formą dwuliniową w przypadku rzeczywistym, a półtoraliniową w przypadku zespolonym.

Odwzorowanie $\mathcal{A} \mapsto f_\mathcal{A}$ z $\mathcal{L}(V)$ do $\mathcal{L}_\theta(V,\mathfrak{K})$ jest oczywiście liniowe, a także różnowartościowe. Istotnie, jeśli $f_\mathcal{A} = 0$, to $(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = 0$ dla dowolnych $\mathbf{x},\mathbf{y} \in V$, w szczególności $(\mathcal{A}\mathbf{x}\,|\,\mathcal{A}\mathbf{x}) = 0$ dla każdego $\mathbf{x} \in V$, tj. $\mathcal{A} = \mathcal{O}$. Ponieważ $\dim\mathcal{L}(V) = \dim\mathcal{L}_\theta(V,\mathfrak{K})$, wynika stąd, że odwzorowanie $\mathcal{A} \mapsto f_\mathcal{A}$ jest bijekcją.

Można się o tym zresztą przekonać, konstruując w sposób jawny, dla danej formy θ-liniowej f, operator $\mathcal{A}_f$ spełniający warunek

$$f(\mathbf{x},\mathbf{y}) = (\mathcal{A}_f\mathbf{x}\,|\,\mathbf{y}). \tag{2}$$

Oto jak należy to zrobić. Niech $(\mathbf{e}_1,\ldots,\mathbf{e}_n)$ będzie bazą ortonormalną w V, a F — macierzą formy f w tej bazie. Jak zwykle, oznaczymy przez $X = [x_1,\ldots,x_n]$ kolumnę współrzędnych wektora $\mathbf{x} = \sum_i x_i\mathbf{e}_i$. Ponieważ baza jest ortonormalna, iloczyn skalarny wektorów $\mathbf{x}$ i $\mathbf{y} = \sum_j y_j\mathbf{e}_j$ jest iloczynem wiersza ${}^tX$ i kolumny $\overline{Y}$: $(\mathbf{x}\,|\,\mathbf{y}) = {}^tX\cdot\overline{Y}$.

Niech $\mathcal{A}_f$ będzie operatorem liniowym o macierzy ${}^tF$. Odpowiada mu więc przekształcenie $X \mapsto {}^tFX$ przestrzeni kolumn. Wzór (2) jest teraz prostą konsekwencją przyjętych oznaczeń:

$$f(\mathbf{x},\mathbf{y}) = {}^tXF\overline{Y} = {}^t({}^tFX)\overline{Y} = (\mathcal{A}_f\mathbf{x}\,|\,\mathbf{y}).$$

Możemy również rozpatrzyć operator liniowy $\mathcal{A}_f^*$ o macierzy $\overline{F}$. Wtedy

$$f(\mathbf{x},\mathbf{y}) = {}^tX(\overline{\overline{F}Y}) = (\mathbf{x}\,|\,\mathcal{A}_f^*\mathbf{y}).$$

Tak więc jeśli macierzą operatora $\mathcal{A}_f$ jest $A := {}^tF$, to macierzą operatora $\mathcal{A}_f^*$ jest $\overline{F} = {}^t(\overline{{}^tF}) = {}^t\overline{A} = A^*$.

Udowodniliśmy w ten sposób

TWIERDZENIE 1. *Niech V będzie przestrzenią liniową z iloczynem skalarnym $(*\,|\,*)$. Wtedy każdy ze wzorów*

$$f_{\mathcal{A}}(\mathbf{x},\mathbf{y}) = (\mathcal{A}\mathbf{x}\,|\,\mathbf{y}), \qquad f_{\mathcal{A}}(\mathbf{x},\mathbf{y}) = (\mathbf{x}\,|\,\mathcal{A}^*\mathbf{y}) \tag{3}$$

wyznacza odpowiedniość wzajemnie jednoznaczną między formami θ-liniowymi i operatorami liniowymi na V. Operator liniowy $\mathcal{A}^* : V \longrightarrow V$ występujący w (3) nazywamy operatorem sprzężonym do $\mathcal{A}$.

Macierz operatora $\mathcal{A}^*$ w dowolnej bazie ortonormalnej otrzymuje się z macierzy operatora $\mathcal{A}$ przez transponowanie i sprzężenie zespolone (w przypadku $\mathfrak{K} = \mathbb{C}$). ■

Definicja

$$(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = (\mathbf{x}\,|\,\mathcal{A}^*\mathbf{y}) \tag{4}$$

usprawiedliwia przyjęcie tej samej nazwy i tego samego oznaczenia, co dla poprzednio rozpatrywanego operatora sprzężonego $\mathcal{A}^* : V^* \longrightarrow V^*$, ponieważ każda forma liniowa na V jest postaci $\mathbf{x} \mapsto (\mathbf{x}\,|\,\mathbf{y})$ dla pewnego ustalonego $\mathbf{y}$.

Warto podkreślić, że w przestrzeni liniowej bez struktury euklidesowej lub unitarnej przypisanie formie θ-liniowej f o macierzy F w pewnej bazie operatora liniowego o macierzy $A = {}^tF$ miałoby charakter przypadkowy. Istotnie, przy przejściu do nowej bazy za pomocą macierzy przejścia B nową macierzą formy f byłoby $F' = {}^tBF\overline{B}$, a więc przypisalibyśmy jej operator o macierzy

$A' = {}^tF' = B^*\,{}^tFB$. Wiemy jednak (rozdz. 2, § 2, twierdzenie 3), że powinno być $A' = B^{-1}AB = B^{-1}({}^tF)B$. Wzory te, ogólnie rzecz biorąc, nie mają ze sobą wiele wspólnego. W wypadku istnienia iloczynu skalarnego rozpatrywaliśmy macierz formy w bazie ortonormalnej; przejście do innej takiej bazy odbywa się za pomocą macierzy unitarnej (lub ortogonalnej), tj. zachodzi równość $B^* = B^{-1}$, dająca pełną zgodność wzorów na A'.

Wypiszmy jeszcze znane własności odwzorowania $\mathcal{A} \mapsto \mathcal{A}^*$:

$$(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \qquad (\alpha\mathcal{A})^* = \overline{\alpha}\mathcal{A}^*, \qquad (\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*, \qquad \mathcal{A}^{**} = \mathcal{A}. \tag{5}$$

Widać tu drobną różnicę w stosunku do wzorów z rozdziału 2 (§ 3, (15)), mianowicie sprzężenie zespolone nad α; wiąże się to z półtoraliniowością formy $(*\,|\,*)$ oraz z faktem, że odwzorowanie $\mathbf{v} \mapsto (*\,|\,\mathbf{v})$ z V do V^* nie jest liniowe, lecz półliniowe (por. § 1, p. 3).

2. Klasy operatorów liniowych. Wśród operatorów liniowych działających na przestrzeni V z iloczynem skalarnym można wyróżnić kilka klas, zależnie od zachowania tych operatorów względem operacji *, wprowadzonej w punkcie 1. Określimy teraz najważniejsze z tych klas.

Definicja 1. Operator liniowy $\mathcal{A}$ na przestrzeni V z iloczynem skalarnym nazywamy *samosprzężonym* lub *hermitowskim*, jeśli $\mathcal{A}^* = \mathcal{A}$. W przypadku rzeczywistym mówimy też wtedy, że operator jest *symetryczny* ([1]).

Z twierdzenia 1 wynika, że operator $\mathcal{A}$ jest samosprzężony wtedy i tylko wtedy, gdy odpowiadająca mu forma θ-liniowa $f_\mathcal{A} : (\mathbf{x},\mathbf{y}) \mapsto (\mathcal{A}\mathbf{x}\,|\,\mathbf{y})$ jest hermitowska. Istotnie, samosprzężoność operatora $\mathcal{A}$ można zapisać w postaci

$$(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = (\mathbf{x}\,|\,\mathcal{A}\mathbf{y}),$$

a hermitowskość formy $f_\mathcal{A}$ — w postaci

$$(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = f_\mathcal{A}(\mathbf{x},\mathbf{y}) = \overline{f_\mathcal{A}(\mathbf{y},\mathbf{x})} = \overline{(\mathcal{A}\mathbf{y}\,|\,\mathbf{x})} = (\mathbf{x}\,|\,\mathcal{A}\mathbf{y})$$

(ostatnia równość wynika z hermitowskości formy $(*\,|\,*)$).

W języku macierzy hermitowskość operatora $\mathcal{A}$ oznacza, że jego macierz A w dowolnej bazie ortonormalnej spełnia warunek ${}^t\overline{A} = A$. Macierze o tej własności nazwaliśmy hermitowskimi, a w przypadku rzeczywistym — symetrycznymi.

Każda macierz rzeczywista jest sumą macierzy symetrycznej i antysymetrycznej (rozdz. 1, § 4, p. 4). Aby podać odpowiednik tej własności w przypadku zespolonym, wprowadzamy następującą definicję:

([1]) Autor przyjmuje (wygodną) konwencję, iż nazwa „operator hermitowski” obejmuje zarówno przypadek zespolony, jak i rzeczywisty (*przyp. tłum.*).

DEFINICJA 2. Operator liniowy $\mathcal{A}$ na przestrzeni z iloczynem skalarnym nazywamy *antyhermitowskim* (lub *antysymetrycznym* dla $\mathfrak{K} = \mathbb{R}$), jeśli $\mathcal{A}^* = -\mathcal{A}$.

Ponieważ $\mathcal{A}^{**} = \mathcal{A}$ dla każdego $\mathcal{A} \in \mathcal{L}(V)$, więc operator $\mathcal{A} + \mathcal{A}^*$ jest hermitowski, zaś $\mathcal{A} - \mathcal{A}^*$ antyhermitowski. Prowadzi to do następującego twierdzenia:

TWIERDZENIE 2. *Każdy operator liniowy $\mathcal{Z}$ na przestrzeni z iloczynem skalarnym można zapisać w postaci*

$$\mathcal{Z} = \mathcal{A} + \mathcal{B},$$

gdzie operator $\mathcal{A}$ jest hermitowski, a $\mathcal{B}$ antyhermitowski. Ponadto jeśli $\mathfrak{K} = \mathbb{C}$, to

$$\mathcal{Z} = \mathcal{X} + i\mathcal{Y}, \tag{6}$$

gdzie operatory $\mathcal{X}$ i $\mathcal{Y}$ są hermitowskie.

Dowód. Przyjmujemy $\mathcal{A} = (\mathcal{Z} + \mathcal{Z}^*)/2$, $\mathcal{B} = (\mathcal{Z} - \mathcal{Z}^*)/2$, $\mathcal{X} = \mathcal{A}$, $\mathcal{Y} = -i\mathcal{B}$. Reszta to bezpośrednie sprawdzenie z użyciem wzorów (5). ■

Zapis (6) przypomina oczywiście zapis liczby zespolonej z w postaci $x + iy$, a więc operatory hermitowskie to dalekie odpowiedniki liczb rzeczywistych. Z kolei operatory antyhermitowskie przypominają liczby urojone $z = iy$, spełniające warunek $\bar{z} = -z$. O ile jednak iloczyn dwóch liczb rzeczywistych jest zawsze liczbą rzeczywistą, to złożenie operatorów hermitowskich może już nie być hermitowskie:

TWIERDZENIE 3. *Złożenie $\mathcal{AB}$ dwóch operatorów hermitowskich jest operatorem hermitowskim wtedy i tylko wtedy, gdy $\mathcal{AB} = \mathcal{BA}$.*

Dowód. Wykorzystując znów wzory (5), otrzymujemy

$$\mathcal{AB} = \mathcal{BA} \;\Rightarrow\; (\mathcal{AB})^* = (\mathcal{BA})^* = \mathcal{A}^*\mathcal{B}^* = \mathcal{AB},$$
$$(\mathcal{AB})^* = \mathcal{AB} \;\Rightarrow\; \mathcal{BA} = \mathcal{B}^*\mathcal{A}^* = (\mathcal{AB})^* = \mathcal{AB}. \;■$$

Rozliczne zastosowania w fizyce i matematyce rodzą potrzebę wprowadzenia w zbiorach wszystkich operatorów hermitowskich lub antyhermitowskich struktury algebry w sensie określonym w rozdziale 2 (§ 2, definicja 1).

Przykład 1. Jak widać z twierdzenia 3, zbiór operatorów hermitowskich nie jest zamknięty względem (łącznego) iloczynu operatorów. Próbując stworzyć ramy algebraiczne dla mechaniki kwantowej, fizyk niemiecki Pascual Jordan wprowadził w latach trzydziestych XX wieku algebry operatorów nad $\mathbb{R}$, noszące dziś jego imię. Iloczyn w tych algebrach, tzw. *iloczyn jordanowski*, jest określony wzorem

$$\mathcal{A} \circ \mathcal{B} = \frac{1}{2}(\mathcal{AB} + \mathcal{BA});$$

jest on (oczywiście) przemienny, ale nie łączny, spełnia jednak *tożsamość Jordana* $(\mathcal{A}^2 \circ \mathcal{B}) \circ \mathcal{A} = \mathcal{A}^2 \circ (\mathcal{B} \circ \mathcal{A})$ (Czytelnik powinien to sprawdzić!). Teoria algebr Jordana, nie tylko skończenie wymiarowych, jest już dziś bardzo rozwinięta.

Przykład 2. Operatory antyhermitowskie tworzą algebrę Liego nad $\mathbb{R}$ (rozdz. 2, § 2, przykład 2) względem zwykłego komutatora operatorów, tzn. jeśli operatory $\mathcal{A}$ i $\mathcal{B}$ są antyhermitowskie, to operator $[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$ jest też antyhermitowski.

Załóżmy, że operator $\mathcal{A}$ ma tę własność, że $(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) = 0$ dla dowolnych wektorów $\mathbf{x}, \mathbf{y} \in V$. Wtedy $\mathcal{A} = \mathcal{O}$ (udowodniliśmy to w punkcie 1). To kryterium trywialności operatora można znacznie wzmocnić:

TWIERDZENIE 4. *Niech V będzie przestrzenią liniową z iloczynem skalarnym. Załóżmy, że operator $\mathcal{A} \in \mathcal{L}(V)$ spełnia warunek $(\mathcal{A}\mathbf{x} \,|\, \mathbf{x}) = 0$ dla każdego $\mathbf{x} \in V$.*

(i) *Jeśli $\mathfrak{K} = \mathbb{C}$, to $\mathcal{A} = \mathcal{O}$.*
(ii) *Jeśli $\mathfrak{K} = \mathbb{R}$ i operator $\mathcal{A}$ jest symetryczny, to również $\mathcal{A} = \mathcal{O}$.*

Dowód. (i) Korzystamy z dwóch tożsamości polaryzacyjnych, które łatwo sprawdzić:

$$(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) + (\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = (\mathcal{A}(\mathbf{x}+\mathbf{y}) \,|\, \mathbf{x}+\mathbf{y}) - (\mathcal{A}\mathbf{x} \,|\, \mathbf{x}) - (\mathcal{A}\mathbf{y} \,|\, \mathbf{y}), \tag{7}$$

$$(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) - (\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = -i(\mathcal{A}(i\mathbf{x}+\mathbf{y}) \,|\, i\mathbf{x}+\mathbf{y}) + i(\mathcal{A}(i\mathbf{x}) \,|\, i\mathbf{x}) + i(\mathcal{A}\mathbf{y} \,|\, \mathbf{y}). \tag{8}$$

Na mocy założenia prawe strony tych tożsamości są równe zeru, co prowadzi do równości

$$(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) + (\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = 0, \qquad (\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) - (\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = 0.$$

Wynika stąd, że $(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) = 0$ dla dowolnych $\mathbf{x}, \mathbf{y} \in V$, a to — jak już wiemy — jest równoważne równości $\mathcal{A} = \mathcal{O}$.

(ii) Z równości (7) (spełnionej w przypadku rzeczywistym i zespolonym) wynika znowu, że $(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) + (\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = 0$. Wobec tego na podstawie symetryczności operatora $\mathcal{A}$ mamy

$$(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) = -(\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = -(\mathbf{y} \,|\, \mathcal{A}^*\mathbf{x}) = -(\mathbf{y} \,|\, \mathcal{A}\mathbf{x}) = -(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}),$$

czyli $(\mathcal{A}\mathbf{x} \,|\, \mathbf{y}) = 0$ dla dowolnych $\mathbf{x}, \mathbf{y} \in V$, co oznacza, że $\mathcal{A} = \mathcal{O}$. ∎

Uwaga 1. W przypadku rzeczywistym symetryczność operatora $\mathcal{A}$ w twierdzeniu 4 jest istotna. Na przykład każdy operator antysymetryczny na przestrzeni euklidesowej V spełnia warunek $(\mathcal{A}\mathbf{x} \,|\, \mathbf{x}) = 0$ dla każdego $\mathbf{x} \in V$, a nie musi być zerowy.

DEFINICJA 3. Operator liniowy $\mathcal{A}$ na przestrzeni z iloczynem skalarnym nazywamy *unitarnym*, jeśli $\mathcal{A}^*\mathcal{A} = \mathcal{E} = \mathcal{A}\mathcal{A}^*$. W przypadku rzeczywistym mówimy też wtedy, że operator jest *ortogonalny* (¹).

Dla $n = 1$ powyższy warunek ma postać $z\bar{z} = 1$, czyli operatory unitarne są odpowiednikami liczb zespolonych o module 1. W postaci macierzowej (względem bazy ortonormalnej) unitarność wyraża się równością (7) z § 2. Macierze o tej własności nazwaliśmy właśnie unitarnymi (lub ortogonalnymi). Pojawiły się one w naturalny sposób jako macierze przejścia od jednej bazy ortonormalnej do drugiej. Wiąże się to z interpretacją geometryczną operatorów unitarnych.

DEFINICJA 4. Operator liniowy $\mathcal{A} : V \to V$ zachowujący odległość (metrykę), tj. spełniający warunek

$$\|\mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in V,$$

nazywamy *izometrią*.

Ponieważ $\mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{y} = \mathcal{A}(\mathbf{x} - \mathbf{y})$, jest oczywiste, że operator $\mathcal{A}$ jest izometrią dokładnie wtedy, gdy $\|\mathcal{A}\mathbf{x}\| = \|\mathbf{x}\|$ dla każdego $\mathbf{x} \in V$. Jeśli ten warunek jest spełniony, to $(\mathcal{A}\mathbf{x} \,|\, \mathcal{A}\mathbf{x}) = (\mathbf{x} \,|\, \mathbf{x})$, skąd $(\mathcal{A}^*\mathcal{A}\mathbf{x} \,|\, \mathbf{x}) = (\mathbf{x} \,|\, \mathbf{x})$, czyli

$$((\mathcal{A}^*\mathcal{A} - \mathcal{E})\mathbf{x} \,|\, \mathbf{x}) = 0$$

dla każdego $\mathbf{x} \in V$. Ponieważ operator $\mathcal{A}^*\mathcal{A} - \mathcal{E}$ jest samosprzężony, więc na mocy twierdzenia 4, zarówno w przypadku zespolonym, jak i rzeczywistym, z powyższego warunku wynika, że $\mathcal{A}^*\mathcal{A} - \mathcal{E} = \mathcal{O}$, tj. jeśli operator $\mathcal{A}$ jest izometryczny, to jest unitarny.

Z drugiej strony, każdy operator unitarny jest izometryczny:

$$(\mathcal{A}\mathbf{x} \,|\, \mathcal{A}\mathbf{x}) = (\mathbf{x} \,|\, \mathcal{A}^*\mathcal{A}\mathbf{x}) = (\mathbf{x} \,|\, \mathcal{E}\mathbf{x}) = (\mathbf{x} \,|\, \mathbf{x}).$$

Udowodniliśmy zatem

TWIERDZENIE 5. *Operatory unitarne, i tylko one, są liniowymi izometriami przestrzeni z iloczynem skalarnym.* ■

Operatory unitarne, czyli izometrie liniowe przestrzeni V z iloczynem skalarnym tworzą *grupę unitarną* $\mathrm{U}(V)$ przestrzeni V dla $\mathfrak{K} = \mathbb{C}$ i *grupę ortogonalną* $\mathrm{O}(V)$ dla $\mathfrak{K} = \mathbb{R}$; grupy te są izomorficzne ze znanymi nam już grupami macierzy unitarnych $\mathrm{U}(n)$ i ortogonalnych $\mathrm{O}(n)$ (§ 2, twierdzenie 2). Mamy tu taką samą sytuację, jak w przypadku grupy $\mathrm{GL}_n(\mathfrak{K})$: można mówić o grupie macierzy albo o grupie $\mathrm{Aut}(V)$ automorfizmów przestrzeni V. Izometrie liniowe to te automorfizmy, które zachowują metrykę.

(¹) W terminologii autora nazwa „operator unitarny" dotyczy zarówno przypadku zespolonego, jak i rzeczywistego (*przyp. tłum.*).

3. Postać kanoniczna operatorów hermitowskich. Istnienie bazy wektorów własnych dla operatora hermitowskiego nie wydaje się, na pierwszy rzut oka, oczywiste. Istotnie, macierze A i A' danego operatora symetrycznego $\mathcal{A} : V \to V$ (dla $\mathfrak{K} = \mathbb{R}$) w różnych bazach ortonormalnych związane są zależnością $A' = B^{-1}AB$, gdzie macierz B jest ortogonalna. Wiemy co prawda, że każdą symetryczną macierz rzeczywistą można sprowadzić do postaci diagonalnej, ale posługując się macierzą B, o której wiemy jedynie, że jest nieosobliwa. Okazuje się, że samosprzężoność operatora $\mathcal{A}$ pozwala na większą „oszczędność" w wyborze B.

Lemat 1. *Wartości własne operatora hermitowskiego są rzeczywiste.*

Dowód. Niech $\mathcal{A} : V \to V$ będzie operatorem hermitowskim, a λ — jego wartością własną; niech $\mathbf{e} \in V$ będzie dowolnym wektorem własnym odpowiadającym λ. Wtedy

$$\lambda(\mathbf{e}\,|\,\mathbf{e}) = (\lambda\mathbf{e}\,|\,\mathbf{e}) = (\mathcal{A}\mathbf{e}\,|\,\mathbf{e}) = (\mathbf{e}\,|\,\mathcal{A}^*\mathbf{e}) = (\mathbf{e}\,|\,\mathcal{A}\mathbf{e}) = (\mathbf{e}\,|\,\lambda\mathbf{e}) = \overline{\lambda}(\mathbf{e}\,|\,\mathbf{e}).$$

Ponieważ $(\mathbf{e}\,|\,\mathbf{e}) \neq 0$, więc $\overline{\lambda} = \lambda$. ∎

Dla operatora symetrycznego (tj. samosprzężonego rzeczywistego) teza lematu 1 nie ma żadnej wartości, gdyż wartości własne z definicji należą do ciała skalarów. Dla takiego operatora z kolei wykażemy, że wszystkie (a priori zespolone) pierwiastki charakterystyczne są rzeczywiste, czyli każdy z tych pierwiastków jest wartością własną.

Lemat 2. *Każdy operator symetryczny na przestrzeni euklidesowej ma wektor własny.*

Dowód. Podobnie jak każdy operator rzeczywisty, operator symetryczny $\mathcal{A}$ ma jednowymiarową lub dwuwymiarową podprzestrzeń niezmienniczą (rozdz. 2, § 3, twierdzenie 7). Istnienie jednowymiarowej podprzestrzeni niezmienniczej to dokładnie teza lematu. Załóżmy więc, że L jest dwuwymiarową podprzestrzenią $\mathcal{A}$-niezmienniczą. Operator $\mathcal{A}$ indukuje operator $\mathcal{A}_L : L \to L$, który jest oczywiście symetryczny względem iloczynu skalarnego ograniczonego do L.

Wybierzmy w L bazę ortonormalną $(\mathbf{e}_1, \mathbf{e}_2)$. Macierz operatora $\mathcal{A}_L$ w tej bazie jest symetryczna:

$$A_L = \begin{Vmatrix} a & b \\ b & d \end{Vmatrix}.$$

Jej wielomianem charakterystycznym jest

$$\chi(t) = \begin{vmatrix} t-a & -b \\ -b & t-d \end{vmatrix} = t^2 - (a+d)t + (ad - b^2).$$

Wyróżnik tego trójmianu kwadratowego wynosi

$$\mathrm{D}(\chi) = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geqslant 0,$$

co oznacza, że trójmian ma pierwiastek rzeczywisty, a operator $\mathcal{A}$ — wartość własną i wektor własny należący do L. ∎

Dalsze rozważania dotyczą już obu przypadków: $\mathfrak{K} = \mathbb{C}$ i $\mathfrak{K} = \mathbb{R}$.

LEMAT 3. *Niech $\mathcal{A}$ będzie operatorem samosprzężonym na przestrzeni liniowej V z iloczynem skalarnym $(*\,|\,*)$, a $L \subset V$ — podprzestrzenią $\mathcal{A}$-niezmienniczą. Wtedy dopełnienie ortogonalne $L^\perp$ jest też $\mathcal{A}$-niezmiennicze.*

Dowód. Istotnie, jeśli $\mathbf{x} \in L$ i $\mathbf{y} \in L^\perp$, to $\mathcal{A}\mathbf{x} \in L$, a więc $(\mathcal{A}\mathbf{x}\,|\,\mathbf{y}) = 0$. Samosprzężoność $\mathcal{A}$ implikuje teraz, że $(\mathbf{x}\,|\,\mathcal{A}\mathbf{y}) = 0$. Wektor $\mathcal{A}\mathbf{y}$ jest więc ortogonalny do każdego wektora $\mathbf{x} \in L$, czyli $\mathcal{A}(L^\perp) \subset L^\perp$. ∎

Teraz możemy już udowodnić *twierdzenie spektralne dla operatorów samosprzężonych*:

TWIERDZENIE 6. *Jeśli operator liniowy $\mathcal{A}$ w przestrzeni V z iloczynem skalarnym jest samosprzężony, to istnieje baza ortonormalna w V, w której macierz operatora $\mathcal{A}$ jest diagonalna. Elementy tej macierzy są liczbami rzeczywistymi.*

Dowód. Na mocy lematów 1 i 2 operator $\mathcal{A}$ ma wektor własny $\mathbf{e}_1$ odpowiadający wartości własnej $\lambda_1 \in \mathbb{R}$. Nie ograniczając ogólności, możemy przyjąć, że $\|\mathbf{e}_1\| = 1$. Dopełnienie ortogonalne $V' = \langle \mathbf{e}_1 \rangle^\perp$ ma wymiar $\dim V - 1$ i jest $\mathcal{A}$-niezmiennicze (lemat 3). Ograniczając operator $\mathcal{A}$ do V' i powtarzając powyższe rozumowanie, znajdujemy unormowany wektor własny $\mathbf{e}_2$ ortogonalny do $\mathbf{e}_1$ i odpowiadający wartości własnej $\lambda_2 \in \mathbb{R}$. Podprzestrzeń $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ jest $\mathcal{A}$-niezmiennicza, jej dopełnienie ortogonalne ma wymiar $\dim V - 2$ i jest też niezmiennicze itd. W ten sposób rozumując przez indukcję względem $\dim V$ lub powtarzając opisaną procedurę dostatecznie wiele razy, znajdziemy $n = \dim V$ parami ortogonalnych unormowanych wektorów własnych $\mathbf{e}_1, \ldots, \mathbf{e}_n$. ∎

Uwaga 2. Wykazaliśmy, że równanie charakterystyczne dowolnej macierzy symetrycznej $A \in M_n(\mathbb{R})$ ma wszystkie pierwiastki rzeczywiste. Aby znaleźć rozmieszczenie tych pierwiastków, można więc zastosować twierdzenia Kartezjusza, Budana–Fouriera i Sturma z części I. Krotność geometryczna każdego pierwiastka jest równa jego krotności algebraicznej, gdyż wszystkie operatory diagonalizowalne mają tę własność (rozdz. 2, § 3, twierdzenie 6).

Uwaga 3. Zgodnie z twierdzeniem 6 każdy operator samosprzężony $\mathcal{A} : V \to V$ ma $n = \dim V$ parami ortogonalnych kierunków własnych. Działanie operatora $\mathcal{A}$ sprowadza się do rozciągnięcia w k-tym kierunku w skali $|\lambda_k|$, gdzie λ_k jest odpowiednią wartością własną, połączonego w przypadku $\lambda_k < 0$ z symetrią względem płaszczyzny ortogonalnej do k-tego kierunku.

4. Sprowadzanie formy kwadratowej do osi głównych. Wiemy (punkt 1), że każdej formie hermitowskiej $f(\mathbf{x}, \mathbf{y})$ na przestrzeni V z iloczynem skalarnym $(* \mid *)$ odpowiada operator samosprzężony $\mathcal{A} = \mathcal{A}_f$, wyznaczony przez warunek

$$f(\mathbf{x}, \mathbf{y}) = (\mathcal{A}\mathbf{x} \mid \mathbf{y}).$$

Na mocy twierdzenia 6 istnieje baza ortonormalna $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni V, złożona z wektorów własnych operatora $\mathcal{A}$: $\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$. Weźmy teraz dowolne dwa wektory

$$\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n, \qquad \mathbf{y} = y_1\mathbf{e}_1 + \ldots + y_n\mathbf{e}_n.$$

Wtedy

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} f(\mathbf{e}_i, \mathbf{e}_j) x_i \overline{y}_j = \sum_{i=1}^{n} \lambda_i x_i \overline{y}_i,$$

ponieważ $f(\mathbf{e}_i, \mathbf{e}_j) = (\mathcal{A}\mathbf{e}_i \mid \mathbf{e}_j) = (\lambda_i \mathbf{e}_i \mid \mathbf{e}_j) = \lambda_i \delta_{ij}$. Biorąc $\mathbf{x} = \mathbf{y}$, dochodzimy do następującego twierdzenia:

TWIERDZENIE 7 (sprowadzanie do osi głównych). *Dla każdej hermitowskiej formy kwadratowej $q(\mathbf{x})$ na n-wymiarowej przestrzeni z iloczynem skalarnym istnieje baza ortonormalna, w której forma przyjmuje postać*

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i |x_i|^2. \ \blacksquare \tag{9}$$

Przykład 3. Dla $\mathfrak{K} = \mathbb{R}$ i $n = 2$ forma kwadratowa q wyznacza krzywą stożkową o równaniu $q(\mathbf{x}) = 1$. W bazie ortonormalnej $(\mathbf{e}_1, \mathbf{e}_2)$, o której mowa w twierdzeniu 7, równanie tej krzywej przyjmuje postać $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$. Wektory $\mathbf{e}_1$ i $\mathbf{e}_2$ wyznaczają kierunki osi głównych elipsy $(\lambda_1\lambda_2 > 0)$ lub hiperboli $(\lambda_1\lambda_2 < 0)$, a wartości λ_1 i λ_2 określają długości półosi.

W języku macierzy twierdzenia 6 i 7 brzmią jednakowo:

Dla każdej hermitowskiej (lub rzeczywistej symetrycznej) macierzy A istnieje macierz unitarna (odpowiednio ortogonalna) B o tej własności, że macierz $A' = B^{-1}AB$ jest diagonalna. Na przekątnej macierzy A' występują wartości własne macierzy A, przy czym każda wartość własna jest powtórzona tyle razy, ile wynosi jej krotność.

Wskazówki praktyczne. Macierzowa interpretacja faktów geometrycznych wskazuje możliwą drogę postępowania przy sprowadzaniu formy kwadratowej

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ij} x_i x_j$$

do postaci kanonicznej (ograniczymy się do przypadku rzeczywistego). Przyjmijmy mianowicie, że $x_1, \ldots, x_n$ to współrzędne wektora $\mathbf{x}$ w przestrzeni euklidesowej V z iloczynem skalarnym

$$(\mathbf{x}\,|\,\mathbf{y}) = \sum_{i=1}^{n} x_i y_i,$$

zatem $\mathbf{e}_1 = (1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, \ldots, 0, 1)$ tworzą bazę ortonormalną. Obliczamy wielomian charakterystyczny $\chi_A(t) = \det(tE - A)$ macierzy $A = (a_{ij})$ i znajdujemy jego pierwiastki (to najtrudniejsza część zadania). Dla każdego pierwiastka λ_i rozwiązujemy układ liniowy jednorodny

$$\begin{aligned}
(a_{11} - \lambda_i)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\
a_{21}x_1 + (a_{22} - \lambda_i)x_2 + \ldots + a_{2n}x_n &= 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda_i)x_n &= 0.
\end{aligned}$$

Przestrzeń rozwiązań tego układu ma wymiar równy algebraicznej krotności pierwiastka λ_i (konsekwencja symetryczności macierzy A). Stosując do bazy tej przestrzeni algorytm ortogonalizacji Grama Schmidta, a następnie łącząc powstałe w ten sposób ortonormalne układy rozwiązań odpowiadające różnym λ_i, otrzymamy bazę ortonormalną przestrzeni V:

$$\mathbf{e}'_j = b_{1j}\mathbf{e}_1 + b_{2j}\mathbf{e}_2 + \ldots + b_{nj}\mathbf{e}_n, \qquad j = 1, \ldots, n.$$

Wykorzystujemy tu znaną już w istocie własność operatora symetrycznego (ogólniej samosprzężonego) $\mathcal{A}$ o macierzy A: wektory własne odpowiadające różnym wartościom własnym λ i μ są ortogonalne. Istotnie, podstawiając $\mathcal{A}\mathbf{u} = \lambda\mathbf{u}$ i $\mathcal{A}\mathbf{v} = \mu\mathbf{v}$ do równości $(\mathcal{A}\mathbf{u}\,|\,\mathbf{v}) = (\mathbf{u}\,|\,\mathcal{A}\mathbf{v})$, otrzymujemy $(\lambda\mathbf{u}\,|\,\mathbf{v}) = (\mathbf{u}\,|\,\mu\mathbf{v})$, a stąd $(\lambda - \mu)(\mathbf{u}\,|\,\mathbf{v})$, czyli $(\mathbf{u}\,|\,\mathbf{v}) = 0$ (wykorzystaliśmy fakt, że μ jest rzeczywiste).

Macierz (b_{ij}), wiążąca stare i nowe układy ortonormalne, jest ortogonalna (tutaj $\mathfrak{K} = \mathbb{R}$), więc nowe współrzędne $x'_1, \ldots, x'_n$ wektora $\mathbf{x}$, dla których

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (x'_i)^2,$$

wyrażają się przez stare według wzorów podanych w § 1 (p. 5).

5. Sprowadzanie pary form kwadratowych do postaci kanonicznej. Przykład form $q(\mathbf{x}) = |x_1|^2 - |x_2|^2$ i $r(\mathbf{x}) = |x_1| \cdot |x_2|$ pokazuje, że nie zawsze można wybrać w przestrzeni taką bazę, by dwie dane formy kwadratowe przyjęły jednocześnie postać kanoniczną. Mimo to w pewnym przypadku, ważnym z praktycznego punktu widzenia, istnienie takiej bazy jest zagwarantowane:

TWIERDZENIE 8. *Załóżmy, że na przestrzeni liniowej V wymiaru n nad $\mathbb{C}$ lub $\mathbb{R}$ dane są dwie formy kwadratowe hermitowskie (tj. formy kwadratowe o wartościach rzeczywistych) $q(\mathbf{x})$ i $r(\mathbf{x})$, przy czym forma $r(\mathbf{x})$ jest dodatnio określona. Wtedy w V istnieje baza, w której obie formy mają postać kanoniczną.*

Dowód. Niech $g(\mathbf{x}, \mathbf{y})$ będzie hermitowską formą θ-liniową odpowiadającą formie kwadratowej $r(\mathbf{x})$. Ponieważ forma $r(\mathbf{x})$ jest dodatnio określona, możemy określić w V iloczyn skalarny wzorem

$$(\mathbf{x} \,|\, \mathbf{y}) := g(\mathbf{x}, \mathbf{y}).$$

Zgodnie z twierdzeniem 7 w przestrzeni V z tym iloczynem skalarnym istnieje baza ortonormalna $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, w której $q(\mathbf{x})$ przyjmuje postać kanoniczną (9). W tej bazie kwadrat skalarny wektora oblicza się według wzoru

$$(\mathbf{x} \,|\, \mathbf{x}) = g(\mathbf{x}, \mathbf{x}) = r(\mathbf{x}) = \sum_{i=1}^{n} |x_i|^2.$$

W ten sposób w bazie $(\mathbf{e}_i)$ obie formy kwadratowe mają postać kanoniczną. ∎

6. Postać kanoniczna izometrii liniowych. Zgodnie z twierdzeniem 5 izometrie liniowe przestrzeni V z iloczynem skalarnym to dokładnie operatory unitarne (ortogonalne dla $\mathfrak{K} = \mathbb{R}$). Będziemy rozpatrywać osobno przypadek rzeczywisty i zespolony, ale najpierw udowodnimy kilka faktów ogólnych.

LEMAT 4. *Wartości własne operatora unitarnego (odpowiednio: ortogonalnego) są liczbami zespolonymi o module 1 (odpowiednio: mogą być równe jedynie ± 1).*

Dowód. Niech $\mathcal{A} : V \to V$ będzie operatorem unitarnym (lub ortogonalnym) i niech $\mathbf{e} \in V$ będzie wektorem własnym odpowiadającym wartości własnej λ. Wtedy

$$(\mathcal{A}\mathbf{e} \,|\, \mathcal{A}\mathbf{e}) = (\lambda\mathbf{e} \,|\, \lambda\mathbf{e}) = \lambda\overline{\lambda}(\mathbf{e} \,|\, \mathbf{e}).$$

Z drugiej strony,

$$(\mathcal{A}\mathbf{e} \,|\, \mathcal{A}\mathbf{e}) = (\mathcal{A}^*\mathcal{A}\mathbf{e} \,|\, \mathbf{e}) = (\mathcal{E}\mathbf{e} \,|\, \mathbf{e}) = (\mathbf{e} \,|\, \mathbf{e}).$$

Stąd $\lambda\overline{\lambda} = 1$, czyli $|\lambda| = 1$. W przypadku rzeczywistym (operatora ortogonalnego) mamy $\lambda^2 = 1$ i są oczywiście tylko dwie możliwości: $\lambda = \pm 1$. ∎

LEMAT 5. *Niech $U \subset V$ będzie podprzestrzenią niezmienniczą dla operatora unitarnego $\mathcal{A} : V \to V$. Wtedy jej dopełnienie ortogonalne $U^{\perp}$ jest też $\mathcal{A}$-niezmiennicze.*

Dowód. Przypominamy, że z definicji

$$U^{\perp} = \{\mathbf{v} \in V \mid (\mathbf{v} \,|\, \mathbf{u}) = 0 \ \forall \mathbf{u} \in U\}.$$

Ograniczenie $\mathcal{A}_U : U \to U$ operatora $\mathcal{A}$ do U jest oczywiście operatorem unitarnym, a w szczególności bijekcją. Niech $\mathbf{u} \in U$. Wtedy $\mathbf{u} = \mathcal{A}\mathbf{u}'$ dla pewnego $\mathbf{u}' \in U$. Jeśli więc $\mathbf{v} \in U^\perp$, to

$$(\mathcal{A}\mathbf{v} \,|\, \mathbf{u}) = (\mathcal{A}\mathbf{v} \,|\, \mathcal{A}\mathbf{u}') = (\mathcal{A}^*\mathcal{A}\mathbf{v} \,|\, \mathbf{u}') = (\mathbf{v} \,|\, \mathbf{u}') = 0,$$

czyli istotnie $\mathcal{A}\mathbf{v} \in U^\perp$. ∎

(A) Operatory unitarne. W języku macierzy chcemy wykazać następujący fakt: *dla każdej macierzy unitarnej A istnieje macierz unitarna B taka, że macierz*

$$C = B^{-1}AB = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

jest diagonalna, przy czym $|\lambda_i| = 1$ dla każdego i.

Okazuje się, że wygodniej jest działać, wykorzystując sens geometryczny operatorów unitarnych.

TWIERDZENIE 9. *Dla $\mathfrak{K} = \mathbb{C}$ każdy operator unitarny jest diagonalizowalny. Inaczej mówiąc, dla każdego operatora unitarnego $\mathcal{A}$ na przestrzeni zespolonej V, gdzie $\dim V = n$, istnieje baza ortonormalna, w której macierz operatora ma postać*

$$A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \qquad |\lambda_i| = 1. \tag{10}$$

Dowód. Niech $\mathbf{e}_1$ będzie dowolnym unormowanym wektorem własnym operatora $\mathcal{A}$. Taki wektor istnieje, gdyż ciało $\mathfrak{K} = \mathbb{C}$ jest algebraicznie domknięte. Na mocy lematu 5 podprzestrzeń $U = \langle \mathbf{e}_1 \rangle^\perp$ wymiaru $n-1$ jest niezmiennicza względem $\mathcal{A}$ i przez indukcję względem n otrzymujemy tezę. Stwierdzenie o modułach wartości własnych udowodniliśmy w lemacie 4. ∎

Zauważmy, że i na odwrót, każdy operator o macierzy postaci (10) jest unitarny, ponieważ

$${}^t\overline{A}A = \operatorname{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n)\operatorname{diag}(\lambda_1, \ldots, \lambda_n) = E.$$

Zwykle diagonalną macierz unitarną zapisuje się w postaci

$$A = \begin{Vmatrix} e^{i\varphi_1} & & & \\ & e^{i\varphi_2} & & \\ & & \ddots & \\ & & & e^{i\varphi_n} \end{Vmatrix},$$

ykorzystując postać trygonometryczną liczb zespolonych.

3) Operatory ortogonalne. Jak już wspominaliśmy, lematy 4 i 5 zachod wnież dla operatorów ortogonalnych $\mathcal{A} : V \to V$, z tym że wartości własne mo az być równe tylko ± 1. Dalej jednak trzeba rozumować nieco inaczej, gdyż o or ortogonalny może nie mieć wektorów własnych — wtedy jednak, podob każdy rzeczywisty operator liniowy, ma dwuwymiarową podprzestrzeń włas

Dlatego stosując lemat 5, możemy rozłożyć V na sumę prostą jedno- i dwuwymiarowych podprzestrzeni niezmienniczych, parami ortogonalnych:

$$V = V_1 \oplus \ldots \oplus V_m; \tag{11}$$

na każdej z nich $\mathcal{A}$ indukuje operator ortogonalny. Załóżmy, że żadna z podprzestrzeni dwuwymiarowych w tym rozkładzie nie da się już rozłożyć na jednowymiarowe podprzestrzenie niezmiennicze. Wtedy sumę baz ortonormalnych w podprzestrzeniach V_i dla $i = 1, \ldots, m$ nazywamy *bazą kanoniczną* dla operatora $\mathcal{A}$.

Zbadajmy, jaka jest macierz operatora $\mathcal{A}$ w bazie kanonicznej. Jeśli $A_i := A_{V_i}$ jest macierzą ograniczenia operatora $\mathcal{A}$ do V_i, to

$$A = A_1 \dotplus \ldots \dotplus A_m = \begin{Vmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{Vmatrix}.$$

Dlatego wystarczy ograniczyć się do przypadku, gdy V nie ma (nietrywialnych) podprzestrzeni niezmienniczych i $\dim V = 1$ lub $\dim V = 2$. Jeśli $\dim V = 1$ i $V = \langle \mathbf{e} \rangle$, przy czym $\|\mathbf{e}\| = 1$, to $\mathcal{A}\mathbf{e} = \lambda \mathbf{e}$ i $\lambda = \pm 1$ (lemat 4). Jeśli $\dim V = 2$ i $V = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$, gdzie $(\mathbf{e}_i \,|\, \mathbf{e}_j) = \delta_{ij}$, to w tej bazie ortonormalnej

$$A = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}, \qquad a, b, c, d \in \mathbb{R}.$$

Przypuśćmy, że $\det A = ad - bc = -1$. Wtedy wielomian charakterystyczny $\chi_A(t) = t^2 - (a + d)t - 1$ ma dwa pierwiastki rzeczywiste, a więc operator $\mathcal{A}$ ma wektor własny, wbrew założeniu. Wobec tego musi być $\det A = 1$. Obliczając macierz odwrotną na podstawie znanych wzorów, znajdujemy

$$A^{-1} = \begin{Vmatrix} d & -b \\ -c & a \end{Vmatrix}.$$

Z drugiej strony, na mocy ortogonalności,

$$A^{-1} = {}^tA = \begin{Vmatrix} a & c \\ b & d \end{Vmatrix}.$$

Porównując te dwa wzory, dochodzimy do wniosku, że

$$A = \begin{Vmatrix} a & -c \\ c & a \end{Vmatrix}, \qquad a^2 + c^2 = 1.$$

Wobec tego $a = \cos\varphi$ i $c = \sin\varphi$ dla pewnego kąta φ, a więc

$$A = \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix},$$

czyli operator $\mathcal{A}$ jest obrotem na płaszczyźnie V.

Z powyższej analizy wynika, że jeśli w rozkładzie (11) pierwsze r składników $V_1, \ldots, V_r$ to nierozkładalne dwuwymiarowe podprzestrzenie niezmiennicze, a pozostałe to podprzestrzenie jednowymiarowe (co zawsze można osiągnąć, przenumerowując wektory bazowe), i jeśli $\varphi_1, \ldots, \varphi_r$ to odpowiednie kąty obrotów, wówczas macierz A przyjmuje postać opisaną w poniższym twierdzeniu.

Twierdzenie 10. *Dla każdego operatora ortogonalnego $\mathcal{A}$ na n-wymiarowej przestrzeni euklidesowej V istnieje baza ortonormalna w V, w której operator $\mathcal{A}$ ma macierz*

$$\left\| \begin{array}{cccccc} \cos\varphi_1 & -\sin\varphi_1 & & & & \\ \sin\varphi_1 & \cos\varphi_1 & & & & \\ & & \ddots & & & \\ & & & \cos\varphi_r & -\sin\varphi_r & & \\ & & & \sin\varphi_r & \cos\varphi_r & & \\ & & & & & -E_k & \\ & & & & & & E_l \end{array} \right\|, \qquad k + l + 2r = n. \blacksquare$$

7. Operatory normalne. Dowody twierdzeń spektralnych 6 i 9 mają ze sobą wiele wspólnego, co nie jest przypadkowe, ponieważ operatory hermitowskie i unitarne należą do ogólniejszej, naturalnej klasy operatorów diagonalizowalnych.

Definicja 5. Niech V będzie przestrzenią unitarną. Operator $\mathcal{A} : V \to V$ spełniający warunek

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A} \tag{12}$$

nazywamy *normalnym*. Jego macierz (w dowolnej bazie) też nazywamy *normalną*.

Przypomnijmy, że na mocy (5) zachodzą równości

$$(\lambda\mathcal{E})^* = \overline{\lambda}\mathcal{E}, \qquad (\mathcal{A} - \lambda\mathcal{E})^* = \mathcal{A}^* - \overline{\lambda}\mathcal{E},$$

więc operator $\mathcal{A}$ jest normalny jednocześnie z $\mathcal{A} - \lambda\mathcal{E}$. Jeśli operator $\mathcal{A}$ jest normalny, to

$$\|\mathcal{A}\mathbf{x}\|^2 = (\mathcal{A}\mathbf{x} \,|\, \mathcal{A}\mathbf{x}) = (\mathbf{x} \,|\, \mathcal{A}^*\mathcal{A}\mathbf{x}) = (\mathbf{x} \,|\, \mathcal{A}\mathcal{A}^*\mathbf{x}) = (\mathcal{A}^*\mathbf{x} \,|\, \mathcal{A}^*\mathbf{x}) = \|\mathcal{A}^*\mathbf{x}\|^2.$$

Podstawiając $\mathcal{A} - \lambda\mathcal{E}$ w miejsce $\mathcal{A}$, otrzymujemy

$$\|\mathcal{A}\mathbf{x} - \lambda\mathbf{x}\| = \|\mathcal{A}^*\mathbf{x} - \overline{\lambda}\mathbf{x}\|,$$

skąd wynika, że

$$\mathcal{A}\mathbf{x} = \lambda\mathbf{x} \;\Leftrightarrow\; \mathcal{A}^*\mathbf{x} = \overline{\lambda}\mathbf{x}. \tag{13}$$

Jest oczywiste, że każdy z warunków $\mathcal{A}^* = \mathcal{A}$ i $\mathcal{A}^* = \mathcal{A}^{-1}$ pociąga za sobą (12). Łatwo jednak podać przykłady operatorów normalnych, które nie są ani hermitowskie (albo antyhermitowskie), ani unitarne, np. operator o macierzy $A = \mathrm{diag}(2i, 2, 1, \ldots, 1)$ w pewnej bazie ortonormalnej. Ponadto definicja operatora normalnego przenosi się na nieskończenie wymiarowe przestrzenie Hilberta i znajduje tam rozliczne zastosowania.

Naszym bezpośrednim celem jest dokładne scharakteryzowanie klasy operatorów normalnych na przestrzeni unitarnej.

TWIERDZENIE 11. *Niech $\mathcal{A}$ będzie operatorem liniowym na przestrzeni unitarnej V. Następujące warunki są równoważne:*

(a) *operator $\mathcal{A}$ jest diagonalizowalny w pewnej bazie ortonormalnej;*
(b) *operator $\mathcal{A}$ jest normalny.*

Dowód. Jeśli $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest bazą ortonormalną, dla której $\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$, to definiujemy operator $\mathcal{B}$ wzorami $\mathcal{B}\mathbf{e}_i = \overline{\lambda}_i \mathbf{e}_i$ dla $i = 1, \ldots, n$. Wtedy oczywiście $\mathcal{AB} = \mathcal{BA}$. Ale

$$(\mathcal{A}\mathbf{e}_i \,|\, \mathbf{e}_j) = \lambda_i \delta_{ij} = \lambda_j \delta_{ij} = (\mathbf{e}_i \,|\, \mathcal{B}\mathbf{e}_j),$$

skąd wynika, że $\mathcal{B} = \mathcal{A}^*$, czyli z (a) wynika (b).

Aby udowodnić implikację odwrotną, weźmy dowolną wartość własną λ operatora $\mathcal{A}$ i rozpatrzmy odpowiadającą mu podprzestrzeń własną

$$V^\lambda = \{\mathbf{x} \in V \mid \mathcal{A}\mathbf{x} = \lambda\mathbf{x}\}.$$

Wtedy z (12) wynika, że

$$\mathcal{A}^*(V^\lambda) \subset V^\lambda,$$

a wobec tego

$$\mathcal{A}((V^\lambda)^\perp) \subset (V^\lambda)^\perp.$$

Istotnie, jeśli $\mathbf{y} \in (V^\lambda)^\perp$, to $(\mathbf{y} \,|\, \mathbf{x}) = 0$ dla każdego $\mathbf{x} \in V^\lambda$. Wtedy jednak $(\mathcal{A}\mathbf{y} \,|\, \mathbf{x}) = (\mathbf{y} \,|\, \mathcal{A}^*\mathbf{x}) = 0$, gdyż $\mathcal{A}^*\mathbf{x} \in V^\lambda$.

Ponieważ $(\mathcal{A}^*)^* = \mathcal{A}$, więc przez symetrię podprzestrzeń $(V^\lambda)^\perp$ jest też $\mathcal{A}^*$-niezmiennicza. Ograniczenia operatorów $\mathcal{A}$ i $\mathcal{A}^*$ do $(V^\lambda)^\perp$ są oczywiście ze sobą przemienne, tj. są operatorami normalnymi. Rozumując przez indukcję względem $n = \dim V$, możemy przyjąć, że na $(V^\lambda)^\perp$ operator $\mathcal{A}$ jest diagonalizowalny w pewnej bazie ortonormalnej. Na V^λ to samo ma miejsce z definicji, a ponieważ $V = V^\lambda \oplus (V^\lambda)^\perp$, dowód został zakończony. ∎

Ponieważ operatory hermitowskie i unitarne są normalne, więc twierdzenia 6 i 9 są szczególnymi przypadkami twierdzenia 11. Przypominamy w tym miejscu o zupełnym układzie operatorów idempotentnych (rozdz. 2, § 3, twierdzenie 1). Ogólne twierdzenie spektralne dla operatorów normalnych można sformułować następująco:

TWIERDZENIE 12. *Dla każdego operatora normalnego $\mathcal{A}$ na n-wymiarowej przestrzeni unitarnej V istnieją różne liczby zespolone $\lambda_1, \ldots, \lambda_m$, gdzie $m \leqslant n$, oraz parami ortogonalne niezerowe projekcje $\mathcal{P}_1, \ldots, \mathcal{P}_m$ takie, że:*

(i) $\sum_j \mathcal{P}_j = \mathcal{E}$;

(ii) *$\sum_j \lambda_j \mathcal{P}_j = \mathcal{A}$ jest rozkładem spektralnym operatora $\mathcal{A}$, tj. $\{\lambda_1, \ldots, \lambda_m\} =$* $\operatorname{Spec} \mathcal{A}$;

(iii) *rozkład w punkcie* (ii) *jest jednoznaczny: jeśli operator $\mathcal{A}$ jest kombinacją liniową $\sum_j \lambda_j \mathcal{P}_j$ parami ortogonalnych niezerowych projekcji spełniających* (i), *to $\{\lambda_1, \ldots, \lambda_m\} =$* $\operatorname{Spec} \mathcal{A}$;

(iv) *istnieją wielomiany zespolone $f_1(t), \ldots, f_m(t)$ takie, że*

$$f_i(\lambda_j) = \delta_{ij}, \qquad f_i(\mathcal{A}) = \mathcal{P}_i$$

(w przypadku operatora samosprzężonego wszystkie liczby λ_i i wielomiany $f_i(t)$ są rzeczywiste).

Dowód. Niech $\lambda_1, \ldots, \lambda_m$ będą wszystkimi różnymi wartościami własnymi operatora $\mathcal{A}$. Na mocy twierdzenia 11 mamy $V = \bigoplus_{i=1}^m V^{\lambda_i}$, przy czym podprzestrzenie własne V^{λ_i} są parami ortogonalne (i oczywiście niezerowe). Niech $\mathcal{P}_i$ będzie rzutem ortogonalnym przestrzeni V na V^{λ_i}, czyli operatorem rzutowania na V^{λ_i} równolegle do $\sum_{j \neq i} V^{\lambda_j}$. Równość (i) wynika teraz z powyższego rozkładu przestrzeni V. Zauważmy przy okazji, że wzajemną ortogonalność rzutów $\mathcal{P}_i$ (tzn. ortogonalność ich obrazów) można zapisać w postaci równości $\mathcal{P}_i \mathcal{P}_j = \mathcal{P}_j \mathcal{P}_i = \delta_{ij} \mathcal{P}_i$.

Jeśli $\mathbf{v} \in V$ i $\mathbf{v}_j = \mathcal{P}_j \mathbf{v}$, to $\mathcal{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$, a wobec tego

$$\mathcal{A}\mathbf{v} = \mathcal{A}\Big(\sum_j \mathcal{P}_j \mathbf{v}\Big) = \sum_j \mathcal{A}\mathbf{v}_j = \sum_j \lambda_j \mathbf{v}_j = \sum_j \lambda_j (\mathcal{P}_j \mathbf{v}) = \Big(\sum_j \lambda_j \mathcal{P}_j\Big)\mathbf{v},$$

co dowodzi (ii).

Jeśli chodzi o stwierdzenie (iii) o jednoznaczności rozkładu spektralnego, rozumujemy tak: Ponieważ $\mathcal{P}_i \neq \mathcal{O}$, więc istnieje niezerowy wektor $\mathbf{x} \in \operatorname{Im} \mathcal{P}_i$. Wtedy $\mathcal{P}_i \mathbf{x} = \mathbf{x}$ i $\mathcal{P}_j \mathbf{x} = \mathbf{0}$ dla $j \neq i$. Wobec tego

$$\mathcal{A}\mathbf{x} = \sum_j \lambda_j \mathcal{P}_j \mathbf{x} = \lambda_i \mathcal{P}_i \mathbf{x} = \lambda_i \mathbf{x},$$

tj. $\lambda_i \in \operatorname{Spec} \mathcal{A}$.

Na odwrót, jeśli $\lambda \in \operatorname{Spec} \mathcal{A}$ i $\mathcal{A}\mathbf{v} = \lambda \mathbf{v}$ dla pewnego $\mathbf{v} \neq \mathbf{0}$, to

$$\mathcal{A}\mathbf{v} = \lambda \mathbf{v} = \lambda \sum_j \mathbf{v}_j, \qquad \mathbf{v}_j = \mathcal{P}_j \mathbf{v},$$

a z drugiej strony,

$$\mathcal{A}\mathbf{v} = \mathcal{A} \sum_j \mathbf{v}_j = \sum_j \mathcal{A}\mathbf{v}_j = \sum_j \lambda_j \mathbf{v}_j.$$

Zatem $\sum_j (\lambda - \lambda_j)\mathbf{v}_j = \mathbf{0}$. Wektory $\mathbf{v}_1, \ldots, \mathbf{v}_m$ są jednak parami ortogonalne (gdyż założyliśmy wzajemną ortogonalność rzutów $\mathcal{P}_1, \ldots, \mathcal{P}_m$), więc te spośród wektorów $\mathbf{v}_i$, które nie są zerowe, są liniowo niezależne. Wobec tego $(\lambda - \lambda_j)\mathbf{v}_j = \mathbf{0}$ dla każdego j, przy czym jeśli $\mathbf{v}_i \neq \mathbf{0}$ (a takie i się znajdzie, gdyż $\mathbf{v} \neq \mathbf{0}$), to

$\lambda = \lambda_i \in \{\lambda_1, \dots, \lambda_m\}$. Zatem rozkład operatora $\mathcal{A}$ na kombinację liniową parami ortogonalnych rzutów jest jednoznaczny.

Wielomiany $f_1(t), \dots, f_m(t)$ występujące w (iv) można określić jawnym wzorem:

$$f_i(t) = \prod_{j \neq i} \frac{t - \lambda_j}{\lambda_i - \lambda_j}$$

(por. wzór interpolacyjny Lagrange'a w części I, rozdz. 6, § 1, (4)). Widać, że $f_i(t) \in \mathbb{R}[t]$, jeśli operator $\mathcal{A}$ jest samosprzężony, oraz $f_i(\lambda_j) = \delta_{ij}$.

Wykorzystując rozkład (ii) i ortogonalność rzutów $\mathcal{P}_j$, otrzymujemy

$$\mathcal{A}^2 = \Big(\sum_i \lambda_i \mathcal{P}_i\Big)\Big(\sum_j \lambda_j \mathcal{P}_j\Big) = \sum_{i,j} \lambda_i \lambda_j \mathcal{P}_i \mathcal{P}_j = \sum_j \lambda_j^2 \mathcal{P}_j,$$

$$\mathcal{A}^3 = \mathcal{A}\mathcal{A}^2 = \Big(\sum_i \lambda_i \mathcal{P}_i\Big)\Big(\sum_j \lambda_j^2 \mathcal{P}_j\Big) = \sum_j \lambda_j^3 \mathcal{P}_j,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\mathcal{A}^k = \sum_j \lambda_j^k \mathcal{P}_j$$

(również dla $k = 0$, na podstawie (i): $\mathcal{A}^0 = \mathcal{E} = \sum_j \mathcal{P}_j = \sum_j \lambda_j^0 \mathcal{P}_j$). Zatem

$$f(\mathcal{A}) = \sum_j f(\lambda_j) \mathcal{P}_j$$

dla dowolnego wielomianu $f(t)$. W szczególności

$$f_i(\mathcal{A}) = \sum_j f_i(\lambda_j) \mathcal{P}_j = f_i(\lambda_i) \mathcal{P}_i = \mathcal{P}_i. \ \blacksquare$$

Podobnie jak każdy operator liniowy na przestrzeni unitarnej, operator normalny $\mathcal{A}$ można zapisać w postaci $\mathcal{A} = \mathcal{B} + i\mathcal{C}$, gdzie operatory $\mathcal{A}$ i $\mathcal{B}$ są hermitowskie (zob. (6)); wtedy $\mathcal{A}^* = \mathcal{B} - i\mathcal{C}$, a więc operatory $\mathcal{B}$ i $\mathcal{C}$ wyrażają się przez $\mathcal{A}$ i $\mathcal{A}^*$:

$$\mathcal{B} = \frac{1}{2}(\mathcal{A} + \mathcal{A}^*), \qquad \mathcal{C} = \frac{1}{2i}(\mathcal{A} - \mathcal{A}^*).$$

Ponieważ $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$, więc $\mathcal{B}\mathcal{C} = \mathcal{C}\mathcal{B}$. Na odwrót, jeśli dla pewnego $\mathcal{A}$ operatory $\mathcal{B}$ i $\mathcal{C}$ są ze sobą przemienne, to $\mathcal{A}$ i $\mathcal{A}^*$ też:

$$\mathcal{A}\mathcal{A}^* = \mathcal{B}^2 + \mathcal{C}^2 = \mathcal{A}^*\mathcal{A},$$

czyli operator $\mathcal{A}$ jest normalny.

Odchodząc na chwilę od operatorów normalnych, zatrzymajmy się na roli przemienności operatorów (mówimy też wtedy, że operatory *komutują*).

Lemat 6. *Jeśli operatory $\mathcal{A}$ i $\mathcal{B}$ na przestrzeni zespolonej V są ze sobą przemienne, to mają wspólny wektor własny.*

Dowód. Niech $\lambda \in \operatorname{Spec} \mathcal{A}$. Rozpatrzmy podprzestrzeń własną $V^\lambda = \{\mathbf{x} \in V \mid \mathcal{A}\mathbf{x} = \lambda\mathbf{x}\}$. Wtedy $\mathcal{B}V^\lambda \subset V^\lambda$. Istotnie, z równości $\mathcal{AB} = \mathcal{BA}$ wynika, że jeśli $\mathbf{x} \in V^\lambda$, to

$$\mathcal{A}(\mathcal{B}\mathbf{x}) = \mathcal{B}(\mathcal{A}\mathbf{x}) = \mathcal{B}(\lambda\mathbf{x}) = \lambda(\mathcal{B}\mathbf{x}),$$

tj. $\mathcal{B}\mathbf{x} \in V^\lambda$.

Operator $\mathcal{B}$ ograniczony do V^λ ma wektor własny $\mathbf{y} \in V^\lambda$: $\mathcal{B}\mathbf{y} = \mu\mathbf{y}$ dla pewnego $\mu \in \operatorname{Spec} \mathcal{B}$. Zatem $\mathcal{A}\mathbf{y} = \lambda\mathbf{y}$ i $\mathcal{B}\mathbf{y} = \mu\mathbf{y}$, czyli $\mathbf{y}$ jest wspólnym wektorem własnym. ∎

TWIERDZENIE 13. *Jeśli $\mathcal{A}$ i $\mathcal{B}$ są dwoma operatorami hermitowskimi lub dwoma operatorami unitarnymi na n-wymiarowej przestrzeni unitarnej V, to macierze obu operatorów są diagonalne w pewnej wspólnej bazie ortonormalnej wtedy i tylko wtedy, gdy operatory $\mathcal{A}$ i $\mathcal{B}$ są ze sobą przemienne.*

Dowód. Jeśli (dowolne) operatory $\mathcal{A}$ i $\mathcal{B}$ są diagonalizowalne we wspólnej bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ (niekoniecznie ortonormalnej) oraz $\mathbf{x} = \sum_i x_i\mathbf{e}_i$, to $\mathcal{A}\mathbf{x} = \sum_i \lambda_i x_i \mathbf{e}_i$ i $\mathcal{B}\mathbf{x} = \sum_i \mu_i x_i \mathbf{e}_i$ dla pewnych skalarów λ_i, μ_i. Wtedy

$$\mathcal{AB}\mathbf{x} = \mathcal{A}\Big(\sum_i \mu_i x_i \mathbf{e}_i\Big) = \sum_i \lambda_i \mu_i x_i \mathbf{e}_i = \sum_i \mu_i \lambda_i x_i \mathbf{e}_i = \mathcal{B}\Big(\sum_i \lambda_i x_i \mathbf{e}_i\Big) = \mathcal{BA}\mathbf{x},$$

czyli operatory $\mathcal{A}$ i $\mathcal{B}$ są przemienne.

Na odwrót, załóżmy, że operatory $\mathcal{A}$ i $\mathcal{B}$ są oba hermitowskie lub oba unitarne oraz $\mathcal{AB} = \mathcal{BA}$. Wtedy na mocy lematu 6 operatory $\mathcal{A}$ i $\mathcal{B}$ mają wspólny wektor własny $\mathbf{e}_1$. Możemy przyjąć, że $\|\mathbf{e}_1\| = 1$. Podprzestrzeń $W = \langle \mathbf{e}_1 \rangle^\perp$ wymiaru $n-1$ jest niezmiennicza względem obu operatorów na podstawie ich hermitowskości (lemat 3) lub unitarności (lemat 5). Ograniczenia obu operatorów do W są znowu przemiennymi operatorami hermitowskimi lub unitarnymi. Indukcja względem wymiaru prowadzi do konstrukcji bazy ortonormalnej, w której macierze obu operatorów mają postać diagonalną. ∎

Uwaga 4. Przypomnijmy, że na mocy twierdzenia 3 przemienność operatorów hermitowskich $\mathcal{A}$ i $\mathcal{B}$ jest równoważna hermitowskości operatora $\mathcal{AB}$.

8. Operatory dodatnio określone. Każdemu operatorowi samosprzężonemu $\mathcal{A}$ na przestrzeni z iloczynem skalarnym odpowiada forma kwadratowa $q(\mathbf{x}) = (\mathcal{A}\mathbf{x} \mid \mathbf{x})$. Pojęcia związane z tą formą, jak dodatnia określoność, półokreśloność itd. (rozdz. 1, § 4, p. 8), można też odnieść do operatora $\mathcal{A}$.

DEFINICJA 6. Operator samosprzężony $\mathcal{A}$ nazywamy *dodatnio określonym*, jeśli $(\mathcal{A}\mathbf{x} \mid \mathbf{x}) > 0$ dla każdego niezerowego wektora $\mathbf{x} \in V$.

Z twierdzenia 6 i towarzyszących mu uwag wynika, że dla każdego dodatnio określonego operatora $\mathcal{A}$ istnieje baza ortonormalna w V, w której macierz A operatora ma postać diagonalną

$$A = \begin{Vmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{Vmatrix} \tag{14}$$

z dodatnimi wartościami własnymi $\lambda_1, \ldots, \lambda_n$. Na odwrót, jeśli operator $\mathcal{A}$ ma w pewnej bazie ortonormalnej macierz (14), to warunek $\lambda_1 > 0, \ldots, \lambda_n > 0$ gwarantuje dodatnią określoność $\mathcal{A}$. Piszemy wtedy $\mathcal{A} > 0$.

Można też mówić o operatorach *dodatnio półokreślonych* (oznaczenie: $\mathcal{A} \geqslant 0$), dla których $\lambda_i \geqslant 0$ dla każdego i. Dla dwóch operatorów samosprzężonych $\mathcal{A}$ i $\mathcal{B}$ piszemy $\mathcal{A} \geqslant \mathcal{B}$, jeśli $\mathcal{A} - \mathcal{B} \geqslant 0$.

Z definicji operator dodatnio określony jest nieosobliwy (odwracalny); widać to też z nierówności Schwarza: jeśli $\mathbf{x} \neq \mathbf{0}$, to

$$0 < |(\mathcal{A}\mathbf{x} \,|\, \mathbf{x})| \leqslant \|\mathcal{A}\mathbf{x}\| \cdot \|\mathbf{x}\|,$$

a więc $\mathcal{A}\mathbf{x} \neq \mathbf{0}$. Na odwrót, nieosobliwość i warunek $\mathcal{A} \geqslant 0$ gwarantują, że operator $\mathcal{A}$ jest dodatnio określony.

Stwierdzenie 1. *Każdy operator dodatnio określony $\mathcal{A}$ jest kwadratem dokładnie jednego dodatnio określonego operatora $\mathcal{B}$ (piszemy $\mathcal{B} = \sqrt{\mathcal{A}}$).*

Dowód. Wystarczy sprowadzić macierz operatora $\mathcal{A}$ do postaci diagonalnej (14) i przyjąć $B = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$ (bierzemy pierwiastki dodatnie). Operator $\mathcal{B}$ o macierzy B w danej bazie ortonormalnej jest dodatnio określony. Warunek $A = B^2$ zachowuje się przy zmianie bazy: $C^{-1}AC = (C^{-1}BC)^2$. Zatem $\mathcal{A} = \mathcal{B}^2$.

Jednoznaczność $\mathcal{B}$ najwygodniej udowodnić, wykorzystując twierdzenie 12 o rozkładzie spektralnym. Jeśli mianowicie $\mathcal{B}' > 0$ i $(\mathcal{B}')^2 = \mathcal{A}$, to uwzględniając rozkład spektralny $\mathcal{B}' = \sum_j \mu_j \mathcal{P}'_j$, otrzymujemy

$$\sum_j \mu_j^2 \mathcal{P}'_j = (\mathcal{B}')^2 = \mathcal{A} = \sum_i \lambda_i \mathcal{P}_i.$$

Liczby $\mu_j > 0$ są parami różne, a więc ich kwadraty też. Jednoznaczność rozkładu spektralnego prowadzi do wniosku, że zbiory $\{\mu_j^2\}$ i $\{\lambda_i\}$ są równe, tj. po odpowiednim przenumerowaniu zachodzą równości $\mu_i^2 = \lambda_i$ oraz $\mathcal{P}'_i = \mathcal{P}_i$, a stąd $\mu_i = \sqrt{\lambda_i}$ i $\mathcal{B}' = \mathcal{B} = \sqrt{\mathcal{A}}$. ∎

Stwierdzenie 2. *Niech $\mathcal{C}$ będzie dowolnym operatorem nieosobliwym na przestrzeni liniowej z iloczynem skalarnym. Wtedy złożenie $\mathcal{A} = \mathcal{C}\mathcal{C}^*$ (a także $\mathcal{C}^*\mathcal{C}$) jest operatorem dodatnio określonym.*

Dowód. Samosprzężoność takiego złożenia już sprawdzaliśmy: $(\mathcal{C}\mathcal{C}^*)^* = \mathcal{C}^{**}\mathcal{C}^* = \mathcal{C}\mathcal{C}^*$. Ponadto jeśli operator $\mathcal{C}$ jest nieosobliwy, to $\mathcal{C}^*$ też, a wobec tego

$$(\mathcal{C}\mathcal{C}^*\mathbf{x} \,|\, \mathbf{x}) = (\mathcal{C}^*\mathbf{x} \,|\, \mathcal{C}^*\mathbf{x}) > 0 \qquad \forall \mathbf{x} \neq \mathbf{0},$$

czyli operator $\mathcal{A} = \mathcal{C}\mathcal{C}^*$ jest dodatnio określony. To samo odnosi się do złożenia $\mathcal{C}^*\mathcal{C}$. ∎

Ze stwierdzeń 1 i 2, uogólnionych na przypadek operatorów osobliwych, wynika natychmiast

TWIERDZENIE 14. *Niech $\mathcal{A}$ będzie operatorem liniowym na przestrzeni V z iloczynem skalarnym. Następujące warunki są równoważne:*

(i) *istnieje operator $\mathcal{B}$ taki, że $\mathcal{A} = \mathcal{B}^2$ i $\mathcal{B}^* = \mathcal{B}$;*
(ii) *istnieje operator $\mathcal{C}$ taki, że $\mathcal{A} = \mathcal{C}\mathcal{C}^*$;*
(iii) $(\mathcal{A}\mathbf{x} \mid \mathbf{x}) \geqslant 0$ *dla każdego* $\mathbf{x} \in V$. ∎

W jednowymiarowym przypadku zespolonym każda z własności (i), (ii) charakteryzuje liczby rzeczywiste nieujemne: $z \geqslant 0$ oznacza zarówno możliwość zapisu $z = \lambda^2$, gdzie $\lambda \in \mathbb{R}$ (odpowiednik (i)), jak i $z = z'\overline{z'}$ (odpowiednik (ii)).

9. Rozkład biegunowy. Wspomniana odpowiedniość między liczbami zespolonymi i operatorami liniowymi na przestrzeni z iloczynem skalarnym sięga jeszcze dalej, aż do zapisu liczby zespolonej w postaci trygonometrycznej $z = |z|e^{i\varphi} = \sqrt{z\bar{z}}e^{i\varphi}$. Świadczy o tym następne

TWIERDZENIE 15. *Każdy nieosobliwy operator $\mathcal{A}$ na przestrzeni V z iloczynem skalarnym można przedstawić w postaci*

$$\mathcal{A} = \mathcal{P}\mathcal{Q}, \tag{15}$$

gdzie operator $\mathcal{P}$ jest dodatnio określony, a $\mathcal{Q}$ jest izometrią (czyli operatorem unitarnym). Rozkład (15) *jest jednoznaczny; nosi on nazwę rozkładu biegunowego operatora $\mathcal{A}$.*

Dowód. Ze stwierdzeń 1 i 2 wynika, że $\mathcal{A}\mathcal{A}^* = \mathcal{P}^2$, gdzie $\mathcal{P}$ jest pewnym operatorem dodatnio określonym (w szczególności odwracalnym). Jeśli przyjmiemy $\mathcal{Q} = \mathcal{P}^{-1}\mathcal{A}$, to równość (15) będzie oczywiście spełniona. Trzeba się tylko przekonać, że $\mathcal{Q}$ jest izometrią.

Istotnie, ponieważ $\mathcal{P}^* = \mathcal{P}$ i $\mathcal{P}\mathcal{P}^{-1} = \mathcal{E} = \mathcal{E}^* = (\mathcal{P}^{-1})^*\mathcal{P}^*$, więc $(\mathcal{P}^{-1})^* = (\mathcal{P}^*)^{-1} = \mathcal{P}^{-1}$, a stąd

$$\mathcal{Q}\mathcal{Q}^* = \mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*(\mathcal{P}^{-1})^* = \mathcal{P}^{-1}\mathcal{P}^2\mathcal{P}^{-1} = \mathcal{E}.$$

Jeśli teraz $\mathcal{P}\mathcal{Q} = \mathcal{A} = \mathcal{P}_1\mathcal{Q}_1$ są dwoma rozkładami postaci (15), to $\mathcal{Q}^*\mathcal{P} = \mathcal{Q}_1^*\mathcal{P}_1$. Wobec tego $\mathcal{P}\mathcal{Q} \cdot \mathcal{Q}^*\mathcal{P} = \mathcal{P}_1\mathcal{Q}_1 \cdot \mathcal{Q}_1^*\mathcal{P}_1$, a stąd $\mathcal{P}^2 = \mathcal{P}_1^2$, czyli $\mathcal{P} = \mathcal{P}_1$ (jednoznaczność pierwiastka kwadratowego). Wtedy również $\mathcal{Q} = \mathcal{Q}_1$ i dowód jednoznaczności rozkładu biegunowego został zakończony. ∎

Uwaga 5. Mamy oczywiście

$$\mathcal{A} = \mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{Q}^{-1}\mathcal{P}\mathcal{Q},$$

a stąd
$$\mathcal{A} = \mathcal{Q}\mathcal{P}_1,$$
gdzie $\mathcal{Q}$ jest izometrią, a $\mathcal{P}_1 = \mathcal{Q}^*\mathcal{P}\mathcal{Q}$ jest operatorem dodatnio określonym.

Rozkład biegunowy (15) (ale bez jednoznaczności $\mathcal{Q}$) zachodzi też dla operatorów osobliwych. Zajmiemy się jednak teraz inną własnością tego rozkładu. Dla liczby zespolonej z kolejność czynników w postaci trygonometrycznej nie ma znaczenia: $|z|e^{i\varphi} = e^{i\varphi}|z|$. Jeśli natomiast $\mathcal{A} = \mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P}$, gdzie operator $\mathcal{P}$ jest dodatnio określony, a $\mathcal{Q}$ — unitarny, to
$$\mathcal{A}\mathcal{A}^* = \mathcal{P}\mathcal{Q}\cdot\mathcal{Q}^*\mathcal{P}^* = \mathcal{P}^2 = \mathcal{P}\mathcal{Q}^*\mathcal{Q}\mathcal{P} = (\mathcal{Q}\mathcal{P})^*\mathcal{Q}\mathcal{P} = \mathcal{A}^*\mathcal{A},$$
czyli operator $\mathcal{A}$ jest normalny.

Na odwrót, jeśli operator normalny $\mathcal{A}$ ma rozkład biegunowy $\mathcal{A} = \mathcal{P}\mathcal{Q}$, to
$$\mathcal{P}^2 = \mathcal{P}\mathcal{Q}\mathcal{Q}^*\mathcal{P} = \mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}\mathcal{Q}^*\mathcal{P}^*\mathcal{P}\mathcal{Q} = \mathcal{Q}^{-1}\mathcal{P}^2\mathcal{Q},$$
czyli $\mathcal{Q}$ jest przemienne z $\mathcal{P}^2$. Pociąga to jednak za sobą przemienność $\mathcal{Q}$ z $\mathcal{P}$, ponieważ $\mathcal{P} = \sqrt{\mathcal{P}^2}$ jest *wielomianem* od $\mathcal{P}^2$ (co wynika z twierdzenia o rozkładzie spektralnym). Wobec tego

operator o rozkładzie biegunowym $\mathcal{A} = \mathcal{P}\mathcal{Q}$ jest normalny $\Leftrightarrow$ $\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P}$.

ĆWICZENIA

1. Kiedy macierz unitarną A wymiarów $n \times n$ można zapisać w postaci (multiplikatywnego) komutatora $A = XYX^{-1}Y^{-1}$ macierzy unitarnych X, Y? Warunkiem koniecznym jest oczywiście $\det A = 1$. Udowodnić, że warunek ten jest też dostateczny. Inaczej mówiąc, w grupie $\mathrm{SU}(n)$ każdy element jest komutatorem.

2. *Macierz Jacobiego* to macierz rzeczywista postaci
$$J = \begin{Vmatrix} a_1 & -b_1 & 0 & \dots & 0 & 0 \\ -c_1 & a_2 & -b_2 & \dots & 0 & 0 \\ 0 & -c_2 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & a_{n-1} & -b_{n-1} \\ 0 & 0 & 0 & \dots & -c_{n-1} & a_n \end{Vmatrix}, \qquad b_ic_i > 0, \qquad i = 1, \dots, n-1$$
Udowodnić, że widmo takiej macierzy jest rzeczywiste i proste.

3. Czy prawdziwy jest odpowiednik twierdzenia 13, w którym jeden z operatorów $\mathcal{A}$, $\mathcal{B}$ jest hermitowski, a drugi jest izometrią?

4. Niech $\mathcal{A}$ i $\mathcal{B}$ będą dwoma przemiennymi operatorami liniowymi na przestrzeni liniowej V nad ciałem $\mathfrak{K}$. Udowodnić, że jeśli każdy z tych operatorów jest diagonalizowalny, to są one wspólnie diagonalizowalne, tzn. istnieje baza w V złożona z wektorów, które są wektorami własnymi dla $\mathcal{A}$ i $\mathcal{B}$ jednocześnie.

5. Udowodnić, że jeśli przemienne operatory liniowe $\mathcal{A}$ i $\mathcal{B}$ są dodatnio określone, to operator $\mathcal{AB}$ jest też dodatnio określony.

6. Udowodnić, że jeśli macierz rzeczywista A jest antysymetryczna (${}^tA = -A$), to macierz A^2 jest symetryczna i niedodatnio określona. Wynika stąd w szczególności, że niezerowe pierwiastki charakterystyczne rzeczywistej macierzy antysymetrycznej są liczbami urojonymi.

7. Niech $\mathcal{A}$ i $\mathcal{B}$ będą operatorami hermitowskimi, przy czym jeden z nich, np. $\mathcal{A}$, jest dodatnio określony. Udowodnić, że widmo $\mathrm{Spec}(\mathcal{AB})$ jest rzeczywiste.

8. Niech $q(\mathbf{x})$ będzie formą kwadratową na przestrzeni euklidesowej V z iloczynem skalarnym $(* \mid *)$. W których punktach sfery jednostkowej $(\mathbf{x} \mid \mathbf{x}) = 1$ forma q osiąga maksimum lub minimum? Ogólniej: które punkty sfery jednostkowej są stacjonarne dla q, tj. pochodne q w tych punktach we wszystkich kierunkach są zerami? Udowodnić następujące stwierdzenie:

Punkt $\mathbf{x}$ *sfery jednostkowej jest punktem stacjonarnym formy kwadratowej* q *wtedy i tylko wtedy, gdy* $\mathbf{x}$ *jest wektorem własnym operatora symetrycznego* $\mathcal{F}$ *określonego wzorem* $q(\mathbf{x}) = (\mathcal{F}\mathbf{x} \mid \mathbf{x})$.

W szczególności największa wartość formy $q(\mathbf{x})$ *na sferze jednostkowej jest równa największemu współczynnikowi w jej postaci kanonicznej, a najmniejsza wartość — najmniejszemu.*

9. Udowodnić następujące uogólnienie lematu 6:

TWIERDZENIE 16. *Każda przemienna rodzina operatorów liniowych na skończenie wymiarowej przestrzeni zespolonej ma wspólny wektor własny.*

10. Udowodnić, że jeśli

$$\mathfrak{S} = \{A_i \in M_n(\mathbb{C}) \mid A_iA_j = A_jA_i \text{ dla } i,j \in J\}$$

jest dowolną przemienną rodziną macierzy, to istnieje taka macierz nieosobliwa C, że zbiór

$$C^{-1}\mathfrak{S}C^{-1} = \{C^{-1}A_jC \mid j \in J\}$$

jest przemienną rodziną macierzy górnotrójkątnych.

11. Niech E będzie (jak zwykle) macierzą jednostkową stopnia n, a E_{ij} dla $i, j = 1, \ldots, n$ — jedynkami macierzowymi (tzn. jedynym elementem niezerowym macierzy E_{ij} jest $e_{ij} = 1$). Sprawdzić, że zbiór

$$\mathfrak{E} = \{E_{ij} \mid 1 \leqslant i \leqslant [n/2],\ [n/2]+1 \leqslant j \leqslant n\} \cup \{E\}$$

ma $[n^2/4] + 1$ elementów i stanowi przemienną rodzinę liniowo niezależnych macierzy górnotrójkątnych. Symbol $[x]$ oznacza tu część całkowitą liczby x.

12. Udowodnić

TWIERDZENIE (I. Schur, 1905). *Największy wymiar przemiennej podalgebry w* $M_n(\mathbb{C})$ *jest równy* $[n^2/4] + 1$.

Inaczej mówiąc, należy wykazać, że największa możliwa liczba liniowo niezależnych i parami przemiennych macierzy kwadratowych stopnia n nad $\mathbb{C}$ wynosi $[n^2/4]+1$.

W rzeczywistości ciało $\mathbb{C}$ można tu zastąpić dowolnym ciałem $\mathfrak{K}$.

13. Niech $(V, (*\,|\,*))$ będzie przestrzenią euklidesową parzystego wymiaru $n = 2m$, a $f(\mathbf{x}, \mathbf{y})$ — niezdegenerowaną formą antysymetryczną na V. Udowodnić, że istnieje rozkład $V = V_1 \oplus V_2$ na sumę prostą dwóch podprzestrzeni m-wymiarowych oraz nieosobliwy operator symetryczny $\mathcal{A} : V \to V$ taki, że jeśli $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ oraz $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$, gdzie $\mathbf{x}_i, \mathbf{y}_i \in V_i$ dla $i = 1, 2$, to

$$f(\mathbf{x}, \mathbf{y}) = (\mathbf{x}_1 \,|\, \mathcal{A}\mathbf{y}_2) - (\mathbf{x}_2 \,|\, \mathcal{A}\mathbf{y}_1).$$

§ 4. KOMPLEKSYFIKACJA I URZECZYWISTNIENIE

Już kilka razy mieliśmy okazję się przekonać, że problem sprowadzania macierzy operatora liniowego $\mathcal{A} : V \to V$ do postaci kanonicznej ma różne rozwiązania w zależności od tego, czy ciało skalarów $\mathfrak{K}$ jest algebraicznie domknięte (np. $\mathfrak{K} = \mathbb{C}$), czy też nie (np. $\mathfrak{K} = \mathbb{R}$). Odnosi się to w szczególności do izometrii, czyli operatorów unitarnych i ortogonalnych. Ponieważ w przypadku zespolonym, kosztem utraty w pewnym stopniu intuicji geometrycznej, sytuacja algebraiczna się upraszcza, często stosuje się do przestrzeni i operatorów rzeczywistych operację (lub, jak również mówimy, funktor) kompleksyfikacji; funktor urzeczywistnienia działa w przeciwnym kierunku. Przyjrzymy się temu nieco dokładniej.

1. Struktura zespolona. Niech V będzie przestrzenią liniową nad $\mathbb{R}$.

Definicja 1. Mówimy, że operator liniowy $\mathcal{J} : V \to V$ wprowadza w przestrzeni V *strukturę zespoloną*, jeśli $\mathcal{J}^2 = -\mathcal{E}$.

Przykład 1. Niech $\dim V = n = 2m$ i rozpatrzmy operator $\mathcal{J}$ o macierzy

$$J = \begin{Vmatrix} 0 & -1 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & -1 & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & -1 \\ & & & & & 1 & 0 \end{Vmatrix} \qquad (1)$$

(zob. rozdz. 1, § 4, wniosek z twierdzenia 9). Jest oczywiste, że $\mathcal{J}$ wprowadza w V strukturę zespoloną, gdyż $J^2 = -E$.

Termin „struktura zespolona” wiąże się z faktem, że za pomocą $\mathcal{J}$ można określić w zbiorze V strukturę przestrzeni liniowej nad $\mathbb{C}$, przyjmując

$$(\alpha + i\beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathcal{J}\mathbf{v}, \qquad \alpha, \beta \in \mathbb{R}, \ \mathbf{v} \in V.$$

Aksjomaty rozdzielności

$$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}, \qquad (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}, \qquad a, b \in \mathbb{C}, \ \mathbf{u}, \mathbf{v} \in V,$$

są spełnione, gdyż $\mathcal{J}$ jest operatorem liniowym. Ponadto z równości $\mathcal{J}^2 = -\mathcal{E}$ wynika, że

$$\begin{aligned}(\alpha + i\beta)[(\gamma + i\delta)\mathbf{v}] &= (\alpha + i\beta)(\gamma\mathbf{v} + \delta\mathcal{J}\mathbf{v}) = \alpha(\gamma\mathbf{v} + \delta\mathcal{J}\mathbf{v}) + \beta\mathcal{J}(\gamma\mathbf{v} + \delta\mathcal{J}\mathbf{v}) \\ &= \alpha\gamma\mathbf{v} + \alpha\delta\mathcal{J}\mathbf{v} + \beta\gamma\mathcal{J}\mathbf{v} - \beta\delta\mathbf{v} = (\alpha\gamma - \beta\delta)\mathbf{v} + (\alpha\delta + \beta\gamma)\mathcal{J}\mathbf{v} \\ &= [\alpha\gamma - \beta\delta + i(\alpha\delta + \beta\gamma)]\mathbf{v} = [(\alpha + i\beta)(\gamma + i\delta)]\mathbf{v}.\end{aligned}$$

Wszystkie pozostałe aksjomaty przestrzeni liniowej są oczywiście spełnione, ponieważ nie wiążą się z mnożeniem przez skalary.

DEFINICJA 2. Przestrzeń zespoloną zdefiniowaną powyżej oznaczamy przez $\widetilde{V}$; nazywamy ją *przestrzenią zespoloną wyznaczoną przez strukturę zespoloną* $\mathcal{J}$ na przestrzeni rzeczywistej V.

Pokażemy, że przykład 1 nie był przypadkowy.

STWIERDZENIE 1. *Skończenie wymiarowa rzeczywista przestrzeń liniowa V ze strukturą zespoloną $\mathcal{J}$ ma wymiar parzysty, a macierz operatora $\mathcal{J}$ w pewnej bazie jest postaci* (1)*. Ponadto*

$$\dim_{\mathbb{C}} \widetilde{V} = \frac{1}{2} \dim_{\mathbb{R}} V. \tag{2}$$

Dowód. Załóżmy, że znaleźliśmy już wektory $\mathbf{e}_1, \ldots, \mathbf{e}_k \in V$ takie, że układ $2k$ wektorów $\mathbf{e}_1, \mathcal{J}\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathcal{J}\mathbf{e}_k$ jest liniowo niezależny. Wtedy albo ich powłoka liniowa

$$V_k = \langle \mathbf{e}_1, \mathcal{J}\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathcal{J}\mathbf{e}_k \rangle$$

jest całą przestrzenią V i dowód jest zakończony, albo istnieje wektor $\mathbf{e}_{k+1} \notin V_k$. W tym ostatnim przypadku układ $\mathbf{e}_1, \mathcal{J}\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathcal{J}\mathbf{e}_k, \mathbf{e}_{k+1}, \mathcal{J}\mathbf{e}_{k+1}$ jest również liniowo niezależny. Istotnie, przypuśćmy, że

$$\mathcal{J}\mathbf{e}_{k+1} = \alpha\mathbf{e}_{k+1} + \mathbf{v}_k, \qquad \alpha \in \mathbb{R}, \qquad \mathbf{v}_k \in V_k,$$

i zastosujmy do obu stron operator $\mathcal{J}$:

$$-\mathbf{e}_{k+1} = \alpha\mathcal{J}\mathbf{e}_{k+1} + \mathcal{J}\mathbf{v}_k.$$

Zauważmy, że podprzestrzeń V_k jest niezmiennicza względem $\mathcal{J}$, więc $\mathcal{J}\mathbf{v}_k \in V_k$. Mnożąc przedostatnią równość przez α i porównując z ostatnią, dochodzimy do wniosku, że

$$(\alpha^2 + 1)\mathbf{e}_{k+1} = -\alpha\mathbf{v}_k - \mathcal{J}\mathbf{v}_k \in V_k,$$

co jednak przeczy wyborowi $\mathbf{e}_{k+1}$, gdyż $\alpha^2 + 1 \neq 0$.

Kontynuując proces dołączania do V_k wektorów liniowo niezależnych, otrzymamy w końcu przy pewnym m całą przestrzeń:

$$V = V_m = \langle \mathbf{e}_1, \mathcal{J}\mathbf{e}_1, \dots, \mathbf{e}_m, \mathcal{J}\mathbf{e}_m \rangle.$$

Zatem $\dim_{\mathbb{R}} V = 2m$, przy czym w bazie $(\mathbf{e}_1, \mathcal{J}\mathbf{e}_1, \dots, \mathbf{e}_m, \mathcal{J}\mathbf{e}_m)$ macierzą operatora $\mathcal{J}$ jest właśnie (1). Równość (2) wyraża dokładnie fakt, że wektory $\mathbf{e}_k$ i $\mathcal{J}\mathbf{e}_k$ są proporcjonalne nad $\mathbb{C}$: $\mathcal{J}\mathbf{e}_k = i\mathbf{e}_k$. ∎

Powyższe rozumowanie powtarza właściwie procedurę sprowadzania formy antysymetrycznej do postaci kanonicznej (rozdz. 1, § 4).

2. Urzeczywistnienie. Niech teraz U będzie dowolną przestrzenią liniową nad $\mathbb{C}$.

Definicja 3. *Urzeczywistnieniem* przestrzeni U nazywamy rzeczywistą przestrzeń liniową $U_{\mathbb{R}}$, równą U jako zbiór, z tą samą strukturą grupy addytywnej oraz z mnożeniem przez liczby rzeczywiste wykonywanym tak samo jak w U, ale w której „zapomnieliśmy" o mnożeniu przez liczby zespolone (nierzeczywiste).

Urzeczywistniając przestrzeń zespoloną U, zubożamy ją: z n-wymiarowej przestrzeni zespolonej

$$U = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle_{\mathbb{C}}$$

otrzymujemy $2n$-wymiarową przestrzeń rzeczywistą

$$U_{\mathbb{R}} = \langle \mathbf{e}_1, i\mathbf{e}_1, \dots, \mathbf{e}_n, i\mathbf{e}_n \rangle_{\mathbb{R}}.$$

Określone w U mnożenie przez $i = \sqrt{-1}$ staje się teraz strukturą zespoloną $\mathcal{J}$ na $U_{\mathbb{R}}$, określoną wzorami

$$\mathcal{J}\mathbf{e}_k = i\mathbf{e}_k, \qquad \mathcal{J}(i\mathbf{e}_k) = -\mathbf{e}_k, \qquad k = 1, \dots, n. \tag{3}$$

Jako przestrzeń zespoloną wyznaczoną przez tę strukturę zespoloną (zob. punkt 1) otrzymujemy z powrotem wyjściową przestrzeń:

$$\widetilde{U_{\mathbb{R}}} = U.$$

Definicja 4. *Urzeczywistnieniem* operatora $\mathbb{C}$-liniowego $\mathcal{A} : U \to U$ nazywamy operator $\mathbb{R}$-liniowy $\mathcal{A}_{\mathbb{R}} : U_{\mathbb{R}} \to U_{\mathbb{R}}$, działający tak samo jak $\mathcal{A}$, tzn. $\mathcal{A}_{\mathbb{R}}\mathbf{u} = \mathcal{A}\mathbf{u}$.

Różnica między $\mathcal{A}$ i $\mathcal{A}_{\mathbb{R}}$ polega na interpretacji ich działania. Rozpatrzmy rozkład

$$U_{\mathbb{R}} = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle_{\mathbb{R}} + \langle i\mathbf{e}_1, \dots, i\mathbf{e}_n \rangle_{\mathbb{R}}.$$

Każdy operator $\mathbb{C}$-liniowy $\mathcal{A} : U \to U$ wyznacza dwa operatory $\mathbb{R}$-liniowe $\mathcal{A}_1$ i $\mathcal{A}_2$ na $\langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle_{\mathbb{R}}$ takie, że

$$\mathcal{A}\mathbf{u} = \mathcal{A}_1\mathbf{u} + i\mathcal{A}_2\mathbf{u} \quad \text{dla } \mathbf{u} \in \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle_{\mathbb{R}}.$$

Niech A_1 i A_2 będą macierzami operatorów (odpowiednio) $\mathcal{A}_1$ i $\mathcal{A}_2$. Ponieważ

$$\mathcal{A}(i\mathbf{e}_k) = i\mathcal{A}\mathbf{e}_k = i(\mathcal{A}_1\mathbf{e}_k + i\mathcal{A}_2\mathbf{e}_k) = -\mathcal{A}_2\mathbf{e}_k + i\mathcal{A}_1\mathbf{u},$$

więc macierzą operatora $\mathcal{A}_{\mathbb{R}}$ w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n; i\mathbf{e}_1, \ldots, i\mathbf{e}_n)$ przestrzeni $U_{\mathbb{R}}$ jest

$$A_{\mathbb{R}} = \begin{Vmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{Vmatrix}. \tag{4}$$

Widać, że daleko nie wszystkie operatory $\mathbb{R}$-liniowe na $U_{\mathbb{R}}$ są urzeczywistnieniami operatorów $\mathbb{C}$-liniowych na U.

Niech $\mathcal{L}(U)_{\mathbb{R}}$ będzie zbiorem wszystkich takich urzeczywistnień, a $\mathcal{L}(U_{\mathbb{R}})$ przestrzenią wszystkich operatorów $\mathbb{R}$-liniowych na $U_{\mathbb{R}}$. Z definicji lub z postaci macierzowej (4) operatorów urzeczywistnionych wynika, że

$$(\mathcal{A}+\mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} + \mathcal{B}_{\mathbb{R}}, \qquad (\mathcal{A}\mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} \circ \mathcal{B}_{\mathbb{R}}, \qquad (\alpha\mathcal{A})_{\mathbb{R}} = \alpha\mathcal{A}_{\mathbb{R}}, \qquad \alpha \in \mathbb{R}.$$

Inaczej mówiąc, $\mathcal{L}(U)_{\mathbb{R}}$ jest podalgebrą w $\mathcal{L}(U_{\mathbb{R}})$. Oczywiście

$$\dim_{\mathbb{R}} \mathcal{L}(U)_{\mathbb{R}} = 2n^2 = \frac{1}{2}(2n)^2 = \frac{1}{2}\dim_{\mathbb{R}} \mathcal{L}(U_{\mathbb{R}}).$$

Macierzą operatora $\mathcal{J}$ (struktury zespolonej) w rozpatrywanej bazie jest

$$J_0 = \begin{Vmatrix} 0 & -E \\ E & 0 \end{Vmatrix}, \qquad E = E_n. \tag{5}$$

Ponieważ operator $\mathcal{A}$ jest $\mathbb{C}$-liniowy, więc $\mathcal{A}_{\mathbb{R}}\mathcal{J} = \mathcal{J}\mathcal{A}_{\mathbb{R}}$, co odpowiada równości macierzowej $A_{\mathbb{R}}J_0 = J_0A_{\mathbb{R}}$, którą łatwo również sprawdzić bezpośrednio. Na odwrót, jeśli

$$\begin{Vmatrix} A_1 & A_3 \\ A_2 & A_4 \end{Vmatrix} \cdot \begin{Vmatrix} 0 & -E \\ E & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & -E \\ E & 0 \end{Vmatrix} \cdot \begin{Vmatrix} A_1 & A_3 \\ A_2 & A_4 \end{Vmatrix},$$

to po pomnożeniu macierzy blokowych otrzymujemy równość

$$\begin{Vmatrix} A_3 & -A_1 \\ A_4 & -A_2 \end{Vmatrix} = \begin{Vmatrix} -A_2 & -A_4 \\ A_1 & A_3 \end{Vmatrix},$$

skąd $A_3 = -A_2$ i $A_4 = A_1$, tj. każda macierz rzeczywista $2n \times 2n$, która jest przemienna z J_0, ma postać (4).

Udowodniliśmy zatem

STWIERDZENIE 2. *Podalgebra $\mathcal{L}(U)_{\mathbb{R}} \subset \mathcal{L}(U_{\mathbb{R}})$ operatorów urzeczywistnionych składa się dokładnie z operatorów przemiennych z $\mathcal{J}$.* ■

Ciekawe jest następujące pytanie. Niech V będzie parzystowymiarową rzeczywistą przestrzenią liniową (powiedzmy $\dim_{\mathbb{R}} V = 2n$) i niech $\mathcal{A} : V \to V$ będzie operatorem $\mathbb{R}$-liniowym. Kiedy na V istnieje struktura zespolona $\mathcal{J}$ przemienna z $\mathcal{A}$, tj. taka, że $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$, gdzie $\mathcal{B}$ jest operatorem $\mathbb{C}$-liniowym na n-wymiarowej przestrzeni zespolonej $\tilde{V}$ (zob. p. 1)? Rozpatrzmy najpierw ważny przypadek szczególny.

TWIERDZENIE 1. *Niech $V = \mathbb{R}^2$ i niech $\mathcal{A} : V \to V$ będzie operatorem $\mathbb{R}$-liniowym niemającym wektorów własnych. Wtedy na V można określić strukturę zespoloną $\mathcal{J}$ przemienną z $\mathcal{A}$.*

Dowód. Według założenia operator $\mathcal{A}$ ma dwa sprzężone pierwiastki charakterystyczne $\lambda, \overline{\lambda}$. Niech $\lambda = \lambda_1 + i\lambda_2$, gdzie $\lambda_1, \lambda_2 \in \mathbb{R}$, przy czym $\lambda_2 \neq 0$. Na mocy twierdzenia Hamiltona–Cayleya $\mathcal{A}^2 - \operatorname{tr} \mathcal{A} + (\det \mathcal{A})\mathcal{E} = \mathcal{O}$, tj.

$$\mathcal{A}^2 - 2\lambda_1\mathcal{A} + (\lambda_1^2 + \lambda_2^2)\mathcal{E} = \mathcal{O}. \tag{6}$$

Definiujemy

$$\mathcal{J} = \lambda_2^{-1}(\mathcal{A} - \lambda_1\mathcal{E}).$$

Wtedy

$$\mathcal{A} = \lambda_1\mathcal{E} + \lambda_2\mathcal{J}.$$

Podstawiając to wyrażenie do (6), otrzymujemy

$$(\lambda_1^2\mathcal{E} + 2\lambda_1\lambda_2\mathcal{J} + \lambda_2^2\mathcal{J}^2) - 2\lambda_1(\lambda_1\mathcal{E} + \lambda_2\mathcal{J}) + (\lambda_1^2 + \lambda_2^2)\mathcal{E} = \mathcal{O},$$

skąd wynika, że

$$\mathcal{J}^2 = -\mathcal{E}.$$

Zgodnie z ogólnymi rozważaniami w punkcie 2 operator ten wprowadza na V strukturę prostej zespolonej $\mathbb{C}^1$. Ponieważ operator $\mathcal{A}$ jest przemienny z $\mathcal{J}$, więc $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$, gdzie $\mathcal{B} : \mathbb{C}^1 \to \mathbb{C}^1$ jest operatorem mnożenia przez pewną liczbę zespoloną. Tą liczbą jest oczywiście λ. ∎

Udowodnimy teraz

STWIERDZENIE 3. *Jeśli $\mathcal{A} : U \to U$ jest operatorem $\mathbb{C}$-liniowym, to* $\det \mathcal{A}_{\mathbb{R}} = |\det \mathcal{A}|^2$.

Dowód. Przeprowadzimy dowód bezpośredni, stosując postać macierzową (4), operacje elementarne na wyznacznikach nad ciałem $\mathbb{C}$ i równości $A = A_1 + iA_2$ oraz $\det \overline{A} = \overline{\det A}$, gdzie kreska oznacza sprzężenie zespolone:

$$\begin{aligned}\det \mathcal{A}_{\mathbb{R}} &= \det A_{\mathbb{R}} = \det \begin{Vmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{Vmatrix} \\ &= \det \begin{Vmatrix} A_1 + iA_2 & -A_2 + iA_1 \\ A_2 & A_1 \end{Vmatrix} = \det \begin{Vmatrix} A_1 + iA_2 & 0 \\ A_2 & A_1 - iA_2 \end{Vmatrix} \\ &= \det \begin{Vmatrix} A & 0 \\ A_2 & \overline{A} \end{Vmatrix} = \det A \cdot \det \overline{A} = |\det A|^2. \ ∎\end{aligned}$$

3. Kompleksyfikacja. Niech V będzie dowolną przestrzenią liniową nad $\mathbb{R}$. Łatwo sprawdzić bezpośrednio, że jeśli na zewnętrznej sumie prostej $V \oplus V$, czyli przestrzeni liniowej par $(\mathbf{u}, \mathbf{v})$ z działaniami

$$\alpha(\mathbf{u}, \mathbf{v}) + \alpha'(\mathbf{u}', \mathbf{v}') = (\alpha\mathbf{u} + \alpha'\mathbf{u}',\ \alpha\mathbf{v} + \alpha'\mathbf{v}'), \qquad \alpha, \alpha' \in \mathbb{R},$$

określimy operator $\mathcal{J}$ wzorem

$$\mathcal{J} : (\mathbf{u}, \mathbf{v}) \mapsto (-\mathbf{v}, \mathbf{u})$$

to $\mathcal{J}$ wprowadza w $V \oplus V$ strukturę zespoloną, zwaną *kanoniczną*.

DEFINICJA 5. Zespoloną przestrzeń liniową $V \widetilde{\oplus} V$ wyznaczoną przez kanoniczną strukturę zespoloną na $V \oplus V$ nazywamy *kompleksyfikacją* (lub *powłoką zespoloną*) przestrzeni V. Wprowadzamy dla niej specjalne oznaczenie

$$V^{\mathbb{C}} := V \widetilde{\oplus} V.$$

Jeśli patrzeć na $\mathbb{C}$ jak na przestrzeń liniową wymiaru 2 nad $\mathbb{R}$, to

$$V^{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$$

jest szczególnym przypadkiem iloczynu tensorowego przestrzeni — ogólnej konstrukcji, szeroko stosowanej w matematyce (wspominaliśmy o niej w rozdziale 1 (§ 4); omówimy ją dokładnie w rozdziale 6).

Jeśli przestrzeń V jest skończenie wymiarowa i $\dim_{\mathbb{R}} V = n$, to ponieważ $\dim_{\mathbb{R}}(V \oplus V) = 2n$, więc z równości (2) wynika, że

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V.$$

Z definicji $i(\mathbf{u}, \mathbf{v}) = \mathcal{J}(\mathbf{u}, \mathbf{v}) = (-\mathbf{v}, \mathbf{u})$, więc $(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \mathbf{0}) + i(\mathbf{v}, \mathbf{0})$. Dlatego naturalne jest oznaczanie pary $(\mathbf{u}, \mathbf{v})$ symbolem $\mathbf{u} + i\mathbf{v}$. Wówczas

$$(\mathbf{u} + i\mathbf{v}) + (\mathbf{u}' + i\mathbf{v}') = (\mathbf{u} + \mathbf{u}') + i(\mathbf{v} + \mathbf{v}')$$

oraz

$$(\alpha + i\beta)(\mathbf{u} + i\mathbf{v}) = (\alpha\mathbf{u} - \beta\mathbf{v}) + (\alpha\mathbf{v} + \beta\mathbf{u}), \qquad \alpha, \beta \in \mathbb{R},$$

ponieważ

$$(\alpha\mathcal{E} + \beta\mathcal{J})(\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{u}, \mathbf{v}) + \beta(-\mathbf{v}, \mathbf{u}) = (\alpha\mathbf{u} - \beta\mathbf{v},\ \alpha\mathbf{v} + \beta\mathbf{u}).$$

Wzory te łatwo zapamiętać, gdyż odpowiadają dokładnie regułom działań na liczbach zespolonych. Wektory $\mathbf{u} + i\mathbf{0}$ będziemy oznaczać po prostu przez $\mathbf{u}$: w ten sposób przestrzeń rzeczywista V staje się podzbiorem swojej kompleksyfikacji $V^{\mathbb{C}}$.

DEFINICJA 6. *Kompleksyfikacją* operatora $\mathbb{R}$-liniowego $\mathcal{A} : V \to V$ nazywamy operator $\mathbb{C}$-liniowy $\mathcal{A}^{\mathbb{C}} : V^{\mathbb{C}} \to V^{\mathbb{C}}$ określony wzorem

$$\mathcal{A}^{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = \mathcal{A}\mathbf{u} + i\mathcal{A}\mathbf{v}.$$

Zauważmy, że zachodzi implikacja

$$V = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle_{\mathbb{R}} \Rightarrow V^{\mathbb{C}} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle_{\mathbb{C}}.$$

Wynika stąd, że macierz A operatora $\mathcal{A}$ w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest jednocześnie macierzą operatora $\mathcal{A}^{\mathbb{C}}$ w tej samej bazie, tj.

$$A^{\mathbb{C}} = A.$$

W szczególności $\det \mathcal{A}^{\mathbb{C}} = \det \mathcal{A}$ i $\operatorname{tr} \mathcal{A}^{\mathbb{C}} = \operatorname{tr} \mathcal{A}$. Ponieważ

$$\begin{aligned}(\mathcal{A}+\mathcal{B})^{\mathbb{C}}(\mathbf{u}+i\mathbf{v}) &= (\mathcal{A}+\mathcal{B})\mathbf{u} + i(\mathcal{A}+\mathcal{B})\mathbf{v} \\ &= (\mathcal{A}\mathbf{u}+\mathcal{B}\mathbf{u}) + i(\mathcal{A}\mathbf{v}+\mathcal{B}\mathbf{v}) = (\mathcal{A}\mathbf{u}+i\mathcal{A}\mathbf{v}) + (\mathcal{B}\mathbf{u}+i\mathcal{B}\mathbf{v}) \\ &= \mathcal{A}^{\mathbb{C}}(\mathbf{u}+i\mathbf{v}) + \mathcal{B}^{\mathbb{C}}(\mathbf{u}+i\mathbf{v}) = (\mathcal{A}^{\mathbb{C}}+\mathcal{B}^{\mathbb{C}})(\mathbf{u}+i\mathbf{v}),\end{aligned}$$

więc $(\mathcal{A}+\mathcal{B})^{\mathbb{C}} = \mathcal{A}^{\mathbb{C}} + \mathcal{B}^{\mathbb{C}}$. Podobnie sprawdza się, że $(\mathcal{A}\mathcal{B})^{\mathbb{C}} = \mathcal{A}^{\mathbb{C}}\mathcal{B}^{\mathbb{C}}$.

Przez analogię do definicji $\mathcal{A}^{\mathbb{C}}$, również formy liniowe, i ogólnie wieloliniowe, można przedłużyć z V na $V^{\mathbb{C}}$. Jeśli np. f jest formą dwuliniową na przestrzeni rzeczywistej V, to definiujemy

$$f^{\mathbb{C}}(\mathbf{x}+i\mathbf{y}, \mathbf{u}+i\mathbf{v}) = f(\mathbf{x},\mathbf{u}) - f(\mathbf{y},\mathbf{v}) + i(f(\mathbf{x},\mathbf{v}) + f(\mathbf{y},\mathbf{u})).$$

Jako ćwiczenie Czytelnik zechce sprawdzić np., że z antysymetryczności f wynika antysymetryczność $f^{\mathbb{C}}$.

Niech teraz V będzie rzeczywistą przestrzenią liniową z iloczynem skalarnym $(* \mid *)$. Wówczas na $V^{\mathbb{C}}$ można określić iloczyn skalarny wzorem

$$(\mathbf{x}+i\mathbf{y} \mid \mathbf{u}+i\mathbf{v})^{\mathbb{C}} := (\mathbf{x}\mid\mathbf{u}) + (\mathbf{y}\mid\mathbf{v}) - i((\mathbf{x}\mid\mathbf{v}) - (\mathbf{u}\mid\mathbf{y})).$$

W ten sposób jeśli $(V, (*\mid*))$ jest przestrzenią euklidesową, to $(V^{\mathbb{C}}, (*\mid*)^{\mathbb{C}})$ staje się przestrzenią unitarną. W szczególności norma $\|*\|^{\mathbb{C}}$ w $V^{\mathbb{C}}$ jest określona wzorem

$$(\|\mathbf{x}+i\mathbf{y}\|^{\mathbb{C}})^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Wracając do przypadku ogólnego, załóżmy, że $\mathcal{A}$ jest operatorem liniowym na V, zaś $\mathbf{a}+i\mathbf{b}$ — wektorem własnym operatora $\mathcal{A}^{\mathbb{C}}$ na $V^{\mathbb{C}}$ odpowiadającym wartości własnej $\alpha + i\beta$ ($\mathbf{a}, \mathbf{b} \in V$, $\alpha, \beta \in \mathbb{R}$). Wtedy, zgodnie z przyjętymi definicjami, $\mathcal{A}\mathbf{a}+i\mathcal{A}\mathbf{b} = \mathcal{A}^{\mathbb{C}}(\mathbf{a}+i\mathbf{b}) = (\alpha+i\beta)(\mathbf{a}+i\mathbf{b}) = (\alpha\mathbf{a}-\beta\mathbf{b})+i(\beta\mathbf{a}+\alpha\mathbf{b})$, tj.

$$\mathcal{A}\mathbf{a} = \alpha\mathbf{a} - \beta\mathbf{b}, \qquad \mathcal{A}\mathbf{b} = \beta\mathbf{a} + \alpha\mathbf{b}.$$

Tak więc $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{R}}$ jest dwuwymiarową podprzestrzenią $\mathcal{A}$-niezmienniczą w V. Ponieważ operator $\mathcal{A}^{\mathbb{C}}$ ma zawsze wektor własny, udowodniliśmy jeszcze raz twierdzenie 7 z rozdziału 2 (§ 3).

Zauważmy ponadto, że każda zespolona przestrzeń liniowa U wymiaru n jest izomorficzna z kompleksyfikacją $V^{\mathbb{C}}$ pewnej rzeczywistej przestrzeni liniowej V. Wystarczy wybrać w U jakąkolwiek bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ i określić V jako zbiór wszystkich kombinacji liniowych $\sum_j \alpha_j \mathbf{e}_j$ z $\alpha_j \in \mathbb{R}$:

$$U = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle_{\mathbb{C}} = (\langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle_{\mathbb{R}})^{\mathbb{C}}.$$

4. Kompleksyfikacja → urzeczywistnienie → kompleksyfikacja. Niech

$$W = (V^{\mathbb{C}})_{\mathbb{R}}$$

oznacza $2n$-wymiarową przestrzeń rzeczywistą otrzymaną z n-wymiarowej przestrzeni rzeczywistej V w wyniku kompleksyfikacji, a następnie urzeczywistnienia. Nietrudno zauważyć, że

$$W = V \oplus iV; \tag{7}$$

podprzestrzeń $V \subset W$ nazywamy *składową rzeczywistą*, a iV – *składową urojoną*. Zgodnie ze wzorem (3) na przestrzeni W określony jest operator $\mathcal{J} = (i\mathcal{E})_{\mathbb{R}}$, urzeczywistnienie operatora $i\mathcal{E}$ mnożenia przez i na $V^{\mathbb{C}}$. Jego macierzą jest J_0 (zob. (5)). Operator $\mathcal{J}$ zamienia miejscami składową rzeczywistą i urojoną.

Rozpatrzmy najprostszy przypadek, gdy $V = \mathbb{R}$, $V^{\mathbb{C}} = \mathbb{C}^1$ i $W = (\mathbb{C}^1)_{\mathbb{R}} = \mathbb{R}^2$. Na prostej zespolonej określona jest operacja sprzężenia zespolonego

$$\alpha + i\beta \mapsto \overline{\alpha + i\beta} = \alpha - i\beta.$$

W ogólnym przypadku na przestrzeni (7) działa analogiczny operator

$$\mathcal{S} : \mathbf{u} + i\mathbf{v} \mapsto \overline{\mathbf{u} + i\mathbf{v}} = \mathbf{u} - i\mathbf{v}$$

o macierzy

$$S = \begin{Vmatrix} E_n & 0 \\ 0 & -E_n \end{Vmatrix}.$$

Uogólniając tę sytuację, rozważmy dowolny operator $\mathbb{C}$-liniowy $\mathcal{A} : V^{\mathbb{C}} \to V^{\mathbb{C}}$ (a nawet, w razie potrzeby, odwzorowanie liniowe $V_1^{\mathbb{C}} \to V_2^{\mathbb{C}}$, gdzie $V_1 \neq V_2$). *Sprzężeniem zespolonym* operatora $\mathcal{A}$ nazywamy operator $\overline{\mathcal{A}} : V^{\mathbb{C}} \to V^{\mathbb{C}}$ określony wzorem

$$\overline{\mathcal{A}}(\mathbf{u} + i\mathbf{v}) = \overline{\mathcal{A}(\mathbf{u} - i\mathbf{v})}.$$

Wtedy

$$(\overline{\mathcal{A}})_{\mathbb{R}} = \mathcal{S}\mathcal{A}_{\mathbb{R}}\mathcal{S}.$$

$\mathbb{C}$-liniowość operatora $\overline{\mathcal{A}}$ jest prostą konsekwencją $\mathbb{C}$-liniowości $\mathcal{A}$ i $\mathbb{R}$-liniowości operatora sprzężenia zespolonego. Jeśli $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest bazą w V, to jest również bazą w $V^{\mathbb{C}}$; w tej bazie operator $\mathcal{A}$ ma macierz $A = A_1 + iA_2$, a operator $\overline{\mathcal{A}}$ macierz $\overline{A} = A_1 - iA_2$, gdzie macierze A_1 i A_2 są rzeczywiste (zob. p. 2). Wynika stąd, że warunek $\mathcal{A} = \overline{\mathcal{A}}$ jest konieczny i dostateczny na to, by operator $\mathcal{A}$ był kompleksyfikacją $\mathcal{B}^{\mathbb{C}}$ pewnego operatora rzeczywistego $\mathcal{B} : V \to V$.

Wykorzystując pojęcie sprzężenia zespolonego operatorów, możemy dla dowolnego operatora postaci $\mathcal{A}_{\mathbb{R}}$ o macierzy (4) odnotować równość

$$\operatorname{tr} \mathcal{A}_{\mathbb{R}} = 2 \operatorname{tr} A_1 = \operatorname{tr}(\mathcal{A} + \overline{\mathcal{A}}) = \operatorname{tr} \mathcal{A} + \operatorname{tr} \overline{\mathcal{A}}.$$

Niech U będzie dowolną zespoloną przestrzenią liniową. W rozdziale 1 (§ 1, p. 1) rozpatrywaliśmy przestrzeń $\overline{U}$, którą nazwiemy teraz *przestrzenią sprzężoną do U w sensie sprzężenia zespolonego*, a która różni się od U tylko mnożeniem przez skalary: $\lambda \odot \mathbf{x} = \overline{\lambda}\mathbf{x}$.

Analogicznie, jeśli $(V, \mathcal{J})$ jest rzeczywistą przestrzenią liniową ze strukturą zespoloną, to operator $-\mathcal{J}$ również wprowadza w V strukturę zespoloną, zwaną *sprzężoną* do poprzedniej. Jeśli przy tym $\widetilde{V}$ jest przestrzenią zespoloną wyznaczoną przez $(V, \mathcal{J})$, to $\overline{\widetilde{V}}$ jest przestrzenią zespoloną wyznaczoną przez $(V, -\mathcal{J})$.

Stosując teraz do zespolonej przestrzeni liniowej V najpierw funktor urzeczywistnienia, a następnie kompleksyfikacji, otrzymamy kanoniczny $\mathbb{C}$-liniowy izomorfizm

$$f : (V_{\mathbb{R}})^{\mathbb{C}} \to V \oplus \overline{V}.$$

Zauważmy mianowicie, że na $(V_{\mathbb{R}})^{\mathbb{C}}$ mamy dwa operatory $\mathbb{R}$-liniowe: operator *kanonicznej* struktury zespolonej $\mathcal{J}(\mathbf{x}, \mathbf{y}) = (-\mathbf{y}, \mathbf{x})$ i operator mnożenia przez $i = \sqrt{-1}$, odpowiadający wyjściowej strukturze zespolonej na V: $i(\mathbf{x}, \mathbf{y}) = (i\mathbf{x}, i\mathbf{y})$. Ponieważ operator $\mathcal{J}$ jest przemienny z i, więc jest on operatorem $\mathbb{C}$-liniowym względem wyjściowej struktury. Skoro $\mathcal{J}^2 = -\mathcal{E}$, to wartościami własnymi operatora $\mathcal{J}$ są ± 1. Wprowadźmy standardowe oznaczenia dla tych dwóch podprzestrzeni własnych:

$$V^{1,0} = \{(\mathbf{x}, \mathbf{y}) \in (V_{\mathbb{R}})^{\mathbb{C}} \mid \mathcal{J}(\mathbf{x}, \mathbf{y}) = i(\mathbf{x}, \mathbf{y})\},$$
$$V^{0,1} = \{(\mathbf{x}, \mathbf{y}) \in (V_{\mathbb{R}})^{\mathbb{C}} \mid \mathcal{J}(\mathbf{x}, \mathbf{y}) = -i(\mathbf{x}, \mathbf{y})\}.$$

Oba te zbiory są też podprzestrzeniami zespolonymi w $(V_{\mathbb{R}})^{\mathbb{C}}$, tj. względem struktury zespolonej wyznaczonej przez $\mathcal{J}$: jest jasne, że są one zamknięte względem dodawania i mnożenia przez liczby rzeczywiste, a zamkniętość względem $\mathcal{J}$ wynika z faktu, że $\mathcal{J}$ jest przemienne z i. Pokażemy, że $V = V^{1,0} \oplus V^{0,1}$ oraz że $V^{1,0}$ jest w naturalny sposób izomorficzne z V, podczas gdy $V^{0,1}$ jest izomorficzne z $\overline{V}$.

Z przyjętych definicji wynika od razu, że $V^{1,0}$ składa się z wektorów postaci $(\mathbf{x}, -i\mathbf{x})$, a $V^{0,1}$ — z wektorów $(\mathbf{y}, i\mathbf{y})$. Dla danych $\mathbf{u}, \mathbf{v} \in V$ równanie $(\mathbf{u}, \mathbf{v}) = (\mathbf{x}, -i\mathbf{x}) + (\mathbf{y}, i\mathbf{y})$ ma jednoznaczne rozwiązanie $\mathbf{x} = (\mathbf{u} + i\mathbf{v})/2$, $\mathbf{y} = (\mathbf{u} - i\mathbf{v})/2$. Zatem $V = V^{1,0} \oplus V^{0,1}$. Odwzorowania $\mathbf{x} \mapsto (\mathbf{x}, -i\mathbf{x})$ oraz $\mathbf{x} \mapsto (\mathbf{x}, i\mathbf{x})$ są $\mathbb{R}$-liniowymi izomorfizmami V na $V^{1,0}$ i (odpowiednio) $\overline{V}$ na $V^{0,1}$. Ponieważ są one ponadto przemienne z działaniem i na V i $\overline{V}$ oraz z działaniem $\mathcal{J}$ na $V^{1,0}$ i $V^{0,1}$ na mocy przyjętych definicji, nasza konstrukcja jest zakończona.

ĆWICZENIA

1. Wykazać, że operator ortogonalny $\mathcal{A}$ na przestrzeni euklidesowej V, niemający wektorów własnych (co jest możliwe tylko w przypadku $\dim V = 2m$) jest urzeczywistnieniem operatora unitarnego $\mathcal{B}$ na przestrzeni zespolonej U wymiaru m, związanej z V. Zauważmy w związku z tym, że urzeczywistnienie przestrzeni unitarnej prowadzi do przestrzeni euklidesowej dwa razy większego wymiaru.

2. Udowodnić równość ze stwierdzenia 3, wybierając w przestrzeni zespolonej U bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, w której macierz A operatora $\mathcal{A} : U \to U$ jest górnotrójkątna z $\lambda_1, \ldots, \lambda_n$ na przekątnej.

3. Udowodnić samosprzężoność operatora różniczkowego S bezpośrednio, nie korzystając z ogólnego wzoru (**) z punktu 5.

4. Udowodnić, że

$$f(x,t)=\frac{1}{\sqrt{1-2tx+x^2}}$$

jest funkcją tworzącą dla wielomianów Legendre'a (tzn. $P_n(t)$ jest współczynnikiem przy x^n w rozwinięciu $f(x,t)$ w szereg względem potęg x).

5. Udowodnić, że

$$\max_{-1\leqslant t\leqslant 1}\left|\frac{1}{2^{n-1}}T_n(t)\right|=\frac{1}{2^{n-1}}.$$

6. Udowodnić, że ***wielomiany Czebyszewa drugiego rodzaju***

$$U_n(t):=\frac{1}{n+1}\frac{d}{dt}T_{n+1}(t)=\sum_{m=0}^{[n/2]}\binom{n+1}{2m+1}t^{n-2m}(t^2-1)^m$$

są ortogonalne w przedziale $[-1,1]$ z wagą $\sqrt{1-t^2}$:

$$\int_{-1}^{1}\sqrt{1-t^2}\,U_m(t)U_n(t)\,dt=\begin{cases}0 & \text{dla } m\neq n,\\ \pi/2 & \text{dla } m=n.\end{cases}$$

7. Udowodnić, że wszystkie pierwiastki wielomianów $T_n(t)$ i $U_n(t)$ są rzeczywiste, jednokrotne i leżą w przedziale $[-1,1]$. Podać jawne wzory na te pierwiastki.