VI MACIERZE ENDOMORFIZMÓW LINIOWYCH („ELEMENTY TEORII SPEKTRALNEJ")

Wstęp.* W tym rozdziale zajmujemy się następującymi ważnymi zagadnieniami: jak znaleźć bazę $\mathcal{V}$ przestrzeni V, w której macierz zadanego operatora liniowego $L : V \to V$ ma możliwie prostą postać? Jak prosta może być ta postać? Gdy przestrzeń jest unitarna, to czy bazę tę można obrać ortonormalną?

Oczywiście, najbardziej pożądane jest, by macierz $[L]_{\mathcal{V}}$ była diagonalna, i niektóre wyniki z §2 i §3 zmierzają do ustalenia, kiedy taka baza $\mathcal{V}$ istnieje i jak ją znaleźć. Badamy to zagadnienie dokładniej gdy przestrzeń jest unitarna, zaś baza ma być ortonormalna. (Ciałem skalarów jest wówczas $\mathbb{R}$ lub $\mathbb{C}$.) Okazuje się, że jeśli operator jest samosprzężony (spełnia warunek $L^* = L$), to żądana baza istnieje – jest to wynik o wielu zastosowaniach w analizie i geometrii. Podamy nawet przejrzysty warunek charakteryzujący operatory „unitarnie diagonalizowalne", czyli te, dla których szukana baza ortonormalna istnieje.

Gdy macierzy zadanego operatora w żadnej bazie nie możemy uczynić diagonalną, to pokusić się można o zbadanie, czy możemy ją uczynić trójkątną i jak prosta może ta macierz trójkątna być. Tu na plan pierwszy wybijają się dwa twierdzenia, oba dotyczące zespolonego ciała skalarów. Pierwsze z nich pochodzi od I. Schura i orzeka, że każdy operator liniowy z $\mathbb{C}^n$ do $\mathbb{C}^n$ ma macierz trójkątną w pewnej ortonormalnej bazie przestrzeni $\mathbb{C}^n$. Drugie zaś pochodzi od C. Jordana i orzeka, że macierz tę możemy nawet uczynić bardzo bliską diagonalnej – lecz wybrana baza niekoniecznie będzie ortonormalna

Wyniki tu uzyskane mają liczne zastosowania. Na przykład, pozwalają one efektywnie wyzaczyć wartość funkcji (np. wielomianowej) na danej macierzy, co omawiamy zarówno w §2, jak i w §3. Równie ważne jest to, że umożliwiają one wniknięcie we własności rozważanego operatora. Komentarz na ten temat znaleźć można w końcowych punktach §3 i w zamykającej rozdział „próbie podsumowania".

Umowy: Rozważamy tu tylko przestrzenie wektorowe skończonego wymiaru. W odniesieniu do przestrzeni unitarnych (rzeczywistych czy zespolonych) używamy wymiennie nazw „izometria liniowa" i „przekształcenie unitarne" czy „izomorfizm unitarny", a w przypadku rzeczywistym zamiast „unitarny" mówimy też „ortogonalny". Dla krótkości piszemy

$$\mathcal{L}(V) := \mathcal{L}(V, V) \ \text{ i } \ [L]_{\mathcal{V}} := [L]_{\mathcal{V}}^{\mathcal{V}} \text{ dla bazy } \mathcal{V} \text{ przestrzeni } V \text{ i } L \in \mathcal{L}(V).$$

Przekształcenia $L \in \mathcal{L}(V)$ nazywamy **endomorfizmami** (liniowymi) przestrzeni V. Często zamiast „przekształcenie liniowe" mówimy „operator liniowy", a nierzadko też słowa „liniowy" czy „liniowe" pomijamy, bo nieliniowych operatorów nie rozpatrujemy.

§ 1. Podobieństwo macierzy bądź operatorów.

1. Podstawowe definicje.

Definicja. a) Operatory $K \in \mathcal{L}(V)$, $L \in \mathcal{L}(W)$ są **podobne**, jeśli istnieje izomorfizm liniowy $S : V \to W$ taki, że $L = SKS^{-1}$. (Tu, V i W to przestrzenie liniowe.)

b) Macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ są podobne, gdy podobne są operatory $L_{\mathbf{A}}, L_{\mathbf{B}} \in \mathcal{L}(\mathbb{F}^k)$.

Niejednokrotnie zależy nam na tym, by powyższy izomorfizm $S : V \to W$ miał jeszcze dodatkowe własności. Najważniejszy przypadek ujmuje następująca

Definicja. Niech V i W będą przestrzeniami unitarnymi. Operatory $K \in \mathcal{L}(V)$ i $L \in \mathcal{L}(W)$ są **unitarnie podobne**, gdy istnieje unitarny izomorfizm $S : V \to W$ taki, że $SKS^{-1} = L$. Tak samo, dla $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ nazwiemy unitarnie podobnymi nad $\mathbb{F}$, jeśli operatory $L_\mathbf{A}, L_\mathbf{B} \in \mathcal{L}(\mathbb{F}^k)$ są unitarnie podobne. (Przestrzeń $\mathbb{F}^k$ rozpatrujemy ze standardowym iloczynem skalarnym.) Zamiast „unitarnie podobne nad $\mathbb{R}$” mówimy też „**ortogonalnie podobne**”.

Lemat 1. *a) Macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ wtedy i tylko wtedy są podobne, gdy $\mathbf{B} = \mathbf{SAS}^{-1}$ dla pewnej nieosobliwej macierzy $\mathbf{S} \in \mathcal{M}_k(\mathbb{F})$.*

b) Macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$, gdzie $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, wtedy i tylko wtedy są unitarnie podobne nad $\mathbb{F}$, gdy $\mathbf{B} = \mathbf{SAS}^{-1}$ dla pewnej unitarnej macierzy $\mathbf{S} \in \mathcal{M}_k(\mathbb{F})$.

Dowód. Dla $F, G, S \in \mathcal{L}(\mathbb{F}^k)$ równość $G = SFS^{-1}$ ma miejsce wtedy i tylko wtedy, gdy $[G] = [S][F][S]^{-1}$. Gdy $F = L_\mathbf{A}$ i $G = L_\mathbf{B}$, to $[F] = \mathbf{A}$ i $[G] = \mathbf{B}$, co daje a).

By dowieść b) korzystamy jeszcze z tego, że unitarność operatora $S : \mathbb{F}^k \to \mathbb{F}^k$ jest równoważna unitarności jego macierzy $[S]$. $\square$

Zadanie 1. Podobieństwo endomorfizmów liniowych czy podobieństwo macierzy są relacjami równoważności, i tak samo jest dla unitarnego podobieństwa endomorfizmów przestrzeni unitarnych czy unitarnego podobieństwa macierzy nad $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Zadanie 2. Jeśli macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ są podobne, to podobne są też macierze $\mathbf{A}^t$ i $\mathbf{B}^t$, jak również macierze $p(\mathbf{A})$ i $p(\mathbf{B})$, dla $p \in \mathbb{F}[x]$; a gdy $\mathbb{F} = \mathbb{C}$, to także macierze $\mathbf{A}^*$ i $\mathbf{B}^*$ są podobne.

Zadania uzupełniające.

1. Każda 2×2-macierz $\mathbf{A}$ jest podobna do macierzy $\mathbf{A}^t$. (Tak samo większa macierz, lecz dowód jest trudny.)

2. Zdefiniować podobieństwo permutacji $\sigma, \tau \in \mathbf{S}_k$ i dowieść, że ma ono miejsce wtedy i tylko wtedy, gdy w przedstawieniu σ w postaci iloczynu cykli rozłącznych występuje dla każdego n tyle cykli długości n, co w przedstawieniu τ w postaci takiego iloczynu.

2. Pewne niezmienniki podobieństwa.

Ważne i wielokrotnie wykorzystywane dalej jest to, że podobne macierze mają zbliżone własności algebraiczne. Ograniczymy się do następującej ilustracji tego stwierdzenia:

Zadanie 1. Gdy macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ są podobne, to
a) dla każdej liczby $n \in \mathbb{N}$, $(\mathbf{A}^n = \mathbf{0}) \Leftrightarrow (\mathbf{B}^n = \mathbf{0})$ oraz $(\mathbf{A}^n = \mathbf{I}) \Leftrightarrow (\mathbf{B}^n = \mathbf{I})$.
b) $\det(\mathbf{A}) = \det(\mathbf{B})$ i $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{B})$.
c) $\operatorname{rk}(\mathbf{A}) = \operatorname{rk}(\mathbf{B})$.

Skrótowo, stwierdzenia od te wyrazić możemy mówiąc, że rozważane w nich własności $\mathbf{A}^n = \mathbf{0}$, $\mathbf{A}^n = \mathbf{I}$, $\det(\mathbf{A}) = \lambda$, $\operatorname{tr}(\mathbf{A}) = \lambda$ oraz $\operatorname{rk}(\mathbf{A}) = n$ są **niezmiennikami podobieństwa macierzy**.

Ćwiczenie. Udowodnić, że posiadanie pierwiastka kwadratowego jest niezmiennikiem podobieństwa: jeśli macierze $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k$ są podobne i $\mathbf{A} = \mathbf{X}^2$ dla pewnej macierzy $\mathbf{X}$, to $\mathbf{B} = \mathbf{Y}^2$ dla pewnej macierzy $\mathbf{Y}$.

Ważne własności podobnych operatorów niejednokrotnie też są takie same:

Zadania.

2. Niech operatory $K \in \mathcal{L}(V)$ i $L \in \mathcal{L}(W)$ oraz izomorfizm $S \in \mathcal{L}(V, W)$ spełniają warunek $K = S^{-1}LS$. Dowieść, że

a) dla wielomianów $p \in \mathbb{F}[x]$ zachodzi $p(K) = S^{-1}p(L)S$;

b) $\ker(L) = S(\ker(K)$ i $\mathrm{im}(L) = S(\mathrm{im}(K))$, i tak samo z $p(K)$ i $p(L)$ w miejsce K i L, odp.

c)* Jeśli $V = \oplus_{i=1}^{t} \ker(p_i(K))$, gdzie $p_1, ..., p_t \in \mathbb{F}[x]$, to $W = \oplus_{i=1}^{t} \ker(p_i(L))$.

3. a) Jeśli jeden z operatorów podobnych P, Q jest rzutem, to drugi też. Ściślej, gdy P jest rzutem przestrzeni V na X wzdłuż Y, a $S : V \to W$ jest izomorfizmem takim, że $Q = SPS^{-1}$, to Q jest rzutem przestrzeni W na $S(X)$ wzdłuż $S(Y)$.
b) Jeśli jeden z unitarnie podobnych operatorów P, Q działających w przestrzeniach unitarnych jest rzutem ortogonalnym, to drugi też.
c) Sformułować i udowodnić analogiczne stwierdzenia dla symetrii względem podprzestrzeni.

4. Jeśli operatory $K, L \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ są unitarnie podobne i jeden z nich jest obrotem, to drugi też.

5. Gdy operatory K, L na przestrzeni unitarnej V są unitarnie podobne i $\langle K(\mathbf{v}), \mathbf{v}\rangle \geq 0$ dla każdego $\mathbf{v} \in V$, to $\langle L(\mathbf{v}), \mathbf{v}\rangle \geq 0$ dla każdego $\mathbf{v} \in V$.

Tak więc o własnościach jednej z dwóch macierzy podobnych wiele można powiedzieć, gdy znamy własności drugiej, i tak samo jest z operatorami. Niestety, pytanie, czy zadane dwie macierze (równoważnie: dwa operatory) są podobne jest na ogół trudne. Dla $\mathbb{F} = \mathbb{C}$, odpowiedź na nie umożliwią dopiero wyniki z §2.2.

Zadanie uzupełniające 1. Wyposażmy przestrzeń $\mathcal{M}_k(\mathbb{C})$ w normę wyznaczoną przez iloczyn skalarny $\langle \mathbf{X}, \mathbf{Y}\rangle = \mathrm{tr}(\mathbf{X}\mathbf{Y}^*)$. Udowodnić, że $||\mathbf{A}|| = ||\mathbf{B}||$ dla unitarnie podobnych macierzy $\mathbf{A}, \mathbf{B} \in \mathbf{M}_k(\mathbb{C})$.

3. Podobieństwo a zapis operatora w bazie.

Jednym z powodów znaczenia podobieństwa macierzy jest to, że pojawia się ono, gdy badamy macierze operatora w różnych bazach.

Stwierdzenie 1. *Niech* $\mathbf{A}$ *będzie macierzą operatora* $L \in \mathcal{L}(V)$ *w bazie* $\mathcal{V}$ *przestrzeni* V*. Wówczas następujące warunki są równoważne:*

a) $\mathbf{B} = [L]_{\mathcal{W}}$ *dla pewnej bazy* $\mathcal{W}$ *przestrzeni* V*;*

b) macierz $\mathbf{B}$ *jest podobna do* $\mathbf{A}$*.*

Dowód. By dowieść implikacji b) $\Rightarrow$ a) oznaczmy przez $\mathbf{S}$ macierz nieosobliwą, dla której $\mathbf{S}^{-1}[L]_{\mathcal{V}}\mathbf{S} = \mathbf{B}$. Baza $\mathcal{W}$ przestrzeni V, dla której $[I]_{\mathcal{V}}^{\mathcal{W}} = \mathbf{S}$, spełnia zarazem warunek $[L]_{\mathcal{W}} = \mathbf{B}$. (Korzystamy z wyników §III.2.2: wniosku 1 i twierdzenia 1.) Podobnie dowodzimy implikacji a)$\Rightarrow$b). $\square$

Wersja stwierdzenia 1 odgrywa też rolę, gdy $V = W = \mathbb{F}^k$.

Wniosek 1. *Dla macierzy* $\mathbf{A}, \mathbf{B} \in \mathbb{F}^k$ *równoważne są warunki:*

a) $\mathbf{B}$ *jest macierzą operatora* $L_{\mathbf{A}}$ *w bazie* $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ *przestrzeni* $\mathbb{F}^k$*;*

b) $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ *dla macierzy* $\mathbf{C}$*, której kolejnymi kolumnami są* $\mathbf{v}_1, ..., \mathbf{v}_k$*.*

Dowód. Oznaczmy przez $\mathcal{E} = (\mathbf{e}_1, ..., \mathbf{e}_k)$ standardową bazę przestrzeni $\mathbb{F}^k$. Mamy $[L_{\mathbf{A}}]_{\mathcal{V}} = ([I]_{\mathcal{E}}^{\mathcal{V}})^{-1}[L_{\mathbf{A}}]_{\mathcal{E}}[I]_{\mathcal{E}}^{\mathcal{V}}$ oraz $[I]_{\mathcal{E}}^{\mathcal{V}} = \mathbf{C}$, $[L_{\mathbf{A}}]_{\mathcal{E}} = \mathbf{A}$. $\square$

Uwaga 1. Przekształcenie pewnej macierzy $\mathbf{A}$ przez podobieństwo tak, by otrzymać macierz zadanej postaci, jest więc równoważne znalezieniu bazy przestrzeni $\mathbb{F}^k$, w której macierz operatora $L_{\mathbf{A}}$ jest tej postaci. Ponadto, gdy $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$ i chcemy by podobieństwo było unitarne, należy żądać, by baza była ortonormalna. Ten punkt widzenia okaże się dalej bardzo użyteczny.

Ważną rolę odgrywa też następująca

Uwaga 2. Niech $\mathbf{A} = [L]_{\mathcal{V}}$ będzie macierzą operatora $L \in \mathcal{L}(V)$ w bazie $\mathcal{V}$ przestrzeni V. Z definicji, $L_{\mathbf{A}} = SLS^{-1}$, gdzie $S : V \to \mathbb{F}^k$ jest mapą wyznaczoną przez $\mathcal{V}$. W szczególności, operator L jest podobny do operatora $L_{\mathbf{A}} : \mathbb{F}^k \to \mathbb{F}^k$ o macierzy $\mathbf{A}$. $\square$

Wniosek 2. *Niech $\mathcal{V}$ będzie bazą przestrzeni V, zaś $\mathcal{W}$ bazą przestrzeni W. Dane operatory $K \in \mathcal{L}(V)$ i $L \in \mathcal{L}(W)$ wtedy i tylko wtedy są podobne, gdy podobne są ich macierze $\mathbf{A} := [K]_{\mathcal{V}}$ i $\mathbf{B} := [L]_{\mathcal{W}}$.*

Dowód. Z uwagi 2 i przechodniości podobieństwa wynika, że $K \sim L \Leftrightarrow L_{\mathbf{A}} \sim L_{\mathbf{B}}$. $\boxminus$.

Uwaga 3. Ustalenia te pozostają prawdziwe, gdy w sformułowaniach zastąpić słowa „baza” przez „baza ortonormalna”, a „podobieństwo” przez „podobieństwo unitarne”. (By się o tym przekonać, wystarcza tak samo zmienić dowody i wykorzystać to, że mapa wyznaczona przez bazę ortonormalną jest izomorfizmem unitarnym.)

Przykład 1. Niech macierz $\mathbf{B}$ powstaje z macierzy $\mathbf{A} \in \mathcal{M}_k$ przez zamianę pierwszych dwóch kolumn $\mathbf{A}$, a następnie pierwszych dwóch wierszy otrzymanej macierzy. Latwo widzieć, że $\mathbf{B}$ jest macierzą operatora $L_{\mathbf{A}}$ w bazie $(\mathbf{e}_2, \mathbf{e}_1, ..., \mathbf{e}_k)$. Stąd macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne. Równie dobrze możemy zamienić inne dwie kolumny i odpowiadające im dwa wiersze. Ogólniej, jeśli poddamy kolumny, a następnie wiersze macierzy $\mathbf{A}$ tej samej permutacji, to otrzymamy macierz podobną do $\mathbf{A}$. (Wynika to z podanego rozumowania; można też przedstawić permutację w postaci iloczynu transpozycji). Gdy ponadto $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, podobieństwo jest unitarne.

Tak więc podobne są dwie macierze takie, że wzdłuż przekątnej pierwszej stoją pewne kwadratowe klatki, wzdłuż przekątnej drugiej –te same klatki, ale być może w innej kolejności, zaś w obu poza klatkami są zera. $\square$

Przykład 2. Rozważając macierz operatora $L_{\mathbf{A}}$ w bazie $(\mathbf{e}_k, \mathbf{e}_{k-1}, ..., \mathbf{e}_1)$ stwierdzamy, że jest ona podobna do macierzy $\mathbf{B}$ otrzymanej z $\mathbf{A}$ przez „odbicie względem środka” (tzn. jej $b_{ij} = a_{k-j+1,k-i+1}$). Ponadto, dla $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ podobieństwo jest unitarne. W szczególności, każda macierz górnie trójkątna jest podobna do pewnej macierzy dolnie trójkątnej (i to unitarnie, gdy $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$).

Zadanie uzupełniające 1. Niech $\mathbf{E}_{ij}$ oznacza $k \times k$–macierz, której ij–ty wyraz jest równy 1, a pozostałe są równe 0. Przyjmując wpierw $k = 2$ dowieść, że:

a) Macierze $\mathbf{E}_{ii}$ i $\mathbf{E}_{jj}$ są podobne.

b) Gdy $i \neq j$, to macierz $\mathbf{E}_{ij}$ jest podobna do $\lambda\mathbf{E}_{ij}$ dla $\lambda \neq 0$.

c) Gdy $\#\mathbb{F} > 2$ i $f : \mathcal{M}_k(\mathbb{F}) \to \mathbb{F}$ jest funkcją liniową, przyjmującą na każdych dwóch macierzach podobnych tę samą wartość, to $f(\mathbf{A}) = f(\mathbf{E}_{11})\mathrm{tr}(\mathbf{A})$ dla $\mathbf{A} \in \mathcal{M}_k$.

4. Własności operatorów wyznaczone przez niezmienniki podobieństwa macierzy. Wielomian charakterystyczny.

Niech V będzie przestrzenią wektorową nad $\mathbb{F}$ wymiaru k, a f funkcją określona na $\mathcal{M}_k(\mathbb{F})$ (zbiór wartości nie jest istotny). Zapytajmy: czy f wyznacza funkcję $\widetilde{f}$ na zbiorze $\mathcal{L}(V)$ wszystkich operatorów liniowych $V \to V$? Okazuje się, że jest tak gdy **funkcja** f jest **niezmiennicza względem podobieństwa macierzy**, tzn. gdy:

$$f(\mathbf{A}) = f(\mathbf{B}) \text{ dla każdych dwóch podobnych macierzy } \mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F}) \tag{1}$$

Dla operatora $L \in \mathcal{L}(V)$ przyjmiemy bowiem $\widetilde{f}(L) := f([L]_{\mathcal{V}})$, gdzie $\mathcal{V}$ jest bazą przestrzeni V. Wynik nie zależy od $\mathcal{V}$, bo gdy $\mathcal{V}$ i $\mathcal{W}$ są bazami, to macierze $[L]_{\mathcal{V}}$ i $[L]_{\mathcal{W}}$ są podobne (stw. 1 w p.3).

Zadanie 1. Zachodzi $\widetilde{f}(K) = \widetilde{f}(L)$ dla podobnych operatorów $K \in \mathcal{L}(V), L \in \mathcal{L}(W)$.

(Materiał poprzedzający następną definicję można uważać za uzupełniający.) Własność macierzy możemy traktować jako funkcję w zbiór dwu–elementowy: macierzy przyporządkowujemy wartość T gdy ma tę własność, a N gdy nie ma. Własności, będącej niezmiennikiem podobieństwa macierzy, odpowiada funkcja spełniająca warunek (1); wyznacza więc ona pewną własność operatorów.

Przykład 1. a) Rozpatrzmy własność $\mathbf{A}^{10} = \mathbf{0}$. Jest ona niezmiennikiem podobieństwa, wobec czego

własność operatora $L \in \mathcal{L}(V)$: $\quad ([L]_{\mathcal{V}})^{10} = \mathbf{0}$, gdzie $\mathcal{V}$ jest bazą w V,

nie zależy od wyboru bazy $\mathcal{V}$. Latwo zauważyć, że L ma tę własność wtedy i tyko wtedy, gdy $L^{10} = 0$.

b) Funkcja $\mathbf{A} \mapsto \mathrm{rk}(\mathbf{A})$ spełnia warunek (1) (patrz przykład 2 w p.1), więc wzór

$$\widetilde{\mathrm{rk}}(L) := \mathrm{rk}([L]_{\mathcal{V}}), \text{ gdzie } \mathcal{V} \text{ jest bazą } V, \text{ zaś } L \in \mathcal{L}(V),$$

określa pewną funkcję $\widetilde{\mathrm{rk}} : \mathcal{L}(V) \to \mathbb{N}$. Z twierdzenia 1 w §III.5.3 wynika, że jest ona identyczna z funkcją rzędu operatora, zdefiniowaną w §III.4.1 wzorem $L \mapsto \dim(L(V))$.

Gdy rozpatrujemy funkcję operatora, wyznaczoną przez funkcję macierzy niezmienniczą względem podobieństwa, to pożytecznie jest wyrazić ją w sposób nie wykorzystujący macierzy rozważanego operatora. W powyższych przykładach było to nietrudne (własności $L^{10} = 0$ i $\dim L(V) = n$ nie zależą już od macierzy operatora L w jakiejkolwiek bazie). Nie jest tak jednak zawsze. Niekiedy funkcję czy własność operatora umiemy definiować tylko przy pomocy macierzy tych operatorów; niekiedy zaś podanie „bezmacierzowej" interpretacji wymaga sporej pomysłowości. Np. funkcja wyznacznika macierzy wyznacza funkcję wyznacznika endomorfizmów liniowych danej przestrzeni, której interpretację umieliśmy podać tylko gdy przestrzeń ta jest euklidesowa.

Korzystając z opisanego przed zadaniem 1 schematu przyporządkujemy teraz macierzom kwadratowym i endomorfizmom liniowym pewne wielomiany.

Definicja. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$. „Pełne rozwinięcie wyznacznika" $\det(\mathbf{A} - x\mathbf{I})$ daje wielomian zmiennej x stopnia k. Wielomian ten będziemy oznaczać przez χ_A i nazywać **wielomianem charakterystycznym** macierzy $\mathbf{A}$.

Zadanie 2. Gdy macierz $\mathbf{A}$ jest trójkątna (górnie lub dolnie), to $\chi_{\mathbf{A}} = \prod_i (a_{ii} - x)$. Ogólniej, gdy wzdłuż przekątnej macierzy $\mathbf{A}$ ustawione są klatki $\mathbf{A}_1, ..., \mathbf{A}_s$, a nad nimi (lub gdy nimi) są same zera, to $\chi_{\mathbf{A}}$ jest iloczynem wielomianów $\chi_{\mathbf{A}_i}$.

Twierdzenie 1. *Macierze podobne mają ten sam wielomian charakterystyczny.*

Dowód. Gdy $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, to

$$\chi_{\mathbf{B}} = \det(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - x\mathbf{I}) = \det(\mathbf{S}^{-1}(\mathbf{A} - x\mathbf{I})\mathbf{S}) = \det(\mathbf{S}^{-1})\det(\mathbf{A} - x\mathbf{I})\det(\mathbf{S}) = \chi_{\mathbf{A}}$$

(Wykorzystano dwukrotnie twierdzenie Cauchy'ego o wyznaczniku iloczynu macierzy.) Możnaby uznać dowód za zakończony, gdyby nie następujący szkopuł. Twierdzenie Cauchy'ego dotyczy macierzy o wyrazach w ciele, a wyżej macierz $\mathbf{A} - x\mathbf{I}$ ma wyrazy będące wielomianami. Można jednak udowodnić, że bez zmiany działań dodawania i mnożenia, $\mathbb{F}[x]$ zanurza się jako podzbiór pewnego (nieskończonego, co czasem jest ważne) ciała $\mathbb{K}$ – wobec czego możemy twierdzenie Cauchyego zastosować, traktując $\mathbf{S}, \mathbf{S}^{-1}$ i $\mathbf{A} - x\mathbf{I}$ jako macierze o wyrazach w $\mathbb{K}$. Wymagane twierdzenie o zanurzaniu udowodnione będzie na wykładzie Algebry. Tu odnotujmy jeszcze, że przeprowadzony rachunek poprawnie dowodzi równości $\chi_{\mathbf{A}}(\lambda) = \chi_{\mathbf{B}}(\lambda)$ dla wszystkich $\lambda \in \mathbb{F}$ – co w przypadku najważniejszych dla nas ciał nieskończonych wystarcza już, by $\chi_{\mathbf{A}} = \chi_{\mathbf{B}}$. (Patrz wniosek 3 w §I.2.2.) $\square$

Definicja. **Wielomian charakterystyczny** χ_L **endomorfizmu** $L \in \mathcal{L}(V)$, gdzie $\dim(V) < \infty$, definiujemy jako $\chi_{\mathbf{A}}$, gdzie $\mathbf{A} = [L]_{\mathcal{V}}$ i $\mathcal{V}$ jest bazą przestrzeni V. Jak już wyjaśniono, ze stwierdzenia wynika poprawność tej definicji. Zaś z zadania 1, zastosowanego przy $f(\mathbf{A}) = \chi_{\mathbf{A}}$, wynika

Wniosek 1. *Twierdzenie 1 pozostaje prawdziwe dla operatorów: gdy operatory $K \in \mathcal{L}(V)$ i $L \in \mathcal{L}(W)$ są podobne, to $\chi_K = \chi_L$.*

Uwaga 1. Dla $\lambda \in \mathbb{F}$ ma miejsce równość $\chi_L(\lambda) = \det(L - \lambda I)$. Istotnie, gdy $\mathcal{V}$ jest bazą przestrzeni V, to $\det(L - \lambda I) := \det([L - \lambda I]_{\mathcal{V}}) = \det([L]_{\mathcal{V}} - \lambda \mathbf{I}) = \chi_L(\lambda)$.

Uwaga 2. Odwrócenie twierdzenia 1 nie jest prawdziwe: macierz o wierszach $(0, 1)$ i $(0, 0)$ ma ten sam wielomian charakterystyczny, co macierz zerowa, a nie jest do niej podobna.

5. * Pewne własności wielomianu charakterystycznego.

Twierdzenie 1. * *Dla macierzy $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ zachodzi równość $\chi_{\mathbf{AB}} = \chi_{\mathbf{BA}}$.*

Dowód. Jest on pouczający, bo korzysta z możliwości powiększenia ciała $\mathbb{F}$. Wyróżnimy 3 przypadki.

i) $\det(\mathbf{A}) \neq 0$. Wtedy macierze $\mathbf{BA}$ i $\mathbf{AB}$ są podobne, bo $\mathbf{BA} = \mathbf{A}^{-1}(\mathbf{AB})\mathbf{A}$, i wystarcza skorzystać z twierdzenia 1 w p.4.

ii) $\mathbb{F}$ jest ciałem nieskończnym. Przyjmijmy $\mathbf{A}_t = \mathbf{A} + t\mathbf{I}$ $(t \in \mathbb{F})$ i oznaczmy przez $c_i(t)$ i przez $d_i(t)$ współczynniki wielomianów charakterystycznych macierzy $\mathbf{A}_t\mathbf{B}$ oraz $\mathbf{BA}_t$, odpowiednio:

$$\chi_{\mathbf{A}_t\mathbf{B}} = \sum_i c_i(t)x^i \quad \text{oraz} \quad \chi_{\mathbf{BA}_t} = \sum_i d_i(t)x^i.$$

Jest jasne, że c_i, d_i oraz $\det(\mathbf{A} + t\mathbf{I})$ są wielomianami zmiennej t. Stąd tylko dla co najwyżej k wartości t mamy $\det(\mathbf{A} + t\mathbf{I}) = 0$, a dla pozostałych t macierz $\mathbf{A}_t$ jest nieosobliwa i $\chi_{\mathbf{A}_t\mathbf{B}} = \chi_{\mathbf{BA}_t}$ na podstawie i). Tak więc $c_i(t) = d_i(t)$ dla nieskończenie wielu t, a tym samym dla wszystkich $t \in \mathbb{F}$. (Korzystamy z wniosku 3 w §I.2.2.) Przy $t = 0$ otrzymujemy tezę.

iii) W przypadku ogólnym oznaczamy przez $\mathbb{K}$ rozszerzenie ciała $\mathbb{F}$, będące ciałem nieskończonym. (Patrz dowód twierdzenia w p.4.) Traktując $\mathbf{A}, \mathbf{B}$ jako elementy $\mathcal{M}_k(\mathbb{K})$ znajdujemy się w przypadku ii) rozstrzygniętym wyżej. $\square$

Zadanie 1. Dla $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ oznaczmy przez $c_i(\mathbf{A})$ współczynniki wielomianu $\chi_{\mathbf{A}}$:

$$\chi_A = (-1)^k x^k + \sum_{i=0}^{k-1} c_i(\mathbf{A})x^i.$$

a) $c_i(\mathbf{A}) = c_i(\mathbf{B})$ dla macierzy podobnych $\mathbf{A}, \mathbf{B}$ (tzn. funkcje c_i są niezmiennikami podobieństwa).

b) $c_0(\mathbf{A}) = \det(\mathbf{A})$ i $c_{k-1}(\mathbf{A}) = (-1)^{k-1}\mathrm{tr}(\mathbf{A})$.

c) Ogólniej, $c_s(\mathbf{A}) = (-1)^s \sum_{\#P=k-s} \det(\mathbf{A}_P)$ dla $s = 0, ..., k-1$, gdzie $\mathbf{A}_P$ oznacza podmacierz wyznaczoną przez wiersze i kolumny o numerach ze zbioru $P \subset \{1, ..., k\}$.

d) Gdy $\chi_{\mathbf{A}}$ ma, uwzględniając krotności, k pierwiastków $\lambda_1, ..., \lambda_k \in \mathbb{F}$, to $\lambda_1 + ... + \lambda_k = \mathrm{tr}(\mathbf{A})$, $\lambda_1\lambda_1...\lambda_k = \det(\mathbf{A})$ oraz, ogólniej,

$$\sum_{\#P=s} \prod_{i \in P} \lambda_i = \sum_{\#P=s} \det(\mathbf{A}_P).$$

Zadania uzupełniające.

1. a) Znaleźć zależności pomiędzy $\chi_{\mathbf{A}}, \chi_{\mathbf{A}^{-1}}$ i $\det(\mathbf{A})$, gdy $\det(\mathbf{A}) \neq 0$.

   b) Dla macierzy unitarnej $\mathbf{A}$ stopnia 2 lub 3, o dodatnim wyznaczniku, wyrazić $\chi_{\mathbf{A}}$ przez $\mathrm{tr}(\mathbf{A})$.

2. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ i $\chi_{\mathbf{A}}(\lambda) = 0$. Dowieść, że:
   a) λ leży w jednym z tzw. **kół Greshgorina** $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$. (Por. zad. uz. 3 w §II.5.2.)
   b) $|\lambda| \leq R := \max_i \sum_j |a_{ij}|$, wobec czego $|\det(\mathbf{A})| \leq R^k$. Ponadto, $R \leq k \max_{i,j} |a_{ij}|$.

6. * Podobieństwo macierzy rzeczywistych nad $\mathbb{C}$ versus nad $\mathbb{R}$.

Dla kwadratowych macierzy zespolonych $\mathbf{A}, \mathbf{B}$ można mówić o ich podobieństwie nad $\mathbb{R}$ (gdy istnieje nieosobliwa macierz rzeczywista $\mathbf{S}$, dla której $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{B}$) lub nad $\mathbb{C}$ (gdy istnieje zespolona taka macierz).

Twierdzenie 1. *Jeśli macierze rzeczywiste* $\mathbf{A}, \mathbf{B}$ *są podobne nad* $\mathbb{C}$*, to są podobne i nad* $\mathbb{R}$. [1]

Dowód. Niech $\mathbf{S} \in \mathcal{M}_k(\mathbb{C})$ będzie macierzą nieosobliwą dla której $\mathbf{SA} = \mathbf{BS}$. Pisząc $\mathbf{S} = \mathbf{P} + i\mathbf{Q}$, gdzie $\mathbf{P}$ i $\mathbf{Q}$ mają wyrazy rzeczywiste, otrzymujemy $\mathbf{PA} = \mathbf{BP}$ i $\mathbf{QB} = \mathbf{BQ}$. Stąd $(\mathbf{P} + t\mathbf{Q})\mathbf{A} = \mathbf{B}(\mathbf{P} + t\mathbf{Q})$ dla wszystkich $t \in \mathbb{R}$. Zauważmy, że wielomian $\det(\mathbf{P} + x\mathbf{Q}) \in \mathbb{C}[x]$ nie jest zerowy, bo w punkcie i jego wartość jest różna od 0. Ma on więc tylko skończenie wiele pierwiastków i istnieje liczba $t \in \mathbb{R}$ nie będąca pierwiastkiem. Macierz $\mathbf{T} := \mathbf{P} + t\mathbf{Q}$ jest nieosobliwa i spełnia równość $\mathbf{TA} = \mathbf{BT}$.

7. ** Podobieństwo a automorfizmy algebry operatorów (problem i zadanie).

Problem 1. Dowieść, że każdy automorfizm algebry $\mathcal{L}(V)$ wynika ze zmiany bazy. Czyli: gdy bijekcja $\Psi : \mathcal{L}(V) \to \mathcal{L}(V)$ jest taka, że $\Psi(cL_1L_2) = c\Psi(L_1)\Psi(L_2)$ i $\Psi(L_1 + L_2) = \Psi(L_1) + \Psi(L_2)$ dla $L_1, L_2 \in \mathcal{L}(V)$ i $c \in \mathbb{F}$, to $\Psi(L) = S^{-1}LS$ dla pewnego izomorfizmu $S : V \to V$ i wszystkich $L \in \mathcal{L}(V)$.

Czytelnik zainteresowany tym problemem może uwzględnić poniższe zadanie jako wskazówkę.

Zadanie uzupełniające 1. Gdy $\Psi : \mathcal{L}(\mathbb{F}^k) \to \mathcal{L}(\mathbb{F}^k)$ jest automorfizmem, to:
a) Dla $L \in \mathcal{L}(V)$ mamy $\dim(\mathrm{im}(\Psi(L)) = \dim(\mathrm{im}(L))$. (Wskazówka: $\dim(\ker(L)) = \max\{s :$ istnieją $P_1, ..., P_s \in \mathcal{L}(\mathbb{F}^k) \setminus \{\mathbf{0}\}$ takie, że $P_1 + ... + P_s = I_V$ oraz $LP_i = 0$ i $P_iP_j = 0$ dla $i, j = 1, ..., s, i \neq j\}$.)
b) Dla $i, j = 1, ..., k$ istnieją wektory kolumnowe $\mathbf{v}_{ij}, \mathbf{w}_{ij} \in \mathbb{F}^k$ takie, że $[\Psi(L_{\mathbf{E}_{ij}})] = \mathbf{v}_{ij}^t \mathbf{w}_{ij}$, gdzie $\mathbf{E}_{ij} \in \mathcal{M}_k$ to macierz o (i, j)–tym wyrazie równym 1, a pozostałych 0.

§ 2. Diagonalizacja i postać kanoniczna Jordana.

1. Macierze i endomorfizmy diagonalizowalne.

Definicja. a) Operator $L \in \mathcal{L}(V)$ jest **diagonalizowalny**, jeśli istnieje baza $\mathcal{V}$ przestrzeni V taka, że $[L]_{\mathcal{V}}$ jest macierzą diagonalną. O bazie $\mathcal{V}$ powiemy, że **diagonalizuje** operator L.
b) Macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ jest diagonalizowalna (nad $\mathbb{F}$), gdy taki jest operator $L_{\mathbf{A}} \in \mathcal{L}(\mathbb{F}^k)$.

Uwaga 1. Macierz diagonalizowalna jest podobna do diagonalnej, i odwrotnie. Co więcej, jeśli baza $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ przestrzeni $\mathbb{F}^k$ diagonalizuje operator $L_{\mathbf{A}}$ i $\mathbf{S}$ jest macierzą o kolumnach $\mathbf{v}_1, ..., \mathbf{v}_k$, to $\mathbf{S}^{-1}\mathbf{AS}$ jest macierzą diagonalną. (Patrz wniosek 1 w §1.3.) Znalezienie takiej bazy $\mathcal{V}$ czy, równoważnie, macierzy $\mathbf{S}$, nazywamy więc zagadnieniem diagonalizacji macierzy $\mathbf{A}$ przez podobieństwo.

[1] Jest to (ważny) przypadek szczególny twierdzenia, dotyczącego dowolnego ciała i jego podciała.

Naszym celem jest ustalenie, które operatory są diagonalizowalne, a także zbadanie własności takich operatorów. Zasadnicza jest

Uwaga 2. Gdy $L \in \mathcal{L}(V)$ i $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ jest bazą przestrzeni V, to warunek $[L]_{\mathcal{V}} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$ jest równoważny temu, by $L(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ dla $i = 1, ..., k$. Tak więc każdy wektor $\mathbf{v}$ bazy diagonalizującej operator L ma tę własność, że $L(\mathbf{v}) \in \mathbb{F}\mathbf{v}$.

Pytanie: Kiedy (i jak) znaleźć można bazę, złożoną z wektorów $\mathbf{v}$ o tej własności?

Definicja. Gdy $L(\mathbf{v}) = \lambda \mathbf{v}$ i $\mathbf{v} \neq \mathbf{0}$, to mówimy, że $\mathbf{v}$ jest **wektorem własnym**, a λ **wartością własną** operatora L, i że **odpowiadają** one każde drugiemu. **Przestrzenią własną operatora** $L \in \mathcal{L}(V)$, odpowiadającą wartości własnej λ, nazywamy zbiór

$$V_L(\lambda) := \{\mathbf{v} \in V : L(\mathbf{v}) = \lambda \mathbf{v}\} = \ker(L - \lambda I).$$

(Tak więc przestrzeń ta składa się z zera i wektorów własnych, odpowiadających wartości λ.)

Stwierdzenie 1. *Dla operatora $L \in \mathcal{L}(V)$ i skalara λ równoważne są warunki: a) λ jest wartością własną operatora L, b) $det(L - \lambda I) = 0, tzn.$ $\chi_L(\lambda) = 0$, c) operator $L - \lambda I$ jest nieodwracalny.*

Dowód. Każdy z tych warunków jest równoważny temu, by $\ker(L - \lambda I) \neq \{\mathbf{0}\}$. $\square$

Definicja. a) Zbiór skalarów λ, spełniających te równoważne warunki, nazywamy **spektrum** lub **widmem** operatora L i oznaczamy $\mathrm{spec}(L)$.

b) Gdy $L = L_{\mathbf{A}}$ dla pewnej $k \times k$–macierzy $\mathbf{A}$, to wektory własne operatora $L_{\mathbf{A}}$ nazywamy **wektorami własnymi macierzy A**, i podobnie czynimy z wartościami własnymi, przestrzeniami własnymi i spektrum. Te ostatnie oznaczamy przez $V_{\mathbf{A}}(\lambda)$ i $\mathrm{spec}(\mathbf{A})$, odpowiednio. Tak więc

$$V_{\mathbf{A}}(\lambda) = \{\mathbf{v} \in \mathbb{F}^k : \mathbf{A}\mathbf{v} = \lambda \mathbf{v}\} \quad \text{i} \quad \mathrm{spec}(\mathbf{A}) = \{\lambda \in \mathbb{F} : \chi_{\mathbf{A}}(\lambda) = 0\}$$

Przykład 1. By znaleźć przestrzeń własną $V_{\mathbf{A}}(\lambda)$ macierzy $\mathbf{A}$ należy więc rozwiązać jednorodny układ równań $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$. Dla przykładu, gdy $\mathbb{F} = \mathbb{R}$ i $\mathbf{A}$ jest macierzą o wierszach ....., to $\chi_{\mathbf{A}} =$ ...., $\mathrm{spec}(\mathbf{A}) =$ .... i $\mathbf{A}$ ma dwie podprzestrzenie własne: $V_{\mathbf{A}}() = \mathbb{R}()$ i $V_{\mathbf{A}}() = \mathbb{R}()$. (Pomijam rachunki.)

Przykład 2. Gdy $\mathbf{A} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, to $V_{\mathbf{A}}(\lambda) = \{(x_i)_{i=1}^k \in \mathbb{F}^k : x_i = 0 \text{ gdy } \lambda_i \neq \lambda\}$.

Uwaga 3. Widmo macierzy zależy od rozpatrywanego ciała skalarów $\mathbb{F}$. Np., dla macierzy o wierszach $(0, 1)$ i $(1, 0)$ jest ono puste przy $\mathbb{F} = \mathbb{R}$, lecz równe $\{-i, i\}$ przy $\mathbb{F} = \mathbb{C}$. Dlatego ciało $\mathbb{F}$ niekiedy zaznaczamy, pisząc $\mathrm{spec}_{\mathbb{F}}(\mathbf{A})$ w miejsce $\mathrm{spec}(\mathbf{A})$.

Stwierdzenie 2. *Układ skończenie wiele wektorów własnych operatora L, odpowiadających jego różnym wartościom własnym, jest liniowo niezależny.*

Dowód. Niech wektorami tymi będą $\mathbf{v}_1, \ldots \mathbf{v}_s$, i niech odpowiadają one (parami) różnym wartościom $\lambda_1, \ldots, \lambda_s \in \mathbb{F}$. Wykorzystamy indukcję względem liczby s. Dla $s = 1$ teza wynika stąd, że wektory własne są różne od $\mathbf{0}$. Niech więc $s > 1$ i przypuśćmy, że $\sum_i c_i \mathbf{v}_i = \mathbf{0}$. Działając na obie strony operatorem L otrzymujemy zależność $\sum \lambda_i c_i \mathbf{v}_i = 0$. Odejmując od niej poprzednią, pomnożoną przez λ_s, otrzymujemy $\sum_{i=1}^{s-1} c_i(\lambda_i - \lambda_s)\mathbf{v}_i = \mathbf{0}$. Ponieważ skalary $\lambda_i - \lambda_s$ są niezerowe, więc z założenia indukcyjnego wynika, że $c_i = 0$ dla $i < s$, skąd też $c_s = 0$ (bo $\sum_i c_i \mathbf{v}_i = \mathbf{0}$). $\square$

Wniosek 1. *Suma algebraiczna $\sum_{\lambda \in \mathrm{spec}(L)} V_L(\lambda)$ wszystkich podprzestrzeni własnych operatora L jest prosta. W szczególności, suma wymiarów tych podprzestrzeni nie przekracza wymiaru przestrzeni V.*

Dowód. Wynika to ze stwierdzenia 2 i definicji, a końcowa konkluzja stąd, że wymiar sumy prostej jest sumą wymiarów składników.

Poprzedzające wyniki pozwalają ustalić, kiedy dany operator jest diagonalizowalny.

Twierdzenie 1. *Następujące warunki są równoważne dla operatora $L \in \mathcal{L}(V)$:*
a) operator ten jest diagonalizowalny;
b) istnieje baza przestrzeni V, złożona z wektorów własnych operatora L;
c) suma wymiarów podprzestrzeni własnych operatora L nie jest mniejsza od wymiaru przestrzeni V, tzn. $\sum_{\lambda\in\ \mathrm{spec}(L)} \dim(V_L(\lambda)) \geq \dim(V)$;
d) V jest sumą prostą podprzestrzeni własnych operatora L, tzn. $V = \oplus_{\lambda\in\mathrm{spec}(L)} V_L(\lambda)$.

Dowód. Równoważność a) $\Leftrightarrow$ b) odnotowano już w uwadze 1.

b) $\Rightarrow$ c). Gdy A jest (nieuporządkowaną) bazą, o której mowa w b), to rozpada się ona na sumę rozłącznych zbiorów $A_\lambda = \{\mathbf{a} \in A : L(\mathbf{a}) = \lambda\mathbf{a}\}$, gdzie λ przebiega wszystkie wartości własne. Każdy zbiór A_λ jest liniowo niezależny (jako podzbiór takiego) i zawarty w $V_\lambda(L)$, skąd $\sum_\lambda \dim(V_L(\lambda)) \geq \sum_\lambda \#A_\lambda = \#A = \dim(V)$.

c) $\Rightarrow$ d). Niech $W = \sum_\lambda V_L(\lambda)$. Jak wiemy, suma po prawej stronie jest prosta, skąd $\dim(W) = \sum_\lambda \dim(V_L(\lambda))$. Jeśli więc zachodzi c), to $\dim(W) \geq \dim(V)$, skąd $W = V$.

d)$\Rightarrow$b). Dla każdej wartości własnej λ obierzmy bazę B_λ odpowiadającej jej podprzestrzeni własnej $V_L(\lambda)$. Gdy zachodzi d), to B jest szukaną bazą przestrzeni V. (Patrz uwaga 1 w §III. 6.2.)

Wniosek 2. *Gdy wielomian χ_L ma $k = \dim(V)$ różnych pierwiastków, to operator L jest diagonalizowalny. (Odwrotna implikacja nie ma miejsca, co pokazuje przykład $L = I_V$.)*

Dowód. Ponieważ wymiar każdej podprzestrzeni własnej jest ≥ 1, więc przy założeniach wniosku spełniony jest warunek c) twierdzenia. □

Uwaga 4. Dowód twierdzenia 2 podsuwa sposób znajdywania bazy diagonalizującej dany operator L, czy diagonalizacji macierzy przez podobieństwo (gdy są one możliwe). Sprowadza się on do wykonania następujących kroków:

1). Znalezienie wszystkich pierwiastków wielomianu charakterystycznego χ_L (czyli wszystkich elementów widma $\mathrm{spec}(L)$).

Choć na ogół możemy pierwiastki te wyznaczyć tylko z pewnym przybliżeniem, to pomijamy dyskusję tego, na ile możliwe błędy wpływają na wynik dalszych kroków. Zakładamy dla uproszczenia, że znamy dokładne wartości wszystkich pierwiastków.

2). Dla każdego $\lambda \in\ \mathrm{spec}(L)$, znalezienie bazy B_λ przestrzeni $V_L(\lambda)$

3). Gdy warunek c) twierdzenia nie jest spełniony, to operator nie jest diagonalizowalny. W przeciwnym razie jest, a $\bigcup_{\lambda\in\mathrm{spec}(L)} B_\lambda$ jest diagonalizującą go bazą.

Gdy diagonalizujemy macierz $\mathbf{A} \in \mathcal{M}_k$, to podobnie tworzymy bazę B_λ każdej podprzestrzeni własnej $V_{\mathbf{A}}(\lambda)$. Jeśli liczebność zbioru $B = \bigcup\{B_\lambda : \lambda \in \mathrm{spec}(\mathbf{A})\}$ jest mniejsza od k, to macierz nie jest diagonalizowalna. W przeciwnym zaś razie jest, a macierz $\mathbf{S}$, której kolumnami są wektory zbioru B, ma tę własność, że $\mathbf{S}^{-1}\mathbf{AS}$ jest macierzą diagonalną. Ostatniego iloczynu nie musimy wyliczać: na i–tym miejscu jego przekątnej stoi wartość własna, odpowiadająca i–tej kolumnie macierzy $\mathbf{S}$, patrz wniosek 1 w §1.3. □

Umiemy więc zbadać, czy dany operator jest diagonalizowalny, a gdy jest, to umiemy wskazać bazę diagonalizującą. Podobnie umiemy rozstrzygnąć, czy diagonalizowalna jest dana macierz $\mathbf{A}$, a gdy jest, to wskazać zarówno podobną do niej macierz diagonalną $\mathbf{D}$, jak i macierz $\mathbf{S}$ taką, że $\mathbf{S}^{-1}\mathbf{AS} = \mathbf{D}$. Czy jednak diagonalizacja jest możliwa, dowiadujemy się w obu przypadkach dopiero po (na ogół niewykonalnym) wyznaczeniu widma. W następnym paragrafie poznamy ogólne twierdzenia, które bez znajomości widma umożliwią rozpoznanie diagonalizowalności pewnych macierzy i operatorów (ale tylko dla $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$).

Uwaga 5. Z przykładu 2 i poniższego zadania 2b) wynika równoważność warunku c) twierdzenia z następującym, łatwiejszym do sprawdzenia:

c)' Wielomian χ_L rozkłada się nad ciałem skalarów czynniki liniowe i wymiar każdej podprzestrzeni własnej $V_L(\lambda)$ jest równy krotności λ jako pierwiastka wielomianu χ_L.

Uwaga 6. Gdy macierz $\mathbf{A}$ jest podobna do macierzy diagonalnej $\mathbf{D}$, to na przekątnej macierzy $\mathbf{D}$ występują pierwiastki wielomianu $\chi_{\mathbf{A}}$, każdy powtórzony tyle razy, ile wynosi jego krotność. (Patrz w §1.4 zadanie 2 i twierdzenie 1.) Macierz $\mathbf{D}$ jest więc przez $\mathbf{A}$ wyznaczona jednoznacznie z dokładnością do kolejności wyrazów przekątnej. Macierzy ustalających podobieństwo $\mathbf{A}$ do $\mathbf{D}$ istnieje jednak wiele.

Podajmy przykłady ilustrujące powyższe wyniki.

Przykład 3. Niech $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$. Wówczas $\operatorname{spec}(\mathbf{A}) = \{1, 4, 6\}$, skąd macierz $\mathbf{A}$ jest na mocy wniosku 1 podobna do macierzy diagonalnej, której przekątna złożona jest z wszystkich elementów $\operatorname{spec}(\mathbf{A})$, np., do macierzy $\mathbf{D} = \operatorname{diag}(1, 4, 6)$. Równości $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{D}$ czyni zadość macierz $\mathbf{S}$, której kolumnami są wektory własne odpowiadające, kolejno, wartościom własnym 1,4,6. (Ćwiczenie: wyznaczyć je.) □

Przykład 4. Niech $\mathbf{A} \in \mathcal{M}_k$ będzie macierzą o wszystkich wyrazach równych 1. Jej wartości własne można znaleźć nie wyliczając wielomianu charakterystycznego. Mamy bowiem $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow v_1 + ... + v_k = \lambda v_i$ dla $i = 1, ..., k$, więc jeśli $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ i $\lambda \neq 0$, to $v_1 = ... = v_k$, co daje $\lambda = k$ i $V_{\mathbf{A}}(k) = \mathbb{F}(1, ..., 1)$; bazę zaś przestrzeni $V_{\mathbf{A}}(0)$ tworzy układ fundamentalny równania $v_1 + ... + v_k = 0$, za który możemy obrać $(-1, 1, 0, ...,), (-1, 0, 1, 0, ..., 0), ..., (-1, 0, ..., 0, 1)$. Stąd $\dim V_{\mathbf{A}}(k) = 1, \dim V_{\mathbf{A}}(0) = k - 1$. Warunek c) twierdzenia 1 jest więc spełniony, a dla macierzy $\mathbf{S}$, której kolejnymi kolumnami są wektory $(1, ..., 1), (-1, 1, 0, ..., 0), ..., (-1, 0, ..., 0, 1)$, zachodzi $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \operatorname{diag}(k, 0, ..., 0)$. (Daje to też $\chi_{\mathbf{A}} = (-x)^{k-1}(k - x)$; dlaczego?)

Zadania.

1. Gdy $\mathbf{A} = \operatorname{diag}(\mathbf{A}_1, \mathbf{A}_2)$, to $V_{\mathbf{A}}(\lambda) = V_{\mathbf{A}_1}(\lambda) \times V_{\mathbf{A}_2}(\lambda)$ i ogólniej $V_{\mathbf{A}}(\lambda) = V_{\mathbf{A}_1}(\lambda) \times \cdots \times V_{\mathbf{A}_s}(\lambda)$ dla $\mathbf{A} = \operatorname{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_s)$.

2. a) Gdy operatory $K \in \mathcal{L}(V)$, $L \in \mathcal{L}(W)$ są podobne i $\lambda \in \mathbb{F}$, to $\dim(V_K(\lambda)) = \dim(V_L(\lambda))$; co więcej, wtedy $S(V_K(\lambda)) = V_L(\lambda)$ dla każdego izomorfizmu $S : V \to W$ takiego, że $L = SKS^{-1}$.

b) Gdy więc $\mathbf{A}$ jest macierzą operatora $K \in \mathcal{L}(V)$ w pewnej bazie $\mathcal{V}$ przestrzeni V, to $\operatorname{spec}(K) = \operatorname{spec}(\mathbf{A})$ oraz $\dim(V_K(\lambda)) = \dim(V_{\mathbf{A}}(\lambda)$ dla $\lambda \in \operatorname{spec}(K)$.

3. Gdy $(\mathbf{v}_1, ..., \mathbf{v}_k)$ jest bazą przestrzeni V, złożoną z wektorów własnych operatora $L \in \mathcal{L}(V)$, to te z wektorów $\mathbf{v}_i$, które odpowiadają danej wartości własnej λ, tworzą bazę przestrzeni $V_L(\lambda)$.

Zadanie uzupełniające 1. Rozpatrzmy przestrzeń $C^\infty(\mathbb{R})$ funkcji posiadających wszystkie pochodne.

a) Dowieść, że funkcja $\exp(\lambda t)$ jest wektorem własnym operatora różniczkowania $f \mapsto Df$, zaś funkcja $\sin(\lambda t)$ jest wektorem własnym operatora $f \mapsto D^2 f$.

b) Wywnioskować, że gdy $\lambda_i \neq \lambda_j$ dla $i \neq j$, to funkcje $\sin(\lambda_1 t), ..., \sin(\lambda_n t)$ są liniowo niezależne, podobnie jak funcje $\exp(\lambda_1 t), ..., \exp(\lambda_n t)$.

Zadanie uzupełniające 2. Niech operator $L \in L(V)$ spełnia warunek $L^3 = I_V$. Zakładamy też, że $\mathbb{F} = \mathbb{C}$ i $\dim V < \infty$. Udowodnić, że operator ten jest diagonalizowalny, a każda jego wartość własna jest pierwiastkiem z jedynki. Uogólnić na przypadek $L^k = I_V$. (Wskazówka: użyteczne może okazać się zadanie 6 w §1.2.2.)

Zadania ze zbioru Kostrykina: od 1 do 5, od 13 do 18 oraz 9, 20 i 21 w §II.3.2.

2. Sformułowanie twierdzenia Jordana i jednoznaczność postaci Jordana.

Definicje. Klatką Jordana, odpowiadającą wartości λ, nazywamy macierz kwadratową, której wyrazy na przekątnej są równe λ, wyrazy bezpośrednio pod nią równe 1, a pozostałe 0. Gdy klatka jest rozmiaru $s \times s$, to oznaczamy ją przez $\mathbf{J}(\lambda, s)$:

$$\mathbf{J}(\lambda,1) = (\lambda) \in \mathcal{M}_1(\mathbb{F}) \ \text{ i } \ \mathbf{J}(\lambda, s) := \begin{pmatrix} \lambda & & & & 0 \\ 1 & \lambda & & & \\ & 1 & \ddots & & \\ & & \ddots & \lambda & \\ 0 & & & 1 & \lambda \end{pmatrix} \in \mathcal{M}_s(\mathbb{F}) \text{ dla } s \geq 2.$$

Macierzą Jordana nazywamy macierz postaci $\operatorname{diag}(\mathbf{J}_1, ..., \mathbf{J}_t)$, gdzie każda macierz $\mathbf{J}_i$ jest klatką Jordana. (Różne klatki mogą odpowiadać różnym wartościom, lecz mogą i tym samym. Mogą mieć też różne lub te same rozmiary.)

Macierze te odgrywają ważną role ze względu na

Twierdzenie 1 (C. Jordana). *Niech $L \in \mathcal{L}(V)$, gdzie $V \neq \{\mathbf{0}\}$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$. Jeśli wielomian charakterystyczny χ_L operatora L rozkłada się nad $\mathbb{F}$ na iloczyn wielomianów liniowych, to istnieje baza $\mathcal{B}$ przestrzeni V taka, że $[L]_\mathcal{B}$ jest macierzą Jordana. (Jak wszędzie zakładamy, że $\dim V < \infty$.)*

Bazę, o której mowa w tezie, nazwiemy **bazą Jordana** dla operatora L, a macierz $[L]_\mathcal{B}$ – **macierzą Jordana operatora** L.

Uwaga 1. Warunek rozkładalności wielomianu χ_L jest dla istnienia bazy Jordana konieczny. Istotnie, wielomiany charakterystyczne operatora L i macierzy $[L]_\mathcal{B}$ są identyczne, a drugi z nich jest iloczynem czynników liniowych, bo $[L]_\mathcal{B}$ jest macierzą trójkątna. (Patrz zadanie 2 w §1.4). Gdy $\mathbb{F} = \mathbb{C}$ (ogólniej: gdy ciało $\mathbb{F}$ jest *algebraicznie domknięte*), to każdy operator spełnia ten warunek.

Gdy twierdzenie 1 i uwagę 1 zastosować wraz z wnioskiem 1 w §1.3 do przekształcenia $L_\mathbf{A}$, o zadanej macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, to otrzymamy równoważne im

Twierdzenie 1'. *Macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ wtedy i tylko wtedy jest nad $\mathbb{F}$ podobna do pewnej macierzy Jordana $\mathbf{J}$, gdy wielomian $\chi_\mathbf{A}$ rozkłada się nad $\mathbb{F}$ na iloczyn czynników liniowych.*

Powyższą macierz $\mathbf{J}$ (gdy istnieje) nazywamy **postacią Jordana** macierzy $\mathbf{A}$.

Dowód twierdzenia 1 podamy w punktach 4 i 5; aż do p.4 zakładamy jego prawdziwość. Obecnie zajmiemy się zbadaniem własności macierzy Jordana i jednoznaczności macierzy $[L]_\mathcal{B}$ z twierdzenia 1.

Zadanie 1. i –tą kolumną n–tej potęgi $\mathbf{J}(0,s)^n$ macierzy $\mathbf{J}(0,s)$ jest $\mathbf{e}_{i+n}$ dla $i \leq s-n$ oraz $\mathbf{0}$ dla pozostałych wartości $i \leq s$. (Patrz zadanie 2 w §II.2.1.) Wobec tego, wymiar przestrzeni kolumn macierzy $\mathbf{J}(0,s)^n$ jest równy $\max(s-n, 0)$.

Wniosek 1. *Gdy $\mathbf{K} = \mathbf{J}(\mu, s)$, to dla $\lambda \in \mathbb{F}$ i $n \in \mathbb{N}$ mamy* $\operatorname{rk}(\mathbf{K} - \lambda\mathbf{I})^{n-1} - \operatorname{rk}(\mathbf{K} - \lambda\mathbf{I})^n = 1$ *jeśli $\lambda = \mu$ i $n \leq s$, a w przeciwnym razie* $\operatorname{rk}(\mathbf{K} - \lambda\mathbf{I})^{n-1} - \operatorname{rk}(\mathbf{K} - \lambda\mathbf{I})^n = 0$.

Dowód. Gdy $\lambda \neq \mu$, to macierz $\mathbf{K} - \lambda\mathbf{I}$ jest nieosobliwa, skąd $\mathrm{rk}(\mathbf{K} - \lambda\mathbf{I})^m = s$ dla każdego m – co daje odpowiednią część tezy. Gdy zaś $\lambda = \mu$, to $\mathbf{K} - \lambda\mathbf{I} = \mathbf{J}(0, s)$ i pozostaje skorzystać z zadania.

Twierdzenie 2. *Niech macierz Jordana* $\mathbf{J}$ *będzie podobna do danej macierzy* $\mathbf{A}$. *Dla* $\lambda \in \mathbb{F}$ *oraz* $n \in \mathbb{N}$ *oznaczmy przez* $p_n(\lambda)$ *(odpowiednio: przez* $q_n(\lambda)$*) liczbę tych jordanowskich klatek macierzy* $\mathbf{J}$*, które odpowiadają wartości* λ *i mają stopień* $\geq n$ *(odpowiednio: są stopnia* n*). Wówczas*

$$p_n(\lambda) = \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n-1} - \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n \quad i \quad q_n(\lambda) = \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n-1} - 2\mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n + \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n+1}.$$

Równoważne sformułowanie: gdy $\mathbf{J}$ *jest macierzą Jordana operatora* L*, to powyższe liczby* $p_n(\lambda)$ *i* $q_n(\lambda)$ *wyrażają się analogicznymi wzorami, z* $\mathbf{A}$ *zmienionym na* L.

Dowód. Udowodnimy wersję operatorową; macierzowa jest jej konsekwencją dzieki stwierdzeniu 1 w §1.3. Ustalmy $n \in \mathbb{N}$ i niech $\mathbf{J}$ będzie sumą zewnętrzną klatek Jordana $\mathbf{K}_1, ..., \mathbf{K}_t$. Dla $m = n, n-1$ macierz $(\mathbf{J} - \lambda\mathbf{I})^m$ jest wtedy sumą zewnętrzną klatek $(\mathbf{K}_i - \lambda\mathbf{I}_i)^m$, a także jest macierzą operatora $(L - \lambda J)^m$ w pewnej bazie, wobec czego $\mathrm{rk}(L - \lambda I)^m = \mathrm{rk}(\mathbf{J} - \lambda\mathbf{I})^m = \sum_{i=1}^t \mathrm{rk}(\mathbf{K}_i - \lambda\mathbf{I})^m$. Stąd $\mathrm{rk}(L - \lambda I)^{n-1} - \mathrm{rk}(L - \lambda I)^n = \sum_{i=1}^t (\mathrm{rk}(\mathbf{K}_i - \lambda\mathbf{I})^{n-1} - \mathrm{rk}(\mathbf{K}_i - \lambda\mathbf{I})^n)$. Na podstawie wniosku 1, pod znakiem sumy pojawiają się prócz zer tylko jedynki, których jest $p_n(\lambda)$. To kończy dowód, bo ponadto $q_n(\lambda) = p_n(\lambda) - p_{n+1}(\lambda)$. $\square$

Wniosek 2. *Macierz Jordana operatora* L *i postać Jordana macierzy* $\mathbf{A}$ *jest przez* L *czy* $\mathbf{A}$ *wyznaczona jednoznacznie z dokładnością do kolejności klatek Jordana.*

Dowód. Macierz czy operator wyznaczają wszystkie liczby $q_n(\lambda)$ $(\lambda \in \mathbb{F}, n \in \mathbb{N})$. $\square$

Uwaga 2. By wskazać macierz Jordana $\mathbf{J}$ operatora L, wystarcza znać liczby $q_n(\lambda)$ dla $n \leq \dim(V)$ i $\lambda \in \mathrm{spec}(L)$. (Klatki Jordana macierzy $\mathbf{J}$ odpowiadają bowiem pierwiastkom wielomianu $\chi_{\mathbf{J}} = \chi_L$ i są rozmiaru $\leq \dim(V)$.) Dla małych n, wystarczające do wyznaczenia $\mathbf{J}$ może być

Zadanie 2. Niech $\mathbf{J}(\lambda, s_1), \mathbf{J}(\lambda, s_2), ..., \mathbf{J}(\lambda, s_t)$ będą tymi klatkami jordanowskimi macierzy Jordana operatora L, które odpowiadają wartości λ. Dowieść, że

a) t jest wymiarem przestrzeni własnej $V_L(\lambda)$.

b) $n_\lambda := s_1 + ... + s_t$ jest krotnością λ jako pierwiastka wielomianu χ_L,

c) liczba $m_\lambda = \max_i s_i$ (wskazująca rozmiar największej spośród wymienionych klatek) jest równa $\inf\{n : \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n = \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n+1}\}$

Ćwiczenie. Niech operator $L \in \mathcal{L}(V)$ i baza $\mathbf{v}_1, ..., \mathbf{v}_4$ przestrzeni V mają następujące własności: $L(\mathbf{v}_1) = \mathbf{v}_2, L(\mathbf{v}_4) = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3$ oraz $L(\mathbf{v}_i) = \mathbf{0}$ dla $i = 2, 3$. W zależności od wartości parametrów a, b, c wyznaczyć macierz Jordana tego operatora. (Wskazówka: wypisać obrazy wektorów $\mathbf{v}_i$ przy L^2 i przy L^3.)

Ćwiczenie. Jaka jest postać Jordana macierzy $\mathbf{J}(0, n)^2$?

Wniosek 3. *Niech macierze* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ *mają wspólny wielomian charakterystyczny* $\chi_{\mathbf{A}} = \chi_{\mathbf{B}}$*, rozkładający się nad* $\mathbb{F}$ *na czynniki liniowe. Do tego, by macierze te były podobne nad* $\mathbb{F}$ *potrzeba i wystarcza, by* $\mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n = \mathrm{rk}(\mathbf{B} - \lambda\mathbf{I})^n$ *dla* $n = 1, ..., k$ *i wszystkich* λ *będących pierwiastkami wielomianu* $\chi_{\mathbf{A}} = \chi_{\mathbf{B}}$.

Dowód. Załóżmy, że ostatni warunek jest spełniony. Z twierdzenia Jordana wynika, że $\mathbf{A}$ i $\mathbf{B}$ są podobne do pewnych macierzy Jordana $\mathbf{J}_{\mathbf{A}}$ i $\mathbf{J}_{\mathbf{B}}$, odpowiednio, a z uwagi 2 – że $\mathbf{J}_{\mathbf{A}}$ i $\mathbf{J}_{\mathbf{B}}$ różnią się tylko kolejnością klatek Jordana, a więc są podobne. (Korzystamy tu z z końcowej części przykładu 1 w §1.3.) W ślad za nimi, podobne są macierze $\mathbf{A}$ i $\mathbf{B}$.

Odwrotnie, gdy macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, to podobne są też macierze $\mathbf{A} - \lambda\mathbf{I}$ i $\mathbf{B} - \lambda\mathbf{I}$, skąd $\mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n = \mathrm{rk}(\mathbf{B} - \lambda\mathbf{I})^n$ $(\lambda \in \mathbb{F}, n \in \mathbb{N})$; por. zadania 1 i 2 w §1.2. $\square$

Uwaga 3. * Rezultaty dotyczące macierzy Jordana dostarczają też wartościowych informacji o diagonalizowalności operatora czy macierzy. I tak:

a) Operator L wtedy i tylko wtedy jest diagonalizowalny, gdy jego wielomian charakterystyczny χ_L rozkłada się na czynniki liniowe i $\mathrm{rk}(L-\lambda I) = \mathrm{rk}(L-\lambda I)^2$, dla każdego $\lambda \in \mathrm{spec}(L)$. (Istotnie, oba warunki są równoważne temu, by macierz Jordana operatora L miała 0 klatek jordanowskich stopnia ≥ 2; patrz twierdzenie 2.)

b) Operator L jest diagonalizowalny wtedy i tylko wtedy, gdy $p(L) = 0$ dla pewnego wielomianu p bez pierwiastków wielokrotnych. (Będzie to treścią zadania 6 w p.3.)

<u>Zadania uzupełniające.</u>

1. Znaleźć postać Jordana macierzy dolnie trójkątnej, mającej na przekątnej wyrazy równe λ, zaś bezpośrednio poniżej wyrazy niezerowe.

2. Udowodnić, że jeśli $\mathbf{A}$ jest macierzą kwadratową o wyrazach w ciele $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, to macierze $\mathbf{A}$ i $\mathbf{A}^t$ są podobne nad $\mathbb{F}$.

3. Niech $k \times k$–macierz $\mathbf{A}$ będzie postaci $\mathrm{diag}(\mathbf{K}_1, ..., \mathbf{K}_s)$, gdzie każda klatka $\mathbf{K}_i$ jest trójkątna, o „stałej" przekątnej $(\lambda_i, \ldots, \lambda_i)$. Oznaczmy przez W_i przestrzeń zerową macierzy $(\mathbf{A} - \lambda_i \mathbf{I})^k$. Dowieść, że $\mathbb{F}^k = \oplus_i W_i$ i wymiar przestrzeni W_i jest równy rozmiarowi l_i klatki $\mathbf{K}_i$ $(i = 1, ..., s)$.

4. Niech $L \in \mathcal{L}(V)$ i wielomian χ_L rozkłada się nad ciałem skalarów przestrzeni V na czynniki liniowe: $\chi_L = \prod_{i=1}^s (\lambda_i - x)^{n_i}$. W oparciu o poprzednie zadanie dowieść, że

$$V = \oplus_{i=1}^s W_i \text{ oraz } \dim(W_i) = n_i, \quad \text{gdzie } W_i := \ker(L - \lambda_i I)^k$$

Uwaga 4. * a) Zbiór $W_\lambda := \ker(L - \lambda I)^k$ nazywany jest **przestrzenią pierwiastkową** operatora L, odpowiadającą wartości λ; teza powyższego zadania zaś – twierdzeniem o rozkładzie na te przestrzenie. Może ono posłużyć do dowodu twierdzenia Jordana. (Tu przyjęto odwrotną kolejność.)

b) Przyjmijmy oznaczenia zadania 2c). Liczba $\dim\ker(L-\lambda I)^n + \mathrm{rk}(L-\lambda I)^n$ nie zależy od n, więc $\ker(L-\lambda I)^n \subsetneq \ker(L-\lambda I)^{n+1}$ dla $n < m_\lambda$ oraz $\ker(L-\lambda I)^n = \ker(L-\lambda I)^{n+1}$ dla $n \geq m_\lambda$. A że $m_\lambda \leq k$, to $\ker(L-\lambda I)^{m_\lambda} = W_\lambda$, co ułatwia znalezienie przestrzeni pierwiastkowej W_λ.

<u>Zadania ze zbioru Kostrykina:</u> 1, 2, 4, 5, 6, 10 w §II.3.3.

3. Funkcje macierzy.

Latwo wyznaczyć wartość wielomianu p na macierzy diagonalnej $\mathrm{diag}(\lambda_1, ..., \lambda_k)$: jest nią macierz $\mathrm{diag}(p(\lambda_1), ..., p(\lambda_k))$. Podobnie wyznaczyć można $p(\mathbf{A})$ gdy macierz $\mathbf{A}$ jest <u>diagonalizowalna</u> i znamy macierz $\mathbf{S}$ taką, że

$$\mathbf{A} = \mathbf{S}\mathrm{diag}(\lambda_1, ..., \lambda_k)\mathbf{S}^{-1} \tag{2}$$

Wtedy bowiem, na podstawie zadania 5a) z §1.2

$$p(\mathbf{A}) = \mathbf{S}\mathrm{diag}(p(\lambda_1), ..., p(\lambda_k))\mathbf{S}^{-1} \tag{3}$$

Ponieważ $\{\lambda_1, ..., \lambda_k\} = \mathrm{spec}(\mathbf{A})$, więc dla $p_1, p_2 \in \mathbb{F}[x]$ zachodzi:

$$\text{gdy } p_1(\lambda) = p_2(\lambda) \text{ dla każdego } \lambda \in \mathrm{spec}(\mathbf{A}), \text{ to } p_1(\mathbf{A}) = p_2(\mathbf{A}). \tag{4}$$

Uwaga 1. Założenie diagonalizowalności jest istotne!

Ćwiczenie. Gdy a, b, c są różnymi pierwiastkami stopnia 18 z 1, to $\begin{pmatrix} a & 2 & 3 \\ 0 & b & 4 \\ 0 & 0 & c \end{pmatrix}^{18} = \mathbf{I}$.

Niech teraz $f : \mathrm{spec}(\mathbf{A}) \to \mathbb{F}$ będzie dowolną funkcją. Ponieważ $\mathrm{spec}(\mathbf{A})$ jest zbiorem skończonym, więc istnieją wielomiany $p \in \mathbb{F}[x]$ takie, że $p(\lambda) = f(\lambda)$ dla wszystkich $\lambda \in \mathrm{spec}(\mathbf{A})$. Wielomianów takich istnieje wiele, lecz wartość $p(\mathbf{A})$ jest dla nich wspólna, na mocy (4). Uzasadnia to celowość następującej definicji:

Definicja i uwaga. Dla diagonalizowalnej macierzy $\mathbf{A}$ obierzmy przedstawienie (2) dowolnie i przyjmijmy

$$f(\mathbf{A}) := \mathbf{S}\mathrm{diag}(f(\lambda_1), ..., f(\lambda_k))\mathbf{S}^{-1} \quad \text{dla } f : \mathrm{spec}(\mathbf{A}) \to \mathbb{F}. \tag{5}$$

Lewa strona jest niezależna od użytego przedstawienia (2) i jest równa $p(\mathbf{A})$, dla każdego wielomianu $p \in \mathbb{F}[x]$ takiego, że $p(\lambda) = f(\lambda)$ dla wszystkich $\lambda \in \mathrm{spec}(A)$. Z (5) wynika, że macierz $f(\mathbf{A})$ jest diagonalizowalna. Nazywamy ją **wartością funkcji f na macierzy $\mathbf{A}$**.

Ćwiczenie. a) Wyznaczyć $\sin(\mathbf{A})$ oraz $e^{\mathbf{A}}$ dla macierzy $\mathbf{A}$ z przykładu 3 w p.1.

b)* Wyznaczyć $e^{\mathbf{A}}$ dla macierzy rzeczywistej $\mathbf{A} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$ (Wskazówka: diagonalizacji dokonać nad $\mathbb{C}$, skorzystać ze *wzorów Eulera*.)

Zadanie 1. Dla diagonalizowalnej macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ i funkcji $f, g : \mathrm{spec}(\mathbf{A}) \to \mathbb{F}$,

a) $(af + bg)(\mathbf{A}) = af(\mathbf{A}) + bg(\mathbf{A})$ dla $a, b \in \mathbb{F}$.

b) $(fg)(\mathbf{A}) = f(\mathbf{A})g(\mathbf{A})$, skąd $f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A})$ oraz $(1/g)(\mathbf{A}) = (g(\mathbf{A}))^{-1}$ gdy $0 \notin g(\mathrm{spec}(\mathbf{A}))$.

c) $f(\mathbf{C}\mathbf{A}\mathbf{C}^{-1}) = \mathbf{C}f(\mathbf{A})\mathbf{C}^{-1}$ dla każdej macierzy nieosobliwej $\mathbf{C} \in \mathcal{M}_k(\mathbb{F})$.

d) $\mathrm{spec}(f(\mathbf{A})) = f(\mathrm{spec}(\mathbf{A}))$.

e) $h(f(\mathbf{A})) = (h \circ f)(\mathbf{A})$ dla każdej funkcji $h : f(\mathrm{spec}(\mathbf{A})) \to \mathbb{F}$.

f) $f(\mathbf{A}) = \sum_{i=0}^{s} c_i \mathbf{A}^i$ dla pewnych $c_0, ..., c_s \in \mathbb{F}$, wobec czego $f(\mathbf{A})\mathbf{B} = \mathbf{B}f(\mathbf{A})$ dla macierzy $\mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ takich, że $\mathbf{AB} = \mathbf{BA}$.

g) Gdy $\mathbb{F} = \mathbb{C}$ i ciąg funkcji $f_n : \mathrm{spec}(\mathbf{A}) \to \mathbb{C}$ jest zbieżny do f, to $f_n(\mathbf{A}) \to f(\mathbf{A})$, tzn. wyrazy macierzy $f_n(\mathbf{A})$ są zbieżne do odpowiadających im wyrazów macierzy $f(\mathbf{A})$.

h) Gdy $\mathbb{F} = \mathbb{C}$ i macierz $\mathbf{A}$ jest unitarnie podobna do diagonalnej, to $(f(\mathbf{A}))^* = \overline{f}(\mathbf{A})$. W szczególności, zachodzi wówczas $\mathbf{A}^* = \sum_{i=0}^{s} c_i \mathbf{A}^i$ dla pewnych $c_0, ..., c_s \in \mathbb{C}$.

Rozszerzymy teraz te wyniki na przypadek niediagonalizowalnej macierzy $\mathbf{A}$. Dla uproszczenia zakładamy, że $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Wykorzystamy następujący

Lemat 1. *Niech macierze $\mathbf{X}_0, \mathbf{X} \in \mathcal{M}_k(\mathbb{F})$ będą przemienne (tzn. $\mathbf{X}_0\mathbf{X} = \mathbf{X}\mathbf{X}_0$). Wówczas dla każdego wielomianu $p \in \mathbb{F}[x]$ prawdziwy jest* **wzór Taylora**, *w którym $D^n p$ oznacza n–tą pochodną wielomianu p (w tym $D^n p = 0$ gdy $n > \deg(p)$):*

$$p(\mathbf{X}_0 + \mathbf{X}) = \sum_{n=0}^{\infty} \frac{(D^n p)(\mathbf{X}_0)}{n!} \mathbf{X}^n \tag{6}$$

Dowód. Zbiór W tych wielomianów p, dla których równość (6) ma miejsce, jest zamknięty względem dodawania wielomianów i mnożenia ich przez skalar. Ponieważ wielomiany $1, x, x^2, ...$ rozpinają liniowo $\mathbb{F}[x]$, więc pozostaje dowieść, że należą one do W. Jednak gdy $p = x^l$, to z przemienności $\mathbf{X}_0$ i $\mathbf{X}$ otrzymujemy łatwo $p(\mathbf{X}_0 + \mathbf{X}) = \sum_{n=0}^{l} \binom{l}{n} \mathbf{X}_0^{l-n} \mathbf{X}^n$, i pozostaje skorzystać z tego, że $\frac{1}{n!} D^n p = \binom{l}{n} x^{l-n}$ dla $n \leq l$. $\square$

Wniosek 1. *Wartość $p(\mathbf{J}(\lambda, s))$ wielomianu p na klatce Jordana $\mathbf{J}(\lambda, s)$ jest macierzą dolnie trójkątną, w której na przekątnej stoją wyrazy równe $p(\lambda)$, bezpośrednio poniżej nich – wyrazy równe $(Dp)(\lambda)/1!$, poniżej–wyrazy $(D^2p)(\lambda)/2!$ itd, aż do wyrazu $(D^{s-1}p)(\lambda)/(s-1)!$, znajdującego się w lewym dolnym rogu. W szczególności, macierz $p(\mathbf{J}(\lambda, s))$ zależy tylko od wartości $p(\lambda), (Dp)(\lambda), ..., (D^{s-1}p)(\lambda)$.*

Dowód. Wynika to ze wzoru (6), przy $\mathbf{X}_0 = \lambda\mathbf{I}$, $\mathbf{X} = \mathbf{J}(0, s)$, i zadania 1 w p.2. □

Uwaga 2. Gdy $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_t)$, gdzie $\mathbf{J}_i$ są klatkami Jordana, to $p(\mathbf{A}) = \mathbf{S}\mathrm{diag}(p(\mathbf{J}_1), ..., p(\mathbf{J}_t))\mathbf{S}^{-1}$, a wartość $p(\mathbf{J}_i)$ dla każdej klatki $\mathbf{J}_i$ wyznaczyć można w oparciu o wniosek. Macierz $p(\mathbf{A})$ zależy więc tylko od wartości $(D^jp)(\lambda)$, dla $\lambda \in \mathrm{spec}(\mathbf{A})$ i $j \in \{0, \ldots, m_\lambda - 1\}$, gdzie m_λ to rozmiar największej spośród klatek $\mathbf{J}_i$, odpowiadających wartości λ. (Z zadania 2c) w p.2 wynika, że $m_\lambda = \inf\{n : \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n = \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n+1}\}$.)

Uwaga 3. Przyjmijmy oznaczenia uwagi 2, lecz niech dodatkowo $\mathrm{spec}_{\mathbb{C}}(\mathbf{A}) \subset \mathbb{R}$ gdy $\mathbb{F} = \mathbb{R}$. Funkcję skalarną f nazwiemy **A–dopuszczalną**, jeśli jest ona określona na pewnym otoczeniu U_f widma $\mathrm{spec}(\mathbf{A})$ w $\mathbb{F}$ i ma w każdym punkcie $\lambda \in \mathrm{spec}(\mathbf{A})$ pochodne $(Df)(\lambda), ..., (D^{m_\lambda-1}f)(\lambda)$. (Przy $\mathbb{F} = \mathbb{C}$ chodzi o pochodne zespolone.) Dla każdej takiej funkcji definiujemy macierz $f(\mathbf{A})$ następująco:

$*$ gdy $\mathbf{A} = \mathbf{J}(\lambda, s)$ jest klatką Jordana, to $f(\mathbf{A})$ jest dolnie trójkątną macierzą opisaną we wniosku 1, z p zamienionym przez f;

$*$ gdy $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_t)$ i $\mathbf{J}_i$ to klatki Jordana, to $f(\mathbf{A}) = \mathbf{S}\mathrm{diag}\left(f(\mathbf{J}_1), ..., f(\mathbf{J}_t)\right)\mathbf{S}^{-1}$.

Wynik jest niezależny od użytej macierzy $\mathbf{S}$ i kolejności klatek $\mathbf{J}_1, \ldots, \mathbf{J}_t$, bo $f(\mathbf{A}) = p(\mathbf{A})$ dla dowolnego wielomianu p takiego, że $D^jp(\lambda) = D^jf(\lambda)$ dla $\lambda \in \mathrm{spec}(\mathbf{A})$ i $0 \leq j < m_\lambda$. (Wielomiany takie istnieją, nawet stopnia $< k$, por. zad. uz. 4 w §III.5.1.) Jak w części g) zadania 1 przekonujemy się, że gdy $(f_n)_{n=0}^{\infty}$ jest ciągiem funkcji **A**–dopuszczalnych i $D^if_n(\lambda) \to D^if_0(\lambda)$ dla $\lambda \in \mathrm{spec}(\mathbf{A})$ i $0 \leq i < m_\lambda$, to $f_n(\mathbf{A}) \to f_0(\mathbf{A})$. Dla przykładu, macierz $e^{\mathbf{A}}$ obliczona zgodnie z tą definicją jest zarazem granicą ciągu $\mathbf{I} + \mathbf{A} + ... + \frac{1}{n!}\mathbf{A}^n$. (Dlaczego?)

Ćwiczenie. Obliczyć $e^{\mathbf{A}}$ dla $\mathbf{A} = \mathbf{J}(2, 3)$.

Zadanie 2. * Udowodnić, że iloczyn **A**–dopuszczalnych funkcji f_1, f_2 też jest funkcją **A**–dopuszczalną i $(f_1 \cdot f_2)(\mathbf{A}) = f_1(\mathbf{A})f_2(\mathbf{A})$. Również inne części zadania 1 są słuszne po podobnych zmianach. (Wskazówka: dla pewnych $p_i \in \mathbb{F}[x]$ zachodzi $p_i(\mathbf{A}) = f_i(\mathbf{A})$, Porównać pochodne f_1f_2 i p_1p_2.)

Zadania uzupełniające. (W zadaniach 1 i 2 zacząć od macierzy Jordana.)

1. Dla kwadratowej macierzy zespolonej $\mathbf{A}$ dowieść, że $\lim_{n\to\infty} \mathbf{A}^n = \mathbf{0} \Leftrightarrow |\lambda| < 1$ dla $\lambda \in \mathrm{spec}(\mathbf{A})$.

2. Niech wielomian charakterystyczny macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ rozkłada się nad $\mathbb{F}$ na czynniki liniowe: $\chi_{\mathbf{A}} = (\lambda_1 - x)...(\lambda_k - x)$, gdzie $\lambda_1, ..., \lambda_k \in \mathbb{F}$. Dowieść, że dla każdego wielomianu $p \in \mathbb{F}[x]$ mamy $\chi_{p(A)} = (p(\lambda_1) - x)...(p(\lambda_k) - x)$; w szczególności, $\mathrm{tr}(p(\mathbf{A})) = \sum_i p(\lambda_i)$ oraz $\det(p(\mathbf{A})) = \prod_i p(\lambda_i)$. Wywnioskować, że $\det(e^{\mathbf{A}}) = e^{\mathrm{tr}A}$ (tu $\mathbb{F} = \mathbb{R}$).

3. Dowieść, że dla $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ i $p \in \mathbb{F}[x]$ zachodzi $p(\mathbf{A}) = r(\mathbf{A})$, gdzie r to reszta z dzielenia p przez $\chi_{\mathbf{A}}$ (lub z dzielenia przez wielomian c taki, że $c(\mathbf{A}) = \mathbf{0}$). Wywnioskować, że gdy $\mathbf{A} \in \mathcal{M}_2(\mathbb{R})$ jest macierzą ortogonalną o dodatnim wyznaczniku, to $\mathrm{tr}(\mathbf{A}^{10})$ można wyrazić przez $\mathrm{tr}(\mathbf{A})$.

4. a) Dowieść, że macierz zespolona, której wszystkie wartości własne są dodatnie, jest kwadratem pewnej macierzy o takiejże własności.

b)* Dowieść, że można wyżej zastąpić „pewnej” przez „jedynej”. (Wskazówka: zbiór rozważanych macierzy oznaczmy przez P. Dowieść, że funkcje $P \ni \mathbf{A} \mapsto \mathbf{A}^2$ i $P \ni \mathbf{A} \mapsto \sqrt{\mathbf{A}}$ mają wartości w P i są wzajemnie odwrotne.)

5. Dla wielomianu $p \in \mathbb{F}[x]$ i macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, której wielomian charakterystyczny rozkłada się nad $\mathbb{F}$ na czynniki liniowe, dowieść równoważności warunków:

a) $p(\mathbf{A}) = \mathbf{0}$;
b) $(D^j p)(\lambda) = 0$ dla $\lambda \in \operatorname{spec}(\mathbf{A})$ i $0 \le j \le m_\lambda - 1$ (liczby m_λ zdefiniowano w uwadze 2);
c) wielomian p jest podzielny przez $q := \prod_{\lambda \in \operatorname{spec}(\mathbf{A})} (x - \lambda)^{m_\lambda}$

Uwaga 4. * Powyższy wielomian q nazywamy **wielomianem minimalnym** macierzy $\mathbf{A}$; ma on najniższy stopień spośród wszystkich, które zerują się na $\mathbf{A}$. Ponieważ $q | \chi_{\mathbf{A}}$, patrz zadanie 2 w p.2, więc otrzymujemy **twierdzenie Cayley'a–Hamiltona**: $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$ dla każdej zespolonej macierzy kwadratowej. (Inny dowód będzie w zad. uz. 6 w p.4.)

6. Udowodnić część c) uwagi 3 z p.2.

Zadania ze zbioru Kostrykina: 3, 17, 18, 19, 20, 43, 45 w §II.3.3.

4. Podprzestrzenie niezmiennicze (przygotowanie do dowodu twierdzenia Jordana).

Niech L będzie operatorem na przestrzeni wektorowej V.

Definicja. a) Podprzestrzeń V_0 przestrzeni V nazwiemy L–**niezmienniczą**, jeśli $L(V_0) \subset V_0$.
b) Na podprzestrzeni takiej, L wyznacza operator $L_0 \in \mathcal{L}(V_0)$, zadany wzorem $L_0(\mathbf{v}) = L(\mathbf{v})$ dla $\mathbf{v} \in V_0$. Oznaczamy go $L_{|V_0}$ i nazywamy **operatorem indukowanym** przez L na V_0.
c) Gdy $V = \mathbb{F}^k$ i $L = L_{\mathbf{A}}$ dla pewnej macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, zamiast o podprzestrzeni $L_{\mathbf{A}}$–niezmienniczej mówimy o $\mathbf{A}$–niezmienniczej.

Przykład 1. Każda podprzestrzeń własna operatora L jest L–niezmiennicza. Gdy L jest obrotem przestrzeni euklidesowej $\mathbb{R}^3$, to oś obrotu i płaszczyzna ortogonalna do niej są L–niezmiennicze. Operatory indukowane to homotetia, identyczność i obrót płaszczyzny, odpowiednio. ⊟

Uwaga 1. a) Gdy podprzestrzezestrzeń U jest L–niezmiennicza i jej bazę $\mathcal{U} = (\mathbf{u}_i)_{i=1}^s$ rozszerzyć do bazy $\mathcal{V} = (\mathbf{u}_1, ..., \mathbf{u}_s, \mathbf{w}_1, ..., \mathbf{w}_t)$ całej przestrzeni, to macierz $[L]_{\mathcal{V}}$ można rozbić na 4 klatki, z których lewą górną klatką jest macierz $[L_0]_{\mathcal{U}}$ operatora indukowanego $L_0 := L_{|U}$, a lewa dolna klatka jest zerowa. (Wynika to z definicji macierzy $[L]_{\mathcal{V}}$.)
b) Gdy zaś baza $\mathcal{V}$ przestrzeni V jest postaci $(\mathbf{u}_1, ..., \mathbf{u}_p, \mathbf{w}_1, ..., \mathbf{w}_q)$, gdzie $\mathcal{U} = (\mathbf{u}_1, ..., \mathbf{u}_p)$ i $\mathcal{W} = (\mathbf{w}_1, ..., \mathbf{w}_q)$ są bazami podprzestrzeni L–niezmienniczych , to $[L]_{\mathcal{V}} = \operatorname{diag}(\mathbf{A}, \mathbf{B})$ dla pewnych klatek $\mathbf{A} \in \mathcal{M}_p, \mathbf{B} \in \mathcal{M}_q$ (równych $[L_{|U}]_{\mathcal{U}}$ i $[L_{|W}]_{\mathcal{W}}$, odpowiednio). Implikacja odwrotna też ma miejsce, i tak samo jest przy rozbiciu $\mathcal{V}$ na większą liczbę baz podprzestrzeni niezmienniczych.

Zadanie 1. Niech $L \in \mathcal{L}(V)$. W oparciu o uwagę 1 dowieść, że:
a) Wielomian charakterystyczny operatora, indukowanego przez L na podprzestrzeni niezmienniczej, jest dzielnikiem wielomianu χ_L.
b) Gdy zaś $V = V_0 \oplus V_1$ i podprzestrzenie V_i są L–niezmiennicze, to $\chi_L = \chi_{L_0} \cdot \chi_{L_1}$, gdzie L_i oznacza operator indukowany przez L na podprzestrzeni V_i.

Stwierdzenie 1. *Gdy operator $L \in \mathcal{L}(V)$ jest osobliwy, to istnieją podprzestrzenie L–niezmiennicze U i W takie, że $U \oplus W = V$, $W \neq V$ i $L^k(U) = \{\mathbf{0}\}$ dla $k = \dim V$. Ponadto, można przyjąć $U := \ker(L^k)$ i żądać, by $L(W) = W$.*

Dowód. Skoro $V \supset L(V) \supset L^2(V) \supset ...$, to $\dim L^n(V) = \dim L^{n+1}(V)$ dla pewnego $n \le k$. Wówczas dla $W := L^n(V)$ zachodzi $L(W) = W$ i wobec tego $L^k(W) = W$. Stąd $W \cap U = \{\mathbf{0}\}$ dla $U := \ker(L^k)$; a że ponadto $\dim \ker(L^k) + \dim \operatorname{im}(L^k) = \dim V$, więc $V = U \oplus W$. (Patrz twierdzenia 1 w §III.5.1 i w §III.6.2.) Pozostaje zauważyć, że $W \neq V$, bo $U = \ker(L^k) \neq \{\mathbf{0}\}$ wobec osobliwości L. □

Zadanie 2. a) Jeśli $V_0 \supset \operatorname{im}(L)$ lub $V_0 \subset \ker(L)$, to podprzestrzeń V_0 jest L-niezmiennicza.
b) Podprzestrzeń L–niezmiennicza jest i $p(L)$–niezmiennicza, dla $p \in \mathbb{F}[x]$.

c) Niech $K, L \in \mathcal{L}(V)$ i $KL = LK$. Gdy podprzestrzeń V_0 jest L–niezmiennicza, to $K(V_0)$ i $K^{-1}(V_0)$ też są takie. W szczególności, L–niezmiennicze są $\ker(K)$, $\operatorname{im}(K)$, czy ogólniej podprzestrzenie $\ker(p(K))$ i $\operatorname{im}(p(K))$, dla $p \in \mathbb{F}[x]$, w tym podprzestrzenie własne operatora K.

Zadania uzupełniające.

1. Niech $\mathbf{J}$ będzie klatką Jordana stopnia k. Dowieść, że jedynymi $\mathbf{J}$ –niezmienniczymi podprzestrzeniami przestrzeni $\mathbb{F}^k$ są $\{\mathbf{0}\}$ i przestrzenie $V_i = \operatorname{lin}(\mathbf{e}_i, \ldots, \mathbf{e}_k)$, $1 \le i \le k$. (Wskazówka: uprościć sobie zadanie odejmując od macierzy $\lambda\mathbf{I}$.)

2. Niech W będzie przestrzenią niezmienniczą diagonalizowalnego operatora $L \in \mathcal{L}(V)$.

   a) Udowodnić, że gdy $V = \oplus_\lambda V_\lambda$ jest rozkładem na podprzestrzenie własne operatora L, to $W = \oplus_\lambda (W \cap V_\lambda)$. (Wskazówka: należy dowieść, że gdy $\sum_\lambda \mathbf{v}_\lambda \in W$ dla pewnych $\mathbf{v}_\lambda \in V_\lambda$, to $\mathbf{v}_\lambda \in W$ $\forall \lambda$. Wykorzystać dowód stwierdzenia 2 z p.1.)

   b) Wywnioskować, że operator indukowany $L_{|W}$ jest diagonalizowalny, i że $V = W \oplus W'$ dla pewnej L–niezmienniczej podprzestrzeni $W' \subset V$.

3. Niech $L \in L(V)$, gdzie χ_L rozkłada się na czynniki liniowe, i niech $L' := L_{|V'} \in \mathcal{L}(V')$ dla pewnej L–niezmienniczej podprzestrzeni $V' \subset V$. Dla $n \in \mathbb{N}$ i $\lambda \in \mathbb{F}$ dowieść, że:

   a) liczba $\operatorname{rk}(L^{n-1}) - \operatorname{rk}(L^n)$ jest równa $\dim\big(\ker(L_{|L^{n-1}(V)})\big)$.

   b) $p_n(\lambda) \ge p'_n(\lambda)$, gdzie obie strony mają to znaczenie, co w tw.2 z p.2.

   c)* Czy zawsze $q_n(\lambda) \ge q'_n(\lambda)$? (Nie znam odpowiedzi.)

4. * a) Dowieść, że gdy $\{L_i\}_{i \in I} \subset \mathcal{L}(V)$ jest przemienną rodziną diagonalizowalnych operatorów, to istnieje baza przestrzeni V, diagonalizująca każdy z nich.

   b) Wywnioskować, że gdy dwie przemienne macierze zespolone są diagonalizowalne i mają tylko nieujemne wartości własne, to ich iloczyn też jest taki.

5. * Niech operatory $L_i \in \mathcal{L}(V)$ $(i \in I)$ będą przemienne, przy czym pewien z nich ma wektor własny. Dowieść, że:

   a) Istnieje (wspólny) wektor, własny dla każdego operatora L_i.

   b) $V = U \oplus W$ dla pewnych różnych od V podprzestrzeni U, W, które są L_i–niezmiennicze dla każdego $i \in I$.

6. * Ustalmy operator $L \in \mathcal{L}(V)$ i niezerowy wektor $\mathbf{v} \in V$.

   a) Dowieść istnienia liczby $s \in \mathbb{N}$ i skalarów $c_0, ..., c_s$ takich, że $L^{s+1}(\mathbf{v}) = \sum_{i=0}^{s} c_i L^i(\mathbf{v})$.

   b) Dowieść, że jeśli $c_0, ..., c_s$ jest najkrótszym z takich ciągów (tzn. liczba s jest najmniejsza z możliwych), to $\mathbf{v}, L(\mathbf{v}), ..., L^s(\mathbf{v})$ jest bazą L–niezmienniczej podprzestrzeni V_0, a wielomian charakterystyczny p indukowanego operatora $L_0 \in \mathcal{L}(V_0)$ jest równy $x^{s+1} - \sum_{i=0}^{s} c_i x^i$ i spełnia warunek $p(L)(\mathbf{v}) = \mathbf{0}$. (Por. przykład 2 w §IV.3.1.)

   c) Wywnioskować, że $\chi_L(L)(\mathbf{v}) = \mathbf{0}$ i wobec tego $\chi_L(L) = 0$, por. uwagę 4 w p.3.

7. * Udowodnić, że każdy operator $L \in \mathcal{L}(\mathbb{R}^k)$ ma podprzestrzeń niezmienniczą wymiaru 1 lub 2.

Zadania ze zbioru Kostrykina: 7, 31 i 32 w §II.3.2.

5. Dowód twierdzenia Jordana

Potrzebne będą definicja i lemat.

Definicja. **Macierz** kwadratowa $\mathbf{B}$ jest **nilpotentna**, jeśli $\mathbf{B}^n = \mathbf{0}$ dla pewnego k; analogicznie, **operator** $L \in \mathcal{L}(V)$ jest **nilpotentny**, jeśli $L^n = 0$ dla pewnego $n \in \mathbb{N}$.

Lemat 1. *Niech operator* $L \in \mathcal{L}(V)$*, gdzie* $V \neq \{\mathbf{0}\}$*, będzie nilpotentny. Wówczas istnieje baza* $\mathcal{V}$ *przestrzeni* V *taka, że macierz* $[L]_{\mathcal{V}}$ *jest postaci* $\mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_t)$*, gdzie* $\mathbf{J}_i$ *to klatki Jordana, odpowiadające wartości 0.*

Udowodnimy wpierw twierdzenie Jordana (=tw. 1 z p.2), przyjmując lemat za prawdziwy.

Dowód twierdzenia Wykorzystamy indukcję względem liczby $k = \dim V$, dzieląc krok indukcyjny na 2 części. (Gdy $k = 1$ nie ma czego dowodzić.)

a) Załóżmy dodatkowo, że $\det(L) = 0$. Na podstawie stwierdzenia z p.4 istnieje rozkład $V = U \oplus W$ na L–niezmiennicze podprzestrzenie $U, W \subset V$ takie, że $L^k(U) = \{\mathbf{0}\}$ i $\dim W < k$. Z lematu wynika istnienie bazy Jordana dla indukowanego operatora $U \to U$, a z założenia indukcyjnego – istnienie jej dla indukowanego operatora $K \in \mathcal{L}(W)$. (Korzystamy z tego, że wielomian χ_K dzieli χ_L, wiec rozkłada się na czynniki liniowe.) Łącznie wzięte, bazy te dają bazę Jordana dla operatora L.

b) Gdy $\det(L) \neq 0$, to założenie z a) stosuje się do operatora $L' = L - \lambda I$, gdzie λ jest dowolnym pierwiastkiem wielomianu χ_L; patrz uwaga 1 w §1.4. Istnieje więc baza Jordana dla L'. Jest ona zarazem bazą Jordana dla operatora $L = L' + \lambda I$, bo $\mathbf{J}(\mu, s) + \lambda \mathbf{I} = \mathbf{J}(\mu + \lambda, s)$ dla $\lambda, \mu \in \mathbb{F}$ i $s \in \mathbb{N}$.

Dowód lematu.* Dla uniknięcia zbędnych indeksów stworzymy pewną bazę nieuporządkowaną A, a potem ją uporządkujemy. Są więc 2 kroki (zasadniczy jest pierwszy):

a) Dowiedziemy istnienia bazy A przestrzeni V takiej, że $L(A) \subset A \cup \{\mathbf{0}\}$ oraz

$$\text{gdy } L(\mathbf{v}) = L(\mathbf{w}) \neq \mathbf{0}, \text{ gdzie } \mathbf{v}, \mathbf{w} \in A, \text{ to } \mathbf{v} = \mathbf{w}. \tag{7}$$

Wyjaśnijmy wpierw sens tych warunków. Każdy wektor $\mathbf{a} \in A \backslash L(A)$ wyznacza ciąg $\mathcal{V}_{\mathbf{a}} = (\mathbf{a}, L(\mathbf{a}), ..., L^d(\mathbf{a}))$, gdzie $d := \max\{s : L^s(\mathbf{a}) \neq \mathbf{0}\} < \infty$ (bo L jest nilpotentny). Ciąg ten nazwiemy **strumieniem**, ze **źródłem a** i **ujściem** $L^d(\mathbf{a})$. Skoro $L(A) \subset A \cup \{\mathbf{0}\}$, to strumień składa się z wektorów zbioru A. Warunek (7) oznacza, że strumienie o różnych źródłach (zawsze z $A \setminus L(A)$) są rozłączne. (Dlaczego?)

Do konstrukcji bazy A wykorzystamy indukcję względem liczby $n := \inf\{i \in \mathbb{N} : L^i = \mathbf{0}\}$. Gdy $n = 1$, to $L = 0$ i A może być dowolną bazą. Dla $n > 1$ rozważamy operator $L_1 : L(V) \to L(V)$, indukowany przez L na L–niezmienniczej podprzestrzeni $L(V)$. Ponieważ $L_1^{n-1} = 0$, więc z założenia indukcyjnego istnieje baza B przestrzeni $L(V)$, spełniająca żądane warunki przy A zastąpionym przez B. Wykorzystując to, że $B \subset L(V)$, obierzmy dla każdego $\mathbf{b} \in B$ pewien wektor $\widetilde{\mathbf{b}} \in L^{-1}(\mathbf{b})$, tak, by $\widetilde{\mathbf{b}} \in B$ gdy $\mathbf{b} \in L(B)$. Wobec (7), wybór $\widetilde{\mathbf{b}}$ jest dla $\mathbf{b} \in L(B)$ jedyny.

Rozszerzmy zbiór $B \cap \ker(L)$ pewnym zbiorem Z_0 do bazy Z przestrzeni $\ker(L)$. Z lematu 1 w §III.5.1 wynika, że poniższy zbiór A jest bazą przestrzeni V:

$$A := Z \cup \{\widetilde{\mathbf{b}} : \ \mathbf{b} \in B\}.$$

Jednak, w ślad za B, zbiór A też rozpada się na rozłączne strumienie: są nimi strumienie o źródłach w $\{\widetilde{\mathbf{b}} : \mathbf{b} \in B \setminus L(B)\}$, prócz źródła płynące w B, oraz strumienie o źródłach w Z_0 (te składają się tylko ze źródła).

b) Uporządkujemy teraz bazę A. Jak odnotowaliśmy, A rozpada się na rozłączne strumienie $\mathcal{V}_i = (\mathbf{a}_i, L(\mathbf{a}_i), ..., L^{d_i}(\mathbf{a}_i))$, gdzie $\mathbf{a}_i$ przebiega zbiór $A \setminus L(A)$ wszystkich źródeł. Za $\mathcal{V}$ obierzmy zbiór A uporządkowany tak, by jako początkowe występowały kolejne wyrazy strumienia $\mathcal{V}_1$, jako następne – kolejne wyrazy strumienia $\mathcal{V}_2$ itd. Jest oczywiste, że $[L]_{\mathcal{V}} = \mathrm{diag}(\mathbf{J}(0, d_1 + 1), \mathbf{J}(0, d_2 + 1), ...)$. $\square$

Zadanie uzupełniające 1. Dowieść, że gdy $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ i $\mathbf{A}^n = \mathbf{0}$, to $\mathrm{rk}(\mathbf{A}) \leq k - k/n$.

Zadania ze zbioru Kostrykina: 18, 20, 25, 26 w §I.4.3.

6. Znajdywanie bazy Jordana.

Niech $\mathcal{V} = (\mathbf{v}_1, ..., \mathbf{v}_k)$ będzie bazą Jordana dla operatora $L \in \mathcal{L}(V)$. O wektorze $\mathbf{v}_i$ tej bazy powiemy, że odpowiada wartości λ, jeśli λ jest i–tym wyrazem przekątnej macierzy $[L]_{\mathcal{V}}$. (Nie musi to być wektor własny, odpowiadający λ.)

Uwaga 1. Dowód twierdzenia Jordana wskazuje, że bazę Jordana dla operatora L konstruować można następująco. Obierzmy $\lambda \in \operatorname{spec}(L)$ dowolnie; zastępując L przez $L - \lambda I$ możemy założyć, że $\lambda = 0$. Niech $V = U \oplus W$ będzie rozkładem, o którym mowa w stwierdzeniu 1 z p.4. Dowód nakazuje szukaną bazę przestrzeni V utworzyć przez połączenie baz Jordana: $\mathcal{U}$ dla operatora $L_{|U} \in \mathcal{L}(U)$, i $\mathcal{W}$ dla operatora $L_{|W} \in \mathcal{L}(W)$. Przy tym, wszystkie wektory bazy Jordana dla $L_{|U}$ odpowiadają wartości 0, zaś żaden wektor bazy $\mathcal{W}$ tej wartości nie odpowiada (bo $L_{|W}$ jest izomorfizmem, skąd jego macierz Jordana nie ma zera na przekątnej). Po wyznaczeniu bazy $\mathcal{U}$ możemy więc zastępować L przez $L - \mu I$ dla innych wartości własnych μ, by w ten sam sposób wyznaczać wektory szukanej bazy, odpowiadające pozostałym wartościom własnym.

Uwaga 2. * Przyjrzyjmy się, jak ostatecznie konstruować można wektory bazy $\mathcal{U}$. Dla małych wartości $\dim V$ wystarczającą może być ta własność, że $\mathcal{U}$ rozpada się na rozłączne strumienie, w liczbie równej $\dim V_L(0)$. (Patrz zadanie 2 w p.2.) W ogólnym przypadku przyjmijmy

$$V_i = L^i(V) \text{ dla } i = 0, 1, \ldots \text{ oraz } n = \max\{i : V_i \neq V_{i+1}\}$$

i niech A_n będzie dowolną bazą przestrzeni $V_n \cap \ker(L)$. Gdy znamy już zbiór A_i wektorów przestrzeni V_i, $i \leq n$, będący sumą rozłącznych strumieni, to

a) rozszerzamy $A_i \cap \ker(L)$ pewnym zbiorem Z_{i-1} do bazy przestrzeni $\ker(L) \cap V_{i-1}$;

b) znajdujemy, dla każdego strumienia $S = (\mathbf{v}, L(\mathbf{v}), \ldots)$ zbioru A_i, wektor $\widetilde{\mathbf{v}} \in L^{-1}(\mathbf{v}) \cap V_{i-1}$, i tworzymy strumień $\widetilde{S} = (\widetilde{\mathbf{v}}, \mathbf{v} = L(\widetilde{\mathbf{v}}), \ldots)$.

Zbiór A_{i-1} tworzą strumienie $\widetilde{S}$, gdzie S przebiega strumienie bazy A_i, oraz strumienie $\{z\}, z \in Z_{i-1}$. Twierdzimy, że postępując w ten sposób uzyskujemy na koniec szukaną bazę Jordana $A_0 = \mathcal{U}$ dla operatora $L_{|U}$.

Istotnie, gdy operator L jest nilpotentny, wynika to wprost z dowodu lematu 1 z p.5. Gdy nie jest, możemy wyznaczyć $\mathcal{U}$ jak wyżej, lecz z V zastąpionym przez $U = \ker(L^k)$, a L przez $L_{|U}$. Znajdywanie przestrzeni U i kolejnych $U_i = L^i(U)$, oraz dokonywanie zastąpień, jest jednak zbędne. Gdy bowiem przyjąć $V_i = L^i(V)$, to wobec $L = L_{|U} \oplus L_{|W}$, gdzie $L_{|W}$ jest izomorfizmem, otrzymujemy $V_i = U_i \oplus W$. Liczba $n := \max\{i : V_i \neq V_{i+1}\}$ jest więc równa $\max\{i : U_i \neq \{\mathbf{0}\}\}$. Ponieważ ponadto $V_i \cap \ker(L) \subset U_i$ i $L^{-1}(U_{i+1}) \cap V_i = U_i$, więc podana recepta (bez zastępowania V_i przez U_i etc.) da zawsze $A_i \subset U_i$, a na końcu otrzymamy szukaną bazę Jordana A_0 dla operatora $L_{|U}$. $\square$

Uwaga 3. * W b) można unikać rozwiązywania równania liniowego $L(\mathbf{x}) = \mathbf{v}$, jeśli każdy element bazy A_i tworzyć tak, by znane było jego przedstawienie w postaci kombinacji liniowej elementów zbioru $L^i(G)$, gdzie G jest pewną bazą przestrzeni V. (Gdy bowiem $\mathbf{v} = \sum_{\mathbf{g} \in G} c_{\mathbf{g}} L^i(\mathbf{g})$, to $L(\widetilde{\mathbf{v}}) = \mathbf{v}$ dla $\widetilde{\mathbf{v}} = \sum_{\mathbf{g} \in G} c_{\mathbf{g}} L^{i-1}(\mathbf{g})$.)

Przykład 1. * Niech V będzie przestrzenią wielomianów rzeczywistych stopnia ≤ 2 zmiennych x, y, z, zaś $L \in \mathcal{L}(V)$ operatorem zadanym wzorem $L(f) := \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z} + \frac{\partial^2 f}{\partial x^2}$ dla $f \in V$. Można też o L myśleć jako o operatorze wyznaczonym tabelą, w której najwyższym wierszu wypisano elementy bazy przestrzeni V, zaś poniżej – ich obrazy przy kolejnych iteracjach operatora L (już 2 najwyższe wiersze wyznaczają L):

nr.0	1	x	y	z	x^2	y^2	z^2	xy	xz	yz
nr.1	0	1	1	1	$2x+2$	$2y$	$2z$	$x+y$	$x+z$	$y+z$
nr.2	0	0	0	0	2	2	2	2	2	2
nr.3	0	0	0	0	0	0	0	0	0	0

Oznaczmy przez G_i zbiór wyrazów wiersza o numerze i. Za bazę przestrzeni $V_2 = \mathrm{lin}(G_2)$ przyjąć można $A_2 = \{2\}$, przy czym 2 przedstawiamy jako kombinację wyrazów wiersza nr 2. Można n.p. za piąty współczynnik kombinacji przyjąć 1, a za pozostałe 0; wtedy $2 = L(2x + 2)$. Następnie $\{2\}$ uzupełniamy do bazy przestrzeni $\ker(L) \cap V_1$. Jak widać z tabeli, przestrzenią $V_1 = \mathrm{lin}(G_1)$ jest zbiór wielomianów stopnia ≤ 1, skąd $\ker(L) \cap V_1 = \{a+bx+cy+dz : b+c+d = 0\}$. (Uwzględniliśmy definicję operatora L.) Jest to przestrzeń wymiaru 3, a elementy zbioru rozszerzającego $\{2\}$ do jej bazy obieramy tak, by ułatwić przedstawienie ich w postaci kombinacji liniowej wyrazów wiersza nr.1; np. przyjmujemy $Z_1 = \{x-y, x-z\}$. Baza A_1 przestrzeni V_1 jest sumą strumienia $\{2x+2, 2\}$ i strumieni $\{x-y\}$ i $\{x-z\}$. By utworzyć A_0 stwierdzamy, że „przodkiem" wektora $2x+2$ jest x^2 (tzn. $L(x^2) = 2x+2$), wektora $x-y$ jest $xy-y^2$ (bo $(x+y)-2y = L(xy)-L(y^2)$), a wektora $x-z$ jest $xz-z^2$. Szukana baza A_0 przestrzeni V jest więc sumą strumieni $\{x^2, 2x+2, 2\}$, $\{xy-y^2, x-y\}$, $\{xz-z^2, x-z\}$ i zbioru Z_0, takiego, że łącznie z $\{2, x-y, x-z\}$ tworzy on bazę $\ker(L)$. Zgodnie z zadaniem .... można przyjąć $Z_0 = \{x^2-xy-xz+yz-2x, y^2-xy-yz+xz, z^2-xz-yz+xy\}$. Baza A_0, po uporządkowaniu, jest więc taka: $(x^2, 2x+2, 2; xy-y^2, x-y; xz-z^2, x-z; x^2-xy-xz+yz-2x; y^2-xy-yz+xz; z^2-xz-yz+xy)$, gdzie średnikami oddzielono strumienie bazy. „Wytwarzają" one klatki Jordana stopni 3, 2, 2, 1, 1, 1, odpowiednio, odpowiadające wartości 0. Macierz operatora L w tej bazie ma więc 100 wyrazów, lecz tylko 4 różne od 0: dwie jedynki w klatce stopnia 3 i po jednej w klatkach stopnia 2. $\square$

7. ** Uogólniona postać Jordana macierzy rzeczywistych.

Założenie rozkładalności wielomianu charakterystycznego na czynniki liniowe, zasadnicze dla p.1, jest przy $\mathbb{F} = \mathbb{R}$ spełnione tylko dla niektórych macierzy czy operatorów. Podamy wersję twierdzenia Jordana, stosowalną do dowolnej macierzy rzeczywistej. W tym celu dla $\lambda = a + bi \in \mathbb{C} \setminus \mathbb{R}$ i $s \in \mathbb{N}$ oznaczmy przez $\mathbf{J}^{\mathbb{R}}(\lambda, 2s)$ macierz rzeczywistą stopnia $2s$, w której wzdłuż przekątnej umieszczone są 2×2-klatki $\mathbf{K}(\lambda)$ (patrz poniżej), pod każdą z nich, prócz ostatniej –klatka jednostkowa $\mathbf{I}_2$, zaś na pozostałych miejscach zera:

$$\mathbf{J}^{\mathbb{R}}(\lambda, 2s) := \begin{pmatrix} \mathbf{K}(\lambda) & & & & 0 \\ \mathbf{I}_2 & \mathbf{K}(\lambda) & & & \\ & \mathbf{I}_2 & \ddots & & \\ & & \ddots & \mathbf{K}(\lambda) & \\ 0 & & & \mathbf{I}_2 & \mathbf{K}(\lambda) \end{pmatrix} \in \mathcal{M}_{2s}(\mathbb{R}), \quad \text{gdzie } \mathbf{K}(\lambda) := \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Uogólnioną macierzą Jordana nazwiemy macierz będącą zewnętrzną sumą klatek, z których każda jest bądź klatką Jordana odpowiadającą rzeczywistej wartości własnej, bądź jest postaci $\mathbf{J}^{\mathbb{R}}(\lambda, 2s)$ dla pewnych $s \in \mathbb{N}$ i $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Lemat 1. ** *Macierz* $\mathbf{J}^{\mathbb{R}}(\lambda, 2s)$ *jest nad* $\mathbb{C}$ *podobna do macierzy* $\mathbf{J} = \mathrm{diag}(\mathbf{J}(\lambda, s), \mathbf{J}(\overline{\lambda}, s))$.

Dowód. Oznaczmy przez $\mathbf{e}_1, ..., \mathbf{e}_{2s}$ wektory standardowej bazy przestrzeni $\mathbb{C}^{2s}$ i przyjmijmy

$$\mathbf{f}_j := \mathbf{e}_{s+j} \text{ oraz } \mathbf{u}_j := \mathbf{e}_j + \mathbf{f}_j,\ \mathbf{v}_j := i\mathbf{e}_j - i\mathbf{f}_j \text{ dla } j = 1, ..., s.$$

Mamy więc $\mathbf{J}\mathbf{e}_j = \lambda\mathbf{e}_j + \mathbf{e}_{j+1}$, $\mathbf{J}\mathbf{f}_j = \overline{\lambda}\mathbf{f}_j + \mathbf{f}_{j+1}$ dla $j < s$ oraz $\mathbf{J}\mathbf{e}_s = \lambda\mathbf{e}_s, \mathbf{J}\mathbf{f}_s = \lambda\mathbf{f}_s$. Stąd, pisząc $\lambda = a + bi$, otrzymujemy łatwo $\mathbf{J}\mathbf{u}_j = a\mathbf{u}_j + b\mathbf{v}_j + \mathbf{u}_{j+1}$, $\mathbf{J}\mathbf{v}_j = a\mathbf{v}_j - b\mathbf{u}_j + \mathbf{v}_{j+1}$ dla $j < s$ oraz $\mathbf{J}\mathbf{u}_s = a\mathbf{u}_s + b\mathbf{v}_s$, $\mathbf{J}\mathbf{v}_s = a\mathbf{v}_s - b\mathbf{u}_s$. Tak więc w bazie $(\mathbf{u}_1, \mathbf{v}_1, ..., \mathbf{u}_s, \mathbf{v}_s)$ przestrzeni $\mathbb{C}^{2s}$, macierzą przekształcenia $L_{\mathbf{J}}$ jest $\mathbf{J}^{\mathbb{R}}(\lambda, 2s)$. $\square$

Twierdzenie 1. ** *Rzeczywista macierz kwadratowa jest nad $\mathbb{R}$ podobna do pewnej uogólnionej macierzy Jordana. (Równoważnie: w pewnej bazie, macierz operatora na rzeczywistej przestrzeni wektorowej jest uogólnioną macierzą Jordana.)*

Dowód. Ustalmy macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$. Jest ona nad $\mathbb{C}$ podobna do macierzy Jordana $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_q)$, gdzie $\mathbf{J}_i = \mathbf{J}(\lambda_i, s_i)$ dla pewnych $s_1, ..., s_q \in \mathbb{N}$ oraz $\lambda_1, ..., \lambda_q \in \mathbb{C}$. Możemy oczywiście wartości λ_i uporządkować tak, by $\lambda_j \in \mathbb{R} \Leftrightarrow j < p$ dla pewnego $p \leq q$ (dopuszczamy możliwość $p = 1$). Ponieważ macierz $\overline{\mathbf{A}}$ jest podobna do macierzy $\overline{\mathbf{J}}$, a $\mathbf{A}$ ma wyrazy rzeczywiste, więc $\overline{\mathbf{A}} = \mathbf{A}$ i zarówno $\mathbf{J}$, jak i $\overline{\mathbf{J}}$ są postaciami Jordana macierzy $\mathbf{A}$. Z jednoznaczności tej postaci wynika, że ciąg klatek $\mathbf{J}_1, ..., \mathbf{J}_q$ tylko kolejnością różni się od ciągu $\overline{\mathbf{J}_1}, ..., \overline{\mathbf{J}_q}$ Dowolna klatka Jordana $\mathbf{J}_i$ pojawia się więc w nim dokładnie tyle razy, co klatka $\overline{\mathbf{J}_i}$. Przy odpowiednim uporządkowaniu par $(\lambda_p, s_p), ..., (\lambda_q, s_q)$ mamy więc dla pewnego r :

$$\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_{p-1}, \mathbf{J}_p, \overline{\mathbf{J}_p}, ..., \mathbf{J}_r, \overline{\mathbf{J}_r})$$

Z lematu, zastosowanego do każdej z macierzy $\mathrm{diag}(\mathbf{J}_j, \overline{\mathbf{J}_j})$, gdzie $p \leq j \leq r$, wynika, że macierz $\mathbf{A}$ jest nad $\mathbb{C}$ podobna do macierzy $\mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_{p-1}, \mathbf{J}^{\mathbb{R}}(\lambda_p, s_p), ..., \mathbf{J}^{\mathbb{R}}(\lambda_r, s_r))$. Jest więc tak i nad $\mathbb{R}$, na mocy twierdzenia 1 w §1.6. □

Zadanie 1. ** Uogólniona macierz Jordana, podobna do zadanej macierzy rzeczywistej, jest jedyna, z dokładnością do kolejności uogólnionych jordanowskich klatek.

§ 3. Macierze i operatory unitarnie diagonalizowalne.

Umowa: O ile nie powiedziano inaczej, w tym pararafie zakładamy, że $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ i rozważane przestrzenie są unitarne. Przestrzeń $\mathbb{F}^k$ rozpatrujemy ze standardowym iloczynem skalarnym. Gdy nie jest ono istotne, określenie ciała $\mathbb{F}$ pomijamy.

1. Twierdzenie Schura o unitarnym podobieństwie do macierzy trójkątnej.

Twierdzenia Jordana stwierdza w szczególności, że każda kwadratowa macierz zespolona jest podobna do macierzy trójkątnej. Taka jego wersja pozostaje prawdziwa dla unitarnego podobieństwa:

Twierdzenie 1 (I. Schura)**.** *Gdy V jest zespoloną przestrzenią unitarną i $L \in \mathcal{L}(V)$, to istnieje ortonormalna baza przestrzeni V, w której macierz operatora L jest dolnie trójkątna.*

Równoważne sformułowanie: Zespolona macierz kwadratowa jest unitarnie podobna do macierzy dolnie trójkątnej.

Dowód. Dowodzić będziemy wersji „operatorowej”, stosując indukcję względem $k := \dim(V)$. Teza jest oczywista dla $k = 1$; niech więc $k > 1$. Obierzmy wektor własny $\mathbf{w}$ operatora L (taki istnieje, bo $\mathbb{F} = \mathbb{C}$) i oznaczmy przez P rzutowanie wzdłuż $\mathbf{w}$ na podprzestrzeń $U := \mathbf{w}^{\perp}$. Z założenia indukcyjnego, zastosowanego do operatora $PL_{|U} : U \to U$, wynika istnienie ortonormalnej bazy $(\mathbf{b}_1, \ldots, \mathbf{b}_{k-1})$ przestrzeni U, w której macierz tego operatora jest dolnie trójkątna. Ale to oznacza, że w bazie $\mathcal{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_{k-1}, \mathbf{w}/||\mathbf{w}||)$ macierz operatora L jest dolnie trójkątna. Istotnie, jej ostatnia kolumna jest postaci $(0, \ldots, 0, \lambda)$, a kolumna i–ta dla $i < k$ zaczyna się od $i-1$ zer (bo ciąg $[PL(\mathbf{b}_i)]_{\mathcal{B}}$ zaczyna się od $i - 1$ zer, a ciąg $[(L - PL)(\mathbf{b}_i)]_{\mathcal{B}}$ – nawet od $k - 1$ zer.) □

Uwaga 1. a) „Równoważność” wersji operatorowej i macierzowej twierdzenia oznacza możliwość łatwego uzasadnienia jednej wersji, gdy przyjmiemy drugą. Wynika ona bezpośrednio z wniosku 1 w §1.3. W dalszej części będziemy formułować podobne równoległe wersje innych tez, dowodząc często tylko jednej z nich.

b) Twierdzenie 1 ma zasadnicze znaczenie i posłuży w następnym punkcie do charakteryzacji macierzy unitarnie diagonalizowalnych. W poniższych zadaniach uzupełniających podajemy dalsze jego zastosowania.

c) Można uzyskać, by macierz podobna do $\mathbf{A}$ była trójkątna górnie (a nie dolnie), i by wzdłuż jej przekątnej wartości własne występowały w zadanej kolejności.

Zadania uzupełniające. (W 1 i 2a) założyć wpierw, że macierz jest trójkątna.)

1. Udowodnić **nierówność Schura**: dla kwadratowej macierzy zespolonej $\mathbf{A}$ mamy $\sum |a_{ij}|^2 \geq \sum |\lambda_i|^2$, gdzie $\lambda_1, \lambda_2, ...$ są wszystkimi pierwiastkami wielomianu $\chi_{\mathbf{A}}$, każdy powtórzony tyle razy, ile wynosi jego krotność.

2. Niech $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{C})$.
   a) Dowieść, że jeśli $\mathbf{Au} \perp \mathbf{u}$ dla każdego $\mathbf{u} \in \mathbb{C}^k$, to $\mathbf{A} = \mathbf{0}$.
   b) Dowieść, że jeśli $\langle \mathbf{Au}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{Bu} \rangle$ dla każdego wektora $\mathbf{u} \in \mathbb{C}^k$, to $\mathbf{B} = \mathbf{A}^*$
   c) Czy tezy te pozostaną prawdziwe, gdy $\mathbb{C}$ zamienić na $\mathbb{R}$?

3. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ i $\mathrm{tr}(\mathbf{A}) = 0$. Zakładamy, że $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ (ogólniej: że $n1_{\mathbb{F}} \neq 0_{\mathbb{F}}$ dla $n \leq k$).
   a) Dowieść podobieństwa $\mathbf{A}$ do macierzy o zerowej przekątnej. (Wskazówka: dowód tw. Schura.)
   b) Dowieść, że macierz $\mathbf{A}$ jest komutatorem, tzn. $\mathbf{A} = \mathbf{XY} - \mathbf{YX}$ dla pewnych macierzy kwadratowych $\mathbf{X}, \mathbf{Y}$. (Wskazówka: zad. uz. 2b) w §II.2.2.)

4. Dowieść, że rzeczywista macierz kwadratowa, której wszystkie wartości własne są rzeczywiste, jest ortogonalnie podobna do macierzy dolnie trójkątnej.

2. Diagonalizacja unitarna.

Definicja. Operator $L \in \mathcal{L}(V)$, gdzie V to przestrzeń unitarna, nazywamy **unitarnie diagonalizowalnym**, jeśli w pewnej ortonormalnej bazie $\mathcal{V}$ jego macierz $[L]_{\mathcal{V}}$ jest diagonalna. Macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ nazywamy unitarnie diagonalizowalną (nad $\mathbb{F}$), jeśli taki jest wyznaczony przez nią operator $L_{\mathbf{A}} : \mathbb{F}^k \to \mathbb{F}^k$.

Uwaga 1. Mapa $S : V \to \mathbb{F}^k$, wyznaczona przez bazę diagonalizującą operator L, zaświadcza o jego podobieństwie do operatora $\mathbb{F}^k \to \mathbb{F}^k$ o macierzy diagonalnej. (Patrz §1.3.) Mapa ta jest izometrią gdy baza jest ortonormalna. Tak więc operator $L \in \mathcal{L}(V)$ wtedy i tylko wtedy jest unitarnie diagonalizowalny, gdy jest unitarnie podobny do operatora $L' \in \mathcal{L}(\mathbb{F}^k)$, zadanego macierzą diagonalną. Podobnie, macierz unitarnie diagonalizowalna to taka, która jest unitarnie podobna do macierzy diagonalnej. (Definicja i najprostsze własności unitarnego podobieństwa są w §1.1–1.3)

Stwierdzenie 1. *Dla operatora $L \in \mathcal{L}(V)$, gdzie V to przestrzeń unitarna, równoważne są warunki:*
a) operator L jest unitarnie diagonalizowalny;
b) przestrzeń V jest ortogonalną sumą podprzestrzeni własnych operatora L, tzn. $V = \bigoplus_\lambda V_L(\lambda)$, gdzie λ przebiega wszystkie wartości własne operatora L;
c) Istnieje ortonormalna baza przestrzeni V, złożona z wektorów własnych operatora L.

Dowód. a)⇒b). Gdy założymy a), to istnieje izomorfim unitarny $S : V \to \mathbb{F}^k$ taki, że macierz $[K]$ operatora $K := SLS^{-1}$ jest diagonalna. Przestrzeń $\mathbb{F}^k$ jest ortogonalną sumą przestrzeni własnych operatora K (patrz przykład 2 w §2.1). Analogiczny warunek b) jest więc spełniony, bo $S(V_L(\lambda)) = V_K(\lambda)$ dla każdego $\lambda \in \mathbb{F}$, a S zachowuje iloczyn skalarny.

b)⇒c). Dla każdego $\lambda \in \mathrm{spec}(L)$ obierzmy pewną ortonormalną bazę B_λ przestrzeni własnej $V_L(\lambda)$. Gdy zachodzi b), to $B := \bigcup_\lambda B_\lambda$ jest ortonormalną bazą przestrzeni $V = \bigoplus_\lambda V_L(\lambda)$.

c)⇒a). Jak wiemy z uwagi 2 w §2.1, baza z c) diagonalizuje operator L. □

Uwaga 2. Dowód wystarczalności jawnie wskazuje, jak utworzyć możemy bazę ortonormalną diagonalizującą dany operator (gdy taka istnieje). Gdy diagonalizujemy macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, możemy B_λ otrzymać ortonormalizując metodą Grama-Schmidta układ fundamentalny rozwiązań równania $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$; macierz zaś $\mathbf{S}$ o kolumnach przebiegających zbiór B jest unitarna i czyni $\mathbf{S}^{-1}\mathbf{AS}$ macierzą diagonalną. Tak więc procedura diagonalizacji unitarnej jest prosta, gdy znamy widmo.

Zaskakujące jest to, że macierze unitarnie diagonalizowalne rozpoznać można bardzo łatwo, bez znajomości widma. Celem naszym jest podanie tyczących się tego ogólnych twierdzeń i ich konsekwencji, umożliwiających zrozumienie własności operatorów na przestrzeniach $\mathbb{R}^k$ i $\mathbb{C}^k$. Najbardziej spektakularny jest wynik końcowy: każdy taki operator jest złożeniem izometrii liniowej z operatorem „rozciągającym" lub „ściągającym" przestrzeń wzdłuż wzajemnie ortogonalnych osi.

Uwaga 3. Gdy macierz $\mathbf{A}$ jest rzeczywista, to przy oznaczeniach uwagi 2 zbiór B_λ składa się dla $\lambda \in \mathbb{R}$ wyłącznie z wektorów o wszystkich współrzędnych rzeczywistych. (Istotnie, dla $\lambda \in \mathbb{R}$ współczynniki układu $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ są rzeczywiste, więc takie są zarówno współrzędne wektorów fundamentalnego układu rozwiązań, jak i układu otrzymanego zeń przy pomocy ortogonalizacji Grama–Schmidta.) Wynika stąd

Wniosek 1. *Macierz rzeczywista, która jest unitarnie diagonalizowalna nad $\mathbb{C}$ i której wszystkie wartości własne są rzeczywiste, jest też unitarnie diagonalizowalna nad $\mathbb{R}$.*

Zadanie uzupełniające 1. * a) Dowieść, że operator unitarnie diagonalizowalny L indukuje takiż operator na każdej podprzestrzeni L–niezmienniczej.

b) Sformułować i udowodnić unitarny odpowiednik zadania uz.5 z §2.4.

Zadania ze zbioru Kostrykina: 5,6 w §II.4.4, oraz 4,7 w §II.4.3.

3. Charakteryzacja zespolonych macierzy unitarnie diagonalizowalnych.

Wszystkie macierze w tym punkcie są zespolone i kwadratowe. Macierz $\mathbf{A}$ nazywamy:

normalną, gdy $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$,

samosprzężoną lub **hermitowską**, gdy $\mathbf{A}^* = \mathbf{A}$

antysamosprzężoną lub **antyhermitowską**, gdy $\mathbf{A}^* = -\mathbf{A}$.

Zadania.

1. a) Suma $\mathbf{A}+\mathbf{B}$ macierzy normalnych jest też taką macierzą, jeśli $\mathbf{AB}^* = \mathbf{B}^*\mathbf{A}$ lub $\mathbf{AB}^* = \mathbf{A}^*\mathbf{B}$.

b) Iloczyn $\mathbf{AB}$ macierzy normalnych jest taką macierzą, jeśli $\mathbf{AB}^* = \mathbf{B}^*\mathbf{A}$.

c) Z a) i b) wynika normalność macierzy $p(\mathbf{A})$, dla każdego wielomianu $p \in \mathbb{C}[x]$ i macierzy normalnej $\mathbf{A}$.

2. a) Każda macierz kwadratowa jest sumą macierzy samosprzężonej i macierzy antysamosprzężonej – a więc sumą dwóch macierzy normalnych.

b) Macierz $\mathbf{A}$ wtedy i tylko wtedy jest samosprzężona, gdy $i\mathbf{A}$ jest macierzą antysamosprzężoną.

3. Gdy macierz $\mathbf{A}$ jest normalna (odp. samosprzężona), to $\mathbf{A}^*$, $\overline{\mathbf{A}}$, $\mathbf{A}^t$ też; tak samo $\mathbf{A}^{-1}$ (jeśli istnieje).

4. $\mathrm{diag}(\mathbf{K}_1, \mathbf{K}_2)$ wtedy i tylko wtedy jest macierzą normalną (odp. samosprzężoną), gdy takie są klatki $\mathbf{K}_1$ i $\mathbf{K}_2$.

Odnotujmy następujące dwie uwagi (pierwsza wynika wprost z przyjętych definicji):

Uwaga 1. W każdym z następujących przypadków macierz $\mathbf{A}$ jest normalna:
a) $\mathbf{A}$ jest macierzą samosprzężoną,
b) $\mathbf{A}$ jest macierzą antysamosprzężoną,
c) $\mathbf{A}$ jest macierzą unitarną;
d) $\mathbf{A}$ jest macierzą diagonalną. ⊟

Uwaga 2. Własności a), b), c) są niezmiennikami unitarnego podobieństwa: w ślad za daną macierzą, ma ją każda podobna do niej unitarnie. Normalność macierzy też jest takim niezmiennikiem.

(Dowód dla a) jest taki: jeśli $\mathbf{A} = \mathbf{A}^*$ i $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$, gdzie macierz $\mathbf{S}$ jest unitarna, to $\mathbf{S}^{-1} = \mathbf{S}^*$ i $\mathbf{B}^* = (\mathbf{S}^*\mathbf{AS})^* = \mathbf{S}^*\mathbf{A}^*(\mathbf{S}^*)^* = \mathbf{B}$, bo $(\mathbf{S}^*)^* = S$. Dowody w innych przypadkach są analogiczne.) □

Dowiedziemy teraz zaskakującej i ważnej charakteryzacji macierzy unitarnie diagonalizowalnych:

Twierdzenie 1 (O.Toeplitza)**.** *Zespolona macierz kwadratowa wtedy i tylko wtedy jest unitarnie diagonalizowalna nad $\mathbb{C}$, gdy jest normalna.*

<u>Dowód.</u> Jeśli macierz $\mathbf{A}$ jest unitarnie diagonalizowalna, to jest normalna na mocy uwag 1d) i 2. Załóżmy teraz, że macierz $\mathbf{A}$ jest normalna. Niech $\mathbf{B}$ będzie macierzą dolnie trójkątną, unitarnie podobną do $\mathbf{A}$; taka istnieje na mocy twierdzenia Schura z p.1. Jak wiemy z uwagi 2, macierz $\mathbf{B}$ jest normalna. Poniższy lemat pokazuje więc, że jest ona diagonalna, co kończy dowód twierdzenia:

Lemat 1. *Każda dolnie trójkątna macierz normalna $\mathbf{B} \in \mathcal{M}_k(\mathbb{C})$ jest diagonalna.*

<u>Dowód.</u> Stosujemy indukcję względem k. Odnotujmy, że $||\mathbf{Bv}||^2 = ||\mathbf{B}^*\mathbf{v}||^2$ dla $\mathbf{v} \in \mathbb{C}^k$, bo wobec normalności macierzy $\mathbf{B}$ obie strony są równe $\langle \mathbf{B}^*\mathbf{Bv}, \mathbf{v}\rangle$. Przyjmując $\mathbf{v} = \mathbf{e}_k$ otrzymujemy $||(0, 0, ..., 0, b_{kk})|| = ||(\overline{b_{k1}}, ..., \overline{b_{kk}})||$, skąd wszystkie wyrazy k-tego wiersza $\mathbf{B}$, poza b_{kk}, są zerowe. Z zadania 4 wynika więc, że klatka, powstała z $\mathbf{B}$ przez wykreślenie ostatniego i ostatniej kolumny, jest normalna. Z założenia indukcyjnego jest ona diagonalna, a w ślad za nią taka jest i macierz $\mathbf{B}$. □

Wniosek 1. *a) Macierz samosprzężona jest unitarnie podobna do rzeczywistej macierzy diagonalnej.*

b) Macierz unitarna jest unitarnie podobna do macierzy diagonalnej, której przekątna ma wyłącznie wyrazy o module 1;

c) Macierz antysamosprzężona jest unitarnie podobna do macierzy diagonalnej, której przekątna ma wyłącznie wyrazy leżące na osi urojonej.

Implikacje przeciwne też są prawdziwe.

<u>Dowód.</u> Ad a) Gdy macierz $\mathbf{A}$ jest samosprzężona, to jest normalna i wobec tego podobna unitarnie do pewnej macierzy diagonalnej $\mathbf{D}$. Na podstawie uwagi 2, macierz $\mathbf{D}$ też jest samosprzężona, wobec czego wyrazy jej przekątnej są rzeczywiste. Implikacja przeciwna również wynika z uwagi 2.

Ad b) i c). Dowody są analogiczne i wykorzystują to, że macierz diagonalna wtedy i tylko wtedy jest unitarna (odp. antysamosprzężona), gdy wszystkie wyrazy jej przekątnej mają moduł 1 (odp. są czysto urojone). ⊟

Wniosek 2. *a) Wartości własne macierzy samosprzężonej są rzeczywiste.*
b) Wartości własne macierzy antysamosprzężonej leżą na osi urojonej.
c) Wartości własne macierzy unitarnej leżą na okręgu $\{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

<u>Dowód.</u> Wynika to z wniosku 1, bo na przekątnej macierzy diagonalnej, podobnej do danej macierzy, występują wszystkie jej wartości własne. □

<u>Zadania uzupełniające.</u> (W zadaniach 1–3 zacząć od macierzy diagonalnych.)

1. Dla macierzy normalnych prawdziwe są implikacje przeciwne do stwierdzonych we wniosku 2.

2. Jeśli $\mathbf{v}$ jest wektorem własnym macierzy normalnej $\mathbf{A}$, odpowiadającym wartości λ, to jest też wektorem własnym macierzy $\mathbf{A}^*$, odpowiadającym wartości $\overline{\lambda}$.

3. Dowieść równoważności warunków:
   a) macierz $\mathbf{A}$ jest normalna i $\mathrm{spec}(\mathbf{A}) \subset [0, \infty)$;
   b) macierz $\mathbf{A}$ jest samosprzężona i dodatnio półokreślona (tzn. $\langle \mathbf{Av}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{Av}\rangle \geq 0$ dla wszystkich $\mathbf{v} \in \mathbb{F}^k$);
   c) macierz $\mathbf{A}$ jest unitarnie podobna do macierzy diagonalnej, której przekątna ma wyłącznie wyrazy ≥ 0;
   d) $\mathbf{A} = \mathbf{B}^2$ dla pewnej macierzy samosprzężonej $\mathbf{B}$;
   e) $\mathbf{A} = \mathbf{B}^*\mathbf{B}$ dla pewnej macierzy kwadratowej $\mathbf{B}$.

4. Dowieść, że normalność macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ jest równoważna temu, by $||\mathbf{Av}|| = ||\mathbf{A}^*\mathbf{v}||$ dla każdego wektora $\mathbf{v} \in \mathbb{C}^k$. (Wskazówka: zad. uz. 2 w p.1.)

Zadania ze zbioru Kostrykina: 16,20,21 w §II.4.2 oraz 6,21,22 w §II.4.4 (z macierzami w miejsce operatorów).

4. Postać ortogonalnie kanoniczna rzeczywistych macierzy normalnych.

Twierdzenie i wniosek z poprzedniego punktu dotyczą diagonalizacji nad $\mathbb{C}$ i nie zapewniają diagonalizowalności nad $\mathbb{R}$ rzeczywistej macierzy normalnej:

Przykład 1. Macierz obrotu $\mathbf{A} := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ jest ortogonalna (a zatem normalna) i jej wartościami własnymi są $e^{i\alpha} := \cos\alpha + i\sin\alpha$ oraz $e^{-i\alpha}$. Gdy $\sin\alpha \neq 0$ nie jest więc ona diagonalizowalna nad $\mathbb{R}$ (bo $\mathrm{spec}_{\mathbb{R}}(\mathbf{A}) = \emptyset$), choć na mocy twierdzenia Toeplitza jest unitarnie podobna (nad $\mathbb{C}$) do macierzy $\mathbf{D} = \mathrm{diag}(e^{i\alpha}, e^{-i\alpha})$.

Zamiast „przekształcenie unitarne nad $\mathbb{R}$" czy „rzeczywista macierz unitarna" zwykło się mówić **przekształcenie ortogonalne** oraz **macierz ortogonalna**, odpowiednio; będziemy tych nazw niekiedy używać. W szczególności, zamiast „macierz unitarnie diagonalizowalna nad $\mathbb{R}$" będziemy też mówić „macierz **ortogonalnie diagonalizowalna**". Podamy teraz charakteryzację takich macierzy.

Twierdzenie 1. *Kwadratowa macierz rzeczywista wtedy i tylko wtedy jest ortogonalnie diagonalizowalna, gdy jest symetryczna.*

Dowód. Wpierw odnotujmy, że gdy jedna z dwóch ortogonalnie podobnych macierzy jest symetryczna i rzeczywista, to druga też. (Wynika to z uwagi 2 w p.3, bo dla macierzy rzeczywistych samosprzężoność oznacza symetrię.) Jeśli więc macierz $\mathbf{A}$ jest ortogonalnie diagonalizowalna, to jest symetryczna (bo jest ortogonalnie podobna do macierzy diagonalnej).

Przeciwnie, niech rzeczywista macierz $\mathbf{A}$ będzie symetryczna. Jest ona wówczas samosprzężona, wobec czego jest unitarnie diagonalizowalna nad $\mathbb{C}$ i jej wartości własne są rzeczywiste. (Patrz p.3.) Pozostaje więc skorzystać z wniosku 1 w p.2. □

Niech $\mathbf{A}$ będzie rzeczywistą macierzą normalną. Gdy $\mathbf{A} \neq \mathbf{A}^t$, macierz ta nie jest ortogonalnie diagonalizowalna. Pokażemy jednak w poniższym materiale uzupełniającym, że jest ona ortogonalnie podobna do macierzy niewiele różniącej się od diagonalnej i wyznaczonej przez pierwiastki wielomianu $\chi_{\mathbf{A}}$ i ich krotności. W dowodzie wygodnie jest użyć oznaczeń z zadań rozdziału II:

$$\overline{\mathbf{v}} := (\overline{v_1}, ..., \overline{v_k}),\ \mathrm{Re}(\mathbf{v}) := \frac{1}{2}(\mathbf{v} + \overline{\mathbf{v}}),\ \mathrm{Im}(\mathbf{v}) := \frac{1}{2\mathbf{i}}(\mathbf{v} - \overline{\mathbf{v}}) \text{ dla } \mathbf{v} = (v_1, ..., v_k) \in \mathbb{C}^k.$$

Zadanie 1. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ oraz $\lambda = a + bi \in \mathbb{C}$ i $\mathbf{v} \in \mathbb{C}^k$ spełniają warunek $\mathbf{Av} = \lambda\mathbf{v}$. Wówczas $\mathbf{A}(\mathrm{Re}(\mathbf{v})) = a\mathrm{Re}(\mathbf{v}) - b\mathrm{Im}(\mathbf{v})$ i $\mathbf{A}(\mathrm{Im}(\mathbf{v})) = b\mathrm{Re}(\mathbf{v}) + a\mathrm{Im}(\mathbf{v})$.

Oznaczenie Dla $\lambda = a + b\mathbf{i} \in \mathbb{C}$ oznaczmy przez $\mathbf{K}(\lambda) \in \mathcal{M}_2(\mathbb{R})$ macierz o pierwszej kolumnie (a, b), zaś drugiej $(-b, a)$. Zauważmy, że gdy $\lambda \neq 0$, to $\frac{1}{|\lambda|}\mathbf{K}(\lambda)$ jest macierzą obrotu, wobec czego operator $\mathbb{R}^2 \to \mathbb{R}^2$ o macierzy $\mathbf{K}(\lambda)$ jest złożeniem obrotu z jednokładnością $\mathbf{v} \mapsto |\lambda|\mathbf{v}$. (Kolejność składania nie jest istotna.)

Umówmy się też, że $(x_s, ..., x_t)$ oznacza dla $t < s$ ciąg pusty.

Twierdzenie 2. * *Rzeczywista macierz normalna* $\mathbf{A}$ *jest ortogonalnie podobna do macierzy* $\mathbf{B}$ *postaci* $\mathrm{diag}(\lambda_1, ..., \lambda_s, \mathbf{K}(\lambda_{s+1}), ..., \mathbf{K}(\lambda_{s+t}))$, *dla pewnych* $\lambda_1, ..., \lambda_s \in \mathbb{R}$ *i* $\lambda_{s+1}, ..., \lambda_{s+t} \in \mathbb{C} \setminus \mathbb{R}$. *(Może się zdarzyć, że* $s = 0$ *lub* $t = 0$.*)*

Dowód. * Niech k to stopień macierzy $\mathbf{A}$, a $\lambda_1, ..., \lambda_k$ to jej wartości własne (z powtórzeniami). Z wartości tych wybierzmy wszystkie, mające nieujemną część rzeczywistą; przy tym niech $\lambda_1, ..., \lambda_s$ leżą na osi rzeczywistej, a $\lambda_{s+1}, ..., \lambda_{s+t}$ poza nią. (Być może $s = 0$ lub $t = 0$.) Korzystać będziemy z tego, że $\mathrm{V}_{\mathbf{A}}(\lambda_i) \perp V_{\mathbf{A}}(\lambda_j)$ gdy $\lambda_i \neq \lambda_j$ – co wynika z twierdzenia Toeplitza i stwierdzenia w p.2. Z tych samych rezultatów wynika też istnienie takiej ortonormalnej bazy $(\mathbf{v}_i)_{i=1}^k$ przestrzeni $\mathbb{C}^k$, że $\mathbf{Av}_i = \lambda_i\mathbf{v}_i$ dla $i = 1, ..., k$ i $\mathbf{v}_i \in \mathbb{R}^k$ dla $i = 1, ..., s$. (Patrz uwaga 3 w p.2.)

Wyróżnimy 3 części dalszego rozumowania:

i) Niech $s < i < j \leq s + t$. Wektory $\overline{\mathbf{v}_i}$ i $\mathbf{v}_j$ odpowiadają różnym wartościom własnym $\overline{\lambda_i}$ i λ_j, odpowiednio, więc są ortogonalne. Tak samo, $\overline{\mathbf{v}_j} \perp \mathbf{v}_i$ i wobec tego $\{\mathbf{v}_i, \overline{\mathbf{v}_i}\} \perp \{\mathbf{v}_j, \overline{\mathbf{v}_j}\}$. A że $\mathrm{Re}(\mathbf{v}_n), \mathrm{Im}(\mathbf{v}_n) \in \mathrm{lin}(\mathbf{v}_n, \overline{\mathbf{v}_n})$ dla $n = i, j$, to $\{\mathrm{Re}(\mathbf{v}_i), \mathrm{Im}(\mathbf{v}_i)\} \perp \{\mathrm{Re}(\mathbf{v}_j), \mathrm{Im}(\mathbf{v}_j)\}$. Podobnie też $\mathbf{v}_i \perp \overline{\mathbf{v}_i}$ oraz $\mathbf{v}_n \perp \{\mathrm{Re}(\mathbf{v}_j), \mathrm{Im}(\mathbf{v}_j)\}$ dla $n \leq s$.

iii) Z przytoczonych własności wektorów $\mathbf{v}_j$ i zadania uzupełniającego 1 w §IV.1.2 wynika, że poniższy układ jest ortonormalny:

$$\mathbf{v}_1, ..., \mathbf{v}_s, \sqrt{2}\mathrm{Re}(\mathbf{v}_{s+1}), -\sqrt{2}\mathrm{Im}(\mathbf{v}_{s+1}), ..., \sqrt{2}\mathrm{Re}(\mathbf{v}_{s+t}), -\sqrt{2}\mathrm{Im}(\mathbf{v}_{s+t}).$$

iii) Układ ten liczy $s+2t$ wyrazów – czyli tyle, ile ciąg $\lambda_1, ..., \lambda_s, \lambda_{s+1}, \overline{\lambda_{s+1}}, ..., \lambda_{s+t}, \overline{\lambda_{s+t}}$ wszystkich pierwiastków wielomianu $\chi_{\mathbf{A}}$ (każdy powtarzany zgodnie ze swą krotnością). Jest to więc układ k wektorów przestrzeni $\mathbb{R}^k$; a że jest ortonormalny, to jest jej bazą. Na podstawie zadania 1, macierzą operatora $L_{\mathbf{A}} : \mathbb{R}^k \to \mathbb{R}^k$ w tej bazie jest macierz $\mathbf{B}$ z tezy twierdzenia –co kończy dowód. $\square$

Uwaga 1. Z dowodu wynika, że liczby $\lambda_{s+1}, ..., \lambda_{s+t}$ można obrać o dodatniej części urojonej. Latwo zauważyć, w oparciu o równość $\chi_{\mathbf{B}} = \chi_{\mathbf{A}}$, że ciąg $(\lambda_1, ..., \lambda_{s+t})$ z twierdzenia 2 jest tym warunkiem wyznaczony jednoznacznie, z dokładnością do kolejności.

Z twierdzenia 2 i wniosku 2 w p.3 wynika natychmiast

Wniosek 1. *a) Macierz ortogonalna* $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ *jest ortogonalnie podobna do macierzy postaci*

$$\mathrm{diag}(\lambda_1, ..., \lambda_s, \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 \\ \sin\alpha_1 & \cos\alpha_1 \end{pmatrix}, ..., \begin{pmatrix} \cos\alpha_t & -\sin\alpha_t \\ \sin\alpha_t & \cos\alpha_t \end{pmatrix}),$$

gdzie $\lambda_j = \pm 1$ *dla* $j = 1, ..., s$ *i* $\alpha_j \in \mathbb{R}$ *dla* $j = 1, ..., t$.

*b)** *Macierz antysymetryczna* $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ *jest ortogonalnie podobna do macierzy postaci*

$$\mathrm{diag}(0, ..., 0, \begin{pmatrix} 0 & -b_1 \\ b_1 & 0 \end{pmatrix}, ..., \begin{pmatrix} 0 & -b_t \\ b_t & 0 \end{pmatrix}) \qquad (b_j \in \mathbb{R} \text{ dla } j = 1, ..., t)$$

Uwaga 2. 2×2-klatki występujące w części a) wniosku są macierzami obrotu. Ponadto, macierz $\pm\mathbf{I}_2$ też jest macierzą obrotu (o 0 lub π radianów). Stąd łącząc w pary te same wartości λ_j dla $j \leq s$ wnosimy, że każda macierz ortogonalna $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ jest ortogonalnie podobna do sumy zewnętrznej klatek, z których wszystkie poza być może jedną są 2×2–macierzami obrotu, a pozostała bądź jest 1×1–klatką, równą $\pm\mathbf{I}_1$, bądź macierzą $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (symetrii osiowej płaszczyzny). □

Ćwiczenie. a) Podać interpretację przekształcenia płaszczyzny o macierzy $\begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$.

Zadania uzupełniające. Dowieść, że:

1. Dwie podobne macierze normalne są unitarnie podobne. Tak samo, dwie podobne macierze, które są rzeczywiste i symetryczne, są ortogonalnie podobne.

2. a) Rząd antysymetrycznej macierzy rzeczywistej jest liczbą parzystą.
 b) Jeśli powyższy rząd jest równy $2m$, to rozważana macierz jest sumą m macierzy antysymetrycznych rzędu 2.

Zadania ze zbioru Kostrykina: 4 i 7 w §II.4.2.

5. Unitarna diagonalizacja operatorów.

W tym punkcie V oznacza zespoloną lub rzeczywistą przestrzeń unitarną. Przeniesiemy na operatory $L \in \mathcal{L}(V)$ twierdzenia, udowodnione wyżej dla macierzy.

Definicja. Operator $L \in \mathcal{L}(V)$ jest:
 normalny, jeśli $L^*L = LL^*$,
 samosprzężony (lub **hermitowski**), jeśli $L^* = L$,
 antysamosprzężony (lub **antyhermitowski**), jeśli $L^* = -L$.
Gdy ciałem skalarów przestrzeni V jest $\mathbb{R}$, zamiast „samosprzężony" czy „antysamosprzężony" używane są też nazwy **symetryczny** i **antysymetryczny**.

Zadanie 1. Własności powyższe są niezmiennikami unitarnego podobieństwa.

Lemat 1. *Niech $\mathcal{V}$ będzie bazą ortonormalną w V. Dla operatora $L \in \mathcal{L}(V)$ następujące warunki są równoważne:*
 a) L jest operatorem normalnym (odp. samosprzężonym, antysamosprzężonym, unitarnym)
 b) $[L]_\mathcal{V}$ jest macierzą normalną (odp. samosprzężoną, antysamosprzężoną, unitarną).

Dowód. Niech $\mathbf{A} := [L]_\mathcal{V}$. Przyporządkowanie operatorowi jego macierzy w bazie $\mathcal{V}$ jest bijekcją $\mathcal{L}(V)$ na $\mathcal{M}_k$ ($k = \dim(V)$), zachowującą sprzężenie i mnożenie, skąd $L^*L = LL^* \Leftrightarrow \mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$. Tak samo, $L = \pm L^* \Leftrightarrow \mathbf{A} = \pm\mathbf{A}^*$ i $L^*L = I \Leftrightarrow \mathbf{A}^*\mathbf{A} = \mathbf{I}$.

Twierdzenie 1. *Operator $L \in \mathcal{L}(V)$, gdzie V jest zespoloną przestrzenią unitarną, jest unitarnie diagonalizowalny wtedy i tylko wtedy, gdy jest normalny.*

Dowód. Załóżmy normalność L i niech $\mathcal{V}$ będzie dowolną bazą ortonormalną w V. Macierz $[L]_\mathcal{V}$ jest normalna, więc istnieje macierz unitarna $\mathbf{C}$ taka, że $\mathbf{D} := \mathbf{C}^{-1}[L]_\mathcal{V}\mathbf{C}$ jest macierzą diagonalną. Baza $\mathcal{W}$, określona równością $[I]_\mathcal{V}^\mathcal{W} = \mathbf{C}$, jest ortonormalna (w ślad za $\mathcal{V}$, wobec unitarności macierzy $\mathbf{C}$) i spełnia warunek $[L]_\mathcal{W} = \mathbf{D}$.

Implikacja przeciwna wynika z lematu 1, bo macierze diagonalne są normalne. □

Tak samo otrzymujemy

Twierdzenie 2. *Dla operatora $L \in \mathcal{L}(V)$ na przestrzeni euklidesowej V równoważne są warunki:*
a) L jest operatorem symetrycznym;
b) L jest operatorem ortogonalnie diagonalizowalnym.

Niech $\{V_i\}_{i=1}^s$ będzie rodziną podprzestrzeni przestrzeni unitarnej V i niech $L_i \in \mathcal{L}(V_i)$ dla każdego i. Operator $L \in \mathcal{L}(V)$ nazwiemy **sumą ortogonalną operatorów** L_i, gdy $V = V_1 \bigoplus ... \bigoplus V_s$ i $L(\mathbf{v}_1 + ... + \mathbf{v}_s) = \sum_i L(\mathbf{v}_i)$ dla $\mathbf{v}_1 \in V_1, ..., \mathbf{v}_s \in V_s$. Piszemy wtedy $L = \bigoplus_i L_i$ oraz $V = \bigoplus_i V_i$.

Twierdzenie 3 (Uogólnione twierdzenie Eulera)**.** *Izometria liniowa przestrzeni euklidesowej jest sumą ortogonalną rodziny operatorów L_i, z których wszystkie poza pierwszym są obrotami wokół* **0** *pewnych płaszczyzn (tj. podprzestrzeni 2–wymiarowych), a L_1 bądź też jest takim obrotem, bądź symetrią osiową płaszczyzny, bądź przekształceniem $\pm I_W$, dla pewnej jednowymiarowej podprzestrzeni W.*

Dowód. * Gdy $V = \mathbb{R}^k$ teza wynika z uwagi 2 w p.4. Gdy nie, rozpatrzmy mapę $S : V \to \mathbb{R}^k$, wyznaczoną przez pewną ortonormalną bazę przestrzeni V. Wówczas $L' := SLS^{-1}$ jest izometrią (jako złożenie izometrii). Jak już wiemy, L' jest więc sumą ortogonalną operatorów L'_i żądanej postaci, z odpowiadającym temu przedstawieniem przestrzeni $\mathbb{R}^k$ jako ortogonalnej sumy podprzestrzeni $V'_i \subset \mathbb{R}^k$. Gdy przyjąć $L_i := S^{-1}L'_iS$ i $V_i = S^{-1}(V'_i)$, otrzymamy szukane rozkłady $V = \bigoplus_i V_i$ oraz $L = \bigoplus_i L_i$. (Por. §1.2, zadanie 2 i 4.)

Definicja. * **Homotetią** (lub: **jednokładnością**) przestrzeni wektorowej W nazywamy przekształcenie postaci $\mathbf{w} \mapsto \lambda\mathbf{w}$; wyznaczający je skalar λ nazywamy **skalą** tej homotetii. Homotetia przestrzeni rzeczywistej jest zatem jej proporcjonalnym „kurczeniem” bądź „rozciąganiem” (zależnie od tego, czy $\lambda \in [0, 1]$, czy też $\lambda \geq 1$), połączonym być może z odbiciem względem **0** (gdy $\lambda < 0$).

Uwaga 1. * Rozumowanie z dowodu twierdzenia 3 pokazuje też, że operator symetryczny L na przestrzeni euklidesowej jest sumą ortogonalną homotetii działających na podprzestrzeniach jednowymiarowych. Warto odnotować, że skale otrzymanych homotetii są wartościami własnymi L.

Uwaga 2. * Również w przypadku zespolonym możemy, na mocy twierdzenia 1, interpretować zadany operator normalny L jako sumę ortogonalną homotetii działających na podprzestrzeniach wymiaru 1. Jednak zespoloną podprzestrzeń wymiaru 1 wyobrażać sobie należy nie jak prostą rzeczywistą, lecz podobnie jak płaszczyznę Gaussa liczb zespolonych; zaś jej homotetię $\mathbf{w} \mapsto \lambda\mathbf{w}$ – jako złożenie obrotu $\mathbf{w} \mapsto (\lambda/|\lambda|)\mathbf{w}$ z homotetią o skali $|\lambda| \geq 0$. Gdy operator L jest samosprzężony, współczynniki λ są już rzeczywiste (bo leżą w $\mathrm{spec}(L)$) i znów mamy do czynienia z „kurczeniem” lub „rozciąganiem” rozważanych podprzestrzeni, złożonym być może z odbiciem względem zera.

Ćwiczenie. Zdefiniujmy operator $L \in \mathcal{L}(\mathbb{C}^k)$ wzorem $L(\mathbf{v}) = (v_2, ..., v_k, v_1)$ dla $\mathbf{v} = (v_1, \dots, v_k) \in \mathbb{C}^k$.
a) Dowieść, że jest on unitarny, a jego widmo to zbiór pierwiastków stopnia k z jedynki.
b) Znaleźć bazę diagonalizującą ten operator i uzasadnić jej ortogonalność.

Zadanie 2. Udowodnić tw. 3 bezpośrednio, bez korzystania z wyników p.4. (Wskazówka: ponieważ $3 \notin 2\mathbb{Z}$, to L ma wektor własny $\mathbf{v}$. Rozpatrzeć operator indukowany na podprzestrzeni $\mathbf{v}^\perp$.)

Zadania uzupełniające. (Założyć wpierw, że $V = \mathbb{C}^k$ i macierz $[L]$ jest diagonalna.)

1. Gdy operator L jest normalny, to $\ker(L) = \ker(L^*)$ i $\mathrm{im}(L) = \mathrm{im}(L^*)$.
2. Niech $\mathbf{v}$ będzie wektorem własnym operatora normalnego L. Dowieść, że:
   a) $\mathbf{v}^\perp$ jest podprzestrzenią niezmienniczą dla L i dla L^*, a operator $L_{|\mathbf{v}^\perp}$ jest normalny.
   b) Jeśli operator L jest samosprzężony, to $L_{|\mathbf{v}^\perp}$ też.
3. Dowieść, że gdy L jest operatorem samosprzężonym, to:
   a) liczba $a := \sup\{\langle L(\mathbf{v}), \mathbf{v}\rangle : ||\mathbf{v}|| = 1\}$ jest jego największą wartością własną;

b) liczba $b := \inf\{\langle L(\mathbf{v}), \mathbf{v}\rangle : ||\mathbf{v}|| = 1\}$ jest jego najmniejszą wartością własną;
c) jeśli $\langle L(\mathbf{v}), \mathbf{v}\rangle = \lambda ||\mathbf{v}||^2$ i $\lambda \in \{a, b\}$, to $\mathbf{v}$ jest wektorem własnym operatora L.

Uwaga 3. Zdanie to może być użyte do (przybliżonego) numerycznego wyznaczenia wartości i wektorów własnych operatora samosprzężonego L: wpierw wyznaczamy liczby a i b oraz odpowiadające im wektory własne $\mathbf{v}_a, \mathbf{v}_b$, po czym czynność tę iterujemy, zastąpiwszy V przez $\{\mathbf{v}_a, \mathbf{v}_b\}^{\perp}$ i obciąwszy L. (Może się zdażyć, że $a = b$, lecz to nie przeszkadza.)

4. Oryginalne twierdzenie Eulera orzeka, że każda izometria liniowa przestrzeni $\mathbb{R}^3$ jest obrotem wokół prostej lub złożeniem obrotu z odbiciem zwierciadlanym względem płaszczyzny, prostopadłej do osi obrotu. Dowieść tego i uzyskać taki rozkład dla izometrii, zadanej macierzą o wierszach $(2/3, 2/3, -1/3), (2/3, -1/3, 2/3), (-1/3, 2/3, 2/3)$.

Zadania ze zbioru Kostrykina: §II.4.2: 15–17,19,21; §II.4.3: 5*,7,10,10,14,16,17; §II.4.4: 5,6,21–23.

6. * Rozkład biegunowy i związki między operatorami samosprzężonymi a unitarnymi.

Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ będzie macierzą unitarnie diagonalizowalną, zaś $f : \mathrm{spec}(\mathbf{A}) \to \mathbb{C}$ pewną funkcją. Ponieważ we wzorach (2) i (5) w §2.3 można za $\mathbf{S}$ obrać macierz unitarną, więc macierz $f(\mathbf{A})$ jest unitarnie diagonalizowalna, zaś jej widmem jest $f(\mathrm{spec}(\mathbf{A}))$. Jeśli więc $\mathbb{F} = \mathbb{C}$ i $f(\mathrm{spec}(\mathbf{A})) \subset \mathbb{R}$, to $f(\mathbf{A})$ możemy rozpoznać jako macierz samosprzężoną, jeśli $f(\mathrm{spec}(\mathbf{A})) \subset [0, \infty)$ –jako nieujemnie określoną, jeśli zaś $f(\mathrm{spec}(\mathbf{A})) \subset \mathbb{R}i$ –jako antysamosprzężoną; wreszcie jeśli $f(\mathrm{spec}(\mathbf{A}))$ składa się wyłącznie z liczb zespolonych o module 1, to $f(\mathbf{A})$ jest macierzą unitarną. (Wykorzystujemy wniosek 1 w p.3.) Daje to sposób konstruowania interesujących przekształceń pomiędzy wymienionymi zbiorami macierzy (bądź odpowiadających im operatorów) i ustanowienia sugestywnej analogii pomiędzy własnościami macierzy samosprzężonych a liczb rzeczywistych, macierzy nieujemnie określonych a liczb nieujemnych, macierzy unitarnych a liczb zespolonych o module 1.

Przykład 1. * Niech $f : [0, \infty) \to [0, \infty)$ będzie pewną bijekcją, niech $g = f^{-1}$ i niech $\mathcal{H}_+$ będzie zbiorem nieujemnie określonych macierzy hermitowskich. Ponieważ $\mathrm{spec}(\mathbf{A}) \subset [0, \infty)$ dla $\mathbf{A} \in \mathcal{H}_+$ (patrz zadanie uzupełniające 3 w p.3), więc funkcje $\mathbf{A} \mapsto f(\mathbf{A})$ i $\mathbf{A} \mapsto g(\mathbf{A})$ są dobrze określone na $\mathcal{H}_+$, a także przyjmują wartości w $\mathcal{H}_+$. Ponadto, jako przekształcenia $\mathcal{H}_+ \to \mathcal{H}_+$ są one wzajemnie odwrotne: $f(g(\mathbf{A})) = \mathbf{A} = g(f(\mathbf{A}))$ dla $\mathbf{A} \in \mathcal{H}_+$, na podstawie zadania 1 e) w §2.3. W szczególności, dla każdej macierzy $\mathbf{B} \in \mathcal{H}_+$ istnieje dokładnie jedna macierz $\mathbf{A} \in \mathcal{H}_+$ taka, że $f(\mathbf{A}) = \mathbf{B}$.

Przyjmując $f(t) = t^n$ stwierdzamy np., że każda macierz $\mathbf{A} \in \mathcal{H}_+$ posiada dokładnie jeden należący do $\mathcal{H}_+$ pierwiastek zadanego stopnia n. $\Box$

Przykład 2. * Oznaczmy przez f obcięcie do $[0, 2\pi)$ przekształcenia $t \mapsto e^{it}$, traktowane jako przekształcenie w okrąg jednostkowy $S^1 \subset \mathbb{C}$, a przez $g : S^1 \to [0, 2\pi)$ przekształcenie odwrotne do f. Jak wyżej, wzory $\mathbf{A} \mapsto \exp(\mathbf{A})$ oraz $\mathbf{A} \mapsto g(\mathbf{A})$ zadają wzajemnie odwrotne odpowiedniości pomiędzy macierzami unitarnymi a samosprzężonymi o widmie w $[0, 2\pi)$. W szczególności, każda macierz unitarna jest postaci $e^{i\mathbf{A}}$ dla pewnej samosprzężonej macierzy nieujemnie określonej $\mathbf{A}$. $\Box$

Przykład 3. * Homografia $z \mapsto (1 - z)/(1 + z)$ jest swą odwrotnością i przeprowadza oś urojoną na $S^1 \setminus \{-1\}$. (Patrz ....) Wynika stąd, że $\mathbf{A} \mapsto (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$ przekształca zbiór macierzy antysamosprzężonych na zbiór macierzy unitarnych o widmie nie zawierającym -1; przekształcenie odwrotne zadane jest tym samym wzorem.

Zauważmy, że gdy macierz $\mathbf{A}$ ma wyrazy rzeczywiste, to macierz $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$ też. Przekształcenie nasze ustala zatem zarazem $1 - 1$ odpowiedniość pomiędzy rzeczywistymi macierzami antysymetrycznymi a macierzami ortogonalnymi, których widmo nie zawiera -1. (Można tego też dowieść bezpośrednio, por. zad. uz. 5 w §II.5.2, przy $\mathbf{A} = \mathbf{I}$.)

Ponieważ każdą $k \times k$–macierz antysymetryczną można wyrazić przy pomocy $k(k-1)/2$ parametrów rzeczywistych (wyrazów nad przekątną), więc korzystając z powyższych tzw. **przekształceń Cayley'a** można przy pomocy tyluż parametrów przedstawić i każdą macierz ortogonalną stopnia k, której widmo nie zawiera -1. $\square$

Do sformułowania zasadniczego rezultatu tego punktu w wersji możliwie ogólnej (obejmującej operatory, nie tylko zaś macierze) potrzebny będzie następujący

Lemat 1. * *Dla operatora samosprzężonego L w przestrzeni unitarnej V równoważne są warunki*
a) $\langle L(\mathbf{v}), \mathbf{v}\rangle \geq 0$ *dla wszystkich* $\mathbf{v} \in V$*;*
b) $\mathrm{spec}(L) \subset [0, \infty)$.

Dowód. Latwo widzieć, że lemat jest prawdziwy gdy $V = \mathbb{F}^k$ i macierz operatora L (w standardowej bazie) jest diagonalna. W przypadku ogólnym pozostaje wykorzystać unitarne podobieństwo L do operatora o tych własnościach. $\boxminus$

Operator samosprzężony spełniający powyższe równoważne warunki nazywamy **nieujemnie określonym** lub **dodatnio półokreślonym**. Z przykładu 1 wynika łatwo, że operator taki posiada jedyny nieujemnie określony pierwiastek.

Twierdzenie 1. * *Każdą macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, można przedstawić w postaci $\mathbf{A} = \mathbf{UH}$, gdzie macierz $\mathbf{U}$ jest unitarna, a $\mathbf{H}$– samosprzężona i nieujemnie określona. (Obie są nad $\mathbb{F}$).*

Równoważnie: Gdy $L \in \mathcal{L}(V)$, gdzie V jest przestrzenią unitarną, to istnieją operatory $U, H \in \mathcal{L}(V)$ takie, że $L = UH$ i operator $H = H^*$ jest nieujemnie określony, a U – unitarny (tzn. jest izometrią liniową).

Dodatek: *W przedstawieniu $L = UH$ operator H jest przez L wyznaczony jednoznacznie. Gdy L jest izomorfizmem, to i operator U jest wyznaczony jdnoznacznie.*

Uwaga 1. * Powyższe przedstawienie $L = UH$ nazywane jest (lewym) **rozkładem biegunowym** operatora L; nazwa bierze się od biegunowgo przedstawienia liczby zespolonej jako iloczynu liczby o module 1 i liczby nieujemnej. Wraz z rezultatami z poprzedniego punktu, twierdzenie daje przejrzysty opis operatora działającego na przestrzeni euklidesowej: jest on złożeniem operatora H, polegającego na „rozciąganiu" przestrzeni we wzajemnie ortogonalnych kierunkach, z izometrią liniową U, której działanie opisano w p.5 (w twierdzeniu 3 i uwadze 2.)

Stosując twierdzenie do operatora L^* i sprzęgając wynik, otrzymujemy prawy rozkład biegunowy: $L = H_1 U_1$, gdzie nadal U_1 jest operatorem unitarnym, zaś $H_1 = H_1^*$ nieujemnie określonym. $\square$

Dowód twierdzenia. Jeśli żądane przedstawienie istnieje, to $L^*L = H^*(U^*U)H = H^*H = H^2$. Definiujemy zatem H jako nieujemnie określony pierwiastek z L^*L. Pierwiastek taki, i to jedyny, istnieje, bo $(L^*L)^* = L^*L$ i $\langle L^*L(\mathbf{v}), \mathbf{v}\rangle = \langle L(\mathbf{v}), L(\mathbf{v})\rangle \geq 0$ dla $\mathbf{v} \in V$. (Patrz przykład 1 i lemat 1.)

By zdefiniować U zauważmy, że $||L(\mathbf{v})|| = ||H(\mathbf{v})||$ dla $\mathbf{v} \in H$, bo

$$||L(\mathbf{v})||^2 = \langle L(\mathbf{v}), L(\mathbf{v})\rangle = \langle L^*L(\mathbf{v}), \mathbf{v}\rangle = \langle H^*H(\mathbf{v}), \mathbf{v}\rangle = ||H(\mathbf{v})||^2 \tag{8}$$

Gdy więc $\ker(L) = \{\mathbf{0}\}$, to $\ker(H) = \{\mathbf{0}\}$ i operator $U := LH^{-1}$ jest unitarny. Dowód twierdzenia, wraz z jednoznaczością H i L, jest wtedy zakończony.

W ogólnym przypadku dla $\mathbf{w} \in H(V)$ przyjmijmy $U_0(\mathbf{w}) := L(\mathbf{v})$, gdzie $\mathbf{v} \in H^{-1}(\mathbf{w})$. Takich wektorów $\mathbf{v}$ może być wiele, lecz definicja jest poprawna: gdy $H(\mathbf{v}_1) = H(\mathbf{v}_2) = \mathbf{w}$, to $||H(\mathbf{v}_1 - \mathbf{v}_2)|| = 0$, skąd $||L(\mathbf{v}_1 - \mathbf{v}_2)|| = 0$ wobec równości (8), zastosowanej do $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$.

Definicja przekształcenia $U_0 : H(V) \to V$ jest więc poprawna i wynika z niej wprost, że jest ono liniowe i spełnia warunek $||U_0(\mathbf{v})|| = ||\mathbf{v}||$ dla $\mathbf{v} \in H(V)$. (Por. (8).) Obieramy za $U : V \to V$

izometrię liniową taką, że $U(\mathbf{v}) = U_0(\mathbf{v})$ dla $\mathbf{v} \in H(V)$; patrz zadanie 4 w §IV.3.1. Z przyjętych określeń, $L = UH$ jest żądanym rozkładem. □

Zadanie uzupełniające 1. a) W oparciu o twierdzenie 1 dowieść następującego **twierdzenia Krasnosielskiego**: operator liniowy $L \in \mathcal{L}(\mathbb{R}^k)$ wtedy i tylko wtedy jest unitarny, gdy dla pewnego $n < k$ zachowuje n–wymiarową miarę w $\mathbb{R}^k$ (tzn. $\mu_n(R) = \mu_n(L(R))$ dla każdego równoległościanu rozpiętego na n wektorach).

b) Czy unitarny jest każdy operator $L \in \mathcal{L}(\mathbb{R}^k)$, zachowujący k–wymiarową miarę?

Zadanie uzupełniające 2. Udowodnić, że gdy dwie samosprzężone macierze są podobne, to są unitarnie podobne.

Zadania ze zbioru Kostrykina: 14, 15*, 17 w §II.4.3.

§ 4. * Próba podsumowania: skąd „teoria spektralna" w tytule rozdziału?

Przy okazji badania macierzy operatorów ustaliliśmy, że niejednokrotnie omal pełna informacja o operatorze $L \in \mathcal{L}(V)$ przekazywana jest przez własności jego widma $\mathrm{spec}(L)$. Ma bowiem miejsce

Wniosek 1. *a) Z dokładnością do podobieństwa, operator diagonalizowalny jest wyznaczony przez swe wartości własne i ich krotności jako pierwiastków wielomianu charakterystycznego.*

b) Z dokładnością do podobieństwa, operator $L \in \mathcal{L}(V)$ o całkowicie rozkładalnym wielomianie charakterystycznym jest wyznaczony przez skończenie wiele liczb $\mathrm{rk}(L - \lambda I)^n$, gdzie $1 \leq n \leq \dim(V)$ i $\lambda \in \mathrm{spec}(L)$.

c) Z dokładnością do podobieństwa unitarnego, operator normalny L na zespolonej lub rzeczywistej przestrzeni unitarnej jest wyznaczony przez pierwiastki wielomianu charakterystycznego χ_L i ich krotności.

Dowód a) wynika bezpośrednio z uwagi 6 w §2.1 i przykładu 1 w §1.3, dowód b) – z wniosku 3 w §2.2 wraz z wnioskiem 2 w §1.3, a dowód c) – z uwagi 3 w §1.3 wraz z twierdzeniem Toeplitza z §3.3 (gdy $\mathbb{F} = \mathbb{C}$) i twierdzeniem 2 w §3.4 (gdy $F = \mathbb{R}$; trzeba to twierdzenie przeformułować jak tw. 3 w §3.5). Przytoczone twierdzenia opisują nawet dokładnie postać macierzy „kanonicznego" operatora $\mathbb{F}^k \to \mathbb{F}^k$, podobnego do L; postać ta zależy tylko od wymienionych we wniosku niezmienników. Niezmienniki takie nazywamy spektralnymi, gdyż zależą od własności widma. Teoria spektralna dotyczy wyrażania własności operatora przez jego niezmienniki spektralne. Badając macierze operatorów zaawansowaliśmy ją znacznie i większość rezultatów tego rozdziału do niej należy.

Dla przykładu, „uogólnione twierdzenie Eulera" z §3.5 pozornie dotyczy własności ortogonalnych przekształceń przestrzeni euklidesowej. Zauważmy jednak, że kąty obrotów $\alpha_2, ..., \alpha_p$ i postać operatora L_1 są (pomijając kolejność kątów) wyznaczone przez wartości własne operatora L i przez ich krotności. Z dokładnością do podobieństwa ortogonalnego, twierdzenie to opisuje więc jednoznacznie operator ortogonalny L poprzez jego niezmienniki spektralne. Podobnie jest z opisem operatorów symetrycznych w twierdzeniu 2 z §3.5.

Z tych względów wyniki §3.5, wzbogacone o powyższe uwagi, bywają nazywne „twierdzeniami spektralnymi" (odpowiednio: dla operatorów ortogonalnych i operatorów symetrycznych na przestrzeni euklidesowej, a także dla n.p. operatorów normalnych czy hermitowskich na zespolonej przestrzeni unitarnej).

Zasadniczego znaczenia teoria spektralna nabiera, gdy badamy operatory na przestrzeniach nieskończonego wymiaru. Wykorzystanie macierzy, nawet nieskończonych, okazuje się wtedy niewystarczające.

§ 5. Przewodnik.

Wyniki tego rozdziału są ważne i powinny być dobrze przyswojone. Mogę wynotować pewne wybijające się hasła:

1. (§1.1–1.4, lecz ważne przykłady też w uwadze 2 w §3.3 i zadaniu 1 w §3.5.)

Podobieństwo macierzy i przekształceń (unitarne i „zwykłe"). Przykłady własności, niezmienniczych względem podobieństwa zwykłego wzgl. unitarnego (macierzy czy przekształceń). Podobieństwo a zapis operatora w bazie. Wielomian charakterystyczny macierzy i operatora; poprawność definicji.

2. (§2.1, włącznie z końcowymi zadaniami 1 i 2.)

Wartości własne. Podprzestrzenie własne i ich niezależność. Warunki konieczne i dostateczne diagonalizowalności macierzy wzgl. operatorów, wyrażone w terminach przestrzeni własnych czy wektorów własnych.

3. (§2.2-2.4)

Sformułowanie twierdzenia Jordana (wersja dla macierzy i dla operatorów; równoważność tych wersji). Sposób wyznaczania postaci Jordana i jej jednoznaczność. Zastosowanie diagonalizowalności lub postaci Jordana do wyznaczania wartości wielomianów i bardziej ogólnych funkcji na danej macierzy. Podprzestrzenie niezmiennicze i ich proste własności. Umiejętność znajdywania bazy Jordana dla małych macierzy metodami „ad hoc" lub opisanymi w §2.6.

4. (§3.1–3.4; fragmenty §3.5)

Twierdzenie Schura z §3.1 (wersja dla macierzy i dla operatorów). Diagonalizacja unitarna i ortogonalność podprzestrzeni własnych operatorów/macierzy unitarnie diagonalizowalnych. Charakteryzacja zespolonych macierzy unitarnie diagonalizowalnych (jako normalnych) i rzeczywistych macierzy ortogonalnie diagonalizowalnych (jako symetrycznych). Równoważne sformułowania dla operatorów.

Uwaga 1. Pewne fragmenty dowodów były relegowane do „zadań" (nie chodzi o uzupełniające czy z gwiazdką). Należy je rozwiązywać, by mieć sprawdzian opanowania minimum umiejętności teoretycznych. (Część była zresztą rozwiązana na wykładzie lub ćwiczeniach.)