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 $L^h = L$), 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 wprowadzić i badać pewne funkcje macierzy, co omawiamy zarówno w §2.3, jak i w §4.2. Równie ważne jest to, że umożliwiają one wniknięcie we własności rozważanego operatora. Komentarz na ten temat znaleźć można w §3.4 i w zamykającej rozdział „próbie podsumowania".

Umowy. Rozważamy tu tylko przestrzenie wektorowe skończonego wymiaru i oznaczamy je literami U, V, W, E (być może z indeksami). Gdy mowa o przekształceniu liniowym z V do W zakładamy, że V i W są nad tym samym ciałem skalarów. W odniesieniu do przestrzeni z iloczynem skalarnym (rzeczywistych czy zespolonych) używamy wymiennie nazw „izometria liniowa" i „przekształcenie unitarne" czy „izomorfizm unitarny". 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 $L \in \mathcal{L}(V)$ nazywane są **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) Endomorfizmy $K \in \mathcal{L}(V)$, $L \in \mathcal{L}(W)$ 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 $\mathbf{B} = \mathbf{SAS}^{-1}$ dla pewnej nieosobliwej macierzy $\mathbf{S} \in \mathcal{M}_k(\mathbb{F})$.

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 z iloczynem skalarnym. 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 $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{F})$ są unitarnie podobne nad $\mathbb{F}$, gdy $\mathbf{B} = \mathbf{SAS}^{-1}$ dla pewnej unitarnej macierzy $\mathbf{S} \in \mathcal{M}_k(\mathbb{F})$.

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

Zadanie 2. a) 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 $\mathbf{A}^h$ i $\mathbf{B}^h$ (gdy $\mathbb{F} = \mathbb{C}$) oraz $p(\mathbf{A})$ i $p(\mathbf{B})$, dla $p \in \mathbb{F}[x]$.

b) Jeśli podobne są macierze $\mathbf{A}$ i $\mathbf{A}'$, a także $\mathbf{B}$ i $\mathbf{B}'$, to podobne są $\operatorname{diag}(\mathbf{A}, \mathbf{B})$ i $\operatorname{diag}(\mathbf{A}', \mathbf{B}')$.

c) Obie części pozostają słuszne dla unitarnego podobieństwa.

Zadania uzupełniające.

1. Każda 2×2-macierz $\mathbf{A}$ jest podobna do macierzy $\mathbf{A}^t$. (Tak samo większa macierz, lecz dowód jest trudny; por. dalej wniosek 1 w §4.1.)

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. 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) $\det(\mathbf{A}) = \det(\mathbf{B})$, $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{B})$ i $\operatorname{rk}(\mathbf{A}) = \operatorname{rk}(\mathbf{B})$.

Skrótowo, stwierdzenia od te wyrazić możemy mówiąc, że rozważane w nich własności $\mathbf{A}^n = \mathbf{0}$, $\mathbf{A}^n = \mathbf{I}$, $\det(\mathbf{A}) = \lambda$, $\mathrm{tr}(\mathbf{A}) = \lambda$ oraz $\mathrm{rk}(\mathbf{A}) = n$ 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. Niech V będzie przestrzenią z iloczynem skalarnym, a operatory $K, L \in \mathcal{L}(V)$ niech będą unitarnie podobne. Dowieść, że:

a) $\langle K(\mathbf{v}), \mathbf{v}\rangle \geq 0 \ \forall \mathbf{v} \in V$, to $\langle L(\mathbf{v}), \mathbf{v}\rangle \geq 0 \ \forall \mathbf{v} \in V$.

b) Jeśli K jest izometrią (odp. zanurzeniem izometrycznym), to L też.

c) Jeśli $V = \mathbb{R}^n$, gdzie $n = 2, 3$, i K jest obrotem, to L też nim jest.

4. a) Jeśli jeden z operatorów podobnych P, Q jest rzutem, to drugi też. Ściślej, gdy $P \in \mathcal{L}(V)$ jest rzutem przestrzeni V na V_0 wzdłuż V_1, a $S : V \to W$ jest izomorfizmem takim, że $Q = SPS^{-1}$, to Q jest rzutem przestrzeni W na $S(V_0)$ wzdłuż $S(V_1)$.

b) Jeśli jeden z unitarnie podobnych operatorów P, Q działających w przestrzeniach z iloczynem skalarnym jest rzutem ortogonalnym, to drugi też.

c) Sformułować i udowodnić analogiczne stwierdzenia dla symetrii względem podprzestrzeni.

Tak więc o własnościach jednego z dwóch operatorów podobnych wiele można powiedzieć, gdy znamy własności drugiego, i tak samo jest z macierzy. Niestety, pytanie, czy zadane dwie macierze (równoważnie: dwa operatory) są podobne jest na ogół trudne. Dla $\mathbb{F} = \mathbb{C}, \mathbb{R}$, odpowiedź na nie umożliwią dopiero wyniki z §4.

Zadania uzupełniające. Udowodnić, że

1. Gdy przestrzeń $\mathcal{M}_k(\mathbb{C})$ wyposażyć w normę wyznaczoną przez iloczyn skalarny $\langle \mathbf{X}, \mathbf{Y}\rangle = \mathrm{tr}(\mathbf{X}\mathbf{Y}^h)$, to $\|\mathbf{A}\| = \|\mathbf{B}\|$ dla unitarnie podobnych macierzy $\mathbf{A}, \mathbf{B} \in \mathbf{M}_k(\mathbb{C})$.

2. Rzuty liniowe $P, Q \in \mathcal{L}(V)$ są podobne wtedy i tylko wtedy, gdy $\mathrm{rk}(P) = \mathrm{rk}(Q)$.

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 dla macierzy kwadratowej* $\mathbf{B}$ *równoważne są warunki:*
a) $\mathbf{B}$ *jest macierzą operatora* L *w pewnej bazie 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). $\boxminus$

Uwaga 1. 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 $\mathbf{A}$. (Podobieństwo jest unitarne dla unitarnych przestrzeni V i ortonormalnych baz $\mathcal{V}$.) $\square$

Wniosek 1. *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. Możemy zakładać, że V i W są wspólnego wymiaru k. (Inaczej ani operatory, ani macierze nie są podobne.) Z uwagi i przechodniości podobieństwa wynika, że operatory K i L wtedy i tylko wtedy są podobne, gdy $L_{\mathbf{A}}S = SL_{\mathbf{B}}$ dla pewnego izomorfizmu $S \in \mathcal{L}(\mathbb{F}^k)$. Pozostaje więc wykorzystać to, że istnienie S jest równoważne istnieniu nieosobliwej macierzy $\mathbf{T} \in \mathcal{M}_k$ takiej, że $\mathbf{AT} = \mathbf{TB}$. $\boxminus$

Uwaga 2. Wniosek ten i stwierdzenie 1 pozostają prawdziwe, z dowodami, gdy zastąpić w nich słowa „baza" przez „baza ortonormalna", a „podobieństwo" przez „podobieństwo unitarne".

Sformułujmy też wersję stwierdzenia 1 dla $V = W = \mathbb{F}^k$ oraz jej konsekwencję:

Wniosek 2. *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) $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ *dla macierzy* $\mathbf{S}$, *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 $[I]_{\mathcal{E}}^{\mathcal{V}} = \mathbf{S}$, $[L_{\mathbf{A}}]_{\mathcal{E}} = \mathbf{A}$. $\square$

Uwaga 3. Znalezienie macierzy danej postaci, podobnej do macierzy $\mathbf{A}$, 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 bardzo użyteczny.

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, jest tak, gdy $\mathbf{B}$ powstaje przez poddanie tej samej permutacji kolumn macierzy $\mathbf{A}$, a następnie wierszy otrzymanej macierzy. (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 w innej kolejności, zaś w obu poza klatkami są zera. (Gdy $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, podobieństwo jest unitarne.) $\square$

Przykład 2. Biorąc macierz operatora $L_\mathbf{A}$ w bazie $(\mathbf{e}_k, \mathbf{e}_{k-1}, ..., \mathbf{e}_1)$ stwierdzamy, że macierz $\mathbf{A}$ jest podobna do macierzy $\mathbf{B}$ otrzymanej z $\mathbf{A}$ przez „odbicie względem środka” (tzn. $b_{ij} = a_{k-j+1,k-i+1}$). 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 $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$). $\square$

Zadanie uzupełniające 1. Niech $\mathbf{E}_{ij}$ oznacza $k \times k$–macierz, której ij–ty wyraz jest równy 1, a pozostałe 0. Dowieść, że:

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

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

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 $\mathcal{M}_k(\mathbb{F})$, której przeciwdziedzinę oznaczmy przez Y. Zapytajmy: czy z f można związać pewną funkcję $\widetilde{f} : \mathcal{L}(V) \to Y$, w sposób niezależny od przypadkowych wyborów (np. wyboru bazy w V)? Okazuje się, że można, 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 (wniosek 1 w p.3).

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

Funkcję $\widetilde{f}$ na ogół nadal będziemy oznaczać przez f.

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.

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

własność operatora $L \in \mathcal{L}(V)$: $([L]_{\mathcal{V}})^{10} = \mathbf{0}$, 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 zadanie 1c) w p.2), więc wzór

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

określa pewną funkcję $\mathcal{L}(V) \to \mathbb{Z}_{\geq 0}$. Z twierdzenia 1 w §III.5.3 wynika, że jest ona identyczna z funkcją rzędu operatora, zdefiniowaną w §III.4.1 wzorem $L \mapsto \dim(L(V))$.

c) Funkcja $\mathbf{A} \mapsto \mathrm{tr}\mathbf{A}$ w ten sam sposób wyznacza funkcję **śladu endomorfizmu liniowego**, oznaczaną nadal przez tr.

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 a) i b) 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 zawsze. Niekiedy funkcję czy własność operatora umiemy definiować tylko przy pomocy macierzy tych operatorów; niekiedy zaś podanie „bezmacierzowej” interpretacji wymaga sporej pomysłowości. Np. funkcji wyznacznika macierzy odpowiada funkcja wyznacznika endomorfizmu liniowego, której interpretację – wiążącą ją z objętością – umieliśmy podać tylko przy $\mathbb{F} = \mathbb{R}$ (patrz §V.4.3); dla funkcji śladu zaś interpretacji takiej jeszcze nie znamy.

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

<u>Definicja.</u> 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 (oznaczamy niżej $\det(\mathbf{X})$ przez $|\mathbf{X}|$):

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

(Wykorzystano dwukrotnie twierdzenie Cauchy'ego o wyznaczniku iloczynu macierzy.) Możnaby uznać dowód za zakończony, gdyby nie następujący szkopuł. Własności wyznaczników, w tym twierdzenie Cauchy'ego, uzasadniono w semestrze I dla macierzy o wyrazach w ciele, a macierz $\mathbf{A} - x\mathbf{I}$ ma wyrazy będące wielomianami. Można jednak dowieść, że bez zmiany działań dodawania i mnożenia, $\mathbb{F}[x]$ zanurza się jako podzbiór pewnego (nieskończonego, siłą rzeczy) ciała $\mathbb{K}$; wolno nam więc twierdzenie Cauchyego zastosować, traktując $\mathbf{S}, \mathbf{S}^{-1}$ i $\mathbf{A} - x\mathbf{I}$ jako macierze o wyrazach w $\mathbb{K}$. (Patrz też §IV.3.4.) Wymagane twierdzenie o zanurzaniu udowodnione będzie na wykładzie Algebry. Odnotujmy, że i bez niego jak w (*) uzyskujemy $\chi_{\mathbf{A}}(\lambda) = \chi_{\mathbf{B}}(\lambda)$ dla wszystkich $\lambda \in \mathbb{F}$ – co w przypadku najważniejszych tu ciał nieskończonych powoduje już, że $\chi_{\mathbf{A}} = \chi_{\mathbf{B}}$. (Patrz wniosek 3 w §I.2.2.) □

Uwaga 1. Odwrócenie twierdzenia 1 nie jest prawdziwe: macierz o wierszach $(0, 1)$ i $(0, 0)$ nie jest podobna do zerowej, a ma wspólny z nią wielomian charakterystyczny.

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

Zadania uzupełniające.

1. a) Dowieść, że ślad rzutu liniowego jest równy rzędowi tego rzutu.

b) Wyznaczyć ślad, wyznacznik i wielomian charakterystyczny symetrii liniowej względem podprzestrzeni U, wzdłuż W.

c)* W oparciu o a) dowieść, że gdy rzuty liniowe $P_i \in \mathcal{L}(V), i = 1, ..., l$, spełniają warunek $\sum_i P_i = I$, to suma $\sum_i P_i(V)$ jest prosta i $P_iP_j = 0$ dla $i \neq j$.

2. Zbadać ślad operatora $V \ni \mathbf{v} \mapsto \varphi(\mathbf{v})\mathbf{w} \in V$, gdzie wektor $\mathbf{w} \in V$ i funkcjonał $\varphi \in V^*$ są dane.

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 $\mathbf{BA} = \mathbf{A}^{-1}(\mathbf{AB})\mathbf{A}$. Pozostaje więc 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 $p := \det(\mathbf{A} + t\mathbf{I})$ są wielomianami zmiennej t, przy czym $\deg(p) > 0$. Stąd $p(t) \neq 0$ dla nieskończenie wielu $t \in \mathbb{F}$; dla tych 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 $t \in \mathbb{F}$. (Dwukrotnie skorzystaliśmy 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 $\mathbf{A}$ i $\mathbf{B}$ jako elementy $\mathcal{M}_k(\mathbb{K})$ znajdujemy się w przypadku ii), rozstrzygniętym wyżej. $\square$

Ćwiczenie. Podać przykład takich macierzy $\mathbf{A}, \mathbf{B} \in \mathcal{M}_2$, że $\mathbf{AB}$ i $\mathbf{BA}$ nie są podobne.

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

$$\chi_{\mathbf{A}} = \sum_{i=0}^{k} c_i(\mathbf{A})x^i. \qquad (*)$$

a) $c_i(\mathbf{A}) = c_i(\mathbf{B})$ gdy $\mathbf{A}$ i $\mathbf{B}$ są podobne. (c_i są więc niezmiennikami podobieństwa).

b) $c_0(\mathbf{A}) = \det(\mathbf{A})$, $c_k(\mathbf{A}) = (-1)^k$ 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 $\sum_i \lambda_i = \mathrm{tr}(\mathbf{A})$, $\prod_i \lambda_i = \det(\mathbf{A})$ oraz, ogólniej,

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

Dalszym własnościom wielomianu charakterystycznego poświęcone są zad. uz. 1 w §2.4 i poniższe.

Zadania uzupełniające. (Patrz też zadanie uzupełniające 1 w §2.4.)

1. Niech operator $L : M_{l,k} \to \mathcal{M}_{l,k}$ będzie zadany wzorem $L(\mathbf{X}) = \mathbf{AX}$, gdzie macierz $\mathbf{A} \in \mathcal{M}_l$ jest ustalona. Znaleźć zależność między χ_L i $\chi_{\mathbf{A}}$.

2. a) Wyrazić wielomian $\chi_{\mathbf{A}^{-1}}$ przez współczynniki wielomianu $\chi_{\mathbf{A}}$, gdy $\det(\mathbf{A}) \neq 0$.
   b) Dowieść, że gdy macierz $\mathbf{A}^{-1}$ jest podobna do $\mathbf{A}$ lub do $\mathbf{A}^t$, to $\det(\mathbf{A}) = \pm 1$ i współczynniki $c_i(\mathbf{A})$ z (*) spełniają zależności $c_{k-i}(\mathbf{A}) = \varepsilon c_i(\mathbf{A})$ dla $i = 0, ..., k$, gdzie $\varepsilon = (-1)^k \det(\mathbf{A}) = \pm 1$. Ponadto, każdy skalar λ ma wtedy tę samą, co $1/\lambda$, krotność jako pierwiastek wielomianu $\chi_{\mathbf{A}}$.
   c) Uzyskać podobną zależność, gdy $\mathbf{A}$ jest macierzą unitarną. Gdy $\mathbf{A}$ jest ponadto stopnia 2 lub 3 i ma dodatni wyznacznik, wyrazić $\chi_{\mathbf{A}}$ przez $\mathrm{tr}(\mathbf{A})$.
3. Niech $L \in \mathcal{L}(V)$, gdzie V jest przestrzenią zespoloną wymiaru k. Oznaczmy przez $V_{\mathbb{R}}$ tę samą przestrzeń, ale z ciałem skalarów ograniczonym do $\mathbb{R}$, zaś przez $L_{\mathbb{R}} \in \mathcal{L}(V_{\mathbb{R}})$ – operator, wyznaczony przez L. Jaki jest związek między χ_L a $\chi_{L_{\mathbb{R}}}$? (Wskazówka: zad. uz. 3b) w §V.4.4.)
4. Niech $K \in \mathcal{L}(V, W)$, $L \in (W, V)$ i $S \in \mathcal{L}(V)$, gdzie $\dim(V) = k$ i $\dim(W) = l$.
   a) Przy $\widetilde{S} \in \mathcal{L}(V \times W)$ określonym jako $\widetilde{S} = S \times \mathbf{0}_W$, wyrazić $\chi_{\widetilde{S}}$ przez χ_S.
   b) Udowodnić, że $(-x)^k \cdot \chi_{KL} = (-x)^l \cdot \chi_{LK}$.
   c) Udowodnić **tożsamość Sylvestera**: $\det(I_W + KL) = \det(I_V + LK)$.
5. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ i $\chi_{\mathbf{A}}(\lambda) = 0$. Dowieść, że (por. zad. uz. 3 w §II.5.2.):
   a) λ leży w jednym z tzw. **kół Greshgorina** $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$.
   b) $|\lambda| \leq R := \max_i \sum_j |a_{ij}|$, wobec czego $|\det(\mathbf{A})| \leq R^k$.

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

Dowód. Niech $\mathbf{S} = \mathbf{P} + i\mathbf{Q}$, gdzie $\mathbf{P}, \mathbf{Q} \in \mathcal{M}_k(\mathbb{R})$, będzie macierzą nieosobliwą dla której $\mathbf{SA} = \mathbf{BS}$. Wtedy $\mathbf{PA} = \mathbf{BP}$ i $\mathbf{QB} = \mathbf{BQ}$, skąd $(\mathbf{P} + t\mathbf{Q})\mathbf{A} = \mathbf{B}(\mathbf{P} + t\mathbf{Q})$ 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 $\mathcal{L}(V)$ ($\dim V < \infty$) wynika ze zmiany bazy. Czyli: gdy bijekcja $\Psi : \mathcal{L}(V) \to \mathcal{L}(V)$ jest taka, że $\Psi(L_1 L_2) = \Psi(L_1)\Psi(L_2)$

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

i $\Psi(cL_1 + L_2) = c\Psi(L_1) + \Psi(L_2)$ 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 $L \in \mathcal{L}(\mathbb{F}^k)$ mamy $\dim(\mathrm{im}(\Psi(L)) = \dim(\mathrm{im}(L))$. (Wskazówka: $\dim(\ker(L)) = \max\{s :$ istnieją rzuty liniowe $P_1, ..., P_s \in \mathcal{L}(\mathbb{F}^k) \setminus \{\mathbf{0}\}$ takie, że $LP_i = 0$ 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. Macierze i operatory diagonalizowalne.

1. Jednokładności i podprzestrzenie własne.

Definicja. **Jednokładnością** (lub: **homotetią**) przestrzeni liniowej V, o **skali** λ, nazywamy przekształcenie λI_V (czyli zadane wzorem $\mathbf{v} \mapsto \lambda\mathbf{v}$ dla $\mathbf{v} \in V$).

Jednokładności są szczególnie prostymi przekształceniami liniowymi. Złożenie dwóch jednokładności znów jest jednokładnością, o skali będącej ilocznem skal wyjściowych jednokładności. W szczególności, dwie jednokładności są przemienne.

Definicja. a) **Przestrzeń własna operatora** $L \in \mathcal{L}(V)$, odpowiadająca wartości λ, określona jest tak:

$$V_L(\lambda) := \{\mathbf{v} \in V : L(\mathbf{v}) = \lambda\mathbf{v}\} = \ker(L - \lambda I). \tag{2}$$

b) Gdy $\underline{\mathbf{v} \neq \mathbf{0}}$ i $L(\mathbf{v}) = \lambda\mathbf{v}$, to mówimy, że $\mathbf{v}$ jest **wektorem własnym**, a λ **wartością własną** operatora L, i że **odpowiadają** one każde drugiemu.

Uwaga 1. Przestrzeń własna $V_L(\lambda)$ składa się więc z zera i wektorów własnych, odpowiadających wartości λ. Na przestrzeni tej, i na każdej jej podprzestrzeni, operator L działa jako homotetia (jednokładność) o skali λ.

Stwierdzenie 1. *Dla skalara λ i operatora $L \in \mathcal{L}(V)$, $\dim V < \infty$, 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 $\ker(L - \lambda I) \neq \{\mathbf{0}\}$. ⊟

Definicja. a) Zbiór skalarów λ, spełniających te równoważne warunki, nazywamy **spektrum** lub **widmem** operatora L i oznaczamy $\mathrm{spec}(L)$. Widmo jest więc skończonym zbiorem skalarów, złożonym ze wszystkich pierwiatków wielomianu χ_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

$$V_{\mathbf{A}}(\lambda) = \{\mathbf{v} \in \mathbb{F}^k : \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\} \ \text{ i } \ \mathrm{spec}(\mathbf{A}) = \{\lambda \in \mathbb{F} : \chi_{\mathbf{A}}(\lambda) = 0\} \tag{3}$$

Oczywiście, stwierdzenie 1 pozostaje słuszne dla macierzy kwadratowej $\mathbf{A}$ w miejsce operatora L. (Wystarcza je zastosować do operatora $L_{\mathbf{A}}$.)

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 $V_{\mathbf{A}}() = \mathbb{R}()$.

Przykład 2. Gdy $\mathbf{A} = \mathrm{diag}(\lambda_1, ..., \lambda_k)$, to $V_{\mathbf{A}}(\lambda) = \{(x_i)_{i=1}^k \in \mathbb{F}^k : x_i = 0 \text{ gdy } \lambda_i \neq \lambda\}$. Wymiar przestrzeni $V_{\mathbf{A}}(\lambda)$ jest więc w przypadku macierzy diagonalnej $\mathbf{A}$ równy krotności λ jako pierwiastka wielomianu $\chi_{\mathbf{A}}$.

Uwaga 2. 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 wielu wektorów własnych operatora L, odpowiadających jego różnym wartościom własnym, jest liniowo niezależny.*

Dowód. Niech wektorami tymi będą $\mathbf{v}_1, \dots \mathbf{v}_s$, i niech odpowiadają one różnym wartościom $\lambda_1, \dots, \lambda_s \in \mathbb{F}$. 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 $\sum_{i=1}^s c_i\mathbf{v}_i = \mathbf{0}$. Działając na obie strony operatorem L otrzymujemy zależność $\sum_i \lambda_i c_i \mathbf{v}_i = 0$. 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 $\sum_i c_i\mathbf{v}_i = \mathbf{0}$). □

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

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

Zadanie 1. Gdy $\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \mathbf{A}_2)$, to $V_{\mathbf{A}}(\lambda) = V_{\mathbf{A}_1}(\lambda) \times V_{\mathbf{A}_2}(\lambda)$. Ogólniej: gdy $\mathbf{A} = \mathrm{diag}(\mathbf{A}_1, \dots, \mathbf{A}_s)$, to $V_{\mathbf{A}}(\lambda) = \prod_i V_{\mathbf{A}_i}(\lambda)$ dla $\lambda \in \mathbb{F}$.

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

b) Gdy $\mathbf{A}$ 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 $\lambda \in \operatorname{spec}(K)$.

Zadanie 3. Jeśli baza przestrzeni V jest złożona z wektorów własnych operatora $L \in \mathcal{L}(V)$, to te z wektorów bazowych, które odpowiadają danej wartości własnej λ, tworzą bazę przestrzeni $V_L(\lambda)$.

Zadanie uzupełniające 1. Niech $V := C^\infty(\mathbb{R})$.

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

b) Wywnioskować, że funkcje $\sin(\lambda_1 t), ..., \sin(\lambda_n t)$ są liniowo niezależne gdy $|\lambda_i| \neq |\lambda_j| \neq 0$ dla $i \neq j$, a funkcje $\exp(\lambda_1 t), ..., \exp(\lambda_n t)$ – gdy $\lambda_i \neq \lambda_j$ dla $i \neq j$.

2. Macierze i endomorfizmy diagonalizowalne.

Definicja. a) **Operator** $L \in \mathcal{L}(V)$ **jest diagonalizowalny**, jeśli dla pewnej bazy $\mathcal{V}$ przestrzeni V, macierz $[L]_{\mathcal{V}}$ jest diagonalna. O bazie $\mathcal{V}$ powiemy, że **diagonalizuje** operator L.

b) Macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ jest diagonalizowalna, 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 $(\mathbf{v}_i)_{i=1}^k$ przestrzeni $\mathbb{F}^k$ diagonalizuje operator $L_{\mathbf{A}}$, a $\mathbf{S}$ jest macierzą o kolumnach $\mathbf{v}_1, ..., \mathbf{v}_k$, to $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ jest macierzą diagonalną. (Patrz wniosek 2 w §1.3.) Znalezienie takiej bazy, czy –równoważnie– macierzy $\mathbf{S}$, nazywamy **zagadnieniem diagonalizacji macierzy A przez podobieństwo**.

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 jest niemniejsza od wymiaru przestrzeni V, tzn. $\sum_{\lambda \in \operatorname{spec}(L)} \dim(V_L(\lambda)) \geq \dim(V)$*;*

d) V jest sumą prostą podprzestrzeni własnych operatora L, tzn. $V = \oplus_{\lambda \in \operatorname{spec}(L)} V_L(\lambda)$.

Dowód. Równoważność a) $\Leftrightarrow$ b) wynika z definicji bazy diagonalizującej.

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 $W := \sum_\lambda V_L(\lambda)$. Na podstawie wniosku 1, suma ta jest prosta. Wraz z c) daje to $\dim(W) = \sum_\lambda \dim(V_L(\lambda)) \geq \dim(V)$, tzn. $V = W = \oplus_\lambda V_L(\lambda)$.

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

Wniosek 1. *Operator* $L \in \mathcal{L}(V)$*, mający tyle różnych wartości własnych, ile wynosi* $\dim V$*, jest diagonalizowalny.*

Dowód. Ponieważ wymiar każdej podprzestrzeni własnej jest ≥ 1, więc spełniony jest warunek c) twierdzenia. $\square$

Uwaga 2. 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 nawet dla $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ pierwiastki te umiemy na ogół wyznaczyć tylko z pewnym przybliżeniem, to tu 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.

Podobnie, gdy diagonalizujemy macierz $\mathbf{A} \in \mathcal{M}_k$, to tworzymy bazy B_λ podprzestrzeni własnych $V_\mathbf{A}(\lambda)$. Jeśli liczebność zbioru $B = \bigcup\{B_\lambda : \lambda \in \mathrm{spec}(\mathbf{A})\}$ jest mniejsza od k, to macierz nie jest diagonalizowalna. W przeciwnym razie jest, a macierz $\mathbf{S}$, której kolumnami są wektory zbioru B, ma tę własność, że $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ jest macierzą diagonalną. Iloczynu $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ nie musimy wyliczać: na i–tym miejscu jego przekątnej stoi wartość własna, odpowiadająca i–tej kolumnie macierzy $\mathbf{S}$, patrz uwaga 1. $\square$

Uwaga 3. Czy jednak diagonalizacja jest możliwa, dowiadujemy się wyżej dopiero po (na ogół niewykonalnym) wyznaczeniu widma. W następnym paragrafie poznamy ogólne twierdzenia, które bez znajomości widma umożliwią rozpoznanie diagonalizowalności niektórych macierzy i operatorów (ale tylko dla $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$).

Uwaga 4. Z przykładu 2 w p. 1 i zadania 2b) tamże, wynika równoważność warunku c) twierdzenia z następującym, łatwiejszym do sprawdzenia:

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

Uwaga 5. Gdy macierz diagonalna $\mathbf{D}$ jest podobna do danej macierzy $\mathbf{A}$, lub jest macierzą danego operatora L w pewnej bazie, to na jej przekątnej występują pierwiastki wielomianu $\chi_\mathbf{A}$ (odp. wielomianu χ_L) , każdy powtarzany zgodnie ze swą krotnością.

(Patrz w §1.4 zadanie 2 i twierdzenie 1.) Tak więc **A** czy L wyznaczają **D** jednoznacznie, z dokładnością do kolejności wyrazów przekątnej. Istnieje jednak wiele macierzy ustalających podobieństwo **A** do **D**, i wiele baz diagonalizujących operator L. □

Podajmy przykłady ilustrujące powyższe wyniki.

Przykład 1. 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 **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 $\mathbf{S}^{-1}\mathbf{AS} = \mathbf{D}$ czyni zadość macierz **S**, której kolumnami są wektory własne odpowiadające, kolejno, wartościom własnym 1,4,6. (Ćwiczenie: wyznaczyć je.) □

Przykład 2. 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 $\mathbf{Av} = \lambda\mathbf{v} \Leftrightarrow v_1 + ... + v_k = \lambda v_i$ dla $i = 1, ..., k$, więc jeśli $\mathbf{Av} = \lambda\mathbf{v}$ 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 **S**, której kolejnymi kolumnami są wektory $(1, ..., 1), (-1, 1, 0, ..., 0), ..., (-1, 0, ..., 0, 1)$, zachodzi $\mathbf{S}^{-1}\mathbf{AS} = \operatorname{diag}(k, 0, ..., 0)$. (Daje to też $\chi_{\mathbf{A}} = (-x)^{k-1}(k - x)$; dlaczego?)

Zadania uzupełniające.

1. a) Czy liniowe rzuty i symetrie są diagonalizowalne?

b) Zbadać diagonalizowalność macierzy $\mathbf{A} \in \mathcal{M}_k$ takiej, że $a_{ij} = 1$ gdy $i + j = k + 1$ i $a_{ij} = 0$ w przeciwnym razie. Gdy macierz **A** jest diagonalizowalna znaleźć macierz nieosobliwą **S**, dla której $\mathbf{S}^{-1}\mathbf{AS}$ jest macierzą diagonalną.

2. Dla permutacji σ zbioru $\{1, ..., k\}$ oznaczmy przez $\mathbf{A}_\sigma \in \mathcal{M}_k(\mathbb{C})$ macierz, której j–tą kolumną jest $\mathbf{e}_{\sigma(j)}$ $(j = 1, ..., k)$. Zbadać, dla jakich σ macierz $\mathcal{A}_\sigma$ jest diagonalizowalna a) nad $\mathbb{C}$, b) nad $\mathbb{R}$.

Uwaga 6. * Niech operator $L \in \mathcal{L}(V)$ będzie diagonalizowalny. Wówczas $V = \oplus_\lambda V_L(\lambda)$, na podstawie twierdzenia 1. (Sumowanie po $\lambda \in \operatorname{spec}(L)$.) Oznaczmy przez P_λ rzutowanie V na $V_L(\lambda)$, wzdłuż $\oplus_{\mu \neq \lambda} V_L(\mu)$. Rodzina rzutów P_λ $(\lambda \in \operatorname{spec}(L))$ nazywana jest **rozkładem spektralnym** operatora diagonalizowalnego L. Nazwa bierze się od rozkładów $I_V = \sum_\lambda P_\lambda$ i $L = \sum_\lambda \lambda P_\lambda$; warto odnotować, że równości te wyznaczają rodzinę rzutów P_λ jednoznacznie.

Zadanie uzupełniające 3. * Uzasadnić ostatnie dwie równości i to, że wraz z warunkami

$P_\lambda^2 = P_\lambda$ wyznaczają one rodzinę P_λ. (Patrz też zad. uz. 1c) w §1.4.)

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

3. Funkcje macierzy diagonalizowalnych.

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 diagonalizowalna i znamy macierz $\mathbf{S}$ taką, że

$$\mathbf{A} = \mathbf{S}\mathrm{diag}(\lambda_1, ..., \lambda_k)\mathbf{S}^{-1} \tag{4}$$

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

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

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

$$\text{gdy } p_1(\lambda) = p_2(\lambda) \text{ dla każdego } \lambda \in \mathrm{spec}(\mathbf{A}), \text{ to } p_1(\mathbf{A}) = p_2(\mathbf{A}). \tag{5}$$

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. Istnieją wielomiany $p \in \mathbb{F}[x]$ takie, że $p(\lambda) = f(\lambda)$ dla wszystkich λ ze (skończonego!) zbioru $\mathrm{spec}(\mathbf{A})$. Choć istnieje ich wiele, to wartość $p(\mathbf{A})$ jest dla nich wspólna, na mocy (5). Podsumujmy to:

Definicja i uwaga. Dla diagonalizowalnej macierzy $\mathbf{A}$ obierzmy przedstawienie (4) dowolnie i przyjmijmy

$$f(\mathbf{A}) := \mathbf{S}\mathrm{diag}(f(\lambda_1), ..., f(\lambda_k))\mathbf{S}^{-1} \text{ dla } f : \mathrm{spec}(\mathbf{A}) \to \mathbb{F}. \tag{6}$$

Lewa strona jest niezależna od użytego przedstawienia (4) 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 (6) wynika, że macierz $f(\mathbf{A})$ jest diagonalizowalna. Nazywamy ją **wartością** (lub: **ewaluacją**) **funkcji f na diagonalizowalnej macierzy A**.

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 $(1/g)(\mathbf{A}) = (g(\mathbf{A}))^{-1}$ gdy $0 \notin g(\mathrm{spec}(\mathbf{A}))$.

c) $f(\mathbf{CAC}^{-1}) = \mathbf{C}f(\mathbf{A})\mathbf{C}^{-1}$ 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) Macierz $f(\mathbf{A})$ jest *wielomianem macierzy* $\mathbf{A}$, tzn. jest równa $\sum_{i=0}^{s} c_i \mathbf{A}^i$, dla pewnych $c_0, ..., c_s \in \mathbb{F}$, i wobec tego jest przemienna z każdą macierzą, przemienną z $\mathbf{A}$.

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 $(f(\mathbf{A}))^h = \overline{f}(\mathbf{A})$. W szczególności, zachodzi wówczas $\mathbf{A}^h = \sum_{i=0}^{s} c_i \mathbf{A}^i$ dla pewnych $c_0, ..., c_s \in \mathbb{C}$.

Ćwiczenie. Wyznaczyć $e^{\mathbf{A}}$ dla macierzy rzeczywistej $\mathbf{A} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$, na 3 sposoby: i) w oparciu o (6), dokonując diagonalizacji nad $\mathbb{C}$, ii) znajdując wartość $p(\mathbf{A})$ dla wielomianu p takiego, że $p(\lambda) = e^{\lambda}$ gdy $\lambda \in \mathrm{spec}(\mathbf{A})$, oraz iii) w oparciu o część g) zadania 1 i rozwijając *exp* w szereg.

Zadanie uzupełniające 1. Dla diagonalizowalnego operatora $L \in \mathcal{L}(V)$ określmy $f(L) \in \mathcal{L}(V)$ obierając w V dowolną bazę $\mathcal{V}$ i żądając, by $[f(L)]_{\mathcal{V}} = f([L]_{\mathcal{V}})$. Dowieść, że:

a) wynik nie zależy od wyboru bazy, a zadanie 1 pozostaje słuszne z L w miejsce $\mathbf{A}$.

b) ma miejsce równość $f(L) = \sum_{\lambda \in \mathrm{spec}(L)} f(\lambda) P_\lambda$, gdzie $\{P_\lambda : \lambda \in \mathrm{spec}(L)\}$ jest rozkładem spektralnym operatora diagonalizowalnego L, patrz w p.2 uwaga 3 i zad. uz. 3.

§ 3. Upraszczanie macierzy operatora wyborem bazy ortonormalnej.

Umowa: O ile nie powiedziano inaczej, w tym paragrafie rozważamy wyłącznie przestrzenie z iloczynem skalarnym, skończonego wymiaru. Gdy nie jest ono istotne, określenie ciała skalarów $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ pomijamy. Przestrzeń $\mathbb{F}^k$ rozpatrujemy ze standardowym iloczynem skalarnym.

1. Diagonalizacja ortonormalna.

Definicja. a) Operator $L \in \mathcal{L}(V)$ nazywamy **ortonormalnie diagonalizowalnym**, jeśli w pewnej ortonormalnej bazie $\mathcal{V}$ jego macierz $[L]_{\mathcal{V}}$ jest diagonalna.

b) Macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ nazywamy **unitarnie diagonalizowalną nad** $\mathbb{F}$, jeśli jest ona unitarnie podobna (nad $\mathbb{F}$) do macierzy diagonalnej. (Definicja i najprostsze własności tego podobieństwa są w §§1.1 – 1.3.)

Uwaga 1. a) 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 operatorem unitarnym gdy baza jest ortonormalna. Operator $L \in$

$\mathcal{L}(V)$ wtedy i tylko wtedy jest więc ortonormalnie diagonalizowalny, gdy jest unitarnie podobny do operatora $L_{\mathbf{D}} \in \mathcal{L}(\mathbb{F}^k)$ z diagonalną macierzą $\mathbf{D}$.

b) Podobnie zauważamy, że gdy macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ jest unitarnie diagonalizowalna nad $\mathbb{F}$, to operator $L_{\mathbf{A}} \in \mathcal{L}(\mathbb{F}^k)$ jest ortonormalnie diagonalizowalny, i vice versa. Możemy więc niekiedy i w odniesieniu do macierzy mówić o ich diagonalizowalności ortonormalnej (w miejsce unitarnej).

Uwaga 2. Jeśli operator L jest ortonormalnie diagonalizowalny, to jego różne podprzestrzenie własne są parami ortogonalne. Istotnie, dla ortonormalnej bazy $\mathcal{V} = (\mathbf{v}_i)_{i=1}^k$ takiej, ze macierz $[L]_{\mathcal{V}}$ jest diagonalna, zachodzi $V_L(\lambda) = \mathrm{lin}\{\mathbf{v}_i : L(\mathbf{v}_i) = \lambda \mathbf{v}_i\}$, wobec czego $V_L(\lambda) \perp V_L(\mu)$ gdy $\lambda \neq \mu$. (Patrz zad. 3 w §2.1.)

Uwaga 3. Możemy więc w następujący sposób badać, czy dany operator $L \in \mathcal{L}(V)$ jest ortonormalnie diagonalizowalny. Dla każdego $\lambda \in \mathrm{spec}(L)$ tworzymy bazę ortonormalną B_λ przestrzeni własnej $V_L(\lambda)$ (np. ortonormalizując metodą Grama–Schmidta dowolną bazę tej podprzestrzeni). Z uwag powyższej i 4 w §2.1 wynika, że odpowiedź jest pozytywna wtedy i tylko wtedy, gdy zbiór $\bigcup_\lambda B_\lambda$ jest ortonormalną bazą przestrzeni V; przy tym jeśli nią jest, to porządkując ją w dowolny sposób uzyskujemy ortonormalną bazę, której macierz operatora $[L]$ jest diagonalna. $\square$

Tak więc procedura diagonalizacji ortonormalnej jest prosta, gdy znamy widmo. Zaskakujące jest to, że macierze ortonormalnie 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.

Wyjaśnijmy też geometryczne znaczenie ortonormalnej diagonalizowalności.

Definicja. 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 $L \in \mathcal{L}(V)$ nazywamy **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 $L = \bigoplus_i L_i$.

Uwaga 4. a) Operator ortonormalnie diagonalizowalny jest sumą ortogonalną homotetii działających na jednowymiarowych podprzestrzeniach (rozpinanych przez poszczególne wektory ortogonalnej bazy diagonalizującej). Warto odnotować, że skale tych homotetii są wartościami własnymi dla L.

b) Przy $\mathbb{F} = \mathbb{C}$, podprzestrzeń wymiaru 1 wyobrażać sobie jednak trzeba nie jak prostą, 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 $|\lambda|$. Gdy skala λ jest rzeczywista, to mamy do czynienia z „kurczeniem" lub „rozciąganiem" podprzestrzeni, złożonym być może z odbiciem względem zera (tak przy $\mathbb{F} = \mathbb{C}$, jak i $\mathbb{F} = \mathbb{R}$). □

Zadanie uzupełniające 1. Udowodnić, że gdy kwadratowa macierz rzeczywista ma tylko rzeczywiste wartości własne, to z jej diagonalizowalności nad $\mathbb{C}$ wynika diagonalizowalność nad $\mathbb{R}$, i tak samo dla ortonormalnej diagonalizowalności.

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

2. Charakteryzacja ortonormalnej diagonalizowalności w przypadku zespolonym.

Kwadratową macierz zespoloną $\mathbf{A}$ nazywamy:

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

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

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

Podobnie, operator L na zespolonej lub rzeczywistej przestrzeni unitarnej nazywamy normalnym (odp. samosprzężonym wzgl. antysamosprzężonym), jeśli $L^hL = LL^h$ (odp. $L^h = L$ wzgl. $L^h = -L$). Gdy $\mathbb{F} = \mathbb{R}$, to w miejsce „(anty)hermitowski" mówi się też **(anty)symetryczny**.

Lemat 1. *Niech $\mathcal{V}$ będzie ortonormalną bazą przestrzeni unitarnej V. Dla operatora $L \in \mathcal{L}(V)$ równoważne są warunki:*

a) L jest operatorem normalnym (odp. samosprzężonym, antysamosprzężonym, unitarnym)

b) $[L]_\mathcal{V}$ jest macierzą normalną (odp. samosprzężoną, antysamosprzężoną, unitarną).

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

Zadanie 1. Macierz $\mathrm{diag}(\mathbf{K}_1, \mathbf{K}_2)$ jest normalna (odp. samosprzężona, antysamosprzężona, unitarna) wtedy i tylko wtedy, gdy takie są klatki $\mathbf{K}_1$ i $\mathbf{K}_2$.

Uwaga 1. a) Gdy operator jest samosprzężony lub antysamosprzężony lub unitarny, to jest normalny, i analogicznie dla macierzy.

b) Macierz diagonalna jest normalna, a gdy jest rzeczywista, to jest samosprzężona.

c) Macierz, unitarnie podobna do normalnej, jest normalna (i tak samo dla macierzy samosprzężonych czy antysamosprzężonych w miejsce normalnych).

d) Część c) pozostaje słuszna dla operatorów. ⊟

Lemat 2. *i-ty wiersz macierzy normalnej $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ ma tę samą normę, co jej i-ta kolumna, dla $i = 1, ..., k$.*

Dowód. Dla każdego wektora $\mathbf{v} \in \mathbb{C}^k$ zachodzi $\|\mathbf{A}\mathbf{v}\|^2 = \langle \mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{v}\rangle = \langle \mathbf{A}^h\mathbf{A}\mathbf{v}, \mathbf{v}\rangle$ i podobnie $\|\mathbf{A}^h\mathbf{v}\|^2 = \langle \mathbf{A}\mathbf{A}^h\mathbf{v}, \mathbf{v}\rangle$, wobec czego $\|\mathbf{A}\mathbf{v}\| = \|\mathbf{A}^h\mathbf{v}\|$ na podstawie normalności macierzy. Stosując to przy $\mathbf{v} = \mathbf{e}_i$ stwierdzamy, że i–te kolumny macierzy $\mathbf{A}$ i $\mathbf{A}^h$ mają tę samą normę, a stąd wynika teza. □

Dowiedziemy teraz zaskakującej i ważnej charakteryzacji:

Twierdzenie 1 (O.Toeplitza). *a) Operator normalny, L, na zespolonej przestrzeni unitarnej jest ortonormalnie diagonalizowalny.*

b) (Wersja macierzowa:) Zespolona macierz normalna jest unitarnie (nad $\mathbb{C}$) podobna do macierzy diagonalnej.

c) Również odwrotnie: ortonormalnie diagonalizowalny operator i ortonormalnie diagonalizowalna macierz są normalne.

Dowód. Ponieważ macierze diagonalne są normalne, więc c) wynika z lematu 1 (dla operatorów) i części c) uwagi 1 (dla macierzy).

Pokażmy następnie, jak z b) wynika a). Niech $\mathbf{C} := [L]_{\mathcal{B}}$, gdzie $\mathcal{B}$ jest dowolną bazą ortonormalną. Korzystając z b) i lematu 1 stwierdzamy, że $\mathbf{C} = \mathbf{S}^{-1}\mathbf{D}\mathbf{S}$, gdzie $\mathbf{S}$ jest macierzą ortonormalną, a $\mathbf{D}$ – diagonalną. Baza $\mathcal{V}$ przestrzeni E taka, że $[I]_{\mathcal{B}}^{\mathcal{V}} = \mathbf{S}$, spełnia warunek $[L]_{\mathcal{V}} = \mathbf{D}$ i jest ortonormalna (patrz wniosek 3 w §V.3.1).

Tezy b) dowiedziemy indukcyjnie względem stopnia k badanej macierzy $\mathbf{A}$. Obierzmy dowolny jej wektor własny $\mathbf{v}$ (taki istnieje, bo $\operatorname{spec}_{\mathbb{C}}\mathbf{A} \neq \emptyset$), unormujmy go i rozszerzmy do ortonormalnej bazy $\mathcal{V}$ przestrzeni $\mathbb{C}^k$, w której występuje jako ostatni. Niech $\mathbf{B} = [L_{\mathbf{A}}]_{\mathcal{V}}$; macierz $\mathbf{B}$ jest unitarnie podobna do $\mathbf{A}$ (patrz §1.3) i normalna (patrz lemat 1). Jej ostatnia kolumna i ostatni wiersz to $(0, ..., 0, b_{kk})$ i $(b_{k1}, ..., b_{kk})$, odpowiednio; a że są one tej samej długości (patrz lemat 2), to $b_{ki} = 0$ dla $i = 1, ..., k-1$. Tym samym $\mathbf{B} = \operatorname{diag}(\mathbf{B}', b_{kk})$ dla pewnej klatki $\mathbf{B}' \in \mathcal{M}_{k-1}(\mathbb{C})$, która jest normalna na podstawie zadania 1. Stosując założenie indukcyjne i zad. 2c) z §1.1, uzyskujemy tezę. □

Ćwiczenie. Zdefiniujmy operator $L \in \mathcal{L}(\mathbb{C}^k)$ wzorem $L(\mathbf{v}) = (v_2, ..., v_k, v_1)$ dla $\mathbf{v} = (v_1, \ldots, v_k) \in \mathbb{C}^k$.

a) Dowieść, że jest on unitarny, a jego widmo to zbiór pierwiastków stopnia k z 1.

b) Znaleźć bazę diagonalizującą ten operator i uzasadnić jej ortogonalność.

Wniosek 1. *Niech L będzie operatorem na zespolonej przestrzeni unitarnej.*

a) Jeśli operator L jest samosprzężony, to w pewnej bazie ortonormalnej jego macierz jest diagonalna i rzeczywista.

b) Jeśli operator L jest unitarny, to w pewnej bazie ortonormalnej jego macierz jest diagonalna i na przekątnej ma wyłącznie wyrazy o module 1;

c)Jeśli operator L jest antysamosprzężony, to w pewnej bazie ortonormalnej jego macierz jest diagonalna i na przekątnej ma wyłącznie wyrazy czysto urojone.

Implikacje odwrotne też są prawdziwe dla operatorów normalnych.

Dowód. Ad a) Gdy operator jest samosprzężony, to jest normalny, więc jego macierz w pewnej bazie ortonormalnej jest diagonalna. Na podstawie lematu 1 jest ona zarazem samosprzężona, wobec czego wyrazy jej przekątnej są rzeczywiste. Implikacja odwrotna również wynika z lematu 1, połączonego z twierdzeniem Toeplitza.

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

Oczywiście, wniosek 1 ma macierzowy odpowiednik:

Wniosek 2. *a) Macierz samosprzężona jest unitarnie podobna do rzeczywistej macierzy diagonalnej. (Tu i niżej, mowa o unitarnym podobieństwie nad $\mathbb{C}$.)*

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 odwrotne też są prawdziwe dla macierzy normalnych. ⊟

Na przekątnej macierzy diagonalnej występują jednak wszystkie elementy jej widma, i tylko one. Z definicji χ_L i twierdzenia 1 w §1.4 wynika więc taka wersja wniosków 1 i 2:

Wniosek 3. *a) Gdy macierz zespolona jest samosprzężona, to jej widmo leży w $\mathbb{R}$, gdy jest unitarna – to leży ono w okręgu $|z| = 1$, a gdy antysamosprzężona, to leży w $i\mathbb{R}$.*

b) Dla zespolonych macierzy normalnych prawdziwe są też implikacje odwrotne.

c) Analogicznie jest dla operatorów na zespolonej przestrzeni unitarnej. □

Zadania. (Formułuję je dla macierzy; podobnie można dla operatorów).

2. a) Suma $\mathbf{A} + \mathbf{B}$ i iloczyn $\mathbf{AB}$ macierzy normalnych jest taką macierzą, jeśli $\mathbf{AB}^h = \mathbf{B}^h\mathbf{A}$. (W przypadku sumy wystarcza też, by $\mathbf{AB}^h = \mathbf{A}^h\mathbf{B}$.)

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

3. a) Każda macierz kwadratowa jest sumą macierzy samosprzężonej i macierzy antysamosprzężonej – a więc sumą dwóch macierzy normalnych, patrz uwaga 1.

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

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

Zadania uzupełniające. Niżej, L jest endomorfizmem zespolonej przestrzeni unitarnej. (Wskazówka: w zadaniach 1 – 5 założyć wpierw, że $V = \mathbb{C}^k$ i macierz $[L]$ jest diagonalna.)

1. Dowieść, że gdy $\mathbf{v}$ jest wektorem własnym operatora normalnego L, odpowiadającego wartości λ, to jest też wektorem własnym dla L^h, odpowiadającym wartości $\overline{\lambda}$.

2. Dla operatora L dowieść równoważności warunków:
 a) operator L jest normalny i $\mathrm{spec}(L) \subset [0, \infty)$;
 b) operator L jest samosprzężony i dodatnio półokreślony (tzn. dla wszystkich wektorów $\mathbf{v}$ jest $\langle L(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, L(\mathbf{v})\rangle \geq 0$);
 c) w pewnej bazie ortonormalnej, operator L ma macierz diagonalną, z (wyłącznie) nieujemnymi wyrazami na przekątnej;
 d) $L = K^2$ dla pewnego samosprzężonego operatora K, który też jest dodatnio półokreślony;
 e) $L = K^h K$ dla pewnego operatora K.

3. Dowieść, że gdy operator L jest normalny, to a) $\ker(L) = \ker(L^h)$ i $\mathrm{im}(L) = \mathrm{im}(L^h)$, oraz b) $\ker(L) = \mathrm{im}(L)^\perp$.

4. Niech $\mathbf{v}$ będzie wektorem własnym operatora normalnego L. Dowieść, że:
 a) $\mathbf{v}^\perp$ jest podprzestrzenią niezmienniczą dla L i dla L^h, a operator $L_{|\mathbf{v}^\perp}$ jest normalny.
 b) Jeśli operator L jest samosprzężony, to $L_{|\mathbf{v}^\perp}$ też.

5. 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 2. Zadanie 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 zawęziwszy L. (Może się zdażyć, że $a = b$, lecz to nie przeszkadza.)

6. Niech macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ będzie normalna. Dowieść, że
 a) Gdy $k = 2$, to macierz $\mathbf{A}$ jest proporcjonalna do macierzy obrotu lub $\mathbf{A} = \mathbf{A}^t$.
 b) Jeśli macierz $\mathbf{A}$ jest trójkątna, to jest diagonalna.

Zadania ze zbioru Kostrykina: 18,19,20,21 w §II.4.2.

3. Operatory normalne na przestrzeniach euklidesowych (= unitarnych rzeczywistych).

W tym punkcie rozważamy operatory działające na przestrzeni euklidesowej (czyli unitarnej rzeczywistej). By to podkreślić, przestrzeń oznaczamy literą E.

Twierdzenie 1. *Operator $L \in \mathcal{L}(E)$ na przestrzeni euklidesowej E jest ortonormalnie diagonalizowalny wtedy i tylko wtedy, gdy jest samosprzężony.*

<u>Wersja macierzowa.</u> *Rzeczywista macierz kwadratowa wtedy i tylko wtedy jest ortogonalnie (czyli unitarnie nad $\mathbb{R}$) podobna do macierzy diagonalnej, gdy jest symetryczna.*

<u>Dowód.</u> Powtarzamy dowód twierdzenia Toeplitza. Różnica jest ta, że operator $L_{\mathbf{A}}$, wyznaczony przez badaną macierz $\mathbf{A}$, traktujemy jako określony na $\mathbb{R}^k$ (a nie na $\mathbb{C}^k$). Na podstawie wniosku 3 w p.2, $\mathbf{A}$ ma pewną <u>rzeczywistą</u> wartość własną λ; niech odpowiada jej jednostkowy wektor własny $\mathbf{v} \in \mathbb{R}^k$. Jak w p.2, rozszerzamy $\mathbf{v}$ do bazy ortonormalnej $\mathcal{V}$ przestrzeni $\mathbb{R}^k$, w której występuje jako ostatni; jak tam, $[L_{\mathbf{A}}]_{\mathcal{V}} = \operatorname{diag}(\mathbf{B}', b_{kk})$ dla pewnej klatki $\mathbf{B}' \in \mathcal{M}_{k-1}(\mathbb{R})$, która tym razem jest samosprzężona (znów na podstawie części b) uwagi 1 z p.2.) Pozostała część rozumowania pozostaje bez zmian. □

Operatory normalne na przestrzeni euklidesowej nie są więc diagonalizowalne ortonormalnie, poza przypadkiem, gdy są samosprzężone. Dla przykładu: różny od identyczności obrót płaszczyzny E^2 nie jest diagonalizowalny, choć jest operatorem normalnym (bo jest izometrią). Jednak i w przypadku rzeczywistym postaramy się zbadać macierze czy operatory normalne, które niekoniecznie są samosprzężone.

<u>Oznaczenie</u> Dla $\lambda = a + b\mathbf{i}$ $(a, b \in \mathbb{R})$ przyjmijmy

$$K(\lambda) = (\lambda) \in \mathcal{M}_1(\mathbb{R}) \text{ gdy } \lambda \in \mathbb{R} \quad \text{i} \quad \mathbf{K}(\lambda) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ gdy } b \neq 0.$$

Zauważmy, że gdy $\lambda \notin \mathbb{R}$, 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.) Odnotujmy, że gdy $\lambda \notin \mathbb{R}$, to $\operatorname{spec}_{\mathbb{C}} K(\lambda) = \{\lambda, \overline{\lambda}\}$.

Twierdzenie 2. *Gdy operator $L \in \mathcal{L}(E)$ na przestrzeni euklidesowej E jest normalny, to w pewnej bazie ortonormalnej jego macierz ma postać*

$$\mathbf{B} = \operatorname{diag}(\mathbf{K}(\lambda_1), ..., \mathbf{K}(\lambda_n)), \ \ \textit{gdzie } \lambda_1, ..., \lambda_n \in \mathbb{C}. \tag{7}$$

<u>Wersja macierzowa.</u> *Rzeczywista macierz normalna, $\mathbf{A}$, jest ortogonalnie podobna do macierzy $\mathbf{B}$ powyższej postaci.*

Uwaga 1. * Zmiana znaku wybranych wektorów bazy pozwala uzyskać, by $\operatorname{Im}(\lambda_i) \geq 0$ dla $i = 1, ..., n$. Ponieważ $\chi_L = \chi_{\mathbf{B}} = \prod_i \chi_{\mathbf{K}(\lambda_i)}$, to wtedy liczby $\lambda_1, ..., \lambda_n$ są wyznaczone jednoznacznie, prócz swej kolejności, jako pierwiastki wielomianu χ_L o nieujemnej części urojonej. (Uwzględniamy krotności.) Tak samo jest dla macierzy. □

Przed podaniem dowodu twierdzenia 2 wskażemy na jego sens geometryczny w ważnym przypadku szczególnym. (Pełną ogólność da zadanie uzupełniające 1.)

Wniosek 1 (Uogólnione twierdzenie Eulera)**.** *Izometria liniowa przestrzeni euklidesowej E jest sumą ortogonalną rodziny operatorów, z których każdy poza być może jednym jest obrotem liniowym płaszczyzny lub identycznościowym przekształceniem prostej, a pozostały (jeśli istnieje) jest odbiciem prostej względem zera.*

Dowód. Obierzmy ortonormalną bazę $\mathcal{V}$, dla której macierz $\mathbf{B} := [L]_{\mathcal{V}}$ ma postać (7). Ponieważ L jest izometrią, to $|\lambda_i| = 1\ \forall i$; patrz wniosek 3 w p.2. W szczególności, $\lambda_i = \pm 1$ dla rzeczywistych λ_i, a dla pozostałych klatki $K(\lambda_i)$ są macierzami obrotu. Jednak $-\mathbf{I}_2$ też jest macierzą obrotu (o π radianów); a że liczby $\lambda_1, ..., \lambda_s$ możemy permutować, przez zmianę kolejności wektorów bazy, to łącząc w pary ujemne wartości λ_i uzyskamy, by macierz $\mathbf{B}$ była sumą zewnętrzną klatek, z których każda poza być może jedną jest równa (1) lub jest macierzą obrotu płaszczyzny; pozostała zaś klatka (jeśli istnieje) jest równa (-1). Stąd już bezpośrednio wynika teza. $\square$

W dowodzie twierdzenia 2 wykorzystamy

Zadanie 1. Dla $\mathbf{v} = (v_1, ..., v_k) \in \mathbb{C}^k$ przyjmijmy

$$\overline{\mathbf{v}} := (\overline{v_1}, ..., \overline{v_k}), \quad \mathrm{Re}(\mathbf{v}) := \frac{1}{2}(\mathbf{v} + \overline{\mathbf{v}}), \quad \mathrm{Im}(\mathbf{v}) := \frac{1}{2\mathbf{i}}(\mathbf{v} - \overline{\mathbf{v}})$$

a) i Jeśli $\mathbf{v} \in \mathbb{C}^k$ i $\mathbf{v} \perp \overline{\mathbf{v}}$, to $\|\mathrm{Re}(\mathbf{v})\| = \|\mathrm{Im}(\mathbf{v})\| = \|\mathbf{v}\|/\sqrt{2}$ i $\mathrm{Re}(\mathbf{v}) \perp \mathrm{Im}(\mathbf{v})$.

b) Gdy $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$ i $\mathbf{v} \in \mathbb{C}^k$ spełniają warunek $\mathbf{A}\mathbf{v} = (a + bi)\mathbf{v}$, gdzie $a, b \in \mathbb{R}$, to $\mathbf{A}(\mathrm{Re}(\mathbf{v})) = a\mathrm{Re}(\mathbf{v}) - b\mathrm{Im}(\mathbf{v})$ i $\mathbf{A}(\mathrm{Im}(\mathbf{v})) = b\mathrm{Re}(\mathbf{v}) + a\mathrm{Im}(\mathbf{v})$.

Dowód twierdzenia 2. Znów jest on podobny do dowodu twierdzenia Toeplitza; jak w nim, zajmiemy się wersją macierzową i wykorzystamy indukcję względem stopnia macierzy. Gdy $\mathbf{A}$ ma rzeczywistą wartość własną, to krok indukcyjny nie różni się od tego w dowodzie twierdzenia Toeplitza.

Niech więc $\lambda = a + bi \notin \mathbb{R}$ i wektor jednostkowy $\mathbf{v} \in \mathbb{C}^k$ będą takie, że $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Wtedy $\lambda \neq \overline{\lambda}$, zaś $\mathbf{v}$ i $\overline{\mathbf{v}}$ są odpowiadającymi tym wartościom wektorami własnymi macierzy $\mathbf{A}$. Tym samym $\overline{\mathbf{v}} \perp \mathbf{v}$, na podstawie twierdzenia Toeplitza i uwagi 2 w p.1. Z zadania wynika więc, że układ $\sqrt{2}\mathrm{Im}(\mathbf{v}), \sqrt{2}\mathrm{Re}(\mathbf{v})$ jest ortonormalny; rozszerzmy go do ortonormalnej bazy $\mathcal{V}$ przestrzeni rzeczywistej $\mathbb{R}^k$, w ktorej te dwa wektory są na końcu. Macierz $[L_{\mathbf{A}}]_{\mathcal{V}}$ jest unitarnie podobna do $\mathbf{A}$ i normalna, a jej ostatnie dwie kolumny to $(0, ..., 0, a, b)$ i $(0, ..., 0, -b, a)$; por. zadanie 1b). Stąd jej ostatnie dwa wiersze są postaci $(x_1, ..., x_{k-2}, a, -b)$ i $(y_1, ..., y_{k-2}, b, a)$, przy czym $x_i = y_i = 0\ \forall i$, wobec uwagi 1d) z p.2. Pozwala to wykonać krok indukcyjny, tym razem przy $[L]_{\mathcal{V}} = \mathrm{diag}(\mathbf{B}', \mathbf{K}(\lambda))$. $\square$

Zadania uzupełniające. Dowieść, że:

1. a) Operator normalny na przestrzeni euklidesowej jest sumą ortogonalną operatorów, z których każdy jest homotetią podprzestrzeni wymiaru 1 lub działa na podprzestrzeni dwuwymiarowej i jest proporcjonalny do pewnego jej obrotu.

b) Gdy operator jest antysymetryczny, to powyższe obroty są o $\pi/2$ radianów, a homotetie są zerowe (lub ich nie ma). W szczególności, operator taki ma wyznacznik nieujemny, a rząd – parzysty. (Rząd operatora rozumiemy jak w rozdziale III.)

c) Jeśli w b) rząd jest równy $2m$, to operator jest sumą m operatorów antysymetrycznych rzędu 2.

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

3. Izometria $L \in \mathcal{L}(E)$ zachowuje orientację wtedy i tylko wtedy, gdy wymiar podprzestrzeni $Fix(L)$ o liczbę parzystą różni się od wymiaru przestrzeni euklidesowej E.

4. a) Izometria liniowa przestrzeni $\mathbb{R}^3$ jest obrotem wokół prostej lub złożeniem obrotu z odbiciem zwierciadlanym wzdłuż osi obrotu. Uzyskać też taki rozkład dla izometrii zadanej macierzą o wierszach $(2/3, 2/3, -1/3), (2/3, -1/3, 2/3)$, $(-1/3, 2/3, 2/3)$.

b) Zmieniająca orientację izometria przestrzeni $\mathbb{R}^4$ jest złożeniem pewnej symetrii zwierciadlanej S, z kanonicznym przedłużeniem pewnego obrotu przestrzeni $Fix(S)$. Dowieść tego i uzyskać taki rozkład, gdy $L(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4, x_1)$.

Zadania ze zbioru Kostrykina: 5 w §II.4.4 („Baza kanoniczna" to taka, w której macierz operatora ma postać (7)).

4. Związki między operatorami (anty)hermitowskimi a unitarnymi; rozkład biegunowy.

Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ będzie macierzą unitarnie diagonalizowalną nad $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, a $f : \mathrm{spec}(\mathbf{A}) \to \mathbb{F}$ pewną funkcją. Ponieważ we wzorach (4) i (6) w §2.3 można za $\mathbf{S} \in \mathcal{M}_k(\mathbb{F})$ obrać macierz unitarną, to macierz $f(\mathbf{A})$ jest unitarnie diagonalizowalna, z widmem $f(\mathrm{spec}(\mathbf{A}))$. Gdy więc $f(\mathrm{spec}(\mathbf{A})) \subset \mathbb{R}$, to $f(\mathbf{A})$ możemy rozpoznać jako macierz samosprzężoną, gdy $f(\mathrm{spec}(\mathbf{A})) \subset [0, \infty)$ –jako nieujemnie określoną, a gdy $f(\mathrm{spec}(\mathbf{A})) \subset \mathbb{R}i$ –jako antysamosprzężoną; wreszcie jeśli $f(\mathrm{spec}(\mathbf{A})) \subset \{z \in \mathbb{C} : |z| = 1\}$, to $f(\mathbf{A})$ jest macierzą unitarną. (Wykorzystujemy wniosek 3 w p.2.) Umożliwia to konstruowanie 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 $\mathcal{H}_+ \subset \mathcal{M}_k(\mathbb{F})$ będzie zbiorem nieujemnie określonych macierzy hermitowskich. Ponieważ $\mathrm{spec}(\mathbf{A}) \subset [0, \infty)$ dla $\mathbf{A} \in \mathcal{H}_+$ (patrz zad. uz. 3 w p.3 lub poniższy lemat 1), 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 $\mathbf{A} \in \mathcal{H}_+$ ma dokładnie jeden należący do $\mathcal{H}_+$ pierwiastek zadanego stopnia n. $\square$

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, nieujemnie określonej macierzy $\mathbf{A}$. □

Przykład 3. Homografia $z \mapsto (1 - z)/(1 + z)$ jest swą odwrotnością i przeprowadza oś urojoną na $S^1 \setminus \{-1\}$. Wynika stąd, że $\mathbf{A} \mapsto (\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$ przekształca zbiór zespolonych 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. (Wyznacznik takiej macierzy siłą rzeczy jest równy 1, co wynika z uogólnionego twierdzenia Eulera z p.3.) □

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 na przestrzeni unitarnej V (rzeczywistej lub zespolonej) równoważne są warunki*

a) $\langle L(\mathbf{v}), \mathbf{v}\rangle \geq 0$ *dla wszystkich* $\mathbf{v} \in V$*;*

b) $\operatorname{spec}(L) \subset [0, \infty)$.

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, udowodnione w pp. 2 i 3. ⊟

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

Twierdzenie 1. *Każdą macierz* $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$*, dla* $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$*, 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 (rzeczywistą lub zespoloną) , to istnieją operatory* $U, H \in \mathcal{L}(V)$ *takie, że* $L = UH$ *i operator* $H = H^h$ *jest nieujemnie określony, a* U *– unitarny (tzn. jest izometrią liniową).*

Dodatek: *W tym przedstawieniu, operator H jest przez L wyznaczony jednoznacznie. Gdy L jest izomorfizmem, to i operator U jest wyznaczony jednoznacznie.*

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 poprzednich punktów, twierdzenie daje przejrzysty opis operatora działającego na przestrzeni euklidesowej: jest on złożeniem operatora H, polegającego na „rozciąganiu" przestrzeni w pewnych wzajemnie ortogonalnych kierunkach, z izometrią liniową U, której działanie opisuje uogólnione twierdzenie Eulera z p.5.

Z lewego rozkładu biegunowego $L = UH$ otrzymujemy prawy taki rozkład: $L = H_1U$, gdzie operator $H_1 := UHU^h$ nadal jest samosprzężony i nieujemnie określony. □

Dowód twierdzenia. Jeśli żądane przedstawienie istnieje, to $L^hL = H^h(U^hU)H = H^hH = H^2$. Definiujemy zatem H jako samosprzężony, nieujemnie określony pierwiastek z L^hL. Pierwiastek taki, i to jedyny, istnieje, bo $(L^hL)^h = L^hL$ i $\langle L^hL(\mathbf{v}), \mathbf{v}\rangle = \langle L(\mathbf{v}), L(\mathbf{v})\rangle \geq 0$ 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

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

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 wnosimy z (8), że $L(\mathbf{v}) = \mathbf{0}$ gdy $H(\mathbf{v}) = \mathbf{0}$, wobec czego istnieje operator liniowy $U_0 : H(V) \to V$ taki, że $L = U_0 \circ H$. (Patrz ....). Z (8) wynika, że $\|U_0(\mathbf{w})\| = \|\mathbf{w}\|$ dla $\mathbf{v} \in H(V)$.

Za $U : V \to V$ obieramy izometrię liniową, równą U_0 na $H(V)$; patrz zadanie 4 w §V.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?

2. a) Dowieść, że każdą macierz $\mathbf{A} \in \mathcal{M}_{l,k}(\mathbb{F})$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, można przedstawić w postaci $\mathbf{A} = \mathbf{U_1DU_2}$, gdzie macierze $\mathbf{U}_1 \in \mathcal{M}_l(\mathbb{F})$ i $\mathbf{U}_2 \in \mathcal{M}_k(\mathbb{F})$ są unitarne oraz $d_{ij} = 0$ dla $i \neq j$, przy czym $d_{ii} \geq 0$ dla wszystkich $i = 1, ..., \min(l, k)$. (Wskazówka: gdy $k \leq l$ obrać $\mathbf{U}_1$ tak, by operator $L_{\mathbf{U}_1}$ przeprowadzał $\mathrm{im}(L_{\mathbf{A}})$ w $\mathbb{F}^k \times \{\mathbf{0}_{l-k}\}$; w przeciwnym razie zastąpić $\mathbf{A}$ przez $\mathbf{A}^h$.)

b) Prócz swej kolejności, skalary d_{ii} są powyżej jednoznacznie wyznaczone, jako nieujemne pierwiastki kwadratowe z wartości własnych macierzy $\mathbf{A}^h\mathbf{A}$. (Ich angielska nazwa to „singular values”, a nazwa rozkładu to **singular value decomposition**.)

3. Niech $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ i $\det(\mathbf{A}) \neq 0$. Dowieść, że gdy $\mathbf{A} = \mathbf{UH}$ jest rozkładem biegunowym, to $\|\mathbf{A} - \mathbf{W}\| > \|\mathbf{A} - \mathbf{U}\|$ dla każdej unitarnej macierzy $\mathbf{W} \neq \mathbf{U}$. (Wskazówka: zadania uzupełniające 1c) i 2 w §V.3.3.)

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

5. Twierdzenie Schura o unitarnym podobieństwie do macierzy trójkątnej.

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. Udowodnimy wersję „operatorową”, 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 równa $(0, \ldots, 0, \lambda)$; zaś kolumna i–ta dla $i < k$, równa $[L(\mathbf{b}_i)]_{\mathcal{B}}$, zaczyna się od $i - 1$ zer – bo tak zaczyna się ciąg $[PL(\mathbf{b}_i)]_{\mathcal{B}}$, różniący się od niej tylko ostatnim wyrazem. □

Zadania uzupełniające. (W 1, 2a), 3a) i 4 założyć wpierw, że macierz jest trójkątna.)

1. Dla macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ dowieść nierówności $\sqrt{\mathrm{rk}(\mathbf{A})} \geq |\mathrm{tr}(\mathbf{A})|/\sqrt{\mathrm{tr}(\mathbf{A}\mathbf{A}^h)}$.

2. a) Udowodnić **nierówność Schura**: dla $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ jest $\sum |a_{ij}|^2 \geq \sum |\lambda_i|^2$, gdzie $\lambda_1, \lambda_2, \ldots$ to wszystkie pierwiastki wielomianu $\chi_{\mathbf{A}}$, powtarzane zgodnie z ich krotnościami.
b) Dowieść, że wyżej równość zachodzi dla macierzy normalnych, i tylko dla nich.
c) Dowieść, że macierz $\mathbf{A}$ jest normalna wtedy i tylko wtedy, gdy $\|\mathbf{A}\mathbf{A}^h\| = \|\mathbf{A}\|^2$, gdzie $\|.\|$ jest jak w zadaniu uz. 1 z §V.3.3.

3. Niech $\mathbf{A}, \mathbf{B} \in \mathcal{M}_k(\mathbb{C})$; wektory $\mathbf{u} \in \mathbb{C}^k$ traktujemy jako kolumny.
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 $\mathbf{B} = \mathbf{A}^h$
c) Dowieść, że gdy $\mathbf{u}^h\mathbf{Au} \in \mathbb{R}$ dla każdego wektora $\mathbf{u} \in \mathbb{C}^k$, to $\mathbf{A}^h = \mathbf{A}$.
d) Czy tezy te pozostaną prawdziwe, gdy $\mathbb{C}^k$ zamienić na $\mathbb{R}^k$?

4. Dowieść, że normalność macierzy $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ jest równoważna temu, by $\|\mathbf{Av}\| = \|\mathbf{A}^h\mathbf{v}\|$ dla każdego wektora $\mathbf{v} \in \mathbb{C}^k$.

5. a) Dowieść, że kwadratowa macierz rzeczywista jest ortogonalnie podobna do macierzy rzeczywistej, wzdłuż przekątnej której stoją klatki stopnia 1 i 2×2–klatki proporcjonalne do macierzy obrotu, a nad tymi klatkami – zera.

b) (Przypisane w [Kostrykin] D. Djokovićowi.) Dowieść, że jeśli wyżej wszystkie klatki są stopnia 2, to macierz komutuje z pewną macierzą, której kwadrat jest równy $-\mathbf{I}$.

6. * $L \in \mathcal{L}(V)$ jest operatorem, który nie jest proporcjonalny do I_V. W tym zadaniu zakładamy, że w ciele skalarów $\mathbb{F}$ jest $n_{\mathbb{F}} \neq 0_{\mathbb{F}}$ $\forall n$. Dowieść że:

a) Dla każdego skalara c istnieje rozkład $V = \mathbb{F}\mathbf{w} \oplus U$ taki, że $\mathbf{w} \neq \mathbf{0}$ i $L(\mathbf{w}) - c\mathbf{w} \in U$.

b) Przekątna macierzy operatora L w pewnej bazie jest równa $(0, ..., 0, \mathrm{tr}(L))$. (Wskazówka: dowód tw. Schura, przy $\mathbf{w}$ i U obranych jak wyżej dla $c = \mathrm{tr}(L)$.)

c) Jeśli $\mathrm{tr}(L) = 0$, to L jest komutatorem, tzn. $L = XY - YX$ dla pewnych $X, Y \in \mathcal{L}(V)$. (Wskazówka: zad. 9 w §II.2.1.)

<u>Problem 2.</u> Nadal, niech $L \in \mathcal{L}(V) \setminus \mathrm{lin}(I_V)$. Dowieść, że przekątną macierzy $[L]_{\mathcal{V}}$ można dla pewnej bazy $\mathcal{V}$ uczynić równą zadanemu ciągowi $(c_i)_{i=1}^k \in \mathbb{F}^k$, o sumie $\mathrm{tr}(L)$.

§ 4. Postać kanoniczna Jordana.

Dowodzone w tym pragrafie twierdzenie Jordana jest pokrewne twierdzeniu Schura z ostatniego punktu, lecz różni się od niego w dwóch istotnych aspektach. Pokazuje ono, że jeśli macierz operatora w pewnej bazie można uczynić trójkątną, to można ją też uczynić bardzo „bliską" diagonalnej. Jednak nawet gdy $\mathbb{F} = \mathbb{C}$, nie można do tego użyć bazy ortonormalnej, jak w twierdzeniu Schura. Z tego względu twierdzenie Jordana gra mniejszą rolę w obliczeniach numerycznych, lecz ma wielkie znaczenie teoretyczne.

1. Sformułowanie twierdzenia Jordana i jednoznaczność postaci Jordana.

Klatką Jordana, odpowiadającą wartości λ, nazywamy macierz kwadratową, której wyrazy na przekątnej są równe λ, wyrazy bezpośrednio pod nią równe 1, a pozostałe 0.[2] Gdy klatka jest rozmiaru $s \times s$, to oznaczamy ją przez $\mathbf{J}(\lambda, s)$:

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

[2] Będziemy tu używać takich własnie **dolnych** klatek Jordana. Często używane są i **górne** klatki Jordana, w których jedynki występują nad przekątną, a nie pod nią.

Macierzą Jordana nazywamy macierz postaci $\mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_t)$, 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 jest przestrzenią wektorową nad ciałem $\mathbb{F}$ i $0 < \dim V < \infty$. 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.*

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.

Jeśli twierdzenie i uwagę 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

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 3 i 4; wcześniej 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* $\mathrm{rk}(\mathbf{K} - \lambda\mathbf{I})^{n-1} - \mathrm{rk}(\mathbf{K} - \lambda\mathbf{I})^n = 1$ *jeśli $\lambda = \mu$ i $n \leq s$, a w przeciwnym razie* $\mathrm{rk}(\mathbf{K} - \lambda\mathbf{I})^{n-1} - \mathrm{rk}(\mathbf{K} - \lambda\mathbf{I})^n = 0$.

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 $\mathbf{J}$ będzie postacią Jordana 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 są stopnia* $\geq n$ *(odp. 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 klatkami Jordana macierzy $\mathbf{J}$ będą $\mathbf{K}_1, ..., \mathbf{K}_t$. Dla $m = n$ i $m = n-1$, macierzą operatora $(L-\lambda J)^m$ w pewnej bazie jest więc $(\mathbf{J}-\lambda\mathbf{I})^m = \mathrm{diag}((\mathbf{K}_1-\lambda\mathbf{I})^m, ..., (\mathbf{K}_t-\lambda\mathbf{I})^m)$, wobec czego

$$\mathrm{rk}(L-\lambda I)^m = \mathrm{rk}(\mathbf{J}-\lambda\mathbf{I})^m = \textstyle\sum_{i=1}^t \mathrm{rk}(\mathbf{K}_i-\lambda\mathbf{I})^m.$$

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, w liczbie $p_n(\lambda)$. To kończy dowód, bo ponadto $q_n(\lambda) = p_n(\lambda) - p_{n+1}(\lambda)$. □

Wniosek 2. *Macierz Jordana operatora* L*, czy 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 $q_n(\lambda)$ $(\lambda \in \mathbb{F}, n \in \mathbb{N})$. □

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 $V_L(\lambda)$;

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 $\inf\{n : \mathrm{rk}(L-\lambda\mathbf{I})^n = \mathrm{rk}(L-\lambda\mathbf{I})^{n+1}\}$

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 były one 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* $\chi_{\mathbf{A}} = \chi_{\mathbf{B}}$*. Tak samo jest dla operatorów.*

Dowód. 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.

Odwrotnie, niech ostatni warunek będzie 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 końcowej części przykładu 1 w §1.3.) W ślad za nimi, podobne są macierze $\mathbf{A}$ i $\mathbf{B}$. □

Wniosek 4. *Dla $\mathbb{F} = \mathbb{R}, \mathbb{C}$, każda macierz $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$ jest nad $\mathbb{F}$ podobna do $\mathbf{A}^t$.*

Dowód. Wynika to z wniosku 3 (wraz z twierdzeniem 1 w §1.6, gdy $\mathbb{F} = \mathbb{R}$). □

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

a) Operator L wtedy i tylko wtedy jest diagonalizowalny, gdy $\mathrm{rk}(L-\lambda I) = \mathrm{rk}(L-\lambda I)^2$, 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 6a) w p.3.)

Zadania uzupełniające.

1. Niech macierz operatora L w bazie $(\mathbf{b}_i)_{i=1}^k$ będzie równa $\mathrm{diag}(\mathbf{J}(0,n_1),...,\mathbf{J}(0,n_s))$, gdzie $n_1 \leq ... \leq n_s$. Jak z tej bazy utworzyć bazę przestrzeni a) $\mathrm{im}(L^n)$, i b) $\ker(L^n)$, dla $n = 1,...,k$?

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

3. Znana jest macierz Jordana operatora L; wyznaczyć ją dla zawężenia $L_{|\mathrm{im}(L)}$.

4. 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 $\mathbb{F}$.

5. 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 $\mathbf{K}_i$ $(i = 1,...,s)$.

6. 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.

Zadania ze zbioru Kostrykina: 33 w §II.3.2 i 1,2,4,5,6,9,15* i 27 w §II.3.3.

2. Funkcje macierzy niediagonalizowalnych.

Rozszerzymy teraz wyniki z §2.3; dla uproszczenia przymujemy do końca tego punktu, że $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Potrzebny jest

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:*

$$p(\mathbf{X}_0 + \mathbf{X}) = \sum_{n=0}^{\infty} \frac{1}{n!}\left((D^n p)(\mathbf{X}_0)\right)\mathbf{X}^n \qquad (tu,\ D^n p = 0 \ dla\ n > \deg(p)). \tag{9}$$

Dowód. Zbiór W tych wielomianów p, dla których równość (9) ma miejsce, jest zamknięty względem dodawania wielomianów i mnożenia ich przez skalar. A że jednomiany $1, x, x^2, \ldots$ rozpinają liniowo $\mathbb{F}[x]$, to pozostaje dowieść, że należą one do W. Jednak gdy $p = x^l$, to $\frac{1}{n!}D^n p = \binom{l}{n}x^{l-n}$ dla $n \leq l$, zaś z przemienności $\mathbf{X}_0$ i $\mathbf{X}$ otrzymujemy łatwo $p(\mathbf{X}_0 + \mathbf{X}) = \sum_{n=0}^{l}\binom{l}{n}\mathbf{X}_0^{l-n}\mathbf{X}^n$ – co daje (9). □

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 $p(\lambda), (Dp)(\lambda), \ldots, (D^{s-1}p)(\lambda)$.*

Dowód. Wynika to ze wzoru (9), przy $\mathbf{X}_0 = \lambda\mathbf{I}$, $\mathbf{X} = \mathbf{J}(0, s)$, i zadania 1 w p.1. □

Uwaga 1. Niech $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathrm{diag}(\mathbf{J}_1, \ldots, \mathbf{J}_t)$, gdzie $\mathbf{J}_i$ są klatkami Jordana. Wtedy $p(\mathbf{A}) = \mathbf{S}\mathrm{diag}(p(\mathbf{J}_1), \ldots, p(\mathbf{J}_t))\mathbf{S}^{-1}$, 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 $(D^j p)(\lambda)$, 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.1 wynika, że $m_\lambda = \inf\{n : \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^n = \mathrm{rk}(\mathbf{A} - \lambda\mathbf{I})^{n+1}\}$.)

Uwaga 2. Przyjmijmy oznaczenia uwagi 1, 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 $f(\mathbf{A}) = \mathbf{S}\mathrm{diag}\,(f(\mathbf{J}_1), ..., f(\mathbf{J}_t))\,\mathbf{S}^{-1}$.

Wynik nie zależny od użytej macierzy $\mathbf{S}$, bo $f(\mathbf{A}) = p(\mathbf{A})$ dla dowolnego wielomianu p takiego, że $D^j p(\lambda) = D^j f(\lambda)$ dla $\lambda \in \mathrm{spec}(\mathbf{A})$ i $0 \le j < m_\lambda$. (Wielomiany takie istnieją, nawet stopnia $< k$, por. zad. uz. 4 w §III.5.1.[3]) Jak w części g) zadania 1 przekonujemy się, że gdy $(f_n)_{n=0}^{\infty}$ jest ciągiem funkcji **A**–dopuszczalnych i $D^i f_n(\lambda) \to D^i f_0(\lambda)$ dla $\lambda \in \mathrm{spec}(\mathbf{A})$ i $0 \le i < m_\lambda$, to $f_n(\mathbf{A}) \to f_0(\mathbf{A})$.

Przykład 1. Macierz $e^{\mathbf{A}}$ obliczona zgodnie z tą definicją jest zarazem granicą ciągu $\mathbf{I} + \mathbf{A} + ... + \frac{1}{n!}\mathbf{A}^n$, a także ciągu $(\mathbf{I} + \frac{1}{n}\mathbf{A})^n$. (Dlaczego?)

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

Zadanie 1. Udowodnić, że części od a) do f) zadania 1 w §2.3 pozostają słuszne dla **A**–dopuszczalnych funkcji f, g i $f(\mathbf{A})$–dopuszczalnej funkcji h. (Wskazówka: dla pewnych $p, q \in \mathbb{F}[x]$ zachodzi $p(\mathbf{A}) = f(\mathbf{A})$ i $q(\mathbf{A}) = g(\mathbf{A})$; porównać pochodne fg i pq korzystając z formuły Leibniza na pochodną iloczynu.)

Uwaga 3. Przy $f(\lambda) := \exp(s\lambda)$, $g(\lambda) := \exp(t\lambda)$ otrzymujemy stąd ważną własność funkcji wykładniczej macierzy: $e^{s\mathbf{A}} e^{t\mathbf{A}} = e^{(s+t)\mathbf{A}}$ dla $s, t \in \mathbb{R}$ i $\mathbf{A} \in \mathcal{M}_k(\mathbb{R})$. (Tak samo jest dla macierzy i skalarów zespolonych; na ogół jednak $e^{\mathbf{A}+\mathbf{B}} \neq e^{\mathbf{A}} e^{\mathbf{B}}$.)

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

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

2. Niech wielomian $\chi_{\mathbf{A}}$ 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 $p \in \mathbb{F}[x]$ mamy $\chi_{p(\mathbf{A})} = (p(\lambda_1) - x)...(p(\lambda_k) - x)$; w szczególności, $\mathrm{tr}(p(\mathbf{A})) = \sum_i p(\lambda_i)$ i $\det(p(\mathbf{A})) = \prod_i p(\lambda_i)$.

   b) Przy $\mathbb{F} = \mathbb{R}$ wywnioskować stąd **tożsamość Liouville'a**: $\det(e^{\mathbf{A}}) = e^{\mathrm{tr}\mathbf{A}}$.

   c) Wywnioskować też, że jeśli $\mathbb{F} = \mathbb{C}$ i $\mathbf{B} \in \mathcal{M}_l(\mathbb{C})$, to wielomiany $\chi_{\mathbf{A}}$ i $\chi_{\mathbf{B}}$ nie mają wspólnych pierwiastków wtedy i tylko wtedy, gdy $\det(\chi_{\mathbf{A}}(\mathbf{B})) \neq 0$.

3. a) Dowieść, że macierz zespolona, której wszystkie wartości własne są dodatnie, jest

[3] Brzmi ono: zadajmy na wielomiany $p \in \mathbb{F}[x]$ warunki postaci $D^0 p(a_i) = b_{i,0}, Dp(a_i) = b_{i,1}, ..., D^{m_i} p(a_i) = b_{i,m_i}$, dla skończenie wielu skalarów b_{ij} i a_i $(a_i \neq a_j)$ i liczb $m_i \in \mathbb{N} \cup \{0\}$. Dowieść, że jeśli warunków tych jest łącznie k, to spełnia je dokładnie jeden wielomian stopnia $\le k - 1$.

kwadratem pewnej macierzy o tejże własności.

b)* Dowieść, że można wyżej zastąpić „pewnej" przez „jedynej". (Wskazówka: przykład 1 w §3.4, z $\mathcal{H}$ zastąpionym przez zbiór macierzy o badanej własności.)

4. Niech $p \in \mathbb{F}[x]$ i $\mathbf{A} \in \mathcal{M}_k(\mathbb{F})$, przy czym wielomian $\chi_{\mathbf{A}}$ rozkłada się nad $\mathbb{F}$ na czynniki liniowe i $n_{\mathbb{F}} \neq 0_{\mathbb{F}}$ $\forall n$. Dowieść równoważności warunków:

a) $p(\mathbf{A}) = \mathbf{0}$;

b) $(D^j p)(\lambda) = 0$ dla $\lambda \in \operatorname{spec}(\mathbf{A})$ i $0 \leq j \leq m_\lambda - 1$ (liczby m_λ są jak w uwadze 2);

c) wielomian p jest podzielny przez $q := \prod_{\lambda \in \operatorname{spec}(\mathbf{A})} (x - \lambda)^{m_\lambda}$

Uwaga 4. Wielomian q z c) 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.)

5. a) Udowodnić część c) uwagi 3 z p.1.

b) Udowodnić, że jeśli $L^n = I_V$ i $x^n - 1$ rozkłada się nad $\mathbb{F}$ na czynniki liniowe, to operator L jest diagonalizowalny. (O $\mathbb{F}$ zakładamy , że $j_{\mathbb{F}} \neq 0$ dla $j = 2, 3, ..., n$.)

6. Udowodnić, że gdy $\mathbf{A} \in \mathcal{M}_k(\mathbb{C})$ i $\operatorname{tr}(\mathbf{A}^n) = 0$ dla $n = 1, ..., k$, to $\mathbf{A}^k = 0$.

7. a) Niech $\mathbf{J}$ będzie macierzą Jordana, a $\mathbf{B}$ – macierzą diagonalną z tą samą przekątną, co $\mathbf{J}$. Dowieść, że $\mathbf{B}$ jest wielomianem macierzy $\mathbf{J}$, tzn. $\mathbf{B} = \sum_{s=1}^{n} c_s \mathbf{J}^s$ dla pewnych skalarów $c_0, ..., c_n$. (Jeśli wygodnie, dodać założenie o $\mathbb{F}$, jak w zad. uz. 6b).)

b) Udowodnić, że każda macierz zespolona jest sumą dwóch macierzy, będących jej wielomianami, z których jedna jest diagonalizowalna, a druga nilpotentna. (Wskazówka: wpierw rozpatrzeć przypadek macierzy Jordana.)

c)* Udowodnić, że opisany w b) **rozkład Jordana** jest jednoznaczny. (Wskazówka: gdy $\mathbf{D} + \mathbf{N} = \mathbf{D}' + \mathbf{N}'$ są dwoma takimi rozkładami, to macierz $\mathbf{D} - \mathbf{D}'$ jest diagonalizowalna, patrz zad. uz. 4 w następnym punkcie.)

Zadania ze zbioru Kostrykina: 3, 17, 18, 19, 20, 43, 45, 46 w §II.3.3.

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

Przypominamy, że wszystkie rozpatrywane obecnie przestrzenie są skończonego wymiaru. Niech $L \in \mathcal{L}(V)$.

Definicja. a) Podprzestrzeń V_0 przestrzeni V jest L–**niezmiennicza**, 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 lub **zawężeniem** L do 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. Podprzestrzenie własne operatora L są L–niezmiennicze. Gdy L jest obrotem przestrzeni euklidesowej E^3, to oś obrotu i płaszczyzna ortogonalna do niej są L–niezmiennicze. Operatory indukowane to homotetie, identyczność i obrót płaszczyzny, odpowiednio. $\boxminus$

Uwaga 1. a) Gdy baza $\mathcal{V}$ przestrzeni V jest postaci $(\mathbf{u}_1, ..., \mathbf{u}_p, \mathbf{w}_1, ..., \mathbf{w}_q)$, gdzie $\mathcal{U} = (\mathbf{u}_1, ..., \mathbf{u}_p)$ rozpina L–niezmienniczą podprzestrzeń U, 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 wyżej $\mathcal{W} = (\mathbf{w}_1, ..., \mathbf{w}_q)$ też rozpina L–niezmienniczą podprzestrzeń W, 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.

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 $k = \dim V$.*
Dodatek: *Można wyżej 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 $n \leq k$. Podprzestrzeń $W := L^n(V)$ spełnia więc warunek $W = L^{n+1}(V) = L^{n+2}(V) = ...$, wobec czego $W = L^k(V)$ i $L^k(W) = W$. Z ostatniej równości wynika, że $W \cap U = \{\mathbf{0}\}$ dla $U := \ker(L^k)$; a że ponadto $\dim\ker(L^k) + \dim \operatorname{im}(L^k) = \dim V$, to $V = U \oplus W$. (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. $\square$

Zadanie 1. a) Podprzestrzeń L–niezmiennicza jest i $p(L)$–niezmiennicza, dla $p \in \mathbb{F}[x]$.

b) Podprzestrzeń zawarta w $\ker(L)$ lub zawierająca $\operatorname{im}(L)$ jest L-niezmiennicza.

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 $\ker(p(K))$ i $\operatorname{im}(p(K))$ dla $p \in \mathbb{F}[x]$, w tym przestrzenie własne operatora K.

Zadania uzupełniające.

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.

c)* Gdy operatory $L' \in \mathcal{L}(V')$ i $S \in \mathcal{L}(V, V')$ są takie, że $L'S = SL$ i $S(V) = V'$, to podprzestrzeń $\ker(S)$ jest L–niezmiennicza i $\chi_L = \chi_{L'} \cdot \chi_{L|\ker(S)}$.

2. Niech $L \in \mathcal{L}(V)$ będzie operatorem rzędu 1 i niech $\mathbf{v} \in L(V) \setminus \{\mathbf{0}\}$.

a) Dowieść, że $\chi_L = (-x)^{k-1}(-x+\lambda)$ dla pewnych $k \in \mathbb{N}$ i $\lambda \in \mathbb{F}$, przy czym $\lambda = 0 \Leftrightarrow L(\mathbf{v}) = \mathbf{0}$.

b) W zależności od λ znaleźć macierz Jordana dla L.

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

4. 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 $W' \subset V$.

c) Wywnioskować też, że jeśli L ma $k = \dim(V)$ różnych wartości własnych, to ma dokładnie 2^k podprzestrzeni niezmienniczych.

5. 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. (Wskazówka: ustalić operator $L_{i_0} \neq I_V$ zauważyć, że jego podprzestrzenie własne są L_i–niezmiennicze, $\forall i$.)

6. a) To samo co w zad. 5, ale dla ortonormalnej diagonalizowalności.

b) Uogólnić podobnie twierdzenie Schura. (Wskazówka: wykazać wpierw istnienie wspólnego wektora własnego rozważanych operatorów.)

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

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

c) Czy zawsze $q_n(\lambda) \ge q'_n(\lambda)$?

8. 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 $\pm(x^{s+1} - \sum_{i=0}^{s} c_i x^i)$ i spełnia warunek $p(L)(\mathbf{v}) = \mathbf{0}$.

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

9. a) Udowodnić, że każda kwadratowa macierz rzeczywista ma podprzestrzeń niezmienniczą wymiaru 1 lub 2. (Wskazówka: zadanie 1 w §3.4.)

b) Uogólnić a) na przemienną rodzinę macierzy rzeczywistych.

c) Wywnioskować, że operator na k-wymiarowej przestrzeni nad $\mathbb{F}$ ma podprzestrzeń niezmienniczą wymiaru $k-1$ jeśli $\mathbb{F}=\mathbb{C}$, zaś wymiaru 1 lub 2, oraz $k-1$ lub $k-2$, jeśli $\mathbb{F}=\mathbb{R}$. Uogólnić na przemienne rodziny operatorów. (Wskazówka: zad. 2 w §V.3.3.)

d)* Udowodnić, że jeśli $k \times k$–macierz rzeczywista ma podprzestrzeń niezmienniczą wymiaru n, to ma i podprzestrzeń niezmienniczą wymiaru $k-n$.

10. Wzmocnić tezę zadania uz. 3 następująco:

a) Niech rozkład $V = U \oplus W$ i operator $L \in \mathcal{L}(V)$ spełniają warunki $U = \ker(L^k)$ i $L(W) = W$. Wówczas dla każdej L–niezmienniczej podprzestrzeni $X \subset V$ zachodzi równość $X = U \cap X + W \cap X$. (Wskazówka: przeciwobraz $W \cap X$ przy $(L_{|W})^k$ jest zawarty w $W \cap X$.)

b) Wywnioskować, że jeśli dla operatora L istnieje baza Jordana, to każda podprzestrzeń L–niemiennicza jest sumą prostą swych części wspólnych z podprzestrzeniami L–pierwiastkowymi.

11. Niech L, L_i $(i \in I)$ będą endomorfizmami przestrzeni V, przy czym $LL_i = L_iL$ $\forall i$. Zakładamy, że nie istnieje nietrywialna podprzestrzeń, L_i–niezmiennicza dla wszystkich operatorów L_i. Udowodnić, że $L = 0$ lub L jest izomorfizmem, a jeśli ciałem skalarów jest $\mathbb{C}$, to L jest mnożeniem przez skalar. (Jest to **lemat I. Schura**.)

Zadania ze zbioru Kostrykina: 22,23,26,27,29,30,31*,32,36*,37* w §II.3.2 i 13 w §II.3.3.

4. Dowód twierdzenia Jordana

Potrzebne będą definicja i lemat.

Definicja. **Macierz** kwadratowa $\mathbf{B}$ jest **nilpotentna**, jeśli $\mathbf{B}^n = \mathbf{0}$ dla pewnego n; 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 $0 < \dim V < \infty$, 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.1), przyjmując lemat za prawdziwy.

Dowód twierdzenia Gdy $\dim V = 0$, to nie ma czego dowodzić. Wykorzystamy więc indukcję względem liczby $k = \dim V$, dzieląc krok indukcyjny na 2 części.

a) Załóżmy dodatkowo, że $\det(L) = 0$. Na podstawie stwierdzenia z p.3 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 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 $s \in \mathbb{N}$. $\square$

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

$$L(A) \subset A \cup \{\mathbf{0}\} \text{ i dla } \mathbf{v}, \mathbf{w} \in A \text{ jeśli } L(\mathbf{v}) = L(\mathbf{w}) \neq \mathbf{0}, \text{ to } \mathbf{v} = \mathbf{w}. \tag{10}$$

Wyjaśnijmy wpierw sens tych warunków. Każdy wektor $\mathbf{a} \in A \setminus L(A)$ wyznacza ciąg $\mathcal{V}_\mathbf{a} = (\mathbf{a}, L(\mathbf{a}), ..., L^d(\mathbf{a}))$, gdzie $d := \sup\{s : L^s(\mathbf{a}) \neq \mathbf{0}\} < \infty$ (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 składa się on z wektorów zbioru A. Warunek (10) powoduje, że strumienie o różnych źródłach (zawsze z $A \setminus L(A)$) są rozłączne i że łącznie wyczerpują zbiór A. (Oto dowód: dla $\mathbf{a} \in A$ przyjmijmy $d(\mathbf{a}) = \max\{i : L^i(\mathbf{a}) \neq \mathbf{0}\}$ i niech $d_{\max} = \max\{d(\mathbf{a}) : \mathbf{a} \in A\}$; wówczas $\{\mathbf{a} \in A : d(\mathbf{a}) = d_{\max}\}$ jest zbiorem źródeł najdłuższych strumieni. W ten sam sposób kolejno wyznaczymy strumienie największej długości poza już znalezionymi, aż wyczerpiemy zbiór A.)

Do konstrukcji bazy A wykorzystamy indukcję względem liczby n takiej, że $L^n = 0$. Gdy $n = 1$, to $L = 0$ i A może być dowolną bazą. Dla $n > 1$ rozważamy operator $L_1 := L_{|L(V)}$, indukowany przez L na L–niezmienniczej podprzestrzeni $L(V)$. Wówczas $L_1^{n-1} = 0$, więc z założenia indukcyjnego istnieje baza B przestrzeni $L(V)$, spełniająca warunek (10) przy A zastąpionym przez B. Wykorzystując to, że $B \subset L(V)$, obierzmy dla każdego $\mathbf{b} \in B$ wektor $\widetilde{\mathbf{b}} \in L^{-1}(\mathbf{b})$ tak, by $\widetilde{\mathbf{b}} \in B$ gdy $\mathbf{b} \in L(B)$.

Rozszerzmy zbiór $B \cap \ker(L)$ pewnym zbiorem Z_0 do bazy Z przestrzeni $\ker(L)$. Z lematu 1 w §III.5.1 wynika,[4] 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

[4] Oto ten lemat: niech $L \in \mathcal{L}(V, W)$ i niech B będzie bazą przestrzeni $L(W)$, a Z–bazą $\ker(L)$. Dla każdego $\mathbf{b} \in B$ obierzmy wektor $\widetilde{\mathbf{b}} \in L^{-1}(\mathbf{b})$. Wówczas $Z \cup \{\widetilde{\mathbf{b}} : \ \mathbf{b} \in B\}$ jest bazą przestrzeni V. Dowód jest polecany jako zadanie.

kolejne wyrazy strumienia $\mathcal{V}_1$, jako następne – kolejne wyrazy strumienia $\mathcal{V}_2$ itd. Jest oczywiste, że $[L]_{\mathcal{V}} = \operatorname{diag}(\mathbf{J}(0, d_1+1), \mathbf{J}(0, d_2+1), ...)$. □

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

Zadania ze zbioru Kostrykina: 18, 20, 25, 26 w §I.4.3 i 8 (dla $\mathbb{F} = \mathbb{C}$) w §II.3.3.

5. 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.3. 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 $\dim V$ wystarczające mogą być metody ad hoc, oparte na twierdzeniu 2 w p.1 czy zadaniu 2 w p.1. 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 $i \leq n$ i znamy już zbiór $A_i \subset V_i$, 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 $\widetilde{\mathbf{v}} \in V_{i-1}$ taki, że $L(\widetilde{\mathbf{v}}) = \mathbf{v}$, 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 uzyskamy na koniec bazę Jordana A_0 dla operatora $L_{|U}$. (Należy ją jeszcze uporządkować.)

Istotnie, jeśli operator L jest nilpotentny, wynika to wprost z dowodu lematu 1 z p.4. Gdy nie jest, możemy wyznaczyć A_0 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 $V_i := L^i(V)$, jak wyżej, to wobec

$L = L_{|U} \oplus L_{|W}$, gdzie $L_{|W}$ jest izomorfizmem, otrzymujemy $V_i = U_i \oplus W$, jak również $L^{-1}(U_i) \cap V_{i-1} = U_{i-1}$ (w tym $\ker(L) \cap V_{i-1} = \ker(L) \cap U_{i-1}$). Podana w a) i b) recepta (bez zastępowania V_i przez U_i etc.) prowadzi więc siłą rzeczy do dołączania wektorów wyłącznie z U_{i-1} i da bazę Jordana A_{i-1} dla operatora $L_{|U_{i-1}}$, w tym dla $i = 1$ – bazę Jordana A_0 dla operatora $L_{|U}$. □

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; wtedy $V_i = \text{lin}(G_i)$. Za bazę przestrzeni V_2 przyjmijmy $A_2 := \{2\}$, po czym 2 przedstawmy jako kombinację wektorów z G_2. Można n.p. za piąty współczynnik kombinacji przyjąć 1, a za pozostałe 0; z tabeli odczytamy wtedy, że $2 = L(2x + 2)$. Następnie $\{2\}$ uzupełniamy do bazy przestrzeni $\ker(L) \cap V_1$. Jak widać z tabeli, przestrzenią $V_1 = \text{lin}(G_1)$ 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 $Z_0 = \{p_1, p_2, p_3\}$, gdzie $p_1 := x^2 - xy - xz + yz - 2x$, $p_2 := y^2 - xy - yz + xz$, $p_3 := z^2 - xz - yz + xy$. Baza A_0, po uporządkowaniu, jest więc taka: $(x^2, 2x + 2, 2; xy - y^2, x - y; xz - z^2, x - z; p_1; p_2; p_3)$, 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. (Rozmiary klatek można też ustalić bez znajdywania bazy Jordana, w oparciu o powyższą tabelę i tw. 2 z p.2.) $\square$

Przykład 2. Niech operator $L : \mathbb{F}^4 \to \mathbb{F}^4$ będzie zadany tak, by $L(\mathbf{e}_1) = \mathbf{e}_2, L(\mathbf{e}_2) = \mathbf{e}_3, L(\mathbf{e}_3) = \mathbf{0}$ i $L(\mathbf{e}_4) = \mathbf{e}_2$. Postępując jak wyżej stwierdzimy, że $V_1 = \mathrm{lin}(\mathbf{e}_2, \mathbf{e}_3)$ i $V_2 = \mathbb{F}\mathbf{e}_3$, co prowadzi do przyjęcia $\{\mathbf{e}_3\}$ jako bazy przestrzeni V_2; daje to ujście najdłuższego strumienia, który będzie następnie rozbudowywany (od ujścia) do strumienia $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Dołączenie elementów jądra nastąpi dopiero na na poziomie przestrzeni $V_0 = \mathbb{F}^4$, np. wektorem $\mathbf{e}_1 - \mathbf{e}_4$ (sprawdzić to!). $\square$

Zadania ze zbioru Kostrykina: 10 w §II.3.3.

6. ** 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 niżej), pod każdą z nich, prócz ostatniej –klatka jednostkowa $\mathbf{I}_2$, a poza nimi 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 $(\mathbf{e}_i)_{i=1}^{2s}$ standardową bazę 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.

§ 5. * 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 L jest wyznaczony przez swój wielomian charakterystyczny χ_L.*

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 wielomian χ_L.

By uzasadnić a) wystarczy zauważyć, że z dokładnością do podobieństwa, operator diagonalizowalny jest wyznaczony przez swą macierz w bazie diagonalizującej (patrz wniosek 1 w §1.3), a ta z dokładnością do podobieństwa jest wyznaczona przez wielomian χ_L; patrz uwaga 5 w §2.2 i przykład 1 w §1.3. By uzasadnić b) wystarczy ponadto uwzględnić twierdzenie 2 z §4.1, a dla c) uwzględnić twierdzenie Toeplitza z §3.2. Przytoczone wyniki i ich dowody pozwalają nawet dokładnie opisać postać macierzy „kanonicznego" operatora $\mathbb{F}^k \to \mathbb{F}^k$, podobnego do L; postać ta zależy tylko od wymienionych we wniosku niezmienników, wyznaczonych przez własności widma. Teoria spektralna dotyczy wyrażania własności operatorów przez takie niezmienniki. Badając macierze operatorów zaawansowaliśmy ją znacznie i większość rezultatów tego rozdziału do niej należy.

Dla przykładu, uogólnione twierdzenie Eulera z §3.3 pozornie dotyczy własności ortogonalnych przekształceń przestrzeni euklidesowej. Zauważmy jednak, że kąty występujących w nim obrotów płaszczyzn, jak również postać pozostałych operatorów, działających na prostej, są wyznaczone (pomijając kolejność) przez wartości własne operatora L i ich krotności. Z dokładnością do podobieństwa ortogonalnego, twierdzenie to opisuje więc jednoznacznie operator ortogonalny L przez jego widmo. Podobnie jest z opisem operatorów symetrycznych w twierdzeniu 1 z §3.3.

Z tych względów wyniki §§3.2 i 3.3, wzbogacone o powyższe uwagi, bywają nazywane „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. Dużą rolę odgrywa natomiast możliwość określenia funkcji operatora, którą w języku macierzy rozwinęliśmy w §§2.3, 3.4 i 4.2.

§ 6. Przewodnik.

skopiowany z wywieszki dla potoku „zwyczajnego".

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 1c),d) z §3.2.)

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 – 3; §2.2 i 2.3.)

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

3. (§3.1–3.3 i 3.5; fragmenty §3.4.)

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. Twierdzenie Schura z §3.5 – wersja dla macierzy i dla operatorów.

4. (§4.)

Sformułowanie twierdzenia Jordana (wersja dla macierzy i dla operatorów; równoważność tych wersji). Sposób wyznaczania postaci Jordana i jej jednoznaczność. Zastosowanie postaci Jordana do ewaluacji, na macierzy, wielomianów i funkcji bardziej ogólnych. Podprzestrzenie niezmiennicze i ich proste własności. Umiejętność znajdywania bazy Jordana dla małych wymiarów metodami „ad hoc" lub opisanymi w §4.5.

Uwaga 1. Pewne fragmenty dowodów były relegowane do „zadań" (nie chodzi o uzupełniające czy z gwiazdką). Należy je rozwiązywać, by sprawdzać rozumienie tego, o czym mowa. (Część była zresztą rozwiązana na wykładzie lub ćwiczeniach.)