

Algebra liniowa 3
2008–2009

Kazimierz Szymiczek

Spis treści

Przedmowa	5
1 Przestrzenie wektorowe	1
1.1 Podstawowe pojęcia	1
1.2 Homomorfizmy	3
1.3 Endomorfizmy i macierze	5
1.4 Moduły nad pierścieniami przemiennymi	6
2 Algebry	9
2.1 K-algebry	9
2.1.1 Wielomian minimalny	15
2.1.2 Elementy odwracalne i dzielniki zera	17
2.2 Endomorfizmy	18
2.2.1 Endomorfizmy odwracalne	18
2.2.2 Podobieństwo	21
2.3 Algebra endomorfizmów	24
2.3.1 Reprezentacja algebry w algebrze endomorfizmów	24
2.3.2 Centrum algebry endomorfizmów	24
2.3.3 Ideały algebry endomorfizmów	26
3 Podprzestrzenie niezmiennicze	29
3.1 Wartości własne endomorfizmu	29
3.2 Szczególne postaci macierzy endomorfizmu	33
3.3 Podejście klasyczne	34
4 Triangularyzacja i diagonalizacja	37
4.1 Endomorfizm indukowany	37
4.2 Triangularyzacja	38
4.3 Diagonalizacja	42
5 Twierdzenie o rozkładzie	45
5.1 Endomorfizmy redukowalne	45
5.2 Twierdzenie o rozkładzie	47
6 Endomorfizmy nilpotentne	51
6.1 Podprzestrzenie cykliczne	51
6.2 Endomorfizmy nilpotentne	52
6.3 Postać kanoniczna	57

7	Postać kanoniczna Jordana	63
7.1	Postać kanoniczna	63
7.2	Jednoznaczność postaci kanonicznej	66
7.3	Wielomian charakterystyczny	69
7.4	Wyznacznik i ślad	71
8	Postać kanoniczna wymierna	75
8.1	Podprzestrzenie cykliczne	75
8.2	Drugie twierdzenie o rozkładzie	78
8.3	Postać kanoniczna	82
9	Moduły nad pierścieniami ideałów głównych	83
9.1	Moduły - definicje i przykłady	83
9.1.1	Operacje na modułach	86
9.1.2	Homomorfizmy modułów	86
9.1.3	Moduły wolne	87
9.2	Moduły torsyjne	88
9.2.1	Moduły torsyjne i ograniczone	88
9.2.2	Składowe prymarne	89
9.2.3	Rozkład prymarny	90
9.2.4	Rozkład prymarny torsyjnych grup abelowych	91
9.2.5	Rozkład prymarny modułu V_τ	92
9.3	Moduły skończenie generowane	93
9.3.1	Struktura modułów skończenie generowanych	93
9.3.2	Struktura skończenie generowanych grup abelowych	96
9.3.3	Struktura $K[X]$ –modułu V_τ	96

Przedmowa

Linear algebra, like motherhood,
has become a sacred cow.
Irving Kaplansky

Jakkolwiek algebrę liniową wykłada się dla wszystkich specjalności matematycznych studiów uniwersyteckich, to jednak usytuowanie tych wykładów na pierwszych semestrach studiów nie pozwala na omówienie bardziej zaawansowanych tematów, które są kluczowe dla zastosowań algebry liniowej w matematyce i poza nią. W szczególności, centralny problem istnienia postaci kanonicznych endomorfizmów przestrzeni wektorowych jest zwykle zaledwie muśnięty wzmianką o diagonalizowalności endomorfizmów samosprężonych przestrzeni euklidesowych.

Niniejszy skrypt jest zapisem wykładu fakultatywnego z algebry liniowej w Uniwersytecie Śląskim w semestrze zimowym roku akademickiego 2008/2009. Koncentruje się on na postaciach kanonicznych macierzy endomorfizmów i przedstawia w miarę kompletnie teorię postaci diagonalnej, trójkątnej, Jordana i Frobeniusa. Akceptując nowe tendencje w wykładzie algebry liniowej staramy się nie korzystać z wyznaczników, nawet tam gdzie są one tradycyjnie uważane za niezbędne (wielomian charakterystyczny i pierwiastki charakterystyczne endomorfizmu).

Literatura

P. R. Halmos, *Finite-dimensional vector spaces*, 2nd edition, Van Nostrand, New York 1958.

I. N. Herstein, *Topics in algebra*, 2nd edition, John Wiley & Sons, New York 1975, Chapters 4 and 6.

S. Axler, *Linear algebra done right*, 2nd edition, Springer, 1997.

Niniejsza wersja wykładu będzie systematycznie rozszerzana i poprawiana. Nowa wersja rozdziału będzie zawierała datę ostatniej zmiany.

Kazimierz Szymiczek

Rozdział 1

Przestrzenie wektorowe

Ostatnie zmiany 9.10.2008 r.

W pierwszym rozdziale przypominamy podstawowe pojęcia algebry liniowej. Zakładamy, że Czytelnik zna te pojęcia i fakty i przytaczamy je jedynie dla ustalenia terminologii, oznaczeń oraz jako wygodne podręczne zestawienie definicji i twierdzeń.

1.1 Podstawowe pojęcia

Niech K będzie dowolnym ciałem i niech $(V, +, \theta)$ będzie addytywną grupą abelową. Grupę V nazywamy *przestrzenią wektorową* nad ciałem K lub K -przestrzenią wektorową, jeśli na grupie V określone jest *działanie zewnętrzne* z ciałem skalarów K :

$$K \times V \rightarrow V, \quad (a, v) \mapsto av$$

i ma ono następujące własności:

$$\begin{aligned} a(v_1 + v_2) &= av_1 + av_2 \\ (a_1 + a_2)v &= a_1v + a_2v \\ (a_1a_2)v &= a_1(a_2v) \\ 1v &= v \end{aligned}$$

dla wszystkich $a, a_1, a_2 \in K$, $v_1, v_2, v \in V$. Elementy przestrzeni V nazywamy *wektorami*. Mówimy, że wektory $v_1, \dots, v_m \in V$ są *liniowo zależne*, jeśli istnieją skalary $a_1, \dots, a_m \in K$ nie wszystkie równe zero takie, że

$$a_1v_1 + \dots + a_mv_m = \theta.$$

Wektory $v_1, \dots, v_m \in V$, które nie są liniowo zależne nazywają się *liniowo niezależne*. Mówimy, że podzbiór $B \subset V$ przestrzeni V jest *liniowo niezależny* jeśli każdy skończony podzbiór zbioru B jest liniowo niezależny.

Mówimy, że podzbiór $B \subset V$ przestrzeni V *rozpina* przestrzeń V jeśli każdy wektor $v \in V$ można przedstawić w postaci kombinacji liniowej wektorów zbioru B ,

$$v = a_1v_1 + \dots + a_mv_m$$

dla pewnych $a_1, \dots, a_m \in K$ oraz $v_1, \dots, v_m \in B$. Piszemy wtedy $V = \text{lin}(B)$.

Podzbiór B przestrzeni V nazywa się *bazą* przestrzeni V jeśli jest liniowo niezależny i rozpina przestrzeń V . Można udowodnić, że każda przestrzeń wektorowa ma bazę i ponadto, każde dwie bazy przestrzeni V są zbiorami równolicznymi. Wobec tego moc

jakiegokolwiek bazy przestrzeni V nazywa się *wymiarem* przestrzeni V i oznacza $\dim V$ lub $\dim_K V$, jeśli chcemy zaznaczyć, że traktujemy V jako przestrzeń wektorową nad ciałem K (a nie nad jakimś podciałem ciała K).

Podgrupę $(U, +, \theta)$ addytywnej grupy $(V, +, \theta)$ przestrzeni wektorowej V nad ciałem K nazywa się *podprzestrzenią* przestrzeni V jeśli U jest zbiorem zamkniętym ze względu na mnożenie przez skalary z ciała K , to znaczy dla każdych $a \in K$ oraz $u \in U$ także $au \in U$. W tej sytuacji U jest także przestrzenią wektorową nad ciałem K .

Jeśli U i W są podprzestrzeniami przestrzeni V to zbiór

$$U + W = \{u + w \in V : u \in U, w \in W\}$$

także jest podprzestrzenią przestrzeni V . Podprzestrzeń $U + W$ nazywamy *sumą* lub *sumą zwykłą* podprzestrzeni U i W . Zauważmy, że $V = U + W$ oznacza iż zbiór $U \cup W$ rozpina przestrzeń V .

Mówimy, że przestrzeń V jest *sumą prostą* podprzestrzeni U i W gdy

$$V = U + W \quad \text{oraz} \quad U \cap W = \{\theta\}.$$

Piszemy wtedy $V = U \oplus W$. Łatwo dowodzi się, że $V = U \oplus W$ wtedy i tylko wtedy gdy każdy wektor $v \in V$ ma *dokładnie jedno* przedstawienie w postaci $v = u + w$ gdzie $u \in U$ oraz $w \in W$.

Pojęcia sumy i sumy prostej dwóch podprzestrzeni są szczególnymi przypadkami ogólniejszych pojęć sumy i sumy prostej dowolnej (niekoniecznie skończonej) rodziny podprzestrzeni przestrzeni V . Niech $\{U_i : i \in I\}$ będzie dowolną rodziną podprzestrzeni przestrzeni wektorowej V . *Sumą* tej rodziny podprzestrzeni nazywamy zbiór wszystkich wektorów $v \in V$, które można przedstawić w postaci skończonych sum postaci

$$v = u_{i_1} + \cdots + u_{i_m}$$

gdzie $u_{i_j} \in U_{i_j}$ oraz $i_j \in I$. Zbiór ten jest podprzestrzenią przestrzeni V i oznaczamy go

$$\sum_{i \in I} U_i.$$

Mówimy, że przestrzeń V jest *sumą prostą* rodziny podprzestrzeni $\{U_i : i \in I\}$ jeśli

$$\sum_{i \in I} U_i = V \quad \text{oraz} \quad U_j \cap \sum_{i \in I, i \neq j} U_i = \{\theta\} \quad \text{dla każdego } j \in I.$$

Piszemy wtedy $V = \bigoplus_{i \in I} U_i$. Podobnie jak w przypadku dwóch podprzestrzeni, $V = \bigoplus_{i \in I} U_i$ wtedy i tylko wtedy gdy każdy wektor $v \in V$ ma *dokładnie jedno* przedstawienie w postaci $v = u_{i_1} + \cdots + u_{i_m}$ gdzie $u_{i_j} \in U_{i_j}$ oraz $i_j \in I$.

Niech U będzie podprzestrzenią przestrzeni V . Wtedy $(U, +, \theta)$ jest podgrupą addytywnej grupy abelowej $(V, +, \theta)$ i wobec tego można rozpatrywać grupę ilorazową V/U , której elementami są warstwy $v + U$ grupy V względem podgrupy (normalnej) U i dodawanie warstw jest określone następująco:

$$(v_1 + U) + (v_2 + U) = v_1 + v_2 + U.$$

Grupa ilorazowa V/U jest grupą abelową i można na niej określić działanie zewnętrzne

$$K \times V/U \rightarrow V/U, \quad (a, v + U) \mapsto av + U,$$

które wyposaża grupę abelową V/U w strukturę przestrzeni wektorowej nad ciałem K . Nazywamy ją *przestrzenią ilorazową* przestrzeni V względem podprzestrzeni U .

Jeśli $\dim V < \infty$, to znaczy jeśli V ma skończoną bazę, to wymiar przestrzeni ilorazowej V/U jest wyznaczony następująco:

$$\dim V/U = \dim V - \dim U.$$

Jeśli bowiem $\{v_1, \dots, v_k\}$ jest dowolną bazą podprzestrzeni U , to uzupełniamy ją do bazy $\{v_1, \dots, v_k, u_1, \dots, u_m\}$ przestrzeni V i łatwo dowodzimy, że warstwy $u_1 + U, \dots, u_m + U$ tworzą bazę przestrzeni ilorazowej V/U .

1.2 Homomorfizmy

Niech teraz V i W będą przestrzeniami wektorowymi nad ciałem K . *Homomorfizmem* przestrzeni V w przestrzeń W nazywamy odwzorowanie $\varphi : V \rightarrow W$, które jest homomorfizmem addytywnej grupy V w addytywną grupę W i zachowuje mnożenie wektorów przez skalary. A więc

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \quad \text{oraz} \quad \varphi(av) = a\varphi(v)$$

dla wszystkich $v_1, v_2, v \in V$ oraz $a \in K$. Homomorfizm $\varphi : V \rightarrow W$ nazywa się *monomorfizmem*, *epimorfizmem*, *izomorfizmem* jeśli φ jest odpowiednio odwzorowaniem iniektywnym, surjektywnym, bijektywnym.

Jądro homomorfizmu $\varphi : V \rightarrow W$ przestrzeni wektorowych nazywa się jądrem homomorfizmu $\varphi : V \rightarrow W$ grupy abelowej V w grupę abelową W . Jądro φ oznacza się $\ker \varphi$. A więc $\ker \varphi = \{v \in V : \varphi(v) = \theta_W\}$. Łatwo sprawdza się, że homomorfizm $\varphi : V \rightarrow W$ jest monomorfizmem wtedy i tylko wtedy gdy $\ker \varphi = 0$.

Jeśli homomorfizm $\varphi : V \rightarrow W$ jest izomorfizmem, to przestrzenie V i W nazywamy izomorficznymi i piszemy $V \cong W$. Łatwo sprawdza się, że izomorfizm $\varphi : V \rightarrow W$ przeprowadza bazę przestrzeni V na bazę przestrzeni W i wobec tego

$$V \cong W \quad \Rightarrow \quad \dim V = \dim W.$$

Jeśli znamy jakąś bazę przestrzeni V , to istnieje bardzo prosty i wygodny sposób określania homomorfizmów $\varphi : V \rightarrow W$ w dowolną przestrzeń wektorową W (nad tym samym ciałem). Jeśli, na przykład, przestrzeń V jest skończenie wymiarowa i ma bazę $\{v_1, \dots, v_n\}$ to dla każdego układu $\{w_1, \dots, w_n\}$ wektorów przestrzeni W istnieje dokładnie jeden homomorfizm $\varphi : V \rightarrow W$ taki, że $\varphi(v_i) = w_i$ dla $i = 1, \dots, n$. Homomorfizm ten działa na wektorach przestrzeni V (które są kombinacjami liniowymi wektorów bazowych) następująco:

$$\varphi(a_1v_1 + \dots + a_nv_n) = a_1w_1 + \dots + a_nw_n.$$

Zatem dla określenia homomorfizmu φ na przestrzeni wektorowej V wystarczy wskazać obrazy poprzez φ wektorów bazowych przestrzeni V . Wykorzystamy to spostrzeżenie dla dowodu następującego faktu:

$$\dim V = \dim W \quad \Rightarrow \quad V \cong W.$$

Jeśli $\mathcal{B}$ i $\mathcal{B}'$ są bazami V i W , odpowiednio, to $\dim V = \dim W$ oznacza, że istnieje bijekcja $f : \mathcal{B} \rightarrow \mathcal{B}'$. Niech $\varphi : V \rightarrow W$ będzie jedynym homomorfizmem, takim, że $\varphi(v) = f(v)$ dla każdego $v \in \mathcal{B}$. Wtedy φ jest izomorfizmem. Rzeczywiście, ponieważ f jest bijekcją, wszystkie wektory bazy $\mathcal{B}'$ przestrzeni W są w obrazie homomorfizmu φ , skąd wynika, że φ jest surjekcją. Z drugiej strony, jeśli $v \in \ker \varphi$, oraz $v = a_1v_1 + \dots + a_mv_m$, gdzie $v_i \in \mathcal{B}$ oraz $a_i \in K$, to z równości

$$0 = \varphi(v) = a_1f(v_1) + \dots + a_mf(v_m)$$

wynika, że $a_1 = \dots = a_m = 0$, gdyż $f(v_1), \dots, f(v_m)$ są liniowo niezależne jako wektory bazy $\mathcal{B}'$ przestrzeni W . Zatem $v = 0$, co oznacza, że $\ker \varphi = 0$ i wobec tego φ jest iniekcją.

Specjalną rolę odgrywa *homomorfizm kanoniczny* $\kappa_U : V \rightarrow V/U$ przestrzeni V w przestrzeni ilorazową V/U względem (dowolnej) podprzestrzeni U . Z definicji mamy $\kappa_U(v) = v + U$ dla każdego $v \in V$. Łatwo zauważyć, że $\ker \kappa_U = U$. Ważne twierdzenie o homomorfizmach przestrzeni wektorowych mówi, że jeśli $\varphi : V \rightarrow W$ jest epimorfizmem przestrzeni wektorowych, to

$$V/\ker \varphi \cong W.$$

Jako proste zastosowanie twierdzenia o homomorfizmach otrzymamy formułę dla wymiaru sumy skończenie wymiarowych podprzestrzeni U i W przestrzeni V :

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Ten rezultat wynika stąd, że epimorfizm

$$U \rightarrow (U + W)/W, \quad u \mapsto u + W$$

ma jądro $U \cap W$. Zatem $U/(U \cap W) \cong (U + W)/W$, skąd porównując wymiary izomorficznych przestrzeni otrzymujemy formułę dla $\dim(U + W)$.

Dla ustalonych przestrzeni wektorowych V, W nad ciałem K zbiór wszystkich homomorfizmów $\varphi : V \rightarrow W$ oznacza się symbolem $\mathbf{Hom}(V, W)$ lub $\mathbf{Hom}_K(V, W)$. Zbiór $\mathbf{Hom}(V, W)$ jest grupą abelową ze względu na zwykłe dodawanie funkcji i grupę tę wyposażamy w strukturę przestrzeni wektorowej nad ciałem K definiując mnożenie homomorfizmów przez skalary następująco: dla $\varphi \in \mathbf{Hom}(V, W)$ i dla $a \in K$ określamy $a\varphi : V \rightarrow W$ następująco:

$$(a\varphi)(v) = a\varphi(v) \quad \text{dla } v \in V.$$

Łatwo sprawdza się, że faktycznie $a\varphi \in \mathbf{Hom}(V, W)$. W ten sposób możemy rozpatrywać *przestrzeń homomorfizmów* $\mathbf{Hom}(V, W)$.

Jeśli $\{v_1, \dots, v_n\}$ oraz $\{w_1, \dots, w_m\}$ są bazami przestrzeni V i W , to definiujemy nm homomorfizmów $\tau_{ij} : V \rightarrow W$ takich, że

$$\tau_{ij}(v_k) = \delta_{jk} w_i$$

dla $1 \leq i \leq m$, $1 \leq j \leq n$. Dowodzi się, że tworzą one bazę przestrzeni $\mathbf{Hom}(V, W)$. Zatem dla przestrzeni skończenie wymiarowych mamy

$$\dim \mathbf{Hom}(V, W) = \dim V \cdot \dim W. \quad (1.1)$$

Jeśli $\varphi : W \rightarrow W'$ jest homomorfizmem, to dla każdego homomorfizmu $f : V \rightarrow W$ otrzymujemy homomorfizm przestrzeni wektorowych $\varphi \circ f : V \rightarrow W'$. W ten sposób definiuje się *homomorfizm indukowany*

$$\varphi_* : \mathbf{Hom}(V, W) \rightarrow \mathbf{Hom}(V, W'), \quad \varphi_*(f) = \varphi \circ f.$$

W języku teorii kategorii można powiedzieć, że przy ustalonej przestrzeni V , przyporządkowania

$$W \mapsto \mathbf{Hom}(V, W), \quad \varphi \mapsto \varphi_*$$

określają funktor kowariantny kategorii przestrzeni wektorowych w siebie. Dwie następujące własności homomorfizmu indukowanego są łatwe do udowodnienia.

STWIERDZENIE 1.2.1. (a) Dla każdego epimorfizmu $\varphi : W \rightarrow W'$ homomorfizm indukowany

$$\varphi_* : \text{Hom}(V, W) \rightarrow \text{Hom}(V, W')$$

jest także epimorfizmem.

(b) Jeśli $0 \rightarrow W \rightarrow W' \rightarrow W'' \rightarrow 0$ jest ciągiem dokładnym, to także

$$0 \rightarrow \text{Hom}(V, W) \rightarrow \text{Hom}(V, W') \rightarrow \text{Hom}(V, W'') \rightarrow 0$$

jest ciągiem dokładnym.

Dla przykładu udowodnimy pierwszą własność. Niech $g \in \text{Hom}(V, W')$ i niech $\{v_i\}$ będzie bazą przestrzeni V . Ponieważ φ jest epimorfizmem, więc istnieją $w_i \in W$ takie, że $\varphi(w_i) = g(v_i)$ dla każdego i . Obieramy teraz homomorfizm $f \in \text{Hom}(V, W)$ taki, że $f(v_i) = w_i$ dla każdego i . Wtedy $g = \varphi f = \varphi_*(f)$. A więc φ_* jest epimorfizmem.

1.3 Endomorfizmy i macierze

Homomorfizm $\tau : V \rightarrow V$ przestrzeni wektorowej V w siebie nazywa się *endomorfizmem* przestrzeni wektorowej V . Endomorfizm $\tau : V \rightarrow V$, który jest izomorfizmem nazywa się *automorfizmem* przestrzeni wektorowej V . Przestrzeń endomorfizmów $\text{Hom}(V, V)$ oznacza się $\text{End } V$ lub $\text{End}_K V$. Jeśli $\dim_K V = n < \infty$, to na podstawie (1.1) otrzymujemy $\dim_K \text{End}_K V = n^2$.

Przypomnijmy teraz fundamentalne pojęcie macierzy endomorfizmu względem bazy przestrzeni.

DEFINICJA 1.3.1. Niech $\mathcal{B} = \{v_1, \dots, v_n\}$ będzie uporządkowaną bazą przestrzeni wektorowej V i niech $\tau \in \text{End}_K V$. Każdy wektor $\tau(v_j)$ przedstawiamy jako kombinację liniową wektorów bazy $\mathcal{B}$:

$$\tau(v_j) = \sum_{i=1}^n b_{ij} v_i,$$

gdzie b_{ij} są jednoznacznie określonymi elementami ciała K .

Macierz $\mathbf{m}(\tau, \mathcal{B}) := [b_{ij}] \in M_n(K)$ nazywa się *macierzą endomorfizmu τ w bazie $\mathcal{B}$* .

Macierz $\mathbf{m}(\tau, \mathcal{B})$ można więc opisać jako macierz, której j -tą kolumnę tworzą współrzędne wektora $\tau(v_j)$ w bazie $\mathcal{B}$.

Niech $M_n(K)$ będzie zbiorem wszystkich $n \times n$ macierzy (macierzy o n kolumnach i n wierszach) o elementach z ciała K . Jak wiadomo z kursu algebry liniowej $M_n(K)$ jest przestrzenią wektorową nad ciałem K z działaniem zewnętrznym i dodawaniem macierzy określonymi następująco:

$$a \cdot [a_{ij}] := [aa_{ij}], \quad [a_{ij}] + [b_{ij}] := [a_{ij} + b_{ij}],$$

gdzie $a \in K$. Zbiór macierzy

$$\{M_{ij} \in M_n(K) : 1 \leq i, j \leq n\}$$

gdzie M_{ij} jest macierzą, która w i -tym wierszu i j -tej kolumnie ma 1 a na pozostałych miejscach zera, jest bazą algebry $M_n(K)$. Rzeczywiście, macierze te rozpinają przestrzeń $M_n(K)$, co wynika z równości

$$[a_{ij}] = \sum_{1 \leq i, j \leq n} a_{ij} M_{ij},$$

oraz są liniowo niezależne, co łatwo otrzymuje się przy pomocy tej samej równości. Zatem

$$\dim_K M_n(K) = n^2.$$

Okazuje się, że każda uporządkowana baza $\mathcal{B}$ przestrzeni V wyznacza przyporządkowanie

$$\mu : \text{End}_K V \longrightarrow M_n(K), \quad \mu(\tau) = m(\tau, \mathcal{B}),$$

które jest izomorfizmem przestrzeni wektorowych. Rutynowy argument pozwala bowiem stwierdzić, że odwzorowanie μ jest bijekcją, a ponadto dla dowolnych $\sigma, \tau \in \text{End}_K V$ oraz dowolnego $a \in K$ mamy

$$\begin{aligned} m(\sigma + \tau, \mathcal{B}) &= m(\sigma, \mathcal{B}) + m(\tau, \mathcal{B}), \\ m(a\tau, \mathcal{B}) &= a m(\tau, \mathcal{B}). \end{aligned}$$

Jeśli bowiem $\mathcal{B} = \{v_1, \dots, v_n\}$ jest bazą V i

$$\sigma(v_j) = \sum_{i=1}^n a_{ij}v_i, \quad \tau(v_j) = \sum_{i=1}^n b_{ij}v_i$$

to

$$(\sigma + \tau)(v_j) = \sum_{i=1}^n (a_{ij} + b_{ij})v_i,$$

oraz

$$a\tau(v_j) = \sum_{i=1}^n ab_{ij}v_i$$

co oznacza, że $m(\sigma + \tau, \mathcal{B}) = m(\sigma, \mathcal{B}) + m(\tau, \mathcal{B})$ oraz $m(a\tau, \mathcal{B}) = a m(\tau, \mathcal{B})$.

1.4 Moduły nad pierścieniami przemiennymi

Pojęcie przestrzeni wektorowej nad ciałem jest szczególnym przypadkiem ogólniejszego pojęcia modułu nad pierścieniem. Niech A będzie dowolnym pierścieniem przemiennym (z jedyneką $1 \in A$) i niech M będzie addytywną grupą abelową.

Grupę M nazywamy *modułem* nad pierścieniem A , lub krótko A -modułem, jeśli na grupie M określone jest *działanie zewnętrzne* z pierścieniem skalarów A :

$$A \times M \rightarrow M, \quad (a, m) \mapsto am$$

i ma ono następujące własności:

$$a(m_1 + m_2) = am_1 + am_2 \tag{1.2}$$

$$(a_1 + a_2)m = a_1m + a_2m \tag{1.3}$$

$$(a_1a_2)m = a_1(a_2m) \tag{1.4}$$

$$1m = m \tag{1.5}$$

dla wszystkich $a, a_1, a_2 \in A$, $m_1, m_2, m \in M$.

Jak widzimy, w przypadku gdy pierścień A jest ciałem, definicja ta redukuje się do definicji przestrzeni wektorowej. Niektóre pojęcia i fakty algebry liniowej przenoszą się bez żadnych zmian do teorii modułów. Na przykład, definicje liniowej niezależności wektorów, bazy i wymiaru przestrzeni, podprzestrzeni, sumy prostej, przestrzeni ilorazowej, homomorfizmu, przestrzeni homomorfizmów, homomorfizmu indukowanego, funkcjonału

liniowego, przestrzeni dualnej, przestrzeni bidualnej - mają swoje oczywiste odpowiedniki w teorii modułów.

Znacznie większej ostrożności wymaga korzystanie w teorii modułów z faktów analogicznych do tych znanych z teorii przestrzeni wektorowych. Na ogół fakty prawdziwe dla przestrzeni wektorowych przestają być prawdziwe w teorii modułów. A więc na przykład, wprowadzić można zdefiniować pojęcie bazy modułu, ale nie można udowodnić, że każdy moduł ma bazę (podczas gdy każda przestrzeń wektorowa ma bazę). Moduł, który ma bazę nazywa się modułem *wolnym*. Okazuje się, że każde dwie bazy modułu wolnego są równoliczne i wobec tego, podobnie jak dla przestrzeni wektorowych można określić wymiar, zwany tutaj rangą modułu wolnego, jako liczbę kardynalną dowolnej bazy modułu wolnego. Nawet jeśli moduł jest wolny, to może się zdarzyć, że ma podmoduły, które nie są modułami wolnymi. Jeśli podmoduły modułu wolnego są modułami wolnymi (na przykład ma to miejsce nad pierścieniami ideałów głównych), to moduły ilorazowe nie są na ogół modułami wolnymi. Wobec tego nie można przenieść na moduły wolne formuły dla wymiaru przestrzeni ilorazowej skończenie wymiarowej przestrzeni wektorowej.

Twierdzenie o homomorfizmach przestrzeni wektorowych ma automatyczne uogólnienie w teorii modułów. Natomiast odpowiedniki twierdzeń o przestrzeniach homomorfizmów, takie jak stwierdzenie 1.2.1 nie są już na ogół prawdziwe. Tym niemniej dla niektórych modułów M prawdziwe są własności:

(a) Dla każdego epimorfizmu $\varphi : N \rightarrow N'$ homomorfizm indukowany

$$\varphi_* : \text{Hom}(M, N) \rightarrow \text{Hom}(M, N')$$

jest także epimorfizmem.

(b) Jeśli $0 \rightarrow N \rightarrow N' \rightarrow N'' \rightarrow 0$ jest ciągiem dokładnym, to także

$$0 \rightarrow \text{Hom}(M, N) \rightarrow \text{Hom}(M, N') \rightarrow \text{Hom}(M, N'') \rightarrow 0$$

jest ciągiem dokładnym.

Przede wszystkim, dla przestrzeni wektorowych własności (a) i (b) można udowodnić posługując się bazami przestrzeni wektorowych. Te same dowody można powtórzyć w przypadku modułów wolnych. A więc każdy moduł wolny M ma własności (a) i (b). Dalej, można pokazać, że dla modułu M , własność (a) jest prawdziwa wtedy i tylko wtedy, gdy jest prawdziwa własność (b). Moduł M nazywa się *projektywny*, jeśli ma jedną (a więc obydwie) z własności (a), (b). Okazuje się, że klasa modułów projektywnych jest istotnie szersza niż klasa modułów wolnych. Można udowodnić, że moduł M jest projektywny wtedy i tylko wtedy gdy jest składnikiem prostym pewnego modułu wolnego, a więc gdy istnieje moduł wolny F i jego podmoduł Q takie, że

$$F = M \oplus Q.$$

Podobnie analizuje się przenoszenie innych własności przestrzeni wektorowych do teorii modułów. Zainteresowanego czytelnika odsyłamy do opracowania autora *Moduły projektywne*, dostępnego na stronie internetowej Instytutu Matematyki UŚ www.math.us.edu.pl w dziale Materiały dydaktyczne.

Rozdział 2

Algebry

Ostatnie zmiany 27.10.2008 r.

2.1 K-algebry

Przypomnimy najpierw podstawowe definicje teorii pierścieni. W rozdziale 1 posługiwaliśmy się już pojęciem pierścienia przemiennego, tutaj natomiast akcentujemy ogólniejszą definicję pierścienia nie zakładającą przemienności mnożenia w pierścieniu.

DEFINICJA 2.1.1. Zbiór P z dwoma działaniami $+$ i $\cdot$ zwanymi *dodawaniem* i *mnożeniem* oraz z dwoma wyróżnionymi elementami 0 i 1 zwanymi *zerem* i *jedynką* nazywa się *pierścieniem*, jeśli spełnione są następujące warunki:

1. $(P, +, 0)$ jest grupą abelową.
2. $(P, \cdot, 1)$ jest monoidem (półgrupą z jedynką).
3. Mnożenie jest rozdzielne względem dodawania, to znaczy

$$a(b + c) = ab + ac \quad \text{oraz} \quad (b + c)a = ba + ca$$

dla każdych $a, b, c \in P$.

Zwracamy uwagę, że każdy pierścień ma jedynkę. Ponadto, mnożenie w pierścieniu P musi być łączne ale może być *nieprzemienne*. Jeśli mnożenie w pierścieniu P jest przemienne, to znaczy jeśli $ab = ba$ dla każdych $a, b \in P$, to pierścień P nazywa się pierścieniem *przemennym*.

Podpierścieniem pierścienia P nazywamy podgrupę P_1 addytywnej grupy pierścienia P zawierającą jedynkę pierścienia P i zamkniętą ze względu na mnożenie. Łatwo sprawdzić, że P_1 jest wtedy także pierścieniem ze względu na działania dodawania i mnożenia będące zacieśnieniami odpowiednich działań w P .

Element a pierścienia P nazywa się *lewostronnie* odwracalny, jeśli istnieje $b \in P$ taki, że $ba = 1$. Element b nazywa się wtedy lewostronnie odwrotnym do elementu a . Podobnie, $a \in P$ jest *prawostronnie* odwracalny, jeśli istnieje $c \in P$ taki, że $ac = 1$. Element c jest wtedy prawostronnie odwrotny do a . W końcu, element $a \in P$ nazywa się *odwracalny*, jeśli jest równocześnie lewostronnie i prawostronnie odwracalny. Zauważmy, że jeśli element $a \in P$ jest odwracalny i $ba = 1$ oraz $ac = 1$, to

$$b = b \cdot ac = ba \cdot c = c.$$

A więc element odwracalny a ma tylko jeden element lewostronnie odwrotny, jak również tylko jeden element prawostronnie odwrotny i elementy te są równe (jedynemu) elementowi odwrotnemu do a . W związku z tą jednoznacznością elementu odwrotnego do a wprowadzamy dla niego oznaczenie a^{-1} .

Stwierdzamy z łatwością, że zbiór $U(P)$ wszystkich elementów odwracalnych pierścienia P tworzy grupę ze względu na mnożenie elementów. Nazywa się ją *grupą elementów odwracalnych* pierścienia P .

Pierścień P nazywa się *pierścieniem z dzieleniem*, jeśli każdy różny od zera element pierścienia P jest odwracalny. Przemienny pierścień z dzieleniem jest więc ciałem.

Element a pierścienia P nazywa się *lewostronnym dzielnikiem zera*, jeśli istnieje $b \in P$, $b \neq 0$, taki, że $ab = 0$. Podobnie, $a \in P$ jest *prawostronnym dzielnikiem zera*, jeśli istnieje $c \in P$, $c \neq 0$, taki, że $ca = 0$. Element $a \in P$ nazywa się *dzielnikiem zera* w P jeśli jest równocześnie lewostronnym i prawostronnym dzielnikiem zera.

Centrum $Z(P)$ pierścienia P nazywamy zbiór wszystkich elementów pierścienia P przemiennych z każdym elementem pierścienia P :

$$Z(P) := \{a \in P : ab = ba \ \forall b \in P\}.$$

Łatwo sprawdzić, że $Z(P)$ jest (przemiennym) podpierścieniem pierścienia P .

DEFINICJA 2.1.2. Niech A będzie przestrzenią wektorową nad ciałem K . Mówimy, że A jest K -algebrą (albo algebrą nad ciałem K) jeśli w A jest określone działanie binarne

$$A \times A \rightarrow A, \quad (a, b) \mapsto a \cdot b$$

zwane mnożeniem lub mnożeniem wewnętrznym w A i spełnione są następujące warunki:

- (I) A z dodawaniem wektorów i mnożeniem wewnętrznym jest pierścieniem.
- (II) Mnożenie wewnętrzne w A i mnożenie wektorów przez skalary spełniają następujący warunek:

$$x(a \cdot b) = xa \cdot b = a \cdot xb \quad \text{dla wszystkich } x \in K, a, b \in A.$$

Jeśli pierścień A jest przemienny, to K -algebrę A nazywamy K -algebrą przemienną.

Bazą K -algebry A nazywamy bazę przestrzeni wektorowej A nad K .

Wymiarem $\dim_K A$ algebry A nazywamy wymiar $\dim_K A$ przestrzeni wektorowej A nad ciałem K .

Jeśli A jest K -algebrą i $B \subseteq A$ jest podprzestrzenią wektorową przestrzeni A i równocześnie B jest podpierścieniem pierścienia A , to B także spełnia warunki (I) i (II) definicji 2.1.2 i wobec tego B jest także K -algebrą. Nazywamy ją *podalgebrą* algebry A .

Jeśli A jest K -algebrą i pierścień A jest pierścieniem z dzieleniem, to algebrę A nazywa się *algebrą z dzieleniem*.

Uwaga 2.1.1. Następująca analiza definicji K -algebry objaśnia rolę warunku (II) w definicji 2.1.2. Fakt, że A jest pierścieniem sprowadza się do spełnienia trzech warunków:

- (I₁) mnożenie wewnętrzne w A jest działaniem łącznym,
- (I₂) mnożenie wewnętrzne w A jest rozdzielne względem dodawania w A ,
- (I₃) istnieje element $1_A \in A$ taki, że $1_A \cdot a = a \cdot 1_A = a$ dla wszystkich $a \in A$.

Niech $\beta : A \times A \rightarrow A$ będzie mnożeniem wewnętrznym w K -algebrze A , to znaczy $\beta(a, b) = ab$ dla $a, b \in A$. Wtedy (I₂) i (II) można zapisać następująco:

$$\begin{aligned} \beta(a, b + c) &= \beta(a, b) + \beta(a, c), \\ \beta(b + c, a) &= \beta(b, a) + \beta(c, a), \\ x\beta(a, b) &= \beta(xa, b) = \beta(a, xb) \end{aligned}$$

dla wszystkich $a, b, c \in A$ oraz $x \in K$. Warunki te stwierdzają, że β jest operacją addytywną (warunek (I₂)) i jednorodną (warunek (II)) ze względu na każdą zmienną. Zatem β jest *odwzorowaniem dwuliniowym* przestrzeni wektorowej A w A .

Tak więc K -algebra A jest przestrzenią wektorową nad ciałem K w której jest określone dwuliniowe mnożenie wewnętrzne spełniające dodatkowo warunki (I₁) i (I₃).

Przykład 2.1.1. Niech K będzie ciałem i niech ciało L będzie rozszerzeniem ciała K . Wtedy L można traktować jako przestrzeń wektorową nad ciałem K (z mnożeniem zewnętrznym elementów L przez elementy ciała K określonym jako zacieśnienie mnożenia $L \times L \rightarrow L$ w ciele L do $K \times L \rightarrow L$) i ponadto, mnożenie w ciele L spełnia warunki (I) i (II) definicji 2.1.2. Zauważmy, że w tym przykładzie jednorodność mnożenia w L (warunek (II) definicji) wynika z przemienności i łączności mnożenia w L . A więc każde rozszerzenie L ciała K można traktować jako przemienną K -algebrę z dzieleniem. W szczególności, K jest 1-wymiarową K -algebrą. Wymiar $\dim_K L$ nazywa się także stopniem ciała L nad ciałem K .

Przykład 2.1.2. Niech $K[X]$ będzie pierścieniem przemiennym wielomianów jednej zmiennej X nad ciałem K . Oczywiście $K[X]$ jest także przestrzenią wektorową nad ciałem K z mnożeniem wielomianów przez skalary jako działaniem zewnętrznym. Ponadto mnożenie (wewnętrzne) wielomianów jest dwuliniowe. Zatem $K[X]$ jest K -algebrą przemienną oraz $\dim K[X] = \infty$. Nie jest to jednak algebra z dzieleniem, gdyż wielomiany stopnia ≥ 1 są nieodwracalnymi elementami pierścienia $K[X]$.

Przykład 2.1.3. Przypomnijmy teraz standardowy przykład K -algebry macierzy. Niech $M_n(K)$ będzie przestrzenią wektorową wszystkich $n \times n$ macierzy o elementach z ciała K . Wiadomo, że w $M_n(K)$ jest określone wewnętrzne działanie mnożenia macierzy

$$[a_{ij}] \cdot [b_{ij}] := [c_{ij}], \quad c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}.$$

To mnożenie jest łączne i dwuliniowe, ponadto macierz jednostkowa $I \in M_n(K)$ jest jedynką mnożenia: $I \cdot M = M \cdot I = M$ dla każdej macierzy $M \in M_n(K)$.

A więc $M_n(K)$ jest K -algebrą i jak stwierdziliśmy w rozdziale 1, $\dim_K M_n(K) = n^2$. Łatwo stwierdzić, że dla $n > 1$ algebra $M_n(K)$ jest nieprzemienna. Ponadto, dla $n > 1$ algebra $M_n(K)$ nie jest algebrą z dzieleniem (gdyż odwracalne są jedynie macierze nieosobliwe).

Zauważmy jeszcze, że definicje działań w algebrze macierzy $M_n(K)$ jak również dowody odpowiednich własności tych działań nie wykorzystują faktu, że mnożenie w K jest przemienne. Wykorzystuje się jedynie fakt, że K jest ciałem. Wobec tego, jeśli A jest dowolną K -algebrą, to te same definicje działań na macierzach określają strukturę K -algebry na zbiorze $M_n(A)$ wszystkich $n \times n$ macierzy o elementach w K -algebrze A . Łatwo stwierdzić, że

$$\dim_K M_n(A) = n^2 \dim_K A.$$

Przykład 2.1.4. Niech K będzie dowolnym ciałem i niech V będzie przestrzenią wektorową nad ciałem K . Rozpatrzmy zbiór $\text{End}_K V$ wszystkich endomorfizmów przestrzeni wektorowej V . Wyróżnimy *endomorfizm zerowy* $0_V : V \rightarrow V$ taki, że $0_V(u) = 0$ dla każdego $u \in V$, oraz *endomorfizm tożsamościowy* $1_V : V \rightarrow V$ taki, że $1_V(u) = u$ dla każdego $u \in V$.

Jak wiemy z rozdziału 1, $\text{End}_K V$ jest przestrzenią wektorową nad K z działaniami dodawania endomorfizmów i mnożenia endomorfizmów przez skalary określonymi następująco: dla $\sigma, \tau \in \text{End } V$

$$\sigma + \tau : V \rightarrow V, \quad (\sigma + \tau)(u) = \sigma(u) + \tau(u),$$

oraz dla $x \in K$ i $\tau \in \text{End } V$

$$x\tau : V \rightarrow V, \quad (x\tau)(u) = x\tau(u)$$

dla każdego $u \in V$. Ponadto określamy działanie *mnożenia* endomorfizmów, jako złożenie odwzorowań:

$$\sigma \cdot \tau : V \rightarrow V, \quad (\sigma \cdot \tau)(u) = \sigma(\tau(u)).$$

Rutynowe rachunki pokazują, że $\text{End}_K V$ z dodawaniem i mnożeniem endomorfizmów jako działaniami jest pierścieniem.

Należy jeszcze zauważyć, że mnożenie endomorfizmów w $\text{End } V$ oraz mnożenie endomorfizmów przez skalary są związane następującą własnością:

$$x(\sigma \cdot \tau) = x\sigma \cdot \tau = \sigma \cdot x\tau$$

dla dowolnych $\sigma, \tau \in \text{End } V$ oraz $x \in K$. W ten sposób pierścień $\text{End}_K V$ jest K -algebrą. Nazywamy ją *algebrą endomorfizmów przestrzeni wektorowej V* . Algebra endomorfizmów przestrzeni o wymiarze > 1 nie jest algebrą z dzieleniem (zob. przykład 2.1.5).

DEFINICJA 2.1.3. Niech A i B będą K -algebrami. *Homomorfizmem K -algebr A i B* nazywamy odwzorowanie $h : A \rightarrow B$ spełniające warunek

$$h(xa) = xh(a), \quad h(a + b) = h(a) + h(b), \quad h(a \cdot b) = h(a) \cdot h(b), \quad h(1_A) = 1_B$$

dla wszystkich $x \in K$, $a, b \in A$. *Izomorfizmem K -algebr* nazywamy bijektywny homomorfizm K -algebr.

A więc homomorfizm algebr jest równocześnie homomorfizmem przestrzeni wektorowych i homomorfizmem pierścieni.

Przykład 2.1.5. Niech V będzie n -wymiarową przestrzenią wektorową nad ciałem K . Wtedy mamy następujący izomorfizm K -algebr:

$$\text{End}_K V \cong M_n(K).$$

Jeśli $\mathcal{B} = \{v_1, \dots, v_n\}$ jest bazą przestrzeni V , to rozpatrujemy n^2 endomorfizmów $\tau_{ij} : V \rightarrow V$ takich, że

$$\tau_{ij}(v_k) = \delta_{jk}v_i$$

dla $1 \leq i, j \leq n$. W §1.1 zauważyliśmy, że tworzą one bazę przestrzeni

$$\text{Hom}(V, V) = \text{End}_K V.$$

Rutynowe rachunki pokazują, że homomorfizm przestrzeni wektorowych

$$\text{End}_K V \rightarrow M_n(K)$$

taki, że $\tau_{ij} \mapsto M_{ij}$, gdzie $M_{ij} \in M_n(K)$ jest macierzą, która w i -tym wierszu i j -tej kolumnie ma 1 a na pozostałych miejscach zera, jest izomorfizmem K -algebr.

Opiszemy jeszcze raz izomorfizm wskazany w przykładzie 2.1.5 używając pojęcia macierzy endomorfizmu względem bazy przestrzeni. Zauważmy, że w oznaczeniach przykładu 2.1.5 mamy

$$\mathbf{m}(\tau_{ij}, \mathcal{B}) = M_{ij}.$$

Posługując się pojęciem macierzy endomorfizmu względem bazy przestrzeni możemy izomorfizm algebr z przykładu 2.1.5 opisać następująco. Obieramy w przestrzeni wektorowej V bazę uporządkowaną $\mathcal{B}$ i każdemu endomorfizmowi τ przestrzeni V przyporządkowujemy macierz $\mathbf{m}(\tau, \mathcal{B})$ endomorfizmu τ względem bazy $\mathcal{B}$. Jak stwierdziliśmy w rozdziale 1, odwzorowanie

$$\mu : \text{End}_K V \longrightarrow M_n(K), \quad \mu(\tau) = \mathbf{m}(\tau, \mathcal{B})$$

jest izomorfizmem przestrzeni wektorowych. Ponadto dla dowolnych $\sigma, \tau \in \text{End}_K V$ mamy

$$\mathbf{m}(\sigma \cdot \tau, \mathcal{B}) = \mathbf{m}(\sigma, \mathcal{B}) \cdot \mathbf{m}(\tau, \mathcal{B}). \quad (2.1)$$

Jeśli bowiem $\mathcal{B} = \{v_1, \dots, v_n\}$ jest bazą V i

$$\sigma(v_j) = \sum_{i=1}^n a_{ij} v_i, \quad \tau(v_j) = \sum_{i=1}^n b_{ij} v_i$$

to

$$\begin{aligned} (\sigma \cdot \tau)(v_j) &= \sigma(\tau(v_j)) = \sigma\left(\sum_{i=1}^n b_{ij} v_i\right) = \sum_{i=1}^n b_{ij} \sigma(v_i) = \sum_{i=1}^n b_{ij} \sum_{k=1}^n a_{ki} v_k \\ &= \sum_{k=1}^n \left(\sum_{i=1}^n a_{ki} b_{ij}\right) v_k, \end{aligned}$$

skąd wynika, że element c_{kj} macierzy endomorfizmu $\sigma \cdot \tau$ w bazie $\mathcal{B}$ ma postać

$$c_{kj} = \sum_{i=1}^n a_{ki} b_{ij}.$$

Jest to więc element k -tego wiersza i j -tej kolumny iloczynu macierzy $[a_{ij}] \cdot [b_{ij}]$. Stąd otrzymujemy (2.1).

Izomorficzne algebry mają oczywiście równe wymiary, zatem jeszcze raz otrzymujemy

$$\dim_K V = n < \infty \quad \Rightarrow \quad \dim \text{End}_K V = \dim M_n(K) = n^2.$$

Odwzorowanie μ przeprowadza grupę elementów odwracalnych $U(\text{End}_K V)$ algebry endomorfizmów na grupę elementów odwracalnych $U(M_n(K))$. Elementy odwracalne w obydwu algebrach nazywa się także *nieosobliwymi*. Tradycyjne oznaczenia dla grup endomorfizmów nieosobliwych i macierzy nieosobliwych (stopnia n nad ciałem K) są następujące: $\text{Aut } V$ i $\mathbf{GL}(n, K)$. A więc $\mu(\text{Aut } V) = \mathbf{GL}(n, K)$.

Jak już zauważyliśmy, najprostszymi przykładami algebr z dzieleniem są rozszerzenia ciał: jeśli K jest podciałem ciała L , to L jest przemienną K -algebrą i oczywiście jest to algebra z dzieleniem. W ten sposób nad ciałem $\mathbb{R}$ liczb rzeczywistych otrzymujemy dwie przemienne $\mathbb{R}$ -algebry z dzieleniem: $\mathbb{R}$ i $\mathbb{C}$. Okazuje się, że nad $\mathbb{R}$ nie ma już innych przemiennych skończenie wymiarowych algebr z dzieleniem. Każda taka algebra jest ciałem i jest skończonym rozszerzeniem ciała $\mathbb{R}$, a więc jest to jedynie $\mathbb{R}$ lub $\mathbb{C}$. Można też udowodnić, że nad $\mathbb{R}$ nie istnieje 3-wymiarowa algebra z dzieleniem. Natomiast nad $\mathbb{R}$ istnieje 4-wymiarowa algebra z dzieleniem $\mathbb{H}$ zwana *algebrą kwaternionów Hamiltona*.

Konstrukcję algebry $\mathbb{H}$ można opisać następująco. W dowolnej 4-wymiarowej przestrzeni wektorowej nad $\mathbb{R}$ obieramy jakąkolwiek bazę i jej elementy oznaczamy $\mathbf{1}, i, j, k$. Definiując wewnętrzne mnożenie w $\mathbb{H}$ zakładamy, że będzie ono rozdzielne względem dodawania i wobec tego wystarczy jedynie wskazać reguły mnożenia elementów bazowych. Kluczowe są następujące definicje:

$$\mathbf{1} = 1_{\mathbb{H}}, \quad i^2 = -\mathbf{1}, \quad j^2 = -\mathbf{1}, \quad i \cdot j = k = -j \cdot i. \quad (2.2)$$

Jeśli, na przykład, chcemy znaleźć $k \cdot j$, to możemy postępować następująco:

$$k \cdot j = (i \cdot j) \cdot j = i \cdot j^2 = i \cdot (-\mathbf{1}) = -i.$$

W ten sposób używając tylko (2.2) sprawdzamy, że

$$j \cdot k = i = -k \cdot j, \quad k \cdot i = j = -i \cdot k, \quad k^2 = -\mathbf{1}.$$

Ponieważ zakładamy rozdzielnosć mnożenia względem dodawania, potrafimy teraz jednoznacznie obliczyć iloczyn dowolnych dwóch *kwaternionów* $(x_0\mathbf{1} + x_1i + x_2j + x_3k)(y_0\mathbf{1} + y_1i + y_2j + y_3k)$. W ten sposób 4-wymiarowa przestrzeń wektorowa $\mathbb{H}$ staje się algebrą nad ciałem $\mathbb{R}$ liczb rzeczywistych.

Istnieje też bardziej geometryczne podejście do definicji algebry kwaternionów Hamiltona. W 4-wymiarowej przestrzeni

$$\mathbb{H} = \mathbb{R}\mathbf{1} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k,$$

której wektory nazywamy kwaternionami, wyróżniamy dwie podprzestrzenie, 1-wymiarową podprzestrzeń $\mathbb{R}\mathbf{1}$ kwaternionów *skalnych* oraz 3-wymiarową podprzestrzeń

$$\mathbb{H}_0 := \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$$

kwaternionów *czystych*. Zauważmy, że $\mathbb{H} = \mathbb{R}\mathbf{1} \oplus \mathbb{H}_0$, zatem każdy kwaternion ma jednoznaczne przedstawienie w postaci sumy kwaternionu skalarnego i kwaternionu czystego. Mnożenie kwaternionów definiujemy w dwóch krokach. Dla kwaternionów $p = x_0\mathbf{1} + p_0$, $q = y_0\mathbf{1} + q_0$, gdzie $x_0, y_0 \in \mathbb{R}$ i p_0, q_0 są kwaternionami czystymi, pierwsza reguła jest następująca:

$$p \cdot q = (x_0\mathbf{1} + p_0) \cdot (y_0\mathbf{1} + q_0) = x_0y_0\mathbf{1} + x_0q_0 + y_0p_0 + p_0 \cdot q_0, \quad (2.3)$$

gdzie iloczyn $p_0 \cdot q_0$ kwaternionów czystych będzie określony drugą regułą mnożenia. Dla wprowadzenia drugiej reguły traktujemy przestrzeń $\mathbb{H}_0$ kwaternionów czystych jako przestrzeń euklidesową ze standardowym iloczynem skalarnym

$$(p_0, q_0) = x_1y_1 + x_2y_2 + x_3y_3$$

dla $p_0 = x_1i + x_2j + x_3k$, $q_0 = y_1i + y_2j + y_3k \in \mathbb{H}_0$. W przestrzeni euklidesowej $\mathbb{H}_0$ mamy także określony *iloczyn wektorowy*

$$[p_0, q_0] := \det \begin{bmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} = (x_2y_3 - x_3y_2)i + (x_3y_1 - x_1y_3)j + (x_1y_2 - x_2y_1)k$$

z pożyteczną interpretacją geometryczną (wektor $[p_0, q_0]$ jest prostopadły do obydwu wektorów p_0 i q_0 oraz jego długość jest równa polu równoległoboku rozpiętego na p_0, q_0). Przy tym jest rzeczą istotną, że iloczyn wektorowy $[\cdot, \cdot] : \mathbb{H}_0 \times \mathbb{H}_0 \rightarrow \mathbb{H}_0$ jest odwzorowaniem dwuliniowym. Wykorzystamy teraz iloczyn skalarny i iloczyn wektorowy p_0 i q_0 by zdefiniować iloczyn czystych kwaternionów:

$$p_0 \cdot q_0 := -(p_0, q_0)\mathbf{1} + [p_0, q_0]. \quad (2.4)$$

Wykorzystując reguły (2.3) i (2.4), określamy zatem iloczyn dowolnych kwaternionów $p, q \in \mathbb{H}$. Jest rzeczą oczywistą, że tak określone mnożenie jest dwuliniowe. Mniej trywialnym ćwiczeniem jest sprawdzenie łączności mnożenia, które pomijamy. Natomiast łatwo stwierdzić, że $\mathbf{1}$ jest jedyneką algebry $\mathbb{H}$. Rzeczywiście,

$$\mathbf{1} \cdot q = (\mathbf{1}\mathbf{1} + 0) \cdot (y_0\mathbf{1} + q_0) = y_0\mathbf{1} + q_0 = q$$

zgodnie z (2.3) i (2.4), i podobnie $q \cdot \mathbf{1} = q$. W ten sposób skonstruowaliśmy 4-wymiarową algebrę kwaternionów Hamiltona. Łatwo sprawdzić, że ta konstrukcja daje dokładnie tę samą algebrę, którą opisaliśmy wcześniej przy pomocy reguł mnożenia (2.2). Rzeczywiście,

$$i \cdot j = -(i, j)\mathbf{1} + [i, j] = [i, j] = k \quad \text{oraz} \quad j \cdot i = -(j, i)\mathbf{1} + [j, i] = [j, i] = -k,$$

$$i^2 = -(i, i)\mathbf{1} + [i, i] = -\mathbf{1}\mathbf{1} = -\mathbf{1} \quad \text{oraz} \quad j^2 = -(j, j)\mathbf{1} + [j, j] = -\mathbf{1}\mathbf{1} = -\mathbf{1}.$$

Niech teraz $p = x_0\mathbf{1} + p_0 \in \mathbb{H}$, gdzie $p_0 = x_1i + x_2j + x_3k \in \mathbb{H}_0$. Kwaternion $\bar{p} := x_0\mathbf{1} - p_0$ nazywa się kwaternionem *sprzężonym* z p , natomiast iloczyn $p \cdot \bar{p}$ nazywa się *normą* kwaternionu p i oznacza $N(p)$. Zatem

$$\begin{aligned} N(p) = p \cdot \bar{p} &= x_0^2\mathbf{1} + x_0(-p_0) + x_0p_0 - p_0^2 = x_0^2\mathbf{1} - p_0^2 \\ &= x_0^2\mathbf{1} + (p_0, p_0)\mathbf{1} - [p_0, p_0] = x_0^2\mathbf{1} + (p_0, p_0)\mathbf{1} \\ &= (x_0^2 + x_1^2 + x_2^2 + x_3^2)\mathbf{1}. \end{aligned}$$

Wykorzystując tę formułę dla normy kwaternionu udowodnimy teraz z łatwością następujący fakt.

STWIERDZENIE 2.1.1. $\mathbb{R}$ –*algebra kwaternionów* $\mathbb{H}$ *jest algebrą z dzieleniem.*

Dowód. Niech $p \in \mathbb{H}$ będzie niezerowym kwaternionem. Wtedy jego norma $p \cdot \bar{p} = x\mathbf{1}$ jest niezerowym kwaternionem skalarnym, a więc x jest różną od zera liczbą rzeczywistą (sumą kwadratów współrzędnych kwaternionu p). Wobec tego

$$p \cdot x^{-1}\bar{p} = \mathbf{1},$$

co dowodzi, że p jest odwracalny (oraz $p^{-1} = x^{-1}\bar{p}$). □

Rolę algebry kwaternionów Hamiltona podkreśla udowodnione w roku 1877 twierdzenie G. Frobeniusa mówiące, że nad ciałem $\mathbb{R}$ liczb rzeczywistych jedynymi skończenie wymiarowymi algebrami z dzieleniem są (z dokładnością do izomorfizmu) $\mathbb{R}$, $\mathbb{C}$ oraz $\mathbb{H}$.

2.1.1 Wielomian minimalny

Rozpatrzmy teraz ważny przykład homomorfizmu przemiennej K –algebry $K[X]$ wielomianów jednej zmiennej nad ciałem K w dowolną K –algebrę A . Homomorfizm ten sprowadza się do operacji podstawiania elementu algebry w miejsce zmiennej wielomianu. Wprawdzie w charakterze A dopuszczamy tu dowolną K –algebrę, ale najważniejszym dla nas szczególnym przypadkiem jest algebra $\text{End}_K V$ endomorfizmów przestrzeni wektorowej V . Poniżej opisujemy zatem własności endomorfizmów postaci $f(\tau)$, gdzie $f \in K[X]$ oraz $\tau \in \text{End}_K V$.

Niech $f = a_0 + a_1X + \dots + a_nX^n$ będzie wielomianem o współczynnikach z ciała K i niech $\tau \in A$ będzie dowolnym elementem algebry A . Wtedy określamy

$$f(\tau) := a_0\mathbf{1}_A + a_1\tau + \dots + a_n\tau^n.$$

Oczywiście dla dowolnego wielomianu $f \in K[X]$ oraz $\tau \in A$ mamy $f(\tau) \in A$. Ponadto, przy ustalonym $\tau \in A$, odwzorowanie

$$\varphi_\tau : K[X] \rightarrow A, \quad \varphi_\tau(f) = f(\tau)$$

jest homomorfizmem K –algebr. Dla dowolnych wielomianów $f, g \in K[X]$ mamy bowiem

$$(f + g)(\tau) = f(\tau) + g(\tau), \quad (fg)(\tau) = f(\tau) \cdot g(\tau),$$

oraz dla jedynki 1 algebry $K[X]$ mamy $\varphi_\tau(1) = \mathbf{1}_A$. Ponadto,

$$\varphi_\tau(af) = (af)(\tau) = af(\tau) = a\varphi_\tau(f)$$

dla każdego $a \in K$.

Ważną konsekwencją tego, że φ_τ jest homomorfizmem pierścieni jest następujący fakt:

Twierdzenie 2.1.1. *Dla każdego elementu $\tau \in A$ i dla dowolnych wielomianów $f, g \in K[X]$ elementy $f(\tau)$ i $g(\tau)$ są przemienne:*

$$f(\tau) \cdot g(\tau) = g(\tau) \cdot f(\tau).$$

Dowód. Mamy bowiem $f(\tau) \cdot g(\tau) = \varphi_\tau(fg) = \varphi_\tau(gf) = g(\tau) \cdot f(\tau)$. □

Zauważmy, że jeśli $\dim A < \infty$, to żaden homomorfizm K -algebr $K[X] \rightarrow A$ nie może być monomorfizmem, w szczególności więc dla żadnego elementu $\tau \in A$ homomorfizm φ_τ nie jest monomorfizmem. Nieco dokładniejszą informację podaje następujące twierdzenie.

Twierdzenie 2.1.2. *Jeśli A jest n -wymiarową algebrą nad ciałem K , to każdy element $\tau \in A$ jest zerem pewnego niezerowego wielomianu o współczynnikach z ciała K stopnia niewiększego od n .*

Dowód. Niech $\dim_K A = n$. Wtedy w przestrzeni wektorowej A każdy układ $n + 1$ elementów jest liniowo zależny. Dla każdego $\tau \in A$ istnieją więc skalary $a_0, a_1, \dots, a_n \in K$, nie wszystkie równe zero, takie, że

$$a_0 1_A + a_1 \tau + \dots + a_n \tau^n = 0_A.$$

Oznacza to, że dla wielomianu $g = a_0 + a_1 X + \dots + a_n X^n \in K[X]$ mamy $g \neq 0$ oraz $g(\tau) = 0_A$. □

Wniosek 2.1.1. *Dla każdego elementu τ skończenie wymiarowej K -algebry A jądro homomorfizmu $\varphi_\tau : K[X] \rightarrow A$ jest niezerowym ideałem w pierścieniu wielomianów $K[X]$.*

Wiemy, że w pierścieniu wielomianów $K[X]$ każdy ideał jest główny. W szczególności $\ker \varphi_\tau = (p)$ jest ideałem głównym generowanym przez pewien wielomian $p \in K[X]$. Prowadzi to do następującej definicji wielomianu minimalnego elementu τ .

Definicja 2.1.4. Niech τ będzie elementem skończenie wymiarowej K -algebry A . Wielomianem minimalnym p_τ elementu τ nazywamy unormowany generator p_τ ideału $\ker \varphi_\tau$ pierścienia $K[X]$.

A więc $p_\tau \in K[X]$ jest wielomianem minimalnym elementu $\tau \in A$ wtedy i tylko wtedy, gdy p_τ jest wielomianem unormowanym (to znaczy, najwyższy współczynnik wielomianu p_τ jest równy 1) oraz $\ker \varphi_\tau = (p_\tau)$. Warunki te są równoważne temu, że p_τ jest wielomianem unormowanym, $p_\tau(\tau) = 0_A$ oraz p_τ dzieli każdy wielomian $f \in K[X]$ taki, że $f(\tau) = 0_A$.

Przykład 2.1.6. Wielomianem minimalnym elementu zerowego 0_A jest wielomian $p_{0_A} = X \in K[X]$. Rzeczywiście, $\ker \varphi_{0_A} = (X)$. Natomiast wielomianem minimalnym jedynki 1_A algebry A jest wielomian $p_{1_A} = X - 1 \in K[X]$, gdyż $\ker \varphi_{1_A} = (X - 1)$.

Wielomian minimalny elementu τ zawiera wiele ważnych informacji o elemencie τ . Na przykład, jeśli wyraz wolny wielomianu minimalnego p_τ elementu τ jest równy zero (to znaczy, jeśli $p_\tau(0) = 0$), to element τ jest dzielnikiem zera w algebrze A . Jeśli bowiem $p_\tau = a_1 X + \dots + a_{m-1} X^{m-1} + X^m$, to

$$0_A = p_\tau(\tau) = a_1 \tau + \dots + \tau^m = \tau(a_1 1_A + \dots + \tau^{m-1}) = (a_1 1_A + \dots + \tau^{m-1})\tau,$$

oraz $a_1 1_A + \dots + \tau^{m-1} \neq 0_A$, gdyż w przeciwnym razie element τ byłby zerem niezerowego wielomianu stopnia $m - 1$ wbrew temu, że wielomian minimalny p_τ elementu τ ma stopień m .

2.1.2 Elementy odwracalne i dzielniki zera

Okazuje się, że w skończonej wymiarowej algebrze A nie może istnieć element, który byłby jednostronnie odwracalny i równocześnie nie był odwracalny. Wynika to z następującej analizy związku między odwracalnością elementu algebry a niezerowaniem się wyrazu wolnego wielomianu minimalnego tego elementu.

Twierdzenie 2.1.3. *Dla elementu τ skończenie wymiarowej algebry A następujące warunki są równoważne.*

- (a) *Wyraz wolny wielomianu minimalnego elementu τ jest różny od zera.*
- (b) *τ jest odwracalny w A .*
- (c) *τ jest lewostronnie odwracalny w A .*
- (d) *τ jest prawostronnie odwracalny w A .*

Dowód. Niech $p_\tau = a_0 + a_1X + \dots + a_{m-1}X^{m-1} + X^m$. Wtedy mamy tożsamość wielomianową

$$p_\tau = a_0 + Xp$$

gdzie $p := a_1 + \dots + a_{m-1}X^{m-2} + X^{m-1}$.

(a) $\Rightarrow$ (b) Jeśli $a_0 \neq 0$, to z równości $p_\tau(\tau) = 0_A$ otrzymujemy $0_A = a_0 1_A + \tau p(\tau)$, skąd

$$1_A = \tau \cdot (-a_0^{-1}p(\tau)) = (-a_0^{-1}p(\tau)) \cdot \tau.$$

A więc τ jest endomorfizmem odwracalnym oraz endomorfizmem odwrotnym do τ jest

$$\tau^{-1} = -a_0^{-1}p(\tau).$$

(b) $\Rightarrow$ (a) Jeśli $a_0 = 0$, to jak już zauważyliśmy τ jest dzielnikiem zera w algebrze A (gdyż $\tau \cdot p(\tau) = p(\tau) \cdot \tau = 0_A$) i wobec tego τ nie jest elementem odwracalnym.

Pozostaje pokazać, że (c) $\Rightarrow$ (b) oraz (d) $\Rightarrow$ (b) (gdyż przeciwne implikacje są trywialne). Załóżmy (c). Wtedy istnieje element τ_1 algebry A taki, że $\tau_1\tau = 1_A$. Jeśli τ nie jest odwracalny, to wobec już udowodnionej równoważności warunków (a) i (b) wyraz wolny wielomianu minimalnego p_τ jest równy zero. Wobec tego dla $\sigma = p(\tau)$ mamy $\sigma \neq 0_A$ oraz $\tau\sigma = 0_A$. Zatem

$$0_A = \tau_1(\tau\sigma) = (\tau_1\tau)\sigma = 1_A\sigma = \sigma,$$

sprzeczność. A więc τ jest odwracalny. Podobnie dowodzi się, że (d) $\Rightarrow$ (b). $\square$

Pokażemy teraz na przykładzie nieskończonej wymiarowej algebry endomorfizmów, że w algebrze mogą istnieć elementy lewostronnie odwracalne, które jednak nie są prawostronnie odwracalne (i w związku z tym nie są odwracalne).

Przykład 2.1.7. Niech $V = \mathbb{R}[X]$ będzie przestrzenią wektorową wielomianów jednej zmiennej nad ciałem $\mathbb{R}$ liczb rzeczywistych i rozpatrzmy algebrę endomorfizmów $\text{End}_{\mathbb{R}} \mathbb{R}[X]$. Niech D oraz I będą odwzorowaniami $V \rightarrow V$ określonymi następująco:

$$D(f) = \frac{d}{dX} f, \quad I(f) = \int_1^X f(t) dt.$$

Tutaj $D(f)$ jest formalną pochodną wielomianu f natomiast $I(f)$ jest formalną całką oznaczoną wielomianu f . Operacje D i I są oczywiście endomorfizmami przestrzeni wektorowej $V = \mathbb{R}[X]$. Zauważamy, że dla dowolnego wielomianu $f \in \mathbb{R}[X]$ mamy

$$DI(f) = D\left(\int_1^X f(t) dt\right) = f(X) = f,$$

oraz z drugiej strony

$$ID(f) = \int_1^X D(f)(t)dt = f(X) - f(1).$$

Zatem $D \cdot I = 1_V$, natomiast $I \cdot D \neq 1_V$, gdyż jeśli tylko wielomian f nie zeruje się w punkcie $X = 1$, to $ID(f) \neq f$. Możemy więc powiedzieć, że na przestrzeni wielomianów $\mathbb{R}[X]$ endomorfizm różniczkowania D jest prawostronnie odwracalny i prawostronnie odwrotnym endomorfizmem jest endomorfizm całkowania I . Natomiast I nie jest endomorfizmem lewostronnie odwrotnym do D . Ponieważ element odwracalny ma tylko jeden element lewostronnie odwrotny, wynika stąd, że operacja różniczkowania (a także operacja całkowania) nie jest odwracalnym endomorfizmem przestrzeni wektorowej $\mathbb{R}[X]$.

TWIERDZENIE 2.1.4. *Dla elementu τ skończenie wymiarowej algebry A następujące warunki są równoważne.*

- (a) *Wyraz wolny wielomianu minimalnego elementu τ jest równy zero.*
- (b) *τ jest dzielnikiem zera w A .*
- (c) *τ jest lewostronnym dzielnikiem zera w A .*
- (d) *τ jest prawostronnym dzielnikiem zera w A .*

Dowód. Każdy z warunków (b), (c), (d) pociąga (a), gdyż jednostronny a tym bardziej obustronny dzielnik zera nie jest elementem odwracalnym. Natomiast jeśli założymy (a), to w oznaczeniach dowodu twierdzenia 2.1.3 mamy $\tau \cdot p(\tau) = p(\tau) \cdot \tau = 0_A$ oraz $p(\tau) \neq 0_A$. Wobec tego (a) pociąga (b) i tym bardziej (c) i (d). $\square$

Według naszej definicji dzielnika zera element τ jest dzielnikiem zera w A jeśli jest równocześnie lewostronnym i prawostronnym dzielnikiem zera. Oznacza to, że istnieją $\sigma_1, \sigma_2 \in A$ takie, że $\tau\sigma_1 = \sigma_2\tau = 0_A$. Zanotujmy zauważony już wcześniej fakt, że w tej sytuacji zawsze można dobrać $\sigma_1 = \sigma_2$.

WNIOSEK 2.1.2. *Jeśli element τ jest dzielnikiem zera w skończenie wymiarowej algebrze A , to istnieje taki element $\sigma \in A$, że*

$$\tau\sigma = \sigma\tau = 0_A.$$

Dowód. W oznaczeniach dowodu twierdzenia 2.1.3 wystarczy przyjąć $\sigma = p(\tau)$. $\square$

2.2 Endomorfizmy

Algebra endomorfizmów skończenie wymiarowej przestrzeni wektorowej jest skończenie wymiarową algebrą i wobec tego stosują się do niej wszystkie konstrukcje i fakty przedstawione w §2.1.1 i §2.1.2. W szczególności, każdy endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny $p_\tau \in K[X]$, przy czym τ jest odwracalnym elementem w algebrze $\text{End}_K V$ wtedy i tylko wtedy gdy $p_\tau(0) \neq 0$ (twierdzenie 2.1.3).

2.2.1 Endomorfizmy odwracalne

Rozpatrywane w rozdziale §2.1 własności endomorfizmów przestrzeni wektorowej V są konsekwencją faktu, że wszystkie endomorfizmy przestrzeni V tworzą K -algebrę i faktycznie własności te przysługują elementom dowolnej skończenie wymiarowej K -algebry. Tymczasem endomorfizmy przestrzeni wektorowej mają swoją specyfikę, są bowiem *funkcjami liniowymi* na przestrzeni V w siebie. W szczególności, dla endomorfizmu τ przestrzeni V możemy rozpatrywać jego *jądro* i *obraz*,

$$\ker \tau = \{v \in V : \tau(v) = 0\}, \quad \text{im } \tau = \{\tau(v) \in V : v \in V\} = \tau(V),$$

podczas gdy dla elementów dowolnej K -algebry pojęcia te nie mają sensu. Pojęcie *nieosobliwości* endomorfizmu jest charakterystycznym przykładem wykorzystania funkcyjnego charakteru endomorfizmu.

DEFINICJA 2.2.1. Endomorfizm τ przestrzeni wektorowej V nazywa się endomorfizmem *nieosobliwym*, jeśli $\ker \tau = 0$.

Endomorfizm τ nazywa się endomorfizmem *osobliwym*, jeśli nie jest nieosobliwy, to znaczy, jeśli istnieje wektor $v \in V$ taki, że $v \neq 0$ i $\tau(v) = 0$.

A więc endomorfizm $\tau : V \rightarrow V$ jest nieosobliwy wtedy i tylko wtedy, gdy jest odwzorowaniem injektywnym (różnowartościowym), natomiast τ jest osobliwy wtedy i tylko wtedy, gdy ma niezerowe jądro.

Podstawą do powiązania nieosobliwości endomorfizmu z własnościami, które zależą od wielomianu minimalnego endomorfizmu jest następujący lemat.

LEMAT 2.2.1. Niech $\tau : V \rightarrow V$ będzie endomorfizmem skończenie wymiarowej przestrzeni wektorowej nad ciałem K . Wtedy

$$\dim_K V = \dim_K \ker \tau + \dim_K \operatorname{im} \tau.$$

Dowód. Rozpatrzmy przestrzeń ilorazową $V/\ker \tau$. Na podstawie twierdzenia o homomorfizmach przestrzeni wektorowych mamy izomorfizm

$$\operatorname{im} \tau \cong V/\ker \tau.$$

Porównując wymiary otrzymujemy

$$\dim_K \operatorname{im} \tau = \dim_K V/\ker \tau = \dim_K V - \dim_K \ker \tau. \quad \square$$

TWIERDZENIE 2.2.1. Dla endomorfizmu $\tau : V \rightarrow V$ skończenie wymiarowej przestrzeni wektorowej V następujące warunki są równoważne:

- (a) τ jest nieosobliwy.
- (b) τ jest monomorfizmem.
- (c) τ jest epimorfizmem.
- (d) τ jest izomorfizmem.
- (e) τ jest odwracalny w $\operatorname{End}_K V$.
- (f) Wyraz wolny wielomianu minimalnego endomorfizmu τ jest różny od zera.

Dowód. Równoważność warunków (a), (b), (c), (d) wynika z lematu 2.2.1, natomiast równoważność warunków (e), (f) wynika z twierdzenia 2.1.3. Jeśli τ jest endomorfizmem odwracalnym i τ^{-1} jest endomorfizmem odwrotnym do τ , to z równości $\tau\tau^{-1} = 1_V$ wynika, że $\tau(\tau^{-1}(v)) = v$ dla każdego $v \in V$. Zatem τ jest epimorfizmem. A więc (e) $\Rightarrow$ (c).

Wystarczy teraz pokazać, że (d) $\Rightarrow$ (e). Jeśli $\tau : V \rightarrow V$ jest izomorfizmem, to każdy wektor przestrzeni V można jednoznacznie przedstawić w postaci $\tau(v)$, gdzie $v \in V$. Wobec tego określamy odwzorowanie

$$\sigma : V \rightarrow V, \quad \sigma(\tau(v)) = v$$

dla każdego $v \in V$. Łatwe sprawdzenie pokazuje, że σ jest endomorfizmem przestrzeni V . Ponadto, z określenia endomorfizmu σ wynika, że $\sigma\tau = 1_V$. A więc endomorfizm τ jest lewostronnie odwracalny i wobec tego, na podstawie twierdzenia 2.1.3, jest odwracalny. $\square$

Twierdzenie 2.2.2. *Dla endomorfizmu $\tau : V \rightarrow V$ skończenie wymiarowej przestrzeni wektorowej V następujące warunki są równoważne:*

- (a) τ jest osobliwy.
- (b) τ nie jest monomorfizmem.
- (c) τ nie jest epimorfizmem.
- (d) τ nie jest izomorfizmem.
- (e) τ jest dzielnikiem zera w $\text{End}_K V$.
- (f) Wyraz wolny wielomianu minimalnego endomorfizmu τ jest równy zero.

Dowód. Równoważność warunków (a), (b), (c), (d) wynika z lematu 2.2.1 natomiast równoważność warunków (e), (f) wynika z twierdzenia 2.1.4. Na podstawie twierdzeń 2.1.3 i 2.1.4 endomorfizm τ jest dzielnikiem zera w $\text{End}_K V$ wtedy i tylko wtedy, gdy nie jest odwracalny. Zatem równoważność (e) i (d) wynika z równoważności warunków (e) i (d) w twierdzeniu 2.2.1. $\square$

Definicja 2.2.2. *Rzędem rank τ endomorfizmu $\tau : V \rightarrow V$ nazywamy wymiar obrazu endomorfizmu τ :*

$$\text{rank } \tau := \dim_K \text{im } \tau.$$

Zauważmy, że $0 \leq \text{rank } \tau \leq \dim_K V$. Przy tym $\text{rank } \tau = \dim_K V$ wtedy i tylko wtedy, gdy endomorfizm τ jest nieosobliwy. Jeśli τ jest osobliwy, to $0 \leq \text{rank } \tau < \dim_K V$. Widzimy więc, że rząd endomorfizmu pozwala nie tylko rozgraniczyć endomorfizmy nieosobliwe od osobliwych, ale jest także swoistą miarą osobliwości endomorfizmu. Można intuicyjnie uważać, że im większy rząd endomorfizmu osobliwego, tym bliższy jest on nieosobliwości. Z drugiej strony, jeśli $\text{rank } \tau = 0$, to $\tau = 0_V$ jest krańcowo osobliwy. Zauważmy, że zgodnie z lematem 2.2.1,

$$\text{rank } \tau = \dim_K V - \dim_K \ker \tau.$$

Twierdzenie 2.2.3. *Niech σ, τ będą endomorfizmami skończenie wymiarowej przestrzeni wektorowej V nad ciałem K . Wtedy*

- (a) $\text{rank } \sigma\tau \leq \text{rank } \tau$.
- (b) $\text{rank } \sigma\tau \leq \text{rank } \sigma$.
- (c) $\text{rank } \sigma\tau \leq \min\{\text{rank } \sigma, \text{rank } \tau\}$.
- (d) *Jeśli endomorfizm σ jest nieosobliwy, to*

$$\text{rank } \sigma\tau = \text{rank } \tau\sigma = \text{rank } \tau.$$

Dowód. (a) Skorzystamy tu z faktu, że dla dowolnej podprzestrzeni U przestrzeni V mamy $\dim \sigma(U) \leq \dim U$. Wtedy dla $U = \tau(V)$ mamy

$$\text{rank } \sigma\tau = \dim \text{im } \sigma\tau = \dim \sigma(\tau(V)) \leq \dim \tau(V) = \text{rank } \tau.$$

(b) Skorzystamy tu z faktu, że dla dowolnych podprzestrzeni U_1, U_2 przestrzeni V , jeśli $U_1 \subseteq U_2$, to $\dim \sigma(U_1) \leq \dim \sigma(U_2)$. Zatem dla $U_1 = \tau(V)$ oraz $U_2 = V$ mamy

$$\text{rank } \sigma\tau = \dim \sigma(\tau(V)) \leq \dim \sigma(V) = \text{rank } \sigma.$$

(c) wynika z (a) i (b).

(d) Endomorfizm nieosobliwy σ jest izomorfizmem i wobec tego przeprowadza podprzestrzeń U przestrzeni V na podprzestrzeń o takim samym wymiarze, a więc $\dim \sigma(U) = \dim U$. Stąd dla $U = \tau(V)$ mamy

$$\text{rank } \sigma\tau = \dim \sigma(\tau(V)) = \dim \tau(V) = \text{rank } \tau.$$

Z drugiej strony, jeśli endomorfizm σ jest nieosobliwy, to $\sigma(V) = V$ i wobec tego

$$\text{rank } \tau\sigma = \dim \tau(\sigma(V)) = \dim \tau(V) = \text{rank } \tau.$$

Dowodzi to części (d) twierdzenia. $\square$

WNIOSEK 2.2.1. *Jeśli σ jest nieosobliwym endomorfizmem skończenie wymiarowej przestrzeni wektorowej V , to dla każdego endomorfizmu τ przestrzeni V mamy*

$$\text{rank } \sigma\tau\sigma^{-1} = \text{rank } \tau.$$

Dowód. Zastosujemy część (d) twierdzenia 2.2.3:

$$\text{rank } \sigma\tau\sigma^{-1} = \text{rank } \sigma(\tau\sigma^{-1}) = \text{rank } \tau\sigma^{-1} = \text{rank } \tau,$$

gdyż wraz z σ także σ^{-1} jest endomorfizmem nieosobliwym. $\square$

2.2.2 Podobieństwo

Zbadamy teraz związek między macierzami $\mathbf{m}(\tau, \mathcal{A})$ oraz $\mathbf{m}(\tau, \mathcal{B})$ endomorfizmu τ w dwóch różnych uporządkowanych bazach $\mathcal{A}$ i $\mathcal{B}$ przestrzeni V . Niech więc $\mathcal{A} = \{u_1, \dots, u_n\}$ i $\mathcal{B} = \{v_1, \dots, v_n\}$ będą uporządkowanymi bazami przestrzeni V i niech $\mathbf{m}(\tau, \mathcal{A}) = [a_{ij}] =: A$ oraz $\mathbf{m}(\tau, \mathcal{B}) = [b_{ij}] =: B$. Zatem

$$\tau(u_j) = \sum_{i=1}^n a_{ij}u_i, \quad \tau(v_j) = \sum_{i=1}^n b_{ij}v_i$$

dla $j = 1, \dots, n$. Obieramy endomorfizm $\sigma \in \text{End}_K V$ taki, że

$$\sigma(u_j) = v_j, \quad j = 1, \dots, n.$$

Wtedy σ jest automorfizmem przestrzeni V oraz

$$\tau\sigma(u_j) = \tau(v_j) = \sum_{i=1}^n b_{ij}v_i = \sum_{i=1}^n b_{ij}\sigma(u_i) = \sigma\left(\sum_{i=1}^n b_{ij}u_i\right).$$

Wynika stąd, że

$$\sigma^{-1}\tau\sigma(u_j) = \sum_{i=1}^n b_{ij}u_i,$$

dla $j = 1, \dots, n$. Równości te pokazują, że endomorfizm $\sigma^{-1}\tau\sigma$ ma w bazie $\mathcal{A}$ macierz B , to znaczy, $\mathbf{m}(\sigma^{-1}\tau\sigma, \mathcal{A}) = B$. Niech $S := \mathbf{m}(\sigma, \mathcal{A})$. Wtedy

$$B = \mathbf{m}(\sigma^{-1}\tau\sigma, \mathcal{A}) = \mathbf{m}(\sigma^{-1}, \mathcal{A}) \mathbf{m}(\tau, \mathcal{A}) \mathbf{m}(\sigma, \mathcal{A}) = S^{-1}AS.$$

Udowodniliśmy więc następujące twierdzenie.

TWIERDZENIE 2.2.4. *Jeśli A i B są macierzami endomorfizmu τ w uporządkowanych bazach $\mathcal{A}$ i $\mathcal{B}$ przestrzeni V i jeśli $S = \mathbf{m}(\sigma, \mathcal{A})$ jest macierzą automorfizmu σ przeprowadzającego bazę $\mathcal{A}$ na bazę $\mathcal{B}$, to $B = S^{-1}AS$, to znaczy*

$$\mathbf{m}(\tau, \mathcal{B}) = \mathbf{m}(\sigma, \mathcal{A})^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}) \cdot \mathbf{m}(\sigma, \mathcal{A}).$$

Uwaga 2.2.1. Macierz $S = \mathbf{m}(\sigma, \mathcal{A})$ można także interpretować jako *macierz przejścia* od bazy $\mathcal{A}$ do bazy $\mathcal{B}$. Jeśli bowiem $S = [s_{ij}]$, to

$$v_j = \sigma(u_j) = \sum_{i=1}^n s_{ij}u_i, \quad j = 1, \dots, n.$$

DEFINICJA 2.2.3. Macierze $A, B \in M_n(K)$ nazywamy *podobnymi* lub *sprzężonymi*, jeśli istnieje macierz odwracalna $S \in M_n(K)$ taka, że

$$B = S^{-1}AS.$$

Podobieństwo macierzy jest relacją równoważnościową w algebrze macierzy $M_n(K)$. Twierdzenie 2.2.4 orzeka, że macierze endomorfizmu τ w różnych bazach przestrzeni V są podobne.

DEFINICJA 2.2.4. Endomorfizmy ρ i τ przestrzeni V nazywamy endomorfizmami *podobnymi* lub *sprzężonymi*, jeśli istnieje endomorfizm odwracalny σ taki, że

$$\rho = \sigma^{-1}\tau\sigma.$$

Łatwo sprawdzić, że relacja podobieństwa endomorfizmów jest relacją równoważności w algebrze $\text{End}_K V$.

TWIERDZENIE 2.2.5. Jeśli ρ i τ są podobnymi endomorfizmami przestrzeni V , to

- (a) $p_\rho = p_\tau$.
- (b) ρ jest nieosobliwy wtedy i tylko wtedy, gdy τ jest nieosobliwy,
- (c) $\text{rank } \rho = \text{rank } \tau$.

Dowód. (a) Niech $\sigma, \tau, \rho \in \text{End}_K V$ oraz niech σ będzie endomorfizmem nieosobliwym. Załóżmy, że $\rho = \sigma^{-1}\tau\sigma$. Zauważmy najpierw, że dla dowolnej liczby naturalnej i mamy

$$(\sigma^{-1}\tau\sigma)^i = \sigma^{-1}\tau\sigma \cdot \sigma^{-1}\tau\sigma \cdots \sigma^{-1}\tau\sigma = \sigma^{-1}\tau^i\sigma.$$

Zatem dla dowolnego wielomianu $g = \sum a_i X^i \in K[X]$ mamy

$$g(\sigma^{-1}\tau\sigma) = \sum a_i (\sigma^{-1}\tau\sigma)^i = \sum a_i \sigma^{-1}\tau^i\sigma = \sigma^{-1}g(\tau)\sigma.$$

Wynika stąd, że $g(\sigma^{-1}\tau\sigma) = 0_V$ wtedy i tylko wtedy, gdy $g(\tau) = 0_V$. A więc dla dowolnego wielomianu g ,

$$g(\rho) = 0_V \iff g(\tau) = 0_V.$$

W szczególności, wobec $p_\rho(\rho) = 0_V$ i $p_\tau(\tau) = 0_V$ otrzymujemy $p_\rho(\tau) = 0_V$ i $p_\tau(\rho) = 0_V$. Zatem $p_\tau \mid p_\rho$ i $p_\rho \mid p_\tau$, skąd wynika już równość wielomianów minimalnych endomorfizmów $\rho = \sigma^{-1}\tau\sigma$ oraz τ .

(b) wynika z twierdzenia 2.2.1 i z części (a). Zauważmy też, że (b) wynika z (c). Natomiast (c) udowodniliśmy już jako wniosek 2.2.1. $\square$

TWIERDZENIE 2.2.6. Dla endomorfizmów ρ i τ przestrzeni V następujące warunki są równoważne:

- (a) ρ i τ są podobne.
- (b) Dla każdej uporządkowanej bazy $\mathcal{A}$ przestrzeni V macierze $\mathbf{m}(\rho, \mathcal{A})$ i $\mathbf{m}(\tau, \mathcal{A})$ są podobne.
- (b') Istnieje uporządkowana baza $\mathcal{A}'$ przestrzeni V taka, że macierze $\mathbf{m}(\rho, \mathcal{A}')$ i $\mathbf{m}(\tau, \mathcal{A}')$ są podobne.
- (c) Dla każdej uporządkowanej bazy $\mathcal{A}$ przestrzeni V istnieje uporządkowana baza $\mathcal{B}$ przestrzeni V taka, że $\mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\tau, \mathcal{B})$.
- (c') Istnieją uporządkowane bazy $\mathcal{A}$ i $\mathcal{B}$ przestrzeni V takie, że $\mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\tau, \mathcal{B})$.

Dowód. (a) $\Rightarrow$ (b) Niech $\rho = \sigma^{-1}\tau\sigma$, gdzie $\sigma \in \text{Aut } V$ i niech $\mathcal{A}$ będzie dowolną bazą przestrzeni V . Wtedy

$$\mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\sigma^{-1}\tau\sigma, \mathcal{A}) = S^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}) \cdot S,$$

gdzie $S = \mathbf{m}(\sigma, \mathcal{A})$. A więc macierze $\mathbf{m}(\rho, \mathcal{A})$ i $\mathbf{m}(\tau, \mathcal{A})$ są podobne.

(b) $\Rightarrow$ (b') jest oczywiste.

(b') $\Rightarrow$ (c) Niech $\mathbf{m}(\rho, \mathcal{A}') = S^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}') \cdot S$, gdzie S jest macierzą odwracalną i niech $\mathcal{A}$ będzie dowolną bazą uporządkowaną przestrzeni V . Macierz $\mathbf{m}(\rho, \mathcal{A})$ jest podobna do macierzy $\mathbf{m}(\rho, \mathcal{A}')$ (na podstawie twierdzenia 2.2.4), zatem istnieje macierz odwracalna T taka, że

$$\mathbf{m}(\rho, \mathcal{A}) = T^{-1} \cdot \mathbf{m}(\rho, \mathcal{A}') \cdot T.$$

Stąd

$$\mathbf{m}(\rho, \mathcal{A}) = T^{-1}S^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}') \cdot ST = (ST)^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}') \cdot (ST).$$

Obieramy automorfizm α przestrzeni V taki, że $ST = \mathbf{m}(\alpha, \mathcal{A}')$. Niech $\mathcal{B} = \alpha(\mathcal{A}')$. Wtedy $\mathcal{B}$ jest bazą przestrzeni V i na podstawie twierdzenia 2.2.4 mamy $(ST)^{-1}\mathbf{m}(\tau, \mathcal{A}')(ST) = \mathbf{m}(\tau, \mathcal{B})$. A więc $\mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\tau, \mathcal{B})$.

(c) $\Rightarrow$ (c') jest oczywiste.

(c') $\Rightarrow$ (a) Niech σ będzie automorfizmem przestrzeni V przeprowadzającym bazę $\mathcal{A}$ na bazę $\mathcal{B}$. Wtedy na podstawie (c') i twierdzenia 2.2.4 mamy

$$\mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\tau, \mathcal{B}) = \mathbf{m}(\sigma, \mathcal{A})^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}) \cdot \mathbf{m}(\sigma, \mathcal{A}) = \mathbf{m}(\sigma^{-1}\tau\sigma, \mathcal{A}).$$

Endomorfizmy ρ i $\sigma^{-1}\tau\sigma$ mają więc równe macierze w bazie $\mathcal{A}$ skąd wynika, że $\rho = \sigma^{-1}\tau\sigma$. A więc (c') $\Rightarrow$ (a). $\square$

Jak widzimy, endomorfizmy podobne mają wiele wspólnych własności. Naturalnym problemem jest więc klasyfikacja endomorfizmów przestrzeni V ze względu na podobieństwo endomorfizmów. Problem ten jest ściśle związany z klasyfikacją macierzy w $M_n(K)$ ze względu na podobieństwo macierzy.

Jeśli bowiem ustalimy uporządkowaną bazę $\mathcal{A}$ przestrzeni V i każdemu endomorfizmowi τ przestrzeni V przyporządkujemy macierz $\mu(\tau) = A = \mathbf{m}(\tau, \mathcal{A})$ endomorfizmu τ względem bazy $\mathcal{A}$, to jak wiemy, otrzymujemy izomorfizm K -algebr

$$\mu : \text{End}_K V \rightarrow M_n(K).$$

Ten izomorfizm przeprowadza klasy endomorfizmów podobnych na klasy macierzy podobnych:

$$\mu\{\sigma^{-1}\tau\sigma : \sigma \in \text{Aut } V\} = \{S^{-1}AS : S \in \mathbf{GL}(n, k)\},$$

gdzie $\mu(\sigma) = S$ dla $\sigma \in \text{Aut } V$. A więc μ przeprowadza klasę endomorfizmów podobnych do τ na zbiór macierzy endomorfizmu τ we wszystkich bazach przestrzeni V (lub równoważnie, na zbiór macierzy wszystkich endomorfizmów podobnych do τ , w ustalonej bazie $\mathcal{A}$ przestrzeni V).

Ta obserwacja jest podstawą klasyfikacji endomorfizmów przestrzeni wektorowej i poszukiwania postaci kanonicznych macierzy endomorfizmów. Z każdym endomorfizmem τ przestrzeni V wiążemy klasę macierzy tego endomorfizmu we wszystkich bazach przestrzeni V i znajdujemy w tej klasie macierze szczególnych postaci, które opisują przejrzyście działanie endomorfizmu τ na odpowiadających tym macierzom bazach przestrzeni wektorowej V . Są to tak zwane postaci kanoniczne macierzy endomorfizmu τ .

2.3 Algebra endomorfizmów

2.3.1 Reprezentacja algebry w algebrze endomorfizmów

W teorii grup dowodzi się, że każda grupa jest izomorficzna z podgrupą grupy symetrycznej (permutacji) pewnego zbioru. Okazuje się, że algebry endomorfizmów odgrywają w teorii algebr taką rolę jak grupy symetryczne w teorii grup. Wynika to z następującego faktu.

TWIERDZENIE 2.3.1. *Każda K -algebra jest izomorficzna z podalgebrą algebry endomorfizmów $\text{End}_K V$ pewnej przestrzeni wektorowej V nad ciałem K .*

Dowód. Niech A będzie K -algebrą. Wtedy A jest przestrzenią wektorową nad ciałem K i wobec tego możemy rozpatrywać algebrę endomorfizmów $\text{End}_K A$ tej przestrzeni wektorowej. Pokażemy, że algebra A jest izomorficzna z podalgebrą algebry endomorfizmów $\text{End}_K A$ przestrzeni wektorowej A .

Dla każdego $a \in A$ rozpatrujemy odwzorowanie $\tau_a : A \rightarrow A$ określone następująco: $\tau_a(v) = av$ dla każdego $v \in A$. Tutaj av jest iloczynem elementów algebry A . Łatwo sprawdza się, że τ_a jest endomorfizmem przestrzeni wektorowej A . Dla $u, v \in A$ mamy bowiem na podstawie rozdzielności mnożenia względem dodawania w A

$$\tau_a(u + v) = a(u + v) = au + av = \tau_a(u) + \tau_a(v),$$

oraz dla $x \in K, v \in A$ mamy

$$\tau_a(xv) = a(xv) = x(av) = x\tau_a(v).$$

Określamy teraz odwzorowanie

$$\varphi : A \rightarrow \text{End } A, \quad \varphi(a) = \tau_a.$$

Rutynowe rachunki pokazują, że φ jest homomorfizmem K -algebr, to znaczy,

$$\varphi(xa + yb) = x\varphi(a) + y\varphi(b), \quad \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b), \quad \varphi(1) = 1_A$$

dla każdych $x, y \in K, a, b \in A$. Jądrem tego homomorfizmu jest zbiór tych $a \in A$ dla których $\varphi(a) = \tau_a$ jest endomorfizmem zerowym 0_A . A więc $a \in \ker \varphi$ pociąga, że $av = 0$ dla każdego $v \in A$. W szczególności więc $a = a1 = 0$, skąd wynika, że $\ker \varphi = \{0\}$. Zatem φ jest różnowartościowym homomorfizmem i wobec tego jest izomorfizmem algebry A na podalgebrę $\varphi(A)$ algebry $\text{End}_K A$. $\square$

2.3.2 Centrum algebry endomorfizmów

W nieprzemiennej K -algebrze A zawsze istnieją elementy, które są przemienne z każdym elementem algebry A . Takim elementem jest jedynka 1_A algebry A a także element zerowy $0 \in A$. Ogólniej, skalarne wielokrotności jedynki algebry mają tę własność. Rzeczywiście, dzięki własności jednorodności wewnętrznego mnożenia w A mamy

$$(x1_A) \cdot a = x(1_A \cdot a) = x(a \cdot 1_A) = a \cdot (x1_A)$$

dla wszystkich $x \in K, a \in A$.

DEFINICJA 2.3.1. *Centrum K -algebry A nazywamy zbiór*

$$Z(A) := \{c \in A : ca = ac \text{ dla wszystkich } a \in A\}$$

wszystkich elementów algebry A przemianych z każdym elementem algebry A .

Jak już zauważyliśmy, zbiór $K1_A := \{x1_A : x \in K\}$ skalnych wielokrotności jedynki algebry A zawsze zawiera się w centrum algebry A : $K1_A \subseteq Z(A)$.

DEFINICJA 2.3.2. K -algebra A nazywa się *centralną* K -algebrą, jeśli $K1_A = Z(A)$.

Można powiedzieć, że centralna K -algebra jest w wysokim stopniu nieprzemienne: jej centrum nie zawiera żadnych elementów spoza zbioru $K1_A$, który koniecznie musi zawierać się w centrum.

TWIERDZENIE 2.3.2. Algebra endomorfizmów $\text{End}_K V$ dowolnej przestrzeni wektorowej nad dowolnym ciałem K jest centralną K -algebrą.

Dowód. Jak w każdym pierścieniu, centrum $Z(\text{End}_K V)$ jest podpierścieniem, a nawet podalgebrą algebry $\text{End}_K V$. Zauważmy najpierw, że

$$0_V, 1_V \in Z(\text{End}_K V)$$

i ogólniej, jak to zauważyliśmy już wcześniej dla dowolnej K -algebry, $Z(\text{End}_K V)$ zawiera wszystkie endomorfizmy skalarne $a1_V$, gdzie $a \in K$. Pokażemy, że centrum jest podciałem algebry $\text{End}_K V$ złożonym z wszystkich endomorfizmów skalnych:

$$Z(\text{End}_K V) = K1_V.$$

Jeśli $\sigma = a1_V$ jest endomorfizmem skalarным, to dla każdego $v \in V$ mamy $\sigma(v) = av \in Kv$. Endomorfizm skalarный σ zachowuje więc wszystkie proste w przestrzeni V (to znaczy $\sigma(Kv) \subseteq Kv$ dla każdego $v \in V$, a dla $\sigma \neq 0$ mamy oczywiście $\sigma(Kv) = Kv$).

Najpierw sprawdzimy, że każdy endomorfizm $\sigma \in Z(\text{End}_K V)$ zachowuje wszystkie proste w przestrzeni V . Przypuśćmy, że endomorfizm τ przestrzeni V nie zachowuje wszystkich prostych przestrzeni V , to znaczy istnieje wektor $v \in V$ taki, że $\tau(v) \notin Kv$. Pokażemy, że τ nie należy do centrum algebry $\text{End}_K V$.

Rzeczywiście, niech $\tau(v) = w \notin Kv$. Wtedy v i w są liniowo niezależne, istnieje więc endomorfizm ρ taki, że $\rho(v) = v$ oraz $\rho(w) \notin Kw$, na przykład $\rho(w) = v + w$. Mamy więc

$$\tau\rho(v) = \tau(v) = w, \quad \rho\tau(v) = \rho(w) = v + w \notin Kw,$$

skąd wynika, że $\tau\rho \neq \rho\tau$ i wobec tego τ nie należy do centrum $\text{End}_K V$.

Dowodzi to, że każdy endomorfizm $\sigma \in Z(\text{End}_K V)$ zachowuje wszystkie proste w przestrzeni V : dla każdego $v \in V$ mamy $\sigma(v) \in Kv$. Teraz łatwo stwierdzamy, że endomorfizm σ jest skalarный.

Niech bowiem $\sigma \in Z(\text{End}_K V)$ i niech u, v będą niezerowymi wektorami przestrzeni V oraz $\sigma(u) = au$, $\sigma(v) = bv$ dla pewnych $a, b \in K$. Obieramy endomorfizm τ przestrzeni V taki, że $\tau(u) = v$. Wtedy

$$\sigma\tau(u) = \sigma(v) = bv \quad \text{oraz} \quad \tau\sigma(u) = \tau(au) = a\tau(u) = av.$$

Ponieważ $\sigma\tau = \tau\sigma$ oraz $v \neq 0$, więc wynika stąd, że $a = b$. Zatem, jeśli $\sigma(u) = au$ dla pewnego niezerowego wektora u , to $\sigma(v) = av$ dla każdego wektora $v \in V$. Oznacza to, że $\sigma = a1_V$ jest endomorfizmem skalarным i dowodzi, że $Z(\text{End}_K V) = K1_V$. $\square$

Uwaga 2.3.1. Łatwo stwierdzić, że jeśli $h : A \rightarrow B$ jest izomorfizmem K -algebr, to $h(Z(A)) = Z(B)$. W szczególności więc, jeśli A jest centralną K -algebrą, to $Z(B) = h(K1_A) = Kh(1_A) = K1_B$, to znaczy, B jest także centralną K -algebrą. Z twierdzenia 2.3.2 i z faktu, że algebra endomorfizmów n -wymiarowej przestrzeni wektorowej jest izomorficzna z algebrą macierzy $M_n(K)$ wynika, że każda algebra $M_n(K)$ jest centralną K -algebrą.

2.3.3 Ideały algebry endomorfizmów

Jak już wspominaliśmy w §2.1, algebra A nazywa się *algebrą z dzieleniem* jeśli każdy niezerowy element algebry A jest odwracalny. Algebry z dzieleniem są więc wśród pierścieni nieprzemiennych odpowiednikami ciał. Jednakże ciała wśród pierścieni przemiennych można wyróżnić nie tylko poprzez fakt, że wszystkie niezerowe elementy są odwracalne. Inna charakterystyka ciał odwołuje się do faktu, że pierścień przemienny K jest ciałem wtedy i tylko wtedy gdy ma tylko dwa ideały, ideał zerowy $0 = 0K$ i ideał jednostkowy $K = 1K$. Uogólniając tę charakterystykę ciał wprowadzimy ważną klasę algebr *prostych*. Najpierw jednak musimy przedyskutować pojęcie ideału w pierścieniach nieprzemiennych.

Lewostronnym ideałem w algebrze A nazywamy podgrupę $\mathcal{I}$ addytywnej grupy algebry A zamkniętą ze względu na mnożenie z lewej strony przez wszystkie elementy algebry A , to znaczy, addytywną podgrupę $\mathcal{I}$ spełniającą warunek

$$a \in A \quad \text{i} \quad b \in \mathcal{I} \quad \Rightarrow \quad ab \in \mathcal{I}.$$

Prawostronnym ideałem algebry A nazywamy podgrupę $\mathcal{J}$ addytywnej grupy algebry A spełniającą warunek

$$a \in A \quad \text{i} \quad b \in \mathcal{J} \quad \Rightarrow \quad ba \in \mathcal{J}.$$

Podgrupa $\mathcal{I}$ addytywnej grupy algebry A , która jest równocześnie lewo- i prawostronnym ideałem algebry A nazywa się *ideałem* algebry A (lub dwustronnym ideałem algebry A). Zauważmy, że w K -algebrze A każdy ideał $\mathcal{I}$ jest podprzestrzenią wektorową przestrzeni A . Rzeczywiście, $\mathcal{I}$ jest podgrupą addytywnej grupy A oraz dla dowolnych $x \in K$ i $a \in \mathcal{I}$ mamy $xa = x(1_A a) = (x1_A)a \in \mathcal{I}$.

Najprostsze przykłady ideałów jednostronnych otrzymuje się następująco. Obieramy dowolny element $a \in A$ i rozpatrujemy zbiór aA wszystkich prawostronnych wielokrotności elementu a . Oczywiście aA jest prawostronnym ideałem w A . Podobnie Aa jest lewostronnym ideałem w A .

Co do (dwustronnych) ideałów, istnieją zawsze dwa oczywiste przykłady w każdej algebrze A . Są to mianowicie $0A = A0 = 0$, ideał zerowy, oraz $1_A A = A1_A = A$, cała algebra A . Można też zauważyć, że jądro $\ker h$ homomorfizmu K -algebr $h : A \rightarrow B$ jest ideałem w A . Rzeczywiście, jeśli $h(a) = 0$, to dla każdego $b \in A$ mamy $h(ab) = 0 \cdot h(b) = 0$ oraz $h(ba) = h(b) \cdot 0 = 0$.

DEFINICJA 2.3.3. Algebra A nad ciałem K nazywa się *prostą K -algebrą*, jeśli ideał zerowy 0 i algebra A są jedynymi ideałami algebry A .

Przykład 2.3.1. Jeśli L jest rozszerzeniem ciała K , to L jest prostą K -algebrą. K -algebra wielomianów $K[X]$ nie jest prosta, gdyż wszystkie ideały główne w $K[X]$ generowane przez wielomiany stopni ≥ 1 są różne od 0 i $K[X]$.

Każda algebra z dzieleniem jest prosta. Załóżmy bowiem, że $\mathcal{I}$ jest niezerowym ideałem w algebrze z dzieleniem A . Niech $a \in \mathcal{I}$ oraz $a \neq 0$. Wtedy a jest elementem odwracalnym w A , zatem istnieje $b \in A$ taki, że $ab = 1_A$. A więc $1_A = ab \in \mathcal{I}$ i stąd wynika, że $A = \mathcal{I}$. Zatem każdy niezerowy ideał algebry z dzieleniem A musi być równy A , skąd wynika, że A jest algebrą prostą. W szczególności, na podstawie stwierdzenia 2.1.1, $\mathbb{R}$ -algebra kwaternionów Hamiltona $\mathbb{H}$ jest algebrą prostą.

Okazuje się jednak, że klasa algebr prostych jest szersza niż klasa algebr z dzieleniem. Wynika to z następującego twierdzenia.

TWIERDZENIE 2.3.3. Niech V będzie skończone wymiarową przestrzenią wektorową nad dowolnym ciałem K . Wtedy algebra endomorfizmów $\text{End}_K V$ jest prostą K -algebrą.

Dowód. Niech $\dim_K V = n$ i niech $\{v_1, \dots, v_n\}$ będzie bazą przestrzeni V . Dla każdej pary liczb naturalnych i, j niewiększych od n obieramy endomorfizm τ_{ij} przestrzeni V taki, że

$$\tau_{ij}(v_k) = \delta_{jk}v_i, \quad k = 1, \dots, n,$$

gdzie δ_{jk} jest "delta Kroneckera," $\delta_{jk} = 1$ gdy $j = k$ oraz $\delta_{jk} = 0$ gdy $j \neq k$.

Wiemy już, że endomorfizmy τ_{ij} tworzą bazę K -algebry $\text{End}_K V$.

Bazę $\{\tau_{ij} : 1 \leq i, j \leq n\}$ nazywamy *bazą standardową* algebry endomorfizmów $\text{End}_K V$.

Łatwo także sprawdzić następującą tabelę mnożenia dla tej bazy:

$$\tau_{ij} \cdot \tau_{k\ell} = \begin{cases} \tau_{i\ell} & \text{gdy } j = k, \\ 0 & \text{gdy } j \neq k. \end{cases} \quad (2.5)$$

Warto zauważyć, że endomorfizm tożsamościowy 1_V ma następujące jednoznaczne przedstawienie jako kombinacja liniowa endomorfizmów bazy standardowej:

$$1_V = \tau_{11} + \dots + \tau_{nn}. \quad (2.6)$$

Niech teraz $\tau \in \text{End}_K V$ będzie dowolnym endomorfizmem i niech

$$\tau = \sum_{k,\ell} a_{k\ell} \tau_{k\ell},$$

gdzie $a_{k\ell} \in K$ są współrzędnymi endomorfizmu τ w bazie standardowej algebry $\text{End}_K V$. Wtedy dla wszystkich i, j , $1 \leq i, j \leq n$ mamy następujące tożsamości:

$$\tau_{ii} \cdot \tau \cdot \tau_{jj} = a_{ij} \tau_{ij}. \quad (2.7)$$

Przypuśćmy teraz, że algebra $\text{End}_K V$ ma niezerowy ideał $\mathcal{I}$ i niech endomorfizm τ będzie niezerowym elementem $\mathcal{I}$. Wtedy τ ma przynajmniej jedną niezerową współrzędną $a_{ij} \neq 0$. Ponieważ $\mathcal{I}$ jest ideałem i $\tau \in \mathcal{I}$, więc także $\tau_{ii} \cdot \tau \cdot \tau_{jj} \in \mathcal{I}$, zatem wobec (2.7) mamy $a_{ij} \tau_{ij} \in \mathcal{I}$. Wiemy także, że $\mathcal{I}$ jest podprzestrzenią przestrzeni wektorowej A , zatem

$$a_{ij}^{-1} \cdot a_{ij} \tau_{ij} = \tau_{ij} \in \mathcal{I}.$$

Pokazaliśmy więc, że jeśli $\tau \in \mathcal{I}$ oraz współrzędna o numerze (i, j) endomorfizmu τ jest niezerowym skalar, to $\tau_{ij} \in \mathcal{I}$. Ponieważ $\mathcal{I}$ jest ideałem, mamy także

$$\tau_{ki} \cdot \tau_{ij} \cdot \tau_{jk} \in \mathcal{I}$$

dla wszystkich $k = 1, \dots, n$. Z drugiej strony, wykorzystując dwukrotnie tożsamość (2.5) otrzymujemy

$$\tau_{ki} \cdot \tau_{ij} \cdot \tau_{jk} = \tau_{kk}.$$

Zatem $\tau_{kk} \in \mathcal{I}$ dla $k = 1, \dots, n$, i wobec (2.6) mamy

$$1_V = \tau_{11} + \dots + \tau_{nn} \in \mathcal{I}.$$

Stąd wynika, że każdy endomorfizm τ należy do $\mathcal{I}$, gdyż $\tau = \tau \cdot 1_V \in \mathcal{I}$. A więc każdy niezerowy ideał algebry endomorfizmów $\text{End}_K V$ jest równy całej algebrze $\text{End}_K V$. Wobec tego $\text{End}_K V$ jest prostą K -algebrą. $\square$

Algebry, które są równocześnie centralne i proste nazywają się *centralnymi prostymi* K -algebrami. Z twierdzeń 2.3.2 i 2.3.3 otrzymujemy zatem następujący rezultat.

TWIERDZENIE 2.3.4. *Algebra endomorfizmów skończenie wymiarowej przestrzeni wektorowej V nad ciałem K jest centralną prostą K -algebrą. Każda algebra macierzy $M_n(K)$ jest centralną prostą K -algebrą.*

Uwaga 2.3.2. Algebra kwaternionów Hamiltona $\mathbb{H}$ jest także centralną prostą $\mathbb{R}$ -algebrą.

Rozdział 3

Podprzestrzenie niezmiennicze

Ostatnie zmiany 6.11.2008 r.

Głównym problemem algebry liniowej jest zrozumienie i opisanie działania endomorfizmów skończone wymiarowych przestrzeni wektorowych. Jeśli $\tau \in \text{End}_K V$, to najważniejszym aspektem działania endomorfizmu τ na przestrzeni V jest działanie τ na podprzestrzeniach przestrzeni V . Dla podprzestrzeni U przestrzeni V obraz $\tau(U)$ jest podprzestrzenią przestrzeni V . Szczególnie ważne dla zrozumienia działania endomorfizmu τ na przestrzeni V są podprzestrzenie U , które τ przeprowadza na siebie.

DEFINICJA 3.0.4. Niech $\tau \in \text{End}_K V$. Podprzestrzeń U przestrzeni V nazywamy podprzestrzenią *niezmienniczą* endomorfizmu τ lub τ -*niezmienniczą*, jeśli $\tau(U) \subseteq U$, to znaczy, jeśli dla każdego $u \in U$ także $\tau(u) \in U$.

Zauważmy, że jeśli U jest podprzestrzenią niezmienniczą endomorfizmu τ przestrzeni V , to zacieśnienie $\tau|_U$ endomorfizmu τ do podprzestrzeni U jest endomorfizmem podprzestrzeni U .

Oczywiście podprzestrzeń zerowa $U = 0$ oraz cała przestrzeń V są podprzestrzeniami niezmienniczymi *każdego* endomorfizmu przestrzeni V . Łatwo stwierdzić, że są to jedyne podprzestrzenie przestrzeni V , które mają tę własność. Podprzestrzeń niezmienniczą endomorfizmu τ nazywamy *nietrywialną*, jeśli jest różna od 0 i V .

Łatwo także wskazać przykład endomorfizmu, który nie ma żadnych nietrywialnych podprzestrzeni niezmienniczych (na przykład, obrót na płaszczyźnie euklidesowej o kąt nie będący całkowitą wielokrotnością kąta półpełnego). Istnienie nietrywialnych podprzestrzeni niezmienniczych endomorfizmu τ jest więc specjalną własnością endomorfizmu τ . W §3.1 rozpatrzmy przypadek, gdy endomorfizm τ ma jednowymiarową podprzestrzeń niezmienniczą. W §3.2 i §3.3 rozpatrujemy wpływ istnienia podprzestrzeni niezmienniczych na macierz endomorfizmu względem odpowiednio dobranej bazy przestrzeni.

W tym rozdziale V oznacza zawsze skończenie wymiarową przestrzeń wektorową nad ciałem K . Będziemy także zawsze zakładać, że V nie jest przestrzenią zerową.

3.1 Wartości własne endomorfizmu

Jako wyjściowy, najłatwiejszy do opisanego typ endomorfizmów uważamy endomorfizmy skalarne postaci $\tau = a1_V$, gdzie $a \in K$. Dla każdego wektora $v \in V$ mamy więc $\tau(v) = av$. Endomorfizm skalarny dokonuje "rozciągania" przestrzeni wektorowej w każdym kierunku w tym samym stopniu.

Jeśli endomorfizm τ nie jest skalarny, to jest rzeczą celową zbadać, jak bardzo różni się on od endomorfizmów skalarnych. Inaczej mówiąc, jeśli endomorfizm $\tau - a1_V$ nie jest

endomorfizmem zerowym dla żadnego $a \in K$, to możemy zapytać, czy nie jest osobliwy dla pewnej wartości $a \in K$. Wtedy endomorfizm τ przynajmniej w jednym kierunku działa tak jak pewien endomorfizm skalarny.

DEFINICJA 3.1.1. Element $a \in K$ nazywa się *wartością własną endomorfizmu τ* przestrzeni wektorowej V , jeśli endomorfizm $\tau - a1_V$ jest osobliwy.

TWIERDZENIE 3.1.1. *Element $a \in K$ jest wartością własną endomorfizmu τ przestrzeni wektorowej V wtedy i tylko wtedy gdy istnieje wektor $v \in V$ taki, że $v \neq 0$ oraz $\tau(v) = av$.*

Dowód. Następujące stwierdzenia są równoważne:

- a jest wartością własną endomorfizmu τ ,
- endomorfizm $\tau - a1_V$ jest osobliwy,
- istnieje wektor $v \in V$ taki, że $v \neq 0$ i $(\tau - a1_V)(v) = 0$,
- istnieje wektor $v \in V$ taki, że $v \neq 0$ i $\tau(v) = av$.

Dowodzi to twierdzenia. □

DEFINICJA 3.1.2. Niech $a \in K$ będzie wartością własną endomorfizmu τ . *Wektorem własnym endomorfizmu τ należącym do wartości własnej a* nazywamy każdy wektor $v \in V$ taki, że

$$v \neq 0 \quad \text{i} \quad \tau(v) = av.$$

Przykład 3.1.1. Jeśli $v \in V$ jest wektorem własnym endomorfizmu τ , to prosta $U = Kv$ jest 1-wymiarową podprzestrzenią niezmienniczą endomorfizmu τ .

Jeśli a jest wartością własną endomorfizmu τ i $v \in V$ jest wektorem własnym endomorfizmu τ należącym do a , to na prostej Kv endomorfizm τ działa tak jak endomorfizm skalarny $a1_V$, to znaczy, $\tau(u) = au$ dla każdego $u \in Kv$.

Dla każdej wartości własnej endomorfizmu τ istnieje oczywiście wektor własny należący do tej wartości własnej. Natomiast istnienia wartości własnych endomorfizmu można oczekiwać tylko w szczególnych sytuacjach. Jedną z nich opisujemy poniżej. Zwykle istnienia wartości własnych endomorfizmu dowodzi się przy użyciu wielomianu charakterystycznego endomorfizmu, którego definicja wymaga pojęcia wyznacznika. Przytoczony tu dowód omija te metody i drastycznie upraszcza klasyczne podejście do problemu istnienia wartości własnych.

TWIERDZENIE 3.1.2. *Niech τ będzie endomorfizmem przestrzeni wektorowej V nad ciałem K . Jeśli wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych, to endomorfizm τ ma przynajmniej jedną wartość własną. Dokładniej, co najmniej jeden z pierwiastków wielomianu p_τ jest wartością własną endomorfizmu τ .*

*Dowód.*¹ Wielomian $p_\tau \in K[X]$ rozkłada się nad K na czynniki liniowe, zatem istnieje rozkład

$$p_\tau = (X - a_1) \cdots (X - a_m), \quad a_1, \dots, a_m \in K.$$

W takim razie dla każdego wektora $v \in V$ mamy

$$0 = 0_V(v) = p_\tau(\tau)(v) = (\tau - a_1 1_V) \cdots (\tau - a_m 1_V)(v).$$

Tak więc iloczyn $(\tau - a_1 1_V) \cdots (\tau - a_m 1_V)$ jest endomorfizmem zerowym. Oznacza to, że przynajmniej jeden z endomorfizmów $\tau - a_i 1_V$ jest osobliwy (gdyby wszystkie były nieosobliwe, to także ich iloczyn byłby nieosobliwy, a jest endomorfizmem zerowym). W takim razie a_i jest wartością własną endomorfizmu τ . □

¹Zobacz Sh. Axler, Down with determinants. *Amer. Math. Monthly* **102** (1995), 139–154.

Twierdzenie 3.1.3. *Jeśli K jest ciałem algebraicznie domkniętym, to każdy endomorfizm $\tau \in \text{End}_K V$ ma przynajmniej jedną wartość własną.*

Dowód. Ponieważ K jest ciałem algebraicznie domkniętym, wielomian p_τ rozkłada się nad K na czynniki liniowe. Istnienie wartości własnej wynika więc z twierdzenia 3.1.2. $\square$

Uwaga 3.1.1. Tak więc nad ciałem liczb zespolonych $K = \mathbb{C}$ każdy endomorfizm ma przynajmniej jedną wartość własną. Natomiast nad ciałem $\mathbb{R}$ liczb rzeczywistych istnieją endomorfizmy, które nie mają wartości własnych. Weźmy wspomniany już na wstępie przykład, czyli płaszczyznę $V = \mathbb{R}^2$ i endomorfizm τ , który jest obrotem o kąt nie będący całkowitą wielokrotnością kąta półpełnego. Wtedy τ nie ma wartości własnych, gdyż nie ma wektorów własnych.

Uwaga 3.1.2. Twierdzenie 3.1.2 i jego dowód przytaczamy przede wszystkim jako ciekawostkę metodologiczną. Faktycznie pełną siłę argumentu użytego w tym dowodzie wykorzystamy w dowodzie twierdzenia 3.1.4 pokazując, że jeśli p_τ ma choć jeden pierwiastek w ciele K , to jest on wartością własną endomorfizmu τ . A więc założenie rozkładalności wielomianu p_τ na czynniki liniowe nad ciałem K okaże się zbędne, wystarczy, że p_τ ma jeden czynnik liniowy nad K .

Lemat 3.1.1. *Niech $\tau \in \text{End}_K V$, $a \in K$, $0 \neq v \in V$ i niech $g \in K[X]$ będzie dowolnym wielomianem. Wtedy*

$$\tau(v) = av \quad \Rightarrow \quad g(\tau)(v) = g(a)v.$$

Dowód. Zauważmy najpierw, że dla każdej liczby naturalnej k ,

jeśli v jest wektorem własnym endomorfizmu τ należącym do wartości własnej a , to v jest wektorem własnym endomorfizmu τ^k należącym do wartości własnej a^k .

Rzeczywiście, jeśli $\tau(v) = av$ dla pewnego niezerowego wektora $v \in V$, to $\tau^2(v) = \tau(\tau(v)) = \tau(av) = a\tau(v) = a^2v$ i łatwa indukcja pokazuje, że $\tau^k(v) = a^k v$. Stąd dla wielomianu $g = c_0 X^m + c_1 X^{m-1} + \dots + c_m$ mamy

$$\begin{aligned} g(\tau)(v) &= c_0 \tau^m(v) + c_1 \tau^{m-1}(v) + \dots + c_m 1_V(v) \\ &= c_0 a^m v + c_1 a^{m-1} v + \dots + c_m v \\ &= g(a)v. \end{aligned}$$

A więc $g(a)$ jest wartością własną endomorfizmu $g(\tau)$ oraz v jest wektorem własnym należącym do wartości własnej $g(a)$. $\square$

Udowodnimy teraz definitywną charakteryzację wartości własnych endomorfizmu jako pierwiastków wielomianu minimalnego endomorfizmu.

Twierdzenie 3.1.4. *Dla $a \in K$ i dla endomorfizmu $\tau \in \text{End}_K V$ następujące warunki są równoważne.*

- (a) *a jest wartością własną endomorfizmu τ .*
- (b) *a jest pierwiastkiem wielomianu minimalnego p_τ endomorfizmu τ .*

Dowód. (a) $\Rightarrow$ (b) Niech $v \in V$ będzie wektorem własnym należącym do wartości własnej a . Zatem $\tau(v) = av$ oraz $v \neq 0$. Na podstawie lematu 3.1.1 mamy $p_\tau(\tau)(v) = p_\tau(a)v$. Ponieważ $p_\tau(\tau) = 0_V$, więc wynika stąd, że $p_\tau(a)v = 0$. Ponieważ zaś wektor v jest niezerowy, otrzymujemy $p_\tau(a) = 0$.

(b) $\Rightarrow$ (a) Jeśli $p_\tau(a) = 0$ oraz $a \in K$, to istnieje wielomian $q \in K[X]$ taki, że $p_\tau = (X - a)q$.

Wtedy wobec $p_\tau(\tau) = 0_V$, endomorfizm $(\tau - a1_V)q(\tau)$ jest endomorfizmem zerowym, ale $q(\tau)$ nie jest endomorfizmem zerowym (gdyż stopień wielomianu q jest mniejszy od stopnia wielomianu minimalnego p_τ). Istnieje więc wektor $v \in V$ taki, że $u := q(\tau)(v) \neq 0$. Wtedy

$$(\tau - a1_V)(u) = (\tau - a1_V)q(\tau)(v) = p_\tau(\tau)(v) = 0,$$

a więc a jest wartością własną endomorfizmu τ . $\square$

WNIOSEK 3.1.1. *Każdy endomorfizm τ ma tylko skończoną liczbę wartości własnych.*

Dowód. Wielomian minimalny p_τ ma tylko skończoną liczbę pierwiastków. $\square$

Uwaga 3.1.3. Na podstawie twierdzenia 2.1.2 wielomian minimalny p_τ endomorfizmu τ przestrzeni n -wymiarowej V ma stopień nie większy niż n^2 (gdyż na podstawie przykładu 2.1.5 wymiar algebry endomorfizmów przestrzeni V jest równy n^2). Zatem wniosek 3.1.1 można wzmocnić stwierdzeniem, że każdy endomorfizm n -wymiarowej przestrzeni wektorowej ma nie więcej niż n^2 wartości własnych. Okazuje się, że łatwo można uzyskać znacznie dokładniejszy rezultat sformułowany we wniosku 3.1.2.

TWIERDZENIE 3.1.5. *Jeśli $a_1, \dots, a_k \in K$ są różnymi wartościami własnymi endomorfizmu τ oraz $v_1, \dots, v_k \in V$ są wektorami własnymi endomorfizmu τ należącymi odpowiednio do wartości własnych $a_1, \dots, a_k$, to wektory $v_1, \dots, v_k$ są liniowo niezależne w przestrzeni V .*

Dowód. Jeśli $v_1, \dots, v_k$ są liniowo zależne, to istnieją skalary $c_1, \dots, c_k$ nie wszystkie równe zero takie, że

$$c_1v_1 + \dots + c_kv_k = 0. \quad (3.1)$$

Faktycznie możemy zakładać, że co najmniej dwa współczynniki spośród $c_1, \dots, c_k$ są różne od zera (gdyż wektory v_i są niezerowe) i wobec tego możemy zakładać, że dla pewnego $i > 1$ mamy $c_i \neq 0$. Biorąc wartości endomorfizmu τ po obydwu stronach otrzymujemy

$$c_1a_1v_1 + \dots + c_ka_kv_k = 0.$$

Mnożąc pierwszą z tych równości przez a_1 i odejmując od niej drugą równość otrzymamy

$$c_2(a_1 - a_2)v_2 + \dots + c_k(a_1 - a_k)v_k = 0.$$

Równość ta pokazuje liniową zależność wektorów $v_2, \dots, v_k$ (gdyż wobec $c_i \neq 0$, także $c_i(a_1 - a_i) \neq 0$). Podobnie z liniowej zależności wektorów $v_2, \dots, v_k$ wydedukujemy liniową zależność wektorów $v_3, \dots, v_k$. Kontynuując dochodzimy do wniosku, że układ jednoelementowy złożony z wektora v_k jest liniowo zależny, sprzeczność (gdyż $v_k \neq 0$).

Drugi dowód. Rozpatrzmy obraz obu stron równości (3.1) przez endomorfizm $(\tau - a_21_V) \cdots (\tau - a_k1_V)$. Wykorzystując fakt, że endomorfizmy $\tau - a_21_V, \dots, \tau - a_k1_V$ są przemienne (zob. twierdzenie 2.1.1), otrzymujemy

$$c_1(a_1 - a_2) \cdots (a_1 - a_k)v_1 = 0.$$

Stąd wynika, że $c_1 = 0$. Podobnie dowodzi się, że $c_j = 0$ dla $j = 2, \dots, k$. $\square$

WNIOSEK 3.1.2. *Każdy endomorfizm τ przestrzeni n -wymiarowej ma co najwyżej n różnych wartości własnych.*

Dowód. Wektory własne należące do różnych wartości własnych endomorfizmu τ są liniowo niezależne, ich liczba nie może więc przekraczać wymiaru przestrzeni. $\square$

WNIOSEK 3.1.3. *Jeśli $\dim_K V = n$ i endomorfizm τ ma n różnych wartości własnych, to istnieje baza przestrzeni V złożona z wektorów własnych endomorfizmu τ .*

Dowód. Na podstawie twierdzenia 3.1.5 endomorfizm τ ma n liniowo niezależnych wektorów własnych. $\square$

3.2 Szczególne postaci macierzy endomorfizmu

Można uważać, że znalezienie macierzy endomorfizmu τ względem jakiejś bazy $\mathcal{B}$ przestrzeni V rozwiązuje już problem opisanego działania endomorfizmu τ na przestrzeni wektorowej V . Jeśli bowiem znamy obrazy elementów bazy $\mathcal{B}$ przestrzeni V , to potrafimy już znaleźć obraz każdego wektora v tej przestrzeni (jeśli tylko znamy współrzędne wektora v w bazie $\mathcal{B}$). Jednakże, jeśli baza $\mathcal{B}$ jest obrana całkowicie przypadkowo, to nie można na ogół z macierzy endomorfizmu odczytać żadnych informacji o geometrycznym charakterze działania endomorfizmu τ . A więc, na przykład, nie można stwierdzić czy endomorfizm zachowuje jakieś proste przestrzeni V , lub ogólniej, czy ma jakieś podprzestrzenie niezmiennicze. Natomiast wybierając bazę przestrzeni V "dopasowaną" do endomorfizmu τ można takie informacje odczytać jednym rzutem oka na macierz endomorfizmu. Nasze główne zadanie polega na tym, żeby dla jak najszerszej klasy przestrzeni wektorowych i ich endomorfizmów wskazać sposób dobierania baz przestrzeni gwarantujący najprostszą postać macierzy endomorfizmu.

A oto przykład takiej sytuacji, kiedy macierz endomorfizmu ma szczególnie prostą postać pozwalającą na pełne zrozumienie działania endomorfizmu na przestrzeni wektorowej.

Twierdzenie 3.2.1. *Jeśli $\dim_K V = n$ oraz endomorfizm τ ma n różnych wartości własnych, to istnieje baza przestrzeni V , względem której endomorfizm τ ma macierz diagonalną.*

Dowód. Niech $a_1, \dots, a_n$ będą wartościami własnymi endomorfizmu τ . Dla każdej wartości własnej a_j wybieramy wektor własny v_j należący do a_j . Na podstawie wniosku 3.1.3, wektory własne $v_1, \dots, v_n$ tworzą bazę $\mathcal{B}$ przestrzeni V . Zatem z równości $\tau(v_j) = a_j v_j$ wynika, że

$$m(\tau, \mathcal{B}) = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & a_n \end{bmatrix}.$$

A więc macierz endomorfizmu τ w bazie $\mathcal{B}$ jest diagonalna. $\square$

Uwaga 3.2.1. W twierdzeniu 3.2.1 mówimy o *różnych* wartościach własnych, by zagwarantować istnienie bazy przestrzeni V złożonej z wektorów własnych endomorfizmu τ . Oczywiście endomorfizm może mieć macierz diagonalną mimo, że nie ma n różnych wartości własnych. Zauważmy bowiem, że endomorfizm τ ma w bazie $\mathcal{B} = \{v_1, \dots, v_n\}$ macierz diagonalną z elementami $a_1, \dots, a_n$ na głównej przekątnej wtedy i tylko wtedy, gdy $a_1, \dots, a_n$ są (niekoniecznie różnymi) wartościami własnymi endomorfizmu τ oraz $v_1, \dots, v_n$ są wektorami własnymi endomorfizmu τ należącymi odpowiednio do wartości własnych $a_1, \dots, a_n$, to znaczy, $\tau(v_j) = a_j v_j$ dla $j = 1, \dots, n$.

Jako drugi przykład macierzy endomorfizmu rozpatrzmy sytuację, gdy endomorfizm τ przestrzeni V ma podprzestrzeń niezmienniczą U . Wtedy zacieśnienie $\tau_1 = \tau|_U$ endomorfizmu τ do podprzestrzeni U jest endomorfizmem podprzestrzeni U . Obieramy jakąkolwiek uporządkowaną bazę $\mathcal{B}_1 = \{v_1, \dots, v_r\}$ podprzestrzeni U i rozpatrujemy macierz $A = m(\tau_1, \mathcal{B}_1)$ endomorfizmu $\tau_1 = \tau|_U$ podprzestrzeni U . Bazę $\mathcal{B}_1$ uzupełniamy następnie do uporządkowanej bazy $\mathcal{B}$ przestrzeni V , a więc

$$\mathcal{B} = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}.$$

Macierz $m(\tau, \mathcal{B})$ ma wtedy postać

$$m(\tau, \mathcal{B}) = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

gdzie B i C są pewnymi macierzami. A więc niezmienniczość podprzestrzeni U względem τ pozwala znaleźć bazę przestrzeni V względem której macierz endomorfizmu τ ma klatkę zer w lewym dolnym rogu o rozmiarach $(n - r) \times r$.

Macierz endomorfizmu τ ma jeszcze prostszą postać, gdy przestrzeń V jest *sumą prostą* dwóch podprzestrzeni niezmienniczych U i W . Obieramy wtedy jakąkolwiek bazę $\mathcal{B}_1 = \{v_1, \dots, v_r\}$ podprzestrzeni U oraz jakąkolwiek bazę $\mathcal{B}_2 = \{v_{r+1}, \dots, v_n\}$ podprzestrzeni W i wobec $V = U \oplus W$ otrzymujemy, że $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ jest bazą przestrzeni V . Ponieważ $\tau(U) \subseteq U$ oraz $\tau(W) \subseteq W$, więc dla $j = 1, \dots, r$ oraz $i = 1, \dots, n - r$ mamy

$$\tau(v_j) \in \text{lin}(\mathcal{B}_1) \quad \text{oraz} \quad \tau(v_{r+i}) \in \text{lin}(\mathcal{B}_2).$$

Zatem macierz endomorfizmu τ w bazie $\mathcal{B}$ ma postać

$$m(\tau, \mathcal{B}) = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix},$$

gdzie $A = m(\tau|_U, \mathcal{B}_1)$ oraz $C = m(\tau|_W, \mathcal{B}_2)$.

Jeśli podprzestrzeń U da się także przedstawić w postaci sumy prostej podprzestrzeni τ -niezmienniczych, to przy odpowiednim wyborze bazy w podprzestrzeni U uzyskamy macierz A endomorfizmu $\tau|_U$ w postaci klatkowej

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

oraz znajdziemy bazę $\mathcal{B}'$ przestrzeni V względem której macierz endomorfizmu τ ma postać

$$\begin{bmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & C \end{bmatrix}.$$

Kontynuując to postępowanie, jeśli jest to możliwe, uzyskujemy macierz endomorfizmu τ coraz bliższą macierzy diagonalnej. Nie należy jednak spodziewać się, że zawsze jest możliwe uzyskanie macierzy diagonalnej. W następnym rozdziale wskażemy warunek konieczny i wystarczający na to by istniała baza przestrzeni V , względem której macierz endomorfizmu τ jest diagonalna. Wynika z niego, że endomorfizmy nie są na ogół diagonalizowalne.

W każdym bądź razie ta obserwacja wskazuje na fundamentalne znaczenie możliwości rozkładu przestrzeni na sumę prostą podprzestrzeni τ -niezmienniczych dla znalezienia macierzy endomorfizmu bliskiej macierzy diagonalnej.

3.3 Podejście klasyczne

Przypomnimy teraz klasyczne podejście do problemu wyznaczania wartości własnych endomorfizmu wykorzystujące pojęcia wyznacznika i wielomianu charakterystycznego macierzy. Zauważamy najpierw, że jeśli $A = [a_{ij}]$ jest macierzą endomorfizmu τ w bazie $\mathcal{B} = \{v_1, \dots, v_n\}$, to

$$\tau(v_j) = \sum_{i=1}^n a_{ij} v_i \quad \text{dla} \quad j = 1, \dots, n$$

i wobec tego dla dowolnego wektora $v \in V$, $v = \sum_{j=1}^n x_j v_j$, mamy

$$\tau(v) = \sum_{j=1}^n x_j \tau(v_j) = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij} v_i = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right) v_i.$$

Jeśli więc $\tau(v) = \sum_{i=1}^n y_i v_i$, to

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad \text{dla } i = 1, \dots, n. \quad (3.2)$$

Jest to zatem związek między współrzędnymi wektora $\tau(v)$ i współrzędnymi wektora v w bazie $\mathcal{B}$. Jeśli współrzędne wektorów traktujemy jako wektory przestrzeni K^n ,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

to relacje (3.2) możemy zapisać w postaci macierzowej następująco:

$$Y = AX.$$

Jest to macierzowy opis działania endomorfizmu τ na przestrzeni V pokazujący jakie współrzędne Y w bazie $\mathcal{B}$ ma obraz wektora v o danych współrzędnych X w bazie $\mathcal{B}$. Zauważamy, że wektor v o współrzędnych X jest wektorem własnym endomorfizmu τ należącym do wartości własnej a wtedy i tylko wtedy gdy $X \neq 0 \in K^n$ oraz $Y = aX$. A więc X jest niezerowym rozwiązaniem następującego układu równań liniowych:

$$aX = AX.$$

Układ ten jest równoważny układowi n równań liniowych jednorodnych o n niewiadomych

$$(A - aI)X = 0,$$

gdzie I oznacza macierz jednostkową stopnia n . Z kursowego wykładu algebry liniowej wiemy, że taki układ ma rozwiązanie niezerowe wtedy i tylko wtedy gdy wyznacznik macierzy współczynników przy niewiadomych jest równy zero:

$$\det(A - aI) = 0.$$

Tutaj skalar $\det(A - aI)$ będziemy traktować jako wartość wielomianu

$$F := \det(A - xI) \in K[x]$$

jednej zmiennej x nad ciałem K w punkcie $x = a$. Wielomian F nazywa się zwykle wielomianem *charakterystycznym* endomorfizmu τ lub macierzy A . W ten sposób otrzymujemy następujące klasyczne kryterium.

TWIERDZENIE 3.3.1. *Niech τ będzie endomorfizmem przestrzeni wektorowej V nad ciałem K i niech F będzie wielomianem charakterystycznym endomorfizmu τ . Skalar $a \in K$ jest wartością własną endomorfizmu τ wtedy i tylko wtedy gdy a jest pierwiastkiem wielomianu charakterystycznego F endomorfizmu τ ,*

$$F(a) = 0.$$

Rozdział 4

Triangularyzacja i diagonalizacja

Ostatnie zmiany 13.11.2008 r.

Zakładamy tak jak w poprzednich rozdziałach, że V jest przestrzenią wektorową nad ciałem K oraz $\dim_K V = n < \infty$.

Postacie kanoniczne macierzy endomorfizmu $\tau \in \text{End}_K V$ zależą w decydującej mierze od wielomianu minimalnego p_τ endomorfizmu τ . Rozpatrzyliśmy już przypadek, gdy wielomian p_τ rozkłada się nad ciałem K na iloczyn $n = \dim_K V$ *różnych* czynników liniowych. Wtedy endomorfizm τ ma n różnych wartości własnych i na podstawie twierdzenia 3.2.1 istnieje baza przestrzeni V , w której τ ma macierz *diagonalną*.

W tym rozdziale rozpatrzymy ogólniejszy przypadek, gdy wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych. A więc nie będziemy wykluczać sytuacji gdy niektóre czynniki liniowe występują w rozkładzie z wykładnikiem > 1 , ani też - w przypadku gdy czynniki są parami różne - nie będziemy wymagać by ich liczba była dokładnie n . Udowodnimy, że wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych wtedy i tylko wtedy gdy istnieje baza przestrzeni V , w której τ ma macierz *trójkątną*.

Nieco trudniejszy do analizy jest przypadek, gdy dodatkowo czynniki liniowe w rozkładzie wielomianu p_τ są jednokrotne (ale ich liczba niekoniecznie jest równa n). Udowodnimy, że wielomian minimalny endomorfizmu τ ma taki rozkład wtedy i tylko wtedy gdy istnieje baza przestrzeni V , w której τ ma macierz *diagonalną*.

Podkreślamy, że założenie o rozkładalności wielomianu p_τ na czynniki liniowe nad ciałem K jest automatycznie spełnione dla wszystkich endomorfizmów jeśli K jest ciałem algebraicznie domkniętym. W szczególności więc udowodnimy, że dla dowolnego endomorfizmu τ przestrzeni wektorowej nad ciałem liczb zespolonych istnieje baza przestrzeni, w której endomorfizm τ ma macierz trójkątną.

Macierz $A = [a_{ij}] \in M_n(K)$ nazywamy macierzą *diagonalną*, jeśli $a_{ij} = 0$ dla $i \neq j$. Piszemy wówczas $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$.

Macierz $A = [a_{ij}] \in M_n(K)$ nazywamy macierzą *trójkątną*, jeśli wszystkie elementy pod główną przekątną są równe zero (to znaczy, jeśli $a_{ij} = 0$ dla $i > j$), lub gdy wszystkie elementy nad główną przekątną są równe zero (to znaczy, $a_{ij} = 0$ dla $i < j$). W pierwszym przypadku mówimy, że A jest górna trójkątna, w drugim, że jest dolna trójkątna.

4.1 Endomorfizm indukowany

Niech τ będzie endomorfizmem przestrzeni V i niech U będzie podprzestrzenią niezmienniczą endomorfizmu τ . Zatem $\tau(u) \in U$ dla każdego wektora $u \in U$. W tej sytuacji dla

dowolnego wektora $v \in V$ mamy

$$\tau(v + U) \subseteq \tau(v) + U.$$

Endomorfizm τ przeprowadza więc warstwę $v + U$ w warstwę $\tau(v) + U$ ale niekoniecznie na warstwę $\tau(v) + U$. Tym niemniej zbiór $\tau(v + U)$ zawiera się w dokładnie jednej warstwie przestrzeni V względem podprzestrzeni niezmienniczej U . Pozwala to na przestrzeni ilorazowej V/U określić odwzorowanie

$$\bar{\tau} : V/U \rightarrow V/U, \quad \bar{\tau}(v + U) = \tau(v) + U.$$

Z łatwością sprawdzamy, że $\bar{\tau}$ jest endomorfizmem przestrzeni wektorowej V/U . Endomorfizm $\bar{\tau}$ nazywamy endomorfizmem *indukowanym* przez endomorfizm τ na przestrzeni ilorazowej V/U .

LEMAT 4.1.1. *Niech τ będzie endomorfizmem przestrzeni V i niech U będzie podprzestrzenią niezmienniczą endomorfizmu τ . Niech $\bar{\tau}$ będzie endomorfizmem indukowanym przez endomorfizm τ na przestrzeni V/U . Wtedy:*

(a) *Jeśli $q \in K[X]$, to $q(\tau)(U) \subseteq U$.*

(b) *Endomorfizmem przestrzeni V/U indukowanym przez $q(\tau)$ jest $q(\bar{\tau})$, to znaczy*

$$\overline{q(\tau)} = q(\bar{\tau}).$$

(c) *Jeśli $q \in K[X]$ oraz $q(\tau) = 0_V$, to $q(\bar{\tau}) = 0_{V/U}$.*

(d) *Wielomian minimalny $p_{\bar{\tau}} \in K[X]$ endomorfizmu $\bar{\tau} \in \text{End}_K V/U$ jest dzielnikiem wielomianu minimalnego p_{τ} endomorfizmu τ .*

Dowód. (a) Dla dowolnego $u \in U$ mamy $\tau^2(u) = \tau(\tau(u)) \in U$, gdyż $\tau(u) \in U$. Podobnie $\tau^i(u) \in U$ dla każdej liczby naturalnej i . Stąd dla $q = \sum c_i X^i \in K[X]$ mamy

$$q(\tau)(u) = \sum c_i \tau^i(u) \in U.$$

(b) Dla dowolnego $v \in V$ i wielomianu $q = \sum c_i X^i \in K[X]$ mamy

$$\begin{aligned} q(\bar{\tau})(v + U) &= \sum c_i \bar{\tau}^i(v + U) \\ &= \sum c_i (\tau^i(v) + U) = \sum c_i \tau^i(v) + U \\ &= q(\tau)(v) + U \\ &= \overline{q(\tau)}(v + U). \end{aligned}$$

(c) Jeśli $q(\tau) = 0_V$, to na podstawie (b) mamy $0_{V/U} = \overline{0_V} = \overline{q(\tau)} = q(\bar{\tau})$.

(d) Ponieważ $p_{\tau}(\tau) = 0_V$, więc na podstawie (c) mamy także $p_{\tau}(\bar{\tau}) = 0_{V/U}$. Zatem wielomian minimalny $p_{\bar{\tau}}$ endomorfizmu $\bar{\tau}$ dzieli wielomian p_{τ} . $\square$

4.2 Triangularyzacja

Endomorfizm τ przestrzeni V nazywamy *triangularyzowalnym*, jeśli istnieje baza przestrzeni V , w której τ ma macierz trójkątną. Jeśli τ ma w bazie $\{v_1, v_2, \dots, v_n\}$ macierz górną trójkątną, to w bazie $\{v_n, \dots, v_2, v_1\}$ ma macierz dolną trójkątną. W związku z tym ograniczymy się do rozpatrywania macierzy górnych trójkątnych i będziemy je nazywać po prostu trójkątnymi. Poniżej wskazujemy warunek wystarczający triangularyzowalności endomorfizmu wyrażony poprzez własności wielomianu minimalnego endomorfizmu. Udowodnimy następnie, że warunek ten jest także konieczny (zob. twierdzenie 4.2.2).

Twierdzenie 4.2.1. *Niech τ będzie endomorfizmem przestrzeni wektorowej V nad ciałem K . Jeśli wielomian minimalny endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych, to przestrzeń V ma bazę, w której macierz endomorfizmu τ jest trójkątna.*

Dowód. Przeprowadzimy dowód indukcyjny ze względu na wymiar n przestrzeni V . Jeśli $n = 1$, to twierdzenie jest trywialnie prawdziwe. Załóżmy więc, że $n > 1$ oraz twierdzenie jest prawdziwe dla endomorfizmów wszystkich przestrzeni wektorowych nad ciałem K o wymiarze $n - 1$.

Niech $\tau \in \text{End}_K V$, $\dim_K V = n$ oraz niech wielomian minimalny p_τ endomorfizmu τ rozkłada się na czynniki liniowe nad ciałem K . Niech $a \in K$ będzie pierwiastkiem wielomianu p_τ . Wtedy na podstawie twierdzenia 3.1.4 skalar a jest wartością własną endomorfizmu τ . Zatem, jeśli u jest wektorem własnym należącym do wartości własnej a , to prosta $U = Ku$ jest 1-wymiarową podprzestrzenią niezmienniczą endomorfizmu τ oraz $\tau(u) = au$.

Rozpatrujemy teraz przestrzeń ilorazową V/U i endomorfizm indukowany $\bar{\tau}$ tej przestrzeni. Ponieważ na podstawie lematu 4.1.1 wielomian minimalny $p_{\bar{\tau}}$ endomorfizmu indukowanego $\bar{\tau}$ jest dzielnikiem wielomianu p_τ , więc wielomian $p_{\bar{\tau}}$ rozkłada się na czynniki liniowe nad ciałem K .

Przestrzeń V/U ma wymiar $n - 1$, wobec tego na podstawie założenia indukcyjnego istnieje baza $\mathcal{B}_1 = \{\bar{v}_2, \dots, \bar{v}_n\}$ przestrzeni V/U , w której endomorfizm $\bar{\tau}$ ma macierz trójkątną:

$$\mathbf{m}(\bar{\tau}, \mathcal{B}_1) = \begin{bmatrix} a_{22} & a_{23} & \dots & a_{2n-1} & a_{2n} \\ 0 & a_{33} & \dots & a_{3n-1} & a_{3n} \\ \vdots & & & & \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}$$

to znaczy dla każdego $j = 2, \dots, n$ mamy

$$\bar{\tau}(\bar{v}_j) = a_{2j}\bar{v}_2 + a_{3j}\bar{v}_3 + \dots + a_{jj}\bar{v}_j.$$

Obieramy teraz wektory $v_j \in V$ tak, by $\bar{v}_j = v_j + U$ dla $j = 2, \dots, n$. Wtedy wektory $u, v_2, \dots, v_n$ tworzą bazę $\mathcal{B}$ przestrzeni V i macierz endomorfizmu τ w tej bazie $\mathbf{m}(\tau, \mathcal{B})$ jest trójkątna. Rzeczywiście,

$$\begin{aligned} \tau(v_j) + U &= \bar{\tau}(\bar{v}_j) \\ &= a_{2j}\bar{v}_2 + \dots + a_{jj}\bar{v}_j \\ &= a_{2j}v_2 + \dots + a_{jj}v_j + U. \end{aligned}$$

Stąd $\tau(v_j) - (a_{2j}v_2 + \dots + a_{jj}v_j) \in U = Ku$. Dla każdego $j = 2, \dots, n$, istnieje więc $a_{1j} \in K$ taki, że

$$\tau(v_j) = a_{1j}u + a_{2j}v_2 + \dots + a_{jj}v_j$$

a to wobec $\tau(u) = au$ oznacza, że endomorfizm τ ma w bazie $\mathcal{B} = \{u, v_2, \dots, v_n\}$ macierz trójkątną. $\square$

Wniosek 4.2.1. *Niech τ będzie endomorfizmem przestrzeni wektorowej V nad ciałem algebraicznie domkniętym K . Wtedy przestrzeń V ma bazę, w której macierz endomorfizmu τ jest trójkątna.*

Możemy teraz udowodnić twierdzenie charakteryzujące endomorfizmy triangularyzowalne poprzez własności ich wielomianów minimalnych.

Twierdzenie 4.2.2. *Endomorfizm $\tau \in \text{End}_K V$ ma macierz trójkątną w pewnej bazie przestrzeni V wtedy i tylko wtedy, gdy wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych:*

$$p_\tau = (X - b_1) \cdots (X - b_k), \quad b_1, \dots, b_k \in K.$$

Dowód. Jedną część tego twierdzenia już udowodniliśmy (twierdzenie 4.2.1). Dla dowodu drugiej części założmy, że $\mathcal{B} = \{v_1, \dots, v_n\}$ jest bazą V w której $\mathbf{m}(\tau, \mathcal{B}) = [a_{ij}]$ jest macierzą trójkątną. Rozpatrzmy wielomian

$$f := (X - a_{11}) \cdots (X - a_{nn}).$$

Jest to iloczyn czynników liniowych nad ciałem K . Udowodnimy, że p_τ dzieli f . Wystarczy zatem pokazać, że $f(\tau) = 0_V$, czyli $f(\tau)(v_j) = 0$ dla każdego $v_j \in \mathcal{B}$. Ponieważ

$$\tau(v_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{jj}v_j$$

dla każdego $j = 1, \dots, n$, więc dla każdego $j > 1$ mamy

$$(\tau - a_{jj}1_V)(v_j) \in \text{lin}(v_1, \dots, v_{j-1}).$$

Ogólniej

$$(\tau - a_{jj}1_V)(\text{lin}(v_1, \dots, v_j)) \subseteq \text{lin}(v_1, \dots, v_{j-1}).$$

Zatem także

$$(\tau - a_{j-1,j-1}1_V)(\tau - a_{jj}1_V)(v_j) \subseteq (\tau - a_{j-1,j-1}1_V)(\text{lin}(v_1, \dots, v_{j-1})) \subseteq \text{lin}(v_1, \dots, v_{j-2})$$

i kontynuując to postępowanie otrzymamy w końcu, że dla dowolnego $v_j \in \mathcal{B}$ mamy

$$(\tau - a_{22}1_V) \cdots (\tau - a_{j-1,j-1}1_V)(\tau - a_{jj}1_V)(v_j) \in \text{lin}(v_1).$$

Ponieważ $(\tau - a_{11}1_V)(v_1) = 0$, więc wobec przemienności endomorfizmów $\tau - a_{11}1_V, \dots, \tau - a_{nn}1_V$ otrzymujemy stąd

$$f(\tau)(v_j) = (\tau - a_{11}1_V) \cdots (\tau - a_{nn}1_V)(v_j) = 0.$$

A więc $f(\tau)(v_j) = 0$ dla każdego $v_j \in \mathcal{B}$ i wobec tego $f(\tau) = 0_V$. Stąd wynika, że p_τ dzieli f i wobec tego rozkłada się na czynniki liniowe nad ciałem K . $\square$

Okazuje się, że dla endomorfizmów triangularizowalnych problem znalezienia wartości własnych sprowadza się do znalezienia ich macierzy trójkątnych. Wynika to z następującej obserwacji.

Twierdzenie 4.2.3. *Niech endomorfizm $\tau \in \text{End}_K V$ ma macierz trójkątną $A = [a_{ij}]$ w pewnej bazie przestrzeni V . Wtedy*

(a) *Wielomian minimalny p_τ endomorfizmu τ jest dzielnikiem wielomianu*

$$(X - a_{11}) \cdots (X - a_{nn}).$$

(b) *Każda wartość własna endomorfizmu τ występuje przynajmniej jeden raz wśród elementów diagonalnych a_{jj} macierzy A .*

(c) *Każdy element diagonalny a_{jj} macierzy A jest wartością własną endomorfizmu τ .*

Dowód. (a) sprawdziliśmy w dowodzie twierdzenia 4.2.2.

(b) Na podstawie (a) mamy $p_\tau \mid (X - a_{11}) \cdots (X - a_{nn})$. Oznacza to, że każdy pierwiastek wielomianu p_τ (a więc każda wartość własna endomorfizmu τ) jest pierwiastkiem wielomianu $(X - a_{11}) \cdots (X - a_{nn})$, a więc jest jednym spośród elementów diagonalnych macierzy A .

(c) Niech $\mathcal{B} = \{v_1, \dots, v_n\}$ będzie bazą V taką, że $m(\tau, \mathcal{B}) = [a_{ij}]$ jest macierzą trójkątną. Ponieważ

$$\tau(v_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{jj}v_j$$

dla każdego $j = 1, \dots, n$, więc $\tau(v_1) = a_{11}v_1$ co oznacza, że a_{11} jest wartością własną endomorfizmu τ . Niech teraz $j > 1$ i niech $U = \text{lin}(v_1, \dots, v_{j-1})$. Wtedy $\tau(U) \subseteq U$ i wobec tego możemy rozpatrzeć endomorfizm indukowany $\bar{\tau} : V/U \rightarrow V/U$ taki, że $\bar{\tau}(v + U) = \tau(v) + U$. Mamy zatem

$$\bar{\tau}(v_j + U) = \tau(v_j) + U = a_{jj}v_j + U = a_{jj}(v_j + U),$$

co oznacza, że a_{jj} jest wartością własną endomorfizmu $\bar{\tau}$. Na podstawie twierdzenia 3.1.4 skalar a_{jj} jest pierwiastkiem wielomianu minimalnego endomorfizmu $\bar{\tau}$, zaś na podstawie lematu 4.1.1(d) wynika stąd, że a_{jj} jest pierwiastkiem wielomianu minimalnego endomorfizmu τ . Zatem a_{jj} jest wartością własną endomorfizmu τ (na podstawie twierdzenia 3.1.4). $\square$

Uwaga 4.2.1. Jakkolwiek twierdzenie 4.2.3 ustala precyzyjnie jakie elementy a_{jj} mogą występować na głównej przekątnej macierzy trójkątnej endomorfizmu triangularyzowanego, to jednak nie rozstrzyga ono *ile razy* na głównej przekątnej występują poszczególne wartości własne endomorfizmu τ . W związku z tym nie potrafimy jeszcze stwierdzić, czy wielomian $f = (X - a_{11}) \cdots (X - a_{nn})$ jest wyznaczony jednoznacznie przez endomorfizm τ . Powrócimy do tego pytania w dalszej części wykładu.

WNIOSEK 4.2.2. *Endomorfizm triangularyzowalny τ z macierzą trójkątną $A = [a_{ij}]$ jest nieosobliwy wtedy i tylko wtedy gdy $a_{11}a_{22} \cdots a_{nn} \neq 0$.*

Dowód. Przypomnijmy, że na podstawie twierdzenia 2.2.1, endomorfizm τ jest odwracalny wtedy i tylko wtedy gdy wyraz wolny wielomianu minimalnego p_τ jest różny od zera, to znaczy gdy $p_\tau(0) \neq 0$. Tymczasem wobec twierdzenia 4.2.3 wielomian minimalny p_τ ma pierwiastek 0 wtedy i tylko wtedy gdy jeden z elementów diagonalnych macierzy A jest równy 0. Zatem

$$p_\tau(0) = 0 \iff a_{11}a_{22} \cdots a_{nn} = 0.$$

Stąd, na podstawie twierdzenia 2.2.1, endomorfizm τ jest odwracalny wtedy i tylko wtedy gdy $a_{11}a_{22} \cdots a_{nn} \neq 0$. $\square$

Wniosek 4.2.2 pokazuje jak bez użycia pojęcia wyznacznika można uzyskać rezultat klasycznie nierozdzielnie związany z teorią wyznaczników.

WNIOSEK 4.2.3. *Niech A będzie macierzą endomorfizmu τ przestrzeni V nad ciałem K w pewnej bazie przestrzeni V . Macierz A jest podobna do macierzy trójkątnej wtedy i tylko wtedy gdy wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn czynników liniowych.*

Dowód. Niech $\mathcal{A}$ będzie bazą przestrzeni V taką, że $m(\tau, \mathcal{A}) = A$. Rozkładalność p_τ na czynniki liniowe nad K jest równoważna istnieniu bazy $\mathcal{B}$ przestrzeni V takiej, że macierz $B = m(\tau, \mathcal{B})$ jest trójkątna (na podstawie twierdzenia 4.2.2). Jeśli S jest macierzą auto-morfizmu przeprowadzającego bazę $\mathcal{A}$ na bazę $\mathcal{B}$, to na podstawie twierdzenia 2.2.4 mamy

$B = S^{-1}AS$. A więc macierz A jest podobna do macierzy trójkątnej B . Z drugiej strony, jeśli macierz $A = \mathbf{m}(\tau, \mathcal{A})$ jest podobna do macierzy trójkątnej B oraz ρ jest endomorfizmem przestrzeni V , który w bazie $\mathcal{A}$ ma macierz $B = \mathbf{m}(\rho, \mathcal{A})$, to endomorfizmy ρ i τ są podobne (twierdzenie 2.2.6). Wobec tego, na podstawie twierdzenia 2.2.5, endomorfizmy ρ i τ mają równe wielomiany minimalne. Ale p_ρ rozkłada się nad K na iloczyn czynników liniowych (na podstawie twierdzenia 4.2.2), zatem p_τ także ma taki rozkład. $\square$

4.3 Diagonalizacja

Endomorfizm τ przestrzeni wektorowej V nazywamy *diagonalizowalnym*, jeśli istnieje baza $\mathcal{B}$ przestrzeni V taka, że macierz $\mathbf{m}(\tau, \mathcal{B})$ jest diagonalna. Twierdzenie 3.2.1 podaje więc warunek wystarczający diagonalizowalności endomorfizmu, który można sformułować następująco: jeśli wielomian minimalny endomorfizmu τ przestrzeni n -wymiarowej ma n różnych pierwiastków w ciele K , to endomorfizm τ jest diagonalizowalny. Okazuje się jednak, że wymaganie by wielomian minimalny miał $n = \dim V$ różnych pierwiastków w ciele K jest zbędne. Istotne jest tylko by wielomian ten rozkładał się na iloczyn *różnych* czynników liniowych nad K . Warunek konieczny i wystarczający diagonalizowalności podajemy w twierdzeniu 4.3.2. Rozpocniemy jednak od wyznaczenia wielomianu minimalnego endomorfizmu diagonalizowalnego.

TWIERDZENIE 4.3.1. *Jeśli endomorfizm $\tau \in \text{End}_K V$ jest diagonalizowalny oraz $b_1, \dots, b_k \in K$ są wszystkimi parami różnymi elementami ciała K występującymi na przekątnej macierzy diagonalnej endomorfizmu τ , to wielomian minimalny p_τ ma postać*

$$p_\tau = (X - b_1) \cdots (X - b_k).$$

Dowód. Niech $\mathcal{B}$ będzie bazą V taką, że $D = \mathbf{m}(\tau, \mathcal{B})$ jest macierzą diagonalną. Wśród elementów ciała K występujących na przekątnej macierzy D niektóre mogą występować więcej niż jeden raz. Niech $b_1, \dots, b_k$ będą wszystkimi parami różnymi elementami ciała K występującymi na przekątnej macierzy D . Rozpatrzmy wielomian

$$f := (X - b_1) \cdots (X - b_k).$$

Udowodnimy, że $f = p_\tau$.

Przed wszystkim $b_1, \dots, b_k$ jako elementy przekątnej diagonalnej macierzy endomorfizmu τ są wartościami własnymi endomorfizmu τ . Na podstawie twierdzenia 3.1.4 elementy $b_1, \dots, b_k$ są pierwiastkami wielomianu minimalnego p_τ i wobec tego f dzieli p_τ .

Z drugiej strony $f(\tau) = 0_V$. Rzeczywiście, weźmy dowolny wektor $v \in \mathcal{B}$ i odpowiadającą mu wartość własną b_i . Zatem $(\tau - b_i 1_V)(v) = 0$ oraz

$$f(\tau)(v) = (\tau - b_1 1_V) \cdots (\tau - b_k 1_V)(v) = 0,$$

gdyż endomorfizmy $\tau - b_1 1_V, \dots, \tau - b_k 1_V$ są przemienne (twierdzenie 2.1.1). A więc $f(\tau)(v) = 0$ dla każdego $v \in \mathcal{B}$ i wobec tego $f(\tau) = 0_V$. Stąd wynika, że p_τ dzieli f . Pokazaliśmy zatem, że $f \mid p_\tau$ oraz $p_\tau \mid f$. Ponieważ obydwa wielomiany f i p_τ są unormowane, wynika stąd, że $f = p_\tau$. $\square$

TWIERDZENIE 4.3.2. *Endomorfizm $\tau \in \text{End}_K V$ jest diagonalizowalny wtedy i tylko wtedy, gdy wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn parami różnych czynników liniowych:*

$$p_\tau = (X - b_1) \cdots (X - b_k), \quad b_1, \dots, b_k \in K, \quad b_i \neq b_j \quad \text{dla} \quad i \neq j.$$

Dowód. Konieczność warunku wynika z twierdzenia 4.3.1, przejdziemy więc do dowodu wystarczalności. Przeprowadzimy dowód indukcyjny ze względu na wymiar przestrzeni. Dla przestrzeni 1-wymiarowych twierdzenie jest oczywiście prawdziwe, zakładamy więc, że $\dim_K V > 1$.

Założmy, że wielomian p_τ ma wskazany rozkład nad ciałem K . Jeśli $k = 1$, to $0_V = p_\tau(\tau) = \tau - b_1 1_V$. Zatem τ jest endomorfizmem skalarnym, $\tau = b_1 1_V$, i w każdej bazie przestrzeni V ma macierz diagonalną. Możemy więc założyć, że $k > 1$.

Niech

$$U := \{v \in V : \tau(v) = b_1 v\} = \ker(\tau - b_1 1_V).$$

Ponieważ b_1 jest pierwiastkiem wielomianu minimalnego endomorfizmu τ , więc jest wartością własną endomorfizmu τ (zob. twierdzenie 3.1.4) i wobec tego do U należą wszystkie wektory własne należące do wartości własnej b_1 . Zatem U jest niezerową podprzestrzenią niezmienniczą endomorfizmu τ i wobec tego $\ell := \dim V/U < \dim V$. Rozpatrujemy endomorfizm $\bar{\tau}$ indukowany na przestrzeni V/U . Pokażemy, że jego wielomian minimalny $p_{\bar{\tau}}$ jest dzielnikiem wielomianu

$$f := (X - b_2) \cdots (X - b_k).$$

Wystarczy pokazać, że $f(\bar{\tau}) = 0_{V/U}$. Przede wszystkim zauważamy, że dla dowolnego $v \in V$ mamy

$$0 = p_\tau(\tau)(v) = (\tau - b_1 1_V)f(\tau)(v),$$

skąd wynika, że $f(\tau)(v) \in \ker(\tau - b_1 1_V) = U$. Zatem dla każdego $v \in V$ mamy

$$f(\bar{\tau})(v + U) = f(\tau)(v) + U = U.$$

Oznacza to, że $f(\bar{\tau}) = 0_{V/U}$ i wobec tego $p_{\bar{\tau}}$ dzieli f . Stąd wynika, że $p_{\bar{\tau}}$ jest iloczynem parami różnych czynników liniowych nad K i wobec tego na podstawie założenia indukcyjnego istnieje baza $\{\bar{v}_1, \dots, \bar{v}_\ell\}$ przestrzeni V/U względem której endomorfizm $\bar{\tau}$ ma macierz diagonalną $\text{diag}(d_1, \dots, d_\ell)$. Przy tym zauważmy, że elementy diagonalne d_j są wartościami własnymi endomorfizmu $\bar{\tau}$, a więc pierwiastkami wielomianu minimalnego $p_{\bar{\tau}}$. Ponieważ ten wielomian jest dzielnikiem wielomianu f , żaden ze skalarów d_j nie jest równy b_1 . Ten fakt wykorzystamy w dalszej części dowodu.

W każdej warstwie bazowej $\bar{v}_j$ wybieramy wektor v_j i wtedy mamy

$$\bar{v}_j = v_j + U, \quad \bar{\tau}(\bar{v}_j) = d_j v_j + U, \quad j = 1, \dots, \ell.$$

Jeśli $\{u_1, \dots, u_m\}$ jest jakąkolwiek bazą podprzestrzeni U , to

$$\mathcal{A} = \{u_1, \dots, u_m, v_1, \dots, v_\ell\}$$

jest bazą przestrzeni V . Niech $A = \mathbf{m}(\tau, \mathcal{A})$ będzie macierzą endomorfizmu τ w bazie $\mathcal{A}$. Ponieważ

$$\tau(u_i) = b_1 u_i \quad \text{dla } i = 1, \dots, m,$$

więc pierwszych m kolumn macierzy A ma identyczną postać jak w macierzy skalarnej $\text{diag}(b_1, \dots, b_1)$. Pozostałe kolumny mają bardziej skomplikowaną postać. Mamy bowiem

$$\tau(v_j) + U = \bar{\tau}(\bar{v}_j) = d_j v_j + U, \quad j = 1, \dots, \ell,$$

zatem $\tau(v_j) - d_j v_j \in U$. Istnieją więc elementy $a_{ij} \in K$ takie, że

$$\tau(v_j) = a_{1j} u_1 + \dots + a_{mj} u_m + d_j v_j, \quad j = 1, \dots, \ell. \quad (4.1)$$

Macierz A ma więc następującą postać:

$$A = m(\tau, \mathcal{A}) = \begin{bmatrix} b_1 & 0 & \dots & 0 & a_{11} & a_{12} & \dots & a_{1\ell} \\ 0 & b_1 & \dots & 0 & a_{21} & a_{22} & \dots & a_{2\ell} \\ & & \dots & & & & \dots & \\ 0 & 0 & \dots & b_1 & a_{m1} & a_{m2} & \dots & a_{m\ell} \\ 0 & 0 & \dots & 0 & d_1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & d_2 & \dots & 0 \\ & & \dots & & & & \dots & \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & d_\ell \end{bmatrix}$$

Pokażemy teraz, że bazę $\mathcal{A}$ można zmodyfikować tak, by w nowej bazie endomorfizm τ miał macierz diagonalną. Przede wszystkim dla każdych $x_{ji} \in K$ zbiór

$$\mathcal{B} := \{u_1, \dots, u_m, v_1 + x_{11}u_1 + \dots + x_{1m}u_m, \dots, v_\ell + x_{\ell 1}u_1 + \dots + x_{\ell m}u_m\}$$

jest bazą przestrzeni V . Udowodnimy teraz, że skalary x_{ji} można tak dobrać, że wszystkie wektory bazy $\mathcal{B}$ są wektorami własnymi endomorfizmu τ . Rzeczywiście, pokażemy, że można tak dobrać $x_{ji} \in K$, że każdy z wektorów $v_j + x_{j1}u_1 + \dots + x_{jm}u_m$ jest wektorem własnym endomorfizmu τ należącym do wartości własnej d_j . Poszukiwane skalary x_{ji} muszą więc spełniać układ równań

$$\tau(v_j + x_{j1}u_1 + \dots + x_{jm}u_m) = d_j(v_j + x_{j1}u_1 + \dots + x_{jm}u_m), \quad j = 1, \dots, \ell.$$

Wykorzystując (4.1) otrzymujemy

$$a_{1j}u_1 + \dots + a_{mj}u_m + d_jv_j + b_1(x_{j1}u_1 + \dots + x_{jm}u_m) = d_j(v_j + x_{j1}u_1 + \dots + x_{jm}u_m),$$

skąd wobec zauważonego już wcześniej faktu, że $d_j \neq b_1$ dla wszystkich $j = 1, \dots, \ell$, mamy

$$x_{j1} = \frac{a_{1j}}{d_1 - b_1}, \dots, x_{jm} = \frac{a_{mj}}{d_j - b_1}.$$

A więc w bazie $\mathcal{B}$ endomorfizm τ ma macierz diagonalną

$$\text{diag}(b_1, \dots, b_1, d_1, \dots, d_\ell),$$

co kończy dowód twierdzenia. □

WNIOSEK 4.3.1. *Niech A będzie macierzą endomorfizmu τ przestrzeni V nad ciałem K w pewnej bazie przestrzeni V . Macierz A jest podobna do macierzy diagonalnej wtedy i tylko wtedy gdy wielomian minimalny p_τ endomorfizmu τ rozkłada się nad ciałem K na iloczyn parami różnych czynników liniowych.*

Dowód. Wystarczy zastosować argumentację użytą w dowodzie wniosku 4.2.3. □

Rozdział 5

Twierdzenie o rozkładzie

Ostatnie zmiany 16.11.2008 r.

W tym rozdziale udowodnimy ogólne twierdzenie o rozkładzie przestrzeni wektorowej na sumę prostą podprzestrzeni niezmienniczych endomorfizmu $\tau \in \mathbf{End}_K V$ wyznaczonych przez czynniki wielomianu minimalnego endomorfizmu τ nierozkładalne nad ciałem K . Twierdzenie to jest podstawą analizy postaci kanonicznych macierzy endomorfizmów i redukuje badanie macierzy endomorfizmu do przypadku, gdy wielomian minimalny endomorfizmu jest potęgą wielomianu nierozkładalnego nad ciałem K .

5.1 Endomorfizmy redukowalne

DEFINICJA 5.1.1. Endomorfizm τ przestrzeni wektorowej V nazywa się *redukowalny*, jeśli istnieją dwie nietrywialne podprzestrzenie niezmiennicze V_1 i V_2 endomorfizmu τ takie, że

$$V = V_1 \oplus V_2.$$

Mówimy także, że endomorfizm τ jest redukowalny przez parę podprzestrzeni niezmienniczych V_1, V_2 .

Zauważmy, że redukowalność endomorfizmu jest własnością silniejszą od istnienia nietrywialnej podprzestrzeni niezmienniczej V_1 endomorfizmu τ . Wprawdzie dla podprzestrzeni niezmienniczej V_1 zawsze można dobrać podprzestrzeń W tak, by $V = V_1 \oplus W$, ale na ogół nie można spodziewać się, że istnieje podprzestrzeń dopełniająca W , która jest *niezmiennicza* względem τ .

Opis działania endomorfizmu redukowalnego sprowadza się do zbadania działania tego endomorfizmu na podprzestrzeniach niezmienniczych V_1, V_2 . Jeśli bowiem $\tau_1 = \tau|_{V_1}$ oraz $\tau_2 = \tau|_{V_2}$ są zacieśnieniami endomorfizmu τ do podprzestrzeni V_1 i V_2 , to są to endomorfizmy V_1 i V_2 , odpowiednio, oraz dla każdego wektora $v \in V$ mamy jednoznaczne przedstawienie w postaci

$$v = v_1 + v_2, \quad v_1 \in V_1, \quad v_2 \in V_2$$

przy czym

$$\tau(v) = \tau_1(v_1) + \tau_2(v_2).$$

Zatem opis działania endomorfizmów τ_1, τ_2 na podprzestrzeniach V_1 i V_2 dostarcza opisu działania endomorfizmu τ na przestrzeni V .

W języku macierzy redukowalność endomorfizmu τ oznacza, że

$$m(\tau, \mathcal{B}) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

gdzie $A = \mathbf{m}(\tau_1, \mathcal{B}_1)$, $B = \mathbf{m}(\tau_2, \mathcal{B}_2)$, $\mathcal{B}_1$ jest dowolną uporządkowaną bazą podprzestrzeni V_1 , $\mathcal{B}_2$ jest dowolną uporządkowaną bazą podprzestrzeni V_2 oraz $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ jest odpowiednią bazą przestrzeni V .

Przechodzimy teraz do badania wielomianu minimalnego endomorfizmu redukowalnego. Dla podprzestrzeni τ –niezmienniczej U przestrzeni V rozpatrywaliśmy już endomorfizm $\bar{\tau}$ indukowany przez τ na przestrzeni ilorazowej V/U (zobacz §4.1). Teraz zajmujemy się endomorfizmem indukowanym przez τ na samej podprzestrzeni niezmienniczej U .

LEMAT 5.1.1. *Niech $\tau \in \text{End}_K V$ i niech $V_1 \subseteq V$ będzie podprzestrzenią niezmienniczą endomorfizmu τ . Wtedy*

- (a) *Odwzorowanie $\tau_1 : V_1 \rightarrow V_1$, $\tau_1(u) = \tau(u)$ jest endomorfizmem przestrzeni V_1 .*
- (b) *Jeśli $q \in K[X]$ oraz $q(\tau) = 0_V$, to $q(\tau_1) = 0_{V_1}$.*
- (c) *Wielomian minimalny endomorfizmu τ_1 dzieli wielomian minimalny endomorfizmu τ .*

Dowód. (a) wynika wprost z definicji podprzestrzeni niezmienniczej endomorfizmu.

(b) Zacieśnienie endomorfizmu $q(\tau)$ do podprzestrzeni V_1 jest endomorfizmem $q(\tau_1)$ przestrzeni V_1 . Stąd dla dowolnego $u \in V_1$ mamy $q(\tau_1)(u) = q(\tau)(u) = 0_V(u) = 0$. Stąd $q(\tau_1) = 0_{V_1}$.

(c) Na podstawie (b) mamy $p_\tau(\tau_1) = 0_{V_1}$. Zatem p_{τ_1} dzieli p_τ na podstawie definicji wielomianu minimalnego endomorfizmu. $\square$

DEFINICJA 5.1.2. Endomorfizm $\tau_1 = \tau|_{V_1}$ podprzestrzeni niezmienniczej V_1 endomorfizmu τ nazywa się endomorfizmem *indukowanym* na V_1 przez endomorfizm τ przestrzeni V .

LEMAT 5.1.2. *Niech $V = V_1 + V_2$, gdzie V_1, V_2 są podprzestrzeniami τ –niezmienniczymi przestrzeni V . Niech τ_1 i τ_2 będą endomorfizmami indukowanymi przez τ na podprzestrzeniach V_1 i V_2 .*

Wtedy wielomian minimalny p_τ endomorfizmu τ jest najmniejszą wspólną wielokrotnością wielomianów minimalnych p_{τ_1} i p_{τ_2} endomorfizmów τ_1 i τ_2 :

$$p_\tau = \text{NWW}(p_{\tau_1}, p_{\tau_2}).$$

Dowód. Wobec $p_\tau(\tau_1) = 0_{V_1}$ oraz $p_\tau(\tau_2) = 0_{V_2}$ wielomian p_τ jest podzielny przez wielomiany p_{τ_1}, p_{τ_2} , zatem także przez ich najmniejszą wspólną wielokrotność $q := \text{NWW}(p_{\tau_1}, p_{\tau_2})$. Z drugiej strony, dla $v_1 \in V_1$ mamy $q(\tau)(v_1) = q(\tau_1)(v_1) = 0$, gdyż wobec $p_{\tau_1} \mid q$ mamy $q(\tau_1) = 0_{V_1}$. Podobnie, dla $v_2 \in V_2$ mamy $q(\tau)(v_2) = 0$.

Wobec tego, dla dowolnego $v \in V = V_1 + V_2$ mamy $v = v_1 + v_2$, gdzie $v_1 \in V_1$, $v_2 \in V_2$ oraz

$$q(\tau)(v) = q(\tau)(v_1) + q(\tau)(v_2) = 0.$$

A więc $q(\tau) = 0_V$ i wobec tego $p_\tau \mid q$. Ponieważ zarówno p_τ jak i q są wielomianami unormowanymi, więc z tego, że $p_\tau \mid q$ i $q \mid p_\tau$ wynika, że $p_\tau = q$. $\square$

TWIERDZENIE 5.1.1. *Niech $V = V_1 + \dots + V_k$, gdzie $V_1, \dots, V_k$ są podprzestrzeniami niezmienniczymi endomorfizmu τ przestrzeni V .*

Niech τ_i będzie endomorfizmem indukowanym przez τ na podprzestrzeni V_i i niech p_{τ_i} będzie wielomianem minimalnym endomorfizmu τ_i dla $i = 1, \dots, k$.

Wtedy wielomian minimalny p_τ endomorfizmu τ jest najmniejszą wspólną wielokrotnością wielomianów $p_{\tau_1}, \dots, p_{\tau_k}$:

$$p_\tau = \text{NWW}(p_{\tau_1}, \dots, p_{\tau_k}).$$

W szczególności, jeśli wielomiany $p_{\tau_1}, \dots, p_{\tau_k}$ są parami względnie pierwsze, to

$$p_\tau = p_{\tau_1} \cdots p_{\tau_k}.$$

Dowód. Indukcja ze względu na k z wykorzystaniem lematu 5.1.2. $\square$

5.2 Twierdzenie o rozkładzie

Twierdzenie 5.1.1 pokazuje, że istnienie rozkładu przestrzeni wektorowej V na sumę prostą podprzestrzeni τ –niezmienniczych ma wyraźny wpływ na postać wielomianu minimalnego endomorfizmu τ . Przystępujemy teraz do dowodu ważnego twierdzenia pokazującego, że także na odwrót, rozkładowi wielomianu minimalnego endomorfizmu τ na iloczyn parami względnie pierwszych czynników odpowiada rozkład przestrzeni V na sumę prostą podprzestrzeni τ –niezmienniczych.

Dla endomorfizmu $\tau \in \text{End}_K V$ będziemy zakładać, że wielomian minimalny p_τ endomorfizmu τ ma rozkład

$$p_\tau = p_1 \cdots p_k, \quad (5.1)$$

gdzie $p_i \in K[X]$ są unormowanymi wielomianami stopni ≥ 1 i wielomiany te są *parami względnie pierwsze*, to znaczy

$$\text{NWD}(p_i, p_j) = 1 \quad \text{dla} \quad i \neq j.$$

Najważniejszym przykładem takiego rozkładu wielomianu p_τ jest *rozkład kanoniczny*

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k}, \quad (5.2)$$

gdzie q_i są różnymi unormowanymi wielomianami nierozkładalnymi nad ciałem K oraz m_i są liczbami naturalnymi. Każdy wielomian $p \in K[X]$ ma tylko jeden rozkład kanoniczny (jeśli rozkłady różniące się tylko porządkiem czynników nie uważać za różne).

W poniższym twierdzeniu pokażemy, że z rozkładem wielomianu p_τ na iloczyn czynników parami względnie pierwszych wiąże się zawsze rozkład przestrzeni V na sumę prostą niezerowych podprzestrzeni niezmienniczych endomorfizmu τ .

Twierdzenie 5.2.1. (Twierdzenie o rozkładzie).

Niech $\tau \in \text{End}_K V$ i niech $p_\tau \in K[X]$ będzie wielomianem minimalnym endomorfizmu τ . Niech (5.1) będzie rozkładem wielomianu p_τ nad ciałem K na iloczyn wielomianów unormowanych, parami względnie pierwszych, stopni ≥ 1 i niech $k \geq 2$. Dla każdego $i = 1, \dots, k$ niech

$$V_i := \ker p_i(\tau) = \{v \in V : p_i(\tau)(v) = 0\}.$$

Wtedy

- (a) V_i jest niezerową podprzestrzenią niezmienniczą endomorfizmu τ .
- (b) $V = V_1 \oplus \cdots \oplus V_k$.
- (c) Endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny p_i .

Dowód. (a) V_i jest jądrem endomorfizmu przestrzeni V , zatem jest podprzestrzenią przestrzeni V .

Łatwo sprawdza się, że V_i jest podprzestrzenią niezmienniczą endomorfizmu τ . Dla dowolnego $v \in V_i$ mamy bowiem $p_i(\tau)(v) = 0$, zatem wykorzystując przemienność endomorfizmów $p_i(\tau)$ i τ mamy także

$$p_i(\tau)(\tau(v)) = (p_i(\tau)\tau)(v) = (\tau p_i(\tau))(v) = \tau(p_i(\tau)(v)) = \tau(0) = 0.$$

Stąd wynika, że $\tau(v) \in V_i$, co dowodzi niezmienniczości podprzestrzeni V_i .

Dla dowodu, że $V_i \neq 0$, dla każdego $i = 1, \dots, k$ bierzemy wielomian

$$h_i = \prod_{\substack{j=1 \\ j \neq i}}^k p_j = \frac{p_\tau}{p_i}.$$

Ponieważ h_i ma mniejszy stopień niż wielomian p_τ , więc $h_i(\tau) \neq 0_V$. Istnieje więc wektor $v \in V$ taki, że $u := h_i(\tau)(v) \neq 0$. Wtedy mamy

$$p_i(\tau)(u) = p_i(\tau)h_i(\tau)(v) = p_\tau(\tau)(v) = 0_V(v) = 0.$$

A więc $u \in \ker p_i(\tau) = V_i$ oraz $u \neq 0$.

(b) Zauważmy, że wielomiany $h_1, \dots, h_k$ są względnie pierwsze. Jeśli bowiem wielomian nierozkładalny q dzieli h_i , to q dzieli pewien wielomian p_j dla $j \neq i$. Ale wtedy q nie dzieli h_j . Wielomiany $h_1, \dots, h_k$ nie mogą więc mieć wspólnego dzielnika nierozkładalnego q . Ponieważ $K[X]$ jest pierścieniem ideałów głównych wynika stąd, że istnieją wielomiany $f_1, \dots, f_k \in K[X]$ takie, że

$$f_1 h_1 + \dots + f_k h_k = 1.$$

W algebrze endomorfizmów $\text{End}_K V$ mamy więc równość

$$f_1(\tau)h_1(\tau) + \dots + f_k(\tau)h_k(\tau) = 1_V$$

i wobec tego dla dowolnego wektora $v \in V$ mamy

$$f_1(\tau)h_1(\tau)(v) + \dots + f_k(\tau)h_k(\tau)(v) = v.$$

Kładąc $v_i := f_i(\tau)h_i(\tau)(v)$ otrzymujemy

$$v_1 + \dots + v_k = v.$$

Zauważmy, że $v_i \in V_i$. Mamy bowiem

$$p_i(\tau)(v_i) = (p_i f_i h_i)(\tau)(v) = (f_i p_\tau)(\tau)(v) = 0_V(v) = 0.$$

Udowodniliśmy więc, że $V = V_1 + \dots + V_k$. Pozostaje udowodnić, że każdy wektor $v \in V$ ma *jednoznaczne* przedstawienie w postaci $v = v_1 + \dots + v_k$, gdzie $v_i \in V_i$.

Wystarczy pokazać, że wektor zerowy ma jednoznaczne przedstawienie, to znaczy, że z równości $0 = v_1 + \dots + v_k$ gdzie $v_i \in V_i$ wynika, że wszystkie $v_i = 0$.

Przypuśćmy więc, że $0 = v_1 + \dots + v_k$, gdzie $v_i \in V_i$. Wtedy dla każdego j mamy

$$0 = h_j(\tau)(0) = h_j(\tau)(v_1) + \dots + h_j(\tau)(v_k).$$

Dla $i \neq j$ wielomian h_j dzieli się przez p_i , zatem $h_j(\tau)(v_i) = 0$. A więc nasza równość redukuje się do

$$0 = h_j(\tau)(v_j).$$

Z tego, że wielomiany p_i są parami względnie pierwsze wynika, że wielomiany h_j i p_j są względnie pierwsze. Jeśli więc $g_1 h_j + g_2 p_j = 1$, gdzie $g_1, g_2 \in K[X]$, to także $g_1(\tau)h_j(\tau) + g_2(\tau)p_j(\tau) = 1_V$, skąd

$$g_1(\tau)h_j(\tau)(v_j) + g_2(\tau)p_j(\tau)(v_j) = v_j.$$

Obydwa składniki po lewej stronie tej równości są wektorami zerowymi, zatem także $v_j = 0$. Dowodzi to (b).

(c) Z określenia podprzestrzeni V_i wynika, że dla każdego $v \in V_i$ mamy

$$p_i(\tau_i)(v) = p_i(\tau)(v) = 0.$$

Innymi słowy $p_i(\tau_i) = 0_{V_i}$, co oznacza, że wielomian minimalny p_{τ_i} endomorfizmu indukowanego τ_i jest dzielnikiem wielomianu p_i . Ponieważ wielomiany p_i są parami względnie pierwsze, także wielomiany p_{τ_i} są parami względnie pierwsze. Na podstawie twierdzenia

5.1.1 i udowodnionego już punktu (b), wielomian minimalny endomorfizmu τ jest najmniejszą wspólną wielokrotnością wielomianów p_{τ_i} , a więc ich iloczynem. Stąd

$$\prod_{i=1}^k p_{\tau_i} = p_{\tau} = \prod_{i=1}^k p_i. \quad (5.3)$$

Niech m_i będzie stopniem wielomianu p_{τ_i} zaś n_i stopniem wielomianu p_i . Ponieważ $p_{\tau_i} \mid p_i$, więc $m_i \leq n_i$. Z drugiej strony, wobec równości (5.3), mamy

$$m_1 + \cdots + m_k = n_1 + \cdots + n_k$$

i wobec tego $m_i = n_i$ dla $i = 1, \dots, k$. Zatem także $p_{\tau_i} = p_i$ dla $i = 1, \dots, k$. □

Zapiszmy jeszcze raz twierdzenie o rozkładzie w szczególnym przypadku gdy rozkład (5.1) wielomianu minimalnego p_{τ} jest rozkładem kanonicznym (5.2).

Twierdzenie 5.2.2. (Twierdzenie o rozkładzie).

Niech $\tau \in \text{End}_K V$ i niech $p_{\tau} \in K[X]$ będzie wielomianem minimalnym endomorfizmu τ . Niech (5.2) będzie rozkładem kanonicznym wielomianu p_{τ} nad ciałem K i niech $k \geq 2$. Dla każdego $i = 1, \dots, k$ niech

$$V_i := \ker q_i^{m_i}(\tau) = \{v \in V : q_i^{m_i}(\tau)(v) = 0\}.$$

Wtedy

- (a) V_i jest niezerową podprzestrzenią niezmienniczą endomorfizmu τ .
- (b) $V = V_1 \oplus \cdots \oplus V_k$.
- (c) Endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny $p_{\tau_i} = q_i^{m_i}$.

Twierdzenie 5.2.1 o rozkładzie ma następującą interpretację macierzową. W oznaczeniach tego twierdzenia, niech $\mathcal{B}_i$ będzie uporządkowaną bazą podprzestrzeni V_i i niech $B_i = \mathbf{m}(\tau_i, \mathcal{B}_i)$ będzie macierzą endomorfizmu indukowanego τ_i podprzestrzeni V_i . Wtedy wobec części (b) twierdzenia zbiór

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$$

jest bazą przestrzeni V i endomorfizm τ przestrzeni V ma w tej bazie macierz klatkową

$$\mathbf{m}(\tau, \mathcal{B}) = \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{bmatrix}. \quad (5.4)$$

Jeśli $\dim V_i = n_i$, to $B_i \in M_{n_i}(K)$ oraz $n = \dim V = n_1 + \cdots + n_k$. Dalsze upraszczanie macierzy endomorfizmu τ sprowadza się do poszukiwania baz w przestrzeniach V_i , względem których endomorfizmy τ_i mają jak najprostszą macierz. Przy tym, jeśli rozkład (5.1) wielomianu p_{τ} był rozkładem kanonicznym (5.2), to twierdzenie o rozkładzie jest już nieprzydatne gdyż wielomian minimalny endomorfizmu τ_i jest potęgą wielomianu nierozkładalnego (i wobec tego nie ma rozkładu na iloczyn czynników parami względnie pierwszych).

Przykład 5.2.1. Jeśli endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny $p_{\tau} = (X - a)^m$, to dla endomorfizmu $\sigma = \tau - a1_V$ mamy

$$0_V = p_{\tau}(\tau) = (\tau - a1_V)^m = \sigma^m.$$

A więc endomorfizm σ jest *nilpotentny*.

Wynika stąd, że dla znalezienia stosownej postaci kanonicznej macierzy endomorfizmu τ , którego wielomian minimalny jest potęgą wielomianu liniowego, wystarczy zbadać postać kanoniczną macierzy endomorfizmów nilpotentnych. Jeśli bowiem $m(\sigma, \mathcal{B}) = A$, to

$$m(\tau, \mathcal{B}) = m(\sigma + a1_V, \mathcal{B}) = m(\sigma, \mathcal{B}) + m(a1_V, \mathcal{B}) = A + aI$$

gdzie I jest macierzą jednostkową. A więc macierz endomorfizmu τ otrzymujemy w bardzo prosty sposób z macierzy A endomorfizmu nilpotentnego $\sigma = \tau - a1_V$. Ta sytuacja motywuje nasze zainteresowanie endomorfizmami nilpotentnymi.

Jako przykład zastosowania twierdzenia o rozkładzie podamy jeszcze jeden dowód warunku wystarczającego diagonalizowalności endomorfizmu (zob. twierdzenie 4.3.2). Dowód, że warunek ten jest konieczny wskazaliśmy w twierdzeniu 4.3.1.

Twierdzenie 5.2.3. *Jeśli wielomian minimalny p_τ endomorfizmu τ przestrzeni V rozkłada się nad ciałem K na iloczyn parami różnych czynników liniowych:*

$$p_\tau = (X - b_1) \cdots (X - b_k), \quad b_1, \dots, b_k \in K, \quad b_i \neq b_j \quad \text{dla} \quad i \neq j,$$

to endomorfizm τ jest diagonalizowalny.

Dowód. Załóżmy, że wielomian p_τ ma wskazany rozkład nad ciałem K . Niech $V_i = \ker(\tau - b_i 1_V)$. Na podstawie twierdzenia o rozkładzie V_i jest niezerową podprzestrzenią τ -niezmienniczą przestrzeni V oraz $V = V_1 \oplus \cdots \oplus V_k$. Zauważmy, że każdy niezerowy wektor podprzestrzeni V_i jest wektorem własnym endomorfizmu τ należącym do wartości własnej b_i . Jeśli $\mathcal{B}_i$ jest dowolną bazą podprzestrzeni V_i , to endomorfizm indukowany τ_i podprzestrzeni V_i ma względem tej bazy macierz diagonalną (faktycznie skalarną) $B_i = b_i I_i$, gdzie I_i oznacza macierz jednostkową stopnia $\dim V_i$. Zatem macierz klatkowa (5.4) jest diagonalna. $\square$

W dalszym ciągu zajmować się będziemy głównie przypadkiem, gdy endomorfizm $\tau \in \text{End}_K V$ przestrzeni wektorowej V ma wielomian minimalny postaci q^m , gdzie q jest wielomianem nierozkładalnym nad ciałem K . Jeśli uda się opisać zadowalająco macierze wszystkich endomorfizmów tego typu, to przypadek ogólny opisany będzie przez zastosowanie twierdzenia o rozkładzie. Jest to strategia prowadząca do głównych twierdzeń o postaci kanonicznej macierzy endomorfizmu takich jak postać kanoniczna Jordana i postać kanoniczna wymierna.

Rozdział 6

Endomorfizmy nilpotentne

Ostatnie zmiany 12.12.2008 r.

6.1 Podprzestrzenie cykliczne

Wprowadzimy najpierw pojęcie podprzestrzeni τ -cyklicznej przestrzeni wektorowej V dla dowolnego endomorfizmu τ przestrzeni V .

DEFINICJA 6.1.1. Niech $\tau \in \text{End}_K V$ będzie dowolnym endomorfizmem. Podprzestrzenia τ -cykliczną generowaną przez wektor $v \in V$ nazywamy podprzestrzeń U przestrzeni V generowaną przez zbiór

$$\{\tau^k(v) : k \in \mathbb{N} \cup \{0\}\}.$$

Podprzestrzeń τ -cykliczna generowana przez wektor $v \in V$ jest więc zbiorem wszystkich kombinacji liniowych postaci

$$a_0 v + a_1 \tau(v) + a_2 \tau^2(v) + \cdots + a_n \tau^n(v), \quad a_0, a_1, \dots, a_n \in K, \quad n \in \mathbb{N}.$$

Inaczej mówiąc, są to wektory postaci

$$f(\tau)(v) \quad \text{gdzie} \quad f = a_0 + a_1 X + \cdots + a_n X^n \in K[X].$$

Dla endomorfizmu τ obraz $\tau(v)$ wektora v zapisuje się także czasem jako τv i wobec tego dla $f \in K[X]$ zamiast $f(\tau)(v)$ pisze się $f(\tau)v$. Używając tego uproszczonego zapisu możemy podprzestrzeń τ -cykliczną U generowaną przez wektor $v \in V$ opisać jako zbiór

$$K[\tau]v := \{f(\tau)v \in V : f \in K[X]\}.$$

LEMAT 6.1.1. Każda podprzestrzeń τ -cykliczna przestrzeni V jest podprzestrzenią niezmienniczą endomorfizmu τ .

Dowód. Dla $k \in \mathbb{N} \cup \{0\}$ mamy $\tau(\tau^k(v)) = \tau^{k+1}(v) \in U$. Zatem τ przeprowadza wszystkie wektory rozpinające podprzestrzeń U na wektory należące do U skąd wynika, że $\tau(U) \subseteq U$. $\square$

TWIERDZENIE 6.1.1. Niech $U = K[\tau]v$ będzie podprzestrzenią τ -cykliczną przestrzeni V generowaną przez wektor $v \in V$ i niech $\tau_1 = \tau|_U$ będzie endomorfizmem indukowanym na U przez endomorfizm τ . Niech $p_{\tau_1} \in K[X]$ będzie wielomianem minimalnym endomorfizmu τ_1 .

- (a) Dla dowolnego wielomianu $f \in K[X]$, jeśli $f(\tau)(v) = 0$, to $p_{\tau_1} \mid f$.
 (b) Jeśli p_{τ_1} ma stopień r , to zbiór

$$\mathcal{B} := \{v, \tau(v), \dots, \tau^{r-1}(v)\}$$

jest bazą podprzestrzeni U i wobec tego $\dim U = r = \deg p_{\tau_1}$.

Dowód. (a) Dla dowolnego wektora $u = g(\tau)(v) \in U$, gdzie $g \in K[X]$, mamy

$$\begin{aligned} f(\tau)(u) &= f(\tau)(g(\tau)(v)) \\ &= (f(\tau)g(\tau))(v) = (g(\tau)f(\tau))(v) = g(\tau)(f(\tau)(v)) = g(\tau)(0) = 0 \in U. \end{aligned}$$

Zatem $f(\tau_1) = f(\tau)|_U = 0_U$ i wobec tego wielomian minimalny p_{τ_1} endomorfizmu τ_1 dzieli f .

(b) Weźmy dowolny wektor $u \in U$. Wtedy $u = f(\tau)(v)$ dla pewnego wielomianu $f \in K[X]$. Na podstawie twierdzenia o dzieleniu z resztą istnieją wielomiany $g, h \in K[X]$ takie, że

$$f = gp_{\tau_1} + h, \quad \deg h < \deg p_{\tau_1} = r.$$

Zauważmy, że wobec $v \in U$ oraz

$$p_{\tau_1}(\tau)(v) = p_{\tau_1}(\tau)|_U(v) = p_{\tau_1}(\tau_1)(v) = 0_U(v) = 0$$

mamy

$$u = f(\tau)(v) = g(\tau)(p_{\tau_1}(\tau)(v)) + h(\tau)(v) = h(\tau)(v).$$

Ponieważ $\deg h < r$ wynika stąd, że dowolny wektor $u \in U$ jest kombinacją liniową wektorów $v, \tau(v), \dots, \tau^{r-1}(v)$. Stąd wynika, że zbiór $\mathcal{B}$ rozpina podprzestrzeń U .

Udowodnimy teraz, że zbiór $\mathcal{B}$ jest liniowo niezależny. Jeśli bowiem

$$a_0v + a_1\tau(v) + \dots + a_{r-1}\tau^{r-1}(v) = 0$$

dla pewnych $a_0, a_1, \dots, a_{r-1} \in K$ nie wszystkich równych zero, to wielomian

$$f = a_0 + a_1X + \dots + a_{r-1}X^{r-1} \in K[X]$$

jest niezerowy oraz $f(\tau)(v) = 0$. Stąd zaś na podstawie (a) wynika, że $p_{\tau_1} \mid f$, co jest niemożliwe, gdyż stopień wielomianu f jest równy co najwyżej $r - 1$. A więc wektory $v, \tau(v), \dots, \tau^{r-1}(v)$ są liniowo niezależne.

Zbiór $\mathcal{B}$ rozpina przestrzeń U i jest liniowo niezależny, zatem jest bazą podprzestrzeni U . Stąd $\dim U = r = \deg p_{\tau_1}$. $\square$

WNIOSEK 6.1.1. Niech $\tau \in \text{End}_K V$. Jeśli $V = K[\tau]v$ jest przestrzenią τ -cykliczną, to $\dim V = \deg p_\tau$.

Dowód. W oznaczeniach twierdzenia 6.1.1 mamy $U = V$, $\tau_1 = \tau$, $p_{\tau_1} = p_\tau$. $\square$

6.2 Endomorfizmy nilpotentne

Przejdziemy teraz do badania endomorfizmów nilpotentnych. Przypomnijmy, że endomorfizm σ przestrzeni V nazywa się *nilpotentny*, jeśli istnieje liczba naturalna m taka, że $\sigma^m = 0_V$.

DEFINICJA 6.2.1. Niech $\sigma \in \text{End}_K V$ będzie niezerowym endomorfizmem nilpotentnym. Jeśli $\sigma^m = 0_V$ i $\sigma^{m-1} \neq 0_V$, to liczbę naturalną m nazywamy *stopniem* endomorfizmu nilpotentnego σ .

LEMAT 6.2.1. $\sigma \in \text{End}_K V$ jest endomorfizmem nilpotentnym stopnia m wtedy i tylko wtedy gdy wielomianem minimalnym endomorfizmu σ jest $p_\sigma = X^m$.

Dowód. Jeśli $p_\sigma = X^m$ to $\sigma^m = p_\sigma(\sigma) = 0_V$ czyli endomorfizm σ jest nilpotentny. Ponadto $\sigma^{m-1} \neq 0_V$, gdyż w przeciwnym razie mielibyśmy $p_\sigma \mid X^{m-1}$, co jest niemożliwe. Z drugiej strony, jeśli σ jest endomorfizmem nilpotentnym stopnia m , to wielomian $X^m \in K[X]$ zeruje się na σ , zatem $p_\sigma \mid X^m$. A więc $p_\sigma = X^\ell$ gdzie $\ell \leq m$. Gdyby $\ell < m$ to wobec $\sigma^\ell = p_\sigma(\sigma) = 0_V$ mielibyśmy także $\sigma^{m-1} = 0_V$ wbrew temu, że σ jest nilpotentny stopnia m . Zatem $\ell = m$ oraz $p_\sigma = X^m$. $\square$

Jeśli σ jest endomorfizmem nilpotentnym stopnia $m > 1$, to podprzestrzeń σ -cykliczna $K[\sigma]v = \text{lin}(\sigma^k(v) : k \in \mathbb{N} \cup \{0\})$ generowana przez wektor $v \in V$ jest faktycznie rozpięta na skończonym zbiorze wektorów $v, \sigma(v), \dots, \sigma^{m-1}(v)$. Przy tym, wobec $\sigma^{m-1} \neq 0_V$ istnieje wektor $v \in V$ taki, że $\sigma^{m-1}(v) \neq 0$. Będziemy rozpatrywać wyłącznie podprzestrzenie σ -cykliczne generowane przez wektor $v \in V$ spełniający ten właśnie warunek. Znaczenie tego warunku wyjaśnia następujący lemat.

LEMAT 6.2.2. Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech $U_1 = K[\sigma]v = \text{lin}(v, \sigma(v), \dots, \sigma^{m-1}(v))$ będzie podprzestrzenią σ -cykliczną przestrzeni V generowaną przez wektor $v \in V$. Wtedy

$$\sigma^{m-1}(v) \neq 0 \iff \dim U_1 = m.$$

Dowód. Jeśli $\sigma^{m-1}(v) \neq 0$, to zauważamy, że endomorfizm indukowany $\sigma_1 = \sigma|_{U_1}$ jest endomorfizmem nilpotentnym stopnia m przestrzeni U_1 . Na podstawie lematu 6.2.1 otrzymujemy, że wielomian minimalny p_{σ_1} endomorfizmu σ_1 jest równy X^m . Na podstawie twierdzenia 6.1.1 otrzymujemy więc $\dim U_1 = \deg p_{\sigma_1} = m$.

Natomiast jeśli $\sigma^{m-1}(v) = 0$, to także $\sigma^k(v) = 0$ dla każdego $k \geq m$ i wobec tego

$$U_1 = \text{lin}(v, \sigma(v), \dots, \sigma^{m-2}(v))$$

jest podprzestrzenią rozpiętą na układzie $m-1$ wektorów a zatem ma wymiar nie większy niż $m-1$. $\square$

Jak już zauważyliśmy, jeśli σ jest endomorfizmem nilpotentnym stopnia m , to podprzestrzeń σ -cykliczna $K[\sigma]v$ jest rozpięta na układzie m wektorów $v, \sigma(v), \dots, \sigma^{m-1}(v)$ i wobec tego ma wymiar $\leq m$. Podprzestrzenie σ -cykliczne o wymiarze m są więc *maksymalnymi* podprzestrzeniami σ -cyklicznymi i na podstawie lematu 6.2.2 podprzestrzeń σ -cykliczna $U_1 = K[\sigma]v$ jest maksymalna wtedy i tylko wtedy gdy $\sigma^{m-1}(v) \neq 0$.

Lemat 6.2.2 pozwala także znaleźć optymalne oszacowanie dla stopnia wielomianu minimalnego endomorfizmu nilpotentnego. Przypomnijmy, że dla dowolnego endomorfizmu τ przestrzeni n -wymiarowej mamy jak dotąd tylko (trywialne) oszacowanie $\deg p_\tau \leq n^2$ (zob. uwagę 3.1.3). Jak wynika z następującego wniosku, dla endomorfizmu nilpotentnego σ stopnia m mamy $m = \deg p_\sigma \leq n$.

WNIOSEK 6.2.1. Jeśli σ jest endomorfizmem nilpotentnym stopnia m przestrzeni n -wymiarowej V , to $m \leq n$.

Dowód. Na podstawie lematu 6.2.2 przestrzeń V ma podprzestrzeń σ -cykliczną U_1 o wymiarze m . Zatem $m \leq n$. $\square$

LEMAT 6.2.3. Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech σ_1 będzie zacieśnieniem σ do maksymalnej podprzestrzeni cyklicznej $U_1 = K[\sigma]v$. Wtedy

$$\mathcal{B}_1 = \{v, \sigma(v), \dots, \sigma^{m-1}(v)\}$$

jest bazą przestrzeni U_1 oraz

$$\mathbf{m}(\sigma_1, \mathcal{B}_1) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ & & & \dots & & \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} =: J_m.$$

Dowód. m -elementowy zbiór $\mathcal{B}_1$ rozpina U_1 oraz $m = \dim U_1$ na podstawie lematu 6.2.2. Zatem $\mathcal{B}_1$ jest bazą podprzestrzeni U_1 . Ponadto, dla $0 \leq i \leq m-2$ mamy $\sigma(\sigma^i(v)) = \sigma^{i+1}(v)$ oraz $\sigma(\sigma^{m-1}(v)) = 0$. $\square$

DEFINICJA 6.2.2. Macierz J_m nazywamy *osobliwą klatką Jordana* stopnia m .

Z wniosku 6.2.1 wiemy, że stopień nilpotentności endomorfizmu $\sigma \in \text{End}_K V$ jest nie większy od wymiaru przestrzeni V . Zauważmy, że w szczególnym przypadku gdy stopień nilpotentności endomorfizmu σ jest równy wymiarowi przestrzeni V , maksymalna podprzestrzeń σ -cykliczna U_1 pokrywa się z V . Zatem jeśli $\sigma^n = 0_V$ i $\sigma^{n-1} \neq 0_V$, gdzie $n = \dim V$, to w przestrzeni V istnieje baza $\mathcal{B}$ taka, że $\mathbf{m}(\sigma, \mathcal{B}) = J_n$. Klatkę Jordana J_n nazywamy wtedy *postacią kanoniczną Jordana* macierzy endomorfizmu nilpotentnego σ .

Przystępujemy teraz do opisu postaci kanonicznej Jordana macierzy dowolnego endomorfizmu nilpotentnego. Ostateczny rezultat podaje twierdzenie 6.3.2, natomiast główny argument potrzebny dla dowodu tego twierdzenia jest zawarty w twierdzeniu 6.2.1.

LEMAT 6.2.4. Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech σ_1 będzie zacieśnieniem σ do maksymalnej podprzestrzeni cyklicznej $U_1 = K[\sigma]v$.

(a) Dla $u \in U_1$ oraz $0 < k \leq m$ mamy

$$\sigma^{m-k}(u) = 0 \iff u = \sigma^k(u_0) \quad \text{dla pewnego wektora } u_0 \in U_1.$$

(b) $\ker \sigma_1^{m-k} = \text{im } \sigma_1^k$.

Dowód. (a) Oczywiście, jeśli $u = \sigma^k(u_0)$, to $\sigma^{m-k}(u) = \sigma^m(u_0) = 0$. Załóżmy teraz, że $\sigma^{m-k}(u) = 0$ dla pewnego $u \in U_1$. Wtedy

$$u = a_0 v + a_1 \sigma(v) + \dots + a_{m-1} \sigma^{m-1}(v)$$

dla pewnych $a_i \in K$ i stąd

$$0 = \sigma^{m-k}(u) = a_0 \sigma^{m-k}(v) + a_1 \sigma^{m-k+1}(v) + \dots + a_{k-1} \sigma^{m-1}(v),$$

gdyż wyższe potęgi σ anihilują wektor v . Na podstawie lematu 6.2.3 wektory

$$\sigma^{m-k}(v), \sigma^{m-k+1}(v), \dots, \sigma^{m-1}(v)$$

są liniowo niezależne, zatem $a_0 = a_1 = \dots = a_{k-1} = 0$. Stąd

$$u = a_k \sigma^k(v) + \dots + a_{m-1} \sigma^{m-1}(v) = \sigma^k(u_0),$$

gdzie $u_0 = a_k v + a_{k+1} \sigma(v) + \dots + a_{m-1} \sigma^{m-k-1}(v) \in U_1$.

(b) jest równoważne z (a). $\square$

WNIOSEK 6.2.2. *Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech $U_1 = K[\sigma]v$ będzie maksymalną podprzestrzenią σ -cykliczną przestrzeni V . Niech σ_1 będzie zacieśnieniem endomorfizmu σ do podprzestrzeni U_1 . Wtedy*

$$\dim \ker \sigma_1 = 1 \quad \text{oraz} \quad \ker \sigma_1 = \text{lin} \left(\sigma^{m-1}(v) \right).$$

Dowód. Jest to przypadek $k = m - 1$ w lemacie 6.2.4. □

Udowodnimy teraz kluczowy fakt o endomorfizmach nilpotentnych pokazujący, że jeśli stopień nilpotentności m endomorfizmu σ spełnia $m < \dim V$, to endomorfizm σ jest redukowalny.

TWIERDZENIE 6.2.1. *Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech U_1 będzie maksymalną podprzestrzenią σ -cykliczną przestrzeni V . Wtedy istnieje podprzestrzeń W przestrzeni V taka, że*

- (a) $V = U_1 \oplus W$
- (b) $\sigma(W) \subseteq W$.

Dowód. Niech W będzie maksymalną podprzestrzenią przestrzeni V spełniającą warunki

$$U_1 \cap W = 0 \quad \text{i} \quad \sigma(W) \subseteq W. \quad (6.1)$$

Istnienie takiej podprzestrzeni W wynika stąd, że w ogóle istnieją podprzestrzenie W spełniające warunki (6.1) (na przykład, podprzestrzeń zerowa $W = 0$) i wobec tego spośród podprzestrzeni spełniających warunki (6.1) wystarczy wybrać podprzestrzeń o największym wymiarze. Udowodnimy, że $V = U_1 + W$.

Przypuśćmy, że tak nie jest. Istnieje zatem wektor $z \in V$ taki, że $z \notin U_1 + W$. Obieramy liczbę naturalną k taką, że

$$\sigma^i(z) \notin U_1 + W \quad \text{dla} \quad i < k \quad \text{oraz} \quad \sigma^k(z) \in U_1 + W. \quad (6.2)$$

Ponieważ $\sigma^m = 0_V$, więc $k \leq m$. Mamy więc przedstawienie

$$\sigma^k(z) = u + w, \quad u \in U_1, \quad w \in W.$$

Stąd $0 = \sigma^m(z) = \sigma^{m-k}(u) + \sigma^{m-k}(w)$. Wobec niezmienniczości podprzestrzeni U_1 i W względem σ mamy

$$\sigma^{m-k}(u) \in U_1, \quad \sigma^{m-k}(w) \in W,$$

a więc

$$\sigma^{m-k}(u) = -\sigma^{m-k}(w) \in U_1 \cap W = 0.$$

Na podstawie lematu 6.2.4 istnieje $u_0 \in U_1$ taki, że

$$u = \sigma^k(u_0).$$

Stąd $\sigma^k(z) = u + w = \sigma^k(u_0) + w$ i dla $z_1 := z - u_0$ mamy $z_1 \notin U_1 + W$ oraz

$$\sigma^k(z_1) = \sigma^k(z) - \sigma^k(u_0) = w \in W.$$

Stąd także

$$\sigma^j(z_1) \in W \quad \text{dla} \quad j \geq k, \quad (6.3)$$

gdyż W jest niezmiennicza względem σ . Rozważmy teraz

$$\sigma^i(z_1) = \sigma^i(z) - \sigma^i(u_0)$$

dla $i < k$. Tutaj $\sigma^i(z) \notin U_1 + W$ na podstawie wyboru liczby k , oraz $\sigma^i(u_0) \in U_1$, bo $u_0 \in U_1$ i $\sigma(U_1) \subseteq U_1$. A więc nie tylko $z_1 \notin U_1 + W$ ale także

$$\sigma^i(z_1) \notin U_1 + W \quad \text{dla} \quad i < k. \quad (6.4)$$

Element z_1 ma zatem te same własności co element z występujący w (6.2) z tym, że mamy dokładniejszą informację co do obrazu elementu z_1 poprzez σ^k , mianowicie $\sigma^k(z_1) \in W$, podczas gdy o $\sigma^k(z)$ wiemy tylko, że należy do większej podprzestrzeni $U_1 + W$.

Definiujemy teraz podprzestrzeń W_1 przestrzeni V jako

$$W_1 := W + \text{lin}(z_1, \sigma(z_1), \dots, \sigma^{k-1}(z_1)).$$

Wobec (6.4) mamy oczywiście $\sigma^i(z_1) \notin W$ dla $i < k$. A zatem

$$W_1 \supset W, \quad W_1 \neq W.$$

Ponadto, ponieważ $\sigma^k(z_1) \in W$, więc W_1 jest niezmiennicza względem σ (W jest także niezmiennicza względem σ). Wobec wyboru W jako maksymalnej podprzestrzeni spełniającej warunki (6.1) wynika stąd, że

$$U_1 \cap W_1 \neq 0.$$

Niech więc

$$0 \neq w_1 := w_0 + a_0 z_1 + a_1 \sigma(z_1) + \dots + a_{k-1} \sigma^{k-1}(z_1) \in U_1 \cap W_1,$$

gdzie $w_0 \in W$ oraz $a_i \in K$. Gdyby wszystkie $a_i = 0$, to $0 \neq w_1 = w_0 \in U_1 \cap W = 0$, sprzeczność. A więc nie wszystkie a_i są równe zero. Niech s będzie najmniejszym wskaźnikiem takim, że $a_s \neq 0$. Wtedy

$$w_1 = w_0 + a_s \sigma^s(z_1) + \dots + a_{k-1} \sigma^{k-1}(z_1) \in U_1 \cap W_1$$

oraz wobec niezmienniczości U_1 i W_1 względem σ także

$$\sigma^{k-1-s}(w_1) = \sigma^{k-1-s}(w_0) + a_s \sigma^{k-1}(z_1) + a_{s+1} \sigma^k(z_1) + \dots + \sigma^{2k-2-s}(z_1) \in U_1 \cap W_1.$$

Tymczasem

$$\sigma^{k-1-s}(w_0) \in W$$

na podstawie niezmienniczości W względem σ , oraz

$$a_s \sigma^{k-1}(z_1) \notin U_1 + W$$

na podstawie (6.4). Natomiast na podstawie (6.3) wszystkie pozostałe składniki należą do W . A więc

$$\sigma^{k-1-s}(w_1) \notin U_1 + W$$

oraz z drugiej strony, $\sigma^{k-1-s}(w_1) \in U_1$ wobec $w_1 \in U_1$ i niezmienniczości U_1 względem σ . Otrzymaliśmy więc sprzeczność.

Przypuszczenie iż $V \neq U_1 + W$ prowadzi zatem do sprzeczności.

Mamy więc $V = U_1 + W$ i wobec (6.1) otrzymujemy $V = U_1 \oplus W$. □

6.3 Postać kanoniczna

TWIERDZENIE 6.3.1. *Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V . Wtedy istnieją podprzestrzenie σ -cykliczne $U_1, \dots, U_r$ przestrzeni V takie, że*

$$V = U_1 \oplus \dots \oplus U_r.$$

Jeśli $m_i = \dim U_i$ oraz $m_1 \geq \dots \geq m_r$, to

- (a) *Endomorfizm σ_i indukowany na podprzestrzeni U_i przez endomorfizm σ jest endomorfizmem nilpotentnym stopnia m_i .*
- (b) *$m_1 = m$.*
- (c) *$r = \dim \ker \sigma$.*

Dowód. Przede wszystkim zauważmy, że twierdzenie jest prawdziwe gdy $m = 1$, to znaczy gdy σ jest endomorfizmem zerowym. Rzeczywiście, jeśli $\{v_1, \dots, v_n\}$ jest jakąkolwiek bazą przestrzeni V , to w charakterze U_i można wziąć $\text{lin}(v_i)$. Punkty (a) i (b) są oczywiste, natomiast (c) wynika stąd, że $V = \ker \sigma$, zatem $r = n = \dim \ker \sigma$. Stąd też wynika, że twierdzenie jest prawdziwe gdy $\dim V = 1$, gdyż wtedy jedynym endomorfizmem nilpotentnym jest endomorfizm zerowy.

Dla dowodu twierdzenia zastosujemy indukcję względem wymiaru przestrzeni V . Niech więc $\dim V > 1$ i niech $m > 1$. Jak wiemy z lematu 6.2.2, jeśli σ jest endomorfizmem nilpotentnym stopnia m przestrzeni V , to istnieje podprzestrzeń σ -cykliczna U_1 przestrzeni V o wymiarze m . Rzeczywiście, jako U_1 można wziąć $\text{lin}(v, \sigma(v), \dots, \sigma^{m-1}(v))$ gdzie $v \in V$ jest dowolnym wektorem spełniającym $\sigma^{m-1}(v) \neq 0$.

Na podstawie twierdzenia 6.2.1 istnieje podprzestrzeń σ -niezmiennicza W przestrzeni V taka, że $V = U_1 \oplus W$. Ponieważ $m > 1$, więc $\dim W < \dim V$. Rozpatrzmy zacieśnienie $\sigma|_W$ endomorfizmu σ do podprzestrzeni W . Wtedy $(\sigma|_W)^m = 0_W$. Niech m_2 będzie stopniem nilpotentności endomorfizmu $\sigma|_W$ przestrzeni W . Wtedy $m_1 := m \geq m_2$. Na podstawie założenia indukcyjnego istnieją podprzestrzenie $\sigma|_W$ -cykliczne $U_2, \dots, U_r$ przestrzeni W takie, że

$$W = U_2 \oplus \dots \oplus U_r$$

oraz $m_2 = \dim U_2 \geq \dots \geq \dim U_r$. Ponadto, endomorfizm σ_i indukowany na podprzestrzeni U_i przez endomorfizm $\sigma|_W$ jest endomorfizmem nilpotentnym stopnia $m_i = \dim U_i$ oraz $r - 1 = \dim \ker \sigma|_W$. Wobec tego $V = U_1 \oplus W = U_1 \oplus \dots \oplus U_r$ oraz dla każdego $i \leq r$ endomorfizm σ_i indukowany na podprzestrzeni U_i przez endomorfizm σ jest endomorfizmem nilpotentnym stopnia $m_i = \dim U_i$, przy czym $m_1 = m$. Pozostaje udowodnić (c). Zauważmy, że wobec rozkładu $V = U_1 \oplus W$ mamy także

$$\ker \sigma = \ker \sigma_1 \oplus \ker \sigma|_W.$$

Wobec tego $\dim \ker \sigma = \dim \ker \sigma_1 + \dim \ker \sigma|_W = 1 + (r - 1) = r$, gdyż na podstawie wniosku 6.2.2 mamy $\dim \ker \sigma_1 = 1$. $\square$

Zanotujmy jeszcze wersję macierzową twierdzenia 6.3.1.

TWIERDZENIE 6.3.2. *Jeśli $\dim V = n$ oraz σ jest endomorfizmem nilpotentnym przestrzeni V stopnia m , to istnieje baza $\mathcal{B}$ przestrzeni V taka, że*

$$\mathbf{m}(\sigma, \mathcal{B}) = \begin{bmatrix} J_{m_1} & 0 & \dots & 0 \\ 0 & J_{m_2} & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & J_{m_r} \end{bmatrix},$$

gdzie $J_{m_1}, \dots, J_{m_r}$ są osobliwymi klatkami Jordana stopni $m_1, \dots, m_r$ oraz

$$m = m_1 \geq m_2 \geq \dots \geq m_r, \quad m_1 + m_2 + \dots + m_r = n, \quad r = \dim \ker \sigma.$$

Dowód. Wystarczy zastosować lemat 6.2.3 do każdej z podprzestrzeni cyklicznych U_i w twierdzeniu 6.3.1. $\square$

Macierz $m(\sigma, \mathcal{B})$ z twierdzenia 6.3.2 nazywamy *postacią kanoniczną Jordana macierzy endomorfizmu nilpotentnego σ* .

Dla zwięzłego formułowania twierdzeń o postaciach kanonicznych macierzy endomorfizmów wprowadzimy jeszcze jedną definicję.

DEFINICJA 6.3.1. Niech $M_1, M_2, \dots, M_k$ będą macierzami dowolnych stopni o elementach z ciała K . Macierz klatkową

$$M = \begin{bmatrix} M_1 & 0 & \dots & 0 & 0 \\ 0 & M_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & M_k \end{bmatrix}$$

nazywamy *sumą prostą* macierzy $M_1, M_2, \dots, M_k$ i oznaczamy

$$M := M_1 \oplus M_2 \oplus \dots \oplus M_k.$$

A więc twierdzenie 6.3.2 mówi, że dla endomorfizmu nilpotentnego σ stopnia m istnieje baza przestrzeni V , w której macierz endomorfizmu σ jest sumą prostą osobliwych klatek Jordana stopni $\leq m$.

Przejdziemy teraz do zagadnienia jednoznaczności postaci kanonicznej Jordana macierzy endomorfizmu nilpotentnego. Rozpocniemy od lematu wyznaczającego wymiar obrazu dowolnej podprzestrzeni σ -cyklicznej poprzez potęgę endomorfizmu nilpotentnego σ .

LEMAT 6.3.1. Niech σ będzie endomorfizmem nilpotentnym stopnia m przestrzeni V i niech U będzie ℓ -wymiarową podprzestrzenią σ -cykliczną przestrzeni V . Wtedy dla każdej liczby naturalnej $k \leq \ell$ zachodzi $\dim \sigma^k(U) = \ell - k$.

Dowód. Jeśli $\{v, \sigma(v), \dots, \sigma^{\ell-1}(v)\}$ jest bazą podprzestrzeni σ -cyklicznej U , to zbiór

$$\{\sigma^k(v), \sigma^{k+1}(v), \dots, \sigma^{\ell-1}(v)\}$$

rozpina podprzestrzeń $\sigma^k(U)$. Ponieważ zbiór ten jest liniowo niezależny (jako podzbiór bazy przestrzeni U), jest on bazą podprzestrzeni $\sigma^k(U)$.

Zatem $\dim \sigma^k(U) = \ell - k$. $\square$

TWIERDZENIE 6.3.3. Niech σ będzie endomorfizmem nilpotentnym przestrzeni V i niech

$$V = U_1 \oplus \dots \oplus U_r = U'_1 \oplus \dots \oplus U'_s$$

będą dwoma rozkładami przestrzeni V na sumę prostą podprzestrzeni σ -cyklicznych. Niech $\dim U_i = m_i$, $\dim U'_j = n_j$, dla $i = 1, \dots, r$, $j = 1, \dots, s$ i załóżmy, że

$$m_1 \geq \dots \geq m_r, \quad n_1 \geq \dots \geq n_s.$$

Wtedy

$$r = s \quad \text{oraz} \quad m_i = n_i, \quad i = 1, \dots, r.$$

Dowód. Udowodnimy najpierw, że $m_1 = n_1$. Przypuśćmy, że tak nie jest. Ze względu na symetrię założeń możemy przyjąć, że $m_1 > n_1$. Rozważmy podprzestrzeń $\sigma^{n_1}(V)$. Ponieważ endomorfizm indukowany przez σ na podprzestrzeni U_i jest nilpotentny i ma stopień $n_i \leq n_1$, więc $\sigma^{n_1}(U_i) = 0$ i wobec tego także $\sigma^{n_1}(V) = 0$. Z drugiej strony, na podstawie lematu 6.3.1, mamy $\dim \sigma^{n_1}(U_1) = m_1 - n_1 > 0$. Zatem $\sigma^{n_1}(U_1)$ jest niezerową podprzestrzenią przestrzeni V wbrew temu, że $\sigma^{n_1}(V) = 0$. Wobec tego musimy mieć $m_1 = n_1$.

Założmy teraz, że $i > 1$ jest najmniejszym wskaźnikiem takim, że $m_i \neq n_i$. Możemy oczywiście założyć, że $m_i > n_i$. Zatem

$$m_1 = n_1, \quad \dots, \quad m_{i-1} = n_{i-1}, \quad m_i > n_i. \quad (6.5)$$

Rozpatrzmy teraz podprzestrzeń $\sigma^{n_i}(V)$ przestrzeni V . Z jednej strony mamy

$$\sigma^{n_i}(V) = \sigma^{n_i}(U_1) \oplus \dots \oplus \sigma^{n_i}(U_i) \oplus \dots \oplus \sigma^{n_i}(U_r).$$

Na podstawie lematu 6.3.1 mamy

$$\dim \sigma^{n_i}(U_j) = m_j - n_i, \quad j = 1, \dots, i.$$

Stąd otrzymujemy

$$\dim \sigma^{n_i}(V) \geq (m_1 - n_i) + (m_2 - n_i) + \dots + (m_i - n_i). \quad (6.6)$$

Z drugiej zaś strony mamy

$$\begin{aligned} \sigma^{n_i}(V) &= \sigma^{n_i}(U'_1) \oplus \dots \oplus \sigma^{n_i}(U'_i) \oplus \dots \oplus \sigma^{n_i}(U'_s) \\ &= \sigma^{n_i}(U'_1) \oplus \dots \oplus \sigma^{n_i}(U'_{i-1}) \end{aligned}$$

gdyż $\sigma^{n_i}(U'_i) = \dots = \sigma^{n_i}(U'_s) = 0$. Na podstawie lematu 6.3.1 mamy

$$\dim \sigma^{n_i}(U'_j) = n_j - n_i, \quad j = 1, \dots, i.$$

Stąd otrzymujemy

$$\dim \sigma^{n_i}(V) = (n_1 - n_i) + (n_2 - n_i) + \dots + (n_{i-1} - n_i)$$

i wobec (6.5) mamy także

$$\dim \sigma^{n_i}(V) = (m_1 - n_i) + (m_2 - n_i) + \dots + (m_{i-1} - n_i). \quad (6.7)$$

Porównując teraz (6.6) i (6.7) otrzymujemy $m_i - n_i \leq 0$, czyli $m_i \leq n_i$, wbrew założeniu, że $m_i > n_i$. A więc mamy równości $m_i = n_i$ dla każdego $i \leq \min(r, s)$. Ponieważ jednak $\sum_{i=1}^r m_i = n = \sum_{j=1}^s n_j$, wynika stąd, że także $r = s$. $\square$

WNIOSEK 6.3.1. Niech σ będzie endomorfizmem nilpotentnym przestrzeni V i niech

$$J_{m_1} \oplus \dots \oplus J_{m_r} \quad \text{oraz} \quad J_{n_1} \oplus \dots \oplus J_{n_s}$$

będą dwiema postaciami kanonicznymi Jordana macierzy endomorfizmu σ . Jeśli

$$m_1 \geq \dots \geq m_r, \quad n_1 \geq \dots \geq n_s,$$

to

$$r = s \quad \text{oraz} \quad m_i = n_i, \quad i = 1, \dots, r.$$

Dowód. Postaci kanoniczne Jordana macierzy endomorfizmu σ odpowiadają rozkładowi przestrzeni V na sumy proste podprzestrzeni σ –cyklicznych, opisanym w twierdzeniu 6.3.3. $\square$

DEFINICJA 6.3.2. Niech σ będzie endomorfizmem nilpotentnym przestrzeni V i niech

$$V = U_1 \oplus \cdots \oplus U_r,$$

będzie rozkładem przestrzeni V na sumę prostą podprzestrzeni σ –cyklicznych. Niech $\dim U_i =: m_i$, $i = 1, \dots, r$ oraz $m_1 \geq \cdots \geq m_r$.

Ciąg liczb $(m_1, \dots, m_r)$ nazywa się *ciągami niezmienników* endomorfizmu σ .

Endomorfizm nilpotentny σ przestrzeni V ma ciąg niezmienników $(m_1, \dots, m_r)$ wtedy i tylko wtedy, gdy istnieje baza uporządkowana $\mathcal{B}$ przestrzeni V taka, że

$$\mathbf{m}(\sigma, \mathcal{B}) = J_{m_1} \oplus \cdots \oplus J_{m_r}.$$

WNIOSEK 6.3.2. *Ciąg niezmienników endomorfizmu nilpotentnego σ jest wyznaczony jednoznacznie przez endomorfizm σ .*

TWIERDZENIE 6.3.4. *Endomorfizmy nilpotentne ρ i τ przestrzeni V są podobne wtedy i tylko wtedy, gdy mają równe ciągi niezmienników.*

Dowód. Endomorfizmy ρ i τ mają równe ciągi niezmienników $(m_1, \dots, m_r)$ wtedy i tylko wtedy gdy w odpowiednich bazach przestrzeni V mają tę samą macierz klatkową $J_{m_1} \oplus \cdots \oplus J_{m_r}$. Na podstawie twierdzenia 2.2.6 ma to miejsce wtedy i tylko wtedy gdy endomorfizmy ρ i τ są podobne. $\square$

Przypomnijmy, że relacja podobieństwa endomorfizmów przestrzeni wektorowej V jest relacją równoważnościową. Pokażemy, że liczba klas abstrakcji tej relacji jest skończona i wyznaczymy liczbę tych klas w zależności od wymiaru przestrzeni.

WNIOSEK 6.3.3. *Liczba klas podobieństwa endomorfizmów nilpotentnych n –wymiarowej przestrzeni wektorowej jest równa liczbie partycji liczby naturalnej n .*

Dowód. Tutaj przez *partycję* liczby naturalnej n rozumiemy każde przedstawienie liczby n w postaci

$$n = m_1 + \cdots + m_r, \quad m_1 \geq \cdots \geq m_r, \quad r \geq 1,$$

gdzie m_i są liczbami naturalnymi. Liczbę wszystkich partycji liczby n oznacza się $p(n)$. A więc $p(1) = 1$, $p(2) = 2$, $p(5) = 7$. Słynny rezultat Hardy’ego i Ramanujana z 1917 roku pozwala obliczyć $p(200) = 3972999029388$.¹

Co do dowodu wniosku zauważmy, że na podstawie twierdzenia 6.3.4 każda klasa podobnych endomorfizmów nilpotentnych wyznacza dokładnie jeden ciąg niezmienników $(m_1, \dots, m_r)$ i każdy taki ciąg wyznacza partycję liczby $n = \dim V$. W ten sposób określamy odwzorowanie zbioru klas podobieństwa endomorfizmów nilpotentnych przestrzeni n –wymiarowej w zbiór wszystkich partycji liczby n . Na podstawie twierdzenia 6.3.4 odwzorowanie to jest różnowartościowe.

Z drugiej strony, każda partycja $n = m_1 + \cdots + m_r$ pozwala skonstruować macierz klatkową $J_{m_1} \oplus \cdots \oplus J_{m_r}$, która w sposób naturalny jest macierzą pewnego endomorfizmu nilpotentnego σ na przestrzeni V (i jest także macierzą każdego endomorfizmu podobnego do σ). Klasa podobieństwa endomorfizmu σ przechodzi więc poprzez nasze odwzorowanie na zadaną partycję $n = m_1 + \cdots + m_r$. A więc nasze odwzorowanie jest surjektywne i wobec tego jest bijekcją. Stąd rozpatrywane dwa zbiory mają tę samą liczbę elementów. $\square$

¹Zob. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford 1960, str. 286.

Przykład 6.3.1. Ponieważ $p(5) = 7$, możemy stwierdzić, że dla 5-wymiarowej przestrzeni wektorowej nad dowolnym ciałem istnieje dokładnie 7 klas podobieństwa endomorfizmów nilpotentnych. Klasy te mają następujące macierze w postaci kanonicznej Jordana:

$$J_5, \quad J_4 \oplus J_1, \quad J_3 \oplus J_2, \quad J_3 \oplus J_1 \oplus J_1, \quad J_2 \oplus J_2 \oplus J_1, \quad J_2 \oplus J_1 \oplus J_1 \oplus J_1, \quad J_1 \oplus J_1 \oplus J_1 \oplus J_1 \oplus J_1.$$

Rozdział 7

Postać kanoniczna Jordana

Ostatnie zmiany 15.01.2009 r.

W tym rozdziale wskażemy postać kanoniczną endomorfizmu $\tau \in \text{End}_K V$ w przypadku, gdy wielomian minimalny endomorfizmu τ rozkłada się na czynniki liniowe nad ciałem K . Zauważmy, że na podstawie twierdzenia 4.2.2 przestrzeń V ma bazę, w której macierz endomorfizmu τ jest trójkątna. Zastosowanie twierdzenia 5.2.2 o rozkładzie i rezultatów o postaci kanonicznej endomorfizmów nilpotentnych prowadzi jednak do znacznie dokładniejszego twierdzenia o postaci kanonicznej Jordana macierzy takiego endomorfizmu.

7.1 Postać kanoniczna

Rozpocniemy od nieznacznego uogólnienia twierdzenia 6.3.2 i wniosku 6.3.1 o istnieniu i jednoznaczności postaci kanonicznej Jordana macierzy endomorfizmu nilpotentnego. Endomorfizm nilpotentny σ stopnia m ma wielomian minimalny $p_\sigma = X^m$, natomiast nasze uogólnienie dotyczy endomorfizmów τ , które mają wielomian minimalny $p_\tau = (X - a)^m$.

Dla dowolnej liczby naturalnej m i dowolnego $a \in K$ wprowadzamy oznaczenie $J_m(a)$ dla macierzy stopnia m postaci

$$J_m(a) := aI_m + J_m = \begin{bmatrix} a & 0 & 0 & \dots & 0 & 0 \\ 1 & a & 0 & \dots & 0 & 0 \\ 0 & 1 & a & \dots & 0 & 0 \\ & & & \dots & & \\ 0 & 0 & 0 & \dots & 1 & a \end{bmatrix}.$$

Macierz $J_m(a)$ nazywa się *klatką Jordana* stopnia m wyznaczoną przez element $a \in K$ albo należącą do $a \in K$.

TWIERDZENIE 7.1.1. *Niech $\tau \in \text{End}_K V$ i niech wielomian minimalny p_τ będzie potęgą wielomianu liniowego nad ciałem K :*

$$p_\tau = (X - a)^m, \quad a \in K, \quad m \geq 1.$$

Wtedy istnieje baza $\mathcal{B}$ przestrzeni V oraz liczby naturalne $m_1, \dots, m_r$ takie, że macierz $\mathbf{m}(\tau, \mathcal{B})$ jest sumą prostą klatek Jordana należących do wartości własnej a endomorfizmu τ :

$$\mathbf{m}(\tau, \mathcal{B}) = J_{m_1}(a) \oplus \dots \oplus J_{m_r}(a).$$

Ponadto, jeśli $m_1 \geq m_2 \geq \dots \geq m_r$, to $m_1 = m$, $r = \dim \ker(\tau - a\mathbf{1}_V)$ oraz ciąg liczb $m_1, \dots, m_r$ jest wyznaczony jednoznacznie przez endomorfizm τ (to znaczy, nie zależy od wyboru bazy $\mathcal{B}$).

Dowód. Ponieważ $p_\tau = (X - a)^m$, więc $(\tau - a\mathbf{1}_V)^m = 0_V$. Wynika stąd, że endomorfizm $\sigma := \tau - a\mathbf{1}_V$ przestrzeni V jest nilpotentny,

$$\sigma^m = 0_V,$$

oraz m jest stopniem nilpotentności endomorfizmu σ . Zauważmy, że jeśli $\mathcal{B}$ jest bazą przestrzeni V oraz $\dim V = n$, to

$$\mathbf{m}(\tau, \mathcal{B}) = \mathbf{m}(a\mathbf{1}_V + \sigma, \mathcal{B}) = aI_n + \mathbf{m}(\sigma, \mathcal{B}),$$

gdzie I_n jest macierzą jednostkową stopnia n .

Wobec tego, na podstawie twierdzenia 6.3.2 o postaci kanonicznej macierzy endomorfizmu nilpotentnego, przestrzeń V ma bazę $\mathcal{B}$ taką, że

$$\mathbf{m}(\tau, \mathcal{B}) = aI_n + \begin{bmatrix} J_{m_1} & 0 & \cdots & 0 \\ 0 & J_{m_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J_{m_r} \end{bmatrix}$$

gdzie $m =: m_1 \geq m_2 \geq \cdots \geq m_r$ oraz $r = \dim \ker \sigma = \dim \ker(\tau - a\mathbf{1}_V)$. Zatem macierz

$$\mathbf{m}(\tau, \mathcal{B}) = \begin{bmatrix} J_{m_1}(a) & 0 & \cdots & 0 \\ 0 & J_{m_2}(a) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J_{m_r}(a) \end{bmatrix}.$$

jest sumą prostą klatek Jordana należących do wartości własnej a endomorfizmu τ .

Przypuśćmy teraz, że w pewnej bazie $\mathcal{B}_1$ przestrzeni V endomorfizm τ ma macierz będącą sumą prostą jakichkolwiek klatek Jordana:

$$\mathbf{m}(\tau, \mathcal{B}_1) = J_{n_1}(b_1) \oplus \cdots \oplus J_{n_s}(b_s),$$

gdzie $b_1, \dots, b_s \in K$ oraz $n_1 \geq n_2 \geq \cdots \geq n_s$. Wtedy $\mathbf{m}(\tau, \mathcal{B}_1)$ jest macierzą trójkątną i na podstawie twierdzenia 4.2.3 każdy element diagonalny tej macierzy jest wartością własną endomorfizmu τ . Endomorfizm τ ma jednak tylko jedną wartość własną a (na podstawie twierdzenia 3.1.4) i wobec tego $b_1 = b_2 = \cdots = b_s = a$. A więc

$$\mathbf{m}(\tau, \mathcal{B}_1) = J_{n_1}(a) \oplus \cdots \oplus J_{n_s}(a),$$

i wobec tego dla endomorfizmu nilpotentnego $\sigma = \tau - a\mathbf{1}_V$ otrzymujemy

$$\mathbf{m}(\sigma, \mathcal{B}_1) = \mathbf{m}(\tau, \mathcal{B}_1) - aI_n = J_{n_1}(0) \oplus \cdots \oplus J_{n_s}(0) = J_{n_1} \oplus \cdots \oplus J_{n_s}.$$

Na podstawie wniosku 6.3.1 otrzymujemy stąd $r = s$ oraz $m_i = n_i$ dla $i = 1, \dots, r$. $\square$

Dla endomorfizmu τ istnieje więc dokładnie jedna suma prosta klatek Jordana będąca macierzą endomorfizmu τ w pewnej bazie przestrzeni V . Tę macierz nazywamy *postacią kanoniczną Jordana macierzy endomorfizmu τ* .

Przykład 7.1.1. Niech $\dim_K V = 5$ i niech endomorfizm τ przestrzeni V ma wielomian minimalny $p_\tau = (X - a)^m$. Wtedy m jest stopniem nilpotentności endomorfizmu $\tau - a\mathbf{1}_V$ i wobec tego $m \leq 5$ (na podstawie wniosku 6.2.1). Endomorfizm τ ma zatem jako macierz w postaci kanonicznej Jordana jedną z następujących siedmiu macierzy:

$$\begin{aligned} J_5(a), \quad J_4(a) \oplus J_1(a), \quad J_3(a) \oplus J_2(a), \quad J_3(a) \oplus J_1(a) \oplus J_1(a), \quad J_2(a) \oplus J_2(a) \oplus J_1(a), \\ J_2(a) \oplus J_1(a) \oplus J_1(a) \oplus J_1(a), \quad J_1(a) \oplus J_1(a) \oplus J_1(a) \oplus J_1(a) \oplus J_1(a) \end{aligned}$$

(zobacz przykład 6.3.1).

Rozpatrzmy teraz przypadek, gdy wielomian minimalny endomorfizmu τ jest iloczynem potęg wielomianów liniowych. Niech więc endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny

$$p_\tau = (X - a_1)^{m_1} \cdots (X - a_k)^{m_k}, \quad (7.1)$$

gdzie $a_1, \dots, a_k \in K$, przy czym $a_i \neq a_j$ dla $i \neq j$ oraz $m_i \geq 1$ dla $i = 1, \dots, k$. Jest to więc rozkład kanoniczny wielomianu p_τ nad ciałem K ,

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k},$$

w którym $q_i = X - a_i$. Na podstawie twierdzenia 5.2.2 o rozkładzie prymarnym, przestrzeń V ma rozkład

$$V = V_1 \oplus \cdots \oplus V_k, \quad (7.2)$$

gdzie V_i jest podprzestrzenią niezmienniczą endomorfizmu τ oraz endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny $q_i^{m_i} = (X - a_i)^{m_i}$. Wobec tego na podstawie twierdzenia 7.1.1 podprzestrzeń V_i ma bazę $\mathcal{B}_i$ taką, że macierz $\mathbf{m}(\tau_i, \mathcal{B}_i)$ jest sumą prostą klatek Jordana należących do wartości własnej a_i endomorfizmu τ_i :

$$\mathbf{m}(\tau_i, \mathcal{B}_i) = J_{m_{i1}}(a_i) \oplus \cdots \oplus J_{m_{ir_i}}(a_i) \quad (7.3)$$

dla pewnych liczby naturalnych $m_{i1}, \dots, m_{ir_i}$, przy czym jeśli założymy, że

$$m_{i1} \geq \cdots \geq m_{ir_i},$$

to $m_{i1} = m_i$ jest krotnością pierwiastka a_i wielomianu minimalnego p_τ endomorfizmu τ . Ponadto, liczba r_i klatek Jordana należących do wartości własnej a_i jest równa $r_i = \dim \ker(\tau_i - a_i \mathbf{1}_{V_i})$. Zauważmy, że $\ker(\tau_i - a_i \mathbf{1}_{V_i}) = \ker(\tau - a_i \mathbf{1}_V)$. Rzeczywiście, inkluzja $\ker(\tau_i - a_i \mathbf{1}_{V_i}) \subseteq \ker(\tau - a_i \mathbf{1}_V)$ jest oczywista, natomiast jeśli $v \in V$ oraz $v \in \ker(\tau - a_i \mathbf{1}_V)$, to także $v \in \ker(\tau - a_i \mathbf{1}_V)^{m_i} = V_i$ i wobec tego $v \in \ker(\tau_i - a_i \mathbf{1}_{V_i})$.

Przechodząc teraz do przestrzeni V rozpatrujemy rozkład (7.2) i związaną z nim bazę uporządkowaną

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$$

przestrzeni V . Dla macierzy endomorfizmu τ w tej bazie mamy

$$\mathbf{m}(\tau, \mathcal{B}) = \mathbf{m}(\tau_1, \mathcal{B}_1) \oplus \cdots \oplus \mathbf{m}(\tau_k, \mathcal{B}_k),$$

czyli macierz o bardzo specjalnej budowie. Jest to właśnie *postać kanoniczna Jordana* macierzy endomorfizmu τ . Udowodniliśmy więc następujące twierdzenie.

TWIERDZENIE 7.1.2. (Postać kanoniczna Jordana macierzy endomorfizmu.)

Niech $\tau \in \text{End}_K V$ i niech wielomian minimalny p_τ endomorfizmu τ rozkłada się na czynniki liniowe nad ciałem K :

$$p_\tau = (X - a_1)^{m_1} \cdots (X - a_k)^{m_k},$$

gdzie $a_1, \dots, a_k \in K$, $a_i \neq a_j$ dla $i \neq j$ oraz $m_i \geq 1$ dla $i = 1, \dots, k$.

Wtedy istnieje baza $\mathcal{B}$ przestrzeni V taka, że

$$\mathbf{m}(\tau, \mathcal{B}) = M_1 \oplus \cdots \oplus M_k,$$

gdzie dla $i = 1, \dots, k$, macierz M_i jest sumą prostą klatek Jordana należących do wartości własnej a_i endomorfizmu τ :

$$M_i = J_{m_{i1}}(a_i) \oplus \cdots \oplus J_{m_{ir_i}}(a_i).$$

Ponadto:

- (a) Jeśli $m_{i1} \geq \dots \geq m_{ir_i}$, to $m_{i1} = m_i$ jest krotnością pierwiastka a_i wielomianu minimalnego p_τ endomorfizmu τ .
 (b) Liczba r_i klatek Jordana należących do wartości własnej a_i jest równa wymiarowi podprzestrzeni wektorów własnych endomorfizmu τ należących do wartości własnej a_i :

$$r_i = \dim \ker(\tau - a_i \mathbf{1}_V), \quad i = 1, \dots, k.$$

Postać kanoniczną Jordana macierzy endomorfizmu τ nazywamy macierz

$$J = \bigoplus_{i=1}^k \bigoplus_{j=1}^{r_i} J_{m_{ij}}(a_i).$$

Natomiast bazę $\mathcal{B}$ przestrzeni V , w której endomorfizm τ ma macierz J nazywamy *bazą kanoniczną Jordana* przestrzeni V . Podkreślmy, że postać kanoniczna Jordana macierzy endomorfizmu $\tau \in \text{End}_K V$ istnieje tylko wtedy, gdy wielomian minimalny endomorfizmu τ rozkłada się na czynniki liniowe nad ciałem K . Istnieją więc ciała K , dla których twierdzenie o istnieniu postaci kanonicznej Jordana jest prawdziwe dla *wszystkich* endomorfizmów *dowolnej* (skończenie wymiarowej) przestrzeni wektorowej nad ciałem K . Są to oczywiście ciała *algebraicznie domknięte*, nad którymi *każdy* wielomian rozkłada się na czynniki liniowe. W szczególności otrzymujemy następujący wniosek.

WNIOSEK 7.1.1. *Każdy endomorfizm dowolnej przestrzeni wektorowej nad ciałem liczb zespolonych ma macierz w postaci kanonicznej Jordana (w odpowiedniej bazie przestrzeni).*

7.2 Jednoznaczność postaci kanonicznej

Przechodzimy teraz do zagadnienia jednoznaczności postaci kanonicznej Jordana macierzy endomorfizmu. Zamierzamy udowodnić, że dwie postacie kanoniczne macierzy endomorfizmu τ mogą się różnić tylko porządkiem klatek. Natomiast liczba klatek należących do danej wartości własnej oraz rozmiary tych klatek są wyznaczone jednoznacznie przez endomorfizm τ i nie zależą od sposobu rozkładu przestrzeni V na sumę prostą podprzestrzeni cyklicznych.

TWIERDZENIE 7.2.1. *Niech $\tau \in \text{End}_K V$ i niech wielomian minimalny p_τ endomorfizmu τ ma rozkład (7.1) na czynniki liniowe nad ciałem K . Niech $\mathcal{A}$ będzie jakkolwiek bazą przestrzeni V , w której endomorfizm τ ma macierz będącą sumą prostą klatek Jordana:*

$$\mathbf{m}(\tau, \mathcal{A}) = \bigoplus_{i=1}^{k'} \bigoplus_{j=1}^{s_i} J_{d_{ij}}(b_i), \quad (7.4)$$

gdzie $b_i \in K$. Niech $b_i \neq b_\ell$ dla $i \neq \ell$ oraz $d_{i1} \geq \dots \geq d_{is_i}$. Wtedy przy oznaczeniach twierdzenia 7.1.2 mamy

$$k' = k, \quad b_i = a_i, \quad s_i = r_i, \quad d_{ij} = m_{ij} \quad (7.5)$$

dla $i = 1, 2, \dots, k$, $j = 1, 2, \dots, r_i$, po ewentualnej zmianie porządku i numeracji elementów b_i .

Dowód. Przedstawienie (7.4) macierzy $\mathbf{m}(\tau, \mathcal{A})$ w postaci sumy prostej klatek Jordana zapiszemy zwięźlej jako

$$\mathbf{m}(\tau, \mathcal{A}) = M'_1 \oplus \dots \oplus M'_{k'}, \quad (7.6)$$

gdzie dla $i = 1, \dots, k'$, macierz M'_i jest sumą prostą klatek Jordana należących do elementu $b_i \in K$:

$$M'_i = J_{d_{i1}}(b_i) \oplus \dots \oplus J_{d_{is_i}}(b_i).$$

Przedstawieniu macierzy endomorfizmu τ w postaci (7.6) odpowiada rozkład przestrzeni V na sumę prostą podprzestrzeni niezmienniczych:

$$V = U_1 \oplus \dots \oplus U_{k'}.$$

Tutaj stopień macierzy M'_i jest równy wymiarowi podprzestrzeni U_i . Ponadto, w każdej podprzestrzeni U_i istnieje baza $\mathcal{A}_i$ taka, że

$$\mathbf{m}(\tau_i, \mathcal{A}_i) = M'_i,$$

gdzie $\tau_i = \tau|_{U_i}$ jest endomorfizmem indukowanym przez τ na podprzestrzeni niezmienniczej U_i . Macierz M'_i jest trójkątna i wobec tego na podstawie twierdzenia 4.2.3(a) wielomian minimalny p_{τ_i} endomorfizmu τ_i jest dzielnikiem wielomianu $(X - b_i)^{d_i}$, gdzie $d_i = d_{i1} + \dots + d_{is_i}$ jest stopniem macierzy M'_i . Zatem

$$p_{\tau_i} = (X - b_i)^{\ell_i}$$

dla pewnej liczby naturalnej $\ell_i \geq 1$. Ponieważ elementy b_i są parami różne, wielomiany p_{τ_i} są parami wzajemnie pierwsze i na podstawie twierdzenia 5.1.1 mamy

$$p_{\tau} = p_{\tau_1} \cdots p_{\tau_{k'}} = (X - b_1)^{\ell_1} \cdots (X - b_{k'})^{\ell_{k'}}.$$

Wobec (7.1) oraz jednoznaczności rozkładu w pierścieniu $K[X]$ wynika stąd, że $k = k'$ oraz układ $b_1, \dots, b_k$ tylko porządkiem może różnić się od układu $a_1, \dots, a_k$. Zmieniając ewentualnie numerację elementów $b_1, \dots, b_k$ można zatem przyjąć, że $b_i = a_i$ dla $i = 1, \dots, k$. Wtedy mamy także $\ell_i = m_i$ i wobec tego

$$p_{\tau_i} = (X - a_i)^{m_i}.$$

Udowodniliśmy w ten sposób część rezultatów zapowiadanych w (7.5).

W dalszej części dowodu będziemy więc zakładać, że macierz $\mathbf{m}(\tau, \mathcal{A})$ ma postać

$$\mathbf{m}(\tau, \mathcal{A}) = M'_1 \oplus \dots \oplus M'_k, \quad (7.7)$$

gdzie $M'_i = \bigoplus_{j=1}^{s_i} J_{d_{ij}}(a_i)$. Przedstawieniu (7.7) odpowiada rozkład przestrzeni V na sumę prostą podprzestrzeni niezmienniczych

$$V = U_1 \oplus \dots \oplus U_k, \quad (7.8)$$

przy czym dla endomorfizmu indukowanego τ_i na U_i mamy $\mathbf{m}(\tau_i, \mathcal{A}_i) = M'_i$ w pewnej bazie $\mathcal{A}_i$ podprzestrzeni U_i .

Zauważmy teraz, że dla wielomianu $q_i = X - a_i$ mamy

$$q_i^{m_i}(\tau)(U_i) = 0$$

gdyż $q_i^{m_i}$ jest wielomianem minimalnym endomorfizmu $\tau_i = \tau|_{U_i}$. Innymi słowy

$$U_i \subseteq \ker q_i^{m_i}(\tau) = V_i,$$

gdzie V_i jest podprzestrzenią występującą w rozkładzie (7.2). Ponieważ

$$U_1 \oplus \dots \oplus U_k = V = V_1 \oplus \dots \oplus V_k,$$

wynika stąd, że $U_i = V_i$ dla wszystkich $i = 1, \dots, k$. Tak więc endomorfizmy τ_i indukowane przez τ na podprzestrzeniach U_i okazują się być endomorfizmami indukowanymi przez τ na podprzestrzeniach V_i . Zatem na podstawie (7.3) mamy

$$\mathbf{m}(\tau_i, \mathcal{B}_i) = J_{m_{i1}}(a_i) \oplus \dots \oplus J_{m_{ir_i}}(a_i)$$

oraz z drugiej strony

$$\mathbf{m}(\tau_i, \mathcal{A}_i) = M'_i = J_{d_{i1}}(a_i) \oplus \dots \oplus J_{d_{is_i}}(a_i).$$

Są to dwie postaci kanoniczne Jordana macierzy endomorfizmu τ_i , którego wielomian minimalny jest równy $(X - a_i)^{m_i}$. Na podstawie twierdzenia 7.1.1 te dwie macierze są identyczne. A więc w szczególności $s_i = r_i$, $d_{ij} = m_{ij}$ dla $i = 1, 2, \dots, k$, $j = 1, 2, \dots, r_i$, co kończy dowód twierdzenia. $\square$

WNIOSEK 7.2.1. (Jednoznaczność postaci kanonicznej Jordana.)

Postać kanoniczna Jordana macierzy endomorfizmu τ jest wyznaczona jednoznacznie przez endomorfizm τ z dokładnością do kolejności występujących w niej klatek Jordana.

WNIOSEK 7.2.2. *Endomorfizm τ jest diagonalizowalny wtedy i tylko wtedy gdy wszystkie klatki Jordana występujące w postaci kanonicznej Jordana macierzy endomorfizmu τ są jednowymiarowe.*

Dowód. Niech $J = \bigoplus_{i=1}^k \bigoplus_{j=1}^{r_i} J_{m_{ij}}(a_i)$ będzie postacią kanoniczną Jordana macierzy endomorfizmu τ . Jeśli wszystkie klatki Jordana $J_{m_{ij}}(a_i)$ są jednowymiarowe, to znaczy $J_{m_{ij}}(a_i) = J_1(a_i)$, to macierz

$$J = J_1(a_1) \oplus \dots \oplus J_1(a_1) \oplus J_1(a_2) \oplus \dots \oplus J_1(a_2) \oplus \dots \oplus J_1(a_k) \oplus \dots \oplus J_1(a_k)$$

jest oczywiście macierzą diagonalną.

Z drugiej strony, jeśli endomorfizm τ jest diagonalizowalny i macierz diagonalna

$$A = [a_1] \oplus \dots \oplus [a_1] \oplus [a_2] \oplus \dots \oplus [a_2] \oplus \dots \oplus [a_k] \oplus \dots \oplus [a_k]$$

jest macierzą endomorfizmu τ w pewnej bazie przestrzeni V , to macierz A jest sumą prostą jednowymiarowych klatek Jordana i wobec tego A jest postacią kanoniczną Jordana macierzy endomorfizmu τ . Wobec jednoznaczności postaci kanonicznej Jordana wynika stąd, że w postaci kanonicznej Jordana macierzy endomorfizmu τ wszystkie występujące tam klatki Jordana są jednowymiarowe. $\square$

TWIERDZENIE 7.2.2. *Niech ρ i τ będą endomorfizmami przestrzeni wektorowej V nad ciałem K i niech wielomiany minimalne endomorfizmów ρ i τ rozkładają się nad ciałem K na iloczyny czynników liniowych. Endomorfizmy ρ i τ są podobne wtedy i tylko wtedy gdy ich postaci kanoniczne Jordana są identyczne.*

Dowód. Jeśli ρ i τ mają w odpowiednich bazach przestrzeni V tę samą macierz, to są podobne na podstawie twierdzenia 2.2.6.

Na odwrót, jeśli $\rho = \sigma^{-1}\tau\sigma$, gdzie $\sigma \in \text{Aut } V$ i $\mathcal{A}$ jest dla endomorfizmu ρ bazą kanoniczną Jordana przestrzeni V oraz $\mathbf{m}(\rho, \mathcal{A}) = J$ jest postacią kanoniczną Jordana macierzy endomorfizmu ρ , to na podstawie twierdzenia 2.2.4 mamy

$$J = \mathbf{m}(\rho, \mathcal{A}) = \mathbf{m}(\sigma, \mathcal{A})^{-1} \cdot \mathbf{m}(\tau, \mathcal{A}) \cdot \mathbf{m}(\sigma, \mathcal{A}) = \mathbf{m}(\tau, \mathcal{B})$$

gdzie $\mathcal{B} = \sigma(\mathcal{A})$. Zatem macierze endomorfizmów ρ i τ mają tę samą postać kanoniczną Jordana. $\square$

7.3 Wielomian charakterystyczny

W tym rozdziale rozpatrujemy wyłącznie endomorfizmy $\tau \in \text{End}_K V$, których wielomiany minimalne rozkładają się nad ciałem K na iloczyn czynników liniowych.

Niech więc τ będzie endomorfizmem przestrzeni wektorowej V i niech

$$p_\tau = (X - a_1)^{m_1} \cdots (X - a_k)^{m_k} \quad (7.9)$$

będzie wielomianem minimalnym endomorfizmu τ , gdzie $a_1, \dots, a_k$ są elementami ciała K , przy czym $a_i \neq a_j$ dla $i \neq j$ oraz $m_i \geq 1$ dla $i = 1, \dots, k$. Wtedy na podstawie twierdzenia 5.2.2 o rozkładzie, przestrzeń V ma rozkład

$$V = V_1 \oplus \cdots \oplus V_k, \quad (7.10)$$

gdzie V_i jest podprzestrzenią niezmienniczą endomorfizmu τ oraz endomorfizm indukowany $\tau_i = \tau|_{V_i}$ na przestrzeni V_i ma wielomian minimalny $(X - a_i)^{m_i}$. Jak wiemy z twierdzenia o rozkładzie,

$$V_i = \ker(\tau - a_i \mathbf{1}_V)^{m_i}.$$

DEFINICJA 7.3.1. *Krotnością algebraiczną $k_{\text{alg}}(a_i)$ wartości własnej a_i endomorfizmu τ nazywamy wymiar podprzestrzeni $V_i = \ker(\tau - a_i \mathbf{1}_V)^{m_i}$:*

$$k_{\text{alg}}(a_i) = \dim \ker(\tau - a_i \mathbf{1}_V)^{m_i}.$$

DEFINICJA 7.3.2. *Krotnością geometryczną $k_g(a_i)$ wartości własnej a_i endomorfizmu τ nazywa się wymiar przestrzeni wektorów własnych endomorfizmu τ należących do wartości własnej a_i :*

$$k_g(a_i) = \dim \ker(\tau - a_i \mathbf{1}_V).$$

Ponieważ $\ker(\tau - a_i \mathbf{1}_V) \subseteq \ker(\tau - a_i \mathbf{1}_V)^{m_i}$ więc dla każdej wartości własnej a_i endomorfizmu τ mamy

$$k_g(a_i) \leq k_{\text{alg}}(a_i).$$

Wracając do rozkładu (7.10), podprzestrzeń V_i ma bazę $\mathcal{B}_i$ taką, że dla endomorfizmu τ_i indukowanego przez τ na V_i mamy

$$\mathbf{m}(\tau_i, \mathcal{B}_i) = J_{m_{i1}}(a_i) \oplus \cdots \oplus J_{m_{ir_i}}(a_i). \quad (7.11)$$

Z twierdzenia 7.2.1 wiemy, że rozmiary $m_{i1} \geq \cdots \geq m_{ir_i}$ występujących tu klatek Jordana są wyznaczone jednoznacznie przez endomorfizm τ . Krotność algebraiczna wartości własnej a_i jest więc stopniem macierzy (7.11):

$$k_{\text{alg}}(a_i) = m_{i1} + \cdots + m_{ir_i}, \quad i = 1, \dots, k.$$

Inaczej mówiąc, krotność algebraiczna wartości własnej a_i wskazuje ile razy wartość własna a_i pojawia się na głównej przekątnej postaci kanonicznej Jordana macierzy endomorfizmu τ . Z drugiej strony, z twierdzenia 7.1.2 wiemy, że liczba r_i klatek Jordana należących do wartości własnej a_i jest równa wymiarowi podprzestrzeni wektorów własnych endomorfizmu τ należących do wartości własnej a_i . Zatem

$$k_g(a_i) = \dim \ker(\tau - a_i \mathbf{1}_V) = r_i, \quad i = 1, \dots, k.$$

Pokażemy teraz, że przy pomocy pojęć krotności algebraicznej i geometrycznej wartości własnych można sformułować warunek konieczny i wystarczający diagonalizowalności endomorfizmu.

Twierdzenie 7.3.1. *Niech wielomian minimalny p_τ endomorfizmu $\tau \in \text{End}_K V$ będzie iloczynem czynników liniowych nad ciałem K i niech $a_1, \dots, a_k \in K$ będą wszystkimi pierwiastkami wielomianu p_τ . Endomorfizm τ jest diagonalizowalny wtedy i tylko wtedy gdy*

$$k_g(a_i) = k_{alg}(a_i) \quad \text{dla każdego } i = 1, \dots, k.$$

Dowód. Na podstawie wniosku 7.2.2, jeśli endomorfizm τ jest diagonalizowalny, to wszystkie klatki Jordana występujące w postaci kanonicznej macierzy endomorfizmu τ są jednowymiarowe. Stąd

$$r_i = k_g(a_i) \leq k_{alg}(a_i) = r_i$$

i wobec tego $k_g(a_i) = k_{alg}(a_i)$.

Z drugiej strony, jeśli $k_g(a_i) = k_{alg}(a_i)$ dla każdego $i = 1, \dots, k$, to $\dim \ker(\tau - a_i \mathbf{1}_V) = \dim \ker(\tau - a_i \mathbf{1}_V)^{m_i}$ skąd wobec $\ker(\tau - a_i \mathbf{1}_V) \subseteq \ker(\tau - a_i \mathbf{1}_V)^{m_i}$ wynika, że

$$\ker(\tau - a_i \mathbf{1}_V) = \ker(\tau - a_i \mathbf{1}_V)^{m_i} = V_i.$$

Wobec tego każda z podprzestrzeni V_i jest przestrzenią wektorów własnych endomorfizmu τ i ponieważ $V = V_1 \oplus \dots \oplus V_k$, endomorfizm τ jest diagonalizowalny w każdej bazie $\mathcal{B}$ postaci $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$, gdzie $\mathcal{B}_i$ jest jakąkolwiek bazą podprzestrzeni V_i . $\square$

Przedstawieniu (7.11) odpowiada rozkład podprzestrzeni V_i na sumę prostą podprzestrzeni τ -niezmienniczych

$$V_i = U_{i1} \oplus \dots \oplus U_{ir_i}$$

takich, że endomorfizm τ_{ij} indukowany na U_{ij} przez τ ma macierz $J_{m_{ij}}(a_i)$. Wyznamy teraz wielomiany minimalne endomorfizmów τ_{ij} .

Lemat 7.3.1. *Jeśli wielomian minimalny endomorfizmu τ ma postać (7.9), to dla wszystkich i, j , $1 \leq i \leq k$, $1 \leq j \leq r_i$ mamy*

$$(a) \quad p_{\tau_i} = (X - a_i)^{m_i}.$$

$$(b) \quad p_{\tau_{ij}} = (X - a_i)^{m_{ij}}.$$

Dowód. (a) zauważyliśmy już w twierdzeniu 5.2.2 o rozkładzie.

(b) Endomorfizm $\sigma_{ij} = \tau_{ij} - a_i \mathbf{1}_{U_{ij}}$ podprzestrzeni U_{ij} ma macierz $J_{m_{ij}}(0)$, jest więc endomorfizmem nilpotentnym stopnia m_{ij} . Zatem σ_{ij} ma wielomian minimalny $X^{m_{ij}}$ natomiast τ_{ij} ma wielomian minimalny $(X - a_i)^{m_{ij}}$. $\square$

Definicja 7.3.3. Wielomiany $p_{\tau_{ij}} = (X - a_i)^{m_{ij}}$, $j = 1, \dots, r_i$ nazywamy *dzielnikami elementarnymi* endomorfizmu τ należącymi do wartości własnej a_i endomorfizmu τ .

Dzielnik elementarny $p_{\tau_{ij}}$ endomorfizmu τ jest zatem wielomianem minimalnym endomorfizmu τ_{ij} indukowanego przez τ na podprzestrzeni U_{ij} .

Krotność algebraiczna $k_{alg}(a_i) = m_{i1} + \dots + m_{ir_i} =: n_i$ wartości własnej a_i endomorfizmu τ jest w następujący sposób związana z dzielnikami elementarnymi endomorfizmu τ , których pierwiastkiem jest a_i :

$$(X - a_i)^{n_i} = \prod_{1 \leq j \leq r_i} p_{\tau_{ij}}.$$

Definicja 7.3.4. Niech wielomian minimalny endomorfizmu τ ma postać (7.9) i niech $n_1, \dots, n_k$ będą krotnościami algebraicznymi wartości własnych $a_1, \dots, a_k$. Wielomian

$$F_\tau = (X - a_1)^{n_1} \dots (X - a_k)^{n_k} \in K[X]$$

nazywamy *wielomianem charakterystycznym* endomorfizmu τ .

Przywołując definicję 7.3.3 możemy więc powiedzieć, że wielomian charakterystyczny endomorfizmu τ jest iloczynem wszystkich dzielników elementarnych endomorfizmu τ :

$$F_\tau = \prod_{1 \leq i \leq k} \prod_{1 \leq j \leq r_i} p_{\tau_{ij}}.$$

Ponadto, krotność algebraiczna wartości własnej a_i endomorfizmu τ jest równa krotności a_i jako pierwiastka wielomianu charakterystycznego endomorfizmu τ .

Twierdzenie 7.3.2. *Wielomian minimalny endomorfizmu τ jest dzielnikiem wielomianu charakterystycznego endomorfizmu τ . Ponadto, stopień wielomianu minimalnego endomorfizmu przestrzeni n -wymiarowej V jest nie większy niż n .*

Dowód. Wielomian p_τ ma rozkład (7.9), zatem wobec $m_i \leq n_i$ dla $i = 1, \dots, k$ mamy oczywiście $p_\tau \mid F_\tau$. Z drugiej strony, na podstawie twierdzenia o rozkładzie mamy

$$n = \dim V = \dim V_1 + \dots + \dim V_k = n_1 + \dots + n_k = \deg F_\tau.$$

A więc $\deg p_\tau = \sum m_i \leq \sum n_i = \deg F_\tau = n$. □

Zauważmy, że podzielność $p_\tau \mid F_\tau$ pociąga natychmiast następujące twierdzenie.

Twierdzenie 7.3.3. (Twierdzenie Cayleya-Hamiltona). $F_\tau(\tau) = 0_V$.

7.4 Wyznacznik i ślad

Także w tym rozdziale rozpatrujemy wyłącznie endomorfizmy $\tau \in \mathbf{End}_K V$, których wielomiany minimalne rozkładają się nad ciałem K na iloczyn czynników liniowych.

Definicja 7.4.1. Niech $a_1, \dots, a_k$ będą wartościami własnymi endomorfizmu τ i niech $n_1, \dots, n_k$ będą ich krotnościami algebraicznymi.

Wyznacznikiem $\det \tau$ endomorfizmu τ nazywamy następujący iloczyn potęg wartości własnych endomorfizmu τ

$$\det \tau := a_1^{n_1} \cdots a_k^{n_k}.$$

Śladem $\mathrm{tr} \tau$ endomorfizmu τ nazywamy następującą sumę wartości własnych endomorfizmu τ

$$\mathrm{tr} \tau := n_1 a_1 + \dots + n_k a_k.$$

Ponieważ a_i jest n_i -krotną wartością własną endomorfizmu τ , można także powiedzieć, że wyznacznik endomorfizmu τ jest iloczynem wartości własnych tego endomorfizmu z uwzględnieniem ich krotności algebraicznych a ślad endomorfizmu τ jest sumą jego wartości własnych z uwzględnieniem ich krotności algebraicznych. Jeśli wielomian charakterystyczny endomorfizmu τ

$$F_\tau = (X - a_1)^{n_1} \cdots (X - a_k)^{n_k}$$

przedstawimy w postaci

$$F_\tau = X^n - c_1 X^{n-1} + \dots + (-1)^n c_n,$$

to na podstawie wzorów Viete'a otrzymamy

$$\mathrm{tr} \tau = c_1, \quad \det \tau = c_n.$$

Twierdzenie 7.4.1. *Endomorfizm τ jest osobliwy wtedy i tylko wtedy, gdy $\det \tau = 0$.*

Dowód. Endomorfizm τ jest osobliwy wtedy i tylko wtedy, gdy endomorfizm

$$\tau - 0 \cdot \mathbf{1}_V$$

jest osobliwy, a więc wtedy i tylko wtedy gdy $0 \in K$ jest wartością własną endomorfizmu τ . Osobliwość endomorfizmu τ jest więc równoważna temu, że $\det \tau = 0$. $\square$

Twierdzenie 7.4.2. *Wielomian charakterystyczny, wyznacznik i ślad endomorfizmu są niezmiennikami podobieństwa endomorfizmów.*

Dowód. Endomorfizmy podobne mają identyczne postacie kanoniczne Jordana (na podstawie twierdzenia 7.2.2), zatem mają te same wartości własne a także krotności wartości własnych. Zatem mają równe wielomiany charakterystyczne, wyznaczniki i ślady. $\square$

Wprowadzimy teraz pojęcie wyznacznika i śladu macierzy o elementach z ciała algebraicznie domkniętego.

Definicja 7.4.2. Niech $A \in M_n(K)$, gdzie K jest ciałem algebraicznie domkniętym i niech $V = K^n$ będzie n -wymiarową przestrzenią wektorową współrzędnych nad ciałem K . Niech $\mathcal{A}$ będzie dowolną bazą przestrzeni V i niech τ będzie endomorfizmem przestrzeni V takim, że

$$m(\tau, \mathcal{A}) = A.$$

- (a) *Wartościami własnymi* macierzy A nazywamy wartości własne endomorfizmu τ .
- (b) *Krotnością algebraiczną* wartości własnej a_i macierzy A nazywamy krotność algebraiczną wartości własnej a_i endomorfizmu τ .
- (c) *Wyznacznikiem* $\det A$ macierzy A nazywamy wyznacznik $\det \tau$ endomorfizmu τ .
- (d) *Śladem* $\text{tr } A$ macierzy A nazywamy ślad $\text{tr } \tau$ endomorfizmu τ .
- (e) *Wielomianem charakterystycznym* F_A macierzy A nazywamy wielomian charakterystyczny F_τ endomorfizmu τ .
- (f) *Postacią kanoniczną Jordana* J_A macierzy A nazywamy postać kanoniczną Jordana J_τ macierzy endomorfizmu τ .

Zauważmy, że jeśli także ρ jest endomorfizmem przestrzeni V i ρ ma w jakiejś bazie przestrzeni V macierz A , to endomorfizmy ρ i τ są podobne, więc mają równe wartości własne, krotności algebraiczne wartości własnych, wyznaczniki, ślady, wielomiany charakterystyczne i postacie kanoniczne. Stąd wynika, że nasza definicja wartości własnych macierzy A , jej wyznacznika, śladu, wielomianu charakterystycznego i postaci kanonicznej Jordana nie zależy od wyboru endomorfizmu τ .

Twierdzenie 7.4.3. *Niech $A \in M_n(K)$, gdzie K jest ciałem algebraicznie domkniętym. Dla każdego $x \in K$ mamy*

$$\det(A - xI) = (-1)^n F_A(x).$$

Dowód. Niech V będzie n -wymiarową przestrzenią wektorową nad ciałem K z bazą $\mathcal{A}$ i niech τ będzie endomorfizmem przestrzeni V takim, że

$$m(\tau, \mathcal{A}) = A.$$

Niech J będzie postacią kanoniczną Jordana macierzy endomorfizmu τ i niech $\mathcal{B}$ będzie bazą przestrzeni V taką, że

$$m(\tau, \mathcal{B}) = J = \bigoplus_{i=1}^k \bigoplus_{j=1}^{r_i} J_{m_{ij}}(a_i).$$

Wtedy

$$\mathbf{m}(\tau - x\mathbf{1}_V, \mathcal{B}) = J - xI = \bigoplus_{i=1}^k \bigoplus_{j=1}^{r_i} J_{m_{ij}}(a_i - x).$$

A więc baza $\mathcal{B}$ jest także bazą kanoniczną Jordana endomorfizmu $\tau - x\mathbf{1}_V$ i endomorfizm ten ma wartości własne $a_i - x$, przy czym krotność algebraiczna wartości własnej $a_i - x$ endomorfizmu $\tau - x\mathbf{1}_V$ jest równa krotności algebraicznej n_i wartości własnej a_i endomorfizmu τ . Ponieważ $\mathbf{m}(\tau - x\mathbf{1}_V, \mathcal{A}) = A - xI$, więc

$$\det(A - xI) = \det(\tau - x\mathbf{1}_V) = \prod_{i=1}^k (a_i - x)^{n_i} = (-1)^n F_A(x).$$

□

Zbadamy teraz związek pomiędzy wprowadzonymi przez nas wyznacznikiem $\det A$ i śladem $\text{tr } A$ macierzy A oraz klasycznymi pojęciami wyznacznika i śladu macierzy. Dla rozróżnienia tych pojęć wprowadźmy następujące oznaczenia dla klasycznych funkcji wyznacznika i śladu macierzy $A = [a_{ij}] \in M_n(K)$:

$$d : M_n(K) \rightarrow K, \quad d(A) = \sum \pm a_{i_1 1} a_{i_2 2} \cdots a_{i_n n},$$

gdzie sumowanie przebiega po wszystkich permutacjach $(i_1, i_2, \dots, i_n)$ zbioru $(1, 2, \dots, n)$ i znak składnika jest $+$ jeśli permutacja numerów wierszowych jest parzysta i $-$ w przypadku przeciwnym, oraz

$$s : M_n(K) \rightarrow K, \quad s(A) = \sum_{i=1}^n a_{ii}.$$

A więc $d(A)$ jest klasycznym wyznacznikiem macierzy A zaś $s(A)$ jest klasycznym śladem macierzy A . Zauważmy, że jeśli J jest macierzą w postaci kanonicznej Jordana, to $\det J = d(J)$ oraz $\text{tr } J = s(J)$.

LEMAT 7.4.1. *Dla dowolnych macierzy $A, B \in M_n(K)$,*

$$d(AB) = d(BA) \quad \text{oraz} \quad s(AB) = s(BA).$$

Dowód. Pierwsza własność wynika z twierdzenia Cauchy'ego znanego z podstawowego kursu algebry liniowej. Na podstawie tego twierdzenia mamy $d(AB) = d(A)d(B)$ i wobec tego

$$d(AB) = d(A)d(B) = d(B)d(A) = d(BA).$$

Drugą dowodzimy następująco. Niech $A = [a_{ij}]$, $B = [b_{ij}]$. Element diagonalny w i -tym wierszu macierzy AB ma postać

$$a_{i1}b_{1i} + \cdots + a_{in}b_{ni},$$

zatem

$$s(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji}a_{ij} = s(BA).$$

□

LEMAT 7.4.2. *Niech $A, B \in M_n(K)$ i niech B będzie macierzą odwracalną. Wtedy*

$$d(B^{-1}AB) = d(A) \quad \text{oraz} \quad s(B^{-1}AB) = s(A).$$

Dowód. Wykorzystując poprzedni lemat i łączność mnożenia macierzy otrzymujemy

$$d(B^{-1}AB) = d((B^{-1}A)B) = d(B(B^{-1}A)) = d((BB^{-1})A) = d(A).$$

Podobnie,

$$s(B^{-1}AB) = s((B^{-1}A)B) = s(B(B^{-1}A)) = s((BB^{-1})A) = s(A).$$

□

Teraz możemy przedstawić metodę obliczania wyznacznika $\det A$ i śladu $\operatorname{tr} A$ macierzy A .

TWIERDZENIE 7.4.4. *Niech K będzie ciałem algebraicznie domkniętym i niech $A \in M_n(K)$. Wtedy*

$$\det A = d(A) \quad \text{oraz} \quad \operatorname{tr} A = s(A).$$

Dowód. Niech J będzie postacią kanoniczną Jordana macierzy A i niech $J = B^{-1}AB$, gdzie B jest pewną macierzą odwracalną. Wtedy

$$\det A = \det J = d(J) = d(B^{-1}AB) = d(A),$$

oraz

$$\operatorname{tr} A = \operatorname{tr} J = s(J) = s(B^{-1}AB) = s(A).$$

□

WNIOSEK 7.4.1. *Niech K będzie ciałem algebraicznie domkniętym i niech $A, B \in M_n(K)$ oraz $a, b \in K$. Wtedy*

$$\operatorname{tr}(aA + bB) = a \operatorname{tr} A + b \operatorname{tr} B.$$

Dowód. $s(aA + bB) = as(A) + bs(B)$.

□

A więc ślad tr jest funkcjonalem liniowym na przestrzeni wektorowej macierzy $M_n(K)$.

Definicje 7.3.4 i 7.4.2 akcentują geometryczny aspekt wielomianu charakterystycznego endomorfizmu i wyznacznika macierzy. Wyznacznik macierzy A jest więc iloczynem wartości własnych endomorfizmu o macierzy A , z uwzględnieniem ich krotności. Dla znalezienia wartości wyznacznika $\det A$ a także wielomianu charakterystycznego F_τ należałoby zatem najpierw znaleźć wszystkie wartości własne endomorfizmu τ a także ich krotności algebraiczne. Klasyczne definicje wyznacznika macierzy i wielomianu charakterystycznego endomorfizmu nie oferują żadnej geometrycznej interpretacji, ale pozwalają znaleźć wartość wyznacznika macierzy i jej wielomian charakterystyczny poprzez arytmetyczne manipulacje na elementach macierzy A .

Rozdział 8

Postać kanoniczna wymierna

Ostatnie zmiany 22.01.2009 r.

W tym rozdziale wskażemy postać kanoniczną endomorfizmu $\tau \in \text{End}_K V$, gdzie K jest dowolnym ciałem i wielomian minimalny endomorfizmu τ jest dowolnym wielomianem nad K . Strategia postępowania będzie podobna do tej, której użyliśmy w rozdziale 7 dla dowodu istnienia i jednoznaczności postaci kanonicznej Jordana. Tam, obok twierdzenia o rozkładzie, główną rolę grał opis postaci kanonicznej macierzy endomorfizmów nilpotentnych, czyli endomorfizmów, których wielomian minimalny ma postać X^m . Dla uzyskania postaci kanonicznej wymiernej znajdziemy przede wszystkim uogólnienia głównych twierdzeń o endomorfizmach nilpotentnych na przypadek endomorfizmów, których wielomian minimalny ma postać q^m , gdzie q jest dowolnym wielomianem nierozkładalnym nad ciałem K .

8.1 Podprzestrzenie cykliczne

Pojęcie podprzestrzeni τ -cyklicznej przestrzeni V wprowadziliśmy w rozdziale 6 (definicja 6.1.1). Przypomnijmy, że podprzestrzenie τ -cykliczną generowaną przez wektor $v \in V$ nazywamy podprzestrzeń $K[\tau]v = \text{lin}(\tau^k(v) : k \in \mathbb{N} \cup \{0\}) = \{f(\tau)(v) \in V : f \in K[X]\}$. W tym rozdziale zbadamy własności podprzestrzeni τ -cyklicznych dla wszystkich endomorfizmów τ , których wielomiany minimalne są potęgami wielomianów nierozkładalnych.

LEMAT 8.1.1. *Niech $\tau \in \text{End}_K V$ i niech wielomian minimalny p_τ endomorfizmu τ będzie potęgą unormowanego wielomianu $q \in K[X]$ nierozkładalnego nad ciałem K :*

$$p_\tau = q^m = c_0 + c_1 X + \cdots + c_{d-1} X^{d-1} + X^d \in K[X]. \quad (8.1)$$

Niech $U = K[\tau]v$ będzie podprzestrzenią τ -cykliczną przestrzeni V generowaną przez wektor $v \in V$ i niech τ_1 będzie endomorfizmem przestrzeni U indukowanym przez τ na U .
(a) Wielomian minimalny p_{τ_1} endomorfizmu τ_1 ma postać

$$p_{\tau_1} = q^\ell, \quad 1 \leq \ell \leq m.$$

(b) Jeśli wektor $v \in V$ spełnia warunek

$$q^{m-1}(\tau)(v) \neq 0, \quad (8.2)$$

to wielomian minimalny p_{τ_1} endomorfizmu τ_1 jest równy wielomianowi minimalnemu endomorfizmu τ :

$$p_{\tau_1} = p_\tau = q^m.$$

(c) Jeśli wektor $v \in V$ spełnia warunek (8.2), to wektory

$$v, \tau(v), \dots, \tau^{d-1}(v) \quad (8.3)$$

tworzą bazę podprzestrzeni U i wobec tego $\dim U = d = \deg p_\tau = \deg p_{\tau_1}$.

Dowód. (a), (b) p_{τ_1} dzieli p_τ na podstawie lematu 5.1.1. Wobec $p_\tau = q^m$ i nierozkładalności wielomianu q wynika stąd, że $p_{\tau_1} = q^\ell$ gdzie $\ell \leq m$. W szczególności, $q^\ell(\tau)(v) = p_{\tau_1}(\tau)(v) = 0$, podczas gdy dla $v \in V$ spełniającego (8.2) mamy $q^{m-1}(\tau)(v) \neq 0$. Zatem (8.2) pociąga, że $\ell > m - 1$ i wobec tego $\ell = m$ oraz $p_{\tau_1} = q^m = p_\tau$.

(c) Na podstawie (b) i twierdzenia 6.1.1 wektory (8.3) tworzą bazę podprzestrzeni U . $\square$

Jest rzeczą godną uwagi, że informacje o podprzestrzeniach cyklicznych jakie uzyskaliśmy w lemacie 8.1.1 pozwalają znaleźć nietrywialne oszacowanie dla stopnia wielomianu minimalnego endomorfizmu τ .

LEMAT 8.1.2. *Jeśli τ jest endomorfizmem przestrzeni V i wielomian minimalny p_τ endomorfizmu τ jest potęgą wielomianu nierozkładalnego nad K , to*

$$\deg p_\tau \leq \dim V.$$

Ponadto, $\deg p_\tau = \dim V$ wtedy i tylko wtedy gdy przestrzeń V jest τ -cykliczna.

Dowód. Endomorfizm τ wyznacza podprzestrzeń τ -cykliczną U generowaną przez wektor $v \in V$ spełniający (8.2). Na podstawie lematu 8.1.1 podprzestrzeń U ma wymiar $\deg p_\tau$. Zatem $\deg p_\tau = \dim U \leq \dim V$.

Jeśli $\deg p_\tau = \dim V$, to wobec $\dim U = \deg p_\tau$ mamy $\dim U = \dim V$. Zatem $U = V$.

Jeśli V jest przestrzenią τ -cykliczną, to $\deg p_\tau = \dim V$ na podstawie wniosku 6.1.1. $\square$

TWIERDZENIE 8.1.1. *Dla każdego endomorfizmu τ przestrzeni V stopień wielomianu minimalnego p_τ endomorfizmu τ jest nie większy od wymiaru przestrzeni V :*

$$\deg p_\tau \leq \dim V.$$

Ponadto, $\deg p_\tau = \dim V$ wtedy i tylko wtedy gdy przestrzeń V jest τ -cykliczna.

Dowód. Endomorfizm τ ma wielomian minimalny

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k},$$

gdzie q_i są unormowanymi wielomianami nierozkładalnymi nad ciałem K . Na podstawie twierdzenia 5.2.2 o rozkładzie, przestrzeń V ma rozkład

$$V = V_1 \oplus \cdots \oplus V_k,$$

gdzie V_i jest podprzestrzenią niezmienniczą endomorfizmu τ oraz endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny $q_i^{m_i}$. Niech

$$d_i = \deg q_i^{m_i} \quad \text{ i } \quad n_i = \dim V_i.$$

Wtedy mamy

$$\deg p_\tau = \sum d_i \leq \sum n_i = \dim V,$$

gdzie wykorzystaliśmy nierówność $d_i \leq n_i$ z lematu 8.1.2.

Założmy teraz, że $\deg p_\tau = \dim V = n$. Wtedy

$$n = \deg p_\tau = \sum d_i \leq \sum n_i = \dim V = n,$$

skąd wynika, że $d_i = n_i$ dla każdego i . Stąd na podstawie lematu 8.1.2 wynika, że V_i jest przestrzenią τ -cykliczną.

Pokażemy teraz, że jeśli $V_i = K[\tau]v_i$ oraz $v = v_1 + \dots + v_k$, to $V = K[\tau]v$. Wystarczy pokazać, że wektory

$$v, \tau(v), \tau^2(v), \dots, \tau^{n-1}(v)$$

są liniowo niezależne. Przypuśćmy więc, że pewna kombinacja liniowa tych wektorów z nie wszystkimi współczynnikami zerowymi jest wektorem zerowym:

$$a_0v + a_1\tau(v) + a_2\tau^2(v) + \dots + a_\ell\tau^\ell(v) = 0.$$

Możemy oczywiście założyć, że $\ell \leq n-1$ oraz $a_\ell \neq 0$. Rozważmy wielomian

$$f = a_0 + a_1X + a_2X^2 + \dots + a_\ell X^\ell \in K[X].$$

Wtedy $\deg f = \ell$ oraz $f(\tau)(v) = 0$. A więc

$$f(\tau)(v_1) + \dots + f(\tau)(v_k) = f(\tau)(v) = 0.$$

Ponieważ przestrzenie V_i są τ -niezmiennicze, są one także $f(\tau)$ -niezmiennicze. Zatem $f(\tau)(v_i) \in V_i$ i ponieważ $V = V_1 \oplus \dots \oplus V_k$, wynika stąd, że $f(\tau)(v_i) = 0$ dla każdego i . Ponieważ podprzestrzeń V_i jest podprzestrzenią τ -cykliczną generowaną przez wektor v_i oraz endomorfizm τ_i indukowany na V_i ma wielomian minimalny $p_{\tau_i} = q_i^{m_i}$, wynika stąd na podstawie twierdzenia 6.1.1(a), że $q_i^{m_i} \mid f$ dla każdego i . Wobec tego także $p_\tau \mid f$ wbrew temu, że $\deg f = \ell \leq n-1 < n = \deg p_\tau$. A więc wektory $v, \tau(v), \dots, \tau^{n-1}(v)$ są liniowo niezależne i wobec tego tworzą bazę przestrzeni V . Zatem

$$V = \text{lin}(v, \tau(v), \dots, \tau^{n-1}(v)) \subseteq K[\tau]v \subseteq V,$$

skąd wynika, że $V = K[\tau]v$ jest przestrzenią τ -cykliczną.

Jeśli V jest przestrzenią τ -cykliczną, to $\dim V = \deg p_\tau$ na podstawie wniosku 6.1.1. $\square$

Przystępujemy teraz do wyznaczenia macierzy endomorfizmu τ_1 podprzestrzeni U w bazie skonstruowanej w lemacie 8.1.1.

LEMAT 8.1.3. *Niech $\tau \in \text{End}_K V$ i niech wielomian minimalny p_τ endomorfizmu τ ma postać (8.1). Niech U będzie przestrzenią τ -cykliczną generowaną przez wektor $v \in V$ spełniający (8.2) i niech τ_1 będzie endomorfizmem przestrzeni τ -cyklicznej U indukowanym przez τ na U . Niech*

$$\mathcal{B} = \{v, \tau(v), \tau^2(v), \dots, \tau^{d-1}(v)\}$$

będzie uporządkowaną bazą przestrzeni U wyznaczoną w lemacie 8.1.1. Wtedy

$$m(\tau_1, \mathcal{B}) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & 0 & \dots & 0 & -c_2 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 1 & -c_{d-1} \end{bmatrix}.$$

Dowód. Dla $0 \leq i \leq d-2$ mamy $\tau(\tau^i(v)) = \tau^{i+1}(v)$. Ponadto

$$\tau(\tau^{d-1}(v)) = \tau^d(v) = \tau^d(v) - p_\tau(\tau)(v) = -c_0v - c_1\tau(v) - \dots - c_{d-1}\tau^{d-1}(v). \quad \square$$

DEFINICJA 8.1.1. Macierz

$$S(p) := \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & 0 & \dots & 0 & -c_2 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 1 & -c_{d-1} \end{bmatrix}$$

nazywa się macierzą *stowarzyszoną* z wielomianem

$$p = c_0 + c_1X + \dots + c_{d-1}X^{d-1} + X^d.$$

Zauważmy, że dla wielomianu $p = X^d$ macierz stowarzyszona $S(p)$ jest osobliwą klatką Jordana J_d stopnia d . Lemat 8.1.3 mówi, że (przy odpowiednich założeniach o endomorfizmie τ i wektorze v generującym podprzestrzeń τ -cykliczną U) endomorfizm $\tau|_U$ przestrzeni τ -cyklicznej U ma w odpowiedniej bazie przestrzeni U macierz stowarzyszoną z wielomianem minimalnym endomorfizmu τ . Dla endomorfizmów nilpotentnych otrzymujemy rezultat znany nam już z lematu 6.2.3.

8.2 Drugie twierdzenie o rozkładzie

Dla uzyskania postaci kanonicznej wymiernej macierzy dowolnego endomorfizmu τ przestrzeni wektorowej V potrzebujemy - obok twierdzenia 5.2.2 o rozkładzie - drugiego twierdzenia o rozkładzie, zastępującego twierdzenie 6.3.1, które użyliśmy w przypadku endomorfizmów nilpotentnych. Twierdzenie 8.2.1 jest uogólnieniem twierdzenia 6.3.1, które otrzymujemy w przypadku gdy $q = X^m$. Jednakże dowód twierdzenia 6.3.1 (oparty na twierdzeniu 6.2.1) nie przenosi się na przypadek ogólny, w związku z czym zastosujemy w dowodzie technikę znaną w teorii modułów nad pierścieniami euklidesowymi.

Dla uproszczenia będziemy opuszczać nawiasy w zapisie obrazu wektora przez endomorfizm. A więc dla endomorfizmu τ przestrzeni V oraz $v \in V$ zamiast $\tau(v)$ będziemy pisać τv . Dla wielomianu $f \in K[X]$ zamiast $f(\tau)(v)$ będziemy zatem pisać $f(\tau)v$.¹

TWIERDZENIE 8.2.1. *Niech τ będzie endomorfizmem przestrzeni V nad ciałem K . Wtedy istnieją podprzestrzenie τ -cykliczne $U_1, \dots, U_r$ przestrzeni V takie, że*

$$V = U_1 \oplus \dots \oplus U_r.$$

Dowód. Najpierw zauważmy, że przestrzeń V zawsze można przedstawić jako (zwykłą) sumę pewnych podprzestrzeni τ -cyklicznych. Rzeczywiście, wybieramy dowolny wektor niezerowy $v_1 \in V$ i rozpatrujemy podprzestrzeń τ -cykliczną V_1 generowaną przez wektor v_1 :

$$V_1 = K[\tau]v_1.$$

Jeśli $V_1 \neq V$, to wybieramy wektor $v_2 \in V \setminus V_1$ i rozpatrujemy $V_2 = K[\tau]v_2$, podprzestrzeń τ -cykliczną generowaną przez v_2 . Wtedy $V_1 \subsetneq V_1 + V_2 \subseteq V$. Jeśli $V_1 + V_2 \neq V$, to wybieramy $v_3 \in V \setminus (V_1 + V_2)$ i rozpatrujemy $V_1 + V_2 + V_3$, gdzie V_3 jest podprzestrzenią τ -cykliczną generowaną przez v_3 . Ponieważ przestrzeń V ma skończony wymiar, kontynuując to postępowanie otrzymamy w końcu $V_1 + \dots + V_r = V$ dla pewnego $r \geq 1$.

¹Bardziej radykalnym uproszczeniem byłoby pisać fv zamiast $f(\tau)v$ ale tego już tutaj nie zrobimy. Przy takim zapisie łatwo zauważyć, że przestrzeń wektorowa V jest modułem nad pierścieniem $K[X]$. Ten punkt widzenia w połączeniu z techniką dowodu twierdzenia 8.2.1 prowadzi do znacznie ogólniejszego twierdzenia: każdy skończony generowany moduł nad pierścieniem euklidesowym jest sumą prostą podmodułów cyklicznych.

Najmniejszą liczbę składników w przedstawieniu przestrzeni V w postaci sumy podprzestrzeni τ -cyklicznych nazywamy τ -rangą przestrzeni V i oznaczamy $\text{rank}_\tau V$ lub $\text{rank } V$, jeśli z kontekstu wiadomo o jaki endomorfizm τ chodzi.

Przeprowadzimy dowód indukcyjny twierdzenia ze względu na $\text{rank } V$.

Jeśli $\text{rank } V = 1$, to przestrzeń V jest τ -cykliczna i twierdzenie jest prawdziwe.

Założmy więc, że $\text{rank}_\tau V = r > 1$ i twierdzenie jest prawdziwe dla każdej podprzestrzeni τ -niezmienniczej W przestrzeni V takiej, że $\text{rank } W = r - 1$.

Niech $V = V_1 + \dots + V_r$ będzie (zwykłą) sumą podprzestrzeni τ -cyklicznych V_i . Niech $v_i \in V_i$ będzie generatorem V_i jako podprzestrzeni τ -cyklicznej, to znaczy $V_i = K[\tau]v_i$. Zbiór wektorów $\{v_1, \dots, v_r\}$ nazywamy *minimalnym zbiorem generatorów* przestrzeni V .

Wobec $V = V_1 + \dots + V_r$ każdy wektor przestrzeni V jest sumą wektorów należących do podprzestrzeni V_i . W szczególności więc istnieją wielomiany $h_1, \dots, h_r \in K[X]$ takie, że

$$h_1(\tau)v_1 + \dots + h_r(\tau)v_r = 0. \quad (8.4)$$

Gdyby wektor $0 \in V$ miał tylko jedno przedstawienie w tej postaci (dla $h_1(\tau)v_1 = \dots = h_r(\tau)v_r = 0$), to stąd wynikałoby także, że *każdy* wektor $v \in V$ ma jednoznaczne przedstawienie w postaci (8.4) i wobec tego przestrzeń V jest *sumą prostą* podprzestrzeni τ -cyklicznych V_i . Musimy więc rozpatrzyć przypadek, gdy nie wszystkie składniki $h_i(\tau)v_i$ w równości (8.4) są równe zero.

Spośród wszystkich minimalnych zbiorów generatorów $\{v_1, \dots, v_r\}$ i spośród wszystkich równości (8.4) wybieramy zbiór generatorów i równość, w której występuje niezerowy wielomian h_i z najmniejszym stopniem. Zmieniając ewentualnie porządek elementów w wybranym zbiorze generatorów $\{v_1, \dots, v_r\}$ możemy zakładać, że h_1 jest niezerowym wielomianem o najmniejszym stopniu we wszystkich równościach postaci (8.4). Udowodnimy teraz dwa pomocnicze fakty stwierdzające własności wielomianu h_1 .

Fakt 1. Jeśli

$$b_1(\tau)v_1 + \dots + b_r(\tau)v_r = 0 \quad (8.5)$$

dla pewnych $b_1, \dots, b_r \in K[X]$, to h_1 dzieli b_1 .

Jeśli h_1 nie dzieli b_1 , to na podstawie twierdzenia o dzieleniu z resztą istnieją wielomiany $g, t \in K[X]$ takie, że

$$b_1 = h_1g + t, \quad t \neq 0, \quad \deg t < \deg h_1.$$

Wtedy z (8.4) otrzymujemy $h_1(\tau)g(\tau)v_1 + \dots + h_r(\tau)g(\tau)v_r = 0$ i odejmując tę równość od równości (8.5) mamy

$$(b_1 - h_1g)(\tau)v_1 + \dots + (b_r - h_rg)(\tau)v_r = 0.$$

W tej równości współczynnik $b_1 - h_1g$ jest równy t i ma stopień mniejszy niż h_1 , co jest sprzeczne z wyborem naszego minimalnego zbioru generatorów i współczynnika h_1 . A więc h_1 dzieli b_1 .

Fakt 2. h_1 dzieli wszystkie współczynniki $h_2, \dots, h_r$ występujące w wybranej równości (8.4).

Przypuśćmy, że h_1 nie dzieli h_2 . Zatem istnieją $g, t \in K[X]$ takie, że

$$h_2 = h_1g + t, \quad t \neq 0, \quad \deg t < \deg h_1.$$

Z równości (8.4) otrzymujemy wtedy

$$h_1(\tau)(v_1 + g(\tau)v_2) + t(\tau)v_2 + h_3(\tau)v_3 + \dots + h_r(\tau)v_r = 0. \quad (8.6)$$

Zauważmy, że $\{v_1 + g(\tau)v_2, v_2, \dots, v_r\}$ jest także minimalnym zbiorem generatorów przestrzeni V . Wystarczy pokazać, że

$$V'_1 + V_2 = V_1 + V_2, \quad \text{gdzie} \quad V'_1 = K[\tau](v_1 + g(\tau)v_2).$$

Inkluzja $\subseteq$ jest oczywista gdyż $v_1 + g(\tau)v_2 \in V_1 + V_2$. Z drugiej strony $v_1 = (v_1 + g(\tau)v_2) - g(\tau)v_2 \in V'_1 + V_2$ i oczywiście $v_2 \in V'_1 + V_2$. Stąd $V'_1 + V_2 \supseteq V_1 + V_2$.

Równość (8.6), w której występuje wielomian t stopnia mniejszego niż $\deg h_1$, jest sprzeczna z wyborem $\{v_1, \dots, v_r\}$ oraz h_1 . Zatem h_1 dzieli h_2 . Zmieniając kolejność generatorów $v_2, \dots, v_r$ pokażemy w ten sam sposób, że h_1 dzieli wszystkie pozostałe wielomiany $h_3, \dots, h_r$.

Na podstawie Faktu 2 istnieją $g_2, \dots, g_r \in K[X]$ takie, że

$$h_2 = h_1g_2, \dots, h_r = h_1g_r.$$

Wykorzystując teraz wielomiany $g_2, \dots, g_r \in K[X]$ rozpatrzmy następujący zbiór wektorów:

$$v_1^* = v_1 + g_2(\tau)v_2 + \dots + g_r(\tau)v_r, \quad v_2, \dots, v_r. \quad (8.7)$$

Jest to minimalny zbiór generatorów przestrzeni V gdyż jak łatwo sprawdzić

$$K[\tau]v_1^* + V_2 + \dots + V_r = V_1 + V_2 + \dots + V_r = V.$$

Niech $U_1 = K[\tau]v_1^*$ będzie podprzestrzenią τ -cykliczną generowaną przez v_1^* oraz niech $W := V_2 + \dots + V_r$. Mamy więc rozkład

$$V = U_1 + W.$$

Pokażemy, że w istocie $V = U_1 \oplus W$. Wystarczy pokazać, że $U_1 \cap W = 0$. Przypuśćmy więc, że

$$v \in U_1 \cap W.$$

Wtedy istnieją $b_1, b_2, \dots, b_r \in K[X]$ takie, że

$$v = b_1(\tau)v_1^* = b_2(\tau)v_2 + \dots + b_r(\tau)v_r,$$

a więc także

$$b_1(\tau)v_1^* - b_2(\tau)v_2 - \dots - b_r(\tau)v_r = 0.$$

Zastępując tutaj v_1^* kombinacją liniową $v_1 + g_2(\tau)v_2 + \dots + g_r(\tau)v_r$ otrzymamy

$$b_1(\tau)v_1 + (g_2 - b_2)(\tau)v_2 + \dots + (g_r - b_r)(\tau)v_r = 0.$$

Na podstawie Faktu 1 wiemy, że h_1 dzieli b_1 , zatem $b_1 = gh_1$ dla pewnego $g \in K[X]$ i w rezultacie otrzymujemy

$$\begin{aligned} v = b_1(\tau)v_1^* &= g(\tau)(h_1(\tau)v_1^*) \\ &= g(\tau)(h_1(\tau)v_1 + h_1(\tau)g_2(\tau)v_2 + \dots + h_1(\tau)g_r(\tau)v_r) \\ &= g(\tau)(h_1(\tau)v_1 + h_2(\tau)v_2 + \dots + h_r(\tau)v_r) \\ &= g(\tau)0 = 0. \end{aligned}$$

A więc $U_1 \cap W = 0$ co wraz z $V = U_1 + W$ dowodzi, że $V = U_1 \oplus W$. Ponieważ $\text{rank } W = r - 1$, więc na podstawie założenia indukcyjnego podprzestrzeń W jest sumą prostą $r - 1$ podprzestrzeni τ -cyklicznych. Wobec tego V jest sumą prostą r podprzestrzeni τ -cyklicznych. $\square$

Sformułujemy teraz dokładniejszą wersję twierdzenia 8.2.1 w przypadku, gdy wielomian minimalny endomorfizmu τ jest potęgą wielomianu nierozkładalnego.

Twierdzenie 8.2.2. *Niech endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny postaci $p_\tau = q^m$, gdzie q jest unormowanym wielomianem nierozkładalnym nad ciałem K . Niech $d = \deg p_\tau$. Wtedy istnieją podprzestrzenie τ -cykliczne $U_1, \dots, U_r$ przestrzeni V takie, że*

(a) $V = U_1 \oplus \dots \oplus U_r$.

(b) Endomorfizm indukowany τ_i na podprzestrzeni U_i ma wielomian minimalny $p_{\tau_i} = q^{s_i}$, przy czym

$$m = s_1 \geq \dots \geq s_r \geq 1.$$

(c) $d = \dim U_1 = \deg q^{s_1} \geq \dim U_2 = \deg q^{s_2} \geq \dots \geq \dim U_r = \deg q^{s_r}$.

Dowód. (a) Przedstawienie przestrzeni V w postaci sumy prostej podprzestrzeni cyklicznych udowodniliśmy już w twierdzeniu 8.2.1.

(b) Postać wielomianu minimalnego endomorfizmu indukowanego na podprzestrzeni cyklicznej opisaliśmy w lemacie 8.1.1(a). Niech $s = \max s_i$. Wtedy $q^s(\tau)(U_i) = 0$ dla każdego $i = 1, \dots, r$ i wobec tego także $q^s(\tau)(V) = 0$. Stąd wynika, że wielomian minimalny q^m endomorfizmu τ dzieli q^s , a więc $m \leq s$. Z drugiej strony, z lematu 8.1.1(a) mamy $s \leq m$. Zatem $m = s$, czyli jeden z wielomianów minimalnych endomorfizmów τ_i jest równy q^m . Zmieniając ewentualnie numerację podprzestrzeni U_i możemy zakładać, że $i = 1$ oraz $m = s_1 \geq \dots \geq s_r$.

(c) Ponieważ endomorfizm indukowany τ_i na podprzestrzeni U_i ma wielomian minimalny q^{s_i} , więc na podstawie lematu 6.1.1 mamy $\dim U_i = \deg q^{s_i}$. Ponadto, nierówności w punkcie (c) wynikają z nierówności w punkcie (b). $\square$

A oto wersja macierzowa twierdzenia 8.2.2.

Twierdzenie 8.2.3. *Niech endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny postaci $p_\tau = q^m$, gdzie q jest unormowanym wielomianem nierozkładalnym nad ciałem K . Wtedy istnieje baza $\mathcal{B}$ przestrzeni V oraz ciąg liczb naturalnych $m = s_1 \geq s_2 \geq \dots \geq s_r \geq 1$ takie, że*

$$\mathbf{m}(\tau, \mathcal{B}) = S(q^{s_1}) \oplus \dots \oplus S(q^{s_r}) = \begin{bmatrix} S(q^{s_1}) & 0 & \dots & 0 & 0 \\ 0 & S(q^{s_2}) & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & S(q^{s_r}) \end{bmatrix}.$$

Dowód. Rozważmy rozkład przestrzeni V z twierdzenia 8.2.2. Ponieważ endomorfizm indukowany τ_i na podprzestrzeni U_i ma wielomian minimalny q^{s_i} , więc istnieje wektor $v_i \in U_i$ taki, że $q^{s_i-1}(\tau)(v_i) \neq 0$. Wobec tego, na podstawie lematu 8.1.3, istnieje baza $\mathcal{B}_i$ przestrzeni U_i taka, że $\mathbf{m}(\tau_i, \mathcal{B}_i) = S(q^{s_i})$. Zatem w bazie $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ endomorfizm τ ma macierz $S(q^{s_1}) \oplus \dots \oplus S(q^{s_r})$. $\square$

Uwaga 8.2.1. W szczególnym przypadku gdy $q = X$ endomorfizm τ jest nilpotentny i dla dowolnego $s \in \mathbb{N}$ macierz $S(q^s) = S(X^s) = J_s(0)$ jest osobliwą kłatką Jordana stopnia s . Twierdzenie 8.2.3 powtarza zatem rezultat twierdzenia 6.3.2 dla endomorfizmów nilpotentnych: $\mathbf{m}(\tau, \mathcal{B}) = J_{s_1} \oplus \dots \oplus J_{s_r}$.

8.3 Postać kanoniczna

Przystępujemy teraz do opisu postaci kanonicznej *wymiernej* (nazywanej też postacią kanoniczną Frobeniusa) dowolnego endomorfizmu $\tau \in \text{End}_K V$, gdzie K jest dowolnym ciałem. Załóżmy, że endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k},$$

gdzie q_i są unormowanymi wielomianami nierozkładalnymi nad ciałem K , $q_i \neq q_j$ dla $i \neq j$ oraz $m_i \geq 1$ są liczbami naturalnymi.

Na podstawie twierdzenia 5.2.2 o rozkładzie przestrzeni V ma rozkład

$$V = V_1 \oplus \cdots \oplus V_k,$$

gdzie V_i jest podprzestrzenią niezmienniczą endomorfizmu τ oraz endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny $q_i^{m_i}$. Opis działania endomorfizmu τ na przestrzeni V redukuje się więc do opisu działania endomorfizmu τ na podprzestrzeniach niezmienniczych V_i . Wykorzystując teraz twierdzenie 8.2.2 otrzymujemy natychmiast następujące ogólne twierdzenie.

TWIERDZENIE 8.3.1. (Postać kanoniczna wymierna macierzy endomorfizmu)

Niech endomorfizm $\tau \in \text{End}_K V$ ma wielomian minimalny

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k},$$

gdzie q_i są unormowanymi wielomianami nierozkładalnymi nad ciałem K , $q_i \neq q_j$ dla $i \neq j$ oraz $m_i \geq 1$ są liczbami naturalnymi.

Wtedy istnieje baza $\mathcal{A}$ przestrzeni V taka, że

$$\mathbf{m}(\tau, \mathcal{A}) = M_1 \oplus \cdots \oplus M_k,$$

gdzie

$$\begin{aligned} M_i &= S(q_i^{s_{i1}}) \oplus \cdots \oplus S(q_i^{s_{ir_i}}) \\ &= \begin{bmatrix} S(q_i^{s_{i1}}) & 0 & \cdots & 0 & 0 \\ 0 & S(q_i^{s_{i2}}) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & 0 & S(q_i^{s_{ir_i}}) \end{bmatrix} \end{aligned}$$

oraz $m_i = s_{i1} \geq s_{i2} \geq \cdots \geq s_{ir_i}$ jest pewnym ciągiem liczb naturalnych.

Rozdział 9

Moduły nad pierścieniami ideałów głównych

Ostatnie zmiany 20.01.2007 r.

W tym rozdziale rozpatrujemy moduły nad pierścieniami przemiennymi. Przedstawimy tutaj dwa główne twierdzenia o strukturze modułów torsyjnych nad pierścieniami ideałów głównych i modułów skończenie generowanych nad pierścieniami euklidesowymi. Pokażemy, że są one uogólnieniami twierdzenia o rozkładzie i twierdzenia o rozkładzie przestrzeni wektorowej na sumę prostą podprzestrzeni τ -cyklicznych.

9.1 Moduły - definicje i przykłady

Niech R będzie dowolnym pierścieniem przemiennym. Będziemy rozpatrywać uogólnienie pojęcia przestrzeni wektorowej nad ciałem K , w którym rolę ciała K przejmie pierścień R .

Niech M będzie addytywną grupą abelową. Grupę abelową M nazywamy R -modułem, jeśli na grupie M określone jest *działanie zewnętrzne* z pierścieniem skalarów R :

$$R \times M \rightarrow M, \quad (a, m) \mapsto am$$

i działanie to ma następujące własności:

$$a(m_1 + m_2) = am_1 + am_2 \quad (9.1)$$

$$(a_1 + a_2)m = a_1m + a_2m \quad (9.2)$$

$$(a_1a_2)m = a_1(a_2m) \quad (9.3)$$

$$1m = m \quad (9.4)$$

dla wszystkich $a, a_1, a_2 \in R$, $m_1, m_2, m \in M$.

Działanie zewnętrzne $R \times M \rightarrow M$, $(a, m) \mapsto am$ nazywamy także *mnożeniem* elementów grupy abelowej M przez elementy pierścienia R .

Tak więc grupa abelowa M jest R -modułem wtedy i tylko wtedy, gdy na M jest określone mnożenie elementów grupy M przez elementy pierścienia R spełniające warunki (9.1)–(9.4).

Zauważmy, że jeśli M jest R -modułem, to działanie zewnętrzne na M ma także następujące własności: dla dowolnych $a, b \in R$ oraz $m \in M$,

$$0m = 0, \quad (-1)m = -m, \quad (a - b)m = am - bm.$$

Mamy bowiem $(a - b)m + bm = (a - b + b)m = am$, skąd wynika, że $(a - b)m$ jest różnicą elementów am i bm w grupie M . Kładąc $a = b = 1$ otrzymujemy $0m = (1 - 1)m =$

$1m - 1m = m - m = 0$, natomiast kładąc $a = 0, b = 1$ otrzymujemy $(-1)m = (0 - 1)m = 0m - 1m = 0 - m = -m$.

Następujące przykłady wskazują jak pojemne jest pojęcie modułu.

Przykład 9.1.1. Każda grupa abelowa M jest $\mathbb{Z}$ –modułem jeśli za działanie zewnętrzne na M przyjąć mnożenie elementów grupy M przez liczby całkowite.

Przykład 9.1.2. Pierścień R jest R –modułem jeśli za działanie zewnętrzne na R przyjąć mnożenie w pierścieniu R . Ogólniej, każdy ideał $\mathcal{J}$ pierścienia R jest R –modułem, jeśli za działanie zewnętrzne na $\mathcal{J}$ przyjąć mnożenie elementów ideału $\mathcal{J}$ przez elementy pierścienia R . Na odwrót, jeśli addytywna podgrupa $\mathcal{J}$ pierścienia R jest R –modułem (z mnożeniem jako działaniem zewnętrznym), to $\mathcal{J}$ jest ideałem pierścienia R .

Przykład 9.1.3. Niech K będzie ciałem. Każda przestrzeń wektorowa V nad K jest K –modułem. Każdy K –moduł V jest przestrzenią wektorową nad ciałem K .

Przykład 9.1.4. Niech $R = K[X]$ będzie pierścieniem wielomianów jednej zmiennej nad ciałem K . Niech $M = V$ będzie przestrzenią wektorową nad ciałem K i niech τ będzie endomorfizmem przestrzeni V . Rozpatrzmy homomorfizm podstawiania endomorfizmu τ przestrzeni V w miejsce zmiennej wielomianu $f \in K[X]$:

$$\varphi_\tau : K[X] \rightarrow \text{End}_K V, \quad \varphi_\tau(f) = f(\tau).$$

Przy pomocy tego homomorfizmu określamy działanie zewnętrzne na V z pierścieniem skalarów $K[X]$ następująco:

$$K[X] \times V \rightarrow V, \quad (f, v) \mapsto f(\tau)(v).$$

A więc $fv := f(\tau)(v)$. Łatwo sprawdza się, że tak zdefiniowane działanie zewnętrzne na V spełnia warunki (9.1)–(9.4). A więc, dla $f \in K[X]$ oraz $v_1, v_2 \in V$ mamy

$$f(v_1 + v_2) = f(\tau)(v_1 + v_2) = f(\tau)(v_1) + f(\tau)(v_2) = fv_1 + fv_2.$$

Dla $f_1, f_2 \in K[X]$ oraz $v \in V$ mamy

$$(f_1 + f_2)v = (f_1 + f_2)(\tau)(v) = (f_1(\tau) + f_2(\tau))(v) = f_1(\tau)(v) + f_2(\tau)(v) = f_1v + f_2v,$$

$$(f_1f_2)v = (f_1f_2)(\tau)(v) = (f_1(\tau)f_2(\tau))(v) = f_1(\tau)(f_2(\tau)(v)) = f_1(\tau)(f_2v) = f_1(f_2v)$$

i wreszcie $1v = 1_V(v) = v$.

W ten sposób przestrzeń wektorowa V jest $K[X]$ –modułem. Tak skonstruowany $K[X]$ –moduł oznaczamy V_τ . Ten przykład stanowi podstawę zastosowań teorii modułów w algebrze liniowej.

DEFINICJA 9.1.1. Niech M będzie R –modułem. Podmodułem N modułu M nazywamy podgrupę N addytywnej grupy M zamkniętą ze względu na mnożenie przez elementy pierścienia R , to znaczy, podgrupę N grupy M spełniającą warunek

$$a \in R, \quad m \in N \implies am \in N.$$

Warunek ten zapisujemy krótko jako

$$RN \subseteq N.$$

Jeśli N jest podmodułem modułu M , to piszemy $N < M$.

Przykład 9.1.5. Każda podgrupa N grupy abelowej M jest podmodułem $\mathbb{Z}$ –modułu M . Każdy ideał $\mathcal{J}$ pierścienia R jest podmodułem R –modułu R .

Każda podprzestrzeń wektorowa przestrzeni wektorowej V nad ciałem K jest podmodułem K –modułu V .

Jeśli $\mathcal{J}$ jest ideałem pierścienia R oraz M jest R –modułem, to

$$\mathcal{J}M := \{a_1m_1 + \cdots + a_nm_n \in M : a_i \in \mathcal{J}, m_i \in M, n \in \mathbb{N}\}$$

jest podmodułem modułu M .

Przykład 9.1.6. Rozpatrzmy $K[X]$ –moduł V_τ z przykładu 9.1.4. Dla endomorfizmu τ przestrzeni wektorowej V , każda podprzestrzeń τ –niezmiennicza U przestrzeni V (to znaczy, podprzestrzeń spełniająca $\tau(U) \subseteq U$) jest podmodułem $K[X]$ –modułu V_τ . Rzeczywiście, jeśli $\tau(U) \subseteq U$ to także $\tau^m(U) \subseteq U$ dla każdej liczby naturalnej m , a stąd otrzymujemy łatwo $f(\tau)(U) \subseteq U$ dla każdego wielomianu $f \in K[X]$. A więc jeśli $u \in U$, to $fu \in U$ dla każdego $f \in K[X]$. Także na odwrót, jeśli U jest podmodułem $K[X]$ –modułu V_τ , to dla każdego wielomianu $f \in K[X]$ i dla każdego $u \in U$ mamy $f(\tau)(u) = fu \in U$, a więc także $f(\tau)(U) \subseteq U$. Zatem w szczególności także $\tau(U) \subseteq U$.

A więc U jest podmodułem $K[X]$ –modułu V_τ wtedy i tylko wtedy, gdy U jest podprzestrzenią niezmienniczą endomorfizmu τ .

DEFINICJA 9.1.2. Niech M będzie R –modułem.

(a) Mówimy, że podmoduł N jest *generowany* przez zbiór $S \subseteq M$, jeśli każdy element podmodułu N można przedstawić w postaci kombinacji liniowej skończonego podzbioru zbioru S ze współczynnikami z pierścienia R . Zbiór S nazywamy *zbiorem generatorów* modułu M .

(b) Mówimy, że podmoduł N jest *skończenie generowany*, jeśli jest generowany przez podzbiór skończony modułu M .

(c) Podmoduł N modułu M nazywamy podmodułem *cyklicznym*, jeśli N jest generowany przez zbiór jednoelementowy, to znaczy, jeśli istnieje element $m \in M$ taki, że

$$N = \{am \in M : a \in R\}.$$

Podmoduł cykliczny N generowany przez element m oznaczamy $N = Rm$.

Przykład 9.1.7. Każda podgrupa cykliczna N grupy abelowej M jest podmodułem cyklicznym $\mathbb{Z}$ –modułu M .

Każdy ideał główny $\mathcal{J}$ pierścienia R jest podmodułem cyklicznym R –modułu R .

Każda jedno-wymiarowa podprzestrzeń wektorowa przestrzeni wektorowej V nad ciałem K jest podmodułem cyklicznym K –modułu V .

Każda skończenie wymiarowa przestrzeń wektorowa V nad ciałem K jest skończenie generowanym K –modułem.

Jeśli V jest skończenie wymiarową przestrzenią wektorową nad ciałem K , to rozpatrywany w przykładzie 9.1.4 $K[X]$ –moduł V_τ jest skończenie generowany. Rzeczywiście, jeśli $\mathcal{B} = \{v_1, \dots, v_n\}$ jest bazą przestrzeni V , to każdy wektor $v \in V$ ma (jednoznaczne) przedstawienie w postaci

$$v = a_1v_1 + \cdots + a_nv_n, \quad a_1, \dots, a_n \in K.$$

Ponieważ $K \subset K[X]$, więc przedstawienie to można także traktować jako przedstawienie elementu $v \in V_\tau$ w postaci kombinacji liniowej elementów zbioru $\mathcal{B}$ ze współczynnikami z pierścienia $K[X]$. Baza $\mathcal{B}$ przestrzeni V jest więc równocześnie zbiorem generatorów $K[X]$ –modułu V_τ . Skończenie wymiarowa przestrzeń wektorowa jest więc skończenie generowanym $K[X]$ –modułem.

9.1.1 Operacje na modułach

Niech M będzie R -modułem i niech N_1 i N_2 będą podmodułami modułu M . Wtedy przekrój $N_1 \cap N_2$ jest podmodułem R -modułu M . Również suma $N_1 + N_2$ podgrup grupy abelowej M jest podmodułem R -modułu M . Nazywamy ją *sumą podmodułów* N_1 i N_2 . Podobnie określa się sumę dowolnej rodziny podmodułów R -modułu M (jako sumę podgrup grupy abelowej M).

W szczególności, jeśli $N_1 + N_2 = M$, to mówimy, że moduł M jest *sumą* podmodułów N_1 i N_2 . Jeśli w dodatku $N_1 \cap N_2 = 0$, to mówimy, że M jest *sumą prostą* podmodułów N_1 i N_2 . Piszemy wtedy $M = N_1 \oplus N_2$. Zauważmy, że tak jak dla grup abelowych, M jest sumą prostą podmodułów N_1 i N_2 wtedy i tylko wtedy gdy każdy element $m \in M$ ma *jednoznaczne* przedstawienie w postaci $m = n_1 + n_2$, gdzie $n_1 \in N_1$, $n_2 \in N_2$.

Mówimy, że podmoduł N_1 modułu M jest *składnikiem prostym* modułu M jeśli istnieje podmoduł N_2 taki, że $M = N_1 \oplus N_2$. W przestrzeni wektorowej V każda podprzestrzeń jest składnikiem prostym przestrzeni V . Jest to wysoce wyjątkowa sytuacja. Na przykład, w $\mathbb{Z}$ -module $\mathbb{Z}$ (czyli w grupie liczb całkowitych) jedynymi składnikami prostymi są podmoduły 0 i $\mathbb{Z}$. Grupa $\mathbb{Z}$ nie ma bowiem nietrywialnych rozkładów na sumę prostą podgrup. Podgrupy $\mathbb{Z}$ są grupami cyklicznymi $a\mathbb{Z}$, gdzie $a \in \mathbb{Z}$. Jeśli $\mathbb{Z} = a\mathbb{Z} \oplus b\mathbb{Z}$, to $a\mathbb{Z} \cap b\mathbb{Z} = 0$ co jest możliwe tylko dla $a = 0$ lub $b = 0$.

Niech M i M' będą R -modułami. Na iloczynie kartezjańskim $M \times M'$ grup abelowych M i M' określamy działanie zewnętrzne z pierścieniem skalarów R kładąc

$$a(m, m') = (am, am')$$

dla $a \in R, m \in M, m' \in M'$. Łatwo sprawdza się, że w ten sposób grupa abelowa $M \times M'$ staje się R -modułem. Nazywamy go *iloczynem prostym* (lub kartezjańskim) modułów M i M' . Podobnie określa się iloczyn prosty dowolnej rodziny R -modułów $M_i, i \in I$.

Jeśli N jest podmodułem R -modułu M , to grupę ilorazową M/N można w sposób naturalny traktować jako R -moduł. Wystarczy zauważyć, że iloczyn kompleksowy warstwy $m + N$ przez element $a \in R$ (jednoelementowy podzbiór pierścienia R) zawiera się w dokładnie jednej warstwie modułu M względem podmodułu N :

$$a(m + N) = \{a(m + n) : n \in N\} = \{am + an : n \in N\} \subseteq am + N.$$

Warstwę $am + N$ nazywamy iloczynem warstwy $m + N$ przez skalar $a \in R$. W ten sposób otrzymujemy działanie zewnętrzne na grupie abelowej M/N z pierścieniem skalarów R . Łatwo sprawdza się, że grupa ilorazowa M/N staje się R -modułem. Moduł ten nazywamy *modułem ilorazowym* modułu M względem podmodułu N .

9.1.2 Homomorfizmy modułów

Niech M i M' będą R -modułami. *Homomorfizmem* $h : M \rightarrow M'$ nazywamy homomorfizm h grupy abelowej M w grupę abelową M' spełniający warunek

$$h(am) = ah(m) \quad \forall a \in R, m \in M.$$

Jądro i obraz homomorfizmu modułów określamy oczywiście jako jądro i obraz homomorfizmu grup abelowych. Również takie terminy jak *epimorfizm*, *monomorfizm*, *izomorfizm* modułów interpretujemy tak jak w teorii grup (jako, odpowiednio, homomorfizm surjektywny, iniektywny, bijektywny).

Niech N będzie podmodułem R -modułu M i niech

$$\kappa : M \rightarrow M/N, \quad \kappa(m) = m + N$$

będzie homomorfizmem kanonicznym grup abelowych. Wtedy $\kappa(am) = am + N = a(m + N) = a\kappa(m)$ dla $a \in R, m \in M$. Zatem κ jest homomorfizmem modułów. Nazywamy go homomorfizmem *kanonicznym* modułów.

Sformułujemy teraz podstawowe twierdzenie o homomorfizmach modułów.

TWIERDZENIE 9.1.1. (Twierdzenie o faktoryzacji.)

Jeśli $h : M \rightarrow M'$ jest homomorfizmem R -modułów, $N = \ker h$ oraz $\kappa : M \rightarrow M/N$ jest homomorfizmem kanonicznym, to istnieje dokładnie jeden iniektywny homomorfizm $h_* : M/N \rightarrow M'$ taki, że $h = h_* \circ \kappa$, to znaczy taki, że następujący diagram jest przemien-ny:

$$\begin{array}{ccc} M & \xrightarrow{h} & M' \\ & \searrow \kappa \quad \nearrow h_* & \\ & M/N & \end{array}$$

Dowód. Można użyć dokładnie takich samych argumentów jak w przypadku przestrzeni wektorowych. $\square$

9.1.3 Moduły wolne

Moduły wolne stanowią uogólnienie przestrzeni wektorowych i grup abelowych wolnych. Charakterystyczną cechą tych obiektów jest istnienie bazy, to znaczy podzbioru liniowo niezależnego generującego obiekt.

DEFINICJA 9.1.3. R -moduł F nazywa się *modułem wolnym*, jeśli istnieje podzbiór $\mathcal{B}$ modułu F o następujących własnościach:

(a) $\mathcal{B}$ jest zbiorem liniowo niezależnym, to znaczy, dla każdych $x_1, \dots, x_n \in R$ i $b_1, \dots, b_n \in \mathcal{B}$ mamy implikację

$$x_1 b_1 + \dots + x_n b_n = 0 \Rightarrow x_1 = \dots = x_n = 0.$$

(b) Każdy element $m \in F$ ma przedstawienie w postaci

$$m = x_1 b_1 + \dots + x_n b_n, \quad b_1, \dots, b_n \in \mathcal{B}, \quad x_1, \dots, x_n \in R. \quad (9.5)$$

Podzbiór $\mathcal{B}$ modułu wolnego F spełniający warunki (a) i (b) nazywa się *bazą* modułu wolnego F . Baza modułu wolnego F jest więc liniowo niezależnym (nad R) podzbiorem F generującym moduł F . Zgodnie z definicją, moduł F jest modułem wolnym wtedy i tylko wtedy gdy ma bazę.

Tak jak w algebrze liniowej stwierdzamy, że zbiór $\mathcal{B} \subseteq F$ jest bazą modułu wolnego F wtedy i tylko wtedy, gdy każdy element $m \in F$ ma i to tylko jedno przedstawienie w postaci (9.5).

Przykład 9.1.8. R -moduł R jest przykładem R -modułu wolnego. Rzeczywiście, zbiór jednoelementowy $\mathcal{B} = \{1\}$ jest bazą R -modułu R .

$\mathbb{Z}$ -moduł wolny F nazywa się *grupą abelową wolną*.

Każda przestrzeń wektorowa V nad ciałem K jest wolnym K -modułem. Jak bowiem wiadomo z algebry liniowej, każda przestrzeń wektorowa ma bazę.

9.2 Moduły torsyjne

Zbadamy najpierw strukturę modułów *torsyjnych* nad całkowitymi pierścieniami ideałów głównych. Nie będziemy tu zakładać, że rozpatrywane moduły są skończenie generowane.

9.2.1 Moduły torsyjne i ograniczone

DEFINICJA 9.2.1. Niech R będzie dowolnym pierścieniem i niech M będzie R –modułem. Element $m \in M$ nazywamy elementem *torsyjnym*, jeśli istnieje element $a \in R$, $a \neq 0$ taki, że $am = 0$.

Moduł M nazywamy modułem *torsyjnym*, jeśli każdy element modułu M jest torsyjny. Moduł M nazywamy modułem *ograniczonym*, jeśli istnieje element $a \in R$, $a \neq 0$ taki, że dla każdego elementu $m \in M$ mamy $am = 0$.

R –moduł M nazywa się modułem *beztorsyjnym*, jeśli jedynym elementem torsyjnym modułu M jest element zerowy $0 \in M$.

Przykład 9.2.1. Każdy moduł ograniczony jest modułem torsyjnym.

W grupie abelowej M element $m \in M$ jest torsyjny jeśli ma skończony rząd. Jeśli w grupie abelowej M każdy element ma rząd skończony, to M jest torsyjnym $\mathbb{Z}$ –modułem. Natomiast jeśli w grupie abelowej M rzędy wszystkich elementów są wspólnie ograniczone, to M jest ograniczonym $\mathbb{Z}$ –modułem. W szczególności każda skończona grupa abelowa jest ograniczonym $\mathbb{Z}$ –modułem.

Przestrzenie wektorowe nie dostarczają przykładów modułów torsyjnych: każda niezerowa przestrzeń wektorowa V nad ciałem K jest beztorsyjnym K –modułem.

Ogólniej, każdy moduł wolny F nad pierścieniem R bez dzielników zera jest beztorsyjny. Rzeczywiście, jeśli dla pewnego $f \in F$ oraz $0 \neq a \in R$ mamy $af = 0$, to przedstawiając f jako kombinację liniową $f = \sum x_i b_i$ elementów b_i pewnej bazy modułu F otrzymamy dla elementu af przedstawienie $0 = af = \sum ax_i b_i$. Stąd $ax_i = 0$ i wobec tego $x_i = 0$ dla każdego i . Zatem $f = 0$.

Natomiast jeśli pierścień R ma dzielniki zera, to moduł wolny nad R zawsze zawiera elementy torsyjne. Na przykład, pierścień $R = \mathbb{Z}/6\mathbb{Z}$ traktowany jako R –moduł jest modułem wolnym, ale element $2 + 6\mathbb{Z}$ jest elementem torsyjnym gdyż $(3 + 6\mathbb{Z})(2 + 6\mathbb{Z}) = 0 \in \mathbb{Z}/6\mathbb{Z}$.

Jeśli τ jest endomorfizmem przestrzeni wektorowej V nad ciałem K , to przestrzeń V traktowana jako $K[X]$ –moduł V_τ (zob. przykład 9.1.4) jest modułem ograniczonym (a więc także torsyjnym). Rzeczywiście, jeśli $p_\tau \in K[X]$ jest wielomianem minimalnym endomorfizmu τ , to dla każdego $v \in V$ mamy

$$p_\tau v = p_\tau(\tau)(v) = 0_V(v) = 0 \in V.$$

Niech M będzie ograniczonym R –modułem. Zbiór

$$\text{Ann } M = \{x \in R : xM = 0\}$$

nazywamy *anihilatorem* modułu M . Oczywiście $0 \in \text{Ann } M$ ale wobec ograniczoności modułu M zbiór $\text{Ann } M$ zawiera także elementy niezerowe pierścienia R . Łatwo stwierdzić, że $\text{Ann } M$ jest niezerowym ideałem pierścienia R . Jeśli R jest pierścieniem ideałów głównych, to istnieje element $a \in R$ taki, że

$$\text{Ann } M = (a) = aR.$$

Element a ma więc następującą własność: $aM = 0$ i jeśli $xM = 0$, to $a \mid x$. Element a nazywamy *minimalnym elementem ograniczającym* (lub *anihilującym*) modułu ograniczonego M . Element ograniczający a jest odwracalny tylko wtedy gdy M jest modułem

zerowym. Zatem anihilator niezerowego modułu ograniczonego jest niezerowym ideałem właściwym pierścienia R .

Przykład 9.2.2. Jeśli τ jest endomorfizmem przestrzeni wektorowej V nad ciałem K , to wielomian minimalny p_τ endomorfizmu τ jest elementem ograniczającym $K[X]$ –moduł V_τ . Faktycznie p_τ jest *minimalnym elementem ograniczającym* $K[X]$ –moduł V_τ . Jeśli bowiem $f \in K[X]$ jest jakimkolwiek wielomianem anihilującym V_τ , to dla każdego $v \in V$ mamy

$$f(\tau)(v) = fv = 0 \in V$$

i wobec tego $f(\tau) = 0_V$ jest endomorfizmem zerowym przestrzeni V . Dzielimy f przez p_τ z resztą,

$$f = gp_\tau + r,$$

gdzie $r = 0$ lub $r \neq 0$ i stopień r jest mniejszy od stopnia p_τ . Wobec $p_\tau(\tau) = 0_V$ otrzymujemy stąd $0_V = r(\tau)$, skąd wynika, że $r = 0$ (gdyż p_τ jest wielomianem najniższego stopnia zerującym się na τ). A więc $p_\tau \mid f$ oraz $\text{Ann } V_\tau = p_\tau K[X]$.

Zanotujmy teraz bardzo prosty fakt, który okaże się użyteczny w kilku miejscach naszej dyskusji modułów torsyjnych nad pierścieniami ideałów głównych. Przyjmujemy umowę, że pierścień ideałów głównych jest automatycznie pierścieniem całkowitym (przeziennym, bez dzielników zera).

LEMAT 9.2.1. Niech R będzie pierścieniem ideałów głównych, $a, b \in R$, $\text{nwd}(a, b) = 1$. Jeśli M jest R –modułem, $m \in M$ oraz $am = 0$, $bm = 0$, to $m = 0$.

Dowód. Wobec naszych założeń istnieją elementy $x, y \in R$ takie, że $ax + by = 1$. Zatem

$$0 = axm + bym = 1m = m. \quad \square$$

9.2.2 Składowe prymarne

Następująca definicja wprowadza do rozważań podmoduły o kluczowym znaczeniu dla opisu struktury modułów torsyjnych.

DEFINICJA 9.2.2. Niech R będzie pierścieniem ideałów głównych i niech M będzie R –modułem. Niech $p \in R$ będzie elementem pierwszym (nierozkładalnym) pierścienia R . Wtedy zbiór

$$T_p(M) := \{m \in M : \exists \ell \in \mathbb{N} \quad p^\ell m = 0\}$$

jest podmodułem modułu M . Podmoduł $T_p(M)$ nazywa się p –prymarną częścią modułu M lub p –prymarną składową modułu M .

LEMAT 9.2.2. Niech R będzie pierścieniem ideałów głównych i niech M będzie R –modułem ograniczonym. Niech $a \in R$ będzie minimalnym elementem ograniczającym modułu M i niech p będzie elementem pierwszym pierścienia R .

(a) Jeśli $p \nmid a$, to $T_p(M) = 0$.

(b) Jeśli $p \mid a$, to $T_p(M) \neq 0$.

Dowód. (a) Niech $m \in T_p(M)$. Wtedy $p^\ell m = 0$ dla pewnej liczby naturalnej ℓ . Ponieważ $p \nmid a$ i p jest elementem pierwszym, mamy $\text{nwd}(p^\ell, a) = 1$ i wobec tego na podstawie lematu 9.2.1 otrzymujemy $m = 0$.

(b) Załóżmy, że $p \mid a$. Wtedy $a = ph$ dla pewnego $h \in R$ oraz $a \nmid h$ (gdyż w przeciwnym razie p byłby elementem odwracalnym). Wobec tego h nie należy do anihilatora modułu M . Istnieje zatem taki element $m \in M$, że $m_1 := hm \neq 0$. Wtedy $pm_1 = phm = am = 0$, co oznacza, że $m_1 \in T_p(M)$. Zatem $T_p(M) \neq 0$. $\square$

A oto dwa szczególne przypadki sytuacji rozpatrywanej w lemacie 9.2.2.

Przykład 9.2.3. Jeśli M jest skończoną grupą abelową rzędu n , to n jest elementem ograniczającym grupę M , ale niekoniecznie jest *minimalnym* elementem ograniczającym M . W każdym razie minimalny element ograniczający M jest dzielnikiem liczby n i wobec tego jeśli liczba pierwsza p nie dzieli n , to nie dzieli także minimalnego elementu ograniczającego M . Wobec tego na podstawie lematu 9.2.2, jeśli liczba pierwsza p nie dzieli n , to p –prymarna składowa grupy M jest podgrupą zerową:

$$p \nmid n \Rightarrow T_p(M) = 0.$$

Jeśli natomiast $p \mid n$, to można udowodnić, że $T_p(M) \neq 0$.

Jeśli M jest grupą cykliczną rzędu n , to sytuacja jest prosta, gdyż n jest minimalnym elementem ograniczającym M i wobec tego dla liczby pierwszej $p \mid n$ mamy $T_p(M) \neq 0$ na podstawie lematu 9.2.2.

Przykład 9.2.4. Niech p_τ będzie wielomianem minimalnym endomorfizmu τ przestrzeni wektorowej V nad ciałem K . Wtedy p_τ jest minimalnym elementem ograniczającym $K[X]$ –modułu V_τ (zob. przykład 9.2.2). Jeśli wielomian nierozkładalny $q \in K[X]$ nie dzieli p_τ , to q –prymarna składowa $K[X]$ –modułu V_τ jest podmodułem zerowym:

$$T_q(V_\tau) = 0.$$

Natomiast jeśli wielomian nierozkładalny $q \in K[X]$ dzieli wielomian p_τ , to q –składowa modułu V_τ jest niezerowa. Obydwa stwierdzenia są konsekwencją lematu 9.2.2.

Zgodnie z definicją, prymarna składowa $T_p(M)$ składa się z elementów modułu M anihilowanych przez jakąkolwiek potęgę elementu pierwszego p . Pokażemy teraz, że dla modułu ograniczonego w składowej p –prymarnej występują tylko elementy anihilowane przez niezbyt wysokie potęgi elementu p .

LEMAT 9.2.3. *Jeśli p^ℓ jest najwyższą potęgą elementu pierwszego p dzielącą minimalny element anihilujący a modułu ograniczonego M , to*

$$T_p(M) = \{m \in M : p^\ell m = 0\}.$$

Dowód. Oczywiście, zbiór $\{m \in M : p^\ell m = 0\}$ zawiera się w $T_p(M)$. Z drugiej strony, jeśli $m \in T_p(M)$ i $p^k m = 0$ dla pewnej liczby naturalnej $k > \ell$, to $\text{nwd}(p^k, a) = p^\ell$ i wobec tego $xp^k + ya = p^\ell$ dla pewnych $x, y \in R$. Stąd $0 = xp^k m + yam = p^\ell m$. $\square$

9.2.3 Rozkład prymarny

Przystępujemy teraz do dowodu głównego twierdzenia strukturalnego dla modułów torsyjnych.

TWIERDZENIE 9.2.1. *Niech M będzie R –modułem, gdzie R jest pierścieniem ideałów głównych. Jeśli M jest modułem torsyjnym, to M jest sumą prostą wszystkich swoich składowych prymarnych:*

$$M = \bigoplus_p T_p(M),$$

gdzie p przebiega wszystkimi niestowarzyszone elementy pierwsze pierścienia R .

Dowód. Udowodnimy najpierw, że $M = \sum_p T_p(M)$. Niech $m \in M$ oraz $am = 0$ dla $0 \neq a \in R$. Gdyby a był elementem odwracalnym, to $m = 0$ i wobec tego $m = 0 + \dots + 0$ jest sumą elementów należących do prymarnych składowych modułu M . W przeciwnym razie

$$a = up_1^{k_1} \cdots p_r^{k_r},$$

gdzie u jest elementem odwracalnym oraz p_i są parami niestowarzyszonymi elementami pierwszymi pierścienia R . Połóżmy

$$P_i = a/p_i^{k_i}.$$

Elementy $P_1, \dots, P_r$ są względnie pierwsze, zatem w pierścieniu (ideałów głównych) R istnieją elementy $a_1, \dots, a_r$ takie, że

$$a_1P_1 + \dots + a_rP_r = 1.$$

Stąd otrzymujemy

$$a_1P_1m + \dots + a_rP_rm = m,$$

przy czym $a_iP_im \in T_{p_i}(M)$ gdyż $p_i^{k_i} \cdot a_iP_im = a_iam = 0$. Zatem $M = \sum_p T_p(M)$. Przypuśćmy teraz, że

$$m \in T_p(M) \cap \sum_{p_i} T_{p_i}(M),$$

gdzie p, p_i są parami niestowarzyszonymi elementami pierwszymi pierścienia R . Wtedy istnieją liczby naturalne $k, k_1, \dots, k_r$ takie, że

$$p^k m = 0 \quad \text{oraz} \quad p_1^{k_1} \cdots p_r^{k_r} m = 0,$$

i wobec $\text{nwd}(p^k, p_1^{k_1} \cdots p_r^{k_r}) = 1$ z lematu 9.2.1 otrzymujemy $m = 0$. Udowodniliśmy zatem, że $T_p(M) \cap \sum_{p_i} T_{p_i}(M) = 0$ i wobec tego $M = \bigoplus_p T_p(M)$. $\square$

9.2.4 Rozkład prymarny torsyjnych grup abelowych

Dla grup abelowych z twierdzenia 9.2.1 wynikają następujące wnioski.

WNIOSEK 9.2.1. *Każda torsyjna grupa abelowa jest sumą prostą swoich składowych prymarnych.*

Inaczej mówiąc, jeśli w grupie abelowej M każdy element ma rząd skończony, to grupa M ma rozkład na sumę prostą podgrup

$$M = \bigoplus_p T_p(M)$$

gdzie p przebiega wszystkie liczby pierwsze i dla każdej liczby pierwszej p podgrupa $T_p(M)$ jest albo podgrupą zerową albo każdy niezerowy element grupy $T_p(M)$ ma rząd będący potęgą liczby pierwszej p . W szczególności otrzymujemy też następujący ważny wniosek.

WNIOSEK 9.2.2. *Każda skończona grupa abelowa jest sumą prostą podgrup, których rzędy są potęgami różnych liczb pierwszych.*

9.2.5 Rozkład prymarny modułu V_τ

Sformułujemy tutaj twierdzenie 9.2.1 o strukturze modułów torsyjnych nad pierścieniami ideałów głównych dla $K[X]$ –modułu V_τ . Uzupełnimy je dodatkowymi informacjami i uzyskamy w ten sposób nowy dowód znanego nam już twierdzenia o rozkładzie przestrzeni wektorowej (twierdzenie 5.2.2) na sumę prostą podprzestrzeni niezmienniczych endomorfizmu $\tau \in \text{End}_K V$ wyznaczonych przez czynniki wielomianu minimalnego endomorfizmu τ nierozkładalne nad ciałem K . Jak wiemy, twierdzenie to jest podstawą analizy postaci kanonicznych macierzy endomorfizmów i redukuje badanie macierzy endomorfizmu do przypadku, gdy wielomian minimalny endomorfizmu jest potęgą wielomianu nierozkładalnego nad ciałem K .

Rozpatrzmy $K[X]$ –moduł V_τ . Jest to moduł ograniczony, gdyż $p_\tau v = p_\tau(\tau)(v) = 0_V(v) = 0$, a więc $p_\tau V = 0$. Ponadto

$$\text{Ann } V_\tau = (p_\tau),$$

to znaczy, anihilator $K[X]$ –modułu V_τ jest ideałem głównym generowanym przez wielomian minimalny endomorfizmu τ .

Moduł V_τ jako moduł *ograniczony* nad pierścieniem $K[X]$, jest na podstawie twierdzenia 9.2.1 sumą prostą swoich prymarnych składowych:

$$V_\tau = \bigoplus T_q(V_\tau),$$

gdzie q przebiega parami niestowarzyszone wielomiany nierozkładalne pierścienia $K[X]$. Składowe prymarne $T_q(V_\tau)$, jako podmoduły V_τ , są podprzestrzeniami τ –niezmienniczymi przestrzeni wektorowej V .

Na podstawie lematu 9.2.3, jeśli q^ℓ jest najwyższą potęgą wielomianu nierozkładalnego q dzielącą wielomian minimalny p_τ endomorfizmu τ , to

$$T_q(V_\tau) = \{v \in V : q^\ell(\tau)(v) = 0\}.$$

A więc możemy napisać

$$T_q(V_\tau) = \ker q^\ell(\tau). \quad (9.6)$$

Możemy teraz sformułować twierdzenie strukturalne (twierdzenie 9.2.1) dla modułów nad pierścieniami ideałów głównych w przypadku $K[X]$ –modułu V_τ . Milcząco zakładamy oczywiście, że przestrzeń V jest niezerowa.

TWIERDZENIE 9.2.2. (Twierdzenie o rozkładzie).

Niech $\tau \in \text{End}_K V$ i niech $p_\tau \in K[X]$ będzie wielomianem minimalnym endomorfizmu τ . Niech

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k}, \quad m_1 \geq 1, \dots, m_k \geq 1$$

będzie rozkładem wielomianu p_τ na iloczyn potęg parami niestowarzyszonych wielomianów nierozkładalnych nad ciałem K . Dla każdego $i = 1, \dots, k$ niech

$$V_i := \ker q_i^{m_i}(\tau) = \{v \in V : q_i^{m_i}(\tau)(v) = 0\}.$$

Wtedy

- (a) V_i jest niezerową podprzestrzenią niezmienniczą endomorfizmu τ .
- (b) $V = V_1 \oplus \cdots \oplus V_k$.
- (c) Endomorfizm indukowany τ_i na przestrzeni V_i ma wielomian minimalny $q_i^{m_i}$.

Dowód. (a) Jak już zauważyliśmy (zob. (9.6)), podprzestrzeń $V_i = \ker q_i^{m_i}(\tau)$ jest faktycznie równa $T_{q_i}(V_\tau)$, jest więc q_i –prymarną składową modułu V_τ . Na podstawie przykładu 9.2.4, jest zatem niezerową podprzestrzenią niezmienniczą przestrzeni V .

(b) wynika z (a) i z twierdzenia 9.2.1.

(c) można teraz udowodnić tak samo jak zrobiliśmy to w dowodzie twierdzenia 5.2.2. Można też argumentować następująco. Stwierdzenie (c) jest równoważne temu, że $\text{Ann } T_{q_i}(V_\tau) = (q_i^{m_i})$. Wystarczy więc udowodnić tę ostatnią równość. Anihilator $\text{Ann } T_{q_i}(V_\tau)$ jest ideałem głównym generowanym przez pewien wielomian $f \in K[X]$. Z (9.6) wynika, że $q_i^{m_i} \in \text{Ann } T_{q_i}(V_\tau)$, zatem f dzieli $q_i^{m_i}$. Wobec nierozkładalności wielomianu q_i wynika stąd, że f jest potęgą wielomianu q_i , $f = q_i^t$ gdzie $t \leq m_i$. Gdyby $t < m_i$, to dla wielomianu p , który powstaje z p_τ przez zamianę czynnika $q_i^{m_i}$ na f mielibyśmy

$$pV = pV_1 \oplus \cdots \oplus pV_k = 0,$$

skąd wynika, że p_τ dzieli p , sprzeczność. A więc $t = m_i$ oraz $\text{Ann } T_{q_i}(V_\tau) = (q_i^{m_i})$, co dowodzi (c). $\square$

9.3 Moduły skończenie generowane

Przypomnijmy, że R –moduł M jest *sumą prostą* swoich podmodułów $N_1, \dots, N_k$, jeśli każdy element $m \in M$ ma *jednoznaczne* przedstawienie w postaci

$$m = n_1 + \cdots + n_k, \quad n_i \in N_i, \quad i = 1, \dots, k.$$

W szczególności, jeśli podmoduły N_i są cykliczne oraz N_i jest generowany przez element $m_i \in M$, to fakt, iż M jest sumą prostą podmodułów N_i oznacza dwie rzeczy: po pierwsze, każdy element $m \in M$ można przedstawić w postaci

$$m = a_1 m_1 + \cdots + a_k m_k, \quad a_1, \dots, a_k \in R$$

oraz, po drugie, przedstawienie to jest jednoznaczne w tym sensie, że jeśli mamy inne przedstawienie

$$m = b_1 m_1 + \cdots + b_k m_k, \quad b_1, \dots, b_k \in R,$$

to $a_1 m_1 = b_1 m_1, \dots, a_k m_k = b_k m_k$. Nie wymagamy więc jednoznaczności współczynników $a_i \in R$ ale jednoznaczności składników $a_i m_i \in N_i$.

9.3.1 Struktura modułów skończenie generowanych

Przechodzimy teraz do głównego twierdzenia o strukturze modułów skończenie generowanych nad pierścieniami euklidesowymi. Przypomnijmy, że pierścień całkowity R nazywa się *euklidesowy*, jeśli dla R prawdziwe jest twierdzenie o dzieleniu z resztą (względem odpowiedniej normy euklidesowej N). A więc istnieje funkcja $N : R \setminus \{0\} \rightarrow \mathbb{N} \cup \{0\}$ taka, że dla każdego elementu $a, b \in R, b \neq 0$ istnieją $q, t \in R$ spełniające równość $a = bq + t$ przy czym albo $t = 0$ albo $N(t) < N(b)$. Łatwo dowodzi się, że każdy pierścień euklidesowy jest pierścieniem ideałów głównych. Rzeczywiście, jeśli $\mathcal{I}$ jest niezerowym ideałem pierścienia euklidesowego R oraz $a \in \mathcal{I}$ jest niezerowym elementem o najmniejszej normie euklidesowej wśród elementów ideału $\mathcal{I}$, to a dzieli każdy element ideału $\mathcal{I}$, skąd wynika, że $\mathcal{I} = aR$ jest ideałem głównym. Przykładami pierścieni euklidesowych są pierścień $\mathbb{Z}$ liczb całkowitych (z normą $N(a) = |a|$) oraz pierścień $K[X]$ wielomianów jednej zmiennej nad dowolnym ciałem K (z normą $N(f) = \deg f$).

Niech M będzie modulem skończenie generowanym i niech r będzie najmniejszą z liczb elementów w zbiorach generujących moduł M . Liczbę r nazywamy *rangą* modułu M i oznaczamy ją

$$\text{rank } M = r.$$

Zbiór generatorów modułu składający się z $r = \text{rank } M$ elementów nazywamy *minimalnym zbiorem generatorów* modułu M .

TWIERDZENIE 9.3.1. *Niech R będzie pierścieniem euklidesowym i niech M będzie skończenie generowanym R -modulem. Wtedy M jest sumą prostą skończonego zbioru podmodułów cyklicznych. Dokładniej, jeśli $\text{rank } M = r$, to moduł M jest sumą prostą r podmodułów cyklicznych.*

Dowód. Przeprowadzimy dowód indukcyjny ze względu na $\text{rank } M$.

Jeśli $\text{rank } M = 1$, to moduł M jest cykliczny i twierdzenie jest prawdziwe.

Założmy więc, że $\text{rank } M = r > 1$ i twierdzenie jest prawdziwe dla R -modułów o randze $r - 1$. Niech $\{m_1, \dots, m_r\}$ będzie minimalnym zbiorem generatorów modułu M . Wtedy $M = Rm_1 + \dots + Rm_r$ jest (zwykłą) sumą podmodułów cyklicznych Rm_i . Jeśli dla $a_1, \dots, a_r \in R$ równość

$$a_1m_1 + \dots + a_rm_r = 0 \tag{9.7}$$

zachodzi tylko wtedy gdy wszystkie jej składniki są 0 ($a_1m_1 = \dots = a_rm_r = 0$), to wynika stąd, że każdy element $m \in M$ ma jednoznaczne przedstawienie w postaci sumy elementów podmodułów cyklicznych $Rm_1, \dots, Rm_r$ i wobec tego moduł M jest sumą prostą podmodułów cyklicznych Rm_i . W tym przypadku twierdzenie jest więc prawdziwe. Pozostaje więc rozpatrzyć przypadek, gdy istnieje równość (9.7), w której nie wszystkie składniki a_im_i są równe zero. Pokażemy, że wtedy można zmodyfikować dany zbiór generatorów tak, by nowy zbiór generatorów prowadził do rozkładu modułu M na sumę prostą podmodułów cyklicznych.

Spośród wszystkich minimalnych zbiorów generatorów $\{m_1, \dots, m_r\}$ i spośród wszystkich równości (9.7) wybieramy zbiór generatorów i równość, w której występuje *niezerowy* współczynnik a_i z najmniejszą normą euklidesową $N(a_i)$. Zmieniając ewentualnie porządek elementów w wybranym zbiorze generatorów $\{m_1, \dots, m_r\}$ możemy zakładać, że a_1 jest niezerowym współczynnikiem o najmniejszej możliwej normie euklidesowej we wszystkich równościach postaci (9.7). Udowodnimy teraz dwa pomocnicze fakty stwierdzające własności elementu a_1 .

Fakt 1. Jeśli

$$b_1m_1 + \dots + b_rm_r = 0 \tag{9.8}$$

dla pewnych $b_1, \dots, b_r \in R$, to a_1 dzieli b_1 .

Jeśli a_1 nie dzieli b_1 , to na podstawie euklidesowości pierścienia R istnieją $q, t \in R$ takie, że

$$b_1 = a_1q + t, \quad t \neq 0, \quad N(t) < N(a_1).$$

Wtedy z (9.7) otrzymujemy $a_1qm_1 + \dots + a_rm_r = 0$ i odejmując tę równość od równości (9.8) mamy

$$(b_1 - a_1q)m_1 + \dots + (b_r - a_rq)m_r = 0.$$

Tutaj $b_1 - a_1q = t \neq 0$ oraz $N(t) < N(a_1)$, co jest sprzeczne z wyborem naszego minimalnego zbioru generatorów i współczynnika a_1 . A więc a_1 dzieli b_1 .

Fakt 2. a_1 dzieli wszystkie współczynniki $a_2, \dots, a_r$ występujące w wybranej równości (9.7).

Przypuśćmy, że a_1 nie dzieli a_2 . Zatem istnieją $q, t \in R$ takie, że

$$a_2 = a_1q + t, \quad t \neq 0, \quad N(t) < N(a_1).$$

Z równości (9.7) otrzymujemy wtedy

$$a_1(m_1 + qm_2) + tm_2 + a_3m_3 + \cdots + a_rm_r = 0. \quad (9.9)$$

Zauważmy, że $\{m_1 + qm_2, m_2, \dots, m_r\}$ jest także minimalnym zbiorem generatorów modułu M i równość (9.9), w której występuje *niezerowy* współczynnik t o normie euklidesowej mniejszej niż $N(a_1)$, jest sprzeczna z wyborem $\{m_1, \dots, m_r\}$ oraz a_1 . Zatem a_1 dzieli a_2 . Zmieniając kolejność generatorów $m_2, \dots, m_r$ pokażemy w ten sam sposób, że a_1 dzieli wszystkie pozostałe współczynniki $a_3, \dots, a_r$.

Na podstawie Faktu 2 istnieją $q_2, \dots, q_r \in R$ takie, że

$$a_2 = a_1q_2, \dots, a_r = a_1q_r.$$

Wykorzystamy teraz elementy $q_2, \dots, q_r \in R$ do konstrukcji następującego zbioru generatorów modułu M :

$$m_1^* = m_1 + q_2m_2 + \cdots + q_rm_r, \quad m_2, \dots, m_r.$$

Zauważmy, że element m_1^* jest torsyjny oraz

$$a_1m_1^* = a_1m_1 + a_2m_2 + \cdots + a_rm_r = 0.$$

Niech $M_1 := Rm_1^*$ będzie podmodułem cyklicznym generowanym przez m_1^* oraz niech $N := Rm_2 + \cdots + Rm_r$ będzie podmodułem generowanym przez zbiór $\{m_2, \dots, m_r\}$. Wtedy

$$M = M_1 + N.$$

Pokażemy, że w istocie $M = M_1 \oplus N$. Wystarczy pokazać, że $M_1 \cap N = 0$. Przypuśćmy więc, że

$$m \in M_1 \cap N.$$

Wtedy istnieją $b_1, b_2, \dots, b_r \in R$ takie, że

$$m = b_1m_1^* = b_2m_2 + \cdots + b_rm_r,$$

a więc także

$$b_1m_1^* - b_2m_2 - \cdots - b_rm_r = 0.$$

Zastępując tutaj m_1^* kombinacją liniową $m_1 + q_2m_2 + \cdots + q_rm_r$ otrzymamy

$$b_1m_1 + (b_1q_2 - b_2)m_2 + \cdots + (b_1q_r - b_r)m_r = 0.$$

Na podstawie Faktu 1 wiemy, że a_1 dzieli b_1 , zatem $b_1 = qa_1$ dla pewnego $q \in R$ i w rezultacie otrzymujemy

$$m = b_1m_1^* = q(a_1m_1^*) = 0.$$

A więc $M_1 \cap N = 0$ co wraz z $M = M_1 + N$ dowodzi, że $M = M_1 \oplus N$. Ponieważ $\text{rank } N = r - 1$, więc na podstawie założenia indukcyjnego podmoduł N jest sumą prostą $r - 1$ podmodułów cyklicznych. Zatem moduł M jest sumą prostą r podmodułów cyklicznych. $\square$

9.3.2 Struktura skończenie generowanych grup abelowych

Szczególnie ważnym typem modułów są $\mathbb{Z}$ –moduły, czyli grupy abelowe. Podmoduł cykliczny $\mathbb{Z}$ –modułu M jest podgrupą cykliczną grupy abelowej M . Z twierdzenia 9.3.1 wynikają następujące wnioski.

WNIOSEK 9.3.1. *Każda skończenie generowana grupa abelowa jest sumą prostą grup cyklicznych.*

Dla opisu struktury skończonych grup abelowych zastosujemy teraz dokładnie taką samą taktykę jakiej użyjemy dla opisu struktury $K[X]$ –modułu V_τ . A więc najpierw zastosujemy twierdzenie o rozkładzie prymarnym przedstawiając skończoną grupę abelową M w postaci $M = \bigoplus T_p(M)$ gdzie $T_p(M)$ są grupami, w których każdy element ma rząd będący potęgą liczby pierwszej p (wniosek 9.2.2). Następnie zaś do każdej skończonej p –grupy abelowej $T_p(M)$ zastosujemy twierdzenie 9.3.1, czyli wniosek 9.3.1. W ten sposób otrzymujemy następujące ważne twierdzenie o strukturze skończonych grup abelowych.

WNIOSEK 9.3.2. *Każda skończona grupa abelowa jest sumą prostą grup cyklicznych, których rzędy są potęgami liczb pierwszych.*

9.3.3 Struktura $K[X]$ –modułu V_τ

Dla każdego endomorfizmu $\tau \in \text{End}_K V$ przestrzeni wektorowej V wprowadziliśmy strukturę $K[X]$ –modułu na przestrzeni V definiując działanie zewnętrzne

$$K[X] \times V \rightarrow V, \quad (f, v) \mapsto fv := f(\tau)(v).$$

Tak skonstruowany $K[X]$ –moduł oznaczamy V_τ (zob. przykład 9.1.4).

Niech $\tau \in \text{End}_K V$. Podmoduł cykliczny w $K[X]$ –module V_τ generowany przez wektor $v \in V$ jest zbiorem wszystkich wektorów przestrzeni V postaci $fv = f(\tau)(v)$, gdzie f jest dowolnym wielomianem o współczynnikach z ciała K . Podmoduł ten oznaczamy $K[X]v$. Jak każdy podmoduł modułu V_τ , podmoduł cykliczny $K[X]v$ jest podprzestrzenią τ –niezmienniczą przestrzeni wektorowej V . Podprzestrzenie $K[X]v$ rozpatrywaliśmy już w rozdziałach 6 i 8, gdzie nazywaliśmy je podprzestrzeniami τ –cyklicznymi przestrzeni wektorowej V .

Ponieważ moduł V_τ jest skończenie generowanym modulem nad pierścieniem euklidesowym $K[X]$, więc jego strukturę opisuje twierdzenie 9.3.1. Zanotujmy ten rezultat.

TWIERDZENIE 9.3.2. *Dla każdego endomorfizmu τ przestrzeni wektorowej V istnieje rozkład przestrzeni V na sumę prostą podprzestrzeni τ –cyklicznych.*

Optymalną informację o strukturze $K[X]$ –modułu V_τ otrzymamy wykorzystując obydwa twierdzenia strukturalne udowodnione w tym rozdziale, a więc twierdzenie 9.2.1 o strukturze modułów torsyjnych i twierdzenie 9.3.1 o strukturze modułów skończenie generowanych nad pierścieniami euklidesowymi. Tak więc jeśli

$$p_\tau = q_1^{m_1} \cdots q_k^{m_k}, \quad m_1 \geq 1, \dots, m_k \geq 1$$

jest rozkładem wielomianu minimalnego endomorfizmu τ na iloczyn potęg parami niestowarzyszonych wielomianów nierozkładalnych nad ciałem K , to zastosowanie twierdzenia 9.2.1 gwarantuje przedstawienie przestrzeni V w postaci sumy prostej podprzestrzeni τ –niezmienniczych

$$V = V_1 \oplus \cdots \oplus V_k,$$

gdzie endomorfizmy τ_i indukowane na V_i przez endomorfizm τ mają wielomiany minimalne $p_{\tau_i} = q_i^{m_i}$ (zobacz twierdzenia 9.2.2 o rozkładzie). Podprzestrzenie V_i są także skończenie generowanymi $K[X]$ –modułami, zatem na podstawie twierdzenia 9.3.2 każda podprzestrzeń V_i jest sumą prostą podprzestrzeni τ –cyklicznych,

$$V_i = U_{i1} \oplus \cdots \oplus U_{ir_i}.$$

W ten sposób otrzymujemy rozkład przestrzeni V na sumę prostą

$$V = \bigoplus_{i=1}^k \bigoplus_{j=1}^{r_i} U_{ij},$$

w którym U_{ij} jest podprzestrzenią τ –cykliczną, a więc także τ –niezmienniczą, z wielomianem minimalnym będącym dzielnikiem $q_i^{m_i}$.

Jeśli wielomiany q_i są liniowe, $q_i = X - a_i$, $a_i \in K$, to w podprzestrzeniach U_{ij} można wybrać bazy, których suma mnogościowa jest bazą podprzestrzeni V_i , względem której endomorfizm indukowany przez τ na V_i ma macierz będącą sumą prostą klatek Jordana $J(a_i)$ odpowiedniego stopnia (zobacz twierdzenie 7.1.1). W ten sposób otrzymamy twierdzenie o istnieniu postaci kanonicznej Jordana macierzy endomorfizmu τ .

W przypadku ogólnym, gdy wiemy tylko, że q_i są wielomianami nierozkładalnymi nad ciałem K , w podprzestrzeniach U_{ij} można wybrać bazy, względem których endomorfizmy indukowane przez τ mają macierze, będące macierzami stowarzyszonymi z potęgami wielomianu q_i o wykładnikach $\leq m_i$ (zobacz lemat 8.1.3). Suma mnogościowa tych baz jest bazą przestrzeni V , względem której endomorfizm τ ma macierz w postaci kanonicznej wymiernej.